Analyzing the Neighborhoods in Chennai for Starting a Restaurant

Applied Data Science Capstone Project

By: Anu Rithiga B

Date: May 10 2021

Table of Contents

Introduction:

Chennai is one of the metropolitan cities of India and is one of the most densely populated cities in the world. It lies on the east coast of India and attracts heavy tourism from all over the globe every year. Personally, I have been brought up in Chennai and have loved the city from the bottom of my heart. It is one of the major hubs of the world and is extremely diverse with people from various ethnicities residing here. It has a rich diversity of food and people are crazy about having authentic foods during festivals and celebrations.

The multi-cultural nature of the city of Chennai has brought along with it numerous cuisines from all over the world. The people of India generally love food and I personally love to try different cuisines and experience different flavors. Thus, the aim of this project is to study the neighborhoods in Chennai to determine possible locations for starting a restaurant. This project can be useful for business owners and entrepreneurs who are looking to invest in a restaurant in Chennai. The main objective of this project is to carefully analyze appropriate data and find recommendations for the stakeholders

Data Collection:

The following data is required for the project:

1) Neighbourhood data of Chennai
2) Geographical coordinates of Chennai and all neighbourhoods in Chennai

3) Venue data for neighbourhoods in Chennai

Neighbourhoods Data:

The data of the neighbourhoods in Chennai was scraped from https://en.wikipedia.org/wiki/List_of_neighbourhoods_of_Chennai. The data is read into a pandas data frame using the read_html() method. The main reason for doing so is that the Wikipedia page provides a comprehensive and detailed table of the data which can easily be scraped using the read_html() method of pandas. The top 10 rows of the dataframe are shown in Figure 1.

	Neighborhood	Location	Latitude	Longitude
0	Adambakkam	South and East Chennai	12.9880	80.2047
1	Adyar	South and East Chennai	13.0012	80.2565
2	Alandur	South and East Chennai	12.9975	80.2006
3	Alapakkam	West Chennai	13.0490	80.1673
4	Alwarthirunagar	West Chennai	13.0426	80.1840
5	Ambattur	West Chennai	13.1143	80.1548
6	Aminjikarai	West Chennai	13.0698	80.2245
7	Anna Nagar	West Chennai	13.0850	80.2101
8	Annanur	West Chennai	13.1184	80.1246
9	Arumbakkam	West Chennai	13.0724	80.2102

Figure 1: Top 10 rows of Chennai neighborhoods data scraped from Wikipedia.

Geographical Coordinates:

The geographical coordinates for Chennai has been obtained from the GeoPy library in python. This data is relevant for plotting the map of Chennai using the

Folium library in python. The code for getting the geographical coordinates of Chennai is shown in Figure 2.

```
[ ]  address = 'Chennai, IN'
     geolocator = Nominatim()
     location = geolocator.geocode(address)
     latitude = location.latitude
     longitude = location.longitude
     print('The geograpical coordinates of Chennai are {}, {}.'.format(latitude, longitude))

     /usr/local/lib/python3.7/dist-packages/geopy/geocoders/osm.py:143: UserWarning: Using Nomi
       UserWarning
     The geograpical coordinates of Chennai are 13.0836939, 80.270186.
```

Figure 2: Obtaining geographical coordinates of Chennai.

The geocoder library in python has been used to obtain latitude and longitude data for various neighborhoods in Chennai . The coordinates of all neighborhoods in Chennai are used to check the accuracy of coordinates given on Wikipedia and replace them in our data frame if the absolute difference is more than 0.001. These refined coordinates are then further used for plotting neighborhoods using the Folium library in python. Figure 3 shows the coordinates of neighbourhoods in Chennai obtained from Wikipedia as ‘Latitude’ and ‘Longitude’ and those obtained from geocoder as ‘Latitude1’ and ‘Longitude1’. Furthermore, it also shows the absolute difference between the two latitude columns and the two longitude columns as ‘Latdiff’ and ‘Longdiff’, respectively. Once again only the top 10 rows are shown.

```
[ ] df['Latdiff'] = abs(df['Latitude'] - df['Latitude1'])
    df['Longdiff'] = abs(df['Longitude'] - df['Longitude1'])
    df.head(10)
```

	Neighborhood	Location	Latitude	Longitude	Latitude1	Longitude1	Latdiff	Longdiff
0	Adambakkam	South and East Chennai	12.9880	80.2047	12.9919	80.206	0.00392	0.00133
1	Adyar	South and East Chennai	13.0012	80.2565	13.003	80.2519	0.00184	0.00463
2	Alandur	South and East Chennai	12.9975	80.2006	13.0001	80.2005	0.00263	0.00011
3	Alapakkam	West Chennai	13.0490	80.1673	13.0461	80.165	0.00287	0.00229
4	Alwarthirunagar	West Chennai	13.0426	80.1840	13.0506	80.184	0.00795	3e-05
5	Ambattur	West Chennai	13.1143	80.1548	13.1188	80.1544	0.00452	0.00038
6	Aminjikarai	West Chennai	13.0698	80.2245	13.0714	80.2226	0.00159	0.00194
7	Anna Nagar	West Chennai	13.0850	80.2101	13.0836	80.2102	0.00141	5e-05
8	Annanur	West Chennai	13.1184	80.1246	13.1121	80.129	0.00628	0.00435
9	Arumbakkam	West Chennai	13.0724	80.2102	13.0731	80.2095	0.00068	0.00068

Figure 3: Absolute difference between latitude and longitude values obtained from Wikipedia and Geo coder.

Figure 4 shows the top 10 rows of the final Chennai neighbourhoods data frame after replacing the latitude and longitude values as mentioned before and dropping unnecessary columns.

```
[ ] df.drop(['Latitude1', 'Longitude1', 'Latdiff', 'Longdiff'], axis=1, inplace=True)
    df.head(10)
```

	Neighborhood	Location	Latitude	Longitude
0	Adambakkam	South and East Chennai	12.9919	80.206
1	Adyar	South and East Chennai	13.003	80.2519
2	Alandur	South and East Chennai	13.0001	80.2006
3	Alapakkam	West Chennai	13.0461	80.165
4	Alwarthirunagar	West Chennai	13.0506	80.184
5	Ambattur	West Chennai	13.1188	80.1548
6	Aminjikarai	West Chennai	13.0714	80.2226
7	Anna Nagar	West Chennai	13.0836	80.2101
8	Annanur	West Chennai	13.1121	80.129
9	Arumbakkam	West Chennai	13.0724	80.2102

Figure 4: Final Chennai neighborhoods dataframe.

Venue Data:

The venue data has been extracted using the Foursquare API. This data contains venue recommendations for all neighbourhoods in Chennai and is used to study the popular venues of different neighbourhoods as well as build the unsupervised learning model to cluster neighbourhoods. The venue recommendations of all neighbourhoods were obtained with a limit of 200, that is, maximum of 200 venue recommendations per neighbourhood and a radius of 1 km around the neighbourhood's geographical coordinates. Figure 5 shows the top 10 rows depicting the results obtained after cleaning the data from Foursquare API.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Adambakkam	12.99192	80.20603	Pizza Republic	12.990987	80.198613	Pizza Place
1	Adambakkam	12.99192	80.20603	St Thomas Mount Railway Station	12.994987	80.200302	Train Station
2	Adambakkam	12.99192	80.20603	Sukkkubai Beef Biryani Shop	12.998769	80.201381	Indian Restaurant
3	Adambakkam	12.99192	80.20603	Loiee	12.992197	80.199000	Bakery
4	Adambakkam	12.99192	80.20603	Venkateshwara Super Market	12.986320	80.205168	Department Store
5	Adambakkam	12.99192	80.20603	Deepam Restaurant	12.985380	80.205281	Indian Restaurant
6	Adambakkam	12.99192	80.20603	St. Thomas mount railway station	12.998494	80.203740	Train Station
7	Adambakkam	12.99192	80.20603	Design Hotel Chennai by jüSTa	12.992068	80.214965	Hotel
8	Adambakkam	12.99192	80.20603	St. Thomas Mount Metro Station	12.994772	80.197556	Metro Station
9	Adambakkam	12.99192	80.20603	Heritage Fresh	13.000499	80.208087	Supermarket

Figure 5: Data obtained from Foursquare API after cleaning.

Methodology:

This section provides details for the methodology used in the project.

Data Visualization

In order to understand the data obtained for Chennai neighbourhoods, basic visualization was carried out. Figure 6 shows a bar plot depicting the number of neighbourhoods in each location in Chennai.

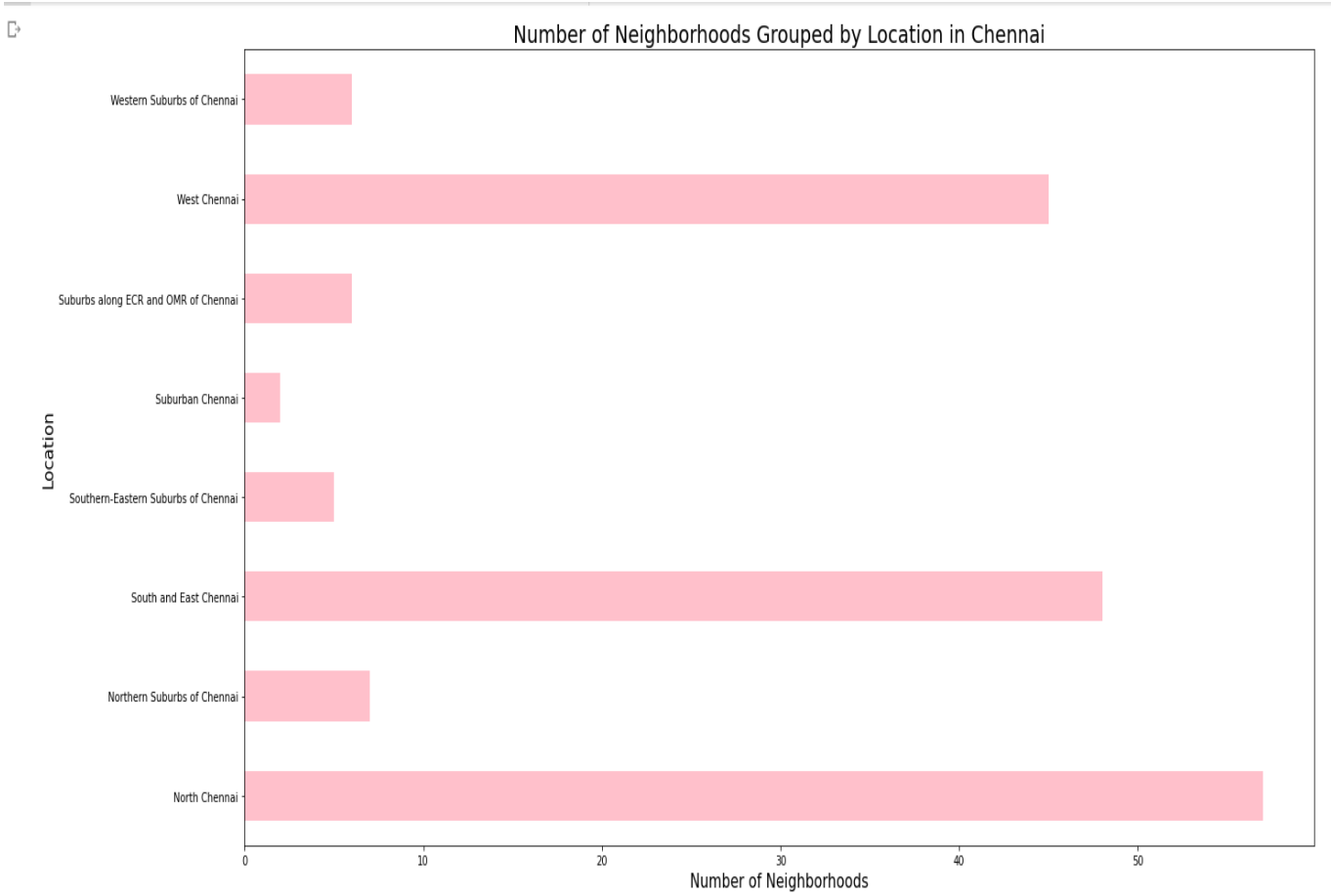

Figure 6: Number of neighbourhoods grouped by location.

It is evident from Figure 6 that South and East Chennai and North Chennaai has the most number of neighbourhoods. Notice how we see one of the locations as Chennai itself? This is because the neighbourhoods contained in this location are located at the outskirts of the city and thus have been termed as just Chennai.

Using folium, a map was plotted to show how the different neighbourhoods are spread all across Chennai. This is shown in Figure 7.

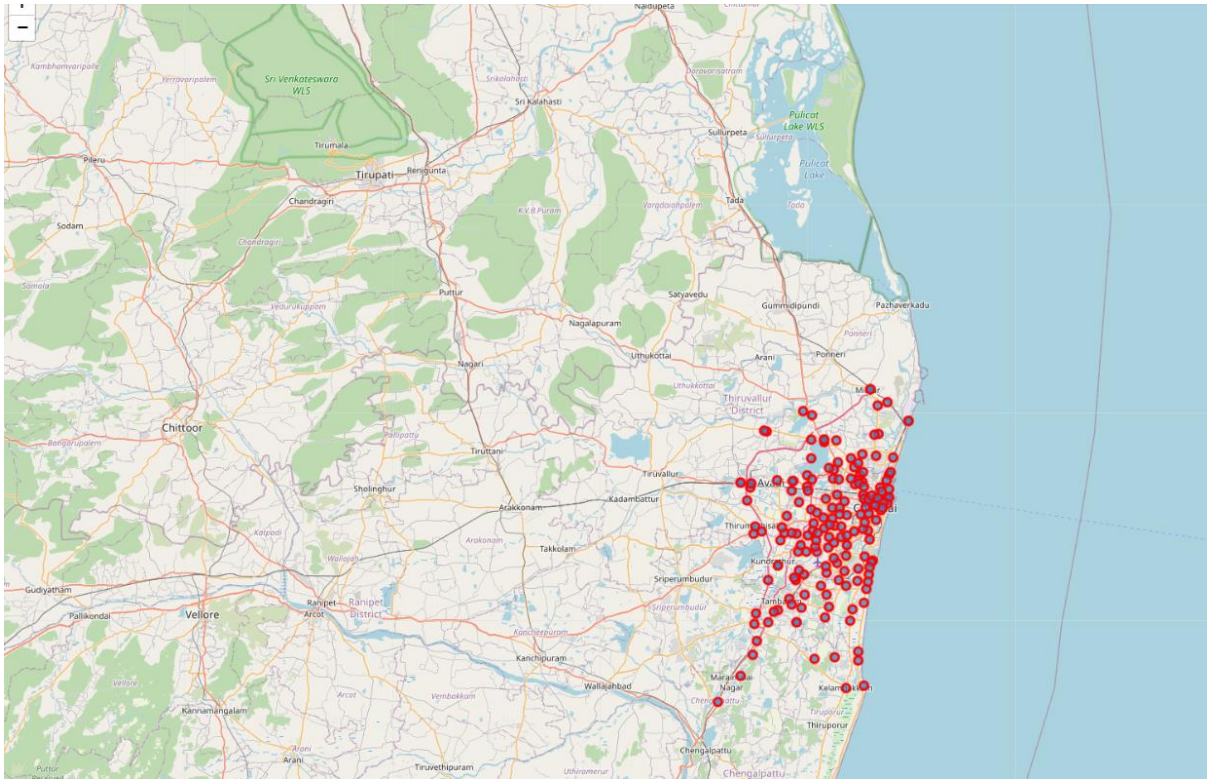

Figure 7: Depicting the neighbourhood spread across Chennai .

Feature Extraction

Feature extraction was carried out to obtain features from the Foursquare API data (as shown in Figure 5) which was used for building the unsupervised learning model. In order to achieve this, the “Venue Category” column had to be converted to some form of numeric value to be used for building the model. This was achieved by the One-hot Encoding method which takes all the unique categories and creates a column for each category. Then, if a neighbourhood venue belongs to that category, it would get a value of 1 for that row in that specific category column and if a neighbourhood venue does not belong to the particular category, the value would be 0. This process was repeated for all venues in all neighbourhoods and the result was a sparse matrix containing the neighbourhood

name and all unique category columns with either 1 or 0 based on whether the neighbourhood venue belonged to that category or not. This data frame was then grouped by the neighbourhood name and the average value was taken for all categories. The result is shown in Figure 8 which shows only the top 10 rows.

(161, 191)

	Neighborhood	ATM	Accessories Store	Afghan Restaurant	African Restaurant	Airport Terminal	American Restaurant	Amphitheater	Andhra Restaurant	Antique Shop	Arcade	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	BBQ Joint	Bakery
0	Adambakkam	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.0	0.000000	0.0	0.000000	0.100000
1	Adyar	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.014706	0.0	0.044118	0.0	0.014706	0.000000
2	Alandur	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.0	0.000000	0.0	0.000000	0.117647
3	Alapakkam	0.0	0.1	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.0	0.100000	0.0	0.000000	0.100000
4	Alwarthirunagar	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.0	0.000000	0.0	0.000000	0.000000
5	Ambattur	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.0	0.100000	0.0	0.000000	0.000000
6	Aminjikarai	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.0	0.000000	0.0	0.000000	0.000000
7	Anna Nagar	0.0	0.0	0.0	0.0	0.0	0.014493	0.0	0.0	0.0	0.000000	0.0	0.014493	0.0	0.014493	0.028986
8	Annanur	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.0	0.000000	0.0	0.000000	0.000000
9	Arumbakkam	0.0	0.0	0.0	0.0	0.0	0.000000	0.0	0.0	0.0	0.000000	0.0	0.000000	0.0	0.000000	0.000000

Figure 8: One-hot Encoding resulting data frame.

Notice that most of the values are 0 since there were a large number of unique categories and not all neighbourhoods had venues belonging to each category. This data was used for the unsupervised learning model with the neighbourhood name dropped. The unsupervised learning model is explained in the next section. A data frame was also created which contained the top 10 most common venues of all neighbourhoods. Though this is not a part of Feature Extraction, it is important to provide a glimpse into what this data frame looks like as it will be used later to combine the results from the unsupervised learning model. The top 10 rows of this data frame are shown in Figure 9.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Adambakkam	Indian Restaurant	Train Station	Supermarket	Department Store	Hotel	Pizza Place	Metro Station	Bakery	Food Truck	Food Service
1	Adyar	Indian Restaurant	Café	Pizza Place	Asian Restaurant	Department Store	Chinese Restaurant	Fast Food Restaurant	Ice Cream Shop	Rock Club	North Indian Restaurant
2	Alandur	Indian Restaurant	Metro Station	Bakery	Train Station	Hotel	Church	Café	Fish Market	Breakfast Spot	Pizza Place
3	Alapakkam	Department Store	Fast Food Restaurant	Bakery	Accessories Store	Sandwich Place	Clothing Store	Business Service	Asian Restaurant	Fried Chicken Joint	Food Truck
4	Alwarthirunagar	Pizza Place	Clothing Store	Fast Food Restaurant	Ice Cream Shop	Tea Room	Chinese Restaurant	Café	Multiplex	Movie Theater	Gym
5	Ambattur	Ice Cream Shop	Movie Theater	Indian Restaurant	Flea Market	Multiplex	River	Asian Restaurant	Department Store	Electronics Store	Event Space
6	Aminjikarai	Fast Food Restaurant	Pizza Place	Electronics Store	Clothing Store	Shopping Mall	Burger Joint	Cosmetics Shop	Bookstore	Furniture / Home Store	Event Space
7	Anna Nagar	Indian Restaurant	Chinese Restaurant	Fast Food Restaurant	Department Store	Vegetarian / Vegan Restaurant	Coffee Shop	Clothing Store	Gym	Bakery	Electronics Store
8	Annanur	Train Station	Platform	Medical Supply Store	Stationery Store	Women's Store	Fast Food Restaurant	Food Truck	Food Service	Food Court	Food & Drink Shop
9	Arumbakkam	Hotel	Vegetarian / Vegan Restaurant	Fast Food Restaurant	Food Court	Scenic Lookout	South Indian Restaurant	Burger Joint	Event Space	Bus Station	Chinese Restaurant
10	Ashok Nagar	Indian Restaurant	Vegetarian / Vegan Restaurant	Pizza Place	Fast Food Restaurant	Sculpture Garden	Coffee Shop	Multicuisine Indian Restaurant	Café	Restaurant	Snack Place

Figure 9: Top 10 most common venues for neighbourhoods.

Unsupervised Learning

K-means unsupervised learning technique was used to cluster the neighbourhoods based on the category of venues near the neighbourhoods. One important aspect of the k-means model is to determine the number of clusters to use in model development. This was determined by the Silhouette score which was calculated for a range of clusters from 2 to 15. The resulting number of clusters and their respective Silhouette scores are shown in Figure 10.

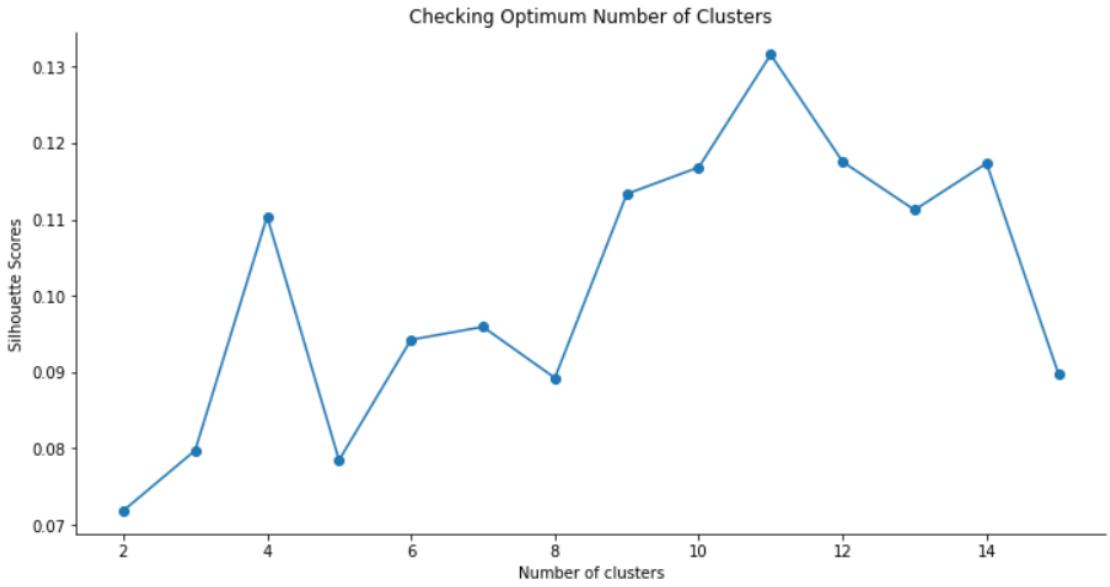

Figure 10: Silhouette scores for different number of clusters.

It is evident that the Silhouette scores are not very high even as the number of clusters increases. This means that the inter-cluster distance is not very high over the range of k-values. Despite this, the data will be clustered to the best possible extent. For this, 11 clusters will be used for the k-means clustering model since it provides the highest silhouette score as seen in Figure 10.

Results

The clustering model then clusters the neighbourhoods in Chennai and provides a label for each neighbourhood which is representative of the cluster it belongs to. The cluster labels were then added to the data frame in Figure 9 along with the Location, Latitude, and Longitude columns to provide a complete summary of the clustering. The top 10 rows are shown in Figure 11.

(176, 15)

	Neighborhood	Location	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
0	Adambakkam	South and East Chennai	12.9919	80.206	0.0	Indian Restaurant	Train Station	Supermarket	Department Store	Hotel	Pizza Place	Metro Station	Bakery	Food Truck
1	Adyar	South and East Chennai	13.003	80.2519	0.0	Indian Restaurant	Café	Pizza Place	Asian Restaurant	Department Store	Chinese Restaurant	Fast Food Restaurant	Ice Cream Shop	Rock Club
2	Alandur	South and East Chennai	13.0001	80.2006	0.0	Indian Restaurant	Metro Station	Bakery	Train Station	Hotel	Church	Café	Fish Market	Breakfast Spot
3	Alapakkam	West Chennai	13.0461	80.165	0.0	Department Store	Fast Food Restaurant	Bakery	Accessories Store	Sandwich Place	Clothing Store	Business Service	Asian Restaurant	Fried Chicken Joint
4	Alwarthirunagar	West Chennai	13.0506	80.184	0.0	Pizza Place	Clothing Store	Fast Food Restaurant	Ice Cream Shop	Tea Room	Chinese Restaurant	Café	Multiplex	Movie Theater
5	Ambattur	West Chennai	13.1188	80.1548	0.0	Ice Cream Shop	Movie Theater	Indian Restaurant	Flea Market	Multiplex	River	Asian Restaurant	Department Store	Electronics Store
6	Aminjikarai	West Chennai	13.0714	80.2226	0.0	Fast Food Restaurant	Pizza Place	Electronics Store	Clothing Store	Shopping Mall	Burger Joint	Cosmetics Shop	Bookstore	Furniture / Home Store
7	Anna Nagar	West Chennai	13.0836	80.2101	0.0	Indian Restaurant	Chinese Restaurant	Fast Food Restaurant	Department Store	Vegetarian / Vegan Restaurant	Coffee Shop	Clothing Store	Gym	Bakery

Figure 11: Clustering neighborhoods in Chennai .

Furthermore, neighbourhoods in each individual cluster can be extracted using cluster labels and thus the details of specific clusters can be seen. This is done below for all clusters with only 10 rows for clusters that contain a high number of neighbourhoods.

	Neighborhood	Location	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Adambakkam	South and East Chennai	Indian Restaurant	Train Station	Supermarket	Department Store	Hotel	Pizza Place	Metro Station	Bakery	Food Truck	Food Service
1	Adyar	South and East Chennai	Indian Restaurant	Café	Pizza Place	Asian Restaurant	Department Store	Chinese Restaurant	Fast Food Restaurant	Ice Cream Shop	Rock Club	North Indian Restaurant
2	Alandur	South and East Chennai	Indian Restaurant	Metro Station	Bakery	Train Station	Hotel	Church	Café	Fish Market	Breakfast Spot	Pizza Place
3	Alapakkam	West Chennai	Department Store	Fast Food Restaurant	Bakery	Accessories Store	Sandwich Place	Clothing Store	Business Service	Asian Restaurant	Fried Chicken Joint	Food Truck
4	Alwarthirunagar	West Chennai	Pizza Place	Clothing Store	Fast Food Restaurant	Ice Cream Shop	Tea Room	Chinese Restaurant	Café	Multiplex	Movie Theater	Gym
5	Ambattur	West Chennai	Ice Cream Shop	Movie Theater	Indian Restaurant	Flea Market	Multiplex	River	Asian Restaurant	Department Store	Electronics Store	Event Space
6	Aminjikarai	West Chennai	Fast Food Restaurant	Pizza Place	Electronics Store	Clothing Store	Shopping Mall	Burger Joint	Cosmetics Shop	Bookstore	Furniture / Home Store	Event Space
7	Anna Nagar	West Chennai	Indian Restaurant	Chinese Restaurant	Fast Food Restaurant	Department Store	Vegetarian / Vegan Restaurant	Coffee Shop	Clothing Store	Gym	Bakery	Electronics Store
9	Arumbakkam	West Chennai	Hotel	Vegetarian / Vegan Restaurant	Fast Food Restaurant	Food Court	Scenic Lookout	South Indian Restaurant	Burger Joint	Event Space	Bus Station	Chinese Restaurant
10	Ashok Nagar	West Chennai	Indian Restaurant	Vegetarian / Vegan Restaurant	Pizza Place	Fast Food Restaurant	Sculpture Garden	Coffee Shop	Multicuisine Indian Restaurant	Café	Restaurant	Snack Place

Figure 12: Cluster 1.

	Neighborhood	Location	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
58	Mangadu	West Chennai	Pharmacy	Furniture / Home Store	Fruit & Vegetable Store	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food & Drink Shop	Food	Flower Shop
61	Mathur MMDA	North Chennai	Pharmacy	Furniture / Home Store	Fruit & Vegetable Store	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food & Drink Shop	Food	Flower Shop

Figure 13: Cluster 2.

	Neighborhood	Location	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
22	Ennore	North Chennai	Train Station	Women's Store	Fast Food Restaurant	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food & Drink Shop	Food	Flower Shop
124	Thirunindravur	West Chennai	Train Station	Pharmacy	Fast Food Restaurant	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food & Drink Shop	Food	Flower Shop
146	Athipattu	Northern Suburbs of Chennai	Train Station	Platform	Women's Store	Fast Food Restaurant	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food & Drink Shop	Food
150	Ennore	Northern Suburbs of Chennai	Train Station	Women's Store	Fast Food Restaurant	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food & Drink Shop	Food	Flower Shop
153	Thirunindravur	Western Suburbs of Chennai	Train Station	Pharmacy	Fast Food Restaurant	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food & Drink Shop	Food	Flower Shop

Figure 14: Cluster 3.

	Neighborhood	Location	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
15	Broadway	North Chennai	Indian Restaurant	Market	Department Store	Restaurant	Snack Place	Video Store	Hotel	Fried Chicken Joint	Food Service	Food Court
17	Chetpet	South and East Chennai	Indian Restaurant	Café	Hotel	Coffee Shop	Chinese Restaurant	Restaurant	Seafood Restaurant	Bakery	Ice Cream Shop	Juice Bar
35	Kadaperi	South and East Chennai	Indian Restaurant	Train Station	Jewelry Store	Restaurant	Light Rail Station	Fish Market	Fried Chicken Joint	Food Truck	Food Service	Food Court
36	Karambakkam	West Chennai	Indian Restaurant	Flea Market	Restaurant	Deli / Bodega	Fish Market	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food & Drink Shop
41	Kolathur	North Chennai	Indian Restaurant	ATM	Restaurant	Bus Stop	Food Truck	Food Service	Food Court	Food & Drink Shop	Food	Flower Shop
43	Kosapet	North Chennai	Indian Restaurant	Department Store	Hotel	Fish Market	Fruit & Vegetable Store	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food & Drink Shop
57	manapakkam	West Chennai	Indian Restaurant	Afghan Restaurant	Trail	Women's Store	Fast Food Restaurant	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food & Drink Shop
59	Manjambakkam	North Chennai	Indian Restaurant	Fish Market	Fruit & Vegetable Store	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food & Drink Shop	Food	Flower Shop
62	Medavakkam	South and East Chennai	Indian Restaurant	Pizza Place	Chinese Restaurant	Women's Store	Fast Food Restaurant	Food Truck	Food Service	Food Court	Food & Drink Shop	Food
85	Pallikaranai	South and East Chennai	Indian Restaurant	Bar	Gaming Cafe	Fruit & Vegetable Store	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food & Drink Shop	Food

Figure 15: Cluster 4.

	Neighborhood	Location	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
11	Avadi	West Chennai	IT Services	Department Store	Concert Hall	Park	Breakfast Spot	Women's Store	Fish Market	Food Truck	Food Service	Food Court
12	Ayappakkam	West Chennai	Pharmacy	Department Store	Clothing Store	Print Shop	Medical Supply Store	Supermarket	Fish Market	Food Truck	Food Service	Food Court
31	Irumbuliyur	South and East Chennai	Asian Restaurant	Motorcycle Shop	Bus Station	Coffee Shop	Restaurant	Fish Market	Food Truck	Food Service	Food Court	Food & Drink Shop
38	Keelkattalai	South and East Chennai	Fast Food Restaurant	Pizza Place	Fried Chicken Joint	Electronics Store	Bus Station	Fish Market	Food Truck	Food Service	Food Court	Food & Drink Shop
46	Koyambedu	West Chennai	Vegetarian / Vegan Restaurant	Bus Station	South Indian Restaurant	Fast Food Restaurant	Metro Station	Department Store	Flower Shop	Food Court	Antique Shop	Bus Line
47	Kundrathur	West Chennai	Soup Place	Bus Station	Furniture / Home Store	Fruit & Vegetable Store	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food & Drink Shop	Food
50	M.G.R. Nagar	West Chennai	Thai Restaurant	Bed & Breakfast	Bus Station	Grocery Store	Furniture / Home Store	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food & Drink Shop
52	Madhavaram	North Chennai	Ice Cream Shop	Park	Food & Drink Shop	Bus Station	Fish Market	Fried Chicken Joint	Food Truck	Food Service	Food Court	Food
63	Minjur	North Chennai	Pharmacy	Train Station	Scenic Lookout	Park	Fast Food Restaurant	Food Truck	Food Service	Food Court	Food & Drink Shop	Food
64	Mogappair	West Chennai	Chinese Restaurant	Bakery	Pizza Place	South Indian Restaurant	Bus Station	Food & Drink Shop	Ice Cream Shop	Accessories Store	Donut Shop	Electronics Store
73	Naravarikuppam	North Chennai	Multiplex	Indian Restaurant	Bus Station	Gym	Fruit & Vegetable Store	Food Truck	Food Service	Food Court	Food & Drink Shop	Food

Figure 16: Cluster 5.

Based on the clusters shown above, the neighbourhoods can once again be plotted on a map of Chennai, however, this time with different colour markers to distinguish between different clusters. This is shown in Figure 17.

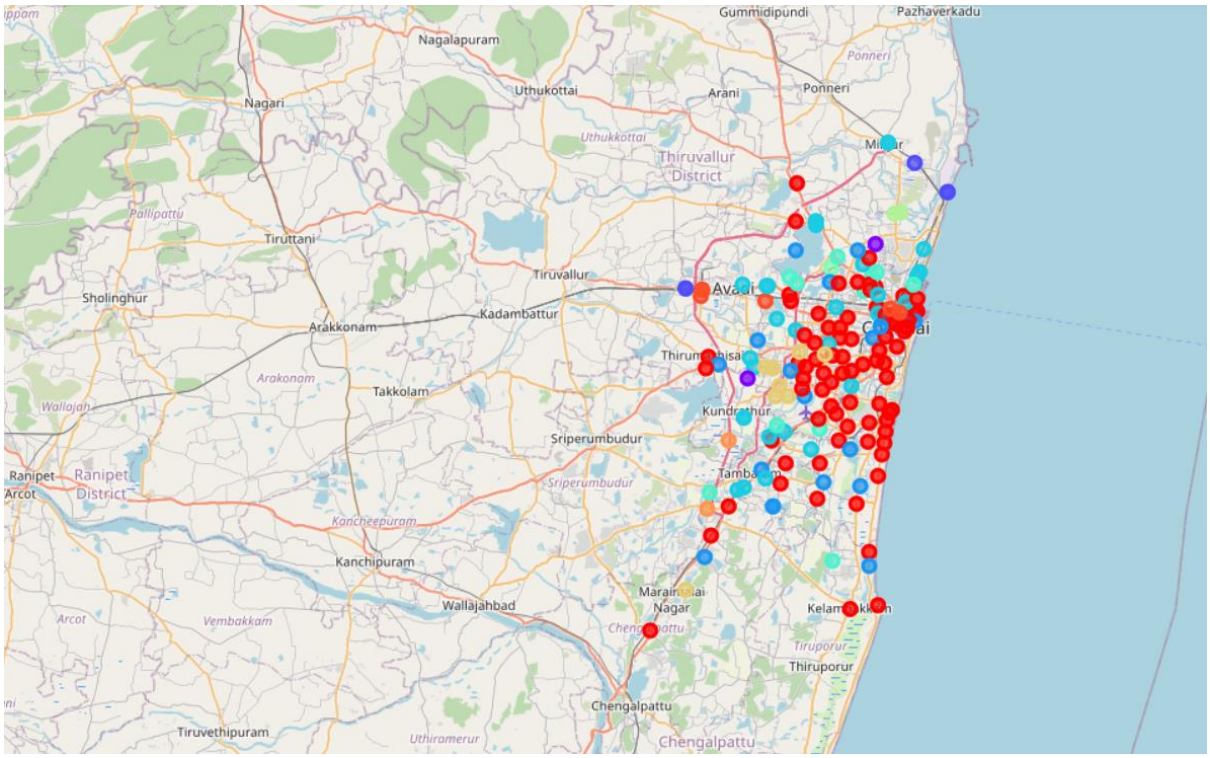

Figure 17: Visualizing the clustering of neighbourhoods in Chennai .

Discussion

By analyzing the five clusters obtained we can see that some of the clusters are more suited for restaurants and hotels, whereas, other clusters are less suited. Ne ighborhoods in clusters 2, 3, and 4 contain a small percentage of restaurants, hot els, cafe and pubs in their top 10 common venues. These clusters contain a high er degree of other venues like train station, bus station, fish market, gym, perfor ming arts venue and smoke shop, to name a few. Thus, they are not well suited f

or opening a new restaurant. On the other hand, neighborhoods in clusters 1 and 5 contain a much higher degree of restaurants, hotels, multiplex, cafes, bars and other food joints. Thus, the neighborhoods in these clusters would be well suite d for opening a new restaurant.

Comparing clusters 1 and 5, neighborhoods in cluster 1 seem to be more suited f or starting a restaurant since they contains a larger percentage of food joints in t he top 10 most common venues than cluster 5. The neighborhoods in cluster 1 c ontain a variety of food joints like restaurants, tea rooms, bakery, cafe, food truc ks and many eateries, and also contain very diverse cuisines like Japanese, India n, Chinese, Italian and seafood restaurants. Most neighborhoods in cluster 5 see m to have Indian Restaurant as their top most common venue; however, on care ful analysis we can see that neighborhoods in cluster 5 also contain other venues like soccer field, flea market, smoke shop, gym, train station, dance studio, mus ic store, cosmetics shop and so on. Thus, it is recommended that the new restaur ant can be opened in the neighborhoods belonging to cluster 1. This neighborho od can be further plotted on a map as shown below.

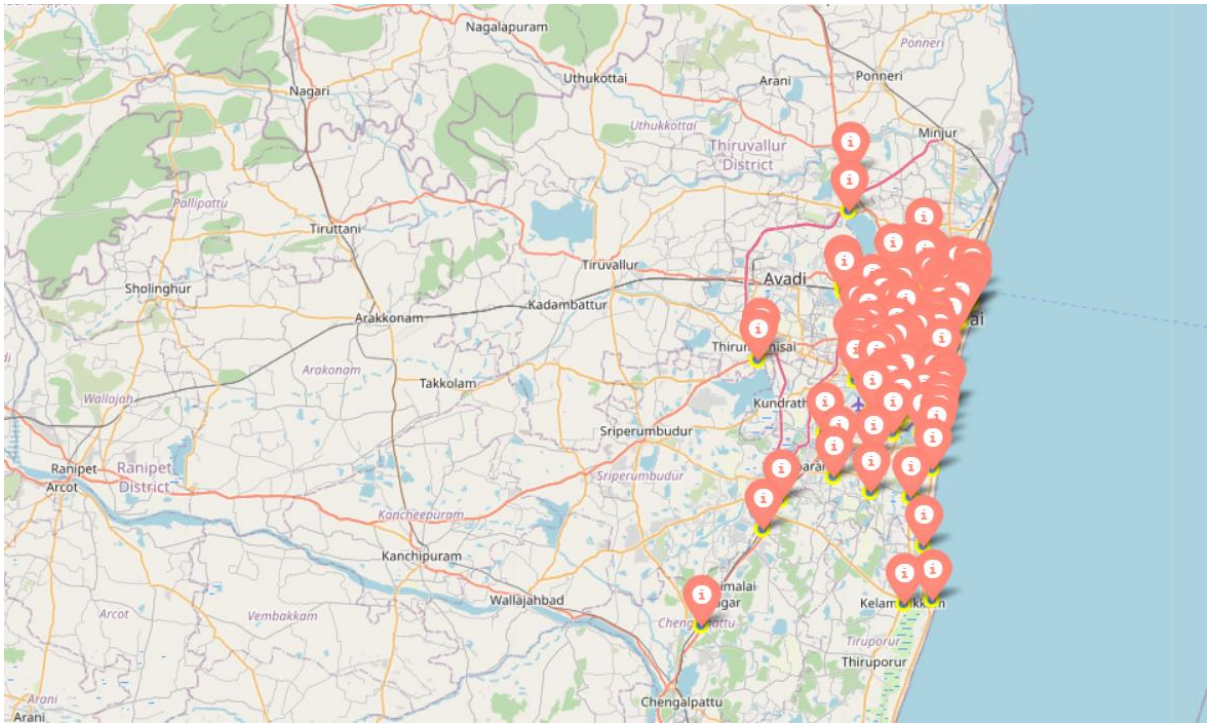

Figure 18: Neighbourhoods most suited for starting a new restaurant.

Conclusion

In this project, the neighbourhoods in Chennai, India have been successfully analysed for determining which would be the best neighbourhoods for opening a new restaurant. Based on the analysis carried out, neighbourhoods in cluster 1 is recommended as locations for the new restaurant. This has also been plotted in the map in Figure 18. The stakeholders and investors can further tune this by considering various other factors like transport, legal requirements, and costs associated. These were out of the scope for this project and thus were not considered.

Final Comments

Note 1: In order to view the code for this project, kindly refer to the notebook on the git hub repository at: https://github.com/iamgrootsh7/Coursera_Capstone