RESTAURANT RECOMMENDER SYSTEM (For Delhi)

Table of contents

1.Introduction :

Problem background :

Delhi is the capital of India and Delhi score 2nd place in the list of most populated cities in the world. Cuisine here consists of a wide variety of regional and traditional cuisines native to the Indian subcontinent. Given the range of diversity in soil type, climate, culture, ethnic groups, and occupations, these cuisines vary substantially from each other and use locally available spices, herbs, vegetables, and fruits. Indian food is also heavily influenced by religion, in particular Hindu, cultural choices and traditions. The cuisine is also influenced by centuries of Islamic rule, particularly the Mughal rule. Samosas and pilafs can be regarded as examples

Problem description :

Food can also attract people around to world to try it out if it were to be the best. In such scenarios, we need to find the right place, at reasonable cost, to serve us the best possible way. So there are few questions that must be addressed, such as :

How many types of foods are available in the restaurant ?

which is the most nearest restaurants to me with good rating ?

location of restaurant?

How many "similar" restaurants are available near by me ?

Do the "similar" restaurants cost more ? if so, what speciality do that have ?

Expectations from this recommender system is to get answer for the questions, and in such a way that it uncovers all the perspective of managing recommendations. It is sighted to show :

What types of restaurants are present in a paeticular area ?

where are the similar restaurant present based on a preference to particular food ? How do different restaurants rank with respect to my preferences ?

2.Data Section :

DATA REQUIREMENTS :

1. To access location of a restaurant, its Latitude and Longitude is to be known so that we can point at its coordinates and create a map displaying all the restaurants with its labels respectively.(referring lab)

2. Population of a neighbourhood is very important factor in determining a restaurant's growth and amount of customers who turn up to eat. Logically, the more the population of a neighbourhood, the more people will be interested to walk openly into a restaurant and less the population, less number of people frequently visit a restaurant. Also if more people visit, better the restaurant is rated because it is accessed by different people with different taste. Hence it is very important factor.

3. Income of a neighbourhood is also very important factor as population was. Income is directly proportional to richness of a neighbourhood. If people in a neighbourhood earns more than an average income, then it is very much possible that they will spend more however not always true with very less probability. So a restaurant assessment is proportional to income of a neighbourhood.

DATA DESCRIPTION :

For this problem, we will get the services of Foursquare API to explore the data of Delhi, in terms of their neighbourhoods. The data also include the information about the places around each neighbourhood like restaurants, hotels, coffee shops, population of neighbourhoods and income by neighbourhoods. We selected one Borough from each city to analyze their neighbourhoods. We will use machine learning technique, "Clustering" to segment the neighbourhoods with similar objects on the basis of each neighbourhood data. Use of foursquare is focused to fetch nearest venue locations so that we can use them to form a cluster. Foursquare api leverages the power of finding nearest venues in a radius(in my case : 500mts) and also

Methodology :

Data Analysis: Scrapping the data from different sources and then combining it to form a single-ton dataset is a difficult task. To do so, we need to explore the current state of dataset and then list up all the features needed to be fetched. Exploring the dataset is important because it gives you initial insights and may help you to get partial idea of the answers that you are looking to find out from the data.
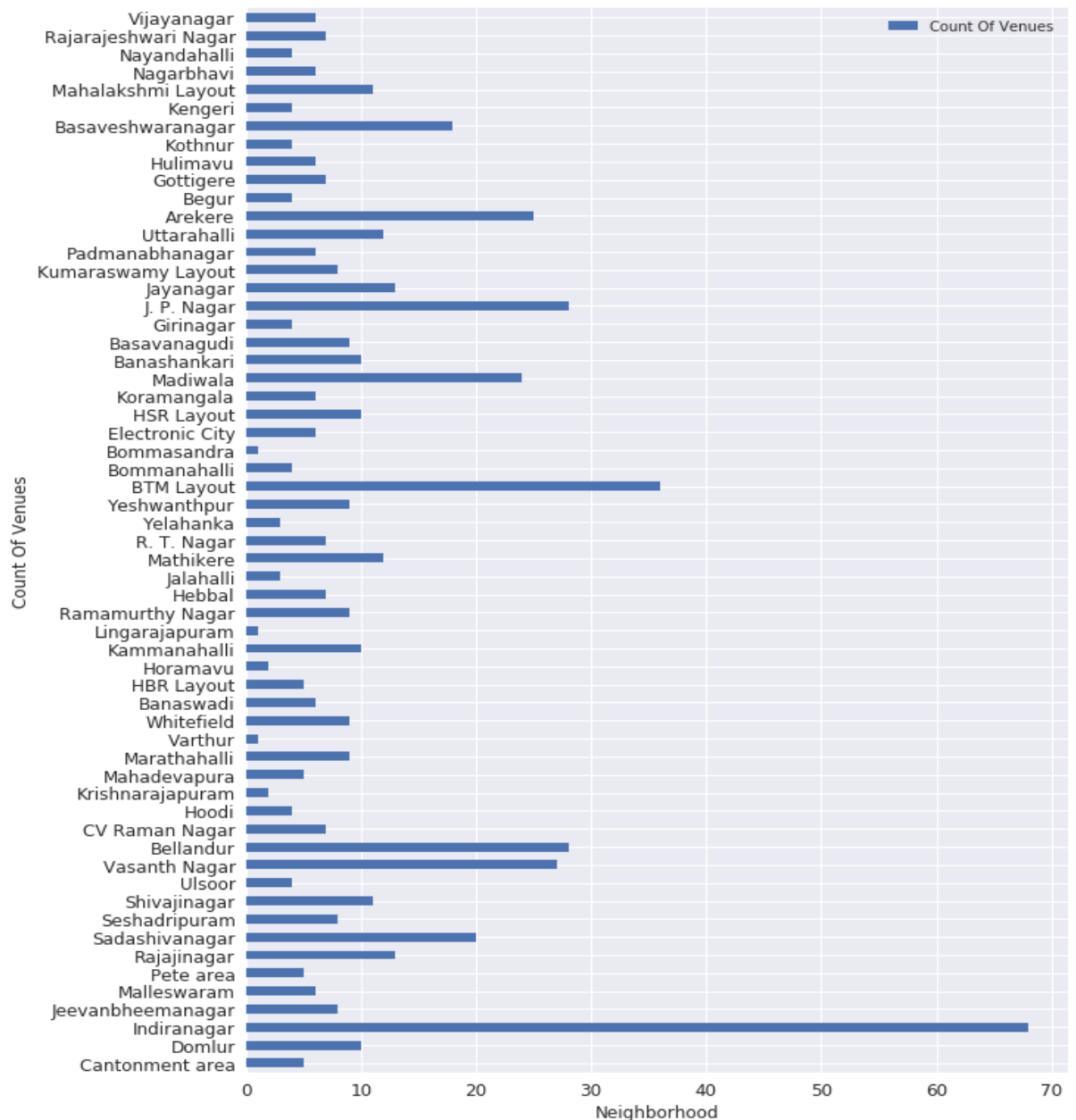
Also while producing graph for number of cluster, I produced a graph to explore all the values for n_clusters and then finding the best by exploring the elbow graph.
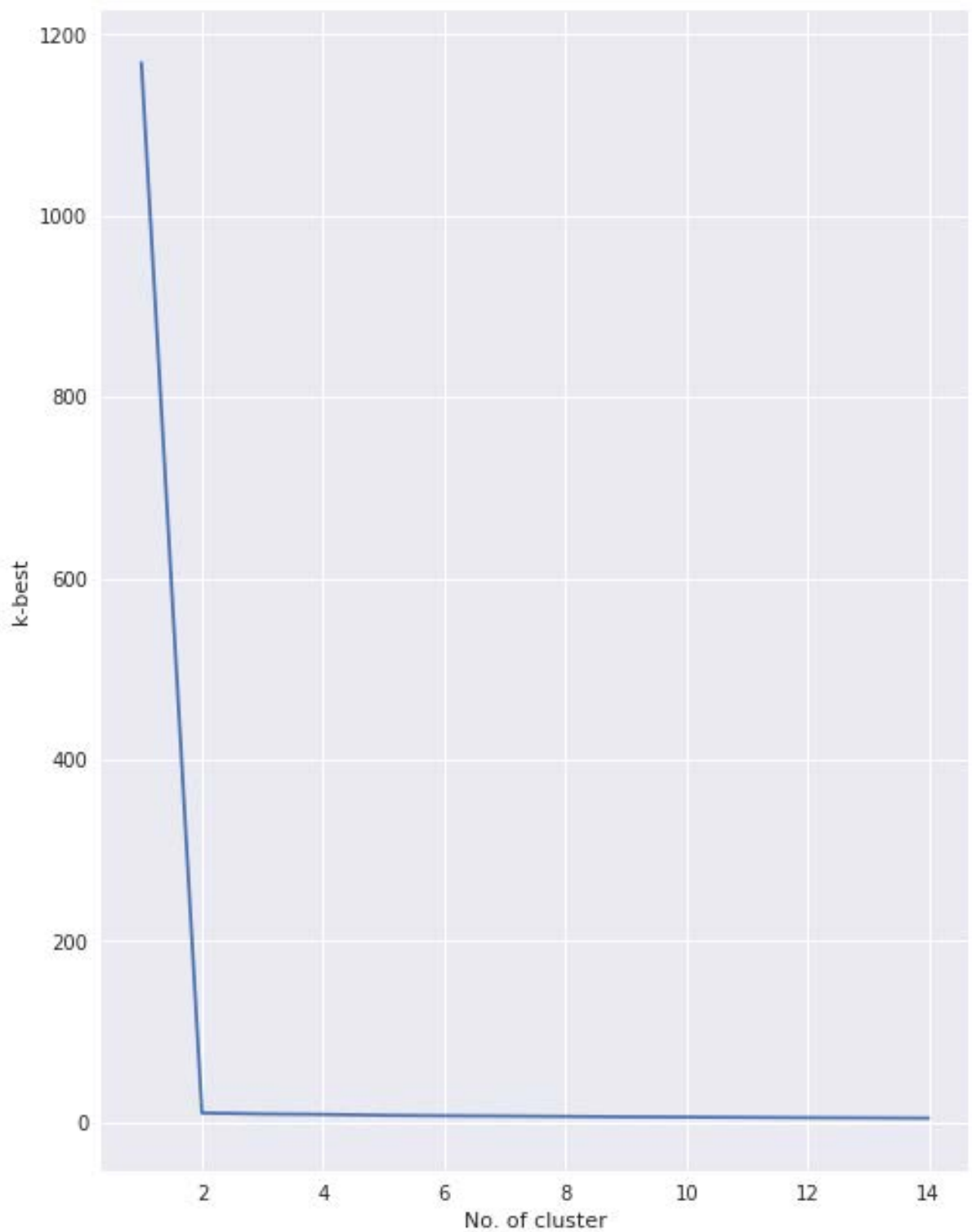
4.Result :

The result of the recommender system is that it produces a list of top restaurants and the most common venue item that the user can enjoy. During the runtime of the model, a simulation was done by taking 'Whitefield' as the neighborhood and then processed through our model so that it could recommend neighborhoods with similar characters as that of 'Whitefield'.

In [81]: top3_df.head(3)

Out[81]:

	Neighborhoods	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	Ranking
0	Begur	Venue Category_Bakery	Venue Category_Indian Sweet Shop	Venue Category_Stadium	[0.7361321887351776]
1	Hulimavu	Venue Category_Bakery	Venue Category_Indian Restaurant	Venue Category_Badminton Court	[0.7638135476902764]
2	Kamakshipalya	Venue Category_South Indian Restaurant	Venue Category_Yoga Studio	Venue Category_Food Truck	[0.80418735993893]

5.Conclusion :

The recommender system is a system that considers factors such as population, income and makes use of Foursquare API to determine nearby venues. It is a powerful data driven model whose efficiency may decrease with more data but accuracy will increase. It will help users to finish their hunger by providing the best recommendation to fulfil all their needs.