APPLIED DATA SCIENCE CAPSTONE PROJECT

MADE BY- NEDIFF AMALA NIXON

MAY 12, 2020

Introduction

- *Fredericton is the Capital City of the only Canadian fully-bilingual Province of New Brunswick.*
- *As the city grows and develops, it becomes increasingly important to examine and understand it quantitatively.*
- *The City of Fredericton provides open data for everyone and encourages entrepreneurial use to develop services for the benefit of its ciitzens.*

What are we looking for?

- *What neighborhoods 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?*

Some sources of data

- *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*
- *Using this data will allow exploration and examination to answer the questions.*

Methodology

- *Loading each data set.*
- *Examine the crime frequency by neighborhood.*
- *Study the crime types and then pivot analysis of crime type frequency by neighborhood.*
- *Understand correlation between crimes and population density.*
- *Perform k-means statistical analysis on venues by locations of interest based on findings from crimes and neighborhood.*

Exploring the data

- *Exploring the count of crimes by neighborhood gives us the first glimpse into the distribution.*
- *One note is the possibility neighborhoods names could change at different times.*

First visualization

- *Once the data was prepared, a choropleth map was created to view the crime count by neighborhood.*
- *Examining the crime types enables us to learn the most frequent occurring crimes which we then plot as a bar chart to see most frequent type.*

Examine 2nd most common crime

- *After exploring the pivot table showing Crime Type by Neighborhood, we drill into a specific type of crime, theft from vehicles and plot the choropleth map to see which area has the greatest frequency.*
- *The Platt neighborhood appears as the most frequent.*
- *Visualizing the population density enables us to determine that the Platt neighborhood has lower correlation to crime frequency than I would have expected*
- .

Understanding connection to venue

- *Loading the "Fredericton Locations" data enables us to perform a statistical analysis on the most common venues by location.*
- *Plotting the latitude and longitude coordinates of the locations of interest onto the crime choropleth map enables us to now study the most common venues by using the Foursquare data.*
- *Putting this data into a pandas dataframe we can then determine the most common venues by location and plot onto a map.*

Result

- *Neighbourhoods in Fredericton.*
- *Crime frequency by neighbourhood.*
- *Crime type frequency and statistics. The mean crime count in the City of Fredericton is 22.*
- *Using k-means, we were able to determine the top 10 most common venues within a 1 km radius of the centroid of the highest crime neighborhood.*
- *We were able to determine the top 10 most common venues by location of interest.*

conclusion

- *Using a combination of datasets from the City of Fredericton Open Data project and Foursquare venue data we were able to analyze, discover and describe neighborhoods, crime, population density and statistically describe quantitatively venues by locations of interest.*
- *The Open Data project is a great start and empowers the need for a "Citizens Like Me" model to be developed where citizens of digital Fredericton are able to share their data as they wish for detailed analysis that enables the creation of valued services.*

THANK YOU