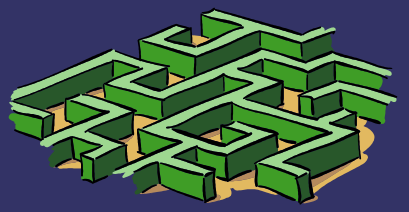
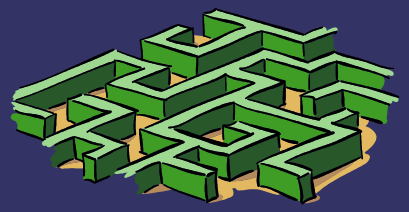
Where to place a new Hotel in central Rome (Italy)?

Andrea Bucci



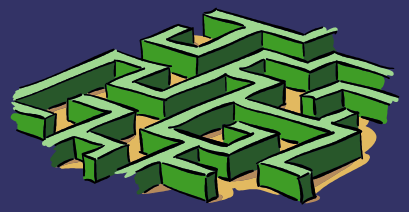
Business Problem

- An enterpreneur wants to know the best location where to start a business regarding tourist accomodation
- The chosen location is the central area of Rome, constituted by 22 Rioni (neighborhoods)



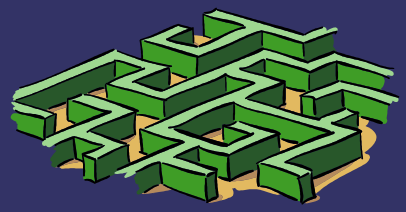
Tools Needed

- Map of the existing businesses at location
- Map of tourist attractions at location
- Ranking of top 10 most occurrent venues in each Rione



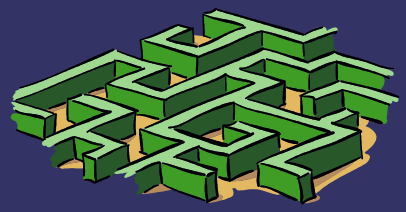
Tools Needed

- Table with demographic numbers to evaluate number of potential private competitors (rented out flats/rooms)
- Table with all the venues in each Rione, divided into categories



Data

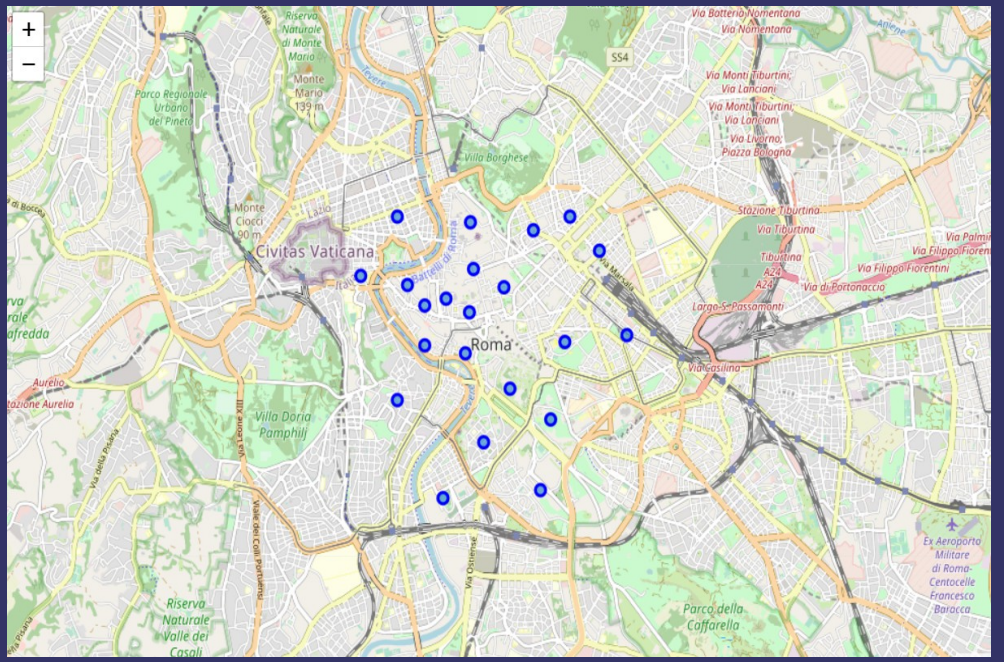
- KML file containing the geocoordinates for each Rione
- Wikipedia page containing info about the number of residents, area (expressed in squared kilometers) and population density (expressed in number of people per squared kilometer) in each Rione (table)
- Foursquare API



Clean Dataframes from Data

- Geocoordinates of Rioni

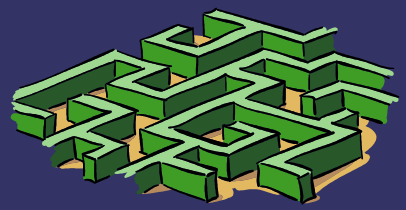
	Name	Latitude	Longitude
0	R.I Monti	41.893556	12.495489
1	R.II Trevi	41.899658	12.485104
2	R.III Colonna	41.901718	12.480018
3	R.IV Campo Marzio	41.907068	12.479482
4	R.V Ponte	41.899993	12.46871
5	R.VI Parione	41.897629	12.471778
6	R.VII Regola	41.893205	12.471714
7	R.VIII Sant'Eustacchio	41.898428	12.475297
8	R.IX Pigna	41.896879	12.47931
9	R.X Campitelli	41.888349	12.486134
10	R.XI Sant'Angelo	41.892311	12.47858



Clean Dataframes from Data

- Basic demographic info

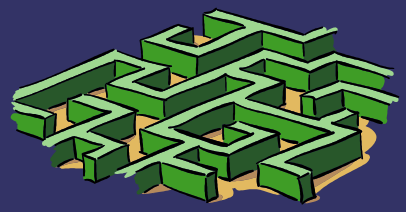
	Name	Latitude	Longitude	Population	Area(Km2)	DensityPerKm2
0	R.I Monti	41.893556	12.495489	13028	1.6508	7891.93
1	R.II Trevi	41.899658	12.485104	2327	0.5503	4228.6
2	R.III Colonna	41.901718	12.480018	2111	0.2689	7850.5
3	R.IV Campo Marzio	41.907068	12.479482	5860	0.8817	6646.25
4	R.V Ponte	41.899993	12.46871	3596	0.3189	11276.26
5	R.VI Parione	41.897629	12.471778	2572	0.1938	13271.41
6	R.VII Regola	41.893205	12.471714	3238	0.3189	10153.65
7	R.VIII Sant'Eustacchio	41.898428	12.475297	1962	0.1688	11623.22
8	R.IX Pigna	41.896879	12.47931	10737	0.2063	52045.56
9	R.X Campitelli	41.888349	12.486134	552	0.5990	921.54
10	R.XI Sant'Angelo	41.892311	12.47858	1084	0.1376	7877.91



Clean Dataframes from Data

- Foursquare API (In picture for one Rione)

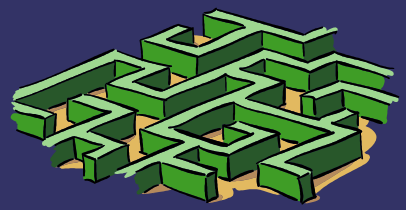
	Name	Category	Latitude	Longitude
0	Piazza del Campidoglio	Plaza	41.893321	12.482956
1	Terrazza delle Quadrighe	Scenic Lookout	41.894346	12.483336
2	Foro di Cesare	Historic Site	41.894128	12.485232
3	Foro di Traiano	Historic Site	41.894729	12.484871
4	Teatro Della Cometa	Theater	41.893399	12.481593



Final Dataframe from Data

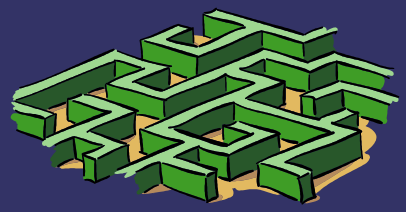
- Rioni locations + Foursquare Venues

	Rione	Rione Latitude	Rione Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	R.I Monti	41.893556	12.495489	Montipalace Hotel	41.895384	12.493839	Hotel
1	R.I Monti	41.893556	12.495489	The Sanctuary Roma	41.893279	12.498310	Lounge
2	R.I Monti	41.893556	12.495489	Mosè di Michelangelo	41.893681	12.492548	Public Art
3	R.I Monti	41.893556	12.495489	Fatamorgana	41.895610	12.493304	Ice Cream Shop
4	R.I Monti	41.893556	12.495489	Zia Rosetta	41.895562	12.493392	Sandwich Place
5	R.I Monti	41.893556	12.495489	Piazza San Martino ai Monti	41.895174	12.497192	Plaza
6	R.I Monti	41.893556	12.495489	Libreria Caffè Bohemien	41.895444	12.492863	Cocktail Bar
7	R.I Monti	41.893556	12.495489	Domus Aurea	41.894033	12.495643	Historic Site
8	R.I Monti	41.893556	12.495489	BrewDog Roma	41.891773	12.493113	Beer Bar



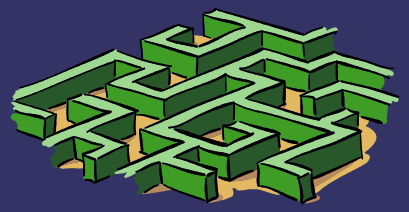
Methodology

- Clustering is used to group existing businesses and find empty spaces in between clusters
- K-Means is the algorithm of choice for its efficiency and simplicity



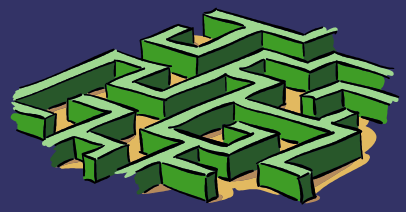
K-Means

- K-Means aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean (cluster centers or cluster centroid).
- Algorithm used is from ScikitLearn library



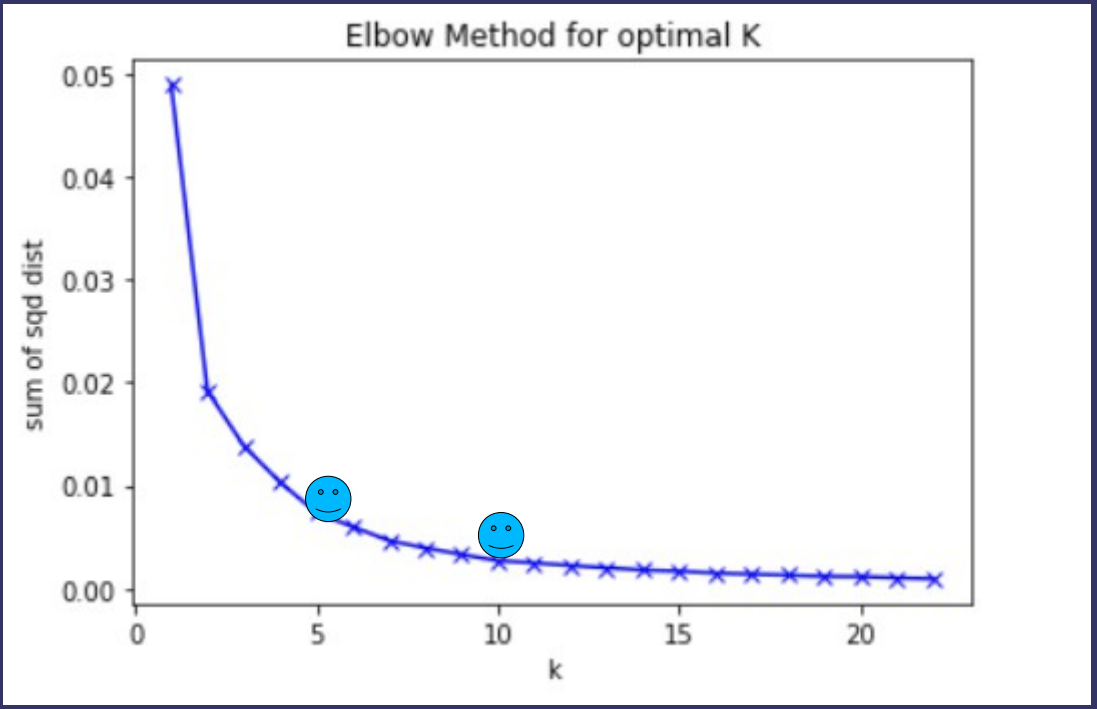
Optimal choice of K

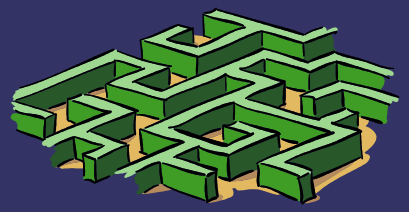
- Before fitting the data, an analysis regarding the optimal choice of k is just preferred in this scenario (but fundamental in others!)



Elbow Method

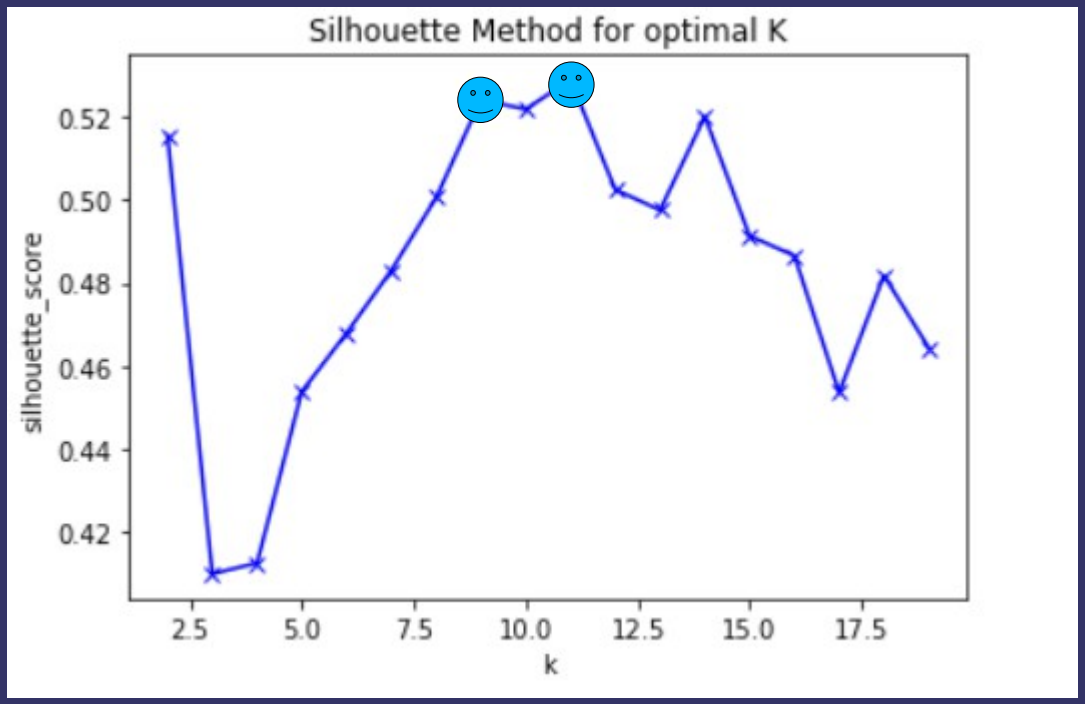
- Heuristic method consisting of plotting the explained variation as a function of the number of clusters, and picking the elbow of the curve as the number of clusters to use

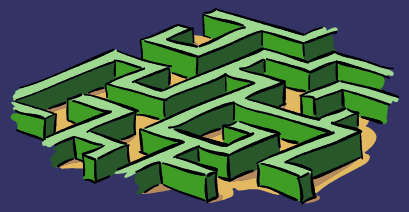




Silhouette Method

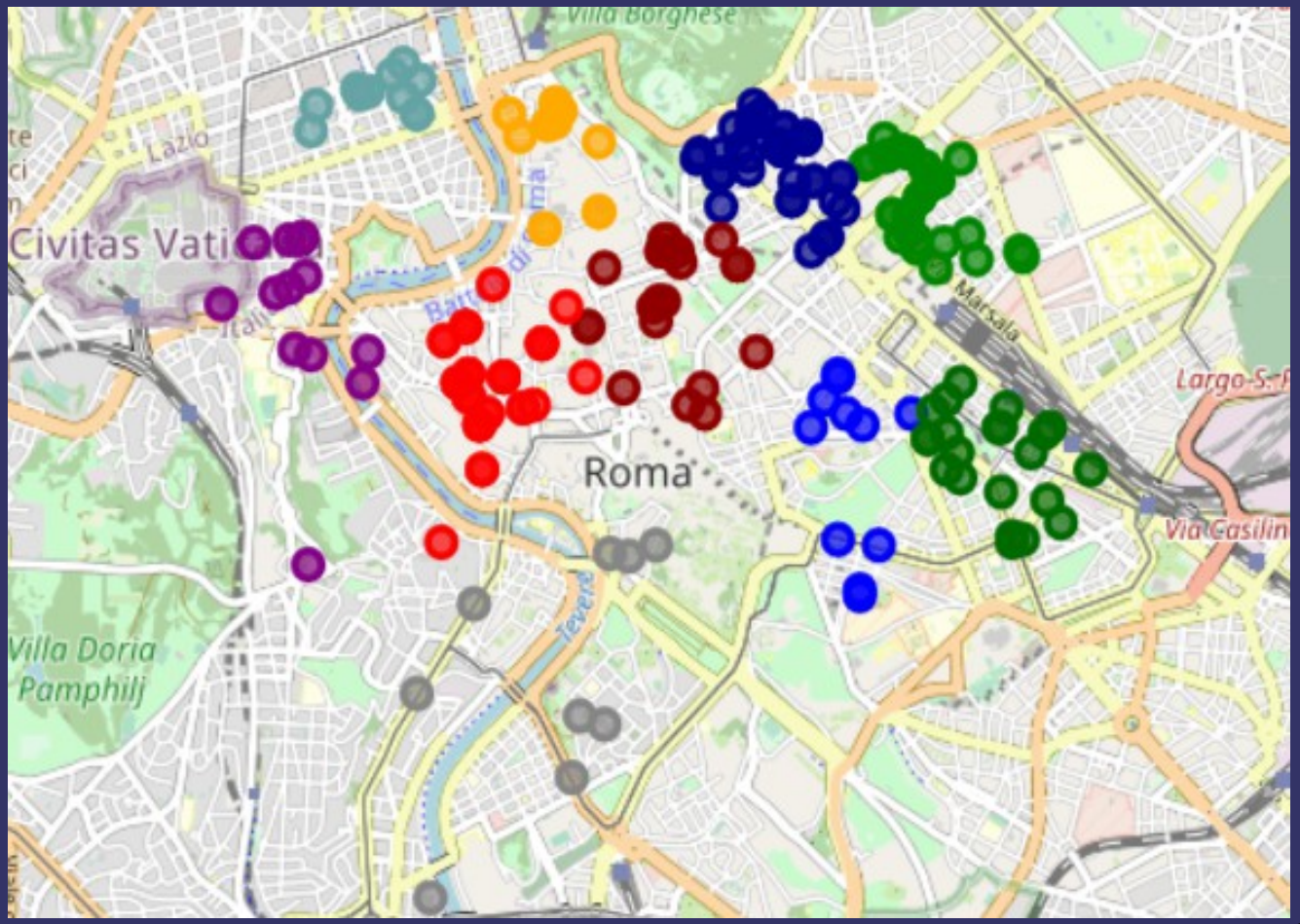
- Measure of how similar an object is to its own cluster (cohesion) compared to other clusters (separation)
- Range from −1 to +1 (high value: object is well matched to own cluster and poorly matched to neighboring clusters)
- Metric used to calculate distances is Euclidean





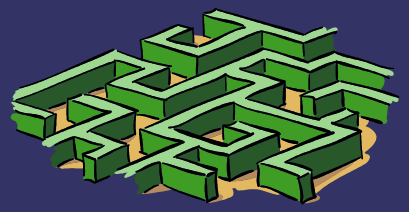
K-Means

- Map with 10 clusters of target businesses



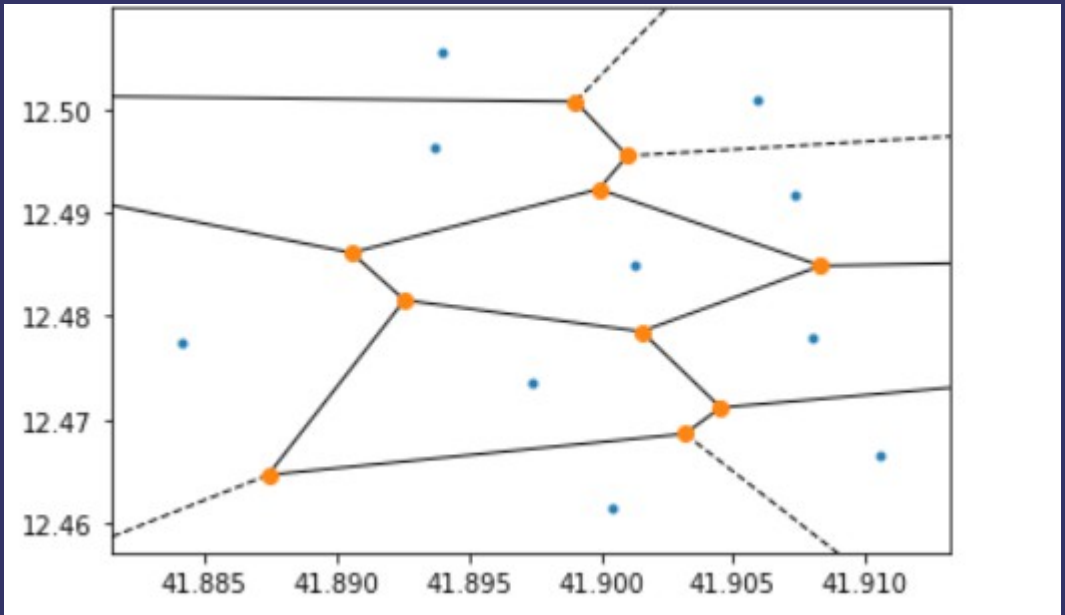
Voronoi

- A Voronoi diagram is a partition of a plane into regions close to each of a given set of objects (in our case, the clusters centroids).
- These regions are called Voronoi cells.

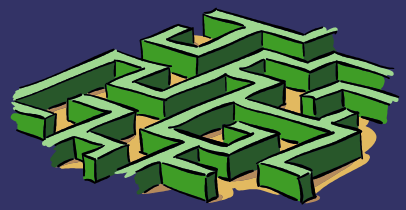


Voronoi

- Centroids represented as blue dots, sectors vertices as yellow dots and sectors edges as continous lines

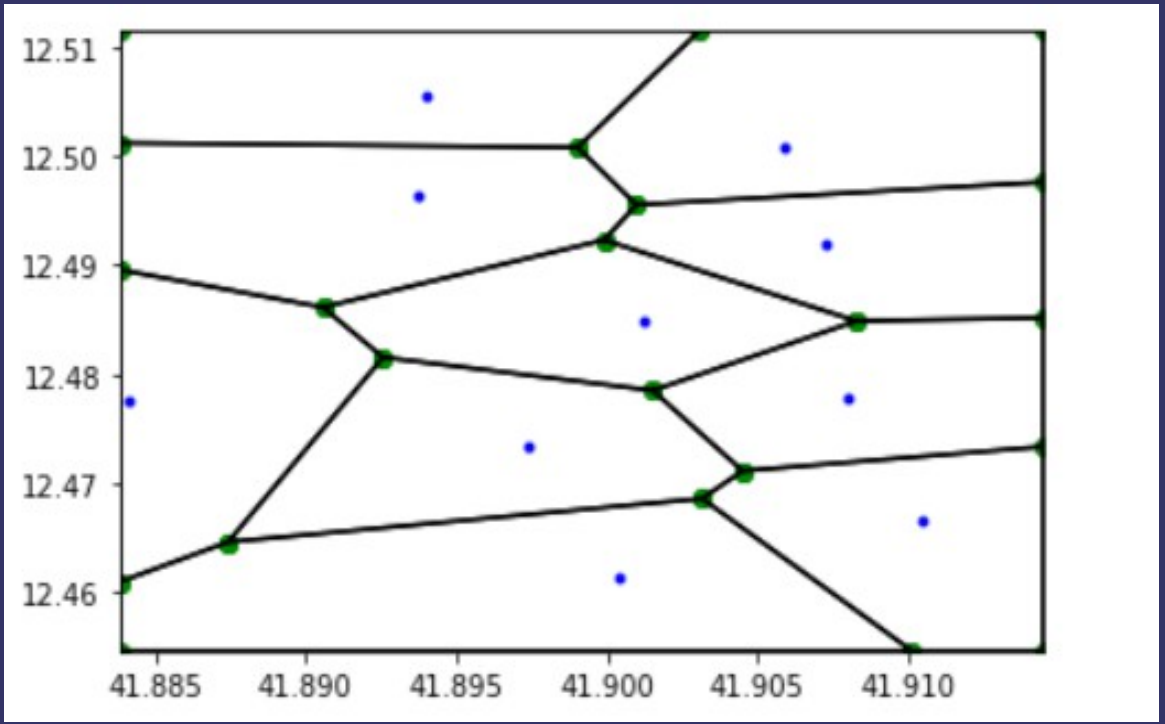


- Dashed lines indicate that some sectors are not closed (the dashed line is an infinite line ad not a side of the polygonal sector)

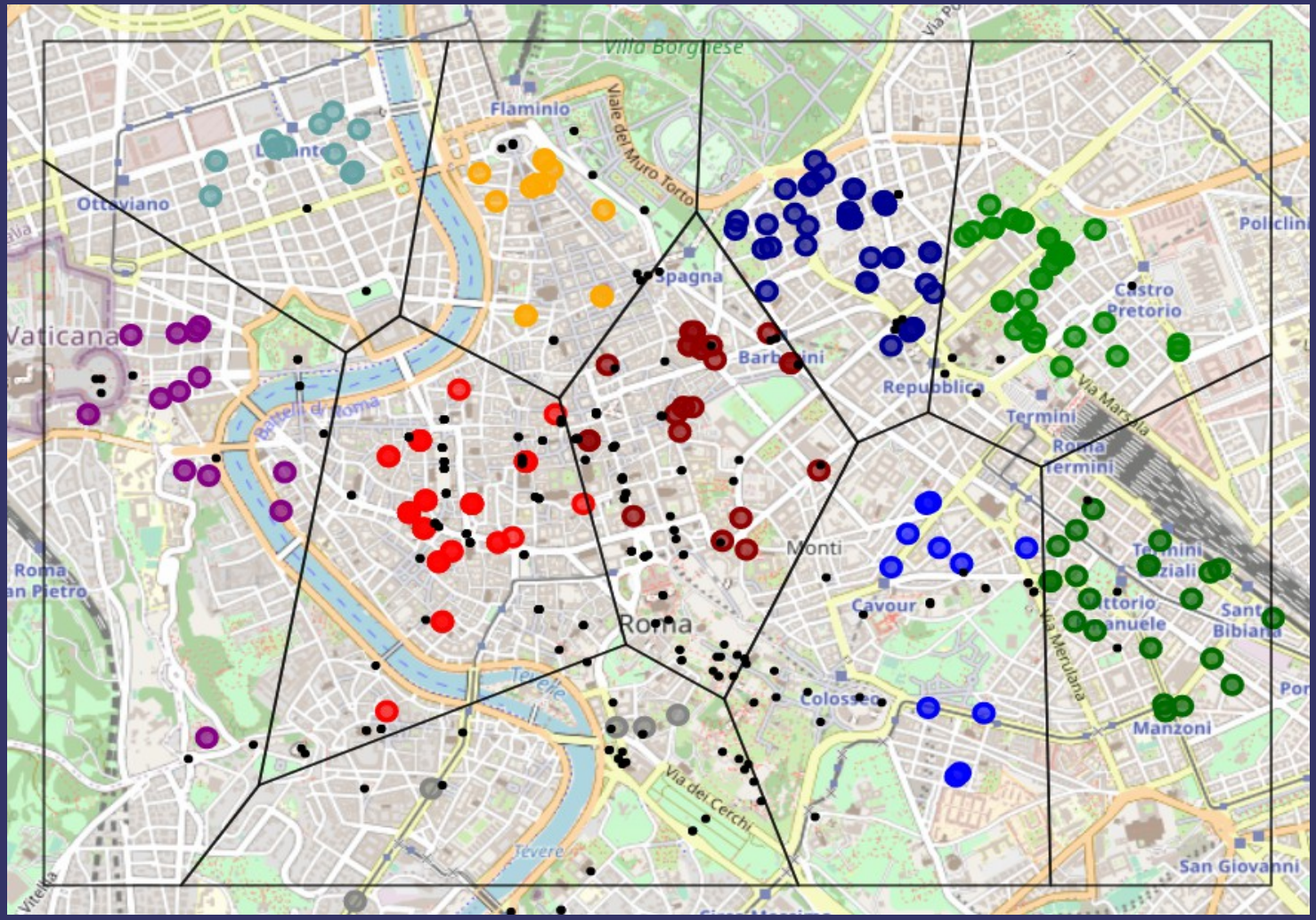


Voronoi

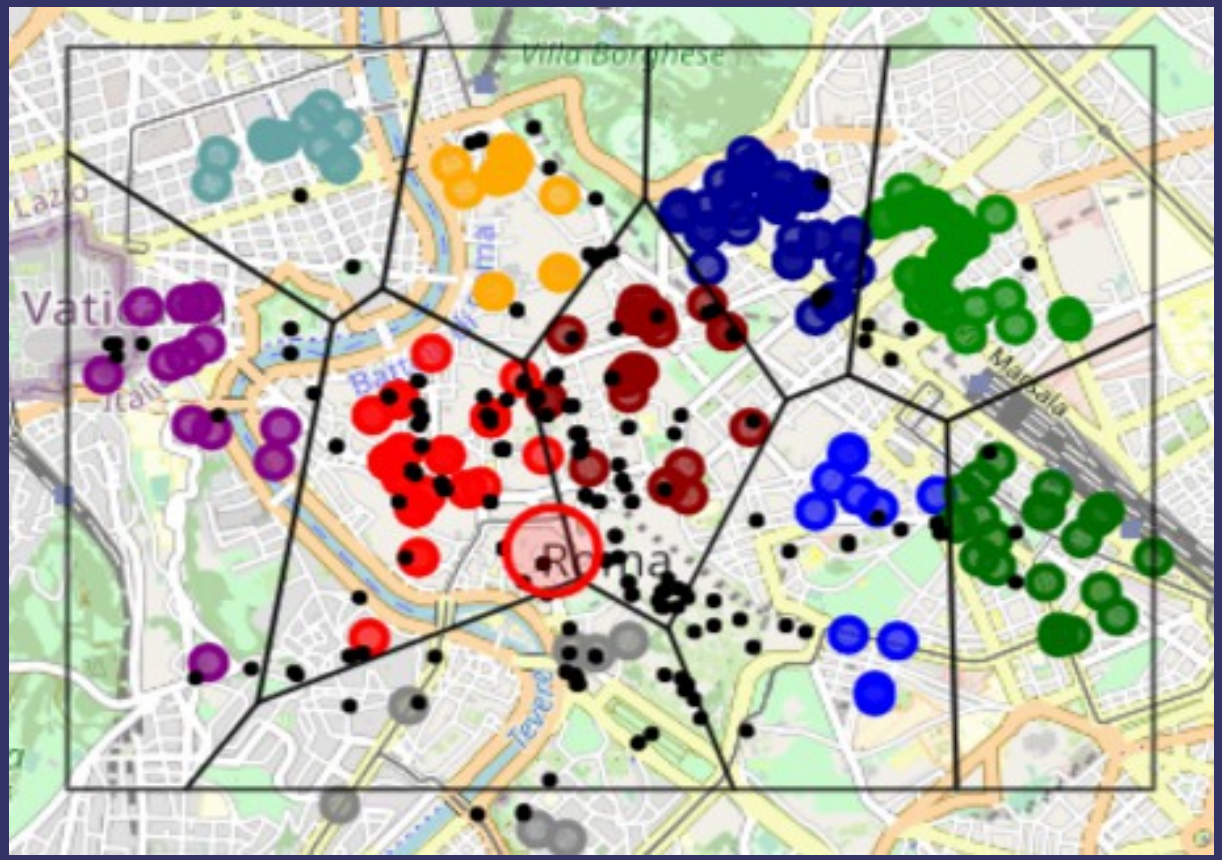
- A “bounding box” is set to close the sectors
- A function is coded that integrates the box

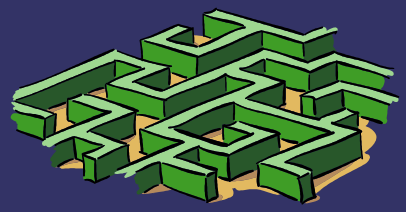


Final Map


Villa Borghese
Flaminio
Viale del Muro Torto
Ottaviano
Spagna
Vaticana
Barberini
Repubblica
Termini
Castro Pretorio
Via Marsala
Monti
Roma
Cavour
Via Merulana
Manzoni
Via dei Cerchi
Tevere
San Giovanni

Final Map + Optimal Zone





Results

- Rione Campitelli and Sant'Angelo remain promising because of:

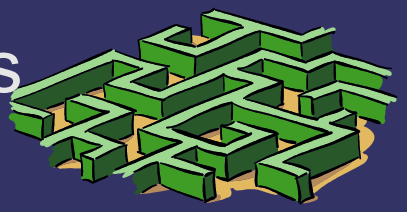
  -Proximity to most attractions

  -Relative scarcity of hotels or similar businesses nearby

  -Position more central wrt other empty areas

  -Proximity to the Train Central Station (close enough to allow fast shuttling)

  -Far enough from it so that attractions are easily reachable on foot

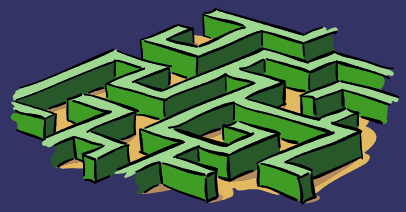


Further Analyses

- The entrepreneur wants a final glance at:

  -Number of specific venues inside the Rioni and around his optimal zone

  -Basic demographics of chosen area and Rioni to evaluate private rents competitors



Further Analyses

- 5 and 10 most common venues in each Rione (one-hot-encoded for eventual further Analysis in the future
- Italian Restaurants and Hotels in large numbers in the other Rioni
- Chosen Rioni have mostly historic sites and very low number of accomodations
- Investigation extended to top 10

Discussion

- Counts of total venues, points of interest, museums and roman restaurants per Rione

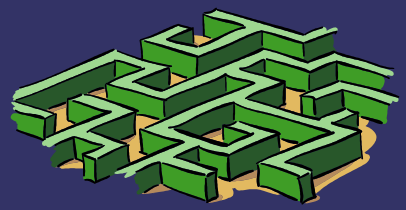
Rione	Venues
R.I Monti	82
R.II Trevi	91
R.III Colonna	92
R.IV Campo Marzio	72
R.IX Pigna	87
R.V Ponte	100
R.VI Parione	100
R.VII Regola	100
R.VIII Sant'Eustacchio	100
R.X Campitelli	56
R.XI Sant'Angelo	45
R.XII Ripa	50
R.XIII Trastevere	74
R.XIV Borgo	100
R.XIX Celio	25
R.XV Esquilino	93
R.XVI Ludovisi	86
R.XVII Sallustiano	100
R.XVIII Castro Pretorio	60
R.XX Testaccio	45
R.XXI San Saba	15
R.XXII Prati	39

Rione	POI
R.I Monti	10
R.II Trevi	27
R.III Colonna	15
R.IV Campo Marzio	11
R.IX Pigna	29
R.V Ponte	11
R.VI Parione	14
R.VII Regola	12
R.VIII Sant'Eustacchio	19
R.X Campitelli	34
R.XI Sant'Angelo	15
R.XII Ripa	7
R.XIII Trastevere	8
R.XIV Borgo	7
R.XV Esquilino	5
R.XVI Ludovisi	4
R.XVII Sallustiano	4
R.XVIII Castro Pretorio	5
R.XX Testaccio	4
R.XXI San Saba	4
R.XXII Prati	2

Rione	Museums
R.I Monti	1
R.II Trevi	9
R.IX Pigna	2
R.V Ponte	4
R.VI Parione	3
R.VII Regola	1
Sant'Eustacchio	2
R.X Campitelli	2
R.XI Sant'Angelo	1
R.XIII Trastevere	2
R.XIV Borgo	2
R.XV Esquilino	1
R.XVI Ludovisi	1
l Castro Pretorio	2
R.XX Testaccio	1

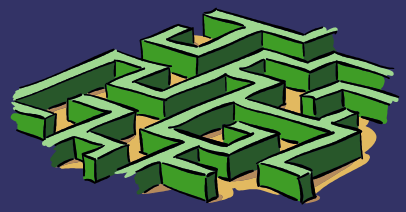
Rione	Roman_Restaurants
R.I Monti	82
R.II Trevi	91
R.III Colonna	92
R.IV Campo Marzio	72
R.IX Pigna	87
R.V Ponte	100
R.VI Parione	100
R.VII Regola	100
R.VIII Sant'Eustacchio	100

R.X Campitelli	56
R.XI Sant'Angelo	45
R.XII Ripa	50
R.XIII Trastevere	74
R.XIV Borgo	100
R.XIX Celio	25
R.XV Esquilino	93
R.XVI Ludovisi	86
R.XVII Sallustiano	100
R.XVIII Castro Pretorio	60
R.XX Testaccio	45
R.XXI San Saba	15
R.XXII Prati	39



Discussion

- This last evaluation strengthens the choice because of:
  - -The relative small presence of businesses wrt the other Rioni
  - -The relative small presence of restaurants other than italian ones
  - -Small-medium size number of residents
  - -Low density per sq. Km (low private competition)
- A more robust presence of traditional roman cuisine would have been better



Conclusions

- The chosen area remains almost the same of the postclustered with further indication that the Voronoi ridge crossing the right side of the optimal zone is both closer to the touristic attractions and roman restaurants in nearby neighborhoods.