

Monza e Brianza restaurant choice - Applied Data Science Capstone Project

Alessandro Fumagalli

December, 2020

Contents

Introduction.....	1
Background.....	1
Problem	1
Data	2
Data Source	2
Data cleaning.....	3
Methodology	4
Exploratory Data analysis	4
Clustering.....	4
Result.....	5
Discussion	7
Conclusion	7

Introduction

Background

I live in Monza, in Northern Italy: Monza is the capital of the "Monza e Brianza" province. For the sake of simplicity let's call the province "MB" for short, like the car plate. MB territory encompasses more or less a circle of 12 Km as radius and it is well placed between Milan and the Lake of Como district, so it attracts many tourists and many business men as well that use MB province as a reference point for their affairs in Milan or spend a couple of days around the lake. Moreover MB hosts different fairs during years and, of course, the National Italian Formula 1 GP. The problem I am going to face and propose as final capstone of this course is to help tourist and business men in finding restaurants of different type in the whole MB province: not only national ones (Italian) but also ethnic restaurants and how they are clustered and distributed over the territory in order to help people to orient themselves to a particular zone of MB province for their living.

Problem

The business question is: how are distributed the restaurants types in MB province, so can meet better foreign people taste and gastronomy palates? This report (with result), the presentation along with the code of this capstone project will be published on my personal https://github.com/alexfum75/Coursera_Capstone account, a hosting service for software project.

Data

First step is where to retrieve the data to work with for this project. The source must be reliable in order to get a valid prediction for our stakeholders (tourist and business men).

Data Source

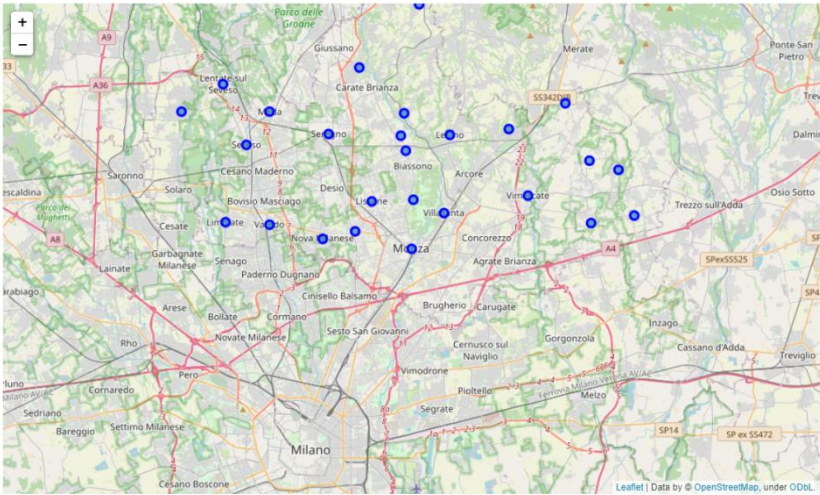
First of I have to recover all the minimum information about the municipality of MB: that is the list of town with their code and geographical position: this can be found following this link: <https://zip-codes.nonsolocap.it/lombardy/20-cap-province-of-monza-and-brianza>

Home > Lombardy ZIP Codes > ZIP codes province of Monza and Brianza			
ZIP regions			
Abruzzo			
Aosta Valley			
Apulia			
Basilicata			
Calabria			
Campania			
Emilia-Romagna			
Friuli Venezia Giulia			
Lazio			
Liguria			
Lombardy			
Marche			
Molise			
Piedmont			
Sardinia			
ZIP Codes province of Monza and Brianza			
List of the Postal codes of the 55 municipalities of province of Monza and Brianza. Links on zip codes lists municipalities and other outlying areas (frazioni) with the same postcode.			
Luca Santambrogio President		Prov. Monza and Brianza Tuttitalia.it	Population 878,267 inh.
Municipality		ZIP	Municipality
Agrate Brianza	20864	Lentate sul Seveso	20823
Alcuzio	20886	Lesmo	20855
Albate	20847	Limbrate	20812
Arcore	20862	Lissone	20851
Barlassina	20825	Macherio	20846
Bellusco	20882	Meda	20821
Bernareggio	20881	Mezzago	20883
Besana in Brianza	20842	Misinto	20826
Blassono	20853	Monza	20900
Bovisio-Masciago	20813	Muggiò	20835
Brinco	20836	Nova Milanese	20834

From this list I removed the columns of the DataFrame not need for our scope, like the ZIP code, in order work only with municipality name and geographical coordinates. Now we need a repository where to find all point of interest, and venues of MB distributed over all municipalities.

	Municipality	City	Latitude	Longitude
0	Lentate sul Seveso	(45.6784049, 9.1181696)	45.678405	9.118170
1	Lesmo	(45.6489742, 9.3056521)	45.648974	9.305652
2	Limbrate	(45.5986047, 9.119731)	45.598605	9.119731
3	Lissone	(45.6107705, 9.2408409)	45.610771	9.240841
4	Macherio	(45.640116, 9.269371)	45.640116	9.269371
5	Meda	(45.6625046, 9.1567532)	45.662505	9.156753
6	Mezzago	(45.628977, 9.44485)	45.628977	9.444850
7	Misinto	(45.662375, 9.083941)	45.662375	9.083941
8	Monza	(45.5834418, 9.2735257)	45.583442	9.273526
9	Muggiò	(45.5931952, 9.2271378)	45.593195	9.227138
10	Nova Milanese	(45.58929, 9.2003403)	45.589290	9.200340
11	Ornago	(45.5984036, 9.4222142)	45.598404	9.422214
12	Renate	(45.724378, 9.279944)	45.724378	9.279944
13	Roncello	(45.6024427, 9.4576108)	45.602443	9.457611
14	Ronco Briantino	(45.6675633, 9.4006983)	45.667563	9.400698
15	Seregno	(45.649723, 9.2053919)	45.649723	9.205392
16	Seveso	(45.6434655, 9.1373787)	45.643465	9.137379
17	Sovico	(45.6485558, 9.2648372)	45.648556	9.264837
18	Sulbiate	(45.6339993, 9.4206842)	45.633999	9.420684
19	Triuggio	(45.6613638, 9.2674406)	45.661364	9.267441
20	Usmate Velate	(45.65245845, 9.354341941243387)	45.652458	9.354342
21	Varedo	(45.5974437, 9.1562926)	45.597444	9.156293
22	Vedano al Lambro	(45.611508, 9.2750718)	45.611508	9.275072
23	Veduggio con Colzano	(45.73362975, 9.262512661573547)	45.733630	9.262513
24	Verano Brianza	(45.6881861, 9.230477)	45.688186	9.230477
25	Villasanta	(45.6039807, 9.3005709)	45.603981	9.300571
26	Vimercate	(45.6139627, 9.37006)	45.613963	9.370060

I will exploit Foursquare, a social network based on geo localization available for mobile device. The developer program allowed me to retrieve the information I need via APIs system (via json file) supply to the developer with its program: <https://developer.foursquare.com/docs/places-api/endpoints/> their technology is used by different companies like Apple's Maps, Uber, Twitter and it is a big repository. Below the distribution of the municipality on MB territory with Folium



Data cleaning

Foursquare give me a lot of venues: gym, bar, pools, restaurants ... Since we are interested only in restaurant I filtered the data in order to get only restaurants in all MB municipalities. Although Foursquare contains data we need, I noticed in the filtered data they are not fully compliant for my work: some data regarding restaurant do not have the necessary consistency we need for our scope. I found about 50 venues (records) misclassified by Foursquare: here are three of them:

153	Lissone	45.610771	9.240841	Kaori sushi	45.609535	9.240886	Restaurant	Misclassified
176	Lissone	45.610771	9.240841	Locanda La Martina	45.631380	9.231245	Restaurant	
205	Lissone	45.610771	9.240841	Pura Brace	45.619190	9.270444	Restaurant	
249	Macherio	45.640116	9.269371	Pura Brace	45.619190	9.270444	Restaurant	Misclassified
260	Macherio	45.640116	9.269371	Pleasel	45.645187	9.306209	Restaurant	
261	Macherio	45.640116	9.269371	Trattoria La Ca'	45.631010	9.304794	Restaurant	
278	Meda	45.662505	9.156753	La Locanda di Mr Brown	45.653522	9.189366	Restaurant	Misclassified
295	Meda	45.662505	9.156753	Tri Basei	45.666292	9.149790	Restaurant	
307	Mezzago	45.628977	9.444850	Bar Valentino	45.620777	9.418321	Restaurant	
349	Misinto	45.662375	9.083941	Osteria del Borgo	45.671070	9.085122	Restaurant	Misclassified
430	Monza	45.583442	9.273526	Cascina del Sole	45.596450	9.278148	Restaurant	
444	Monza	45.583442	9.273526	Altrove	45.577836	9.266700	Restaurant	
447	Monza	45.583442	9.273526	Atmosfera	45.572796	9.265312	Restaurant	Misclassified
511	Muggiò	45.593195	9.227138	Kaori sushi	45.609535	9.240886	Restaurant	
569	Nova Milanese	45.589290	9.200340	Redsky	45.605050	9.173604	Restaurant	
594	Ornago	45.598404	9.422214	Bar Valentino	45.620777	9.418321	Restaurant	Misclassified
627	Renate	45.724378	9.279944	Locanda dell'Erba Matta	45.719010	9.277160	Restaurant	
629	Renate	45.724378	9.279944	Trattoria Maurizi	45.748698	9.270067	Restaurant	
687	Ronco Briantino	45.667563	9.400698	Ristorante Il Granero	45.662636	9.418066	Restaurant	Misclassified
694	Ronco Briantino	45.667563	9.400698	Bar Trattoria Pizzeria Castello	45.659407	9.379736	Restaurant	
719	Seregno	45.649723	9.205392	La Locanda di Mr Brown	45.653522	9.189366	Restaurant	
750	Seregno	45.649723	9.205392	Locanda La Martina	45.631380	9.231245	Restaurant	Misclassified
798	Seveso	45.643465	9.137379	Ristorante Nesis	45.621224	9.150226	Restaurant	
860	Sovico	45.648556	9.264837	Ses Culon	45.639383	9.285753	Restaurant	
871	Subiate	45.633999	9.420684	Bar Valentino	45.620777	9.418321	Restaurant	Misclassified
913	Truggio	45.661364	9.267441	Ses Culon	45.639383	9.285753	Restaurant	
921	Usmate Velate	45.652458	9.354342	Kangaroo	45.635265	9.347463	Restaurant	
930	Usmate Velate	45.652458	9.354342	Bar Trattoria Pizzeria Castello	45.659407	9.379736	Restaurant	

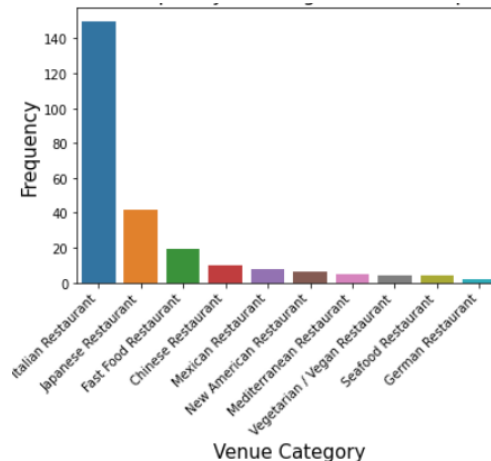
Japanese, Australian, Argentinian, New American and See Food restaurants where classified as Italian ones or generic restaurant. In order to avoid wrong insights from the data I decided to create a python script that works on the DataFrame in order to properly classify them. The script was the right choice since we have few venues so it was not a big problem.

Methodology

In this section, I proceed with a first exploratory data analysis and then apply the methodology to yield the result: in particular we clustered data according to the restaurant type.

Exploratory Data analysis

I exploit Folium library (<https://python-visualization.github.io/folium/>) to have an overview about how and where restaurants are distributed the municipalities in MB according with its area size. Just to play around, I filtered the top 10 restaurant type from the DataFrame and then identified 259 different restaurants in the areas, as shown in the below histogram:



Clustering

In order to classify restaurant type per area I used a clustering analysis, in particular the k-means. It is an algorithm to group a set of object (restaurants) in k-groups based on some attributes. First of all I transformed the DataFrame with the restaurant venues, associated to municipalities, by one-hot encoding;

Neighborhood		American Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Fast Food Restaurant	German Restaurant	Greek Restaurant	Italian Restaurant	Japanese Restaurant	Kebab Restaurant	Lorr Resta
1	Lentate sul Seveso	0	0	0	0	0	0	0	0	1	0	0	
2	Lentate sul Seveso	0	0	0	0	0	0	0	0	1	0	0	
3	Lentate sul Seveso	0	0	0	0	0	0	0	0	0	1	0	
4	Lentate sul Seveso	0	0	0	0	0	0	0	0	1	0	0	
5	Lentate sul Seveso	0	0	0	0	0	0	0	0	1	0	0	

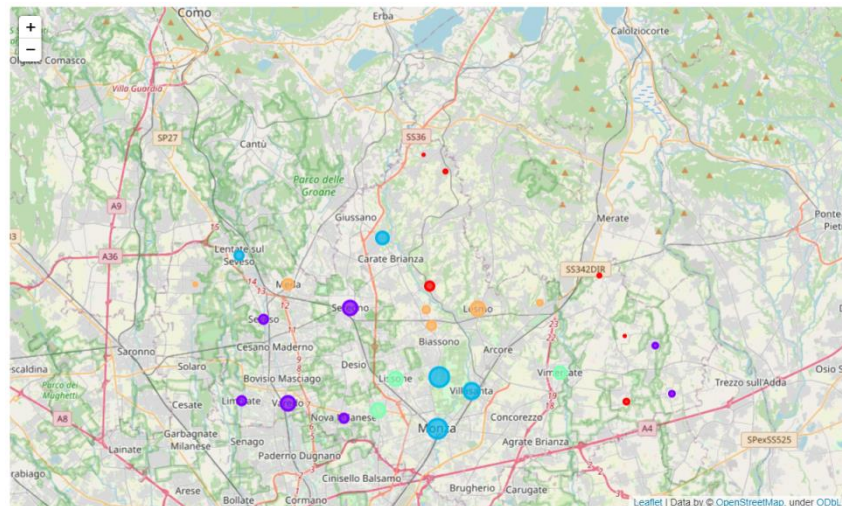
Next, I used grouping to show the frequency of each category of restaurants in each city district. I used this information to create a data frame in which you can see the most common restaurant venue types for each city district.

	Neighborhood	American Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Fast Food Restaurant	German Restaurant	Greek Restaurant	Italian Restaurant	Japanese Restaurant	Kebab Restaurant	Lo Rest
0	Lentate sul Seveso	0.111111	0.000000	0.000000	0.000000	0.000000	0.222222	0.000000	0.000000	0.555556	0.111111	0.000000	0.0
1	Lesmo	0.000000	0.000000	0.000000	0.133333	0.000000	0.000000	0.000000	0.000000	0.866667	0.000000	0.000000	0.0
2	Limbate	0.000000	0.000000	0.000000	0.222222	0.000000	0.111111	0.000000	0.000000	0.333333	0.333333	0.000000	0.0
3	Lissone	0.000000	0.000000	0.000000	0.000000	0.066667	0.000000	0.000000	0.000000	0.400000	0.200000	0.000000	0.0
4	Macherio	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.800000	0.000000	0.000000	0.0
5	Meda	0.000000	0.000000	0.000000	0.090909	0.000000	0.000000	0.000000	0.000000	0.727273	0.181818	0.000000	0.0
6	Mezzago	0.000000	0.000000	0.000000	0.000000	0.000000	0.250000	0.000000	0.000000	0.250000	0.500000	0.000000	0.0
7	Misinto	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.333333	0.666667	0.000000	0.000000	0.0
8	Monza	0.000000	0.000000	0.000000	0.047619	0.000000	0.047619	0.000000	0.000000	0.571429	0.190476	0.000000	0.0
9	Muggiò	0.000000	0.000000	0.000000	0.000000	0.066667	0.066667	0.000000	0.000000	0.266667	0.333333	0.000000	0.0
10	Nova Milanese	0.000000	0.000000	0.000000	0.000000	0.000000	0.333333	0.000000	0.000000	0.222222	0.333333	0.000000	0.0
11	Ornago	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	0.000000	0.000000	0.0
12	Renate	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	0.000000	0.000000	0.0
13	Roncello	0.000000	0.000000	0.000000	0.000000	0.000000	0.250000	0.000000	0.000000	0.250000	0.500000	0.000000	0.0
14	Ronco Briantino	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	0.000000	0.000000	0.0
15	Seregno	0.000000	0.066667	0.000000	0.000000	0.000000	0.133333	0.066667	0.000000	0.333333	0.266667	0.066667	0.0
16	Seveso	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.666667	0.333333	0.000000	0.0
17	Sovico	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.857143	0.000000	0.000000	0.0
18	Subiate	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	0.000000	0.000000	0.0
19	Triuggio	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	0.000000	0.000000	0.0
20	Usmate Velate	0.000000	0.000000	0.166667	0.000000	0.000000	0.000000	0.000000	0.000000	0.833333	0.000000	0.000000	0.0
21	Varedo	0.000000	0.000000	0.000000	0.066667	0.000000	0.200000	0.000000	0.000000	0.400000	0.133333	0.000000	0.0
22	Veduggio al Lambro	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.545455	0.136364	0.000000	0.0
23	Veduggio con Colzano	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	0.000000	0.000000	0.0

Finally I applied from scikit-learn package the k-means algorithm (unsupervised machine learning algorithm). I did not applied any optimization on the k values, I used the value k=5 as default, inspired also by the Coursera course.

Result

For each municipality (the column Neighborhood) the algorithm assigned a cluster with the most common venue (restaurant type). The cluster interval span from 0 to 4 (five cluster in total). The clusters can be shown in the following colored geographic map (color = cluster code), again, with Folium.



There are about 20 bubbles with 5 different colors. Now let's show the content of each cluster trying to identify the discriminant venue restaurant type.

Cluster 1

	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
11	Ornago	Italian Restaurant	Vegetarian / Vegan Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant	Fast Food Restaurant	German Restaurant	Greek Restaurant
12	Renate	Italian Restaurant	Vegetarian / Vegan Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant	Fast Food Restaurant	German Restaurant	Greek Restaurant
14	Ronco Briantino	Italian Restaurant	Vegetarian / Vegan Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant	Fast Food Restaurant	German Restaurant	Greek Restaurant
18	Sulbiate	Italian Restaurant	Vegetarian / Vegan Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant	Fast Food Restaurant	German Restaurant	Greek Restaurant
19	Triuggio	Italian Restaurant	Vegetarian / Vegan Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant	Fast Food Restaurant	German Restaurant	Greek Restaurant
23	Veduggio con Colzano	Italian Restaurant	Vegetarian / Vegan Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant	Fast Food Restaurant	German Restaurant	Greek Restaurant

This cluster the most common restaurant types are: *Italian, Vegetarian and Argentinian*

Cluster 2

	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	Limbiate	Italian Restaurant	Japanese Restaurant	Chinese Restaurant	Fast Food Restaurant	Argentinian Restaurant	Australian Restaurant	Comfort Food Restaurant	Falafel Restaurant	German Restaurant	Greek Restaurant
6	Mezzago	Italian Restaurant	Japanese Restaurant	Fast Food Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant	German Restaurant	Greek Restaurant
10	Nova Milanese	Japanese Restaurant	Italian Restaurant	Fast Food Restaurant	New American Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant	German Restaurant
13	Roncello	Italian Restaurant	Japanese Restaurant	Fast Food Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant	German Restaurant	Greek Restaurant
15	Seregno	Italian Restaurant	Japanese Restaurant	Fast Food Restaurant	Kebab Restaurant	Argentinian Restaurant	German Restaurant	Vegetarian / Vegan Restaurant	Mediterranean Restaurant	Lombard Restaurant	Mexican Restaurant
16	Seveso	Italian Restaurant	Japanese Restaurant	Falafel Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Fast Food Restaurant	German Restaurant	Greek Restaurant
21	Varedo	Italian Restaurant	Japanese Restaurant	Fast Food Restaurant	New American Restaurant	Chinese Restaurant	Mediterranean Restaurant	Seafood Restaurant	German Restaurant	Argentinian Restaurant	Australian Restaurant

This cluster the most common restaurant types are: *Italian, Japanese and Fast Food*

Cluster 3

	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	Lentate sul Seveso	Italian Restaurant	Fast Food Restaurant	Japanese Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant	German Restaurant	Greek Restaurant
8	Monza	Italian Restaurant	Japanese Restaurant	Seafood Restaurant	Chinese Restaurant	Fast Food Restaurant	Vegetarian / Vegan Restaurant	Roman Restaurant	Mexican Restaurant	Mediterranean Restaurant	Lombard Restaurant
22	Vedano al Lambro	Italian Restaurant	Japanese Restaurant	Mexican Restaurant	Seafood Restaurant	Mediterranean Restaurant	Vegetarian / Vegan Restaurant	Modern European Restaurant	New American Restaurant	Lombard Restaurant	Kebab Restaurant
24	Verano Brianza	Italian Restaurant	Fast Food Restaurant	Japanese Restaurant	Mexican Restaurant	Seafood Restaurant	Greek Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant
25	Villasanta	Italian Restaurant	Chinese Restaurant	Fast Food Restaurant	Vegetarian / Vegan Restaurant	Mexican Restaurant	Argentinian Restaurant	Japanese Restaurant	New American Restaurant	Mediterranean Restaurant	Lombard Restaurant

This cluster the most common restaurant types are: *Italian, Fast Food and Japanese*

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
3	Lissone	Italian Restaurant	Japanese Restaurant	Mexican Restaurant	New American Restaurant	Mediterranean Restaurant	Lombard Restaurant	Comfort Food Restaurant	Falafel Restaurant	German Restaurant	Argentinian Restaurant
9	Muggiò	Japanese Restaurant	Italian Restaurant	Mexican Restaurant	Roman Restaurant	Mediterranean Restaurant	Comfort Food Restaurant	Fast Food Restaurant	German Restaurant	Argentinian Restaurant	Australian Restaurant
26	Vimercate	Italian Restaurant	Japanese Restaurant	Mediterranean Restaurant	Australian Restaurant	Chinese Restaurant	German Restaurant	Kebab Restaurant	American Restaurant	Lombard Restaurant	Mexican Restaurant

This cluster the most common restaurant types are: *Italian, Japanese and Mexican*

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
1	Lesmo	Italian Restaurant	Chinese Restaurant	Vegetarian / Vegan Restaurant	Argentinian Restaurant	Australian Restaurant	Comfort Food Restaurant	Falafel Restaurant	Fast Food Restaurant	German Restaurant	Greek Restaurant
4	Macherio	Italian Restaurant	Paella Restaurant	New American Restaurant	Vegetarian / Vegan Restaurant	Greek Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant
5	Meda	Italian Restaurant	Japanese Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant	Fast Food Restaurant	German Restaurant	Greek Restaurant
7	Misinto	Italian Restaurant	Greek Restaurant	Vegetarian / Vegan Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant	Fast Food Restaurant	German Restaurant
17	Sovico	Italian Restaurant	Paella Restaurant	Vegetarian / Vegan Restaurant	Argentinian Restaurant	Australian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant	Fast Food Restaurant	German Restaurant
20	Usmate Velate	Italian Restaurant	Australian Restaurant	Vegetarian / Vegan Restaurant	Argentinian Restaurant	Chinese Restaurant	Comfort Food Restaurant	Falafel Restaurant	Fast Food Restaurant	German Restaurant	Greek Restaurant

This cluster the most common restaurant types are: *Italian and Mediterranean (Spanish, Greek)*

Discussion

In all cluster the generic Italian restaurant is at first place, but the DataFrame identifies that there are many other choices and municipalities where other restaurants types goes for more, in particular Japanese restaurants. For sure in these years, the openings of these last types of restaurants have significant increase in MB province and the analysis confirm that. The bubbles show also that the distribution of the restaurant types split the province in four macro areas: the south (Monza city), the East-West as a unique zone (same cluster), the Center (Biassono, Lesmo municipality) and North-Est (Renate, Triuggio, Ornago ...)

Conclusion

In MB province there is enough choice of restaurant types to satisfy the palate tourists or business man: the choice depends on where you are going to stay or which city are within your reach. Here are some ideas for further improvement:

- Include in the analysis also the ratings of the place.
- Split the generic Italian restaurant into regional Italian ones. For simplicity we can identify three macro areas; North, center and South.
- Try to identify, if any, a correlation between restaurant type and local communities.