Data for Capstone Project

- As we are building this project from a sketch, the most basic step is to collect the data regarding neighbourhoods of Mumbai and their respective pin codes.

- We have collected the basic data from here and copied it in to an excel file.

- We will use python's pandas library to read this data and store in to a data frame as shown in the figure below.

```
In [54]: df=pd.read_excel('Mumbai_Data.xlsx')
         df.head()
```

Out[54]:

	Neighborhood	PostalCode
0	August Kranti Marg	400036
1	Aarey Milk Colony	400065
2	Andheri (East)	400069
3	Andheri (West)	400058
4	Antop Hill	400037

Figure 1 Primary Data

- Now to add location data like Longitude and Latitude, we will use geocoder package of Python and prepare a data frame with features like: pin codes, neighbourhoods, longitude and latitude as shown in the figure below.

```
In [63]: df.head(10)
```

Out[63]:

	Neighborhood	PostalCode	Latitude	Longitude
0	August Kranti Marg	400036	18.964005	72.807983
1	Aarey Milk Colony	400065	19.161085	72.884394
2	Andheri (East)	400069	19.119298	72.851100
3	Andheri (West)	400058	19.122935	72.840610
4	Antop Hill	400037	19.020313	72.868280
5	Anu Shakti Nagar	400094	19.033945	72.925200
6	B A R C	400085	19.016345	72.926988
7	Ballard Estate	400038	18.940170	72.834830
8	Bandra (East)	400051	19.060715	72.854564
9	Bandra (West)	400050	19.052259	72.829405

Figure 2 Data with Longitude and Latitude

Four Square API

- Four Square API is a power full tool to retrieve location data for any location. This includes various venues by their categories, their exact location, user reviews, recommendations and other venue specific details.
- We will use this opportunity and get list of venues with their location and category for each of the neighbourhoods.
- We will then apply exploratory data analysis techniques to ultimately prepare a data frame with each neighbourhood and their top 10 most frequent venues category wise.
- This data frame will be used for carrying out clustering and analysis of different clusters.

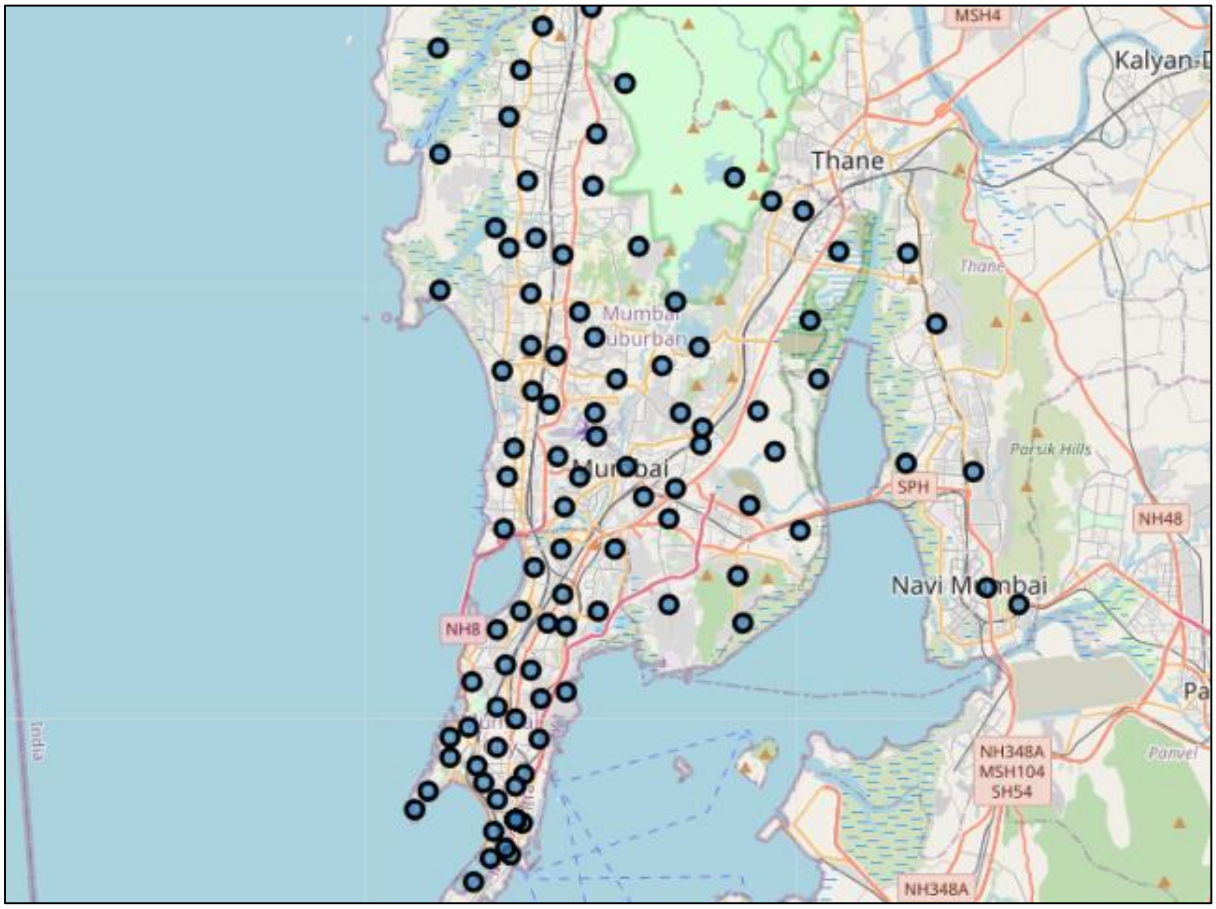

Figure 3 Map of Mumbai