Settle down in Lyon, France

Segmentize districts using Python

The problem

I will move myself into Lyon, in order to pursue an MBA. The program will be on a campus located at Ecully district, which is placed outside of Lyon center.

So, in order to be around campus and have a full French experience, I need to segmentize districts based on their distance to campus and their trending places around.
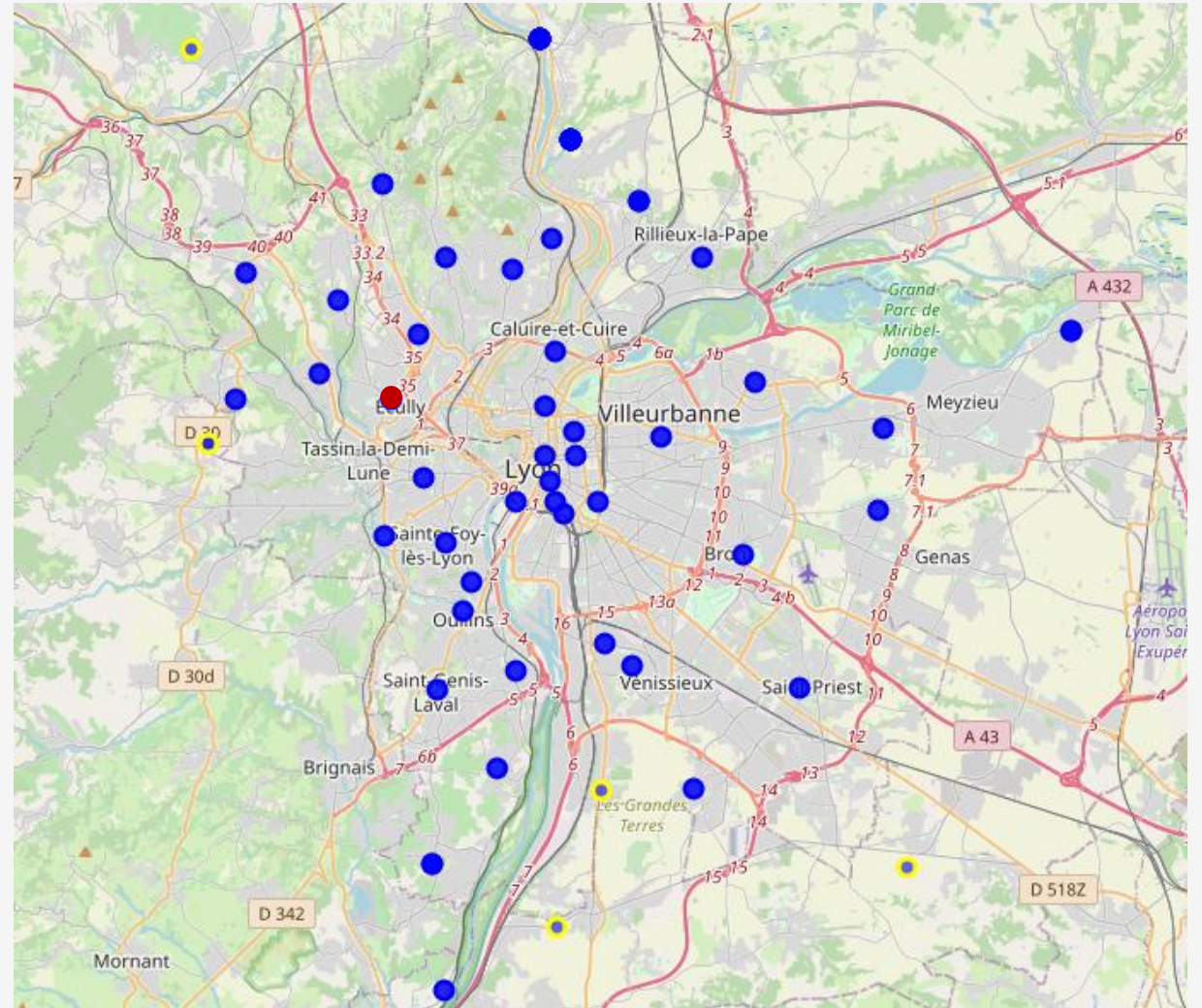
First overlook

Lyon has 66 districts

Ecully is located at the west of the center of the city, which it is supposed to have greater number of trending places.

The mission is to segregate these districts using kmeans in order to get a short list of district to look at to rent a place, that makes compatible studying and enjoying in Lyon

Methodology

1 Get the data from *Wikipedia* and then deploy it before using.

2 Use *pgeocode* to get the **coordinates** and **distance** to campus district for each district

3 Use *foursquare* query to rank each district with the **number of venues**.

4 Use *kmeans* to segmentize the district using **distance** and **number of venues**.

5 Map the results to analyze and conclude.

Data acquisition and cleaning

https://en.wikipedia.org/wiki/Communes_of_the_Lyon_Metropolis

Data deployment

Getting coordinates and distance using geocode

Old DataFrame

In [87]: df.head(10)

Out[87]:

	Postal code	Commune
0	69250	Albigny-sur-Saône
1	69500	Bron
2	69270	Cailloux-sur-Fontaines
3	69300	Caluire-et-Cuire
4	69410	Champagne-au-Mont-d'Or
5	69260	Charbonnières-les-Bains
6	69390	Charly
7	69680	Chassieu
8	69660	Collonges-au-Mont-d'Or
9	69960	Corbas

New DataFrame

In [92]: df.head(10)

Out[92]:

	Postal code	Commune	latitude	longitude	Distance
0	69250	Albigny-sur-Saône	45.8782	4.834533	12.136730
1	69500	Bron	45.7333	4.916700	12.047306
2	69270	Cailloux-sur-Fontaines	45.8500	4.846833	9.824269
3	69300	Caluire-et-Cuire	45.7904	4.841100	5.343529
4	69410	Champagne-au-Mont-d'Or	45.7955	4.785700	2.193151
5	69260	Charbonnières-les-Bains	45.7842	4.746300	2.339509
6	69390	Charly	45.6469	4.791625	14.558567
7	69680	Chassieu	45.7458	4.970900	15.610449
8	69660	Collonges-au-Mont-d'Or	45.8220	4.839800	7.077836
9	69960	Corbas	45.6679	4.896300	15.392917

Variable for segmentizing

Use Foursquare to get the number of Venues of each district

	Postal code	Commune	latitude	longitude	Distance	N_Venues
0	69250	Albigny-sur-Saône	45.8782	4.834533	12.136730	5.0
1	69500	Bron	45.7333	4.916700	12.047306	7.0
2	69270	Cailloux-sur-Fontaines	45.8500	4.846833	9.824269	1.0
3	69300	Caluire-et-Cuire	45.7904	4.841100	5.343529	4.0
4	69410	Champagne-au-Mont-d'Or	45.7955	4.785700	2.193151	2.0
5	69260	Charbonnières-les-Bains	45.7842	4.746300	2.339509	4.0
6	69390	Charly	45.6469	4.791625	14.558567	1.0
7	69680	Chassieu	45.7458	4.970900	15.610449	3.0
8	69660	Collonges-au-Mont-d'Or	45.8220	4.839800	7.077836	1.0
9	69960	Corbas	45.6679	4.896300	15.392917	4.0
10	69270	Couzon-au-Mont-d'Or	45.8500	4.846833	9.824269	1.0
11	69290	Craponne	45.7647	4.701675	5.842184	NaN
12	69250	Curis-au-Mont-d'Or	45.8782	4.834533	12.136730	5.0
13	69570	Dardilly	45.8051	4.753400	3.508371	4.0
14	69150	Décines-Charpieu	45.7690	4.972800	15.384323	4.0
15	69130	Écully	45.7773	4.774800	0.000000	5.0
16	69320	Feyzin	45.6673	4.859700	13.894110	NaN
17	69250	Fleurieu-sur-Saône	45.8782	4.834533	12.136730	5.0
18	69270	Fontaines-Saint-Martin	45.8500	4.846833	9.824269	1.0
19	69270	Fontaines-sur-Saône	45.8500	4.846833	9.824269	1.0
20	69340	Francheville	45.7388	4.772200	4.285760	5.0
21	69730	Genay	45.8976	4.837300	14.226074	3.0
22	69700	Givors	45.5519	4.730812	25.295394	NaN
23	69520	Grigny	45.6113	4.795900	18.530990	2.0
24	69540	Irigny	45.6736	4.817500	11.997868	1.0
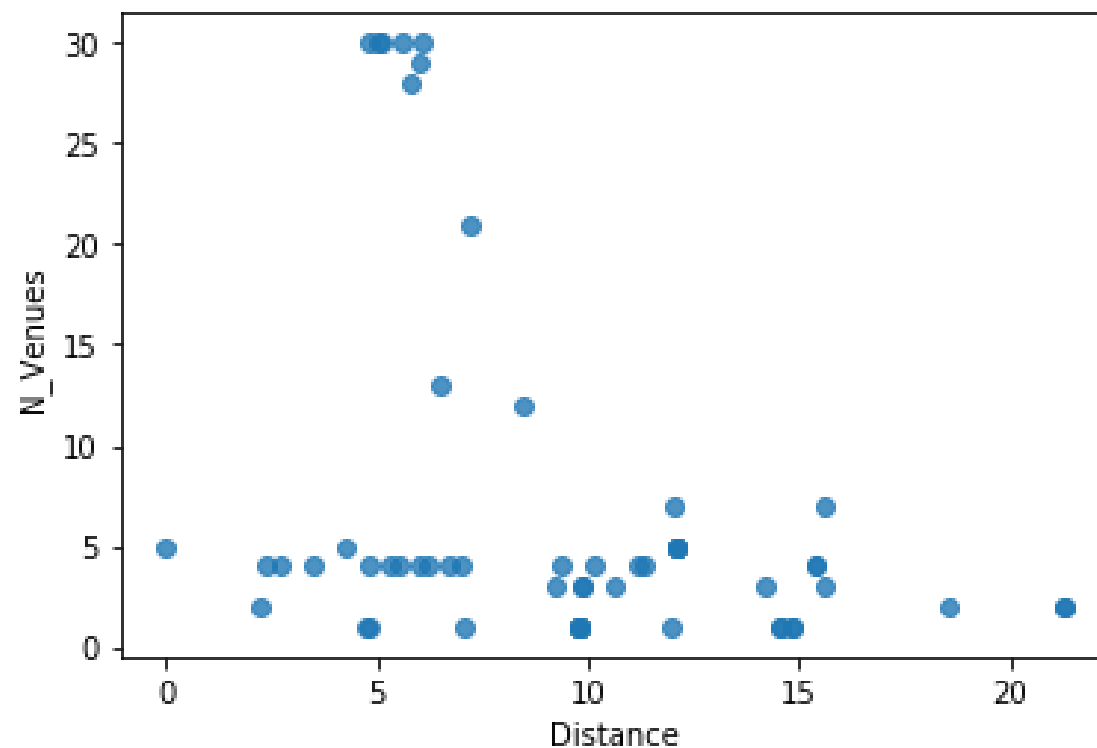
Exploring at the data, it will be prioritized the districts which have at least 3 venues and a distance lesser than 20 kilometers.