**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](#_Toc19451309)

[**Methodology used** 2](#_Toc19451310)

[**Data** 2](#_Toc19451311)

[**Results/ Findings** 2](#_Toc19451312)

[**Discussions** 6](#_Toc19451313)

[**Conclusion** 7](#_Toc19451314)

# **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.

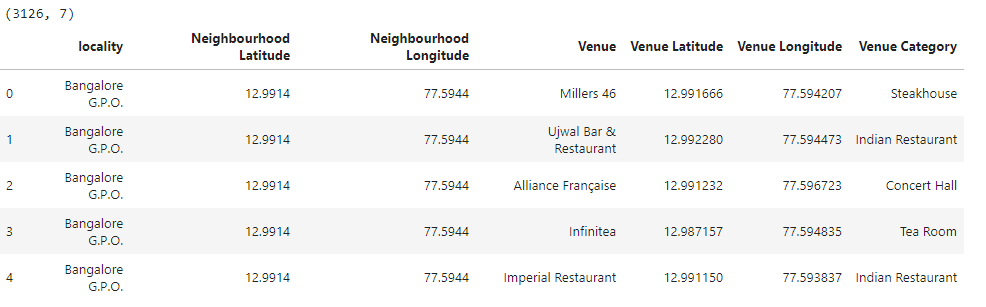
Will also be using foursquare APIs to collect venue details.

# **Data**

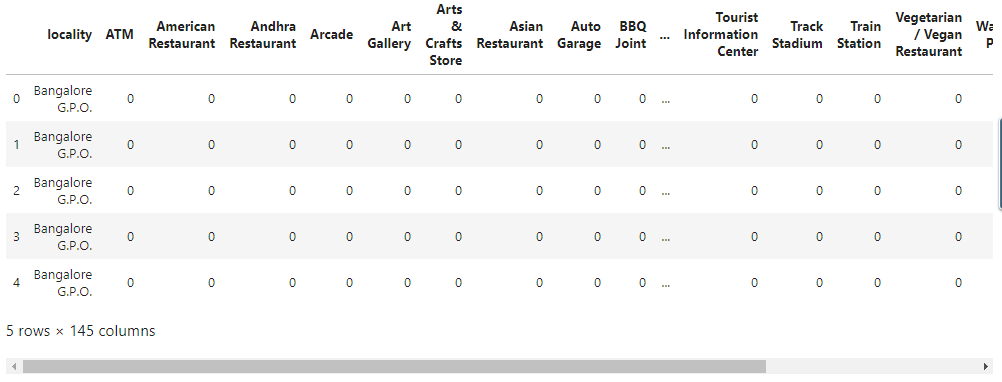
1. Country wise geo data of location from web <http://www.geonames.org/export/zip/>
2. Data of Bangalore ZIP codes from CSV file

# **Results/ Findings**

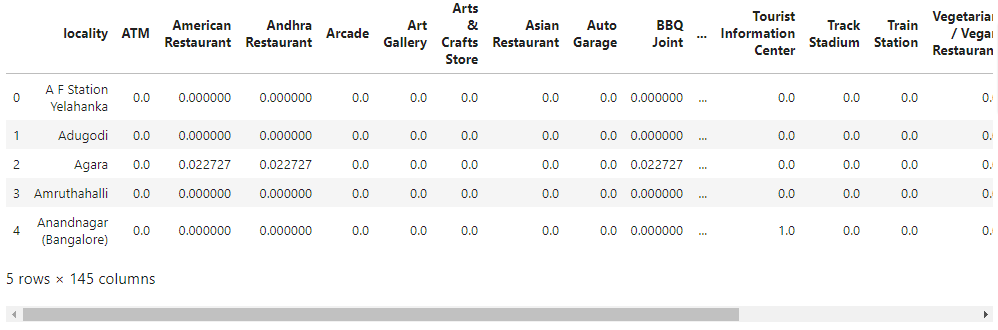
Get list of venues in Bangalore and store the data into a dataframe.



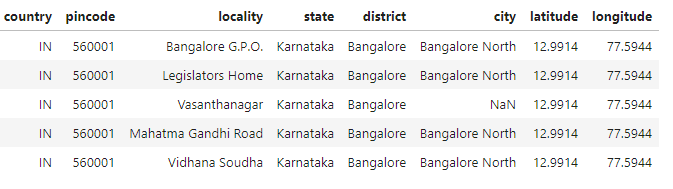
### Now encode the data



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



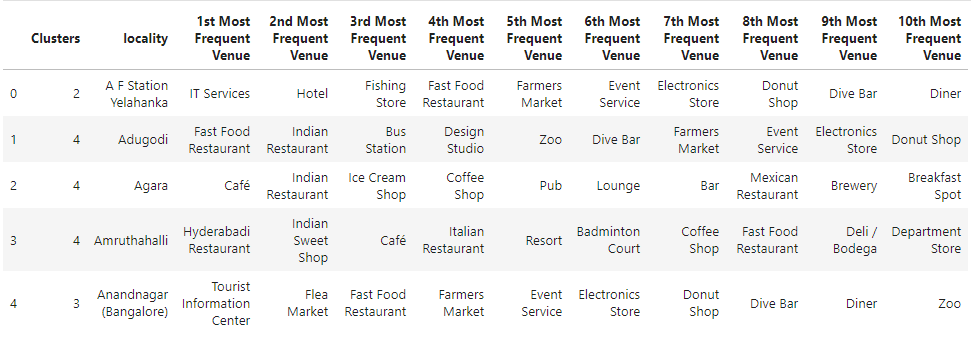




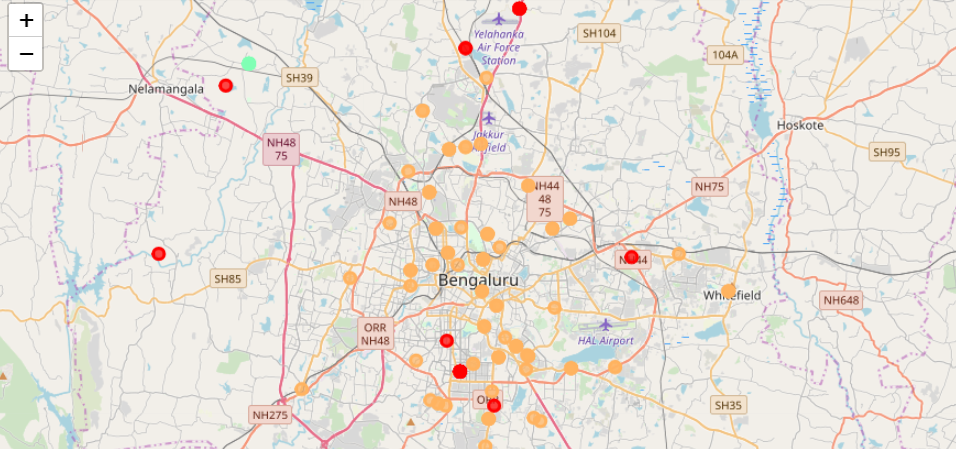
## Now create a new dataframe and display the top 10 venues for each neighborhood.



### Run k-means to cluster the neighbourhoods



### Visualize the clusters





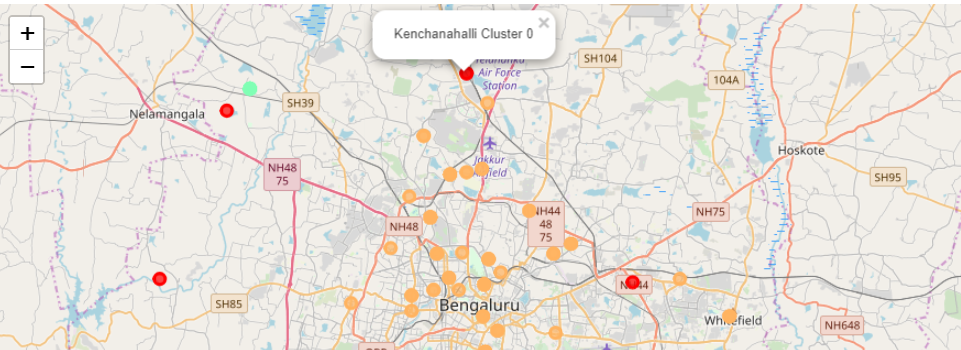
By reading the data of 10 most common venues out of top 5 clusters, we see that there is low number of “coffee shops” at locality “A F Station Yelahanka”. And there are not many “fast food restaurants” in the same locality.

# **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 are most suitable would be between localities “Kenchanahali” and “A F Station Yelahanka” in cluster 0.