Capstone Project – Battle of Neighbourhoods

Leverage the FourSquare location data to explore or compare neighbourhoods or cities of your choice

Project Title: Determine best locations to establish new stores for a fashion retailer in Paris

Introduction

A renowned fashion retailer, with a substantial e-commerce footprint, has begun the rollout of their new stores, which begins in Paris. As part of this project, we are going to make data-driven decisions on the new locations/ districts that are more suitable for their new stores in Paris.

This retailer brand is not what can be considered as high-end, but they are positioned in the mid-range of the fashion industry. So they do not seek to establish their stores in premium/ plush localities, but rather, in high traffic areas where consumers go for shopping, visit restaurants or go for entertainment. We will be using Foursquare data to make decisions about the best of the areas.

Observing the data from the other retail stores suggests that the best locations to open new stores may not only be where there are other fashion outlets, but in fact, the best areas are near French restaurants, cafes, wine bars. Since Parisians are very social in nature and visit such places quite often, opening new stores in such locations is getting popular.

So, the analysis will focus on districts in Paris with such establishments – narrowing down to the best district options, and then further research can be done.

Data Requirements:

The main districts in Paris are 20. The data regarding these 20 districts needs to be researched.
Then, the cleansed data will be used along side the FourSquare data, which is readily available. FourSquare data will be used to explore or compare the districts in Paris, identifying the most populous areas where most of the fashion brands are set up.

Data for the districts should be collected - to select the most suitable of these areas for new stores. Obtained this data from:

https://opendata.paris.fr/explore/dataset/arrondissements/table/?dataChart

Download the .csv file from the website and upload it to the GitHub repository.

Exploring, Wrangling and Cleaning the Data ¶

```
# Rename the necessary columns

paris.rename(columns={'NAME': 'Neighborhood ', 'CAR': 'Dist_Num', 'Geometry_X ': 'Latitude', 'Geometry_Y': 'Longitude',  'LAR
paris
```

	Dist_Num	Neighborhood	NSQAR	CARINSEE	French_Name	NSQCO	SURFACE	PERIMETRE	Latitude	Longitude
0	3	Temple	750000003	75103	3eme Ardt	750001537	1.170883e+06	4519.263648	48.862872	2.360001
1	2	Bourse	750000002	75102	2eme Ardt	750001537	9.911537e+05	4554.104360	48.868279	2.342803
2	17	Batignolles-Monceau	750000017	75117	17eme Ardt	750001537	5.668835e+06	10775.579520	48.887327	2.306777
3	7	Palais-Bourbon	750000007	75107	7eme Ardt	750001537	4.090057e+06	8099.424883	48.856174	2.312188
4	4	Hotel-de-Ville	750000004	75104	4eme Ardt	750001537	1.600586e+06	5420.908434	48.854341	2.357630

```
# Clean up the dataset to remove unnecessary columns

paris.drop(['NSQAR','CARINSEE','NSQCO','SURFACE', 'PERIMETRE' ], axis=1, inplace=True)
paris
```

	Dist_Num	Neighborhood	French_Name	Latitude	Longitude
0	3	Temple	3eme Ardt	48.862872	2.360001
1	2	Bourse	2eme Ardt	48.868279	2.342803
2	17	Batignolles-Monceau	17eme Ardt	48.887327	2.306777
3	7	Palais-Bourbon	7eme Ardt	48.856174	2.312188
4	4	Hotel-de-Ville	4eme Ardt	48.854341	2.357630
5	8	elysee	8eme Ardt	48.872721	2.312554
6	18	Buttes-Montmartre	18eme Ardt	48.892569	2.348161

Goal

- **Provide recommendations and results based on the data analysis which will help the fashion retailer setup new stores that yield high profits**

- **Discussion of any limitations and how the results can be used, and any conclusions that can be drawn.**