



That Italian Place

From A to Z about venues in Milan, Italy

Introduction

Milan has a vast array of activities to offer that one can do, be it a tourist or a local who wants to explore new places or food.

The purpose of this project is to analyze the most common places in Milan and bundle them into distinct categories based on a list of properties. Every category would fit a single or a series of venue types that are similar between each other, offering the option to choose for a venue by the kind and location.

Success Criteria

- Generate a good user recommendation of a district.
- User's activities/places of choice are frequently met in the chosen area of Milan.
- Vice-versa – suggest to the user the most frequent activities/places for a specific district.

About Milan

- Milan, the capital of Lombardy region in Italy, has a population of 1.3 million people
- Biggest industrial city of Italy.
- Milan has an ancient city centre with high and interesting buildings and palazzos.
- Famous for restaurants recommended by design-world insiders, chic bars where you can sample the classic Milanese aperitivo.
- The city of top designer boutiques and typical Italian local shops.



Data

- The analysis will be implemented for all the districts in the city of Milan, the source of this being the Wikipedia web-page that contains this information.
- Three frameworks are integrated into the project for the subsequent purposes:
 - Beautiful Soup.
 - HERE.
 - Foursquare API.

Pages in category "Districts of Milan"

The following 76 pages are in this category, out of 76 total. This list may not reflect recent changes ([learn more](#)).

A

- [Affori](#)
- [Assiano](#)

B

- [Baggio \(district of Milan\)](#)
- [Barona](#)
- [Bicocca \(district of Milan\)](#)
- [Bovisa](#)
- [Bovisasca](#)
- [Brera \(district of Milan\)](#)
- [Bruzzano](#)

C

- [Calvairate](#)
- [Centro Direzionale di Milano](#)
- [Chiaravalle \(district of Milan\)](#)
- [Chinatown, Milan](#)
- [Cimiano](#)
- [Città Studi](#)
- [Comasina](#)
- [Crescenzago](#)

D

- [Dergano](#)

F

- [Figino \(district of Milan\)](#)
- [Forlanini \(district of Milan\)](#)

G

- [Gorla](#)
- [Gratosoglio](#)
- [Greco \(district of Milan\)](#)

L

- [Lambrate](#)
- [Lampugnano](#)

M

- [Milano Santa Giulia](#)
- [Monluè](#)
- [Morivione](#)
- [Muggiano \(district of Milan\)](#)

N

- [Niguarda](#)
- [Noseda](#)

O

- [Ortica](#)

P

- [Ponte Lambro \(district of Milan\)](#)
- [Porta Garibaldi \(Milan\)](#)
- [Porta Genova](#)
- [Porta Lodovica](#)
- [Porta Magenta](#)
- [Porta Monforte](#)
- [Porta Nuova \(Milan\)](#)
- [Porta Romana \(Milan\)](#)
- [Porta Sempione](#)

Porta Volta

- [Portello \(district of Milan\)](#)
- [Prato Centenaro](#)
- [Precotto](#)

Q

- [QT8](#)
- [Quadrilatero della moda](#)
- [Quartiere Feltre](#)
- [Quartiere Musocco](#)
- [Quarto Cagnino](#)
- [Quarto Oggiaro](#)
- [Quinto Romano](#)
- [Quintosole](#)

R

- [Rogoredo](#)
- [Ronchetto sul Naviglio](#)
- [Rosario](#)

S

- [San Cristoforo sul Naviglio \(district of Milan\)](#)
- [San Siro \(district\)](#)
- [Segnano](#)
- [Stazione di Milano Centrale](#)

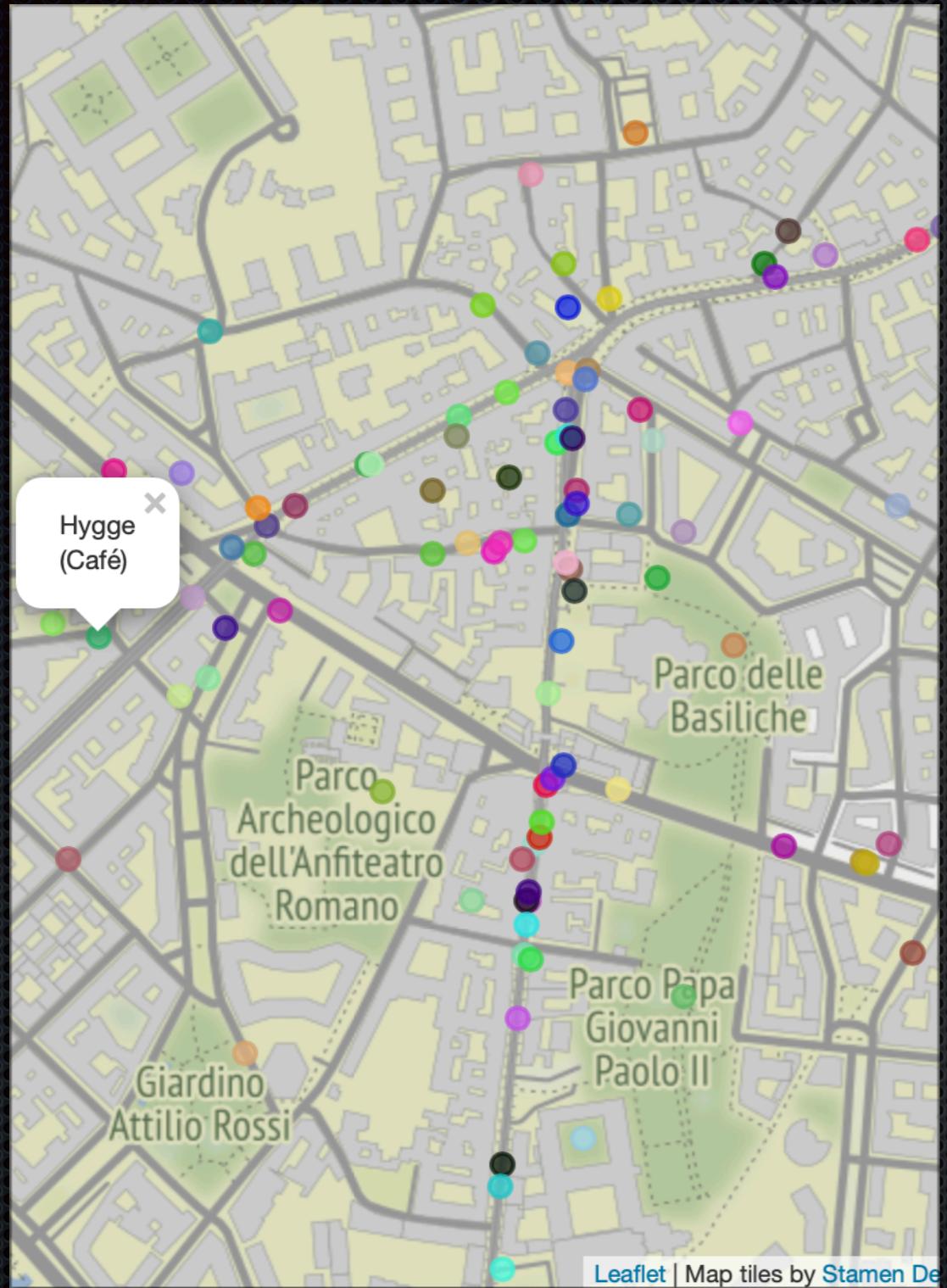
T

- [Taliedo](#)
- [Trenno](#)
- [Turro](#)

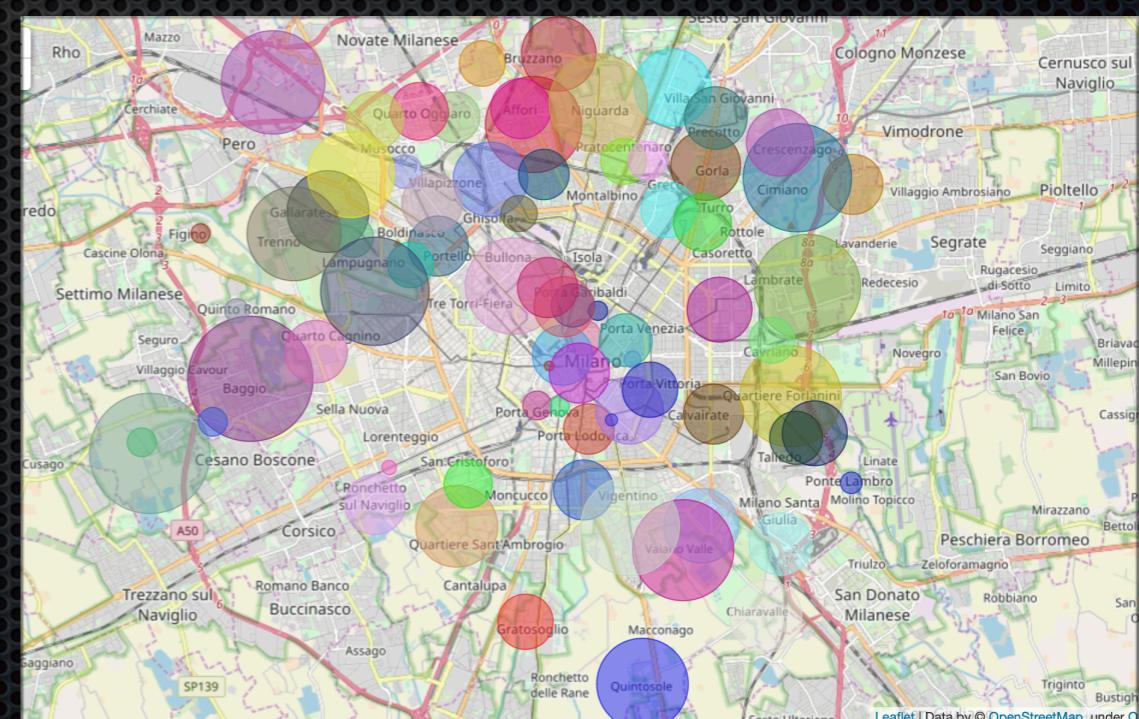
Districts of Milan from Wikipedia page (source – Wikipedia)

Processing Stages

- Web scraping (Beautiful Soup and custom cleaning methods).
- Geocoding (HERE).
- Venues extraction (Foursquare API).



	District Name	District Latitude	District Longitude	Venue Name	Venue Category	Venue Latitude	Venue Longitude
2020	Porta Vigentina, Milan, Italy	45.45500	9.19577	Viva	Salad Place	45.456250	9.195423
437	Chinatown, Milan, Italy	45.46796	9.18178	Six Inch Hair & Spa	Salon / Barbershop	45.470829	9.185033
1826	Porta Ticinese, Milan, Italy	45.45771	9.18096	Scout	Clothing Store	45.459396	9.180609
1766	Porta Tenaglia, Milan, Italy	45.47766	9.18224	Gam	Korean Restaurant	45.481481	9.184667
1885	Porta Ticinese, Milan, Italy	45.45771	9.18096	Tokuyoshi	Italian Restaurant	45.458383	9.176125
2751	Vigentino, Milan, Italy	45.43373	9.20105	Esselunga	Supermarket	45.443121	9.197132
1277	Porta Magenta, Milan, Italy	45.46582	9.17800	Gay Odin	Chocolate Shop	45.466187	9.180801
798	Greco, Milan, Italy	45.49702	9.21213	Alvin's Bar Pasticceria	Café	45.494939	9.207228
2489	San Siro, Milan, Italy	45.48075	9.12824	Tennis Lido	Tennis Court	45.480909	9.142764
614	Dergano, Milan, Italy	45.50412	9.17647	Ac Garibaldina 1932	Soccer Field	45.505813	9.172451
2767	Vigentino, Milan, Italy	45.43373	9.20105	Bar Chopin	Coffee Shop	45.425268	9.204760
2374	Rogoredo, Milan, Italy	45.43017	9.24401	Igiban	Japanese Restaurant	45.426923	9.250569
737	Goria, Milan, Italy	45.50591	9.22265	Seven - La Casa dei Ciliegi	Italian Restaurant	45.503713	9.222976
427	Chinatown, Milan, Italy	45.46796	9.18178	L.O.V.E. Maurizio Cattelan	Monument / Landmark	45.464834	9.183314
2765	Vigentino, Milan, Italy	45.43373	9.20105	Barrio Alto	Gastropub	45.441364	9.201348
2100	Porta Vittoria, Milan, Italy	45.46106	9.20677	Piadiniamo	Piadineria	45.456505	9.205159
2638	Trenno, Milan, Italy	45.49225	9.10518	McCafé	Fast Food Restaurant	45.496345	9.113852
732	Giambellino-Lorenteggio, Milan, Italy	45.44548	9.13233	Fermata ATM Piazza Tirana / San Cristoforo [49]	Bus Stop	45.444730	9.130393
1615	Porta Romana, Milan, Italy	45.45691	9.20096	Osteria Delizie del Mare	Seafood Restaurant	45.452097	9.203307
1505	Porta Nuova, Milan, Italy	45.47685	9.19192	Panini Durini	Sandwich Place	45.475532	9.194549
242	Brera, Milan, Italy	45.47149	9.18773	La Prosciutteria	Sandwich Place	45.474152	9.183449

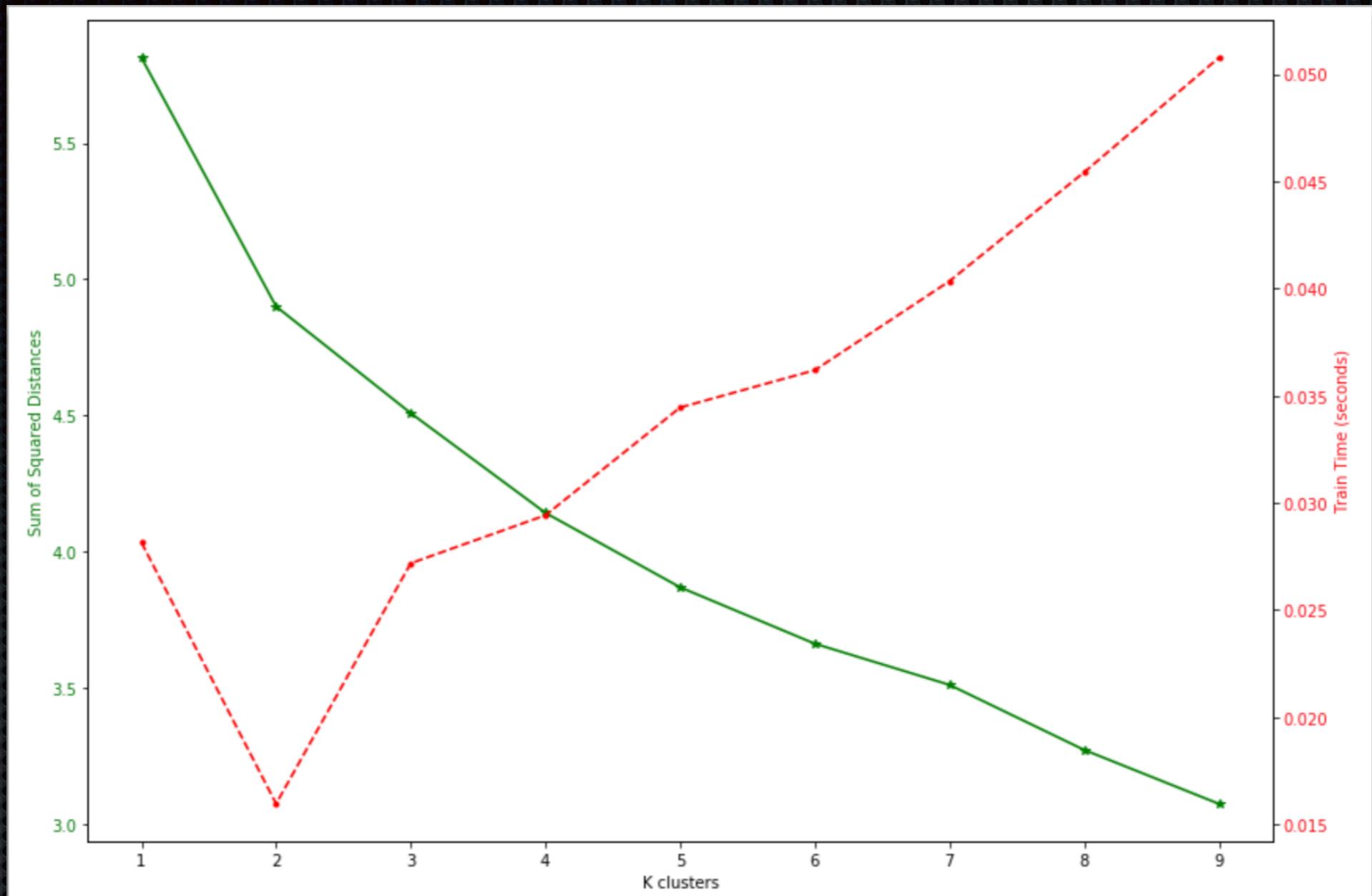


Methodology

- Analytic approach:
 - Milan has 75 districts in total.
 - Venues are filtered from Foursquare.
 - Clustering the venues by category type.
- Algorithm:
 - K-means.
 - Elbow method to find the optimal K-value.

Cluster Label	Accessories Store	Adult Education Center	African Restaurant	Airport	Airport Terminal	American Restaurant	Amphitheater	Argentinian Restaurant	Art Gallery	...	Video Game Store	Vietnamese Restaurant	Volleyball Court	Watch Shop
0	0.001311	0.002621	0.001311	0.001311	0.002621	0.000000	0.001311	0.000000	0.005242	...	0.001311	0.001311	0.003932	0.000000
1	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000
2	0.003405	0.000000	0.000973	0.000000	0.000000	0.002432	0.000000	0.002432	0.008268	...	0.000000	0.000000	0.000000	0.000486
3	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000
4	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.000000	0.000000	0.000000

- The project scope implies clustering the main venues from Milan's districts using Machine Learning algorithms.
- Venue category type serves as the main feature when running the clustering algorithm.
- The custom clustering algorithm is composed of several stages.
- The clusters capture all the venue categories that are similar between each other, based on the type, frequency and quantity per district.



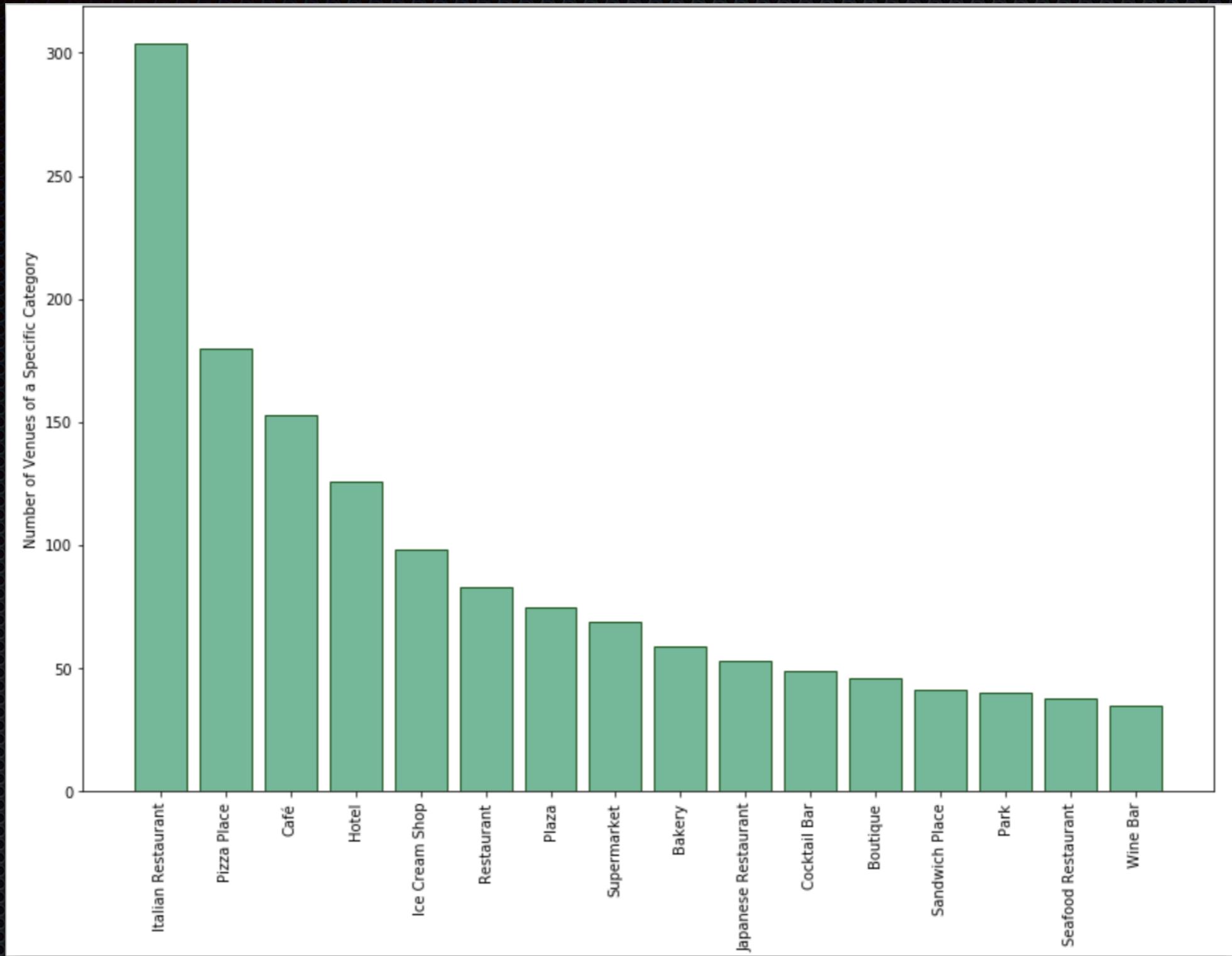
Elbow method results when running the data from Milan's venues.

Results

- Using Milan's districts and venues, these were clustered using K-means algorithm into groups based on the mean occurrence of the venue category in each district.
- The clustered data has been utilized to create the most frequent venue types per each district, represented by a table.
- An interesting aspect is to analyze all the venues throughout the entire city, so a general overview of the most frequent venue categories in Milan can be observed.

District Name	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	11th Most Common Venue
0 Affori, Milan, Italy	Hotel	Pizza Place	Park	Soccer Field	Italian Restaurant	Café	Skate Park	Shipping Store	Salon / Barbershop	Rock Club	Cocktail Bar
1 Assiano, Milan, Italy	Pizza Place	Yoga Studio	Filipino Restaurant	Event Space	Fabric Shop	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Film Studio	Dumpling Restaurant
2 Baggio, Milan, Italy	Supermarket	Café	Plaza	Italian Restaurant	Bus Station	Campground	Park	Japanese Restaurant	Bar	Volleyball Court	Pizza Place
3 Barona, Milan, Italy	Soccer Field	Café	Japanese Restaurant	Theater	Bakery	Athletics & Sports	Trattoria/Osteria	Brewery	Tennis Stadium	Pub	Event Space
4 Bicocca, Milan, Italy	Café	Italian Restaurant	Sandwich Place	Steakhouse	Sushi Restaurant	Pizza Place	Hotel	Multiplex	Restaurant	Plaza	Supermarket
5 Bovisa, Milan, Italy	Italian Restaurant	Café	Pizza Place	Snack Place	Piadineria	Plaza	Ice Cream Shop	Smoke Shop	Sicilian Restaurant	Kebab Restaurant	Steakhouse
6 Bovisasca, Milan, Italy	Clothing Store	Shoe Store	Soccer Field	Restaurant	Café	Cosmetics Shop	Pizza Place	Supermarket	Park	Health & Beauty Service	Chinese Restaurant
7 Brera, Milan, Italy	Italian Restaurant	Hotel	Boutique	Ice Cream Shop	Wine Bar	Plaza	Restaurant	Café	Theater	Pizza Place	Arts & Crafts Store
8 Bruzzano, Milan, Italy	Italian Restaurant	Bakery	Pizza Place	Football Stadium	Gift Shop	Diner	Train Station	Gym / Fitness Center	Bus Station	Ice Cream Shop	Fish & Chips Shop
9 Calvairate, Milan, Italy	Italian Restaurant	Park	Pizza Place	Supermarket	Hotel	Ice Cream Shop	Bakery	Market	Other Nightlife	Tanning Salon	Bar
10 Centro Direzionale di Milano, Milan, Italy	Hotel	Italian Restaurant	Outdoors & Recreation	Lake	Rest Area	Café	Farmers Market	Fabric Shop	Falafel Restaurant	Farm	Fast Food Restaurant

Most frequent venue type per each district.



Venue frequency in Milan based on the category type.

Conclusion

- Analysis performed on data returned by Foursquare, that is limited only to the top venues in a specific region.
- To update the results the notebook must be run again.
- Milan has a vast selection of Italian restaurants and pizza places – this can be further explored.
- Framework limitations – HERE geocoding could not identify the location of one district in Milan.

Thank you!