

Analysis of Toronto Break and Enter Crime Data in 2019

B&E Crimes and Geographical Factors

Feifei Li, Abdulrahman Alzaabi, Stuti Sekhri, Yan Mezhiborsky,
TUT0203, 203-1

March 30, 2020

Section 1

Introduction and Methods

Introduction

Occurrences of B&E crimes in Toronto during 2019 are normally distributed for spatial variables such as longitude, and latitude. Unusual frequencies of B&E crimes for some neighbourhoods were also observed. We attempt to find the correlations between the occurrences of B&E crimes and the abovementioned geographical variables.

Objectives

- Determine the correlation between B&E crime occurrence and geographical factors
 - Longitude
 - Latitude
 - Neighbourhood
 - And underlying variables in neighbourhoods such as:
 - Premise types
 - Population
- Estimate the location where B&E crimes occur most frequently
- Provide directions for future Toronto B&E crime analyses
- Provide advice for B&E crime prevention in Toronto

Data Summary (Longitude and Latitude)

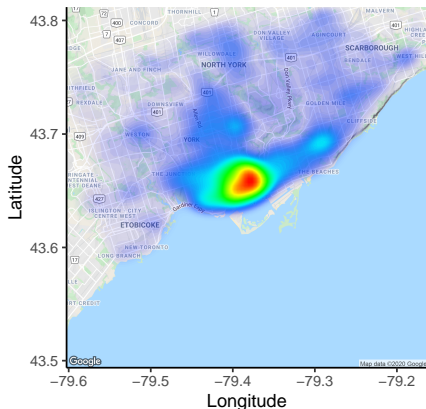
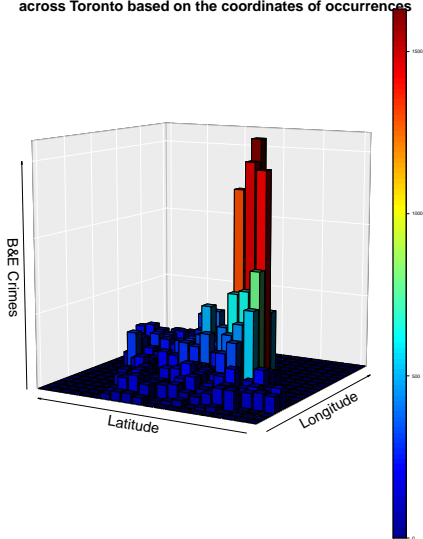


Figure 1. A heat map presenting the intensity of recorded B&E crimes committed in 2019 across Toronto.

Figure2. B&E crimes distribution across Toronto based on the coordinates of occurrences



Data Summary (Neighbourhood)

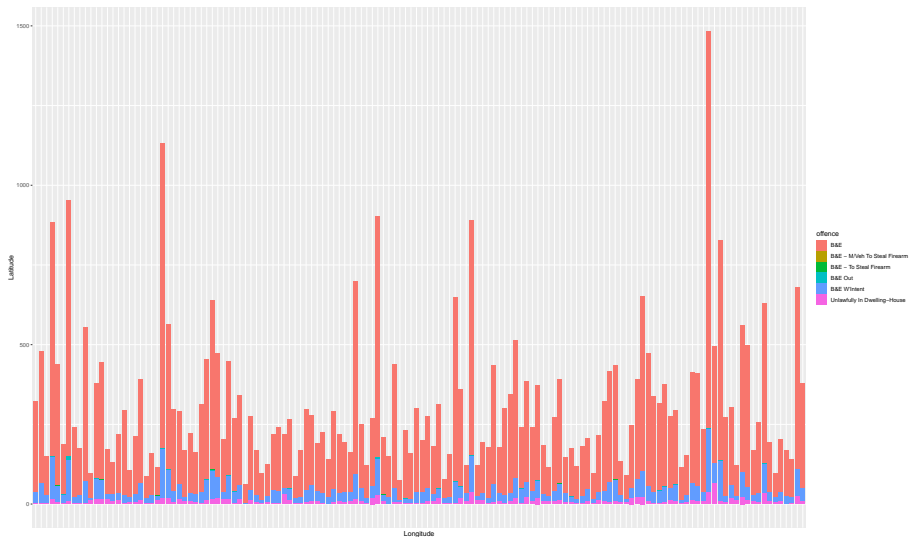


Figure 3. The distribution of B&E crimes in each neighbourhood of Toronto.

Statistical Methods

- Hypothesis testing for finding statistical evidence that B&E crimes are more inclined to take place in some neighbourhoods and less in some others.
 - Null hypothesis: All 140 neighbourhoods have the same probability of B&E crime occurrence i.e. $H_0 = 1/140$
- Bootstrapped the mode of coordinate data to estimate the location where B&E crimes are committed most frequently.
- Linear regression model to determine correlations between numbers of each premise type and occurrences of B&E crimes in neighbourhoods.
- Linear regression model to determine correlations between population and occurrences of B&E crimes in neighbourhoods.
 - Data source: 2016 Census of Population data from Statistics Canada

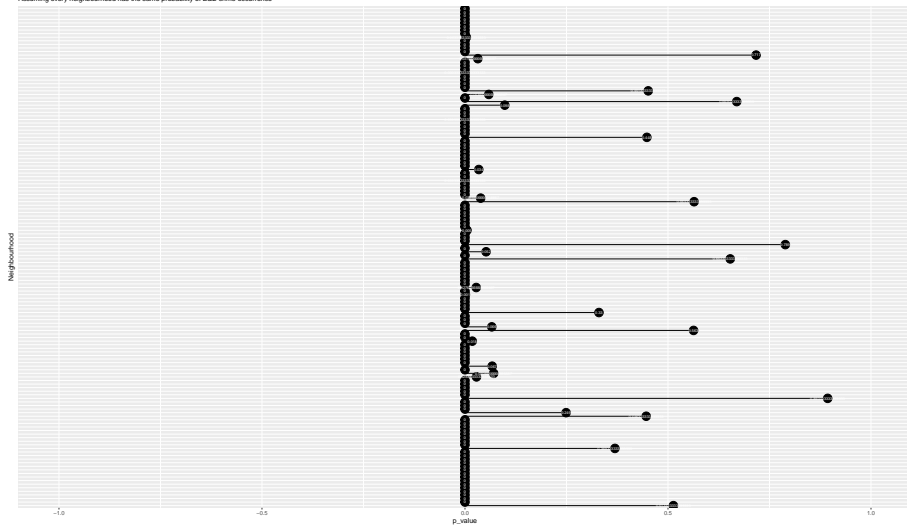
Section 2

Results

Hypothesis Testing

P-values for 140 neighbourhoods in Toronto

Assuming every neighbourhood has the same probability of B&E crime occurrence



Hypothesis Testing

- 108 of the 140 neighbourhoods have p-values of **0**, and 6 of the neighbourhoods have p-values not equal to 0 but less than 0.01.
 - This means that, if *the B&E crime rate in every neighbourhood is the same*, then, for 108 of the neighbourhoods, the probability of observing data that *is at least as unusual as the reported B&E crime rates* in those neighbourhoods, is **0%**, and for 6 of the neighbourhoods, the probability of observing data that *is at least as unusual as the reported B&E crime rates* in those neighbourhoods, is *less than 1%*
- Only 14 of the 140 neighbourhoods have p-values greater than 0.1
 - We have no evidence against the hypothesis that every neighbourhood has the same B&E crime rates for only $\frac{1}{10}$ of the neighbourhoods.

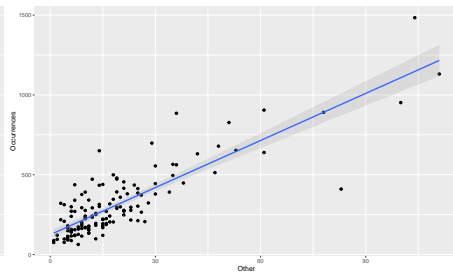
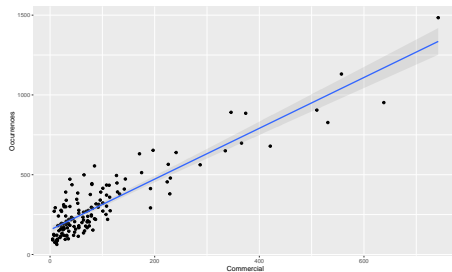
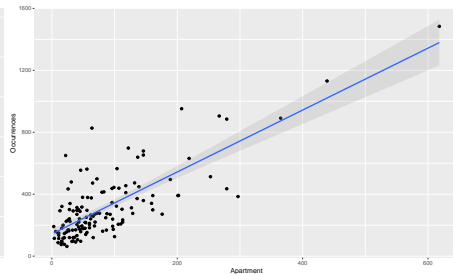
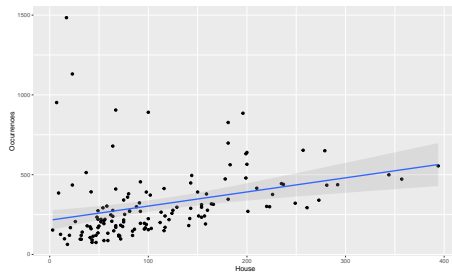
Bootstrapping

- An estimate for the location where B&E crimes occur most frequently in Toronto
- We are **95%** confident that the most frequent B&E crimes occurring location across Toronto in 2019 is in the area enclosed by the longitude between -79.3962 and -79.1862, and the latitude between 43.6533 and 43.7795.

Table 1: 95% confidence intervals for longitudes and latitudes of the most frequent B&E crime occurrence

coordinate	2.5%	97.5%
Longitude	-79.39616	-79.18623
Latitude	43.65335	43.77955

Linear Regression (Premise Type)



Linear Regression (Premise Type)

Table 2: Linear Regression Coefficients for House

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	214.6095797	30.9840389	6.926456	0.0000000
House	0.8853051	0.2306591	3.838154	0.0001891

Table 3: Linear Regression Coefficients for Apartment

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	143.713618	16.6140795	8.65011	0
Apartment	1.999268	0.1383825	14.44741	0

Positive correlation between the number of apartments in a neighbourhood and the occurrences of B&E crimes.

- p value for the t-test ($\text{Pr}(>|t|)$) being 0 tells us we have strong evidence against the null hypothesis that there is no correlation between the number of apartments in a neighbourhood and occurrences B&E crimes
- We are also seeing a positive correlation between the number of houses in a neighbourhood and occurrences B&E crimes. However the estimate of the slope is not as positive as that of the linear regression for the number of apartments in a neighbourhoods.
- Hence, a weaker correlation for the premise type of houses.

Linear Regression (Premise Type)

Table 4: Linear Regression Coefficients for Commercial Premises

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	153.954107	10.3230239	14.91366	0
Commercial	1.592232	0.0646892	24.61358	0

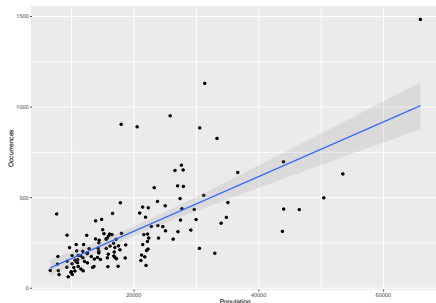
Table 5: Linear Regression Coefficients for Other Premises

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	124.325277	14.1837335	8.765342	0
Other	9.835915	0.5272135	18.656418	0

Positive correlation for both the numbers of commercial premises and other premises in a neighbourhood.

- Since both of their p-values are 0, and their estimates of the slope parameters for their simple linear regression models are positive, we have strong evidence against the null hypothesis that there is no correlations between the numbers of these two premise types and occurrences of B&E crimes.
- Strong correlation for other premise types.

Linear Regression (Population)



- Very weak positive correlation between population of a neighbourhood and occurrences of B&E crimes
- High $P_{intercept}$ value
- It's OK since it makes no sense to observe that 13 B&E crimes occur in a neighbourhood of zero population

Table 6: Linear Regression Coefficients for Population of 137 Neighbourhoods

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	13.258250	30.4312716	0.4356785	0.6637658
Population	0.015088	0.0013793	10.9386200	0.0000000

Conclusion

- According to our results, we are 95% confident that the most frequent B&E crimes occurring across Toronto in 2019 is in the area enclosed by the longitude between -79.3962 and -79.1862, and the latitude between 43.6533 and 43.7795. Perhaps the TPS should focus on this area for future B&E crime research.
- We found strong correlations between the number of apartments and commercial buildings in a neighbourhood, and the number of B&E crimes
 - For these property managers, installing more cameras inside these buildings could be a plausible measure for preventing B&E crimes from occurring.