Restaurant@Bangalore, INDIA

FINAL REPORT

A Data Science Capstone Project by NikunjaBihari Padhy

Contents:

Introduction / Business Problem

We have a friend who wish to setup a new restaurant in the city of Bangalore, India which is considered as the IT city of the country. Although he has some moderate level of knowledge about the city's neighborhoods/areas and its' famous venues, but not quite sure in details of each and every neighborhoods & can't really compare every neighborhoods and venues manually like visiting every places personally and exploring to figure out which neighborhood would be most suitable to open up a restaurant and what type of restaurant would be best which is popular amongst Bangalorians and will easily fetch customers as well good profit.

So, the 2 main business problem here with us are as follows:

1. Which neighborhood/area would be best to open a new restaurant in Bangalore.
2. What type of restaurant is most popular & best in Bangalore to easily get customer & earn profit.

Data Requirement

- To explore each and every areas of Bangalore, 1st we need to collect all the neighborhood names & it's Pin codes / Postal codes.
- https://www.mapsofindia.com/pincode/india/karnataka/bangalore - We found this website link by searching in the internet which lists all the neighborhoods & its respective Pincode of Bangalore city.
- We will use the beautiful soup library to extract the neighborhood names and its Pincode from the table in the website.
- If there are multiple neighborhoods for any Pincode, we will merge them into one as these fall into the same area.
- By using the Pincode i.e. postal code of each neighborhood, we will use the google geocoder to get the latitude & longitude for each Pincode.
- Once we have the data of neighborhoods with its respective latitude and longitude values. We can use the Foursquare API to explore each neighborhood area to get top 100 common venues.
- Then we can analyze the common venues around the neighborhoods, segment and cluster the neighborhoods based on the category of venues which are more common to a specific neighborhood.
- We will see the most common type of restaurants found in Bangalore to choose our type of restaurant to open.
- We will segregate the neighborhoods based on the restaurant density and find an area which has less restaurant density at the same time has other business venues and market places which will ensure that there are enough crowd in those areas to easily get customers for the restaurant and eventually to turn it into profit in terms revenue.

Data Science Methodologies Used

We followed the Data requirement methodology to identify required data for this exercise which is listed in the previous section.

Then we collected the data from identified sources and validated it by going through the collection and plotting maps of Bangalore using the Data visualization.

Cleaned the Data which has no relevant information like no venues returned by Foursquare API for some neighborhoods.

Then we grouped & sorted the dataset to find out the no. unique venue categories found and their frequency in each neighbourhood as well the top 10 most common venue categpries for each neighbourhood.

One-hot encoding was used for **Normalization** of the dataset to use the machine learning techniques

The **K-means clustering** machine learning was used for clustering the neighborhoods **because** -> we wanted to classify the neighborhoods among different categories as per the nearby common venue categories and their frequency found for each neighbourhood area.

Then we further used the **venue category density** for each neighbourhood to find out the neighborhoods where there is very less no. of restaurants.

We again plotted the map via **folium data visualization** for final set of recommended neighborhoods to find out the nearest and relevant place within the Bangalore city.

Results

Based on the 99 neighborhood areas of Bangalore city collected form the data source, the nearby venues are collected for each neighbourhood area limited to a max of 100 no.s within 1000 metres.

No. of venues found across all neighborhoods – 2366
No. of unique category of venues – 207
No Venues found for neighborhoods – 3 (Removed from the Dataset in Data clean-up)

Top 10 venue categories across all the neighborhoods -

Indian Restaurant	**428** ->>> Preferred Restaurant Type Recommendation
Café	161
Fast Food Restaurant	85
Ice Cream Shop	72
Pizza Place	70
Department Store	70
Coffee Shop	68
Hotel	64
Bakery	59
Chinese Restaurant	58

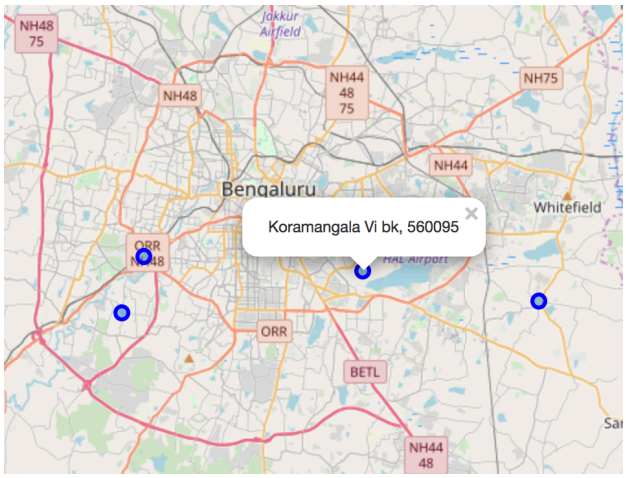
Segmentation of neighbourhood as per the common venue categories and frequency.

	Count of Neighborhoods	
Cluster 0	44	Place of restaurants and eateries
Cluster 1	1	Less restaurants and eateries, outskirt neighborhood
Cluster 2	**50**	**Mix-N-Match of all types of venue categories**
Cluster 3	1	Less restaurants and eateries, outskirt neighborhood

We found 4 neighborhood areas in Cluster 2 with '0' restaurant density that means there were no restaurants types found in top 10 common venue categories in these 4 areas.

	Neighborhood	Restaurant Density
0	Gunjur, Muthusandra, Panathur, Vartur	0
1	Kenchanahalli, Rajarajeshwarinagar	0
2	Koramangala Vi bk	0
3	Nayandahalli	0

Then the final four on the map ->



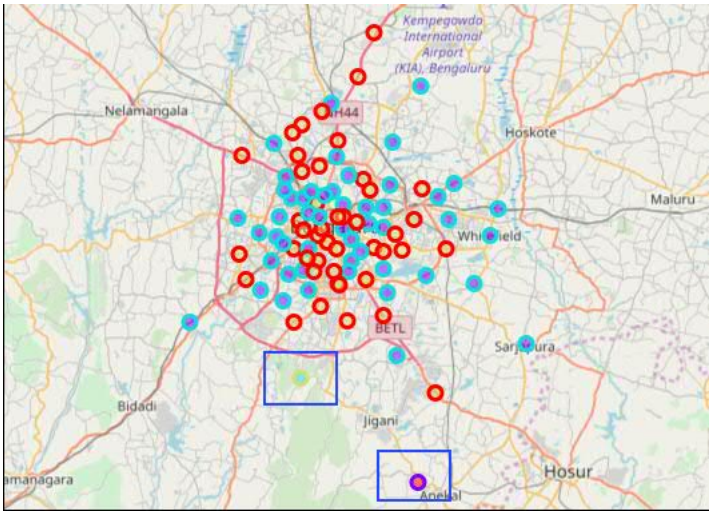
Discussion

We have found venue category “Indian Restaurant” being the top contributor with 428 venues out of the 2366 total venues found all across Bangalore. Also, this category is the major contributor for all the neighborhoods in Cluster 0 which was flagged as places of restaurants & eateries.

The finding is “Indian Restaurants” are most popular among Bangalorians, so to open up an Indian Restaurant in Bangalore will be a safe bet.

When we got the four clusters plotted, it seemed like Cluster 1 & 3 are most perfect places as there were less restaurants and eateries. But those were too far away and in outskirts of Bangalore as identified in the blue boxes in below map.

Map of four Clusters ->



Cluster 1 & 3 -

There are only one 'Pincode Area' in each of these 2 clusters and they are looking perfect for opening a restaurant as there are very less no. of restaurants and other places like business areas and flea markets which are indication of potential customer-based areas for restaurant. But this is good if someone is looking to open a restaurant in outskirts of Bangalore as these locations are very far from the actual city market area.

Cluster 0 -

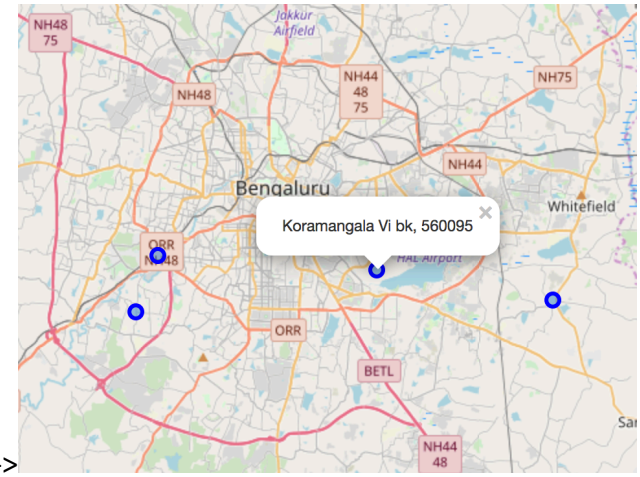
These neighborhoods have lot of restaurants and eateries. Although these places would have high business turn around that’s why there are so many eateries already established, but competition will be very very tough for a new restaurant to open and settle.

Cluster 2 -

These neighborhoods have mix of all types of venues, let’s explore and analyse which areas have less restaurants and eateries and other business venues which will fetch potential customers for a restaurant.

As we found the final four from Cluster 2 with '0' restaurant density, plotted in the results section. By looking at the geographical position, "Koramangala VI bk – 560095" is 1st preference since it's within the prime city limit boundary. And 2nd will be the one on the borderline ORR of the Bangalore city i.e. "Nayandahalli – 560039" to open a restaurant.

Then the final four on the map ->



Conclusion

As per the data available from mapofindia.com website & foursquare API, after plotting the final set of neighborhoods in the Map above, we can confidently say that below two in our list are spot on & are near around the core city as well for opening a restaurant

Preferred Locations for Opening a Restaurant ->

1. Koramangala Vi bk - 560095
2. Nayandahalli - 560039

Type of Restaurant –> Indian Restaurant