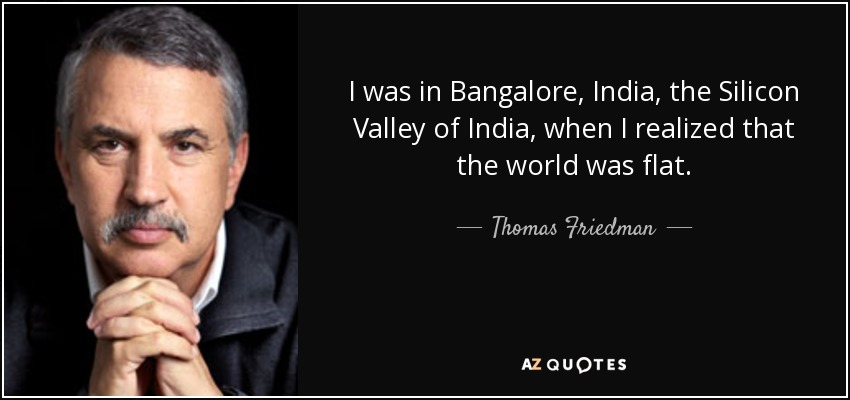
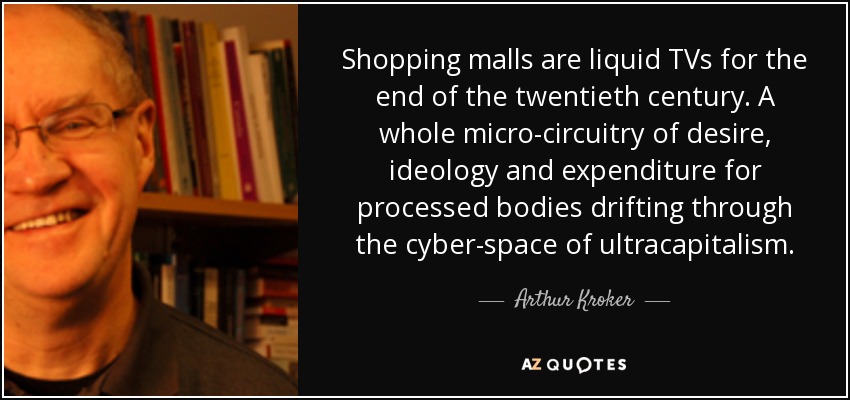
INTRODUCTION

**Bangalore?**

Bangalore(the city where I reside) is the capital of the Indian state Karnataka. It is known as ‘The Garden City’ and is the third most populated city in India. It is a large metropolis still undergoing growth and is a home to various scientific institutes and prestigious colleges. It is the leading IT exporter of country giving it the title ‘ Silicon Valley Of India’ .Many foreigners reside here because of presence of various multinational companies. With a diverse demographic, it is a huge economic and cultural hub.



**What is a mall and why is it important?**

****

According to Wikipedia, a shopping mall can be termed as ‘ a large indoor shopping centre anchored by department stores’ .Nowadays we see huge establishments consisting of theatres, restaurants and many stores(retailing clothing and electronics to name few categories) termed as malls. The main purpose of a mall is to get people outside their homes to enjoy comforts provided by it. In the current scenario, many youngsters prefer to hang out with their friends at such establishments. They are also very popular due to the fact that all your needs can be addressed in one location through them. As stated in the pic above, malls are considered to be the manifest of capitalism itself.

In Bangalore , there are many popular malls like Mantri Square in Malleshwaram, Central in JP Nagar and PVR Forum in Koramangala. With a piece of the economy of Bangalore being tempting to many prospective mall owners, it would be confusing for them to select a location to build one in this huge city.

BUSINESS PROBLEM

The main objective is to find an area for prospective mall owners to build a mall with the following features kept in mind:

1.Accessibility i.e Availability of public transport close to candidate area

2.Lack of competing venues like theatres and other malls nearby

I also considered areas with popular open grounds like fields and parks satisfying above as risky because of public backlash. Nevertheless, those areas have been included in another category where interested parties should look before they leap.

AUDIENCE

This project has been designed for those businessmen who are interested in building a mall in Bangalore. It is helpful by narrowing down list where said businessmen can survey area and then take a suitable decision.

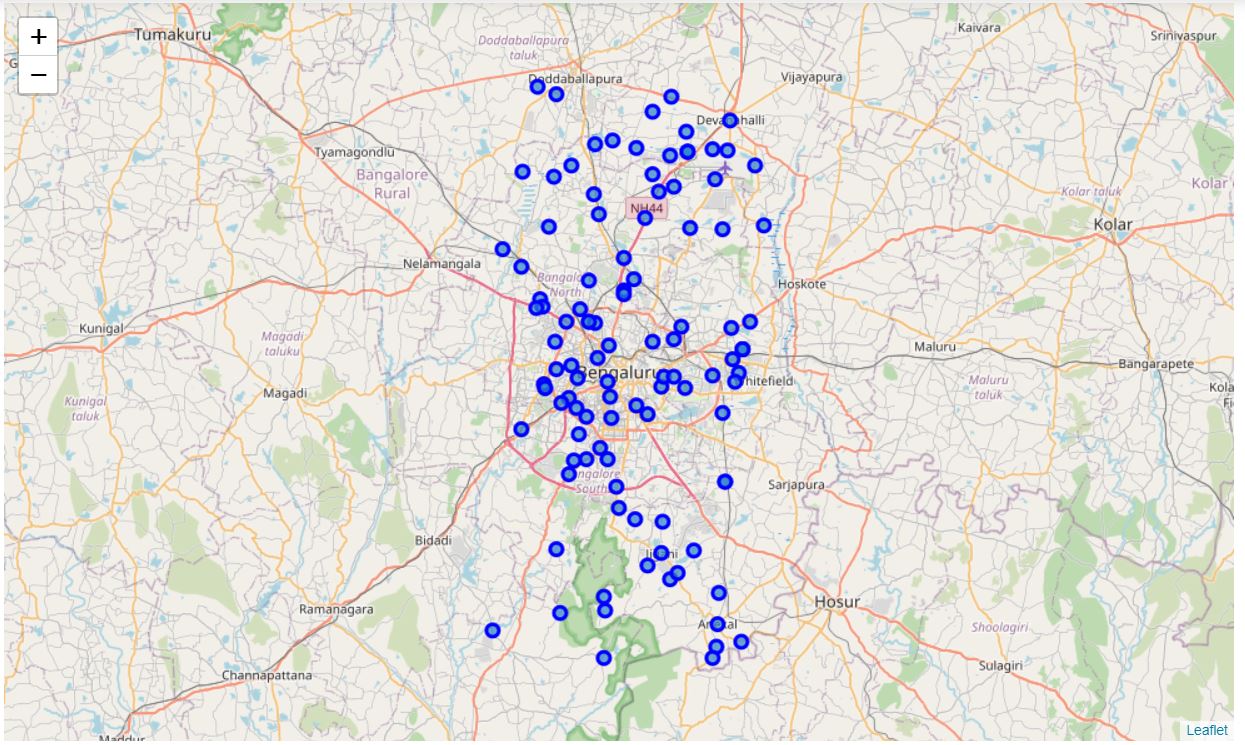
DATA

1.Bangalore neighbourhoods dataset obtained from [Kaggle](https://www.kaggle.com/rmenon1998/bangalore-neighborhoods) consisting of co-ordinates(latitude and longitude) of each of the 109 neighbourhoods of Bangalore.

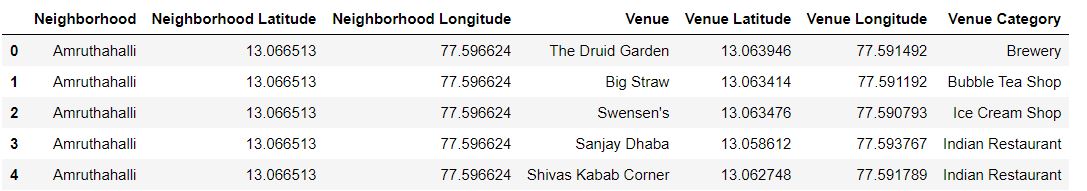
2.Data of public transport stations and nearby competing venues like theatres or malls close to each neighbourhood obtained by use of [Foursquare API](https://developer.foursquare.com/)’s explore service with which we cluster neighbourhoods based on similarity in popular venues within a given radius(taken to be 1 km here)

METHODOLOGY

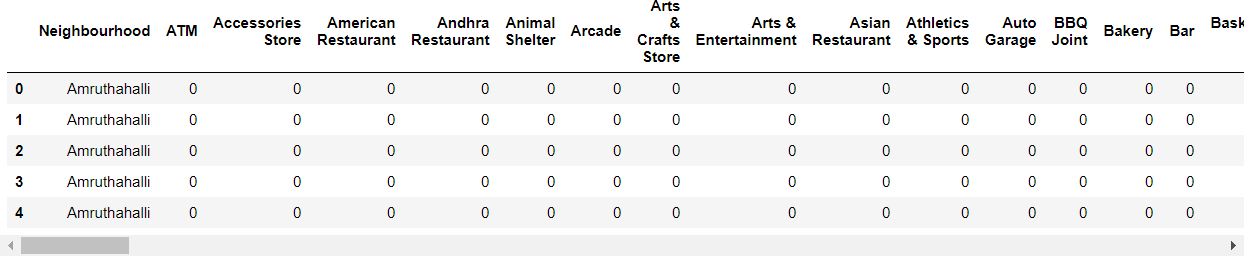
The Kaggle data set mentioned above was downloaded and stored as a dataframe. On using Folium, visualization of each neighbourhood in Bangalore was done. However,many areas outside Bangalore were also included so by removing neighbourhoods with co-ordinates outside a certain range, I concentrated on areas of Urban Bangalore.



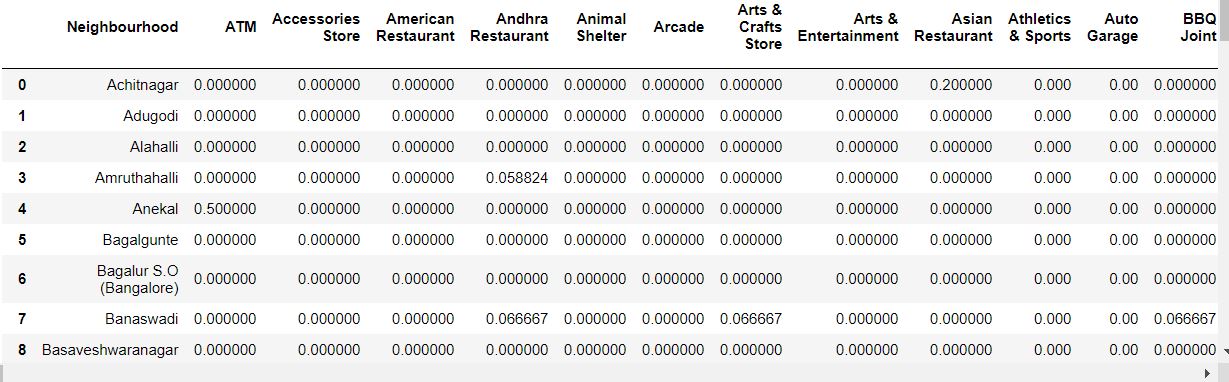
After that, I used the Foursquare API to explore popular venues within a 1 km radius limited to a max of 100 results returned.(I chose 1 km as radius as it was a popular answer when I inquired a few people about distance they were willing to travel from a public transport station to a mall).The API returned 1270 venues of 181 unique types from the 110 neighbourhoods fed to it. Below is a picture of first 5 rows of dataframe containing information about each venue returned by the API for each neighbourhood given as input.

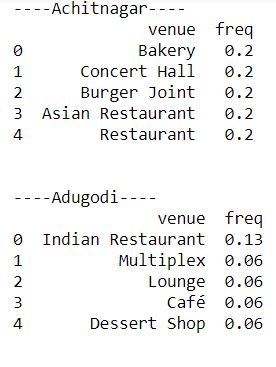


To get a summary of popular venues,one hot encoding was done for each venue in each neigbourhood.



Now all venues under each neighbourhood are grouped and mean value for each venue category is obtained.



Top 5 venues in some neighbourhoods:  


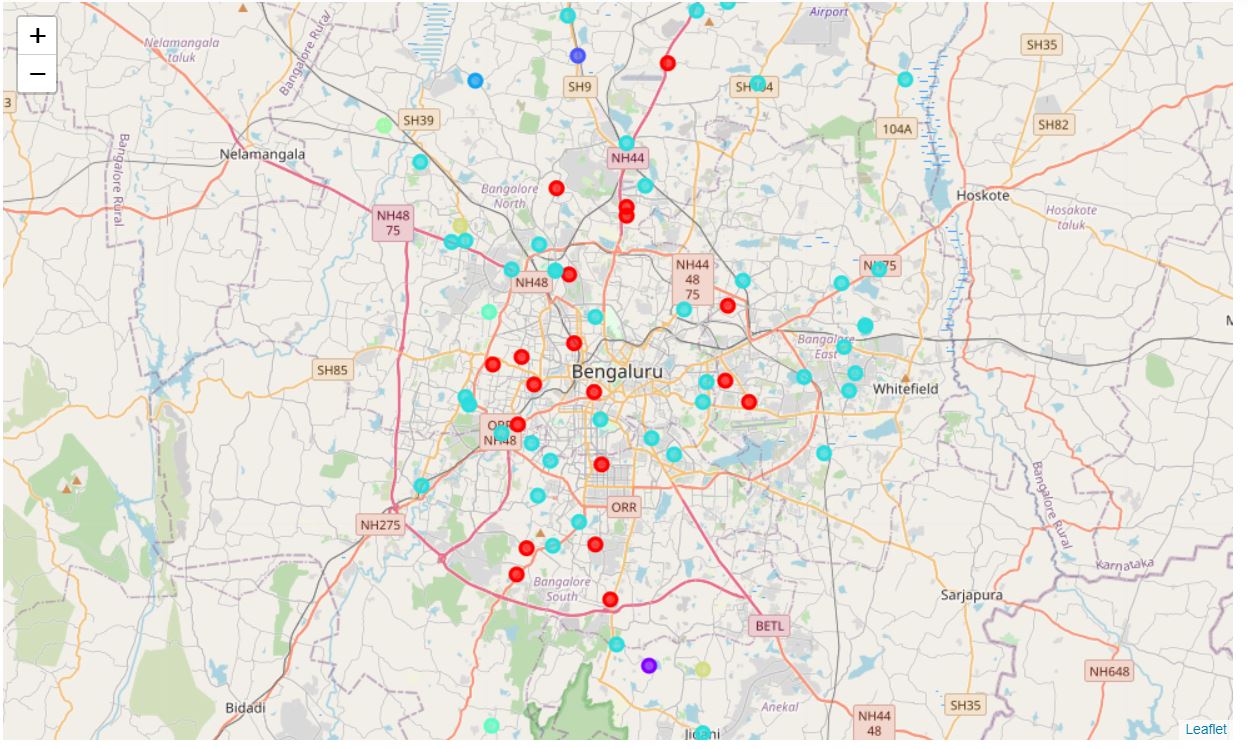
Now top 10 popular venues in each neighbourhood was found by creating a function formatting a dataframe with this information and passing each neighbourhood to this function.



Now K Means clustering algorithm is applied on above.After varying the value of number of clusters and observing the results,I decided to go with 10 as it gave 2 dominant clusters with information I needed and less number of neighbourhoods I didn’t need to look into further.



Visualizing using Folium:



Now I merged 2 clusters in which public transport was popular and created a new dataframe.