The battles of cities

Capstone Project

Tiziana Bassi

The problem

Context
We are in the age of easy travels, people move easly and quickly from one city to another one in a different country.

Problem statement
● How to choose where to go?   ● Which are the facilities you will find?

Analysis steps

Step 1	Step 2	Step 3

Retrieve the data

- Google API
- Foursquare API

Clean the data

- *Select the category* of the venues using the items key in the *json* file
- Split the main data frame in four sub-dataframe basing on the city where the venue is.

Analysis

- compare the kind of venue present in each city, identify the first 10 venues for each city.
- evaluate how many venues are linked to food and how many to culture.

Step 1
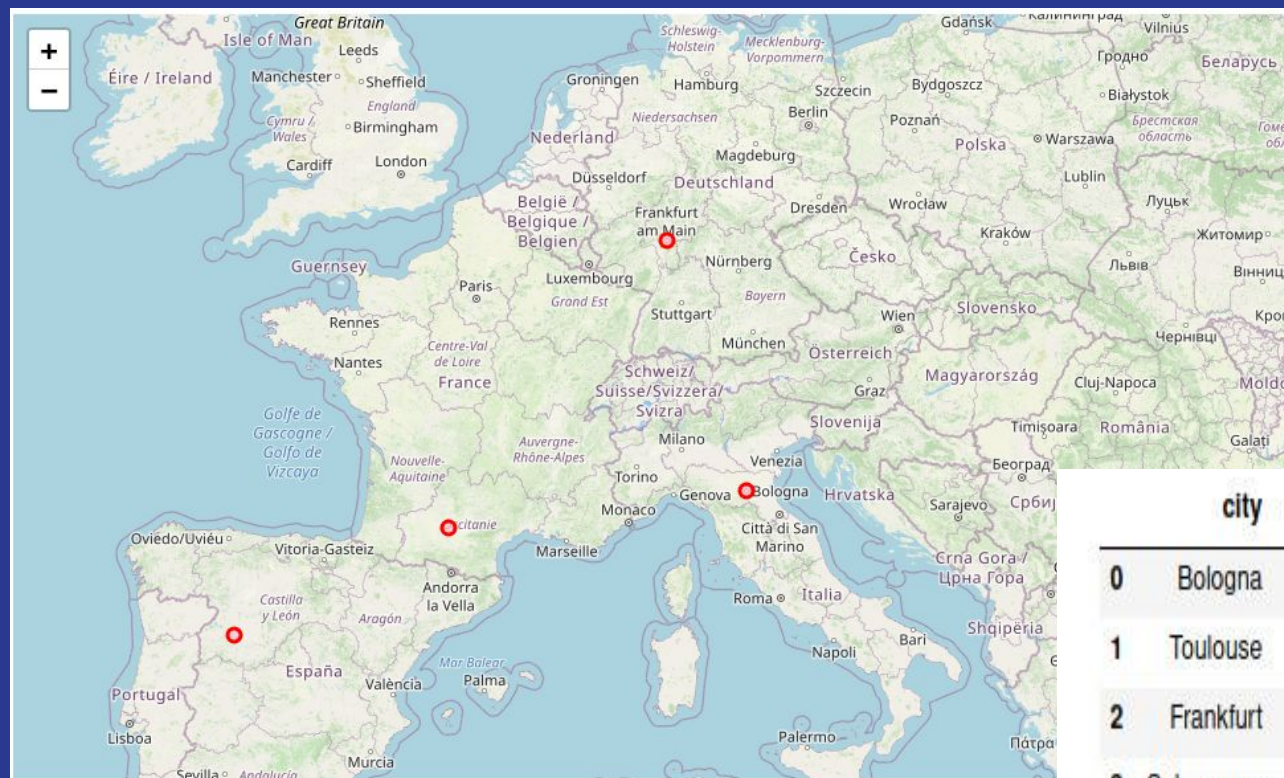
Retrieve the data

- **Google API Data:**

  Once the cities has been selected, we can retrieve google's geocoded data from *Google Geocoding API* using Geocoder. The process convert the addresses (in our case the format will be 'city name, country name') into geographic coordinates (e.g latitude and longitude).

- **Foursquare API Data:**

  In order to gain data about different venues in each cities we will use *Foursquare* locational information. *Foursquare* is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos.

-

	city	country	area	Latitude	Longitude
0	Bologna	IT	140.9	44.4949	11.3426
1	Toulouse	FR	118.3	43.6047	1.44421
2	Frankfurt	DE	248.3	50.1109	8.68213
3	Salamanca	ES	38.6	40.9701	-5.66354

Step 2

Clean the data

- **Google API Data:**

  Once the cities has been selected, we can retrieve google's geocoded data from *Google Geocoding API* using Geocoder. The process convert the addresses (in our case the format will be 'city name, country name') into geographic coordinates (e.g latitude and longitude).

- **Foursquare API Data:**

  In order to gain data about different venues in each cities we will use *Foursquare* locational information. *Foursquare* is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos.

-

Using the *Foursquare API* we obtain:

1. City
2. Venue
3. Name of the venue e.g. the name of a store or restaurant
4. Venue Latitude
5. Venue Longitude
6. Venue Category

	City	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Bologna	44.494887	11.342616	Piazza Nettuno	44.494685	11.342880	Plaza
1	Bologna	44.494887	11.342616	La Prosciutteria Firenze	44.496025	11.345934	Trattoria/Osteria
2	Bologna	44.494887	11.342616	Pescheria Pavaglione	44.493691	11.344831	Fish Market
3	Bologna	44.494887	11.342616	La Baita Formaggi	44.493746	11.344242	Cheese Shop
4	Bologna	44.494887	11.342616	Biblioteca Comunale dell'Archiginnasio	44.492041	11.343266	College Library

|	Venues	Bologna	Toulouse	Frankfurt	Salamanca
0	Art Gallery	0.01	0.00	0.00	0.000000
1	Art Museum	0.02	0.02	0.04	0.012821
2	Beer Bar	0.02	0.01	0.01	0.000000
3	Beer Garden	0.01	0.00	0.01	0.000000
4	Bistro	0.02	0.00	0.01	0.000000

1. Split the main data frame in four sub-dataframe basing on the city where the venue is.
2. Group the rows in the main dataframe by city and by taking the mean of the frequency of occurrence of each category.
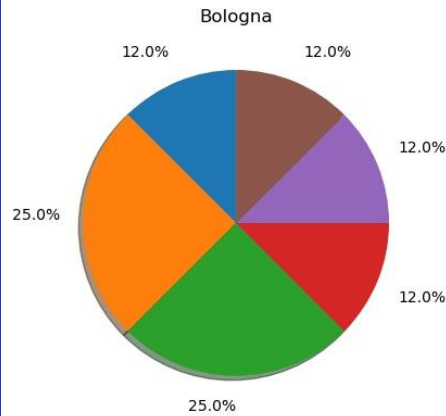
Step 3

Analysis

- Compare the kind of venue present in each city, identify the first 10 venues for each city.

- Evaluate how many venues are linked to food and how many to culture.

Venues distribution
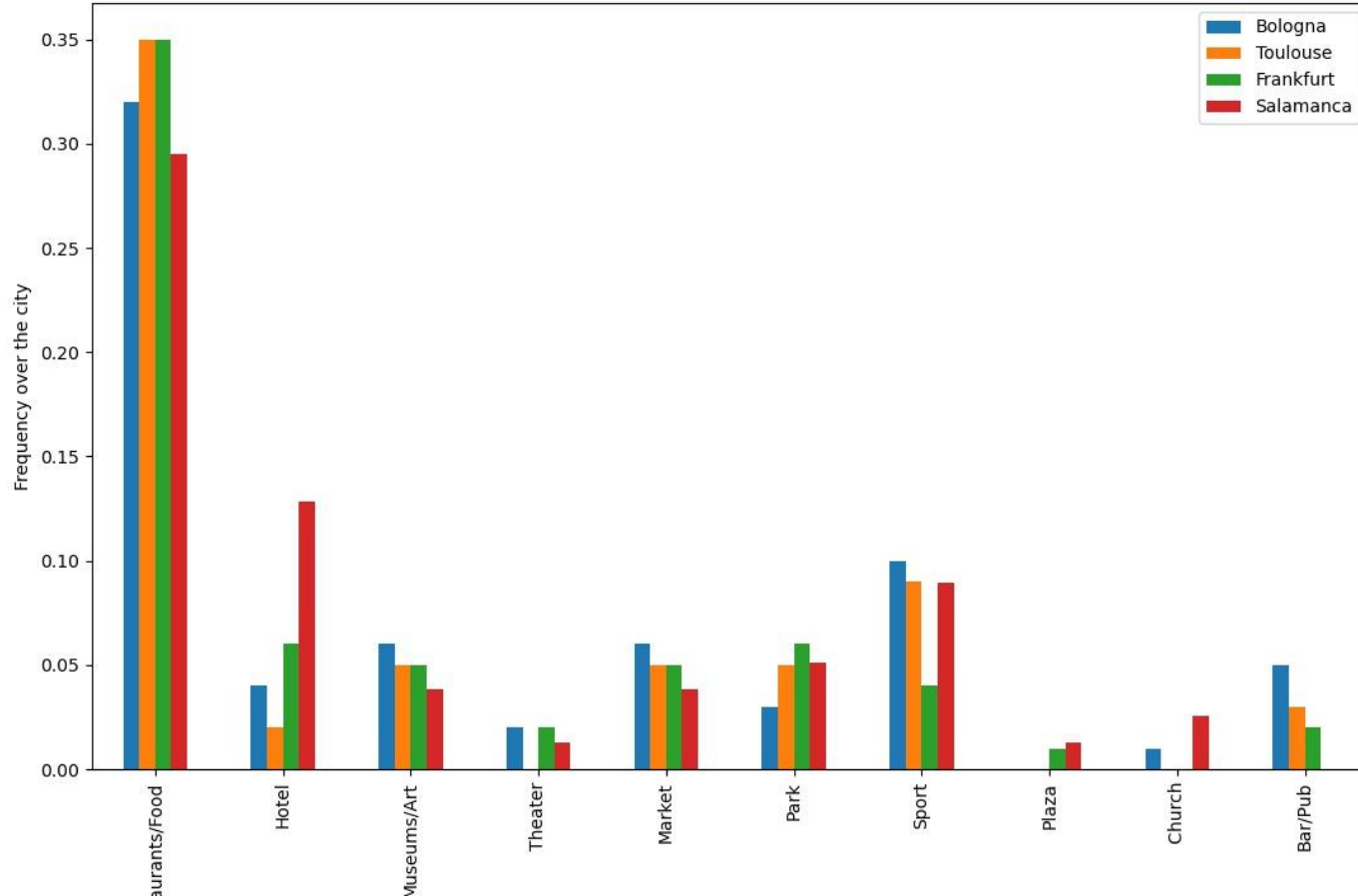
Category venues per city

The most frequent venues in the four cities concerning food. This is followed by the venues about sport ('gym', 'fitness places').
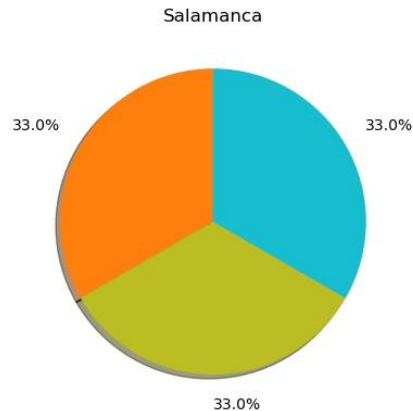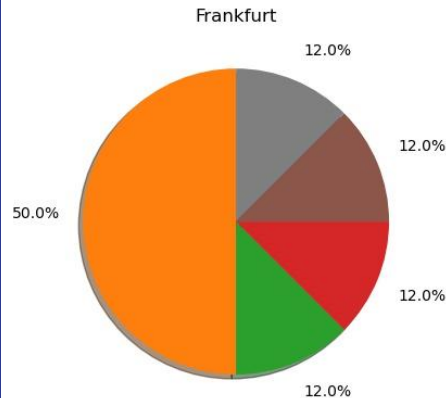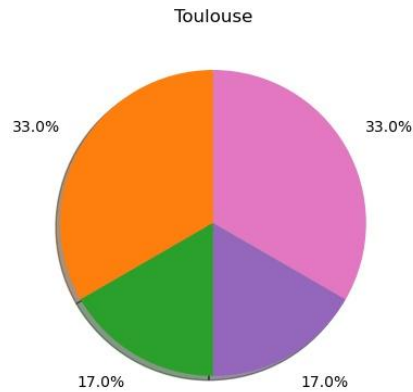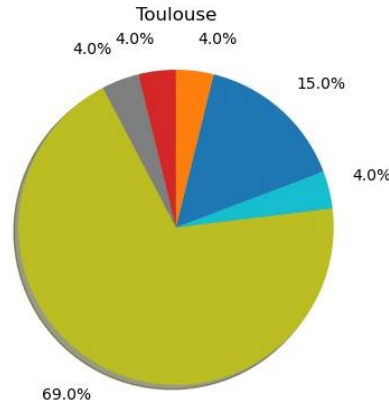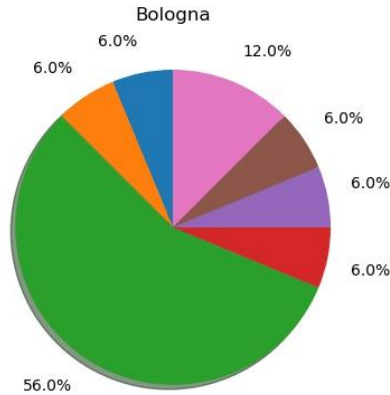
Culture

- One of the most frequent venue present in all the cities is the *Art Museum*
- The distribution of the venues shows that **Bologna** is a city with *Theaters*
- In **Toulouse** the preference is for the *Concert Hall*.
- In **Salamanca** the cultural venues are equally distributed between *Art Museum, Indie Theater, Historic Center*.

Restaurant

- **Bologna** and **Toulouse** are characterised by a dominant number of *Italian* and *French* restaurant.
- **Salamanca** even though *Spanish restaurant* are the most numerous, it is possible to find also different place where to eat something different.
- **Frankfurt** looks populated by different kind of restaurant that can satisfy all the tastes.
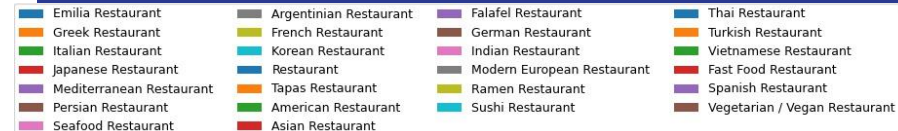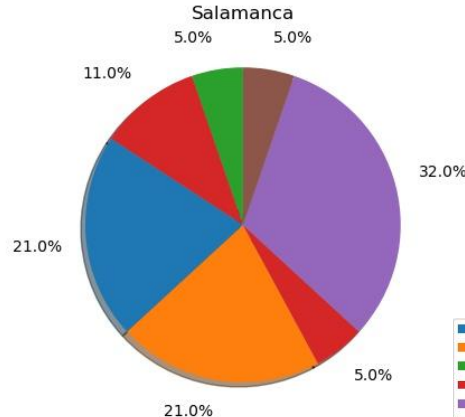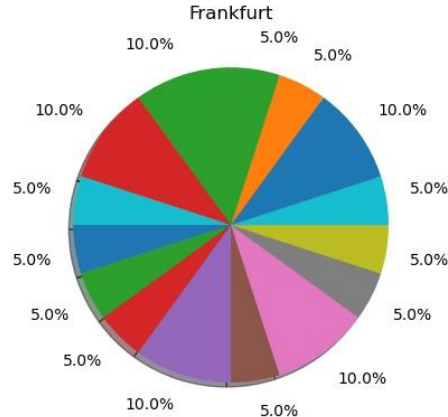
Conclusion

- People can have nice typical or exothic meal in all the cties analysed, no one will remain without food.
- All the cities offer *'Art Museum'* to entertain the curious people and in addition **Salamanca** offers also *'Historic Places'*.