COURSERA-IBM DATA SCIENCE SPECIALIZATION

Capstone Project - Analysing Mumbai Neighborhoods

Introduction:

The City of Mumbai, also known as the City of Dreams is the financial capital of India. In recent times, the price of real estate in the city has sky-rocketed while income equality has plummeted. This has made house hunting a significantly greater challenge than it already was. In this project therefore, we try to ease this process by comparing the different neighborhood of Mumbai on the basis of a variety of factors.

Problem Statement: Exploring and Clustering the neighborhoods of Mumbai city on the basis of several factors in order to find similarity among them thus helping anyone looking to find homes.

Target Audience: Anyone out in the market looking for homes can be aided through this project. It can also help real estate developers and investors to find facilities lacking in the area which can be improved upon so as to make them more attractive to potential customers.

Data:

For a list of neighborhoods along with their rental prices we scrape data from the following website:

[https://www.makaan.com/price-trends/property-rates-for-rent-in-mumbai]

We use the column 'Avg Rent' for 2BHK flats as a feature to build our dataset. The coordinates of the different neighborhoods are found using the Nominatim library in the code itself.

Next, we use two different APIs, namely Foursquare API and HERE API to find out different venues and facilities available near each locality. We will use them to find out the most popular venues in the neighborhood. The category search available in both these APIs will be used to find the essential ameneties while looking for houses, such as hospitals, schools and colleges, shopping services etc.

All the data collected would then have to be cleaned, standardized and one-hot encoded before we can use them for clustering. We would use the Folium library several times in our code to visualize the different locations using maps.