

# Report\_Madrid

July 28, 2019

## 1 The Battle Of Neighborhoods - Madrid

### 1.1 Introduction

Madrid is the capital of Spain and the largest city of Spain. The city has a population of almost 3.3m but the metropolitan area has approximately 6.5 million inhabitants. The municipality covers 604.3 km<sup>2</sup> (233.3 sq mi) and its metropolitan area is the third largest in the EU, after London and Paris.

The city has 21 boroughs and 131 neighborhoods but in this project we are going to focus on the area known as “Madrid Central” and its extension to the ring inside the M-30 motorway. The meaning of “Madrid Central” is that there are traffic restrictions to this area so it is very important a good location near public transport.

### 1.2 Business Problem

A foreign investor is looking to open a new restaurant in Madrid, so he has contacted me in order to help him in finding the best location for the business. On the first conversation, he asked me some questions of my knowledge of the city and I have explained that from my own experience, the offer in Madrid is huge and you can find restaurants of all types, but mainly Spanish cuisine.

After this first approach, he let me know his intentions of opening a fusion of Spanish with Mediterranean and Asian cuisine. He is looking to compete with tapas style bars but from a restaurant approach.

In order for me to complete this task, he has provided me with some guidelines of what he is looking for:

- An area where potential clients are tourists but also locals.
- The area has good transport connection: underground stations, parking.
- Busy area where the value added of his restaurant could make an impact “borrowing” customers from restaurants around.

### 1.3 Data description

The data used to approach this problem is explained as per below:

- In order to get the coordinates of Madrid, I used the data from the City Council of Madrid at their website [datos.madrid.es](https://datos.madrid.es) and with the help of Geopandas and Geocoding API from Google, I have created “coordenadas.csv” file in a different project. I also used the Geojson file for area delimitation on the map.
- I used Foursquare API to get the most common venues of Madrid Central.

## 1.4 Methodology

Using the data with the coordinates of the different neighborhoods of Madrid, we will explore them using Foursquare API.

We will use the explore function to retrieve the following data:

- Most common venue categories in each neighborhood
- Most common travel and transports venues in each neighborhood
- Most common food venues in each neighborhood

We will use these features to group the neighborhoods into clusters. I will use k-means clustering algorithm to complete this. I will visualize the neighborhoods in Madrid and their emergin clusters using Folium library.

**Madrid Central** We can see below the map of the area we are going to study:



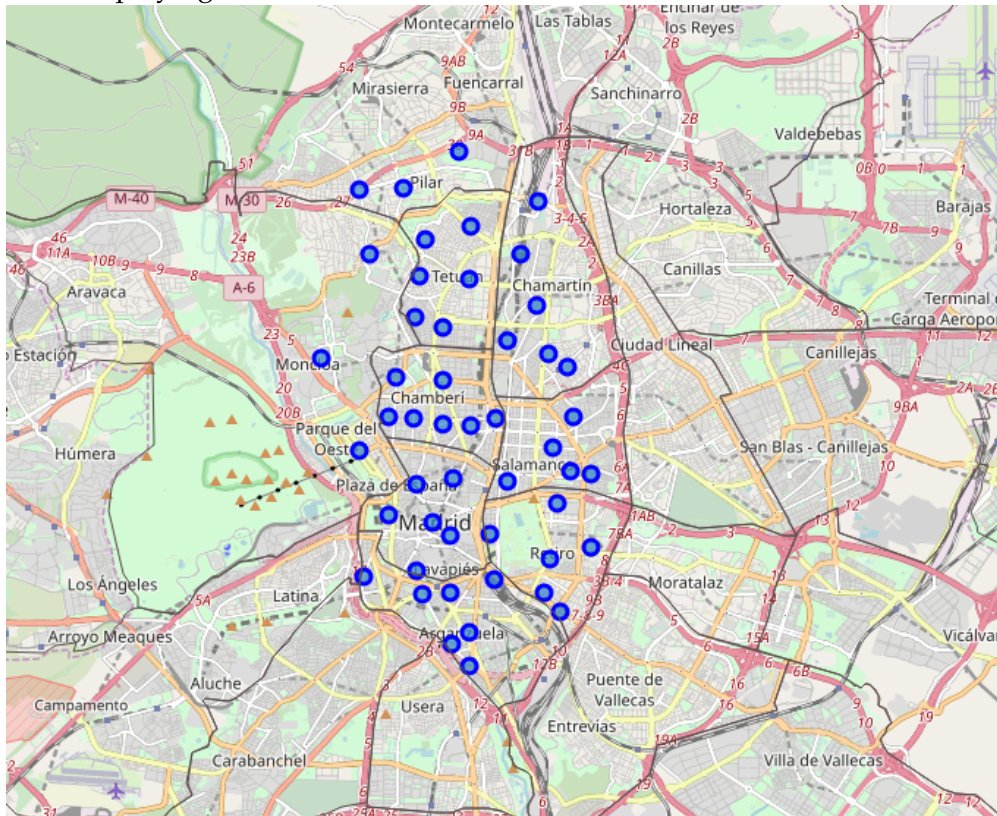
### Data Cleaning

**Neighborhood coordinates** In order to get a clean file for the Boroughs and Neighborhoods we have performed a data cleaning using geopandas and pandas. This can be found in the attached notebook in my github. In this [project](#), we have used also the Geocoding API from Google to get the coordinates of each neighborhood.

We started with a data of 21 boroughs and 131 neighborhoods but we needed just the area inside the M-30. So, after performing some manipulation of the data, we have finally created our dataframe for Madrid Central, with 9 boroughs and 49 neighborhoods.

**Folium and Madrid Central** So, in order to have a first view, we have visualized the boroughs to be analysed in a map using Folium. Firstly, we have retrieved the coordinates of Madrid to center the map. We have used Nominatim retrieving the following coordinates: 40.4167047, -3.7035825.

So, playing with both, our coordinates and Folium we have the following map:



**Clustering** In order to do our data segmentation, we have used Foursquare API to get the venues of Madrid. We have divided our data into 3 parts: - All venues - Travel and Transports - Food

#### All venues

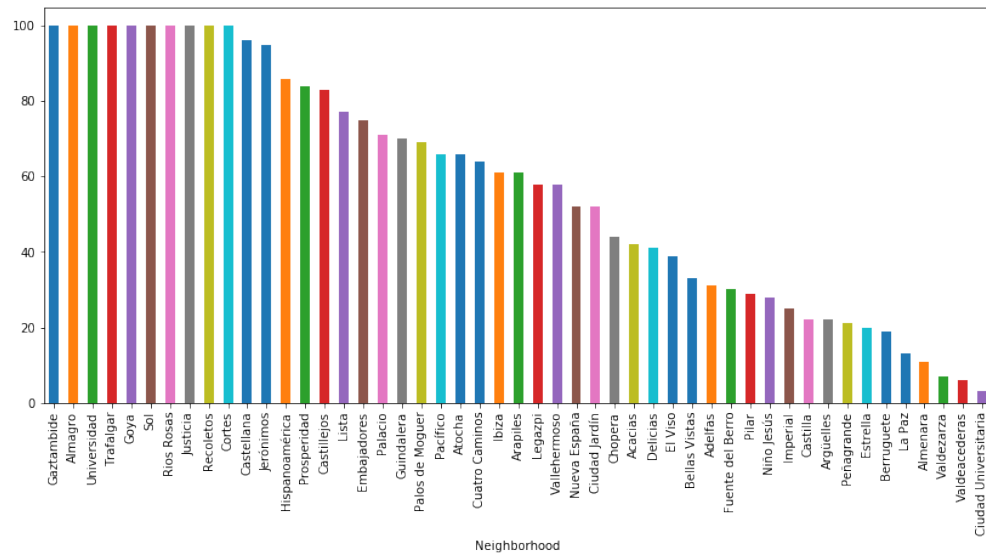
- We have retrieved in total 2830 venues:

#### Travel and Transports

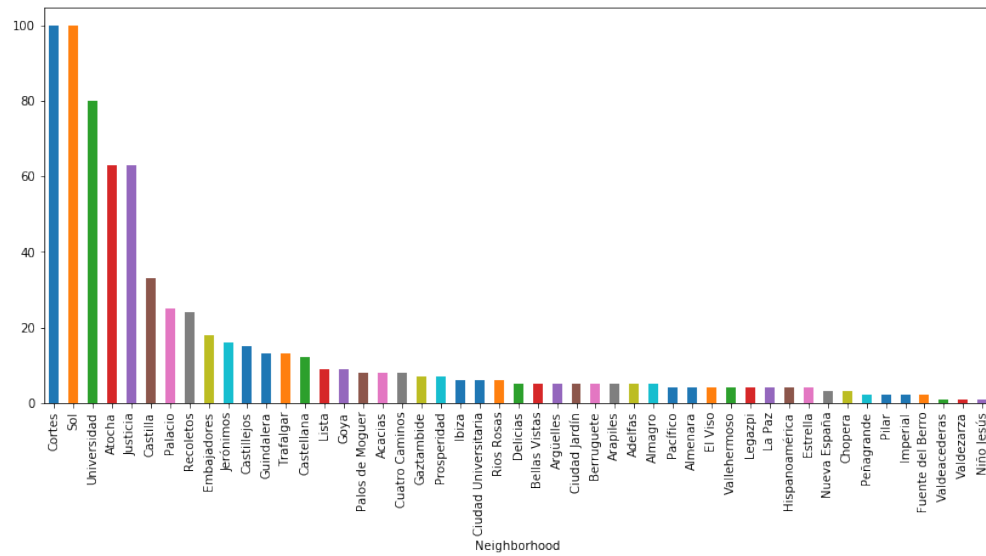
- We have retrieved in total 738 venues:

#### Food

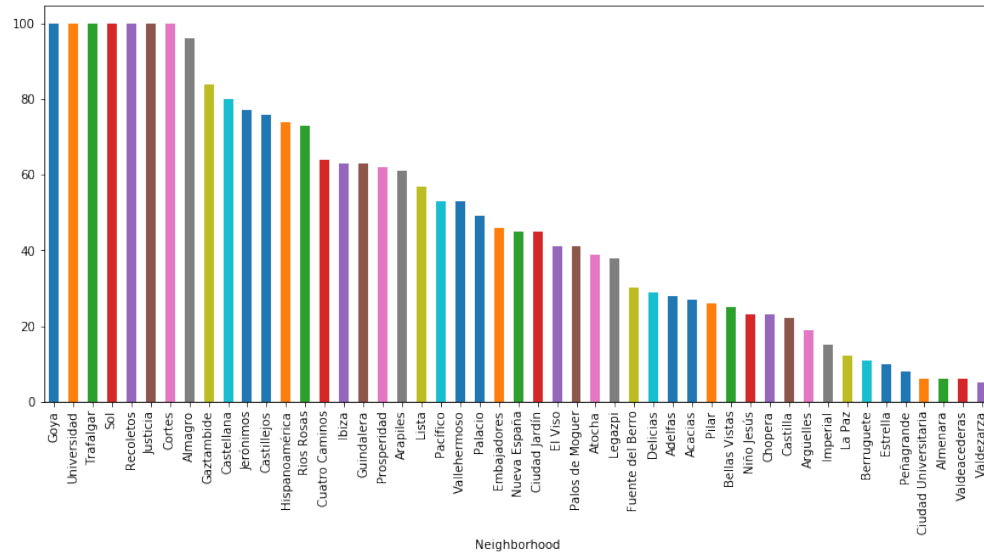
- We have retrieved in total 2411 venues:



All venues



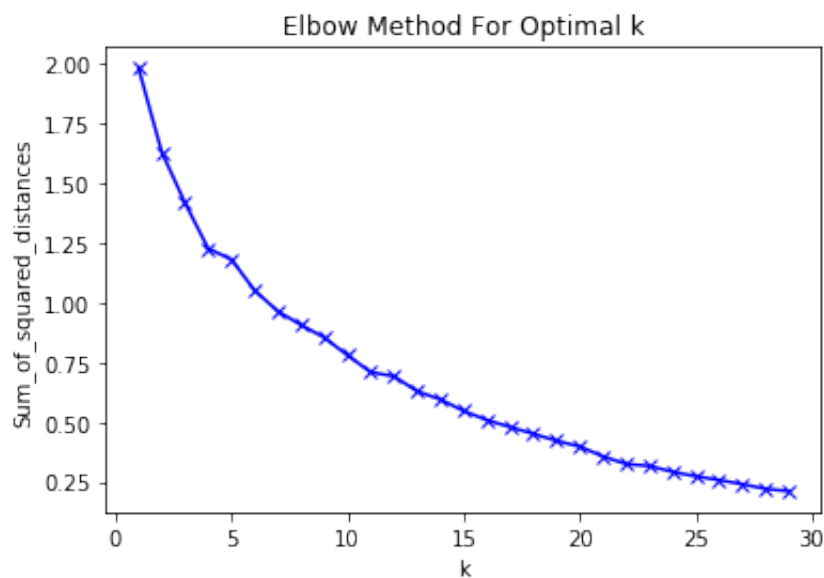
Travel and transport venues



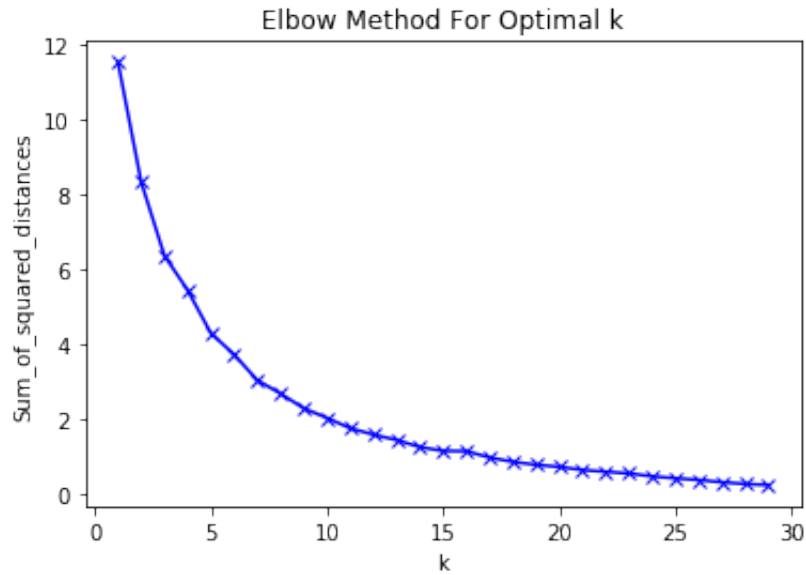
Food venues

**K-means** We have used K-means algorithm to perform our segmentation. We have performed an elbow analysis of the data to get the best number of clusters for our segmentation:

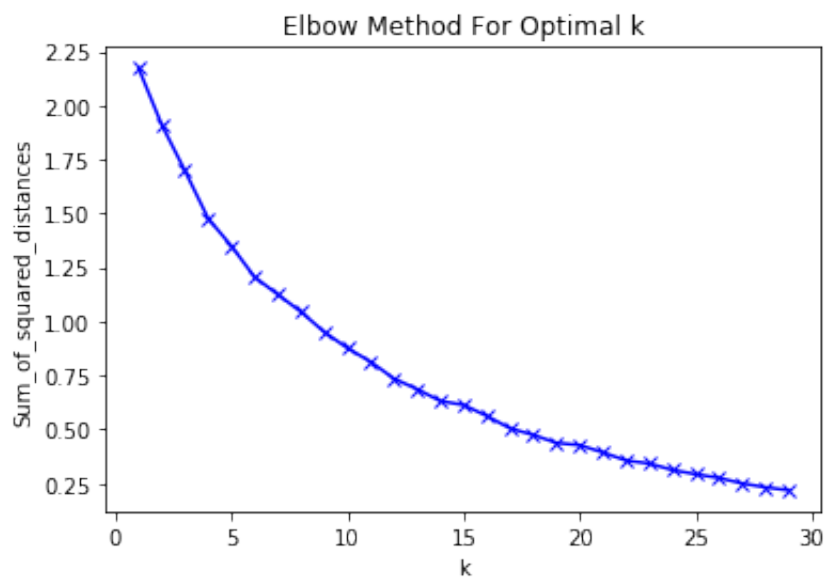
- for all venues we have decided that 11 was the optimum



- for travel and transport we have decided that 6 was the optimum



- and for food we have decided that 11 was the optimum.

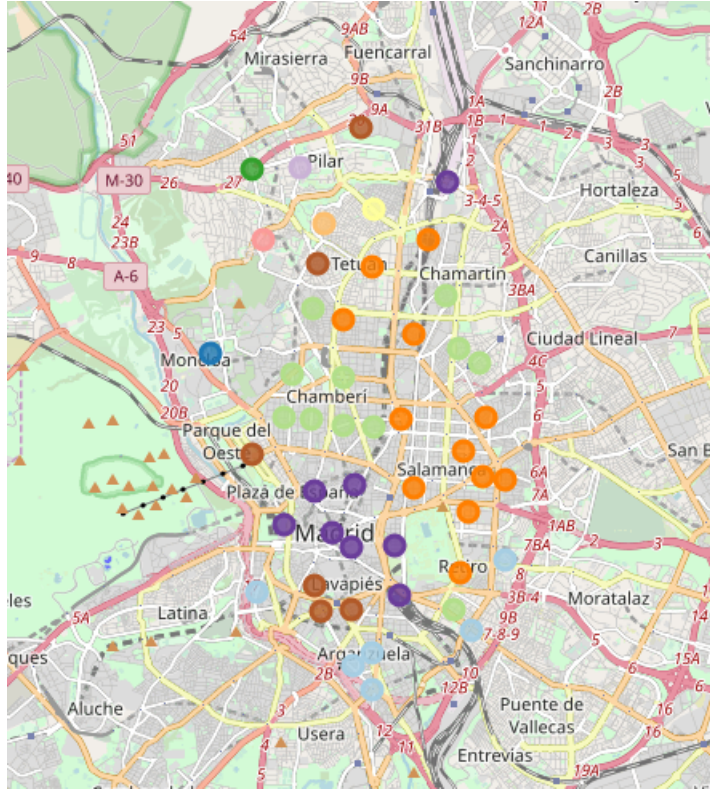


**Clusters** So, using Folium again we have visualized the clusters into the map and we have just analysed those clusters that we think are more relevant to the business requirement:

#### All Venues

- Cluster 3 (ligh green): Spanish Restaurants and Tapas, with bars and cafes.
- Cluster 7 (orange): More variaty of restaurants, more expensive area
- Cluster 9 (violet): Tourist area with Hotels, museums and also Spanish Restaurants and

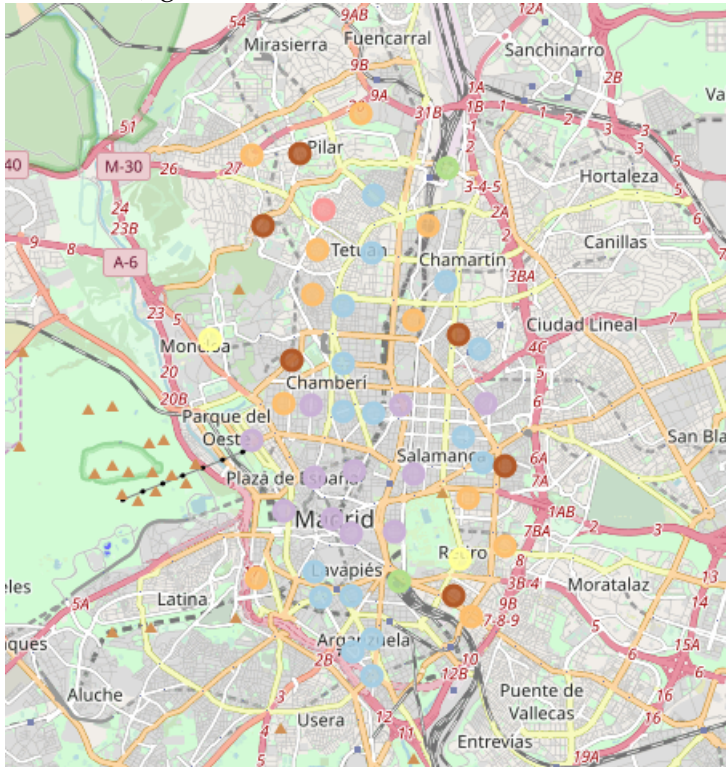




tapas

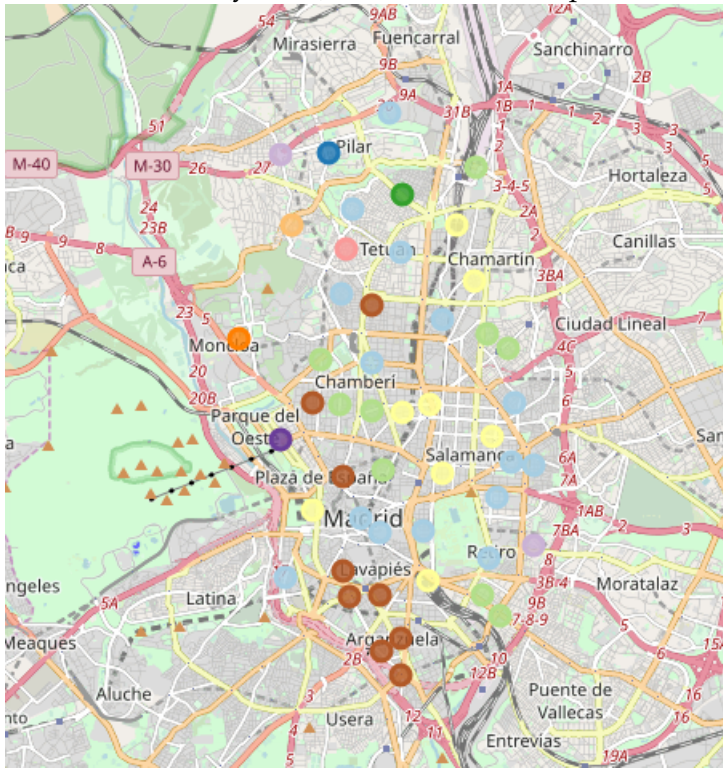
### Travel and Transport

- Cluster 1 (light blue): Transport links and hotels
- Cluster 5 (light violet): Tourist area: Hotels, bed and breakfasts and good transport links.



## Food

- Cluster 0 (brown): Middle class area with a lot of tapas restaurants
- Cluster 1 (light blue): Area with lots of restaurants and cafes.
- Cluster 3 (green): Cafes and restaurants
- Cluster 10 (yellow): More expensive restaurants and big variety



## 1.5 Results

So, we can get to the conclusion of the below results: - Tourist area with plenty of accommodation, places to visit and eat: Sol, Cortes, Embajadores and Justicia - Area full of bars, tapas and cafes: Almagro, Rio Rosas and Trafalgar - Area with restaurants more expensive and also more variety: Castellana, Lista and Recoletos

## 1.6 Discussion

As expected, the area of Sol and Cortes, is a tourist area where you can find plenty of accommodation, places to visit but also places to eat. This could be a really good area to open the new restaurant. Many tourism but also frequented by locals.

Another area that would be really good for opening this business would be Castellana, Lista and Recoletos. This area is well known because there are plenty of shops and restaurants are usually more expensive. You can find Seafood, Japanese, Italian, Argentinian but also Spanish restaurants.

It is good to understand that Madrid is full of offices all around the area we have studied and it is very common for people working at these offices, having a lunch break from around 1 to 2 hours. It is very common that many restaurants near the offices are specialized in what is called "Menu del Dia", a 3 course menu for around 10 euros. This fact is something we need to take



into consideration, as our business will not be focused on this clients but the data is probably “contaminated” of this type of restaurante, so after selecting the area of opening the new business would be very usefull to do a new evaluation taking some more particular variables.

## **1.7 Conclusion**

Madrid is a really good city to open a restaurant for several reasons: more than 6 million people visit the city from abroad and millions coming from the rest of Spain.

Also, it is interesting because it is full of offices and many people go out after work to enjoy the night life of Madrid. Usually the weather is nice and shops are open usually until 9 or 10 pm.

It is a city where a good business idea would really make a difference and will have a lot of chances of success.