

A wide-angle photograph of the Mumbai skyline at sunset. The sky is a mix of orange, yellow, and grey clouds. The city is filled with numerous high-rise buildings, many of which are under construction, with cranes visible on top. The buildings are reflected in the calm water of the harbor in the foreground. A few small boats are visible on the water.

# Coursera Capstone

## IBM Applied Data Science

# Battle of Cafes - Mumbai

by  
Shashank Wankhede



# Introduction

Coffee being a most consumed beverage in the world. This project is for those who love coffee much and always wanted to explore coffee shops around cities. Basically it will help tourists/coffee to explore nearby coffee shops and also people who want to open a new coffee shop and looking for a less or more crowded place in Mumbai. With the help of data science we can group places or neighborhoods according to number of cafes available also with advanced business analysis we can advice any one who wants to open a new coffee shop



## **Business Problem:**

The objective of this capstone project is to guide tourists so that they can stay nearby most of coffee shops or explore most of the coffee shops and also to analyse a proper location to open new coffee shop with less competition around using the power of data and machine learning



# Data Description

To solve the problem, we will need the following data:

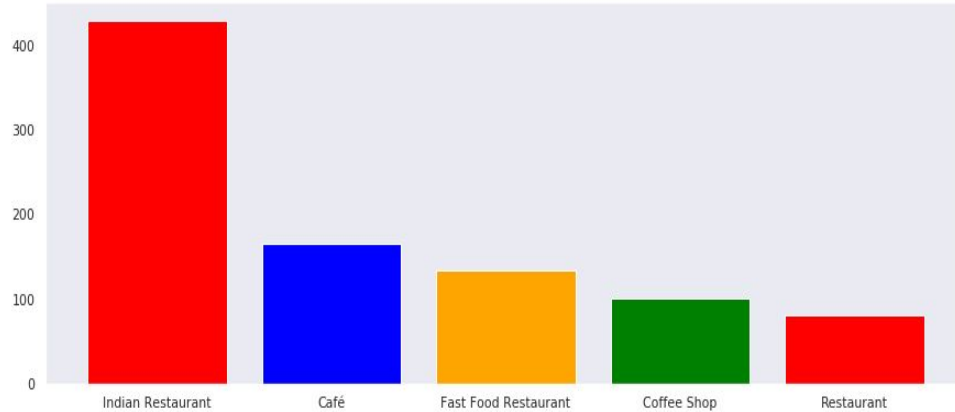
- List of neighbourhoods in Mumbai .This defines the scope of this project which is confined to the city of Mumbai
- Latitude and longitude coordinates of those neighbourhoods.This is required in order to plot the map and also to get the venue data.
- Venue data, particularly data related to venues.



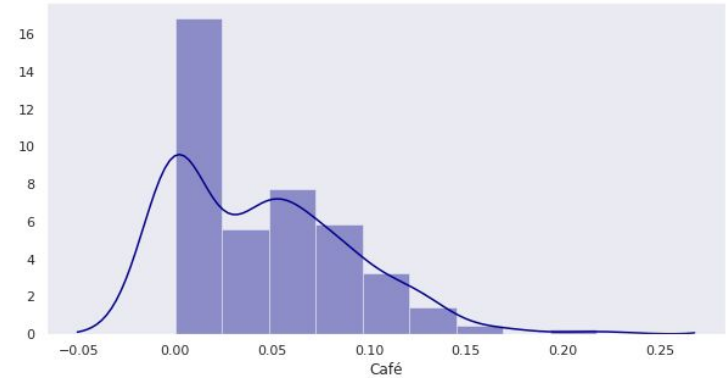
# Dataset

	Index	Pincode	Neighborhood	Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Category
0	0	400001		18.935293	72.835751	Starbucks	18.932190	72.833959	Coffee Shop
1	1	400001		18.935293	72.835751	Taste of Kerala	18.934205	72.833215	Indian Restaurant
2	2	400001		18.935293	72.835751	Mahesh Lunch Home	18.934121	72.833821	Indian Restaurant
3	3	400001		18.935293	72.835751	Yazdani Bakery	18.933191	72.833591	Bakery
4	4	400001		18.935293	72.835751	Café Universal	18.936021	72.837453	Irani Cafe
5	5	400001		18.935293	72.835751	Ideal Corner	18.934961	72.834050	Indian Restaurant
6	6	400001		18.935293	72.835751	Sher-E-Punjab	18.937944	72.837853	Indian Restaurant
7	7	400001		18.935293	72.835751	Pratap Lunch Home	18.933605	72.832854	Seafood Restaurant
8	8	400001		18.935293	72.835751	Cafe Excelsior	18.937701	72.833566	Café
9	9	400001		18.935293	72.835751	Britannia & Co.	18.934683	72.840183	Parsi Restaurant

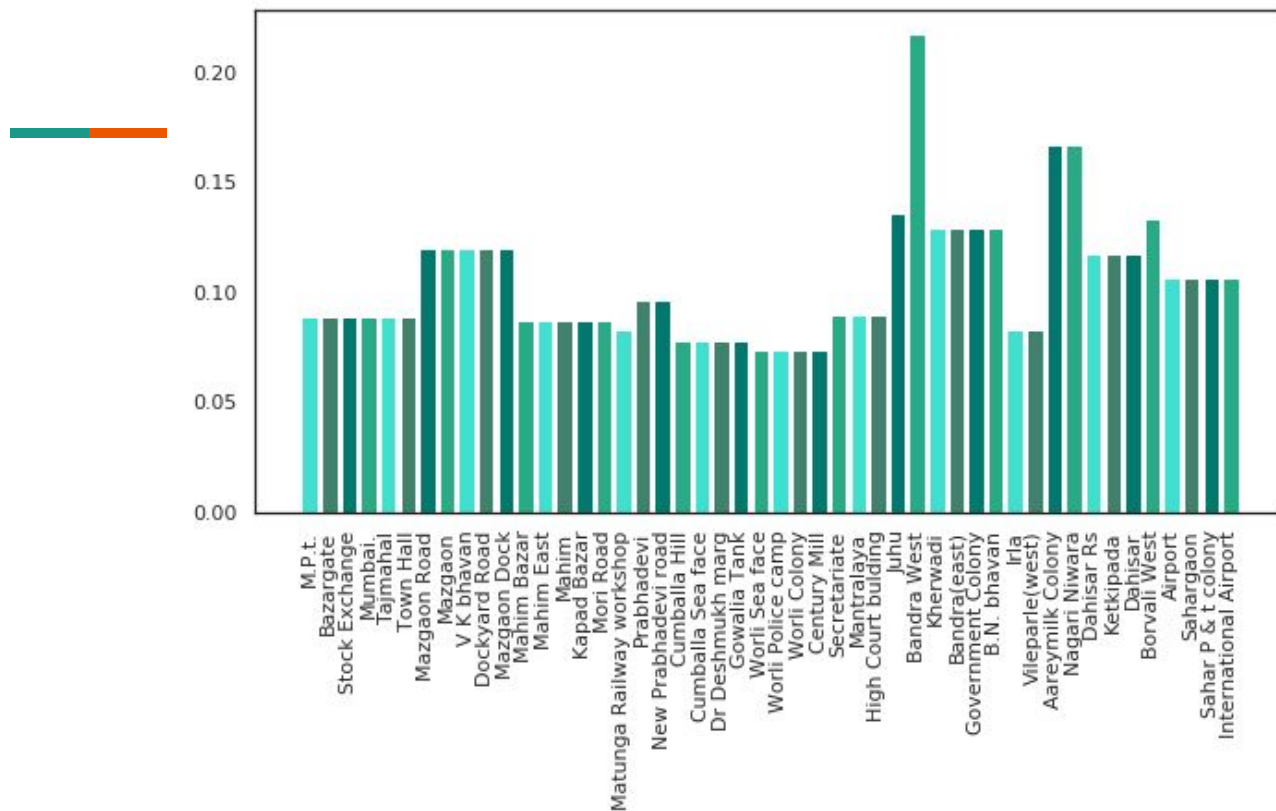
# Exploratory Data Analysis:



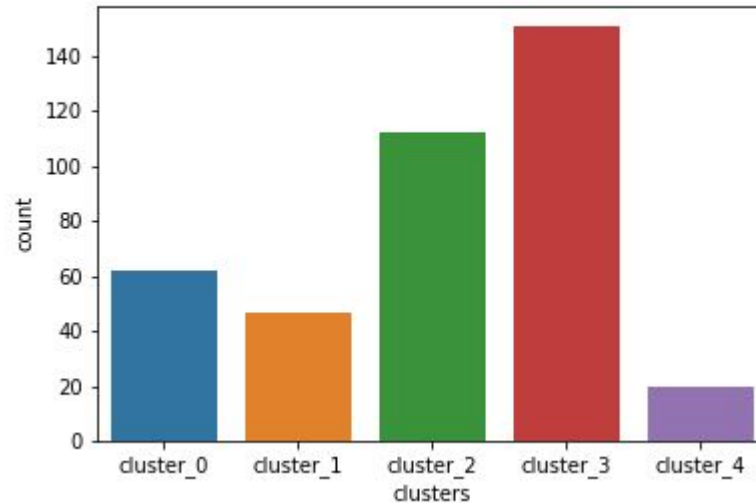
Count of Categories



Frequency Distribution of Cafes



Neighborhoods with frequency



### Cluster Count

- from the above figure we can say that cluster 3 has more cafes
- Neighborhood belonging to cluster 3 has frequency 0.08



# Conclusion:

- Neighborhood Juhu is repeated as because there are more cafes and also the cafe called Grub shop is also repeated as it is in radius of more than one neighborhood that is why location is same . Which means that if you are in Juhu you may find many cafes around you and if you are in Grub shop you are near to these Neighborhoods
- cluster 1 has 0.1 we means that in cluster 1 no of Neighborhood are less but in that small area there are more cafes its concentrated in small area on the other hand Neighborhoods in cluster 3 are widely spread over the region or mumbai
- so we can conclude that if you want to explore more cafes in less time you should choose neighborhoods belonging to cluster 2 and 3
- also if you want to open a new cafe consider cluster 4 and cluster 0 neighborhoods where compitation seems less



Project by shashank wankhede