

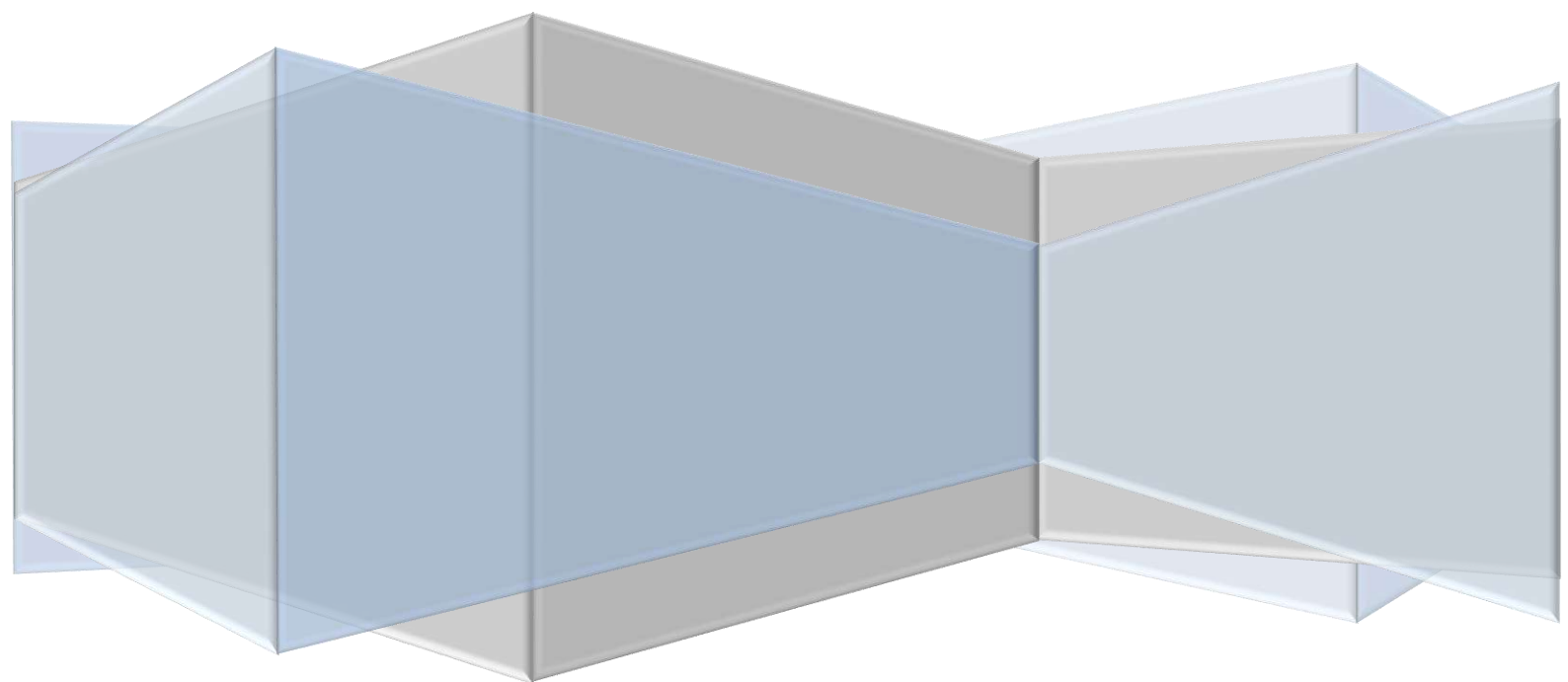
Promo 2015



Analyse de Données

TAI : Qualité de vie en fonction du département

DAVID Aurélien – FABRE Maxime
LEPOT Florian – RAMOS Adrien



Sommaire

Introduction.....	2
Choix d'étude.....	2
Description de nos données	2
Données choisies et utilisées.....	3
Analyse des données	6
Calcul de la matrice de corrélation	6
Calcul des valeurs propres.....	9
Calcul des vecteurs propres.....	9
Calcul des coordonnées des points.....	9
Représentations graphiques.....	11
Coordonnées des individus	12
Interprétation de chaque axe	16
Ouverture sur les résultats de la présidentielle de 2012.	17
Synthèses personnelles.....	19
Aurélien	19
Florian	19
Adrien.....	19
Maxime.....	19

Introduction

Le but de ce rapport de TAI est de nous faire interpréter une analyse de données, dans le cadre de l'approfondissement de nos cours.

Choix d'étude

« La qualité de vie varie-t-il entre les différents départements de France métropolitaine et a-t-elle eu une influence sur la Présidentielle de 2012 ? »

Nous avons choisi d'étudier la qualité de vie en France, en fonction des différents départements. Le contexte des élections de 2012, c'est particulièrement bien prêté au choix de ce sujet puisque chaque département choisi un candidat à la présidentielle, ce qui peut dans de nombreux cas, se relier à la qualité de vie du département. La notion de *qualité de vie* est différente de celle du niveau de vie qui ne prend pas en compte des facteurs plus subjectifs qui participent à la vie humaine, comme les loisirs, la sécurité, les ressources culturelles, etc.

Le sujet consiste en l'étude d'un fichier de données publié le 05/12/2011 par l'Institut national de la statistique et des études économiques (INSEE). En ligne, sont figurés les 95 départements analysés. Nous allons naturellement effectuer une ACP puisque nous sommes dans le cas de variables quantitatives.

Description de nos données

- **Nombre de délits** enregistrés dans le département en 2010, c'est temoins de la sécurité du département.
- **Taux de chômage** moyen de 1981 à 2007, qui présente la situation économique et reflète la main d'œuvre du département.
- **Nombre entreprises créées**, qui montre le dynamisme du marché du travail, de l'innovation.
- **Pourcentage d'étudiants**, montre la qualité de l'éducation et l'attrait pour les jeunes du département ainsi que la possibilité d'accéder à l'enseignement supérieur.
- **Nombre de musées**, l'importance accordée à la culture.

Données choisies et utilisées

Département	Nombre de délits enregistrés dans le département en 2010	Taux de chômage moyen de 1981 à 2007	Nombre entreprises créées	Pourcentage d'étudiants	Nombre de musées
1	22626	6,33	214,67	20,8	14
2	24051	12,43	143,53	21,5	15
3	11650	11,20	103,97	18,1	8
4	7022	9,87	93,17	18,8	9
5	5293	7,87	88,03	19,3	2
6	93200	11,36	930,43	20,9	32
7	11010	10,27	133,60	19,4	7
8	10750	13,48	84,45	20,7	7
9	4813	10,30	75,58	17,4	6
10	14378	10,24	94,50	21,5	11
11	19933	12,45	195,09	18,7	7
12	7295	6,58	98,52	18,4	11
13	172211	14,28	1184,78	23,5	41
14	29314	11,07	257,59	23,9	22
15	2984	8,77	42,17	17,2	5
16	13949	10,39	121,88	18,5	5
17	32620	12,49	292,33	19,4	19
18	11539	9,55	96,01	18,1	9
19	7975	8,04	81,38	17,7	5
21	19267	12,51	192,25	24,6	21
22	2092	8,75	36,89	20,2	3
23	13410	9,50	182,86	15,3	2
24	23316	8,95	174,64	16,8	14
25	25560	9,96	246,66	24,1	8
26	24729	8,81	193,25	21,2	4
27	17627	11,31	134,27	21,0	9
28	33119	10,09	269,90	20,6	14
29	6714	8,52	401,38	22,0	18
30	7910	9,49	664,05	21,0	7
31	47125	14,73	80,38	17,9	13
32	79929	10,91	761,92	17,8	7
33	5436	7,17	722,72	23,8	15
34	85286	11,56	333,14	23,8	19
35	76914	15,35	62,68	24,6	10
36	40027	8,49	208,09	26,6	8
37	7770	8,94	543,08	17,6	19
38	26115	9,74	90,59	23,4	23
39	65088	9,16	175,04	24,9	13
40	8319	6,66	104,69	20,3	11
41	14498	8,90	287,77	24,6	8
42	13851	8,55	73,62	29,2	21

Département	Nombre de délits enregistrés dans le département en 2010	Taux de chômage moyen de 1981 à 2007	Nombre entreprises créées	Pourcentage d'étudiants	Nombre de musées
43	34741	11,09	512,97	25,7	4
44	6183	8,34	227,07	17,8	14
45	71279	11,25	76,45	22,8	12
46	31823	8,08	132,20	22,5	11
47	4768	8,85	26,49	19,5	6
48	12608	10,59	241,90	16,9	2
49	2079	6,22	132,11	18,4	13
50	26984	9,74	180,77	20,8	20
51	15124	9,10	45,80	25,8	9
52	33286	9,75	74,47	20,3	7
53	7761	9,86	229,88	19,9	8
54	8878	6,41	50,03	24,4	11
55	36686	9,45	264,49		9
56	6787	9,29	342,67	21,8	10
57	25840	10,16	64,03	24,8	12
58	45331	9,70	803,99	19,6	11
59	7514	9,46	258,75	21,4	31
60	165411	13,74	84,17	20,8	15
61	41135	9,13	377,92	18,1	7
62	10047	9,10	215,22	22,1	15
63	73833	13,76	294,12	20,2	13
64	25336	9,27	100,15	23,1	
65	27868	10,10	306,07	21,2	7
66	9045	11,29	429,04	22,8	10
67	32383	14,54	261,09	23,3	24
68	49339	7,35	908,15	20,5	
69	32240	6,92	72,83	20,5	17
70	116078	9,11	170,64	19,1	6
71	7368	9,20	152,83	20,4	25
72	18519	9,71	243,08	27,9	5
73	19184	10,02	381,43	18,8	10
74	25650	8,11	2663,51	19,7	6
75	35317	7,16	350,69	26,6	56
76	238038	10,66	524,66	21,5	33
77	65791	12,53	601,53	22,5	19
78	80808	7,43	99,30	20,3	17
79	84309	6,59	139,08	19,3	8
80	11937	8,98	153,88	24,0	8
81	27086	12,10	109,47	24,5	13
82	13738	10,32	728,15	24,5	6
83	10559	10,11	354,17	22,4	16
84	62959	13,82	218,47	24,3	18
85	36998	11,99	132,01	19,2	14

Département	Nombre de délits enregistrés dans le département en 2010	Taux de chômage moyen de 1981 à 2007	Nombre entreprises créées	Pourcentage d'étudiants	Nombre de musées
86	20087	8,36	118,44	22,4	12
87	16551	9,57	125,83	19,3	4
88	13427	8,04	115,70	20,5	7
89	13163	10,17	45,08	24,2	12
90	15926	9,36	488,42	19,5	4
91	6394	10,07	898,42	23,8	8
92	69801	6,94	772,25	20,5	20
93	96476	8,30	611,32	23,8	5
94	142034	12,08	491,15	19,5	8
95	90179	8,67	601,53	24,0	8

Analyse des données

Calcul de la matrice de corrélation

Pour réaliser l'ACP, nous allons travailler avec une matrice de Corrélation R. En effet les valeurs de notre matrice de départ ont parfois un écart très important, il faut leur donner le même poids. On commencera donc par calculer la matrice centrée et réduite.

Pour d'obtenir notre matrice centrée et réduite, il faut enlever à chaque valeur la moyenne de sa colonne et diviser le tout par l'écart type de la colonne encore une fois.

Dans un souci d'efficacité tous nos calculs seront faits sous Excel.

Voici les moyennes et écart-type de chaque colonne :

	Colonne 1	Colonne 2	Colonne 3	Colonne 4	Colonne 5
Moyenne	35419,160	9,861	300,803	21,320	12,478
Ecart-type	40606,026	2,021	346,546	2,739	8,575

On obtient la matrice centrée et réduite suivante :

Nombre de délits enregistrés dans le département en 2010	Taux de chômage moyen de 1981 à 2007	Nombre entreprises créées	Pourcentage d'étudiants	Nombre de musées
-0,315	-1,746	-0,249	-0,182	0,177
-0,280	1,270	-0,454	0,078	0,294
-0,585	0,662	-0,568	-1,171	-0,522
-0,699	0,003	-0,599	-0,923	-0,406
-0,742	-0,986	-0,614	-0,751	-1,222
1,423	0,741	1,817	-0,155	2,277
-0,601	0,201	-0,482	-0,700	-0,639
-0,608	1,793	-0,624	-0,218	-0,639
-0,754	0,219	-0,650	-1,419	-0,755
-0,518	0,190	-0,595	0,075	-0,172
-0,381	1,281	-0,305	-0,954	-0,639
-0,693	-1,622	-0,584	-1,084	-0,172
3,369	2,189	2,551	0,798	3,326
-0,150	0,598	-0,125	0,929	1,110
-0,799	-0,541	-0,746	-1,510	-0,872
-0,529	0,263	-0,516	-1,031	-0,872
-0,069	1,302	-0,024	-0,690	0,761
-0,588	-0,152	-0,591	-1,166	-0,406
-0,676	-0,903	-0,633	-1,339	-0,872
-0,398	1,311	-0,313	1,212	0,994
-0,821	-0,552	-0,762	-0,417	-1,105

-0,542	-0,180	-0,340	-2,184	-1,222
-0,298	-0,453	-0,364	-1,661	0,177
-0,243	0,048	-0,156	1,033	-0,522
-0,263	-0,519	-0,310	-0,028	-0,989
-0,438	0,717	-0,481	-0,132	-0,406
-0,057	0,114	-0,089	-0,255	0,177
-0,707	-0,665	0,290	0,254	0,644
-0,677	-0,182	1,048	-0,134	-0,639
0,288	2,410	-0,636	-1,245	0,061
1,096	0,518	1,331	-1,278	-0,639
-0,738	-1,330	1,217	0,896	0,294
1,228	0,840	0,093	0,891	0,761
1,022	2,714	-0,687	1,182	-0,289
0,113	-0,681	-0,268	1,929	-0,522
-0,681	-0,457	0,699	-1,353	0,761
-0,229	-0,061	-0,607	0,771	1,227
0,731	-0,346	-0,363	1,325	0,061
-0,667	-1,583	-0,566	-0,364	-0,172
-0,515	-0,477	-0,038	1,205	-0,522
-0,531	-0,648	-0,656	2,889	0,994
-0,017	0,609	0,612	1,594	-0,989
-0,720	-0,751	-0,213	-1,302	0,177
0,883	0,685	-0,647	0,551	-0,056
-0,089	-0,882	-0,487	0,424	-0,172
-0,755	-0,500	-0,792	-0,659	-0,755
-0,562	0,358	-0,170	-1,597	-1,222
-0,821	-1,803	-0,487	-1,053	0,061
-0,208	-0,059	-0,346	-0,192	0,877
-0,500	-0,378	-0,736	1,640	-0,406
-0,053	-0,055	-0,653	-0,376	-0,639
-0,681	0,001	-0,205	-0,505	-0,522
-0,654	-1,709	-0,724	1,107	-0,172
0,031	-0,204	-0,105	-7,784	-0,406
-0,705	-0,281	0,121	0,163	-0,289
-0,236	0,148	-0,683	1,254	-0,056
0,244	-0,079	1,452	-0,639	-0,172
-0,687	-0,196	-0,121	0,040	2,160
3,201	1,921	-0,625	-0,199	0,294
0,141	-0,363	0,223	-1,171	-0,639
-0,625	-0,375	-0,247	0,302	0,294
0,946	1,930	-0,019	-0,417	0,061
-0,248	-0,294	-0,579	0,646	-1,455
-0,186	0,120	0,015	-0,038	-0,639
-0,650	0,706	0,370	0,555	-0,289
-0,075	2,314	-0,115	0,738	1,344
0,343	-1,242	1,753	-0,315	-1,455

-0,078	-1,456	-0,658	-0,286	0,527
1,986	-0,372	-0,376	-0,804	-0,755
-0,691	-0,329	-0,427	-0,337	1,460
-0,416	-0,076	-0,167	2,386	-0,872
-0,400	0,081	0,233	-0,929	-0,289
-0,241	-0,868	6,818	-0,576	-0,755
-0,003	-1,336	0,144	1,930	5,075
4,990	0,394	0,646	0,068	2,393
0,748	1,318	0,868	0,445	0,761
1,118	-1,201	-0,581	-0,362	0,527
1,204	-1,621	-0,467	-0,751	-0,522
-0,578	-0,437	-0,424	0,962	-0,522
-0,205	1,106	-0,552	1,176	0,061
-0,534	0,226	1,233	1,164	-0,755
-0,612	0,122	0,154	0,390	0,411
0,678	1,958	-0,238	1,074	0,644
0,039	1,055	-0,487	-0,773	0,177
-0,378	-0,741	-0,526	0,402	-0,056
-0,465	-0,145	-0,505	-0,741	-0,989
-0,542	-0,903	-0,534	-0,293	-0,639
-0,548	0,151	-0,738	1,043	-0,056
-0,480	-0,246	0,541	-0,682	-0,989
-0,715	0,105	1,725	0,907	-0,522
0,847	-1,444	1,360	-0,314	0,877
1,504	-0,773	0,896	0,923	-0,872
2,626	1,100	0,549	-0,657	-0,522
1,349	-0,591	0,868	0,974	-0,522

A présent, on calcule la matrice de Corrélation grâce à la formule

$$R = \frac{1}{n} * t_{AC} * A^C$$

On trouve alors la matrice suivante :

Variables	Nombre de délits	Taux de chômage	Nombre entreprises créées	Pourcentage d'étudiants	Nombre de musées
Nombre de délits enregistrés dans le département en 2010	1	0,330	0,237	0,091	0,343
Taux de chômage moyen de 1981 à 2007	0,330	1	-0,009	0,075	0,125
Nombre entreprises créées	0,237	-0,009	1	0,037	0,135
Pourcentage d'étudiants	0,091	0,075	0,037	1	0,300
Nombre de musées	0,343	0,125	0,135	0,300	1

Les valeurs en gras sont celles étant le plus éloignées de 0.

Calcul des valeurs propres

On calcule les valeurs propres de la matrice, en passant par le déterminant = 0 :

$$Det(\varepsilon - \lambda I) = \begin{vmatrix} 1 - \lambda & 0,330 & 0,237 & 0,091 & 0,343 \\ 0,330 & 1 - \lambda & 0,009 & 0,075 & 0,125 \\ 0,237 & 0,009 & 1 - \lambda & 0,037 & 0,135 \\ 0,091 & 0,075 & 0,037 & 1 - \lambda & 0,300 \\ 0,343 & 0,125 & 0,135 & 0,300 & 1 - \lambda \end{vmatrix} = 0$$

On obtient les valeurs propres suivantes, en utilisant un calculateur :

	λ_1	λ_2	λ_3	λ_4	λ_5
Valeur	1,729	1,043	0,995	0,707	0,526
Qualité (%)	34,583	20,852	19,903	14,139	10,523
Qualité cumulée (%)	34,583	55,435	75,337	89,477	100,000

Comme la qualité cumulée doit être supérieure à 70%, on choisit les 3 premiers axes : λ_1, λ_2 et λ_3 pour une qge = 75,337%.

Calcul des vecteurs propres

Un vecteur propre est déterminé pour chacun des valeurs propres. Ce sont les axes envisageables pour notre étude graphique.

De ces vecteurs propres, nous en déduisons la matrice de changement de base U :

$$U = \begin{pmatrix} 0,577 & -0,341 & -0,012 & 0,268 & 0,692 \\ 0,393 & -0,414 & -0,602 & -0,404 & -0,386 \\ 0,311 & -0,258 & 0,791 & -0,405 & -0,216 \\ 0,351 & 0,728 & -0,093 & -0,518 & 0,265 \\ 0,541 & 0,341 & 0,056 & 0,576 & -0,506 \end{pmatrix}$$

Calcul des coordonnées des points

On calcule ensuite la matrice des composantes principales, qui nous servira pour obtenir les coordonnées des points sur les axes.

La formule est la suivante :

$$F = A^S * U$$

Une fois cette matrice obtenue, on détermine la matrice de saturation qui contiendra les coordonnées des points :

$$S = \frac{1}{n} t_{A^S} * F * \Delta^{-\frac{1}{2}}$$

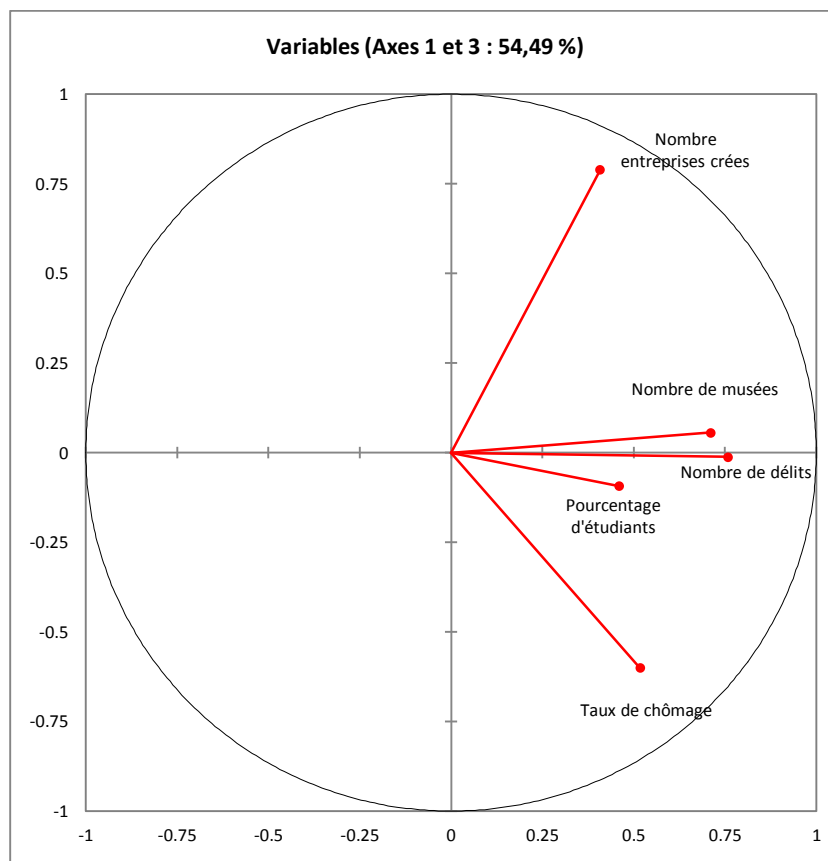
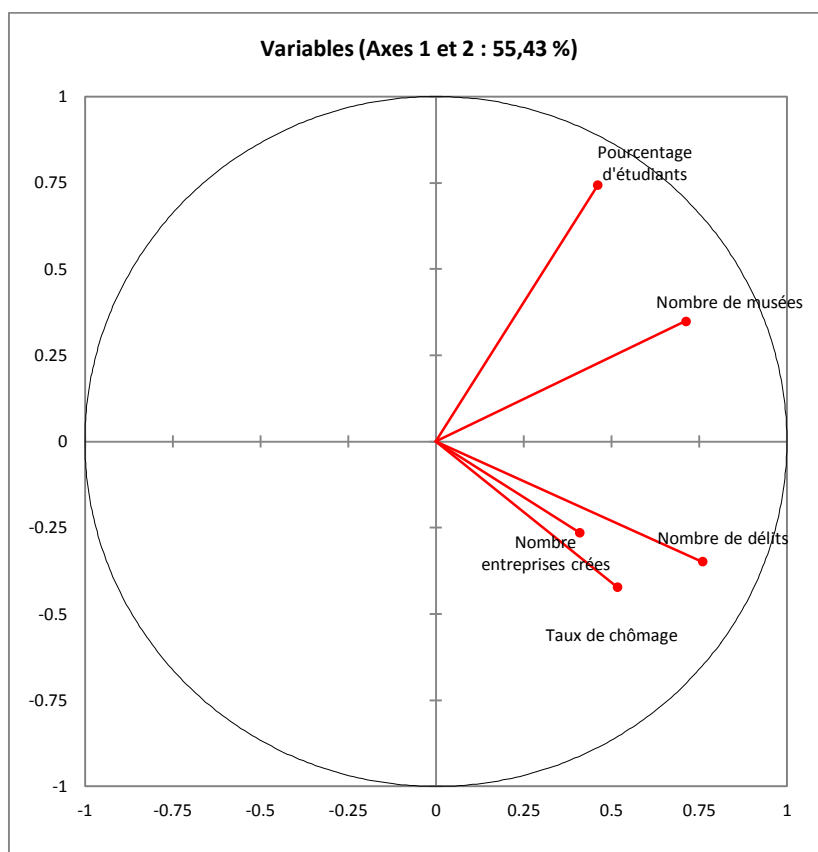
Dans cette formule, Δ est la matrice diagonale suivante, contenant les valeurs propres par ordre décroissantes :

$$\Delta = \begin{pmatrix} 1,729 & 0 & 0 & 0 & 0 \\ 0 & 1,043 & 0 & 0 & 0 \\ 0 & 0 & 0,995 & 0 & 0 \\ 0 & 0 & 0 & 0,707 & 0 \\ 0 & 0 & 0 & 0 & 0,526 \end{pmatrix}$$

On obtient alors les données suivantes :

	Axe 1	Axe 2	Axe 3	Axe 4	Axe 5
Nombre de délits	0,577	-0,341	-0,012	0,268	0,692
Taux de chômage	0,393	-0,414	-0,602	-0,404	-0,386
Nombre entreprises créées	0,311	-0,258	0,791	-0,405	-0,216
Pourcentage d'étudiants	0,351	0,728	-0,093	-0,518	0,265
Nombre de musées	0,541	0,341	0,056	0,576	-0,506

Représentations graphiques



Cordonnées des individus

Le placement de chaque individu par rapport à chaque axe demande la connaissance de ses coordonnées. Pour cela, on utilise la formule :

$$F = A^s * U$$

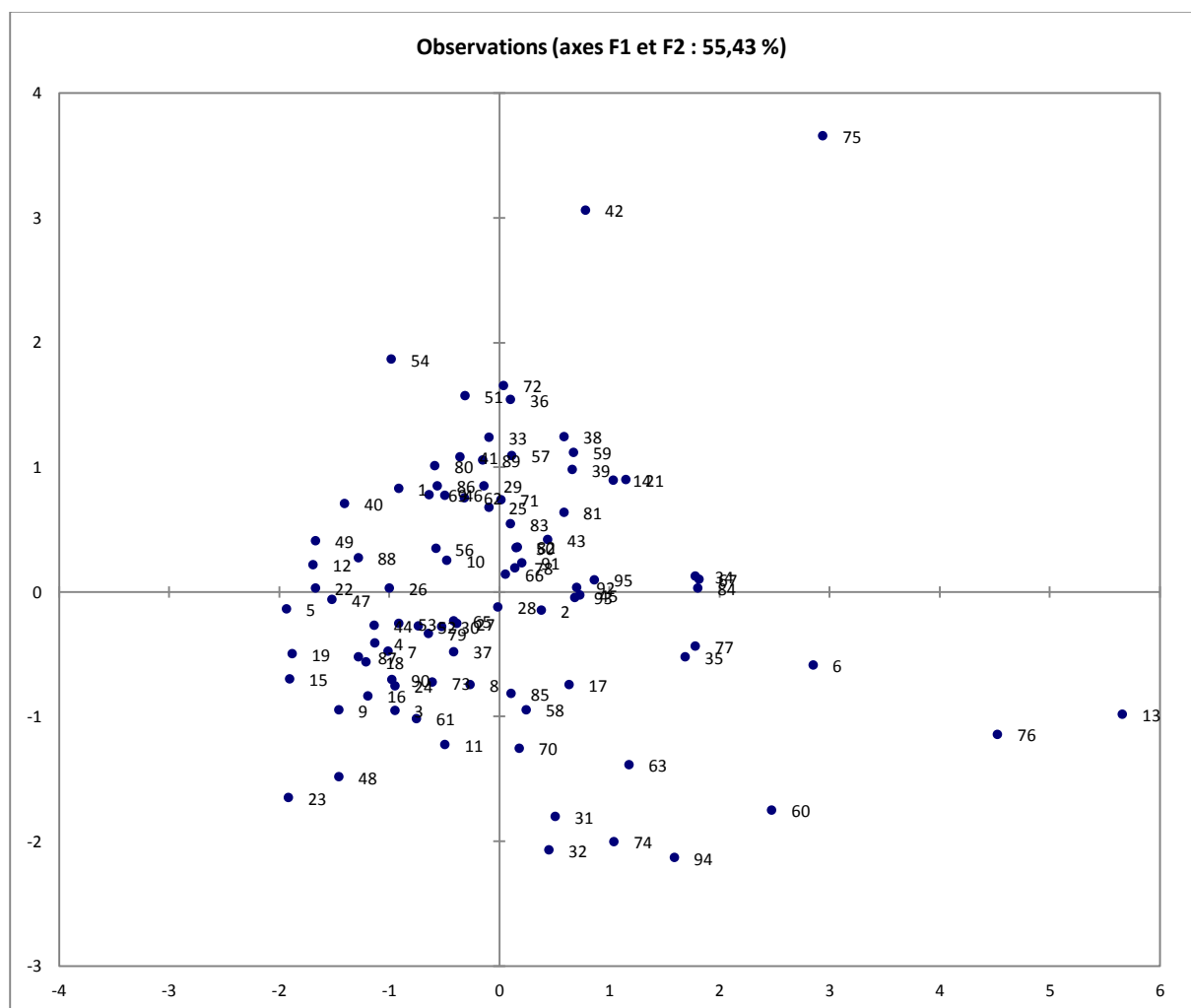
On obtient la matrice suivante :

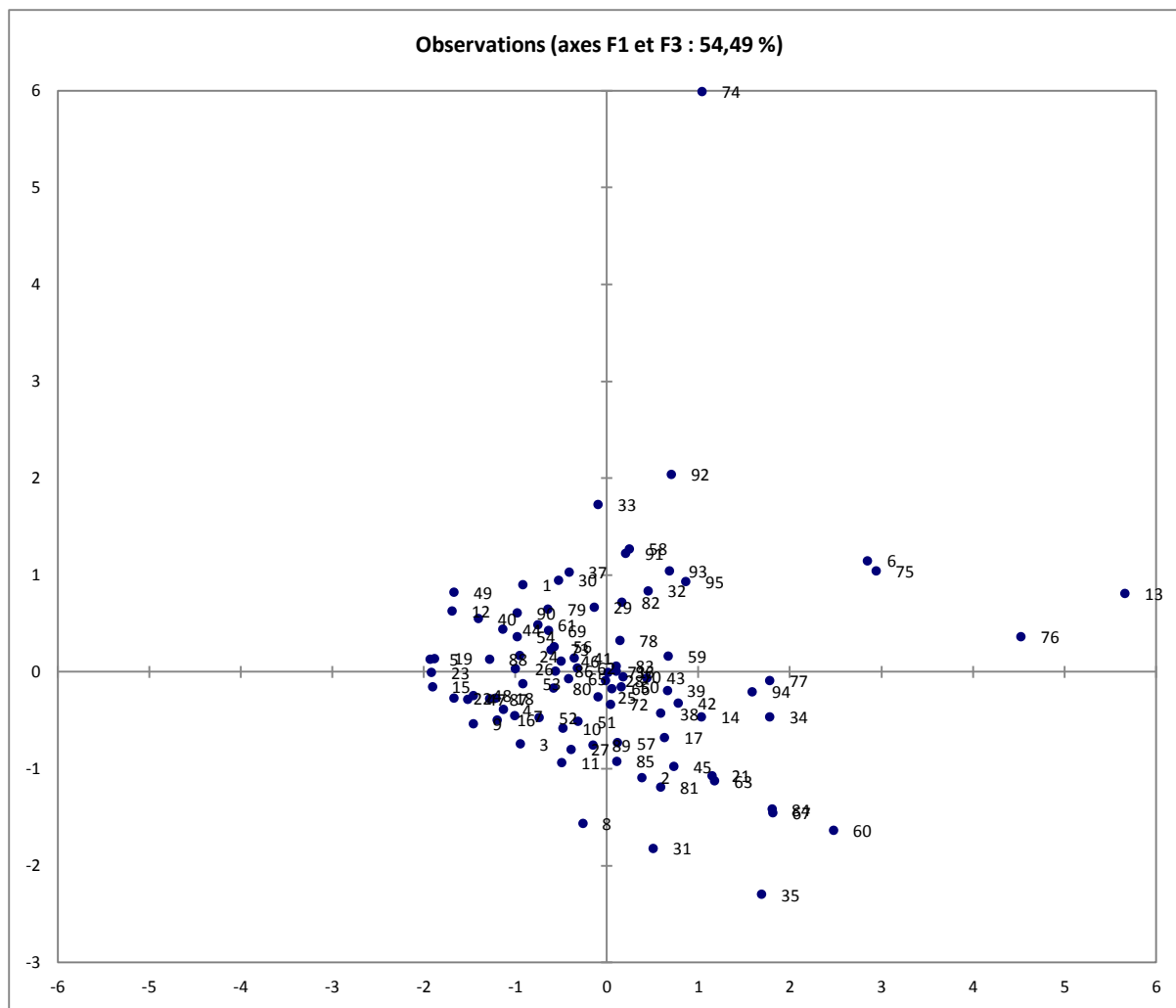
Département	F1	F2	F3	F4	F5
1	-0,916	0,827	0,902	0,918	0,382
2	0,379	-0,150	-1,092	-0,274	-0,701
3	-0,949	-0,951	-0,744	0,112	-0,569
4	-1,132	-0,412	-0,388	0,299	-0,380
5	-1,932	-0,139	0,133	0,134	0,434
6	2,848	-0,589	1,150	0,728	-0,890
7	-1,011	-0,476	-0,449	-0,051	-0,237
8	-0,265	-0,744	-1,563	-0,887	-0,696
9	-1,458	-0,949	-0,531	0,273	-0,443
10	-0,480	0,252	-0,579	-0,110	-0,183
11	-0,494	-1,227	-0,935	-0,370	-0,608
12	-1,692	0,216	0,628	1,169	0,086
13	5,660	-0,984	0,813	0,474	-0,553
14	1,032	0,895	-0,461	-0,072	-0,613
15	-1,907	-0,701	-0,150	0,587	-0,126
16	-1,197	-0,836	-0,496	-0,007	-0,174
17	0,631	-0,744	-0,674	0,260	-1,102
18	-1,212	-0,564	-0,267	0,514	-0,310
19	-1,884	-0,498	0,141	0,632	0,118
21	1,147	0,898	-1,069	-0,561	-0,883
22	-1,673	0,028	-0,269	-0,107	0,274
23	-1,917	-1,651	-0,002	0,491	-0,179
24	-0,951	-0,757	0,170	1,212	-0,471
25	-0,095	0,679	-0,256	-0,856	0,396
26	-1,001	0,032	0,035	-0,290	0,589
27	-0,389	-0,252	-0,799	-0,376	-0,292
28	-0,013	-0,124	-0,085	0,209	-0,212
29	-0,141	0,848	0,672	0,201	-0,541
30	-0,527	-0,279	0,950	-0,835	-0,325
31	0,506	-1,805	-1,819	0,041	-0,944
32	0,450	-2,071	0,840	-0,170	0,256
33	-0,095	1,238	1,732	-0,451	-0,161
34	1,780	0,126	-0,464	-0,074	0,357
35	1,687	-0,523	-2,296	-1,322	0,272
36	0,101	1,540	0,006	-0,885	1,182
37	-0,415	-0,480	1,029	0,854	-1,177
38	0,585	1,244	-0,427	0,520	-0,410

39	0,660	0,979	-0,190	-0,169	1,042
40	-1,405	0,709	0,552	0,781	0,275
41	-0,360	1,082	0,142	-0,854	0,431
42	0,781	3,059	-0,320	-0,533	0,299
43	0,439	0,419	-0,061	-1,896	0,553
44	-1,136	-0,270	0,441	0,973	-0,583
45	0,729	-0,027	-0,973	-0,096	0,665
46	-0,498	0,775	0,114	0,211	0,593
47	-1,519	-0,064	-0,283	0,229	0,064
48	-1,458	-1,485	-0,244	-0,105	-0,281
49	-1,669	0,411	0,826	1,287	-0,063
50	0,153	0,351	-0,151	0,714	-0,530
51	-0,315	1,573	-0,509	-0,762	0,611
52	-0,738	-0,276	-0,468	0,100	0,360
53	-0,916	-0,256	-0,118	-0,139	-0,283
54	-0,982	1,865	0,365	0,139	0,757
56	-0,579	0,348	0,263	-0,375	-0,202
57	0,114	1,093	-0,731	-0,525	0,299
58	0,242	-0,945	1,272	-0,267	-0,191
59	0,672	1,117	0,167	1,170	-1,443
60	2,474	-1,750	-1,635	0,602	1,396
61	-0,752	-1,018	0,487	0,329	0,210
62	-0,321	0,755	0,044	0,099	-0,290
63	1,176	-1,388	-1,123	-0,270	-0,223
65	-0,418	-0,232	-0,070	-0,454	0,145
66	0,055	0,140	-0,169	-1,063	-0,495
67	1,812	0,099	-1,453	-0,515	-1,393
69	-0,641	0,777	0,427	1,287	0,316
70	0,177	-1,255	-0,047	0,810	1,764
71	0,011	0,740	-0,001	1,139	-1,073
72	0,038	1,654	-0,333	-1,749	0,863
73	-0,609	-0,727	0,232	0,078	-0,446
74	1,041	-2,005	5,994	-2,637	-1,073
75	2,938	3,655	1,043	2,407	-1,568
76	4,525	-1,143	0,368	2,246	1,941
77	1,780	-0,435	-0,087	-0,480	-0,442
78	0,140	0,193	0,328	1,510	1,001
79	-0,644	-0,333	0,648	1,251	1,625
80	-0,586	1,011	-0,167	-0,603	0,392
81	0,584	0,635	-1,188	-0,850	-0,157
82	0,164	0,358	0,723	-1,775	-0,021
83	0,101	0,545	0,064	-0,241	-0,596
84	1,802	0,029	-1,417	-0,698	-0,270
85	0,106	-0,817	-0,921	0,284	-0,559
86	-0,566	0,848	0,010	0,173	0,285
87	-1,279	-0,522	-0,277	-0,047	0,160

88	-1,284	0,270	0,135	0,221	0,348
89	-0,155	1,058	-0,753	-0,478	0,040
90	-0,980	-0,705	0,613	-0,468	-0,023
91	0,202	0,234	1,228	-1,707	-0,391
92	0,704	0,034	2,041	0,919	0,322
93	0,682	-0,044	1,047	-0,635	1,828
94	1,588	-2,133	-0,202	0,066	1,355
95	0,863	0,097	0,932	-0,563	1,494

Cette matrice nous permet d'obtenir les graphiques suivants, qui nous serviront à finir l'analyse :





Interprétation de chaque axe

Chaque graphique nous permet d'identifier la signification de chaque axe. On peut se baser sur le graphique pour déterminer l'importance de chaque donnée par rapport à l'axe. Ainsi nous pouvons donner une signification à chaque axe.

Comme on le voit sur le graphique représentant les axes 1 et 2, le nombre d'entreprises créées est vers le centre. Nous pouvons déduire de cette observation que cette donnée n'a pas une forte influence sur notre étude comparée aux autres.

Sur les axes 1 et 3 cependant, on peut voir que c'est la donnée correspondant au pourcentage d'étudiants qui n'est pas en bordure du graphique. On peut, comme précédemment, conclure que pour ces 2 axes, cette donnée a une influence assez faible comparée aux 4 autres.

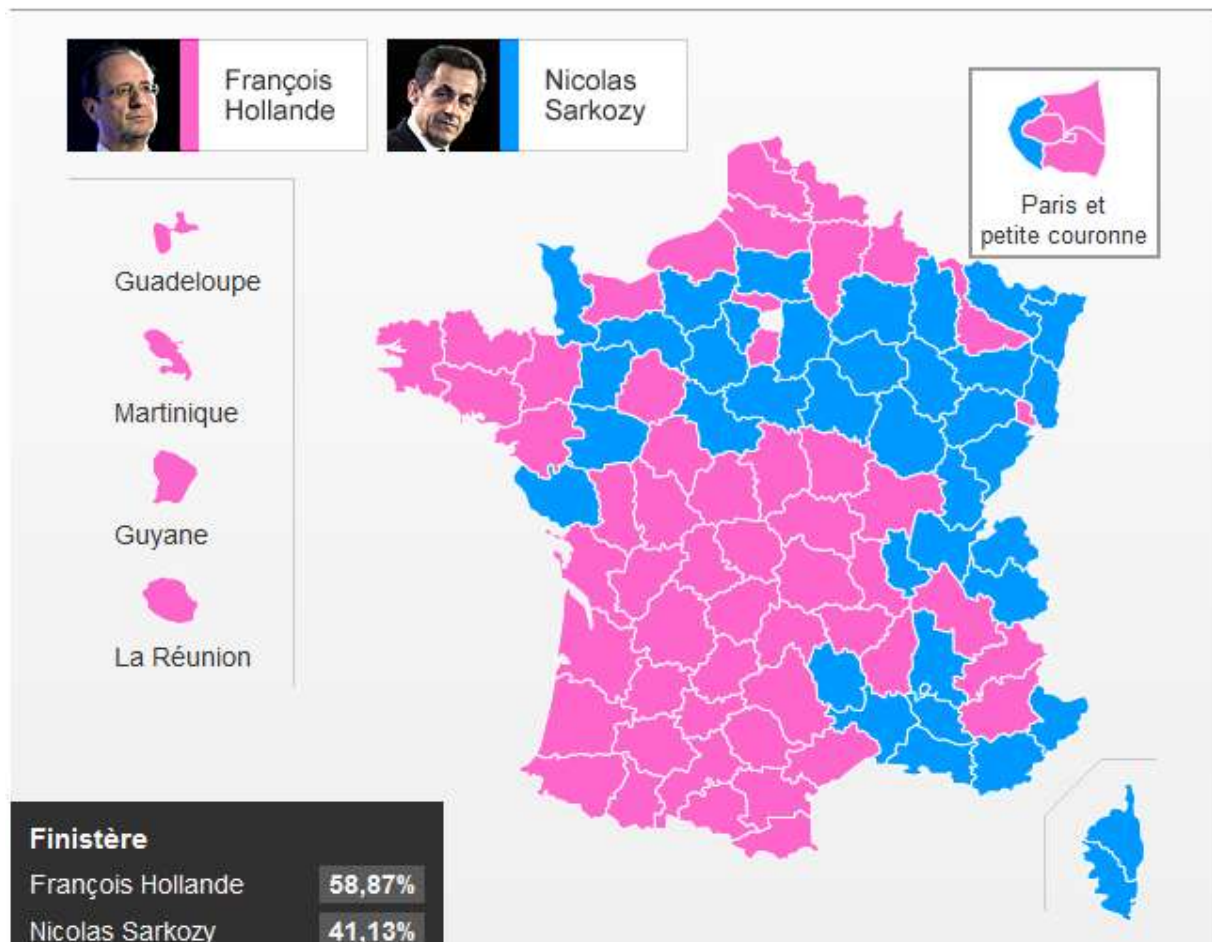
L'axe 1 oppose le pourcentage d'étudiant, élevé dans des départements comme Paris, la Loire et Meurthe-et-Moselle, au taux de chômage élevé d'autres départements comme le Val-de-Marne, la Haute-Savoie, la Haute-Garonne, et le Gers.

L'axe 2 oppose le nombre d'entreprise créé qui est important dans des départements comme Paris, les Hauts-de-Seine et la Gironde au taux de chômage élevé d'autres départements comme l'Ille-et-Vilaine, la Haute-Garonne et l'Oise.

Le premier graphique nous permet de mettre en relation le Nombre de délits avec le nombre d'entreprises créées. Les entreprises sont créées en majorité dans les grandes villes qui sont sujettes à la délinquance.

Le second graphique permet de mettre en relation le nombre de musées avec le pourcentage d'étudiants et le nombre de délit. On peut expliquer cela par le fait que les villes étudiantes sont en générales de grandes agglomérations, et les grandes agglomérations sont les villes les plus sujettes à la délinquance. Mais ces agglomérations sont souvent les mieux équipées dans le domaine de la culture et donc des musées.

Ouverture sur les résultats de la présidentielle de 2012.



Dimanche 6 mai, 61 départements français sur 101 ont accordé une majorité de leurs suffrages à François Hollande. En métropole, c'est en Seine-Saint-Denis que le nouveau président obtient son meilleur score (65,3%), devant la Corrèze (64,9%) et l'Ariège (64,7%). Viennent ensuite la Haute-Vienne, les Hautes-Pyrénées, la Guyane, le Lot, la Creuse et le Puy-de-Dôme.

La liste des départements où Hollande recule par rapport au score de Mitterrand en 1981 est intéressante à analyser. Près de 11 points de perdus dans le Var et le Vaucluse, -10 dans les Alpes-Maritimes, -9 dans les Bouches-du-Rhône et le Gard, -7 dans l'Aube, l'Aude et l'Oise, -6 en Haute-Marne et dans le Territoire de Belfort, -5 points dans les Pyrénées-Orientales, le Lot-et-Garonne, l'Hérault et l'Ain... En trois décennies, la gauche a fortement reculé sur toute la façade méditerranéenne, où la population a beaucoup progressé en s'enrichissant nettement, ainsi que dans le nord-est de la France, du bassin parisien aux portes de l'Alsace. Ce sont les départements pour lesquels la qualité de la vie a fortement augmentée.

Dans toute la France, François Hollande a eu du succès dans les grandes villes, c'est-à-dire généralement les départements ayant une qualité de vie plus importante. Nicolas Sarkozy

réalise ses meilleurs scores dans des départements traditionnellement ancrés très à droite, ayant souvent un niveau de vie plus important.

Synthèses personnelles

Aurélien

Ce projet de nous a permis d'utiliser une méthode du cours, l'ACP de façon complète et ainsi d'appliquer une démarche théorique dans un cas concret. Je me suis particulièrement intéressé à la détermination des composantes principales, ainsi qu'à la qualité de chacun des individus et à la justification du choix de l'étude et des données. Le sujet choisi était particulièrement intéressant à étudier puisqu'il nous a permis d'observer quels étaient les départements où il faisait bon vivre. Cette étude m'a donné envie de découvrir de nouveaux départements de la France, comme par exemple le Gers et la Haute-Garonne.

Florian

Durant ce projet nous avons dû effectuer une ACP complète, comme vu à la fois en cours et en TD. Nous avons donc appliqué les mêmes méthodes tout du long. Je me suis personnellement occupé du calcul des valeurs propres. La difficulté principale de cette étape a été la taille de la matrice ($5 * 5$) et la difficulté à la factoriser pour obtenir les valeurs propres. N'ayant pu le faire, je me suis servi d'un calculateur qui m'a donné les valeurs recherchées. Les données choisies ont été plutôt intéressantes puisqu'elles nous permettaient de nous rendre compte, selon quelques critères précis, dans quels départements la qualité de vie est la plus haute.

Adrien

Ce TAI nous a permis de mettre en pratique dans un cas concret le cours d'analyse de donnée vu ce semestre et surtout l'ACP, l'analyse en composante principale. J'étais chargé du calcul des vecteurs propres, étant donné que cette opération est habituellement réalisée avec des valeurs moins complexes et une matrice moins grande, j'ai dû m'aider d'un calculateur pour trouver les vecteurs. Les chiffres que nous avons choisis porte sur la qualité de vie des différents départements, ce qui nous a permis de déjà choisir l'endroit parfait pour y faire sa vie.

Maxime

Durant ce TAI, je me suis occupé de commencer l'ACP. C'est-à-dire, dans un premier temps d'analyser et choisir de réaliser l'ACP avec une matrice de Variance-covariance ou de Corrélation, puis dans un deuxième temps de la calculer. Cette matrice a servi, plus tard, aux autres membres du groupe dans la suite de l'analyse. Les calculs sont longs à faire mais simple à réaliser.