# IBM Data Science Capstone Project

## Exploring Infrastructure and Neighborhood in Mumbai

January 16 2021

## Introduction:

Mumbai, traditionally known as Bombay, is the commercial capital of India [1]. It is the second-most populous city in India. It has a highly diversified infrastructure. It is also known as the center attraction of Bollywood cinema.

Despite various infrastructure, Mumbai suffers from air and water pollution, widespread areas of substandard housing, and overcrowding [2]. It is always difficult to identify a prosperous area in Mumbai.

This project will help any newcomers in exploring Mumbai City and its neighborhood. It will help people to make a decision for finding the best infrastructure place to stay in the city.

Every year's millions of people migrate from different parts of India to Mumbai for finding new opportunities and a better life [3]. It will help individuals to acquaint themselves with the city and neighborhood.

## Problem Statement:

In this project I would like to answer the following questions.

1. Visualize the Mumbai
2. Identify the best infrastructure locations in Mumbai
3. Identify the potential areas for future development in Mumbai
4. Identify the city locations with poor infrastructure in Mumbai
5. Identify the best place to live within Mumbai

## Data

I will use following data of the Mumbai.

- Pin code : https://mumbai7.com/postal-codes-in-mumbai/. It has city, postal office, and pin code information's. Latitude and longitude information are added using google map manually.

- Different kinds of infrastructures in each neighborhood of Mumbai City: Foursquare API

## Method

I will use Foursquare API to get all venues in each postal office in Mumbai. Following are the main steps to address our problem.

1) Data cleaning
2) Data wrangling and map visualization
3) K-means clustering to identify the neighborhood

# Results

Figure 1 visualizes the available infrastructure in Mumbai.

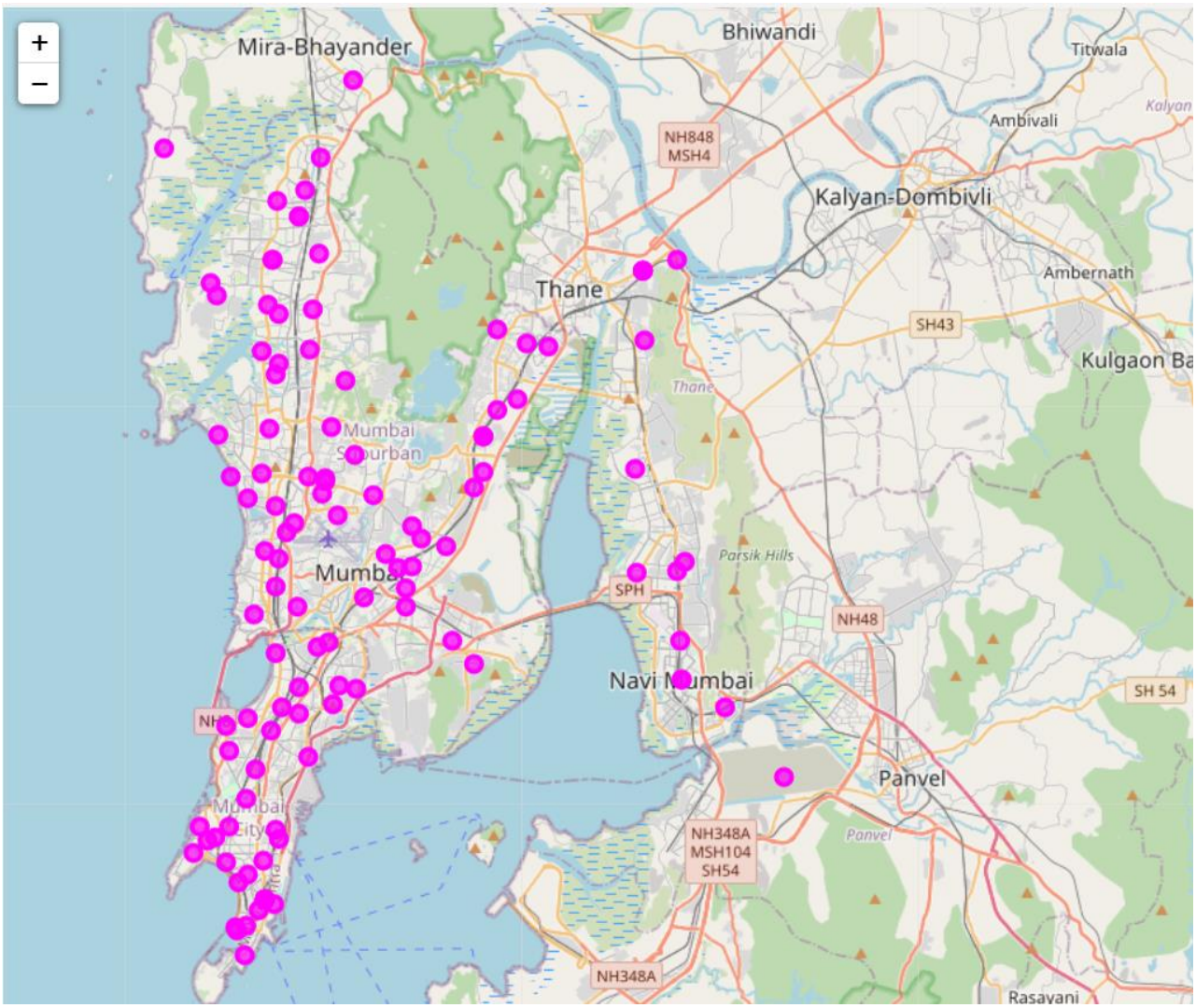

Figure 1. Location of infrastructure in Mumbai

Table 1 provides best infrastructure information's within Mumbai.

Table 1. Best infrastructure details in Mumbai

| Post Office | Bandra (West) |
|---|---|
| Pin Code | 400050 |
| City | Mumbai |
| Airport Terminal | 0 |
| Bank | 0 |
| Bus Station | 0 |
| Business Service | 0 |
| Café | 10 |
| College Auditorium | 1 |
| Electronics Store | 1 |
| Farmers Market | 1 |
| Garden | 0 |
| Government Building | 0 |
| Gym / Fitness Center | 3 |
| Hotel | 1 |
| Indie Movie Theater | 1 |
| Light Rail Station | 0 |
| Market | 0 |
| Monument / Landmark | 0 |
| Park | 1 |
| Pharmacy | 0 |
| Playground | 0 |
| Resort | 0 |
| Restaurant | 1 |
| Shopping Mall | 1 |
| Theater | 0 |
| Train Station | 0 |
| Total infrastructure | 21 |

Table 2 provides the place information that have poor infrastructure within Mumbai.

Table 2. Poor infrastructure locations in Mumbai

| Post Office | Pin Code | City |
|---|---|---|
| Agashi | 401301 | Thane |
| Anu Shakti Nagar | 400094 | Mumbai |
| Bassien | 401201 | Thane |
| Bhandup (East) | 400042 | Mumbai |
| Bhayander (East) | 401105 | Thane |
| Boisar | 401501 | Thane |
| Ghansoli | 400701 | Navi Mumbai |
| Jacob Circle | 400011 | Mumbai |
| Jakegram | 400606 | Thane |
| Jawhar | 401603 | Thane |
| Kopri Colony | 400603 | Thane |
| Krishi Utpanna Bazar | 400705 | Navi Mumbai |
| Mahim | 400016 | Mumbai |
| Nerul Mode | 400706 | Navi Mumbai |
| Santacruz P&T Colony | 400029 | Mumbai |
| Sopara | 401203 | Thane |
| Tagore Nagar | 400083 | Mumbai |
| Talasari | 401606 | Thane |
| Umbarpada | 401102 | Thane |
| Uran | 400702 | Navi Mumbai |
| Vasai East I/E | 401208 | Thane |
| Wadala | 400031 | Mumbai |

Table 3 provides the infrastructures that have that have highest potential within Mumbai.

Table3. List of highest potential infrastructures in Mumbai

| **Infrastructures** |
|---|
| Airport Terminal |
| Bank |
| Bus Station |
| Business Service |
| College Auditorium |
| Farmers Market |
| Garden |
| Government Building |
| Indie Movie Theater |
| Light Rail Station |
| Market |
| Monument / Landmark |
| Park |
| Pharmacy |
| Playground |
| Resort |
| Train Station |

Table 4 provides the top 5 post offices located near the best places in Mumbai.

Table 4 Post office located in best places in Mumbai

| Pin Code | Latitude | Longitude | Post Office |
|---|---|---|---|
| 400065 | 19.15613 | 72.870722 | Aarey Milk Colony |
| 400708 | 19.17298 | 73.003532 | Airoli Mode |
| 400069 | 19.11588 | 72.854202 | Andheri (East) |
| 400058 | 19.11725 | 72.833968 | Andheri (West) |
| 400037 | 19.02076 | 72.865256 | Antop Hill |

Figure 2 shows clustering based on total number of infrastructures. Firstly, the location marked with blue colors refers Neighborhoods with a low number of infrastructures (Cluster 1). Secondly, the location marked with red colors refers neighborhoods with a high number of infrastructures (Cluster 0). Finally, the location marked with green colors refers neighborhoods with a low number of infrastructures (Cluster 2).

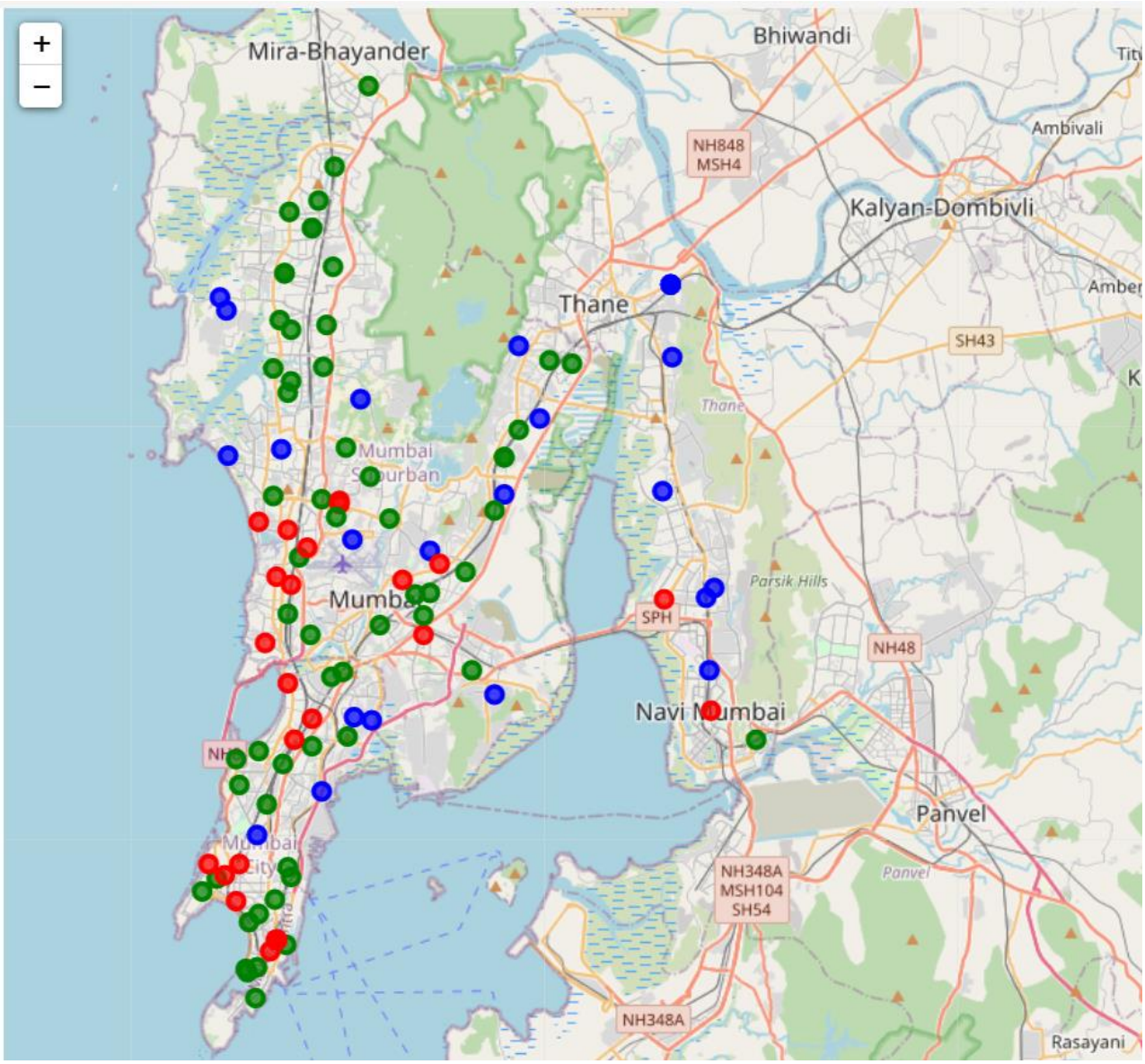


Figure 2. Clustering based on total number of infrastructures in Mumbai.

## References


1. https://en.wikipedia.org/wiki/Mumbai#:~:text=Upon%20India's%20independence%20in%201947,the%20entertainment%20capital%20of%20India.
2. https://www.britannica.com/place/Mumbai
3. https://www.gatewayhouse.in/how-mumbai-became-a-magnet-for-migrants/