



Il en faut deux pour être heureux

Mémoire de Statistiques descriptives sur le lien entre nombre
d'enfants et bonheur

Auteurs :

Pauline CULIOLI
Jérôme PIVARD
Rémi TURQUIER

Référent :

Maël GINSBURGER

Janvier 2020

Résumé

Ce mémoire étudie le lien entre le nombre d'enfants d'une personne en couple et son bonheur. À partir des données de la vague 6 de l'European Social Survey, nous étudions les réponses de 1090 personnes habitant avec leur conjoint en France. Nous trouvons un lien non-linéaire entre le bonheur déclaré et le nombre d'enfants ($\chi^2 = 25$, $p = 0.003$). Les personnes déclarant le plus de bonheur étant celles qui ont deux enfants. Cela ne permet pas de dire qu'avoir deux enfants permet d'être plus heureux, car notre échantillon a sans doute un biais d'auto-sélection. Lorsque nous construisons un indicateur de bonheur prenant aussi en compte la satisfaction par rapport à la vie au quotidien, et non seulement en général, ce lien disparaît. Nous trouvons aussi que le nombre d'enfants ne semble pas influencer le sentiment d'être bien entouré. Nous étudions finalement la corrélation en considérant des variables confondantes telles que le revenu, l'âge de la personne interrogée et le genre. L'utilisation de méthodes statistiques plus avancées permettrait d'obtenir des résultats plus rigoureux.

Table des matières

1	Choix et construction des variables	3
1.1	Délimitation de la population, construction de la variable nombre d'enfants et profil de la population étudiée	3
1.2	Autres variables utilisées pour étudier la corrélation	4
2	Etude de la corrélation entre bonheur déclaré et nombre d'enfants par une analyse bivariée	6
2.1	Représentation graphique du tri croisé entre niveau de bonheur et nombre d'enfants	6
2.2	Statistiques bivariées sur les deux variables	7
2.3	Conclusions et interprétations	7
3	Score synthétique de bonheur	9
3.1	Construction du score	9
3.2	Résultats et interprétations	9
4	Pas de corrélation entre le nombre d'enfants et le sentiment d'être bien entouré	11
5	Etude de variables confondantes pour mieux comprendre la corrélation	12
5.1	Le revenu	12
5.2	L'âge de la personne interrogée	14
5.3	Le genre	14
6	Conclusion	16
	Annexes	19
	Annexe 1 : Variables utilisées	19
	Annexe 2 : Tableau et graphiques complémentaires	21
	Annexe 3 : Script SAS	24

Introduction

Selon un sondage TNS Sofres-Philosophie Magazine réalisé en 2009 auprès de 1 000 hommes et femmes, si 69% des Français affirment faire des enfants par devoir, 73% déclarent le faire par plaisir. La parentalité serait une expérience unique qui rendrait la vie plus belle, notamment par le bonheur de la transmission. De plus, la sociabilité familiale permet de ne pas se sentir seul, en particulier lorsque l'on vieillit. Cependant, la parentalité représente également un poids financier important, une charge de travail supplémentaire, moins de liberté... Quel est donc le lien entre famille et bonheur ?

On définit une famille comme un ensemble de personnes apparentées considérées comme socialement proches. Entre les membres d'une famille, les relations sont fréquentes, et prennent des formes multiples (Bengtson 2001 [4], Bonvalet 2006 [5]). Elles comprennent des transferts monétaires, des échanges en nature, ainsi que des relations affectives et du soutien psychologique. A priori, pour une famille, habiter sous un même toit augmente la fréquence des interactions. Nous pouvons alors émettre l'hypothèse que le nombre de membres d'une famille qui vivent dans un même ménage influence significativement leur bonheur, c'est-à-dire leur satisfaction par rapport à leur vie en général.

La question du lien entre structure familiale et bonheur a déjà été abordée dans un certain nombre d'articles, issus le plus souvent de travaux de chercheurs américains en économie, en sociologie et en psychologie (eg. Barnett et Baruch, 1985 ; Nomaguchi et Milkie, 2001). Ces recherches ont déjà mis en évidence plusieurs aspects importants du lien entre le fait d'avoir des enfants et le sentiment de bonheur.

Ainsi, Nomaguchi et Milkie (2001) [9] ont attiré l'attention sur l'existence d'un "Parental happiness gap" : les parents seraient moins heureux que les autres adultes du fait des responsabilités qui leur reviennent et des tâches qui découlent de leur statut, contribuant à réduire leur temps libre. De leur côté, Barnett et Baruch (1985) [2] ont souligné que la forte implication de certaines femmes dans leur vie familiale et au travail, et que la multiplicité des rôles à assumer était pour elles une source potentielle d'anxiété et de surcharge mentale et psychologique. On serait tenté d'en conclure qu'avoir une famille de grande taille pourrait réduire le bonheur.

Cependant, tous les articles de recherche ne vont pas dans le même sens et certains ont parfois abouti à des conclusions diamétralement opposées. Cela témoigne de la complexité de la question et de l'ambivalence de ce lien, sans doute en raison de la pluralité des effets et conséquences de la parentalité. Par exemple, dans une étude plus récente, Herbst et Ifcher (2016) [6] affirment au contraire que les parents sont plus heureux que les adultes sans enfants, quelque soit l'âge des enfants, et qu'ils sont de plus en plus heureux à mesure que leurs enfants grandissent. De plus, il n'y aurait pas de nombre maximal d'enfants au-delà duquel le supplément de bonheur résultant de la présence d'un enfant en plus deviendrait négatif : plus on aurait d'enfants, plus on serait heureux, ce qui est à l'opposé des conclusions des travaux que nous avons mentionné ci-dessus !

Dans ce contexte, on peut penser qu'une approche plus quantitative, fondée sur les données d'une enquête statistique, l'European Social Survey, couplée à des interprétations sociologiques et/ou économiques, est susceptible de contribuer à dévoiler ce qui se joue vraiment dans le lien entre taille du ménage et bonheur des individus

Nous restreignons notre champ d'étude à la France métropolitaine. Nous utilisons les données issues de la vague 6 de l'European Social Survey (ESS), menée en 2012. Il s'agit d'une enquête menée dans 29 pays européens, pour laquelle les données ont été recueillies par questionnaire sur un échantillon représentatif des individus de 15 ans et plus. Ces individus ont

été sélectionnés aléatoirement.

En particulier, il paraît pertinent de tester statistiquement le lien entre famille et bonheur en se focalisant sur le nombre d'enfants d'un couple, dont dépend fortement la structure du ménage. On définit "enfant" par deux conditions : avoir moins de 16 ans, et être le fils ou la fille de la personne de référence du ménage. Ceci nous amène à poser la question suivante : Y a-t-il une corrélation entre le nombre d'enfants d'une personne en couple et son bonheur ?

Dans une première partie, nous revenons sur notre choix et notre construction de variables mesurant le nombre d'enfants et le niveau de bonheur exprimé, notamment en présentant la population étudiée. Ensuite, nous étudions l'existence d'un lien entre le niveau de bonheur et la structure du ménage en nous appuyant à la fois sur une première analyse bivariable des variables retenues. Dans notre troisième partie, nous constatons que ce lien semble disparaître si l'on prend aussi en compte des mesures moins directes du bonheur. Puis nous créons un indicateur du sentiment d'être bien entouré pour essayer de voir si une explication au bonheur apporté par la parentalité réside dans le fait de se sentir bien entouré. Dans un dernier temps, nous essayons de prendre en compte trois autres variables permettant de raffiner l'analyse de la corrélation, à savoir le revenu du ménage, l'âge de la première personne du ménage et le genre. En effet, ces variables sont susceptibles de jouer le rôle de variables confondantes, i.e. jouant à la fois sur la variable expliquée (le niveau de bonheur) et sur la variable explicative (le nombre d'enfants).

1 Choix et construction des variables

Commençons par nous intéresser aux variables qui vont nous permettre d’explorer l’existence d’un lien entre le niveau de bonheur et le nombre d’enfants d’un couple.

La liste complète des variables que nous avons retenues est disponible en annexe.

1.1 Délimitation de la population, construction de la variable nombre d’enfants et profil de la population étudiée

Tout d’abord, se poser la question de l’existence d’une corrélation entre niveau de bonheur exprimé et nombre d’enfants implique de retenir une variable expliquée, à savoir le niveau de bonheur déclaré par les personnes interrogées.

Nous nous appuyons donc dans un premier temps sur une variable nommée “happy”, qui regroupe les réponses à la question “A quel point êtes-vous heureux ?” sur une échelle de 1 à 10.

Nous retenons également une variable explicative principale, à savoir le nombre d’enfants de la personne interrogée. Pour cela, nous utilisons les variables “yrbrn” et les variables “rshipa”. Les premières nous permettent de compter le nombre d’individus de moins de 16 ans, et ayant un lien de filiation avec la première personne. Les variables “rshipa” représentent le lien de la personne avec la première personne du ménage. Nous ne comptons que les personnes qui sont des enfants de la première personne du ménage, et pas les enfants qui ne sont pas ceux de la personne de référence. Nous restreignons notre étude aux personnes en couple, en retenant la variable “icpart1”. On peut supposer que ne pas le faire aurait amené à devoir prendre en compte des situations familiales (plus) complexes, qui sont de nature à pouvoir influencer elles aussi sur le niveau de bonheur, et donc potentiellement à brouiller la corrélation étudiée.

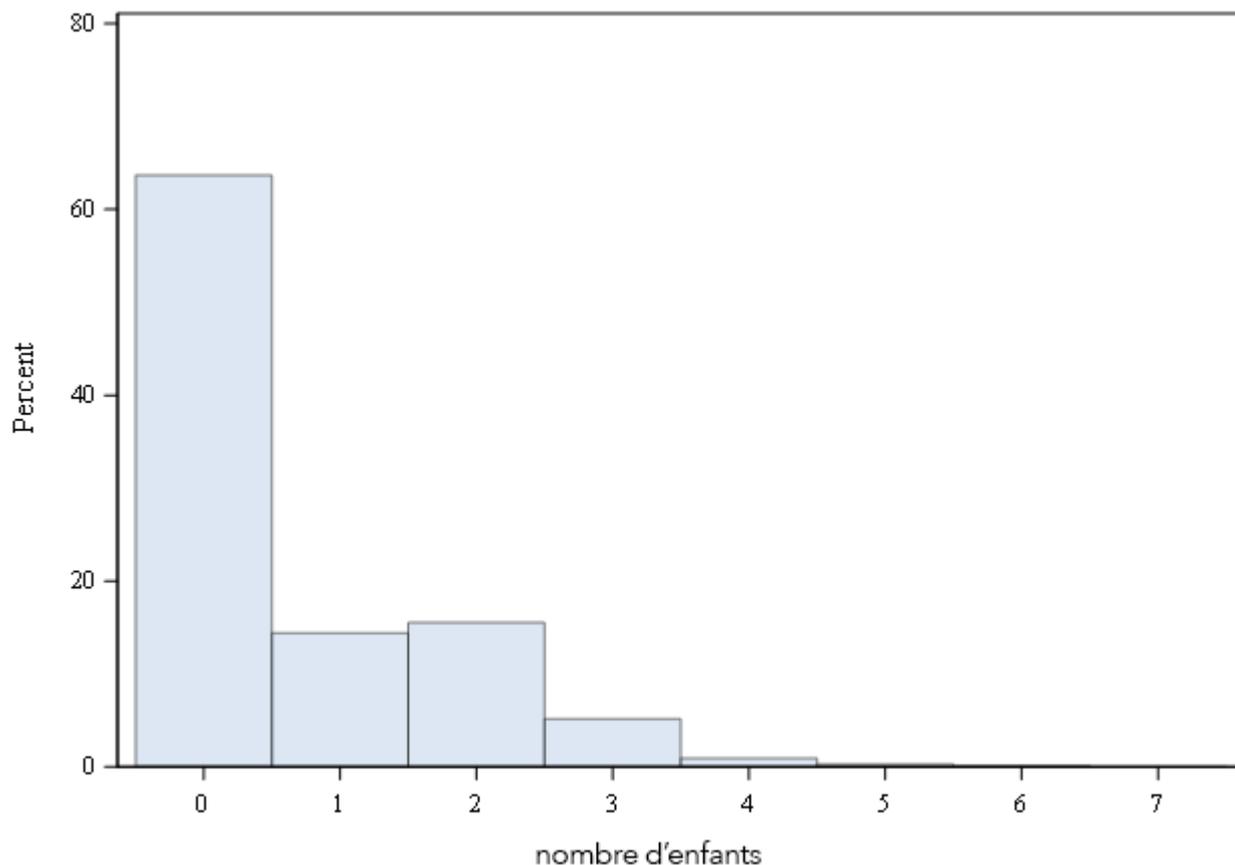
A partir de ces deux catégories de variables, nous avons créé la variable codant le nombre d’enfants “nbenfants”, comme somme de variables indicatrices ; “enfant(i)” prend la valeur 1 si la n-ième personne du ménage est un enfant , 0 sinon ; puis on somme ces variables pour obtenir “nbenfants”.

A partir de cette variable, il est possible de commencer à esquisser le profil de la population étudiée grâce à une analyse univariée (voir histogramme ci-dessous).

Plus de 60% des individus en couple interrogés lors de l’enquête n’ont pas d’enfant, et une proportion significative des répondants ont entre un et trois enfants, ce qui nous permet d’avoir assez de données pour travailler sur une éventuelle corrélation entre le nombre d’enfants et le bonheur exprimé. La haute proportion de couples sans enfants est sans doute liée au fait que nous avons inclus tous les couples, y compris ceux qui n’ont pas l’âge d’avoir des enfants qui habitent avec eux.

Un travail similaire sur la variable quantitative “happyqti”, issue du recodage de “happy”, permet d’observer qu’une majorité des individus se déclare heureuse voire très heureuse, mais qu’il y a un nombre non négligeable d’individus se déclarant peu heureux.

Nombre d'enfants des personnes en couple interrogées



Source : Enquête ESS 2012 vague 6

Champ : Personnes françaises de plus de 15 ans vivant dans un ménage avec un conjoint

Lecture : Environ 15% des personnes en couple interrogées ont un seul enfant.

1.2 Autres variables utilisées pour étudier la corrélation

Ainsi, il paraît légitime de procéder à un découpage de cette population en différentes modalités, que ce soit pour le niveau de bonheur ou pour le nombre d'enfants, en essayant d'obtenir des classes d'effectifs de taille à peu près équivalente lorsque c'est possible. Cela permettra de gagner à la fois en lisibilité et en robustesse sur le plan statistique. On obtient ainsi deux variables qualitatives ordonnées : "nbenfants2", prenant 4 modalités (pas d'enfant, un enfant, 2 enfants, 3 enfants et plus), et "happy2", prenant elle-aussi 4 modalités (bonheur exprimé sur 10 : 0-2, 3-5, 6-8, 9-10).

Les variables obtenues sont celles sur lesquelles nous avons procédé à une analyse bivariée, que nous abordons dans la partie suivante.

Par ailleurs, nous avons aussi retenu des variables qui nous permettent de construire un indice de bonheur exprimé et un indicateur du sentiment d'être bien entouré. La construction et l'analyse de ces indices sont l'objet de nos troisième et quatrième parties.

Enfin, nous avons retenu un dernier ensemble de variables permettant d'approfondir dans un dernier temps l'analyse et l'interprétation de la corrélation entre le nombre d'enfants et le bonheur exprimé. Chacune de ces variables a la particularité d'influer potentiellement à la fois sur le niveau de bonheur exprimé, et sur le nombre d'enfants.

Premièrement, la variable "hinctnta", qui représente le décile de revenu net du ménage, nous permet d'essayer de contrôler l'effet du revenu sur le niveau de bonheur exprimé. Ce dernier est peut-être plus fort et n'agit pas forcément dans le même sens que celui du nombre d'en-

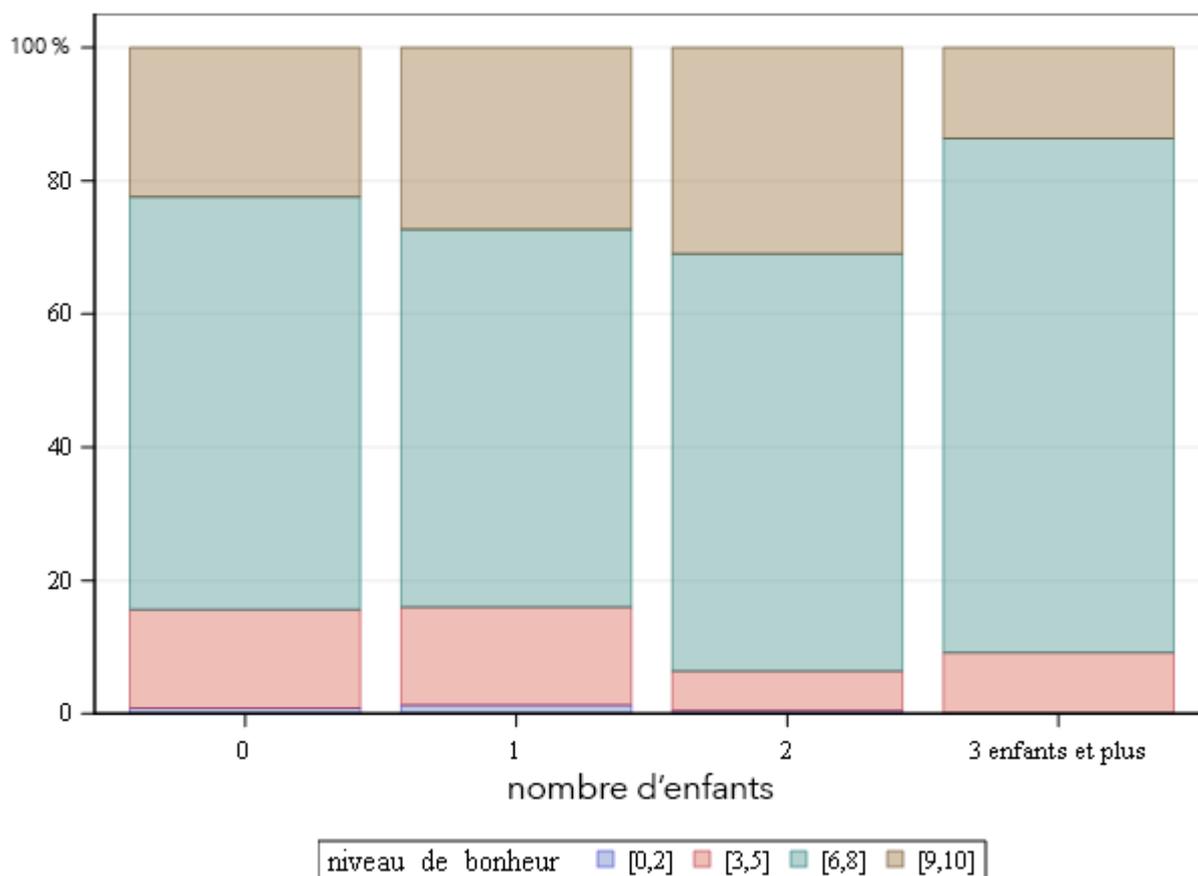
fants, et il est susceptible d'influer sur la corrélation explorée. De plus, il peut y avoir un biais de structure lié au fait que les individus ayant le plus d'enfants ont aussi tendance à être ceux disposant de revenus plus élevés, d'où la nécessité de prendre en compte cette troisième variable.

Ensuite, prendre en compte la variable "agea" (âge de la personne interrogée) est un moyen de réfléchir à la question suivante : l'âge de la personne en question agit-il sur la corrélation étudiée et sur son intensité ? Autrement dit, est-ce que le fait d'habiter avec des enfants nous rend aussi heureux lorsque nous avons 30 ans que lorsque nous avons 60 ans ?

Finalement, l'une des questions les plus importantes lorsqu'on traite du bonheur apporté par la famille est peut-être celle du genre de la personne interrogée, d'où l'intérêt de la variable "gndr" : avoir des enfants, et en avoir beaucoup plutôt qu'un seul, apporte-t-il autant de bonheur à une mère qu'à un père, (ou davantage à une mère qu'à un adulte cohabitant, dans le cas où l'homme du couple n'est pas le père de l'enfant) ? La question de la différence des rôles sociaux qui découlent de la maternité et de la paternité nous semble être particulièrement importante pour mieux comprendre la corrélation étudiée.

2 Etude de la corrélation entre bonheur déclaré et nombre d'enfants par une analyse bivariée

2.1 Représentation graphique du tri croisé entre niveau de bonheur et nombre d'enfants



Source : Enquête ESS 2012 vague 6

Champ : Personnes françaises de plus de 15 ans vivant dans un ménage avec un conjoint

Lecture : Environ 60% des personnes n'ayant pas d'enfants déclarent un niveau de bonheur compris entre 6 et 8.

Pour rendre l'analyse plus parlante et les résultats plus lisibles, nous avons créé les nouvelles variables "happy2" et "nbenfants2" qui regroupent respectivement les différents niveaux de bonheur exprimés et les différents nombres d'enfants en quatre modalités.

La proportion de personnes exprimant le niveau de bonheur maximal (en marron clair) augmente en général avec le nombre d'enfants (20% des sans enfants contre 30% des personnes ayant deux enfants). Au contraire, la proportion de ceux exprimant un niveau de bonheur compris entre trois et cinq diminue quand le nombre d'enfants augmente (en rose).

Ainsi, une corrélation positive entre nombre d'enfants et bonheur exprimé semble se dessiner. Une enquête récente de l'INSEE[11] confirme et nuance à la fois ce résultat.

D'une part, elle montre que la satisfaction dans la vie des personnes interrogées varie en fonction du nombre d'enfants. D'autre part, elle n'aboutit pas exactement à la même conclusion :

en 2017, seulement 10% des personnes seules déclarent que leur niveau de satisfaction est élevé, contre 20% des couples sans enfants et 28% des couples avec trois enfants. Cette relation n'est cependant pas linéaire selon nos résultats : elle n'est plus vraie à partir de trois enfants. On peut faire l'hypothèse qu'à partir de trois enfants, la charge de travail liée aux obligations parentales ou le poids financier compense ou atténue la hausse du bonheur qui semble accompagner la parentalité. Plus généralement, il peut s'agir de tout un ensemble de contraintes matérielles liées à la parentalité et se trouvant renforcées dans une famille plus nombreuse (à partir de trois enfants) qui expliqueraient pourquoi les individus ayant plus de deux enfants seraient tendanciellement moins heureux que ceux n'en ayant que deux.

2.2 Statistiques bivariées sur les deux variables

On obtient un khi-deux de 25, avec une p-valeur de 0,0033 (voir tableau 1 en annexe). On peut donc rejeter l'hypothèse d'indépendance au seuil de 99%. Le lien entre le nombre d'enfants et le bonheur déclaré provient principalement des personnes qui ont deux enfants et qui déclarent un bonheur entre 3 et 5 (khi-deux de cellule : 6,89) et des personnes qui ont deux enfants ou plus et un bonheur de 9 ou plus. Cela semble suggérer que l'association entre les deux variables est relativement faible certes, mais qu'elle existe et est assez fiable statistiquement.

Comme le tau b de Kendall est plus de deux fois plus grand que son écart type (0.06 et 0.02), on peut supposer qu'il est statistiquement significatif au seuil de 5%.

Il est plus robuste que le coefficient de corrélation linéaire de Pearson qui suggérerait déjà l'existence d'une très faible corrélation positive entre nombre d'enfants et niveau de bonheur. Dans notre cas, il vaut 0.06 , ce qui signifie que les paires d'observations ont une très légère tendance à être plutôt concordantes (Kendall positif). Néanmoins, la proximité relative du coefficient par rapport à 0 doit nous conduire à être prudents : s'il y a une liaison statistique significative entre les deux variables (ce que semble suggérer le test du khi-deux), plus précisément une corrélation positive entre le niveau de bonheur et le nombre d'enfants, celle-ci est certainement d'une ampleur assez faible si l'on se fonde sur ces résultats.

2.3 Conclusions et interprétations

L'analyse bivariée suggère ainsi qu'il y a une relation positive entre le nombre d'enfants et le bonheur exprimé, d'intensité assez faible certes, mais statistiquement robuste. On peut trouver plusieurs explications à cette corrélation.

La première est l'effet d'auto-sélection. Comme l'écrit Tanya Basu [3] : "Les parents deviennent probablement parents plus parce qu'ils veulent l'être qu'à cause de la pression sociale. Ces adultes sont plus susceptibles d'être un groupe auto-sélectionné, de désirer leurs enfants, et donc de tirer plus de bonheur du fait d'avoir les enfants qu'ils voulaient." Nous pouvons ensuite envisager qu'avoir des enfants permet de se sentir mieux entouré, et donc d'être plus heureux. Nous testons cette hypothèse dans la quatrième partie de notre mémoire.

Une troisième explication peut être que les parents tirent de la réussite sociale de leurs enfants une reconnaissance aux yeux des autres au sens où cette réussite leur incombe en partie ; ils ont bien élevé leurs enfants, leur ont offert les moyens de réussir et, s'ils ne dégagent pas de leur investissement éducatif un profit matériel, ils en perçoivent néanmoins des rétributions symboliques sous la forme de fierté et reconnaissance. De plus, comme le relèvent Herbst et Ifcher [6], "les parents ont tendance à vouloir passer plus de temps avec leurs amis, à s'informer, à faire plus confiance à autrui. [Ils] restent de fait plus attachés à la société(...)".

En outre, plusieurs enquêtes ont illustré un refus de l'enfant unique assez répandu et en corollaire une préférence pour un nombre d'enfants égal à deux. Selon une enquête menée par l'INED et l'INSEE [7] citée par Régnier-Loilier [10], en France, moins de 3% des femmes âgées de 15 à 44 ans estiment idéal de n'avoir qu'un seul enfant, 41% deux enfants, 11% entre deux et trois, et 30% trois enfants, proportions qui se comprennent à la lumière des représentations que se font la plupart des femmes et des hommes de la famille ; celle-ci doit compter au moins deux enfants. Soulignons qu'au vu de nos résultats, cela semble relever davantage d'un idéal que de la réalité concrète : dans les faits, il y a beaucoup de familles où l'enfant est enfant unique.

Avoir plusieurs enfants permettrait "d'augmenter la diversité" [10], que chacun des conjoints se retrouve sous certains traits dans ses enfants. En avoir un deuxième, ce peut être aussi un moyen de se donner l'expérience possible de l'autre sexe et de parvenir à un équilibre familial, harmonie conférée par une sorte de symétrie dans la famille : deux parents et deux enfants, signe de conformité sociale et source potentielle de plaisir intergénérationnel (Bourguignon, 1987). Comme l'écrit Régnier-Loilier [10], "les raisons d'offrir un(e) petit(e) frère/soeur à un premier enfant relèvent également de motivations altruistes, orientées dans intérêt de l'enfant : crainte que l'enfant se sente seul, qu'il se trouve au centre de toutes les attentions et ainsi investi des attentes de ses parents alimente le désir d'élargir la famille". Les notions de partage ou de complicité sont fréquemment évoquées pour justifier le refus de l'enfant unique.

Cependant, le tri croisé que nous avons effectué semble indiquer que cette relation ne serait valable que jusqu'à deux enfants, nombre au-delà duquel la satisfaction moyenne aurait tendance à légèrement baisser (mais à rester supérieure que chez les personnes sans enfant). On peut penser que ce renversement est possiblement lié au poids de certaines contraintes matérielles telles que de nouvelles dépenses qui rendent la parentalité plus contraignante dans une famille nombreuse, d'où une possible baisse de la satisfaction exprimée. A titre d'exemple, citons la taille du logement ou celle de la voiture. S'occuper d'une famille où il y a plus de deux enfants impliquerait pour les parents une surcharge de travail pouvant réduire le bonheur déclaré.

3 Score synthétique de bonheur

3.1 Construction du score

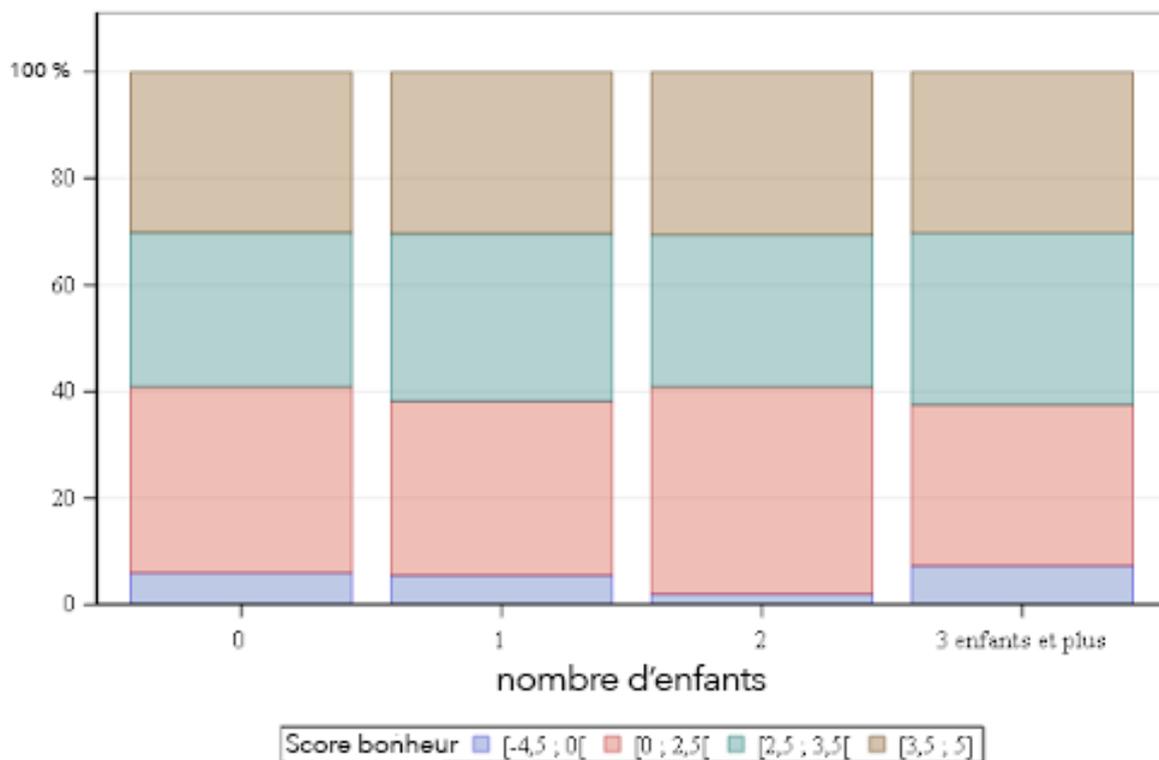
Mesurer le bonheur n'est pas évident. Plusieurs problèmes se posent. Premièrement, tout le monde ne s'accorde pas sur sa définition. Deuxièmement, être satisfait de sa vie en général ne signifie pas que l'on est satisfait au quotidien. On peut ainsi se sentir triste souvent, et tout de même se considérer heureux. Troisièmement, les réponses à la question "à quel point êtes-vous heureux?" risquent d'être biaisées par un effet de désirabilité sociale. Puisqu'il est socialement valorisé de se considérer heureux, les personnes interrogées peuvent avoir tendance à se déclarer plus heureuses qu'elles ne se considèrent vraiment.

Pour tenter de dépasser ces problèmes, nous construisons un score de bonheur qui prend aussi en compte la réponse à neuf autres questions. Ce score, que nous appelons score1, est construit en deux parties. L'une compte positivement, l'autre négativement. Chacune agrège les réponses à cinq questions. Le détail de la construction du score est résumé par un tableau en annexe. Dans les variables positives, "happy" est celle à laquelle nous avons donné le plus de poids, car elle est directement la réponse à la question "à quel point êtes vous heureux?". Les variables "a apprécié la vie la semaine dernière" et "a été heureux.se la semaine dernière" nous semblent très proches de la définition du bonheur. Les deux autres variables positives, "a généralement une très bonne opinion de soi" et "enthousiaste à propos de ce que je fais" sont moins immédiatement liées au bonheur. Parmi les cinq variables que nous avons compté négativement, celle à laquelle nous avons donné le plus de poids est "s'est senti.e déprimé.e la semaine dernière". "S'est senti.e triste la semaine dernière" et "a parfois le sentiment d'être un échec" nous semblent un peu moins fortes. Nous considérons que "ressentir tout ce qu'on fait comme un effort" et "se sentir anxieux.se" sont des obstacles moins immédiats au bonheur. Théoriquement, le score de bonheur peut prendre des valeurs entre -10 et 10. Une fois construit, la valeur minimale est -4, et la valeur maximale est 5. La médiane vaut 2,8. La majorité des individus ont un score positif (voir graphique ci-dessous).

3.2 Résultats et interprétations

L'indicateur de bonheur que nous avons construit ne mesure pas la même chose que la question "à quel point êtes vous heureux", mais ces deux variables sont fortement et significativement corrélées ($r = 0.59$, $p < 0.0001$). Comme nous avons trouvé une corrélation entre le bonheur déclaré et le nombre d'enfants, on pourrait s'attendre à aussi en trouver notre nouvel indicateur de bonheur. Pourtant, le coefficient de corrélation linéaire entre le score de bonheur et le nombre d'enfants est très faible ($r = 0,035$), et la p-valeur associée est très élevée ($p = 0.25$). Il n'y a pas de corrélation entre ce score de bonheur et le nombre d'enfants. Cela se voit bien sur le graphique ci-dessous, à comparer avec le graphique en page 6.

Niveau de bonheur codé par notre score en fonction du nombre d'enfants



Source : Enquête ESS 2012 vague 6

Champ : Personnes françaises de plus de 15 ans vivant dans un ménage avec un conjoint

Lecture : 30% des personnes n'ayant pas d'enfants ont un score de bonheur compris entre 3,5 et 5.

Nous pouvons envisager deux manières d'expliquer que la corrélation n'est pas transitive. Les différentes variables que nous avons agrégées sont différentes, et sont déterminées par des éléments différents. Pour que la corrélation soit conservée, il faudrait que ces différentes variables aient chacune une corrélation aussi forte que notre première mesure du bonheur. A priori, construire un score synthétique risque d'affaiblir une corrélation. La deuxième explication est la suivante : peut-être que les parents de deux enfants se considèrent plus heureux en général, mais qu'il ne sont pas plus heureux que les autres au quotidien, ni moins anxieux, tristes et déprimés que les autres au quotidien.

4 Pas de corrélation entre le nombre d'enfants et le sentiment d'être bien entouré

Dans cette partie, nous cherchons à tester une possible explication au lien entre bonheur déclaré et nombre d'enfants que nous avons trouvé dans notre première partie. Cette hypothèse explicative est la suivante : habiter avec ses enfants permettrait de se sentir mieux entouré, et donc d'être plus heureux.

Pour la tester, nous créons un indice pour mesurer le sentiment d'être bien entouré. Cet indice est construit comme la moyenne de trois variables : le sentiment de solitude la semaine passée, le fait de recevoir de l'aide et du soutien de ses proches, et le fait de se sentir apprécié par son entourage (voir annexe). L'indice que nous obtenons prend des valeurs entre 0 et 10. 75% des individus ont un score supérieur à 7,9.

Nous trouvons une corrélation entre notre indicateur du sentiment d'être bien entouré et le bonheur déclaré ($r = 0,34$, $p < 0,0001$). En revanche, nous ne trouvons pas de lien entre le nombre d'enfants et notre indicateur ($r = -0,02$, $p > 0,5$). Le nombre d'enfants n'influence visiblement pas le fait de se sentir moins seul, aidé par ses proches et apprécié de son entourage.

5 Etude de variables confondantes pour mieux comprendre la corrélation

Les variables confondantes sont des variables qui jouent à la fois sur la variable expliquée (ici le niveau de bonheur), et sur la variable explicative (le nombre d'enfants). Ainsi, elles peuvent potentiellement influencer la corrélation que l'on étudie. Ne pas les prendre en compte signifierait tirer des conclusions hâtives voire erronées, et conduirait à ignorer des réalités fondamentales de la vie sociale et familiale. Nous nous intéresserons ici au revenu, au genre et à l'âge. Il faut noter que si notre échantillon présente presque 50% d'hommes et 50% de femmes, le premier décile de revenu est sous-représenté (environ 5%), et que la moyenne d'âge est élevée (51,6 ans contre 41,2 ans pour la population française en 2016 d'après l'INSEE).

5.1 Le revenu

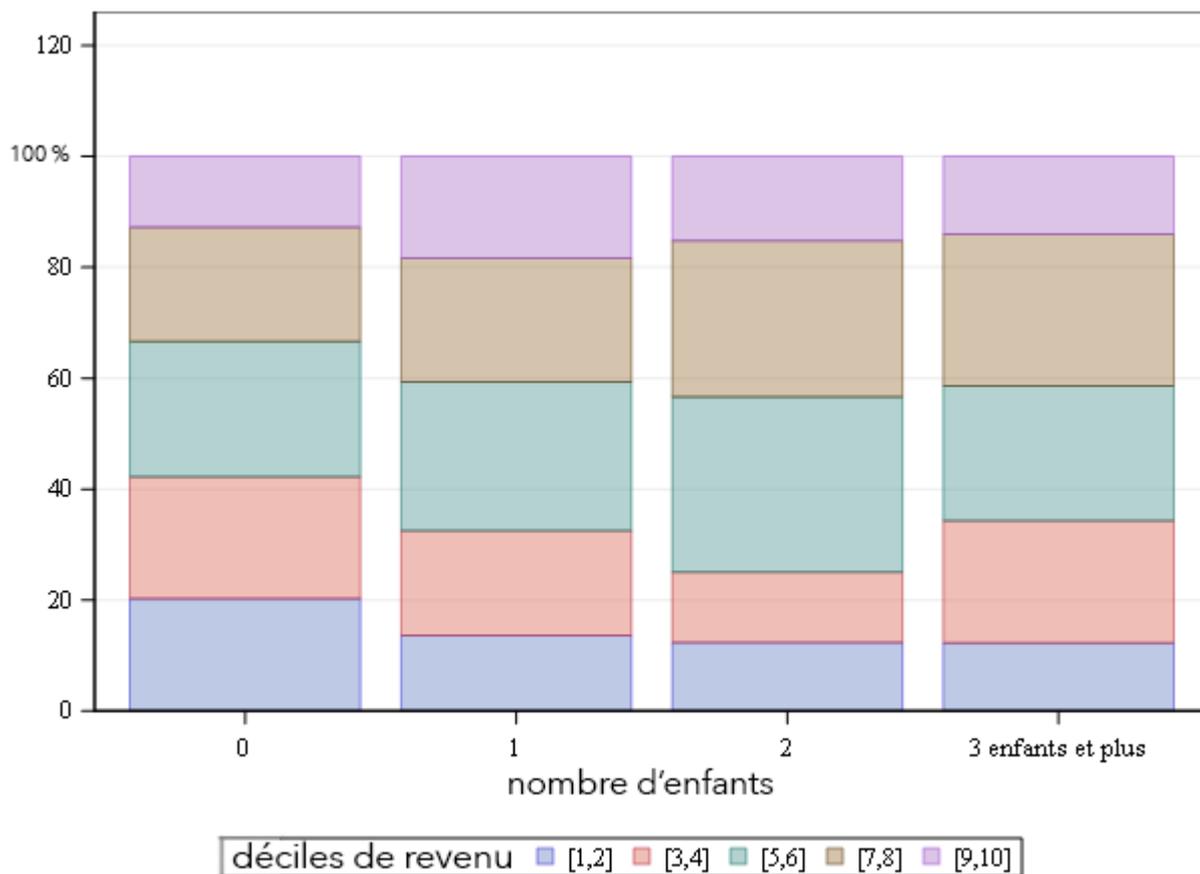
Si, comme le dit le célèbre adage, "l'argent ne fait pas le bonheur", il semble raisonnable de penser que les individus ayant une situation financière et matérielle plus favorable ont plus de chances d'exprimer une satisfaction générale à propos de leur existence que les individus les moins aisés. Nous vérifions cette hypothèse en étudiant la corrélation entre la variable "hinctnta" codant les différents déciles de revenu, et le niveau de bonheur exprimé par "happyqti". On obtient un coefficient de corrélation linéaire d'environ 0.14 et une faible P-value : il y a bien une corrélation positive entre revenu et bonheur exprimé. La question est de savoir à quel point elle influe sur la corrélation que l'on étudie.

Parallèlement, le niveau de revenus est également susceptible d'influer positivement sur le nombre d'enfants, étant donné notamment le poids des contraintes matérielles dans la parentalité. Un tri croisé entre la variable codant le revenu et la variable "nbenfants2" permet de corroborer cette hypothèse (voir graphique ci-dessous). Afin d'affiner cette observation et de confirmer nos conclusions, on recode la variable de revenu en cinq modalités, ce qui nous permet de réaliser un test d'indépendance entre les deux variables. Le test du khi-deux, et en particulier la faiblesse de la P-valeur (0.016) permettent de rejeter l'hypothèse d'indépendance entre revenu et bonheur au seuil de 5%.

Plusieurs conclusions semblent émerger du tri croisé entre la variable de revenus recodée et le nombre d'enfants. Premièrement, on note une légère surreprésentation des plus faibles niveaux de revenus dans les personnes sans enfant (20% contre environ 15% chez les personnes ayant au moins un enfant). Les 40% les moins aisés sont plus nettement sur-représentés parmi les individus n'ayant pas d'enfants (plus de 40% personnes des sans enfant contre à peine plus de 25% des personnes ayant deux enfants). Symétriquement, les plus aisés sont donc sur-représentés dans les personnes ayant deux enfants. Pour les personnes ayant trois enfants ou plus, cette différence est moins nette. Une des conclusions que l'on peut donc en tirer est la mise en évidence du coût économique lié au fait d'avoir des enfants.

Néanmoins, au vu des deux dernières analyses, une question subsiste alors : les plus riches sont-ils plus heureux en moyenne parce qu'ils ont plus d'enfants ou parce qu'ils sont plus riches ? Pour essayer de commencer à répondre à cette question, nous proposons de refaire un test de corrélation entre nombre d'enfants et bonheur exprimé, mais en prenant la variable revenu comme variable contrôlée. Concrètement, on refait deux tests de corrélation entre "nbenfants2" et "happyqti" : dans le premier, on ne considère que les six premiers déciles (les 60% les moins aisés de la population) ; dans le second, les six derniers (les 60% les plus aisés).

Déciles de revenu en fonction du nombre d'enfants



Source : Enquête ESS 2012, vague 6

Champ : Personnes françaises de plus de 15 ans vivant dans un ménage avec un conjoint.

Lecture : 20% des individus n'ayant pas d'enfant se trouvent dans les deux premiers déciles de revenu, alors que c'est le cas d'à peine plus de 10% des personnes ayant au moins un enfant.

Cette construction est faite de manière à avoir assez d'individus pris en compte (ce ne serait pas le cas en se limitant à deux déciles par exemple), tout en essayant de distinguer la situation des plus riches et celle des plus modestes. La faiblesse des effectifs ne nous permet plus d'interpréter les tests du khi-deux. Cependant, la comparaison des coefficients de corrélation (0.07 puis 0.036) suggère que la corrélation positive entre nombre d'enfants et bonheur exprimé est deux fois plus élevée chez les moins aisés. La faible P-valeur atteste de la fiabilité statistique de ces tests.

Ainsi, avoir plus d'enfants serait davantage corrélé positivement avec le bonheur chez les moins aisés. On peut donc penser que la parentalité contribue en moyenne davantage au bonheur des personnes ayant de plus faibles revenus, même si l'on doit rester prudent dans l'interprétation car notre observation ne s'effectue toujours pas toutes choses égales par ailleurs.

5.2 L'âge de la personne interrogée

On peut raisonnablement penser que le nombre d'enfants vivant avec le répondant est corrélé à son âge : par exemple, au dessus d'un certain âge, les enfants du foyer ont déménagé pour faire leurs études. Or on peut également penser que le niveau de bonheur dépend de l'âge : ainsi, la corrélation positive trouvée entre nombre d'enfants et bonheur n'est-elle pas en fait une conséquence de l'âge plutôt que des bénéfices de la parentalité ?

Commençons par regarder la relation entre nombre d'enfants et âge. Le test du khi-deux nous donne une p-value inférieure à 0.0001 : on peut donc rejeter l'hypothèse d'indépendance au seuil de 5%. On obtient ensuite un coefficient de corrélation de -0.46 soit une corrélation négative assez nette : plus les personnes sont âgées, moins le nombre d'enfants vivant dans leur ménage a tendance à être important.

Regardons maintenant la relation entre le bonheur et l'âge. On trouve une p-value de 0,01 pour le test du khi-deux : on peut donc rejeter l'hypothèse d'indépendance au seuil de 5%. On obtient un coefficient de corrélation de -0.09 soit une faible corrélation négative : le niveau de bonheur a tendance à diminuer quand l'âge augmente.

Finalement, on a trouvé qu'en général, plus les personnes sont âgées, moins elles sont heureuses et moins elles ont d'enfants. Il faut donc être prudent : on avait noté une corrélation positive entre niveau de bonheur et nombre d'enfants, mais celle-ci pourrait être due au fait que les personnes vivants avec des enfants sont plus jeunes que les autres, et le niveau de bonheur élevé peut être une conséquence de l'âge et non du nombre d'enfants.

5.3 Le genre

Enfin, une dernière variable est susceptible de jouer un rôle fondamental dans la corrélation entre nombre d'enfants et niveau de bonheur : il s'agit du genre de la personne interrogée. En effet, comme l'ont documenté de nombreux travaux sociologiques, le temps de travail domestique et parental est très inégalement réparti : chaque jour en moyenne, les mères d'enfants de moins de 15 ans consacrent 4h13 au travail domestique et 1h35 aux activités parentales, contre respectivement 2h05 et 31 minutes pour les pères (Algava,2002) [5]. Ainsi, la présence d'enfants, a fortiori d'enfants en bas âge, ferait reposer en moyenne une charge de travail beaucoup plus importante sur les épaules des femmes, ce qui pourrait réduire la satisfaction qu'elles tirent de la parentalité. Afin de vérifier si cette hypothèse se concrétise dans les faits, on regarde la corrélation entre nombre d'enfants et bonheur en fixant le genre de la personne interrogée.

De manière plutôt surprenante, il ne semble pas y avoir d'effet notable du genre sur la corrélation étudiée. La seule différence significative semble être les effectifs des répondants ayant un niveau de bonheur élevé et plus de trois enfants : parmi les femmes ayant plus de trois enfants, 20% déclarent un niveau de bonheur de 9 ou 10, contre 6% des hommes ayant plus de trois enfants. (voir graphiques 1 et 2 en annexe).

Pour tenter d'expliquer pourquoi notre hypothèse est ainsi en partie invalidée, deux hypothèses peuvent être avancées. D'une part, les inégalités entre les hommes et les femmes persistent mais tendent à perdre en intensité au sein des nouvelles générations. Les répondants ont donc peut-être le sentiment que les charges incombant aux parents sont mieux réparties que chez leurs parents ou grands parents, et sont donc globalement satisfaits de la répartition, quand bien même elle demeurerait assez inégalitaire au détriment des femmes. L'étude de la variable "prhlpl" semble abonder dans ce sens : en moyenne, il n'y a pas de différence

significative entre la propension qu'ont les femmes et les hommes à déclarer s'occuper des autres (5.3 pour les hommes contre 5.4 pour les femmes). D'autre part, femmes et hommes participent de façon plus ou moins égale au travail parental et coordonnent leur tâches domestiques en fonction de leur participation au marché du travail, ou encore de ce que le sociologue Kellerhals a appelé le "style conjugal" [8]. Or, Kellerhals explique que ce "style conjugal" dépend en grande partie du milieu social : en particulier, à mesure que le niveau d'éducation (fortement lié au milieu social) du couple s'élève, la norme de répartition des tâches tend à devenir plus égalitaire. Par conséquent, on peut affirmer que l'idée que la maternité est plus contraignante que la paternité doit être nuancée puisqu'elle n'est pas vraie pour tous les couples. Plus précisément, elle serait valable davantage dans les milieux populaires que dans les milieux plus aisés. Toutefois, lorsque l'on compare la corrélation entre genre et bonheur exprimé en contrôlant le revenu, on ne constate pas de différence majeure. De plus, toutes ces analyses ne parviennent toutefois pas à répondre à la question suivante : Comment expliquer la différence de niveau de bonheur entre les hommes ayant beaucoup d'enfants et les femmes ayant beaucoup d'enfants que nous avons relevée ?

6 Conclusion

Il en faut deux pour être heureux ! C'est du moins ce qui semble d'abord ressortir de notre étude de la corrélation entre nombre d'enfants et bonheur déclaré. En effet, les personnes interrogées qui déclarent un niveau de bonheur le plus important sont celles qui ont deux enfants. Viennent ensuite celles qui n'ont qu'un enfant, puis celles qui n'en ont aucun ou qui en ont plus de deux. Lorsque nous construisons un indicateur de bonheur prenant aussi en compte la satisfaction par rapport à la vie au quotidien, et non seulement en général, ce lien disparaît. Une raison envisagée est que les parents pourraient se sentir plus satisfaits en général sans être plus heureux au quotidien. Nous trouvons aussi que le nombre d'enfants ne semble pas influencer le sentiment de se sentir bien entouré.

Précisons-le bien : nous ne pouvons pas dire que c'est le fait d'avoir deux enfants qui rend heureux. L'échantillon des couples habitant avec leurs enfants a toutes les chances d'être biaisé par un effet d'auto-sélection : les parents sont effectivement heureux d'avoir d'enfants car il s'agit précisément d'individus qui souhaitaient avoir des enfants. De plus, nous ne disposons pas d'outils qui nous permettraient de trouver des liens de causalité, comme la régression.

Même en s'en tenant à l'étude de la corrélation, des limites méthodologiques émergent. Tout d'abord, il est possible que les individus "sur-déclarent" leur bonheur par un effet de désirabilité sociale. Ensuite, les choix des variables et des pondérations utilisées pour construire nos indices contiennent une part d'arbitraire. Ce que nous assimilons à "être heureux" et "se sentir bien entouré" peut être discuté. Finalement, l'étude de variables confondantes telles que le revenu, l'âge de la personne interrogée ou le genre nous conduit à plusieurs questions auxquelles on ne peut répondre qu'en partie et de manière incertaine. Est-ce que les plus aisés sont plus heureux en moyenne parce qu'ils font plus d'enfants ou parce qu'ils sont plus aisés ? Est-ce que les plus jeunes sont plus heureux parce qu'ils vivent avec plus d'enfants ou parce qu'ils sont plus jeunes ? Comment expliquer que les femmes ayant au moins trois enfants sont tendanciellement plus heureuses que les hommes dans la même situation, alors même que femmes et hommes en général semblent exprimer à peu près le même niveau de bonheur ?

Pour prolonger notre approche de la corrélation et tenter de mieux répondre à certaines de ces interrogations, il pourrait être pertinent de regarder l'effet d'autres variables telles que le lieu d'habitation ou encore la place de l'individu dans le "cycle de vie" (étudiant, jeune actif, retraité...) qui apporterait sans doute d'autres enseignements sur la relation entre bonheur et nombre d'enfants, et sur ce qui se joue dans les relations sociales au sein de la famille.

Bibliographie

- [1] E. Algava. “quel temps pour les activités parentales?”. *Etudes et Résultats*, n° 162, 2002.
- [2] R. C. Barnett and G. K. Baruch. ‘women’s involvement in multiple roles and psychological distress’. *Journal of Personality and Social Psychology*, 1985. 49, 135–145.
- [3] T. Basu. ‘does having kids make parents happy after all?’. *The Atlantic*, 2014.
- [4] V. Bengtson. ‘beyond the nuclear family’. 2001.
- [5] C. Bonvalet and J. Ogg. *Enquêtes sur l’entraide familiale en Europe*. INED, 2006.
- [6] C. M. Herbst and J. Ifcher. ‘the increasing happiness of american parents’. *Review of Economics of the Household*, 2016. 14(3), 529–551.
- [7] Enquête INED-Insee. Enquête intentions de fécondité, 1ère vague, 1998.
- [8] J. Kellerhals et al. *Mariages au quotidien*. Editions Pierre - Marcel Favre. Regards sociologiques., 1982.
- [9] K.M. Nomaguchi and M.A. Milkie. ‘costs and rewards of children : The effects of becoming a parent on adults’ lives’. *Journal of Marriage and Family*, 2004.
- [10] A. Régnier-Loilier. *Avoir des enfants en France, Désirs et réalités*. INED, 2007.
- [11] Enquête SRCV-Insee Résultats. Satisfaction dans la vie en 2017. insee.fr/fr/statistiques/3637041, 2019.

Annexes

Annexe 1 - Liste des variables utilisées

Variable expliquée : le bonheur déclaré

nom de la variable	signification	justification
happy	Réponse sur une échelle de 1 à 10 à la question "How happy are you?"	La mesure la plus directe du bonheur parmi les variables disponibles.

Variable explicative : le nombre d'enfants du couple

nom de la variable	signification	justification
icpart1	La personne interrogée vit-elle en couple ?	pour ne sélectionner que les couples
yrbrn2 à yrbrn24	année de naissance des personnes 2 à 24 du ménage	pour avoir le nombre d'enfants du couple
rshipa2 à rshipa24	lien des personnes 2 à 24 du ménage avec la personne interrogée	pour éliminer les relations non pertinentes comme les frères et sœurs de la personne en couple interrogée

Variations confondantes envisagées

nom de la variable	signification	justification
gndr	genre de la personne interrogée	pour voir comment l'influence du nombre d'enfants sur le bonheur varie selon le genre, dont découlent des rôles sociaux et des responsabilités sexuellement différenciées pouvant influencer sur la satisfaction exprimée(5)
agea	âge de la personne interrogée	pour ne sélectionner que les couples qui ont l'âge d'avoir des enfants qui habitent avec eux
hinctnta	décile de revenu net du ménage	nous nous attendons à ce que le revenu ait une influence significative sur le bonheur, et nous aimerions contrôler cet effet

Variables utilisées pour le score de bonheur

variable	signification	normalisation	inversion	coefficient	signe
happy	Sentiment de bonheur en général			3	+
pstvms	Bonne opinion de soi-même	oui	oui	1	+
wrhpp	Sentiment de bonheur la semaine passée			2	+
enjl	Fait d'avoir apprécié sa vie la semaine passée	oui		2	+
tmendng	Enthousiasme à propos de ce qu'on fait (fréquence)			1	+
ftsd	Fréquence du sentiment de tristesse la semaine passée	oui		2	-
ftdpr	Fréquence du sentiment de déprime la semaine passée	oui		3	-
firms	Impression occasionnelle d'être un échec	oui	oui	2	-
fteeff	Fréquence de l'impression que tout est un effort la semaine passée	oui		1	-
fitanx	Fréquence du sentiment d'anxiété la semaine passée	oui		1	-

Variables utilisées pour la mesure du sentiment d'être bien entouré

variable	signification	normalisation	inversion	coefficient	signe
ftlnl	Fréquence du sentiment de solitude la semaine passée	oui	oui	1	+
rehlpl	Recevoir de l'aide et du soutien de ses proches	oui		1	+
flapppl	Se sentir apprécié par son entourage			1	+

Annexe 2 - Tableau et graphiques complémentaires

Tableau 1 : test du χ^2 pour le tri croisé entre le nombre d'enfants et le niveau de bonheur

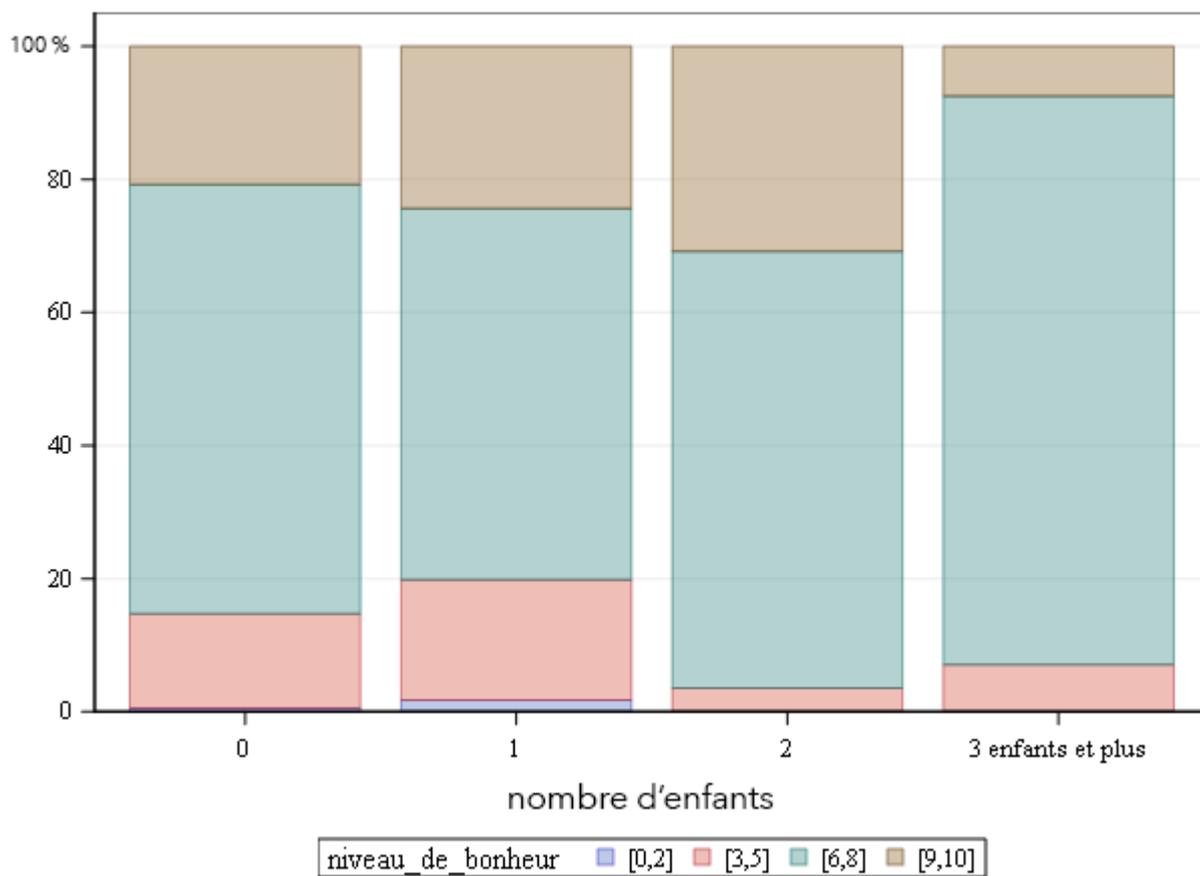
Niveau de bonheur	Nombre d'enfants				Total
	0	1	2	3 et plus	
[0,2]	0,03	0,51	0,32	0,69	1,55
[3,5]	1,65	0,42	6,89	0,99	9,95
[6,8]	0,01	0,99	0	3,06	4,06
[9,10]	0,66	0,99	3,78	3,74	9,17
Total	2,35	2,91	10,98	8,48	24,73

Source : enquête ESS 2012, vague 6

Champ : personnes françaises de plus de 15 ans vivant dans un ménage avec un conjoint

Lecture : Les nombre de personnes qui ont 2 enfants et déclarent un niveau de bonheur de 9 ou 10 contribuent à hauteur de 3,78 au 2 total, qui vaut 24,73.

Graphique 1 : Niveau de bonheur des hommes en fonction du nombre d'enfants

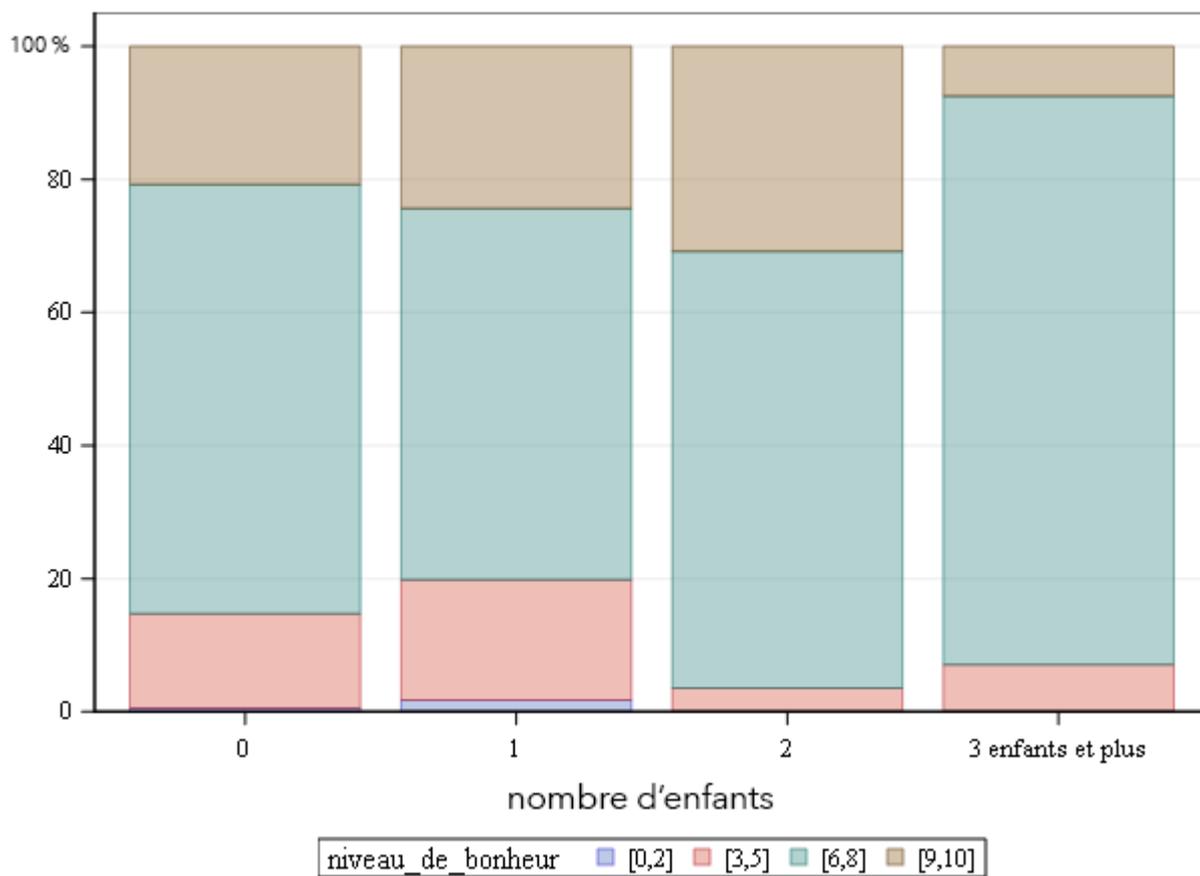


Source : Enquête ESS 2012 vague 6

Champ : Hommes français de plus de 15 ans vivant dans un ménage avec un conjoint

Lecture : 20% des hommes n'ayant pas d'enfants déclarent un niveau de bonheur de 9 ou de 10.

Graphique 2 : Niveau de bonheur des femmes en fonction du nombre d'enfants



Source : Enquête ESS 2012 vague 6

Champ : Hommes français de plus de 15 ans vivant dans un ménage avec un conjoint

Lecture : 20% des femmes ayant 3 enfants ou plus déclarent un niveau de bonheur de 9 ou de 10.

Annexe 3 - Script SAS

```
/* =====  
Titre : Il en faut deux pour etre heureux – Memoire de statistiques descriptives  
Sujet : Famille et bonheur  
  
Janvier 2020  
  
Pauline Culioli  
Jerome Pivard  
Remi Turquier  
  
Referent : Mael Ginsburger  
===== */  
  
/* ===== PLAN DU SCRIPT =====  
# I. Construction de la table et des variables  
# II. Correlation entre le bonheur declare et le nombre d enfants  
# III. D autres manieres de mesurer le bonheur  
# IV. etre bien entoure : une explication possible au lien entre nombre d enfants et bonheur ?  
# V. etude de variables confondantes  
  
===== */  
  
/* ===== I. Construction de la table et des variables ===== */  
  
/*importation des donnees de l ESS dans la librairie work */  
PROC IMPORT datafile="W:\Bureau\SAS\ESS\ESS6e02_4.sav"  
  out = ess;  
RUN;  
/*creation d une table ne contenant que les donnees qui concernent la France */  
DATA essfrance;  
  SET ess;  
  WHERE cntry="FR";  
RUN;  
/*on ne garde que les variables qui nous interessent*/  
DATA essfrance;  
  set essfrance  
  (keep = happy  
  icpart1  
  yrbrn yrbrn2 yrbrn3 yrbrn4 yrbrn5 yrbrn6 yrbrn7 yrbrn8 yrbrn9 yrbrn10 yrbrn11 yrbrn12 yrbrn13  
  yrbrn14  
  yrbrn15 yrbrn16 yrbrn17 yrbrn18 yrbrn19 yrbrn20 yrbrn21 yrbrn22 yrbrn23 yrbrn24  
  rshipa2 rshipa3 rshipa4 rshipa5 rshipa6 rshipa7 rshipa8 rshipa9 rshipa10 rshipa11 rshipa12  
  rshipa13  
  rshipa14 rshipa15 rshipa16 rshipa17 rshipa18 rshipa19 rshipa20 rshipa21 rshipa22 rshipa23  
  rshipa24  
  
  hhmmb  
  gndr  
  hinctnta  
  pspwght  
  pweight  
  wrhpp  
  enjlf  
  tmendng  
  pstvms  
  fltsd  
  fltdpr  
  flrms  
  flteeff  
  fltanx  
  fltlnl  
  flapppl  
  rehlppl  
  agea  
  prhlppl);  
RUN;  
  
/*definition d une nouvelle variable de ponderation */  
DATA essfrance;  
  SET essfrance;  
  pond = pspwght*pweight*10000;  
RUN;  
  
/*la variable qualitative "happy" est issue des reponses donnees a la question "How happy are you?",  
  ses modalites sont "Extremely unhappy",  
  2, 3, 4, 5, 6, 7, 8, 9, "Extremely happy".  
  Pour pouvoir effectuer des manipulations numeriques, il faut la transformer en variable quantitative.  
  On appelle cette variable happyqti*/  
DATA essfrance;  
  SET essfrance;  
  happyqti = INPUT(happy,BEST12.);  
RUN;
```

```

/* on fait de meme pour la variable wrhpp */
DATA essfrance;
SET essfrance;
LENGTH wrhppqti 5;
IF wrhpp>4 THEN wrhppqti = .;
ELSE wrhppqti = INPUT (wrhpp,BEST12.);
RUN;

/* On ne garde que les couples */
DATA essfrance;
SET essfrance;
WHERE icpart1=1;
RUN;

/* Creation de la variable nombre d enfants :
- nous creons d abord une indicatrice enfant[n] qui vaut 1 si la n-ieme personne du menage est un
  enfant et 0 sinon
- le nombre d enfants est la somme des indicatrices
*/

DATA essfrance;
SET essfrance;
IF rshipa2 = 2 AND 2012 - yrbrn2 <= 15 THEN enfant2 = 1;
ELSE enfant2 = 0 ;
IF rshipa3 = 2 AND 2012 - yrbrn3 <= 15 THEN enfant3 = 1;
ELSE enfant3 = 0 ;
IF rshipa4 = 2 AND 2012 - yrbrn4 <= 15 THEN enfant4 = 1;
ELSE enfant4 = 0 ;
IF rshipa5 = 2 AND 2012 - yrbrn5 <= 15 THEN enfant5 = 1;
ELSE enfant5 = 0 ;
IF rshipa6 = 2 AND 2012 - yrbrn6 <= 15 THEN enfant6 = 1;
ELSE enfant6 = 0 ;
IF rshipa7 = 2 AND 2012 - yrbrn7 <= 15 THEN enfant7 = 1;
ELSE enfant7 = 0 ;
IF rshipa8 = 2 AND 2012 - yrbrn8 <= 15 THEN enfant8 = 1;
ELSE enfant8 = 0 ;
IF rshipa9 = 2 AND 2012 - yrbrn9 <= 15 THEN enfant9 = 1;
ELSE enfant9 = 0 ;
IF rshipa10 = 2 AND 2012 - yrbrn10 <= 15 THEN enfant10 = 1;
ELSE enfant10 = 0 ;
IF rshipa11 = 2 AND 2012 - yrbrn11 <= 15 THEN enfant11 = 1;
ELSE enfant11 = 0 ;
IF rshipa12 = 2 AND 2012 - yrbrn12 <= 15 THEN enfant12 = 1;
ELSE enfant12 = 0 ;
IF rshipa13 = 2 AND 2012 - yrbrn13 <= 15 THEN enfant13 = 1;
ELSE enfant13 = 0 ;
IF rshipa14 = 2 AND 2012 - yrbrn14 <= 15 THEN enfant14 = 1;
ELSE enfant14 = 0 ;
IF rshipa15 = 2 AND 2012 - yrbrn15 <= 15 THEN enfant15 = 1;
ELSE enfant15 = 0 ;
IF rshipa16 = 2 AND 2012 - yrbrn16 <= 15 THEN enfant16 = 1;
ELSE enfant16 = 0 ;
IF rshipa17 = 2 AND 2012 - yrbrn17 <= 15 THEN enfant17 = 1;
ELSE enfant17 = 0 ;
IF rshipa18 = 2 AND 2012 - yrbrn18 <= 15 THEN enfant18 = 1;
ELSE enfant18 = 0 ;
IF rshipa19 = 2 AND 2012 - yrbrn19 <= 15 THEN enfant19 = 1;
ELSE enfant19 = 0 ;
IF rshipa20 = 2 AND 2012 - yrbrn20 <= 15 THEN enfant20 = 1;
ELSE enfant20 = 0 ;
IF rshipa21 = 2 AND 2012 - yrbrn21 <= 15 THEN enfant21 = 1;
ELSE enfant21 = 0 ;
IF rshipa22 = 2 AND 2012 - yrbrn22 <= 15 THEN enfant22 = 1;
ELSE enfant22 = 0 ;
IF rshipa23 = 2 AND 2012 - yrbrn23 <= 15 THEN enfant23 = 1;
ELSE enfant23 = 0 ;
IF rshipa24 = 2 AND 2012 - yrbrn24 <= 15 THEN enfant24 = 1;
ELSE enfant24 = 0 ;
RUN;

/* Creation de la variable nbenfants en faisant la somme des indicatrices */
DATA essfrance;
SET essfrance;
nbenfants = enfant2 + enfant3 + enfant4 + enfant5 + enfant6 + enfant7 + enfant8 + enfant9 + enfant10
+ enfant11
+ enfant12 + enfant13 + enfant14 + enfant15 + enfant16 + enfant17 + enfant18 + enfant19 +
enfant20 + enfant21
+ enfant22 + enfant23 + enfant24;
RUN;

/* Construction du graphique representant le nombre d enfants des personnes interrogees */
ODS RTF FILE="W:\Bureau\SAS\ESS\graphnbenfants.rtf";
TITLE "Nombre d enfants des personnes en couple interrogees";
PROC UNIVARIATE DATA = essfrance;
VAR nbenfants;
HISTOGRAM / MIDPOINTS = 0 to 7;
RUN; QUIT;
TITLE;
ODS RTF CLOSE;

/* ===== */

```

```

/* ===== II. Correlation entre le bonheur declare et le nombre d enfants
===== */

/***** Statistiques bivariees *****/

/* tri croise entre le niveau de bonheur et le nombre d enfants */
PROC FREQ DATA = essfrance;
TABLES happyqti * nbenfants / NOROW NOCOL
NOPERCENT;
RUN;

/* pour rendre les resultats plus lisibles , on va separer le nombre d enfants en 4 modalites: pas d
enfants , 1 enfants , 2 enfants , 3 enfants et plus */
DATA essfrance;
SET essfrance;
LENGTH nbenfants2 $20;
IF nbenfants=0 then nbenfants2 = "0";
ELSE IF nbenfants=1 then nbenfants2 = "1";
ELSE IF nbenfants=2 then nbenfants2 = "2";
ELSE IF nbenfants>2 then nbenfants2= "3 enfants et plus";
RUN;

/* on fait de meme pour le niveau de bonheur */
DATA essfrance;
SET essfrance;
LENGTH happy2 $20;
IF happyqti<3 then happy2=" [0,2]";
ELSE IF happyqti<6 then happy2=" [3,5]";
ELSE IF happyqti<9 then happy2=" [6,8]";
ELSE IF happyqti = 9 OR happyqti = 10 THEN happy2=" [9,10]";
RUN;

/* tri croise entre le niveau de bonheur et le nombre d enfants codes par nos nouvelles variables */
PROC FREQ DATA = essfrance;
TABLES happy2 * nbenfants2 / NOROW NOCOL
NOPERCENT;
RUN;

/* Analyse de la correlation entre le nombre d enfants et le niveau de bonheur exprime , et calculs
des coefficients de Pearson, Spearman et Kendall */
PROC CORR DATA = essfrance SPEARMAN KENDALL ;
VAR nbenfants happy;
WEIGHT pspwght;
RUN;

/* Calcul du tau b de Kendall */
PROC FREQ DATA = essfrance;
TABLES happy*nbenfants / MEASURES;
RUN;

/* On obtient que le tau de Kendall n est pas deux fois plus grand que son ecart type ( qui vaut
0.0245) .
Il y a donc de fortes chances qu il ne soit pas statistiquement significatif au seuil de 5%. */

/* Tri croise entre le niveau de bonheur et le nombre d enfants et obtention du khi-deux */
PROC FREQ DATA = essfrance;
TABLES happyqti * nbenfants / CHISQ CELLCHI2;
RUN;

/*-----tri croise et test CHI2 avec les nouvelles variables
-----*/
PROC FREQ DATA = essfrance;
TABLES happy2 * nbenfants2 / CHISQ CELLCHI2;
RUN;
/* cette fois , il n y a que 3 cellules ayant un effectif inferieur a 5.
On obtient un chi2 de 0.0862*/

/* Test du khi-deux entre le niveau de bonheur et le nombre d enfants codes par nos nouvelles
variables avec ponderation*/
PROC FREQ DATA = essfrance;
TABLES happy2 * nbenfants2 / CHISQ CELLCHI2 NOPERCENT NOROW NOCOL;
WEIGHT pspwght;
RUN;

/* Calculs des coefficients de Pearson et Spearman, ainsi que du tau b de Kendall avec les nouvelles
variables qualitatives */
PROC FREQ DATA = essfrance;
TABLES happy2*nbenfants2 / MEASURES;
RUN;

/*
On obtient le coefficient de Pearson : 0.06
Le coefficients de correlation des rangs de Spearman vaut 0.06

```

```

Enfin , le tau b de Kendall vaut 0.05.
*/

/* Tri croise entre les deux variables apres ponderation.
Calculs des coeffcients de Pearson , Spearman et Kendall */

PROC FREQ DATA = essfrance;
    TABLES happy2*nbenfants2 / MEASURES;
    WEIGHT pspwght;
RUN;

/* Export du tableau dans Excel */

ODS EXCEL FILE= "W:\Bureau\sas\analysevariables.xls ";
PROC FREQ DATA = essfrance;
    TABLES happy2*nbenfants2 / MEASURES;
    WEIGHT pspwght;
    title "Niveau de bonheur en fonction du nombre d enfants";
    ODS OUTPUT OneWayFreqs(match_all)=freqs;
RUN;
ODS EXCEL CLOSE;

/* ----- Resultats -----
Coefficient de Pearson de correlation lineaire : 0.06.
Coefficient de Spearman : 0.06.
Tau b de Kendall : 0.06.
*/

/* Representation graphique du niveau de bonheur en fonction du nombre d enfants */

/*recodage des variables pour la representation graphique*/
DATA essfrance;
    SET essfrance;
    LENGTH niveau_de_bonheur $20;
    niveau_de_bonheur=happy2;
    LENGTH nombre_d_enfants $20;
    nombre_d_enfants=nbenfants2;
RUN;

ODS RTF FILE = "W:\Bureau\SAS\ess\graphpourcentage.rtf ";
PROC FREQ Data=essfrance;
    TABLES niveau_de_bonheur*nombre_d_enfants / plots=FREQPLOT(twoway=stacked scale=group);
    WEIGHT pond;
RUN;
ODS RTF CLOSE;

/*
=====
*/

/* ===== III. D autres manieres de mesurer le bonheur =====
*/

/*-----Creation d un score synthetique de bonheur-----*/

/*Recodage des variables en variables quantitatives*/

/*enjlif: Enjoyed life , how often past week
1 None or almost none of the time
2 Some of the time
3 Most of the time
4 All or almost all of the time
7 Refusal
8 Don t know
9 No answer */

DATA essfrance;
    SET essfrance;
    LENGTH enjlfqti 5;
    IF enjlf > 4 THEN enjlfqti = .;
    ELSE enjlfqti = INPUT(enjlf ,BEST12.);
RUN;

/* tmendng - Enthusiastic about what you are doing , how much of the time
0 None of the time
1 1
2 2
3 3
4 4
5 5
6 6
7 7
8 8
9 9
10 All of the time
77 Refusal
88 Don t know

```

```

99 No answer */
DATA essfrance;
SET essfrance;
LENGTH tmendngqti 5;
IF tmendng > 10 THEN tmendngqti = .;
ELSE tmendngqti = INPUT(tmendng,BEST12.);
RUN;

/*pstvms - In general feel very positive about myself
1 Agree strongly
2 Agree
3 Neither agree nor disagree
4 Disagree
5 Disagree strongly
7 Refusal
8 Don t know
9 No answer */
DATA essfrance;
SET essfrance;
LENGTH pstvmsqti 5;
IF pstvms > 5 THEN pstvmsqti = .;
ELSE pstvmsqti = INPUT(pstvms,BEST12.);
RUN;

/* fltsd - Felt sad, how often past week
1 None or almost none of the time
2 Some of the time
3 Most of the time
4 All or almost all of the time
7 Refusal
8 Don t know
9 No answer*/
DATA essfrance;
SET essfrance;
LENGTH fltsdqti 5;
IF fltsd > 4 THEN fltsdqti = .;
ELSE fltsdqti = INPUT(fltsd ,BEST12.);
RUN;

/* fltdpr - Felt depressed, how often past week
1 None or almost none of the time
2 Some of the time
3 Most of the time
4 All or almost all of the time
7 Refusal
8 Don t know
9 No answer*/
DATA essfrance;
SET essfrance;
LENGTH fltdprqti 5;
IF fltdpr > 4 THEN fltdprqti = .;
ELSE fltdprqti = INPUT(fltdpr ,BEST12.);
RUN;

/* flrms - At times feel as if I am a failure
1 Agree strongly
2 Agree
3 Neither agree nor disagree
4 Disagree
5 Disagree strongly
7 Refusal
8 Don t know
9 No answer */
DATA essfrance;
SET essfrance;
LENGTH flrmsqti 5;
IF flrms > 5 THEN flrmsqti = .;
ELSE flrmsqti = INPUT(flrms ,BEST12.);
RUN;

/* flteeff - Felt everything did as effort, how often past week
1 None or almost none of the time
2 Some of the time
3 Most of the time
4 All or almost all of the time
7 Refusal
8 Don t know
9 No answer */
DATA essfrance;
SET essfrance;
LENGTH flteeffqti 5;
IF flteeff > 5 THEN flteeffqti = .;
ELSE flteeffqti = INPUT(flteeff ,BEST12.);
RUN;

/*fltanx - Felt anxious, how often past week
1 None or almost none of the time
2 Some of the time
3 Most of the time
4 All or almost all of the time
7 Refusal

```

```

8 Don t know
9 No answer*/

DATA essfrance;
SET essfrance;
LENGTH fltanxqti 5;
IF fltanx > 4 THEN fltanxqti = .;
ELSE fltanxqti = INPUT(fltanx,BEST12.);
RUN;

/* macro pour normaliser les variables */
%macro normaliser(varqti= , a= ,b= );
DATA essfrance;
SET essfrance;
LENGTH norm&varqti. 5 ;
%do indice= &a. %to &b. ;
IF &varqti. = &indice. THEN norm&varqti. = ((&indice. - &a.)/(&b. - &a.))*10 ;
%end;
%mend normaliser;

options mprint ; /* On voit dans le journal ce qui a ete execute au fur et a mesure */
%normaliser(varqti= wrhppqti , a= 1, b= 4);
%normaliser(varqti= enjlfqti , a= 1, b= 4);
%normaliser(varqti= pstvmsqti , a= 1, b= 5);
%normaliser(varqti= fltsdqti , a= 1, b= 4);
%normaliser(varqti= fltdprqti , a= 1, b= 4);
%normaliser(varqti= flrmsqti , a= 1, b= 5);
%normaliser(varqti= flteeffqti , a= 1, b= 4);
%normaliser(varqti= fltanxqti , a= 1, b= 4);

options nomprint; /* On clot pour pas qu il ne se reutilise sur les autres macros */

/* Deux variables sont codees dans le sens contraire de celui qui nous arrange.
Par exemple, une petite valeur de la variable pstvms signifie que la personne interrogee a souvent
une bonne opinion d elle-meme,
alors que nous voulons qu une petite valeur corresponde a moins de bonheur.
Nous creons donc un macro qui nous permet d inverser les valeurs de ces deux variables
(apres qu elles ont ete recodees en variables quantitatives , puis renormalisees sur une echelle de 0
a 10.*/

%macro flip(var= );
DATA essfrance;
SET essfrance;
LENGTH flip&var. 5 ;
flip&var. = 10 - &var. ;
%mend flip;

options mprint;
%flip(var=normpstvmsqti);
%flip(var=normflrmsqti);
options nomprint;

/* On regarde combien il y a de non-reponses et de valeurs manquantes prises par les variables
recodees et normalisees */

DATA essfrance;
SET essfrance;
nbnr = (happyqti = .) + (flipnormpstvmsqti = .) + (normwrhppqti = .) + (normenjlfqti = .) + (
tmendngqti = .)
+ (normfltsdqti = .) + (normfltdprqti = .) + (flipnormflrmsqti = .) + (normflteeffqti = .) +
(normfltanxqti = .);
RUN;

/* On n a pas de non-reponse */

/* Construction du score */

DATA essfrance;
SET essfrance;
LENGTH score1 5;
IF nbnr = 0 THEN score1 = ( 3*happyqti + 1*flipnormpstvmsqti + 2*normwrhppqti + 2*
normenjlfqti + 1*tmendngqti - 2*normfltsdqti - 3*normfltdprqti - 2*flipnormflrmsqti -
1*normflteeffqti - 1*normfltanxqti)/18;
ELSE score1 = . ;
RUN;

/* On fait la moyenne des contributions des differentes variables en affectant une valeur positive a
ce qui connote le bonheur et une valeur negative a ce qui connote le malheur.
Nous avons choisi de ponderer cette moyenne en affectant aux differentes composantes un coefficient
prenant les valeurs 1, 2, ou 3, en fonction de ce que l on estime etre leur importance relative
dans l indicateur ,
choix assurement discutable mais que nous justifions dans le corps du memoire. Nous nous sommes
notamment efforces d affecter autant de coefficients aux "variables positives" et aux "
variables negatives" vis-a-vis du bonheur exprime.*/

PROC MEANS DATA = essfrance;
VAR score1;

```

```

RUN;

/* On fixe le nombre de decimales a une seule en reformatant la variable "score1" */
DATA essfrance;
    SET essfrance;
    FORMAT score1 5.1;
RUN;

/* ===== */

/* ----- Recodage de score1 en classes ----- */

/* La PROC UNIVARIATE permet de choisir un decoupage en classes */
PROC UNIVARIATE DATA = essfrance;
    VAR score1;
    HISTOGRAM / MIDPOINTS = -10 to 10;
RUN; QUIT;

/* La distribution des personnes selon la variables score1 est une courbe en cloche de mode 3
    environ.
Nous decoupons donc avec un intervalle plus resserre autour de 3.
*/

DATA essfrance;
    SET essfrance;
    LENGTH score1_class $20 ;
    IF -10 <= score1 < 0 THEN score1_class = "[-4,5 ; 0[";
    ELSE IF 0 <= score1 < 2.5 THEN score1_class = "[0 ; 2,5[";
    ELSE IF 2.5 <= score1 < 3.5 THEN score1_class = "[2,5 ; 3,5[";
    ELSE IF 3.5 <= score1 < 5.5 THEN score1_class = "[3,5 ; 5]";
    ELSE score1_class = "";
RUN;

/* Une PROC FREQ pour savoir si notre decoupage est assez equilibre. */
PROC FREQ DATA = essfrance;
    TABLES score1_class;
RUN; QUIT;

/* Les 4 classes obtenues sont d effectifs comparables (entre 300 et 400),
    sauf pour les personnes avec un score1 negatif, qui ne sont que 51. */

/* Representation graphique du niveau de bonheur en fonction du nombre d enfants */
PROC FREQ DATA=essfrance;
    TABLES score1_class*nbenfants2 / plots=FREQPLOT(twoway=stacked scale=group);
    WEIGHT pspwght;
RUN;
/* Le lien entre niveau de bonheur et nombre d enfants semble dispara[tre] lorsque le bonheur est
    code avec notre score:
on obtient des distributions qui sont presque identiques quel que soit le nombre d enfants. */

/* Est-ce qu il y a une correlation entre score1 et nbenfants ? */
PROC CORR DATA=essfrance spearman vardef=wt;
    VAR nbenfants score1;
    WEIGHT pspwght;
RUN;
/* Resultat : coefficient de correlation tres faible (0,035) et p-valeur tres elevee (p > 0.25).
    Il n y a effectivement pas de correlation entre le score de bonheur et le nombre d enfants. */

/* Essayons de regarder si notre score1 represente bien le bonheur*/
proc corr data=essfrance spearman kendall pearson vardef=wt;
var happyqti score1;
weight pspwght;
run;
/* Il y a une correlation forte et tres certaine. Notre score represente bien le niveau de bonheur. */

/* ===== IV. etre bien entoure : une explication possible au lien entre nombre d enfants et bonheur
    ===== */

/* ===== Score d integration sociale===== */

/* ----- Recodage en variables quantitatives ----- */

/* la variable "fltlnl" est reponse a la question "Felt lonely, how often past week" selon les
    modalites suivantes:
1 None or almost none of the time
2 Some of the time
3 Most of the time
4 All or almost all of the time
7 Refusal
8 Don t know
9 No answer
Il faut donc exclure les reponses 7,8,9 dans les traitements. On va aussi transformer la variable en
    variable quantitative */

```

```

DATA essfrance;
SET essfrance;
LENGTH fltlnlqti 5;
IF fltlnl>4 THEN fltlnlqti = .;
ELSE fltlnlqti = INPUT (fltlnl,BEST12.);
RUN;

/* Variables pour le score de non solitude: */

/* rehlpppl – Receive help and support from people you are close to
0 Not at all
1 1
2 2
3 3
4 4
5 5
6 Completely
7 Refusal
8 Don t know
9 No answer */

DATA essfrance;
SET essfrance;
LENGTH rehlppplqti 5;
IF rehlpppl > 6 THEN rehlppplqti = .;
ELSE rehlppplqti = INPUT(rehlpppl,BEST12.);
RUN;

/* flapppl – Feel appreciated by people you are close to
0 Not at all
1 1
2 2
3 3
4 4
5 5
6 6
7 7
8 8
9 9
10 Completely
77 Refusal
88 Don t know
99 No answer */

DATA essfrance;
SET essfrance;
LENGTH flappplqti 5;
IF flapppl > 10 THEN flappplqti = .;
ELSE flappplqti = INPUT(flapppl,BEST12.);
RUN;

/* ----- Normalisation -----*/
options mprint ;
%normaliser(varqti= rehlppplqti , a= 0, b= 6);
%normaliser(varqti= fltlnlqti , a= 1, b= 4);
options nomprint;

/* La variable fltlnl est codee dans le sens inverse a celui des deux autres variables du score d
integration sociale.
Nous utilisons la macro flip pour la mettre dans le bon sens. */
options mprint;
%flip(var=normfltlnlqti);
options nomprint;

/* ----- Creation du score -----*/
DATA essfrance;
SET essfrance;
LENGTH score2 5;
IF nbnr = 0 THEN score2 = (flappplqti + normrehlppplqti + flipnormfltlnlqti) / 3 ;
ELSE score2= . ;
RUN;

/* Limiter le nombre de decimales a 1 */
DATA essfrance;
SET essfrance;
FORMAT score2 5.1;
RUN;

/* ----- Repartition de la population selon le score d integration sociale ----- */
PROC UNIVARIATE DATA = essfrance;
VAR score2;
HISTOGRAM / MIDPOINTS = 0 to 10;
RUN; QUIT;

/*
score2 prend des valeurs entre 0 et 10.
La valeur moyenne de score2 est 8,5.
L ecart-type vaut 1,3.
75% des individus ont un score superieur 7,9.

```

```

*/

/* ----- Correlation entre le score d integration sociale et le bonheur ----- */
proc corr data=essfrance spearman kendall pearson vardef=wt;
    var happyqti score2;
    weight pspwght;
run;

/* Le coefficient de correlation lineaire (Pearson) vaut 0.34 (p < 0.0001).
Il y a visiblement un lien entre le bonheur declare et le sentiment d etre bien entoure.
*/

/* ----- Correlation entre le score d integration sociale et le nombre d enfants ----- */
proc corr data=essfrance spearman kendall pearson vardef=wt;
    var score2 nbenfants;
    weight pspwght;
run;

/* Le coefficient de correlation lineaire (Pearson) vaut -0.02, avec une p-valeur tres elevee (p >
0.5).
Il n y a visiblement pas de lien entre le nombre d enfant d une personne en couple et son sentiment d
etre bien entouree.
*/

/* ===== */

/* ===== V. Variables confondantes ===== */

/* ===== A. Revenu ===== */

/* Regardons la distribution du revenu au sein de notre echantillon. */
PROC FREQ data=essfrance;
    TABLES hinctnta;
RUN;
/*On note une nette sous-representation du premier decile de revenu (environ 5% de l echantillon). */

/* Test de la correlation entre niveau de bonheur declare et niveau de revenu */
PROC CORR DATA = essfrance;
    VAR happyqti hinctnta;
    WEIGHT pspwght;
RUN;

/* On obtient un coefficient de correlation d environ 0.14 et une faible P-value : il y a une
correlation positive entre revenu et bonheur.
La question est de savoir a quel point elle influe sur la correlation que l on etudie */

/* tri croise entre le niveau de revenus et le nombre d enfants code par notre nouvelle variable */
PROC FREQ DATA = essfrance;
    TABLES hinctnta * nbenfants2 / NOROW NOCOL
    NOPERCENT;
RUN;

/*Recodage de la variable hinctnta*/
DATA essfrance;
    SET essfrance;
    LENGTH hinctnta5 $ 1;
    IF hinctnta IN(1,2) THEN hinctnta5 = "1";
    ELSE IF hinctnta IN(3,4) THEN hinctnta5 = "2";
    ELSE IF hinctnta IN(5,6) THEN hinctnta5 = "3";
    ELSE IF hinctnta IN(7,8) THEN hinctnta5 = "4";
    ELSE IF hinctnta IN(9,10) THEN hinctnta5 = "5";
    ELSE IF hinctnta IN(77,88) THEN hinctnta5 = "";
RUN;

PROC FREQ DATA = essfrance;
    TABLES hinctnta5 * nbenfants2 / NOCOL CHISQ CELLCHI2 MEASURES;
    WEIGHT pspwght;
    WHERE hinctnta5 NE "";
RUN;

/* La faible P-value ( environ 0.016) indique que le test du khi-deux entre les deux variables est
significatif :
on peut rejeter l hypothese d independance entre nombre d enfants et niveau de revenu au seuil de 5%.
*/

/* Representation graphique du tri croise */

/*recodage de hinctnta5*/
DATA essfrance;
    SET essfrance;
    LENGTH decile_de_revenu $10;
    IF hinctnta5="1" then decile_de_revenu=" [1,2]";
    ELSE IF hinctnta5="2" then decile_de_revenu=" [3,4]";
    ELSE IF hinctnta5="3" then decile_de_revenu=" [5,6]";

```

```

ELSE IF hinctnta5="4" then decile_de_revenu="[7,8]";
ELSE IF hinctnta5="5" then decile_de_revenu="[9,10]";
RUN;

/*creation et exportation du graphique*/
TITLE "Niveau de revenus et nombre d enfants";
ODS RTF FILE = "W:\Bureau\SAS\ESS\graphrevenu-nbenfants2.rtf";
PROC FREQ DATA=essfrance;
    TABLES decile_de_revenu*nombre_d_enfants / plots=FREQPLOT(twoway=stacked scale=group);
    WEIGHT pspwght;
RUN;
TITLE;
ODS RTF CLOSE;

/* On fixe le revenu et on regarde la correlation entre nombre d enfants et bonheur exprime */
/* On regarde cette correlation d abord chez les 60% les plus pauvres ( deciles 1 a 6) */
PROC FREQ DATA = essfrance;
    TABLES happyqti * nbenfants2 / NOCOL CHISQ CELLCHI2 MEASURES;
    WEIGHT pspwght;
    WHERE hinctnta IN (1,2,3,4,5,6) ;
RUN;

/* Pas assez de monde pour avoir un khi-deux significatif. Coefficient de correlation lineaire : 0.07
( ie legere relation positive) */

/* On regarde cette correlation d abord chez les 60% les plus riches ( deciles 5 a 10) */
PROC FREQ DATA = essfrance;
    TABLES happyqti * nbenfants2 / NOCOL CHISQ CELLCHI2 MEASURES;
    WEIGHT pspwght;
    WHERE hinctnta IN (5,6,7,8,9,10) ;
RUN;

/* Meme probleme pour le khi-deux. Coefficient de correlation lineaire : 0.036, ie deux fois plus
faible que chez les moins aises.
On a toujours une relation positive mais moins forte que chez les moins aises.
Conclusion : avoir plus d enfants serait davantage correle positivement avec le bonheur chez les
moins aises. */

/* ===== B. Age ===== */

/* Commencons par creer des classes d ages d effectifs proches. */
DATA essfrance;
    SET essfrance;
    LENGTH agea2 $20;
    IF agea<35 THEN agea2="Moins de 35 ans";
    ELSE IF agea<45 THEN agea2="Entre 35 et 44 ans";
    ELSE IF agea<55 THEN agea2="Entre 45 et 54 ans";
    ELSE IF agea<65 THEN agea2="Entre 55 et 64 ans";
    ELSE IF agea NE . THEN agea2="Plus de 64 ans";
RUN;

/* Regardons la distribution de l age dans notre echantillon. */
PROC FREQ DATA = essfrance;
    TABLES agea2;
RUN;

PROC MEANS DATA = essfrance;
    VAR agea;
RUN;

/*trace de la h[U+FFFD] de Tukey*/
PROC SORT DATA = essfrance;
    BY agea;
RUN;
DATA essfrance;
    SET essfrance;
    ensemble="Ensemble";
RUN;
PROC BOXPLOT DATA = essfrance;
    PLOT agea*ensemble;
RUN;

/* On trouve que la moyenne d age de notre echantillon est de 51,7 ans et que la moitie des
repondants ont entre 40 et 65 ans environ. */

/*On regarde s il existe un lien statistiquement significatif entre les variables agea2 et happy2*/
PROC FREQ DATA = essfrance;
    TABLES happy2*agea2 / CHISQ CELLCHI2;
    WEIGHT pspwght;
RUN;

/* On trouve une p-valeur de 0,01 pour le test du Khi2: on peut donc rejeter l hypothese d
independance au seuil de 5%. */

/*Calcul du coefficient de correlation*/
PROC CORR DATA = essfrance;
    VAR agea happy;
    WEIGHT pspwght;
RUN;

```

```

/*On obtient un coefficient de -0.09 soit une corrélation négative: le niveau de bonheur a tendance a
diminuer quand l age augmente*/

/*Représentation graphique du niveau de bonheur en fonction de l age*/
PROC FREQ Data=essfrance;
  TABLES happy2*agea2 / plots=FREQPLOT(twoway=stacked scale=groupct);
  WEIGHT pspwght;
RUN;

/* On regarde maintenant la relation entre le nombre d enfants et l age */
PROC FREQ DATA = essfrance;
  TABLES nbenfants*agea2 / CHISQ CELLCHI2;
  WEIGHT pspwght;
RUN;
/* On trouve une p-value de <.0001 donc on peut largement rejeter l hypothese d independance au seuil
de 5% */

/*Calcul du coefficient de corrélation*/
PROC CORR DATA = essfrance;
  VAR agea nbenfants;
  WEIGHT pspwght;
RUN;
/*On obtient un coefficient de -0.46 soit une corrélation négative: plus les personnes sont agees ,
moins le nombre d enfants
vivant dans leur menage est important.*/

/* ===== C. Genre ===== */

/*Regardons la distribution de l age dans notre échantillon. */
PROC FREQ data = essfrance;
  TABLES gndr;
RUN;
/* Il y a 546 hommes et 544 femmes: on a donc pas de surreprésentation des hommes ou des femmes. */

/*Commen cons par comparer la corrélation entre niveau de bonheur et nombre d enfant selon le genre
du repondant (variable gndr)
Gender:
1 Male
2 Female
9 No answer*/

/* On regarde seulement les hommes */
TITLE "Niveau de bonheur des hommes en fonction du nombre d enfants";
ODS RTF FILE = "W:\Bureau\SAS\ESS\graphhommes2.rtf";
PROC FREQ Data=essfrance;
  WHERE gndr=1;
  TABLES niveau_de_bonheur*nombre_d_enfants / plots=FREQPLOT(twoway=stacked scale=groupct);
  WEIGHT pspwght;
RUN;
TITLE;
ODS RTF CLOSE;

/*On regarde seulement les femmes */
TITLE "Niveau de bonheur des femmes en fonction du nombre d enfants";
ODS RTF FILE = "W:\Bureau\SAS\ESS\graphfemmes2.rtf";
PROC FREQ Data=essfrance;
  WHERE gndr=2;
  TABLES niveau_de_bonheur*nombre_d_enfants / plots=FREQPLOT(twoway=stacked scale=groupct);
  WEIGHT pspwght;
RUN;
TITLE;
ODS RTF CLOSE;

/* On va maintenant regarder si les femmes ont en effet tendance a remplir le role de "caregiver"
grace a la variable prhlpl:
"Provide help and support to people you are close to"
0 Not at all
1 1
2 2
3 3
4 4
5 5
6 Completely
7 Refusal
8 Don t know
9 No answer */

PROC FREQ DATA = essfrance;
  WHERE prhlpl<7;
  TABLES prhlpl*gndr / NOROW NOCOL;
RUN;

/*recodage de prhlpl pour calculer une moyenne*/
DATA essfrance;
  SET essfrance;
  WHERE prhlpl<7;
  prhlplqti = INPUT(prhlpl,BEST12.);

```

```

RUN;

/*calcul de la moyenne pour chaque genre */
PROC SORT DATA=essfrance;
  BY gndr;
RUN;

PROC MEANS DATA=essfrance;
  VAR prhlpplqti;
  BY gndr;
RUN;

/* Resultat: on trouve une moyenne de 5.29 pour les hommes et de 5.42 pour les femmes, ce qui est
trop proche pour
conclure.
Comment alors expliquer la difference de niveau de bonheur entre les hommes ayant beaucoup d enfants
et les femmes ayant
beaucoup d enfants?*/

/***** Influence du genre sur la correlation en fonction du milieu social ( donc du revenu )
*****/

/* On regarde le lien entre genre et bonheur exprime dans les milieux plus modestes */
PROC CORR DATA = essfrance KENDALL;
  VAR happyqti gndr;
  WHERE hinctnta IN (1,2,3,4,5,6);
RUN;

/* Le tau b de Kendall vaut un peu plus de 0.07. */

/* On regarde le lien entre genre et bonheur exprime dans les milieux plus aises */
PROC CORR DATA = essfrance KENDALL;
  VAR happyqti gndr;
  WHERE hinctnta IN (5,6,7,8,9,10);
RUN;

/* Le tau b de Kendall vaut un peu moins de 0.07 */

/* Conclusion : Pas de difference notable */

```