**Clustering of Chennai Neighbourhoods**

1. Introduction

**Chennai is the capital of Tamilnadu, a southern state in India.**

**According to the 2011 Indian census,** it is the [sixth-most populous city](https://en.wikipedia.org/wiki/List_of_most_populous_cities_in_India) and [fourth-most populous urban agglomeration in India](https://en.wikipedia.org/wiki/List_of_million-plus_urban_agglomerations_in_India). The city together with the adjoining regions constitutes the [Chennai Metropolitan Area](https://en.wikipedia.org/wiki/Chennai_Metropolitan_Area), which is the [36th-largest urban area by population in the world](https://en.wikipedia.org/wiki/List_of_urban_areas_by_population).

As part of this project, we will be analysing the Chennai neighbourhood by taking the Chennai neighbourhood data from Wikipedia and scrapping the data for further analysis.

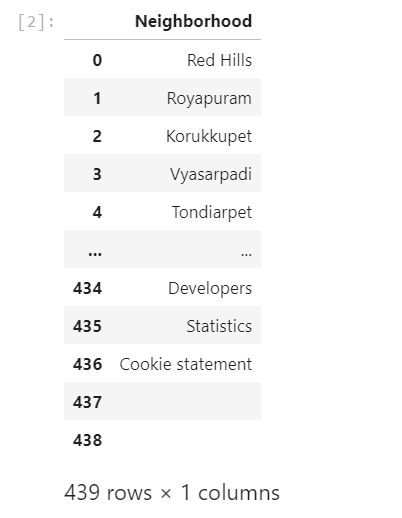
By using Four square API, the geographic data will also be taken and it will be merged with scrapped Wikipedia data on neighbourhoods.

Once it is done, clustering the neighbourhood data using K-Means clustering Algorithm will be carried out.. Finally these clusters help the various stake holders to establish their businesses, to sell the houses so on.

1. Data Section:

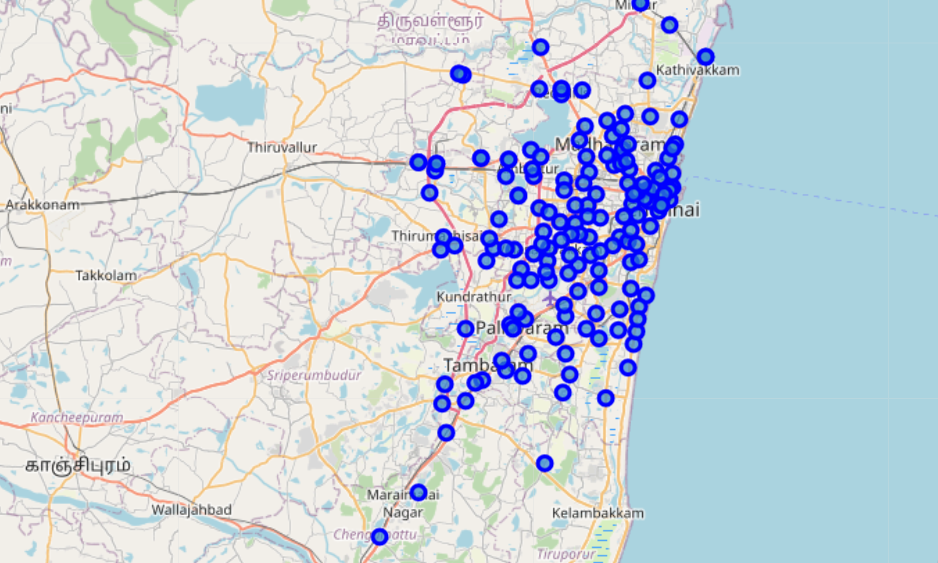
These are the sources from which the data will be taken:

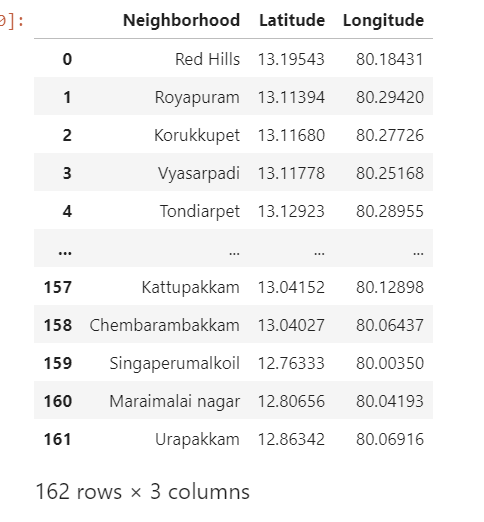
1. From the Wikipedia page about the Chennai neighbourhoods
2. From the four square API about the Geographic data
3. Chennai flood data from the crowd sourced sources
4. Clustering the Chennai neighbourhood using K-Means clustering Algorithm
5. Neighbourhood details:



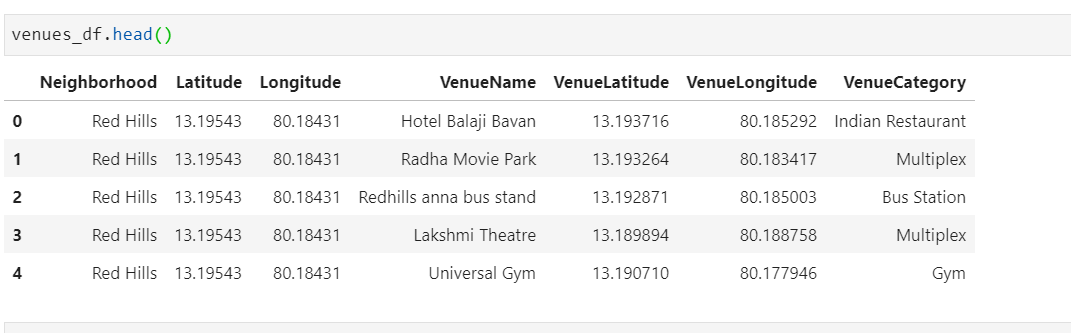
1. Chennai Map

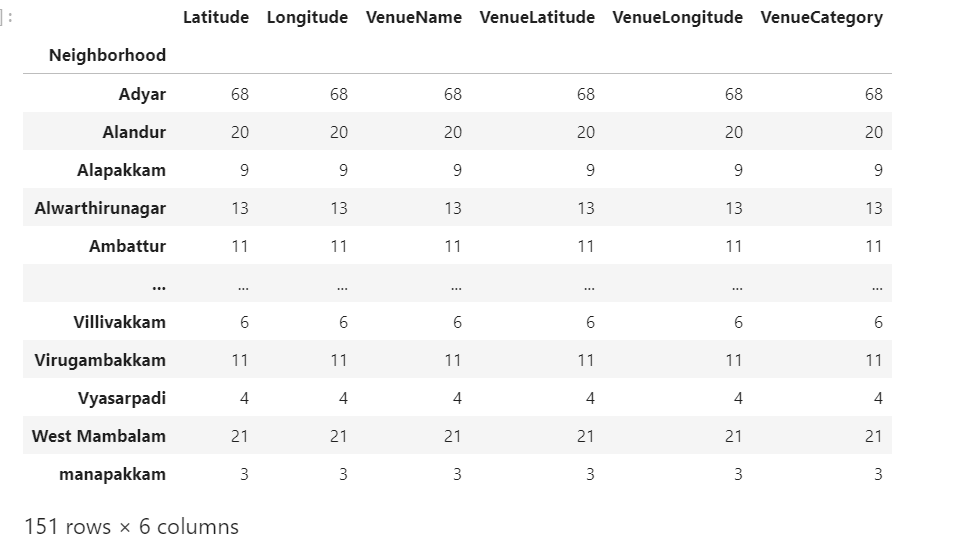
Chennai Map :





1. Methodology Section:
2. From the Wikipedia page , the neighbourhood data is taken
3. From the four square API about the Geographic data is taken and merged with neighbourhood data
4. Clustering the Chennai neighbourhood using K-Means clustering Algorithm



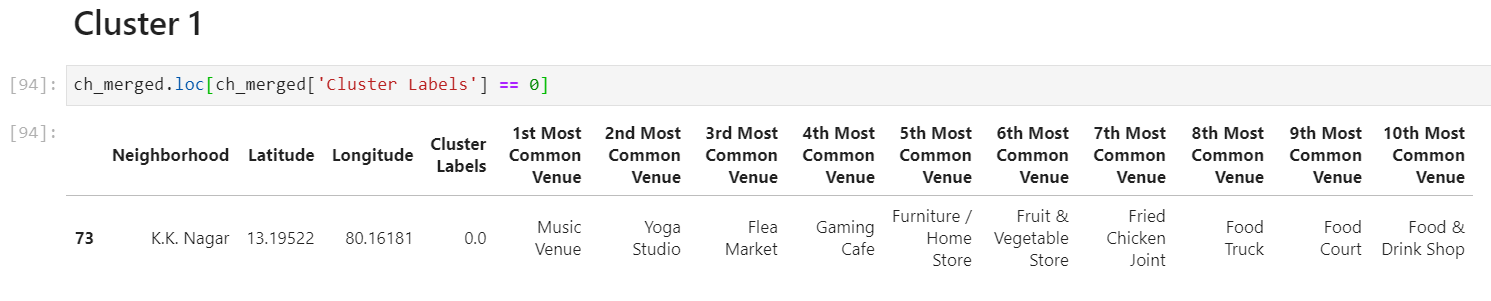


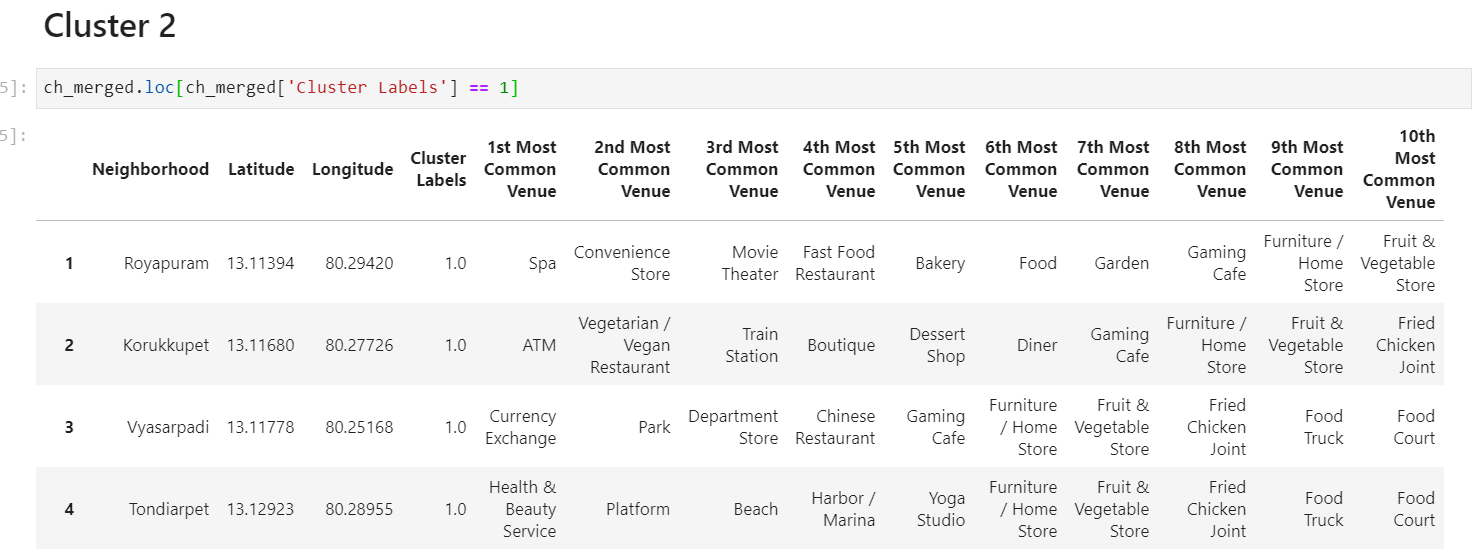


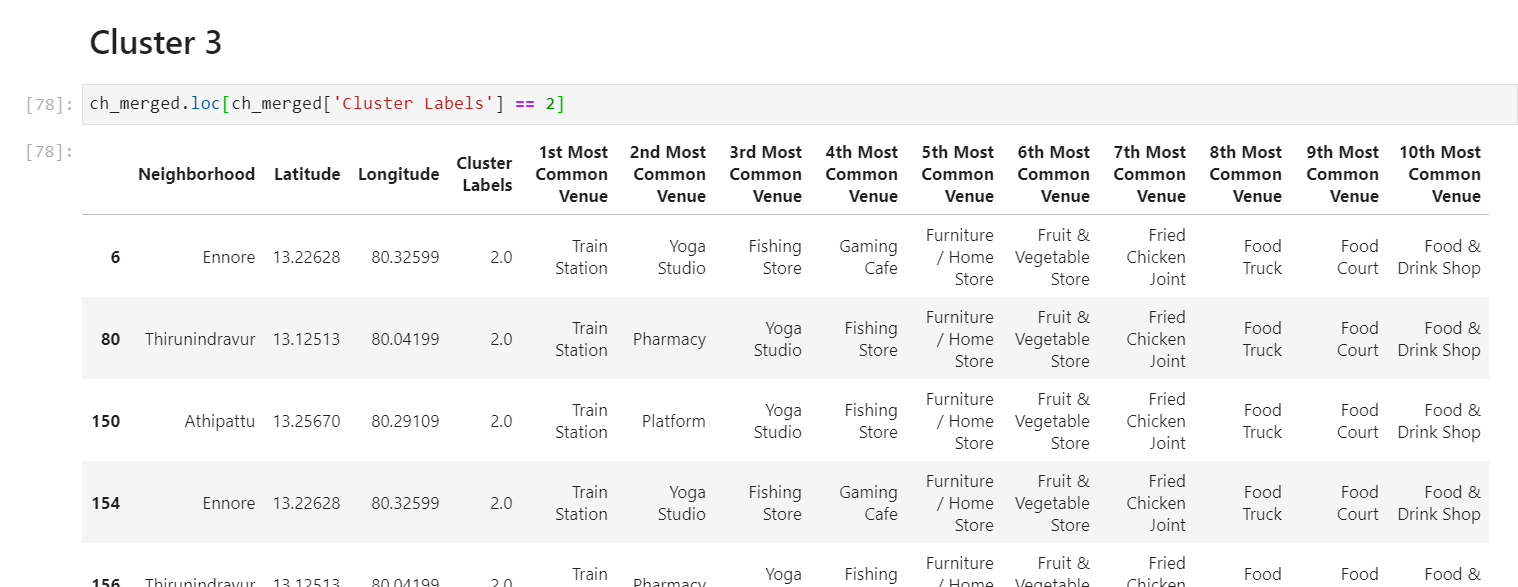
1. Results section:

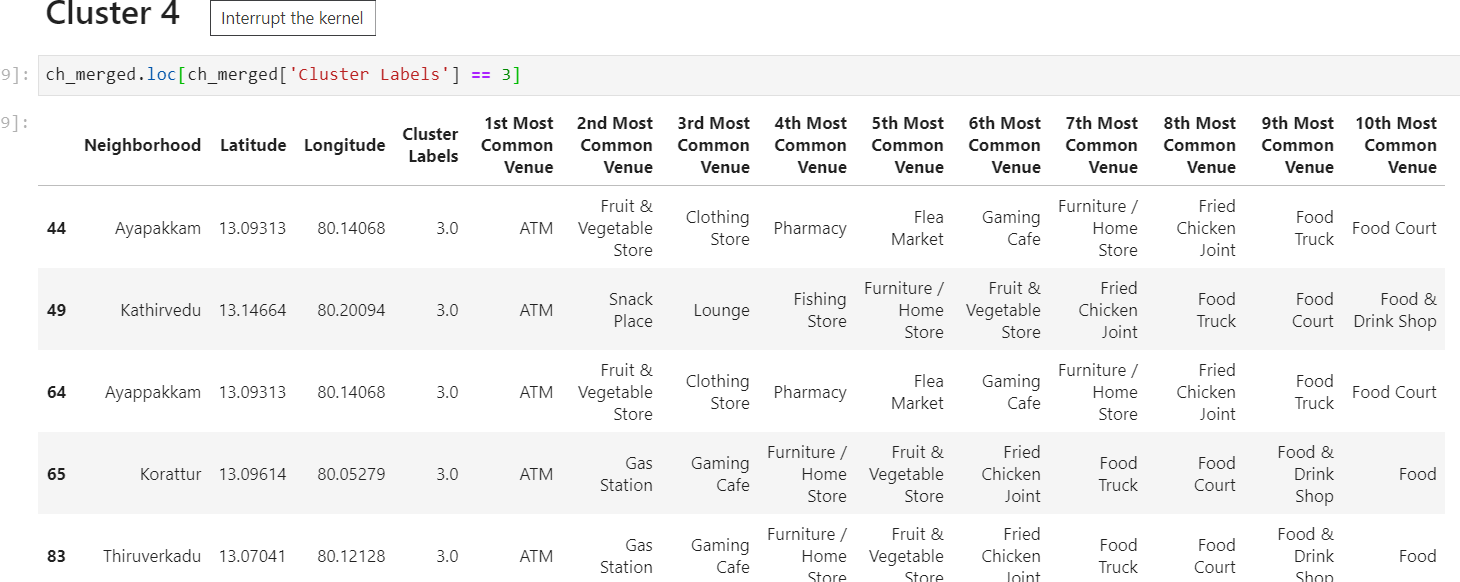
The entire neighbourhood is clustered into 5 clusters using K-Means clustering algorithm

These are the clusters:











1. Discussion and conclusion

From the cluster information, we can conclude that the North west region of the Chennai neighbourhood will find more activities with respect to Infrastructure development, institutions and shopping places.