MOUILLON Vincent
TAILLIEU Victor
VAIO Luca

Data analysis

Voting in French presidential elections according to socio-professional category
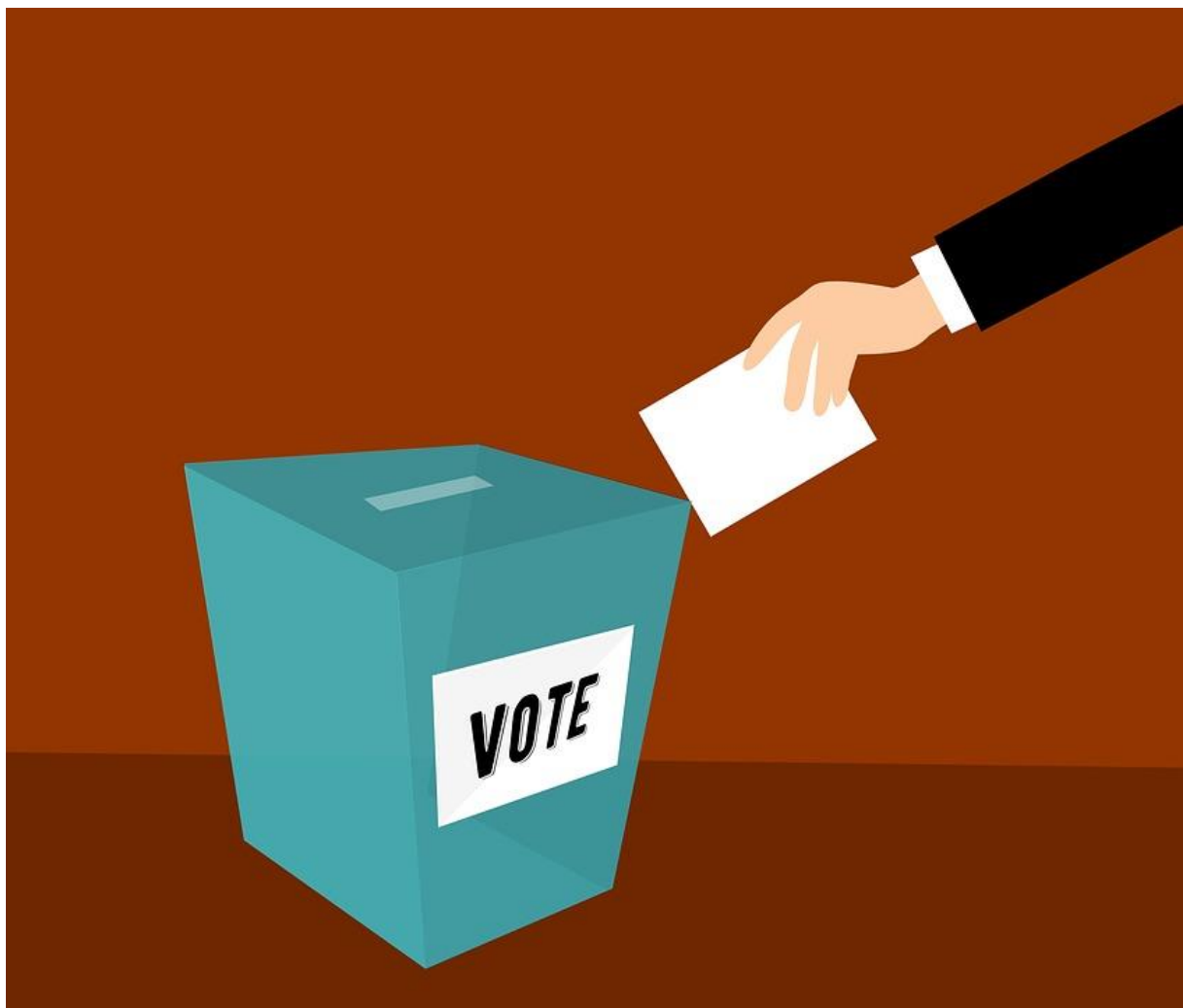
Table of contents

Introduction

We decided to analyse the results of French presidential elections and a study on socio-professional categories. The goal is to see if the socio-professional category has an impact on the voting choice. To do this, we performed an analysis on the 95 French departments (Metropolitan France). We expect less qualified people to vote mainly for extreme parties and more qualified people to vote for more moderate parties (left, right or centre).

Data

As raw data, we used results of presidential elections from 2012 and 2017 (1st turn) as well as a study on socio-professional categories. Departments 2A and 2B of Corsica were merged and renamed 20 for simplicity reasons.

Source : https://www.data.gouv.fr/fr/posts/les-donnees-des-elections/

The first step was to clean the raw data. Indeed, we only take variables that interest us for the analysis. We decided to take the abstention rate, the rate of blank and void votes, as well as results for each candidate. We took the proportions (in %) so that there is no weight influence (departments more populated than others). The following images are extracts from the cleaned datasets (you can find the csv files in the folder "data").

Presidential elections 2012:

	ABS	BN	JOLY	LE.PEN	SARKOZY	MELENCHON	POUTOU	ARTHAUD	CHEMINADE	BAYROU	DUPONT.AIGNAN	HOLLANDE
1	16.76	1.97	2.26	20.71	30.41	9.61	1.03	0.56	0.27	10.16	2.24	22.75
2	19.39	1.71	1.16	26.33	24.20	10.19	1.30	0.84	0.25	6.68	1.96	27.10
3	17.66	2.40	1.57	18.32	24.02	13.58	1.25	0.72	0.22	8.65	1.98	29.68
4	16.97	2.05	2.91	20.71	25.47	15.15	1.38	0.48	0.28	7.42	1.83	24.36
5	17.07	2.08	3.63	17.70	26.11	14.03	1.33	0.56	0.24	9.86	2.05	24.49
6	20.58	1.53	2.15	23.50	37.19	8.49	0.69	0.27	0.21	6.69	1.59	19.21
7	15.85	1.99	2.80	20.04	23.76	14.07	1.37	0.63	0.26	9.15	1.94	25.98
8	20.61	1.59	1.22	24.50	24.43	9.28	1.29	0.77	0.24	7.52	1.81	28.93
9	15.90	1.86	2.85	16.79	18.72	16.86	1.45	0.55	0.23	6.67	1.51	34.36
10	18.64	2.09	1.41	25.12	30.33	7.93	1.01	0.61	0.25	8.37	2.17	22.79
11	16.51	1.95	2.09	23.22	21.63	13.15	1.26	0.52	0.21	6.04	1.47	30.41
12	14.70	2.44	2.34	14.10	25.51	12.26	1.33	0.53	0.24	12.55	1.70	29.44
13	19.58	1.58	2.11	23.38	27.50	13.43	0.80	0.35	0.22	6.35	1.35	24.51
14	16.82	1.62	2.13	16.22	27.34	10.56	1.34	0.66	0.25	10.10	2.05	29.34
15	16.84	2.18	1.56	15.12	28.61	8.98	1.24	0.60	0.24	11.16	1.66	30.84
16	17.89	2.18	1.83	17.72	23.06	11.32	1.47	0.70	0.25	8.81	2.04	32.79
17	18.11	2.00	2.21	17.45	28.10	10.38	1.42	0.57	0.23	9.08	2.09	28.47
18	19.62	2.16	1.58	19.73	24.96	13.81	1.22	0.76	0.24	8.84	2.11	26.76
19	14.64	2.05	1.38	13.31	21.59	10.55	1.01	0.42	0.18	6.93	1.65	42.97
20	25.73	1.78	2.29	24.39	31.41	9.85	1.17	0.31	0.21	5.01	1.07	24.28
21	16.99	1.85	2.10	18.84	28.57	9.51	1.06	0.56	0.24	9.37	1.95	27.79
22	14.29	1.80	2.78	13.58	23.86	12.20	1.33	0.67	0.24	10.60	1.71	33.02
23	17.77	2.39	1.76	16.27	22.23	13.01	1.49	0.72	0.25	8.33	1.93	34.02
24	15.16	2.20	2.13	17.01	22.93	13.71	1.37	0.57	0.24	8.07	1.87	32.09
25	17.20	2.02	2.35	19.19	28.49	10.96	1.18	0.65	0.26	8.80	1.84	26.28
26	16.87	1.79	2.87	20.96	26.12	12.10	1.11	0.71	0.26	8.88	1.94	25.05
27	18.49	1.79	1.66	22.75	27.78	10.34	1.37	0.65	0.27	8.41	2.11	24.65
28	19.00	1.82	1.60	20.72	29.39	8.96	1.22	0.67	0.26	9.22	2.24	25.71
29	15.84	1.67	2.95	11.98	24.46	11.52	1.48	0.60	0.27	11.27	1.78	33.70
30	17.15	1.68	2.12	25.51	24.86	13.33	1.01	0.47	0.23	6.91	1.46	24.11

Presidential elections 2017:

	ABS	BN	MÉLENCHON	MACRON	FILLON	LE.PEN	HAMON	DUPONT.AIGNAN	POUTOU	LASSALLE	ASSELINEAU	ARTHAUD	CHEMINADE
1	19.60	2.57	15.88	22.62	21.43	25.00	5.13	6.07	0.95	1.06	1.11	0.57	0.18
2	21.34	2.52	16.99	17.94	16.30	35.67	4.24	5.08	1.10	0.79	0.75	0.96	0.19
3	21.41	3.40	19.91	23.72	18.94	22.34	5.52	5.09	1.21	1.55	0.74	0.80	0.18
4	19.22	2.49	22.51	20.02	18.49	24.53	5.00	4.87	1.18	1.73	0.93	0.52	0.21
5	18.94	2.42	21.62	21.80	19.15	21.25	5.88	5.68	1.21	1.85	0.90	0.47	0.19
6	21.25	1.99	14.96	19.04	27.39	27.75	3.58	4.28	0.62	0.90	1.03	0.29	0.16
7	18.70	3.49	21.80	21.64	17.30	23.17	6.01	5.08	1.33	1.81	1.01	0.67	0.19
8	22.68	2.33	17.83	18.33	17.22	32.41	4.93	5.32	1.15	0.96	0.74	0.94	0.18
9	18.22	2.82	26.76	20.92	12.75	21.70	7.85	3.61	1.26	3.54	0.86	0.60	0.15
10	19.11	2.32	13.97	18.98	23.05	30.33	4.06	6.35	0.86	0.72	0.82	0.69	0.18
11	19.27	2.63	21.52	20.07	15.06	28.26	6.35	3.86	1.17	2.17	0.83	0.56	0.16
12	16.67	2.99	19.66	25.83	20.78	16.20	6.16	4.85	1.32	3.66	0.74	0.63	0.18
13	21.87	2.11	22.02	19.37	19.76	27.28	4.53	3.85	0.76	0.96	0.90	0.39	0.17
14	18.02	2.24	18.83	24.83	20.50	20.36	6.80	5.00	1.31	0.72	0.70	0.77	0.17
15	19.39	3.26	15.91	26.73	23.58	18.17	5.25	4.42	1.29	3.10	0.58	0.76	0.22
16	21.15	2.91	20.48	25.07	16.96	21.40	6.32	5.03	1.37	1.55	0.77	0.83	0.23
17	20.22	2.54	18.97	23.91	20.59	21.14	5.76	5.32	1.28	1.38	0.79	0.67	0.18
18	22.23	3.02	19.51	22.05	19.09	24.18	5.30	5.53	1.22	1.11	0.86	0.94	0.20
19	18.51	3.55	20.85	26.93	17.46	17.34	6.36	4.84	1.59	2.92	0.75	0.72	0.23
20	31.96	2.91	13.81	18.48	25.56	27.88	3.74	2.89	0.89	5.64	0.63	0.32	0.16
21	19.04	2.34	17.84	23.65	21.26	22.52	5.90	5.25	0.99	0.92	0.86	0.63	0.18
22	15.71	2.43	20.27	27.99	18.38	16.46	8.60	4.26	1.46	0.95	0.66	0.81	0.17
23	21.39	4.02	21.11	22.50	17.99	19.88	7.82	5.01	1.72	1.92	0.83	0.97	0.25
24	18.32	3.04	22.97	22.49	17.09	20.93	6.35	4.59	1.44	2.43	0.86	0.65	0.20
25	20.68	2.75	17.88	22.50	21.09	23.45	5.74	5.18	1.25	0.89	1.10	0.72	0.19
26	19.81	2.55	20.10	21.88	18.50	23.90	6.02	5.19	1.10	1.34	1.03	0.76	0.18
27	19.23	2.47	17.47	19.89	18.84	29.31	5.05	5.67	1.17	0.77	0.87	0.78	0.18
28	20.06	2.58	16.20	21.74	21.84	25.08	5.25	6.01	1.13	0.83	0.96	0.77	0.19
29	17.47	2.24	19.67	29.45	17.94	13.89	10.91	4.08	1.48	1.11	0.66	0.65	0.15
30	19.94	2.31	21.61	18.78	17.20	29.30	4.87	4.23	0.96	1.41	0.97	0.50	0.16

Study on socio-professional categories:

	JEUN	VIEU	AGRI	ACCE	CPIS	PINT	EMPL	OUVR	CHOM
1	20.6	25.0	0.8	3.7	8.1	15.7	16.7	16.6	8.7
2	19.6	27.5	1.2	2.7	4.5	11.5	16.8	18.3	15.0
3	15.7	35.2	2.0	3.4	4.2	10.8	15.8	14.4	11.6
4	16.9	34.4	1.6	5.2	5.4	12.7	16.0	11.7	12.0
5	17.3	31.2	2.0	5.1	5.4	14.5	17.6	11.2	8.3
6	16.0	32.6	0.2	4.7	8.6	13.1	17.8	9.5	11.3
7	17.7	32.6	1.6	4.3	5.1	12.7	15.3	14.3	11.5
8	19.0	28.3	1.4	3.0	4.2	11.3	16.0	18.2	14.9
9	16.4	34.7	1.8	4.1	4.6	11.9	16.6	12.6	13.1
10	18.3	28.8	1.9	3.0	5.1	12.0	16.2	16.9	12.9
11	17.1	34.6	1.9	4.4	4.8	11.3	16.5	11.9	15.9
12	15.8	36.1	4.5	4.4	4.4	11.4	14.9	12.6	8.2
13	17.8	27.8	0.3	3.4	8.8	14.2	16.5	10.4	14.3
14	18.3	27.2	1.0	3.5	6.7	13.5	17.0	14.5	11.3
15	14.6	35.6	5.7	4.2	3.8	10.0	15.3	13.3	7.7
16	16.3	32.5	1.9	3.6	5.1	11.4	15.7	15.7	11.9
17	16.1	35.1	1.6	4.3	5.1	11.4	16.6	12.7	12.6
18	16.5	33.2	1.6	3.3	5.2	11.8	16.4	14.8	11.8
19	14.9	36.0	2.2	3.9	4.6	11.7	16.0	13.4	8.6
20	16.3	34.2	1.6	3.8	5.2	11.6	16.5	13.8	12.2
21	17.0	26.9	1.2	3.3	7.7	14.7	16.4	14.1	9.5
22	17.7	34.6	2.4	3.7	5.2	11.6	14.7	14.5	9.6
23	14.0	39.1	4.8	3.6	4.0	9.2	15.2	12.0	10.7
24	15.2	37.4	2.1	4.7	4.1	10.2	15.8	13.6	11.8
25	18.6	26.5	0.9	3.0	7.1	13.8	15.3	18.1	11.0
26	18.9	29.5	1.5	4.1	6.4	13.8	15.3	14.5	12.3
27	20.4	25.9	0.8	3.4	6.4	13.9	16.5	18.3	11.7
28	19.8	26.7	1.1	3.1	6.8	14.4	17.4	16.3	10.5
29	17.7	30.9	1.4	3.3	6.6	13.2	16.0	13.7	10.4
30	18.1	30.7	0.9	4.4	6.1	13.0	16.5	12.2	15.9

Here is the signification of the variables:

JEUN	under 15 years %
VIEU	more than 60 years %
AGRI	farmers % relative to the labour force
ACCE	craftsmen, traders and entrepreneurs %
CPIS	executives, higher intellectual professions %
PINT	intermediate professions %
EMPL	employees %
OUVR	working class %
CHOM	unemployed people %

Basic statistics

These files are now ready to be imported in R to perform the analysis. First, we can look at basic statistics (min, quartiles, median, mean and max):

```
> summary(elec2012)
      ABS              BN             JOLY            LE.PEN          SARKOZY         MELENCHON          POUTOU          ARTHAUD
 Min.   :13.86   Min.   :1.200   Min.   :1.160   Min.   : 6.20   Min.   :18.72   Min.   : 7.22   Min.   :0.670   Min.   :0.2700
 1st Qu.:16.46   1st Qu.:1.785   1st Qu.:1.760   1st Qu.:15.56   1st Qu.:24.20   1st Qu.: 9.85   1st Qu.:1.105   1st Qu.:0.4950
 Median :17.66   Median :1.910   Median :2.100   Median :19.17   Median :25.83   Median :11.17   Median :1.270   Median :0.6100
 Mean   :17.94   Mean   :1.948   Mean   :2.168   Mean   :18.88   Mean   :26.44   Mean   :11.29   Mean   :1.227   Mean   :0.5966
 3rd Qu.:19.32   3rd Qu.:2.100   3rd Qu.:2.510   3rd Qu.:22.36   3rd Qu.:28.42   3rd Qu.:12.46   3rd Qu.:1.360   3rd Qu.:0.7050
 Max.   :26.54   Max.   :2.660   Max.   :4.180   Max.   :27.03   Max.   :37.19   Max.   :16.99   Max.   :1.730   Max.   :0.8900
   CHEMINADE         BAYROU        DUPONT.AIGNAN      HOLLANDE
 Min.   :0.1800   Min.   : 5.01   Min.   :1.000   Min.   :18.89
 1st Qu.:0.2300   1st Qu.: 8.02   1st Qu.:1.665   1st Qu.:24.86
 Median :0.2500   Median : 9.18   Median :1.900   Median :27.65
 Mean   :0.2479   Mean   : 9.21   Mean   :1.884   Mean   :28.05
 3rd Qu.:0.2600   3rd Qu.:10.29   3rd Qu.:2.105   3rd Qu.:31.30
 Max.   :0.3300   Max.   :15.67   Max.   :3.410   Max.   :42.97

> summary(elec2017)
      ABS              BN            MÉLENCHON         MACRON           FILLON          LE.PEN            HAMON         DUPONT.AIGNAN
 Min.   :15.71   Min.   :1.360   Min.   :13.81   Min.   :17.73   Min.   :12.75   Min.   : 4.99   Min.   : 3.440   Min.   :1.670
 1st Qu.:18.54   1st Qu.:2.340   1st Qu.:17.02   1st Qu.:20.58   1st Qu.:17.27   1st Qu.:18.18   1st Qu.: 5.150   1st Qu.:4.435
 Median :19.80   Median :2.580   Median :19.37   Median :22.50   Median :18.60   Median :22.85   Median : 5.840   Median :5.010
 Mean   :19.80   Mean   :2.628   Mean   :19.24   Mean   :23.10   Mean   :19.58   Mean   :22.54   Mean   : 6.085   Mean   :5.008
 3rd Qu.:21.26   3rd Qu.:2.850   3rd Qu.:20.98   3rd Qu.:25.57   3rd Qu.:21.27   3rd Qu.:27.02   3rd Qu.: 6.820   3rd Qu.:5.680
 Max.   :31.96   Max.   :4.020   Max.   :34.02   Max.   :34.83   Max.   :29.14   Max.   :35.67   Max.   :10.910   Max.   :7.180
    POUTOU          LASSALLE        ASSELINEAU         ARTHAUD         CHEMINADE
 Min.   :0.620   Min.   :0.510   Min.   :0.5800   Min.   :0.2700   Min.   :0.1300
 1st Qu.:1.040   1st Qu.:0.835   1st Qu.:0.7500   1st Qu.:0.5450   1st Qu.:0.1700
 Median :1.170   Median :1.100   Median :0.8300   Median :0.7100   Median :0.1800
 Mean   :1.166   Mean   :1.531   Mean   :0.8774   Mean   :0.6799   Mean   :0.1853
 3rd Qu.:1.315   3rd Qu.:1.735   3rd Qu.:0.9600   3rd Qu.:0.8100   3rd Qu.:0.2000
 Max.   :1.720   Max.   :7.550   Max.   :1.6100   Max.   :1.0600   Max.   :0.2500

> summary(socioeco)
      JEUN             VIEU            AGRI             ACCE            CPIS             PINT            EMPL             OUVR
 Min.   :14.00   Min.   :20.10   Min.   :0.000   Min.   :2.300   Min.   : 3.800   Min.   : 9.20   Min.   :13.30   Min.   : 5.00
 1st Qu.:16.65   1st Qu.:26.35   1st Qu.:0.700   1st Qu.:3.000   1st Qu.: 4.800   1st Qu.:11.75   1st Qu.:15.70   1st Qu.:12.35
 Median :17.80   Median :29.80   Median :1.400   Median :3.400   Median : 5.700   Median :12.70   Median :16.20   Median :14.30
 Mean   :17.82   Mean   :29.73   Mean   :1.459   Mean   :3.508   Mean   : 6.887   Mean   :13.09   Mean   :16.36   Mean   :14.22
 3rd Qu.:18.90   3rd Qu.:33.10   3rd Qu.:1.900   3rd Qu.:4.050   3rd Qu.: 7.600   3rd Qu.:14.35   3rd Qu.:16.75   3rd Qu.:16.30
 Max.   :22.30   Max.   :39.10   Max.   :5.700   Max.   :5.200   Max.   :27.600   Max.   :18.10   Max.   :21.40   Max.   :18.40
     CHOM
 Min.   : 7.50
 1st Qu.: 9.95
 Median :11.30
 Mean   :11.35
 3rd Qu.:12.20
 Max.   :16.90
```

Correlation matrices

Basic statistics only tell us how values are distributed. We can compute the correlation matrix on raw data or centred and reduced data (it is the same) to find some relationships between variables:

	ABS	BN	JOLY	LE.PEN	SARKOZY	MELENCHON	POUTOU	ARTHAUD	CHEMINADE	BAYROU	DUPONT.AIGNAN	HOLLANDE
ABS	1.00000000	-0.46031614	-0.16234094	0.24867384	0.25039553	-0.11509294	-0.44470757	-0.16357847	0.06932515	-0.38513883	-0.1411244	-0.18858719
BN	-0.46031614	1.00000000	-0.26064157	-0.02897423	-0.32573247	0.07248912	0.52806601	0.46567118	0.02069426	0.24555298	0.3002745	0.13577072
JOLY	-0.16234094	-0.26064157	1.00000000	-0.52714270	0.19499877	0.23082687	-0.20334382	-0.51675869	0.16647781	0.37180132	-0.2375012	0.02263634
LE.PEN	0.24867384	-0.02897423	-0.52714270	1.00000000	0.03064053	-0.24633879	0.02478287	0.29707018	0.03896017	-0.56617458	0.1217789	-0.64920750
SARKOZY	0.25039553	-0.32573247	0.19499877	0.03064053	1.00000000	-0.72696911	-0.44801230	-0.29155766	0.21578626	0.26952321	0.1082046	-0.65742452
MELENCHON	-0.11509294	0.07248912	0.23082687	-0.24633879	-0.72696911	1.00000000	0.06587362	-0.18551551	-0.24762837	-0.30215174	-0.3422821	0.53389672
POUTOU	-0.44470757	0.52806601	-0.20334382	0.02478287	-0.44801230	0.06587362	1.00000000	0.58148143	0.12493452	0.28218192	0.2644890	0.13596814
ARTHAUD	-0.16357847	0.46567118	-0.51675869	0.29707018	-0.29155766	-0.18551551	0.58148143	1.00000000	0.25278594	0.10459907	0.4951324	-0.05191478
CHEMINADE	0.06932515	0.02069426	0.16647781	0.03896017	0.21578626	-0.24762837	0.12493452	0.25278594	1.00000000	0.30704745	0.3904033	-0.31180691
BAYROU	-0.38513883	0.24555298	0.37180132	-0.56617458	0.26952321	-0.30215174	0.28218192	0.10459907	0.30704745	1.00000000	0.2898429	-0.02451697
DUPONT.AIGNAN	-0.14112444	0.30027452	-0.23750115	0.12177892	0.10820460	-0.34228205	0.28448900	0.49513240	0.39040332	0.28984292	1.0000000	-0.26107390
HOLLANDE	-0.18858719	0.13577072	0.02263634	-0.64920750	-0.65742452	0.53389672	0.13596814	-0.05191478	-0.31180691	-0.02451697	-0.2610739	1.00000000

However, this table is difficult to read because there are many decimals and it is not easy to find extreme values (near -1 or 1). So that it is easier to read, we wrote a function to display only extreme values according to a threshold and with only two decimals.

```
highCor<-function(table,threshold){
    for(i in seq(1,length(table[,1]))){
        for(j in seq(1,length(table[1,]))){
            if(table[i,j]>-threshold && table[i,j]<threshold || i==j){
                table[i,j]<-NA
            }
        }
    }
    print(table,digits=2,na.print=".")
}
```

So, we obtain this for the 3 datasets:

```
> highCor(elec2012_cor,0.5)
                ABS    BN  JOLY LE.PEN SARKOZY MELENCHON POUTOU ARTHAUD CHEMINADE BAYROU DUPONT.AIGNAN HOLLANDE
ABS              .      .    .     .       .        .       0.53     .        .       .          .          .
BN               .      .    .     .       .        .       0.53     .        .       .          .          .
JOLY             .      .    .   -0.53     .        .        .     -0.52      .       .          .          .
LE.PEN           .      . -0.53    .       .        .        .       .        .     -0.57        .        -0.65
SARKOZY          .      .    .     .       .      -0.73      .       .        .       .          .        -0.66
MELENCHON        .      .    .     .     -0.73      .        .       .        .       .          .         0.53
POUTOU           .    0.53   .     .       .        .        .     0.58       .       .          .          .
ARTHAUD          .      . -0.52    .       .        .       0.58     .        .       .          .          .
CHEMINADE        .      .    .     .       .        .        .       .        .       .          .          .
BAYROU           .      .    .   -0.57     .        .        .       .        .       .          .          .
DUPONT.AIGNAN    .      .    .     .       .        .        .       .        .       .          .          .
HOLLANDE         .      .    .   -0.65   -0.66     0.53      .       .        .       .          .          .

> highCor(elec2017_cor,0.5)
                ABS    BN MÉLENCHON MACRON FILLON LE.PEN HAMON DUPONT.AIGNAN POUTOU LASSALLE ASSELINEAU ARTHAUD CHEMINADE
ABS              .      .     .        .      .      .     .        .           .        .         .        .        .
BN               .      .     .        .      .      .     .        .          0.68      .         .      0.57     0.52
MÉLENCHON        .      .     .      -0.63    .    0.54    .      -0.50         .        .         .        .        .
MACRON           .      .     .        .    -0.93  0.78    .        .           .        .         .        .        .
FILLON           .      .   -0.63      .      .      .     .        .           .        .         .        .        .
LE.PEN           .      .     .      -0.93    .      .   -0.79       .           .        .         .        .        .
HAMON            .      .   0.54     0.78     .    -0.79   .        .           .        .         .        .        .
DUPONT.AIGNAN    .      .   -0.50      .      .      .     .        .           .        .         .      0.52       .
POUTOU           .    0.68    .        .      .      .     .        .           .        .         .      0.59       .
LASSALLE         .      .     .        .      .      .     .        .           .        .         .        .        .
ASSELINEAU       .      .     .        .      .      .     .        .           .        .         .        .        .
ARTHAUD          .    0.57    .        .      .      .     .        .          0.52     0.59       .        .       0.54
CHEMINADE        .    0.52    .        .      .      .     .        .           .        .         .      0.54       .

> highCor(socioeco_cor,0.6)
      JEUN   VIEU  AGRI ACCE   CPIS   PINT EMPL   OUVR CHOM
JEUN    .  -0.81    .    .       .    0.64   .      .    .
VIEU -0.81    .   0.69 0.63  -0.63  -0.85   .      .    .
AGRI    .   0.69    .    .       .   -0.67   .      .    .
ACCE    .   0.63    .    .       .      .    .      .    .
CPIS    .  -0.63    .    .       .    0.71   .   -0.61   .
PINT  0.64 -0.85 -0.67   .     0.71    .     .      .    .
EMPL    .     .     .    .       .      .    .      .    .
OUVR    .     .     .    .    -0.61     .    .      .    .
CHOM    .     .     .    .       .      .    .      .    .
```

We can clearly see some relations between variables. We notice for example that Mélenchon is negatively correlated with Sarkozy, Macron is negatively correlated with Le Pen and farmers are positively correlated with old people.

These correlation matrices are the basis of the analysis. They are used to compute eigen values and vectors.

Eigen vectors

These are the corresponding eigen values and vectors calculated from correlation matrices:

```
> eigen(elec2012_cor)
eigen() decomposition
$values
 [1] 3.048053e+00 2.932263e+00 2.306867e+00 9.590417e-01 8.331776e-01 5.406769e-01 4.635538e-01 3.566225e-01 2.641855e-01 1.715202e-01
[11] 1.240371e-01 9.855940e-07


> eigen(elec2017_cor)
eigen() decomposition
$values
 [1] 3.844025e+00 2.870357e+00 1.955948e+00 1.550500e+00 9.346727e-01 6.183691e-01 3.872251e-01 2.939055e-01 1.831087e-01 1.608214e-01
[11] 1.061530e-01 9.491346e-02 1.008854e-06


> eigen(socioeco_cor)
eigen() decomposition
$values
[1] 4.16554760 1.85488086 1.20173728 0.84043736 0.46230224 0.28633901 0.08733250 0.08644695 0.01497618
```

```
> elec2012_pca$loadings

Loadings:
                Comp.1 Comp.2 Comp.3
ABS              0.336  0.105  0.249
BN              -0.444
JOLY             0.189 -0.203 -0.460
LE.PEN                  0.331  0.469
SARKOZY          0.346  0.345 -0.267
MELENCHON              -0.467  0.135
POUTOU          -0.475
ARTHAUD         -0.417  0.255  0.148
CHEMINADE               0.289 -0.239
BAYROU          -0.161        -0.571
DUPONT.AIGNAN   -0.250  0.352 -0.107
HOLLANDE               -0.184 -0.467
```

```
> elec2017_pca$loadings

Loadings:
                Comp.1 Comp.2 Comp.3
ABS              0.189  0.204  0.404
BN               0.213 -0.425
MÉLENCHON       -0.256 -0.161  0.511
MACRON          -0.399 -0.170 -0.260
FILLON                  0.266 -0.541
LE.PEN           0.431  0.156  0.175
HAMON           -0.407 -0.260
DUPONT.AIGNAN    0.374        -0.209
POUTOU                 -0.536
LASSALLE        -0.123 -0.168  0.191
ASSELINEAU              0.254  0.288
ARTHAUD          0.327 -0.339 -0.120
CHEMINADE        0.263 -0.251
```

Here, we only show the 3 first eigen vectors but there are as many eigen vectors as eigen values and variables. The goal of PCA is to reduce the number of variables/dimensions. We generally take at most the 3 most important PC because it is easy to represent on 2D graphs (one 2D graph for 2 PC and three 2D graphs for 3 PC). Eigen values and vectors are classified in decreasing order (the first one has the most information and the last one the least). An eigen vector contains the weights of initial variables to make the corresponding principal component. The latter is obtained by multiplying the respective eigen vector with the table of centered and reduced data.

To choose how many PC to take, we can use elbow curves (see next page). We can also look at the proportion of information contained in the k first components with the formula:

$$\frac{\lambda_1 + \ldots + \lambda_k}{\lambda_1 + \ldots + \lambda_n} \times 100$$

λ: eigen value, n: number of eigen values

```
> socioeco_pca$loadings

Loadings:
     Comp.1 Comp.2 Comp.3
JEUN  0.388  0.271  0.162
VIEU -0.471        -0.141
AGRI -0.402         0.314
ACCE -0.292 -0.305 -0.232
CPIS  0.328 -0.473
PINT  0.434 -0.239  0.132
EMPL  0.284  0.187 -0.301
OUVR         0.652  0.335
CHOM         0.300 -0.761
```

Elbow curves

Elbow curves are plots of eigen values in decreasing order. Their goal is to see where the "elbow" is to know how many principal components we should take. The idea is to take enough PC to have a good proportion of the information but not too much to not make the analysis too complex.

```
cat("\nInformation comp 1 and 2:", (elec2012_pca$sdev[1]^2+elec2012_pca$sdev[2]^2)/length(elec2012_pca$sdev)*100, "%\n")
```
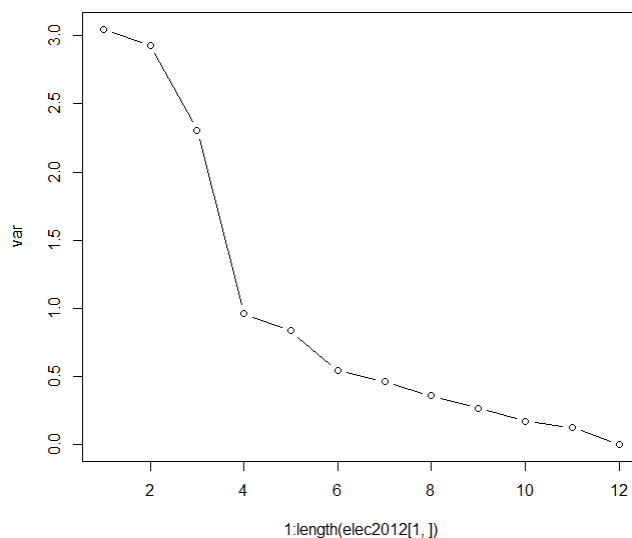

Elections 2012:

```
Information comp 1, 2: 49.8 %
Information comp 1, 2, 3: 69.1 %
```

We choose 3 PC. There is a clear elbow for the fourth eigen value.
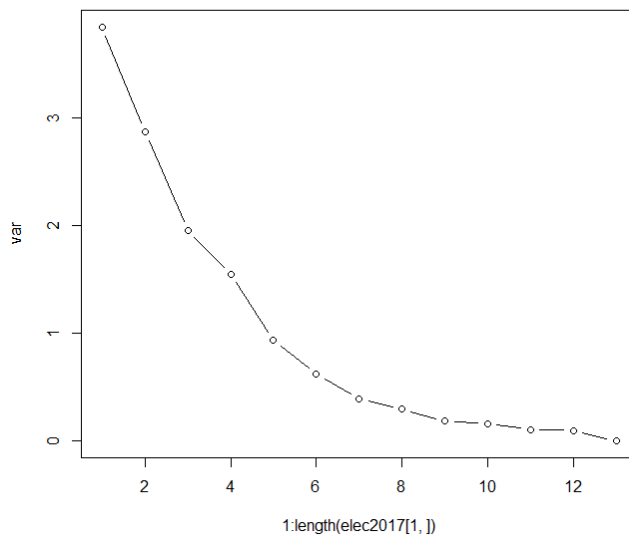
Elections 2017:

```
Information comp 1, 2: 51.6 %
Information comp 1, 2, 3: 66.7 %
```

We choose 3 PC. There is no clear elbow, but we don't have enough information with only 2 PC and 4 PC would be too complicated to analyze.
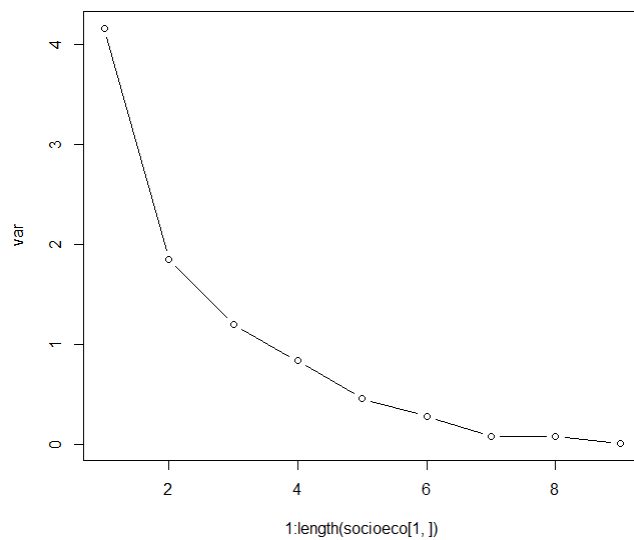
Socio-professional categories:

```
Information comp 1, 2: 66.9 %
Information comp 1, 2, 3: 80.2 %
```

We choose 2 PC. We can see an elbow for the second eigen value, but 1 PC is not enough to make an analysis. So, we take one more.

For each dataset, we managed to keep at least 66% of the initial information. This is not very high, but it is enough to perform a good analysis.

Meaning of the axes

Then, we can sort values in decreasing order in eigen vectors so that it is easier to see high positive and negative values. This will allow us to determine the meaning of the axes. Correlation circles of variables can also be helpful.

Elections 2012:

Comp1		Comp2		Comp3	
	[,1]		[,1]		[,1]
SARKOZY	0.34602593	DUPONT.AIGNAN	0.35172522	LE.PEN	0.469099803
ABS	0.33581140	SARKOZY	0.34451679	ABS	0.248633240
JOLY	0.18885206	LE.PEN	0.33094223	ARTHAUD	0.147721086
LE.PEN	0.03565760	CHEMINADE	0.28879059	MELENCHON	0.135395579
CHEMINADE	-0.05335058	ARTHAUD	0.25465435	HOLLANDE	-0.008632064
MELENCHON	-0.09406844	ABS	0.10529948	POUTOU	-0.020280390
BAYROU	-0.16050026	BAYROU	0.09252199	BN	-0.024359187
HOLLANDE	-0.18423796	POUTOU	0.04182432	DUPONT.AIGNAN	-0.107388435
DUPONT.AIGNAN	-0.25014465	BN	0.03121155	CHEMINADE	-0.238961153
ARTHAUD	-0.41652918	JOLY	-0.20288838	SARKOZY	-0.266906381
BN	-0.44425696	MELENCHON	-0.46701531	JOLY	-0.459820931
POUTOU	-0.47510833	HOLLANDE	-0.46710515	BAYROU	-0.570600749
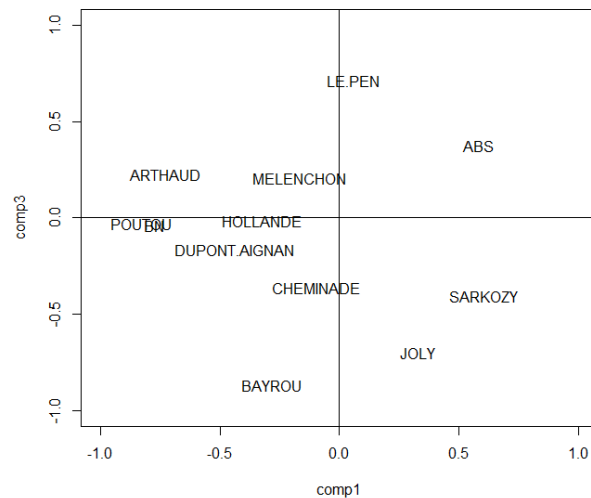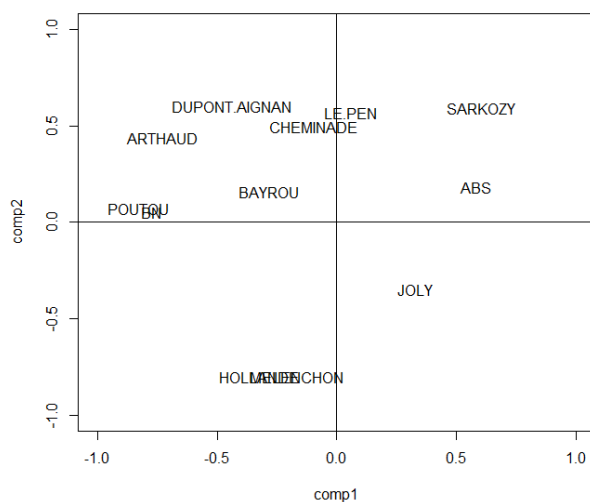
With this, we can make a table:

PC	-	+
1	POUTOU, BN, ARTHAUD	SARKOZY, ABS
2	HOLLANDE, MELENCHON	DUPONT-AIGNAN, SARKOZY, LE PEN
3	BAYROU, JOLY	LE PEN

For the first axis, we can say that a negative value means that people vote for far-left workers parties and a positive value vote for Sarkozy or abstain. The

second axis seems to oppose left (negative) and right (positive) candidates. The third one opposes center candidates (negative) and Le Pen (positive).

Elections 2017:

Comp1		Comp2		Comp3	
	[,1]		[,1]		[,1]
LE.PEN	0.43078087	FILLON	0.26584524	MÉLENCHON	0.51071625
DUPONT.AIGNAN	0.37371580	ASSELINEAU	0.25393027	ABS	0.40410204
ARTHAUD	0.32733496	ABS	0.20393398	ASSELINEAU	0.28806825
CHEMINADE	0.26305121	LE.PEN	0.15598605	LASSALLE	0.19108086
BN	0.21317623	DUPONT.AIGNAN	-0.04173796	LE.PEN	0.17510274
ABS	0.18869870	MÉLENCHON	-0.16110796	BN	0.06305554
POUTOU	0.09686105	LASSALLE	-0.16833015	HAMON	0.02468262
ASSELINEAU	0.01701032	MACRON	-0.17042287	POUTOU	-0.00255383
FILLON	-0.04048242	CHEMINADE	-0.25143803	CHEMINADE	-0.05386474
LASSALLE	-0.12331022	HAMON	-0.26010236	ARTHAUD	-0.11998439
MÉLENCHON	-0.25569071	ARTHAUD	-0.33869341	DUPONT.AIGNAN	-0.20865540
MACRON	-0.39943349	BN	-0.42541002	MACRON	-0.26023924
HAMON	-0.40734312	POUTOU	-0.53599156	FILLON	-0.54089121
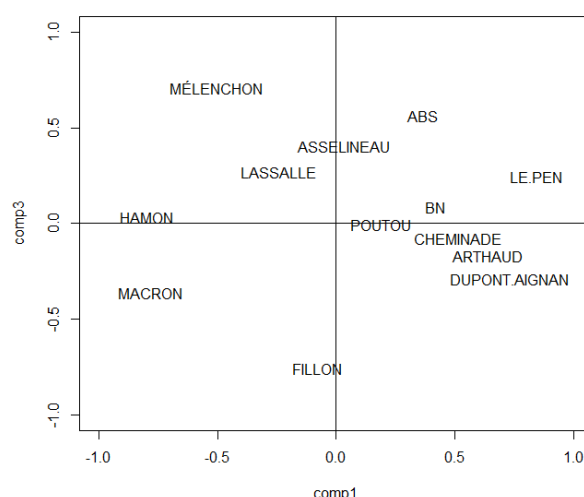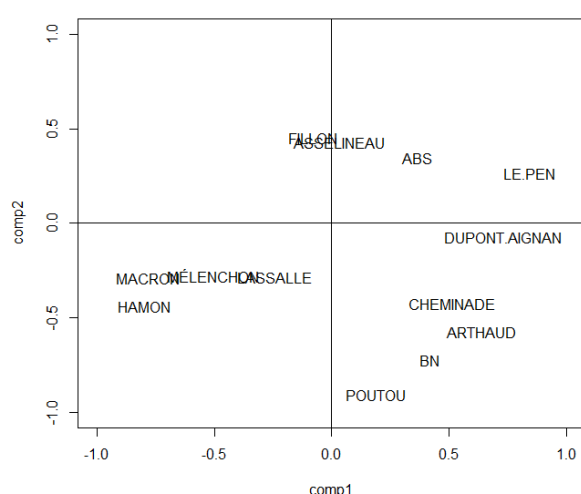
PC	-	+
1	HAMON, MACRON	LE PEN, DUPONT-AIGNAN, ARTHAUD
2	POUTOU, BN, ARTHAUD	FILLON, ASSELINEAU
3	FILLON	MELENCHON, ABS

For the first axis, a negative value means that people vote for center-left parties and a positive value for extreme parties. The second one opposes far-left workers parties (negative) and right parties (positive). The last one confronts Fillon and Mélenchon/abstention.

Socio-professional categories:

```
               [,1]                   [,1]
PINT    0.43350058   OUVR    0.65244023
JEUN    0.38761995   CHOM    0.30003245
CPIS    0.32827862   JEUN    0.27110614
EMPL    0.28360863   EMPL    0.18726332
CHOM    0.05075403   AGRI   -0.01730830
OUVR   -0.04894830   VIEU   -0.04547124
ACCE   -0.29228908   PINT   -0.23926368
AGRI   -0.40230595   ACCE   -0.30457113
VIEU   -0.47050231   CPIS   -0.47260796

Comp1                Comp2
```
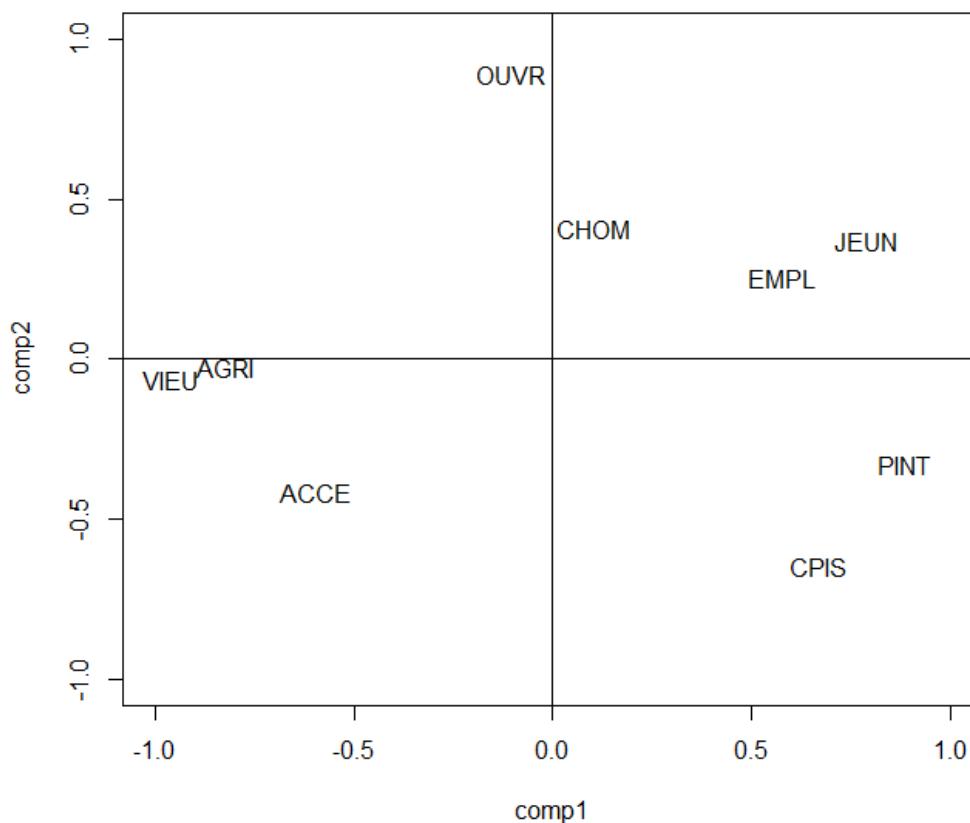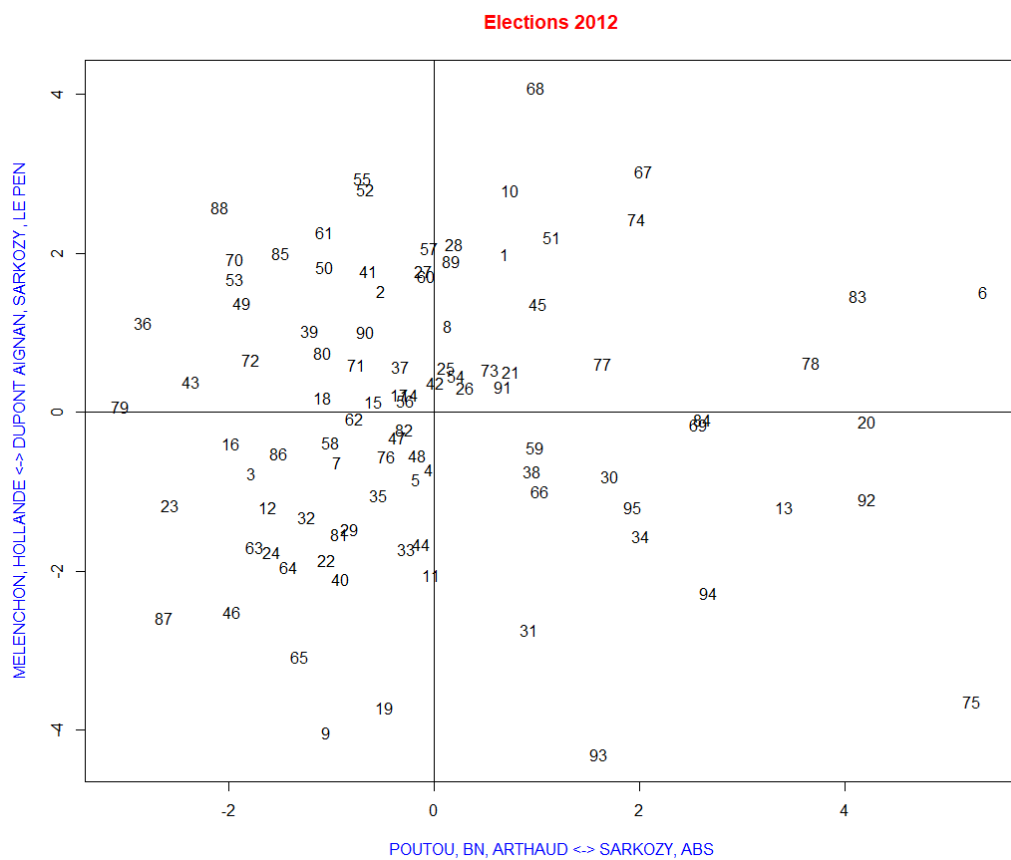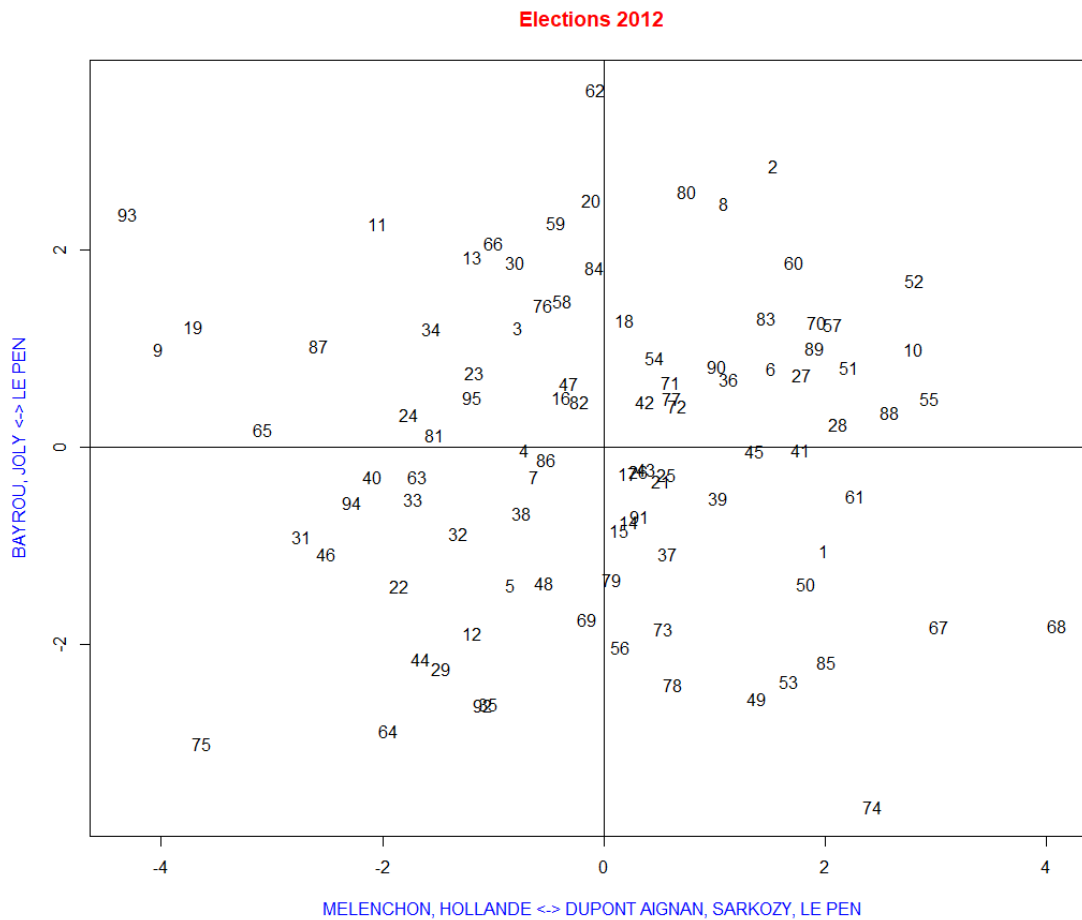

PC	-	+
1	VIEU, AGRI	PINT, JEUN
2	CPIS	OUVR

The first axis opposes old farmers and intermediate professions with children. We can interpret it as the axis of the activity of the population. The other one confronts very qualified people and the working class. It illustrates the qualification of people.

Graphs

Elections 2012

Elections 2012

Elections 2012
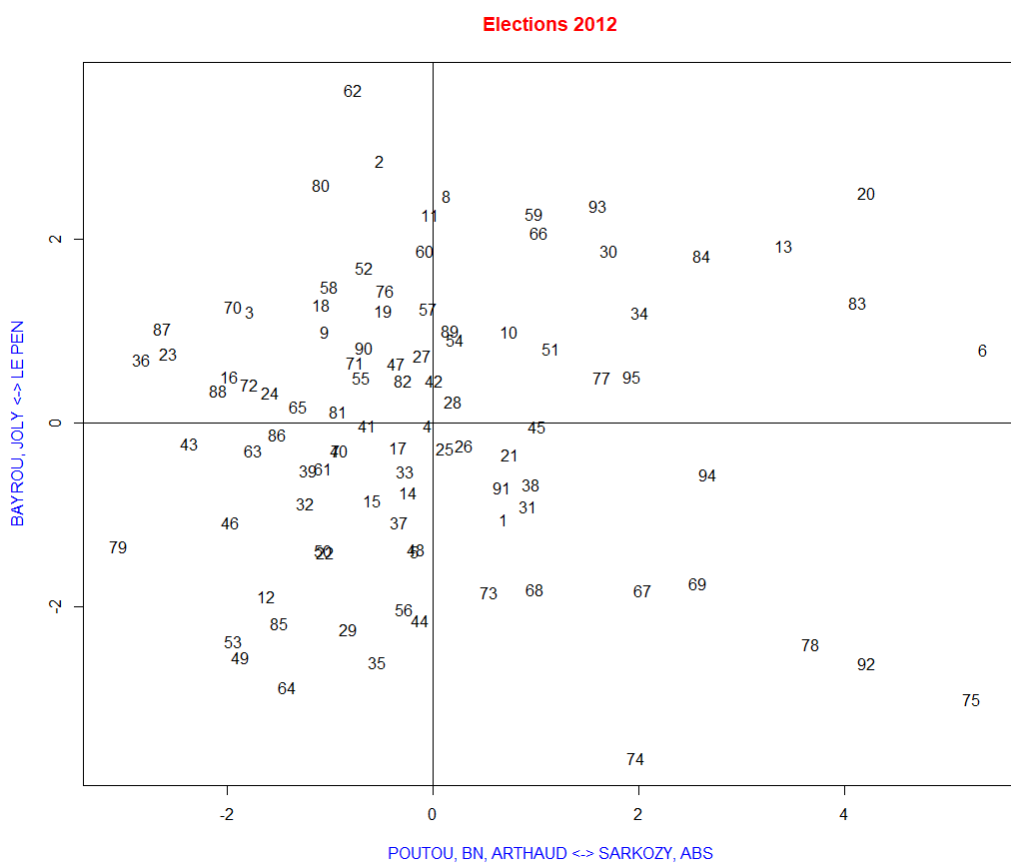
From these graphs (representing departments according to principal components), we can make some conclusions. On the first axis, we remark that departments 79, 36, 87, 23 and 43 vote for far-left workers parties as opposed to departments 6, 75, 92, 20, 83, 78 and 13 that vote for Sarkozy and have a high rate of abstention. On the second axis, we can see the classic left-right opposition, respectively departments 93, 9, 19, 75, 65 and departments 68, 67, 55, 52 and 10. Finally, the third axis confronts departments voting for center candidates: 74, 75, 64, 92, 35, 49 and the ones voting for Le Pen: 62, 2, 80, 20 and 8.

14

Elections 2017
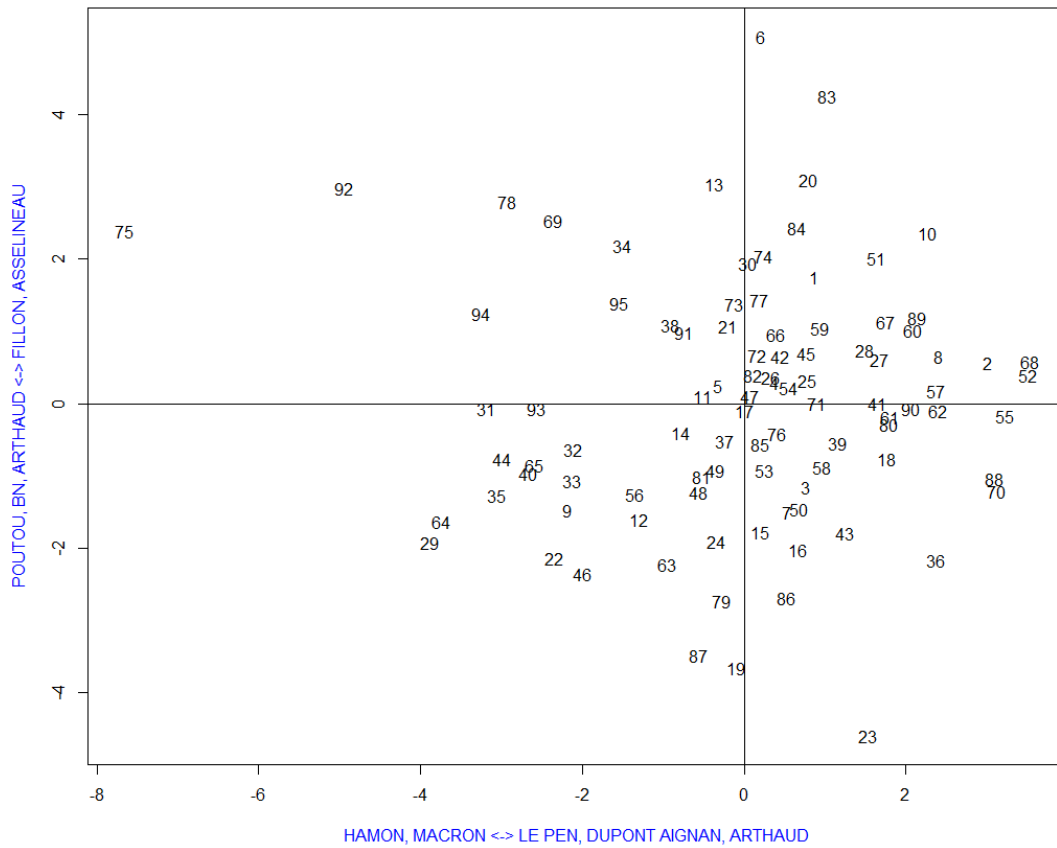
Elections 2017
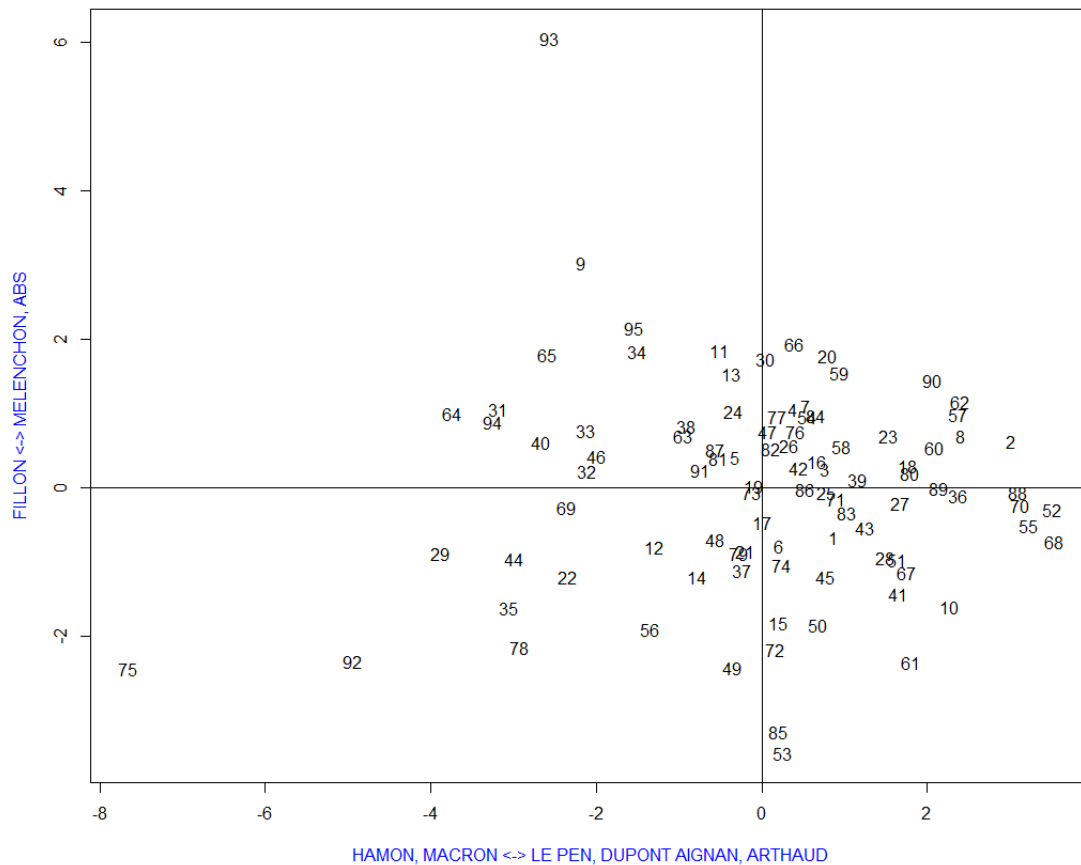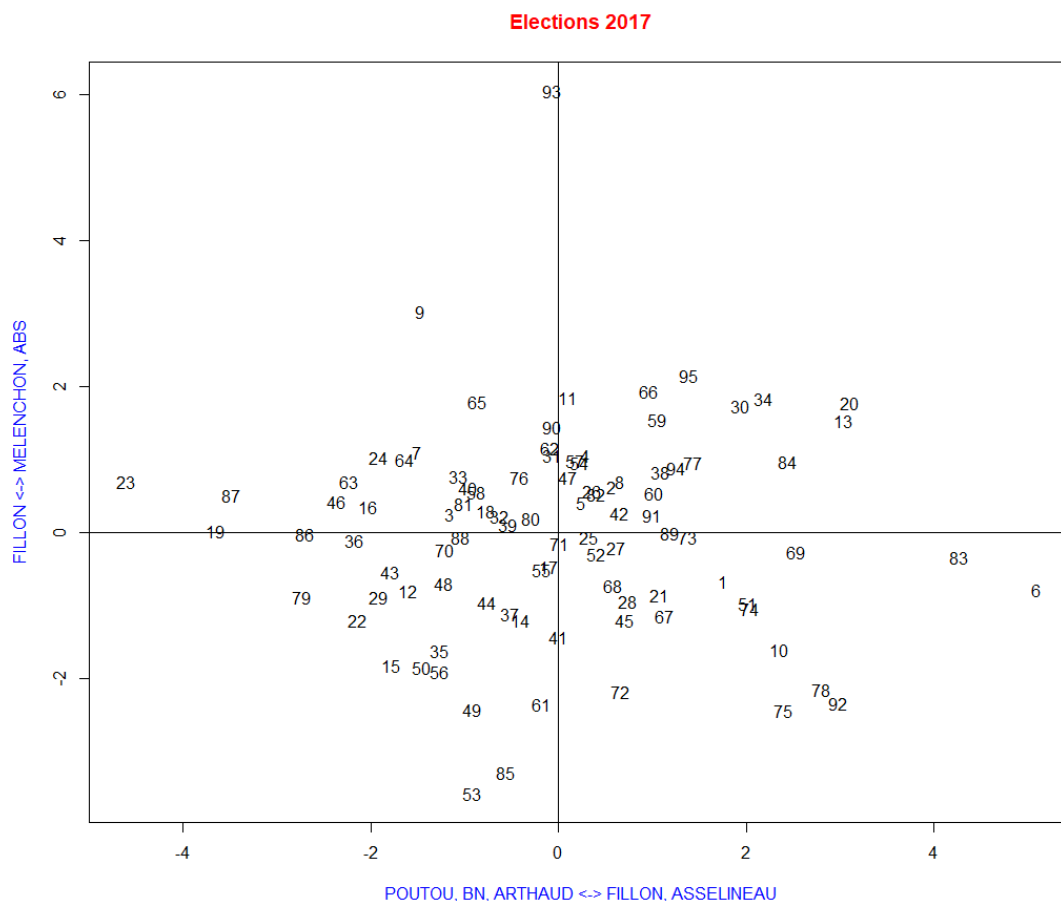
Elections 2017

For the 2017 presidential elections, axes are a bit different from the 2012 one. But we still have some similarities. The first axis (that looks like the second one of the 2012 elections) contrasts between departments voting for center-left parties: 75, 92, 29, 64 and others voting for extreme parties: 68, 52, 55, 70, 88 and 2. The second axis, which can be assimilated to the first one of the 2012 elections, opposes far-left workers parties: 23, 19, 87, 79 and 86 to right parties: 6, 83, 20, 13, 92 and 78. The last one confronts departments voting for Fillon: 53, 85, 75, 49, 61, 92, 72, 78 and those voting for Mélenchon and having a high rate of abstention: 93, 9, 95 and 66.

Etude socio-économique
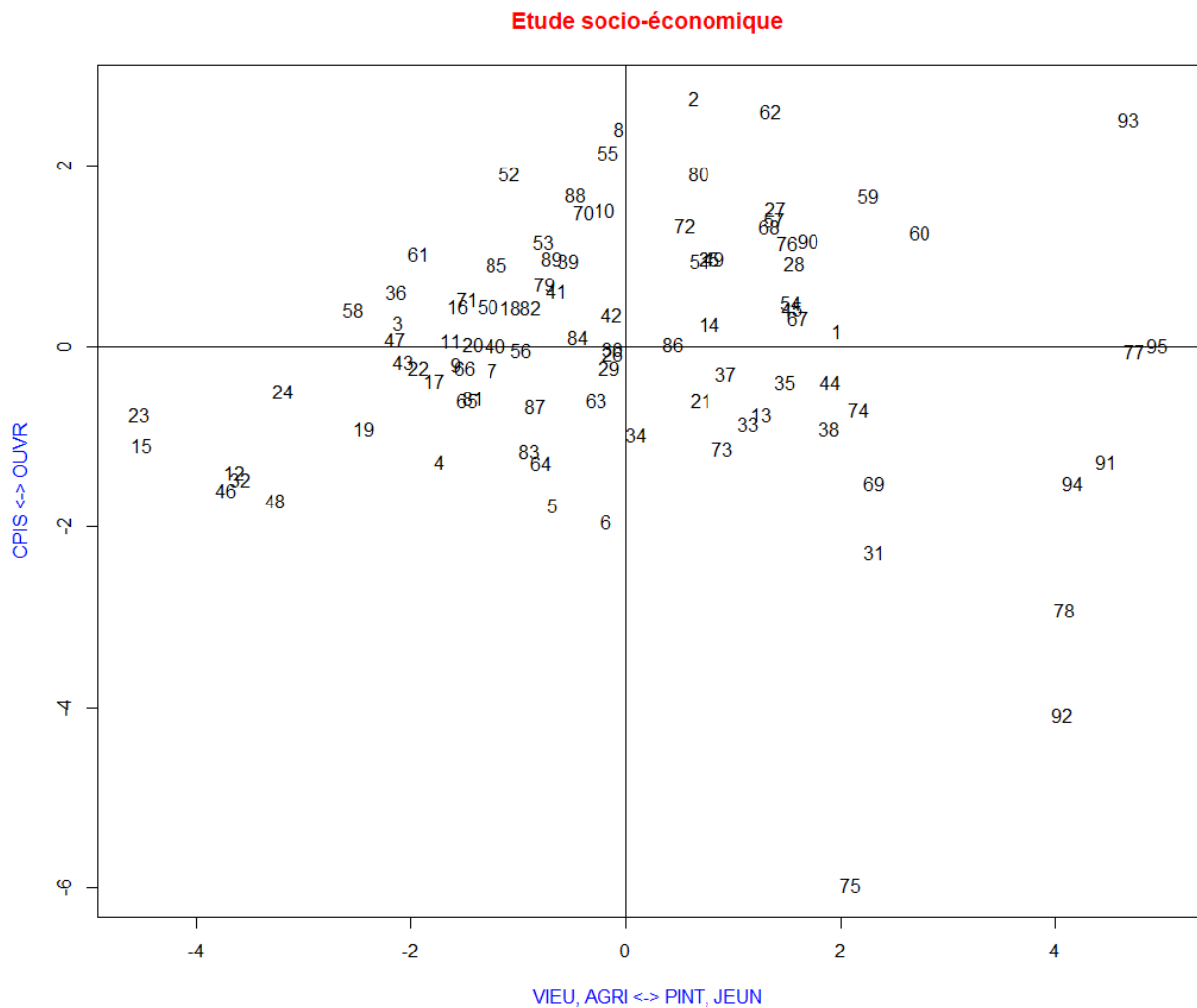
Now, let's look at the socio-professional analysis. The first axis reflects the activity of the departments. It opposes departments with a high proportion of old farmers: 23, 15, 46, 12, 32, 48 and 24 and those having a high rate of intermediate professions with children: 95, 77, 93, 91, 94, 78 and 92. The other one represents the qualification of the population, confronting very qualified people: 75, 92, 78, 31 and 6 with the working class: 2, 62, 93, 8, 55, 52 and 80.

Conclusion

From these, we can make categories of departments:

🔴 Departments voting for far-left workers parties: 79, 36, 87, 23, 43, 19 and 86

🔵 Departments voting for right parties: 6, 20, 83, 78, 13, 53, 85, 61 and 72

🟣 Departments voting for left to far-left parties: 93, 9, 65, 95 and 66

🔵 Departments voting for far-right parties: 68, 67, 55, 52, 10, 62, 2, 80, 8, 70 and 88

🟢 Departments voting for center parties: 74, 64, 35, 49 and 29

🟣 Special departments: 75 and 92

As departments 75 and 92 are cosmopolitan, they have not a clear orientation. They vote for right, left and center parties.

From these categories, we can color them on a France map:
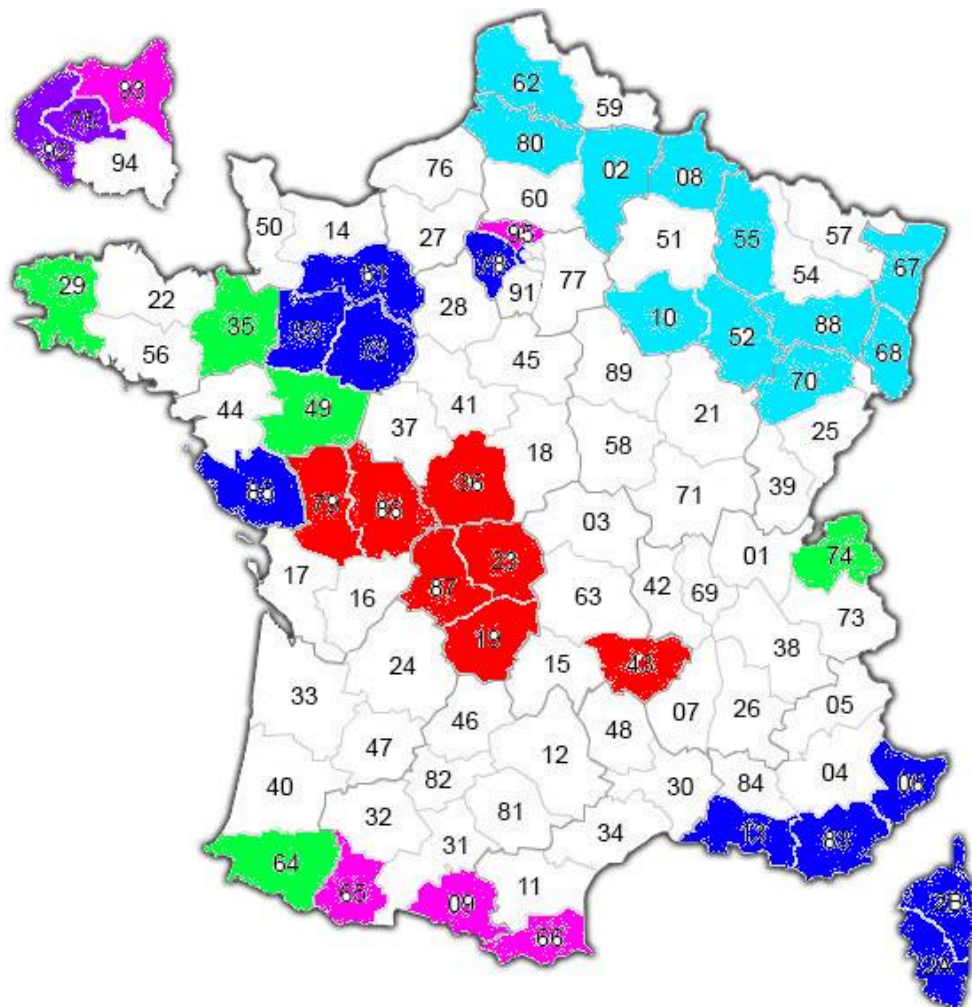
We notice concentrations of departments with the same political orientation. Indeed, north-east departments tend to vote for far-right parties. Likewise, south-east departments vote for right parties, south-west departments for center or left parties, center departments for far-left workers parties and north-west departments for center of right parties. The Ile-de-France region is a bit special. As it is a cosmopolitan region, people have no clear orientation and vote for right, left and center parties.

To make a link with the socio-professional study, we can look if a specific voting category matches with a socio-professional status. For instance, we see that department 23, which votes for far-left workers parties, has a high proportion of old farmers. In the same way, intermediate professions with children vote for left and right candidates as well as very qualified people (which contains special departments). Finally, the working class vote mainly for far-right candidates.

To conclude, we have shown that the socio-professional category can have a real impact on the voting choice. It confirms our expectations. Moreover, we remark zones of departments having the same political orientation. That can be explained: for example, the north-east of France is known to have a high rate of working class and they feel like they are the forgotten ones of globalization. That can explain their tendency to vote for far-right parties.