



Capstone Project - The Battle of Neighborhoods

Coursera IBM Data Science Certification

RISHI K S

July 2020



INTRODUCTION



- I am creating a hypothetical concept scenario for American Restaurants in Busy Toronto Area
- Designing this project to help the entrepreneur to find the most suitable location in Toronto



DATA Acquisition and cleaning

- List of neighborhoods in Toronto, Canada
- Latitude and Longitude of these neighborhoods
- Venue data related to American restaurants. This will help us find neighborhoods that are more suitable to open an American Restaurant.

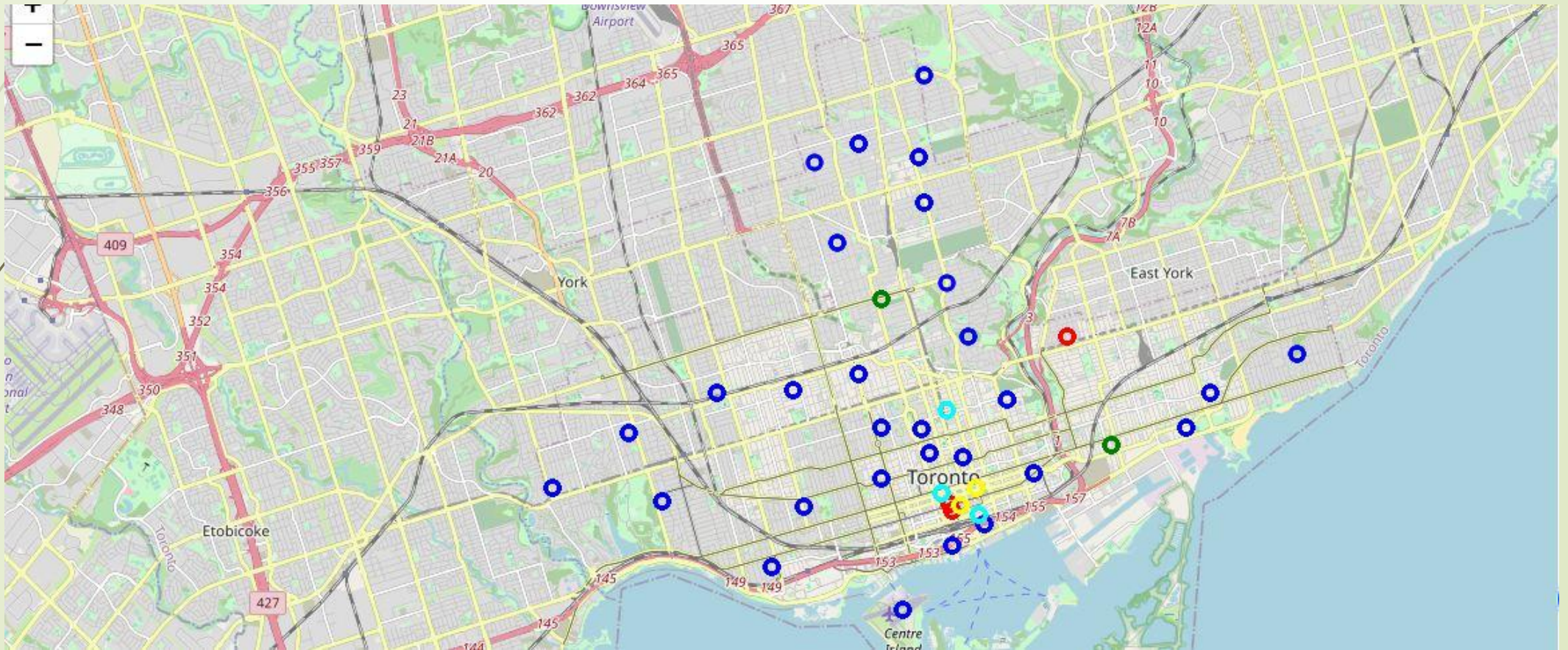


Methodolgy



- The scrapping of Toronto neighborhoods via Wikipedia
- Getting Latitude and Longitude data of these neighborhoods via Geocoder package
- Using Foursquare API to get venue data related to these neighborhoods
- Performing the clustering method by using k-means clustering
- Clustered the neighborhoods in Toronto into 5 clusters based on their frequency

Results



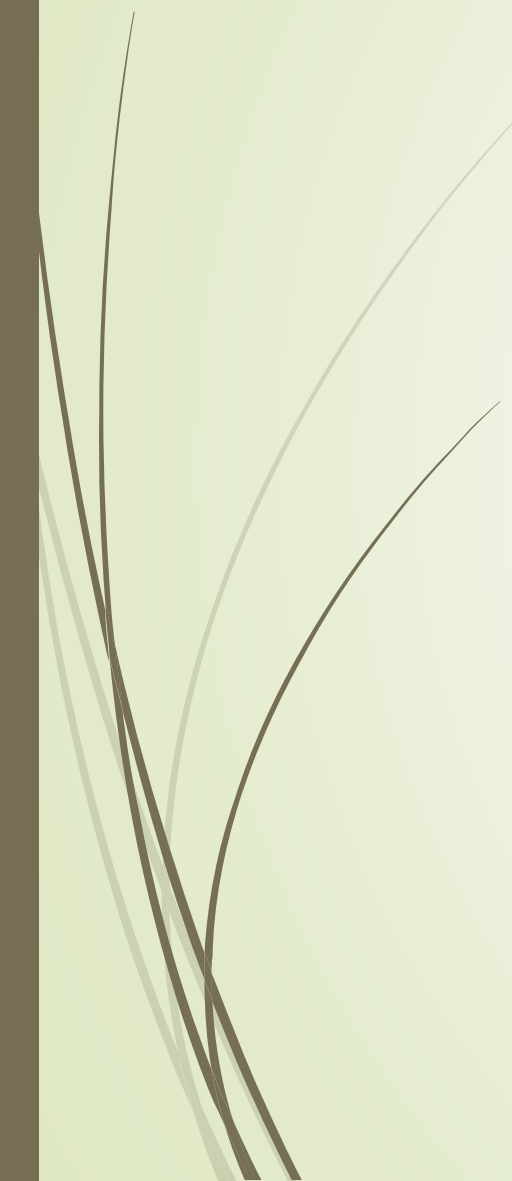


Result

- Using The scrapping of Toronto neighborhoods via Wikipedia
- Getting Latitude and Longitude data of these neighborhoods via CSV File
- Using Foursquare API to get venue data related to these neighborhoods
- Performing the clustering method by using k-means clustering
- Unsing Clustered the neighborhoods in Toronto into 5 clusters based on their frequency
- Cluster 0: Neighborhoods with good number of American restaurants.
- Cluster 1: Neighborhoods with no American restaurants.
- Cluster 2: Neighborhoods with a more number of American restaurants
- Cluster 3: Neighborhoods with a more number of American restaurants
- Cluster 4: Neighborhoods with a few number of American restaurants



CONCLUSION

- Most of the American restaurants are in cluster 0, cluster 3 and cluster 4
 - cluster 1 might be a good location as there are no American restaurants in these areas
- 



Thank you