

Analyzing the Neighborhoods in Athens for Starting a Cafe

Applied Data Science Capstone Project

By Dimitrios Christodoulou

Introduction/Business Problem

Starting a new business can be a really demanding process and requires a vast amount of money to invest in advance, so the selection of the store location can have a significant impact on the prosperity of the store. This project can be useful for people who are interesting in opening a new coffee store in London since this project focus to determine possible locations for starting a café.

Data Collection

The implementation of this project requires a variety of data, which must be collected from multiply sources. The required data are the following:

- 1) Neighborhoods of London (Name, Post Code).
- 2) Geographical coordinates of neighborhoods.
- 3) Venue data for London neighborhoods.

1. Neighborhoods Data

The list of the neighborhoods in London will be taken from Wikipedia using web scraping (https://en.wikipedia.org/wiki/List_of_areas_of_London). Wikipedia's Neighborhoods Data format is a table that can easily be read by using the `read_html()` method of pandas.

2. Geographical Coordinates

Using the GeoPy library in python, the London neighborhoods' geographical coordinates can be easily obtained. To be more precise, using this library we can obtain the latitude and longitude for all neighborhoods in Athens. Using these coordinates, we can get tips and even plot the neighborhoods on the map using the Folium Library.

3. Venue Data

The venue data will be extracted using the Foursquare API. The venue recommendations for all neighborhoods in London which are prerequisites to implementing the project are extracted using the Foursquare API.