

# Café Recommendation

By Jaynik Gaglani

12 - 02 - 2020

# Introduction:

- For business to be successful, the location of the store matters a lot.
- A place with many businesses of the same kind may result into less profit or even a loss due to high competition.
- Using machine learning algorithms, we try to find out the best place to open up a café based on the number of cafés present in that area in the city of Mumbai.
- The system can also be used for other businesses too by just a change of parameter.

# Data Acquisition and Cleaning:

- The data of neighborhoods of Mumbai has been scraped from a Wikipedia page.

| Neighborhood |   |
|--------------|---|
| 0            | Western Suburbs\n\nAndheri\nBhayandar\nBandra |
| 1            | Andheri                                       |
| 2            | Bhayandar                                     |
| 3            | Bandra  |
| 4            | Borivali                                      |
| 5            | Dahisar                                       |
| 6            | Goregaon                                      |
| 7            | Jogeshwari                                    |
| 8            | Juhu  |
| 9            | Kandivali west                                |
| 10           | Kandivali east                                |

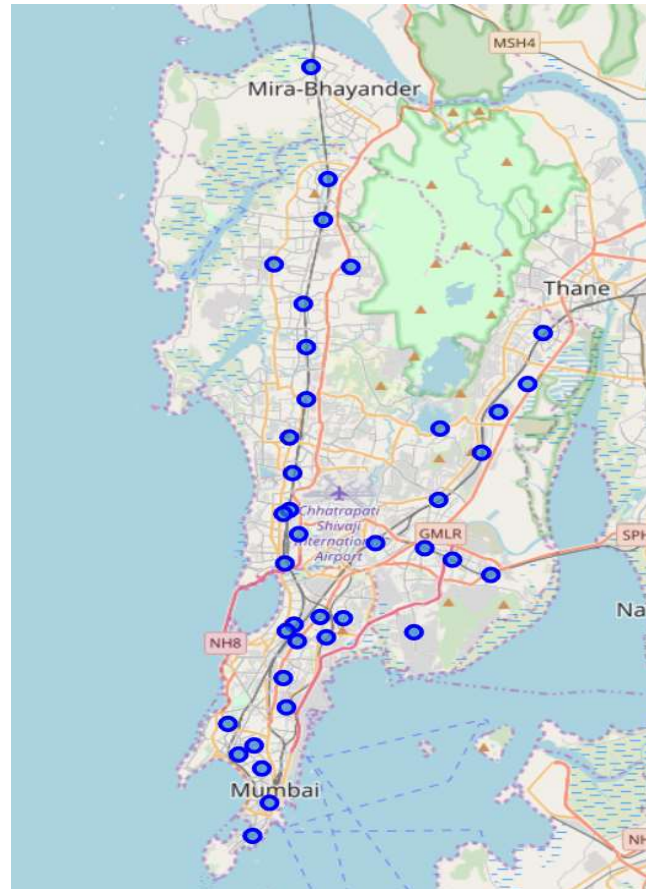
# Data Acquisition and Cleaning:

- The data is cleaned by removing all the unnecessary rows from the data.
- After removing the rows, the latitude and longitude of each location is found and added to the data to form the final dataframe.

Out[16]:

|   | Neighborhood | Latitude  | Longitude |
|---|--------------|-----------|-----------|
| 0 | Andheri      | 19.118459 | 72.841763 |
| 1 | Bhayandar    | 19.307430 | 72.851840 |
| 2 | Bandra       | 19.054370 | 72.840170 |
| 3 | Borivali     | 19.229360 | 72.857510 |
| 4 | Dahisar      | 19.250030 | 72.859070 |
| 5 | Goregaon     | 19.164550 | 72.849460 |
| 6 | Jogeshwari   | 19.137920 | 72.849410 |
| 7 | Juhu         | 19.014920 | 72.845220 |

# Map of Mumbai Neighborhoods:



# Data Cleaning:

- The venue categories for each neighborhood are found out with the help of Foursquare API.
- A new dataframe is created consisting of the venue along with the neighborhoods.

Out[21]:

|   | Neighborhood | Latitude  | Longitude | VenueName                    | VenueLatitude | VenueLongitude | VenueCategory     |
|---|--------------|-----------|-----------|------------------------------|---------------|----------------|-------------------|
| 0 | Andheri      | 19.118459 | 72.841763 | Merwans Cake shop            | 19.119300     | 72.845418      | Bakery            |
| 1 | Andheri      | 19.118459 | 72.841763 | Radha Krishna Veg Restaurant | 19.115130     | 72.843060      | Indian Restaurant |
| 2 | Andheri      | 19.118459 | 72.841763 | Naturals                     | 19.111204     | 72.837255      | Ice Cream Shop    |
| 3 | Andheri      | 19.118459 | 72.841763 | Narayan Sandwich             | 19.121398     | 72.850270      | Sandwich Place    |
| 4 | Andheri      | 19.118459 | 72.841763 | Temptations                  | 19.113767     | 72.841337      | Ice Cream Shop    |

# Data Selection:

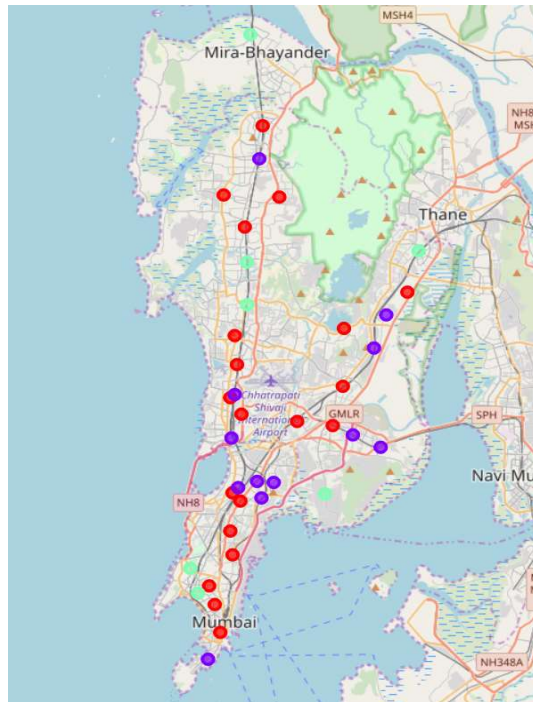
- From all the venue categories, café is selected.
- The mean of number of cafes in a neighborhood is found out.
- The dataframe for neighbourhood and café is created for clustering.

Out[33]:

|   | Neighborhoods | Café     |
|---|---------------|----------|
| 0 | Andheri       | 0.040000 |
| 1 | Antop Hill    | 0.064103 |
| 2 | Bandra        | 0.100000 |
| 3 | Bhandup       | 0.040000 |
| 4 | Bhayandar     | 0.000000 |

# Clustering:

- Number of clusters selected is three.
- The following is obtained on clustering the data:





## Observation & Results:

- The light blue points are the ones with least number of cafes present.
- The red points are the ones with moderate number of cafes present.
- The dark blue points are the ones with highest number of cafes present
- Hence, the café must be opened where there are very less or moderate number of cafes.
- This will help in reducing the competition and increasing profits. It must also be noted that other factors like locality, popularity of the place must also be considered before opening of a café.

Thank You!