Coursera Capstone

Opening of a new café in Delhi, India

By: Shashank Drolia

Introduction:

Café is a small eating and drinking establishment, historically a coffee house, usually featuring a limited menu. Originally these establishments served only coffee.

There are many café in Delhi and more are being built. It has become a trend mainly among the youths and as a result the location of the café is of utmost importance or else it will be a complete failure.

Business Problem:

The primary objective of the project is to analyse and select the best locations in the city of Delhi. The project is timely as there is an increasing demand of the café.

Target Audience:

This project is useful for the investors and developers looking to invest in the café in the city of Delhi.

Data:

The above mentioned page was used to extract the neighbourhood in Delhi. We then used web scrapping techniques to extract the data the from the mentioned page. We used beautifulSoup package of python to do so.

After this we used the Python geocoder package to find the location of all the extracted neighbourhood. After this we use

foursquare API to find the venues in the mentioned neighbourhood. IN the next 'Methodology' section I will explain the steps taken in the completion of this project.

Methodology:

Firstly we need to get a list of all the neighbourhood in Delhi. This was made available to us through Wikipedia page.

We then use the BeautifulSoup package to then extract this data. Now to find the location of all this neighbourhood we use another python package- geocoder package.

We then find the top venues in each of the neighbourhood. We then analyse the data for "café". And lastly, we will perform clustering on the data by using k-means algorithm (machine learning).

It is one of the most popular unsupervised learning algorithm. We divide the neighbourhood in three clusters based of the frequency of the occurance of the café in the neighbourhood
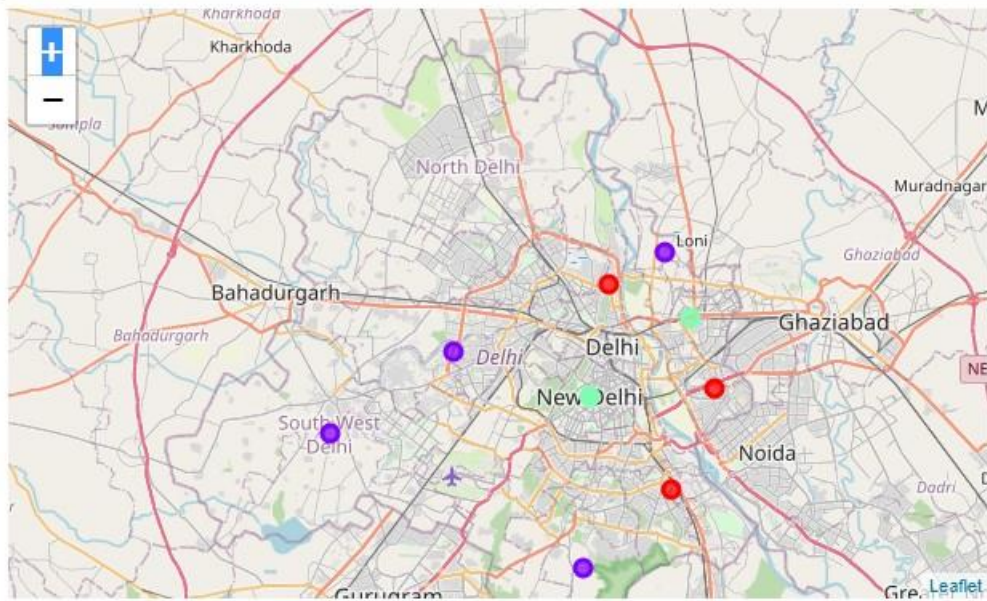
Results:

~ Cluster 0: these neighbourhood have high concentration of café.

~Cluster 1: these have no existence of Café.

~Cluster 2: these have moderate no of cafes.

The map below depicts the clustering of Neighbourhood in Delhi.

Here red colour corresponds to Cluster 0

Purple colour Cluster 1

And Green colour Cluster 2.

Discussion:

We can observe that there is no existence of café in the neighbourhood belonging to cluster 1 , So it might be a good opportunity to open an café there.

Note:

For this project we have only considered one factor i.e. the frequency of occurrence of café. There are many factors which can affect this decision like the Offices nearby, Youth population etc. But these data were unavailable to us.

Conclusion:

We conclude that neighbourhood in cluster 1 is the best location to open a café as we would have competition there. The findings of this this project is very useful to the investors who are eyeing to open a café in Delhi. no

Reference:

The above mention Wikipedia page.