Analyzing neighborhoods of Madrid, Spain

Introduction

Madrid, capital of Spain

Where to open a new restaurant?

Pros: more opportunities

Cons: hard to choose location

Methodology

- **Data:**

Wikipedia: list of neighborhoods

Geopy: latitude and longitude

Foursquare: nearby venues

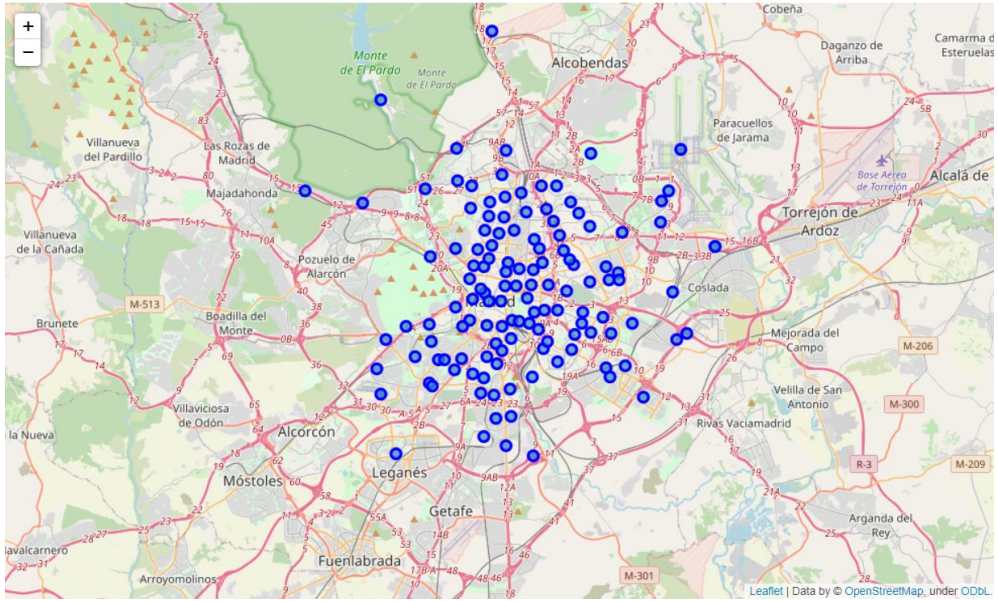
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Palacio	40.417821	-3.715111	Palacio Real de Madrid	40.417940	-3.714259	Palace
1	Palacio	40.417821	-3.715111	Plaza de la Almudena	40.416320	-3.713777	Plaza
2	Palacio	40.417821	-3.715111	Santa Iglesia Catedral de Santa María la Real ...	40.415767	-3.714516	Church
3	Palacio	40.417821	-3.715111	Plaza de Oriente	40.418326	-3.712196	Plaza
4	Palacio	40.417821	-3.715111	Jardines de Sabatini	40.419954	-3.713126	Garden

Methodology

- **Exploratory data analysis:**

Normal distribution on a map

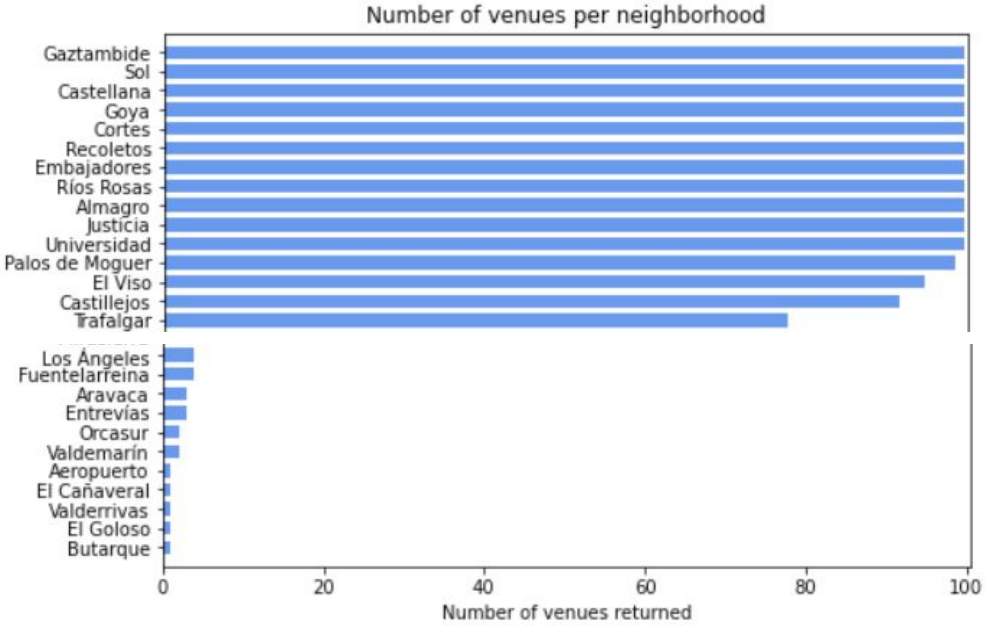
(not really helpful)



Methodology

- **Exploratory data analysis:**

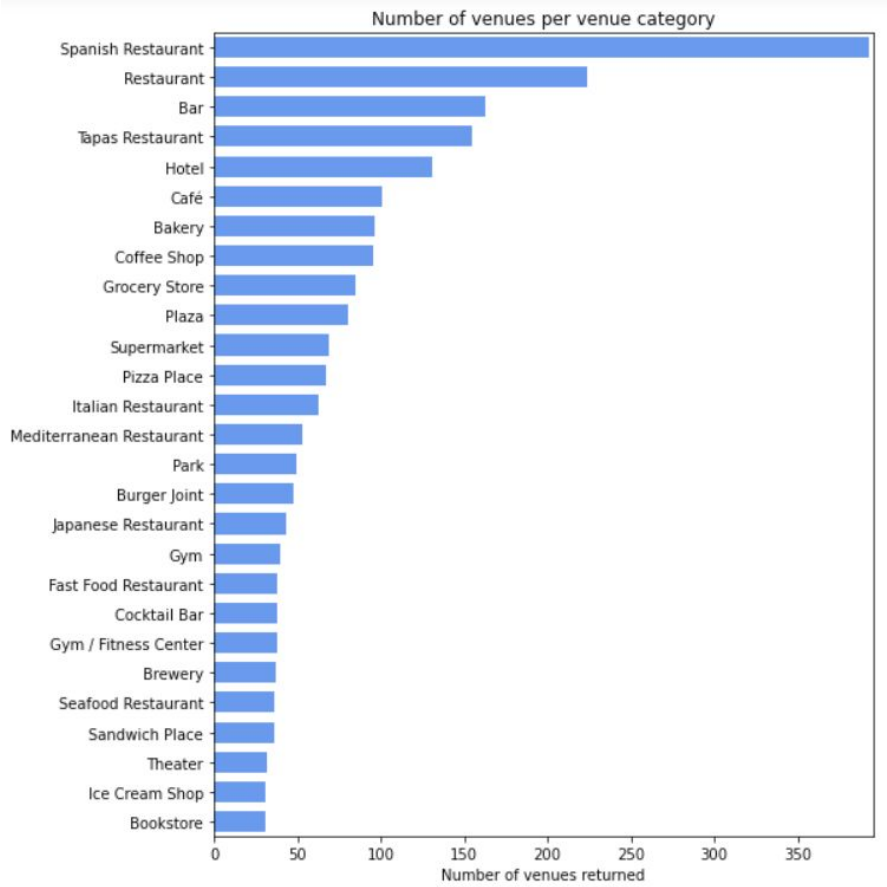
Some neighborhoods are bigger and more important than others



Methodology

- **Exploratory data analysis:**

The most common venue categories are related to restaurants



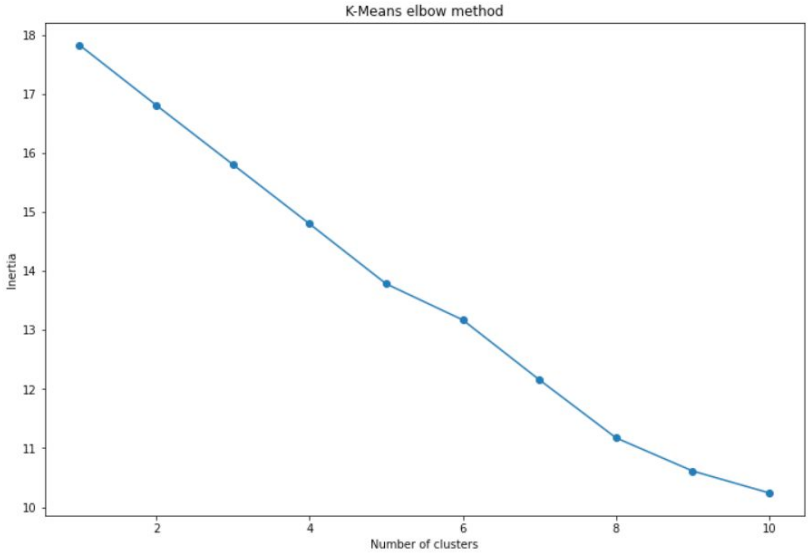
Methodology

- **Machine learning**

k-means clustering

Optimal k ≈ 8

Final data frame:



	Neighborhood	1st most common venue	2nd most common venue	3rd most common venue	4th most common venue	5th most common venue	Latitude	Longitude	Label
0	Abrantes	Bakery (0.2)	Soccer Field (0.2)	Gym / Fitness Center (0.1)	Pizza Place (0.1)	Park (0.1)	40.380998	-3.727985	3
1	Acacias	Bar (0.09)	Tapas Restaurant (0.07)	Spanish Restaurant (0.07)	Pizza Place (0.07)	Art Gallery (0.05)	40.404075	-3.705957	0
2	Adelfas	Supermarket (0.06)	Grocery Store (0.06)	Tapas Restaurant (0.06)	Bar (0.06)	Fast Food Restaurant (0.06)	40.401903	-3.670958	0
3	Aeropuerto	Business Service (1.0)	Women's Store (0.0)	Escape Room (0.0)	Food Court (0.0)	Food & Drink Shop (0.0)	40.494838	-3.574081	7
4	Alameda de Osuna	Restaurant (0.08)	Tapas Restaurant (0.08)	Hotel (0.08)	Bakery (0.08)	Cocktail Bar (0.04)	40.457581	-3.587975	0

Results and discussion

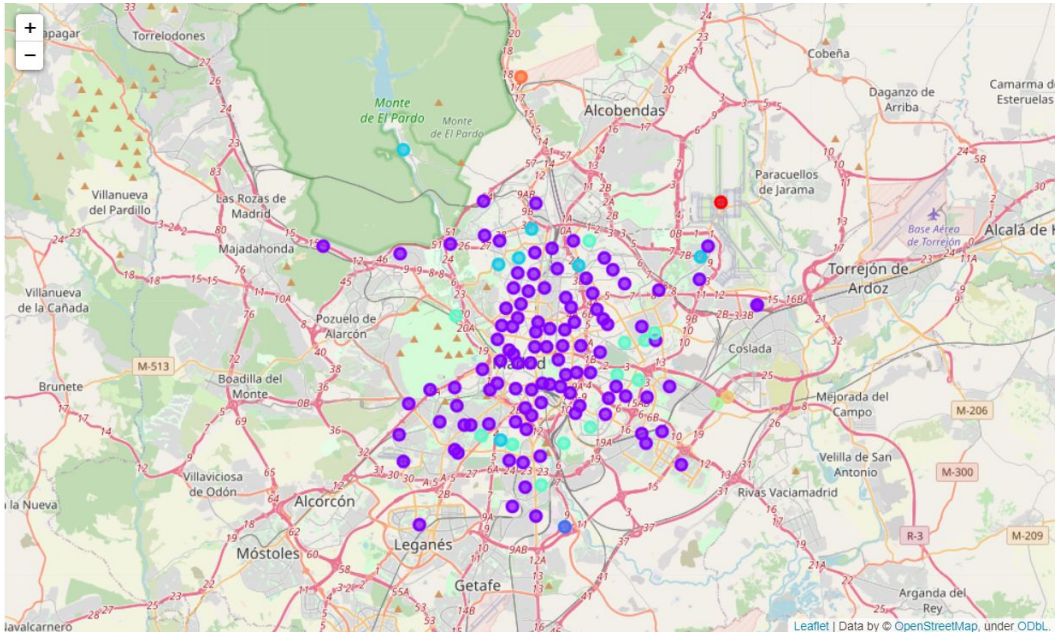
Cluster 0: general (purple)

Cluster 1, 4, 5, 6: “noise”
(dark blue, green, yellow, orange)

Cluster 2: restaurants (blue)
Best one to open restaurant

Cluster 3: sport (cyan)

Cluster 7: airport (red)



Conclusion

- Analysis of neighborhoods of Madrid to open a restaurant
- Data from Wikipedia and Foursquare (among others)
- Clusterization using k-means
- **Best location:** neighborhoods of cluster number 2