The Battle of two cities

A comparative study of popular venues in New York City and London

The Problem

Tourist’s Selection:
Choice of places to visit for a tourist

- As a first-time visitor, one would be interested in knowing about the popular venues around

- Selection of venues is difficult especially in a time-crunch

Solution

Exploratory Analysis of the locations of interest

Using Foursquare Location API with Python

- *Foursquare* gives a useful insight of the popular venues of your preferred location

- It’s a reliable application in terms of location-specific information like user reviews, location tips and much more

Implementation

Fetch Geolocation parameters (latitude, longitude) of the location and access the *Foursquare* location API to **get info of top 100 venues**

Visualize the analysis using libraries like:
Folium (for geographically visualizing the venues),
Seaborn (for visualizing statistical calculations)

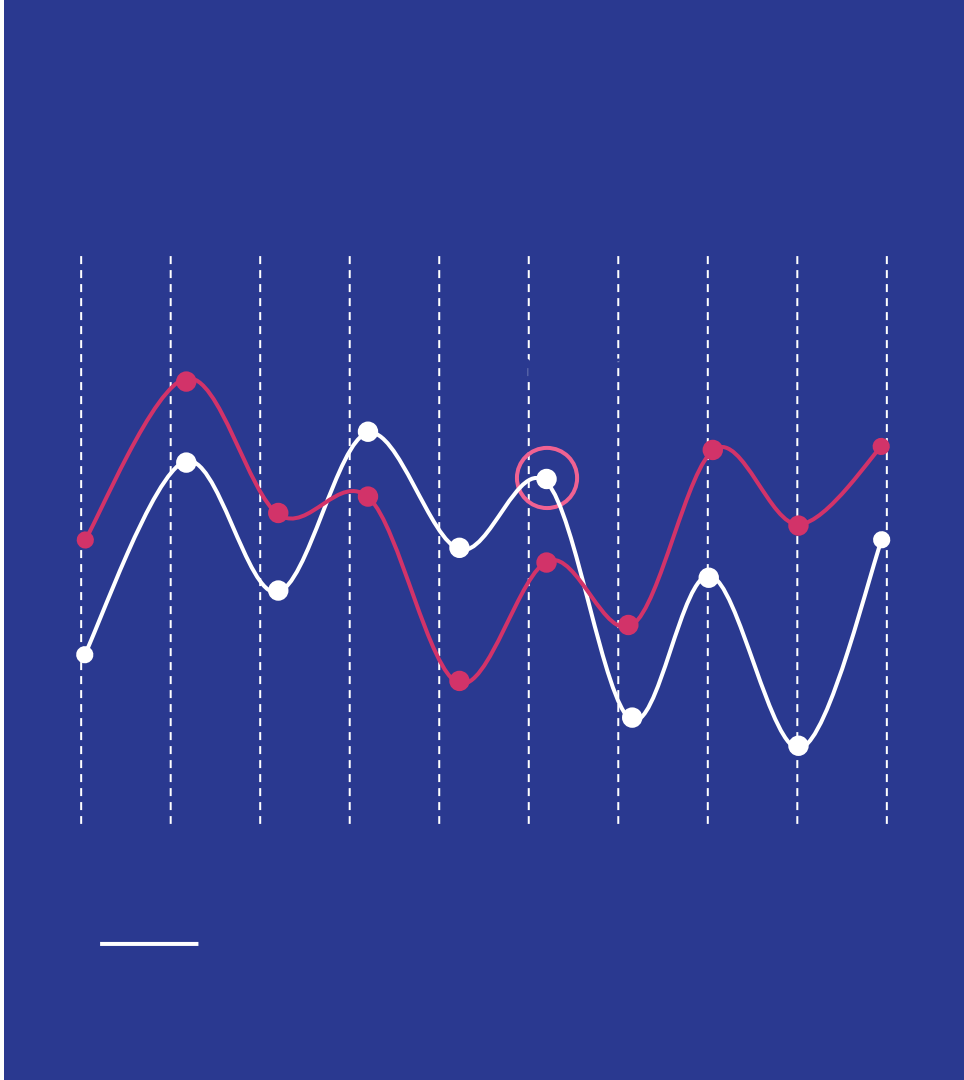
Draw **conclusion** on the information collected from this analysis

For both the cities, **group the venues** according to their categories and **gather statistical information** (e.g. frequency of occurrence, Mean of these frequencies, and K-means clustering)

For a **comparative analysis**, combine the data of the two cities on the basis of their *common venue categories*, and collect statistical information for this combined collection

Observations

Sample observations recorded for the analysis of a sample of ***100*** venues

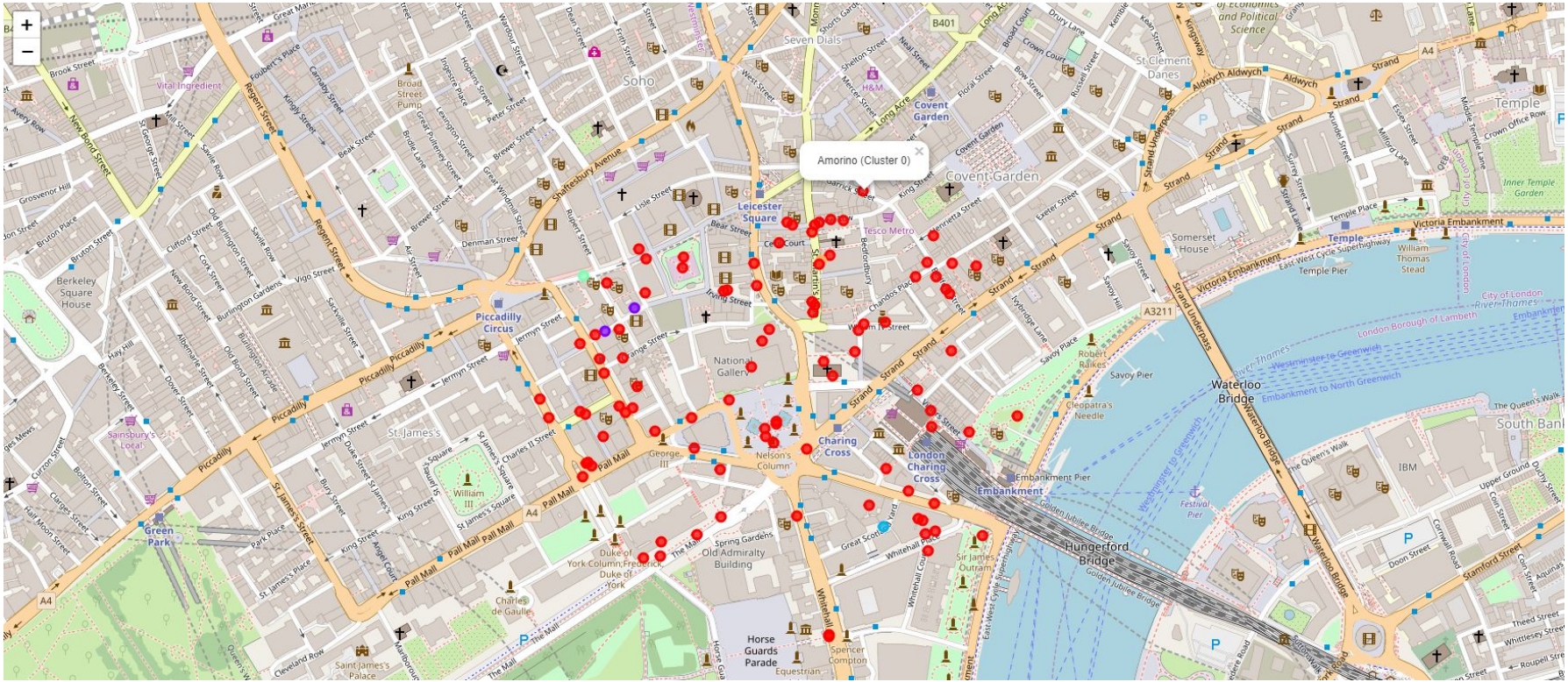


The Leaflet slides showcase the clusters of popular venues in both the cities

Leaflet view (Folium) of NYC

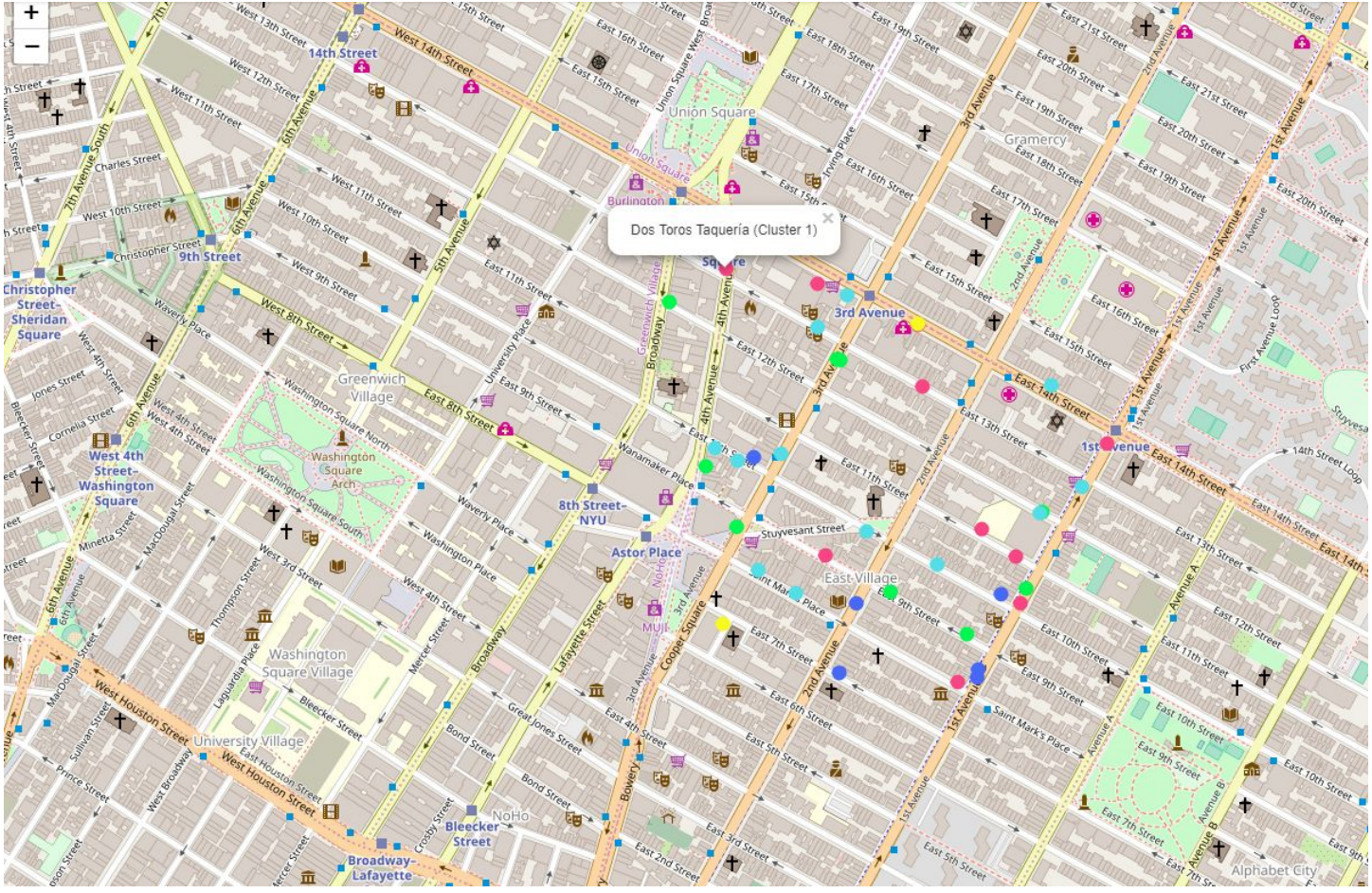
The Popular Venues in New York City - clustered using K-means algorithm

Leaflet view (Folium) of London



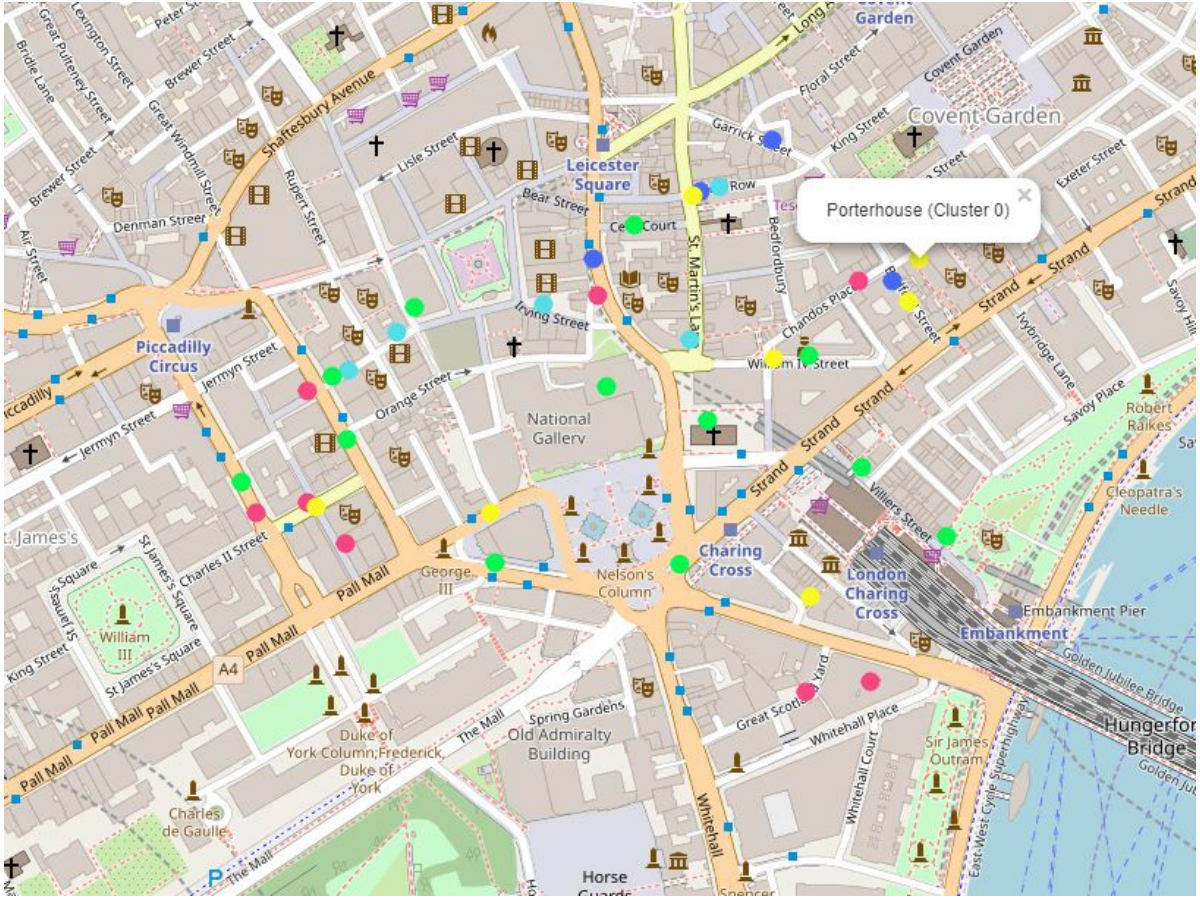
The Popular Venues in the city of London - clustered using K-means algorithm

Leaflet view (Folium) of Common Venue Categories (NYC)



The venues displayed here share common veue categories with the venues in London.

Leaflet view (Folium) of Common Venue Categories (London)

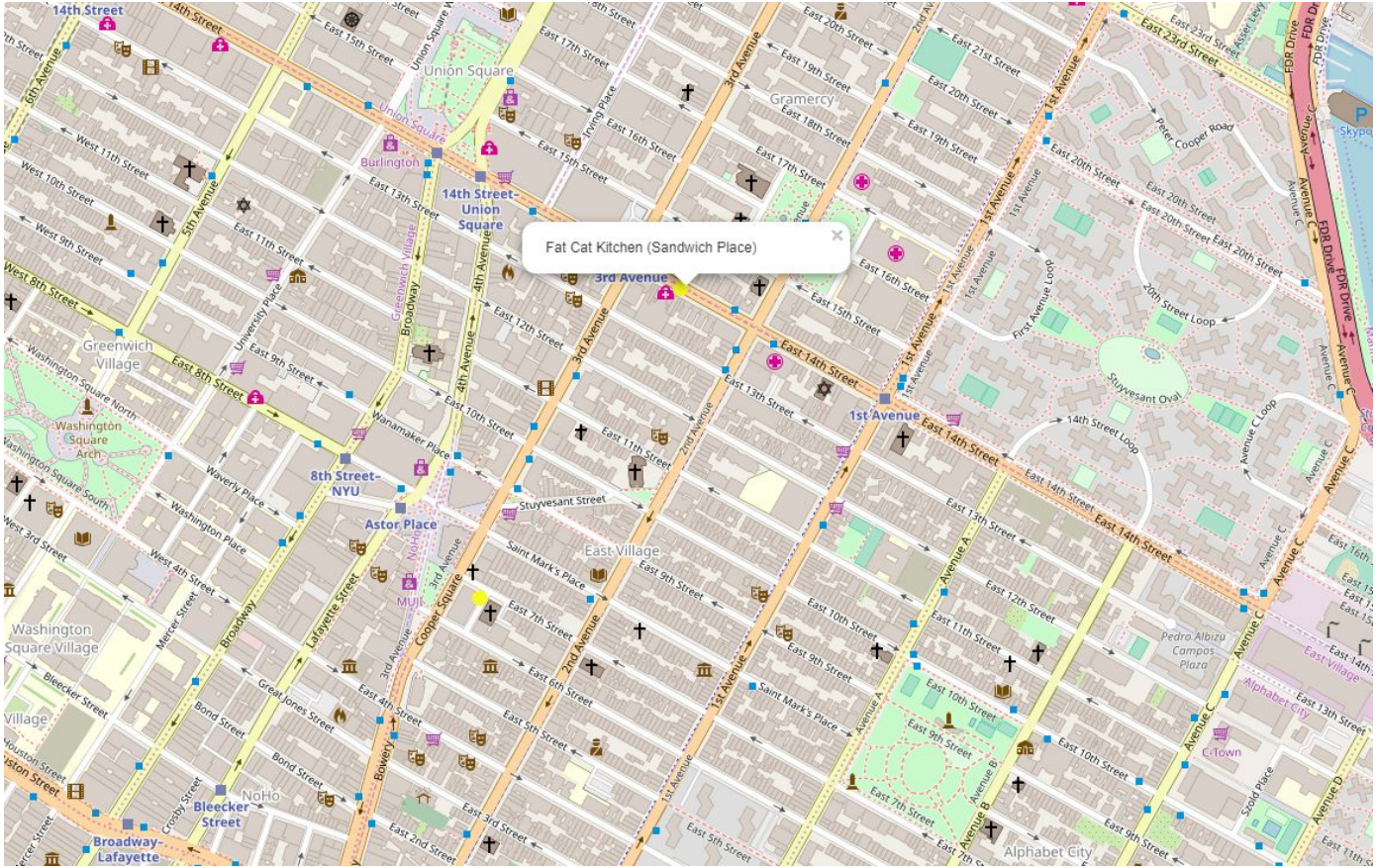


The venues displayed here share common veue categories with the venues in New York City.

Cluster Analysis (New York City)

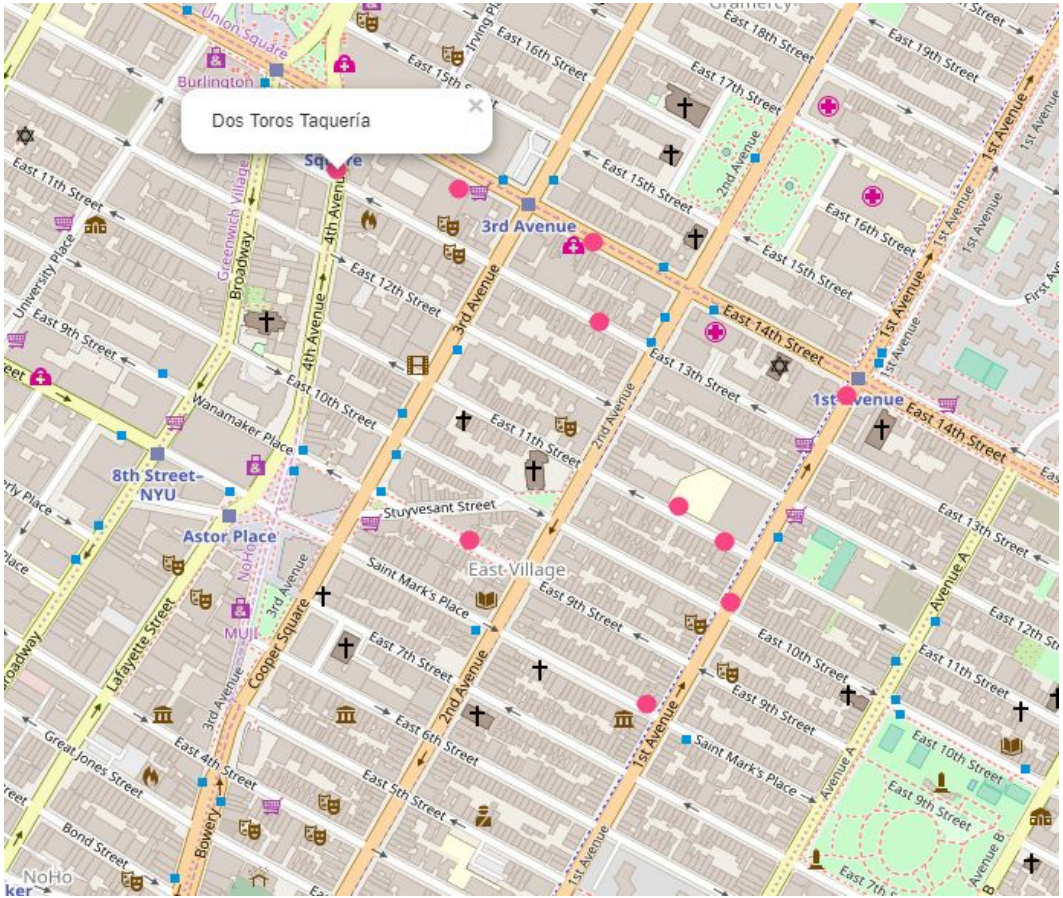
The following gives us an insight of the venues in NYC, commonly categorized with same type of venues in London

Map of Clustered Common Venues - New York City



- The venues displayed here represent the **first** cluster of common -categorized venues in New York City
- Cluster information: There are **2 categories**, each having **one venue**, a Sandwich place and a Bar, as see here

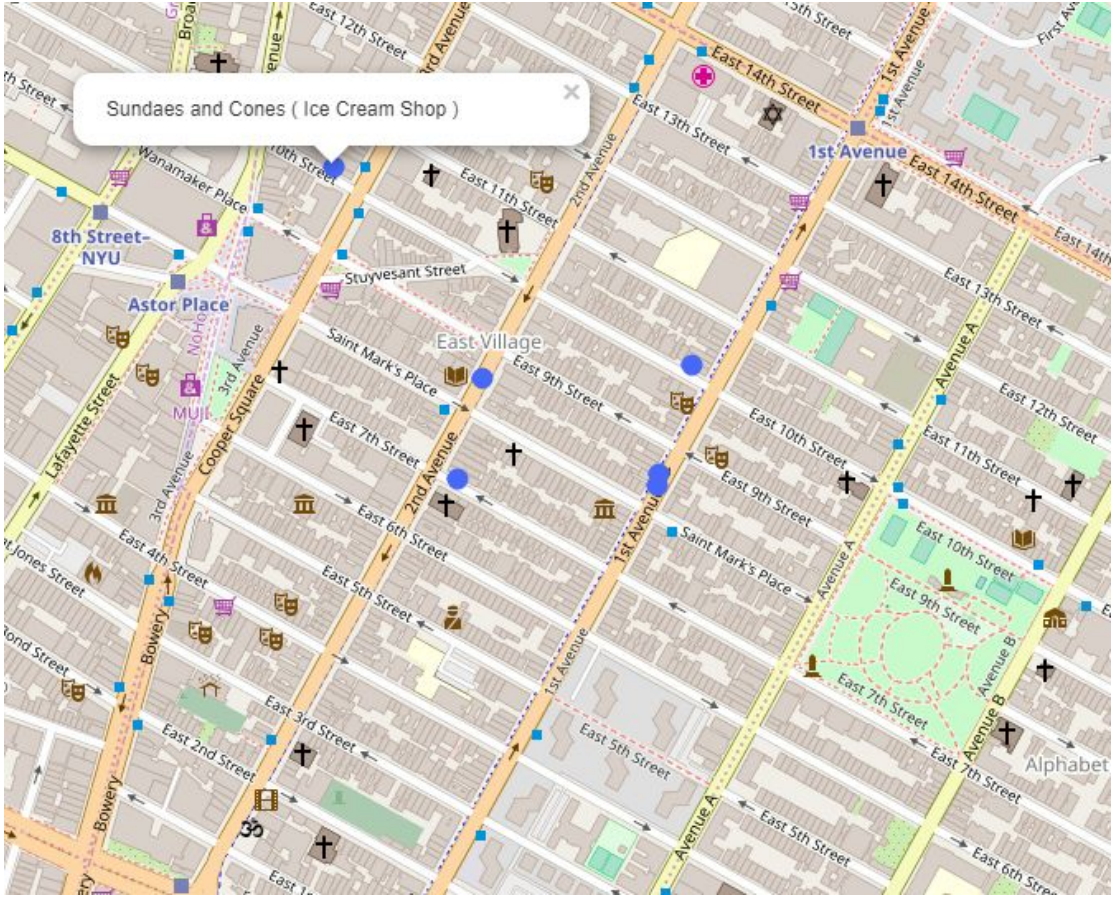
Map of Clustered Common Venues - New York City



- The venues displayed here represent the **second** cluster of common-categorized venues in New York City

- Cluster information:
  There are **8 categories** here, with a total of **10 venues**, mostly comprising of Bars, Wine Shops and Taqueria, along with 2 restaurants, one tea room and bakery

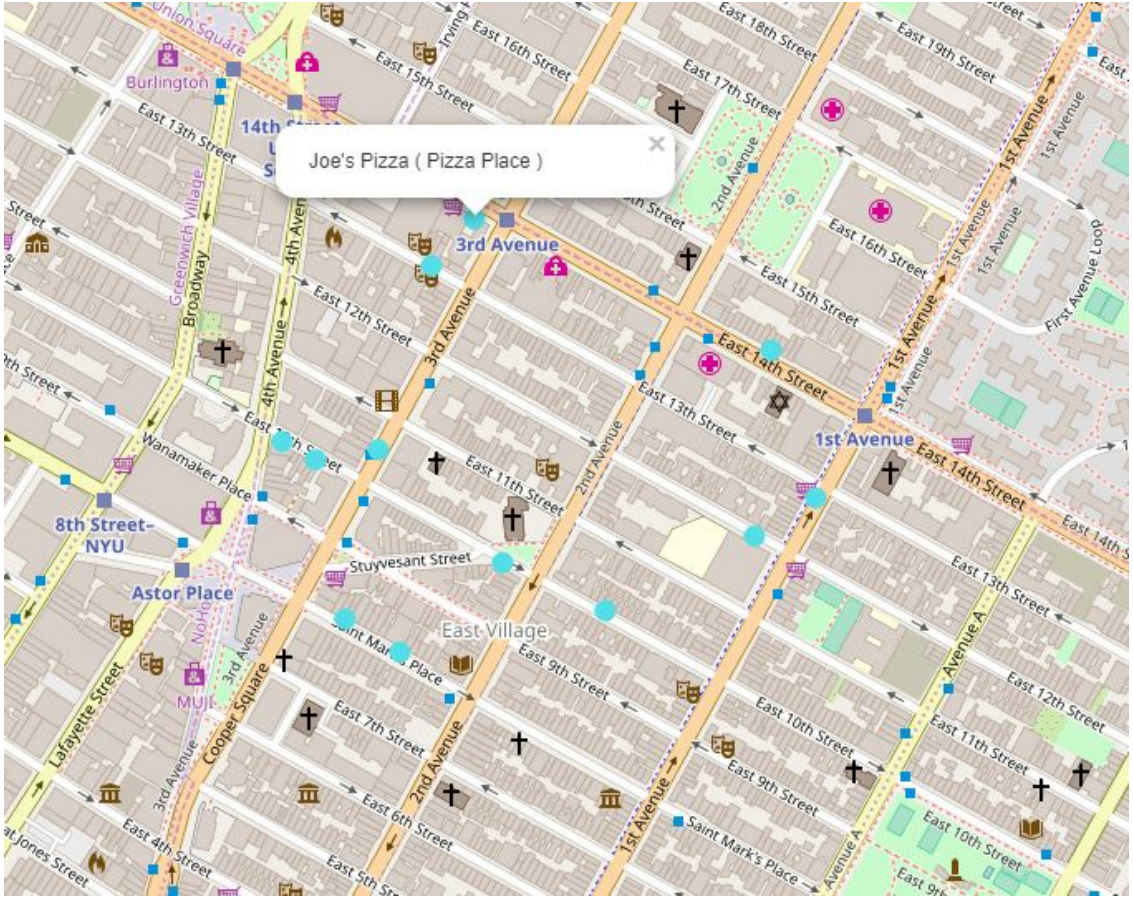
Map of Clustered Common Venues - New York City



- The venues displayed here represent the **third** cluster of common-categorized venues in New York City

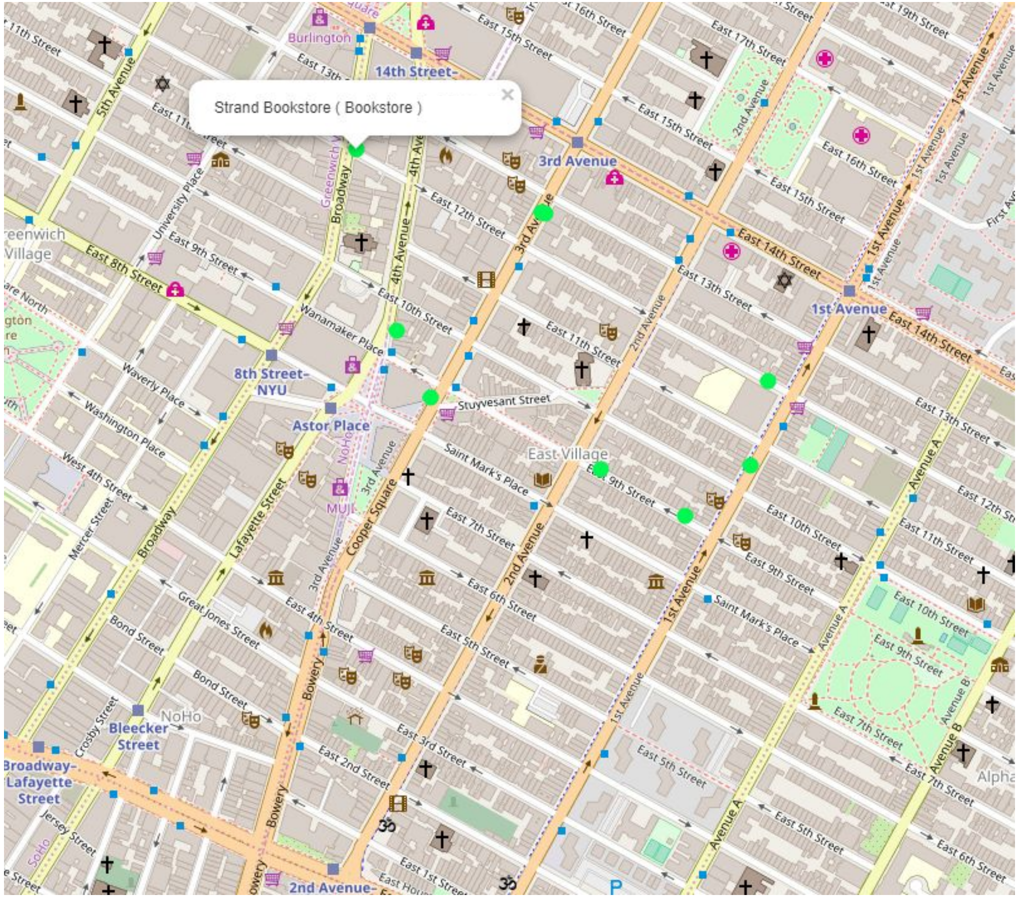
- Cluster information:
  There is **1 category** here, (*Ice-cream Shop*) with a total of **6 venues**, spread across New York City

Map of Clustered Common Venues - New York City



- The venues displayed here represent the **fourth** cluster of common-categorized venues in New York City
- Cluster information:
  There are **3 categories** here, (*Coffee Shop, Pizza Place, and Japanese Restaurant*) with a total of **12 venues**, spread across New York City

Map of Clustered Common Venues - New York City



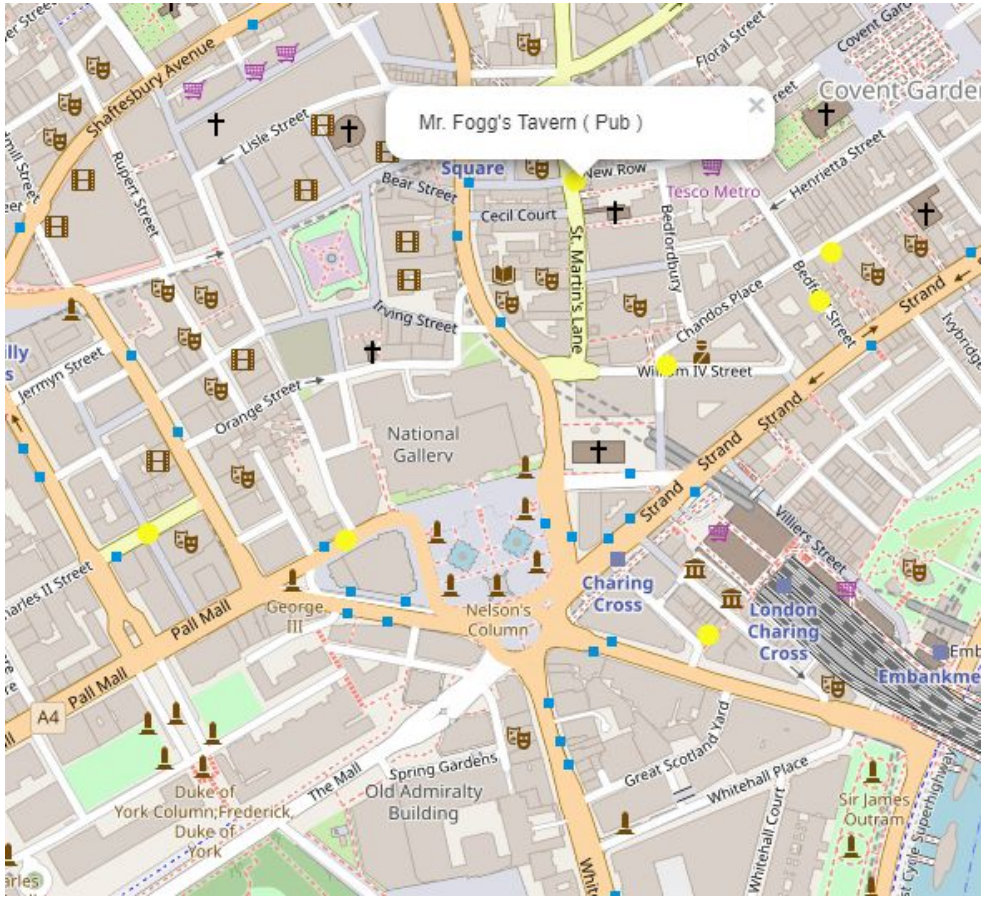
- The venues displayed here represent the **fifth** cluster of common-categorized venues in New York City

- Cluster information:
There are **6 categories** here, with a total of **9 venues**, spread across New York City

Cluster Analysis (London)

The following gives us an insight of the venues in London, commonly categorized with same type of venues in London

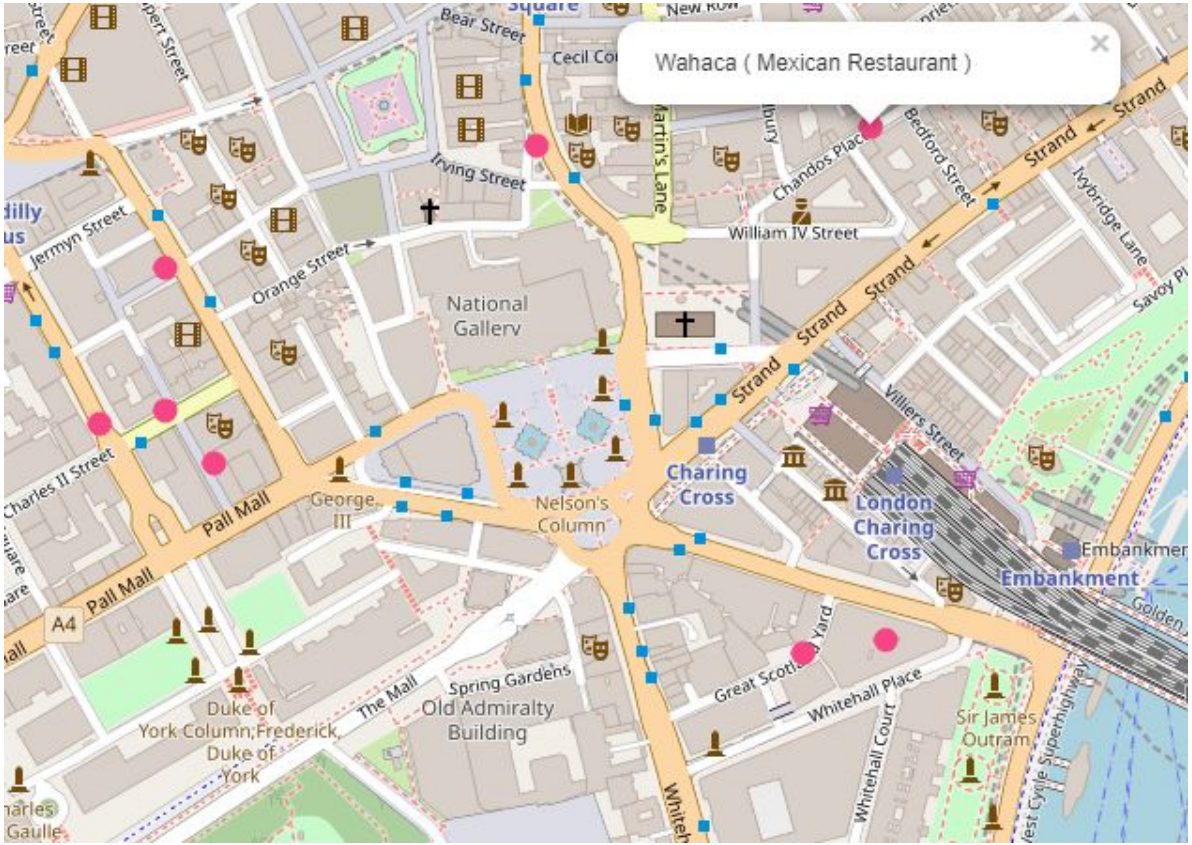
Map of Clustered Common Venues - London



- The venues displayed here represent the **first** cluster of common -categorized venues in London

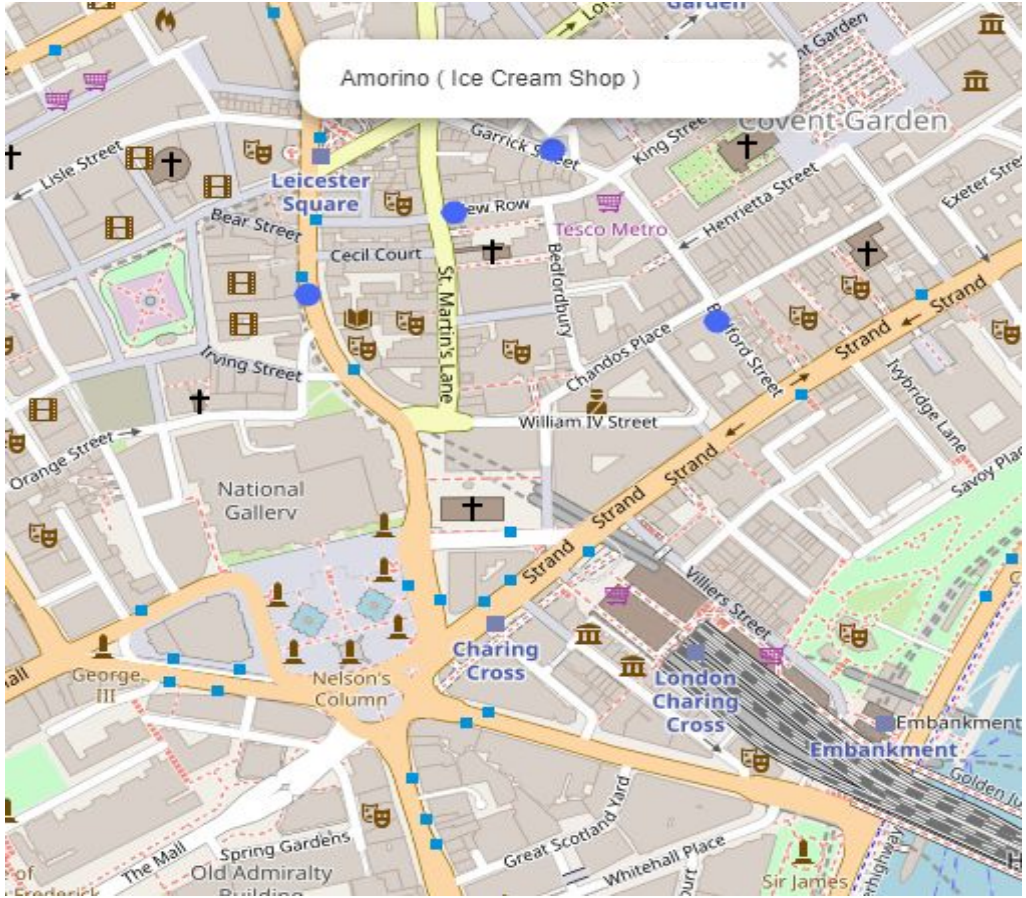
- Cluster information:
  There are **2 categories (***Pub, and Sandwich Place***)**, with a total of **7 venues -** the pubs being clustered in the right cluster, as seen here

Map of Clustered Common Venues - London



- The venues displayed here represent the **second** cluster of common-categorized venues in London
- Cluster information:
  There are **8 categories** here, with a total of **8 venues**, with a bar and wine shop clustered together with a Greek Restaurant and a Bakery

Map of Clustered Common Venues - London

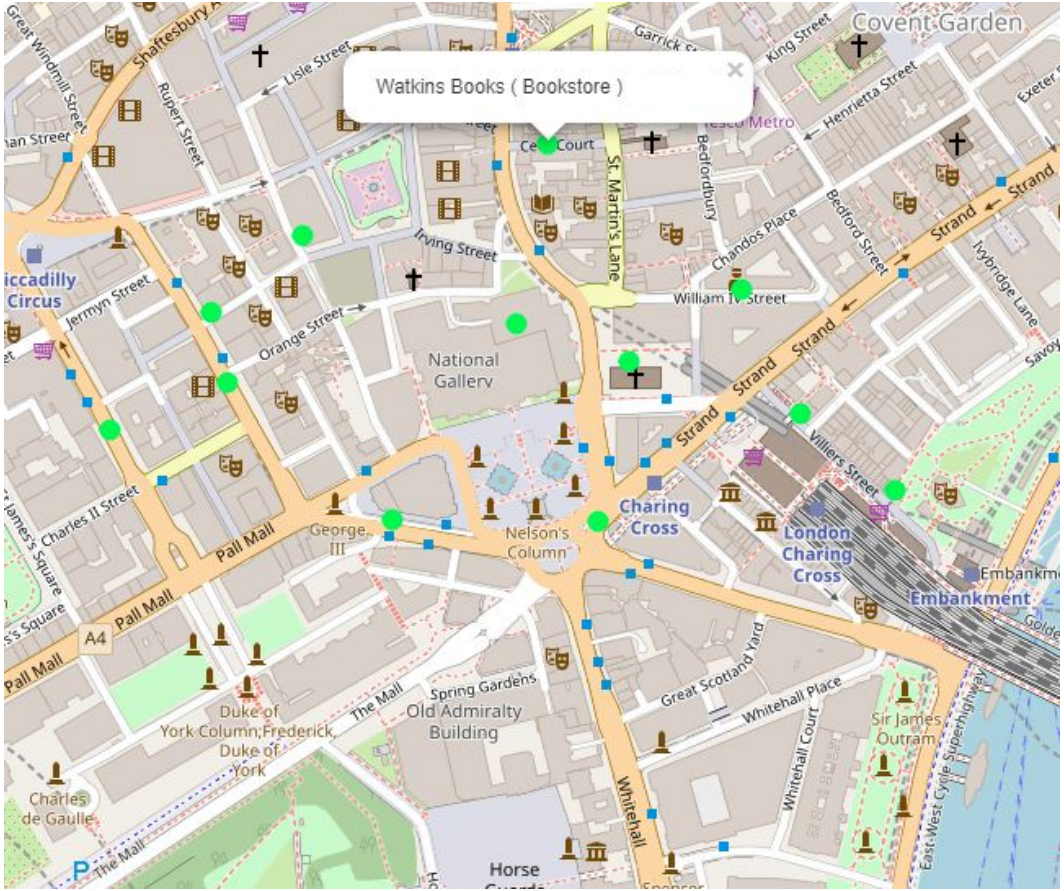


- The venues displayed here represent the **third** cluster of common-categorized venues in London
- Cluster information:
  There is **1 category** here, (*Ice-cream Shop*) with a total of **4 venues**

Map of Clustered Common Venues - London

- The venues displayed here represent the **fourth** cluster of common-categorized venues in London
- Cluster information:
  There are **3 categories** here, (*Coffee Shop, Pizza Place, and Japanese Restaurant*) with a total of **5 venues**, spread across London, as seen here

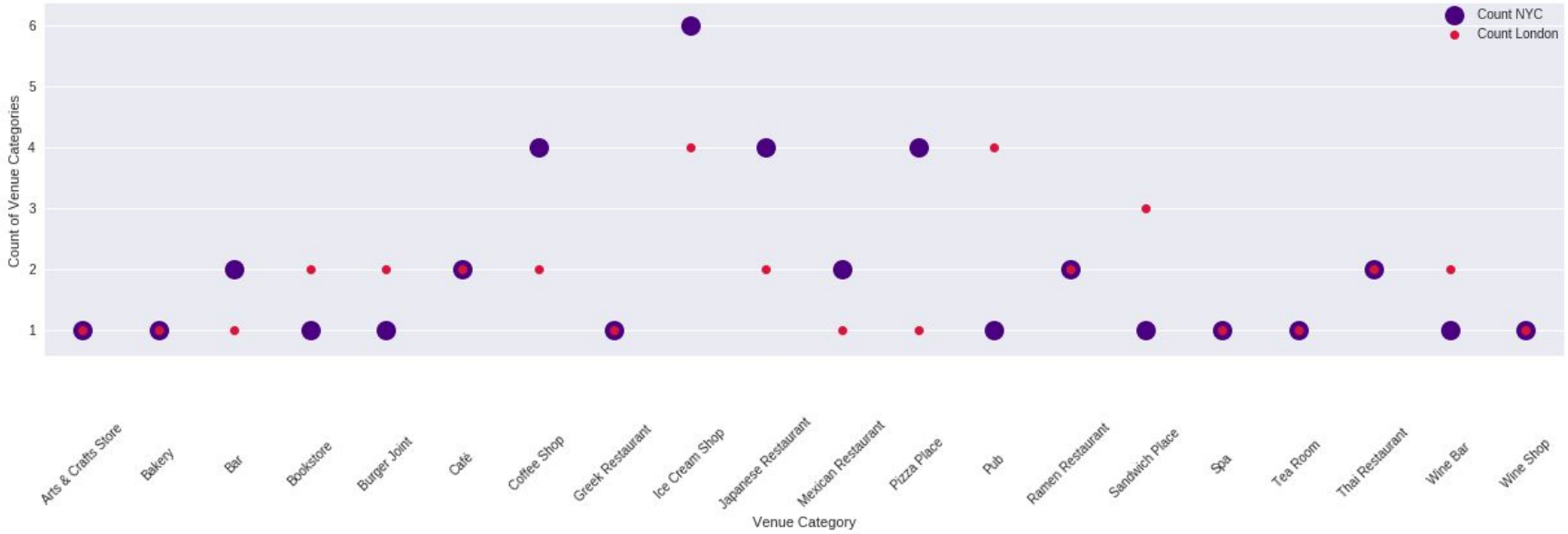
Map of Clustered Common Venues - London



- The venues displayed here represent the **fifth** cluster of common-categorized venues in London
- Cluster information:
  There are **6 categories** here, with a total of **12 venues**, spread across London

The next few slides focus on individual counts of venue categories in both the cities

Frequency of occurrence of venues in both the cities

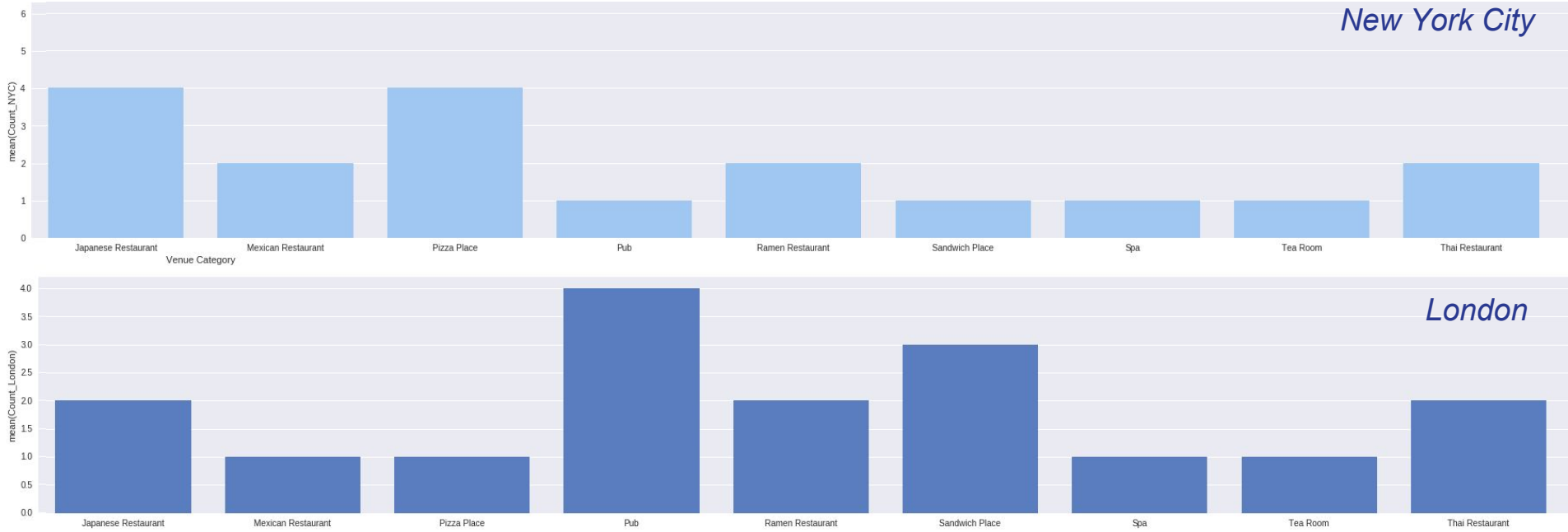


This strip-plot shows the count of each venue category common in both the cities.

- The number of ice-cream shops rank the highest (in NYC)
- For London, the highest occurrences of venues were Pubs and ice-cream shops
- There's a higher probability of stumbling across a **pub in London** than a bar, while you are highly likely of hitting **coffee shops and pizza joints in NYC**

Mean of the frequency of occurrence of venues in both the cities (*Part 1*)

Mean of the frequency of occurrence of venues in both the cities (*Part 2*)



Mean of the frequency of occurrence of venues in both the cities (*Part 3*)

Inferences

The conclusions drawn from the *geospatial* and *statistical* analysis of the sample data-set

The geospatial and statistical analyses are done on the clusters of the venues, in the aforementioned set of data, which gives us the following insights on the venues of these prime cities

New York City

- Highest number of Ice-cream shops - ***6*** around East Village
- ***1*** Pub and ***2*** Bars - spread across Astor Place and East Village
- ***4*** Pizza Places and ***4*** Coffee Shops, spread across East Village *(again!)*
- *Ramen and Thai restaurants* (***4***) are more likely to be found sandwiched between 1st and 3rd Avenue, East Village
- Unlikeliness, or the least number of occurrences (***1***) of common venue categories are
    - Arts and Crafts Store
    - Bakery
    - Book Store
    - Burger Joint
    - Greek Restaurant
    - Pub
    - Sandwich Place
    - Spa
    - Tea Room
    - Wine Shop

London

- ***4*** Ice-cream shops - around Leicester Square
- ***4*** Pubs and ***1*** Bar - spread across Leicester Square, London Charing Cross, and Piccadilly Circus
- Probability of finding *Pizza Places* (***1***) is lesser as compared to a *Bar* or *Pub*
- *Ramen and Thai restaurants* (***4***) are more likely to be found sandwiched between Leicester Square and Piccadilly Circus
- Unlikeliness, or the least number of occurrences (***1***) of common venue categories are
    - Bar
    - Bakery
    - Book Store
    - Burger Joint
    - Greek Restaurant
    - Mexican Restaurant
    - Pizza Place
    - Spa
    - Tea Room
    - Wine Shop