1 Introduction :

1.1 Business Problem :

India is known for it's wide variety of cuisine.Each state in India has it's own cuisine, which reflects social and economical diversity of India.Restaurants in each part of India has their own speciality, like if you travel in south states you will find Udupi cuisine, where as if you travel in nothern states like Punjab you will find totally different kinf of cuisine.

Suppose you are a traveler and you visit lot many places then it becomes very difficult to find a right restaurant that suits your taste and which is available at resonable rates.To find solution of this problem we build a recommender system hat can help in recommending new places based on their rankings compared to the previously visited by user.

1.2 Target Audience :

Target audiences for this project does not limit to a person who keeps travelling but everyone. People could simply decide to look for a similar restaurant all the time because they are addicted to a specific category of food. People who rarely use restaurants would prefer to have the most rated restaurants nearby them and all this could be easily handled by our recommender system. So target for this project is basically everyone who is exploring different places or similar places.

2 Data :

2.1 Data Requirement :

To find a solution to the questions and build a recommender model, we need data and lots of data. Data can answer question which are unimaginable and non-answerable by humans because humans do not have the tendency to analyze such large dataset and produce analytics to find a solutions.

Let's consider the base scenario:
Suppose I want to find a restaurant, then logically, I need 3 things:
1. Its geographical coordinates (latitude and longitude) to find out where exactly it is located.
2. Population of the neighborhood where the restaurant is located.
3. Average income of neighborhood to know how much is the restaurant worth.

Let's take a closer look at each of these:
1. To access location of a restaurant, it's 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.
2. Population of a neighborhood 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 neighborhood, 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 is very important factor.
3. Income of a neighborhood is also very important factor as population was. Income is directly proportional to richness of a neighborhood. If people in a neighborhood 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 neighborhood.

2.2 Data Collection :

Instead of gathering data of entire India, this project will focus on data gathered from city of Bangelor which is 3rd largest city in India and a major IT hub.

1. Initially I scrapped list of neighbor's using beautifulSoup4 from `[wikipedia]` `(https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Bang alore).` The table headings becoming the boroughs and data becoming the neighborhoods. Bangalore has 8 boroughs and 64 neighborhoods. So i manually googled each neighborhood to find its corresponding latitude and longitude. After doing so, I produced the following data frame.

Borough	Neighborhoods	Latitude	Longitude
Central	Cantonment area	12.972442	77.580643
Central	Domlur	12.960992	77.638726
Central	Indiranagar	12.971891	77.641151
Central	Jeevanbheemanagar	12.962900	77.659500
Central	Malleswaram	13.003100	77.564300
Central	Pete area	12.962700	77.575800
Central	Rajajinagar	12.990100	77.552500
Central	Sadashivanagar	13.006800	77.581300
Central	Seshadripuram	12.993500	77.578700
Central	Shivajinagar	12.985700	77.605700

2. Population by neighborhood is again easy to find out given that it's readily available. But in case of Bangalore, it is again not the case. i was able to find population data for few cities. [Here is the link](https://indikosh.com/dist/655489/bangalore). Rest other neighborhood population is assumed and may be inaccurate but since this is a demonstrating project, the main idea to get the working model. The data frame for Bangalore neighborhood population looks like:

	Borough	Neighborhoods	Population	Normalized_population
0	Central	Cantonment area	866377	0.880810
1	Central	Domlur	743186	0.755567
2	Central	Indiranagar	474289	0.482190
3	Central	Jeevanbheemanagar	527874	0.536668
4	Central	Malleswaram	893629	0.908516

3. Income by neighborhood is again easy to find out given that it's readily available. But in case of Bangalore, it is again not the case. i was able to find Income data for main city.(https://en.wikipedia.org/wiki/List_of_Indian_cities_by_GDP_per_capita). Neighborhood Income is assumed and may be inaccurate but since this is a demonstrating project, the main idea to get the working model. The data frame for Bangalore neighborhood population looks like:

	Borough	Neighborhoods	AverageIncome	Normalized_income
0	Central	Cantonment area	18944.099792	0.293051
1	Central	Domlur	56837.022198	0.879225
2	Central	Indiranagar	41991.817435	0.649581
3	Central	Jeevanbheemanagar	6667.447632	0.103140
4	Central	Malleswaram	53270.063892	0.824047

4. 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 corresponding coordinates, venue location and names. After calling, the following data frame is created:

|	Neighborhood	Borough	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Cantonment area	Central	12.972442	77.580643	Hotel Fishland	12.975569	77.578592	Seafood Restaurant
1	Cantonment area	Central	12.972442	77.580643	Sapna Book House	12.976355	77.578461	Bookstore
2	Cantonment area	Central	12.972442	77.580643	Vasudev Adigas	12.973707	77.579257	Indian Restaurant
3	Cantonment area	Central	12.972442	77.580643	Adigas Hotel	12.973554	77.579161	Restaurant
4	Cantonment area	Central	12.972442	77.580643	Kamat Yatrinivas	12.975985	77.578125	Indian Restaurant

3 Methodology :

3.1 Exploratory analysis:

Scrapping the data from different sources and then combining it to form a single-ton dataset is a difficult task.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.

3.2 Inferential analysis:

Most important factors while building the recommender system were population and income. They are the most import factor because they have a nonlinear relationship according to our dataset.

It needed to make some inferential analysis to understand this nonlinear relationship. As the amount of population increases, it does not necessarily mean that average income of a neighborhood will also increase. It is true to most of the case but also many cases differ to follow this trend. Similarly, a neighborhood with less number of people may not necessarily have less average income. It is possible to have less number of people and more income and vice versa. This can be inferred from the following graph:
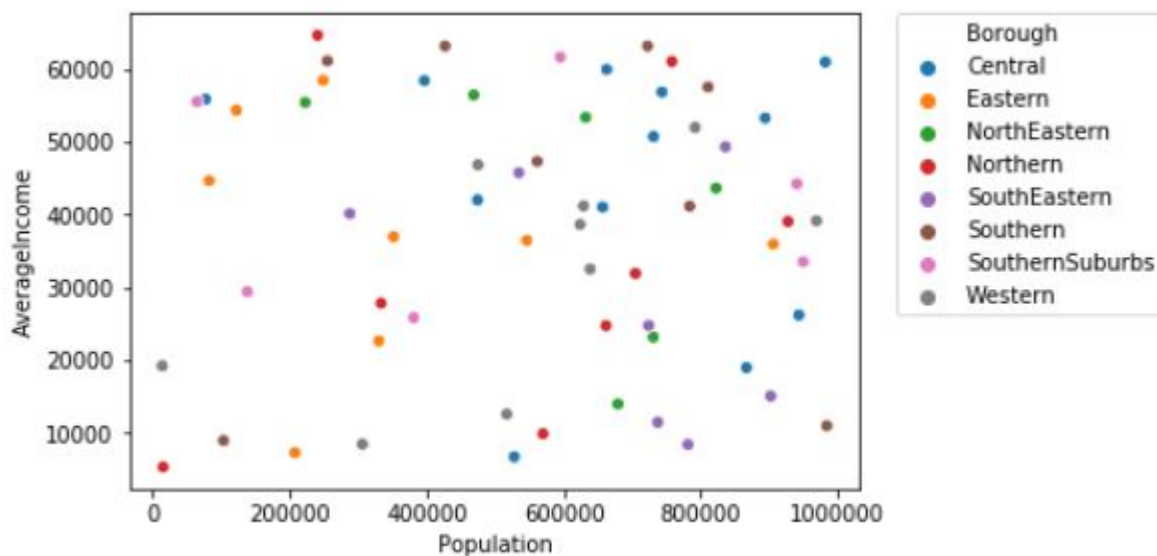
4 Result :

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'.
The following image shows the result:

	Neighborhoods	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	Ranking
0	Basaveshwaranagar	Venue Category_Ice Cream Shop	Venue Category_Indian Restaurant	Venue Category_Fast Food Restaurant	[0.6426377807870477]
1	Begur	Venue Category_Indian Restaurant	Venue Category_Indian Sweet Shop	Venue Category_Food Court	[0.7361321887351776]
2	Electronic City	Venue Category_Outlet Store	Venue Category_Furniture / Home Store	Venue Category_Bus Stop	[0.5423513638809381]