Best Areas to Open New Shopping malls and Shopping plazas in Texas

By: Moushumi Zaman

Introduction

For many shoppers, visiting shopping malls is a great way to relax and enjoy themselves during weekends and holidays. Shopping malls and shopping plazas are one – stop destinations for all types of shoppers. For retailers, the central location and the large crowd at the mall provide a great opportunity to market their products and services. Property developers are also taking the advantage of this trend to build more shopping malls to cater this demand. Opening shopping malls allows property developers to earn a consistent rental income. But as many other business decisions, opening a shopping mall also requires a serious consideration and is lot more complicated than it seems. Specially, 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 object of this capstone project is to analyze and select the best locations to open new shopping malls and shopping plazas in the state of Texas, USA. Using data science methodologies and machine learning techniques like clustering and with help of Foursquare API this project aims to answer the business question: In Texas, USA, if a developer is looking to open a new shopping mall, where would be your recommendation to open it?

Target Audience

This project is particularly targeted to the property developers and investors looking to open new shopping malls and plazas in the state of Texas, USA. Texas is a huge state with many cities. The state gained 187,545 people from migration between July 2017 and July 2018 — even after accounting for people leaving the state, according to U.S. Census data. In 2018, the majority of migrants to Texas — 104,976 people — came from other countries, with the rest arriving from other U.S. states. Though many of this cities already

have shopping malls and plazas, with the increasing population which eventually leads to an increasing demand, opening shopping malls and plazas is a great investment idea.

Data

To solve the problem, I used the following data:

- List of cities in Texas, USA. This defines the scope of the project.
- Latitude, longitude coordinates of the cities. This is required in order to plot the map and to get the venue data.
- Venue data, specially data related to shopping malls and shopping plazas. I used this data to perform clustering on the cities.

Sources of data:

As the most populous cities has more demand for shopping malls and plazas, I used the list of Texas cities ranked by population from the Wikipedia page https://en.wikipedia.org/wiki/List_of_cities_in_Texas_by_population. This list consists of 68 cities. I used the top 40 cities as my dataset. I used pandas 'read_html' method to extract the table from the Wikipedia page.

Then I used geographical coordinates of the cities by using Pyhton Geocoder package which gave me the latitude and longitude coordinates of the cities.

Finally, I used the Foursquare API to get the venue data for those cities. Foursquare has one of the largest databases of 105+ million places and is used by over 125,000 developers all around the world. Foursquare API provided me with many categories of venue data. But in this project I am particularly interested with the data of shopping malls and shopping plazas. Therefore, I built a new data frame with only the shopping malls and shopping plaza information.