nwxjgi2ko

July 19, 2023

0.1 Capstone Project - The Battle of Neighborhoods

0.1.1 A PIECE OF ART IN THE CAPITAL CITIES OF ALL COUNTRIES

0.1.2 INTRODUCTION

Where would you recommend an art adventurer to visit in order to fulfill her/his hunger for various art pieces? The world is full of wonders and maybe the most ambiguous ones are considered as art. Many people dedicated themselves to travel around the world and discover these ambiguities. This project seeks for art related venues in the capitals of all countries, cluster them and reveals the different cities in terms of art venues although they have closer geography. By the nature of this project, travel companies and their costumers might be interested in this project to find similar and different places around the world.

0.1.3 DATA

In order to accomplish this goal, latitudes and longitudes of the capital cities of all countries in the world are required. The "simplemaps.com" offers a simple, accurate and up-to-date database of the world's cities and their locations. From this data, I need to select the capital cities and using their latitude and longitude values, I need to explore the venues with "art" section around these cities by utilizing Foursquare API.

The "simplemap" data contains city name, corresponding latitude, longitude and country name along with many other features for all cities in the world. I only use the features that I mentioned in this data and then filtered the capital cities for each country. After cleaning the data I have 225 countries with 225 capital cities.

Then I used the acquired location data to explore the nearby art venues from the Foursquare API. Using the "explore" option, I look for top 25 art venues in 10 km radius for each city center and I get "Venue", "Venue Latitude", "Venue Longitude" and "Venue Category" columns with 3316 row in total.

0.1.4 METHODOLOGY

After pivoting and groping the data at hand, we get a new version of the data which includes 41 unique categories as columns. The cells contains the numbers of the corresponding category in the city. Utilizing this data, I applied DBSCAN from sklearn library which runs DBSCAN clustering from vector array or distance matrix. After the tuning of the parameters of the model the result contains 5 different clusters.

0.1.5 RESULT

The resulting map is as follows:

It can be seen that there are two dominant clusters with red and orange marks. So, going to these places will offer somehow similar taste to their visitors. However other colors will serve different taste in terms of art venues.

0.1.6 DISCUSSION

Different epsilon parameters in the DBSCAN Model return different number of clusters with similar results. The best result is as presented. However I believe that the results are quite satisfying.

0.1.7 CONCLUSION

To sum up, the adventurer can go to purple cities such as Dublin with a lot of theater opportunities at the center of the city.