KV TORONTO RESTAURANT PROBLEM

PROBLEM STATEMENT-

- Opening of a new restaurant in Toronto
- Requirement of maximum footfall
- Good crowd with perishable income
- Finding of best area/neighborhood

DATA-

Our problem clearly requires a certain type of data set to come upon the best possible outcomes.

- Wikipedia will be used to get the postal codes of various locations
- Foursqaure API will be used to get exact locations,no. of restaurants and type of restaurants.

We are getting following data from these credible sources:

- Postal codes
- Longitudes and latitudes
- Distance from centre of the city
- Age group of people
- Income of people
- Existing restaurants in the neighbourhood

METEHODOLOGY

We have taken following steps:

- We have collected data from aforementioned sources.
- Location (in terms of longitude and latitude) along with the category of restaurant.
- We have identified restaurants’ density in various areas.
- We’ll use K-mean to cluster areas with promising potential.
- Chosen clusters are made to be close to less than a km away from centre of the city.
- Map using Folium will be displayed to give a practical outlook of the chosen locations.

ANALYSIS

1. Identification and cleaning

   We have to identify and capture the data from all mentioned sources.Some portion of the data is missing ,so we need to clean that portion out of our dataset.

2. Combining data sources

   With postal address and longitude-latitude from different data source present with us,we need to combine all of them .

3. Sorting neighbourhood on the basis of latitude and longitude

   From previous step,a resulting data set will contain data about neighbourhood ,its postal code and latitude-longitudes

4. Clustering

We have used K-cluster algorithm to cluster various neighbourhoods.Each cluster is analsyed on the basis of distinguishing features. No. of restaurants and types are our determining variables.

CONCLUSION

In our result we have found 65 neighbourhoods inside the gerographical co-ordinates obtained using Foursqaure API.Out of the 5 clusters ,1 cluster shows the perfect density for opening of the restaurant.