NEW STORE LOCATION FOR FASHION RETAILER

LOCATIONS FOR NEW FASHION STORES IN HIGH TRAFFIC AREAS FOR PARIS

TASK ON HAND

- Data-driven decisions on the new locations that are most suitable for their new stores

- The ground qualitative analysis of districts once the data and report are reviewed and studied.

- They do not seek stores in the premium upmarket strips, 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 suggests that the best locations to open new fashion retail stores may not only be where other clothing is located, rather the best places are in fact areas that are near *French Restaurants, Cafés and Wine Bars.*

- 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.

OUTCOME

- The goal is to identify the best districts 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.

DATA RESEARCH AND PREPARATION

- **Import the Paris District Data**

  - Paris is divided into 20 administrative districts. They and normally referenced by the arrondissement number rather than a name.

  - Data for the districts is necessary to select the most suitable of these areas for new stores.

  - After much research, this data is available on the web and can be manipulated and cleansed to provide a meaningful dataset to use

  - Data from Open|DATA: https://opendata.paris.fr/explore/dataset/arrondissements/table/?dataChart

  - Data From Opendatasoft: https://data.opendatasoft.com/explore/dataset/arrondissements%40parisdata/export/

EXPLORING, WRANGLING AND CLEANING THE DATA

Rename the columns 'Geometry_X' and 'Geometry_Y', "CAR' to 'Arrondissement' etc...

```
# Rename the necessary columns 'Geometry_X and Geometry_Y' etc...

# District : name of the central District for the Arrondissement
# Arrondissement : the Arrondissement or district number which is used to identify it
# Arrondissement_Fr : the descriptive French label for each Arrondissement

paris.rename(columns={'NAME': 'Neighborhood ', 'CAR': 'Arrondissement_Num', 'Geometry_X': 'Latitude', 'Geometry_Y': 'Longitude',
'LAR': 'French_Name'}, inplace=True)
paris
```

	Arrondissement_Num	Neighborhood	NSQAR	CAR.1	CARINSEE	French_Name	NSQCO	SURFACE	PERIMETRE	Latitude	Longi
0	3	Temple	750000003	3	3	3eme Ardt	750001537	1170882828	4519264	48.862872	2.3600
1	19	Buttes-Chaumont	750000019	19	19	19eme Ardt	750001537	6792651129	11253182	48.887076	2.3848
2	14	Observatoire	750000014	14	14	14eme Ardt	750001537	5614877309	10317483	48.829245	2.3265
3	10	Entrepot	750000010	10	10	10eme Ardt	750001537	2891739442	6739375	48.876130	2.3607
4	12	Reuilly	750000012	12	12	12eme Ardt	750001537	16314782637	24089666	48.834974	2.4213
5	16	Passy	750000016	16	16	16eme Ardt	750001537	16372542129	17416110	48.860392	2.2619
6	11	Popincourt	750000011	11	11	11eme Ardt	750001537	3665441552	8282012	48.859059	2.3800
7	2	Bourse	750000002	2	2	2eme Ardt	750001537	991153745	4554104	48.868279	2.3428
8	4	Hotel-de-Ville	750000004	4	4	4eme Ardt	750001537	1600585632	5420908	48.854341	2.3576
9	17	Batignolles-Monceau	750000017	17	17	17eme Ardt	750001537	5668834504	10775580	48.887327	2.3067

Clean up the dataset to remove unnecessary columns

```
# Clean up the dataset to remove unnecessary columns.
# Some of the columns are for mapping software - not required here.

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

	Arrondissement_Num	Neighborhood	French_Name	Latitude	Longitude
0	3	Temple	3eme Ardt	48.862872	2.360001
1	19	Buttes-Chaumont	19eme Ardt	48.887076	2.384821
2	14	Observatoire	14eme Ardt	48.829245	2.326542
3	10	Entrepot	10eme Ardt	48.876130	2.360728
4	12	Reuilly	12eme Ardt	48.834974	2.421325
5	16	Passy	16eme Ardt	48.860392	2.261971
6	11	Popincourt	11eme Ardt	48.859059	2.380058
7	2	Bourse	2eme Ardt	48.868279	2.342803
8	4	Hotel-de-Ville	4eme Ardt	48.854341	2.357630
9	17	Batignolles-Monceau	17eme Ardt	48.887327	2.306777
10	18	Buttes-Montmartre	18eme Ardt	48.892569	2.348161
11	1	Louvre	1er Ardt	48.862563	2.336443

METHODOLOGY AND EXPLORATORY DATA ANALYSIS

- **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 visualization and statistical analysis.
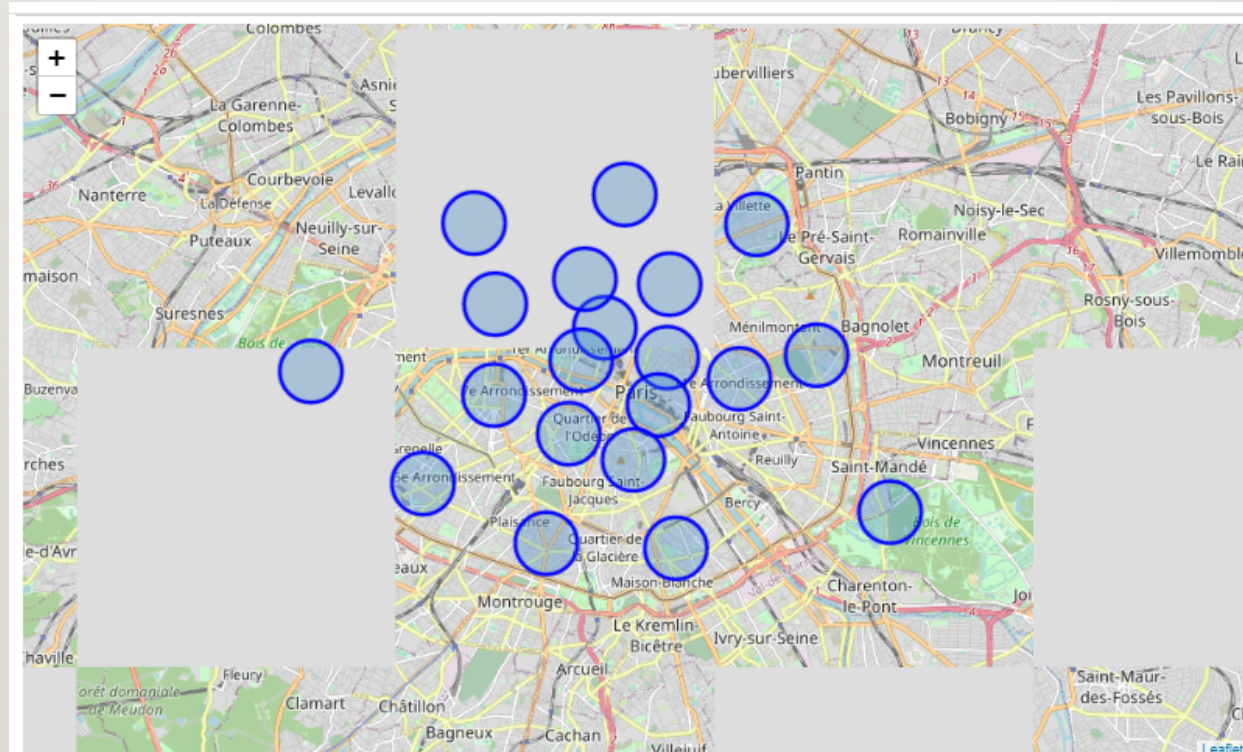
- **Visualization:**
  - Analysis and plotting visualizations.
  - Data visualization using various mapping libraries.

- **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.

# CREATE A MAP WITH DISTRICTS SUPERIMPOSED

DATA ANALYSIS

- **Use the Foursquare API to explore the Districts (Neighborhoods)**
  - Define Foursquare Credentials and Version

- **Exploratory data analysis**
  - Explore the first district in our data frame to become familiar with the data.
  - The first district is identified as *3eme Ardt*
  - Get the top 100 venues that are in the neighborhood *3eme Ardt* within a radius of 500 meters
  - Define the function that extracts the category of the venue
  - Structure the json file into a pandas data frame
  - Create a nearby venues function for all the neighborhoods in Paris
  - Create a new data frame called for the venues of Paris called *paris-venues*

THE BUSINESS TYPES CRITERIA SPECIFIED BY THE CLIENT! *'RESTAURANTS', 'CAFÉS' & 'WINE BARS'*

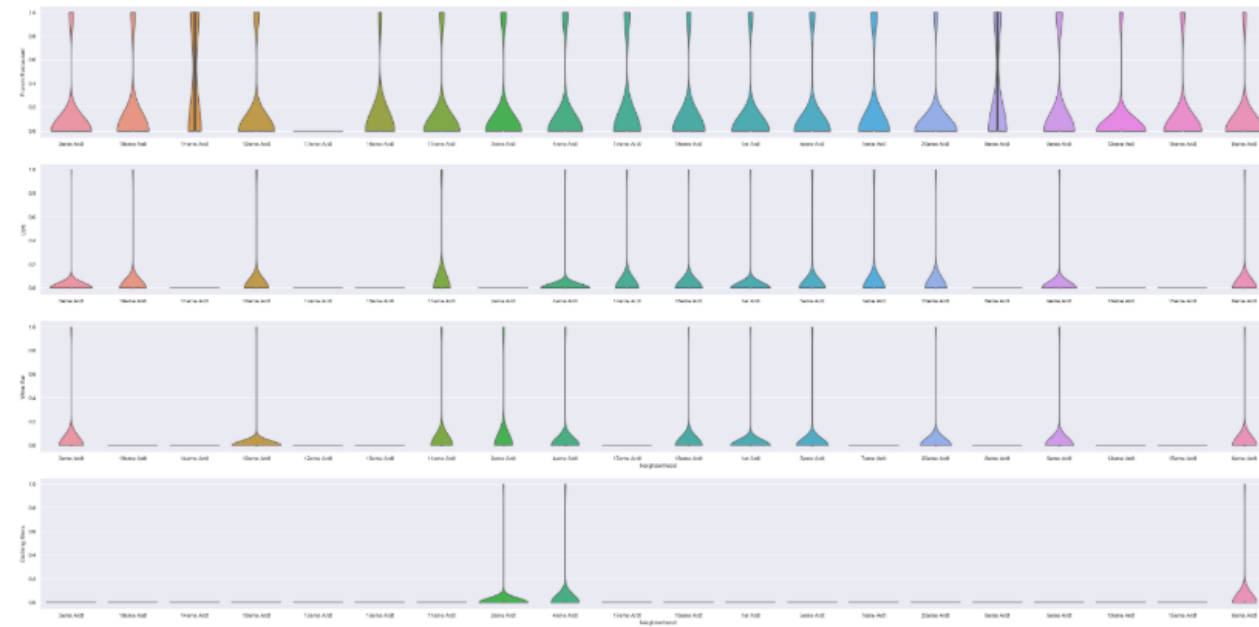- Frequency distribution for top 3 venue categories for each neighborhood

FREQUENCY OF CLOTHING STORE

Frequency of Clothing stores for each neighborhood
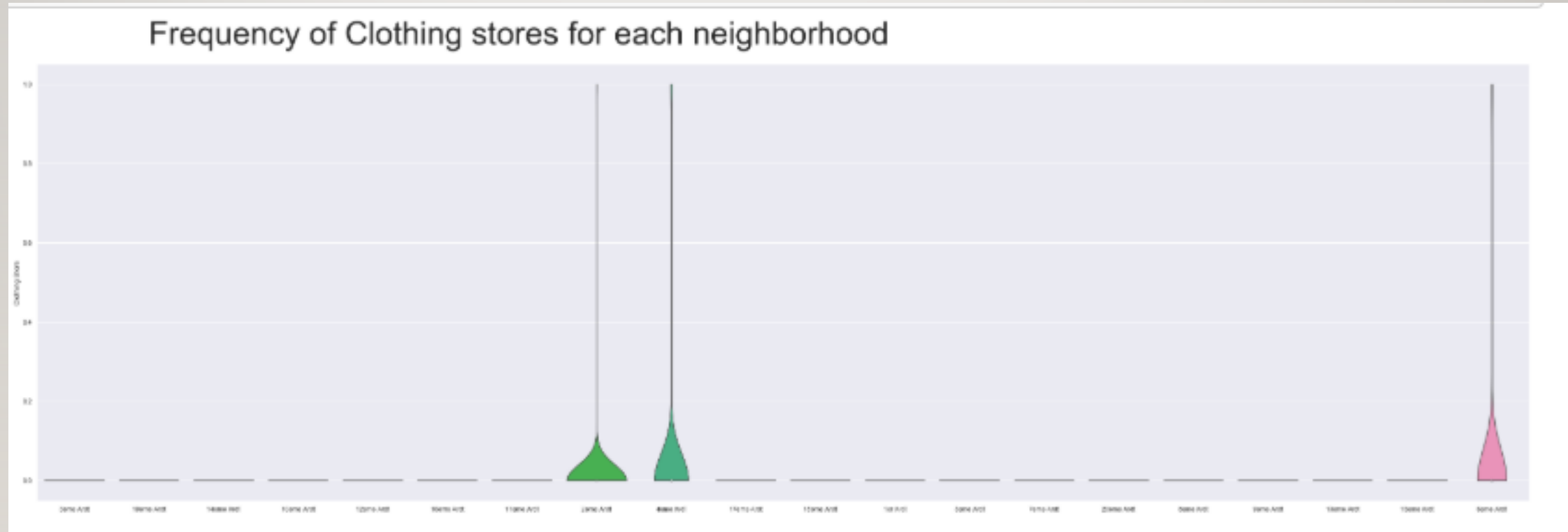
FREQUENCY DISTRIBUTION FOR TOP 3 CATEGORIES INCLUDING CLOTHING

Frequency distribution for the top 3 venue categories for each neighborhood (includes clothing)

CHOSEN NEIGHBORHOODS - RESULTS

- Inferential analysis using the data, as well as domain knowledge of retail and marketing, allow the list to be focused to just 3 neighborhoods.

- The reasoning being that if the 3 criteria have been met - identifying neighborhoods that are lively with Restaurants, Cafés and Wine Bars - adding Clothing Stores into the mix of stores in the area is a significant bonus. Having some of the same category of stores in the same area - especially in fashion retail - is very desirable as a retailer.

- So we can increase the criteria to include *Restaurants, Cafés, Wine Bars and Clothing Stores* - which narrows down and focuses the suggested districts for new stores to be located, and at the same time provides better locations for the brand.

- So the final 3 prospective neighborhoods for new store locations are where 4 criteria are met

- **3eme Ardt : Arrondissement 3, Temple**

- **- 4eme Ardt : Arrondissement 4, Hotel-de-Ville**

- **- 6eme Ardt : Arrondissement 6, Luxembourg**

LET'S LOOK AT THE 3 DISTRICTS ON A PARIS MAP

FOR A CLOSER VIEW OF CENTRAL PARIS

CONCLUSIONS

- There are many ways this analysis could have been performed based on different methodology and perhaps different data sources.

- I selected the above as it was a straight forward way to narrow down the options, not complicating what is actually simple in many ways – meeting the criteria for the surrounding venues.

- The analysis and results are not an end point, but rather a starting point that will guide the next part of the process to find specific store locations.

- The next part will involve domain knowledge of the industry, and perhaps, of the city itself. But the data analysis and resulting recommendations have greatly narrowed down the best district options based on data and what we can infer from it.

- Without leveraging data to make focused decisions, the process could have been drawn out and resulted in new stores opening in sub-standard areas for this retailer.

- Data has helped to provide a better strategy and way forward, these data-driven decisions will lead to a better solution in the end.