Introduction

For the Applied Data Science Capstone Project, this attempts to determine the potential neighborhoods for a new shopping mall in the city of Mumbai in India.

Mumbai is India's financial capital and the seventh most populous city in the world. For an investor looking to invest in a new shopping mall, Mumbai can prove to be very profitable. With the growing population of the city and its increasing standard of living, a shopping mall can be a good investment for entrepreneurs looking to profit from the economic situation of the city. As the demand for luxurious items increases, the footfall for the malls increases and therefore, this project aims to identify neighborhoods with such profitable opportunities.

However, it is important to tap into an unexplored location in Mumbai. This project aims at finding such neighborhoods for these investors.

Business Problem

In a city like Mumbai, a good shopping experience is the optimal one, where a consumer's needs can be met in a single place with options for meals and recreation as well. Shopping malls prove to get high footfall even during weekdays and profits from investing in such ventures are guarantees. In the project, similar neighborhoods are clustered together to find viable options for a new shopping mall.

Data

The data required for the analysis is obtained in three steps.

1. List of neighborhoods in Mumbai. They are categorised by their pin codes obtained from a Mumbai Tour Guide page.
2. The geographical coordinates of the neighborhoods.
3. Venues present in these neighborhoods within 500 m radius of their geographical coordinates.

Data Sources

The list of neighborhoods is obtained using the library Beautiful Soup from a Mumbai Tour Guide page (https://mumbai7.com/postal-codes-in-mumbai/). The Python Geocoder package gives the latitudes and the longitudes of each of those neighborhoods. Then using the Foursquare API, upto 100 venues within 500 m of radius of the neighborhood coordinates are extracted. Venue names, categories and coordinates are fetched for further understanding.