IBM Data Science

Capstone Project Report

By: Ismail Raslan

Where to Open a Shopping Mall in Riyadh

Introduction

Going to shopping malls is a great way to relax and enjoy oneself during weekends and holidays. There is a great deal of activities to be done including eating, shopping, and watching movies. For retailers, the central location and the large crowd at the shopping malls provides a great distribution channel to market their products and services. For shops and sellers, it is a great environment for their business to flourish due to the high-density population that visits. Opening a shopping mall requires a tremendous amount of consideration and this report and project exist to even slightly alleviate the tedious research that must be done. The project is done for the city of Riyadh, Saudi Arabia.

Problem

The goal of the project is to analyse possible locations in Riyadh to open a new shopping mall. Using data science and machine learning techniques such as gathering & wrangling the data and clustering respectively, this project aims to provide solutions to answer the question of where to open a shopping mall in Riyadh, Saudi Arabia.

Target Audience

This project is particularly aimed at investors looking to open new shopping malls in the capital city of Saudi Arabia. This project is useful nowadays as the country is welcoming more and more tourists by the day, and shopping malls are always a go-to location when visiting a new city.

Data Required

To solve the problem, we will need the following data:

- List of neighbourhoods in Riyadh.
- Latitude & longitude coordinates of those neighbourhoods. Needed to get data of nearby venues.
- Venue data related to project.

Sources of data and methods to extract them

- Wikipedia provides a list of populated areas:
    - List available through:
        - https://en.wikipedia.org/wiki/Category:Populated_places_in_Riyadh_Province
    - Web scraping techniques are used to extract and store the data from the webpage.

- Foursquare will be used to get the venue data for those neighbourhoods.
    - Particularly venues related to shopping malls (restaurants, theatres, etc…).
    - Done through Foursquare API.

Methodology

We first get the list of neighbourhoods in Riyadh from Wikipedia (https://en.wikipedia.org/wiki/Category:Populated_places_in_Riyadh_Province). Web-scraping was done using Python's "requests" and "BeautifulSoup" modules. Geodata (latitude and longitude coordinates) were acquired through the Foursquare API by using the "Geocoder" package. After gathering the data, we turn it into a "pandas" dataframe so that we are able to visualize it using the "folium" package. We Then gather the top 100 venues that are within a 2000-meter radius of the neighbourhoods. Then, we will analyse each neighbourhood by grouping the rows by neighbourhood and taking the mean of the frequency of occurrence of each venue category. By doing so, we are also preparing the data for use in clustering. Finally, we perform k-means clustering.

Results & Discussion

As shown below, we first notice that there are not a lot of malls in the capital city of Riyadh. The concentration of malls lies within cluster 0, with the availability of a smaller the number of malls in cluster 2. The city is indeed in need of shopping malls to aid its efforts in being a tourist zone.

Figure 1: Clusters 0, 1, and 2 in red, purple, and green respectively

Conclusion

This project was overall very eye-opening and enjoyable. I found everything from picking a topic to work on, looking for, collecting, cleaning, and visualizing the data & using machine learning techniques (k-means clustering here) to perform this analysis quite the ride. I hopefully managed to somewhat answer the question of “Where to open a mall in Riyadh?” and maybe provided a little visual insight on how much the capital city needs shopping malls in its time of growth.

References

Neighbourhoods in Riyadh, Wikipedia. Retrieved from

https://en.wikipedia.org/wiki/Category:Populated_places_in_Riyadh_Province

Foursquare Developers Documentation. Foursquare. Retrieved from

https://developer.foursquare.com/docs