**MUMBAI REAL ESTATE PRICING AND CORRELATION WITH NEARBY FACILITIES**

Capstone Project- Business Understanding & Data Description Phase

*By*

**ARJUN C.**

[iamarjunchandra@gmail.com](mailto:iamarjunchandra@gmail.com)

**IBM – COURSERA**

**DATA SCIENCE PROFESSIONAL CERTIFICATE SPECIALIZATION**

**JANUARY 2019**

**BUSSINESS UNDERSTANDING**

Mumbai is one of the largest metropolitan cities of India and also the economical capital of India. The price of properties in Mumbai is sky rocket high and the price of some portion of Mumbai is par with that of the rate of properties in many U.S cities. In this capstone project for the ‘IBM Data Science Professional Certificate’ specialization course, I am building a clustering and segmentation model that will sort the various places in Mumbai based on its real estate price and amenities and venues available nearby.

Do the nearby venues and amenities available near a place affect the pricing of that place? What are the posh places in Mumbai? What are the facilities and venues nearby to each place? These are the main questions that will be addressed in this project. By the end of this project, answers for all these questions will be derived.

This project will be useful for a number of people for a number of purposes. Some of the main beneficiaries of this project are:

* People who are interested to buy a property in Mumbai.

In this project, the various places in Mumbai will be classified in to various categories depending on their average price per sq.ft. So people can target respective places depending on their budget.

* Companies looking for a perfect place to open their outlet/office.

This project will be useful for those entrepreneurs who are searching for an ideal place in Mumbai as in this project all nearby attractions of different places in Mumbai will be determined. So companies can choose places which have nearby attractions that can enhance their operations.

**DATA** **DESCRIPTION**

For the project, Data regarding various places in Mumbai is necessary. A real estate website called ‘Makaan’ has a tabular webpage containing details regarding all major places in Mumbai and their real estate price.

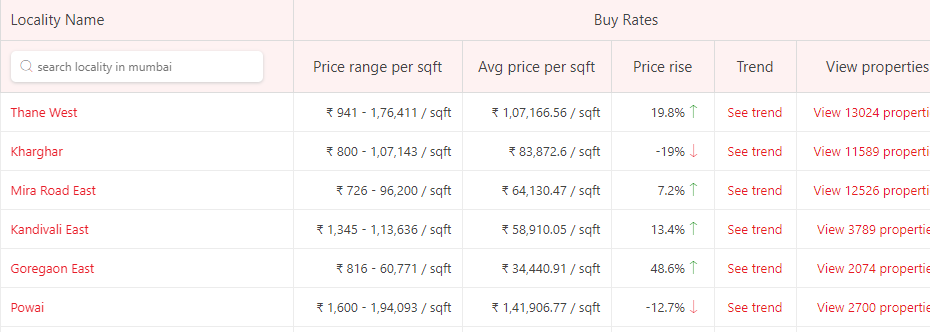


Fig.1: Web page containing necessary data

From the web page shown above, details regarding various places in Mumbai and their average price per sqft. are extracted using HTML parsing techniques. The extracted data frame’s sample is provided below.

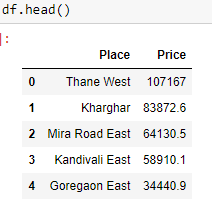


Fig 2. Data Frame’s first 5 rows

The latitude and longitude of each place will be extracted using geopy API, then the nearby attractions of each place will be determined with the help of foursquare API. Segmentation and clustering algorithm will be modeled to group places according to their real estate value.

**THANK YOU**