**Introduction :**

**Problem background:**

Bangalore is the capital and largest city of the Indian state of Karnataka. With a population of over 15 million (as of January 2016), Bangalore is the third largest city in India and 27th largest city in the world. It’s a metropolitan city where a group of people with diversity live in the city.

**Problem description:**

Food can attract people around the world to try it out if it is 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:

1. How many types of foods are available in the restaurant?

2. Which is the nearest to me with good rating?

3. Do the "similar" restaurants near me cost more? If so, what specialty 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:

1. What types of restaurants are present in a particular area?

2. Where are the similar restaurant present based on a preference to particular food?

3. How do different restaurants rank with respect to my preferences?

**Target Audience:**

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.

**Success rate:**

Computers are good at remembering things, and with Machine learning to its peak, it high time technology will by our personal guidance and help us personally based on our likes and dislikes. So, people would care about this project as their personal assistance and success rate could certainly increase with time.

**Data :**

**Data requirements:**

To find a solution to the questions and build a recommender model, we need data and lots of data.

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.

**Data collection:**

I decided to use Google API to fetch latitude and longitude but Google API has limited number of calls that I could make with my free account.

Initially I scrapped list of neighbor's using beautifulSoup4 from <https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Bangalore>. The table headings becoming the boroughs and data becoming the neighborhoods. Bangalore has 8 boroughs and 64 neighborhoods.



Population by neighborhood is again easy to find out given that it’s readily available at <https://indikosh.com/dist/655489/bangalore>



Income by neighborhood is again easy to find out given that it’s readily available at <https://en.wikipedia.org/wiki/List_of_Indian_cities_by_GDP_per_capita.>



Foursquare API:

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:



1. **Methodology :**

**Exploratory analysis:**

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.

While exploring the dataset, I found out that Inderanagar has most number of venues while Varthur has the least.



**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.



**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:



**Discussion :**

Since there was a nonlinear relationship between income and population, it can be concluded that we must always perform inferential approach to find relationship among different set of features. Also during clustering, similar neighborhoods must be dumped into the right cluster.

The following graph shows the clusters:



**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.