Capstone_Project_Rome_Restaurant

May 30, 2020

1 A new restaurant in Rome

1.1 Introduction

1.1.1 Background

Rome is the largest city in Italy and it is one of the favourite touristic destinations for people all over the world. Apart from hosting an outstanding number of museums and other touristic attractions, Rome is home to a large number of restaurants. Restaurants' revenues are highly dependent on the number of clients, that tend to prefer certain neighbours over others due to many reasons: public transport, the number of restaurants to chose from, nearness to attractions and so on.

1.1.2 Problem

An investor from Brazil would like to invest in Rome by opening several restaurants in the city. Being new to the city he would like to receive some advice regarding the best locations where to open his restaurants. The main idea is to open his Brazilian restaurants in areas where there is already a high number of people going out for lunch or dinner and offering them something different from Italian restaurants, that represent the majority in Rome.

1.2 Data

The data used for the problem will be extracted from the Foursquare database and it will regard the venues in 35 neighbourhoods of Rome. In particular, the data imported from the database for each venue will be:

- name
- latitude
- longitude
- category

The neighbourhoods' names and locations (i.e. latitude and longitude) have been obtained from Wikipedia and they will be imported as a csv file.

The data from Foursquare will be used to cluster the neighbourhoods in Rome in order to identify the most promising locations where to open a restaurant: the core concept is to identify the cluster characterized by a high number of restaurants or similar venues since the high number of restaurants denotes a lively area with many clients.

1.3   Methodology

1.3.1   Exploratory analysis

First of all, we want to visualize the neighbourhoods on a map to see where they are located in Rome.
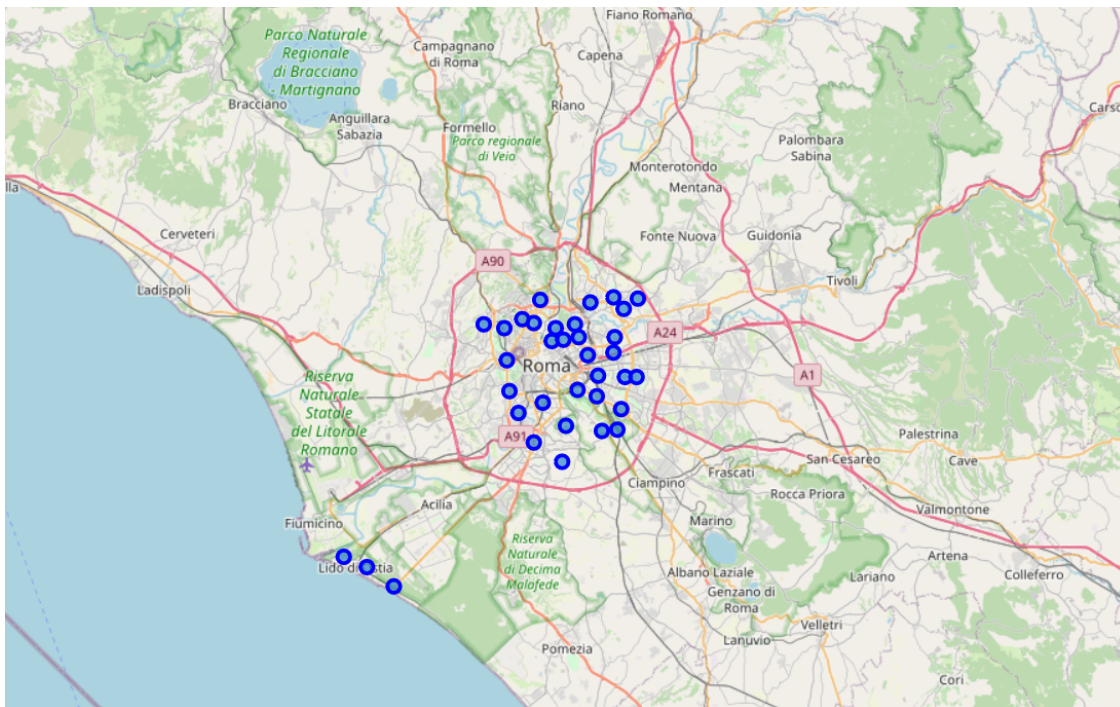

For each neighbourhood, a maximum of 200 venues in the radius of 1 km have been imported from the Foursquare dataset. In order to have a first idea of the activities that characterize the neighbourhoods, for each of them the 10 most common venues have been listed in descending order, from most to least common. From a first glace, restaurants and Italian restaurants, in particular, are very common, an expected result.

1.3.2   Clustering

In order to find similar neighbourhoods with the same lively atmosphere and a great number of restaurants, a clustering approach will be used. In clustering, the most important problem to tackle is the selection of the right number of clusters. Two approaches have been used to make this evaluation: the elbow method and the silhouette method. In the elbow method, we try to find the number of clusters for which a sharp change in the slope of the sum of squared errors (SSE) is observed. Unfortunately, in this case, this method is not very helpful since there is not an abrupt change in the slope of the SSE.
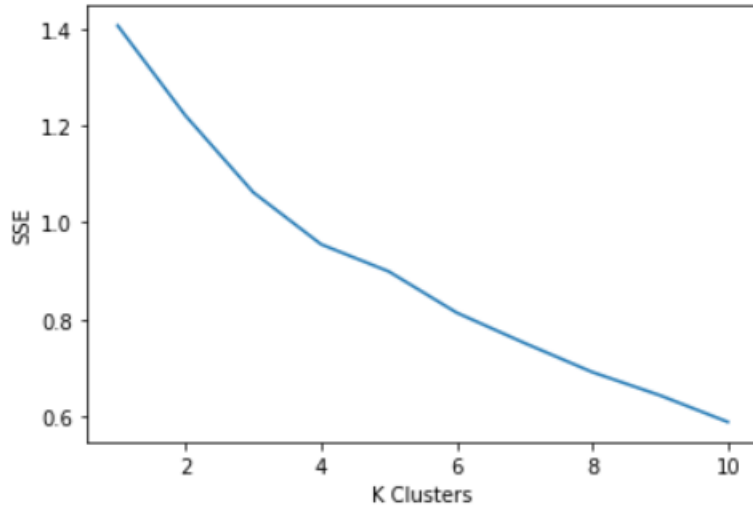
Therefore, a second method can be used, the silhouette method, that measures how similar a point is to its own cluster (cohesion) compared to other clusters (separation) with values ranging from -1 (worst) to +1 (best). In this case, we see that the maximum values are reached for 3, 2 and 6 clusters. We exclude k=2 and k=3 since they appear to low and they will not provide many insights. Therefore, we select k = 6.
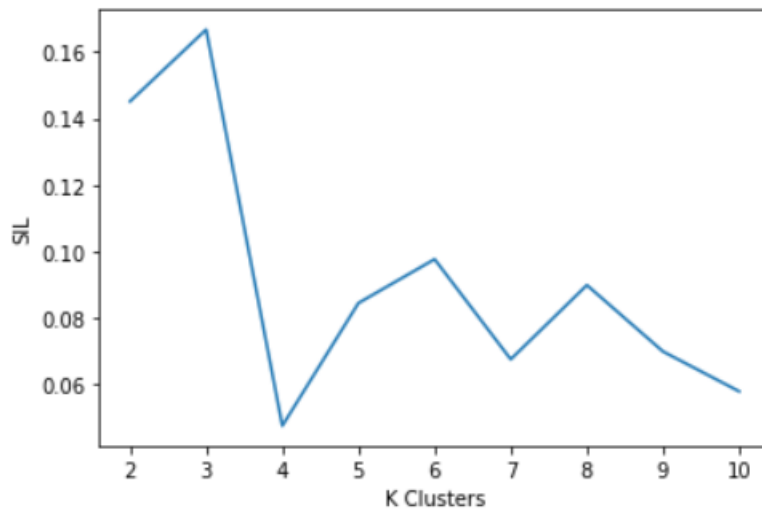
After running the k-means algorithm for k=6 we visualize the results on a map with colour code to distinguish neighbourhoods belonging to different clusters.
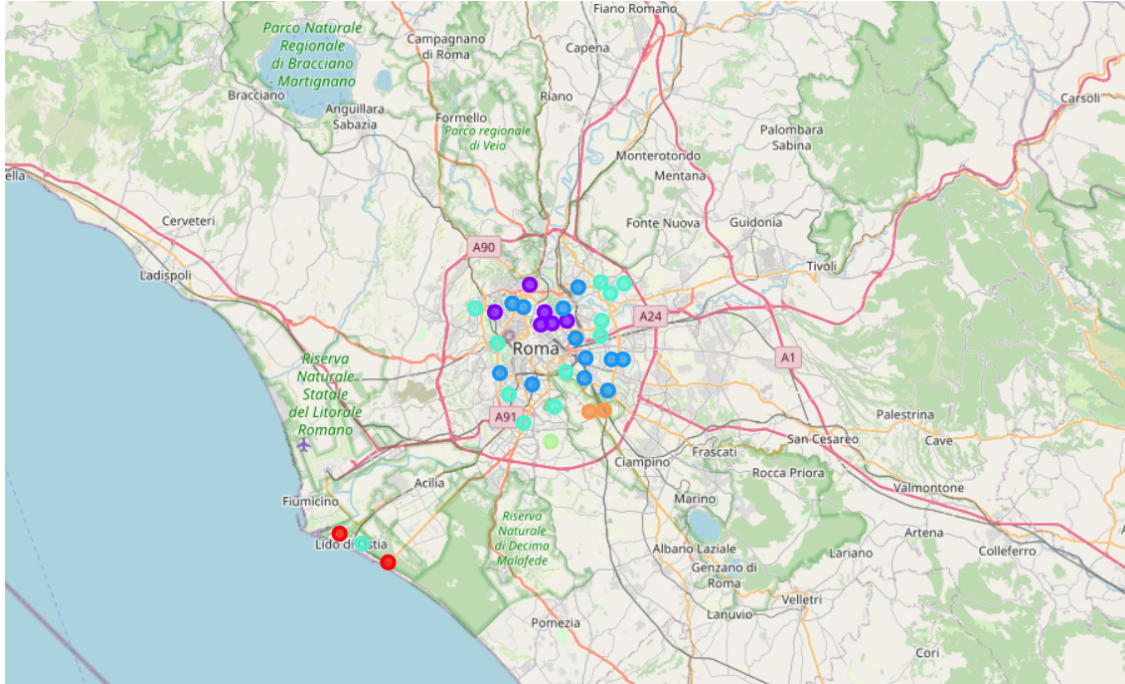
The six clusters identified by the algorithm can be categorized as follows:

- Cluster 1 - Beaches (red)
- Cluster 2 - Italian restaurants and hotels (centre-north of Rome, violet)
- Cluster 3 - Italian restaurants and pizza places (dark blue)
- Cluster 4 - Cafés (light blue)
- Cluster 5 - Hotels and Italian restaurants (suburbs of south Rome, green)
- Cluster 6 - Hotel and steakhouses (orange)

1.4 Results

The cluster that appears as the most suitable for the Brazilian investor to open his restaurants is the cluster number two, composed of the following neighbourhoods:

- Q. II Parioli
- Q. III Pinciano
- Q. IV Salario
- Q. V Nomentano
- Q. XIV Trionfale
- Q. XVIII Tor di Quinto

1.5 Discussion

The selected neighbourhoods are located in the centre-north of Rome and are considered to be fairly trendy and expensive. It is no surprise that they host a high number of restaurants and hotels. A further analysis, out of the scope of this study, would be to consider financial data regarding the rent or cost to acquire a place where to open a restaurant. Furthermore, a more detailed study would analyze the differences inside the same neighbourhood, for example, the nearness to bus stops or metro stations. Moreover, the Brazilian investor might have a preference for places

that are more popular during lunchtime or dinner time or both. Therefore, we can state that the neighbourhoods presented above represent a good starting point for further analysis.

1.6   Conclusion

In this study, we have analyzed the neighbourhoods of Rome in order to identify the best neighbourhoods where a Brazilian investor could open new Brazilian restaurants in the capital of Italy. The objective was to select the most promising neighbourhoods by analyzing data from the Foursquare database regarding the venues in 1 km radius from the centre of the neighbourhoods. The neighbourhoods with the highest number of restaurants and hotels appear to be most promising due to the high number of clients that go out in these areas for lunch or dinner. The analysis led to the conclusion that the best neighbourhoods where to open a restaurant are Parioli, Pinciano, Salario, Nomentano, Trionfale and Tor di Quinto. These neighbourhoods are located in the centre-north of Rome and are considered to be fairly trendy and expensive. This conclusion could be a good starting point a further analysis that could be performed by taking into account other aspects, for example, rent prices or public transports nearby.