IBM Data Science Capstone Project – Week 4

Author: Chandrasekhar Dwivedi

Introduction:

Shopping malls are like a one-stop destination for all types of shoppers. For retailers, the central location and the large crowd at the shopping malls provides a great distribution channel to market their products and services. Property developers are also taking advantage of this trend to build more shopping malls to cater to the demand. As a result, there are many shopping malls in the city of Paris and many more are being built. Opening shopping malls allows property developers to earn consistent rental income. Of course, as with any business decision, opening a new shopping mall requires serious consideration and is a lot more complicated than it seems. Particularly, the location of the shopping mall is one of the most important decisions that will determine whether the mall will be a success or a failure.

Business Problem:

The objective of this capstone project is to analyse and select the best locations in the city of Paris to open a new shopping mall. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the business question: In the city of Paris, if a property developer is looking to open a new shopping mall, where would you recommend that they open it?

Target Audience:

This project is particularly useful to property developers and investors looking to open or invest in new shopping malls in the city of Paris. This project is timely as the city is currently suffering from oversupply of shopping malls. Data from various studies released last year showed that an additional 10 per cent will be added to existing mall space, and the agency predicted that total occupancy may dip below 84 per cent. One of the local newspapers also reported in March last year that the true occupancy rates in malls may be as low as 40 per cent in some areas, quoting a Financial Times (FT) article cataloguing the country's continued obsession with building more shopping space despite chronic oversupply.

Data Requirements:

The below mentioned data is required to solve the above mentioned business problem.

• List of districts in Paris. This defines the scope of this project which is confined to the city of Paris.

• Latitude and longitude coordinates of those neighbourhoods. This is required in order to plot the map and also to get the venue data.

• Venue data, particularly data related to shopping malls. We will use this data to perform clustering on the neighbourhoods.