Coursera Capstone

IBM Applied Data Science Capstone

Opening a New Restaurant in Miami, Florida

By: Haroldo Ale Filho

June 2020

Introduction

Miami is known for its tourism. Travelers come and go from the city during all seasons and from multiple geographies. We see a consistent influx of LATAM, European and American tourists that come to enjoy our beaches, restaurants and vibrant culture. This is a business opportunity for small restauranteurs and hospitality groups to open successful restaurants in the city's best neighborhoods. The problem is finding the best neighborhood where restaurants are more frequently visited. This includes finding their nearness to hotels, which will determine their rate of failure or success. The closer to hotels, more traffic and a better chance of success.

Business problem

The objective of this capstone project is to analyse and select the best locations in the city of Miami, Florida to open a new restaurant. Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the business question: If a restauranteur is looking to open a restaurant in Miami, where would you recommend that they open it?

Target audience of this project

The target audience are smaller hospitality groups fighting against Groot Hospitality and SBE that dominate the South Florida restaurant scene. The big players have saturated the market with good restaurants that have the wherewithal to bring in different customers with they social marketing skills.

Data

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

- List of neighborhoods in Miami. This defines he scope of this project as the neighborhoods of the city of Miami.
- Latitude and longitude of these neighborhoods, required to perform the map clustering and get the hotel data.
- Venue data, related to the hotels in the city of Miami, FL. It will be use this datapoint to show our clusters.

Sources of data and methods to extract them

This Wikipedia page (https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Miami) contains a list of neighbourhoods in Miami. We will use web scraping techniques to extract the data from the Wikipedia page, with the help of Python requests and beautifulsoup packages. The list also contains the geographical 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 Hotel 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.