Examining the Correlation between Neighbourhood Property Crime Rates and Social Factors in Toronto, ON

Ashmitha Bala *400148743*
Cassis Lindsay *400194383*
CJ Murphy *400195804*
Hayat Shariff *400258691*
Zoe Spronk *400231466*

This paper reports our analysis of property crime rates in Toronto, ON and their correlation with population density, income, race, and age by neighbourhood. Data were retrieved from The City of Toronto Open Data Portal.

Keywords: crime, social factors, spatial analysis

Introduction

In most cities worldwide, property crime is listed as the most common criminal offence. (Thatcher 2015). Larceny consisting of burglary, auto theft, and other forms of theft are seen the most frequently when gauging property crime. The conflict theory denotes that crime as an offence is a result of various types of material inequality, often governed by disparities within different socioeconomic strata. (Nickerson 2019). Continually, the social disorganization theory is derived from the idea that physical signs of neighbourhood disorder such as broken windows, inadequate lighting, and vacant housing, play a role in the presence of crime within an area. (He, Páez, and Liu 2017). Both theories highlight the link of socioeconomic factors and disparitie and the presence of crime. The aim of this study is to examine the following research question: What is the relative risk of property crime in various Toronto, ON neighbourhoods, and how strong is this correlation when compared to socioeconomic factors such as, income, age, race, and population density. Open datasets obtained from the City of Toronto as well as The Toronto Open Data Portal will be visualized using a combination of regression modelling, Moran's I tests, and spatial data visualizations will be used to help quantify the above question.

Background

Within Toronto, the gap between economic classes has been at a steady incline. Between the years of 1980 to 2005, income inequality within the city had risen nearly 31% percent, the largest increase in any Canadian major city. (Hulchanski 2011). The distinct economic divide found within the city is just one of the reasons Toronto has been nicknamed "the city of neighbourhoods" (Van Ingen, Khandor, and Fleiszer 2015), the well defined social, and economic variability amongst Toronto's neighbourhoods are polarizing realities from the outside looking in. However, this differentiability between neighbourhood hints to further correlations between income and additional social factors. In a study done by David Hulchanski at the University of Toronto, 2016 census data was used to investigate the ineluctable link between income polarization, and racial demographics in Toronto. Findings from the study modelled that 48% of Toronto's census tracts were considered low income with annual average income being $32,000 before taxes, additionally, it was found that 68% of individuals living in these low-income tracts were part of a racialized group, or other visible minority. (Hulchanski 2011).

Over the past few decades, theories of crime have gone through extensive development to gain valuable insight on the constituents of criminal offences and to better understand what might influence individuals of rather normal biological, and psychological nature to commit various crimes. (Longley 2020). Most notably, social disorganization theory, and conflict theory have highlighted the association of socioeconomic factors, like, race, income, and physical environment, to crime rates. These theories back up the notion that the presence of crime is not evenly distributed spatially, but rather, varies as a result of external factors.

The basis of the social disorganization theory suggests that crime is a result of one's physical and social environment. (Bond 2015). A case study conducted in Chicago, Illinois by Robert Park and Ernest Burgess of the Chicago School of Criminology , found that crime rates were unevenly distributed around the city, with a heavy concentration of crime occurring in and around the inner-city. These inner city neighbourhoods were also found to be transitional areas that had generally low socio-economic status, and a large racialized population. In concluding the study, Park and Burgess suggested that the increase rates of crime in these pockets of the city were not a result of "personal attributes" of the occupants of these neighbourhoods, rather the "the structural factors of poverty, high heterogeneity, and high mobility causing what was dubbed "social disorganization." (Kitchen 2006). In understanding race, income, and gender as potential drivers of high crime rates within certain areas of the city, there allows space for the argument of elevated presence of law enforcement and over policing of these boroughs. The duality of this topic allows for multiple vantage points to be investigated implementing varying theories to support.

Study area

The relationship between property crime rates and social factors were studied across the divisional neighbourhoods of Toronto, ON, Canada (see Figure 1). There are 140 neigbourhoods in total at the time of this study. This area was selected due to its noted polarization between neighbourhood demographics (Van Ingen, Khandor, and Fleiszer 2015), and public availability of datasets.

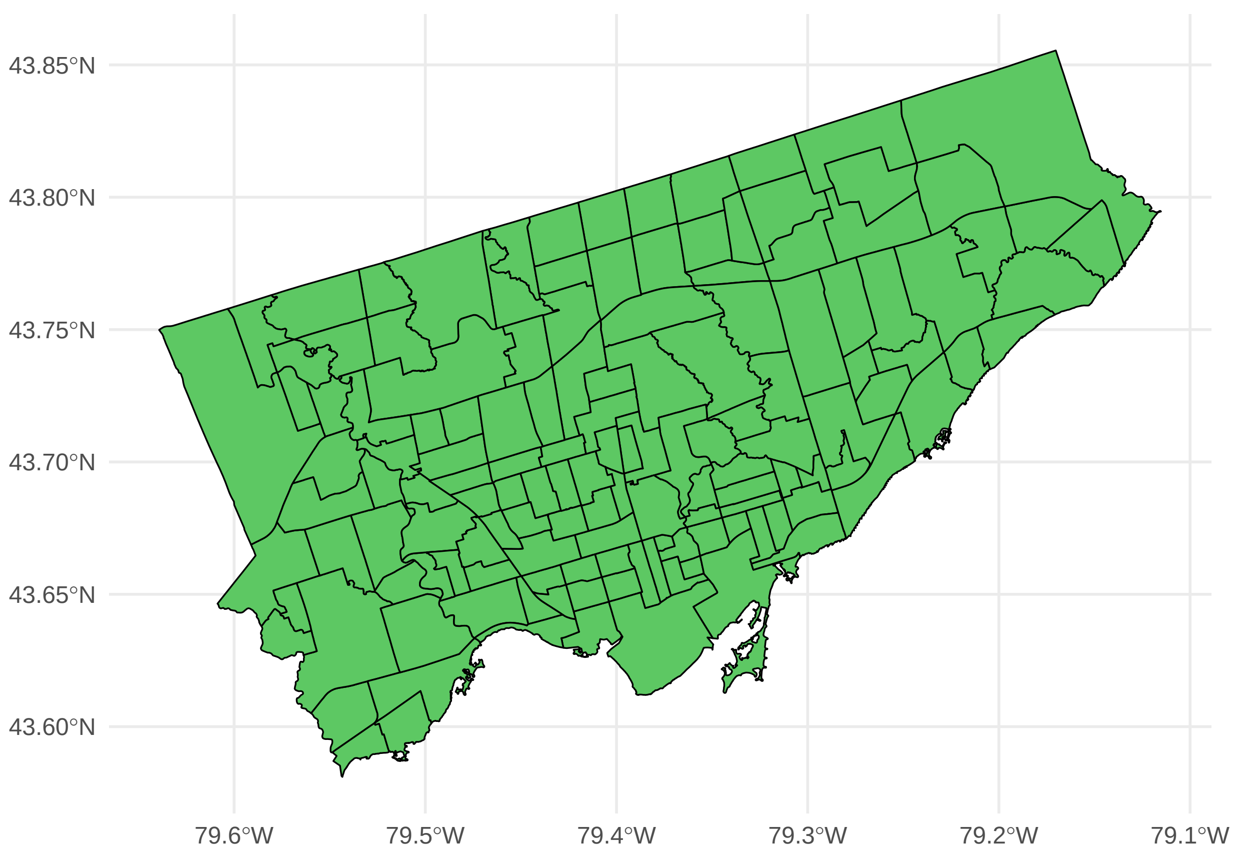

Figure 1: Census Neighbourhoods of Toronto, ON, Canada

Data

The data used in this study comes from the City of Toronto Open Data Portal, accessed using the **opendataportal** package in R. (Gelfand 2020). Two separate datasets were used: *Neighbourhood Crime Rates* and *Neighbourhood Profiles*. The *Neighbourhood Crime Rates* dataset reports both the raw count and rates per 100,000 of seven different types of crime within the 140 neighbourhoods of Toronto. (Toronto Police Service 2019). Of the seven types of crime, the four reported types of property crime — auto theft, breaking and entering, robbery, and theft over $5000 — were the focus of this study. The *Neighbourhood Profiles* dataset reports the demographic, social, and economic findings of the 2016 census for each of Toronto's 140 neighbourhoods. (Statistics Canada 2021). Select data regarding population density, income, race, and age were utilised as the social factors of interest for this study.

Methods

This study was conducted using the open software language **R** to carry out statistical analyses of crime rates and social factors in the city of Toronto, ON, and visualise these results through the production of maps and figures. Property crime rates by neighbourhood in 2016 were used as the independent variable. This was compared against 4 other variables from 2016 census data at the neighbourhood level: population density, percentage of the population designated as 'low income,' percentage of the population belonging to a visible minority group, and percentage of the population aged 15-29. The areal data of this study was analysed using choropleth mapping to visualize geographic trends, boolean maps relative to a selected values, scatterplots, and regression analyses in order to evaluate the relationship between crime and selected social factors.

Results

First the data of interest was visualised through choropleth maps to gain a general view of any potential trends and create a starting point for further analysis. Figure 2 illustrates the rates of auto theft, breaking and entering, robbery, and theft over $5000 in Toronto's neighbourhoods in 2016, and figure 3 shows the aggregated rate of these 4 types of property crimes.

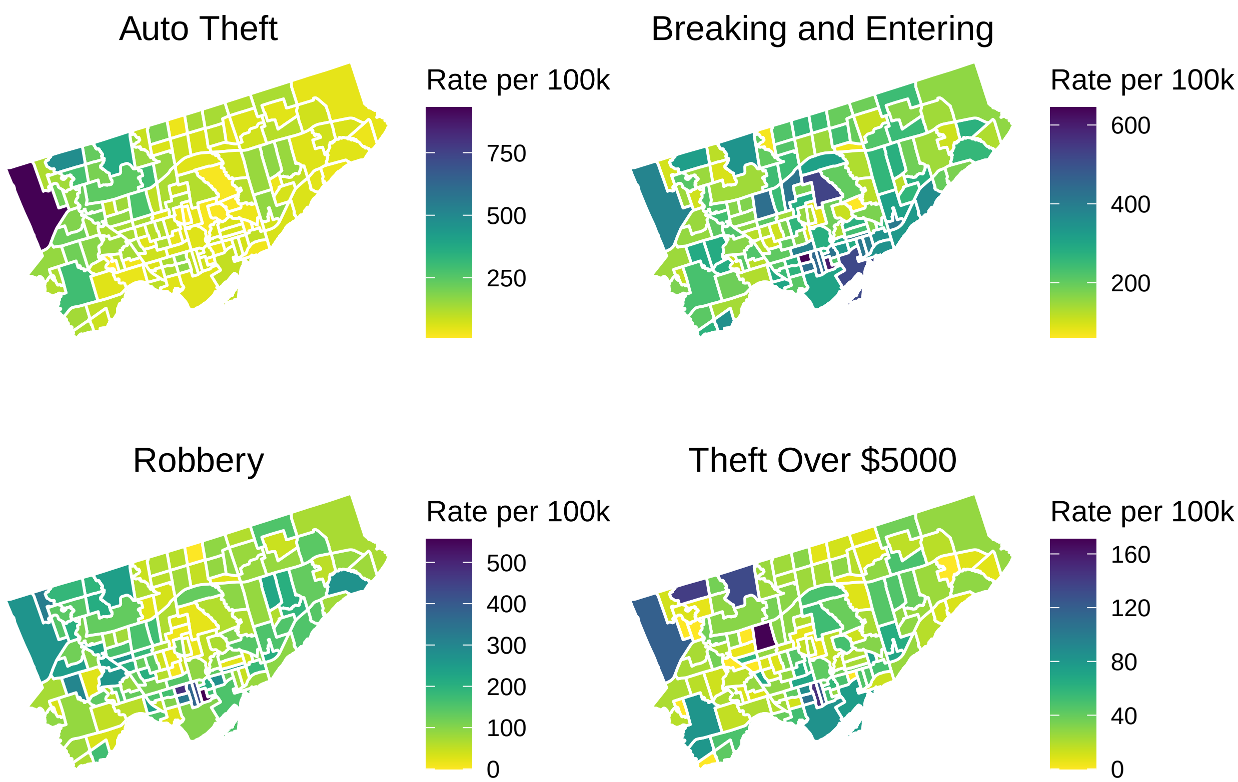

Figure 2: Four categories of property crime rates by neighbourhood (2016)

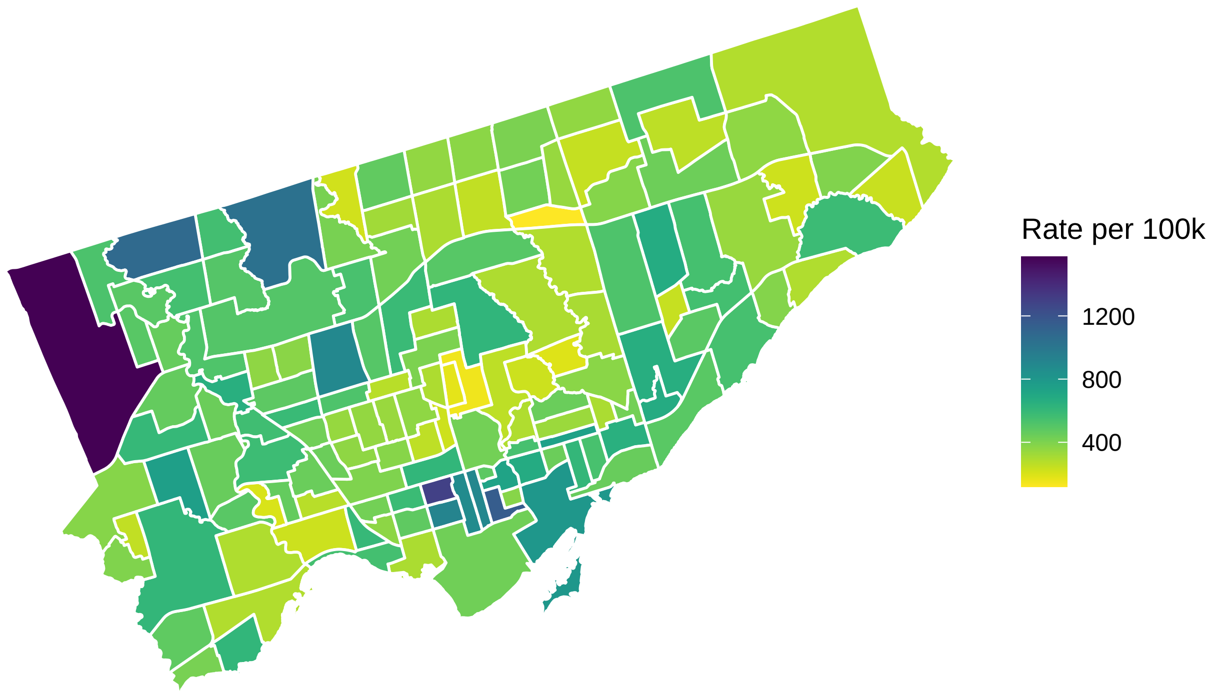

Figure 3: Aggregated property crime rates by neighbourhood (2016)

Figure 4 through figure 7 are choropleth maps denoting the social factors of interest for this study in 2016. Figure 4 denotes population density by neighbourhood. Figure 5 shows the percentage of individuals in each neighbourhood who are categorized as low income after taxation. Figure 6 illustrates the percentage of individuals in each neighbourhood who belong to a visible minority group. Finally, Figure 7 illustrates the percentage of individuals in each neighbourhood who are aged 15-29.

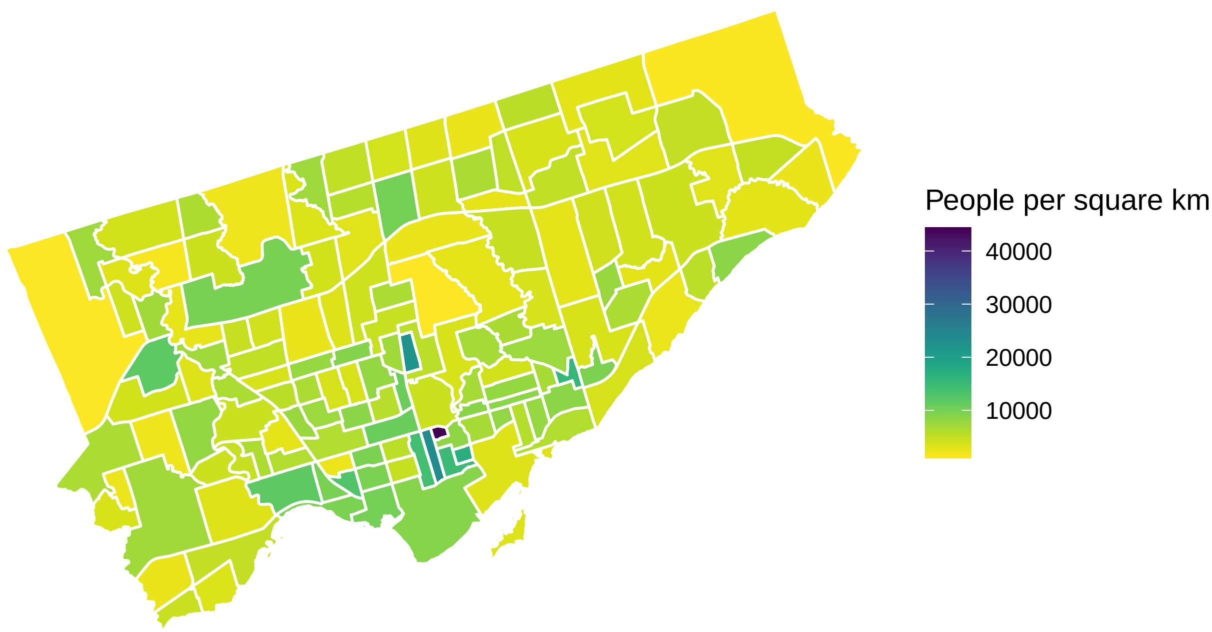

Figure 4: Population density by neighbourhood (2016)

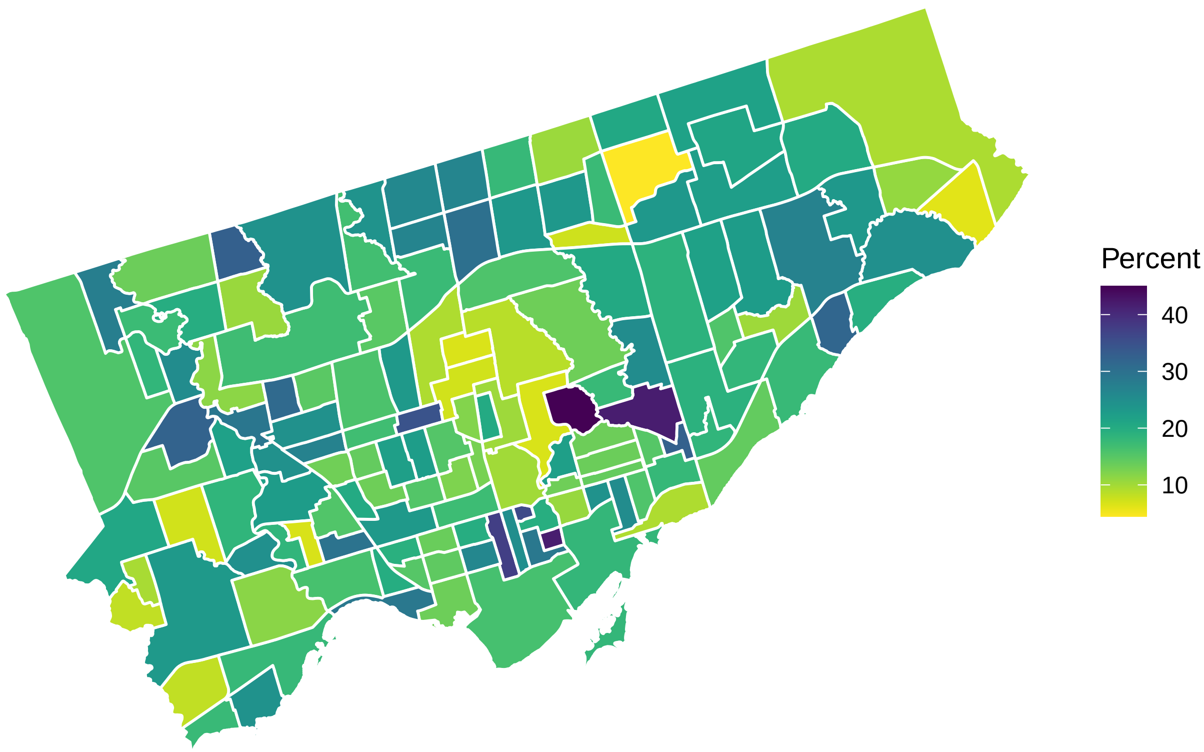

Figure 5: Percentage of inhabitants who have been identified as 'Low Income After Taxation' by neighbourhood (2016)

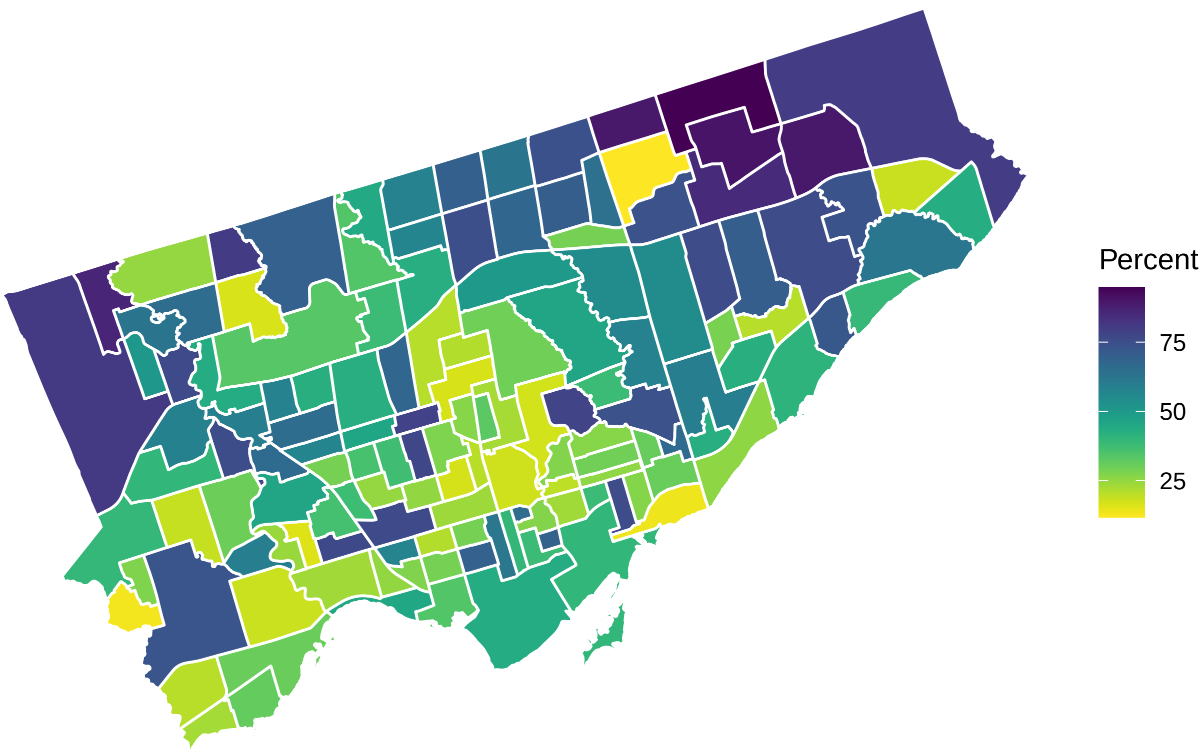

Figure 6: Percentage of inhabitants who belong to visible minority groups by neighbourhood (2016)

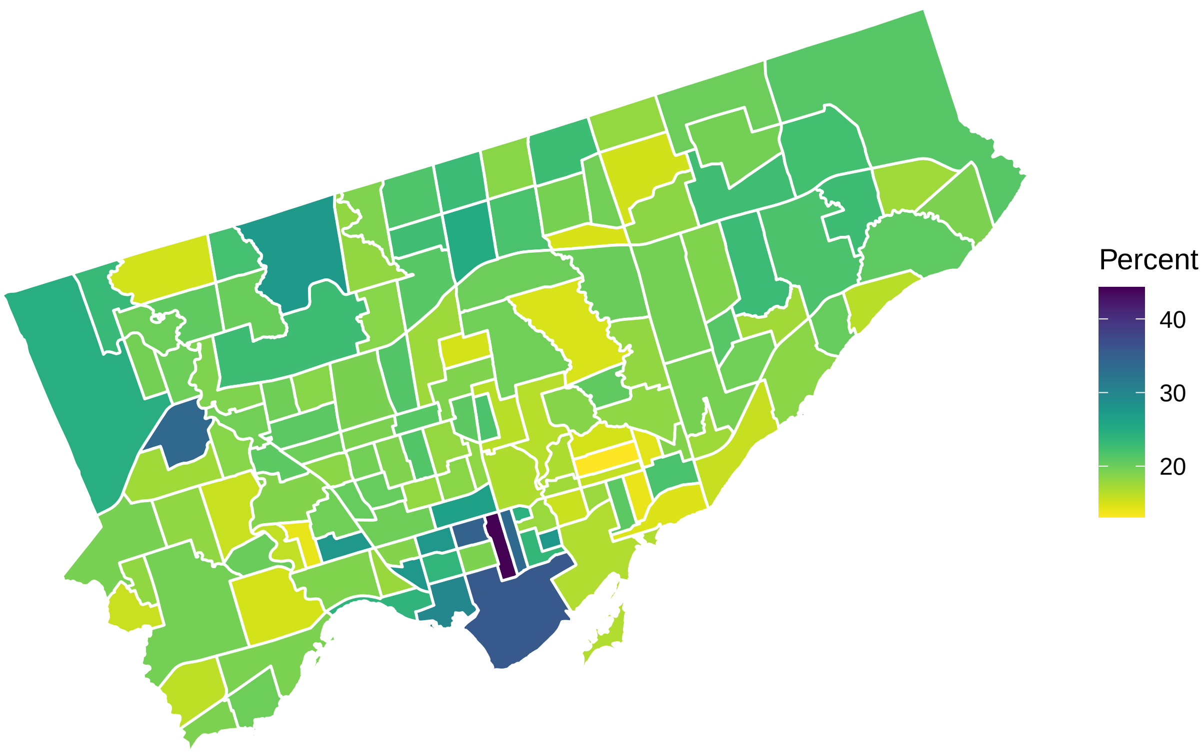

Figure 7: Percentage of inhabitants aged 15-29 by neighbourhood (2016)

Figure 8 is a boolean map indicating which neighbourhoods have property crime rates that are above and below the citywide mean in 2016. This is useful in illustrating potential trends in the data with high-crime or low-crime neighbourhoods being clustered in some form.

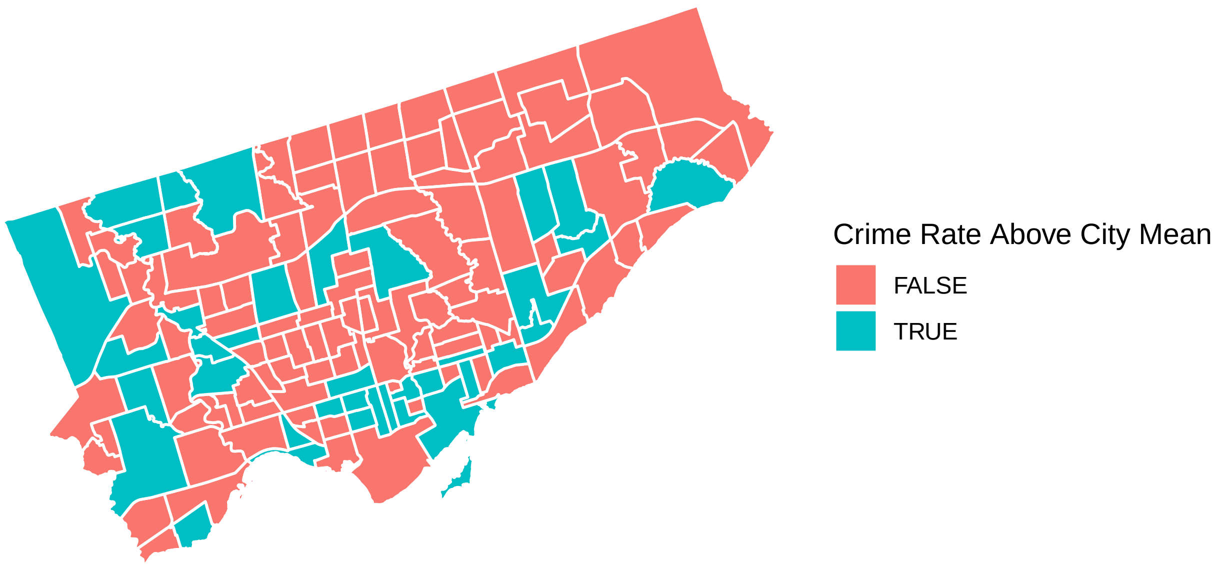

Figure 8: Neighbourhood crime rates relative to mean crime rate of the whole of Toronto (2016)

Figure 9 is a Moran scatterplot for neighbourhood property crime rate in 2016. The value of Moran's I for this dataset is 0.246 with a p-value of 6.197e-08. With these statistics, we can reject the null hypothesis of spatial independence, as Moran's I is large enough to indicate some spatial autocorrelation, and the p-value is small enough to indicate a high degree of confidence.

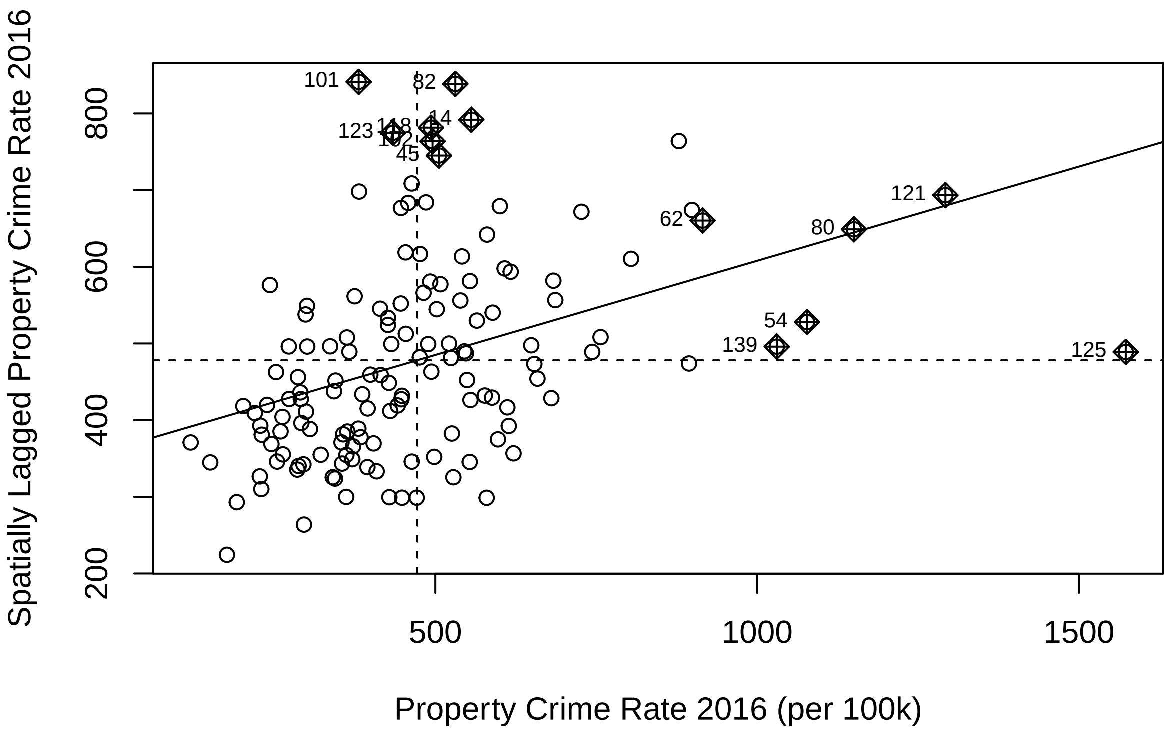

Figure 9: Moran's plot for spatial autocorrelation regarding neighbourhood property crime rates (2016)

Table 1: Property crime rate regressed on population density

	Dependent variable:
	propertyCrime2016
population_density	0.0004
	(0.004)
Constant	469.292***
	(30.192)
Observations	140
R^2	0.0001
Adjusted R^2	−0.007
Residual Std. Error	218.002 (df = 138)
F Statistic	0.011 (df = 1; 138)
Note:	*p<0.1; **p<0.05; ***p<0.01

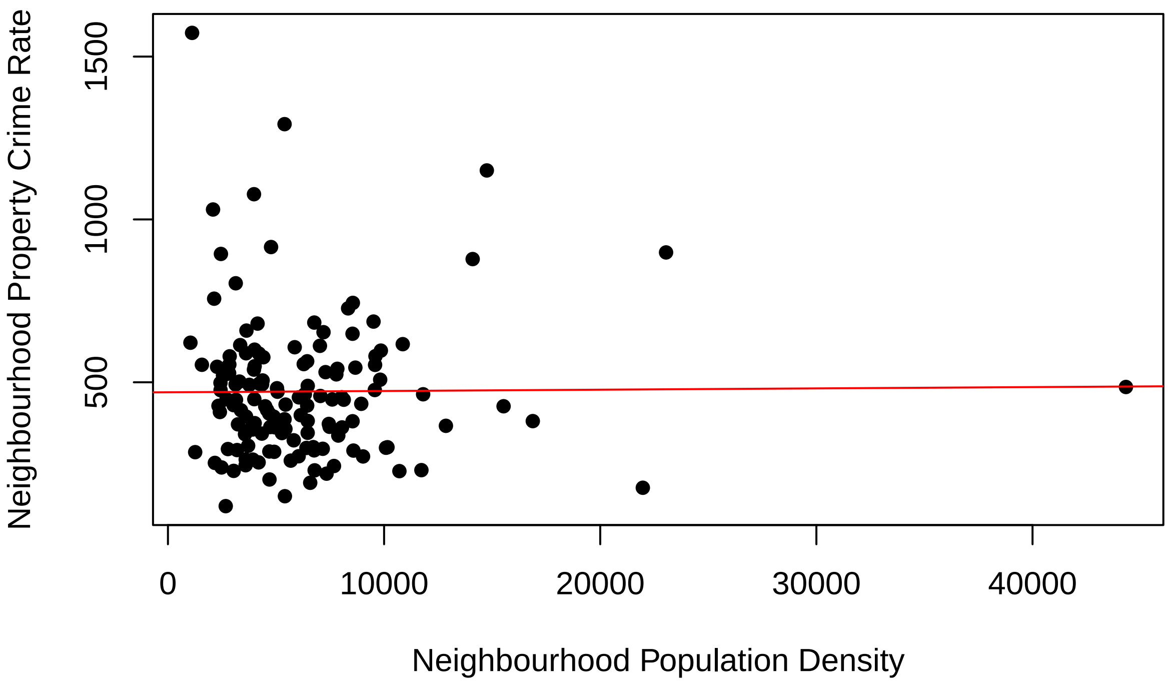

Figure 10: Regression model comparing population density and property crime rates by neighbourhood (2016)

Table 2: Property crime rate regressed on percentage of individuals designated as 'low income'

	Dependent variable:
	propertyCrime2016
low_income_percent	2.307
	(2.379)
Constant	427.491***
	(49.239)
Observations	140
R^2	0.007
Adjusted R^2	−0.0004
Residual Std. Error	217.272 (df = 138)
F Statistic	0.940 (df = 1; 138)
Note:	*p<0.1; **p<0.05; ***p<0.01

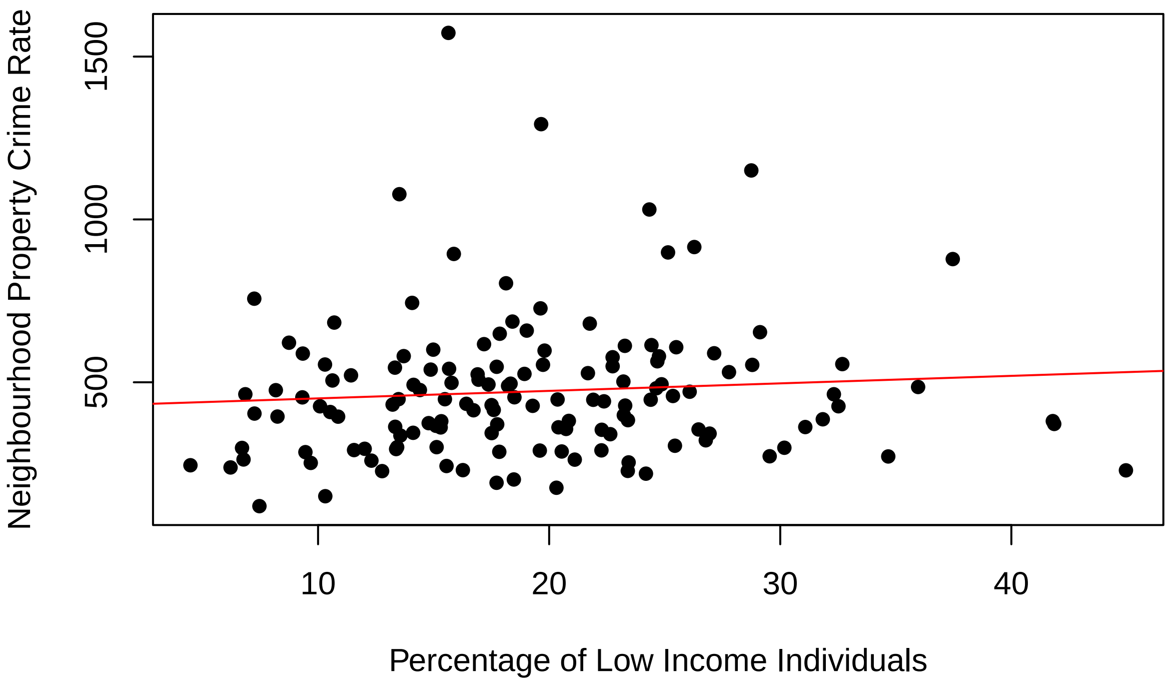

Figure 11: Regression model comparing percentage of low income individuals and property crime rates by neighbourhood (2016)

Table 3: Property crime rate regressed on percentage of individuals who are a visible minority

	Dependent variable:
	propertyCrime2016
visible_minority_percent	0.481
	(0.841)
Constant	449.410***
	(43.203)
Observations	140
R^2	0.002
Adjusted R^2	−0.005
Residual Std. Error	217.752 (df = 138)
F Statistic	0.328 (df = 1; 138)
Note:	*p<0.1; **p<0.05; ***p<0.01

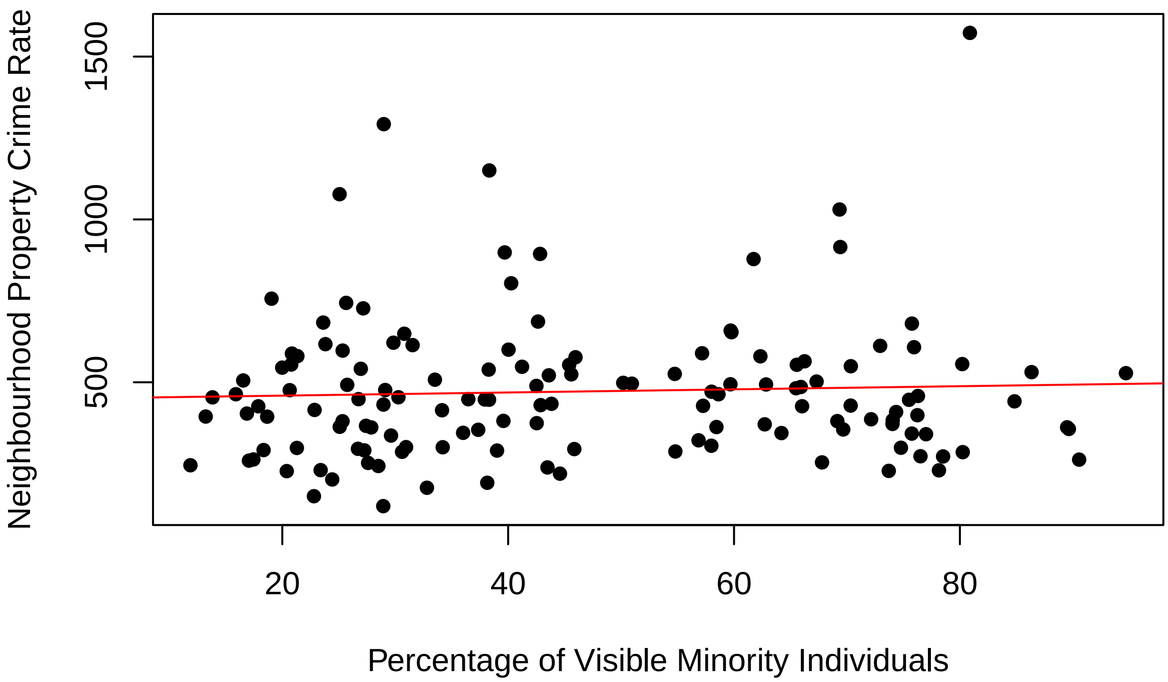

Figure 12: Regression model comparing percentage of visible minority individuals and property crime rates by neighbourhood (2016)

Table 4: Property crime rate regressed on percentage of individuals who are aged 15-29'

	Dependent variable:
	propertyCrime2016
percent_aged_15to29	14.014***
	(3.997)
Constant	188.206**
	(82.790)
Observations	140
R^2	0.082
Adjusted R^2	0.075
Residual Std. Error	208.905 (df = 138)
F Statistic	12.292*** (df = 1; 138)
Note:	*p<0.1; **p<0.05; ***p<0.01

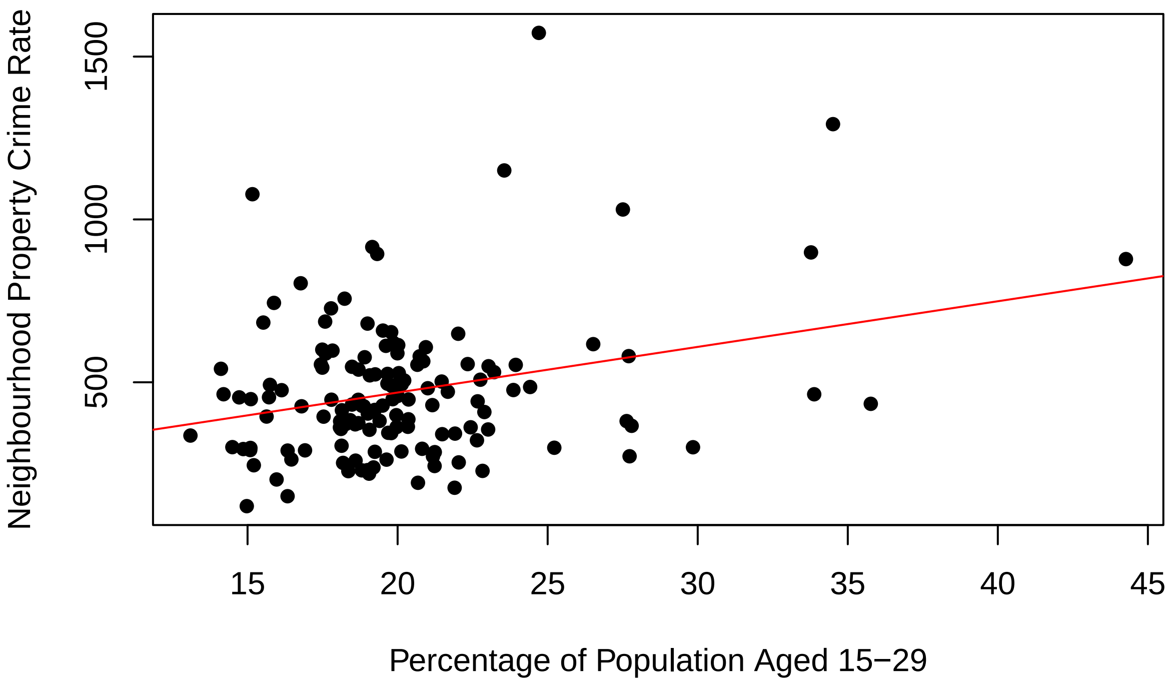

Figure 13: Regression model comparing percentage of young (aged 15-29) individuals and property crime rates by neighbourhood (2016)

Table 1 thorough table 4 summarize the determined regression models of property crime rate against the four social factors of interest, and figure 10 through figure 13 display these regression models in graphical form. These were used to quantify any potential relationship between property crime and select social factors. Table 1 and figure 10 show population density regressed against property crime rate. Table 2 and figure 11 show percentage of low income individuals regressed against property crime rate. Table 3 and figure 12 show percentage of visible minority individuals regressed against property crime rate. Table 4 and figure 13 show percentage of individuals aged 15-29 regressed against property crime rate.

Analysis

The individualised choropleth maps show 'auto theft' increasing towards the left of the map, demonstrating the most obvious pattern of all four types of property crime. The auto theft rates also seem to show a larger proportion of yellow neighbourhoods, or those with rates of auto theft under 250 per 100 thousand people, and less of a gradual distribution of higher values (figure 2). Both 'breaking and entering' and 'robbery' show similar distribution, with higher values focused around the downtown/south-central neighbourhoods. On the other hand, the 'theft over 5000$' and 'breaking and entering' are similar in the sense that they both have distributions that are alike and at first glance show a rather random (at least to our current grasp of knowledge) juxtaposition of low- to mid-range values throughout. In general, for all four maps in Figure 2, there seems to be a trend showing higher rates in the top left corner, and the small downtown neighbourhoods in south-central (with the exception of auto theft). The choropleth map in Figure 3 illustrating aggregate rates of total crime in 2016 support my prior statement, showing the higher rates of crime, or the purple neighbourhoods, in the top-left block and a cluster of higher values in the south-central area. There are the most eye-catching areas. The lowest crime rates, the light-yellow colours, for all property crime are found in the centre of Toronto. This information is fairly vague; we have not yet compared our findings to any of the four variables we identified earlier, which include population density, percentage of population designated as 'low income,' percentage of the population belonging to a visible minority group, and the percentage of population aged 15-29.

The choropleth map denoting population density in Figure 4 is rather easy to interpret. There are a few clear neighbourhoods in the south-central/downtown Toronto that are the dark purple and dark blue colours that represent 20,000 people per square kilometre. The scale ranges from 10,000 to 40,000 people per km^2 but three quarters of these people are found in those few neighbourhoods, indicating that they are significantly more populated. As seen in our aggregated crime rate map (Figure 3), this same location has been shown to experience high rates of crime, approximately above 1000 events per 100 thousand people. In the individual choropleth maps (Figure 2), we can see that most of these exact neighbourhoods also show high rates of robbery, theft over $5000, and breaking and entering. This suggests a relationship between population density and robbery, theft over $5000, and breaking and entering.

The percentage of inhabitants who have been identified as 'low income after taxation' in Figure 5 shows a bit more of a distribution. However, those same few neighbourhoods that also showed high population density and higher general rates of crime also number amongst some of the highest percentage of low-income inhabitants. That being said, the highest percent neighbourhood is outside of this little downtown/south central region. The highest-ranking neighbourhood for this variable is not shown distinctly on any of the property crime choropleths. This suggests that there may be some relationship between low-income inhabitants and some crime like breaking and entering, robbery and theft over $5000 (see figure 2), but it could also be a situation where high population density leads to an inevitable population of low-income residents that happen to be in proximity to crime. This is suggested because the highest percent of low-income inhabitants occurs in a neighbourhood that has low rates of all types of property crime.

When interpreting the Boolean map in Figure 8, we can see that the areas identified as higher rates of the types of property crime on the choropleths also show crime rates that are above the city mean. Some of these neighbourhoods are the ones that also host larger population density and percentage of low-income inhabitants.

The Moran's I plot in Figure 9 indicates that the property crime values are spatially independent, with more points on the quadrants of high-low and low-high quadrants. This suggests that the neighbourhoods with high and low concentrations of property crime rates are not located nearby each other but in a fairly random manner. Except for the lower left quadrant which is fairly clustered indicating the low property crime rate, neighbourhoods are surrounded by neighbourhoods with low crime rate as observed in Figure 8. Next, we will compare our four variables through regression analysis. The 'population density' estimated standard is 469.292 compared to the variance which is 30.192 (Table 1). Similarly, the 'low income' estimated standard of 427.491 and variance of 49.239 (Table 2). Since the estimated standard and the variance do not vary largely for both the independent variable against the property crime rate it is possible to see no strong correlation to the low income variable. This can be visually observed in the regression model graph in Figure 10 for population density as the regression line on the graph is almost zero, indicating that there is little to no relationship between the two variables. The line from Figure 11 has a slope, indicating some relationship. The p-value for population density was 0.9188 (Table 1). This is not a significant value. Therefore, we can say that the residuals show no correlation because the p-value is rather high. We cannot be confident that the estimated standard is far from zero. Similarly, the p-value for low-income percentage is 0.3518 (Table 2). This is another non-significant result, meaning that we can say that the residuals here are also non-spatially correlated.

The two other independent variables that were of interest for the analysis of property crime rates against social economic factors were the percentage of the population belonging to a visible minority group and the percentage of the population aged 15-29. The following age group was selected because according to previous studies this age group had the most correlation to crime rates in neighbourhoods in Toronto. (Rocque, Posick, and Hoyle 2015). The first stage of the analysis is comparing the generated choropleth maps for the dependent (property crime rates) and independent variables to visualise any geographic trends (3, 6, & 7).

In Figure 3 a conspicuous trend can be seen for the property crime rates where there is a higher concentration in the west and southern region of the city and a lower concentration in the north and eastern regions of the city. This trend approximately correlates with the spatial distribution of percentage of the population belonging to a visible minority group in Figure 6 because there is a high concentration of the variable in regions where property crime rates were observed to be highly concentrated as well. However, regions of low property crime rates in Figure 3 such as the east and north parts of the city are also shown to have high concentrations of the variable which contradicts with the prediction of a positive correlation between the variables.

A similar observation is made for the percentage of the population aged 15-29 in Figure 7. A high concentration is evident in the southern and western region of the city but not exactly at the neighbourhoods where a high concentration of property crime rates were observed in Figure 3. Hence the observations made using the choropleth maps do not suggest a strong correlation between the dependent and independent variables.

The second stage of the analysis is through inquiring the results of linear regression models for property crime rates against the independent variables to further enhance the visual observations of the variables. The summary of the linear regression model for property crime rates against percentage of the population belonging to a visible minority group in Table 3 suggests there is no correlation between the variables. This was concluded based on the very high p value for the F-statistic value which is in turn very low. These high and low values respectively, suggest that the coefficients such as the (slope and intercept) of this linear regression model is zero implying there is no relationship between the two variables. This can also be observed in Figure 12 where the slope of the regression line is horizontal which is slope of zero.

This similar trend is observed for the linear regression analysis for property crime rates against percentage of the population aged 15-29 where the F-statistic value is relatively low and the p value that is less than 0.05. This observation suggests that there might be a correlation between the variable and property crime rates but not a strong relation as it can be observed by the zero slope regression line in Figure 13.

Conclusion

The issue of property crime in Toronto is one that individuals devote a lot of energy to predicting and explaining. In the scope of our project, property crime consists of auto theft, breaking and entering, robber, and theft over $5000. The goal of this study was to analyse the relationship between property crime occurrence and four variables: population density, percentage of the population designated as 'low income,' percentage of the population belonging to a visible minority group, and percentage of the population aged 15-29. After creating choropleths maps to observe trends, a Boolean map to show relative risk above mean per neighbourhood, and regression analyses, it was determined that while there is a general, visual trend, there is no significant spatial correlation between property crime and three of the selected variables. To clarify, although there is overlap in neighbourhoods and crime type to be seen in the choropleth maps, it is not significant enough to warrant a firm conclusion. In response to our research question, the correlation between social economic factors and property crime in Toronto neighbourhoods is not strong enough to be significant. The percentage of population aged 15-29 is the only variable to show a fairly significant correlation to property crime in Toronto but still not strong enough to suggest a positive correlation between property crime rates and age.

References

10 Bond, Mark. 2015. "Criminology: Social Disorganization Theory Explained." https://www.linkedin.com/pulse/criminology-social-disorganization-theory-explained-mark-bond/?fbclid=IwAR1fCDHVXX-JLyw5Ov5JGoMtiOl4482

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

He, Li, Antonio Páez, and Desheng Liu. 2017. "Built Environment and Violent Crime: An Environmental Audit Approach Using Google Street View." *Computers, Environment and Urban Systems* 66: 83–95.

Hulchanski, John David. 2011. "The Three Cities Within Toronto: Income Polarization Among Toronto's Neighbourhoods, 1970-2005."

Kitchen, Peter Francis. 2006. *Exploring the Link Between Crime and Socio-Economic Status in Ottawa and Saskatoon: A Small-Area Geographical Analysis*. Department of Justice, Research; Statistics Division.

Longley, Robert. 2020. "Criminology: Definition and History." https://www.thoughtco.com/the-history-of-criminology-part-1-974579#:~:text=Sociological%20T.

Nickerson, Charlotte. 2019. "Conflict Theory." https://www.simplypsychology.org/conflict-theory.html.

Rocque, Michael, Chad Posick, and Justin Hoyle. 2015. "Age and Crime." *The Encyclopedia of Crime and Punishment*, 1–8.

Statistics Canada. 2021. "Census Profile, 2016 Census: Toronto, City [Census Subdivision], Ontario and Canada [Country]." Government of Canada. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/Page.cfm?Lang=E&Geo1=CSD&Code1=3520005&Geo2=PR&Data=Count&B1=All.

Thatcher, Amelia. 2015. "Property Crime." *RCMP Gazzette* 77 (3).

Toronto Police Service. 2019. "Public Safety Data Portal." https://data.torontopolice.on.ca/pages/open-data.

Van Ingen, Trevor, Erika Khandor, and Paul Fleiszer. 2015. *The Unequal City 2015: Income and Health Inequities in Toronto*. Toronto Public Health.