Coursera Capstone
IBM Applied Data Science Capstone

Opening a New Shopping Mall in Istanbul, turkey

Content

- Introduction :The business problem to be solved in this project
- Data section : Describe Data sources needed to solve the problem

Introduction

- The importance of a shopping mall is mainly for people to get out of the house for a while and do something entertaining. Shopping malls can provide the best shopping experiences such as social gatherings, entertainment, performances, product launches, promotions and festivals. The events list at shopping malls goes on and on for any, particular, person to be entertained for a number of hours.

- Any shopping mall can be a great place to hang out with friends, eat, shop, and more. You can go to all your favorite stores and personally I believe that parents enjoy it just as much as kids. Malls provide you with the opportunity to send the ladies to look at makeup, perfume or clothing. The gentleman can look at electronics and sports equipment and so on. Malls can be very helpful because all the needs are in one building.

- Shopping malls tend to be a major tourist attraction. The mall can be more convenient, for a tourist, to have one central location to do all their shopping; rather than to have to drive many miles just to buy different types of products for their personal needs.

Business Problem :

- The objective of this capstone project is to analyze and select the best locations in the city of Istanbul, turkey 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 Istanbul, turkey , if a property developer is looking to open a new shopping mall, where would you recommend that they open it?

Data :

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

- List of neighborhoods in Istanbul. This defines the scope of this project which is confined to the city

- 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.

Data sources :

- 1- This Wikipedia page (https://en.wikipedia.org/wiki/Category:Neighbourhoods_of_Istanbul) contains a list of neighborhoods in Istanbul , We will use web scraping techniques to extract the data from the Wikipedia page.

- 2- 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.

- 3- we will use Foursquare API (https://foursquare.com/) to get the venue data for those neighborhoods , 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.