



# NEW VETERINARY CLINIC IN MADRID

The Battle of the Neighbourhoods  
Applied Data Science Capstone by IBM/Coursera

Eva Porras

Madrid, September 2019

# Business Problem

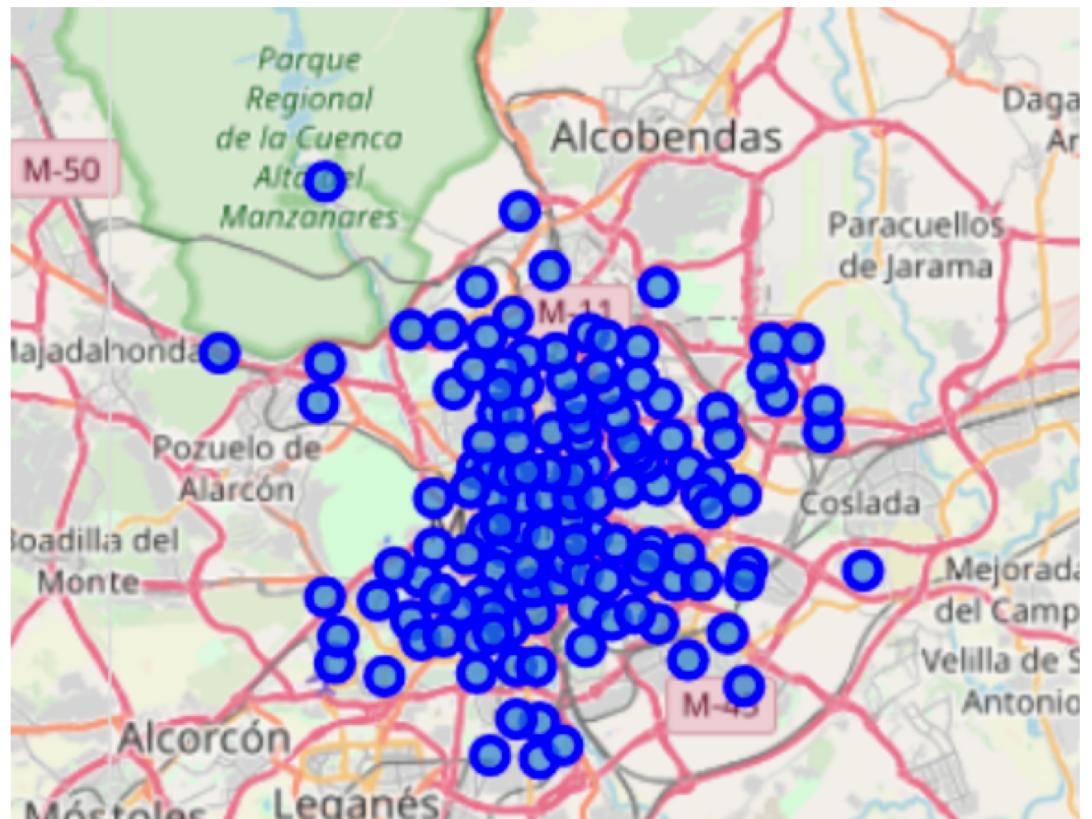
- Find an optimal location for a new veterinarian clinic in Madrid, Spain.
- Analysis of the number of pets and veterinarians per neighbourhood to detect those neighborhoods with the highest ratios of pets per veterinarian.
- Analysis of the ascending trend of the number of pets in Madrid in the last five years and analysis those boroughs with highest increase ratios.
- With these two criteria, we will cluster the neighbourhoods to detect the most promising neighborhoods to install a new veterinary clinic (and the most "saturated" neighbourhood to avoid).



# Data (1/4)

Following data sources will be needed to extract/generate the required information:

- name and location of the boroughs and neighbourhoods of Madrid: City Hall Public Data web publishes the street guide including numbering of all urban premises (206866 premises). We will extract the name of the boroughs, neighbourhoods and their location.



Madrid has 21 boroughs and 131 neighbourhoods.

## Data (2/4)

- number of pets in the last five years: City Hall Public Data web publishes the number of dogs and cats per borough in the last 5 years. We will extract the increase of the number of total pets per borough. On the other hand, since the number of pets is published per borough, we use the proportion of population to distribute the number of pets per neighbourhood.

There are 370,018 registered pets in Madrid.

The increase of the number of pets in the last 5 years in Madrid is 22.0%.

	Neighbourhood	Borough	Longitude	Latitude	Total_Pets_Neighbourhood	Inc_5_y (%)
0	PALACIO	CENTRO	-3.711270	40.414430	4055.0	43.7
1	EMBAJADORES	CENTRO	-3.702885	40.409738	8070.0	43.7
2	CORTES	CENTRO	-3.697530	40.414064	1857.0	43.7
3	JUSTICIA	CENTRO	-3.697507	40.423651	3074.0	43.7
4	UNIVERSIDAD	CENTRO	-3.706039	40.425119	5711.0	43.7
5	SOL	CENTRO	-3.704920	40.416991	1300.0	43.7
6	IMPERIAL	ARGANZUELA	-3.718203	40.407663	2223.0	31.7
7	ACACIAS	ARGANZUELA	-3.706246	40.401900	3587.0	31.7
8	CHOPERA	ARGANZUELA	-3.698228	40.395881	1961.0	31.7
9	LEGAZPI	ARGANZUELA	-3.689785	40.388806	1920.0	31.7

## Data (3/4)

- number of veterinary in every neighborhood: City Hall Public Data web publishes the active venues in Madrid (163,251 venues). We will extract the number of veterinary clinics per neighbourhood.

	Neighbourhood	Number_of_Vets
0	GUINDALERA	13
1	PUEBLO NUEVO	10
2	CANILLAS	10
3	ALUCHE	9
4	PACIFICO	9
5	ARAVACA	7
6	VALVERDE	7
7	ACACIAS	7
8	PEÑA GRANDE	6
9	TRAFalgar	6

The total number of Veterinary Clinics in Madrid is 403

# Data (4/4)

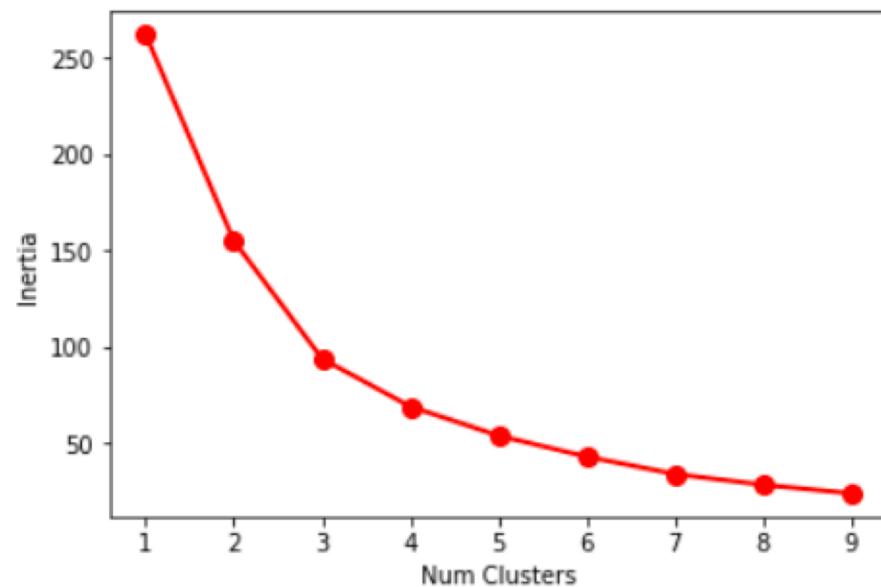
As final dataframe we have per each neighbourhood:

- Borough,
- Location: Latitude and Longitude,
- Total Number of pets,
- Increase of the number of pets in the last 5 years,
- Number of veterinarian clinics and
- Ratio pets/vet

	Neighbourhood	Borough	Longitude	Latitude	Total_Pets_Neighbourhood	Inc_5_y (%)	Number_of_Vets	pets/vet
0	PALACIO	CENTRO	-3.711270	40.414430	4055.0	43.7	6.0	676.0
1	EMBAJADORES	CENTRO	-3.702885	40.409738	8070.0	43.7	5.0	1614.0
2	CORTES	CENTRO	-3.697530	40.414064	1857.0	43.7	3.0	619.0
3	JUSTICIA	CENTRO	-3.697507	40.423651	3074.0	43.7	1.0	3074.0
4	UNIVERSIDAD	CENTRO	-3.706039	40.425119	5711.0	43.7	5.0	1142.0
5	SOL	CENTRO	-3.704920	40.416991	1300.0	43.7	1.0	1300.0
6	IMPERIAL	ARGANZUELA	-3.718203	40.407663	2223.0	31.7	3.0	741.0
7	ACACIAS	ARGANZUELA	-3.706246	40.401900	3587.0	31.7	7.0	512.0
8	CHOPERA	ARGANZUELA	-3.698228	40.395881	1961.0	31.7	3.0	654.0
9	LEGAZPI	ARGANZUELA	-3.689785	40.388806	1920.0	31.7	2.0	960.0

# Clustering Model

**Determination of the number of clusters K**  
by Elbow Method: the idea is compute k-means clustering for different values of k. For each k, calculate the inertia (sum of squared distances of samples to their closest cluster center) and plot the curve of inertia depending of the number of clusters k. The location of the elbow in the plot is considered as the appropriate number of clusters.



According to the plot, the optimal number of clusters is set in 4

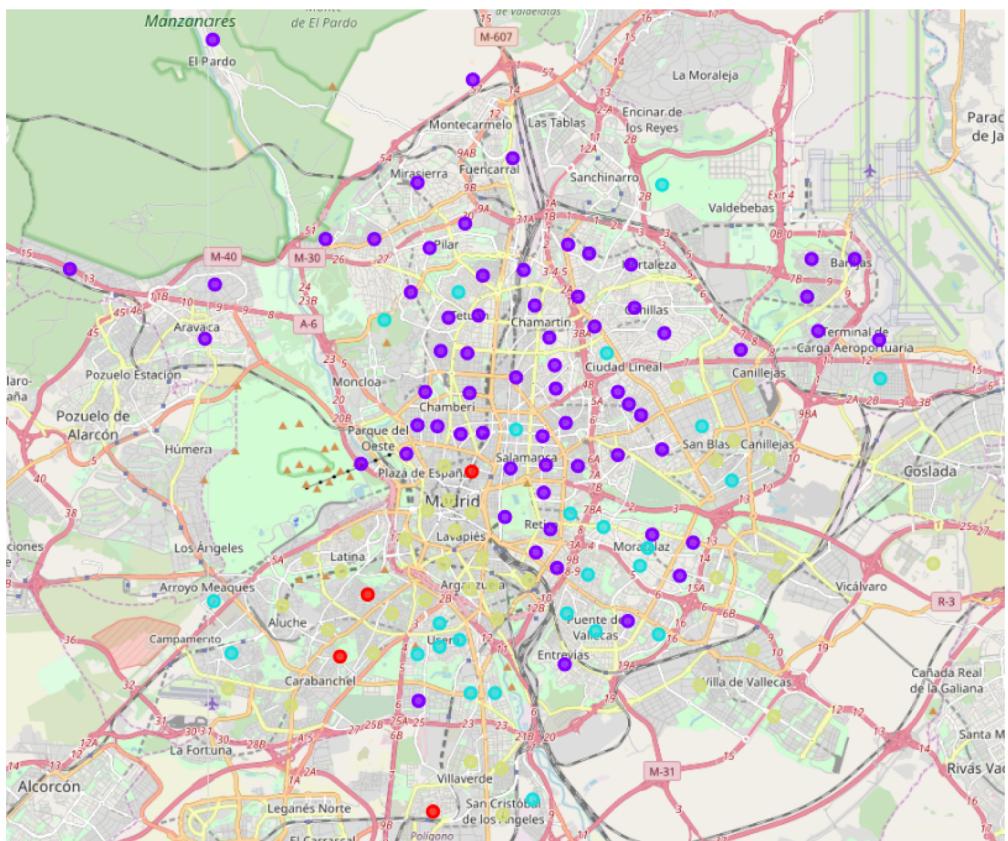
# Clustering Model Results

**Cluster 0:** The 4 neighbourhoods included in the cluster 0 are located in boroughs with a high increase in the number of pets in the last 5 years and with a high ratio pets/vets.

**Cluster 1:** The 65 neighbourhoods included in the cluster 1 are located in boroughs with a low increase in the number of pets in the last 5 years and with very low ratio pets/vets.

**Cluster 2:** The 25 neighbourhoods included in the cluster 2 are located in boroughs with a low increase in the number of pets in the last 5 years but with high ratio pets/vets.

**Cluster 3:** The 37 neighbourhoods included in cluster 3 are located in boroughs with a high increase in the number of pets in the last 5 years but with a low ratio pets/vets.



# Clustering Model Results

## Cluster 0:

	Neighbourhood	Total_Pets_Neighbourhood	Inc_5_y (%)	Number_of_Vets	pets/vet
3	JUSTICIA	3074.0	43.7	1.0	3074.0
67	SAN ISIDRO	3966.0	29.7	1.0	3966.0
69	PUERTA BONITA	3555.0	29.7	1.0	3555.0
106	VILLAVERDE ALTO C.H.	4550.0	32.6	0.0	4550.0

## Cluster 1:

	Neighbourhood	Total_Pets_Neighbourhood	Inc_5_y (%)	Number_of_Vets	pets/vet
13	PACIFICO	3015.0	6.0	9.0	335.0
14	ADELFA'S	1680.0	6.0	3.0	560.0
16	IBIZA	1947.0	6.0	3.0	649.0
17	LOS JERONIMOS	626.0	6.0	2.0	313.0
18	NIÑO JESUS	1380.0	6.0	2.0	690.0
19	RECOLETOS	1907.0	21.0	3.0	636.0
20	GOYA	3620.0	21.0	5.0	724.0
21	FUENTE DEL BERRO	2572.0	21.0	4.0	643.0
22	GUINDALERA	5100.0	21.0	13.0	392.0
23	LISTA	2563.0	21.0	4.0	638.0
25	EL VISO	1778.0	1.9	2.0	889.0
26	PROSPERIDAD	3755.0	1.9	4.0	939.0
27	CIUDAD JARDIN	1926.0	1.9	5.0	385.0
28	HISPANOAMERICA	3263.0	1.9	5.0	653.0
29	NUEVA ESPAÑA	2548.0	1.9	2.0	1274.0
30	CASTILLA	1747.0	1.9	3.0	582.0
31	BELLAS VISTAS	3190.0	19.1	3.0	1063.0
32	CUATRO CAMINOS	3761.0	19.1	6.0	627.0
33	CASTILLEJOS	2233.0	19.1	4.0	558.0
34	ALMENARA	2464.0	19.1	4.0	616.0
36	BERRUQUETE	2750.0	19.1	4.0	688.0

37	GAZTAMBIDE	2925.0	12.2	3.0	975.0
38	ARAPILES	3106.0	12.2	5.0	621.0
39	TRAFalgar	3129.0	12.2	6.0	522.0
40	ALMAGRO	2501.0	12.2	3.0	834.0
41	RIOS ROSAS	3468.0	12.2	5.0	694.0
42	VALLEHERMOSO	2573.0	12.2	3.0	858.0
43	EL PARDO	325.0	11.4	0.0	325.0
44	FUENTELARREINA	308.0	11.4	1.0	308.0
45	PEÑA GRANDE	4187.0	11.4	6.0	698.0
46	EL PILAR	4361.0	11.4	5.0	872.0
47	LA PAZ	3138.0	11.4	4.0	784.0
48	VALVERDE	5906.0	11.4	7.0	857.0
49	MIRASIERRA	3134.0	11.4	4.0	784.0
50	EL GOLOSO	1754.0	11.4	2.0	877.0
51	CASA DE CAMPO	1733.0	14.1	4.0	433.0
52	ARGUELLES	3203.0	14.1	3.0	1068.0
54	VALDEZARZA	3974.0	14.1	5.0	795.0
55	VALDEMARIN	907.0	14.1	0.0	907.0
56	EL PLANTIO	368.0	14.1	0.0	368.0
57	ARAVACA	3597.0	14.1	7.0	514.0
72	ORCASITAS	2417.0	21.2	3.0	806.0
79	ENTREVIAS	4225.0	18.8	4.0	1056.0

83	PORAZGO	3451.0	18.8	3.0	1150.0
85	PAVONES	887.0	15.9	0.0	887.0
86	HORCAJO	613.0	15.9	0.0	613.0
87	MARROQUINA	2602.0	15.9	6.0	434.0
91	VENTAS	5479.0	19.2	5.0	1096.0
92	PUEBLO NUEVO	7083.0	19.2	10.0	708.0
93	QUINTANA	2815.0	19.2	3.0	938.0
94	CONCEPCION	2367.0	19.2	5.0	473.0
95	SAN PASCUAL	2039.0	19.2	3.0	680.0
97	COLINA	727.0	19.2	1.0	727.0
98	ATALAYA	176.0	19.2	0.0	176.0
99	COSTILLARES	2497.0	19.2	4.0	624.0
100	PALOMAS	848.0	11.9	0.0	848.0
101	PIOVERA	1857.0	11.9	4.0	464.0
102	CANILLAS	5056.0	11.9	10.0	506.0
103	PINAR DEL REY	6584.0	11.9	5.0	1317.0
104	APOSTOL SANTIAGO	1926.0	11.9	3.0	642.0
126	ALAMEDA DE OSUNA	2579.0	10.4	4.0	645.0
127	AEROPUERTO	252.0	10.4	0.0	252.0
128	CASCO H.BARAJAS	1004.0	10.4	2.0	502.0
129	TIMON	1660.0	10.4	3.0	553.0
130	CORRALEJOS	1009.0	10.4	1.0	1009.0

# Clustering Model Results

## Cluster 2:

	Neighbourhood	Total_Pets_Neighbourhood	Inc_5_y (%)	Number_of_Vets	pets/vet
15	ESTRELLA	2081.0	6.0	1.0	2081.0
24	CASTELLANA	2080.0	21.0	1.0	2080.0
35	VALDEACEDERAS	2843.0	19.1	0.0	2843.0
53	CIUDAD UNIVERSITARIA	2134.0	14.1	1.0	2134.0
62	CAMPAMENTO	2232.0	32.1	1.0	2232.0
64	LAS AGUILAS	5887.0	32.1	3.0	1962.0
73	ORCASUR	1473.0	21.2	1.0	1473.0
74	SAN FERMIN	2514.0	21.2	1.0	2514.0
75	ALMENDRALES	2299.0	21.2	1.0	2299.0
76	MOSCARDO	2831.0	21.2	2.0	1416.0
77	ZOFIO	1501.0	21.2	1.0	1501.0
78	PRADOLONGO	1858.0	21.2	1.0	1858.0
80	SAN DIEGO	5180.0	18.8	3.0	1727.0
81	PALOMERAS BAJAS	4879.0	18.8	2.0	2440.0
82	PALOMERAS SURESTE	5164.0	18.8	3.0	1721.0
84	NUMANCIA	5765.0	18.8	2.0	2882.0
88	MEDIA LEGUA	1720.0	15.9	0.0	1720.0
89	FONTARRON	1665.0	15.9	1.0	1665.0
90	VINATEROS	1624.0	15.9	1.0	1624.0
96	SAN JUAN BAUTISTA	1419.0	19.2	0.0	1419.0
105	VALDEFUENTES	7490.0	11.9	3.0	2497.0
108	BUTARQUE	1923.0	32.6	1.0	1923.0
118	SIMANCAS	3306.0	25.7	2.0	1653.0
121	ARCOS	2858.0	25.7	2.0	1429.0
123	REJAS	1979.0	25.7	1.0	1979.0

## Cluster 3:

	Neighbourhood	Total_Pets_Neighbourhood	Inc_5_y (%)	Number_of_Vets	pets/vet
0	PALACIO	4055.0	43.7	6.0	676.0
1	EMBAJADORES	8070.0	43.7	5.0	1614.0
2	CORTES	1857.0	43.7	3.0	619.0
4	UNIVERSIDAD	5711.0	43.7	5.0	1142.0
5	SOL	1300.0	43.7	1.0	1300.0
6	IMPERIAL	2223.0	31.7	3.0	741.0
7	ACACIAS	3587.0	31.7	7.0	512.0
8	CHOPERA	1961.0	31.7	3.0	654.0
9	LEGAZPI	1920.0	31.7	2.0	960.0
10	DELICIAS	2733.0	31.7	2.0	1366.0
11	PALOS DE MOGWER	2540.0	31.7	5.0	508.0
12	ATOCHA	116.0	31.7	0.0	116.0
58	LOS CARMENES	1992.0	32.1	2.0	996.0
59	PUERTA DEL ANGEL	4765.0	32.1	5.0	953.0
60	LUCERO	4143.0	32.1	4.0	1036.0
61	ALUCHE	7590.0	32.1	9.0	843.0
63	CUATRO VIENTOS	664.0	32.1	1.0	664.0
65	COMILLAS	2278.0	29.7	3.0	759.0
66	OPAÑEL	3352.0	29.7	6.0	559.0
68	VISTA ALEGRE	4711.0	29.7	5.0	942.0
70	BUENAVENTURA	4783.0	29.7	5.0	957.0
71	ABRANTES	3144.0	29.7	3.0	1048.0
107	SAN CRISTOBAL	1643.0	32.6	0.0	1643.0
109	LOS ROSALES	3733.0	32.6	4.0	933.0
110	LOS ANGELES	3158.0	32.6	2.0	1579.0
111	CASCO H.VALLECAS	4445.0	53.2	6.0	741.0
112	SANTA EUGENIA	2678.0	53.2	2.0	1339.0
113	ENSANCHE DE VALLECAS	5201.0	53.2	4.0	1300.0
114	CASCO H.VICALVARO	3220.0	31.1	4.0	805.0
115	VALDEBERNARDO	1614.0	31.1	0.0	1614.0
116	VALDERRIVAS	1655.0	31.1	0.0	1655.0
117	EL CANAVERAL	182.0	31.1	0.0	182.0
119	HELLIN	1081.0	25.7	0.0	1081.0
120	AMPOSTA	1018.0	25.7	0.0	1018.0
122	ROSAS	3696.0	25.7	4.0	924.0
124	CANILLEJAS	3366.0	25.7	3.0	1122.0
125	EL SALVADOR	1332.0	25.7	4.0	333.0

# Conclusions

- **Best location** for the new Veterinary Clinic based in the analyzed criteria are the 4 neighbourhoods included in **Cluster 0**.
- The neighbourhoods included in the Cluster 2 are not so good as the included in cluster 0 but are good location to install a new veterinary clinic yet.
- The neighbourhoods included in Cluster 3 are not good location to install a new veterinary clinic at the present but can be studied at the future to see how the increase of pets evolves.
- **Worst location** to install a new veterinary clinic are the neighbourhoods included in the **Cluster 1**.