

IBM CAPSTONE PROJECT



INTRODUCTION

- The city of Mumbai consists of a large number of coffee, but still scope for new ones.
- Finding a suitable location for it to flourish is the most important factor for a coffee shop.
- It has to be set up in a location where one can attract a good crowd as well as it must be located in an area where there is little or no competition.

DATA

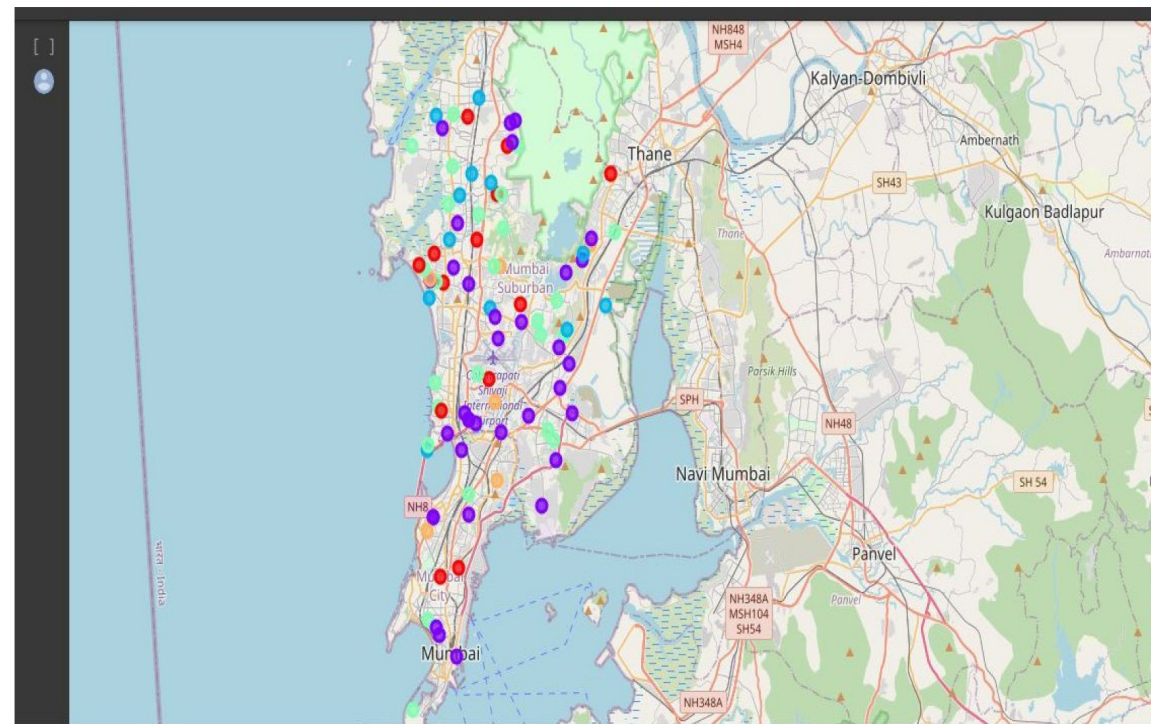
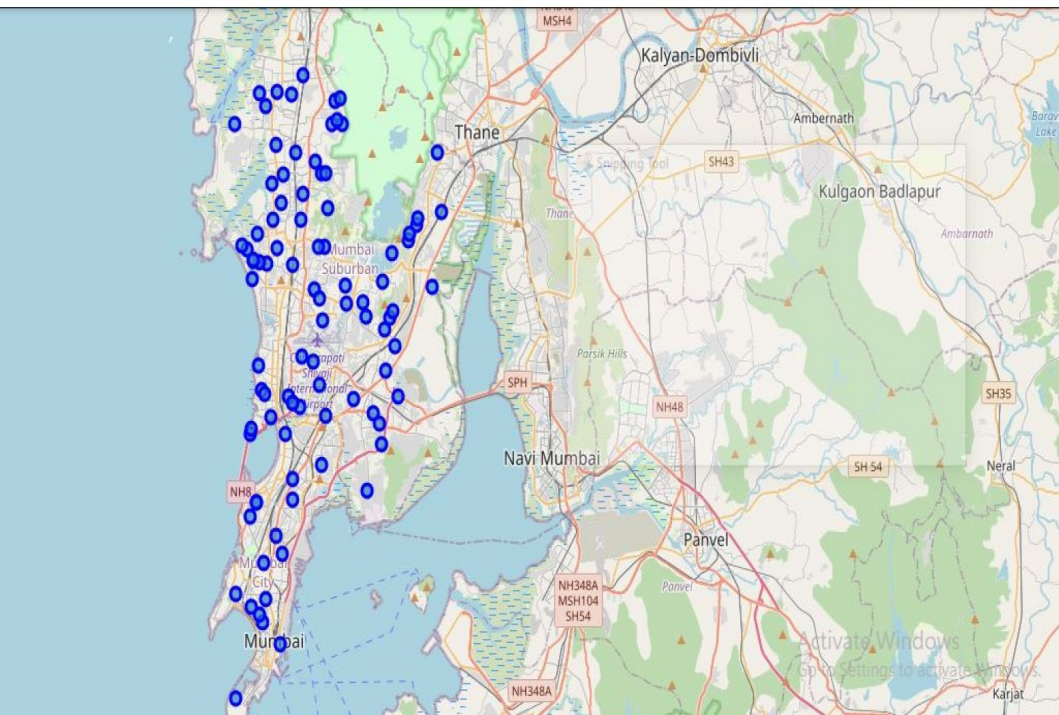
- The data that will be used in this projects is a csv file having data related to all neighborhoods in the city of Mumbai. Getting from Wikipedia page of Mumbai neighborhoods
- Finding a suitable location for a coffee shop.
- We explore the neighborhoods using Foursquare API to find the avenues within 1000 meters of each neighborhood.
- The Foursquare API that will be used to explore the neighborhoods is <https://api.foursquare.com/v2/venues/explore>.
- This API returns json response which will be transformed into a Data Frame, taking only the required details into consideration.

METHODOLOGY

- Using the Foursquare API venue information is obtained in nearby vicinity of locations in Mumbai.
- A radius is set to cover large neighborhoods in a particular area in Mumbai.
- **Getting lag lat data from** geolocator.
- This data is then merge with the neighborhoods dataset and a clustering algorithm is applied to the data.
- K-Means Clustering : The data points are clustered into 4 clusters using K-Means algorithm.

RESULTS

- By exploring the requirements we found only two neighborhoods that match the requirements.



CONCLUSION & DISCUSSION

- According to results we observe that most common venues (Top 10) come out to be restaurants and café, coffee shop and like Snack places, which means any new chef/business man can start a restaurant provided that they need to compete with existing restaurants (Other Types) but only if he provides top class facilities to get to top.
- Based on the both Clusters formed it would be a good idea to open a new coffee shop in Clusters 0 or 1 & 2 since the other clusters already have coffee shop and Restaurants in their vicinities. cluster 0 and "NAN" value having no competition has possibility is there people are not interested at coffee shop.
- cluster 1 & 2 have many restaurants in the vicinity (Pizza restaurants, European, Intercontinental, Indian) so one will be able to attract a good crowd.