



Applied Data Science Capstone Project

SEGMENTING AND CLUSTERING NEIGHBOURHOODS OF
FREDERICTON CITY

Project Scope

What neighbourhoods have the highest crime?

Is population density correlated to crime level?

Using Foursquare data, what venues are most common in different locations within the city?

Does the Knowledge Park really need a coffee shop?

Datasets of the project



1. OPEN DATA



2. FOURSQUARE
DATA FOR VENUES



3. FREDERICTON
CRIME BY
NEIGHBOURHOOD



4. FREDERICTON
LOCATIONS OF
INTEREST

Methodology



Loading each data set



Determine crime frequency by neighbourhood



Study the crime types based on neighbourhood



Identify correlation between crimes and population density



Perform k-means statistical analysis on venues by locations of interest based on findings from crimes and neighbourhood



Determine which venues are most common statistically in the region of greatest crime count then in all other locations of interest.



Determine if an area, such as the "Knowledge Park" needs a coffee shop.

Results

Crime frequency by neighbourhood

Total crimes in the city

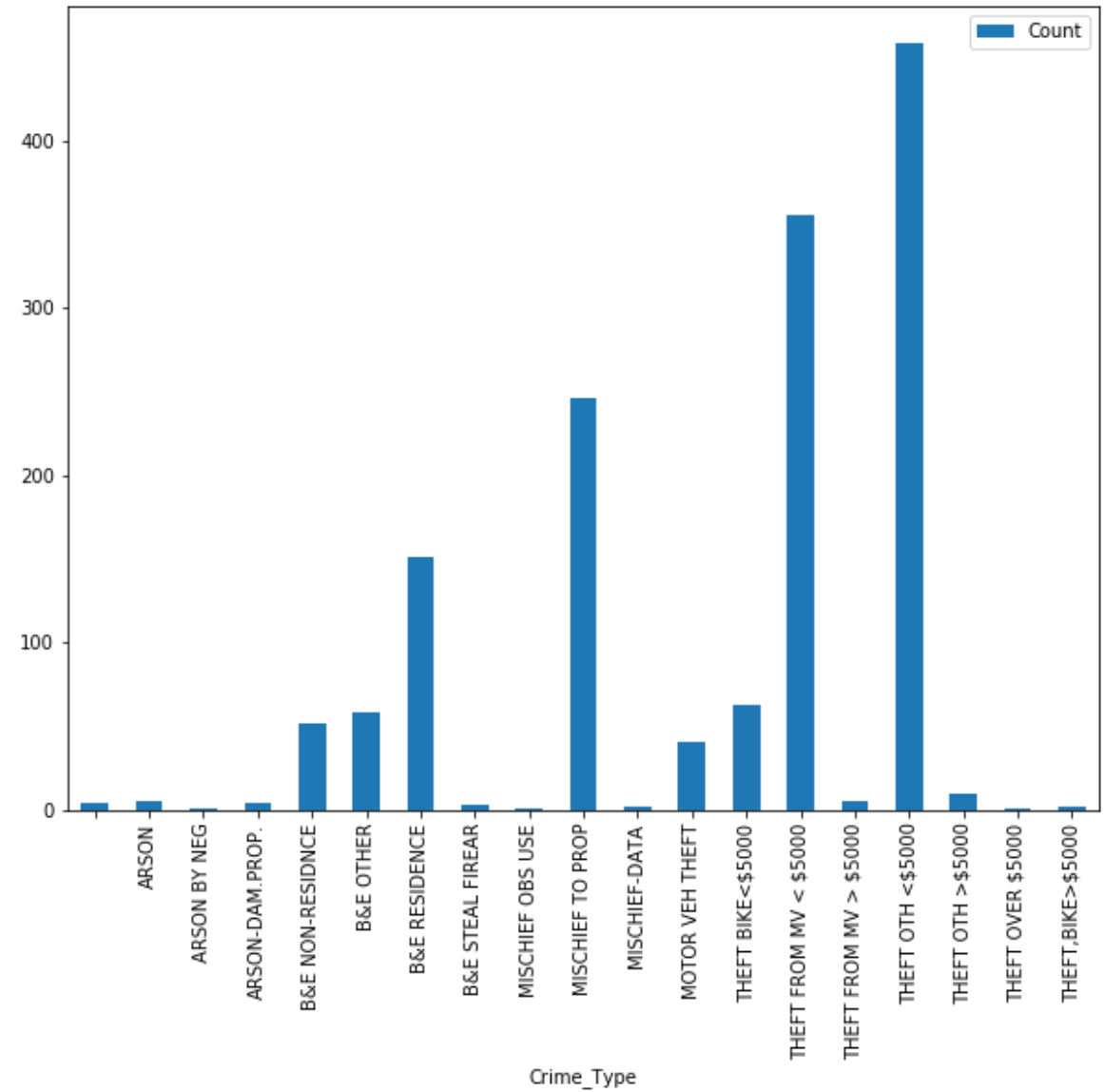
Crime type count by neighbourhood.

Theft from motor vehicles is most prevalent.

Using k-menas, we were able to determine the top 10 most common venues within a 1 km radius of the centroid of the highest crime neighbourhood.

The most common venues in the highest crime neighbourhood are **coffee shops followed by Pubs and Bars **

Plot





1. All in all, the City of Fredericton Open Data is interesting, it misses the details required for true valued quantitative analysis and predictive analytics which would be most valued by investors and developers to make appropriate investments and to minimize risk.



2. However, the inference from open dataset and foursquare API is a good starting point to know more the various neighborhoods in the city and understand the crime incidents.

Conclusion

THANKS

