

THE BATTLE OF NEIGHBORHOODS IN GENOVA



Matteo Zangani

PROBLEM

The goal of this project is designed to analyze the opportunity of relocation for our family in one of the neighborhood in the city of Genova since a new job opportunity arise.

Each neighborhood will be explored to analyze their features and try to be able to identify the best places to relocate.

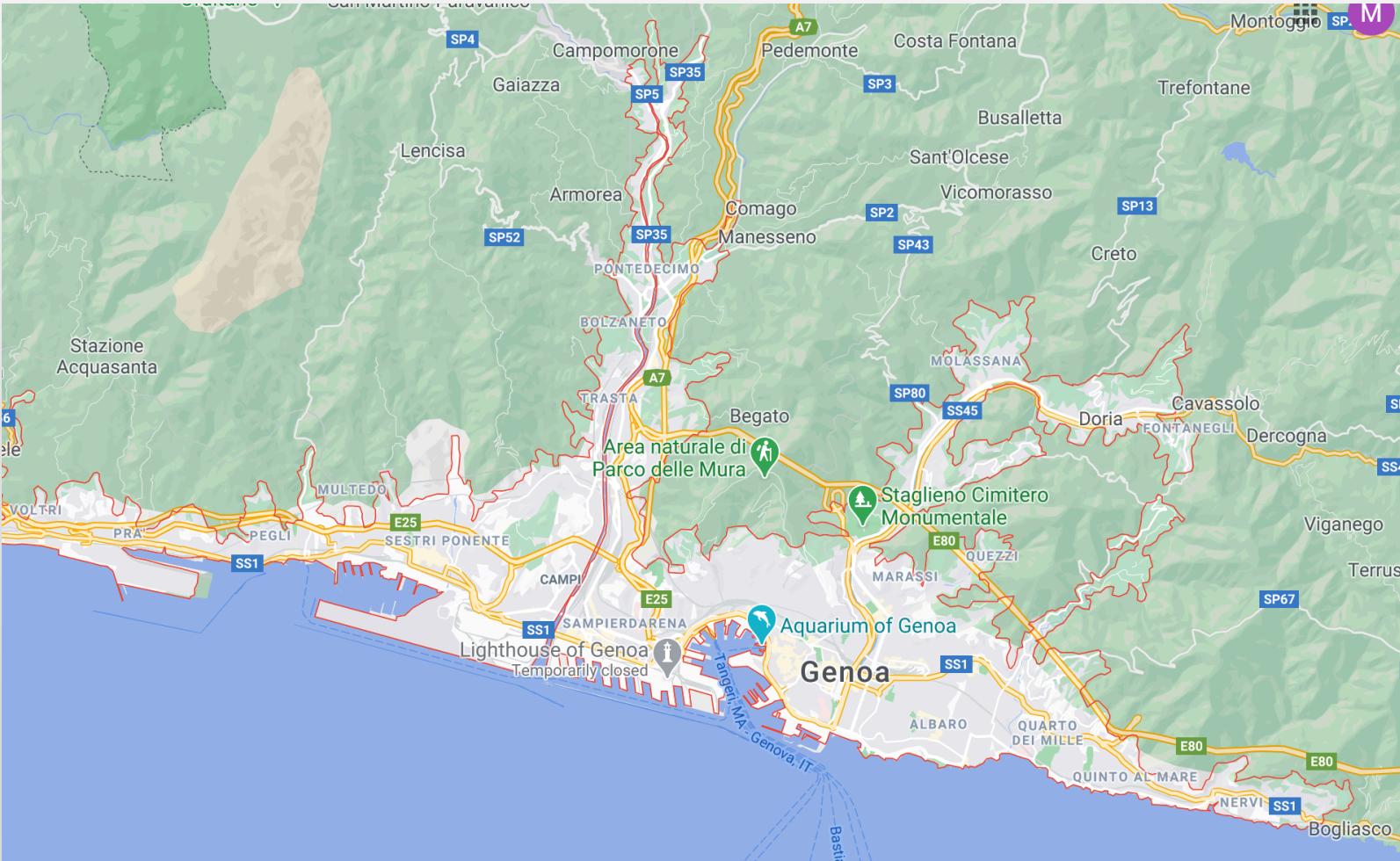
INTRODUCTION

Genova is a city in the North of Italy with sustainable industrial development.

The city is on the Mediterranean Sea and has a wonderful climate to live.



MAP OF GENOVA



Data from google map

DATA

The name and number of Neighborhoods in Genova will be taken from the website mercato-immobiliare.info

The website also gives the average sell prices for houses, which will help the final decision on which area is more convenient within the chosen clusters.

From wikipedia website we can assign the municipalities values to each Neighborhood

Using Foursquare we will explore venues of each Neighborhood and understand what are the main characteristics of each area.

METHODOLOGY

The data will be imported and analyzed to identify the most interesting areas of Genova.

By segmenting and clustering the different Neighborhoods we will be able to identify common characteristics between areas which will give us more options to choose the best places

Housing price and location will drive the final decision on which is the best area for searching a new house

DATA EXAMPLE

	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
0	Albaro	Café	Hotel	Pizza Place	Beach	Plaza	Ice Cream Shop	Cocktail Bar	Restaurant	Harbor / Marina
1	Apparizione	Pizza Place	Bakery	Sports Club	Gym / Fitness Center	Convenience Store	Gym Pool	Plaza	Soccer Field	Electronics Store
2	Bavari	Italian Restaurant	Diner	Stables	Café	Food & Drink Shop	Garden	Furniture / Home Store	Fried Chicken Joint	French Restaurant
3	Bolzaneto	Pizza Place	Toy / Game Store	Stadium	Coffee Shop	Health Food Store	Wine Shop	Flea Market	Fried Chicken Joint	French Restaurant
4	Borzoli	Soccer Field	Rock Club	Train Station	Furniture / Home Store	Bakery	Stadium	Garden	Fried Chicken Joint	French Restaurant

After the data has been cleaned and imported, all the available venues have been retrieved and merged. Here is a table summary with displayed the most common venues. A searching radius of 500m have been used from the center of each Neighborhood
The price value was not used for modeling as it would interfere too much in the clustering process

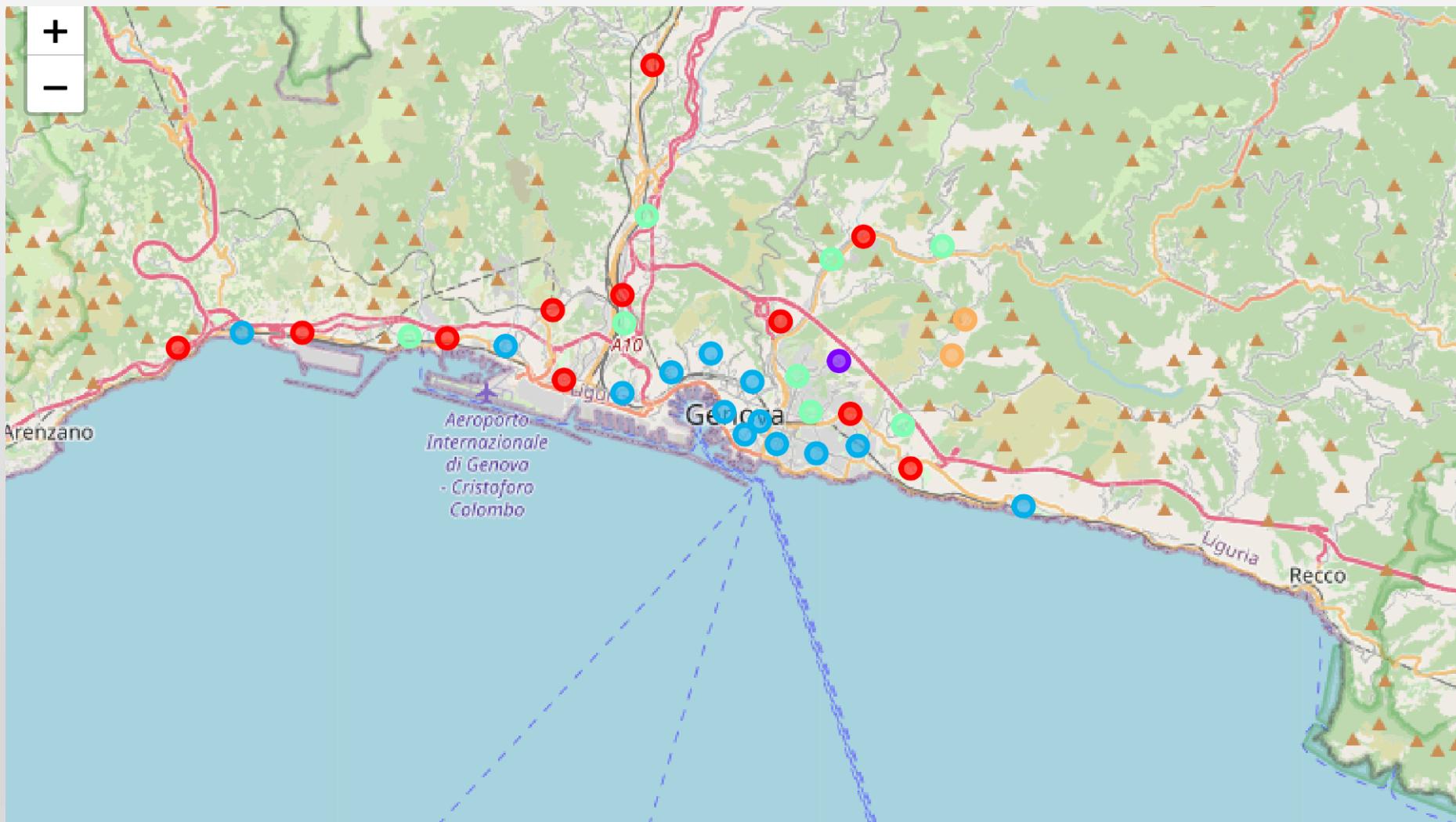
CLUSTERING

The algorithm used is **k-nearest neighbors algorithm (k-NN)** is a non-parametric classification method. The k-nearest neighbors algorithm is a simple, supervised machine learning algorithm that can be used to solve both classification and regression problems.

After inspecting the dataset several k segmentation values were tested.
An optimal segmentation that yields a sensitive result is k=5

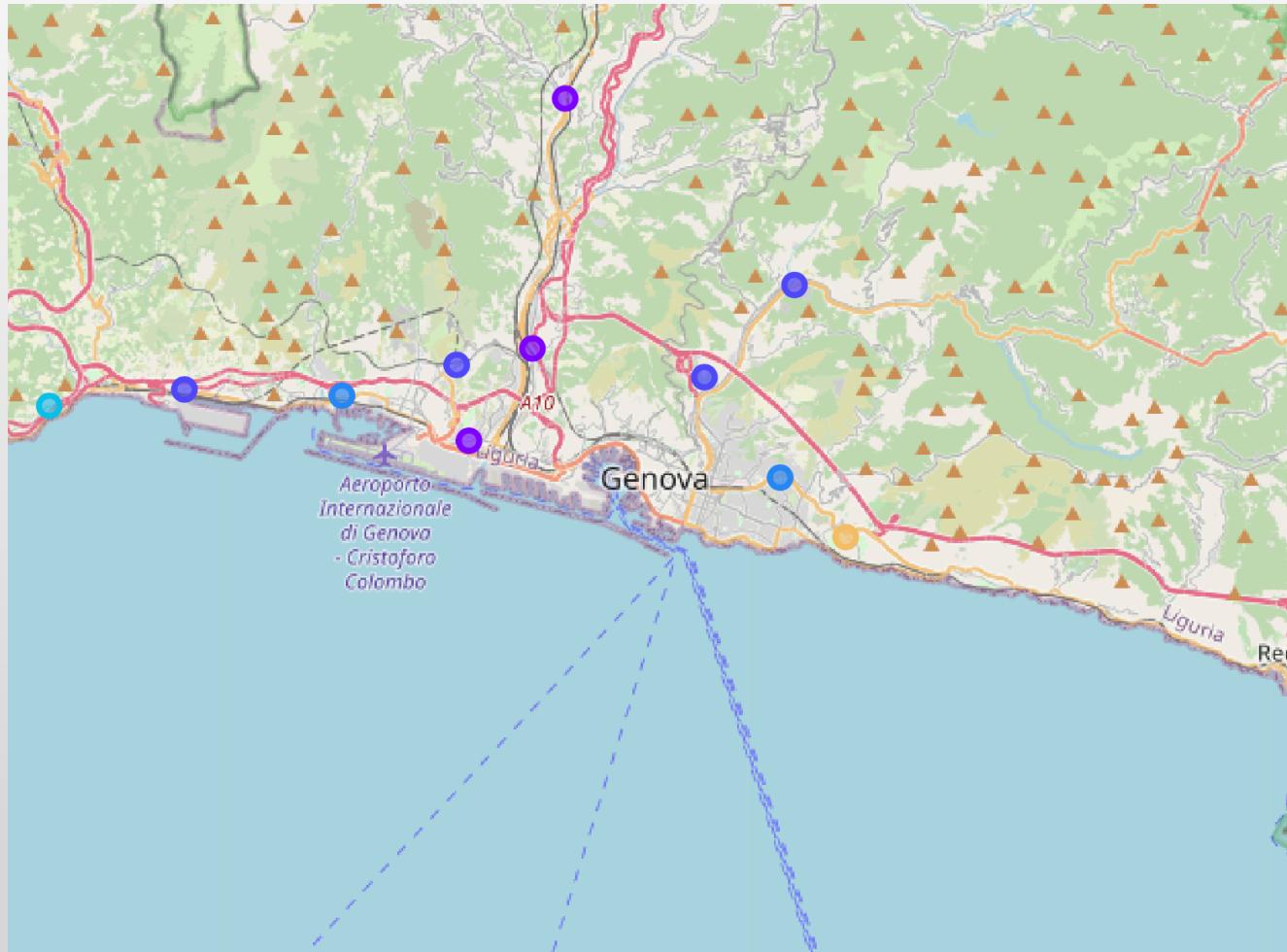
Data was segmented by using all the venues which reflect the character of each Neighborhood

MAP OF NEIGHBORHOODS AFTER CLUSTERING



The color of each neighborhood corresponds to the Cluster differentiation

DATA ANALYSIS



Cluster 1

Cluster 1 shows medium population of Neighborhoods

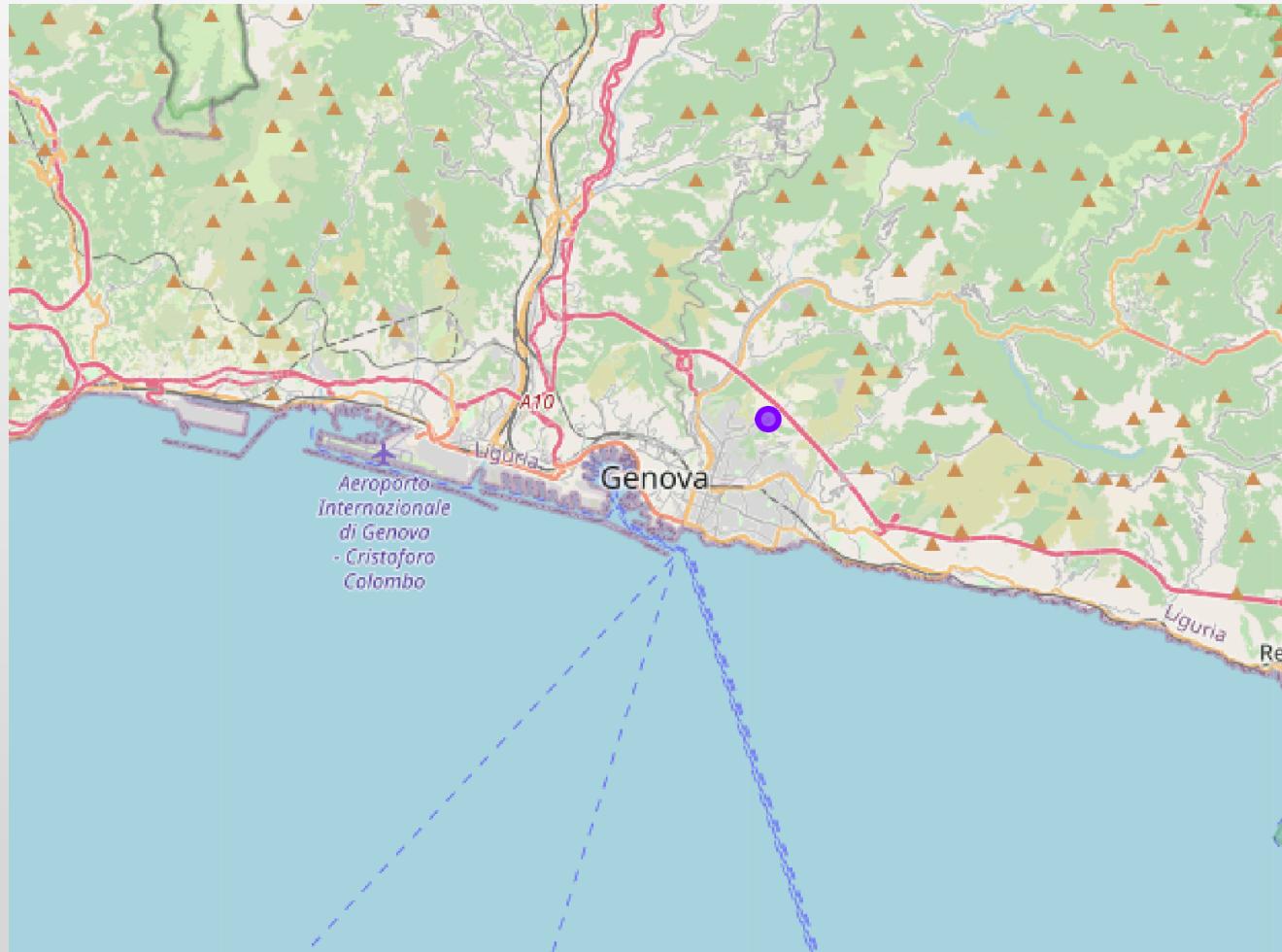
The color shows the price per square meter of the housing market

As we can see from the venues data frame and the value of the property per square meter this cluster is quite homogeneous.

By the activities explored we can see this is mostly residential, however is very visible that the housing price is lower.

This can be an index of a lower level, so we do not consider it for relocation

DATA ANALYSIS



Cluster 2

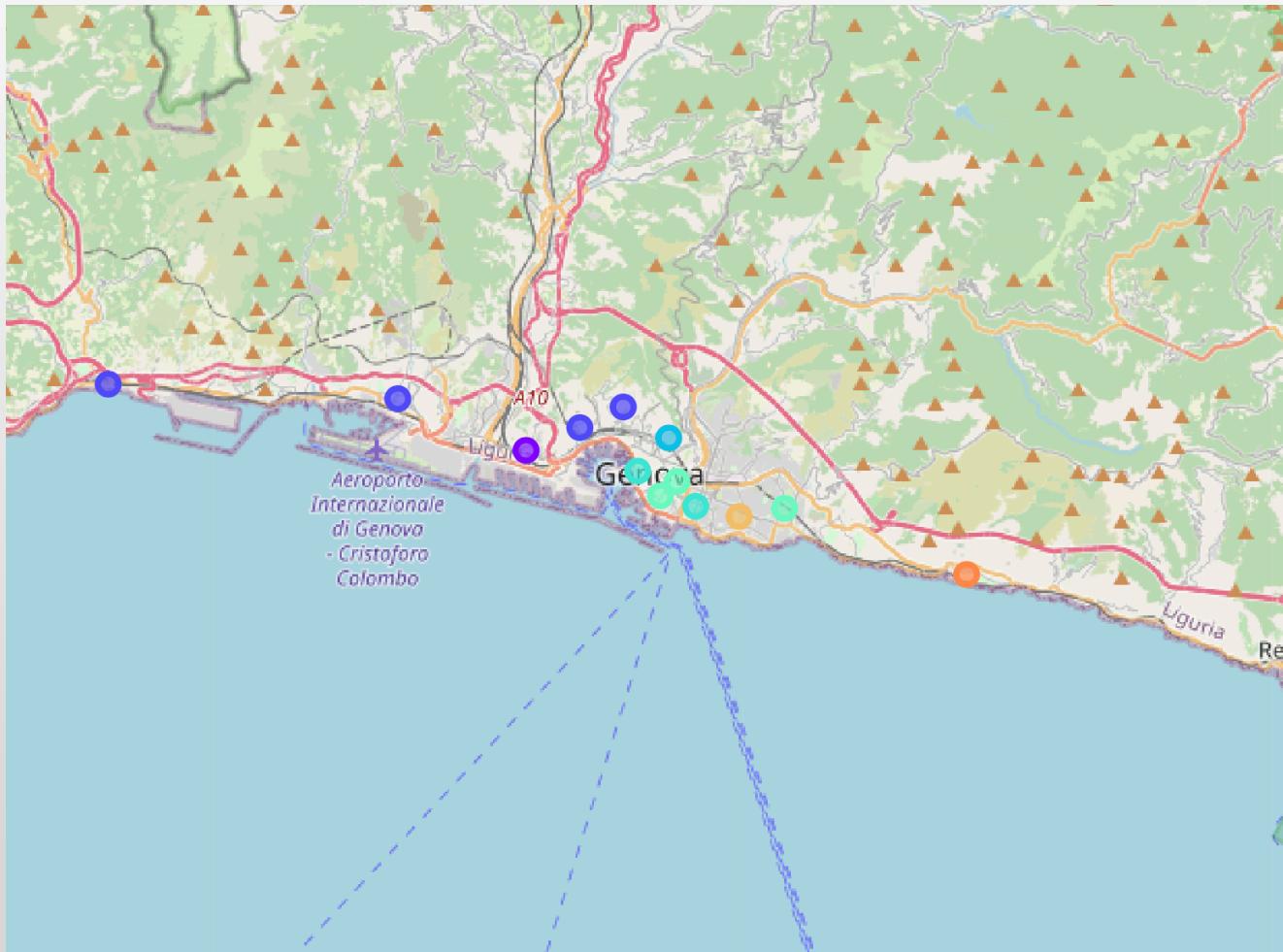
Cluster 2 shows a single population of Neighborhoods

The color shows the price per square meter of the housing market

This cluster is very isolated

Most likely the Neighborhood is very small and the number of venues are limited

DATA ANALYSIS



Cluster 3

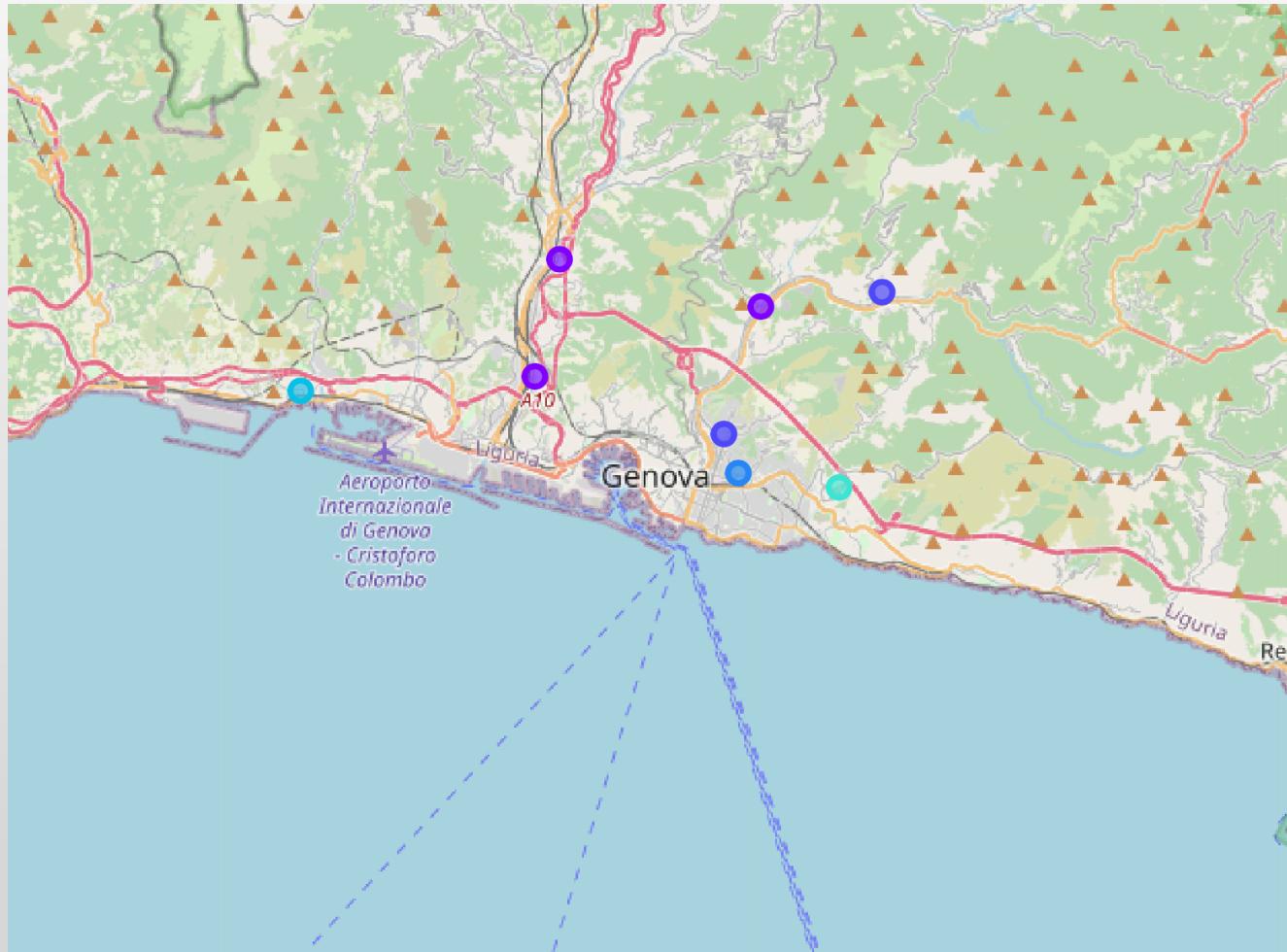
Cluster 3 shows a larger population of Neighborhoods

The color shows the price per square meter of the housing market

As we can see from the venues data frame and the value of the property per square meter this cluster is quite heterogeneous.

By the activities explored we can see this is mostly residential with the property more expensive on the eastern end. Most likely d/t to a better access to the coast and the beaches

DATA ANALYSIS



Cluster 4

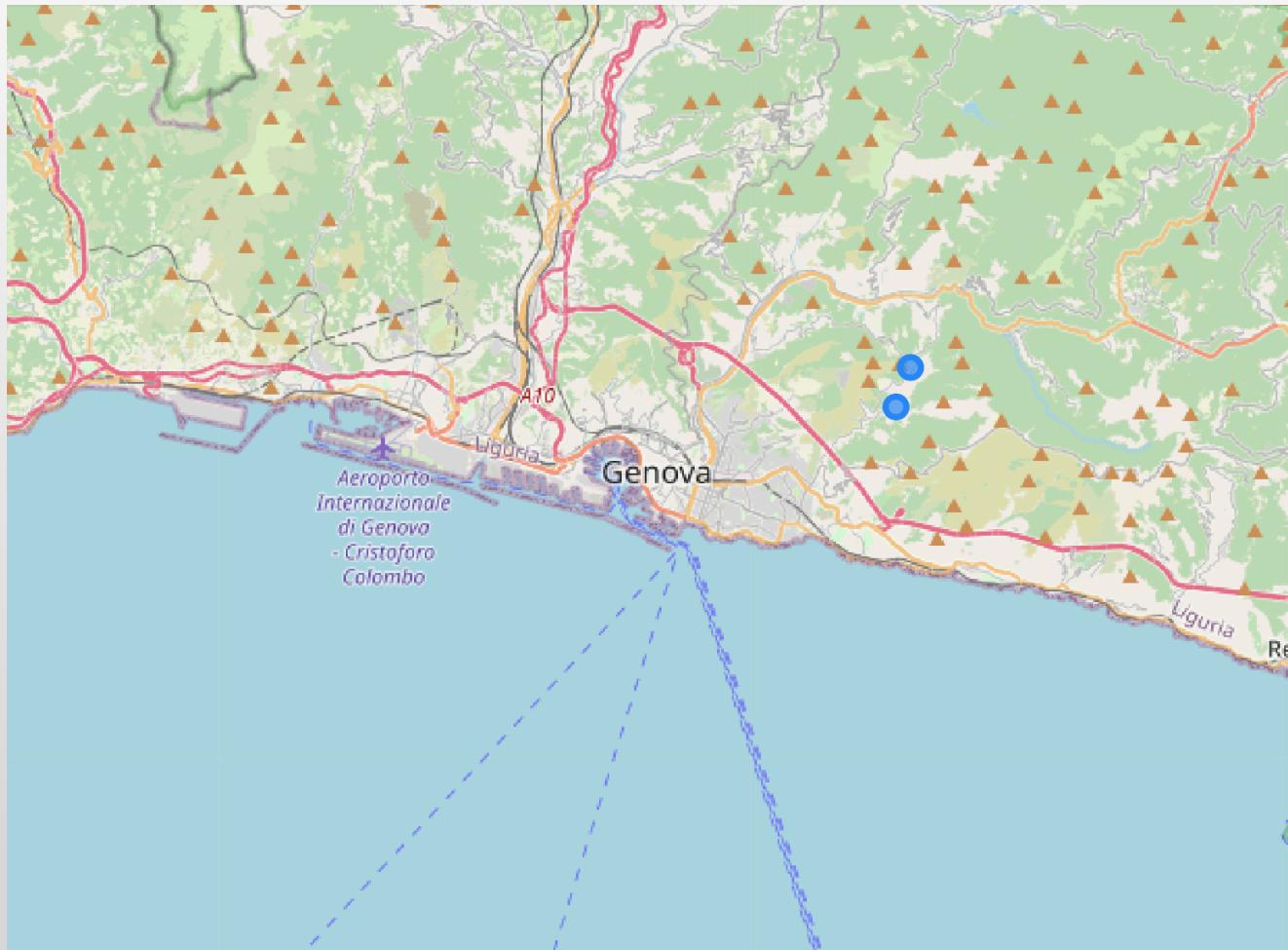
Cluster 4 shows a medium population of Neighborhoods

The color shows the price per square meter of the housing market

As we can see from the venues data frame and the value of the property per square meter this cluster is fairly homogeneous.

By the activities explored and from the information gathered, we can see this area is mostly driven by industrial activities and is not very suitable for relocation

DATA ANALYSIS



Cluster 5

Cluster 5 shows a small population of Neighborhoods

The color shows the price per square meter of the housing market

As we can see from the venues data frame and the value of the property per square meter this cluster is very homogeneous.

By the activities explored we can see this is only residential .

However is quite far from seaside, so we decided to exclude it from final choice

CONCLUSIONS

By analyzing the data we can see that cluster 4 is the most interesting and most neighborhood sits next to the seaside.

However the housing price can be very high and the crowd city center is avoided.

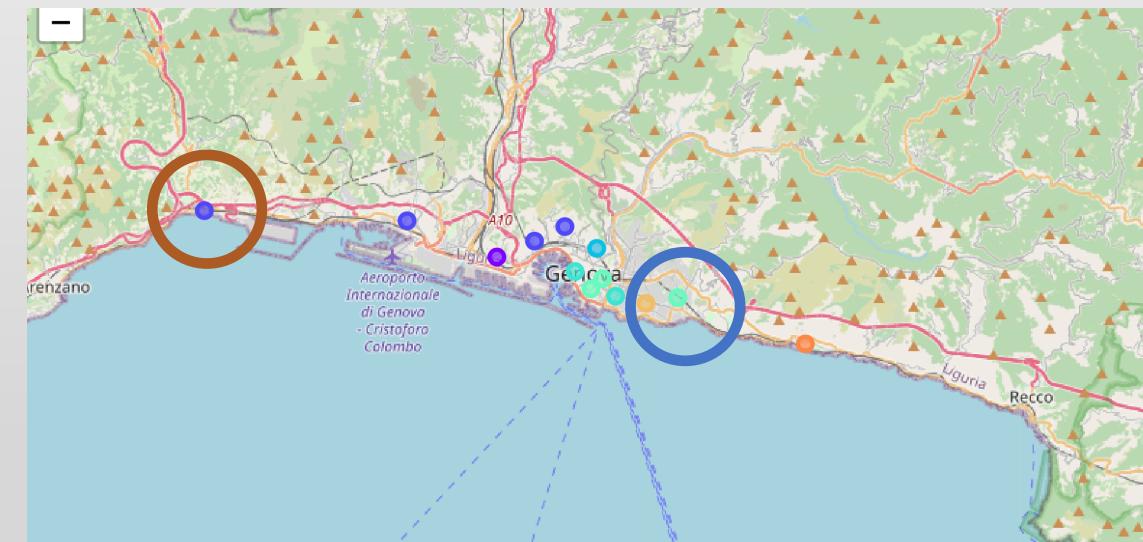
So two interesting areas are identified for relocation.

First area of interest is delimited by the blue circle (Sturla Neighborhood).

This area is very interesting: it is close to the city center but a bit away from the main crowd places. Bigger limitation is the average price which is fairly high (2450 Euro/m²)

The second area of interest is marked in the orange circle (Voltri Neighborhood).

It sits outside the main harbor area and is on the seaside. Housing price is far more affordable (1450 Euro/m²). Main limitation is the distance from city center which needs to be considered



Cluster 3