

1.Introduction Section:

In this section, we would define the scope of this project and the core idea of this project. Primary question for any business owner would be selecting the location for their operations. Here I would consider a restaurant owner who is looking for expanding in the city of Toronto. The investors/owners would need certain proofs to consider a particular neighborhood in Toronto for their new venture which would result in profits. One can understand the neighborhoods by exploring the venues around and can get an idea on the type of locality.

Business problem:

Challenge is to find a suitable location in Toronto for a new restaurant along with recommendations regarding the theme and cuisine that would comply with the demands of the restaurant owner.

Target audience:

Target audience for this project would be

- Business executives for exploring options in and around a particular location
- Investors/ potential owners of restaurants, coffee shops, gyms
- Real estate dealers
- Data science enthusiasts

2.Data Section:

Data that is required to solve above business problem:

- List of neighborhoods in Toronto
- Geospatial data regarding the neighborhoods
- Venues around neighborhood

Data sources:

- "List of Postal code of Canada: M"
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M) wiki page to get all the information about the neighborhoods present in Toronto
- "https://cocl.us/Geospatial_data" csv file and geocoder to get geospatial data of neighborhoods
- Foursquare API calls for getting information regarding venues in Toronto

Data cleaning/preprocessing:

1. Data from wiki regarding the neighborhoods is extracted into data frame which include neighborhood name, postal code and borough.
2. Only cells with assigned borough are processed ignoring others
3. Latitude and longitude data is obtained from CSV file and added to the data frame to complete the data frame
4. Foursquare API call is utilized to obtain venues around each neighborhood and the data is merged into a dataframe along with venue categories

With the data obtained which has neighborhood information along with venues around, we would perform next steps to converge to a solution