RENTAL HOUSE BUSINESS IN MUMBAI CITY

1. INTRODUCTION:

In this report we will discuss about the House Rentals Business in Mumbai city. As Mumbai is one of the metropolitan cities on India there is a lot of migration of people from other regions towards this small city. House accommodation is quite difficult in this city and the house prices go through the roof. Hence rental houses are life saver to many common people and are generally high in demand in this city. We will try to study the area of our interest for figuring out best location candidates for setting up houses for rental services.

- Problem

  It is very tedious work for a person looking renting houses in the city. He/she must look for various aspects such as house what is the carpet area? Is the area safe? Are there any public commutes nearby? Are there any convenience stores, grocery stores nearby, Is the place closer to his/her workplace? and much more. They must see to it that all the members of family are comfortable in the new environment. Are there facilities in their vicinity they might be requiring? So we work on the problem of locating an area where estate properties will be fit for House Rental Business.

- Interest

  This report is mainly targeted to the Rental House Property Business stakeholders who are willing to purchase estates in Mumbai city, India. It will be useful for stakeholders who would make a pool in high investment since house rentals comprises of buying a big estate to make profits in good return. This report will be also useful for people who are looking for rental houses.

2. DATA INTEPRETATION:

- ## Source

  The data set for this project was acquired from [MagicBricks](#) website. MagicBricks is an online House Rental Service platform where you can buy, sell, rent properties. The dataset used for the project is dated for January 2020.

- ## Data Cleaning

  i.   The data we require for now would be for renting out houses i.e. apartments, small row houses. Hence, we it was needed to clear out other data instances comprising of penthouses, condominium, and other extravagantly big houses.

  ii. There were many instances wherein the apartments were just area plots which had to be removed too.

  iii. There were features such as name of the broker, email of the broker, windows in the houses which were not necessary for our project.

- ## Features Selection

  i.   Carpet area of the house in square feet.
  ii.   No. of bathrooms in the house.
  iii. No. of bedrooms in the house.
  iv.   Name of developer of the building.
  v.   Locality of the house.
  vi.   Monthly rent of the house.
  vii. House type.
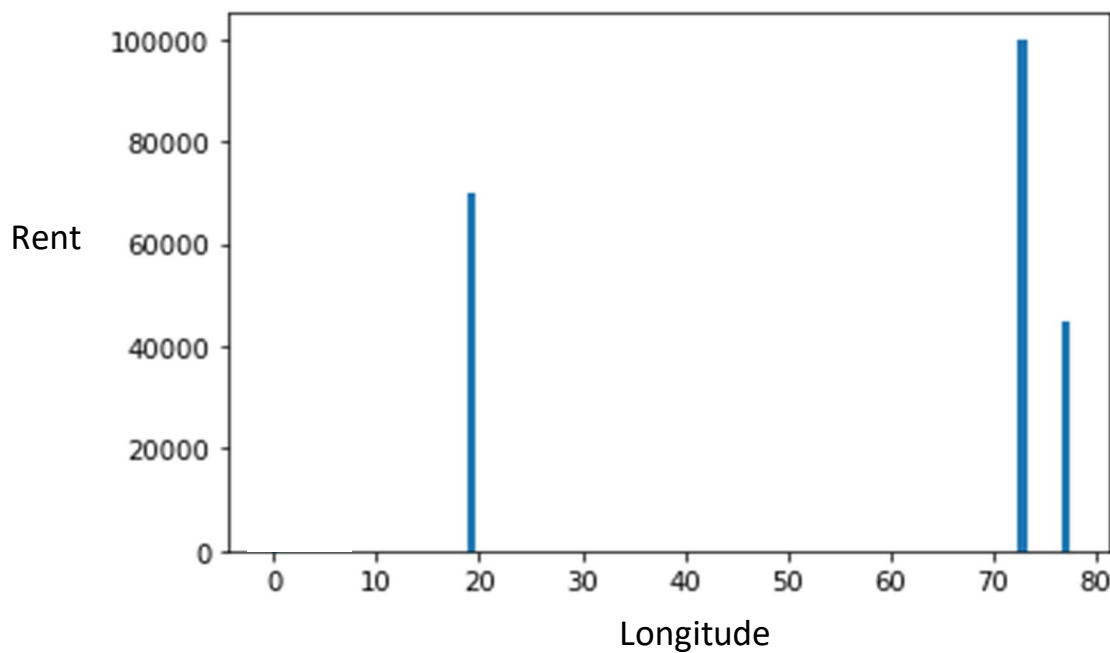
3. <u>DATA ANALYSIS:</u>

Analysis of the data for this project is done different sections of the project. We will discuss them below –

Rent

Area sq. Ft.
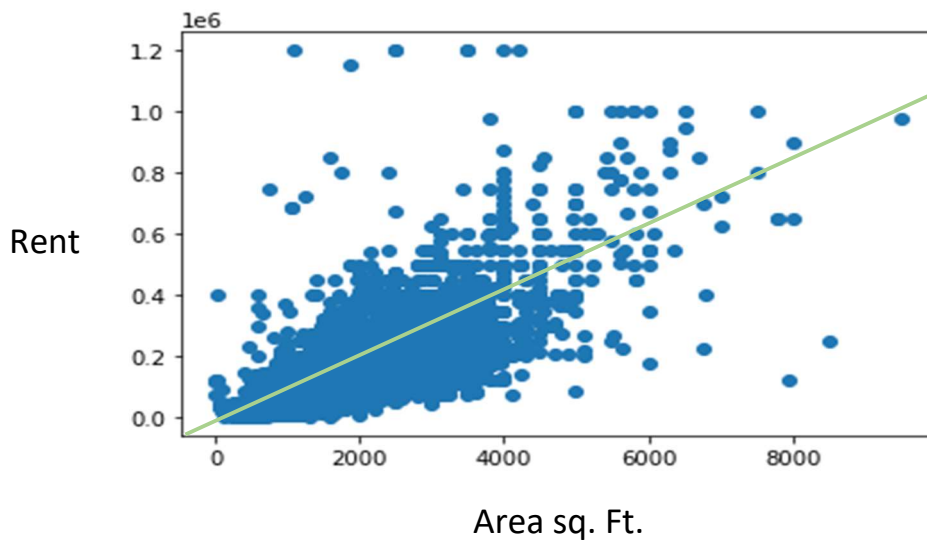
- The above relation shows that there is an uptrend in its own way. This is because even if carpet area of a house is bigger, it does not ensure higher rent price. There are many factors affecting the rent prices.
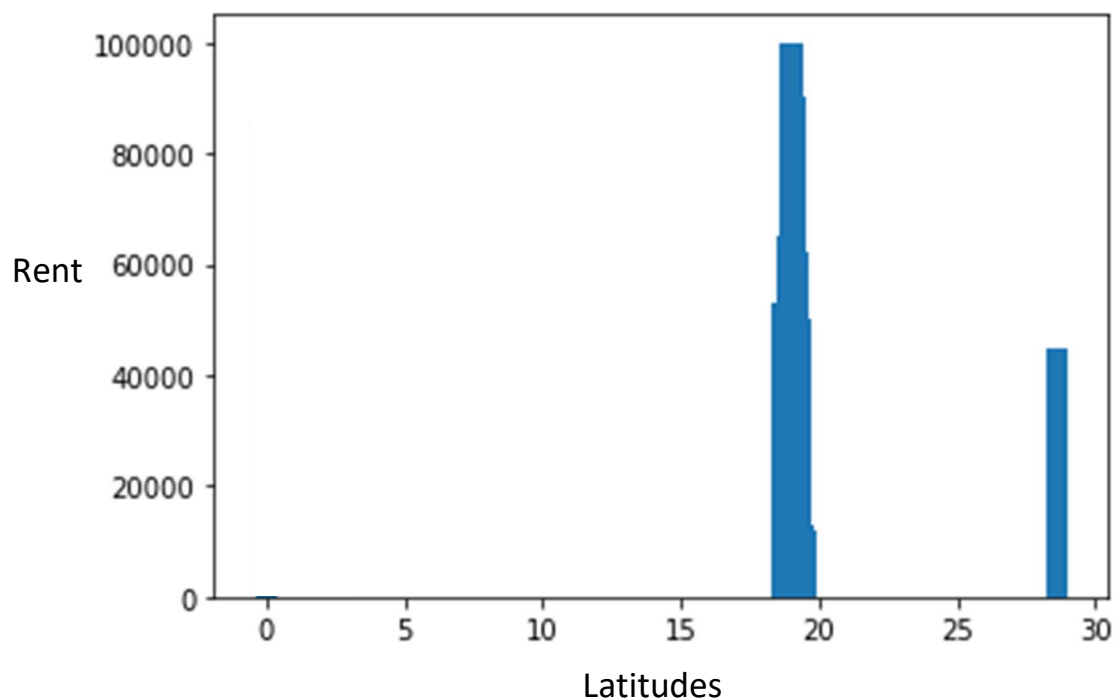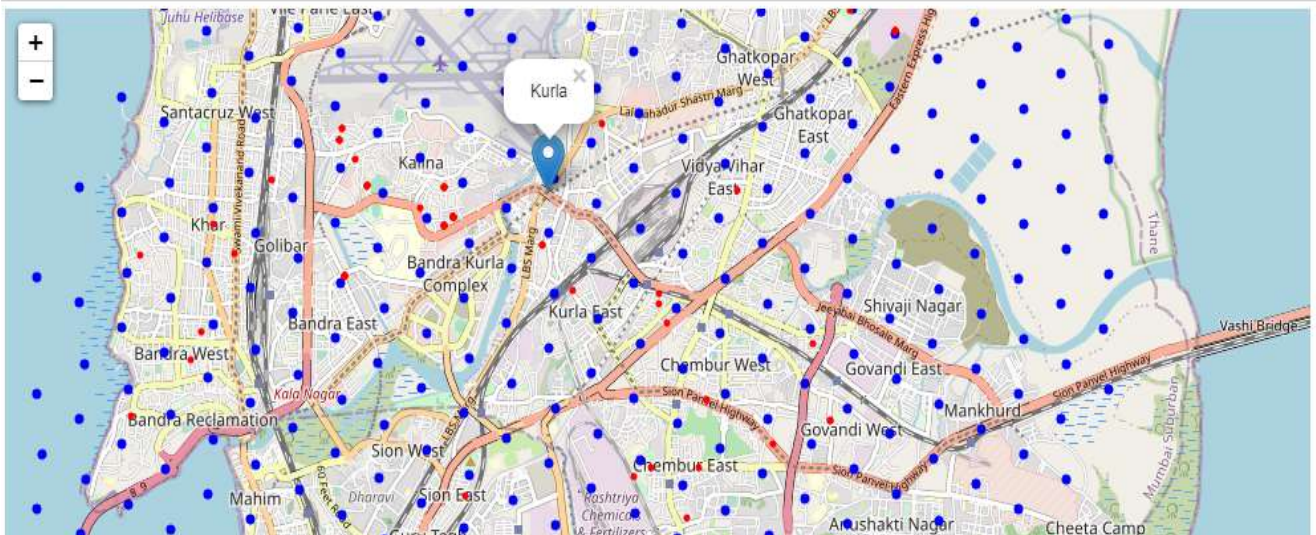

Rent

Latitudes

- In our data set we have locations of areas in terms of latitude and longitude. By plotting each parameter of location against rent prices we get the flowing graph.

- We notice that the rent prices of houses between 19~20 latitude, 72~77 longitude is the highest. Centre location of the Mumbai city is Kurla West having co-ordinates:
  longitude=72.877, latitude = 19.075.
  We use this area as our workshop for our project.

- We create a grid of location points on the map of Mumbai city using Folium. Then converting this location points into cartesian coordinates for acquiring distance of the grid point from the centre of the city. We use Google Maps API for this operation.

- We use the longitudes and latitudes of each location instance of our original data set to get venue categories in the surrounding area using Foursquare API. The API sends json response for each request and the category type for each venue needs to be filtered out. These venue categories are stored in the original data frame corresponding to their location.

```python
map_mumbai = folium.Map(location=mumbai_center, zoom_start=13)
folium.Marker(mumbai_center, popup='Kurla').add_to(map_mumbai)
for lat, lon in zip(latitudes, longitudes):
    folium.CircleMarker([lat, lon], radius=2, color='blue', fill=True, fill_color='blue', fill_opacity=1).add_to(map_mumbai)

for lat, lon in zip(list(df['latitude']), list(df['longitude'])):
    folium.CircleMarker([lat,lon], radius=1,color='red',fill=True,fill_color='red',fill_opacity=1).add_to(map_mumbai)

map_mumbai
```


- From the total collected venue categories, we pick out the items with high number of frequency and list them as popular venue categories. This list is used as a feature in the data frame as one of the factors that influence the rent pricing of houses in a region.

- Since we know that certain places like hospitals, public transport, police stations, etc are necessary we append these places directly inside most popular venue categories. Similarly, we create another list of most unpopular venue categories and use them to filter out location instances consisting of only them.

Frequency of the occurring venue category.
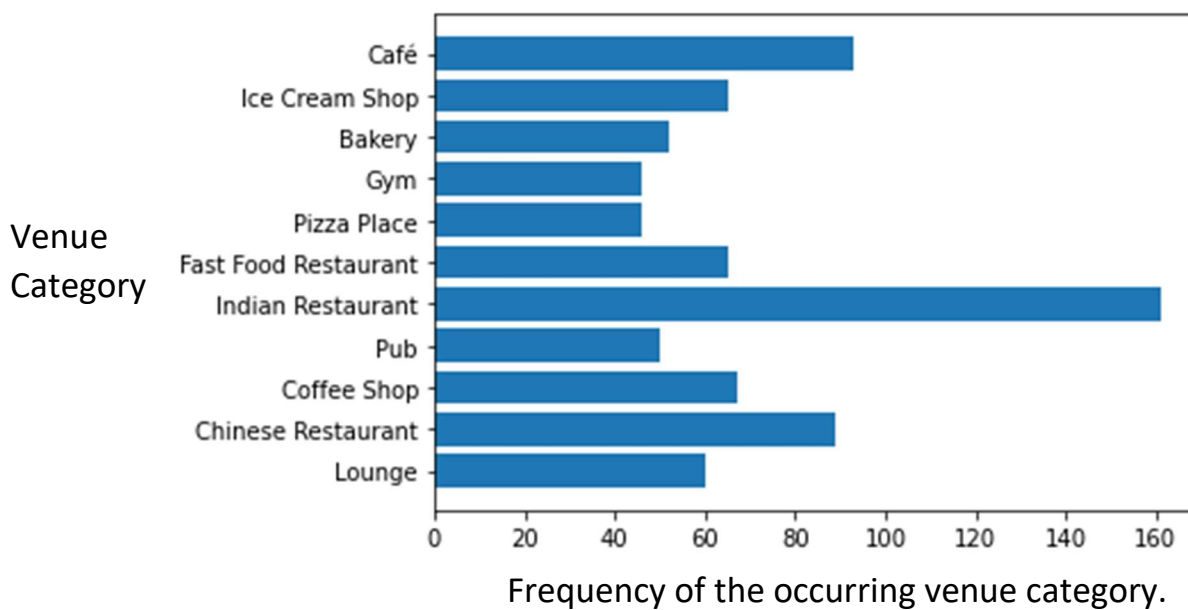
- We use Foursquare API again to get the surrounding venues of the location. We then make sure that the venues explored in this search are bit closer. We give area to check for venues a radius of 300 meters with 10 venues as the limit. Then the grid locations are then matched to their respective nearby venues.

- After this process we make use of the popular venues list we acquired from the previous data set and use them to filter our location grids. The filtered locations now consist of instances with many popular venue categories. We then plot these 63 location points on our map. We use heatmap to enhance extension of our point areas. These locations are then clustered into groups using KMeans Clustering.

4. Clustering Model:

- Previous filtration left us with 63 location candidates with high potential for Rental House Business. Now these locations point out a specific area which satisfies the popular venues condition. We need an approximate extent of an area for our business to look for estate property.
- We use KMeans clustering model to get an overall spread of our 63 pinned candidates.
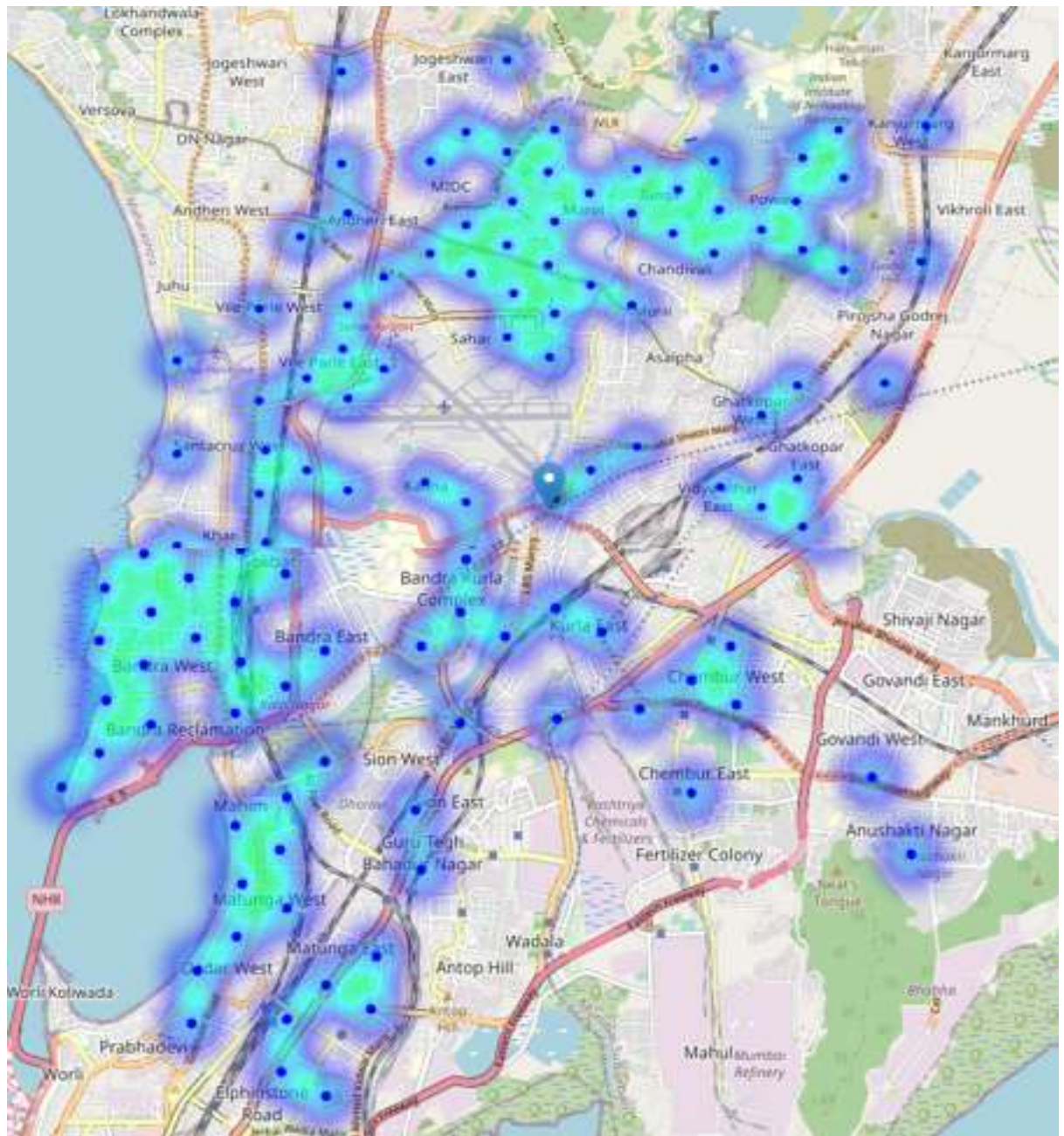
- Above are the 63 candidates for our clustering model. Data frame for this model consists of location co-ordinates of the candidates. As you can see the extent of an area is quite displayed by the heatmap in this folium map, however heatmap is relative to the scale of zoom and is not reliable.

- Therefore, clustering proves to be a good option for us in this situation. For choosing the right number of clusters for our model we plot graph for certain range of clusters and their respective amount of distortion.
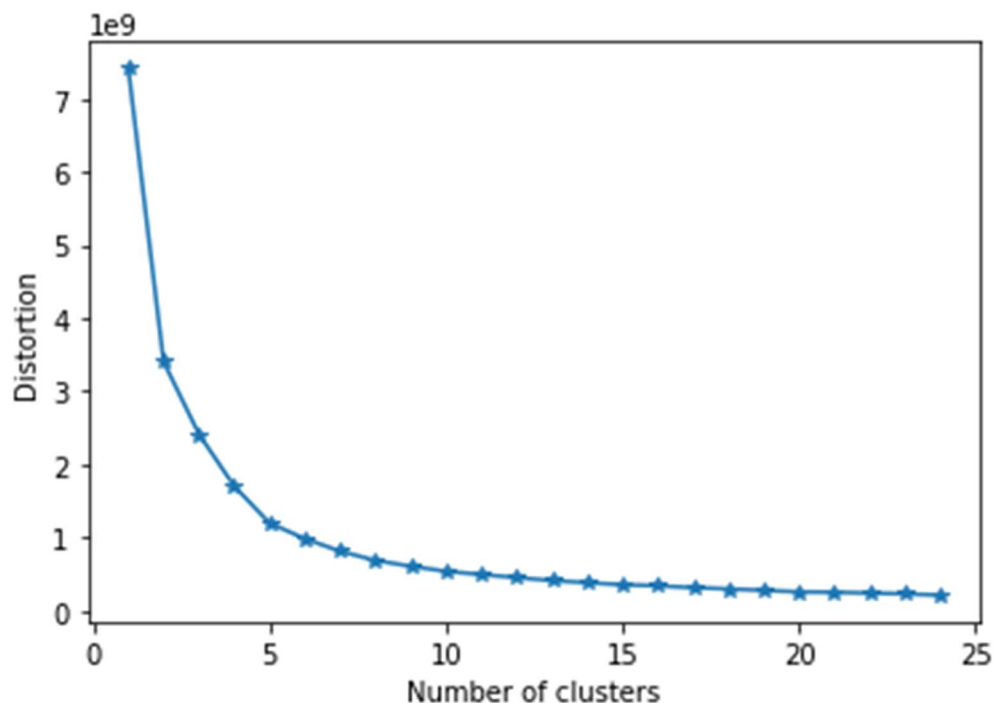
- From the above graph 15 seem to be quite good number of clusters for our model since we don't want to overfit it.

- Using no. of clusters = 15, we then plot the locations on our map with the green clusters showing potential zones for House Rental Business.

- This is the resulting display of our cluster model. These zones are can now be fetched to sweep them for estate properties by the stakeholders.
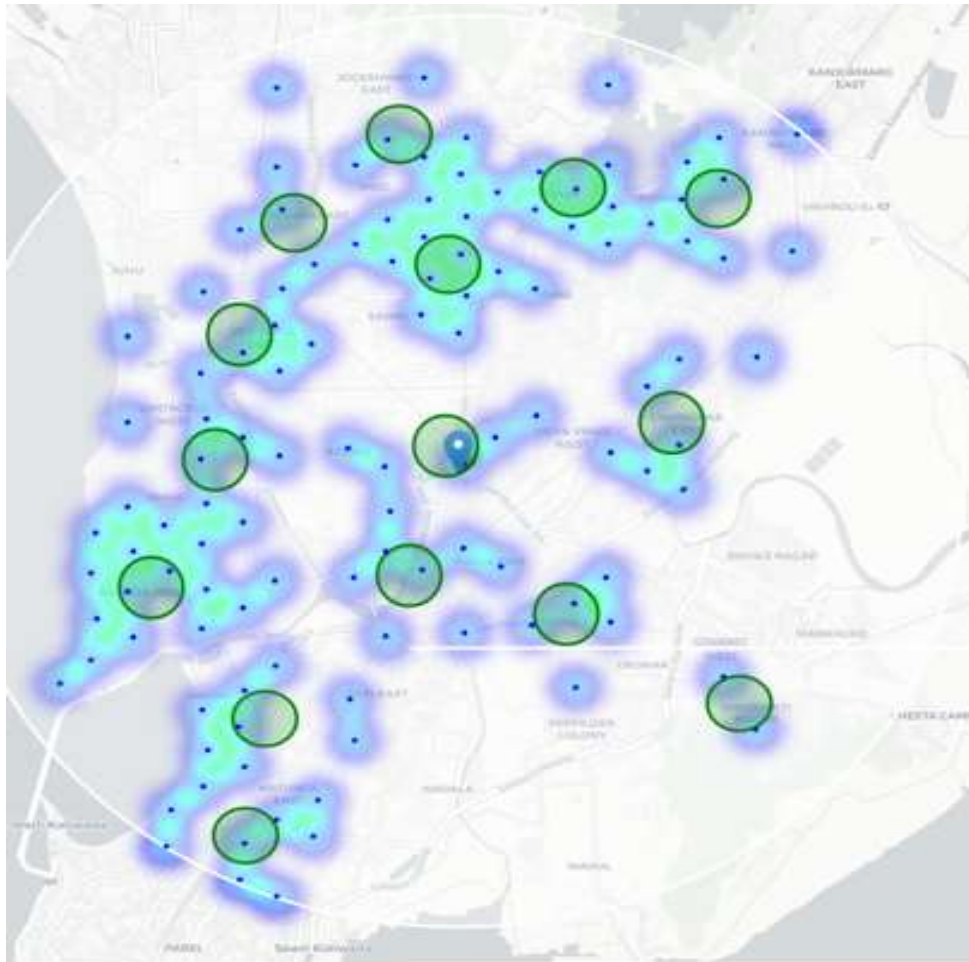
- Depending on the number of clusters, we will get the exact same number of addresses corresponding to the location of cluster and its cartesian distance from the centre of the city i.e. Kurla West, Mumbai.

```
Addresses of centers of areas recommended for further analysis:

1) plot no55, 1st FL,flat no.7.Neelam Bldg, 1st lane, Dadar East, Hindu Colony, Dadar, Mumbai, Maharashtra 400014 => 11.9km f
rom Kurla, Mumbai

2) 308/A, AAPLI EKTA SOCIETY, Navpada, Marol, Andheri East, Mumbai, Maharashtra 400059 => 5.9km from Kurla, Mumbai

3) 1582, Rd Number 4, TPS III, Sen Nagar, Santacruz East, Mumbai, Maharashtra 400055 => 6.4km from Kurla, Mumbai

4) 101-240, Babli Mahadev Kanekar Marg, Saibaba Nagar, Pant Nagar, Ghatkopar East, Mumbai, Maharashtra 400077 => 5.8km from K
urla, Mumbai

5) 2-3, KN Gaikwad Marg, Ganesh Nagar, Postal Colony, Chembur, Mumbai, Maharashtra 400071 => 4.8km from Kurla, Mumbai
```

Siddhesh Kadam   17/06/2020

5. Conclusion:

The goal of this project was to acquire list of possible location in Mumbai city for stake holders of house rentals business. These locations were chosen on basis of factors which applied to a reliable data set of rental houses i.e. MagicBricks. With the help of Google Maps API, we were able to perform operations on location of interest. We explored and collected data of a venue and used its category with the help of Foursquare API to shortlist the number instances generated by Google Maps API. This shortlisted instance was hence cultivated from features of MagicBricks data set. The number of suitable clusters were evaluated from Distortion Graph. These clusters pointed out in this study are created as major potential locations and are recommended for Rental Houses stakeholders. These locations then need to be further processed by the stakeholders themselves based on the characteristics of the candidate's locations as per their own needs.

Siddhesh Kadam   17/06/2020