



Pietro Contegiacomo

Capstone Project

The Battle of Neighborhoods of Torino

1. Introduction

1.1 - Business problem

Last month Toshio, a Japanese friend of mine who lives in Tokyo, asked me: "is Torino the right city for starting my Japanese restaurant chain in Italy?"

Now, that I have the skills and the tools to use location data to explore a geographical location, I can answer the question.

For the uninitiated, Torino is the capital of Piedmont, in the northwest Italy, and was the first capital of the Kingdom of Italy from 1861 to 1865. Ranked third in Italy, after Milan and Rome, for economic strength, Torino is the world's 78th richest city by purchasing power, with a GDP of \$58 billion. The city has a rich culture and history, being known for its numerous art galleries, restaurants, churches, palaces, opera houses, piazzas, parks, gardens, theatres, libraries, museums and other venues. Turin is well known for its Renaissance, Baroque, Rococo, Neo-classical, and Art Nouveau architecture. Turin's attractions make it one of the world's top 250 tourist destinations.

The population of the city is estimate about 875,698 and I am a resident of this amazing city.

1.2 Target audience

This analysis could be useful for all business entrepreneurs who wants start a new business in Torino. They will prefer, I guess, the districts where there is a lower real estate cost and the type of business is less intense.

2 - Data

The city is split up into 8 boroughs, called "circoscrizioni", and 34 neighborhood, as you can see on this map:

https://github.com/pie88tro/Coursera_Capstone/blob/master/neighborhood-borough.png

The list of neighborhoods with the correspondent borough could be find in this PDF:

<http://www.comune.torino.it/statistica/osservatorio/annuario/2018/pdf/CAP%20%20Toponomastica%20ed%20Edilizia%202018.pdf>

Unfortunately, a dataset for Torino with boroughs-neighborhoods correlation, latitude, longitude and average land prices doesn't exist so I set-up my own data table and uploaded it on my github repository:

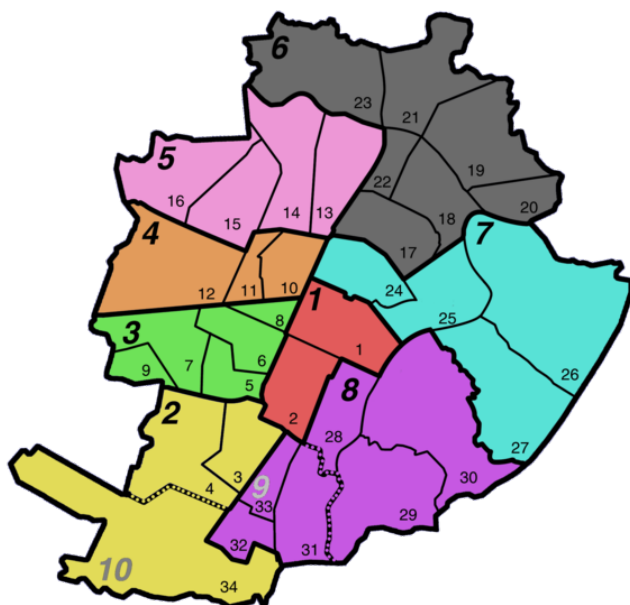
https://raw.githubusercontent.com/pie88tro/Coursera_Capstone/master/Torino_Neighborhoods.csv

I used Google Map, 'Search Nearby' option to get the center coordinates of the each Borough ("Latitude" and "Longitude" columns).

The column "AVG price (€/m²)" is the average land price for every neighborhood and I've taken it from those two websites:

<https://www.immobiliare.it/mercato-immobiliare/piemonte/torino/>

<https://www.mercato-immobiliare.info/piemonte/torino/torino.html>



Borough	Neighborhood	Latitude	Longitude	AVG Price (€/m²)
1	Centro	45.071767	7.678488	3263.5
1	Crocetta	45.058193	7.665198	2541.0
2	Santa Rita	45.046291	7.643089	1739.5
2	Mirafiori Nord	45.042951	7.629056	1689.5
2	Mirafiori Sud	45.021725	7.619190	1295.5
3	Borgo San Paolo	45.060186	7.642546	1891.0
3	Cenisia	45.069506	7.649560	1941.0
3	Pozzo Strada	45.068565	7.623186	1708.0
3	Cit Turin	45.074472	7.658915	2106.0
3	Borgata Lesna	45.055226	7.620605	3050.0
4	San Donato	45.082516	7.661298	1906.0
4	Parella	45.080732	7.626679	1683.0
4	Campidoglio	45.083313	7.648933	1881.0
5	Vallette	45.102862	7.631450	1143.0
5	Madonna di Campagna	45.104990	7.660830	1118.0
5	Borgata Vittoria	45.102334	7.674091	1146.0
5	Lucento	45.097289	7.649093	1193.0
6	Barriera di Milano	45.088973	7.693471	1020.0
6	Falchera	45.129054	7.710557	1140.0
6	Regio Parco	45.098183	7.709007	1728.5
6	Barca	45.109947	7.727843	1240.0
6	Bertolla	45.099267	7.744301	1315.0
6	Rebaudengo	45.106829	7.697552	1045.0
6	Villaretto	45.132154	7.681473	1050.0
7	Aurora	45.078634	7.684076	1245.0
7	Vanchiglia	45.070486	7.700462	2153.5
7	Madonna del Pilone	45.070867	7.740251	2350.0
7	Sassi	45.079092	7.730270	2200.0
8	San Salvario	45.053193	7.677227	2500.5
8	Nizza Millefonti	45.027288	7.665156	1594.5
8	Lingotto	45.031674	7.651271	1569.5
8	Borgo Po	45.049365	7.704225	2900.0
8	Cavoretto	45.031806	7.690380	2589.0

3. Methodology

3.1. Business Understanding

The aim of my project is to find the best neighborhoods in Torino for opening a restaurant chain.

3.2. Analytical Approach

For my analysis, for extracting the venues of each neighborhood, I will use the Foursquare API:

<https://developer.foursquare.com/docs/places-api/>

After Data Exploration, I will use a K-Means algorithm to extract the clusters, produce a map and make an argument on the final result.

The total number of neighborhoods in Torino are 34 so I need to find a way to cluster them based on their similarities (number and kind of restaurant).

3.3. Data Exploration

First, I define a function `getNearbyVenues()` to get venues info, using the Foursquare API with radius 500 and LIMIT 100, for all the neighborhoods in Torino.

```
def getNearbyVenues(names, latitudes, longitudes):  
    radius=500  
    LIMIT=100  
    venues_list=[]  
    for name, lat, lng in zip(names, latitudes, longitudes):  
        print(name)  
  
        # create the API request URL  
        url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{&radius={}&limit={}'.format(  
            CLIENT_ID,  
            CLIENT_SECRET,  
            VERSION,  
            lat,  
            lng,  
            radius,  
            LIMIT)  
  
        # make the GET request  
        results = requests.get(url).json()["response"]["groups"][0]["items"]
```

Then I analyze each neighborhood with one hot encoding.

Neighborhood	Abruzzo Restaurant	Asian Restaurant	Auditorium	Auto Garage	Auto Workshop	Automotive Shop	Bakery	Bar	Beer Bar	...	Theater	Toy / Game Store	Train Station	Tram Station	Trattoria/Osteria	Vegetarian / Vegan Restaurant
0	Centro	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0
1	Centro	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0
2	Centro	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0
3	Centro	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0
4	Centro	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0

5 rows x 142 columns

I group rows by neighborhood and by taking the mean of the frequency of occurrence of each category

	Neighborhood	Abruzzo Restaurant	Asian Restaurant	Auditorium	Auto Garage	Auto Workshop	Automotive Shop	Bakery	Bar	Beer Bar	...	Theater	Toy / Game Store	Train Station	Tram Station	Trattoria/Osteria
0	Aurora	0.00	0.014925	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0	...	0.000000	0.000000	0.00	0.000000	0.000000
1	Barca	0.00	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0	...	0.000000	0.000000	0.00	0.000000	0.000000
2	Barriera di Milano	0.00	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0	...	0.000000	0.071429	0.00	0.000000	0.000000
3	Bertolla	0.00	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0	...	0.000000	0.000000	0.00	0.000000	0.000000
4	Borgata Lesna	0.00	0.000000	0.000000	0.000000	0.000000	0.2	0.200000	0.200000	0.0	...	0.000000	0.000000	0.00	0.000000	0.000000
5	Borgata Vittoria	0.00	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0	...	0.000000	0.000000	0.00	0.000000	0.000000
6	Borgo Po	0.00	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.2	...	0.000000	0.000000	0.00	0.000000	0.000000
7	Borgo San Paolo	0.00	0.000000	0.000000	0.000000	0.000000	0.0	0.071429	0.000000	0.0	...	0.000000	0.000000	0.00	0.000000	0.000000
8	Campidoglio	0.00	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0	...	0.000000	0.000000	0.00	0.000000	0.000000
9	Cavoretto	0.00	0.000000	0.000000	0.000000	0.000000	0.0	0.000000	0.000000	0.0	...	0.000000	0.000000	0.00	0.000000	0.000000
10	Genisia	0.00	0.052632	0.000000	0.000000	0.000000	0.0	0.000000	0.052632	0.0	...	0.000000	0.000000	0.00	0.000000	0.000000

And finally, I create a new dataframe for displaying the top 10 venues for each neighborhood

	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	Aurora	Italian Restaurant	Cocktail Bar	Café	Pizza Place	Piedmontese Restaurant	Historic Site	Plaza	Vegetarian / Vegan Restaurant	Restaurant	Tapas Restaurant
1	Barca	Pizza Place	Italian Restaurant	Café	Soccer Field	Fish Market	French Restaurant	Food Stand	Food Court	Food & Drink Shop	Flea Market
2	Barriera di Milano	Pizza Place	Hostel	Pet Store	IT Services	Jewelry Store	Dive Bar	Toy / Game Store	Cultural Center	Café	Bus Stop
3	Bertolla	Piercing Parlor	Restaurant	Gym / Fitness Center	Pizza Place	Diner	Dive Bar	Dessert Shop	Eastern European Restaurant	Electronics Store	Emilia Restaurant
4	Borgata Lesna	Soccer Field	Sporting Goods Shop	Automotive Shop	Bakery	Bar	Fish Market	Fried Chicken Joint	French Restaurant	Food Stand	Food Court

3.4. Clustering

I use a K-means clustering, a type of unsupervised learning used with unlabeled data. The goal of this algorithm is to assign each data point (each neighborhood) to one of K groups based on the similarity features.

I use *KElbowVisualizer* to find the best k-number of cluster (called “the elbow method”).

```

#!pip install yellowbrick

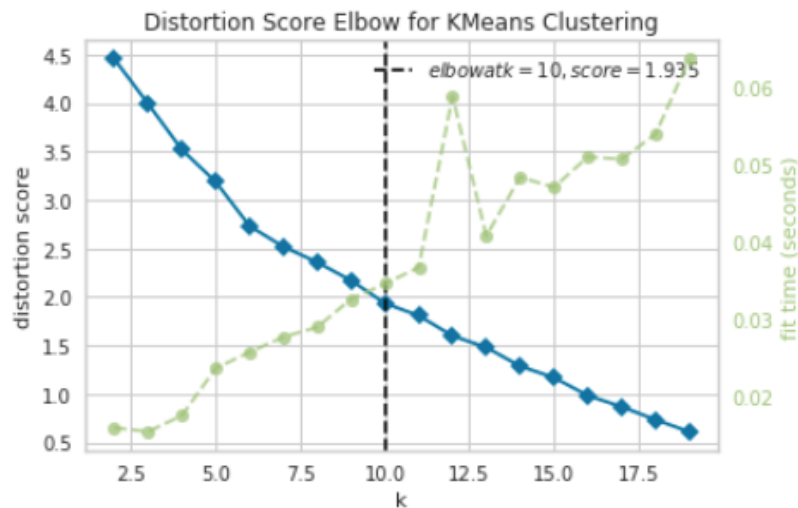
from yellowbrick.cluster import KElbowVisualizer

grouped_clustering = grouped.drop('Neighborhood', 1)

model = KMeans()
visualizer = KElbowVisualizer(model, k=(2,20))

visualizer.fit(grouped_clustering)
visualizer.poof()

```



The best K is 10. Now I can cluster neighborhoods, based on the venue categories, with K-Means clustering.

```

# set number of clusters
kclusters = 10

grouped_clustering = grouped.drop('Neighborhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]

```

Finally, we can visualize the results with a folium map:


```

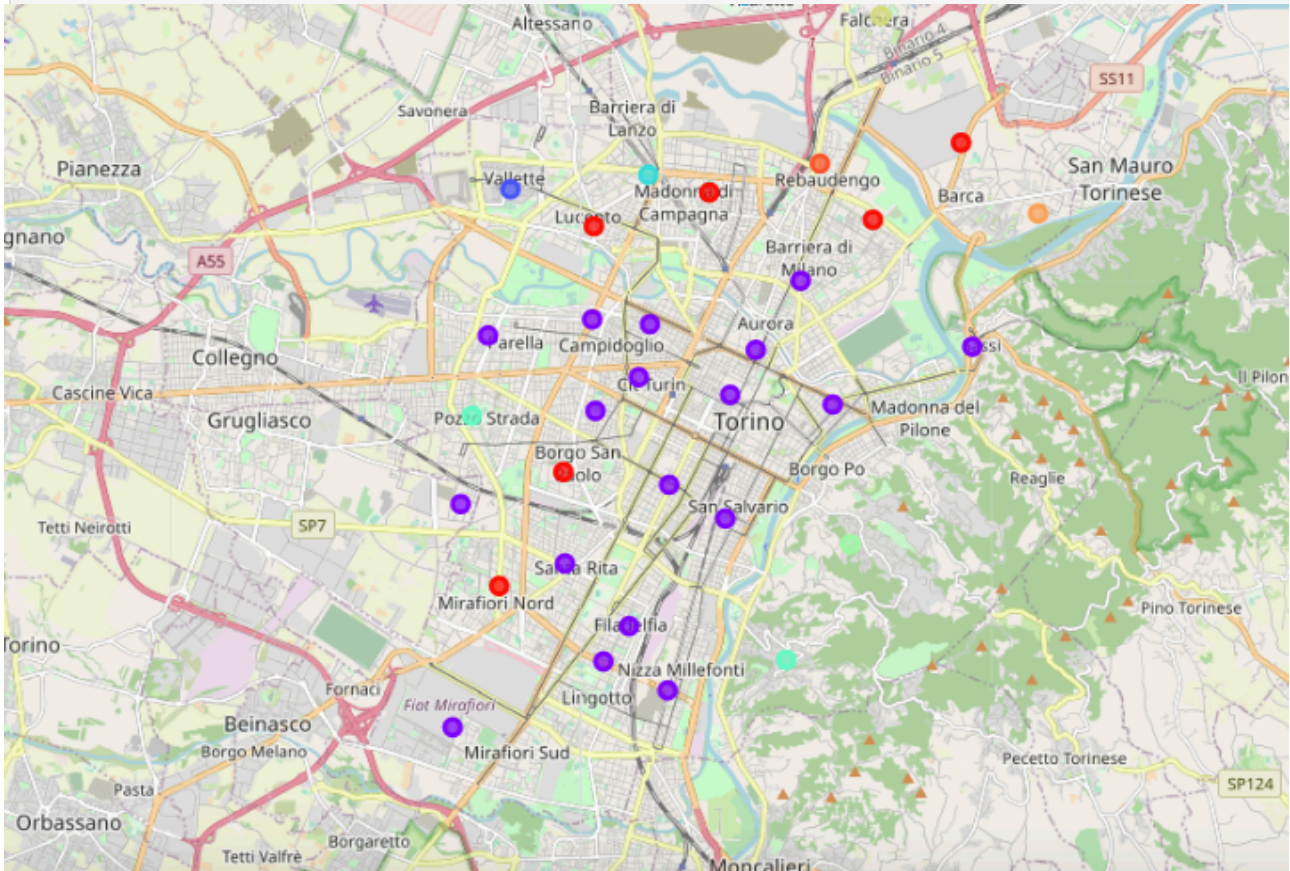
# create map
map_clusters = folium.Map(location=[45.071767, 7.678488], zoom_start=11)

# set color scheme for the clusters
x = np.arange(kclusters)
ys = [i + x + (i*x)**2 for i in range(kclusters)]
colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
rainbow = [colors.rgb2hex(i) for i in colors_array]

# add markers to the map
markers_colors = []
for lat, lon, poi, cluster in zip(merged['Latitude'], merged['Longitude'], merged['Neighborhood'], merged['Cluster Labels']):
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
        color=rainbow[cluster-1],
        fill=True,
        fill_color=rainbow[cluster-1],
        fill_opacity=0.7).add_to(map_clusters)

map_clusters

```



4. Result

Cluster 1

```
merged.loc[merged['Cluster Labels'] == 0, merged.columns[[1] + list(range(4,merged.shape[1]))]]
```

	Neighborhood	AVG Price (€/m²)	Cluster Labels	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	Mirafiori Nord	1689.5	0	Pizza Place	Health Food Store	Event Space	Seafood Restaurant	Supermarket	Gym	Ice Cream Shop	Italian Restaurant	History Museum	Hostel
5	Borgo San Paolo	1891.0	0	Pizza Place	Italian Restaurant	Bakery	Park	Café	Supermarket	Kids Store	Diner	Hobby Shop	Vegetarian / Vegan Restaurant
15	Borgata Vittoria	1146.0	0	Pizza Place	Gym	Italian Restaurant	Cupcake Shop	Food Stand	Food Court	Food & Drink Shop	Flea Market	Fish Market	Fast Food Restaurant
16	Lucento	1193.0	0	Italian Restaurant	Plaza	Soccer Stadium	Pizza Place	Diner	Dive Bar	Eastern European Restaurant	Electronics Store	Emilia Restaurant	Dessert Shop
19	Regio Parco	1728.5	0	Garden	Grocery Store	Supermarket	Furniture / Home Store	Italian Restaurant	Dive Bar	Eastern European Restaurant	Electronics Store	Emilia Restaurant	Event Space
20	Barca	1240.0	0	Pizza Place	Italian Restaurant	Café	Soccer Field	Fish Market	French Restaurant	Food Stand	Food Court	Food & Drink Shop	Flea Market

Cluster 2

```
merged.loc[merged['Cluster Labels'] == 1, merged.columns[[1] + list(range(4,merged.shape[1]))]]
```

	Neighborhood	AVG Price (€/m²)	Cluster Labels	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
0	Centro	3263.5	1	Italian Restaurant	Ice Cream Shop	Piedmontese Restaurant	Bistro	Hotel	Japanese Restaurant	Plaza	Pizza Place	Seafood Restaurant
1	Crocetta	2541.0	1	Café	Pizza Place	Ice Cream Shop	Women's Store	Sushi Restaurant	Coffee Shop	Pub	Restaurant	Monument / Landmark
2	Santa Rita	1739.5	1	Pizza Place	Ice Cream Shop	Dessert Shop	Café	Supermarket	Brewery	Market	Fried Chicken Joint	Movie Theater
4	Mirafiori Sud	1295.5	1	Garden	Hotel	Bus Station	Café	Women's Store	Fish Market	Food Stand	Food Court	Food & Drink Shop
6	Cenisia	1941.0	1	Ice Cream Shop	Pizza Place	Sushi Restaurant	Piedmontese Restaurant	Movie Theater	Food & Drink Shop	Coffee Shop	Cocktail Bar	Chinese Restaurant
8	Città Turin	2106.0	1	Café	Plaza	Pizza Place	Italian Restaurant	Piedmontese Restaurant	Chinese Restaurant	Market	Furniture / Home Store	Food Stand
9	Borgata Lesna	3050.0	1	Soccer Field	Sporting Goods Shop	Automotive Shop	Bakery	Bar	Fish Market	Fried Chicken Joint	French Restaurant	Food Stand
10	San Donato	1906.0	1	Pizza Place	Hotel	Café	Gift Shop	Eastern European Restaurant	Roman Restaurant	Food Court	Bus Stop	Sicilian Restaurant

11	Parella	1683.0	1	Ice Cream Shop	Supermarket	Theater	Furniture / Home Store	Bookstore	Miscellaneous Shop	Gym / Fitness Center	Tram Station	Auto Garage
12	Campidoglio	1881.0	1	Irish Pub	Supermarket	Japanese Restaurant	Jewelry Store	Park	Hotel	Rock Club	Café	Pizza Place
17	Barriera di Milano	1020.0	1	Pizza Place	Hostel	Pet Store	IT Services	Jewelry Store	Dive Bar	Toy / Game Store	Cultural Center	Café
24	Aurora	1245.0	1	Italian Restaurant	Cocktail Bar	Café	Pizza Place	Piedmontese Restaurant	Historic Site	Plaza	Vegetarian / Vegan Restaurant	Restaurant
25	Vanchiglia	2153.5	1	Italian Restaurant	Cocktail Bar	Café	Kebab Restaurant	Diner	Bar	Wine Bar	Trattoria/Osteria	Sandwich Place
26	Sassi	2200.0	1	Pizza Place	Hotel	Bus Stop	Cable Car	Trattoria/Osteria	Café	Food Stand	German Restaurant	Mexican Restaurant
27	San Salvario	2500.5	1	Italian Restaurant	Café	Science Museum	Pizza Place	Mexican Restaurant	Bistro	Bus Stop	Coffee Shop	Pub
28	Nizza Millefonti	1594.5	1	Pizza Place	Lingerie Store	Buffet	Auditorium	Chinese Restaurant	Optical Shop	Food Stand	Food Court	Food & Drink Shop
29	Lingotto	1569.5	1	Pizza Place	Hotel	Supermarket	Japanese Restaurant	Plaza	Pool	Rock Club	Chinese Restaurant	Bus Stop
32	Filadelfia	1750.0	1	Pizza Place	Pool	Light Rail Station	Supermarket	Chinese Restaurant	Sushi Restaurant	Pub	Soccer Stadium	Plaza

Cluster 3

```
merged.loc[merged['Cluster Labels'] == 2, merged.columns[[1] + list(range(4,merged.shape[1]))]]
```

	Neighborhood	AVG Price (€/m²)	Cluster Labels	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
13	Vallette	1143.0	2	Lounge	Women's Store	Garden	Fried Chicken Joint	French Restaurant	Food Stand	Food Court	Food & Drink Shop	Flea Market	Fish Market

Cluster 4

```
merged.loc[merged['Cluster Labels'] == 3, merged.columns[[1] + list(range(4,merged.shape[1]))]]
```

	Neighborhood	AVG Price (€/m²)	Cluster Labels	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
23	Villaretto	1050.0	3	Rest Area	Restaurant	Women's Store	Farmers Market	Food Stand	Food Court	Food & Drink Shop	Flea Market	Fish Market	Fast Food Restaurant

Cluster 5

```
merged.loc[merged['Cluster Labels'] == 4, merged.columns[[1] + list(range(4,merged.shape[1]))]]
```

	Neighborhood	AVG Price (€/m²)	Cluster Labels	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
14	Madonna di Campagna	1118.0	4	Abruzzo Restaurant	Bus Station	Train Station	Chinese Restaurant	Fish Market	French Restaurant	Food Stand	Food Court	Food & Drink Shop	Flea Market

Cluster 6

```
merged.loc[merged['Cluster Labels'] == 5, merged.columns[[1] + list(range(4,merged.shape[1]))]]
```

	Neighborhood	AVG Price (€/m²)	Cluster Labels	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
7	Pozzo Strada	1708.0	5	Gym	Pool	Italian Restaurant	Trattoria/Osteria	Café	Women's Store	Food Stand	Food Court	Food & Drink Shop	Flea Market
31	Cavoretto	2589.0	5	Italian Restaurant	Park	Bus Station	Café	Women's Store	Food Stand	Food Court	Food & Drink Shop	Flea Market	Fish Market

Cluster 7

```
merged.loc[merged['Cluster Labels'] == 6, merged.columns[[1] + list(range(4,merged.shape[1]))]]
```

	Neighborhood	AVG Price (€/m²)	Cluster Labels	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
30	Borgo Po	2900.0	6	Plaza	Italian Restaurant	Mini Golf	Restaurant	Beer Bar	Fast Food Restaurant	Food Stand	Food Court	Food & Drink Shop	Flea Market

Cluster 8

```
merged.loc[merged['Cluster Labels'] == 7, merged.columns[[1] + list(range(4,merged.shape[1]))]]
```

	Neighborhood	AVG Price (€/m²)	Cluster Labels	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
18	Falchera	1140.0	7	Bus Station	Bus Stop	Women's Store	Dessert Shop	French Restaurant	Food Stand	Food Court	Food & Drink Shop	Flea Market	Fish Market

Cluster 9

```
merged.loc[merged['Cluster Labels'] == 8, merged.columns[[1] + list(range(4,merged.shape[1]))]]
```

	Neighborhood	AVG Price (€/m²)	Cluster Labels	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
21	Bertolla	1315.0	8	Piercing Parlor	Restaurant	Gym / Fitness Center	Pizza Place	Diner	Dive Bar	Dessert Shop	Eastern European Restaurant	Electronics Store	Emilia Restaurant

Cluster 10

```
merged.loc[merged['Cluster Labels'] == 9, merged.columns[[1] + list(range(4,merged.shape[1]))]]
```

	Neighborhood	AVG Price (€/m²)	Cluster Labels	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
22	Rebaudengo	1045.0	9	Café	Plaza	Motorcycle Shop	Supermarket	Fish Market	Food Stand	Food Court	Food & Drink Shop	Flea Market	Fast Food Restaurant

5. Conclusion

1. The most common venues in Torino are Pizza Place and Italian Restaurant (obviously!)
2. Japanese and Sushi Restaurants are not very common in Torino, so that could be a great business for Toshio. They are frequent only in Cluster 2 (Cenisia, Campidoglio, Lingotto, Filadelfia, Crocetta and Centro Neighborhoods). Crocetta is an elegant residential neighborhood, Centro is the center of the city (the most expensive neighborhood in Torino) and Campidoglio is a multicultural neighborhood (perfect for walking, with a lot of artworks and murales). Lingotto and Filadelfia are bordering (in this area there were the FIAT factories now re-evaluated)
3. After my analysis, I suggest to my friend starting his business at Borgo Po, an expensive but strategic neighborhood. It is near the center but less expensive and less mess (here customers can find parking easily), elegant, with a lot of history and tourist places and, last but not least, it's the neighborhood where I live :)
4. Later, he could expand his restaurant chain in Vanchiglia, Cit Turin and San Salvario neighborhoods, wich have a good quality/price ratio and strategic position for nightlife.

Thank you for your time!

Pietro Contegiacomo