

IBM Data Science Capstone Project

Venues of Thessaly, an overview



Manolis Efthimiou

12/04/2021



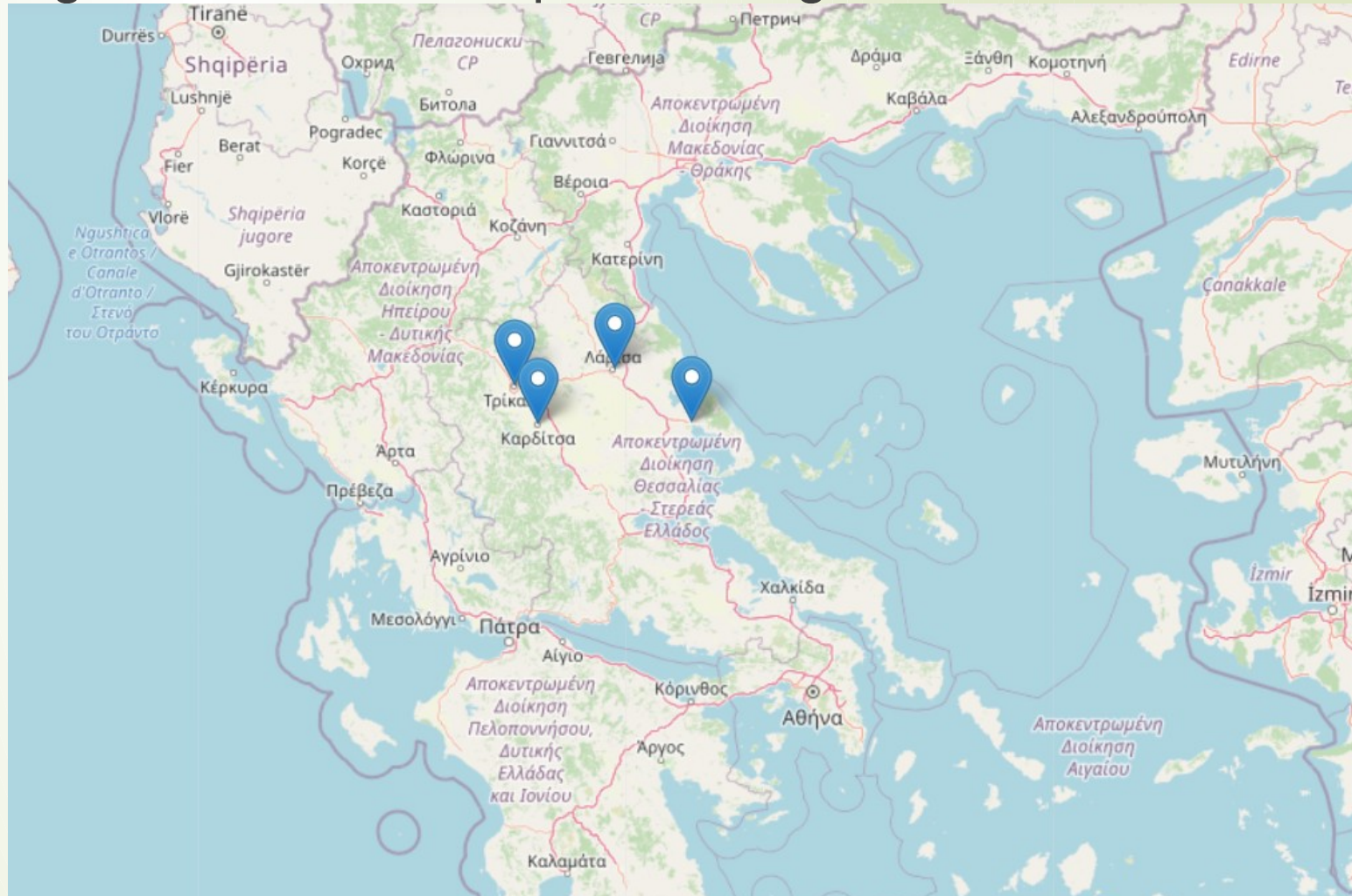
Business Problem

Thessaly is a geographic and modern administrative region in central Greece, a country known for Tourism. In this project we will make an overview of the venues of the 4 biggest cities in Thessaly.

Volos, Larisa, Karditsa and Trikala.

- Which type of venues are found in these cities, and how good are their ratings?
- How similar or dissimilar are these 4 cities that are geographically close, in regard to the venues they offer?

- Which type of venues are found in these cities, and how good are their ratings?
- A tool for seeing any gaps that can be filled by any businessman willing to make an attempt in the region





Data acquired

From the Foursquare API were about each venue's

- location, geographical coordinates about latitude and longitude
 - category of venue, like cafeteria, bar, etc
 - ratings that Foursquare users gave them
-
- From the Geopy API to obtain latitude and longitude of cities automatically, for ease of receptiveness for other locations



Methodology

- Data collection
 - 272 venues info acquired from Foursquare
- Data cleaning
 - 257 venues kept,
 - split to 7 venue types (food, cafes, alcohol, leisure, cultural, athletic, sweets)
- Data exploration
 - metrics in total, mean ratings per type
 - top rated venues per type
- Visualization
 - per city: all venues, top venues
- Clustering
 - depending on venue category & rating

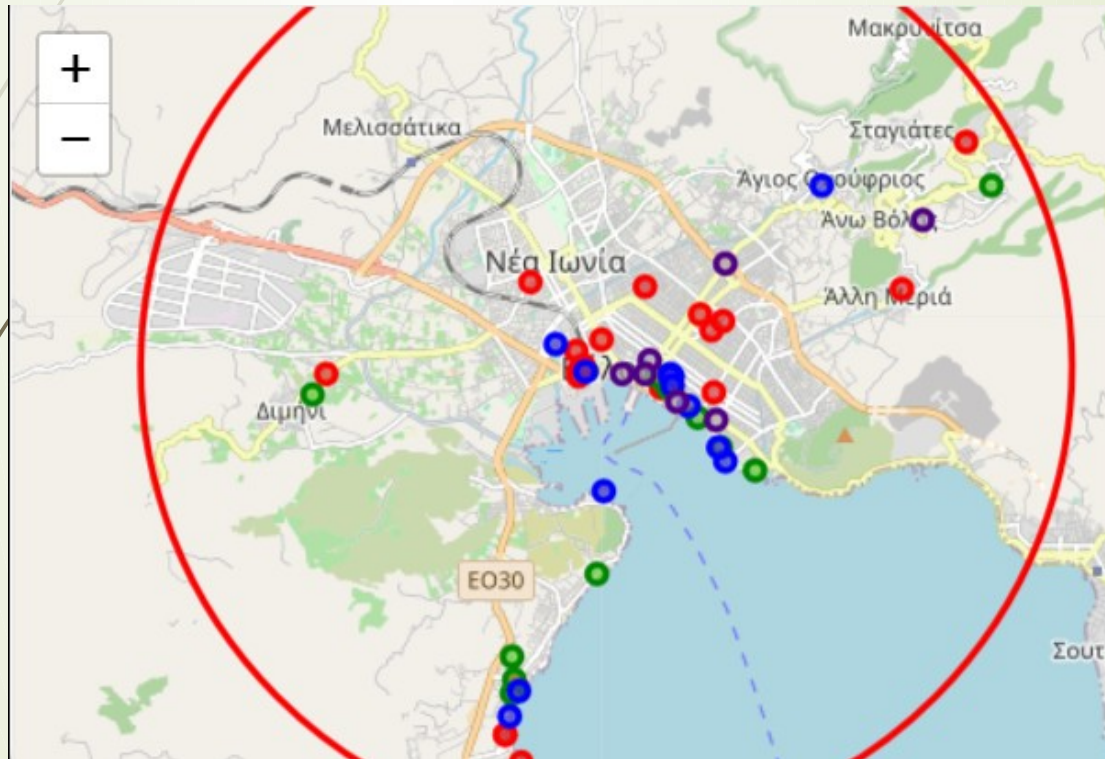
Metrics per venue type

Total venues - top rated venue - Mean rating

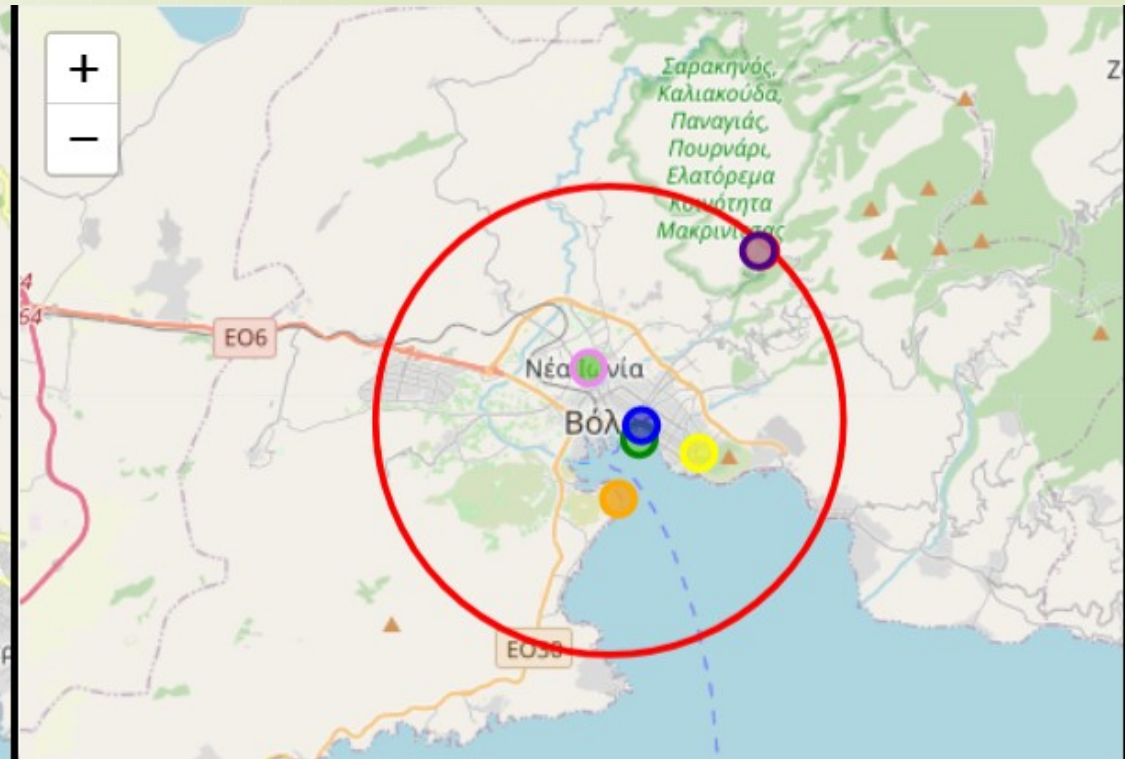
	karditsa	karditsa best	karditsa score	larisa	larisa best	larisa score	trikala	trikala best	trikala score	volos	volos best	volos score
food	15	8.1	7.48	20	8.8	7.82	22	8.3	7.67	23	9.2	8.42
caffes	11	8.8	7.81	18	8.9	7.88	15	8.3	7.58	8	9.0	8.36
alcohol	9	8.3	7.49	10	8.8	8.27	10	8.4	7.72	14	8.7	8.30
leisure	6	9.0	7.90	10	8.8	7.89	15	8.8	7.88	12	9.5	7.87
athletic	2	8.4	8.40	7	8.6	7.76	4	8.5	7.75	1	7.4	7.40
cultural	2	8.2	8.05	5	9.2	7.56	3	8.9	8.17	1	6.9	6.90
sweets	5	7.8	7.38	3	8.9	8.33	5	8.0	7.90	1	8.4	8.40

Map for every city

Info by popup when clicking on map



All venues

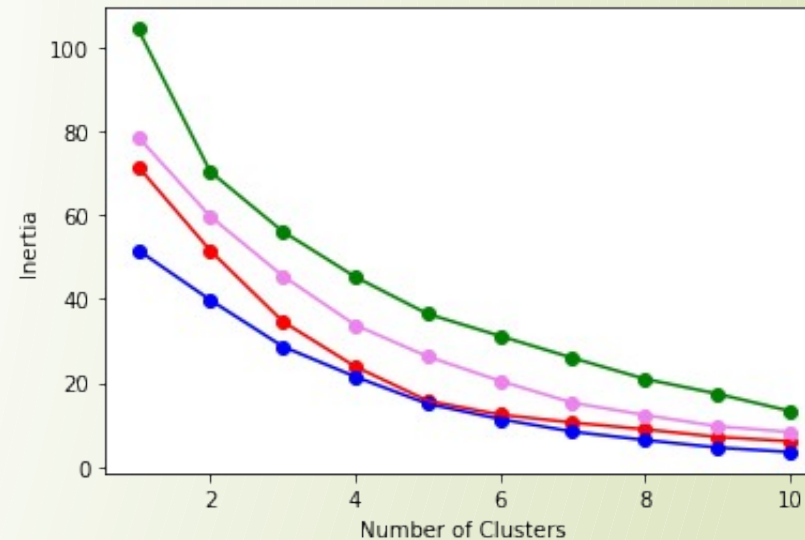


top rated venues

Clustering

Discover natural grouping for the venues of every city

- K-means algorithm
- 3 clusters



Elbow method

- Metrics per city per cluster

Metrics

		ratings	
		count	mean rating
city	cluster		
karditsa	0	11	7.809091
	1	23	7.660870
	2	14	7.478571
larisa	0	16	6.643750
	1	39	8.348718
	2	12	8.158333
trikala	0	33	7.842424
	1	18	7.672222
	2	13	7.576923
volos	0	27	8.470370
	1	8	6.837500
	2	21	8.476190

total venues, and mean rating found per city and per cluster

Metrics

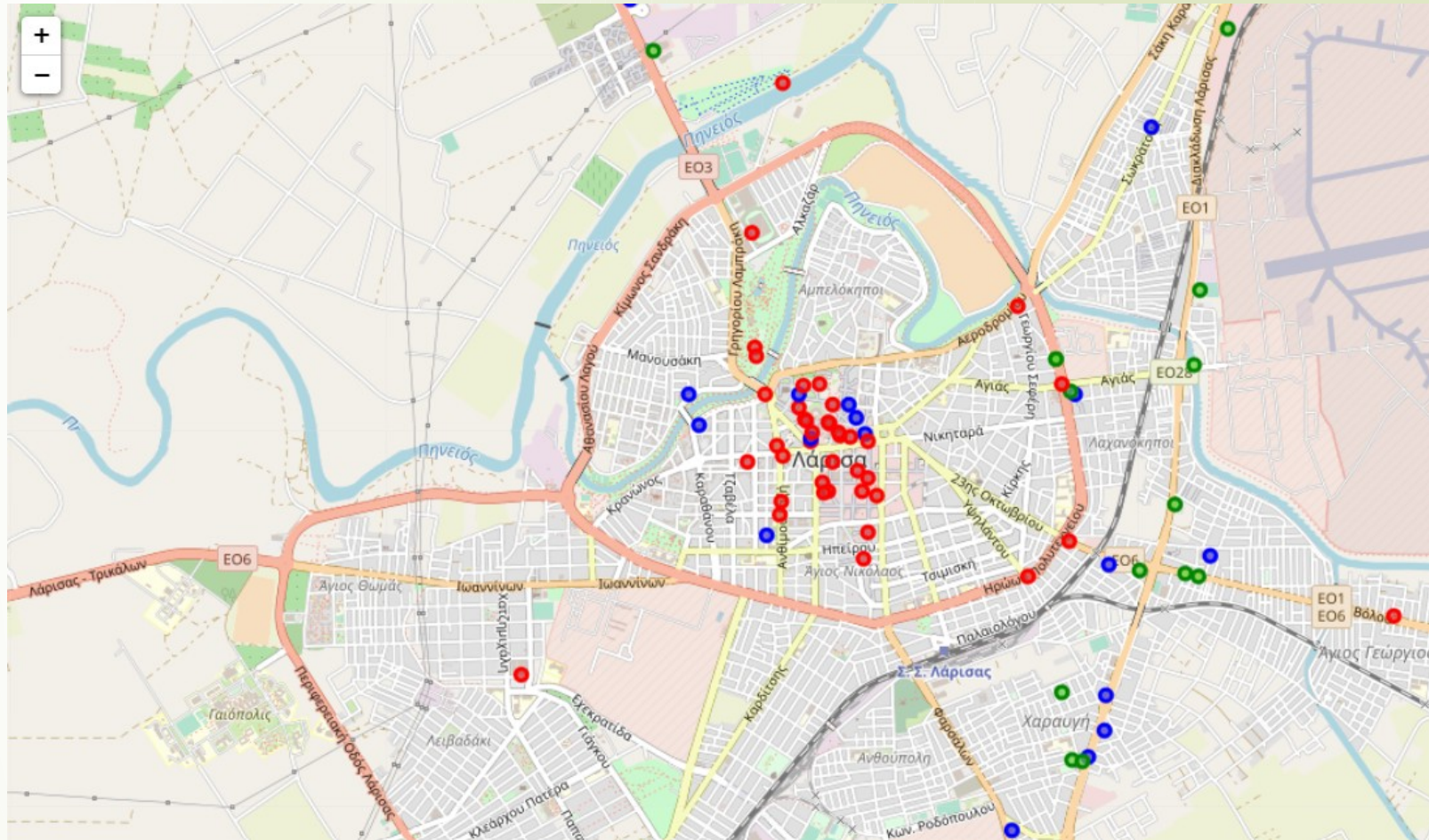
venues per category

				count
city	cluster	type		
karditsa	0	alcohol		9
		athletic		2
		cultural		2
		leisure		6
		sweets		5
	1	food		15
	2	caffe		11
larisa	0	alcohol		10
		athletic		6
		caffe		13
		cultural		2
		leisure		7
		sweets		3
	1	athletic		1
		caffe		5
		cultural		3
		food		3
		leisure		3

trikala	0	alcohol	10
		athletic	4
		cultural	3
		leisure	15
		sweets	5
volos	1	food	22
	2	caffe	15
	0	alcohol	13
		caffe	8
		leisure	8
		sweets	1
	1	food	22
	2	alcohol	1
		athletic	1
		cultural	1
		food	1
		leisure	4

total venues per category per city per cluster

Clustered venues map per city





Info by popup when clicking on map



Results and Discussion

- Most frequent venue categories are explained by the nature of Greece's economy which relies a lot on Tourism, local or foreign
- Venues of type 'leisure have higher ratings for what was done from the local government, like parks and plazas, in comparison to what is private economic activity, like nightclubs and resorts
- High rating, so interest too, for venues of historical interest
- Biggest similarity between the 4 cities is that the food venues were always a distinct cluster.
- The cafes had to be a separate type of venue due to the Greek population culture

- 
- 
- There is a correlation with a venues geographic proximity to the city's center or at least it's accessibility with main roads, and its rating
 - The highest ratings are for the city of Larisa, the biggest in population and known for it's night life.
 - The highest rated venues, were found in Volos, for the category of food (mean of 8.42). That is explained by the fact that region is well known for it's restaurants with tsipouro (a drink with alcohol).
 - For Volos, the highest ratings for the venue types of food and cafes, are inside the radius of 3-6 km, to the mountain Pilio