Analyzing neighborhoods for an opening restaurant in Milan

Tommaso Duratorre

October 2, 2019

Contents

1 Introduction: Business Problem

In this project we will try to find the best location to open a restaurant in the city of Milan.

In Milan there a lot of restaurants, each specialized in a particular cuisine, like Italian, Japanese, Chinese, Indian, Brazilian, etc. Our objective is to find locations that are not well covered with restaurants, and to discover the type of cuisine, which is still missing or is less present in those locations, in order to have the least competition when opening the new restaurant.

We will analyze the top venues around the locations and determine what type of cuisine our restaurant should serve in each neighborhood, and decide the best one based on the number of potential customers, i.e. how much popular is the type of cuisine around the city and how close to the city center each location is. We will assume the popularity of a certain type on cuisine as proportional to the number of corresponding locations that serve it.

This project may be of interest for potential restaurateurs that want to open a new business in Milan.

2 Data

We will take into consideration the following aspects:

- number of already existing restaurants around each location;
- most common restaurants and their type of cuisine in each neighborhood;
- distance of each neighborhood from the city center.

We will divide the city into neighborhoods, based on a geojson file provided by the municipality's website of the city of Milan, with the coordinates of the neighborhoods already given. Using the coordinate values, we will be able to determine the center of each neighborhood and its distance from the city center.

We will use Foursquare API to retrieve the number of restaurants along with their type in each neighborhood, and to get the most popular restaurants in every location, in order to choose the candidate cuisine for the new restaurant.

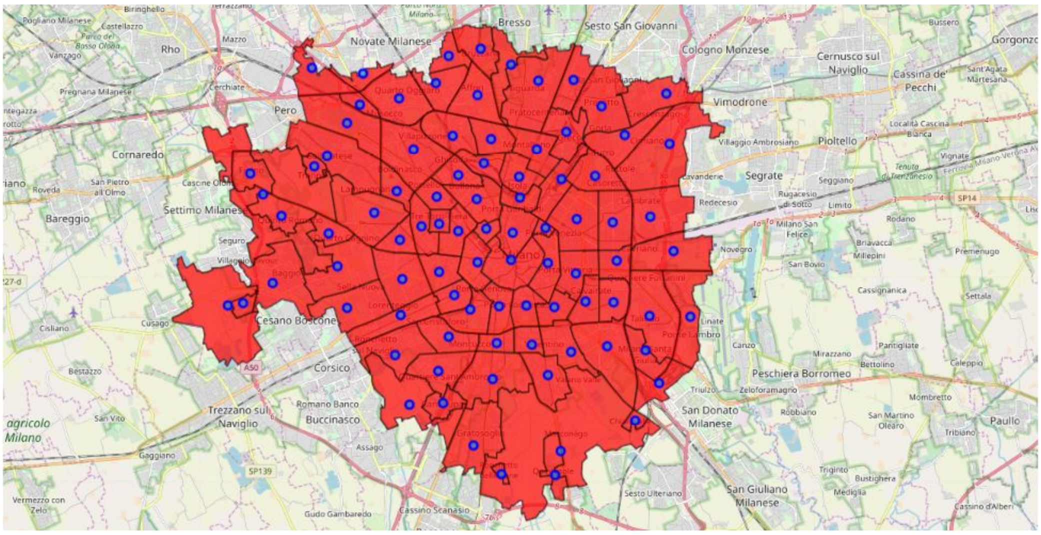

Figure 1: Map of Milan with its neighborhoods

3 Methodology

3.1 Downloading and exploring the dataset

First step consists in creating the database containing the neighborhoods of Milan along with their coordinates. Since the geojson file given by the municipality of Milan contains the coordinates of the borders of each neighborhood, it is necessary to obtain the coordinates of the centers. These values are obtained by means of a spatial mean between the border coordinates. Displaying the results, the map of Milan with its neighborhoods looks like Figure 1.

Some of the center coordinates do not respect the neighborhood division. Therefore, a new approach is recommended. The new values of the city center will be calculated using the geopy library, inserting the name of each neighborhood in the request. With this method we have the Figure 2.

The new values look more reasonable, even though there are still some center coordinates that do not allow a proper coverage of the area. Since the location of the centers of these neighborhoods, whose centers do not represent well the corresponding neighborhoods, was better in the first map, we will proceed by replacing the new values who do not match the neighborhood division with the old values obtained by means of the spatial mean. Eventually, the map looks like Figure 3.

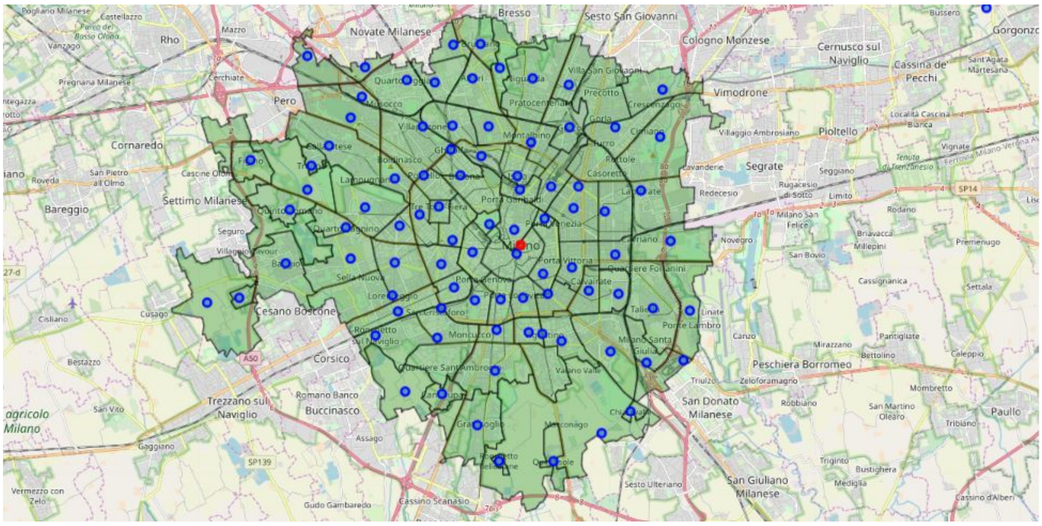

Figure 2: Map of Milan with its neighborhoods, obtained with geopy

The next step is about the calcuation of the distance from the city center of each neighborhood. We are going to assume that the city center is located at the coordinates of Milan given by geopy. The distance is simply calculated starting from the values of latitude and longitude of the neighborhoods. Once the distances are calculated, we will insert the values in a dataframe. The resulting dataframe looks like Figure 4.

3.2 Using Foursquare

Using Foursquare, we will explore the top 100 venues in each neighborhood, setting the radius of search equal to 1km around each neighborhood center. We are also going to put the results given by Foursquare in a pandas dataframe. In order to search for restaurants around neighborhoods, we will include in our API calls the category command, and look for the specific category “4d4b7105d754a06374d81259”, which is the category related to food in Foursquare.

Since this category includes more places other than restaurants, like bakeries, bagel shops, and more, we will make sure that our research only consider restaurant categories, by removing everything whose category name does not contain the word “Restaurant”. Eventually, the new dataframe is the one showed in Figure 5.

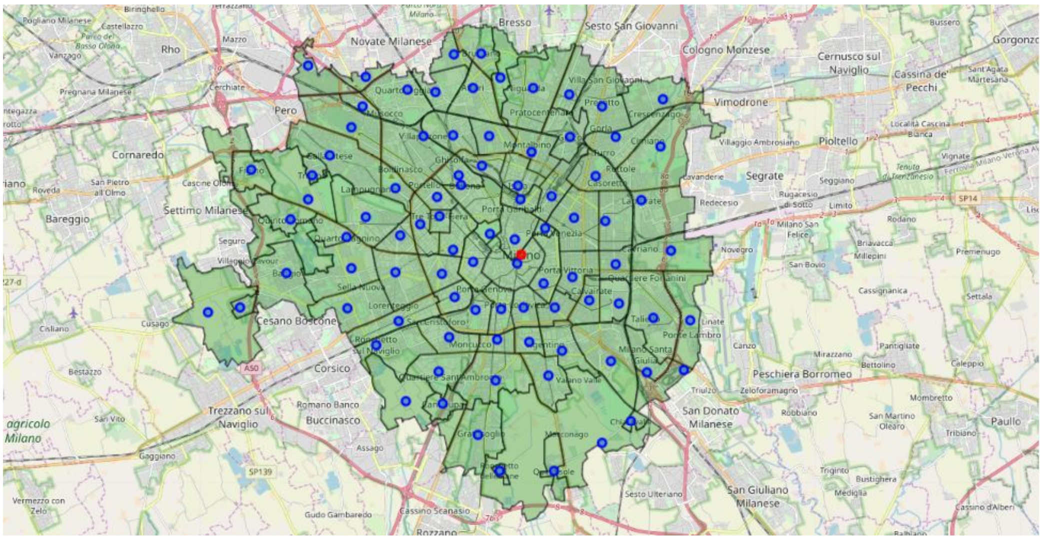

Figure 3: Map of Milan

	Neighborhood	Latitude	Longitude	Dist_from_center
0	SACCO	45.460146	9.123899	5.260875
1	COMASINA	45.460146	9.161565	2.380538
2	STEPHENSON	45.460146	9.122565	5.364142
3	QT 8	45.460146	9.136660	4.275062
4	ORTOMERCATO	45.460146	9.232368	3.356946
5	MAGGIORE - MUSOCCO	45.460146	9.117462	5.759688
6	PARCO LAMBRO - CIMIANO	45.460146	9.250219	4.728826
7	GALLARATESE	45.460146	9.108251	6.474792
8	S. SIRO	45.460146	9.123964	5.255810
9	GHISOLFA	45.460146	9.160657	2.448132
10	BAGGIO	45.460146	9.089843	7.906750
11	QUARTO CAGNINO	45.460146	9.115668	5.898915
12	LORENTEGGIO	45.460146	9.135620	4.355129
13	GIAMBELLINO	45.460146	9.137871	4.181792
14	S. CRISTOFORO	45.460146	9.156440	2.764360
15	RONCHETTO SUL NAVIGLIO	45.460146	9.128216	4.926809
16	TIBALDI	45.460146	9.180175	1.094696
17	CASCINA TRIULZA - EXPO	45.460146	9.099068	7.188747
18	QUARTO OGGIARO	45.460146	9.141661	3.890415
19	AFFORI	45.460146	9.169653	1.790041

Figure 4: Dataframe with neighborhoods and their distances from the center

	Neighborhood	Venue	Venue Category
0	SACCO	Shi So Restaurant Sushi	Japanese Restaurant
1	COMASINA	McDonald's	Fast Food Restaurant
2	COMASINA	McDonald's	Fast Food Restaurant
3	STEPHENSON	Rossopomodoro	Italian Restaurant
4	QT 8	Ristorante Ribot	Italian Restaurant
5	QT 8	Unico Restaurant	Italian Restaurant
6	QT 8	Ristorante Pizzeria Monte Stella	Italian Restaurant
7	QT 8	McDonald's	Fast Food Restaurant
8	QT 8	L'Arca	Italian Restaurant
9	ORTOMERCATO	Oste Italiano	Italian Restaurant
10	ORTOMERCATO	Piccolo Sogno	Italian Restaurant
11	ORTOMERCATO	Trattoria del Nuovo Macello	Italian Restaurant
12	ORTOMERCATO	Syderfood	Italian Restaurant
13	MAGGIORE - MUSOCCO	Osteria da Salvo	Seafood Restaurant
14	PARCO LAMBRO - CIMIANO	Angolo nascosto	Italian Restaurant
15	PARCO LAMBRO - CIMIANO	El magna gatt	Italian Restaurant
16	GALLARATESE	Yun Quick Fusion Cusine	Asian Restaurant
17	GALLARATESE	McCafé	Fast Food Restaurant
18	GALLARATESE	McDonald's	Fast Food Restaurant
19	GALLARATESE	Mishi-Mishi	Sushi Restaurant

Figure 5: Neighborhoods and their venues

After that, it is necessary to get the number of venues for each neighborhood, since one indicator for our decision will be the number of existing locations around the neighborhoods. This can be done by grouping the previous dataframe on “Neighborhoods” and using the ”count” command.

3.3 Determing the recommendation in each neighborhood

Then, we use the one hot encoding technique to calculate both the popularity of each restaurant type, and also their frequency of occurrence in each neighborhood.

We will assume the popularity of each kind of restaurant to be equal to the total number of the different types of restaurant around the city. Having said that, we can calculate the total number of each type of restaurant in the city, starting from the one hot encoding dataframe. Sorting the resulting dataframe in descending order for the number of resturants, we obtain the dataframe in Figure 6.

To determine the frequency of occurrence of each type of restaurant in the neighborhoods, starting from the one hot encoding, it is necessary to group

	Category	Popularity
0	Italian Restaurant	738
1	Japanese Restaurant	138
2	Seafood Restaurant	111
3	Sushi Restaurant	100
4	Chinese Restaurant	81
5	Asian Restaurant	42
6	Kebab Restaurant	37
7	Indian Restaurant	33
8	Vegetarian / Vegan Restaurant	30
9	Mediterranean Restaurant	25
10	Fast Food Restaurant	25
11	American Restaurant	13
12	Argentinian Restaurant	13
13	Ramen Restaurant	13
14	Mexican Restaurant	12

Figure 6: Dataframe with type of restaurants and their popularity

	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	ADRIANO	Italian Restaurant	Vietnamese Restaurant	Latin American Restaurant	Korean Restaurant	Kebab Restaurant	Japanese Restaurant	Indian Restaurant	Himalayan Restaurant	Hawaiian Restaurant	Greek Restaurant
1	AFFORI	Italian Restaurant	Japanese Restaurant	Vietnamese Restaurant	Latin American Restaurant	Korean Restaurant	Kebab Restaurant	Indian Restaurant	Himalayan Restaurant	Hawaiian Restaurant	Greek Restaurant
2	BAGGIO	Italian Restaurant	Japanese Restaurant	Vietnamese Restaurant	Latin American Restaurant	Korean Restaurant	Kebab Restaurant	Indian Restaurant	Himalayan Restaurant	Hawaiian Restaurant	Greek Restaurant
3	BANDE NERE	Italian Restaurant	Sushi Restaurant	Japanese Restaurant	Seafood Restaurant	Falafel Restaurant	Fast Food Restaurant	German Restaurant	Filipino Restaurant	French Restaurant	Vietnamese Restaurant
4	BARONA	Japanese Restaurant	Vietnamese Restaurant	Latin American Restaurant	Korean Restaurant	Kebab Restaurant	Italian Restaurant	Indian Restaurant	Himalayan Restaurant	Hawaiian Restaurant	Greek Restaurant
5	BICOCCA	Italian Restaurant	Sushi Restaurant	Kebab Restaurant	Seafood Restaurant	American Restaurant	Sardinian Restaurant	Vietnamese Restaurant	French Restaurant	Falafel Restaurant	Fast Food Restaurant
6	BOVISA	Italian Restaurant	Vegetarian / Vegan Restaurant	Fast Food Restaurant	Kebab Restaurant	Doner Restaurant	Sicilian Restaurant	German Restaurant	Filipino Restaurant	French Restaurant	Hawaiian Restaurant
7	BOVISASCA	Italian Restaurant	Chinese Restaurant	Fast Food Restaurant	Vietnamese Restaurant	Ethiopian Restaurant	Korean Restaurant	Kebab Restaurant	Japanese Restaurant	Indian Restaurant	Himalayan Restaurant
8	BRERA	Italian Restaurant	Japanese Restaurant	Seafood Restaurant	Asian Restaurant	Mediterranean Restaurant	Sushi Restaurant	Modern European Restaurant	Shabu-Shabu Restaurant	French Restaurant	Puglia Restaurant
9	BRUZZANO	Italian Restaurant	Fast Food Restaurant	Vietnamese Restaurant	Latin American Restaurant	Korean Restaurant	Kebab Restaurant	Japanese Restaurant	Indian Restaurant	Himalayan Restaurant	Hawaiian Restaurant

Figure 7: Neighbrohoods along with the ten top existing locations

the dataframe on "Neighborhood" and then perform a mean. In this way, the frequency of each type of restaurant on the total number of restaurants in every neighborhood is calculated. After that, we will create a dataframe with the ten top most common venues in each neighborhood, taking the first ten type of restaurants with the highest frequency, sorted in descending order. The resulting dataframe looks like Figure 7.

Once we got the first top ten venues, we can determine the candidate type of restaurant in each neighborhood, by looking at the dataframe containing the types of restaurant sorted by popularity and choosing the first element, which is not present in the top ten venues for each neighborhood. The addition of a new column "Recommended" in the dataframe, containing the recommendation results, leads to the dataframe in Figure 8.

3.4 Selecting the candidate neighborhood along with the type of restaurant

To select the neighborhood where to open the new restaurant and to choose the type of cuisine, we will focus on two parameters, the distance from the center and the number of existing restaurants. In particular, the smaller the distance and the fewer existing restaurants, the better it will be.

As a first step, it is necessary to merge all the previous dataframes in order to have all the necessary information displayed. The resulting dataframe is the on in Figure 9.

	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	Recommendation
0	ADRIANO	Italian Restaurant	Vietnamese Restaurant	Latin American Restaurant	Korean Restaurant	Kebab Restaurant	Japanese Restaurant	Indian Restaurant	Himalayan Restaurant	Hawaiian Restaurant	Greek Restaurant	Seafood Restaurant
1	AFFORI	Italian Restaurant	Japanese Restaurant	Vietnamese Restaurant	Latin American Restaurant	Korean Restaurant	Kebab Restaurant	Indian Restaurant	Himalayan Restaurant	Hawaiian Restaurant	Greek Restaurant	Seafood Restaurant
2	BAGGIO	Italian Restaurant	Japanese Restaurant	Vietnamese Restaurant	Latin American Restaurant	Korean Restaurant	Kebab Restaurant	Indian Restaurant	Himalayan Restaurant	Hawaiian Restaurant	Greek Restaurant	Seafood Restaurant
3	BANDE NERE	Italian Restaurant	Sushi Restaurant	Japanese Restaurant	Seafood Restaurant	Falafel Restaurant	Fast Food Restaurant	German Restaurant	Filipino Restaurant	French Restaurant	Vietnamese Restaurant	Chinese Restaurant
4	BARONA	Japanese Restaurant	Vietnamese Restaurant	Latin American Restaurant	Korean Restaurant	Kebab Restaurant	Italian Restaurant	Indian Restaurant	Himalayan Restaurant	Hawaiian Restaurant	Greek Restaurant	Seafood Restaurant
5	BICOCCA	Italian Restaurant	Sushi Restaurant	Kebab Restaurant	Seafood Restaurant	American Restaurant	Sardinian Restaurant	Vietnamese Restaurant	French Restaurant	Falafel Restaurant	Fast Food Restaurant	Japanese Restaurant
6	BOVISA	Italian Restaurant	Vegetarian / Vegan Restaurant	Fast Food Restaurant	Kebab Restaurant	Doner Restaurant	Sicilian Restaurant	German Restaurant	Filipino Restaurant	French Restaurant	Hawaiian Restaurant	Japanese Restaurant
7	BOVISASCA	Italian Restaurant	Chinese Restaurant	Fast Food Restaurant	Vietnamese Restaurant	Ethiopian Restaurant	Korean Restaurant	Kebab Restaurant	Japanese Restaurant	Indian Restaurant	Himalayan Restaurant	Seafood Restaurant
8	BRERA	Italian Restaurant	Japanese Restaurant	Seafood Restaurant	Asian Restaurant	Mediterranean Restaurant	Sushi Restaurant	Modern European Restaurant	Shabu-Shabu Restaurant	French Restaurant	Puglia Restaurant	Chinese Restaurant
9	BRUZZANO	Italian Restaurant	Fast Food Restaurant	Vietnamese Restaurant	Latin American Restaurant	Korean Restaurant	Kebab Restaurant	Japanese Restaurant	Indian Restaurant	Himalayan Restaurant	Hawaiian Restaurant	Seafood Restaurant

Figure 8: Neighborhoods and their restaurant recommendation

From the dataframe above it is not easy to make a decision about the best location for the new restaurant, i.e. it is difficult to set threshold values for the distance from the center and for the number of existing venues in order to choose the best neighborhood. For this purpose, it may be helpful to use clustering in order to simplify the choice.

We will cluster the neighborhoods into 4 clusters, using k-means clustering. Once clustering is performed, we add the corresponding label to the previous dataframe, obtaining the dataframe in Figure 10.

So clustering led to 4 clusters named Cluster 0, Cluster 1, Cluster 2, and Cluster 3. Then, it is necessary to analyze all the clusters, in order to define their unique features. Starting from Cluster 0, we have something like Figure 11.

We can see that Cluster 0 contains those neighborhoods that are far from the city center and not well covered with restaurants. If the second factor is good when opening the new restaurant, the fact that these neighborhoods are far from the city center, with the closest being 5 km far, is a disadvantage. Furthermore, the recommendations are pretty much the same, since there are not many restaurants in the area and it is difficult to produce a recommendation due to the lack of data.

	Neighborhood	Recommendation	Latitude	Longitude	Dist_from_center	Number_venues
0	ADRIANO	Seafood Restaurant	45.513572	9.251202	7.038687	3
1	AFFORI	Seafood Restaurant	45.517029	9.169653	5.815828	4
2	BAGGIO	Seafood Restaurant	45.461384	9.089843	7.894986	3
3	BANDE NERE	Chinese Restaurant	45.461504	9.136484	4.265065	13
4	BARONA	Seafood Restaurant	45.431686	9.155134	4.783389	1
5	BICOCCA	Japanese Restaurant	45.514917	9.211138	5.586139	15
6	BOVISA	Japanese Restaurant	45.502770	9.161264	4.605204	13
7	BOVISASCA	Seafood Restaurant	45.515842	9.153778	6.160506	4
8	BRERA	Chinese Restaurant	45.471519	9.187735	0.567514	56
9	BRUZZANO	Seafood Restaurant	45.527369	9.173292	6.865102	4
10	BUENOS AIRES - VENEZIA	Asian Restaurant	45.477892	9.212902	2.142291	61
11	CANTALUPA	Seafood Restaurant	45.421965	9.156848	5.635435	1
12	CASCINA TRIULZA - EXPO	Seafood Restaurant	45.523592	9.099068	9.535147	1
13	CENTRALE	Asian Restaurant	45.484352	9.203372	2.195455	54
14	CHIARAVALLE	Seafood Restaurant	45.416697	9.237421	6.669467	1
15	CITTA' STUDI	Asian Restaurant	45.477056	9.226575	3.042737	29
16	COMASINA	Seafood Restaurant	45.526930	9.161565	7.055480	2
17	CORSICA	Asian Restaurant	45.463909	9.230802	3.168273	13
18	DE ANGELI - MONTE ROSA	Kebab Restaurant	45.476130	9.147302	3.533522	24
19	DERGANO	Sushi Restaurant	45.502513	9.176784	4.111718	19

Figure 9: Neighborhoods and all the relevant information

	Neighborhood	Recommendation	Latitude	Longitude	Dist_from_center	Number_venues	Cluster
0	ADRIANO	Seafood Restaurant	45.460146	9.251202	4.804740	3	0
1	AFFORI	Seafood Restaurant	45.460146	9.169653	1.790041	4	0
2	BAGGIO	Seafood Restaurant	45.460146	9.089843	7.906750	3	0
3	BANDE NERE	Chinese Restaurant	45.460146	9.136484	4.288545	13	3
4	BARONA	Seafood Restaurant	45.460146	9.154670	2.897954	1	0
5	BICOCCA	Japanese Restaurant	45.460146	9.211138	1.775423	15	3
6	BOVISA	Japanese Restaurant	45.460146	9.161264	2.402933	13	3
7	BOVISASCA	Seafood Restaurant	45.460146	9.153778	2.965476	4	0
8	BRERA	Chinese Restaurant	45.460146	9.187735	0.770184	56	2
9	BRUZZANO	Seafood Restaurant	45.460146	9.173292	1.535396	4	0
10	BUENOS AIRES - VENEZIA	Asian Restaurant	45.460146	9.212902	1.901684	61	2
11	CANTALUPA	Seafood Restaurant	45.460146	9.156848	2.733604	1	0
12	CASCINA TRIULZA - EXPO	Seafood Restaurant	45.460146	9.099068	7.188747	1	0
13	CENTRALE	Asian Restaurant	45.460146	9.203372	1.249052	54	2
14	CHIARAVALLE	Seafood Restaurant	45.460146	9.237421	3.743452	1	0

Figure 10: Neighborhoods and their clusters

Looking at Cluster 1 instead, we have Figure 12. Cluster 1 contains neighborhoods that are moderately near to the city center, with a number of exisiting venues greater than Cluster 0.

Cluster 2 is showed in Figure 13. From what we can see, Cluster 2 includes those neighborhoods that are close to the city center, but with a high number of existing locations, as one can expect. Lastly, regarding Cluster 3, it is showed in Figure 14

In Cluster 3 there are neighborhoods with a number of existing restaurants which is higher than Cluster 0 but smaller than Cluster 1. These neighborhoods are moderately far from the city center.

4 Results

From the clusters, we can see that the ideal neighborhoods are those in Cluster 1. These neighborhoods are not as crowded with restaurants as Cluster 2, and they are generally located farther from the city center with respect to Cluster 1, but still within an acceptable value.

	Neighborhood	Recommendation	Latitude	Longitude	Dist_from_center	Number_venues	Cluster
0	ADRIANO	Seafood Restaurant	45.513572	9.251202	7.038687	3	0
1	AFFORI	Seafood Restaurant	45.517029	9.169653	5.815828	4	0
2	BAGGIO	Seafood Restaurant	45.461384	9.089843	7.894986	3	0
4	BARONA	Seafood Restaurant	45.431686	9.155134	4.783389	1	0
7	BOVISASCA	Seafood Restaurant	45.515842	9.153778	6.160506	4	0
9	BRUZZANO	Seafood Restaurant	45.527369	9.173292	6.865102	4	0
11	CANTALUPA	Seafood Restaurant	45.421965	9.156848	5.635435	1	0
12	CASCINA TRIULZA - EXPO	Seafood Restaurant	45.523592	9.099068	9.535147	1	0
14	CHIARAVALLE	Seafood Restaurant	45.416697	9.237421	6.669467	1	0
16	COMASINA	Seafood Restaurant	45.526930	9.161565	7.055480	2	0
23	FIGINO	Seafood Restaurant	45.492287	9.074819	9.477839	1	0
24	GALLARATESE	Italian Restaurant	45.496641	9.108251	7.235443	4	0
29	GRATOSOGLIO - TICINELLO	Seafood Restaurant	45.412666	9.171992	6.187957	1	0
34	LODI - CORVETTO	Seafood Restaurant	45.434814	9.228742	4.646016	6	0
35	LORENTEGGIO	Seafood Restaurant	45.450799	9.115979	6.093701	7	0
39	MAGGIORE - MUSOCCO	Italian Restaurant	45.505177	9.117462	7.127205	1	0
40	MECENATE	Seafood Restaurant	45.447779	9.247091	4.905244	3	0
41	MUGGIANO	Seafood Restaurant	45.451031	9.069758	9.604874	1	0
43	NIGUARDA - CA' GRANDA	Japanese Restaurant	45.516961	9.195574	5.589422	3	0
44	ORTOMERCATO	Seafood Restaurant	45.452223	9.232368	3.653469	4	0
47	PARCO BOSCO IN CITTÀ	Japanese Restaurant	45.483447	9.099050	7.386161	3	0
48	PARCO DEI NAVIGLI	Seafood Restaurant	45.422774	9.140946	6.242434	1	0
49	PARCO FORLANINI - ORTICA	Seafood Restaurant	45.467985	9.254705	5.022883	5	0
50	PARCO LAMBRO - CIMIANO	Seafood Restaurant	45.499372	9.250219	5.908261	2	0
51	PARCO MONLUE' - PONTE LAMBRO	Seafood Restaurant	45.447044	9.262690	6.058418	3	0
52	PARCO NORD	Seafood Restaurant	45.519945	9.181313	5.950382	1	0
56	QT 8	Seafood Restaurant	45.486940	9.136660	4.767919	5	0
57	QUARTO CAGNINO	Sushi Restaurant	45.472079	9.115668	5.881169	6	0
58	QUARTO OGGIARO	Seafood Restaurant	45.516567	9.141661	6.720965	3	0
59	QUINTO ROMANO	Seafood Restaurant	45.477421	9.091601	7.823020	1	0
60	QUINTOSOLE	Seafood Restaurant	45.401867	9.204357	7.297329	3	0
61	RIPAMONTI	Japanese Restaurant	45.430404	9.202077	4.144944	8	0
62	ROGOREDO	Italian Restaurant	45.431573	9.244480	5.758400	3	0
63	RONCHETTO DELLE RANE	Seafood Restaurant	45.401890	9.181023	7.251771	3	0
64	RONCHETTO SUL NAVIGLIO	Japanese Restaurant	45.439712	9.128216	5.726894	9	0
66	S. SIRO	Italian Restaurant	45.478200	9.123964	5.354870	5	0
67	SACCO	Seafood Restaurant	45.520365	9.123899	7.908696	1	0
71	STADERA	Seafood Restaurant	45.429045	9.179334	4.285763	6	0
72	STEPHENSON	Seafood Restaurant	45.511392	9.122565	7.264144	1	0
77	TRIULZO SUPERIORE	Italian Restaurant	45.432155	9.260206	6.675305	3	0
81	VILLAPIZZONE	Seafood Restaurant	45.502553	9.148469	5.156383	7	0

Figure 11: Cluster 0

	Neighborhood	Recommendation	Latitude	Longitude	Dist_from_center	Number_venues	Cluster
15	CITTA' STUDI	Asian Restaurant	45.477056	9.226575	3.042737	29	1
26	GHISOLFA	Indian Restaurant	45.490783	9.163536	3.398630	41	1
36	LORETO	Indian Restaurant	45.490510	9.222467	3.632286	44	1
38	MAGENTA - S. VITTORE	Seafood Restaurant	45.464689	9.169665	1.646077	32	1
46	PAGANO	Seafood Restaurant	45.468285	9.161100	2.304995	41	1
55	PORTELLO	Asian Restaurant	45.484139	9.154279	3.425705	43	1
73	TIBALDI	Seafood Restaurant	45.441302	9.180175	2.946347	36	1
76	TRE TORRI	Chinese Restaurant	45.478374	9.155361	3.033995	27	1
78	UMBRIA - MOLISE	Seafood Restaurant	45.453199	9.219422	2.720618	33	1
80	VIGENTINA	Vegetarian / Vegan Restaurant	45.451087	9.191564	1.748021	35	1
82	WASHINGTON	Kebab Restaurant	45.461206	9.156310	2.745026	35	1

Figure 12: Cluster 1

	Neighborhood	Recommendation	Latitude	Longitude	Dist_from_center	Number_venues	Cluster
8	BRERA	Chinese Restaurant	45.471519	9.187735	0.567514	56	2
10	BUENOS AIRES - VENEZIA	Asian Restaurant	45.477892	9.212902	2.142291	61	2
13	CENTRALE	Asian Restaurant	45.484352	9.203372	2.195455	54	2
20	DUOMO	Seafood Restaurant	45.464138	9.188555	0.332283	54	2
25	GARIBALDI REPUBBLICA	Sushi Restaurant	45.483527	9.189933	1.859916	63	2
28	GIARDINI PORTA VENEZIA	Asian Restaurant	45.474727	9.200750	1.191326	53	2
31	GUASTALLA	Chinese Restaurant	45.458252	9.200023	1.206977	48	2
32	ISOLA	Asian Restaurant	45.487565	9.188972	2.311273	52	2
42	NAVIGLI	Asian Restaurant	45.450176	9.170897	2.400625	61	2
53	PARCO SEMPIONE	Seafood Restaurant	45.473033	9.176970	1.264677	50	2
54	PORTA ROMANA	Chinese Restaurant	45.451098	9.204923	2.077810	53	2
68	SARPI	Asian Restaurant	45.487855	9.164604	3.094640	48	2
74	TICINESE	Chinese Restaurant	45.450596	9.181951	1.920714	53	2
75	TORTONA	Fast Food Restaurant	45.454124	9.162046	2.633599	55	2
83	XXII MARZO	Kebab Restaurant	45.460146	9.212421	1.867083	49	2

Figure 13: Cluster 2

	Neighborhood	Recommendation	Latitude	Longitude	Dist_from_center	Number_venues	Cluster
3	BANDE NERE	Chinese Restaurant	45.461504	9.136484	4.265065	13	3
5	BICOCCA	Japanese Restaurant	45.514917	9.211138	5.586139	15	3
6	BOVISA	Japanese Restaurant	45.502770	9.161264	4.605204	13	3
17	CORSICA	Asian Restaurant	45.463909	9.230802	3.168273	13	3
18	DE ANGELI - MONTE ROSA	Kebab Restaurant	45.476130	9.147302	3.533522	24	3
19	DERGANO	Sushi Restaurant	45.502513	9.176784	4.111718	19	3
21	EX OM - MORIVIONE	Japanese Restaurant	45.440539	9.193754	2.929383	18	3
22	FARINI	Sushi Restaurant	45.493650	9.173480	3.267605	20	3
27	GIAMBELLINO	Seafood Restaurant	45.446969	9.137871	4.669172	10	3
30	GRECO	Japanese Restaurant	45.502184	9.211233	4.253921	10	3
33	LAMBRATE	Japanese Restaurant	45.483148	9.241998	4.417985	16	3
37	MACIACHINI - MAGGIOLINA	Asian Restaurant	45.497704	9.194891	3.452164	25	3
45	PADOVA	Japanese Restaurant	45.502142	9.230765	5.034007	14	3
65	S. CRISTOFORO	Sushi Restaurant	45.441999	9.159403	3.675857	14	3
69	SCALO ROMANA	Seafood Restaurant	45.438067	9.208088	3.476935	20	3
70	SELINUNTE	Seafood Restaurant	45.472688	9.138665	4.105937	15	3
79	VIALE MONZA	Chinese Restaurant	45.511446	9.225178	5.654620	13	3

Figure 14: Cluster 3

	Neighborhood	Recommendation	Latitude	Longitude	Dist_from_center	Number_venues	Cluster
0	MAGENTA - S. VITTORE	Seafood Restaurant	45.464689	9.169665	1.646077	32	1
1	TIBALDI	Seafood Restaurant	45.441302	9.180175	2.946347	36	1
2	UMBRIA - MOLISE	Seafood Restaurant	45.453199	9.219422	2.720618	33	1
3	VIGENTINA	Vegetarian / Vegan Restaurant	45.451087	9.191564	1.748021	35	1
4	WASHINGTON	Kebab Restaurant	45.461206	9.156310	2.745026	35	1

Figure 15: Cluster 1 with restrictions

Having said that, let's analyze further this cluster, by putting some restrictions on it. Let's say, we want those neighborhoods whose distance is not greater than 3 km and that have a number of existing restaurant which less than or equal to 40. Eventually, we have something like Figure 15.

By giving a look at the dataframe above, it is clear that the best neighborhoods are 'MAGENTA - S.VITTORE' and 'VIGENTINA'. In fact, they are the closest neighborhoods to the city center. Instead, the number of existing locations is the same among all 5 neighborhoods.

The two neighborhoods have also a similar distance from the city center, but looking at the "Recommendation" column, we can say, based on the “milangrouped” dataframe, that the recommended restaurant for 'MAGENTA - S.VITTORE' has a higher rank in “Popularity”, since is a Seafood restaurant, than a Vegetarian/Vegan restaurant, which is recommended for “VIGENTINA”.

Therefore, the best location is 'MAGENTA - S.VITTORE', where the ideal cuisine should be a restaurant that serves seafood.

5 Discussion and Conclusion

The results obtained are characterized by some simplifications. For example, we assumed that the popularity of a certain type of cuisine is proportional to the number of corresponding existing restaurants. In reality, one should have a database containing all the restaurants and their ratings, in order to decide which type of cuisine is the most popular among people.

Another aspect to take into consideration is how many people frequent the candidadte neighborhoods, which appear at first sight to be good candidates. In fact, it might be possible that in that neighborhood, where the distance from the city center and the number of existing venues have really good values, not so many people would be eager to pass their night. This is because it is also necessary to take into account the safety of the neighborhood and the presence of other places where to spend time before or after eating at the restaurant.

The costs of opening the new restaurant should also be accounted. In fact, land prices vary according to the place, and this must be considered. It is also possible to divide the neighborhoods in more clusters, in order to find further connections between neighborhoods in the same cluster.

We considered a radius of 800 metres, since this can be assumed on average as the distance one would be willing to walk after reaching the place. Of course, considering such a value for the radius implies taking into consideration certain venues not just for one neighborhood, but also for the other surrounding neighborhoods.

In general, we can see that is possible to make a decision based on two simple factors like the distance from the city center and the number of existing restaurants.