USING CREDIT SCORES TO PREDICT AND ANALYZE AREAS OF GENTRIFICATION AND URBAN DEVELOPMENT IN TORONTO, CANADA - A GEO-SPATIAL APPROACH

By
Nicholas Pulsone
Honours B.A., Geographic Analysis, Ryerson University, 2018

A Major Research Paper
presented to Ryerson University
in partial fulfillment of the
requirements for the degree of
Master of Spatial Analysis
in the program of
Spatial Analysis

Toronto, Ontario, Canada, 2019

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 revisions.

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.

Using Credit Scores to Predict and Analyze Areas of Gentrification and Urban Development in Toronto, Canada - A Geo-Spatial Approach

Nicholas Pulsone, 2019

Master of Spatial Analysis
in
Spatial Analysis

Ryerson University

Abstract

This study examines the use of financial well-being indicators such as credit scores to identify gentrification. This study is a response to the redevelopment of neighbourhoods in the City of Toronto through gentrification. This study also explores both theoretical and analytical frameworks outlined in literature to identify correlations between financial wellbeing indicators and gentrification. Comparing the observations in this study to areas experience gentrification such as Regent Park revealed large implications that gentrification is largely associated with financial wellbeing. The study also found that the average credit scores in the City of Toronto seem to be increasing. The analysis determined that the credit score changes reflected the development in the Regent Park development zone.

Key words: Gentrification, credit scores, spatial analysis, urban development

Acknowledgements

I would like to thank all the people that have helped or inspired me to conduct this study and write this MRP. I would not have been motivated nor able to complete this study without the following people.

First, I would like to thank my academic supervisor Dr. Brian Ceh for his help, ideas and inspiration. I found his comments, insights and suggestions very helpful. Additionally, his research and understanding about urban development sparked my interest and kept my interest in this study.

Next I would like to thank my friends and colleagues for the support they have given me. The academic and non-academic experiences I have shared with my classmates have been very important to me and I thank them for making my time at Ryerson University very memorable.

Finally, I would like to thank my family for the care and support they have given me through the years. Without their support, none of this would have been possible and I thank them for all that they have done for me.

Table of Contents

List of Tables

List of Figures

1.0 Introduction Statement

In 2005, the City of Toronto proposed a plan to demolish and redevelop low-income areas throughout the next decade. The 1-billion-dollar revitalization project approved by the City of Toronto demolished numerous buildings and households (Johnson & Schippling, 2009). The redevelopment project falls under a premise, in which, a new type of "*mixed-housing*" will be created to better accommodate the income of residents in the area. The premise of such redevelopment projects, however, also creates a displacement in the population that must, therefore, relocate according to the proposed development. The concentrated poverty of these low-income areas throughout the 1990s to early 2000s led to **stigmatization** but also benefitted from a planning perspective. In 2013, the University of Toronto published an article which identified the perception of low-income areas in the City of Toronto as a *ghetto's* (Kelly, 2013). The term *ghetto* refers to an area in which the localized population is subject to economic marginalization (Sung, 2015; August, 2014). The transition of such economically marginalized areas into middle or upper-class housing shows urban development, and in many cases, identifies gentrification.

Credit scores can identify such areas through financial well-being indicators, as a comprehensive quantitative rating of each person's financial status based on bank transactions, loans and other financial payments (Dean & Nicholas, 2018). However, to examine trends in gentrification and development, the following research questions will be addressed:

1. **Do credit scores of similar values cluster together spatially?**
2. **Do socio-economic indicators correlate with credit scores?**
3. **Can financial well-being indicators such as credit scores accurately identify relationships and areas of gentrification or urban development?**

These three research questions will be the basis of the analyses conducted within the research paper. Specifically, the goal of the research paper is to examine these research questions within the context of the **City of Toronto**. The assumption is that each research question will be able to identify spatial or a-spatial relationships. One should expect credit scores of similar values to cluster together. As income and financial well-being are largely correlated together, it is also likely that a correlation exists between socio-economic indicators such as income and credit scores. Therefore, financial well-being indicators such as credit scores should be able to identify relationships and areas of gentrification.

2.0 Preliminary context and relevant literature

2.1 Gentrification and New Development

In the 1960s, Ruth Glass coined the term *gentrification* which referred to changes in social structure and housing markets within inner-city districts or neighbourhoods. However, Martin & Beck (2018) argue that this definition of gentrification is biased to inner-city geographies. The focus of urban gentrification in media and academia is largely due to the turnover of *culture and economy* within inner-city geographies (Phillips, 2004). Phillips (2004) states that such comparisons between spatial trends in gentrification across rural and urban regions should be avoided due to the different social and economic circumstances. Ideally, spatial trends of gentrification should be examined between inner-city and comparable geographies such as districts or neighbourhoods (Phillips, 2004; Glass 1964).

Although gentrification research can differ between levels of geography, such transformations are often fueled by large-scale economic developments and government policy. (Slater, 2006; Lees et al., 2007; Hackworth and Smith, 2001). Traditionally, gentrification research is concerned with the processes in which working-class residential neighbourhoods become comprised of a middle-class or a higher income demographic.

Various studies have examined housing patterns across various geographies, which examine the financial well-being of neighbourhoods. Studies such as Atuesta & Hewings (2019) have determined that neighbourhoods with low or seemingly *unstable* financial well-being may cause displacement in neighbourhoods. Generally, such assumptions would indicate that neighbourhoods with stable or excellent financial well-being are not prone to such displacement pressures.

Arguably, research can often analyze the effect of gentrification on indicators such as housing and rent prices, income and racial turnover, but rarely identify the consequences of the processes. Specifically, Knapp & Dean (2018) has indicated that single direct indicators such as income or housing prices are relevant with the displacement of the low-income population; however holistic indicators such as credit scores are more appropriate to identify if displacement due to financial or economic reasons may occur. It is also important to indicate that displacement is not entirely financially dependent. As various studies have identified gentrification through racial turnover; the effects and causes of gentrification through socio-economic influences should not be ignored (Phillips, 2004; Glass 1964).

2.2 What is a Credit Score?

Fair, Isaac and Company developed a standardized index in 1989 which would evaluate consumer credit information in the United States (Carrns, 2012). The consumer credit information would represent the creditworthiness, or, lack thereof, when applying for loans, mortgages and other financial investments. The **FICO** model was initially developed to use the credit information from the three major national credit bureaus: **Equifax**, **TransUnion** and **Experian**. These *scores* would vary between industry with specific scores for mortgages and credit cards, along with a total financial credit score.

In 2008, the United States experienced a nation-wide financial crisis with a vast majority of the population experiencing housing delinquencies and foreclosures. Many academics deem the cause of the financial crisis to be mainly due to the fraud committed by the Credit rating agencies which did not correctly evaluate loans and *creditworthiness* (Feldkircher, 2014; Dwyer, Tabak & Vilmunen, 2012; Soros, 2008). However, as credit scores are a composite score of credit behaviour, banks and other financial institutions still use credit scores such as FICO, or other modelled equivalents, to measure the *risk* associated with each credit rating (Arya, Eckel & Wichman, 2013).

Additionally, common industry practices follow the assumption that FICO or credit scoring models across each of the three major national credit bureaus follow different scoring methods. Specifically, such "black-box" information is scarcely found in literature, however, it is common practice for specific banks, lenders or borrowers to exclusively partner with a single credit bureau.

Although many factors can determine creditworthiness or financial well-being, credit is inherently linked to a person's ability or behavioural tendency to manage debt. More specifically, the borrowers or person's willingness to pay and the ability to pay such debt heavily influence a person's financial well-being. Additionally, credit scores are more concerned with the present and future probability of a borrower failing to pay their debts and ultimately *defaulting*.

2.3 Determining Financial Wellbeing

Traditional gentrification research and literature has identified links with gentrification, affordability and financial wellbeing (Lees and Demeritt, 1998; Gibbs and Kreuger, 2007, Hagerman, 2007; Jonas and While, 2007; Kear, 2007; Krueger, 2007; Krueger and Savage, 2007). However, research into the determinants of financial wellbeing and the correlations with gentrification has been scarce. More specifically, existing theoretical and analytical approaches are often concerned with identifying economic relationships in settlement such as **income, rent affordability and dwelling value** to examine correlations with credit scores.

Similarly, literature has identified financial well-being through indicators which can determine economic management on a long-term basis. For example, variables such as income are often connected to research indicating financial well-being. Such research often identifies income as a large indicator of financial well-being, as the ability to pay off various debt is often linked with the inherent income and savings someone may have. However, a recent study by Citizens Financial Group (2016), has indicated that college graduates under the age of 35 spend approximately 20% of their annual salary on student loan payments, which, therefore, limits spending and consumption. Furthermore, “having limited financial reserves can cause great difficulty when unexpected financial emergencies arise and may prompt individuals to suffer from financial hardship.” (Brüggen, Hogreve, Holmlund, Kabadayi & Löfgren, p.229, 2017)

However, financial wellbeing cannot be measured through such a binary approach. Joo and Grable (2004) conducted a study to determine factors that influence financial wellbeing. The study determined that financial activity and consumer behaviour records lead to a direct correlation with financial wellbeing. Such a relationship indicates that consumers with poor financial wellbeing are likely to be vulnerable to economic risk.

Other studies tested correlations with financial wellbeing and demographic variables (Cude, 2010; Dvorak & Hanley, 2010; Joo & Grable, 2004). However, such studies were not able to prove that demographic variables are more representative of financial wellbeing. Similarly, credit scores are directly linked to consumer behaviour habits and financial activity and are more representative of financial wellbeing than demographic or socio-economic variables (Westover, 2013).

2.4 Comparing Urban development indicators

Sustainable development practices are commonly defined as the systematic approach of integrating a city subsystem to avoid decreasing levels of financial wellbeing (United Nations, 1987; Tran, 2016). Tran (2016) defines sustainable development through the changing characteristics of inner-city geographies which improve social, economic and financial wellbeing. Various studies have also used linear regression models to identify if socio-economic or demographic variables are related to urban development (Sheng, Han, & Zhou, 2017; Shazmeen, Mirza Mustafa Ali Baig, & Pawar, 2013). However, urban development is heavily influenced by government policy and zoning by-law changes (MacLaran, 2003). Simply defining urban development as renewal should be avoided as this ignores changes in social structure and policy changes.

As urban development and urban renewal is often linked to the gentrification and resettlement of the upper-middle class into low-income settlements some studies show that various empirical method for analyzing urban development and renewal can successfully employ both quantitative and qualitative approaches (Sheng, Han, & Zhou, 2017; Shazmeen, Mirza Mustafa Ali Baig, & Pawar, 2013). However, it is important to understand that urban renewal can be defined through two different concepts. First, such definitions can be explored in the sense of housing and socioeconomic influence. More specifically, literature often refers to urban renewal as the literal transition of housing through large-scale renovations or development, leading to an increase in housing or rent prices (Martin and Beck, 2016; Helms, 2003; Rosenthal and Ross, 2015). However, such definitions ignore political and socio-economic changes and should be avoided when comparing neighbourhoods within a social or demographic context.

2.5 Comparing Credit Bureaus

There is common thought within the industry that the major credit bureaus in Canada all provide different credit reports and scores within the Canadian Market. However, due to the confidentiality and nature of proprietary credit score data, this creates an inconsistency amongst research and data. The first inconsistency comes from the date and time at which these scores are calculated. It can be argued that credit bureaus calculate scores on different dates, therefore resulting in different scores as the methods for classification change with time. Secondly, credit scores for a person can vary among credit bureaus. For example, the FICO scoring model may be used by each of the 3 major credit bureaus, but some lenders may not report to each of the major credit bureaus. Therefore, the credit bureau may be missing information that could change credit scores. Additionally, each credit bureau uses different ranges to define poor credit, fair credit and good credit. Equifax provides specific ranges for such analysis however; such ranges are rarely used in academic research. For example, Ding, Hwang & Divringi (2016) used Equifax credit scores, however, did not use the Equifax provided ranges. Using non-standardized ranges aggregates populations with significantly different financial wellbeing into the same group. Such aggregations should be avoided as each credit range or *credit band* can help define each individual's financial situation and changing the range of credit scores will ultimately affect the resulting comparative analysis.

2.6 Using Credit Scores as a Geo-Spatial Predictor

A study conducted by Israel et al. (2014), measured credit score values and associated cardiovascular risk. The results indicated a link between high credit risk scores and lower cardiovascular disease. Although used as a predicative entity, the credit risk data was not examined geographically to determine if a trend exists spatially.

Ding, Hwang & Divringi (2016), used credit score information obtained from Equifax, along with socioeconomic indicators representing gentrification such as low-income households, rent prices and dwelling values to visually identify areas of gentrification. This study was coherent with various literature identifying the socio-economic properties for areas of gentrification (Pattillo, 2008; Freeman, 2005). However, other studies suggest that empirical measures of gentrification should identify the changing distribution of race and ethnicity within a neighbourhood (Lees, 2008; Charles, 2000). Generally, empirical measurements of gentrification follow these previously mentioned theoretical frameworks. Ding, Hwang & Divringi (2016), did not indicate if specific areas with various aggregate credit scores were likely to experience gentrification. Although the study did mention the potential use of regression for such a comparison, no statistical analysis of the variables was carried out.

It is important to separate simple neighbourhood change from gentrifying or non-gentrifying areas. Early empirical measures of gentrification examine changing house prices and socio-economic transition of the neighbourhood. Such changes are argued to displace existing inhabitants and therefore stimulate the arrival of a new, wealthier population into the neighbourhood. Reades, De Souza, & Hubbard (2019), argue that such definitions are too narrow to be defined as gentrification. Instead, such linear comparisons are more appropriately labelled as *neighbourhood change*. In the USA and UK, many studies have indicated the appropriate use of spatial data through geodemographic analysis to measure specific changes of socio-economic status and income. Although such comprehensive analysis can group and classify such changes according to the characteristics of the neighbourhood, geo-spatial attributes which measure increasing tax or potential economic stress should be explored (Reades, De Souza & Hubbard, 2019; Galster, 2001).

2.7 Empirical measures of gentrification

Causes and consequences of gentrification are different across communities, neighbourhoods and cities and must be assessed using a holistic approach of theoretical and empirical analysis (Atkinson & Bridge, 2005). Specifically, cities often contain large economic disparities between neighbourhoods due to economic shifts and urban development policies. To accommodate such disparities of gentrified and non-gentrified areas, Atkinson & Bridge (2005), argue that the degree in which gentrification is observed can often be identified with socio-economic indicators. Such changes are also identifiable through populations with increasing or high credit scores.

2.8 Predicting areas of new development and gentrification

Predicting gentrification or new development is quite complex as gentrification and new development is heavily influenced by social and political changes. Walker (2018), identifies gentrification in a study through social networking. However, additional quantitative analysis can be conducted into the changes of socio-economic indicators (Ding, Hwang & Divringi, 2016; Freeman, 2005). Although gentrification can be linked to social and political influence, Martin & Beck (2018), argue that gentrification is easily identifiable due to the financial well-being of the area that follows gentrification. Specifically, the change of property tax and policy in many gentrifying areas within the United States *forced* various families out of neighbourhoods. These involuntary moves were directly correlated with the increase in liabilities, debt and property taxes (Martin & Beck, 2018). This increase and change in taxation can also be reflected through credit scores, which can indicate the financial well-being of the population throughout the geographic region.

2.8.1 Common statistical practices

Many statistical analyses that aim to predict gentrification; to some degree, often use robust methods, along with appropriate variables specified in literature or through the conceptual framework and geographic context. Specifically, the most common techniques include regression modelling to predict neighbourhood change and gentrification at a granular level.

Other statistical practices include machine learning and complex algorithms to determine the most optimal model when using regression. However, such analysis isolates the geographic context of the neighbourhood and ignores theoretical frameworks from literature. It can be shown, however, that such analysis is more suited to appropriately dissect correlations with variables that may indicate neighbourhood change.

Various studies have also used PCA / FA when conducting correlation and regression tests for gentrification. However, it should be noted that such analysis is not recommended at a large scale; and therefore, requires granular data. The implication is that comparing areas of similar geographies will show specific differences and changes after the PCA. However, such analysis requires specific variables to be chosen. Unlike the approach of using various algorithms to determine the most optimal variables, careful consideration is needed to determine which indicators will be used. More specifically, such variables can indicate if the neighbourhood is simply undergoing minor change, or potential gentrification according to the definitions provided in literature.

2.9 Literature Review Overview

As shown through the supporting literature, currently, no literature exists that have proven or disproven the use of credit scores for predicting gentrification. However, a consensus exists that indicates such a variable should be tested. Although studies have previously tested the relationship with similar variables that measure financial well-being, as a comparison with gentrification, credit scores are not frequently used due to confidentiality and the nature of the data.

However, as credit scores are more comprehensive than traditional financial well-being indicators such as income and housing value; various literature recommends using credit scores to more appropriately indicate wealth management and financial wellbeing. As indicated in literature, this is shown through,

> properties located in neighbourhoods with different levels of income or housing prices at baseline. For instance, neighbourhoods that, in 1990, had already begun the gentrification process, could have greater housing prices and higher income than neighborhoods that began experiencing gentrification during the 1990s. The effect of gentrification on housing prices after 2000 could be different for these two types of neighborhoods. (Atuesta & Hewings, 2019, p.36)

Such changes are often neglected in most empirical measures of gentrification, which, therefore, fail to measure the changes between simple neighbourhood changes and gentrification. However, it should also be understood that many empirical tests of gentrification often avoid commenting on the policy changes that may influence such changes. Such changes are difficult to measure empirically and are often not included in empirical or statistical gentrification analysis. Moreover, a gap in literature also exists as most of the research or empirical analyses on gentrification using predictor variables do not analyze spatial autocorrelation. " One of the main reasons why **spatial** auto-**correlation is important** is because statistics relies on observations being independent from one another. If **autocorrelation** exists in a map, then this violates the fact that observations are independent from one another." (GIS & Geography, 2018)

As credit scores can appropriately indicate the financial well-being of a neighbourhood, such measurements of economic stress are more indicative of the creditworthiness of a population. Many statistical practices regarding gentrification have avoided using credit scores due to the confidential nature of the data and the inability to obtain such information. Despite these issues, geospatial analyses using credit scores are not taboo. Studies such as Ding, Hwang & Divringi (2016), indicate the importance of financial well-being concepts such as credit scores when measuring gentrification. However, this paper will look to examine the claims that such financial well-being indicators are, indeed, clear indicators of gentrification.

3.0 Data & Methodology

3.1 Data

Credit score data is collected at a very granular level. Although such information may be able to provide statistically significant results, such information is confidential. Therefore, the credit score data was aggregated to the Dissemination Area level to maintain confidentiality. The aggregated credit scores will ultimately show the average (mean) credit score in each of the **3702 DAs** across the City of Toronto. This, therefore, introduces two assumptions into the data, the Ecological fallacy and the Modifiable areal unit problem.

3.1.1 Problems with aggregation : Ecological fallacy and MAUP assumptions

The first assumption is known as the ecological fallacy. The ecological fallacy is described as a phenomenon where the entire population within a geographic area, must contain the same characteristics. However, this is not the case and ultimately is an incorrect assumption as population characteristics can differ heavily in both large and small scales of geography.

The second assumption is the Modifiable Areal Unit Problem (MAUP). MAUP pertains to the sensitivity and bias of statistical results to differing levels of geography. As geographic or demographic analysis is often conducted using modifiable areal units, MAUP contains two major issues of concern. Specifically, these include effects of scale and the effects of zoning.

Effects of scale in MAUP are related to the idea that using geography and geographic units are subject to be influenced by the scale of the areal. For example, when looking at average household income; different values will be shown across the region, depending on the units of geography that are used. Generally, complex research analysis should avoid analysis of areas using different or numerous geographic scale(s). If necessary, analysis conducted using very granular level of data can be aggregated, however, it is not recommended to disaggregate the data (Leventhal, 2016).

The effects of zoning in MAUP are associated with the understanding of the area and the choice of a geographic region. If a large number of smaller units are aggregated; manipulating the geographic region and data should be avoided. An example of this is “manipulating an electoral area’s shape for political gain” (Leventhal, 2016). When interpreting geographic data, we should keep in mind that the demographic description will not typically reflect every individual in that area, regardless of the unit of geography or location.

3.2 Anatomy of a credit score

The anatomy of credit scores is calculated through the Premier Equifax scoring model which uses information from lenders to divulge the specific scores in a range from 300 - 900. Such scoring models are used by the 3 major credit bureaus and in this study, Equifax's proprietary credit scores will be used.

Although the models and specific algorithms used are typically confidential information, Equifax states how the credit scores are calculated. Equifax has identified key characteristics and factors involving financial well-being that are used to calculate credit scores.

The main factors involved in calculating a credit score are:

- Your payment history
- Your used credit vs. your available credit
- The length of your credit history
- Public records
- Number of inquiries into your credit file

As shown in Figure 1, Equifax uses the following percentages when calculating credit scores.

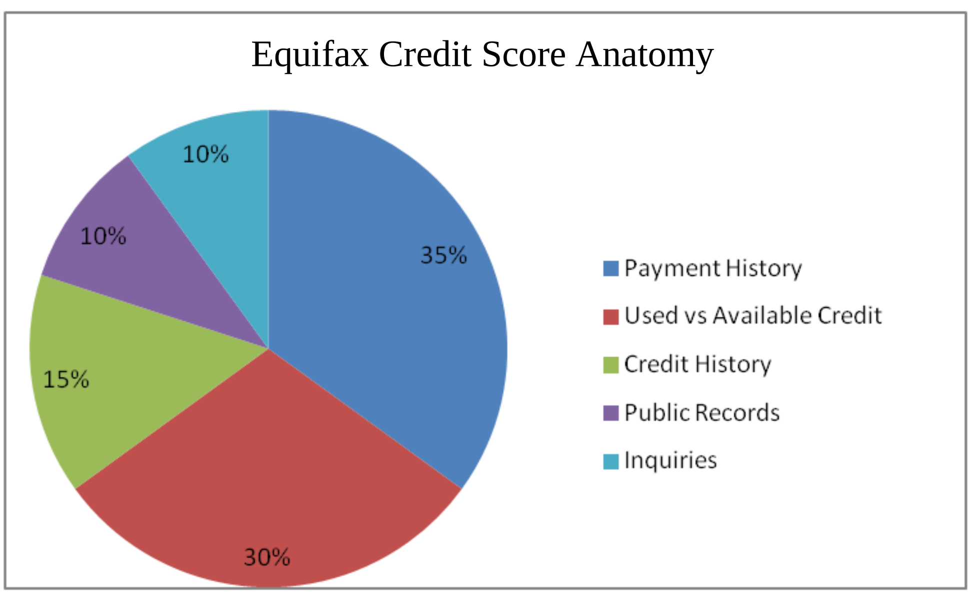


Figure 1 Main factors and percentages used in calculating credit scores

3.2.1 Payment History

The payment history is the largest influence for calculating credit scores, accounting for approximately 35% of the credit score. Payment history includes information such as repaying lines of credit such as credit cards, loans and potential investment property mortgages. However, this information not only includes how these payments have been made but also the time in which a person takes to make a specific payment.

Similarly, information such as late or missed payments are also included. Equifax states "Credit scoring models look at how late your payments were, how much was owed, and how recently and how often you missed a payment"(Equifax, 2019). Additionally, models often detail all credit accounts and separate the number of delinquent accounts compared to the total number of credit accounts on file.

3.2.2 Used vs Available Credit

When creating a credit score, 30% of the score is dependent on the total amount of credit that is available. For example, if a large sum of money is being owed on multiple credit cards, the available credit may decrease, depending on the lines of credit. This typically focuses on revolving lines of credit, which according to Equifax is a “line of credit is a type of loan that allows you to borrow, repay, and then reuse the credit line up to its available limit"(Equifax, 2019).

3.2.3 Credit History

Credit history accounts for 15% of the credit score calculation and details information about how long various credit accounts have been active and the ability for an individual to handle payments with the various accounts. Ideally, individuals will show good management of credit accounts across a lengthy period of time.

3.2.4 Public Records

Public records are specifically a collection of issues indicating problems with an individual's well-being and account for 10% of the credit score. The public records are strictly indicators that can show any indication of bankruptcy or behaviour that may be deemed "financially risky". A higher presence of these records can negatively impact the credit score value.

3.2.5 Inquiries

Inquiries account for 10% of the credit score calculation and are typically separated into two categories; *soft pulls* and *hard pulls*. Hard pulls are defined as inquiries that may impact the calculation of a credit score. Hard pulls are inquiries such as seeking a new credit card or applying for a new loan by an individual. Soft pulls are specific inquiries about a credit score or requests for credit / transaction history within an account. However, credit scores generally do not account for soft pulls and only measure hard pulls.

Many studies using credit scores have scarcely defined the ranges used for the analysis (Ding, Hwang & Divringi, 2016). Equifax has provided specific ranges recommended for analysis. Studies such as Ding, Hwang & Divringi (2016), do not use the Equifax recommended ranges which are fundamental in credit analysis and are industry standard. Using non-regulation ranges are problematic as many different groups may be incorrectly aggregated together. As such, the following ranges will be used:

Table 1

Credit Ranges

Credit Ranges	Implication
≤559	Poor
560 - 659	Below Average
660 - 724	Fair
725 - 760	Good
≥761	Excellent

3.3 Percent Change and linear differences

To better understand the depreciating or increasing credit scores within Toronto, linear differences will be calculated to understand such changes. Specifically, the **2018 - 2014** data will be used to calculate the linear difference for the 2010 year as percentages, using the following formula shown in equation 1.

$$\textbf{Raw Change} = \textbf{X} - \textbf{Y} \tag{1}$$

The **X** variable in the formula represents the newer year in the formula (2018), while the **Y** variable represents the older data (2014). However, as this simply calculates linear raw change, the data must be further converted using the formula in equation 2 to find the percentage change.

$$\textbf{Percentage Change} = \textbf{Raw Change} \div \textbf{Y} \tag{2}$$

Using the percentage change, a linear calculation will be used to predict the data for 2010, using the provided formula. However, it must also be understood that such equations are linear and cannot predict exponential changes or anomalies in data. The percentage change calculation was also used to examine change among the census variables for 2011 and 2016.

3.4 Variables

When choosing variables for this research, theoretical and analytical research was used to explore various frameworks used in academia. Additionally, research such as Tran (2016) and Sheng, Han & Zhou (2017), specified the use of variables to help understand gentrification concepts using various economic and demographic indicators. Therefore, a variable list was created to further examine such concepts and identify if such correlations exist:

- **Dwelling Value**
- **Average Household Income**
- **Visible Minority Population**
- **Housing Ownership** (Renters)
- **Shelter Costs** (Rented households)
- **Shelter Costs** (Owned households)

Each of these variables were obtained for the 2011 and 2016 years at the DA level for the City of Toronto through Statistics Canada via CHASS census analyzer. Moreover, each census variable was used to show the specific changes representative of gentrification.

3.4.1 Dwelling Value (% change)

As frequently reported in academic literature, one of the most common economic changes associated with gentrification is dwelling value (Eldaidamony & Shetawy, 2016). Gentrification traditionally is associated with property development as gentrification also includes changes and development in property landscapes. Such transformations are often followed by new housing. This creates a change in observed areas where older housing for the middle-low income is removed for modern and upscale housing alternatives.

3.4.2 Average Household Income (% change)

As indicated with the increase in Dwelling Value, such changes and displacement often influence a new demographic to uproot the previous inhabitants. Therefore, as these changes are common in gentrification, such variables can indicate a population change. Additionally, the increase in living costs are attributable to the higher income demographic entering the neighbourhood(s). Literature often shows that higher dwelling value and associated property tax increases have also led to the displacement and, therefore, the changing demographic in gentrified areas (Martin and Beck, 2016; Rosenthal and Ross, 2015).

3.4.3 Visible Minority Population (% change)

Literature has also emphasized that non-socio-economic variables can be more appropriate for indicating change. Specifically, racial turnover or visible minority turnover has been emphasized by many non-parametric gentrification research (Colburn & Jepson, 2012; Galster & Peacock, 1986). However, as racial turnover can be very difficult to measure, the visible minority census variable will be used. Moreover, this measurement will be created by using census data for different years to determine the change in percentage of visible and non-visible minorities.

3.4.4 Housing Ownership - Renters (% change)

Gentrification is characterized by a change in housing characteristics. This includes a decrease in renters that will often reflect a change in the neighbourhood. This change is often dependent on the development and housing changes that occur. However, literature has emphasized the importance of housing ownership when measuring gentrification by commonly stating that the number of renters will often likely decrease (Frank, 2017; Qian, He & Liu 2013; Róka-Madarász & Mályusz; 2013).

3.4.5 Shelter costs (Rented units)

Shelter costs increase in areas of gentrification which is partly attributed to more expensive housing being built in gentrified areas. Additionally, redevelopment of currently existing units is something that also frequently occurs in developed areas. However, redevelopment and new housing is likely to cause a change in shelter cost (Frank, 2017; Qian, He & Liu 2013). Specifically, when looking at rented units, this change also includes rent prices along with condo and maintenance fees.

3.4.6 Shelter costs (Owned units)

Similarly, a change is also expected when looking at owned units. As literature indicates that changes in property taxation and maintenance fees for owned units can be indicators of gentrification; such changes are often shown through the changes in shelter costs (Immergluck, 2009; Martin & Beck, 2018). Large changes in shelter costs are often representative of targeted development, and therefore, indicative of gentrification.

3.4.7 Variable transformation

To create a value that can determine the likelihood that an area has experienced gentrification, a composite score was needed to visualize changes between census periods. However, prior to the gentrification scoring, the variables were changed using the appropriate methods to convert each variable into a similar 0 - 100 score range.

3.5 Index creation

Indices were created to initially transform the data into a unique value. The resulting transformation would, therefore, show a percentage difference from the mean. To show this, each value was divided by the mean value of that variable, for the City of Toronto, and then further multiplied by 100. Therefore, values of 100 would indicate that the data is equivalent to the mean. Values above or below 100 would indicate a percentage difference from the mean. This transformation was done for each socio-economic variable. However, the newly created indices included negative values due to negative percentage change values. As each of the variables still included negative values which showed percentage decrease instead of decrease, it would be very difficult to measure these changes using traditional scoring methods. Therefore, an additional transformation was conducted to transform the variables into a scored value indicating a high or low percentage change.

3.5.1 Min / Max Scaling

After the creation of the index, values were divided by the maximum value to normalize each of the socio-economic variables relative to the highest value.

$$\boldsymbol{x}_{\mathbf{i}} = \frac{(x - x\mathrm{min})}{(x\mathrm{max} - x\mathrm{min})} \tag{3}$$

The formula shown in equation 3 is a linear transformation which uses the minimum and maximum values to transform the converted data values into a score range. In equation 1, the $\boldsymbol{x}_{\mathrm{i}}$

value in the equation represents the rescaled value. The *x* value in the formula is the index value that was created for each DA in the City of Toronto. The formula was able to recognize the lowest negative value as the minimum and therefore helped to convert each index value into a standardized unique value. Although this formula rescales the index values, an additional transformation was needed to convert each value into a score.

3.5.2 Score-Range transformation

A score range transformation was used to convert each value to a score from 0 - 100. Therefore, each value would be a score indicating high or low values and be representative of the maximum value. Therefore, the following formula was used:

$$\boldsymbol{x}_{\mathbf{ii}} = \frac{xi}{x\max} * 100 \tag{4}$$

The $\boldsymbol{x_{ii}}$ value in equation 4 represents the newly scored value. Each rescaled value from the min-max scaling formula was divided by the maximum value and multiplied by 100 to create a score range from 0 - 100. The value of 100 would be representative of the maximum value and therefore show the highest score.

3.5.3 Gentrification Scoring

The final gentrification scores were created using each of the socio-economic variables. As each variable was chosen as a representative variable for gentrification; the scores were then combined and weighted to create a composite gentrification score. Additionally, as the variables used in this analysis are both positive and negative indicators of gentrification, each variable was weighted separately. The variables used also contained specific indications of gentrification. Therefore, each variable was given a weight with the relative importance in the study, and the potential indication of gentrification.

Table 2

Census Variables & Weights

Variable	Indication of Gentrification	Weight
Average Household Income	Positive Percent Change	16.67
Average Dwelling Value	Positive Percent Change	16.67
Visible Minority Population	Negative Percent Change	16.67
Housing Ownership (Renters)	Negative Percent Change	16.67
Average Shelter cost (rented households)	Positive Percent Change	16.67
Average Shelter cost (owned households)	Positive Percent Change	16.67
Total		100

As shown in Table 2, each variable and the associated indication of gentrification was noted to understand the likely effect gentrification would have on a neighbourhood and therefore identify the potential influence each variable should have on the gentrification index. In this study, each variable was assumed to have equal importance. The variables were then multiplied with the according weights and summed evenly to represent the final index.

Using the frameworks identified by the relevant literature, as well as the geospatial census variables, the regression will indicate the level of association with the **independent** census variables and the credit scores as the **dependent** variable. As this study strictly examines an empirical measurement of gentrification; assumptions based on data such as zoning by-laws and non-empirical data will not be used in this analysis.

The index is a composite variable used to measure the likelihood that an area has experienced gentrification. Therefore, specific benchmark values were used to indicate varying likelihoods. As shown in table 3, the values were converted into 4 separate categories indicating how unlikely or likely a DA would be to have experienced gentrification.

Table 3

Gentrification Index values

Index Value	Likelihood to have experienced Gentrification
100 - 86	Very Likely
85 - 81	Likely
80 - 76	Unlikely
75 - 0	Very Unlikely

The benchmark values were created using the **Jenks natural breaks** classification. Although other options such as equal interval, quantile, defined interval, geometric interval, manual interval and standard deviation are feasible, the natural breaks method was chosen to identify ranges that best group similar values and that maximize the differences between classes (ESRI, 2019[a]).

3.6 Chi-squared analysis

To examine if there is a relationship between gentrification and credit scores, chi-squared testing was used to compare the likelihood that a DA has experienced gentrification, to each of the credit bands. The chi-squared test is a measurement of independence for non-parametric data.

Specifically, the test measures group differences between categorical variables and determines if an association exists. It should be understood that chi-squared testing only tests observations for evidence of association or no association. Chi-squared testing does not produce estimated levels of correlations and confidence intervals.

Additionally, the contingency coefficient will be used when examining the chi-squared results to determine the level of association between the categorical variables. As the contingency coefficient indicates the strength of the association, the conventions and benchmark values shown in table 4 will be used to measure the level of association.

Table 4

Contingency Coefficient benchmarks (Blaikie, 2003)

Coefficient Value	Strength of Association
0.00	None
0.01 - 0.09	Negligible
0.10 - 0.29	Weak
0.30 - 0.59	Moderate
0.60 - 0.74	Strong
0.75 - 0.99	Very Strong
1.00	Perfect

The categorical variables that will be examined will be the **credit score band** and **the likelihood that a DA has experienced gentrification**.

3.6.1 Hypothesis testing

Chi-squared analysis also requires testing via a hypothesis to determine the presence or absence of a tested association. The hypothesis testing follows a specific format where a null hypothesis and an alternate hypothesis is used to specify outcomes. Generally, the null hypothesis is defined by a denoted H_0, while the alternate hypothesis is defined by a denoted H_a or H_1.

Through each band of credit scores; the chi-squared analysis will indicate if there is a higher level of association with the credit scores and gentrification. Although testing for such associations can be conducted through Pearson's correlation, literature has indicated that there are varying levels of association with bands of credit and gentrification.

3.7 Regression

To examine if financial well-being indicators such as credit scores accurately identify areas of gentrification or urban development, regression will be used to determine the correlation amongst the credit scores and selected census variables that act as a surrogate for gentrification, while also indicating the percentage of explained variance by each compared variable. As previously mentioned, the independent census variables for **2011** and **2016** were used. The percentage change among the observed years, deviate in numeric format. Therefore, each variable was standardized to avoid the previously mentioned issues. To examine the influence of each census variable on the credit scores (dependent variable), the (β) beta value will be used.

3.7.1 Standardization

The data manipulation for the regression analysis required each variable to be standardized using two methods: **Standard Score** and **Logarithm**. Both methods of data manipulation will be used when creating the regression model.

The logarithm function is typically described as the inverse of the exponent function. To avoid the influence of skewed distributions across each variable, a logged and unlogged regression model will be heavily influenced by skewed data.

The regression analysis was conducted using standardized variables as z-scores. Each socio-demographic variable was standardized into z-scores (standard score) to address issues within the data. These issues are similar to the issues mentioned in MAUP in which data could be representative of differing scales. For example, the average household income variable is in a different scale of measurement when compared to other variables. To avoid comparing aggregate percentages and averages, each variable was standardized into z-scores to show standard deviations from the mean, by using the following formula shown in equation 5.

$$z = \frac{x - \mu}{\sigma}$$

$$\mu = \text{Mean}$$
$$\sigma = \text{Standard Deviation} \qquad (5)$$

3.7.2 Linear Regression

As variations of linear regression can be traced back to the early 19th century, many academic studies often use regression models, along with predictor variables, for various observations. Traditionally, the linear regression model would follow the following formula:

$$\mathbf{Y} = a + bX1 + bX2 .... bXn \tag{6}$$

The above equation indicates a linear predicted output. Credit scores are the dependent Y variable which the regression model is trying to predict. The b indicates the slope of the model, while the X represents the estimated values of each variable used for the predicted Y.

However, such regression models are linear; and therefore, contain associated limitations. The first of which is the output. Linear models do not account for exponential changes in the data distribution. Such assumptions are often overlooked in research and various studies do not identify this assumption as a limitation.

The regression analysis was used to create a model indicating gentrification. As such, the variables used as the independent variables will be defined by exploratory analyses along with the composite indicators of gentrification. Additionally, the dependent variable will be the credit scores; to indicate if such composite analysis is better represented by the financial well-being indicator(s).

To measure the various models and results, the R and R^2 values will be used to indicate model suitability. The R^2 is often described as the coefficient of determination in mathematics. The R^2 value indicates the percentage of variation of the dependent variable, that can be explained by the independent variables. To indicate model strength, a scale of R^2 values

Additionally, regression can be conducted using different methods. Traditional regression models use only a specific variation of variables that are user determined to be included in the model. However, the selection of the significant variables can also be determined automatically through the ***Stepwise*** method. Stepwise regression tests each independent variable *one-by-one*. The stepwise regression can either be conducted by including all potential independent variables and removing the variables that are not statistically significant, or by only adding the most statistically significant variables into the model one-by-one starting with the most significant variable. More specifically, the variables used will include the 2016 and 2011 years along with the percentage change across each of the census variables.

3.7.3 Multicollinearity

Although this study, along with relevant literature, has indicated the importance of using very similar variables, it may heavily influence specific results. Such high correlations amongst the variables, both dependent and independent, along with high R^2 values may be artificial. For example, using the same income variable across varying years, along with a percentage change variable may heavily influence the results. The repetition of the same variable is rarely used due to such heavy influence of those specific variables. However, as this analysis distinctly requires variables across varying years, high correlation values across multiple variables are expected.

3.8 Spatial Autocorrelation

To examine if credit scores of similar values cluster together, spatial autocorrelation will be used to identify the spatial patterns and distribution of credit scores across the City of Toronto. Spatial autocorrelation is typically measured to describe the presence of clustering or systematic variation within an entity or variable. (ESRI, 2019[b]) Variables that are measured typically contain attributes indicating high or low values either within a specific region or along a continuous surface. Spatial autocorrelation can be both positive or negative. Positive spatial autocorrelation indicates the spatial presence of similar values clustering together, while negative spatial autocorrelation is typically defined through dissimilar values clustering together. Testing for spatial autocorrelation can identify if spatial data should be further investigated. Moreover, unique patterns of spatial variability may indicate underlying political, economic or social influences.

3.8.1 Hot-Spot Analysis

Getis-Ord Gi* is a local statistical calculation which highlights significant clustering of high or low values. It is a single measurement which can identify hot and cold spots spatially distributed across a geographic region. The hotspot analysis was conducted to identify if statistically significant high or low values are likely to cluster together within the city of Toronto. Hot spots indicate areas of statistically significant high values clustering together, while cold spots indicate statistically significant low values clustering together. To accommodate the various confidence levels, and therefore significance levels, the **90**, **95** and **99** confidence levels were all shown through the Getis-Ord Gi*. The Getis-Ord Gi*statistics was calculated through the following formula:

The Getis-Ord local statistic is given as:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j}x_j - \bar{X}\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{\left[n\sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}} \tag{7}$$

where x_j is the attribute value for feature j, $w_{i,j}$ is the spatial weight between feature i and j, n is equal to the total number of features and:

$$\bar{X} = \frac{\sum_{j=1}^{n} x_j}{n}$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \left(\bar{X}\right)^2} \tag{8}$$

The G_i^* statistic is a z-score so no further calculations are required.

Additionally, the Getis-Ord Gi* calculation shown in equation 7 and 8 also provides p values and z-score statistics to measure statistical significance. Generally, spatial clustering of high values can be indicated by high z-scores and small p-values while spatial clustering of low values is represented by low z-scores and high p-values.

3.8.2 Moran's-I

To indicate if similar credit scores are likely to cluster together, spatial autocorrelation will also be tested to examine the correlation of a variable, within itself through space. Using Global Moran's I, spatial autocorrelation will be measured by examining both feature locations and values simultaneously (ESRI, 2019[d]). The Global Moran's I calculation creates an index identifying the type of spatial autocorrelation present within the data set. The index is a unique value between -1 to +1. Values closer to -1 indicate clustering of dissimilar values while values close to +1 indicate clustering of similar values. Generally, the value of 0 represents no autocorrelation and therefore indicates complete randomness within the dataset.

However, to further evaluate the index and therefore identify its significance as a statistic, hypothesis testing must be carried out. Similar to Getis-Ord Gi*, the Moran's I calculation also includes a p-value and z-score values.

The Moran's I statistic for spatial autocorrelation is given as:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} z_i z_j}{\sum_{i=1}^{n} z_i^2} \tag{9}$$

where z_i is the deviation of an attribute for feature i from its mean $(x_i - \bar{X})$, $w_{i,j}$ is the spatial weight between feature i and j, n is equal to the total number of features, and S_0 is the aggregate of all the spatial weights:

$$S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} \tag{10}$$

The z_I-score for the statistic is computed as:

$$z_I = \frac{I - \mathrm{E}[I]}{\sqrt{\mathrm{V}[I]}} \tag{11}$$

where:

$$\begin{aligned} \mathrm{E}[I] &= -1/(n-1) \\ \mathrm{V}[I] &= \mathrm{E}[I^2] - \mathrm{E}[I]^2 \end{aligned} \tag{12}$$

Global Moran's I shown in equation 9 follows hypothesis testing using the traditional null and alternate hypothesis identified in literature. Hypothesis testing, in this case, is also heavily dependent on the p-value. Observations where p-values are not statistically significant cannot reject the null hypothesis, while observations where p-values are statistically significant generally reject the null hypothesis. Additionally, z-scores also indicate the spatial distribution. For example, a statistically significant p-value with a positive z-score will indicate high spatial clustering of high values or low values, while negative z-scores indicate a heavily spatially dispersed pattern where high values repel other high values and low values repel other low values.

Additionally, Moran's I is divided into two parts, local and global. Global Moran's I is a measure of general autocorrelation within the entire study area. However, this assumes homogeneity amongst the entire study space. Therefore, to avoid such assumptions, local

Moran's I is also used to evaluate clustering on an individual (micro) level, instead of examining the entire study area. Anselin local Moran's I identifies statistically significant clusters of hot or cold spots and spatial outliers.

The Anselin local Moran's I identifies spatial clusters of significantly hot or cold spots, along with outliers with z-scores and p-value similar to global moran's I. However, there are a few assumptions that are consistent with Anselin local Moran's I. Firstly, Anselin local Moran's I only identifies significant outliers for a 95% confidence interval, unless otherwise stated. Secondly, there are varying distance calculations and thresholds that can be used. Thirdly, a varying number of permutations can be used to "determine how likely it would be to find the actual spatial distribution of the values you are analyzing" (ESRI, 2019[e], para. 6). In this specific case, the maximum number of permutations **9999** was used to identify all the statistically significant spatial outliers.

Additionally, the Anselin local Moran's I also separates each cluster or values by using associated z-scores and p-values. High positive z-scores indicate spatial clustering of similar values amongst the neighbouring features. Therefore, this will most likely result in a Low-Low cluster (cold spot of significantly low values) or a High-High cluster (hot spot of significantly high values). Similarly, low-negative z-score values indicate spatial clustering of dissimilar values amongst the neighbouring features (ESRI, 2019[e]). Traditionally, these z-score values must be less than -3.96. These values are shown as outliers through a HL designation where a feature of high value is surrounded by features of low values or LH where a feature with a low value is surrounded by various features with high values. Although literature identifies that both Getis-Ord Gi* and Anselin local Moran's I are both methods of examining spatial autocorrelation, they are quite different. Getis-Ord Gi* considers three elements of spatial analysis:

1. **A study area**
2. **Features with values, such as polygons**
3. **A surrounding area or *neighbourhood* for each feature**

Getis-ord identifies if each feature can be categorized as part of a hot or cold spot depending on the values. However, Getis-Ord GI* specifically looks at each feature and identifies if its neighbouring features are significantly different from the study area. Anselin local Moran's I also uses the same elements of spatial analysis, however, the Moran's I analysis also identifies cluster-outliers by examining if each feature is significantly different from its neighbourhood.

Therefore, both methods of spatial autocorrelation are needed to examine the presence of spatial autocorrelation in the data. Additionally, examining spatial autocorrelation through Getis-Ord Gi* and Anselin Local Moran's I requires two assumptions to be met:

- **Assumption of established spatial relationships**
- **Assumption of distance**

The first assumption requires spatial relationships to be modelled and conceptualized through one of the various available methods:

- Inverse Distance (Impedance)
- Distance Band (Sphere of influence)
- Zone of indifference
- Polygon Conginuity (First Order)
- K Nearest Neighbours
- Delaunay Triangulation (Natural Neighbours)
- Space-Time window

Each of the available methods uses specific parameters to identify the spatial relationships between each entity and its neighbours. However, this analysis will strictly use the Inverse Distance method which conceptualizes spatial relationships through a distance decay function (ESRI, 2019[f]). The specific distances between features were measured through Euclidean distance in ArcGIS.

4.0 Analysis and Discussion

4.1 Credit Score Analysis

First, it is important to establish that credit scores are not the same as income variables. Literature often states that income is closely related to credit scores, however, they are also very different as mentioned throughout this study. Although unlikely, a household of high-income can have poor credit. Similarly, it is also completely possible for a household of low income to have excellent credit.

As variations in credit scores are largely indicative of changes in financial well-being, we must first examine where these changes are occurring. As shown in figures 4 & 5, the 2018 & 2014 credit scores for DAs in the City of Toronto show clustered patterns of high credit scores in the Rosedale - Bridle Path area, along with the Kingsway & Princess Margaret area in Etobicoke.

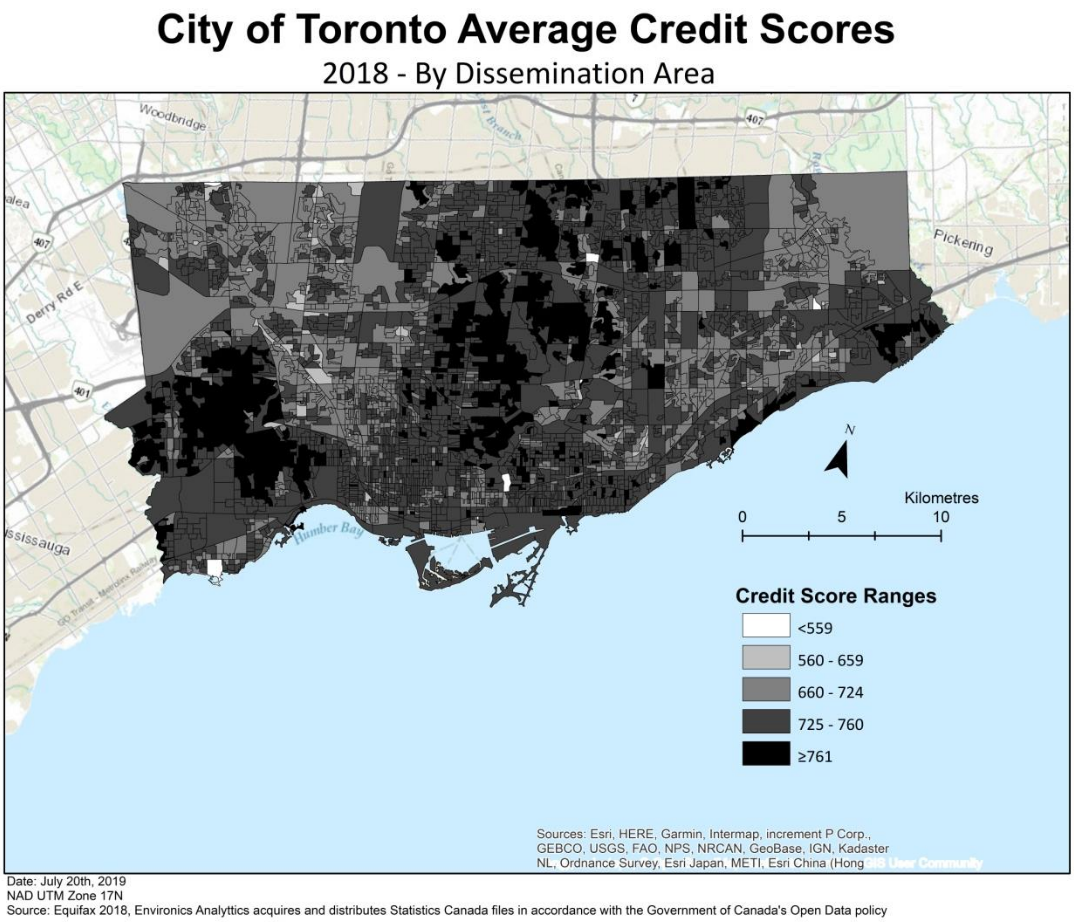


Figure 2 City of Toronto 2018 Credit Scores by DA

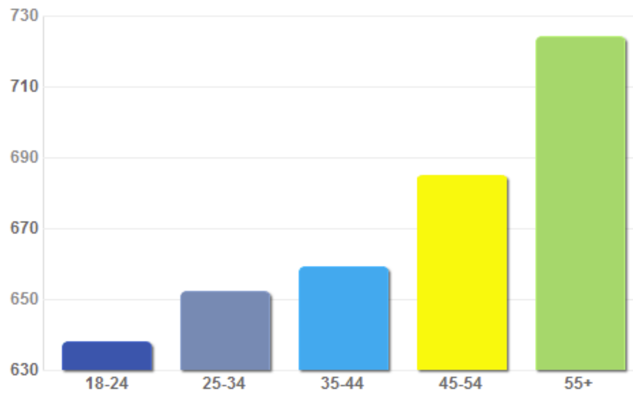


Figure 3 Average Credit Scores by Age (Credit Karma, 2015)

As shown in figure 3, credit scores do show correlations with various demographic variables. However, it was important to examine if trends in credit scores exist spatially across various years and with census variables that are common with gentrification.

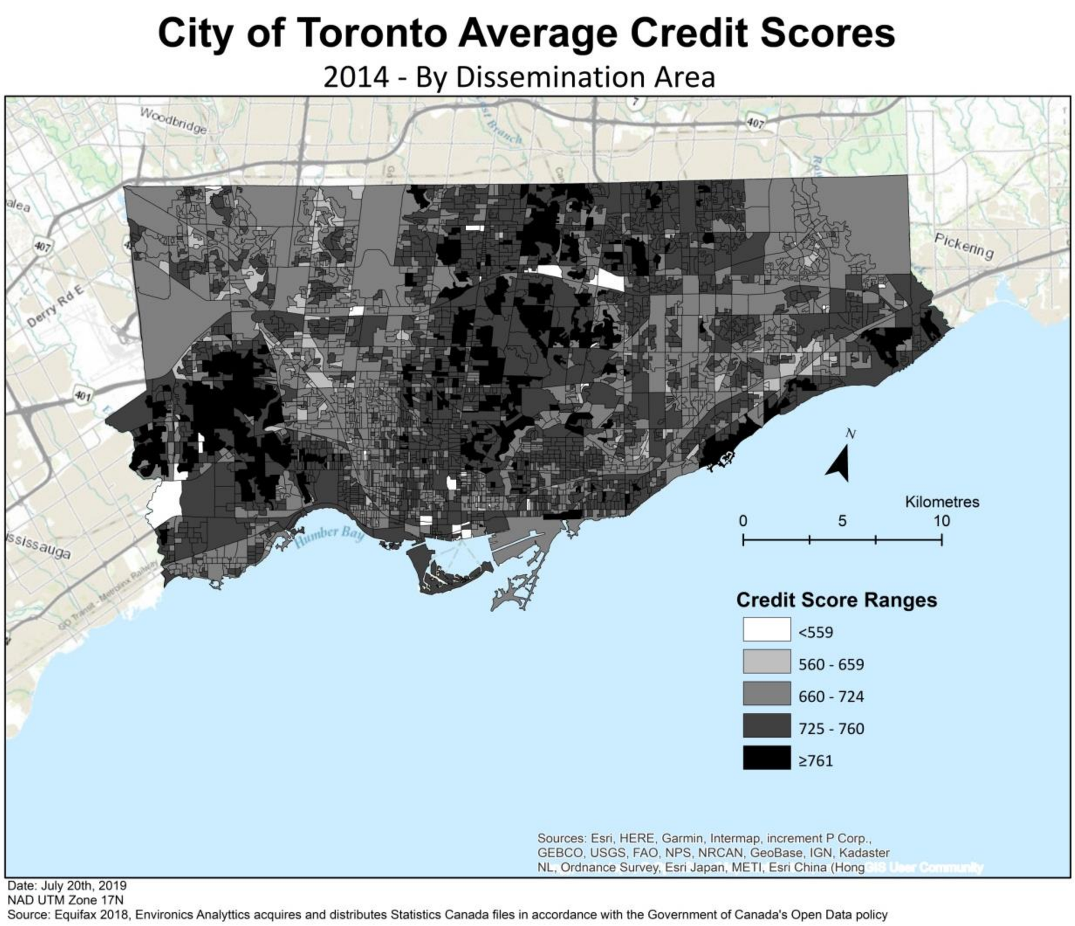


Figure 4 City of Toronto 2014 Credit Scores by DA

Both the 2018 & 2014 data show a clustered pattern of credit scores; however, the 2018 data shows a large, yet dispersed, increase in values across many DAs scattered within the City of Toronto. Shown in figure 5, the 2010 values also indicate this trend. However, the 2010 values show a more dispersed spatial distribution, with lower credit ratings. There is also an increase in credit scores across the DAs in North York.

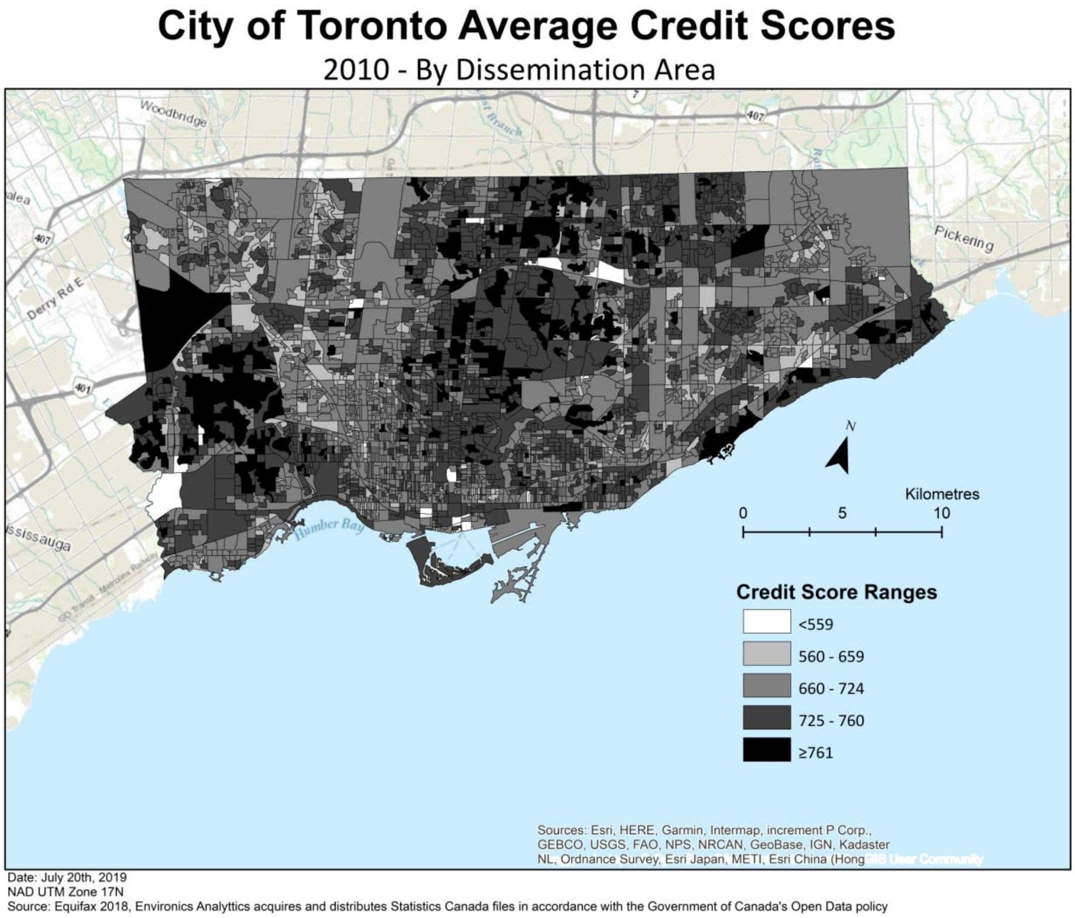


Figure 5 City of Toronto 2010 Credit Scores by DA

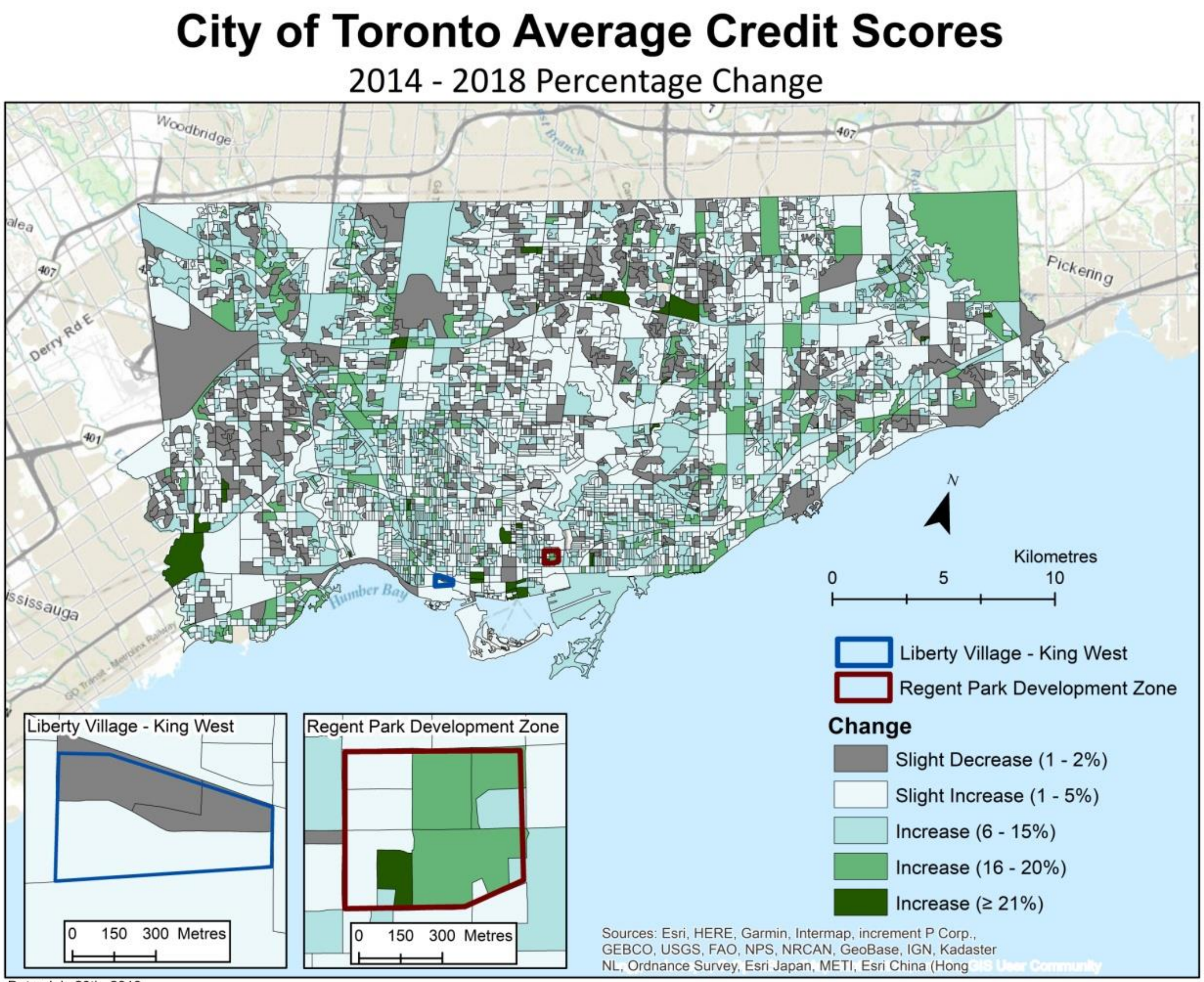
Table 5 shows the difference in count and percentage of DAs within a specific credit band.

Table 5

Credit Bands by Dissemination Area

	Dissemination Areas					
Credit Band	2018		2014		2010	
<560	9	0%	45	1%	47	1%
560 - 660	56	2%	140	4%	312	8%
660 - 725	1171	32%	1520	41%	1632	44%
725 - 760	1954	53%	1594	43%	1332	36%
>760	512	14%	403	11%	379	10%

As shown, the credit scores for the City of Toronto in 2018 had the largest number of DAs within the two best credit bands. As these two bands accounted for almost 70% of the DAs within the City of Toronto, the remaining DAs were split across the remaining 3 bands. 2014 & 2010 are different as the 3 lowest credit bands account for approximately 50% of the DAs within the City of Toronto. These changes can also be shown spatially in figure 6 and 7.



Date: July 20th, 2019
NAD UTM Zone 17N

Figure 6 City of Toronto Credit Score changes 2014 - 2018

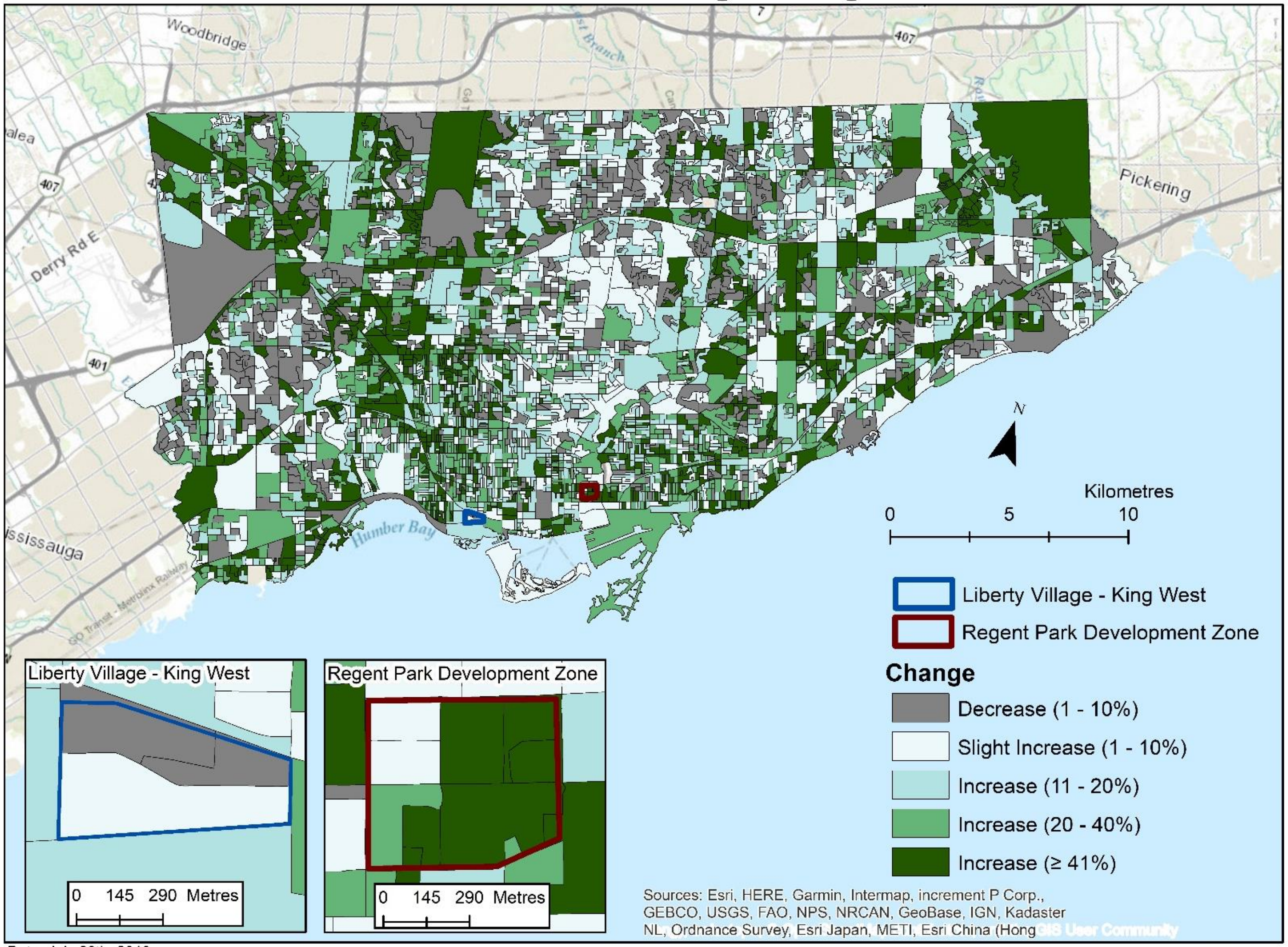
Source: Equifax 2018, Environics Analyttics acquires and distributes Statistics Canada files in accordance with the Government of Canada's Open Data policy

As shown in figure 6, most of the DAs in the City of Toronto experienced an increase in average credit score between 2014 to 2018. The areas with the largest increase were clustered in the downtown core, with few outliers in Etobicoke and North York. However, few DAs across the City of Toronto also experienced a slight decrease of approximately **1 – 5 %.** These areas were dispersed throughout the City of Toronto but largely clustered in Etobicoke and North York as well.

As shown in figure 7 the largest increase in credit scores were shown in Regent Park. The 2010 - 2018 changes showed larger clusters of credit decrease in neighbourhoods such as Humber Valley and Willowdale East. Additionally, larger clusters of 40% increase were shown in the Weston and Beechborough-Greenbrook neighbourhoods. Areas with the largest average household income such as Bridle Path and Rosedale only showed marginal credit score change (slight increase and decrease) across 2010 to 2018. As shown, the credit scores in the City of Toronto have been increasing in the past decade, with outliers showing credit score decreases dispersed throughout parts of the city. However, to examine if the average credit scores increased throughout the City of Toronto, the mean and median values of each year was compared.

City of Toronto Average Credit Scores

2010 - 2018 Percentage Change



Date: July 20th, 2019
NAD UTM Zone 17N

Figure 7 City of Toronto Credit Score changes 2010 - 2018

Source: Equifax 2018, Environics Analyttics acquires and distributes Statistics Canada files in accordance with the Government of Canada's Open Data policy

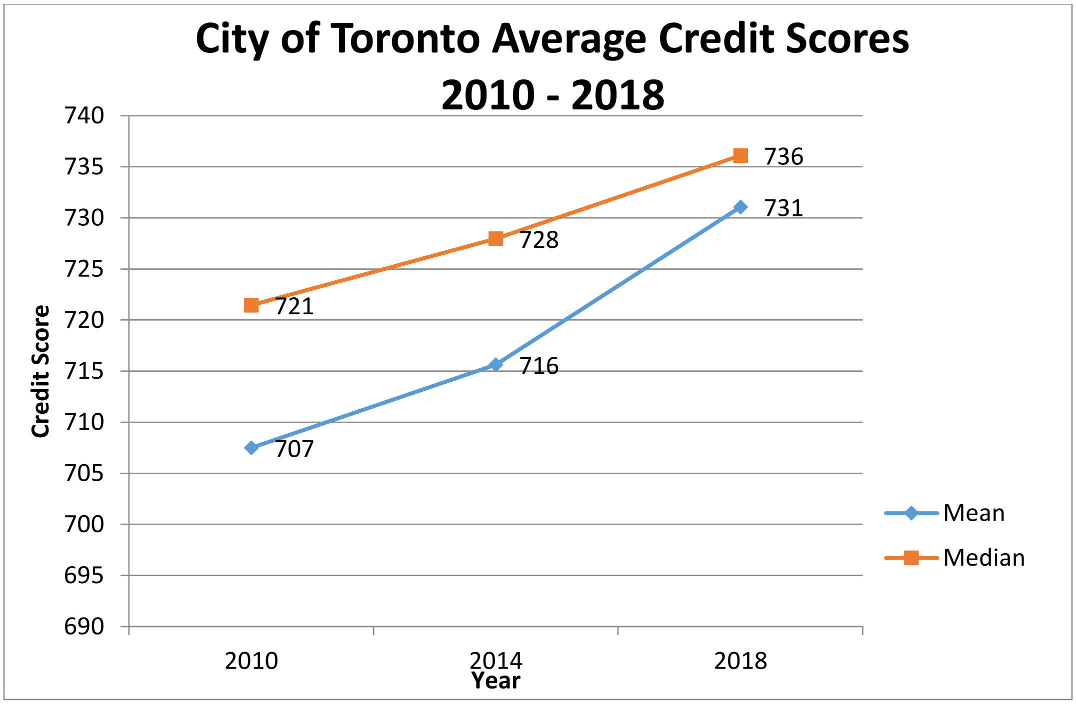


Figure 8 City of Toronto Average Credit Scores

Figure 8 shows the change in the median and mean credit scores across each of the observed years. Noticeably, the largest increase was the 2014 - 2018 years, which showed an increase in mean credit score by 8 and median credit score by 15 across the City of Toronto.

4.2 Spatial Autocorrelaton

4.2.1 Global Moran's I

The spatial autocorrelation results were split into two separate analyses, the Getis Ord GI* analysis and the Morans-I. The Morans-I or Morans index analysis was measured through ArcGIS by using the 2016 credit score values for the City of Toronto. However, to interpret the Global Morans-I, the null and alternate hypothesis must first be stated. The null hypothesis denoted by $\mathbf{H_0}$ indicates that there will be complete spatial randomness and the alternate hypothesis denoted by $\mathbf{H_a}$ indicates that there is not complete spatial randomness among the 2018 credit scores.

Table 6

Global Moran's I Summary

Global Moran's I Summary	
Moran's Index	0.069756
Expected Index	-0.000266
Variance	0.000009
Z-score	23.490969
P-value	0

The Global Morans-I analysis indicated a p-value of 0, a z-score value of 23.49 and an Index value of 0.069. Therefore, we can reject the null hypothesis at the 95% confidence interval as the p-value of $0 < 0.05$. Additionally, as there is a positive z-score, the spatial distribution of high / low values indicates a presence of spatial clustering. Moreover, the Moran's Index value indicated a value of approximately **0.07** which shows slight **spatial clustering of similar values**. However, as Global Morans-I only uses a statistic to examine the spatial distribution of values across the entire study area, Getis-Ord GI* examines the specific areas where hot and cold spots are located at a specific confidence level.

*4.2.2 Getis-Ord GI**

Getis-Ord GI* was used to examine the spatial distribution of high and low values of credit scores. The **90**, **95** & **99%** confidence levels were used to examine the hot spots in the Getis-Ord analysis.

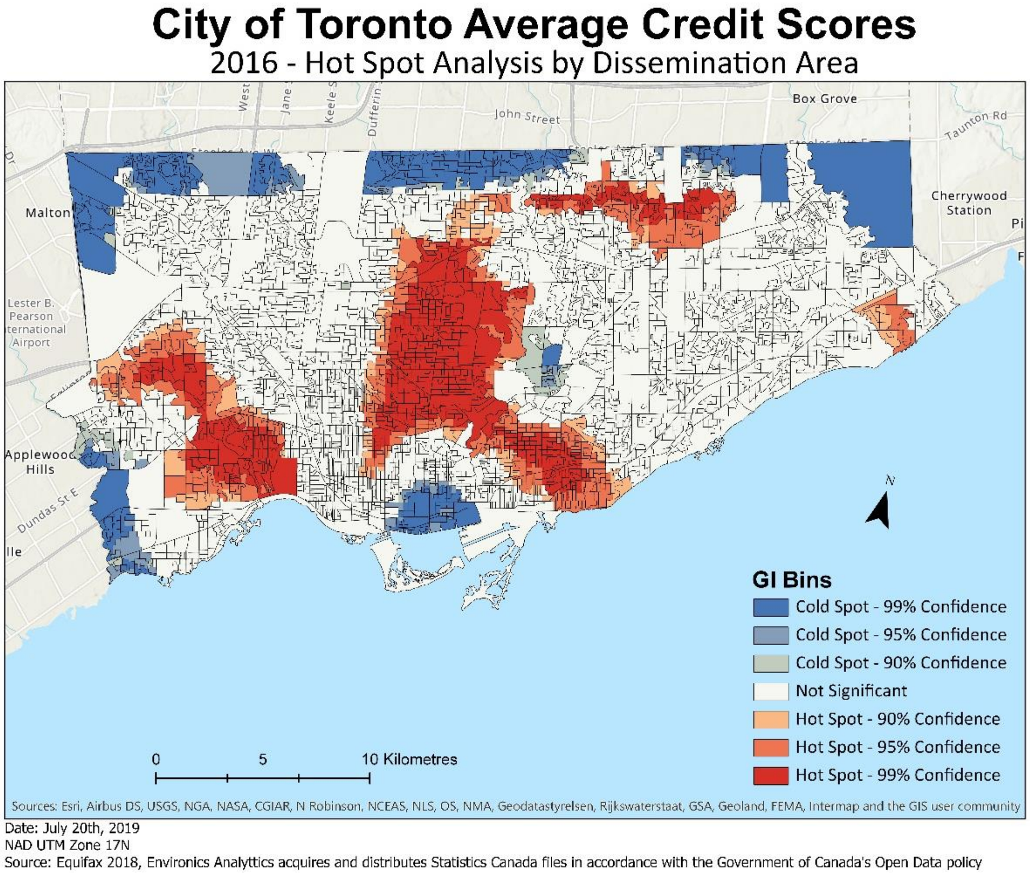


Figure 9 City of Toronto 2016 Credit Scores - Hot Spot analysis

Specifically, in figure 9, the GI bins were used to examine each of the hot and cold spots within the City of Toronto. The largest hot-spot are clustered around the Rosedale-Bridle Path area and extending to the lower beaches area. The second hot-spot is clustered around the Humber valley - Kingsway area. The cold spots were mostly located across the border of the City of Toronto along Steeles avenue, as well as in parts of the downtown core. Many of the areas across the eastern part of the city do not show statistically significant hot or cold spots. However, this does not show spatial outliers.

4.2.3 Anselin Local Moran's I

The Anselin Local Moran's I identifies the spatial clustering of outliers such as High-Low DAs and Low-High DAs, as well as High-High and Low-Low clusters. As shown in figure 10, spatial outliers are heavily dispersed across the City of Toronto.

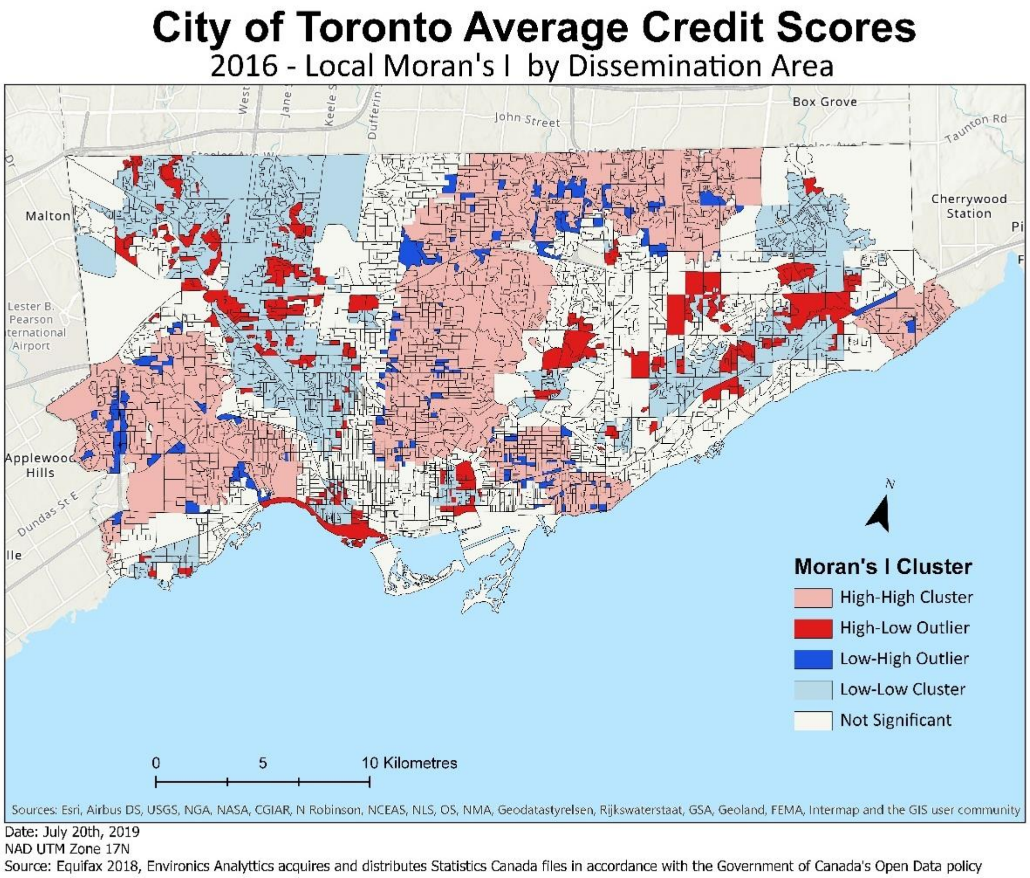
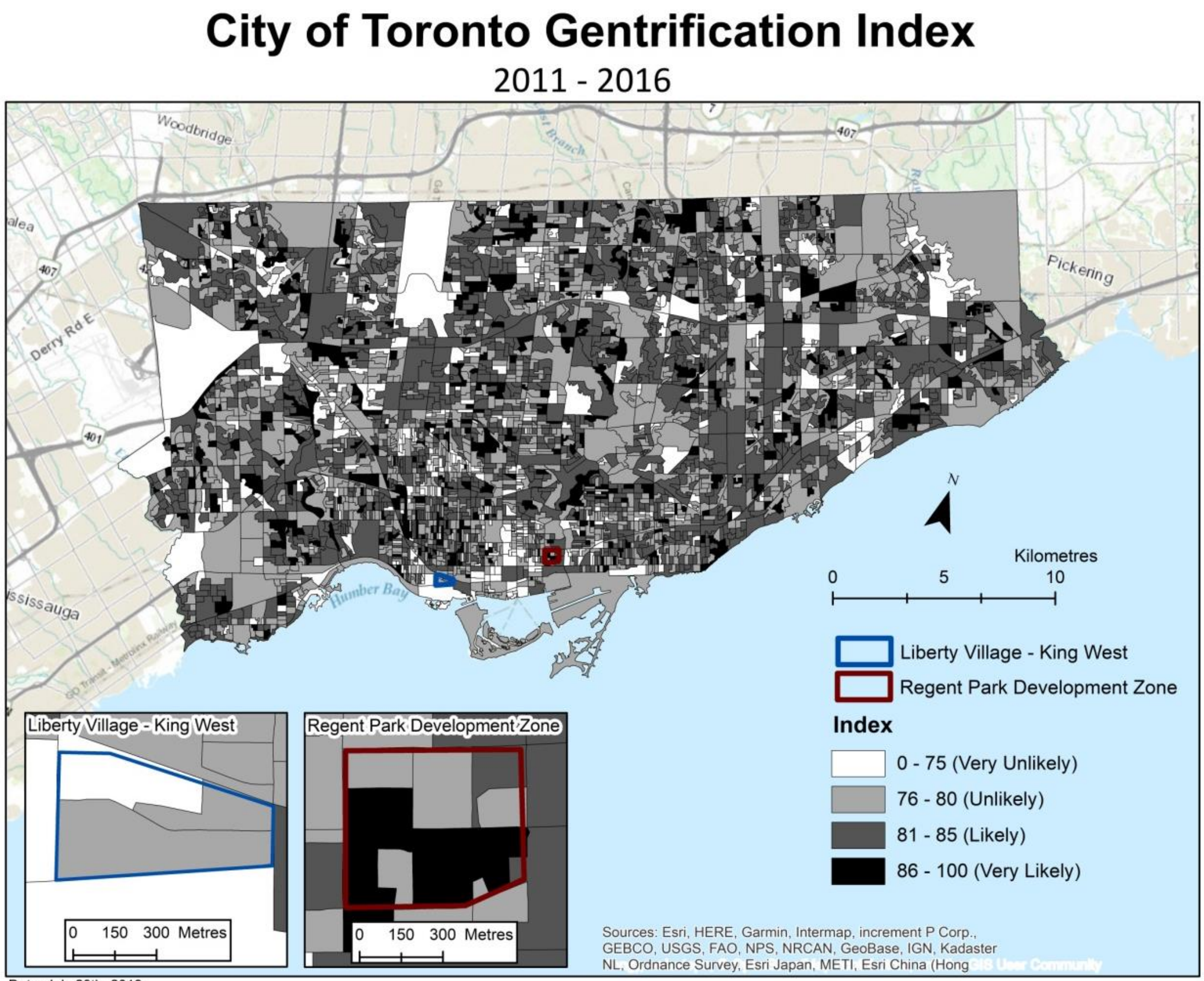


Figure 10 City of Toronto 2016 Credit Scores - Local Moran's I Analysis

Although there are many spatial outliers, the spatial distribution of high-high clusters and low-low clusters are very similar to the results of the Getis-Ord GI* hot spot analysis. However, there is also a large low-low cluster in the Humber Summit - Downsview area, extending to the high park area. Therefore, as shown through the spatial autocorrelation analysis as well as the statistical analysis, there is a presence of spatial clustering of similar values of credit scores within the City of Toronto.

4.3 Gentrification Analysis

As previously mentioned in this report, the gentrification analysis used a created index to examine the potential likelihood that an area has experienced gentrification by using the percentage change of specific variables. The created index was scaled from 0 - 100 and is shown in Figure 10 across the City of Toronto.



Date: July 20th, 2019
NAD UTM Zone 17N
Source: Equifax 2018, Environics Analyttics acquires and distributes Statistics Canada files in accordance with the Government of Canada's Open Data policy

Figure 11 City of Toronto 2016 Gentrification Index

The created index in figure 11 shows a largely random and dispersed pattern with clusters of high and low index values across various parts of the City of Toronto. However, there is also a noticeable cluster of low values across the southern border of the City of Toronto.

The created index was compared to the 2016 credit scores, which was the base year for the percentage change calculation. As shown in figure 12, the 2016 credit scores show large clusters of each credit band.

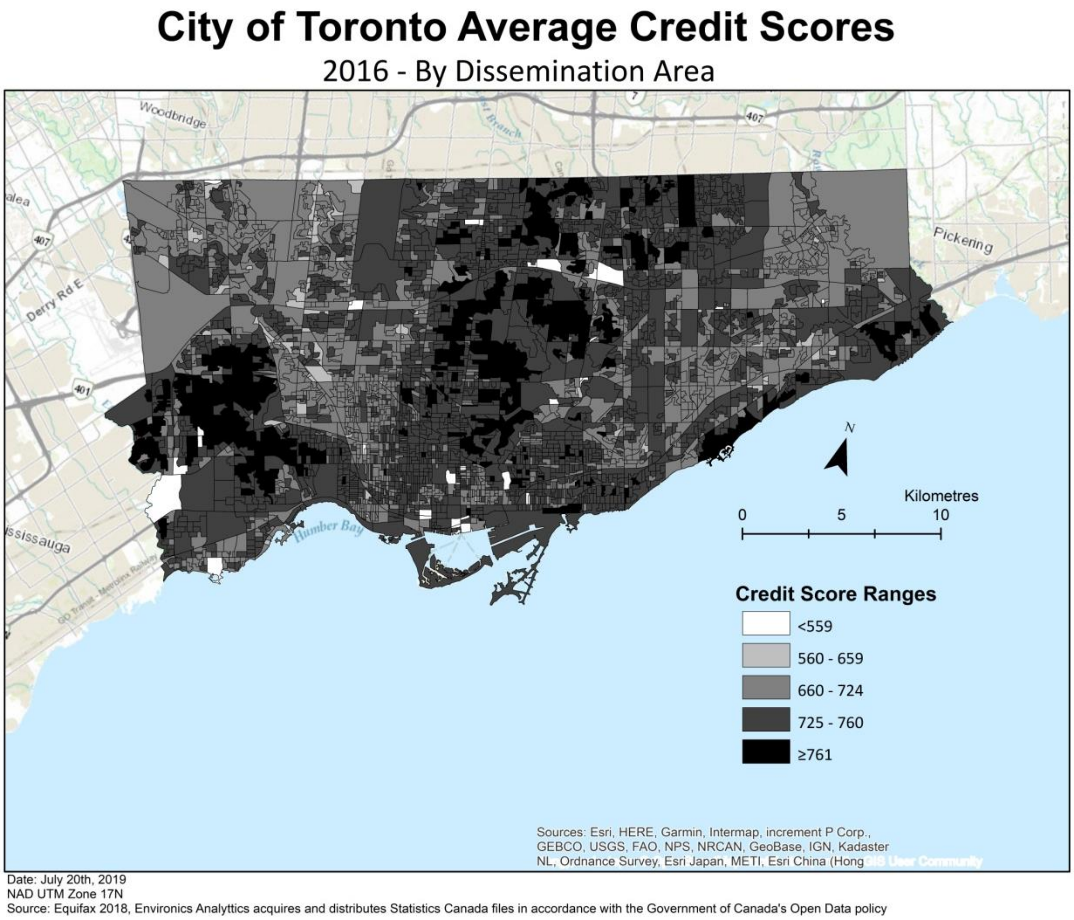


Figure 12 City of Toronto 2016 Credit Scores

A cluster of low credit scores was shown across parts of the downtown core, Birchmount and West Humber-Clairville. There were also outliers of DAs with low average credit score values dispersed across the City of Toronto. The largest credit scores were clustered around the Kingsway - Princess Margaret area as well as parts of Bridle Path and Rosedale. However, to examine if an association exists between both the credit scores and the created gentrification index. Chi-squared and regression was used to measure if credit scores can be used to examine gentrification.

4.3.1 Chi-squared

The first part of the chi-squared analysis was to examine if credit scores and the gentrification index identify any significant results in the crosstab analysis. As shown in table 7 approximately 90% of the lowest credit band, fall within the **Very Unlikely** category to have experienced gentrification.

Table 7

Chi-squared Cross-Tabulation

Likelihood to have experienced gentrification		Credit band				
		≤559	560 - 659	660 - 724	725 - 760	≥761
Very Unlikely	Count	35.00	31	157	171	194
	Expected Count	4.59	9.5261	153.565	206.704	173.01
	% within Likelihood	8%	7%	37%	40%	13%
	% within Credit Band	88%	37%	12%	9%	44%
	% of Total	1%	1%	4%	5%	5%
Unlikely	Count	5.00	0.45	534.00	685.00	174.00
	Expected Count	15587.00	32.34	521.40	701.82	171.46
	% within Likelihood	0%	3.00%	37.00%	47.00%	12.00%
	% within Credit Band	13%	54.00%	40.00%	38.00%	40.00%
	% of Total	0%	1.00%	14.00%	18.00%	5.00%
Likely	Count	0.00	7.00	518.00	737.00	194.00
	Expected Count	15728.00	32.64	526.10	708.14	173.01
	% within Likelihood	0%	0.00%	36.00%	51.00%	13.00%
	% within Credit Band	0%	8.00%	39.00%	41.00%	44.00%
	% of Total	0%	0.00%	14.00%	20.00%	5.00%
Very Likely	Count	0.00	0.00	129.00	208.00	41.00
	Expected Count	4.08	8.47	136.58	183.85	44.92
	% within Likelihood	0.00%	0.00%	34.00%	55.00%	11.00%
	% within Credit Band	0.00%	0.00%	10.00%	12.00%	9.00%
	% of Total	0.00%	0.00%	3.00%	6.00%	1.00%
Total	Count	40.00	83.00	1338.00	1801.00	440.00
	Expected Count	40.00	83.00	1338.00	1801.00	440.00
	% within Likelihood	1.00%	2.00%	36.00%	49.00%	12.00%
	% within Credit Band	100.00%	100.00%	100.00%	100.00%	100.00%
	% of Total	1.00%	2.00%	36.00%	49.00%	12.00%

Additionally, **50%** of the **725 - 760** credit band also fell within the category of **likely** to have experienced gentrification.

4.3.2 Contingency Coefficient

The second part of the chi-squared analysis was to examine if any significant association exists. However, the null and alternate hypothesis must first be stated when examining the chi-squared statistics. The null hypothesis denoted by H_0 states that credit scores are independent of gentrification. The alternate hypothesis denoted by H_a states that credit scores are not independent of gentrification. As shown in table 8, the significance value of **0.00 < 0.05**. Therefore, assuming a 95% confidence interval, we can reject the null hypothesis.

Table 8

Chi-squared Contingency Coefficient

Symmetric Measures			
		Value	Significance
Nominal by Nominal	Contingency Coefficient	0.722	0.000
N of Valid Cases		3703	

As a significant association exists, the strength of the association must also be determined. Additionally, due to the contingency coefficient value of **.722**, there is a strong association between the credit score bands and the gentrification index.

4.4 Regression

The regression analysis examined the credit score values as the dependent value, and the census variables used to create the gentrification index, as the independent variables. Although the variables were transformed using standard score & logarithm, all three regression models were used and compared.

Table 9

Regression Model Comparison

Regression Type	R	R^2	Durbin-Watson	Significance
Stepwise	0.744	0.554	1.045	0.05
Standardized Stepwise	0.727	0.528	1.011	0.00 (<0.01)
Logged & Standardized Stepwise	0.727	0.528	1.398	0.00 (<0.01)

Note. The variables included in the regression model included percentage change of the following variables: Average Household Income, Average Dwelling Value, Average Shelter Cost (owned), Average Shelter Cost (rented), Visible Minority Population, Rented households

The logged and standardized model produced an R^2 value of 0.528. Although the non-standardized regression model produced a higher R^2 value, the standardized and logged regression model was used due to the significance value and the previously mentioned issues of scale within the variables. The poor R^2 value can also be attributed to the non-linear data. After logging and standardizing the data, the Durbin-Watson value indicated positive autocorrelation. As shown with the regression model, **52.8%** of the variation within the data can be explained by the regression model. Additionally, the logged and standardized regression model only included 3 variables in total. **Average Household Income**, **Average Dwelling Value**, **Average Shelter cost for owned households**.

As shown in table 10, the regression coefficients for the average household income, average dwelling value and shelter cost for owned households indicated the largest correlated variable was dwelling value with a .660 beta value. As the beta value compares the strength of each independent variable on the dependent variable (credit scores), income was shown to have the largest strength amongst each of the variables included in the regression model.

Table 10

Regression Coefficients

Variable	*B*	β	t	p
(Constant)	721.223		751.486	0.000
Average Household Income	37.022	0.660	35.624	0.000
Average Dwelling Value	44.955	0.559	27.117	0.000
Average Shelter cost (Owned Households)	-24.691	-0.406	-15.552	0.000

4.5 Discussion

To identify if credit scores can be predictive of gentrification, credit score changes must first be examined. As shown with the credit score analysis, credit scores have changed over time across the City of Toronto. This can be partly explained by a few factors. The first of which is the changing population. Moreover, a growing population can heavily impact the ratio of accounts likely to experience a *default*. As the Office of the Superintendent of Bankruptcy reported a slight increase in credit defaults and bankruptcies, the rate of defaults and bankruptcies decreased in the past few years as shown in figure 13.

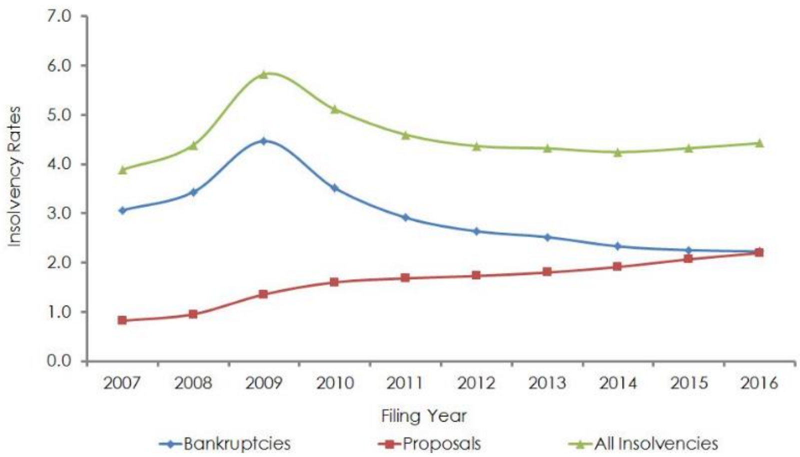


Figure 13 Changing nature of credit scores (Office of the Superintendent of Bankruptcy, 2019)

The second potential reason for these changes in credit scores is the constantly changing credit model calculations. Although credit scores can be standardized as a score between 300 - 900, the way these scores are calculated have changed over time. For example, Vantage scores or FICO scores which are two of the largest industry-standard credit score solutions have released multiple scoring models. Although the specific methodologies of these calculations are largely deemed as *black box* information, the decrease in bankruptcy rates and changing credit score models can help to explain the increasing credit scores.

The second part of the analysis was to examine if credit scores of similar or dissimilar values are likely to cluster together. The spatial autocorrelation testing with Getis-Ord GI* and the Local and Global Moran's I identified that there is a spatial correlation amongst similar values, with various outliers across the city. As shown in figures 9 & 10 the spatial autocorrelation of these values indicated that credit scores show clustering of similar values. The reason that many areas of similair credit values cluster together are largely influenced by a few factors. Crocco, Santos and Amaral (2010), determined that financial development is also largely associated to urban and socio-economic development. The socio-spatial distribution of Toronto was examined by Walks, Dinca-Panaitescu & Simone (2016) determined that large disparities of income and housing values exist in Toronto due to polarization and inequality. Additionally, this concept, along with the regression testing identified relationships **between income, dwelling value and shelter cost**. Financial well-being indicators such as credit scores are likely to follow a similar spatial distribution to closely linked variables such as income and housing value. This reflected what was outlined in literature.

The spatial autocorrelation analysis was also important to examine if these areas or clusters, can show to some extent, a correlation or association with gentrification. The created gentrification index was used to show a likelihood that specific dissemination areas have experienced gentrification. In figure 11, the index was shown to have very high values in various parts of Regent Park. Specifically, these regions can be identified by the development regions in figure 14. The City of Toronto's plan to redevelop regent part was separated into a timeline with 5 phases:

- 2007 - City Council passes Social Development Plan, developed with residents
- 2009 - Phase 2 construction begins
- 2012 - Phase 1 completed
- 2014 - Phase 3 construction begins
- 2018 - Phase 2 completed
- 2018 - RFP process begins for to select a developer partner for Phases 4 & 5
- 2021 - Phase 3 completed (estimated)

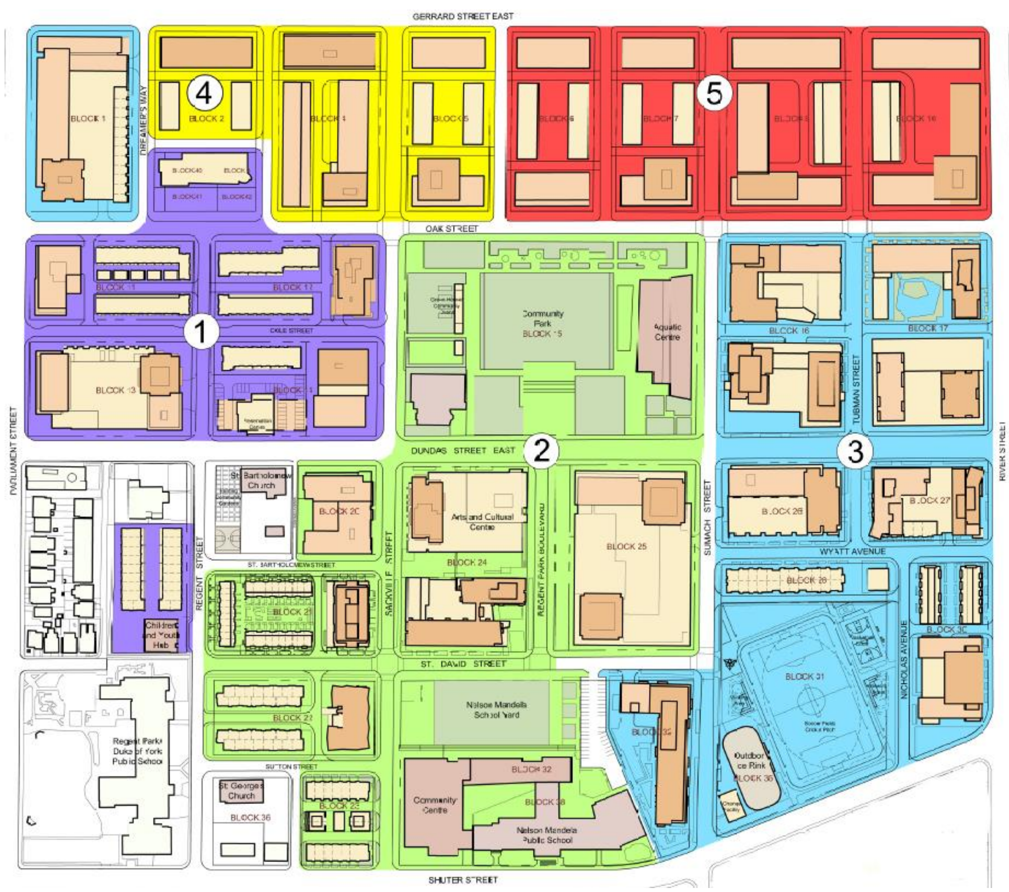


Figure 14 Regent Park development zone (Toronto Housing, 2019)

As previously mentioned, the index used census data which examined the changes between 2011 - 2016. To examine if the index can identify gentrification, the regent park area was used as this area has experienced large development within the past decade. Specifically, according to the City of Toronto, phase 1 development has finished, and the phase 2 development is close to completion. These socio-demographic changes associated with the development plan were shown through the index which identified that this area has experienced gentrification. The rising credit scores in the neighbourhood reflect the development and the changing financial well-being status of the population.

Using individual variables would be hard to identify such composite changes, but it would also be hard to discern if these areas simply experienced a demographic or socio-economic change or experienced gentrification. Therefore, the index created was a surrogate indicator of gentrification and compared to the credit scores. The first thing to note is that, inherently, financial well-being indicators, to some extent are correlated. As the chi-squared analysis revealed a strong association between the credit scores and gentrification, this proved the hypothesis that financial well-being is strongly linked to gentrification. Moreover, the credit bands specifically indicated a statistically significant relationship with the likelihood that a DA has experienced gentrification. As previously mentioned, approximately **90% of the lowest credit band were Very Unlikely to have experienced gentrification**. Therefore, the financially vulnerable population were less likely to have experienced gentrification when compared to populations that are more financially stable.

However, creating such an index can often be problematic for inter-census years which would use population estimates as well as containing various data inconsistencies, therefore being unable to explain true percentage and raw changes. Therefore, due to the heavy association amongst the credit scores and created index, credit scores can be a large composite indicator of gentrification.

5.0 Conclusion and Future Recommendations

5.1 Additional insights

Empirical measures of gentrification often contain a mix of both economic indicators and socio-demographic data, it is important to identify if such measurements can identify known areas of development and change. As shown throughout the analysis and observations presented, it is apparent that credit scores in Toronto are changing. Specifically, changes in credit scores showed a large increase in average credit rating across the City of Toronto. This ultimately proves the first hypothesis correct and answers the first research question, that credit scores of similar values clustered together. Identifying the spatial clustering of similar credit scores helps to not only indicate the financial well-being of the population within the DAs, but also the surrounding neighbourhood. The changing financial well-being of a neighbourhood can, therefore, be representative of adverse changes in income, dwelling value and rent which are products of urban development. As mentioned throughout this report, such changes are also linked to gentrification. However, it is important to identify such changes both spatially and a-spatially. For example, it is very unlikely that areas in the Bridle Path or Rosedale regions will gentrify even though they may have a large change in income or dwelling value across specific years. As mentioned in the literature review, gentrification in this context can be shown through **rising credit scores in low to middle-income areas**. The regression testing indicated that the largest correlated variables with credit scores are income, dwelling value and shelter cost variables. The regression testing also identified that socio-economic indicators correlate with credit scores which answered the second research question. Additionally, it was also shown that financial well-being indicators such as credit scores can identify relationships and areas of gentrification. The chi-squared testing identified that there is a relationship between credit scores and gentrification. Additionally, the credit score percentage change and gentrification index both showed similar spatial distributions and specifically identified the development that has occurred in Regent Park. As Regent Park underwent large socio-demographic and financial well-being changes, as shown in the empirical measurement of gentrification, the study was able to identify the gentrification changes that occurred across the observed years. Although such changes can represent gentrification, the rising credit scores in such areas can also be indicative of the displaced **low to middle income population** that previously lived in the neighbourhood.

5.2 Limitations

A large limitation of this study is the linear data calculations and values that were used. For example, the estimated 2010 credit scores may not be representative of true credit score values in 2010. Credit score methodologies also change frequently. The number of accounts used to calculate credit scores are different across each year as credit bureaus use different information from various companies. Therefore, the specific calculations and methodologies for credit scores are different across each year. Additionally, gentrification is typically hard to measure unless the analysis conducted uses known areas of development. For example, the Regent Park area was used to compare the created index to a known area of development and gentrification. As standardized measurements of gentrification do not exist, empirical measurements of gentrification are created through theoretical and analytical frameworks explained in literature.

5.3 Future Recommendations

As the gentrification index that was created in this report implies equal importance of each of the socio-demographic variables used, varying weighting schemes should be used to better measure the importance of specific factors on gentrification. Although the literature does not indicate a consensus among the variables that are most important when investigating gentrification, varying weighting schemes may be better suited to measure such changes.

Additionally, traditional methods of variable weighting are often manual and typically defined through theoretical frameworks and guided by relevant literature. However, deterministic methods of weighting are also potential alternatives. Some of these methods include the Rank-Exponent, Rank-Reciprocal and Rank-Exponent methods which are common within multi-criteria and decisions analysis (**MCDA**) models.

Literature has also defined the potential use of social development variables or wellbeing indicators to be included in empirical measures of gentrification. However, very few research and empirical analyses exist with such frameworks. Specifically, changes in education or unemployment rate can also be used to identify large changes in social development or class structure.

This report has also examined that credit scores of similar values are likely to cluster together. However, future research should examine if credit score changes are likely to cluster together. As the research in this report was able to identify that credit scores can be predictive of gentrification, large clusters of increases in credit scores in middle to low-income neighbourhoods would, therefore, be representative of gentrification. Additionally, future research or analyses should be used to evaluate the gentrification index used in this research. Specifically, it should be noted if areas that were not likely to gentrify according to the index, did or did not experience gentrification.

Literature has specified in using PCA or discriminant analysis to identify if independent variables were likely to predict if areas were gentrified. Additionally, Pearson's correlation analysis could also be conducted to further evaluate the correlation between the index representing gentrification and the credit scores or percentage change in credit scores.

5.4 Concluding statement

As the series of results and observations listed in this research identify that credit scores are changing. Furthermore, the relationship of credit scores and gentrification indicators such as the index used in this report identified a strong relationship between financial well-being indicators and areas of gentrification. Additionally, the statistical and spatial analysis identified areas of gentrification. Such research can also be used to identify population displacement of the vulnerable. However, the extent to which population displacement can be identified by credit scores has yet to be measured. Rather, identifying gentrifying neighbourhoods through credit scores and credit score changes indicated that financial well-being is closely linked to urban development and the changing structure of the built environment. The socio-economic impact of gentrification can be identified when examining specific bands of credit scores or credit score changes.

References

Arya, S., Eckel, C., & Wichman, C. (2013). Anatomy of the credit score. *Journal of Economic Behavior and Organization, 95*, 175-185. doi:10.1016/j.jebo.2011.05.005

Atkinson, R., & Bridge, G. (2005). *Gentrification in a global context: The new urban colonialism* Routledge.

Atuesta, L. H., & Hewings, G. J. D. (2019). Housing appreciation patterns in low-income neighborhoods: Exploring gentrification in chicago. *Journal of Housing Economics, 44*, 35-47. doi:10.1016/j.jhe.2018.08.005

August, M. (2014). Challenging the Rhetoric of Stigmatization: The Benefits of Concentrated Poverty in Toronto’s Regent Park. Environment and Planning A: Economy and Space, 46(6), 1317–1333. doi:10.1068/a45635

Blaikie, N. (2003). *Analyzing quantitative data: From description to explanation*. London: SAGE Publications.

Brüggen, E. C., Hogreve, J., Holmlund, M., Kabadayi, S., & Löfgren, M. (2017). Financial well-being: A conceptualization and research agenda. *Journal of Business Research, 79*, 228-237. doi:10.1016/j.jbusres.2017.03.013

Carrns, A. (2012). Is That Credit Score a FICO, or a FICO 8? Retrieved from https://bucks.blogs.nytimes.com/2012/05/10/is-that-credit-score-a-fico-or-a-fico-8/

Charles, C. (2000). Neighborhood racial-composition preferences: Evidence from multiethnic metropolis. Social Problems 47(3), 379-407.

Colburn, L. L., & Jepson, M. (2012). Social indicators of gentrification pressure in fishing communities: A context for social impact assessment. *Coastal Management, 40*(3), 289-300. doi:10.1080/08920753.2012.677635

Credit Karma. (2015). Credit Trends. Retrieved from https://www.creditkarma.com/trends/age

Crocco, M., Santos, F., & Amaral, P. V. (2010). The Spatial Structure of Financial Development in Brazil. *Spatial Economic Analysis*, *5*(2), 181-203. doi:10.1080/17421770903511973

Dean, L. T., ScD., & Nicholas, Lauren Hersch,PhD., M.P.P. (2018). Using credit scores to understand predictors and consequences of disease. *American Journal of Public Health*, *108*(11), 1503-1505. doi:http://dx.doi.org.ezproxy.lib.ryerson.ca/10.2105/AJPH.2018.304705

Ding, L., Hwang, J., & Divringi, E. (2016). Gentrification and residential mobility in philadelphia. *Regional Science and Urban Economics*, *61*, 38-51. doi:10.1016/j.regsciurbeco.2016.09.004

Dwyer, G. P., Tabak, B. M., & Vilmunen, J. (2012). The financial crisis of 2008, credit markets and effects on developed and emerging economies. *Journal of Financial Stability*, *8*(3), 135-137. doi:10.1016/j.jfs.2012.03.001

Eldaidamony, M., & Shetawy, A. (2016). Gentrification indicators in the historic city of cairo.*Procedia - Social and Behavioral Sciences*, *225*, 107-118. doi:10.1016/j.sbspro.2016.06.013

ESRI. (2019[a]). Data classification methods. Retrieved from https://pro.arcgis.com/en/pro-app/help/mapping/layer-properties/data-classification-methods.htm

ESRI. (2019[b]). Spatial Autocorrelation (Global Moran's I). Retrieved from https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/spatial-autocorrelation.htm

ESRI. (2019[c]). How Hot Spot Analysis (Getis-Ord Gi*) works. Retrieved from https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/h-how-hot-spot-analysis-getis-ord-gi-spatial-stati.htm

ESRI. (2019[d]). Global Moran's I & Spatial Autocorrelation. Retrieved from http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/h-how-spatial-autocorrelation-moran-s-i-spatial-st.htm

ESRI. (2019[e]). How Cluster and Outlier Analysis (Anselin Local Moran's I) works. Retrieved from https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/h-how-cluster-and-outlier-analysis-anselin-local-m.htm

ESRI. (2019[f]). Modeling spatial relationships. Retrieved from https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/modeling-spatial-relationships.htm

Equifax. (2019). What is a Good Credit Score? Retrieved from https://www.consumer.equifax.ca/personal/education/credit-score/what-is-a-good-credit-score/

Frank, A. (2017). *What does it take to see gentrification before it happens?*. Washington: NPR.

Feldkircher, M. (2014). The determinants of vulnerability to the global financial crisis 2008 to 2009: Credit growth and other sources of risk. *Journal of International Money and Finance, 43*, 19-49. doi:10.1016/j.jimonfin.2013.12.003

Freeman, L. (2005). Displacement or succession?: Residential mobility in gentrifying neighborhoods. *Urban Affairs Review, 40*(4), 463-491. doi:10.1177/1078087404273341

Galster, G (2001) On the nature of neighbourhood. Urban Studies 38(12): 2111–2124.

Galster, G., & Peacock, S. (1986). Urban gentrification: Evaluating alternative indicators. *Social Indicators Research, 18*(3), 321.

Glass, R. (1964). *London: Aspects of change*. London: MacGibbon & Kee.

GIS & Geography. (2018, February 16). Spatial Autocorrelation and Moran's I in GIS. Retrieved from https://gisgeography.com/spatial-autocorrelation-moran-i-gis/

Immergluck, D. (2009). Large redevelopment initiatives, housing values and gentrification: The case of the atlanta beltline. *Urban Studies, 46*(8), 1723-1745. doi:10.1177/0042098009105500

Israel, S., Caspi, A., Belsky, D. W., Harrington, H., Hogan, S., Houts, R., . . . Moffitt, T. E. (2014). Credit scores, cardiovascular disease risk, and human capital. *Proceedings of the National Academy of Sciences, 111*(48), 17087-17092. doi:10.1073/pnas.1409794111

Johnson, L. C., & Schippling, R. (2009). *Regent Park revitalization: Young peoples experience of relocation from public housing redevelopment*. Ottawa, Ont.: Canada Mortgage and Housing Corporation.

Kelly, S. (2013). The New Normal: The Figure of the Condo Owner in Torontos Regent Park. *City & Society, 25*(2), 173-194. doi:10.1111/ciso.12015

Knapp, E. A., & Dean, L. T. (2018). Consumer credit scores as a novel tool for identifying health in urban U.S. neighborhoods. *Annals of Epidemiology, 28*(10), 724-729. doi:10.1016/j.annepidem.2018.07.013

Lees, L. (2008). Gentrification and social mixing: Towards an inclusive urban renaissance?*Urban Studies, 45*(12), 2449-2470. doi:10.1177/0042098008097099

Leventhal, B. (2016). *Geodemographics for marketers using location analysis for research and marketing*. Kogan Page.

MacLaran, A. (2003). *Making space: Property development and urban planning* Arnold.

Martin, I. W., & Beck, K. (2018). Gentrification, Property Tax Limitation, and Displacement. Urban Affairs Review, 54(1), 33–73. https://doi.org/10.1177/1078087416666959

Pattillo, M. E. (2008). *Black on the block: The politics of race and class in the city* (1st pbk. ed.) University of Chicago Press.

Phillips, M. (2004). Other geographies of gentrification. *Progress in Human Geography*, *28*(1), 5-30. doi:10.1191/0309132504ph458oa

Qian, J., He, S., & Liu, L. (2013). Aestheticisation, rent-seeking, and rural gentrification amidst china's rapid urbanisation: The case of xiaozhou village, guangzhou. *Journal of Rural Studies, 32,* 331-345. doi:10.1016/j.jrurstud.2013.08.002

Róka-Madarász, L., & Mályusz, L. (2013). evaluating indicators for renewal of properties via gentrification in budapest. *Open House International, 38*(2), 66.

Reades, J., De Souza, J., & Hubbard, P. (2019). Understanding urban gentrification through machine learning. *Urban Studies, 56*(5), 922-942. doi:10.1177/0042098018789054

Shazmeen, S. F., Mirza Mustafa Ali Baig, & Pawar, M. R. (2013). Regression analysis and statistical approach on socio-economic data. *International Journal of Advanced Computer Research*, *3*(3), 347.

Sheng, J., Han, X., & Zhou, H. (2017). Spatially varying patterns of afforestation/reforestation and socio-economic factors in china: A geographically weighted regression approach. *Journal of Cleaner Production, 153,* 362-371. doi:10.1016/j.jclepro.2016.06.055

Soros, G. (2008). *The new paradigm for financial markets: The credit crisis of 2008 and what it means* (1st ed.) PublicAffairs.

Sung, K. K. (2015). "hella ghetto!" (dis)locating race and class consciousness in youth discourses of ghetto spaces, subjects and schools. *Race Ethnicity and Education*, *18*(3), 363-395.

Toronto Housing. (2019). Regent Park Development Project. Retrieved from https://www.torontohousing.ca/regentpark

Walker, L. A., Ince, D., Riphenburg-Reese, A., & Littman, J. (2018). Predicting Resident Involvement in Neighborhood and Voluntary Groups in a Latino/Latina Neighborhood

Walks, A., Dinca-Panaitescu, M., & Simone, D. (2016). Income Inequality and Polarization, City of Toronto & York Region. Retrieved from http://neighbourhoodchange.ca/homepage/income-inequality-polarization-toronto-york-region-levels-trends/

Westover, L. (2013). Improving financial wellbeing. *Charter*, *84*(8), 48.Undergoing Transit-Oriented Development and Gentrification. Nonprofit and Voluntary Sector Quarterly, 47(6), 1135–1154. https://doi.org/10.1177/0899764018783093