

# IBM APPLIED DATA SCIENCE CAPSTONE PROJECT

## The Battle of the Neighbourhoods

Title: *Setting up a coffee shop in  
Mumbai*

Submitted by Aayushi Gupta

# *INTRODUCTION*

- This is a part of the capstone project offered by IBM Applied Data Science Course on Coursera. In this project I have selected Mumbai as my target city. Mumbai is one of the most populous cities and also the financial, commercial and entertainment capital of India. The tremendous opportunities offered by this city attracts a lot of youth here.
- Coffee is a beverage that never goes out of style and is consumed daily by millions of people specially the youth. Moreover, coffee shops are a great place for people to sit and relax or get their work done. When set up in the right location it can be a great business. One can always add some creativity to make the shop stand out amongst others. Therefore, I decided to pick a coffee shop business in Mumbai for my project.



# BUSINESS PROBLEM

The aim of this project is to find a suitable location in the city of Mumbai to set up a coffee shop business. Now, there are various factors to be considered when trying to set up a coffee shop like the cost of living of that place, the competition in the market, the population there, the kind of neighborhood of that location etc. In this project I have solved the problem using two parameters-

- The competition in the market
- The neighborhood i.e. the kind of places present there like bookstores, malls etc.



A location having places like bookstores, banks, shops and movie theaters is more likely to offer a larger number of customers. This is because people like to have a cup of coffee while reading a book, during a movie or while waiting at the shops or bank. This is why analyzing the neighborhood becomes an important factor while deciding on the location. So keeping the competition and the kind of neighborhood in mind, I have tried to formulate a solution using datasets from Wikipedia and Foursquare API and machine learning algorithms like K-means clustering.



# TARGET AUDIENCE

The target audience of this project is anyone who is looking to set up a coffee shop in the city of Mumbai. It can act as the main business or even as an extra source of income. So it can also be useful to people looking for some extra money. This project will be helpful in narrowing down the choices of locations which will benefit the entrepreneur immensely as location is one of the major factors to be considered while setting up the shop.



# DATA SECTION

- List of neighborhoods in Mumbai

Data source: The dataset was obtained from Wikipedia:

[https://en.wikipedia.org/wiki/List\\_of\\_neighbourhoods\\_in\\_Mumbai](https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai)

Data was scraped from the web using BeautifulSoup and extracted into a dataframe.

- The location co-ordinates (Latitudes and Longitudes)

Data Source: The data was already present in the dataset extracted from Wikipedia.

- The venues present in Mumbai

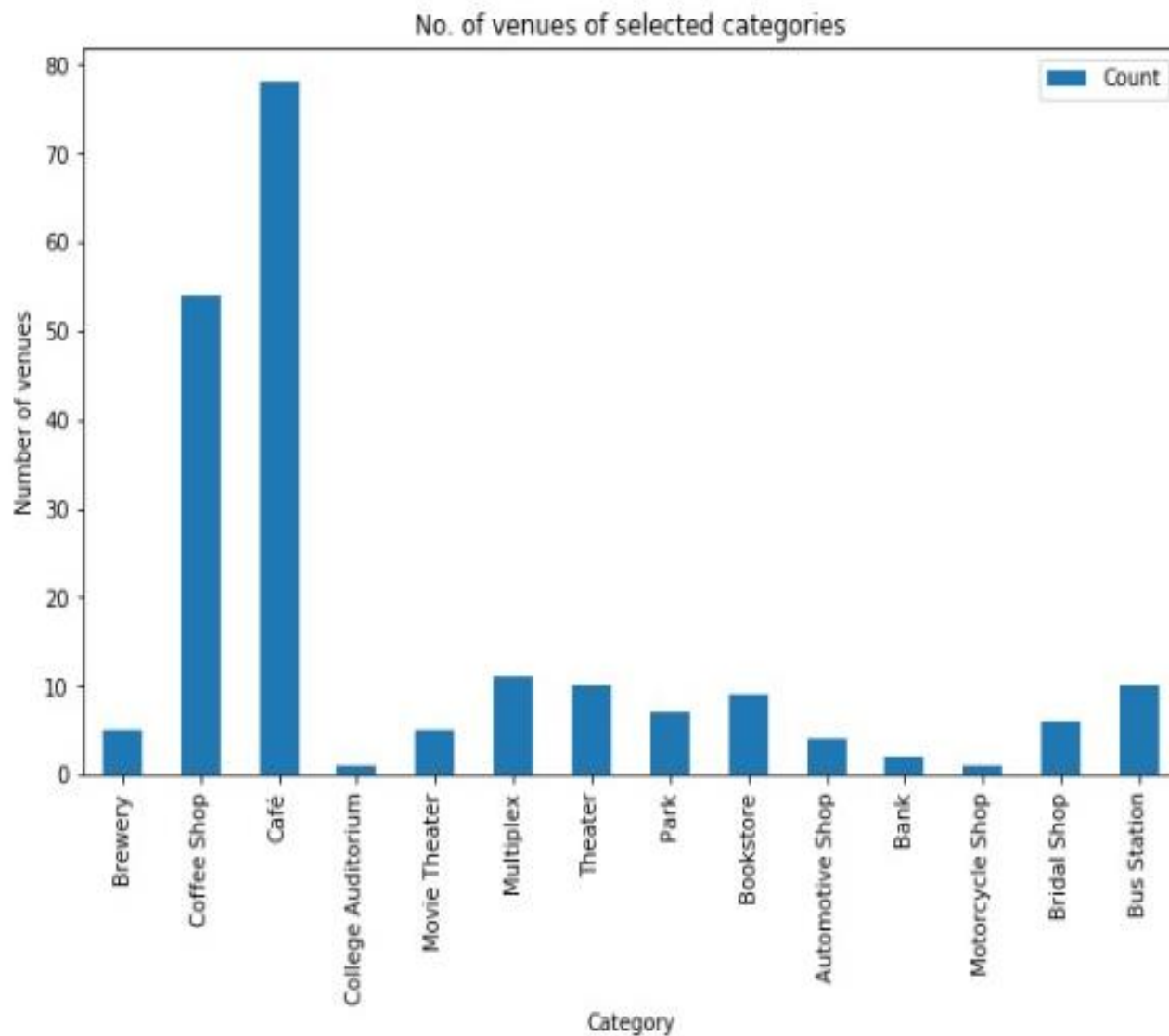
Data Source: Foursquare API



# METHODOLOGY

- **Exploratory data analysis used-** *Groupby(), describe()*
- **Inferential Statistics used-** *Bar Chart*
- **Machine Learning Algorithm used-** *K-Means Clustering*
- ❖ The data was first scraped from the web and extracted into a dataframe.
- ❖ Foursquare API was used to fetch venue data.
- ❖ Groupby() function was used to find number of venues in each category.







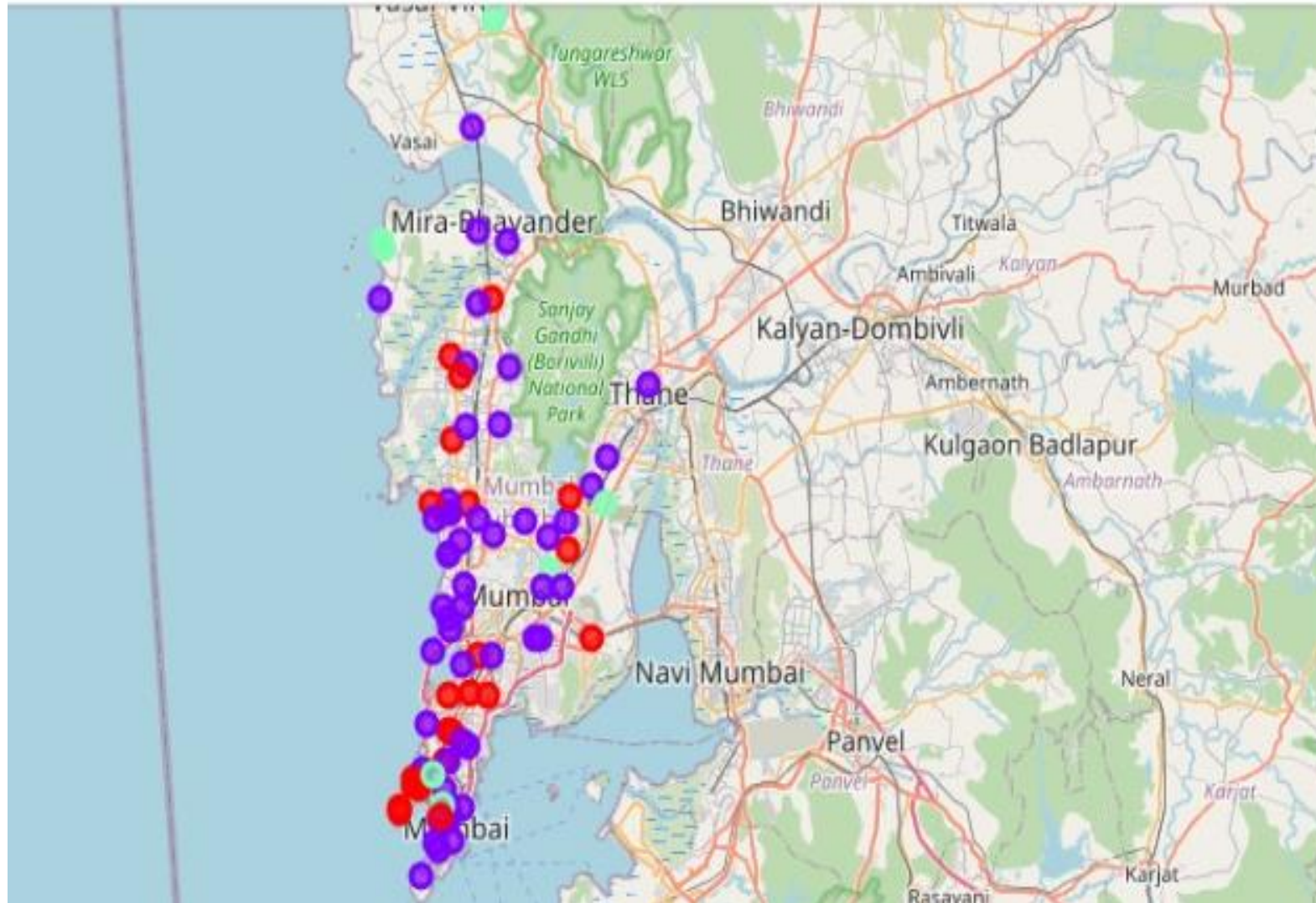
- ❖ Then the venue data was grouped by neighbourhood and K-Means clustering was performed.
- ❖ The clusters were displayed on the map of Mumbai using Folium.
- ❖ And finally the clusters were examined to find a solution to the business problem.



EN-1-1-Types.ipyn X

own v Submit Notebook ...

Python O



Support

# RESULTS

- Cluster 0 has an average competition 0.22 and favorable venues 0.03
- Cluster 1 has an average competition 0.05 and favorable venues 0.03
- Cluster 2 has an average competition 0.00 and favorable venues 0.23



# DISCUSSIONS

- Mumbai is the financial, commercial and entertainment capital of India which makes it the perfect choice for setting up a coffee shop business. It's immense population is a plus point. Through this project I have tried to examine some factors that can affect the coffee shop business. A location having less competition and more venues to attract people to the shop would be the ideal spot.
- The bar graph clearly showed that Mumbai is a happening place as it is full of different kind of places that attract foot traffic with coffee shop topping the chart. This shows great possibility of a successful coffee business. The statistics returned by the cluster showed promising results. A cluster was identified to be a possible set of ideal locations.



# CONCLUSION

- Cluster 0 has the highest competition.
- Cluster 1 has less competition but the number of favorable venues is also less.
- Cluster 2 has negligible competition and moreover has a fair amount of favorable venues to bring customers into the shop.

So a preferred neighborhood to set up a coffee shop would be one present in Cluster 2.



Selecting a location is one of the major factors when setting up the business. However, there are other factors too which need to be taken into consideration for example the cost of living in that area, the population of that area, finding the space that suits the business needs etc. Though this project doesn't cover every factor but does provide an useful insight which can help the entrepreneurs.



*THANK YOU*

