Segmenting & Clustering Neighbourhoods in Fredericton

Introduction

Developers, investors, policy makers and/or city planners have an interest in answering the following questions as the need for additional services and citizen protection:

1. What neighbourhoods have the highest crime?

2. Is population density correlated to crime level?

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

4. Does the Knowledge Park really need a coffee shop?

Data Acquisition

To understand and explore we will need the following Open Data sources:

- Open Data Site: http://data-fredericton.opendata.arcgis.com/

- Fredericton Neighbourhoods: http://data-fredericton.opendata.arcgis.com/datasets/neighbourhoods--quartiers

- Fredericton Crime by Neighbourhood: http://data-fredericton.opendata.arcgis.com/datasets/crime-by-neighbourhood-2017--crime-par-quartier-2017

- Fredericton Census Tract Demographics: http://data-fredericton.opendata.arcgis.com/datasets/census-tract-demographics--donn%C3%A9es-d%C3%A9mographiques-du-secteur-de-recensement

- Fredericton locations of interest: https://github.com/JasonLUrquhart/Applied-Data-Science-Capstone/blob/master/Fredericton%20Locations.xlsx

Methodology

The methodology will include:

1. Loading each data set

2. Examine the crime frequency by neighbourhood

3. Study the crime types and then pivot analysis of crime type frequency by neighbourhood

4. Understand correlation between crimes and population density

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

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

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

Results
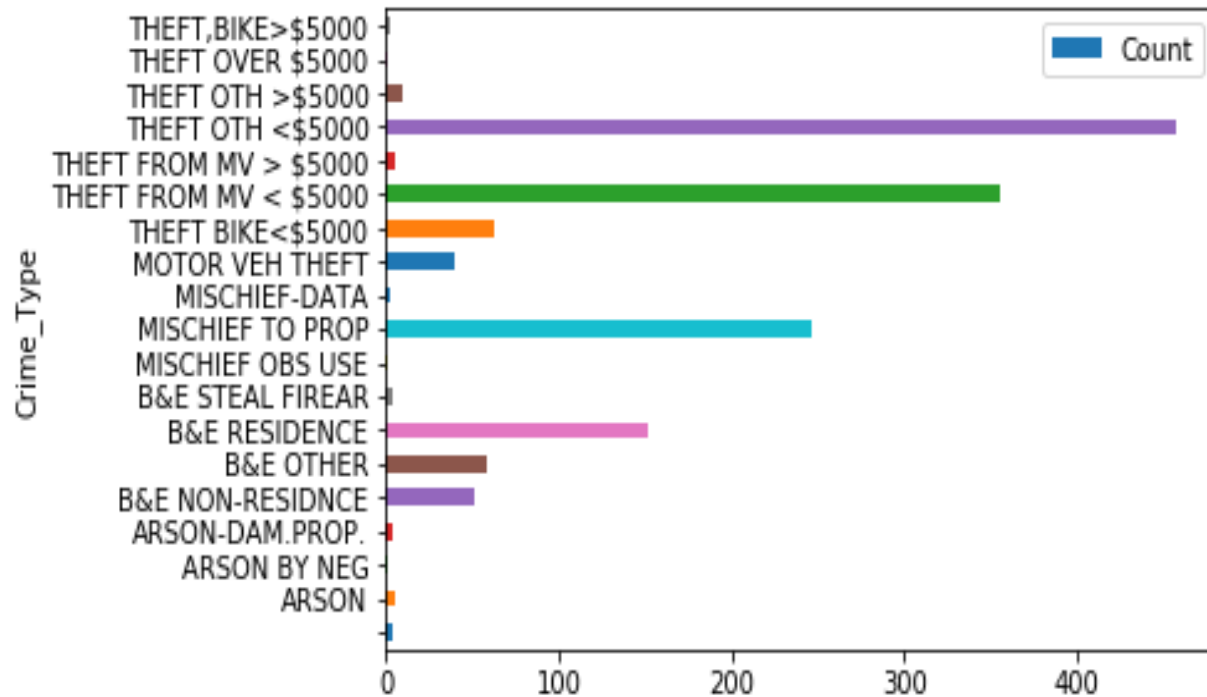
- Other theft less than $5000 was the most common crime

- Motor vehicle theft less than $5000 was the second most common

Results

- Analysis of crime by suburb resulted in the suburb of 'Plat'

- Population density was not strongly correlated to crime frequency

26	Monteith / Talisman	3
27	Montogomery / Prospect East	3
28	Nashwaaksis	9
29	Nethervue Minihome Park	1
30	North Devon	17
31	Northbrook Heights	5
32	Plat	40
33	Poet's Hill	2
34	Prospect	11
35	Rail Side	2
36	Saint Mary's First Nation	1
37	Saint Thomas University	1
38	Sandyville	3

Results

- 73 unique venue categories

- 153 unique venues

- Able to print the top 5 most common venues in each location

- Currently **no** coffee shops in Knowledge park therefore it may be appropriate to open one

```
----Knowledge Park----
                      venue  freq
0     Fast Food Restaurant   0.13
1           Clothing Store   0.10
2             Liquor Store   0.06
3               Restaurant   0.06
4  Furniture / Home Store   0.06
```

Conclusion and Future Directions

- Able to understand crime type and frequency by area but we do not understand the distribution properties

- Theft from motor vehicles is the most prevalent crime in these areas

- Knowledge park **does not** currently have a coffee shop, so it may be beneficial to open one there

- Future Direction / Ideas
  - Issues with open data sources, more scope to fix this for more accurate data
  - Issues with combining data sets between crime neighbourhoods – more specific coding (timestamp and date) may be beneficial