Battle of Neighborhoods (Week 1 & Week 2)

1. Description of the Problem and Discussion of the Background

The main idea is compare the neighborhoods from Bilbao to explore the city's businesses

2. Initial Data Preparation (Week 1)

1. Loading Bilbao Open Data about neighborhoods

The Wikipedia page of Bilbao contains a table from the neighborhoods from the city, but in order to get data info, the Bilbao Open Data website is a portal where the city council of Bilbao upload free public data in order to facilitate the development of businesses.

2. Tidy data and splitting info
3. Get the coordinates of the neighborhoods

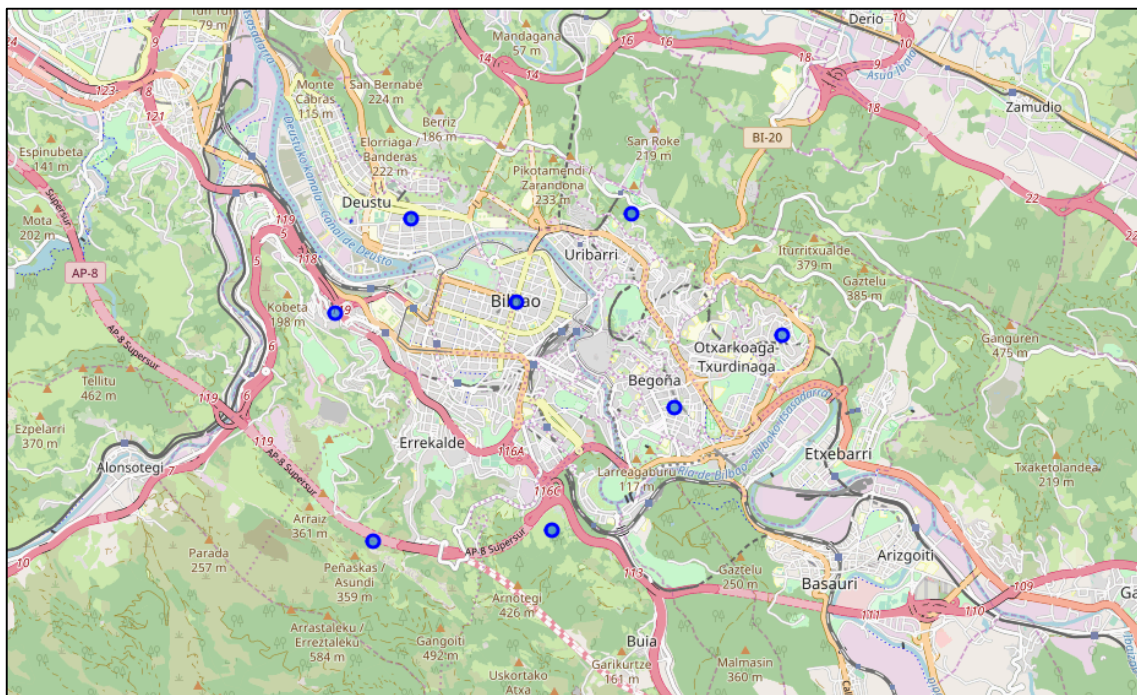|	Neighborhood	Totalhab	Latitude	Longitude
0	DEUSTU	49193	43.271387	-2.949581
1	URIBARRI	36147	42.991078	-2.615515
2	OTXARKOAGA	25706	43.259411	-2.897683
3	BEGOÑA	40563	40.480207	-3.685955
4	IBAIONDO	62619	43.239543	-2.929882
5	ABANDO	51507	43.262788	-2.934947
6	ERREKALDE	48799	42.869074	-2.229884
7	BASURTU	32549	43.261691	-2.960225

Some coordinates returned by Geopy are wrong and it's because the name of the districts are different. I found the right ones in Google and replaced it with Pandas.

4. Check information

3. Exploratory Data Analysis. (Week 2)

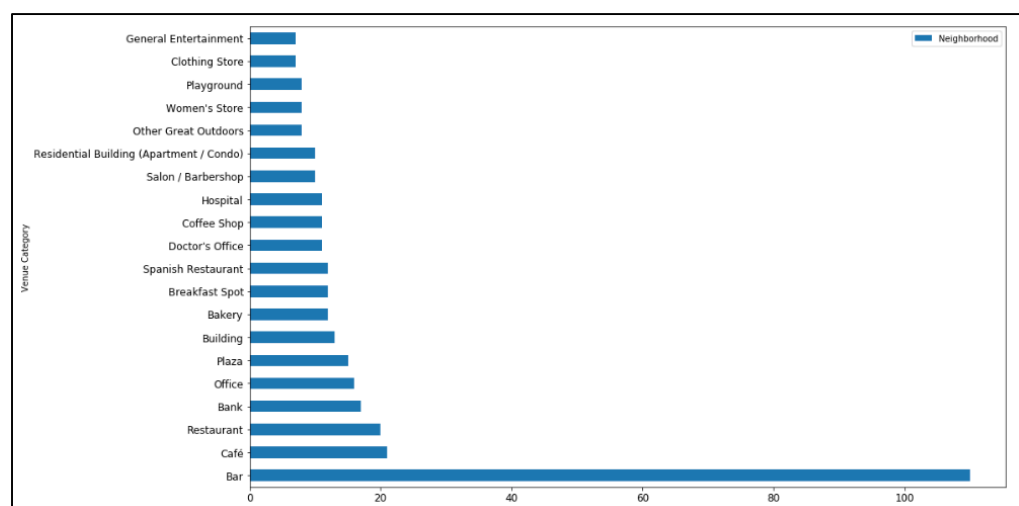1. Map of Bilbao's Neighboorhods

I have used Folium library to plot a map where to display all the Bilbao neighborhoods.

4.Getting Ready to Use FourSquare API

1. Considering the best Bilbao venues

Using Foursquare API credentials, i have managed to get the top 100 venues that are in Bilbao within a radius of 3000 meters.

5. One Hot Encoding

I proceed to create a dataframe with pandas hot encoding fo the venue categories…

	Neighborhood	Accessories Store	American Restaurant	Antique Shop	Arcade	Arts & Crafts Store	Arts & Entertainment	Asian Restaurant	Assisted Living	Athletics & Sports	...	Urgent Care Center	Vacation Rental	Veterinarian
0	ABANDO	0.032967	0.000000	0.000000	0.000000	0.000000	0.000000	0.010989	0.000000	0.000000	...	0.000000	0.000000	0.000000
1	BASURTU	0.000000	0.000000	0.010309	0.010309	0.000000	0.010309	0.000000	0.000000	0.010309	...	0.010309	0.000000	0.000000
2	BEGOÑA	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.010526	0.000000	...	0.000000	0.000000	0.000000
3	DEUSTU	0.000000	0.010204	0.000000	0.000000	0.010204	0.000000	0.000000	0.000000	0.000000	...	0.000000	0.010204	0.000000
4	ERREKALDE	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.010309	0.000000	...	0.000000	0.000000	0.000000
5	IBAIONDO	0.000000	0.000000	0.000000	0.000000	0.010870	0.000000	0.000000	0.000000	0.010870	...	0.000000	0.000000	0.000000
6	OTXARKOAGA	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.011111	0.011111	0.000000	...	0.000000	0.000000	0.011111
7	URIBARRI	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.010526	...	0.000000	0.000000	0.000000

8 rows × 206 columns

… and grouped by Bilbao's top ten most common Venues

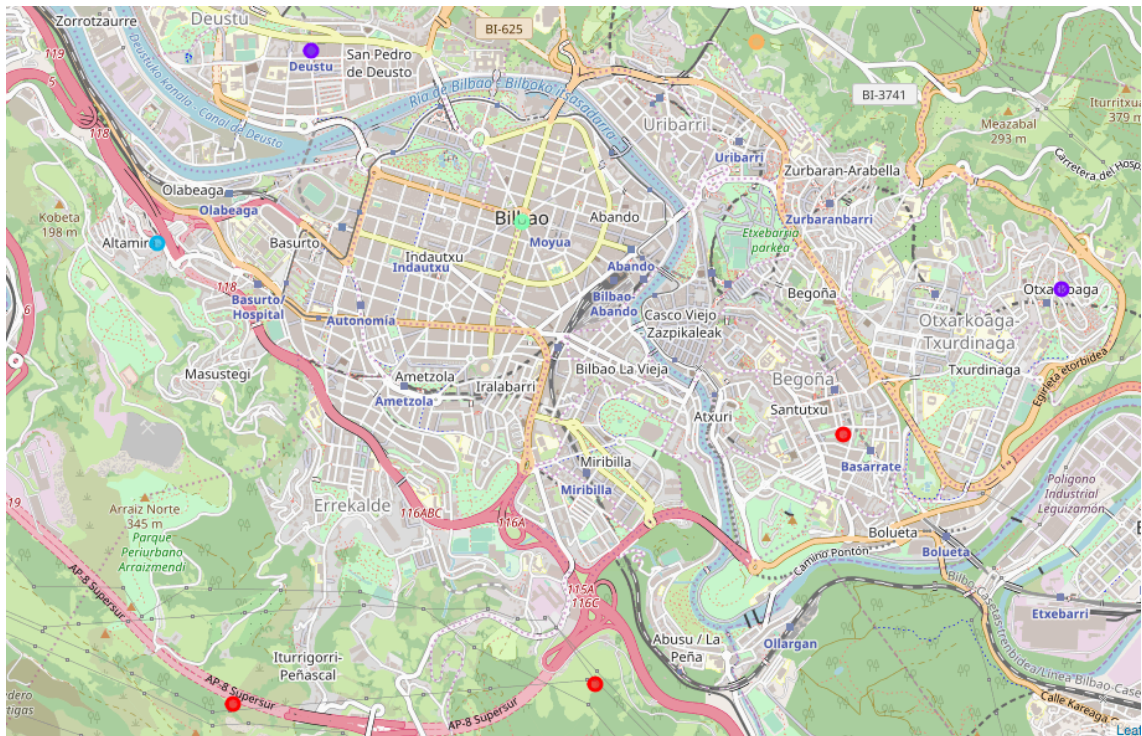|	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	ABANDO	Women's Store	Clothing Store	Bank	Restaurant	Boutique	Dessert Shop	Hotel	Accessories Store	Hotel Bar	Café
1	BASURTU	Hospital	Office	Building	Brewery	Music Venue	Restaurant	BBQ Joint	Coworking Space	Doctor's Office	Café
2	BEGOÑA	Bar	Coffee Shop	Restaurant	Playground	Plaza	Optical Shop	Café	Gift Shop	Snack Place	Salon / Barbershop
3	DEUSTU	Bar	Bank	Residential Building (Apartment / Condo)	Café	Breakfast Spot	Shopping Mall	Spanish Restaurant	Plaza	Coffee Shop	Automotive Shop
4	ERREKALDE	Bar	Bakery	Butcher	Other Great Outdoors	Grocery Store	Convenience Store	Restaurant	Breakfast Spot	Residential Building (Apartment / Condo)	Spanish Restaurant
5	IBAIONDO	Bar	Bank	Office	Salon / Barbershop	Doctor's Office	Dentist's Office	Bus Stop	Bus Line	Kebab Restaurant	Bakery
6	OTXARKOAGA	Bar	Café	General Entertainment	Bank	Police Station	Student Center	Salon / Barbershop	Dessert Shop	Bike Rental / Bike Share	Bakery
7	URIBARRI	Bar	Restaurant	Diner	Plaza	Playground	Spanish Restaurant	Other Great Outdoors	Supermarket	Office	Café

6. K-means

To get better information about these locations, i've used clustering (K-means) setting as 5 the number of clusters.

	Cluster_Labels	Neighborhood	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	ABANDO	Women's Store	Clothing Store	Bank	Restaurant	Boutique	Dessert Shop	Hotel	Accessories Store	Hotel Bar	Café
1	2	BASURTU	Hospital	Office	Building	Brewery	Music Venue	Restaurant	BBQ Joint	Coworking Space	Doctor's Office	Café
2	0	BEGOÑA	Bar	Coffee Shop	Restaurant	Playground	Plaza	Optical Shop	Café	Gift Shop	Snack Place	Salon / Barbershop
3	1	DEUSTU	Bar	Bank	Residential Building (Apartment / Condo)	Café	Breakfast Spot	Shopping Mall	Spanish Restaurant	Plaza	Coffee Shop	Automotive Shop
4	0	ERREKALDE	Bar	Bakery	Butcher	Other Great Outdoors	Grocery Store	Convenience Store	Restaurant	Breakfast Spot	Residential Building (Apartment / Condo)	Spanish Restaurant

7. Explore the final clusters

Finally, here's a map using Folium library that represent these 5 clusters

8. Results

The results of the exploratory data analysis and clustering are summarized below:

- Except Cluster 3 (Basurtu) and Cluster 4 (Abando) , the most common venue is bar, and this is not a surprise due that Bilbao is famous for its high concentration of outstanding restaurants, alongside its renowned pintxos bars.

- The average 2st most common venue are banks and this is also an evidence due that the GDP in the Basque Country is 24,1 % above the spanish average.

- Cluster 4 (Abando) is the commercial district full of clothing stores

	Neighborhood	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
5	ABANDO	Women's Store	Clothing Store	Bank	Restaurant	Boutique	Dessert Shop	Hotel	Accessories Store	Hotel Bar	Café

- The 10th most common venue in Cluster 3 (Basurtu), Cluster 4 (Abando) and Cluster 5 (Uribarri) are Café's and this is reamarkable due that are clusters almost in the city centre. Fort he next capstone, should be interesting the idea of establish a Café in one of these districts.

9. Conclusion

Finally to close this project, i need to express how much i've learned of how real life data-science projects look like. I really learnt things along all these months of hard work that i'd never have imagined; use Python libraries, Folium for maps, Foursquare API for explore venues...and quite a lot more. I'd for sure continue learning data science because this is not the future, is the present for all of us.