PARIS'S BATTLE OF NEIGHBORHOODS

IBM Applied Data Science Capstone

Submitted By: Youssra Saadeddine

Introduction:

Paris, renowned as the city of light and love, is simply Europe's most enthralling capital. Beyond the boulevards and classical monuments, you'll find fascinating small museums, scores of family-run restaurants and neighborhoods filled with restaurant, food and drink shops, boutiques and bakeries.

Paris is the best city of food in 2019, in Paris you'll find the best places to eat and drink.

Having a restaurant or a food and drink shop at Paris is absolutely a good project but before having this you should do an analysis of location.

Business Problem:

The objective of this capstone project is to analyze and select the best locations in Paris to open a restaurant. Using data science methodology and machine learning techniques like clustering.

if someone is looking to open a new restaurant or a new drink and food shop, where would you recommend that they open it?

Target Audience of this project

This project is useful to property all people looking to open a restaurant in Paris.

DataSet:

To solve the problem, we will need the following data:

• List of neighborhoods in Paris.

• Latitude and longitude coordinates of those neighborhoods.

This is required in order to plot the map and also to get the venue data

• Venue data, particularly data related to restaurant and food shop in Paris. We will use this data to perform clustering on the neighborhoods.

Sources of data

This Wikipedia page (https://en.wikipedia.org/wiki/Quarters_of_Paris) contains a list of neighborhoods in Paris.

methodology

We will use web scraping techniques to extract the data from the Wikipedia page, with the help of Python requests and beautifulsoup packages. Then we will get the geographical coordinates of the neighborhoods using Python Geocoder package which will give us the latitude and longitude coordinates of the neighborhoods.
After that, we will use Foursquare API to get the venue data for those neighborhoods.

This is a project that will make use of many data science skills, from web scraping (Wikipedia), working with API (Foursquare), data cleaning, data wrangling, to machine learning (K-means clustering) and map visualization (Folium).