**Capstone Project - The Battle of Neighborhoods**

**Suitable New Store Locations in Paris for a Fashion Retailer**

.

***1 Introduction and Discussion of the Business Objective and Problem***

**Locations for New Fashion Stores in High Traffic Areas in Paris France**

**The Task At Hand**

A digitally native vertical fashion retailer, with a substantial e-commerce footprint, has begun the rollout of brick and mortar stores as part of their omnichannel retail strategy. After rolling out stores in a few select cities by guessing where the best locations were to open, as part of their store expansion for Paris they've decided to be more informed and selective, and take the time to do some research.

I've been given the exciting task of assisting them to make data-driven decisions on the new locations that are most suitable for their new stores in Paris. This will be a major part of their decision-making process, the other being on the ground qualitative analysis of districts once this data and report are reviewed and studied.

The fashion brand is not what is considered high-end, they are positioned in upper end of the fast fashion market. As such, they do not seek stores in the premium upmarket strips like Avenue Montaigne, but rather, in high traffic areas where consumers go for shopping, restaurants and entertainment. Foursquare data will be very helpful in making data-driven decisions about the best of those areas.

**Criteria**

Qualitative data from another retailer that they know, suggests that the best locations to open new fashion retail stores may not only be where other clothing is located. This data strongly suggests that the best places are in fact areas that are near ***French Restaurants, Cafés and Wine Bars***. Parisians are very social people that frequent these place often, so opening new stores in these locations is becoming popular.

The analysis and recommendations for new store locations will focus on general districts with these establishments, not on specific store addresses. Narrowing down the best district options derived from analysis allows for either further research to be conducted, advising agents of the chosen district, or on the ground searching for specific sites by the company's personnel.

**Why Data?**

Without leveraging data to make decisions about new store locations, the company could spend countless hours walking around districts, consulting many real estate agents with their own district biases, and end up opening in yet another location that is not ideal.

Data will provide better answers and better solutions to their task at hand.

**Outcomes**

The goal is to identify the best districts - *Arrondissements* - to open new stores as part of the company's plan. The results will be translated to management in a simple form that will convey the data-driven analysis for the best locations to open stores.

***2 The Data Science Workflow***

**Data Requirements**

The main districts in Paris are divided into 20 *Arrondissements Municipaux* (administrative districts), shortened to *arrondissements*.

The data regarding the districts in 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 - the areas where the fashion brand are most interested in opening new stores.

**The Data Science Workflow for Part 1 & 2 includes the following:**

* **Outline the initial data that is required:**
  + District data for Paris including names, location data if available, and any other details required.
* **Obtain the Data:**
  + Research and find suitable sources for the district 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.

**The Data Science Workflow for parts 3 & 4 includes:**

* **Data Analysis and Location Data:**
  + Foursquare location data will be leveraged to explore or compare districts around Paris.
  + Data manipulation and analysis to derive subsets of the initial data.
  + Identifying the high traffic areas using data visualisation and tatistical nalysis.
* **Visualization:**
  + Analysis and plotting visualizations.
  + Data visualization using various mapping libraries.
* **Discussion and Conclusions:**
  + Recomendations 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.