Capstone Project Report :

Analyzing the locations in Paris to open a restaurant

Introduction

The primary objective of this project is to analyze the locations in Paris to choose a location that is ideal to open a restaurant. The project is mainly targeted towards investors who are looking for ideal places with less competition to open a restaurant. The task is accomplished using simple data analysis techniques in Python.

Data

The data used for the analysis are :

(i) Paris location data accessible from “https://lvdesign.com.fr/coursera_data/paris.csv”

(ii) Further location data is accessed through FourSquare API.

Methodology

Firstly the list of locations and their latitudes and longitudes are acquired through the Paris location data. Further, using these attributes the searches are made on FourSquare. The connection to FourSquare is established by giving the client_id, client_secret of the developer account and the version. The geocoder package is used to convert the address into geographical coordinates in the form of latitudes and longitudes.

The data is gathered into a pandas dataframe and the locations are visualized in the map. The map is generated using folium package. Further, the foursquare API is used to search for restaurants within a given radius. Foursquare returns the data in JSON format and hence is converted to a

pandas dataframe. Then, the locations are analyzed by the number of restaurants at a particular postal code. This is done by using the group by function. The best location is chosen accordingly.

Conclusion

The project mainly aims on how the FourSquare API can be accessed and the how the analysis is made using the location data. According to the results acquired, it is understood that the areas with zip code 75006, 75014, 75015 have the least number of restaurants. Therefore, these areas would be ideal to open a restaurant with less competition.

Future research

The current models only considers one factor i.e., considering places with less competition. Gradually, the project can be improved to analyze many other factors like the population, type of crowd, cost of living, surroundings, etc.