Capstone Project - The Battle of Neighborhoods Suitable New Coffee Shop Locations in Paris

This notebook contains multiple parts:
- A description of the problem and a discussion of the background
- A description of the data and how it will be used to solve the problem
- Methodology and Exploratory Data Analysis
- Inferences and Discussion

1. Introduction and Discussion of the Business Objective and Problem

Problem:

A coffee shop owner wants to open a new coffee shop in Paris. The owner wants to analysis every data to find out best place to open shop.

Criteria

One of the criteria is to open coffee shop near restaurants.

The other criteria would be to find place densly clustered shop and restaurant areas.

Aim

The aim is to use owner's resources effectively by choosing correct place. This analysis provide essential data for this purpose.

Result

The results will be conveyed to owner in a simple and understandable way.

2. Data Science Methodology

Data Requirements

The data of Paris needs to be researched and a suitable useable source identified. If it is found but is not in a useable form, data wrangling and cleaning will have to be performed.

The cleansed data will then be used alongside Foursquare data, which is readily available. Foursquare location data will be leveraged to explore or compare districts around Paris, identifying the high traffic areas where consumers go for shopping, dining and entertainment and most interested in opening new stores.

Roadmap:

Outline the initial data that is required:

- City data for Paris including names, location data if available, and any other details required.

Obtain the Data:

- Research and find suitable sources for the city data for Paris.
- Access and explore the data to determine if it can be manipulated for our purposes.

Initial Data Wrangling and Cleaning:

- Clean the data and convert to a useable form as a dataframe.

Data Analysis and Location Data:

- Foursquare location data will be leveraged to explore or compare districts around Virignia Beach.
- Data manipulation and analysis to derive subsets of the initial data.
- Identifying the high traffic areas using data visualisation.
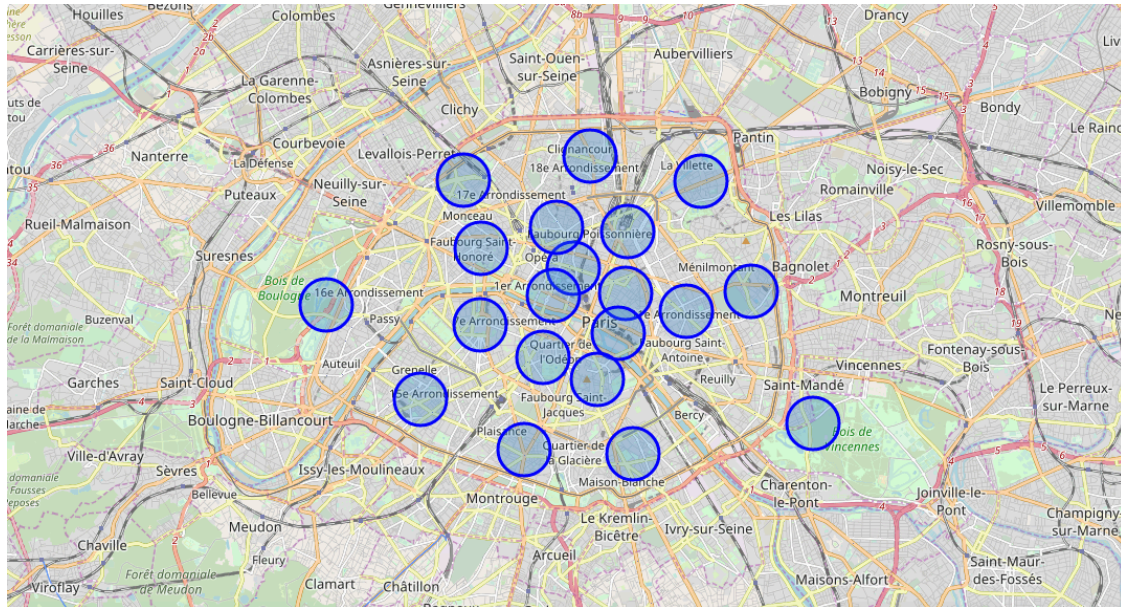
Visualization:

- Analysis and plotting visualizations.
- Data visualization using various mapping libraries.
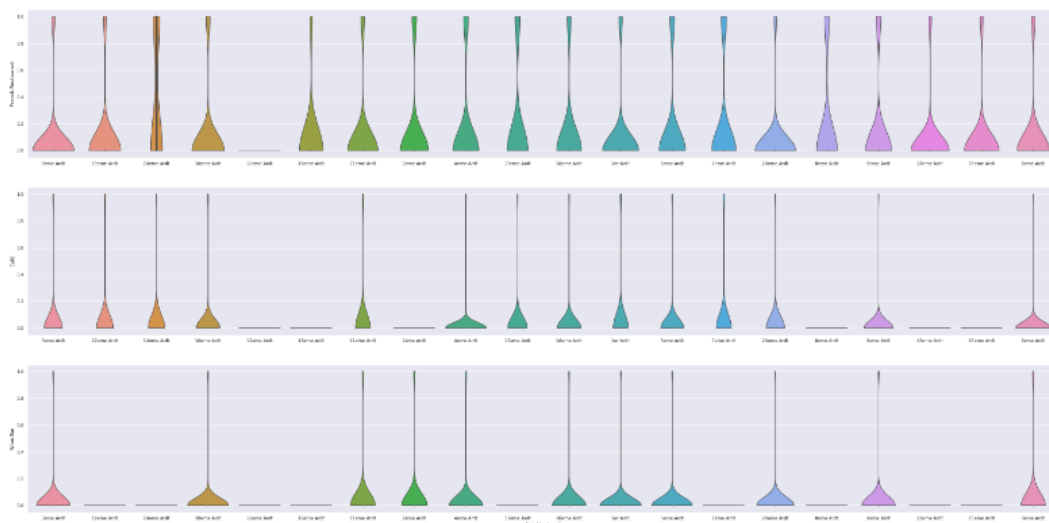
Discussion and Conclusions:

- Recommendations and results based on the data analysis.
- Discussion of any limitations and how the results can be used, and any conclusions that can be drawn.
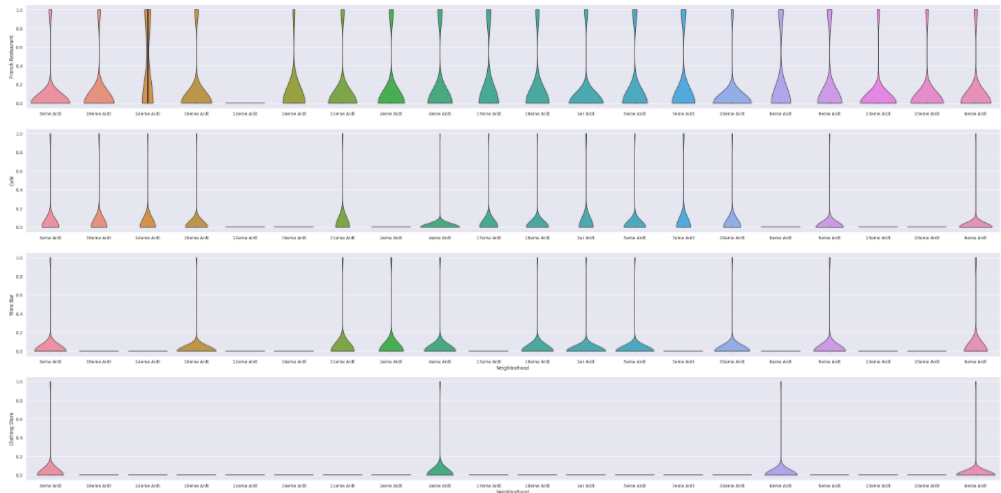
3. Discussion and Conclusions

Neighborhood Locations of Paris:

Frequency distribution for the top 3 venue categories for each neighborhood (click to enlage)

Frequency distribution for the top 3 venue categories for each neighborhood

Possible Coffee Shop Locations:

	Arrondissement_Num	Neighborhood	French_Name	Latitude	Longitude
0	3	Temple	3eme Ardt	48.862872	2.360001
1	4	Hotel-de-Ville	4eme Ardt	48.854341	2.357630
2	6	Luxembourg	6eme Ardt	48.849130	2.332898