Capstone Project Week 5

New Spanish Restaurant on a great city of India

By aw 2021

1. INTRODUCTION

The objective of this project is to determine, through the study of the neighborhood data in one of the greatest towns on India, trying to check good locations for the start-up of a Spanish Café.

We need a multicultural and tourist character, with a great image in gastronomy, considering all the possible aspects, as population, financial, to extract a geographical candidate for our new restaurant.

Searching on data sources, Mumbai is our city for the study. It´s financial, commercial, and the entertainment capital of India. It is also one of the world's top ten centers of commerce in terms of global financial flow, generating 6.16% of India's GDP, and accounting for 25% of industrial output, and 70% of maritime trade in India. Mumbai Port trust over 70% of capital transactions to India´s economy.

Mumbai has the eighth highest number of billionaires of any city in the world, and Mumbai's billionaires had the highest average wealth of any city in the world in 2008. The city houses important financial institutions and the corporate headquarters of numerous Indian companies and multinational corporations. It is also home to some of India's premier scientific and nuclear institutes.

The city is also home to Bollywood and Marathi cinema industries. Mumbai's business opportunities attract migrants from all over India, and a great group of tourists, for all of this we considered it a great candidate for the study and the startup of a new Spanish Café.

2. DATA RESOURCES

The essential data that we are going to require for the project will be:

2.1. New York neighborhood data source

2.2. Geographical data and coordinates within New York for those neighborhoods

2.3. Data management with recommendations

2.1. Data of the neighborhoods in Mumbai

The data of the neighborhoods in Mumbai was scraped from https://en.wikipedia.org/wiki/List of neighborhoods in Mumbai. The data is read into a **pandas** data frame **using the read_html_method**. The main reason for doing so is that the Wikipedia page provides a comprehensive and detailed table of the data which can easily be **scraped using the read HTML method of pandas**.

2.2. Geographical data and coordinates

The geographical coordinates for Mumbai had been obtained from the **GeoPy library in python**. This data is relevant for plotting the map of Mumbai using the **Folium library** in python.

2.3. Data management with recommendations

The recommendations data has been extracted using the **Foursquare API.** This data contains venue recommendations for all neighborhoods in Mumbai and is used to study the popular recommendations of different neighborhoods as well as build the unsupervised learning model to cluster neighborhoods.