**COURSERA-IBM DATA SCIENCE SPECIALIZATION**

**Capstone Project by Hiral K**

**Mumbai Neighborhood Analysis**

**Introduction**

Mumbai is one of the metro cities of India. It is also known as the financial capital of India. Mumbai is one of the most expensive city of the world in terms of real estate both residential and commercial. In addition to the real estate prices, the Mumbai is also famous for its heavy traffic conditions. Travelling in Mumbai from one location to another is difficult task.

This project explores different localities of Mumbai between Andheri to Dahisar locations for availability of restaurants and buy rates of residential properties. This project can help someone who is looking to open restaurant in locality between Andheri to Dahisar based on the number of restaurants in particular locality. Here I have assumed that person also wants to purchase a house in the same or nearby locality to avoid travelling considering the heavy traffic situations in Mumbai.

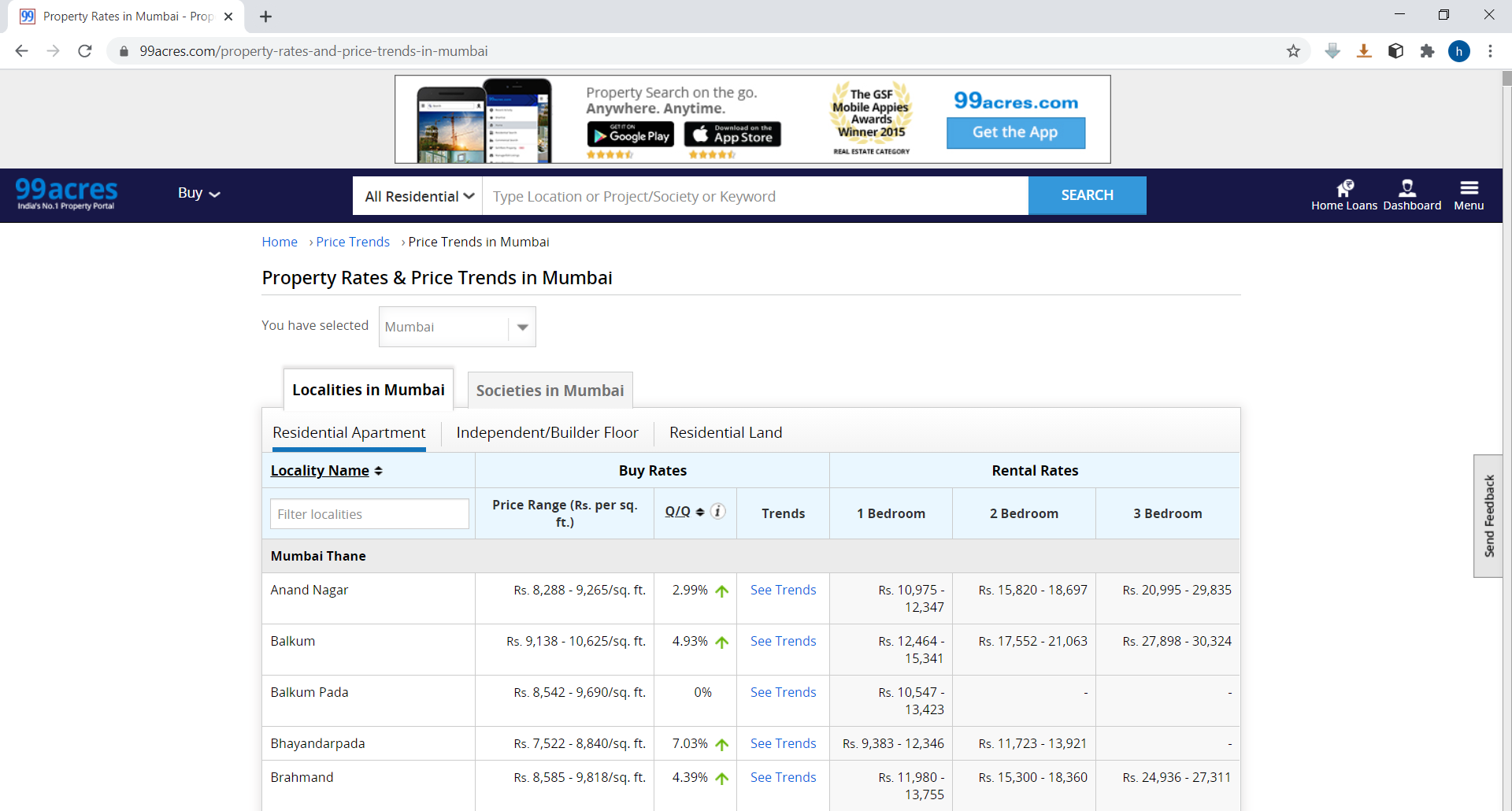
**Problem Statement:** Exploring the localities of the Mumbai from Andheri to Dahisar locations to find suitable locality to open a restaurant. Also exploring buy rates of residential properties in the same localities with the assumption that person wants to purchase house in the same locality where he/she opens a restaurant. Clustering the localities for better observation of number of restaurants and buy rates of houses.

**Target Audience:** Anyone who is looking to open restaurant and/or purchasing house between Andheri to Dahisar locations in Mumbai city. The project can also help real estate developers for new constructions and giving offers to buyers. Real estate agents can also use the system to suggest particular localities to the customers with the same need.

**Data:**

The main data required to build the system is details of localities between Andheri to Dahisar in Mumbai city along with their average buy rates. This data is scraped from 99acres.com website. The URL for the same is:

https://www.99acres.com/property-rates-and-price-trends-in-mumbai



As shown above, the table on the above website contains many features. We use name of localities and buy rates. The table is read into panda’s dataframe using read\_html. Then dataframe is processed to get required features (localities and average buy rates) in required format.

The coordinates of the different localities are found using the Nominatim library. Nominatim library is an Open-source geocoding technology which can find the coordinates of a place using the its address. Then we have used Foursquare API to find out different venues in each locality.