

Capstone Project - The Battle of Neighborhoods (Week 2)

1. Introduction – The problem

Millions of people worldwide are travelling every day for different locations. With the world wide web everyone can easily have access to billions of travelling data. However, sometimes the enormous load of information is inconvenient, especially for the elders. Attractions, hotels, restaurants and beautiful sceneries are searched every day from travelers, for lots and lots of cities or countries. So, this project is an approach for keeping such search relatively simple. I believe that the project could be easily used from tourists, travel agencies or even businessmen/businesswomen from all over the world. Tourists and travel agencies could easily find the desired vacation spot while business people could search cities or countries for venues and possible investments. In my scenario, a travel agency used the code in order to find the best result for a demanding customer.

2. Data collection

The data used in this project are collected from 3 different websites. First, I found a csv file from Simple Maps (<https://simplemaps.com/data/world-cities>) that includes cities from all over the world along with some useful data. From the csv file, I kept:

- City names
- Country
- Latitude
- Longitude

Also, I used data from Foursquare <https://foursquare.com/city-guide>, using their API. Specifically, I used:

- Ratings (in order to find the most visited and liked venues)
- Longitude and latitude of all venues
- Names of venues
- Category of venues (bars, restaurants, attractions, etc)

Lastly, I used another API that provided available hotels and prices. The API was created by Makcorps (<https://www.makcorps.com/index.html>)

3. Methodology

In my case study, the project was used by a traveling agency. A demanding customer wanted the agency to find him the best spot for his vacation. Specifically, the customer emailed the agency and said: “I would like you to find the best location for me. I am interested in travelling to Greece or Italy, but I don’t want it to be too crowded. I think population of 160000-165000 people would be more than enough. Also, I would like to see what attractions I could possibly experience there. Oh! Please check for a hotel room too please, 200€ per night would be good, thanks!

The methodology is quite simple, so I am presenting it as a list:

- a) Importing the csv file from simple maps (worldcities.csv).
- b) Analysing the imported data.
- c) Dataframe shorting by: city, country, population, latitude, longitude.
- d) Checking the size of the dataframe.
- e) Shorting the dataframe in order to find the cities in Greece or Italy.
- f) Shorting the cities in both countries according to the population.
- g) Collecting the data from the result city.
- h) Calling foursquare data, by geo coordinates for this specific city.
- i) Shorting the data and keeping only venue name, venue category, geo coordination.

- j) Building a map with data from foursquare. The venues can be seen as blue dots.
- k) Calling the Makcorps API.
- l) Getting results for the available hotels.
- m) Resulting to the hotel with best price.

4. Results and Discussion

So, after checking all the results, it is clear that the customer should travel to the city of Patras (my hometown obviously), since this city covers all of his needs: Desired country, desired population, interesting venues, hotel room. It's a really interesting fact that the results can be adapted to everyone's needs. As an example, a photographer would maybe like to travel to a city with high population. So, we could check for high population values, and sort the dataframes in a descending population order. The point is, that according to the data that you have, the results can be adapted in lots of people's needs. That's what Data Science is all about.

5. Conclusion

To conclude, data analysis on cities all over the world was successfully done. Of course, in the future more analysis can be done even in the same set of data. The project can be used by travellers, travel agencies or business people that would like to invest in some city, or even find the city that they would like to invest in. This project is a strong tool for real life problems and can be easily give even more results by adding or changing the used data.