



Prediction on location to open Italian restaurant

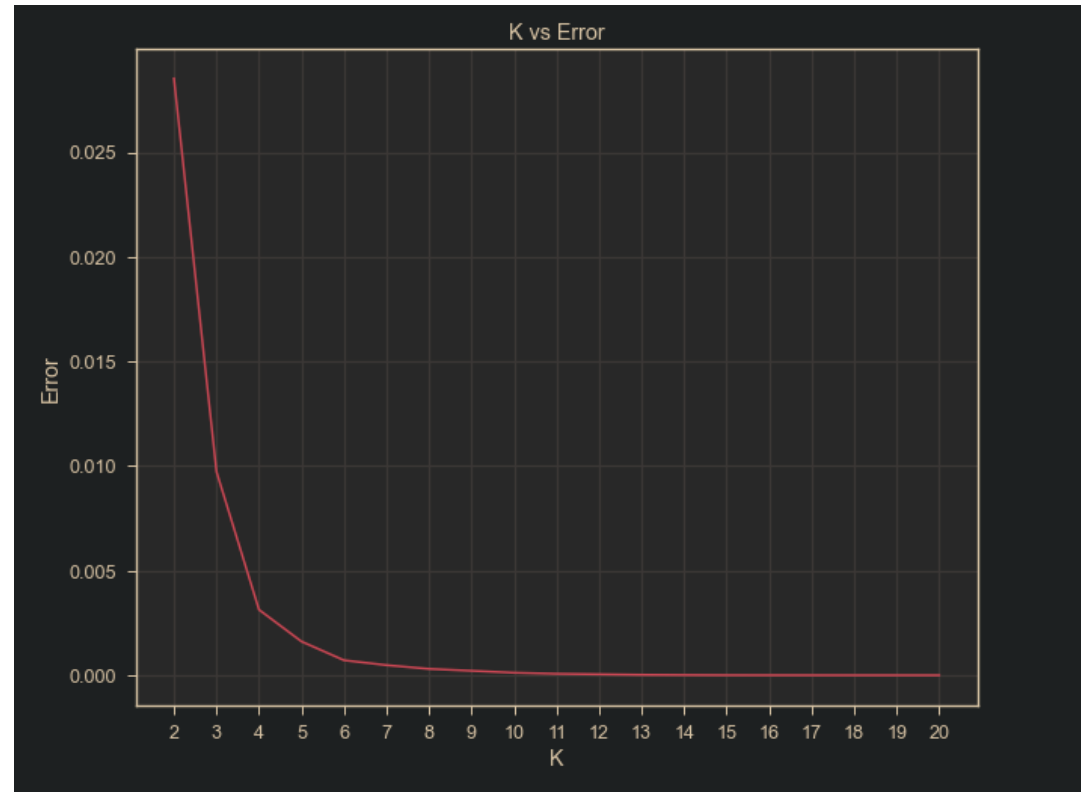
Predicting location for restaurant is valuable for business

1. Minimizing risk when starting up a new business
2. Maximizing revenue base on location of business
3. Minimizing failure of business

Data acquisition and cleaning

1. Wikipedia dataset on Toronto neighbourhoods
2. CSV file on geospatial data
3. Foursquare API on venues

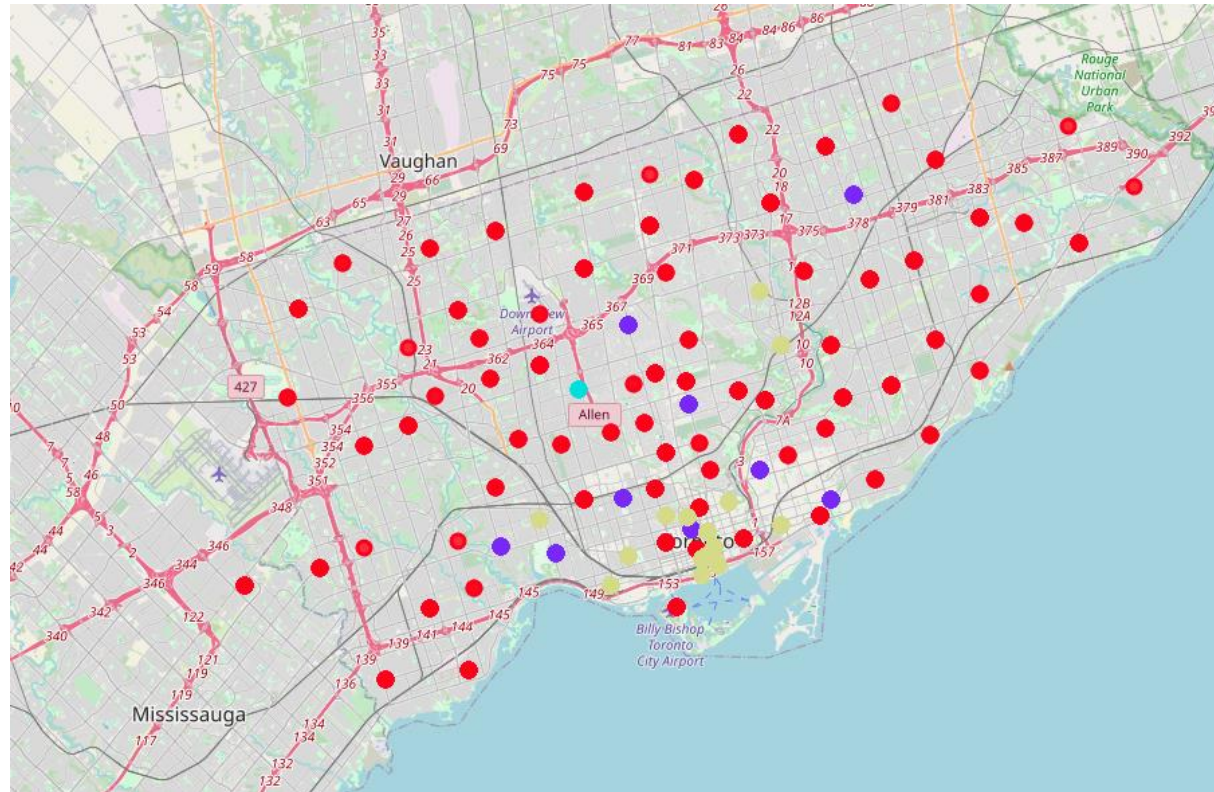
K-means Clustering



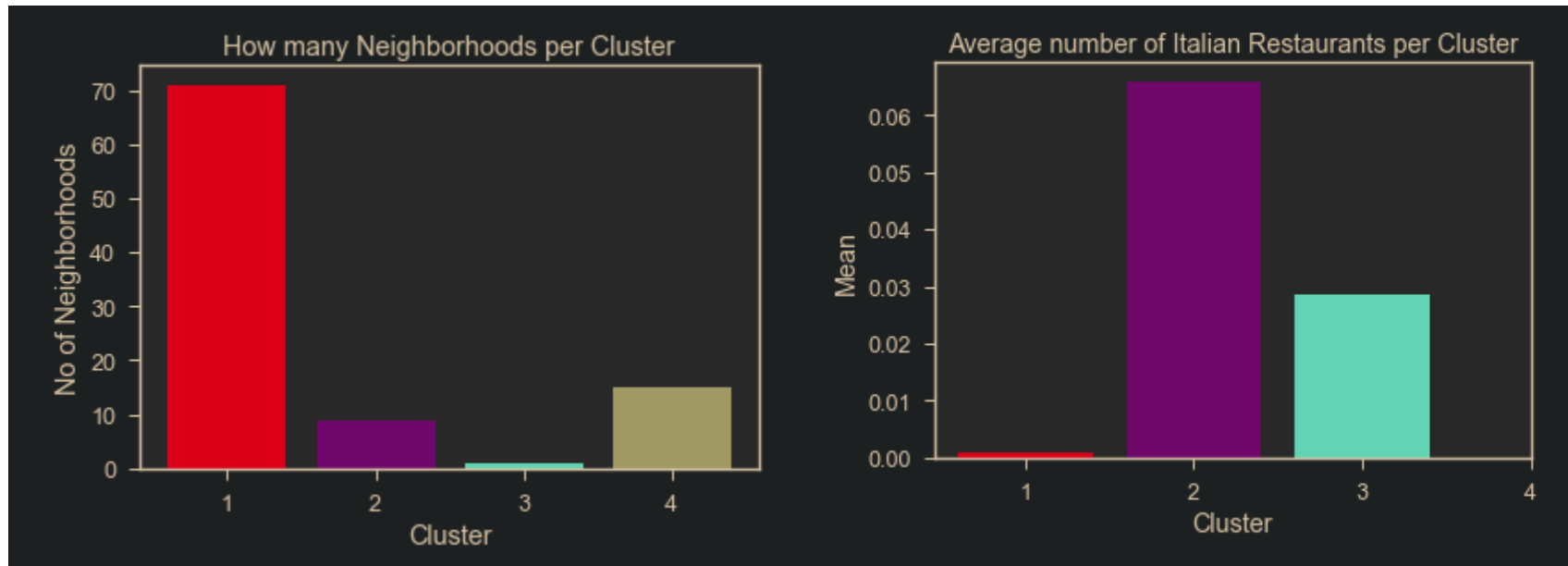
K-means Clustering

1. Best K to use is shown on the graph (K=4) (Elbow point)
2. Group neighbourhoods into clusters (4 clusters)

Visualizing the clusters



Visualizing the clusters



Cluster 1 (Red)

71 unique neighbourhoods and only 1 Italian Restaurant in First Canadian Place, Underground city

Cluster 2 (Purple)

Cluster 2 has 9 unique neighbourhoods but they have 17 Italian Restaurants

Cluster 3 (Blue)

Cluster 3 has 15 unique neighbourhoods and 26 Italian Restaurants

Cluster 4 (Turquoise)

No neighbourhoods and no Italian restaurants

Discussion

1. Setting up a new shop at cluster 1 area (eg: Parkwoods, Victoria Village)
2. 71 neighbourhoods but only 1 Italian restaurant
3. Less competition with many neighbourhoods

Conclusion and future direction

1. Add population data over the neighbourhoods
2. Find out what other restaurants are there in the cluster