Determining best locations to open a new Souvlaki Shop in: Athens - Greece

Kostas Smonos

November 11, 2020

1. Introduction

1.1 Background

Athens is the capital of Greece populated by more than 3.150.000 people, and a must-go for tourists visiting the country. It has a large number of restaurants of any kind, with many of those enjoying huge success. Souvlaki is a hot commodity in Greece, both for tourists and locals but it seems like bad practice to have many such Restaurants in a small area. Opening a Souvlaki Shop in a good spot can be highly profitable. Determining the best spot to open will be a key factor in its success!

1.2 Problem

Data that might contribute to determining the best spot to start a Souvlaki Shop might include: Location data of existing Souvlaki Shops in Athens, Venue population per-neighborhood, centered around nearby subway stations. The project aims to determine the best spots to start up a new Restaurant based on these data.

1.3 Interest

The Restaurant Managers that want to open up the Souvlaki Shop would be obviously very interested in determining the best spot for their investment. Others who will be interested are the people in the vicinity of the new Shop, that will certainly enjoy the new addition to their area.

2. Data acquisition and cleaning

2.1 Data Sources

Highly accurate Venue locations and data will be amassed utilizing the Foursquare API. I will also scrape this page to get all subway stations in Athens: https://en.wikipedia.org/wiki/List_of_Athens_Metro_stations. I will attempt to find accurate location data for those using Wikipedia as well.

2.2 Data Cleaning

Data scraped from the Wikipedia page needed quite a lot of tidying up. First I followed a separate link found in the table for each subway station, where location data for each station can be found. This method worked well enough for most of the entries except a few that ended up with NaN values, as their Wiki did not actually contain any location information. I looked up the location of the problematic stations and input it manually from Google Maps.

After fixing these problems that occurred from scraping Wikipedia, I grabbed the venue information, in a range of 400 meters around each subway station. We want our spot to be easily accessible.

2.3 Feature Selection

After data cleaning there were 2295 samples in the venues dataset, of 249 unique categories. The features we are most interested in to make our report are; the location of the venues (we have already kept only those 400 meters from any subway station), the number of venues assigned to each subway station, and the number of Souvlaki Restaurants around each station in particular.

3. Exploratory Data Analysis

3.1 Calculation of target variable

Number of Souvlaki Restaurants for each neighborhood (subway station area) had to be calculated by summing the correct values on the dataset. The number of total venues followed a similar procedure. I filtered the final dataset to only contain areas with more than 45 venues as we wish to find an area with good venue density that is guaranteed to have a good flow of potential customers.

(the mean value of venues within 400 meters of all subway stations in Athens was about 35)

4. Clustering

4.1 K-means clustering

I used k-means clustering to cluster my data into subsets. Since I had already filtered my data according to my specifications, the only independent variable was the amount of Souvlaki Shops in the area, so it is quite evident that k-means produced the correct result.

5. Conclusions

In this study I looked for the best spot to open a new Souvlaki Shop in Athens, Greece. I analyzed data of all the subway stations in Athens, as well as Foursquare data of Venues in close range to stations. The models I have built are easily applicable to any other commodity so it can prove useful into analyzing the potential of many other new shop ideas as well as Souvlaki Shops.

The areas I identified as the best for opening up a new Souvlaki Shop from my model were within 400 meters of the following subway stations:

- Omonoia
- Petralona
- Sintagma
- Evangelismos
- Megaro Mousikis

These spots are guaranteed to have a good flow of people, only one other Souvlaki Shop, as well as good distance from the primary means of transportation in Athens, the subway.