DATA

The data used for this capstone was obtained from
https://datos.madrid.es/portal/site/egob/menuitem.c05c1f754a33a9fbe4b2e4b284f1a5a0/?vgnextoid=b3c41f3cf6a6c410VgnVCM2000000c205a0aRCRD&vgnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnextfmt=default It is official data from a public website of the government it has multiple data that it is too detailed and useless for my study, but I can extract the coordinates form the different Postal Codes, which is the information that I actually will use.
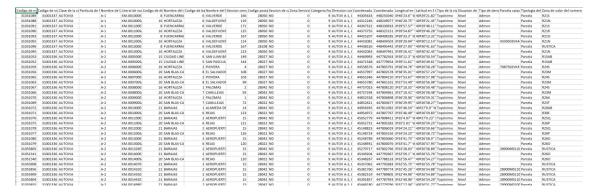

Figure 1 Example of the original Data

After cleaning this is the data that I will be using:

	Barrio	CP	Lon	Lat
0	CASTELLANA	28001	-3.681086	40.429061
1	CIUDAD JARDIN	28002	-3.666369	40.451381
2	ALMAGRO	28003	-3.699867	40.438308
3	ALMAGRO	28004	-3.695542	40.428072
4	ACACIAS	28005	-3.703536	40.404592
5	CASTELLANA	28006	-3.680981	40.430419
6	ADELFAS	28007	-3.673142	40.402025
7	ARGUELLES	28008	-3.717186	40.431500
8	GOYA	28009	-3.680008	40.422697
9	ALMAGRO	28010	-3.692111	40.426914
10	ALUCHE	28011	-3.761086	40.401814
11	ACACIAS	28012	-3.702308	40.404411
12	CORTES	28013	-3.699481	40.419275
13	CORTES	28014	-3.693883	40.412444
14	ARAPILES	28015	-3.706394	40.429808
15	CASTILLA	28016	-3.671872	40.467547
16	PUEBLO NUEVO	28017	-3.637119	40.433400
17	PALOMERAS BAJAS	28018	-3.657567	40.378314
18	ABRANTES	28019	-3.720228	40.385092

Figure 2: Consolidated Data