IMPROVING TRAVEL DEMAND MODELLING FOR SENIORS BY USING ADMINISTRATIVE HEALTH CARE DATA

by

Brianna J. Morehouse

BScE, University of New Brunswick, 2016

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Master of Science in Engineering

in the Graduate Academic Unit of Civil Engineering

Supervisor: Trevor Hanson, PhD, P.Eng, Department of Civil Engineering

Examining Board: Eric Hildebrand, PhD, P.Eng, Department of Civil Engineering
Jim Christie, PhD, P.Eng, Department of Civil Engineering
Rose McCloskey, PhD, Department of Nursing and Health Sciences

This thesis is accepted by the
Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK
June 2021

©Brianna J. Morehouse, 2021

ABSTRACT

Understanding travel mobility of rural older adults (aged 65+) is essential to adequately plan for their future travel demands and needs given that driving one's own vehicle may be difficult or impossible for many in the long term. Historically, Canadian data sets have not been able to provide detailed information regarding travel behaviour beyond "Journey to Work", yet for some older adults, finding transportation to healthcare remains a major challenge. The literature identifies a need for understanding travel burden for healthcare related travel and measuring inequalities related to geographic location as there is currently no standard approach for defining and measuring travel impacts. The measure of travel impacts is complicated by a lack of data to conduct analyses. Administrative health data has considerable potential to permit the estimation of health travel demand, but its use for this purpose in Canada has not been widespread.

This research presents a novel methodology to analyze and use administrative health data from the New Brunswick Institute for Research, Data and Training (NB-IRDT) for transportation planning purposes. Datasets from the NB-IRDT were analyzed with statistical and spatial analysis tools to estimate the travel demands and distances of rural older adults associated with accessing health care facilities within New Brunswick. Data were organized in cross-classification tables and explored through regression models and the gravity model.

ACKNOWLEDGEMENTS

I have many people to thank for helping me finish this research project. A special thank you to NB-IRDT for making it possible to have access to the provincial health care data, and to the NB-IRDT analysts and team for their assistance and guidance throughout the process to complete this research. I also owe thanks to NSERC for research funding which allowed me to put full effort into completing this research project.

I am thankful for my family and friends who have supported me throughout this experience. My father Don has always encouraged me to achieve my goals. Since I was little, he always knew I would follow in his footsteps and take the career path towards engineering. My mother Sarah has been my biggest fan and supported me throughout this process. My brother Tristan has always made sure I had a good laugh and plenty of lobster along the way.

My fiancé Mike, thank you for always being by my side and encouraging my decision to return to school - you have always given me the motivation to do my best. To Parker (our loveable “little” bear) for his endless dedication of laying by my side while I worked to complete this project.

Lastly, a big thank you to Dr. Trevor Hanson. As my supervisor, Trevor has provided me with an enormous amount of guidance and encouragement throughout this process. He has always made himself available to answer questions and provide advice. His enthusiasm towards the field of transportation and his willingness to allow me to take on this research has made this entire project an enjoyable and valuable experience.

TABLE OF CONTENTS

ABSTRACT ..... **ii**

ACKNOWLEDGEMENTS ..... **ii**

TABLE OF CONTENTS ..... **iii**

LIST OF TABLES ..... **v**

LIST OF FIGURES ..... **viii**

CHAPTER 1 : INTRODUCTION **1**

1.1 Problem Statement 2
1.2 Purpose 3
1.3 Significance of the Study 4
1.4 Research Questions and Corresponding Hypotheses 5
1.5 Thesis Organization 5

CHAPTER 2 : LITERATURE REVIEW **7**

2.1 Aging Population in a Car Dependent Society 7
2.2 Approaches to Understanding Travel Demand 12
2.3 Issues with Rural Transportation Data to Support Modelling 19
2.4 Potential of Administrative Data to Address Rural Transportation Data Gaps 23
2.5 Administrative Health Data to Support Transportation Planning for Rural Older Adults 24
2.6 Techniques to Analyze Administrative Health Data for Transportation Planning 27
2.7 Analytical Applications Used to Measure Access to Health Care 29
2.8 Summary 40

CHAPTER 3 : METHODOLOGY **43**

3.1 NB-IRDT Process to Access Data 44
3.2 Working with the Administrative Data 49
3.3 Health Variables 72
3.4 Determining Descriptive Statistics 77
3.5 Linear Regression Models 81
3.6 Releasing the Data from the Lab 92
3.7 Trip Generation: Creating Classification Tables 93
3.8 Trip Distribution: Gravity Models 95

CHAPTER 4 : ANALYSIS & RESULTS **103**

4.1 Descriptive Statistics 103
4.2 Regression Analysis 135
4.3 Trip Generation: Cross Classification Tables 157
4.4 Trip Distribution: Gravity Model 176

4.5 Research Challenges and Limitations 199

CHAPTER 5 : CONCLUSIONS & RECOMMENDATIONS 202

5.1 Summary of Work.....202
5.2 Recommendations.....211

LIST OF REFERENCES 214

APPENDIX A: NB-IRDT APPLICATIONS AND DOCUMENTS.....216

APPENDIX B: SAS SYNTAX 254

APPENDIX C: HEALTH DATA & GRAVITY MODEL ITERATIONS.....311

CURRICULUM VITAE

LIST OF TABLES

Table 3-1 - Physician Billing Variables 56
Table 3-2 – Summary of Arc Map OD-Matrix Outputs 66
Table 3-3 – ArcMap Final Output Structure 68
Table 3-4 - Final ArcMap Distance Table for SAS 69
Table 3-5 - Frequency Table Example 79
Table 3-6 - Independent Variables Used to Develop DAD Health Model 84
Table 3-7 - Independent Variables Used to Develop Physician Billing Health Model 85
Table 3-8 - DAD Data - Correlation of Dependent Variables 86
Table 3-9 - Physician Billing Data - Correlation of Dependent Variables 87
Table 3-10 - Census Model Independent Variables 91
Table 3-11 - Total Health Trips for 2011 from the DAD Data 93
Table 3-12 - Population of New Brunswick from Census 2011 93
Table 3-13 - Average Population Between Census Years 2006-2011 94
Table 3-14 - Population by County from 2011 Census 95
Table 3-15 - Intrazonal Categories for Health Trips 99
Table 3-16 - Zone-to-Zone Categories for Health Trips 100
Table 4-1 - Census Data for 2011 104
Table 4-2 - Comparing Distance Travelled and Frequency for Type of Health Trip 107
Table 4-3 - General Practice Health Trips by Age Group 107
Table 4-4 - Specialty Practice Health Trips by Age Group 108
Table 4-5 - Comparison Table: Mean Distance Traveled by Origin County 109
Table 4-6 - Comparing Average Distance Travelled for General Practice Health Trips 112
Table 4-7 - Comparing Average Distance Travelled for Specialty Practice Health Trips 113
Table 4-8 - Specialty Group Services Used by Older Adults 114
Table 4-9 - Model 1A Regression Variables 138
Table 4-10 - Model 1B Regression Variables 140
Table 4-11 - Model 2A Regression Variables 149
Table 4-12 – Model 2B Regression Variables 151
Table 4-13 – Rate of Health Trips per person per year for New Brunswick Taken in 2011 158
Table 4-14 – Average Rate of Health Trips per person per year for New Brunswick Between 2005-2011 158

Table 4-15 - Rate of Health Trips per Individual That Took a Health Trip in 2011 159

Table 4-16 - Rate of Health Trips per Individual That Took a Health Trip Between 2005-2011 160

Table 4-17 - Average New Brunswick Trip Rate for General Practice per person per year Between 2005-2011 160

Table 4-18 - Average New Brunswick Trip Rate for Specialty Practice per person per year Between 2005-2011 161

Table 4-19 - Average Trip Rate for Health Trips by County per person per year Between 2005-2011 162

Table 4-20 - Average New Brunswick Trip Rate for General Practice by County per person per year Between 2005-2011 163

Table 4-21 - Average New Brunswick Trip Rate for Specialty Practice by County per person per year Between 2005-2011 164

Table 4-22 - Rate for New Brunswick Health Trips Taken in 2011 167

Table 4-23 - Rate for the Average New Brunswick Between 2006-2011 167

Table 4-24 - Rate of Health Trips per Individual That Took a Health Trip in 2011 168

Table 4-25 - Rate of Health Trips per Individual That Took a Health Trip Between 2006-2011 169

Table 4-26 - Average New Brunswick Trip Rate for General Practice Between 2006-2011 169

Table 4-27 - Average New Brunswick Trip Rate for Specialty Practice by County Between 2006-2011 170

Table 4-28 - Average Trip Rate for Total Health Trips by County 170

Table 4-29 - Average New Brunswick Trip Rate for General Practice by County Between 2006-2011 172

Table 4-30 - Average New Brunswick Trip Rate for Specialty Practice by County Between 2006-2011 173

Table 4-31 - Variation Between Internal Trips 178

Table 4-32 - Variation of County-to-County Trips 179

Table 4-33 - Distance (km) - Cost Matrix 181

Table 4-34 – DAD Data - Origin-Destination Matrix for 2011 Health Trips 182

Table 4-35 – DAD Data – Traditional Model - Trip Matrix for Health Trips in 2011 183

Table 4-36 – Traditional Model - Difference Between Actual and Predicted Health Trips 184

Table 4-37 - Health Factor Values for Zone-to-Zone Trips 185

Table 4-38 - Health Factor Cost for Zone-to-Zone Travel 188

Table 4-39 - DAD Data – Health Model - Trip Matrix for Health Trips in 2011 189

Table 4-40 – DAD - Difference of Predicted to Actual Trips 190

Table 4-41 - 2011 Trip matrix from Physician Billing Data 193

Table 4-42 – Traditional Gravity Model using Physician Billing Health Trips 194
Table 4-43 - Variation between Actual and Predicted 195
Table 4-44 – Health-based Gravity Model using Physician Billing Data 196
Table 4-45 – Variation between Actual and Predicted 197

LIST OF FIGURES

Figure 3-1 - Summarized NB-IRDT Data Approval 49
Figure 3-2 - Example SAS Code: Creating Default Dates for Missing Termination Dates 51
Figure 3-3 - Example SAS Code: DAD Admission Date Format 52
Figure 3-4 - Example SAS Code: Joining Tables with Identifiers 52
Figure 3-5 - Example SAS Code: Merging DAD and Citizen Datasets 53
Figure 3-6 - Example SAS Code: Applying conditions Using if-then-else Statement 54
Figure 3-7 - Example SAS Code: Creating Establishment ID Numbers from Institution Numbers 55
Figure 3-8 - Example SAS Code: Condition to Assign Origin Postal Codes to Physician Billing Data 57
Figure 3-9 - Example SAS Code: if-then Statement to Assign Destination Postal Codes to Provider ID 59
Figure 3-10 - Example SAS Code: Assigning a DA Number to Origin Postal Codes 60
Figure 3-11 - SAS Dataset Process - Part 1 61
Figure 3-12 - Example Code for ArcMap Text File to SAS Dataset 70
Figure 3-13 - Example SAS: Correcting Distance Data to Merge with SAS Datasets 71
Figure 3-14 - SAS Example: Combining ArcMap Distances to Health Trips 72
Figure 3-15 - Process to Combine Health Variables with DAD and Physician Billing Data 77
Figure 3-16 - Example SAS: DAD Data - Creating Origin-Destination Tables 78
Figure 3-17 - Example SAS: DAD Data - Distance Statistics 80
Figure 3-18 - Example SAS: Creating Dummy Variables 82
Figure 3-19 - Health Factors Cost Matrix Categories 101
Figure 4-1 - Comparison of Population (65+) to 2011 Province Population 104
Figure 4-2 - Percentage of Health Tips Compared to Population by County 105
Figure 4-3 – Percent of Health Trips per Specialty Type 106
Figure 4-4 – Average Distance Traveled (km) VS Total Percent of Health Trips per County ... 111
Figure 4-5 - DAD - Travel Distance for Percent of Health Trips Specialized in Cardiology 116
Figure 4-6 - PHB - Travel Distance for Percent of Health Trips Specialized in Cardiology 117
Figure 4-7 - DAD - Travel Distance for Percent of Health Trips Specialized in Diagnostic Radiology 118
Figure 4-8 - PHB - Travel Distance for Percent of Health Trips Specialized in Diagnostic Radiology 119

Figure 4-9 - DAD - Travel Distance for Percent of Health Trips Specialized in General Practice 120
Figure 4-10 - PHB - Travel Distance for Percent of Health Trips Specialized in General Practice 121
Figure 4-11 - DAD – Travel Distance for Percent of Health Trips Specialized in General Surgery 122
Figure 4-12 - PHB - Travel Distance for Percent of Health Trips Specialized in General Surgery 123
Figure 4-13 - DAD - Travel Distance for Percent of Health Trips Specialized in Ophthalmology 124
Figure 4-14 - PHB - Travel Distance for Percent of Health Trips Specialized in Ophthalmology 125
Figure 4-15 - DAD - Travel Distance for Percent of Health Trips Specialized in Orthopedic Surgery 126
Figure 4-16 - PHB - Travel Distance for Percent of Health Trips Specialized in Orthopedic Surgery 128
Figure 4-17 - DAD - Travel Distance for Percent of Health Trips Specialized in Urology 129
Figure 4-18 - PHB - Travel Distance for Percent of Health Trips Specialized in Urology 131
Figure 4-19 - DAD Regression Models 1A to 1C 137
Figure 4-20 - Physician Billing Regression Models 2A to 2C 148
Figure 4-21 – County Map of New Brunswick (GNB 2021) 177

CHAPTER 1 : INTRODUCTION

In jurisdictions with rural geography and aging populations with high car dependence, there has proven to be a greater need for transportation engineering research to better understand the safety and mobility implications of older drivers. When travel options other than driving are limited in rural locations, it can be difficult for older adults to meet their needs independently. This can sometimes result in older adults depending on others, such as friends or family, or not making a trip due to driving being the only available, or practical option.

The province of New Brunswick (NB), Canada, is an example of a jurisdiction with an aging population, an extensive rural population, and high automobile dependence. A total of 20.8% of its population is aged 65 years and older, making it one of the "oldest" provinces in Canada by population (Statistics Canada, 2019). This trend is also evident nationally as it is expected that by year 2024, one in five Canadians will be aged 65+ (Statistics Canada, 2019).

Changes in health and mobility that can come with aging can present older adults with a decision to either continue, limit or stop driving altogether (Marottoli & Mendes de Leon, 2000). The 2009 Canadian Community Health Survey – *Healthy Aging*, shows that driving is the primary method of transportation for older adults (aged 65 years and older) in New Brunswick (Statistics Canada, 2010). The number one choice for older adults that do not drive is to travel with friends and family (Hanson & Hildebrand, 2011). According to the 2018 General Social Survey of Canada, approximately 25% of Canadians provide care to friends or family that have problems related to aging, including a long-term condition or a physical or mental health disability. "Assistance with transportation" was reported by 84%

of caregivers, including travel for errands, shopping, or medical appointments (Statistics Canada, 2018). Given that the aging of NB's population is mainly a result of "interprovincial migration loss" along with low fertility rates (Statistics Canada, 2019), it is unclear whether sufficient informal support will exist in the future to support safety and mobility needs?

1.1 Problem Statement

Canadian data sets, particularly Government of Canada data and Statistics Canada datasets, do not provide enough information to understand the detailed travel behaviour of older adults (65+) to support operational planning aspects for transportation systems (Hanson, 2018). Similarly, rural jurisdictions may not be conducting travel surveys or detailed transportation service planning to sufficiently understand the magnitude of local needs. There is potential for administrative datasets, which are government-maintained databases on citizen use of government services, to be a source of information for estimating travel demand for those services. Quantifying the use of "health services" by rural older adults is an opportunity to translate this use into an estimated travel demand for these services.

Datasets from the New Brunswick Institute for Research, Data and Training (NB-IRDT) were analysed to determine the feasibility of using administrative data to quantify travel demands of rural older adults accessing health care facilities within New Brunswick. These administrative datasets have yet to be explored for travel demand modelling purposes, yet their attributes and variables suggest there is a potential opportunity to quantify older adult travel behaviour for health purposes, and then to model the travel impacts on them due to a change in a rural healthcare facility or service. Understanding these impacts can help transportation engineers and planners to address current deficiencies within the

transportation network and plan for future demands for those rural areas. The data sets that were explored include: The Citizen Database, Discharge Abstract Database, Physician Billing Database, and the Provider Registry Database.

Currently, Volunteer Driver Programs (VDP) in rural New Brunswick have been able to help close a gap in terms of providing assistance with health care travel demands; however, with an aging population there is the possibility that these programs exclusively might not be a sustainable solution in the long term. This is an example of how the results from predicting travel demands and quantifying these impacts for rural areas could be used as a tool for transportation planners and potentially health facility planners as well.

1.2 Purpose

The first goal of this research was to determine how to use administrative health data in New Brunswick to quantify travel demands associated with healthcare use by older adults (aged 65 years and older) living in rural areas. This required developing an approach to accurately associate a healthcare destination with an origin using New Brunswick health data sets from NB-IRDT. By understanding how administrative data can be used to comprehend travel demands from rural communities, the impacts associated with changes to healthcare services and/or facilities can be quantified through a modelling approach. This would assist with filling a major gap within the field of transportation planning.

The second goal was to determine how to forecast changes in travel demands associated with demographic changes in the use of health facilities or services. The 4-step travel demand model (trip generation, trip distribution, mode choice and trip assignment) was explored for its suitability as a forecasting approach in predicting health trip distribution to

health facilities based on actual health record data. The underlying trip distribution model was the gravity model.

1.3 Significance of the Study

Older adults in rural areas often require travelling long distances and may require various modes of transportation to meet specific healthcare needs, but "transportation planning and services" is not within the mandate of health authorities. Quantifying the impacts on travel due to a change in rural healthcare facilities can be used to determine secondary effects, such as the ability of a person to meet their travel needs independently, impacts on caregivers or assistants, and roadway safety. This opens up opportunities for operational planning for both transportation planning and capital planning purposes.

Transportation Planners: By using historical and present travel demands within rural communities, transportation planners can use this information to forecast future demand for services for healthcare-related transportation for older adults. This will contribute in understanding what additional transportation requirements or services may be required for communities based on their demographics and geographic barriers.

Department of Health and Capital Planners: Facing a structural deficit of about $600 million, New Brunswick's Department of Health is considering the closure of various health care facilities across the province (The Canadian Press, 2016). This research will provide Government a potential tool to make evidence-based decisions and provide a better understanding of the additional travel demands of rural patients due to health care facility closures. This could potentially assist in identifying any additional transportation modes or services that would alleviate these travel challenges to rural patients. The results of this

study focused on impacts to rural older adults, but provides lessons that will be helpful to understand the impact for all people living in rural areas.

1.4 Research Questions and Corresponding Hypotheses

The first research question asks how administrative data regarding older adult use of health services can be translated into a way that aligns with transportation planning metrics, including the frequency of trip making, the origin and destination, time of day, etc. It is believed that applicable transportation data can be extracted from the administrative data sets and be used for understanding how older adults access health services and translating this into a measure of travel demand.

The second research question asks how administrative health data that has been translated into transportation planning metrics can be used for travel demand modeling of "health trips." It is believed that a relationship exists between the attributes of a health facility (e.g. number of beds, types of services), and the number, type and length of health trips attracted to those facilities. It is hypothesized that it is possible to adjust the attributes of a health facility and quantify how health trips redistribute to other facilities. Understanding these travel demands will lead to a more accurate understanding of the travel impacts associated with changes to a health care service or facility.

1.5 Thesis Organization

The remainder of this thesis will cover the following material: Chapter two will consist of a literature review exploring the aging demographic in Canada and New Brunswick, as well as challenges that seniors (aged 65+) experience in relation to transportation, more specifically in rural areas. From a health care perspective, the barriers for rural

communities and older adults are discussed, along with that main issue being presented for measuring travel impacts. Furthermore, different methods of travel demand modelling are explored as well as any data availability challenges. Chapter three gives a detailed methodology describing the steps that were taken to complete this research. Methodology steps will include data processing, GIS distance calculations, descriptive statistics, regression modelling and a trip distribution application using the gravity model. Chapter four will analyse the results obtained from the analysis of this research and any conclusions and recommendations will be provided within chapter five.

CHAPTER 2 : LITERATURE REVIEW

This chapter outlines the findings from the literature review including older adults in a car dependent society and health effects on driving, approaches to measuring travel demand, issues with rural data, the potential of administrative health data for transportation planning, and techniques used for analyzing spatial data. Additionally, various approaches others have used to measure access to health care are discussed. This chapter also identifies two main research needs found from the literature review including how administrative health data can address gaps in rural transportation planning for older adults and how transportation planning tools can integrate with health data to estimate impacts of health policy decisions on transportation.

2.1 Aging Population in a Car Dependent Society

Travel mobility is important for older adults, especially within rural areas as the communities are more spread out than in urban areas and often lack alternatives to driving one's own vehicle. Driving can give older adults the opportunity to participate socially within the community and meet their needs for activities such as grocery shopping and healthcare. The broader issue facing a car dependent society is how to address mobility issues for a growing population of older people when health effects of aging making driving difficult or impossible.

2.1.1 Growth in Older Adult Population

Between 2011 and 2018, the percentage of the Canadian population aged 65 years and older increased from 14.8% to 17.2%. This growth is expected to continue as the baby boomers (generation born in the years 1946-1965) reach the ages of 65 and older (Statistics

Canada, 2019). Population projections estimate that 1 in 5 Canadians will be aged 65 and older by 2024 (Statistics Canada, 2019).

While Canada's population is aging, the growth in older adult population is not uniformly distributed amongst the provinces and territories; in 2018 the province of New Brunswick had one of the highest percentages of the population aged 65 years and older (20.8%) (Statistics Canada, 2019). This is mainly a result of "interprovincial migration loss" along with low fertility rates within NB (Statistics Canada, 2019). As out-migration tends to be common in rural areas, this means that younger residents are moving out of NB, while families that remain are having fewer children, contributing to the increase in average age of the population. As the Canadian population ages, the challenges and limitations that older drivers may face could end up having significant societal impacts, in particular in rural areas where car dependence is high and essential services are primarily accessible by automobile.

2.1.2 Health Effects of Aging on Driving Ability

The ability for older adults to drive safely may be limited due to chronic medical conditions increasing with age (Syed & Holland Jr., 2016). Syed and Holland, from the *New Direction in Geriatric Medicine*, state that "there are three key functions for safe driving: vision, cognition, and motor/somatosensory function" (p.98).

Older adults may experience a decline in their effective depth perception, and central and peripheral vision as they age. They may also develop limited night vision driving due to glare and contrast sensitivity (Syed & Holland Jr., 2016). This may limit their ability to see important infrastructure, and obstacles or hazards in various lighting and weather conditions. This may also affect their judgment of their own, or other vehicle movements

such as distance and speeds. (Syed & Holland Jr., 2016). In addition, Syed and Holland (2016) state that "Drivers with dual sensory impairment are at a greater risk of motor vehicle accidents than those with only hearing or vision impairment" (p.96).

2.1.3 Rural and Urban Safety Impacts Associated with Aging

A study completed by Hildebrand, Myrick and Creed looked at collision rates between urban and rural older adults within New Brunswick, Canada. The study demonstrates that "*older drivers will face greater hazards on the city streets and are involved with a disproportionate amount of urban content collisions*" (Hildebrand, Myrick, & Creed, 2000). Overall, the data presented a trend where older adults (both rural and urban) are avoiding highway travel, driving during non-peak hours, and more than half are seldomly driving after dark. This suggests that older drivers are self-restricting their driving and recognize the deterioration of their driving skills (Hildebrand, Myrick, & Creed, 2000).

When comparing older adult driving behaviour between rural and urban areas, it was found that rural drivers are less likely to drive after dark than their equal drivers in urban areas. This could be due to the lack of streetlights that illuminate the rural roads which assists with some of the vision challenges that older adults experience when aging (Hildebrand, Myrick, & Creed, 2000).

In addition, Hildebrand and Myrick reported that older adults in rural areas had a greater demand for driving than their urban counterparts (2000). However, even though the rural older adults avoided urban driving and apply self-restrictive measures, they were most likely to have a collision on an urban road (57.8%) or a provincial highway (37%) (Hildebrand & Myrick, 2001).

Overall, rural drivers have a lower collision rate compared to urban areas due to lower volumes and opportunity for conflict points in a rural location. However, when looking at older adults aged 80+ years, the collision rate of rural drivers exceeds the rate for urban drivers of the same age group (Hildebrand & Myrick, 2001). Hildebrand and Myrick found that the collision rates (per veh/km) of older adults aged 80+ were 4.8 times higher than the group of rural drivers aged 46-65 (Hildebrand & Myrick, 2001).

When making decisions that influence travel, it is important to understand how older adults in rural communities operate differently than the same demographic in urban areas. From the results found by Hildebrand and Myrick, it is evident that older adults, specifically those aged 80+ are still highly reliant on driving. Hildebrand and Myrick make this conclusion based on the over representation of older adults with licences; this is because, even when experiencing declining driving skills, rural drivers tend to keep their driver's licence longer than older adults in urban areas (2000).

As explained by Syed and Holland, as older adults age their ability to drive independently may be reduced, which can be a challenge for rural seniors that rely on a vehicle as a main use of mobility. This may result in a situation postulated by Hildebrand and Myrick that some rural older drivers are driving longer than they should, while in other cases as described by Vézina and Turcotte, family members and friends step up to provide transportation.

2.1.4 Primary Options for Non-driving Older Adults

Statistics Canada (2010) reports that the main method of transportation for older adults is driving which raises a question as to how these users would meet their travel needs if driving was not an option. Hanson and Hildebrand (2011) used a Global Positioning

System (GPS) travel diary survey to better understand the transportation alternatives available to a convenience sample of older drivers in rural areas in New Brunswick. In addition, they sought out to understand the difficulty of securing an alternative mode of transportation when the option of driving oneself was inaccessible (Hanson & Hildebrand, 2011). It was found that when faced with the challenge of not being able to drive a vehicle, survey participants reported they would modify their driving behaviour by relying on friends and family in 52% of their trips and they would not have taken 34% of the trips they took (Hanson & Hildebrand, 2011). Hanson and Hildebrand explain that work trips and medical visits were the two categories that had the highest proportion of participants who reported they would still seek to take the trip without access to a vehicle (86% and 89%, respectively) (2011). From stated adaptation results, participants indicated that out of the remaining trips they would still take, they would fulfill 79% of their trip needs by relying on friends and family which results in them being self-sufficient for only 21% of their trip needs. Hanson and Hildebrand noted that family and friends appeared to be the number one travel choice for all trip lengths and that a participant's dependence on family and friends was tended to be the highest for trips greater than 20km. Participants also rated trip categories based on difficulty levels associated with finding travel alternatives; the highest rated difficulty was medical visits with 75% of participants considering medical trips either *"difficult but doable"* or a *"major inconvenience"*. This indicated that participants considered life maintenance trips/medical visits the most difficult trips to make without vehicle access (Hanson & Hildebrand, 2011).

While the research by Hanson & Hildebrand focused on older adults with independent mobility through driving, Vézina and Turcotte (2010) completed a study published by

Statistics Canada regarding caregiving assistance to older adults. This study determined that the main source of informal assistance to people who have chronic health problems are family members. Older adults who require support but do not have a partner to help assist them with care were found to be highly dependent on their children (Vézina & Turcotte, 2010). Vézina and Turcotte (2010) suggest that there will be an increasing number of adults required to support parents as the population ages due to a chronic illness or reduced independence from aging.

In 2018, Statistics Canada reported that one in four Canadians were providing care to friends or family, not only for issues related to aging, but also for long-term conditions and physical or mental disabilities as well. This also applies to assisting friends of family with transportation related needs. Statistics Canada states that 47% of caregivers reported caring for a parent or parent-in-law, followed by 13% providing care for a close friend or neighbour, 10% providing for extended family and 9% for a grandparent. For adults assisting their parents or parent-in-laws, 84% reported that the main type of help provided was transportation. Transportation assistance included running errands, shopping or medical appointments (Statistics Canada, 2018).

2.2 Approaches to Understanding Travel Demand

There is a general understanding that there is a growing population of older adults, who of which will eventually require assistance and not be able to drive independently. Many of these individuals will require alternate methods of transportation, typically receiving that assistance from family and friends. The level of demand for alternative transportation, particularly in rural areas, is not well understood. This section describes current approaches

for understanding travel demand, as well as the limitations and opportunities for the methods to assist with transportation planning.

Travel demand models are often used as an engineering tool for predicting travel patterns and behaviours of an area of interest. An assessment is completed, providing a measurement of advantages and disadvantages for travel demand alternatives for decision makers. These can consist of policy decisions, capital investments, socio-economic decisions and demographic understanding (Castiglione, Bradley, & Gliebe, 2015). Inputs are used to develop the travel demand model which can then be used to evaluate proposed alternatives.

The Institute of Transportation Engineering (ITE) Transportation Planning Handbook, an important reference for transportation planning, explains that travel demand models can be broken into two main groups: activity-based models and trip-based models (2016). Activity-based models focus on an individuals' activity and travel behaviour while, the trip-based models use a traditional four-step model approach (Transportation Planning Handbook, 2016). A detailed review of the two groups of travel models is presented below along with their associated limitations.

2.2.1 Activity-based Models

Activity-based models use behavioural theories about how individuals make decisions and schedule priorities for daily activities they participate in. The model suggests that travel demands are due to socio- economic activities and behavioural decisions consisting of constraints such as where the activity is, when to participate in the desired activity, and how the individual will get there. Because travel is based on a sequence of activities or routine, the individuals' ability to participate in additional activities is a function of the

individual's schedule. Therefore, as the schedule becomes filled, the likelihood of participating in additional activities will decrease. This ultimately represents the fundamental aspect that, behavioural decisions derive travel demands from individual needs and desired activities (Castiglione, Bradley, & Gliebe, 2015).

The development of the activity-based model was introduced as a tool for estimating travel demands with the individuals' goal of accomplishing personal activities throughout the day, producing a connected travel demand (Transportation Planning Handbook, 2016). Behavioural models use time and space when analysing traveler behaviour. The two goals of behavioural models are to explain traveler behaviour and to incorporate the values of time into a traveler's choice (Bhat & Koppelman, 1999).

Activity-based models share common attributes to the 4-step model approach such as identifying, generating and predicting traveler trip activities, destinations, modes and network facility routes for travel; the activity-based model incorporates a connection among travelers activities and a constraint of time and space that realistically represents the effect of travel conditions and choices (Castiglione, Bradley, & Gliebe, 2015).

This method is a more accurate representation of an individuals' travel decisions which leads to a more efficient tool that can predict how policies, investments or new infrastructure may affect travel behaviour (Castiglione, Bradley, & Gliebe, 2015). Activity-based models are able to represent how an individual or household travels throughout the day which can often be used to evaluate alternative policies or capital investments that are difficult to evaluate using a trip-based demand model. Because of the nature of these activity-based models, details of how different populations experience travel benefits or, non-benefits, can be estimated. Evaluations which are more sensitive to daily fluctuations

in pricing policies can be assessed, such as a bridge toll (Castiglione, Bradley, & Gliebe, 2015). Policies with a direct relationship to travel behaviour such as land use, congestion, pricing and travel demand management strategies can also be evaluated.

Activity-based models are known to overcome the conventional challenges that are associated with the four-step model approach such as the lack of a consistent human decision-making representation that is based off a strong behavioural framework. However, a limitation associated with activity-based models is the heavily detailed data requirements of travelers' activities and new data collection for verification of existing travel demands. These data is often obtained through data collection methods such as travel diaries, where the 4-step model requires data sets often captured through household surveys and census data (Transportation Planning Handbook, 2016).

2.2.2 Trip-Based Model

The trip-based model, a 4-step modelling approach, is used to measure 4 main objectives: how many people travel, what a study area's travel patterns are, which modes of travel are used, and which paths or routes will be used to travel the transportation network (Transportation Planning Handbook, 2016).

The ITE Transportation Planning Handbook 4th edition (2016) describes the trip-based model using 4 steps which are used to determine the above objectives:

(1) *Trip Generation*: Predict the number of trips that are generated by/ attracted to each defined zone in the study area. This includes estimating both trip productions and attractions; where trip productions are related from home and employment or other

non-home activities related to the trip attraction. The handbook suggests that trip productions are often used as a control point that trip attractions can be adjusted.

(2) *Trip Distribution*: Connect the production and attraction trips estimated within the trip generation model to determine the trip interchanges between zones. Trip distribution uses the productions and attractions developed within each zone and translates them into origin and destinations and identifies the travel time or cost associated. This creates an origin-destination trip as an output for trip-by-trip purposes.

(3) *Mode Choice:* Determine modes that will be used to travel on each zonal interchange using diversion curves or logit functions as predictive models. Mode choice models look at each origin-destination pair and estimate a number of trips that will use an available mode of transportation. This could include automobile, transit, walking, etc.

(4) *Trip Assignment*: Use attributes of each link for specific highway and transit routes such as capacity, speed limits, turning restrictions, etc. to produce an estimate of expected highway volumes and transit ridership. This method uses behavioural assumptions that suggest that users will choose a route to minimize their travel time and travel costs. This comes to an equilibrium when switching routes does not save a user any travel costs.

This four-step approach has evolved over time but has maintained its fundamental structure since the 1950s. The complexity of the model can vary depending on the study area's size and the overall understanding of the application from the technical staff (Transportation Planning Handbook, 2016).

Using actual travel behaviour observations, mathematical relationships to describe travel can be derived. This approach is used to help professionals analyze hypothetical scenarios such as alternative demographic forecasts, major highway and transit investments, etc. However, the trip-based models' principal limitation is its ability to associate an individual travel behaviour with the decision-making process for making a trip (Transportation Planning Handbook, 2016).

2.2.3 The Gravity Model

Gravity model is widely used as a form of trip distribution step within the 4-step model, therefore is used to measure the trip productions and attractions between zones within a study area. (Transportation Planning Handbook, 2016).

This method is based off the Law of Gravity; it assumes that the number of trips from zone A to zone B has a direct relationship with the trip productions within zone A and the trip attractions within zone B. It assumes that these productions and attractions are proportional to each other based on a friction factor that captures the spatial aspect of travel impedance between zones. This friction factor is considered inversely proportional to the number of trips and can be represented as a form of distance, travel time or some form of cost (Transportation Planning Handbook, 2016). This means that as a trip's distance increases between zones, the "attractiveness" or the draw to that zone decreases.

$$T_{ij} = P_i \times \frac{A_j \times f(C_{ij}) \times k_{ij}}{\sum A_j \times f(C_{ij}) \times k_{ij}} \tag{1}$$

Trip distribution considers both zone-to-zone and intrazonal travel impedance; zone-to-zone captures the impedance factor between two different zones while, intrazonal is used

to describe trips that take place within a zone (Transportation Research Board, 2012). Travel time or distance between each zone is a simple method to determine zone-to-zone impedance and intrazonal impedance can be determined from using a "neighbouring method" which takes the average of the nearest zones and then takes ½ of the average by dividing this average by two (Transportation Research Board, 2012).

Travel impedance is used to help determine friction factors, or as sometimes referred to as a deterrence function, that which deter trips from zones based on the cost or distance required for travel; these friction factors are then applied within the gravity model to produce a number of trips. There are two methods to determining the friction factors including mathematical formulas and fitted curves (Transportation Research Board, 2012). Shown below in equations (2), (3) and (4), are three common formulas that are often used to determine a friction factor: including a power function, exponential function and a gamma function.

$$f(c_{ij}) = c^{-n} \tag{2}$$

$$f(c_{ij}) = e^{(-\beta \times c)} \tag{3}$$

$$f(c_{ij}) = c^{n} \times e^{(-\beta \times c)} \tag{4}$$

Currently, there is not a preferred method or "best practice" to determine friction factors (Transportation Research Board, 2012); therefore, it is left up to judgment of which method is most appropriate for the model.

Characteristics of trip distributions can differ depending on the purpose of travel; therefore, it is recommended that the gravity model be applied to each travel purpose separately (Transportation Planning Handbook, 2016).

2.3 Issues with Rural Transportation Data to Support Modelling

The ITE Transportation Planning Handbook states*, "Rural communities can be defined in many different ways. In some cases, the term is used to denote a small city, town or village that is outside of a metropolitan area, while in others it represents areas of a state defined as "rural' by the U.S. Bureau of Census."* (2016).

Statistics Canada defines rural (for purposes of the census) as:

- *"Census rural is the population outside settlements with 1,000 or more population density of 400 or more inhabitant per square kilometre"* (Statistics Canada, 2015).

Transportation service planning requires data inputs to be effective. These data can include land-use information, traffic counts, road geometry, crash data, travel surveys, etc. These types of data collection can be a financial challenge for some rural communities which means they often do not collect the proper data to support an effective transportation planning system. In addition, if they do manage to collect the data, and are not able to maintain the data, the data becomes outdated quickly in any high growth areas (Transportation Planning Handbook, 2016).

During a review of the national datasets available in Canada, Hanson (2018) explains that Canada currently collects personal and household travel data in 3 main ways:

a. **Commuting Travel Data**: Statistics Canada provides a "Journey to Work" data which is a national dataset including small and rural communities. These data can provide valuable information regarding travel behaviours, commuting patterns and modal choice transportation; however, these data is fairly limited for transportation planning. The data from Statistics Canada do not provide information on trip

frequency (how often a trip is taken per week or per day) or information on modal split, which could indicate what percent of the time a person is a driver or a passenger during their commute (Hanson, 2018). Hanson further explains that the longer the distance between a place of residence and a place of work, the less likely that a "usual" commute will be represented from the data which can present difficulties when interpretating the data in rural areas. This is because residents from rural areas may be typically required to travel longer distances to work in comparison to someone from an urban area (2018).

Hanson explains that the 2011 National Household Survey (NHS) had low participation rates in past making the data unreliable in many small and rural communities (2018). This means that the travel surveys may be the only origin-destination information available for small or rural communities; however, the data may not always be available or reliable due to a small sample size from low participation rates (Hanson, 2018).

b. **National Surveys - travel behaviour for understanding regional and nation trends**: Hanson states *"The Census of Canada is the only national data source that collects and disseminates data at the community level, and the only travel behaviour characteristics presented are commuting-based"* (2018) . Hanson explains that General Social Survey on Time use is the most relevant data source for understanding travel behaviour. This survey is based on an interview style, 24hr time-use diary which provides information on the time spent for various activities. This includes information on travel between activities, and the type of transportation. Another national survey that asks transportation questions includes

the Canadian Community Health Survey (CCHS). Although the focus is on access to healthcare services and waiting times, this survey can provide relevant information about transportation and the aging population (Hanson, 2018). Hanson notes that questions regarding difficulties to accessing healthcare were asked with an option of transportation problems (2018).

c. **Metropolitan surveys - travel behaviour for service and infrastructure planning**: Metropolitan or urban travel surveys typically focus on measuring, understanding, and forecasting transportation (Hanson, 2018). Hanson states that *"While the Census "Journey to Work" data looks at the number of workers and where they live and work most often, an urban travel survey will look at all elements of a journey, and the one-way trips generated from Home and Non-home, and include more categories than "Work", since the goal is often to understand travel impacts on infrastructure capacity and service levels."* (Hanson, 2018).

   Historically, metropolitan transportation plans for New Brunswick cities included household travel surveys; however, they typically did not cover rural or suburban areas and were focused mainly on understanding the impacts of highway infrastructure (Hanson, 2018). Although the City of Moncton is moving towards approaches to target active transportation and transit usage, and forecasting automobile usage, a large percentage of the population resides outside of the 3 communities within study area. Hanson states that this population is likely *"highly automobile dependent, and primarily living in smaller, rural communities, yet fall outside any transportation planning framework that could assist them planning for a future where access to the automobile is reduced"* (2018).

The Transportation Planning Handbook suggests three possible approaches to developing a more usable database (2016):

(1) Using generic data collected from similar communities, although caution should be used with generic data because average values can hide many of the differences in the data that might be important within a rural aspect. The data is proportional to the number of samples within the average and the relevance of the sample to the study context. These types of datasets should only be used when validating local information and when no other data is available.

(2) Applying processes with quick response default values that are incorporated into the analysis approaches

(3) Collect the required data by combining resources with other agencies or communities.

The National Cooperative Highway Research Program (NCHRP) Report 365: Travel Estimation Techniques for Urban Planning. [Martin and McGuckin, 1998] is a useful tool when sourcing generic data for modeling small communities, it provides default values for variable and factor inputs for transportation analysis (Transportation Planning Handbook, 2016). However, it should be noted that a small community in the United States (U.S) may consist of a much larger population than a small community in Canada, or more specifically, a small community in New Brunswick, therefore, these values may not be appropriate for all jurisdictions.

Often in rural areas, transportation modeling is minimal and because of small populations and large distances for some transportation systems, it can be difficult to justify, which creates financial barriers to rural transportation solutions (Maikut, 2011).

Canadian national survey data can provide relevant and useful information for transportation planners in metropolitan or urban locations; however, there appears to be a gap in rural travel data due to the lack of available and reliable data, as well as often low participation from small communities and rural areas. For the information that is available from the Statistics Canada Census data, there is origin-destination data; however, this information does not provide trip frequency or other trip type information to better understand travel behaviour in rural communities, specifically older adults who are likely retired and not making a work trip.

2.4 Potential of Administrative Data to Address Rural Transportation Data Gaps

Administrative data, specifically health data, has the potential to address existing data limitations for transportation planning. The use of health data for transportation planning provides the opportunity to better understand the travel behaviour of older adults seeking healthcare. This also presents an opportunity to better understand rural travel behaviour for health-related travel as rural communities typically do not have enough data to provide reliable information.

Patient medical records would be considered one source type of health data, which using the information provided by health data, are able account for a person's trip origin-destination travel, how often health trips are taken and provide appropriate variables to indicate trip purpose. These travel attributes that are collected with health data provide an opportunity to better understand and predict travel patterns for health-related travel

purposes. Some variables available in the health data, which are not typically used in transportation modelling, may have the potential to provide insight on travel behaviour of rural older adults.

Travel surveys can be a burden to participants and can have a low participation rate, data collection can be limited and time consuming for researchers. Therefore, the use of health data for transportation planning can provide a reliable dataset that is already being collected with proper processes in place to ensure proper data collection and storage for records. Existing rural data sources are not strong resources for planning; therefore, the use of health data could provide transportation planners with reliable data to help close a gap for rural travel data, specifically for health-related travel needs.

2.5 Administrative Health Data to Support Transportation Planning for Rural Older Adults

Administrative health data provides a complete record for the use of health services including which individual accesses a service and on which date. This type of data can be interpreted to provide insight on travel behaviour for rural older adults that may need transportation alternatives to driving. Hanson and Hildebrand (2011) had found that among the sample of older adults in their study, health-related travel appeared to be the most difficult trips to find alternative transportation for. This would likely suggest that older adults would be seeking alternative transportation for those healthcare trips.

Hanson & Goudreau (2017) found that health trips (for non-emergency medical purpose) are one of the most frequently generated trips among seven Volunteer Driver Programs (VDP) in rural communities in New Brunswick, many of which have an older adult clientele. Health trips accounted for more than half of all drives within the one year of data

collection but had lowest rate of passengers per drive of all the known trip types and consumed the most resources. This is because health trips averaged a total of 1.1 persons per drive and averaged 81.7 km per drive meaning that users of the VDP for health trips are typically traveling alone and driving long distances (Hanson & Goudreau, How Many New Brunswickers Rely on Organized Volunteer Driver Programs? Preliminary Tools and Approaches to Help Us Find the Answer, 2017).

The 2017 Aging Strategy for New Brunswick reports there are 47.4% of New Brunswick seniors which live in a rural community (Province of New Brunswick, 2017). With New Brunswick's aging demographic and almost half of adults 65+ living in rural areas, the need for these VDPs (or a similar travel substitute) will likely increase to meet the needs of this age group, and it can be expected that "health trips" will continue to be the top trip purpose. The Canadian Institute for Health Information (CIHI) has found that seniors are typically the highest-cost users of health care services as they are considered the most frequent users. Senior costs are often higher with a more complex care required due to being affected by multiple chronic conditions (Canadian Institute for Health Information, 2019).

The 2019 document from the Canadian Institution for Health Information (CIHI) states that "*Access to health care services for people living in rural locations may be challenged by geographic remoteness, long distances to facilities and inclement weather conditions, among other factors*" (2019). Rural areas typically have higher mortality rates and a lower rate of self-reported health compared to urban areas. In order to improve rural residents' quality of care and access to health services, there is a need for a more in-depth understanding of the rural population (Canadian Institute for Health Information, 2019).

In comparison to urban communities, rural areas in New Brunswick present a challenge where there are fewer service providers for seniors (Province of New Brunswick, 2017). It is difficult for rural communities to attract a young workforce of health care professionals because they often seek work in a more urban centre, resulting in few providers and having an aging workforce in rural communities (Province of New Brunswick, 2017). These rural communities are faced with a combination of aging health care workforce and an older demographic with increasing health care needs; this means that some of these older adults are required to drive long distances to seek specialized care for their health needs.

In 2015, Horizon Health, one of two provincial health authorities in New Brunswick, had stated an objective to develop a sustainable health system in New Brunswick by optimizing their health care professionals, health facilities and specialized equipment by regionalizing health care in some areas (Horizon Health Network, 2015). Given that there is potential for more centralizing of health care in New Brunswick, this may result in an increased travel demands for rural users, including for rural older adults which is a population demographic that may require the most assistance with traveling.

Unlike some surveys or sample diary results, administrative health data provides a complete record of travel for health-related trips including origin, destination, trip purpose and demographic variables. Although these data method does not provide a complete record of all travel by an older adult, understanding health transportation needs would address the trip type that is expected to have a high demand for alternative transportation.

Using administrative health data to better understand travel behaviours of rural older adults provides the opportunity to quantify the impacts transportation decisions have regarding travel demand for health care purposes. It is important to better understand what type of

impacts these decisions have, at what magnitude and how it will change how people use transportation.

2.6 Techniques to Analyze Administrative Health Data for Transportation Planning

The CIHI states that an "Urban and rural/remote designations can be defined using a range of measures, including population size, population density and distance/travel burden to an urban centre or to an essential service" (Canadian Institute for Health Information, 2018). Currently CIHI is able to use place of residence postal codes from administrative health data to measure health inequalities between urban and rural/remote areas (Canadian Institute for Health Information, 2018).

Statistical Area Classification (SAC) and Population Centre (POPCTR) are two geographic related measures that use Statistics Canada's postal code conversion file (PCCF/PCCF+) to help categorize individuals. SAC combines information from population size and commuter flows to determine which census subdivisions (CSD) belong to a component of census metropolitan area (CMA), Census Aggregation (CA) or metropolitan influenced zone (MIZ) (Canadian Institute for Health Information, 2018). POPCTR combines information on size and density to identify areas which fall within 3 categories based on the population size; these include small population centres, medium population centres and large urban population centres (Canadian Institute for Health Information, 2018).

Currently these two measures are used as a base-line approach and informative tool to better identify locations that are considered urban, rural or remotes areas, as well as to assist with measuring health inequalities however, the SAC tool presents limitations when determining between rural and remote areas (Canadian Institute for Health Information,

2018). CIHI states that there is a need for measuring travel burden analysis to measure health inequalities related to geographic location and that "***there is a currently no standard approach for defining and measuring travel burden at a pan-Canadian level***" (Canadian Institute for Health Information, 2018).

2.6.1 Two-Step Floating Catchment Area

The 2-step floating catchment area (2SFCA) model is related to the trip-based modeling group and is another form of gravity model that was developed as a measurement of spatial accessibility to primary healthcare (McGrail & Humphreys, 2009). The gravity model integrates the demand from a population of a limited service with the concept of diminishing "attractiveness" as distance increases (Luo & Wang, 2003; McGrail & Humphreys, 2014).

As a form of the gravity model, the 2-step floating catchment area (2SFCA) model includes levels of supply and demand as additional accessibility factors (Frew, Higgs, Harding, & Langford, 2017). The 2SFCA model uses catchment sizes to help delimit how far a service is being delivered to a user and how far populations are prepared to travel to access a particular service (McGrail & Humphreys, 2014). The 2SFCA method also measures population-to-provider ratios within a defined distance of each supply location (Frew, Higgs, Harding, & Langford, 2017).

McGrail and Humphreys (2009) define two 2 steps as follows:

(1) Service Catchment: Determines what population falls within a threshold distance of the catchment area of each service providers. Population-to-provider ratios are then calculated.

(2) Population Catchment: Find all services that fall within the threshold distance of the catchment area of each population and allocates available services to populations. Population-to-provider ratios are then summed.

An accessibility measure can be determined for each demand centre by the sum of ratios for all supply locations found within the distance threshold of each demand center (Frew, Higgs, Harding, & Langford, 2017). With the 2SFCA method, it possible for two catchment areas to overlap resulting in a higher accessibility rate for the residents within the overlapping catchment area (Luo & Wang, 2003).

2.6.2 GIS Modelling

Increases in rural or urban travel demands can be modeled in various ways including special modeling programs such as Geographic Information Systems (GIS), Gravity Models and Two-step floating catchment areas (2STCA) (Frew, Higgs, Harding, & Langford, 2017). GIS can determine all possible combinations of distances between origin nodes, which would be the patient's location, and destination nodes, which would be the desired health care facility (Brabyn & Skelly, 2002). Using known travel distances and national road network information it is possible to determine travel times based on speeds of rural or urban roads (Brabyn & Skelly, 2002). An example of tools that have been used with GIS Modeling for travel demands are Esri Arc Map and ArcGIS Network analyst extension (Frew, Higgs, Harding, & Langford, 2017).

2.7 Analytical Applications Used to Measure Access to Health Care

This section will discuss the various research applications conducted by others which explored modelling healthcare travel for spatial applications.

2.7.1 Illinois Study (Wang and Luo 2003) - Measures of Spatial Accessibility to Health Care in a GIS Environment: Synthesis and a Case Study in the Chicago Region

In the state of Illinois, Luo and Wang chose ten counties within the Chicago Consolidated Metropolitan Statistical Area (CMSA) as the study area for spatial accessibility to healthcare. The study area represented both urban and rural areas in the region and included a 15-mile (24.1km) buffer zone around each county to account for "edge-effects" (Luo & Wang, 2003). The 2000 Census data files were used to extract population data and boundary files for census tracts and census blocks. Census tracts were then assigned a centroid which are typically calculated geographically, however because the population is not typically distributed in a homogenous manner, population-weighted centroids were calculated (Luo & Wang, 2003). This was to better represent rural areas that have large census tracts and a more concentrated population in a smaller area (Luo & Wang, 2003).

The Physician Master File of the American Medical Association (AMA) provided all primary care physician data of Illinois in 2000 which included all family physicians, general practitioners, general internalist, general pediatricians and some obstetricians. Address of a physician's practice was used to determine the location of the physician; when addresses were not available, the zip code was used. Similarly, to the census tract centroids, a population-weighted centroid was determined to better represent the likelihood of a physician's location. This centroid method was used because physicians typically choose to practice in populated locations (Luo & Wang, 2003).

The road network was derived obtained from the 2000 census data file, which was used to derive travel time estimations between each population and physician locations instead of

a straight-line travel path (Luo & Wang, 2003). Arc/Info network analysis was used to determine the shortest travel times between locations, assuming users would take the fastest path. All users were assumed to travel at the assigned speed limits which were retrieved from Illinois Department of transportation speed limit maps (Luo & Wang, 2003). When determining the network's shortest paths, impedance factors were applied using the travel time based on speed limits (Luo & Wang, 2003).

Luo and Wang then apply the two-step floating catchment (2SFCA) area method which applies a time threshold to the catchment areas instead of a straight-line radius, this helps to represent residents seeking health care locations that are on the quickest path. Population-to-provider ratios are applied within the catchment areas representing how many physicians are available to populations of census tracts within a time threshold (2003).

The 2SFCA method is later compared to the gravity-based method for determining accessibility, this method has been used in past studies for evaluating job accessibility. The larger the accessibility value (A) suggests better accessibility. When comparing the two methods it was determined that the 2SFCA method is simply a "special case" of the gravity-based accessibility method (Luo & Wang, 2003).

The two accessibility methods are then compared by standard deviations of their parameters. The 2SFCA methods uses a threshold travel time as parameters while the gravity-based method uses travel-friction coefficients (Luo & Wang, 2003). It was found that the gravity-based method leads to larger variance in accessibility when higher travel-friction coefficients are applied, and when higher time thresholds are applied the 2SFCA method leads to smaller variability of accessibility. In addition, the larger threshold of

travel time in the 2SFCA method corresponds with a lower travel-friction coefficient in the gravity-based method. This means that users would still travel a long distance to seek a physician if travel friction was less (Luo & Wang, 2003).

2.7.2 Illinois Study (Wang and Luo 2005) - Assessing Spatial and Nonspatial Factors for Healthcare Access: Towards an Integrated Approach to Defining Health Professional Shortage Areas

Wang and Luo describe how there are two aspects that are associated with accessing healthcare, spatial and non-spatial factors. Spatial factors account for the variance of access because of the distribution of healthcare providers as well as consumers of healthcare. This would include the geographic barriers between healthcare providers and users such as, distance or time. Non-spatial factors consider different socioeconomic and demographic characteristics that influence the variance among population groups (2005). Typically, these two factors are studied separately; however, Wang and Luo have highlighted three challenges to successfully integrating these two aspects (2005):

1. "*Implementing the measure of spatial accessibility in a reasonably simple process.*
2. *Aggregating various sociodemographic variables (often correlated) to independent (or uncorrelated) indicators of nonspatial accessibility; and*
3. *Integrating spatial and nonspatial factors into one framework for assessing healthcare access and identifying physician shortage areas*."

Similarly, to their previous study in 2003, Wang and Luo use 2000 population Census data and locations of physicians from the Physician Master File to determine accessibility to primary healthcare in Illinois using the two-step floating catchment area (2SFCA) method

(2005). In addition, Wang and Luo used the same methods to determine weighted centroids of populations and primary physician locations as they did in 2003.

Wang and Luo then used the factor analysis method (FA) to combine significant sociodemographic variables into selected factors. Different loadings were applied to determine if a factor were to be used as a primary or secondary influence for healthcare accessibility (2005).

Statistical data was obtained from the Census 2000 summary file concerning the following 11 sociodemographic variables, within 5 main categories that Wang and Luo had chosen were influential to healthcare accessibility (2005):

1. Demographic Variables: *population with high needs* (seniors aged 65+, children aged 0-4 and women aged 15-44).
2. Socioeconomic Status: *population in poverty, female headed households, home ownership* and *median income*.
3. Environment: *households with an average of more than 1 person per room* and *housing units lack of basic amenities.*
4. Linguistic Barrier and Service Awareness: *non-white minorities, population without a high school diploma* and *households linguistically isolated*.
5. Transportation Mobility: *households without vehicles*.

The factor analysis method was then used to better understand the correlation of the sociodemographic variables to show their relative importance when determining their influence on healthcare accessibility (Wang & Luo, 2005).

2.7.3 Australia Study (McGrail and Humphry 2009) - The Index of Rural Access: An Innovative Integrated Approach for Measuring Primary Care Access

McGrail and Humphreys look at the accessibility issue related to healthcare from an Australian perspective; they have recognized that rural and remote communities often have high health care needs while facing both spatial and aspatial barriers when seeking healthcare (2009).

To identify spatial and aspatial barriers that these areas experience, McGrail and Humphreys describe that "*Access to health facilities is a function of several factors including appropriate supply (availability), reasonable distance/time impedance to available services (proximity), the level and nature of need for those seeking care (health needs) and the ability of individuals to access care at a time of need (mobility)*" (McGrail & Humphreys, 2009).

McGrail and Humphreys suggest, populations that experience higher health needs require more services in relation to similar sized populations with lower health needs to be able to maintain an equal level of access (2009). In addition, mobility of a population is considered another significant factor when considering their ability to surmount a barrier such as distance (2009). They found that "*The key factor of mobility is an individual's access to transportation, which enables them to transcend distance when accessing services*." (McGrail & Humphreys, 2009).

This study aimed to provide an *Index of Rural Access* that can be applied to measure health access of rural areas by integrating spatial and aspatial barriers, this is accounting for rural accessibility barriers such as service availability, proximity, health needs and mobility (McGrail & Humphreys, 2009).

When measuring accessibility, to help capture proximity and availability together, McGrail and Humphreys used both "location of primary health care services and population, as well as the number of services and the size of the population at each location" (2009). Information from Australian Bureau of statistics (ABS), census population and housing information of Census Collection Districts (CD) was utilized, as well as, ABS boundary files, to calculate geometric centroids within CDs (McGrail & Humphreys, 2009).

The Medical Directory of Australia (MDA) was used to obtain a list of service providers consisting of all General Practitioners (GP), these GPs were being used to represent primary care services. Service locations of these GPs were geo-coded using ArcGIS software; the *closest facility tool* function within the Network Analysis tool in ArcView was then used to determine network routes and general proximity between the GP locations and populations. Rather then using distance, travel time was used to measure impedance by combining road section lengths and speed limits. Proximity data was then transferred to be used within the 2SFCA model (McGrail & Humphreys, 2009).

McGrail and Humphreys used regression modelling to determine a list of predictor variables to be used to help measure health needs within the *Index of Rural Access*. Most of these variables were obtained from an existing index retrieved from the ABS called the Socio-Economic Indexes for Areas (SEIFA); this was used because of its adequate selection of indicators which are used for predicting health needs. These variables included family type and income, education, occupation and employment status, and household type (2009).

When measuring for mobility, McGrail and Humphreys found that spatial accessibility is low in areas with low populations and typically higher in areas with greater population as

health services most likely tend to be located in those areas (2009). In addition, areas that experience poorer health, socio-economic status and higher health needs tend to have a close correspondence with the largest decrease in accessibility (McGrail & Humphreys, 2009). McGrail and Humphreys determined that using strictly geography classifications were not an accurate way of determining resource allocation to address problems related to health care accessibility (2009).

2.7.4 Montreal Study (Ngui and Vanasse 2012) - Assessing Spatial Accessibility to Mental Health Facilities in an Urban Environment

A study by Ngui and Vanasse intended to identify areas of southwest Montreal that were underserved, or at risk of being underserved, by focusing on potential accessibility of mental health services using the 2SFCA method to map disparities (2012). Ngui and Vanasse state that "Accessibility varies across space because neither health services nor service users are uniformly distributed" (2012).

Ngui and Vanasse used three main parameters when applying the 2SFCA method to the Montreal study. That included determining the supply and demand of users, as well as measuring the travel distance between the supply and demand locations (2012).

To determine the demands of the mental health facility users, information from the 2006 Census retrieved from Statistics Canada and dissemination areas (DA) were applied within the study area. The study area DAs were then used as catchment areas for capturing demands as they are the smallest spatial unit opposed to individual postal codes (Ngui & Vanasse, 2012). A list of 38 mental health facilities were obtained to determine supply of this service, these were then geocoded with ArcGIS using the facilities postal codes (Ngui & Vanasse, 2012). Lastly, Ngui and Vanasse used CanMap street files from DMTI to

calculate the shortest network distances between supply and demand locations using the Network Analysis tool in ArcGIS. Four thresholds including 500 m, 1 km, 2 km and 3 km were applied when determining the demand (Ngui & Vanasse, 2012). Like in a number of studies, Ngui and Vanasse found that as distance increased, a score of accessibility decreased substantially (2012).

2.7.5 Gaps Identified in Previous Approaches

Each of the above studies, designate a patient's residence as an origin, as well as a health facility as a destination when attempting to determine an area's health care accessibility. In addition, they each use impedance factors of time or distance thresholds when determining travel demand. These approaches provide a general understanding of accessibility to health care, assuming that travelers are going to the closest destination however, the following three concepts are not considered within these studies:

Doctors practicing at multiple locations:

It is recognized by Luo and Wang that the physician data used for the study does not account for the time a physician spends at a specific location or moving to multiple practices because that method would require extensive surveys and fieldwork (2003). Similarly, the other studies discussed by McGrail and Humphreys, as well as Ngui and Vanasse, do not account for this concept either. Knowing the location of a doctor providing a service can be beneficial when determining an accurate destination for travellers seeking health care. However, as stated by Luo and Wang, this method would take a lot of time and resources therefore, administrative data would be an optimal solution to address this missing concept.

Historical travel behaviour of patients seeking healthcare:

Understanding the historical health trips being traveled can help transportation planners and health officials better understand which areas are experiencing higher travel demands in relation to seeking health care. The destination of a health care trip can vary depending on what type of health care service is being sought after and which services are offered at facilities within a traveler's community. These studies account for the general assumption that travelers will take the shortest route and pick the closest destination similar to a typical trip; however, health care trips can have more influential factors that may attribute to longer trips for health-related reasons. Factors such as facility proximity, capabilities and resources (eg. Number of hospital beds) may influence travel for health care, as well as various trip purposes such as, emergency services, routine check-ups, specialized visits, etc. This may cause expected trip destinations to differ meaning that distance or time may not be the only variables influencing these trips.

Additional healthcare variables that influence travel:

The studies completed by Wang and Luo in 2005 and McGrail and Humphreys in 2009 aimed to identify variables that may have an influence on travel in relation to health care accessibility; however, because these studies do not look at the users historic travel behaviour, health information, or more specific facility variables and trip purposes, they do not consider what additional variables are influencing extended travel distances that may be occurring for areas with "poor accessibility". Traditional variables were identified in these studies such as demographic and socioeconomic information including education, income, etc. However, additional variables related to health trips such as health needs, including chronic illness and purpose of health trip, health facility information such as

available hospital beds, specialty care, etc. may all be variables that are influencing extended travel for healthcare users in rural areas.

These variables could be strong indicators as to why distance for health trips may be significantly greater in some locations. With an increasing age demographic in rural areas and increasing health needs from this population, understanding the variables that demonstrate they have an influence on health trips (outside of the expected trip distributions) will be important for transportation planning and forecasting aspects, as well as planning for health services.

2.8 Summary

To summarize the findings for the literature review, notable findings are as follows:

- Research indicates that there is a need to better understand the travel behaviour of older adults, including how they are able to make health trips on their own or with the help of others. This would help to support transportation alternatives for older drivers in rural areas who lack travel options.
- There is a lack of data, nationally and locally, that can help support the understanding of travel behaviour for older adults both. Rural data is difficult to capture due to lack of resources, budgets, etc.
- Driving is the main method of transportation for older adults and their ability to drive safely may be limited due to chronic medical conditions increasing with age. As older adults age their ability to drive independently may be reduced, which can be a challenge for rural seniors that rely on a vehicle as a main use of mobility.
- Older adults are typically the most frequent users of the healthcare system. They have been found to rely on friends and family for assistance with related health travel.
- Approximately 47% of New Brunswick seniors reside in a rural location. Rural areas typically experience fewer service providers resulting in more travel for older adults seeking healthcare.
- CIHI has indicated that there is a need for a standard approach to measuring travel burden related to geographic location for healthcare related travel.

2.8.1 Research Need #1: How administrative health data can address gaps in rural transportation planning for older adults

Administrative health data has the potential to provide additional opportunities for transportation planners that traditional census data cannot provide. Administrative health data uses patient health records to provide information on origins, destinations, type of visit, trip purpose, etc. in a secure and pseudonymous way to protect the identity and privacy of individuals.

Administrative data can provide accurate travel origins and destinations of a health trip by using unidentified data retrieved from health records. These data include various data bases that provide information regarding an origin, destination, what date a service is being provided to a patient as well as which physician was providing a service. These data need to be processed and linked together to demonstrate the proper flow of information; however, this allows researchers to better understand at which locations patients are being served, how often and at what distance cost. It is believed that administrative data provides a new platform for transportation planners to use health related information to plan for travel demands more effectively.

2.8.2 Research Need #2: How transportation planning tools can integrate with health data to estimate impacts of health policy decisions on transportation

The gravity model is an effective tool to determine trip distribution; as well, by applying an impedance factor, it is possible to represent the effect of increasing distance being less attractive to users. Determining which variables have a greater influence on travel related to health care can provide a more in depth prospective as to why older adults in rural areas may be traveling further than expected for health care trips. It may be possible to include

these influencing variables within the gravity model to more accurately distribute trips to appropriate facilitates for planning purposes. Using administrative health data to determine historic travel behaviour, specific to health trips, is a way to accurately demonstrate travel demands and trip distributions occurring in rural areas.

Existing travel data can be used not only to understand current travel behaviours but to also predict future travel behaviours using forecasting methods. Forecasting applies these past trends and uses them to estimate future travel behaviours. Using variables that influence health travel can also provide a better representation of travel demands when later determining the effects on a specific location if a facility or service was altered.

Due the novelty of using administrative health data sets for travel studies in New Brunswick, fundamental tools such as the gravity model are sufficient for the current stage of research.

CHAPTER 3 : METHODOLOGY

This chapter describes the methodology to complete the thesis work including processes to access administrative data, processing health datasets, determining travel distances, developing descriptive statistics and regression models to predict kilometers traveled for a health trip. This chapter also outlines the methodology of applying a base application of the first two steps of the four-step travel model, including trip generation and trip distribution.

The objective of this approach is to identify and work with administrative data sources that can permit quantifying the extent of travel someone aged 65 years and older undertakes when accessing the New Brunswick healthcare system. These administrative data sources were used to establish baselines for transportation planning by the development of models to forecast travel distance as demographics change and healthcare services are altered. This involves the identification of geographic origin, destination, frequency of health access, travel distances, and purpose of health trips. Administrative data was used to develop cross-classification tables for New Brunswick and a base application of the gravity model was applied to demonstrate trip distribution for New Brunswick health trips. Finally, a preliminary health factor was developed for the gravity model to improve the trip distribution of health trips for rural areas that are required to travel long distances for healthcare services localized in more urban locations. The following sections of this chapter detail the various processes used to develop this approach.

3.1 NB-IRDT Process to Access Data

The New Brunswick Institution of Research, Data, and Training (NB-IRDT) securely hold many databases from the Province of New Brunswick's Department of Health (DOH). These databases contain a variety of information regarding the citizens of New Brunswick and their personal health records. To access these data for research purposes, a formal process was required to be followed to ensure the privacy of the sensitive information. Figure 3-1, on page 49, represents the processes followed prior to gaining access to the lab at NB-IRDT.

3.1.1 Feasibility Application

This process began by submitting a feasibility application provided by NB-IRDT. This application consisted of preparing information about the project team and research information such as project summary, research goals, project methodology, and a list of potential datasets that would be requested. Prior to being able to submit a full data access application, the feasibility application needed to be reviewed and approved by the director of NB-IRDT.

The majority of databases available to researchers at NB-IRDT have a reference guide that gives a brief introduction to that database and a summary of all data variables that are available within the database. The reference guide then elaborates on details of what information each variable can provide based on that specific database category. Multiple database reference guides were reviewed before determining which databases were applicable to this project and which variables would be requested to make up each dataset. Variables were determined based on their expected value to the project from a travel demand modelling perspective.

Four main databases were used to reference variables which included: The Citizen Database, the Discharge Abstract Database (DAD), the Provider Registry Database, and the Physician Billing Database. The Citizen Database includes information about general residence and location variables for all residents in New Brunswick with a Medicare Card between the years of 1997-2016. The DAD Database is a provincial database that holds information capturing clinical and demographic information for health institutions within New Brunswick between the years of 1999–2014. The Provider Registry Database holds information between 1997-2014 and captures information regarding service providers and physicians practicing in New Brunswick. The importance of this database for the study was specifically the provider's location, dates of service, and specialty. Lastly, the Physician Billing Database was used to obtain information related to the dates of service, the purpose of the health trip, and provider numbers indicating the location of the hospital or clinic. These reference guides can be found on the NB-IRDT website.

3.1.2 Data Access Application

The data access application included four main sections and an additional document that outlined each individual data variable that was being requested for access. Section 1 of the application included basic information about the researcher and supervisor, while section 2 inquired about resource requirements and timelines. Section 3 requested further project details including project summary, research goals, methodology, preliminary results disclosure, and geocoding. Section 4 included a list of datasets that were being requested access to, as well as the time frames being requested. The datasets requested were from the following databases: The Citizen Database, the Discharge Abstract Database (DAD), Physician Billing Database, and the Provider Registry Database. All datasets were

requested between the dates of January 1, 2006, and December 31, 2011, to be consistent with Canadian census data.

The “Request of Data - list of variables” application was submitted with the main data access application to gain access to specific variables within the requested datasets that were thought to have value from a transportation aspect. This included variables such as age, location and postal codes, the purpose for health trips, etc. Each requested variable required a definition of that variable and a short rationale for access. Variable definitions were obtained from the applicable database reference guide and a rationale for access was assigned to each variable based on the specific use intended for the project. The Citizen, DAD, and physician billing database variables were requested for specifically New Brunswick residents aged 65+, while the provider registry was requested as all information between the years of 2006 to 2011. By providing all physician information data between the years of 2006 to 2011 without a physician age restriction, health physicians younger than 65 years of age were also included.

The difference between the feasibility application and the full data access application is the depth of detail for each inquiry. The feasibility application is a preliminary application that consisted of 3 pages followed by an invitation from the director to submit the full application. The full data access application was an 8-page document with an additional 15-page document of variables being requested. This application can be referenced within Appendix A.

3.1.3 Data Access Application Process

The data access application was submitted with the requested variable list, resumes and a letter of support to the NB-IRDT Project Coordinator to be reviewed by the NB-IRDT team

before moving on to the additional review processes. After the initial review, a project meeting was held with the review team at NB-IRDT to discuss the submitted application as well, to achieve a better understanding of the project and additional resources that may be required from NB-IRDT. Review members consisted of the NB-IRDT Manager, Project Coordinator, Privacy Officer, and several Data Analysts. This meeting was followed by a total of 4 additional revisions to the original application prior to being submitted to the Department of Health (DOH), the data owners.

A Data Research Committee (DRC) meeting was held July 19, 2019, where the data access application was reviewed with the UNB project team, NB-IRDT team, and members of DOH. This was an opportunity for DOH to gain a better understanding of the project's intent, as well as ask any questions regarding the project. The project was accepted based upon the outcome of the DRC meeting and minor updates to the application concerning additional variables were completed prior to the application moving forward.

3.1.4 Research Ethics Board Application

This research project was required to have approval from the Research Ethics Board (REB) before moving forward due to the nature of the sensitive information of individuals' health records. It was determined that there was no more than minimal risk to the individuals used in this research because there was no interaction with the individuals and these data was considered secondary information from an institution with safeguards already in place.

No surveys or questionnaires were distributed to individuals, and all personal data was obtained through the institution, a secondary source. All secondary data has already been made pseudonymous by NB-IRDT. This means that any direct identifiers, such as names,

Medicare numbers, or addresses, were removed and/or scrambled prior to receiving the information to preserve the privacy of individuals.

There is always a possibility of linking individuals due to the small population within rural areas in New Brunswick; however, pseudonymous data mitigates the concern for identifiable information. As well, the data were used throughout the research at an aggregate level to avoid identifying or linking individuals in that way.

A letter of recommendation from NB-IRDT was submitted to REB simultaneously with the submittal of the REB application. After the review process, a letter of approval from REB's Steven Turner was received with reference number 2019-109. This letter can be found in Appendix A.

With a project acceptance from REB, NB-IRDT completed a final privacy review before the NB-IRDT director granted approval to proceed. The final application was then submitted to the University's Office of Research Services (ORS) for approval before access to the requested data was permitted. During this time, the required criminal record checks were submitted, privacy training with Donna Curtis-Maillet and security training with Zikuan Liu were completed to allow access to the lab. The UNB Ucard was set up as a secure access pass to enter the lab.

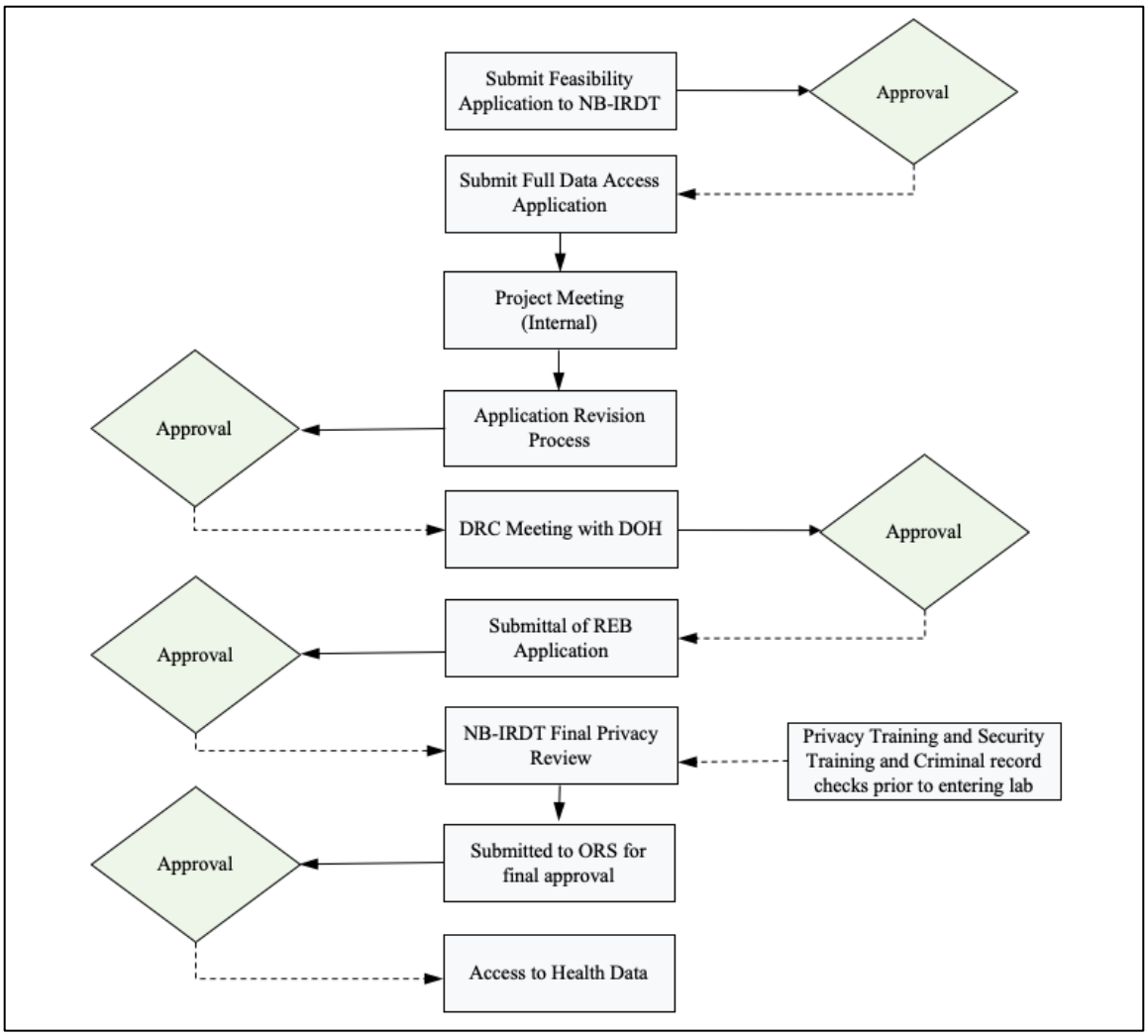

Figure 3-1 - Summarized NB-IRDT Data Approval

3.2 Working with the Administrative Data

A project profile (P0036) was created by NB-IRDT which allowed access to the requested datasets. All requested datasets and variables were approved; however, access to specific variables were staggered throughout the project to ensure no linkage of any individuals information.

NB-IRDT does not typically release individual postal code information to researchers for privacy purposes. However, to complete these travel distance calculations, postal codes of

individuals and establishments were required. To complete this portion of the project, postal code information was provided for distance calculation purposes only and was then removed once the distance calculations were completed. As a safeguard strategy, NBIRDT held back part of the requested information while travel distances were being determined; once the distance calculations were complete, the access to postal code information was removed and the remaining variables requested were then released.

Statistical Analysis Software (SAS) was used for most of the sorting and processing of the available information within the datasets. This programming language was recommended by the data analysts at NB-IRDT due to its strong ability to handle large amounts of data. Figure 3-11 (page 61), represents the process used to integrate the different datasets to create proper origin/ destination paths for health visits.

3.2.1 Determining Travel Distances

3.2.1.1 Processing Health Using SAS

There were various major steps identified when creating the final SAS datasets to be used for distance calculations; however, before these steps were completed, the following background and sorting steps took place.

The Postal Code Conversion File Plus (PCCF+) software agreement was obtained through the UNB Libraries. This file was uploaded to the NB-IDRT system into the project file P0036 by NB-IRDT's lead analyst. The PCCF+ file was used to link postal codes and individual ID numbers with geographical information from Statistics Canada's census data. This file was used in later steps when linking postal codes to corresponding dissemination areas (DA).

The citizen dataset supplied the individual scrambled ID numbers as well as an associated postal code with two dates, the effective date (*eff_date*) and the termination date (*term_date*). This meant that the postal code assigned to the ID number of an individual was considered active with the individual between these two dates. Some ID numbers were assigned multiple postal codes which meant that the individual had moved locations. For all postal codes that do not have a termination date, it was assumed that these postal codes were still active with the corresponding ID number therefore, the date of December 31, 2019, was used.

Figure 3-2 shows the code used to assign a present date to the termination date variables.

```
Data bmoreh.addr_var4dist_2;
Set bmoreh.addr_var4dist;
IF TERM_DATE = . THEN TERM_DATE = mdy(12,31,2019);
run;
```

Figure 3-2 - Example SAS Code: Creating Default Dates for Missing Termination Dates

The DAD datasets provided the individual scrambled ID number of a patient with an admission date of when that individual had checked into the hospital. To later compare the admission dates to the effective and termination dates within the citizen dataset, the admission date format was required to be changed from a character format to a numeric format as shown in Figure 3-3. For the SAS program to properly compare the variable inputs, all variable formats must be identical, this process was applied to all dataset years from 2005-2011.

```
/* DAD data 2005 */
Data bmoreh.dad2005_var4dist_2;
Set Source.dad2005_var4dist;
NewAdmdate = Input(Admdate,YYMMDD10.);
FORMAT NewAdmdate YYMMDD10.;
Rename indvid = ID;
run;
```

Figure 3-3 - Example SAS Code: DAD Admission Date Format

3.2.1.1.1 Working with the DAD datasets:

All DAD datasets from years 2005-2011 were combined into one table using the *proc seq* command in SAS, this command was used to join the tables rather than merge them because the merge command would mix up the postal codes and remove the observation lines that had "NB" as a postal code. Figure 3-4 shows the code that was used to complete this union of tables. When joining these tables, the sequence number provided within the original DAD datasets were kept within the new table as a well, a fiscal year variable was added to the table. These two variables were required to be included as identifiers for the lead analyst to later recall the information.

```
proc sql;
Create table bmoreh.newdad_2005_2011 as
select  ID,Postalcode,Instnum,NewAdmdate,seqn,  2005  as  fiscalyr  From
bmoreh.dad2005_var4dist_2
union
…
select  ID,Postalcode,Instnum,NewAdmdate,seqn,  2011  as  fiscalyr  From
bmoreh.dad2011_var4dist_2;
run;
```

Figure 3-4 - Example SAS Code: Joining Tables with Identifiers

The first step when combining the health data was to use the citizen database to complete any missing variables within the DAD datasets as some of the postal code variables were incomplete or were substituted with "NB". This process was completed by first combining the new DAD dataset (2005-2011) with the citizen dataset by individual ID number using the *merge* command. Within the Citizen dataset there are identification numbers that do

not appear within the DAD dataset because while they may live in New Brunswick, they have not been to a health care facility between 2005-2011. To account for this, a condition of a=1 and b=1 was used within the code to ensure that all identification numbers being linked were present within both data sets. This can be seen in Figure 3-5

below.

```
data bmoreh.dadaddr_2005_2011;
MERGE bmoreh.newdad_2005_2011(in=a) bmoreh.addr_var4dist_2(in=b);
by ID;
if a=1 and b=1;
run;
```

Figure 3-5 - Example SAS Code: Merging DAD and Citizen Datasets

The citizen and DAD datasets were then sorted by ID numbers and an *if-then-else* statement was used to complete the missing postal codes within the DAD dataset. This statement was based on the condition that the admission date of hospitalization (*NewAdmdate*) fell within the provided effective (*Eff_date*) and termination (*Term_date*) dates of the postal code. If this condition shown in Equation (5 was proved to be true, the postal code from the DAD dataset reading "NB" was replaced with the postal code within the citizen dataset. This was completed for all applicable observations.

$$Eff_{date} \leq NewAdmdate \leq Term_{date} \tag{5}$$

The following code within Figure 3-6 below demonstrates how these conditions were applied to the datasets using *if-then-else* statements.

```
data bmoreh.dadaddr_origins_2005_2011;
set bmoreh.dadaddr_2005_2011;
by id;
if postalcode ="NB" then
do;
 if NewAdmdate ge eff_date and NewAdmdate le term_date then
do;
 postalcode=pcode;
 success = '1' ;
end;
end;
run;
```

Figure 3-6 - Example SAS Code: Applying conditions Using if-then-else Statement

The postal codes within the DAD dataset are considered to be the origin location of the health trip. To determine the destination of the health trip, the postal code of the establishment had to be determined.

This was done by assigning the appropriate postal code to the institution number from the DAD data. To do so, the last three characters of the institution number were linked with the three-digit establishment IDs from the *establishment* dataset. This administrative postal code is used to associate a destination postal code with the health institution that was visited for the individual trip.

The institution number variable (*instnum*) from the DAD dataset consists of a four-digit code. The four digits that make up the administrative number are split into two parts: the first number representing the treatment received at the health establishment, while the last three digits represent which health establishment was providing the service.

Figure 3-7 shows the code used in SAS to determine the establishment ID number from the institution number. Using a similar process that was used for the admission dates within the DAD data, the institution number was also converted from a character format to a numeric format.

```
Data bmoreh.Final_dadaddr_origins_2005_2011;
set bmoreh.dadaddr_origins_2005_2011;
est_id = INPUT(SUBSTR(Instnum,2,3),4.);
run;
```

Figure 3-7 - Example SAS Code: Creating Establishment ID Numbers from Institution Numbers

The establishments dataset contains information regarding healthcare establishments for New Brunswick. This included establishment ID numbers, external establishment ID numbers, postal codes, city names, counties, etc. The establishment ID (*est_id*) is a unique three-digit code that is assigned to a health care location throughout the province with the exception of the code "0". A code of "0" refers to an office location such as a physician's office or a clinic that a physician is working. This code for physician offices is later presented in the physician billing datasets.

To begin linking the DAD dataset with the establishment dataset, the establishment ID number was retrieved from the institution number to create a new variable within the DAD dataset. This variable was renamed to represent the establishment ID number; this then allowed the DAD and establishment datasets to be merged by the common establishment ID variable, which assigned the proper postal code to the corresponding health location.

The DAD dataset specifically used the establishment ID numbers that are associated with the health care facility, while the external establishment ID numbers are specifically related to the Physician Billing datasets. During this process it was found that there were a number of observations that were not properly coded into the DAD database. This can be one of the challenges when using administrative datasets and because this was a small number of observations in comparison to the total observations, they were advised to be removed from the dataset.

3.2.1.1.2 Working with the Physician Billing datasets:

The next step was to input the patients' postal code with the physician billing data for years 2006-2011 by matching the individual ID and ensuring the following condition in equation (6) holds true.

$$Eff_{date} \leq Phys.Billing_{servicedate} \leq Term_{date} \tag{6}$$

Similar to the process above, combining multiple DAD datasets into one, the physician billing datasets from years 2006-2011 were combined into one dataset. As well, because the establishment ID variable within the physician billing dataset is the external establishment ID within the establishment dataset, this variable was changed to read *ext_est_id*. When combining these datasets, the following variables shown in Table 3-1 were kept in the new dataset created. As done previously with the DAD datasets, the sequence number (*seqn*) and fiscal year were variables added to the new table as identifiers.

Table 3-1 - Physician Billing Variables

VARIABLE	DESCRIPTION
Servdate	The service date that a physician was treating a patient
Ext_est_id	The external establishment ID for the facility/office the service is being billed
Provider ID	The ID number of the physician providing the service
ID	The scrambled identification number of the patient
Seqn	The sequence number used as an identifier
Fiscal year	Year of dataset being used as an identifier

The citizen and physician datasets were then sorted by ID numbers and merged based on the condition that, if the physician billing service date (*servdate*) fell within the provided effective (*Eff_date*) and termination (*Term_date*) dates of the postal codes, the postal code from the citizen dataset would be linked to the corresponding individual ID within the

physician dataset. This assigned an origin postal code within the physician billing dataset to the individual ID number that received a service from a physician. The coding process for this is shown in Figure 3-8 below; similar to the DAD dataset, the condition a=1 and b=1 was used.

```
data bmoreh.Final_addBill_2006_2011;
MERGE bmoreh.billing_2006_2011(in=a) bmoreh.addr_var4dist_2(in=b);
by ID;
if a=1 and b=1;
if servdate ge eff_date and servdate le term_date;
run;
```

Figure 3-8 - Example SAS Code: Condition to Assign Origin Postal Codes to Physician Billing Data

Next, the postal codes associated with the external establishment ID within the *establishments* dataset were inputted into the physician billing dataset by linking the external establishment IDs. This added a new variable to the physician billing dataset that included the corresponding postal codes, these postal codes were considered the destination postal codes. The external establishment IDs within the physician database were linked to available postal codes; however, many external establishment ID numbers were either coded with a "*0*" or a ".". Any external ID numbers that consisted of a "." were converted to the code "*0*", this meant that the physician provided a service at an office location that is not included within the establishment dataset.

The provider registry dataset holds information containing the provider ID numbers of the physicians, the practice location's postal code that a service was provided as well as the effective and termination dates of the practice location. Similar to the citizen database, some of these practice locations do not have a termination date; therefore, it was assumed that this practice is still at the same location or, was during the 2006-2011 time period. To address this, missing termination dates were completed by inputting a date of December

31, 2019. In addition, there were a number of missing provider ID numbers within the provider registry as well as multiple missing service dates within the data. This again, is part of the challenges experienced when working with administrative data; however, the number of missing provider numbers and service dates were minimal therefore the locations were removed from the dataset before completing the process.

To ensure that the proper destination location is associated with the correct physician providing a service on a specific date, the following condition shown in equation (8) must be met.

$$Eff\ Prac.Loc_{date} \leq Phys.Billing_{servicedate} \leq Term\ Prac.loc_{date} \tag{8}$$

The physician billing and provider registry datasets were then sorted by *provider ID* numbers and merged, using an *if-then* statement, based on the condition that if the physician billing service date (*servdate*) fell within the provided effective (*Eff_date_pract_loc*) and termination (*Term_date_pract_loc*) dates of the practice location. This linked the postal codes from the provider registry dataset to the corresponding *provider ID* number on a specific service date within the physician dataset; this coding process is shown in Figure 3-9 below.

```
data bmoreh.final_dist4phybill;
MERGE                                bmoreh.final_bill_estab_dest(in=a)
bmoreh.NB_practice_loc_var4dist(in=b);
by provider_ID;
if a=1 and b=1;
if est_id = '0' then
do;
if servdate ge EFF_DATE_PRACT_LOC and servdate le TERM_DATE_PRACT_LOC
then
do;
est_pcode = ploc_codes;
end;
end;
IF ID = '' THEN DELETE;
IF provider_ID = '' THEN DELETE;
IF servdate = '.' THEN DELETE;
run;
```

Figure 3-9 - Example SAS Code: if-then Statement to Assign Destination Postal Codes to Provider ID

It should be noted that the physician billing and DAD datasets are two very different datasets and cannot be merged; these datasets need to work independently of each other and always be two separate data tables. They could later be joined instead of merged to become one table (this puts the data on top of one another); however, if this is done then an additional identifier similar to the sequence number and fiscal years must be included. To avoid adding any additional identifiers, these two datasets were kept separate throughout this process.

The PCCF+ file was then used in conjunction with the DAD dataset to assign a DA number (*DAUID*) to the origin postal codes within the DAD dataset. This was done by merging the datasets by ID numbers if the postal codes from each dataset were equal. Figure 3-10 shows how this command was completed within SAS.

```
Data bmoreh.dad_dist4travel;
MERGE bmoreh.final2_dist4DAD bmoreh.Org_hlth;
BY ID;
IF origin_pcode = PCODE;
IF ID = '' THEN DELETE;
label origin_pcode = 'Origin Postal Code';
label dest_pcode = 'Destination Postal Code';
label est_id = 'Establishment ID';
label DAuid = 'Origin DAuid';
run;
```

Figure 3-10 - Example SAS Code: Assigning a DA Number to Origin Postal Codes

The same process within Figure 3-10 was completed using the physician billing dataset and the PCCF+ file except, in addition to linking the DA number to the origin postal codes, the corresponding coordinates (*latitude, longitude*) from the PCCF+ file were assigned to the destination postal codes. This allowed the physician billing data to later be exported into a csv file to be used as physician office locations within ArcGIS. Prior to exporting, a second file was created where all duplicates were removed, this is to avoid having duplicate offices in ArcGIS because multiple visits are occurring at the same location.

The following Figure 3-11 represents the process used to integrate the different datasets to create a proper file that could be used within Arc GIS to determine origin/ destination paths for health visits.

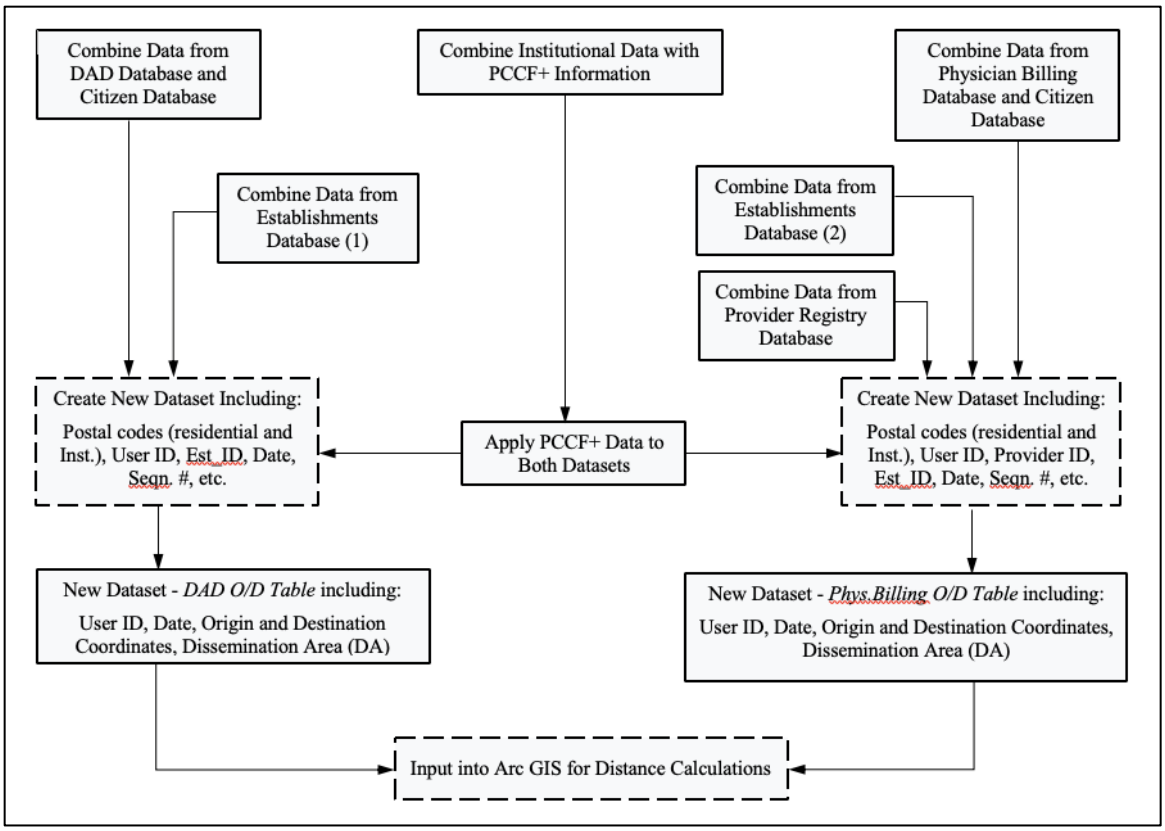

Figure 3-11 - SAS Dataset Process - Part 1

3.2.1.2 Processing Travel Distances using ArcGIS Software

New Brunswick data files and layers were obtained from the DMTI CanMap 2011 files provided by NB-IRDT. The information within the DMTI package used for this analysis included the New Brunswick 2006 and 2011 census DA layers, various points of interest which include the NB health care facility locations, and the route logistics network layers for the province.

The Provincial layers were imported into the ArcGIS program, ArcMap, to create the foundation of the GIS model. All layers, except the 2006 DA layer, were imported into the program with a default coordinate system of GCS_North_American_1983, this is also known as NAD 1983. The 2006 DA layer's default coordinate system was in the

GCS_North_American_1983_CSRS. All layers must be operating on the same coordinate system to ensure that geographic locations are being placed properly using the correct coordinates; therefore, the 2006 DA layer was transformed into the NAD 1983 coordinate system to match the coordinate system of all other layers. Using the projections and transformations tool within the data management toolbox, this created a new 2006 DA layer that operated with the NAD 1983 coordinate system.

The features to point tool within the data management toolbox was then used to determine geographical centroid locations within each DA. The 2006 DA layer was used as the input feature layer within the *feature to point* tool and the box for "*inside (optional)*" was checked, this was to ensure that all the centroids were created inside the DA polygons. This tool created a new layer of 2006 DA centroids that were added to the map called *DA_2006_centroids layer*. There was minimal difference between the 2006 and 2011 DA layers therefore because the health data dated back to 2006, the 2006 DA layer was used to develop the DA based centroids. It was recognized that determining population-based centroids may provide a more accurate representation of rural areas with larger DA polygons; however, for the purpose of this project, geographic centroids were used to expedite analysis.

The provincial road shapefile (NBrte) was converted into a network database file using the Arc Catalog. Typically, a shapefile is converted into a feature class before it can be transformed into a database; however, this shapefile was already in a feature class format. Attributes of this new road network included globally restricted turns, measurements in meters, and no elevations were able to be selected. This network database was then added to the map being created in ArcMap.

New Brunswick health care facilities layer included all hospitals and major clinics within the province. The information within the facilities table matched the information provided within the NB-IRDT *establishments* dataset used previously during the SAS processing. This layer included essential identification variables such as establishment ID, external establishment ID, name of the facility, postal codes, City, etc. This layer was added to the map populating with a total of 104 locations. The establishment IDs within the facilities table will link with the DAD datasets previously associated with the corresponding health facility. Some of the physician datasets were able to be linked with external establishment IDs; however, the majority of those IDs were a "0" which again, indicates the physician was at their office. This means that some of the physician data would be linked with the facilities layer; however, a second layer was required to be created to capture the physicians' offices that would be missed within the facilities layer.

To create the physician location layer in ArcMap, the final version of the physician billing dataset was first sorted to remove duplicates and then exported into a csv file containing the DA ID number, known as the DAuid, corresponding postal codes, and the latitude and longitude coordinates. Arc Catalogue was used to create a future class using the csv file via the XY table option using the longitudes and latitudes respectively, to populate the x and y table. A shapefile was then created to be used in ArcMap and then created into a layer. No clinic or office names were available to be linked with the new layer; therefore, the physician's location was labeled with the associated postal code. The new physician location layer populated a total of 803 locations over the province.

The OD Matrix tool within the Network Analysis toolbox was used to determine all possible distances between DA centroids (origins) and health care locations (destinations).

This tool uses the road network to determine distances between the DA centroids and health care locations by routing the travel path on the road network. The DMTI road file that was converted into a network database was used within the OD Matrix tool and included the following attributes: U-turns were restricted to travelers and the impedance factor was set to the length of travel (km).

A total of 1439 origin locations were added to the origin feature within the OD Matrix window using the *DA_2006_centroids layer*. The location analysis table was used to alter the name field and set it to DAuid from the *DA_2006_centroids layer*. This assigned the DA number to the centroid name rather than assigning it a generic location number. This was done to easily identify a DA location and to provide a common identifier when linking tables in the future. All other default values within the fields of the Location Analysis tables remained the same. The field for location position was set to geometry and a search radius of 400km was applied to ensure that all distances would be accounted for when searching for destinations and none would be omitted based on a distance limitation. To determine an appropriate search radius, approximate distances across the province were estimated. An offset from road networks was set to 1km to try and include any centroids that might be out of range of the road network.

When adding the origin locations, a warning appeared for a total of 34 origin locations that were on a *non-traversable network position*. These DA locations were not connected to the network because they were "*a location with a non-traverse element position"*. The majority of these 34 locations are all within major city limits, such as the Saint John or Fredericton area. It was decided that these 34 locations were able to be omitted from the analysis because these locations were all within a minimal travel distance to a nearby hospital. In

addition, because these centroids were in urban areas, all of these error locations were within close proximity to many other DA centroids. For the purpose of this project, it was assumed that were would be a minimal difference between distances calculated using the missing DA centroid and a nearby DA centroid that was included in the analysis. This research is focused on rural areas in New Brunswick; therefore, it is assumed that these 34 missing locations within the City's CMAs should not have a large impact on these results.

The healthcare facilities layer and physician locations layers were added to the destination locations within the OD Matrix. This populated the destinations with a total of 907 locations including, 104 facility destinations and 803 physician locations. Warnings 38 destinations that were unavailable for routing as they were also on "*a non-traversable network element position"*. Of the 38 destinations, 1 destination was the *PEM unité de Grand-Sault* from the facilities layer. It was decided that for this application, the small percentage of error when adding the destinations was acceptable.

The OD Matrix tool was run after all origin and destination locations were added, this resulted in a total of 1,200,769 successful distance combinations between the total origin and destinations and a total of 43 errors due to the previous origin and destination errors. A summary of these outcomes is shown below in Table 3-2.

Table 3-2 – Summary of Arc Map OD-Matrix Outputs

Total Origins (DA)	1,439
Health Facilities	104
Physician Billing Locations	803
Total Destinations	**907**
Errors or Origins	34
Errors for Destinations	38
Errors for Travel	43
Total successful Combinations from ArcMap	**1,200,769**

A "lines" table was created from running the OD Matrix tool, this table included all the travel combinations between origin and destinations with a total length (km) column that had been calculated from routing the distances. In addition, a lines layer was created in the map as a result of this process, even though the distances were calculated using a routing tool, this layer used straight lines to represent distances between the origin and destination locations.

The lines table included serval variables; however, for this purpose, the important variables being used were the name of the travel locations (*DA number – name of destination facility*), the total distance between the two locations as well as the software's default origin and destination ID numbers. These ID numbers were used to join the origin, destination, NB facilities, and lines tables together to create a table that could be linked with the SAS datasets.

To create a table that was compatible with the SAS datasets previously created, the NB facilities, origin, destination, and the lines tables were all combined. This was completed in the following steps.

1. The lines table was joined with the common variable from the origins table, the *Origin ID* (from the lines table), and the *Object ID* (from the origin table). This assigned the *DAuid* number to the distance calculations calculated from the lines table. Prior to joining these tables, the variables were both sorted in ascending order to ensure they were linked properly. This assigned the proper variables from the origin table to the lines table ensuring a DA was assigned to the correct distance.
2. The lines table was then joined with the destinations table by a common variable, the *Destination ID* (from the lines table) and the *Object ID* (From the designations table). This assigned destination locations, depicted as a postal code, to the appropriate distance calculations from the lines table. These variables were also sorted in ascending order to ensure they would be linked properly.

This table now has Origin DA numbers, destination postal codes and the associated distance traveled from origin to destination. The physician locations and facilities table was also joined to add additional information to the destinations such as destination city names, county names, destination DA numbers, and coordinates using the common variable, *destination postal code*.

Before each join was processed, a "*join validation*" was run to check for any common errors when joining. When joining these tables, it was important to make sure that each variable being linked was sorted in the same order as the other variable being linked. The data within the ArcMap output was not able to be released due to the presence of postal codes; however, the output structure of the final table is shown in Table 3-3.

Table 3-3 – ArcMap Final Output Structure

Obj.ID	Travel location	Org. ID	Org. DA	Dest. ID	Dest. Pcode	Dest. County	Dest. CMA	Dest. DA	Length (km)

It is also important to note that when joining these tables by fields in other tables, the "fields" are required to be the same structure type. This means that object ID fields can be joined with other object ID fields and string fields, such as text (names) can be joined with other string fields; however, object ID fields are unable to join with string fields.

Important variables that were linked to the lines table from the NB facilities table were the facility's *postal codes, establishment IDs, and external establishment IDs*. These three variables, along with the *DAUID* variable from the origin table, can be linked with the SAS data sets to assign distances to the actual health trips taken. The final table consisted of many variables from all four tables, the non-relevant variables were hidden and the remaining variables within the table were exported into a text (csv) file. The following Table 3-4 shows which variables were included within the final table for linking with the SAS datasets as well as explains what each variable is used for. The variables *POST_CODE* and *EstblPosta* were removed because they provided identical postal code information as the *POSTALCODE* variable, because these were removed, the status is "drop". The *POSTALCODE* variable was kept over the other two variables as it had consistent formatting as the postal code variables in the SAS datasets.

Table 3-4 - Final ArcMap Distance Table for SAS

Variable Name	Description	Status
Object ID	Object ID from lines table	KEEP
Name	Locations traveled for distance calculations "*DA# – Location Name*"	KEEP
Origin ID	Origin ID from lines table	KEEP
Destination ID	Destination ID from lines table	KEEP
Total Length	Total length or distance traveled between location (m)	KEEP
Object ID_1	Object ID from origin table	KEEP
Name_1	DAUID	KEEP
Object ID_12	Object ID from the destination table	KEEP
Name_12	Facility Location Name	KEEP
POSTALCODE	Postal code with no Spaces between FSA and LDU	KEEP
POST_CODE	Postal code with space between FSA and LDU	DROP
Establ ID	Establishment ID of Facility – used within DAD data	KEEP
City	The city the facility is located in	KEEP
EstblPosta	Postal code of establishment no spaces between FSA and LDU	DROP
External_E	External Establishment ID of facility – often used with physician billing data if not a physician office	KEEP

The exported text file was then imported to the SAS software using the following code shown in Figure 3-12, this code created a dataset table with all the variables shown above in Table 3-4.

```
/* B. Bring Text file 'arcmapdistances'
in from arc map
*/
filename                          arcmap
"K:\p0036_travel_demand\users\Brianna
Morehouse\ArcMap\arcmapdistances.txt";

data bmoreh.arcmapdistances;
/*format table to read properly*/
format
      ObjectID
      Name              $200.
      OriginID          $200.
      DestinationID     $200.
      Total_Length      14.4
      ObjectID_1        $200.
      Name_1            $200.
      Name_12           $200.
;

infile  arcmap  dsd  dlm=","  truncover
firstobs= 2;
/*input file is what is read from the
txt file*/
input
      ObjectID          $
      Name              $
      OriginID          $
      DestinationID     $
      Total_Length
      ObjectID_1        $
      Name_1            $
      Name_12           $
;
/*drop the following variables from the
table (to clean it up)*/
DROP
      ObjectID
      Name
      OriginID
      DestinationID
      ObjectID_1;

Rename Name_1 = DAuid;
Rename Name_12 = dest_pcode;
Rename Total_Length = distance
;
run;
```

Figure 3-12 - Example Code for ArcMap Text File to SAS Dataset

Both the facility and physician locations files were used to calculate the total number of distance combinations; therefore a check for duplicates based on *DAuid* and *dest_pcode*

was run to remove any duplicated distance combination. This removed a total of 108,654 distances leaving 1,902,115 distance combinations to be applied to the health datasets.

The distances were formatted as shown below in Figure 3-13 - Example SAS: Correcting Distance Data to Merge with SAS Datasets to be able to link the distances with the associated distance.

```
/* Make sure variables are the correct length and data type to match
dad data set
Dauid = length of 8
dest_pcode length of 8 */

data bmoreh.arcmapdistances2;
length dauid $8. dest_pcode $8.;
set bmoreh.arcmapdistances;
len = round(distance / 1000);
DAuid = compress(DAuid);
dest_pcode = compress(dest_pcode);
run;
```

Figure 3-13 - Example SAS: Correcting Distance Data to Merge with SAS Datasets

The DAD dataset was then merged with the distance file from ArcMap to assign a distance traveled for each health trip in the dataset that has the corresponding origin and destination combination. When merging the DAD dataset with the distance file it is important to merge by both the *DAuid* and the *dest_pcode* variables rather than just with one. This ensures that the correct origin-destination combination is linked with the proper distance. Figure 3-14 below shows the example for combining the ArcMap distances with the DAD dataset. When combining these files, a total of 276,565 health trips out of 291,353 trips were able to be linked with a corresponding distance. This provided approximately a 95% match with the DAD data.

```
Data bmoreh.Total_dist4dad;
MERGE bmoreh.dad_dist4travel2 bmoreh.arcmapdistances2;
BY DAuid dest_pcode;
label org_pcode = 'Origin Postal Code';
label DAuid = 'Origin DAuid';
label distance = 'Travel Distance';
label len = 'Disance (km)';
IF ID = '' THEN DELETE;
IF distance = '.' THEN DELETE;
run;

/* dad_dist4travel2: 291,353
arcmapdistances: 1,092,115
Total_dist4dad Observations: 276,565 */
```

Figure 3-14 - SAS Example: Combining ArcMap Distances to Health Trips

Similarly, to the DAD dataset, the Physician Billing dataset was combined with the ArcMap distances resulting in matching 3,575,991 health trips out of 4,465,759 being able to be linked to ArcMap distances. This provided an approximately 80% match with the physician billing data which was found to be acceptable within the scope of this project.

This process produced two final datasets with various variables, most importantly including an identification number, an origin, a destination, a date of visit, and an associated travel distance in kilometers. It was then possible for NB-IRDT to remove all access to datasets that included postal code information and remove any postal codes or identifiers from the two final datasets to be used in the next steps.

3.3 Health Variables

Once the postal code information was removed, the second half of the DAD and Physician Billing datasets were released to the project. These datasets contained the requested health variables which were associated with the health trips from the original DAD and Physician Billing datasets which were later combined with additional datasets.

3.3.1 Health Variables from the DAD Data

The requested health variables found in Appendix A were included in the second part of the DAD dataset. Similar to the process with the initial DAD dataset, the datasets from years 2005 to 2011 were combined by the individual's ID number and sequence number, into one dataset including 2005-2011 fiscal year data (*dad2005_2011*). This new dataset included multiple variables outlining the details of each health trip, including a chain of different provider ID numbers (1-8), provider type (1-8), provider service (1-8), intervention codes (1-20), intervention provider service (1-20), diagnostic type (1-25) and diagnostic codes (1-25). These variables are used to navigate how the individual moves throughout the health system and what type of care they received from various physicians. Many of these provider, intervention, and diagnostic codes were found to be missing observations as the variable number increased. This meant that that typically, after a few transfers between physicians or several diagnoses, the individual did not continue to move to different physicians, or a diagnosis was determined. These variables were determined to be too in-depth for the scope of this project; therefore, provider ID numbers, type and services 3-8, diagnostic types and codes 3-25, and intervention codes and provider services 2-20 were removed. The initial provider ID number, service and type, diagnostic type and code, and intervention codes and provider type were included in the dataset to determine the individual's original purpose of travel. The previous dataset with the DAD health trips and associated travel distances (*dist4dad_nopc*) were then combined with the new DAD dataset (*dad2005_2011*) containing the requested health variables. This created a new DAD dataset (*dad_dist*) that was then combined with additional datasets provided by NB-IRDT to link important data for travel purposes as shown in Figure 3-15.

The new DAD dataset was first combined with the *Register* dataset (*reg_final*) and the *address* dataset (*addr_final*). This allowed for information about the individual making a health trip and their origin, to be linked to the according health trip observations. This included variables such as geolocation, originating health region, QAIPPE, etc. Before combining the address data with the DAD data, if a geolocation code was = 0 or 99997 it was removed as a code of 0 represents an individual from out of province and a code of 99997 represents an unknown location. If a termination date was missing, the date of December 31, 2019, was inputted to assume the address is still valid for that time period. In order for the second half of the address data to be linked to the appropriate individual, the same condition used in equation (5) was applied. This ensured that the individual's address information was valid during the time of the admission date of the health trip.

A new dataset, Specialty History 2018, was used to assign physician specialties to a provider number within the DAD dataset. This provided the effective dates and the termination dates of which physicians were practicing in specific fields. If a termination date was missing, the most recent date of December 31, 2019, was used to represent the physician still practicing within that specialty. If a provider ID number was missing, the observation was deleted. Provider ID "1" within the DAD dataset was changed to *provider ID* to be able to be combined with the Physician Specialty dataset. This assumed that the specialty of the first physician assigned to the individual was the purpose of the health trip.

To combine the specialty history with the existing DAD dataset, the two were combined by the provider ID variable, if the admission date was greater than or equal to the effective date of specialty and less than or equal to the termination date of specialty.

The physician provider database was then combined with the most recent dataset by the provider ID variable to assign additional information about the physicians providing a health service.

Finally, the dataset was linked with the *establishments* dataset by the establishment ID number. This assigned the maximum number of beds that were available at the facility of the health trip. If the number of beds was indicated as missing with a "." it was assumed to be a total of 0 beds. Any remaining observations that missing individual ID numbers missing were deleted. Individuals who were less than 65 years of age were deleted.

The final DAD dataset consisted of 171,467 observations between the fiscal years of 2005-2011. This included observations that were able to be linked with a travel distance for individuals 65 years or older, making a health trip to the hospital. Health trips to physician offices and clinics are not included within this dataset, nor are individuals who were unsuccessfully matched through the ArcMap process. Due to the nature of using administrative data, some observations were removed during the linking process because of missing data, distance traveled unavailable or unsuccessful links from merging conditions. This results in multiple observations present in the original datasets to be missing from the final datasets through various merging conditions, missing data. Full detail of the syntax used to process this dataset using SAS can be found within Appendix B.

3.3.2 Health Variables from the Physician Billing Data

The final Physician Billing dataset was linked with the requested health variables from Appendix A, and the associated datasets in the same way previously described for the DAD

data as shown in Figure 3-15. The Physician Billing data did not provide the same type of variables as the DAD data such as intervention codes, diagnostic types, etc. However, requested variables that would help determine the purpose of travel, origin and destination information, etc. were processed to create a Physician Billing dataset with health variables. Like the DAD dataset, any persons less than 65 years of age were removed from the Physician Billing data set

The final Physician Billing dataset consisted of 2,882,220 observations between the calendar years of 2006-2011. This included observations that were able to be linked with a travel distance for individuals 65 years or older, making a health trip to a physician's office/ clinic location. Health trips for individuals who were unsuccessfully matched through the ArcMap process or did not have a provider ID associated with their health trip were not included in this dataset. Due to the nature of using administrative data, some observations were removed during the linking process because of missing data, distance traveled unavailable or unsuccessful links from merging conditions. This results in multiple observations present in the original datasets to be missing from the final datasets through various merging conditions, missing data. In addition, some observations within this dataset were kept in the data as valid based on the presence of the major variables, however, some observations may have been missing a variable. Although a minimal amount, this created some "missing observations" later in the analysis when creating origin-destination tables and classification tables. Full detail of the syntax used to process this dataset using SAS can be found within Appendix B.

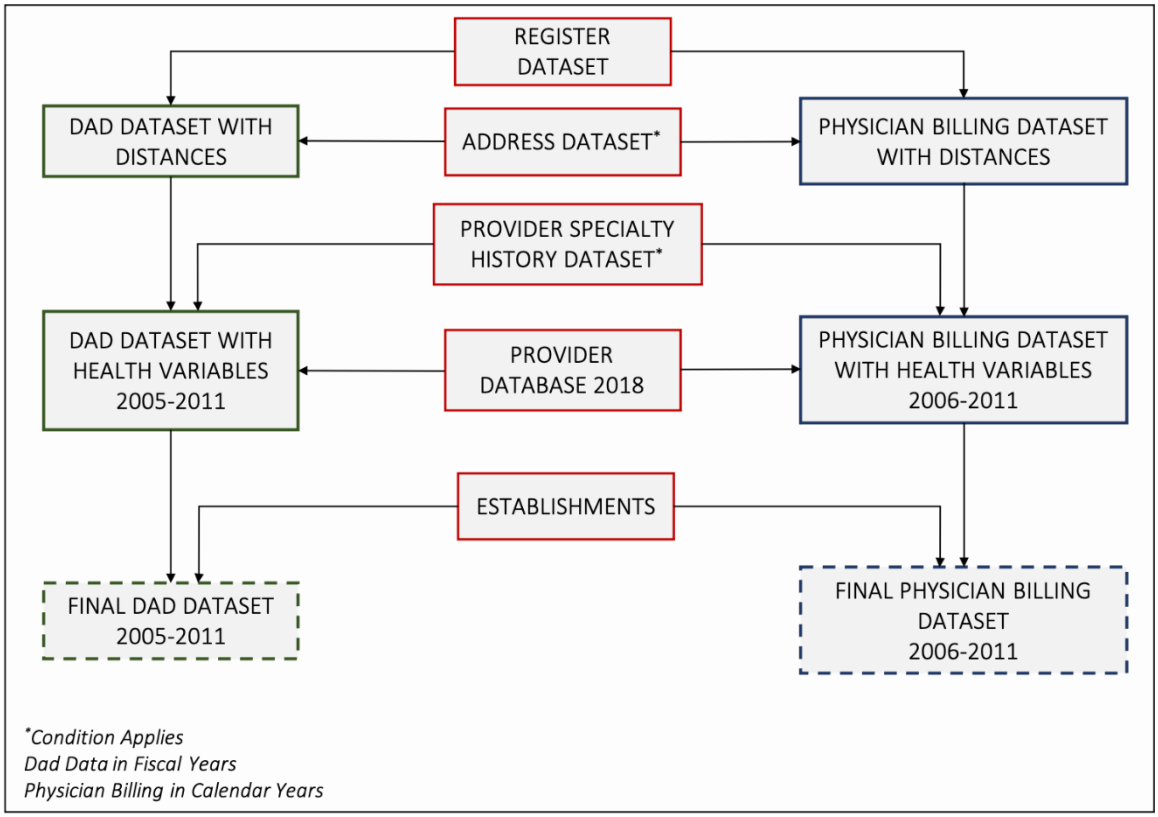

Figure 3-15 - Process to Combine Health Variables with DAD and Physician Billing Data

3.4 Determining Descriptive Statistics

Five main questions were used to develop descriptive statistics using the final data sets from both the DAD data and the Physician Billing data. The questions were as follows:

1. **Origins:** Where are health trips being generated from and how many?
2. **Destinations:** Where are health trips being generated to and how many?
3. **Frequency:** How often are health trips being taken and by who?
4. **Distance:** How far are individuals travelling for a health trip?
5. **Trip Purpose:** What type of health trips are individuals travelling for?

3.4.1 Origin-Destination (OD) Matrix:

To answer the first two questions, origin-destination matrix (OD matrix) tables were created with the data for each of the DAD and Physician Billing datasets. This was done

using a frequency function within the SAS program to generate frequency tables showing the number of trips between origin and destination areas as shown in Figure 3-16.

```
Proc freq data = bmoreh.dad_prvbase_est;
tables CDname*dest_city/ NOROW NOCOL NOPERCENT;
tables HRename*d_hlth_region_cur/ NOROW NOCOL NOPERCENT;
run;
```

Figure 3-16 - Example SAS: DAD Data - Creating Origin-Destination Tables

The trip tables for the DAD dataset were created using the origin county variable and the destination city variable. This showed the number of health trips originating from a county to a hospital within a destination city. Later on, to complete the origin-destination table for the DAD data, the destination city was grouped into the corresponding county to create a county-to-county matrix. The Physician Billing data set used the origin county and destination county to create an origin-destination trip table.

OD matrix tables were also created using the origin health region variable and the destination health region variable for both the DAD data and the Physician Billing data.

A total of seven origin-destination tables were created for each fiscal year between the years 2005 to 2011 and a single origin-destination table was also created using the combined 2005-2011 data for the DAD dataset. A total of six OD matrix tables were created for each calendar year between the years 2006 to 2011 and a single OD matrix table was also created using the combined 2006-2011 data for the Physician Billing dataset.

This was completed for both the origin-destination counties and the origin-destination health regions.

3.4.2 Frequency Tables:

The frequency of trips between origin and destinations was captured within the OD matrix; however, trip tables were also created to demonstrate the frequency of health trips for various age groups and genders from the data.

A total of two genders, male (M) and female (F) were used, as well as a total of six age groups, including 65-69, 70-74, 75-79, 80-84, 85-89, and 90+. A sample table is shown below to demonstrate how frequencies were determined.

Table 3-5 - Frequency Table Example

	65-69	70-74	75-79	80-84	85-89	90+	Total
M							
F							
Total							

In addition to these two main groupings, additional factors were included to better understand trip frequency including the type of health trip, general practice or specialty practice, and origin county of health trip.

3.4.3 Distance and Purpose of Travel Tables:

Distance distributions were determined using the frequency function based on length of travel in kilometers. This provided a table of distances ranging from 0km to 437km, where 0km means less than 1km, with a number of frequencies related to each distance. These distributions were created for the following four scenarios for both the DAD and Physician Billing datasets:

1. General frequency of distances (km) for New Brunswick.
2. Frequency of distances (km) base on New Brunswick Age Groups

3. Frequency of distances (km) based on type of trip (GP vs SP)

4. Frequency of distance (km) based on origin county

Distance statistics were also generated using the proc means function in SAS as shown below in Figure 3-17.

```
Proc means data = bmoreh.dad_prvbase_est MEAN MEDIAN MODE RANGE MAX MIN
N NMISS;
class Specialty_Type_Cur;
var len;
run;
```

Figure 3-17 - Example SAS: DAD Data - Distance Statistics

This created tables showing the total number of observations for each class category (eg. Specialty type) and calculated the mean distance travel (km) for that category based on the number of individuals making that health trip. This was created for the following scenarios for the DAD dataset:

1. Mean distance (km) traveled by age group and gender
2. Mean distance (km) traveled by County and age group
3. Mean distance (km) traveled by type of specialty and County
4. Mean distance (km) traveled by Major specialty and age group
5. Mean distance (km) traveled by minor specialty
6. Mean distance (km) traveled by specialty group
7. Mean distance (km) traveled by health region
8. Mean distance (km) traveled by income

Similarly, to the scenarios for the DAD data, the following scenarios were created using the variables from the Physician Billing Dataset:

1. Mean distance (km) traveled by age group and gender
2. Mean distance (km) traveled by County and age group
3. Mean distance (km) traveled by type of specialty and County
4. Mean distance (km) traveled by Major specialty and age group
5. Mean distance (km) traveled by minor specialty
6. Mean distance (km) traveled by specialty group
7. Mean distance (km) traveled by service type
8. Mean distance (km) traveled by health region
9. Mean distance (km) traveled by income

Although the DAD and Physician Billing Datasets do not have the exact variables when creating descriptive statistics identical variables between the two datasets were selected to be used to create statistics that could be compared. The descriptive statistics scenarios for both the DAD and Physician Billing datasets were later sorted and compared to each other with matching variables.

3.5 Linear Regression Models

Linear regression models were developed using the DAD dataset and the Physician Billing datasets with a common dependent variable being distance traveled in kilometers. Two models were created, each using forward stepwise regression, to predict distance traveled for a health care trip based on various variables included in the model. Two separate models were developed to represent each dataset because the DAD data are exclusive to health trips travelling to hospitals, while the Physician Billing data include health trips to a physician's office or clinic.

3.5.1 Creating Regression Variables

Variables were selected from the DAD and Physician Billing datasets to create two datasets with potential independent variables for regression. Similar variables were selected from each the DAD and Physician Billing data to make the models similar for comparison; however, some variables were not available in both datasets. Continuous variables such as age, the maximum number of beds, and length of stay were all included in the same format that they appeared in the data.

New variables were developed for categorical data as shown in Table 3-6 and Table 3-7. Variables with two options, for example, the specialty being general practice (GP) or

specialty practice (SP), were given either a 0 or 1 to indicate which variable was present. For variables with more than two categories, a new variable was created for each category except the last, which was removed from the list of variables as a "dummy variable". A value of 0 or 1 was used for each variable to indicate if the variable was present, if all values were 0 for a category, it is indicated that the "dummy variable" is present. This was done using IF-THEN statements in SAS as shown in the example in Figure 3-18. The remaining syntax for creating regression variables can be found within Appendix B.

```
data stats.dadvar4reg;
set stats.dadvar4reg;

/* Only two options - create new variable, 1 or 0 */
IF Specialty_Typ_Cd_Cur = '100' THEN DO PHY_SP1 = 0;
      end;
IF Specialty_Typ_Cd_Cur = '200' THEN DO PHY_SP1 = 1;
      end;

/* Total of 5 income Categories - create new variables Q1-Q4.
QAIPPE n=4 */
IF QAIPPE = '1' THEN DO Q1 = 1;
      end;
IF QAIPPE = '2' THEN DO Q2 = 1;
      end;
IF QAIPPE = '3' THEN DO Q3 = 1;
      end;
IF QAIPPE = '4' THEN DO Q4 = 1;
      end;
run;

/* If all new variable have '.' Then category 5 (dummy variable)
Make 0 for all categories to create dummy variable. */
data stats.dadvar4reg;
set stats.dadvar4reg;
IF Q1 = '.' THEN DO Q1=0;
      end;
IF Q2 = '.' THEN DO Q2=0;
      end;
IF Q3 = '.' THEN DO Q3=0;
      end;
IF Q4 = '.' THEN DO Q4=0;
      end;
run;
```

Figure 3-18 - Example SAS: Creating Dummy Variables

A correlation function was run with the new independent variables and the “len” variable as the dependent variable for distance traveled (km). From this matrix variables that have a strong correlation with distance traveled for health trips were represented.

The independent variables and their correlation with the dependent variable, distance (km), from the DAD data are shown in Table 3-6. The independent variables and their correlation with the dependent variable, distance (km), from the Physician Billing data are shown in Table 3-7. These correlations were used to determine the order of variables to first be applied to the regression analysis.

Table 3-6 - Independent Variables Used to Develop DAD Health Model

DAD DATA				
HEALTH MODEL INDEPENDENT VARIABLES				
Variable	**Correlation**	**Regression Variable Description**	**Variable Meaning**	**Type**
Age	-0.131	Age of Individual at time of visit	Numeric Age	Continuous
Max_beds	0.250	Max number of Beds at facility	Number of beds	Continuous
len_of_stay	-0.042	Length of stay for individual during visit	Number of days	Continuous
gender 2	0.076	Variable created for gender	Male = 1, Female = 0	Categorical
Q1	0.007	Variable created for income, QAIPPE = 1	1 = Low Income	Categorical
Q2	0.004	Variable created for income, QAIPPE = 2	2 = Low-Medium Income	Categorical
Q3	0.012	Variable created for income, QAIPPE = 3	3 = Medium Income	Categorical
Q4	0.006	Variable created for income, QAIPPE = 4	4 = Medium-High Income	Categorical
Q5[1]	n/a	Dummy variable for income, QAIPPE = 5	5 = High Income	Categorical
PHY_SP1	0.267	Variable created for "Specialty_Typ_Cd_Cur"	Code 100 = 0, Code 200 = 1	Categorical
Admitcat	-0.059	Variable created for "Admitcat"	Elective (L) = 0, Urgent (U) = 1	Categorical
MAJOR1	-0.267	Variable created for Specialty_Major_Grp_Code_Cur = 11000	11000 = General Prac.	Categorical
MAJOR2	0.308	Variable created for Specialty_Major_Grp_Code_Cur = 12000	12000 = Med. Specialty	Categorical
MAJOR 3[1]	n/a	Dummy Variable for Specialty_Major_Grp_Code_Cur = 13000	13000 = Surgical Specialty	Categorical
dCITYv3	0.198	Variable created for "v3_City_Urb_Rur_Ind_Cur"	Rural (R) = 0, Urban (U) = 1	Categorical
dCITYv7	0.122	Variable created for "v7_City_Urb_Rur_Ind_Cur"	Rural (R) = 0, Urban (U) = 1	Categorical
CD1	0.023	Variable created for Origin County name	Albert	Categorical
CD2	0.035	Variable created for Origin County name	Carleton	Categorical
CD3	0.171	Variable created for Origin County name	Charlotte	Categorical
CD4	0.121	Variable created for Origin County name	Gloucester	Categorical
CD5	0.036	Variable created for Origin County name	Kent	Categorical
CD6	-0.027	Variable created for Origin County name	Kings	Categorical
CD7	-0.068	Variable created for Origin County name	Madawaska	Categorical
CD8	0.058	Variable created for Origin County name	Northumberland	Categorical
CD9	0.046	Variable created for Origin County name	Queens	Categorical
CD10	0.070	Variable created for Origin County name	Restigouche	Categorical
CD11	-0.089	Variable created for Origin County name	Saint John	Categorical
CD12	-0.046	Variable created for Origin County name	Sunbury	Categorical
CD13	0.070	Variable created for Origin County name	Victoria	Categorical
CD14	-0.196	Variable created for Origin County name	Westmorland	Categorical
CD15[1]	n/a	Dummy variable for Origin County Name	York	Categorical
HR1	-0.141	Variable created for health_region = 1	Moncton	Categorical
HR2	0.013	Variable created for health_region = 2	Saint John	Categorical
HR3	0.003	Variable created for health_region = 3	Fredericton	Categorical
HR4	-0.030	Variable created for health_region = 4	Edmundson	Categorical
HR5	0.061	Variable created for health_region = 5	Campbellton	Categorical
HR6	0.122	Variable created for health_region = 6	Bathurst	Categorical
HR7[1]	n/a	Dummy variable for health_region = 7	Miramichi	Categorical
INST_TYPE1	0.047	Variable created for Insttype = A	Day Surgery	Categorical
INST_TYPE2	-0.028	Variable created for Insttype = 1	Acute Care	Categorical
INST_TYPE3	-0.014	Variable created for Insttype = 2	Chronic Care	Categorical
INST_TYPE4	-0.049	Variable created for Insttype = 3	General rehab facility	Categorical
INST_TYPE5[1]	n/a	Dummy variable created for Insttype = 5	Psychiatric Facility	Categorical

[1]*Dummy Variable*

Table 3-7 - Independent Variables Used to Develop Physician Billing Health Model

PHYSICIAN BILLING DATA				
HEALTH MODEL INDEPENDENT VARIABLES				
Variable	**Correlation**	**Description**	**Variable Meaning**	**Variable Type**
Age	-0.040	Age of Individual at time of visit	Numeric Age	Continuous
Max_beds	0.214	Max number of Beds at a facility	Number of beds	Continuous
gender 2	0.031	Variable created for income, QAIPPE = 1	Male = 1, Female = 0	Categorical
Q1	0.015	Variable created for income, QAIPPE = 2	1 = Low Income	Categorical
Q2	0.015	Variable created for income, QAIPPE = 3	2 = Low-Medium Income	Categorical
Q3	0.008	Variable created for income, QAIPPE = 4	3 = Medium Income	Categorical
Q4	0.003	Dummy variable for income, QAIPPE = 5	4 = Medium-High Income	Categorical
Q5[1]	n/a	Variable created for income, QAIPPE = 1	5 = High Income	Categorical
PHY_SP1	0.30233	Variable created for "Specialty_Typ_Cd_Cur"	Code 100 = 0, Code 200 = 1	Categorical
MAJOR1	-0.302	Variable created for Specialty_Major_Grp_Code_Cur = 11000	11000 = General Prac.	Categorical
MAJOR2	0.234	Variable created for Specialty_Major_Grp_Code_Cur = 12000	12000 = Med. Specialty	Categorical
MAJOR 3[1]	n/a	Dummy Variable for Specialty_Major_Grp_Code_Cur = 13000	13000 = Surgical Specialty	Categorical
dCITYv3	0.122	Variable created for "v3_City_Urb_Rur_Ind_Cur"	Rural (R) = 0, Urban (U) = 1	Categorical
dCITYv7	0.10312	Variable created for "v7_City_Urb_Rur_Ind_Cur"	Rural (R) = 0, Urban (U) = 1	Categorical
CD1	0.028	Variable created for Origin County name	Albert	Categorical
CD2	0.027	Variable created for Origin County name	Carleton	Categorical
CD3	0.169	Variable created for Origin County name	Charlotte	Categorical
CD4	-0.003	Variable created for Origin County name	Gloucester	Categorical
CD5	0.031	Variable created for Origin County name	Kent	Categorical
CD6	-0.026	Variable created for Origin County name	Kings	Categorical
CD7	0.002	Variable created for Origin County name	Madawaska	Categorical
CD8	0.077	Variable created for Origin County name	Northumberland	Categorical
CD9	0.013	Variable created for Origin County name	Queens	Categorical
CD10	0.080	Variable created for Origin County name	Restigouche	Categorical
CD11	-0.049	Variable created for Origin County name	Saint John	Categorical
CD12	-0.012	Variable created for Origin County name	Sunbury	Categorical
CD13	0.031	Variable created for Origin County name	Victoria	Categorical
CD14	-0.147	Variable created for Origin County name	Westmorland	Categorical
CD15[1]	n/a	Dummy variable for Origin County Name	York	Categorical
HR1	-0.101	Variable created for health_region = 1	Moncton	Categorical
HR2	0.03591	Variable created for health_region = 2	Saint John	Categorical
HR3	-0.015	Variable created for health_region = 3	Fredericton	Categorical
HR4	0.012	Variable created for health_region = 4	Edmundston	Categorical
HR5	0.076	Variable created for health_region = 5	Campbellton	Categorical
HR6	-0.003	Variable created for health_region = 6	Bathurst	Categorical
HR7[1]	n/a	Dummy variable for health_region = 7	Miramichi	Categorical
SERVICE1	0.001	Variable created for SERVICE_TYPE_GROUP_CODE = 1110	SERVICE1 = 1110 = Hospital Inpatient	Categorical
SERVICE2	-0.20342	Variable created for SERVICE_TYPE_GROUP_CODE = 1130	SERVICE2 = 1130 = Office Visit	Categorical
SERVICE3	0.140	Variable created for SERVICE_TYPE_GROUP_CODE = 1300	SERVICE3 = 1300 = Consultation	Categorical
SERVICE4	0.045	Variable created for SERVICE_TYPE_GROUP_CODE = 1500	SERVICE4 = 1500 = Emergency	Categorical
SERVICE5	0.121	Variable created for SERVICE_TYPE_GROUP_CODE = 2200	SERVICE5 = 2200 = Diagnostic/ Therapeutic	Categorical
SERVICE6[1]	n/a	Variable created for all other SERVICE_TYPE_GROUP_CODE	SERVICE6 = ALL Other (under 4%)	Categorical
SERV_CAT1	-0.18571	Variable created for SERVICE_CATEGORY_CODE = 1000	SERV_CAT1 = 1000 = Visit	Categorical
SERV_CAT2	0.140	Variable created for SERVICE_CATEGORY_CODE = 2000	SERV_CAT2 = 2000 = Consultation	Categorical
SERV_CAT3[1]	n/a	Variable created for SERVICE_CATEGORY_CODE = 3000	SERV_CAT3 = 3000 = Procedure	Categorical

[1]*Dummy Variable*

3.5.2 Multicollinearity Between Independent Variables

To avoid the risk of multicollinearity between dependent variables, two common indicators were used. The first being a correlation matrix, where dependent variables considered to be used within the linear regression models were compared together in a correlation matrix. This was to indicate which variables may have a high correlation with each other which may influence the model if used together. The full correlation matrix of the dependent variables was unable to be released from the lab; however, variables that had a high correlation with each other were noted.

Literature suggests that variables of 0.7-0.9 are considered to have a correlation; therefore, the following matrixes, shown in Table 3-8 and Table 3-9, indicate which variables were noted to be highly correlated with each other.

- A correlation of 0.9 or above was indicated as a strong correlation shown in red
- A correlation of 0.7-0.9 was considered a high correlation shown in orange
- A correlation of 0.5-0.6 was considered a lower correlation however still notable and highlighted yellow.
- All other correlations below 0.5 were considered not highly correlated.

Table 3-8 - DAD Data - Correlation of Dependent Variables

	Max_beds	**PHY_SP1**	**ADMIT1**	**MAJOR1**	**CD7**	**CD10**
CityV3	OK	n/a	n/a	n/a	n/a	n/a
City V7	OK	n/a	n/a	n/a	n/a	n/a
MAJOR1	n/a	STRONG	n/a	n/a	n/a	n/a
INST_TYPE1	n/a	OK	HIGH	OK	n/a	n/a
INST_TYPE2	n/a	n/a	HIGH	n/a	n/a	n/a
HR4	n/a	n/a	n/a	n/a	HIGH	n/a
HR5	n/a	n/a	n/a	n/a	n/a	STRONG

Table 3-9 - Physician Billing Data - Correlation of Dependent Variables

	Max_beds	PHY_SP1	MAJOR1	MAJOR2	CD4	CD6	CD14
SERIVCE2	OK	n/a	OK	OK	n/a	n/a	n/a
SERVICE5	n/a	n/a	OK	OK	n/a	n/a	n/a
MAJOR1	n/a	STRONG	n/a	OK	n/a	n/a	n/a
MAJOR2	n/a	OK	OK	n/a	n/a	n/a	n/a
SERV_CAT1	n/a	n/a	n/a	OK	n/a	n/a	n/a
PHYS_SP1	n/a	n/a	n/a	OK	n/a	n/a	n/a
HR6	n/a	n/a	n/a	n/a	STRONG	n/a	n/a
HR2	n/a	n/a	n/a	n/a	n/a	OK	n/a
HR1	n/a	n/a	n/a	n/a	n/a	n/a	STRONG

Categories that indicated a high or strong correlation between variables were not used together within the regression model and an “all or nothing” approach was used. As an example, a variable that showed high correlation with each other in both datasets was the variable PHY_TYPE1 which indicated if the health trip was general practice or specialty practice, and the major variables MAJOR 1, MAJOR 2, MAJOR 3 which further broke the health trips into general practice, medical specialty and surgical specialty. PHY_TYPE1 and MAJOR 1 presented a very high correlation with each other since both were able to indicate if general practice was the health trip type. Due to the high correlation with each other, these variables were not used together within the regression models. When using the PHY_TYPE1 variable, all MAJOR variables were removed as well as when MAJOR variables were used in the model, PHY_TYPE1 was removed.

In addition, because the correlation matrix does not always present concerns for multicollinearity, the Variance Inflation Factor (VIF) was monitored throughout iterations of forwards regression to determine if multicollinearity was present. The higher the VIF

value, the higher indication that multicollinearity may be present. There is not a clear threshold for the VIF value; however, a common "rule of thumb" suggests that a VIF value of 10 or greater indicates severe multicollinearity and a VIF between 5-10 suggests that there is a possible concern for mild multicollinearity. Typical thresholds for determining multicollinearity are a VIF of 4 or 5, with some sources suggesting 2.5 being a threshold of multicollinearity (O'Brien, 2007; Chang & Mastrangelo, 2011). O'Brien suggests that although a VIF is a good indication of multicollinearity, they need to be interpreted by using the judgment of other factors before removing variables from the model. Other factors include consideration of sample size, t-values, and/ or confidence intervals (2007). The VIF was monitored when completing the regression steps for both datasets and a VIF of 4 was used as a threshold; however, ideally, a VIF of 2.5 was achieved. When variables were greater than a VIF of 2.5 consideration of factors was evaluated before accepting or rejecting the variable in the model.

3.5.3 Forward Stepwise Regression

Independent variables were added to the regression analysis one at a time, beginning with the "*major*" variables. These were chosen to be first applied to the regression analysis because "*major 2*" appeared to have the largest correlation with distance traveled as shown in Table 3-6. The "*major*" variables (1-3) are categorical variables representing the three major specialties that physicians are categorized in, *major 1* being general practice, *major 2* being medical specialty and *major 3* being a dummy variable representing surgical specialties.

County variables where trips were originating were then added to the model ranging from *CD1* to *CD14* representing each of the 15 New Brunswick Counties. The 15th county variable, York County was used as a dummy variable.

Variables were continued to be added to the regression models one by one to determine their significance to the model. This included variables such as the maximum number of beds, DA origin income levels, age, gender, and destination city type (*dCITYv3*). The variable *dCITYv3* indicates if a health trip destination is to an urban or rural location with urban locations including the 3 major New Brunswick Cities determined by NB-IRDT, Moncton, Fredericton, and Saint John. To keep the health models consistent with each other, variables were kept similar except for some variables which were only available in one of the two datasets such as admission type and length of stay which were variables included in the DAD data and not the Physician Billing Data. Detailed specialty groups were decided to not be included in the regression model as it was determined to be too in-depth for this application and linear regression may not be the correct way to demonstrate those relationships.

As variable groups were being added to the regression analysis, key indicators to determine the value of the variable in the model were the variable coefficients, the adjusted R^2 values, parameter estimates "PR>[t]" and the "variance inflation" factor (VIF). As variables were added, it was noted if the variable coefficients made sense, if the adjusted R^2 value increased or decreased, and if any changes occurred to the parameter estimates. In particular, the confidence interval used during this regression was 95%; therefore, if the PR>[t] exceeded 0.05 after adding a set of variables, it was noted as insignificant. The VIF was also monitored throughout iterations of adding new variables to the model as this factor

is an indicator of multicollinearity between dependent variables flagging variables that were exceeding a VIF of 2.5 or approaching the threshold of a VIF of 4.

As mentioned previously, some independent variables are similar variable types and are unable to be used in the model together due to a high correlation between the two. An example of this would be the physician specialty variable (PHY_SP1), which also displayed one of the higher correlated variables to distance. This variable can be either GP or SP in the data; however because the physician specialty variable and the *major (1-3)* variables are a similar variable type and highly correlated, they are unable to be used together in the model. To determine which variable should be included in the regression model, the "*major(1-3)*" variables were removed from the model and the physician specialty variable was added.

Other variables were also removed and replaced with their similar variable type such as the destination city variables *dCITYv3* vs. *dCITYv7* to determine which variable fit the model best. The *dCITY3* and *dCITY7* variables are both rural/ urban indicators for the destination location; however, *dCITY3* considers any location within 40km of the 3 major New Brunswick cities to be "urban", making the rest of the destinations considered "rural". The *dCITY7* variable considers any location within a 40km radius of the 7 major cities in the province to be "urban" and the remaining destinations to be considered "rural". These variables are highly correlated therefore were applied to the model separately. Originating health region variables (*H1-7*) and originating county variables (*CD1-15*) are another example where both categories were unable to be analyzed together in the model due to the similar relationship and correlation that was shared. When analyzing one category, the opposing category was removed.

Multiple iterations of regression analysis were run to determine a model with reasonable coefficients, parameter estimates, VIF values, and a higher adjusted R^2 value.

3.5.4 Comparing with Census Data Regression Model

In addition to models developed using health variables, models using variables that may be available from census data were also developed for each data set. This included ages, gender, county of origin, and origin DA income. This allowed the models with a health component to be compared to a model using typically available census information. The same datasets were used to create both models to ensure that the results would reflect the addition of health variables appropriately.

The following variables in Table 3-10 were used to develop census models for each dataset:

Table 3-10 - Census Model Independent Variables

CENSUS MODEL INDEPENDENT VARIABLES			
Variable	**Description**	**Variable Meaning**	**Variable Type**
Age	Age of Individual at time of visit	Numeric Age	Continuous variable
gender2	Regression variable created for gender	Male = 1, Female = 0	Categorical Variable
Q1	Regression variable created for QAIPPE = 1	1 = Low Income	Categorical Variable
Q2	Regression variable created for QAIPPE = 2	2 = Low-Medium Income	Categorical Variable
Q3	Regression variable created for QAIPPE = 3	3 = Medium Income	Categorical Variable
Q4	Regression variable created for QAIPPE = 4	4 = Medium-High Income	Categorical Variable
Q5[1]	Dummy Variablefor QAIPPE = 5	5 = High Income	Categorical Variable
CD1	Regression Variable created for Origin County name	Albert	Categorical Variable
CD2	Regression Variable created for Origin County name	Carleton	Categorical Variable
CD3	Regression Variable created for Origin County name	Charlotte	Categorical Variable
CD4	Regression Variable created for Origin County name	Gloucester	Categorical Variable
CD5	Regression Variable created for Origin County name	Kent	Categorical Variable
CD6	Regression Variable created for Origin County name	Kings	Categorical Variable
CD7	Regression Variable created for Origin County name	Madawaska	Categorical Variable
CD8	Regression Variable created for Origin County name	Northumberland	Categorical Variable
CD9	Regression Variable created for Origin County name	Queens	Categorical Variable
CD10	Regression Variable created for Origin County name	Restigouche	Categorical Variable
CD11	Regression Variable created for Origin County name	Saint John	Categorical Variable
CD12	Regression Variable created for Origin County name	Sunbury	Categorical Variable
CD13	Regression Variable created for Origin County name	Victoria	Categorical Variable
CD14	Regression Variable created for Origin County name	Westmorland	Categorical Variable
CD15[1]	Dummy Variable for Origin County Name	York	Categorical Variable

3.6 Releasing the Data from the Lab

Prior to being released from the lab, random rounding was applied to tables and matrixes that included a cell with less than a value of 5. This meant that if a cell had fewer than 5 trips it was randomly rounded to 5 or 0 to ensure that data was not able to be identified. To complete this, a matrix of randomly generated numbers was created in excel providing a number between 0 and 1.0. The MROUND function in Excel was then applied to the trip matrix with a condition if the random number associated with the trip cell was greater than 0.5 then round up if less than, round down. This produced a trip table with all randomly rounded cells ending in a 0 or 5. This was a requirement to release the data and, in some locations, made a minimal difference in the end trip number; however, it should be noted that in some areas where trips were less frequent, the random rounding increased the number of trips. For example, if county A attracted 1 trip from each of the 15 counties, the total trips attracted to that county should be 15. However, the total trips attracted to that county could be 75 total trips if each cell was rounded up to 5. For the purposes of this project, this circumstance was deemed to be acceptable; however, it was taken into consideration when creating trip distribution models.

A data release form was submitted and reviewed for the release of data and project notes. After a review process, a total of 16 spreadsheets was released from the lab which included regression models, classification table information, O-D matrixes, distance distributions, distance statistics, age statistics, regression models, and correlation matrixes from both the DAD dataset and the Physician Billing datasets.

3.7 Trip Generation: Creating Classification Tables

Trip tables were created in the lab both as a single combination of years and by separate years. These trips tables included total health trips taken by individuals which captured the total number of health trips taken by a single individual for that time period. This was completed for both the DAD dataset and the Physician Billing Datasets. Table 3-11 below shows an example of health trips taken in 2011 from the DAD dataset. These trips include only health trips to a hospital.

Table 3-11 - Total Health Trips for 2011 from the DAD Data

DAD DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	2,850	2,615	2,400	2,080	1,405	945	12,295
M	3,380	2,895	2,570	1,905	1,050	465	12,265
Total	6,230	5,510	4,970	3,985	2,455	1,410	24,560

Population, age groups, and gender data were then collected from Statistics Canada for the census years of 2006 and 2011. This included data for the province of New Brunswick and each of the 15 Counties. These data were put into a similar table as the trip tables, displaying the total population of an age group within a gender category for that census year. An example of the census data is shown below in Table 3-12, displaying New Brunswick's population in 2011.

Table 3-12 - Population of New Brunswick from Census 2011

2011 New Brunswick Population							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	19,695	15,190	12,455	9,795	6,930	4,295	68,360
M	19,415	14,060	10,025	6,530	3,715	1,525	55,270
Total	39,110	29,250	22,480	16,325	10,645	5,820	123,630

Trip rates were then calculated for each age group and gender by dividing the total health trips by the total population as shown below in equation 9, where x is the year the rates are being calculated. This provided a rate presented as a health trip per person for 2011.

$$Rate\ of\ Health\ Trip_x = \frac{Heath\ Trip_x}{Population_x} \quad \textbf{(9)}$$

This process was repeated using the 2006 census data and the corresponding year from the health datasets. Two additional trip rate tables were created to demonstrate trip rates from all of the data years. This was done by creating a table that represented the average population from the 2006 and 2011 census populations. The average population between 2006 and 2011 is shown below in Table 3-13.

Table 3-13 - Average Population Between Census Years 2006-2011

Average New Brunswick Population							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	17,825	14,445	12,088	9,858	6,575	3,965	64,755
M	17,370	12,960	9,395	6,405	3,350	1,395	50,875
Total	35,195	27,405	21,483	16,263	9,925	5,360	115,630

Trip rates were then determined using equation 10 by dividing the total health trips by the average population and number of years of heath trips. For the DAD data (2005-2011), the number of years was considered to be seven, and the number of years for Physician Billing data (2006-2011) was considered to be six. This was created in an attempt to use data from the remaining years between 2006 and 2011 to produce trip rates using all data.

$$Rate\ of\ Heath\ Trip = \frac{Health\ Trip_{Total}/_{year}}{Population_{Average}} \quad \textbf{(10)}$$

The same process was used to determine rates for health trips by county using the population data per county from Statistics Canada. Table 3-14 shows an example of the 2011 population data used to develop rates.

Table 3-14 - Population by County from 2011 Census

2011 Census Populations							
County	**65-69**	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
Albert	1,585	1,095	755	570	405	205	4,615
Carleton	1,235	990	760	610	410	230	4,235
Charlotte	1,430	1,145	855	660	400	260	4,750
Gloucester	4,895	3,505	2,605	1,750	1,150	590	14,495
Kent	1,980	1,490	1,105	700	460	265	6,000
Kings	3,415	2,420	1,660	1,185	720	425	9,825
Madawaska	1,865	1,395	1,085	795	550	195	5,885
Northumberland	2,785	2,265	1,730	1,185	700	410	9,075
Queens	445	335	245	195	145	95	1,460
Restigouche	1,020	810	730	575	360	220	3,715
Saint John	3,650	2,700	2,305	1,930	1,240	780	12,605
Sunbury	915	785	550	365	180	75	2,870
Victoria	1,095	775	680	475	285	130	3,440
Westmorland	6,745	5,065	4,065	3,030	2,105	1,265	22,275
York	4,585	3,425	2,525	1,850	1,255	620	14,260
Total	37,645	28,200	21,655	15,875	10,365	5,765	119,505

It should be noted that the DAD data from 2006 and 2011 are health trips taken in the fiscal years of 2006 and 2011. This means that the census data from 2006 and 2011 are not exactly the same as the DAD data years as census data is not captured in fiscal years. However, for the purposes of this analysis, it was deemed acceptable to develop baseline rates.

3.8 Trip Distribution: Gravity Models

Origin-destination (OD) matrix tables were generated from the DAD and Physician Billing datasets. This was done for each year, as well as for one large OD matrix including all trips. When generating the OD matrix table for the DAD data, the cities of where the health trip took place were used as the destinations therefore, they were manually grouped into the

associated county by adding the trips from the destination cities within one county together. This provided an OD matrix with 15 origin counties and 15 destination counties resulting in a total number of productions and attractions, as well as the actual distribution of health trips between counties. The productions and attractions of the health trips were equal therefore no balancing was required.

3.8.1 Traditional Gravity Model

A 15x15 cost matrix was created based on kilometers (km) traveled between each county (zone). Distances between counties were determined using google maps and rounded to the nearest 5. Internal travel between zones was assumed to range between 30-40 km except for Charlotte county and Westmorland County. Charlotte county assumed an average health trip within Charlotte county to be 60km. This was to account for the three islands in Charlotte county that experience additional travel, Deer Island, Campobello Island, and Grand Manan Island. Westmoreland County assumed an average travel distance of 50km due to its inclusion of outer locations such as Sackville which appear to require travel to Moncton for health trips. It should be noted that for the DAD data OD matrix output, Albert county produced health trips but did not attract any health trips; therefore, a destination column for Albert county was created to complete the trip matrix and populated with zeros.

A deterrence function was then created using the exponential function shown in equation (11). This used an assumed B of 0.1 and the distance of travel cost (c) from the cost matrix created previously. The deterrence function was created to apply a *friction factor* to the gravity model. This function makes a trip with a further distance less attractive than one that requires less travel.

$$F_{ij} = \exp\left(-B * c\right) \quad \textbf{(11)}$$

The traditional form of the gravity model was used to demonstrate trip distribution for health trips using the following equation (12) from NCHRP Report 716.

$$T_i^P = P_i^P * \frac{A_j^P * f(t_{ij}) * K_{ij}}{\sum_{j' \in zones} A_{j'}^P * f(t_{ij'}) * K_{ij'}} \quad \textbf{(12)}$$

Where;

T_i^P	Trips produced in zone *i* and attracted to zone *j*
P_i^P	Production of trip ends for purpose p in zone *i*
A_j^P	Attraction of trip ends for purpose p in zone *j*
$f(t_{ij})$	Friction factor, a function of travel impedance between zone *i* and zone *j*
K_{ij}	K-factor, optional adjustment factor used to account for effects of variables other than travel impedance

The K-factor was assumed to be 1 for the traditional method of trip distribution as it is typically used to represent socioeconomic barriers.

This was applied to each trip cell in the trip matrix and then summed to produce new totals for productions and attractions. An attraction ratio and production ratio were determined by dividing the target by the new total. The attraction ratios were then used as an attraction growth factor to create a set of new attractions by multiplying the target attractions by the attraction ratio. The trips were then recalculated using the new calculated attractions and the process was then repeated but adjusting for productions instead of attractions. This completed one iteration, a total of two iterations were completed.

3.8.2 Health-Based Gravity Model

To create a health base model for trip distribution, a health factor was applied to the traditional method in the same way a K-factor or deterrence function would be applied. Equation (13) below demonstrates how this factor would be applied.

$$T_i^P = P_i^P * \frac{A_j^P * f(t_{ij}) * K_{ij} * HF_{ij}}{\sum_{j' \in zones} A_{j'}^P * f(t_{ij'}) * K_{ij'} * HF_{ij}} \quad \textbf{(13)}$$

Where;

T_i^P	Trips produced in zone *i* and attracted to zone *j*
P_i^P	Production of trip ends for purpose p in zone *i*
A_j^P	Attraction of trip ends for purpose p in zone *j*
$f(t_{ij})$	Friction factor, a function of travel impedance between zone *i* and zone *j*
K_{ij}	K-factor, optional adjustment factor used to account for effects of variables other than travel impedance
HF_{ij}	Health Factor, a function of impedance based on a health factor between zones *i* and zone *j*

3.8.3 Health Factors

After analyzing the regression outputs for both the DAD and Physician Billing models, variables that were considered good candidates for health factors were determined. This was done by using the coefficients from the regression models to determine which variables appeared to have the largest influence on distance traveled for health trips. The variables for consideration were the following:

1. County of Origin
2. Type of Health Trip (general practice vs. specialist)
3. Urban/ Rural for hospital destination

Both regression models from the DAD and Physician Billing data presented relatively high coefficients for rural counties and negative coefficients for the urban counties such as Saint John and Westmorland County. This indicates that typically trips originating from rural

counties have an increase in average distance traveled compared to urban counties that would experience a decrease in average distance traveled for health trips. Rural counties that are within close proximity to an urban county tended to have a lower coefficient, however still positive indicating that they still increase the distance traveled but much less of a distance in comparison to rural counties in further proximity to an urban county.

A rural/urban health factor was created to demonstrate the likelihood of a person traveling from one county to another county based on their origin county being rural or urban and the destination county being rural or urban. An exception was made for Sunbury County as it was deemed an urban county for trips originating from York County. This is because it is home to the Oromocto Hospital which is in close proximity to the Dr. Everett Chalmers Hospital in York County and typically receives health trips from York County. For all other counties, Sunbury was considered a rural destination. To categorize the counties into rural/urban indicators for health travel, the following situations outlined in Table 3-15 and Table 3-16 were considered:

Table 3-15 - Intrazonal Categories for Health Trips

Interzonal Trips		
Origin	Destination	Description
Urban	Urban	Urban-to-urban internal trips were considered the most likely trip and were given a value of 0. This is because the urban counties are expected to have the most health care resources for both general and specialty practices.
Rural	Rural	Rural-to-rural internal trips were considered likely however less likely than an urban-to-urban internal trip due to fewer recourses and specialist physicians in rural hospitals.

Table 3-16 - Zone-to-Zone Categories for Health Trips

Zone-to-Zone		
Origin	Destination	Description
Urban	Urban	An originating urban trip to another urban trip was considered significantly more likely than an urban to rural health trip as most rural hospitals do not have the equivalent resources as an urban hospital.
Urban	Rural	Urban-to-rural trips were considered the least likely trip since it would be unlikely that a resource in a rural county would not be available in an urban county.
Rural	Rural	Rural-to-rural trips were considered a moderate likelihood for health trips. It was assumed that rural counties would have similar resources and if rural trips were leaving a county it would be to seek out a higher level of care in an urban County.
Rural	Urban	Rural-to-urban trips were considered to be a likely trip as rural counties tended to travel further distances on average. Rural-to-urban trips were highly likely for specialty health trips as rural counties appeared to travel further distances on average for specialty trips.
A destination with no Hospital or small Health Centre with no beds		Counties without a regional hospital or only a small health care center to provide basic physician services were considered highly unlikely to attract health trips to those counties.

A 15x15 cost matrix was created for the health factor by applying estimations of cost to an origin-destination pair. Each color in Figure 3-19 represents one of the categories that were created to represent a health factor from the data.

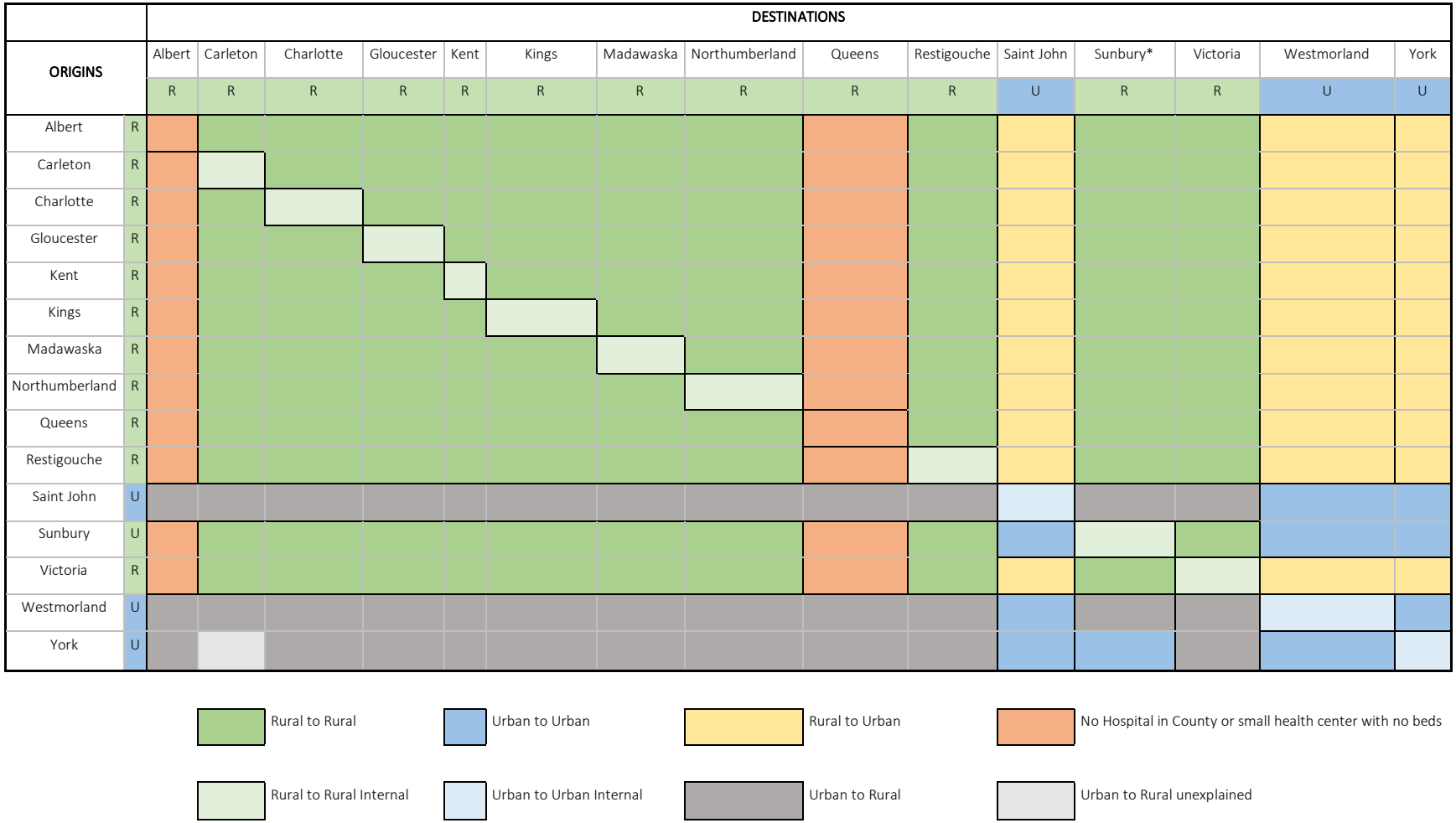

Figure 3-19 - Health Factors Cost Matrix Categories

The health factors were then created in a matrix using the exponential function, similar to the deterrence function, and used b = 0.1. Instead of using distance (km) between counties, a value was given to apply how attractive or unattractive a health trip would be based on where the person was originating from. This was used to provide the model with a better representation of how the health trips are actually distributed throughout the province. This factor made a health trip to one of the 3 urban counties with major hospitals more attractive when they were originally less attractive based solely on distances.

This health factor was applied with the deterrence function in the traditional gravity model formula to ensure that counties, that were traveling to major cities for health care, were still travelling to the closest urban county to a major hospital. For example, health trips from Charlotte County should be more attracted to Saint John County before travelling to York County or Westmorland County.

Once the health factor was calculated it was applied to the traditional gravity model formula the way a K-Factor would normally be applied. The model was then continued through the gravity model steps for two iterations. This was completed for both the DAD dataset and the Physician Billing dataset using data from 2011. Each dataset has a traditional gravity model and a health-based gravity model with two iterations. The results of the two trip distribution models were then able to be compared against the true trip distributions for health trips. The result from these models will be discussed in the following chapter.

CHAPTER 4 : ANALYSIS & RESULTS

This chapter summarizes, discusses, and compares results from the analysis completed using the administrative health data from both the DAD and the Physician Billing datasets. The analysis will include descriptive statistics and linear regression analysis, as well as base applications of trip generation and trip distribution models, using classification tables and the traditional form of the gravity model with an applied health factor.

4.1 Descriptive Statistics

Descriptive statistics were determined using the total distance traveled (km), trip type, county of origin, age, and gender for health trips made by older adults aged 65+ within the DAD and Physician Billing datasets. It should be noted that the DAD data captures health trips to a physician at a hospital location, while the Physician Billing dataset includes health trips to where a physician is located and providing the health service. This would include mostly physician offices and clinic locations; however, a physician billing trip could sometimes include a trip to a hospital location. This was assumed because some physicians have clinics or offices within a hospital.

4.1.1 Age and Population Statistics

Populations for New Brunswick, including older adults ages 65+, were obtained from Statistics Canada from the years of 2006 and 2011 to represent the years of data analyzed in the lab. As shown in Table 4-1, the older adult population as a percentage of the overall New Brunswick population for the year 2011 was 16%, but ranged from 11% to 24% depending on the county as shown in Figure 4-1. The counties with the 3 major cities, Fredericton, Moncton, and Saint John, located in York County, Westmorland County, and

Saint John County, all had between 15%-16% of their county population aged 65+. The three major counties considered as "*urban*" make up a total of 40% of the older adult population while the remaining 60% of older adults were in the remaining counties.

Table 4-1 - Census Data for 2011

2011 CENSUS DATA			
County Name	**Total Population**	**Population 65+**	**% Senior Population**
Albert	28,845	4,615	16%
Carleton	27,020	4,235	16%
Charlotte	26,550	4,750	18%
Gloucester	77,795	14,495	19%
Kent	30,835	6,000	19%
Kings	69,665	9,825	14%
Madawaska	33,425	5,885	18%
Northumberland	48,355	9,075	19%
Queens	11,085	2,705	24%
Restigouche	32,595	6,655	20%
Saint John	76,550	12,605	16%
Sunbury	27,145	2,870	11%
Victoria	19,920	3,440	17%
Westmorland	144,160	22,275	15%
York	97,240	14,260	15%
New Brunswick	751,185	123,690	16%

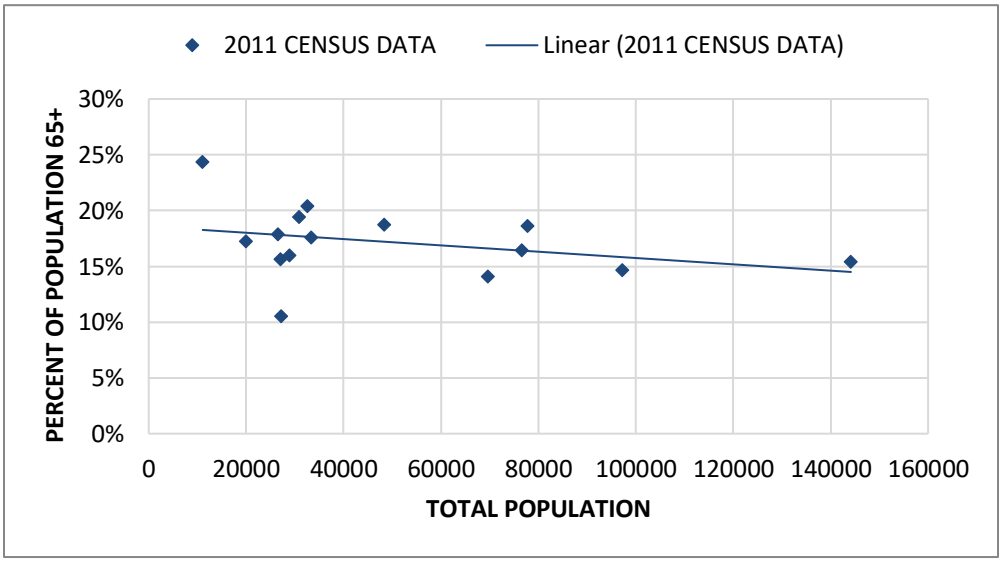

Figure 4-1 - Comparison of Population (65+) to 2011 Province Population

The final health dataset created using the DAD data included a total of 123,490 individuals between the fiscal years of 2005-2011 totaling in 171,150 health trips to a hospital facility.

The final physician billing dataset represents a total 103,670 individuals between the years of 2006-2011 totaling in 2,618,810 health trips to a physician's office, clinic, etc. It should be noted that not every person 65 years and older will take a health trip within the study time period. As well, due to the nature of using raw, administrative data, not all health trips were captured in the data sets or had relevant information for this study therefore were removed. In addition, for instances where multiple services were recorded for one health visit, duplicated services were removed during data processing and one health trip was counted. Figure 4-2, demonstrates the percentage of total health trips produced from each county in comparison to the percent of the New Brunswick population that county represents. As shown below, most health trips are proportional to the percent of the population for that county.

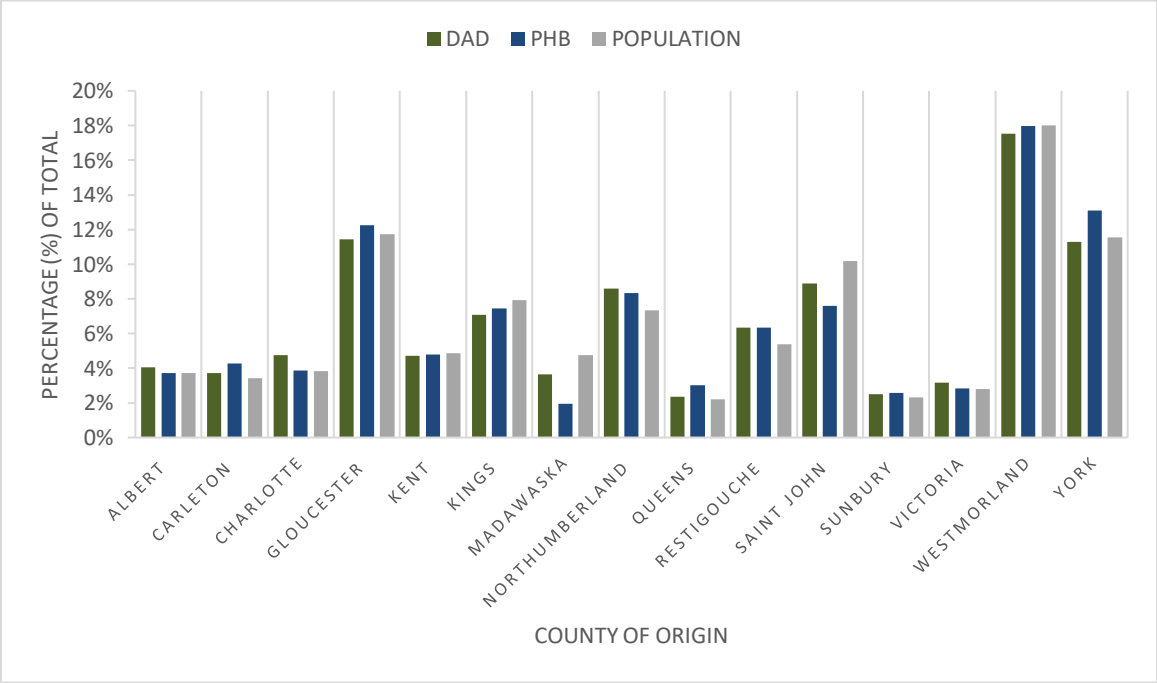

Figure 4-2 - Percentage of Health Tips Compared to Population by County

4.1.2 Distance Statistics for Health Trips

Figure 4-3 illustrates the percentage of health trips that were distributed based on the three major types of health trips for the province of New Brunswick. Generally, 39.5% of health trips from the DAD dataset were designated to general practice, while the remaining 60.5% were specialist health trips, specifically, with 11.5% in a medical specialty and 49.0% to a surgical specialty. Physician Billing demonstrated an opposite trend with 65.3% of trips assigned to general practice, and 34.7% of trips designated to specialty practice with 19.2% medical specialty and 15.5% surgical specialty.

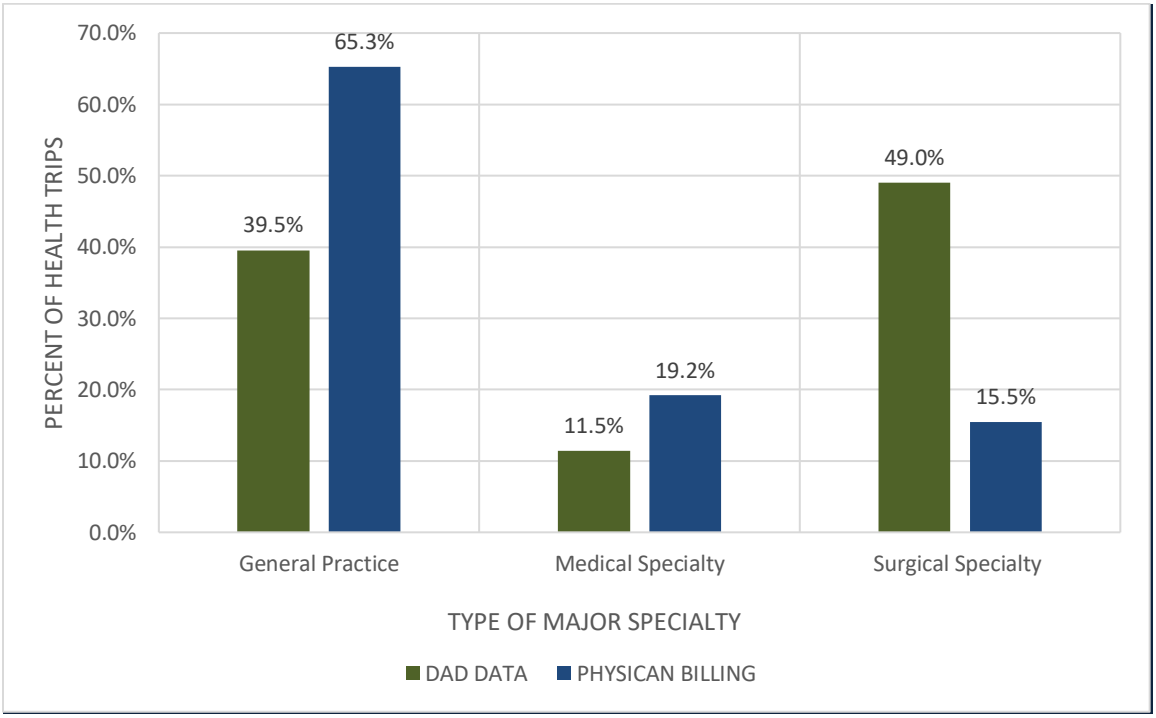

Figure 4-3 – Percent of Health Trips per Specialty Type

As shown in Table 4-2, older adults seeking specialty practices (medical specialty and surgical specialty) in New Brunswick, on average, will travel greater than 50km. The DAD data shows that those seeking a medical specialty were travelling an average of 97.2km, surgical specialties an average of 50.6km and an average of 26.9km for general practice.

Form the Physician Billing data, on average, those seeking medical specialty are travelling 65.3km, surgical specialty an average of 57.3 km and general practice are travelling an average of 27.2km.

Table 4-2 - Comparing Distance Travelled and Frequency for Type of Health Trip

	DAD DATA			PHYSICAN BILLING		
Description of Health Trip	**Number of Observations**	**Percent (%) Total Trips**	**Mean (km)**	**Number of Observations**	**Percent (%) Total Trips**	**Mean (km)**
General Practice	67,800	39.5%	26.9	1,882,430	65.3%	27.2
Medical Specialty	19,640	11.5%	97.2	553,345	19.2%	65.3
Surgical Specialty	84,030	49.0%	50.6	446,620	15.5%	57.3

Trips from Table 4-2 were broken up into age groups as shown in Table 4-3 and Table 4-4 to demonstrate how far different age groups were travelling for type of trip.

Table 4-3 - General Practice Health Trips by Age Group

11000 - GENERAL PRACTICE (GP)						
	DAD DATA			PHYSICAN BILLING		
Age Group	**Number of Observations**	**Percent (%) Trips**	**Mean (km)**	**Number of Observations**	**Percent (%) Trips**	**Mean (km)**
65-69	9,940	14.7%	30.75	453,265	26.4%	27.39
70-74	11,370	16.8%	29.46	389,045	22.7%	27.30
75-79	13,125	19.4%	27.49	334,630	19.5%	26.36
80-84	14,015	20.7%	25.51	264,035	15.4%	25.96
85-89	11,520	17.0%	24.09	176,295	10.3%	25.93
90+	7,785	11.5%	23.68	97,220	5.7%	28.92

As shown in Table 4-3, the provincial average distance travelled for general practice (GP) health trips does not exceed an average of 30km. The DAD data show that between the fiscal years of 2005 and 2011, adults 80-84 years of age were producing the most GP health trips to a hospital, followed by adults 75-79 and 85-84 years of age respectively. The Physician Billing data show between the calendar years of 2006 and 2011, adults 65-69 produced the most GP health trips to a physician's location with 26.4% of trips, followed

by adults 70-74 who produced 22.7% of trips. These two age groups generated almost half of the GP health trips to a physician's location.

Table 4-4 breaks down the specialty trips into medical specialty and surgical specialty trips for each age group within the two datasets. The age groups generating the most health trips for both the medical specialties and the surgical specialties are adults aged 65-69 and 70-74. These age groups also have the greatest travel distance on average for these specialty groups from both the DAD and Physician Billing datasets.

Table 4-4 - Specialty Practice Health Trips by Age Group

	12000 - MEDICAL SPECIALTY (MSP)						13000 - SURGICAL SPECIALTY (SSP)					
	DAD DATA			PHYSICAN BILLING			DAD DATA			PHYSICAN BILLING		
Age Group	**No. of Obs.**	**Percent (%) Trips**	**Mean (km)**	**No. of Obs.**	**Percent (%) Trips**	**Mean (km)**	**No. of Obs.**	**Percent (%) Trips**	**Mean (km)**	**No. of Obs.**	**Percent (%) Trips**	**Mean (km)**
65-69	5,760	29.3%	105.11	138,235	27.6%	69.16	23,515	28.1%	52.60	113,145	28.0%	58.56
70-74	4,885	24.9%	103.84	117,645	23.5%	69.67	21,355	25.5%	51.79	99,415	24.6%	59.02
75-79	4,110	20.9%	97.91	101,120	20.2%	67.20	18,235	21.8%	49.76	84,045	20.8%	57.80
80-84	3,080	15.7%	88.02	80,290	16.0%	58.36	12,600	15.0%	47.11	60,280	14.9%	55.63
85-89	1,420	7.2%	71.62	46,045	9.2%	53.11	6,010	7.2%	45.98	33,415	8.3%	51.26
90+	375	1.9%	54.46	17,285	3.5%	54.78	2,065	2.5%	46.31	13,385	3.3%	50.86

The DAD data shows that on average, people between 65 and 89 years are travelling 70km or more for a medical specialty health trip to a hospital. Physician Billing data shows that age groups 65-69, 70-74 and 75-79 make up approximately 71% of the medical specialty trips and travel over 65km to make a health trip to a medical specialty physician. Previously, Table 4-2 stated that 11.5% of health trips from the DAD data set were designated to medical specialties while, 49% were assigned to surgical specialties. This means that although older adults in the province are traveling greater distance for a Medical specialty health trip, they do not make up the majority of health trips. However, surgical specialties make up almost half of the DAD data health trips and the average distance travelled throughout the province ranges between 45-53km for a health trip to the hospital

for a surgical specialty. The Physician Billing data shows that on average older adults in the province are required to travel 50-60km to seek a health care physician for a surgical specialty. Although this is a high average travel distance for all age groups, surgical specialties make up approximately 16% of physician billing data as shown in Table 4-2.

It is important to analyze these data from a county perspective in addition to provincial perspective to identify areas within the province that are experiencing high travel distances for healthcare needs. The mean travel distances were analyzed by county to identify which counties are required to travel the furthest distances on average and which counties produce the most health trips. When comparing the mean distance traveled in kilometers (km) for a health trip by county, DAD and Physician Billing datasets were compared together.

Table 4-5 shows the total number of observations, the percentage (%) of total trips and the mean km traveled for a health trips broken down by each county.

Table 4-5 - Comparison Table: Mean Distance Traveled by Origin County

	DAD DATA			PHYSICAN BILLING		
Origin County	**Number of Observations**	**Percent (%) Total Trips**	**Mean (km)**	**Number of Observations**	**Percent (%) Total Trips**	**Mean (km)**
Albert	6,905	4%	53.09	107,230	4%	46.97
Carleton	6,370	4%	57.02	122,960	4%	46.19
Charlotte	8,130	5%	91.78	111,275	4%	85.21
Gloucester	19,555	11%	66.46	353,075	12%	38.70
Kent	8,100	5%	56.04	138,255	5%	46.78
Kings	12,090	7%	40.54	214,815	7%	34.09
Madawaska	6,205	4%	25.61	56,085	2%	39.82
Northumberland	14,720	9%	57.73	240,365	8%	53.07
Queens	4,025	2%	64.03	87,150	3%	43.14
Restigouche	10,855	6%	62.44	182,275	6%	55.91
Saint John	15,190	9%	29.45	218,850	8%	29.86
Sunbury	4,280	3%	29.34	73,825	3%	35.01
Victoria	5,405	3%	69.54	81,380	3%	49.17
Westmorland	30,020	18%	21.20	517,575	18%	22.05
York	19,305	11%	36.05	377,260	13%	31.05

As shown in Table 4-5, the three counties that travel the furthest on average for the DAD dataset are Charlotte County (91.78km), Victoria County (69.54km) Gloucester County (66.46km). Six of 15 counties have average one-way travel distances less than 50km to make health trips.

The three counties which travel the greatest average distance on for the Physician Billing data are Charlotte County (85.21km), Restigouche County (55.91km) and Northumberland County (53.07km). These three counties represent 18% of the health trips taken while the remaining counties are all travelling on average less then 50km.

When comparing these mean distances between the two datasets it is evident that older adults are having to travel farther for hospital trips than for trips to a physician's office. However, older adults in some counties are travelling greater than 50km on average to receive care at a physician's office/clinic. Charlotte County is an outlier in the Physician Billing data, as it demonstrates that there is significantly higher travel on average related to a general office visit. Due to the geographic make-up of Charlotte County, it is expected that the three islands could be influencing the travel distances.

Table 4-5 also shows that from the DAD dataset, the three urban counties represent a total of 38% of the health trips, which are on average less than 30km. This means that 62% of health trips represented in the DAD data are originating from the more rural counties. The rural counties have smaller populations; however, some counties such as Gloucester County are generating more health trips (11%) than the urban counties. Similarly, in the Physician Billing data, the three urban counties make up 39% of the total health trips taken. Gloucester County also generated the third highest number of health trips (12%) provincially.

Figure 4-4 shows a linear relationship in the health data where the travel distance decreases as the percent of health trips from an origin increases. The DAD health trips are shown in green while the Physician Billing health trips are shown in blue. Although there are some notable outliers, there is a general linear relationship shown.

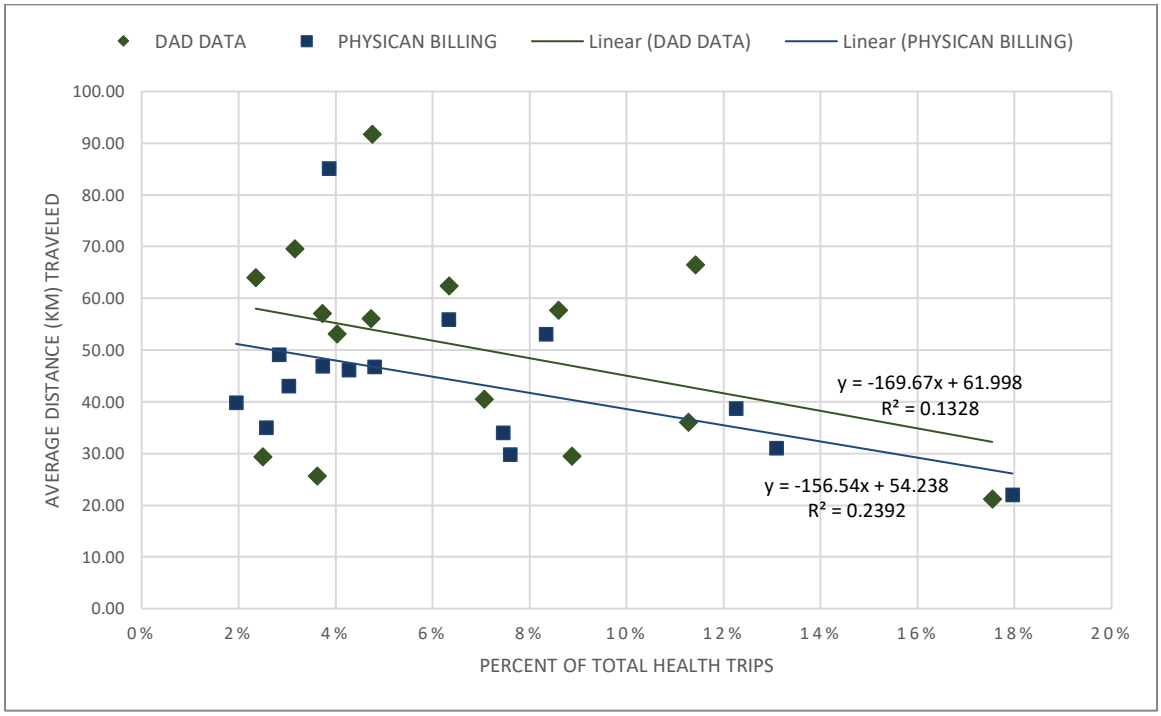

Figure 4-4 – Average Distance Traveled (km) VS Total Percent of Health Trips per County

The most notable outliers within Figure 4-4 are for Charlotte County with an average travel distance of 91.78km (DAD) and 85.21km (Physician) and representing a total of 5% and 4% of all trips. As well as the green outlier in the middle is Gloucester County generating 11% of trips however travelling an average of 66.46km for a DAD health trip.

Health trips taken by county were then broken down into two tables by type of health trip including General Practice health trips (GP) and Specialty Practice health trips (SP). Table 4-6 and Table 4-7 show the number of observations of trips generated from each county, the percentage of the total and the mean distance traveled for GP and SP trips respectively.

Table 4-6 - Comparing Average Distance Travelled for General Practice Health Trips

	DAD DATA			PHYSICAN BILLING		
	GENERAL PRACTICE HEALTH TRIPS (GP)					
County Name	**Number of Observations**	**Percent of GP Trips**	**Mean (km)**	**Number of Observations**	**Percent of GP Trips**	**Mean (km)**
Albert	2,440	3.6%	47.59	62,530	3.3%	40.77
Carleton	3,020	4.5%	21.74	79,290	4.2%	26.12
Charlotte	3,705	5.5%	40.77	85,385	4.5%	68.59
Gloucester	8,315	12.3%	31.04	246,625	13.1%	21.69
Kent	2,065	3.0%	35.75	78,740	4.2%	34.56
Kings	3,145	4.6%	31.21	158,170	8.4%	30.08
Madawaska	2,680	4.0%	19.24	41,955	2.2%	17.41
Northumberland	6,775	10.0%	32.05	159,710	8.5%	27.80
Queens	1,675	2.5%	55.23	62,085	3.3%	28.23
Restigouche	5,655	8.3%	25.48	98,445	5.2%	24.33
Saint John	5,345	7.9%	27.80	143,680	7.6%	29.33
Sunbury	1,705	2.5%	21.80	45,660	2.4%	24.31
Victoria	2,485	3.7%	21.46	60,120	3.2%	22.14
Westmorland	10,575	15.6%	13.24	316,315	16.8%	20.08
York	8,165	12.1%	21.23	243,735	12.9%	22.51
Total	67,750		29.71	1,882,445		29.20

Table 4-6 above, demonstrates that within the DAD data users are typically not required to travel more than 50km on average for a GP health trip except for Queens County. Queens county is travelling on average 55.23km for a GP health trip to a hospital. The second highest travel distance is Albert County, travelling an average of 47.59km for the GP health trip at a hospital. This is most likely because Albert County and Queens County do not have hospital facilities therefore people travel out of county to make a hospital trip. The third highest travel distance is Charlotte County. This is likely due to the location of the Charlotte County Hospital at the western extent of the county rather than at centroid of the population.

Within the Physician Billing data, the average distance traveled is typically 30 km or less on average for a GP health trip except for Albert County, Charlotte County and Kent County. On average older adults are travelling 40.77km one-way in Albert County, travelling 68.59km in Charlotte County and travelling 34.56km in Kent for a GP trip to a physician's office or clinic.

Table 4-7, below presents health trips for Specialty Practice (SP) which is a combination of medical and surgical specialties.

Table 4-7 - Comparing Average Distance Travelled for Specialty Practice Health Trips

	DAD DATA			PHYSICAN BILLING		
	SPECIALTY PRACTICE HEALTH TRIPS (SP)					
County Name	**Number of Observations**	**Percent of SP Trips**	**Mean (km)**	**Number of Observations**	**Percent of SP Trips**	**Mean (km)**
Albert	4,465	4.3%	56.10	44,705	4.5%	55.65
Carleton	3,345	3.2%	88.94	43,675	4.4%	82.63
Charlotte	4,425	4.3%	134.40	25,885	2.6%	140.03
Gloucester	11,240	10.9%	92.64	106,450	10.6%	78.11
Kent	6,035	5.8%	62.97	59,520	6.0%	62.95
Kings	8,940	8.6%	43.82	56,650	5.7%	45.30
Madawaska	3,525	3.4%	30.46	14,130	1.4%	106.35
Northumberland	7,945	7.7%	79.60	80,650	8.1%	103.10
Queens	2,345	2.3%	70.30	25,065	2.5%	80.07
Restigouche	5,200	5.0%	102.61	83,835	8.4%	93.00
Saint John	9,845	9.5%	30.34	75,170	7.5%	30.89
Sunbury	2,570	2.5%	34.35	28,165	2.8%	52.37
Victoria	2,925	2.8%	110.27	21,260	2.1%	125.59
Westmorland	19,445	18.8%	25.53	201,265	20.1%	25.13
York	11,150	10.8%	46.90	133,535	13.4%	46.65
Total	103,400		67.28	999,960		75.19

The DAD data and Physician Billing data indicate that older adults in three counties are travelling greater than an average of 100km for SP trips to hospitals, while older adults in four counties travel greater than 100km to SP trips to physicians. Older adults in the three urban counties of Saint John, Westmorland and York County average 30-45km when seeking a SP at a hospital and 30-47km for a SP to a physician's location.

Older adults from Sunbury County appear to follow a similar average travel distance of those from York County for both GP and SP health trips. This can also be shown in the previous Figure 4-4, which shows that Sunbury County trips represent a smaller proportion of provincial health trips at 3% with average travel distance ranging between 30-35km on from both datasets. This is expected as the Oromocto Public Hospital in Sunbury County is approximately 20km from Fredericton's Dr. Everett Chalmers Hospital in York County making them in close proximity to each other.

4.1.3 Distance Traveled for Specialty Health Trips

Health trips from both datasets were broken into physician specialty groups which provided a breakdown of heath trips by county into 43 different physician specialties. Out of the 43 specialty groups, a total of 7 common specialty groups had a significant percent of health trips being generated from older adults for each county. The following Table 4-8, includes the lists of specialty groups that had common travel associated with them from older adults.

Table 4-8 - Specialty Group Services Used by Older Adults

SPECIALTY GROUP	DESCRIPTION
Cardiology	Cardiovascular system – Heart, blood, etc.
Diagnostic Radiology	Diagnostic imaging - MRI, x-ray etc.
General Practice	General/ primary care
General Surgery	Broad spectrum of diseases - focus on abdominal contents
Ophthalmology	Eye disorders and related surgeries
Orthopedic Surgery	Musculoskeletal system - bones, joints, ligaments, etc.
Urology	Urinary-tract system and reproductive organs

The remaining physician specialties did have trips present; however, minimal trips were associated with the other specialty groups. A total list of these specialty groups by county can be found within Appendix C.

For each specialty group shown in Table 4-8, trip percentages per county were created and the associated average travel distances were included. Trips originating from New Brunswick counties were converted into a trip percentage which was used to bring values to a common scale.

Each county is represented by a weighted bubble (circle) which represents the total number of health trips produced by that county. The size of this circle on the graph illustrates a weighted representation of which counties are producing the most health trips.

The following Figure 4-5 to Figure 4-18 include the 7 specialty groups and the travel behaviours for each count from the DAD and Physician Billing datasets (PHB). Using the following figures, it can be better understood which counties are travelling the furthest for the 7 specialty groups compared to others, what percentage of total health trips that county produces compared to others and what percentage of trips from a county is designated towards a health trip for a health specialty groups in comparison to other counties.

4.1.3.1 Cardiology

Specialty trips for cardiology are shown in Figure 4-5 and Figure 4-6. These figures show the percent of cardiology trips were made per county on the x-axis and the distance traveled for cardiology per county on the y-axis.

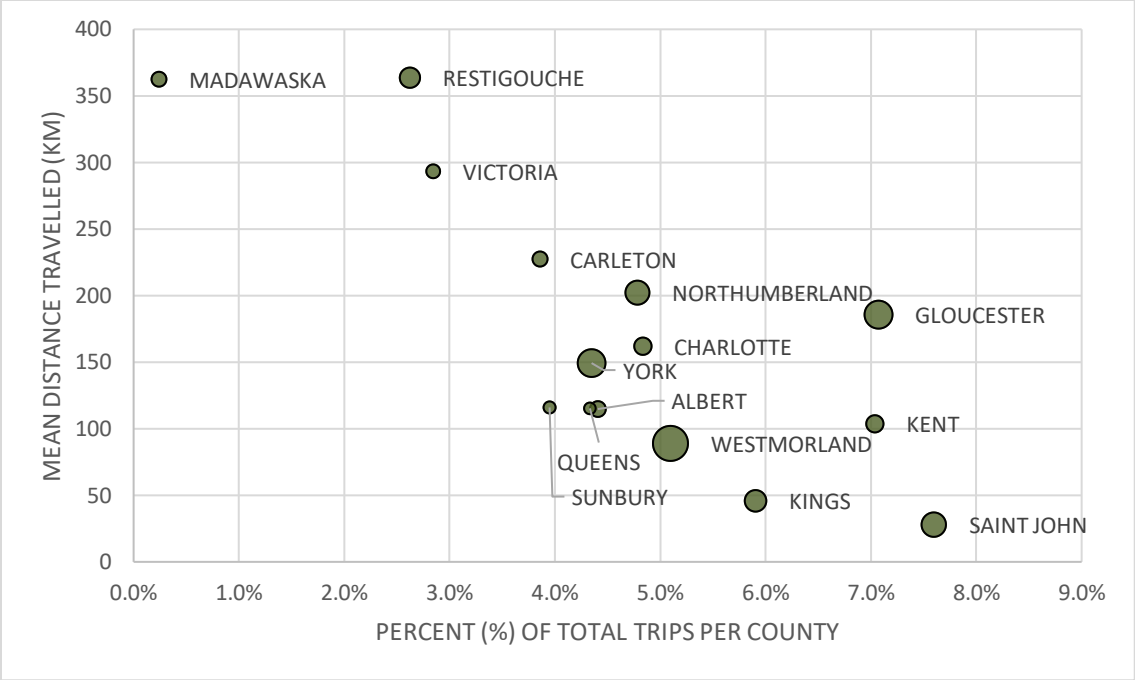

Figure 4-5 - DAD - Travel Distance for Percent of Health Trips Specialized in Cardiology

When analyzing health trips from the DAD data in Figure 4-5, it shows a cluster of counties where 4%-8% of the county's health trips are for the cardiology specialty. Counties such as Gloucester and Northumberland have between 4%-7% of trips in cardiology, but higher average travel distances (150km-200km). Westmorland produces the largest number of health trips and had 5% of health trips for cardiology, while travelling on average just under 100km to a hospital for a cardiology specialist.

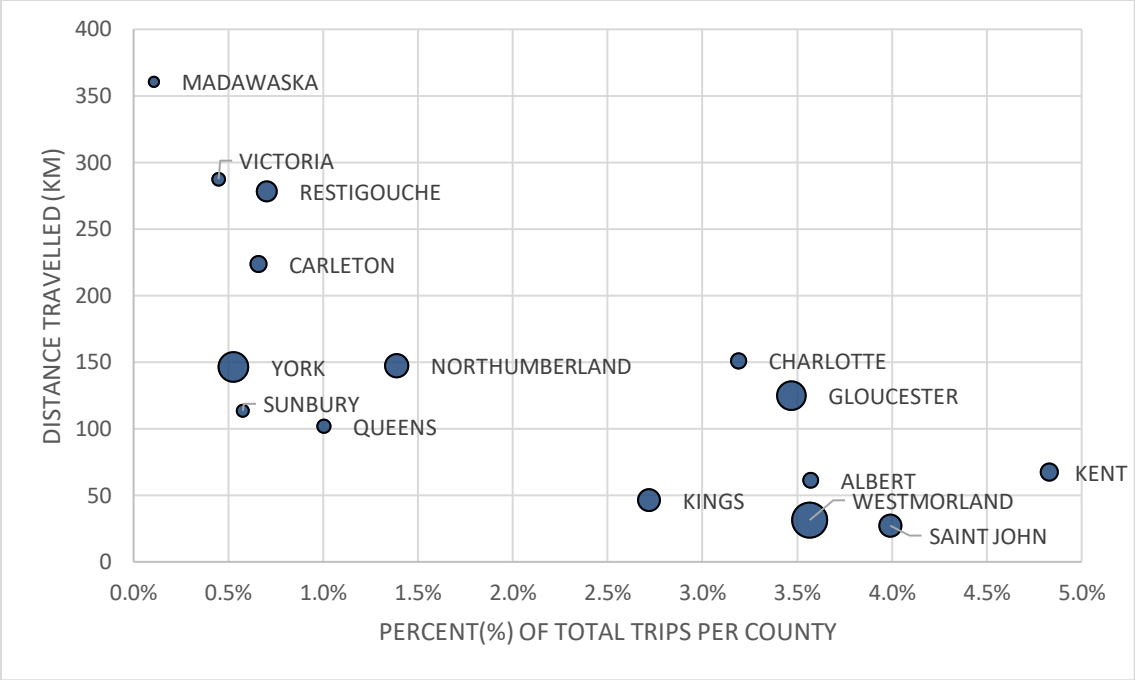

Figure 4-6 - PHB - Travel Distance for Percent of Health Trips Specialized in Cardiology

Fewer older adults travelled to a physician's office/location than a hospital for cardiology; however, Figure 4-6 shows the counties with the highest proportion of cardiology trips appear to have the shortest travel distances. This is expected as the majority of health trips for the Physician Billing dataset are general practice trips. cariology health trips make up 3%-3.5% of Gloucester and Charlotte County total health trips and they are required to travel between 100km-150km to meet this need. It should be noted that Madawaska, Victoria, Restigouche and Carleton County all have an average travel distance greater than 200km; however, cardiology trips for these four counties make up less than 1% of their total health trips per county.

4.1.3.2 Diagnostic Radiology

Specialty trips for diagnostic radiology are shown in Figure 4-7 and Figure 4-8. These figures show the percent of diagnostic radiology trips were made per county on the x-axis and the distance traveled for diagnostic radiology per county on the y-axis.

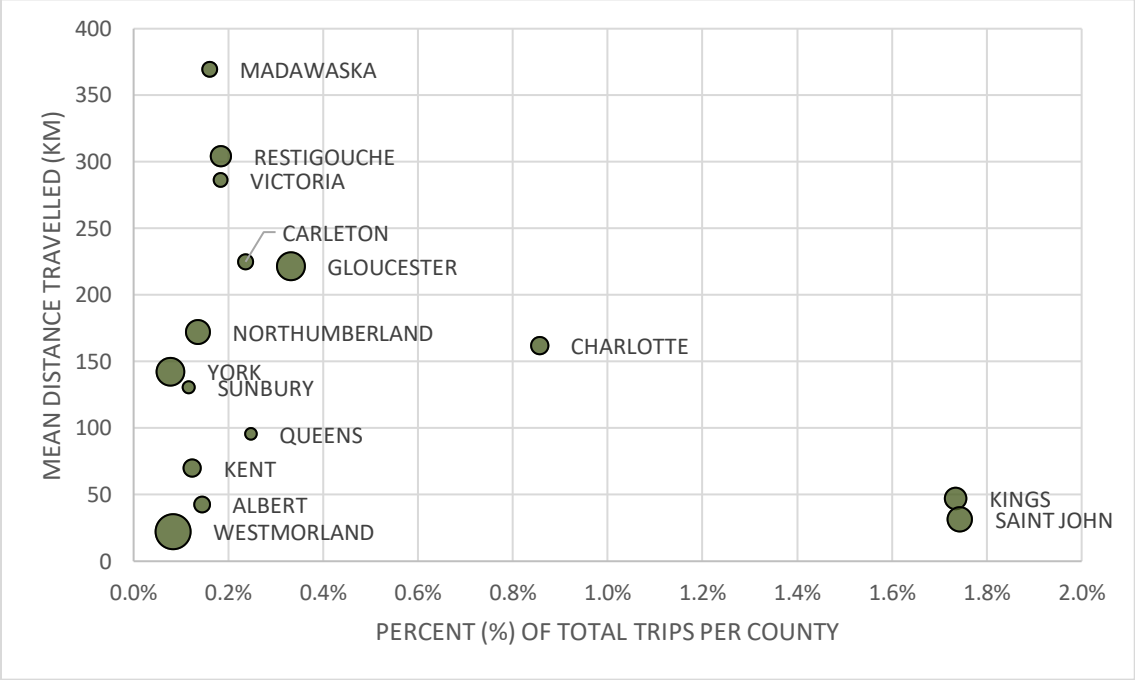

Figure 4-7 - DAD - Travel Distance for Percent of Health Trips Specialized in Diagnostic Radiology

The DAD data in Figure 4-7, shows that diagnostic radiology trips to the hospital are generally quite low, fewer than 1% of trips for each county, except for Kings County and Saint John County who fall within 1.6%-1.8% of trips. It can also be concluded from this figure that although at a significantly lower percentage of trips, older adults in Madawaska, Restigouche, Victoria, Carleton and Gloucester County are all travelling over 200km for diagnostic radiology services at a hospital.

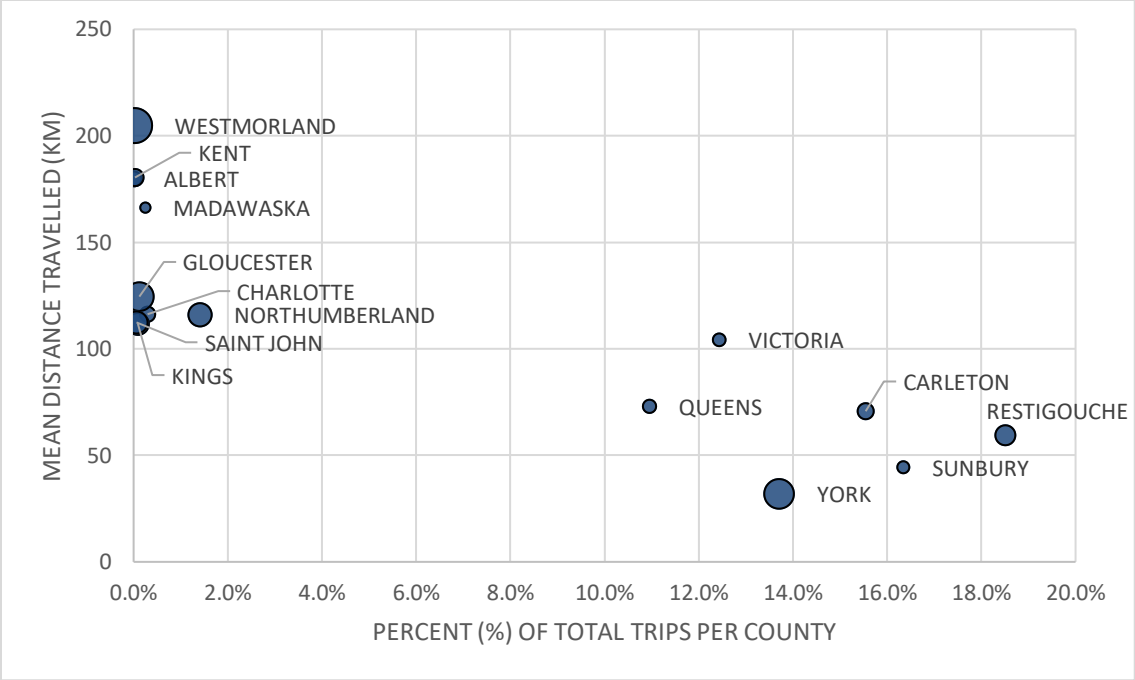

Figure 4-8 - PHB - Travel Distance for Percent of Health Trips Specialized in Diagnostic Radiology

Figure 4-8 shows a significant split in the amount of diagnostic radiology health trips being generated in counties. Approximately 10%-20% of health trips from Queens, Victoria, York, Carleton, Sunbury and Restigouche County are to a physician's office/location for diagnostic radiology and are generally travel less then 100km on average for a trip. The remaining nine counties all have average travel distances greater than 100km to a physician's office/ location for diagnostic radiology; however, overall percentage of health trips for this purpose is close to zero.

4.1.3.3 General Practice

Specialty trips for general practice are shown in Figure 4-9 and Figure 4-10. These figures show the percent of general practice trips were made per county on the x-axis and the distance traveled for general practice per county on the y-axis.

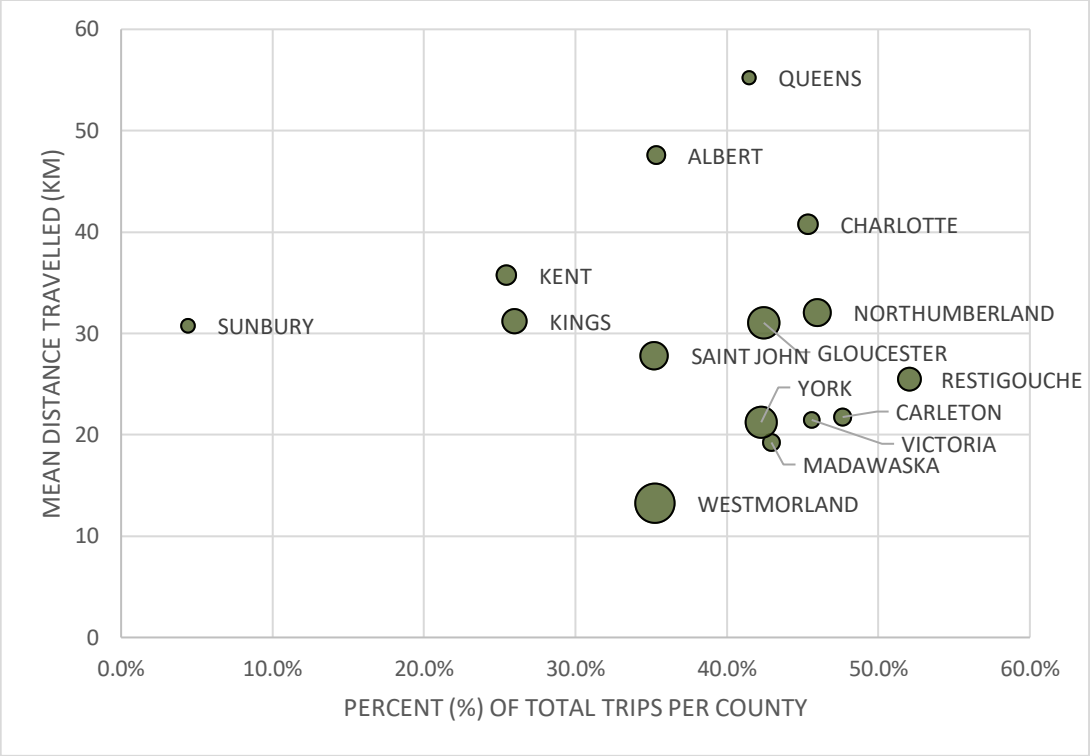

Figure 4-9 – DAD – Travel Distance for Percent of Health Trips Specialized in General Practice

Figure 4-9 shows the general practice specialty trips from the DAD dataset. This figure shows that most of these health trips are less then 40km on average except for Queens, Albert and Charlotte County where older adults travel on average between 40km-60km for a general practice health trip at a hospital. General practice trips account for 30% -50% of total health trips to a hospital from 11 counties. Exceptions to this would include Restigouche County which has general practice trips make up greater than 50% of its health trips to a hospital. General practice health trips make up approximately 20%-30% of Kent County and Kings County health trips and less than 10% of Sunbury County's health trips.

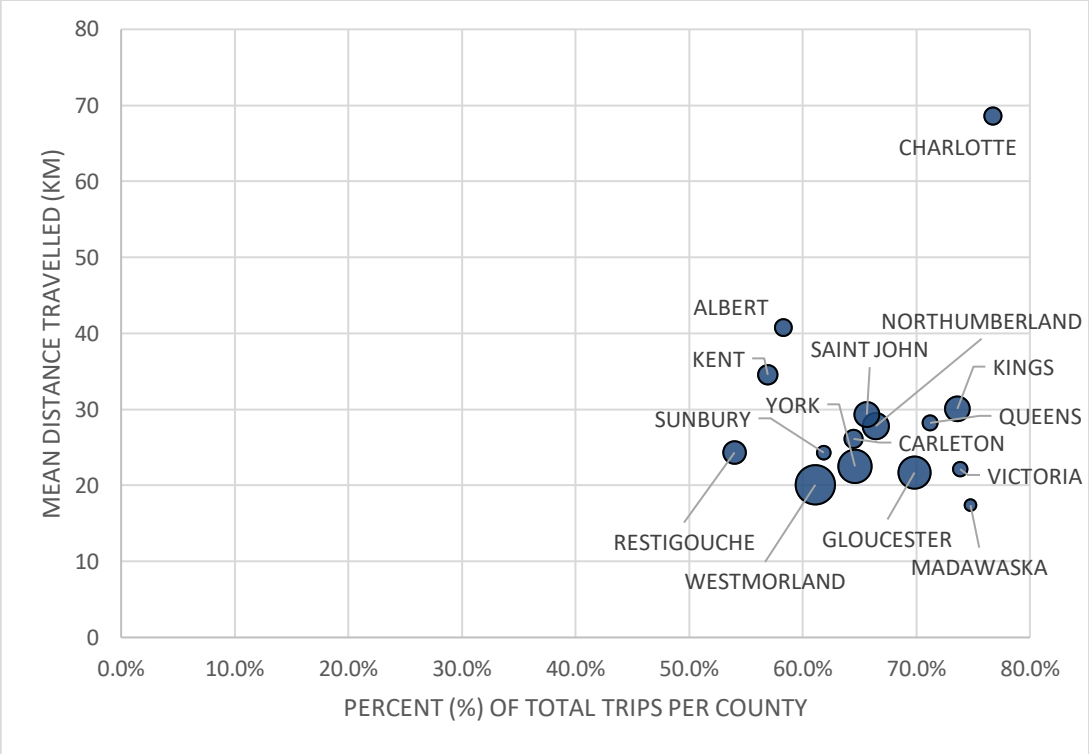

Figure 4-10 - PHB - Travel Distance for Percent of Health Trips Specialized in General Practice

When looking at Figure 4-10, it shows a cluster of counties where older adults travel fewer than 30km to make a general practice health trip to a physician's office/location and where general practice health trips make up 60%-80% of total health trips. One outlier that is obvious when looking at Figure 4-10 is Charlotte County where over 70% of its health trips are for general practice specialties, with an average travel distance of almost 70km.

4.1.3.4 General Surgery

Specialty trips for general surgery are shown in Figure 4-11 and Figure 4-12. These figures show the percent of general surgery trips were made per county on the x-axis and the distance traveled for general surgery per county on the y-axis.

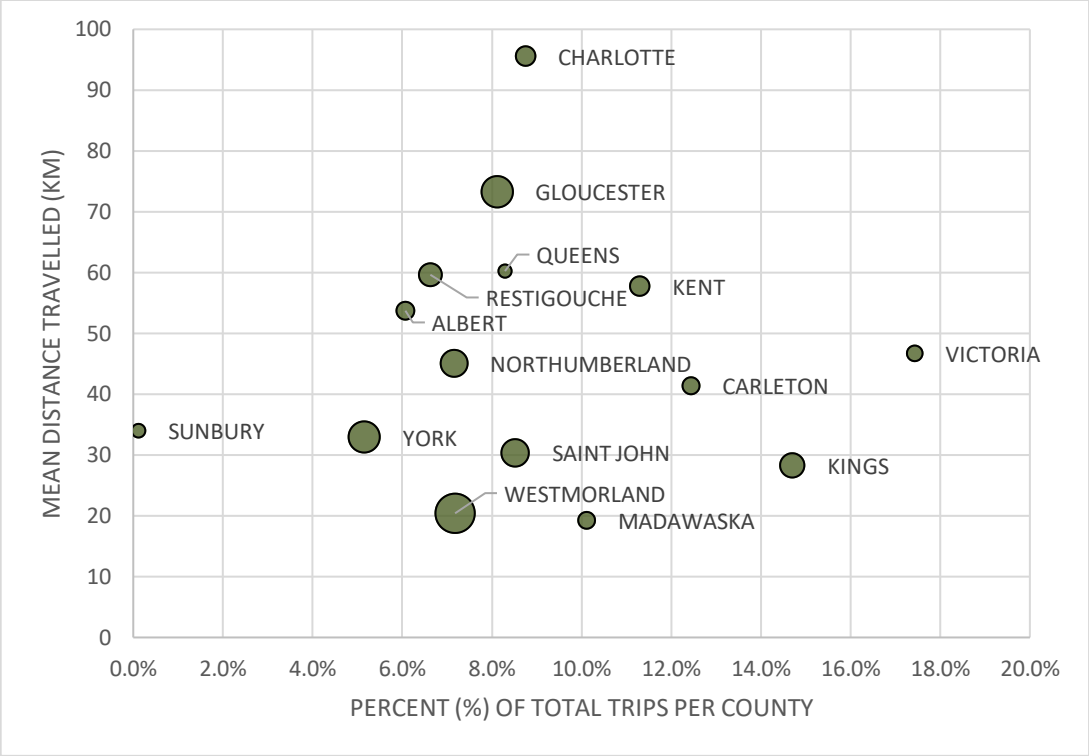

Figure 4-11 - DAD – Travel Distance for Percent of Health Trips Specialized in General Surgery

The majority of general surgery trips from the DAD data range between approximately 6%-12% of all health trips per county, with some outliers within the 14%-18% range and Sunbury County with fewer than 2% of trips. Older adults in 10 counties have an average travel distance less than 50km for general surgery. Notable outlines include Gloucester and Charlotte Counties, where on average older adults travel more than 70km and 90km, respectively, for general surgery.

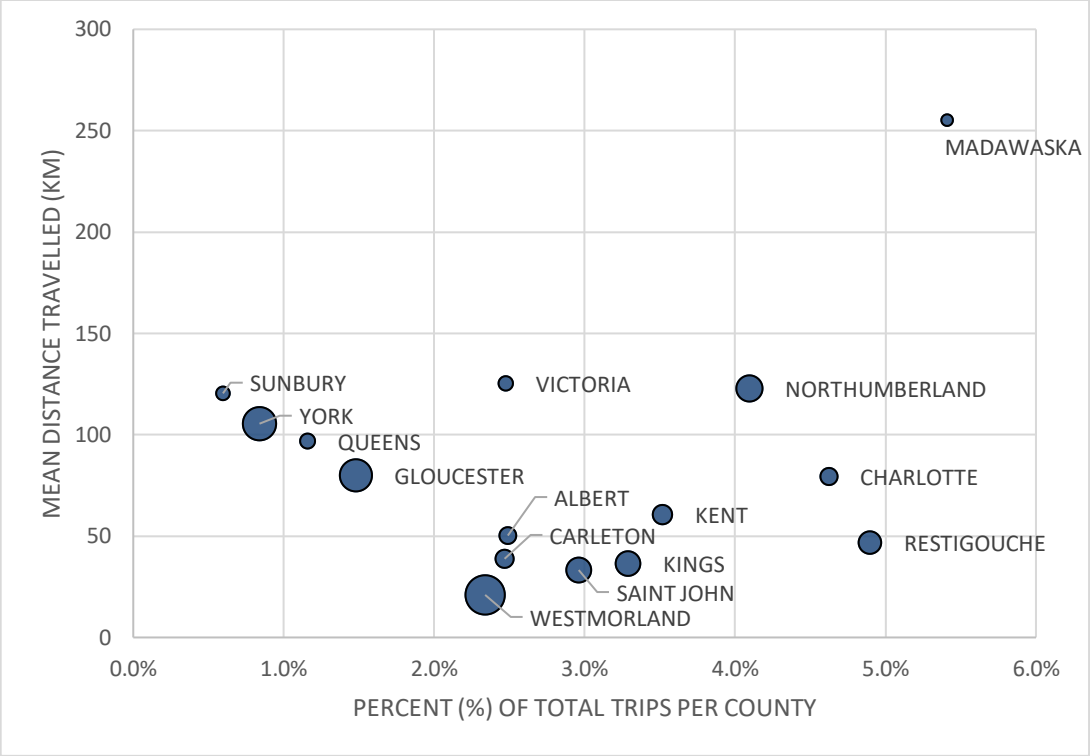

Figure 4-12 - PHB - Travel Distance for Percent of Health Trips Specialized in General Surgery

For the Physician Billing dataset, Figure 4-12 shows a cluster of counties that range between 2%-5% of general surgery trips and where older adults travel 50km or less to a physician's office/location. Notable extremes include Sunbury and York County, where fewer than 1% of trips are for general surgery and older adults travel over 100km one-way on average.

A significant outlier for these data is Madawaska County which has relatively a small percentage of overall health trips (2%), however has General Surgery represent more than 5% of its health trips where older adults travel on average more than 250km to a physician's office/clinic.

4.1.3.5 Ophthalmology

Specialty trips for ophthalmology are shown in Figure 4-13 and Figure 4-14. These figures show the percent of ophthalmology trips were made per county on the x-axis and the distance traveled for ophthalmology per county on the y-axis.

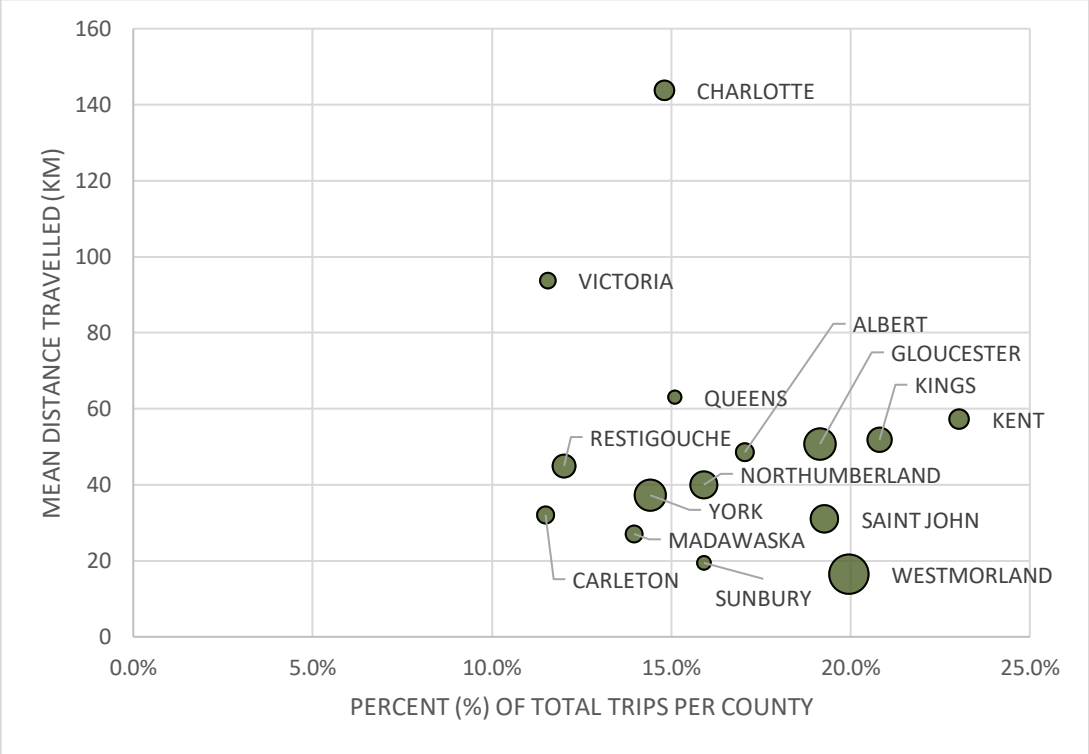

Figure 4-13 - DAD - Travel Distance for Percent of Health Trips Specialized in Ophthalmology

Figure 4-13 presents a cluster of counties where average travel distance to Ophthalmology visits at a hospital is less than 60km. Notable outliers include Victoria County and Charlotte County. Restigouche, Carleton, Madawaska and York County have ophthalmology represent 10-15% of the county's total health trips. Sunbury, Northumberland, Albert, Gloucester, and Saint John County have 15%-20% of their health trips represented by ophthalmology. Ophthalmology represents more than 20% of the health trips to hospitals for older adults in Westmorland, Kings and Kent Counties.

A significant outlier for these data is Charlotte County. Just under 15% of health trips from Charlotte County are for ophthalmology specialty and on average this county travels over 140km to a hospital for this service.

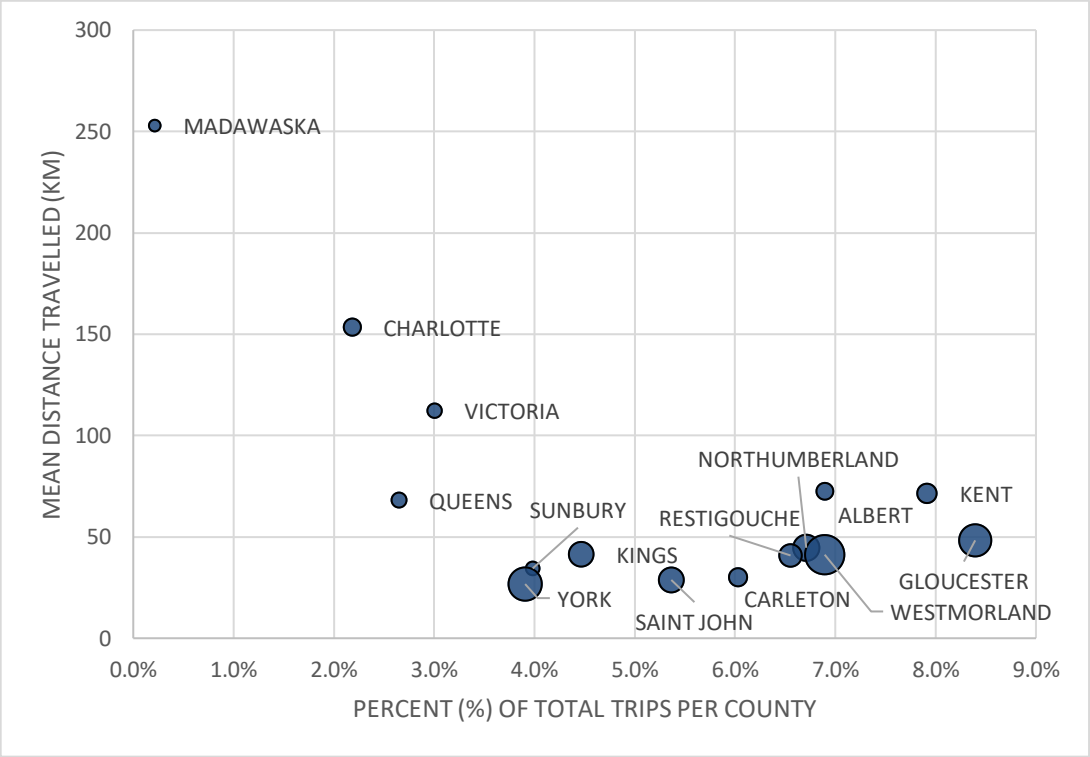

Figure 4-14 - PHB - Travel Distance for Percent of Health Trips Specialized in Ophthalmology

For the Physician Billing data, the majority of counties have average travel distances less than 50km or less to a physician's office/location for ophthalmology which represents 4%-9% of health trips

Queens County has 2%-3% of its health trip designated as ophthalmology and Albert and Kent County have 7%-8% designated to ophthalmology with each travelling between 50km-100km on average. Victoria County and Charlotte County have 2%-3% of their health trips designated to ophthalmology and travel over 100km, and 150km respectively.

4.1.3.6 Orthopedic Surgery

Specialty trips for orthopedic surgery are shown in Figure 4-15 and Figure 4-16. These figures show the percent of orthopedic surgery trips were made per county on the x-axis and the distance traveled for orthopedic surgery per county on the y-axis.

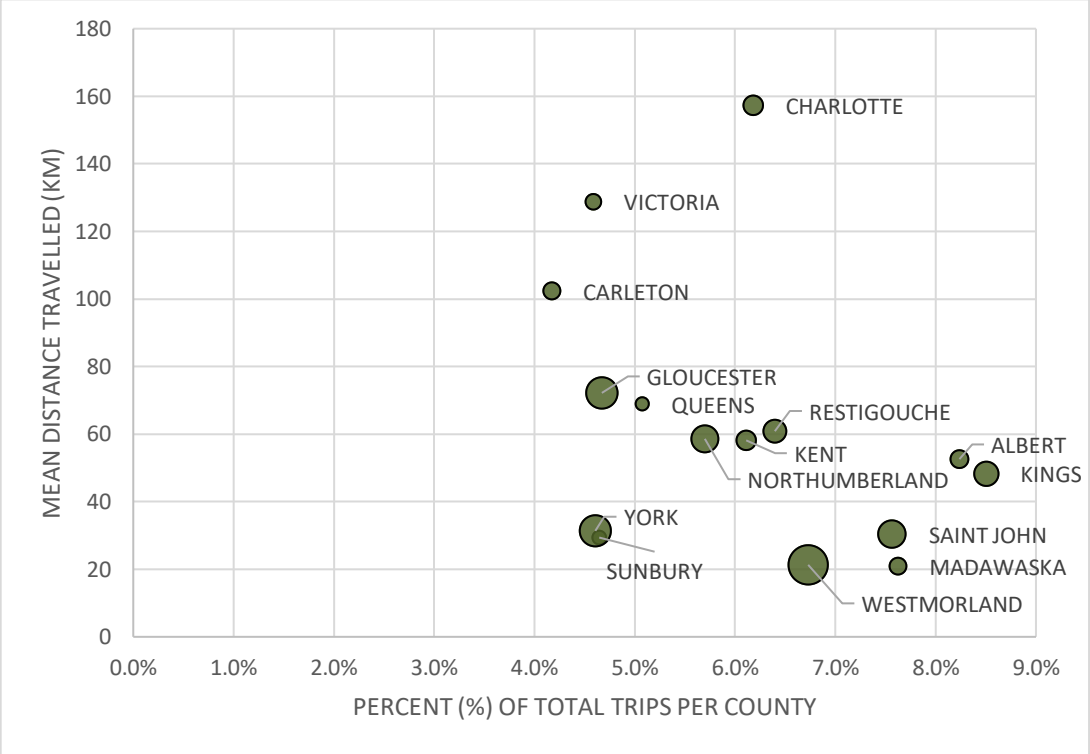

Figure 4-15 - DAD - Travel Distance for Percent of Health Trips Specialized in Orthopedic Surgery

Figure 4-15 above, shows the following observations:

a. Westmorland County, Saint John County and Madawaska County have 6%-8% of their health trips designated to orthopedic surgery and are required to travel on average between 20-40km to a hospital.

b. Orthopedic surgery makes up 4%-5% of health trips for York County and Sunbury County and are required to travel on average between 20-40km to a hospital.

c. Orthopedic surgery makes up 8%-9% of health trips for Albert County and Kings County which is the highest percentage of all the counties. These counties are required to travel between 40-60km to a hospital for orthopedic surgery.

d. Orthopedic surgery makes up 5.5%-6.5% of health trips for Northumberland, Kent and Restigouche County who are also required to travel an average of 60km to a hospital for orthopedic surgery.

e. Queens County and Gloucester County travel on average approximately 70km to a hospital for orthopedic surgery which makes up 4.5%-5% of their health trips

f. Orthopedic Surgery makes up 4%-5% of health trips for Carleton County and Victoria County and they travel on average approximately 100km and 130km respectively to a hospital.

g. Charlotte County experiences a little over 6% of their health trips to be orthopedic surgery for which they are required to travel on average just below 160km.

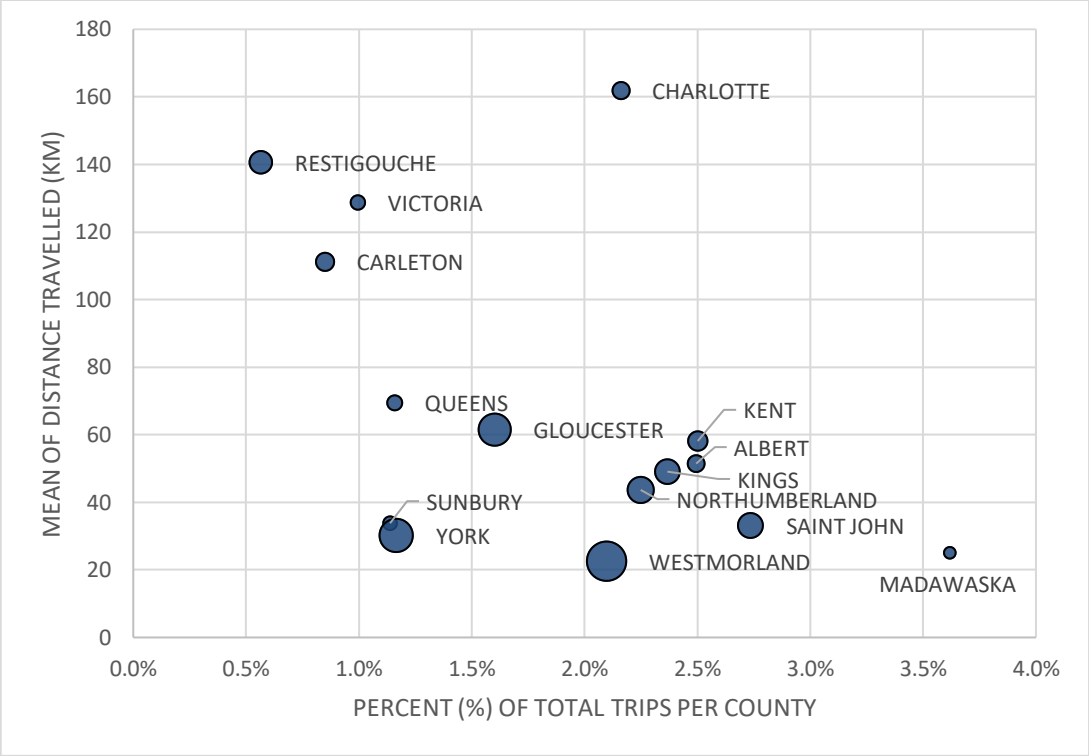

Figure 4-16 - PHB - Travel Distance for Percent of Health Trips Specialized in Orthopedic Surgery

Figure 4-16 above, shows the following observations:

a. Older adults in York, Sunbury, Westmorland, Saint John and Madawaska County travel on average 20km-40km to a physician's office/ location for orthopedic surgery. Orthopedic surgery makes up 1%-4% of these county health trips.

b. Orthopedic surgery makes up 1%-2.5% of health trips for Queens, Gloucester, Northumberland, Kings, Albert and Kent County which are required to travel on average between 40-70km to a physician's office/ location.

c. Orthopedic surgery makes up 0.5%-1% of health trips for Carleton, Victoria and Restigouche County which are required to travel approximately 110km, 130km and 140km on average to a physician's office/ location for orthopedic surgery.

d. Older adults in Charlotte County travel on average 160km to a physician's office/ clinic for orthopedic surgery which makes up 2%-2.5% of their total health trips.

4.1.3.7 Urology

Specialty trips for urology are shown in Figure 4-17 and Figure 4-18. These figures show the percent of urology trips were made per county on the x-axis and the distance traveled for urology per county on the y-axis.

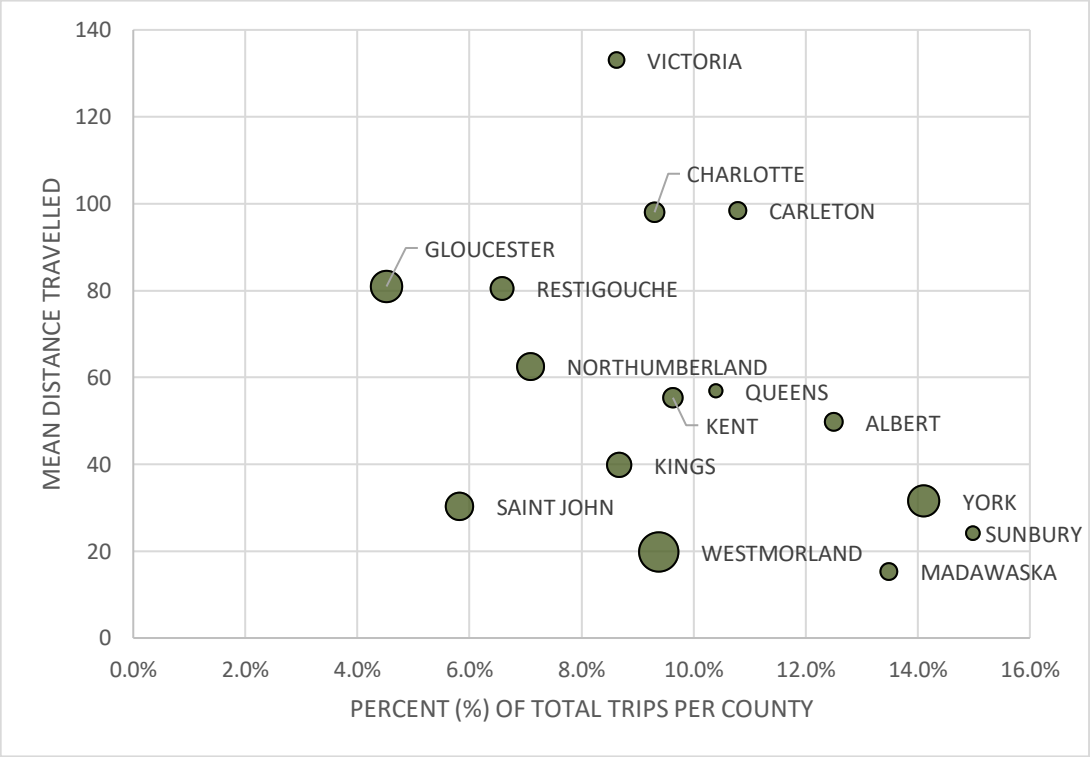

Figure 4-17 - DAD - Travel Distance for Percent of Health Trips Specialized in Urology

Figure 4-17 above, shows the following observations:

a. York County and Sunbury County have 14%-16% of their health trips associated with Urology with older adults travelling on average 20km-40km to a hospital.

b. Older adults from Madawaska County and Westmorland County travel 20km or less on average to a hospital for urology, accounting for 12%-14% and 8%-10% of their health trips respectively.

c. Older adults from Saint John County and Kings County travel between 30km-40km on average for urology trips making up 6%-9% of their health trips.

d. Older adults from Northumberland, Kent, Queens and Albert County travel an average of 40-50km to a hospital for urology. Urology makes up approximately 7% of Northumberland County health trips, 9%-11% of Kent County and Queens County health trips, and 12%-13% of Albert County health trips.

e. Older adults from Gloucester and Restigouche County are required to travel on average 80km to a hospital for urology which makes up 4%-7% of county health trips.

f. Urology makes up approximately 9%-11% of health trips for Charlotte and Carleton County. These counties are required to travel an average of 100km to a hospital for urology.

g. Victoria County has 8%-9% of health trips designated to urology for which they are required to travel just below 140km on average to a hospital.

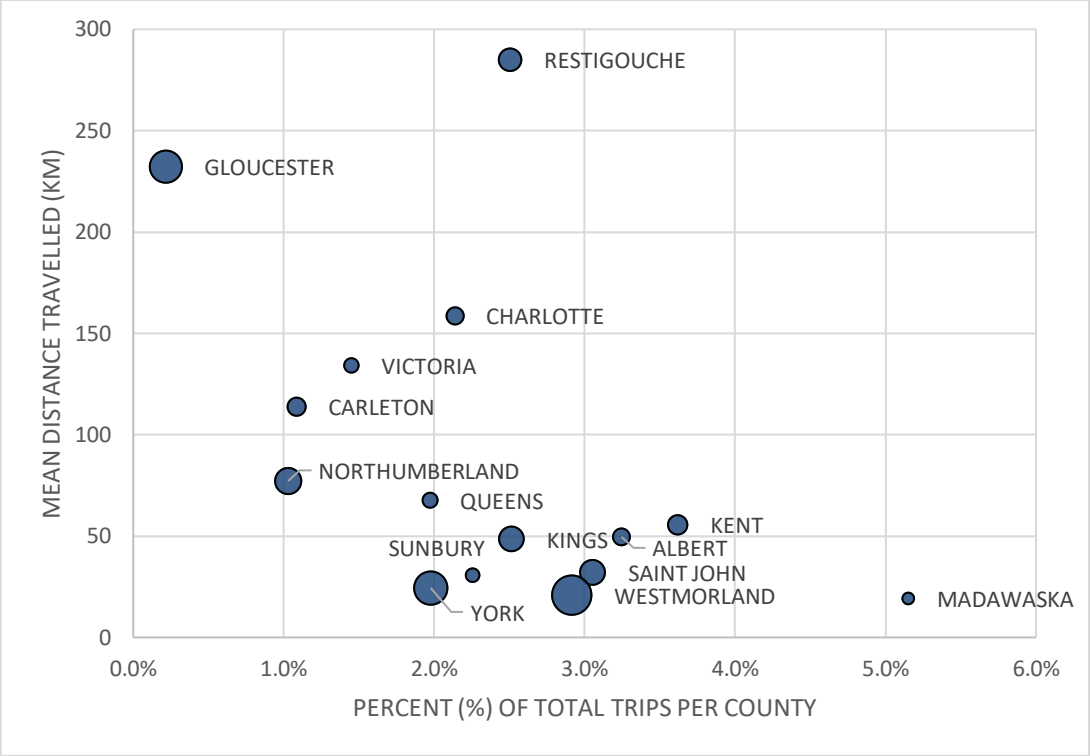

Figure 4-18 - PHB - Travel Distance for Percent of Health Trips Specialized in Urology

Figure 4-18 above, shows the following conclusions:

a. The majority of counties have average travel distances of 50km or less to see a physician regarding urology. Urology makes up 2%-4% of trips for these counties including Sunbury, York, Kings, Westmorland, Saint John, Albert, and Kent County. Madawaska County designates 5%-6% of trips to a physician for urology.

b. Older adults from Northumberland and Queens County travel between 50km-100km to a physician for urology which makes up 1%-2% of their health trips.

c. Older adults from Carleton and Victoria County travel between 100km-150km to a physician for urology which makes up 1%-2% of their health trips.

d. Urology makes up approximately 2% of health trips to a physician for Charlotte County of which they travel over 150km on average.

e. Older adults from Gloucester County travels an average of 200-250km to a physician for urology; however, this specialty makes up less than 1% of their trips.

f. Restigouche County has urology make up approximately 2.5% of their health trips to a physician's office/ location and are required to travel almost 300km on average.

4.1.4 Summary:

The following conclusions were obtained from analyzing the descriptive statistics for both the DAD and Physician Billing datasets:

a. From the 2011 Census population data, the New Brunswick counties with the greatest percentage of older adults aged 65+ include: Queens County and Restigouche County with 24% and 20% respectively, followed by Gloucester County, Kent County and Northumberland County. Charlotte County held the next greatest percent population with 18%.

b. The percentage of health trips made from each county was comparable to the population of older adults within the county for both datasets.

c. The DAD dataset indicates that:

- 39.5% of health trips to a hospital were for general practice
- 11.5% of health trips were for medical specialty
- 49.0% of health trips were for surgical specialty

d. The Physician Billing dataset indicates that:

- 65.3% of health trips to see a physician (at office, clinic, etc.) are for general practice reasons
- 19.2% of health trips to see a physician (at office, clinic, etc.) are for medical specialties
- 15.5% of health trips to see a physician (at office, clinic, etc.) are for surgical specialties

e. Trips for medical specialties and surgical specialties travel greater distances than general practice trips for both datasets.

f. The 40% of general practice health trips to hospital trips are produced from age groups 75-79 and 80-84. The 49% of general practice trips to a physician are produced from age groups 65-69 and 70-74

g. The majority of specialty health trips are produced from age groups 65-69, 70-74 and 75-79 for both datasets ranging between 70% -75%.

h. On average, the counties with the highest average one-way travel distance are originating from:

- Charlotte County – 92 km
- Victoria County – 69.5 km
- Gloucester County – 66.5 km

These 3 counties account for 19% of health trips to hospitals for adults aged 65+ for the province.

i. Approximately 27% of older adults travel an average of over 60km for a health trips to a hospital.

j. Approximately 49% of older adult health trips to the hospital will be over an average length of 50km.

k. For General practice health trips, Queens County and Albert County have the longest average travel distances to a hospital, likely due to there not being a hospital in those counties. For health trips to a physician, Charlotte County has an average distance of 68 km for general practice trips to a physician.

l. On average, the counties with the greatest average one-way travel distance to a physician are:

- Charlotte County – 85.0 km
- Restigouche – 56.0 km
- Northumberland – 53.0 km

These 3 counties account for approximately 18% of adults aged 65+ travelling to a physician for a health trip.

m. For Specialty trips to a hospital, older adults in Charlotte County, Restigouche County and Victoria County all travel over 100km on average. For specialty trips to a physician, Charlotte County, Northumberland County and Victoria County all travel over 100km for a health trip.

n. There is a linear relationship between the average distance traveled for a health trips per county vs. the percent of health trips a county produces for both datasets. It appears that the counties travelling the greatest distances produce

fewer percentages of health trips whereas the counties with the greater percentage of trips tend to have a lower average travel distance.

o. Diagnostic Radiology accounts for 10%-20% of all trips to a physician for York, Queens, Victoria, Carleton, Restigouche and Sunbury Counties.

p. Ophthalmology makes up 10%-25% of all hospital trips for all counties. This particular specialty group is significant because typically patients are unable to drive themselves to/from the health trip.

4.2 Regression Analysis

4.2.1 Linear Regression Analysis using DAD Data

Linear regression analysis was completed using the DAD dataset with a total of 171,150 observation points. This model was created using variables from the health data to predict the distance (km) an older adult (65+) is expected to travel for a health trip to a hospital. Three models were produced as shown in Figure 4-19, which included two models with health variables and one model with variables which would be available from census data. Forward stepwise regression was used to build the models while looking at significant values within the regression analysis. This included a combination of the adjust R^2 value to determine how well the model fits the data, the p-value for statistical significance, the variance inflation factor (VIF) to indicate multicollinearity and the coefficients of the independent variables to indicate if they were justified in relation to the dependent variable, distance (km). The p-values were determined using a 95% confidence interval meaning that if less than 0.05 the variable was deemed to be statistically significant. This means that the null hypothesis, that the independent variable has no correlation, is rejected. VIF was

used to determine the presence of multicollinearity using a scale between 1-10. A VIF of 1.0 is the smallest possible value which indicates absolutely no multicollinearity is present, a VIF limit of 4 was used and an ideal VIF threshold of 2.5 was used to indicate minimal multicollinearity was present.

DAD DATA																		
	MODEL 1A - DAD HEALTH VARIABLES						**MODEL 1B - DAD HEALTH VARIABLES**						**MODEL 1C - DAD CENSUS VARIABLES**					
	TRIP DISTANCE - LINEAR REGRESSION						TRIP DISTANCE - LINEAR REGRESSION						TRIP DISTANCE - LINEAR REGRESSION					
Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	18.656	1.327	14.060	<.0001	.	0.000	-43.319	1.192	-36.340	<.0001	.	0.000	100.752	1.447	69.640	<.0001	.	0.000
MAJOR1	-18.899	0.351	-53.860	<.0001	0.466	2.144	-10.539	0.308	-34.250	<.0001	0.466	2.148	-	-	-	-	-	-
MAJOR2	27.653	0.431	64.150	<.0001	0.731	1.367	30.426	0.375	81.050	<.0001	0.737	1.358	-	-	-	-	-	-
CD1	-2.923	0.697	-4.200	<.0001	0.744	1.343	3.744	0.606	6.180	<.0001	0.742	1.348	15.951	0.778	20.500	<.0001	0.764	1.309
CD2	47.424	0.729	65.070	<.0001	0.747	1.339	83.901	0.653	128.520	<.0001	0.691	1.448	21.280	0.802	26.530	<.0001	0.777	1.286
CD3	64.295	0.653	98.420	<.0001	0.720	1.388	88.492	0.579	152.940	<.0001	0.696	1.437	56.331	0.737	76.420	<.0001	0.728	1.373
CD4	35.563	0.499	71.250	<.0001	0.550	1.818	96.198	0.523	184.040	<.0001	0.381	2.623	30.256	0.562	53.850	<.0001	0.560	1.784
CD5	7.064	0.654	10.800	<.0001	0.733	1.363	15.847	0.567	27.940	<.0001	0.727	1.376	18.709	0.733	25.510	<.0001	0.738	1.355
CD6	-6.067	0.576	-10.540	<.0001	0.651	1.536	7.860	0.501	15.690	<.0001	0.640	1.563	4.185	0.646	6.480	<.0001	0.654	1.528
CD7	3.123	0.722	4.320	<.0001	0.773	1.293	69.973	0.702	99.660	<.0001	0.612	1.634	-11.130	0.810	-13.750	<.0001	0.782	1.279
CD8	31.346	0.540	58.000	<.0001	0.606	1.651	78.824	0.523	150.590	<.0001	0.489	2.043	20.877	0.608	34.360	<.0001	0.617	1.621
CD9	31.575	0.850	37.160	<.0001	0.844	1.185	36.317	0.738	49.240	<.0001	0.844	1.185	26.982	0.960	28.100	<.0001	0.846	1.182
CD10	39.534	0.591	66.850	<.0001	0.666	1.502	96.123	0.583	164.990	<.0001	0.523	1.913	26.439	0.665	39.730	<.0001	0.681	1.469
CD11	-29.792	0.545	-54.680	<.0001	0.578	1.730	-21.881	0.474	-46.210	<.0001	0.581	1.721	-6.199	0.603	-10.290	<.0001	0.610	1.640
CD12	2.214	0.834	2.650	0.008	0.827	1.209	-1.340	0.723	-1.850	0.064	0.828	1.207	-6.890	0.940	-7.330	<.0001	0.832	1.202
CD13	56.972	0.768	74.170	<.0001	0.771	1.298	104.603	0.705	148.420	<.0001	0.693	1.442	32.577	0.855	38.100	<.0001	0.800	1.250
CD14	-29.481	0.460	-64.150	<.0001	0.459	2.180	-25.459	0.400	-63.680	<.0001	0.456	2.193	-14.670	0.512	-28.680	<.0001	0.473	2.114
Max_Beds	0.152	0.001	160.580	<.0001	0.607	1.647	0.070	0.001	77.940	<.0001	0.512	1.952	-	-	-	-	-	-
Q1	4.700	0.396	11.870	<.0001	0.526	1.902	5.763	0.343	16.780	<.0001	0.530	1.887	4.666	0.447	10.430	<.0001	0.531	1.884
Q2	3.198	0.389	8.230	<.0001	0.531	1.882	4.782	0.337	14.190	<.0001	0.534	1.871	3.116	0.439	7.100	<.0001	0.535	1.869
Q3	3.848	0.399	9.630	<.0001	0.561	1.782	4.295	0.346	12.430	<.0001	0.564	1.774	4.757	0.450	10.560	<.0001	0.564	1.773
Q4	3.552	0.398	8.930	<.0001	0.554	1.805	4.314	0.344	12.530	<.0001	0.556	1.798	4.014	0.449	8.950	<.0001	0.556	1.797
ADMIT1	-8.135	0.328	-24.810	<.0001	0.526	1.901	-3.614	0.288	-12.540	<.0001	0.509	1.965	-	-	-	-	-	-
Age	-0.200	0.016	-12.460	<.0001	0.894	1.118	-0.202	0.014	-14.440	<.0001	0.897	1.114	-0.933	0.017	-53.410	<.0001	0.979	1.021
gender2	3.398	0.239	14.190	<.0001	0.976	1.024	2.988	0.208	14.380	<.0001	0.977	1.024	6.961	0.270	25.770	<.0001	0.982	1.019
len_of_stay	0.007	0.003	2.140	0.032	0.957	1.045	-	-	-	-	-	-	-	-	-	-	-	-
dCITYv3	-	-	-	-	-	-	80.622	0.381	211.610	<.0001	0.317	3.151	-	-	-	-	-	-
Adj R-Sq	0.357	-	-	-	-	-	0.487	-	-	-	-	-	0.129	-	-	-	-	-
F Value	3578.83	-	-	-	-	-	6503.07	-	-	-	-	-	1264.28	-	-	-	-	-
Pr > F	<.0001	-	-	-	-	-	<.0001	-	-	-	-	-	<.0001	-	-	-	-	-

Note: these models only include persons ages 65+ and their recorded trips

Figure 4-19 - DAD Regression Models 1A to 1C

Linear regression models were created with health variables and census variables using forward regression. This process resulted in two models that were considered as final models, **Model 1A** and **Model 1B**.

The following variables, with corresponding dummy variables, in Table 4-9 were used as to develop the **Model 1A**.

Table 4-9 - Model 1A Regression Variables

DAD DATA	
Variable	Description
MAJOR 1	General Practice
MAJOR 2	Medical Specialty
Dummy Variable	Surgical Specialty
CD1	Albert County
CD2	Carleton County
CD3	Charlotte County
CD4	Gloucester County
CD5	Kent County
CD6	Kings County
CD7	Madawaska County
CD8	Northumberland County
CD9	Queens County
CD10	Restigouche County
CD11	Saint John County
CD12	Sunbury County
CD13	Victoria County
CD14	Westmorland County
Dummy Variable	York County
Max Beds	Number of Beds at Destination Location
Q1	Low Income for Origin DA
Q2	Low-Medium Income for Origin DA
Q3	Medium Income for Origin DA
Q4	Medium-High Income for Origin DA
Dummy Variable	High Income for Origin DA
ADMIT1 Category	Urgent or Elective Admission
Age	Age in years
Gender	Female =0, Male =1
Length of Stay	Length of Stay in number of days

Model 1A produced an adjusted R^2 of 0.357 which is a relatively good fit to the data for this type of model due to socioeconomic factors and high behaviour variability. All independent variables produced a p-value of less than 0.05 meaning they are statistically significant in the model. The VIF of the variables all fall below a value of 2.5; however, *Major 1* and *CD14* produced a VIF of 2.144 and 2.18 which was flagged as approaching

the ideal threshold of 2.5 however still less than a VIF of 4. The values of the independent variable coefficients were reviewed to ensure they were justified based on their relationship with the dependent variable, distance traveled (km).

The *Major 1* variable produced a -18.65 coefficient for general practice trips meaning that the distance traveled (km) decreased if present in the model which is likely as adults 65+ tend to travel less distances for general practice trips than specialty trips. Two counties provided large negative numbers which suggest significantly lower km traveled if originating from that county. This was accurate when looking at the two counties, Westmorland County (-29.48) and Saint John County (-29.79) which travel some of the lowest distances for hospital trips.

Age produced a coefficient of -0.20 which suggests that as age increases, the distance of travel to a hospital decreases by 0.2 km per year of age.

County variables *CD3* for Charlotte County (64.29) and *CD13* for Victoria County (56.97) produced large positive coefficients, this suggests that the distance travel for a health trip increases if originating from these counties. This would be the kilometers traveled if a value of 1 is used for the county variable, indicating the health trip is originating from that county.

The following variables, with corresponding dummy variables, in Table 4-10 were used as to develop the **Model 1B**.

Table 4-10 - Model 1B Regression Variables

DAD DATA	
Variable	Description
MAJOR 1	General Practice
MAJOR 2	Medical Specialty
Dummy Variable	Surgical Specialty
CD1	Albert County
CD2	Carleton County
CD3	Charlotte County
CD4	Gloucester County
CD5	Kent County
CD6	Kings County
CD7	Madawaska County
CD8	Northumberland County
CD9	Queens County
CD10	Restigouche County
CD11	Saint John County
CD12	Sunbury County
CD13	Victoria County
CD14	Westmorland County
Dummy Variable	York County
Max Beds	Number of Beds at Destination Location
Q1	Low Income for Origin DA
Q2	Low-Medium Income for Origin DA
Q3	Medium Income for Origin DA
Q4	Medium-High Income for Origin DA
Dummy Variable	High Income for Origin DA
ADMIT1 Category	Urgent or Elective Admission
Age	Age in years
Gender	Male or Female
dCityv3	Rural/ Urban Destination (3 major Cities)

Model 1B produced an adjusted R^2 of 0.487 which suggests a greater fit to the data than Model 1A. All independent variables produced a p-value of less than 0.05 meaning they are statistically significant in the model except for *CD12* which produced a value of 0.064. This means that it is a greater likelihood that the *CD12* variable, Sunbury County is equal to 0. The VIF of the variables all fall below a value of 2.5 however except for *CD4*, Gloucester County and *dCITYv3*, the rural/urban indictor for the three major cities. Variable *CD4* produced a VIF value of 2.62 and *dCITYv3* produced a VIF of 3.15 which were both noted to be greater than the ideal 2.5 threshold however still less than a VIF of 4. The values of the independent variable coefficients were reviewed to ensure they were justified based on their relationship with the dependent variable, distance traveled (km).

Major 1 produced a -10.54 coefficient for general practice trips meaning that the distance traveled (km) decreased if present in the model which is likely as adults 65+ tend to travel less distances for general practice trips then specialty trips. Two counties, Saint John County (*CD11*) and Westmorland County (*CD12)* provided large negative numbers which suggest significantly lower distance traveled if originating from that county. This was accurate when looking at the two counties, Westmorland County (-25.46) and Saint John County (-21.88) which travel some of the lowest distances for hospital trips.

Age produced a coefficient of -0.20 which suggests that as age increased, the distance of travel to a hospital decreased by 0.2 km per year of age.

The last regression model was created using variables that would be available from census data which included age, gender, income level and origin county. This model was created to be compared with the regression model including health variables. The census **Model 1C** produced a model with an adjusted R^2 value of 0.129. All p-values were less than 0.05 meaning the variables were statistically significant and the greatest value of VIF was 2.11 for Westmorland County (*CD14)* which is below the ideal value of 2.5.

4.2.2 Final Model

Model 1B was determined to be the final model for the DAD data which resulted in an R^2 value of 0.487. Even though model 1B produced a higher adjusted R^2 than model 1A, it also included indicators that needed to be evaluated before determining it as the final model.

The variable for Sunbury County (*CD12*) was determined to not be statistically significant based on a 95% confidence interval with a p-value of 0.064. This means that the null

hypothesis is rejected and there is a greater probability that the value is zero and there is no correlation to the independent variable. Consideration was given to rejecting Model 1B due to the statistical significance of this variable; however, because Sunbury County has the lowest percentage of adults 65+ and generate a low number of trips, the remaining confidence of 93.6% was determined to be acceptable. The rural/urban indicator variable *dCITYv3* appeared to have a significant influence on the adjusted R^2 value and resulted in a VIF higher than the ideal value of 2.5; however, because it was still less than 4, it was considered to be acceptable.

The following equation (14) represents the final regression model for estimating distance traveled for a health trip. This model was developed using the DAD data and includes health related variables as well as geographic and demographic variables.

MODEL 1B: adj. R^2 = 0.487 **(14)**

$$
\begin{aligned}
Distance\ (km)_{DAD-1B} \\
&= -43.32 - 10.54\boldsymbol{m1} + 30.43\boldsymbol{m2} + 3.74\boldsymbol{CD1} + 83.9\boldsymbol{CD2} + 88.49\boldsymbol{CD3} + 96.20\boldsymbol{CD4} \\
&+ 15.85\boldsymbol{CD5} + 7.86\boldsymbol{CD6} + 69.97\boldsymbol{CD7} + 78.82\boldsymbol{CD8} + 36.32\boldsymbol{CD9} + 96.12\boldsymbol{CD10} \\
&- 21.88\boldsymbol{CD11} - 1.34\boldsymbol{CD12} + 104.60\boldsymbol{CD13} - 25.46\boldsymbol{CD14} + 0.07\boldsymbol{maxbeds} \\
&+ 5.76\boldsymbol{Q1} + 4.78\boldsymbol{Q2} + 4.29\boldsymbol{Q3} + 4.31\boldsymbol{Q4} - 3.61\boldsymbol{Admit1} - 0.20\boldsymbol{Age} + 2.99\boldsymbol{gender} \\
&+ 80.62\boldsymbol{dCITYv3}
\end{aligned}
$$

The following examples were used to demonstrate the final **Model 1B:**

Example 1:

Gender:	Female	**Estimated Travel Distance** **142 km**
Age:	86	
Origin County:	Gloucester County	
Health Trip Type:	Surgical Specialty	
Income DA Level:	Low	
Destination Type (R/U):	Urban	
Number of beds:	300	
Admission Category	Elective (non-urgent)	

$$Distance\ (km)_{DAD-1B}$$

$$= -43.32 - 10.54(\mathbf{0}) + 30.43(\mathbf{0}) + 3.74(\mathbf{0}) + 83.9(\mathbf{0}) + 88.49(\mathbf{0}) + 96.20(\mathbf{1})$$

$$+ 15.85(\mathbf{0}) + 7.86(\mathbf{0}) + 69.97(\mathbf{0}) + 78.82(\mathbf{0}) + 36.32(\mathbf{0}) + 96.12(\mathbf{0}) - 21.88(\mathbf{0})$$

$$- 1.34(\mathbf{0}) + 104.60(\mathbf{0}) - 25.46(\mathbf{0}) + 0.07(\mathbf{300}) + 5.76(\mathbf{1}) + 4.78(\mathbf{0}) + 4.2(\mathbf{0})$$

$$+ 4.31(\mathbf{0}) - 3.61(\mathbf{0}) - 0.20(\mathbf{86}) + 2.99(\mathbf{0}) + 80.62(\mathbf{1})$$

When comparing the outcome of Example 1 to the distance statistics, the average travel for Gloucester is 93km for specialty trips. Although the average travel distance for specialty trips includes both medical specialty and surgical specialty trips, this is still a good indication that an older adult from Gloucester County is expected to travel a great distance. This model indicates that the above attributes will estimate a high average travel distance.

Example 2a:

Gender:	Male	**Estimated Travel Distance** **37 km**
Age:	65	
Origin County:	Charlotte County	
Health Trip Type:	General Practice	
Income DA Level:	Low	
Destination Type (R/U):	Rural	
Number of beds:	100	
Admission Category	Elective (non-urgent)	

In example 2a, an older adult is travelling for general practice to another rural county or within their own county which results in a lower travel rate. Looking at the descriptive statistics for Charlotte County for general practice trips, the data indicates on average that older adults will travel 40km. This model appears to estimate trip distance well for the above attributes.

It should be noted that if the same attributes are held true however, the type of destination changes to urban, the estimated trip distance for the attributes shown in Example 2b increases significantly.

Example 2b:

Gender:	Male	**Estimated Travel Distance** **121 km**
Age:	65	
Origin County:	Charlotte County	
Health Trip Type:	General Practice	
Income DA Level:	Low	
Destination Type (R/U):	Urban	
Number of beds:	100	
Admission Category	Elective (non-urgent)	

Example 3:

Lastly, Example 3 demonstrated an older adult making a general practice health trip originating from one of the three urban counties, York County, and travelling to an urban hospital. When comparing back to the descriptive statistics, the average travel distance for older adults from York County is 21 km. Based on these attributes, this model appears to accurately estimate a reasonable travel distance.

Gender:	Male	**Estimated Travel Distance** **26 km**
Age:	76	
Origin County:	York County	
Health Trip Type:	General Practice	
Income DA Level:	Medium	
Destination Type (R/U):	Urban	
Number of beds:	100	
Admission Category	Elective (non-urgent)	

It should be noted that when applying examples to the regression model, that realistic situations should be considered. For example, it is unlikely that someone will be travelling to a rural county to seek a surgical specialty or that that the hospital for a surgical specialty will have 0 hospital beds. It is also unlikely that a person originating from an urban area will travel to a rural area for a hospital trip.

The following equation (15) represents the regression model for estimating distance traveled for a health trip using only demographic and geographic variables which would be available from census data. This model was also developed using the DAD dataset.

MODEL 1C: adj. R^2 = 0.129 **(15)**

$$Distance\ (km)_{DAD-1C} = 100.75 + 15.95\boldsymbol{CD1} + 21.28\boldsymbol{CD2} + 56.33\boldsymbol{CD3} + 30.26\boldsymbol{CD4} + 18.71\boldsymbol{CD5} + 4.19\boldsymbol{CD6} - 11.13\boldsymbol{CD7} + 20.88\boldsymbol{CD8} + 26.98\boldsymbol{CD9} + 26.44\boldsymbol{CD10} - 6.20\boldsymbol{CD11} - 6.89\boldsymbol{CD12} + 32.58\boldsymbol{CD13} - 14.67\boldsymbol{CD14} + 4.67\boldsymbol{Q1} + 3.12\boldsymbol{Q2} + 4.76\boldsymbol{Q3} + 4.01\boldsymbol{Q4} - 0.93\boldsymbol{Age} + 6.96\boldsymbol{gender2}$$

4.2.3 Summary

The following conclusions were made from the linear regression analysis using the DAD data:

a. The final health **model 1B** for the DAD data resulted in an adjusted R^2 value of 0.487 which means the data fit the model relatively well considering the socioeconomic nature and factors of the data.

b. The census **model 1C** for the DAD data resulted in an adjusted R^2 value of 0.129 which was based off of only variables which would be available within the census data such as New Brunswick counties, age, gender, DA income.

c. All values within the models resulted in a VIF of 4 or less which indicates a low probability of multicollinearity being present within the models.

d. The age variable had a very low coefficient and a negative value meaning that as age increased, the travel distance decreased by 0.02km per year of age. It is expected that the age variable would have more of a contribution if other ages below 65 years were included in the model.

e. The coefficients for the urban origin county variables appeared to decrease travel significantly while the more rural counties increased travel significantly the more rural the county.

f. The type of health trip resulted in significantly more travel for specialty trips as opposed to general practice trips.

g. The destination rural/urban indicator had a significant influence on the R^2 for the model and resulted in a large positive coefficient.

4.2.4 Linear Regression Analysis using Physician Billing Data

Linear regression analysis was competed using the Physician Billing dataset with a total of 2,882,390 observation points. Three models were produced as shown in Figure 4-20, which included two models with health variables and one model with variables available from census data. Forward linear regression was used to build the models while looking at significant values within the regression analysis in the same way as discussed previously using the DAD data.

PHYSCIAN BILLING DATA																		
	MODEL 2A - PHYSCIAN BILING HEALTH VARIABLES						**MODEL 2B - PHYSCIAN BILING HEALTH VARIABLES**						**MODEL 2C - PHYSCIAN BILING CENSUS VARIABLES**					
	TRIP DISTANCE - LINEAR REGRESSION						TRIP DISTANCE - LINEAR REGRESSION						TRIP DISTANCE - LINEAR REGRESSION					
Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	-21.354	0.321	-66.570	<.0001	.	0.000	-0.408	0.328	-1.250	0.213	.	0.000	45.518	0.343	132.790	<.0001	.	0.000
PHY_TYPE1	21.370	0.074	289.110	<.0001	0.675	1.482	-	-	-	-	-	-	-	-	-	-	-	-
MAJOR1	-	-	-	-	-	-	-20.901	0.087	-241.610	<.0001	0.493	2.030	-	-	-	-	-	-
MAJOR2	-	-	-	-	-	-	0.945	0.103	9.170	<.0001	0.515	1.942	-	-	-	-	-	-
CD1	9.356	0.173	54.210	<.0001	0.808	1.238	9.380	0.173	54.340	<.0001	0.807	1.239	15.130	0.194	77.900	<.0001	0.812	1.232
CD2	55.032	0.176	313.270	<.0001	0.652	1.534	55.036	0.176	312.790	<.0001	0.653	1.532	14.936	0.181	82.570	<.0001	0.785	1.274
CD3	80.935	0.172	469.430	<.0001	0.753	1.328	80.444	0.173	464.930	<.0001	0.754	1.327	54.230	0.189	286.190	<.0001	0.798	1.253
CD4	54.602	0.141	387.660	<.0001	0.394	2.535	54.568	0.141	387.430	<.0001	0.394	2.537	7.298	0.128	56.880	<.0001	0.586	1.705
CD5	17.366	0.155	112.240	<.0001	0.766	1.305	17.408	0.155	112.460	<.0001	0.765	1.308	15.141	0.174	87.250	<.0001	0.770	1.298
CD6	16.133	0.134	120.790	<.0001	0.671	1.489	16.194	0.134	121.100	<.0001	0.669	1.494	3.612	0.149	24.230	<.0001	0.684	1.461
CD7	58.980	0.235	251.240	<.0001	0.766	1.306	58.939	0.235	250.890	<.0001	0.764	1.308	7.807	0.247	31.590	<.0001	0.883	1.132
CD8	60.046	0.148	404.660	<.0001	0.511	1.959	60.030	0.148	404.450	<.0001	0.510	1.961	21.519	0.144	149.260	<.0001	0.661	1.513
CD9	40.229	0.189	212.370	<.0001	0.781	1.281	40.174	0.189	212.170	<.0001	0.781	1.281	10.758	0.207	52.090	<.0001	0.836	1.196
CD10	65.416	0.163	402.550	<.0001	0.525	1.906	65.328	0.163	401.990	<.0001	0.524	1.908	24.375	0.156	156.380	<.0001	0.713	1.403
CD11	-6.310	0.130	-48.400	<.0001	0.666	1.502	-6.261	0.131	-47.940	<.0001	0.663	1.508	-1.631	0.146	-11.170	<.0001	0.672	1.489
CD12	8.657	0.198	43.700	<.0001	0.851	1.175	8.636	0.198	43.600	<.0001	0.851	1.175	4.055	0.223	18.200	<.0001	0.854	1.171
CD13	62.954	0.203	309.730	<.0001	0.720	1.389	62.844	0.203	309.220	<.0001	0.720	1.390	16.879	0.213	79.230	<.0001	0.837	1.195
CD14	-14.029	0.105	-133.320	<.0001	0.510	1.960	-13.995	0.105	-132.870	<.0001	0.509	1.965	-9.074	0.118	-76.920	<.0001	0.516	1.939
Max_Beds	0.026	0.000	108.850	<.0001	0.670	1.493	0.025	0.000	103.550	<.0001	0.627	1.594	-	-	-	-	-	-
Q1	7.865	0.097	81.170	<.0001	0.551	1.814	7.965	0.097	82.160	<.0001	0.551	1.816	5.583	0.108	51.490	<.0001	0.553	1.808
Q2	5.653	0.094	60.350	<.0001	0.544	1.837	5.765	0.094	61.500	<.0001	0.543	1.841	3.834	0.105	36.510	<.0001	0.546	1.832
Q3	4.834	0.095	50.850	<.0001	0.567	1.763	4.939	0.095	51.920	<.0001	0.566	1.767	4.308	0.107	40.400	<.0001	0.568	1.761
Q4	4.088	0.096	42.680	<.0001	0.567	1.762	4.220	0.096	44.020	<.0001	0.566	1.766	3.063	0.108	28.490	<.0001	0.568	1.760
dCITYv3	53.673	0.099	541.190	<.0001	0.348	2.873	53.609	0.099	540.280	<.0001	0.348	2.876	-	-	-	-	-	-
Age	-0.154	0.004	-40.340	<.0001	0.981	1.019	-0.155	0.004	-40.610	<.0001	0.981	1.019	-0.256	0.004	-60.210	<.0001	0.990	1.010
gender2	1.449	0.058	25.010	<.0001	0.991	1.009	1.449	0.058	25.020	<.0001	0.991	1.009	2.981	0.065	45.900	<.0001	0.993	1.007
Adj R-Sq	0.254	-	-	-	-	-	0.253	-	-	-	-	-	0.069	-	-	-	-	-
F Value	38080.1	-	-	-	-	-	36348.2	-	-	-	-	-	9725.36	-	-	-	-	-
Pr > F	<.0001	-	-	-	-	-	<.0001	-	-	-	-	-	<.0001	-	-	-	-	-

Note: these models only include persons ages 65+ and their recorded trips

Figure 4-20 - Physician Billing Regression Models 2A to 2C

Linear regression models were created with health variables and census variables using forward regression. This process resulted in two models that were considered as final models, **Model 2A** and **Model 2B**.

The following variables, with corresponding dummy variables, in Table 4-11 were used to develop the **Model 2A**.

Table 4-11 - Model 2A Regression Variables

PHYSICIAN BILLING DATA	
Variable	Description
PHY_TYPE1	General Practice or Specialty Practice
CD1	Albert County
CD2	Carleton County
CD3	Charlotte County
CD4	Gloucester County
CD5	Kent County
CD6	Kings County
CD7	Madawaska County
CD8	Northumberland County
CD9	Queens County
CD10	Restigouche County
CD11	Saint John County
CD12	Sunbury County
CD13	Victoria County
CD14	Westmorland County
Dummy Variable	York County
Max Beds	Number of Beds at Destination Location
Q1	Low Income for Origin DA
Q2	Low-Medium Income for Origin DA
Q3	Medium Income for Origin DA
Q4	Medium-High Income for Origin DA
Dummy Variable	High Income for Origin DA
dCITYv3	Rural/Urban destination Indicator
Age	Age in years
Gender	Male or Female

Model 2A produced an adjusted R^2 of 0.254. Due to the variability in this type of data compared to the DAD data trips, an adjusted R^2 of 0.254 was considered a relatively good fit to the data. All independent variables produced a p-value of less than 0.05 meaning they are all considered statistically significant in the model. The VIF of the variables all fall below a value of 2.5 with the exception of CD4, Gloucester County which produced a VIF of 2.535 and dCITYv3, the rural-urban indicator which produced a VIF of 2.873. These

variables were over the ideal 2.5 threshold for VIF values but still below a VIF of 4. The values of the independent variable coefficients were reviewed to ensure they were justified based on their relationship with the dependent variable, distance traveled (km).

The physician type being general practice or specialty practice produced a positive coefficient of 21.37. This indicates an increase in 21.37 km if travelling to a doctor's office/location for a specialty trip.

Two county variables provided negative numbers which suggests significantly lower distance (km) traveled if originating from that county. This was accurate when looking at the two counties, Westmorland County (-14.03) and Saint John County (-6.31) where patients travel some of the lowest distances to physicians for health trips. County variables *CD3* for Charlotte County (80.94), *CD10* for Restigouche County, and *CD13* for Victoria County (62.95) produced large positive coefficients. This suggests that the distance traveled (km) for a health trip to see a physician increases significantly if originating from these counties.

Age produced a coefficient of -0.15 which suggests that as age increases, the distance of travel to a hospital decreases by 0.15 km per year of age.

The following variables, with corresponding dummy variables, in Table 4-12 were used to develop the **Model 1B**.

Table 4-12 – Model 2B Regression Variables

PHYSICIAN BILLING DATA	
Variable	Description
MAJOR 1	General Practice
MAJOR 2	Medical Specialty
Dummy Variable	Surgical Specialty
CD1	Albert County
CD2	Carleton County
CD3	Charlotte County
CD4	Gloucester County
CD5	Kent County
CD6	Kings County
CD7	Madawaska County
CD8	Northumberland County
CD9	Queens County
CD10	Restigouche County
CD11	Saint John County
CD12	Sunbury County
CD13	Victoria County
CD14	Westmorland County
Dummy Variable	York County
Max Beds	Number of Beds at Destination Location
Q1	Low Income for Origin DA
Q2	Low-Medium Income for Origin DA
Q3	Medium Income for Origin DA
Q4	Medium-High Income for Origin DA
Dummy Variable	High Income for Origin DA
dCITYv3	Rural/Urban destination Indicator
Age	Age in years
Gender	Male or Female

Model 2B produced an adjusted R^2 of 0.253 using the major variable groups instead of the physician type variable. All independent variables produced a p-value of less than 0.05 meaning they are all considered statistically significant in the model; however, the intercept produced a value of 0.213 which is much greater than 0.05. Like **Model 2A**, the VIF of the variables all fall below a value of 2.5 with the exception of CD4, Gloucester County which produced a VIF of 2.537 and dCITYv3, the rural-urban indicator which produced a VIF of 2.876. The VIF values of these variables were very similar to the variables in **Model 2A** which experienced a higher VIF. These variables were over the ideal 2.5 threshold for VIF values but still below a VIF of 4. The values of the independent variable coefficients were reviewed to ensure they were justified based on their relationship with the dependent variable, distance traveled (km).

All variables in **Model 2B** are identical to **Model 2A** with exception of the trip type variables. **Model 2B** includes the *major1-3* variables (*major 3* being a dummy variable) instead of the PHY_TYPE1 variable indicating general practice or specialty practice.

The *major1* variable, representing a general practice health trip, produced a negative coefficient of -20.90 and a positive coefficient of 0.94 for *major2*, which represents medical specialty. This indicates a decrease of 20.9 km if travelling to a doctor's office/ location for a general practice specialty trip and a slight increase of less than 1km if travelling to a doctor's office/ location for a medical specialty trip. An increase of 1km for a medical specialty trip does not seem high; however, in relation to the difference with the general practice trips, it appears that there is significantly more travel associated with that trip type.

The same as in **Model 2A**, the two county variables that provided negative coefficients for travel distance were Saint John County and Westmorland County. The 3 highest coefficients for travel distance also remained the same as **Model 2A** consisting of Charlotte County, Restigouche County, and Victoria County. These coefficients suggest that the distance traveled (km) for a health trip to see a physician increases significantly if originating from these counties.

The Census model, **Model 2C** included the same variables as **Model 1C**, the census model developed using the DAD data which included origin county, DA income level, age and gender. **Model 2C** resulted in an adjusted R^2 of 0.069 which is significantly lower than the health models **2A** and **2B**. Like Model 1C, all variables in **Model 2C** were considered statistically significant with a p-value less than 0.05 and all variables produced a VIF less than 2.0.

4.2.5 Final Model

Model 2A was determined to be the final model for the Physician Billing data which resulted in an R^2 value of 0.254. Although the two models were very similar, **Model 2A** produced all statistically significant variables and included the PHY_TYPE variable which was thought to be a more simplistic variable for determining trip type.

The following equation 16 represents the final regression model, **Model 2A** for estimating distance travelled for a physician visit. This model was developed using the Physician Billing dataset and includes health related variables as well as geographic and demographic variables.

MODEL 2A: adj. R2 = 0.254 **(16)**

$$\begin{aligned} Distance\ (km)_{PHB-2A} &= -21.35 + 21.37\boldsymbol{PHY_{TYPE}} + 9.36\boldsymbol{CD1} + 55.03\boldsymbol{CD2} + 80.94\boldsymbol{CD3} + 54.60\boldsymbol{CD4} \\ &+ 17.37\boldsymbol{CD5} + 16.13\boldsymbol{CD6} + 58.98\boldsymbol{CD7} + 60.05\boldsymbol{CD8} + 40.23\boldsymbol{CD9} + 65.42\boldsymbol{CD10} \\ &- 6.31\boldsymbol{CD11} + 8.66\boldsymbol{CD12} + 62.95\boldsymbol{CD13} - 14.03\boldsymbol{CD14} + 0.03\boldsymbol{maxbeds} + 7.81\boldsymbol{Q1} \\ &+ 5.65\boldsymbol{Q2} + 4.83\boldsymbol{Q3} + 4.09\boldsymbol{Q4} - 53.67\boldsymbol{dCITYv3} - 0.15\boldsymbol{Age} + 1.45\boldsymbol{geneder2} \end{aligned}$$

The following examples were used to demonstrate the final **Model 2A:**

Example 1:

Gender:	Male	**Estimated Travel Distance** **53 km**
Age:	72	
Origin County:	Albert County	
Health Trip Type:	Specialty Practice	
Income DA Level:	High	
Destination Type (R/U):	Urban	
Number of beds:	0	

Example 1 estimates that the above attributed for a health trip will result in a travel distance of 53km to an urban location. When looking at the descriptive statistics, older adults from Albert County are estimated to travel on average 55km for a specialty practice health trip for a visit to a physician. Although the R^2 value for this model is lower, it appears that the model can still make some basic estimates that are within reason.

Example 2:

Gender:	Female	**Estimated Travel Distance** **12 km**
Age:	90	
Origin County:	Westmorland County	
Health Trip Type:	General Practice	
Income DA Level:	Low	
Destination Type (R/U):	Urban	
Number of beds:	0	

Example 2 estimates that the above attributes are likely to result in a travel distance of 12km to a physician for a general practice health trip in an urban location. Descriptive statistics for general practice health trips indicate that on average, older adults from Westmorland County will travel 25km.

Similar to the DAD data, it should be noted that when choosing attributes in the model, scenarios that would not be likely to occur should be avoided. For example, a person travelling from an urban location to a rural location for a specialty health trip is unlikely.

The following equation (17) represents the regression model for estimating distance traveled for a health trip using only demographic and geographic variables which would be available from census data. This model was also developed from the Physician Billing dataset.

MODEL 2C: adj. R2 = 0.069 **(17)**

$$Distance\ (km)_{PHB-2C} = 45.52 + 15.13\boldsymbol{CD1} + 14.94\boldsymbol{CD2} + 54.23\boldsymbol{CD3} + 7.30\boldsymbol{CD4} + 15.14\boldsymbol{CD5} + 3.61\boldsymbol{CD6} + 7.81\boldsymbol{CD7} + 21.52\boldsymbol{CD8} + 10.76\boldsymbol{CD9} + 24.38\boldsymbol{CD10} - 1.63\boldsymbol{CD11} + 4.05\boldsymbol{CD12} + 16.88\boldsymbol{CD13} - 9.07\boldsymbol{CD14} + 5.58\boldsymbol{Q1} + 3.83\boldsymbol{Q2} + 4.31\boldsymbol{Q3} + 3.06\boldsymbol{Q4} - 0.26\boldsymbol{Age} + 2.98\boldsymbol{gender2}$$

4.2.6 Summary

The following conclusions were made from the linear regression analysis using the Physician Billing data:

a. The final health **model 2A** for the Physician Billing data resulted in an adjusted R^2 value of 0.254. This R^2 is lower than the R^2 predicted using the DAD data however, due to the nature of high variability within the Physician Billing dataset, a lower R^2 value was expected.

b. The census **model 2C** for the Physician Billing dataset included only variables available within the census data such as New Brunswick counties, age, gender, DA income and resulted in an adjusted R^2 value of 0.069. The R^2 for model 2C was also below the R^2 value of **model 1C** using the DAD data which again, was expected due to the nature of the Physician Billing data.

c. All values within the models resulted in a VIF of 4 or less which indicates a low probability of multicollinearity being present within the models. Most variables produced a VIF of 2.5 or less which was indicated as the ideal threshold value.

d. The age variable had a very low coefficient and a negative value meaning that has age increased, the travel distance decreased by 0.15 km per year of age. Like with the DAD data, it is expected that the age variable would have more of a contribution if other ages below 65 years were included in the model.

e. Like the models developed with the DAD data, the coefficients for variables originating from the urban counties appeared to decrease travel significantly while the more rural counties increased travel significantly. On average, the more rural the county, the higher the coefficient for that county indicating further travel to access health needs.

f. Like the DAD data, the type of health trip resulted in significantly more travel for specialty trips as opposed to general practice trips.

g. The destination rural/urban indicator had a significant influence on the R^2 for the model and resulted in a large positive coefficient.

4.3 Trip Generation: Cross Classification Tables

Cross classification tables are one of the methods that transportation planners use for trip generation. To generate health specific trips for older adults, classification tables were created using the DAD and Physician Billing datasets and population data to create rates for health trips generated in New Brunswick. Health trips rates created using the DAD data are trip rates for adults 65+ travelling to the **hospital** for a health trip. All other health trips are captured within the Physician Billing Data which are rates for adults 65+ travelling to **a physician's location** which includes offices, clinics, etc. This can sometimes bring trips to a hospital if the physician is providing the service from a clinic or office within a hospital.

Rates are classified based on age groups, gender and origin county. Health trip tables and New Brunswick population tables from 2006 and 2011 obtained from Statistics Canada (Catalogue no.97-551-XCB2006006 and no. 98-311-XCB2011018) were used to create classification tables.

4.3.1 Trip Rates for Health Trips Based on the DAD Data

Trip rates were developed for adults aged 65+ using the DAD dataset and the census population. Rates for the year of 2011 health data was created as well as an average rate between 2005-2011 for age and gender, specialty practice, and for each county.

4.3.1.1 New Brunswick Rate

Table 4-13 provides New Brunswick health trip rates. This table was generated using 2011 health trips and census populations for the year of 2011. These rates can be applied to a gender and age group population to determine an approximate number of health trips for

New Brunswick. This shows that for 2011 data, men typically had a higher trip rate than women in all age groups. These data also show that after 80 years of age, the rate of health trips for the province are expected to increase for both women and men.

Table 4-13 – Rate of Health Trips per person per year for New Brunswick Taken in 2011

DAD DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	0.14	0.17	0.19	0.21	0.20	0.22	0.18
M	0.17	0.21	0.26	0.29	0.28	0.30	0.22
Total	0.16	0.19	0.22	0.24	0.23	0.24	0.20

**Health trips from DAD data in 2011 fiscal year*

Table 4-14 below, displays the average trip rate for a health trip per person per year to a hospital using data from 2005-2011 fiscal years. An average rate was completed to capture all trips from the health dataset. When comparing the 2011 trip rates from Table 4-13 and the average trip rates in Table 4-14, both tables produce a similar trip rate however, the average rates are slightly higher for some categories. Some categories have more of an increased average rate such as age groups 85-89 and 90+ however these increased rates are still considered a minimal increase.

Table 4-14 – Average Rate of Health Trips per person per year for New Brunswick Between 2005-2011

DAD DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	0.14	0.17	0.21	0.23	0.25	0.25	0.19
M	0.18	0.22	0.27	0.30	0.33	0.33	0.24
Total	0.16	0.20	0.24	0.26	0.27	0.27	0.21

**Health trips from DAD data in 2011 fiscal year*

When looking at the average trip rates for 2005-2011 data, men still typically have a higher trip rate than women in all age groups. These data also show that after 85 years of age, the rate of health trips for the province are expected to increase for both women and men and remain the same for 90+ years of age.

4.3.1.2 Health Data Rate

Trip rates were determined for individuals who were a part of the DAD dataset to better understand the frequency of health trips for those who took health trips in 2005-2011 fiscal years. These were created using the total health trips and the total individuals in the datasets.

Table 4-15 provides rates for frequency of individuals that took health trips in 2011. This rate ranges between 1.24 to 1.45 trips per person in the 2011 dataset. These data show that in all age groups, men took more health trips than women. On average, women who took health trips in 2011, took the most health trips between the ages of 70-84. On average, men who took health trips in 2011 took the most health trips for the age group of 75-79. The age group 75-79 shows the highest rate of health trips for both women and men.

Table 4-15 - Rate of Health Trips per Individual That Took a Health Trip in 2011

DAD DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	1.28	1.30	1.30	1.30	1.24	1.24	1.28
M	1.32	1.37	1.45	1.44	1.38	1.35	1.38
Total	1.30	1.34	1.38	1.36	1.30	1.28	1.33

**Health trips from DAD data in 2011 fiscal year*

Table 4-16 provides rates for frequency of individuals that took health trips between 2005-2011. This rate ranges between 1.31 to 1.48 trips per person per year in the 2005-2011 dataset. When looking at Table 4-15 for 2011 health trips and Table 4-16 for the average of 2005-2011 health trips, although not drastically, the rates increase between all age categories by approximately 0.02 to 0.11 for women and 0.02 to 0.08 for men.

Table 4-16 - Rate of Health Trips per Individual That Took a Health Trip Between 2005-2011

DAD DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	1.31	1.34	1.37	1.38	1.35	1.32	1.35
M	1.37	1.43	1.47	1.48	1.45	1.40	1.43
Total	1.34	1.39	1.42	1.42	1.39	1.35	1.39

**Health trips from DAD data in 2011 fiscal year*

Again, these data show that in all age groups, men took more health trips than women. On average, the age group 80-84 shows the highest frequency rate of health trips for both women and men between 2005-2011.

4.3.1.3 Trip Rates by Age and Type of Health Trip

Table 4-17 and Table 4-18 shown below, provide rates for average health trips for general practice and specialty practice health trips using the 2005-2011 dataset.

Table 4-17 - Average New Brunswick Trip Rate for General Practice per person per year Between 2005-2011

DAD DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	0.04	0.05	0.08	0.11	0.16	0.20	0.08
M	0.04	0.06	0.10	0.14	0.18	0.24	0.08
Total	0.04	0.06	0.09	0.12	0.17	0.21	0.08

**Health trips from DAD data in 2005- 2011 fiscal year*

Table 4-17 shows the average rate for general practice health trips increases in both women and men as an individual's age increases. Age group 65-69 begins at a provincial rate of 0.04 for both women and men and gradually increases to a rate of 0.20 for women and 0.24 for men by the age of 90+.

Table 4-18 - Average New Brunswick Trip Rate for Specialty Practice per person per year Between 2005-2011

DAD DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	0.10	0.12	0.13	0.12	0.09	0.05	0.11
M	0.13	0.16	0.18	0.17	0.14	0.09	0.15
Total	0.12	0.14	0.15	0.14	0.11	0.07	0.13

**Health trips from DAD data in 2005-2011 fiscal year*

Table 4-18 provides a provincial trip rate for specialty practice health trips. The rates for specialty practice trend in the opposing direction where the rates for ages 65-69 are greater than the rates of ages 90+. Trip rates for women taking specialty practice health trips decrease from a rate at 0.10 trips person at ages 65-69 to a rate of 0.05 trips per person at ages 90+. Trip rates for men decrease from 0.13 for ages 65-69 to a rate of 0.09 for ages 90+. Trip rate peak for both women and men at ages 75-79.

Overall, specialty trip rates are greater than general practice trip rates for the DAD dataset which is capturing trips travelling to the hospital.

4.3.1.4 Trip Rates by County Level

Trip type was then broken into rates per person for each county by age group for an average trip rate between 2005-2011 as shown in Table 4-19.

Looking at the overall trips per county Restigouche has the greatest trip rate per person with a rate of 0.42 trips/person/year, followed by Queens County with a trip rate of 0.40 trips/person/year. The third highest trip rate is from Charlotte County with a trip rate of 0.24 trips/person/year which is almost half of the rate produced from Restigouche County. The county with the lowest overall trip rate is Madawaska County with a trip rate of 0.15 trips/person/year.

Table 4-19 – Average Trip Rate for Health Trips by County per person per year Between 2005-2011

DAD DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
Albert	0.16	0.19	0.27	0.27	0.26	0.25	0.21
Carleton	0.15	0.21	0.26	0.27	0.25	0.22	0.22
Charlotte	0.18	0.20	0.27	0.33	0.32	0.32	0.24
Gloucester	0.13	0.18	0.22	0.27	0.27	0.30	0.19
Kent	0.16	0.18	0.21	0.25	0.23	0.22	0.19
Kings	0.13	0.16	0.20	0.25	0.25	0.20	0.18
Madawaska	0.12	0.14	0.17	0.19	0.16	0.22	0.15
Northumberland	0.18	0.21	0.24	0.31	0.33	0.30	0.23
Queens	0.32	0.38	0.50	0.46	0.44	0.34	0.40
Restigouche	0.33	0.40	0.46	0.50	0.47	0.42	0.42
Saint John	0.12	0.17	0.19	0.21	0.22	0.20	0.17
Sunbury	0.16	0.19	0.26	0.27	0.30	0.28	0.21
Victoria	0.16	0.24	0.25	0.26	0.29	0.30	0.23
Westmorland	0.14	0.18	0.22	0.25	0.24	0.22	0.19
York	0.13	0.18	0.23	0.27	0.24	0.28	0.19
Total	0.15	0.19	0.23	0.27	0.26	0.25	0.20

Persons aged 80-89 have the highest trip rate overall; however, there is variability in this rate among counties. Restigouche County has a trip rate of 0.50 trip per person per year for ages 80-84 with the lowest rate being 0.33 for ages 65-69. Queens County has the greatest trip rate for ages 75-79 at 0.50 and the lowest rate of 0.32 at ages 65-69. Lastly, Charlotte County experiences the greatest trip rate for age 80-84 with a rate of 0.33 and the smallest rate being 0.18 for ages 65-69. Madawaska County has the smallest trip rate at 0.12 for ages 65-69.

4.3.1.5 Trip Rate by Trip Type and County Level

Table 4-20 and Table 4-21 provide trips rates by county for general practice trips and specialty practice trips respectively.

Table 4-20 – Average New Brunswick Trip Rate for General Practice by County per person per year Between 2005-2011

DAD DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
Albert	0.04	0.05	0.09	0.11	0.15	0.19	0.08
Carleton	0.05	0.08	0.12	0.16	0.16	0.17	0.10
Charlotte	0.06	0.07	0.12	0.17	0.21	0.25	0.11
Gloucester	0.04	0.06	0.09	0.13	0.18	0.25	0.08
Kent	0.02	0.03	0.05	0.08	0.11	0.15	0.05
Kings	0.02	0.03	0.05	0.08	0.12	0.14	0.05
Madawaska	0.03	0.05	0.07	0.10	0.11	0.17	0.06
Northumberland	0.06	0.08	0.11	0.17	0.21	0.23	0.11
Queens	0.08	0.12	0.20	0.26	0.27	0.27	0.17
Restigouche	0.13	0.17	0.24	0.29	0.32	0.35	0.22
Saint John	0.02	0.04	0.06	0.09	0.13	0.15	0.06
Sunbury	0.04	0.06	0.11	0.13	0.17	0.23	0.08
Victoria	0.05	0.09	0.11	0.14	0.19	0.24	0.10
Westmorland	0.03	0.05	0.07	0.10	0.14	0.16	0.07
York	0.04	0.06	0.09	0.13	0.15	0.22	0.08
Total	0.04	0.06	0.09	0.13	0.16	0.19	0.08

Table 4-20 provides an average trip rate per person for general practice trips for each county. Looking at the overall trip rate for general practice trips, Restigouche County has the greatest trip rate per person per year with 0.22 followed by Queens County at 0.17. Northumberland County and Charlotte County both have a trip rate of 0.11 per person while Carleton County and Victoria County have a trip rate of 0.10 per person. Trips for general practice appear to increase as an individual ages with most peak rates within the 90+ age group.

Table 4-21 – Average New Brunswick Trip Rate for Specialty Practice by County per person per year Between 2005-2011

DAD DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
Albert	0.12	0.14	0.18	0.16	0.11	0.06	0.14
Carleton	0.10	0.13	0.14	0.12	0.09	0.05	0.11
Charlotte	0.13	0.13	0.15	0.16	0.12	0.07	0.13
Gloucester	0.09	0.12	0.14	0.13	0.10	0.05	0.11
Kent	0.14	0.15	0.16	0.17	0.12	0.07	0.14
Kings	0.11	0.13	0.16	0.17	0.13	0.06	0.13
Madawaska	0.08	0.09	0.10	0.09	0.05	0.05	0.09
Northumberland	0.12	0.13	0.14	0.14	0.11	0.07	0.13
Queens	0.24	0.26	0.30	0.20	0.18	0.07	0.23
Restigouche	0.20	0.23	0.23	0.21	0.14	0.07	0.20
Saint John	0.10	0.13	0.13	0.12	0.09	0.05	0.11
Sunbury	0.11	0.13	0.15	0.14	0.13	0.05	0.13
Victoria	0.11	0.15	0.14	0.12	0.10	0.06	0.12
Westmorland	0.11	0.13	0.15	0.15	0.10	0.07	0.12
York	0.09	0.12	0.14	0.13	0.09	0.06	0.11
Total	0.11	0.13	0.15	0.14	0.10	0.06	0.12

Table 4-21 provides an average trip rate per person for specialty practice trips for each county. Looking at the overall trip rate for general practice trips, Queens County has the greatest trip rate per person per year at 0.23 followed by Restigouche County at 0.20. Albert County and Kent County both have a trip rate of 0.14 trips per person while Charlotte County, Northumberland and Sunbury County have a trip rate of 0.13 per person for a specialist visit. Trips for specialty practice appear to decrease as an individual ages however, peak rates tend to be within age groups 75-79 and 80-84.

4.3.1.6 Summary

The following conclusions were made from creating trips rates for hospital trips using the DAD dataset:

a. Overall, men were found to have a higher trip rate per person than women for all age groups. As age increased, the provincial trip rates for men exceeded trip rates for women. Population data shows that as age increases, the population of women are almost double the population of men, meaning that women are living longer.

b. The age groups with the highest trips rates per person were ages 80-84 and 85-89.

c. Trip rates for health trips to a hospital in New Brunswick for those aged 65+ ranged between 0.14 to 0.33 per year for the fiscal years of 2005-2011.

d. Older adults who took a health trip between 2005-2011 fiscal years, averaged between a rate of 1.31 to 1.48 health trips per person per year .

e. Overall, the DAD data showed smaller trip rates per person for general practice trips compared to specialty practice trips.

f. Trip rates per person for general practice trips increased as age increased. The peak rate for general practice trips was for the 90+ age group with a rate of 0.20 for women and 0.24 for men.

g. Trip rates per person for specialty practice trips decreased overall as age increased; however, the peak trip rate for specialty practice trips was for ages 75-79 with rates of 0.13 for women and 0.18 for men.

h. The greatest county level trip rates per person were generated from Restigouche and Queens County with trip rates of 0.42 and 0.40 trips per person per year.

i. Charlotte County was the third highest trip rate for health trips per person to a hospital with a rate of 0.24 which is significantly lower than rates from Restigouche and Queens County.

4.3.2 Trip Rates for Health Trips Based on the Physician Billing Data

Physician Billing data was used to generate trip rates for older adults making a health trip to a physician's location such as an office or a clinic. It should be noted that with Physician Billing dataset, some trips were not captured when running frequency trips for age groups, genders, counties and trip types. This was considered to be part of the nature of using administrative data to estimate trip rates for physician trips, it should be noted that not all data was available for rates. A total of 263,606 trips were missing for health trips between 2006-2011; therefore, a total of 2,618,810 trips were used to develop these rates. The number of trips missing from the data used to generate the following trip rates are noted for each trip rate table.

Table 4-22 provides New Brunswick health trip rates for health trips to see a physician practicing at an office, clinic, etc. This table was generated using 2011 health trips from the Physician Billing dataset and census populations for the year of 2011. These rates can be applied to a gender and age group population to determine an approximate number of health trips for New Brunswick. Table 4-22 shows that for 2011 data, men typically have a higher trip rate per person than women in all age groups. These data also show that as

age increases, the rate of health trips increases excluding the age groups of 80-84 with a rate of 3.93 and age group 85-89 with a rate of 3.94 per person for both women and men.

Table 4-22 - Rate for New Brunswick Health Trips Taken in 2011

PHYSICIAN BILLING DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	3.13	3.28	3.22	3.24	3.36	3.35	3.23
M	3.24	3.87	4.60	4.96	5.02	5.40	4.03
Total	3.19	3.56	3.84	3.93	3.94	3.88	3.59

Frequency Missing = 45,209

Table 4-23 below, displays the average trip rate for a health trip to the location of a physician for 2006-2011 calendar years. An average rate was completed to capture all trips from the health dataset. When comparing the 2011 trip rates from Table 4-22 and the average trip rates in Table 4-23, both tables produce a similar trend with rates increasing with age with a peak rate for age group 85-89 at 4.29 trips per person per year for both women and men.

Table 4-23 - Rate for the Average New Brunswick Between 2006-2011

PHYSICIAN BILLING DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	3.25	3.35	3.50	3.57	3.74	3.53	3.43
M	3.42	4.06	4.72	5.04	5.38	5.25	4.21
Total	3.34	3.69	4.03	4.15	4.29	3.98	3.77

Frequency Missing = 263,606

When looking at the average trip rates for 2005-2011 data, men have a significantly higher trip rate than women in all age groups. These data also show that trip rates for physician visits peak between 85-89 years of age..

4.3.2.1 Health Data Rate

Trip rates were determined for individuals who were a part of the Physician Billing dataset to better understand the frequency of health trips for those who took health trips in 2005-

2011 fiscal years. These were created using the total health trips and the total individuals in the datasets.

Table 4-24 provides rates for frequency of individuals that took health trips in 2011. This rate ranges between 7.65 to 11.63 trips per person in the 2011 dataset. These data shows that in all age groups, men took more health trips then women except for age group 65-69 where women had a rate of 7.65 trips per person and men produced a rate of 7.51 trips per person. On average, women who took health trips in 2011 took the most health trips between the ages of 85-89 with a rate of 10.18 trips per person. On average, men who took health trips in 2011 also took the most health trips for the age group of 85-89 with a rate of 11.63 trips per person.

Table 4-24 - Rate of Health Trips per Individual That Took a Health Trip in 2011

PHYSICIAN BILLING DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	7.65	8.45	9.26	9.94	10.18	9.10	8.72
M	7.51	8.65	10.19	10.74	11.63	10.16	9.04
Total	7.58	8.55	9.74	10.33	10.78	9.46	8.88

Health trips from PHB data in 2006-2011 calendar year
Frequency missing = 45,209

Table 4-25 provides rates for frequency of individuals that took health trips between the calendar years of 2006- 2011. This rate ranges between 7.36 to 10.28 trips per person per year in the 2006-2011 dataset. When looking at Table 4-24 for 2011 health trips and Table 4-25 for the average of 2005-2011 health trips, rates for 2011 health trips were greater than the average rate.

Again, these data shows that in all age groups, men took more health trips than women with exception of 65-69 where women have a trip rate of 7.37 trips per person that took a health trip per year and men have a rate of 7.36 per person that took a health trip per year.

On average, the age group 80-84 shows the highest rate of health trips for both women and men between 2005-2011.

Table 4-25 - Rate of Health Trips per Individual That Took a Health Trip Between 2006-2011

PHYSICIAN BILLING DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	7.37	7.97	8.49	8.99	9.09	8.00	8.15
M	7.36	8.53	9.49	10.28	10.49	9.18	8.71
Total	7.37	8.25	8.97	9.56	9.63	8.37	8.42

Health trips from PHB data in 2006-2011 calendar year
Frequency missing = 263,606

4.3.2.2 Trip Rates by Age and Type of Health Trip

Table 4-26 and Table 4-27 shown below, provide rates for average health trips for general practice and specialty practice health trips using the 2006-2011 Physician Billing dataset.

Table 4-26 shows the average rate for general practice health trips increases for total age groups as an individual's age increases; however, within the gender categories, women's trip rate decreases and men's trip rate increases. Age group 65-69 begins at a provincial rate of 2.11 trips/person/year for women and 2.18 trips/person/year for men. By age group 90+ trip rates decrease for women to a rate of 1.38 trips/person/year while trip rates for men increase drastically to 7.70 trips/person. The average trip rate for general practice health trips for both women and men increases from 2.15 trips/person/year for ages 65-69 to 3.02 trips/person/year for ages 90+.

Table 4-26 - Average New Brunswick Trip Rate for General Practice Between 2006-2011

PHYSICIAN BILLING DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	2.11	2.28	2.31	2.10	1.81	1.38	2.11
M	2.18	2.46	2.96	3.65	5.22	7.70	2.93
Total	2.15	2.37	2.60	2.71	2.96	3.02	2.47

Health trips from PHB data in 2006-2011 calendar year

Frequency Missing = 167,943

Table 4-27 provides a provincial trip rate for specialty practice health trips. The rates for specialty practice are for both women and men are the lowest for ages 90+ with a rate of 0.95 trips/person/year and the greatest for ages 75-79 and 80-84 with a rate of 1.44 trips/person/year. Trips rates for women tend to be much lower in the age group 90+ with a rate of 0.47 trips/person/year compared to men with a rate of 2.23 trips/person/year for specialty trips. Women have the greatest rate for specialty trips at ages 70-74 and 75-79 where the rate is 1.36 trips/person/year. The greatest rate for men taking specialty trips is 2.23 trips/person/year for ages 90+.

Table 4-27 - Average New Brunswick Trip Rate for Specialty Practice by County Between 2006-2011

PHYSICIAN BILLING DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
F	1.23	1.36	1.36	1.18	0.93	0.47	1.20
M	1.15	1.28	1.54	1.85	2.13	2.32	1.44
Total	1.19	1.32	1.44	1.44	1.33	0.95	1.30

Health trips from PHB data in 2006-2011 calendar year
Frequency Missing = 95,663

Overall, specialty trip rates are less than general practice trip rates for the Physician Billing dataset which is capturing trips travelling to the location of a physician.

4.3.2.3 Trip Rates by County Level

Trip type was then broken into rates per person for each county by age group for an average trip rate between 2006-2011 as shown in Table 4-28.

Table 4-28 - Average Trip Rate for Total Health Trips by County

PHYSICIAN BILLING DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
Albert	2.98	3.25	3.94	3.73	3.70	2.82	3.35
Carleton	3.78	4.24	5.13	5.16	4.11	4.21	4.38

Charlotte	3.09	3.19	3.66	4.12	3.86	4.29	3.49
Gloucester	3.19	3.58	3.97	4.37	4.40	4.66	3.72
Kent	3.06	3.51	3.65	3.96	3.62	3.33	3.44
Kings	2.59	3.05	3.85	4.40	4.32	3.45	3.30
Madawaska	1.27	1.56	1.53	1.72	1.35	1.62	1.47
Northumberland	3.35	3.66	4.13	4.97	5.08	5.01	3.99
Queens	8.34	9.43	11.03	9.67	8.22	5.76	9.04
Restigouche	7.22	7.98	8.28	7.44	7.43	6.11	7.58
Saint John	2.33	2.91	2.98	2.95	3.14	2.29	2.74
Sunbury	3.41	3.68	4.29	4.26	4.63	4.10	3.85
Victoria	2.95	3.95	3.93	3.78	4.21	3.12	3.60
Westmorland	2.94	3.31	3.66	4.17	4.16	3.55	3.47
York	3.42	4.04	4.57	4.72	4.17	3.73	4.02
Total	3.12	3.58	4.00	4.25	4.11	3.70	3.65

Health trips from PHB data in 2005-2011 calendar year
Frequency Missing = 263,606

Looking at the overall trips per county Queens County has the greatest trip rate per person with a rate of 9.04 trips/person/year, followed by Restigouche County with a trip rate of 7.58 trips/person/year. The third highest trip rate is from Carleton County with a trip rate of 4.38 trips/person/year which is almost half of the rate produced from Queens County. The county with the lowest overall trip rate is Madawaska County with a trip rate of 1.47 trips/person/year.

Ages 80-84 have the highest trip rate overall; however, trip rates vary by age and county. Queens County has a trip rate of 11.03 trips/person/year for ages 75-79 with the lowest rate being 5.76 trips/person/year for ages 90+. Restigouche County experiences the greatest trip rate at ages 75-79 with a rate of 8.28 trips/person/year and the lowest rate of 6.11 trips/person/year for the 90+ age group. Lastly, Carleton County experiences the greatest trip rate for age 80-84 with a rate of 5.16 trips/person/year and the smallest rate being 3.78 trips/person/year for ages 65-69. Madawaska County has the smallest trip rate at 1.27 for ages 65-69.

4.3.2.4 Trip Rate by Trip Type and County Level

Table 4-29 and Table 4-30 provide trips rates by county for general practice trips and specialty practice trips respectively.

Table 4-29 - Average New Brunswick Trip Rate for General Practice by County Between 2006-2011

PHYSICIAN BILLING DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
Albert	1.82	1.87	2.22	2.02	2.22	1.93	1.96
Carleton	2.36	2.71	3.32	3.40	2.78	3.01	2.84
Charlotte	2.29	2.37	2.82	3.21	3.13	3.76	2.68
Gloucester	2.18	2.42	2.80	2.98	3.21	3.99	2.60
Kent	1.70	1.92	1.98	2.35	2.34	2.40	1.96
Kings	1.89	2.25	2.86	3.26	3.36	2.65	2.45
Madawaska	0.90	1.17	1.12	1.38	1.13	1.34	1.10
Northumberland	2.09	2.37	2.69	3.42	3.71	3.95	2.66
Queens	5.73	6.72	7.95	7.20	6.18	4.37	6.48
Restigouche	3.83	4.20	4.35	4.09	4.43	4.03	4.12
Saint John	1.54	1.91	1.88	1.83	2.15	1.76	1.80
Sunbury	2.11	2.24	2.63	2.72	3.02	3.23	2.41
Victoria	2.19	2.86	2.91	2.80	3.21	2.37	2.66
Westmorland	1.76	1.98	2.18	2.47	2.64	2.61	2.12
York	2.21	2.57	2.90	3.09	2.80	2.60	2.60
Total	2.01	2.30	2.58	2.77	2.83	2.81	2.39

Health trips from PHB data in 2005-2011 calendar year
Frequency Missing = 167,943

Table 4-29 provides an average trip rate per person/year for general practice trips for each county. Looking at the overall trip rate for general practice trips, Queens County has the greatest trip rate with 6.48 trips/person/year followed by Restigouche County at 4.12 trips/person/year. Carleton County has a trip rate of 2.84 trips/person/year while Charlotte County has a rate of 2.68 trips/person/year. Trips for general practice appear to increase as an individual ages until ages 85-89 where the trip rate peaks for general practice trips at 2.83 trips/person/year.

Table 4-30 provides an average trip rate per person/year for specialty practice trips for each county. Looking at the overall trip rate for specialty practice trips, Restigouche County has the greatest trip rate with 3.46 trips/person/year followed by Queens County at 2.55

trips/person/year. Carleton County has a rate of 1.54 trips/person/year and Kent County has a trip rate of 1.48 trips/person/year. Trips for specialty practice appear to have the greatest trip rates produced from age groups 80-84 with 1.48 trips/person/year and ages 75-79 with 1.43 trips/person/year.

Table 4-30 - Average New Brunswick Trip Rate for Specialty Practice by County Between 2006-2011

PHYSICIAN BILLING DATA							
	65-69	**70-74**	**75-79**	**80-84**	**85-89**	**90+**	**Total**
Albert	1.15	1.37	1.72	1.71	1.48	0.89	1.39
Carleton	1.41	1.53	1.81	1.76	1.33	1.20	1.54
Charlotte	0.80	0.82	0.83	0.91	0.73	0.53	0.80
Gloucester	1.01	1.16	1.17	1.39	1.19	0.68	1.12
Kent	1.36	1.60	1.67	1.61	1.28	0.94	1.48
Kings	0.70	0.80	0.99	1.15	0.96	0.79	0.85
Madawaska	0.37	0.39	0.41	0.34	0.22	0.28	0.36
Northumberland	1.25	1.29	1.44	1.55	1.37	1.07	1.34
Queens	2.60	2.71	3.08	2.47	2.05	1.39	2.55
Restigouche	3.39	3.78	3.94	3.36	3.00	2.08	3.46
Saint John	0.79	1.00	1.10	1.11	0.99	0.52	0.94
Sunbury	1.30	1.43	1.65	1.53	1.61	0.87	1.44
Victoria	0.76	1.09	1.02	0.98	1.00	0.75	0.94
Westmorland	1.18	1.32	1.48	1.70	1.52	0.94	1.35
York	1.21	1.47	1.67	1.63	1.36	1.13	1.42
Total	1.11	1.28	1.43	1.48	1.28	0.89	1.26

Health trips from PHB data in 2005-2011 calendar year
Frequency Missing = 95,663

4.3.2.5 Summary

The following conclusions were made from creating trips rates for physician visits using the Physician Billing dataset:

a. Overall, men were found to have a higher trip rate of trip/person/year than women for all age groups except 65-69 for which men were sometimes slightly lower than women. As age increased, the provincial trip rates for men exceeded trip rates for women.

b. The most common age group that generated the greatest trips rates per person were ages 84-85.

c. Trip rates for physician visit health trips in New Brunswick for those aged 65+ ranged between 3.25 to 5.38 per year for the calendar years of 2006-2011.

d. Older adults who took a health trip between 2006-2011 calendar years averaged between a rate of 7.36 to 10.49 health trips/person who took a health trip/year.

e. Overall, the Physician Billing data produced larger trip rates per person for general practice trips compared to specialty practice trips.

f. Trip rates for general practice trips increased as age increased for men and decreased as age increased for women. The peak rate for men's general practice trips was for the 90+ age group with a rate of 7.70 trips/person/year and 2.31 trips/person/year for women in age groups 75-79. This is a significant difference in trip rates; therefore, it should be noted that the difference in trip rates could be due to missing data when running frequencies on the Physician Billing dataset.

g. Trip rates per person for specialty practice trips decreased overall as age increased; however, the peak trip rate for specialty practice trips peaked at 1.36 trips/person/year for women ages 75-79 and 2.32 trips/person/year for men ages 90+.

h. The greatest county level trip rates were generated from Queens County and Restigouche County with trip rates of 9.04 trips/person/year and 7.58 trips/person/year.

i. Carleton County was the third highest trip rate for health with a rate of 4.38 trips/person/year which is significantly lower than rates from Queens County and Restigouche County.

j. When running frequency tables for trip types and age groups, a total of 263,606 observations were missing for 2006-2011 Physician Billing dataset. This left 2,618,810 total trips used to create trip rate tables. Due to the nature of using administrative data it is recognized that this could affect trips rates; however, rates were determined as a starting point for New Brunswick health trips.

4.4 Trip Distribution: Gravity Model

The origin and destination trips for 2011 were applied to the traditional form of the gravity model to demonstrate a base application of the distribution of health trips. This was completed using both the DAD data and the Physician Billing data and comparing the actual trip distributions with the predicted distributions from the gravity model. Iterations for the gravity model can be found within Appendix C.

The following cost matrix shown in Table 4-33, was used in all applications of the gravity model for each dataset. When looking at the cost matrix, it is evident that with New Brunswick's low population density, that moving from county-to-county increases distance (cost) significantly.

4.4.1 Traditional Gravity Model with the DAD Data

The traditional gravity model was applied to the following origin-destination matrix shown in Table 4-34, to determine trip distributions using the 2011 health trips from the DAD dataset. Although Albert County did not attract health trips as a destination, a value of 1 was used for an internal trip for the analysis to avoid receiving an error during calculations. With an attraction of 1 trip to Albert County, the traditional gravity model distributed the rest of the health trips produced from Albert County to the next closest County, Westmorland County.

The origin-destination matrix shown in Table 4-34, indicates that Westmorland and Saint John County attract the highest total of health trips to hospitals with totals of 6,785 and 5,155 respectively. These two counties include two of the three urban cities and are home

to major hospitals in New Brunswick therefore it is expected that they attract a high number of health trips.

Figure 4-21 – County Map of New Brunswick (GNB 2021)

Figure 4-21 illustrates how Westmorland County and Saint John County are both located on the South-East end of the province. Although Westmorland and Saint John County produce a number of internal health trips, the location of these major hospitals cause those

travelling from other counties, specifically the North-West end of the province, to travel long distances to seek health services from Westmorland or Saint John County.

The Origin-destination Matrix also indicates that Gloucester and York County attract the next highest number of health trips totaling to 2,690 and 2,330 respectively. With Gloucester located in Northern New Brunswick and York County located in Central/ Southern New Brunswick.

The gravity model distributes trips based on distance with the idea that a greater cost, or distance, to another zone will deter the amount of travel to that zone. When looking at Table 4-35, it is evident that the traditional gravity model assumes high interzonal trips due to cost of travel to other counties and then distributes trips to the next closest county. However, with hospital health trips, travel is dependent on the type of trip which causes people to travel longer distances than expected to seek a specific health service rather than travelling to the next closest county.

Table 4-36 shows the difference in actual and predicted distribution of health trips using the gravity model. Some highlighted areas where there are too many trips estimated for the interzonal trips are in the following counties:

Table 4-31 - Variation Between Internal Trips

County	Difference in Trips
Carleton	+419
Charlotte	+511
Gloucester	+515
Kent	+537
Northumberland	+655
Restigouche	+500
Victoria	+412
York	-567

As shown in Table 4-31, the gravity model has overestimated the number of internal trips being made for health trips to a hospital and underestimated the number of internal trips for York County.

There are also various trips between zones that have a significant variation from actual and predicted trips. Variation between county-to-county trips was present in various situations however the instances with significant difference in trips are shown below in Table 4-32.

Table 4-32 - Variation of County-to-County Trips

Origin County	Destination County	Difference in Trips
Charlotte	Saint John	-395
Gloucester	Westmorland	-245
Kent	Westmorland	-452
Northumberland	Westmorland	-215
York	Saint John	-228
York	Sunbury	+975

As shown in Table 4-32, hospital trips to Saint John County and Westmorland County were underestimated for a number of counties and overestimated with York County travelling to Sunbury County.

The counties which underestimated trips to Saint John County and Westmorland County are located further away which makes travel to these locations less attractive. However, because Westmorland and Saint John County are home to major New Brunswick hospitals, health trips are taken here regardless of distance traveled.

York County and Sunbury County are within close proximity with one another however, York County consists of a larger radius as shown in Figure 4-21. The Oromocto Public Hospital, in Sunbury County, is within 24km of the Dr. Everett Chalmers Hospital in York County. Due to Fredericton, New Brunswick being one of the three major urban cities with

a higher population density for York County, the distance between York County and Sunbury County was made much lower in comparison to other counties which resulted in the majority of York County health trips being attracted to Sunbury County. Although Sunbury does attract a number of health trips from York County, the small distance from the cost matrix inflated the number of distributed trips.

Table 4-33 - Distance (km) - Cost Matrix

	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York
Albert	30	340	330	320	130	115	480	220	150	400	160	190	350	70	215
Carleton	340	30	150	300	250	250	160	200	200	250	240	150	100	300	130
Charlotte	330	150	40	380	250	180	300	280	190	410	110	140	240	300	120
Gloucester	320	300	380	50	200	310	330	150	300	200	380	300	300	260	280
Kent	130	250	250	200	30	130	330	100	100	280	200	330	220	80	140
Kings	115	250	180	310	130	30	400	200	50	400	50	100	330	80	120
Madawaska	480	160	300	330	330	400	30	240	350	220	380	300	100	450	280
Northumberland	220	200	280	150	100	200	240	40	160	240	260	180	140	180	150
Queens	150	200	190	300	100	50	350	160	40	380	90	55	260	130	80
Restigouche	400	250	410	200	280	400	220	240	380	40	470	400	190	350	380
Saint John	160	240	110	380	200	50	380	260	90	470	30	90	300	180	115
Sunbury	190	150	140	300	330	100	300	180	55	400	90	30	230	180	30
Victoria	350	100	240	300	220	330	100	140	260	190	300	230	30	340	215
Westmorland	70	300	300	260	80	80	450	180	130	350	180	180	340	50	200
York	215	130	120	280	140	120	280	150	80	380	115	30	215	200	40

Kilometers traveled estimated from Google Maps

All internal county trips assumed to be within 30-40km with exceptions of Gloucester and Westmorland County having a 50km radius

Table 4-34 – DAD Data - Origin-Destination Matrix for 2011 Health Trips

ORIGIN	DESINATION															
	Albert[1]	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens[2]	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York	Total
Albert	0	5	0	10	5	15	0	0	5	5	45	5	10	940	0	1,046
Carleton	0	640	0	10	0	5	0	5	5	0	95	80	70	15	135	1,060
Charlotte	0	5	530	5	0	5	0	5	0	5	530	80	10	5	15	1,195
Gloucester	0	15	5	2,270	0	5	0	65	5	15	140	5	10	245	5	2,785
Kent	0	10	5	5	100	5	0	30	0	0	40	0	0	1,010	5	1,210
Kings	0	0	0	10	5	285	0	5	5	5	1,360	10	5	65	10	1,765
Madawaska	0	10	5	0	0	0	880	0	0	10	5	5	45	45	0	1,005
Northumberland	0	0	5	135	5	0	0	1,570	5	5	130	50	10	215	95	2,225
Queens	0	5	5	5	0	35	0	5	5	5	105	215	10	15	220	630
Restigouche	0	5	5	165	5	0	95	10	5	1,050	65	0	5	135	5	1,550
Saint John	0	5	5	15	5	5	5	5	5	10	2,165	0	10	5	0	2,240
Sunbury	0	5	5	20	0	5	5	5	5	5	45	325	15	15	225	680
Victoria	0	75	5	20	5	5	165	0	0	0	35	40	355	15	50	770
Westmorland	0	10	10	15	0	15	0	0	0	10	165	5	10	4,035	0	4,275
York	0	110	10	5	5	0	5	0	5	10	230	870	5	25	1,565	2,845
Total	0	900	595	2,690	135	385	1,155	1,705	50	1,135	5,155	1,690	570	6,785	2,330	25,281

[1]No Hospital Facility in Albert County

[2]No Longer a Hospital in Queens County

All cell values below 5 were randomly rounded to either a 0 or a 5 for privacy restrictions

Table 4-35 – DAD Data – Traditional Model - Trip Matrix for Health Trips in 2011

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS															Productions Total	Productions Target	Production Ratio
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York			
Albert	8	0	0	0	0	1	0	0	0	0	0	0	0	1,038	0	1,047	1,046	0.9990
Carleton	0	1,059	0	0	0	0	0	0	0	0	0	0	1	0	0	1,060	1,060	1.0000
Charlotte	0	0	1,041	0	0	0	0	0	0	0	135	2	0	0	18	1,195	1,195	1.0000
Gloucester	0	0	0	2,785	0	0	0	0	0	0	0	0	0	0	0	2,785	2,785	1.0000
Kent	0	0	0	0	637	0	0	13	0	0	0	0	0	558	0	1,208	1,210	1.0017
Kings	0	0	0	0	0	520	0	0	7	0	1,173	2	0	63	0	1,765	1,765	1.0000
Madawaska	0	0	0	0	0	0	1,005	0	0	0	0	0	0	0	0	1,005	1,005	1.0000
Northumberland	0	0	0	0	0	0	0	2,225	0	0	0	0	0	0	0	2,225	2,225	1.0000
Queens	0	0	0	0	0	146	0	0	39	0	45	356	0	1	43	630	630	1.0000
Restigouche	0	0	0	0	0	0	0	0	0	1,550	0	0	0	0	0	1,550	1,550	1.0000
Saint John	0	0	0	0	0	18	0	0	0	0	2,220	1	0	0	0	2,240	2,240	1.0000
Sunbury	0	0	0	0	0	0	0	0	1	0	3	274	0	0	403	680	680	1.0000
Victoria	0	1	0	0	0	0	2	0	0	0	0	0	767	0	0	770	770	1.0000
Westmorland	0	0	0	0	2	12	0	0	0	0	0	0	0	4,261	0	4,276	4,275	0.9998
York	0	0	0	0	0	0	0	0	0	0	2	1,845	0	0	998	2,845	2,845	1.0000
Attractions Total	8.0	1,060	1,041	2,785	639	697	1,007	2,238	47	1,550	3,578	2,480	768	5,921	1,462			
Attractions Target	1	900	595	2,690	135	385	1,155	1,705	50	1,135	5,155	1,690	570	6,785	2,330		25,281	
Attraction Ratio	0.13	0.85	0.57	0.97	0.21	0.55	1.15	0.76	1.06	0.73	1.44	0.68	0.74	1.15	1.59			

Table 4-36 – Traditional Model - Difference Between Actual and Predicted Health Trips

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS														
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York
Albert	7	-5	0	-10	-5	-14	0	0	-5	-5	-45	-5	-10	98	0
Carleton	0	419	0	-10	0	-5	0	-5	-5	0	-95	-80	-69	-15	-135
Charlotte	0	-5	511	-5	0	-5	0	-5	0	-5	-395	-78	-10	-5	3
Gloucester	0	-15	-5	515	0	-5	0	-65	-5	-15	-140	-5	-10	-245	-5
Kent	0	-10	-5	-5	537	-5	0	-17	0	0	-40	0	0	-452	-5
Kings	0	0	0	-10	-5	235	0	-5	2	-5	-187	-8	-5	-2	-10
Madawaska	0	-10	-5	0	0	0	125	0	0	-10	-5	-5	-45	-45	0
Northumberland	0	0	-5	-135	-5	0	0	655	-5	-5	-130	-50	-10	-215	-95
Queens	0	-5	-5	-5	0	111	0	-5	34	-5	-60	141	-10	-14	-177
Restigouche	0	-5	-5	-165	-5	0	-95	-10	-5	500	-65	0	-5	-135	-5
Saint John	0	-5	-5	-15	-5	13	-5	-5	-5	-10	55	1	-10	-5	0
Sunbury	0	-5	-5	-20	0	-5	-5	-5	-4	-5	-42	-51	-15	-15	178
Victoria	0	-74	-5	-20	-5	-5	-163	0	0	0	-35	-40	412	-15	-50
Westmorland	0	-10	-10	-15	2	-3	0	0	0	-10	-165	-5	-10	226	0
York	0	-110	-10	-5	-5	0	-5	0	-5	-10	-228	975	-5	-25	-567

The results from the traditional gravity model demonstrate that the gravity model can be applied when distributing heath trips; however, the distance between counties due to the geography of New Brunswick creates a strong variation between predicted and actual trips.

To improve the distribution of trips for the gravity model, a "health factor" was developed to indicate the likelihood of rural and urban travel for hospital trips.

This factor used coefficients of the health-based regression model to indicate which variables should be considered. Categories were developed to indicate the likelihood of health trips being internal and the likelihood of health trips leaving the county for travel. Similar to the deterrence function, the exponential function was used with a value of b=0.1 to develop the health factor.

Westmorland, Saint John and York County were considered urban counties because the three major New Brunswick cities were located in these counties. The following Table 4-37 shows the category of travel and the assigned value to be applied as a cost to the exponential function.

Table 4-37 - Health Factor Values for Zone-to-Zone Trips

Zonal Trips	Value
Rural-to-Rural	30
Rural-to-Rural Internal	1
Urban-to-Urban	15
Urban-to-Urban Internal	0
Rural-to-Urban	-50
Urban-to-Rural	200
No Hospital Facility	40
Urban-to-Rural Unexplained	-70

Urban-to-urban internal health trips were assigned a value of zero meaning they were the most likely trip to occur. Rural-to-rural internal trips were assigned a value of 1 making them likely to occur but still less likely than an urban-to-urban internal trip.

County-to-County Trips:

Urban-to-urban trips were assigned a value of 15 while rural-to-rural trips were assigned a value of 30. This was because it was more likely to travel from one urban county to another rather than travel from one rural county to another. It was expected that if a trip produced from a rural county is leaving the internal zone, that it was most likely travelling to the next closest urban county. To account for this, rural-to-urban trips were assigned a value of -50. A negative value was applied to create a stronger attraction from a rural county to an urban county while still considering the distance cost. While still considering the deterrence function for distance, this allows the proximity of the nearest urban location to be the desired trip destination instead while still making internal trips the most appealing.

Urban-to-rural trips were given a value of +200 because these trips were unlikely. It was assumed that rural locations would not have more resources than an urban location therefore urban trips would be unlikely to travel to rural locations for healthcare. Although trips from York County to Sunbury County is an Urban-to-rural trip, Sunbury County was made an “urban” county for this application because of its close proximity with Fredericton.

Albert County and Queens county were given an attraction value of 40 in attempt to deter trips from being attracted to these locations. This is because these counties do not have a Hospital Facility therefore health trips to these locations are highly unlikely. It should be

noted that Queens County used to have a hospital therefore some internal trips were noted; however other trips should not be attracted here.

An outlier value was applied to York-to-Carleton County as a value of -70. This was to help trips from York County be attracted to Carleton County even though this is considered an Urban-to-rural trip. The cost matrix for the health factor is shown below in Table 4-38 and the new trip matrix after applying the health factor and deterrence factor is shown in Table 4-39.

Table 4-38 - Health Factor Cost for Zone-to-Zone Travel

		DESTINATIONS														
	ORIGINS	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury*	Victoria	Westmorland	York
R	Albert	40	30	30	30	30	30	30	30	40	30	-50	30	30	-50	-50
R	Carleton	40	1	30	30	30	30	30	30	40	30	-50	30	30	-50	-50
R	Charlotte	40	30	1	30	30	30	30	30	40	30	-50	30	30	-50	-50
R	Gloucester	40	30	30	1	30	30	30	30	40	30	-50	30	30	-50	-50
R	Kent	40	30	30	30	1	30	30	30	40	30	-50	30	30	-50	-50
R	Kings	40	30	30	30	30	1	30	30	40	30	-50	30	30	-50	-50
R	Madawaska	40	30	30	30	30	30	1	30	40	30	-50	30	30	-50	-50
R	Northumberland	40	30	30	30	30	30	30	1	40	30	-50	30	30	-50	-50
R	Queens	40	30	30	30	30	30	30	30	40	30	-50	30	30	-50	-50
R	Restigouche	40	30	30	30	30	30	30	30	40	1	-50	30	30	-50	-50
U	Saint John	200	200	200	200	200	200	200	200	200	200	0	200	200	15	15
U	Sunbury	40	30	30	30	30	30	30	30	40	30	15	1	30	15	15
R	Victoria	40	30	30	30	30	30	30	30	40	30	-50	30	1	-50	-50
U	Westmorland	200	200	200	200	200	200	200	200	200	200	15	200	200	0	15
U	York	200	-70	200	200	200	200	200	200	200	200	15	15	200	15	0
		R	R	R	R	R	R	R	R	R	R	U	R	R	U	U

Table 4-39 - DAD Data – Health Model - Trip Matrix for Health Trips in 2011

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS															Productions Total	Productions Target	Production Ratio
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York			
Albert	**0**	0	0	0	0	0	0	0	0	0	0	0	0	1,046	0	1,046	1,046	1.000
Carleton	0	**1,035**	0	0	0	0	0	0	0	0	0	0	0	0	25	1,060	1,060	1.000
Charlotte	0	0	**477**	0	0	0	0	0	0	0	624	0	0	0	94	1,195	1,195	1.000
Gloucester	0	0	0	**2,785**	0	0	0	0	0	0	0	0	0	0	0	2,785	2,785	1.000
Kent	0	0	0	0	**23**	0	0	0	0	0	0	0	0	1,186	1	1,210	1,210	1.000
Kings	0	0	0	0	0	**6**	0	0	0	0	1,660	0	0	99	1	1,766	1,765	0.999
Madawaska	0	0	0	0	0	0	**1,005**	0	0	0	0	0	0	0	0	1,005	1,005	1.000
Northumberland	0	0	0	0	0	0	0	**2,212**	0	0	0	0	0	2	11	2,225	2,225	1.000
Queens	0	0	0	0	0	0	0	0	**0**	0	295	1	0	6	328	630	630	1.000
Restigouche	0	0	0	0	0	0	0	0	0	**1,550**	0	0	0	0	0	1,550	1,550	1.000
Saint John	0	0	0	0	0	0	0	0	0	0	**2,240**	0	0	0	0	2,240	2,240	1.000
Sunbury	0	0	0	0	0	0	0	0	0	0	1	**434**	0	0	245	680	680	1.000
Victoria	0	0	0	0	0	0	0	0	0	0	0	0	**770**	0	0	770	770	1.000
Westmorland	0	0	0	0	0	0	0	0	0	0	0	0	0	**4,275**	0	4,275	4,275	1.000
York	0	66	0	0	0	0	0	0	0	0	0	1,162	0	0	**1,616**	2,844	2,845	1.000
Attractions Total	0	1,101	477	2,785	23	6	1,005	2,212	0	1,550	4,820	1,597	770	6,614	2,321			
Attractions Target	1	900	595	2,690	135	385	1,155	1,705	50	1,135	5,155	1,690	570	6,785	2,330		**25,281**	
Attraction Ratio	0	0.82	1.25	0.97	5.87	64.17	1.15	0.77	0	0.73	1.07	1.06	0.74	1.03	1.00			

Table 4-40 – DAD - Difference of Predicted to Actual Trips

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS														
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York
Albert	-1	-5	0	-10	-5	-15	0	0	-5	-5	-45	-5	-10	106	0
Carleton	0	395	0	-10	0	-5	0	-5	-5	0	-95	-80	-70	-15	-110
Charlotte	0	-5	-54	-5	0	-5	0	-5	0	-5	95	-80	-10	-5	78
Gloucester	0	-15	-5	515	0	-5	0	-65	-5	-15	-140	-5	-10	-245	-5
Kent	0	-10	-5	-5	-77	-5	0	-30	0	0	-40	0	0	176	-4
Kings	0	0	0	-10	-5	-279	0	-5	-5	-5	299	-10	-5	34	-9
Madawaska	0	-10	-5	0	0	0	125	0	0	-10	-5	-5	-45	-45	0
Northumberland	0	0	-5	-135	-5	0	0	642	-5	-5	-130	-50	-10	-213	-84
Queens	0	-5	-5	-5	0	-35	0	-5	-5	-5	190	-214	-10	-9	107
Restigouche	0	-5	-5	-165	-5	0	-95	-10	-5	500	-65	0	-5	-135	-5
Saint John	0	-5	-5	-15	-5	-5	-5	-5	-5	-10	75	0	-10	-5	0
Sunbury	0	-5	-5	-20	0	-5	-5	-5	-5	-5	-44	187	-15	-15	-58
Victoria	0	-75	-5	-20	-5	-5	-165	0	0	0	-35	-40	415	-15	-50
Westmorland	0	-10	-10	-15	0	-15	0	0	0	-10	-165	-5	-10	240	0
York	0	-32	-10	-5	-5	0	-5	0	-5	-10	-229	-1	-5	-25	333

When looking at the difference between actual and predicted trips from the health model, some county trips were affected positively by this factor and some remained unchanged. The positive numbers indicate that the model is predicting more trips to a location than the actual number and a negative number indicates that the model is predicting less trips than actual.

These factors had the following effects on the variation in county trips:

a. Charlotte County internal trips improved significantly dropping from a difference of +511 to a -54. The Trips were also properly distributed from Charlotte County to Saint John County changing the difference in trips from a -395 to a +53.

b. Kent County internal trips improved significantly from a difference in trips +537 to -77. Trips from Kent to Westmorland changed from -452 to a +176

c. Kings County went from a -187 trips to Saint John County to a +299.

   Internal trips in Kings County changed from +235 to -279

d. Internal trips for York County changed from -567 to +333. Health trips travelling from York County to Sunbury County improved from a difference of +957 to -1. York County trips to Carleton County trips improved from a difference of -110 to -32.

e. Sunbury County internal trips changed from a difference of -51 to +187.

This demonstrates that the health factors applied did have some beneficial influence on the health trips, however, still have variance in other locations therefore should be explored more.

4.4.2 Traditional Gravity Model with the Physician Billing Data

The identical process was applied with the 2011 health trips from the physician billing data as shown in Table 4-41. As seen in the trip matrix, the health trips to a physician's office or location are much more distributed from county to county. This is because these trip types are not travelling to one healthcare facility but to many different locations depending on the health need and the location of the physician. When applying the gravity model to the attractions and productions the traditional form of the gravity model does not properly distribute the physician billing health trips for the 2011 data as shown in Table 4-42 below.

Westmorland County attracted the most health trips to a physician's office or location with a total of 126,600 trips in 2011, followed by York, Gloucester and Saint John County.

Table 4-41 - 2011 Trip matrix from Physician Billing Data

ORIGIN	DESTINATION															
	Albert	**Carleton**	**Charlotte**	**Gloucester**	**Kent**	**Kings**	**Madawaska**	**Northumberland**	**Queens**	**Restigouche**	**Saint John**	**Sunbury**	**Victoria**	**Westmorland**	**York**	**Total**
Albert	**3,155**	0	5	10	45	45	15	10	0	0	535	5	5	15,540	35	19,405
Carleton	10	**13,895**	1,480	0	5	55	65	60	5	15	965	25	875	110	1,930	19,495
Charlotte	5	210	**9,230**	0	0	2,075	20	0	10	5	6,310	30	0	100	370	18,365
Gloucester	0	15	0	**53,615**	5	30	10	715	0	335	500	0	5	3,110	695	59,035
Kent	255	35	0	30	**5,195**	10	20	475	10	5	415	5	5	15,515	35	22,010
Kings	35	5	1,115	15	10	**13,595**	20	5	265	15	13,040	30	5	1,205	420	29,780
Madawaska	0	10	5	20	5	15	**6,850**	5	5	60	15	15	615	445	25	8,090
Northumberland	20	975	5	2,205	30	155	30	**28,315**	0	15	510	10	5	5,875	2,635	40,785
Queens	5	460	35	0	5	505	270	75	**5,700**	30	910	470	0	250	3,555	12,270
Restigouche	0	85	5	5,005	5	10	350	85	5	**22,130**	425	5	0	1,780	675	30,565
Saint John	5	0	805	5	5	5,975	30	0	25	30	**29,265**	15	5	395	295	36,855
Sunbury	5	135	60	5	0	35	550	50	1,510	25	415	**3,195**	20	110	5,120	11,235
Victoria	0	1,365	5	35	5	35	435	45	5	30	295	15	**8,655**	225	810	11,960
Westmorland	2,490	10	0	55	525	95	210	45	35	10	1,955	30	25	**81,480**	95	87,060
York	5	2,465	1,410	5	5	220	565	1,120	300	150	2,920	4,925	45	460	**51,410**	66,005
Total	5,990	19,665	14,160	61,005	5,845	22,855	9,440	31,005	7,875	22,855	58,475	8,775	10,265	126,600	68,105	472,915

Table 4-42 – Traditional Gravity Model using Physician Billing Health Trips

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS															Productions Total	Productions Target	Production Ratio
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York			
Albert	**8,589**	0	0	0	0	13	0	0	0	0	1	0	0	10,802	0	19,405	19,405	1.000000
Carleton	0	**19,484**	0	0	0	0	0	0	0	0	0	0	8	0	3	19,495	19,495	1.000000
Charlotte	0	4	**17,251**	0	0	0	0	0	0	0	815	2	0	0	293	18,365	18,365	1.000000
Gloucester	0	0	0	**59,034**	0	0	0	1	0	0	0	0	0	0	0	59,035	59,035	1.000000
Kent	1	0	0	0	**13,220**	6	0	160	70	0	0	0	0	8,543	10	22,010	22,010	1.000000
Kings	0	0	0	0	0	**17,215**	0	0	1,308	0	10,166	3	0	1,078	9	29,779	29,780	1.000034
Madawaska	0	0	0	0	0	0	**8,084**	0	0	0	0	0	6	0	0	8,090	8,090	1.000000
Northumberland	0	0	0	2	8	0	0	**40,772**	0	0	0	0	1	0	2	40,785	40,785	1.000000
Queens	0	0	0	0	3	4,156	0	0	**6,341**	0	332	543	0	13	882	12,270	12,270	1.000000
Restigouche	0	0	0	0	0	0	0	0	0	**30,565**	0	0	0	0	0	30,565	30,565	1.000000
Saint John	0	0	2	0	0	1,108	0	0	11	0	**35,722**	4	0	0	7	36,854	36,855	1.000027
Sunbury	0	0	0	0	0	2	0	0	114	0	27	**534**	0	0	10,558	11,235	11,235	1.000000
Victoria	0	24	0	0	0	0	14	1	0	0	0	0	**11,921**	0	0	11,960	11,960	1.000000
Westmorland	170	0	0	0	45	463	0	0	2	0	0	0	0	**86,380**	0	87,060	87,060	1.000000
York	0	2	3	0	0	5	0	0	140	0	33	7,952	0	0	**57,871**	66,006	66,005	0.999985
Attractions Total	8,760.0	19,514	17,256	59,036	13,276	22,968	8,098	40,934	7,986	30,565	47,096	9,038	11,936	106,816	69,635			
Attractions Target	5,990	19,665	14,160	61,005	5,845	22,855	9,440	31,005	7,875	22,855	58,475	8,775	10,265	126,600	68,105		**472,915**	
Attraction Ratio	0.68	1.01	0.82	1.03	0.44	1.00	1.17	0.76	0.99	0.75	1.24	0.97	0.86	1.19	0.98			

Table 4-43 - Variation between Actual and Predicted

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS														
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York
Albert	5,434	0	-5	-10	-45	-32	-15	-10	0	0	-534	-5	-5	-4,738	-35
Carleton	-10	5,589	-1,480	0	-5	-55	-65	-60	-5	-15	-965	-25	-867	-110	-1,927
Charlotte	-5	-206	8,021	0	0	-2,075	-20	0	-10	-5	-5,495	-28	0	-100	-77
Gloucester	0	-15	0	5,419	-5	-30	-10	-714	0	-335	-500	0	-5	-3,110	-695
Kent	-254	-35	0	-30	8,025	-4	-20	-315	60	-5	-415	-5	-5	-6,972	-25
Kings	-35	-5	-1,115	-15	-10	3,620	-20	-5	1,043	-15	-2,874	-27	-5	-127	-411
Madawaska	0	-10	-5	-20	-5	-15	1,234	-5	-5	-60	-15	-15	-609	-445	-25
Northumberland	-20	-975	-5	-2,203	-22	-155	-30	12,457	0	-15	-510	-10	-4	-5,875	-2,633
Queens	-5	-460	-35	0	-2	3,651	-270	-75	641	-30	-578	73	0	-237	-2,673
Restigouche	0	-85	-5	-5,005	-5	-10	-350	-85	-5	8,435	-425	-5	0	-1,780	-675
Saint John	-5	0	-803	-5	-5	-4,867	-30	0	-14	-30	6,457	-11	-5	-395	-288
Sunbury	-5	-135	-60	-5	0	-33	-550	-50	-1,396	-25	-388	-2,661	-20	-110	5,438
Victoria	0	-1,341	-5	-35	-5	-35	-421	-44	-5	-30	-295	-15	3,266	-225	-810
Westmorland	-2,320	-10	0	-55	-480	368	-210	-45	-33	-10	-1,955	-30	-25	4,900	-95
York	-5	-2,463	-1,407	-5	-5	-215	-565	-1,120	-160	-150	-2,887	3,027	-45	-460	6,461

Table 4-44 – Health-based Gravity Model using Physician Billing Data

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS															Productions Total	Productions Target	Production Ratio
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York			
Albert	**5,262**	0	0	0	0	5	0	0	0	0	1	0	0	14,137	0	19,405	19,405	1.000000
Carleton	0	**19,468**	0	0	0	0	0	0	0	0	0	0	5	0	22	19,495	19,495	1.000000
Charlotte	0	2	**14,281**	0	0	0	0	0	0	0	2,840	1	0	0	1,241	18,365	18,365	1.000000
Gloucester	0	0	0	**59,034**	0	0	0	1	0	0	0	0	0	0	0	59,035	59,035	1.000000
Kent	0	0	0	0	**6,948**	3	0	28	19	0	0	0	0	14,994	18	22,010	22,010	1.000000
Kings	0	0	0	0	0	**12,980**	0	0	331	0	14,648	1	0	1,805	16	29,781	29,780	0.999966
Madawaska	0	0	0	0	0	0	**8,086**	0	0	0	0	0	4	0	0	8,090	8,090	1.000000
Northumberland	0	0	0	1	8	0	0	**40,759**	0	0	0	0	0	1	14	40,783	40,785	1.000049
Queens	0	0	0	0	2	3,469	0	0	**4,835**	0	874	230	0	40	2,821	12,271	12,270	0.999919
Restigouche	0	0	0	0	0	0	0	0	0	**30,565**	0	0	0	0	0	30,565	30,565	1.000000
Saint John	0	0	0	0	0	588	0	0	3	0	**36,256**	1	0	0	5	36,853	36,855	1.000054
Sunbury	0	0	0	0	0	7	0	0	185	0	20	**1,308**	0	0	9,715	11,235	11,235	1.000000
Victoria	0	15	0	0	0	0	8	0	0	0	0	0	**11,937**	0	0	11,960	11,960	1.000000
Westmorland	80	0	0	0	13	209	0	0	0	0	0	0	0	**86,757**	0	87,059	87,060	1.000011
York	0	0	1	0	0	2	0	0	33	0	16	7,827	0	0	**58,125**	66,004	66,005	1.000015
Attractions Total	5,342.0	19,485	14,282	59,035	6,971	17,263	8,094	40,788	5,406	30,565	54,655	9,368	11,946	117,734	71,977			
Attractions Target	5,990	19,665	14,160	61,005	5,845	22,855	9,440	31,005	7,875	22,855	58,475	8,775	10,265	126,600	68,105		**472,915**	
Attraction Ratio	1.12	1.01	0.99	1.03	0.84	1.32	1.17	0.76	1.46	0.75	1.07	0.94	0.86	1.08	0.95			

Table 4-45 – Variation between Actual and Predicted

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS														
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York
Albert	2,107	0	-5	-10	-45	-40	-15	-10	0	0	-534	-5	-5	-1,403	-35
Carleton	-10	5,573	-1,480	0	-5	-55	-65	-60	-5	-15	-965	-25	-870	-110	-1,908
Charlotte	-5	-208	5,051	0	0	-2,075	-20	0	-10	-5	-3,470	-29	0	-100	871
Gloucester	0	-15	0	5,419	-5	-30	-10	-714	0	-335	-500	0	-5	-3,110	-695
Kent	-255	-35	0	-30	1,753	-7	-20	-447	9	-5	-415	-5	-5	-521	-17
Kings	-35	-5	-1,115	-15	-10	-615	-20	-5	66	-15	1,608	-29	-5	600	-404
Madawaska	0	-10	-5	-20	-5	-15	1,236	-5	-5	-60	-15	-15	-611	-445	-25
Northumberland	-20	-975	-5	-2,204	-22	-155	-30	12,444	0	-15	-510	-10	-5	-5,874	-2,621
Queens	-5	-460	-35	0	-3	2,964	-270	-75	-865	-30	-36	-240	0	-210	-734
Restigouche	0	-85	-5	-5,005	-5	-10	-350	-85	-5	8,435	-425	-5	0	-1,780	-675
Saint John	-5	0	-805	-5	-5	-5,387	-30	0	-22	-30	6,991	-14	-5	-395	-290
Sunbury	-5	-135	-60	-5	0	-28	-550	-50	-1,325	-25	-395	-1,887	-20	-110	4,595
Victoria	0	-1,350	-5	-35	-5	-35	-427	-45	-5	-30	-295	-15	3,282	-225	-810
Westmorland	-2,410	-10	0	-55	-512	114	-210	-45	-35	-10	-1,955	-30	-25	5,277	-95
York	-5	-2,465	-1,409	-5	-5	-218	-565	-1,120	-267	-150	-2,904	2,902	-45	-460	6,715

After applying the health factors in the same way as the DAD data, there was not much change in variation of predicted and actual as shown in Table 4-44 and Table 4-45.

4.4.2.1 Summary

The following conclusions were determined from applying a base application of the gravity model to both the DAD and Physician Billing Datasets:

a. The geography of New Brunswick causes the cost between counties, in terms of distance (km), to be significant from one county to another. This resulted in the majority of trips being assigned to internal trips making travel outside of the county to be "unattractive".

b. The trip matrix from the 2011 DAD data indicates that typically trips to hospitals are kept internally; however, a number of counties are travelling to other counties for hospital trips even though the distance is significant.

c. The DAD data indicates that Westmorland County and Saint John County attract the most health trips to a hospital even though they are both located within the South-East end of the province.

d. The trip matrix from the 2011 Physician Billing data indicates that health trips to physician offices/ locations are being distributed throughout the province to different counties. These data includes many trip types to offices, clinics, etc. therefore trips are not always going to specific facilities.

e. The Physician Billing data indicates that Westmorland County attracts the most health trips, followed by York County, Gloucester and Saint John.

f. When applying the traditional gravity model to the two trip matrixes, the model was able to predict health trip distributions more accurately using the DAD data opposed to using the Physician Billing data.

g. When the "health factor" was applied to the gravity model using the DAD data, some counties experienced significant improvements to the predicted distributions.

For some counties, the health factor had no effect and for some counties the health factor was too strong and caused the opposite outcome to occur.

h. When the "health factor" was applied to the gravity model using the Physician Billing data, there was minimal changes or improvements to the model.

i. Overall, the gravity model was generally successful in predicting health trips to hospitals; however additional effort is required to improve the current health factor. An additional health factor for the type of health trip should also be explored. The gravity model was only partially successful when predicting the Physician Billing Data and the health factors had minimal effect on improving the prediction of these trips. Therefore, the physician billing data will require additional effort to improve a health factor specific to that data.

4.5 Research Challenges and Limitations

When completing the analysis, several notable challenges and limitations were identified. The following limitations were found when completing the research:

1. **County Size and Boundary Trips**: For Counties with large areas, trips originating at the edge of a county boundary may have been closer to the hospital located in the next county. Even though distance calculations were determined using the DA centroids, it was difficult to capture the boundary trips when analysing the data as the trips were categorized and evaluated from a county-to-county level. Additionally, the deterrence function used in the gravity model assumed a general travel cost in distance (km) which was estimated from the general center of the County or Hospital location.
2. **DA Centroids in Rural Areas**: When determining a DA centroid, the centroid was placed in the geographic center of the DA using ArcMap. For urban areas this was a good representation of the DA due to small size and high population. However,

in rural areas, the size of the DA was much larger than the size of DA in the urban areas and the population within the DA was not as proportional as an urban DA. To address this limitation, a population-weighted centroid should be considered to better represent the DA in rural areas.

3. **Origin Assumptions**: It was assumed that all trips where originating from home based on the person's address. This is a good initial assumption and does capture the extent of travel for what a health trip would require when originating from home; however, there are likely trips not originating from home (or in a chain of trips) therefore those distances were not captured. In addition, long distance health trips that may have occurred over a series of days and require a person to stay overnight in the destination city of the health trip were not captured. This would not capture what is required of the person to stay overnight for a health trip; however, by originating from home does show the amount of travel required if a person did not have the means to stay overnight for reoccurring health visits.
4. **Trip Assumptions**: The health data at NB-IRDT provided a series of information for a health trip including each type of physician the person saw during the hospital visit. When completing this analysis, the first trip type was used as trip purpose as it was assumed to be the reason for the person making a hospital trip. When further investigating this data, consideration could be given to using the last health trip type noted for that health trip. This may provide a more insight as to what health purpose the trip ended up being for.
5. **Hospital Status**: When analysing the data, the sample years were chosen to correspond with the available census information however, the history of hospital

closures and openings were not considered. Though aggregating multiple years of data should help smooth out any temporary changes in service, this could have had an impact on the number of trips being attracted to a county due to the change in hospital availability throughout the sample years.

6. **Fiscal Years vs. Calendar years**: The DAD data is reported by fiscal year while the Physician Billing data was reported by calendar year. For the purposes of this research the DAD data included part of 2005 health trips to capture trips that would be considered in the 2006 calendar year however, having this data in a fiscal year made it more difficult to compare with the census data as it is collected as a calendar year. If the DAD data was able to be reported in calendar years or was able to be converted to a calendar year format it would improve the ability to compare both datasets to the census data.
7. **Restrictive Access to Data**: When working with this data it was important to go into the lab at NB-IRDT with a specific plan as the lab holds sensitive data and is highly secure. There is no internet access within the lab, and notes are not able to be brought in or out of the lab unless gone through an approval process. This made it challenging to research the best approaches for analysis using specific programs or troubleshoot errors that may have occurred. This required leaving the lab to research and then proceeding with another plan when re-entering the lab.

CHAPTER 5 : CONCLUSIONS & RECOMMENDATIONS

This chapter summarizes the findings determined from analyzing the use of administrative health data for transportation purposes. As well, recommendations and opportunities for further research will be discussed.

5.1 Summary of Work

This research explored the use of administrative health data from the NBIRDT for the Province of New Brunswick for quantifying older adult (65 years and older) travel behaviour for health purposes. It included: the development of a methodology to develop origin/destination data from administrative health records; to create descriptive statistics associated with health use and travel by geography; to develop regression and cross-classification models to connect demographic and geographic attributes to travel behaviour; and the development of a health trip distribution model based upon the gravity model for the purposes of estimating how changes in health services impact travel.

Overall, this research proved that administrative health data can be used for transportation planning purposes. Although applications are in preliminary stages, it was determined that these data can be used for trip generation and be applied to the gravity model for trip distribution. This research provides baseline predictive modelling for expected distance travelled for health trips, trip rates and distribution modelling for New Brunswick health trips by older adults 65 years or older.

5.1.1 Methodology development

Two databases from NB-IRDT were explored which included information from DAD and Physician Billing. Two datasets were created from these databases including the fiscal years 2005-2011 and the calendar years 2006-2011. These datasets were used in conjunction with the citizen dataset, the physician registry and the PCCF+ file to create two datasets that were able to be explored.

These datasets presented patient health records which needed to be processed in order to isolate relevant variables which could be used to calculate travel distances. Using ArcMap and relevant shape files/ databases provided by NB-IRDT, a trip distance was calculated for an individual health trip from each of the final datasets. These data was combined with other variables relevant to transportation travel demand modelling such as demographics, income level, DA location, etc.

Currently New Brunswick does not have a standard method for understanding these health-related travel needs for the age group that experience the most health needs. This provided information available for transportation planning in New Brunswick to use health record data to better understand the travel needs and behaviours of older adults aged 65+ travelling for a health trip.

In addition, cross classification tables were created which provide trip rates for health trips to a hospital and health trips to a physician's office/ clinic location for various age groups 65+ and originating counties. These are baseline trip rates that can help provide a better understanding of which geographic location health trips are being generated from and by which age group.

A regression model was developed to predict the expected distance (km) travelled by an older adult in New Brunswick based on various health and geographic variables. This was built on existing data which can improve the understanding of older adults travelling for healthcare.

A base application of the gravity model was applied using both the DAD and Physician Billing Data. In addition, a preliminary health factor was developed to demonstrate the potential of a health factor (or additional factor) could have on predicting trip distribution.

Overall, the major contribution of this research is that it was determined that health data can be used for transportation planning after proper processing. Health data presents a strong potential to understanding the travel behaviours of older adults who travel for a health trip. These data also provide information to understanding which type of health trips require the most travel and which trips may require more assistance with travel than other types of health trips. Understanding these travel behaviours of older adults have the potential for planners to identify areas which may have needs for travel assistance, additional services, etc.

5.1.2 Development of Descriptive Statistics

The final datasets used for this research sampled a total of 171,155 health trips to a hospital and 2,882,375 health trips designated for physician visits.

The number of hospital trips and physician visits were quantified by age, gender, origin county and destination information, as well as type of health trip, including by specialty group type. This provided relevant information from a transportation perspective that can identify an estimated number of health trips that older adults would be expected to need

assistance with travel depending on the type of health trip. Some of the major findings regarding descriptive statistics included:

- On average, approximately 27% of the province aged 65+ travelled over 60km for a health trip to a hospital, with those in Charlotte County travelling the farthest on average to a hospital (92 km) and to a physician (85 km) one way.
- On average 49% of the trips taken to the hospital by those aged 65+ in New Brunswick were over 50km (one way).
- Older adults travel farther for medical specialties and surgical specialties than general practice, with those in Charlotte, Restigouche and Victoria Counties all travelling over 100km one way on average.
- There is a negative linear relationship between the average distance traveled for health trips per county vs. the percent of health trips originating in a county for both hospital data and physician billing datasets; the greater the number of health trips produced in a county, the shorter the distance travelled.
- Ophthalmology makes up at least 10% of trips for all counties travelling to a hospital. Many counties experience 15-20% of health trips from older adults travelling for ophthalmology and some experience greater than 20%. It should be noted that older adults travelling for an ophthalmology health trips are seeking care from an ophthalmologist which is a medical physician specialized in ophthalmology. This includes medical conditions/ surgical treatments related to the eye. It should be noted that an ophthalmologist is different than an optometrist which are basic eye care specialists. It is expected that typically older adults

travelling for ophthalmology are unable to drive themselves to appointments/treatments.

5.1.3 Development of Regression models

Two regression models for the DAD and Physician Billing datasets were created which were able to predict an expected travel distance (km) for health trips in New Brunswick. This used demographic and geographic variables as well as type of health trip to predict travel distances for older adults.

The DAD model resulted in a model with a best fit of an adjusted R^2 = 0.487 when considering health data and geographic data together however, the adjusted R^2 reduced to a best fit of R^2 = 0.129 when health variables were removed, and geographic and demographic variables were reduced to those which could be found within census data. This indicates that for health trips, the health variables had a large influence on predicting distance travelled. When using only the variables which would be available through the census data, predicting travel distance was proven to be more difficult. This is likely due to the fact that health travel in New Brunswick is heavily influenced on the type of health trip being taken rather than the distance required to be travelled. This is expected to be due to the more urban counties having more resources for specialty type trips which typically cause the surrounding counties to travel.

The model produced from the Physician Billing data resulted in a best fit with an adjusted R^2 = 0.254, which is a lower R^2 than that of the DAD data; however, due to the nature of high variability within the Physician Billing dataset (i.e. more dispersed destinations than in the hospital data), a lower R^2 value was expected. Similarly, when the health-related

variables were removed, the adjusted R^2 of the model was reduced to an $R^2 = 0.069$. Other notable findings include:

- The age variable in models for both sets of data (DAD, Physician Billing) had a negative, and very small, coefficient (e.g., 0.02 km and 0.15 per year of age, respectively) meaning its contribution to travel distance in the model was small. It is expected that the age variable would have more of a contribution if other ages below 65 years were included in the model.
- In both datasets, variables relating to rural origins contributed to longer travel distances
- All values within the models created for both datasets resulted in a VIF of 4 or less which indicates a low probability of multicollinearity being present within the models.
- Specialty trips contributed to significantly longer travel distances in the models than general practice trips.

5.1.4 Development of Cross-Classification Tables for Trip Generation

Cross-classification tables were created to generate trip rates using the final DAD and Physician Billing datasets. The average health trip frequency, by older adults 65+ in New Brunswick, were able to permit the estimation of trips and demonstrate the relationship between trip rates, patient attributes, and geographic location. Notable findings from the data include:

- Overall trip rates for older adults in New Brunswick travelling to hospitals between the fiscal years of 2005-2011 ranged between 0.14 to 0.33 health trips per person

(65+) per year. Overall trips rates for older adults travelling for a physician visit in New Brunswick ranged between 3.25 to 5.38 health trips per person (65+) per year.

- Older adults who took a hospital trip between the fiscal years of 2005-2011 averaged between a trip rate of 1.31 to 1.48 per person (65+) per year, while those making a physician visit between the calendar years 2005-2011 averaged a trip rate between 7.36 to 10.49 health trips per person (65+) per year.
- Overall, men were found to have a higher hospital trip rate per person than women for all age groups.
- The most common age groups that generated the highest trip rates per person for hospital trip rates were ages 80-84 and 85-89, while the most common Physician Billing trip rates were generated from persons aged 85-89 years.
- Trip rates per person for general practice trips increased as age increased. The peak rate for general practice trips was for the 90+ age group with a rate of 0.20 trips per person per year for women and 0.24 for men.
- Trip rates per person for specialty practice trips decreased overall as age increased; however, the peak trip rate for specialty practice trips was ages 75-79 with rates of 0.13 for women and 0.18 for men.
- Trip rates for general practice physician trips increased as age increased for men and decreased as age increased for women. The peak rate for men's general practice trips was for the 90+ age group with a rate of 7.70 trips/person/year and 2.31 trips/person/year for women in age groups 75-79. This is a sizeable difference in trip rates; therefore, it should be noted that the difference in trip rates could be due to missing data when running frequencies on the Physician Billing dataset.

- The greatest county level trip rates per person were generated from Queens County, with trip rates of 0.4 and 9.04 trips/person/year for hospitals and physicians, respectively, and Restigouche County with trip rates of 0.42 and 7.58 trips/person/year per year for hospitals and physicians, respectively.
- Trip rates per person for specialty practice trips decreased overall as age increased; however, the peak trip rate for specialty practice trips within genders peaked at 1.36 trips/person/year for women ages 75-79 and 2.32 trips/person/year for men ages 90+.
- When running frequency tables for trip types and age groups, a total of 263,606 observations were missing for 2006-2011 Physician Billing dataset. This left 2,618,810 total trips used to create trip rate tables. Due to the nature of using administrative data it is recognized that this could affect trips rates; however, rates were determined a starting point for New Brunswick health trips.

5.1.5 Development of a Health Trip Distribution Model

The gravity model, a typical tool used with trip distribution, has the potential to quantify the impacts of health travel due to the changes to health facilities using attributes which attract trips to health facilities such as type of trips, hospital beds or "costs" such as distance.

Although preliminary applications of the model were explored it was discovered that a significant level of effort should be considered to properly calibrate the gravity model for health trips. The gravity model uses distance as a main cost when distributing trips as the theory suggests that a trip would become unattractive as the distance increased. This proves to be true for typical travel; however, for health trips it was found that trips are attracted to

zones even when a large travel distance is required. This indicates that travel for health trips is attracted to locations with more than distance in consideration. It was challenging to attract health trips to areas with a large distance cost due to the geography of New Brunswick. New Brunswick counties were found to have a large distance between most other counties that made the gravity model distribute most of the health trips internally. However, due to the nature of the types of health trips, some health trips required more resources or services than origin counties were able to provide. This required further travel by older adults that the gravity model had difficulty predicting.

An opportunity for a health factor that represents the type of health trip, type of facility or number of beds would be an opportunity to explore with next steps. There should be a factor developed that can better attract these health trips to the appropriate destination that will fit the need of the health trip. This would then present the opportunity of trips to be redistributed based on purpose or health need when a service or facility is affected.

Overall, the gravity model was generally successful in predicting health trips to hospitals; however additional effort is required to improve the current health factor. An additional health factor for the type of health trip should also be explored. The Gravity model was only partially successful when predicting the Physician Billing Data and the health factors had minimal effect on improving the prediction of these trips. Therefore, the physician billing data will require additional effort to improve a health factor specific to that data. Notable observations from the initial gravity distribution model:

- The geography of New Brunswick (low population density) causes the cost between counties, in terms of distance (km), to be significant from one county to another.

This resulted in the majority of trips being assigned to internal trips making travel outside of the county to be “unattractive”.

- The trip matrix from the 2011 DAD data indicates that typically trips to hospitals are kept internally; however, a number of counties are travelling to other counties for hospital trips even though the distance is significant.
- When applying the traditional gravity model to the two trip matrixes, the model was able to predict health trip distributions more accurately using the DAD data as opposed to using the Physician Billing data

5.2 Recommendations

This research has provided a baseline approach to better understanding health travel in New Brunswick; however, further research is required to improve results and explore further opportunities for administrative health data. Going forward, there are six main recommendations that should continue with this research that are listed below:

1. ***Further Development of the Health Factor:***

   A preliminary health factor has been developed and applied as a base application; however, further development of this factor should be explored. The type of trip should be integrated into the health factor as it is now known that trip type influences the distance travelled depending on the origin county.

   The health models should be further explored independently of each other and have health factors specific to each of the data sources: from the DAD dataset and Physician Billing dataset. This would be beneficial since each dataset is providing different health trip needs; therefore, a universal health factor may not be the best

approach. Hospital trips require more resources than a trip to physician's office or clinic.

Different functions other than the exponential function for cost in the health factor should be explored, as well as different ways for applying the health factor. For this research the health factor was applied in the similar way the K-factor is and the deterrence factor. An alternative method may be more beneficial such as a ratio or a measurable attribute such as number of beds or level of facility.

2. ***Applying an application to reduce a service or facility:***

   A modelling approach should be explored where a reduction in service occurs the redistribution of health trips can be predicted by the model. This service reduction would represent a facility closure or removal of service to a location and the forecasting model would demonstrate the results of that colure. This may be dependent on a specific health trip factor to indicate the proportion of trips that would be affected by a service disruption and to where they would then be attracted.

3. ***Explore Travel Modes:***

   The health data was unable to provide the model of travel for those taking health trips other than when arriving at a hospital as emergency or non-emergency. Due to health trips being a large part of trips for older adults, further research into how adults 65 years and older are travelling to destinations for health trips would be a beneficial aspect to understanding older adult travel.

4. ***Further develop a County Specific Model:***

   A county specific model would allow a better understanding of health trips for older adults in rural locations as they tend to be the ones who travel the greatest distances.

The Physician Billing dataset may be the better dataset to use for a county specific model to understand the general travel of rural areas; whereas the DAD dataset captures hospital trips which would have minimal destination locations within one county.

5. ***Explore additional variables available within the Administrative Datasets:***

   There were more variables available in the DAD and Physician Billing datasets which could be explored further for their applicability to transportation.

6. ***Explore destination counties and how travellers are attracted to hospitals:***

   Further explore where types of health trips are travelling to. This could better represent which counties are attracting which specialty types and if there are noticeable specialties that older adults are travelling to one place to receive care.

LIST OF REFERENCES

Bhat, C., & Koppelman, F. (1999). Activiy Based Modeling of Travel Damand. In R. Hall, *Handbook of Transportation Science. International Series in Operations Research & Managment Science. Volume 23.* Boston, MA: Springer.

Canadian Institute for Health Information. (2018). *In Pursuit of Health Equity: Defining Stratifiers for Measuring Health Inequality — A Focus on Age, Sex, Gender, Income, Education and Geographic Location.* Ottawa: CIHI.

Canadian Institute for Health Information. (2019). *Your Health System: In Depth — Technical Notes for Contextual Measures.* Ottawa: CIHI.

Castiglione, J., Bradley, M., & Gliebe, J. (2015). *Activity-Based Travel Demand Mdoels: A Primer.* Washington, D.C: Tranportation Research Board.

Chang, Y.-C., & Mastrangelo, C. (2011). Addressing Multicollinearity in Semiconductor Manufacturing. *Quality and Reliability Engineering International*, 843-854.

Hanson, T. (2018). Can We Better Leverage National Data Collection Initiatives to Support Transportation Planning in Smaller and Rural Areas in Canada? A View from New Brunswick. *Building Tomorrow's Society* (pp. TR43-1-TR43-8). Fredericton, NB: Canadian Society for Civil Engineers (CSCE).

Hanson, T. (2018). Can We Better Leverage National Data Collection Initiatives to Support Transportation Planning in Smaller Citites and Rural Areas in Canada? A View From New Brunswick. *CSCE Conference.* Fredericton: Canadian Society of Civil Engineering.

Hanson, T., & Goudreau, M. (2017). How Many New Brunswickers Rely on Organized Volunteer Driver Programs? Preliminary Tools and Approaches to Help Us Find the Answer. *52nd Annual Conference - Canadian Transportation Research Forum.* Canadian Transportation Research Forum.

Hanson, T., & Hildebrand, E. D. (2011). Can Rural Older drivers meet their needs without a car? Stated adaptation responses from a GPS travel diary survey. *Transportation*, 975-992.

Hildebrand, E. D., & Myrick, B. (2001). Collision Experience and Mobility Concerns of the Rural Elderly. *Canadian Multidisipinary Road Safety Conference XII.* London, Ontario.

Hildebrand, E. D., Myrick, B., & Creed, T. (2000). The Rural Elderly: Driving Patterns and Accident Involvment. *Canadian Transportation Research Forum - Bridging the Gaps* (pp. 175-191). Charlottetown, PEI: CTRF.

Institute of Transportation Engineers. (2016). *Transportation Planning Handbook.* New Jersey: Institute of Transportation Engineers.

Luo, W., & Wang, F. (2003). Measures of Spatial Accessibility to Health Care in a GIS Environment: Synthesis and a case study in the Chicago region. *Environment and Planning B: Planning and Design, 30*, 865-884.

Maikut, K. (2011). Rural Transportation Issues and Strategies. Canada.

Marottoli, R. A., & Mendes de Leon, F. C. (2000). Consequences of Driving Cessation: Decreased Out-of-Home Activity Levels. *Journal of Gerontology: SOCIAL SCIENCES*, S334-S340.

McGrail, M. R., & Humphreys, J. S. (2009). The Index of Rural Access: An Innovative Integrated Approach for Measuring Primary Care Access. *BMC Health Services Research*.

McGrail, M., & Humphreys, J. S. (2009). Measureing Spatial accessability to primary care in rural areas: Improving the effectiveness of the two-step floating catchment area method. *Applied Geography*, 533-541.

Ngui, A. N., & Vanasse, A. (2012). Assessing spatial accessibility to mental health facilities in an urban environment. *Spatial and Saptio-temoral Epidemiology*, 195-203.

O'Brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflaction Factors. *Quality & Quanity*, 673-690.

Province of New Brunswick. (2017). *We are all in this together: An Aging Strategy for New Brunswick.* Fredericton: Province of New Brunswick .

Statistics Canada. (2001, November). *Rural and Small Town Canada Analysis Bulletin.* Retrieved from Statistics Canada: https://www150.statcan.gc.ca/n1/en/pub/21-006-x/21-006-x2001003-eng.pdf?st=qThNOlcy

Statistics Canada. (2010). *The use of Transportation by Seniors in Canada - EnviroStats, Winter 2010, Vol. 4, no. 4* . Retrieved from https://www150.statcan.gc.ca/n1/pub/16-002-x/2010004/part-partie3-eng.htm

Statistics Canada. (2015, 11 27). *Data and Definitions*. Retrieved from Statistics Canda: https://www150.statcan.gc.ca/n1/pub/21-006-x/2008008/section/s2-eng.htm

Statistics Canada. (2018). *Caregivers in Canada 2018.* Statistics Canada.

Statistics Canada. (2019). *Canada's population estimates: Age and sex, July 1, 2018.* Statistics Canada.

Syed, Q., & Holland Jr., N. W. (2016). Should Your Older Patient Be Driving? In *New Directions in Geriatric Medicine* (pp. 91-103). Springer International Publishing Switzerland .

The Canadian Press. (2016, January 11). *Government warns of possible hospital closures without alternate major change like tax hikes*. Retrieved from Global News: https://globalnews.ca/news/2446253/government-warns-of-possible-hospital-closures-without-alternate-major-change-like-tax-hikes/

Transportation Research Board. (2012). *NCHRP Report 716 - Travel Demand Forecasting: Parameters and Techniques.* Washington: National Academy of Sciences.

Vézina, M., & Turcotte, M. (2010). *Caring for a parent who lives far away: The Consequences.* Statistics Canada.

Wang, F., & Luo, W. (2005). Assessing Spatial and Nonspatial Factors for Healthcare Access: Towards an integrated approach to defining health professional shortage areas. *Health and Place*, 131-146.

APPENDIX A: NB-IRDT APPLICATIONS AND DOCUMENTS

A-1: NB-IRDT Project Feasibility Application

developed to accurately quantify current travel demands and forecast for future demands. Understanding these travel demands will lead to our goal of a more accurate understanding of the travel impacts associated with changes to a health care service or facility.

This research is considered to be exploratory research using existing datasets from the NB-IRDT to help quantify the travel demands/ impacts for senior patients which will consist of a volume of trips from various origins, attracted to a specific "zone" or destination. These data sets will include the *Citizen database, Discharge Abstract database and the Physician Billing database*.

Two statistical techniques will be considered to analyze this data including *Descriptive statistics* and *Regression Analysis*.
Descriptive statistics will be used to summarize the data sets and to better describe the information of travel behaviors based on trip purpose and trip generations (production and attractions). This may include determining the distance traveled, mode choice for travel, purpose of travel for healthcare purposes, etc.

By exploring the data sets for significant variables that may be associated with health care trips, *regression analysis* will be used to potentially identify a correlation with the data. Specific variables that will be explored are age, sex, postal code (origin), type of health care trip, destination and disciplines for specialist consultations. These relationships would use past data sets to determine trip generations (production and attractions) which could be used to develop trends that can be used for forecasting opportunities and explanatory purposes.

Using the NB-IRDT data sets, the following variables can be determined:

- **Destinations:** locations of health facilities traveled to can be determined using the facility geo-locations found within the *Physician Billing* database from NB-IRDT.
- **Origins:** Postal codes of the patient will be used to determine their trip origins.
- **Trip Purpose:** Services provided at health care facilities can be determined with The *Physician Billing* database and the *Discharge Abstract* database. These will include the type of institution a particular facility is categorized as, and the type of service provided to the patient.
- **Modes of transportation**: The *Discharge Abstract* database can provide the information of patients arrive to the health facility by an emergency (ambulance) arrival or non-ambulance arrival. For the non-ambulance arrivals, methods such as diversion curves will be used to proportionally estimate mode choice.
- **Category of health visit**: The *Discharge Abstract* database and the *Physician Billing* database will be used to determine the nature of the health visit/ trip purpose to a health care facility.

Other:
Include any additional information beneficial to reviewing your application not captured above
Click to enter other related information here

NB-IRDT
New Brunswick Institute for Research, Data and Training

Project Feasibility Form ***Application Version Date: Feb 05, 2019***

Do not modify contents of this application. Direct questions to NB-IRDTprojects@unb.ca

Project:					
Full project title:	Feasibility of institutional health data for modelling travel demands for health care by rural older adults in New Brunswick				
Requested access to data in secure lab	From:	Mar 31, 2019	To:	Mar 31, 2020	
Project PI:					
First name:	Brianna	Last name:	Morehouse		
Mailing address:	256 Inglewood Dr. Fredericton, NB E3B 2K6				
Email:	b.morehouse@unb.ca	Phone:			
Affiliation:	Academic	If other, describe:			
Name of affiliation:	University of New Brunswick	Position:	Graduate Student		
Accessing data in secure Lab?	Yes	If yes, location of data access:	Fredericton NB-IRDT		
Data Requirements:					
Project data sets:	Only data sets currently available on the NB-IRDT platform				
Funding:					
Funding source:	NSERC Discovery	Award amount:	Ongoing		
Intended Output: (check all that apply)					
☒ Conference or Symposium Presentation/Poster	☒ Peer-reviewed Publication				
☐ Report for government or public sector organization	☒ Academic Requirement (e.g. Thesis)				
☐ Report for non-government agency or group	☐ Edited Volume				
☐ Other (describe):					
Date output anticipated/required	May 01, 2020				

Project Summary:
(Max 500 words) Provide a brief overview of your project in lay language.
Rural areas face unique challenges where patients often travel long distances and require various modes of transportation to meet healthcare needs. Increasing financial pressures to centralize health care services is likely to result in an increase in travel demand for rural older adults (65+), but the degree of this increase is not understood. Understanding the degree to which this population relies on transportation to meet their healthcare needs can inform planners and policy makers on making decisions regarding transportation service. Older patients may not be able to meet their travel needs independently, requiring them to rely on others, such as family or friends. Impacts on transportation demand created by an individual's health care needs are not currently quantified or measured. Volunteer Driver Programs (VDP) in rural New Brunswick have been able to close existing travel supply gaps; however, with an aging population this might not be a sustainable long-term solution. The proposed research has two primary objectives. First, to analyze if the data sets from the New Brunswick Institute for Research, Data and Training (NB-IRDT) can be used for understanding travel demands for health care access. If the data is proven to be feasible, actual travel behaviors of older adults accessing health care services in the province will be quantified and modelled. This will require developing an approach to accurately associate a health care destination with an origin using NB-IRDT health data sets.

These databases have yet to be explored for their applicability to transportation planning, making this potentially a novel approach to quantify and model the travel behavior of older adults using institutional data. Such a model could be used to forecast changes in travel demands associated with demographic changes or changes in facilities or services. Understanding the travel demands from rural communities will allow transportation planners and policy makers to quantify the impacts associated with changes to health care facilities.

Research Goals:

(Max 200 words)

- What are the research goals for this project?
- What is the anticipated policy or clinical impact of your research?
- Who is the intended audience?

The goals of this research are to determine how to use institutional health data in New Brunswick to quantify travel demands associated with healthcare use by older adults (aged 65 years and older) living in rural areas. This will require developing an approach to accurately associate a health care destination with an origin using New Brunswick health data sets from NB-IRDT. In addition, how to forecast changes in travel demands associated with demographic changes and changes in facilities or services.

Rural areas can face a unique challenge where patients often require travelling long distances and require various modes of transportation to meet specific healthcare needs. By understanding these travel demands from rural communities in New Brunswick, the impacts associated with changes to health care facilities can be quantified and would assist with filling a major gap in transportation planning. This opens up opportunities for operational planning for both transportation planning and capital planning purposes.

Methodology:

(Max 500 words) Explain the methodology for this project including:

- How the data requested, and methods employed will be used to answer your research question
- The type of study (retrospective or prospective; RCT, cohort or observation based; etc.)
- Type of analysis (descriptive and/or statistical; cross-sectional or longitudinal; etc.)
- Data: how the data will be used to answer your research question.
- If applicable, description of any derived variables that you be defining
- Disease/Drug/Procedure Classification codes: Specify classification codes you are using in the study (e.g., ICD-10-CA/CCI) and tariff codes for diagnoses, tests, and procedures and Drug Identification Numbers (DINs) for drugs.

The first research question asks how institutional data can be analyzed and used for transportation, more specifically understanding older adult travel demands for health services. It is believed that applicable transportation data (destination, origin, time, frequency, and trip purpose) can be extracted from the institutional data sets to be used for understanding older adult travel demands.

The second research question asks how institutional health data can be used for travel demand modeling. It is believed that a correlation can be determined between variables (e.g. Trips purpose, origin, mode of travel, etc.) associated with travel demands and accessing services; by exploring these relationships, a model can be

A-2: NB-IRDT Data Access Application

Do not modify contents of this application. Direct questions to NB-IRDTprojects@unb.ca
By submitting this application, you are agreeing to all terms and conditions contained within.

SECTION 1: Basic Information

Purpose of Submission:
Application - No prior agreement
If purpose is other than "Application" provide a summary of the reason for amendment:
n/a

Amendments level may be modified at the discretion of the NB-IRDT Director.
Direct questions regarding amendments to NB-IRDTprojects@unb.ca

Project:					
Full project title:	Feasibility of institutional health data for modelling travel demands for health care by rural older adults in New Brunswick				
Requested access to data in secure lab	From:	Jun 01, 2019	To:	Mar 31, 2020	
Project PI:					
First name:	Trevor	Last name:	Hanson		
Mailing address:	Gillian Hall D129				
Email:	thanson@unb.ca	Phone:	506-452-6372		
Affiliation:	Academic	If other, describe:			
Name of affiliation:	University of New Brunswick	Position:	Assistant Professor		
Accessing data in secure Lab?	Yes	If yes, location of data access:	Fredericton NB-IRDT		
Data Requirements:					
Project data sets:	Only data sets currently available on the NB-IRDT platform				
Funding:					
Funding source:	NSERC Discovery	Award amount:	Ongoing		
Intended Output: (check all that apply)					
☒ Conference or Symposium Presentation/Poster		☒ Peer-reviewed Publication			
☐ Report for government or public sector organization		☒ Academic Requirement (e.g. Thesis)			
☐ Report for non-government agency or group		☐ Edited Volume			
☐ Other (describe):					
Date output anticipated/required	May 01, 2020				

Members of research team performing analysis portion of project in the secure NB-IRDT lab
Identify all persons who will be accessing the secure facility, their contact information, affiliation, role in the project, justification for access and location of access. Restrict your list to only those who are necessary. All persons named below: • will be incur a data access fee • must complete NB-IRDT privacy and security training (refresher/update required annually) • must sign a confidentiality agreement • must submit a criminal record check

First name:	Brianna	Last name:	Morehouse
Email:	b.morehouse@unb.ca	Phone:	
Affiliation:	Academic	If other, describe:	n/a
Name of affiliation:	University of New Brunswick	Position:	Graduate Student
Location of data access:	Fredericton NB-IRDT	Role in project:	Research Lead
Specific Purpose for Access:	Data Analysis		

First name:	Trevor	Last name:	Hanson
Email:	thanson@unb.ca	Phone:	
Affiliation:	Academic	If other, describe:	n/a
Name of affiliation:	University of New Brunswick	Position:	Assistant Professor
Location of data access:	Fredericton NB-IRDT	Role in project:	Investigator/Collaborator
Specific Purpose for Access:	Data Analysis/ Verification		

First name:		Last name:	
Email:		Phone:	
Affiliation:	Choose an item.	If other, describe:	
Name of affiliation:		Position:	
Location of data access:	Choose an item.	Role in project:	Choose an item.
Specific Purpose for Access:			

First name:		Last name:	
Email:		Phone:	
Affiliation:	Choose an item..	If other, describe:	
Name of affiliation:		Position:	
Location of data access:	Choose an item.	Role in project:	Choose an item.
Specific Purpose for Access:			

Section 2: Resource Requirements and Timelines

NB-IRDT will provide a quote for data access to its facilities for those listed above using the secure lab. NB-IRDT also has a team of highly qualified personnel available on a cost recovery basis to perform duties beyond making the datasets available. Should you require any of these additional supports from NB-IRDT staff please indicate below.

Request for contracted services:
NB-IRDT will provide a quote and timeline for any contract service work based on the following: • Data extraction: building and testing an analytical data set • Data processing: installation, transferring and understanding new data

• Data analysis including code development: from basic descriptives, like frequencies and means to more complex analysis, like hierarchical linear modeling and exploration of interaction effects
Data extraction: building and testing an analytical data set Data processing: installation and transferring data

Section 3: Project Details

Project Summary:
(Max 500 words) Provide a brief overview of your project in lay language.
Rural areas face unique challenges where patients often travel long distances and require various modes of transportation to meet healthcare needs. Increasing financial pressures to centralize health care services is likely to result in an increase in travel demand for rural older adults (65+), but the degree of this increase is not understood. Understanding the degree to which this population relies on transportation to meet their healthcare needs can inform planners and policy makers on making decisions regarding transportation service. Older patients may not be able to meet their travel needs independently, requiring them to rely on others, such as family or friends. Impacts on transportation demand created by an individual's health care needs are not currently quantified or measured. Volunteer Driver Programs (VDP) in rural New Brunswick have been able to close existing travel supply gaps; however, with an aging population this might not be a sustainable long-term solution. The proposed research has two primary objectives. First, to analyze if specific variables within data sets from the New Brunswick Institute for Research, Data and Training (NB-IRDT) can be used for understanding travel demands for health care access. If the data is proven to be feasible, actual travel behaviors of older adults accessing health care services in the province will be quantified and modelled. This will require developing an approach to accurately associate a health care destination with an origin using NB-IRDT health data sets. These databases have yet to be explored for their applicability to transportation planning, making this potentially a novel approach to quantify and model the travel behavior of older adults using institutional data. Such a model could be used to forecast changes in travel demands associated with demographic changes or changes in facilities or services. Understanding the travel demands from rural communities will allow transportation planners and policy makers to quantify the impacts associated with changes to health care facilities.

Research Goals:
(Max 200 words) • What are the research goals for this project? • What is the anticipated policy or clinical impact of your research? • Who is the intended audience?
The goals of this research are to determine how to use institutional health data in New Brunswick to quantify travel demands associated with healthcare use by older adults (aged 65 years and older) living in rural areas. This will require developing an approach to accurately associate a health care destination with an origin using New

Brunswick health data sets from NB-IRDT. In addition, how to forecast changes in travel demands associated with demographic changes and changes in facilities or services.

Rural areas can face a unique challenge where patients often require travelling long distances and require various modes of transportation to meet specific healthcare needs. By understanding these travel demands from rural communities in New Brunswick, the impacts associated with changes to health care facilities can be quantified and would assist with filling a major gap in transportation planning. This opens up opportunities for operational planning for both transportation planning and capital planning purposes.

Methodology:

(Max 500 words) Explain the methodology for this project including:

- How the data requested and methods employed will be used to answer your research question
- The type of study (retrospective or prospective; RCT, cohort or observation based; etc.)
- Type of analysis (descriptive and/or statistical; cross-sectional or longitudinal; etc.)
- Data: how the data will be used to answer your research question.
- If applicable, description of any derived variables that you be defining
- Disease/Drug/Procedure Classification codes: Specify classification codes you are using in the study (e.g., ICD-10-CA/CCI) and tariff codes for diagnoses, tests, and procedures and Drug Identification Numbers (DINs) for drugs.

The first research question asks how institutional data can be analyzed and used for transportation, more specifically understanding older adult travel demands for health services. It is believed that applicable transportation data (destination, origin, time, frequency, and trip purpose) can be extracted from the institutional data sets to be used for understanding older adult travel demands.

The second research question asks how institutional health data can be used for travel demand modeling. It is believed that a correlation can be determined between variables (e.g. Trips purpose, origin, mode of travel, etc.) associated with travel demands and accessing services; by exploring these relationships, a model can be developed to accurately quantify current travel demands and forecast for future demands. Understanding these travel demands will lead to our goal of a more accurate understanding of the travel impacts associated with changes to a health care service or facility.

This research is considered to be exploratory research using existing datasets from the NB-IRDT to help quantify the travel demands/ impacts for senior patients which will consist of a volume of trips from various origins, attracted to a specific "zone" or destination. These data sets will include the *Citizen database, Discharge Abstract database, Physician Billing database* and the *Provider Registry database*.

Two statistical techniques will be considered to analyze this data including *Descriptive statistics* and *Regression Analysis*.

Descriptive statistics will be used to summarize the data sets and to better describe the information of travel behaviors based on trip purpose and trip generations (production and attractions). This may include determining the distance traveled, mode choice for travel, purpose of travel for healthcare purposes, etc.

By exploring the data sets for significant variables that may be associated with health care trips, *regression analysis* will be used to potentially identify a correlation with the data. Specific variables that will be explored are age, sex, postal code (origin), type of health care trip, destination and disciplines for specialist consultations. These relationships would use past data sets to determine trip generations (production and attractions) which could be used to develop trends that can be used for forecasting opportunities and explanatory purposes.

Using the NB-IRDT data sets, the following variables can be determined:

- **Destinations:** locations of health facilities traveled to can be determined using the facility geo-locations found within the *Physician Billing* database and the *Provider Registry* database from NB-IRDT.
- **Origins:** Postal codes of the patient will be used to determine their trip origins using the *Citizen Database*.
- **Trip Purpose:** Services provided at health care facilities can be determined with the *Physician Billing* database, *Provider Registry* database and the *Discharge Abstract* database. These will include the type of institution a particular facility is categorized as, and the type of service provided to the patient.
- **Modes of transportation**: The *Discharge Abstract* database can provide the information of patients arrive to the health facility by an emergency (ambulance) arrival or non-ambulance arrival. For the non-ambulance arrivals, methods such as diversion curves will be used to proportionally estimate mode choice.
- **Category of health visit**: The *Discharge Abstract* database, the *Physician Billing* database and the *Provider Registry* database will be used to determine the nature of the health visit/ trip purpose to a health care facility.

Preliminary Results Disclosure:
If preliminary results disclosure is planned for small group consult or patient/public engagement, include plan below demonstrating • The purpose for this disclosure • How privacy will be maintained • Who will view the data
A possible release of some early results within a technical meeting however, information would be aggregated for trip purposes therefore no one would be identified, and privacy would be maintained.

Geocoding: If this project requires geocoding, use the below table to indicate the type needed and justify the purpose in the space provided below the table.	
Geocode type	Examples
☒ Geographic codes	Rural/Urban, City, Statistics Canada geographic codes (e.g. CD/CSD/DA), Health Zone, longitude, latitude

☒ Distance/Travel	ArcGIS Travel distance/time, proximity to health centres, etc.
☒ Census Neighborhoods	Median Income, Education Profile, Language Profile, Immigrant Profile
☐ Environmental	Air Quality, Green Space, Blue Space
☐ Other	
Geographic codes – data for urban/ rural city, health zones, etc. will be used for understanding geographical placement of users and health facilities with a potential for a visual component to the model. *Distance/ Travel* – data for travel distance, time and health facility proximity will be used for modelling and forecasting purposes *Census Neighborhoods* - data for income, education, etc. will also be used for modelling and forecasting purposes	

Section 4: Requested Data Sets

- All data sets for this project must be identified. Details on the below data sets, including available dates, dictionaries and data set info can be found at the below web address
- Projects using data outside of NB-IRDT's current data holdings must complete and submit a Request to Import External Data form found at the below web address
- Each data set listed must be accompanied by a list of variables. A variable list form is available at the below web address.

All forms and data set information are available at www.unb.ca/nb-irdt

NB-IRDT Data Set		Start Date	End Date
☐	ALS Registry	Start date	End date
☐	CCDSS COPD Data	Start date	End date
☐	CCDSS Diabetes Data	Start date	End date
☐	CCDSS Hypertension Data	Start date	End date
☐	CCDSS MS Data	Start date	End date
☐	CCDSS Parkinson's Data	Start date	End date
☒	Citizen Database	Jan 01, 2006	Dec 31, 2011
☒	Discharge Abstract Data	Jan 01, 2006	Dec 31, 2011
☐	Healthy Toddler Assessment	Start date	End date
☐	NB Breast Cancer Screening	Start date	End date
☐	NB Cancer Registry	Start date	End date
☐	NB Trauma Data	Start date	End date
☐	NRS Data	Start date	End date
☒	Physician Billing	Jan 01, 2006	Dec 31, 2011
☒	Provider Registry	Jan 01, 2006	Dec 31, 2011
☐	Vital Statistics	Start date	End date
☐	Enter Dataset name if not found above.	Start date	End date
☐	Enter Dataset name if not found above.	Start date	End date

Other:
Include any additional information beneficial to reviewing your application not captured above
Click to enter other related information here

ROLES AND RESPONSIBILITIES:

Principal Investigator (PI):

- There may be only one PI.
- Data Access Agreements must be signed by the PI.
- All amendment requests to the approved project must be made by the PI.
- Responsible to oversee the data analysis of the research project and to ensure the protection of Personal Information and/or Personal Health Information disclosed to the research team as required by NB legislation and NB-IRDT policies and procedures.

Research Lead: (If different than PI)

- Responsible to oversee the analysis for the project and ensure the Personal Information and the Personal Health Information disclosed to the research team is protected as required by NB legislation and NB-IRDT policies and procedures.

Research Team:

- Responsible to familiarize themselves with NB-IRDT's policies available at www.unb.ca/nb-irdt and undertake to abide by them.

TERMS OF THIS APPLICATION:

All applications for access to data in the custody of the New Brunswick Institute for Research, Data and Training (NB-IRDT), and related projects must comply with the NB Personal Health Information Privacy and Access Act, SNB 2009, c P-7.05 (PHIPAA), the NB Right to Information and Protection of Privacy Act, SNB 2009, c R-10.6 (RTIPPA) and other provincial and federal legislation pertaining to the protection of Personal Information and Personal Health Information as well as NB-IRDT policies and procedures.

The Applicant understands and agrees:

- by submitting this application the Applicant warrants the truth and accuracy of the information contained herein and, further, acknowledges that NB-IRDT will be relying on the information contained herein in assessing the application.
- all requests for access to data shall be reviewed carefully with regard to compliance with the PHIPAA, RTIPPA, other relevant provincial and federal legislation and NB-IRDT policies and procedures, and that the application may be denied by NB-IRDT for any reason.
- to undertake to advise NB-IRDT of any significant change to the information contained herein should any such change take place while the application is being considered by NB-IRDT or, if the application is granted, during the life of the project.
- the information provided herein shall form the basis for a data access agreement between the Principal Investigator and NB-IRDT that must be signed by the appropriate parties should this application be granted.

- if this application is granted, project team members accessing the secure facility must provide a Criminal Record Check, participate in privacy and security training sessions and sign a confidentiality agreement prior to data access. Access to the secure facility requires agreement to abide by all relevant rules and regulations of NB-IRDT and UNB.
- NB-IRDT reserves the right to require further information from the Applicant in considering the application or to specify further conditions on the provision of the data requested herein.
- if provided, NB-IRDT data is to be used solely for the purpose identified in this application.
- linking data sets is limited only to what has been outlined in the application, approved by NB-IRDT for access, and covered in the Data Access Agreement signed between the parties. Any additional linkage requests will require additional review by NB-IRDT and may require a full review if it constitutes a major revision to the original approved application.
- that in the event that data matching is intended, this process will be controlled entirely by NB-IRDT staff. Individuals will not be permitted to match their own data under any circumstances.
- all current or future contemplated uses and or disclosures of data results must be identified either in this application or in a future communication and receive NB-IRDT's written consent before such uses or disclosures occur. Such request may require the applicant to submit a revised application to UNB REB. NB-IRDT reserves the right to withhold its approval of any such request for any reason.

By submitting this application, you are agreeing that you have read and understood these terms as set out above and agree to be bound by such terms.

A-3: NB-IRDT Request for Data List of Variables

SPECIALTY_MINOR_GRP_CUR	Provider current specialty minor group code	To determine type of service being offered for purpose of trip
SPECIALTY_MAJOR_GRP_CUR	Provider current specialty major group code	To determine type of service being offered for purpose of trip
SPECIALTY_TYP_CD_CUR	Service Provider current specialty type code	To determine type of service being offered for purpose of trip
ESTABL_ID_CUR	Provider current primary establishment identifier	To determine the facility being traveled to
PRMRY_ESTABL_HLTH_REGN	Provider current primary establishment health region	To determine the facility being traveled to within a health region
SCRAM_SPID_NPDB	Service provider scrambled Identifier	To match attributes to services being provided
HEALTH_REGION_PRACT_LOC	Practice health region location code	To determine region of service provided for location
GEO_LOCATION_CODE_PRATC_LOC	Practice geographical location code	To determine location of service provided
POSTAL_CODE_PRACT_LOC	Practice location postal code	To determine location of service provided
V_3_CITY_URBAN_RURAL_IND_PRACT_LOC	Practice V3 urban/ rural location code	To determine urban/ rural distinction
V_7_CITY_URBAN_RURAL_IND_PRACT_LOC	Practice V7 urban/rural location code	To determine urban/ rural distinction
CME_URBAN_RURAL_IND_PRACT_LOC	Practice CME urban/ rural code	To determine urban/ rural distinction
EFF_DATE_PRACT_LOC	Practice location effective date	To determine location of each service

Data Set Title:	Provider Registry	
Requested Years:	2006 – 2011	
Sample Universe:	Service Providers Serving Patients Age 65+	
Dates & Geographic Area:	New Brunswick	
Requested Variable(s)	**Definition**	**Rationale for Access**
SCRAM_SPID_NPDB	Service scrambled identifier	To match attributes to services being provided
SERV_PRVD_TYPE_CODE	Service provider type code	To determine type of service provided for travel purposes
HLTH_REGION_CUR	Service provider current health region	To determine region of service provided for location
CNTY_CD_CUR	Service provider county location	To determine county of service provided
GEO_LOCATION_CD_CUR	Service county of location code	To determine location of service provided
V3_CITY_URB_RUR_IND_CUR	Service current V3 city rural/urban indicator	To determine urban/ rural distinction
V7_CITY_URB_RUR_IND_CUR	Service current V7 city rural/urban indicator	To determine urban/ rural distinction
CME_URBAN_RUR_IND_CUR	Service CME rural/urban classification	To determine urban/ rural distinction
SPECIALTY_TYPE_CUR	Service provider current specialty type	To determine type of service being offered for purpose of trip
MAIN_SPECIALY_CD_CUR	Service provider main specialty code	To determine type of service being offered for purpose of trip
SPECIALTY_GRP_CUR	Service provider current specialty group code	To determine type of service being offered for purpose of trip

Dates & Geographic Area:	New Brunswick	
Requested Variable(s)	**Definition**	**Rationale for Access**
SCRAM_ID	tba	Make a connection between a type of physician to a service being received
Scrambled Physician ID	A random number that uniquely identifies a service provider	To determine type of service being received
SERVICE_TYPE_CODE	Type of service being identified	To determine type of service being received
SERVICE_TYPE_GROUP_CODE	Type of grouping for service	To determine type of service group being received
SERVICE_CATEGORY_CODE	Identified broad grouping of service	To determine a category of service being received
SERVICE_CODE_CLAIMED	tba	To determine type of service being received
SERVICE_CLAIMED_NUMERIC	Service provider's primary specialty (numeric)	To determine type of service being received
SERVICE_CLAIMED_DESC	Service provider's primary specialty (description)	To determine type of service being received
SPECIALTY_TYPE_CODE	Code to identify current specialty	To determine which specialty of care is being received
SERVICE_DATE	Service delivery date	To determine when the service was received
ESTABLISHMENT_ID	tba	To determine what location a service was provided
FOR OFFICE USE ONLY – Variables used to link, validate and for calculations only (variables that will not visible to researcher)		
Data Set Title:		
Years:		
Sample Universe:		
Dates & Geographic Area:		
Variable(s)	**Definition**	**Purpose**
		Choose an item

Intervention 11 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 12 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 13 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 14 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 15 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 16 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 17 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 18 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 19 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 20 code	See "Intervention 1 code"	To determine service associated with provider.
FOR OFFICE USE ONLY – Variables used to link, validate and for calculations only (variables that will not visible to researcher)		
Data Set Title:		
Years:		
Sample Universe:		
Dates & Geographic Area:		
Variable(s)	**Definition**	**Purpose**
		Choose an item

Data Set Title:	Physician Billing
Requested Years:	2006 – 2011
Sample Universe:	Age 65+

Intervention 13 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 14 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 15 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 16 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 17 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 18 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 19 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 20 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 1 code	Code which identifies an intervention that was performed during the patient's hospital/institution stay	To determine service associated with provider.
Intervention 2 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 3 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 4 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 5 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 6 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 7 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 8 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 9 code	See "Intervention 1 code"	To determine service associated with provider.
Intervention 10 code	See "Intervention 1 code"	To determine service associated with provider.

Intervention 14 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 15 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 16 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 17 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 18 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 19 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 20 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 1 provider service	A number used to identify the principal service associated with the intervention provider	To determine service associated with provider.
Intervention 2 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 3 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 4 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 5 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 6 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 7 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 8 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 9 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 10 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 11 provider service	See "Intervention 1 provider service"	To determine service associated with provider.
Intervention 12 provider service	See "Intervention 1 provider service"	To determine service associated with provider.

Provider 5 Type	See "Provider 2 Type"	To determine purpose of health trip
Provider 6 Type	See "Provider 2 Type"	To determine purpose of health trip
Provider 7 Type	See "Provider 2 Type"	To determine purpose of health trip
Provider 8 Type	See "Provider 2 Type"	To determine purpose of health trip
Provider 1 Service – Provider 8 Service	Identifies the training/ specialty of the provider responsible for the patient	To determine purpose of health trip – Provider's specialty
Intervention 1 provider number	Scrambled facility assigned number of the principal provider associated with the intervention performed	To determine service provider.
Intervention 2 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 3 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 4 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 5 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 6 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 7 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 8 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 9 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 10 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 11 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 12 provider number	See "Intervention 1 provider number"	To determine service provider.
Intervention 13 provider number	See "Intervention 1 provider number"	To determine service provider.

	patient's care during hospitalization	
Provider 2 Number	Scrambled facility assigned number of a provider who was responsible for the patient's care during hospitalization	To determine type of service and where service is being provided
Provider 3 Number	See "Provider 2 Number"	To determine type of service and where service is being provided
Provider 4 Number	See "Provider 2 Number"	To determine type of service and where service is being provided
Provider 5 Number	See "Provider 2 Number"	To determine type of service and where service is being provided
Provider 6 Number	See "Provider 2 Number"	To determine type of service and where service is being provided
Provider 7 Number	See "Provider 2 Number"	To determine type of service and where service is being provided
Provider 8 Number	See "Provider 2 Number"	To determine type of service and where service is being provided
Provider 1 Type	A code which identifies the role of the provider responsible for the care of the patient during hospitalization	To determine purpose of health trip
Provider 2 Type	A code which identifies the role of the provider responsible for the care of the patient during hospitalization	To determine purpose of health trip
Provider 3 Type	See "Provider 2 Type"	To determine purpose of health trip
Provider 4 Type	See "Provider 2 Type"	To determine purpose of health trip

Diagnosis 14 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 15 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 16 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 17 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 18 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 19 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 20 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 21 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 22 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 23 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 24 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 25 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Provider 1 Number	Scrambled facility assigned number of the provider who was most responsible for the	To determine type of service and where service is being provided

	hospitalization (most responsible)	
Diagnosis 2 Type	A Code which indicates the relationship of the diagnosis to the patient's hospitalization	To help determine trip purpose and location of service
Diagnosis 3 Type	See "Diagnosis 1 Type"	To help determine trip purpose and location of service
Diagnosis 4 Type	See "Diagnosis 1 Type"	To help determine trip purpose and location of service
Diagnosis 5 Type	See "Diagnosis 1 Type"	To help determine trip purpose and location of service
Diagnosis 6 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 7 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 8 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 9 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 10 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 11 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 12 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service
Diagnosis 13 Type	See "Diagnosis 2 Type"	To help determine trip purpose and location of service

Diagnosis 5 Code	See "Diagnosis 1 Code"	To determine trip purpose
Diagnosis 6 Code	See "Diagnosis 1 Code"	To determine trip purpose
Diagnosis 7 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 8 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 9 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 10 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 11 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 12 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 13 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 14 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 15 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 16 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 17 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 18 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 19 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 20 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 21 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 22 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 23 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 24 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 25 Code	See "Diagnosis 2 Code"	To determine trip purpose
Diagnosis 1 Type	Code which indicates the relationship of the diagnosis to the patient's	To help determine trip purpose and location of service

Admission Category	Patient Classification on admission to a health care facility	To determine purpose of trip.
Admission Date	Calendar date that the patient was officially registered	To select appropriate data samples
Institution from Type	Identifies health facility or level of care within reporting facility	Determine Type of care traveling for.
Ambulance Code	Code indicates whether a patient arrives via ambulance and type of ambulance	Determine travel mode for emergencies.
Entry Code	Code which indicates the point of entry to the health care facility	To determine trip purpose and destination.
Age Group	CIHI assigned Age Group for admission	To determine age group.
Age Code	Code indicates how the patient's age is recorded	To determine age of patient
Age Unit	A derived variable that represents how old the patients is at the time of admission – calculated using birth date	To determine age group of patient
Institution Type	Indicated level of care of the Facility	Determine level of care facility for travel.
Total Length of Stay	Total number of days the patient was hospitalized	To determine length of travel impact/ durations.
Diagnosis 1 Code	ICD code which identifies the condition considered to be most responsible for patient's hospitalization	To determine primary reason for trip (purpose).
Diagnosis 2 Code	ICD code which identifies the condition which contributed to the patient's hospitalization	To determine trip purpose
Diagnosis 3 Code	See "Diagnosis 1 Code"	To determine trip purpose
Diagnosis 4 Code	See "Diagnosis 1 Code"	To determine trip purpose

FOR OFFICE USE ONLY – Variables used to link, validate and for calculations only (variables that will not visible to researcher)		
Data Set Title:		
Years:		
Sample Universe:		
Dates & Geographic Area:		
Variable(s)	**Definition**	**Purpose**
		Choose an item

Data Set Title:	Discharge Abstract Database	
Requested Years:	2006 – 2011	
Sample Universe:	Age 65+	
Dates & Geographic Area:	New Brunswick	
Requested Variable(s)	**Definition**	**Rationale for Access**
Province	Province the Facility is located	To analyze facilities in NB.
Institution Number	Identifying the Facility and level of care to a reporting facility	To identify facility and Level of care facility covers.
Scrambled ID	Scrambled Patient ID	To match patient with medical travel attributes.
Residence Code	Identifies patient's area of residence	Identify NB Resident.
Postal code	Postal Area	To determine Origin.
Gender	Sex of the individual	To determine a pattern with a gender variable.
Admission Time	Time patient was registered (24-hour clock)	To determine length of stay for travel impacts.

NB-IRDT
New Brunswick Institute for Research, Data and Training

Request for Data – List of Variables

Using a separate table for each data set, complete this form to list all variables required for your research. *This includes each data set outlined in Section 7 of the* ***Data Access Application*** *and each data set outlined in the* ***Request to Import External Data****.*

Data Set Title:	Citizen Database
Requested Years:	2006 – 2011
Sample Universe:	Ages 65+
Dates & Geographic Area:	New Brunswick

Requested Variable(s)	Definition	Rationale for Access
SCRAM_ID	Uniquely Identifies Individual	To connect an individual's travel behavior with other attributes
GENDER	Sex of the individual	To determine a pattern with a gender variable
YEAR_OF_BIRTH	Year of birth	To determine age
PROV	Province of residence	To confirm origin of individual from New Brunswick
POSTAL_CODE	Individual Postal Code	To determine origin
PROV_STATE_CODE	Individual province of residence code	To confirm origin of individual from New Brunswick
HEALTH_REGION	Individual Health Region Code	To determine the health region the individual is located
GEO_LOCATION_CODE	Individual geographical location code	To determine the geographic location
EFF_DATE	Date which the address information of individual became effective	To determine the exact location at a specific service
TERM_DATE	Date which the address information of individual was terminated or no longer active.	To determine the exact location at a specific service

TERM_DATE_PRACT_LOC	Practice location termination date	To determine the location of each service
SPECIALTY_CODE_HISTORY	Provider specialty history code	To determine a provider's specialty at service
SPECIALTY_PRIORITY_HISTORY	Provider specialty priority history code	To determine a provider's specialty at service
SPECIALTY_GROUP_CODE_HISTORY	Provider specialty group history code	To determine a provider's specialty at service
EFF_DATE_HISTORY	Provider effective date history	To determine a provider's specialty at service
TERM_DATE_HISTORY	Provider termination date history	To determine a provider's specialty at service

FOR OFFICE USE ONLY – Variables used to link, validate and for calculations only (variables that will not visible to researcher)		
Data Set Title:		
Years:		
Sample Universe:		
Dates & Geographic Area:		
Variable(s)	**Definition**	**Purpose**
		Choose an item

A-4: Research Ethics Board Application for Review of Research Involving Humans

FOR REB USE ONLY: **File No:** _____
Date Complete App'n. Rec'd: _____
Approved/Approved with: _____
Modification/Rejected: _____
Date: _____

University of New Brunswick
Application for Review of Research Involving Humans

Principal Investigator(s): Name(s); Academic Status (Faculty, Undergraduate Student or Graduate Student); Academic Unit, e-mail Address, Office Telephone, Home Telephone:
Brianna Morehouse; Graduate Student, Civil Engineering, b.morehouse@unb.ca,

Title of Proposed Research:

Feasibility of institutional health data for modelling travel demands for health care by rural older adults in New Brunswick

Commencement Date: September 2018 Completion Date: May 2020

Co-Investigator(s): Academic Unit, e-mail address, Office Telephone

Supervisor(s) (if Principal Investigator is a student); Academic Unit, e-mail address, Office Telephone:
Trevor Hanson; Faculty; Civil Engineering, thanson@unb.ca, 506-452-6374

The undersigned parties certify that they have read, and undertake to comply fully with, the Tri-Council Policy Statement "Ethical Conduct for Research Involving Humans." (*Original signatures required*)

Principal Investigator(s): _____

Co-Investigator(s): _____

Supervisor(s): _____

The undersigned certifies that the proposed research has been reviewed by, and is acceptable *in all respects* to, the academic unit(s) responsible. (*Original signature required*)

Dean/Director/Chair(s) Typed Name _____Eric Hildebrand_____

Signature _____

Date submitted to the REB: August 6, 2019

1. Summary: Provide here, in approximately 300 words, a summary of the proposed research, indicating clearly the role of the research subjects and any procedures to which they will be subjected.

Rural areas face unique challenges where patients often travel long distances and require various modes of transportation to meet healthcare needs. Increasing financial pressures to centralize health care services is likely to result in an increase in travel demand for rural older adults (65+), but the degree of this increase is not understood. Understanding the degree to which this population relies on transportation to meet their healthcare needs can inform planners and policy makers on making decisions regarding transportation service.

Older patients may not be able to meet their travel needs independently, requiring them to rely on others, such as family or friends. Impacts on transportation demand created by an individual's health care needs are not currently quantified or measured. Volunteer Driver Programs (VDP) in rural New Brunswick have been able to close existing travel supply gaps; however, with an aging population this might not be a sustainable long-term solution.

The proposed research has two primary objectives. First, to analyze if specific variables within data sets from the New Brunswick Institute for Research, Data and Training (NB-IRDT) can be used for understanding travel demands for health care access. If the data is proven to be feasible, actual travel behaviors of older adults accessing health care services in the province will be quantified and modelled. This will require developing an approach to accurately associate a health care destination with an origin using NB-IRDT health data sets.

These databases have yet to be explored for their applicability to transportation planning, making this potentially a novel approach to quantify and model the travel behavior of older adults using institutional data. Such a model could be used to forecast changes in travel demands associated with demographic changes or changes in facilities or services. Understanding the travel demands from rural communities will allow transportation planners and policy makers to quantify the impacts associated with changes to health care facilities.

2. Risk: In your opinion, does this research pose more than minimal risk (Tri-Council Policy, Chapter 2, Section B) to participating subjects? *Yes* *No* ✓

If yes, provide here a statement that describes in detail the aspects of the research procedure that pose a risk to subjects, and provide your assessment of the risk of harm (probability and severity). Note that not only physical injury but also anxiety or embarrassments are included in the concept of harm. Describe means adopted to minimize risk, and means (such as provision of counseling) to deal with harms, which subjects may experience. Describe as well the potential benefit, which will result from this research, which justifies the above risk of harm.

n/a – secondary data

3. Deception: Does this research involve deception or partial disclosure? *Yes* *No* ✓

If yes, refer to the Tri-Council Policy, Chapter 3, specifically Article 3.7 and subsequent commentary, and provide here an explanation of how you plan to comply with the requirements of that Section for debriefing. Describe as well the potential benefit, which will result from this research, which justifies waiving the normal requirements for full disclosure.

n/a – secondary data

4. Funding: Has funding been received for this research? ***Yes ✓No***

If yes, ***from*** what agency and for what period?

NSERC Discovery Funding; September 2018 – May 2020

If yes, ***from*** what agency and for what period?

5. Research Subjects:

5.1 Number of Subjects: How many subjects will participate in this research? n/a

5.2 Recruitment: How will they be recruited, and from what population?
n/a – secondary data

Using specific variables within approved data sets from NB-IRDT, travel information involving New Brunswick citizens ages 65 + will be used for research purposes.

6. Informed Consent:

6.1 Informing Subjects: How will the nature of the research be explained to potential subjects, in compliance with Chapter 3, specifically article 3.2 and the subsequent commentary of the Tri-Council Policy? Attach a copy of any document(s), such as an explanatory letter, to be used for this purpose.

n/a – secondary data

6.2 Consent: If written evidence of informed consent will be obtained, attach a copy of the consent form. (See Requirements for Informed Consent Forms.) **If written evidence of informed consent will not be used, explain here, in detail, how you intend to comply with the requirements of Chapter 3, Article 3.12 and the subsequent commentary of the Tri-Council Policy**:

n/a – secondary data

6.3 Children as Research Subjects: If the proposed research involves children as subjects, provide here a statement indicating how compliance with Chapter 3, Section C, and specifically with Articles 3.9 and 3.10 of the Tri-Council Policy, will be achieved.

n/a – secondary data & no children, (ages 65+ being requested)

6.4 Incompetent Adults as Research Subjects: If the research involves adults of diminished competence as subjects, provide a statement indicating how compliance with Chapter 3, Section C, and specifically with Articles 3.9 and 3.10 of the Tri-Council Policy, will be achieved.

n/a – secondary data

7. Inducements: Will any inducements (money, grade points, etc.) be offered to encourage participation? *Yes* *No* ✓

If yes, indicate here how compliance with Chapter 3, Article 3.1 and the subsequent commentary of the Tri-Council Policy (concerning voluntariness) will be achieved. If academic rewards are to be used, give details of alternative means of achieving equivalent rewards.

n/a – secondary data, no inducements

8. Private Information: Does the proposed research involve accessing identifiable personal information about subjects by means of surveys, questionnaires, etc.? *Yes* ✓ *No*

If yes, indicate here, in detail, how you propose to meet the requirements of the Tri-Council Policy, Chapter 5, Section B and C. A copy of any questionnaire, survey document or interview schedule to be used should be attached as well.

No Identifiable information related to surveys, questionnaires, etc. being used within research. However, accessing personal information obtained through approved data variables within NB-IRDT Data Sets.

All data being accessed is secondary data and has already been made pseudonymous with scrambled patient ID numbers and scrambled service provider numbers by the institution.

All secondary data, therefore, no relationship with patients or any requirement to disclose information for purposes such as abuse, etc.

Data to be used at an aggregate level to avoid identifying or linking individuals.

9. Feedback: Describe the measures, which you propose for providing feedback to research subjects concerning the outcome of the research.

n/a – secondary data

10. Data Security: Describe the measures, which you propose for ensuring the security of any identifiable personal data, which will be retained after completion of the research.
Research maintaining an aggregate approach ensuring identifiable information be secured and non-identifiable for analysis and results.

Unique ID numbers for individuals and service providers are scrambled by NB-IRDT to maintain privacy. Postal codes of individuals are used to determine travel distances prior the release of all specific variables requested. After distances are calculated, postal codes are removed from accessible data sets to ensure no risk of individuals being identified due to origin. When postal codes are removed, remainder of data variables will be provided for research.

Data set variables are only accessed within the NB-IRDT lab for research purposes. After completion of the project, data variables stay within institution and a results summary will be provided in an acceptable format to include within Masters' thesis report.

Data sets and results from this research project will remain within the Institution for data security and safeguarding purposes.

Overall, the research project is requesting secondary data from the Institution and will be following the NB-IRDT procedures for safeguarding information and maintaining individual's privacy.

11. Continuing Review: All research requires brief annual reports and a brief report upon completion of the research. Suitable report forms are included at the end of this file. **Research involving more than minimal risk may require additional measures for continuing review.** If your research involves more than minimal risk, describe here the measures you propose for facilitating continuing review of this research, in compliance with Article 2.8 of the Tri-Council Policy.

n/a

12. Additional Information: Please feel free to append any additional information, which you feel may be helpful to the REB in evaluating this application.

n/a

A-5: Research Ethics Board Acceptance Letter

REB 2019-109

R. Steven Turner
Mon 2019-08-12 2:28 PM

To: Brianna Morehouse <b.morehouse@unb.ca>
Cc: Ethics, Department of <ethics@unb.ca>; Trevor Hanson <thanson@unb.ca>; Jennifer Moorcraft <Jennifer.Moorcraft@unb.ca>; NB-IRDT Project Coordination <nb-irdtprojects@unb.ca>

Ms Brianna Morehouse
Civil Engineering
Dear Ms Morehouse,

As Chair of the Research Ethics Board (REB) at the University of New Brunswick in Fredericton, I have reviewed your ethics application for the project entitled "Feasibility of institutional health data for modelling travel demands for health care by rural older adults in New Brunswick," which has been assigned the file number REB #2019-109. On the basis of the review, I consider your project to be eligible for delegated review, since any risk to participants that might exist appears not to exceed the "minimal risk" outlined in the Tri-Council Policy Statement, 2nd edition (TCPS2). I am also pleased to inform you that, in my opinion, your project is in compliance with TCPS2 and the University Policy on Research Involving Humans (UPRIH). Accordingly, please consider this E-mail to represent official notification of REB approval of your project for a period of three years from the date of this E-Mail.

This project involves the secondary analysis of data on New Brunswick residents held in the NBHC and in various administrative databases (Citizen Database, Hospital Discharge Abstract Data, Physician Billing, Provider Registry), and the latter databases are being accessed through the NB-IRDT as the legal data custodian. The REB has received from the NB-IRDT their letter of approval and endorsement as well as the list of variables to be accessed. I find that the information to be accessed potentially allows individuals to be identified. I believe that the project satisfies the six requirements laid out in Article 5.5 of the TCPS2 (2014) for secondary use of identifiable information without consent of the original contributors, and am extending REB approval under Article 5.5.

Please note that, in the future, if you find that you must make any changes to your protocol, those changes must be considered and approved by the REB before they are implemented. Please submit the REB Case Modification Request form, available online through the Research Ethics page of the Office of the VP (Research).

In the event that you have not already done so, please take a moment to send, by e-mail, a copy of your summary (your answer to question # 1 of the ethics application form) at your earliest convenience. It should be sent to ethics@unb.ca. I thank you in advance for your co-operation in this matter.

Best wishes for the successful completion of your research project.

Steven Turner, Chair
UNBF Research Ethics Board

APPENDIX B: SAS SYNTAX

B1: Syntax for Processing Data and Calculating Travel Distances

```
/* Create Libraries to save and Source Data */
libname moreh "K:\p0036_travel_demand\users\Brianna
Morehouse";
libname bmoreh "K:\p0036_travel_demand\users\Brianna
Morehouse\bmoreh";
libname source "K:\p0036_travel_demand\src";
libname arcmap "K:\p0036_travel_demand\users\Brianna
Morehouse\ArcMap";

/* PRE-WORK BEFORE DATA PROCESSING STEPS */

/* STEP 1: Create data set named 'addr_var4dist' in bmoreh
library.
Using data set 'addr_var4dist' from source library.
Extract a substring from argument (postal_code) at position
(1) for n characters (1).
This tells SAS to show only NB postal codes (All postal codes
beginning with 'E').
Rename variables to match variables in PCCF+ file: postal_code
= pcode, indivd = id */

Data bmoreh.addr_var4dist;
Set source.addr_var4dist;
where substr(postal_code,1,1)='E';
rename
postal_code = pcode
indvid=id;
run;

/* STEP 2: Final_hlthout was created in moreh library using
the PCCF+ file.
To avoid overwriting original file, create data set named
'Final_hlthout' in bmoreh library using data set
'Final_hltout' from moreh library.
Set length of 'indvid' to '9' characters to match the length
of 'indvid' characters in 'addr_var4dist' data set. */

data bmoreh.Final_hlthout;
set moreh.Final_hlthout;
length indvid $9.;
indvid=compress(id);
run;

/* STEP 3: Make sure there is no missing data in addr_var4dist
(civ. dataset.
Some of the 'term_date' values are missing from dataset
add_var4dist and are replaced with a '.'.
```

```
To assign a 'term_date', assume no there is not yet a
'term_date' therefore, set variable 'term_date' to present
date December 31, 2019'
(term_date = mdy(12,13,2019);) */

Data bmoreh.addr_var4dist_2;
Set bmoreh.addr_var4dist;
IF TERM_DATE = . THEN TERM_DATE = mdy(12,31,2019);
run;
/* Sort addr_var4dist_2 Data by ID to later merge.*/

PROC SORT
DATA = bmoreh.addr_var4dist_2;
BY ID;
Run;



/* STEP 4: Formatting the DAD Datasets:
1. Change 'Admdate' in the DAD dataset to the be the same date
informat (YYMMDD10.) as the 'eff_date' and 'term_date'
variables in the addr_var4dist_2 data set,
This will allow the IF AND THEN Statments to be able to read
the dates properly.
2. Rename 'indvid' variable in DAD data sets to 'ID'.
Note: This will allow SAS to read the 'ID' variable in the DAD
data base the same way as it does in the addr_var4dist_2
database and final_hlthout. */

/* DAD data 2005 */
Data bmoreh.dad2005_var4dist;
Set Source.dad2005_var4dist;
NewAdmdate = Input(Admdate,YYMMDD10.);
FORMAT NewAdmdate YYMMDD10.;
Rename indvid = ID;
run;

/* DAD data 2006 */
Data bmoreh.dad2006_var4dist;
Set Source.dad2006_var4dist;
NewAdmdate = Input(Admdate,YYMMDD10.);
FORMAT NewAdmdate YYMMDD10.;
Rename indvid = ID;
run;

/* DAD data 2007 */
Data bmoreh.dad2007_var4dist;
Set Source.dad2007_var4dist;
NewAdmdate = Input(Admdate,YYMMDD10.);
```

```
FORMAT NewAdmdate YYMMDD10.;
Rename indvid = ID;
run;

/* DAD data 2008 */
Data bmoreh.dad2008_var4dist;
Set Source.dad2008_var4dist;
NewAdmdate = Input(Admdate,YYMMDD10.);
FORMAT NewAdmdate YYMMDD10.;
Rename indvid = ID;
run;

/* DAD data 2009 */
Data bmoreh.dad2009_var4dist;
Set Source.dad2009_var4dist;
NewAdmdate = Input(Admdate,YYMMDD10.);
FORMAT NewAdmdate YYMMDD10.;
Rename indvid = ID;
run;

/* DAD data 2010 */
Data bmoreh.dad2010_var4dist;
Set Source.dad2010_var4dist;
NewAdmdate = Input(Admdate,YYMMDD10.);
FORMAT NewAdmdate YYMMDD10.;
Rename indvid = ID;
run;
/* DAD data 2011 */
Data bmoreh.dad2011_var4dist;
Set Source.dad2011_var4dist;
NewAdmdate = Input(Admdate,YYMMDD10.);
FORMAT NewAdmdate YYMMDD10.;
Rename indvid = ID;
run;

/*Sort all DAD datasets (2005-2011) by ID number before
combining into one dataset */

/* Sort 2005 */
PROC SORT
DATA = bmoreh.dad2005_var4dist;
BY ID;
FORMAT NewAdmdate YYMMDD10.;
run;

/* Sort 2006 */
PROC SORT
DATA = bmoreh.dad2006_var4dist;
BY ID;
```

```
FORMAT NewAdmdate YYMMDD10.;
run;

/* Sort 2007 */
PROC SORT
DATA = bmoreh.dad2007_var4dist;
BY ID;
FORMAT NewAdmdate YYMMDD10.;
run;

/* Sort 2008 */
PROC SORT
DATA = bmoreh.dad2008_var4dist;
BY ID;
FORMAT NewAdmdate YYMMDD10.;
run;

/* Sort 2009 */
PROC SORT
DATA = bmoreh.dad2009_var4dist;
BY ID;
FORMAT NewAdmdate YYMMDD10.;
run;

/* Sort 2010 */
PROC SORT
DATA = bmoreh.dad2010_var4dist;
BY ID;
FORMAT NewAdmdate YYMMDD10.;
run;

/* Sort 2011 */
PROC SORT
DATA = bmoreh.dad2011_var4dist;
BY ID;
FORMAT NewAdmdate YYMMDD10.;
run;

/* STEP 5: Create a new dataset table named 'dad2005_2011'
that includes all DAD data set years (2005-2011)
- Do this by "joining" the tables, not merging them
- Sequence number and fiscal year variables were included as
identifiers */

proc sql;
Create table bmoreh.dad2005_2011 as
select ID,Postalcode,Instnum,NewAdmdate,seqn, 2005 as fiscalyr
From bmoreh.dad2005_var4dist
union
```

```
select ID,Postalcode,Instnum,NewAdmdate,seqn, 2006 as fiscalyr
From bmoreh.dad2006_var4dist
union
select ID,Postalcode,Instnum,NewAdmdate,seqn, 2007 as fiscalyr
From bmoreh.dad2007_var4dist
union
select ID,Postalcode,Instnum,NewAdmdate,seqn, 2008 as fiscalyr
From bmoreh.dad2008_var4dist
union
select ID,Postalcode,Instnum,NewAdmdate,seqn, 2009 as fiscalyr
From bmoreh.dad2009_var4dist
union
select ID,Postalcode,Instnum,NewAdmdate,seqn, 2010 as fiscalyr
From bmoreh.dad2010_var4dist
union
select ID,Postalcode,Instnum,NewAdmdate,seqn, 2011 as fiscalyr
From bmoreh.dad2011_var4dist;
run;


/* DATA PROCESSING STEPS TO DETERMINE DISTANCES */

/* STEP 1: "In the DAD dataset, if a postal_code is not
complete or missing - use the addr_var4dist to input missing
postal codes by matching the indvid
IF the following condition is met: admdate>= eff_date and
admdate < term_date"
(This will make sure that the postal code being assigned was
active with that individual during this time)*/

/* A. Combine (merge) the new DAD data set 'dad2005_2011' and
the new civilian data set 'addr_var4dist_2'.
New File Name: 'dadaddr2005_2011'

- This dataset will include all columns and be sorted by ID
number.
- There will be multiple number of 'ID' variable duplicates
due to duplicates from both data sets and multiple visits for
one ID number.
- Also Double checked that 'NB' Postal code is still within
dataset - good.
- By setting data set 'dad2005_2011' equal to 'a' and dataset
'addr_var4dist_2' equal to 'b' we can merge the two data sets
together by ID and
set both = 1.
- This will tell SAS to only include variables that are
included in both data sets 'a' and 'b'*/

data bmoreh.dadaddr2005_2011;
```

```
MERGE bmoreh.dad2005_2011(in=a) bmoreh.addr_var4dist_2(in=b);
by ID;
if a=1 and b=1;
run;

/* B. If a postal code is missing from the DAD data (i.e 'NB')
then have to complete dataset by importing missing postal code
from 'addr_var4dist_2' dataset.
- Replace missing DAD Postal Code ('NB') with pcode from
'addr_var4dist_2' based on the following conditions:
     IF 'ID' numbers are equal AND IF 'EFF_Date' <=
'NewAdmdate' < 'TERM_Date'

Notes:
- 'postalcode' = Postal code in DAD Data
- 'pcode' = Postal code from 'addr_var4dist_2' Data
- 'EFF_Date' = effective date of pcode from addr_var4dist_2
data
- 'NewAdmdate' = Admission Date to institution from DAD data
- 'TERM_Date' = Termination date of pcode from addr_var4dist_2
data

- ge = Greater than/ equal to
- le = less than or equal to
- ne = not equal to */

data bmoreh.org_dadaddr2005_2011;
set bmoreh.dadaddr2005_2011;
by id;
if postalcode ="NB" then
do;
 if NewAdmdate ge eff_date and NewAdmdate le term_date then
do;
 postalcode=pcode;
 success = '1' ;
end;
end;
run;

/* Sort data set 'org_dadaddr2005_2011' */

proc sort
data = bmoreh.org_dadaddr2005_2011;
by id eff_date;
run;
```

```
/* STEP 2: Assign institution postal code to dad data by
matching last 3-char of instnum with establishment_id (in
Establishments)
Note: The 'institution number' (instnum) is made up of 4-char,
the first being the level of care being used at the
institution, the last 3-char
representing which institution is being used.*/

/* A. Create dataset 'establishments' from source library
within bmoreh library to ensure nothing happens to original
dataset. */

Data bmoreh.establishments;
Set Source.establishments;
run;

/* B. Create a new table to remove variables that are not
needed.
New table includes:
- 'Establishment_Id'
- 'Postal Code'
- 'City'
- 'External_Estbl_Id'

Notes:
- Rename 'Establishment_Id' to 'est_id' to match 'est_id'
Variable in 'org_dadaddr2005_2011' data set
- Rename Postal_Code to est_Postal_code to be able to
determine between establishment and individual postal codes
when merging tables. */

data bmoreh.new_estab;
Set bmoreh.establishments (KEEP = Establishment_Id
External_Estbl_Id Postal_Code City);
rename Establishment_Id = est_id;
rename Postal_Code = est_pcode;
rename External_Estbl_Id = ext_est_id;
run;

/* Sort by est_id */
Proc sort
data = bmoreh.new_estab;
by est_id;
run;

/* C. Create a new variable called 'est_id' (Establishment_Id)
in the 'org2_dadaddr2005_2011' data set using the
'org_dadaddr2005_2011' data set
from STEP 1.
```

```
- Tell SAS to take the last 3-char from the Institutional ID
Variable (Instnum) with the SUBSTR command.
- Convert the variable from character to numeric by the INPUT
Statement - because In 'new_estab' data set, variable 'est_id'
is in Numeric format. */

Data bmoreh.org2_dadaddr2005_2011;
set bmoreh.org_dadaddr2005_2011;
est_id = INPUT(SUBSTR(Instnum,2,3),4.);
run;

/* D. Combine data set 'org2_dadaddr2005_2011' and dataset
'new_estab' by 'ID variable.
- This imports the 'est_pcode' variable to the new data set to
assign the est_id a postal code (destination)
Note: Prior to data step, must sort data by same variable
(est_id) */

Proc sort
data = bmoreh.org2_dadaddr2005_2011;
by est_id;
run;

data bmoreh.Final_dist4DAD;
MERGE bmoreh.org2_dadaddr2005_2011 bmoreh.new_estab;
by est_id;
run;


/* STEP 3: Combine patient’s postal codes from
'addr_var4dist_2' with the billing dataset BY 'ID' of Patient
IF the following condition is true:
     Condition: addr_var4dist.eff_date <=
billxxxx_var4dist.servdate < addr_var4dist.term_date
(Service date from physician billing falls between the
EFF_Date and TERM_date of postal codes from the Citizen data
set)*/

/* A. Create new table with all Billing data years (2006-
2011).
- Change Variable name 'indvid' to ID for merging purposes.
- Establishment_Id changed to est_id to match other datasets
later. */

/* Billing data 2006 */
Data bmoreh.bill2006_var4dist;
Set Source.bill2006_var4dist;
Rename indvid = ID;
```

```
Rename EstablishmentId = ext_est_id;
run;

/* Billing data 2007 */
Data bmoreh.bill2007_var4dist;
Set Source.bill2007_var4dist;
Rename indvid = ID;
Rename EstablishmentId = ext_est_id;
run;
/* Billing data 2008 */
Data bmoreh.bill2008_var4dist;
Set Source.bill2008_var4dist;
Rename indvid = ID;
Rename EstablishmentId = ext_est_id;
run;

/* Billing data 2009 */
Data bmoreh.bill2009_var4dist;
Set Source.bill2009_var4dist;
Rename indvid = ID;
Rename EstablishmentId = ext_est_id;
run;

/* Billing data 2010 */
Data bmoreh.bill2010_var4dist;
Set Source.bill2010_var4dist;
Rename indvid = ID;
Rename EstablishmentId = ext_est_id;
run;

/* Billing data 2011 */
Data bmoreh.bill2011_var4dist;
Set Source.bill2011_var4dist;
Rename indvid = ID;
Rename EstablishmentId = ext_est_id;
run;

/* Sort all billing data sets prior to merge to one large data
set (2006-2011) */

/* sort 2006 by ID */
PROC SORT
DATA = bmoreh.bill2006_var4dist;
BY ID;
run;

/* sort 2007 by ID */
PROC SORT
DATA = bmoreh.bill2007_var4dist;
```

```
BY ID;
run;

/* sort 2008 by ID */
PROC SORT
DATA = bmoreh.bill2008_var4dist;
BY ID;
run;

/* sort 2009 by ID */
PROC SORT
DATA = bmoreh.bill2009_var4dist;
BY ID;
run;

/* sort 2010 by ID */
PROC SORT
DATA = bmoreh.bill2010_var4dist;
BY ID;
run;

/* sort 2011 by ID */
PROC SORT
DATA = bmoreh.bill2011_var4dist;
BY ID;
run;

/* Create a table that combines all Billing data years into
one dataset.
- Sequence Number and fiscal year variables were included in
table as identifiers. */

proc sql;
Create table bmoreh.bill2006_2011 as
select ID,servdate,provider_ID,ext_est_id, seqn, 2006 as
fiscalyr From bmoreh.bill2006_var4dist
union
select ID,servdate,provider_ID,ext_est_id, seqn, 2007 as
fiscalyr From bmoreh.bill2007_var4dist
union
select ID,servdate,provider_ID,ext_est_id, seqn, 2008 as
fiscalyr From bmoreh.bill2008_var4dist
union
select ID,servdate,provider_ID,ext_est_id, seqn, 2009 as
fiscalyr From bmoreh.bill2009_var4dist
union
select ID,servdate,provider_ID,ext_est_id, seqn, 2010 as
fiscalyr From bmoreh.bill2010_var4dist
union
```

```
select ID,servdate,provider_ID,ext_est_id, seqn, 2011 as
fiscalyr From bmoreh.bill2011_var4dist;
run;

/* Sort 'bill2006_2011' data set
note: billing data registers each service received separately,
even if at same location.
- By using 'nodupkey', should be able to remove duplicates
that might be repeated

/* PROC SORT
DATA = bmoreh.bill2006_2011 nodupkey;
BY ID servdate provider_id est_id fiscalyr;
run; */

/* Sort 'addr_var4dist_2' data set by ID for Merge */
PROC SORT
DATA = bmoreh.addr_var4dist_2;
BY ID;
run;


/* Sort 'billing_2006_2011' data set by ID for Merge */
PROC SORT
DATA = bmoreh.bill2006_2011;
BY ID;
run;

/* B. Combine 'bill2006_2011' data set and 'addr_var4dist_2'
(Citizen) data set to link ID and Postal Codes (billing data
has no postal codes originally).
- Assign Postal code to ID number IF conditions are met.
     Condition: addr.eff_date <= bill.servdate <=
addr.term_date

Note:
'eff_date' = Effective Date of postal code from
'addr_var4dist_2'
'term_date' = Termination Date of postal code from
'addr_var4dist_2'
'servdate' = Date of Service from 'bill2006_2011' */

data bmoreh.addr_bill2006_2011;
MERGE bmoreh.bill2006_2011(in=a) bmoreh.addr_var4dist_2(in=b);
by ID;
if a=1 and b=1;
if servdate ge eff_date and servdate le term_date;
run;
```

```
/* STEP 4: Merge Billing data 'est_id' with external
Establishments ID (ext_est_id) to determine a postal code for
the Establishment ID (Destination) */

/* A. Many 'unknown' establishment id that is placed with a
'.'
- Assume '.' or no est_id is associated with a '0'
- Change all "." or " " ext_est_id variables to "0" to be able
to link with with provider ID number and practice location.
- Remove any missing provider_id numbers because they cannot
be used to link */

data bmoreh.addr2_bill2006_2011;
set bmoreh.addr_bill2006_2011;
IF ext_est_id = '.' THEN DO
ext_est_id = '0';
 IF ext_est_id = '' THEN DO
 ext_est_id = '0';
end;
 end;
IF provider_id = '.' THEN DELETE;
IF provider_id = '' THEN DELETE;
run;

/*Sort 'new_estab' by est_id
note: some establishments have no postal code associated with
the est_id or the ext_est_id */

proc sort
data = bmoreh.new_estab;
by ext_est_id;
run;

/*Sort 'Final_addBill_2006_2011' By est_id */

proc sort
data = bmoreh.addr2_bill2006_2011;
by ext_est_id;
run;

/* B. Merge data sets 'new_estab' and 'addr2_bill2006_2011' BY
the 'est_id' variable.  */

data bmoreh.bill_estab_dest;
MERGE bmoreh.addr2_bill2006_2011 bmoreh.new_estab;
by ext_est_id;
run;
```

```
/* Merge creates a large number of duplicates because SAS
tries to match with everything therefore need to remove
duplicates */

proc sort
data = bmoreh.bill_estab_dest nodupkey;
BY ID servdate provider_ID ext_est_id pcode fiscalyr
est_pcode;
run;

/* STEP 5: For Dataset 'bill_estab_dest', the 'ext_est_id'
variable have a '0' code to represent a service visit to an
office or clinic.
- Using the provider registry data set
'practice_loc_var4dist', link the postal code (ploc_pcode)
associated with the office/clinic (Practice location)
  to the 'service date' by the 'provider_id' number IF the
following condition is met:
     (Condition: servdate >= EFF_DATE_PRACT_LOC AND servdate
<= TERM_DATE_PRACT_LOC)

- 'servdate' = date of service from 'final2_bill_est_dest'
- 'EFF_DATE_PRACT_LOC' = Effective date of practice location
from 'practice_loc_var4dist'
- 'TERM_DATE_PRACT_LOC' = Termination date of practice
location from 'practice_loc_var4dist' */

/* A. Complete the following notes to get dataset in proper
format.
FILE NAME: 'practice_loc_var4dist'

- Make sure no missing data:
Some 'term_date' values are missing from dataset
practice_loc_var4dist and are replaced with a '.' For these
values, assume no 'term_date' to current date
and set variable 'term_date' to December 31, 2019' (term_date
= mdy(12,13,2019);)
- Missing provider_id Numbers from dataset need to be deleted
because they cannot be used.
- Provider_id Numbers with no Postal codes need to be deleted
because they cannot be used. */

Data bmoreh.practice_loc_var4dist;
Set Source.Practice_loc_var4dist;
IF TERM_DATE_PRACT_LOC = . THEN TERM_DATE_PRACT_LOC =
mdy(12,31,2019);
IF provider_id = '' THEN DELETE;
IF POSTAL_CODE_PRACT_LOC = '' THEN DELETE;
Label POSTAL_CODE_PRACT_LOC = 'Practice Location pcode';
```

```
run;

/* B. practice_loc_var4dist = FULL Data set including areas
NOT in New Brunswick
Create a separate data set named 'NBpractice_loc_var4dist'

- NBpractice_loc_var4dist = Dataset with ONLY NB Locations
- Set Extract a substring from argument
(POSTAL_CODE_PRACT_LOC) at position (1) for n characters (1)=
'E'.
- This tells SAS to show only NB postal codes (All postal
codes beginning with 'E').
- Note: Will later need to rename variables to match variables
in PCCF+ file */

Data bmoreh.NBpractice_loc_var4dist;
Set bmoreh.Practice_loc_var4dist;
where substr(POSTAL_CODE_PRACT_LOC,1,1)='E';
run;
/* Total observations = 19900 */

/*  Sort both 'final2_bill_est_dest' and
'practice_loc_var4dist' by provider_ID */

proc sort
data = bmoreh.NBpractice_loc_var4dist;
by provider_id;
run;

proc sort
data = bmoreh.bill_estab_dest;
by provider_id;
run;

/* C. Combine 'bill_estab_dest' and 'NBpractice_loc_var4dist'
by provider_id variable IF est_id = "0" AND the following
conditions are met:
     Conditions: servdate >= EFF_DATE_PRACT_LOC and servdate
<= TERM_DATE_PRACT_LOC

Notes:
POSTAL_CODE_PRACT_LOC = Practice location pcode (desination)
pcode = Patient address postal code (Origin)
est_pcode = Postal code for external/establishment number from
Establishment File. */

/* If Missing or incomplete data, delete. */
```

```
/* When missing or incomplete data from ID, provider_id and
date is removed */

data bmoreh.dist4phybill;
MERGE bmoreh.bill_estab_dest(in=a)
bmoreh.Nbpractice_loc_var4dist(in=b);
by provider_ID;
if a=1 and b=1;
if ext_est_id = '0' then
do;
if servdate ge EFF_DATE_PRACT_LOC and servdate le
TERM_DATE_PRACT_LOC then
do;
est_pcode = POSTAL_CODE_PRACT_LOC;
end;
end;
IF ID = '' THEN DELETE;
IF provider_ID = '' THEN DELETE;
IF servdate = '.' THEN DELETE;
run;

/* STEP 6: Combine datasets 1: DAD data, and dataset 2:
Billing data, with PCCF+ file to link postal codes with DA.
PCCF+ File: 'Final_hlthout'
Dataset 1: Combine DAD data, Citizen Data and Establishment
Data - FILE NAME: 'Final_dist4DAD'
Dataset 2: Combined Billing Data, Civilian Data, Provider
Registry Data and Establishment Data - FILE NAME:
'dist4phybill'

Note: Have to keep DAD dataset and Billing dataset separate
because if not, will have to add in another identifier.
Therefore, combine 'Final_hlthout' with'Final_dist4DAD' and
'dist4phybill' in separate operations.
Because these are separate operations, distances from ArcMap
were done separately as well.

A. Make 'Final_hlthout' more manageable by removing
unnecessary variables.
Keep the following terms:
- PCODE
- indvid
- DAuid
- CMAname
- CDname
- LAT
- LONG
- Comm_Name
```

```
- Rename indvid to ID to match other datasets.

- Delete any with blank coordinates (LAT,LONG) */

data bmoreh.hlth_2;
set bmoreh.Final_hlthout (KEEP = indvid PCODE LAT LONG DAuid
CMAname CDname Comm_Name );
Rename indvid = ID;
IF LAT = "." THEN DELETE;
IF LONG = "." THEN DELETE;
run;

/* B. Need to determine which variable to KEEP that are common
between Dataset 1 'Final_dist4DAD' and Dataset 2
'final_dist4phybill'.
Rename variables for each dataset to be consistent with each
other. */

/* Dataset 1: 'final_dist4DAD'
Keep the following Variables:
ID = Personal ID#
Postalcode = Origin Postal code/ place of residence)
NewAdmdate = Date admitted to Hospital
est_id = Establishment ID (#1) - For DAD data
City = City of Hospital
est_pcode = Establishment Postal code
Sqn = Sequence Number (Identifier for real data)
fiscalyr = Fiscal year of DAD data (Identifier)

Rename the following:
Postalcode = origin_pcode
NewAdmdate = DATE
est_pcode = dest_pcode

NEW DATASET 1: 'final2_dist4DAD' */

data bmoreh.Final2_dist4DAD;
set bmoreh.Final_dist4DAD (KEEP = ID Postalcode NewAdmdate
est_id est_pcode City seqn fiscalyr);
Rename Postalcode = org_pcode;
Rename NewAdmdate = DATE;
Rename est_pcode = dest_pcode;
IF ID = '' THEN DELETE;
run;

/* Dataset 2: 'dist4phybill'
Keep the following Variables:
ID = Personal ID#
pcode = Origin Postal code/ place of residence
```

```
provider_ID = Provider Number
servdate = Date of service being received
ext_est_id = Establishment ID (#2) - For Billing Data
City = City of Office, clinic, etc.
ploc_codes = Practice location Postal code (Including
hospital, clinic or office)
Sqn = Sequence Number (Identifier for real data)
fiscalyr = Fiscal year of DAD data (Identifier)
POSTAL_CODE_PRACT_LOC = Postal code of practice location

Rename the following:
pcode = origin_pcode
servdate = DATE
ploc_codes = dest_pcode

NEW DATASET 2: 'final_dist4phybill' */

data bmoreh.final_dist4phybill;
set bmoreh.dist4phybill (KEEP = ID pcode servdate ext_est_id
City provider_ID POSTAL_CODE_PRACT_LOC est_pcode seqn
fiscalyr);
Rename pcode = org_pcode;
Label POSTAL_CODE_PRACT_LOC = 'Provider Postal Code';
Label est_pcode = 'Establishment Postal Code';
Label ext_est_id = 'External Establishment ID';
Label servdate = 'Service Date';
run;

/* C. Combine Dataset 1 'final2_dist4DAD' and Dataset 2
'final2_dist4phybill' with PCCF+ dataset 'hlth_2' to determine
DAuid and geocoding attributes */

/* DATASET 1: 'final2_dist4DAD'
PCCF + File: 'hlth_2' */

Proc sort
Data = bmoreh.hlth_2;
By ID;
run;

Proc sort
Data = bmoreh.final2_dist4DAD;
By ID;
run;
/* Combine files 'final2_dist4DAD' and 'hlth_2'
Link 'origin_pcode' with 'DAuid' BY 'ID number. This will give
DA information associated with the postal code for the origin
postal codes (patient residence).
```

```
This allows us to assign a DA (origin) to each health trip in
the DAD dataset */

Data bmoreh.dad_dist4travel;
MERGE bmoreh.final2_dist4DAD bmoreh.hlth_2;
BY ID;
IF org_pcode = PCODE;
IF ID = '' THEN DELETE;
label org_pcode = 'Origin Postal Code';
label dest_pcode = 'Destination Postal Code';
label est_id = 'Establishment ID';
label DAuid = 'Origin DAuid';
label CDname = 'Origin CDname';
label Comm_Name = 'Origin Comm_Name';
label CMAname = 'Origin CMAname';
run;

/* Remove any duplicates that are present */
Proc sort
data = bmoreh.dad_dist4travel OUT = bmoreh.dad_dist4travel
nodupkey;
by ID DATE org_pcode est_id dest_pcode DAuid seqn fiscalyr;
run;


/* DATASET 2: 'final_dist4phybill'
PCCF + File: 'Org_hlth' */

Proc sort
Data = bmoreh.hlth_2;
By ID;
run;

Proc sort
Data = bmoreh.final_dist4phybill;
By ID;
run;
```

```
/* Combine files 'final_dist4phybill' and 'hlth_2'
Link 'origin_pcode' with 'DAuid' BY 'ID number. This will give
DA information associated with the postal code for the origin
postal codes (patient residence).
We do not need to do this for 'dest_pcode' because there are a
limited number of postal code destinations due to the # of
hospitals, clinics, doctor offices */

Data bmoreh.phybill_dist4travel;
MERGE bmoreh.final_dist4phybill bmoreh.hlth_2;
BY ID;
IF org_pcode = PCODE;
IF ID = '' THEN DELETE;
label org_pcode = 'Origin Postal Code';
label ext_est_id = 'Ext Establ ID';
label DAuid = 'Origin DAuid';
label CDname = 'Origin CDname';
label Comm_Name = 'Origin Comm_Name';
label CMAname = 'Origin CMAname';
run;

/* Remove any duplicates that are present */
Proc sort
data = bmoreh.phybill_dist4travel OUT =
bmoreh.phybill_dist4travel nodupkey;
by ID servdate org_pcode ext_est_id est_pcode DAuid seqn
fiscalyr POSTAL_CODE_PRACT_LOC;
run;


/* STEP 7: Map distance for each row in DADxxxx_var4dist and
billxxxx_var4dist by matching dauid and pcode_dest */

/* A. Use ArcMap to determine distance combinations between
origin DAs and destination postal codes (Hospital locations/
office locations)
          - Hospital/ Facility Locations: file in DMTI folder
used in ArcMap as a facility layer
          - Physician office/ clinic Locations: file created
from physician billing and PCCF+ file. LAT, LONG from PCCF+
file was linked with the postal code from
            the physician's practice location. - SAS file
name: 'Physician Locations -  ArcMap'
          - OD-Matrix was used to determine distance
combinations.

Dataset 1: 'dad_dist4travel'
Dataset 2: 'phybill_dist4travel' */
```

```
/* B. Bring Text file 'arcmapdistances' in from arc map
note that there could be duplicates if physicians reference
hospitals therefore run nodupkey */

filename arcmap "K:\p0036_travel_demand\users\Brianna
Morehouse\ArcMap\arcmapdistances.txt";
data bmoreh.arcmapdistances;
/*formatting the table to read properly as sas output */
format
     ObjectID
     Name            $200.
     OriginID        $200.
     DestinationID   $200.
     Total_Length    14.4
     ObjectID_1      $200.
     Name_1          $200.
     Name_12         $200.
;
infile arcmap dsd dlm="," truncover firstobs= 2;
/*input file is what sas reads from the arcmap txt file*/
input
     ObjectID        $
     Name            $
     OriginID        $
     DestinationID   $
     Total_Length
     ObjectID_1      $
     Name_1          $
     Name_12         $
;
/*drop the following variables from the table (to clean it
up)*/
DROP
     ObjectID
     Name
     OriginID
     DestinationID
     ObjectID_1;

Rename Name_1 = DAuid;
Rename Name_12 = dest_pcode;
Rename Total_Length = distance
;
run;
Proc sort
data = bmoreh.arcmapdistances OUT = bmoreh.arcmapdistances
nodupkey;
by DAuid dest_pcode;
run;
```

```
/* C. LINK DAD DATA 'dad_dist4travel' WITH TRAVEL DISTANCES TO
Destinations 'arcmapdistances' */

/* The following datasets were merged to assign total lengths
(m) to each origin/destination trip
(these are distance combinations between origin centroids (DA)
and destinations locations (Establishments/ hospitals) created
within arc map. */


/* Minimize 'dad_dist4travel' for merge to link distances */
data bmoreh.dad_dist4travel2;
set bmoreh.dad_dist4travel;
DROP
Pcode
LAT
LONG
Comm_Name
CMAname;
run;

/* Make sure variables are the correct length and data type to
match dad data set
Dauid = length of $8.
dest_pcode length of $8. */

data bmoreh.arcmapdistances2;
length dauid $8. dest_pcode $8.;
set bmoreh.arcmapdistances;
len = round(distance / 1000);
DAuid = compress(DAuid);
dest_pcode = compress(dest_pcode);
run;

/* Merge Dataset 1: 'dad_dist4travel' with 'arcmapdistances'
By DAuid if EstblPosta = dest_pcode */

Proc sort
data = bmoreh.dad_dist4travel2;
by DAuid dest_pcode;
run;

Proc sort
data = bmoreh.arcmapdistances2;
by DAuid dest_pcode;
run;

Data bmoreh.Total_dist4dad;
MERGE bmoreh.dad_dist4travel2 bmoreh.arcmapdistances2;
```

```
BY DAuid dest_pcode;
label org_pcode = 'Origin Postal Code';
label DAuid = 'Origin DAuid';
label distance = 'Travel Distance';
label len = 'Disance (km)';
IF ID = '' THEN DELETE;
IF distance = '.' THEN DELETE;
run;

/* D. LINK PHYSICAN BILLING DATA 'phybill_dist4travel' WITH
TRAVEL DISTANCES TO Destinations 'arcmap_dist4travel' */

/* The following datasets were merged to assign total lengths
(m) to each origin/destination trip
(these are distance combinations between origin centroids (DA)
and destinations locations (Establishments/ hospitals) created
within arc map. */


/* Minimize 'phybill_dist4travel' for merge to link distances
*/
data bmoreh.phybill_dist4travel2;
set bmoreh.phybill_dist4travel;
Rename POSTAL_CODE_PRACT_LOC = dest_pcode;
DROP
City
est_pcode
Pcode
LAT
LONG
Comm_Name
CMAname;
run;

/* Make sure variables are the correct length and data type
Dauid = length of $8.
dest_pcode length of $6. */

data bmoreh.arcmapdistances2;
length dauid $8. dest_pcode $6.;
set bmoreh.arcmapdistances;
len = round(distance / 1000);
DAuid = compress(DAuid);
dest_pcode = compress(dest_pcode);
run;

/* Merge Dataset 2: 'phybill_dist4travel2' with
'arcmapdistances' By DAuid if prac_pcode = dest_pcode */
```

```
Proc sort
data = bmoreh.phybill_dist4travel2;
by DAuid dest_pcode;
run;

Proc sort
data = bmoreh.arcmapdistances2;
by DAuid dest_pcode;
run;

Data bmoreh.Total_dist4phybill;
MERGE bmoreh.phybill_dist4travel2 bmoreh.arcmapdistances2;
BY DAuid dest_pcode;
label dest_pcode = 'Practice Location Postal Code';
label DAuid = 'Origin DAuid';
label distance = 'Distance (m)';
label len = 'Disance (km)';
IF ID = '' THEN DELETE;
IF distance = '.' THEN DELETE;
run;

/* Add Destination Name to file based on destination postal
code */

data bmoreh.pccf_dest (KEEP = dest_pcode dest_CDname
dest_common);
set bmoreh.hlth_2;
rename PCODE = dest_pcode;
rename comm_name = dest_common;
rename CDname = dest_CDname;
run;

proc sort
data = bmoreh.pccf_dest nodupkey;
BY dest_pcode;
run;

proc sort
data = bmoreh.Total_dist4phybill;
BY dest_pcode;
run;

data bmoreh.total_dist4phybill2;
MERGE bmoreh.Total_dist4phybill bmoreh.pccf_dest;
BY dest_pcode;
IF ID = '' THEN DELETE;
run;
/*remove postal codes */
data moreh.dist4dad_nopc;
```

```
set bmoreh.Total_dist4dad;
rename city = dest_city;
DROP
org_pcode
dest_pcode;
run;

data moreh.dist4phybill_nopc;
set bmoreh.Total_dist4phybill2;
label dest_common = 'Common Name (destination)';
label dest_CDname = 'County - Destination';
label CDname = 'County - Origin';
DROP
org_pcode
dest_pcode;
run;

/* remove postal code and address from establishments */

data moreh.estab_nopc;
set source.establishments;
DROP
Street_Address
Postal_Code
Effective_End_Date
Inst_Open_Date
Inst_Close_Date;
run;
```

Adding Physician Billing Locations into ArcMap

```
/* Libraries */
libname moreh "K:\p0036_travel_demand\users\Brianna
Morehouse";
libname bmoreh "K:\p0036_travel_demand\users\Brianna
Morehouse\bmoreh";
libname source "K:\p0036_travel_demand\src";
libname arcmap "K:\p0036_travel_demand\users\Brianna
Morehouse\ArcMap";

/* PCCF+ Data */

data arcmap.Final_hlthout;
set moreh.Final_hlthout;
length indvid $9.;
indvid=compress(id);
run;

data arcmap.hlth_2;
set arcmap.Final_hlthout (KEEP = indvid PCODE LAT LONG DAuid
CMAname CDname Comm_Name );
Rename indvid = ID;
IF LAT = "." THEN DELETE;
IF LONG = "." THEN DELETE;
run;

/* Postal codes have multiple variations of LAT/LONG therefore
Remove duplicates by sorting with postal code and DAuid */

Proc sort
data = arcmap.hlth_2 nodupkey;
BY pcode DAuid ;
run;


/* Physician Practice Locations

- Make sure no missing data:
Some 'term_date' values are missing from dataset
practice_loc_var4dist and are replaced with a '.' For these
values, assume no 'term_date' to current date
and set variable 'term_date' to December 31, 2019' (term_date
= mdy(12,13,2019);)
- Missing provider_id Numbers from dataset need to be deleted
because they cannot be used.
- Provider_id Numbers with no Postal codes need to be deleted
because they cannot be used. */
```

```
Data arcmap.practice_loc_var4dist;
Set Source.Practice_loc_var4dist;
IF TERM_DATE_PRACT_LOC = . THEN TERM_DATE_PRACT_LOC =
mdy(12,31,2019);
IF provider_id = '' THEN DELETE;
IF POSTAL_CODE_PRACT_LOC = '' THEN DELETE;
Label POSTAL_CODE_PRACT_LOC = 'Practice Location pcode';
run;

/* Create a separate data set named 'NBpractice_loc_var4dist'

- NBpractice_loc_var4dist = Dataset with ONLY NB Locations
- Set Extract a substring from argument
(POSTAL_CODE_PRACT_LOC) at position (1) for n characters (1)=
'E'.
- This tells SAS to show only NB postal codes (All postal
codes beginning with 'E').
- Note: Will later need to rename variables to match variables
in PCCF+ file */
Data arcmap.NBpractice_loc_var4dist;
Set arcmap.Practice_loc_var4dist;
where substr(POSTAL_CODE_PRACT_LOC,1,1)='E';
run;
/* Remove any duplicate Postal codes
For ArcMap purposes, just looking for one location per office,
do not need multiple variations of doctors sharing locations
*/

proc sort
data = arcmap.NBpractice_loc_var4dist nodupkey;
by POSTAL_CODE_PRACT_LOC;
run;

 /* Rename POSTAL_CODE_PRACT_LOC to pcode to match hlth_2
variable name */
Data arcmap.NBpractice_loc_var4dist;
Set arcmap.Practice_loc_var4dist;
rename POSTAL_CODE_PRACT_LOC = pcode;
run;

/* NB Establishments and External Establishments */

Data arcmap.establishments;
Set Source.establishments;
IF Postal_Code = '' THEN DELETE;
run;

/* Rename POSTAL_CODE_PRACT_LOC to pcode to match hlth_2
variable name */
```

```
data arcmap.new_estab;
Set arcmap.establishments (KEEP = Establishment_Id
External_Estbl_Id Postal_Code City);
rename Establishment_Id = est_id;
rename Postal_Code = pcode;
rename External_Estbl_Id = ext_est_id;
run;

/* Remove any duplicated establishment postal codes
For ArcMap purposes just looking at single location, not
various services within one location */

Proc sort
data = arcmap.new_estab nodupkey;
by pcode;
run;

/* Merge Practice Locations with the PCCF+ file to assign LAT
LONG coordinates.
Remove any variable that are not needed for Arc Map */

Proc sort
data = arcmap.NBpractice_loc_var4dist;
by pcode;
run;

Proc sort
data = arcmap.hlth_2;
by pcode;
run;

data arcmap.hlth_pracloc (DROP = EFF_DATE_PRACT_LOC
TERM_DATE_PRACT_LOC provider_id ID) ;
MERGE arcmap.NBpractice_loc_var4dist (in=a) arcmap.hlth_2
(in=b);
IF a=1 and b=1;
by pcode;
IF LAT = '.' THEN DELETE;
IF LONG = '.' THEN DELETE;
run;

/*Check and see if we want to sort by just pcode if a lot of
duplicated pcodes with different DA*/

Proc sort
data = arcmap.hlth_pracloc nodupkey;
by pcode;
run;
```

```
/* use file 'hlth_pracloc' to create arcmap layer Part 1 */

/* External Establishment files with the PCCF+ file to assign
LAT LONG coordinates.
Remove any variable that are not needed for Arc Map */

Proc sort
data = arcmap.New_estab;
by pcode;
run;

Proc sort
data = arcmap.hlth_2;
by pcode;
run;

data arcmap.hlth_estab (DROP = est_id ID) ;
MERGE arcmap.New_estab (in=a) arcmap.hlth_2 (in=b);
IF a=1 and b=1;
by pcode;
IF LAT = '.' THEN DELETE;
IF LONG = '.' THEN DELETE;
run;

/* Delete duplicates sorting by pcode and DAuid
Check and see if we want to sory by just pcode if alot of
douplicated pcodes with different DA*/
Proc sort
data = arcmap.hlth_estab nodupkey;
by pcode;
run;
/* use file 'hlth_estab' to create arcmap layer Part 2 */
```

Part 2 – Combining Travel Distances with Associated Health Variables

Physician Billing Health Variables:

```
libname moreh "K:\p0036_travel_demand\users\Brianna
Morehouse";
libname bmoreh "K:\p0036_travel_demand\users\Brianna
Morehouse\bmoreh";
libname source "K:\p0036_travel_demand\src";
libname sourceb "K:\p0036_travel_demand\src\briana";
libname stats "K:\p0036_travel_demand\users\Brianna
Morehouse\stats";

/* Combine Phy.Billing Datasets
     note: in calendar year */
/* Billing 2006 */
```

```
data bmoreh.bill2006_final;
set source.bill2006_final;
run;

proc sort
data = bmoreh.bill2006_final;
by indvid;
run;

/* Billing 2007 */
data bmoreh.bill2007_final;
set source.bill2007_final;
run;

proc sort
data = bmoreh.bill2007_final;
by indvid;
run;

/* Billing 2008 */
data bmoreh.bill2008_final;
set source.bill2008_final;
run;

proc sort
data = bmoreh.bill2008_final;
by indvid;
run;

/* Billing 2009 */
data bmoreh.bill2009_final;
set source.bill2009_final;
run;

proc sort
data = bmoreh.bill2009_final;
by indvid;
run;

/* Billing 2010 */
data bmoreh.bill2010_final;
set source.bill2010_final;
run;
proc sort
data = bmoreh.bill2010_final;
by indvid;
run;

/* Billing 2011 */
```

```
data bmoreh.bill2011_final;
set source.bill2011_final;
run;

proc sort
data = bmoreh.bill2011_final;
by indvid;
run;

proc sql;
Create table bmoreh.bill2006_2011 as
select
indvid,seqn,servdate,DAuid,fsa,provider_ID,SERVICE_TYPE_CODE,S
ERVICE_TYPE_GROUP_CODE,SERVICE_TYPE_GROUP_DESC,SERVICE_CATEGOR
Y_CODE,SPECIALTY_TYPE_CODE,EstablishmentId,EstablishmentDescE,
is_institution, 2006 as Year From bmoreh.bill2006_final
union
select
indvid,seqn,servdate,DAuid,fsa,provider_ID,SERVICE_TYPE_CODE,S
ERVICE_TYPE_GROUP_CODE,SERVICE_TYPE_GROUP_DESC,SERVICE_CATEGOR
Y_CODE,SPECIALTY_TYPE_CODE,EstablishmentId,EstablishmentDescE,
is_institution, 2007 as Year From bmoreh.bill2007_final
union
select
indvid,seqn,servdate,DAuid,fsa,provider_ID,SERVICE_TYPE_CODE,S
ERVICE_TYPE_GROUP_CODE,SERVICE_TYPE_GROUP_DESC,SERVICE_CATEGOR
Y_CODE,SPECIALTY_TYPE_CODE,EstablishmentId,EstablishmentDescE,
is_institution, 2008 as Year From bmoreh.bill2008_final
union
select
indvid,seqn,servdate,DAuid,fsa,provider_ID,SERVICE_TYPE_CODE,S
ERVICE_TYPE_GROUP_CODE,SERVICE_TYPE_GROUP_DESC,SERVICE_CATEGOR
Y_CODE,SPECIALTY_TYPE_CODE,EstablishmentId,EstablishmentDescE,
is_institution, 2009 as Year From bmoreh.bill2009_final
union
select
indvid,seqn,servdate,DAuid,fsa,provider_ID,SERVICE_TYPE_CODE,S
ERVICE_TYPE_GROUP_CODE,SERVICE_TYPE_GROUP_DESC,SERVICE_CATEGOR
Y_CODE,SPECIALTY_TYPE_CODE,EstablishmentId,EstablishmentDescE,
is_institution, 2010 as Year From bmoreh.bill2010_final
union
select
indvid,seqn,servdate,DAuid,fsa,provider_ID,SERVICE_TYPE_CODE,S
ERVICE_TYPE_GROUP_CODE,SERVICE_TYPE_GROUP_DESC,SERVICE_CATEGOR
Y_CODE,SPECIALTY_TYPE_CODE,EstablishmentId,EstablishmentDescE,
is_institution, 2011 as Year From bmoreh.bill2011_final;
run;
```

```
/* Add register data to dist4phybill */

data bmoreh.dist4phybill;
set sourceb.dist4phybill_nopc;
rename ID = indvid;
rename fiscalyr = year;
run;

proc sort
data = bmoreh.dist4phybill;
BY indvid;
run;

proc sort
data = bmoreh.reg_final;
BY indvid;
run;

data bmoreh.phybill2;
MERGE bmoreh.dist4phybill bmoreh.reg_final;
BY indvid;
IF servdate = '.' THEN DELETE;
run;

proc sort
data = bmoreh.phybill2 nodupkey;
BY indvid seqn servdate year;
run;


/* add address info to phybill2 */

Data bmoreh.addr2_final;
Set bmoreh.addr_final;
IF TERM_DATE = . THEN TERM_DATE = mdy(12,31,2019);
run;

proc sort
data = bmoreh.addr2_final;
BY indvid DAuid;
run;

proc sort
data = bmoreh.phybill2;
BY indvid DAuid;
run;
```

```
data bmoreh.phybill3;
MERGE bmoreh.phybill2 bmoreh.addr2_final;
by indvid DAuid;
if servdate ge EFF_DATE and servdate le TERM_DATE;
run;

/* LINK phybill3 with Billfinal2006_2011 */
proc sort
data= bmoreh.phybill3;
by indvid provider_id seqn year;
run;

proc sort
data = bmoreh.bill2006_2011;
by indvid provider_id seqn year;
run;

data bmoreh.phybill4;
MERGE bmoreh.phybill3 (in=a) bmoreh.bill2006_2011 (in=b);
by indvid provider_id seqn year;
if a=1 and b=1;
run;

proc sort data = bmoreh.phybill4 nodupkey;
by indvid provider_id seqn year;
run;

/* clean up phybill4 */
data bmoreh.phybill_4;
set bmoreh.phybill4;
rename CSDname = O_CSDname;
rename GEO_LOCATION_DESC = O_geoloc_desc;
rename fsa = O_fsa;
rename dest_CDname = D_CDname;
rename CDname = O_CDname;
     IF EstablishmentDescE = 'Establishment is missing' THEN
     do EstablishmentDescE = 'Office/Clinc';
     end;
DROP
     ext_est_id
     dest_common
     month_of_birth
     PROV
     GEO_LOCATION_CODE
     EFF_DATE
     TERM_DATE
     CMAname
     is_institution;
run;
```

```
data bmoreh.Spec_hist_provid;
set source.Specialty_hist2018_fin;
IF TERM_DATE_HISTORY = . THEN TERM_DATE_HISTORY =
mdy(12,31,2019);
IF provider_id = '' THEN DELETE;
run;

proc sort data = bmoreh.phybill_4;
BY provider_id;
run;


proc sort data = bmoreh.Spec_hist_provid;
BY provider_id;
run;

data bmoreh.phb_prv_hist;
MERGE bmoreh.phybill_4 bmoreh.Spec_hist_provid;
BY provider_id;
IF servdate ge EFF_DATE_HISTORY and servdate le
TERM_DATE_HISTORY;
run;

/* combine phybill4 with provider_base2018 */

data bmoreh.provbase2018;
set source.Provider_base2018_fin;
IF provider_id='' THEN DELETE;
rename Geo_Location_Cd_Cur = D_geoloc_code;
rename v3_City_Urb_Rur_Ind_Cur = D_v3city_ind;
rename v7_City_Urb_Rur_Ind_Cur = D_v7city_ind;
rename Hlth_Region_Cur = d_hlth_region_cur;
run;

proc sort
data = bmoreh.phb_prv_hist;
by provider_id;
run;

proc sort
data= bmoreh.provbase2018;
by provider_id;
run;

data bmoreh.phybill_prv2018;
MERGE bmoreh.phb_prv_hist (in=a) bmoreh.provbase2018 (in=b);
BY provider_id;
IF a=1 and b=1;
run;
```

```
data bmoreh.phybill_prv2018;
set bmoreh.phybill_prv2018;
Age = (year - year_of_birth );
rename EstablishmentId = External_Estbl_Id;
IF age = '60' THEN DELETE;
IF age = '61' THEN DELETE;
IF age = '62' THEN DELETE;
IF age = '63' THEN DELETE;
IF age = '64' THEN DELETE;
run;

PROC FREQ DATA = bmoreh.phybill_prv2018;
TABLES AGE;
RUN;

proc sort data = bmoreh.phybill_prv2018;
by External_Estbl_Id;
run;


proc sort data = bmoreh.estab;
by External_Estbl_Id;
run;

data bmoreh.phyb_estab_prv2018;
MERGE bmoreh.phybill_prv2018 bmoreh.estab;
by External_Estbl_Id;
IF indvid = '' THEN DELETE;
drop
est_id
Health_Region_Cd;
run;

/* create age groups similar to DAD dataset for descriptive
statistics */

proc format;
value agecat
low-<65='<65'
65-<70='65-69'
70-<75='70-74'
75-<80='75-79'
80-<85='80-84'
85-<90='85-89'
90-high='90+'
;
value agenum
65-<70='15'
```

```
70-<75='16'
75-<80='17'
80-<85='18'
85-<90='19'
90-high='20'
other = ' '
;
value agegroup
15='65-69'
16='70-74'
17='75-79'
18='80-84'
19='85-89'
20='90+'
;
run;

data bmoreh.phyb_estab_prv2018;
set bmoreh.phyb_estab_prv2018;
/*FORMAT agegroup agecat.;
agegroup = age;*/
FORMAT agegroup agegroup.;
agegroup = INPUT (PUT(age,agenum.),8.);
run;

data bmoreh.phyb_estab_prv2018;
set bmoreh.phyb_estab_prv2018;
IF QAIPPE = '9' THEN DELETE;
IF D_v3city_ind = 'X' THEN DELETE;
IF D_v7city_ind = 'X' THEN DELETE;
IF HEALTH_REGION = '9' THEN DELETE;
run;
```

DAD Health Variables:

```
libname moreh "K:\p0036_travel_demand\users\Brianna
Morehouse";
libname bmoreh "K:\p0036_travel_demand\users\Brianna
Morehouse\bmoreh";
libname source "K:\p0036_travel_demand\src";
libname sourceb "K:\p0036_travel_demand\src\briana";
libname stats "K:\p0036_travel_demand\users\Brianna
Morehouse\stats";


/* Combine DAD Data sets using the sql union function */

/* DAD 2005 */
Data bmoreh.Dad2005_final;
Set Source.Dad2005_final;
```

```
NewAdmdate = Input(Admdate,YYMMDD10.);
FORMAT NewAdmdate YYMMDD10.;
run;

PROC SORT
DATA = bmoreh.Dad2005_final;
BY indvid;
run;

/* DAD 2006 */
Data bmoreh.Dad2006_final;
Set Source.Dad2006_final;
NewAdmdate = Input(Admdate,YYMMDD10.);
FORMAT NewAdmdate YYMMDD10.;
run;

PROC SORT
DATA = bmoreh.Dad2006_final;
BY indvid;
run;

/* DAD 2007 */
Data bmoreh.Dad2007_final;
Set Source.Dad2007_final;
NewAdmdate = Input(Admdate,YYMMDD10.);
FORMAT NewAdmdate YYMMDD10.;
run;

PROC SORT
DATA = bmoreh.Dad2007_final;
BY indvid;
run;

/* DAD 2008 */
Data bmoreh.Dad2008_final;
Set Source.Dad2008_final;
NewAdmdate = Input(Admdate,YYMMDD10.);
FORMAT NewAdmdate YYMMDD10.;
run;

PROC SORT
DATA = bmoreh.Dad2008_final;
BY indvid;
run;

/* DAD 2009 */
Data bmoreh.Dad2009_final;
Set Source.Dad2009_final;
NewAdmdate = Input(Admdate,YYMMDD10.);
```

```
FORMAT NewAdmdate YYMMDD10.;
run;

PROC SORT
DATA = bmoreh.Dad2009_final;
BY indvid;
run;

/* DAD 2010 */
Data bmoreh.Dad2010_final;
Set Source.Dad2010_final;
NewAdmdate = Input(Admdate,YYMMDD10.);
FORMAT NewAdmdate YYMMDD10.;
run;

PROC SORT
DATA = bmoreh.Dad2010_final;
BY indvid;
run;

/* DAD 2011 */

Data bmoreh.Dad2011_final;
Set Source.Dad2011_final;
NewAdmdate = Input(Admdate,YYMMDD10.);
FORMAT NewAdmdate YYMMDD10.;
run;

PROC SORT
DATA = bmoreh.Dad2011_final;
BY indvid;
run;

/* Combine tables and remove the following because variables
too indepth or have many missing observations:
     - Diagnosis type 3-25
     - Diagnosis code 3-25
     - intervention code 2-20
     - invention provider service 2-20
     - Provider type 3-8
     - Provider service 3-8
     - intv_pid 2-18
     - provider id 3-8
     - DAuid06 - already have 'DAuid' variables */

proc sql;
Create table bmoreh.dad2005_2011 as
select
indvid,seqn,NewAdmdate,Instnum,Resdcode,Agecode,Ageunits,Gende
```

```
r,Admitcat,Entrycode,Agegroup,Insttype,diagcode1,diagtype1,dia
gcode2,diagtype2,intvcode1,intvproviderservice1,Totlenofstay,p
rovidertype1,providerservice1,providertype2,providerservice2,i
ntv_pid1,providerid_1,providerid_2,DAuid,fsa, 2005 as fiscalyr
From bmoreh.Dad2005_final
union
select
indvid,seqn,NewAdmdate,Instnum,Resdcode,Agecode,Ageunits,Gende
r,Admitcat,Entrycode,Agegroup,Insttype,diagcode1,diagtype1,dia
gcode2,diagtype2,intvcode1,intvproviderservice1,Totlenofstay,p
rovidertype1,providerservice1,providertype2,providerservice2,i
ntv_pid1,providerid_1,providerid_2,DAuid,fsa, 2006 as fiscalyr
From bmoreh.Dad2006_final
union
select
indvid,seqn,NewAdmdate,Instnum,Resdcode,Agecode,Ageunits,Gende
r,Admitcat,Entrycode,Agegroup,Insttype,diagcode1,diagtype1,dia
gcode2,diagtype2,intvcode1,intvproviderservice1,Totlenofstay,p
rovidertype1,providerservice1,providertype2,providerservice2,i
ntv_pid1,providerid_1,providerid_2,DAuid,fsa, 2007 as fiscalyr
From bmoreh.Dad2007_final
union
select
indvid,seqn,NewAdmdate,Instnum,Resdcode,Agecode,Ageunits,Gende
r,Admitcat,Entrycode,Agegroup,Insttype,diagcode1,diagtype1,dia
gcode2,diagtype2,intvcode1,intvproviderservice1,Totlenofstay,p
rovidertype1,providerservice1,providertype2,providerservice2,i
ntv_pid1,providerid_1,providerid_2,DAuid,fsa, 2008 as fiscalyr
From bmoreh.Dad2008_final
union
select
indvid,seqn,NewAdmdate,Instnum,Resdcode,Agecode,Ageunits,Gende
r,Admitcat,Entrycode,Agegroup,Insttype,diagcode1,diagtype1,dia
gcode2,diagtype2,intvcode1,intvproviderservice1,Totlenofstay,p
rovidertype1,providerservice1,providertype2,providerservice2,i
ntv_pid1,providerid_1,providerid_2,DAuid,fsa, 2009 as fiscalyr
From bmoreh.Dad2009_final
union
select
indvid,seqn,NewAdmdate,Instnum,Resdcode,Agecode,Ageunits,Gende
r,Admitcat,Entrycode,Agegroup,Insttype,diagcode1,diagtype1,dia
gcode2,diagtype2,intvcode1,intvproviderservice1,Totlenofstay,p
rovidertype1,providerservice1,providertype2,providerservice2,i
ntv_pid1,providerid_1,providerid_2,DAuid,fsa, 2010 as fiscalyr
From bmoreh.Dad2010_final
union
select
indvid,seqn,NewAdmdate,Instnum,Resdcode,Agecode,Ageunits,Gende
r,Admitcat,Entrycode,Agegroup,Insttype,diagcode1,diagtype1,dia
```

```
gcode2,diagtype2,intvcode1,intvproviderservice1,Totlenofstay,p
rovidertype1,providerservice1,providertype2,providerservice2,i
ntv_pid1,providerid_1,providerid_2,DAuid,fsa, 2011 as fiscalyr
From bmoreh.Dad2011_final;
run;

/* LINK reg_dad with dad2005_2011 */

data bmoreh.Dist4dad_nopc;
set sourceb.Dist4dad_nopc;
rename ID = indvid;
rename DAuid = OrgDAuid;
run;

Proc sort
data = bmoreh.Dist4dad_nopc;
BY indvid seqn;
run;

Proc sort
data = bmoreh.dad2005_2011;
BY indvid seqn;
run;


Data bmoreh.dad_dist;
MERGE bmoreh.dad2005_2011 (in=a) bmoreh.Dist4dad_nopc (in=b);
BY indvid seqn;
IF a=1 and b=1;
run;


/* LINK Patient Information from Addr_final
0 = out of province code
99997 = unknown location code */

data bmoreh.addr_final;
set source.addr_final;
IF GEO_LOCATION_CODE = '0' THEN DELETE;
DROP
DAIPPE
DNIPPE
HRuid
CSDuid
CMA
DA06uid
DA01uid
EA96uid
EA91uid
```

```
CSD;
run;

Data bmoreh.addr2_final;
Set bmoreh.addr_final;
IF TERM_DATE = . THEN TERM_DATE = mdy(12,31,2019);
run;

proc freq
data = bmoreh.addr_final;
tables GEO_LOCATION_CODE;
run;

proc sort
data = bmoreh.addr2_final;
BY indvid DAuid;
run;
 proc sort
 data = bmoreh.dad_dist;
 BY indvid DAuid;
 run;

data bmoreh.dad2_dist;
MERGE bmoreh.dad_dist bmoreh.addr2_final;
by indvid;
if NewAdmdate ge EFF_DATE and NewAdmdate le TERM_DATE;
run;

/* FINAL FILE NAME:dad2_dist */

data bmoreh.Spec_hist_prov1;
set source.Specialty_hist2018_fin;
IF TERM_DATE_HISTORY = . THEN TERM_DATE_HISTORY =
mdy(12,31,2019);
IF provider_id = '' THEN DELETE;
rename provider_id = providerid_1;
run;

proc sort data = bmoreh.dad2_dist;
BY providerid_1;
run;

proc sort data = bmoreh.Spec_hist_prov1;
BY providerid_1;
run;

data bmoreh.dad_spec_prv1;
MERGE bmoreh.dad2_dist bmoreh.Spec_hist_prov1;
BY providerid_1;
```

```
IF NewAdmdate ge EFF_DATE_HISTORY and NewAdmdate le
TERM_DATE_HISTORY;
run;

data bmoreh.provbase_dad;
set bmoreh.Provbase2018;
rename provider_id = providerid_1;
run;

proc sort data = bmoreh.dad_spec_prv1;
by providerid_1;
run;

proc sort data = bmoreh.provbase_dad;
by providerid_1;
run;

data bmoreh.dad_prvbase_spec;
MERGE bmoreh.dad_spec_prv1 bmoreh.provbase_dad;
BY providerid_1;
if indvid = '' THEN DELETE;
run;

data bmoreh.dad_prvbase_spec;
set bmoreh.dad_prvbase_spec;
Age= Input(Ageunits,BEST3.);
FORMAT Age BEST3.;
len_of_stay = input (Totlenofstay,BEST5.);
FORMAT  len_of_stay BEST5.;
run;
/* Double check that it's okay to add provider base data to
DAD data based on provider ID */

/* Remove ages below 65 years old */

proc freq data = bmoreh.dad_prvbase_spec;
tables age;
run;

data bmoreh.dad_prvbase_spec;
set bmoreh.dad_prvbase_spec;
IF age = '1' THEN DELETE;
IF age = '20' THEN DELETE;
IF age = '28' THEN DELETE;
IF age = '45' THEN DELETE;
IF age = '51' THEN DELETE;
IF age = '55' THEN DELETE;
IF age = '56' THEN DELETE;
IF age = '57' THEN DELETE;
```

```
IF age = '58' THEN DELETE;
IF age = '59' THEN DELETE;
IF age = '60' THEN DELETE;
IF age = '61' THEN DELETE;
IF age = '62' THEN DELETE;
IF age = '63' THEN DELETE;
IF age = '64' THEN DELETE;
run;

data bmoreh.estab;
set source.Establishments;
DROP
Organization_Id
Province_Cd
Street_Address
Postal_Code
Effective_End_Date
Inst_Open_Date
Inst_Close_Date
Max_Bassinets;
IF max_beds = '.' THEN DO max_beds = 0;
end;
rename Establishment_Id= est_id;
run;

proc sort data = bmoreh.dad_prvbase_spec;
by est_id;
run;

proc sort data = bmoreh.estab;
by est_id;
run;

data bmoreh.dad_prvbase_est;
MERGE bmoreh.dad_prvbase_spec bmoreh.estab;
BY est_id;
IF indvid = '' THEN DELETE;
rename city = est_city;
drop
Health_Region_Cd
External_Estbl_Id;
run;
```

B2: Syntax for Calculating Descriptive Statistics and Regression Analysis

Part 3 – Creating Descriptive Statistics and Regression Modelling

Descriptive Statistics for the DAD Dataset

```
/* Distance statistics */
Proc means data = bmoreh.dad_prvbase_est MEAN MEDIAN MODE
RANGE MAX MIN N NMISS;
class Specialty_Type_Cur;
var len;
run;

proc freq data = bmoreh.dad_prvbase_est;
tables D_v7city_ind*Specialty_Type_Cur/ NOROW NOCOL NOPERCENT
;
run;

proc freq data = bmoreh.dad_prvbase_est;
tables Specialty_Type_Cur*CDname*agegroup / NOROW NOCOL
NOPERCENT;
run;

/* Origin Destination Tables */
Proc freq data = bmoreh.dad_prvbase_est;
tables CDname*Cnty_Cd_Cur/ NOROW NOCOL NOPERCENT;
tables CDname*dest_city/ NOROW NOCOL NOPERCENT;
tables CSDname*dest_city/ NOROW NOCOL NOPERCENT;
tables HRename*d_hlth_region_cur/ NOROW NOCOL NOPERCENT;
run;
```

Creating Regression Variables for the DAD Dataset

```
libname moreh "K:\p0036_travel_demand\users\Brianna
Morehouse";
libname bmoreh "K:\p0036_travel_demand\users\Brianna
Morehouse\bmoreh";
libname source "K:\p0036_travel_demand\src";
libname sourceb "K:\p0036_travel_demand\src\briana";
libname stats "K:\p0036_travel_demand\users\Brianna
Morehouse\stats";

/* Create data set with variables for regression */

proc sql;
Create table bmoreh.dadvar4reg as
select
indvid,Age,Gender,NewAdmdate,seqn,fiscalyr,CDname,OrgDAuid,fsa
,QAIPPE,HEALTH_REGION,HRename,dest_city,est_id,Admitcat,Entryc
ode,Agegroup,
```

```
Insttype,diagcode1,diagtype1,providerid_1,providertype1,provid
erservice1,SPECIALTY_GROUP_CODE_HISTORY,Specialty_Type_Cur,Spe
cialty_Grp_Cur,Specialty_Typ_Cd_Cur,
Specialty_Minor_Grp_Code_Cur,Specialty_Major_Grp_Code_Cur,d_hl
th_region_cur,D_v3city_ind,D_v7city_ind,Prmry_Establ_Hlth_Regn
,len_of_stay,max_beds,distance,len
 from bmoreh.dad_prvbase_est;
run;


proc freq data = stats.dadvar4reg;
tables CDname;
run;

/* Create Dummy variables */

data stats.dadvar4reg;
set stats.dadvar4reg;

/* Physician Specialty n=1*/
IF Specialty_Typ_Cd_Cur = '100' THEN DO PHY_SP1 = 0;
     end;
IF Specialty_Typ_Cd_Cur = '200' THEN DO PHY_SP1 = 1;
     end;

/* QAIPPE n=4 */
IF QAIPPE = '1' THEN DO Q1 = 1;
     end;
IF QAIPPE = '2' THEN DO Q2 = 1;
     end;
IF QAIPPE = '3' THEN DO Q3 = 1;
     end;
IF QAIPPE = '4' THEN DO Q4 = 1;
     end;

/* Specialty Major n=2 */
IF Specialty_Major_Grp_Code_Cur = '11000' THEN DO MAJOR1 = 1;
     end;
IF Specialty_Major_Grp_Code_Cur = '12000' THEN DO MAJOR2 = 1;
     end;

/* Urban/rural n=1 */
IF D_v3city_ind = 'U' THEN DO dCITYv3 = 1;
     end;
IF D_v3city_ind = 'R' THEN DO dCITYv3 = 0;
     end;

/* Urban/rural n=1 */
IF D_v7city_ind = 'U' THEN DO dCITYv7 = 1;
```

```
     end;
IF D_v7city_ind = 'R' THEN DO dCITYv7 = 0;
     end;

/* gender n=1 */
IF gender = 'M' THEN DO gender2 = 1;
     end;
IF gender = 'F' THEN DO gender2 = 0;
     end;

/* Admit category n=1 */
IF Admitcat = 'U' THEN DO ADMIT1=1;
     end;
IF Admitcat = 'L' THEN DO ADMIT1=0;
     end;
/* Origin County n=14 */
IF CDname = 'Albert' THEN DO CD1=1;
     end;
IF CDname = 'Carleton' THEN DO CD2=1;
     end;
IF CDname = 'Charlotte' THEN DO CD3=1;
     end;
IF CDname = 'Gloucester' THEN DO CD4=1;
     end;
IF CDname = 'Kent' THEN DO CD5=1;
     end;
IF CDname = 'Kings' THEN DO CD6=1;
     end;
IF CDname = 'Madawaska' THEN DO CD7=1;
     end;
IF CDname = 'Northumberland' THEN DO CD8=1;
     end;
IF CDname = 'Queens' THEN DO CD9=1;
     end;
IF CDname = 'Restigouche' THEN DO CD10=1;
     end;
IF CDname = 'Saint John' THEN DO CD11=1;
     end;
IF CDname = 'Sunbury' THEN DO CD12=1;
     end;
IF CDname = 'Victoria' THEN DO CD13=1;
     end;
IF CDname = 'Westmorland' THEN DO CD14=1;
     end;

/* Origin health region n=7 */

IF health_region = '1' THEN DO HR1=1;
     end;
```

```
IF health_region = '2' THEN DO HR2=1;
     end;
IF health_region = '3' THEN DO HR3=1;
     end;
IF health_region = '4' THEN DO HR4=1;
     end;
IF health_region = '5' THEN DO HR5=1;
     end;
IF health_region = '6' THEN DO HR6=1;
     end;

/* Institution type n=4 */

IF Insttype = 'A' THEN DO INST_TYPE1 = 1;
     end;
IF Insttype = '1' THEN DO INST_TYPE2 = 1;
     end;
IF Insttype = '2' THEN DO INST_TYPE3 = 1;
     end;
IF Insttype = '3' THEN DO INST_TYPE4 = 1;
     end;

run;

data stats.dadvar4reg;
set stats.dadvar4reg;

IF Q1 = '.' THEN DO Q1=0;
     end;
IF Q2 = '.' THEN DO Q2=0;
     end;
IF Q3 = '.' THEN DO Q3=0;
     end;
IF Q4 = '.' THEN DO Q4=0;
     end;


IF MAJOR1 = '.' THEN DO MAJOR1 =0;
     end;
IF MAJOR2 = '.' THEN DO MAJOR2 =0;
     end;

IF CD1 = '.' THEN DO CD1=0;
     end;
IF CD2 = '.' THEN DO CD2=0;
     end;
IF CD3 = '.' THEN DO CD3=0;
     end;
IF CD4 = '.' THEN DO CD4=0;
```

```
      end;
IF CD5 = '.' THEN DO CD5=0;
      end;
IF CD6 = '.' THEN DO CD6=0;
      end;
IF CD7 = '.' THEN DO CD7=0;
      end;
IF CD8 = '.' THEN DO CD8=0;
      end;
IF CD9 = '.' THEN DO CD9=0;
      end;
IF CD10 = '.' THEN DO CD10=0;
      end;
IF CD11 = '.' THEN DO CD11=0;
      end;
IF CD12 = '.' THEN DO CD12=0;
      end;
IF CD13 = '.' THEN DO CD13=0;
      end;
IF CD14 = '.' THEN DO CD14=0;
      end;

IF HR1 = '.' THEN DO HR1=0;
      end;
IF HR2 = '.' THEN DO HR2=0;
      end;
IF HR3 = '.' THEN DO HR3=0;
      end;
IF HR4 = '.' THEN DO HR4=0;
      end;
IF HR5 = '.' THEN DO HR5=0;
      end;
IF HR6 = '.' THEN DO HR6=0;
      end;

IF INST_TYPE1 = '.' THEN DO INST_TYPE1 = 0;
      end;
IF INST_TYPE2 = '.' THEN DO INST_TYPE2 = 0;
      end;
IF INST_TYPE3 = '.' THEN DO INST_TYPE3 = 0;
      end;
IF INST_TYPE4 = '.' THEN DO INST_TYPE4 = 0;
      end;
run;
```

Correlation and Regression for the DAD Dataset:

```
/* New Brunswick */

proc corr data = stats.dadvar4reg ;
     var Age max_beds len_of_stay gender2
     Q1 Q2 Q3 Q4
     PHY_SP1
     ADMIT1
     MAJOR1 MAJOR2
     dCITYv3 dCITYv7
     CD1 CD2 CD3 CD4 CD5 CD6 CD7 CD8 CD9 CD10 CD11 CD12 CD13
CD14
     HR1 HR2 HR3 HR4 HR5 HR6
     INST_TYPE1 INST_TYPE2 INST_TYPE3 INST_TYPE4  ;
     with len;
run;

proc reg data = stats.dadvar4reg;
     model len =
     Age
     max_beds
     gender2
     Q1 Q2 Q3 Q4
     ADMIT1
     MAJOR1 MAJOR2
     CD1 CD2 CD3 CD4 CD5 CD6 CD7 CD8 CD9 CD10 CD11 CD12 CD13
CD14
     /tol vif ;
     title DAD Data - New Brunswick Regression Model;
run;
```

Descriptive Statistics for the Physician Billing Dataset:

```
/* Distance statistics */
Proc means data = bmoreh.phyb_estab_prv2018 MEAN MEDIAN MODE
RANGE MAX MIN N NMISS;
class O_CDname agegroup Specialty_Type_Cur;
var len;
run;

proc freq data = bmoreh.phyb_estab_prv2018;
tables Specialty_Type_Cur*agegroup*gender/ NOROW NOCOL ;
run;
/* Origin-Destination */
proc freq data = bmoreh.phyb_estab_prv2018;
tables O_CDname*D_CDname/ NOROW NOCOL NOPERCENT;
run;
```

Creating Regression Variables for the Physician Billing Dataset:

```
libname moreh "K:\p0036_travel_demand\users\Brianna
Morehouse";
libname bmoreh "K:\p0036_travel_demand\users\Brianna
Morehouse\bmoreh";
libname source "K:\p0036_travel_demand\src";
libname sourceb "K:\p0036_travel_demand\src\briana";
libname stats "K:\p0036_travel_demand\users\Brianna
Morehouse\stats";

/* Create data set with variables for regression */

proc sql;
Create table bmoreh.Phybvar4reg as
select
indvid,servdate,provider_ID,seqn,year,O_CDname,D_CDname,GENDER
,age,HEALTH_REGION,QAIPPE,SERVICE_TYPE_GROUP_CODE,SERVICE_CATE
GORY_CODE,
SPECIALTY_TYPE_CODE,Specialty_Grp_Cur,Specialty_Major_Grp_Code
_Cur,Specialty_Minor_Grp_Code_Cur,D_v3city_ind,D_v7city_ind,ma
x_beds,distance,len from bmoreh.phyb_estab_prv2018;
run;

proc freq data = bmoreh.Phybvar4reg;
tables Specialty_Grp_Cur;
tables SERVICE_CATEGORY_CODE;
tables age;
run;

/* Create dummy variables */

data stats.Phybvar4reg1;
set bmoreh.Phybvar4reg;

/* QAIPPE n=4 */
IF QAIPPE = '1' THEN DO Q1=1;
     end;
IF QAIPPE = '2' THEN DO Q2=1;
     end;
IF QAIPPE = '3' THEN DO Q3=1;
     end;
IF QAIPPE = '4' THEN DO Q4=1;
     end;

/* Specialty type n=1 */
IF SPECIALTY_TYPE_CODE = '100' THEN DO PHY_TYPE1 = 0;
     end;
IF SPECIALTY_TYPE_CODE = '200' THEN DO PHY_TYPE1 = 1;
     end;
```

```
/* Specialty Major n=2 */
IF Specialty_Major_Grp_Code_Cur = '11000' THEN DO MAJOR1 = 1;
     end;
IF Specialty_Major_Grp_Code_Cur = '12000' THEN DO MAJOR2 = 1;
     end;

/* Service type n=5 */

IF SERVICE_TYPE_GROUP_CODE = '1110' THEN DO SERVICE1 = 1;
     end;
IF SERVICE_TYPE_GROUP_CODE = '1130' THEN DO SERVICE2 = 1;
     end;
IF SERVICE_TYPE_GROUP_CODE = '1300' THEN DO SERVICE3 = 1;
     end;
IF SERVICE_TYPE_GROUP_CODE = '1500' THEN DO SERVICE4 = 1;
     end;
IF SERVICE_TYPE_GROUP_CODE = '2200' THEN DO SERVICE5 = 1;
     end;

/* Service Category Code n=2 */

IF SERVICE_CATEGORY_CODE = '1000' THEN DO SERV_CAT1 = 1;
     end;
IF SERVICE_CATEGORY_CODE = '2000' THEN DO SERV_CAT2 = 1;
     end;

/* Origin CDname n=14 */

IF O_CDname = 'Albert' THEN DO CD1=1;
     end;
IF O_CDname = 'Carleton' THEN DO CD2=1;
     end;
IF O_CDname = 'Charlotte' THEN DO CD3=1;
     end;
IF O_CDname = 'Gloucester' THEN DO CD4=1;
     end;
IF O_CDname = 'Kent' THEN DO CD5=1;
     end;
IF O_CDname = 'Kings' THEN DO CD6=1;
     end;
IF O_CDname = 'Madawaska' THEN DO CD7=1;
     end;
IF O_CDname = 'Northumberland' THEN DO CD8=1;
     end;
IF O_CDname = 'Queens' THEN DO CD9=1;
     end;
IF O_CDname = 'Restigouche' THEN DO CD10=1;
     end;
```

```
IF O_CDname = 'Saint John' THEN DO CD11=1;
     end;
IF O_CDname = 'Sunbury' THEN DO CD12=1;
     end;
IF O_CDname = 'Victoria' THEN DO CD13=1;
     end;
IF O_CDname = 'Westmorland' THEN DO CD14=1;
     end;

/* Urban/rural n=1 */
IF D_v3city_ind = 'U' THEN DO dCITYv3 = 1;
     end;
IF D_v3city_ind = 'R' THEN DO dCITYv3 = 0;
     end;

/* Urban/rural n=1 */
IF D_v7city_ind = 'U' THEN DO dCITYv7 = 1;
     end;
IF D_v7city_ind = 'R' THEN DO dCITYv7 = 0;
     end;

/* gender n=1 */
IF gender = '1' THEN DO gender2 = 1;
     end;
IF gender = '2' THEN DO gender2 = 0;
     end;

/* health region n=6 */
IF health_region = '1' THEN DO HR1=1;
     end;
IF health_region = '2' THEN DO HR2=1;
     end;
IF health_region = '3' THEN DO HR3=1;
     end;
IF health_region = '4' THEN DO HR4=1;
     end;
IF health_region = '5' THEN DO HR5=1;
     end;
IF health_region = '6' THEN DO HR6=1;
     end;
run;

data stats.Phybvar4reg1;
set stats.Phybvar4reg1;
IF Q1 = '.' THEN DO Q1=0;
     end;
IF Q2 = '.' THEN DO Q2=0;
     end;
IF Q3 = '.' THEN DO Q3=0;
```

```
     end;
IF Q4 = '.' THEN DO Q4=0;
     end;

IF MAJOR1 = '.' THEN DO MAJOR1 =0;
     end;
IF MAJOR2 = '.' THEN DO MAJOR2 =0;
     end;

IF SERVICE1 = '.' THEN DO SERVICE1 = 0;
     end;
IF SERVICE2 = '.' THEN DO SERVICE2 = 0;
     end;
IF SERVICE3 = '.' THEN DO SERVICE3 = 0;
     end;
IF SERVICE4 = '.' THEN DO SERVICE4 = 0;
     end;
IF SERVICE5 = '.' THEN DO SERVICE5 = 0;
     end;

IF SERV_CAT1 = '.' THEN DO SERV_CAT1 = 0;
     end;
IF SERV_CAT2 = '.' THEN DO SERV_CAT2 = 0;
     end;

IF CD1 = '.' THEN DO CD1=0;
     end;
IF CD2 = '.' THEN DO CD2=0;
     end;
IF CD3 = '.' THEN DO CD3=0;
     end;
IF CD4 = '.' THEN DO CD4=0;
     end;
IF CD5 = '.' THEN DO CD5=0;
     end;
IF CD6 = '.' THEN DO CD6=0;
     end;
IF CD7 = '.' THEN DO CD7=0;
     end;
IF CD8 = '.' THEN DO CD8=0;
     end;
IF CD9 = '.' THEN DO CD9=0;
     end;
IF CD10 = '.' THEN DO CD10=0;
     end;
IF CD11 = '.' THEN DO CD11=0;
     end;
IF CD12 = '.' THEN DO CD12=0;
     end;
```

```
IF CD13 = '.' THEN DO CD13=0;
     end;
IF CD14 = '.' THEN DO CD14=0;
     end;

IF HR1 = '.' THEN DO HR1=0;
     end;
IF HR2 = '.' THEN DO HR2=0;
     end;
IF HR3 = '.' THEN DO HR3=0;
     end;
IF HR4 = '.' THEN DO HR4=0;
     end;
IF HR5 = '.' THEN DO HR5=0;
     end;
IF HR6 = '.' THEN DO HR6=0;
     end;
run;
```

Correlation and Regression for the Physician Billing Dataset:

```
/* New Brunswick */
proc corr data = stats.Phybvar4reg1 ;
     var age max_beds gender2
     Q1 Q2 Q3 Q4
     PHY_TYPE1
     MAJOR1 MAJOR2
     dCITYv3 dCITYv7
     CD1 CD2 CD3 CD4 CD5 CD6 CD7 CD8 CD9 CD10 CD11 CD12 CD13
CD14
     HR1 HR2 HR3 HR4 HR5 HR6
     SERVICE1 SERVICE2 SERVICE3 SERVICE4 SERVICE5
     SERV_CAT1 SERV_CAT2 ;
     with len;
run;

proc corr data = stats.Phybvar4reg1;
var MAJOR1 MAJOR2 SERVICE1 SERVICE2 SERVICE3 SERVICE4 SERVICE5
SERV_CAT1 SERV_CAT2;
with PHY_TYPE1 ;
run;

proc reg data = stats.Phybvar4reg1;
     model len =
     MAJOR1 MAJOR2
     CD1 CD2 CD3 CD4 CD5 CD6 CD7 CD8 CD9 CD10 CD11 CD12 CD13
CD14
     max_beds
     Q1 Q2 Q3 Q4
     dCityv7
     age
     gender2
          /tol vif;
     title New Brunswick Regression Model;
run;
```

APPENDIX C: HEALTH DATA & GRAVITY MODEL ITERATIONS

C1: Health Data for Specialty Group Health Trips

	ALBERT COUNTY						CARLETON COUNTY					
	FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011			FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011		
Analysis Variable : len Disance (km)	DAD DATA			PHYSICAN BILLING			DAD DATA			PHYSICAN BILLING		
Specialty Group (Current)	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**
Anatomical Pathology	-	-	-	-	-	-	-	-	-	0	0	*
Anesthesiology	20	0.3%	57.40	1,645	1.5%	57.58	10	0.2%	154.50	1,330	1.1%	124.15
Cardiac Electrophysiology	5	0.1%	143.17	15	0.0%	141.09	0	0.0%		20	0.0%	230.68
Cardiac Surgery	**55**	**0.8%**	**145.88**	**245**	**0.2%**	**146.82**	**40**	**0.6%**	**229.98**	**230**	**0.2%**	**228.04**
Cardiology	**305**	**4.4%**	**114.78**	**3,830**	**3.6%**	**61.27**	**245**	**3.9%**	**227.47**	**810**	**0.7%**	**223.72**
Dermatology	5	0.1%	39.00	3,335	3.1%	49.45	-	-	-	350	0.3%	138.55
Diagnostic Radiology	**10**	**0.1%**	**42.60**	**10**	**0.0%**	**179.79**	**15**	**0.2%**	**224.90**	**19,115**	**15.5%**	**70.72**
Emergency Medicine	-	-	-	0	0.0%	*	-	-	-	10	0.0%	24.92
Endocrinology & Metabolis	0	0.0%	*	70	0.1%	55.86	15	0.2%	35.40	1,180	1.0%	23.08
Gastroenterology	20	0.3%	46.37	2,180	2.0%	49.64	10	0.2%	200.38	175	0.1%	138.79
General Internal Medicine	-	-	-	0	0.0%	*	0	0	*	50	0.0%	101.96
General Pathology	-	-	-	5	0.0%	43.00	-	-	-	-	-	-
General Practice	**2,445**	**35.3%**	**47.59**	**62,525**	**58.3%**	**40.77**	**3,025**	**47.6%**	**21.74**	**79,290**	**64.5%**	**26.12**
General Surgery	**420**	**6.1%**	**53.73**	**2,670**	**2.5%**	**50.26**	**790**	**12.4%**	**41.39**	**3,035**	**2.5%**	**38.80**
Geriatric Medicine	0	0.0%	*	75	0.1%	53.30	5	0.1%	108.50	100	0.1%	112.37
Geriatric Psychiatry	-	-	-	-	-	-	-	-	-	-	-	-
Gynecologic Oncology	15	0.2%	57.75	65	0.1%	52.22	5	0.1%	275.71	25	0.0%	279.45
Hematological Pathology	-	-	-	-	-	-	-	-	-	-	-	-
Hematology	80	1.2%	49.23	1,180	1.1%	51.11	10	0.2%	217.00	50	0.0%	279.80
Infectious Disease	15	0.2%	48.58	265	0.2%	46.56	-	-	-	5	0.0%	247.33
Internal Medicine	95	1.4%	49.61	2,995	2.8%	49.39	50	0.8%	114.35	540	0.4%	117.82
Maternal-Fetal Med	-	-	-	-	-	-	-	-	-	-	-	-
Medical Microbiology	-	-	-	50	0.0%	43.62	0	0		10	0.0%	223.88
Medical Oncology	5	0.1%	45.50	205	0.2%	51.10	5	0.1%	235.56	370	0.3%	155.19
Nephrology	105	1.5%	50.25	5,720	5.3%	51.01	5	0.1%	227.00	2,410	2.0%	230.42
Neurology	0	0.0%		645	0.6%	51.46	30	0.5%	229.73	125	0.1%	114.08
Neurosurgery	40	0.6%	47.62	145	0.1%	48.54	-	-	-	170	0.1%	227.36
Nuclear Medicine	-	-	-	5	0.0%	260.00	-	-	-	765	0.6%	21.13
Obstetrics & Gynecology	45	0.7%	48.04	415	0.4%	49.10	60	0.9%	110.44	475	0.4%	99.28
Ophthalmology	**1,180**	**17.1%**	**48.58**	**7,395**	**6.9%**	**72.50**	**730**	**11.5%**	**32.06**	**7,415**	**6.0%**	**30.09**
Orthopedic Surgery	**570**	**8.2%**	**52.60**	**2,675**	**2.5%**	**51.45**	**265**	**4.2%**	**102.40**	**1,045**	**0.8%**	**111.18**
Otol-Head & Neck Surgery	80	1.2%	49.71	1,480	1.4%	50.66	45	0.7%	177.53	420	0.3%	105.74
Pediatrics	-	-	-	5	0.0%	82.00	-	-	-	-	-	-
Physical Medicine & Rehab	-	-	-	105	0.1%	40.10	-	-	-	65	0.1%	109.02
Plastic Surgery	**185**	**2.7%**	**55.95**	**920**	**0.9%**	**49.54**	**70**	**1.1%**	**114.95**	**35**	**0.0%**	**217.70**
Psychiatry	10	0.1%	55.60	725	0.7%	47.88	5	0.1%	104.88	85	0.1%	173.21
Public Health & Preventiv	-	-	-	-	-	-	-	-	-	-	-	-
Radiation Oncology	15	0.2%	51.33	145	0.1%	50.46	25	0.4%	233.17	150	0.1%	233.22
Respirology	25	0.4%	47.96	445	0.4%	51.54	10	0.2%	147.43	125	0.1%	232.72
Rheumatology	15	0.2%	46.33	590	0.6%	49.04	-	-	-	545	0.4%	104.71
Thoracic Surgery	55	0.8%	49.21	240	0.2%	50.73	80	1.3%	110.59	270	0.2%	106.65
Urology	**865**	**12.5%**	**49.77**	**3,480**	**3.2%**	**49.62**	**685**	**10.8%**	**98.41**	**1,335**	**1.1%**	**113.83**
Vascular Surgery	**235**	**3.4%**	**49.68**	**750**	**0.7%**	**48.39**	**115**	**1.8%**	**114.17**	**845**	**0.7%**	**105.93**
Total	6,920			107,250			6,350			122,975		

	CHARLOTTE COUNTY						GLOUCESTER COUNTRY					
	FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011			FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011		
Analysis Variable : len Disance (km)	DAD DATA			PHYSICAN BILLING			DAD DATA			PHYSICAN BILLING		
Specialty Group (Current)	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**
Anatomical Pathology	-	-	-	-	-	-	-	-	-	0	0	*
Anesthesiology	10	0.1%	141.20	1,690	1.5%	166.67	25	0.1%	118.96	3,840	1.1%	102.56
Cardiac Electrophysiology	15	0.2%	166.57	45	0.0%	165.26	15	0.1%	338.37	35	0.0%	335.41
Cardiac Surgery	**40**	**0.5%**	**165.05**	**245**	**0.2%**	**164.46**	**120**	**0.6%**	**339.43**	**545**	**0.2%**	**341.17**
Cardiology	**395**	**4.8%**	**161.97**	**3,550**	**3.2%**	**150.94**	**1,385**	**7.1%**	**185.68**	**12,250**	**3.5%**	**124.83**
Dermatology	-	-	-	910	0.8%	137.06	-	-	-	1,645	0.5%	252.24
Diagnostic Radiology	**70**	**0.9%**	**161.96**	**330**	**0.3%**	**116.25**	**65**	**0.3%**	**221.58**	**430**	**0.1%**	**124.52**
Emergency Medicine	-	-	-	-	-	-	5	0.0%	29.80	-	-	-
Endocrinology & Metabolis	15	0.2%	158.67	180	0.2%	163.13	-	-	-	65	0.0%	180.57
Gastroenterology	15	0.2%	163.94	315	0.3%	135.71	325	1.7%	82.15	10,985	3.1%	41.05
General Internal Medicine	5	0.1%	135.00	5	0.0%	109.67	5	0.0%	88.33	15	0.0%	256.15
General Pathology	-	-	-	-	-	-	-	-	-	0	0.0%	*
General Practice	**3,705**	**45.3%**	**40.77**	**85,385**	**76.7%**	**68.59**	**8,310**	**42.4%**	**31.04**	**246,625**	**69.8%**	**21.69**
General Surgery	**715**	**8.8%**	**95.61**	**5,145**	**4.6%**	**79.41**	**1,590**	**8.1%**	**73.29**	**5,225**	**1.5%**	**79.98**
Geriatric Medicine	5	0.1%	140.00	145	0.1%	170.79	5	0.0%	239.50	130	0.0%	216.28
Geriatric Psychiatry	-	-	-	-	-	-	-	-	-	-	-	-
Gynecologic Oncology	0	0.0%	*	5	0.0%	267.56	30	0.2%	226.74	170	0.0%	222.68
Hematological Pathology	-	-	-	-	-	-	-	-	-	-	-	-
Hematology	20	0.2%	158.81	0	0.0%	*	100	0.5%	223.69	1,305	0.4%	224.14
Infectious Disease	-	-	-	30	0.0%	168.85	10	0.1%	218.60	25	0.0%	218.77
Internal Medicine	100	1.2%	159.47	210	0.2%	132.19	310	1.6%	71.20	2,965	0.8%	102.06
Maternal-Fetal Med	0	0	*	10	0.0%	141.33	-	-	-	-	-	-
Medical Microbiology	5	0.1%	150.00	55	0.0%	164.00	5	0.0%	240.00	90	0.0%	236.56
Medical Oncology	35	0.4%	163.26	555	0.5%	162.19	5	0.0%	269.67	110	0.0%	235.58
Nephrology	30	0.4%	155.00	1,335	1.2%	159.49	265	1.4%	65.14	15,935	4.5%	39.12
Neurology	0	0.0%	*	60	0.1%	151.20	5	0.0%	288.00	735	0.2%	223.60
Neurosurgery	45	0.6%	176.46	250	0.2%	166.12	115	0.6%	227.59	375	0.1%	226.37
Nuclear Medicine	-	-	-	45	0.0%	149.19	5	0.0%	36.14	5	0.0%	289.63
Obstetrics & Gynecology	90	1.1%	144.30	525	0.5%	182.88	205	1.0%	69.88	565	0.2%	128.36
Ophthalmology	**1,210**	**14.8%**	**143.79**	**2,430**	**2.2%**	**153.48**	**3,750**	**19.1%**	**50.72**	**29,640**	**8.4%**	**48.28**
Orthopedic Surgery	**505**	**6.2%**	**157.33**	**2,405**	**2.2%**	**161.87**	**915**	**4.7%**	**72.19**	**5,655**	**1.6%**	**61.51**
Otol-Head & Neck Surgery	70	0.9%	154.64	155	0.1%	112.66	380	1.9%	54.44	7,285	2.1%	47.69
Pediatrics	-	-	-	45	0.0%	141.30	-	-	-	70	0.0%	328.86
Physical Medicine & Rehab	40	0.5%	159.95	135	0.1%	142.19	0	0.0%	*	35	0.0%	231.19
Plastic Surgery	**100**	**1.2%**	**164.72**	**805**	**0.7%**	**147.88**	**215**	**1.1%**	**62.83**	**2,350**	**0.7%**	**90.43**
Psychiatry	10	0.1%	207.33	175	0.2%	165.23	85	0.4%	57.55	600	0.2%	61.21
Public Health & Preventiv	-	-	-	-	-	-	-	-	-	-	-	-
Radiation Oncology	-	-	-	385	0.3%	164.72	130	0.7%	223.09	490	0.1%	226.97
Respirology	50	0.6%	162.76	820	0.7%	145.31	30	0.2%	227.76	395	0.1%	229.58
Rheumatology	10	0.1%	155.07	90	0.1%	151.56	10	0.1%	232.67	845	0.2%	230.36
Thoracic Surgery	25	0.3%	164.50	10	0.0%	103.71	10	0.1%	222.40	60	0.0%	213.22
Urology	**760**	**9.3%**	**98.01**	**2,380**	**2.1%**	**158.63**	**885**	**4.5%**	**80.93**	**765**	**0.2%**	**232.25**
Vascular Surgery	**75**	**0.9%**	**154.31**	**400**	**0.4%**	**147.62**	**270**	**1.4%**	**222.59**	**845**	**0.2%**	**225.02**
Total	8,170			111,260			19,585			353,105		

	KENT COUNTY						KINGS COUNTY					
	FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011			FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011		
Analysis Variable : len Disance (km)	DAD DATA			PHYSICAN BILLING			DAD DATA			PHYSICAN BILLING		
Specialty Group (Current)	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**
Anatomical Pathology	-	-	-	-	-	-	-	-	-	5	0.0%	55.25
Anesthesiology	15	0.2%	42.08	2,065	1.5%	72.33	20	0.2%	60.00	3,155	1.5%	50.02
Cardiac Electrophysiology	10	0.1%	203.00	25	0.0%	193.96	25	0.2%	41.52	95	0.0%	39.66
Cardiac Surgery	**40**	**0.5%**	**196.12**	**270**	**0.2%**	**196.12**	**80**	**0.7%**	**48.93**	**415**	**0.2%**	**45.91**
Cardiology	**570**	**7.0%**	**103.68**	**6,680**	**4.8%**	**67.45**	**715**	**5.9%**	**45.84**	**5,840**	**2.7%**	**46.38**
Dermatology	-	-	-	2,550	1.8%	56.02	-	-	-	2,425	1.1%	25.67
Diagnostic Radiology	**10**	**0.1%**	**70.00**	**40**	**0.0%**	**180.43**	**210**	**1.7%**	**47.13**	**195**	**0.1%**	**111.70**
Emergency Medicine	-	-	-	-	-	-	-	-	-	0	0.0%	*
Endocrinology & Metabolis	5	0.1%	47.25	85	0.1%	47.74	35	0.3%	37.47	330	0.2%	43.80
Gastroenterology	120	1.5%	55.06	1,780	1.3%	57.18	40	0.3%	50.77	745	0.3%	44.06
General Internal Medicine	0	0.0%	*	30	0.0%	126.37	0	0.0%	*	10	0.0%	94.17
General Pathology	-	-	-	-	-	-	-	-	-	0	0.0%	*
General Practice	**2,060**	**25.4%**	**35.75**	**78,740**	**56.9%**	**34.56**	**3,145**	**26.0%**	**31.21**	**158,170**	**73.6%**	**30.08**
General Surgery	**915**	**11.3%**	**57.78**	**4,865**	**3.5%**	**60.61**	**1,780**	**14.7%**	**28.30**	**7,065**	**3.3%**	**36.45**
Geriatric Medicine	5	0.1%	40.50	90	0.1%	52.73	0	0.0%	*	555	0.3%	48.51
Geriatric Psychiatry	5	0.1%	51.00	5	0.0%	55.00	-	-	-	-	-	-
Gynecologic Oncology	20	0.2%	61.56	230	0.2%	56.75	5	0.0%	124.00	5	0.0%	110.63
Hematological Pathology	-	-	-	-	-	-	-	-	-	5	0.0%	36.00
Hematology	30	0.4%	58.38	1,175	0.8%	53.51	60	0.5%	44.68	65	0.0%	74.13
Infectious Disease	5	0.1%	45.33	75	0.1%	60.50	0	0.0%	*	120	0.1%	55.03
Internal Medicine	235	2.9%	57.55	3,615	2.6%	60.28	190	1.6%	42.26	620	0.3%	56.38
Maternal-Fetal Med	-	-	-	-	-	-	0	0	*	30	0.0%	29.74
Medical Microbiology	5	0.1%	31.00	360	0.3%	56.71	15	0.1%	35.85	160	0.1%	43.69
Medical Oncology	0	0.0%		170	0.1%	53.27	75	0.6%	50.46	1,230	0.6%	47.57
Nephrology	95	1.2%	61.82	5,300	3.8%	63.07	60	0.5%	46.16	4,855	2.3%	55.43
Neurology	5	0.1%	51.14	615	0.4%	57.18	5	0.0%	25.50	190	0.1%	82.12
Neurosurgery	50	0.6%	60.80	245	0.2%	60.48	105	0.9%	47.49	590	0.3%	52.39
Nuclear Medicine	-	-	-	5	0.0%	233.33	-	-	-	15	0.0%	188.14
Obstetrics & Gynecology	60	0.7%	57.49	705	0.5%	54.06	145	1.2%	45.79	1,015	0.5%	63.14
Ophthalmology	**1,865**	**23.0%**	**57.28**	**10,945**	**7.9%**	**71.46**	**2,520**	**20.8%**	**51.85**	**9,595**	**4.5%**	**41.43**
Orthopedic Surgery	**495**	**6.1%**	**58.16**	**3,460**	**2.5%**	**58.15**	**1,030**	**8.5%**	**48.26**	**5,085**	**2.4%**	**49.08**
Otol-Head & Neck Surgery	155	1.9%	56.19	2,775	2.0%	58.02	175	1.4%	39.17	45	0.0%	67.45
Pediatrics	-	-	-	5	0.0%	50.50	-	-	-	70	0.0%	36.14
Physical Medicine & Rehab	5	0.1%	218.00	40	0.0%	59.25	65	0.5%	41.71	535	0.2%	31.17
Plastic Surgery	**185**	**2.3%**	**64.75**	**970**	**0.7%**	**57.71**	**255**	**2.1%**	**44.20**	**2,495**	**1.2%**	**39.68**
Psychiatry	65	0.8%	57.43	675	0.5%	52.60	10	0.1%	37.92	545	0.3%	43.00
Public Health & Preventiv	-	-	-	-	-	-	-	-	-	-	-	-
Radiation Oncology	40	0.5%	57.48	265	0.2%	55.57	40	0.3%	69.95	730	0.3%	48.11
Respirology	60	0.7%	57.38	2,110	1.5%	55.78	50	0.4%	47.45	1,310	0.6%	33.20
Rheumatology	5	0.1%	49.78	1,250	0.9%	60.02	25	0.2%	57.08	310	0.1%	65.30
Thoracic Surgery	15	0.2%	59.06	110	0.1%	55.05	10	0.1%	73.60	30	0.0%	74.61
Urology	**780**	**9.6%**	**55.32**	**5,005**	**3.6%**	**55.56**	**1,050**	**8.7%**	**39.86**	**5,400**	**2.5%**	**48.62**
Vascular Surgery	**170**	**2.1%**	**59.32**	**990**	**0.7%**	**52.95**	**170**	**1.4%**	**56.73**	**800**	**0.4%**	**60.15**
Total	8,100			138,320			12,110			214,855		

	MADAWASKA COUNTY						NORTHUMBERLAND COUNTY					
	FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011			FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011		
Analysis Variable : len Disance (km)	DAD DATA			PHYSICAN BILLING			DAD DATA			PHYSICAN BILLING		
Specialty Group (Current)	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**
Anatomical Pathology		0.0%			0.0%		-	-	-	10	0.0%	41.82
Anesthesiology	5	0.1%	18.00	615	1.1%	122.64	5	0.0%	54.00	3,190	1.3%	132.43
Cardiac Electrophysiology	-	-	-	-	-	-	15	0.1%	238.95	55	0.0%	234.68
Cardiac Surgery	**5**	**0.1%**	**365.50**	**15**	**0.0%**	**381.12**	**85**	**0.6%**	**246.19**	**450**	**0.2%**	**242.35**
Cardiology	**15**	**0.2%**	**362.53**	**60**	**0.1%**	**360.56**	**705**	**4.8%**	**202.21**	**3,335**	**1.4%**	**147.33**
Dermatology	-	-	-	5	0.0%	355.14	-	-	-	1,975	0.8%	168.66
Diagnostic Radiology	**10**	**0.2%**	**369.50**	**140**	**0.2%**	**166.30**	**20**	**0.1%**	**172.17**	**3,390**	**1.4%**	**115.99**
Emergency Medicine	-	-	-	5	0.0%	48.00	-	-	-	-	-	-
Endocrinology & Metabolis	-	-	-	0	0.0%	*	5	0.0%	207.00	50	0.0%	165.52
Gastroenterology	-	-	-	15	0.0%	358.29	60	0.4%	128.11	1,180	0.5%	120.85
General Internal Medicine	5	0.1%	9.00	75	0.1%	15.22	35	0.2%	30.19	1,680	0.7%	179.18
General Pathology	-	-	-	-	-	-	-	-	-	0	0.0%	*
General Practice	**2,675**	**42.9%**	**19.24**	**41,950**	**74.8%**	**17.41**	**6,775**	**46.0%**	**32.05**	**159,710**	**66.4%**	**27.80**
General Surgery	**630**	**10.1%**	**19.26**	**3,035**	**5.4%**	**255.20**	**1,055**	**7.2%**	**45.08**	**9,845**	**4.1%**	**122.84**
Geriatric Medicine	-	-	-	5	0.0%	327.00	-	-	-	120	0.0%	134.70
Geriatric Psychiatry	-	-	-	-	-	-	-	-	-		0.0%	
Gynecologic Oncology	0	0.0%		15	0.0%	348.92	30	0.2%	143.89	175	0.1%	141.06
Hematological Pathology	-	-	-	-	-	-	-	-	-	-	-	-
Hematology	0	0.0%		25	0.0%	365.15	150	1.0%	144.37	1,710	0.7%	137.78
Infectious Disease	-	-	-	-	-	-	10	0.1%	126.20	100	0.0%	133.50
Internal Medicine	190	3.0%	17.03	1,345	2.4%	175.49	175	1.2%	77.41	6,450	2.7%	111.38
Maternal-Fetal Med	-	-	-	-	-	-	-	-	-	-	-	-
Medical Microbiology	-	-	-	30	0.1%	17.07	5	0.0%	146.00	115	0.0%	144.67
Medical Oncology	5	0.1%	359.00	10	0.0%	349.60	5	0.0%	147.88	450	0.2%	138.42
Nephrology	5	0.1%	369.00	35	0.1%	331.47	210	1.4%	131.04	12,680	5.3%	129.46
Neurology	-	-	-	710	1.3%	19.53	-	-	-	545	0.2%	141.08
Neurosurgery	25	0.4%	368.81	80	0.1%	366.18	95	0.6%	152.47	350	0.1%	155.71
Nuclear Medicine	-	-	-	5	0.0%	77.22	-	-	-	605	0.3%	138.78
Obstetrics & Gynecology	45	0.7%	14.06	160	0.3%	175.34	180	1.2%	37.29	680	0.3%	64.35
Ophthalmology	**870**	**14.0%**	**27.06**	**120**	**0.2%**	**252.90**	**2,345**	**15.9%**	**39.98**	**16,130**	**6.7%**	**44.62**
Orthopedic Surgery	**475**	**7.6%**	**20.93**	**2,030**	**3.6%**	**25.06**	**840**	**5.7%**	**58.60**	**5,405**	**2.2%**	**43.66**
Otol-Head & Neck Surgery	60	1.0%	32.68	420	0.7%	69.37	145	1.0%	88.93	2,385	1.0%	71.69
Pediatrics	-	-	-	0	0.0%	*	-	-	-	15	0.0%	246.64
Physical Medicine & Rehab	-	-	-	10	0.0%	268.92	5	0.0%	214.50	50	0.0%	122.04
Plastic Surgery	**10**	**0.2%**	**243.09**	**20**	**0.0%**	**348.13**	**140**	**0.9%**	**120.07**	**460**	**0.2%**	**121.63**
Psychiatry	45	0.7%	19.10	95	0.2%	20.09	70	0.5%	30.73	1,200	0.5%	37.73
Public Health & Preventiv	-	-	-	-	-	-	-	-	-	-	-	-
Radiation Oncology	35	0.6%	364.13	95	0.2%	369.53	80	0.5%	149.13	380	0.2%	147.70
Respirology	-	-	-	15	0.0%	354.93	25	0.2%	141.30	625	0.3%	143.02
Rheumatology	0	0.0%		55	0.1%	268.84	15	0.1%	171.93	780	0.3%	149.90
Thoracic Surgery	5	0.1%	226.00	5	0.0%	218.33	125	0.8%	135.28	435	0.2%	136.02
Urology	**840**	**13.5%**	**15.30**	**2,890**	**5.2%**	**19.23**	**1,045**	**7.1%**	**62.49**	**2,475**	**1.0%**	**77.17**
Vascular Surgery	**275**	**4.4%**	**23.29**	**2,015**	**3.6%**	**14.43**	**285**	**1.9%**	**137.21**	**1,200**	**0.5%**	**131.19**
Total	6,230			56,105			14,740			240,390		

	QUEENS COUNTY						RESTIGOUCHE COUNTY					
	FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011			FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011		
Analysis Variable : len Disance (km)	DAD DATA			PHYSICAN BILLING			DAD DATA			PHYSICAN BILLING		
Specialty Group (Current)	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**
Anatomical Pathology	-	-	-	-	-	-	-	-	-	-	-	-
Anesthesiology	15	0.4%	67.30	1,200	1.4%	74.16	160	1.5%	15.60	3,030	1.7%	91.66
Cardiac Electrophysiology	10	0.2%	131.55	30	0.0%	122.20	5	0.0%	398.25	15	0.0%	410.05
Cardiac Surgery	**30**	**0.7%**	**124.07**	**150**	**0.2%**	**117.61**	**50**	**0.5%**	**402.17**	**205**	**0.1%**	**404.64**
Cardiology	**175**	**4.3%**	**115.28**	**875**	**1.0%**	**101.87**	**285**	**2.6%**	**363.67**	**1,280**	**0.7%**	**278.29**
Dermatology	-	-	-	325	0.4%	79.28	-	-	-	175	0.1%	299.84
Diagnostic Radiology	**10**	**0.2%**	**95.69**	**9,550**	**11.0%**	**73.01**	**20**	**0.2%**	**304.24**	**33,740**	**18.5%**	**59.44**
Emergency Medicine	-	-	-	5	0.0%	50.00	-	-	-	-	-	-
Endocrinology & Metabolis	0	0.0%	*	30	0.0%	89.07	80	0.7%	25.65	1,385	0.8%	23.25
Gastroenterology	5	0.1%	98.60	360	0.4%	77.63	30	0.3%	196.94	875	0.5%	99.22
General Internal Medicine	20	0.5%	63.04	265	0.3%	60.83	0	0.0%	*	15	0.0%	172.93
General Pathology	-	-	-	-	-	-	-	-	-	5	0.0%	70.00
General Practice	**1,675**	**41.5%**	**55.23**	**62,090**	**71.2%**	**28.23**	**5,655**	**52.0%**	**25.48**	**98,440**	**54.0%**	**24.33**
General Surgery	**335**	**8.3%**	**60.25**	**1,010**	**1.2%**	**96.89**	**720**	**6.6%**	**59.65**	**8,925**	**4.9%**	**46.80**
Geriatric Medicine	-	-	-	195	0.2%	62.93	-	-	-	255	0.1%	306.22
Geriatric Psychiatry	-	-	-	-	-	-	-	-	-	-	-	-
Gynecologic Oncology	0	0.0%	*	5	0.0%	126.00	20	0.2%	296.82	70	0.0%	300.89
Hematological Pathology	-	-	-	-	-	-	-	-	-	-	-	-
Hematology	10	0.2%	126.86	25	0.0%	98.14	100	0.9%	293.25	1,020	0.6%	290.76
Infectious Disease	0	0.0%	*	5	0.0%	114.00	-	-	-	20	0.0%	311.10
Internal Medicine	70	1.7%	72.57	640	0.7%	71.32	225	2.1%	59.25	2,305	1.3%	92.54
Maternal-Fetal Med	-	-	-	0	0.0%	*	-	-	-	-	-	-
Medical Microbiology	-	-	-	0	0.0%	*	5	0.0%	312.00	45	0.0%	296.67
Medical Oncology	20	0.5%	116.00	265	0.3%	75.92	10	0.1%	335.14	85	0.0%	323.28
Nephrology	5	0.1%	100.14	225	0.3%	129.83	65	0.6%	155.88	5,365	2.9%	176.68
Neurology	-	-	-	180	0.2%	65.37	0	0.0%	*	410	0.2%	227.96
Neurosurgery	25	0.6%	115.18	145	0.2%	108.57	45	0.4%	286.60	215	0.1%	290.97
Nuclear Medicine	-	-	-	1,365	1.6%	145.82	10	0.1%	81.80	10	0.0%	256.67
Obstetrics & Gynecology	50	1.2%	61.60	510	0.6%	92.91	25	0.2%	98.71	135	0.1%	95.31
Ophthalmology	**610**	**15.1%**	**63.06**	**2,310**	**2.6%**	**68.17**	**1,305**	**12.0%**	**44.93**	**11,945**	**6.6%**	**40.83**
Orthopedic Surgery	**205**	**5.1%**	**68.97**	**1,010**	**1.2%**	**69.44**	**695**	**6.4%**	**60.87**	**1,030**	**0.6%**	**140.65**
Otol-Head & Neck Surgery	45	1.1%	80.32	320	0.4%	62.44	150	1.4%	85.87	1,555	0.9%	79.27
Pediatrics	-	-	-	10	0.0%	74.83	-	-	-	60	0.0%	402.94
Physical Medicine & Rehab	10	0.2%	84.40	115	0.1%	66.67	5	0.0%	362.00	15	0.0%	299.08
Plastic Surgery	**85**	**2.1%**	**65.07**	**150**	**0.2%**	**84.54**	**70**	**0.6%**	**117.45**	**790**	**0.4%**	**137.03**
Psychiatry	5	0.1%	58.20	200	0.2%	63.94	45	0.4%	8.69	1,715	0.9%	12.88
Public Health & Preventiv	-	-	-	-	-	-	-	-	-	-	-	-
Radiation Oncology	25	0.6%	117.76	175	0.2%	109.62	50	0.5%	295.33	255	0.1%	301.45
Respirology	10	0.2%	79.40	255	0.3%	159.46	20	0.2%	70.50	890	0.5%	54.87
Rheumatology	10	0.2%	115.00	375	0.4%	71.95	0	0.0%	*	235	0.1%	299.44
Thoracic Surgery	70	1.7%	59.61	290	0.3%	65.00	55	0.5%	296.77	205	0.1%	290.99
Urology	**420**	**10.4%**	**56.92**	**1,720**	**2.0%**	**67.70**	**715**	**6.6%**	**80.47**	**4,565**	**2.5%**	**285.00**
Vascular Surgery	**90**	**2.2%**	**70.38**	**805**	**0.9%**	**66.57**	**245**	**2.3%**	**251.14**	**1,000**	**0.5%**	**227.25**
Total	4,040			87,185			10,865			182,285		

	SAINT JOHN COUNTY						SUNBURY COUNTY					
	FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011			FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011		
Analysis Variable : len Disance (km)	DAD DATA			PHYSICAN BILLING			DAD DATA			PHYSICAN BILLING		
Specialty Group (Current)	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**
Anatomical Pathology	-	-	-	0	0.0%	*	-	-	-	-	-	-
Anesthesiology	10	0.1%	24.33	3,765	1.7%	50.40	10	0.2%	28.00	1,185	1.6%	41.32
Cardiac Electrophysiology	35	0.2%	35.13	150	0.1%	32.38	10	0.2%	125.80	10	0.0%	113.67
Cardiac Surgery	**115**	**0.8%**	**26.68**	**560**	**0.3%**	**31.54**	**25**	**0.6%**	**113.82**	**135**	**0.2%**	**114.89**
Cardiology	**1,155**	**7.6%**	**27.96**	**8,735**	**4.0%**	**27.18**	**170**	**3.9%**	**115.99**	**425**	**0.6%**	**113.48**
Dermatology		0.0%		4,125	1.9%	23.14	-	-	-	315	0.4%	50.30
Diagnostic Radiology	**265**	**1.7%**	**31.50**	**135**	**0.1%**	**112.39**	**5**	**0.1%**	**130.67**	**12,065**	**16.3%**	**44.41**
Emergency Medicine	-	-	-	-	-	-	-	-	-	5	0.0%	34.00
Endocrinology & Metabolis	70	0.5%	28.60	555	0.3%	29.46	0	0.0%	*	5	0.0%	107.88
Gastroenterology	55	0.4%	32.69	1,040	0.5%	27.67	5	0.1%	103.00	305	0.4%	46.88
General Internal Medicine	0	0.0%	*	5	0.0%	124.00	40	0.9%	24.78	640	0.9%	30.88
General Pathology	-	-	-	-	-	-	1,705	0.3960511	21.80	-	-	-
General Practice	**5,350**	**35.2%**	**27.80**	**143,675**	**65.6%**	**29.33**	**190**	**4.4%**	**30.75**	**45,665**	**61.9%**	**24.31**
General Surgery	**1,295**	**8.5%**	**30.35**	**6,480**	**3.0%**	**33.26**	**5**	**0.1%**	**34.00**	**440**	**0.6%**	**120.42**
Geriatric Medicine	10	0.1%	40.29	850	0.4%	33.27	-	-	-	160	0.2%	25.38
Geriatric Psychiatry	-	-	-	0	0.0%	*	-	-	-	-	-	-
Gynecologic Oncology	5	0.0%	158.50	10	0.0%	158.50	0	0.0%	*	10	0.0%	166.75
Hematological Pathology	-	-	-	0	0.0%		-	-	-	-	-	-
Hematology	55	0.4%	25.89	0	0.0%	*	0	0.0%	*	15	0.0%	155.40
Infectious Disease	10	0.1%	53.70	120	0.1%	37.44	-	-	-	10	0.0%	166.00
Internal Medicine	350	2.3%	31.66	735	0.3%	38.61	90	2.1%	31.13	645	0.9%	44.00
Maternal-Fetal Med	5	0.0%	19.75	55	0.0%	21.27	-	-	-	0	0.0%	*
Medical Microbiology	15	0.1%	26.31	195	0.1%	33.67	0	0.0%	*	-	-	-
Medical Oncology	75	0.5%	27.39	1,595	0.7%	28.59	15	0.3%	80.25	260	0.4%	45.49
Nephrology	95	0.6%	28.14	11,030	5.0%	27.68	10	0.2%	117.60	1,135	1.5%	116.32
Neurology	0	0.0%	*	195	0.1%	64.13	-	-	-	185	0.3%	33.34
Neurosurgery	95	0.6%	32.73	510	0.2%	30.51	15	0.3%	125.38	130	0.2%	121.22
Nuclear Medicine	-	-	-	10	0.0%	191.60	-	-	-	1,305	1.8%	116.19
Obstetrics & Gynecology	200	1.3%	30.30	1,100	0.5%	53.93	55	1.3%	20.91	570	0.8%	59.25
Ophthalmology	**2,930**	**19.3%**	**31.05**	**11,740**	**5.4%**	**28.75**	**685**	**15.9%**	**19.46**	**2,940**	**4.0%**	**34.44**
Orthopedic Surgery	**1,150**	**7.6%**	**30.42**	**5,985**	**2.7%**	**33.14**	**200**	**4.6%**	**29.38**	**840**	**1.1%**	**33.82**
Otol-Head & Neck Surgery	240	1.6%	26.59	0	0.0%	*	35	0.8%	64.75	425	0.6%	28.31
Pediatrics	-	-	-	135	0.1%	21.96	-	-	-	0	0.0%	*
Physical Medicine & Rehab	105	0.7%	30.54	690	0.3%	25.96	5	0.1%	122.00	115	0.2%	38.31
Plastic Surgery	**275**	**1.8%**	**29.17**	**2,975**	**1.4%**	**24.38**	**130**	**3.0%**	**25.98**	**25**	**0.0%**	**104.16**
Psychiatry	40	0.3%	26.57	985	0.5%	25.72	5	0.1%	26.00	45	0.1%	59.53
Public Health & Preventiv	-	-	-	-	-	-	-	-	-	-	-	-
Radiation Oncology	40	0.3%	27.30	1,005	0.5%	29.67	15	0.3%	125.07	105	0.1%	122.14
Respirology	55	0.4%	29.98	1,730	0.8%	25.68	5	0.1%	53.75	265	0.4%	227.35
Rheumatology	30	0.2%	30.59	355	0.2%	37.96	5	0.1%	159.00	445	0.6%	35.10
Thoracic Surgery	-	-	-	-	-	-	120	2.8%	27.05	365	0.5%	33.28
Urology	**885**	**5.8%**	**30.31**	**6,680**	**3.1%**	**32.10**	**645**	**15.0%**	**24.14**	**1,665**	**2.3%**	**30.71**
Vascular Surgery	**190**	**1.2%**	**38.88**	**945**	**0.4%**	**36.30**	**105**	**2.4%**	**24.03**	**970**	**1.3%**	**30.87**
Total	15,205			218,855			4,305			73,825		

	VICTORIA						WESTMORLAND					
	FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011			FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011		
Analysis Variable : len Disance (km)	DAD DATA			PHYSICAN BILLING			DAD DATA			PHYSICAN BILLING		
Specialty Group (Current)	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**
Anatomical Pathology	-	-	-	-	-	-	-	-	-	5	0.0%	11.67
Anesthesiology	0	0.0%	*	810	1.0%	163.32	105	0.3%	23.75	6,695	1.3%	35.25
Cardiac Electrophysiology	5	0.1%	274.00	5	0.0%	288.13	25	0.1%	171.14	85	0.0%	167.94
Cardiac Surgery	**25**	**0.5%**	**293.85**	**110**	**0.1%**	**295.11**	**145**	**0.5%**	**168.86**	**800**	**0.2%**	**167.95**
Cardiology	**155**	**2.8%**	**293.36**	**365**	**0.4%**	**287.37**	**1,530**	**5.1%**	**88.97**	**18,455**	**3.6%**	**31.35**
Dermatology	-	-	-	60	0.1%	198.20	10	0.0%	11.67	12,890	2.5%	18.30
Diagnostic Radiology	**10**	**0.2%**	**286.40**	**10,120**	**12.4%**	**104.24**	**25**	**0.1%**	**22.19**	**105**	**0.0%**	**204.81**
Emergency Medicine	0	0.0%	*	10	0.0%	79.30	-	-	-	0	0.0%	*
Endocrinology & Metabolis	10	0.2%	92.40	85	0.1%	109.99	15	0.0%	19.12	340	0.1%	19.25
Gastroenterology	5	0.1%	302.67	50	0.1%	180.53	285	0.9%	19.56	7,900	1.5%	15.90
General Internal Medicine	10	0.2%	85.40	55	0.1%	111.18	0	0.0%	*	15	0.0%	142.23
General Pathology	-	-	-	-	-	-	-	-	-	0	0.0%	*
General Practice	**2,485**	**45.6%**	**21.46**	**60,120**	**73.9%**	**22.14**	**10,580**	**35.2%**	**13.24**	**316,310**	**61.1%**	**20.08**
General Surgery	**950**	**17.4%**	**46.72**	**2,015**	**2.5%**	**125.34**	**2,155**	**7.2%**	**20.41**	**12,105**	**2.3%**	**20.99**
Geriatric Medicine	15	0.3%	172.00	20	0.0%	165.05	15	0.0%	15.68	420	0.1%	20.71
Geriatric Psychiatry	-	-	-	-	-	-	-	-	-	-	-	-
Gynecologic Oncology	-	-	-	5	0.0%	289.50	80	0.3%	17.78	585	0.1%	18.97
Hematological Pathology	-	-	-	-	-	-	-	-	-	-	-	-
Hematology	5	0.1%	290.75	10	0.0%	293.17	195	0.6%	18.78	5,570	1.1%	19.54
Infectious Disease	5	0.1%	298.50	-	-	-	60	0.2%	23.00	665	0.1%	13.76
Internal Medicine	65	1.2%	97.52	425	0.5%	164.36	565	1.9%	20.99	13,680	2.6%	16.70
Maternal-Fetal Med	-	-	-	0	0.0%	*	-	-	-	-	-	-
Medical Microbiology	-	-	-	10	0.0%	165.22	5	0.0%	23.17	750	0.1%	22.94
Medical Oncology	15	0.3%	285.20	175	0.2%	256.58	25	0.1%	21.86	880	0.2%	14.94
Nephrology	5	0.1%	308.33	145	0.2%	298.83	370	1.2%	17.32	25,200	4.9%	16.36
Neurology	-	-	-	175	0.2%	103.34	5	0.0%	26.20	2,990	0.6%	20.49
Neurosurgery	20	0.4%	294.96	85	0.1%	289.71	200	0.7%	19.48	840	0.2%	21.38
Nuclear Medicine	-	-	-	200	0.2%	78.64	5	0.0%	214.00	0	0.0%	*
Obstetrics & Gynecology	35	0.6%	124.20	225	0.3%	148.96	175	0.6%	14.43	2,460	0.5%	17.76
Ophthalmology	**630**	**11.6%**	**93.72**	**2,445**	**3.0%**	**112.26**	**5,990**	**20.0%**	**16.49**	**35,680**	**6.9%**	**41.21**
Orthopedic Surgery	**250**	**4.6%**	**128.77**	**810**	**1.0%**	**128.75**	**2,020**	**6.7%**	**21.31**	**10,855**	**2.1%**	**22.57**
Otol-Head & Neck Surgery	30	0.6%	125.64	235	0.3%	137.90	460	1.5%	15.14	8,250	1.6%	15.32
Pediatrics	-	-	-	5	0.0%	256.00	-	-	-	25	0.0%	28.43
Physical Medicine & Rehab	5	0.1%	288.00	20	0.0%	169.11	0	0.0%	*	205	0.0%	29.69
Plastic Surgery	**65**	**1.2%**	**182.07**	**30**	**0.0%**	**268.77**	**755**	**2.5%**	**21.08**	**3,660**	**0.7%**	**19.87**
Psychiatry	0	0.0%	*	15	0.0%	260.84	160	0.5%	10.00	2,630	0.5%	15.35
Public Health & Preventiv	-	-	-	-	-	-	-	-	-	-	-	-
Radiation Oncology	25	0.5%	292.46	100	0.1%	292.89	100	0.3%	19.34	755	0.1%	20.09
Respirology	5	0.1%	243.00	65	0.1%	271.33	160	0.5%	17.81	3,195	0.6%	21.50
Rheumatology	-	-	-	275	0.3%	166.53	80	0.3%	18.91	3,290	0.6%	22.51
Thoracic Surgery	35	0.6%	171.81	100	0.1%	162.31	160	0.5%	18.19	700	0.1%	17.79
Urology	**470**	**8.6%**	**133.07**	**1,180**	**1.4%**	**134.21**	**2,815**	**9.4%**	**19.81**	**15,085**	**2.9%**	**20.88**
Vascular Surgery	**115**	**2.1%**	**117.77**	**830**	**1.0%**	**121.40**	**745**	**2.5%**	**18.55**	**3,515**	**0.7%**	**19.68**
Total	5,450			81,390			30,020			517,590		

	YORK					
	FISCAL YEARS: 2005-2011			CALENDAR YEARS: 2006-2011		
Analysis Variable : len Disance (km)	DAD DATA			PHYSICAN BILLING		
Specialty Group (Current)	**N Obs**	**%**	**Mean (km)**	**N Obs**	**%**	**Mean (km)**
Anatomical Pathology	-	-	-	-	-	-
Anesthesiology	30	0.2%	42.69	5,620	1.5%	39.42
Cardiac Electrophysiology	30	0.2%	144.26	110	0.0%	143.26
Cardiac Surgery	**135**	**0.7%**	**146.99**	**615**	**0.2%**	**147.60**
Cardiology	**840**	**4.3%**	**149.30**	**1,985**	**0.5%**	**146.39**
Dermatology	-	-	-	1,880	0.5%	32.12
Diagnostic Radiology	**15**	**0.1%**	**142.33**	**51,705**	**13.7%**	**31.92**
Emergency Medicine	-	-	-	65	0.0%	16.05
Endocrinology & Metabolis	25	0.1%	40.60	265	0.1%	62.34
Gastroenterology	15	0.1%	118.86	1,665	0.4%	36.10
General Internal Medicine	115	0.6%	25.81	1,830	0.5%	20.75
General Pathology	-	-	-	-	-	-
General Practice	**8,165**	**42.3%**	**21.23**	**243,730**	**64.6%**	**22.51**
General Surgery	**995**	**5.2%**	**32.97**	**3,165**	**0.8%**	**105.44**
Geriatric Medicine	5	0.0%	38.00	950	0.3%	23.76
Geriatric Psychiatry	-	-	-	-	-	-
Gynecologic Oncology	10	0.1%	199.22	60	0.0%	190.08
Hematological Pathology	-	-	-	5	0.0%	137.00
Hematology	20	0.1%	153.86	55	0.0%	183.69
Infectious Disease	-	-	-	10	0.0%	180.11
Internal Medicine	380	2.0%	32.77	3,835	1.0%	34.24
Maternal-Fetal Med	-	-	-	5	0.0%	126.00
Medical Microbiology	0	0.0%	*	20	0.0%	152.35
Medical Oncology	15	0.1%	119.67	1,490	0.4%	49.01
Nephrology	10	0.1%	140.27	7,350	1.9%	148.73
Neurology	-	-	-	1,105	0.3%	21.81
Neurosurgery	85	0.4%	148.34	400	0.1%	147.40
Nuclear Medicine	-	-	-	6,315	1.7%	84.50
Obstetrics & Gynecology	245	1.3%	21.25	2,965	0.8%	51.72
Ophthalmology	**2,785**	**14.4%**	**37.27**	**14,740**	**3.9%**	**26.73**
Orthopedic Surgery	**890**	**4.6%**	**31.38**	**4,395**	**1.2%**	**30.20**
Otol-Head & Neck Surgery	140	0.7%	66.10	2,155	0.6%	21.05
Pediatrics	-	-	-	5	0.0%	185.00
Physical Medicine & Rehab	0	0.0%	*	575	0.2%	26.01
Plastic Surgery	**530**	**2.7%**	**34.66**	**80**	**0.0%**	**138.45**
Psychiatry	35	0.2%	26.74	855	0.2%	47.16
Public Health & Preventiv	-	-	-	5	0.0%	10.25
Radiation Oncology	110	0.6%	162.94	610	0.2%	152.68
Respirology	35	0.2%	54.84	905	0.2%	238.22
Rheumatology	10	0.1%	123.27	2,815	0.7%	25.06
Thoracic Surgery	450	2.3%	27.65	1,660	0.4%	23.37
Urology	**2,725**	**14.1%**	**31.57**	**7,460**	**2.0%**	**24.36**
Vascular Surgery	**475**	**2.5%**	**23.93**	**3,820**	**1.0%**	**24.82**
Total	19,320			377,280		

C2: DAD Data – Gravity Model Iterations with Health Factor

DAD DATA - 2011 HEALTH TRIPS
GRAVITY MODEL - HEALTH FACTORS

TRIP MATRIX
DAD Data - 2011 Health Trips

	Destination															
Origin	**Albert**	**Carleton**	**Charlotte**	**Gloucester**	**Kent**	**Kings**	**Madawaska**	**Northumberland**	**Queens**	**Restigouche**	**Saint John**	**Sunbury**	**Victoria**	**Westmorland**	**York**	**Total**
Albert	**1**	5	0	10	5	15	0	0	5	5	45	5	10	940	0	1,046
Carleton	0	**640**	0	10	0	5	0	5	5	0	95	80	70	15	135	1,060
Charlotte	0	5	**530**	5	0	5	0	5	0	5	530	80	10	5	15	1,195
Gloucester	0	15	5	**2,270**	0	5	0	65	5	15	140	5	10	245	5	2,785
Kent	0	10	5	5	**100**	5	0	30	0	0	40	0	0	1,010	5	1,210
Kings	0	0	0	10	5	**285**	0	5	5	5	1,360	10	5	65	10	1,765
Madawaska	0	10	5	0	0	0	**880**	0	0	10	5	5	45	45	0	1,005
Northumberland	0	0	5	135	5	0	0	**1,570**	5	5	130	50	10	215	95	2,225
Queens	0	5	5	5	0	35	0	5	**5**	5	105	215	10	15	220	630
Restigouche	0	5	5	165	5	0	95	10	5	**1,050**	65	0	5	135	5	1,550
Saint John	0	5	5	15	5	5	5	5	5	10	**2,165**	0	10	5	0	2,240
Sunbury	0	5	5	20	0	5	5	5	5	5	45	**325**	15	15	225	680
Victoria	0	75	5	20	5	5	165	0	0	0	35	40	**355**	15	50	770
Westmorland	0	10	10	15	0	15	0	0	0	10	165	5	10	**4,035**	0	4,275
York	0	110	10	5	5	0	5	0	5	10	230	870	5	25	**1,565**	2,845
Total	1	900	595	2,690	135	385	1,155	1,705	50	1,135	5,155	1,690	570	6,785	2,330	25,281

NOTES:
Any cell below 5 was randomly rounded between 0 and 5 therefore there is some error within areas with small trips
No Hospital in Albert County - Only Clincs/ Doctor Office
No Hospital in Queens County anymore - 50 trips may actually be much less considering rounding requirement

COST MATRIX
COST = DISTANCE (KM)

	Albert	**Carleton**	**Charlotte**	**Gloucester**	**Kent**	**Kings**	**Madawaska**	**Northumberland**	**Queens**	**Restigouche**	**Saint John**	**Sunbury**	**Victoria**	**Westmorland**	**York**
Albert	30	340	330	320	130	115	480	220	150	400	160	190	350	70	215
Carleton	340	30	150	300	250	250	160	200	200	250	240	150	100	300	130
Charlotte	330	150	40	380	250	180	300	280	190	410	110	140	240	300	120
Gloucester	320	300	380	50	200	310	330	150	300	200	380	300	300	260	280
Kent	130	250	250	200	30	130	330	100	100	280	200	330	220	80	140
Kings	115	250	180	310	130	30	400	200	50	400	50	100	330	80	120
Madawaska	480	160	300	330	330	400	30	240	350	220	380	300	100	450	280
Northumberland	220	200	280	150	100	200	240	40	160	240	260	180	140	180	150
Queens	150	200	190	300	100	50	350	160	40	380	90	55	260	130	80
Restigouche	400	250	410	200	280	400	220	240	380	40	470	400	190	350	380
Saint John	160	240	110	380	200	50	380	260	90	470	30	90	300	180	115
Sunbury	190	150	140	300	330	100	300	180	55	400	90	30	230	180	25
Victoria	350	100	240	300	220	330	100	140	260	190	300	230	30	340	215
Westmorland	70	300	300	260	80	80	450	180	130	350	180	180	340	50	200
York	215	130	120	280	140	120	280	150	80	380	115	25	215	200	40

Travel Estimates From Google Maps
Assume all Internal Trips are within 30-40km

DETERENCE / HEALTH FUNCTIONS

DETERANCE FUNCTION

Deterance function = EXP(cost*-B)

Assume B = 0.1

	DESTINATIONS														
ORIGINS	**Albert**	**Carleton**	**Charlotte**	**Gloucester**	**Kent**	**Kings**	**Madawaska**	**Northumberland**	**Queens**	**Restigouche**	**Saint John**	**Sunbury**	**Victoria**	**Westmorland**	**York**
Albert	0.049787068	1.7139E-15	4.6589E-15	1.2664E-14	2.2603E-06	1.013E-05	1.4252E-21	2.78947E-10	3.05902E-07	4.2484E-18	1.1254E-07	5.6028E-09	6.3051E-16	0.000911882	4.5991E-10
Carleton	1.71391E-15	0.04978707	3.059E-07	9.3576E-14	1.3888E-11	1.3888E-11	1.1254E-07	2.06115E-09	2.06115E-09	1.3888E-11	3.7751E-11	3.059E-07	4.54E-05	9.35762E-14	2.2603E-06
Charlotte	4.65889E-15	3.059E-07	0.01831564	3.1391E-17	1.3888E-11	1.523E-08	9.3576E-14	6.9144E-13	5.6028E-09	1.5629E-18	1.6702E-05	8.3153E-07	3.7751E-11	9.35762E-14	6.1442E-06
Gloucester	1.26642E-14	9.3576E-14	3.1391E-17	0.00673795	2.0612E-09	3.4425E-14	4.6589E-15	3.05902E-07	9.35762E-14	2.0612E-09	3.1391E-17	9.3576E-14	9.3576E-14	5.10909E-12	6.9144E-13
Kent	2.26033E-06	1.3888E-11	1.3888E-11	2.0612E-09	0.04978707	2.2603E-06	4.6589E-15	4.53999E-05	4.53999E-05	6.9144E-13	2.0612E-09	4.6589E-15	2.7895E-10	0.000335463	8.3153E-07
Kings	1.01301E-05	1.3888E-11	1.523E-08	3.4425E-14	2.2603E-06	0.04978707	4.2484E-18	2.06115E-09	0.006737947	4.2484E-18	0.00673795	4.54E-05	4.6589E-15	0.000335463	6.1442E-06
Madawaska	1.42516E-21	1.1254E-07	9.3576E-14	4.6589E-15	4.6589E-15	4.2484E-18	0.04978707	3.77513E-11	6.30512E-16	2.7895E-10	3.1391E-17	9.3576E-14	4.54E-05	2.86252E-20	6.9144E-13
Northumberland	2.78947E-10	2.0612E-09	6.9144E-13	3.059E-07	4.54E-05	2.0612E-09	3.7751E-11	0.018315639	1.12535E-07	3.7751E-11	5.1091E-12	1.523E-08	8.3153E-07	1.523E-08	3.059E-07
Queens	3.05902E-07	2.0612E-09	5.6028E-09	9.3576E-14	4.54E-05	0.00673795	6.3051E-16	1.12535E-07	0.018315639	3.1391E-17	0.00012341	0.00408677	5.1091E-12	2.26033E-06	0.00033546
Restigouche	4.24835E-18	1.3888E-11	1.5629E-18	2.0612E-09	6.9144E-13	4.2484E-18	2.7895E-10	3.77513E-11	3.13913E-17	0.01831564	3.874E-21	4.2484E-18	5.6028E-09	6.30512E-16	3.1391E-17
Saint John	1.12535E-07	3.7751E-11	1.6702E-05	3.1391E-17	2.0612E-09	0.00673795	3.1391E-17	5.10909E-12	0.00012341	3.874E-21	0.04978707	0.00012341	9.3576E-14	1.523E-08	1.013E-05
Sunbury	5.6028E-09	3.059E-07	8.3153E-07	9.3576E-14	4.6589E-15	4.54E-05	9.3576E-14	1.523E-08	0.004086771	4.2484E-18	0.00012341	0.04978707	1.0262E-10	1.523E-08	0.082085
Victoria	6.30512E-16	4.54E-05	3.7751E-11	9.3576E-14	2.7895E-10	4.6589E-15	4.54E-05	8.31529E-07	5.10909E-12	5.6028E-09	9.3576E-14	1.0262E-10	0.04978707	1.71391E-15	4.5991E-10
Westmorland	0.000911882	9.3576E-14	9.3576E-14	5.1091E-12	0.00033546	0.00033546	2.8625E-20	1.523E-08	2.26033E-06	6.3051E-16	1.523E-08	1.523E-08	1.7139E-15	0.006737947	2.0612E-09
York	4.59906E-10	2.2603E-06	6.1442E-06	6.9144E-13	8.3153E-07	6.1442E-06	6.9144E-13	3.05902E-07	0.000335463	3.1391E-17	1.013E-05	0.082085	4.5991E-10	2.06115E-09	0.01831564

K- FACTORS

Optimal Adjustment Factor used to account for factors other than travel impedance

	DESTINATIONS														
ORIGINS	**Albert**	**Carleton**	**Charlotte**	**Gloucester**	**Kent**	**Kings**	**Madawaska**	**Northumberland**	**Queens**	**Restigouche**	**Saint John**	**Sunbury**	**Victoria**	**Westmorland**	**York**
Albert	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Carleton	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Charlotte	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Gloucester	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Kent	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Kings	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Madawaska	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Northumberland	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Queens	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Restigouche	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Saint John	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Sunbury	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Victoria	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Westmorland	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
York	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

ORIGIN COUNTY FACTORS (h2)

R = County Travels on Average more then 50km for a Health Trip

	DESTINATIONS														
ORIGINS	**Albert**	**Carleton**	**Charlotte**	**Gloucester**	**Kent**	**Kings**	**Madawaska**	**Northumberland**	**Queens**	**Restigouche**	**Saint John**	**Sunbury***	**Victoria**	**Westmorland**	**York**
Albert	40	30	30	30	30	30	30	30	40	30	-50	30	30	-50	-50
Carleton	40	1	30	30	30	30	30	30	40	30	-50	30	30	-50	-50
Charlotte	40	30	1	30	30	30	30	30	40	30	-50	30	30	-50	-50
Gloucester	40	30	30	1	30	30	30	30	40	30	-50	30	30	-50	-50
Kent	40	30	30	30	1	30	30	30	40	30	-50	30	30	-50	-50
Kings	40	30	30	30	30	1	30	30	40	30	-50	30	30	-50	-50
Madawaska	40	30	30	30	30	30	1	30	40	30	-50	30	30	-50	-50
Northumberland	40	30	30	30	30	30	30	1	40	30	-50	30	30	-50	-50
Queens	40	30	30	30	30	30	30	30	40	30	-50	30	30	-50	-50
Restigouche	40	30	30	30	30	30	30	30	40	1	-50	30	30	-50	-50
Saint John	200	200	200	200	200	200	200	200	200	200	0	200	200	15	15
Sunbury	40	30	30	30	30	30	30	30	40	30	15	1	30	15	15
Victoria	40	30	30	30	30	30	30	30	40	30	-50	30	1	-50	-50
Westmorland	200	200	200	200	200	200	200	200	200	200	15	200	200	0	15
York	200	-70	200	200	200	200	200	200	200	200	15	15	200	15	0

	DESTINATIONS														
ORIGINS	**Albert**	**Carleton**	**Charlotte**	**Gloucester**	**Kent**	**Kings**	**Madawaska**	**Northumberland**	**Queens**	**Restigouche**	**Saint John**	**Sunbury**	**Victoria**	**Westmorland**	**York**
Albert	0.018315639	0.04978707	0.04978707	0.04978707	0.04978707	0.04978707	0.04978707	0.049787068	0.018315639	0.04978707	148.413159	0.04978707	0.04978707	148.4131591	148.413159
Carleton	0.018315639	0.90483742	0.04978707	0.04978707	0.04978707	0.04978707	0.04978707	0.049787068	0.018315639	0.04978707	148.413159	0.04978707	0.04978707	148.4131591	148.413159
Charlotte	0.018315639	0.04978707	0.90483742	0.04978707	0.04978707	0.04978707	0.04978707	0.049787068	0.018315639	0.04978707	148.413159	0.04978707	0.04978707	148.4131591	148.413159
Gloucester	0.018315639	0.04978707	0.04978707	0.90483742	0.04978707	0.04978707	0.04978707	0.049787068	0.018315639	0.04978707	148.413159	0.04978707	0.04978707	148.4131591	148.413159
Kent	0.018315639	0.04978707	0.04978707	0.04978707	0.90483742	0.04978707	0.04978707	0.049787068	0.018315639	0.04978707	148.413159	0.04978707	0.04978707	148.4131591	148.413159
Kings	0.018315639	0.04978707	0.04978707	0.04978707	0.04978707	0.90483742	0.04978707	0.049787068	0.018315639	0.04978707	148.413159	0.04978707	0.04978707	148.4131591	148.413159
Madawaska	0.018315639	0.04978707	0.04978707	0.04978707	0.04978707	0.04978707	0.90483742	0.049787068	0.018315639	0.04978707	148.413159	0.04978707	0.04978707	148.4131591	148.413159
Northumberland	0.018315639	0.04978707	0.04978707	0.04978707	0.04978707	0.04978707	0.04978707	0.904837418	0.018315639	0.04978707	148.413159	0.04978707	0.04978707	148.4131591	148.413159
Queens	0.018315639	0.04978707	0.04978707	0.04978707	0.04978707	0.04978707	0.04978707	0.049787068	0.018315639	0.04978707	148.413159	0.04978707	0.04978707	148.4131591	148.413159
Restigouche	0.018315639	0.04978707	0.04978707	0.04978707	0.04978707	0.04978707	0.04978707	0.049787068	0.018315639	0.90483742	148.413159	0.04978707	0.04978707	148.4131591	148.413159
Saint John	2.06115E-09	2.0612E-09	2.0612E-09	2.0612E-09	2.0612E-09	2.0612E-09	2.0612E-09	2.06115E-09	2.06115E-09	2.0612E-09	1	2.0612E-09	2.0612E-09	0.22313016	0.22313016
Sunbury	0.018315639	0.04978707	0.04978707	0.04978707	0.04978707	0.04978707	0.04978707	0.049787068	0.018315639	0.04978707	0.22313016	0.90483742	0.04978707	0.22313016	0.22313016
Victoria	0.018315639	0.04978707	0.04978707	0.04978707	0.04978707	0.04978707	0.04978707	0.049787068	0.018315639	0.04978707	148.413159	0.04978707	0.90483742	148.4131591	148.413159
Westmorland	2.06115E-09	2.0612E-09	2.0612E-09	2.0612E-09	2.0612E-09	2.0612E-09	2.0612E-09	2.06115E-09	2.06115E-09	2.0612E-09	0.22313016	2.0612E-09	2.0612E-09	1	0.22313016
York	2.06115E-09	1096.63316	2.0612E-09	2.0612E-09	2.0612E-09	2.0612E-09	2.0612E-09	2.06115E-09	2.06115E-09	2.0612E-09	0.22313016	0.22313016	2.0612E-09	0.22313016	1

Deterance function = EXP(cost*-B)

Assume B = 0.1

TRIP MATRIX - ITERATION 1

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS															Productions Total	Productions Target	Production Ratio
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York			
Albert	**0**	0	0	0	0	0	0	0	0	0	0	0	0	1,046	0	1,046	1,046	1.000
Carleton	0	**1,040**	0	0	0	0	0	0	0	0	0	0	0	0	20	1,060	1,060	1.000
Charlotte	0	0	**476**	0	0	0	0	0	0	0	617	0	0	0	103	1,195	1,195	1.000
Gloucester	0	0	0	**2,785**	0	0	0	0	0	0	0	0	0	0	0	2,785	2,785	1.000
Kent	0	0	0	0	**21**	0	0	0	0	0	0	0	0	1,188	1	1,210	1,210	1.000
Kings	0	0	0	0	0	**6**	0	0	0	0	1,651	0	0	108	1	1,765	1,765	1.000
Madawaska	0	0	0	0	0	0	**1,005**	0	0	0	0	0	0	0	0	1,005	1,005	1.000
Northumberland	0	0	0	0	0	0	0	**2,215**	0	0	0	0	0	1	8	2,225	2,225	1.000
Queens	0	0	0	0	0	0	0	0	**0**	0	279	1	0	7	343	630	630	1.000
Restigouche	0	0	0	0	0	0	0	0	0	**1,550**	0	0	0	0	0	1,550	1,550	1.000
Saint John	0	0	0	0	0	0	0	0	0	0	**2,240**	0	0	0	0	2,240	2,240	1.000
Sunbury	0	0	0	0	0	0	0	0	0	0	1	**435**	0	0	244	680	680	1.000
Victoria	0	0	0	0	0	0	0	0	0	0	0	0	**770**	0	0	770	770	1.000
Westmorland	0	0	0	0	0	0	0	0	0	0	0	0	0	**4,275**	0	4,275	4,275	1.000
York	0	84	0	0	0	0	0	0	0	0	0	1,161	0	0	**1,600**	2,845	2,845	1.000
Attractions Total	0	1,124	476	2,785	21	6	1,005	2,215	0	1,550	4,788	1,597	770	6,625	2,320			
Attractions Target	1	900	595	2,690	135	385	1,155	1,705	50	1,135	5,155	1,690	570	6,785	2,330		**25,281**	
Attraction Ratio	962.58	0.80	1.25	0.97	6.31	64.81	1.15	0.77	684.07	0.73	1.08	1.06	0.74	1.02	1.00			

IOWA STATE EXAMPLE

ATTRACTION GROWTH FACTORS

Determine new attractions (A) by multiplying by the ratio of the Target/ Total

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15
963	721	744	2,598	851	24,952	1,327	1,312	34,203	831	5,551	1,789	422	6,949	2,340

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS															Productions Total	Productions Target	Production Ratio
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York			
Albert	**1**	0	0	0	0	0	0	0	0	0	0	0	0	1,045	0	1,046	1,046	1.000
Carleton	0	**1,035**	0	0	0	0	0	0	0	0	0	0	0	0	25	1,060	1,060	1.000
Charlotte	0	0	**522**	0	0	0	0	0	0	0	583	0	0	0	90	1,195	1,195	1.000
Gloucester	0	0	0	**2,785**	0	0	0	0	0	0	0	0	0	0	0	2,785	2,785	1.000
Kent	0	0	0	0	**121**	0	0	0	0	0	0	0	0	1,088	1	1,210	1,210	1.000
Kings	0	0	0	0	0	**282**	0	0	1	0	1,394	0	0	87	1	1,765	1,765	1.000
Madawaska	0	0	0	0	0	0	**1,005**	0	0	0	0	0	0	0	0	1,005	1,005	1.000
Northumberland	0	0	0	0	0	0	0	**2,212**	0	0	0	0	0	2	11	2,225	2,225	1.000
Queens	0	0	0	0	0	22	0	0	**30**	0	266	1	0	6	305	630	630	1.000
Restigouche	0	0	0	0	0	0	0	0	0	**1,550**	0	0	0	0	0	1,550	1,550	1.000
Saint John	0	0	0	0	0	0	0	0	0	0	**2,240**	0	0	0	0	2,240	2,240	1.000
Sunbury	0	0	0	0	0	0	0	0	14	0	1	**434**	0	0	231	680	680	1.000
Victoria	0	0	0	0	0	0	0	0	0	0	0	0	**770**	0	0	770	770	1.000
Westmorland	0	0	0	0	0	0	0	0	0	0	0	0	0	**4,275**	0	4,275	4,275	1.000
York	0	66	0	0	0	0	0	0	0	0	0	1,204	0	0	**1,575**	2,845	2,845	1.000
Attractions Total	1	1,101	522	2,785	121	305	1,005	2,212	45	1,550	4,484	1,639	770	6,503	2,239			
Attractions Target	1	900	595	2,690	135	385	1,155	1,705	50	1,135	5,155	1,690	570	6,785	2,330		**25,281**	
Attraction Ratio	1.03	0.82	1.14	0.97	1.12	1.26	1.15	0.77	1.11	0.73	1.15	1.03	0.74	1.04	1.04			

PRODUCTION GROWTH FACTORS

Determine new attractions (A) by multiplying by the ratio of the Target/ Total

P1	1046
P2	1060
P3	1195
P4	2785
P5	1210
P6	1765
P7	1005
P8	2225
P9	630
P10	1550
P11	2240
P12	680
P13	770
P14	4275
P15	2845

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS															Productions Total	Productions Target	Production Ratio
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York			
Albert	1	0	0	0	0	0	0	0	0	0	0	0	0	1,045	0	1,046	1,046	1.000
Carleton	0	1,035	0	0	0	0	0	0	0	0	0	0	0	0	25	1,060	1,060	1.000
Charlotte	0	0	522	0	0	0	0	0	0	0	583	0	0	0	90	1,195	1,195	1.000
Gloucester	0	0	0	2,785	0	0	0	0	0	0	0	0	0	0	0	2,785	2,785	1.000
Kent	0	0	0	0	121	0	0	0	0	0	0	0	0	1,088	1	1,210	1,210	1.000
Kings	0	0	0	0	0	282	0	0	1	0	1,394	0	0	87	1	1,765	1,765	1.000
Madawaska	0	0	0	0	0	0	1,005	0	0	0	0	0	0	0	0	1,005	1,005	1.000
Northumberland	0	0	0	0	0	0	0	2,212	0	0	0	0	0	2	11	2,225	2,225	1.000
Queens	0	0	0	0	0	22	0	0	30	0	266	1	0	6	305	630	630	1.000
Restigouche	0	0	0	0	0	0	0	0	0	1,550	0	0	0	0	0	1,550	1,550	1.000
Saint John	0	0	0	0	0	0	0	0	0	0	2,240	0	0	0	0	2,240	2,240	1.000
Sunbury	0	0	0	0	0	0	0	0	14	0	1	434	0	0	231	680	680	1.000
Victoria	0	0	0	0	0	0	0	0	0	0	0	0	770	0	0	770	770	1.000
Westmorland	0	0	0	0	0	0	0	0	0	0	0	0	0	4,275	0	4,275	4,275	1.000
York	0	66	0	0	0	0	0	0	0	0	0	1,204	0	0	1,575	2,845	2,845	1.000
Attractions Total	1.0	1,101	522	2,785	121	304	1,005	2,212	45	1,550	4,484	1,639	770	6,503	2,239			
Attractions Target	1	900	595	2,690	135	385	1,155	1,705	50	1,135	5,155	1,690	570	6,785	2,330		25,281	
Attraction Ratio	1.00	0.82	1.14	0.97	1.12	1.27	1.15	0.77	1.11	0.73	1.15	1.03	0.74	1.04	1.04			

ITERATION 2

ATTRACTION GROWTH FACTORS

Determine new attractions (A) by multiplying by the ratio of the Target/ Total

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15
1	736	678	2,598	151	488	1,327	1,314	56	831	5,926	1,743	422	7,079	2,425

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS															Productions Total	Productions Target	Production Ratio
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York			
Albert	**0**	0	0	0	0	0	0	0	0	0	0	0	0	1,046	0	1,046	1,046	1.000
Carleton	0	**1,035**	0	0	0	0	0	0	0	0	0	0	0	0	25	1,060	1,060	1.000
Charlotte	0	0	**477**	0	0	0	0	0	0	0	624	0	0	0	94	1,195	1,195	1.000
Gloucester	0	0	0	**2,785**	0	0	0	0	0	0	0	0	0	0	0	2,785	2,785	1.000
Kent	0	0	0	0	**23**	0	0	0	0	0	0	0	0	1,186	1	1,210	1,210	1.000
Kings	0	0	0	0	0	**6**	0	0	0	0	1,660	0	0	99	1	1,765	1,765	1.000
Madawaska	0	0	0	0	0	0	**1,005**	0	0	0	0	0	0	0	0	1,005	1,005	1.000
Northumberland	0	0	0	0	0	0	0	**2,212**	0	0	0	0	0	2	11	2,225	2,225	1.000
Queens	0	0	0	0	0	0	0	0	**0**	0	295	1	0	6	328	630	630	1.000
Restigouche	0	0	0	0	0	0	0	0	0	**1,550**	0	0	0	0	0	1,550	1,550	1.000
Saint John	0	0	0	0	0	0	0	0	0	0	**2,240**	0	0	0	0	2,240	2,240	1.000
Sunbury	0	0	0	0	0	0	0	0	0	0	1	**434**	0	0	245	680	680	1.000
Victoria	0	0	0	0	0	0	0	0	0	0	0	0	**770**	0	0	770	770	1.000
Westmorland	0	0	0	0	0	0	0	0	0	0	0	0	0	**4,275**	0	4,275	4,275	1.000
York	0	66	0	0	0	0	0	0	0	0	0	1,162	0	0	**1,616**	2,845	2,845	1.000
Attractions Total	0	1,101	477	2,785	23	7	1,005	2,212	0	1,550	4,819	1,596	770	6,614	2,321			
Attractions Target	1	900	595	2,690	135	385	1,155	1,705	50	1,135	5,155	1,690	570	6,785	2,330		**25,281**	
Attraction Ratio	1,004.34	0.82	1.25	0.97	5.90	58.32	1.15	0.77	661.09	0.73	1.07	1.06	0.74	1.03	1.00			

PRODUCTION GROWTH FACTORS

Determine new attractions (A) by multiplying by the ratio of the Target/ Total

P1	1046
P2	1060
P3	1195
P4	2785
P5	1210
P6	1765
P7	1005
P8	2225
P9	630
P10	1550
P11	2240
P12	680
P13	770
P14	4275
P15	2845

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS															Productions Total	Productions Target	Production Ratio
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York			
Albert	0	0	0	0	0	0	0	0	0	0	0	0	0	1,046	0	1,046	1,046	1.000
Carleton	0	1,035	0	0	0	0	0	0	0	0	0	0	0	0	25	1,060	1,060	1.000
Charlotte	0	0	477	0	0	0	0	0	0	0	624	0	0	0	94	1,195	1,195	1.000
Gloucester	0	0	0	2,785	0	0	0	0	0	0	0	0	0	0	0	2,785	2,785	1.000
Kent	0	0	0	0	23	0	0	0	0	0	0	0	0	1,186	1	1,210	1,210	1.000
Kings	0	0	0	0	0	6	0	0	0	0	1,660	0	0	99	1	1,766	1,765	0.999
Madawaska	0	0	0	0	0	0	1,005	0	0	0	0	0	0	0	0	1,005	1,005	1.000
Northumberland	0	0	0	0	0	0	0	2,212	0	0	0	0	0	2	11	2,225	2,225	1.000
Queens	0	0	0	0	0	0	0	0	0	0	295	1	0	6	328	630	630	1.000
Restigouche	0	0	0	0	0	0	0	0	0	1,550	0	0	0	0	0	1,550	1,550	1.000
Saint John	0	0	0	0	0	0	0	0	0	0	2,240	0	0	0	0	2,240	2,240	1.000
Sunbury	0	0	0	0	0	0	0	0	0	0	1	434	0	0	245	680	680	1.000
Victoria	0	0	0	0	0	0	0	0	0	0	0	0	770	0	0	770	770	1.000
Westmorland	0	0	0	0	0	0	0	0	0	0	0	0	0	4,275	0	4,275	4,275	1.000
York	0	66	0	0	0	0	0	0	0	0	0	1,162	0	0	1,616	2,844	2,845	1.000
Attractions Total	0.0000	1,101	477	2,785	23	6	1,005	2,212	0	1,550	4,820	1,597	770	6,614	2,321			
Attractions Target	1	900	595	2,690	135	385	1,155	1,705	50	1,135	5,155	1,690	570	6,785	2,330		25,281	
Attraction Ratio	#DIV/0!	0.82	1.25	0.97	5.87	64.17	1.15	0.77	#DIV/0!	0.73	1.07	1.06	0.74	1.03	1.00			

COMPARING BETWEEN PREDICTED AND ACTUAL

DIFFERENCE (PREDITCTED - ACTUAL)

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS															Productions Total	Productions Target
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York		
Albert	0	-5	0	-10	-5	-15	0	0	-5	-5	-45	-5	-10	105	0		
Carleton	0	395	0	-10	0	-5	0	-5	-5	0	-95	-80	-70	-15	-110		
Charlotte	0	-5	-8	-5	0	-5	0	-5	0	-5	53	-80	-10	-5	75		
Gloucester	0	-15	-5	515	0	-5	0	-65	-5	-15	-140	-5	-10	-245	-5		
Kent	0	-10	-5	-5	21	-5	0	-30	0	0	-40	0	0	78	-4		
Kings	0	0	0	-10	-5	-3	0	-5	-4	-5	34	-10	-5	22	-9		
Madawaska	0	-10	-5	0	0	0	125	0	0	-10	-5	-5	-45	-45	0		
Northumberland	0	0	-5	-135	-5	0	0	642	-5	-5	-130	-50	-10	-213	-84		
Queens	0	-5	-5	-5	0	-13	0	-5	25	-5	161	-214	-10	-9	85		
Restigouche	0	-5	-5	-165	-5	0	-95	-10	-5	500	-65	0	-5	-135	-5		
Saint John	0	-5	-5	-15	-5	-5	-5	-5	-5	-10	75	0	-10	-5	0		
Sunbury	0	-5	-5	-20	0	-5	-5	-5	9	-5	-44	109	-15	-15	6		
Victoria	0	-75	-5	-20	-5	-5	-165	0	0	0	-35	-40	415	-15	-50		
Westmorland	0	-10	-10	-15	0	-15	0	0	0	-10	-165	-5	-10	240	0		
York	0	-44	-10	-5	-5	0	-5	0	-5	-10	-230	334	-5	-25	10		
Attractions Total																	
Attractions Target																	

C3: Physician Billing Data – Gravity Model Iterations with Health Factor

PHYSICIAN BILLING DATA - 2011 HEALTH TRIPS
GRAVITY MODEL - HEALTH FACTORS

TRIP MATRIX

DAD Data - 2011 Health Trips

	Destination															
Origin	**Albert**	**Carleton**	**Charlotte**	**Gloucester**	**Kent**	**Kings**	**Madawaska**	**Northumberland**	**Queens**	**Restigouche**	**Saint John**	**Sunbury**	**Victoria**	**Westmorland**	**York**	**Total**
Albert	**3,155**	0	5	10	45	45	15	10	0	0	535	5	5	15,540	35	19,405
Carleton	10	**13,895**	1,480	0	5	55	65	60	5	15	965	25	875	110	1,930	19,495
Charlotte	5	210	**9,230**	0	0	2,075	20	0	10	5	6,310	30	0	100	370	18,365
Gloucester	0	15	0	**53,615**	5	30	10	715	0	335	500	0	5	3,110	695	59,035
Kent	255	35	0	30	**5,195**	10	20	475	10	5	415	5	5	15,515	35	22,010
Kings	35	5	1,115	15	10	**13,595**	20	5	265	15	13,040	30	5	1,205	420	29,780
Madawaska	0	10	5	20	5	15	**6,850**	5	5	60	15	15	615	445	25	8,090
Northumberland	20	975	5	2,205	30	155	30	**28,315**	0	15	510	10	5	5,875	2,635	40,785
Queens	5	460	35	0	5	505	270	75	**5,700**	30	910	470	0	250	3,555	12,270
Restigouche	0	85	5	5,005	5	10	350	85	5	**22,130**	425	5	0	1,780	675	30,565
Saint John	5	0	805	5	5	5,975	30	0	25	30	**29,265**	15	5	395	295	36,855
Sunbury	5	135	60	5	0	35	550	50	1,510	25	415	**3,195**	20	110	5,120	11,235
Victoria	0	1,365	5	35	5	35	435	45	5	30	295	15	**8,655**	225	810	11,960
Westmorland	2,490	10	0	55	525	95	210	45	35	10	1,955	30	25	**81,480**	95	87,060
York	5	2,465	1,410	5	5	220	565	1,120	300	150	2,920	4,925	45	460	**51,410**	66,005
Total	5,990	19,665	14,160	61,005	5,845	22,855	9,440	31,005	7,875	22,855	58,475	8,775	10,265	126,600	68,105	472,915

COST MATRIX

COST = DISTANCE (KM)

	Albert	**Carleton**	**Charlotte**	**Gloucester**	**Kent**	**Kings**	**Madawaska**	**Northumberland**	**Queens**	**Restigouche**	**Saint John**	**Sunbury**	**Victoria**	**Westmorland**	**York**
Albert	30	340	330	320	130	115	480	220	150	400	160	190	350	70	215
Carleton	340	30	150	300	250	250	160	200	200	250	240	150	100	300	130
Charlotte	330	150	60	380	250	180	300	280	190	410	110	140	240	300	120
Gloucester	320	300	380	50	200	310	330	150	300	200	380	300	300	260	280
Kent	130	250	250	200	30	130	330	100	100	280	200	330	220	80	140
Kings	115	250	180	310	130	30	400	200	50	400	50	100	330	80	120
Madawaska	480	160	300	330	330	400	30	240	350	220	380	300	100	450	280
Northumberland	220	200	280	150	100	200	240	40	160	240	260	180	140	180	150
Queens	150	200	190	300	100	50	350	160	40	380	90	55	260	130	80
Restigouche	400	250	410	200	280	400	220	240	380	40	470	400	190	350	380
Saint John	160	240	110	380	200	50	380	260	90	470	30	90	300	180	115
Sunbury	190	150	140	300	330	100	300	180	55	400	90	30	230	180	25
Victoria	350	100	240	300	220	330	100	140	260	190	300	230	30	340	215
Westmorland	70	300	300	260	80	80	450	180	130	350	180	180	340	50	200
York	215	130	120	280	140	120	280	150	80	380	115	25	215	200	40

Travel Estimates From Google Maps
Assume all Internal Trips are within 30-40km

DETERENCE / HEALTH FUNCTIONS

DETERANCE FUNCTION

Deterance function = EXP(cost*-B)

Assume B = 0.1

	DESTINATIONS														
ORIGINS	**Albert**	**Carleton**	**Charlotte**	**Gloucester**	**Kent**	**Kings**	**Madawaska**	**Northumberland**	**Queens**	**Restigouche**	**Saint John**	**Sunbury**	**Victoria**	**Westmorland**	**York**
Albert	0.04978707	1.7139E-15	4.6589E-15	1.26642E-14	2.2603E-06	1.013E-05	1.42516E-21	2.78947E-10	3.059E-07	4.24835E-18	1.1254E-07	5.6028E-09	6.3051E-16	0.000911882	4.5991E-10
Carleton	1.7139E-15	0.04978707	3.059E-07	9.35762E-14	1.3888E-11	1.3888E-11	1.12535E-07	2.06115E-09	2.0612E-09	1.38879E-11	3.7751E-11	3.059E-07	4.54E-05	9.35762E-14	2.2603E-06
Charlotte	4.6589E-15	3.059E-07	0.00247875	3.13913E-17	1.3888E-11	1.523E-08	9.35762E-14	6.9144E-13	5.6028E-09	1.56288E-18	1.6702E-05	8.3153E-07	3.7751E-11	9.35762E-14	6.1442E-06
Gloucester	1.2664E-14	9.3576E-14	3.1391E-17	0.006737947	2.0612E-09	3.4425E-14	4.65889E-15	3.05902E-07	9.3576E-14	2.06115E-09	3.1391E-17	9.3576E-14	9.3576E-14	5.10909E-12	6.9144E-13
Kent	2.2603E-06	1.3888E-11	1.3888E-11	2.06115E-09	0.04978707	2.2603E-06	4.65889E-15	4.53999E-05	4.54E-05	6.9144E-13	2.0612E-09	4.6589E-15	2.7895E-10	0.000335463	8.3153E-07
Kings	1.013E-05	1.3888E-11	1.523E-08	3.44248E-14	2.2603E-06	0.04978707	4.24835E-18	2.06115E-09	0.00673795	4.24835E-18	0.00673795	4.54E-05	4.6589E-15	0.000335463	6.1442E-06
Madawaska	1.4252E-21	1.1254E-07	9.3576E-14	4.65889E-15	4.6589E-15	4.2484E-18	0.049787068	3.77513E-11	6.3051E-16	2.78947E-10	3.1391E-17	9.3576E-14	4.54E-05	2.86252E-20	6.9144E-13
Northumberland	2.7895E-10	2.0612E-09	6.9144E-13	3.05902E-07	4.54E-05	2.0612E-09	3.77513E-11	0.018315639	1.1254E-07	3.77513E-11	5.1091E-12	1.523E-08	8.3153E-07	1.523E-08	3.059E-07
Queens	3.059E-07	2.0612E-09	5.6028E-09	9.35762E-14	4.54E-05	0.00673795	6.30512E-16	1.12535E-07	0.01831564	3.13913E-17	0.00012341	0.00408677	5.1091E-12	2.26033E-06	0.00033546
Restigouche	4.2484E-18	1.3888E-11	1.5629E-18	2.06115E-09	6.9144E-13	4.2484E-18	2.78947E-10	3.77513E-11	3.1391E-17	0.018315639	3.874E-21	4.2484E-18	5.6028E-09	6.30512E-16	3.1391E-17
Saint John	1.1254E-07	3.7751E-11	1.6702E-05	3.13913E-17	2.0612E-09	0.00673795	3.13913E-17	5.10909E-12	0.00012341	3.874E-21	0.04978707	0.00012341	9.3576E-14	1.523E-08	1.013E-05
Sunbury	5.6028E-09	3.059E-07	8.3153E-07	9.35762E-14	4.6589E-15	4.54E-05	9.35762E-14	1.523E-08	0.00408677	4.24835E-18	0.00012341	0.04978707	1.0262E-10	1.523E-08	0.082085
Victoria	6.3051E-16	4.54E-05	3.7751E-11	9.35762E-14	2.7895E-10	4.6589E-15	4.53999E-05	8.31529E-07	5.1091E-12	5.6028E-09	9.3576E-14	1.0262E-10	0.04978707	1.71391E-15	4.5991E-10
Westmorland	0.00091188	9.3576E-14	9.3576E-14	5.10909E-12	0.00033546	0.00033546	2.86252E-20	1.523E-08	2.2603E-06	6.30512E-16	1.523E-08	1.523E-08	1.7139E-15	0.006737947	2.0612E-09
York	4.5991E-10	2.2603E-06	6.1442E-06	6.9144E-13	8.3153E-07	6.1442E-06	6.9144E-13	3.05902E-07	0.00033546	3.13913E-17	1.013E-05	0.082085	4.5991E-10	2.06115E-09	0.01831564

K- FACTORS

Optimal Adjustment Factor used to account for factors other than travel impedance

	DESTINATIONS														
ORIGINS	**Albert**	**Carleton**	**Charlotte**	**Gloucester**	**Kent**	**Kings**	**Madawaska**	**Northumberland**	**Queens**	**Restigouche**	**Saint John**	**Sunbury**	**Victoria**	**Westmorland**	**York**
Albert	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Carleton	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Charlotte	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Gloucester	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Kent	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Kings	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Madawaska	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Northumberland	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Queens	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Restigouche	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Saint John	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Sunbury	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Victoria	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Westmorland	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
York	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

ORIGIN COUNTY FACTORS (h2)

	DESTINATIONS														
ORIGINS	**Albert**	**Carleton**	**Charlotte**	**Gloucester**	**Kent**	**Kings**	**Madawaska**	**Northumberland**	**Queens**	**Restigouche**	**Saint John**	**Sunbury***	**Victoria**	**Westmorland**	**York**
Albert	0	5	5	5	5	5	5	5	5	5	-20	5	5	-20	-20
Carleton	5	0	5	5	5	5	5	5	5	5	-20	5	5	-20	-20
Charlotte	5	5	0	5	5	5	5	5	5	5	-20	5	5	-20	-20
Gloucester	5	5	5	0	5	5	5	5	5	5	-20	5	5	-20	-20
Kent	5	5	5	5	0	5	5	5	5	5	-20	5	5	-20	-20
Kings	5	5	5	5	5	0	5	5	5	5	-20	5	5	-20	-20
Madawaska	5	5	5	5	5	5	0	5	5	5	-20	5	5	-20	-20
Northumberland	5	5	5	5	5	5	5	0	5	5	-20	5	5	-20	-20
Queens	5	5	5	5	5	5	5	5	0	5	-20	5	5	-20	-20
Restigouche	5	5	5	5	5	5	5	5	5	0	-20	5	5	-20	-20
Saint John	20	20	20	20	20	20	20	20	20	20	0	20	20	5	5
Sunbury	5	5	5	5	5	5	5	5	5	5	5	0	5	5	5
Victoria	5	5	5	5	5	5	5	5	5	5	-20	5	0	-20	-20
Westmorland	20	20	20	20	20	20	20	20	20	20	5	20	20	0	5
York	20	20	20	20	20	20	20	20	20	20	5	5	20	5	0

	DESTINATIONS														
ORIGINS	**Albert**	**Carleton**	**Charlotte**	**Gloucester**	**Kent**	**Kings**	**Madawaska**	**Northumberland**	**Queens**	**Restigouche**	**Saint John**	**Sunbury**	**Victoria**	**Westmorland**	**York**
Albert	1	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	7.3890561	0.60653066	0.60653066	7.389056099	7.3890561
Carleton	0.60653066	1	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	7.3890561	0.60653066	0.60653066	7.389056099	7.3890561
Charlotte	0.60653066	0.60653066	1	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	7.3890561	0.60653066	0.60653066	7.389056099	7.3890561
Gloucester	0.60653066	0.60653066	0.60653066	1	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	7.3890561	0.60653066	0.60653066	7.389056099	7.3890561
Kent	0.60653066	0.60653066	0.60653066	0.60653066	1	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	7.3890561	0.60653066	0.60653066	7.389056099	7.3890561
Kings	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	1	0.60653066	0.60653066	0.60653066	0.60653066	7.3890561	0.60653066	0.60653066	7.389056099	7.3890561
Madawaska	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	1	0.60653066	0.60653066	0.60653066	7.3890561	0.60653066	0.60653066	7.389056099	7.3890561
Northumberland	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	1	0.60653066	0.60653066	7.3890561	0.60653066	0.60653066	7.389056099	7.3890561
Queens	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	1	0.60653066	7.3890561	0.60653066	0.60653066	7.389056099	7.3890561
Restigouche	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	1	7.3890561	0.60653066	0.60653066	7.389056099	7.3890561
Saint John	0.13533528	0.13533528	0.13533528	0.135335283	0.13533528	0.13533528	0.135335283	0.135335283	0.13533528	0.135335283	1	0.13533528	0.13533528	0.60653066	0.60653066
Sunbury	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	1	0.60653066	0.60653066	0.60653066
Victoria	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	0.60653066	7.3890561	0.60653066	1	7.389056099	7.3890561
Westmorland	0.13533528	0.13533528	0.13533528	0.135335283	0.13533528	0.13533528	0.135335283	0.135335283	0.13533528	0.135335283	0.60653066	0.13533528	0.13533528	1	0.60653066
York	0.13533528	0.13533528	0.13533528	0.135335283	0.13533528	0.13533528	0.135335283	0.135335283	0.13533528	0.135335283	0.60653066	0.60653066	0.13533528	0.60653066	1

Deterance function = EXP(cost*-B)

Assume B = 0.1

TRIP MATRIX - ITERATION 1

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS															Productions Total	Productions Target	Production Ratio
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York			
Albert	**5,026**	0	0	0	0	2	0	0	0	0	1	0	0	14,376	0	19,405	19,405	1.000
Carleton	0	**19,467**	0	0	0	0	0	0	0	0	0	0	6	0	23	19,495	19,495	1.000
Charlotte	0	1	**14,193**	0	0	0	0	0	0	0	2,918	2	0	0	1,250	18,365	18,365	1.000
Gloucester	0	0	0	**59,034**	0	0	0	1	0	0	0	0	0	0	0	59,035	59,035	1.000
Kent	0	0	0	0	**10,563**	1	0	31	8	0	0	0	0	11,391	15	22,010	22,010	1.000
Kings	0	0	0	0	0	**7,704**	0	0	218	0	19,711	2	0	2,125	21	29,780	29,780	1.000
Madawaska	0	0	0	0	0	0	**8,085**	0	0	0	0	0	5	0	0	8,090	8,090	1.000
Northumberland	0	0	0	1	12	0	0	**40,760**	0	0	0	0	0	1	11	40,785	40,785	1.000
Queens	0	0	0	0	4	2,369	0	0	**3,658**	0	1,352	552	0	54	4,281	12,270	12,270	1.000
Restigouche	0	0	0	0	0	0	0	0	0	**30,565**	0	0	0	0	0	30,565	30,565	1.000
Saint John	0	0	0	0	0	262	0	0	2	0	**36,584**	2	0	0	5	36,855	36,855	1.000
Sunbury	0	0	0	0	0	2	0	0	57	0	13	**1,274**	0	0	9,889	11,235	11,235	1.000
Victoria	0	13	0	0	0	0	6	0	0	0	0	0	**11,941**	0	0	11,960	11,960	1.000
Westmorland	75	0	0	0	27	106	0	0	0	0	0	0	0	**86,852**	0	87,060	87,060	1.000
York	0	0	0	0	0	1	0	0	14	0	14	17,113	0	0	**48,862**	66,005	66,005	1.000
Attractions Total	5,102	19,481	14,194	59,035	10,606	10,446	8,091	40,792	3,957	30,565	60,593	18,944	11,952	114,798	64,358			
Attractions Target	5,990	19,665	14,160	61,005	5,845	22,855	9,440	31,005	7,875	22,855	58,475	8,775	10,265	126,600	68,105		**472,915**	
Attraction Ratio	1.17	1.01	1.00	1.03	0.55	2.19	1.17	0.76	1.99	0.75	0.97	0.46	0.86	1.10	1.06			

ATTRACTION GROWTH FACTORS

Determine new attractions (A) by multiplying by the ratio of the Target/ Total

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15
7,033	19,851	14,126	63,041	3,221	50,003	11,014	23,566	15,674	17,090	56,431	4,065	8,816	139,615	72,070

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS															Productions Total	Productions Target	Production Ratio
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York			
Albert	**5,262**	0	0	0	0	5	0	0	0	0	1	0	0	14,137	0	19,405	19,405	1.000
Carleton	0	**19,466**	0	0	0	0	0	0	0	0	0	0	5	0	24	19,495	19,495	1.000
Charlotte	0	1	**14,209**	0	0	0	0	0	0	0	2,826	1	0	0	1,328	18,365	18,365	1.000
Gloucester	0	0	0	**59,034**	0	0	0	1	0	0	0	0	0	0	0	59,035	59,035	1.000
Kent	0	0	0	0	**6,948**	3	0	28	19	0	0	0	0	14,993	19	22,010	22,010	1.000
Kings	0	0	0	0	0	**12,978**	0	0	334	0	14,646	1	0	1,804	17	29,780	29,780	1.000
Madawaska	0	0	0	0	0	0	**8,086**	0	0	0	0	0	4	0	0	8,090	8,090	1.000
Northumberland	0	0	0	1	8	0	0	**40,758**	0	0	0	0	0	1	15	40,785	40,785	1.000
Queens	0	0	0	0	1	3,416	0	0	**4,799**	0	860	168	0	39	2,986	12,270	12,270	1.000
Restigouche	0	0	0	0	0	0	0	0	0	**30,565**	0	0	0	0	0	30,565	30,565	1.000
Saint John	0	0	0	0	0	588	0	0	3	0	**36,256**	1	0	0	6	36,855	36,855	1.000
Sunbury	0	0	0	0	0	4	0	0	114	0	12	**593**	0	0	10,512	11,235	11,235	1.000
Victoria	0	15	0	0	0	0	8	0	0	0	0	0	**11,937**	0	0	11,960	11,960	1.000
Westmorland	80	0	0	0	13	209	0	0	0	0	0	0	0	**86,757**	0	87,060	87,060	1.000
York	0	0	1	0	0	2	0	0	31	0	15	8,767	0	0	**57,189**	66,005	66,005	1.000
Attractions Total	5,343	19,483	14,210	59,036	6,971	17,205	8,095	40,787	5,300	30,565	54,617	9,531	11,945	117,731	72,096			
Attractions Target	5,990	19,665	14,160	61,005	5,845	22,855	9,440	31,005	7,875	22,855	58,475	8,775	10,265	126,600	68,105		**472,915**	
Attraction Ratio	1.12	1.01	1.00	1.03	0.84	1.33	1.17	0.76	1.49	0.75	1.07	0.92	0.86	1.08	0.94			

PRODUCTION GROWTH FACTORS

Determine new attractions (A) by multiplying by the ratio of the Target/ Total

P1	19405
P2	19495
P3	18365
P4	59035
P5	22010
P6	29780
P7	8090
P8	40785
P9	12270
P10	30565
P11	36855
P12	11235
P13	11960
P14	87060
P15	66005

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS															Productions Total	Productions Target	Production Ratio
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York			
Albert	**5,262**	0	0	0	0	5	0	0	0	0	1	0	0	14,137	0	19,405	19,405	1.00000
Carleton	0	**19,466**	0	0	0	0	0	0	0	0	0	0	5	0	24	19,495	19,495	1.00000
Charlotte	0	1	**14,209**	0	0	0	0	0	0	0	2,826	1	0	0	1,328	18,365	18,365	1.00000
Gloucester	0	0	0	**59,034**	0	0	0	1	0	0	0	0	0	0	0	59,035	59,035	1.00000
Kent	0	0	0	0	**6,948**	3	0	28	19	0	0	0	0	14,993	19	22,010	22,010	1.00000
Kings	0	0	0	0	0	**12,978**	0	0	334	0	14,646	1	0	1,804	17	29,780	29,780	1.00000
Madawaska	0	0	0	0	0	0	**8,086**	0	0	0	0	0	4	0	0	8,090	8,090	1.00000
Northumberland	0	0	0	1	8	0	0	**40,758**	0	0	0	0	0	1	15	40,783	40,785	1.00005
Queens	0	0	0	0	1	3,416	0	0	**4,799**	0	860	168	0	39	2,986	12,269	12,270	1.00008
Restigouche	0	0	0	0	0	0	0	0	0	**30,565**	0	0	0	0	0	30,565	30,565	1.00000
Saint John	0	0	0	0	0	588	0	0	3	0	**36,256**	1	0	0	6	36,854	36,855	1.00003
Sunbury	0	0	0	0	0	4	0	0	114	0	12	**593**	0	0	10,512	11,235	11,235	1.00000
Victoria	0	15	0	0	0	0	8	0	0	0	0	0	**11,937**	0	0	11,960	11,960	1.00000
Westmorland	80	0	0	0	13	209	0	0	0	0	0	0	0	**86,757**	0	87,059	87,060	1.00001
York	0	0	1	0	0	2	0	0	31	0	15	8,767	0	0	**57,189**	66,005	66,005	1.00000
Attractions Total	5,342.0	19,482	14,210	59,035	6,970	17,205	8,094	40,787	5,300	30,565	54,616	9,531	11,946	117,731	72,096			
Attractions Target	5,990	19,665	14,160	61,005	5,845	22,855	9,440	31,005	7,875	22,855	58,475	8,775	10,265	126,600	68,105		**472,915**	
Attraction Ratio	1.12	1.01	1.00	1.03	0.84	1.33	1.17	0.76	1.49	0.75	1.07	0.92	0.86	1.08	0.94			

COMPARING BETWEEN PREDICTED AND ACTUAL

DIFFERENCE (PREDITCTED - ACTUAL)

ORIGINS/ PRODUCTIONS	DESTINATIONS/ ATTRACTIONS															Productions Total	Productions Target
	Albert	Carleton	Charlotte	Gloucester	Kent	Kings	Madawaska	Northumberland	Queens	Restigouche	Saint John	Sunbury	Victoria	Westmorland	York		
Albert	2,107	0	-5	-10	-45	-40	-15	-10	0	0	-534	-5	-5	-1,403	-35		
Carleton	-10	5,571	-1,480	0	-5	-55	-65	-60	-5	-15	-965	-25	-870	-110	-1,906		
Charlotte	-5	-209	4,979	0	0	-2,075	-20	0	-10	-5	-3,484	-29	0	-100	958		
Gloucester	0	-15	0	5,419	-5	-30	-10	-714	0	-335	-500	0	-5	-3,110	-695		
Kent	-255	-35	0	-30	1,753	-7	-20	-447	9	-5	-415	-5	-5	-522	-16		
Kings	-35	-5	-1,115	-15	-10	-617	-20	-5	69	-15	1,606	-29	-5	599	-403		
Madawaska	0	-10	-5	-20	-5	-15	1,236	-5	-5	-60	-15	-15	-611	-445	-25		
Northumberland	-20	-975	-5	-2,204	-22	-155	-30	12,443	0	-15	-510	-10	-5	-5,874	-2,620		
Queens	-5	-460	-35	0	-4	2,911	-270	-75	-901	-30	-50	-302	0	-211	-569		
Restigouche	0	-85	-5	-5,005	-5	-10	-350	-85	-5	8,435	-425	-5	0	-1,780	-675		
Saint John	-5	0	-805	-5	-5	-5,387	-30	0	-22	-30	6,991	-14	-5	-395	-289		
Sunbury	-5	-135	-60	-5	0	-31	-550	-50	-1,396	-25	-403	-2,602	-20	-110	5,392		
Victoria	0	-1,350	-5	-35	-5	-35	-427	-45	-5	-30	-295	-15	3,282	-225	-810		
Westmorland	-2,410	-10	0	-55	-512	114	-210	-45	-35	-10	-1,955	-30	-25	5,277	-95		
York	-5	-2,465	-1,409	-5	-5	-218	-565	-1,120	-269	-150	-2,905	3,842	-45	-460	5,779		
Attractions Total																	
Attractions Target																	

CURRICULUM VITAE

Brianna Jean Mae Morehouse

UNIVERSITIES ATTENDED

Graduate 2018-Present	**Master of Science in Civil Engineering** University of New Brunswick
Undergraduate 2012-2016	**Bachelor of Science in Civil Engineering** University of New Brunswick

PUBLICATIONS: None

CONFERENCE PRESENTATIONS: None