



## Le rêve des Français.

Si vous gagniez au Loto que feriez vous ? L'analyse de cet extrait d'une étude menée par le Credoc en 1992 permet d'illustrer les méthodes d'analyse quantitatives et qualitatives disponibles dans Sphinx :

**Analyse quantitative** : tri à plats, tris croisés, analyses multivariées et typologie

**Analyse qualitative** : approximation lexicale, sémiométrie et analyse de contenu par code book.

## Extrait de l'étude ménage du Credoc 1992

Si vous gagniez le gros lot au Loto, que feriez-vous ?

### Sexe

- Homme  
 Femme

### Diplôme

- |                                   |                                     |
|-----------------------------------|-------------------------------------|
| <input type="checkbox"/> Aucun    | <input type="checkbox"/> Bac+2 BTS  |
| <input type="checkbox"/> CEP BEPC | <input type="checkbox"/> Bac+4 DESS |
| <input type="checkbox"/> CAP BEP  | <input type="checkbox"/> Refus      |
| <input type="checkbox"/> Bac      | <input type="checkbox"/> Autres     |

### Situation matrimoniale

- |  |                                     |
|--|-------------------------------------|
| <input type="checkbox"/> Célibataire         | <input type="checkbox"/> Veuf(ve)   |
| <input type="checkbox"/> Marié(e)            | <input type="checkbox"/> Divorcé(e) |
| <input type="checkbox"/> Vivant maritalement | <input type="checkbox"/> Séparé(e)  |

### Catégorie socio-professionnelle

- |  |  |
|--|--|
| <input type="checkbox"/> Agriculteurs            | <input type="checkbox"/> Employés        |
| <input type="checkbox"/> Commerçant, artisan     | <input type="checkbox"/> Ouvriers        |
| <input type="checkbox"/> Cadre.Prof.Intell. Sup. | <input type="checkbox"/> Retraités       |
| <input type="checkbox"/> Prof.Intermédiaires     | <input type="checkbox"/> Inactifs, Autre |

### Taille de la commune

- |   |  |
|---|--|
| <input type="checkbox"/> - de 2 000       | <input type="checkbox"/> + de 100 000      |
| <input type="checkbox"/> 2 000-20 000     | <input type="checkbox"/> Région parisienne |
| <input type="checkbox"/> 20 000 - 100 000 |  |

## Un échantillon représentatif des français en 1992

### AGE

Moyenne = 41.28

18-24	158	15.9%
25-34	276	27.8%
35-49	280	28.2%
50-64	161	16.2%
65+	118	11.9%
<b>Total</b>	<b>993</b>	<b>100.0%</b>

### SEXE

Homme	444	44.7%
Femme	549	55.3%
<b>Total</b>	<b>993</b>	<b>100.0%</b>

### CSP

Agriculteurs	47	4.7%
Commerçant, artisan	47	4.7%
Cadre.Prof.Intell. Sup.	152	15.3%
Prof.Intermédiaires	78	7.9%
Employés	277	27.9%
Ouvriers	132	13.3%
Retraités	176	17.7%
Inactifs, Autre	84	8.5%
<b>Total</b>	<b>993</b>	<b>100.0%</b>

### DIPLOME

Aucun	101	10.2%
CEP BEPC	245	24.7%
CAP BEP	240	24.2%
Bac	141	14.2%
Bac+2 BTS	122	12.3%
Bac+4 DESS	135	13.6%
Refus	1	0.1%
Autres	8	0.8%
<b>Total</b>	<b>993</b>	<b>100.0%</b>

### SIT. MATRIMONIALE

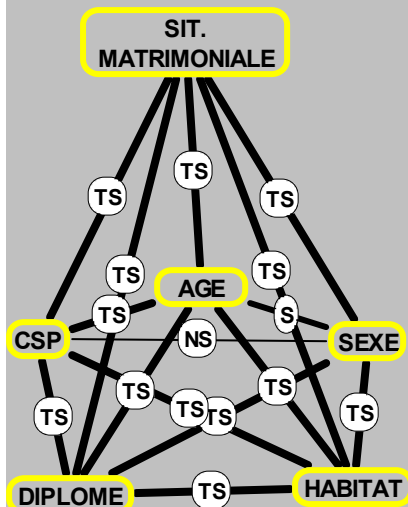
Célibataire	226	22.8%
Marié(e)	538	54.2%
Vivant maritalement	108	10.9%
Veuf(ve)	75	7.6%
Divorcé(e)	33	3.3%
Séparé(e)	13	1.3%
<b>Total</b>	<b>993</b>	<b>100.0%</b>

### HABITAT

- de 2 000	335	33.7%
2 000-20 000	154	15.5%
20 000 - 100 000	196	19.7%
+ de 100 000	177	17.8%
Région parisienne	131	13.2%
<b>Total</b>	<b>993</b>	<b>100.0%</b>

**Tableaux croisés :  
analyse bi variée**

*Graphes des relations entre les variables socio-démographique*



Chi2, Anova : TS <= 1% < S <= 5% < PS <= 15%  
Corr. : TS >= 0.8 > S >= 0.6 > PS >= 0.4 > NS

*Il faut examiner 15 tableaux pour savoir comment les caractéristiques de cette population se combinent entre elles...  
L'analyse multivariée qui suit permet une représentation beaucoup plus synthétique.*

**CSP x Diplôme**

	Aucun	CEP BEPC	CAP BEP	Bac	Bac+2 BTS	Bac+4 DESS	Refus	Autres	Total
Agriculteurs	8.5%	48.9%	14.9%	17.0%	8.5%	0.0%	0.0%	2.1%	100.0%
Commerçant, artisan	14.9%	27.7%	36.2%	14.9%	4.3%	2.1%	0.0%	0.0%	100.0%
Cadre.Prof.Intell. Sup.	3.9%	8.6%	10.5%	13.2%	20.4%	41.4%	0.0%	2.0%	100.0%
Prof.Intermédiaires	1.3%	9.0%	12.8%	19.2%	28.2%	29.5%	0.0%	0.0%	100.0%
Employés	10.8%	19.5%	35.7%	17.0%	11.6%	5.1%	0.0%	0.4%	100.0%
Ouvriers	15.9%	35.6%	37.9%	8.3%	2.3%	0.0%	0.0%	0.0%	100.0%
Retraités	12.5%	42.0%	19.9%	8.0%	6.8%	8.5%	0.6%	1.7%	100.0%
Inactifs, Autre	11.9%	16.7%	7.1%	22.6%	19.0%	22.6%	0.0%	0.0%	100.0%
<b>Total</b>	<b>10.2%</b>	<b>24.7%</b>	<b>24.2%</b>	<b>14.2%</b>	<b>12.3%</b>	<b>13.6%</b>	<b>0.1%</b>	<b>0.8%</b>	<b>100.0%</b>

p = <0.1% ; chi2 = 372.04 ; ddl = 49 (TS)

**Age x Sexe**

	AGE
Homme	39.92
Femme	42.38
<b>Total</b>	<b>41.28</b>

p = 1.9% ; F = 5.41 (S)

**Habitat x Age**

	AGE
- de 2 000	43.37
2 000-20 000	40.32
20 000 - 100 000	41.64
+ de 100 000	37.59
Région parisienne	41.51
<b>Total</b>	<b>41.28</b>

p = 0.5% ; F = 3.73 (TS)

**Age x Sexe**

	AGE
Aucun	45.87
CEP BEPC	50.46
CAP BEP	37.75
Bac	34.50
Bac+2 BTS	34.39
Bac+4 DESS	39.53
Refus	57.00
Autres	59.88
<b>Total</b>	<b>41.28</b>

p = <0.1% ; F = 25.20 (TS)

**Age x CSP**

	AGE
Agriculteurs	43.32
Commerçant, artisan	39.77
Cadre.Prof.Intell. Sup.	39.20
Prof.Intermédiaires	35.49
Employés	35.13
Ouvriers	36.46
Retraités	63.05
Inactifs, Autre	32.37
<b>Total</b>	<b>41.28</b>

p = <0.1% ; F = 91.08 (TS)

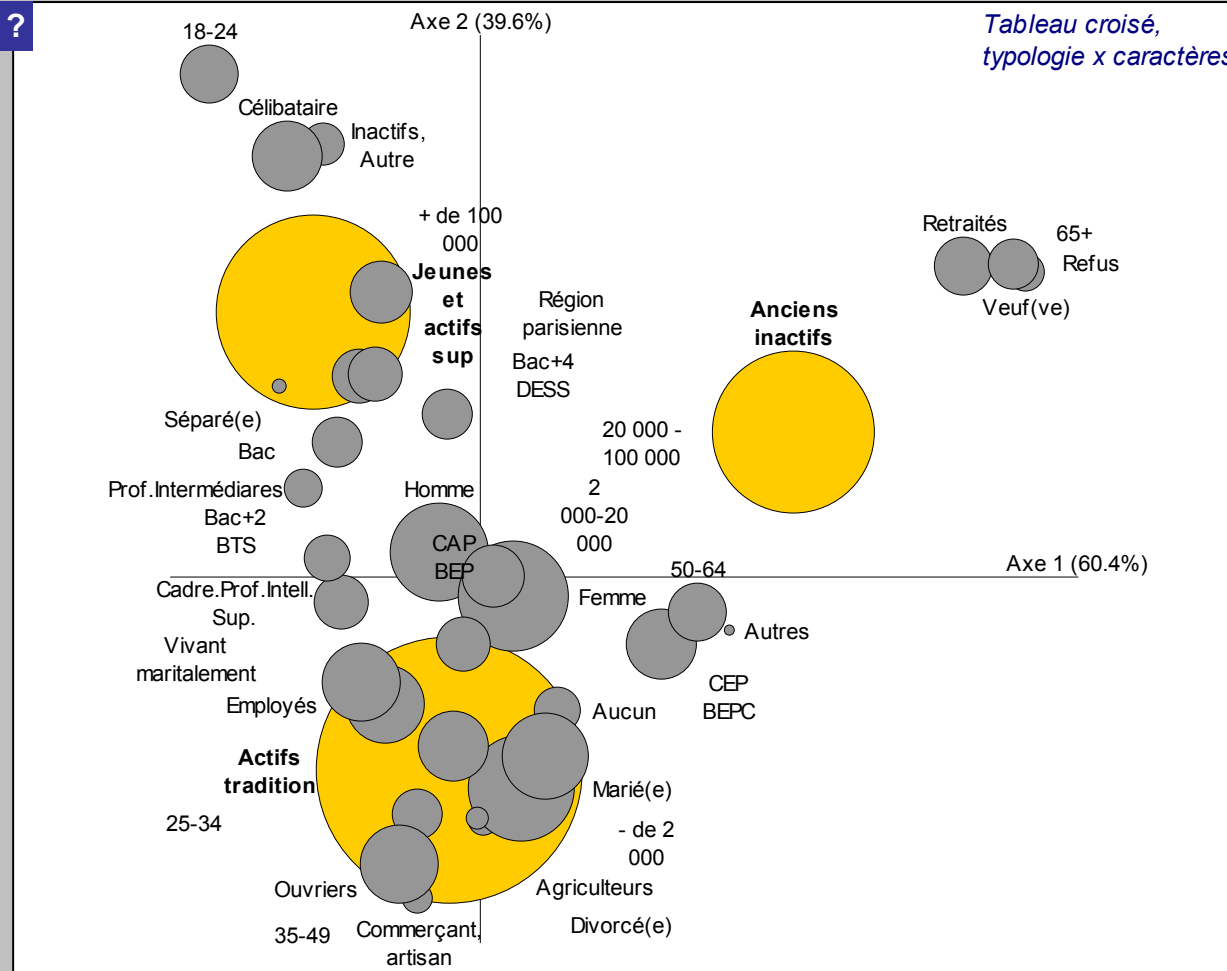


### Typologie : 3 Frances ?

De la position des individus sur la carte précédente on construit la typologie décrite ci contre

#### Âges moyens

	AGE
Anciens inactifs	66.02
Jeunes et actifs sup	27.50
Actifs tradition	39.22
Homme	39.92
Femme	42.38
Célibataire	29.65
Marié(e)	44.42
Vivant maritalement	31.42
Veuf(ve)	67.08
Divorcé(e)	44.48
Séparé(e)	38.31
- de 2 000	43.37
2 000-20 000	40.32
20 000 - 100 000	41.64
+ de 100 000	37.59
Région parisienne	41.51



## Les lexiques

### Si vous gagniez au Loto, que feriez vous ?

On peut se faire une idée du contenu des réponses par l'examen des termes les plus fréquemment utilisés.

Le lexique brut exprime la stimulation de l'égo (je, j) sur un mode plutôt indéfini (une, un) que défini (le la l)...

Le lexique lemmatisé (chaque terme est ramené à sa forme racine) renvoie l'essentiel du contenu

Les substantifs donnent les objets évoqués et les verbes les actions...

On accède ainsi très rapidement au contenu des réponses comme les mots clé donnent celui d'un livre.

## Lexique brut

je	1182
j	653
de	613
une	566
en	464
à	412
le	385
des	362
un	362
maison	355
enfants	286
la	286
pour	280
l	257
mes	244
les	240
et	208
voiture	198
dans	194
voyage	192
ferais	182
voyages	158
ma	156
achèterais	150
aux	137
famille	122
pas	120
argent	118

## Sans mots outils

maison	355
enfants	286
voiture	198
voyage	192
ferais	182
voyages	158
achèterais	150
famille	122
argent	118
acheter	115
donnerais	113
placerais	111
faire	107
reste	97
monde	95
placer	84
travailler	79
vacances	79
achète	69
voyagerais	68
immobilier	67
partie	64
vie	64
dons	62
aider	61
côté	61
appartement	58
profiter	58

## Lemmatisé

maison	354
acheter	350
voyage	310
faire	297
enfant	264
placer	230
voiture	194
donner	187
voyager	148
aider	122
argent	118
famille	117
partir	91
travailler	91
monde	81
profiter	81
investir	76
don	75
immobilier	74
beau	72
mettre	71
placement	71
appartement	65
reste	65
changer	63
aller	59
arrêter	58
association	57

## Les substantifs

maison	354
voyage	310
enfant	259
voiture	194
argent	118
famille	117
monde	81
don	75
placement	71
appartement	65
reste	65
association	57
immobilier	55
partie	54
oeuvre	51
vie	49
tour	45
vacance	43
achat	41
loisir	39
besoin	36
placer	35
gens	34
ami	33
cadeau	33
plaisir	33
cancer	27
enfants	24

## Les verbes

acheter	350
faire	296
placer	195
donner	187
voyager	148
aider	122
partir	91
travailler	91
profiter	81
investir	76
mettre	71
changer	63
aller	59
arrêter	58
améliorer	48
prendre	38
vivre	37
partager	36
payer	33
rester	33
épargner	32
avoir_besoin	30
pouvoir	28
dépenser	27
distribuer	26
faire_plaisir	26
garder	26
faire_le_tour	25

## Les segments répétés

*L'examen du seul lexique est mutilant.  
Le sens des mots est aussi défini par  
leur contexte.*

*Les segments répétés soulèvent bien  
des ambiguïtés (arrêter de travailler) et  
révèlent les leitmotiv du corpus*

### Segments plus de 10 fois

acheter maison	140	14.1%
placer argent	58	5.8%
faire voyage	55	5.5%
arrêter travailler	52	5.2%
tour monde	44	4.4%
faire profiter	43	4.3%
donner enfant	40	4.0%
acheter voiture	37	3.7%
aider enfant	33	3.3%
investir immobilier	33	3.3%
mettre de_côté	30	3.0%
acheter appartement	29	2.9%
changer voiture	29	2.9%
faire don	28	2.8%
partir voyage	27	2.7%
faire_plaisir plaisir	26	2.6%
partir en_vacances	25	2.5%
beau voiture	21	2.1%
beau voyage	21	2.1%
placer reste	20	2.0%
faire_le_tour tour monde	19	1.9%
acheter beau	17	1.7%
partager enfant	17	1.7%
achat maison	15	1.5%
argent de_côté	15	1.5%





### Les usages selon les contextes

Les analyses intrinsèques qui précèdent sont menées sans référence au contexte et au statut des locuteurs.

Elles peuvent être complétées par l'analyse des spécificités lexicales.

Par exemple les tableaux ci contre montrent que les hommes et les femmes ont tendance à sur utiliser certains termes.....

Substantifs selon le sexe			
	Homme	Femme	
maison	126	211	37.4% 62.6%
<b>voyage</b>	<b>104</b>	<b>205</b>	33.7% 66.3%
enfant	84	162	34.1% 65.9%
voiture	82	111	42.5% 57.5%
argent	43	73	37.1% 62.9%
famille	48	68	41.4% 58.6%
<b>monde</b>	<b>41</b>	<b>39</b>	51.3% 48.8%
don	30	40	42.9% 57.1%
<b>placement</b>	<b>38</b>	<b>31</b>	55.1% 44.9%
appartement	23	40	36.5% 63.5%
association	21	35	37.5% 62.5%
<b>immobilier</b>	<b>31</b>	<b>24</b>	56.4% 43.6%
oeuvre	18	33	35.3% 64.7%
vie	25	23	52.1% 47.9%
<b>tour</b>	<b>27</b>	<b>18</b>	60% 40%
vacance	22	21	51.2% 48.8%
achat	14	23	37.8% 62.2%
loisir	19	20	48.7% 51.3%
besoin	13	23	36.1% 63.9%
gens	14	20	41.2% 58.8%
ami	11	22	33.3% 66.7%
cadeau	13	20	39.4% 60.6%
plaisir	10	23	30.3% 69.7%
cancer	7	20	25.9% 74.1%
enfants	9	15	37.5% 62.5%
logement	8	16	33.3% 66.7%
<b>pauvre</b>	<b>15</b>	<b>8</b>	65.2% 34.8%
banque	11	11	50% 50%
pays	13	9	59.1% 40.9%

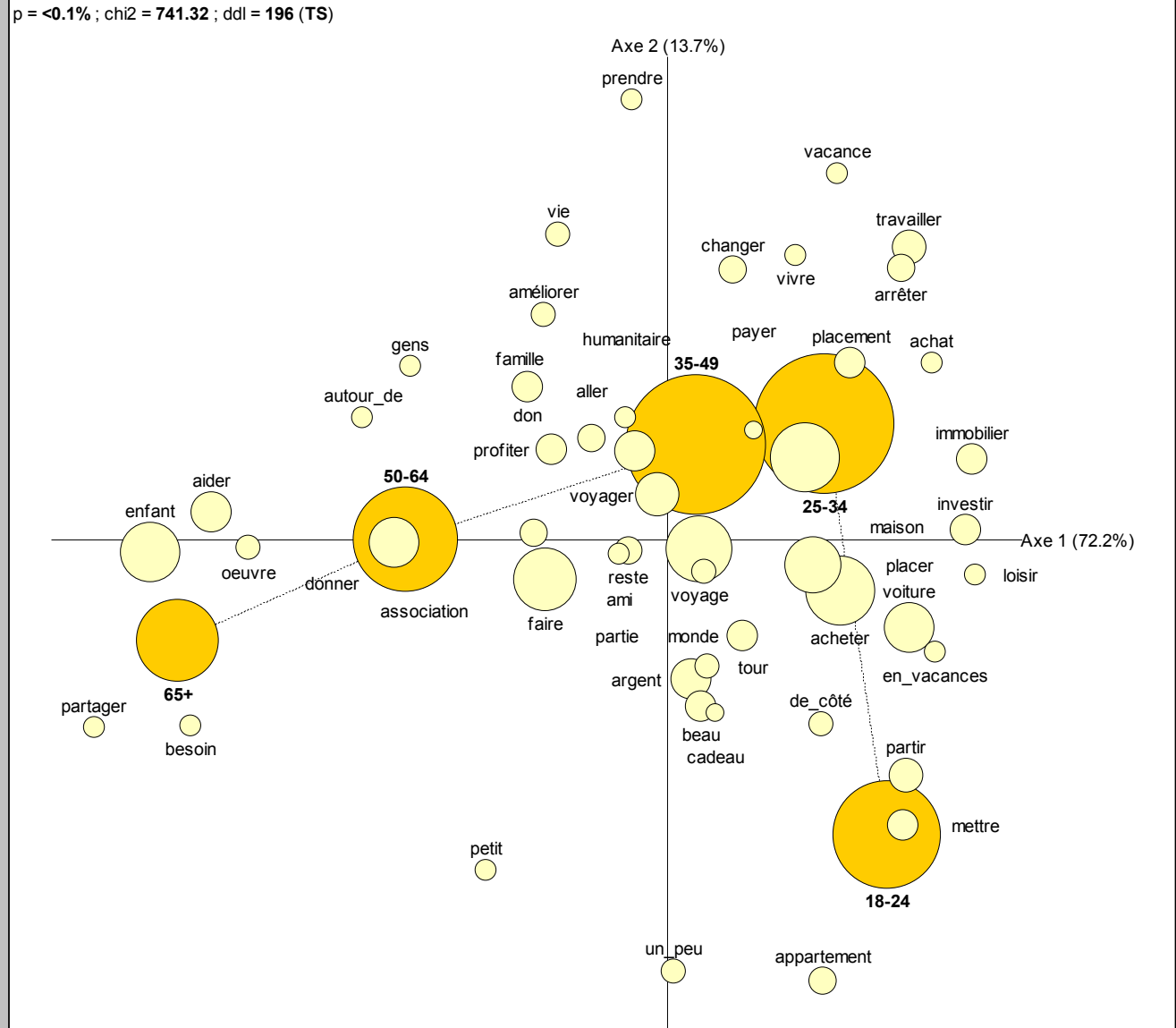
p = <0.1% ; chi2 = 57.71 ; ddl = 28 (TS)

Verbes selon le sexe			
	Homme	Femme	
acheter	135	191	41.4% 58.6%
faire	116	141	45.1% 54.9%
placer	91	101	47.4% 52.6%
<b>donner</b>	<b>57</b>	<b>114</b>	33.3% 66.7%
voyager	70	77	47.6% 52.4%
<b>aider</b>	<b>40</b>	<b>71</b>	36.0% 64.0%
partir	34	56	37.8% 62.2%
<b>travailler</b>	<b>53</b>	<b>37</b>	58.9% 41.1%
<b>profiter</b>	<b>43</b>	<b>36</b>	54.4% 45.6%
<b>investir</b>	<b>46</b>	<b>30</b>	60.5% 39.5%
mettre	26	45	36.6% 63.4%
changer	29	32	47.5% 52.5%
<b>aller</b>	<b>30</b>	<b>24</b>	55.6% 44.4%
<b>arrêter</b>	<b>33</b>	<b>25</b>	56.9% 43.1%
améliorer	19	29	39.6% 60.4%
prendre	20	17	54.1% 45.9%
vivre	17	19	47.2% 52.8%
<b>partager</b>	<b>10</b>	<b>26</b>	27.8% 72.2%
payer	11	21	34.4% 65.6%
<b>rester</b>	<b>19</b>	<b>13</b>	59.4% 40.6%
épargner	13	18	41.9% 58.1%
avoir_besoin	10	20	33.3% 66.7%
pouvoir	9	17	34.6% 65.4%
dépenser	15	12	55.6% 44.4%
distribuer	9	17	34.6% 65.4%
faire_plaisir	8	18	30.8% 69.2%
garder	9	16	36% 64%
faire_le_tour	13	12	52% 48%
essayer	14	9	60.9% 39.1%

p = <0.1% ; chi2 = 63.91 ; ddl = 28 (TS)

### Les usages selon les contextes : zones de langage selon l'âge

*L'analyse factorielle des correspondances permet de généraliser la recherche des effets de contexte. Ici le choix des mots et des préoccupations apparaissent comme clairement déterminés par l'âge*





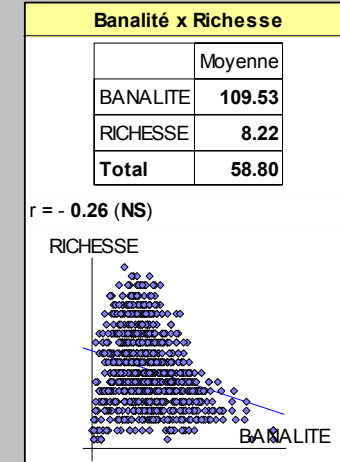
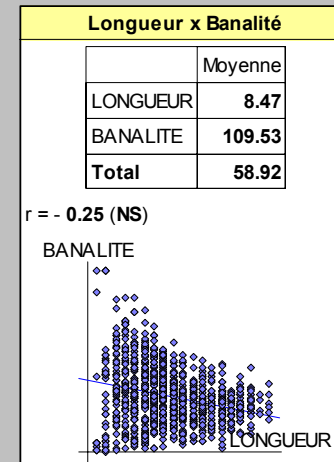
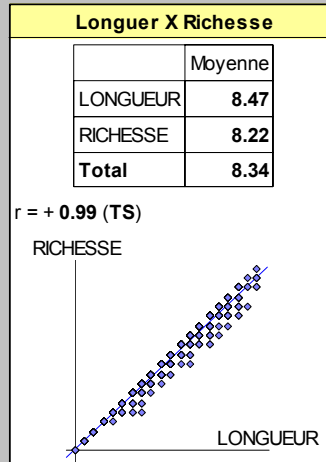
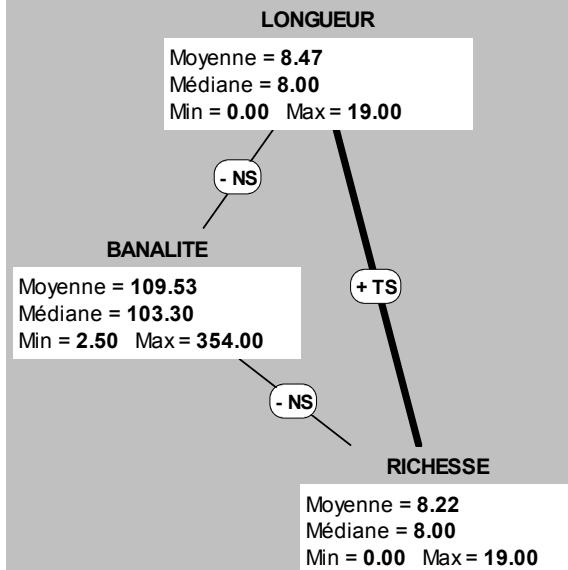


## Mesures lexicales : caractériser l'énonciation

La longueur des réponses, leur richesse (nombre de mots différents) et leur banalité (fréquence moyenne dans le corpus des mots utilisés par chacun) permet de caractériser les réponses.

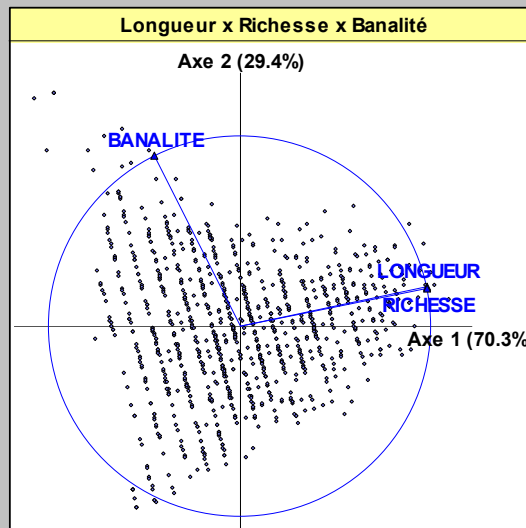
Les réponses longues sont les plus riches, mais la banalité est indépendante de la longueur ou de la richesse.

Les inactifs, ouvriers et agriculteurs font les réponses les plus courtes et les plus convenues (banalité).



Les nuages de point ci dessus illustrent bien la mesure du coefficient de corrélation.

Ci dessous l'ACP donne une image de ces corrélations. Longueur et richesse sont fortement liées. La banalité varie d'une manière totalement indépendante.



**Longueur x Richesse x Banalité selon la CSP**

	LONGUEUR	RICHESSE	BANALITE
Commerçant, artisan	9.32	9.15	98.85
Employés	8.91	8.64	113.40
Cadre.Prof.Intell. Sup.	8.55	8.33	103.05
Prof.Intermédiaires	8.38	8.18	104.76
Retraités	8.34	8.07	101.51
Inactifs, Autre	8.27	8.07	110.42
Ouvriers	7.86	7.60	125.10
Agriculteurs	7.34	7.13	110.70
Total	8.47	8.22	109.53

CSP / LONGUEUR  $p = 8.0\%$  ;  $F = 1.82$  (PS)  
CSP / RICHESSE  $p = 5.3\%$  ;  $F = 1.99$  (PS)  
CSP / BANALITE  $p = 0.7\%$  ;  $F = 2.80$  (TS)

## Sémiométrie : la mesure des contenus

Les dictionnaires classent les termes utilisés en répertoires thématiques.

Les intensités lexicales indiquent le pourcentage des mots d'une réponse appartenant à chaque thème.

On peut ainsi caractriser chaque catégorie par l'importance que les individus qui la composent accordent à chacun des thèmes

On retrouve les oppositions déjà mises en évidence par les cartes lexicales.

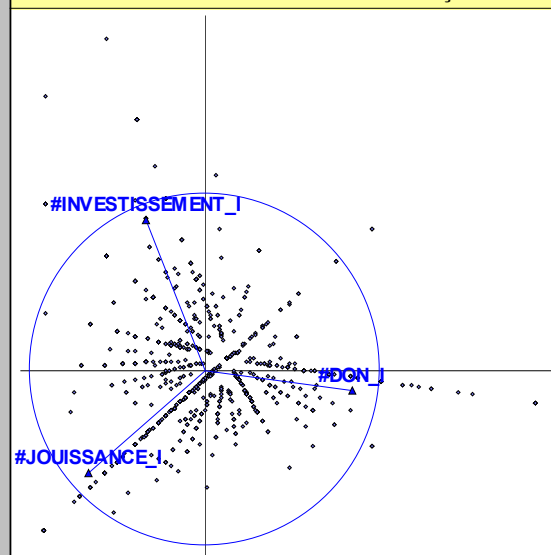
### DON INVESTISSEMENT JOUISSANCE

enfant	placer	maison
donner	investir	acheter
famille	de_côté	voyage
don	épargner	voiture
association	construire	voyager
ami	crédit	partir
partager	épargne	monde
distribuer	rembourser	améliorer
malheureux	fructif	tour
enfants	intérêt	vacance
femme	rapporter	plaisir
handicapé	rente	faire_plaisir
mari		faire_le_tour
parent		pays
		bateau
		étranger
		robe

### Intensités lexicales

poids du thème dans la réponse	
	Moyenne
#JOUISSANCE_I	22.76
#DON_I	10.47
#INVESTISSEMENT_I	6.79
<b>Total</b>	<b>13.34</b>

### Les 3 dimensions du rêve des français



### Présence des thèmes

Nombre de réponses évoquant un des mots du thème

#JOUISSANCE	770	77.5%
#DON	494	49.7%
#INVESTISSEMENT	411	41.4%
<b>Total</b>	<b>993</b>	

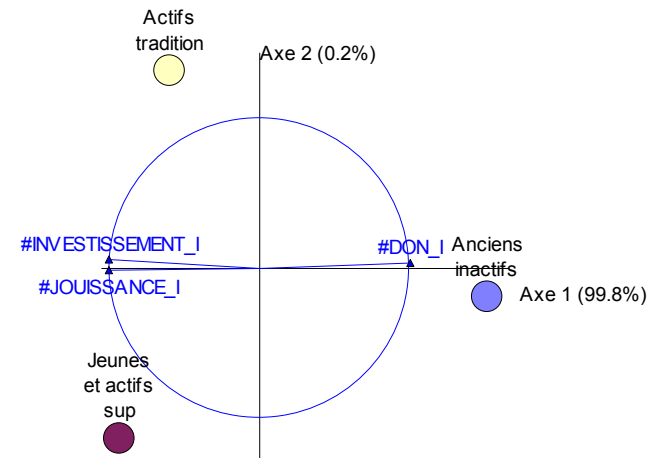
### Les contenus selon les classes de la typologie

	#JOUISSANCE_I	#DON_I	#INVESTISSEMENT_I
Anciens inactifs	14.00	21.48	2.56
Jeunes et actifs sup	26.12	5.92	8.12
Actifs tradition	24.25	8.76	7.68
<b>Total</b>	<b>22.76</b>	<b>10.47</b>	<b>6.79</b>

Typologie / #JOUISSANCE\_I p = <0.1% ; F = 25.30 (TS)

Typologie / #DON\_I p = <0.1% ; F = 79.07 (TS)

Typologie / #INVESTISSEMENT\_I p = <0.1% ; F = 21.16 (TS)



## Analyse de contenu

Le code book préparé dans Sphinx permet « d'industrialiser le travail de lecture ». Cette une approche alternative à l'analyse lexicale. Elle a le mérite de reposer sur une lecture intelligente (mais peut être biaisé) et de reposer sur une réflexion modélisatrice préalable.

Code book		Reste A coder : 785 observations	
arrêter de travailler / une maison sur la Côte d'Azur / aider les gens qui sont malheureux autour de moi / aider les associations / en conserver pour faire ma vie.			
<i>Femme, Cadre.Prof.Intell. Sup., Célibataire, CEP BEPC, 50-64, 20 000 - 100 000</i>			
<b>Thème</b>			
<input checked="" type="checkbox"/> Jouissance	<input type="checkbox"/> Investissement	<input checked="" type="checkbox"/> Don	
<input type="checkbox"/> Maison		<input checked="" type="checkbox"/> Argent	
<input checked="" type="checkbox"/> Vacances		<input type="checkbox"/> Bien	
<input type="checkbox"/> Voyages		<input type="checkbox"/> Temps	
<input type="checkbox"/> Automobile			
<input type="checkbox"/> Bien équipement			
<input type="checkbox"/> Loisir divertissement			
<b>Acteurs</b>		<b>Tonalité</b>	
<input checked="" type="checkbox"/> Moi seul	<input type="checkbox"/> Enfants	<input type="checkbox"/> Neutre	<input type="checkbox"/> Sceptique
<input type="checkbox"/> Proche	<input checked="" type="checkbox"/> Humanité	<input type="checkbox"/> Drole	<input type="checkbox"/> Enthousiaste
<input type="checkbox"/> Conjoint	<input type="checkbox"/> Parents	<input type="checkbox"/> Amis	<input type="checkbox"/> Famille proche
<b>Remarquable</b>			
aider les gens qui sont malheureux autour de moi			

