Swaggy Intelligent Canines and Kitties (SICK Co.) Company

Expansion Analysis – Venue Scoring and Toronto Crime Analysis
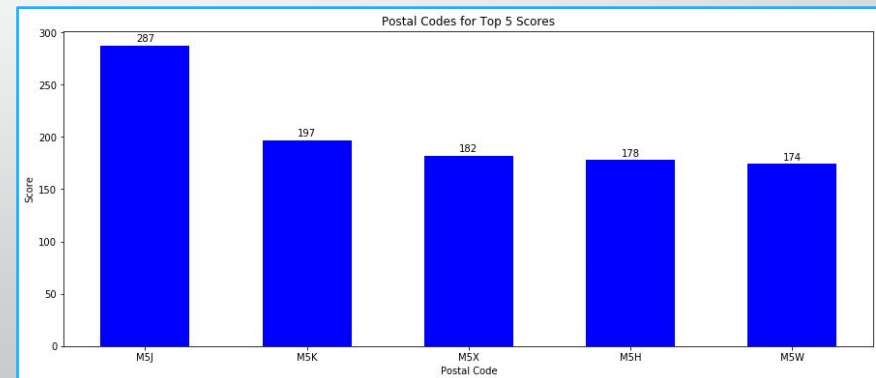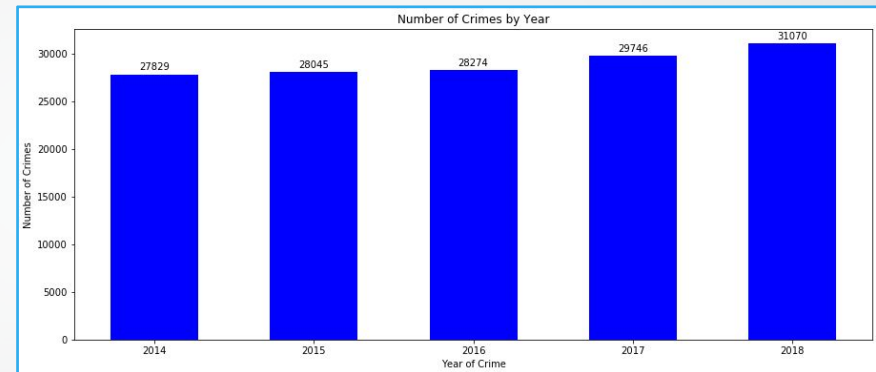
Introduction – Business Problem

- As a consultant to SICK Co., our objective is to identify the best location for SICK Co. to expand within the Toronto metropolitan area. Since this expansion is critical to the Company's success, SICK Co. created an Expansion Committee (EC) solely for the purpose of spearheading the proper selection of the site locale.

- Per instruction of the EC, geographic areas for expansion would take into account the types of venues (e.g, shops, entertainment centers, and various other points of interest). In addition, the EC devised a scoring/weighting system based on the type of venue. Total venue score, as well as the prevalence of crime in the area, will play a material part in the final decision on where to add the new location.
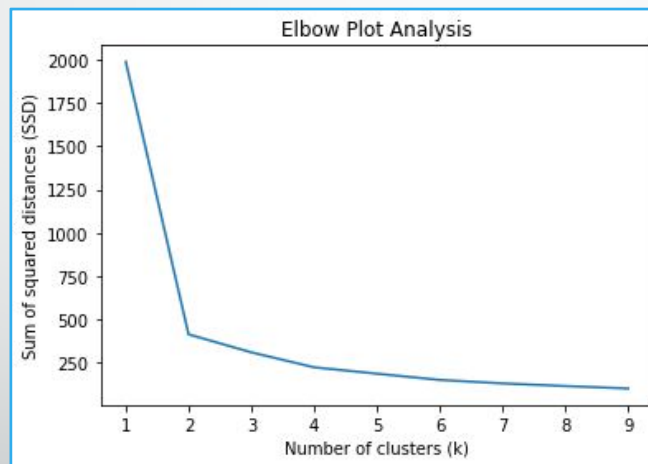
Data Sources, Extraction and Preparation

The data sources for site expansion analysis came from a variety of websites and data tables, including the following:

- FourSquare Places API (https://developer.foursquare.com/docs/api)
  - Contains venue data in order to score potential expansion locations in the Toronto area.
- Toronto Police Service Public Safety Data Portal (http://data.torontopolice.on.ca/datasets/mci-2014-to-2018/geoservice)
  - Deleted duplicate id's and 40 NULL records (0.03% of the data set) leaving 145,777 rows and 29 features detailing crime activity in the Toronto area. In addition, records prior to year 2014 were deleted based on limited number of data points (crime totals for 2014-2018 shown in bar chart to the right).
- Wikipedia List of Postal Codes (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
  - Data enables us to connect crime activity from venue data from FourSquare Places API with crime data from Toronto Police Service Public Safety Data Portal.
- Scores / weightings for venue categories provided by the Expansion Committee for use in this expansion project (toronto_venues_scores_only.xlsx).
  - Top five Postal Codes by scores (M5J, M5K, M5X, M5H, and M5W) are shown in the bar chart to the right.

Number of Crimes by Year

Postal Codes for Top 5 Scores

Methodology – K-Means Clustering

- After normalizing the one hot coding for the crime data, we used the "elbow" plot to find the optimal number of clusters (k=2).

K-Means Clustering Results

- Grouping the data based on the resulting cluster labels (and taking the mean of the normalized features) results in the table below with the following interpretation:

  - The negative values mean "lower than most" and positive values mean "higher than most."

  - Label 0 is the "not so lucky" label that has Neighborhoods with higher than average MCI counts (i.e., more crime than average).

  - Label 1 has Neighborhoods with lower than average MCI counts (e.g., assault, auto theft, break and enter, robbery and theft over). So, it has less crime (on average) than other cohorts.

Labels	Assault	Auto_Theft	Break_and_Enter	Robbery	Theft_Over
0	0.266645	0.312836	0.200536	0.25538	0.172422
1	-0.175672	-0.206103	-0.132118	-0.16825	-0.113595

Recommendation and Conclusion

- Comparatively, Bay Street Corridor (Label = 1) has lower than average crime counts while Waterfront Communities-The Island (Label = 0) has higher than average crime counts. Therefore, Bay Street Corridor would be more desirable for the expansion location due to lower relative rates of crime.

- Recommendation is to choose the Bay Street Corridor Neighborhood (Postal Code M5K) for the expansion location. Ultimately, employee safety should be one of the most important factors when deciding where to expand corporate locations.

- Choosing the expansion location in the Bay Street Corridor Neighborhood (Postal Code M5K) is consistent with the goal of selecting an area with one of the highest scores as well as relatively lower crime occurrences (when comparing the areas with the top five scores).

Postal Code	Toronto Police Department Neighborhood	Score	Label
M5J	Waterfront Communities-The Island (77)	287	0
M5K	Bay Street Corridor (76)	197	1
M5X	Bay Street Corridor (76)	182	1
M5H	Bay Street Corridor (76)	178	1
M5W	Waterfront Communities-The Island (77)	174	0