Coursera Capstone

IBM Applied Data Science Capstone

# *Opening a New Restaurant In Mumbai,India*

By: Sumit

October 2019

# Business Problem

* Location of the restaurant is one of the most important decisions
* Objective: To analyse and select the best locations in the city of Mumbai, New Delhi to open a new restaurant
* Business question

➢In the city of Mumbai, if a property developer is looking to open a new restaurant, where would you recommend that they open it?

# Data

* Data required

➢List of neighbourhoods in Mumbai

➢Latitude and longitude coordinates of the neighbourhoods

➢Venue data, particularly data related to shopping malls

* Sources of data

➢Wikipedia page for neighbourhoods

(<https://en.wikipedia.org/wiki/Category:Suburbs_of_Mumbai>)

➢Geocoder package for latitude and longitude coordinates

➢Foursquare API for venue data

# Methodology

* Web scraping Wikipedia page for neighbourhoods list
* Get latitude and longitude coordinates using Geocoder
* Use Foursquare API to get venue data
* Group data by neighbourhood and taking the mean of the frequency of occurrence of each venue category
* Filter venue category by Restaurant
* Perform clustering on the data by using k-means clustering
* Visualize the clusters in a map using Folium

# Results

K-means clustering shows 5 clusters based on the frequency of occurrence for “Restaurant”:

Cluster 0 and cluster 4 have the most number of restaurants while clusters 1,2 and 3 have moderate no of clusters.Restaurants in cluster 0 and 4 have the most competitionwhile restaurants in clusters 1,2 and 3 have moderate competition.So this represents a great opportunity for a person to open a restaurant in these areas.While Clusters 0 and 4 have a oversupply of restaurants they may have adequate amount of customers due to that area being well developed.But the other areas are also well developed and have a undersupply of restaurants.So it is a great opportunuty to open a restaurant there.

# Discussion

Group 0 and group 4 have the most number of restaurant while bunches 1,2 and 3 have moderate no of clusters.Restaurants in bunch 0 and 4 have the most challenge while restaurant s in clusters 1,2 and 3 have moderate competition.So this speaks to an extraordinary open door for an individual to open an restaurant in these areas.While Clusters 0 and 4 have an oversupply of restaurant they may have sufficient measure of clients because of that region being admirably developed.But different regions are additionally very much created and have an undersupply of restaurants.So it is an incredible opportunuty to open a restaurant there.

# Conclusion

* The neighbourhoods in cluster 1 or 3 are the most preferred locations to open a new restaurant
* Findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open a new shopping mall

Thank you!