Traitement statistique d'un gros fichier de données. Enchaînement de quelques méthodes. Première partie

Joseph SAINT PIERRE,
Centre Interuniversitaire
de Calcul de Toulouse
118, Route de Narbonne
31062 Toulouse Cedex 04
Tél: 05 61 36 60 21
Fax: 05 61 52 14 58
e-mail: Joseph.Saint-Pierre@cict.fr

Vendredi 15 Janvier 1999

Cet exposé et le suivant seront consacrés à l'exploration statistique d'un même fichier de données. Le but de ces exposés, plus que de montrer des méthodes statistiques est de montrer une démarche possible en utilisant plusieurs types de méthodes.

Ce document correspond à la première partie de l'exposé et est donc incomplet, le plan et la rédaction de cette partie reste très informels. Éventuellement un document complet sera rédigé en intégrant les deux parties et en faisant une structuration. Il faut donc voir ce document comme un outil permettant de suivre l'exposé oral correspondant. De plus ces données font l'objet d'une étude plus poussée par une équipe de recherche à laquelle appartiennent les données.

En premier lieu il faut définir le principe de l'étude qui a amené la constitution du fichier de données. Il s'agit d'une étude menée par l'équipe Psychologie Sociale du Développement et de la Santé de l'Université de Toulouse le Mirail. L'étude a été faite sur la base d'un questionnaire particulièrement long et fermé, il y a 5 questions «biographiques» sexe, âge, année d'étude, situation familiale et ressources, et il y a 211 questions qui correspondent à des notes sur des échelles de 1 à 5, ces notes sont des appréciations sur l'importance que le répondant accorde à un mot ou un groupe de mots.

Une étude similaire avait déjà été traitée au CICT, sous une forme différente et surtout sur une population différente avec une langue différente. Il y avait 177 mots mais les échelles étaient avec 7 niveaux. Il m'avait été demandé en Novembre 1996 de faire une analyse factorielle des correspondances multiples. Comme il y avait des réponses manquantes cela faisait $8 * 177$ soit 1416 modalités. Ce qui constituait une sorte de record, il m'avait fallu «recompiler» le logiciel SPAD pour pouvoir faire passer ce fichier de données.

Pour le fichier de données que je vais utiliser, il n'y aurait pour l'AFCM que (si j'ose dire) 1266 modalités... Mais il ne pas été demandé de faire une AFCM mais une ACP ce qui ramène le problème à une diagonalisation d'une matrice de seulement 211 colonnes ce qui est quand même considérable.

Avant de se lancer dans le traitement des 211 questions, il paraît intéressant de regarder les variables dites biographiques pour connaître la population d'intérêt.

Pour la variable âge, la moyenne est de 21,5 ans, la médiane est de 20 ans l'écart type vaut 5 ans, le minimum est 18 ans le maximum est de 54 ans. L'histogramme schématique montre que cette population est extrêmement homogène en âge par rapport à la population française. Qui plus est elle représente une population relativement peu importante dans la pyramide des âges du pays.

```
270       20 |**********************************
104       24 |*************
 10       28 |*
  3       32 |
  5       36 |*
  6       40 |*
  8       44 |*
  0       48 |
  1       52 |
             +----+----+----+----+----+----+----+----+----+----+
             0         80        160       240       320       400
```

	effectif	%
Sexe		
Garçons	57	14.0%
Filles	350	86.0%

Pour la variable sexe la disproportion est phénoménale. Sur l'ensemble de la population française il doit y avoir 52% de femmes essentiellement en raison de la durée de vie, mais comme il s'agit d'une population jeune il devrait y avoir une proportion proche de 50% pour chacun des sexes.

	effectif	%
Scolarité		
1ère année	213	52.3%
2ème année	69	17.0%
3ème année	25	6.1%
4ème année	99	24.3%
Doctorat	1	.2%

La variable scolarité montre qu'il s'agit d'étudiants et donc d'une population particulière. Dans une population d"étudiants l'âge est corrélé positivement avec le niveau d'étude alors que c'est le contraire avec la population adulte française, cela étant dû à l'allongement de la durée des études depuis très longtemps. Les gens âgés ont en principe un niveau d'études inférieur aux jeunes. Cela est heureusement faux pour la population scolarisée et particulièrement les étudiants.

	effectif	%
Fam		
célibataire	117	28.7%
marié(e)	11	2.7%
vie en couple	50	12.3%
vit chez parents	43	10.6%
vit seul(e)	22	5.4%
cél+Vie en couple	4	1.0%
cél.+ vit parents	69	17.0%
cél+vit seul(e)	91	22.4%

	effectif	%
Ressources		
indépendant fin.	110	27.2%
ress. familiales	239	59.0%
les deux	56	13.8%

Les 2 variables suivantes sont de manière intuitive liées aux variables précédentes; la poursuite d'études et le jeune âge des répondants explique en partie la dépendance des parents pour les financements et aussi la faible importance du nombre de personnes vivant en couple ou mariées.

La disproportion entre les effectifs de garçons et de filles est explicable par l'origine disciplinaire des étudiants, les répondants sont soit des étudiants de psychologie, soit des étudiants d'IUT, soit des étudiants de première année du Mirail suivant des enseignements de psychologie. La psychologie est une discipline où les filles sont particulièrement nombreuses. Sachant où les étudiants ont été interrogés une autre variable à deux modalités a été créée il s'agit de la filière (Mirail ou IUT)

	effectif	%
Filiere d'etudes		
Mirail	294	72.2%
IUT	113	27.8%

Il y a beaucoup plus d'étudiants venant du Mirail que de l'IUT. Il est à noter que les étudiants de l'IUT sont tous en première année. Par principe il ne peut y avoir d'étudiants d'IUT au dessus de la deuxième année. À partir d'une constatation de ce style on peut dire qu'il y a danger de tomber dans des situations du type «paradoxe de Simpson». C'est à dire que l'on peut attribuer à l'année d'étude ce qui est en fait plutôt explicable par la filière.

À titre d'exemple un tableau croisé entre le sexe et la filière:

```
SEXE  Sexe  by  FILIERE  Filière d'études

                    FILIERE
            Count  |
           Exp Val |
           Row Pct |Mirail   IUT
           Col Pct |
           Std Res |       1|       2|
SEXE       --------+--------+--------+
                 1 |    24  |    33  |    57
  Garçons          |  41.2  |  15.8  | 14.0%
                   | 42.1%  | 57.9%  |
                   |  8.2%  | 29.2%  |
                   |  -2.7  |   4.3  |
                   +--------+--------+
                 2 |   270  |    80  |   350
  Filles           | 252.8  |  97.2  | 86.0%
                   | 77.1%  | 22.9%  |
                   | 91.8%  | 70.8%  |
                   |   1.1  |  -1.7  |
                   +--------+--------+
                      294      113      407
```

Chi-Square	Value	DF	Significance
Pearson	30.00406	1	.00000

La liaison entre les deux variables est très forte, elle n'a rien de structurel mais avec ces deux variables on peut aussi avoir un effet de liaison croisée. Par exemple on peut attribuer à la variable sexe ce qui est dû à la filière ou réciproquement.

Bien évidemment il se peut que d'autres variables aient un rôle explicatif sur les réponses aux 211 questions et qui ne soient pas dans le questionnaire (niveau socio-économique d'origine, nationalité, appartenance religieuse etc.) Le choix des variables explicatives est souvent dépendant des questions de recherche.

L'étape la plus importante du traitement fut de faire le tableau de fréquences (appelé tri à plat) et de le présenter simplement. Comme toutes les variables ont les mêmes modalités ont les mêmes modalités l'idée a été de mettre les modalités en colonne et les variables en ligne avec au croisement ligne colonne l'effectif.

L'obtention d'un tel tableau avec SPSS fut très délicat: Voici la syntaxe utilisée, avec le détail suivant il fallait écrire les noms des 211 variables et pas des points. De plus on peut gérer que 100 variables à la fois. Le plus dur à trouver était la partie «by (LABELS)» j'ai dû faire appel aux «news» comp.soft-sys.stat.spss pour qu'un guru SPSS me donne cette astuce.

```
tables
  / format zero
  / table = v1+v2+V3+...+v211  BY (LABELS)
  / stat = COUNT ('')
```

J'ai décidé aussitôt de mettre la moyenne et la variance de chacune des variables sur la même ligne, ce qui se fait avec une syntaxe similaire. Mais attention cela met en évidence que les échelles de 1 à 5 peuvent être traitées en quantitatif ou en qualitatif, il n'y a beaucoup de méthodes spécifiques au traitement des données qualitatives ordonnées ce qui est le cas de ce type d'échelle.

L'intérêt premier de ce tableau est de donner la liste des mots qui ont été présentés aux répondants.

Pas	Peu	Moy	Imp	Très	NR	Moy	Var	Num	Intitulé
3	8	57	141	198	0	4.285	0.692	V001	Loyauté
10	40	178	133	42	4	3.390	0.795	V002	Charité
0	5	26	124	251	1	4.530	0.452	V003	Fidélité
1	11	82	180	131	2	4.059	0.654	V004	persévérance
1	1	26	87	289	3	4.639	0.405	V005	Liberté
19	42	142	132	68	4	3.467	1.080	V006	charme
28	55	150	101	34	39	3.158	1.076	V007	Décentration
3	12	76	200	112	4	4.007	0.654	V008	Détermination
1	3	21	106	275	1	4.603	0.417	V009	Franchise
100	86	141	42	14	24	2.436	1.210	V010	piété
7	15	86	162	137	0	4.000	0.852	V011	Loisir
68	118	131	64	22	4	2.638	1.217	V012	révolution
1	4	31	176	195	0	4.376	0.476	V013	volonté
0	0	27	120	259	1	4.571	0.378	V014	sentiment
13	33	142	130	78	11	3.573	1.005	V015	développement
28	56	149	76	25	73	3.042	1.031	V016	mansuétude
164	101	113	20	3	6	1.995	0.960	V017	célibat
142	114	100	43	7	1	2.160	1.152	V018	patriotisme
1	4	44	160	196	2	4.348	0.534	V019	affection
6	22	100	173	105	1	3.860	0.837	V020	progrès
8	23	86	151	138	1	3.956	0.956	V021	fraternité
239	68	66	23	4	7	1.713	1.002	V022	abstinence
250	71	65	12	5	4	1.638	0.883	V023	tout-puissance
35	59	141	119	51	2	3.227	1.240	V024	vocation
18	50	108	158	73	0	3.536	1.121	V025	sensualité
4	15	90	172	124	2	3.980	0.766	V026	professionnali
148	89	92	39	29	10	2.275	1.584	V027	désintéresseme
4	7	56	155	185	0	4.253	0.687	V028	solidarité
16	41	106	135	98	11	3.652	1.174	V029	acceptation
115	89	123	49	23	8	2.439	1.418	V030	communion
218	79	76	20	14	0	1.853	1.210	V031	virginité
0	12	65	182	145	3	4.139	0.620	V032	sensibilité
3	13	72	190	127	2	4.049	0.685	V033	Accueil
13	39	167	140	38	10	3.380	0.822	V034	clémence
1	3	25	122	256	0	4.545	0.445	V035	dialogue

Pas	Peu	Moy	Imp	Très	NR	Moy	Var	Num	Intitulé
3	5	35	139	225	0	4.420	0.579	V036	plaisir
33	59	152	121	39	3	3.183	1.128	V037	vertu
36	53	148	86	59	25	3.207	1.319	V038	pluralisme
1	10	67	194	135	0	4.111	0.606	V039	dynamisme
6	41	118	162	76	4	3.648	0.900	V040	innovation
12	40	145	147	56	7	3.488	0.912	V041	conciliation
18	59	105	128	97	0	3.558	1.282	V042	écologie
3	25	101	168	109	1	3.874	0.816	V043	hospitalité
20	59	142	133	50	3	3.332	1.061	V044	méthode
10	39	126	141	87	4	3.635	1.008	V045	prudence
6	31	98	171	98	3	3.802	0.888	V046	sérénité
51	57	126	104	57	12	3.149	1.483	V047	unité
67	97	157	52	9	25	2.579	1.011	V048	repentance
26	58	164	106	32	21	3.155	1.004	V049	tempérance
0	2	12	73	319	1	4.746	0.278	V050	amour
92	88	112	76	38	1	2.704	1.601	V051	compétitivité
31	67	126	134	47	2	3.244	1.210	V052	don
35	48	118	118	87	1	3.429	1.431	V053	idéal
15	44	118	154	76	0	3.570	1.059	V054	discrétion
134	103	111	48	11	0	2.260	1.252	V055	pouvoir
1	2	10	78	316	0	4.735	0.303	V056	santé
20	33	159	133	56	6	3.429	0.990	V057	pratique
7	18	111	181	89	1	3.805	0.789	V058	action
32	32	103	145	95	0	3.587	1.342	V059	combativité
38	80	132	108	47	2	3.114	1.294	V060	méticulosité
21	51	112	136	86	1	3.530	1.237	V061	honneur
58	91	167	67	12	12	2.706	1.025	V062	rénovation
6	13	101	178	108	1	3.909	0.769	V063	spontanéité
5	20	99	200	81	2	3.820	0.722	V064	coopération
1	9	39	117	241	0	4.445	0.602	V065	égalité
3	22	116	162	104	0	3.840	0.799	V066	intelligence
42	63	153	94	50	5	3.117	1.296	V067	sobriété
3	26	99	166	110	3	3.876	0.828	V068	réceptivité
2	8	55	126	216	0	4.342	0.673	V069	complicité
5	19	88	178	117	0	3.941	0.799	V070	disponibilité

Pas	Peu	Moy	Imp	Très	NR	Moy	Var	Num	Intitulé
1	11	39	134	222	0	4.388	0.622	V071	intimité
7	32	100	175	90	3	3.765	0.890	V072	organisation
188	91	73	35	20	0	2.037	1.430	V073	religion
15	45	137	139	66	5	3.488	1.029	V074	surêté
38	37	88	94	149	1	3.687	1.697	V075	fécondité
4	8	44	136	215	0	4.351	0.681	V076	justice
5	17	69	147	167	2	4.121	0.849	V077	espérance
97	91	105	83	31	0	2.656	1.571	V078	coquetterie
53	88	172	79	14	1	2.786	1.028	V079	autorité
9	19	79	175	125	0	3.953	0.887	V080	autocontrôle
2	3	78	203	121	0	4.076	0.558	V081	connaissance
1	13	56	174	163	0	4.192	0.652	V082	douceur
2	5	49	146	204	1	4.342	0.601	V083	émotion
9	36	114	161	84	3	3.681	0.947	V084	observation
77	79	137	52	16	46	2.587	1.226	V085	rédemption
0	6	46	179	172	4	4.283	0.521	V086	sympathie
45	80	147	101	33	1	2.993	1.217	V087	beauté
1	2	13	107	284	0	4.649	0.351	V088	honnêté
3	8	63	172	160	1	4.177	0.664	V089	maîtrise de so
14	46	136	144	65	2	3.494	1.008	V090	rigueur
37	71	149	90	52	8	3.123	1.284	V091	transparence
2	17	77	192	118	1	4.002	0.693	V092	partage
0	5	28	116	258	0	4.541	0.460	V093	joie
2	21	131	168	83	2	3.763	0.725	V094	cohérence
7	24	88	193	94	1	3.845	0.817	V095	courage
149	91	97	43	9	18	2.157	1.261	V096	mysticisme
19	26	85	157	120	0	3.818	1.149	V097	féminité
71	79	106	86	59	6	2.958	1.711	V098	spiritualité
7	24	126	182	65	3	3.678	0.764	V099	raison
25	22	109	169	80	2	3.635	1.109	V100	curiosité
25	51	147	130	49	5	3.316	1.089	V101	alliance
2	11	60	188	146	0	4.143	0.639	V102	générosité
3	8	35	161	198	2	4.341	0.606	V103	affectivité
32	62	149	103	61	0	3.243	1.263	V104	chance
113	126	124	31	10	3	2.255	1.054	V105	infaillibilité

Pas	Peu	Moy	Imp	Très	NR	Moy	Var	Num	Intitulé
4	24	123	166	87	3	3.762	0.792	V106	objectivité
12	11	77	175	132	0	3.993	0.889	V107	simplicité
0	2	17	118	270	0	4.612	0.351	V108	vérité
165	114	97	19	3	9	1.947	0.916	V109	austérité
42	74	167	84	40	0	3.015	1.197	V110	pudeur
10	12	68	139	178	0	4.138	0.926	V111	démocratie
8	21	95	169	112	2	3.879	0.883	V112	dignité
51	89	169	75	23	0	2.828	1.103	V113	obéissance
24	57	123	116	86	1	3.451	1.310	V114	secret
21	31	106	167	82	0	3.634	1.100	V115	serviabilité
3	1	20	35	348	0	4.779	0.374	V116	vie
33	79	154	99	30	12	3.035	1.090	V117	expansivité
3	8	30	94	271	1	4.532	0.605	V118	famille
146	80	96	12	3	70	1.950	0.940	V119	ascèse
36	60	136	106	48	21	3.181	1.271	V120	caressant
38	75	165	101	26	2	3.005	1.069	V121	économie
2	8	42	157	195	3	4.324	0.606	V122	indépendance
8	15	69	197	118	0	3.988	0.785	V123	patience
6	14	57	157	173	0	4.172	0.807	V124	profession
6	8	40	129	224	0	4.369	0.725	V125	rêve
0	4	31	154	218	0	4.440	0.458	V126	tendresse
7	33	142	178	46	1	3.549	0.742	V127	vigilance
58	66	152	50	23	58	2.754	1.203	V128	affabilité
3	14	65	190	135	0	4.081	0.690	V129	culture
5	27	110	162	103	0	3.813	0.866	V130	calme
5	15	80	158	149	0	4.059	0.819	V131	harmonie
51	61	150	76	26	43	2.904	1.217	V132	opiniâtreté
9	28	138	159	72	1	3.633	0.860	V133	soin
56	86	117	105	42	1	2.978	1.439	V134	tradition
16	53	135	140	62	1	3.441	1.052	V135	précaution
2	10	45	169	181	0	4.270	0.626	V136	Autonomie
5	15	122	171	94	0	3.821	0.758	V137	Bonté
1	2	41	179	184	0	4.334	0.484	V138	Éducation
3	4	23	100	276	1	4.581	0.505	V139	Paix
8	31	128	171	69	0	3.644	0.840	V140	Ténacité

Pas	Peu	Moy	Imp	Très	NR	Moy	Var	Num	Intitulé
38	80	145	104	35	5	3.045	1.185	V141	Propriété
70	59	74	95	109	0	3.280	2.064	V142	Mariage
3	13	93	187	108	3	3.950	0.692	V143	Amabilité
83	100	146	40	11	27	2.463	1.073	V144	Conversion
10	20	106	200	70	1	3.739	0.786	V145	Indulgence
70	86	116	78	44	13	2.848	1.554	V146	Paternalisme
12	33	104	169	87	2	3.706	0.980	V147	Modestie
106	129	111	34	14	13	2.292	1.133	V148	Béatitude
0	9	47	203	147	1	4.202	0.527	V149	Compréhension
11	24	87	157	126	2	3.896	0.999	V150	Authenticité
8	10	58	171	160	0	4.143	0.792	V151	Imagination
1	5	30	147	224	0	4.445	0.498	V152	Tolérance
60	72	162	51	15	47	2.692	1.083	V153	Rectitude
6	26	109	176	87	3	3.772	0.826	V154	Réconciliation
18	42	152	140	52	3	3.411	0.977	V155	Compassion
35	61	148	113	47	3	3.188	1.210	V156	Fougue
137	101	73	54	42	0	2.418	1.805	V157	Foi
5	16	113	183	89	1	3.825	0.737	V158	Lucidité
220	84	63	32	7	1	1.823	1.139	V159	Noblesse
4	8	60	186	149	0	4.150	0.659	V160	Travail
60	67	109	116	54	1	3.091	1.569	V161	Pureté
6	20	89	175	116	1	3.924	0.831	V162	Sécurité
6	18	85	176	122	0	3.958	0.818	V163	Créativité
34	63	112	136	60	2	3.309	1.328	V164	Intérêt
19	51	129	131	74	3	3.470	1.153	V165	Moralité
10	28	74	167	128	0	3.921	0.988	V166	Ambition
27	82	160	102	34	2	3.084	1.052	V167	Discipline
16	55	132	133	71	0	3.462	1.106	V168	Intuition
11	23	110	168	95	0	3.769	0.921	V169	Originalité
1	3	21	130	252	0	4.545	0.425	V170	Respect
83	105	122	79	17	1	2.611	1.290	V171	Ruse
3	16	87	183	115	3	3.968	0.726	V172	Conscience
70	75	101	110	43	8	2.952	1.598	V173	Masculinité
6	24	68	143	166	0	4.079	0.939	V174	Passion
11	46	141	148	57	4	3.481	0.926	V175	Clairvoyance

Pas	Peu	Moy	Imp	Très	NR	Moy	Var	Num	Intitulé
20	44	115	122	86	20	3.543	1.233	V176	Altruisme
3	49	102	149	103	1	3.739	0.988	V177	Sagesse
34	71	168	92	23	19	2.997	1.010	V178	Réhabilitation
12	29	125	140	101	0	3.710	1.024	V179	Légalité
25	41	133	141	63	4	3.437	1.137	V180	Humilité
147	122	96	28	12	2	2.101	1.136	V181	Censure
21	53	133	152	43	5	3.356	1.023	V182	Concentration
0	1	20	149	237	0	4.528	0.363	V183	Sincérité
8	19	104	174	101	1	3.840	0.845	V184	Protection
17	52	145	133	55	5	3.391	1.027	V185	Subtilité
34	72	146	97	53	5	3.157	1.265	V186	Droiture
124	104	89	56	34	0	2.440	1.636	V187	Croyance
24	56	145	131	43	8	3.283	1.063	V188	Libéralité
5	31	118	191	61	1	3.670	0.750	V189	Décision
10	28	125	180	63	1	3.635	0.829	V190	Assurance
28	85	158	104	31	1	3.062	1.046	V191	Exigence
15	19	88	174	111	0	3.853	0.988	V192	Pardon
22	55	144	137	48	1	3.330	1.056	V193	Compromis
7	14	123	198	64	1	3.734	0.684	V194	Compétence
13	38	96	139	121	0	3.779	1.148	V195	Promesse
27	71	117	131	58	3	3.302	1.253	V196	Force
25	43	132	133	67	7	3.435	1.174	V197	Don de soi
9	27	106	200	65	0	3.700	0.796	V198	Diplomatie
14	27	103	164	95	4	3.742	1.008	V199	Intégrité
0	7	60	196	142	2	4.168	0.541	V200	Responsabilité
40	48	165	76	23	55	2.983	1.071	V201	Magnanimité
14	51	147	148	46	1	3.397	0.926	V202	Devoir
261	84	45	6	6	5	1.537	0.743	V203	Chasteté
44	91	167	83	17	5	2.846	1.019	V204	Modération
242	85	55	12	9	4	1.663	0.945	V205	Sainteté
22	58	140	144	42	1	3.310	1.034	V206	Dévouement
8	19	120	174	85	1	3.761	0.814	V207	Sérieux
46	44	126	117	70	4	3.300	1.464	V208	Privé
6	34	108	160	98	1	3.764	0.921	V209	Séduction
9	34	141	166	50	7	3.535	0.805	V210	Gratitude
92	80	102	83	49	1	2.796	1.753	V211	Méditation

En «butinant» ou «surfant» sur ces tableaux certains points sautent vite aux yeux, à titre d'exemple la colonne des non-réponses a beaucoup de 0 mais aussi des nombres relativement importants 70 pour ascèse mot numéro 119. On peut déduire que certains mots ne sont peut être pas bien connus par tous les répondants en ce qui me concerne c'est le mot décentration, numéro 7, que je ne connaissais pas du tout; il y a 39 non-réponses cela m'a paru peu étant donné que je ne l'ai pas trouvé sur les dictionnaires.

Une fois ce tableau obtenu, je suis passé à divers tri de manière à classer les mots suivants les notes qu'ils ont obtenu. Cette présentation tend d'ailleurs à considérer les mots comme des sujets sachant que les répondants sont les membres d'un jury.

Le premier tri et le plus simple consiste à classer les mots par la moyenne. Cela est très facile à faire avec des tas de logiciels et même au niveau du système ou avec le traitement de texte utilisé pour écrire ce texte (emacs). Les mots qui ont la plus forte moyenne sont vie (numéro 116), amour (50) et santé (56), ce sont les seuls qui dépassent $4,7$ ce sont les seuls à obtenir plus de 300 modalités 5, les mots suivants sont honnêteté (88), liberté (5) etc. Les mots à plus faible moyenne sont Chasteté (203), toute-puissance (203), Sainteté (205), abstinence (22), Noblesse (159) et virginité (31) etc. ces 6 mots sont les seuls à avoir obtenu une majorité absolue de réponses 1 (pas importants) soit plus de 205. On peut voir que seule la modalité 5 a obtenu pour certains mots des scores supérieurs à 75% et que les modalités 1 et 5 sont les deux seules à avoir obtenu des majorités absolues pour certains mots. La modalité 2 n'est la valeur modale que pour deux mots, Béatitude (148) et infaillibilité (105) et c'est de très peu.

Il est intéressant de regarder le classement par la variance, il est même amusant de regarder le mot le plus variable et le moins variable. Le mot à plus forte variance est de loin Mariage (142), la variance de ce mot est $2,064$ et c'est le seul mot dont la variance dépasse la valeur 2. Par contre le mot à plus faible variance est amour (50) c'est le seul mot dont la variance est inférieure à $0,3$. L'interprétation d'une forte variance est que le mot ne réalise pas d'accord sur son importance cela peut être obtenu, par exemple, avec une forte occurrence de la modalité 1 et de la modalité 5. Les mots à plus forte variance après Mariage sont Foi (157), Méditation (211), spiritualité (98), fécondité (75), Croyance (187) etc. Les moins variables sont après amour, santé (56), vérité (108), honnêteté (88), Sincérité (183), vie(116) etc. Les mots à trop forte moyenne ou trop faible moyenne ne peuvent pas avoir de forte variance, un mot qui aurait la plus forte moyenne possible soit 5, ou la plus faible possible soit 0, aurait une variance nulle. On pourrait obtenir des classements pour toutes les valeurs possibles de l'échelle, y compris la réponse manquante, mais une idée pour résumer les deux mesures moyenne et variance est de faire une représentation en deux dimensions ce qui permet de visualiser rapidement les mots par leur niveau d'importance et leur niveau de variabilité. Un graphique haute définition se trouve en fin de document, mais voici une représentation alphanumérique obtenu avec SPAD:

```
                           moyenne  *  variance
variance
      ---------- 187 157 211  98 161 142----  75---------------------
     !        27  51  78      173         !                           !
     !                   146              !                           !
1.5  +                                    !                           !
     !                            47  53  !                           !
     !                       134     208  !                           !
     !    73      30                      !                           !
     !                                    !                           !
     !                        38  67 164  59                          !
     !                       104  60 114  !                           !
     !    96         171      91 120 186  42  61                      !
     !        55      85 128  87  24 196 176                          !
1.2   31          10  12 132 110  52 156  !                           !
      18                     141 165 197  29  97                      !
     159 181 148                  37 180 101 195                      !
     !                   113 117 121 168  25 100                      !
     !           144 153  16 201   7  44  54 115                      !
     !       105      62  79 191 188 182  74 179 199 192              !
      22              48 204 178  49  57  15 177 166 150              !
      17 119------------------------ 155  45 147  21 111-------------!
     205 109                         202 175 209 169 174              !
 .9   23                              41  40  72 112  80              !
     !                               140 133 130  11  77              !
     !                                34 190 184  95 131 151          !
     203                               2 210 198 123 124              !
     !                                   207 137  63  70              !
     !                                   127  64 172  26 125  28      !
     !                                   158 194  92  33  76   1      !
     !                                    !   143   8  89  69  71     !
     !                                    !        82 102 136 122  36 !
 .6  +                                    !        32  39 103  65 118 !
     !                                    !        81 200  83  86     !
     !                                    !           149  19 138     !
     !                                    !                13 152 139 !
     !                                    !                35  93 126   9
     !                                    !                 3 170   5 !
     !                                    !                    14 108 116
     !                                    !                   183  88 !
     !                                    !                        56 !
 .3  +-----------+---------------+---------------+--------------  50-
                2.4             3.2             4.0             4.8
                                                                 moyenne
```

Attention certains points ne sont pas représentés ou sont sont mal représentés voici les informations concernant ces points:

Points éloignés :

IDENTIFICATEUR	ABSCISSE	ORDONNEE
22	1.713	1.002
23	1.638	.883
31	1.853	1.210
75	3.687	1.697
98	2.958	1.711
142	3.280	2.064
157	2.418	1.805
159	1.823	1.139
203	1.537	.743
205	1.663	.945
211	2.796	1.753

11 points ont ete ramenés sur le bord du graphique.

point vu	abscisse approchée	ordonnee approchée	nb. de cachés	points cachés
44	3.38	1.07	2	6 193
191	3.02	1.04	1	167
182	3.38	1.04	3	206 135 185
15	3.57	1.00	1	90
209	3.75	.93	1	84
72	3.75	.90	1	46
112	3.93	.90	1	107
184	3.75	.83	1	154
95	3.93	.83	5	163 20 43 68 162
198	3.75	.80	4	58 66 145 106
137	3.75	.76	2	99 189
64	3.75	.73	1	94
33	4.11	.69	1	129
89	4.11	.66	2	4 160

Une autre façon de synthétiser un tel tableau est de faire une analyse factorielle des correspondances simple (AFC). C'est ce que j'ai fait ensuite. Cela va permettre de voir les associations (correspondances) entre les mots et les notes. Bien évidemment il faut se souvenir que la population étudiée est particulièrement homogène et que les variables «biographiques» sont absentes de ces traitements.

```
                            AXE 1  *  AXE 2
AXE 2
     56  116-------------------------------------------- 31  205 159 73
     50  88                  !                           22  203PAS1 !
     5   139                 !                                       !
 .6  9   118                 !                                       !
     108 93                  !                                       !
     14  35                  !                                       !
     170 3                   !                           157         !
     !                       !                                   109 !
     183 65  125             !                               27  17  !
     36  152 71              !                               181     !
    TRE5 76  126             !                       187 18          !
     103 69                  !   142                         96      !
 .3  122 83  13              !                           55          !
     !   19  1   28      75  !                                       !
     !   138 111 136         !                                       !
     !       124             !               211     30              !
     !   77  151 174         !                   78      105         !
     !   86  82              !                   51                  119
     !   160 89  149         !           98              148         !
     !   200 102 131 11      !                                       !
     !   129 32  107 97  150 !                                       !
 .0  +------ 39  80  195 59 -+------ 161 173 146 171 10 -------------!
     !       4   33  192     53                                      !
     !       81  70  123 42  !   208     134    PEU2                 !
     !       26  8   112 199 114     47          12                  !
     !       172 63  20  115 61  164                 144             !
     !       143 68IMP4  147 25  176 196 67                          !
     !           209 137 84  165 104 91  87          85              !
     !           72  158 58  180 52  156 110 113                     !
     !           64  145 40  6   168 120 38  79  62  48              !
-.3  +           198 94  133 90  101 37  121 204                     !
     !               194 189 175MOY3 167 117 132 128                 !
     !               99  185 41  155 206 191     153                 !
     !                   127 210 202     178                         !
     !                       !       49  7                           !
     !                       2   34                                  !
     !                       !               201                     !
     !                       !                                       !
     !                       !                      NR6              !
-.6  +-------+---------------+-------------- 16 -------------+-------
            -.6              .0              .6             1.2
                                                                  AXE 1
```

Le graphique de l'AFC et le graphique de représentation simple donnent des interprétations similaires attention tous les points ne sont pas représentés dans le graphique de l'AFC. Le graphe de l'AFC a une forme en U, au centre du graphique en dehors du U se trouve les quelques mots notamment Mariage (142) qui ont de fortes variances, ce mots se trouvent entre les modalités 1 et 5 et assez loin des autres. Le point très important de l'AFC qui n'est pas dans la représentation simple est la place de la modalité non-réponse. L'interprétation de cette position m'a amené à l'hypothèse suivante: la modalité 1 qui était intitulée «pas important» a été peut être utilisée comme une modalité de rejet, on aurait dû s'attendre à ce que la non-réponse soit quelque chose proche des modalités de faible valeur, or non-réponse est plus proche de 2 que de 1. Il me semble qu'une telle hypothèse devrait être explorée. L'énorme paquets de mots dans le coin inférieur gauche du graphique correspond à la très grande prédominance des modalités 3 et surtout 4 pour de très nombreux mots.

Tous ces traitements, tableau de fréquences, tris, puis graphique de dispersion et enfin AFC simple n'ont servi qu'à regarder les 211 variables mots de façon interne et sans aucune volonté de les synthétiser. Compte tenu des données en jeu deux choix classiques sont devant nous, l'un deux consiste à essayer de déterminer si il est possible de résumer les 211 variables en un ensemble plus réduit, cela est souvent fait à partir de méthodes de type analyse factorielles, ACP ou variantes, mais bien sûr en revenant au tableau initial avec les sujets. Une des constatations de l'AFC simple peut amener à considérer que les modalités 1 à 5 ne sont pas des scores mais bien des modalités de variables qualitatives, auquel cas c'est plutôt une analyse factorielle des correspondances multiples qu'une analyse en composantes principales qui peut être intéressante. Peut être que les analyses en composantes principales non linéaire du type princals seraient opportunes. L'autre type de questions que l'on se pose souvent dans ce type de situation est de savoir quels sont parmi les 211 variables celle qui permettraient le mieux de distinguer des sous populations définies par des variables dites biographiques. Cela peut être fait de diverses manières, analyse de variances systématiques, ou croisements systématiques ou même analyses discriminantes. In fine il s'agit souvent d'arriver à comparer divers sous ensembles de données même lorsque l'on procède par ACP ou AFCM, mais la comparaison se fait une fois le regroupement des variables effectué.

Dans cette étude, il me semble qu'une élimination de variables aurait pu être faite à partir de la lecture du tableau de fréquences et des analyses qui ont suivi, mais il se trouve que la première demande fut d'effectuer une ACP et c'est donc une ACP qui fut l'analyse faite. La première ACP donnait un très fort premier axe avec un coefficient *alpha* de Cronbach extrêmement élevé, supérieur à $0,98$ pour éviter cela j'ai recentré les notes par rapport aux répondants, c'est à dire que j'ai calculé la moyenne des notes données par chaque étudiant et j'ai retranché cette moyenne à chacune des notes données par ce même étudiant. Ainsi la moyenne des notes recentrées pour chaque étudiant vaut 0. C'est l'ACP faites sur ces notes recentrées que je présente ici. Compte tenu de la taille des fichiers il est hors de question que je présente ici tous les résultats de l'ACP, on ne doit pas agir avec un fichier de 211 variables comme avec un fichier de 211, cela utiliserait beaucoup trop de papier.

Voici donc les représentations graphiques de l'ACP, je n'ai gardé que les deux premiers axes pour cette présentation alors que je pense que les 5 premiers axes pouvaient avoir un intérêt d'interprétation, mais ce n'est pas le but de cet exposé. Le premier graphe correspond aux variables actives de l'ACP, les 211 mots. Le second aux variables illustratives, biographiques.

```
                             AXE 1  *  AXE 2
       ------------------------------------------ 73  142-------------
      !                               !                               !
      !                       206     !                               !
      !                       192     !                       205     !
      !                               !       134 157     113         !
      !                   115         118         146 203 187         !
      !                               !                               !
      !           107          3      !           31  162             !
 .25  +           108 2       170     85      10  161 165             !
      !           103 88          148 !   110     22                  !
      92      50          155 101 114 48  144 184 109     18          !
      102 14  183 82      197 195 75  !   30      37              167 !
      32      137 143     174 154     !   179         67              !
      86  65  126 139 9   210         !           78      207     79  !
      28  83                  34  56  54      119 135 45  202         !
      !       145 19  70  71  147 116 !   96  128         160         !
      !       69  29  123             !       124                     !
      !       93  33          120 180 49  181 173 74              141 !
      !           52          188 76  97      204             61      !
 .00  +-  43------------  21  16  138 1   17------------  186---  23--!  AXE1
      !   152     35      77  47  133 130     104     121 60      105 !
      !                       27  177 209     87  208 153             !
      !               193         178 24      127 112     166     90  !
      !           111 176 156 125 117 46  201 132 190                 55
      !               36  131 198 199 !   91  26          159         !
      !           64          41  98  13      4                       !
      !           5           150 7  106 6   164         59          51
      !       63      11      25  200 95  99  20  89  191             !
      !           149     38      122 211 140 171 62      72          !
      !       42      12          39  136     80                      !
-.25  +                   68      15  !   172     196                 !
      !                   129 94      189     8                       !
      !                       53     Âge  158 182                     !
      !                       100 84  175     194 66                  !
      !                           168 !                   44          !
      !           169                 57                              !
      !           151         58      !                               !
      !                               !   185                         !
      !                           40  !                               !
      +---------------+------ 163-----+-- 81----------+---------------+
     -.50            -.25             .00             .25             .50
```

```
                             AXE 1  *  AXE 2
       ------------------------------fam7-----------------------------
      !                           !                                   !
      !                           !                                   !
      !             parents       !                                   !
      !                           !                                   !
  .6  +                           !                                   !
      !                           !                                   !
      !                           !                                   !
      !                           !  1ère                             !
      !              2ème        RFam    seul                         !
      !                           !                                   !
      !                           !                                   !
      !              MirF   filles!              IUTf                 !
      !                           !                                   !
      !                  Mirail   !                                   !
  .0  +----------fam6-------------+--fam8-----------------------------!  AXE1
     3ème                         !                                   !
      !                           !                                   !
      !                           !                      IUT          !
      !                           !                                   !
      !                     Célib RF&A                                !
      !                           !                                   !
      !                           !                                   !
      !                           !                                   !
      !                  Indp     !                                   !
 -.6  +                           !                                   !
      !                           !                                   !
      !                           !                                   !
      !                           !                                   !
      !                           !  mari                             !
      !                  4ème     !                                   !
      !                           !                                   !
      !                           !          MirG            garçons  IUTg
      !                           !                                   !
      !                  couple   !                                   !
-1.2  +                           !                                   !
      !                           !                                   !
      !                           !                                   !
      !                           !                                   !
       -----------+--Doct---------+---------------+------  7_-----+---
                  -1               0               1               2
```

L'interprétation des deux premiers axes de cette ACP n'est pas exactement de mon ressort mais une idée générale n'est pas trop difficile à imaginer le premier axe est un axe «politique» et le deuxième est un axe «religieux». Dans les représentations des variables qualitatives illustratives, j'ai, après quelques essais, introduit une variable à 4 valeurs en croisant sexe et filière. IUTg représente les garçons de l'IUT, IUTf les filles de l'IUT, MirG les garçons du Mirail et MirF les filles du Mirail. Ce sont ces deux variables qui m'ont paru les simples à gérer pour montrer ce que pouvait être un équivalent du paradoxe de Simpson, Même si l'on peut voir à partir de ce graphique certaines des corrélations structurelles déjà mentionnées dès l'analyse du tableau de fréquence de ces variables, concernant d'autres variables que le sexe et la filière, ainsi on peut constater que le deuxième axe est assez bien lié avec l'âge et aussi avec d'autres variables liées à l'âge, ce sont les gens mariés ou vivant en couple qui sont sur le bas du graphique mais ce sont aussi les quatrièmes années, ou ceux qui sont indépendants financièrement.

Attention, une remarque s'impose concernant l'interprétation politique du premier axe, les opinions de droite et surtout d'extrême droite sont très souvent censurées par les répondants aux sondages. Les instituts de sondages «redressent» très fortement les intentions de vote pour le Front National, qui d'après les chiffres bruts n'obtiendrait que 5% des voix quand il en obtient 15% aux élections. Il n'est pas impossible que des mots tels que autorité (79) ou discipline (167) subissent un effet comparable.

L'ACP peut, de cette façon, masquer des liaisons qui peuvent être intéressantes, à titre d'exemple, je me suis posé la question de savoir si le mot le plus variable : Mariage (142) était lié à la situation matrimoniale des répondants. Mais la variable FAM a beaucoup trop de modalités alors j'ai regroupé toutes les modalités correspondant à une situation sans mariage ni vie de couple. Voilà ce que donne un tableau de χ^2 avec l'effectif observé, l'effectif théorique, le pourcentage ligne, le pourcentage colonne et le résidu standardisé.

```
                  FAM   Situation matrimoniale

                  |célibat  marié(e) Couple
V142              |       1|       2|       3|
Mariage   --------+--------+--------+--------+
                1 |    54  |     0  |    16  |    70
  Pas important   |  59.5  |   1.9  |   8.6  | 17.2%
                  | 77.1%  |   .0%  | 22.9%  |
                  | 15.6%  |   .0%  | 32.0%  |
                  |   -.7  |  -1.4  |   2.5  |
                  +--------+--------+--------+
                2 |    48  |     1  |    10  |    59
  Peu important   |  50.2  |   1.6  |   7.2  | 14.5%
                  | 81.4%  |  1.7%  | 16.9%  |
                  | 13.9%  |  9.1%  | 20.0%  |
                  |   -.3  |   -.5  |   1.0  |
                  +--------+--------+--------+
                3 |    69  |     2  |     3  |    74
  Moyennement     |  62.9  |   2.0  |   9.1  | 18.2%
  important       | 93.2%  |  2.7%  |  4.1%  |
                  | 19.9%  | 18.2%  |  6.0%  |
                  |    .8  |    .0  |  -2.0  |
                  +--------+--------+--------+
                4 |    86  |     1  |     8  |    95
  Important       |  80.8  |   2.6  |  11.7  | 23.3%
                  | 90.5%  |  1.1%  |  8.4%  |
                  | 24.9%  |  9.1%  | 16.0%  |
                  |    .6  |  -1.0  |  -1.1  |
                  +--------+--------+--------+
                5 |    89  |     7  |    13  |   109
  Très important  |  92.7  |   2.9  |  13.4  | 26.8%
                  | 81.7%  |  6.4%  | 11.9%  |
                  | 25.7%  | 63.6%  | 26.0%  |
                  |   -.4  |   2.4  |   -.1  |
                  +--------+--------+--------+
                     346       11       50      407
            Total  85.0%     2.7%    12.3%   100.0%

Khi2= 22.98692 ddl=8  Significativité:0.00338
```

La liaison est très forte et cela semble cohérent avec un certain bon sens, mais ce n'est pas là le but du jeu... Dans l'ACP le Mariage (142) se retrouve en haut à droite les sujets mariés ou vivant en couple se retrouvent opposés au Mariage vers le bas, il y a donc une contradiction apparente et donc bien une situation type «paradoxe de Simpson». Il serait possible d'essayer de démonter le paradoxe mais il y a plusieurs cas similaires dans le traitement et c'est plus donner l'ordre d'idée que résoudre explicitement des paradoxes qui me semble intéressant sur ce fichier, compte tenu de sa taille. Ce qui se passe entre les variables sexe et filière est beaucoup plus simple et surtout graphiquement facile à identifier. Il y a 4 groupes qui se détachent clairement sur le graphique, sur le premier axe on voit que les deux filières, IUT et Mirail sont distinctes, on voit aussi que les garçons et les filles se distinguent mais lorsqu'on regarde les quatre groupes on constate les différences deviennent moins évidentes. Cela est dû en grande partie au premier tableau croisé entre sexe et filière et qui apparaît sur le graphique. Il y a relativement beaucoup plus de garçons en IUT qu'en psychologie et le fait que les garçons soient sur la droite du graphique est surtout dû à cette liaison. Mais bien sûr on pourrait invoquer le fait que le choix de la filière d'étude est plutôt une conséquence qu'une cause de la position sur cet axe, mais il s'agit là d'une interprétation trop fine pour cet exposé.

D'après ces constatations j'ai pensé qu'il était possible d'utiliser ces interprétations pour revenir en arrière et essayer l'autre solution évoquée plus haut. Ayant vu 4 groupes faciles à comprendre et bien distincts sur les 2 axes de l'ACP, pourquoi ne pas prendre ces 4 groupes pour essayer de les distinguer à partir des 211 mots de départ, plutôt que d'utiliser les facteurs de l'ACP. Cela ne correspond à aucun a priori méthodologique, mais plutôt à la volonté de montrer que l'on ne doit pas, à mon avis, rester figé dans une démarche.

Mais ici aussi, j'ai décidé de faire simple et même ultra simple. Une première démarche à consisté à faire des analyses de variance systématiques des 211 variables (mots) sur la variable groupe créée en croisant sexe et filière, cela a donné des choses intéressantes que je ne présente pas pour des raisons de place et de temps. Puis j'ai opté pour un classement des mots, d'après la moyenne obtenue par le mot dans chacun des groupes. Les résultats des analyses de variance sont «consistants» avec la lecture de ce tableau les mots pour qui il y a une différence entre les groupes obtiennent des classements très contrastés suivant les groupes. Ce qui est très différent des résultats de l'ACP c'est que les variables qui discriminent les groupes dans le même sens ne sont pas nécessairement corrélées entre elles et peuvent donc être éloignées dans l'ACP.

Dans la lecture de ce tableau il y a le classement du mot pour l'ensemble de la population, puis le classement pour le Mirail, puis pour l'IUT, puis pour les filles du Mirail, puis pour celles de l'IUT, puis pour les garçons du

Mirail et enfin pour ceux de l'IUT.

Num	Intitulé	Tot	Mir	IUT	Fil	Gar	FM	FI	GM	GI
V001	Loyauté	30	29	36	33	19	32	39	10	30
V002	Charité	135	132	136	132	144	131	128	139	147
V003	Fidélité	14	13	13	10	25	12	11	31	23
V004	persévérance	50	55	34	51	41	57	34	46	38
V005	Liberté	5	6	4	7	2	7	4	1	3
V006	charme	124	131	101	125	105	135	97	110	98
V007	Décentration	154	140	188	151	175	137	194	163	182
V008	Détermination	52	62	41	57	37	62	45	48	31
V009	Franchise	7	7	7	6	9	6	7	13	8
V010	piété	193	194	189	193	188	193	195	198	180
V011	Loisir	54	67	28	61	26	70	37	40	18
V012	révolution	186	185	187	186	180	185	187	186	175
V013	volonté	21	22	20	24	15	23	21	16	19
V014	sentiment	9	11	6	9	7	13	9	8	6
V015	développement	111	112	113	111	99	112	123	115	90
V016	mansuétude	165	161	172	163	179	161	166	179	178
V017	célibat	203	203	203	203	202	203	207	204	200
V018	patriotisme	199	202	196	200	193	202	197	196	191
V019	affection	24	20	31	21	39	19	33	43	35
V020	progrès	71	84	52	76	46	87	52	67	37
V021	fraternité	60	61	61	56	90	55	58	103	81
V022	abstinence	208	208	209	208	210	208	209	210	207
V023	tout-puissance	210	210	208	210	204	210	208	207	203
V024	vocation	149	150	149	149	145	148	157	169	120
V025	sensualité	116	118	103	120	72	122	110	64	82
V026	professionnalisme	57	64	48	55	63	63	43	68	68
V027	désintéressement	196	195	204	196	199	195	206	187	204
V028	solidarité	33	31	42	31	69	27	38	73	71
V029	acceptation	102	105	102	100	120	101	96	127	105
V030	communion	192	191	197	192	194	191	193	188	199
V031	virginité	206	205	206	206	208	205	201	201	211
V032	sensibilité	42	38	56	41	57	35	59	63	60
V033	Accueil	51	50	51	49	56	52	50	83	46
V034	clémence	136	129	147	135	139	129	143	114	153
V035	dialogue	10	10	15	14	12	8	17	9	17

Num	Intitulé	Tot	Mir	IUT	Fil	Gar	FM	FI	GM	GI
V036	plaisir	19	26	5	20	6	25	13	21	2
V037	vertu	152	156	148	153	155	156	145	154	154
V038	pluralisme	150	148	159	157	112	154	163	87	132
V039	dynamisme	45	53	25	42	59	53	20	90	45
V040	innovation	103	110	86	109	84	110	95	124	57
V041	conciliation	121	119	123	119	129	119	119	109	139
V042	écologie	113	116	100	116	74	118	112	81	76
V043	hospitalité	70	69	77	74	50	69	87	58	44
V044	méthode	138	142	132	141	137	143	132	129	131
V045	prudence	106	101	116	105	119	102	109	95	140
V046	sérénité	82	74	97	80	87	76	93	65	101
V047	unité	157	157	160	156	165	157	154	158	166
V048	repentance	189	186	191	188	187	187	191	184	189
V049	tempérance	156	153	166	155	164	153	162	144	171
V050	amour	2	2	3	2	3	1	3	5	4
V051	compétitivité	183	188	154	185	158	188	164	189	113
V052	don	147	144	161	148	135	144	169	131	125
V053	idéal	130	127	129	129	123	128	127	104	133
V054	discrétion	112	108	131	110	124	106	136	123	119
V055	pouvoir	197	200	184	198	192	199	183	197	184
V056	santé	3	3	2	3	4	3	2	4	5
V057	pratique	131	138	105	131	114	138	108	136	93
V058	action	81	90	58	82	80	89	64	116	59
V059	combativité	110	130	50	112	83	126	54	142	34
V060	méticulosité	160	162	145	159	154	162	150	176	135
V061	honneur	118	125	83	122	61	130	91	78	61
V062	rénovation	182	181	183	184	177	182	186	174	176
V063	spontanéité	66	63	81	60	111	61	65	100	122
V064	coopération	78	75	87	71	118	75	72	101	130
V065	égalité	16	18	18	16	29	18	12	17	36
V066	intelligence	75	78	67	81	43	77	85	66	32
V067	sobriété	159	155	171	158	170	155	168	167	172
V068	réceptivité	69	68	79	70	58	67	83	51	66
V069	complicité	25	23	27	23	34	21	29	55	25
V070	disponibilité	63	54	82	59	78	54	81	72	84

Num	Intitulé	Tot	Mir	IUT	Fil	Gar	FM	FI	GM	GI
V071	intimité	20	19	21	19	24	20	23	32	16
V072	organisation	86	95	62	86	94	98	55	85	97
V073	religion	202	201	205	202	206	200	205	206	206
V074	surêté	120	115	138	117	143	113	135	138	146
V075	fécondité	98	89	120	88	152	85	101	141	161
V076	justice	23	24	23	26	14	26	24	11	20
V077	espérance	44	40	59	45	35	41	69	26	58
V078	coquetterie	185	182	186	182	200	181	181	199	198
V079	autorité	180	180	174	179	178	179	172	185	173
V080	autocontrôle	61	60	64	66	32	64	70	14	63
V081	connaissance	48	46	54	50	31	46	60	25	47
V082	douceur	35	36	40	34	51	37	35	33	74
V083	émotion	26	25	24	27	18	24	28	18	21
V084	observation	99	91	117	102	75	93	129	75	80
V085	rédemption	188	187	190	187	189	186	189	192	185
V086	sympathie	31	30	33	30	52	29	31	57	53
V087	beauté	170	172	162	171	156	170	161	165	150
V088	honnêté	4	4	9	4	10	4	8	15	10
V089	maîtrise de soi	36	35	43	40	33	38	47	28	42
V090	rigueur	119	122	114	123	96	121	124	106	87
V091	transparence	158	158	165	160	146	159	167	135	151
V092	partage	53	48	78	52	97	47	68	71	110
V093	joie	12	14	10	11	16	15	6	24	12
V094	cohérence	88	80	107	89	81	79	114	80	88
V095	courage	73	77	66	78	49	78	77	54	49
V096	mysticisme	200	199	200	201	190	201	200	175	196
V097	féminité	79	81	73	73	106	83	51	60	145
V098	spiritualité	173	171	173	173	159	172	173	150	168
V099	raison	100	103	89	106	66	105	98	76	64
V100	curiosité	105	107	106	107	101	111	88	53	134
V101	alliance	140	136	156	136	171	134	142	153	179
V102	générosité	41	37	55	38	76	36	48	52	92
V103	affectivité	27	21	29	22	38	22	27	35	48
V104	chance	148	152	140	147	136	151	139	134	124
V105	infaillibilité	198	196	199	197	201	197	198	195	201

Num	Intitulé	Tot	Mir	IUT	Fil	Gar	FM	FI	GM	GI
V106	objectivité	89	86	92	92	62	88	99	62	69
V107	simplicité	55	51	69	53	92	51	66	89	94
V108	vérité	6	5	8	5	13	5	5	19	11
V109	austérité	205	204	201	204	203	204	202	208	202
V110	pudeur	167	160	182	165	183	160	178	173	187
V111	démocratie	43	41	53	44	30	42	61	38	24
V112	dignité	68	70	71	72	47	71	84	50	50
V113	obéissance	178	178	179	178	184	178	174	178	188
V114	secret	126	128	118	124	133	127	111	117	143
V115	serviabilité	108	102	111	104	122	103	105	96	144
V116	vie	1	1	1	1	1	2	1	3	1
V117	expansivité	166	167	163	166	162	166	165	172	148
V118	famille	13	12	14	15	17	11	16	23	14
V119	ascèse	204	206	198	205	197	207	199	200	193
V120	caressant	153	154	158	152	157	152	153	151	167
V121	économie	168	173	157	167	163	171	155	168	155
V122	indépendance	29	32	19	29	21	31	22	36	15
V123	patience	56	49	80	54	70	49	82	77	70
V124	profession	37	42	38	36	60	40	36	84	55
V125	rêve	22	28	16	25	8	30	18	12	7
V126	tendresse	18	17	22	17	28	17	19	22	33
V127	vigilance	114	114	108	113	125	114	100	121	126
V128	affabilité	181	183	169	180	186	183	160	190	183
V129	culture	46	47	45	47	54	48	42	42	65
V130	calme	80	82	76	79	79	84	71	70	85
V131	harmonie	49	44	74	46	91	43	62	61	107
V132	opiniâtreté	175	175	178	174	172	174	176	157	177
V133	soin	109	99	122	103	128	99	120	118	128
V134	tradition	172	163	181	169	176	164	182	164	181
V135	précaution	127	126	127	127	117	124	133	122	108
V136	Autonomie	32	33	35	32	40	33	30	34	56
V137	Bonté	77	65	112	75	89	66	116	56	109
V138	Éducation	28	27	26	28	23	28	26	20	28
V139	Paix	8	8	11	8	5	9	10	2	9
V140	Ténacité	104	109	90	108	93	108	94	119	77

Num	Intitulé	Tot	Mir	IUT	Fil	Gar	FM	FI	GM	GI
V141	Propriété	164	170	143	170	127	173	156	146	96
V142	Mariage	146	149	125	140	174	146	107	177	170
V143	Amabilité	62	57	72	58	95	56	63	74	106
V144	Conversion	190	190	195	190	195	190	190	191	197
V145	Indulgence	93	87	99	87	116	86	90	120	116
V146	Paternalisme	176	177	177	176	181	177	180	193	162
V147	Modestie	96	97	93	98	85	94	104	111	75
V148	Béatitude	195	197	193	195	198	196	196	203	190
V149	Compréhension	34	34	49	35	36	34	49	30	52
V150	Authenticité	67	59	95	68	64	60	102	47	79
V151	Imagination	40	43	46	43	27	44	53	27	29
V152	Tolérance	17	16	32	18	20	16	32	6	41
V153	Rectitude	184	184	180	183	185	184	175	182	186
V154	Réconciliation	84	83	91	85	100	80	89	113	91
V155	Compassion	132	135	119	130	141	136	115	133	138
V156	Fougue	151	147	164	150	168	149	159	149	174
V157	Foi	194	193	194	194	191	194	192	180	195
V158	Lucidité	76	71	96	77	71	72	106	79	73
V159	Noblesse	207	207	207	207	207	206	204	202	209
V160	Travail	39	45	30	37	67	45	25	82	62
V161	Pureté	161	165	144	161	169	165	138	171	163
V162	Sécurité	64	66	57	65	65	65	67	105	39
V163	Créativité	59	58	63	63	44	59	76	45	43
V164	Intérêt	142	143	142	144	140	142	141	156	117
V165	Moralité	123	121	126	118	147	116	122	162	141
V166	Ambition	65	79	37	64	77	73	40	147	27
V167	Discipline	162	164	153	162	149	163	158	170	142
V168	Intuition	125	120	137	134	53	125	151	29	78
V169	Originalité	85	93	70	93	55	96	73	44	72
V170	Respect	11	9	17	13	11	10	15	7	22
V171	Ruse	187	189	176	189	160	189	184	181	136
V172	Conscience	58	56	65	62	45	58	78	49	40
V173	Masculinité	174	174	170	175	132	176	177	126	127
V174	Passion	47	52	39	48	48	50	41	93	26
V175	Clairvoyance	122	117	139	121	131	117	131	97	149

Num	Intitulé	Tot	Mir	IUT	Fil	Gar	FM	FI	GM	GI
V176	Altruisme	115	104	151	114	126	107	148	69	156
V177	Sagesse	92	76	121	94	82	81	125	41	115
V178	Réhabilitation	169	168	168	172	150	169	171	140	164
V179	Légalité	95	92	104	95	102	91	103	98	99
V180	Humilité	128	123	141	128	121	120	146	128	100
V181	Censure	201	198	202	199	205	198	203	205	205
V182	Concentration	137	139	135	139	115	141	137	107	123
V183	Sincérité	15	15	12	12	22	14	14	37	13
V184	Protection	74	72	84	69	108	68	74	108	104
V185	Subtilité	134	133	133	138	86	139	144	88	89
V186	Droiture	155	159	152	154	166	158	147	155	169
V187	Croyance	191	192	192	191	196	192	188	194	194
V188	Libéralité	145	146	130	143	167	147	117	160	165
V189	Décision	101	96	110	99	103	97	113	91	112
V190	Assurance	107	111	88	101	130	109	79	130	114
V191	Exigence	163	166	155	164	161	167	152	161	158
V192	Pardon	72	73	75	67	110	74	56	86	129
V193	Compromis	139	137	150	137	153	133	149	159	152
V194	Compétence	94	100	68	96	73	100	75	92	67
V195	Promesse	83	85	85	83	98	82	80	102	95
V196	Force	143	145	124	145	142	145	126	166	118
V197	Don de soi	129	124	134	126	134	123	134	125	137
V198	Diplomatie	97	94	98	97	104	95	92	94	111
V199	Intégrité	91	88	94	91	88	90	86	59	103
V200	Responsabilité	38	39	44	39	42	39	44	39	54
V201	Magnanimité	171	169	167	168	173	168	170	183	160
V202	Devoir	133	134	128	133	138	132	130	148	121
V203	Chasteté	211	211	211	211	211	211	211	211	208
V204	Modération	177	176	185	177	182	175	185	152	192
V205	Sainteté	209	209	210	209	209	209	210	209	210
V206	Dévouement	141	141	146	142	148	140	140	137	159
V207	Sérieux	90	98	60	84	113	92	57	143	86
V208	Privé	144	151	109	146	109	150	121	132	83
V209	Séduction	87	106	47	90	68	104	46	99	51
V210	Gratitude	117	113	115	115	107	115	118	112	102
V211	Méditation	179	179	175	181	151	180	179	145	157

Il est évident qu'on ne peut retenir tous les résultats d'un tel tableau et qu'un travail de butinage est encore nécessaire mais cela est beaucoup plus simple à lire et à comprendre que des corrélations variables facteurs ou des contributions dans une AFC. Il est assez rare qu'un mot ait de très grandes variations de classement. Le mot intuition (168) doit être un de ceux ayant la plus grande variabilité, il est classé 29ème chez les garçons du Mirail, 78ème chez les garçons de l'IUT, 125ème chez les filles du Mirail et seulement 151ème chez les filles de l'IUT. Les changements, même faibles, dans les premières places sont intéressants à regarder. La vie (116) qui est en première position pour la population entière, l'est aussi pour les l'ensemble des filles et l'ensemble des garçons, elle est aussi pour l'ensemble du Mirail et l'ensemble de l'IUT, mais pour les garçons du Mirail elle ne vient qu'en troisième position derrière la liberté (5) et la paix (139), et pour les filles du Mirail elle est en deuxième position derrière l'amour (50). La liberté, la paix et l'amour sont tous les 3 des mots très bien classés dans tous les groupes. La deuxième place la plus originale est celle obtenue par le plaisir (36) chez les garçons de l'IUT, car ce mot est seulement 13ème chez les filles de l'IUT, 21ème chez les garçons du Mirail et seulement 25ème chez les filles du Mirail.

Ici se situe la fin de la première partie de cet exposé en deux parties. La suite aura lieu le Vendredi 26 Février, même lieu, même heure. J'insiste sur le côté inachevé de ce document qui sera intégré à celui plus global contenant les deux parties, en tenant compte des remarques et critiques qui seront plus que bienvenues. dans la deuxième partie les traitements utiliseront une option plus qualitative, AFCM et croisements systématiques, puis modèle log-linéaire pour mettre en évidence une situation similaire au paradoxe de Simpson. Il y aura donc une certaine ressemblance avec l'exposé du 16 Janvier 1998 qui peut être récupéré sur le web à l' adresse suivante:

http://www.cict.fr/cict/personnel/stpierre/expose-16-01-98/expose.html

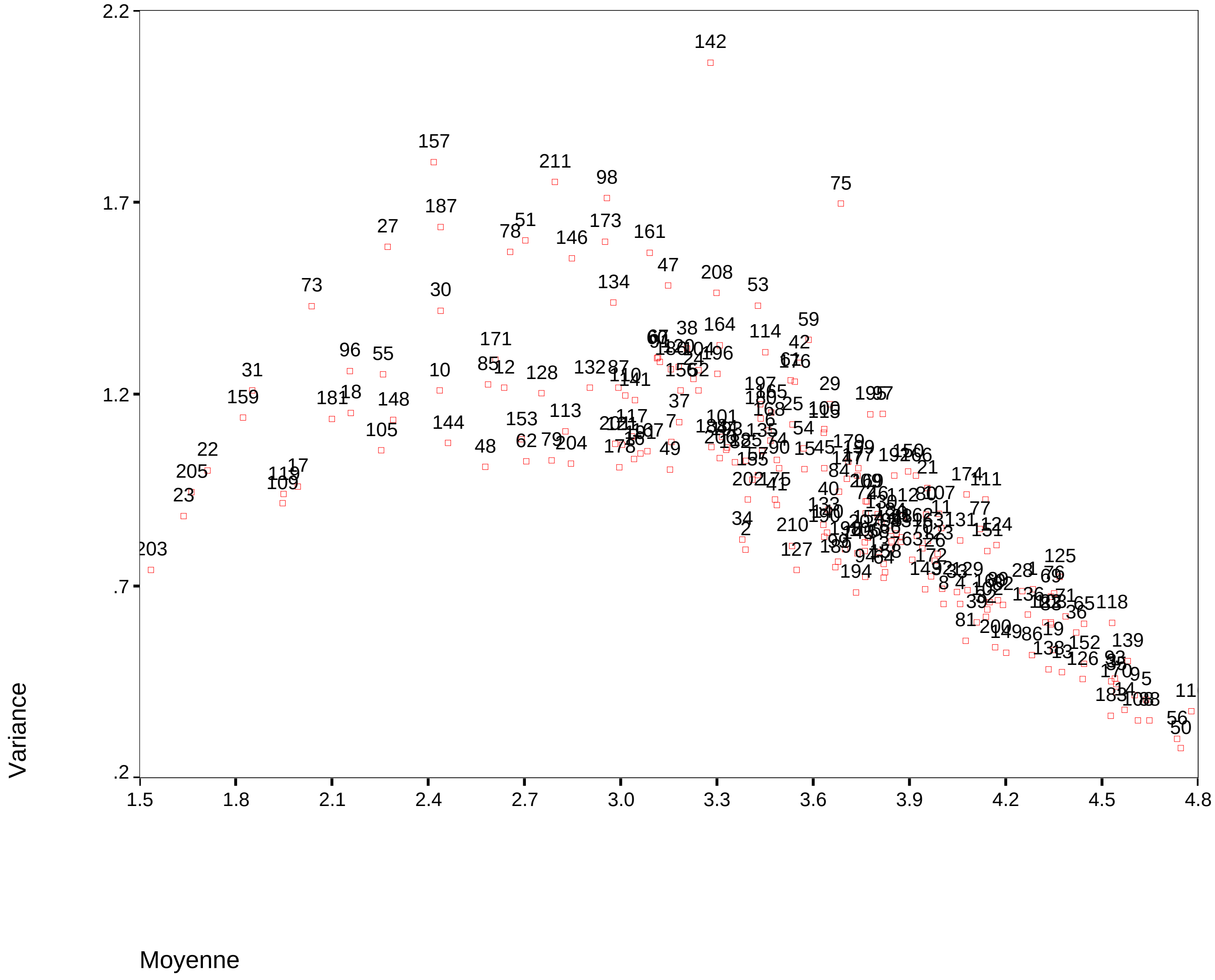

FIG. 1 – .