2.Data Acquisition and Cleaning

2.1 Data Sources

The data, including postal codes, Boroughs and Neighborhoods scrapped from the Wikipedia page: <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>

The latitude and longitude values of different Neighborhoods of Toronto was available from: <https://cocl.us/Geospatial_data>

The location data was accessed from Foursquare location data.

2.2 Data Cleaning

The Toronto data including Postal code, Borough and Neighborhood was scrapped from Wikipedia and changed into a pandas Data Frame.

As part of Data Cleaning, first removed all rows in which Borough= “Not Assigned”.

More than one Neighborhood can exist in one Postal Code. Combined these rows into one row.

If Neighborhood= “Not Assigned”, then, value of corresponding Borough is given to Neighborhood.

Downloaded the Geospatial data of Toronto and merged this data to the existing Data Frame.

2.3 Location Analysis

Accessed the location data of Toronto including nearby venues using Foursquare API. Collected the information regarding Indian restaurants in different Neighborhoods.