<div align="center">**The battle of Neighborhoods**</div>

Week 4:

Part b) –

Title: Restaurant Recommender System in India the Place Chennai

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 analyses 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, 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.

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 visits a restaurant. Also if more people visit, better the restaurant is rated because it is accessed by different people with different taste. Hence is 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:

Collecting geographical coordinates is not difficult but after googling for more than 5 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 10 - 20 days to fetch location of all the neighborhoods in Chennai. Initially i scrapped list of neighbor's using beautifulSoup4 from wikipedia. The table headings becoming the boroughs and data becoming the neighborhoods. Chennai has 8 boroughs and 175 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
Adambakkam	South and East Chennai	12.98800	80.20470
Adyar	South and East Chennai	13.00120	80.25650
Alandur	South and East Chennai	12.99750	80.20060
Alapakkam	West Chennai	13.04900	80.16730
Alwarthirunagar	West Chennai	13.04260	80.18400
Ambattur	West Chennai	13.11430	80.15480
Aminjikarai	West Chennai	13.06980	80.22450
Anna Nagar	West Chennai	13.08500	80.21010
Annanur	West Chennai	13.11840	80.12460
Arumbakkam	West Chennai	13.07240	80.21020
Ashok Nagar	West Chennai	13.03730	80.21230

Population by neighborhood is again easy to find out given that it's readily available. But in case of Chennai, it is again not the case. i was able to find population data for few cities. Here is the link. 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 Chennai neighborhood population looks like:

Unnamed: 0	Borough	Neighborhoods	Latitude	Longitude	Population	City	AverageIncome
0	Adambakkam	South and East Chennai	12.98800	80.20470	82848	Chennai	18944.099792
1	Adyar	South and East Chennai	13.00120	80.25650	147143	Chennai	56837.022198
2	Alandur	South and East Chennai	12.99750	80.20060	47983	Chennai	41991.817435
3	Alapakkam	West Chennai	13.04900	80.16730	527874	Chennai	6667.447632
4	Alwarthirunagar	West Chennai	13.04260	80.18400	893629	Chennai	53270.063892
5	Ambattur	West Chennai	13.11430	80.15480	730999	Chennai	50712.430215

Income by neighborhood is again easy to find out given that its readily available. But in case of Chennai, it is again not the case. i was able to find Income data for main city. Here is the link. 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 Chennai neighborhood population looks like:

Unnamed: 0	Borough	Neighborhoods	Latitude	Longitude	Population	City	AverageIncome
0	Adambakkam	South and East Chennai	12.98800	80.20470	82848	Chennai	18944.099792
1	Adyar	South and East Chennai	13.00120	80.25650	147143	Chennai	56837.022198
2	Alandur	South and East Chennai	12.99750	80.20060	47983	Chennai	41991.817435
3	Alapakkam	West Chennai	13.04900	80.16730	527874	Chennai	6667.447632
4	Alwarthirunagar	West Chennai	13.04260	80.18400	893629	Chennai	53270.063892
5	Ambattur	West Chennai	13.11430	80.15480	730999	Chennai	50712.430215

1. 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	South and East Chennai	Adambakkam	12.9880	80.2047	Venkateshwara Super Market	12.986320	80.205168	Department Store
1	South and East Chennai	Adambakkam	12.9880	80.2047	Ibaco	12.988729	80.205646	Dessert Shop
2	South and East Chennai	Adambakkam	12.9880	80.2047	Deepam Restaurant	12.985380	80.205281	Indian Restaurant
3	South and East Chennai	Adambakkam	12.9880	80.2047	ibaco Adambakkam	12.987358	80.200504	Ice Cream Shop
4	South and East Chennai	Adyar	13.0012	80.2565	Zha Cafe	12.999730	80.254806	Café

The following map is produced by marking all the neighborhoods in Chennai city.