Coursera Capstone

IBM Applied Data Science Capstone

New Shopping Mall Opening in Brooklyn, United States of America

Done by: Hasan Ali Mohamed Taresh

April, 2020

1. Introduction

Nowadays, shopping malls are considered one of the main targets to be visited, despite your needs. You can do grocery shopping, dine at restaurants, watch movies, shop for clothes from various fashion outlets, and much more. In 2017, shopping malls accounted for 8% of retailing space in the United States. As a result, there are many shopping malls in the United States, and many more are being built these days. Opening shopping malls allows property developers to earn consistent rental income. Hence, opening a new shopping mall requires serious consideration and deep study, putting in mind certain factors. One of the main factors to consider is the location of the shopping mall, which could determine whether the mall will be a success or not.

2. Business Problem

The objective of this project is to analyze and choose the best locations in Brooklyn (a borough of New York City, USA) to open a new shopping mall. With the power of Data Science methodologies, along with machine learning techniques such as clustering, this project aims to provide answers for the following business problem: If a property developer is planning to open a new shopping mall in Brooklyn, in New York City, USA, where is the best recommended area that should be chosen to build the mall on?

3. Targeted Audience

This project is certainly useful to property developers along with investors that are interested in opening or investing in a new shopping mall in Brooklyn.

4. Data Description

In order to solve the problem described in the previous chapter, we need to gather the following data:

- A list of neighborhoods in Brooklyn.
- The latitude and longitude coordinates of each neighborhood in Brooklyn.
- Venue data, which is related to shopping malls.

5. Data Sources

1. Wikipedia:

We are going to obtain a list of neighborhoods of Brooklyn from the following page: https://en.wikipedia.org/wiki/Category:Neighborhoods_in_Brooklyn. After extracting the list (using web scraping techniques), we are going to obtain the geographical coordinates for each neighborhood using Python Geocoder package.

2. Foursquare API:

We will use Foursquare API to retrieve venue details for each neighborhood extracted from the Wikipedia page. It is good to mention here that we are going to focus on the “Shopping Mall” category from various categories that Foursquare will provide for us, in order to solve the problem described.