Data :

Data requirements: 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.

Data collection:

1. Collecting geographical coordinates is not difficult but after googling for more than 2 days, it was not available on open source data websites such as Wikipedia, India gov website, census report websites etc. So I decided to use Google maps API to fetch latitude and longitude but google API has limited number of calls that I could make with my free account. So it would take around 15 - 20 days to fetch location of all the neighborhoods in Bangalore. Initially I scrapped list of neighbor's using beautifulSoup4 from [wikipedia](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. So i manually googled each neighborhood to find its corresponding latitude and longitude. After doing so, I produced the following data frame.

	Borough	Neighborhood	Latitude	Longitude
0	Scarborough	Malvern / Rouge	43.806686	-79.194353
1	Scarborough	Rouge Hill / Port Union / Highland Creek	43.784535	-79.160497
2	Scarborough	Guildwood / Morningside / West Hill	43.763573	-79.188711
3	Scarborough	Woburn	43.770992	-79.216917
4	Scarborough	Cedarbrae	43.773136	-79.239476
5	Scarborough	Scarborough Village	43.744734	-79.239476
6	Scarborough	Kennedy Park / Ionview / East Birchmount Park	43.727929	-79.262029
7	Scarborough	Golden Mile / Clairlea / Oakridge	43.711112	-79.284577
8	Scarborough	Cliffside / Cliffcrest / Scarborough Village West	43.716316	-79.239476
9	Scarborough	Birch Cliff / Cliffside West	43.692657	-79.264848

2. 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:

	Neighborhood	Borough	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Malvern / Rouge	Scarborough	43.806686	-79.194353	Wendy's	43.807448	-79.199056	Fast Food Restaurant
1	Rouge Hill / Port Union / Highland Creek	Scarborough	43.784535	-79.160497	RIGHT WAY TO GOLF	43.785177	-79.161108	Golf Course
2	Rouge Hill / Port Union / Highland Creek	Scarborough	43.784535	-79.160497	Royal Canadian Legion	43.782533	-79.163085	Bar
3	Guildwood / Morningside / West Hill	Scarborough	43.763573	-79.188711	G & G Electronics	43.765309	-79.191537	Electronics Store
4	Guildwood / Morningside / West Hill	Scarborough	43.763573	-79.188711	Big Bite Burrito	43.766299	-79.190720	Mexican Restaurant