**Data Science - Capstone Project : Pizza Eat-Out place data analysis of Bengaluru city**

**A. Introduction**

**A.1. Description & Discussion of the Background**

Bangalore, officially known as Bengaluru is the capital of the Indian state of Karnataka. It has a population of about 10 million and a metropolitan population of about 8.52 million, making it the third most populous city and fifth most populous urban agglomeration in India.[12] Located in southern India on the Deccan Plateau, at a height of over 900 m (3,000 ft) above sea level, Bangalore is known for its pleasant climate throughout the year. Its elevation is the highest among the major cities of India.Bangalore is widely regarded as the "Silicon Valley of India" (or "IT capital of India") because of its role as the nation's leading information technology (IT) exporter.

Bengaluru is a city with a high population and population density. Being such a crowded city leads the fast-food chain business and social sharing places in the city where the population is dense. When we think of it by the investor, we expect from them to prefer the areas or suburban places where there is a lower competition and the type of business they want to install is less intense. If we think of the city residents, they may want to choose the regions where options are more and prices are economical. However, it is difficult to obtain information that will guide investors in this direction, nowadays.

Considering below hypothetical scenario, we can take help of data-science to nail down this problem to some extent.

Some of my entrepreneur friends wants to open new pizza eat-out place in Bangalore city. One day when we met for casual meet, they explained me about their plans. There basic requirements includes,

1. To open pizza eat-out in a place, where there are less numbers of pizza places, so that competition will be less and profit will be more.

2. The place should be economical. In sense investment price should be less and profit should be more.

In this project we will try to find an optimal location to set-up pizza eat-out. Specifically, this report will be targeted to stakeholders interested in opening Pizza Eat-out in Bangalore, Karnataka.

Since there are lots of pizza eat-out in Bangalore, we will try to detect locations that are not already crowded with Pizza places.

We will use our data science powers to generate a few most promising neighbourhoods based on this criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

When we consider all these requirement points, we can create a map and information chart where the pizza place index is placed on Bengaluru and each area is clustered according to the number of counts.

**A.2. Data Description**

Based on definition of our problem, factors that will influence our decision are:

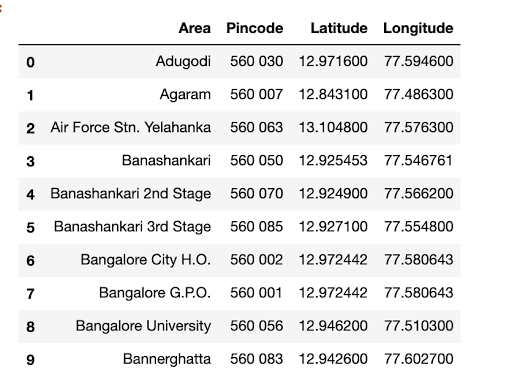
* Number of existing pizza places in the neighbourhood Following data sources will be needed to extract/generate the required information: I have prepared .xls file which has main details including area pin-code, latitude and longitude. This data set I have committed in kaggle.

[Kaggle-Dataset](https://www.kaggle.com/hegdetapan/bangaloreareaspincodewithlatitudelongitude)

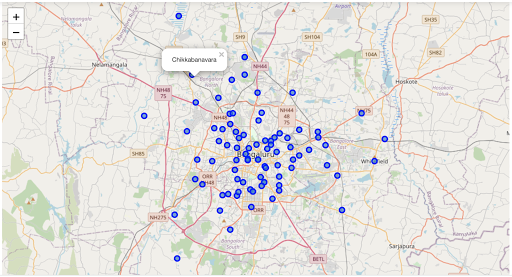
* Number of pizza places and location in every neighbourhood will be obtained using Foursquare API

**B. Methodology**

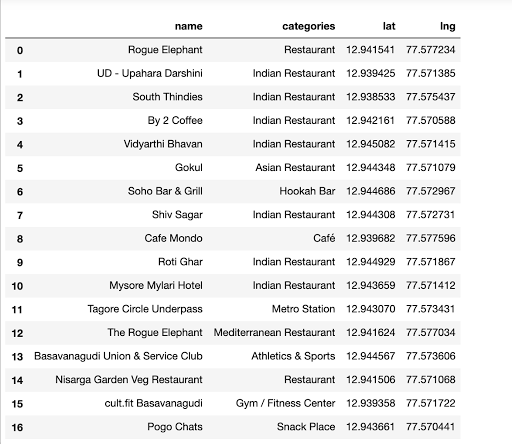
As a database, I used CSV file which I have prepared. My master data which has the main components Area, pin-code, Latitude and Longitude informations of the city.

[](https://1.bp.blogspot.com/-MZ8GUm7-vcI/XzN9monG2VI/AAAAAAAADpE/81fBEwGRoP0SPPWOSuU4h9JrtsTFO5zCQCLcBGAsYHQ/s836/Screen%2BShot%2B2020-08-12%2Bat%2B10.45.11%2BAM.png)

I used python **folium** library to visualise geographic details of Bengaluru city and its suburban areas and I created a map of Bengaluru with areas superimposed on top. I used latitude and longitude values to get the visual as below:

[](https://1.bp.blogspot.com/-C00N02xsDqk/XzN_YvUWHWI/AAAAAAAADpQ/ETdSP1lHrk0SwgPdIb2BM8lTkMpekLzhQCLcBGAsYHQ/s1988/Screen%2BShot%2B2020-08-12%2Bat%2B11.02.44%2BAM.png)

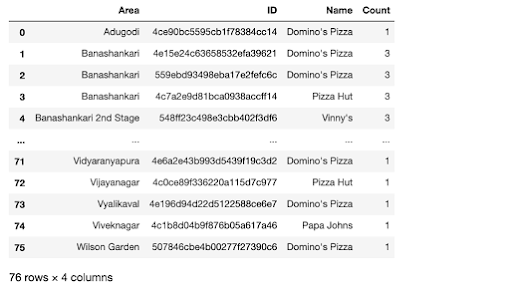
I Used the Foursquare API to explore neighbourhood areas and segment them. Here as an example I used 'Basavanagudi' area and tried to list down restaurants, eat-out places using Foursquare API.

[](https://1.bp.blogspot.com/-RgX_CBgp1U8/XzOAu7qH3GI/AAAAAAAADpc/rAPNAseohmU7DC5OmCaEdrqa6INX3YHTgCLcBGAsYHQ/s1174/Screen%2BShot%2B2020-08-12%2Bat%2B11.06.50%2BAM.png)

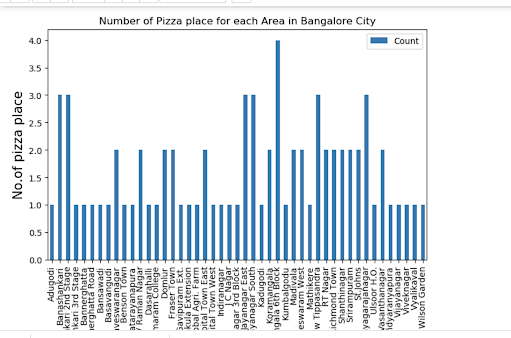
Next, Prepare  neighbourhood list that contains pizza places.

[](https://1.bp.blogspot.com/-JxBnAgnWCd4/XzOBa3vz0II/AAAAAAAADpk/yZCB73y2Erg9kcBgZrri3a_dVeFBHunVACLcBGAsYHQ/s1498/Screen%2BShot%2B2020-08-12%2Bat%2B11.12.47%2BAM.png)

We can see that, there 76 places which are related to pizza.

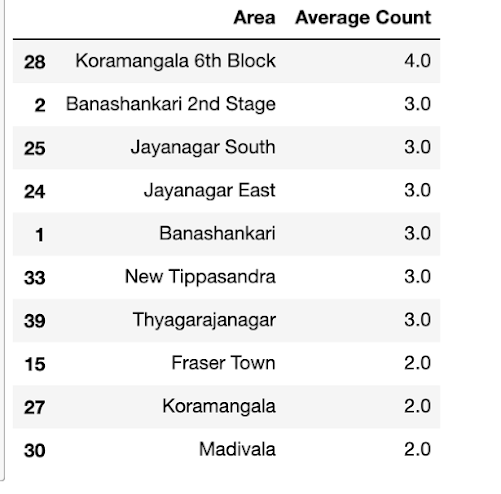
[](https://1.bp.blogspot.com/-zdXRI5J6D8Y/XzOCNV8LbmI/AAAAAAAADpw/vCHbCbRl5PAce_8F6sy7A5SXDnbuwMNQgCLcBGAsYHQ/s1282/Screen%2BShot%2B2020-08-12%2Bat%2B11.13.49%2BAM.png)

Let's plot the bar graph to get to know which area has more number of pizza places and which has less.(comparatively)

[](https://1.bp.blogspot.com/-GT9aQkZ4zLs/XzOCvX9lARI/AAAAAAAADp4/lWVe4xPh8WMTGuLclWHWf9Pq8IZ21QLIwCLcBGAsYHQ/s1672/Screen%2BShot%2B2020-08-12%2Bat%2B11.18.17%2BAM.png)

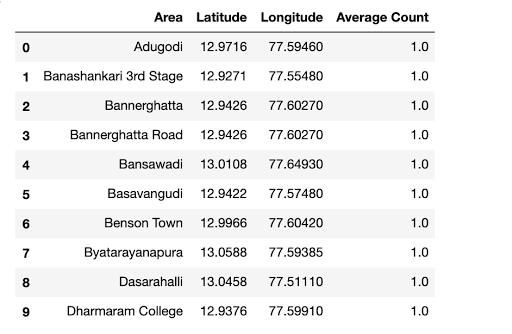
From above graph, it clear that 'Koramangala 6th Block' has highest number of pizza places.

Now, Let's calculate average count for each area.

[](https://1.bp.blogspot.com/-PfC_085Ejxs/XzODfQeVxRI/AAAAAAAADqE/Um95pKDURAoABGA2Xstxil3RcfmOqb9pgCLcBGAsYHQ/s626/Screen%2BShot%2B2020-08-12%2Bat%2B11.21.19%2BAM.png)

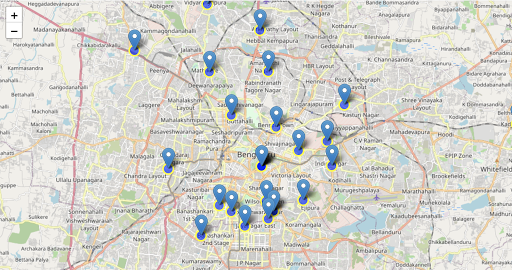
we will consider all the areas with average count lesser than 2.0 to visualise on map.

Let's get list of areas in which pizza places counts are less than 2.

[](https://1.bp.blogspot.com/-XRDR5-mhPws/XzOEABa3sEI/AAAAAAAADqM/UrapZDkKMggu4_Gr-SIOhk8p3NvDYCrrgCLcBGAsYHQ/s942/Screen%2BShot%2B2020-08-12%2Bat%2B11.23.44%2BAM.png)

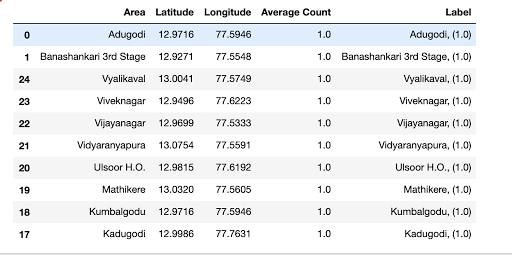
Now above we have the pizza places which has count lesser  2.0 and their Geo-Co-ordinates

Lets Visualise it!



**C. Results**

Let’s merge those new variables with related information in our main master table



As we can see below are the list of areas which has less pizza places and it would be good for business, if we consider these areas to open new pizza eat-out.

1. Adugodi

2. Banashankari 3rd Stage

3. Vyalikaval

4. VivekNagar

5. Vidyaranyapura

6. Ulsoor

7. Peenya

8. Mathikere

9. Kumbalgodu

10. Kadugodi

**Limitations :**

The Count is highly dependent on Foursquare API details.

There might be some other pizza eat-outs which are not listed in foursquare database.

**D. Discussion**

As I mentioned before, Bengaluru is a big city with a high population density with many of fast-food options.

For more detailed and accurate guidance, the data set can be expanded and the details of the areas or street can also be drilled.

I also performed data analysis through this information by adding the coordinates of areas. In future studies, these data can also be accessed dynamically from specific platforms or packages.

**F. Conclusion**

Being pizza lover, I would like to explore more places which serves varieties of flavours. I believe this small hypothetical project would help some of us.(Either of the cases, If you are planning to open new pizza restaurant or just want to try out some new pizza places).

This project can be modified for more options.

For eg : I have plan to extend this project to get details about Udupi south Indian restaurant chains.

**G. References:**

[1] Bengaluru — Wikipedia - [Bengaluru-wikipedia](https://en.wikipedia.org/wiki/Bangalore)

[2] Kaggle - [Kaggle Dataset](https://www.kaggle.com/hegdetapan/bangaloreareaspincodewithlatitudelongitude)

[3] Forsquare API - [Foursquare API](https://foursquare.com/)

[4] Google Maps

[5] Distancefrom.in - [Latitude-Longitude](https://www.distancesfrom.com/in)