Analysis of COVID-19 in Madrid

Contents

- *Background*
- *Stakeholders*
- *Datasets*
- *Classifying districts: procedure*
- *Classifying districts: visualization*
- *Analysing data from FourSquare*
- *Applying machine learning*
- *Results*
- *Discussion*
- *Conclusion*

Background

- *COVID-19* pandemic has marked a *milestone* in our lives
- Spain, and especially its capital, *Madrid*, was one of the places where the situation was *critical* during the first months of the pandemic
- Factors such as *ageing population*, *precarious jobs*, high *population density*, *travellers* continously coming and going and *overcrowded* places help the virus to spread more fastly
- The *study and the analysis of these factors* may *predict, slow down* or even *avoid* future outbreaks

3

Stakeholders

- *Experience* and *data* adquired during the first wave of the pandemic can be used to *avoid another similar situation* in the future

- *Districts of Madrid* can be *labelled* and *clustered* using their *number of infections* and their *most popular venues*

- Labelling and clustering districts may be *helpful for governments*: *measures* can be taken as soon the number of infections starts to grow in any of *these districts* or in *districts with similar characteristics*

Datasets *(1/4)*

1. ### *COVID-19 data by district*
   - Provided by the ***Community of Madrid***
   - Contains ***COVID-19 cumulative incidence*** on specific dates by district
   - After ***cleaning*** this dataset (***translate*** headers, ***remove data*** previous to July, 1st, ***remove towns*** different then Madrid city, ***change data types*** and ***remove columns*** which will not be used), the resulting dataframe looks like this:

	town_district	total_positive_tests	total_cumulative_rate	district_code
0	Madrid-Retiro	1691.0	1417.23	79603
1	Madrid-Salamanca	1781.0	1218.99	79604
2	Madrid-Centro	1282.0	950.70	79601
3	Madrid-Arganzuela	1769.0	1150.17	79602
4	Madrid-Chamartín	1800.0	1234.15	79605
5	Madrid-Tetuán	2233.0	1413.09	79606
6	Madrid-Chamberí	1739.0	1247.68	79607
7	Madrid-Fuencarral-El Pardo	3290.0	1337.73	79608
8	Madrid-Moncloa-Aravaca	1568.0	1313.02	79609
9	Madrid-Latina	3080.0	1292.93	79610
10	Madrid-Carabanchel	2957.0	1168.32	79611
11	Madrid-Usera	1505.0	1076.99	79612
12	Madrid-Puente de Vallecas	3348.0	1425.55	79613
13	Madrid-San Blas - Canillejas	1584.0	1001.63	79620
14	Madrid-Barajas	464.0	948.00	79621
15	Madrid-Moratalaz	1517.0	1604.10	79614
16	Madrid-Ciudad Lineal	2431.0	1123.99	79615
17	Madrid-Hortaleza	1885.0	1001.66	79616
18	Madrid-Villaverde	1588.0	1066.16	79617
19	Madrid-Villa de Vallecas	1318.0	1194.22	79618
20	Madrid-Vicálvaro	929.0	1288.65	79619

Datasets *(2/4)*

2. ***Districts' coordinates***
   - "*Handmade*" searching coordinates in Google
   - Contains the ***coordinates of each district***

	district_code	latitude	longitude
0	79604	40.430000	-3.677778
1	79601	40.415278	-3.707500
2	79602	40.402778	-3.695278
3	79605	40.453333	-3.677500
4	79606	40.460556	-3.700000
5	79607	40.432778	-3.697222
6	79608	40.478611	-3.709722
7	79609	40.435470	-3.731700
8	79610	40.402500	-3.741389
9	79611	40.383611	-3.728056
10	79612	40.381389	-3.706944
11	79613	40.393540	-3.662000
12	79620	40.426111	-3.612778
13	79621	40.470278	-3.585000
14	79614	40.410000	-3.644444
15	79615	40.445050	-3.651320
16	79616	40.469444	-3.640556
17	79617	40.345833	-3.709444
18	79618	40.379722	-3.621389
19	79619	40.404167	-3.608056
20	79603	40.408056	-3.676667

Datasets *(3/4)*

3. *Merging datasets*

- Using the column "*district_code*" the ***previous dataframes will be merged*** in one.
- The ***resulting dataframe*** is shown below

	town_district	total_positive_tests	total_cumulative_rate	district_code	latitude	longitude
0	Madrid-Retiro	1691.0	1417.23	79603	40.408056	-3.676667
1	Madrid-Salamanca	1781.0	1218.99	79604	40.430000	-3.677778
2	Madrid-Centro	1282.0	950.70	79601	40.415278	-3.707500
3	Madrid-Arganzuela	1769.0	1150.17	79602	40.402778	-3.695278
4	Madrid-Chamartín	1800.0	1234.15	79605	40.453333	-3.677500
5	Madrid-Tetuán	2233.0	1413.09	79606	40.460556	-3.700000
6	Madrid-Chamberí	1739.0	1247.68	79607	40.432778	-3.697222
7	Madrid-Fuencarral-El Pardo	3290.0	1337.73	79608	40.478611	-3.709722
8	Madrid-Moncloa-Aravaca	1568.0	1313.02	79609	40.435470	-3.731700
9	Madrid-Latina	3080.0	1292.93	79610	40.402500	-3.741389
10	Madrid-Carabanchel	2957.0	1168.32	79611	40.383611	-3.728056
11	Madrid-Usera	1505.0	1076.99	79612	40.381389	-3.706944
12	Madrid-Puente de Vallecas	3348.0	1425.55	79613	40.393540	-3.662000
13	Madrid-San Blas - Canillejas	1584.0	1001.63	79620	40.426111	-3.612778
14	Madrid-Barajas	464.0	948.00	79621	40.470278	-3.585000
15	Madrid-Moratalaz	1517.0	1604.10	79614	40.410000	-3.644444
16	Madrid-Ciudad Lineal	2431.0	1123.99	79615	40.445050	-3.651320
17	Madrid-Hortaleza	1885.0	1001.66	79616	40.469444	-3.640556
18	Madrid-Villaverde	1588.0	1066.16	79617	40.345833	-3.709444
19	Madrid-Villa de Vallecas	1318.0	1194.22	79618	40.379722	-3.621389
20	Madrid-Vicálvaro	929.0	1288.65	79619	40.404167	-3.608056

4. ## *Most popular venues*

- Using the coordinates of each district, **most popular venues** in each one will be got through **calls to** **FourSquare API**.
- We started **getting the nearby venues** in a radius of 500 meters for each district
- After **executing some operations**, we will have a **dataframe with the 10 most common venue category** for each district:

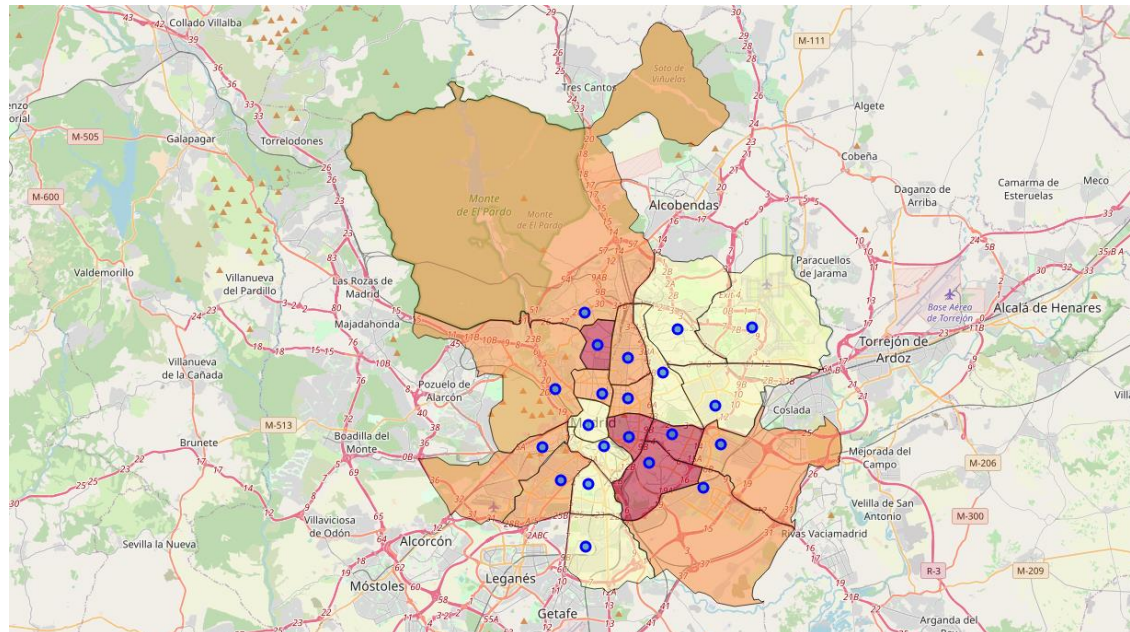|	District	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Madrid-Arganzuela	Spanish Restaurant	Restaurant	Grocery Store	Tapas Restaurant	Bakery	Gym / Fitness Center	Burger Joint	Hotel	Sandwich Place	Falafel Restaurant
1	Madrid-Barajas	Hotel	Restaurant	Spanish Restaurant	Supermarket	Tapas Restaurant	Coffee Shop	Brewery	Fast Food Restaurant	Bistro	Japanese Restaurant
2	Madrid-Carabanchel	Metro Station	Soccer Field	Burger Joint	Fast Food Restaurant	Tapas Restaurant	Nightclub	Bakery	Plaza	Pizza Place	Pharmacy
3	Madrid-Centro	Tapas Restaurant	Spanish Restaurant	Plaza	Hostel	Bar	Ice Cream Shop	Restaurant	Cocktail Bar	Coffee Shop	Pastry Shop
4	Madrid-Chamartín	Spanish Restaurant	Restaurant	Grocery Store	Park	Tapas Restaurant	Bakery	Café	Japanese Restaurant	Gastropub	Coffee Shop
5	Madrid-Chamberí	Spanish Restaurant	Bar	Restaurant	Café	Tapas Restaurant	Brewery	Italian Restaurant	Plaza	Bakery	Coffee Shop
6	Madrid-Ciudad Lineal	Spanish Restaurant	Grocery Store	Hotel	Pizza Place	Italian Restaurant	Tapas Restaurant	Bar	Motorcycle Shop	Fast Food Restaurant	Park
7	Madrid-Fuencarral-El Pardo	Clothing Store	Italian Restaurant	Fast Food Restaurant	Tapas Restaurant	Burger Joint	Sandwich Place	Restaurant	Chocolate Shop	Pizza Place	Pharmacy
8	Madrid-Hortaleza	Pizza Place	Breakfast Spot	Supermarket	Spanish Restaurant	Dance Studio	Pharmacy	Plaza	Clothing Store	Donut Shop	Chinese Restaurant
9	Madrid-Latina	Pizza Place	Falafel Restaurant	Metro Station	Train Station	Food & Drink Shop	Park	Grocery Store	Asian Restaurant	Fast Food Restaurant	Bakery
	scroll output; double click to hide	Spanish Restaurant	Gym	Pool	Pub	Stadium	Wine Shop	Falafel Restaurant	Food	Flea Market	Fish Market
11	Madrid-Moratalaz	Park	Bar	Bakery	Food Truck	Café	Nightclub	Soccer Field	Ice Cream Shop	Pizza Place	Plaza
12	Madrid-Puente de Vallecas	Gym	Pub	Pizza Place	Music Venue	Sandwich Place	Spanish Restaurant	Concert Hall	Performing Arts Venue	Camera Store	Café
13	Madrid-Retiro	Spanish Restaurant	Supermarket	Tapas Restaurant	Bar	Grocery Store	Mediterranean Restaurant	Diner	Brewery	Burger Joint	Food & Drink Shop
14	Madrid-Salamanca	Spanish Restaurant	Restaurant	Seafood Restaurant	Mediterranean Restaurant	Coffee Shop	Burger Joint	Supermarket	Tapas Restaurant	Mexican Restaurant	Bakery
15	Madrid-San Blas - Canillejas	Metro Station	Asian Restaurant	Shopping Mall	Snack Place	Supermarket	Gas Station	Pizza Place	Gym	Grocery Store	Dance Studio
16	Madrid-Tetuán	Spanish Restaurant	Bakery	Coffee Shop	Supermarket	Chinese Restaurant	Brazilian Restaurant	Grocery Store	Farmers Market	Pizza Place	Gym / Fitness Center
17	Madrid-Usera	Seafood Restaurant	Bubble Tea Shop	Spanish Restaurant	Asian Restaurant	Theater	Noodle House	Fast Food Restaurant	Chinese Restaurant	Wine Shop	Farmers Market
18	Madrid-Vicálvaro	Pizza Place	Spanish Restaurant	Beer Bar	Fast Food Restaurant	Sandwich Place	Café	Breakfast Spot	Metro Station	Grocery Store	Ice Cream Shop
19	Madrid-Villa de Vallecas	Metro Station	Grocery Store	Spanish Restaurant	Food	Soccer Field	Train Station	Department Store	Dessert Shop	Fast Food Restaurant	Food Court
20	Madrid-Villaverde	Pizza Place	Spanish Restaurant	Train	Diner	Falafel Restaurant	Food & Drink Shop	Food	Flea Market	Fish Market	Fast Food Restaurant

Classifying districts: procedure

- Districts will be *labelled* using their *COVID-19 total cumulative rate*
- Three categories: "*Low*", "*Medium*" and "*High*"
- Using the range generated by the function "*linspace*" from "*numpy*", the *districts will be labelled* as follows:

	town_district	total_positive_tests	total_cumulative_rate	district_code	latitude	longitude	total_cumulative_rate_label
0	Madrid-Retiro	1691.0	1417.23	79603	40.408056	-3.676667	2
1	Madrid-Salamanca	1781.0	1218.99	79604	40.430000	-3.677778	1
2	Madrid-Centro	1282.0	950.70	79601	40.415278	-3.707500	0
3	Madrid-Arganzuela	1769.0	1150.17	79602	40.402778	-3.695278	0
4	Madrid-Chamartín	1800.0	1234.15	79605	40.453333	-3.677500	1
5	Madrid-Tetuán	2233.0	1413.09	79606	40.460556	-3.700000	2
6	Madrid-Chamberí	1739.0	1247.68	79607	40.432778	-3.697222	1
7	Madrid-Fuencarral-El Pardo	3290.0	1337.73	79608	40.478611	-3.709722	1
8	Madrid-Moncloa-Aravaca	1568.0	1313.02	79609	40.435470	-3.731700	1
9	Madrid-Latina	3080.0	1292.93	79610	40.402500	-3.741389	1
10	Madrid-Carabanchel	2957.0	1168.32	79611	40.383611	-3.728056	1
11	Madrid-Usera	1505.0	1076.99	79612	40.381389	-3.706944	0
12	Madrid-Puente de Vallecas	3348.0	1425.55	79613	40.393540	-3.662000	2
13	Madrid-San Blas - Canillejas	1584.0	1001.63	79620	40.426111	-3.612778	0
14	Madrid-Barajas	464.0	948.00	79621	40.470278	-3.585000	0
15	Madrid-Moratalaz	1517.0	1604.10	79614	40.410000	-3.644444	2
16	Madrid-Ciudad Lineal	2431.0	1123.99	79615	40.445050	-3.651320	0
17	Madrid-Hortaleza	1885.0	1001.66	79616	40.469444	-3.640556	0
18	Madrid-Villaverde	1588.0	1066.16	79617	40.345833	-3.709444	0
19	Madrid-Villa de Vallecas	1318.0	1194.22	79618	40.379722	-3.621389	1
20	Madrid-Vicálvaro	929.0	1288.65	79619	40.404167	-3.608056	1

Classifying districts: visualization

- Districts need to **be seen over a map**:
    - To see their **classification by number of infections** (where *yellow* corresponds to *low* incidence, *orange* to *medium* and *red* to *high*)
    - To check whether **coordinates are right** (each district has its own marker over the map)
- Using **folium** and **choropleth**, this is the resulting map:

Analysing data from FourSquare

- Dataframe which contains **FourSquare** information looks like this:

	District	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	10th Most Common Venue
0	Madrid-Arganzuela	Spanish Restaurant	Restaurant	Grocery Store	Tapas Restaurant	Bakery	Gym / Fitness Center	Burger Joint	Hotel	Sandwich Place	Falafel Restaurant
1	Madrid-Barajas	Hotel	Restaurant	Spanish Restaurant	Supermarket	Tapas Restaurant	Coffee Shop	Brewery	Fast Food Restaurant	Bistro	Japanese Restaurant
2	Madrid-Carabanchel	Metro Station	Soccer Field	Burger Joint	Fast Food Restaurant	Tapas Restaurant	Nightclub	Bakery	Plaza	Pizza Place	Pharmacy
3	Madrid-Centro	Tapas Restaurant	Spanish Restaurant	Plaza	Hostel	Bar	Ice Cream Shop	Restaurant	Cocktail Bar	Coffee Shop	Pastry Shop
4	Madrid-Chamartín	Spanish Restaurant	Restaurant	Grocery Store	Park	Tapas Restaurant	Bakery	Café	Japanese Restaurant	Gastropub	Coffee Shop
5	Madrid-Chamberí	Spanish Restaurant	Bar	Restaurant	Café	Tapas Restaurant	Brewery	Italian Restaurant	Plaza	Bakery	Coffee Shop
6	Madrid-Ciudad Lineal	Spanish Restaurant	Grocery Store	Hotel	Pizza Place	Italian Restaurant	Tapas Restaurant	Bar	Motorcycle Shop	Fast Food Restaurant	Park
7	Madrid-Fuencarral-El Pardo	Clothing Store	Italian Restaurant	Fast Food Restaurant	Tapas Restaurant	Burger Joint	Sandwich Place	Restaurant	Chocolate Shop	Pizza Place	Pharmacy
8	Madrid-Hortaleza	Pizza Place	Breakfast Spot	Supermarket	Spanish Restaurant	Dance Studio	Pharmacy	Plaza	Clothing Store	Donut Shop	Chinese Restaurant
9	Madrid-Latina	Pizza Place	Falafel Restaurant	Metro Station	Train Station	Food & Drink Shop	Park	Grocery Store	Asian Restaurant	Fast Food Restaurant	Bakery
		Spanish Restaurant	Gym	Pool	Pub	Stadium	Wine Shop	Falafel Restaurant	Food	Flea Market	Fish Market
11	Madrid-Moratalaz	Park	Bar	Bakery	Food Truck	Café	Nightclub	Soccer Field	Ice Cream Shop	Pizza Place	Plaza
12	Madrid-Puente de Vallecas	Gym	Pub	Pizza Place	Music Venue	Sandwich Place	Spanish Restaurant	Concert Hall	Performing Arts Venue	Camera Store	Café
13	Madrid-Retiro	Spanish Restaurant	Supermarket	Tapas Restaurant	Bar	Grocery Store	Mediterranean Restaurant	Diner	Brewery	Burger Joint	Food & Drink Shop
14	Madrid-Salamanca	Spanish Restaurant	Restaurant	Seafood Restaurant	Mediterranean Restaurant	Coffee Shop	Burger Joint	Supermarket	Tapas Restaurant	Mexican Restaurant	Bakery
15	Madrid-San Blas - Canillejas	Metro Station	Asian Restaurant	Shopping Mall	Snack Place	Supermarket	Gas Station	Pizza Place	Gym	Grocery Store	Dance Studio
16	Madrid-Tetuán	Spanish Restaurant	Bakery	Coffee Shop	Supermarket	Chinese Restaurant	Brazilian Restaurant	Grocery Store	Farmers Market	Pizza Place	Gym / Fitness Center
17	Madrid-Usera	Seafood Restaurant	Bubble Tea Shop	Spanish Restaurant	Asian Restaurant	Theater	Noodle House	Fast Food Restaurant	Chinese Restaurant	Wine Shop	Farmers Market
18	Madrid-Vicálvaro	Pizza Place	Spanish Restaurant	Beer Bar	Fast Food Restaurant	Sandwich Place	Café	Breakfast Spot	Metro Station	Grocery Store	Ice Cream Shop
19	Madrid-Villa de Vallecas	Metro Station	Grocery Store	Spanish Restaurant	Food	Soccer Field	Train Station	Department Store	Dessert Shop	Fast Food Restaurant	Food Court
20	Madrid-Villaverde	Pizza Place	Spanish Restaurant	Train	Diner	Falafel Restaurant	Food & Drink Shop	Food	Flea Market	Fish Market	Fast Food Restaurant

- Most common categories in all districts **are related with restaurant business**, what can mean a **handicap** for this project's goal.
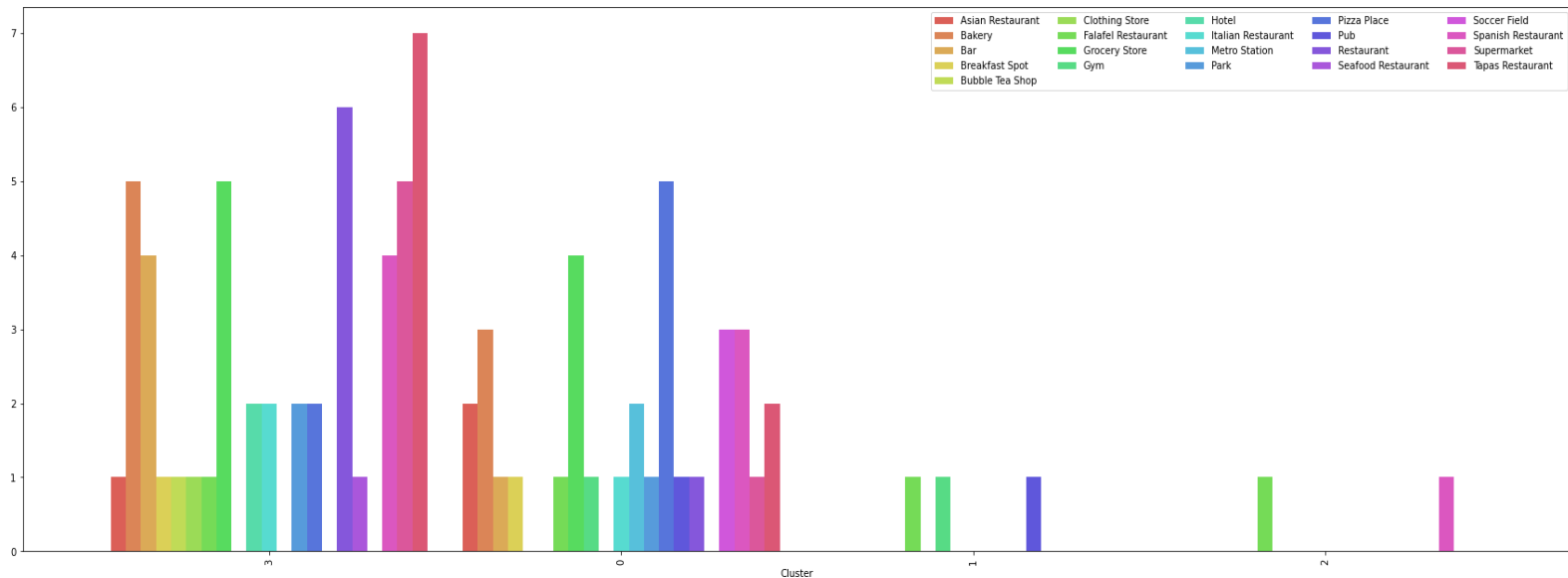
Applying machine learning *(1/3)*

- Using their most common venues, we will classify the districts using **K-means algorithm**
- Having **four clusters**, the result after applying K-means looks as follows:

	Cluster Labels	District	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	10th Most Common Venue
0	3	Madrid-Arganzuela	Spanish Restaurant	Restaurant	Grocery Store	Tapas Restaurant	Bakery	Gym / Fitness Center	Burger Joint	Hotel	Sandwich Place	Falafel Restaurant
1	3	Madrid-Barajas	Hotel	Restaurant	Spanish Restaurant	Supermarket	Tapas Restaurant	Coffee Shop	Brewery	Fast Food Restaurant	Bistro	Japanese Restaurant
2	0	Madrid-Carabanchel	Metro Station	Soccer Field	Burger Joint	Fast Food Restaurant	Tapas Restaurant	Nightclub	Bakery	Plaza	Pizza Place	Pharmacy
3	3	Madrid-Centro	Tapas Restaurant	Spanish Restaurant	Plaza	Hostel	Bar	Ice Cream Shop	Restaurant	Cocktail Bar	Coffee Shop	Pastry Shop
4	3	Madrid-Chamartín	Spanish Restaurant	Restaurant	Grocery Store	Park	Tapas Restaurant	Bakery	Café	Japanese Restaurant	Gastropub	Coffee Shop
5	3	Madrid-Chamberí	Spanish Restaurant	Bar	Restaurant	Café	Tapas Restaurant	Brewery	Italian Restaurant	Plaza	Bakery	Coffee Shop
6	3	Madrid-Ciudad Lineal	Spanish Restaurant	Grocery Store	Hotel	Pizza Place	Italian Restaurant	Tapas Restaurant	Bar	Motorcycle Shop	Fast Food Restaurant	Park
7	0	Madrid-Fuencarral-El Pardo	Clothing Store	Italian Restaurant	Fast Food Restaurant	Tapas Restaurant	Burger Joint	Sandwich Place	Restaurant	Chocolate Shop	Pizza Place	Pharmacy
8	3	Madrid-Hortaleza	Pizza Place	Breakfast Spot	Supermarket	Spanish Restaurant	Dance Studio	Pharmacy	Plaza	Clothing Store	Donut Shop	Chinese Restaurant
9	0	Madrid-Latina	Pizza Place	Falafel Restaurant	Metro Station	Train Station	Food & Drink Shop	Park	Grocery Store	Asian Restaurant	Fast Food Restaurant	Bakery
10	1	Madrid-Moncloa-Aravaca	Spanish Restaurant	Gym	Pool	Pub	Stadium	Wine Shop	Falafel Restaurant	Food	Flea Market	Fish Market
11	0	Madrid-Moratalaz	Park	Bar	Bakery	Food Truck	Café	Nightclub	Soccer Field	Ice Cream Shop	Pizza Place	Plaza
12	0	Madrid-Puente de Vallecas	Gym	Pub	Pizza Place	Music Venue	Sandwich Place	Spanish Restaurant	Concert Hall	Performing Arts Venue	Camera Store	Café
13	3	Madrid-Retiro	Spanish Restaurant	Supermarket	Tapas Restaurant	Bar	Grocery Store	Mediterranean Restaurant	Diner	Brewery	Burger Joint	Food & Drink Shop
14	3	Madrid-Salamanca	Spanish Restaurant	Restaurant	Seafood Restaurant	Mediterranean Restaurant	Coffee Shop	Burger Joint	Supermarket	Tapas Restaurant	Mexican Restaurant	Bakery
15	0	Madrid-San Blas - Canillejas	Metro Station	Asian Restaurant	Shopping Mall	Snack Place	Supermarket	Gas Station	Pizza Place	Gym	Grocery Store	Dance Studio
16	3	Madrid-Tetuán	Spanish Restaurant	Bakery	Coffee Shop	Supermarket	Chinese Restaurant	Brazilian Restaurant	Grocery Store	Farmers Market	Pizza Place	Gym / Fitness Center
17	3	Madrid-Usera	Seafood Restaurant	Bubble Tea Shop	Spanish Restaurant	Asian Restaurant	Theater	Noodle House	Fast Food Restaurant	Chinese Restaurant	Wine Shop	Farmers Market
18	0	Madrid-Vicálvaro	Pizza Place	Spanish Restaurant	Beer Bar	Fast Food Restaurant	Sandwich Place	Café	Breakfast Spot	Metro Station	Grocery Store	Ice Cream Shop
19	0	Madrid-Villa de Vallecas	Metro Station	Grocery Store	Spanish Restaurant	Food	Soccer Field	Train Station	Department Store	Dessert Shop	Fast Food Restaurant	Food Court
20	2	Madrid-Villaverde	Pizza Place	Spanish Restaurant	Train	Diner	Falafel Restaurant	Food & Drink Shop	Food	Flea Market	Fish Market	Fast Food Restaurant

- As we can see, "**3**" is the **most common cluster**, followed by cluster "**0**" and leaving "**1**" and "**2**" in last places.

Applying machine learning *(2/3)*

- In order to make out what each cluster have in common, we will try to *visualize each cluster in a graph*.

- Using *matplotlib* and the *1ˢᵗ*, the *2ⁿᵈ* and *the 3ʳᵈ most common venues* for each district, the result is shown below:

Applying machine learning *(3/3)*

- Studying each cluster separately, we can *come to some conclusions*:
  - Cluster "*3*" groups the *most varied categories of venues* and would probably correspond to *districts very touristic*: apart from places usually *crowded by residents* (bakery and grocery stores), there are also *places crowded by tourist*
  - Cluster "*0*" has *a lot in common with cluster "3"* but, in this case, *parks* are the most common venues. The districts from these clusters would probably be *less touristic* than districts of cluster "3"
  - Cluster "*1*" has a kind of *restaurant*, but also *gyms* and *pubs*. We can suppose that this cluster correspond to *districts with more residents than tourists* and venues *less overcrowded*
  - Cluster "*2*" only has *2 kinds of restaurants* as its most popular venues
- Knowing how COVID-19 spreads, we can expect that those districts which belong to cluster "*3*" and "*0*" have a *high total cumulative incidence*
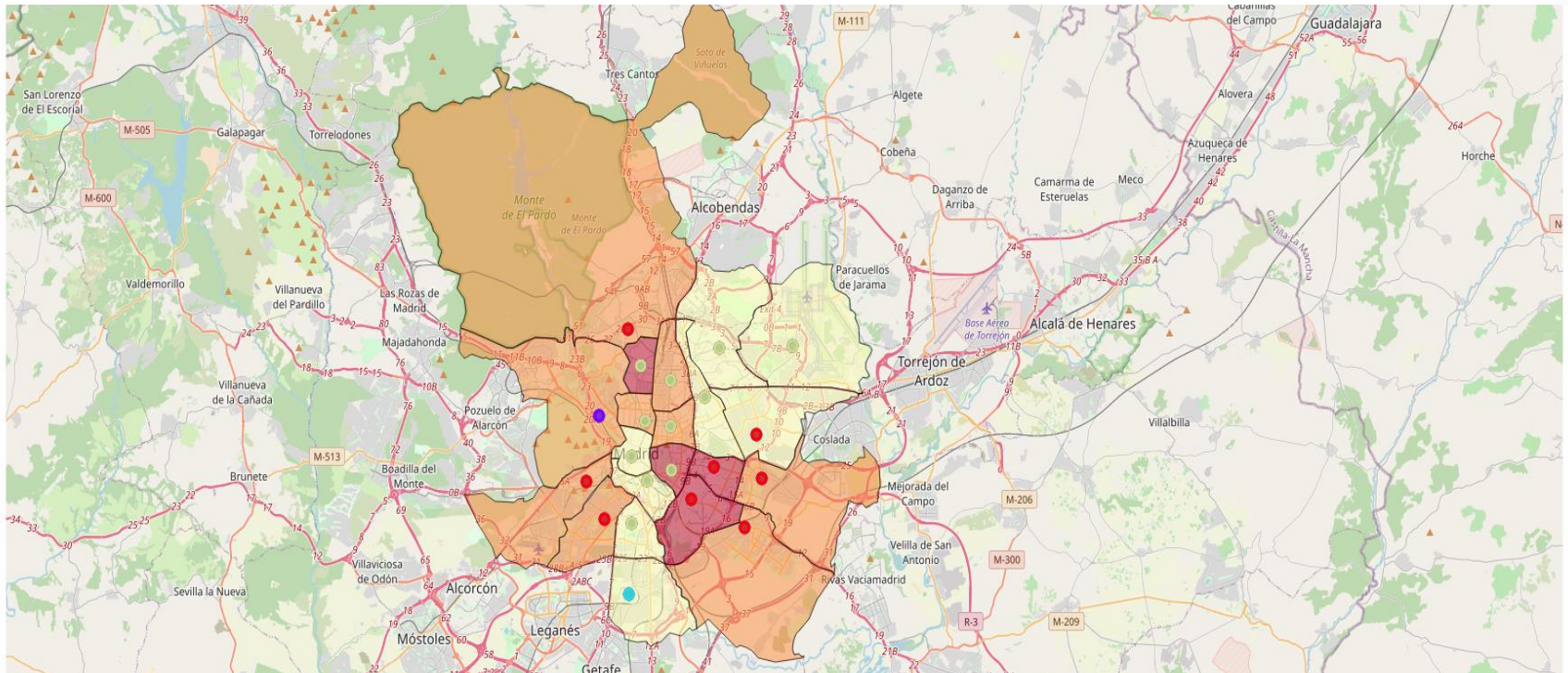
Results *(1/2)*

- In order to know if **COVID-19 infections are related with some kind of places**, we need to **see together the districts labelled by COVID-19 incidence** and the **districts clustered by their most common venues**:

	town_district	total_positive_tests	total_cumulative_rate	district_code	latitude	longitude	total_cumulative_rate_label	Cluster Labels	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	10th Most Common Venue
0	Madrid-Retiro	1691.0	1417.23	79603	40.408056	-3.676667	2	3	Spanish Restaurant	Supermarket	Tapas Restaurant	Bar	Grocery Store	Mediterranean Restaurant	Diner	Brewery	Burger Joint	Food & Drink Shop
1	Madrid-Salamanca	1781.0	1218.99	79604	40.430000	-3.677778	1	3	Spanish Restaurant	Restaurant	Seafood Restaurant	Mediterranean Restaurant	Coffee Shop	Burger Joint	Supermarket	Tapas Restaurant	Mexican Restaurant	Bakery
2	Madrid-Centro	1282.0	950.70	79601	40.415278	-3.707500	0	3	Tapas Restaurant	Spanish Restaurant	Plaza	Hostel	Bar	Ice Cream Shop	Restaurant	Cocktail Bar	Coffee Shop	Pastry Shop
3	Madrid-Arganzuela	1769.0	1150.17	79602	40.402778	-3.695278	0	3	Spanish Restaurant	Restaurant	Grocery Store	Tapas Restaurant	Bakery	Gym / Fitness Center	Burger Joint	Hotel	Sandwich Place	Falafel Restaurant
4	Madrid-Chamartín	1800.0	1234.15	79605	40.453333	-3.677500	1	3	Spanish Restaurant	Restaurant	Grocery Store	Park	Tapas Restaurant	Bakery	Café	Japanese Restaurant	Gastropub	Coffee Shop
5	Madrid-Tetuán	2233.0	1413.09	79606	40.460556	-3.700000	2	3	Spanish Restaurant	Bakery	Coffee Shop	Supermarket	Chinese Restaurant	Brazilian Restaurant	Grocery Store	Farmers Market	Pizza Place	Gym / Fitness Center
6	Madrid-Chamberí	1739.0	1247.68	79607	40.432778	-3.697222	1	3	Spanish Restaurant	Bar	Restaurant	Café	Tapas Restaurant	Brewery	Italian Restaurant	Plaza	Bakery	Coffee Shop
7	Fuencarral-El Pardo	3290.0	1337.73	79608	40.478611	-3.709722	1	0	Clothing Store	Italian Restaurant	Fast Food Restaurant	Tapas Restaurant	Burger Joint	Sandwich Place	Restaurant	Chocolate Shop	Pizza Place	Pharmacy
8	Madrid-Moncloa-Aravaca	1568.0	1313.02	79609	40.435470	-3.731700	1	1	Spanish Restaurant	Gym	Pool	Pub	Stadium	Wine Shop	Falafel Restaurant	Food	Flea Market	Fish Market
9	Madrid-Latina	3080.0	1292.93	79610	40.402500	-3.741389	1	0	Pizza Place	Falafel Restaurant	Metro Station	Train Station	Food & Drink Shop	Park	Grocery Store	Asian Restaurant	Fast Food Restaurant	Bakery
10	Madrid-Carabanchel	2957.0	1168.32	79611	40.383611	-3.728056	1	0	Metro Station	Soccer Field	Burger Joint	Fast Food Restaurant	Tapas Restaurant	Nightclub	Bakery	Plaza	Pizza Place	Pharmacy
11	Madrid-Usera	1505.0	1076.99	79612	40.381389	-3.706944	0	3	Seafood Restaurant	Bubble Tea Shop	Spanish Restaurant	Asian Restaurant	Theater	Noodle House	Fast Food Restaurant	Chinese Restaurant	Wine Shop	Farmers Market
12	Madrid-Puente de Vallecas	3348.0	1425.55	79613	40.393540	-3.662000	2	0	Gym	Pub	Pizza Place	Music Venue	Sandwich Place	Spanish Restaurant	Concert Hall	Performing Arts Venue	Camera Store	Café
13	Madrid-San Blas - Canillejas	1584.0	1001.63	79620	40.426111	-3.612778	0	0	Metro Station	Asian Restaurant	Shopping Mall	Snack Place	Supermarket	Gas Station	Pizza Place	Gym	Grocery Store	Dance Studio

Results *(2/2)*

- Visualizing this data frame **over a map**, where "**green**" markers correspond to cluster "**3**", "**red**" markers to cluster "**0**", "**purple**" markers to cluster "**1**" and "**blue**" markers to cluster "**2**":

Discussion

- According to the obtained results, it is ***not possible to state by them which kinds of venues are the most dangerous*** when we talk about getting infected by COVID-19:
  - As we had supposed, districts which belong to clusters "***3***" and "***0***" are those which have ***more infections by COVID-19***, probably to the ***great variety of venues*** in those clusters and the ***traffic of people*** between them (restaurants, hotels and metro stations)
  - On the contrary, districts which belong to clusters "***1***" and "***2***" (these were the clusters with ***less variety of venues***) do ***not have high number of infections*** by COVID-19

17

Conclusion

- The ***results were not the expected***: the more the project progressed, the less enthusiastic I was
- Some ***problems*** during its development:
    - ***Lack of datasets*** from public sources
    - FourSquare does not provide ***most popular venues in a period of time***
    - FourSquare does not collect ***the information that would have best fit for this project***: tourists are the most common users of this application (a resident does not check in the restaurant where he/she eat every day), but they do not count in Madrid's statistics about COVID-19
    - ***Spanish culture***: restaurants and bars are the most common businesses, what makes ***impossible to differentiate between kind of venues provided by FourSquare*** (four our purpose, a Italian restaurant makes no difference to a Chinese restaurant, for example: they are both restaurants, but FourSquare divides them in different categories
- Suggestions to improve the results:
    - Using Google Maps instead of FourSquare
    - Using the information collected by Madrid underground
    - Using the population density, mean income and/or number of residents per home by district