

IBM Capstone Project

The Battle of Neighborhoods

Ines Ben Ameer

May 12, 2019

Contents

1	Introduction	2
2	Business Problem	2
3	Explored Data	2
3.1	Data specification	2
3.2	Data Sources	2
3.3	Data Samples	2

1 Introduction

This report details the specifications of my final capstone project for IBM Applied Data Science Capstone course in Coursera.

For this project, I'm going to use the Foursquare location data and machine learning tools in order to segment and cluster neighborhoods in Athens, Greece.

Being a fan of both Italian food and the grandiose Greek capital, I chose to setup a tool that helps an Italian restaurateur to open an Italian restaurant in Athens. Since Italian food is very spread worldwide, it is quite common to come across a Neapolitan or Sicilian restaurant in any neighborhood or city. Therefore finding the most suitable location to open such a restaurant is one of the most important decisions for this entrepreneur as a good location is one of the pillars of a profitable business.

2 Business Problem

The objective of this capstone project is to help an Italian entrepreneur to look for the most fitting location to open a new Italian restaurant in Athens, Greece. By using machine learning methods such as clustering and a location data platform like Foursquare, we will extract features and characteristics of neighborhoods in Greece, cluster them and finally determine the ideal place for opening the restaurant.

3 Explored Data

The first step of building the tool would be naturally to specify the data we're going to need and then collect it from the adequate sources.

3.1 Data specification

In order to be able to cluster the neighborhoods of Athens we need:

- List of neighborhoods in Athens: It is necessary to collect the neighborhoods of the city of Athens to apply the clustering.
- Latitude and Longitude of the neighborhoods: We need the geolocation coordinates of each neighborhood in order to extract the venues located in each one.
- Venues: Finally we need to extract the list of Italian restaurants in each neighborhood to complete our data collection.

3.2 Data Sources

We used the following data sources to collect the features specified above:

- List of neighborhoods in Athens: This part is collected from this [Wikipedia page](#) [1].
- Latitude and Longitude of the neighborhoods: The coordinates are extracted via **geopy.geocoders** python package .
- Venues: The venues and their attributes are provided by the Foursquare API [2]

3.3 Data Samples

- List of neighborhoods in Athens and their coordinates:

Neighborhood	Latitude	Longitude
Sepolia	38.006449	23.717277
Skouze Hill	37.977363	23.728436
Thiseio	37.976766	23.720329
Treis Gefyres	38.014098	23.718425
Vathi	35.359413	23.595234

- Venues: The venues and their attributes are provided by the Foursquare API:

Neigh	Neigh.Lat	Neigh.Long	Venue	V.Lat	V.Long	V.Cat
Aerides	36.147765	22.989737	Belvedere	36.148118	22.989718	Italian Restaurant
Akadimia	37.980285	23.734528	Piadena L' Umbro	37.979424	23.735630	Italian Restaurant
Akadimia	37.980285	23.734528	La Pasteria	37.978384	23.736132	Italian Restaurant
Akadimia	37.980285	23.734528	Frankie	37.980583	23.737988	Italian Restaurant
Akadimia	37.980285	23.734528	Il Postino	37.982145	23.736906	Italian Restaurant

References

- [1] Wikipedia, “Category:neighbourhoods in athens.” [Online]. Available: https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Athens
- [2] Foursquare, “Foursquare api.” [Online]. Available: <https://foursquare.com/>