ANALYSING BENGALURU NEIGHBORHOODS

FINAL ASSIGNMENT: PROFESSIONAL DATA SCIENCE CAPSTONE PROJECT

- ROBIN VARGHEESE

INTRODUCTION

The IT boom in India has made a Bengaluru (Bangalore) a cosmopolitan city. This is a city where we see a lot of variety, especially in food and culture. One can find people from many nationalities either working in Benagaluru or often visiting Bengaluru. This is in addition to a large number of people from other states of India. This has made Bengaluru a food lover's paradise. We can have a taste of both Italy and China in the same neighborhood itself. Hence it is very interesting to analyze the restaurants of Bengaluru.

Business Problem:

Consider some businessman probing the scope of opening restaurants in Bengaluru. The primary concern would be choosing the neighborhood best suited for the type of restaurant he plans to open. By analyzing the available restaurant types and other common venues in the different neighborhoods an understanding of food habits and social structure of the people of each locality can be obtained. There would be some neighborhoods where dance bar and pubs are common occurrences. Some neighborhoods may have fast food restaurants more commonly. There may be interesting correlations like more number of café and ice cream parlors in neighborhoods where there are more educational institutes. Upon analysis, segmenting the neighborhoods sharing common characteristics into different clusters would be the best way to study the neighborhoods. This is what I have attempted to do in this project.

The focus would be on gathering the information like the following:

- The distribution of restaurants across the neighborhoods.
- The different types of restaurants based on the cuisines they serve.
- Top dishes, ratings and approximate costs by neighborhood.
- Most widely distributed restaurants and larger chains.
- Getting the geographical coordinates of the neighborhoods and plotting them in a map.
- Differences and similarities between neighborhoods based on food habits.
- Common venues other and other places in each neighborhood.
- Finding neighborhoods with similar characteristics and segmenting them into clusters.

DATA

The restaurant data listed based on the city's neighborhoods is derived from this kaggle dataset. The uploader has scraped the data from the website of Zomato – one of the most popular restaurant review and online food ordering service in India.

Data Acquisition:

The data used in the project are acquired from three sources:

- Restaurant data from the kaggle dataset mentioned above
- Geographical coordinates obtained using geopy package
- Venues and interests for the neighborhood locations using Foursquare API

The restaurant dataset from kaggle is in comma separated value format and is of size 547 MB. It contains the following fields:

Field name	**Number of unique values**	**Description**
url	51,717	url of the restaurant's page in the zomato website
address	11,495	restaurant address as listed in the zomato website
name	8,792	name of the restaurant
online order	2	whether online ordering is available in the restaurant or not
book table	2	table book option available or not
rate	64	overall rating of the restaurant out of 5
votes	2328	total number of votes for rating
phone	64	restaurant phone number
location	93	neighborhood in which the restaurant is located
rest type	93	restaurant category (like quickbites, bakery etc.)
dish liked	5271	dishes people liked in the restaurant
cuisines	2723	cuisine types available in the restaurant
approx cost (for two people)	70	approximate cost of meal for two people
reviews list	22513	list of tuples containing customer reviews and ratings for the restaurant

menu item	9098	menu items available in the restaurant
listed in(type)	7	restaurant category in which the restaurant is listed
listed in(city)	30	neighborhood in which the restaurant is listed

Data Preparation and Cleaning:

Initially pandas was not able to read the original 547 MB CSV dataset because of the hardware incapacity of my PC. I had used an ETL tool, pentaho kettle, to read the CSV file and reduce the size by removing some fields which are not much relevant to our analysis like URL, address and review list fields.

The resultant dataset is read into a pandas data frame and is checked for duplicates and null values. The ambiguous data from certain fields are corrected. The data is checked for other inconsistencies.

METHODOLOGY

Exploratory Data Analysis:

Once the data is cleaned and prepared various charts are drawn and analyses are made with interesting observations. For plotting the data in the map and to run k-means clustering algorithm, the latitude and longitude coordinates for each neighborhood in the dataset is obtained by calling the geocoder in the geopy package. After getting the coordinate values it is used to get the venue details from the Foursquare API. Exploratory data analysis on the venues data is performed and the clustering models are made.

The latitude and longitude coordinates for each neighborhood in the dataset is obtained by using the geopy library which is a free python library for geographical data.

Foursquare is one of the most popular location data providers. We can access the venue and places for any city through their API.

RESULTS AND INFERENCES

Let us have some interesting observations made from the restaurant data from the following charts.

Chart1 :- Neighborhoods with most number of restaurants

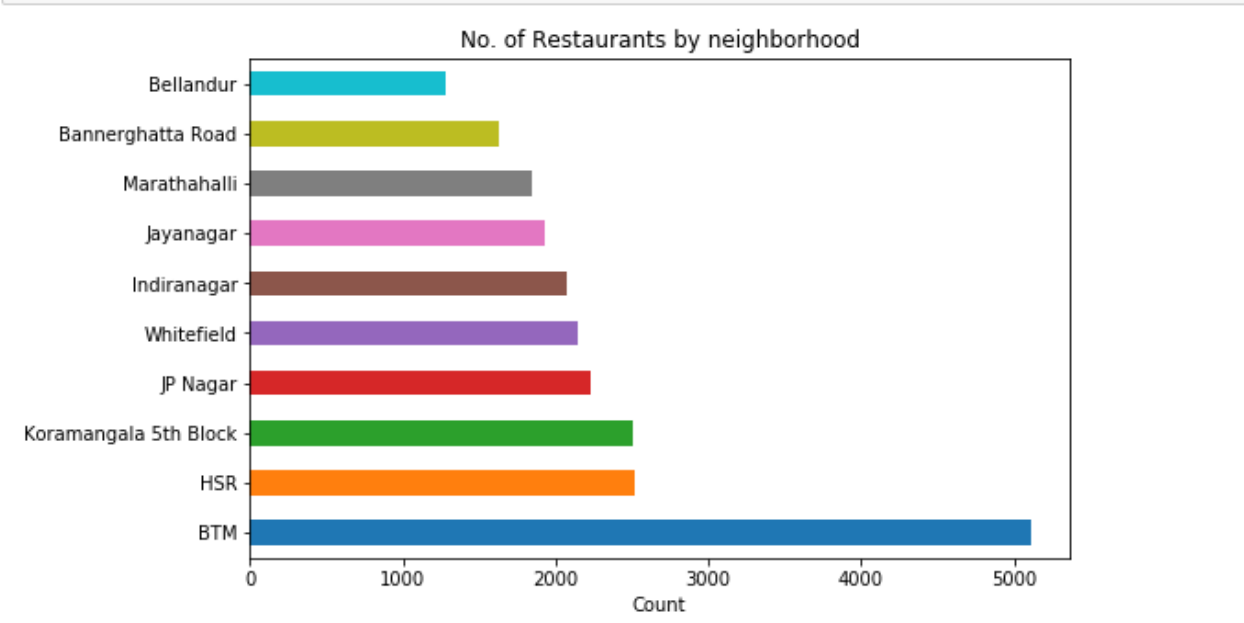


Chart2 :- Number of restaurants by food type.

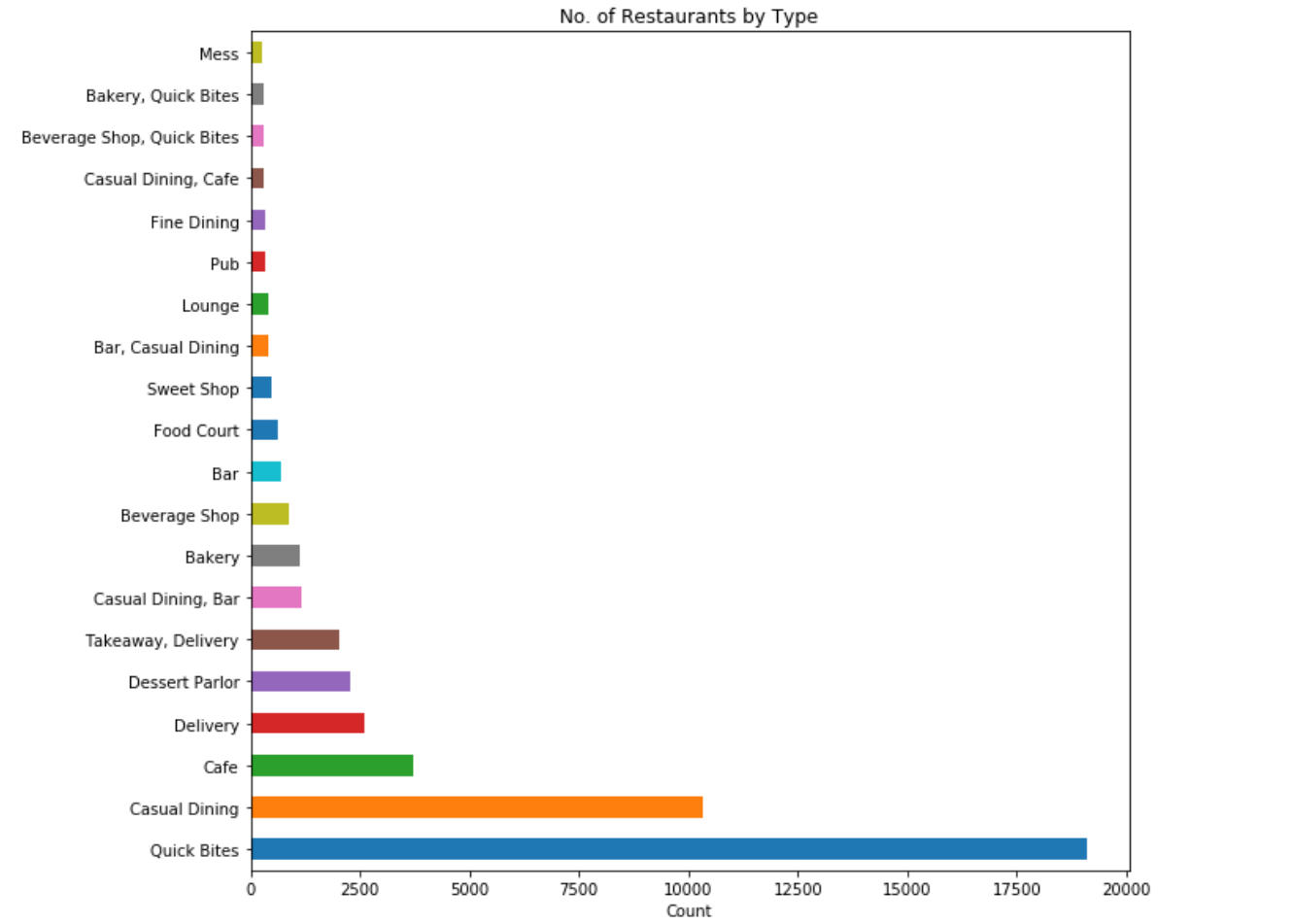


Chart3 :- Top restaurants by cuisine served

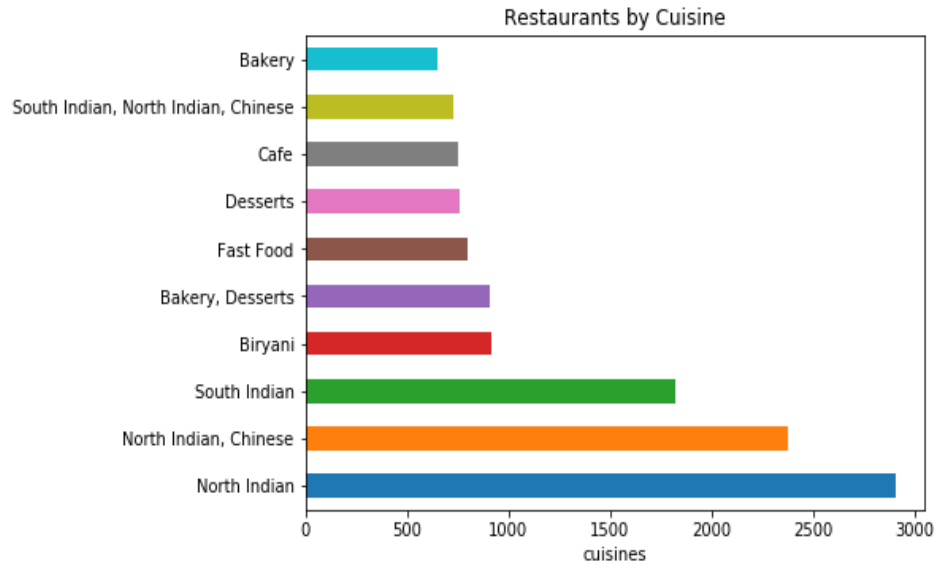


Chart4 :- Top restaurant chains by number of branches in the city

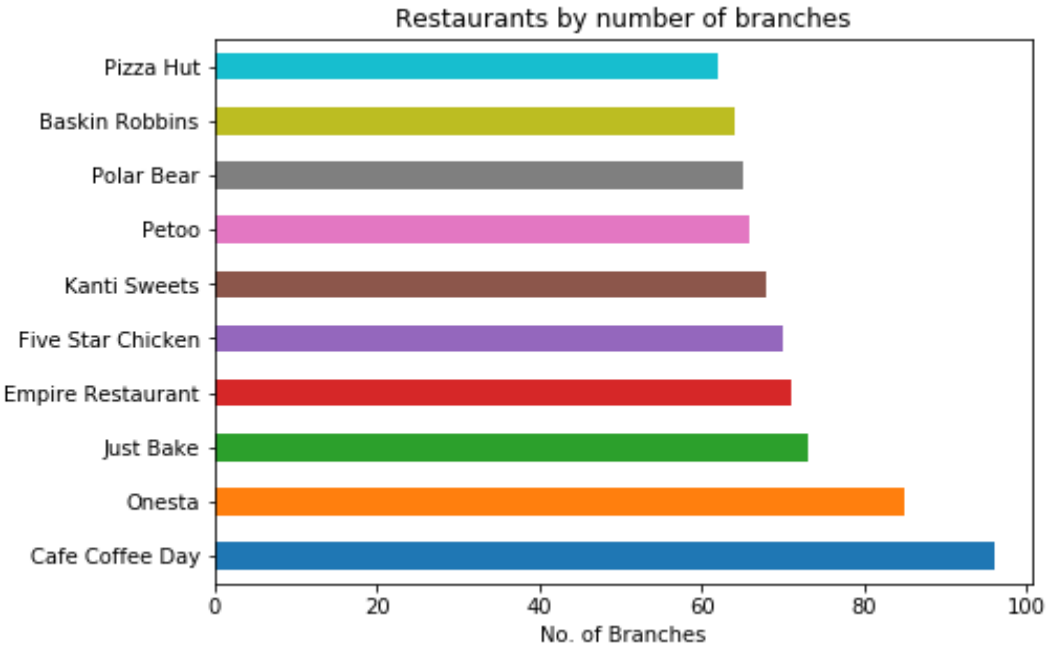
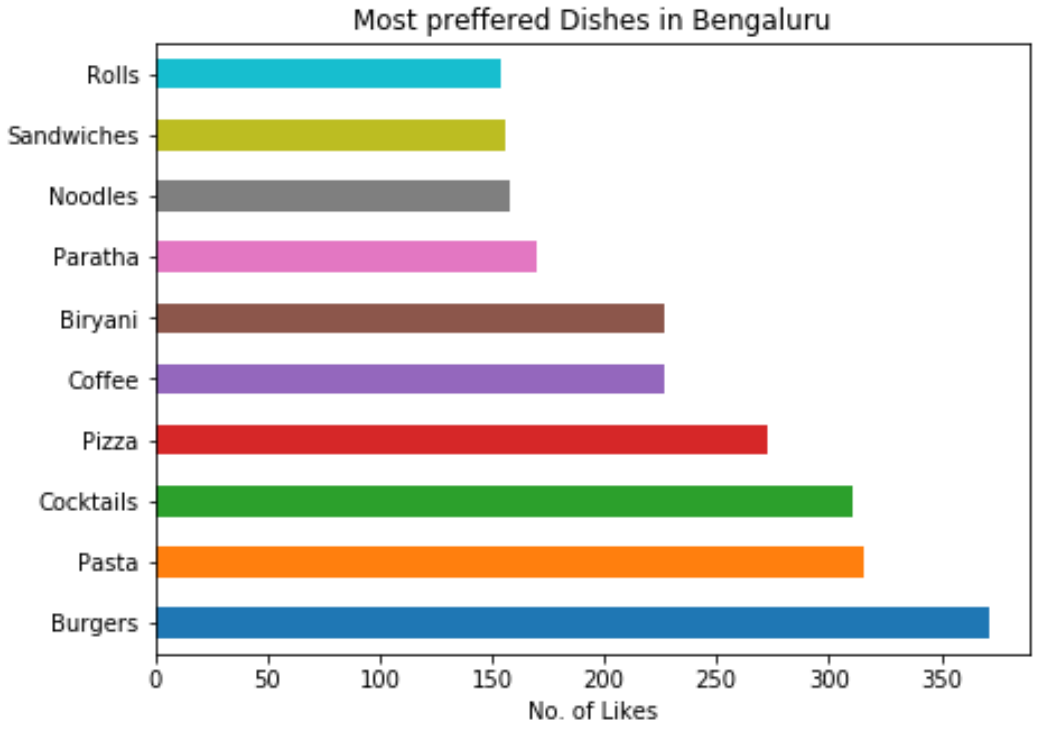
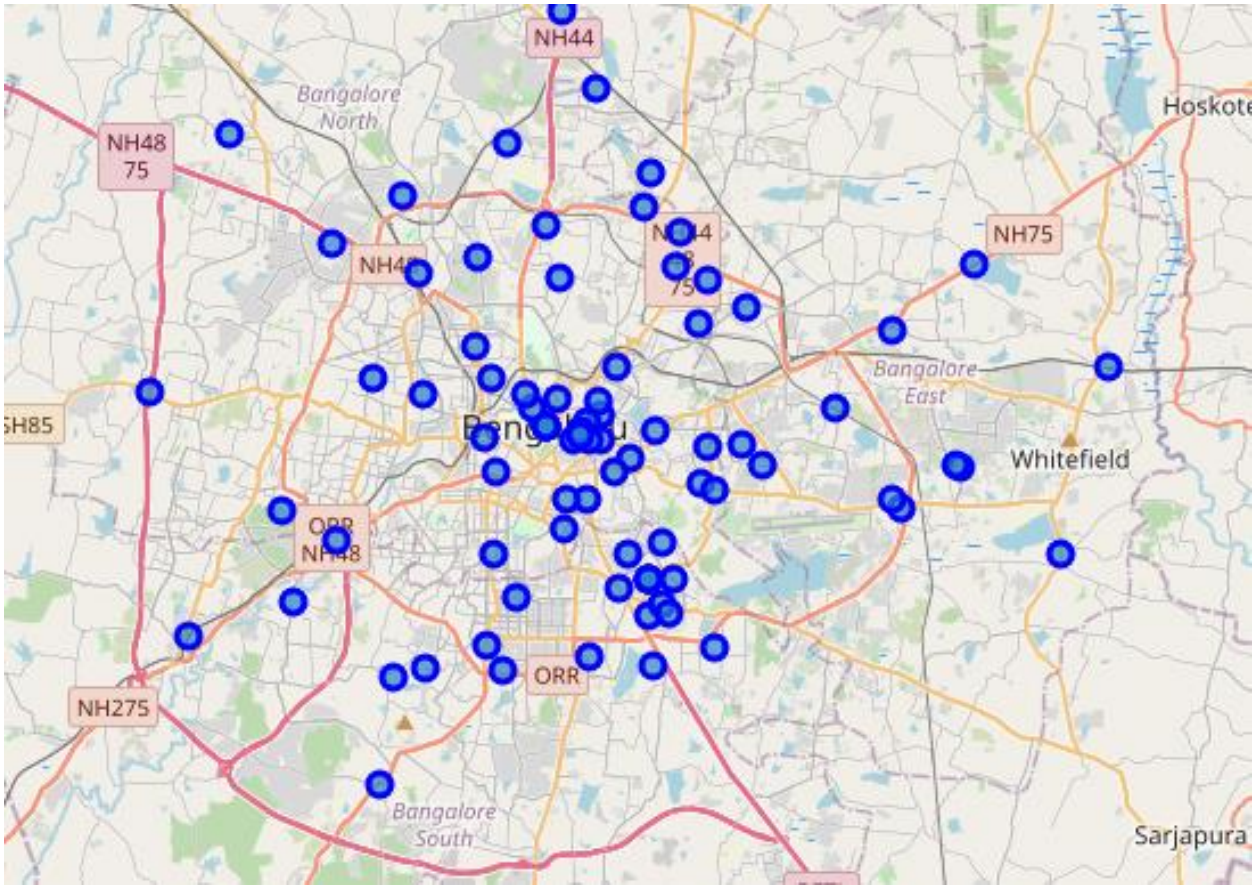


Chart5 :- Most preferred dishes by number of likes posted



Bengaluru Neighborhoods Map : plotted using the geographical coordinates from geopy.



The following is a snapshot of the venue data obtained from **Foursquare** :

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
BTM	12.911276	77.604565	Ginger Tea House	12.912498	77.606373	Bakery
BTM	12.911276	77.604565	Ashirwad Departmental Store	12.915237	77.606057	Department Store
BTM	12.911276	77.604565	Cafe Coffee Day	12.912062	77.606156	Café
BTM	12.911276	77.604565	Flavours of China	12.913776	77.607662	Chinese Restaurant
BTM	12.911276	77.604565	Baskin-Robbins	12.913697	77.607022	Ice Cream Shop
BTM	12.911276	77.604565	Disney Bakery	12.913697	77.607225	Bakery
BTM	12.911276	77.604565	Balaji's Veg	12.913810	77.608180	Vegetarian / Vegan Restaurant
BTM	12.911276	77.604565	Nandi Chat House	12.913762	77.607803	Snack Place
BTM	12.911276	77.604565	Faaso's	12.913709	77.607115	Fast Food Restaurant
BTM	12.911276	77.604565	Leo's Paratha Rolls	12.913787	77.607815	Snack Place

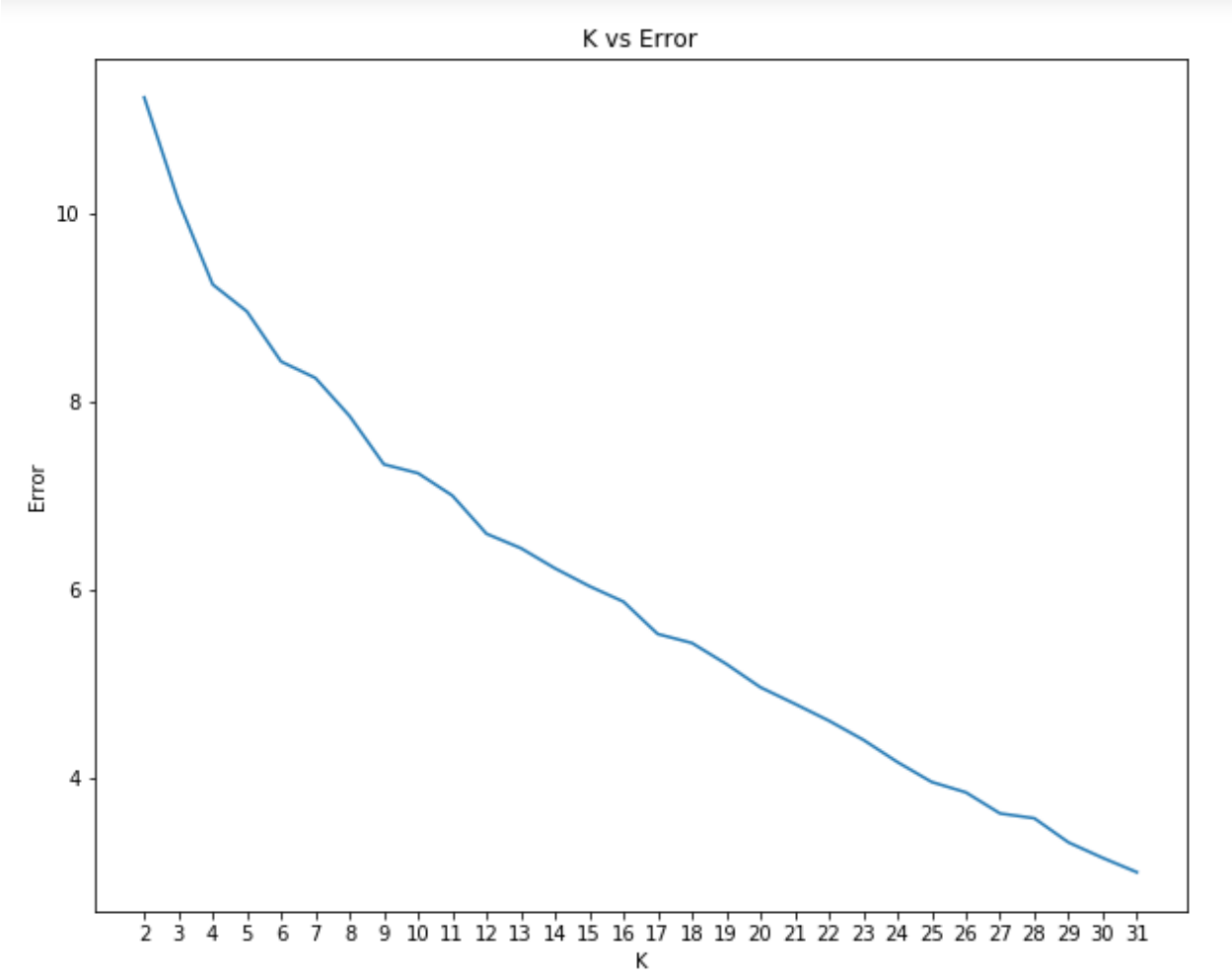
Grouped by count of venues available for each neighborhood

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
BTM	14	14	14	14	14	14
Banashankari	4	4	4	4	4	4
Banaswadi	8	8	8	8	8	8
Basavanagudi	11	11	11	11	11	11
Basaveshwara Nagar	19	19	19	19	19	19
Bellandur	4	4	4	4	4	4
Bommanahalli	4	4	4	4	4	4
Brigade Road	93	93	93	93	93	93
Brookefield	13	13	13	13	13	13
CV Raman Nagar	5	5	5	5	5	5
Central Bangalore	4	4	4	4	4	4
Church Street	32	32	32	32	32	32
City Market	11	11	11	11	11	11
Commercial Street	29	29	29	29	29	29
Cunningham Road	40	40	40	40	40	40

CLUSTERING AND ANALYSIS OF VENUE DATA

K- means clustering is a segmentation technique which groups similar data into different clusters after running the selection according to the algorithm. K – means clustering analysis on the foursquare venue data is performed for segmentation of the neighborhoods into different clusters each having some common characteristics.

Finding the best K value (ie. Number of clusters) is important in k-means. We have derived the best k value as 8 by making and error plot as shown below:



K –MEANS CLUSTERING:

After obtaining the optimal value for number of clusters (K=8) K-means clustering is performed on the dataset produces eight different clusters having similar neighborhoods grouped into each clusters.

The top venues of the eight clusters, their differences and similarities etc. can then be analyzed. A snapshot of data forming the different clusters are shown in the following pages:

Table1 : First Cluster

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
Banashankari	Indian Restaurant	Metro Station	Gym / Fitness Center	Liquor Store	Fast Food Restaurant	Food Court	Food & Drink Shop	Food	Flea Market	Fish Market
Banaswadi	Indian Restaurant	Vegetarian / Vegan Restaurant	Juice Bar	Kerala Restaurant	BBQ Joint	Convenience Store	Andhra Restaurant	Financial or Legal Service	French Restaurant	Food Truck
Basavanagudi	Indian Restaurant	Café	Snack Place	Hookah Bar	Mediterranean Restaurant	Athletics & Sports	Restaurant	Tea Room	Yoga Studio	Food
Bommanahalli	Indian Restaurant	Restaurant	Lake	Auto Garage	Department Store	Dessert Shop	Food Court	Food & Drink Shop	Food	Flea Market
City Market	Indian Restaurant	Plaza	Diner	Middle Eastern Restaurant	Food Truck	South Indian Restaurant	Bookstore	Market	Miscellaneous Shop	Historic Site
Commercial Street	Indian Restaurant	Clothing Store	Hotel	Women's Store	Juice Bar	Bar	Donut Shop	Fast Food Restaurant	Candy Store	Burger Joint
Cunningham Road	Indian Restaurant	Coffee Shop	Chinese Restaurant	Hotel	Café	Pizza Place	Italian Restaurant	Hookah Bar	Middle Eastern Restaurant	Mediterranean Restaurant

Table2 : Second Cluster

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
Bellandur	Restaurant	Park	Capitol Building	Yoga Studio	Electronics Store	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service
Central Bangalore	Restaurant	Park	Capitol Building	Yoga Studio	Electronics Store	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service
East Bangalore	Restaurant	Park	Capitol Building	Yoga Studio	Electronics Store	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service
Electronic City	Restaurant	Park	Capitol Building	Yoga Studio	Electronics Store	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service
North Bangalore	Restaurant	Park	Capitol Building	Yoga Studio	Electronics Store	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service
West Bangalore	Restaurant	Park	Capitol Building	Yoga Studio	Electronics Store	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service

Table3 : Third Cluster

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
Koramangala 3rd Block	Indian Restaurant	Hotel Bar	Hotel	Breakfast Spot	Electronics Store	Food & Drink Shop	Food	Flea Market	Fish Market	Fish & Chips Shop
Langford Town	Indian Restaurant	Hotel	Juice Bar	Electronics Store	Food & Drink Shop	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service
Race Course Road	Indian Restaurant	Hotel	Planetarium	Juice Bar	Coffee Shop	Racetrack	Hotel Bar	Vietnamese Restaurant	Cricket Ground	Eastern European Restaurant
Ulsoor	Indian Restaurant	Light Rail Station	Hotel	Café	Food & Drink Shop	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service
Vasanth Nagar	Indian Restaurant	Hotel	Nightclub	Golf Course	Coffee Shop	Art Gallery	Restaurant	Lounge	Art Museum	Hotel Bar

Table4 : Fourth Cluster

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
Varthur Main Road, Whitefield	Indian Restaurant	Candy Store	Cricket Ground	Cupcake Shop	Food & Drink Shop	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service

Table5 : Fifth Cluster

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
Old Madras Road	Café	Restaurant	Yoga Studio	Food & Drink Shop	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service	Fast Food Restaurant

Table6 : Sixth Cluster

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
Nagarbhavi	Gym	Food Truck	Food & Drink Shop	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service	Fast Food Restaurant	Farmers Market

Table7 : Seventh Cluster

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
Jeevan Bhima Nagar	Indian Restaurant	Café	Department Store	Chinese Restaurant	Food & Drink Shop	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service
Koramangala 8th Block	Indian Restaurant	Ice Cream Shop	Market	Electronics Store	Food & Drink Shop	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service
Magadi Road	Indian Restaurant	Resort	Electronics Store	Food & Drink Shop	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service	Fast Food Restaurant
Majestic	Indian Restaurant	Asian Restaurant	Fast Food Restaurant	Bus Station	Metro Station	Platform	Hotel	Hookah Bar	Electronics Store	Flea Market
Shivajinagar	Indian Restaurant	Fast Food Restaurant	Shoe Store	South Indian Restaurant	Clothing Store	Furniture / Home Store	Juice Bar	Donut Shop	Market	Tea Room

Table8 : Eighth Cluster

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
BTM	Bakery	Snack Place	Indian Restaurant	Department Store	Ice Cream Shop	Coffee Shop	Chinese Restaurant	Café	Fast Food Restaurant	Shopping Mall
Basaveshwara Nagar	Fast Food Restaurant	Ice Cream Shop	Café	Pizza Place	Liquor Store	Bakery	Sporting Goods Shop	Bus Station	Gym	Clothing Store
Brigade Road	Indian Restaurant	Café	Pub	Chinese Restaurant	Lounge	Bar	Coffee Shop	Hotel	Bookstore	Fast Food Restaurant
Brookefield	Indian Restaurant	Fast Food Restaurant	Ice Cream Shop	Breakfast Spot	Department Store	Dessert Shop	Asian Restaurant	Salon / Barbershop	Gym	Multiplex
CV Raman Nagar	Fried Chicken Joint	Electronics Store	Health & Beauty Service	Breakfast Spot	Department Store	Financial or Legal Service	Cupcake Shop	Food Court	Food & Drink Shop	Food
Church Street	Indian Restaurant	Café	Coffee Shop	Lounge	Pub	Bookstore	Music Store	Plaza	Clothing Store	Steakhouse
HBR Layout	Café	Road	South Indian Restaurant	Coffee Shop	North Indian Restaurant	Food	Flea Market	Fish Market	Fish & Chips Shop	Financial or Legal Service
Indiranagar	Pub	Lounge	Cocktail Bar	Indian Restaurant	Café	Restaurant	Bakery	Italian Restaurant	Cupcake Shop	Ice Cream Shop
JP Nagar	Department Store	Indian Restaurant	Pub	Boarding House	Bus Station	Chinese Restaurant	Brewery	Fast Food Restaurant	Bakery	Yoga Studio
Kalyan Nagar	Indian Restaurant	BBQ Joint	Ice Cream Shop	Korean Restaurant	Fast Food Restaurant	Coffee Shop	Café	Italian Restaurant	Electronics Store	Clothing Store

OBSERVATIONS:

We have made an analysis of bengaluru restaurant data according to its neighborhoods and also found out the characteristics of different neighborhoods by getting their most common venue details. We have also segmented the neighborhoods into eight clusters and analysed them. Some interesting findings are:

- BTM is the neighborhood that have most number of restaurants.
- Quick bites and casual dining are more popular types of restaurants in B'lore.
- North Indian cuisine is the most preferred cuisine of the city.
- Cafe coffee day is the largest restaurant chain by number of branches followed by Onesta.
- Top three among the most liked dishes/drinks are Burger, Pasta and Cocktails.
- We have segemented the neighborhoods into eight clusters using k-means technique.
- First cluster and eighth cluster are bigger ones with more number of neighborhoods while remaining are smaller ones.
- First, third and seventh cluster have a large number of retaurants and have them as most common venues

CONCLUSION:

We have performed an analysis of Bengaluru neighborhood data based on restaurant data from kaggle and venues data from foursquare. We have also segmented the different neighborhoods into eight clusters based on the similarities in their venues. I hope this throws some light into the characteristics of the Bengaluru city neighborhoods: especially the food habits. Thank you.