# The best neighborhoods in Toronto

Mithun Chakraborty

August 1, 2019

**1.Introduction**

**1.1Background(Business Problem)**

Toronto is the ultimate multicultural destination, as underlined by its many ethnic enclaves, all peppered with a variety of delicious restaurants, bars etc.

Although each neighborhood offers a deep dive into its dominant culture’s retail and culinary offerings, they all somehow also cater to a variety of urban tribes, from the creative to the foodie type. Together, they tell the tale of a city both resolutely heritage-minded and full-throttle dedicated to its burgeoning future. Real-estate is booming industry now in Toronto, people are immigrating from different countries and making Toronto as their new home. Hence it will be advantages for them to know which place is safest for them to stay in Toronto neighborhoods. This project aims to figure out where to stay when in Town.

**2. Data**

Data that might contribute to determining safest places in the neighborhoods might include different crime related statistics for 2018. I will be using Toronto police website to solve the problem. I will provide rank to each neighborhood based on above mentioned data and then share the result of top ten neighborhoods in Toronto.

**3. Methodology**

In this project we will direct our efforts on detecting neighborhoods of Toronto that have low crime rates.

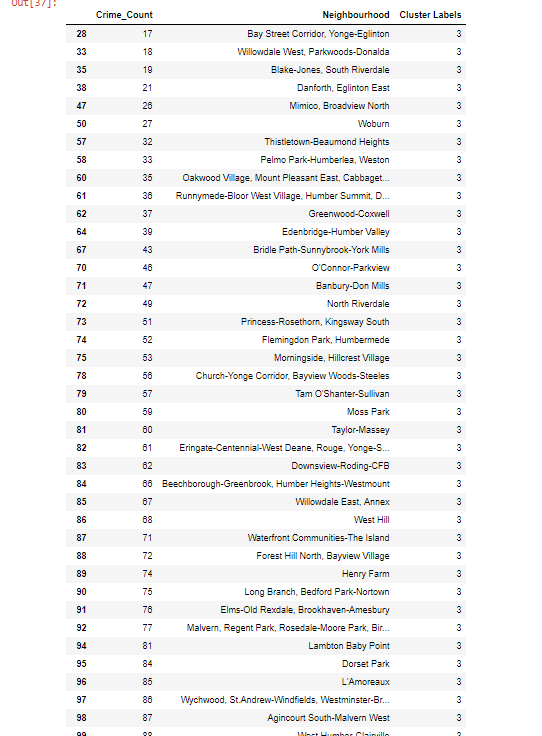
In first step we have collected the total **crime count(**assaults+Homicide) **in 2018 for each neighborhood in Toronto**.

Then, I have used Kmeans Clustering for exploratory analysis.

K-Means is one of the most popular “clustering” algorithms. It is the process of partitioning groups of data points into a small number of clusters. Using our crime data, as we measure the number of assaults+Homicide, the neighborhoods with high numbers of crimes will be grouped together. The goal of K-Means clustering is to assign a cluster to each data point (neighborhoods). We first partition datapoints (neighborhoods) into k clusters in which each neighborhoods belongs to the cluster with the nearest mean (serving as a prototype of the cluster).

As one of the unsupervised learning algorithms,we use K-Mean to build models that help us understand our data better. It enables us to group unlabeled data points.

**4. Result:** Based on Kmeans clustering algorithm, I have come to the conclusion that Safest neighborhoods in Toronto should be places which has clustered label as 3



Please refer the notebook for complete list.

**5. Discussion:**

During determining the safest places in Toronto, I have considered that less crime with be more safe place to live for the newcomer. That’s why I have used data for different types of crime from Toronto police website for 2018 and grouped/clustered the neighborhoods based on the total crime count. We have used KCluster value of 5, as there 140 neighborhoods so clustered them under different clustered labels. It is projected from the exploratory analysis clustered with label 3 is the cluster for safest neighborhoods.

**6. Conclusion:**

In this study, I analyzed the relationship between Crime count and Neighborhood within Toronto. I identified different crimes committed in 2018.I built Cluster using Kmeans algorithm which is a method of vector quantization to determine which are areas are most safest in Toronto to stay. Any crime count between 0-103 are considered as safest neighborhoods in Toronto as derived by the algorithm. This will be helpful to determine the safest neighborhoods for each year. Based on the suggested model new comer can choose the area the want to live with their family.