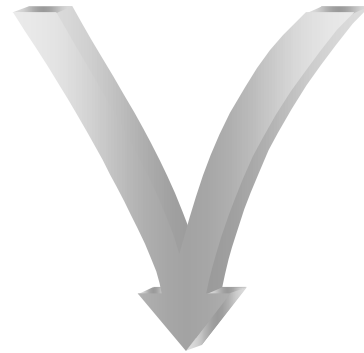


A.F.C.

Analyse Factorielle des Correspondances



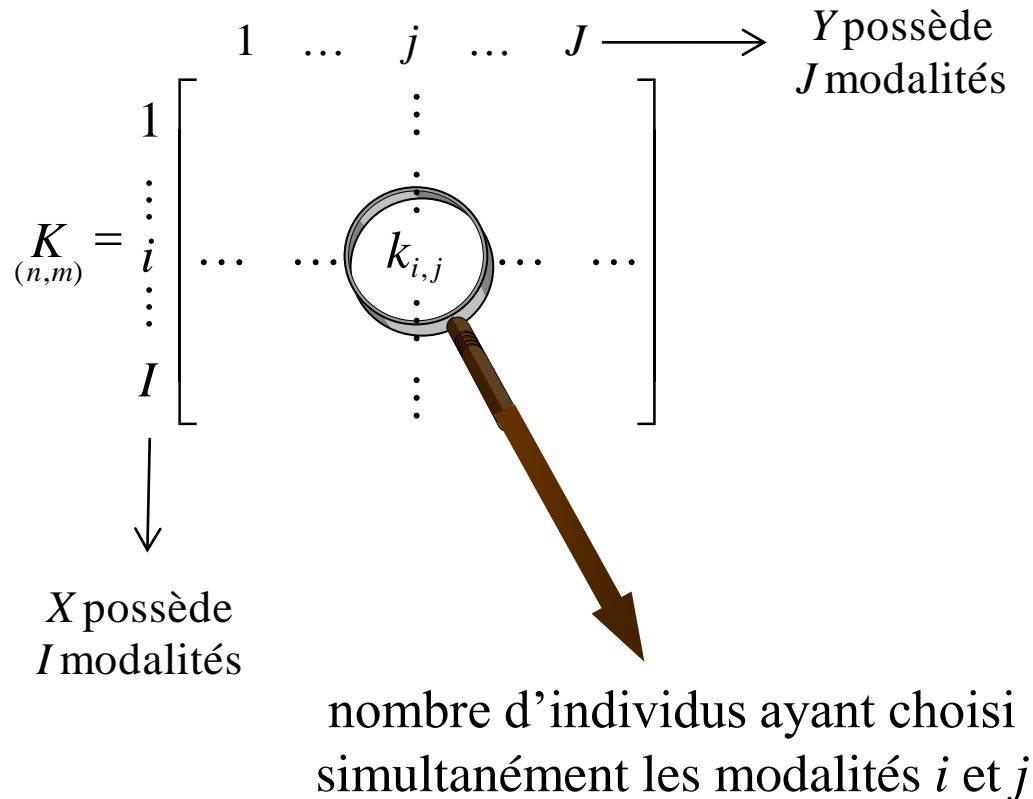
Généralisation de l'A.C.P. adaptée au traitement de **données qualitatives** se présentant sous la forme d'un **tableau de contingence**.

Le tableau de données initial

A.F.C.

Soient X et Y deux variables qualitatives ayant respectivement I et J modalités.

Le tableau de contingence K formé à partir de ces deux variables aura autant de lignes (colonnes) que la variable X a de modalités (I) et autant de colonnes (lignes) que la variable Y a de modalités (J).



Le tableau de données initial

A.F.C.

Exemple : Résultats du Baccalauréat 76

	A	B	C	D	E	F	G	H	Effectifs marginaux
La Basse-Normandie	4 446	202	100	100	100	100	100	100	4 446

La Basse-Normandie représente $4\,446/202\,100=2.20\%$ au niveau national
 et le Bac E $100/4\,446=2.25\%$ des bacs de cette région

Centre	1 482	667	1 020	1 535	173	629	989	26	6 521
Basse-Normandie	1 033	509	553	1 063	100	433	742	13	4 446
Bourgogne	1 272	527	861	1 116	219	769	1 232	13	6 009
Nord-Pas de Calais	2 549	1 141	2 164	2 752	587	1 660	1 951	41	12 845
Lorraine									8 903
Alsace									5 688
Franche-Comté									3 757
Pays de la Loire									10 140
Bretagne									10 245
Poitou-Charentes									5 505
Aquitaine	2 757	873	1 466	2 296	215	789	1 459	17	9 872
Midi-Pyrénées	2 493	1 120	1 494	2 329	254	855	1 565	28	10 138
Limousin	551	297	386	663	67	334	378	12	2 688
Rhône-Alpes	3 951	2 127	3 218	4 743	545	2 072	3 018	36	19 710
Le bac E					5 333				4 871
et la région Basse-Normandie					100				7 287
Corse	327	31	85	178	9	27	79	0	15 625

Les individus n'apparaissent qu'au travers de leurs effectifs

Effectifs marginaux	45 593	21 563	32 738	46 017	5 333	19 656	30 749	451	202 100
----------------------------	--------	--------	--------	--------	--------------	--------	--------	-----	----------------

Dans une AFC, les lignes et les colonnes jouent le même rôle

L'AFC consiste à considérer successivement les lignes et les colonnes comme les individus d'une ACP (les colonnes et les lignes étant successivement les variables)

AFC

=

double ACP

(sur les profils lignes
et les profils colonnes)

Méthodologie de calcul

A.F.C.

Le nombre de valeurs propres est égal au minimum entre le nombre de lignes moins 1 et le nombre de colonnes moins 1

$$\text{Min}(I-1, J-1) = \text{Min}(22-1, 8-1) = 7$$

Numéro	Valeur propre	Pourcent.	Pourcent. Cumulé
1	0.0121	56.04	56.04
2	0.0052	24.26	80.30
3	0.0017	7.84	88.14
4	0.0011	5.26	93.41
5	0.0008	3.49	96.89
6	0.0005	2.52	99.42
7	0.0001	0.58	100.00

Toutes les valeurs propres sont comprises entre 0 et 1

Méthodologie de calcul

A.F.C.

Le nombre de valeurs propres est égal au minimum entre le nombre de lignes moins 1 et le nombre de colonnes moins 1

$$\text{Min}(I-1, J-1) = \text{Min}(22-1, 8-1) = 7$$

Numéro	Valeur propre	Pourcent.	Pourcent. Cumulé
1	0.0121	56.04	56.04
2	0.0052	24.26	80.30

Pour $\alpha=5\%$, $\chi^2(147)=176.3$

5	0.0008	3.49	96.89
6	0.0005	2.52	99.42
7	0.0001	0.58	100.00

Somme	0.0215
Effectif	202100
Produit	4345

$\sim \chi^2$ à $(22-1)*(8-1)=147$ ddl

La somme des valeurs propres multipliée par le nombre d'individus est égale au χ^2 calculé du test du même nom.

Méthodologie de calcul

A.F.C.

Identificateur
Ile-de-France
Champagne-Ardennes
Picardie
Haute-Normandie
Centre
Basse-Normandie
Bourgogne
Nord-Pas de Calais
Lorraine
Alsace
Franche-Comté
Pays de la Loire
Bretagne
Poitou-Charentes
Aquitaine
Midi-Pyrénées
Limousin
Rhône-Alpes
Auvergne
Languedoc-Roussillon
Provence-Alpes-Côte d'Azur
Corse

Poids relatif	Distance à l'origine
21.33	0.02635
2.10	0.00944
2.43	0.00510
2.40	0.02808
3.23	0.00204
2.20	0.01229
2.97	0.04605
6.36	0.03237
4.41	0.04984
2.81	0.07061
1.86	0.02385
5.02	0.01775
5.07	0.00654
2.72	0.02367
4.88	0.02126
5.02	0.00524
1.33	0.01637
9.75	0.00434
2.41	0.00820
3.61	0.02008
7.73	0.00807
0.36	0.32607

Coordonnées

Axe 1	Axe 2
-0.14	-0.07
0.06	0.03
-0.03	0.02
0.07	-0.02
0.01	0.02
0.03	0.07
0.19	-0.01
0.13	-0.08
0.21	-0.04
0.24	-0.07
0.13	0.04
0.07	0.08
-0.01	-0.03
0.06	0.14
-0.04	0.12
-0.03	0.06
0.06	-0.03
0.03	-0.03
-0.03	0.00
-0.12	0.08
-0.05	0.07
-0.19	0.45

Contributions

Axe 1	Axe 2
36.01	22.40
0.64	0.49
0.20	0.18
1.00	0.26
0.01	0.21
0.12	1.98
9.22	0.05
9.15	7.51
16.20	1.20
13.12	2.75
2.60	0.54
2.18	5.88
0.07	1.07
0.72	9.67
0.73	14.49
0.33	3.17
0.36	0.18
0.53	2.09
0.15	0.01
3.96	4.23
1.56	7.24
1.12	14.41

Cosinus carrés

Axe 1	Axe 2
0.77	0.21
0.39	0.13
0.20	0.07
0.18	0.02
0.02	0.16
0.05	0.38
0.81	0.00
0.54	0.19
0.89	0.03
0.80	0.07
0.71	0.06
0.30	0.35
0.03	0.17
0.13	0.78
0.08	0.73
0.15	0.63
0.20	0.04
0.15	0.26
0.09	0.00
0.66	0.31
0.30	0.61
0.11	0.63

Méthodologie de calcul

$$\frac{43115}{202100} = 21.33\%$$

	A	B	C	D	E	F	G	H	Effectifs marginaux
Ile-de-France	9 724	5 650	8 679	9 432	839	3 353	5 355	83	43 115
Corse	327	31	85	178	9	27	79	0	736
Effectifs marginaux	45 593	21 563	32 738	46 017	5 333	19 656	30 749	451	202 100

$$\frac{327}{736} \quad \frac{45593}{202100}$$

Poids relatif

	A	B	C	D	E	F	G	H	
Total	22.56%	10.67%	16.20%	22.77%	2.64%	9.73%	15.21%	0.22%	
Corse	44.43%	4.21%	11.55%	24.18%	1.22%	3.67%	10.73%	0.00%	
	0.0478	0.0042	0.0022	0.0002	0.0002	0.0037	0.0020	0.0000	Total
	0.2120	0.0391	0.0133	0.0009	0.0076	0.0377	0.0132	0.0022	0.32607

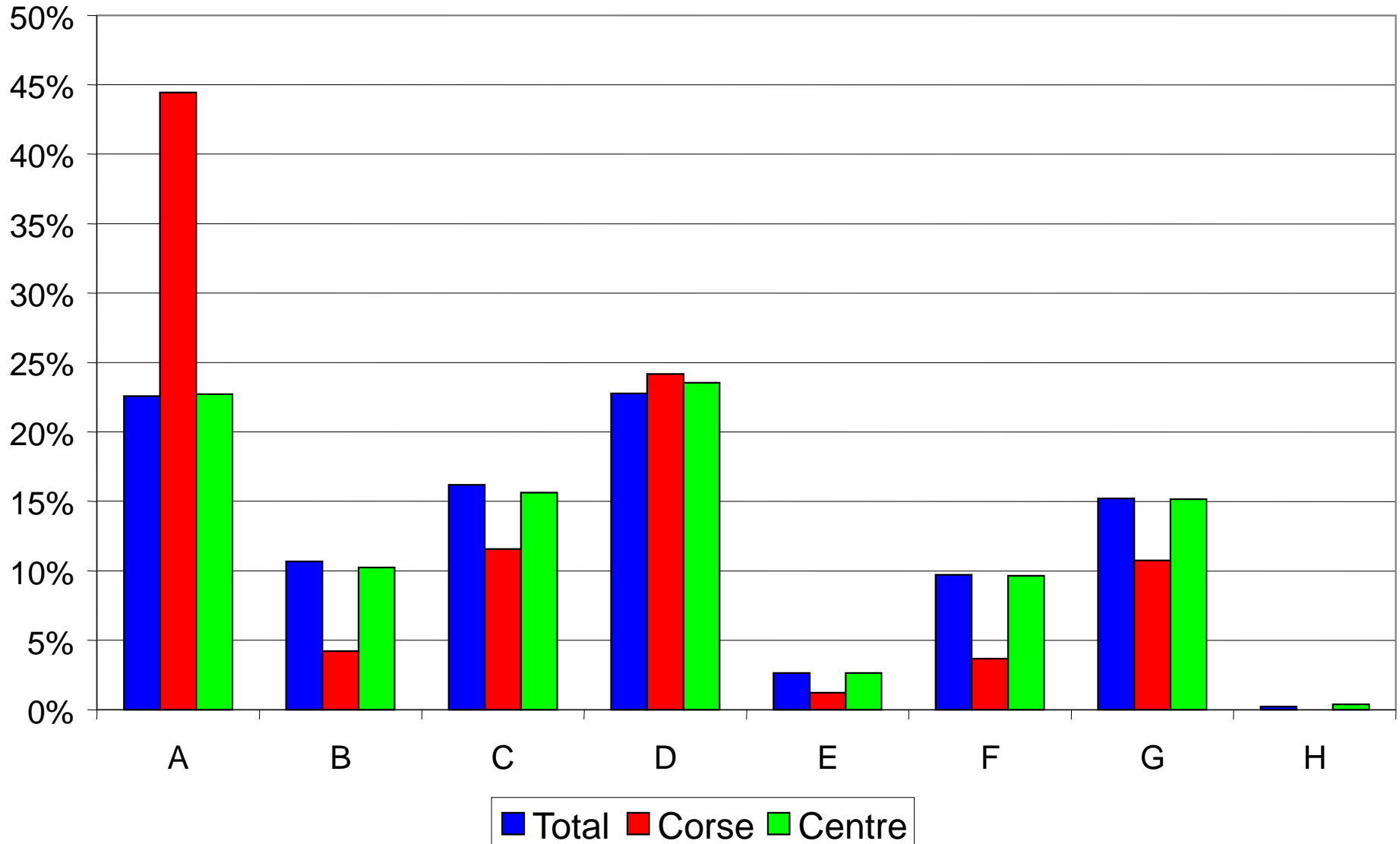
Distance à l'origine

$$\frac{(0.2256 - 0.4443)^2}{0.2256}$$

Centre	22.73%	10.23%	15.64%	23.54%	2.65%	9.65%	15.17%	0.40%	
	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	Total
	0.0000	0.0002	0.0002	0.0003	0.0000	0.0000	0.0000	0.0014	0.00204

Méthodologie de calcul

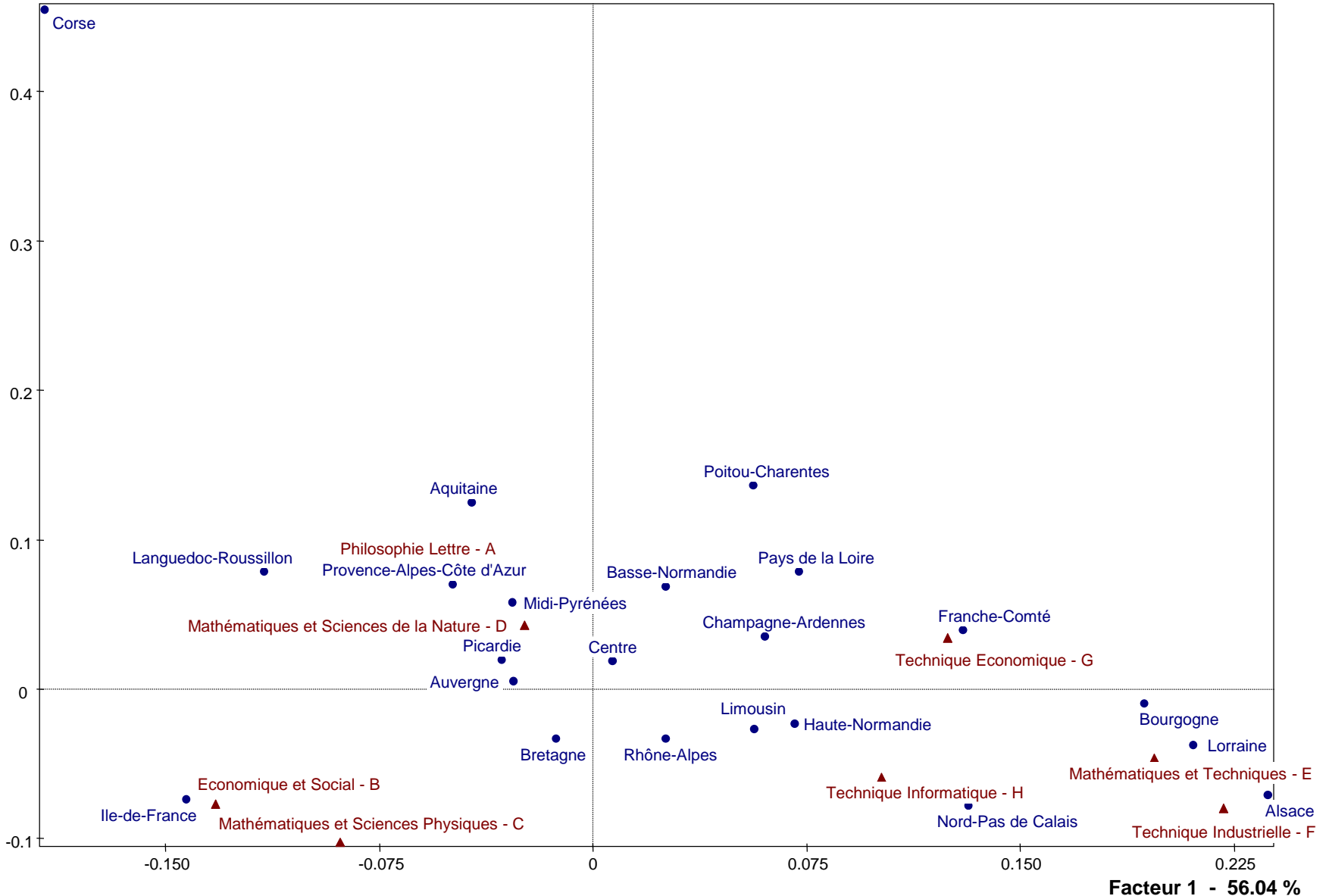
A.F.C.



Méthodologie de calcul

Représentation graphique

Facteur 2 - 24.26 %



Méthodologie de calcul

A.F.C.

Identificateur
Ile-de-France
Champagne-Ardennes
Picardie
Haute-Normandie
Centre
Basse-Normandie
Bourgogne
Nord-Pas de Calais
Lorraine
Alsace
Franche-Comté
Pays de la Loire
Bretagne
Poitou-Charentes
Aquitaine
Midi-Pyrénées
Limousin
Rhône-Alpes
Auvergne
Languedoc-Roussillon
Provence-Alpes-Côte d'Azur
Corse

Poids relatif	Distance à l'origine
21.33	0.02635
2.10	0.00944
2.43	0.00510
2.40	0.02808
3.23	0.00204
2.20	0.01229
2.97	0.04605
6.36	0.03237
4.41	0.04984
2.81	0.07061
1.86	0.02385
5.02	0.01775
5.07	0.00654
2.72	0.02367
4.88	0.02126
5.02	0.00524
1.33	0.01637
9.75	0.00434
2.41	0.00820
3.61	0.02008
7.73	0.00807
0.36	0.32607

Coordonnées

Axe 1	Axe 2
-0.14	-0.07
0.06	0.03
-0.03	0.02
0.07	-0.02
0.01	0.02
0.03	0.07
0.19	-0.01
0.13	-0.08
0.21	-0.04
0.24	-0.07
0.13	0.04
0.07	0.08
-0.01	-0.03
0.06	0.14
-0.04	0.12
-0.03	0.06
0.06	-0.03
0.03	-0.03
-0.03	0.00
-0.12	0.08
-0.05	0.07
-0.19	0.45

Contributions

Axe 1	Axe 2
36.01	22.40
0.64	0.49
0.20	0.18
1.00	0.26
0.01	0.21
0.12	1.98
9.22	0.05
9.15	7.51
16.20	1.20
13.12	2.75
2.60	0.54
2.18	5.88
0.07	1.07
0.72	9.67
0.73	14.49
0.33	3.17
0.36	0.18
0.53	2.09
0.15	0.01
3.96	4.23
1.56	7.24
1.12	14.41

Axe 1	Axe 2
1.69	1.05
0.30	0.23
0.08	0.07
0.42	0.11
0.00	0.06
0.06	0.90
3.10	0.02
1.44	1.18
3.68	0.27
4.66	0.98
1.40	0.29
0.44	1.17
0.01	0.21
0.26	3.55
0.15	2.97
0.07	0.63
0.27	0.14
0.05	0.21
0.06	0.00
1.10	1.17
0.20	0.94
3.07	39.56

Méthodologie de calcul

A.F.C.

Identificateur
Ile-de-France
Champagne-Ardennes
Picardie
Haute-Normandie
Centre
Basse-Normandie
Bourgogne
Nord-Pas de Calais
Lorraine
Alsace
Franche-Comté
Pays de la Loire
Bretagne
Poitou-Charentes
Aquitaine
Midi-Pyrénées
Limousin
Rhône-Alpes
Auvergne
Languedoc-Roussillon
Provence-Alpes-Côte d'Azur
Corse

Poids relatif	Distance à l'origine
21.33	0.02635
2.10	0.00944
2.43	0.00510
2.40	0.02808
3.23	0.00204
2.20	0.01229
2.97	0.04605
6.36	0.03237
4.41	0.04984
2.81	0.07061
1.86	0.02385
5.02	0.01775
5.07	0.00654
2.72	0.02367
4.88	0.02126
5.02	0.00524
1.33	0.01637
9.75	0.00434
2.41	0.00820
3.61	0.02008
7.73	0.00807
0.36	0.32607

Coordonnées

Axe 1	Axe 2
-0.14	-0.07
0.06	0.03
-0.03	0.02
0.07	-0.02
0.01	0.02
0.03	0.07
0.19	-0.01
0.13	-0.08
0.21	-0.04
0.24	-0.07
0.13	0.04
0.07	0.08
-0.01	-0.03
0.06	0.14
-0.04	0.12
-0.03	0.06
0.06	-0.03
0.03	-0.03
-0.03	0.00
-0.12	0.08
-0.05	0.07
-0.19	0.45

Contributions

Axe 1	Axe 2
36.01	22.40
0.64	0.49
0.20	0.18
1.00	0.26
0.01	0.21
0.12	1.98
9.22	0.05
9.15	7.51
16.20	1.20
13.12	2.75
2.60	0.54
2.18	5.88
0.07	1.07
0.72	9.67
0.73	14.49
0.33	3.17
0.36	0.18
0.53	2.09
0.15	0.01
3.96	4.23
1.56	7.24
1.12	14.41

Cosinus carrés

Axe 1	Axe 2
0.77	0.21
0.39	0.13
0.20	0.07
0.18	0.02
0.02	0.16
0.05	0.38
0.81	0.00
0.54	0.19
0.89	0.03
0.80	0.07
0.71	0.06
0.30	0.35
0.03	0.17
0.13	0.78
0.08	0.73
0.15	0.63
0.20	0.04
0.15	0.26
0.09	0.00
0.66	0.31
0.30	0.61
0.11	0.63

$$1/22=4.55 \%$$

Méthodologie de calcul

A.F.C.

Libellé de la variable	Poids relatif	Distance à l'origine	Coordonnées		Contributions		Cosinus carrés	
			Axe 1	Axe 2	Axe 1	Axe 2	Axe 1	Axe 2
Philosophie Lettre - A	22.56	0.01231	-0.05	0.08	5.18	30.67	0.23	0.58
Economique et Social - B	10.67	0.02904	-0.13	-0.08	15.53	11.98	0.61	0.20
Mathématiques et Sciences Physiques - C	16.20	0.02090	-0.09	-0.10	10.58	32.81	0.38	0.51
Mathématiques et Sciences de la Nature - D	22.77	0.00650	-0.02	0.04	1.08	8.00	0.09	0.28
Mathématiques et Techniques - E	2.64	0.07497	0.20	-0.05	8.47	1.08	0.52	0.03
Technique Industrielle - F	9.73	0.05788	0.22	-0.08	39.47	11.90	0.85	0.11
Technique Economique - G	15.21	0.01934	0.12	0.03	19.48	3.40	0.80	0.06
Technique Informatique - H	0.22	0.11392	0.10	-0.06	0.19	0.15	0.09	0.03

$$1/8=12.5 \%$$

Résultats sans la Corse

A.F.C.

Numéro	Valeur propre	Pourcentage	Pourcentage cumulé
1	0.0120	58.71	58.71
2	0.0045	22.25	80.96
3	0.0016	7.74	88.70
4	0.0010	4.76	93.46
5	0.0007	3.64	97.09
6	0.0005	2.34	99.43
7	0.0001	0.57	100.00

Somme	0.0204
Effectif	202100
Produit	4124.07

Identificateur
Ile-de-France
Champagne-Ardennes
Picardie
Haute-Normandie
Centre
Basse-Normandie
Bourgogne
Nord-Pas de Calais
Lorraine
Alsace
Franche-Comté
Pays de la Loire
Bretagne
Poitou-Charentes
Aquitaine
Midi-Pyrénées
Limousin
Rhône-Alpes
Auvergne
Languedoc-Roussillon
Provence-Alpes-Côte d'Azur

1/21=4.76 %

Poids relatif	Distance à l'origine
21.41	0.02629
2.11	0.00934
2.44	0.00515
2.41	0.02792
3.24	0.00207
2.21	0.01236
2.98	0.04571
6.38	0.03197
4.42	0.04943
2.82	0.07000
1.87	0.02374
5.04	0.01781
5.09	0.00636
2.73	0.02395
4.90	0.02190
5.03	0.00546
1.33	0.01612
9.79	0.00413
2.42	0.00818
3.62	0.02046
7.76	0.00839

Coordonnées	
Axe 1	Axe 2
-0.15	0.07
0.06	-0.04
-0.03	-0.03
0.07	0.03
0.01	-0.02
0.03	-0.08
0.19	0.02
0.13	0.09
0.21	0.05
0.23	0.08
0.13	-0.04
0.07	-0.08
-0.02	0.02
0.06	-0.14
-0.04	-0.12
-0.03	-0.06
0.05	0.02
0.02	0.03
-0.03	-0.01
-0.11	-0.09
-0.05	-0.07

Contributions	
Axe 1	Axe 2
38.06	22.06
0.64	0.78
0.21	0.35
0.97	0.44
0.01	0.30
0.14	2.81
9.22	0.21
8.78	10.39
16.11	2.50
12.89	4.34
2.66	0.56
2.32	7.39
0.10	0.70
0.84	11.66
0.56	15.02
0.30	4.05
0.33	0.15
0.43	1.62
0.17	0.10
3.83	5.85
1.42	8.71

Cosinus carrés	
Axe 1	Axe 2
0.81	0.18
0.39	0.18
0.20	0.13
0.17	0.03
0.02	0.20
0.06	0.47
0.81	0.01
0.52	0.23
0.88	0.05
0.78	0.10
0.72	0.06
0.31	0.37
0.04	0.10
0.15	0.81
0.06	0.64
0.13	0.67
0.18	0.03
0.13	0.18
0.10	0.02
0.62	0.36
0.26	0.61

Libellé de la variable
Philosophie Lettre - A
Economique et Social - B
Mathématiques et Sciences Physiques - C
Mathématiques et Sciences de la Nature - D
Mathématiques et Techniques - E
Technique Industrielle - F
Technique Economique - G
Technique Informatique - H

Poids relatif	Distance à l'origine
22.48	0.00897
10.69	0.02768
16.22	0.02064
22.76	0.00652
2.64	0.07390
9.75	0.05641
15.23	0.01906
0.22	0.10986

Coordonnées	
Axe 1	Axe 2
-0.05	-0.07
-0.14	0.06
-0.09	0.10
-0.02	-0.05
0.19	0.05
0.22	0.09
0.12	-0.04
0.09	0.04

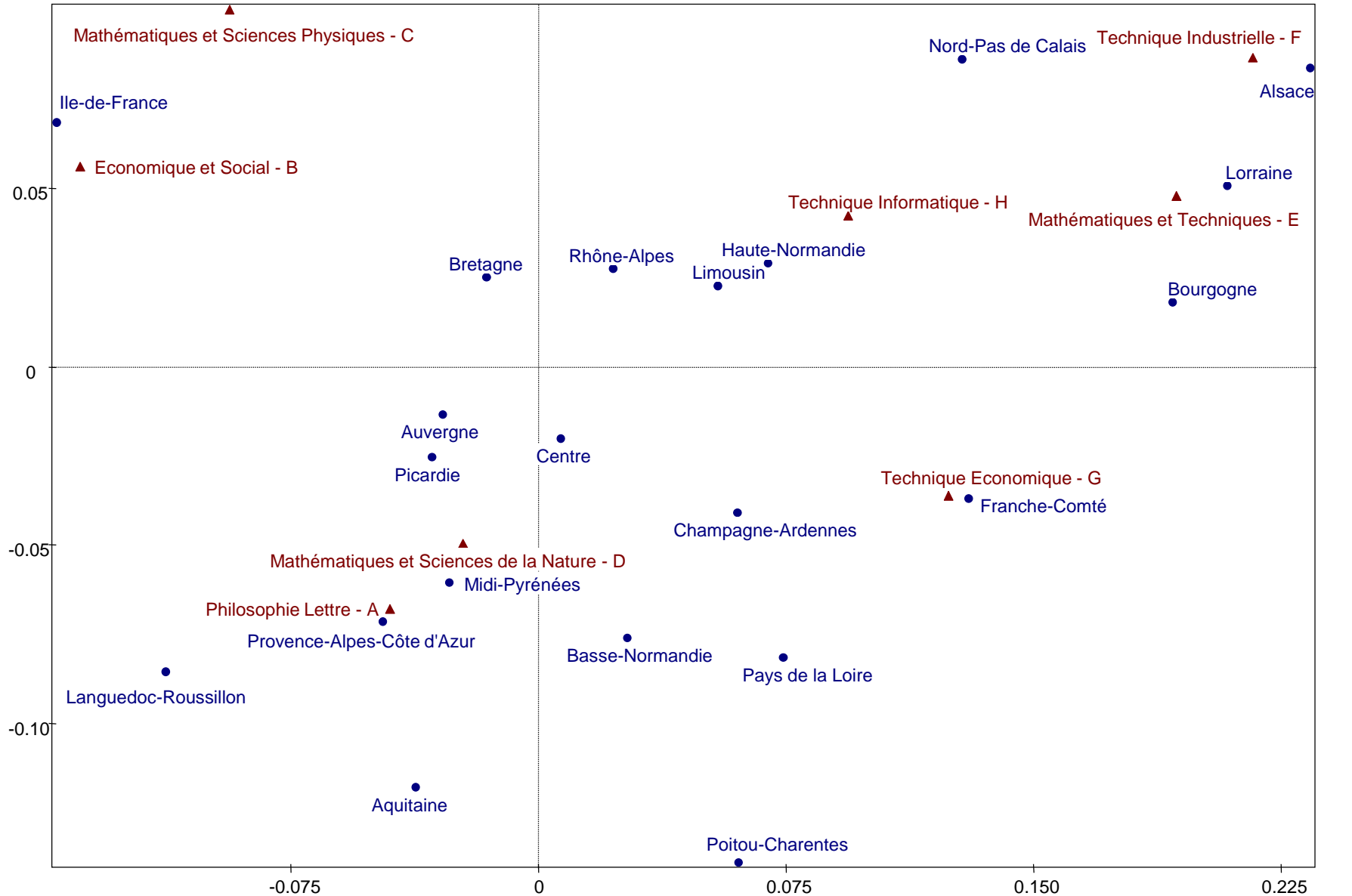
Contributions	
Axe 1	Axe 2
3.82	22.80
17.19	7.40
11.83	35.64
0.99	12.25
8.25	1.33
38.17	16.13
19.59	4.37
0.16	0.09

Cosinus carrés	
Axe 1	Axe 2
0.23	0.51
0.70	0.11
0.42	0.48
0.08	0.38
0.51	0.03
0.83	0.13
0.81	0.07
0.08	0.02

1/8=12.5 %

Représentation graphique sans la Corse A.F.C.

Facteur 2 - 22.25 %



Facteur 1 - 58.71 %