



Battle of Neighborhoods: Mumbai

Analytical Approach to Indian Cuisine

Dhaval Modi

The Problem

Exploring **Mumbai** to Select Best Location to Open an Indian Cuisine Restaurant

Why Mumbai:

Selecting **Mumbai** for opening a restaurant has many merits, below are a few to start with:

- Financial Capital of India.
- Seventh most populous city in the world with 20 Million people, providing **volume**.
- A cultural melting pot for diverse Indian ethnicities, providing **variety**.

Why Do It:

Along with displaying the power of data analytics to problem solving, this project can also be of value to:

- **Entrepreneurs** who want to invest in Restaurant Business.
- **Restaurant Chains** attracted to Mumbai because of its diversity and potential.
- While the analysis presented is geared towards Indian Cuisine Restaurants, it can be expanded to any kind of Cuisine or Venue Category.

Methodology

The Approach Taken



Data Acquisition & Cleaning

Data Acquired from the Web

Neighborhood Data:

- Post Offices (PO) are located across the city, hence selecting them as representatives of each neighborhood.
- Data on PO Names and PIN codes was scrapped from:
 - pincode.india-server.com
- Data of location coordinates for POs was sourced through Geocode function.

	Post Office	Pin Code	Latitude	Longitude
0	Agripada	400011	18.9753	72.8249
1	Airport	400099	19.0901	72.8637
2	Ambewadi	400004	18.9907	72.8413
3	Andheri East	400069	19.1159	72.8542
4	Andheri	400053	19.1197	72.8464

Rental Data:

- Rent of 1 BHK apartments was taken as representative of rent values in the neighborhoods.
- Data on rent values was sourced from:
 - magicbricks.com, 99acres.com, makaan.com
- Data on popular venues and venue categories was sourced using FourSquare API.
- Resulting dataset consisted of **176 neighborhoods** from Mumbai.

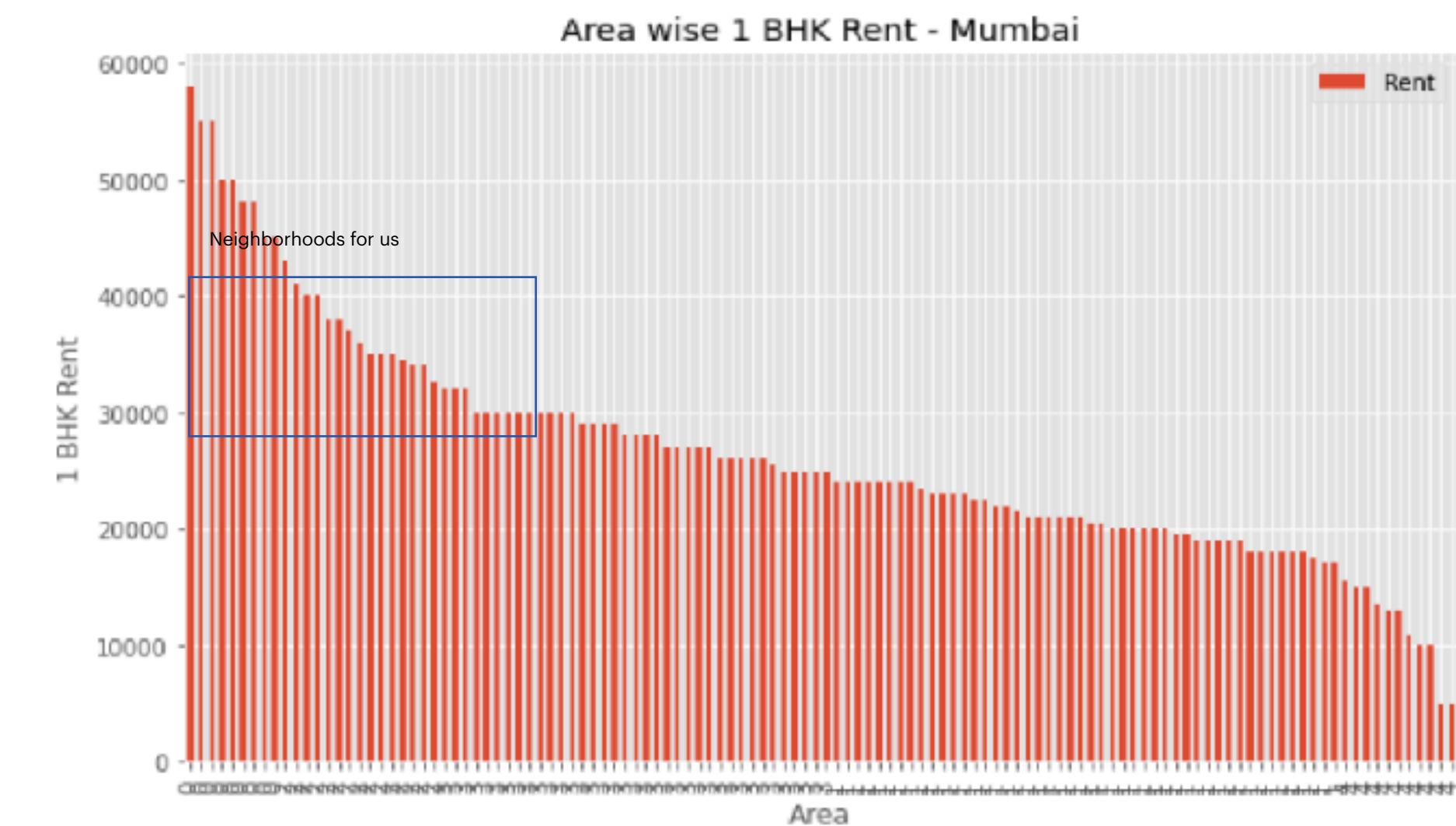
	Area	Rent
0	Agripada	28000
1	Andheri	30000
2	Andheri East	23000
3	Andheri Railway Station	13000
4	Anushakti Nagar	30000

Popular Venue Categories in Neighborhoods

Exploring Data

- Using FourSquare API, top 10 venue categories were sourced for each neighborhood, from a dataset of **4065 venue categories**.
- Neighborhoods were selected where “Indian Cuisines” related venue categories featured in top 3, retaining **115 Neighborhoods**.
- Rent distribution of 1 BHK apartment across neighborhoods was explored.
- Focus was kept on “upper middle class” neighborhoods, with rent ranging from **INR 30K to 40K**.

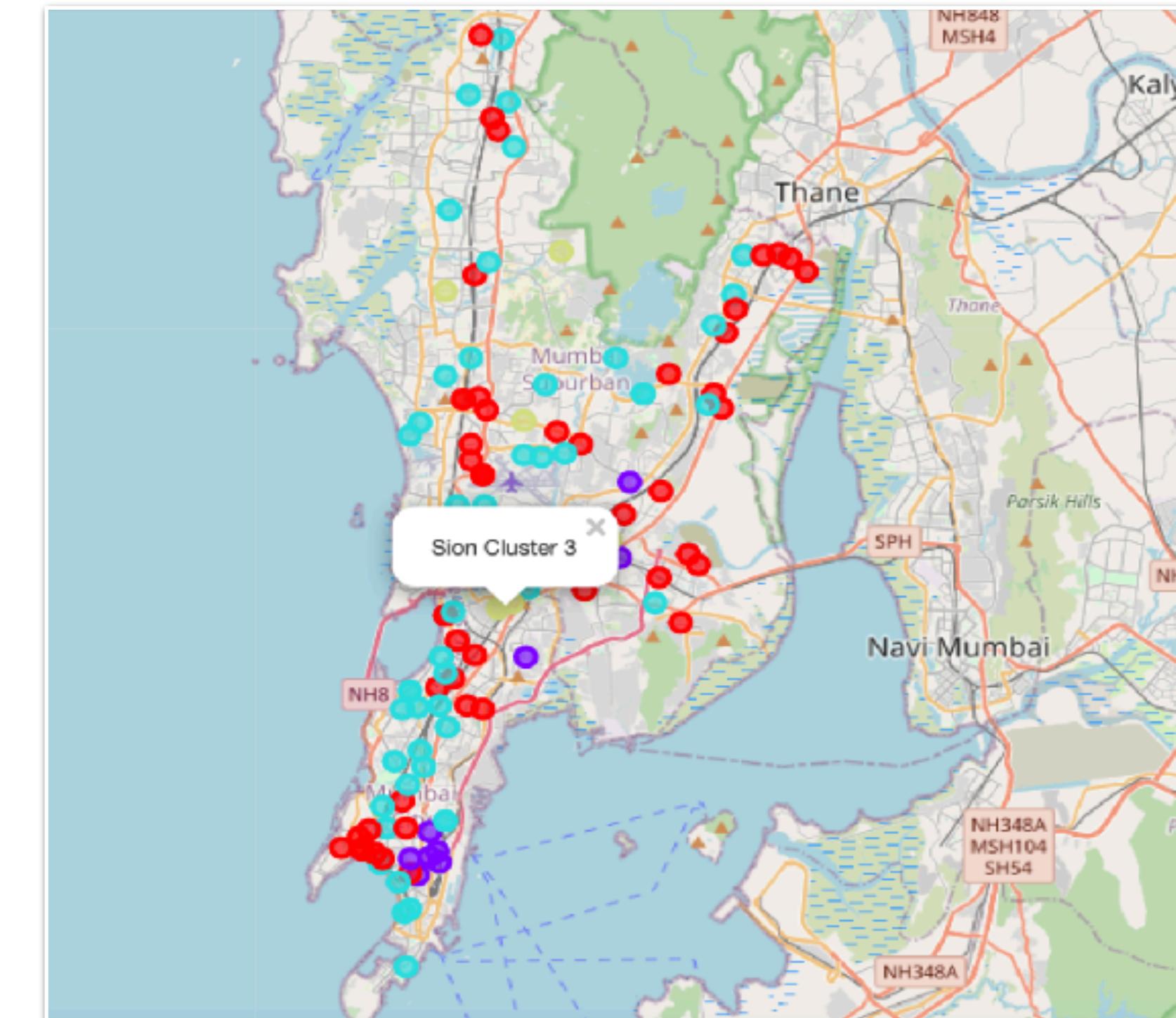
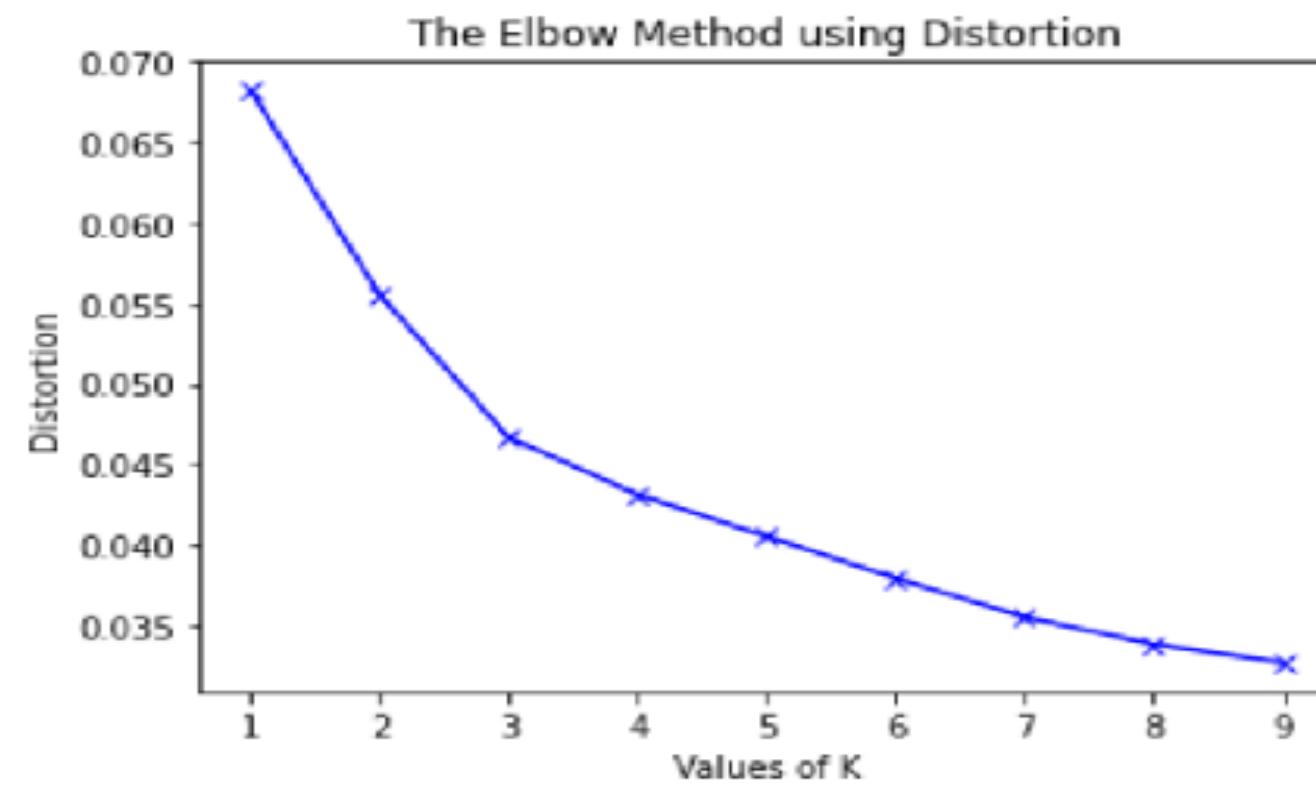
	Area	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	True Count
0	Agripada	Indian Restaurant	Fast Food Restaurant	Bakery	1
3	Andheri	Indian Restaurant	Sandwich Place	Vegetarian / Vegan Restaurant	2
4	Andheri East	Fast Food Restaurant	Indian Restaurant	Hotel	1
5	Andheri Railway Station	Indian Restaurant	Sandwich Place	Vegetarian / Vegan Restaurant	2
7	Anushakti Nagar	Food	Indian Restaurant	Fast Food Restaurant	1



Clustering Neighborhoods: K Means

Based on *Venue Categories* related to “Indian Cuisine”

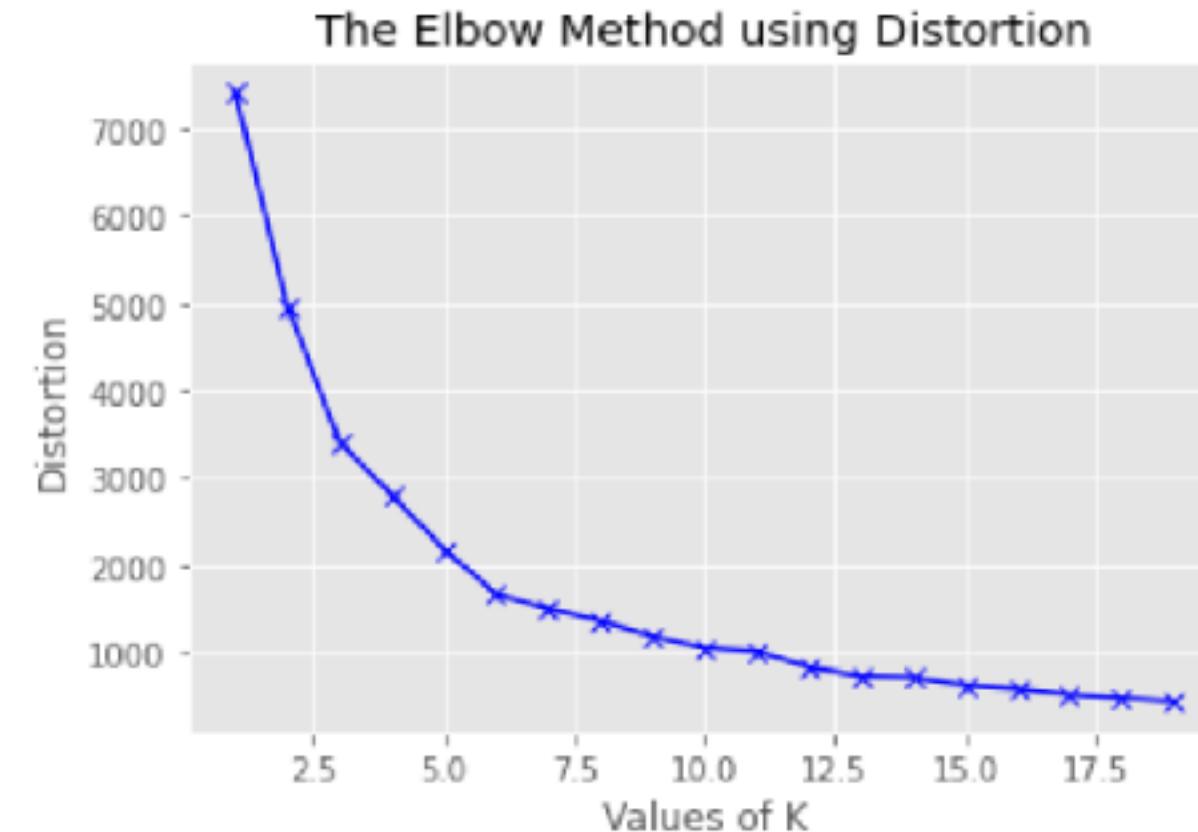
- The 115 neighborhoods were then segmented into 4 clusters using K Means.
- The mean of frequency of occurrence of Venue Categories related to “Indian Cuisine” in each of the neighborhoods was used for clustering.
- Optimum k (number of clusters) was selected using Elbow Plot.
- Resulting clusters were plotted on the map of Mumbai using folium library in Python.



Clustering Neighborhoods: K Means

Based on *Rent Values* of 1 BHK Apartments

- The 115 neighborhoods were then segmented into 5 clusters using K Means.
- The rent value of 1 BHK apartment in each of these neighborhoods was used for clustering.
- Optimum k (number of clusters) was selected using Elbow Plot.
- Resulting clusters displayed based on median rent value in each cluster

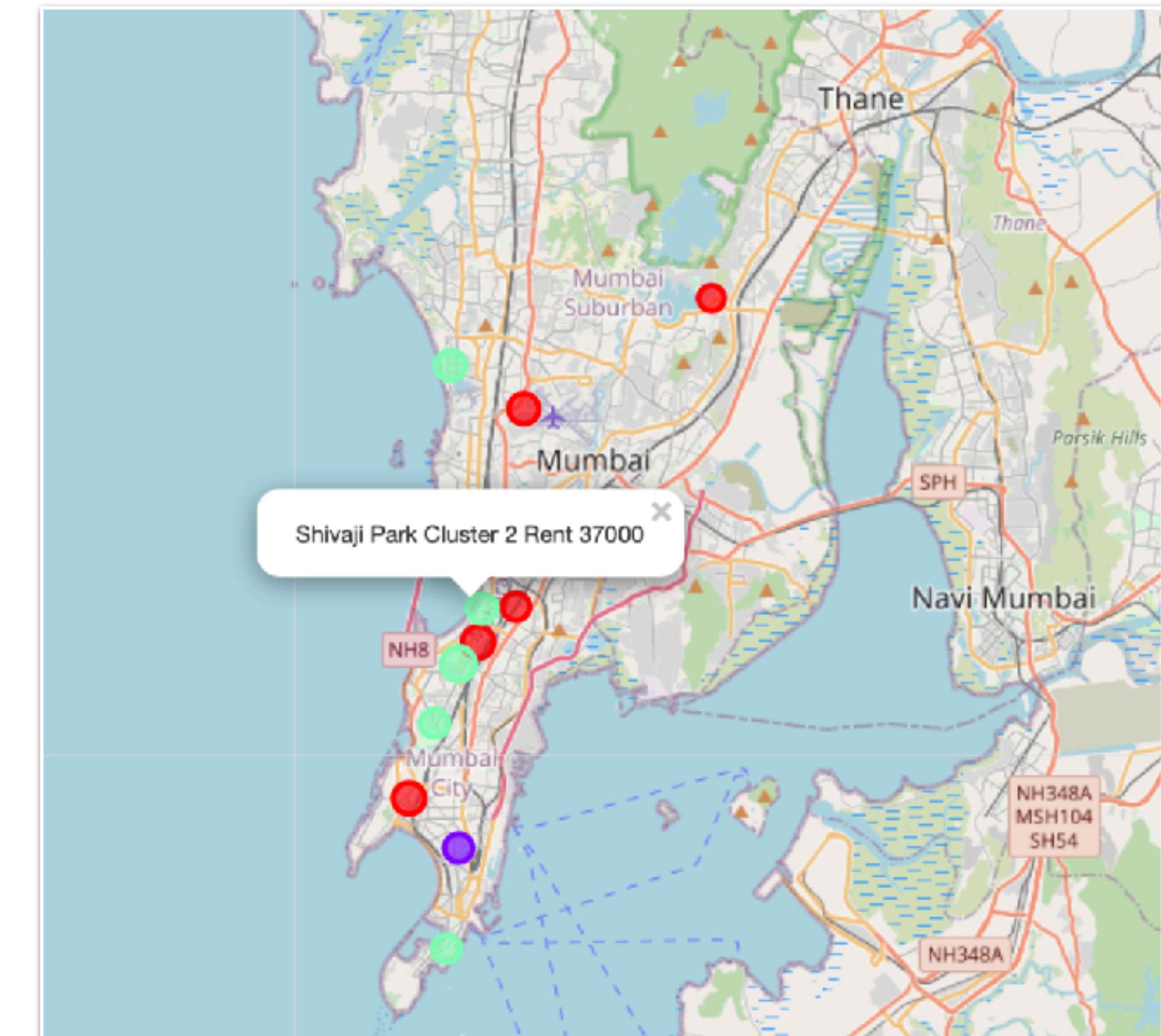


Rent	Cluster Labels_rent
0	50000
1	20750
2	36000
3	28000
4	13000

Arriving at Results

Selecting “upper middle class” Neighborhoods

- Based on the rent distribution, neighborhoods with rents between INR 30K to 40K should be selected.
- These Neighborhoods belong to cluster 2 as shown in previous slide,
- Combining this selection with our dataset of 115 neighborhoods, we arrive at **14 shortlisted neighborhoods**.
- Finally, based on a preferred type of “Indian Cuisine”, specific neighborhood for opening the restaurant can be selected.
- These 14 neighborhoods have been plotted on the Map of Mumbai.



Conclusion & Future Direction

Conclusion

With the use of machine learning approach (K Means) in Python, I was able to analyze and segment different neighborhoods of Mumbai based on popularity of venue categories and their rent value. Based on the analysis optimum neighborhoods were selected within specific rental budget for opening an **Indian Cuisine restaurant**.

Future Direction

The Analysis can be expanded through:

- Using different approaches to clustering (Hierarchical, DBSCAN).
- Increasing scope to include different Venue Categories.
- Increasing effectiveness by including actual footfall data per Venue Category type.

Above approaches would highly improve the effectiveness of selecting the best venue suited to the requirement.

Annexure

Shortlisted 14 Neighborhoods with Coordinates

	Area	Rent	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	Latitude	Longitude
0	Bhawani Shankar	40000	0	Indian Restaurant	Coffee Shop	Fast Food Restaurant	19.015929	72.837460
1	Colaba	34500	2	Indian Restaurant	Diner	Hotel	18.915091	72.825969
2	Gokhale Road	43000	2	Lounge	Indian Restaurant	Bakery	19.008791	72.830730
3	Grant Road	40000	0	Indian Restaurant	Bakery	Snack Place	18.964447	72.813573
4	Haines Road	35000	2	Indian Restaurant	Hotel	Shopping Mall	18.989172	72.822429
5	Hanuman Road	36000	1	Indian Restaurant	Bakery	Fast Food Restaurant	18.948282	72.830243
6	Juhu	38000	2	Hotel	Indian Restaurant	Seafood Restaurant	19.107021	72.827528
8	Mahim	34000	0	Indian Restaurant	Fast Food Restaurant	Bakery	19.042314	72.839834
9	Mahim East	34000	0	Indian Restaurant	Fast Food Restaurant	Bakery	19.042314	72.839834
10	Matunga	35000	0	Indian Restaurant	Fast Food Restaurant	South Indian Restaurant	19.027436	72.850147
11	Powai lit	32500	0	Indian Restaurant	Café	Coffee Shop	19.129055	72.918227
12	Santacruz(West)	38000	0	Indian Restaurant	Hotel	Coffee Shop	19.092336	72.853201
13	Shivaji Park	37000	2	Chinese Restaurant	Indian Restaurant	Movie Theater	19.027236	72.838348