

**Understanding and Modelling the Impacts of a New Competitor on the Sales of Established  
Businesses Within Similar Markets**

By

Scott Norman

A Major Research Paper  
presented to Ryerson University

in partial fulfilment of the

requirement for the degree of  
Master of Spatial Analysis (MSA)  
in the program of  
Spatial Analysis

Toronto, Ontario, Canada, 2017

© (Scott Norman) 2017

## **Author's Declaration**

I hereby declare that I am the sole author of this MRP. This is a true copy of the MRP, including any required final revision.

I authorize Ryerson University to lend this MRP to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this MRP by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my MRP may be made electronically available to the public.

## **Abstract**

The research conducted for this study presents a Geographic Information Systems (GIS) and spatial analysis to understanding and modelling the impacts of a new competitor on the sales of established businesses within similar markets. Data for this analysis was provided by Canadian Tire Corporation (CTC). This study will outline and create trade areas for all stores in the Part Source network to determine which stores will be the focus of this study. Once focus stores have been determined, sales data was analyzed in two ways. The first analysis was conducted to determine the distance decay effects of the new competitor on sales. The second analysis utilized geographic weighted regression (GWR) models on expenditure and vehicle data.

## **Acknowledgement**

This research was made possible by Canadian Tire Corporation (CTC). I would also like to acknowledge my supervisor, Dr. Eric Vaz, for his support throughout the research and writing of this paper. In Addition, I would also like to acknowledge my manager at CTC, Christopher Curry, for ensuring that data was made available in a timely matter and for being flexible with my work hours so I could attend classes.

## Table of Contents

Author's Declaration .....	i
Abstract .....	ii
Acknowledgement .....	iii
Table of Contents .....	iv
List of Tables .....	vi
List of Figures .....	vii
List of SQL Queries .....	viii
List of Python Scripts.....	ix
Table of Contents .....	v
1. Introduction .....	1
2. Literature Review .....	4
2.1 Effects of International Expansion of Companies into Canada .....	5
2.2 Assessing the Factors Impacting Market Share .....	8
2.3 Defining Trade Delineation .....	10
2.4 Geographic Weighted Regression .....	11
3. Study Areas .....	12
3.1 Markham, Ontario .....	13
3.2 Medicine Hat, Alberta .....	15
4. Methodology .....	18
4.1 Data .....	18
4.1.1 Sales Data Validation and Transformation .....	19
4.1.2 Creating Dominant Sales Layer .....	21

4.2 Creating Trade Areas .....	21
4.3 Determining Competitors within Trade Areas .....	22
4.4 Geographic Weighted Regression .....	22
5. Results and Discussion .....	24
5.1 Initial Assessment and Statistics of Competitor Impacts .....	26
5.1.1 Medicine Hat, Alberta .....	27
5.1.2 Markham, Ontario .....	29
5.2 Determining Factors of the Effects of Competitors .....	32
5.2.1 Medicine Hat, Alberta .....	33
5.2.2 Markham, Ontario .....	36
6. Conclusions .....	39
6.1 Findings .....	39
6.2 Limitations .....	41
6.3 Future Research .....	42
References .....	45
Appendix A .....	49
Appendix B .....	62

## List of Tables

<b>Table 1.</b> Summary of Target locations acquired by each major retailer .....	<b>6</b>
<b>Table 2.</b> Summary of stores whose trade areas border the focus store .....	<b>16</b>
<b>Table 3.</b> Summary of automotive retailers in Canada by province .....	<b>19</b>
<b>Table 4.</b> Summary of Census Subdivisions that border Markham.....	<b>24</b>
<b>Table 5.</b> Summary of Census Subdivisions that border Medicine Hat.....	<b>25</b>
<b>Table 6.</b> Summary of trade areas that were utilized .....	<b>25</b>
<b>Table 7.</b> Summary of sales included in the analysis .....	<b>26</b>
<b>Table 8.</b> Summary of sales by distance to competitor in Medicine Hat.....	<b>28</b>
<b>Table 9.</b> Summary of sales by distance to focus store in Medicine Hat .....	<b>28</b>
<b>Table 10.</b> Summary of sales by distance to competitor in Markham.....	<b>30</b>
<b>Table 11.</b> Summary of sales by distance to focus store in Markham .....	<b>31</b>
<b>Table 12.</b> Summary statistics of GWR of expenditure variables in Medicine Hat .....	<b>32</b>
<b>Table 13.</b> Summary statistics of GWR of top five TVIO variables in Medicine Hat.....	<b>32</b>
<b>Table 14.</b> Summary statistics of GWR of top three categories of imported cars in Medicine Hat.....	<b>33</b>
<b>Table 15.</b> Summary statistics of GWR of top three categories of domestic cars in Medicine Hat.....	<b>33</b>
<b>Table 16.</b> Summary statistics of GWR of top three categories of imported trucks in Medicine Hat .....	<b>34</b>
<b>Table 17.</b> Summary statistics of GWR of top three categories of domestic trucks in Medicine Hat .....	<b>34</b>
<b>Table 18.</b> Summary statistics of GWR of expenditure variables in Markham .....	<b>35</b>
<b>Table 19.</b> Summary statistics of GWR of top five TVIO variables in Markham .....	<b>35</b>
<b>Table 20.</b> Summary statistics of GWR of top three categories of imported cars in Markham .....	<b>36</b>
<b>Table 21.</b> Summary statistics of GWR of top three categories of domestic cars in Markham .....	<b>36</b>
<b>Table 22.</b> Summary statistics of GWR of top three categories of imported trucks in Markham .....	<b>37</b>
<b>Table 23.</b> Summary statistics of GWR of top three categories of domestic trucks in Markham .....	<b>37</b>

## List of Figures

<b>Figure 1.</b> Map of focus store's and intersecting Part Source trade areas in Markham .....	<b>14</b>
<b>Figure 2.</b> Map of focus store's trade area and the new competitor in Markham .....	<b>15</b>
<b>Figure 3.</b> Map of focus store's and intersecting Part Source trade areas in Medicine Hat .....	<b>16</b>
<b>Figure 4.</b> Map of focus store's trade area and the new competitor in Medicine Hat .....	<b>17</b>
<b>Figure 5.</b> Outline transformation and analysis methodology .....	<b>18</b>
<b>Figure 6.</b> Map of the change in sales between 2014 and 2016 in Medicine Hat .....	<b>27</b>
<b>Figure 7.</b> Map of the change in sales between 2014 and 2016 in Markham .....	<b>29</b>



## List of SQL Queries

<b>Figure A1.</b> Validate postal codes .....	<b>49</b>
<b>Figure A2.</b> Create unique FSA table .....	<b>49</b>
<b>Figure A3.</b> Create unique postal code table .....	<b>49</b>
<b>Figure A4.</b> Create table of positive sales by postal code .....	<b>50</b>

## List of Python Scripts

<b>Figure A6.</b> Postal code sales data transformation .....	<b>51</b>
<b>Figure A7.</b> Dominant postal code sales data compiler .....	<b>55</b>
<b>Figure A8.</b> Spatial join dominant sales to FSA shapefile.....	<b>57</b>
<b>Figure A9.</b> Trade area creation .....	<b>58</b>

## 1. Introduction

The international expansion of companies into various markets has led to many researchers to try and understand two main by-products. The first by-product is to understand why a company succeeds and fails in a market. The second by-product is what happens to markets if a company succeeds in entering into it. Shostack (1977), and later Booms and Bitner (1981), identified the 4Ps of marketing; Product, Place, Promotion and Price. Researchers later identified that products compared to services had significant difference which led to the addition of three Ps (Participants, Physical Evidence, and Process) into the model (Anitsal et. al, 2013). According to Foote and Margaret (1995) outlined the issues of analyzing the spatial components due to a lack of technology to effectively collecting and analyze spatial data. While the late 1950s to early 1970s saw the development of sophisticated mathematical algorithms, it wasn't until the early 1990s that the software was being refined and relied upon by companies and organizations (Foote & Margaret 1995). Jeffrey Star and John Estes, defined a Geographic Information System (GIS) as:

“an information system that is designed to work with data referenced by spatial or geographic coordinates. In other words, a GIS is both a database system with specific capabilities for spatially-reference data, as well [as] a set of operations for working with data”

With technology ever increasing its able to handle larger and larger quantities of data, companies are becoming more and more reliant on spatially referenced data in order to make their decisions (Mishra, 2009). This includes, but is not limited to, site selection, customer and

product analysis, determining market share of current locations, and assessing impacts of competitors.

In order to assessing impacts of competitors, researchers have sought out various methods to model these effects. One of the most widely used spatial methods is a gravity model. A gravity model is considered to be one of the basic spatial analysis models. Its initial intent to be able to describe the migration of population between two regions but was later adopted for use in understanding and measuring attraction between any two geographic locations (Chen, 2013). Another spatial method that is commonly used is a huff model, which is a more advanced gravity model. ESRI defines a huff model as “the principle that the probability of a given consumer visiting and purchasing at a given site is a function of the distance to that site, its attractiveness, and the distance and attractiveness of competing sites” (How Original Huff Model Works, 2016). The final method that is regularly used is a regression model. A regression model is defined by HBR (Harvard Business Review) as:

“A way of mathematically sorting out which of those variables does indeed have an impact. It answers the questions: Which factors matter most? Which can we ignore? How do those factors interact with each other? And, perhaps most importantly, how certain are we about all of these factors?”

For the purpose of this study, a sample data of Part Source’s postal code sales data will be utilized for the study. With international companies, such as NAPA and Carquest, further expanding into the Canadian market along with competition from other Canadian retailers, such as Princess Auto, Part Source is an excellent candidate to assess the effects of new competitors. Since Part Source was established, it has opened an average of just over 5 locations per year (About Us, 2016). Part Source has had to conservatively expand its network

in order to ensure that the markets that they enter into have enough market potential so the new Part Source store can obtain a profitable market share. While Part Source was only established 18 years ago, it is one of five banners, which are Canadian Tire Retail, Canadian Tire Petroleum, Marks, Part Source, and FGL Sports (formerly “The Forzani Group Ltd.”) under the CTC (Canadian Tire Corporation) which was established in 1922 (Canadian Tire Corporation, 2016). This means that the site selection process has been rigorously tested and refined and can be utilized in this study with limited concerns about other factors, such as poor site selection or limitations of funds to open multiple stores, affecting the results.

Before any focus stores can be determined, it is important to understand how markets are being defined within this study. In many cases, researchers utilize census boundaries such as Census Subdivisions or Census Metropolitan Areas to define markets. While this is beneficial by allowing for easy understanding of boundaries by a larger audience, it limits the potential scope of the analysis. An example of this is utilizing Census Subdivisions to determine competitors within Ontario. If a Part Source was located in the west end of Etobicoke and a competitor opened in the east end of Mississauga, the two would not be considered to be within the same market while the sales would indicate that the competitor has had an impact. While substituting the Census Subdivision boundaries with larger boundaries, such as Census Divisions or Census Metropolitan Areas, would reduce these limitations, there would still be an impact on the amount of potential focus stores. Due to this, developing trade areas, sometimes referred to as catchment areas, was determined to be the best method to mitigate any of these impacts.

Trade areas can be defined using a wide variety of geographies, methods and variables. Some geography boundaries that are used include, but are not limited to, census boundaries, road networks, or custom grids such as squares or hexagons. Many common methods include building buffers using a fixed or varying distance(s), creating Thiessen polygons, or amalgamation of smaller geographic boundaries based on a set of criteria. The variables that are most frequently used are population or household counts, counts of vehicles, or sales data. For this study, trade areas will be defined using an amalgamation of Forward Sortation Areas based on sales associated with each store. This study aims to:

- Determine the spatial patterns of the impacts a new competitor on sales
- Determine factors behind the spatial pattern using Geographic Weighted Regression (GWR)

## **2. Literature Review**

The study of new competitor's effects on an existing stores performance in the Canadian retail industry has received interest in theoretical and scholar research. Furthermore, there has been a larger focus on four major topics including: the effects of the international expansion of companies into Canada, factors impacting market share, trade area delineation, and Geographically Weighted Regression. This literature review will explore existing studies to assist and evaluate current theories and practices to assess the impacts of new competitors on existing store performances.

## **2.1 Effects of International Expansion of Companies into Canada**

As major markets continue to open up to international expansion, businesses and researchers alike try to understand two main by-products. The first by-product is the successes and failures of companies due to the competitiveness of the markets. The second by-product is the impacts on markets when a company successfully expands into international markets. With regards to the successes and failures of companies' expansion, Walmart has been the focal point of many studies due to its success. According to Mun's research (2012), one of Walmart's main objectives was to have a significant presence when they expanded into the Canadian market. To do that, Walmart purchased 120 of 142 Woolco Discount stores in 1994 (Mun & Yazdanifard, 2012). Walmart also acquired 39 of the Zellers locations from Target when Target purchased the majority of Zeller locations in Canada (Canada Newswire, 2011). While Walmart succeeded in integrating itself into the Canadian market through purchasing stores from another retailer, Target did not have the same success using a similar method. Target attempted to enter the Canadian market by purchasing 146 of 277 Zellers locations (CTC Competitor Database, 2017). It was found that while Target had succeeded in securing a large network of stores, it failed to obtain a larger market share than Zellers had previously in the same markets. When Target had decided to exit the Canadian market, Canadian Tire, Lowes, and Walmart each selected Target locations to obtain (See Table 1). The remaining locations were either returned to the landlord or sold on a case-by-case situation.

	Canadian Tire	Lowes	Walmart	Grand Total
Alberta		2		2
British Columbia	4	4	4	12
Manitoba	1		1	2
New Brunswick	1			1
Nova Scotia	1			1
Ontario	2	7	5	14
Quebec	3		4	7
Saskatchewan		1		1
Grand Total	12	14	14	40

**Table 1.** Summary of Target locations acquired by each major retailer

With Target's failure to embed itself within the Canadian market, it is important to understand other approaches taken by companies. According to Dawson et al. (2006), department stores such as Walmart or Target, have a large enough variety of products, it is easier for them to enter a larger quantity of markets. This is due to department stores being able to revise their inventory based on the needs of the customers (Dawson et. al., 2006). While department stores have the ability to stock various products based on customer needs, specialty retailers are not able to adapt in the same way. This, in turn, means that their expansion into new markets is predominantly slower than that of department stores (Dawson et. al., 2006). Home Depot and Lowes, who are specializing home improvement/Do It Yourself (DIY) projects, are two examples in which companies slowly enter the market. Home Depot entered the Canadian market in 1994 by acquiring Aikenhead's Home Improvement Warehouse, which consisted of five locations (History, 2017). Currently, Home Depot operates 182 stores within ten provinces in Canada (History, 2017). This means that Home Depot expands its store network by approximately 8 locations per year. Like Home Depot, Lowes entered the Canadian market by opening stores in Brantford, Brampton, and Hamilton in



December of 2007 (CTC Competitor Database, 2017). Lowes' plan was to select a few store locations that appeared to have high market potential rather than purchasing a large store network from another company. This was done to avoid the risk and cost associated with purchasing a store network that may have poor site selection, such as Target purchasing Zellers locations.

Up until 2016, Lowes had slowly increased its network to 51 locations. This was a slow expansion to ensure that the integration into the Canadian market was successful. In May of 2016, Lowes announced that it would be acquiring RONA Inc., which currently has a store network of 539 locations (Lowes, 2017). This acquisition is continuously being monitored and assessed due to Lowes acquiring a large specialized retailer in a foreign market.

While Lowes has only recently acquired Rona Inc., Walmart has effectively embedded itself within various international markets, such as Canada and Mexico, making it the focal point of many studies regarding the international expansion's effects on a market. According to Clarks (2000) and Kumar and Karanda (2000), Walmart's Supercentre format, which includes a grocery department, has an advantage over the majority of other grocery stores due to its convenience of a one-stop shop, low-cost, and is open 24-hours. Because of this success, Walmart has switched its focus to building Supercentres and updating older Walmarts to the Supercentre format. Since 2013, Walmart has built 35 stores, 29 of which were Supercentre format (CTC Competitor Database, 2017).

This has led to many businesses struggling to maintain a profitable market share due to a combination of trying to compete with low prices, only selling a limited range of products and

hours of operation. For businesses to compete with having to lower prices, businesses must make the decision to have less of a profit margin on each item. The second factor of only selling a limited range of products has led to stores expanding their products or locating within shopping centers, to appeal to customers who require items from various retail departments, such as plumbing, automotive, or sporting goods (Clark, 2000). The final factor of hours of operation has led to businesses requiring hiring more staff to compete with big box operation hours. All of the factors have led to many businesses to close. According to Clark (2000), these factors have more of an effect on retailers with a smaller store network than retailers who are more established. This is due to limitations such as smaller retailers not being able to purchase products in large enough quantities to obtain the same discounts per unit or not having the finances to hire more employees to stay open later.

## **2.2 Assessing the Factors Impacting Market Share**

To accurately interpret and predict the effects of competitive retail facilities, the attractiveness of a retail facility must be measured and implemented into a given model (Drezner & Drezner, 2002). These models can include trade area estimation, market share, site selection, or assessing impacts on new competitors. According to Drezner et al. (2002), an interrelationship exists between in an area of four factors, one dependent variable, and three independent variables. The dependent variable is market share, which is a unit of measurement derived sales figures of competing for retail facilities within an area. The first independent variable is the buying power of an area. The second independent variable is the cost (usually

measured in distance or time) between customers and retail facilities. The final independent variable is the attractiveness of retail facilities.

The buying power of an area, more commonly referred to as market potential, is the "[Amount of money] dedicated to purchases in a specific type of retail facilities" (Drezner & Drezner, 2002). Due to most retail companies not releasing sales figures, competing companies and researchers alike attempt to use a combination of datasets and variables to infer the market potential of an area. Some of the common datasets used to determine the market potential are Census data, DemoStats, and CensusPlus, which include variables about income, expenditure, and household characteristics. To increase the granularity of assessing market potential, and subsequent market share(s), areas are divided into smaller zones. These zones include but are not limited to cities, CMA (Census Metropolitan Areas), FSALDU (more commonly known as Postal Codes), or DA (Dissemination Areas) which are determined by Statistics Canada. All subsequent data, such as census data and third party datasets, are aligned with these areas.

To accurately assess costs such as distance or time in a retail environment, it is important to utilize a road network as well a public transportation systems, such as subway tracks, rather than straight line distance (Dolega et al., 2016). This is due retail customers requiring a mode of transportation to shop; this does not include online shopping if the company delivers to customer's homes.

The attractiveness of a location includes but is not limited to size, products offered, and if it has other retailers within easy access (such as a mall or plaza). According to Hollander

(2002), the option for one-stop shopping, such as a mall or plaza can significantly impact a customer's decision in where to shop for a given item. Another method for retailers to make their location more appealing, companies will make an agreement with other retailers, who are not direct competitors to open up location. Some examples of this are home Depot locations having a Harvey's location within the store or Lowe's locations having a Country Style location within or on its property.

## **2.3 Defining Trade Delineation**

When determining the boundaries of a trade area, there are a few methods which can be implemented. According to Lea (1998), a retail trade area is defined as "[An] area, typically around the store, from which the store derives most of its patronage." Early work by Huff (1963) outlined defining retail trade areas by "(1) dividing the surrounding area into small statistical units; (2) calculating the probability of consumers from each of these units going to a particular shop-ping center; and (3) drawing lines around all statistical units having like probabilities."

Popular techniques for creating trade area include Simple Rings, Hand Drawn Areas, Drive-time Areas, Hand-drawn areas, Non-overlapping Rings, Threshold Areas, Customer Derived Areas, and utilization of a Huff Model. The most simplistic and widely used method is the Simple Rings Method, which a simple radius is created around a store based on a Euclidean (Straight Line) Distance. Drive Areas are similar to a Simple Rings Method however it utilizes a road network rather than a straight line distance. Distance is also regularly substituted for time as many researchers have determined that the time it takes to get to a location can be more of

a determining factor than the distance to the location. Threshold Areas are areas that are created by continuously expanding the radius or drive areas until a specified limit is met. Customer Derived Areas are comprised of a primary, secondary, and tertiary markets which capture 30 percent, 50 percent, and 70 percent respectively of a variable. Flater (2012) defined the Huff Model as “A spatial interaction model that calculates gravity-based probabilities of consumers at each origin location patronizing each store in the store dataset.” The Huff Model, unlike the other techniques, utilizes multiple variables, such as sales, store size, and distance to determine the trade area.

## **2.4 Geographic Weighted Regression**

To quantify various relationships that exist between different variables, researchers have developed various mathematical models. One of the most commonly used models is a regression model when researchers are attempting to estimate the effects of independent variables on the dependent variable (Brundson, Charlton, and Fotheringham, 2006). OLS (Ordinary Least Square) is a linear modeling technique that is used to estimate regression model at a global level, and due to its simplicity, OLS is one of the most utilized techniques (Dismuke and Lindrooth, 2006). In many cases, however, OLS is misused for local analysis and can, therefore, be misleading based on the findings (Chumney, 2006). One technique that has become more regularly used for local modeling is GWR (Geographically Weighted Regression) (Brunsdon, Fotheringham, and Charlton, 2003). With many researchers in various fields recognizing the importance of local modeling and the inclusion of weighting utilizing spatial component(s), GWR has been regarded as being more reliable and accurate when determining

the spatial relationships of variables (Zhang et al., 2004). Research surrounding Health and Disease (Hu et. al., 2012), Criminology (Brunsdon, et. al., 2003), Environmental (Harris et. al, 2010) and Economic Growth (Öcal and Yildirim, 2010) have started regularly relying on GWR to interpret the spatial relationship of various factors.

### **3. Study Area**

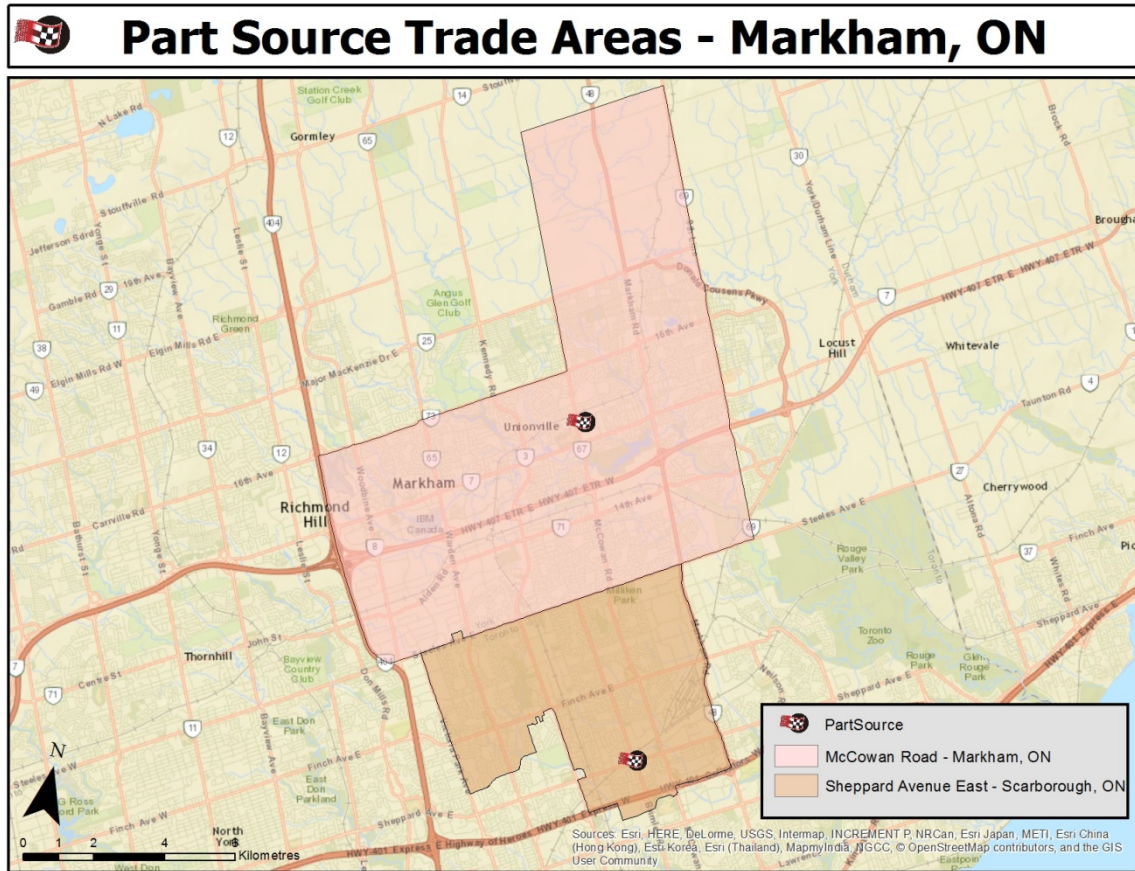
The study areas were chosen based on three requirements. The first requirement was that the store did not have any renovations done to the store after 2012. This was done to ensure that expansion, relocation, or renovation did not influence the change in sales of a store. The second requirement is that the store must have one competitor that has opened within the trade area between 2013 and 2015. This requirement is to ensure that the sales data that is being analyzed has had time to adjust to the new competitor within the market. To accurately assess, sales from one fiscal year prior and one year after the competitor opened. The final requirement is that adjacent stores have not been renovated, relocated, expanded, or built between 2013 and 2015. This is to ensure that no sales data was being influenced by a store within its network.

Based on these requirements, two stores, in Markham, Ontario and Medicine Hat, Alberta were chosen. According to Vaz and Arsanjani (2015), the Toronto Census Metropolitan Area is continuously expanding its urban land use and, subsequently, there is rapid economic growth. Because of this, the focus stores in Markham, Ontario and Medicine Hat, Alberta must be analyzed separately as the markets themselves are too different to be considered comparable.

### **3.1 Markham Study Area**

The first store that was analysed is located in Markham, Ontario which has a population of over 353,000 and covers 212 square kilometres (City of Markham, 2014). According to Statistics Canada, Markham had a population growth of 15.3% between 2006 and 2011. This growth rate is significantly higher than the national average of 5.9% within the same years (Statistics Canada, 2011).

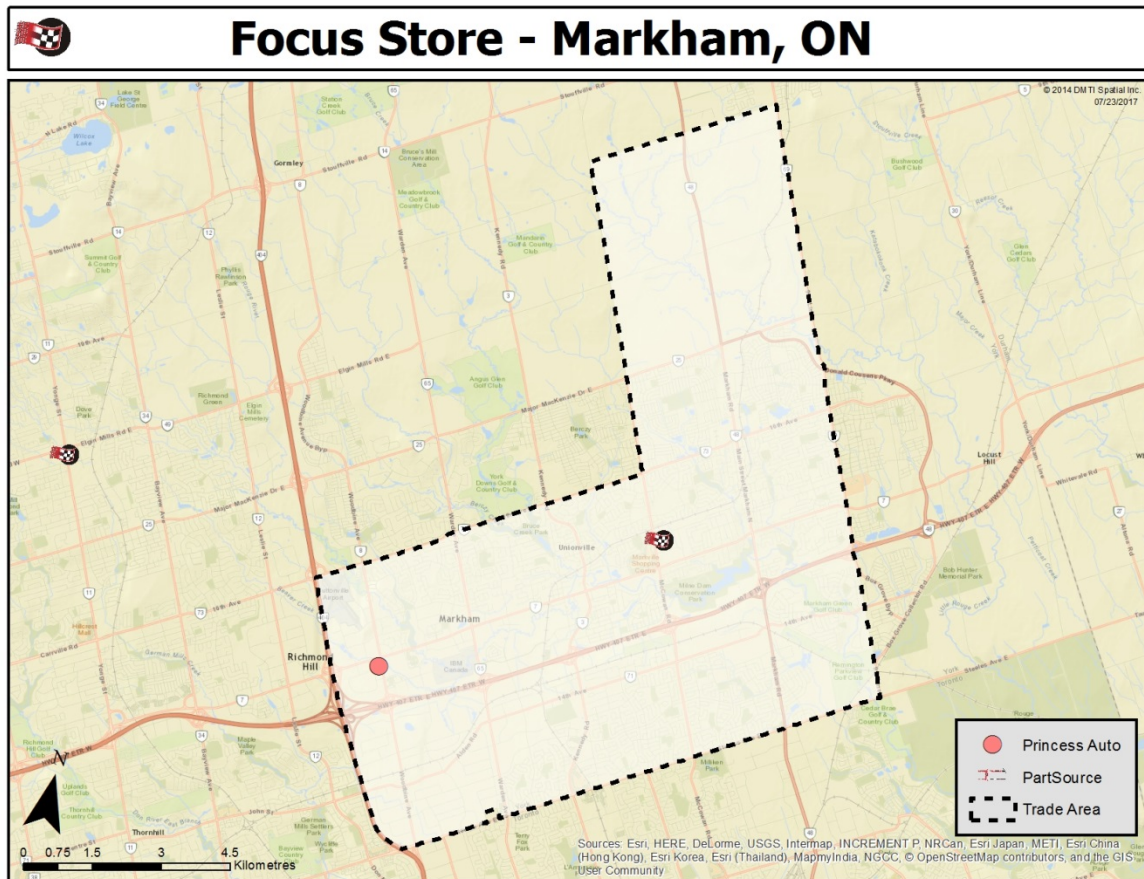
The Part Source store that was selected as the focus store for this area is located at 8675 McCowan Road in Markham. The store opened in December of 2008 and is approximately 9,000 square feet (CTC, 2017). Based on Census Subdivision boundaries, this store is considered to be the only Part Source store in the market. Based on trade areas however, the Markham store's trade area is bordering the trade area of a Part Source store in Scarborough which is located at 2650 Lawrence Ave. East. The location in Scarborough has not been altered since it was built in 2004. Figure 1 shows both the Part Source stores in Markham and Scarborough.



**Figure 1.** Map of focus store's and intersecting Part Source trade areas in Markham

The competitor that was analyzed was the Princess Auto which is located at 8500 Woodbine Ave in Markham. The store is approximately 38,000 square feet and opened in February of 2015 (CTC, 2017). This Princess Auto location is approximately 7 km away from the focus store (See Figure 2).





**Figure 2.** Map of focus store’s trade area and the new competitor in Markham

### 3.2 Medicine Hat Study Area

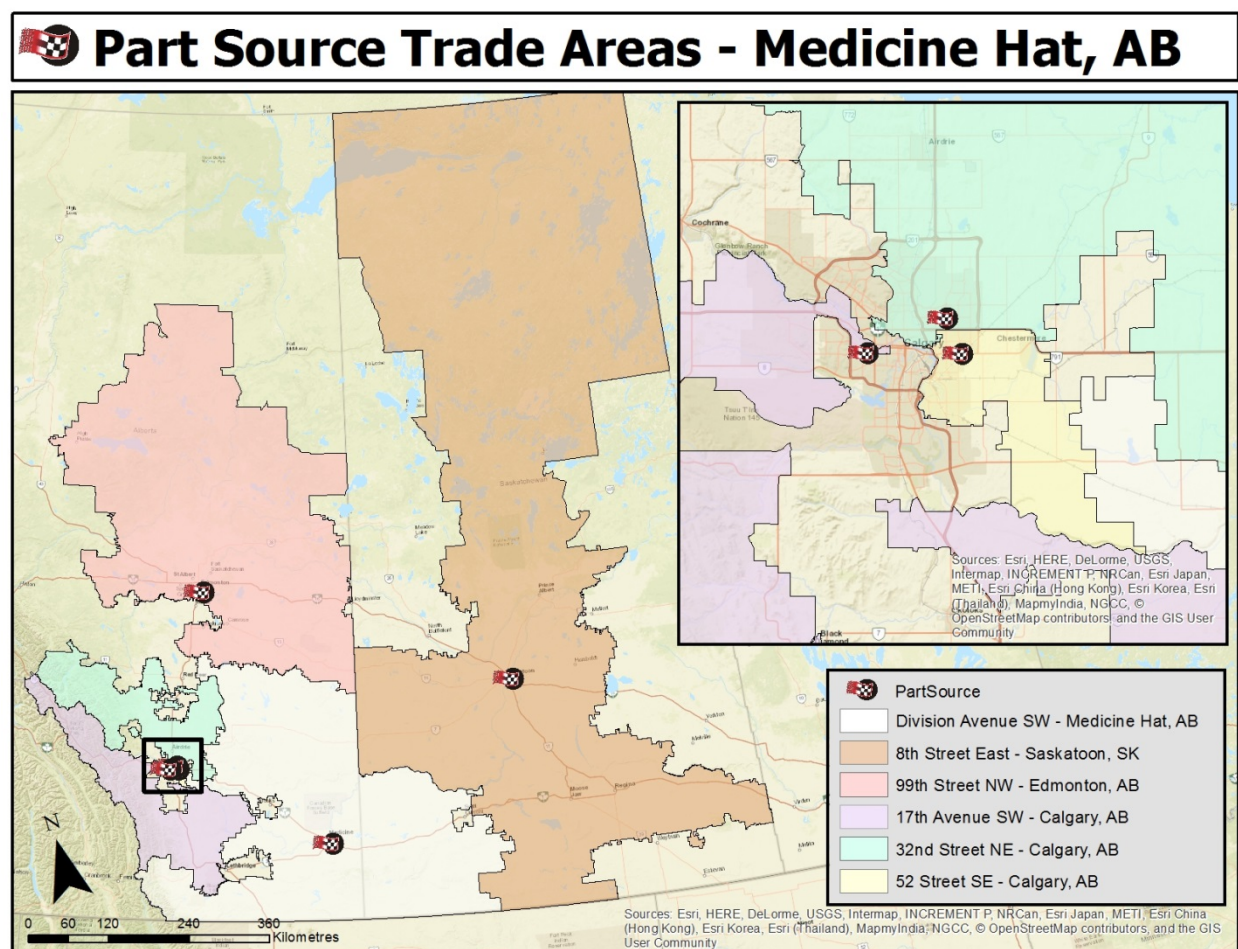
The second store that was analyzed is located in Medicine Hat, Alberta which covers approximately 112 square kilometres within its municipal boundary (City of Medicine Hat, 2012). According to Statistics Canada, Medicine Hat currently has a population of approximately 65,000 and a population growth of 5.3% (2017).

The store that will be the Part Source store to be focused on is located at 310 Division Avenue SW. The store is approximately 12,000 square feet and opened in September of 2009 (CTC, 2017). Based on Census Divisions, this store is the only Part Source location within

Medicine Hat. Based on trade areas however, it borders the trade areas of five store in Edmonton, Calgary and Saskatoon (See Table 2 and Figure 3).

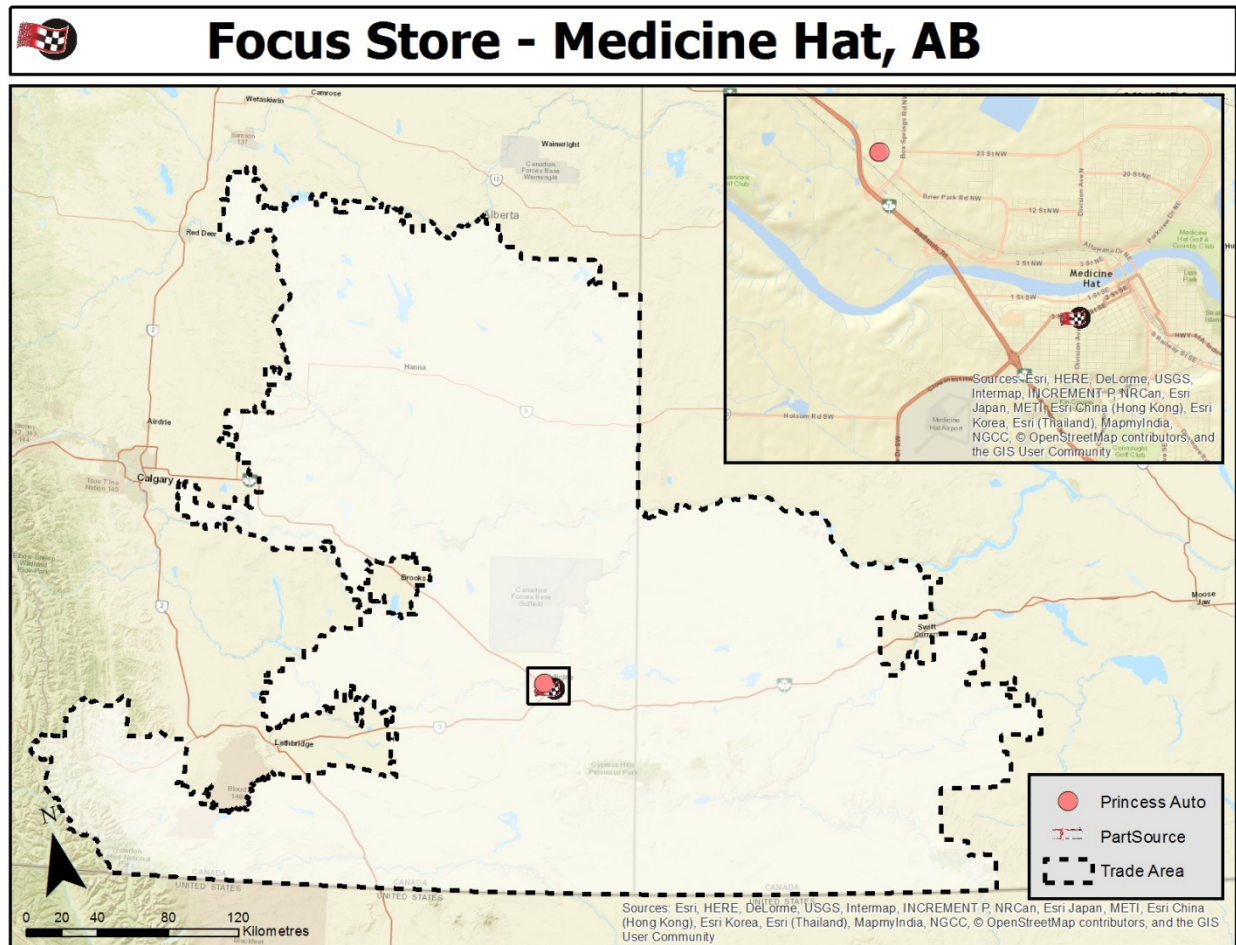
Address	City	Province	Last Alteration Type	Last Alteration Date	Approximate size
8TH STREET EAST	SASKATOON	SASKATCHEWAN	New	Sept. 1999	7,500
32ND STREET NE	CALGARY	ALBERTA	New	June 1999	7,500
52ND STREET SE	CALGARY	ALBERTA	New	June 2000	7,500
17TH AVENUE SW	CALGARY	ALBERTA	New	Sept 2000	7,500
99TH STREET NW	EDMONTON	ALBERTA	New	March 2001	7,500

**Table 2.** Summary of stores whose trade areas border the focus store



**Figure 3.** Map of focus store's and intersecting Part Source trade areas in Medicine Hat

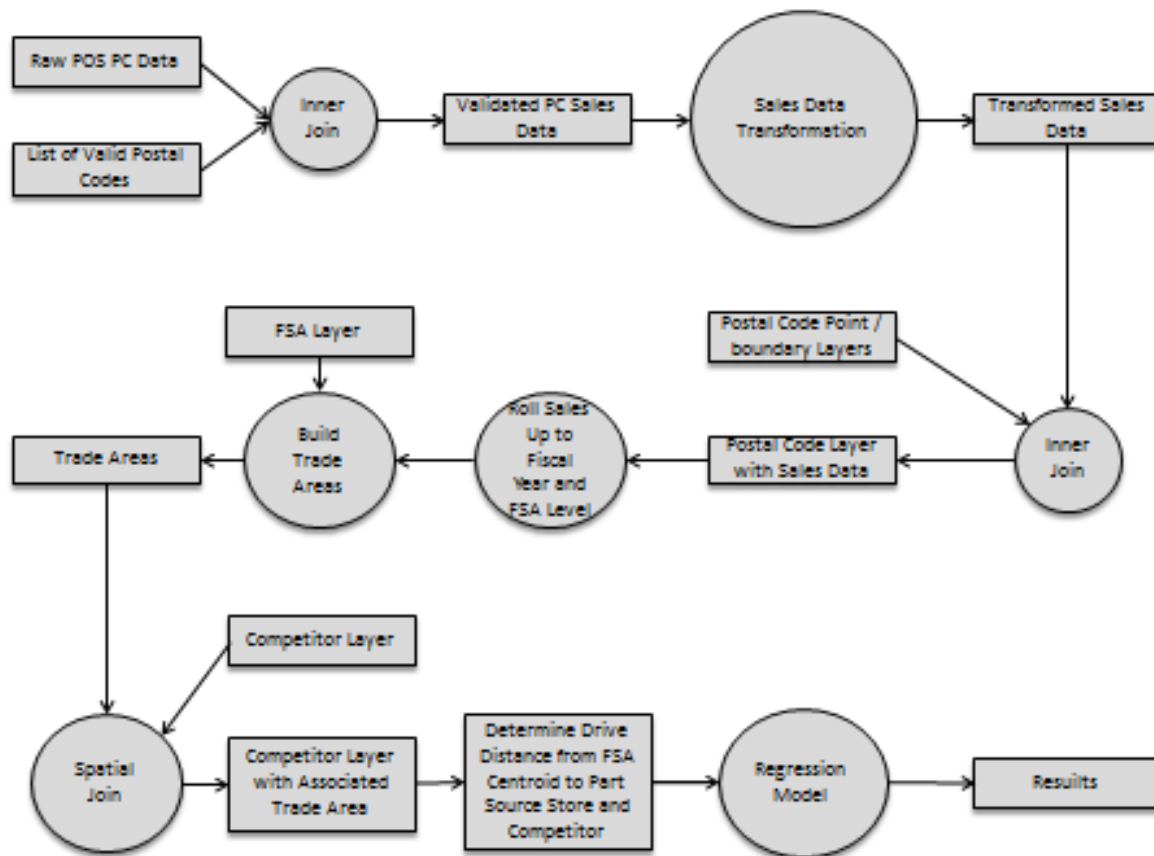
Princess Auto, which opened in February of 2015, is approximately 20,000 square feet and is located at 2251 Box Spring Road NW in Medicine Hat, Alberta (CTC, 2017). This store is approximately 6 Kilometres away from the Part Source store that is being analyzed (See Figure 4).



**Figure 4.** Map of focus store's trade area and the new competitor in Medicine Hat



## 4. Methodology



**Figure 5.** Outline transformation and analysis methodology

### 4.1 Data

Competitor locations and associated data were partially provided by CTC (Canadian Tire Corporation) and collected by web scraping. The associated data included the opening dates, retail square footage, renovation projects, and closing dates. The competitors that were analyzed to identify potential Part Source stores to focus on were Auto Value Parts store, Carquest, Lordco, Napa, Piston Ring, and Princess Auto (See Table 3).

	AB	BC	MB	NB	NF	NWT	NS	ON	PEI	QC	SK	Grand Total
Auto Value Parts Stores	15											15
CARQUEST				22	6		15	115	4	96		258
Lordco		85										85
NAPA	84	68	16	31	19	1	30	155	6	122	36	568
Piston Ring			37					5			1	43
Princess Auto	9	6	3	3	1		1	18	1		2	44
<b>Grand Total</b>	<b>108</b>	<b>159</b>	<b>56</b>	<b>56</b>	<b>26</b>	<b>1</b>	<b>46</b>	<b>293</b>	<b>11</b>	<b>218</b>	<b>39</b>	<b>1013</b>

**Table 3.** Summary of automotive retailers in Canada by province

While other businesses such as Walmart or Costco sell automotive products, these retailers do not have a primary focus on automotive products, and therefore were excluded from this study. Canadian Tire will also be excluded as it is Part Source's parent company, and accumulate sales from Canadian Tire locations via the sales of parts.

Postal Code, road network, and FSA boundary shapefiles were acquired from Stats Canada. The Part Source locations, their associated information, and sales data were provided by CTC.

#### **4.1.1 Sales Data Validation and Transformation**

Sales data is collected by a combination of staff manually at the POS (Point of Sales) system and utilizing customers who have an account. In the case of customers who have an account, the postal codes associated with each account are rarely inaccurate as the account is regularly updated. It should be noted that Canadian Tires have accounts with Part Source but the sales associated with any Canadian Tire store have been removed. This was done due to the market share provided by Canadian Tire stores will not be impacted by a competitor of Part Source entering a market because of Part Source being a banner under the CTC. In the case of

customers' postal codes being manually entered; there is a greater chance of there being an error while it is being entered. This includes, but is not limited to, postal codes not being entered, being entered incorrectly, or a postal code such as the store's postal code or a random valid postal code (e.g. "A0A0A0") are entered in place of the customer's postal code. In order to correct for the not entered and some incorrectly entered postal codes, a SQL (Structured Query Language) query was created to confirm that all postal codes that remained in the table were found in the list of postal codes from Statistics Canada (See Appendix).

In the case of random postal codes being entered, the amount of misrepresented postal codes can be reduced by putting a restriction such as postal codes having to be within 1000 kilometres of a store. While this type of restriction is applied when looking at the entire network of sales, the stores chosen for this study did not require one. In the case of the store postal code being entered, this is a limitation of the data as the stores that were chosen to be used for this study are not self-contained within their own respected postal code.

Once all records that did not have a valid associated Postal Code were removed, the remaining sales data was then grown out to represent each store's total sales for each fiscal year. The following equation was utilized to grow out the sales:

DSAPC = Distributed Sales Amount per Postal Code

ASAPC = Actual Sales Amount per Postal Code

TSSM = Total Store Sales per Month

TVPCSM = Total Valid Postal Code Sales per Month

$$DSAPC = (ASAPC * \frac{TSSM}{TVPSM})$$

Due to the data transformations requiring to be subdivided, a python script was developed to ensure efficiency (See Appendix).

#### **4.1.2 Creating Dominant Sales Layer**

In order to reduce both the amount of time and limit the amount of memory required to generate each trade area, a secondary table was created. This table consists of unique postal codes with an associated store that has the highest sales within that postal code. This was done by creating a python script to rank store sales by postal code, create and populate dominant sales tables, and create dominant sales shapefiles for each fiscal year (See Appendix). It should be noted that while postal codes are being utilized for the analysis, postal codes were reduced to the FSA level to ensure that sensitive data was not released.

### **4.2 Creating Trade Areas**

For this analysis, a threshold area methodology will be utilized, and the boundaries will be FSA rather than FSALDU. This was done to ensure that sensitive data was not released. The dominant sales layer will then be joined to FSA layer, and each store will be assigned the FSA in which it falls within. While many stores may fall within an FSA in which the sales are dominant to that store, the area may not consist of only businesses, therefore, sales would not be associated to that FSA. Once each store has been assigned to the FSA that it falls within, the surrounding FSA boundaries are then analyzed and added to the trade area if the sales are

dominant to the store that the trade area is being built for. This process continues until new FSA boundaries are no longer being added to the trade area. Once the trade area is determined, all FSA boundaries that are contained within the trade area but not assigned to the trade area will be assigned to the trade area. In order to ensure that trade areas aren't limited due to boundaries such as rivers, digital boundaries will be utilized.

### **4.3 Determining Competitors within Trade Areas**

Competitors were subdivided into stores that opened since 2013 and other. The stores that opened since 2013 were then associated with the trade area in which it falls within. While new competitors can affect multiple locations, for the purpose of this study, new competitor's effects will only be analyzed for the store which is most likely to impact. The purpose of the contiguous trade areas were to ensure that competitors that fell within the trade area were properly identified.

### **4.4 Geographic Weighted Regression**

According to Drezner et al. (2002), an interrelationship exists between the market share, market potential, cost of traveling to the site, and attractiveness of the site. In this analysis, the market share is represented by the subtracting the percent of sales per each postal code in 2016 from the percent of sales per each postal code in 2014. Market potential was calculated using the expenditure data while the cost was represented by the distance to the competitor or store. For the purpose of this study, the attractiveness was represented by the total vehicles in operation (TVIO) data derived at the postal code level. The TVIO data was chosen as both the



Part Source and Princess Auto can be more or less attractive based on the parts in which they regularly stock. A store that stocks more parts, such as batteries or headlight bulbs, for a specific customer's vehicle is considered to be more attractive than a store that does not.

In order to understand how these factors are related, a geographic weighted regression (GWR) was utilized. A GWR is considered to be an extension of an ordinary least square (OLS) (Gao and Li, 2011). Charlton and Fotheringham (2011) define the simplest form of linear regression model as:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \text{ for } i=1 \dots n$$

where  $y_i$  represents the dependent variable located at  $i$ , the independent variable is represented by  $x_i$ ,  $\varepsilon_i$  represents the error term, and  $\beta_0$  and  $\beta_1$  are the parameters.

Research has shown that findings can be misleading when an OLS is utilized for a local analysis (Chumney, 2006). A GWR has the ability to perform a local analysis with greater precision than that of an OLS (Dismuke and Lindrooth, 2006). A GWR was chosen due to this analysis requiring a more local approach. Lu et. al, (2016) define a geographic weighted regression as:

$$y_i = \beta_0(\mu_i, v_i) + \sum_{k=1}^p \beta_k(\mu_i, v_i) x_{ik} + \varepsilon_i$$

In this equation  $y_i$  represents the dependent variable located at  $i$ , the intercept parameter is represented by  $\beta_0(\mu_i, v_i)$  located at  $i$ ,  $\beta_k(\mu_i, v_i)$  represents the local regression

coefficient for the  $k$ th explanatory variable located at  $i$ , the coordinate of located at  $i$  is represented by  $(\mu_i, v_i)$ , and  $\varepsilon_i$  represents the random error located at  $i$  (Lu et. al., 2016).

An OLS and GWR are similar in the fact that they both utilize a  $R^2$  value to express the goodness of fit of the model. The  $R^2$  value has a minimum value of 0 and a maximum value of 1.

## 5. Results and Discussion

The two focus stores, located in Markham, Ontario and Medicine Hat, Alberta, will be assessed separately and compared due to two key differences in their markets. The first key difference is that Markham, Ontario has a population density of approximately 1,665 people per square kilometre while Medicine hat, Albert has a population density of 580 people per square kilometre. The second key difference is the markets that surround the focus market. Based on Census Subdivision boundaries, Markham is shares its borders with five other Census Subdivisions totaling a population of approximately 3.5 million (See Table 4).

CSD Name	Total Population	Area Sq Km	Population Density (per sq. km.)
Pickering	97,885	244	400.65
Richmond Hill	207,709	107	1,941.14
Toronto	2,849,222	665	4,284.38
Vaughan	322,525	288	1,119.93
Whitchurch-Stouffville	44,180	219	201.92

**Table 4.** Summary of Census Subdivisions that border Markham

While Markham, Ontario is surrounded by larger and smaller markets, Medicine Hat, Alberta's Census Subdivision boundaries only intersect with two other Census Subdivisions

which have a total population of approximately 15,000. Both of these markets are smaller and have a lower population density (See Table 5).

CSD Name	Total Population	Area Sq Km	Population Density (per sq. km.)
Cypress County	8,215	13,382	0.6
Redcliff	6,242	16	384.4

**Table 5.** Summary of Census Subdivisions that border Medicine Hat

With these two key differences, the trade areas for the two focus stores are significantly different in total area coverage and population density. As shown in Table , the population of the two trade areas are similar, the trade area size and subsequent population density is significantly different.

Trade Area	Total Population	Trade Area Size (Sq. Km.)	Population Density (per Sq. Km.)
Markham	193,213	101	1,913.0
Medicine Hat	210,816	119,628	1.8

**Table 6.** Summary of trade areas that were utilized

This means that these stores are utilized more effectively if they are treated as separate entities and contrasted rather than building one model to represent the changes in sales data for both stores.

Before analyzing changes within each specific postal code, there are a few key pieces of information which should be addressed. This information includes the opening date of competitor, sales data that will be utilized and the transiency of sales. For the purpose of this study, transient sales are defined as all sales that fall outside of the trade area boundaries for the focus store. The competitor in each of the focus stores` trade areas, the competitor`s grand

opening date was in 2015; based on Part Source's fiscal calendar dates. Because of this, sales from Part Sources fiscal 2014 and 2016 years will be utilized. The transient sales for the store in Markham were calculated to be approximately 0.8% in 2014 and increased to approximately 8.8% in 2016. The store in Medicine Hat had approximately 4.7% transient sales in 2014 and approximately 5.4% transient sales in 2016. The difference in total area covered by each trade area is the primary factor to explain the additional 7.3% increase in transient sales that Markham had compared to Medicine Hat between 2014 and 2016.

### 5.1 Initial Assessment and Statistics of Competitor Impacts

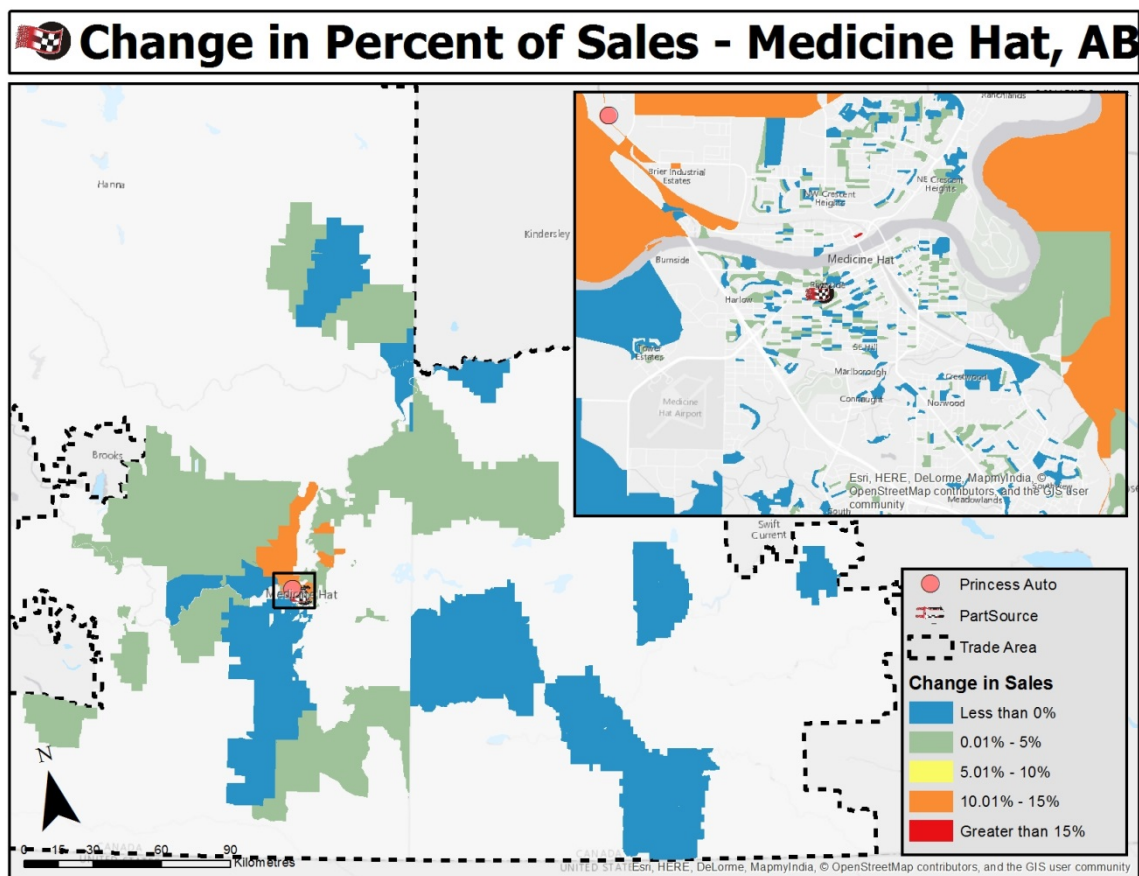
In order to ensure that trade areas cannot be further derived from the FSA level, all subsequent analyses will only include postal codes which had sales present in both 2014 and 2016. The table below outlines the statistics of postal codes which were included within the study.

Focus Store	2014		2016	
	% of Sales	% of Postal Codes With Sales	% of Sales	% of Postal Codes With Sales
Markham	59.0%	75.0%	63.6%	36.0%
Medicine hat	74.0%	50.0%	72.5%	6.9%

**Table 7.** Summary of sales included in the analysis

### 5.1.1 Medicine Hat, Alberta

In the case of the store in Medicine Hat, Alberta, the highest negative change in sales between 2014 and 2016 was approximately -3.8% and the highest positive change in sales within the same years was approximately 29.2%. The average change in sales was -0.003% and the median change was 0.001%.



**Figure 6.** Map of the change in sales between 2014 and 2016 in Medicine Hat

Due to the size of the sample of changes in sales for this focus store, there were a few spatial patterns that were observed. With regards to postal codes within the Census Subdivision boundaries it was observed that sales remain relatively unchanged with the exception of postal

codes surrounding the new competitor. These postal codes had a positive increase of 12.5 percent of total sales within each of these postal codes. Due to this phenomenon, further analyses were conducted utilizing the distance to the competitor and contrasting the findings against the same analyses utilizing distance to the focus store.

In the case of distances from the competitor to surrounding postal codes with sales data, it was found that the sales from postal codes closer to the competitor represented more of the stores total sales after the competitor opened than they had prior. While postal codes closer to the competitor were increased once the competitor opened, postal codes represented less the further away they were from the competitor.

	Within 5 kilometres		Within 10 kilometres		Within 20 kilometres		Greater than 20 kilometres	
	Distance (KM)	Change in Sales (%)	Distance (KM)	Change in Sales (%)	Distance (KM)	Change in Sales (%)	Distance (KM)	Change in Sales (%)
Average	4.3	0.3	6.9	0.0	11.1	-0.1	123.4	-0.3
Minimum	2.1	-1.2	5.0	-3.4	10.0	-1.9	28.4	-3.8
Maximum	5.0	13.0	10.0	29.2	14.5	0.3	244.5	0.3
Sum		20.4		2.3		-6.9		-8.7

**Table 8.** Summary of sales by distance to competitor in Medicine Hat

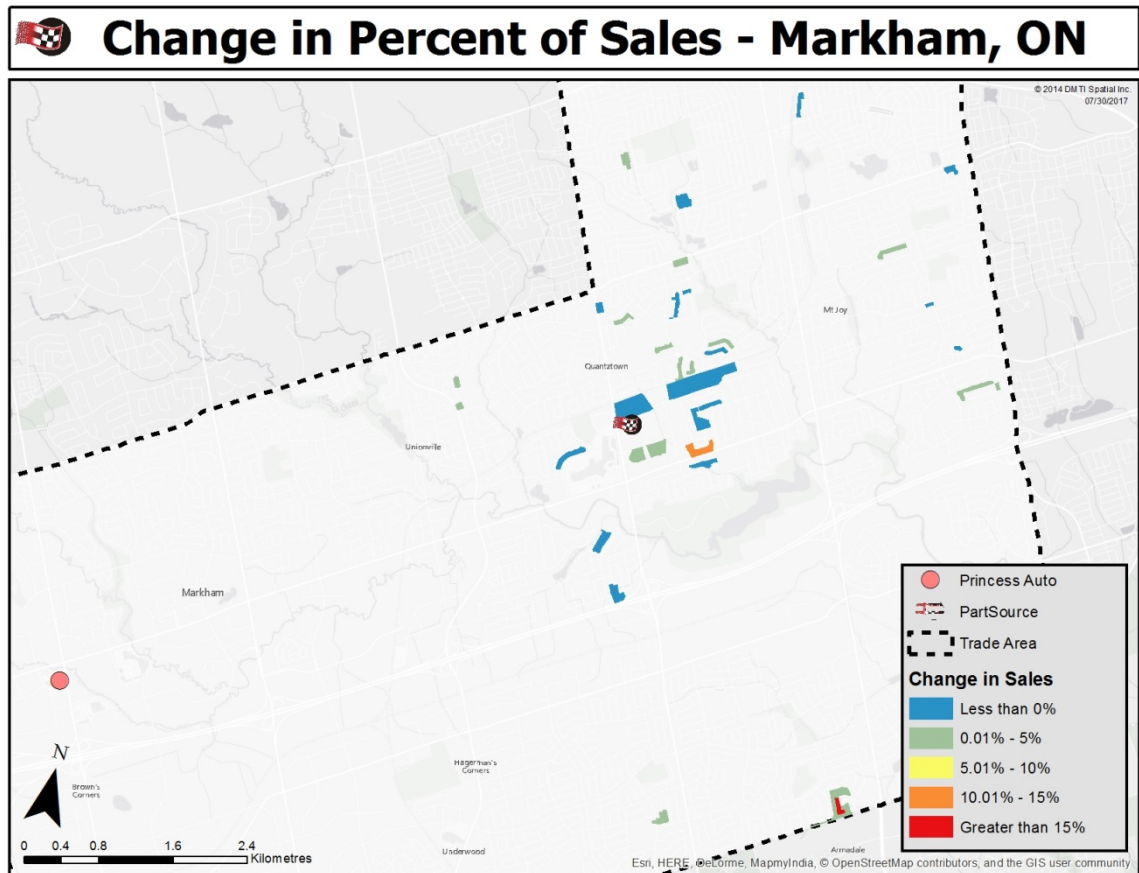
Unlike the analysis above, the distance from Part Source to postal codes with associated changes in sales did not follow the same spatial patterns. Sales within five kilometres and greater than 20 kilometres were found to represent fewer sales after the competitor opened while sales between 10 and 20 kilometers represented a larger percent of sales.

	Within 5 kilometres		Within 10 kilometres		Greater than 20 kilometres		
	Distance (KM)	Change in Sales (%)	Distance (KM)	Change in Sales (%)	Distance (KM)	Change in Sales (%)	
Average	2.6	-0.01	6.2	0.3	121.6	-0.3	
Minimum	0.1	-3.4	5.0	-1.9	25.2	-3.8	
Maximum	5.0	29.2	9.9	13.0	243.9	0.3	
Sum		-2.9		18.7		-8.7	

**Table 9.** Summary of sales by distance to focus store in Medicine Hat

### **5.1.2 Markham, Ontario**

In the case of the store in Markham, Ontario the highest negative change in sales between 2014 and 2016 was approximately -5.1 percent and the highest positive change in sales within the same years was approximately 22.7 percent. The average and median change in sales were both positive and were 0.4 percent and 0.3 percent respectively.



**Figure 7.** Map of the change in sales between 2014 and 2016 in Markham

Due to the limitations of the data available for analysis, the initial visual inspection does not have any spatial patterns that are easily identified. Like the analysis for Medicine Hat, the following analysis was conducted initially using distance from postal codes to competitor and then was redone using the distance from postal codes to the focus store. Along with the visual inspection being limited due to the data available, the data available does not allow for the same distance breaks to be utilized for both of these analyses. Because of this, natural breaks in the distance calculated for each of the postal code to the competitor or focus store were utilized.



In the case of the distance from postal codes with sales data to the Princess Auto, it was found that the further away the postal codes were, the more they positively increased. The closer the postal codes were to the Princess Auto, the more the sales negatively increased. The table below is based on natural breaks within the data.

	Within 7.5 Kilometres		Within 10 kilometres		Greater than 10 kilometres	
	Distance (KM)	Change in Sales (%)	Distance (KM)	Change in Sales (%)	Distance (KM)	Change in Sales (%)
Average	6.7	-1.2	8.6	0.1	10.9	2.3
Minimum	6.3	-3.1	7.9	-5.1	10.4	-2.8
Maximum	7.4	0.1	9.9	14.6	12.1	22.7
Sum		-7.4		1.3		20.3

**Table 10.** Summary of sales by distance to competitor in Markham

Unlike the focus store in Medicine Hat, utilization of the distance from postal code to the Part Source location produced a similar trend to the analysis above. Postal codes that are close to the Part Source had a negative increase in represented sales. As the distance increased, the represented sales positively increased as well.

	Within 750 Meters		Within 1 kilometres		Greater than 1 kilometres	
	Distance (KM)	Change in Sales (%)	Distance (KM)	Change in Sales (%)	Distance (KM)	Change in Sales (%)
Average	1.3	-0.7	2.2	-0.05	5.1	2.6
Minimum	0.7	-3.3	1.6	-5.1	4.0	-2.1
Maximum	1.6	4.4	3.2	14.6	6.2	22.7
Sum		-5.4		-0.5		23.2

**Table 11.** Summary of sales by distance to focus store in Markham

## 5.2 Determining Factors of the effect of Competitors

In order to further understand the effects of a competitor influencing postal code sales, it is important to understand what factors are correlated. To ensure that the variables that are selected have a clear reason for being selected, the correlation between the change in sales represented by each postal code, the cost to the customer to the competitor or the focus stores, the attractiveness of a store and the market potential must be understood. While the cost is determined by the drive distance to a location, the attractiveness and market potential will be determined through expenditure data and Total Vehicles In Operation (TVIO) data.

The three expenditure variables that were selected were; “Accessories for automobiles, vans and trucks”, “tires, batteries, and other parts and supplies for vehicles”, and “Maintenance and repairs of vehicles”. The TVIO data has been divided into domestic and import, sedan or pickup, and the year the vehicle was manufactured; based on five year intervals. Each of these variables will be modeled and compared. All models that are created for the subsequent analyses of the markets in Medicine Hat and Markham will be geographic weighted regression models. The interpretation column was determined based on the difference between the minimum and maximum adjusted  $R^2$  values and divided based on equal breaks into five classifications; Very Weak (less than 0.313), Weak (0.314 – 0.466), Moderate (0.467 – 0.619), Strong (0.62 – 0.772), and Very Strong (greater than 0.723).

### 5.2.1 Medicine Hat, Alberta

	Variables	Adjusted R <sup>2</sup>	Interpretation
Expenditure Data	HSTR058	0.925	Very Strong
	HSTR033	0.922	Very Strong
	HSTR034	0.899	Very Strong

**Table 12.** Summary statistics of GWR of expenditure variables in Medicine Hat

Based on the results presented in Table 12, Accessories for automobiles, vans, and trucks variable had the highest R<sup>2</sup> value out of the three expenditure variables and explained 92.5 percent of the model. The expenditure variable “Maintenance and repairs of vehicles” explained 89.9 percent of the model which was the lowest percent compared to the other two variable in the same category. All three variables were considered to have very strong interpretation (See Table 12).

	Variables	Adjusted R <sup>2</sup>	Interpretation
Top 5 TVIO Variables	Import Truck 1986 - 1990	0.761	Strong
	Import Car 2011 - 2015	0.751	Strong
	Import Truck 2011 - 2015	0.744	Strong
	Domestic Car 2011 - 2015	0.743	Strong
	Import Truck 2006 - 2010	0.742	Strong

**Table 13.** Summary statistics of GWR of top five TVIO variables in Medicine Hat

The results presented in table 13 are the five highest correlated. Imported cars that were built from 1986 to 1990 and 1996 to 2000 and Domestic trucks that were built from 2001 and 2005 had the lowest R-squared values at 0.628, 0.686, and 0.696 respectively. The average R-square value for all 58 variables was 0.724 (see appendix).

	Variables	Adjusted R <sup>2</sup>	Interpretation
Top 3 Imported Car Segments	2011 - 2015	0.751	Strong
	2006 - 2010	0.739	Strong
	1991 - 1995	0.738	Strong

**Table 14.** Summary statistics of GWR of top three categories of imported cars in Medicine Hat

The results presented in table 14 are the top 3 segments of imported cars based on the R-squared values. The three variables that had the lowest R-squared values were import cars built from 2011 to 2015, 2006 to 2010, and 1991 to 1995 (See Appendix). The average R-squared value was 0.712.

	Variables	Adjusted R <sup>2</sup>	Interpretation
Top 3 Domestic Car Segments	2011 - 2015	0.743	Strong
	1991 - 1995	0.737	Strong
	1981 - 1985	0.734	Strong

**Table 15.** Summary statistics of GWR of top three categories of domestic cars in Medicine Hat

The results presented in table 15 are the top 3 segments of domestic cars based on the R-squared values. The three variables that had the lowest R-squared values were import cars built from 2006 to 2010, 1986 to 1990, and 1996 to 2000 (See Appendix). The average R-squared value was 0.730.

	Variables	Adjusted R <sup>2</sup>	Interpretation
Top 3 Imported Truck Segments	1986 - 1990	0.761	Strong
	2011 - 2015	0.744	Strong
	2006 - 2010	0.742	Strong

**Table 16.** Summary statistics of GWR of top three categories of imported trucks in Medicine Hat

The results presented in table 16 are the top 3 segments of imported trucks based on the R-squared values. The three variables that had the lowest R-squared values were import cars built from 1981 to 1985 and 1991 to 2000 (See Appendix). The average R-squared value was 0.733.

	Variables	Adjusted R <sup>2</sup>	Interpretation
Top 3 Domestic Trucks Segments	2011 - 2015	0.734	Strong
	1991 – 1995	0.728	Strong
	2006 – 2010	0.721	Strong

**Table 17.** Summary statistics of GWR of top three categories of domestic trucks in Medicine Hat

The results presented in table 17 are the top 3 segments of imported trucks based on the R-squared values. The three variables that had the lowest R-squared values were import cars built from 1981 to 1985, and 1996 to 2005 (See Appendix). The average R-squared value was 0.715.

### 5.2.2 Markham, Ontario

	Variables	Adjusted R <sup>2</sup>	Interpretation
Expenditure Data	HSTR058	0.160	Very Weak
	HSTR033	0.269	Very Weak
	HSTR034	0.643	Strong

**Table 18.** Summary statistics of GWR of expenditure variables in Markham

Unlike the R<sup>2</sup> values that were calculated for expenditure variables in Medicine Hat, the variable that explained the least percent of the model was the “Accessories for automobiles, vans, and trucks”. “Maintenance and repairs of vehicles” explained 64.3 percent which was the highest percent of the model to be explained by any of the expenditure variables (See Table 18).

	Variables	Adjusted R <sup>2</sup>	Interpretation
Top 5 TVIO Variables	Domestic Car 1996 - 2000	0.503	Moderate
	Domestic Truck 2001 – 2005	0.499	Moderate
	Domestic Truck 1996 - 2000	0.486	Moderate
	Domestic Car 2001 – 2005	0.486	Moderate
	Domestic Truck 2006 - 2010	0.485	Moderate

**Table 19.** Summary statistics of GWR of top five TVIO variables in Markham

The results presented in table 19 are the five highest correlated. Imported cars that were built from 1981 to 1986, domestic cars that were built from 1986 to 1990 and Import trucks that were built from 1981 to 1985 had the lowest R-squared values at 0.007, 0.076, and 0.081 respectively (See Appendix). The average R-square value for all 58 variables was 0.724.

	Variables	Adjusted R <sup>2</sup>	Interpretation
Top 3 Imported Car Segments	1986 – 1990	0.468	Moderate
	2001 – 2005	0.450	Weak
	1996 – 2000	0.434	Weak

**Table 20.** Summary statistics of GWR of top three categories of imported cars in Markham

The results presented in table 20 are the top 3 segments of imported cars based on the R-squared values. The three variables that had the lowest R-squared values were import cars built from 1981 to 1985, 1991 to 1995, and 2011 to 2015 (See Appendix). The average R-squared value was 0.338.

	Variables	Adjusted R <sup>2</sup>	Interpretation
Top 3 Domestic Car Segments	1996 - 2000	0.503	Moderate
	2001 - 2005	0.486	Moderate
	2011 - 2015	0.418	Weak

**Table 21.** Summary statistics of GWR of top three categories of domestic cars in Markham

The results presented in table 21 are the top 3 segments of domestic cars based on the R-squared values. The three variables that had the lowest R-squared values were import cars built from 1981 to 1995 (See Appendix). The average R-squared value was 0.322

	Variables	Adjusted R <sup>2</sup>	Interpretation
Top 3 Imported Truck Segments	1996 – 2000	0.451	Weak
	2001 – 2005	0.342	Weak
	1986 – 1990	0.189	Very Weak

**Table 22.** Summary statistics of GWR of top three categories of imported trucks in Markham



The results presented in table 22 are the top 3 segments of imported trucks based on the R-squared values. The three variables that had the lowest R-squared values were import cars built from 1981 to 1985, 1991 to 1995, and 2006 to 2010 (See Appendix). The average R-squared value was 0.215.

	Variables	Adjusted R <sup>2</sup>	Interpretation
Top 3 Domestic Trucks Segments	2011 - 2015	0.734	Strong
	1991 - 1995	0.728	Strong
	2006 - 2010	0.721	Strong

**Table 23.** Summary statistics of GWR of top three categories of domestic trucks in Markham

The results presented in table 23 are the top 3 segments of imported trucks based on the R-squared values. The three variables that had the lowest R-squared values were import cars built from 1981 to 1985, and 1996 to 2005 (See Appendix). The average R-squared value was 0.715.

## 6. Conclusion

### 6.1 Findings

Based on each of the competitor that opened in relation to the drive distance a given postal code, the initial findings were that the Princess Auto that opened in Medicine Hat, Alberta had increased sales between 2014 and 2016 within 5 km of the Princess Auto. As

further away postal codes with a change in sales between 2014 and 2016 are analyzed, the sales are reduced and become negative. This is the opposite effect that the Princess Auto that opened in Markham, Ontario had. Postal codes with associated sales data that are close to this Princess Auto were more likely to be negative than postal codes that were further away.

Based on each of the focus stores in relation to the drive distance a given postal code, the initial findings were that the Princess Auto that opened in Medicine Hat, Alberta had less of a linear change in sales compared to the competitor focused analysis. Within five kilometres, sales were found to have had a slightly negative change between 2014 and 2016. Between five and 10 kilometres, sales had a positive change between 2014 and 2016. Sales greater than 10 kilometres, had a negative change between 2014 and 2016. While the focus store in Medicine Hat, Alberta had a different spatial pattern than that of its competitor, the focus store in Markham, Ontario had fundamentally the same results as the competitor within its market.

Based on the analyses that had a spatial linear pattern, the differences in the two markets can be explained by the size of the trade area. The trade area that was generated for Medicine Hat, which covered approximately 119,628 square kilometres, was significantly larger than the trade area that was generated for Markham, Ontario which covered approximately 101 square kilometres.

The geographic weighted regression models that were created for each of the TVIO data allowed for an understanding about which vehicles within the market were most correlated with the positive or negative change in sales. In the case of Medicine Hat, Alberta, Imported Trucks built from 1986 to 1990 and 2006 to 2015 and both domestic and import cars built from

2011 to 2015 were the top five categories with the highest correlation. Unlike imported trucks and new cars being the highest correlated categories with the sales changes, the focus store's changes in sales in Markham, Ontario had the highest correlation with domestic cars built from 1996 to 2005 and domestic trucks built from 1996 to 2010.

## **6.2 Limitations**

While the findings outline above allowed for understanding and modeling the impacts of new competitors, it is important to understand the limitations of the data as well. The first limitation was the sales data. The sales data that was utilized was a sample of the data collected by a combination of staff manually at the POS (Point of Sales) system and utilizing customers who have an account. As outlined in section 4.1.1 Sales Data Validation, the most common problems that occur are postal codes not being entered, being entered incorrectly, or a postal code such as the stores postal code or a random valid postal code "A0A0A0" are entered in place of the customer's postal code. Along with these limitations, there is also the limitation of the data being a sample size of the total store sales. This means that there are potential postal codes that should have sales associated with them. This limits the amount of postal codes that were analyzed and the postal codes that were analyzed had the potential to be an inaccurate representation of sales associated with each postal code.

The next limitation is the FSA boundaries that were utilized when the trade areas were created. While these boundary files were advantageous in ensuring that more competitors were captured within a trade area, it also increased the probability that trade areas would intersect each other. A few potential focus stores were excluded due to the requirement that

all trade areas that intersect with the focus store's trade area are not representative of stores that have been renovated, relocated, expanded, or built from 2013 and 2015.

Along with the trade areas limiting potential study areas, the competitor data that was used within the study only had 1.3% of the records with opening dates. Due to this, both focus stores were analyzed based on a Princess Auto opening. This means that the findings within this study may only be applicable to Princess Auto in relation to Part Source rather than automotive competitors in general.

The final limitation was the TVIO data which was derived from data at the dissemination area level. Because of this, the data had been evenly distributed based on number of households in each postal code. While there was no analysis done to calculate how this skewed the results, it was determined that this had an impact on the analysis.

### **6.3 Future Research**

Any future research that is conducted should attempt to minimize or eliminate all the limitations that were stated above and conduct the methodology of the study again. With this in mind, there are a few recommended approaches for future research. The first is to transform the sales data to the dissemination area level and rerun the trade area methodology. This will limit the size of the trade area, which will be less likely to intersect with other trade areas and allow for the TVIO data to be analyzed without having to be transformed. While the sales data will become less accurate, it is believed that the increase in accuracy of the TVIO data will outweigh the increased inaccuracy of the sales data at the dissemination area level. The next is

to purchase or conduct a more thorough sweep of various media outlets to increase the amount of competitor records that have opening dates.

This study and the addition of the recommendations above can be useful in many different retail applications. To start, Part Source could utilize this information to help better predict the impact of a competitor and could attempt to mitigate the effects of a new competitor by creating a promotion that targets specific customers. An example of this would be if the store in Markham, Ontario had a promotion of 10% off parts to all customers that own a vehicle that was built from 1996 to 2010. Another example would be if the store in Medicine Hat, Alberta had a promotion of 10% off parts to all customers that own an import vehicle. In both scenarios, the targeted customers would be less likely to go to the new Princess Auto. While Part Source could utilize this information to interpret and take precautions to minimize the effects of a new competitor, the same competitors can utilize this information to understand which markets have the potential to be more or less difficult to enter. An example of this would be if Princess Auto ensured that they have parts for import truck and car are always in stock in their store in Medicine Hat, Alberta to maximize the amount of sales captured from Part Source sales. While this study specifically pertains to the effects of a new competitor on Part Source, the methodology can be applied and utilized by the majority of retailers. If Canadian Tire were to utilize the same methodology to determine the effects of a competitor, such as Walmart, opening up within a given market, the sales data could be subdivided based on department. This, in combination with a larger demographic dataset, could be used to develop a refined definition of which customers are most at risk to be effected by a new Walmart.

This methodology can also be applied to other sectors such as health or public safety. In the case of health, if a hospital were to be shut down then the effects could be measured using this methodology and substituting sales with a variable such as the count of patients with a specific health issue. In the case of public safety, if a new precinct were to be built, this methodology could be applied to understand both the general effects on crime and specific crimes being impacted. An example of this would be to substitute the sales variable for the number of cars that have been broken into.

It is also important to note that while the python scripts were developed using the module “arcpy”, which is only used for ArcMap software, the module is not unique and the scripts can be easily modified for use in software such as QGIS. The SQL queries are able to be used with modifications to the column names. In both cases, the advantage to develop custom programs in an open source language is that it can be easily adopted across platforms. This is advantageous in situations where refined methodologies must be preserved during a change such as a company switching software due to budget cuts.

## References

- About Us. (2016). Retrieved July 04, 2017, from <http://partsource.ca/content/microsites/partsource/en/about.html>
- Anderson, E., & Coughlan, A. T. (1987). International market entry and expansion via independent or integrated channels of distribution. *The Journal of Marketing*, 71-82.
- Anitsal, I., Anitsal, M. M., & Girard, T. (2013). Retail Mission Statements: Top 100 Global Retailers. *Academy of Strategic Management Journal*, 12(1), 1.
- Ayad A. (2008). Optimizing inventory and store results in big box retail environment. *International Journal of Retail & Distribution Management*. 36(3): 180-191.
- Brunsdon, C.; Fotheringham, A.S.; Charlton, M.E. (1996). Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*, 28(4):281-298.
- Brunsdon, C.; Fotheringham, S.; Charlton, M. (1998). Geographically Weighted Regression-Modelling Spatial Non-Stationarity. *Journal of the Royal Statistical Society*, 47(3):431-443.
- Brunsdon, C.; Fotheringham, S.; Charlton, M. (2003). *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons.
- Canada, N. (2011). Walmart canada announces its next québec supercentres. Canada Newswire.
- Canada, N. (2011). Walmart canada completes acquisition of leases for 39 stores. Canada Newswire
- Canadian Tire Corporation. (2016). Retrieved July 04, 2017, from <http://corp.canadiantire.ca/EN/AboutUs/Pages/default.aspx>
- Charlton, M., & Fotheringham, A. S. Geographically Weighted Regression.
- Chen, Y. (2015). The distance-decay function of geographical gravity model: Power law or exponential law?. *Chaos, Solitons & Fractals*, 77, 174-189.
- City of Markham - Facts & Stats. (2014). Retrieved July 23, 2017, from <https://www.markham.ca/wps/portal/Markham/AboutMarkham/FactsStats/>
- Chumney, E. C. (2006). *Methods and designs for outcomes research*. ASHP.
- Clarke, I. (2000). Retail power, competition and local consumer choice in the UK grocery sector. *European Journal of Marketing*. 34(8): 975-1002.

- Dawson, J., & Mukoyama, M. (2006). The international transfer of key success factors. *Dawson, J, Larke, R & Jukoyama, M (2006) Strategic issues in international retailing. Routledge*, 196-213.
- Dismuke, C., & Lindrooth, R. (2006). Ordinary least squares. *Methods and Designs for Outcomes Research*, 93, 93-104.
- Dolega, L., Pavlis, M., & Singleton, A. (2016). Estimating attractiveness, hierarchy and catchment area extents for a national set of retail centre agglomerations. *Journal of Retailing and Consumer Services*. 28(2016): 78-90.
- Drezner, T., & Drezner, Z. (2002). Validating the gravity-based competitive location model using inferred attractiveness. *Annals of Operations Research*, 111(1), 227. Retrieved from <http://ezproxy.lib.ryerson.ca/login?url=http://search.proquest.com.ezproxy.lib.ryerson.ca/docview/214508176?accountid=13631>
- Goi, C. L. (2009). A review of marketing mix: 4Ps or more?. *International journal of marketing studies*, 1(1), 2.
- Flater, D. (2012, January 23). Huff Model. Retrieved July 01, 2017, from <https://www.arcgis.com/home/item.html?id=f4769668fc3f486a992955ce55caca18>
- Foote, K, E., & Margaret, L. (1995). Context, Concepts, and Definitions. Retrieved June 19, 2017, from <http://www.colorado.edu/geography/gcraft/notes/intro/intro.bak16>
- Gao, J.; Li, S. (2011). Detecting Spatially Non-Stationary and Scale-Dependent Relationships Between Urban Landscape Fragmentation and Related Factors Using Geographically Weighted Regression. *Applied Geography*, 31:292-302.
- Harris, P., Fotheringham, A., & Juggins, S. (2010). Robust Geographically Weighted Regression: A Technique for Quantifying Spatial Relationships Between Freshwater Acidification Critical Loads and Catchment Attributes. *Annals of the Association of American Geographers*, 100(2), 286-306. Retrieved from <http://www.jstor.org/stable/40645355>
- History. (2017). Retrieved June 22, 2017, from <https://www.homedepot.ca/en/home/corporate-information/history.html>
- Hollander, S. C. (2002). Notes on the retail accordion. *Retailing: The evolution and development of retailing*, 1, 322.
- How Original Huff Model Works. (2016). Retrieved June 20, 2017, from <http://desktop.arcgis.com/en/arcmap/10.3/tools/business-analyst-toolbox/how-original-huff-model-works.htm#GUID-8F1142E0-D329-409B-BC37-98DF9FEB2453>
- Hu, M., Li, Z., Wang, J., Jia, L., Liao, Y., Lai, S., ... & Yang, W. (2012). Determinants of the incidence of hand, foot and mouth disease in China using geographically weighted regression models. *PloS one*, 7(6), e38978.



- Hughes, R., Hallsworth, A. G., & Clarke, G. (2009). Testing the effectiveness of the proposed UK 'competition test'. *The Service Industries Journal*, 29(5), 569-590.
- Kerr, J., Frank, L., Sallis, J.F., Saelens, B., Glanz, K., Chapman, J. (2012). Predictors of trips to food destinations. *International Journal of Behavioral Nutrition and Physical Activity*. 9(58): 1-10.
- Lowes (2017). LOWE'S COMPLETES ACQUISITION OF RONA. Retrieved August 07, 2017, from [https://www.lowes.ca/articles/lowes-completes-acquisition-of-rona\\_a6524.html](https://www.lowes.ca/articles/lowes-completes-acquisition-of-rona_a6524.html)
- Lu, B., Charlton, M., & Fotheringham, A. S. (2011). Geographically weighted regression using a non-Euclidean distance metric with a study on London house price data. *Procedia Environmental Sciences*, 7, 92-97.
- Medicine Hat - City, Alberta, Canada. (2012). Retrieved July 23, 2017, from <http://www.city-data.com/canada/Medicine-Hat-City.html>
- Mishra, S. (2009). GIS in Indian retail industry-a strategic tool. *International Journal of Marketing Studies*, 1(1), 50.
- Mun, L. Y., & Yazdanifard, R. (2012). Walmart success in Mexico, Canada and China: global expansion, strategies, entry modes, threats and opportunities.
- Öcal, N., & Yildirim, J. (2010). Regional effects of terrorism on economic growth in Turkey: A geographically weighted regression approach. *Journal of Peace Research*, 47(4), 477-489. Retrieved from <http://www.jstor.org/stable/20752202>
- Singh, V., Hansen, K., & Blattberg, R. (2006). Market Entry and Consumer Behavior: An Investigation of a Wal-Mart Supercenter. *Marketing Science*, 25(5): 457-476.
- Statistics Canada (2016). Focus on Geography Series, 2011 Census. Retrieved July 23, 2017, from <http://www12.statcan.gc.ca/census-recensement/2011/as-sa/fogs-spg/Facts-csd-eng.cfm?LANG=Eng&GK=CSD&GC=3519036>
- Statistics Canada (2017, April 14). Census Profile, 2016 Census Medicine Hat [Population centre], Alberta and Saskatchewan [Province]. Retrieved July 23, 2017, from <http://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/page.cfm?Lang=E&Geo1=POPC&Code1=0523&Geo2=PR&Code2=47&Data=Count&SearchText=Medicine Hat&SearchType=Begins&SearchPR=01&B1=All>
- Stabell, C., & Fjeldstad, Ø. (1998). Configuring Value for Competitive Advantage: On Chains, Shops, and Networks. *Strategic Management Journal*, 19(5), 413-437. Retrieved from <http://www.jstor.org/stable/3094221>

- Vaz, E., & Arsanjani, J. J. (2015). Predicting Urban Growth of the Greater Toronto Area- Coupling a Markov Cellular Automata with Document Meta-Analysis. *Journal of Environmental Informatics*, 25(2), 71-80.
- Zhang, L.; Bi, H.; Cheng, P.; Davis, C.J. (2004). Modeling Spatial Variation in Tree Diameter-Height Relationships. *Forest Ecology and Management*, 189:317-329.

## Appendix A

**Figure A1.** Validate Postal Codes

```
SELECT PS_Sales_By_Year.FISCAL_YEAR, PS_Sales_By_Year.STORE_NUMBER,  
PS_Sales_By_Year.FSA, PS_Sales_By_Year.SumOfRETAIL_TOTAL_SALES_AMT INTO  
PS_Sales_By_Year_Checked  
  
FROM PS_Sales_By_Year INNER JOIN Unique_PCs ON PS_Sales_By_Year.FSA =  
Unique_PCs.POSTALCODE  
  
GROUP BY PS_Sales_By_Year.FISCAL_YEAR, PS_Sales_By_Year.STORE_NUMBER,  
PS_Sales_By_Year.POSTAL_CODE, PS_Sales_By_Year.SumOfRETAIL_TOTAL_SALES_AMT  
  
HAVING (((PS_Sales_By_Year.SumOfRETAIL_TOTAL_SALES_AMT)>0));
```

**Figure A2.** Create New Unique FSA Table

```
SELECT Left([POSTALCODE],3) AS FSA INTO CANMEP_FSA  
  
FROM GISOWNER_POSTAL_CANMEP  
  
GROUP BY Left([POSTALCODE],3);
```

**Figure A3.** Create unique postal code table

```
SELECT GISOWNER_POSTAL_CANMEP.POSTALCODE,  
GISOWNER_POSTAL_CANMEP.POST_CODE INTO Unique_PCs  
  
FROM GISOWNER_POSTAL_CANMEP  
  
WHERE (((GISOWNER_POSTAL_CANMEP.SLI)=1))  
  
GROUP BY GISOWNER_POSTAL_CANMEP.POSTALCODE,  
GISOWNER_POSTAL_CANMEP.POST_CODE;
```

**Figure A4.** Create table of positive sales by postal code

```
SELECT PS_Sales_By_Year.FISCAL_YEAR, PS_Sales_By_Year.STORE_NUMBER,  
PS_Sales_By_Year.FSA, PS_Sales_By_Year.SumOfRETAIL_TOTAL_SALES_AMT INTO  
PS_Sales_By_Year_Checked  
  
FROM PS_Sales_By_Year INNER JOIN Unique_PCs ON PS_Sales_By_Year.FSA =  
Unique_PCs.POSTALCODE  
  
GROUP BY PS_Sales_By_Year.FISCAL_YEAR, PS_Sales_By_Year.STORE_NUMBER,  
PS_Sales_By_Year.POSTAL_CODE, PS_Sales_By_Year.SumOfRETAIL_TOTAL_SALES_AMT  
  
HAVING (((PS_Sales_By_Year.SumOfRETAIL_TOTAL_SALES_AMT)>0));
```

**Figure A6.** Postal code sales data transformation

#Developer: Scotty Norman

#Purpose: Sales Data Transformation

#Usage: Major Research Paper

```
import arcpy
from arcpy import env
import os
import time
import traceback
from datetime import datetime
import sys
arcpy.AddMessage("Script Started- 1")
arcpy.AddMessage("Modules have been imported")

CrtDB = "Database Connections\\MRP_PS_SLS_DATA_PC.odc\\" #Connect to P2CTT Database
tableExists = arcpy.Exists(CrtDB) #Function for checking if a table exists
if tableExists == False: #If it DOES NOT exist
    raise NameError(CrtDB + "Does not exist!") #Add message raising error
else:
    arcpy.AddMessage("Connection to DB Created") #If it DOES exist

PC_SALES_Table = CrtDB + "PS_Sales_By_Year_Checked" #Creating variable to identify
table
TOTAL_SALES_Table = CrtDB + "Store_Totals" #Creating variable to identify table

if arcpy.Exists("M:/gis/Scott Norman/MRP/GDBs/04 - Apr/PS_5Yr_Sales_Transformed.gdb"):
    arcpy.AddMessage("File Found, Deleting Table Now")
    arcpy.Delete_management("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb")
else:
    pass
if arcpy.Exists("M:/gis/Scott Norman/MRP/GDBs/04 - Apr/PS_5Yr_Sales_Transformed.gdb"):
    arcpy.AddMessage("Delete didn't work")
else:
    arcpy.AddMessage("Delete complete")

arcpy.CreateFileGDB_management("M:/gis/Scott Norman/MRP/GDBs/04 - Apr/",
"PS_5Yr_Sales_Transformed.gdb")
if arcpy.Exists("M:/gis/Scott Norman/MRP/GDBs/04 - Apr/PS_5Yr_Sales_Transformed.gdb"):
    arcpy.AddMessage("Created GDB file")
else:
    arcpy.AddMessage("GDB Not Created")
```

```

arcpy.CreateTable_management("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb","Converted_Sales")
if arcpy.Exists("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb/Converted_Sales"):
    arcpy.AddMessage("Converted_Sales created")
else:
    arcpy.AddMessage("Converted_Sales did not create")

arcpy.AddField_management("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb/Converted_Sales","FISCAL_YEAR","LONG")
arcpy.AddField_management("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb/Converted_Sales","STORE_NO","LONG")
arcpy.AddField_management("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb/Converted_Sales","POSTAL_CODE","TEXT")
arcpy.AddField_management("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb/Converted_Sales","NET_SALES_PC","FLOAT")
arcpy.AddField_management("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb/Converted_Sales","Dist_Sales_PC","FLOAT")
arcpy.AddField_management("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb/Converted_Sales","Percent_of_Sales","FLOAT")
INRows = arcpy.InsertCursor("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb/Converted_Sales")

Year1Dict = {}
Year2Dict = {}
Year3Dict = {}
Year4Dict = {}
Year5Dict = {}
PS_PC_Table = arcpy.SearchCursor(PC_SALES_Table)
for totals in PS_PC_Table:
    Year = totals.getValue("FISCAL_YEAR")
    Store_Code = int(totals.getValue("STORE_NUMBER"))
    Postal_Code = totals.getValue("FSA")
    Sales = totals.getValue("SumOfRETAIL_TOTAL_SALES_AMT")
    if Year == 2012:
        if Year1Dict.has_key(Store_Code):
            Year1Dict[Store_Code] = float(Year1Dict[Store_Code]) + float(Sales)
        else:
            Year1Dict[Store_Code] = float(Sales)
    if Year == 2013:
        if Year2Dict.has_key(Store_Code):
            Year2Dict[Store_Code] = float(Year2Dict[Store_Code]) + float(Sales)
        else:
            Year2Dict[Store_Code] = float(Sales)

```

```

    if Year == 2014:
        if Year3Dict.has_key(Store_Code):
            Year3Dict[Store_Code] = float(Year3Dict[Store_Code]) + float(Sales)
        else:
            Year3Dict[Store_Code] = float(Sales)
    if Year == 2015:
        if Year4Dict.has_key(Store_Code):
            Year4Dict[Store_Code] = float(Year4Dict[Store_Code]) + float(Sales)
        else:
            Year4Dict[Store_Code] = float(Sales)
    if Year == 2016:
        if Year5Dict.has_key(Store_Code):
            Year5Dict[Store_Code] = float(Year5Dict[Store_Code]) + float(Sales)
        else:
            Year5Dict[Store_Code] = float(Sales)
    else:
        pass
arcpy.AddMessage(Year1Dict)
arcpy.AddMessage("FYB - Years Compiled into Dictionaries")
TotalSalesDict = {}
TOTAL_SALES_Tbl = arcpy.SearchCursor(TOTAL_SALES_Table)
for totSales in TOTAL_SALES_Tbl:
    Store_Code = int(totSales.getValue("STR_N"))
    Y2016 = totSales.getValue("2016")
    Y2015 = totSales.getValue("2015")
    Y2014 = totSales.getValue("2014")
    Y2013 = totSales.getValue("2013")
    Y2012 = totSales.getValue("2012")
    TotalSalesDict[Store_Code]={"2012":Y2012,"2013":Y2013,"2014":Y2014,"2015":Y2015,"
2016":Y2016}
arcpy.AddMessage(TotalSalesDict)
y = 0
PS_PC_Table2 = arcpy.SearchCursor(PC_SALES_Table)
for sls in PS_PC_Table2:
    #arcpy.AddMessage("Fguncik Tynoaub")
    Yr = str(sls.getValue("FISCAL_YEAR"))
    Store = int(sls.getValue("STORE_NUMBER"))
    PostCo = sls.getValue("FSA")
    Saless = float(sls.getValue("SumOfRETAIL_TOTAL_SALES_AMT"))
    #arcpy.AddMessage("2")
    if Yr == "2012":
        Dis_Sls = Saless * float(TotalSalesDict[Store][Yr])/float(Year1Dict[Store])
    if Yr == "2013":
        Dis_Sls = Saless * float(TotalSalesDict[Store][Yr])/float(Year2Dict[Store])

```

```

if Yr == "2014":
    Dis_Sls = Saless * float(TotalSalesDict[Store][Yr])/float(Year3Dict[Store])
if Yr == "2015":
    Dis_Sls = Saless * float(TotalSalesDict[Store][Yr])/float(Year4Dict[Store])
if Yr == "2016":
    Dis_Sls = Saless * float(TotalSalesDict[Store][Yr])/float(Year5Dict[Store])
#arcpy.AddMessage(TotalSalesDict[Store][Yr])
Per_Dis_Sls = Dis_Sls / TotalSalesDict[Store][Yr] * 100
Updaterow = INRows.newRow()
Updaterow.FISCAL_YEAR = Yr
Updaterow.STORE_NO = Store
Updaterow.POSTAL_CODE = PostCo
Updaterow.NET_SALES_PC = Saless
Updaterow.Dist_Sales_PC = Dis_Sls
Updaterow.Percent_of_Sales = Per_Dis_Sls
if y%10000 == 0:
    arcpy.AddMessage(y)
y+=1
INRows.insertRow(Updaterow)

```



**Figure A7.** Dominant postal code sales data compiler

#Developer: Scotty Norman

#Purpose: Dominant Sales Layer

#Usage: Major Research Paper

```
import arcpy
from arcpy import env
import os
import time
import traceback
from datetime import datetime
import sys
arcpy.AddMessage("Script Started- Dom Sales")
arcpy.AddMessage("Modules have been imported")

listofYears = ["2012", "2013", "2014", "2015", "2016"]

i = 0
for years in listofYears:
    if arcpy.Exists("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb/Dom_Sales_FSA" + str(years)):
        arcpy.AddMessage("File Found, Deleting Table Now")
        arcpy.Delete_management("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb/Dom_Sales_FSA" + str(years))
    else:
        pass
    if arcpy.Exists("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb/Dom_Sales_FSA" + str(years)):
        arcpy.AddMessage("Delete didn't work")
    else:
        arcpy.AddMessage("Delete complete: " + (years))

Sales = {}
SearchCons = arcpy.SearchCursor("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb/Converted_Sales_FSA")

for row in SearchCons:
    FiscalYear = row.getValue("FISCAL_YEAR")
    StoreNum = row.getValue("STORE_NO")
    PostalCode = row.getValue("POSTAL_CODE")
    NetSales = row.getValue("NET_SALES_PC")
    DistSales = row.getValue("Dist_Sales_PC")
    PercSales = row.getValue("Percent_of_Sales")
    UniqueKey = str(FiscalYear) + " - " + str(PostalCode)
```

```

if Sales.has_key(UniqueKey):
    if Sales[UniqueKey][0] < PercSales:
        Sales[UniqueKey] = [NetSales,StoreNum,DistSales,PercSales]
    else:
        Sales[UniqueKey] = [NetSales,StoreNum,DistSales,PercSales]
if i % 100000 == 0:
    arcpy.AddMessage("So far " + str(i) + " have been read")
    i += 1
for year in listofYears:
    saving = "Dom_Sales_FSA" + str(year)
    save = "M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb/Dom_Sales_FSA" + str(year)
    arcpy.CreateTable_management("M:/gis/Scott Norman/MRP/GDBs/04 -
Apr/PS_5Yr_Sales_Transformed.gdb/",saving)
    arcpy.AddField_management(save,"FISCAL_YEAR","LONG")
    arcpy.AddField_management(save,"STORE_NO","LONG")
    arcpy.AddField_management(save,"POSTAL_CODE","TEXT")
    arcpy.AddField_management(save,"NET_SALES_PC","DOUBLE")
    arcpy.AddField_management(save,"DIST_SALES_PC","DOUBLE")
    arcpy.AddField_management(save,"PERC_OF_SALES","DOUBLE")
    INSRows = arcpy.InsertCursor(save)
    i = 0
    for PCs in Sales:
        rowinsert = INSRows.newRow()
        infoSplit = PCs.split(" - ")
        if infoSplit[0]== year:
            rowinsert.FISCAL_YEAR = infoSplit[0]
            rowinsert.STORE_NO = Sales[PCs][1]
            rowinsert.POSTAL_CODE = infoSplit[1]
            rowinsert.NET_SALES_PC = Sales[PCs][0]
            rowinsert.DIST_SALES_PC = Sales[PCs][2]
            rowinsert.PERC_OF_SALES = Sales[PCs][3]
            INSRows.insertRow(rowinsert)
        else:
            pass
    if i % 10000 == 0:
        arcpy.AddMessage("So far " + str(i) + " have been read")
        i += 1
arcpy.AddMessage("K Tru")

```

**Figure A8.** Spatial join dominant sales to FSA shapefile

#Developer: Scotty Norman

#Purpose: Joining Dominant Sales Layer to FSA Shapefile

#Usage: Major Research Paper

```
import arcpy
from arcpy import env
import os
import time
import traceback
from datetime import datetime
import sys
arcpy.AddMessage("Script Started- Spatial Join")
arcpy.AddMessage("Modules have been imported")
for y in xrange(2012,2017):
    if arcpy.Exists("M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\02_Dom_Sales_" + str(y) +
"\\gfsa000a11a_e.shp"):
        arcpy.AddMessage("Dom Sales " + str(y) + " File Found, Deleting Table Now")
        arcpy.Delete_management("M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\02_Dom_Sales_" + str(y) +
"\\gfsa000a11a_e.shp")
    else:
        pass
Dom_Sales = "M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\02_Dom_Sales_"
for x in xrange(2012,2017):
    arcpy.AddMessage("Starting to create Dominant Sales: " + str(x))
    gfsa000a11a_e_shp = "M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\01_Boundaries\\gfsa000a11a_e\\gfsa000a11a_e.shp"
    gfsa000a11a_e_shp__2_ = gfsa000a11a_e_shp
    Dom_Sales_FSA2012 = "M:\\Gis\\Scott Norman\\MRP\\GDBs\\04 -
Apr\\PS_5Yr_Sales_Transformed.gdb\\Dom_Sales_FSA" + str(x)
    v02_Dom_Sales = Dom_Sales + str(x)
    v02_Dom_Sales__2_ = v02_Dom_Sales
    newFile = arcpy.JoinField_management(gfsa000a11a_e_shp, "CFSAUID",
Dom_Sales_FSA2012, "POSTAL_CODE")
    newFile
    arcpy.FeatureClassToShapefile_conversion("M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\01_Boundaries\\gfsa000a11a_e\\gfsa000a11a_e.shp",
v02_Dom_Sales)
    arcpy.DeleteField_management("M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\01_Boundaries\\gfsa000a11a_e\\gfsa000a11a_e.shp",["FISCAL_YE
A","STORE_NO","POSTAL_COD","NET_SALES_","DIST_SALES","PERC_OF_SA"])
    del newFile
```

**Figure A9. Trade area creation**

#Developer: Scotty Norman  
#Purpose: Trade Area Creation  
#Usage: Major Research Paper

```
import arcpy
from arcpy import env
import os
import time
import traceback
from datetime import datetime
import sys
arcpy.AddMessage("Script Started- Trade Areas")
arcpy.AddMessage("Modules have been imported")

for year in xrange(2012,2017):
    if arcpy.Exists("M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\02_Dom_Sales_" + str(year) +
"\\Dom_Sales_Start_Point.shp"):
        arcpy.AddMessage("File Found, Deleting Table Now")
        arcpy.Delete_management("M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\02_Dom_Sales_" + str(year) +
"\\Dom_Sales_Start_Point.shp")
    else:
        pass

    if arcpy.Exists("M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\02_Dom_Sales_" + str(year) +
"\\Trade_Areas.shp"):
        arcpy.AddMessage("File Found, Deleting Table Now")
        arcpy.Delete_management("M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\02_Dom_Sales_" + str(year) +
"\\Trade_Areas.shp")
    else:
        pass

    arcpy.AddMessage("Identifying Files")
    PS_shp = "M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\04_Lambert_Files\\Part_Source_Locations_NEW.shp"
    Dom_Sales = "M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\02_Dom_Sales_" + str(year) +
"\\gfsa000a11a_e.shp"
```

```

        PS_Shp_Dom = "M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\02_Dom_Sales_" + str(year) +
"\\Dom_Sales_With_Stores.shp"
        Assign_Start = "M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\02_Dom_Sales_" + str(year) +
"\\Dom_Sales_Start_Point.shp"
        memory = "M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\00_Working_Files\\"
        Pclist = "PC_list.shp"
        memPclist = "M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\00_Working_Files\\PC_list.shp"
        memPclist2 = "M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\00_Working_Files\\PC_list_Diss.shp"
        dissolvedPCs = "M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\00_Working_Files\\PC_dissolved"

        out = "M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\02_Dom_Sales_" + str(year) + "\\
        TAFile = "Trade_Areas.shp"
        ListOfStore = []
        count3 = 0
        arcpy.SpatialJoin_analysis(Dom_Sales,PS_shp,Assign_Start,"JOIN_ONE_TO_ONE","KEEP
_COMMON","", "INTERSECT","", "")

        Search_SJ = arcpy.SearchCursor(Assign_Start)
        arcpy.AddMessage("Starting Trade Area Generation")
        PCListing = {}
        for store in Search_SJ:
            count = 1
            if arcpy.Exists("M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\00_Working_Files\\PC_list.shp"):
                arcpy.AddMessage("File Found, Deleting Table Now")
                arcpy.Delete_management("M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\00_Working_Files\\PC_list.shp")
            else:
                pass
            if arcpy.Exists("M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\00_Working_Files\\PC_list_Diss.shp"):
                arcpy.AddMessage("File Found, Deleting Table Now")
                arcpy.Delete_management("M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\00_Working_Files\\PC_list_Diss.shp")
            else:
                pass
            PolyShape = store.getValue("Shape")

```

```

    arcpy.CreateFeatureclass_management(memory,PCList,"POLYGON",Assign_Start,"","","
M:\\Gis\\Scott Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\02_Dom_Sales_" + str(year)
+ "\\Dom_Sales_Start_Point.prj")
    arcpy.Append_management(PolyShape,memPCList,"NO_TEST","","")
    arcpy.Dissolve_management(memPCList,memPCList2)
    Store_Number = int(store.getValue("Store_Numb"))
    arcpy.AddMessage("Generating Trade Area for: " + str(Store_Number))
    ListOfStore.append(Store_Number)
    while count > 0:
        arcpy.MakeFeatureLayer_management(Dom_Sales, "Domin_Sales")
        arcpy.MakeFeatureLayer_management(memPCList2, "PCList2mem")
        selectFSAs =
arcpy.SelectLayerByLocation_management("Domin_Sales","WITHIN_A_DISTANCE","PCList2me
m","5 Meters","NEW_SELECTION")

        count = 0
        SC_selectFSAs = arcpy.SearchCursor(selectFSAs)
        area = False
        for build in SC_selectFSAs:
            FSAtest      = build.getValue("CFSAUID")
            shapes       = build.getValue("Shape")
            StoreNo      = int(build.getValue("STORE_NO"))
            NetSls       = float(build.getValue("NET_SALES_"))
            DistSLS      = float(build.getValue("DIST_SALES"))
            PEROSLS      = float(build.getValue("PERC_OF_SA"))
            if StoreNo    == Store_Number and FSAtest not in PCListing:

arcpy.Append_management(shapes,memPCList,"NO_TEST","","")
                count += 1
                PCListing[FSAtest] = [StoreNo]
            else:
                pass
        if arcpy.Exists("M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\00_Working_Files\\PC_list_Diss.shp"):
            arcpy.AddMessage("File Found, Deleting Table Now")
            arcpy.Delete_management("M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\00_Working_Files\\PC_list_Diss.shp")
        else:
            pass
        arcpy.Dissolve_management(memPCList,memPCList2)
        #count2 += 1
        arcpy.Delete_management("Domin_Sales")
        arcpy.Delete_management("PCList2mem")

```

```

        if count3 > 0:
            arcpy.Append_management(memPClist2,out+TAFile,"NO_TEST","", "")
        else:

            arcpy.CreateFeatureclass_management(out,TAFile,"POLYGON","M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\02_Dom_Sales_" + str(year) +
"\\Dom_Sales_Start_Point.shp","", "", "M:\\Gis\\Scott
Norman\\MRP\\Shapefiles\\03_Trade_Area_Files\\02_Dom_Sales_" + str(year) +
"\\Dom_Sales_Start_Point.prj")
            arcpy.Append_management(memPClist2,out+TAFile,"NO_TEST","", "")
            count3 += 1
        cursor = arcpy.UpdateCursor(out+TAFile)
        countrow = 0
        for row in cursor:
            if row.STORE_NO == 0:
                row.STORE_NO = int(ListOfStore[countrow])
                cursor.updateRow(row)
                countrow += 1
arcpy.AddMessage("Completed Trade Areas")

```

## Appendix B

Variables	R <sup>2</sup>	Interpretation
Domestic Car 1981 - 1985	0.737167	Strong
Domestic Car 1986 - 1990	0.719357	Strong
Domestic Car 1991 - 1995	0.739766	Strong
Domestic Car 1996 - 2000	0.73375	Strong
Domestic Car 2001 - 2005	0.736033	Strong
Domestic Car 2006 - 2010	0.718609	Strong
Domestic Car 2011 - 2015	0.745898	Strong
Domestic Truck 1981 - 1985	0.709501	Strong
Domestic Truck 1986 - 1990	0.713173	Strong
Domestic Truck 1991 - 1995	0.731328	Strong
Domestic Truck 1996 - 2000	0.71037	Strong
Domestic Truck 2001 - 2005	0.699851	Strong
Domestic Truck 2006 - 2010	0.72444	Strong
Domestic Truck 2011 - 2015	0.737258	Strong
Import Car 1981 - 1985	0.713252	Strong
Import Car 1986 - 1990	0.632352	Strong
Import Car 1991 - 1995	0.740318	Strong
Import Car 1996 - 2000	0.689315	Strong
Import Car 2001 - 2005	0.736721	Strong
Import Car 2006 - 2010	0.740982	Strong
Import Car 2011 - 2015	0.7524	Strong
Import Truck 1981 - 1985	0.729523	Strong
Import Truck 1986 - 1990	0.763417	Strong
Import Truck 1991 - 1995	0.721254	Strong
Import Truck 1996 - 2000	0.707294	Strong
Import Truck 2001 - 2005	0.735687	Strong
Import Truck 2006 - 2010	0.743827	Strong
Import Truck 2011 - 2015	0.745328	Strong

Table of TVIO Variables for Medicine Hat, Ablerta



Variables	R <sup>2</sup>	Interpretation
Domestic Car 1981 - 1985	0.174046	Very Weak
Domestic Car 1986 - 1990	0.161679	Very Weak
Domestic Car 1991 - 1995	0.445878	Weak
Domestic Car 1996 - 2000	0.612733	Moderate
Domestic Car 2001 - 2005	0.588434	Moderate
Domestic Car 2006 - 2010	0.545748	Moderate
Domestic Car 2011 - 2015	0.532466	Moderate
Domestic Truck 1981 - 1985	0.209258	Weak
Domestic Truck 1986 - 1990	0.163762	Weak
Domestic Truck 1991 - 1995	0.540732	Moderate
Domestic Truck 1996 - 2000	0.607267	Moderate
Domestic Truck 2001 - 2005	0.598732	Moderate
Domestic Truck 2006 - 2010	0.602727	Moderate
Domestic Truck 2011 - 2015	0.573374	Moderate
Import Car 1981 - 1985	0.161256	Very Weak
Import Car 1986 - 1990	0.586901	Moderate
Import Car 1991 - 1995	0.405763	Weak
Import Car 1996 - 2000	0.560818	Moderate
Import Car 2001 - 2005	0.580805	Moderate
Import Car 2006 - 2010	0.508052	Moderate
Import Car 2011 - 2015	0.439933	Weak
Import Truck 1981 - 1985	0.163268	Very Weak
Import Truck 1986 - 1990	0.211551	Very Weak
Import Truck 1991 - 1995	0.181288	Very Weak
Import Truck 1996 - 2000	0.569014	Moderate
Import Truck 2001 - 2005	0.487938	Moderate
Import Truck 2006 - 2010	0.30291	Very Weak
Import Truck 2011 - 2015	0.313036	Weak

Table of TVIO Variables for Markham, Ontario