

**Pramit Samanta**

## **IBM Applied Data Science Capstone**

### ***City Of Opportunity: Mumbai***

#### **Introduction:**

Mumbai is the commercial capital of India. It is also known as the city that never sleeps. Mumbai is the perfect blend of culture, customs, and lifestyles. Mumbai is India's most cosmopolitan city, its financial powerhouse, and the nerve center of India's fashion industry. Mumbai is also dotted with plenty of architectural landmarks from the Victorian era and the days of Raj. Mumbai is also the birthplace of Indian Cinema.

Located on Maharashtra's coast, Mumbai is India's most-populous city, and it is one of the largest and most densely populated urban areas in the world. Mumbai developed a highly diversified infrastructure.

It suffers, however, from some of the perennial problems of many large expanding industrial cities: air and water pollution, widespread areas of substandard housing, and overcrowding. With its diverse society, comes diverse infrastructure which decides the quality of living. There are many infrastructures in Mumbai, each belonging to different categories like Drinking Water Plant, Waste Water/ Sewage, Hospitals, Schools, Colleges, Railway Network, Electricity Power Plants, Telecommunication Support, Bank, Shopping malls, Supermarket, Gas Station, Hotels, Police Station, Café, medical shops, grocery shops, theatre, etc. One of the main problems, when one moves to a new city, is where to find a good area to build and grow prosperously.

## **Business Problem:**

The questions I aim to answer in this project are the following:

- 
- 1. List and visualize all major parts of Mumbai City with top existing infrastructure.*
  - 2. What are the best locations in Mumbai as per infrastructure?*
  - 3. Which areas have the potential for the development of infrastructure of different kinds?*
  - 4. Which all areas lack the infrastructure facilities?*
  - 5. What is the best place to stay within a city for all vital infrastructure facilities?*
- 

## **Target Audience:**

The purpose of this project is to help people in exploring better facilities around their neighborhood. It will help people making a smart and efficient decision on selecting great neighborhoods out of numbers of other postal areas in Mumbai, India.

Lots of people are migrating from various states of India and needed lots of research for good housing prices, new business, and reputed professional places for their children. This project is for those people who are looking for better neighborhoods and businesses.

It will help people to get the awareness of the area and neighborhood before moving to a new city, state, country, or place for their work or to start a new fresh life.

## Data Description:

Mumbai City's demographics show that it is a large and ethnically diverse metropolis. With its diverse society, comes diverse infrastructure. There are many different kinds of infrastructure in Mumbai City, each belonging to different categories like Hospitals, Schools, Colleges, Hotels, etc.

For this project we need the following data:

- Mumbai Pincode ( Scraped from web source)
  - Data source: <https://mumbai7.com/postal-codes-in-mumbai/>
  - Description: Contain a list of pin codes, postal office names, city which can be used to discover all postal office of Mumbai.
- Mumbai City data contain list pin codes, postal office names, city along with their latitude and longitude.
  - Data Source :
  - Description: This data set contains the required information. And we will use this data set to explore various neighborhoods of Mumbai City.
- Different kinds of infrastructures in each neighborhood of Mumbai City.
  - Data source: Foursquare API
  - Description: By using this API we will get all the venues in each postal office. We can filter these venues to get different infrastructures and venues.
- GeoSpace data
  - Data source : <https://github.com/geospace-code/pymap3d>
  - Description: By using this geospace data, we will get the latitude and longitude coordinates of the postal office of Mumbai.

Using this data will allow exploration and examination to answer the questions. This is a project that will make use of many data science skills, from web scraping (mumbai7.com), working with API (Foursquare), data cleaning, data wrangling and map visualization (Folium) and to machine learning (K-means clustering).