Neighbourhood Analysis – Chennai, South India

1. Introduction & Business Problem:

1.1 Background:

This report deals with analysis of neighbourhoods of **Chennai, South India**. Chennai also known as Madras is the capital city of Indian State of **TamilNadu**. Located on the Coromandel Coast of the Bay of Bengal, it is one of the largest cultural, economic and educational centres of south India. According to the 2011 Indian census, it is the sixth-most populous city and fourth-most populous urban agglomeration in India. The city together with the adjoining regions constitutes the Chennai Metropolitan Area, which is the **36th-largest urban area by population** in the world.

Chennai is divided into four broad regions: North, Central, South, and West. North Chennai is primarily an industrial area. South Chennai and West Chennai, previously mostly residential, are fast becoming commercial, home to a growing number of information technology firms, financial companies and call centres. The city is expanding quickly along the Old Mahabalipuram Road and the Grand Southern Trunk Road (GST Road) in the south and towards Ambattur, Koyambedu and Sriperumpudur in the west.

1.2 Problem Statement:

Having said this, I want to analyse the neighbourhoods of Chennai for its **shopping malls**. Chennai hosts a large number of shopping malls across the city. This report provides a detail analysis of existing shopping malls in the city. **The main purpose of this project is intended to find out few prime neighbourhoods in Chennai that will be appealing for an entrepreneur to open a new shopping mall.**

1.3 Who will be interested? :

Deciding where to put your business is as important as the business you decide to go. Intended audience of this report would be **large scale entrepreneurs** who wish to know about prime locations in Chennai city to open their **new** business specifically a shopping mall. This report also gives them basic understanding about the existing malls.

Note: This project can easily be extended for other businesses such as restaurants or electronic shops with little modifications.

2. Data:

2.1 Data Source:

1. **Locations** of **existing shopping malls** are obtained using **FoursquareAPI**.
2. Premium details such as rating and reviews for each shopping mall are gathered using **FoursquareAPI premium services**.
3. Population distribution of Chennai :
   Latest census data for Chennai is available for the year 2011. Ward wise population (A city is divided into wards for administrative purposes – similar to boroughs) for Chennai is obtained from District Census Hand Book (DCHB) from census website of India
4. Choropleth map will be a good choice to visualize the population distribution. To plot choropleth we need to get shape/geojson file for Chennai. Thanks to https://github.com/mickeykedia/India-Maps for the the geojson, shape and allied files for Chennai city!
5. Details of prime localities and popular IT parks are collected from Wikipedia. Localities include major IT parks and large residential localities.

2.2 Data cleaning:

- Foursquare API returns lot of details. Extract only the name, geographical coordinates and id of the shopping malls.
- From Foursquare premium call, extract only rating and reviews for the malls.
- Population Data:

Challenge here was, during 2011 when census was taken, Chennai city had 155 wards. Soon afterwards it was extended to have 200 wards. Population data for the newly included wards were collected from DCHB of Kancheepuram. Population data along with ward numbers are extracted to 'Chennai_wardwise_population_2011.csv'.

2.3 How this data will be used to solve the problem? :

- We first get the locations of existing shopping malls using FoursquareAPI. Here we search for shopping malls in Chennai city as a whole.

    - Use **'Search' endpoint**
    - Latitude and Longitude values is obtained using **geocoder** library for the location 'Chennai'
    - Use **CategoryId** parameter to pass the category ID of 'Shopping Mall' which is '4bf58dd8d48988d1fd941735'. (Get the category id from foursquare documentation https://developer.foursquare.com/docs/build-with-foursquare/categories/)
    - Use **'query'** parameter to list only the category **'Shopping Mall'**
    - We will not pass radius so that 'city-wide' search is performed
    - We will set limit to the maximum value i.e 1000 to list all the existing malls

  API will look like:
  https://api.foursquare.com/v2/venues/search?&client_id={}&client_secret={}&v={}&ll={},{}&categoryId={}&query={}&limit={}
  From the resulting JSON extract name, venue id, latitude and longitude for all the returned venues.

- To fetch ratings and reviews from Foursquare API use its premium call for vanues with venue id being endpoint:
  https://api.foursquare.com/v2/venues/{}?&client_id={}&client_secret={}&v={}
  Here venue id from previous result is passed to get the ratings for individual venues. Since this is a premium call, we will be able to execute this API only once per venue per day. From the resulting JSON extract ratings and tips.

- **KMeans** clustering machine learning model is used to group the venues. Feature set includes ratings and location of the venues. On plotting the data using folium library we will get a picture of locations of existing malls and groupings of similar kind of malls. At this point stakeholders will get an idea about the existing business situation in the city.

- Success of a shopping mall highly depends on the customer traffic. Two criterions are taken into account to solve our problem. Fix areas:

    - ✓ Based on the **population distribution**:
      Locations with high population are better places to get more customers
    - ✓ Based on **purchasing power of the localities**:
      Localities having big residential properties, large IT parks, prominent companies attracts affluent customers.

- Ward wise population of Chennai city is visualized using choropleth map. Now we can get an idea about densely populated localities
- The prime localities where potential customers are expected are also plotted in the choropleth map along with the existing malls locations.
- Now coming to the final part, find out the areas which
    a) Are densely populated
    b) Enclose large number of localities with potential customers
    c) Have no shopping malls within vicinity of say 10 kms.
- Areas which satisfy all the three criteria will be recommended as a target location to start a new shopping mall.