Capstone Project

on

Find suitable location to open new restaurant or coffee shop

Project by Puranjan Pathak

Date: 15-Sep-2019

Introduction/ Business problem 2
Methodology used 2
Data 2
Results/ Findings 2
Discussions 6
Conclusion 7

Introduction/ Business problem

Bangalore is one of the hi-tech cities of India. There are a number of IT parks and business establishments setup in Bangalore.

Because of the growing number of establishments in this city, there is good a demand for eateries in locations where there are less number of restaurants or coffee shops.

In this project, the aim is to identify locations of Bangalore where there are less number of restaurants or coffee shops and suggest some of these locations as candidate for opening up new restaurants or coffee shops.

This project is part of report for fulfillment of the course offered by coursera.

Methodology used

I will be using the K-means clustering to analysis the neighborhoods.
Will also be using foursquare APIs to collect venue details.

Data

1. Country wise geo data of location from web http://www.geonames.org/export/zip/
2. Data of Bangalore ZIP codes from CSV file

Results/ Findings

Get list of venues in Bangalore and store the data into a dataframe.

(3126, 7)

	locality	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Bangalore G.P.O.	12.9914	77.5944	Millers 46	12.991666	77.594207	Steakhouse
1	Bangalore G.P.O.	12.9914	77.5944	Ujwal Bar & Restaurant	12.992280	77.594473	Indian Restaurant
2	Bangalore G.P.O.	12.9914	77.5944	Alliance Française	12.991232	77.596723	Concert Hall
3	Bangalore G.P.O.	12.9914	77.5944	Infinitea	12.987157	77.594835	Tea Room
4	Bangalore G.P.O.	12.9914	77.5944	Imperial Restaurant	12.991150	77.593837	Indian Restaurant

Now encode the data

	locality	ATM	American Restaurant	Andhra Restaurant	Arcade	Art Gallery	Arts & Crafts Store	Asian Restaurant	Auto Garage	BBQ Joint	...	Tourist Information Center	Track Stadium	Train Station	Vegetarian / Vegan Restaurant	Wa P
0	Bangalore G.P.O.	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
1	Bangalore G.P.O.	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
2	Bangalore G.P.O.	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
3	Bangalore G.P.O.	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
4	Bangalore G.P.O.	0	0	0	0	0	0	0	0	0	...	0	0	0	0	

5 rows × 145 columns

Now group rows by neighborhood taking mean of the frequency of occurrence of each category

	locality	ATM	American Restaurant	Andhra Restaurant	Arcade	Art Gallery	Arts & Crafts Store	Asian Restaurant	Auto Garage	BBQ Joint	...	Tourist Information Center	Track Stadium	Train Station	Vegetarian / Vegan Restaurant
0	A F Station Yelahanka	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.0
1	Adugodi	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.0
2	Agara	0.0	0.022727	0.022727	0.0	0.0	0.0	0.0	0.0	0.022727	...	0.0	0.0	0.0	0.0
3	Amruthahalli	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.000000	...	0.0	0.0	0.0	0.0
4	Anandnagar (Bangalore)	0.0	0.000000	0.000000	0.0	0.0	0.0	0.0	0.0	0.000000	...	1.0	0.0	0.0	0.0

5 rows × 145 columns

Let's print each neighbourhood along with top 5 common venues

```
topVenues = 5

for v in group_by_venue['locality']:
    print("----"+v+"----")
    temp = group_by_venue[group_by_venue['locality'] == v].T.reset_index()
    temp.columns = ['venue','freq']
    temp = temp.iloc[1:]
    temp['freq'] = temp['freq'].astype(float)
    temp = temp.round({'freq': 2})
    print(temp.sort_values('freq', ascending=False).reset_index(drop=True).head(topVenues))
    print('\n')
```

```
----A F Station Yelahanka----
                venue  freq
0         IT Services   1.0
1  American Restaurant   0.0
2      Motorcycle Shop   0.0
3        Movie Theater   0.0
4            Multiplex   0.0


----Adugodi----
```

country	pincode	locality	state	district	city	latitude	longitude
IN	560001	Bangalore G.P.O.	Karnataka	Bangalore	Bangalore North	12.9914	77.5944
IN	560001	Legislators Home	Karnataka	Bangalore	Bangalore North	12.9914	77.5944
IN	560001	Vasanthanagar	Karnataka	Bangalore	NaN	12.9914	77.5944
IN	560001	Mahatma Gandhi Road	Karnataka	Bangalore	Bangalore North	12.9914	77.5944
IN	560001	Vidhana Soudha	Karnataka	Bangalore	Bangalore North	12.9914	77.5944

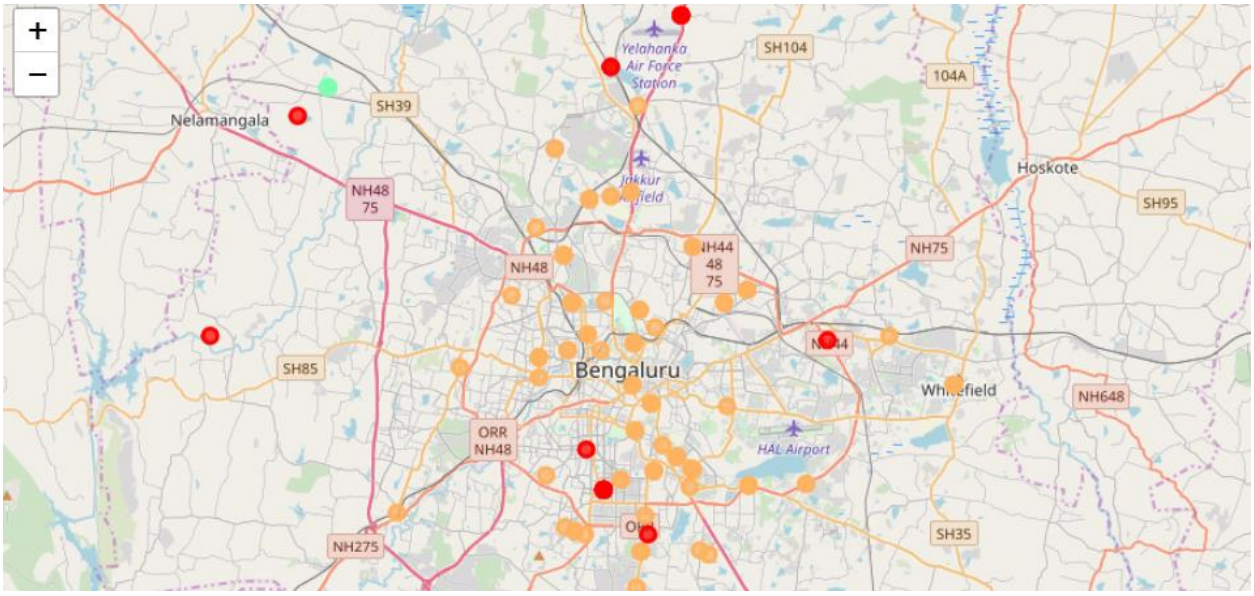
Now create a new dataframe and display the top 10 venues for each neighborhood.

	locality	1st Most Frequent Venue	2nd Most Frequent Venue	3rd Most Frequent Venue	4th Most Frequent Venue	5th Most Frequent Venue	6th Most Frequent Venue	7th Most Frequent Venue	8th Most Frequent Venue	9th Most Frequent Venue	10th Most Frequent Venue
0	A F Station Yelahanka	IT Services	Hotel	Fishing Store	Fast Food Restaurant	Farmers Market	Event Service	Electronics Store	Donut Shop	Dive Bar	Diner
1	Adugodi	Fast Food Restaurant	Indian Restaurant	Bus Station	Design Studio	Zoo	Dive Bar	Farmers Market	Event Service	Electronics Store	Donut Shop
2	Agara	Café	Indian Restaurant	Ice Cream Shop	Coffee Shop	Pub	Lounge	Bar	Mexican Restaurant	Brewery	Breakfast Spot
3	Amruthahalli	Hyderabadi Restaurant	Indian Sweet Shop	Café	Italian Restaurant	Resort	Badminton Court	Coffee Shop	Fast Food Restaurant	Deli / Bodega	Department Store
4	Anandnagar (Bangalore)	Tourist Information Center	Flea Market	Fast Food Restaurant	Farmers Market	Event Service	Electronics Store	Donut Shop	Dive Bar	Diner	Zoo

Run k-means to cluster the neighbourhoods

	Clusters	locality	1st Most Frequent Venue	2nd Most Frequent Venue	3rd Most Frequent Venue	4th Most Frequent Venue	5th Most Frequent Venue	6th Most Frequent Venue	7th Most Frequent Venue	8th Most Frequent Venue	9th Most Frequent Venue	10th Most Frequent Venue
0	2	A F Station Yelahanka	IT Services	Hotel	Fishing Store	Fast Food Restaurant	Farmers Market	Event Service	Electronics Store	Donut Shop	Dive Bar	Diner
1	4	Adugodi	Fast Food Restaurant	Indian Restaurant	Bus Station	Design Studio	Zoo	Dive Bar	Farmers Market	Event Service	Electronics Store	Donut Shop
2	4	Agara	Café	Indian Restaurant	Ice Cream Shop	Coffee Shop	Pub	Lounge	Bar	Mexican Restaurant	Brewery	Breakfast Spot
3	4	Amruthahalli	Hyderabadi Restaurant	Indian Sweet Shop	Café	Italian Restaurant	Resort	Badminton Court	Coffee Shop	Fast Food Restaurant	Deli / Bodega	Department Store
4	3	Anandnagar (Bangalore)	Tourist Information Center	Flea Market	Fast Food Restaurant	Farmers Market	Event Service	Electronics Store	Donut Shop	Dive Bar	Diner	Zoo

Visualize the clusters



	Clusters	locality	1st Most Frequent Venue	2nd Most Frequent Venue	3rd Most Frequent Venue	4th Most Frequent Venue	5th Most Frequent Venue	6th Most Frequent Venue	7th Most Frequent Venue	8th Most Frequent Venue	9th Most Frequent Venue	10th Most Frequent Venue
0	2	A F Station Yelahanka	IT Services	Hotel	Fishing Store	Fast Food Restaurant	Farmers Market	Event Service	Electronics Store	Donut Shop	Dive Bar	Diner
1	4	Adugodi	Fast Food Restaurant	Indian Restaurant	Bus Station	Design Studio	Zoo	Dive Bar	Farmers Market	Event Service	Electronics Store	Donut Shop
2	4	Agara	Café	Indian Restaurant	Ice Cream Shop	Coffee Shop	Pub	Lounge	Bar	Mexican Restaurant	Brewery	Breakfast Spot
3	4	Amruthahalli	Hyderabadi Restaurant	Indian Sweet Shop	Café	Italian Restaurant	Resort	Badminton Court	Coffee Shop	Fast Food Restaurant	Deli / Bodega	Department Store
4	3	Anandnagar (Bangalore)	Tourist Information Center	Flea Market	Fast Food Restaurant	Farmers Market	Event Service	Electronics Store	Donut Shop	Dive Bar	Diner	Zoo

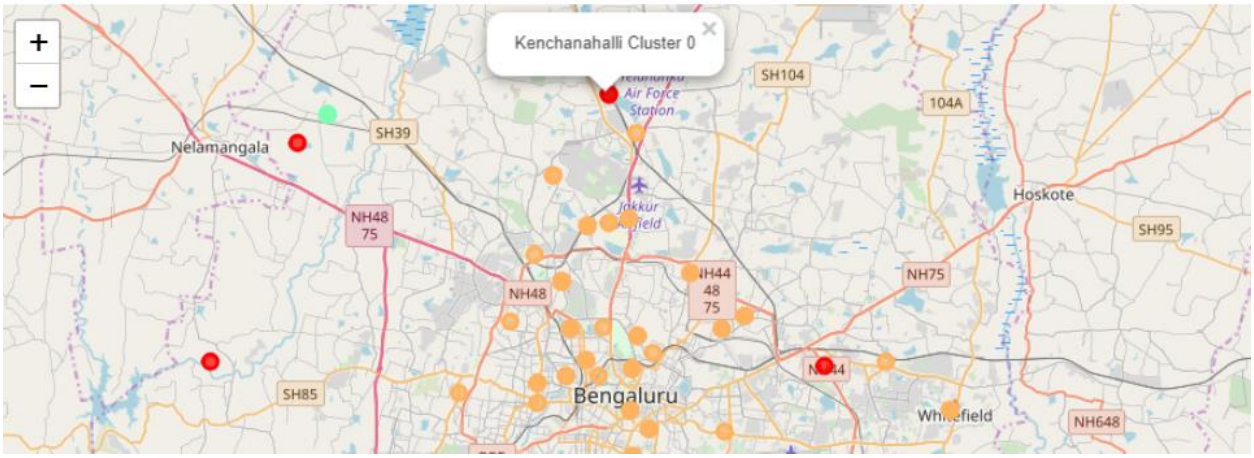
By reading the data of 10 most common venues out of top 5 clusters, we see that there is low number of “coffee shops” at locality “A F Station Yelahanka”. And there are not many “fast food restaurants” in the same locality.

Discussions

The top 5 clusters captured most of the venues which includes “IT Service”, “Hotel”, “Fast Food Restaurant”, “Resort”, “Pub”, “Coffee shop”, “Electronic store” and many different types of venues.

Conclusion

Considering the above clustered data, we can conclude that it will be better to start a new “coffee shop” at cluster 0 location than opening a new “restaurant”.



The most suitable location would be between localities “Kenchanahali” and “A F Station Yelahanka” in cluster 0.