IBM Applied Data Science Capstone

Chennai Metro Stations Clustering

Chennai, Tamil Nadu, India

By: Sanjay Kannu Periyasami

November 2020

CHENNAI METRO RAIL LIMITED

Introduction

Chennai is the capital of the Indian state of **Tamil Nadu**. The Quality of Living Survey rated **Chennai** as the safest **city** in India. As per the estimates of UN World Population Prospects, the Chennai Population in 2020 is **10.9 Million** (1.09 Crore). Chennai is fifth largest urban agglomeration city in India and 30th largest city in the World

The Chennai Metro Rail Limited (CMRL) started Chennai Metro in 2015 to handle its traffic problem. Chennai Metro is now lifeline of Chennai and its working population is depending on Chennai Metro.

Foursquare.com has very rich database of venues around the world it also provides tips and photos in detail. Foursquare provide APIs to developers to explore these data. We will use Foursquare APIs in this work to get Chennai Metro information like name/latitude/longitude and different venue details around each Chennai Metro stations.

Business Problem:

Due to high volume of traffic on road and very densely populated area, people often spend lot of time in traffic and avoiding traffic is main agenda for people. Staying near Chennai Metro is most convenient, especially in early morning while going to Office and School and other times when traffic on road is high.

A residential location near Chennai Metro is most wanted. We have to find out which metro station is best suited for residential purpose.

To decide best residential area, we have considered five categories listed below.

1. Food & Drink Shops
2. School
3. Medical Centre
4. Gym
5. Restaurants
6. Shopping Mall

In this project we must identify which metro station is best considering above 6 conditions and categorize them and plot them on map to visual identification.

Target Audience of this project

This project is particularly useful for expatriates from other places of India looking for the residence in the neighborhoods of Chennai.

Data Section

To solve this problem, we will use Foursquare APIs and collect data of Chennai Metro Stations. Find out Chennai Metro Station Details including Name, Latitude, and Longitude. Using search endpoint with category id of “Metro Station” we will get all metro station details. We will convert JSON into dataframe. Once dataframe is received we will remove unwanted columns except Name, Latitude and Longitude.

Gathering all the Metro Station information, we have total 34 rows.
For each metro station we will find venue details of each category listed below with search endpoint using category id from Foursquare.

S. No.	Category	Category Id
1	Food & Drink Shops	4bf58dd8d48988d1f9941735
2	Schools	4bf58dd8d48988d13b941735
3	Medical Center	4bf58dd8d48988d104941735
4	Fitness Center	4bf58dd8d48988d175941735
5	Restaurants	4d4b7105d754a06374d81259
6	Shopping Mall	4bf58dd8d48988d1fd941735

Now we have metro station details and venue for each metro station in radius of 500 meters. We will group venue detail by Name of Metro station and Venue Category to get number of venues for each category at each metro station.