**Prediction of a suitable location for opening a restaurant in Mumbai**

*By*

Ice Bear

**1.Introduction:**

* 1. **Background:**

**Mumbai** also known as **Bombay**  is the capital of the Indian state of Maharashtra. According to United Nations, as of 2018, Mumbai was the second city in India after Delhi and the seventh most populous city in the world with a population of roughly 20 million. As per Indian government population census of 2011, Mumbai was the most populous in India with an estimated city proper  population of 12.5 million living under Municipal Corporation of Greater Mumbai . Mumbai is the centre of the Mumbai Metropolitan Region, the sixth most populous metropolitan area in the world with a population of over 23 million. Mumbai is also the Financial, Commercial, and the Entertainment capital of India generating about 6.16% of India's GDP alone. So it wouldn’t be a surprise if a large corporate or even an individual would be interested in opening a restaurant or a restaurant chain here given the fact that there is a good sizable population that is providing a healthy contribution to this growing financial workhorse of India with a respectable purchasing power.

* 1. **Problem:**

With a total area of 603.4 the city has it’s fair share of economic inequality and not all areas and corners are same in terms of economic and urban growth. This makes it quite a challenge to choose locations that may increase the success of the restaurant based solely on these locations. Therefore, This project aims to mitigate this problem by providing several locations and inferring a best possible location for the head start of the restaurant.

* 1. **Interest:**

Individuals that are interested in opening a restaurant or a large corporate that wants to start a restaurant chain would be interested in this project.

**2.Data acquisition and cleaning:**

* 1. **Data Sources:**

For the Project, we first need to have a data that shows all the Suburban and Urban areas in Mumbai along with their respective postal codes which can be found [here](https://www.nativeplanet.com/india-pin-codes/maharashtra/mumbai/). After this, by web-scraping using BeautifulSoup , the data was obtained in the notebook in the form of a list but it was not in a tabular form that I expected which was needed in order to carry out analysis.

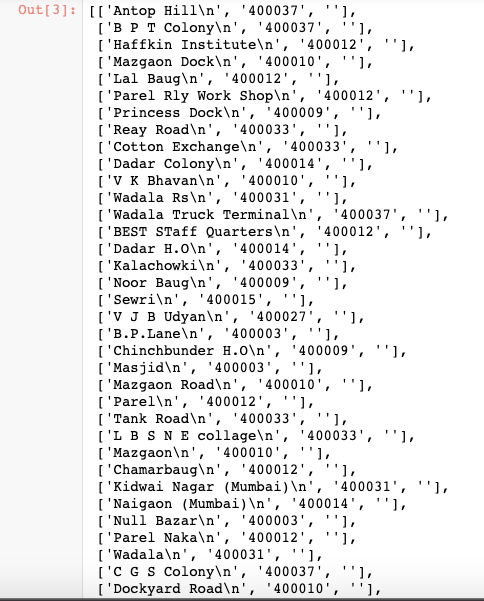
**A close up of text on a white background

Description automatically generated**

**Fig 1:Format of all Areas/Neighborhoods of Mumbai in the list**

* 1. **Data Cleansing:**

As the integers and strings were together in the form of a string in the list, the objective was to first separate them and represent them separately which was achieved by importing and using re library for splitting of strings and integers in the form of list and thus I was able to obtain a much more refined look.



**Fig 2:List within a List**

From here on, using For Loop I was able to form a separate list for both string and integers i.e a list for Postal Code and a list for Neighborhood.

A screenshot of a cell phone

Description automatically generated

**Fig 3:Separate lists for Neighborhoods and Postal Codes**

The only thing left is to arrange them in a tabular form which would then be used as a basis for further analysis.

A screenshot of a cell phone

Description automatically generated

**Fig 4: Table containing Postal Code and Area**

* 1. **Adding Features:**

After the data has been cleansed, we need to find longitude and latitude for each of our neighbors in order to carry out our analysis which can easily be obtained by using various geocoding APIs which are out there but for this experiment I have decided to use OpenCage geocoding API. So, by using the appropriate code and passing down the keys I was able to obtain the list which contains both longitude and latitude.

A screenshot of a cell phone

Description automatically generated

**Fig 5:List of Longitude and Latitude**

After this just append these into the original table so that we can move towards our analysis part.

A screenshot of a cell phone

Description automatically generated

**Fig 6:Final Base Table**

**3.Data Analysis:**

**3.1 Nearby Values:**

In order to find the suitable location for a restaurant we need to first define our ideal location. An ideal location for our restaurant startup would be where we can expect customers that have come primarily to eat in the neighborhood. Since our restaurant is new, we can’t expect our customers to go out of their way and try it out but what we can expect is that if customers are in the neighborhood to eat at their usual places and they get to see our restaurant then they may be tempted to try it out. So our primary objective should be to find neighborhoods that has food related venues as their most common venues.

For this, we need to first find all the nearby venues for each neighborhood. This was done using Foursquare API by defining a function. By passing all the required credentials of the API in the function I was able to find all the nearby venues of each neighborhood with limit set to 100 venues along with their longitude and latitude followed by the category under which they fall.

A screenshot of a cell phone

Description automatically generated

**Fig 7:Table consisting of all nearby venues**

After this, we can group each neighborhood and count the total number of venue category under each of these neighborhoods followed by finding out the frequency of each this category with which they occur. With this, the top 10 most common venues was found out for each neighborhood.

A screenshot of a cell phone

Description automatically generated

**Fig 8:Table containing the most common venues.**

**3.2 Clustering:**

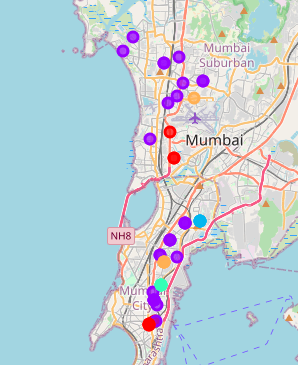
After finding out the top 10 venues for each neighborhood, we can cluster them using K-means and through each cluster point we can visualize these venues and try to find out the most optimum location based on features that each location offers and thereby narrowing down our search. For clustering, we first need to set a cluster label for each of this neighborhoods that can be seen as follows:

A screenshot of a cell phone

Description automatically generated

**Fig 9:Setting Cluster Labels for each neighborhood**

After this we can plot a map using Folium and visualize our cluster points to make it easier in narrowing down our search for selection of location.



**Fig 9:Map containing all Cluster Points**

**4.Result:**

Now since we have a clear map of all cluster points, it has now come down to observation and having a basic geographic knowledge about locations thereby aiming for factors that we would need in order to make our restaurant a success.

First we need to narrow down our search to locations that have most food related venues as their most common venues for reasons we discussed earlier and second we need our cluster point to be close to railway stations because when we talk about Mumbai, it is also famous for it’s ‘Local trains’ which are cheap and easy means of travel around the city for almost everybody and lakhs of people use it almost every single day. Now, there are 3 lines through which these local trains travel. These are Western Line, Central Line & Harbour Line. Now, We would want to go with Western Line cause it has a good mix of Urban and Suburban Neighborhoods.

**A picture containing text, map

Description automatically generated**

**Fig 10:Western Line Cluster Points**

As we can see from the map, we do have some potential locations that are relatively close to the railway line meaning they do have good connectivity. But we would want to focus on the area named ‘Vile Parle’ because when we look closely there are two cluster points that are quite close to the actual ‘Vile Parle Station’.

A close up of a map

Description automatically generated

**Fig 11:Vile Parle Cluster points**

If we look at the most common venues that these points contain then we can observe that both of them have for most food related venues meaning it can be the food locality that we were looking for. On further searching through the area we can also find out that the west side has some well-known colleges which are recognised throughout the city and country as well and there are Hostels and Rented apartments in this area.Apart from this, there is also a famous tradtional drama theatre for which many people attend. So it can be inferred that this can an ideal location for setting up our restaurant.



**Fig 12.Most Common Venues of Vile Parle**

**5.Conclusion:**

Thus, based on analysis and then further narrowing the results down based on several factors. I am able to infer that ‘Vile Parle’ might be a suitable location if one wishes to start a restaurant since it does provides an opportunity to grow and does looks promising.