Capstone Project- The Battle Of Neighborhoods

Around The World

Background:

- World as we all know and would agree is a vivid concept possessing its fair share of enigmas and wonders, of cultures and diversities, of art and beauty. With the increase in art forms and appreciation, we find ourselves amongst many of those who seek travelling to such places of myriad art and culture.While this real life situation is common and relatable, one who would always look puzzled in decided on such art places.

- Through this project, I hope to go through the world cities and pick out the art venues located in different cities.Cluster these palces and indicate the most preferrable art places to visit.One might find this information analysis handy while considering their next Destination that spells Art.

Problem/Analytic Approach:

- Since the question is to find various cities that have art aspects in them, the problem deals with identifying the art aspects in the world and grouping them based on their geographical location into actual cities that they belong. This is a problem that can approached through the model of Clustering.

Interest:

- One might find this information analysis handy while considering their next Destination that spells Art.Travellers might be interested in using such clustering analysis to easily decide or prioritize their route map of art venues.

Purpose Of Data:

- In this project, we try to locate cities that are home for art places. So, we require the geographical locations that are specified by key attributes like Latitude and Longitude coordinates.Using this data, I will be able to explore the art areas belonging to these specific cities through Foursquare API.

- It is evident that a comprehensive dataset is required that possesses these values for a wide range of cities.I was able to procure a suitable dataset through 'SimpleMaps'.This dataset consists of over 200 countries with their respective capital cities, provided with their Latitudinal and Longitudinal Coordinates.the cities would be filtered for art areas and then applied through Foursquare API to further explore individual venues for further nearby places for art.

Features of the applying Dataset:

- Resource:'www.simplemaps.com'
- Number of rows:225 - i.e, there are 225 Countries with respective Capitals
- Latitude Coordinates Provided
- Longitudinal Coordinates Provided
- Population
- ID

Further Steps:

- The Dataset will be filtered for necessary columns

- These Details will be put through Foursquare API

- Then we try to apply the acquired location data to explore the surrounding art venues utilizing the Foursquare API. Through the "explore" option, we check for top 25 art venues within approximately 10 km radius around every city center and receive "Venue", "Venue Latitude", "Venue Longitude" and "Venue Category" columns for all 3316 rows.

Analyzing the procured dataset, we obtain a redefined version including 41 unique categories into columns. The cells contain the number of entities belonging to the respective category in the city. Utilizing the above dataset, we need to apply the 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.

Characteristics:

- Acquired knowledge and skills from week 3.

- Foursquare API to retrieve all areas to explore in each neighborhood.

- Grouping the art places by neighborhoods.

- Counting the number of venues before filtering the top 10 most popular kinds of places in neighborhoods

- Providence has more venues than Hartford (512 vs 460)
- Federal Hill has the most venues in Providence
- Downtown has the most venues in Hartford
- Most of venues in Federal Hill are restaurants, bars
- Venues in Downtown Hartford are more diverse includes: banks, concert hall, gyms…
- Bars, Restaurants, Bakeries, and Diners are very common in Providence
- Hartford is more about commercial

- After following the adequate methodology and procedure, we curate the datasets into clusters. Inculcating Various Epsilon parameters in to applied DBSCAN Model returns different number of clusters comprising similar results. The best result is as shown above.I hope that these results as evident to cluster the common art venues which are more profound, preferred and aid in travellers and venturists to develop theri understanding and aid in their decision to choose the next art destination.

- Hence from the course of project and obtained outputs, we can conclude that most travellers can visit the purple marked cities such as Dublin which has a lot of theatric opportunities at the heart of the city. This project may deem useful to the art venues visitors or travellers or travelling agencies etc to make recommendations, choose destinations or explore more places nearby.

For travellers:

- Providence may be congenial for spending a considerable amount of time relishing the local food and cuisine.

- Hartford is not as adept in cuisine but is more commercial.

For business owners:

- Providence is lucrative and welcoming for startups or businesses pertaining to food and travel industry

- Hartford is profitable for investing the insurance industry