The Battle of the Neighborhoods - Rome VS Paris

Thomas KLEIN
October 27, 2019

1. Introduction/Business Problem

1.1 Background

Rome and Paris are two major capitals of important European countries, respectively France and Italy. Both capitals are known to be major tourist destinations worldwide as well places famous for their cuisine and gastronomy. As shown by the numbers below, they have a comparable size and population as well as are ranked 2nd and 3rd on worldatlas.com in cities attracting the highest number of tourists in Europe in 2017 (1st is London).

	Rome	Paris	Source
Official population (city limits)	2,857,321	2,140,526	*(Wikipedia)*
Tourists (international arrivals 2017)	9.5 million	14.26 million	*(worldatlas.com)*
Number of restaurants	13,721 (2017)	44,896 (2017)	*(worldcitiesculturefforum.com)*

As a Frenchman and a foodie, I feel engaged in the existing rivalry between these two cities and decided to analyze and compare them further in order to explore how similar or different they are as well as conclude on which city has the most diverse restaurant scene and the best ratings on Foursquare.

Although Foursquare ratings are only partial, subjective and dependent on many variables, we can consider that since Foursquare has an international user base and that both cities are visited by important numbers of tourists, the users and the ratings are comparable.

1.2 Target group

While there are many different popular cuisines all over the world, the two main ones that often come back in comparison in discussions in western Europe are the French and the Italian ones regarding which one is the most popular. Foodies and gastronomy tourists from different countries would be interested in the results of a study comparing the ratings of both cities.

For most also, Paris is considered as more multicultural than Rome would be. An outcome of the present study will be to verify which city has the most diverse food scene, which can also be of interest of both local inhabitants and international gastronomy tourists looking to decide which city to visit.

1.3 Problem setting

The objective of this study and this workbook will be to answer 3 questions:

1. How similar are the neighborhoods of Paris and Rome and if they are dissimilar, what distinguishes them?
2. Which of the two cities has the most diverse food scene?

3. Which one of the two cities has the best ratings (both in total and for their national cuisine)?

In the next section we will cover the data needed to answer these questions.

2. Data

In order to solve the problems set above, we will need different sources of data.

- A breakdown of the two cities into "Neighborhoods"

The first type of data will be way to divide and separate the different areas in Rome and Paris in order to answer the first question which is to compare the neighborhoods of both cities and find similarities or differences. To do so the following lists and sources of information will be used:
- the Wikipedia page describing the 'Administrative subdivision of Rome' into 15 municipi (https://it.wikipedia.org/wiki/Municipi_di_Roma),
- the Wikipedia page describing the 'Arrondissements of Paris' into 20 arrondissements (https://en.wikipedia.org/wiki/Arrondissements_of_Paris).

- GPS coordinates to be used with the Foursquare API and database

The second type of data are GPS coordinates to be able to locate each neighborhood and search venues around using the Foursquare API. To do so we can use the geolocator package but in case we do not obtain correct information we can use the following websites:
- The website coordinatesmarker.com (https://www.coordinatesmarker.com/results/x5CIL-z14eDjgYvrFSXe) provides the GPS coordinates of the all 'municipi' of the city of Rome. A CSV file containing has been created and stored on Github.
- The same website (https://www.coordinatesmarker.com/results/YJtUk9HYiG8XQX0k28q9) has been used to get the coordinates of the different 'arrondissements' of Paris. A CSV file containing has been created and stored on Github.

- The Foursquare database - list of venues

The third source of information and data is the Foursquare database and API which contains venues in both cities. Accessing the database using the GPS coordinates listed in the files mentioned above we can characterize the neighborhoods by obtaining the categories of the venues. Further exploring the venues and their categories we can isolate the types of restaurants and their numbers to answer the second question.

- The Foursquare database - restaurants ratings

Finally, still using the Foursquare database and isolating the restaurants in both cities we can extract their ratings to be able to compare them between Rome and Paris.

3. Methods

3.1 Obtaining the list of neighborhoods in Rome and Paris

In order to first obtain the lists of 15 municipi in Rome and 20 arrondissements in Paris we use the beautiful soup package in order to scrape the two tables we want from the Wikipedia pages mentioned in the section above. We store them in dataframes, with the number of the neighborhood (Municipio in Rome and Arrondissement in Paris) in the first column, and the name of the neighborhood in the second.

3.2. Merging it with the coordinates lists

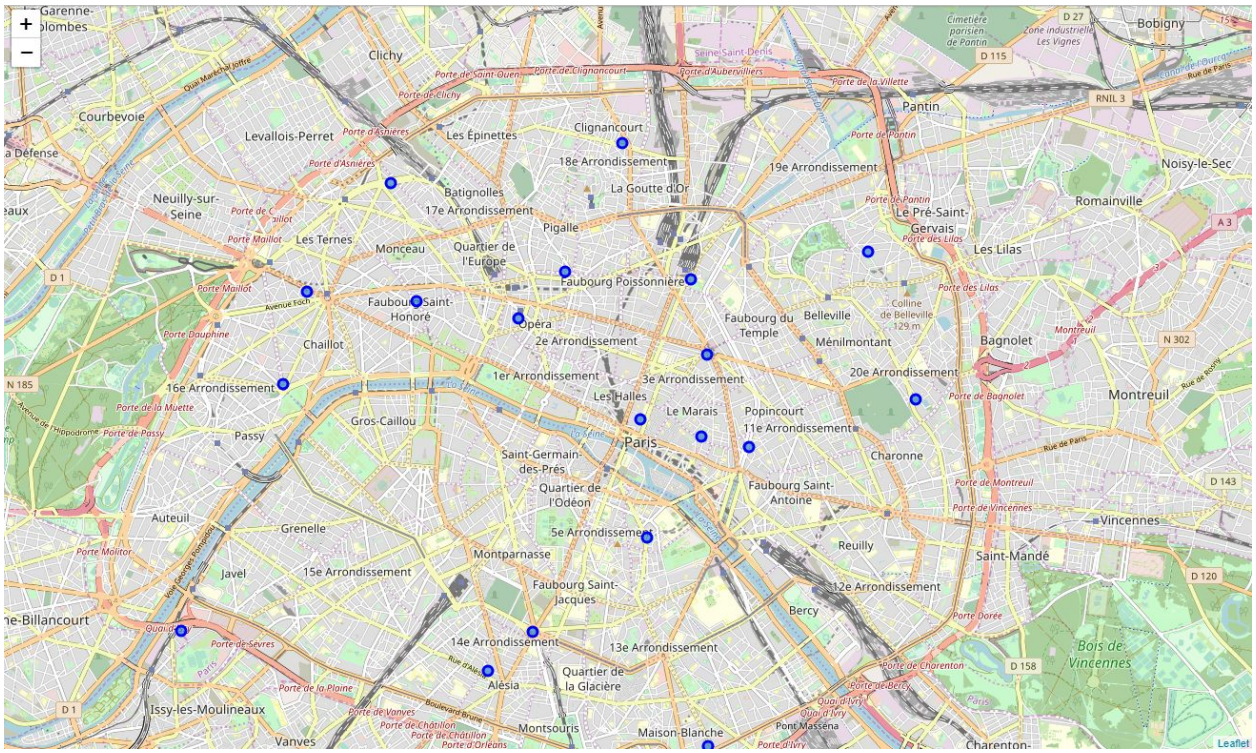
As mentioned in the Data section, we have already obtained and stored csv files of the coordinates of the neighborhoods of Rome and Paris. Now we need to import them, reorganize them, verify them, and finally merge them with the dataframes created in 3.1.

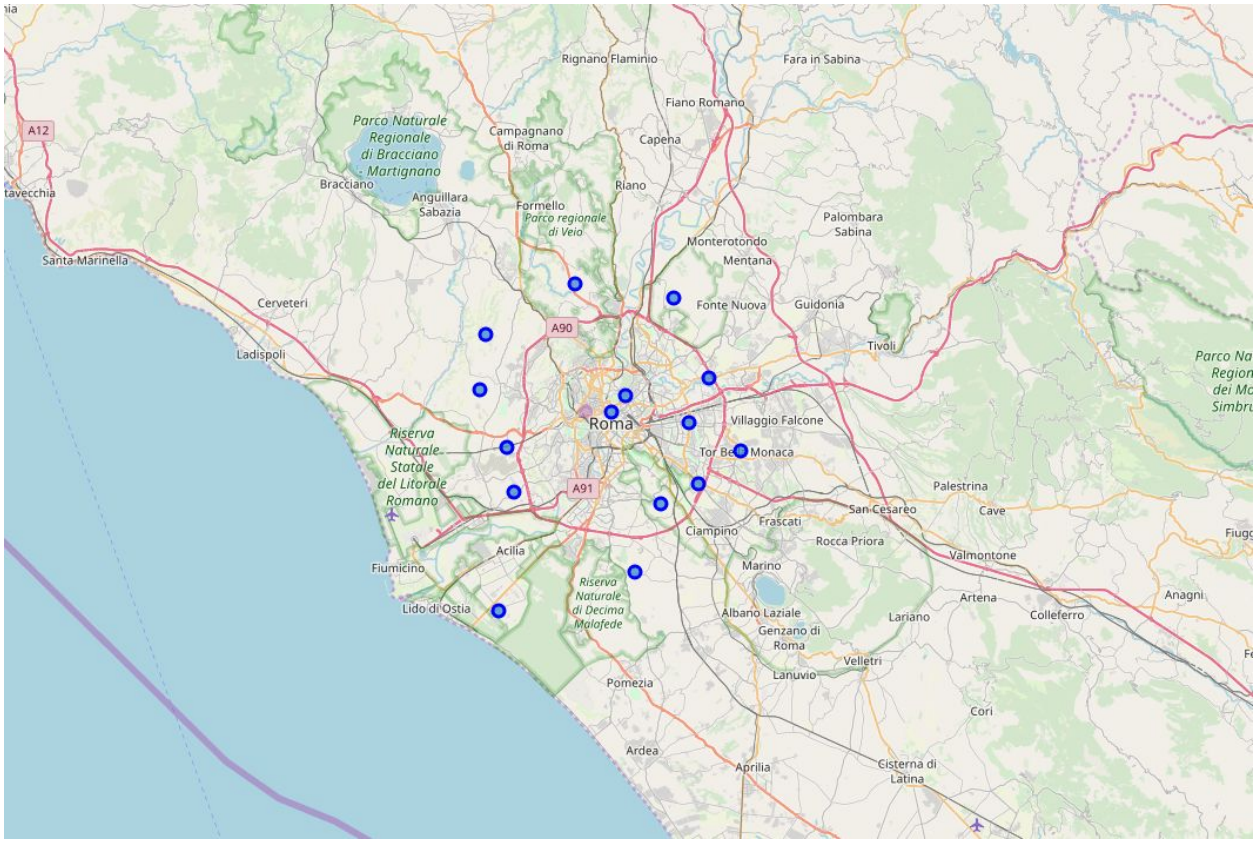
In order to merge both tables we need the Municipi to have the same names in both tables Rome_df and Rome_coordinates_df. We will settle on the naming "Municipio Roma xxx" as the final name to identify the Municipi. An extra column "Neighborhood_Name" will contain the more descriptive name of the Neighboorhood. We do the same for Paris. We will choose to name the neighborhoods "Paris arrondissement xxx" and another column will indicate a more descriptive name for the neighborhoods.

The data source for neighborhoods coordinates does not look right for arrondissements V, X, XI which have completely different coordinates from the other ones. Other sources of data gives us the following information so we can correct the coordinates for the erroneous neighborhoods.

3.3 Visualizing the Neighborhoods of both cities on maps

We can first create Folium maps to visualize the neighborhood of both cities on maps to make sure we have correct coordinates.





We can see that the Municipi in Roma are more spread out than the arrondissement in Paris. However, if we compare the totals of the populations in Paris's arrondissements and in Roma's Municipi we get a comparable figure (between 2 and 3 millions). Therefore we will assume that this is a comparable level for our analysis and we will just be careful to use different radius for our searches with the Foursquare API.

3.4 Exploring the neighborhoods with the Foursquare API

Now that we have our lists of neighborhoods with their coordinates we can start analyzing and exploring them with the Foursquare API.

3.4.1. Extracting the lists of venues and their category types for every neighborhood

We start by preparing our Foursquare API credentials, a function "get_category_type()" to extract the category of the venues, and another function "get_nearby_venues()", to extract the lists of venues in each neighborhood. Once used with the coordinates prepared before we obtain two dataframe with the list of venues and their names and categories.
In Rome we obtain a list of 1403 venues:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Municipio Roma I	41.902915	12.482536	Pane e Salame	41.900604	12.481798	Sandwich Place
1	Municipio Roma I	41.902915	12.482536	La Sandwicheria	41.902901	12.483336	Sandwich Place
2	Municipio Roma I	41.902915	12.482536	Louis Vuitton	41.903849	12.478822	Boutique
3	Municipio Roma I	41.902915	12.482536	Venchi	41.900042	12.480883	Ice Cream Shop
4	Municipio Roma I	41.902915	12.482536	La Prosciutteria	41.901888	12.484467	Italian Restaurant

In Paris we obtain a list of 1893 venues:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Paris arrondissement I	48.834188	2.331955	Amorino	48.833763	2.329871	Ice Cream Shop
1	Paris arrondissement I	48.834188	2.331955	Poissonnerie Daguerre Marée	48.833492	2.330609	Fish Market
2	Paris arrondissement I	48.834188	2.331955	Swann et Vincent	48.833882	2.330880	Italian Restaurant
3	Paris arrondissement I	48.834188	2.331955	La Lingerie	48.837353	2.334107	Bar
4	Paris arrondissement I	48.834188	2.331955	Les Grands Voisins	48.837244	2.334895	Art Gallery

We have reached the maximum number of 100 venues per neighborhood in most neighborhoods. We will continue our analysis with these lists.

3.4.2 Analysis of the neighborhoods and the categories of venues.

Now that we have our lists of max 100 venues describing each neighborhood we will start analyzing them. First we will do a One hot encoding and group the venues category by the number of venues in each category. This will enable us to see, after some sorting, the 10 most frequent types of venues in each neighborhood.

	Neighborhood	1stMost common venue	2ndMost common venue	3rdMost common venue	4th Most common venue	5th Most common venue	6th Most common venue	7th Most common venue	8th Most common venue	9th Most common venue	10th Most common venue
0	Municipio Roma I	Plaza	Ice Cream Shop	Historic Site	Sandwich Place	Hotel	Wine Bar	Monument / Landmark	Fountain	Church	Pizza Place
1	Municipio Roma II	Plaza	Ice Cream Shop	Hotel	Historic Site	Sandwich Place	Monument / Landmark	Pizza Place	History Museum	Park	Italian Restaurant
2	Municipio Roma III	Italian Restaurant	Pizza Place	Ice Cream Shop	Cocktail Bar	Park	Café	Clothing Store	Steakhouse	Hotel	Gym / Fitness Center
3	Municipio Roma IV	Ice Cream Shop	Italian Restaurant	Plaza	Hotel	Pizza Place	Park	Art Museum	Sandwich Place	Historic Site	Wine Bar
4	Municipio Roma IX	Pizza Place	Café	Hotel	Italian Restaurant	Ice Cream Shop	Clothing Store	Restaurant	Fast Food Restaurant	Pool	Gym / Fitness Center

	Neighborhood	1stMost common venue	2ndMost common venue	3rdMost common venue	4th Most common venue	5th Most common venue	6th Most common venue	7th Most common venue	8th Most common venue	9th Most common venue	10th Most common venue
0	Paris arrondissement I	French Restaurant	Hotel	Italian Restaurant	Pizza Place	Bakery	Bar	Vietnamese Restaurant	Creperie	Wine Shop	Coffee Shop
1	Paris arrondissement II	French Restaurant	Italian Restaurant	Wine Bar	Restaurant	Bistro	Coffee Shop	Hotel	Korean Restaurant	Breakfast Spot	Seafood Restaurant
2	Paris arrondissement III	French Restaurant	Hotel	Italian Restaurant	Plaza	Art Museum	Bakery	Museum	Burger Joint	Restaurant	Shopping Mall
3	Paris arrondissement IV	French Restaurant	Art Gallery	Clothing Store	Café	Hotel	Italian Restaurant	Tea Room	Cocktail Bar	Ice Cream Shop	Pastry Shop
4	Paris arrondissement IX	French Restaurant	Hotel	Wine Bar	Italian Restaurant	Pizza Place	Bakery	Cocktail Bar	Bar	Restaurant	Gourmet Shop

3.4.3 Clustering

Now that we have a description of the most common types of venues in each neighborhood we can start the clustering. First we will merge both lists as the goal is to see if any clusters can be formed between neighborhoods in Paris and in Rome. We use the KMeans package, with 6 clusters, to see what clusters can be formed.

3.4.4 Clusters analysis

Finally, let's display and observe the clusters formed. It is interesting to see if the KMeans Clustering method has formed any clusters grouping neighborhoods belong to Rome and Paris.

- Cluster 0:

	Neighborhood	Neighborhood_Name	1stMost common venue	2ndMost common venue
0	Municipio Roma I	Centro Storico	Plaza	Ice Cream Shop
1	Municipio Roma II	Parioli/Nomentano	Plaza	Ice Cream Shop
3	Municipio Roma IV	Tiburtina	Ice Cream Shop	Italian Restaurant
4	Municipio Roma V	Prenestino/Centocelle	Ice Cream Shop	Historic Site
7	Municipio Roma VIII	Appia Antica	Ice Cream Shop	Historic Site
11	Municipio Roma XII	Monte Verde	Italian Restaurant	Ice Cream Shop

- Cluster 1:

	Neighborhood	Neighborhood_Name	1stMost common venue	2ndMost common venue
16	Paris arrondissement II	Bourse	French Restaurant	Italian Restaurant
18	Paris arrondissement IV	Hôtel-de-Ville	French Restaurant	Art Gallery
19	Paris arrondissement V	Panthéon	French Restaurant	Bakery
20	Paris arrondissement VI	Luxembourg	French Restaurant	Hotel
21	Paris arrondissement VII	Palais-Bourbon	Hotel	French Restaurant
23	Paris arrondissement IX	Opéra	French Restaurant	Hotel
24	Paris arrondissement X	Entrepôt	French Restaurant	Coffee Shop
25	Paris arrondissement XI	Popincourt	French Restaurant	Pizza Place
29	Paris arrondissement XV	Vaugirard	French Restaurant	Japanese Restaurant
32	Paris arrondissement XVIII	Butte-Montmartre	French Restaurant	Bar
33	Paris arrondissement XIX	Buttes-Chaumont	French Restaurant	Bar
34	Paris arrondissement XX	Ménilmontant	French Restaurant	Bar

- Cluster 2:

	Neighborhood	Neighborhood_Name	1stMost common venue	2ndMost common venue
26	Paris arrondissement XII	Reuilly	Bus Stop	Hotel

- Cluster 3:

	Neighborhood	Neighborhood_Name	1stMost common venue	2ndMost common venue
2	Municipio Roma III	Monte Sacro	Italian Restaurant	Pizza Place
5	Municipio Roma VI	Roma delle Torri	Italian Restaurant	Pizza Place
6	Municipio Roma VII	Appio-Latino/Tuscolana/Cinecittà	Pizza Place	Historic Site
8	Municipio Roma IX	Eur	Pizza Place	Café
9	Municipio Roma X	Ostia/Acilia	Beach	Italian Restaurant
10	Municipio Roma XI	Arvalia/Portuense	Pizza Place	Ice Cream Shop
12	Municipio Roma XIII	Aurelia	Italian Restaurant	Hotel
13	Municipio Roma XIV	Monte Mario	Italian Restaurant	Café
14	Municipio Roma XV	Cassia/Flaminia	Italian Restaurant	Pizza Place

- Cluster 4:

	Neighborhood	Neighborhood_Name	1stMost common venue	2ndMost common venue
15	Paris arrondissement I	Louvre	French Restaurant	Hotel
17	Paris arrondissement III	Temple	French Restaurant	Hotel
22	Paris arrondissement VIII	Élysée	Hotel	French Restaurant
28	Paris arrondissement XIV	Observatoire	French Restaurant	Hotel
30	Paris arrondissement XVI	Passy	French Restaurant	Hotel
31	Paris arrondissement XVII	Batignolles-Monceau	French Restaurant	Italian Restaurant

- Cluster 5:

	Neighborhood	Neighborhood_Name	1stMost common venue	2ndMost common venue
27	Paris arrondissement XIII	Gobelins	Vietnamese Restaurant	Asian Restaurant

The results will be discussed in the results section.

3.5 Food scene diversity

The second question we posed was to estimate the food scene diversity in both capitals. To do so, we will filter out dataframes containing the venues categories to only keep the categories corresponding to restaurants types. We will calculate in each city:

- the total number of restaurants,
- the number of different types of restaurants,

- the number of types of French Restaurants in Paris and of Italian Restaurants in Rome
- the total number of restaurants of these types in Paris and Rome
- the number of types of restaurant outside of French Restaurants in Paris and of Italian Restaurants in Rome
- the total number of restaurants of these types in Paris and Rome.

To do so we use the venue categories information from the Foursquare API and filter the categories related to restaurants. We also filter manually on French regional and national restaurants (such as Alsatian, Savoyard, Southwestern restaurant etc.) and on Italian restaurants types (such as Roman restaurant, Pizza place and Trattoria/Osteria).

3.6 Restaurants ratings

Now we want to answer the third problem we posed for our study which is to decide on which city has the highest ratings for their restaurants. To do so, we first need to gather the lists of restaurants in both cities, with their Foursquare unique ID. Once we have the lists of restaurants, we will run another series of API calls to obtain the ratings of the restaurants listed.
Note: Since we are limited in the number of API calls we can perform every day, we will reduce the perimeter of our study. We will only consider restaurants in the 5 most center neighborhoods of Rome, meaning restaurants in the neighborhoods called Municipio Roma I to V. Likewise, we will reduce the perimeter of the analysis in Paris to the 5 most center arrondissements (being Paris arrondissement 1 to 5). Finally, another limitation we will make to our study is to only compare French Restaurants' ratings in Paris to Italian Restaurants' ratings in Rome.

3.6.1 Scope definition

Redefining the scope according to these limitations lets us create two filtered dataframe, with 78 restaurants remaining in the scope in Paris and 41 in Rome. We create a function, getIDs(), to extract the IDs of the restaurants in this scope from the Foursquare API. We will need the IDs of the restaurants to run API calls to extract their ratings on Foursquare.

3.6.2 Ratings extraction

Finally we create a function getRating(), which uses the IDs of the restaurants to make API calls to extract the rating (when available). 47 out of the 78 restaurants in scope in Paris have ratings available and 41 out of the 41 in scope in Rome have ratings. We finish by calculating the average rating on these reduced scopes.

4. Results

4.1 Similarity between Paris's and Rome's neighborhoods

The results of the KMeans clustering analysis carried out in section 3.4 show that with 6 clusters created, the neighborhoods in Rome and Paris are too different to be classified together. Indeed, the clusters created only contain neighborhoods of one city, but never of both.

4.2 Food scene diversity

The results of the food scene diversity analysis are summarized in the table below:

	Rome	Paris
Total number of restaurants	379	863
Number of restaurants types	28	72
Number of national restaurants types	4	11
Total number of restaurants of national restaurants types	251	309
Number of restaurants types outside of national restaurants types	24	61
Total number of restaurants outside of national restaurants types	128	554

This shows that even when taking out the different types of French regional restaurant types into the calculations (11 types representing 309 restaurants)and doing the same with Italian national restaurant types (4 types representing 251 restaurants) there are still 61 types of restaurants in Paris (representing 554 restaurants) versus 24 types in Rome (representing 128 restaurants). We can infer that Paris has the most diverse food scene.

4.3 Restaurants ratings in Paris and Rome

The following table summarizes the results of the analysis of the ratings, with a reminder of the scope used and the calculation of the average rating over that scope.

	Rome	Paris
Number of restaurants in the scope	41	47
Average rating of the restaurants withing that scope	8.21	8.29

As we can see, the results are very close over scopes that are in fact quite similar in size and can therefore be compared. The ratings seems to be slightly higher in Paris than in Rome but it would be interesting to know if the trend remains if we extend the scope by including more neighborhoods and more types of restaurants.

5. Discussion

5.1 Clustering analysis

The results of the clustering analysis are however partial. Indeed, it would be needed to observe the results of the KMeans clustering approach when trying with different numbers of clusters (in the present analysis, a number of 6 clusters has been used). It is however unlikely that with higher number of clusters we would start seeing clusters containing neighborhoods of both cities, as the level of detail gets higher. Therefore, it would be interesting to do such a story but on focusing on only certain parts of both cities, and breaking them down into more precise subdivisions that arrondissements and municipi. For example, by breaking down Municipio I in Roma into quartieri, an existing subdivision, we could potentially focus on some more specific

areas of the very center of Rome, such as the historical center. We could then compare them to the same type of areas in Paris using the subdivision into 'quartiers' and comparing to Paris's historical center as well to see how both areas differ. It could be interesting also to do a similar analysis on the business centers of both cities.

5.2 Food scene diversity

The analysis on the types of restaurants appearing in the categories defined in Foursquare seem to show a way more diverse food scene in Paris than in Rome. This can be observed through the number of different types of restaurants, including or not regional and national types. This does not necessarily come as a surprise as while both cities attract very important numbers of international tourists as mentioned in the Introduction section, Paris is known to be a more multi-cultural city than Rome, which would explain the higher diversity in restaurants created in Paris. However, it is important to note that a limit to the conclusion of this study is that the system of categories defined on Foursquare is not necessarily homogeneous and some categories might in fact intersect or on the contrary should be broken down to more details. For example, in the list of categories of restaurants found in Paris we can see examples of French regional cuisines such as 'Alsatian', 'Savoyard', 'Basque' or 'Southwestern', while in the list of categories of Rome's restaurants there are only 4 categories referring to Italian cuisine which are 'Italian restaurant', 'Pizza Place', 'Roman restaurant' and 'Trattoria/Osteria'. We could have imagined more categories in Rome such as 'Sicilian Restaurant', and other regional cuisines. The point of the comment here is that food scene diversity remains something that is difficult to analyze analytically if not defined and measured precisely and in the case of the present study, the definition of "venue categories" in Foursquare does not seem to be precise and homogeneous enough to do so.

5.3 Rome VS Paris restaurants

The results of the analysis are also very partial. Due to the daily number of API calls constraints we were not able to expand the scope of the study, therefore resulting on a very limited scope, both geographically and in terms of restaurants included. It would be interesting to expand the scope of the study to see if the trend observed remains.

Furthermore, Foursquare users' ratings is a very subjective and partial measure of the actual quality of a restaurant. As mentioned in the Introduction section, there might not be the same userbase in both cities and the ratings given depend on very diverse factors. Although it is not possible to improve the quality of the results and conclusions here with the data used, it is important to keep these limitations in mind when analyzing the results.

6. Conclusion

This analysis shows that the different neighborhoods in Paris and Rome are in fact quite dissimilar although inside of each city, clusters can be formed. The food scene seems to be more diverse in Paris than it is in Rome, which makes sense since Paris is known to be a more multi-cultural city. Finally, while very limited, the results of the analysis on the restaurants' ratings seem to show a trend according to which Paris's restaurants have slightly higher ratings

than Rome, providing French people (at least Parisians) something to be proud of over their Roman counterparts.