

APPLIED DATA SCIENCE CAPSTONE - First Week

Franco Zanini

Customers of bed and breakfast companies, such as Airbnb, choose their apartments looking at different parameters. Price and interiors are of course the first thing to consider, but location is also important. Using Foursquare location data to compare different neighborhoods could be a good opportunity to choose, for example, an area with good restaurants or theatres and cinemas. I was thinking to design my capstone project analyzing the data of my own town, Trieste, a very lively place with a lot of venues for tourists.

Unfortunately, the division of Trieste in seven boroughs ("*circoscrizioni*") and twentyeight neighborhoods ("*rioni*") is rather cumbersome, and some *rioni* are divided between different *circoscrizioni*, as shown in figure 1.

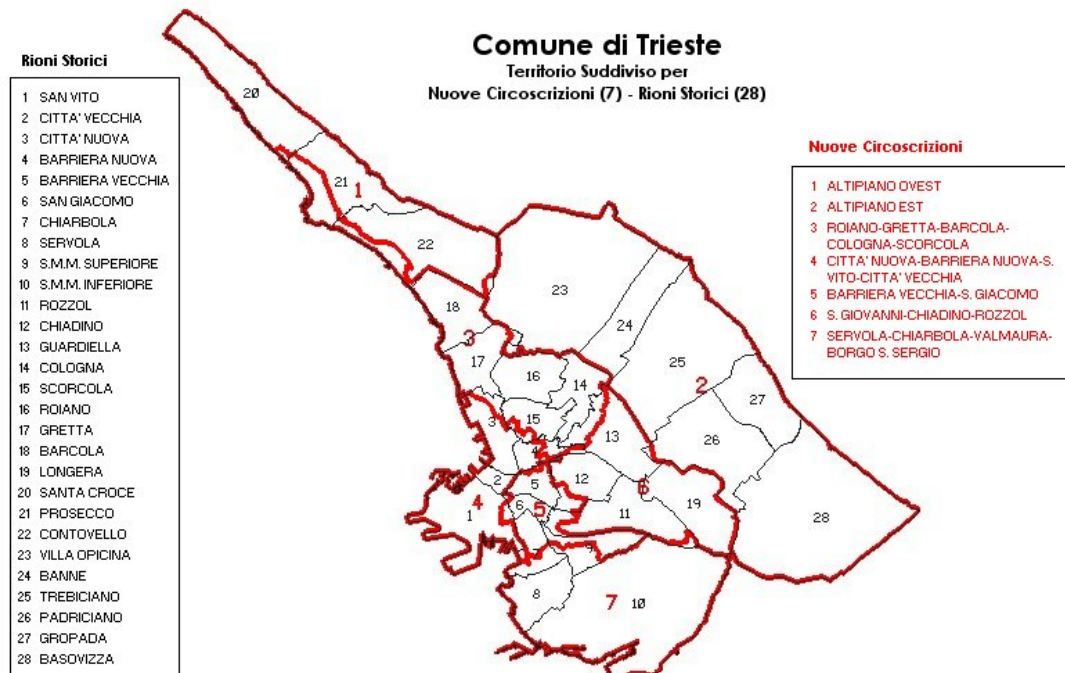


Fig. 1. Trieste divided by circoscrizioni (in red) and rioni (in black).

A more recent division of the town in 43 *zone di quartiere* could solve this problem, but a database with geographical coordinates of the zone has not been provided yet by the municipality (see figure 2).

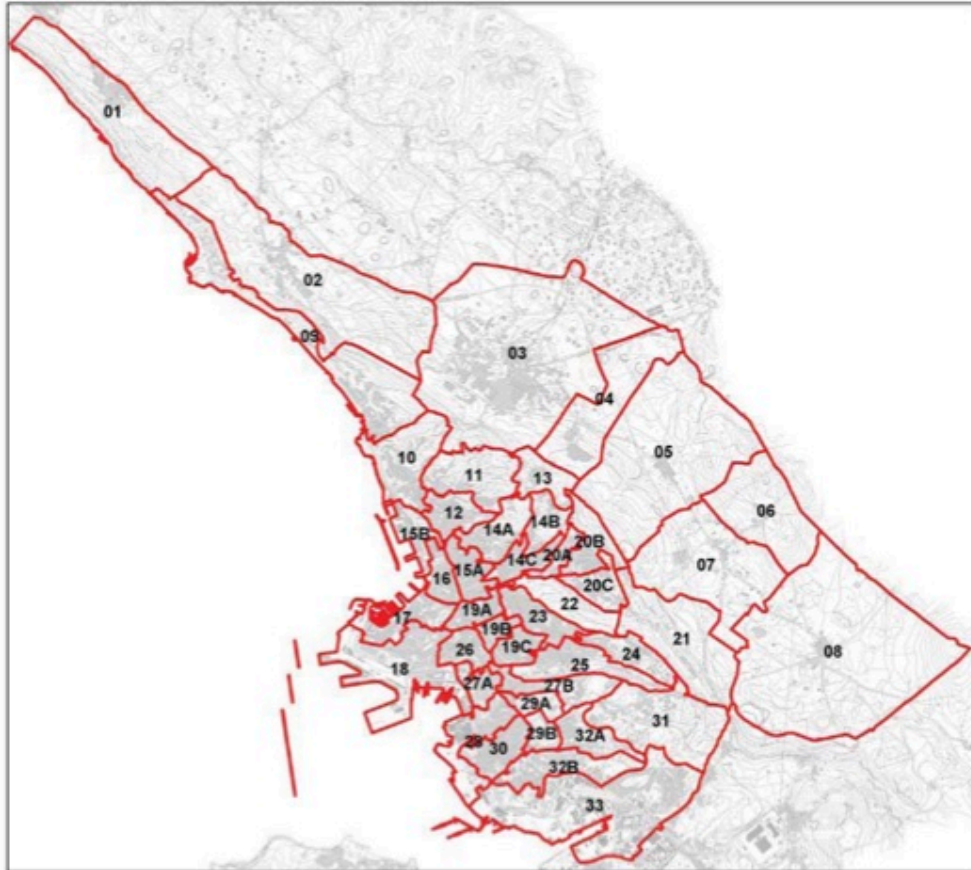


Fig. 1. Trieste divided by zone di quartiere.

The first part of my work will be to retrieve the geographical data of the *zone di quartiere* from some software like Google Maps and prepare a file to be transformed in a *pandas* dataset for subsequent analysis. I will also check if there is a useful correlation between geographical data and postcodes of the different areas.

When the geographical data will be ordered in a suitable way, I will be able to start exploring, segmenting and clustering the neighbourhoods of Trieste.

Borough number	Borough name	Neighbourhood name	Number
Circoscrizione 1	Altipiano Ovest	Santa Croce	1
		Prosecco	2
Circoscrizione 2	Altipiano Est	Opicina	3
		Banne	4
		Trebiciano	5
		Gropada	6
		Padriciano	7
		Basovizza	8
Circoscrizione 3	Roiano Gretta Barcola Cologna Scorcola	Barcola	9
		Gretta	10
		Roiano Nord	11
		Roiano Sud	12
		Conconello	13
		Cologna Nord	14b
		Cologna Sud	14c
		Scorcola	14a
Circoscrizione 4	Città Nuova Barriera Nuova S. Vito Città Vecchia	Barriera Nuova	15a
		Porto Vecchio	15b
		Borgo Teresiano	16
		Città Vecchia	17
		San Vito	18
Circoscrizione 5	Barriera Vecchia San Giacomo	Ospedale Maggiore	19a
		Piazza Perugino	19b
		Ippodromo	19c
Circoscrizione 6	S. Giovanni Chiadino Rozzol	Guardiella Inferiore	20a
		Guardiella Superiore	20b
		Sottolonga	20c
		Longera	21
		Boschetto	22
		Chiadino	23
		Melara	24
		Rozzol	25
		San Giacomo	26
Circoscrizione 7	Servola Chiarbola Valmaura Borgo S. Sergio	Chiarbola Superiore	27a
		Campanelle	27b
		Chiarbola inferiore	28
		Cimiteri non Cattolici	29a
		Cimiteri	29b
		Servola	30
		S.M. Maddalena Inferiore	31
		Poggi S. Anna	32a
		Valmaura	32b
		Borgo San Sergio	33