Coursera Capstone

IBM Applied Data Science Capstone

Opening a New Shopping Mall in Chennai, TN, India

Description: Shopping Mall Analysis in Chennai city neighborhoods

By: SRIRAM

May 2020

Introduction

Malls are not only a shopping place but a place to rejuvenate, socialize and entertain. In big retail stores you get everything under one roof from branded clothes, grocery and electronics to foot wear. Without a doubt malls have changed the shopping experience of Indians. Doing shopping in the scorching heat of the sun has been replaced by AC shopping. Youth take this as a status symbol. Visiting malls and buying branded products satisfy their thirst for better quality of life. Teenagers do come to show off. Certainly shopping malls are bringing in a new culture in India which is different from the traditional culture as far as shopping is concerned. Chennai offers an immense market opportunity because of increased income and changed lifestyle of middle class families. Property developers are also taking advantage of this trend to build more shopping malls to cater to the demand. As a result, there are currently many shopping malls in the city of Chennai 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 analyze and select the best locations in the city of Chennai 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 Chennai, if a property developer is looking to open a new shopping mall, where would you recommend that they open it?

Target Audience of this project

This project is particularly useful to property developers and investors looking to open or invest in new shopping malls in the capital city of Tamil Nadu i.e. Chennai. This project is timely as the city is currently suffering from oversupply of shopping malls. Data from last year showed that an additional 15 per cent will be added to existing mall space, and the agency predicted that total occupancy may dip below 86 per cent. The local newspaper It is also reported in August 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

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

- List of neighborhoods in Chennai. This defines the scope of this project which is confined to the city of Chennai, the capital city of the country of Tamil Nadu in India.
- Latitude and longitude coordinates of those neighborhoods. 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 neighborhoods.

Sources of data and methods to extract them

This Wikipedia page (https://en.wikipedia.org/wiki/Category:Suburbs_of_Chennai) contains a list of neighborhoods in Chennai, with a total of 61 neighborhoods. We will use web scraping techniques to extract the data from the Wikipedia page, with the help of Python requests and beautifulsoup4 packages. Then we will get the geographical coordinates of the neighborhoods using Python Geocoder package which will give us the latitude and longitude coordinates of the neighborhoods.

After that, we will use Foursquare API to get the venue data for those neighborhoods. Foursquare has one of the largest database of 105+ million places and is used by over 125,000 developers. Foursquare API will provide many categories of the venue data, we are particularly interested in the Shopping Mall category in order to help us to solve the business problem put forward. This is a project that will make use of many data science skills, from web scraping (Wikipedia), working with API (Foursquare), data cleaning, data wrangling, to machine learning (K-means clustering) and map visualization (Folium). In the next section, we will present the Methodology section where we will discuss the steps taken in this project, the data analysis that we did and the machine learning technique that was used.