



HISTOIRE DE VIE, INTRODUCTION A SAS

Rapport d'Enquête & Sondages

Gilliard Tallulah

Enquête & Sondages

L2 MIAASH – Université de Bordeaux

Sommaire

TD1

Les enquêtes de l'INSEE

TD2

Tris à plat sur les données HdV 2003

TD3

Tris croisés avec PROC FREQ

Les formats dans les procédures 1/2

TD4

Les formats dans les procédures 2/2

Les tabulations avancées avec PROC TABULATE

TD5

Association de tables et sélection d'observations

Tableaux disjonctifs complets

Les enquêtes de l'INSEE

Question 1.1 : Missions de l'INSEE

(Questions 1.1 réalisée sur le même ordinateur que Mélody Labarchède)

« L'Institut national de la statistique et des études économiques collecte, produit, analyse et diffuse des informations sur l'économie et la société françaises. »

Les missions sont séparées en six grands thèmes :

- ◆ Collecter et produire : l'INSEE réalise régulièrement des enquêtes statistiques sur différents sujets (ménage, emploi, condition de vie...) et peut utiliser pour y parvenir les dossiers administratifs.
- ◆ Analyser : l'INSEE effectue des études nationales particulièrement au niveau du système productif, il fait des études sur la société française (démographie économique), les dimensions spatiales (échange entre territoires), l'évolution de l'économie à court terme.
- ◆ Diffuser : Les données produites par l'INSEE sont en open data sur le site, certaines données particulières sont disponibles aux entreprises, administrations, chercheurs, étudiants... De plus, l'INSEE propose des produits adaptés aux besoins particuliers des utilisateurs.
- ◆ Coordonner : L'INSEE réalise un regroupement de travaux statistiques venant de services différents afin d'obtenir des données plus complètes. Il s'occupe du CNIS, des diverses formations, de gérer les répertoires de personnes, le développement des outils, la création d'un cadre favorable au développement et s'occupe d'un portail de statistiques pratiques.
- ◆ Enseigner et développer la recherche : Le Genes (Groupe des écoles nationales d'économie et de statistique) forme des statisticiens, économistes et spécialistes du traitement de l'information dans le cadre de la recherche supérieure.
- ◆ Contribuer à l'édification d'un espace statistique international : L'INSEE est en lien ou en partenariat avec des instituts de statistique de l'UE, de l'ONU, Banque Mondiale... Il propose aussi sa coopération aux instituts de recherche des pays en voie de développement.

Question 1.2 : Quelles sont les différents types de bases de données que l'Insee met à la disposition du public ?

L'INSEE propose des données à l'échelle française, locales par territoire selon divers types d'analyses de base de données au fil du temps. Ces données peuvent être séparées en différentes sous catégories de bases de données, qui sont :

- ◆ Comportements démographique (naissances, décès, migrations) : recensement de la population, état civil.
- ◆ Education et formation : recherche et développement
- ◆ Emploi et chômage : emploi en continu, dénombrement des entreprises, emploi et rémunération. L'ESANE propose des statistiques annuelles pour les entreprises
- ◆ Revenu et pauvreté :

◆ Condition de vie

Question 1.3 : Décrire les principaux thèmes couverts par les fichiers détail

Les fichiers détail concernent le recensement de la population et le dénombrement des entreprises et établissements en France, ainsi que des bases de données sur les équipements en France, ou des enquêtes en continu sur l'emploi.

Question 1.4 : Quelles logiciels peut-on utiliser pour analyser ce fichier détail ?

On peut utiliser les fichiers détails qui sont en formats dBase (dbf) et Beyond20/20. Pour le format Beyond20/20, on utilise le browser de Beyond qui permet une lecture plus universelle mais reste limité du fait qu'il est gratuit. Le format dBase est plus couramment utilisé en statistique car il est applicable à beaucoup de logiciels de statistiques tels que SAS, de plus il est compatible avec des fichiers d'analyse statistiques comme SPSS de IBM, c'est pourquoi il est assez complexe et payant.

Question 1.5 : indiquez la nature des activités de chacun des organismes impliqués dans cette enquête.

Les organismes impliqués dans cette enquête sont les suivants :

- ◆ L'Institut national d'études démographiques (L'INED) est impliqué au niveau de l'étude des questions démographiques
- ◆ La Direction de la Recherche, des Etudes, de l'Evaluation et des Statistiques (DREES) rattachée au Ministère des affaires sociales et de la santé est impliquée dans le cadre des analyses sur les populations et les politiques sanitaires et sociales.
- ◆ La Direction de l'Animation, de la Recherche, des Etudes et des Statistiques (DARES) rattachée au Ministère du travail et de l'emploi s'implique dans les statistiques du marché du travail.
- ◆ Le Département des Etudes et de la Prospective (DEP), lié au Ministère de la culture et de la communication s'occupe des statistiques dans le domaine culturel.
- ◆ La Délégation Interministérielle à la Ville (DIV) s'implique dans le financement.
- ◆ L'Institut National de La Santé et de la recherche médicale (INSERM) est chargé de la recherche médicale.

Question 1.6 : Quelles sont les caractéristiques de l'enquête ?

L'enquête comporte 8403 individus de 18 ans et plus, 817 variables réparties dans 4 tables (pour le format dBase) et 808 dans une table (pour le format Beyond)

Question 1.7 : Qu'est-ce qui différencie le format .dbf du format Beyond 20/20 ?

	Format dBase	Format Beyond 20/20
Nombre de variables	817	808
Nombre de tables	4	1
Géré par SAS	oui	non
Géré par le browser Beyond	non	oui
Extension	.dbf	.IVX ou .IVD

Question 1.8 : Quelles sont les principaux thèmes abordés par cette enquête ?

Les principaux thèmes de l'enquête sont les trajectoires de vie et construction identitaire de la population française du XXI^{ème} siècle, regroupant parcours professionnel, familial, social et géographique. Les enquêtés sont aussi interrogés sur leur « identité » c'est-à-dire sur leurs activités, leur situation, leurs implications dans leurs groupes sociaux et la manière dont ils se sentent perçus.

Question 1.9 : Parmi les variables du secteur J2, quels sont les tableaux statistiques qu'il vous semblerait utile de réaliser ?

On pourrait réaliser deux types de tableaux pour étudier différents phénomènes de société :

- ◆ Un tableau croisant la fréquence des rencontres avec des collègues de travail (en colonne) avec la présence de personnes issues du même milieu professionnel parmi les amis (en ligne) Cela pourrait permettre d'étudier les relations amicales en fonction des études effectuées et du milieu socio professionnel. Cela permettrait de comparer les amis issus de ces milieux et ceux issus des milieux de loisirs ou familiaux et voir où se fait t on le plus d'amis et faire un rapport entre ces deux valeurs.
- ◆ Un autre tableau pourrait être effectué avec l'âge des amis (en ligne) et les liens avec ces amis (en colonne. Est un ami d'enfance, vient-il du même milieu professionnel ou social ou partage-t-il les mêmes valeurs, mêmes manière de vivre.
On pourrait observer les caractéristiques qui rapprochent des amis d'âges éloignés et chercher qu'est ce qui peut conduire une personne à s'approcher de personnes ayant un âge assez différent du sien.

Question 1.10 : Dressez une liste de cinq questions prises dans l'ensemble de l'enquête qu'il vous semble intéressant d'analyser plus tard.

-L'histoire familiale d'une personne est faite de mises en couple, de séparations, de naissances d'enfants et de petits-enfants... Aujourd'hui, personnellement, diriez-vous que vous êtes avant tout:

-Actuellement, diriez-vous que dans votre vie en général, votre travail :

-Quelle(s) langue(s) vous parlaient d'habitude vos parents (y compris langues régionales ou locales) quand vous étiez enfant (vers l'âge de 5 ans)?

-Au cours des 12 derniers mois, à quelle fréquence avez-vous rencontré, chez vous ou chez eux ou lors de sorties communes des personnes de votre propre famille proche ou éloignée (parents, enfants, frères et soeurs, oncles et tantes, cousins, grands-parents...) ?

-Faites-vous partie d'une association à but humanitaire, ou de défense de l'environnement ou une autre association de défense d'intérêts, etc. ?

Tris à plat sur les données HdV 2003

Editeur

```
x cd "E:\EetS";
proc dbf db5= hdv1 out=hdv1;

%macro importDBF(tableDBF);
    proc dbf db5=&tableDBF out=&tableDBF;
    run;
    proc contents data=&tableDBF varnum short;
    run;
%mend;

%importDBF(hdv1);
%importDBF(hdv2);
%importDBF(hdv3);
%importDBF(hdv4);
```

Journal

```
137 x cd "E:\EetS"
137!      ;
138 proc dbf db5= hdv1 out=hdv1;
139 %macro importDBF(tableDBF);
140     proc dbf db5=&tableDBF out=&tableDBF;
141     run;
142     proc contents data=&tableDBF varnum short;
143     run;
144 %mend;
145 %importDBF(hdv1);
```

NOTE: 8403 observations written to the output SAS data set.

NOTE: PROCEDURE DBF used (Total process time):

real time	0.03 secondes
cpu time	0.03 secondes

NOTE: 8403 observations written to the output SAS data set.

NOTE: PROCEDURE DBF used (Total process time):

real time	0.03 secondes
cpu time	0.03 secondes

NOTE: PROCEDURE CONTENTS used (Total process time):

real time	0.08 secondes
cpu time	0.01 secondes

```
146 %importDBF(hdv2);
```

NOTE: 8403 observations written to the output SAS data set.

NOTE: PROCEDURE DBF used (Total process time):

real time	0.03 secondes
cpu time	0.03 secondes

NOTE: PROCEDURE CONTENTS used (Total process time):

real time	0.04 secondes
cpu time	0.01 secondes

```
147 %importDBF(hdv3);
```

NOTE: 8403 observations written to the output SAS data set.

NOTE: PROCEDURE DBF used (Total process time):

real time	0.03 secondes
cpu time	0.03 secondes

NOTE: PROCEDURE CONTENTS used (Total process time):

real time	0.04 secondes
cpu time	0.00 secondes

148 %importDBF(hdv4);

NOTE: 8403 observations written to the output SAS data set.

NOTE: PROCEDURE DBF used (Total process time):

real time	0.02 secondes
cpu time	0.01 secondes

NOTE: PROCEDURE CONTENTS used (Total process time):

real time	0.05 secondes
cpu time	0.00 secondes

Editeur

```
libname mabib "E:\EetS";
data sasuser.hdv1;
    set hdv1;
run;
data sasuser.hdv2;
    set hdv2;
run;
data sasuser.hdv3;
    set hdv3;
run;
data mabib.hdv4;
    set hdv4;
run;

libname mabib "E:\EetS";
data mabib.hdv1;
    set hdv1;
run;
data mabib.hdv2;
    set hdv2;
run;
data mabib.hdv3;
    set hdv3;
run;
data mabib.hdv4;
    set hdv4;
run;

proc freq data=mabib.hdv1;
    tables AGEE/out=tri;
    weight POIDSF;
run;
```

Journal

```
237 libname mabib "E:\EetS";
NOTE: Libref MABIB was successfully assigned as follows:
      Engine:      V9
      Physical Name: E:\EetS
238 data sasuser.hdv1;
239     set hdv1;
240     run;
```


NOTE: There were 8403 observations read from the data set WORK.HDV1.
NOTE: The data set SASUSER.HDV1 has 8403 observations and 205 variables.
NOTE: L'étape DATA used (Total process time):
 real time 0.38 secondes
 cpu time 0.01 secondes

```
241 data sasuser.hdv2;  
242     set hdv2;  
243     run;
```

NOTE: There were 8403 observations read from the data set WORK.HDV2.
NOTE: The data set SASUSER.HDV2 has 8403 observations and 203 variables.
NOTE: L'étape DATA used (Total process time):
 real time 0.31 secondes
 cpu time 0.03 secondes

```
244 data sasuser.hdv3;  
245     set hdv3;  
246     run;
```

NOTE: There were 8403 observations read from the data set WORK.HDV3.
NOTE: The data set SASUSER.HDV3 has 8403 observations and 206 variables.
NOTE: L'étape DATA used (Total process time):
 real time 0.29 secondes
 cpu time 0.04 secondes

```
247 data mabib.hdv4;  
248     set hdv4;  
249     run;
```

NOTE: There were 8403 observations read from the data set WORK.HDV4.
NOTE: The data set MABIB.HDV4 has 8403 observations and 203 variables.
NOTE: L'étape DATA used (Total process time):
 real time 1.41 secondes
 cpu time 0.00 secondes

```
250 libname mabib "E:\EetS";  
NOTE: Libref MABIB was successfully assigned as follows:  
    Engine:      V9  
    Physical Name: E:\EetS  
251 data mabib.hdv1;  
252     set hdv1;  
253     run;
```

NOTE: There were 8403 observations read from the data set WORK.HDV1.
NOTE: The data set MABIB.HDV1 has 8403 observations and 205 variables.
NOTE: L'étape DATA used (Total process time):
 real time 1.41 secondes
 cpu time 0.00 secondes

```
254 data mabib.hdv2;  
255     set hdv2;  
256     run;
```

NOTE: There were 8403 observations read from the data set WORK.HDV2.
NOTE: The data set MABIB.HDV2 has 8403 observations and 203 variables.
NOTE: L'étape DATA used (Total process time):
 real time 1.16 secondes
 cpu time 0.01 secondes

```
257 data mabib.hdv3;  
258     set hdv3;  
259     run;
```

NOTE: There were 8403 observations read from the data set WORK.HDV3.
NOTE: The data set MABIB.HDV3 has 8403 observations and 206 variables.
NOTE: L'étape DATA used (Total process time):
 real time 1.33 secondes
 cpu time 0.03 secondes

```
260 data mabib.hdv4;
```

```

261 set hdv4;
262 run;

```

NOTE: There were 8403 observations read from the data set WORK.HDV4.
NOTE: The data set MABIB.HDV4 has 8403 observations and 203 variables.
NOTE: L'étape DATA used (Total process time):
real time 0.88 secondes
cpu time 0.03 secondes

```

263 proc freq data=mabib.hdv1;
264 tables AGEE/out=tri;
265 weight POIDSF;
266 run;

```

NOTE: There were 8403 observations read from the data set MABIB.HDV1.
NOTE: The data set WORK.TRI has 80 observations and 3 variables.
NOTE: PROCEDURE FREQ used (Total process time):
real time 0.15 secondes
cpu time 0.01 secondes

Le Système SAS

Procédure FREQ

AGEE	Fréquence	Pourcentage	Fréquence cumulée	Pctage. cumulé
17	1	0.01	1	0.01
18	95	1.13	96	1.14
19	106	1.26	202	2.40
20	111	1.32	313	3.72
21	98	1.17	411	4.89
22	127	1.51	538	6.40
23	136	1.62	674	8.02
24	93	1.11	767	9.13
25	90	1.07	857	10.20
26	101	1.20	958	11.40
27	97	1.15	1055	12.56
28	108	1.29	1163	13.84
29	143	1.70	1306	15.54
30	145	1.73	1451	17.27
31	139	1.65	1590	18.92
32	140	1.67	1730	20.59
33	132	1.57	1862	22.16
34	136	1.62	1998	23.78
35	161	1.92	2159	25.69
36	154	1.83	2313	27.53
37	159	1.89	2472	29.42
38	131	1.56	2603	30.98
39	188	2.24	2791	33.21
40	184	2.19	2975	35.40
41	162	1.93	3137	37.33
42	184	2.19	3321	39.52

43	150	1.79	3471	41.31
44	159	1.89	3630	43.20
45	175	2.08	3805	45.28
46	167	1.99	3972	47.27
47	190	2.26	4162	49.53
48	221	2.63	4383	52.16
49	178	2.12	4561	54.28
50	184	2.19	4745	56.47
51	166	1.98	4911	58.44
52	179	2.13	5090	60.57
53	180	2.14	5270	62.72
54	201	2.39	5471	65.11
55	171	2.03	5642	67.14
56	191	2.27	5833	69.42
57	161	1.92	5994	71.33
58	139	1.65	6133	72.99
59	149	1.77	6282	74.76
60	112	1.33	6394	76.09
61	103	1.23	6497	77.32
62	81	0.96	6578	78.28
63	105	1.25	6683	79.53
64	113	1.34	6796	80.88
65	103	1.23	6899	82.10
66	92	1.09	6991	83.20
67	120	1.43	7111	84.62
68	91	1.08	7202	85.71
69	100	1.19	7302	86.90
70	95	1.13	7397	88.03

71	88	1.05	7485	89.08
72	94	1.12	7579	90.19
73	100	1.19	7679	91.38
74	76	0.90	7755	92.29
75	81	0.96	7836	93.25
76	69	0.82	7905	94.07
77	75	0.89	7980	94.97
78	79	0.94	8059	95.91
79	57	0.68	8116	96.58
80	46	0.55	8162	97.13
81	52	0.62	8214	97.75
82	44	0.52	8258	98.27
83	29	0.35	8287	98.62
84	17	0.20	8304	98.82
85	14	0.17	8318	98.99
86	10	0.12	8328	99.11
87	17	0.20	8345	99.31
88	19	0.23	8364	99.54
89	10	0.12	8374	99.65
90	7	0.08	8381	99.74
91	8	0.10	8389	99.83
92	4	0.05	8393	99.88
93	4	0.05	8397	99.93
94	1	0.01	8398	99.94
96	4	0.05	8402	99.99
97	1	0.01	8403	100.00

Question 2.7 : Examinez le contenu de la fenêtre « Result Viewer. » Quelle est l'utilité de cette liste ? Décrivez chacun de ses éléments

Le logiciel SAS génère un tri à plat sur un ensemble de données en fonction d'un des champs de la table. Cette liste indique le nombre d'individus ayant un âge donné et la proportion d'individus ayant cet âge sur l'ensemble de l'échantillon.

Question 2.8 : De combien d'étapes se compose ce programme ? Quelle est la fonction de chacune de ces étapes ?

Ce programme est composé de trois étapes `proc (freq, contents et print)`. `freq` nous permet d'effectuer un tri à plat sur le champ `AGEE`, `contents` affiche des informations à propos de la table et `print` affiche cette table (tableau « le système SAS » ci-dessus).

Question 2.9 : Décrivez les résultats des procédures `CONTENTS` et `PRINT`.

Le Système SAS			
Procédure CONTENTS			
Variables par ordre de création			
AGEE COUNT PERCENT			

Le Système SAS			
Obs.	AGEE	COUNT	PERCENT
1	17	1	0.01190
2	18	95	1.13055
3	19	106	1.26145
4	20	111	1.32096
5	21	98	1.16625

La procédure `contents` permet d'obtenir des informations à propos d'une table donnée, en ce qui concerne le nom des champs de la table.

La procédure `print` affiche les éléments de la table. Le paramètre `obs` correspond au nombre d'observations à afficher voulues.

Question 2.10 : La répartition par âge des enquêtés est-elle conforme avec la population effectivement ciblée par l'enquête ? Pourquoi ?

La répartition par âge des enquêtés n'est pas conforme avec la population ciblée par l'enquête car l'échantillon sélectionné n'est pas pondéré et n'est donc pas réellement représentatif de la population ciblée.

Question 2.11 : A l'aide de la documentation de l'enquête, présentez les critères ayant permis de calculer les valeurs de `POIDSF`.

Dans la documentation, on apprend que `POIDSF` dépend de la pondération dû au tirage de l'échantillon et de la surreprésentation de certaines parties de la population. On donne donc un poids à l'individu qui a répondu au questionnaire en fonction de son âge, son sexe, sa classe sociale, et d'autres variables. Cet individu ainsi standardisé représentera alors l'ensemble de la population française ayant des caractéristiques similaires en tant qu' « individu lambda ».

Question 2.12 : Dans la documentation de la procédure `freq`, recherchez comment on doit être utilisé l'instruction `WEIGHT` afin de pondérer les fréquences du tri à plat par la variable `POIDSF`

À l'aide de la documentation, on sait que `POIDSF`

TD3

Tris croisés avec PROC FREQ

Question 3.1 : A l'aide de la documentation de l'enquête, expliquez comment sont construites ces deux variables, listez pour chacune d'elles les modalités de réponse. Identifiez dans quelle table se trouvent ces deux variables (hdv1, hdv2, hdv3 ou hdv4 ?)

LT1AMI3 = ait gardé contact ou non avec ses amis d'enfance. Elle prend la valeur '1' si c'est le cas, '0' sinon.

QUALIFE = position professionnel de l'emploi de l'enquêté.

Cette variable a pour modalités :

1	Manœuvre ou ouvrier spécialisé
2	Ouvrier qualifié ou hautement qualifié, technicien(ne) d'atelier
3	Technicien(ne) (non cadre)
4	Agent de maîtrise, maîtrise administrative ou commerciale, VRP (non cadre), personnel de catégorie B de la fonction publique
5	Ingénieur ou cadre, personnel de catégorie A de la fonction publique
6	Employé de bureau, employé de commerce, personnel de service, personnel de catégorie C ou D de la fonction publique
7	Autre cas

Ce sont deux variables de la table hdv4.

Editeur

```
libname mabib "E:\EetS";

proc freq data=mabib.hdv4;
    tables QUALIFE * LT1AMI3;
run;
```

Journal

```
27 libname mabib "E:\EetS";
NOTE: Libref MABIB was successfully assigned as follows:
      Engine:      V9
      Physical Name: E:\EetS
28 proc freq data=mabib.hdv4;
29     tables QUALIFE * LT1AMI3;
30 run;

NOTE: Writing HTML Body file: sashtml.htm
NOTE: There were 8403 observations read from the data set MABIB.HDV4.
NOTE: PROCEDURE FREQ used (Total process time):
      real time      1.22 secondes
      cpu time       0.14 secondes
```

Procédure FREQ

Fréquence	Table de QUALIFE par LT1AMI3			
Pourcentage				
Pctage en ligne				
Pctage en col.				
	QUALIFE	LT1AMI3		Total
		0	1	
1		591	201	792
		8.56	2.91	11.47
		74.62	25.38	
		12.08	10.00	
2		857	304	1161
		12.41	4.40	16.81
		73.82	26.18	
		17.51	15.12	
3		273	128	401
		3.95	1.85	5.81
		68.08	31.92	
		5.58	6.36	

4	504	231	735
	7.30	3.35	10.64
	68.57	31.43	
	10.30	11.49	
5	702	350	1052
	10.17	5.07	15.24
	66.73	33.27	
	14.34	17.40	
6	1812	740	2552
	26.24	10.72	36.96
	71.00	29.00	
	37.02	36.80	
7	155	57	212
	2.24	0.83	3.07
	73.11	26.89	
	3.17	2.83	
Total	4894	2011	6905
	70.88	29.12	100.00
Valeur(s) manquante(s) = 1498			

Tableau modifié pour l'ensemble de la population :

Editeur

```
libname mabib "E:\EetS";

proc freq data=mabib.hdv4;
    tables QUALIFE * LT1AMI3;
    weight POIDSF;
run;
```

Journal

```
31 libname mabib "E:\EetS";
NOTE: Libref MABIB was successfully assigned as follows:
      Engine:      V9
      Physical Name: E:\EetS
32 proc freq data=mabib.hdv4;
33   tables QUALIFE * LT1AMI3;
34   weight POIDSF;
35 run;

NOTE: There were 8403 observations read from the data set MABIB.HDV4.
NOTE: PROCEDURE FREQ used (Total process time):
      real time      0.19 secondes
      cpu time       0.00 secondes
```

Procédure FREQ

Fréquence	Table de QUALIFE par LT1AMI3			
Pourcentage	QUALIFE	LT1AMI3		
Pctage en ligne		0	1	Total
Pctage en col.				
1		2657573	1113340	3770913
		7.59	3.18	10.77
		70.48	29.52	
		11.23	9.82	
2		4060119	1729568	5789686
		11.60	4.94	16.54
		70.13	29.87	
		17.15	15.26	
3		1438040	719996	2158035
		4.11	2.06	6.16
		66.64	33.36	
		6.07	6.35	
4		2405224	1363786	3769011
		6.87	3.90	10.77
		63.82	36.18	
		10.16	12.03	
5		3632478	1928439	5560917
		10.38	5.51	15.88
		65.32	34.68	
		15.34	17.02	
6		8852408	4101349	1.295E7
		25.29	11.72	37.00
		68.34	31.66	
		37.39	36.19	
7		629199	376030	1005229
		1.80	1.07	2.87
		62.59	37.41	
		2.66	3.32	
Total		2.368E7	1.133E7	3.501E7
		67.63	32.37	100.00
Valeur(s) manquante(s) = 10028450.752				

Question 3.2 : Identifiez en clair chacune des cases du tableau en indiquant son numéro de ligne et de colonne (quelle est la signification des modalités croisées)

[LT1AMI3, QUALIFE]

[0,1] : Personne de catégorie 1 qui n'ont pas parmi leurs amis des amis d'enfance.
[1,1] : Personne de catégorie 1 qui ont parmi leurs amis des amis d'enfance.
[0,2] : Personne de catégorie 2 qui n'ont pas parmi leurs amis des amis d'enfance.
[1,2] : Personne de catégorie 2 qui ont parmi leurs amis des amis d'enfance.
[0,3] : Personne de catégorie 3 qui n'ont pas parmi leurs amis des amis d'enfance.
[1,3] : Personne de catégorie 3 qui ont parmi leurs amis des amis d'enfance.
[0,4] : Personne de catégorie 4 qui n'ont pas parmi leurs amis des amis d'enfance.
[1,4] : Personne de catégorie 4 qui ont parmi leurs amis des amis d'enfance.
[0,5] : Personne de catégorie 5 qui n'ont pas parmi leurs amis des amis d'enfance.
[1,5] : Personne de catégorie 5 qui ont parmi leurs amis des amis d'enfance.
[0,6] : Personne de catégorie 6 qui n'ont pas parmi leurs amis des amis d'enfance.
[1,6] : Personne de catégorie 6 qui ont parmi leurs amis des amis d'enfance.
[0,7] : Personne de catégorie 7 qui n'ont pas parmi leurs amis des amis d'enfance.
[1,7] : Personne de catégorie 7 qui ont parmi leurs amis des amis d'enfance.

Question 3.3 : Chaque case contient quatre nombre. Expliquez ce que sont ces nombres en vous appuyant sur la case 4/1.

Le 1er nombre est le nombre de personnes selon la catégorie [colonne] et la catégorie [ligne]

Le 2eme nombre est le pourcentage de personnes dans la catégorie [colonne] et la catégorie [ligne] par rapport aux nombres total de personnes.

Le 3eme nombre est le pourcentage de personnes dans la catégorie [colonne] et la catégorie [ligne] par rapport aux nombre de personnes de la catégorie [ligne]

Le 3eme nombre est le pourcentage de personnes dans la catégorie [colonne] et la catégorie [ligne] par rapport aux nombre de personnes de la catégorie [colonne]

Question 3.4 : Que sont les nombres de la colonne et de la ligne Total ?

Le 1er nombre est le nombre total de personnes dans la catégorie [colonne] ou [ligne] qui correspond.

Le 2eme nombre est le pourcentage de personnes dans la catégorie [colonne] ou [ligne] qui correspond par rapport aux nombre total de personnes.

Question 3.5 : Comment faut-il comprendre l'expression « Fréquence manquante ? » ?

L'expression « Fréquence manquante ? » correspond aux enquêtés qui n'ont pas répondu ou qui sont en dehors des catégories proposées.

Question 3.6 : Que pouvez-vous dire sur ce tableau en rapport avec l'hypothèse initiale ?

Il semblerait que 70% environ des personnes des classes les plus basses n'ont pas parmi leurs amis des amis d'enfance, contre environ 65% pour les classes les plus élevées. Cet écart est minime et surement pas très significatif mais il semble contredire l'hypothèse de départ.

Statistiques pour la table de QUALIFE par LT1AMI3

Statistique	DDL	Valeur	Prob
Khi-2	6	84610.4297	<.0001
Test du rapport de vraisemblance	6	84166.3818	<.0001
Khi-2 de Mantel-Haenszel	1	17464.6617	<.0001
Coefficient Phi		0.0492	
Coefficient de contingence		0.0491	
V de Cramer		0.0492	

Taille réelle de l'échantillon = 35007549.248
Valeur(s) manquante(s) = 10028450.752

WARNING: 22% des données sont manquantes.

Question 3.7 : Ce rejet de l'hypothèse nulle n'est pas contradictoire avec l'assertion initiale de votre ami d'enfance. Pourquoi ? Et pourtant, il a tort. Pourquoi ?

Un Khi-deux aussi faible (inférieur à 0.01) montre qu'il y a sûrement une corrélation importante entre la catégorie de métier et le fait de garder des amis d'enfance ou non. En revanche, ce sont, les personnes appartenant à une classe aisée ont plus tendance à garder leurs amis d'enfance.

Les formats dans les procédures 1/2

Editeur

```
proc format library=mabib;
  value $FMT_RSATIS
    1="aimé + tard"
    2="aimé + tôt"
    3="bon moment"
    4=" ";
  value $FMT_PREGRE
    1="souvent"
    2="parfois"
    3="rarement"
    4="jamais";
run;
```

Journal

```
77 proc format library=mabib;
78   value $FMT_RSATIS
79     1="aimé + tard"
80     2="aimé + tôt"
81     3="bon moment"
82     4=" ";
NOTE: Format $FMT_RSATIS is already on the library.
NOTE: Format $FMT_RSATIS has been written to MABIB.FORMATS.
83   value $FMT_PREGRE
84     1="souvent"
85     2="parfois"
86     3="rarement"
87     4="jamais";
NOTE: Format $FMT_PREGRE is already on the library.
```


NOTE: Format \$FMT_PREGRE has been written to MABIB.FORMATS.
88 run;

NOTE: PROCEDURE FORMAT used (Total process time):
real time 0.02 secondes
cpu time 0.00 secondes

Question 3.8 : Pourquoi les noms `FMT_RSATIS` et `FMT_PREGRE` sont-ils précédés du caractère \$?

Le caractère \$ devant un nom de format signifie que c'est un format « Chaîne de caractère » (valeurs alphanumériques) pour les différencier des valeurs numériques au niveau de l'encodage.

Le Système SAS

FORMAT NAME: \$FMT_PREGRE LENGTH: 8			
MIN LENGTH:	1	MAX LENGTH: 40	DEFAULT LENGTH: 8 FUZZ: 0
START	END	LABEL (VER. 9.3 23OCT2014:14:45:26)	
1	1	souvent	
2	2	parfois	
3	3	rarement	
4	4	jamais	

Le Système SAS

FORMAT NAME: \$FMT_RSATIS LENGTH: 11			
MIN LENGTH:	1	MAX LENGTH: 40	DEFAULT LENGTH: 11 FUZZ: 0
START	END	LABEL (VER. 9.3 23OCT2014:14:45:26)	
1	1	aimé + tard	
2	2	aimé + tôt	
3	3	bon moment	
4	4		

Question 3.9 : Que contient la bibliothèque `Work` ? D'où provient son contenu ? Pourquoi ces éléments se trouvent-ils dans la bibliothèque `Work` ? Que deviennent ces deux éléments en fin de session SAS ?

La bibliothèque `Work` contient ces deux tables de format issues des procédures `proc format` qui sont exécutées ci-contre (issues de l'éditeur un peu plus haut). La bibliothèque `Work` est la bibliothèque par défaut de SAS ; n'ayant pas spécifié de bibliothèque pour enregistrer nos tables, SAS les enregistre dans `Work`. Ces deux tables seront supprimées à la fin de la session SAS ; pour les enregistrer, il faut spécifier une bibliothèque où les enregistrer (de type `Sasuser` ou `Mabib`).

Question 3.10 : Que contient de nouveau la bibliothèque `Mabib` ? D'où provient son contenu ? Décrivez-le.

`Mabib` contient maintenant les formats `$FMT_RSATIS` et `$FMT_PREGRE`, générés par l'ajout du paramètre `library=mabib` dans la procédure `format`. Ces formats sont des tables qui contiennent une clé et un alias. Lorsqu'une table s'affiche sous ce format, les occurrences des clés présentes sont remplacées par leurs alias.

TD4

Les formats dans les procédures 2/2

Question 4.1 : Rédigez d'un point de vue sociologique, un commentaire de ce tableau d'une vingtaine de ligne

Editeur

```
libname mabib "E:\EetS";

options fmtsearch=(mabib);
proc freq data=mabib.hdv4;
    tables RREGRE*RSATIS;
    weight poidsf;
    title 'LES RETRAITES';
    title2 'Moment du départ en retraite / Regret du travail';
    label RREGRE='Regret du travail ?'
          RSATIS='Satisfaction départ retraite ?';
    format RREGRE $FMT_PREGRE.
          RSATIS $FMT_RSATIS.;
run;
```

Journal

```
1  libname mabib "E:\EetS";
NOTE: Libref MABIB was successfully assigned as follows:
    Engine:      V9
    Physical Name: E:\EetS
2
3  options fmtsearch=(mabib);
4  proc freq data=mabib.hdv4;
5      tables RREGRE*RSATIS;
6      weight poidsf;
7      title 'LES RETRAITES';
8      title2 'Moment du départ en retraite / Regret du travail';
9      label RREGRE='Regret du travail ?'
10         RSATIS='Satisfaction départ retraite ?';
11      format RREGRE $FMT_PREGRE.
12         RSATIS $FMT_RSATIS.;
13 run;
```

NOTE: Writing HTML Body file: sashtml.htm

NOTE: There were 8403 observations read from the data set MABIB.HDV4.

NOTE: PROCEDURE FREQ used (Total process time):

real time	0.25 secondes
cpu time	0.15 secondes

LES RETRAITES

Moment du départ en retraite / Regret du travail

Procédure FREQ

Fréquence	Table de RREGRE par RSATIS				
Pourcentage	RREGRE(Regret du travail ?)	RSATIS(Satisfaction départ retraite ?)			
Pctage en ligne Pctage en col.		aimé + tard	aimé + tôt	bon moment	Total
souvent		512825	67890.6	442718	1023434
		5.13	0.68	4.43	10.25
		50.11	6.63	43.26	
		25.63	6.88	6.33	
parfois		543468	163725	1538727	2245920
		5.44	1.64	15.41	22.49
		24.20	7.29	68.51	
		27.16	16.58	21.98	
rarement		231174	62354.6	885640	1179168
		2.31	0.62	8.87	11.81
		19.60	5.29	75.11	
		11.55	6.31	12.65	
jamais		713320	693489	4131995	5538804
		7.14	6.94	41.37	55.46
		12.88	12.52	74.60	
		35.65	70.23	59.04	
Total		2000787	987459	6999081	9987326
		20.03	9.89	70.08	100.00
Valeur(s) manquante(s) = 35048673.747					

Ce tableau a pour sujet le départ à la retraite. Il confronte le regret du travail d'une personne à la satisfaction qu'il a de son moment de départ. Il concerne les retraités uniquement, d'où les trente-cinq millions de valeurs manquantes (en bas)

Ce tableau compare la satisfaction du départ à la retraite ; Si l'enquêté aurait voulu partir plus tard, partir plus tôt ou s'il trouve que c'était le bon moment (en ligne) face au regret du travail de la personne : Souvent/Parfois/Rarement/Jamais

On constate qu'une grande partie, 70% de la population en France, pense avoir quitté son travail au bon moment contre 20% pensent qu'ils auraient préféré quitter leur emploi plus tardivement et 10% auraient préféré le faire plus tôt. En ce qui concerne le regret du travail, 67% des personnes à la retraite ne regrettent jamais ou rarement leur travail face à seulement 33% qui regrettent souvent ou parfois.

On remarque que les personnes qui regrettent souvent leur travail auraient aimé une retraite plus tardive (50%) ou la trouvaient au bon moment (43%), ce qui semble plutôt cohérent.

Ceux qui regrettent parfois leur travail, vont majoritairement trouver le départ à la retraite au bon moment (69%) et de même pour ceux qui regrettent rarement (75%).

Nous pourrions nous attendre à ce que les gens qui ne regrettent jamais leur travail soient plus enclins à préférer partir en retraite plus tôt. Pourtant, parmi les personnes qui ne regrettent jamais leur travail, la majorité pensent être partis en retraite au bon moment.

Pour conclure, on peut dire que pour les gens, la retraite arrive pour la plupart au bon moment (70%) et que le travail n'est plutôt jamais regretté (55%). S'il l'est, les gens ont tendance tout de même à penser que la retraite arrivait au bon moment sauf quand ils regrettent celui-ci souvent leur travail.

Question 4.2 : Retournez à la fiche d'ED n°2. Réalisez un tri à plat sur l'âge de la population recodé en groupe d'âges décennaux à l'aide la procédure `format` et `freq`. Notez que la définition d'intervalles dans la procédure `format` se fait de la façon suivante (: 10-19='10-19 ans')

Editeur

```
libname mabib "E:\EetS";

options fmtsearch=(mabib);

proc format library=mabib;
    value FMT_AGE
        0-9='0-9 ans'
        10-19='10-19 ans'
        20-29='20-29 ans'
        30-39='30-39 ans'
        40-49='40-49 ans'
        50-59='50-59 ans'
        60-69='60-69 ans'
        70-79='70-79 ans'
        80-89='80-89 ans'
        90-99='90-99 ans';
run;

proc freq data=mabib.hdv1;
    tables AGEE;
    weight poidsf;
    title "Tri à plat sur l'âge";
    title2 "Par décennies";
    label AGEE="Age de l'enquêté";
    format AGEE FMT_AGE.;
run;
```

Journal

```
114 libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";
NOTE: Libref MABIB was successfully assigned as follows:
      Engine:          V9
      Physical Name: E:\S3\EetS\EetS-SASjournalEnregistrements
115
116 options fmtsearch=(mabib);
117
118 proc format library=mabib;
119     value FMT_AGE
120         0-9='0-9 ans'
121         10-19='10-19 ans'
122         20-29='20-29 ans'
123         30-39='30-39 ans'
124         40-49='40-49 ans'
125         50-59='50-59 ans'
126         60-69='60-69 ans'
127         70-79='70-79 ans'
128         80-89='80-89 ans'
129         90-99='90-99 ans';
NOTE: Format FMT_AGE is already on the library.
NOTE: Format FMT_AGE has been written to MABIB.FORMATS.
130 run;

NOTE: PROCEDURE FORMAT used (Total process time):
      real time        0.16 secondes
      cpu time         0.00 secondes

131
132 proc freq data=mabib.hdv1;
133     tables AGEE;
134     weight poidsf;
135     title "Tri à plat sur l'âge";
```

```

136 title2 "Par décennies";
137 label AGEE="Age de l'enquête";
138 format AGEE FMT_AGE.;
139 run;

```

NOTE: There were 8403 observations read from the data set MABIB.HDV1.

NOTE: PROCEDURE FREQ used (Total process time):

```

real time    0.06 secondes
cpu time     0.00 secondes

```

Tri à plat sur l'âge Par décennies

Procédure FREQ

Age de l'enquête				
AGEE	Fréquence	Pourcentage	Fréquence cumulée	Pctage. cumulé
10-19 ans	1605897	3.57	1605897	3.57
20-29 ans	8241429	18.30	9847326	21.87
30-39 ans	8639908	19.18	18487234	41.05
40-49 ans	8108796	18.01	26596029	59.06
50-59 ans	7223801	16.04	33819830	75.10
60-69 ans	5186536	11.52	39006367	86.61
70-79 ans	4414416	9.80	43420782	96.41
80-89 ans	1395075	3.10	44815857	99.51
90-99 ans	220143.3	0.49	45036000	100.00

Les tabulations avancées avec PROC TABULATE

Question 4.3 : En vous aidant de la documentation SAS, indiquez ce que fait chaque ligne du programme.

Editeur

```
libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";

options nodate pageno=1 linesize=140 pagesize=80;

proc format;
  value $FMT_GNIV
    1="région france"
    2="france"
    3="europe"
    4="autre pays ou continent"
    5="autre pays ou continent"
    8="refus de répondre"
    9="ne sait pas";
  value $FMT_FLMAT
    1="seulement le français"
    2="seulement une autre langue"
    3="français et une autre langue"
    4="deux autres langues";
run;

proc tabulate data=mabib.hdv2 format=10.0;
  class GNIV FLMAT;
  tables GNIV(all="Total origine ressentie"),
         FLMAT(all="total langue");
  freq poidsf;
  format GNIV $FMT_GNIV.
         FLMAT $FMT_FLMAT.;
  label gniv="Origine ressentie..."
        flmat="Langue parlée dans l'enfance";
run;
```

Journal

```
140 libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";
NOTE: Libref MABIB was successfully assigned as follows:
      Engine:      V9
      Physical Name: E:\S3\EetS\EetS-SASjournalEnregistrements
141
142 options nodate pageno=1 linesize=140 pagesize=80;
143
144 proc format;
145   value $FMT_GNIV
146     1="région france"
147     2="france"
148     3="europe"
149     4="autre pays ou continent"
150     5="autre pays ou continent"
151     8="refus de répondre"
152     9="ne sait pas";
NOTE: Format $FMT_GNIV has been output.
153   value $FMT_FLMAT
154     1="seulement le français"
155     2="seulement une autre langue"
156     3="français et une autre langue"
```

```

157      4="deux autres langues";
NOTE: Format $FMT_FLMAT has been output.
158 run;

```

NOTE: PROCEDURE FORMAT used (Total process time):

real time	0.03 secondes
cpu time	0.01 secondes

```

159
160
161 proc tabulate data=mabib.hdv2 format=10.0;
162   class GNIV FLMAT;
163   tables GNIV(all="Total origine ressentie"),
164          FLMAT(all="total langue");
165   freq poids;
166   format GNIV $FMT_GNIV.
167          FLMAT $FMT_FLMAT.;
168   label gniv="Origine ressentie..."
169          flmat="Langue parlée dans l'enfance";
170 run;

```

NOTE: There were 8403 observations read from the data set MABIB.HDV2.

NOTE: PROCEDURE TABULATE used (Total process time):

real time	0.76 secondes
cpu time	0.06 secondes

Tri à plat sur l'âge Par décennies

	Langue parlée dans l'enfance				total langue
	seulement le français	seulement une autre langue	français et une autre langue	deux autres langues	
	N	N	N	N	N
Origine ressentie...					
région france	8557861	1314468	1441914	44112	11358355
france	21710516	2391933	2670288	111421	26884158
europe	3054890	670952	551707	70695	4348244
autre pays ou continent	252389	1329752	428647	82651	2093439
refus de répondre	8393	.	3556	.	11949
ne sait pas	155535	92943	63635	23559	335672
Total origine ressentie	33739584	5800048	5159747	332438	45031817

La première ligne permet d'assigner la librairie usuelle `mabib`. L'instruction `option` permet de paramétrer les options d'affichage du résultat dans le `Result Viewer`, ces options seront utilisées par la procédure `tabulate` pour dessiner le tableau désiré.

Des formats sont ensuite créés dans la bibliothèque `Work`, ils serviront à indiquer les champs utilisés dans la procédure `tabulate`.

La procédure `tabulate` s'applique sur la table `mabib.hdv2`.

Le paramètre `class` concerne les champs sur lesquelles la procédure `tabulate` s'applique dans la table `hdv2`. Le paramètre `tables` indique la manière dont la procédure traite ces champs. (tableau croisé). `freq` nous permet alors de pondérer les résultats suivant la variable `poids` pour appliquer les données à la population. `format` nous permet d'appliquer les formats aux variables `gniv` et `flmat`. `label` donne un intitulé à `gniv` et `flmat`.

Question 4.4 : En vous aidant de la documentation SAS, indiquez ce que fait chaque ligne de l'étape

PROC TABULATE

Editeur

```
libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";

options fmtsearch= (mabib);
options nodate pageno=1 linesize=140 pagesize=80;

proc format;
  value $FMT_GNIV
    1="région france"
    2="france"
    3="europe"
    4="autre pays ou continent"
    5="autre pays ou continent"
    8="refus de répondre"
    9="ne sait pas";
  value $FMT_FLMAT
    1="seulement le français"
    2="seulement une autre langue"
    3="français et une autre langue"
    4="deux autres langues";
  value $FMT_F1NACTA
    99="ne sait pas"
    10-71="étrangère"
    " " ="française";
run;

proc tabulate data=mabib.hdv2 format=10.0 missing;
  class GNIV FLMAT F1NACTA;
  tables GNIV*F1NACTA(all=" total langue"),
         FLMAT(all="Total origine ressentie ");
  freq poidsf;
  format GNIV $FMT_GNIV.
         FLMAT $FMT_FLMAT.;
         F1NACTA $FMT_F1NACTA
  label gniv="Origine ressentie..."
        flmat="Langue parlée dans l'enfance";
        f1nacta="Nationalité";
run;
```

Journal

```
171 libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";
NOTE: Libref MABIB was successfully assigned as follows:
      Engine:          V9
      Physical Name: E:\S3\EetS\EetS-SASjournalEnregistrements
172
173 options fmtsearch=(mabib);
174 options nodate pageno=1 linesize=140 pagesize=80;
175
176 proc format;
177   value $FMT_GNIV
178     1="région france"
179     2="france"
180     3="europe"
181     4="autre pays ou continent"
182     5="autre pays ou continent"
183     8="refus de répondre"
184     9="ne sait pas";
NOTE: Format $FMT_GNIV is already on the library.
```



```

NOTE: Format $FMT_GNIV has been output.
185   value $FMT_FLMAT
186   1="seulement le français"
187   2="seulement une autre langue"
188   3="français et une autre langue"
189   4="deux autres langues";
NOTE: Format $FMT_FLMAT is already on the library.
NOTE: Format $FMT_FLMAT has been output.
190   value $FMT_FINACTA
191   99="ne sait pas"
192   10-71="étrangère"
193   " "="française";
NOTE: Format $FMT_FINACTA has been output.
194 run;

NOTE: PROCEDURE FORMAT used (Total process time):
      real time      0.00 secondes
      cpu time       0.00 secondes

195
196 proc tabulate data=mabib.hdv2 format=10.0 missing;
197   class GNIV FLMAT FINACTA;
198   tables GNIV*FINACTA(all=" total langue"),
199     FLMAT(all="Total origine ressentie ");
200   freq poids;
201   format GNIV $FMT_GNIV.
202     FLMAT $FMT_FLMAT.
203     FINACTA $FMT_FINACTA.;
204   label gniv="Origine ressentie..."
205     flmat="Langue parlée dans l'enfance"
206     FINACTA="Nationalité";
207 run;

NOTE: There were 8403 observations read from the data set MABIB.HDV2.
NOTE: PROCEDURE TABULATE used (Total process time):
      real time      0.31 secondes
      cpu time       0.06 secondes

```

Tri à plat sur l'âge Par décennies						
		Langue parlée dans l'enfance				Total origine ressentie
		seulement le français	seulement une autre langue	français et une autre langue	deux autres langues	
		N	N	N	N	N
Origine ressentie...	Nationalité					
région france	française	8554482	1180345	1377133	38010	11149970
	étrangère	3379	134123	64781	6102	208385
france	française	21704274	2106513	2576063	75302	26462152
	étrangère	6242	285420	94225	36119	422006
europe	française	3025302	348492	497196	35707	3906697
	étrangère	29588	322460	54511	34988	441547
autre pays ou continent	française	212244	396747	301408	15015	925414
	étrangère	40145	911281	127239	67636	1146301
	ne sait pas	.	21724	.	.	21724
refus de répondre	française	8393	.	3556	.	11949
ne sait pas	française	152754	31100	61163	23559	268576
	étrangère	2781	61843	2472	.	67096
total langue		33739584	5800048	5159747	332438	45031817

On remplace cette procédure

Editeur

```
proc tabulate data=mabib.hdv2 format=10.0 missing;
  class GNIV FLMAT F1NACTA;
  tables GNIV*F1NACTA(all=" total langue"),
         FLMAT(all="Total origine ressentie ");
  freq poidsf;
  format GNIV $FMT_GNIV.
         FLMAT $FMT_FLMAT.;
         F1NACTA $FMT_F1NACTA. ;
  label gniv="Origine ressentie..."
        flmat="Langue parlée dans l'enfance";
        F1NACTA="Nationalité";
run;
```

Par celle-ci pour obtenir les résultats en pourcentages

Editeur

```
proc tabulate data=mabib.hdv2 format=10.0 missing;
  class GNIV FLMAT F1NACTA;
  tables GNIV*F1NACTA*(colpctn="%");
         FLMAT(all="Total origine ressentie ");
  freq poidsf;
  format GNIV $FMT_GNIV.
         FLMAT $FMT_FLMAT.;
         F1NACTA $FMT_F1NACTA. ;
  label gniv="Origine ressentie..."
        flmat="Langue parlée dans l'enfance";
        F1NACTA="Nationalité";
run;
```

Et on obtient le tableau ci-contre

Tri à plat sur l'âge Par décennies						
		Langue parlée dans l'enfance				total origine ressentie
		seulement le français	seulement une autre langue	français et une autre langue	deux autres langues	
Origine ressentie...	Nationalité					
région france	française	% 25.35	20.35	26.69	11.43	24.76
	étrangère	% 0.01	2.31	1.26	1.84	0.46
france	française	% 64.33	36.32	49.93	22.65	58.76
	étrangère	% 0.02	4.92	1.83	10.86	0.94
europe	française	% 8.97	6.01	9.64	10.74	8.68
	étrangère	% 0.09	5.56	1.06	10.52	0.98
autre pays ou continent	française	% 0.63	6.84	5.84	4.52	2.06
	étrangère	% 0.12	15.71	2.47	20.35	2.55
	ne sait pas	% .	0.37	.	.	0.05
refus de répondre	française	% 0.02	.	0.07	.	0.03
ne sait pas	française	% 0.45	0.54	1.19	7.09	0.60
	étrangère	% 0.01	1.07	0.05	.	0.15

La procédure `tabulate` s'applique sur la table `mabib.hdv2`. L'option `format=10.0` permet d'afficher les éléments manquants (`missings`) dans le tableau.

Le paramètre `class` concerne les champs sur lesquelles s'applique la procédure.

Le paramètre `tables` indique la manière dont la procédure traite ces champs (tableau croisé), il est précisé que les champs sont traités en pourcentages. `freq` nous permet alors de pondérer les résultats suivant la variable `poidsf` pour appliquer les données à la population. `format` nous permet d'appliquer les formats aux variables `gniv`, `flmat` et `flnacta`. Le paramètre `label` donne un intitulé à `gniv`, `flmat` et `flnacta`. Run lance le programme.

Question 4.5 : Que pouvez-vous dire du tableau réalisé du point de vue du sociologue ?

En premier lieu, nous constatons grâce à ce tableau que la plupart des personnes de nationalité française se sentent français ou d'une région de France. Les personnes ayant la nationalité d'un autre pays, se sentent appartenant à « autre pays » mais il y a peu de données donc cela est difficilement utilisable. Ensuite, pour les personnes de nationalité française qui se sentent d'une région de France, la langue maternelle ne semble pas avoir influencé leurs choix. En effet, seul le fait d'avoir parlé deux langues autre que le français semble avoir une influence, les autres pourcentages restent globalement les mêmes. De plus, les français ayant parlé au moins le français durant leur enfance se ressentent plus de France que les français ayant parlé seulement une autre langue que le français dans leur enfance. En effet, le pourcentage de personnes n'ayant pas parlé français durant leur enfance et qui se sentent français est plus bas. Le fait de parler français joue donc ici une influence dans le sentiment d'appartenance au pays. En revanche, le fait de se sentir européen parmi les gens ayant la nationalité française ne semble pas être dû à la langue parlée, les pourcentages sont assez proches.

On conclut que le sentiment d'appartenance est très lié à la langue parlée et utilisée pendant l'enfance. En effet, le pourcentage de personnes se sentant d'un autre pays ou continent est bien plus élevé quand cette personne a parlé une autre langue durant l'enfance.

Association de tables et sélection d'observations

Question 5.1 : En vous aidant de la documentation de SAS, indiquez ce que fait chaque ligne de ce programme.

Editeur

```
libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";

data sexes_ages;
    merge mabib.hdv1 mabib.hdv4;
    by ident;
    keep poidsf agee sexee;
run;
```

Journal

```
1 libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";
NOTE: Libref MABIB was successfully assigned as follows:
      Engine:          V9
      Physical Name: E:\S3\EetS\EetS-SASjournalEnregistrements
2
3 data sexes_ages;
4     merge mabib.hdv1 mabib.hdv4;
5     by ident;
6     keep poidsf agee sexee;
7 run;
```

NOTE: There were 8403 observations read from the data set MABIB.HDV1.
NOTE: There were 8403 observations read from the data set MABIB.HDV4.
NOTE: The data set WORK.SEXES_AGES has 8403 observations and 3 variables.
NOTE: L'étape DATA used (Total process time):
 real time 0.89 secondes
 cpu time 0.09 secondes

VIEWTABLE: Work.Sexes_ages							
	POIDSF	AGEE	SEXEE				
1	4766.86520130	60	2	15	7442.02005900	72	2
2	12381.58974600	50	1	16	4994.92379170	49	2
3	14685.43134400	33	2	17	5603.27830200	61	1
4	9516.04993880	82	1	18	4865.79093210	47	2
5	3925.90758810	58	1	19	7482.77897720	73	1
6	9423.81070500	53	1	20	6242.18861580	27	1
7	8793.52510700	39	1	21	13338.34929700	52	1
8	2218.72399590	22	1	22	5860.24274170	54	1
9	7390.94118860	37	2	23	7176.33732540	31	2
10	4099.14242280	26	2	24	4237.51396940	27	2
11	4468.39108490	25	2	25	2467.52268960	75	2
12	12855.25096500	78	2	26	2072.63595910	43	2
13	7579.94880100	40	2	27	3865.35394720	23	1
14	4454.33180170	23	2	28	2478.11428480	54	2
				29	4144.03151530	44	1
				30	4302.71127400	27	2
				31	6861.69278610	56	1
				32	2997.05954480	37	2
				33	2239.61924270	48	2
				34	9450.91996570	45	1

`sexes_ages` est le nom d'une table SAS. L'instruction `merge` permet de fusionner les observations de tables SAS. `by` donne la variable commune définissant les entrées à fusionner. `ident` permet d'associer `hdv1` et `hdv4`. `keep` définit les champs à conserver après la fusion et conserve dans la table `sexes_ages` les `poidsf`.

Question 5.2 : En vous aidant de la documentation de SAS, indiquez ce que fait chaque ligne de ce programme.

Editeur

```
libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";

option fmtsearch=(mabib);
data hommes;
    set sexes_ages (where=(sexee='1'));
run;

data femmes;
    set sexes_ages (where=(sexee='2'));
run;
```

Journal

```
NOTE: The data set WORK.FEMMES has 4616 observations and 3 variables.
NOTE: L'étape DATA used (Total process time):
      real time      0.00 secondes
      cpu time       0.00 secondes
```

Set permet de sélectionner les champs correspondant à une condition particulière grâce au mot-clé where. Ainsi, la table work.hommes contient toutes les entrées de sexes_ages ayant pour valeur sexee='1' et work.femmes contient les entrées de sexes_ages ayant pour valeur sexee='2'.

Question 5.3 : En vous aidant de la documentation de SAS, expliquez comment l'instruction SET telle qu'elle est écrite dans ce programme.

1ere méthode :

(On peut repartir de la proc freq de la question 4.2 du TD4)

Editeur

```
libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";

option fmtsearch=(mabib);

proc freq data=hommes;
    tables agee;
    weight poidsf;
    title "Classe d'âge chez les hommes";
    label agee="Age";
    format agee fmt_age.;
run;

proc freq data=femmes;
    tables agee;
    weight poidsf;
    title "Classe d'âge chez les femmes";
    label agee="Age";
    format agee fmt_age.;
run;
```

Journal

```
97 libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";
NOTE: Libref MABIB was successfully assigned as follows:
      Engine:      V9
      Physical Name: E:\S3\EetS\EetS-SASjournalEnregistrements
98
99 option fmtsearch=(mabib);
100
```

```

101 proc freq data=hommes;
102   tables agee;
103   weight poidsf;
104   title "Classe d'âge chez les hommes";
105   label agee="Age";
106   format agee fmt_age.;
107
108

```

NOTE: Writing HTML Body file: sashtml.htm

NOTE: There were 3787 observations read from the data set WORK.HOMMES.

NOTE: PROCEDURE FREQ used (Total process time):

```

      real time      1.55 secondes
      cpu time       0.21 secondes

```

```

109 proc freq data=femmes;
110   tables agee;
111   weight poidsf;
112   title "Classe d'âge chez les femmes";
113   label agee="Age";
114   format agee fmt_age.;

```

Classe d'âge chez les femmes					Classe d'âge chez les hommes				
Procédure FREQ					Procédure FREQ				
Age					Age				
AGEE	Fréquence	Pourcentage	Fréquence cumulée	Pctage. cumulé	AGEE	Fréquence	Pourcentage	Fréquence cumulée	Pctage. cumulé
10-19 ans	824608.9	3.52	824608.9	3.52	10-19 ans	781288.3	3.61	781288.3	3.61
20-29 ans	4061667	17.34	4886276	20.86	20-29 ans	4179761	19.34	4961050	22.95
30-39 ans	4377099	18.69	9263376	39.56	30-39 ans	4262808	19.72	9223858	42.67
40-49 ans	4050951	17.30	13314326	56.85	40-49 ans	4057845	18.77	13281703	61.44
50-59 ans	3722203	15.89	17036529	72.75	50-59 ans	3501598	16.20	16783301	77.64
60-69 ans	2740294	11.70	19776823	84.45	60-69 ans	2446243	11.32	19229544	88.95
70-79 ans	2645860	11.30	22422683	95.75	70-79 ans	1768555	8.18	20998099	97.14
80-89 ans	838630	3.58	23261313	99.33	80-89 ans	556444.6	2.57	21554544	99.71
90-99 ans	157406.9	0.67	23418720	100.00	90-99 ans	62736.41	0.29	21617280	100.00

2eme méthode:

Editeur

```

libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";

proc sort data=sexes_ages;
  by sexee;
run;

proc format;
  value $fmt_sexe
    '1'="Homme"
    '2'="Femmes";
run;

proc freq data=sexes_ages;
  tables agee;
  by sexee;
  weight poidsf;
  title "Classe d'âge chez les français";
  label agee="Age"
        sexee="Sexe";
  format agee fmt_age.
        sexee $fmt_sexe.;
run;

```

Journal

```
132 libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";
```

NOTE: Libref MABIB was successfully assigned as follows:

```
Engine: V9
```

```
Physical Name: E:\S3\EetS\EetS-SASjournalEnregistrements
```

```
133
```

NOTE: There were 4616 observations read from the data set WORK.FEMMES.

NOTE: PROCEDURE FREQ used (Total process time):

```
real time 11:29.73
```

```
cpu time 4.44 secondes
```

```
134 proc sort data=sexes_ages;
```

```
135 by sexee;
```

```
136 run;
```

NOTE: There were 8403 observations read from the data set WORK.SEXES_AGES.

NOTE: The data set WORK.SEXES_AGES has 8403 observations and 3 variables.

NOTE: PROCEDURE SORT used (Total process time):

```
real time 0.03 secondes
```

```
cpu time 0.00 secondes
```

```
137
```

```
138 proc format;
```

```
139 value $fmt_sexe
```

```
140 '1'="Homme"
```

```
141 '2'="Femmes";
```

NOTE: Format \$FMT_SEXE has been output.

```
142 run;
```

NOTE: PROCEDURE FORMAT used (Total process time):

```
real time 0.04 secondes
```

```
cpu time 0.00 secondes
```

```
143
```

```
144 proc freq data=sexes_ages;
```

```
145 tables agee;
```

```
146 by sexee;
```

```
147 weight poidsf;
```

```
148 title "Classe d'âge chez les français";
```

```
149 label agee="Age"
```

```
150 sexee="Sexe";
```

```
151 format agee fmt_age.
```

```
152 sexee $fmt_sexe.;
```

```
153 run;
```

NOTE: There were 8403 observations read from the data set WORK.SEXES_AGES.

NOTE: PROCEDURE FREQ used (Total process time):

```
real time 0.11 secondes
```

```
cpu time 0.01 secondes
```

Classe d'âge chez les français					Classe d'âge chez les français				
Procédure FREQ					Procédure FREQ				
Sexe=Femmes					Sexe=Homme				
Age					Age				
AGEE	Fréquence	Pourcentage	Fréquence cumulée	Pctage. cumulé	AGEE	Fréquence	Pourcentage	Fréquence cumulée	Pctage. cumulé
10-19 ans	824608.9	3.52	824608.9	3.52	10-19 ans	781288.3	3.61	781288.3	3.61
20-29 ans	4061667	17.34	4886276	20.86	20-29 ans	4179761	19.34	4961050	22.95
30-39 ans	4377099	18.69	9263376	39.56	30-39 ans	4262808	19.72	9223858	42.67
40-49 ans	4050951	17.30	13314326	56.85	40-49 ans	4057845	18.77	13281703	61.44
50-59 ans	3722203	15.89	17036529	72.75	50-59 ans	3501598	16.20	16783301	77.64
60-69 ans	2740294	11.70	19776823	84.45	60-69 ans	2446243	11.32	19229544	88.95
70-79 ans	2645860	11.30	22422683	95.75	70-79 ans	1768555	8.18	20998099	97.14
80-89 ans	838630	3.58	23261313	99.33	80-89 ans	556444.6	2.57	21554544	99.71
90-99 ans	157406.9	0.67	23418720	100.00	90-99 ans	62736.41	0.29	21617280	100.00

Tableaux disjonctifs complets

Question 5.4 : En vous aidant de la seconde étape DATA du programme ci-dessus, rédigez et exécutez le tableau disjonctif complet de la variable GLVIE.

Editeur

```
libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";

data test;
    input num sexe;
    cards;
    1 1
    2 2
    3 1
    4 1
    5 2
    ;
run;

data disjonctif;
    set test;

    array ar_sexe(sexe) sex1 sex2;
    array ar_modal sex1--sex2;
    array disjonct ar_sexe;
    do over ar_modal;
        ar_modal=0;
    end;
    do over disjonct;
        disjonct=1;
    end;
run;

proc print;
run;
```


Journal

```
154 libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";
```

NOTE: Libref MABIB was successfully assigned as follows:

Engine: V9

Physical Name: E:\S3\EetS\EetS-SASjournalEnregistrements

```
155
```

```
156 data test;
```

```
157   input num sexe;
```

```
158   cards;
```

NOTE: The data set WORK.TEST has 5 observations and 2 variables.

NOTE: L'étape DATA used (Total process time):

real time 0.01 secondes

cpu time 0.00 secondes

```
164   ;
```

```
165 run;
```

```
166
```

```
167 data disjonctif;
```

```
168   set test;
```

```
169
```

```
170   array ar_sexe(sexe) sex1 sex2;
```

```
171   array ar_modal sex1--sex2;
```

```
172   array disjonct ar_sexe;
```

```
173   do over ar_modal;
```

```
174     ar_modal=0;
```

```
175   end;
```

```
176   do over disjonct;
```

```
177     disjonct=1;
```

```
178   end;
```

```
179 run;
```

NOTE: There were 5 observations read from the data set WORK.TEST.

NOTE: The data set WORK.DISJONCTIF has 5 observations and 4 variables.

NOTE: L'étape DATA used (Total process time):

real time 0.01 secondes

cpu time 0.00 secondes

```
180
```

```
181 proc print;
```

```
182 run;
```

NOTE: There were 5 observations read from the data set WORK.DISJONCTIF.

NOTE: PROCEDURE PRINT used (Total process time):

real time 0.10 secondes

cpu time 0.04 secondes

Classe d'âge chez les français

Obs.	num	sexe	sex1	sex2
1	1	1	1	0
2	2	2	0	1
3	3	1	1	0
4	4	1	1	0
5	5	2	0	1

Classe d'âge chez les français

Obs.	num	sexe	age	note	sex1	sex2	age1	age2	age3	not1	not2	not3
1	1	1	1	2	1	0	1	0	0	0	1	0
2	2	2	3	2	0	1	0	0	1	0	1	0
3	3	1	1	3	1	0	1	0	0	0	0	1
4	4	1	3	1	1	0	0	0	1	1	0	0
5	5	2	2	1	0	1	0	1	0	1	0	0

Classe d'âge chez les français

Obs.	num	sexe	age	note	sex1	sex2	age1	age2	age3	not1	not2	not3
1	1	1	1	2	1	0	1	0	0	0	1	0
2	2	2	3	2	0	1	0	0	1	0	1	0
3	3	1	1	3	1	0	1	0	0	0	0	1
4	4	1	3	1	1	0	0	0	1	1	0	0
5	5	2	2	1	0	1	0	1	0	1	0	0

Editeur

```

libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";

data test;
    input num sexe age note;
    cards;
1 1 1 2
2 2 3 2
3 1 1 3
4 1 3 1
5 2 2 1
;
run;
data disjonctif;
    set test;
    array ar_sexe(sexe) sex1 sex2;
    array ar_age(age) age1 age2 age3;
    array ar_not(note) not1 not2 not3;
    array ar_modal sex1--not3;
    array disjonct ar_sexe ar_age ar_not;
    do over ar_modal;
        ar_modal=0;
    end;
    do over disjonct;
        disjonct=1;
    end;
run;

proc print;
run;

proc print data=disjonctif(obs=50);
run;

```

Journal

```
270 libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";
NOTE: Le nom de la biblio. MABIB se rapporte à la même biblio. physique que TMP1.
NOTE: Libref MABIB was successfully assigned as follows:
```

```
Engine:      V9
Physical Name: E:\S3\EetS\EetS-SASjournalEnregistrements
```

```
271 data test;
272   input num sexe age note;
273   cards;
```

NOTE: The data set WORK.TEST has 5 observations and 4 variables.

NOTE: L'étape DATA used (Total process time):

```
real time    0.00 secondes
cpu time     0.00 secondes
```

```
279   ;
280 run;
281 data disjonctif;
282   set test;
283   array ar_sexe(sexe) sex1 sex2;
284   array ar_age(age) age1 age2 age3;
285   array ar_not(note) not1 not2 not3;
286   array ar_modal sex1--not3;
287   array disjonct ar_sexe ar_age ar_not;
288   do over ar_modal;
289     ar_modal=0;
290   end;
291   do over disjonct;
292     disjonct=1;
293   end;
294 run;
```

NOTE: There were 5 observations read from the data set WORK.TEST.

NOTE: The data set WORK.DISJONCTIF has 5 observations and 12 variables.

NOTE: L'étape DATA used (Total process time):

```
real time    0.00 secondes
cpu time     0.00 secondes
```

```
295
296 proc print;
297 run;
```

NOTE: There were 5 observations read from the data set WORK.DISJONCTIF.

NOTE: PROCEDURE PRINT used (Total process time):

```
real time    0.05 secondes
cpu time     0.00 secondes
```

```
298
299
300 proc print data=disjonctif(obs=50);
301 run;
```

NOTE: There were 5 observations read from the data set WORK.DISJONCTIF.

NOTE: PROCEDURE PRINT used (Total process time):

```
real time    0.04 secondes
cpu time     0.00 secondes
```

Editeur (on remplace la variable sexe par la variable glvien)

```
libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";

data disjonctif;
  set test;

  array ar_glvien(glvien) glvie1 glvie2 glvie3 glvie4;
  array ar_modal glvie1--glvie4;
  array disjonct ar_glvien;
  do over ar_modal;
    ar_modal=0;
  end;
  do over disjonct;
```

```

        disjonct=1;
    end;
run;

proc print data=disjonctif(obs=50);
run;

```

Question 5.5 : En vous aidant de la seconde étape DATA du programme ci-dessus, rédigez et exécutez le tableau disjonctif complet de la variable GLVIE (lieu où l'enquête aimerait vivre plus tard) et de la variable GRFAM (membres de la famille dans la même région que celle du lieu de résidence actuelle)

Editeur

```

libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";

data test;
    set mabib.hdv2;
    length glvien 4.;
    glvien = glvie;
    length grfamn 4.;
    grfamn = grfam;
    keep glvien grfamn;
run;

data disjonctif;
    set test;
    array ar_glvie(glvien) glvie1 glvie2 glvie3 glvie4;
    array ar_grfam(grfamn) grfam1 grfam2 grfam3 grfam4 grfam5;
    array ar_modal glvie1--grfam5;
    array disjonct ar_glvie ar_grfam;

    do over ar_modal;
        ar_modal=0;
    end;
    do over disjonct;
        disjonct=1;
    end;
run;

proc print data=disjonctif(obs=50);
run;

```

Journal

```

183 libname mabib "E:\S3\EetS\EetS-SASjournalEnregistrements";
NOTE: Libref MABIB was successfully assigned as follows:
      Engine:          V9
      Physical Name: E:\S3\EetS\EetS-SASjournalEnregistrements
184
185 data test;
186   set mabib.hdv2;
187   length glvien 4.;
188   glvien = glvie;
189   length grfamn 4.;
190   grfamn = grfam;
191   keep glvien grfamn;
192 run;

NOTE: Character values have been converted to numeric values at the places given by: (Line):(Column).
      188:14   190:14
NOTE: There were 8403 observations read from the data set MABIB.HDV2.
NOTE: The data set WORK.TEST has 8403 observations and 2 variables.
NOTE: L'étape DATA used (Total process time):

```

real time	0.26 secondes
cpu time	0.03 secondes

```
193
194 data disjonctif;
195   set test;
196   array ar_glvie(glvien) glvie1 glvie2 glvie3 glvie4;
197   array ar_grfam(grfamn) grfam1 grfam2 grfam3 grfam4 grfam5;
198   array ar_modal glvie1--grfam5;
199   array disjonct ar_glvie ar_grfam;
200
201   do over ar_modal;
202     ar_modal=0;
203   end;
204   do over disjonct;
205     disjonct=1;
206   end;
207 run;
```

NOTE: There were 8403 observations read from the data set WORK.TEST.

NOTE: The data set WORK.DISJONCTIF has 8403 observations and 11 variables.

NOTE: L'étape DATA used (Total process time):

real time	0.01 secondes
cpu time	0.01 secondes

```
208
209 proc print data=disjonctif(obs=50);
210 run;
```

NOTE: There were 50 observations read from the data set WORK.DISJONCTIF.

NOTE: PROCEDURE PRINT used (Total process time):

real time	0.07 secondes
cpu time	0.03 secondes

Obs.	glvien	grfamn	glvie1	glvie2	glvie3	glvie4	grfam1	grfam2	grfam3	grfam4	grfam5
1	1	5	1	0	0	0	0	0	0	0	1
2	1	3	1	0	0	0	0	0	1	0	0
3	1	2	1	0	0	0	0	1	0	0	0
4	3	3	0	0	1	0	0	0	1	0	0
5	1	3	1	0	0	0	0	0	1	0	0
6	1	2	1	0	0	0	0	1	0	0	0
7	1	2	1	0	0	0	0	1	0	0	0
8	2	2	0	1	0	0	0	1	0	0	0
9	4	3	0	0	0	1	0	0	1	0	0
10	4	2	0	0	0	1	0	1	0	0	0
11	4	5	0	0	0	1	0	0	0	0	1
12	3	2	0	0	1	0	0	1	0	0	0
13	2	3	0	1	0	0	0	0	1	0	0
14	2	3	0	1	0	0	0	0	1	0	0
15	1	3	1	0	0	0	0	0	1	0	0
16	1	5	1	0	0	0	0	0	0	0	1
17	3	5	0	0	1	0	0	0	0	0	1
18	1	3	1	0	0	0	0	0	1	0	0
19	4	3	0	0	0	1	0	0	1	0	0
20	4	3	0	0	0	1	0	0	1	0	0
21	2	2	0	1	0	0	0	1	0	0	0
22	4	5	0	0	0	1	0	0	0	0	1
23	2	2	0	1	0	0	0	1	0	0	0
24	4	4	0	0	0	1	0	0	0	1	0
25	2	2	0	1	0	0	0	1	0	0	0
26	2	3	0	1	0	0	0	0	1	0	0
27	3	3	0	0	1	0	0	0	1	0	0
28	3	2	0	0	1	0	0	1	0	0	0
29	2	3	0	1	0	0	0	0	1	0	0
30	3	2	0	0	1	0	0	1	0	0	0
31	2	5	0	1	0	0	0	0	0	0	1
32	4	2	0	0	0	1	0	1	0	0	0
33	4	3	0	0	0	1	0	0	1	0	0
34	1	2	1	0	0	0	0	1	0	0	0
35	1	3	1	0	0	0	0	0	1	0	0
36	2	4	0	1	0	0	0	0	0	1	0
37	4	2	0	0	0	1	0	1	0	0	0
38	3	5	0	0	1	0	0	0	0	0	1
39	4	3	0	0	0	1	0	0	1	0	0
40	1	3	1	0	0	0	0	0	1	0	0
41	4	2	0	0	0	1	0	1	0	0	0
42	2	5	0	1	0	0	0	0	0	0	1
43	4	2	0	0	0	1	0	1	0	0	0
44	4	3	0	0	0	1	0	0	1	0	0
45	4	2	0	0	0	1	0	1	0	0	0
46	1	2	1	0	0	0	0	1	0	0	0
47	1	3	1	0	0	0	0	0	1	0	0
48	1	3	1	0	0	0	0	0	1	0	0
49	4	2	0	0	0	1	0	1	0	0	0
50	2	2	0	1	0	0	0	1	0	0	0

