



Valencia data by districts to begin a business

By Luis Enrique Palma



Table of Contents

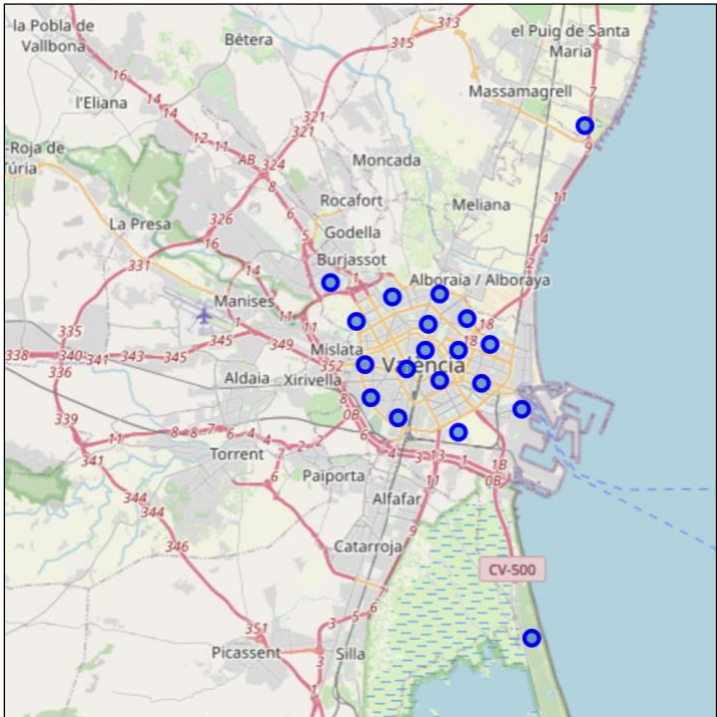
1. INTRODUCTORY SECTION	3
2. METODOLOGY SECTION.....	4
2.1. FIRTST STAGE: DATA MINNING	4
2.2. SECOND STAGE: DATA PROCESSING	6
2.2.1 <i>Analyzing the data</i>	6
2.2.2 <i>Data visualization</i>	7
2.2.3 <i>Clustering the venues</i>	11
3. RESULTS SECTION.....	13
3.1 FOURSQUARE TO FIND THE MOST POPULAR VENUES IN VALENCIAN DISTRICTS	13
4. DISCUSSION SECTION	14
5. CONCLUSIONS.....	14
6. REFERENCES.....	15

1. Introductory Section

Valencia is the third largest city in Spain, it is the capital of the province of Valencia which is located in the Valencian Community, this report will show relevant demographic data, analysis of interest and an ideal tool that will serve as support for those interested in starting a business in the city.

Valencia is located in the east of Spain, facing the Mediterranean Sea, it is crossed by the Turia river, the greatest contribution to its economy is by tourism and services, on a smaller scale there are industrial, construction and agriculture. Valencia capital is made up of 19 districts, which will be presented in this report according to the average characteristics of its habitants to have an adequate idea of what type of business will be more worth starting in each of the districts.

Table 1. Valencian Districts

Valencia Districts	Map
Ciutat Vella l'Eixample Extramurs Campanar la Saïdia el Pla del Real l'Olivereta Patraix Jesús Quatre Carreres Poblat Marítims Camins al Grau Algirós Benimaclet Rascanya Benicalap Pobles del Nord Pobles de l'Oest Pobles del Sud	 <p>The map shows the city of Valencia and its surrounding districts. The city is divided into several districts, each marked with a blue dot. The districts are: Ciutat Vella, l'Eixample, Extramurs, Campanar, la Saïdia, el Pla del Real, l'Olivereta, Patraix, Jesús, Quatre Carreres, Poblat Marítims, Camins al Grau, Algirós, Benimaclet, Rascanya, Benicalap, Pobles del Nord, Pobles de l'Oest, and Pobles del Sud. The map also shows the Turia river, the Mediterranean Sea, and various roads and highways. The CV-500 highway is highlighted in red.</p>

Source: self made

2. Metodology Section

The methodology is divided into two stages, the first consists of collecting relevant and current quantitative information with which to achieve greater precision in the estimates suggested or made with the help of the results. The second stage consists of data processing with the use of MS Excel, MS Word and Python with its different libraries to achieve the results.

2.1. Firts Stage: Data Minning

Demographic and descriptive information was collected of the year 2020 of the inhabitants of Valencia by their districts through the official page of the Valencia City Council, Gross average income of 2019, from the official page of the Institute of Statistics of Spain, coordinates of their districts in Google Maps, information of current most popular businesses on Foursquare and other qualitative information on Wikipedia (not being scientific).

The Data Frame (DF) was built in MS Excel from the tables collected in the sources mentioned in the previous paragraph and the following table resulted:

Table 2. Initial Data frame

District	Area (he)	Pop/km2	Population	Men population	Women population	Age 0-19	Age 20-39	Age 40-59	Age 60 greater	Foreign s	Foreigns EU	Foreigns Rest Europe	Foreigns Africa	Foreigns North Central America	Foreign South America	Foreign Asia	Foreign Other	Unemployed	Tourism	Bus	Truck	Tractor	Trailer	Motorcycle	Moped	Average Gross income 2018
Ciutat Vella	169	16,225	27,418	13,034	14,384	4,362	7,116	8,582	7,358	4,784	2,316	292	452	518	771	427	8	1,640	14,073	526	1,883	484	274	3,504	745	25064
l'Eixample	173.3	24,709	42,826	19,769	23,057	7,822	9,379	13,135	12,490	4,675	2,178	244	303	589	926	422	13	2,333	19,745	2	1,171	447	387	4,375	1,011	26142
Extramurs	197.2	24,789	48,877	22,440	26,437	8,380	11,258	14,659	14,580	5,705	2,018	274	341	654	1,557	853	8	3,140	21,073	73	1,124	311	305	4,191	970	20247
Campanar	523.8	7,395	38,736	18,370	20,366	7,868	8,620	11,775	10,473	5,178	1,666	523	507	430	1,309	740	3	2,485	17,422	90	725	213	158	2,866	716	17988
la Saïdia	194.4	24,431	47,491	21,996	25,495	7,815	11,570	14,131	13,975	7,222	2,027	689	760	631	2,216	894	5	3,793	19,351	5	1,207	223	271	3,190	1,177	14429
el Pla del Real	169.3	18,101	30,644	14,098	16,546	5,979	7,350	8,208	9,107	2,898	1,183	127	154	248	729	449	8	1,433	15,167	9	716	249	202	3,282	765	26930
l'Olivereta	198.9	24,764	49,250	23,386	25,864	8,342	11,783	15,357	13,768	8,935	1,725	634	1,038	772	2,738	2,024	4	4,485	19,989	31	1,203	176	204	2,830	1,144	12427
Patraix	287.3	20,209	58,053	27,461	30,592	10,235	13,030	18,523	16,265	5,857	1,378	430	588	472	1,815	1,170	4	4,354	27,448	11	1,624	605	334	4,272	1,282	14907
Jesús	298.5	17,730	52,917	25,679	27,238	8,946	12,686	16,697	14,588	7,990	1,554	423	1,119	705	2,586	1,600	3	4,516	22,696	23	1,403	245	243	3,689	1,461	13065
Quatre Carreres	1,132.50	6,580	74,518	35,749	38,769	13,683	17,255	23,742	19,838	10,707	2,989	702	930	1,043	3,143	1,878	22	6,741	32,231	9	2,036	819	1,049	5,317	2,286	13704
Poblat Marítim	396.7	14,049	55,725	26,841	28,884	10,073	13,093	17,479	15,080	7,917	3,105	598	902	534	1,801	969	8	6,194	23,796	26	1,568	985	1,291	3,991	2,058	13064
Camins al Grau	236.7	27,834	65,890	31,667	34,223	13,084	15,274	21,546	15,986	10,070	2,954	766	870	585	2,534	2,351	10	5,151	26,504	38	1,614	539	565	4,790	1,563	15962
Algirós	295.9	12,388	36,657	17,250	19,407	5,509	9,372	9,993	11,783	4,269	1,211	231	384	274	1,221	947	1	2,626	16,895	8	643	129	131	2,755	868	17502
Benimaclet	157	18,268	28,686	13,381	15,305	4,456	7,695	8,049	8,486	3,391	1,090	165	303	390	850	589	4	1,981	12,576	0	500	126	112	2,051	723	17122
Rascanya	262.9	20,631	54,233	26,412	27,821	11,105	13,138	17,374	12,616	9,500	1,722	697	1,889	837	2,868	1,484	3	5,177	21,361	2	1,220	225	187	3,311	1,507	12660
Benicalap	221.6	21,401	47,421	22,947	24,474	9,515	11,535	15,242	11,129	6,960	1,531	562	988	544	2,091	1,243	1	4,015	19,719	9	1,105	298	235	2,873	1,270	12761
Pobles del Nord	1,519.60	433	6,582	3,179	3,403	1,333	1,334	2,102	1,813	472	185	44	79	56	94	11	3	396	3,452	0	315	189	132	668	301	17773
Pobles de l'Oest	201.1	7,250	14,581	7,078	7,503	2,916	3,450	4,617	3,598	1,865	589	199	203	159	583	131	1	1,374	6,661	0	529	94	70	995	443	11576
Pobles del Sud	3,242.60	649	21,040	10,380	10,660	3,853	4,874	6,605	5,708	2,277	961	82	239	187	619	188	1	1,721	10,858	34	960	546	616	1,549	680	13576

Source: self-made

2.2. Second Stage: Data Processing

To process the data, the previous DF will be imported in Jupyter labs and with the use of Python language and its libraries such as pandas for Data frame manipulation ; the folium library for the creation of interactive, informative and choropletic maps; the json library for the extraction of coordinates through the BeautifulSoup library, geocoder, geopy and geopandas; the seaborn and matplotlib libraries for creating frequency and other statistical data graphs; the sklearn library for the K Nearest Neighbors implementation of the most popular sites in their districts collected by Foursquare developer account and other libraries will be included to process or complement the information.

2.2.1 Analyzing the data

First of all, the data was cleaned and the elements of number columns that was in string type were converted to float type, then the data was analyzed with info and describe methods to see a type and statistic summary respectively. See figure (1) and (2):

```
RangeIndex: 19 entries, 0 to 18
Data columns (total 28 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   District                              19 non-null     object
1   Area (he)                             19 non-null     float64
2   Pop/km2                               19 non-null     float64
3   Population                             19 non-null     float64
4   Men population                         19 non-null     float64
5   Women population                       19 non-null     float64
6   Age 0-19                              19 non-null     float64
7   Age 20-39                             19 non-null     float64
8   Age 40-59                             19 non-null     float64
9   Age 60 greater                         19 non-null     float64
10  Foreigns                              19 non-null     float64
11  Foreigns EU                           19 non-null     float64
12  Foreigns Rest Europe                  19 non-null     float64
13  Foreigns Africa                       19 non-null     float64
14  Foreigns North Central America        19 non-null     float64
15  Foreign South America                 19 non-null     float64
16  Foreign Asia                          19 non-null     float64
17  Foreign Other                         19 non-null     float64
18  Unemployed                            19 non-null     float64
19  Tourism                               19 non-null     float64
20  Bus                                    19 non-null     float64
21  Truck                                 19 non-null     float64
22  Tractor                               19 non-null     float64
23  Trailer                               19 non-null     float64
24  Motorcycle                            19 non-null     float64
25  Moped                                 19 non-null     float64
26  Average Gross income 2018             19 non-null     float64
27  DISTRICT                              19 non-null     object
dtypes: float64(26), object(2)
```

Figure 1. Dtype product of .info method

Table 3. Statistic summary snipped table

	Area (he)	Pop/km2	Population	Men population	Women population	Age 0-19	Age 20-39	Age 40-59	Age 60 greater
count	19.000000	19.000000	19.000000	19.000000	19.000000	19.000000	19.000000	19.000000	19.000000
mean	519.910526	16201.894737	42186.578947	20058.789474	22127.789474	7646.105263	9990.105263	13042.947368	11507.421053
std	749.226919	8351.209160	17489.813109	8399.449179	9108.808488	3348.386260	4037.707998	5743.886368	4599.216555
min	157.000000	433.000000	6582.000000	3179.000000	3403.000000	1333.000000	1334.000000	2102.000000	1813.000000
25%	195.800000	9891.500000	29665.000000	13739.500000	15925.500000	4982.500000	7522.500000	8395.000000	8796.500000
50%	236.700000	18101.000000	47421.000000	21996.000000	24474.000000	7868.000000	11258.000000	14131.000000	12490.000000
75%	347.600000	22916.000000	53575.000000	26045.500000	27529.500000	9794.000000	12858.000000	17035.500000	14584.000000
max	3242.600000	27834.000000	74518.000000	35749.000000	38769.000000	13683.000000	17255.000000	23742.000000	19838.000000

Foreigns	Foreigns EU	Foreigns Rest Europe	Foreigns Africa	Foreigns North Central America	Foreign South America	Foreign Asia	Foreign Other	Unemployed	Tourism
19.000000	19.000000	19.000000	19.000000	19.000000	19.000000	19.000000	19.000000	19.000000	19.000000
5824.842105	1704.315789	404.315789	634.157895	506.736842	1602.684211	966.842105	5.789474	3345.000000	18474.578947
2925.109158	786.472579	237.082980	449.001641	250.194778	906.544419	664.067957	5.191649	1772.800923	7046.036256
472.000000	185.000000	44.000000	79.000000	56.000000	94.000000	11.000000	1.000000	396.000000	3452.000000
3830.000000	1197.000000	215.000000	303.000000	332.000000	810.500000	438.000000	3.000000	1851.000000	14620.000000
5705.000000	1666.000000	423.000000	507.000000	534.000000	1557.000000	894.000000	4.000000	3140.000000	19719.000000
7953.500000	2102.500000	616.000000	916.000000	642.500000	2375.000000	1363.500000	8.000000	4500.500000	22028.500000
10707.000000	3105.000000	766.000000	1889.000000	1043.000000	3143.000000	2351.000000	22.000000	6741.000000	32231.000000

Bus	Truck	Tractor	Trailer	Motorcycle	Moped	Average Gross income 2018
19.000000	19.000000	19.000000	19.000000	19.000000	19.000000	19.000000
47.157895	1134.000000	363.315789	356.105263	3184.157895	1103.684211	16678.894737
118.574526	481.763081	245.296658	322.135047	1233.011997	514.093599	4774.948166
0.000000	315.000000	94.000000	70.000000	668.000000	301.000000	11576.000000
3.500000	720.500000	201.000000	172.500000	2792.500000	734.000000	13064.500000
9.000000	1171.000000	249.000000	243.000000	3282.000000	1011.000000	14907.000000
32.500000	1485.500000	511.500000	360.500000	4091.000000	1371.500000	17880.500000
526.000000	2036.000000	985.000000	1291.000000	5317.000000	2286.000000	26930.000000

These three last tables are all the same, this is very important to have an idea of the behavior of the data recollected and how it describes the generality of the habitants of Valencia.

2.2.2 Data visualization

As we know the table data can be hard to understand, in this unit some plots will be shown to have a better familiarity of the data resulted from the Valencian Districts.

The first plot is the population, there are to plots to know which district have more density and Population per area.

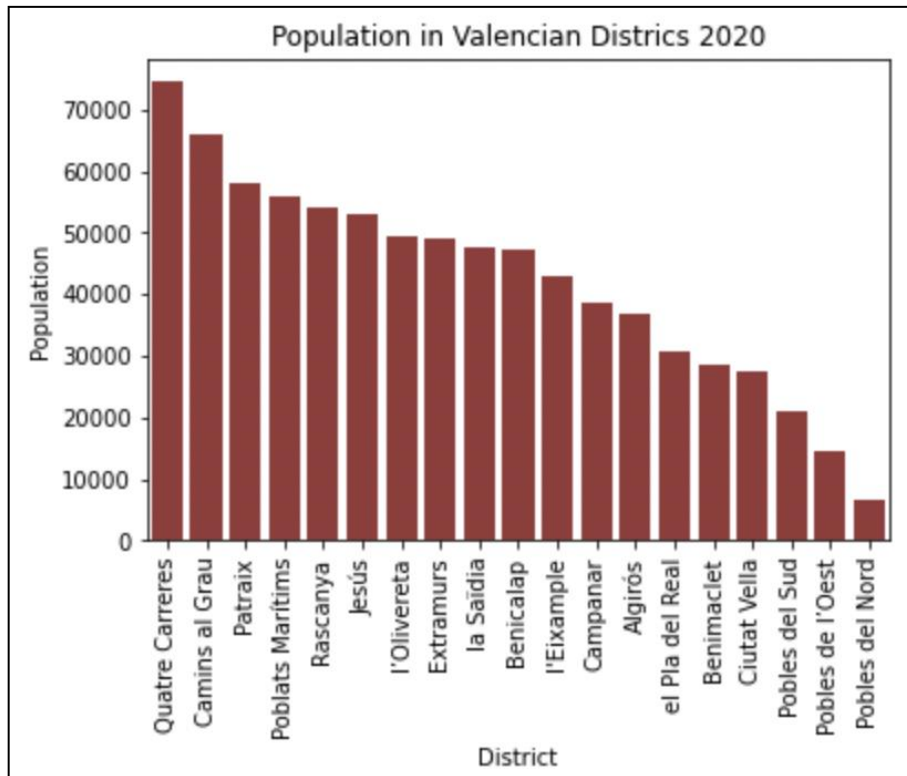


Figure 2. Population in Valencian Districts 2020

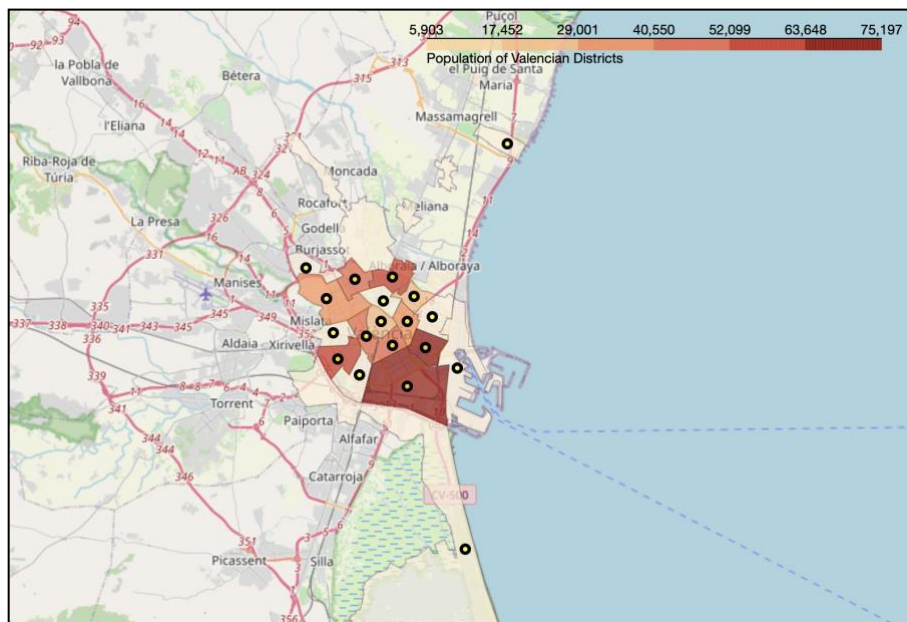


Figure 3. Population of Valencian Districts in map

Now we know that Districts named: Quatre Carreres, Camins al Grau, Patraix, Poblats Maritims and Rascanya are the most populated, there are some businesses that need lots of people like supermarkets, fast food restaurants, schools and other. Let's look what districts have the most and less gross income.

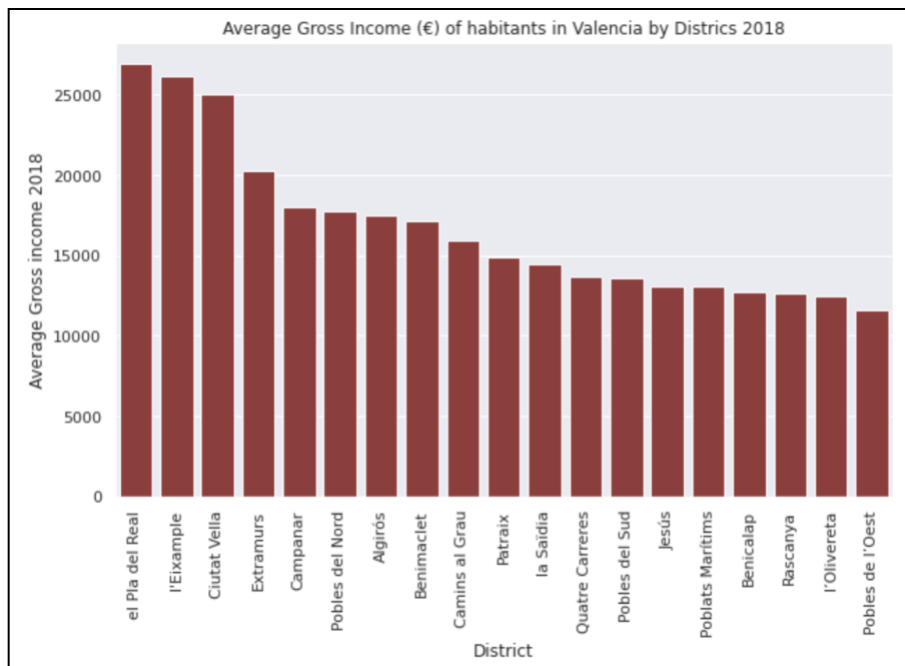


Figure 4. Average gross income of habitants in Valencia by Districts 2018

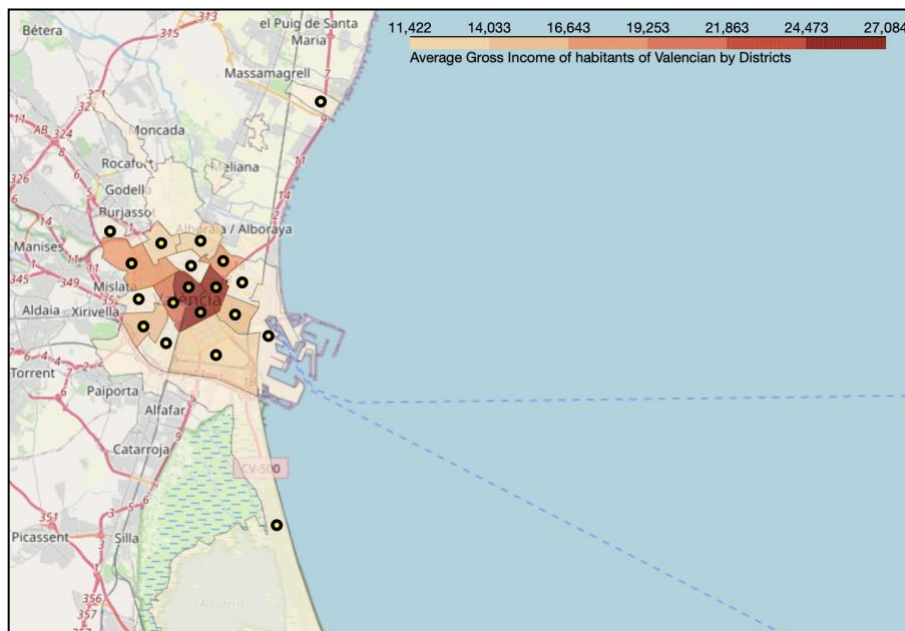


Figure 5. Average gross income of habitants in Valencia by District 2018 map

Now we know that Districts named: el Pla del Real, l'Eixample and Ciutat Vella are the district which habitants have the greatest economic power in Valencia, as we can see they tend to be at the center of city. This can be relevant to begin deluxe business like gourmet restaurants, spa (it could be better near the beach), wine store and other. There could be some business related to vehicles, in the next plots, we can see the districts with the most vehicle fleet. It's divided in vehicle and heavy vehicle fleet.

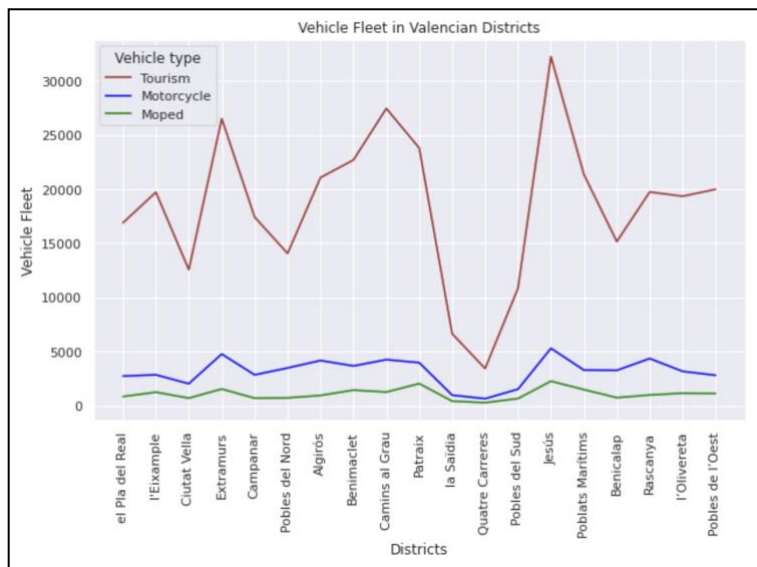


Figure 6. Vehicle fleet in Valencian Districts

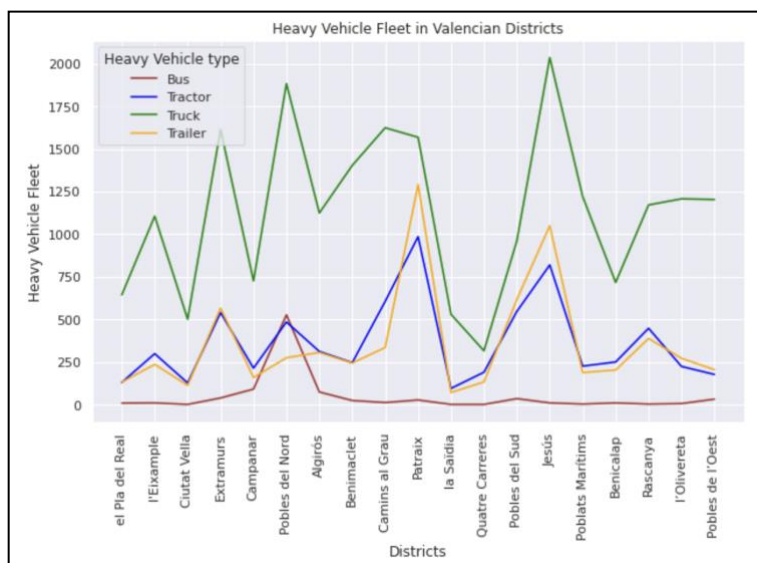


Figure 7. Heavy Vehicle fleet in Valencian Districts

There are a lot of plots that could be shown, but I think there are the most relevant to see, to know some of Valencian districts inhabitants.

2.2.3 Clustering the venues

Foursquare have a venue database, which helps to find the most popular venues in a city, using the Foursquare API, I set a Radius of 750, and limit to 100 venues. We obtain a table with 597 rows like we see in Table 4

Table 4. Snipped table of Venues obtained from Foursquare API

	District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Ciutat Vella	39.474652	-0.375633	Colmado Lalola	39.475241	-0.375624	Tapas Restaurant
1	Ciutat Vella	39.474652	-0.375633	San Tommaso	39.475282	-0.377144	Italian Restaurant
2	Ciutat Vella	39.474652	-0.375633	Catedral de Valencia	39.475724	-0.375005	Church
3	Ciutat Vella	39.474652	-0.375633	Plaça de la Reina	39.474792	-0.375501	Plaza
4	Ciutat Vella	39.474652	-0.375633	Véneta Food & Gelato Italiano	39.475332	-0.375657	Ice Cream Shop
5	Ciutat Vella	39.474652	-0.375633	Gelateria La Romana	39.473107	-0.376200	Ice Cream Shop
6	Ciutat Vella	39.474652	-0.375633	La Nocciola Toscana	39.474042	-0.374516	Ice Cream Shop
7	Ciutat Vella	39.474652	-0.375633	Oslo - The Vegetarian	39.475949	-0.377397	Vegetarian / Vegan Restaurant
8	Ciutat Vella	39.474652	-0.375633	Creperie Bretonne	39.475236	-0.375981	French Restaurant
9	Ciutat Vella	39.474652	-0.375633	La Pappardella	39.475327	-0.376086	Italian Restaurant

To cluster the common venues, I will use K-mean, an unsupervised learning algorithm, starting with 2 k-clusters and using the elbow method to kind the optimal k-cluster, we'll have the graphic in figure 8. It helps us to corroborate the correct assumed k-cluster.

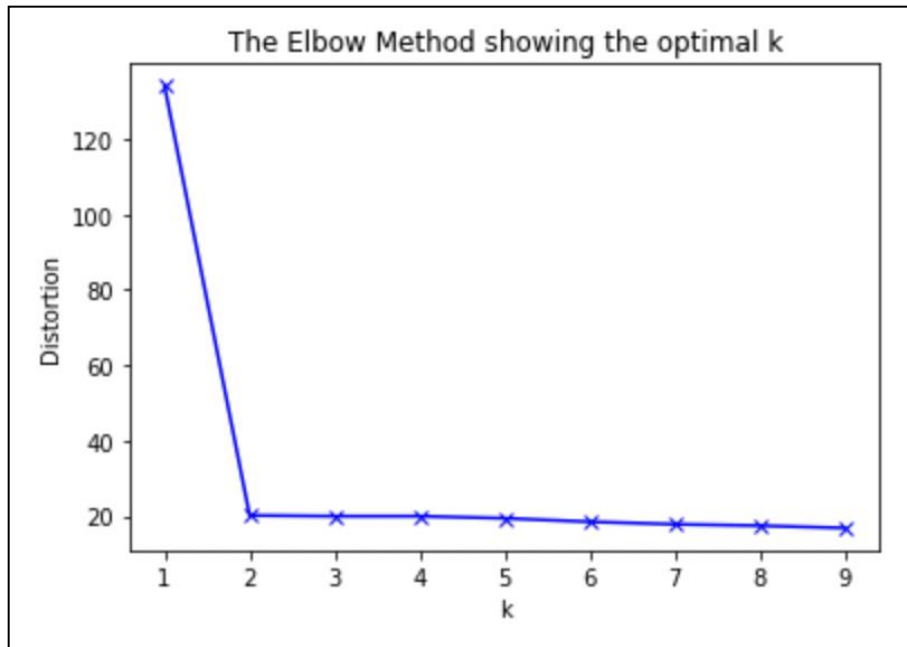


Figure 8. Elbow method to find optimal k-clusters

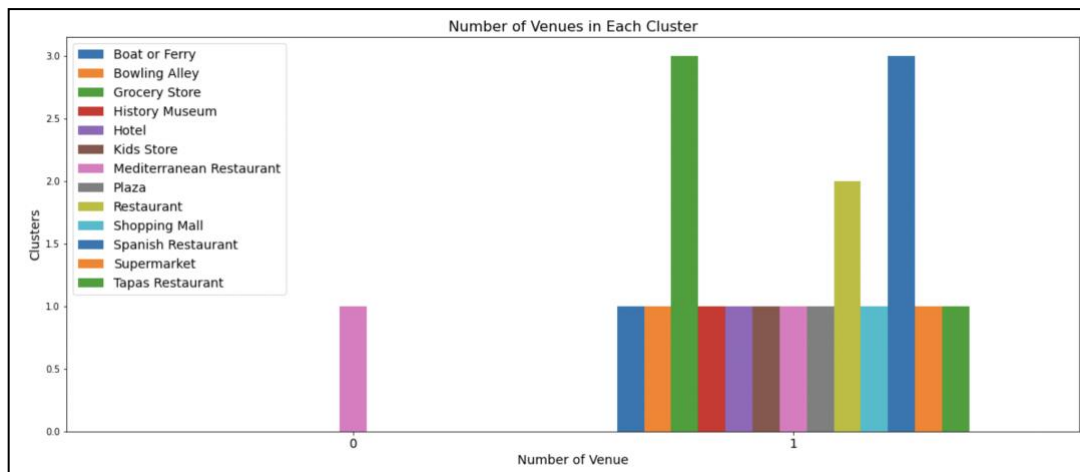


Figure 9.. Number of venues in each cluster

3. Results Section

3.1 Foursquare to find the most popular venues in Valencian Districts

Now we can obtain Table 4, it gives an idea of what is working well per district. A particular visible characteristic is that all the districts less Pobles de l'Oest and l'Olivereta have restaurants as most popular venue in its district.

Table 5. Most popular type venues in Valencian Districts

	District	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Algirós	Spanish Restaurant	Bar	Tapas Restaurant	Italian Restaurant	Supermarket
1	Benicalap	Shopping Mall	Grocery Store	Brewery	Restaurant	Pool
2	Benimaclet	Tapas Restaurant	Café	Bar	Pizza Place	Cocktail Bar
3	Camins al Grau	Hotel	Seafood Restaurant	Pub	Supermarket	Tapas Restaurant
4	Campanar	Grocery Store	Gas Station	Restaurant	Pool	Event Space
5	Ciutat Vella	Plaza	Tapas Restaurant	Spanish Restaurant	Ice Cream Shop	Hotel
6	Extramurs	Spanish Restaurant	Bakery	Tapas Restaurant	Plaza	Hotel
7	Jesús	Restaurant	Grocery Store	Coffee Shop	Supermarket	Performing Arts Venue
8	Patraix	Supermarket	Bakery	Miscellaneous Shop	Spanish Restaurant	Fast Food Restaurant
9	Poblets Marítims	Boat or Ferry	Athletics & Sports	Cruise Ship	Southern / Soul Food Restaurant	Spanish Restaurant
10	Pobles de l'Oest	Grocery Store	Hotel	Tapas Restaurant	Metro Station	Park
11	Pobles del Nord	Kids Store	Asian Restaurant	Wine Bar	Frozen Yogurt Shop	French Restaurant
12	Pobles del Sud	Mediterranean Restaurant	Nudist Beach	Falafel Restaurant	Frozen Yogurt Shop	French Restaurant
13	Quatre Carreres	Mediterranean Restaurant	Dive Bar	Restaurant	Auto Workshop	Playground
14	Rascanya	History Museum	Tapas Restaurant	Café	Soccer Stadium	Spanish Restaurant
15	el Pla del Real	Spanish Restaurant	Pub	Café	Tapas Restaurant	Breakfast Spot
16	la Saïdia	Grocery Store	Tapas Restaurant	Bakery	Spanish Restaurant	Café
17	l'Eixample	Restaurant	Spanish Restaurant	Mediterranean Restaurant	Italian Restaurant	Tapas Restaurant
18	l'Olivereta	Bowling Alley	Bakery	Metro Station	Plaza	Cafeteria

4. Discussion Section

Valencia as I mentioned at the beginning of this report, most economic activity is the service: tourism as well, which include restaurant, bars, plazas, beaches. It can be verifiable in the result plots and tables.

The clustering algorithm used was k-mean, starting with k-cluster 2, and confirmed with the elbow method.

The habitants with the greatest economic power use to live in the center districts, where are the historical places, it could be considered some luxor restaurants to take advantages of the lots of people doing tourism and the rich neighbors.

5. Conclusions

Valencia's most principles existing businesses are the restaurants, bars, cafeterias, plazas. There are a lot of national and international tourist that every year goes to enjoy the city. I recommend adding some qualitative data to future analysis.

6. References

- Valencia Town Hall*. (December de 2020). Obtenido de Ajuntament de Valencia:
<https://www.valencia.es/es/cas/estadistica/mapa-barrios>
- Enterat. (2021). *Enterat*. Obtenido de <https://www.enterat.com/actualidad/habitantes-valencia.php>
- INE. (2020). *Instituto Nacional de Estadística*. Obtenido de <https://ine.es>
- Wikipedia. (2021). *Wikipedia.org*. Obtenido de <https://es.wikipedia.org/wiki/Valencia>
- Foursquare. (2021). *Foursquare*. Obtenido de <https://foursquare.com>
- Google. (2021). *Google maps*. Obtenido de <https://www.google.com/maps/place/Valencia,+Espa%C3%B1a/>
- Espana, D. d. (2021). *Datos de gobierno de Espana*. Obtenido de <http://mapas.valencia.es/lanzadera/opendata/Distritos/JSON?srsName=EPSG%3A4326>