



MÉMOIRE DE STATISTIQUES DESCRIPTIVES

---

# **Comment le niveau d'études influence-t-il la situation familiale ?**

---

**Situation familiale et éducation**

Gabriel WATKINSON  
Mohamed GUENNOUNI HASSANI

Encadré par M. CUBILLÉ

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>Présentation des données</b>	<b>2</b>
<b>1 Éducation et situation maritale</b>	<b>4</b>
1.1 Niveau d'études et célibat . . . . .	4
1.2 Niveau d'études et âge de la mise en couple . . . . .	6
<b>2 Homogamie et reproduction sociale sous le spectre de l'éducation</b>	<b>8</b>
2.1 Liens entre le niveau de la personne et des parents . . . . .	8
2.2 Liens entre le niveau de la personne et du partenaire . . . . .	9
<b>3 Propension à fonder un foyer</b>	<b>10</b>
3.1 Éducation et paternité / maternité . . . . .	10
3.2 Éducation et nombre de personnes dans le foyer . . . . .	11
<b>Conclusion</b>	<b>13</b>
<b>Bibliographie</b>	<b>14</b>
<b>Annexe</b>	<b>i</b>
Graphiques et tableaux . . . . .	i
Code et commentaires . . . . .	iii

# Introduction

L'éducation correspond à la mise en œuvre de moyens propres à assurer la formation et le développement d'un être humain. Ainsi, l'éducation d'une personne dépend nécessairement de son environnement. Elle débute à la naissance dans le cocon familial et se poursuit tout au long de la vie en fonction des interactions entre la personne et son environnement. De fait, famille et éducation sont profondément intriquées. Toutefois, l'étude de certains paramètres permet de dresser des liens entre les deux entités. S'il est vrai que définir l'éducation d'une personne reviendrait à connaître la personne dans sa totalité, le niveau d'études atteint par la personne reste un bon indicateur pour pouvoir juger de l'éducation au sens commun du terme. En s'appuyant sur cette approche, nous comprenons aisément que le niveau d'études d'une personne peut avoir des répercussions économiques et sociales qui expliqueraient potentiellement sa situation familiale, sans oublier les contraintes inhérentes à l'éducation scolaire comme par exemple le fait de devoir quitter son domicile familial pour poursuivre ses études.

La situation familiale, quant à elle, doit être comprise au sens large. Elle englobe tout aussi bien la position dans la hiérarchie familiale (mère, fils, beau-père,...) que la situation de couple (marié, divorcé,...) ou encore le fait de vivre ou non au sein du même domicile. Inversement, la situation familiale pourrait également expliquer le niveau d'éducation d'une personne. En effet, la position de la personne au sein de sa fratrie pourrait par exemple avoir des répercussions sur son parcours scolaire [4]. Néanmoins, les données recueillies dont nous disposons nous invitent essentiellement à n'étudier que l'impact de l'éducation sur la situation familiale. Ainsi, au cours de notre étude, nous nous attacherons tout particulièrement à étudier les répercussions du niveau d'études sur la situation familiale de personnes majeures. Subséquemment, se pose alors la question de savoir dans quelle mesure le niveau d'éducation influence la situation familiale.

En utilisant les données fournies par le *round 6 de l'European Social Survey*, nous étudierons dans le cadre de la France le lien entre le niveau d'études et plusieurs variables qui rendent compte de la situation familiale. Malgré les limites intrinsèques à notre démarche qui rendent difficiles une analyse *ceteris paribus*, nos résultats sont en cohérence avec les travaux antérieurs de sociologues et statisticiens dans ce domaine.

Dans un premier temps, nous étudierons le lien entre l'éducation et le statut marital, puis nous mettrons en exergue dans un deuxième temps l'homogamie sociale en France à travers le niveau d'études des parents et enfin, nous nous intéresserons aux liens entre éducation et propension à fonder un foyer.

# Présentation des données

Les données que nous utilisons proviennent toutes du round 6 de l'European Social Survey de 2012. La base initiale tronquée uniquement à la France nous fournit un échantillon de 1968 personnes. Nous présentons ci-dessous l'ensemble des variables utilisées, leurs explications selon l'ESS ainsi que la manière avec laquelle nous les avons traitées. Dans notre approche, nous ne prenons pas en considération les réponses manquantes et les refus de réponses car nous n'observons pas de corrélation entre les non-réponses et les autres paramètres étudiés.

## Variables relatives à l'éducation :

Nous avons utilisé les quatre variables **EDLVDFR**, **EDLVFDFR**, **EDLVMDFR**, **EDLVPDFR** qui représentent respectivement le plus haut niveau d'éducation atteint par la personne sondée, par son père, par sa mère et par son partenaire. Ces variables sont qualitatives ordinales et présentent toutes les mêmes modalités au nombre de 31. Par souci de clarté et de représentativité, nous avons décidé de recoder ces variables en les regroupant en 5 modalités : Brevet, Bac, Licence, Master, Doctorat. Nous leur avons assigné un type numérique, respectivement 1, 2, 3, 4 et 5, pour une plus grande commodité dans leur utilisation.

## Variables relatives à la situation familiale :

La variable **MARSTS** représente le statut marital de la personne. Elle présente 5 modalités. Néanmoins, nous recodons cette variable pour regrouper ces modalités en deux nouvelles modalités distinctes : avoir eu une union civile légale et ne jamais en avoir eu.

La variable **LVGPTNEA** renseigne sur le fait que la personne sondée a déjà vécu ou non avec un partenaire.

La variable **CHLDHHE** renseigne sur le fait que la personne a déjà eu un enfant vivant dans son foyer. Cet enfant peut être le sien, adopté ou l'enfant de son partenaire.

La variable **CHLDHM** renseigne sur le fait qu'il y a ou non des enfants qui vivent dans le foyer de la personne sondée.

La variable **HHMMB** renseigne sur le nombre de personnes qui vivent régulièrement dans le foyer de la personne sondée. Pour avoir des modalités d'effectif comparable et statistiquement significatives, nous avons décidé de recoder cette variable en 6 modalités : 1, 2, 3, 4, 5, 6 et plus.

## Variable transversale :

La variable **AGEA** représente l'âge de la personne. Elle nous permettra de cibler certaines tranches d'âge en particulier et voir dans quelle mesure l'éducation retarde la mise en place d'une situation familiale.

## Variables de pondération :

La variable **PSPWGHT** (post-stratification weight) permet de pondérer en réduisant les erreurs liées à l'échantillonnage et les potentiels biais de non-réponse. Elle a été construite à partir de l'âge, du genre, de l'éducation et de la région de la personne sondée.

La variable **POND** a été construite à partir de **PSPWGHT** et de **PWEIGHT** pour avoir des résultats qui correspondent à l'échelle de la France.

# 1 Éducation et situation maritale

L'éducation impose de nombreuses contraintes, que ce soit au niveau de l'éloignement géographique par rapport à ses parents ou bien par rapport à la durée et le temps investi dans les études. D'un autre côté, les études peuvent également améliorer le statut social de l'individu et donc lui ouvrir de plus grandes opportunités de rencontre. L'objectif de cette partie est de pouvoir quantifier, s'il existe, l'impact de l'éducation sur la situation maritale de l'individu.

## 1.1 Niveau d'études et célibat

Nous nous intéressons aux liens entre niveau d'études et célibat dans le sens légal du terme, c'est-à-dire si la personne a déjà eu une situation maritale officielle ou non. Pour cela, nous traçons tout d'abord un graphique qui représente la situation maritale en fonction du niveau d'études atteint. Nous nous restreignons aux personnes qui ont 26 ans ou plus pour s'assurer que toutes les personnes étudiées ont bien fini leurs études. Dans le cas inverse, une personne de 24 ans qui est en train de préparer un doctorat aura comme plus haut niveau atteint un master.

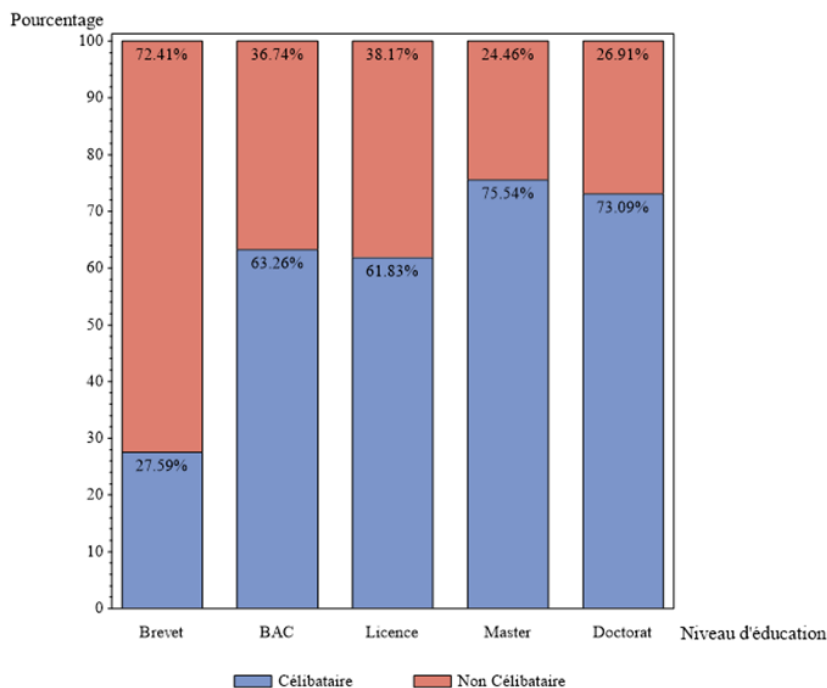


FIGURE 1.1 – Situation maritale en fonction de l'éducation pour les individus de plus de 26 ans

Les résultats obtenus peuvent être regroupés en trois catégories : les titulaires de Brevet dont environ 28% d'entre eux sont encore célibataires, les titulaires du Bac et de licence dont 62% d'entre eux environ sont encore célibataires et finalement les titulaires de Master et Doctorat dont le pourcentage de célibataires s'élève à 74% environ.

Statistic	DF	Value	Prob
Chi-Square	4	86.6036	<.0001
Cramer's V		0.3586	

TABLE 1.1 – Mesure de l'association entre le niveau d'études et le statut marital pour les individus de plus de 26 ans

Nous observons alors une corrélation entre durée des études et célibat. En s'intéressant au test du chi-2, on remarque une p-valeur inférieure à 1%. Nous considérons alors que les deux variables ne sont pas indépendantes. Par ailleurs, le V de Cramer de 0,36 souligne un lien modéré entre ces deux variables. Toutefois, il serait possible que la durée des études ne soit qu'un facteur parmi d'autres. Il est concevable que, au-delà de leurs durées, les études contribuent à modifier la perception de la personne par rapport au couple. Cependant, ces résultats peuvent être biaisés, car le nombre de personnes qui entreprennent de longues études est plus élevé maintenant qu'il y a 50 ans (FIGURE 1 en annexe).

En s'intéressant à savoir si la personne a déjà été en couple avant de se marier, nous constatons qu'à mesure que les études entreprises sont longues, la personne est plus susceptible d'avoir vécu en couple hors mariage.

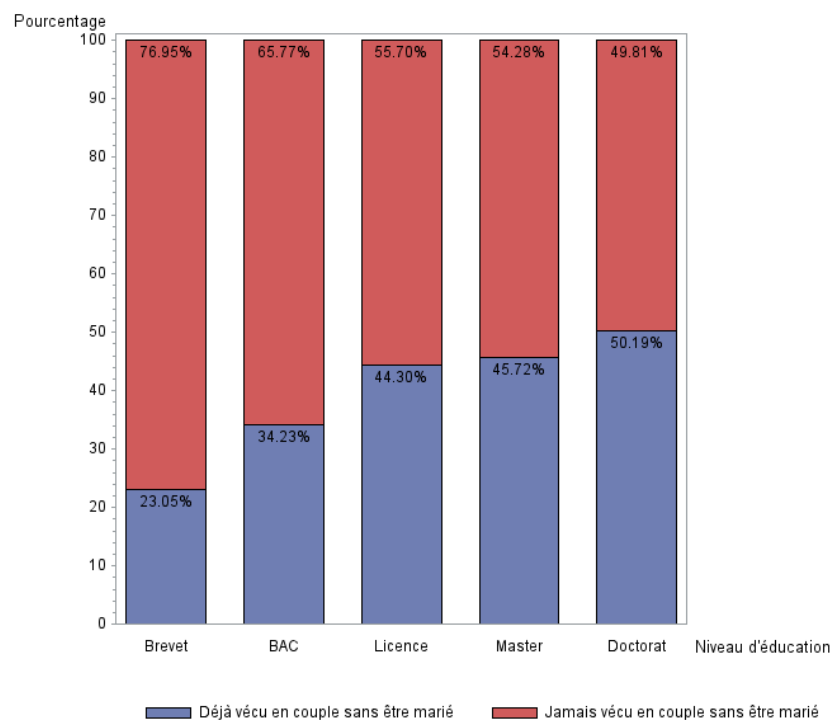


FIGURE 1.2 – Pourcentage de personne ayant déjà vécu en couple sans être marié ou non en fonction de l'éducation pour les individus de plus de 26 ans

Statistic	DF	Value	Prob
Chi-Square	4	47.6593	<0.0001
Cramer's V		0.1815	

TABLE 1.2 – Mesure de l'association entre le niveau d'études et le fait d'avoir déjà vécu en couple sans être marié ou non pour les individus de plus de 26 ans

Le test de chi-2 nous révèle là encore que le niveau d'études et le fait d'avoir vécu en couple avant mariage ne sont pas indépendants à un seuil inférieur à 1%. Ainsi la mise en couple, qu'elle soit officielle ou non, ne dépend pas significativement du niveau d'études. Nonobstant, les personnes qui effectuent de plus longues études préfèrent moins s'engager officiellement, ce qui pourrait s'expliquer par un manque de temps ou un manque d'envie, choisissant de se concentrer plutôt sur l'accomplissement de leurs études.

## 1.2 Niveau d'études et âge de la mise en couple

Si nous avons observé qu'il y a plus de personnes qui sont légalement célibataires à mesure que le niveau des études augmente pour les personnes de plus de 26 ans, il serait intéressant de voir si cette tendance s'applique pour les individus de plus de 50 ans. Nous avons choisi arbitrairement l'âge de 50 ans car il correspond à un âge à partir duquel la plupart des liaisons de couple se sont déroulées antérieurement et ont été officiellement signalées.

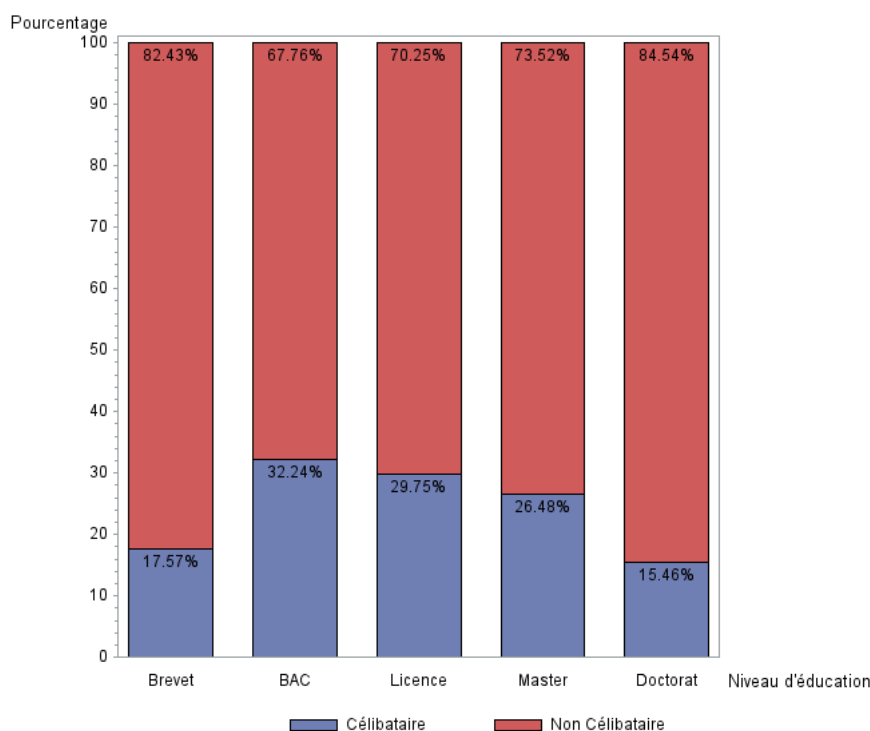


FIGURE 1.3 – Situation maritale en fonction de l'éducation pour les individus de plus de 50 ans



Statistic	DF	Value	Prob
Chi-Square	4	7.1735	0.1270
Cramer's V		0.1547	

TABLE 1.3 – Mesure de l'association entre le niveau d'études et le fait d'avoir déjà vécu en couple sans être marié ou non pour les individus de plus de 50 ans

Le graphique nous montre alors que le pic des personnes célibataires officiellement se situe au niveau des titulaires au plus du Bac avec 32,24% alors que les minimas se situent chez les titulaires d'au plus le brevet avec 17,57% et surtout chez les personnes ayant obtenu un doctorat avec 15,46%. La p-valeur assez élevée (12,70%) souligne le fait que la situation maritale tend à devenir indépendante par rapport au niveau d'études quand les individus deviennent plus âgés. Ainsi, cela pourrait s'expliquer par le fait que les docteurs finissent par avoir une relation maritale officielle dans des proportions élevées mais avec un retard d'années par rapport aux autres. À l'inverse, les titulaires d'au plus le Bac ont plus de mal à se mettre en relation officielle à mesure que le temps passe. Les personnes titulaires d'au plus le brevet profitent du fait qu'ils se sont mis dans des relations officielles assez vite. À travers cette comparaison, il apparaît que les docteurs profitent de la situation sociale acquise tardivement pour contrebalancer leur célibat antérieur ayant pour cause manifeste les études qu'ils ont entreprises.

## 2 Homogamie et reproduction sociale sous le spectre de l'éducation

Après avoir étudié les liens entre éducation et situation maritale, nous nous proposons maintenant de s'intéresser à comprendre dans quelle mesure l'éducation influe sur le choix du partenaire et participe à une certaine reproduction sociale.

### 2.1 Liens entre le niveau de la personne et des parents

En considérant le plus haut niveau d'études atteint comme une variable qualitative ordonnée en fonction de la durée des études, nous pouvons alors dresser des corrélations entre le niveau atteint par la personne et celui de ses parents.

	edlvdf	edlvmdfr	edlvdf
Highest level of education, France edlvdf	1.00000	0.38109 <.0001	0.41074 <.0001
Mother's highest level of education, France edlvmdfr	0.38109 <.0001	1.00000	0.58563 <.0001
Father's highest level of education, France edlvdf	0.41074 <.0001	0.58563 <.0001	1.00000

TABLE 2.1 – Coefficients du  $\tau_b$  de Kendall

Nous remarquons alors une corrélation modérée entre le niveau d'éducatons des parents et celui de la personne sondée. Le  $\tau_b$  de Kendall est de 0,38 entre la personne et sa mère et de 0,41 entre la personne et son père. La p-valeur nous permet de rejeter l'hypothèse d'indépendance au seuil de 1%. Mais si nous nous intéressons maintenant au lien entre le niveau d'études du père et de la mère, nous observons une corrélation plus forte avec un  $\tau_b$  de Kendall de 0,59. Là encore, nous pouvons rejeter l'hypothèse d'indépendance au seuil de 1%. Si la personne tend à avoir un niveau d'études plus ou moins analogue de celui de ses parents, il semble que l'homogamie horizontale est plus importante. Néanmoins, ces résultats sont en partie biaisés par le fait que les jeunes générations entreprennent de plus longues études que les précédentes.

Personne	Médianes du plus haut niveau d'étude
Sujet sondée	Baccalauréat
Père	Scolarité suivie de la 6 <sup>ème</sup> à la 3 <sup>ème</sup>
Mère	Scolarité suivie de la 6 <sup>ème</sup> à la 3 <sup>ème</sup>

TABLE 2.2 – Médianes du plus haut niveau d'étude de la personne sondée et de ses parents

L'étude des médianes du plus haut niveau d'études atteint montre bien que les plus jeunes se lancent dans des études plus longues. Le phénomène de reproduction sociale à travers l'éducation reste néanmoins présent au sein de la famille nucléaire. Le meilleur niveau d'études des enfants pourrait s'expliquer par un changement de paradigme par rapport aux études entre les générations mais aussi peut-être une envie des enfants de faire mieux que leurs parents.

## 2.2 Liens entre le niveau de la personne et du partenaire

Après s'être intéressé aux parents, nous pouvons maintenant comparer le niveau d'études de la personne avec son partenaire.

	edlvdfr	edlvpdfr
Highest level of education, France edlvdfr	1.00000	0.48532 <.0001
Partner's highest level of education, France edlvpdfr	0.48532 <.0001	1.00000

TABLE 2.3 – Coefficients du  $\tau_b$  de Kendall

Nous observons alors une corrélation forte entre le plus haut niveau d'études atteint par la personne et celui de son partenaire avec un  $\tau_b$  de Kendall de 0,49. La p-valeur nous invite à rejeter l'hypothèse d'indépendance au seuil de 1%. Cette corrélation est plus forte que la corrélation précédente entre la personne et ses parents ( $\tau_b$  de Kendall de l'ordre de 0,4). Ainsi, l'homogamie en termes de niveaux de diplôme est plus forte que l'effet de reproduction sociale entre les parents et l'enfant. Cela s'explique par le fait qu'un nombre important de couples se forment au moment des études ou dans un cadre professionnel dans lequel les études de chacun sont assez homogènes [3].

### 3 Propension à fonder un foyer

Après avoir montré l'impact de l'éducation sur le choix du partenaire, il s'agit d'étudier dans cette partie dans quelle mesure le niveau d'éducation influe sur le fait d'avoir des enfants et sur la formation d'un foyer

#### 3.1 Éducation et paternité / maternité

L'éducation, ne serait-ce que par la longue durée de certaines études, peut avoir des conséquences sur le choix d'avoir des enfants et l'âge du premier enfant. Nous nous intéressons à la présence d'enfants dans le ménage selon le niveau d'études pour les individus de plus de 26 ans. Cet enfant peut être celui de la personne, de son partenaire ou un enfant adopté.

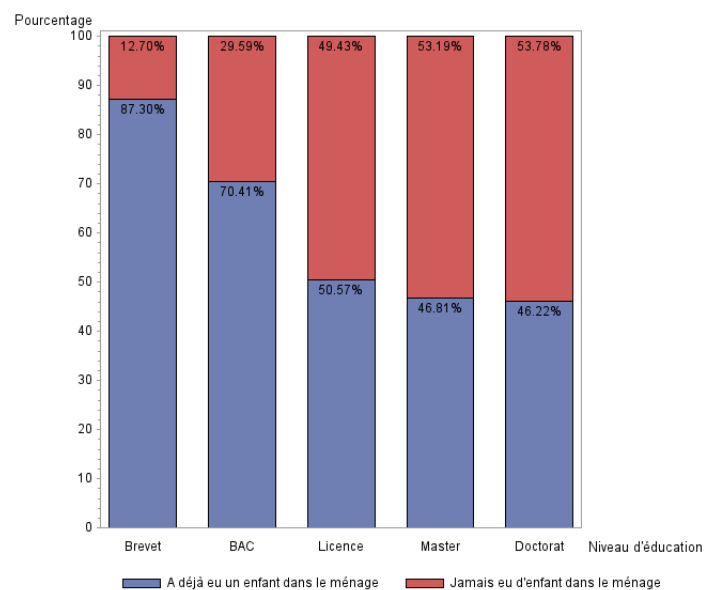


FIGURE 3.1 – Pourcentage de personne de plus de 26 ans ayant déjà eu un enfant dans le ménage selon le niveau d'éducation

Statistic	DF	Value	Prob
Chi-Square	4	103.0988	<.0001
Cramer's V		0.3327	

TABLE 3.1 – Mesure de l'association entre le niveau d'études et le fait d'avoir eu un enfant dans le ménage ou non pour les individus de plus de 26 ans

En effet, nous remarquons dans ce graphique que le pourcentage de personnes ayant un enfant dans leur ménage décroît à mesure que les études entreprises sont longues. Ces résultats vont dans le même sens que ceux de la situation maritale en fonction de l'éducation (Figure 1.1). Nous observons que 87,30% des titulaires d'au plus le brevet ont déjà eu un enfant dans le ménage contre 46,22% des personnes titulaires d'un doctorat. Le test du chi-2 nous invite à rejeter l'hypothèse d'indépendance au seuil de 1%. Ainsi, les personnes qui effectuent de plus longues études s'engagent moins dans des relations officielles et préfèrent attendre avant d'avoir un enfant dans le ménage. Comme avec la situation maritale, nous remarquons que les personnes qui font de longues études finissent par « rattraper » leur retard par rapport aux autres (Figure 2 de l'annexe). Cela corrobore l'idée que les individus qui font de longues études n'ont pas une vision différente du couple et du fait d'être parents mais qu'ils préfèrent plutôt se concentrer sur leur éducation lorsqu'ils sont plus jeunes.

### 3.2 Éducation et nombre de personnes dans le foyer

Nous allons maintenant étudier le rapport entre éducation et propension de créer un grand foyer à travers le nombre de personnes vivant dans le même foyer que la personne. Nous choisissons de nous intéresser qu'aux personnes qui ont entre 36 et 52 ans pour avoir une fourchette qui permet à la fois de s'assurer que la majorité des personnes qui souhaitent des enfants sont devenus parents, que les enfants ne sont pas suffisamment grands pour avoir quitté le domicile familiale et que l'échantillon soit statistiquement significatif.

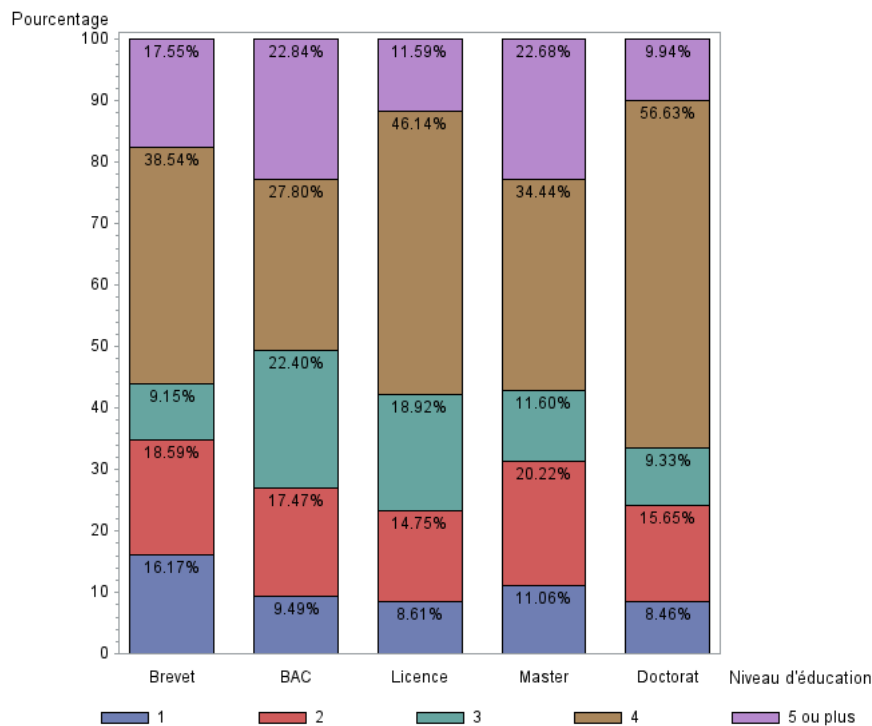


FIGURE 3.2 – Nombre de personnes vivant régulièrement dans le ménage selon le niveau d'éducation pour les individus de 36 à 52 ans

Statistic	DF	Value	Prob
Chi-Square	16	29.7850	0.0191
Cramer's V		0.1164	

TABLE 3.2 – Mesure de l'association entre le niveau d'études et nombre de personnes vivant régulièrement dans le ménage pour les individus de 36 à 52 ans

Le test du chi-2 nous invite tout d'abord à rejeter l'hypothèse d'indépendance au seuil de 5% (p-valeur inférieure à 0.02).

Nous remarquons que le pourcentage de personnes vivant seul est le plus élevé chez les personnes titulaires d'au plus le brevet (16,17%). La difficulté de ces personnes pour vivre avec quelqu'un pourrait s'expliquer par leur situation sociale plus précaire. Leurs professions peuvent nécessiter une plus grande mobilité s'ils travaillent par exemple dans le bâtiment ou les travaux publics.

Pour les foyers à 3 personnes, ils sont sous-représentés chez les titulaires d'au plus le brevet ainsi que les titulaires de doctorat. Nous observons cependant que pour ces deux catégories, les foyers avec 4 personnes ou plus sont bien représentés. Ainsi, les titulaires d'au plus le brevet, quand ils peuvent construire un foyer, préfèrent avoir plus qu'un enfant. Les docteurs, quant à eux, ont une situation sociale qui leur permet plus aisément de construire un foyer. Néanmoins, la majorité d'entre eux préfère avoir 2 enfants plutôt que 3 enfants ou plus. En effet, seulement 9,94% des docteurs ont 3 enfants ou plus. Certains docteurs semblent donc choisir intentionnellement de n'avoir que 2 enfants. Cette caractéristique se retrouve aussi chez les titulaires d'au plus une licence mais peut-être que le choix de ces derniers est plus motivé par une réalité financière plutôt qu'une question d'envie.

En comparant les titulaires d'au plus le Bac et ceux de licence, nous remarquons que les premiers ont plus tendance à avoir des foyers à 3 enfants ou plus. Cela peut être dû au fait que les titulaires d'au plus une licence sont plus soucieux de fournir une meilleure qualité de vie aux enfants, quitte à en avoir que 2. Les titulaires de master, quant à eux, profitent de leur situation pour avoir une grande famille (22,68% ont 3 enfants ou plus).

Ces interprétations s'appuient sur le fait qu'à mesure que les études sont longues, la situation sociale et financière de la personne s'améliore, ce qui n'est pas toujours vrai. Elles s'appuient aussi sur le fait qu'un foyer est constitué de deux parents et d'enfants alors qu'il est envisageable qu'un seul parent vive avec ses enfants (parent divorcé sans remariage). Néanmoins, notre grille de lecture semble fournir des résultats satisfaisants et conformes au bon sens. Des données statistiques plus fournies nous auraient permis d'affirmer certaines interprétations avec moins de réserves.

# Conclusion

À travers notre étude statistique, nous constatons que le niveau d'éducation influe de diverses façons sur la situation familiale de l'individu. De longues études ont ainsi un impact direct sur la formation officielle de couples ainsi que sur le fait de devenir parents. Par ailleurs, l'éducation participe considérablement à l'homogamie et à la reproduction sociale en France. Les corrélations obtenues tendent à montrer que l'éducation est un facteur social discriminant important dans la société française. Enfin, l'éducation joue un rôle central dans la formation des foyers et le nombre d'enfants de la personne. Cela s'explique à la fois par la situation sociale et financière obtenues à l'issue des études mais également par une prise de conscience et une appréhension différente des risques selon le niveau d'études atteint.

Notre étude statistique, bien qu'elle fournisse des résultats probants, gagnerait en profondeur si les questions étaient plus complexes et si l'échantillon était plus important. Elle reste cependant en accord avec les études antérieures de sociologues et statisticiens et dévoile surtout des résultats intéressants sur le rapport explicite mais aussi implicite entre niveau d'études et situation familiale.

# Bibliographie

- [1] Christian BAUDELLOT et François LECLERCQ. “Les Effets de l’éducation sur la formation et le fonctionnement des ménages”. In : *La Documentation française* (mar. 2005).
- [2] BOUCHET-VALAT et MILAN. “Plus diplômées, moins célibataires. L’inversion de l’hypergamie féminine au fil des cohortes en France”. In : *Population* 70.n° 4 (2006), p. 705-730.
- [3] Michel BOZON et François HÉRAN. “La formation du couple. Textes essentiels pour la sociologie de la famille”. In : *La Découverte* (2006).
- [4] Monja SCHMITT, Claudia KARWATH et Iona RELIKOWSKI. “Sibling Structure and Educational Achievement : How Do the Number of Siblings, Birth Order, and Birth Spacing Affect Childrens Vocabulary Competences?” In : *Journal of Family Research* 26 (déc. 2014), p. 374-398. DOI : 10.3224/zff.v26i3.18993.



# Annexe

## Graphiques et tableaux

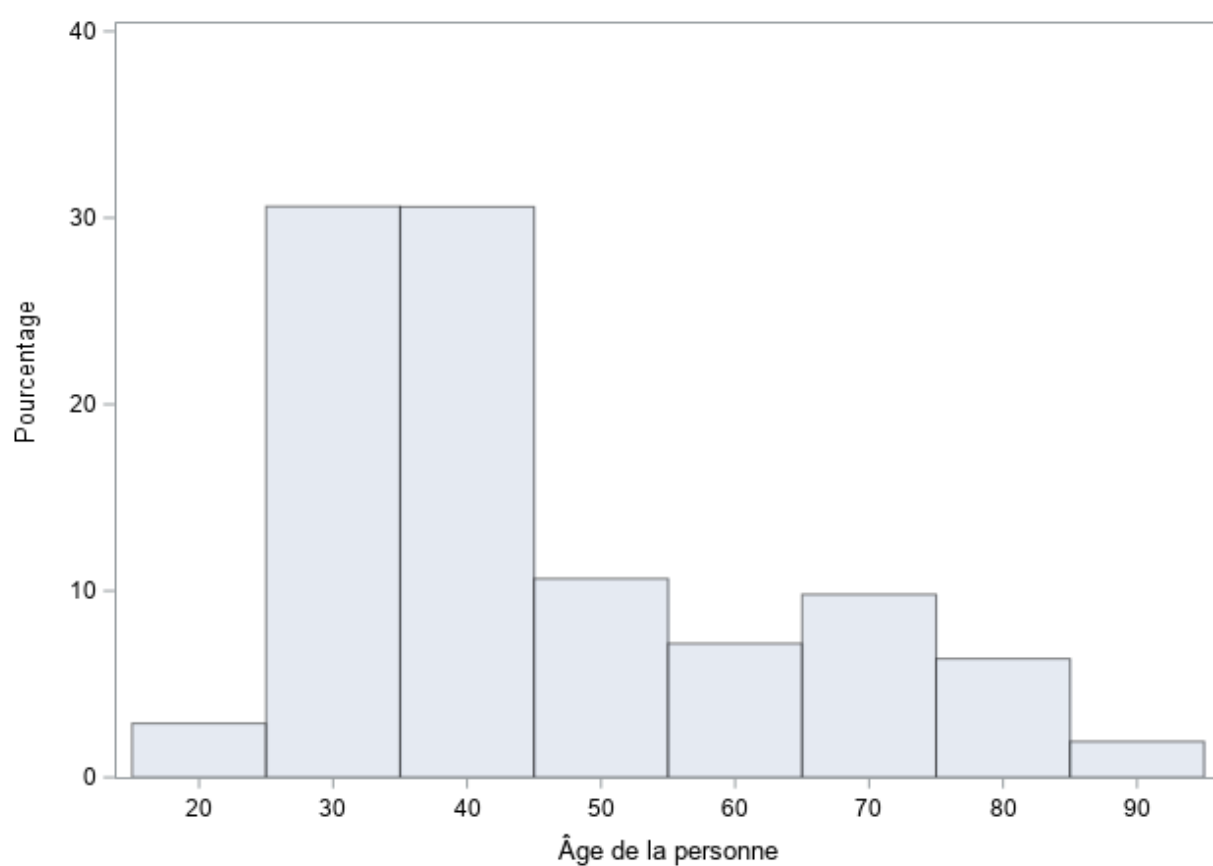


FIGURE 1 – Distribution des doctorants par tranche de 10 années

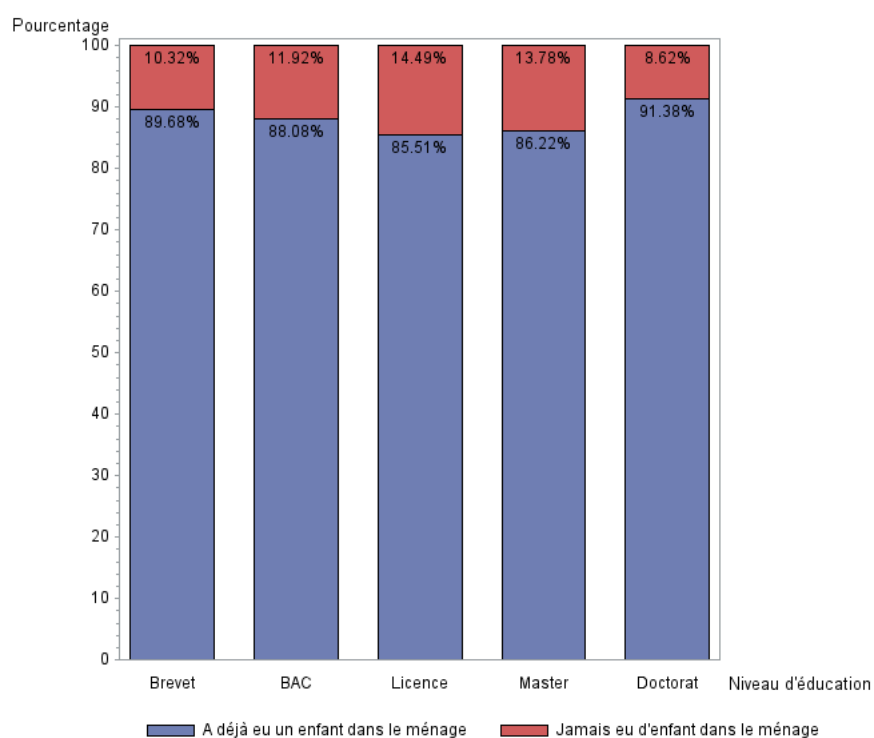


FIGURE 2 – Pourcentage de personne de plus de 50 ans ayant déjà eu un enfant dans le ménage selon le niveau d'éducation

## Code et commentaires

```
/* Gabriel WATKINSON et Mohamed GUENNOUNI HASSANI */
/* Sujet : Situation familiale et éducation */

/*~~~~~ Création de la table de travail ~~~~~*/

/*A partir du fichier .sav, on exporte la table ESS dans la work */
PROC IMPORT OUT= WORK.ess
DATAFILE= "W:\Bureau\Memoire SAS\ESS6e02_4.sav"
DBMS=SPSS REPLACE;
RUN;

/* On crée une nouvelle base ess2 dans la librairie work */
DATA ess2;
SET ess ( WHERE = ( cntry = "FR" ));
pond = pspwght * pweight * 10000;
/* On crée la variable pond qui nous sera utile par la suite
pour notre pondération */
LABEL pond ="Ponderation";
KEEP chldhhe chldhm hhmmb marsts edlvdfr edlvfdfr edlvmdfr
edlvpdfr pspwght dweight pweight pond lvgptnea agea idno;
/* On garde toutes les variables utiles pour notre étude */
RUN;

/* On sauvegarde notre table dans une librairie permanente */

LIBNAME memoire "W:\Bureau\Memoire SAS";
DATA memoire.ess2;
SET ess2;
RUN;

/*~~~~~ Recodage des variables ~~~~~*/

/*On créer la variable edperso pour le niveau d'éducation personnelle
pour regrouper les 26 modalités de edlvdfr en 5 modalités :
Brevet, Bac, Licence, Master, Doctorat */

DATA ess2;
SET ess2;
LENGTH edperso $ 20;
IF edlvfdfr <=5 THEN edperso=" Brevet";
ELSE IF edlvfdfr <=11 THEN edperso=" BAC";
ELSE IF edlvfdfr <=18 THEN edperso=" Licence";
ELSE IF edlvfdfr <=24 THEN edperso=" Master";
ELSE IF edlvfdfr <=26 THEN edperso=" Doctorat";
```

RUN;

/\*On créer la variable edpere pour le niveau d'éducation du père  
pour regrouper les 26 modalités de edlvdfdr en 5 modalités :  
Brevet , Bac , Licence , Master , Doctorat \*/

```
DATA ess2;  
SET ess2;  
LENGTH edpere $ 20;  
IF edlvdfdr <=5 THEN edpere="    Brevet";  
ELSE IF edlvdfdr <=11 THEN edpere="    BAC";  
ELSE IF edlvdfdr <=18 THEN edpere="    Licence";  
ELSE IF edlvdfdr <=24 THEN edpere="    Master";  
ELSE IF edlvdfdr <=26 THEN edpere="Doctorat";  
RUN;
```

/\*On créer la variable edmere pour le niveau d'éducation personnelle  
pour regrouper les 26 modalités de edlvdfdr en 5 modalités :  
Brevet , Bac , Licence , Master , Doctorat \*/

```
DATA ess2;  
SET ess2;  
LENGTH edmere $ 20;  
IF edlvdmfr <=5 THEN edmere="    Brevet";  
ELSE IF edlvdmfr <=11 THEN edmere="    BAC";  
ELSE IF edlvdmfr <=18 THEN edmere="    Licence";  
ELSE IF edlvdmfr <=24 THEN edmere="    Master";  
ELSE IF edlvdmfr <=26 THEN edmere="Doctorat";  
RUN;
```

/\*On créer la variable edpartenaire pour le niveau d'éducation  
personnelle pour regrouper les 26 modalités de edlvdfdr en  
5 modalités : Brevet , Bac , Licence , Master , Doctorat \*/

```
DATA ess2;  
SET ess2;  
LENGTH edpartenaire $20;  
IF edlvdpfr <=5 THEN edpartenaire="    Brevet";  
ELSE IF edlvdpfr <=11 THEN edpartenaire="    BAC";  
ELSE IF edlvdpfr <=18 THEN edpartenaire="    Licence";  
ELSE IF edlvdpfr <=24 THEN edpartenaire="    Master";  
ELSE IF edlvdpfr <=26 THEN edpartenaire="    Doctorat";  
RUN;
```

/\*On créer la variable marital pour la situation maritale de  
la personne. On regroupe les modalités de marsts pour avoir  
deux modalités : légalement célibataire ou non \*/

DATA ess2;

```

SET ess2;
IF marsts in (1,2,3,4,5) THEN marital="Non Célibataire ";
ELSE If marsts=6 THEN marital="Célibataire ";
RUN;

/*On créer la variable veccouple pour savoir si la personne a déjà
vécu en couple sans être marié. On regroupe les modalités de lvgptnea
pour avoir deux modalités :
A déjà vécu en couple sans être marié ou non*/

DATA ess2;
SET ess2;
LENGTH veccouple $80;
IF lvgptnea=1 THEN avacouple="Déjà vécu en couple sans être marié";
ELSE IF lvgptnea=2 THEN avacouple="Jamais vécu en couple sans être marié";
RUN;

/*On créer la variable enfant pour savoir si la personne a déjà eu un
enfant dans le ménage. On regroupe les modalités de chldhhe pour avoir
deux modalités : A déjà eu un enfant ou non*/

DATA ess2;
SET ess2;
LENGTH enfant $ 60;
IF chldhhe=1 THEN enfant="A déjà eu un enfant dans le ménage";
ELSE If chldhhe=2 THEN enfant="Jamais eu d'enfant dans le ménage";
RUN;

/*On créer la variable foyer pour savoir le nombre de personnes
vivant dans le foyer. On regroupe les modalités de hhmb pour avoir
5 modalités : 1, 2, 3, 4, 5 ou plus*/

DATA ess2;
SET ess2;
LENGTH foyer $ 15;
IF hhmb=1 THEN foyer = "1";
ELSE IF hhmb=2 THEN foyer = "2";
ELSE IF hhmb=3 THEN foyer = "3";
ELSE IF hhmb=4 THEN foyer = "4";
ELSE IF hhmb>=5 THEN foyer = "5 ou plus";
RUN;

/* On crée le format age pour regrouper l'âge des personnes sondées
de plus de 20 ans en tranche de 10 années */

PROC FORMAT;
VALUE age
20<-30 = "De 20 a 30 ans"
30<-40 = "De 30 a 40 ans"

```

```

40<-50 = "De 40 a 50 ans"
50<-60 = "De 50 a 60 ans"
60<-70 = "De 60 a 70 ans"
70<-80 = "De 70 a 80 ans"
80<-high = "Plus de 80 ans";
RUN;

```

```

/*~~~~~ Éducation et situation familiale ~~~~~*/

```

```

/* On trace le graphique qui représente la situation maritale
en fonction du niveau d'études atteint pour les plus de 26 ans */

```

```

PROC GCHART DATA = ess2 (WHERE = (marital ne "" AND agea >=26));
/* On ne s'intéresse qu'aux individus de plus de 26 ans */
TITLE "Situation maritale en fonction de l'éducation pour
les individus de plus de 26 ans";
AXIS1 LABEL = ("Pourcentage");
AXIS2 LABEL = ("Niveau d'éducation");
VBAR edperso /FREQ = pond RAXIS=AXIS1 MAXIS=AXIS2
/* On pondère par pond pour avoir un graphique représentatif de la
population française */
DISCRETE GROUP = edperso
SUBGROUP = marital
G100 NOZERO TYPE=PERCENT
INSIDE=SUBPCT;
RUN;

```

```

/* On réalise le test de chi-2 entre la situation maritale et
le niveau d'études */

```

```

PROC FREQ DATA =ess2
(WHERE = (marital NE "" AND edperso NE "" AND agea >=26));
TITLE "Mesure de l'association entre le niveau d'études et
le statut marital pour les individus de plus de 26 ans";
TABLES edperso*marital / CHISQ;
WEIGHT pspwght;
/* On veille à utiliser pspwght pour vérifier si l'échantillon peut
permettre de rejeter l'hypothèse d'indépendance */
RUN;

```

```

/* On réalise l'histogramme qui représente la distribution de docteurs
par rapport à des tranches d'âge de 10 années */

```

```

PROC UNIVARIATE DATA = ess2 (WHERE= (edperso="Doctorat")) VARDEF=WGT;
TITLE "Distribution des doctorants par tranche de 10 années";
VAR agea;
HISTOGRAM/MIDPOINTS = 20 TO 90 BY 10;

```

```

FREQ pond;
LABEL agea = "Âge par tranche de 10 années";
RUN;

/* On trace le graphique qui représente si la personne a déjà vécu en
couple en fonction du niveau d'études atteint pour les plus de 26 ans */

PROC GCHART DATA = ess2 (WHERE = (avacouple NE "" AND agea >=26));
/* On ne s'intéresse qu'aux individus de plus de 26 ans */
TITLE " Pourcentage de personne ayant déjà vécu en couple
sans être marié ou non en fonction de l'éducation pour les
individus de plus de 26 ans";
AXIS1 LABEL = ("Pourcentage");
AXIS2 LABEL =("Niveau d'éducation");
VBAR edperso /FREQ = pond RAXIS=axis1 MAXIS=axis2
/*On pondère par pond pour avoir un graphique représentatif
de la population française*/
DISCRETE GROUP = edperso
SUBGROUP = avacouple
G100 NOZERO TYPE=PERCENT
INSIDE=SUBPCT;
RUN;

/* On réalise le test de chi-2 entre le fait que la personne a déjà
vécu en couple et le niveau d'études */

PROC FREQ DATA =ess2
(WHERE = (avacouple NE "" AND edperso NE "" AND agea >=26));
TITLE " Mesure de l'association entre le niveau d'études et
le fait d'avoir déjà vécu en couple sans être marié ou non
pour les individus de plus de 26 ans";
TABLES edperso*avacouple /CHISQ ;
WEIGHT pspwght;
/* On veille à utiliser pspwght pour vérifier si l'échantillon peut
permettre de rejeter l'hypothèse d'indépendance */
RUN;
RUN;

/* On trace le graphique qui représente la situation maritale en fonction
du niveau d'études atteint pour les plus de 50 ans */

PROC GCHART DATA = ess2 (WHERE = (marital ne "" AND agea >=50));
/* On ne s'intéresse qu'aux individus de plus de 50 ans */
TITLE " Situation maritale en fonction de l'éducation pour
les individus de plus de 50 ans";
AXIS1 LABEL = ("Pourcentage");
AXIS2 LABEL =("Niveau d'éducation");
VBAR edperso /FREQ = pond RAXIS=AXIS1 MAXIS=AXIS2
/* On pondère par pond pour avoir un graphique représentatif

```

```

de la population française */
DISCRETE GROUP = edperso
SUBGROUP = marital
G100 NOZERO TYPE=PERCENT
INSIDE=SUBPCT;
RUN;

```

```

/*On réalise le test de chi-2 entre le fait que la personne
a déjà vécu en couple et le niveau d'études*/

```

```

PROC FREQ DATA =ess2
(WHERE = (avacouple NE "" AND edperso NE "" AND agea >=50));
TITLE "Mesure de l'association entre le niveau d'études et
le fait d'avoir déjà vécu en couple sans être marié ou non
pour les individus de plus de 50 ans";
TABLES edperso*marital / CHISQ ;
WEIGHT pspwght;
/* On veille à utiliser pspwght pour vérifier si l'échantillon
peut permettre de rejeter l'hypothèse d'indépendance */
RUN;

```

```

/*~~~ Homogamie et reproduction sociale sous le spectre de l'éducation ~~~*/

```

```

/*On réalise le tableau de corrélation de Kendall
entre edlvdfr, edlvmdfr, edlvfdfr.
Ces variables représentent éducation personnelle,
celle du père et celle de la mère */

```

```

PROC CORR DATA = ess2 (WHERE =(edlvdfr <5000 AND edlvdfr NE .
AND edlvmdfr<5000 AND edlvmdfr NE . AND edlvfdfr <5000 AND edlvfdfr NE .
AND edlvpdfr <5000 AND edlvpdfr NE .)) KENDALL;
/*On inclue dans les conditions l'éducation du partenaire pour avoir le
même échantillon et comparer avec le tableau de corrélation ci-après */
TITLE "Coefficients du tau b de Kendall";
VAR edlvdfr edlvmdfr edlvfdfr;
RUN;

```

```

/*On réalise le tableau de corrélation de Kendall
entre edlvdfr, edlvmdfr, edlvfdfr.
Ces variables représentent éducation personnelle et celle du partenaire */

```

```

PROC CORR DATA = ess2 (WHERE =(edlvdfr <5000 and edlvdfr NE .
AND edlvmdfr<5000 and edlvmdfr NE . AND edlvfdfr <5000 AND edlvfdfr NE .
AND edlvpdfr <5000 and edlvpdfr NE .)) KENDALL;
TITLE "Coefficients du tau b de Kendall";
VAR edlvdfr edlvpdfr;
RUN;

```



```
/*~~~~~ Propension à fonder un foyer ~~~~~*/
```

```
/* On trace le graphique qui représente le fait d'avoir eu un enfant dans le ménage en fonction du niveau d'études atteint pour les plus de 26 ans */
```

```
PROC GCHART DATA = ess2 (WHERE = (enfant NE "" AND agea >=26));  
/* On ne s'intéresse qu'aux individus de plus de 26 ans */  
TITLE "Pourcentage de personne de plus de 26 ans ayant déjà eu un enfant dans le ménage selon le niveau d'éducation";  
AXIS1 LABEL = ("Pourcentage");  
AXIS2 LABEL = ("Niveau d'éducation");  
VBAR edperso /FREQ = pond RAXIS=AXIS1 MAXIS=AXIS2  
/* On pondère par pond pour avoir un graphique représentatif de la population française */  
DISCRETE GROUP = edperso  
SUBGROUP = enfant  
G100 NOZERO TYPE=PERCENT  
INSIDE=SUBPCT;  
RUN;
```

```
/* On réalise le test de chi-2 entre le fait d'avoir eu un enfant dans le ménage et le niveau d'études pour les plus de 26 ans */
```

```
PROC FREQ DATA = ess2  
(WHERE = (enfant NE "" AND edperso NE "" AND agea >=26));  
TABLES edperso*enfant /CHISQ ;  
WEIGHT pspwght;  
TITLE "Mesure de l'association entre le niveau d'études et le fait d'avoir eu un enfant dans le ménage ou non pour les individus de plus de 26 ans";  
/* On veille à utiliser pspwght pour vérifier si l'échantillon peut permettre de rejeter l'hypothèse d'indépendance */  
RUN;
```

```
/* On trace le graphique qui représente le fait d'avoir eu un enfant dans le ménage en fonction du niveau d'études atteint pour les plus de 50 ans */
```

```
PROC GCHART DATA = ess2 (WHERE = (enfant NE "" AND agea >=50));  
/* On ne s'intéresse qu'aux individus de plus de 50 ans */  
TITLE "Pourcentage de personne de plus de 50 ans ayant déjà eu un enfant dans le ménage selon le niveau d'éducation";  
AXIS1 LABEL = ("Pourcentage");  
AXIS2 LABEL = ("Niveau d'éducation");  
VBAR edperso /FREQ = pond RAXIS=AXIS1 MAXIS=AXIS2  
/* On pondère par pond pour avoir un graphique représentatif de la population française */  
DISCRETE GROUP = edperso
```

```

SUBGROUP = enfant
G100 NOZERO TYPE=PERCENT
INSIDE=SUBPCT;
RUN;

```

```

/* On réalise le test de chi-2 entre le fait d'avoir eu un enfant
dans le ménage et le niveau d'études pour les plus de 50 ans */

```

```

PROC FREQ DATA =ess2
(WHERE = (enfant NE "" AND edperso NE "" AND agea >=50));
TABLES edperso*enfant /CHISQ ;
/* On veille à utiliser pspwght pour vérifier si l'échantillon
peut permettre de rejeter l'hypothèse d'indépendance */
WEIGHT pspwght;
RUN;

```

```

/* On trace le graphique qui représente le nombre de personnes
dans le foyer en fonction du niveau d'études atteint pour
les personnes entre 36 et 52 ans */

```

```

PROC GCHART DATA = ess2
(WHERE = (foyer NE "" AND edperso ne "" AND agea>=36 and agea<=52));
/* On ne s'intéresse qu'aux individus entre 36 et 52 ans */
TITLE "Nombre de personnes vivant régulièrement dans le ménage
selon le niveaud'éducation pour les individus de 36 à 52 ans";
AXIS1 LABEL = ("Pourcentage");
AXIS2 LABEL =("Niveau d'éducation");
VBAR edperso /FREQ = pond RAXIS=AXIS1 MAXIS=AXIS2
/* On pondère par pond pour avoir un graphique représentatif
de la population française */
DISCRETE GROUP = edperso
SUBGROUP = foyer
G100 NOZERO TYPE=PERCENT
INSIDE=SUBPCT;
RUN;

```

```

/* On réalise le test de chi-2 entre le nombre de personnes dans
le ménage et le niveau d'études pour les personnes
entre 36 et 52 ans */

```

```

PROC FREQ DATA =ess2
(WHERE = (foyer NE "" AND edperso NE "" AND agea>=36 AND agea<=52));
TITLE "Mesure de l'association entre le niveau d'études et nombre de
personnes vivant régulièrement dans le ménage pour les individus
de 36 à 52 ans";
TABLES edperso*foyer /CHISQ;
WEIGHT pspwght;
RUN;

```