Open a Restaurant in Rome - Where's better?

Introduction

If someone is looking to open a restaurant, where would you recommend that they open it?

Rome is the capital city of Italy as well as the capital of the Lazio region. The city has been a major human settlement for over two millennial. With 2,860,009 residents in 1,285 km2 (496.1 sq mi),it is also the country's most populated comune. It is the third most populous city in the European Union by population within city limits. It offers many business opportunities for new bees so is competitive. The analysis from this report help for the new businesses strategically target the market and help in a high return on investment so is low risk.

Description

Opening a new restaurant can be daunting. There are so many things to consider - from choosing the right location to finding financing to selecting the right name. And, of course, buying equipment and hiring staff. The first step in opening a new restaurant is deciding what type of restaurant it is going to be. Do you want to open a high-end fine dining restaurant? A casual 1950s-style diner? Do you have a specific type of cuisine in mind, such as Italian, French, or Indian? But a good question could be: **Is the restaurant located in a good position?** For example, is the location busy with plenty of foot traffic? If not, is there enough parking? Have several other restaurants opened and closed in the same spot?

We will focus on the types of different restaurants which are opened or closed in a particular place and then decide if it is a good place to open a new restaurant based on the popular cuisine around that place.

Target Audience

The objective of this project is to recommend which neighborhood of Rome is a good choice for a new restaurant business to open.

Data

We will be using a list of Rome Neighborhood data taken from Wikipedia (https://en.wikipedia.org/wiki/Quarters_of_Rome (https://en.wikipedia.org/wiki/Quarters_of_Rome)) for analysis. There is a total of 35 neighborhoods in Rome. We will find latitude and longitude of each neighborhood, then we will use Rome geographical data as input to the Foursquare location data and fetch top 50 restaurants nearby to each neighborhood within 1000 meters. Then we will make a decision examining each cluster of neighborhoods.

Sample records from Wikipedia: (Q. I, Q. II, etc.. stands for Neighborhood 1, Neighborhood 2, and so on...)

- **Q. I** Flaminio
- **Q. II** Parioli
- **Q. III** Pinciano
- **Q. IV** Salario
- **Q. V** Nomentano
- **Q. VI** Tiburtino
- **Q. VII** Prenestino-Labicano
- **Q. VIII** Tuscolano
- **Q. IX** Appio-Latino
- **Q. X** Ostiense
- **Q. XI** Portuense
- **Q. XII** Gianicolense
- **Q. XIII** Aurelio
- **Q. XIV** Trionfale
- **Q. XV** Della Vittoria
- **Q. XVI** Monte Sacro
- **Q. XVII** Trieste
- **Q. XVIII** Tor di Quinto
- **Q. XIX** Prenestino-Centocelle
- **Q. XX** Ardeatino
- **Q. XXI** Pietralata
- **Q. XXII** Collatino
- **Q. XXIII** Alessandrino
- **Q. XXIV** Don Bosco
- **Q. XXV** Appio Claudio
- **Q. XXVI** Appio-Pignatelli
- **Q. XXVII** Primavalle
- **Q. XXVIII** Monte Sacro Alto
- **Q. XXIX** Ponte Mammolo
- **Q. XXX** San Basilio
- **Q. XXXI** Giuliano-Dalmata
- **Q. XXXII** Europa
- **Q. XXXIII** Lido di Ostia Ponente
- **Q. XXXIV** Lido di Ostia Levante
- **Q. XXXV** Lido di Castel Fusano

Then we will find the latitude and longitude of each neighborhood and map to the data frame.

	Neighborhood	Name	Latitude	Longitude
1	Q. II	Parioli	41.925949	12.487320
2	Q. III	Pinciano	41.918781	12.485332
3	Q. IV	Salario	41.940398	12.508896
4	Q. V	Nomentano	41.917002	12.512918
5	Q. VI	Tiburtino	41.896633	12.511238
6	Q. VII	Prenestino-Labicano	41.891549	12.516829
7	Q. VIII	Tuscolano	41.867638	12.539368

We will then use this Rome geographical data as input to the Foursquare location data and fetch top 50 restaurants nearby to each neighborhood within 1000 meters. The sample foursquare location data looks like this:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
1	Q. II	41.925949	12.487320	Metamorfosi Restaurant	41.924077	12.484641	Italian Restaurant
6	Q. II	41.925949	12.487320	La Pariolina	41.928589	12.486173	Italian Restaurant
7	Q. II	41.925949	12.487320	Ristorante Molto	41.927481	12.484284	Italian Restaurant
12	Q. II	41.925949	12.487320	Il Caminetto	41.928184	12.486310	Italian Restaurant
13	Q. II	41.925949	12.487320	Chez Cocò	41.929623	12.485371	Italian Restaurant
...	...	...	...	...	...	...	...
959	Q. XXXV	41.716500	12.323709	Sporting Beach	41.713333	12.317398	Italian Restaurant
960	Q. XXXV	41.716500	12.323709	La Mariposa	41.709623	12.326375	Italian Restaurant
961	Q. XXXV	41.716500	12.323709	Kursaal	41.715174	12.315824	Italian Restaurant
962	Q. XXXV	41.716500	12.323709	kursal	41.715134	12.315835	Italian Restaurant
964	Q. XXXV	41.716500	12.323709	Da rutilio a mare	41.715249	12.314477	Italian Restaurant

Methodology

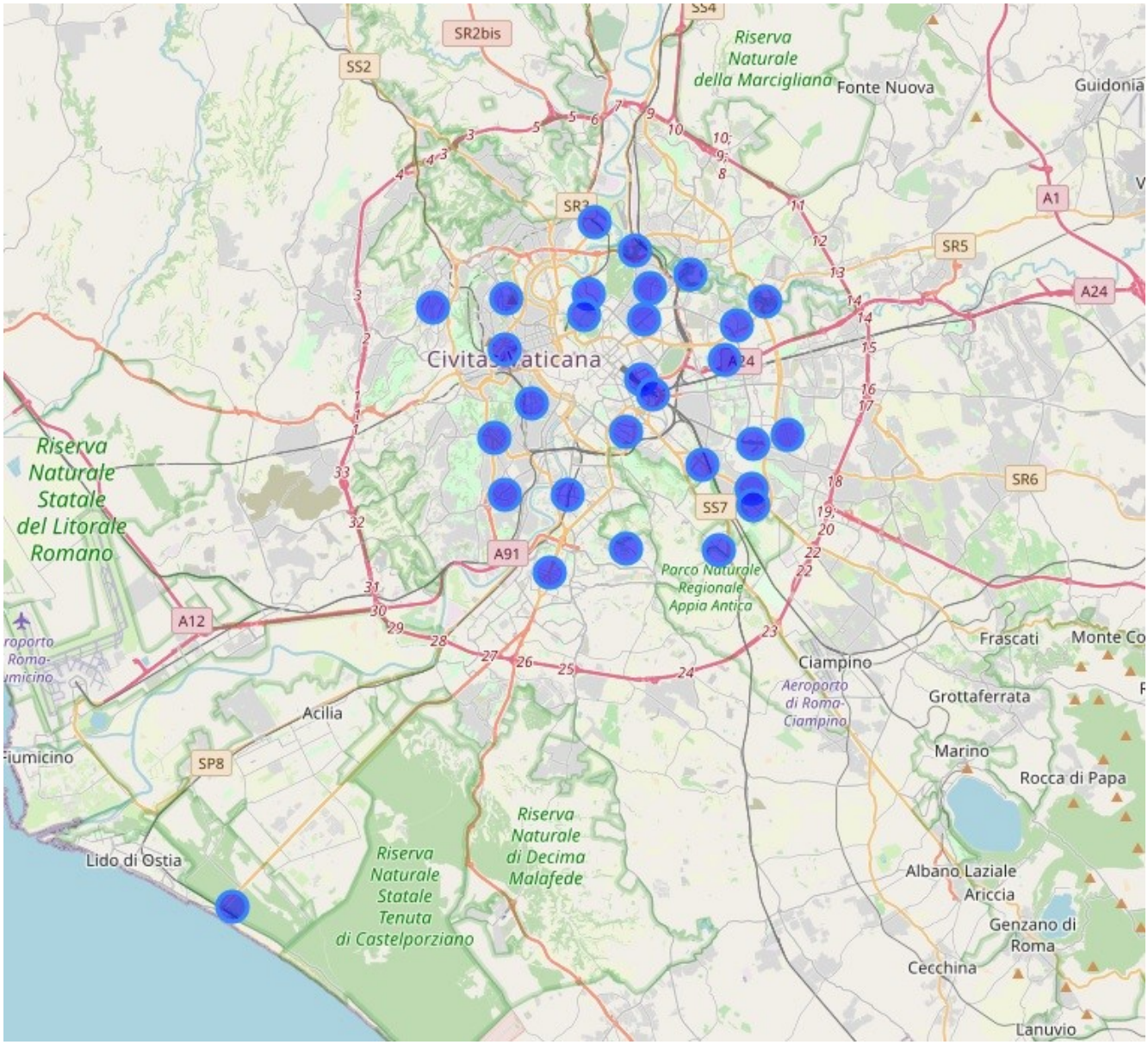
The objective of this project is to find which neighbourhood of Houston is a good choice for a new restaurant business to open. i.e finding the best location for a new restaurant to start.

We will follow this step:

1. We first take the neighborhood information from Wikipedia and load that into a data frame dropping unnecessary columns.
2. Then we perform data wrangling to convert the data into the analysis-ready form.
3. Then we use the geopy library to get the latitude and longitude information for each neighborhood.
4. Then we use this data to get the 50 nearby restaurants within 1000 meters of a neighborhood using foursquare location data.
5. Calculate and perform unique restaurant categories in each neighborhood and find the top 10 most common restaurants.
6. Then we will cluster all the restaurants in all the neighborhoods using the k-means clustering algorithm.

Rome Neighbourhood visualization

Here we can see a map of Rome with neighborhoods superimposed on top. The following depiction is a map generated using python ‘folium’ library.



Total number of venues in each Neighbourhood

There are 68 different types of restaurants in all the neighbourhood and total of 1156 restaurants.

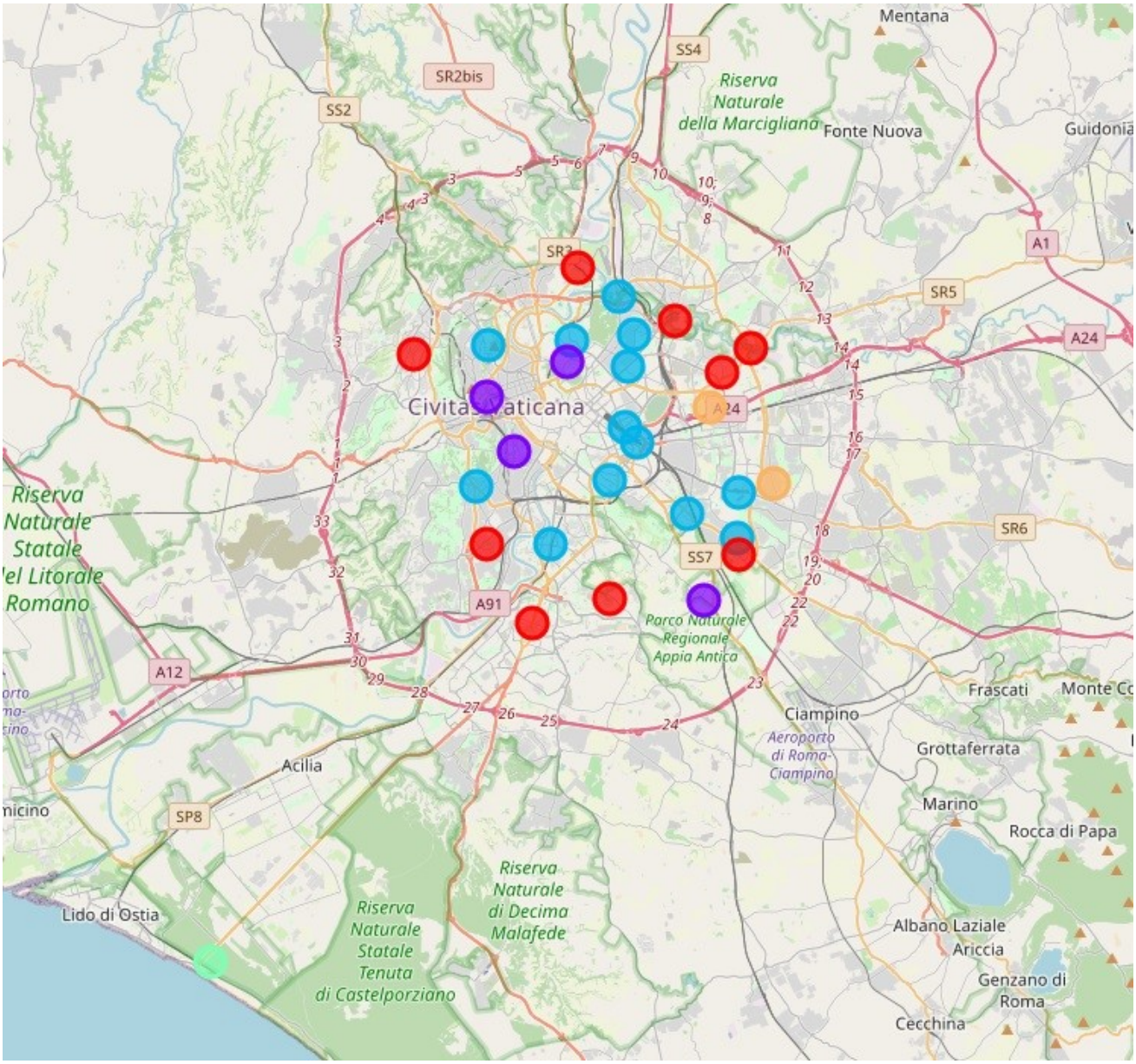
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
3	Q. II	41.925949	12.487320	Metamorfosi Restaurant	41.924077	12.484641	Italian Restaurant
12	Q. II	41.925949	12.487320	Bucavino	41.917442	12.499237	Italian Restaurant
15	Q. II	41.925949	12.487320	Enoteca La Torre a Villa Laetitia	41.916261	12.469070	Italian Restaurant
18	Q. II	41.925949	12.487320	Larys	41.916530	12.499190	Italian Restaurant
19	Q. II	41.925949	12.487320	Ristorante Andrea	41.909045	12.489692	Italian Restaurant
...	...	...	...	...	...	...	...
2692	Q. XXXV	41.716500	12.323709	La Querida	41.733805	12.334911	Italian Restaurant
2696	Q. XXXV	41.716500	12.323709	cinghiale	41.731761	12.350988	Italian Restaurant
2697	Q. XXXV	41.716500	12.323709	Fraschetta Da Matteo	41.735674	12.345803	Italian Restaurant
2698	Q. XXXV	41.716500	12.323709	La Mariposa	41.709623	12.326375	Italian Restaurant
2700	Q. XXXV	41.716500	12.323709	La bussola	41.707346	12.332224	Italian Restaurant

Find 10 most common restaurants in each neighbourhood:

	Cluster Labels	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most 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	2	Q. II	Italian Restaurant	Restaurant	Pizza Place	Café	Sandwich Place	Diner	Japanese Restaurant	Trattoria/Osteria	Bistro	Bakery
1	1	Q. III	Italian Restaurant	Sandwich Place	Turkish Restaurant	Restaurant	Café	Korean Restaurant	Bistro	Vietnamese Restaurant	Ethiopian Restaurant	Falafel Restaurant
2	2	Q. IV	Italian Restaurant	Pizza Place	Café	Japanese Restaurant	Chinese Restaurant	Restaurant	Gastropub	Trattoria/Osteria	Steakhouse	Sandwich Place
3	2	Q. IX	Italian Restaurant	Pizza Place	Café	Restaurant	Roman Restaurant	Sandwich Place	Gastropub	Diner	Trattoria/Osteria	Fish & Chips Shop
4	2	Q. V	Pizza Place	Italian Restaurant	Café	Mediterranean Restaurant	Trattoria/Osteria	Gastropub	Seafood Restaurant	Sandwich Place	Salad Place	Roman Restaurant

Results

From the foursquare location data, we got 50 restaurants in each neighbourhood and found most common restaurants. Performing the k-means clustering on the most common restaurants for each neighbourhood will partition the data set into 5 clusters. The 5 clusters are partitioned based on similar type of restaurants that belong to neighbourhoods. The folium representation of the clustered data will look like below.



Examining the Clusters:

Cluster 1 (Red)

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most 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
10	Q. XI	Italian Restaurant	Pizza Place	Café	Restaurant	Sandwich Place	Food	Vegetarian / Vegan Restaurant	Fast Food Restaurant	Modern European Restaurant	Chinese Restaurant
15	Q. XVI	Italian Restaurant	Café	Pizza Place	Sushi Restaurant	Restaurant	Diner	Japanese Restaurant	Mexican Restaurant	Thai Restaurant	Gastropub
17	Q. XVIII	Italian Restaurant	Café	Pizza Place	Restaurant	Bakery	Sandwich Place	Seafood Restaurant	Steakhouse	Asian Restaurant	Gastropub
19	Q. XX	Café	Italian Restaurant	Pizza Place	Restaurant	Japanese Restaurant	Bistro	Sandwich Place	Moroccan Restaurant	Fast Food Restaurant	Chinese Restaurant
20	Q. XXI	Pizza Place	Café	Italian Restaurant	Japanese Restaurant	Chinese Restaurant	Restaurant	Steakhouse	Seafood Restaurant	Sandwich Place	African Restaurant
24	Q. XXV	Café	Pizza Place	Italian Restaurant	Restaurant	Bakery	Noodle House	Middle Eastern Restaurant	Fried Chicken Joint	Trattoria/Osteria	Chinese Restaurant
26	Q. XXVII	Café	Pizza Place	Italian Restaurant	Chinese Restaurant	Restaurant	Fast Food Restaurant	Bakery	Sandwich Place	Vegetarian / Vegan Restaurant	Peruvian Restaurant
28	Q. XXIX	Café	Italian Restaurant	Pizza Place	Seafood Restaurant	Food Court	Restaurant	Fast Food Restaurant	Burger Joint	Diner	Bistro
31	Q. XXXII	Café	Italian Restaurant	Pizza Place	Restaurant	Fast Food Restaurant	Sandwich Place	Chinese Restaurant	Sushi Restaurant	Steakhouse	Bakery

‘Coffee shop’ holds a larger accountability for this cluster with 5 occurrences in ‘1st Most Common Venue’ across different neighborhoods followed by ‘Italian Restaurant’ with 3 occurrences. Also, ‘Coffee shop’ occurs 3 times in ‘2nd Most Common Venue’. So, Cluster – 1 is a ‘Coffee Shop’ dominant cluster.

Cluster 2 (Violet)

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most 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
2	Q. III	Italian Restaurant	Sandwich Place	Turkish Restaurant	Restaurant	Café	Korean Restaurant	Bistro	Vietnamese Restaurant	Ethiopian Restaurant	Falafel Restaurant
12	Q. XIII	Italian Restaurant	Sandwich Place	Restaurant	Bakery	Café	Falafel Restaurant	Trattoria/Osteria	Pizza Place	Romagna Restaurant	Fondue Restaurant
13	Q. XIV	Italian Restaurant	Café	Sandwich Place	Pizza Place	Diner	Vegetarian / Vegan Restaurant	Restaurant	Chinese Restaurant	Salad Place	Vietnamese Restaurant
25	Q. XXVI	Italian Restaurant	Sandwich Place	Gastropub	Café	Vietnamese Restaurant	Fondue Restaurant	Ethiopian Restaurant	Falafel Restaurant	Fast Food Restaurant	Fish & Chips Shop

In this second cluster we can see that the "Italian Restaurant" has 4/4 occurrences in '1st Most Common Venue' followed by ‘Sandwich Place’ with 3 occurrences. So, Cluster – 2 is a ‘Italian Restaurant’ dominant cluster.

Cluster 3 (Light blue)

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most 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
1	Q. II	Italian Restaurant	Restaurant	Pizza Place	Café	Sandwich Place	Diner	Japanese Restaurant	Trattoria/Osteria	Bistro	Bakery
3	Q. IV	Italian Restaurant	Pizza Place	Café	Japanese Restaurant	Chinese Restaurant	Restaurant	Gastropub	Trattoria/Osteria	Steakhouse	Sandwich Place
4	Q. V	Pizza Place	Italian Restaurant	Café	Mediterranean Restaurant	Trattoria/Osteria	Gastropub	Seafood Restaurant	Sandwich Place	Salad Place	Roman Restaurant
5	Q. VI	Italian Restaurant	Pizza Place	Restaurant	Bistro	Trattoria/Osteria	Sandwich Place	Greek Restaurant	Food Court	Soup Place	Bakery
6	Q. VII	Italian Restaurant	Pizza Place	Bakery	Greek Restaurant	Gastropub	Restaurant	Steakhouse	Noodle House	Fish & Chips Shop	Ethiopian Restaurant
7	Q. VIII	Pizza Place	Italian Restaurant	Café	Trattoria/Osteria	Restaurant	Steakhouse	Chinese Restaurant	Sandwich Place	Bakery	Bistro
8	Q. IX	Italian Restaurant	Pizza Place	Café	Restaurant	Roman Restaurant	Sandwich Place	Gastropub	Diner	Trattoria/Osteria	Fish & Chips Shop
9	Q. X	Pizza Place	Italian Restaurant	Café	Japanese Restaurant	Bakery	Restaurant	Mediterranean Restaurant	Roman Restaurant	Diner	Gastropub
11	Q. XII	Italian Restaurant	Pizza Place	Café	Restaurant	Chinese Restaurant	Burger Joint	Steakhouse	Breakfast Spot	Donut Shop	Diner
14	Q. XV	Italian Restaurant	Pizza Place	Restaurant	Japanese Restaurant	Sandwich Place	Mexican Restaurant	Creperie	Café	Burger Joint	Roman Restaurant
16	Q. XVII	Pizza Place	Seafood Restaurant	Italian Restaurant	Japanese Restaurant	Fast Food Restaurant	Vietnamese Restaurant	Donut Shop	Ethiopian Restaurant	Falafel Restaurant	Fish & Chips Shop
18	Q. XIX	Italian Restaurant	Pizza Place	Café	Fast Food Restaurant	Chinese Restaurant	Restaurant	Sandwich Place	Japanese Restaurant	Kebab Restaurant	Donut Shop
23	Q. XXIV	Italian Restaurant	Café	Pizza Place	Fast Food Restaurant	Bakery	Fried Chicken Joint	Japanese Restaurant	Steakhouse	Indian Restaurant	Seafood Restaurant

Italian Restaurant is even dominating in this cluster, in fact it is always in '1st Most Common Venue' with a huge accountability for this cluster. Cluster – 2 can be termed as ‘Italian Restaurant’ dominant cluster.

Cluster 4 (Green)

```
cluster_4 =rome_merged.loc[rome_merged['Cluster Labels'] == 3, rome_merged.columns[[0] + list(range(5, rome_merged.shape[1]))]]
cluster_4
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most 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
34	Q. XXXV	Italian Restaurant	Seafood Restaurant	Mediterranean Restaurant	Vietnamese Restaurant	Fondue Restaurant	Donut Shop	Ethiopian Restaurant	Falafel Restaurant	Fast Food Restaurant	Fish & Chips Shop

Cluster 5 (Orange)

```
cluster_5 = rome_merged.loc[rome_merged['Cluster Labels'] == 4, rome_merged.columns[[0] + list(range(5, rome_merged.shape[1]))]]
cluster_5
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most 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
21	Q. XXII	Pizza Place	Restaurant	Sushi Restaurant	Spanish Restaurant	Kebab Restaurant	Fast Food Restaurant	Deli / Bodega	Diner	Donut Shop	Ethiopian Restaurant
22	Q. XXIII	Pizza Place	Café	Fast Food Restaurant	Roman Restaurant	Vietnamese Restaurant	Donut Shop	Ethiopian Restaurant	Falafel Restaurant	Fish & Chips Shop	Fondue Restaurant

As we can see, for the last two clusters we have 1 and 2 neighborhoods respectively, so we aren't able to do a specific analysis about these two clusters.

Discussion

Based on the current results it can be observed that in the entire city of Rome there are many different types of restaurant, where the Italian Restaurant dominate the scene. In fact we have:

1. Cluster 1 – Coffee Shops
2. Cluster 2 – Italian Restaurant
3. Cluster 3 – Italian Restaurant
4. Cluster 4 and 5 - Italian Restaurant and Pizza Place

But, regarding the 2nd Most Common Venue in all the clusters, we can also say from the analysis that Pizza Place is the most common venue across all the clusters or neighborhoods. So, Pizza Place is a ready-to-go place for Rome.

Conclusion

On application of Clustering Algorithms, to a multi-dimensional dataset, a very inquisitive results can be curated which helps to understand and visualize the data. The neighborhoods of Rome were very briefly segmented into five clusters and upon analysis it was possible to rename them basis upon the categories of venues in and around that neighborhood. Italian Restaurant and Coffee Shops are very dominant in Rome. The results of this project can be improved and made more inquisitive by using a current Rome’s dataset along with API platforms which are more interested in Food Venues (like Yelp, etc.) The scope of this project can be expanded further to understand the dynamics of each neighborhood and suggest a new vendor a profitable location to open his or her food place.