Car sharing and neighborhoods in Milan, Italy

Finding new locations for parking areas for car sharing.

Paolo Cavadini, August 2020

Introduction

Milan, Italy, radically evolved during the last decade into the development of a sustainable touristic economy.

Car sharing is quite popular.

Suppose that potential stakeholders are looking for an area which is not in the high end of the real estate value, is as close as possible to the city center but at the same time has low density of parking areas for car sharing, what neighborhoods should they consider?

Definition of the problem: distance of neighborhood from the city center (~3Km). Number of existing parking areas for car sharing in the range of 300m from the neighborhood (max 1).

Data

[1] The website Immobiliare (www.immobiliare.it) publishes real estate values for Italy, aggregated from the regions through borough down to the street level (here). It was used to find the neighborhoods and the related real estate index.

[2] The Openstreetmap Nominatim API for geocoding (here) have been used to find the neighborhoods' latitude and longitude, and the coordinates of Milan center.

[3] The Foursquare Explore API have been used to retrieve the list of venues for the selected neighborhoods.

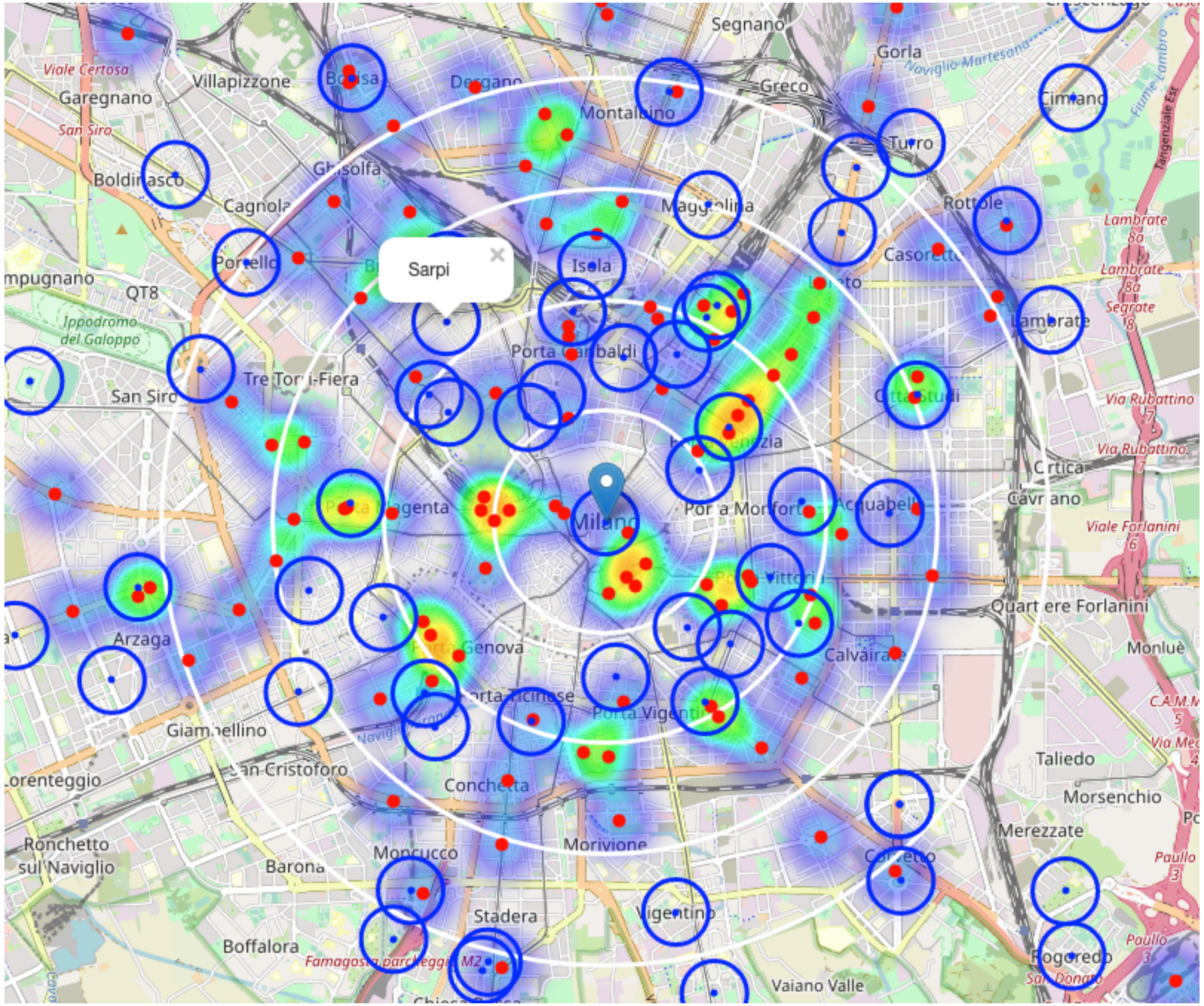
[4] The open dataset portal of the Municipality of Milan (https://dati.comune.milano.it/dataset) have been used to find the parking areas for car sharing (here).

Methodology
1st step

I am directing our efforts to detect neighborhoods of Milan that have low density of parking areas for car sharing.

In first step I have collected the required data: the neighborhoods obtained scarping data from a website and completing them using the **Nominatim API**, the parking areas for car sharing obtained from an open dataset published by the Municipality of Milan.

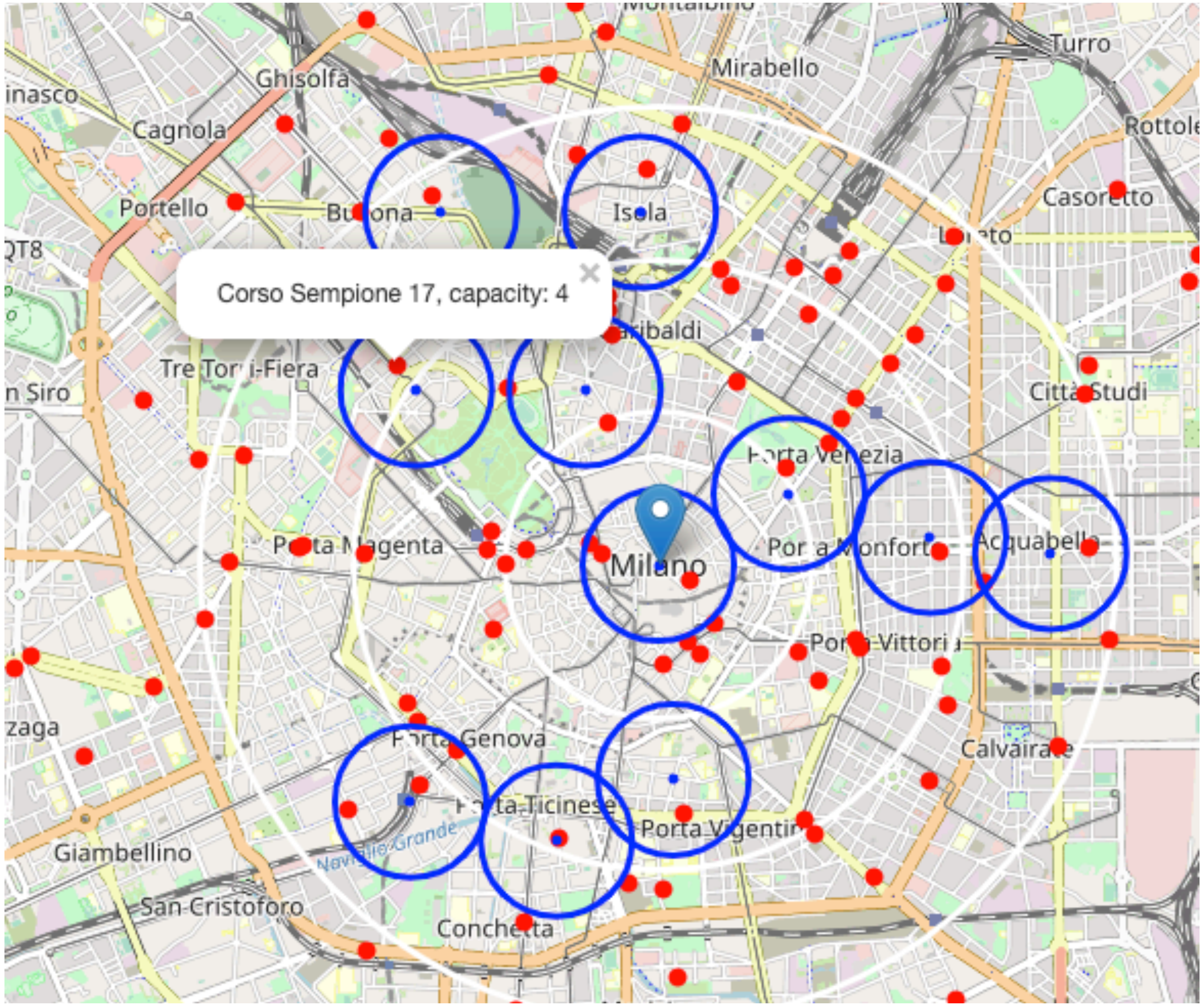
The map (**folium**) shows parking areas for car sharing (red dots), and the neighborhoods (blue circles, 300m radius). The heatmap shows the density of the parking areas.



Methodology
2nd step

In a second step, in agreement with the stakeholders I **limited my analysis to neighborhoods located ~3km from the city center, and with no more than 1 parking area for car sharing in the range of 300m.** For the selected neighborhoods I used the **Fourquare API** to obtain the venues.

The map shows how the 11 finally selected neighborhoods look like, in blue. A popup shows the neighborhood name. Each red dot is a parking area for car sharing, and a popup displays the parking name and capacity.



The venues

After finding 11 **neighborhoods candidates**, the **Foursquare API** provides the venues for each of the selected neighborhoods:

The **Foursquare API** found 705 venues.

```
[48]: Milan_venues
```

[48]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Susa	45.467484	9.223430	La Ciribiciaccola	45.469540	9.223775	Dessert Shop
1	Susa	45.467484	9.223430	Locanda del Menarost	45.465184	9.221471	Italian Restaurant
2	Susa	45.467484	9.223430	Pizzeria Ciak	45.467878	9.225039	Pizza Place
3	Susa	45.467484	9.223430	Fujiyama	45.469193	9.223766	Sushi Restaurant
4	Susa	45.467484	9.223430	Ristorante Il Grissino	45.471248	9.221944	Italian Restaurant
...	...	...	...	...	...	...	...
700	Ticinese	45.450596	9.181951	Ristorante Da Teresa	45.447448	9.178658	Italian Restaurant
701	Ticinese	45.450596	9.181951	Parco Via Tabacchi	45.447163	9.182943	Park
702	Ticinese	45.450596	9.181951	Loop	45.448168	9.186667	Café
703	Ticinese	45.450596	9.181951	Prince Cafè	45.448768	9.177262	Cocktail Bar
704	Ticinese	45.450596	9.181951	Q.P. Chiùpizza	45.447493	9.178139	Pizza Place

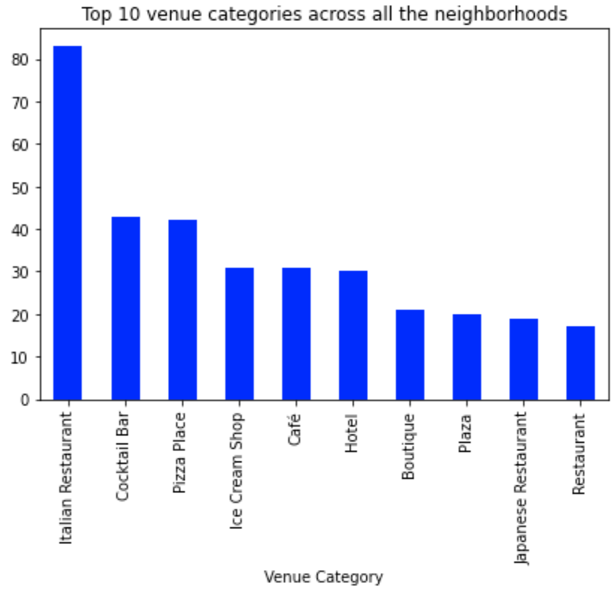
705 rows × 7 columns

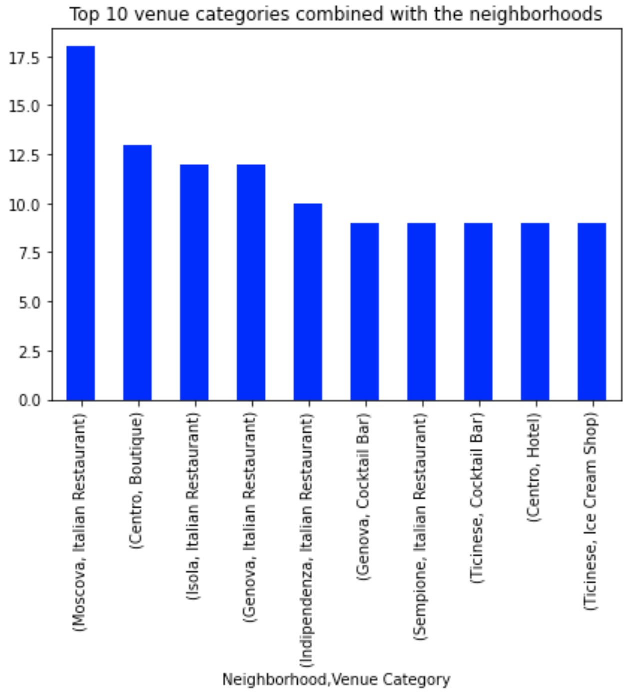
```
[49]: print('There are {} uniques categories.'.format(len(Milan_venues['Venue Category'].unique())))
There are 147 uniques categories.
```

I finally found **147 unique venue categories**, which looks ok for clustering.

Analysis

The top 10 venue categories, and the **top 10 neighborhood/venue category combinations**. Food dominates with Italian restaurants and cocktail bars, but we also have hotels and boutiques. We should be able to cluster **neighborhoods characterized by lifestyle, shopping, and hospitality districts.**





	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Cenisio	Pizza Place	Hotel	Gym / Fitness Center	Thai Restaurant	Seafood Restaurant
1	Centro	Boutique	Hotel	Italian Restaurant	Plaza	Ice Cream Shop
2	Genova	Italian Restaurant	Cocktail Bar	Pizza Place	Seafood Restaurant	Bar
3	Indipendenza	Italian Restaurant	Ice Cream Shop	Pizza Place	Café	Dessert Shop
4	Isola	Italian Restaurant	Pizza Place	Café	Bistro	Cocktail Bar
5	Moscova	Italian Restaurant	Ice Cream Shop	Wine Bar	Café	Restaurant
6	Palestro	Boutique	Cocktail Bar	Women's Store	Plaza	Art Gallery
7	Quadronno	Wine Bar	Hotel	Pizza Place	Salad Place	Café
8	Sempione	Italian Restaurant	Pizza Place	Cocktail Bar	Japanese Restaurant	Ice Cream Shop
9	Susa	Italian Restaurant	Hotel	Ice Cream Shop	Café	Convenience Store
10	Ticinese	Ice Cream Shop	Cocktail Bar	Italian Restaurant	Café	Pizza Place

Analysis

Onehot encoding to obtain the dataframe for clustering, and the 5 top common venues for each of the 11 selected neighborhoods.

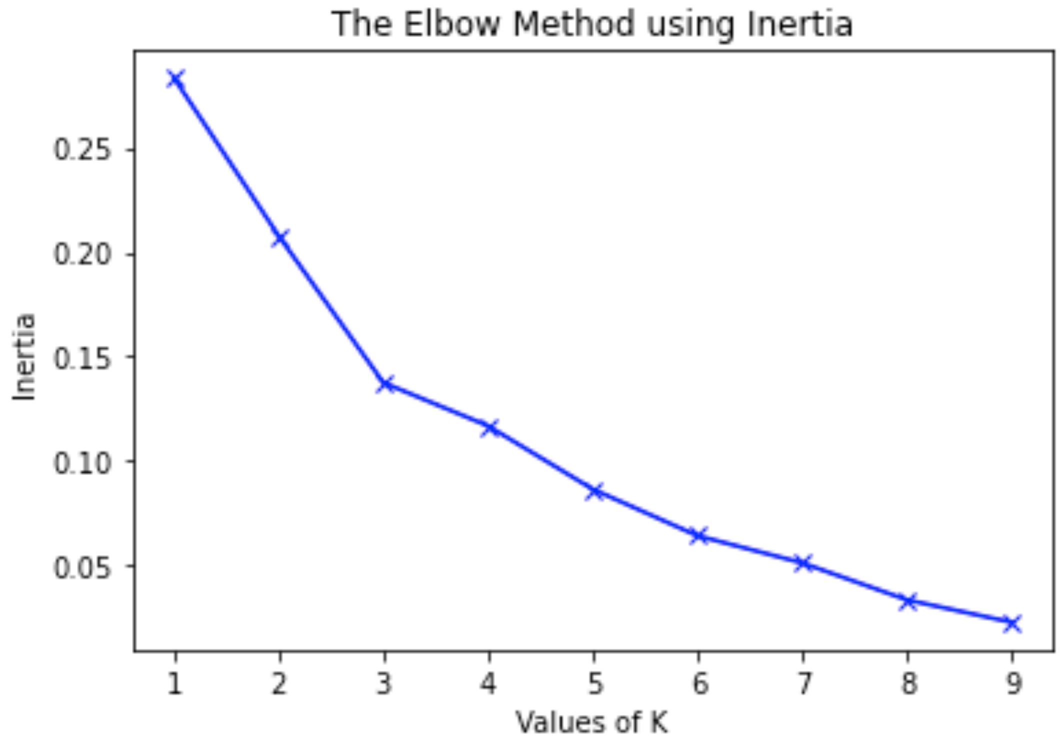
Methodology
3rd step

In the last step I will create clusters (using **k-means**) of neighborhoods, I will present them on a map to identify general zones which should be a starting point for further detailed exploration for the stakeholders. The map will show the clusters together with a heatmap which will represent the density of the parking areas for car sharing.

Clustering

I used GridSearCV to find the optimal parameters, and the Elbow method with inertia to confirm the number of clusters.

```
{'init': 'random', 'n_clusters': 3, 'random_state': 0}
```



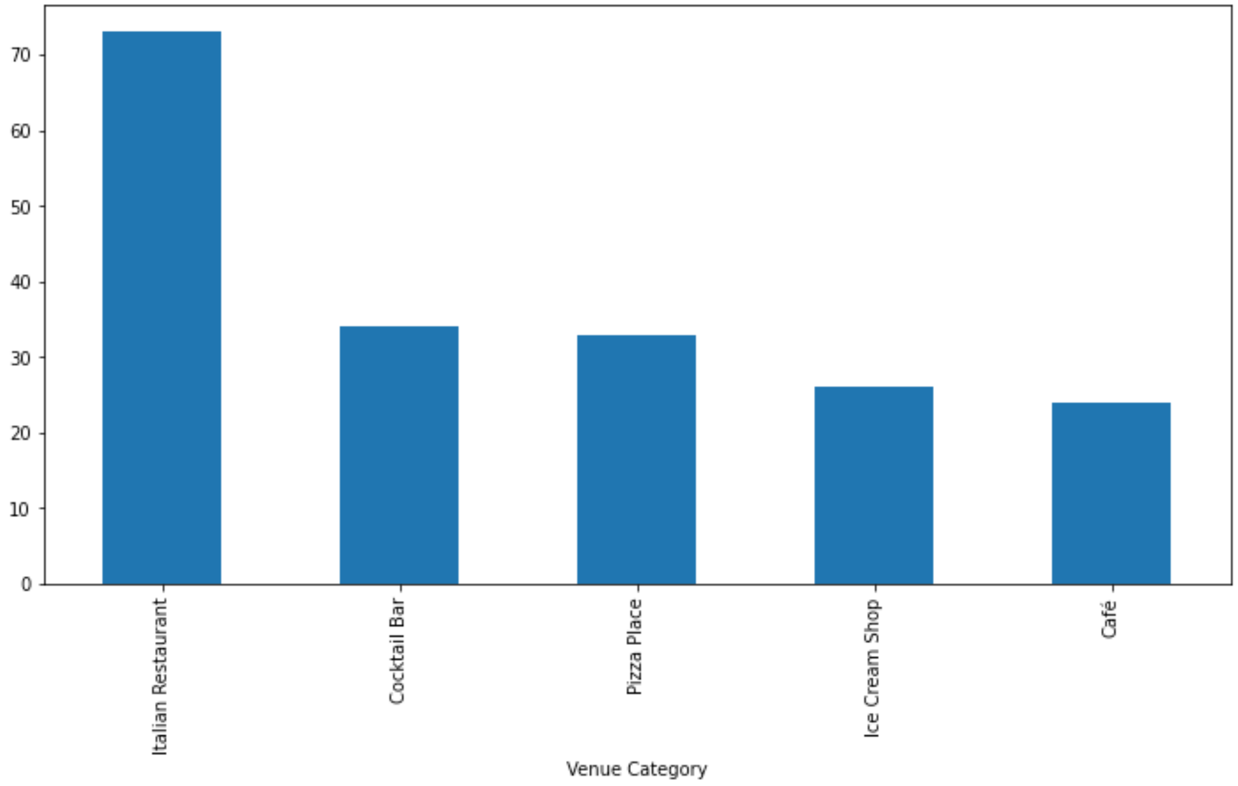
Clustering with K=3.

Results

The clustered neighborhoods shown on a **folium** map and a **heatmap** to show the density of parking areas for car sharing. Each neighborhood shows a popup to indicate the name of the neighborhood, the cluster to which it belongs, and the real estate index.

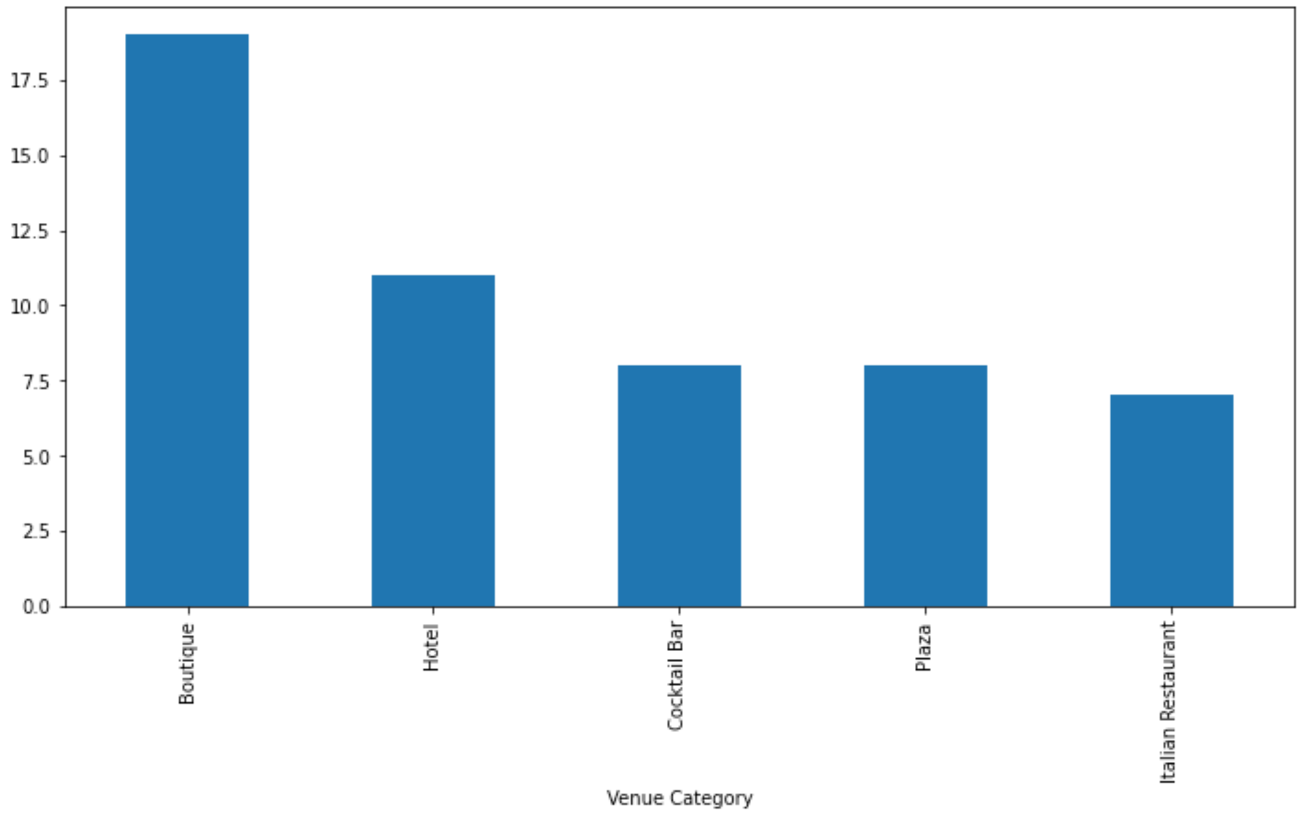
Results

Cluster 0 is primarily a **lifestyle area.** Looking at the map, the neighborhoods **Sempione, Indipendenza, Susa and Isola are suburbs and show a lower real estate index. Moscova is central and expensive, Porta Genova and Ticinese are still central and lifestyle.**



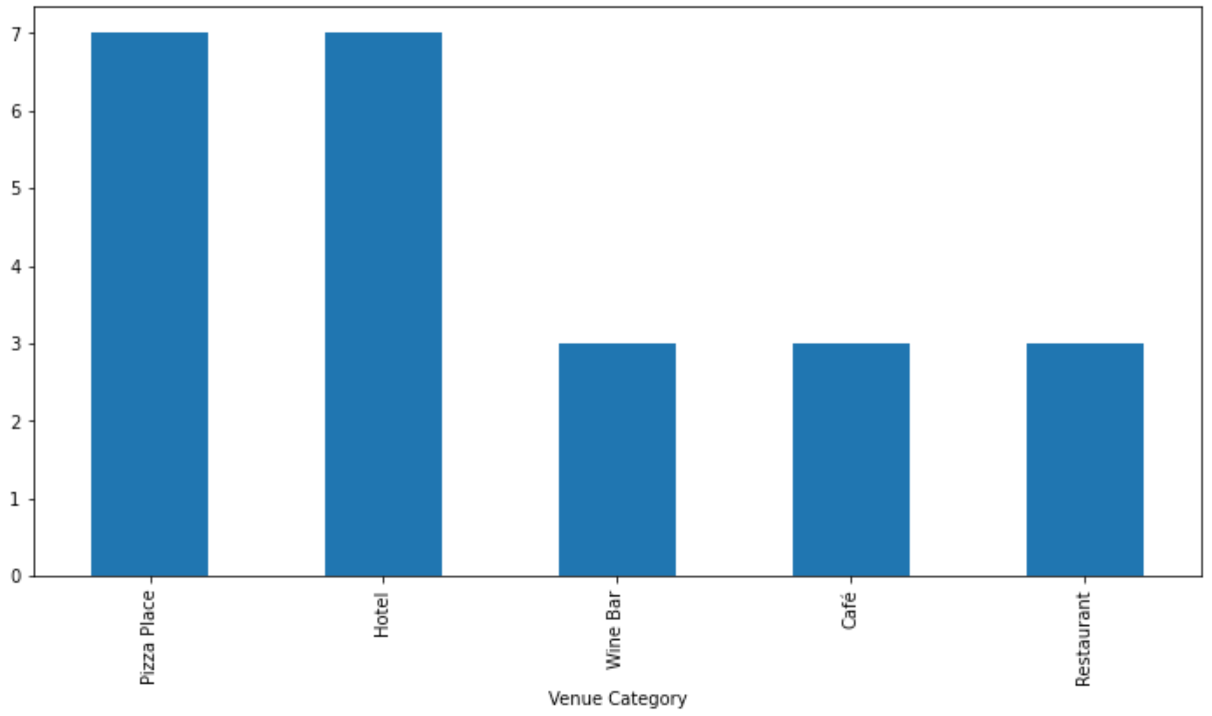
Results

Cluster 1 is primarily a **shopping district**, it's smaller in terms of venues that cluster Zero. The neighborhoods are clearly central and have a high real estate index.



Results

Cluster 2 is primarily a **suburb hotels area**, it's smaller in terms of venues than cluster Zero and One. The neighborhoods are different because *Cenisio* is suburbs and has a low real estate index, while *Quadronno* is still central and has a high real estate index.



Results and discussion

The analysis shows that although parking areas for car sharing are present in the center of Milan (~3Km), there are zones of low parking spot density, no more than 1 in range of 300m, fairly close to city center. The Highest concentration of parking areas was detected East and close to the city center. I focused my attention to areas East and less close to the city center (Indipendenza, Susa) because they also combine a lower real estate value. Similarly, other neighborhoods were identified as potentially interesting North from the city center (Cenisio, Sempione, Isola).

The neighborhood of Cenisio is close to hotels and intercepts traffic arriving from North, North West (Malpensa airport, directions arriving from the A8 and A9 highways -Switzerland). The neighborhood of Sempione and Isola (lifestyle) are near Cenisio, and Sempione is actually slightly closer to the city center. The neighborhoods of Indipendenza and Susa, both lifestyle, intercept traffic arriving from East (Linate airport, Bergamo airport and directions along the A4 highway -Venice).

Those neighborhoods candidates have been clustered to create zones of interest. They will become starting points for more detailed local analysis based on other factors. As I learned during the course, this does not imply that those zones are actually optimal locations for a new parking area. It is possible that there is a very good reason for small number of parking for car sharing in any of those areas, regardless of lack of competition in the area. Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in locations which have not only no nearby competition, but also other factors taken into account and all other relevant conditions met.

Conclusions

Purpose of this project was to identify Milan areas within 3Km from the city center with no more than 1 parking area for car sharing within a 300m radius. This will help the stakeholders to narrow down the search for optimal location for new parking areas. Clustering of those neighborhoods was then performed in order to create major zones of interest which will be used as starting points for final exploration by stakeholders.

A final decision on optimal parking areas location will be made by stakeholders based on specific characteristics of neighborhoods in every recommended zone, taking into consideration additional factors.