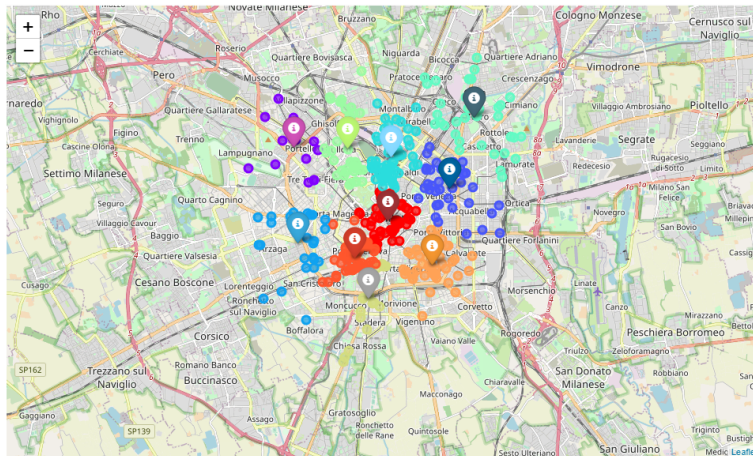


Food Delivery Optimization in Milan

Daniele Fistolera



Contents

1	Introduction	2
1.1	Business Problem	2
2	Data Gathering	2
3	Methodology	4
4	Results	5
5	Discussion	10
6	Conclusion	10

1 Introduction

In this project we will find the best placements for a number, in this case 10, of food couriers in Milan to optimize the delivery service with faster and better deliveries and also improve working conditions of couriers who will have to travel less distance.

1.1 Business Problem

In Milan there are many food delivery services and many more restaurants, so the optimal organization between these two fundamental components of the food business is essential to create an efficient and profitable system.

The target audience interested in this project are the food delivery companies who work in the area of Milan because finding the optimal location in which the couriers will wait the next order is useful to reduce the time necessary to reach the restaurant, so the meal will be better, the client will wait less and the courier will travel less distance.

2 Data Gathering

Using the Foursquare API we will retrieve the data of the restaurants and food services in the area of Milan, which has a location of $45^{\circ}28'01''N$ and $09^{\circ}11'24''E$. For each activity we will retrieve name, position and type of cuisine.

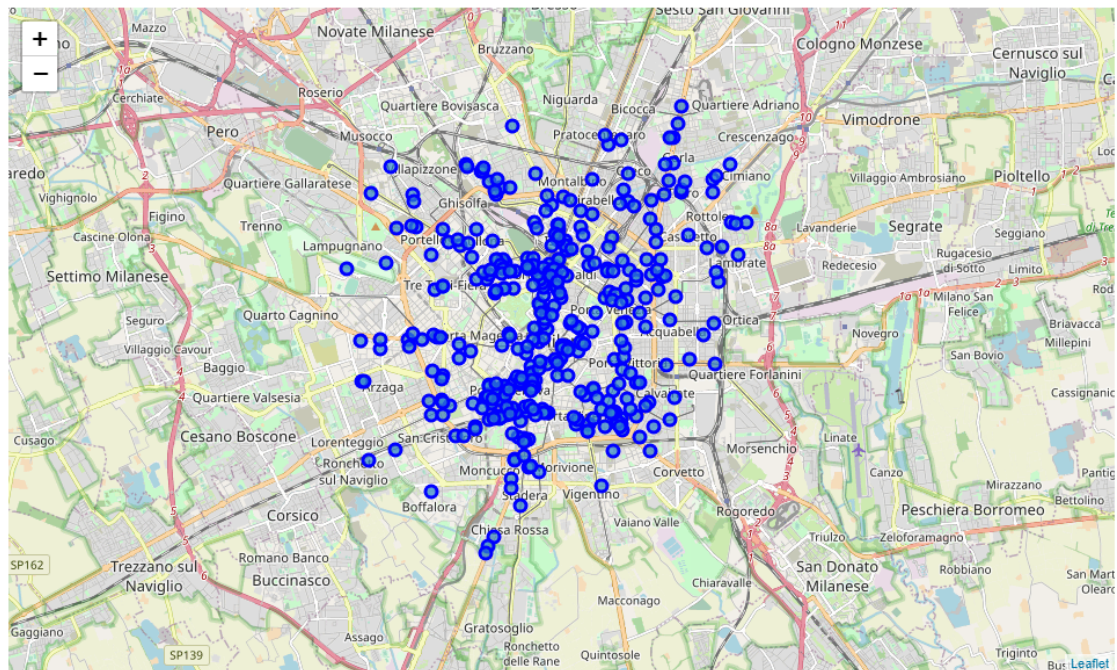
The Foursquare urls have a maximum number of results of 100, so are created nine different urls requests at different locations in the area of Milan to maximize the results.

An example of the dataset is showed below:

	Name	Category	Latitude	Longitude
0	Luini	Bakery	45.465707	9.191431
1	Spontini	Pizza Place	45.465655	9.191222
2	Langosteria Café	Bistro	45.464760	9.194899
3	La Veranda	Mediterranean Restaurant	45.469293	9.195544
4	Galleria Ristorante Pizzeria	Italian Restaurant	45.465881	9.190025
...
92	Ristorante Erba Brusca	Italian Restaurant	45.422867	9.167307
93	BA Asian Mood	Asian Restaurant	45.468122	9.151817
94	Cacio e Pepe	Italian Restaurant	45.452780	9.181500
98	Trapizzino	Sandwich Place	45.467483	9.154074
99	Zibo - Cuochi Itineranti, Campo Base	Bistro	45.459180	9.178136

388 rows × 4 columns

The map of Milan with the venues is showed below:



3 Methodology

Through a K-means algorithm with 10 point we will define the locations which minimize the distance between the couriers and the restaurants and finally we will analyze the composition of each cluster of restaurants by their cuisine.

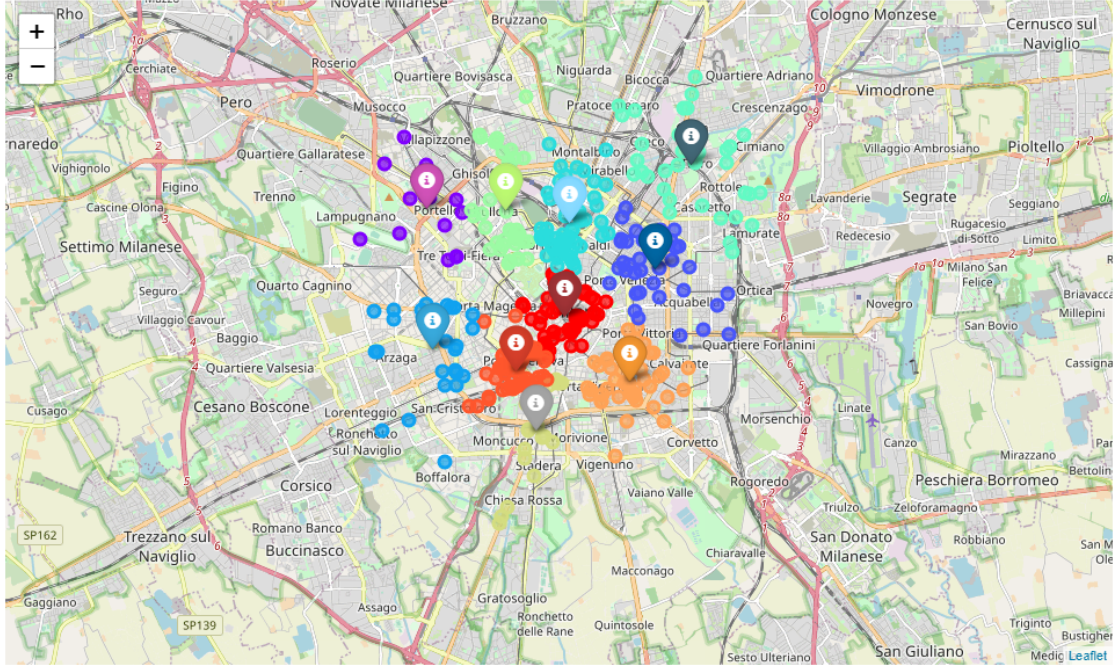
After data selection from the Foursquare result dictionary, conversion to a dataframe and elimination of duplicates the coordinates of the venues are clustered.

The K-means algorithm is implemented using the method “sklearn.cluster.KMeans” with 10 centers initiated on a 3x3 grid of points with the center at the Milan coordinates and spaced by a coordinate value corresponding to $2.5Km$. The algorithm produce 10 pairs of coordinates that localize the cluster centers.

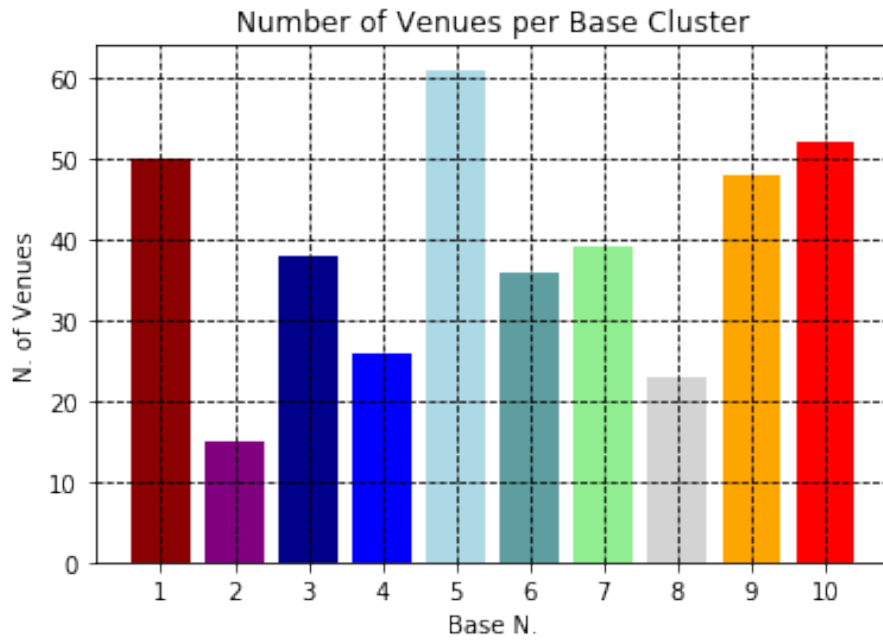
Later each cluster is analysed by counting the number of venues, by calculating the mean distance between the locations of the centers and the venues and finally by counting the occurrences of cuisine categories for each base cluster.

4 Results

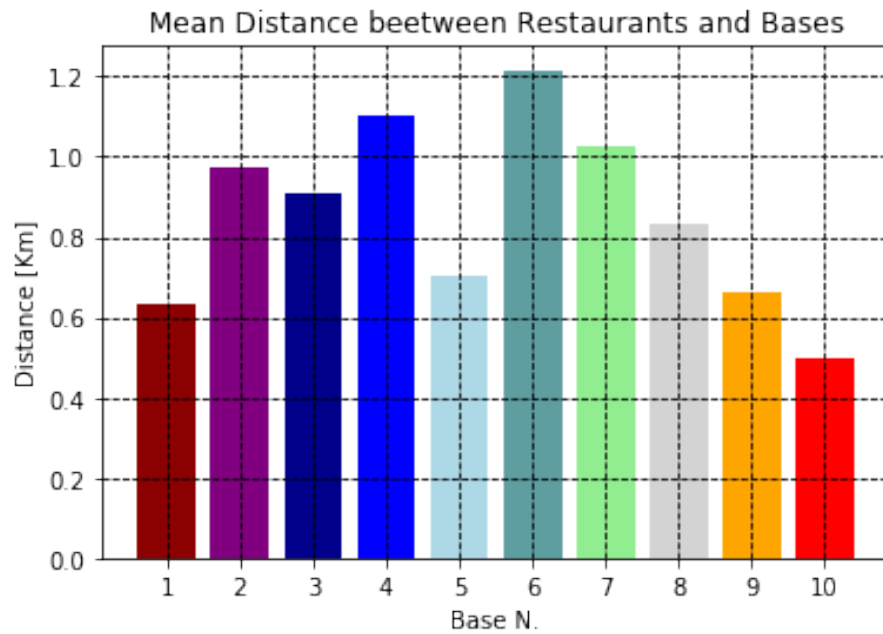
The k-means cluster algorithms permit to localize 10 different courier bases which maximize the distances between each other, but minimize the distances between the cluster centers and the venues. Each courier could be assigned a base location which will be used as a reference hub to wait for next order. A map of milan showing the clusters and the base locations is showed below:



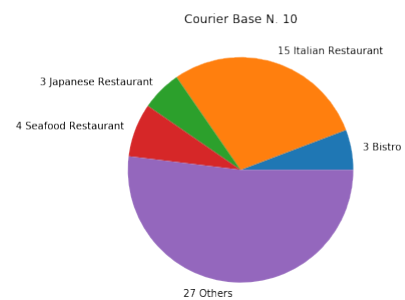
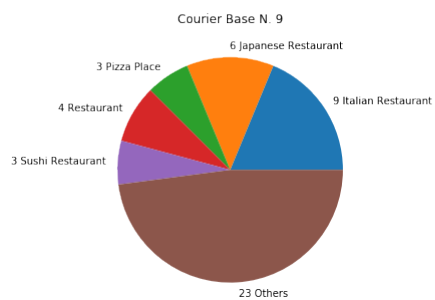
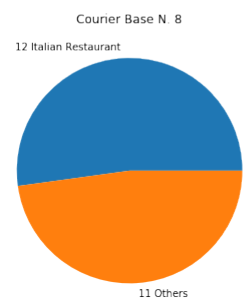
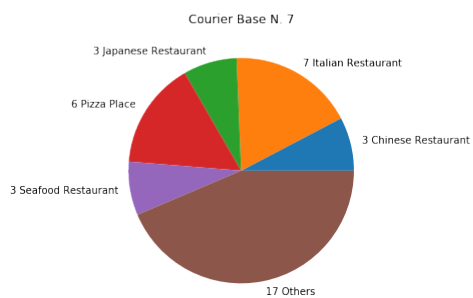
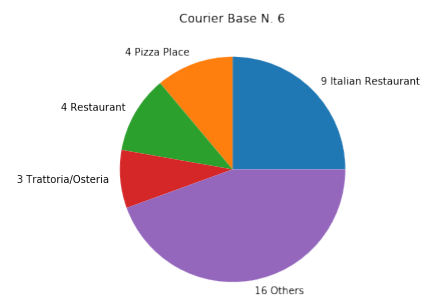
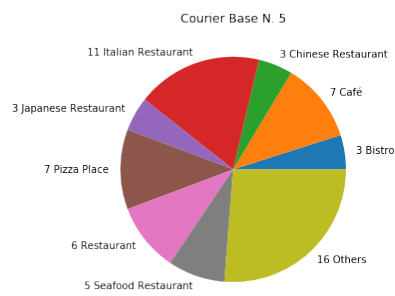
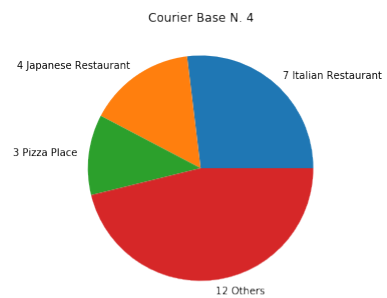
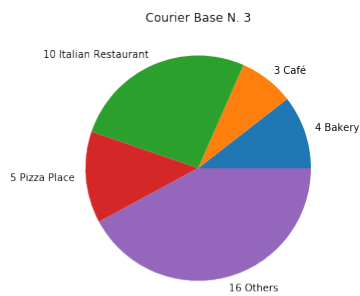
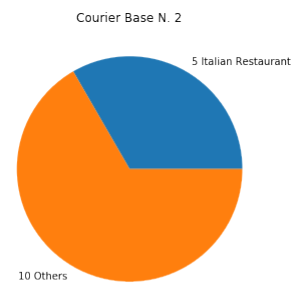
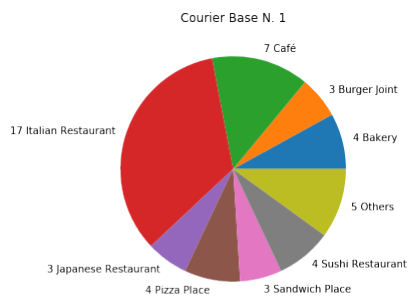
The clusters have different numbers of venues: the base cluster number 2 has only 15 venues, instead the base number 5 is up to 61 venues. This difference is an important characteristic of each cluster because is an indicator of the number of orders that each base will receive.



Another important feature of each cluster is the mean distance between bases and venues because it is an index of how long it will take for the courier to reach the restaurant and of the length of the route. Clusters number 2, 4, 6 and 7 are the ones with the biggest mean distance.



The pie charts shows that there are basically two types of clusters: the ones with a predominant type of cuisine, as bases 2 and 8 with "Italian Restaurants", and the ones with a very diverse set of important categories.



5 Discussion

The results of the clusterization showed that is possible to locate an arbitrary number of bases, in this case 10, to optimize the connection between food couriers and restaurants. The number of venues per cluster varies for each base, so is a good suggestion to assign more couriers to cluster centers with more venues.

At the same time the graph of the mean distance between bases and venues per cluster shows that in general the cluster with less venues have a greater mean distance, for example clusters 2, 4 and 8. The most critique situations are clusters 3 and 7 which have both a large number of venues and a considerable mean distance, so it is possible that they are the bases that need more staff.

The pie charts could be used to optimize the gear of the couriers assigned to a specific base, for example if the cluster is characterized by a majority of pizza services provide the courier with a bigger or different case to improve the service.

6 Conclusion

In conclusion the principal objective of finding the position of 10 courier basis in Milan to optimize the food delivery service has been satisfied. By creating this bases for the couriers the food delivery companies could improve their service through the reduction of delivery times and also improving the working condition of couriers by reducing the distances traveled. With the additional information obtained from the analysis of each cluster is possible to decide the correct number of staff necessary for each base and provide them with the best gear for the predominant type of cuisine of the cluster.