

Master 1 - Econométrie et Statistiques, parcours Econométrie Appliquée

Analyse économétrique de l'impact de la gestion budgétaire
des étudiants sur leur moyenne générale (2023–2024)

Dossier réalisé par :

QUINTIN DE KERCADIO Pierre
CROCHET Florian

Année universitaire 2024 - 2025

Résumé

La moyenne générale représente un enjeu central dans la réussite académique des étudiants. Cette étude a pour objectif de mieux comprendre les facteurs financiers qui influencent la variation de la moyenne générale des étudiants. Dans cette perspective, nous avons conçu un questionnaire afin de recueillir des données auprès d'étudiants de Loire-Atlantique. Ensuite, nous avons soumis notre modèle à des tests économétriques appropriés pour en évaluer la validité et la robustesse. En conclusion, notre projet met en évidence l'impact de variables financières sur la moyenne générale, comme le nombre de restaurants mensuels, ou le fait de recevoir de l'argent autre que les aides nationales ou de travailler. Notre étude met également en évidence l'impact de la santé et du stress financier sur les résultats académiques. En réponse, nous proposons des solutions comme le fait d'aller aux événements gratuits organisés par l'université pour rencontrer des professionnels de la santé. Ainsi que de diminuer le nombre de sorties au restaurant.

Abstract

The overall grade point average is a central issue in students' academic success. This study aims to better understand the financial factors that influence variations in students' overall grade averages. To this end, we designed a questionnaire to collect data from students in the Loire-Atlantique region. We then subjected our model to appropriate econometric tests to assess its validity and robustness. In conclusion, our project highlights the impact of financial variables on the overall grade average, such as the number of monthly restaurant visits, receiving money other than national aid, or working. Our study also underscores the influence of health and financial stress on academic performance. In response, we propose solutions such as attending free events organized by the university to meet health professionals, as well as reducing the number of restaurant outings

Sommaire

I - Introduction	2
II - Analyse détaillée de la gestion budgétaire dans la réussite académique et présentation des variables	4
III - Méthode	23
IV - Analyse de la base de donnée	28
V - Estimations économétriques	44
VI - Prévisions	57
VII - Discussion	59
VIII - Conclusion	62
IX - Bibliographie	64
X - Annexes	71
XI - Table des matières	108

I - Introduction

La gestion budgétaire a toujours été une préoccupation primordiale des individus, que ce soit les professionnels, les personnes âgées ou les étudiants. Dans un contexte économique actuellement très difficile, une mauvaise gestion budgétaire peut entraîner un stress financier significatif, affectant la santé mentale et le bien-être général des étudiants, et par conséquent, leur capacité à se concentrer sur leurs études. Actuellement, environ 26 % des étudiants occupent un emploi à côté de leurs études¹. De ce fait, ces étudiants disposent de moins de temps pour étudier et, s'ils n'ont pas une organisation optimale, peuvent obtenir des résultats moins satisfaisants.

À l'heure actuelle, les étudiants font face à de nombreuses dépenses, allant des frais de scolarité aux fournitures scolaires, en passant par le logement, la nourriture et les loisirs. En France, en 2023, le coût de la vie étudiante a connu une augmentation de 6,47 %, marquant une hausse sans précédent en 19 ans. Cela représente un reste à charge annuel supplémentaire de 594,76 euros, soit 49,56 euros par mois². La capacité à bien gérer ses finances peut influencer divers aspects de la vie étudiante, y compris les résultats académiques.

L'intérêt du sujet réside dans le fait qu'une gestion efficace du budget permettrait aux étudiants de minimiser le stress financier, d'avoir par conséquent une meilleure concentration en cours, et de leur offrir un environnement propice à la réussite. De plus, dans un contexte où l'accès à l'éducation supérieure devient de plus en plus coûteux, comprendre comment la maîtrise des finances personnelles impacte les performances académiques devient un enjeu majeur tant pour les étudiants que pour les institutions éducatives.

Cette étude se limite à un cadre bien défini. En effet, nous nous consacrerons seulement sur les étudiants des Pays de la Loire d'un niveau minimum Bac +1, sur l'année 2023-2024.

¹ Vie Publique. (27/03/2023). Jobs étudiants : 146 000 étudiants occupent un emploi hors du cadre de leurs études.

² Unef. (2023). Enquête sur l'explosion du coût de la vie étudiante.

Par gestion budgétaire, nous chercherons à montrer que les étudiants qui gèrent bien leur budget auront, par conséquent, de meilleurs résultats académiques, et cela engendrera une baisse du stress financier.

Nous nous demanderons, comment les étudiants qui gèrent efficacement leur budget parviennent-ils à obtenir de meilleurs résultats académiques ?

Pour mener à bien notre analyse économétrique, nous avons élaboré un questionnaire que nous avons diffusé exclusivement sur nos réseaux sociaux, ainsi qu'au sein d'étudiants de l'IAE et des sciences par la messagerie universitaire. Nous avons ciblé les étudiants car ils représentaient la population la plus accessible pour ce type de questionnaire.

Dans l'optique de répondre à notre problématique, nous appliquons diverses techniques de régression linéaire afin d'étudier comment notre variable dépendante est déterminée par des variables explicatives, tout en effectuant des prévisions. L'analyse économétrique réalisée nous permettra d'identifier les facteurs clés qui influencent les résultats académiques des étudiants et de proposer des recommandations pertinentes.

Notre étude est structurée autour de quatre axes principaux. Dans un premier temps, nous réaliserons une analyse détaillée du sujet en examinant de manière approfondie les différentes études existantes, ce qui nous permettra de sélectionner des variables appropriées pour notre analyse. Dans un second temps, nous présenterons la méthode économétrique utilisée ainsi que la méthode de collecte des données. Dans un troisième temps, nous ferons une analyse détaillée de la base de données collectées. Et pour finir, nous effectuerons les estimations économétriques.

II - Analyse détaillée de la gestion budgétaire dans la réussite académique et présentation des variables

Dans cette partie nous avons décidé de nous pencher sur la pertinence de la moyenne générale (A), puis nous nous intéresserons aux variables explicatives selon la littérature (B) et pour finir nous ferons un récapitulatif du modèle (C) .

A. Pertinence de la variable à expliquer : la moyenne générale lors de l'année 2023-2024

Dans le cadre de notre étude visant à comprendre comment la gestion budgétaire des étudiants influence leurs résultats académiques, nous avons choisi d'utiliser la moyenne générale lors de l'année 2023-2024 comme variable à expliquer. Ce choix repose sur plusieurs raisons clés.

En effet, la moyenne générale est un indicateur qui synthétise les résultats de l'ensemble des matières étudiées. Elle permet de refléter la performance académique globale d'un étudiant, en prenant en compte ses réussites dans toutes les disciplines de son domaine d'étude sans exception. En évitant de se limiter à une seule matière, nous obtenons une vue d'ensemble plus représentative de la réussite étudiante.

En analysant les moyennes générales, nous sommes en mesure d'identifier des tendances significatives. Par exemple, une amélioration des résultats académiques pourrait être corrélée à une meilleure gestion financière des étudiants, ou à un étudiant qui a déjà pris des cours particuliers.

La moyenne générale facilite une analyse comparative entre différents groupes d'étudiants, tels que ceux qui gèrent efficacement leur budget par rapport à ceux qui rencontrent des difficultés, ou ceux qui sont assidus aux cours. Cette segmentation entre les différents groupes d'étudiants permet de mettre en évidence des différences significatives dans la performance académique.

De plus, la moyenne générale est une mesure largement acceptée et reconnue dans le milieu académique. Son utilisation renforce la validité et la fiabilité de nos résultats, rendant notre étude crédible aux yeux des étudiants, les institutions académiques et des employeurs.

En intégrant les variations individuelles des performances académiques, la moyenne générale nous permet d'analyser les étudiants sous différents angles. Cela inclut l'examen de ceux qui parviennent à maintenir de bonnes notes tout en travaillant, par exemple, ce qui peut illustrer l'importance d'une gestion budgétaire efficace.

En résumé, le choix de la moyenne générale comme variable économétrique est fondamental pour notre étude sur l'impact de la gestion budgétaire sur la réussite académique. Cette approche nous permet non seulement de répondre à notre problématique, mais aussi de formuler des recommandations pour aider les étudiants à améliorer leur performance académique par une gestion financière optimale. Les résultats de cette analyse pourraient également éclairer les établissements d'enseignement sur la nécessité de soutenir davantage les étudiants dans leur gestion budgétaire, contribuant ainsi à leur succès académique.

B. Pertinence des variables explicatives

Nous allons à présent étudier l'impact de chaque variable explicative sur notre variable à expliquer, Y. Nous nous aiderons d'études ainsi que de la littérature afin de justifier le choix de chaque variable.

1. Variables sociales

1.1. Genre

Le genre d'un individu, qu'il soit homme ou femme, peut avoir un impact sur les résultats académiques ainsi que sur la gestion budgétaire. Selon les chiffres de l'INSEE, les étudiantes sont plus susceptibles d'occuper un emploi que leurs homologues masculins, avec

un taux de 6,1 % contre 4,2 %. Cet écart se retrouve à tous les niveaux de diplôme, ce qui indique qu'il ne s'explique pas uniquement par des études plus longues.

Les étudiantes qui travaillent occupent souvent des postes tels que hôtesse de caisse, un métier qu'elles exercent plus fréquemment que les étudiants (15,0 % contre 9,4 %). Elles sont également plus nombreuses à travailler comme vendeuses (12,9 % contre 9,8 %). Par ailleurs, 7,2 % des étudiantes sont employées dans la famille professionnelle des assistants maternels, qui inclut le baby-sitting, tandis que seulement 0,8 % des étudiants occupent ce type de poste.

À l'inverse, les étudiants sont légèrement plus souvent employés dans le domaine culturelle, sportive et des surveillants, avec un taux de 9,1 %, contre 8,1 % pour les étudiantes³.

De plus, cette variable nous a semblé pertinente car il y a davantage de scolarisation chez les femmes que chez les hommes. En effet, en 2011, à l'âge de 21 ans, 37,2 % des femmes étaient scolarisées, contre seulement 30 % des hommes⁴.

On peut aussi noter que les études montrent que les femmes obtiennent, en moyenne, de meilleurs résultats académiques que les hommes à l'université, ce qui se traduit par des taux de passage à l'année supérieur plus élevés. Divers facteurs pourraient expliquer ce phénomène. Des recherches suggèrent que les femmes adoptent des méthodes de travail et des habitudes d'étude plus efficaces, favorisant la rigueur, la persévérance et l'organisation, des éléments cruciaux pour réussir dans le cadre universitaire. De plus, les étudiantes tendent à participer plus activement aux cours, posant des questions et recherchant des interactions avec leurs professeurs, ce qui peut améliorer leur compréhension des sujets abordés et leur performance académique⁵.

³ Solène Hilary, Laurent Lefèvre, Laurence Pen. (2024). Les « jobs » étudiants début 2020 : quatre fois sur dix, des emplois de serveurs, caissiers ou vendeurs, INSEE.

⁴ Angela Tirroloni. (2015). Meilleure scolarité mais insertion professionnelle plus difficile pour les filles, INSEE.

⁵ Maetz, I. (2016). Parcours et réussite aux diplômes universitaires. Note Flash – Enseignement supérieur & Recherche, 15.

Pour la variable genre, nous pensons que le sexe féminin va influencer positivement la moyenne générale.

1.2. Age

La variable âge peut en effet avoir un impact significatif sur la moyenne générale. Effectivement, une personne plus âgée aura acquis plus de maturité dans la vie, avec une plus grande expérience des périodes d'examens ou de gestion budgétaire.

Un individu plus âgé aura également plus tendance à avoir des objectifs clairs dans sa vie, au contraire d'un individu plus jeune qui peut encore chercher sa voie.

Nous nous attendons ainsi à une corrélation positive entre l'âge et la moyenne académique.

1.3. Santé

La santé d'un étudiant joue un rôle crucial dans ses résultats académiques. Un étudiant en bonne santé aura tendance à avoir une meilleure concentration, plus d'énergie et une motivation supplémentaire pour suivre les enseignements.

De plus, les étudiants en bonne santé ont tendance à maintenir une routine plus régulière, incluant un sommeil adéquat et une alimentation équilibrée, qui sont des facteurs déterminants pour la performance cognitive. À l'inverse, les étudiants souffrant de problèmes de santé peuvent rencontrer des obstacles tels que la fatigue, l'absentéisme ou des difficultés de concentration, ce qui peut nuire à leurs résultats académiques⁶.

⁶ Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation. (2013). La vie étudiante: La santé des étudiants

La santé mentale est également très importante, en effet elle est un élément primordial pour un étudiant dans son cursus pour réussir. Selon l'Organisation mondiale de la Santé (OMS), environ 10 à 20 % des enfants et adolescents connaissent des troubles mentaux. Parmi les plus courants figurent l'anxiété, la dépression, les troubles de l'attention et les troubles du comportement. Les facteurs de risque incluent le stress académique, les pressions sociales, les changements familiaux, et l'exposition aux médias numériques⁷.

Par conséquent, nous nous attendons à observer une influence de la santé sur la moyenne générale : en effet, un étudiant en bonne santé aura tendance à obtenir de meilleures notes.

2. Variables académiques

2.1. Tutorat

Cette variable indique si un étudiant a participé ou non au tutorat durant l'année académique 2023-2024. Le tutorat est un système de soutien éducatif où un tuteur, qu'il s'agisse d'un étudiant avancé ou d'un enseignant, fournit une assistance personnalisée à un étudiant ou un groupe d'étudiants pour les aider à comprendre et maîtriser certains concepts académiques. Ce soutien permet aux étudiants de combler leurs lacunes et, parfois, de bénéficier d'un retour d'expérience précieux.

D'après une étude réalisée par Sandoss Ben Abid-Zarrouk et Marc Weisser, qui compare les étudiants ayant suivi le tutorat à ceux ne l'ayant pas fait, les résultats montrent qu'en 2008, le taux de passage de la première année de licence (L1) à la deuxième (L2) était de 70 % pour les étudiants suivant le tutorat, contre seulement 45 % pour ceux qui n'en bénéficiaient pas (cf. annexe 1).

⁷ Éducacentre (2024). L'impact de santé mentale sur la réussite scolaire

Nous pensons qu'il existe une influence positive du tutorat sur la moyenne générale. Assurément, un étudiant qui suit le tutorat aura plus de chance d'avoir une meilleure moyenne générale.

2.2. Taux d'assiduité

Le taux d'assiduité d'un étudiant, évalué sur une échelle de 0 à 10, peut avoir un impact significatif sur sa réussite académique. Des études montrent qu'une présence régulière en cours est corrélée à de meilleures performances scolaires, en moyenne 10 à 15 % de meilleurs résultats aux examens que ceux qui n'y participaient pas⁸. Nous pouvons affirmer que les étudiants qui assistent systématiquement aux cours sont plus susceptibles de comprendre les concepts abordés et de s'engager activement dans leur apprentissage.

Une faible assiduité peut entraîner un décrochage scolaire. Les étudiants qui manquent régulièrement des cours risquent de se retrouver en difficulté, car ils n'ont pas accès à toutes les informations et les ressources nécessaires à leur réussite⁹. Cela peut créer un cercle vicieux, où le manque de compréhension des sujets étudiés les amène à se désengager davantage, aggravant ainsi leurs résultats académiques.

Il convient également de noter que le taux d'assiduité peut varier en fonction de divers facteurs, tels que les contraintes personnelles, le travail étudiant ou la santé. Les étudiants confrontés à des défis extérieurs peuvent avoir des difficultés à maintenir une assiduité élevée, ce qui peut affecter leur performance scolaire.

Par conséquent, nous nous attendons à observer une corrélation positive entre un taux d'assiduité élevé et les résultats académiques.

⁸ Transkriptor. (2023). La présence aux cours augmente-t-elle les performances académiques ?

⁹ Heather. (2022). L'importance de l'assiduité des élèves dans l'éducation, SeatsSoftware

2.3. Filière d'étude

A l'heure actuelle, nous pouvons affirmer que l'environnement de l'étudiant et notamment le choix de la filière est un facteur crucial dans la moyenne générale. En effet, chaque formation impose ses propres exigences allant de la complexité des concepts exposés à l'intensité de la charge de travail en passant par les compétences requises pour la bonne compréhension du domaine.

Comme nous pouvons le voir dans l'annexe 2, certaines filières comme la santé ou les sciences demandent beaucoup plus de temps de travail. En effet, en moyenne les étudiants en santé travaillent plus de 55 heures hebdomadaires (cf. annexe 2) contre 25 heures pour les étudiants en sociologie¹⁰.

Ainsi, nous avons décidé de demander aux étudiants leur formation, en leur proposant une liste de choix pertinents : Économie et gestion, Droit, Sciences, Baccalauréat, École de commerce, Autre. Après avoir recueilli les réponses, nous avons exclu les étudiants ayant uniquement le Baccalauréat, car leur inclusion risquait de biaiser nos résultats. En effet, ils étaient généralement mineurs, avaient une vie sociale moins développée et ne pouvaient, par conséquent, pas encore travailler.

Pour la variable filière d'étude, nous nous attendons à ce que la variable Économie et gestion impacte positivement la moyenne.

2.4. Temps dédié au travail personnel pour les cours

Cette variable nous permet de distinguer les individus qui travaillent leurs cours à côté de leurs études, et ceux qui ne les travaillent pas suffisamment. En effet, un apprentissage régulier dans l'étude permet de mieux assimiler les concepts abordés en classe et de renforcer

¹⁰ Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation. (2023). Tableaux de réussite et passage par discipline et établissement.

la compréhension des matières. Cela offre également l'opportunité de réaliser des exercices supplémentaires, d'approfondir ses connaissances et de se préparer efficacement aux examens. Plus on consacre de temps à réviser et à travailler sur les sujets traités, plus on augmente ses chances de réussir et d'obtenir de meilleures notes. Ainsi, le lien entre le temps consacré au travail des cours et la performance académique est indéniable, car une gestion proactive de son emploi du temps peut conduire à une amélioration notable de la moyenne générale. Dans l'enquête de Pagis et Farges (2020), nous pouvons remarquer que un étudiant sur trois a une perspective de performance¹¹ et pour réussir, il travaille ses cours à côté des cours magistraux et des travaux dirigés.

Pour cette variable, nous avons décidé de poser la question suivante : “Combien de temps, en moyenne, avez-vous dédié chaque jour à vos études en dehors des cours magistraux et des travaux dirigés durant l'année 2023-2024 ?” Les réponses possibles sont “0-1h”, “1-2h”, “2-3h” et “3h et plus”.

Nous nous attendons à avoir une influence positive du temps à travailler ses cours sur les résultats académiques.

Cependant, nous reconnaissions l'existence d'une endogénéité potentielle entre le temps consacré au travail personnel et la moyenne générale. Autrement dit, plus un étudiant travaille ses cours, meilleure sera sa moyenne ; et inversement, plus sa moyenne est bonne, plus il sera incité à travailler pour la maintenir. Afin de traiter cette possible endogénéité, nous avons décidé d'utiliser trois variables instrumentales : la structure de formation, le temps de sommeil et le temps de trajet.

Nous considérons que la structure de formation peut servir d'instrument car elle influence les méthodes d'apprentissage et la quantité de travail personnel exigée, sans pour autant être directement liée à la moyenne générale. Nous pensons également en effet que le temps de sommeil a un impact direct sur le temps de travail personnel : plus un étudiant dort, mieux il se repose, ce qui améliore sa concentration et réduit sa fatigue, lui permettant ainsi de travailler plus efficacement. De plus, le temps de trajet joue également un rôle, car un

¹¹ Pagis, J., & Farges, G. (2020). Quand des élèves de milieux populaires s'engagent pour leurs études au lycée. Revue Française de Pédagogie, 20.

étudiant qui habite loin de son établissement scolaire disposera de moins de temps pour travailler ses cours.

2.5. Cours particulier

La variable des cours particuliers nous permet de distinguer les étudiants ayant eu accès à ce soutien personnalisé de ceux qui n'en ont pas bénéficié. Les cours particuliers renforcent la compréhension des sujets abordés en classe et comblent les éventuelles lacunes. Les étudiants qui consacrent une partie de leur budget à ce type d'accompagnement peuvent ainsi se préparer plus efficacement aux examens et améliorer leurs performances scolaires. En effet, les étudiants ayant suivi des cours particuliers voient leurs notes augmenter en moyenne de 71,83 % (cf. annexe 3).

Contrairement aux autres variables, nous avons décidé de ne pas restreindre celle-ci à l'année 2023-2024, car une personne qui a suivi des cours particuliers durant son cursus scolaire a plus de chances de réussir et présente en principe moins de lacunes qu'une autre n'en ayant pas suivis.

Nous pensons qu'un étudiant qui suit des cours particuliers est plus susceptible d'avoir une meilleure moyenne, donc le fait de prendre des cours particuliers a un impact positif sur la réussite académique.

Nous suspectons que cette variable est endogène, car si un étudiant a une moyenne faible, il est probable qu'il choisisse de prendre des cours particuliers. En revanche, si sa moyenne est élevée et qu'il suit déjà des cours particuliers, il sera plus enclin à les arrêter. Pour traiter cette possible endogénéité, nous utiliserons le statut social des parents comme variable instrumentale.

En effet, le statut social a un lien direct avec les cours particuliers, puisque si le statut social est “famille aisée”, alors l’individu aura plus tendance à prendre des cours particuliers pour augmenter sa moyenne.

3. Variables financières

3.1. Stress Financier

Cette variable nous permet de distinguer les étudiants qui ressentent un stress financier important et ceux qui n’en sont pas affectés. En effet, le stress financier peut affecter de manière significative la concentration des étudiants en cours, entraînant une assimilation réduite des informations. Par conséquent, les étudiants stressés financièrement doivent consacrer plus de temps à revoir leurs cours pour rattraper leur retard, en plus de devoir souvent travailler parallèlement à leurs études pour subvenir à leurs besoins. Cette double charge peut avoir des répercussions sur leur performance académique, leur bien-être, et même leur santé mentale¹².

Ainsi, le lien entre la pression financière et les résultats académiques est notable : plus un étudiant se trouve en difficulté financière, plus il est susceptible de faire face à des obstacles qui peuvent freiner sa réussite. Des données internationales montrent l’ampleur de ce phénomène. Le rapport annuel de 2015 du Center for Collegiate Mental Health (CCMH) indique que 37,9 % des étudiants universitaires aux États-Unis qualifient leur situation financière de source de stress fréquent ou constant. De même, une enquête au Royaume-Uni en 2013 révèle que 47 % des étudiants ressentent une pression financière¹³.

¹² Lassarre, D., Giron, C., & Paty, B. (2003). Stress des étudiants et réussite universitaire : les conditions économiques, pédagogiques et psychologiques du succès. *Revue française des affaires sociales*, 32(4), 669-691.

¹³ Lucia Romo, Stéphanie Nann, Elisabetta Scanferla, José Esteban, Hélène Riazuelo et Laurence Kern. (2019). La santé des étudiants à l’université comme déterminant de la réussite académique. Volume 40, numéro 2, 2019, p. 187–202

Nous avons décidé de demander aux étudiants sur une échelle de 1 à 10 leur niveau de stress financier. Nous nous attendons à observer une corrélation négative entre le niveau de stress financier et les performances académiques.

Cependant, nous reconnaissions l'existence d'une potentielle endogénéité entre le stress financier et la moyenne générale. En d'autres termes, plus un étudiant subit un stress financier, plus il sera amené à travailler pour compenser son manque de ressources, mais inversement, une bonne moyenne pourrait également motiver certains étudiants à moins travailler pour des raisons financières. Pour traiter cette possible endogénéité, nous avons décidé d'utiliser quatres variables instrumentales : les aides de la CAF, le loyer 2023-2024, le temps de sommeil en 2023-2024 et les commandes Uber Eats.

En effet, nous pensons que les aides de la CAF allègent le stress financier pour les bénéficiaires ce qui donne plus de temps pour travailler ses cours. Ce choix s'explique par le fait que payer un loyer exerce une pression financière directe sur l'étudiant, ce qui peut augmenter son niveau de stress sans être directement lié à sa performance académique. En effet, les étudiants qui paient un loyer ont généralement plus de dépenses fixes, ce qui peut impacter leur bien-être financier et leur concentration, indépendamment de leurs résultats scolaires. Cet instrument permet ainsi de mieux isoler l'effet du stress financier sur les résultats académiques en tenant compte de cette charge économique spécifique.

Le temps de sommeil est également un facteur pertinent, car un bon sommeil réduit la fatigue, améliore la concentration, et permet aux étudiants de gérer plus efficacement leur travail personnel¹⁴. Un étudiant bien reposé peut en effet mieux résister au stress et étudier dans de meilleures conditions, indépendamment de ses résultats académiques.

Nous avons également décidé de choisir les commandes Uber Eats, car le fait d'utiliser fréquemment cette application peut révéler soit une certaine aisance financière ou, à l'inverse, un budget serré en raison du coût élevé de ces services. En moyenne, une commande sur Uber Eats coûte environ 33,94 USD¹⁵(environ 30,50 EUR) par commande.

¹⁴ Pedrelli, L., & Peretti, N. (2013). L'impact du stress sur la performance et la mémoire des élèves. Revue Française de Pédagogie, 13, 13-25.

¹⁵ Beyrouthy, L. (2024). Average food delivery spend per order in the U.S. Statista.

3.2. Budget transport

La variable du budget transport peut avoir un impact significatif sur la moyenne générale, car un individu qui dépense beaucoup pour ses déplacements est susceptible de ressentir davantage de stress financier. En effet, s'il doit allouer une somme importante pour se déplacer, il pourrait réduire ses déplacements et, par conséquent, manquer des activités en lien avec sa formation (tutorat, cours non obligatoires, etc.), ou se rendre moins souvent à la bibliothèque.

Nous avons donc posé la question suivante aux étudiants : “Quel était, en moyenne, votre budget transport par mois durant l'année 2023-2024 ? Rappel : le prix d'un ticket de bus à Nantes est de 1,80 €.” Les réponses possibles étaient : “0-15 €”, “15-30 €”, “30-50 €”, et “50 € et plus”. voir annexe (mettre le questionnaire)

Nous nous attendons à observer une influence négative du budget transport sur la moyenne générale. En effet, plus le budget alloué aux transports est élevé, plus il est probable que la moyenne générale soit faible.

3.3. Travail étudiant

Par cette variable, nous visons à démontrer qu'un étudiant qui travaille a moins de chances d'obtenir une moyenne élevée. En 2020, 26 % des étudiants occupaient un emploi en parallèle de leurs études¹⁶, ce qui peut entraîner des conséquences multiples sur leur réussite académique.

Les effets du travail étudiant sont variés. D'une part, un étudiant qui travaille dispose de moins de temps pour réviser ses cours et, dans certains cas, il peut être contraint de manquer des cours pour faire ses heures de travail et subvenir à ses besoins financiers. D'autre part, le fait de combiner études et emploi peut lui permettre de développer des

¹⁶ Vie Publique. (27/03/2023). Jobs étudiants : 146 000 étudiants occupent un emploi hors du cadre de leurs études.

compétences en organisation et en discipline, nécessaires pour gérer ces deux engagements¹⁷. Toutefois, les chiffres montrent un impact significatif sur la réussite : le taux de succès pour les étudiants travaillant plus de 16 heures par semaine est de 37,9 %, tandis que pour ceux qui ne travaillent pas, il atteint 66 % (cf. annexe 4).

De plus, un étudiant qui travaille est souvent plus fatigué, ce qui réduit sa capacité de concentration en cours. Cette fatigue accrue l'oblige à fournir un effort supplémentaire pour maintenir sa réussite académique.

Ainsi, bien que le travail étudiant puisse renforcer certaines compétences transversales, il présente également des défis notables pour les étudiants en termes de disponibilité et de performance académique.

Nous anticipons une influence négative du travail étudiant sur la moyenne générale. En effet, un étudiant qui travaille aura tendance à obtenir une moyenne plus faible.

3.4. Nombre d'heures travaillées

Le nombre d'heures travaillées par semaine est un facteur susceptible d'influencer la moyenne générale pour plusieurs raisons. D'après la littérature, divers seuils d'heures hebdomadaires à partir desquels la moyenne commence à décroître ont été identifiés. Le seuil le plus souvent retenu est celui de 15 heures, comme l'indique Montmarquette¹⁸. Cependant, une étude de Froment¹⁹, menée en 2012, suggère que le seuil critique serait plutôt de 8 heures. En effet, Froment a démontré une différence notable entre les étudiants travaillant moins de 8 heures par semaine et ceux travaillant davantage : en dessous de 8 heures, le taux de réussite est supérieur à la moyenne, atteignant 97,9 % (cf. annexe 5), contre 90,8 % pour ceux qui ne travaillent pas du tout. Cette meilleure réussite peut s'expliquer par une

¹⁷ Beffy, M., Fougère, D., & Maurel, A. (2013). L'impact du travail salarié des étudiants sur la réussite et la poursuite des études universitaires. INSEE.

¹⁸ Montmarquette, C., Viennot-Briot, N., & Dagenais, M. (2007). Dropout, school performance, and working while in school. *The Review of Economics and Statistics*, 89(4), 752-760.

¹⁹ Froment, B. (2012). Les effets du travail salarié en première année universitaire. *Sociologies*.

organisation et une autonomie accrues chez les étudiants travaillant quelques heures par semaine.

En revanche, pour les étudiants travaillant entre 8 et 16 heures, le taux de réussite est de seulement 85,4 %, et au-delà de 16 heures, il diminue encore à 78,4 %. La probabilité de réussite diminue progressivement avec l'augmentation des heures de travail hebdomadaire. Les étudiants sont ainsi confrontés à un arbitrage, comme le souligne l'étude de Lévy-Garboua de 1976²⁰, entre le besoin de travailler et l'impact de ce travail sur leur performance académique.

De plus, travailler parallèlement aux études entraîne souvent de la fatigue supplémentaire, ce qui peut affecter la concentration en cours et nécessiter plus de temps pour rattraper les informations manquées, augmentant le risque de stress pour certains.

En résumé, nous anticipons une corrélation négative entre le temps de travail hebdomadaire et la moyenne générale : plus un étudiant consacre d'heures à un emploi en dehors de ses cours, plus sa moyenne risque de baisser.

3.5. Budget loisirs

Avec cette variable, nous cherchons à évaluer le budget mensuel des étudiants pour leurs loisirs, incluant le shopping, les sorties dans des bars, le cinéma, le sport, etc.

A travers ce facteur, nous souhaitons démontrer que des dépenses excessives allouées aux loisirs peuvent expliquer la faible moyenne générale observée pour certains étudiants. Nous supposons que si ces derniers consacrent davantage d'argent à leurs loisirs, ils passeront inévitablement plus de temps à en profiter. Par conséquent, ils auront moins de temps à

²⁰ Lévy-Garboua, L. (1976). Les demandes de l'étudiant ou les contradictions de l'université de masse. *Revue française de sociologie*, 17(1), 53.

consacrer à leurs études, et s'ils manquent d'organisation, leurs résultats scolaires risquent de baisser.

Pour le budget loisir, nous avons posé la question suivante aux étudiants : “Quelle est votre estimation de vos dépenses mensuelles dans les loisirs durant l'année 2023-2024 ? * Magasins, bar, cinéma, sport, etc. **”, avec comme réponse possible : “0-100€”, “100-250€” et “250€ et plus”.

Nous pensons qu'un budget excessif consacré aux loisirs influence négativement la moyenne générale. En effet, un étudiant qui dépense beaucoup pour ses loisirs est susceptible d'avoir une moyenne plus basse.

3.6. Argent autre que la bourse, la CAF et le travail étudiant

Pour cette variable, nous avons cherché à identifier les individus recevant un soutien financier autre que les bourses, les aides de la CAF, ou les revenus issus d'un travail étudiant. Il s'agit donc des étudiants bénéficiant d'une aide financière de leur famille ou de revenus provenant d'investissements (actions, immobilier, etc.). En 2000, environ 37,1% des étudiants bénéficiaient d'une aide financière de leurs parents²¹.

Selon une étude de Verley et Zilloniz (2011)²², les étudiants bénéficiant d'un soutien financier parental ont plus de chances de poursuivre des études longues. En effet, lorsque les étudiants reçoivent de l'argent de leurs parents ou de leurs investissements, ils subissent moins de stress financier, ce qui les incite moins à travailler parallèlement à leurs études. Ils disposent ainsi de davantage de temps pour se consacrer pleinement à leurs études, ce qui devrait, en théorie, favoriser de meilleurs résultats académiques.

²¹ Lassarre, D., Giron, C., & Paty, B. (2003). Stress des étudiants et réussite universitaire : les conditions économiques, pédagogiques et psychologiques du succès. Revue française des affaires sociales, 32(4), 669-691.

²² Verley E. et Zilloniz S., « Fragilités économiques, fragilités studieuses », in Observatoire national de la vie étudiante (OVE), Les mondes étudiants –enquête conditions de vie 2010, La documentation française, 2011, pp. 117-136.

Nous pensons donc que des revenus financiers externes (autres que les bourses, les aides CAF, ou les revenus de travail étudiant) influencent positivement la moyenne générale.

3.7. Bourse

Une bourse universitaire est un soutien financier accordé aux étudiants pour les aider à couvrir les frais de scolarité, de logement, de livres et autres dépenses liées à leurs études. Les bourses peuvent être attribuées sur la base du mérite académique, de la situation financière, des compétences sportives, ou d'autres critères spécifiques²³.

Nous observons que les étudiants boursiers qui ont plus de problèmes financiers réussissent davantage leur première année que les étudiants non boursiers. Par exemple, parmi les étudiants confrontés à des difficultés financières, 43 % des boursiers ont réussi leur licence contre 33 % des non-boursiers (cf. annexe 6). Inversement, nous constatons que, parmi les étudiants ayant moins de difficultés qui ont réussi leur première année, le pourcentage des boursiers est plus faible que celui des non-boursiers.

Au regard des résultats, nous considérons que l'attribution de bourses a une influence positive sur la réussite des étudiants en situation financière précaire.

3.8. Nombre de restaurants par mois

À travers cette question sur le nombre moyen de sorties au restaurant par mois, nous souhaitons démontrer qu'une fréquentation élevée des restaurants peut refléter une mauvaise gestion financière. En France, en 2023, le coût moyen d'un repas au restaurant s'élevait à environ 15 €, ce qui représente un budget non négligeable pour les étudiants²⁴.

²³ Service Public. (2024). Étudiant : bourse sur critères sociaux.

²⁴ Combien-coute.net.(2024). Prix d'un repas au restaurant en France.

Aller souvent au restaurant peut entraîner des dépenses importantes, augmentant ainsi le risque de stress financier. De plus, si l'individu privilégie des choix peu équilibrés, cette habitude pourrait également nuire à sa santé²⁵. Par ailleurs, des sorties tardives au restaurant peuvent conduire à une réduction des heures de sommeil, ce qui risque d'affecter la concentration et la performance académique.

Nous anticipons donc une corrélation négative entre la fréquence des repas au restaurant et la moyenne générale. En effet, plus un étudiant fréquente les restaurants, plus sa moyenne pourrait en pâtir.

C. Récapitulatif du modèle

En somme, notre étude comporte 17 variables : 1 variable expliquée (la moyenne générale durant l'année 2023-2024), ainsi que 16 variables explicatives sans compter les 7 variables instrumentales qui seront utilisées pour contrer de potentiels problèmes d'endogénéité (cf. annexe 42).

Tableau 1 : Tableau récapitulatif de la revue de littérature sur la moyenne

Variables quantitatives	Correspondance sur R	Corrélation attendue
Moyenne	MOYENNE	Variable expliquée
Taux d'assiduité	ASSISUITE	Positive
Nombre d'heures travaillées	HEMPLOI	Négative
Nombre de restaurants par mois	RESTAURANT	Négative
Age	AGE	Indéterminée
Stress Financier	STRESS	Négative

²⁵ Harris Interactive. (2017). Les étudiants et l'alimentation : Rapport Allo Resto.

Variables qualitatives	Correspondance sur R	Modalité suspectées d'impacter positivement la moyenne générale
Formation	FORMATION	Économie et gestion
Nombre d'heures de révisions	REVISIONS	2 - 3h
Cours particuliers	PARTICULIERS	Oui
Tutorat	TUTORAT	Oui
Bourse	BOURSE	Oui
Job étudiant	EMPLOI	Non
Argent autre que la CAF, bourse et job étudiant	ARGENT	Oui
Dépenses mensuelles dans les loisirs	DEPENSES	0 - 100€
Budget transport	TRANSPORT	0 - 15 €
Genre	GENRE	Femme
Santé	SANTE	Non

Instruments	Correspondance sur R	Variable associée
Aide CAF	CAF	Stress Financier
Commande Uber Eats	UBER	Stress Financier
Loyer 2023-2024	LOGEMENT	Stress Financier
Temps de sommeil	SOMMEIL	Stress Financier & Temps dédié au travail personnel des cours
Temps de trajet	TRAJET	Temps dédié au travail personnel des cours
Statut social	STATUT	Cours particulier

Structure	STRUCTURE	Temps dédié au travail personnel des cours
-----------	-----------	--

III - Méthode

Dans le but de modéliser la moyenne générale durant l'année 2023-2024 en fonction des différentes variables énumérées ci-dessus, nous avons réalisé une régression linéaire à partir des données recueillies à l'aide d'un questionnaire. Ainsi, dans cette partie, nous présenterons d'abord le questionnaire que nous avons élaboré (A), puis nous aborderons la méthode utilisée (B).

A. Méthode de recueil de données

Pour recueillir les données nécessaires à notre modélisation, nous avons opté pour la mise en place d'un questionnaire en utilisant la plateforme Google Forms. Ce questionnaire a été diffusé sur une période de deux semaines (du 16 au 28 octobre), en partageant son lien via la messagerie universitaire et notre réseau personnel, auprès des étudiants de l'IAE, des sciences, ainsi que ceux présents sur nos réseaux sociaux.

Concernant notre questionnaire, nous l'avons structuré en trois parties : la première portait sur les caractéristiques du répondant, la deuxième sur sa scolarité en 2023-2024, et la troisième sur ses dépenses budgétaires.

Tableau 2 : Tableau représentant les questions et les modalités de réponses

Variables quantitatives	Question associée	Modalités de réponse
MOYENNE	Quelle était votre moyenne pour l'année 2023-2024 ?	Réponse libre (nombre)
ASSIDUITE	Quel était, selon vous, votre taux d'assiduité aux cours magistraux et aux travaux dirigés sur une échelle de 1 à 10 ?	Échelle de 0 (Aucune) à 10 (Extrêmement élevée)

HEMPLOI	Si c'était le cas, combien d'heures, en moyenne, par semaine y consaciez-vous ?	Réponse libre (nombre)
RESTAURANT	Quel était, en moyenne, le nombre de fois par mois où vous alliez au restaurant ou au fast-food, durant l'année 2023-2024 ?	Réponse libre (nombre)
AGE	Quel âge avez-vous ?	Réponse libre (nombre)
STRESS	Quel était votre degré de stress financier durant l'année 2023-2024 ?	Échelle de 0 (Aucun) à 10 (Extrêmement élevé)

Variables qualitatives	Question associée	Modalités de réponse
FORMATION	Quelle formation faisiez-vous en 2023-2024 ?	Économie et gestion Droit Sciences Baccalauréat Autres
REVISIONS	Combien de temps, en moyenne, avez-vous dédié chaque jour à vos études en dehors des cours magistraux et des travaux dirigés durant l'année 2023-2024 ?	0 - 1h 1 - 2h 2 - 3h 3h et plus
PARTICULIERS	Avez-vous déjà pris des cours particuliers durant votre formation ?	Oui Non

TUTORAT	Avez-vous assisté au tutorat durant l'année 2023-2024 ?	Oui Non
BOURSE	Bénéficiez-vous d'une bourse durant l'année 2023-2024 ?	Oui Non
EMPLOI	Aviez-vous un emploi étudiant durant l'année 2023-2024 ?	Oui Non
ARGENT	Receviez-vous de l'argent en dehors de votre travail étudiant, de votre bourse et des aides de la CAF durant l'année 2023-2024 ?	Oui Non
DEPENSES	Quelle est votre estimation de vos dépenses mensuelles dans les loisirs durant l'année 2023-2024 ?	0 - 100 € 100 - 250 € 250 € et plus
TRANSPORT	Quel était, en moyenne, votre budget transport par mois, durant l'année 2023-2024 ?	0 - 15 € 15 - 30 € 30 - 50 € 50 € et plus
GENRE	Quel est votre genre ?	Homme Femme
SANTE	Avez-vous eu des problèmes de santé durant l'année 2023-2024 ?	Oui Non

Instruments	Question associée	Modalités de réponse
CAF	Perceviez-vous des aides de la CAF durant l'année 2023-2024 ?	Oui Non
UBER	Combien de fois, en moyenne, commandiez-vous Uber Eats par mois, durant l'année 2023-2024 ?	Réponse libre (nombre)
LOGEMENT	Durant l'année 2023-2024, payiez-vous un logement ?	Oui Non
SOMMEIL	En 2023-2024, combien d'heures dormiez-vous par nuit en moyenne ?	Réponse libre (nombre)
TRAJET	Combien de temps mettiez-vous pour vous rendre à votre établissement pendant l'année 2023-2024 ?	Réponse libre (nombre)
STATUT	Comment décrivez-vous le statut social de vos parents ?	Classe aisée Classe moyenne Classe populaire
STRUCTURE	Quelle était la structure de votre formation durant l'année 2023-2024 ?	Université Ecole de commerce Ecole d'ingénieur Lycée Autres

B. Méthode économétrique utilisée

En premier lieu, nous analysons notre base de données grâce à des représentations graphiques et des statistiques. Nous commençons par une analyse univariée, en examinant d'abord les variables quantitatives, notamment pour identifier les valeurs aberrantes ou atypiques, puis les variables qualitatives. Nous procérons ensuite à une analyse bivariée. Par la suite, nous débutons par l'étude de deux variables quantitatives en évaluant en particulier leur corrélation pour détecter et éviter la multicolinéarité. Ensuite, nous examinons ensuite les relations entre variables qualitatives, et terminons par l'analyse des interactions entre une variable quantitative et une qualitative.

En second lieu, nous utilisons diverses méthodes de régression linéaire pour examiner la relation entre la variable dépendante et les variables explicatives, en étudiant leur influence et en effectuant des prévisions. L'estimation du modèle débute par la sélection des variables à l'aide de la méthode stepwise.

Les variables explicatives sont, dans un premier temps, supposées exogènes. Nous commençons par appliquer la méthode des moindres carrés ordinaires. Plusieurs tests sont réalisés afin de confirmer que les conditions nécessaires à la validation du modèle sont respectées. En effet, nous vérifions la normalité des résidus, la forme fonctionnelle, l'absence de multicolinéarité, l'absence d'observations influentes, ainsi que l'homoscédasticité des erreurs. Si nécessaire, les méthodes des moindres carrés pondérés ou quasi généralisés, ou encore celle de White, sont appliquées pour corriger l'hétérosclélasticité.

Dans un second temps, nous considérons les variables explicatives suspectées d'endogénéité. La méthode des doubles moindres carrés est réalisée à l'aide de nos instruments jugés appropriés sur le plan économique. Nos deux premiers tests vérifient la pertinence et la validité des instruments. Un dernier test évalue leur endogénéité et permet de sélectionner la méthode appropriée. Nos résultats peuvent ensuite être interprétés à partir des tests évaluant la significativité globale du modèle et celle spécifique à chaque variable explicative. Enfin, nous évaluons la capacité du modèle à prédire notre variable dépendante.

IV - Analyse de la base de donnée

Après avoir soumis notre questionnaire et récolté les données, nous avons pu premièrement nous questionner sur la représentativité de notre échantillon de répondants (A) puis nous avons commencé l'analyse descriptive à l'aide de graphiques et de statistiques, en procédant d'abord à une analyse univariée (B), puis à une analyse bivariée des relations entre variables quantitatives et qualitatives (C).

A. Représentativité de l'échantillon

Pour vérifier la représentativité de l'échantillon comprenant 150 observations, nous devons vérifier si la variable GENRE est représentative ou non.

Tableau 3 : Répartition des étudiants en France vs observé du genre français

	Hommes	Femmes	Total
Répartition théorique	44,1	55,9	100
Échantillon théorique	66	84	150
Échantillon observé	68	82	150

Le test du khi-deux d'homogénéité a pour objectif de vérifier si les distributions observées (dans ce cas, les proportions des hommes et des femmes dans l'échantillon) sont similaires à la répartition théorique dans différents groupes.

L'hypothèse nulle stipule que la répartition observée des hommes et des femmes dans l'échantillon est homogène par rapport à la répartition théorique.

Calcul des fréquences théoriques :

- pour les hommes : $150 * 44,1 / 100 = 66,2$
- pour les femmes : $150 * 55,9 / 100 = 83,9$

Calcul de la statistique de test :

- Pour les hommes : $(68 - 66,2)^2 / 66,2 = 0,049$
- Pour les femmes : $(82 - 84)^2 / 84 = 0,043$

La statistique de test est donc : $\text{khi-2} = 0,049 + 0,043 = 0,092$

Calcul du degré de liberté : $\text{ddl} = 2 - 1 = 1$

Pour un seuil de risque de $\alpha = 0,05$ et un degré de liberté de 1, la valeur critique du khi-2 est:

3,841

Nous avons $0,092 < 3,841$.

La statistique de test étant bien inférieure à la valeur critique, nous ne rejetons pas l'hypothèse nulle et concluons que la répartition observée des hommes et des femmes dans l'échantillon ne diffère pas de manière significative de la répartition théorique.

Ainsi, pour un seuil de risque de 5 %, il apparaît que la répartition par genre dans notre échantillon est conforme à la répartition théorique des étudiants en France.

B. Statistiques descriptives univariées

Désormais, nous allons procéder à une analyse descriptive univariée des variables explicatives et expliquée. Cette étape est cruciale pour nous permettre de mesurer la dispersion des données selon chaque variable. En effet, une analyse univariée a pour but de décrire et mesurer la répartition des données d'une seule variable. Nous débuterons par l'étude des variables quantitatives (1), suivie de l'examen des valeurs atypiques dans ces variables (2), et enfin, nous analyserons les variables qualitatives (3).

1. Variables quantitatives

Les statistiques descriptives pour nos variables quantitatives ont pour but de mesurer les tendances, les relations et les distributions, en particulier grâce à de nombreux indicateurs statistiques allant de la moyenne à l'écart-type.

Figure 1: Distribution des valeurs de la variable expliquée

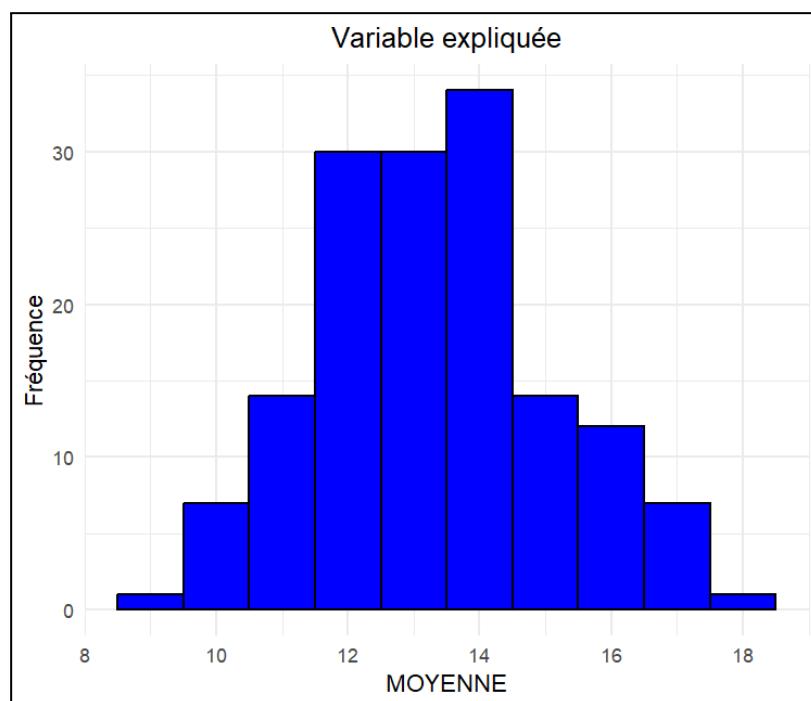
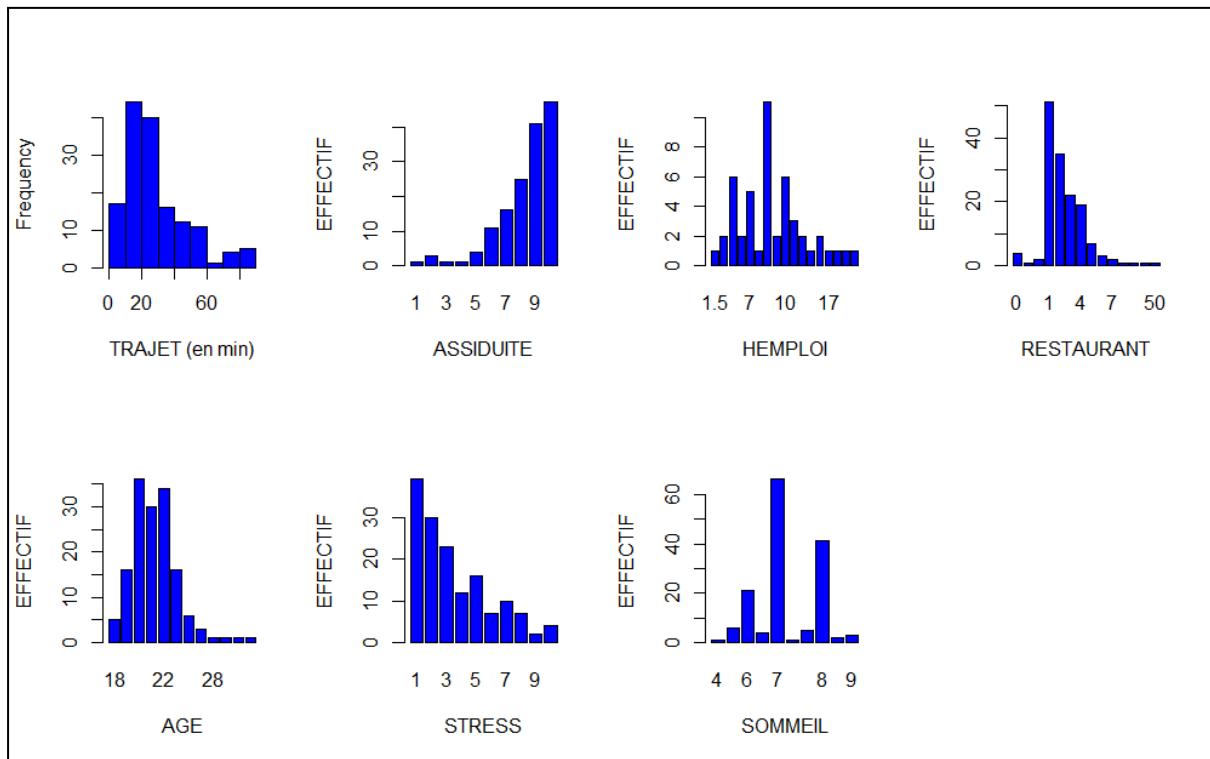


Figure 2 : Distribution des valeurs des variables quantitatives



Avec ces graphiques, nous pouvons dans la mesure du possible identifier la tendance de la distribution des variables quantitatives. Cependant, ces graphiques ne sont pas précis, nous avons donc besoin d'un tableau avec des indicateurs statistiques (moyenne, médiane, minimum...).

Tableau 4 : Statistiques descriptives des variables quantitatives

Variables	Min	1Q	Médiane	Moyenne	3Q	Max	Sd
MOYENNE	9	12,06	13,5	13,45	14,5	17,53	1,78
ASSIDUITE	1	8	9	8,34	10	10	1,84
HEMPLOI	1,5	7	8	9,479	10,5	30	NA
RESTAURANT	0	1	2	2,8	3	50	4,5
AGE	18	20	21	21,41	22	40	2,58
STRESS	1	1	3	3,48	5	10	2,4

SOMMEIL	4	7	7	7,1	8	9	0,87
TRAJET	5	15	30	31,63	40	90	19,88

(cf. annexe 7 et 8)

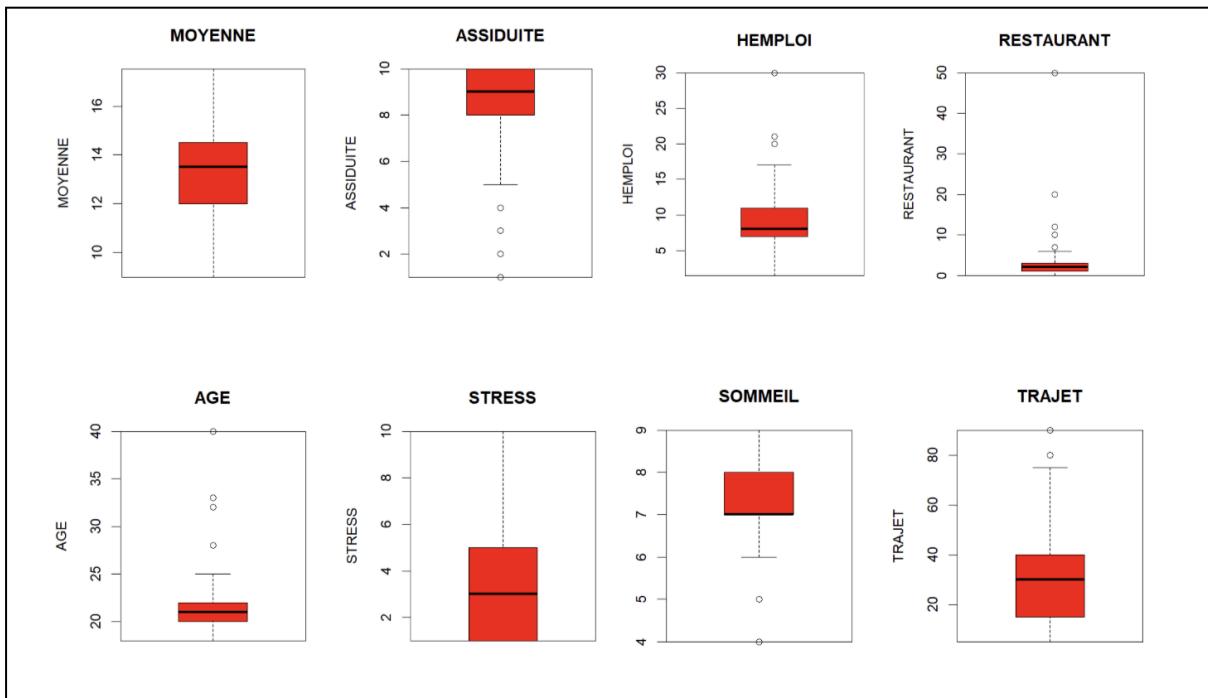
Tout d'abord, nous pouvons observer que les médianes ainsi que les moyennes des variables quantitatives sont proches. Cela suggère que les données de ces variables suivent une distribution symétrique, ce qui implique par conséquent un impact limité des valeurs atypiques.

Nous interprétons à présent les écarts-types de chacune de nos variables afin de savoir si elles sont homogènes ou non. Un écart-type élevé indique une grande variabilité des données par rapport à la moyenne, ce qui explique qu'une variable soit hétérogène. Au contraire, un écart-type réduit traduit la faible volatilité des données par rapport à la moyenne, ce qui implique qu'une variable soit homogène. Nous constatons que c'est le cas de toutes nos variables quantitatives sauf pour RESTAURANT. En effet, cela est sans doute dû à des variables atypiques comme nous pouvons le voir avec la valeur maximum de 50.

2. Atypicité des variables quantitatives

Dans la continuité de notre analyse descriptive, nous examinons la possibilité de valeurs atypiques dans les variables quantitatives. Cette étude vise à prévenir l'influence significative qu'elles peuvent avoir sur les statistiques descriptives de nos variables quantitatives.

Figure 3: Boxplot des variables explicatives quantitatives et de la variable expliquée



Grâce à ces graphiques, nous pouvons visualiser les possibles valeurs atypiques dans nos variables quantitatives. Ainsi, nous observons des points potentiellement atypiques sur les valeurs suivantes : ASSIDUITE, HEMPLOI, RESTAURANT, AGE, SOMMEIL et TRAJET. Le nombre de points visibles dans les boxplots de ces 6 variables étant supérieur à 1 et inférieur à 10, nous pouvons appliquer le test de Rosner pour chaque variable (cf. annexe 9, 10, 11, 12, 13, 14 et 15).

Tableau 5 : Tests des valeurs atypiques

Variables	Nombre de valeurs atypiques potentielles	Nombre de valeurs atypiques réelles	Valeurs atypiques	Test appliqué	Individus Atypiques
ASSIDUITE	4	3	1, 2 et 3	Rosner	39, 44, 48 113 et 128
HEMPLOI	3	2	21 et 30	Rosner	131* et 118
RESTAURAN	5	4	10, 12, 20 et	Rosner	25, 53, 58 et

T			50		88
AGE	4	4	28, 32, 33 et 40	Rosner	58, 65, 67 et 117
SOMMEIL	2	1	4	Rosner	74
TRAJET	2	0	-	Rosner	-

* Individu atypique détecté lors de la deuxième vérification de l'absence d'observations atypiques (individu 116 dans le seconde base).

Ainsi, avec les tests effectués, nous avons décidé d'éliminer les valeurs atypiques de notre étude. Nous passons, par conséquent, de 150 à 135 individus. Pour cette raison, il est important de procéder à une nouvelle analyse des statistiques descriptives de nos variables quantitatives.

Figure 4: Distribution des valeurs de la variable expliquée après retrait des variables atypiques

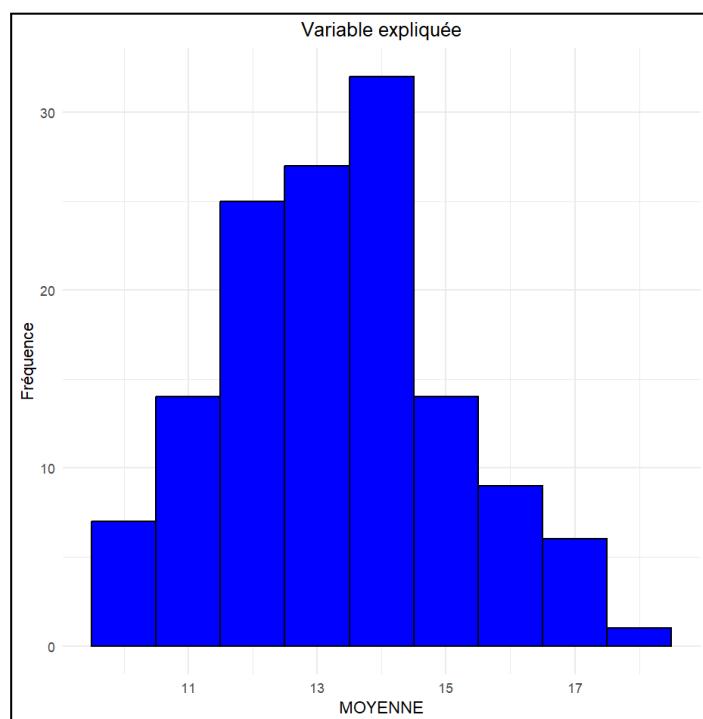


Figure 5 : Distribution des valeurs des variables quantitatives après retrait des variables atypiques

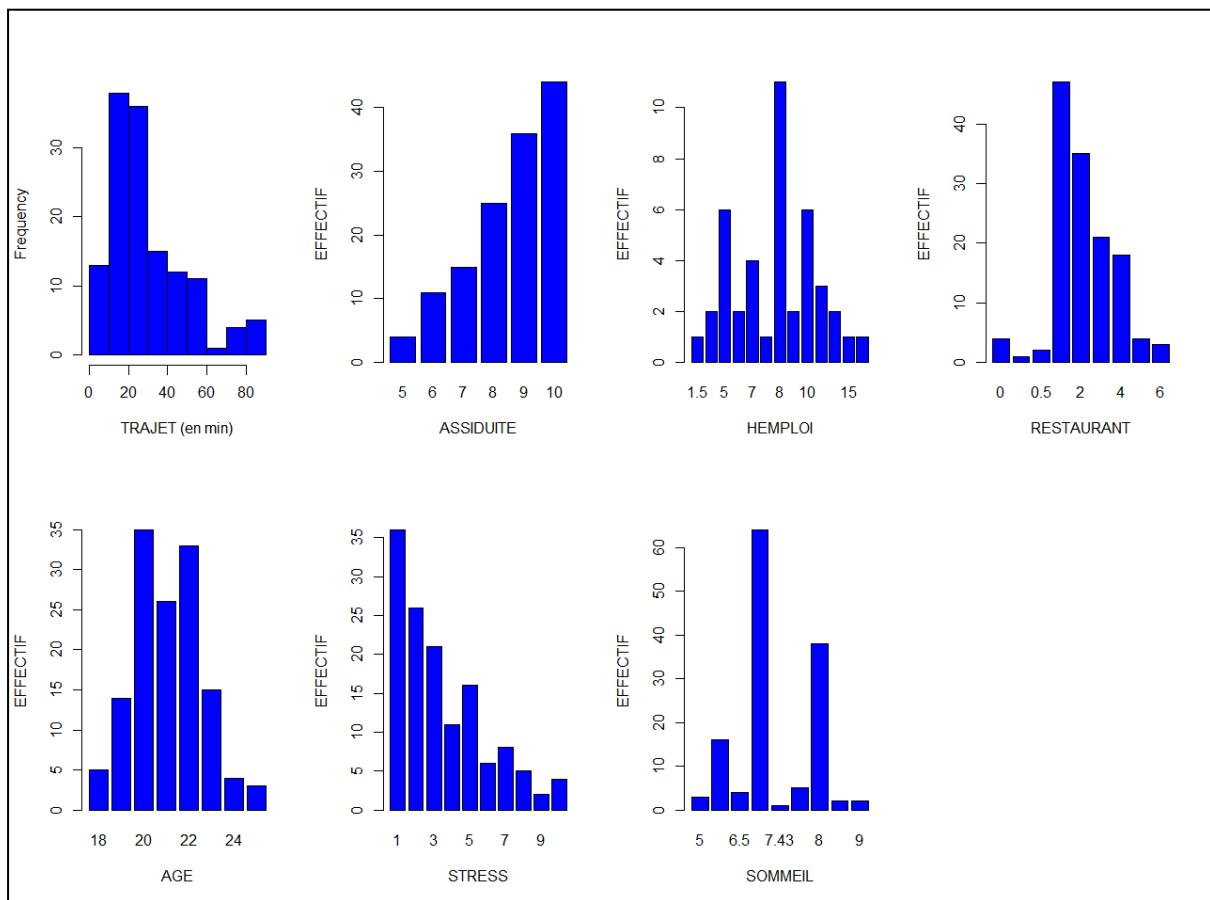


Tableau 6 : Statistiques descriptives des variables quantitatives après retrait des valeurs atypiques

Variables	Min	1Q	Médiane	Moyenne	3Q	Max	Sd
MOYENNE	9	12,06	13,5	13,45	14,5	17,53	1,75
ASSIDUITE	5	8	9	8,556	10	10	1,40
HEMPLOI	1,5	6,25	8	8,2	10	17	NA
RESTAURANT	0	1	2	2,158	3	6	1,34
AGE	18	20	21	21,07	22	25	1,52
STRESS	1	1	3	3,48	5	10	2,43

SOMMEIL	5	7	7	7,18	8	9	0,75
TRAJET	5	20	30	32,98	40	90	20,27

(cf. annexe 16 et 17)

Nous pouvons noter, la baisse significative de l'écart-type pour la variable RESTAURANT, ainsi qu'une légère baisse pour AGE. Cela est dû à la suppression des valeurs atypiques qui étaient extrêmes dans ce cas là. Les moyennes et les médianes n'ont que très peu changé et restent cohérentes.

La suppression de ces valeurs atypiques nous a permis de préserver la représentativité, et de réduire l'influence de celles-ci, le tout sans altérer de manière significative notre première analyse statistique.

Après toutes ces analyses, nous avons décidé de supprimer la variable HEMPLOI car celle-ci ne présentait que 42 observations sur 135. En effet, ceux qui ont répondu qu'ils ne travaillaient pas n'ont pas répondu à cette question. Cela pourrait compromettre l'interprétation et la précision des corrélations et des régressions, justifiant sa mise à l'écart de l'analyse.

3. Variables qualitatives

À présent, nous allons effectuer l'étude descriptive des variables qualitatives de notre modèle.

Tableau 7 : Tableau des regroupements effectués pour les modalités des variables qualitatives

Variables	Modalités initiales	Modalités finales
FORMATION	Économie et gestion Droit Sciences Baccalauréat Autres	Économie et gestion Autres
REVISIONS	0 - 1h 1 - 2h 2 - 3h 3h et plus	0 - 1h 1 - 2h 2h et plus
DEPENSES	0 - 100 € 100 - 250 € 250 € et plus	0 - 100 € 100 € et plus
TRANSPORT	0 - 15 € 15 - 30 € 30 - 50 € 50 € et plus	0 - 15 € 15 - 30 € 30 € et plus
STRUCTURE (instrument)	Université Ecole de commerce Ecole d'ingénieur Lycée Autres	Université Autres

Concernant FORMATION, nous avons fait le choix de tout regrouper dans Autres parce que la formation Économie et gestion représente 76 % des répondants. Par conséquent, pour obtenir de meilleurs résultats, nous avons pris la décision de regrouper Droit, Sciences, Baccalauréat et Autres.

Pour REVISIONS, nous avons fusionné “2 - 3h” et “3h et plus” parce que “3h et plus” représente seulement 8,8 % des répondants.

En ce qui concerne TRANSPORT, nous avons décidé de rassembler “30 - 50 €” et “50 € et plus” parce que “50 € et plus” représente 17,5 % et “30 - 50 €” représente 15,9 %.

Pour finir, l'instrument STRUCTURE, nous avons regroupé en seulement Université et Autres car après analyse des réponses l'Université représente 80 % de l'échantillon. Nous avons également décidé d'enlever tous ceux qui étaient au lycée l'année dernière parce qu'en général les élèves au lycée ne font pas de sortie seul au restaurant, n'ont pas de dépenses annexes, et ont généralement une meilleure moyenne. Cela aurait pu entraîner un biais dans nos résultats.

Après avoir regroupé ces modalités nous pouvons analyser et comprendre les caractéristiques de notre échantillon.

Tableau 8 : Distribution des variables qualitatives

FORMATION		RÉVISIONS		PARTICULIERS		TUTORAT	
Économie et gestion	114	0 - 1h	53	Non	127	Non	101
Autres	21	1 - 2h	50	Oui	8	Oui	34
		2h et plus					
BOURSE	92	EMPLOI	93	LOGEMENT	63	ARGENT	75
Non	92	Non	93	Non	63	Non	75
Oui	43	Oui	42	Oui	72	Oui	60
UBER		DEPENSES					
Non	98	0 - 100 €	80	Non	98	0 - 100 €	80
Oui	37	100 € et plus	55	Oui	37	100 € et plus	55
CAF		TRANSPORT		GENRE		STATUT	
Non	61	0 - 15 €	28	Femme	74	Classe moyenne	106
Oui	74	15 € - 30 €	61	Homme	61	Classe populaire	14
		30 € et plus				Classe aisée	
SANTÉ							
Non	111			Non	111		
Oui	24			Oui	24		

(cf. annexe 18)

Ce tableau de la distribution des variables qualitatives, nous permet de remarquer que certaines variables sont dominées par des modalités. Nous constatons que la variable FORMATION est dominée par la modalité “Économie et gestion” avec 114 réponses sur 135, ou encore que la variable SANTE est dominée par la modalité “Non” avec 111 réponses sur 135. Cependant, certaines variables sont homogènes, comme la variable LOGEMENT où l’on retrouve la modalité “Oui” avec 72 réponses et la modalité “Non” avec 63.

C. Statistiques descriptives bivariées

Nous allons désormais aborder l’analyse bivariée, qui consiste à explorer les variations d’une variable en fonction d’une autre afin de comprendre leur relation. Nous réaliserons ainsi trois analyses en fonction de la nature de la variable : une entre les variables quantitatives (1), une deuxième entre les variables qualitatives (2) et une dernière entre les variables quantitatives et qualitatives (3).

1. Analyse des variables Quantitatives - Quantitatives

Pour commencer notre analyse, nous allons examiner les relations entre les variables explicatives quantitatives et la variable à expliquer à travers le tableau 9 (cf. annexe 19).

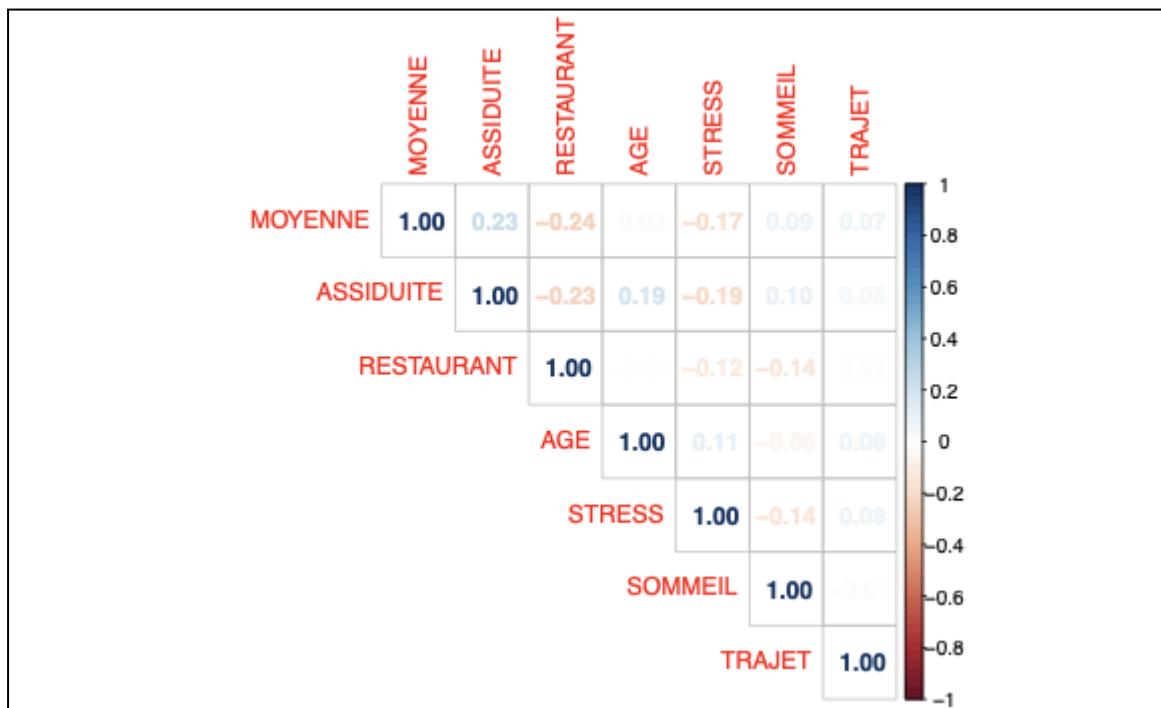
Tableau 9 : Relation attendue (selon la littérature) et observée dans nos données

Variables quantitatives	Relation attendue	Relation observée
ASSISUITE	Positive	Positive
RESTAURANT	Négative	Négative
AGE	Indéterminée	Indéterminée
STRESS	Négative	Négative
SOMMEIL	-	Indéterminée
TRAJET	-	Indéterminée

(cf. annexe 19, 20, 21, 22, 23 et 24)

Ce tableau récapitulatif nous permet de conclure que les relations observées confirment les hypothèses formulées à partir de la littérature. De plus, celui-ci nous fournit des informations sur la direction des relations entre les variables quantitatives et la variable à expliquer. En effet, nous remarquons qu'une augmentation de l'assiduité est corrélée à une augmentation de la moyenne. Concernant le stress, qui a une relation négative, une augmentation du stress financier est corrélée à une diminution de la moyenne.

Figure 6: Corrélation des variables quantitatives avec la moyenne



(cf. annexe 25, 26 et 27)

Nos variables quantitatives semblent peu corrélées entre elles, nous ne relevons pas de problème de multicolinéarité. Nous pouvons par conséquent toutes les prendre en compte pour notre étude.

2. Analyse des variables Qualitatives - Qualitatives

Nous avons débuté cette analyse en nous intéressant aux relations entre certaines variables qualitatives pertinentes, à travers des tableaux de contingence (cf. annexe 28) et des cartes de points chauds (annexe 29).

ARGENT STATUT	Oui	Non	PARTICULIERS TUTORAT	Oui	Non	LOGEMENT CAF	Oui	Non	BOURSE EMPLOI	Oui	Non
Classe populaire	1	13	Oui	6	28	Oui	65	9	Oui	19	23
Classe moyenne	50	56	Non	2	99	Non	7	54	Non	24	69
Classe aisée	9	6									

Il apparaît que certaines variables semblent être liées entre elles, comme c'est le cas pour CAF et LOGEMENT, où une relation positive est observable. Un test du khi-deux entre ces variables est alors nécessaire pour vérifier leur dépendance.

En analysant les p-values du test de khi-deux d'indépendance (cf. annexe 30), des différents couples de variables, nous observons qu'elles sont toutes inférieures au seuil de 5 %. De ce fait, cela signifie qu'il existe une relation statistiquement significative entre ces deux variables.

Les résultats du premier tableau indiquent une relation positive entre le statut socio-économique et le fait de recevoir de l'argent, bien que le faible nombre d'individus dans les classes populaire et aisée empêche une affirmation claire.

Le deuxième tableau révèle que les étudiants n'ayant pas suivi de cours particuliers sont moins susceptibles de participer au tutorat, et ceci dans les deux sens.

À partir du troisième tableau, nous constatons que les étudiants payant un logement sont plus souvent bénéficiaires de la CAF, contrairement à ceux qui n'en paient pas ; ce qui est cohérent du point de vue administratif.

Enfin, selon le dernier tableau, les non-boursiers, ayant probablement un meilleur soutien financier, paraissent moins disposés à occuper un emploi pendant leurs études.

Figure 7: Matrice d'indépendance des variables qualitatives

Pearson's Chi2-Test of Independence																
	ARGENT	BOURSE	CAF	DEPENSES	EMPLOI	FORMATION	GENRE	LOGEMENT	PARTICULIER	REVISIONS	SANTE	STATUT	STRUCTURE	TRANSPORT	TUTORAT	UBER
UBER	0.425	0.595	0.005	1.000	0.095	0.692	0.009	0.145	0.285	0.711	0.587	0.098	1.000	0.425	1.000	0.000
TUTORAT	0.297	0.572	0.211	0.023	0.410	0.328	0.254	0.801	0.003	0.390	0.423	0.265	0.542	0.502	0.000	1.000
TRANSPORT	0.094	0.442	0.979	0.670	0.537	0.929	0.693	0.590	0.143	0.255	0.207	0.248	0.994	0.000	0.502	0.425
STRUCTURE	0.015	0.080	0.764	0.413	0.333	0.000	0.694	0.836	1.000	0.946	0.928	0.997	0.000	0.994	0.542	1.000
STATUT	0.008	0.000	0.112	0.087	0.282	0.525	0.635	0.353	0.408	0.509	0.477	0.000	0.997	0.248	0.265	0.098
SANTE	1.000	0.944	1.000	0.558	0.615	0.086	0.876	0.299	0.941	0.193	0.000	0.477	0.928	0.207	0.423	0.587
REVISIONS	0.745	0.716	0.850	0.014	0.830	0.328	0.000	0.993	0.304	0.000	0.193	0.509	0.946	0.255	0.390	0.711
PARTICULIERS	0.488	0.412	0.933	0.357	1.000	0.207	1.000	1.000	0.000	0.304	0.941	0.408	1.000	0.143	0.003	0.285
LOGEMENT	0.009	0.187	0.000	0.178	0.068	1.000	0.495	0.000	1.000	0.993	0.299	0.353	0.836	0.590	0.801	0.145
GENRE	0.238	0.145	0.473	0.562	0.581	0.629	0.000	0.495	1.000	0.000	0.876	0.635	0.694	0.693	0.254	0.009
FORMATION	0.381	0.033	0.343	0.140	0.596	0.000	0.629	1.000	0.207	0.328	0.086	0.525	0.000	0.929	0.328	0.692
EMPLOI	0.021	0.041	0.091	0.016	0.000	0.596	0.581	0.068	1.000	0.830	0.615	0.282	0.333	0.537	0.410	0.095
DEPENSES	0.469	0.994	0.199	0.000	0.016	0.140	0.562	0.178	0.357	0.014	0.558	0.087	0.413	0.670	0.023	1.000
CAF	0.003	0.145	0.000	0.199	0.091	0.343	0.473	0.000	0.933	0.850	1.000	0.112	0.764	0.979	0.211	0.005
BOURSE	0.014	0.000	0.145	0.994	0.041	0.033	0.145	0.187	0.412	0.716	0.944	0.000	0.080	0.442	0.572	0.595
ARGENT	0.000	0.014	0.003	0.469	0.021	0.381	0.238	0.009	0.488	0.745	1.000	0.008	0.015	0.094	0.297	0.425

(cf. annexe 28, 29 et 30)

L'analyse de la matrice d'indépendance des variables qualitatives est primordiale pour vérifier s'il existe des liens significatifs entre les variables qualitatives. Le coefficient varie entre 0 et 1, une valeur proche de 0 indiquera une forte association entre les deux variables, entraînant une dépendance plus élevée. Réciproquement, une valeur proche de 1 signifiera une faible association.

Nous pouvons constater que plusieurs intensités sont proches de 0, voire égale 0. Cela révèle une certaine dépendance entre les variables qualitatives, à laquelle il faudra prêter attention lors de l'estimation économétrique de notre modèle.

3. Analyse des variables Qualitatives - Quantitatives

Passons à l'analyse bivariée entre une variable qualitative et une variable quantitative. L'objectif est d'étudier l'influence des différentes catégories d'une variable qualitative sur une variable quantitative, spécifiquement notre variable dépendante, MOYENNE.

Tableau 10 : Modalités attendues selon la littérature et les modalités observées qui impactent la moyenne

Variables qualitatives	Modalité attendue qui impacte positivement la moyenne	Modalité observée qui impacte positivement la moyenne
FORMATION	Économie et gestion	Autres
REVISIONS	2h et plus	2h et plus
PARTICULIERS	Oui	Indéterminé
TUTORAT	Oui	Indéterminé
BOURSE	Oui	Oui
EMPLOI	Non	Oui
ARGENT	Oui	Oui
DÉPENSES	0 - 100 €	Indéterminé
TRANSPORT	0 - 15 €	0 - 15 € / 15 - 30 €
GENRE	Femme	Femme
SANTE	Non	Non
UBER	-	-
CAF	-	-
LOGEMENT	-	-
STRUCTURE	-	-
STATUT	-	-

(cf. annexe 31)

Le tableau 10 des modalités attendues selon la littérature et les modalités observées qui impactent la moyenne a été construit via la visualisation des boîtes à moustaches. Cette représentation révèle la distribution du jeu de données en affichant le premier quartile, la médiane et le troisième quartile qui respectivement délimitent les 25 %, 50 % et 75 % inférieurs des données. La répartition des données est également illustrée à travers les courbes de ces graphiques.

Ces graphiques confirment que la majorité de nos hypothèses, fondées sur la littérature, sont vérifiées. Par exemple, le fait d'être une femme impacte positivement la moyenne d'après la littérature et, selon les résultats observés, cela se confirme.

En revanche, certaines prédictions étaient fausses, comme celle relative à la variable DEPENSES. Cela est peut-être dû au fait que nous n'avions pas un échantillon assez grand ou à une sous-estimation des réponses de la part des participants.

V - Estimations économétriques

A. Estimations préliminaires

1. Sélection des variables explicatives significatives

Grâce à ses trois techniques, sélection ascendante, descendante et combinée, la fonction step va nous permettre de déterminer les variables les plus pertinentes pour notre modèle.

Notre modèle de base se compose de 15 variables explicatives.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \beta_8 X_{8i} + \beta_9 X_{9i} + \beta_{10} X_{10i} + \beta_{11} X_{11i} + \beta_{12} X_{12i} + \beta_{13} X_{13i} + \beta_{14} X_{14i} + \beta_{15} X_{15i} + \varepsilon_i$$

Avec : X_{1i} = GENRE, X_{2i} = AGE, X_{3i} = SANTE, X_{4i} = TUTORAT, X_{5i} = ASSIDUITE, X_{6i} = FORMATION, X_{7i} = REVISIONS, X_{8i} = PARTICULIERS, X_{9i} = STRESS, X_{10i} = TRANSPORT, X_{11i} = EMPLOI, X_{12i} = DEPENSES, X_{13i} = ARGENT, X_{14i} = BOURSE, X_{15i} = RESTAURANT

Notre fonction se base sur le critère d'information d'Akaike (AIC), proposé par Hirotugu Akaike en 1973, pour choisir notre modèle. Par sa formule $AIC = -2 * \log(L) + 2 * k$, avec L, la vraisemblance du modèle, et k, le nombre total de paramètres estimés, ce critère permet de comparer les modèles en évaluant lequel s'ajuste le mieux aux données tout en évitant le surajustement. Les trois méthodes de sélection minimisent ainsi l'AIC pour nous proposer le meilleur modèle.

Après avoir effectué la fonction step, nous obtenons $AIC = 136,19$ pour la méthode ascendante, $AIC = 136,04$ concernant la sélection descendante et $AIC = 136,19$ pour la méthode à double sens. Nous retenons donc le modèle ayant un critère d'Akaike de 136,04, car, par définition, plus ce critère est faible, meilleur est le modèle (cf. annexe 32, 33 et 34).

Les variables explicatives de cette méthode semblent également plus pertinentes au regard de notre analyse économique.

Tableau 11 : Tableau récapitulatif de notre modèle après stepwise

	Variables
Y	MOYENNE
X ₁	RESTAURANT
X ₂	TRANSPORT 15 - 30 €
X ₃	TRANSPORT 30 € et plus
X ₄	SANTE
X ₅	STRESS
X ₆	ARGENT
X ₇	GENRE

2. Choix du meilleur modèle

Après sélection de notre modèle grâce à la méthode Stepwise, nous pouvons effectuer la première régression linéaire de notre modèle.

Tableau 12 : Régression linéaire multiple de notre modèle (MCO)

	Estimate	Std. Error	t value	Pr(> t)
Constante	14.67	0.47	30.84	< 2e-16
RESTAURANT	-0.32	0.10	-3.014	0.0031 **
TRANSPORT 15 - 30€	-0.07	0.37	-0.189	0.85
TRANSPORT 30€ et plus	0.588	0.398	1.477	0.142

SANTE	-0.849	0.371	-2.285	0.0239 *
STRESS	-0.163	0.0599	-2.71	0.00757 **
ARGENT	0.45	0.29	1.537	0.127
GENRE (Homme)	-0.43	0.282	-1.526	0.129
Multiple R-squared: 0.203		Adjusted R-squared: 0.1591		p-value: 0.0001237

(cf. annexe 35)

Après vérification des hypothèses, ce tableau nous permettra d'interpréter le test de Fisher. Celui-ci nous permettra de valider ou non l'hypothèse H_0 selon laquelle aucune des variables indépendantes n'a d'effet significatif sur la variable dépendante. Dans ce cas présent, la p-value inférieure à 0,05 nous permettrait de rejeter H_0 ; ce qui signifierait qu'il y a au moins une variable qui explique la moyenne.

Ensuite, nous pourrons effectuer une vérification de la significativité individuelle de nos variables explicatives avec un test de Student. En effet, il apparaît que la variable RESTAURANT serait significative au seuil de 1 % et la variable SANTE le serait au seuil de 5 %. Enfin, la variable STRESS serait significative à 1 %. Nous effectuerons l'interprétation des coefficients après validation des tests statistiques vérifiant les hypothèses de la méthode des MCO.

En ce qui concerne le R^2 , nous obtenons seulement 0.1591. Cela impliquerait que notre modèle explique seulement 15.91 % de la moyenne générale des étudiants, ce qui ne serait pas satisfaisant.

B. Tests Statistiques

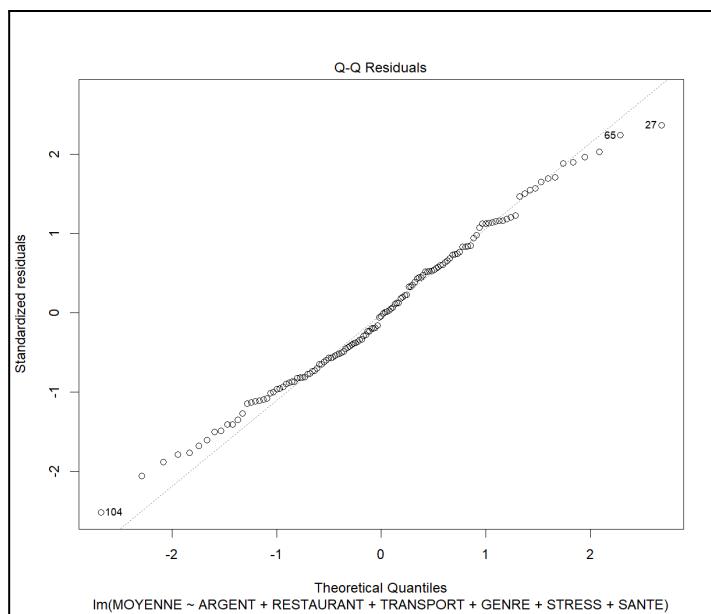
Nous allons maintenant réaliser plusieurs tests pour vérifier les hypothèses nécessaires à l'application de la méthode des moindres carrés ordinaires et trouver un modèle linéaire valide. Pour cela, nous allons commencer par un test de normalité des résidus (1), ensuite

nous vérifierons la forme fonctionnelle (2), puis l'absence de multicolinéarité (3), l'homogénéité des erreurs (4) et pour finir la distance de cook (5).

1. Normalité des résidus

La normalité des résidus est primordiale pour garantir la fiabilité des tests statistiques suivants. Dans cette optique, nous observerons la distribution des résidus à l'aide d'un Q-Q plot des résidus.

Figure 8: Q-Q plot des résidus



Dans cette figure, nous remarquons que la majorité des points semblent s'aligner sur la droite, ce qui suggère que les résidus standardisés suivent une distribution normale. Cela reste subjectif et un test plus approfondi est nécessaire, nous allons par conséquent procéder au test de Shapiro-Wilk.

Tableau 13 : Test de Shapiro-Wilk

Test de Shapiro-Wilk	
p-value	0.5421

(cf. annexe 36)

L'hypothèse nulle du test de Shapiro-Wilk est la suivante, les résidus suivent une loi normale. Ici, la p-value est de 0.5421. Elle est supérieure à 0.05, ce qui signifie que H_0 n'est pas rejetée au seuil de 5 % et donc que les résidus suivent bien une distribution normale.

En revanche, le test de Shapiro-Wilk est par définition efficace pour les petits échantillons. 135 observations étant supérieures à 100, il peut être plus approprié d'effectuer le test de Kolmogorov-Smirnov.

Tableau 14 : Test de Kolmogorov-Smirnov

Test de Kolmogorov-Smirnov	
p-value	0.7022

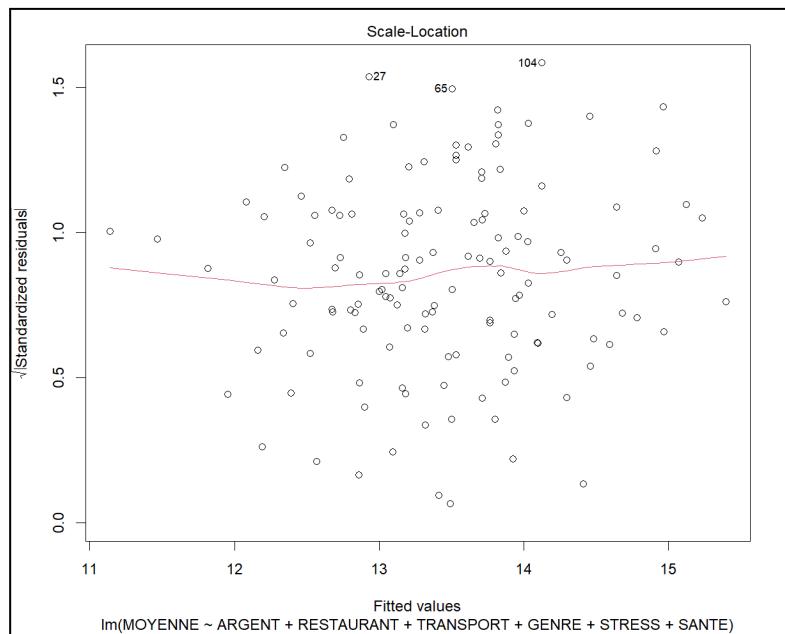
(cf. annexe 37)

Ici, la p-value de 0.7022 est bien supérieure à 0.05. Par conséquent, H_0 n'est toujours pas rejetée au seuil de 5 %. Les résidus suivent une distribution normale, ce qui confirme notre première analyse.

2. Forme fonctionnelle

Nous allons maintenant vérifier si la forme fonctionnelle choisie est bien linéaire, en commençant par un graphique sur la représentation de la distribution des résidus en fonction des valeurs prédictes.

Figure 9 : Graphique de la distribution des résidus en fonction des valeurs prédictes



Le modèle semble bien spécifié puisqu'aucune tendance ne se dégage. Cela reste subjectif et un test de Ramsey est nécessaire.

Tableau 15 : Test de Ramsey

Test de Ramsey	
p-value	0.3399

(cf. annexe 38)

La p-value est de 0.3399, ce qui signifie que H_0 n'est pas rejetée au seuil de 5 %. Le modèle est donc bien spécifié : la forme fonctionnelle linéaire est appropriée.

3. Absence de multicolinéarité

Il s'agit désormais d'éviter toute multicolinéarité entre les variables explicatives. En effet, celle-ci affecte l'estimation et l'interprétation des coefficients. Il est difficile d'évaluer l'effet individuel de chaque variable sur la variable dépendante car la multicolinéarité rend les

coefficients instables et sensibles aux variations des données. Ainsi, déterminer quelle variable explicative a réellement l'impact le plus significatif sur la variable dépendante est compliqué tout comme l'interprétation de leur coefficient.

Par conséquent, nous utiliserons le facteur d'inflation de la variance (VIF) pour identifier cette multicolinéarité et la quantifier. Cet indicateur évalue dans quelle mesure la variance estimée du coefficient d'une variable est amplifiée par la colinéarité avec les autres variables explicatives. Nous admettrons qu'une variable présentant un VIF inférieur à 10 est acceptable.

Tableau 16 : Test de multicolinéarité

	GVIF	Df	GVIF ^{(1/(2*df))}
RESTAURANT	1.066001	1	1.032473
TRANSPORT	1.124012	2	1.029657
GENRE	1.030033	1	1.014905
STRESS	1.101970	1	1.049748
SANTE	1.053461	1	1.026383
ARGENT	1.108109	1	1.052668

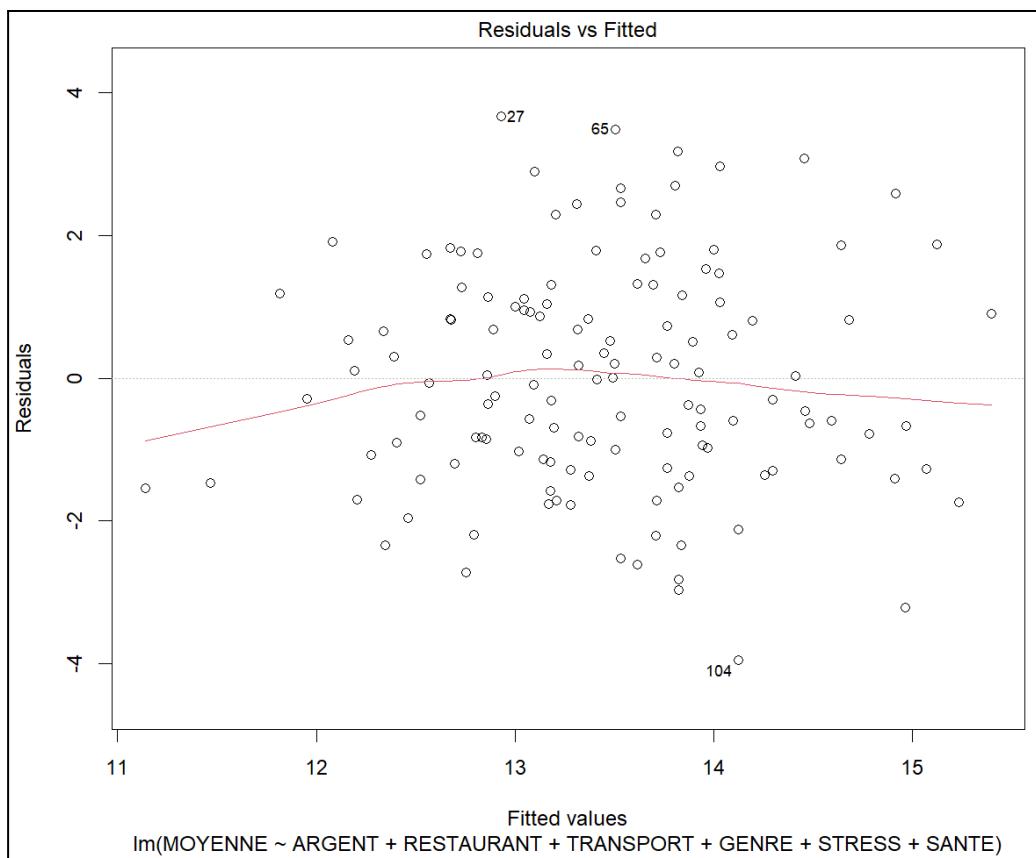
Les VIF ajustés montrent des valeurs proches de 1 pour chaque variable. Il n'existe donc pas de problème de multicolinéarité entre les variables explicatives du modèle. Par conséquent, aucune des variables n'a besoin d'être supprimée (cf. annexe 39).

4. Homoscédasticité des erreurs

À présent, nous devons vérifier l'hypothèse d'homoscédasticité des erreurs, selon laquelle la variance des résidus reste constante pour tous les individus.

Tout d'abord, examinons le tableau ci-dessous, qui montre la distribution des résidus en fonction des valeurs ajustées.

Figure 10 : Répartition des résidus en fonction des valeurs ajustées



Nous constatons une répartition aléatoire des résidus autour de zéro, avec une légère courbure de la ligne rouge. Cette dernière observation peut susciter des doutes concernant la validité de l'hypothèse d'homoscédasticité.

Par conséquent, un test de Breusch-Pagan doit être effectué pour s'assurer que la variance des erreurs est constante à travers toutes les observations.

Tableau 17 : Test de Breusch-Pagan

Test de Breusch-Pagan	
p-value	0.8171

(cf. annexe 40)

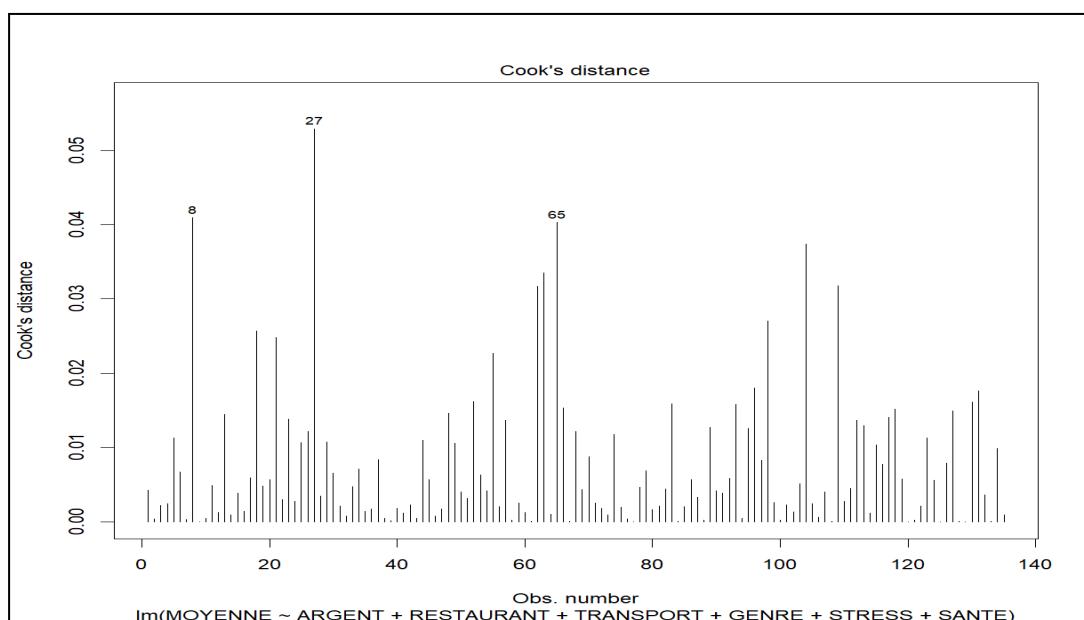
La p-value de 0.8171, qui est largement supérieure à 0,05, indique que l'hypothèse nulle ne peut être rejetée au seuil de 5 %, ce qui suggère que les résidus sont homoscédastiques.

5. Distance de Cook

Pour terminer, nous allons exploiter la distance de Cook pour identifier les observations susceptibles d'exercer une influence significative sur l'ajustement du modèle de régression. En effet, une distance de Cook élevée (supérieure à 1) indique que la suppression de l'individu concerné pourrait affecter considérablement les coefficients du modèle.

Examinons la distance de Cook à l'aide de son graphique.

Figure 11 : Distance de Cook



D'après le graphique ci-dessus, étant donné que la distance de Cook est nettement inférieure à 1, aucune observation n'exerce une grande influence sur les estimations des paramètres du modèle ou sur les prédictions. Il n'est donc pas nécessaire de supprimer des observations.

En conclusion, toutes les hypothèses des Moindres Carrés Ordinaires (MCO) semblent être respectées, à savoir : la normalité des résidus, la forme fonctionnelle linéaire appropriée, l'absence de multicolinéarité, l'homoscédasticité des erreurs et l'absence d'autocorrélation des résidus. Cette dernière hypothèse étant confirmée par la nature des données en coupe transversale qui ne présentent pas d'autocorrélation temporelle par définition. De plus, le nombre d'observations est supérieur à celui des variables, ce qui assure que la matrice des variables explicatives est bien définie. Ces éléments montrent ainsi que le modèle est robuste et fiable pour cette analyse.

C. Vérification de l'endogénéité

Nous supposons désormais que certaines variables explicatives sont endogènes, c'est-à-dire qu'elles sont corrélées avec le terme d'erreur du modèle de régression.

L'endogénéité peut survenir dans trois situations principales qui remettent en cause l'hypothèse d'exogénéité des variables explicatives, nécessaire pour une estimation correcte par MCO. Premièrement, l'omission d'une variable explicative expliquant de manière significative y et corrélée avec d'autres variables explicatives. Deuxièmement, les erreurs de mesure sur les variables explicatives peuvent biaiser les estimations. Enfin, le biais de simultanéité qui se produit lorsque la variable dépendante et l'une ou plusieurs des variables explicatives sont déterminées simultanément ou lorsque la relation de causalité est bidirectionnelle, créant une corrélation entre les variables explicatives et le terme d'erreur du modèle.

L'endogénéité peut donc provoquer des biais dans les estimations des coefficients et compliquer l'interprétation des résultats. Pour résoudre ce problème, il est nécessaire d'utiliser des variables instrumentales, en appliquant la méthode des variables instrumentales ou celle des doubles moindres carrés.

1. Choix des instruments de notre modèle

Le choix des variables instrumentales est primordial pour corriger les problèmes d'endogénéité. Elles permettent d'estimer de manière fiable les coefficients des variables endogènes, assurant ainsi des estimations consistantes et sans biais du modèle.

Dans notre modèle, nous suspectons trois variables d'être endogènes, la variable PARTICULIERS, la variable RÉVISIONS et la variable STRESS. Par conséquent, nous avons sélectionné, d'après notre analyse économique, des instruments qui pourraient être fortement corrélés avec la variable endogène et non corrélés avec l'erreur du modèle ; garantissant la consistance et le fait que les estimations soient non biaisées. Les variables instrumentales choisies sont donc : les aide de la CAF (CAF), les commandes UBER EAT (UBER), le loyer 2023-2024 (LOGEMENT), le temps de sommeil (SOMMEIL), le temps de trajet pour aller en étude (TRAJET), le statut social (STATUT) et la structure d'étude (STRUCTURE).

Tableau 18 : Récapitulatif des variables instrumentales pour chaque variable suspecté d'endogénéité dans notre modèle initial.

Variables suspecté d'endogénéité	Variables instrumentale
STRESS	CAF / UBER / LOGEMENT / SOMMEIL /
RÉVISIONS	SOMMEIL / STRUCTURE / TRAJET
PARTICULIERS	STATUT

Toutefois, nous nous concentrerons uniquement sur le modèle obtenu par la méthode descendante, pour lequel toutes les hypothèses liées à la méthode des MCO ont été vérifiées. Dans ce modèle, la variable STRESS semble être la seule susceptible de poser un problème d'endogénéité. Nous devons donc vérifier la pertinence et l'exogénéité de nos instruments via des tests statistiques.

2. Tests Statistiques

Pour vérifier les hypothèses concernant nos instruments et leur influence sur la variable endogène nous avons procédé à une régression des doubles moindres carrés, représentée par le tableau ci-dessous.

Tableau 19 : Résultats des doubles moindres carrés et des tests d'endogénéité

	Estimate	Std. Error	t value	Pr(> t)
Constante	15.667	0.833	18.8008	< 2e-16 ***
RESTAURANT	-0.367	0.11783	-3.115	0.00227 **
TRANSPORT 15 - 30€	-0.03144	0.39823	-0.079	0.93721
TRANSPORT 30€ et plus	0.77345	0.44351	1.744	0.08359 .
SANTE	-0.58740	0.43384	-1.354	0.17816
STRESS	-0.41831	0.18198	-2.299	0.02316 *
ARGENT	0.251	0.34	0.737	0.46234
GENRE (Homme)	-0.53462	0.30963	-1.727	0.08667 .
Multiple R-Squared : 0.0888	Adjusted R-Squared: 0.03866		p-value: 0.0007381	

	df1	df2	Statistic	p-value
Weak instruments	4	124	4.389	0.00236 **
Wu-Hausman	1	126	2.607	0.10886
Sargan	3	NA	0.865	0.83384

(cf. annexe 41)

Premièrement, nous nous demandons si les instruments expliquent suffisamment notre variable endogène, STRESS. Le test Weak instruments, dont l'hypothèse nulle est que les instruments sont faiblement corrélés avec la variable suspectée d'endogénéité (coefficients proches de zéro ou non significatifs), répond à notre interrogation.

Ici, pour la variable STRESS, la p-value est de 0.0026, ce qui signifie que H_0 est rejetée au seuil de 5 %. En conséquence, nos instruments sont pertinents : ils ont un impact significatif sur la variable endogène.

Nous vérifions ensuite que les instruments que nous avons choisis ne sont pas corrélés avec les termes d'erreur de l'équation estimée à l'aide du test de Sargan. La p-value associée à notre modèle de 0.833854 est supérieure à 0,05 ce qui implique que l'hypothèse nulle d'exogénéité des instruments n'est pas rejetée au seuil de risque de 5 %. Cela signifie que nos instruments sont valides : ils sont orthogonaux, c'est-à-dire non corrélés, aux résidus du modèle.

Pour finir, nous allons déterminer quelle méthode est la plus appropriée pour estimer notre modèle, celle des Moindres Carrés Ordinaires (MCO) ou celle des Doubles Moindres Carrés (DMC). Pour ce faire, nous considérons le résultat du test d'exogénéité de Wu-Hausman, dont l'hypothèse nulle est que la variable testée est supposée exogène. Ici, la p-value de 0.10886 est nettement supérieure à 0,05. L'hypothèse nulle n'est donc pas rejetée au seuil de risque de 5 %, ce qui signifie que le terme résiduelle de l'équation auxiliaire n'a pas d'impact sur notre variable dépendante, MOYENNE. Par conséquent, notre variable explicative suspectée d'être endogène, STRESS, peut être traitée comme exogène ; ce qui signifie que l'estimation par la méthode des MCO doit être privilégiée.

D. Interprétation des résultats

Maintenant, après la vérification de tous les tests ci-dessus, nous pouvons interpréter les résultats de la régression linéaire multiple estimée par la méthode des MCO du tableau 12 (cf. annexe 30).

Ainsi, le test de Fisher-Snedecor nous permet de savoir si au moins une des variables est significative. Pour notre modèle, la p-value associée est égale à 0.0001237, ce qui veut dire que H_0 est rejetée au seuil de 5 %. Par conséquent, notre modèle est globalement explicatif : au moins une variable est statistiquement significative.

Ensuite, nous pouvons examiner le test de Student pour chaque variable explicative afin de savoir si elles sont significatives ou non. Dans notre modèle, nous avons les variables RESTAURANT et STRESS qui sont significatives au seuil de 1 %, ainsi que la variable SANTE qui est significative au seuil de 5 %.

Nous pouvons, avec ces variables statistiquement significatives, interpréter les coefficients :

Nombre de restaurants (- 0.32) \Rightarrow L'augmentation d'une unité du nombre de restaurants fait baisser la moyenne de 0.32 point, toutes choses égales par ailleurs.

Stress financier (- 0.163) \Rightarrow L'augmentation d'une unité du stress financier fait baisser la moyenne de 0.163 point, toutes choses égales par ailleurs.

Santé (- 0.84) \Rightarrow Avoir des problèmes de santé fait baisser la moyenne de 0.84 point, toutes choses égales par ailleurs.

Concernant le R^2 , nous obtenons seulement 0.1591. Cela signifie que notre modèle explique seulement 15.91 % de la moyenne générale des étudiants, révélant une qualité d'ajustement du modèle qui n'est pas satisfaisante.

VI - Prévisions

Concernant la prévision, nous avons construit un graphique et un histogramme (figure 12 et 13) sur les prédictions du modèle. La droite rouge représente le cas où la moyenne générale observée chez un individu est égale à celle calculée par le modèle. La distance entre l'observation et la droite rouge représente ainsi les résidus. En effet, tous les individus étant au-dessus de cette droite ont une moyenne observée supérieure à celle prédite et ceux en dessous ont une moyenne inférieure à celle prédite. Nous pouvons remarquer qu'un grand nombre de points sont éloignés de la droite rouge, ce qui révèle une limite de notre modèle.

Figure 12: Représentation graphique des prédictions du modèle

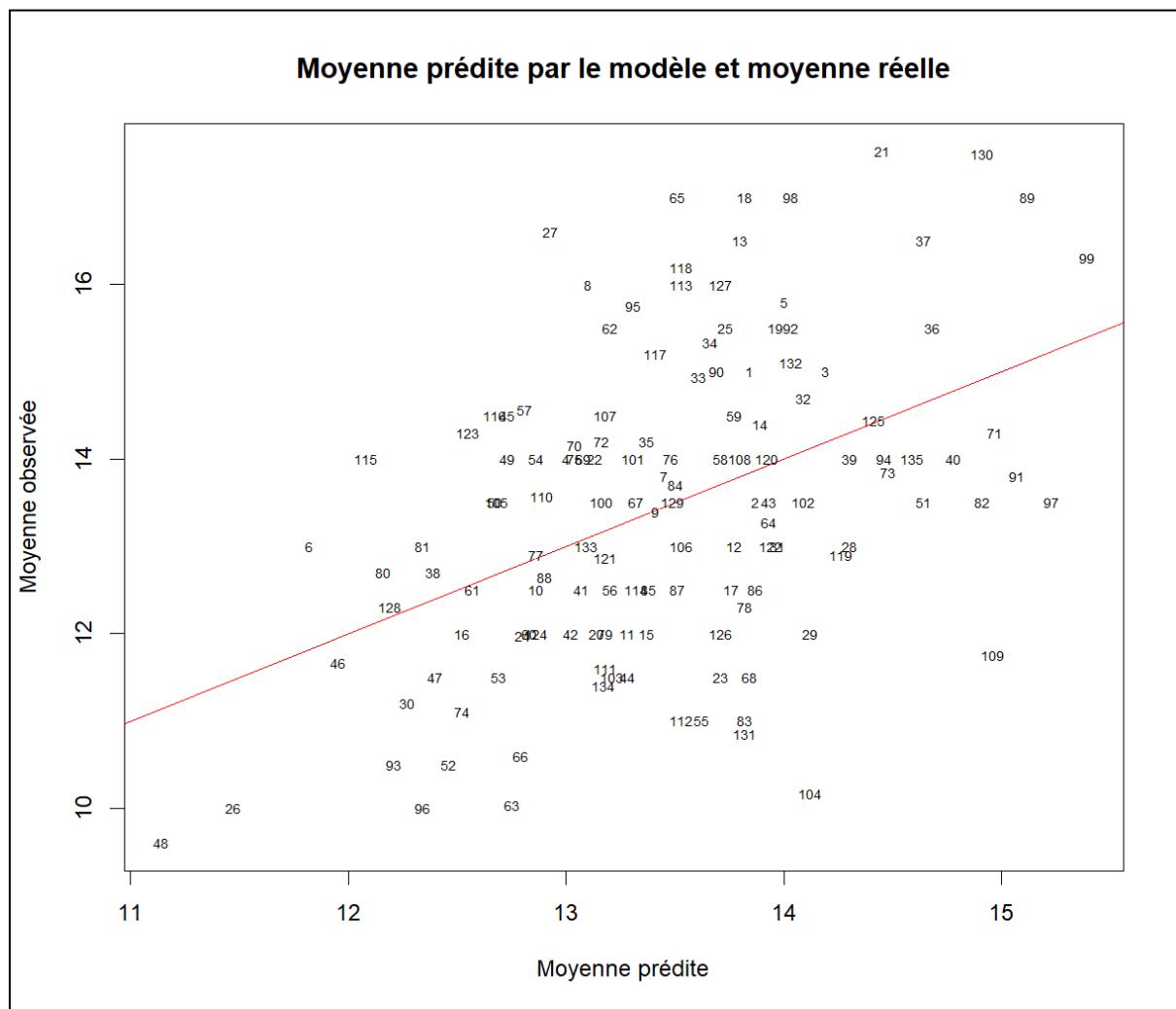
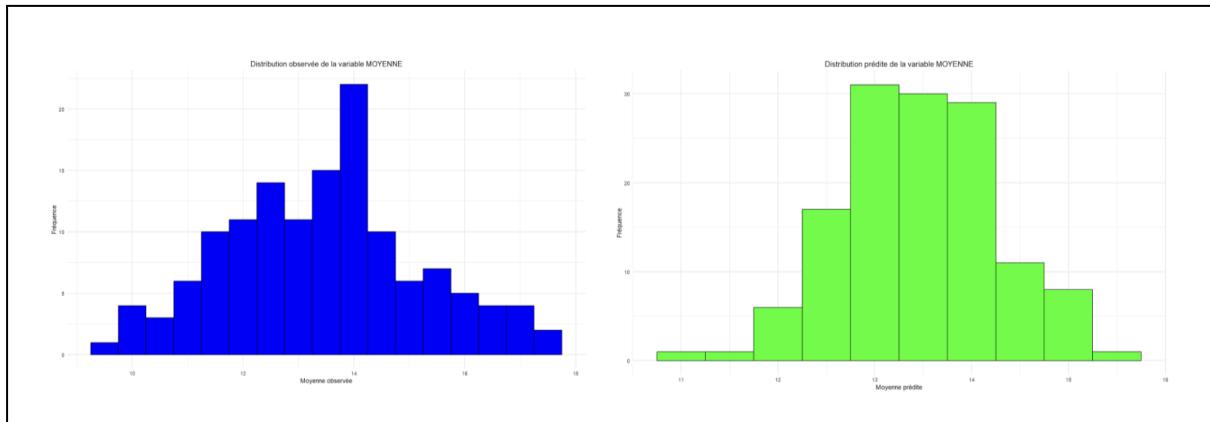


Figure 13: Comparaison entre la distribution de la moyenne prédictive et observée



VII - Discussion

Pour la discussion nous allons commencer par discuter de nos résultats en les comparants à la littérature (A) puis nous parlerons des limites rencontrées (B)

A. Comparaison de nos résultats avec la littérature

Nous allons à présent faire une comparaison des résultats observés avec ce qui était attendu par la littérature. En effet, certaines relations attendues étaient vraies, comme le fait d'être assidu en cours. Effectivement, selon la littérature avoir une assiduité élevée lors des cours magistraux et des travaux dirigés est corrélé à de meilleures performances scolaires et que, en moyenne ils obtiennent 10 à 15 % de meilleurs résultats aux examens que ceux qui ne sont pas assidus et cela s'est plutôt vérifié dans les résultats observés. Nous pouvons également parler du fait du nombre de fois d'aller au restaurant, qui d'après la littérature cela a un impact négatif sur la moyenne et d'après les résultats obtenus, il est vrai que le fait d'aller un grand nombre de fois au restaurant impacte négativement la moyenne. De plus, cette variable est significative au seuil de 5 % dans notre modèle.

Cependant, il y a également des variables où la modalité attendue de la littérature n'étaient pas les mêmes avec ceux observés.

En effet, pour la variable des cours particuliers, d'après la littérature la modalité attendue qui devrait impacter positivement la moyenne est "vrai", mais d'après les résultats aucune réponse ne se distingue . Cela peut être dû à notre question qui peut être mal posée. Par ailleurs, cela peut être dû au peu de réponses qu'on a reçues pour cette variable, en effet nous avons reçu seulement 8 réponses positives pour cette variable. Ce nombre de réponses positives est trop juste pour pouvoir avoir des résultats cohérents avec la littérature.

Nous pouvons également parler des dépenses mensuelles dans les loisirs (bar, cinéma, magasin, sport...), qui d'après la littérature la modalité qui devrait impacter positivement la

moyenne est “0 - 100 €”, mais d’après les résultats observés aucune ne se distingue. Nous avons peut-être mis de mauvaises tranches.

Pour la variable transport, la modalité attendue était “0 - 15 €” et d’après les résultats obtenus, la modalité qui impacte positivement la moyenne est “0 - 15 €” et “15 - 30 €” ce qui confirme notre prédiction mais en ajoutant 15 - 30 €”. Cela est sans doute dû au fait que nos tranches étaient très rapprochées.

De plus, nous pouvons nous intéresser aux coefficients des résultats. En effet, même si la variable GENRE n’est pas significative au seuil de 10 % (12,9 %), nous remarquons que le fait d’être un homme impact négativement la moyenne, et d’après la littérature la modalité qui impacte positivement est la modalité Femme, ce qui confirme ce qui est observé. Nous pouvons également faire la même analyse pour le fait de recevoir de l’argent, en effet on s’attendait à avoir un coefficient positif et dans notre MCO le coefficient est positif.

Pour la variable santé, après étude de la littérature on s’attendait à avoir un coefficient négatif, et en effet le coefficient était négatif. Ce qui confirme la littérature entre le fait d’être en bonne santé et d’avoir plus de chance d’avoir une meilleure moyenne.

Enfin pour la variable du stress financier, d’après la littérature on s’attendait également à un coefficient négatif, et dans notre régression linéaire multiple nous avons obtenu un coefficient négatif, ce qui confirme la littérature.

B. Limites

Après avoir tenté d’expliquer les différences entre nos résultats et la littérature, nous pouvons à présent évoquer les limites de notre étude.

Notre enquête a été diffusée uniquement sur 181 individus, et après le nettoyage des données, nous avons retenu 150 individus. L’échantillon est composé de plus de 76 %

d'étudiants issus des filières Économie et Gestion, dont les études, notamment en finance et en économie, peuvent probablement influencer leur gestion budgétaire. Ce questionnaire a été diffusé exclusivement via l'adresse e-mail générale de la faculté de l'IAE Nantes, auprès de certaines promotions en sciences, ainsi que sur nos réseaux sociaux, ce qui peut introduire des biais.

Des étudiants ont pu être amenés à ignorer notre sondage étant donné que nous avons envoyé un premier questionnaire deux semaines auparavant. De plus, le thème de la gestion budgétaire a pu freiner quelques personnes en les dissuadant de participer. En France, les discussions autour de l'argent peuvent être délicates, bien que la génération Z semble aborder ce sujet avec plus de liberté.

Certaines réponses peuvent également ne pas refléter la réalité. Premièrement, les réponses peuvent être subjectives, notamment lorsqu'il est préférable de mesurer certaines variables sur une échelle. Deuxièmement, nous pensons que notre sujet peut être perçu comme sensible, tout du moins en France. Des étudiants peuvent ne pas vouloir révéler qu'ils reçoivent de l'argent de leurs parents, bénéficient de revenus d'investissements ou que leurs parents appartiennent à une classe sociale modeste. Concernant notre variable dépendante, certains peuvent avoir tendance à surestimer leur moyenne générale, par crainte d'être jugés, même si le questionnaire est anonyme. Dans notre échantillon, la moyenne générale de tous les individus s'élève à 13,45/20. Ces réponses ont donc pu altérer l'exactitude des résultats.

VIII - Conclusion

La visée de notre projet était de réussir à trouver et analyser de manière approfondie les facteurs économiques qui pouvaient influencer la moyenne générale des étudiants. Pour arriver à notre but, nous avons transmis un questionnaire qui nous a permis de récolter 135 réponses en enlevant les données atypiques et en faisant un nettoyage de données.

Notre modèle économétrique de base se composait de 16 variables explicatives et 5 instruments. Après la méthode de Stepwise nous sommes passés à 6 variables explicatives, qui sont GENRE, SANTE, TRANSPORT, STRESS, RESTAURANT et ARGENT. En revanche le R^2 de ce modèle est de seulement 15.9%. Ce modèle a été soumis à une série de tests statistiques et ils ont tous été validés. Les tests sont les suivants: la normalité des résidus, la forme fonctionnelle, la multicolinéarité, l'homoscédasticité des erreurs et la distance de Cook. Dans notre modèle, après le stepwise, nous avions qu'une seule variable suspecté d'endogénéité, la variable du stress financier qui était traité par 4 instruments: les aides la CAF en 2023-2024, le nombre de commande moyen Uber Eat en 2023-2024, le fait de payer un loyer ou non en 2023-2024 et le temps de sommeil. Toutefois, les tests de Weak-instruments, Wu-Hausman et Sargan nous ont confirmé la pertinence du modèle des moindres carrés ordinaires. Par conséquent, nous avons interprété ce modèle.

De notre modèle, seulement trois variables sont ressorties significatives, la variable du stress financier, la variable de la santé et la variable du nombre de restaurants. Nous pouvons également noter que les variables sur le genre et de recevoir de l'argent ont des p-values respectives de 0.129 et 0.127 ont pu avoir un impact sur la moyenne.

Après analyse de nos résultats, nous pouvons donner certaines recommandations pour améliorer la moyenne générale. En effet, le fait de diminuer le nombre de restaurant par mois fera améliorer la moyenne générale, dû au fait que si l'individu sort moins pour aller au restaurant il aura plus de temps pour réviser, aura moins de stress financier dû au prix élevé des restaurants, aura une meilleure alimentation et aura sans doute un meilleur sommeil. Nous vous recommandons également d'aller régulièrement faire des rendez-vous médicaux pour prévenir certaines maladies. Des événements organisés par l'université en collaboration

avec des médecins universitaires, gratuits pour les participants, pourraient constituer une solution pour aider les étudiants confrontés à des difficultés financières. Pour diminuer le stress financier, nous préconisons d'aller évidemment moins au restaurant ou au bar, de travailler à côté de ses études sans faire plus de 15 heures hebdomadaire(cela amène une certaine organisation et responsabilise les étudiants sans, normalement les pénaliser dans les résultats académiques), de faire du sport, et hiérarchiser ses dépenses.

En conclusion, il serait pertinent d'étendre cette étude à l'ensemble des universités, ainsi qu'aux écoles de commerce et d'ingénierie, afin de pouvoir comparer les résultats en fonction des structures et des filières d'études.

IX - Bibliographie

Vie Publique. (27/03/2023). Jobs étudiants : 146 000 étudiants occupent un emploi hors du cadre de leurs études.

<https://www.vie-publique.fr/en-bref/293511-jobs-etudiants-146-000-etudiants-ont-un-emploi-hors-de-leurs-etudes>

Ministère de l'enseignement supérieur et de la recherche. (09/2024). Les boursiers sur critères sociaux en 2023-2024.

<https://www.enseignementsup-recherche.gouv.fr/fr/les-boursiers-sur-criteres-sociaux-en-2023-2024-97464#:~:text=La%20proportion%20de%20boursiers%20au,de%20formations%20et%20entre%20acad%C3%A9mies.>

Jean-Paul Caille. (12/2010) Les cours particuliers en première année de collège : un entrant en sixième sur dix bénéficie de soutien scolaire payant, Éducation & formations n° 79.

https://www.education.gouv.fr/sites/default/files/imported_files/document/DEPP-REF-2010-79-Les_cours_particuliers_en_premiere_annee_de_college_167238.pdf

David Corceiro et Richard Lioger. (12/2021). PAR LA COMMISSION DES AFFAIRES ÉCONOMIQUES sur le logement et la précarité des étudiants, des apprentis et des jeunes actifs.

<https://www.lesepl.fr/wp-content/uploads/2021/12/rapport-logement-etudiant.pdf>

Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation. (2023). Conditions de logement des étudiants.

https://publication.enseignementsup-recherche.gouv.fr/eestr/FR/T239/conditions_de_logement_des_etudiants/

Laura Castell, Insee, Mickaël Portela et Raphaëlle Rivalin. (06/2016). Les principales ressources des 18-24 ans, INSEE

<https://www.insee.fr/fr/statistiques/2019048#:~:text=Lorsqu'ils%20sont%20en%20cours,mon tant%20moyen%20de%20220%20euros.>

Ali Mâalej et Sami Boudabbous. (2020). Caractéristiques individuelles et taux de réussite à l'université : fondement théorique et vérification empirique sur une cohorte d'étudiants, ResearchGate.

https://www.researchgate.net/profile/Ali-Maalej/publication/339570209_Caracteristiques_individuelles_et_taux_de_reussite_a_l'universite_fondement_theorique_et_verification_empirique_sur_une_cohorte_d'etudiants/links/633ffa1a76e39959d6a7b975/Caracteristiques-individuelles-et-taux-de-reussite-a-l'universite-fondement-theorique-et-verification-empirique-sur-une-cohorte-detudiants.pdf

April R. Jumawan-Powao, Krizia T. Tubat, Rio Cyrel L Cellero, Ma. Khym T. Capilitan. (2024). Family Income in Relation to Budgeting of Accounting Students in A University, ResearchGate.

https://www.researchgate.net/publication/383403484_Family_Income_in_Relation_to_Budgeting_of_Accounting_Students_in_A_University

Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation. (2014). La réussite des étudiants selon les difficultés financières et la perception d'une allocation d'études.

https://www.enseignementsup-recherche.gouv.fr/sites/default/files/content_migration/document/NI_ESR_14_05_340913.pdf

<https://www.enseignementsup-recherche.gouv.fr/fr/la-reussite-des-etudiants-selon-les-difficultes-financieres-et-la-perception-d'une-allocation-d-47382>

Unef. (2023). Enquête sur l'explosion du coût de la vie étudiante.

<https://unef.fr/wp-content/uploads/2024/08/Enquete-sur-le-cout-de-la-vie-etudiante-2023.pdf>

Charente Libre. .(04/2024). Infographie, plus d'un quart des étudiants français travaillent à côté de leurs études.

<https://www.charentelibre.fr/education/etudes-superieures/infographie-plus-d-un-quart-des-etudiants-francais-travaillent-a-cote-de-leurs-etudes-19268799.php?csnt=19266c21502>

Solène Hilary, Laurent Lefèvre, Laurence Pen. (2024). Les « jobs » étudiants début 2020 : quatre fois sur dix, des emplois de serveurs, caissiers ou vendeurs, INSEE.

<https://www.insee.fr/fr/statistiques/7927621>

Angela Tirroloni. (2015). Meilleure scolarité mais insertion professionnelle plus difficile pour les filles, INSEE.

<https://www.insee.fr/fr/statistiques/1285441#:~:text=absentes%20en%20Corse.-,Les%20jeunes%20femmes%20plus%20dipl%C3%B4m%C3%A9es%20que%20les%20hommes,elles%20r%C3%A9ussissent%20 davantage%20au%20baccalaur%C3%A9at.>

Transkriptor. (2023). La présence aux cours augmente-t-elle les performances académiques ?
<https://transkriptor.com/fr/lassiduite-aux-cours-magistraux-augmente-t-elle-les-performances-academiques/>

Heather. (2022). L'importance de l'assiduité des élèves dans l'éducation, SeatsSoftware
<https://www.seatssoftware.com/fr/the-importance-of-student-attendance-in-education/>

LEMAIRE, S. 2000. Étude sur les facteurs de la réussite dans les deux premières années d'enseignement supérieur (DEUG, D.U.T, B.T.S.), Note d'information, D.P.D, 00.25, Août.

Perrot, J. (1988). L'influence de l'utilisation du temps sur la réussite des étudiants en France. L'Actualité économique, 64(1), 44–67.

<https://www.erudit.org/fr/revues/ae/1988-v64-n1-ae2717/601436ar.pdf>

Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation. (2013). La vie étudiante: La santé des étudiants

https://publication.enseignementsup-recherche.gouv.fr/eessr/8/EESR8_ES_14-la_vie_etudiant_e_la_sante_des_etudiants.php

Lucia Romo, Stéphanie Nann, Elisabetta Scanferla, José Esteban, Hélène Riazuelo et Laurence Kern. (2019). La santé des étudiants à l'université comme déterminant de la réussite académique. Volume 40, numéro 2, 2019, p. 187–202

<https://www.erudit.org/fr/revues/rqpsy/2019-v40-n2-rqpsy04976/1065909ar/>

Denis Fougère. (2010). Les méthodes économétriques d'évaluation. Pages 105 à 128, Cairn.info

[Les méthodes économétriques d'évaluation | Cairn.infoshs.cairn.info](#)

Educacentre (2024). L'impact de santé mentale sur la réussite scolaire

<https://educacentre.com/limpact-de-la-sante-mentale-sur-la-reussite-scolaire-2/>

Jean-Jacques Quintin. L'impact du tutorat sur les performances des étudiants. Environnements Informatiques pour l'Apprentissage Humain, Jun 2007, Lausanne, Suisse. pp.221-231. □edutice-00349026□

https://edutice.hal.science/edutice-00349026v1/file/Quintin_2007_L_impact_du_tutorat_sur_les_performances_EIAH.pdf

National Union of Students (NUS). (2013). Mental Distress Overview.

<https://www.nus.org.uk/global/campaigns/20130517%20mental%20distress%20survey%20%20overview.pdf>

Chudzik, J. (2023). Les facteurs modérateurs de la relation entre le stress financier subjectif et la détresse psychologique chez les personnes étudiantes de l'Université de Sherbrooke (Mémoire de maîtrise, Université de Sherbrooke). https

https://savoirs.usherbrooke.ca/bitstream/handle/11143/20440/chudzik_jessica_MSc_2023.pdf?sequence=4&isAllowed=y

Lassarre, D., Giron, C., & Paty, B. (2003). Stress des étudiants et réussite universitaire : les conditions économiques, pédagogiques et psychologiques du succès. Revue française des affaires sociales, 32(4), 669-691.

<https://journals.openedition.org/osp/2642>

Mulder, A. M. et Cashin, A. (2015). Health and wellbeing in students with very high psychological distress from a regional Australian university. *Advances in Mental Health*, 13(1), 72-83.

<https://doi.org/10.1080/18374905.2015.1035618>

Pedrelli, L., & Peretti, N. (2013). L'impact du stress sur la performance et la mémoire des élèves. *Revue Française de Pédagogie*, 13, 13-25.

[L'impact du stress sur la performance et la mémoire ...Patrinumhttps://patrinum.ch/.../mp8924_ms2_p46754_p47...](https://patrinum.ch/.../mp8924_ms2_p46754_p47...)

Beyrouthy, L. (2024). Average food delivery spend per order in the U.S. Statista.

<https://www.statista.com/statistics/1358873/average-food-delivery-spend-per-order-us/>

Bouhnik, D. (2015). L'impact du travail salarié des étudiants sur la réussite et la poursuite des études universitaires. Academia.edu.

https://www.academia.edu/14939961/L_impact_du_travail_salari%C3%A9_des_%C3%A9tudiants_sur_la_r%C3%A9ussite_et_la_poursuite_des_%C3%A9tudes_universitaires

Giret, J.-F., & Moullet, S. (2017). Activité rémunérée des étudiants et réussite académique : Les effets contrastés des emplois en cours d'études. *Travail et Emploi*, (150), 89-110.

<https://shs.cairn.info/revue-travail-et-emploi-2017-4-page-89?lang=fr>

Givord, P., & Goux, D. (2017). Travail étudiant et réussite dans l'enseignement supérieur : Une analyse économétrique des données françaises. *Revue Économique*, 68(6), 1005-1032.

<https://shs.cairn.info/revue-economique-2017-6-page-1005?lang=fr>

Giret, J.-F., & Moullet, S. (2017). Travail étudiant et réussite dans l'enseignement supérieur. *Mesure et évaluation en éducation*, 40(3), 5-26.

<https://www.erudit.org/fr/revues/mee/2017-v40-n3-mee03805/1048911ar/>

Maetz, I. (2016). Parcours et réussite aux diplômes universitaires. Note Flash – Enseignement supérieur & Recherche, 15.

Montmarquette, C., Viennot-Briot, N., & Dagenais, M. (2007). Dropout, school performance, and working while in school. *The Review of Economics and Statistics*, 89(4), 752-760.

<https://direct.mit.edu/rest/article-abstract/89/4/752/57722/Dropout-School-Performance-and-Working-while-in?redirectedFrom=fulltext>

Froment, B. (2012). Les effets du travail salarié en première année universitaire. *SociologieS*.

<https://journals.openedition.org/sociologies/4006>

Lévy-Garboua, L. (1976). Les demandes de l'étudiant ou les contradictions de l'université de masse. *Revue française de sociologie*, 17(1), 53. doi: 10.2307/3321303

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1), 89-125.

<https://www.jstor.org/stable/3321303>

Verley E. et Zilloniz S., « Fragilités économiques, fragilités studieuses », in Observatoire national de la vie étudiante (OVE), *Les mondes étudiants –enquête conditions de vie 2010*, La documentation française, 2011, pp. 117-136.

Harris Interactive. (2017). Les étudiants et l'alimentation : Rapport Allo Resto.

<https://harris-interactive.fr/wp-content/uploads/sites/6/2017/06/Rapport-Harris-Les-etudiants-et-lalimentation-Allo-Resto.pdf>

Labojka, S. (2012). Enquête Sorbonne Universités. Santé Étudiant.

<https://santetudiant.com/wp-content/uploads/2013/08/Enquete-Sorbonne-Universit%C3%A9-S.-Labojka-2012.pdf>

Observatoire de la Vie Étudiante (OVE). (2019). Conditions de vie des étudiants : 2016.

https://www.ove-national.education.fr/wp-content/uploads/2019/01/oi6_oi6.pdf

Ben Abid-Zarrouk, S., & Weisser, M. (2013). Efficacité du tutorat et étude des profils « efficaces » des tuteurs. Recherches en éducation, 16, 89-110.

<https://journals.openedition.org/ree/7813>

Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation. (2023). Tableaux de réussite et passage par discipline et établissement.

<https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.enseignementsup-recherche.gouv.fr%2Fsites%2Fdefault%2Ffiles%2F2023-11%2Fnf-sies-2023-26-tableaux-russite-et-passage-par-discipline-et-tableau-30240.xlsx&wdOrigin=BROWSELINK>

Pagis, J., & Farges, G. (2020). Quand des élèves de milieux populaires s'engagent pour leurs études au lycée. Revue Française de Pédagogie, 206, 71-87.

<https://journals.openedition.org/rfp/7546>

Ernst & Young. (2023). Proof of Concept..

<https://a.storyblok.com/f/192322/x/fde3fae721/fr-proof-of-concept.pdf>

Beffy, M., Fougère, D., & Maurel, A. (2013). L'impact du travail salarié des étudiants sur la réussite et la poursuite des études universitaires. INSEE.

<https://www.google.com/url?q=https://www.insee.fr/fr/statistiques/fichier/1376784/ES422B.pdf&sa=D&source=docs&ust=1730304512048229&usg=AOvVaw3pHIE-G6Cvuj0Nua7eBT7j>

Service Public. (2024). Étudiant : bourse sur critères sociaux.

<https://www.service-public.fr/particuliers/vosdroits/F12214>

Combien-coute.net.(2024). Prix d'un repas au restaurant en France.

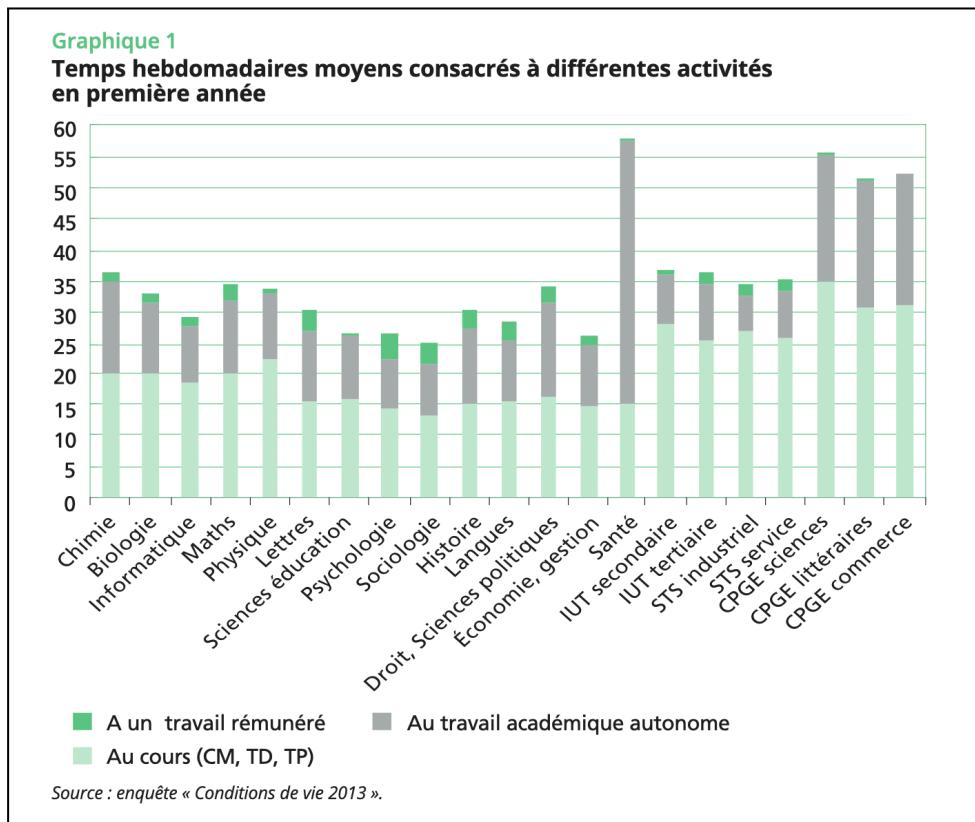
<https://www.combien-coute.net/restaurant/france/>

X - Annexes

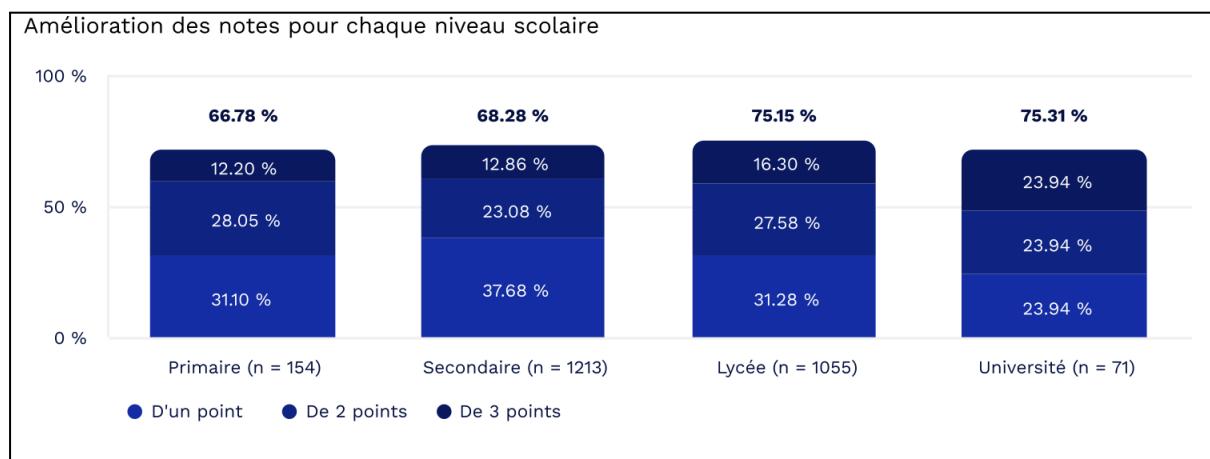
Annexe 1 : Évolution des effectifs des étudiants tutorés et de leurs taux de réussite (2000-2008)

Années Universitaires	Total inscrits	Total tutorés	Proportion de tutorés	Taux de réussite à l'Université	Taux de réussite des tutorés
2000-2001	1265	205	16,20	50,12	88,78
2001-2002	1136	217	19,10	40,30	65,40
2002-2003	1164	155	13,31	34,91	61,29
2003-2004	1236	144	11,65	31,41	60
2004-2005	1190	196	16,47	37	53
2005-2006	1200	256	21,33	39	67
2006-2007	1239	240	19,37	39	64
2007-2008	1162	242	20,82	45	70

Annexe 2 : Temps hebdomadaires moyens consacrés à différentes activités en première année



Annexe 3 : Amélioration des notes pour les élèves qui suivent des cours particuliers



Annexe 4 : Activités salariés et taux de réussite

	Activité salariée des étudiants	Taux de réussite à l'examen	Taux de poursuite des études	En %
Étudiants salariés (plus de 16 heures)	7,2	37,9	74,1	
Étudiants salariés (moins de 16 heures)	5,4	55,8	86,1	
Étudiants non salariés	87,4	66,0	89,9	
Total	100,0	63,4	88,5	

Lecture : 7,2 % des étudiants de l'échantillon sont salariés et travaillent plus de 16 heures.
Champ : étudiants préparant un diplôme universitaire de premier ou de second cycle (Deug, licence ou maîtrise), âgés au plus de 29 ans l'année de l'enquête, et nés en France métropolitaine, à l'exclusion de ceux dont les études vont de pair avec un emploi.
Source : enquêtes Emploi, 1992 à 2002, Insee.

Annexe 5 : Calcul des probabilités conditionnelles en fonction des caractéristiques des emplois étudiants

Conditionnellement à	Probabilité estimée de la			
	Réussite totale		Réussite partielle	
	%	Écart-type	%	Écart-type
Étudiant non salarié	90,8	5,8	5,4	2,3
Étudiant salarié	85,7	16,9	8,6	7,9
Étudiant salarié avec				
aménagement du temps universitaire	91,7	11,7	5,7	5,9
sans aménagement du temps universitaire	77,5	19,8	12,5	8,5
Étudiant salarié dont le temps hebdomadaire				
du travail est				
inférieur à 8 h	97,9	3,2	1,8	2,6
compris entre 8 et 16 h	85,4	15,9	9,1	7,7
plus de 16 h	78,4	19,1	12,0	7,8
Étudiant salarié dans				
le secteur public	95,3	8,7	3,3	5,6
chez des particuliers	88,0	15,8	7,4	7,9
dans le secteur privé	82,1	17,8	10,4	7,6
chez des particuliers ou dans le secteur privé	84,3	17,3	9,3	7,9

Note. La probabilité moyenne estimée de valider complètement son année universitaire sachant qu'on n'a pas occupé d'emploi durant l'année universitaire est de 90,8 %.

Annexe 6 : Difficultés financières des étudiants

		Etudiants avec le moins de difficultés financières estimées (65 %)		Etudiants avec le plus de difficultés financières estimées (35 %)	
		Non boursiers (50 %)	Boursiers (15 %)	Boursiers (22 %)	Non boursiers (13 %)
BTS	Sortis de l'enseignement supérieur	3	4	6	13
	Redoublement ou réorientation	10	10	12	13
	Réussite première année	87	86	82	74
Licence	Sortis de l'enseignement supérieur	5	6	10	19
	Redoublement ou réorientation	32	38	47	48
	Réussite première année	63	56	43	33
DUT et CPGE (1)	Sortis de l'enseignement supérieur	1	1	2	5
	Redoublement ou réorientation	16	20	17	22
	Réussite première année	83	79	81	73

Sources : Panel de bacheliers 2008 et fichiers AGLAE.

Annexe 7 : Statistiques descriptives de la base de donnée d'origine

FORMATION		ASSIDUITE		REVISONS	PARTICULIERS
Autres	: 24	Min.	: 1.00	0 - 1h	:59 Non:139
Economie et gestion	:126	1st Qu.:	8.00	1 - 2h	:56 Oui: 11
		Median :	9.00	2h et plus:	35
		Mean :	8.34		
		3rd Qu.:	10.00		
		Max. :	10.00		
TUTORAT		MOYENNE	BOURSE	EMPLOI	HEMPLOI
Non:115		Min. : 9.00	Non:103	Non:102	Min. : 1.500 Non:68
Oui: 35		1st Qu.:12.06	Oui: 47	Oui: 48	1st Qu.: 7.000 Oui:82
		Median :13.50			Median : 8.000
		Mean :13.45			Mean : 9.479
		3rd Qu.:14.50			3rd Qu.:10.500
		Max. :17.53			Max. :30.000
					NA's :102
ARGENT		RESTAURANT	UBER	DEPENSES	CAF
Non:83		Min. : 0.000	Non:103	0 - 100 € :85	Non:69
Oui:67		1st Qu.: 1.000	Oui: 47	100 € et plus:65	Oui:81
		Median : 2.000			
		Mean : 2.822			
		3rd Qu.: 3.000			
		Max. :50.000			
TRANSPORT		GENRE	AGE	STRESS	
0 - 15 €	:32	Femme:82	Min. :18.00	Min. :	1.00
15 - 30 €	:64	Homme:68	1st Qu.:20.00	1st Qu.:	1.00
30 € et plus:	54		Median :21.00	Median :	3.00
			Mean :21.41	Mean :	3.48
			3rd Qu.:22.00	3rd Qu.:	5.00
			Max. :40.00	Max. :	10.00
STATUT		SOMMEIL	SANTE	STRUCTURE	
Classe aisée	: 18	Min. :4.0	Non:125	Autres :	11
Classe moyenne	:114	1st Qu.:7.0	Oui: 25	Université:	139
Classe populaire:	18	Median :7.0			
		Mean :7.1			
		3rd Qu.:8.0			
		Max. :9.0			
TRAJET					
Min. :	5.00				
1st Qu.:	15.00				
Median :	30.00				
Mean :	31.63				
3rd Qu.:	40.00				
Max. :	90.00				

Annexe 8 : Écart-types des variables quantitatives

MOYENNE[1] 1.78341
ASSIDUITE[1] 1.842197
HEMPLOI[1] NA
RESTAURANT[1] 4.505059
AGE[1] 2.58334
STRESS[1] 2.448722
SOMMEIL[1] 0.8687292
TRAJET[1] 19.88239

Annexe 9 : RosnerTest pour les valeurs de la variable ASSIDUITE

Number of Outliers Detected: 5								
i	Mean.i	SD.i	Value	Obs.Num	R.i+1	lambda.i+1	Outlier	
1	0	8.340000	1.842197	1	44	3.984372	3.517009	TRUE
2	1	8.389262	1.746469	2	48	3.658388	3.514898	TRUE
3	2	8.432432	1.670725	2	113	3.850085	3.512770	TRUE
4	3	8.476190	1.589061	2	128	4.075483	3.510625	TRUE
5	4	8.520548	1.500433	3	39	3.679303	3.508462	TRUE
6	5	8.558621	1.433119	4	65	3.180908	3.506282	FALSE
7	6	8.590278	1.386313	5	10	2.589804	3.504084	FALSE
8	7	8.615385	1.357935	5	17	2.662414	3.501868	FALSE
9	8	8.640845	1.328047	5	52	2.741502	3.499634	FALSE
10	9	8.666667	1.296515	5	69	2.828093	3.497381	FALSE

Annexe 10 : RosnerTest pour les valeurs de la variable HEMPLOI

Number of Outliers Detected: 1								
	i	Mean.i	SD.i	Value	Obs.Num	R.i+1	lambda.i+1	Outlier
1	0	9.479167	5.213483	30.0	118	3.936108	3.111796	TRUE
2	1	9.042553	4.292093	21.0	130	2.785924	3.103243	FALSE
3	2	8.782609	3.947787	20.0	113	2.841438	3.094456	FALSE
4	3	8.533333	3.607757	17.0	123	2.346795	3.085425	FALSE
5	4	8.340909	3.407879	16.0	25	2.247466	3.076135	FALSE
6	5	8.162791	3.234341	16.0	48	2.423124	3.066572	FALSE
7	6	7.976190	3.030239	15.0	94	2.317906	3.056723	FALSE
8	7	7.804878	2.854553	1.5	72	2.208709	3.046571	FALSE
9	8	7.962500	2.704193	14.0	31	2.232644	3.036097	FALSE
10	9	7.807692	2.553675	14.0	68	2.424861	3.025284	FALSE

Annexe 11 : RosnerTest pour les valeurs de la variable RESTAURANT

Number of Outliers Detected: 4								
	i	Mean.i	SD.i	Value	Obs.Num	R.i+1	lambda.i+1	Outlier
1	0	2.822000	4.505059	50	25	10.472228	3.517009	TRUE
2	1	2.505369	2.300611	20	58	7.604343	3.514898	TRUE
3	2	2.387162	1.797969	12	53	5.346499	3.512770	TRUE
4	3	2.321769	1.617901	10	88	4.745797	3.510625	TRUE
5	4	2.269178	1.492074	7	74	3.170635	3.508462	FALSE
6	5	2.236552	1.444040	7	128	3.298695	3.506282	FALSE
7	6	2.203472	1.392860	6	79	2.725707	3.504084	FALSE
8	7	2.176923	1.360701	6	83	2.809639	3.501868	FALSE
9	8	2.150000	1.326743	6	139	2.901842	3.499634	FALSE
10	9	2.122695	1.290812	5	10	2.229066	3.497381	FALSE

Annexe 12 : RosnerTest pour les valeurs de la variable AGE

Number of Outliers Detected: 4								
	i	Mean.i	SD.i	Value	Obs.Num	R.i+1	lambda.i+1	Outlier
1	0	21.41333	2.583340	40	58	7.194820	3.517009	TRUE
2	1	21.28859	2.090181	33	67	5.603061	3.514898	TRUE
3	2	21.20946	1.859900	32	65	5.801676	3.512770	TRUE
4	3	21.13605	1.637068	28	117	4.192828	3.510625	TRUE
5	4	21.08904	1.539911	25	78	2.539730	3.508462	FALSE
6	5	21.06207	1.510246	25	132	2.607477	3.506282	FALSE
7	6	21.03472	1.479053	25	143	2.680958	3.504084	FALSE
8	7	21.00699	1.446202	18	21	2.079235	3.501868	FALSE
9	8	21.02817	1.428900	18	80	2.119231	3.499634	FALSE
10	9	21.04965	1.410806	18	93	2.161633	3.497381	FALSE

Annexe 13 : RosnerTest pour les valeurs de la variable SOMMEIL

Number of Outliers Detected: 1								
	i	Mean.i	SD.i	Value	Obs.Num	R.i+1	lambda.i+1	Outlier
1	0	7.099533	0.8687292	4	74	3.567894	3.517009	TRUE
2	1	7.120336	0.8333313	5	25	2.544409	3.514898	FALSE
3	2	7.134662	0.8175419	5	52	2.611074	3.512770	FALSE
4	3	7.149184	0.8009553	5	55	2.683275	3.510625	FALSE
5	4	7.163904	0.7835053	5	67	2.761825	3.508462	FALSE
6	5	7.178828	0.7651157	5	88	2.847710	3.506282	FALSE
7	6	7.193958	0.7456993	5	123	2.942149	3.504084	FALSE
8	7	7.209301	0.7251543	9	53	2.469404	3.501868	FALSE
9	8	7.196690	0.7118119	9	93	2.533408	3.499634	FALSE
10	9	7.183901	0.6977840	9	108	2.602667	3.497381	FALSE

Annexe 14 : RosnerTest pour les valeurs de la variable TRAJET

Number of Outliers Detected: 0								
	i	Mean.i	SD.i	Value	Obs.Num	R.i+1	lambda.i+1	Outlier
1	0	31.62667	19.88239	90	85	2.935932	3.517009	FALSE
2	1	31.23490	19.35982	90	89	3.035416	3.514898	FALSE
3	2	30.83784	18.80695	90	106	3.145760	3.512770	FALSE
4	3	30.43537	18.22052	90	111	3.269096	3.510625	FALSE
5	4	30.02740	17.59661	90	132	3.408191	3.508462	FALSE
6	5	29.61379	16.93049	80	2	2.976063	3.506282	FALSE
7	6	29.26389	16.45506	80	133	3.083313	3.504084	FALSE
8	7	28.90909	15.95059	80	143	3.203074	3.501868	FALSE
9	8	28.54930	15.41372	75	54	3.013595	3.499634	FALSE
10	9	28.21986	14.95856	70	124	2.793059	3.497381	FALSE

Annexe 15 : Deuxième RosnerTest pour les valeurs de la variable TRAJET

Number of Outliers Detected: 1								
	i	Mean.i	SD.i	Value	Obs.Num	R.i+1	lambda.i+1	Outlier
1	0	8.511628	3.824247	21.0	116	3.265577	3.066572	TRUE
2	1	8.214286	3.329731	17.0	110	2.638566	3.056723	FALSE
3	2	8.000000	3.063903	15.0	84	2.284668	3.046571	FALSE
4	3	7.825000	2.887972	1.5	64	2.190118	3.036097	FALSE
5	4	7.987179	2.734976	14.0	30	2.198491	3.025284	FALSE
6	5	7.828947	2.584455	14.0	60	2.387757	3.014109	FALSE
7	6	7.662162	2.403857	2.0	33	2.355449	3.002552	FALSE
8	7	7.819444	2.236556	2.0	48	2.601967	2.990585	FALSE
9	8	7.985714	2.030958	12.0	45	1.976548	2.978183	FALSE
10	9	7.867647	1.935744	12.0	80	2.134763	2.965315	FALSE

Annexe 16 : Statistiques descriptives après suppression des valeurs atypiques

FORMATION		ASSIDUITE		REVISIONS		PARTICULIERS	
Autres	: 21	Min.	: 5.000	0 - 1h	:53	Non:127	
Economie et gestion	:114	1st Qu.:	8.000	1 - 2h	:50	Oui: 8	
		Median	: 9.000	2h et plus:	32		
		Mean	: 8.556				
		3rd Qu.:	10.000				
		Max.	:10.000				
TUTORAT		MOYENNE	BOURSE	EMPLOI	HEMPLOI	LOGEMENT	ARGENT
Non:101	Min. : 9.60	Non:92	Non:93	Min. : 1.500	Non:63	Non:75	
Oui: 34	1st Qu.:12.00	Oui:43	Oui:42	1st Qu.: 6.250	Oui:72	Oui:60	
	Median :13.50			Median : 8.000			
	Mean :13.45			Mean : 8.214			
	3rd Qu.:14.47			3rd Qu.:10.000			
	Max. :17.53			Max. :17.000			
				NA's :93			
RESTAURANT		UBER	DEPENSES		CAF	TRANSPORT	
Min. :0.000	Non:98	0 - 100 €	:80	Non:61	0 - 15 €	:28	
1st Qu.:1.000	Oui:37	100 € et plus:	55	Oui:74	15 - 30 €	:61	
Median :2.000					30 € et plus:	46	
Mean :2.158							
3rd Qu.:3.000							
Max. :6.000							
GENRE		AGE	STRESS		STATUT		
Femme:74	Min. :18.00	Min. : 1.000	Classe aisée : 15				
Homme:61	1st Qu.:20.00	1st Qu.: 1.000	Classe moyenne :106				
	Median :21.00	Median : 3.000	Classe populaire: 14				
	Mean :21.07	Mean : 3.444					
	3rd Qu.:22.00	3rd Qu.: 5.000					
	Max. :25.00	Max. :10.000					
SOMMEIL		SANTE	STRUCTURE		TRAJET		
Min. :5.000	Non:111	Autres : 9	Min. : 5.00				
1st Qu.:7.000	Oui: 24	Université:126	1st Qu.:20.00				
Median :7.000			Median :30.00				
Mean :7.177			Mean :32.98				
3rd Qu.:8.000			3rd Qu.:40.00				
Max. :9.000			Max. :90.00				

Annexe 17 : Écart-types des variables quantitatives

MOYENNE[1]	1.753759
ASSIDUITE[1]	1.40716
HEMPLOI[1]	NA
RESTAURANT[1]	1.349487
AGE[1]	1.526874
STRESS[1]	2.433187
SOMMEIL[1]	0.7565292
TRAJET[1]	20.27367

Annexe 18 : Effectif des modalités de chaque variable qualitative

\$FORMATION	\$UBER		
Autres Economie et gestion	Non Oui		
21	98	37	
\$REVISIONS	\$DEPENSES		
0 - 1h	1 - 2h	2h et plus	0 - 100 €
53	50	32	100 € et plus
			80
			55
\$PARTICULIERS	\$CAF		
Non Oui	Non Oui		
127	8		61
			74
\$TUTORAT	\$TRANSPORT		
Non Oui	0 - 15 €	15 - 30 €	30 € et plus
101	28	61	46
\$BOURSE	\$GENRE		
Non Oui	Femme	Homme	
92	74	61	
\$EMPLOI	\$STATUT		
Non Oui	Classe aisée	Classe moyenne	Classe populaire
93	15	106	14
42			
\$LOGEMENT	\$SANTE		
Non Oui	Non Oui		
63	111	24	
72			
\$ARGENT	\$STRUCTURE		
Non Oui	Autres	Université	
75	9	126	
60			

Annexe 19 : Régressions linéaires simples dont les hypothèses ont été validées

```
RLS: MOYENNE_ASSIDUITE
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 10.7214    0.9057 11.837 < 2e-16 ***
ASSIDUITE    0.3188    0.1045  3.052  0.00275 **

```



```
RLS: MOYENNE_LNRESTAURANT
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 14.6043    0.3917 37.286 < 2e-16 ***
log(RESTAURANT + 1) -1.0889    0.3426 -3.178  0.00184 **

```



```
RLS: MOYENNE AGE2
Call:
lm(formula = MOYENNE ~ AGE, data = Budget3, weights = 1/AGE)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 13.97379   2.12383   6.580 9.92e-10 ***
AGE         -0.02489   0.10108  -0.246   0.806

```



```
RLS: MOYENNE_STRESS
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 13.95714   0.25768  54.16 <2e-16 ***
STRESS      -0.14743   0.06118  -2.41  0.0173 *

```



```
RLS: MOYENNE SOMMEIL
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 11.2151    1.4375   7.802 1.59e-12 ***
SOMMEIL     0.3113    0.1992   1.563     0.12

```



```
RLS: MOYENNE TRAJET
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 13.331996   0.289827 46.000 <2e-16 ***
TRAJET      0.003558   0.007495  0.475     0.636

```

Annexe 20 : Test de Shapiro-Wilk pour vérifier la normalité des résidus

RLS: MOYENNE_ASSIDUITE	RLS: MOYENNE_STRESS
Shapiro-Wilk normality test	Shapiro-Wilk normality test
data: residus W = 0.98914, p-value = 0.373	data: residus W = 0.98818, p-value = 0.303
RLS: MOYENNE_LNRESTAURANT	RLS: MOYENNE_SOMMEIL
Shapiro-Wilk normality test	Shapiro-Wilk normality test
data: residus W = 0.99147, p-value = 0.5866	data: residus W = 0.98739, p-value = 0.2544
RLS: MOYENNE_AGE2	RLS: MOYENNE_TRAJET
Shapiro-Wilk normality test	Shapiro-Wilk normality test
data: residus W = 0.98771, p-value = 0.2729	data: residus W = 0.98968, p-value = 0.4167

Annexe 21 : Test de Ramsey afin de vérifier la forme fonctionnelle du modèle

```
RLS: MOYENNE_ASSIDUITE

RESET test

data: get(i)
RESET = 0.27512, df1 = 2, df2 = 131, p-value = 0.7599


RLS: MOYENNE_LNRESTAURANT

RESET test

data: MOYENNE_LNRESTAURANT
RESET = 1.8176, df1 = 2, df2 = 131, p-value = 0.1665


RLS: MOYENNE_AGE2

RESET test

data: MOYENNE_AGE2
RESET = 1.898, df1 = 2, df2 = 131, p-value = 0.154


RLS: MOYENNE_STRESS

RESET test

data: get(i)
RESET = 2.4654, df1 = 2, df2 = 131, p-value = 0.08891


RLS: MOYENNE_SOMMEIL

RESET test

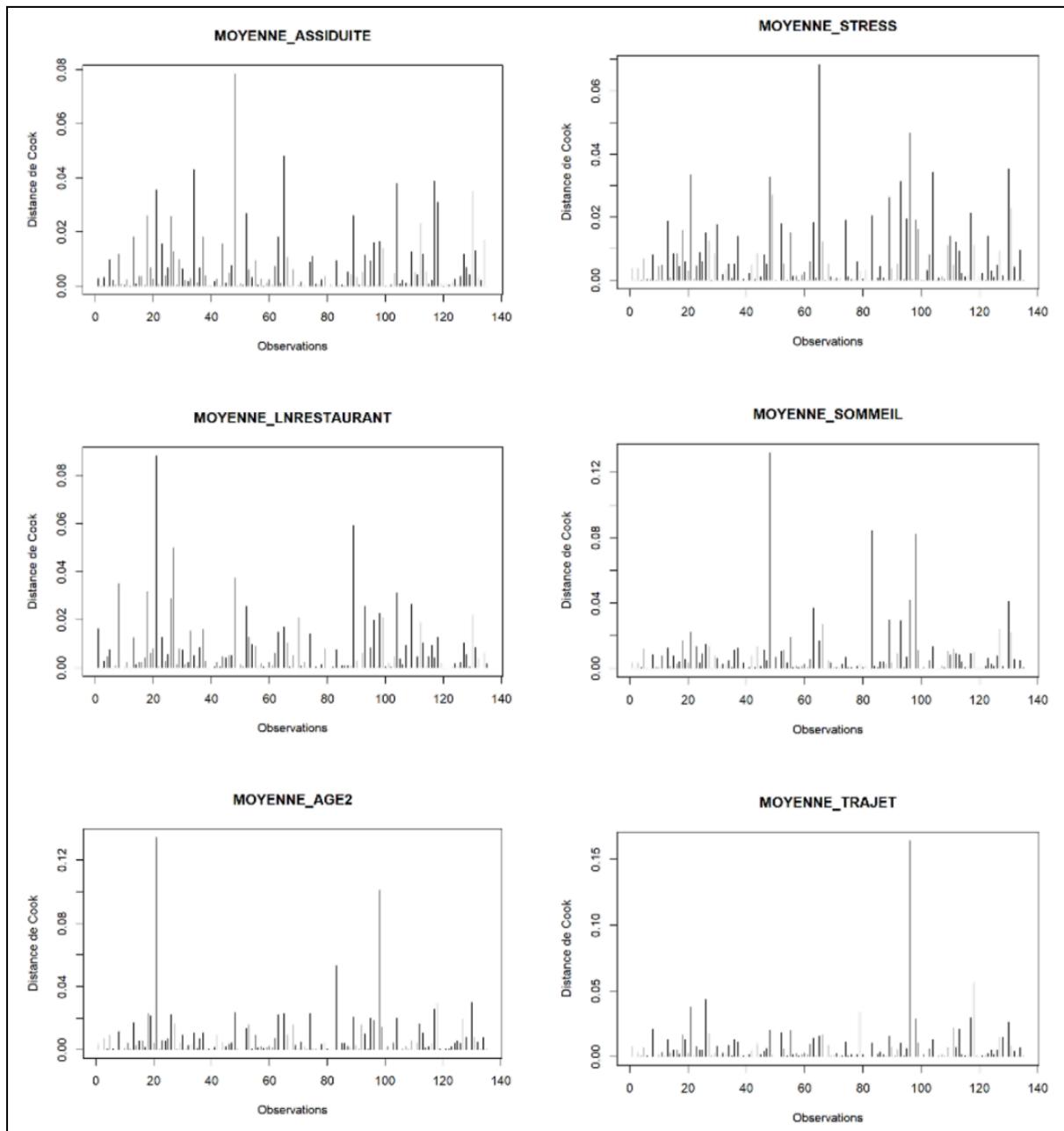
data: get(i)
RESET = 0.97424, df1 = 2, df2 = 131, p-value = 0.3802


RLS: MOYENNE_TRAJET

RESET test

data: get(i)
RESET = 0.83272, df1 = 2, df2 = 131, p-value = 0.4372
```

Annexe 22 : Distance de Cook pour vérifier l'absence d'observations influentes



Annexe 23 : Test de Breusch-Pagan afin de vérifier l'homoscédasticité des résidus

```
RLS: MOYENNE_ASSIDUITE

    studentized Breusch-Pagan test

data: get(i)
BP = 0.50455, df = 1, p-value = 0.4775

RLS: MOYENNE_LNRESTAURANT

    studentized Breusch-Pagan test

data: get(i)
BP = 0.18724, df = 1, p-value = 0.6652

RLS: MOYENNE_AGE2

    studentized Breusch-Pagan test

data: MOYENNE_AGE2
BP = 2.5533, df = 1, p-value = 0.1101

RLS: MOYENNE_STRESS

    studentized Breusch-Pagan test

data: get(i)
BP = 0.13445, df = 1, p-value = 0.7139

RLS: MOYENNE_SOMMEIL

    studentized Breusch-Pagan test

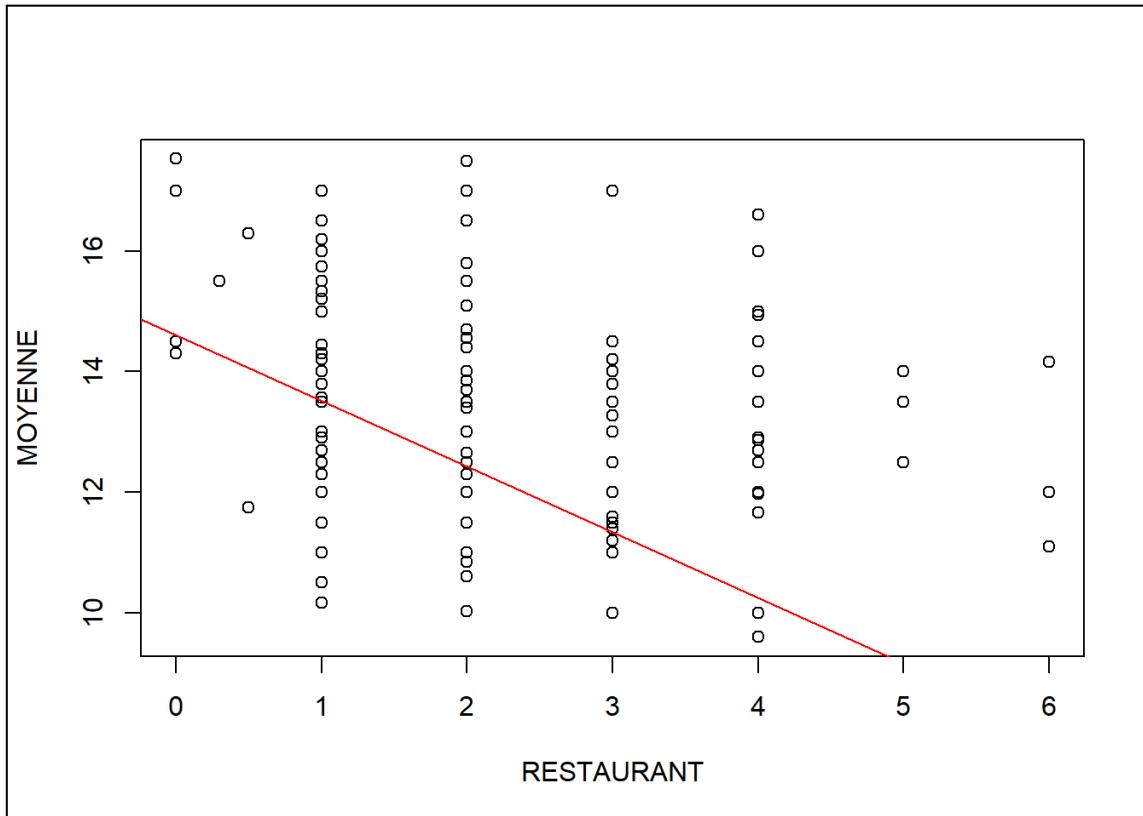
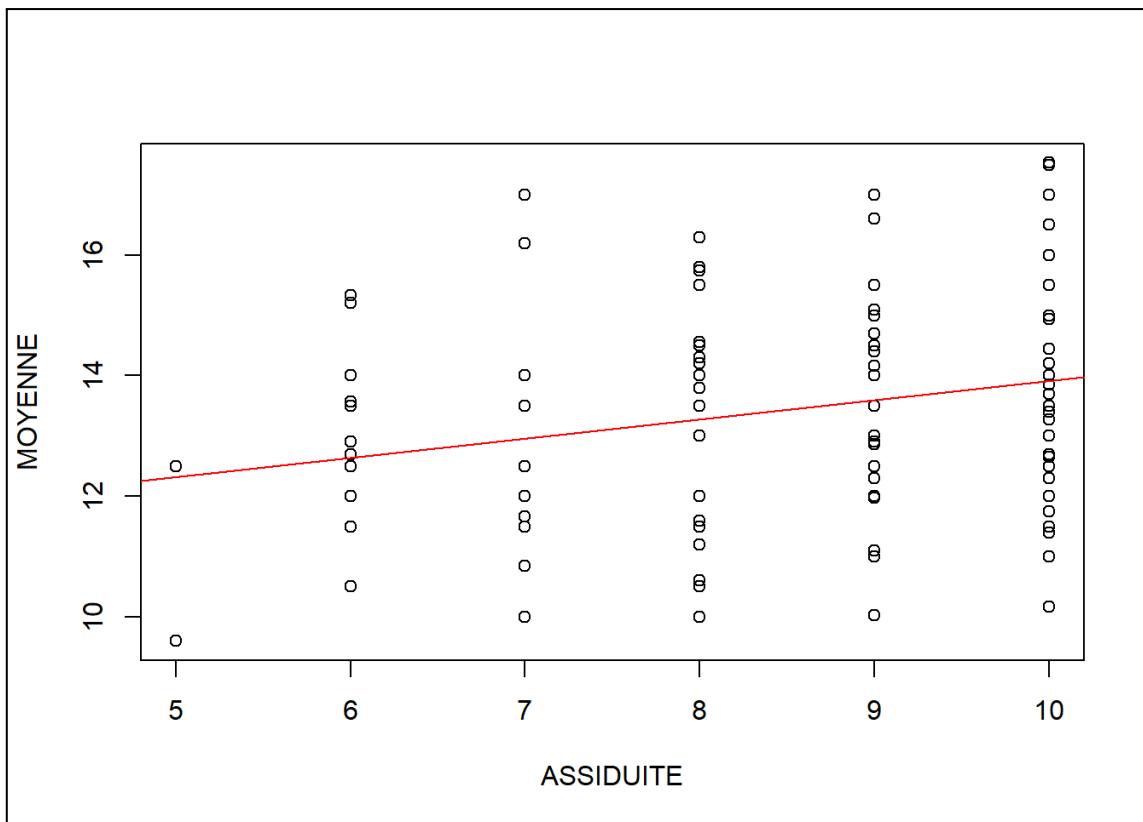
data: get(i)
BP = 0.00061085, df = 1, p-value = 0.9803

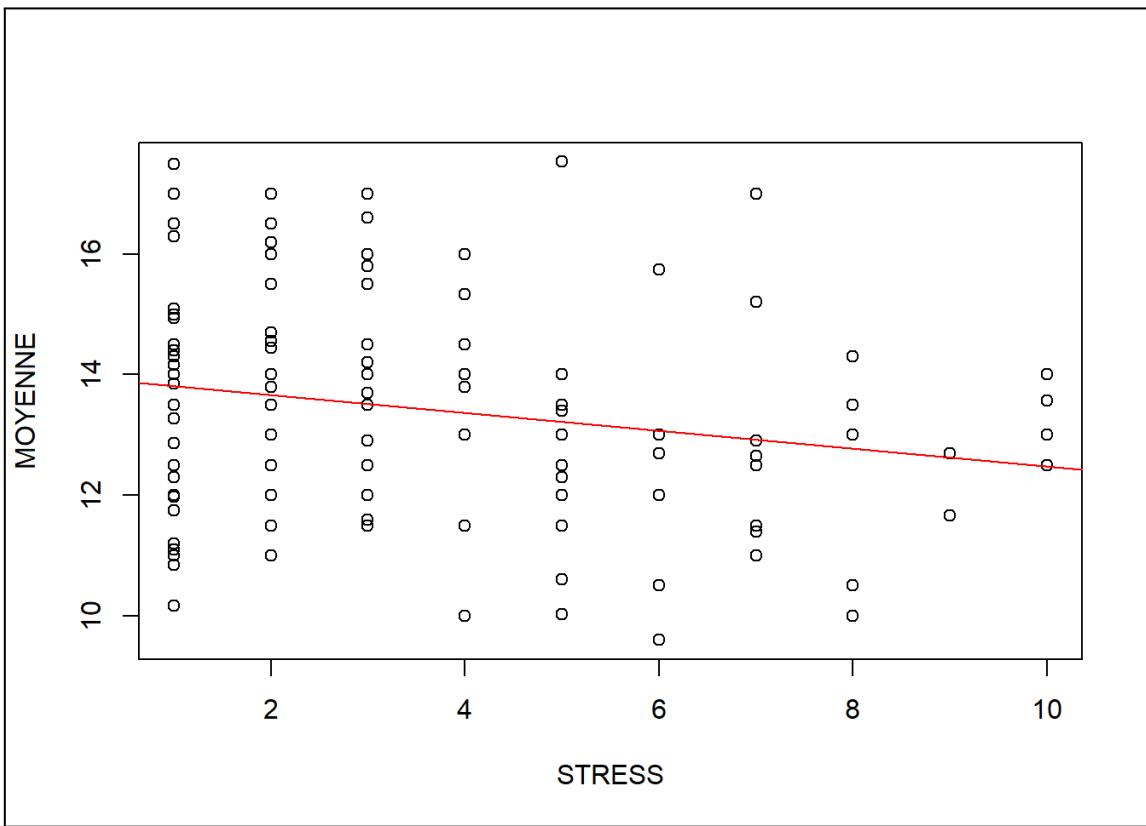
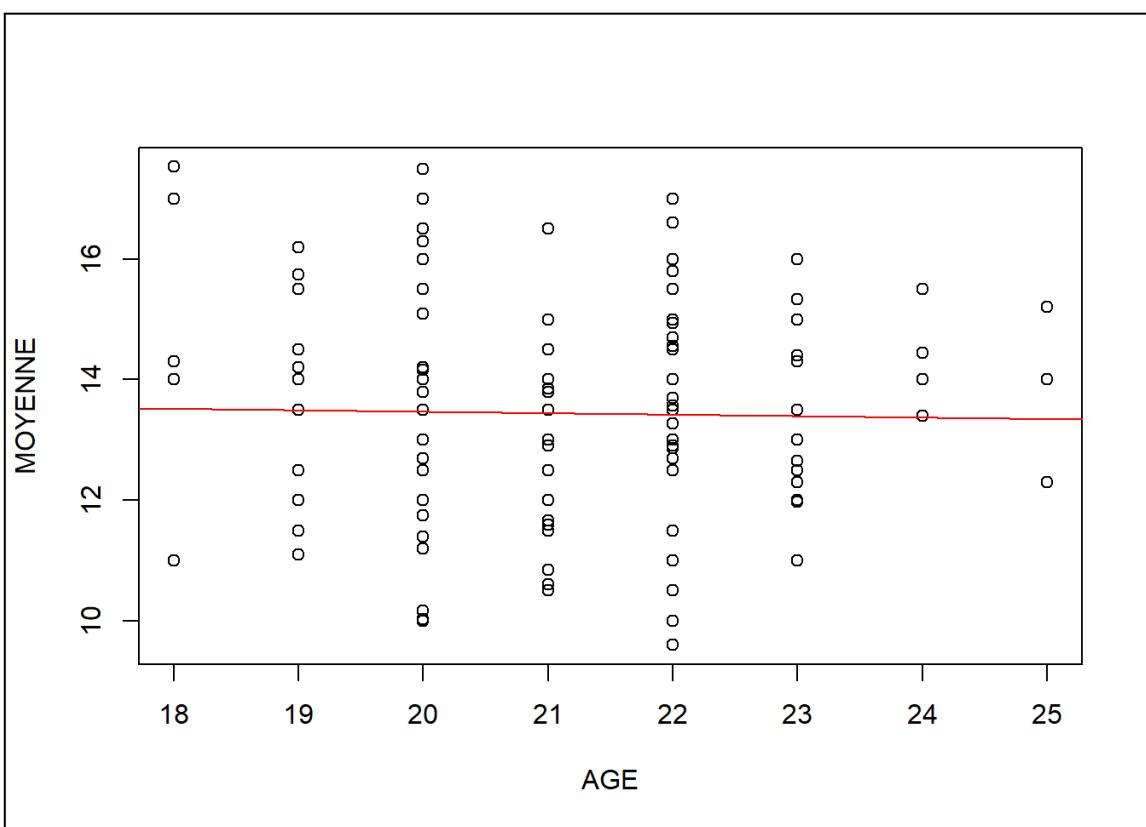
RLS: MOYENNE_TRAJET

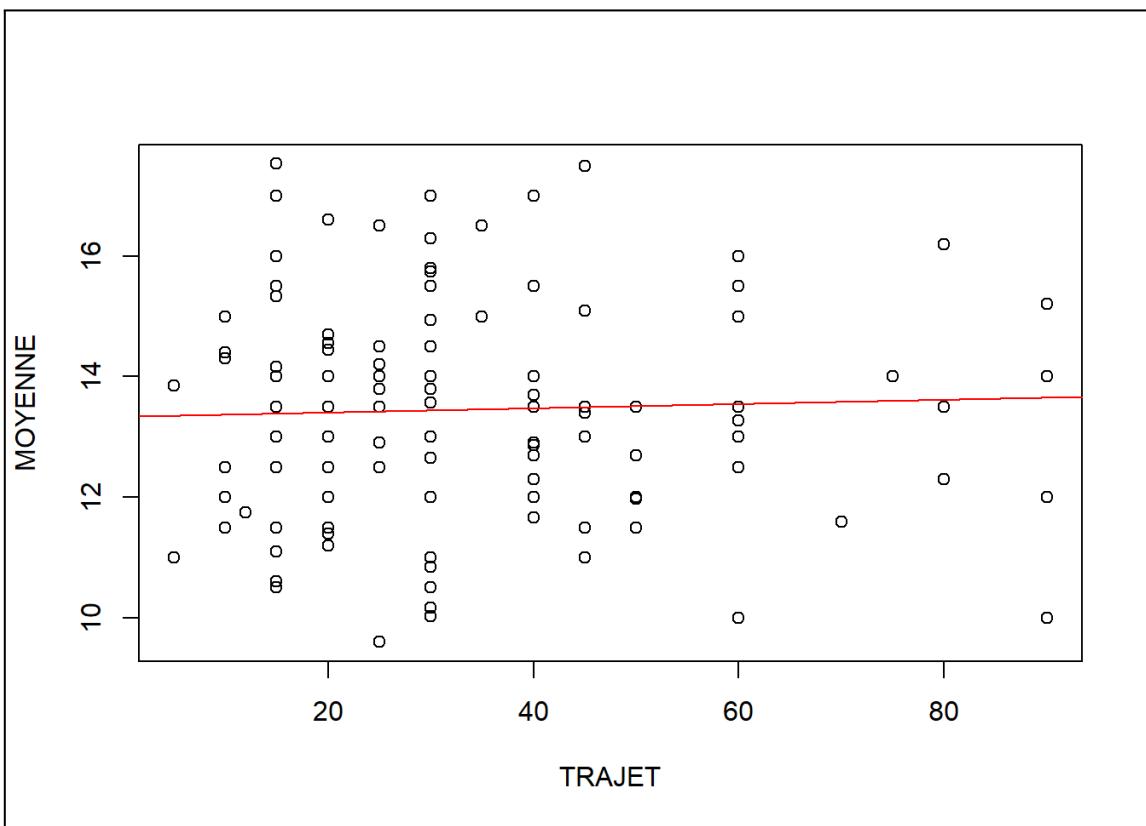
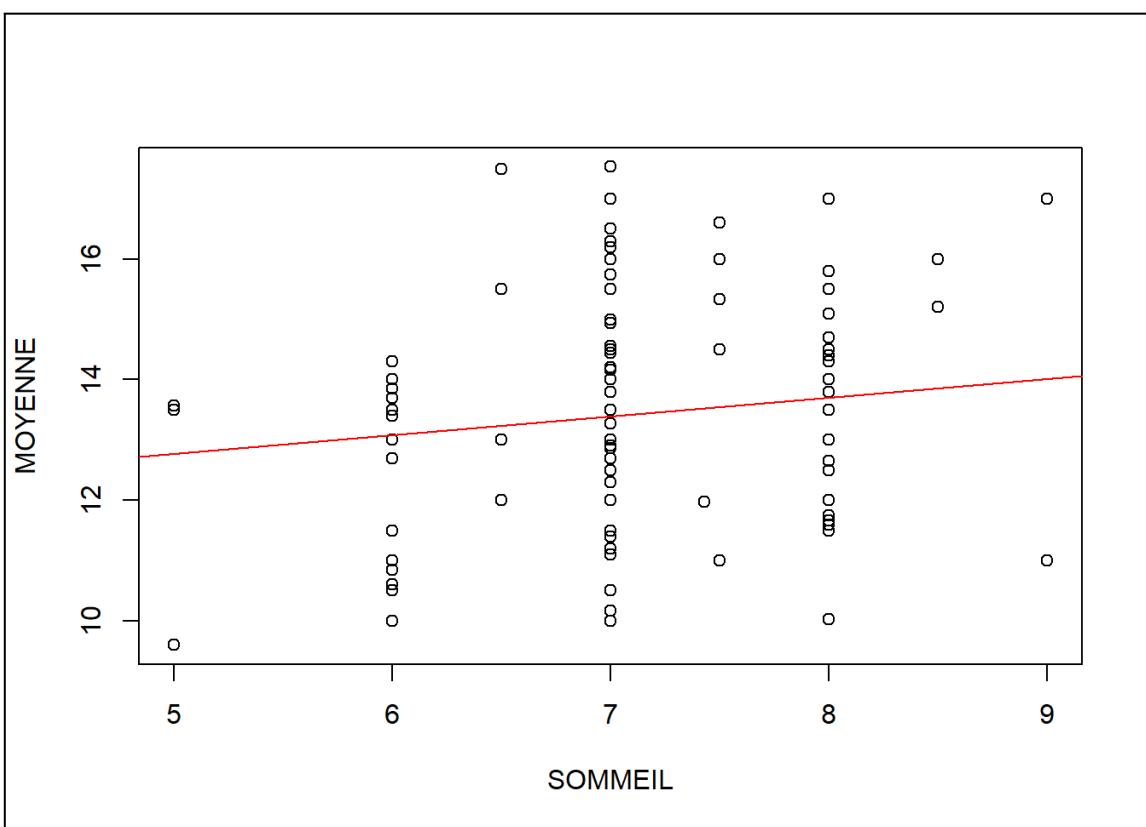
    studentized Breusch-Pagan test

data: get(i)
BP = 0.13688, df = 1, p-value = 0.7114
```

Annexe 24 : Nuage de points illustrant la répartition de chaque variable quantitative par rapport à la variable MOYENNE.



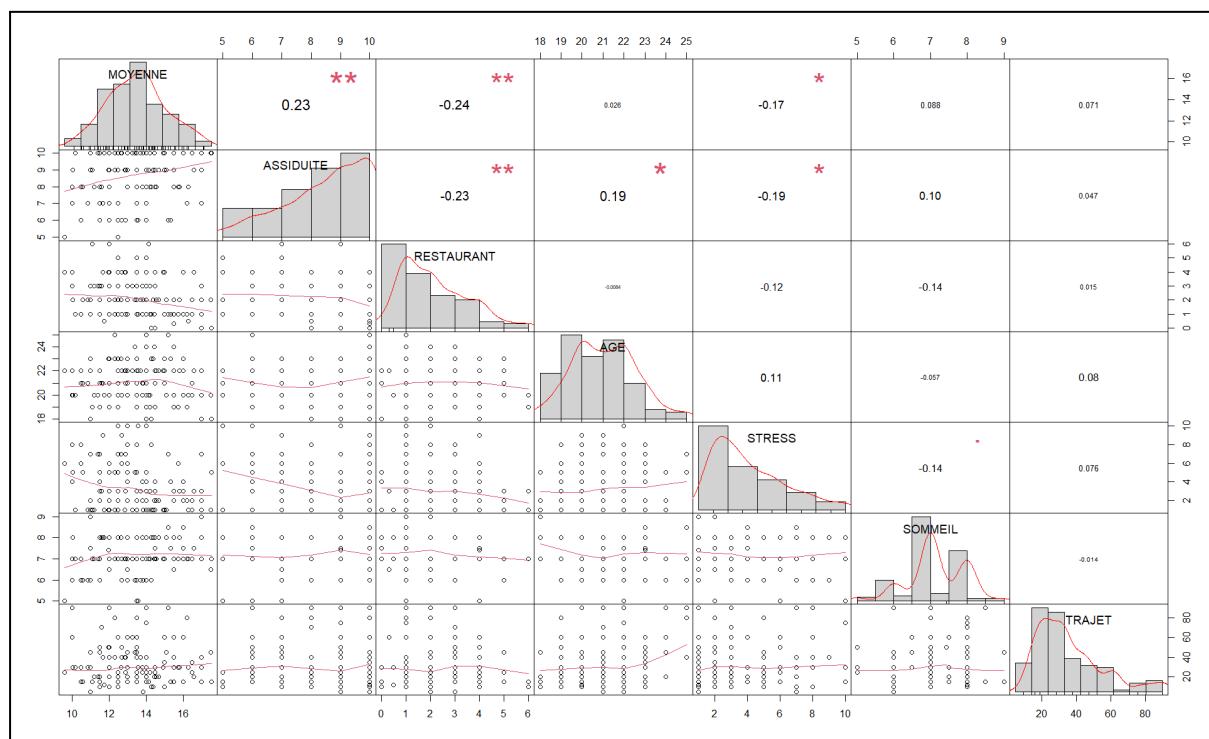




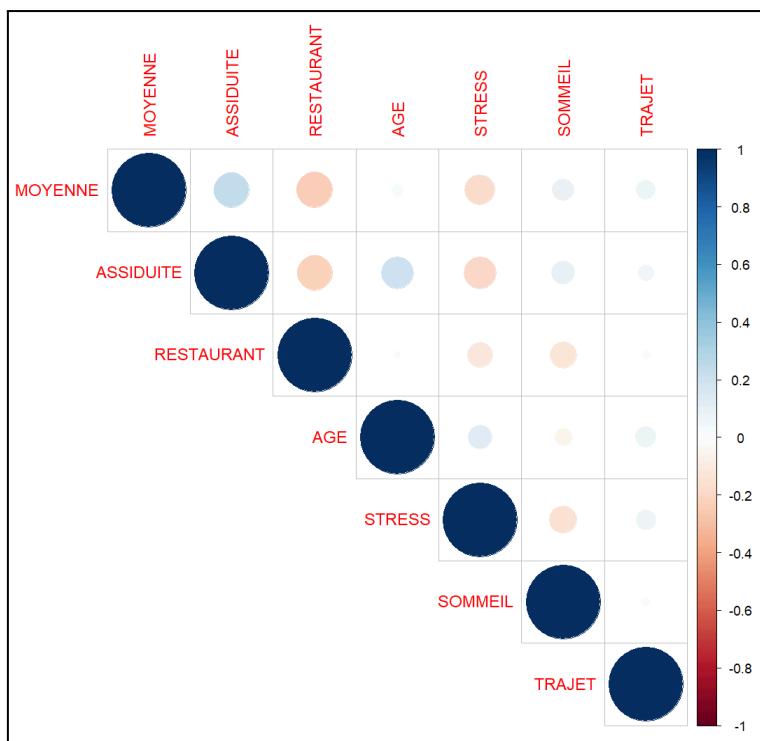
Annexe 25 : Corrélation de Spearman entre variables quantitatives

	MOYENNE	ASSIDUITE	RESTAURANT	AGE	STRESS
MOYENNE	1.00000000	0.23182403	-0.239765703	0.025757908	-0.17301474
ASSIDUITE	0.23182403	1.00000000	-0.226843529	0.192969746	-0.19288877
RESTAURANT	-0.23976570	-0.22684353	1.000000000	-0.008443852	-0.11995179
AGE	0.02575791	0.19296975	-0.008443852	1.000000000	0.11071464
STRESS	-0.17301474	-0.19288877	-0.119951790	0.110714640	1.00000000
SOMMEIL	0.08844256	0.10463386	-0.135247334	-0.057081542	-0.14310863
TRAJET	0.07118009	0.04737948	0.015448710	0.079644454	0.07587723
		SOMMEIL	TRAJET		
MOYENNE	0.08844256	0.07118009			
ASSIDUITE	0.10463386	0.04737948			
RESTAURANT	-0.13524733	0.01544871			
AGE	-0.05708154	0.07964445			
STRESS	-0.14310863	0.07587723			
SOMMEIL	1.00000000	-0.01353375			
TRAJET	-0.01353375	1.00000000			

Annexe 26 : Graphique des corrélations selon la méthode de Spearman



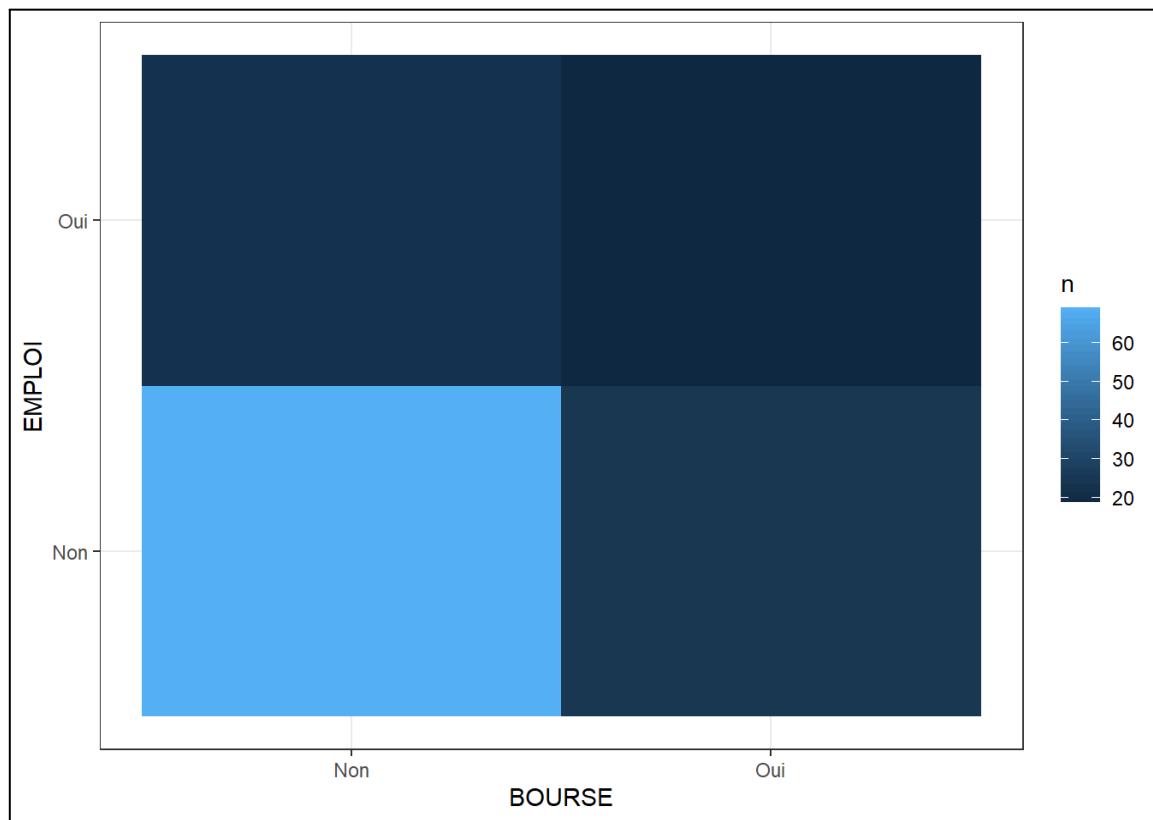
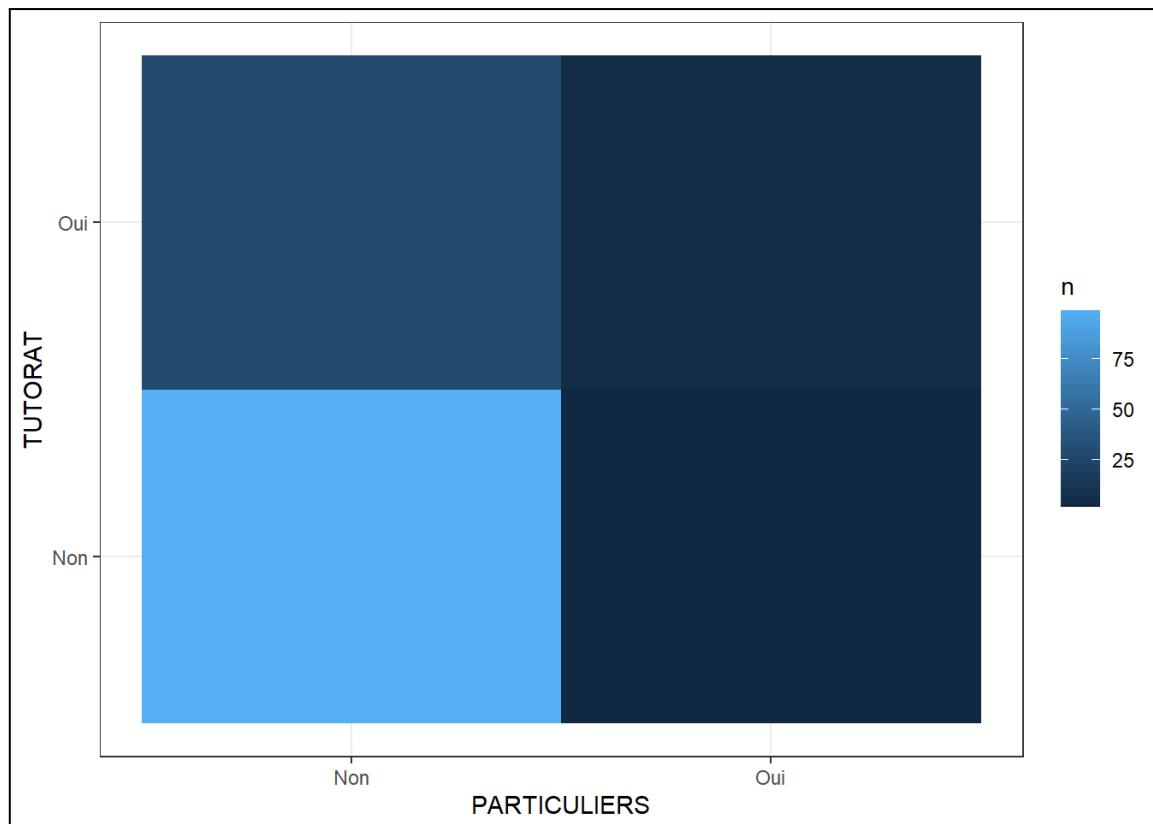
Annexe 27 : Représentation alternative des corrélations en appliquant également la méthode de Spearman.

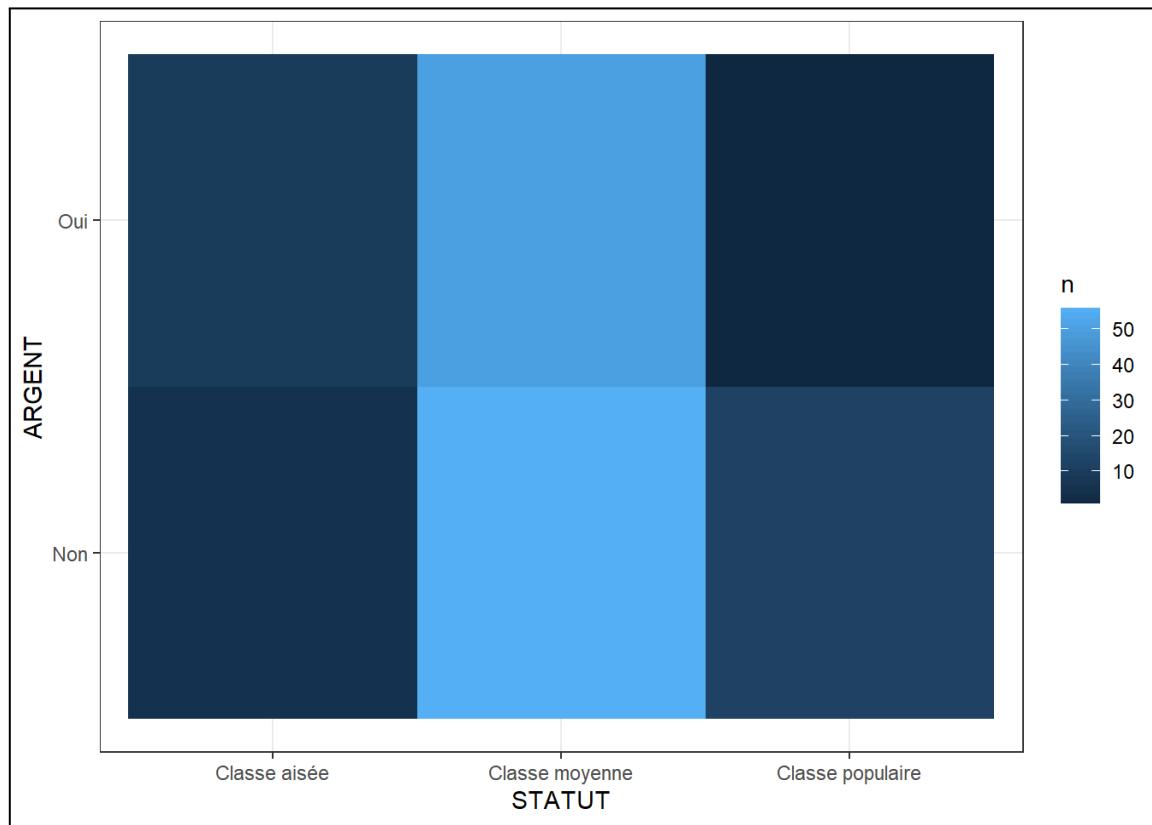
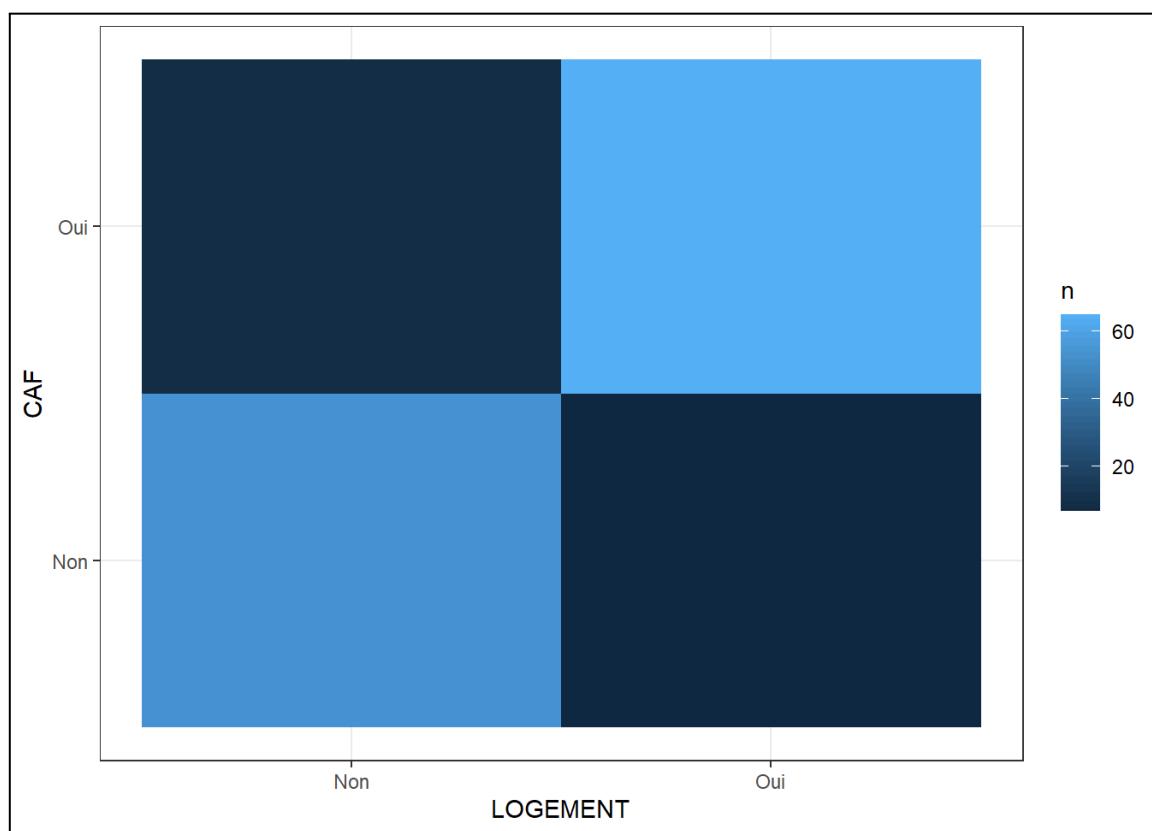


Annexe 28 : Tableau de contingence permettant de visualiser la relation entre deux variables qualitatives

PARTICULIERS			TUTORAT	n	BOURSE			EMPLOI	n
1	Non	Non	99		1	Non	Non	69	
2	Non	Oui	28		2	Non	Oui	23	
3	Oui	Non	2		3	Oui	Non	24	
4	Oui	Oui	6		4	Oui	Oui	19	
STATUT ARGENT			n		LOGEMENT CAF			n	
1	Classe aisée	Non	6		1	Non	Non	54	
2	Classe aisée	Oui	9		2	Non	Oui	9	
3	Classe moyenne	Non	56		3	Oui	Non	7	
4	Classe moyenne	Oui	50		4	Oui	Oui	65	
5	Classe populaire	Non	13						
6	Classe populaire	Oui	1						

Annexe 29 : Cartes des points chauds illustrant le niveau de concentration des individus pour une paire de modalités

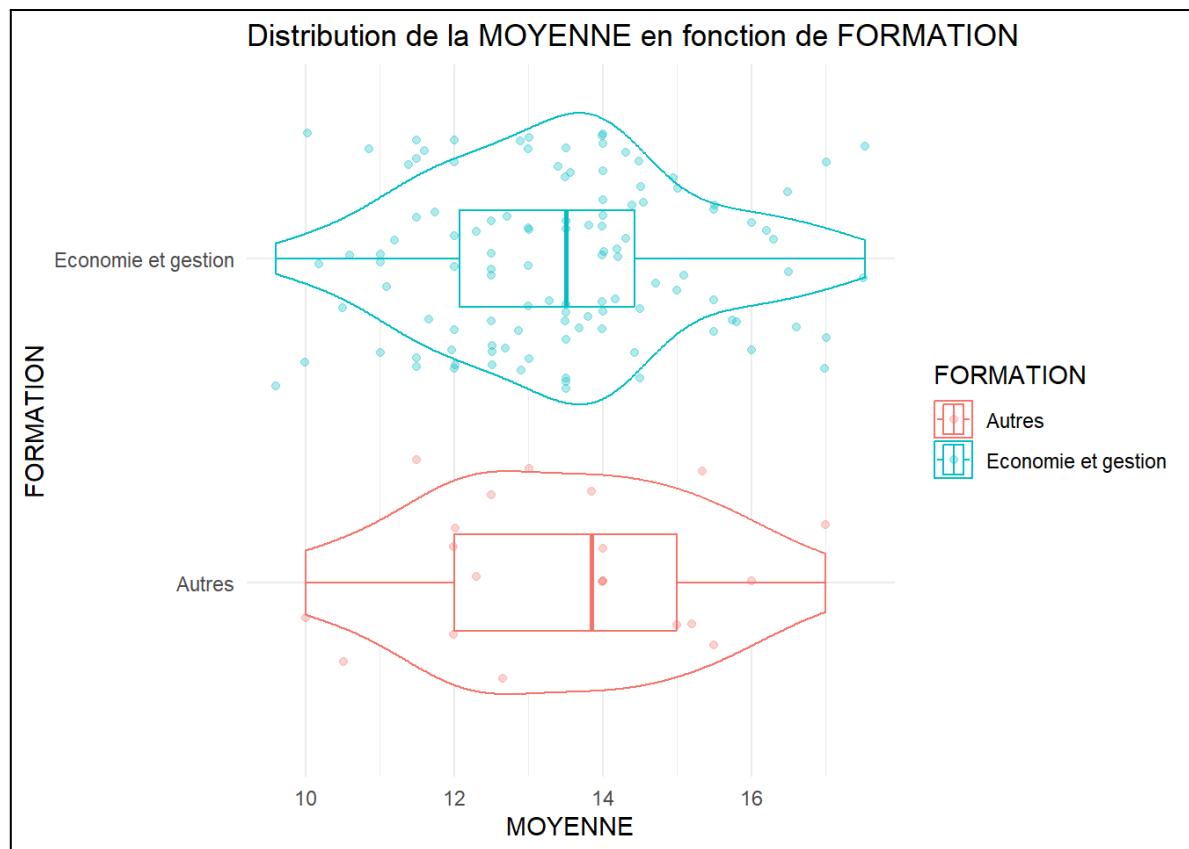


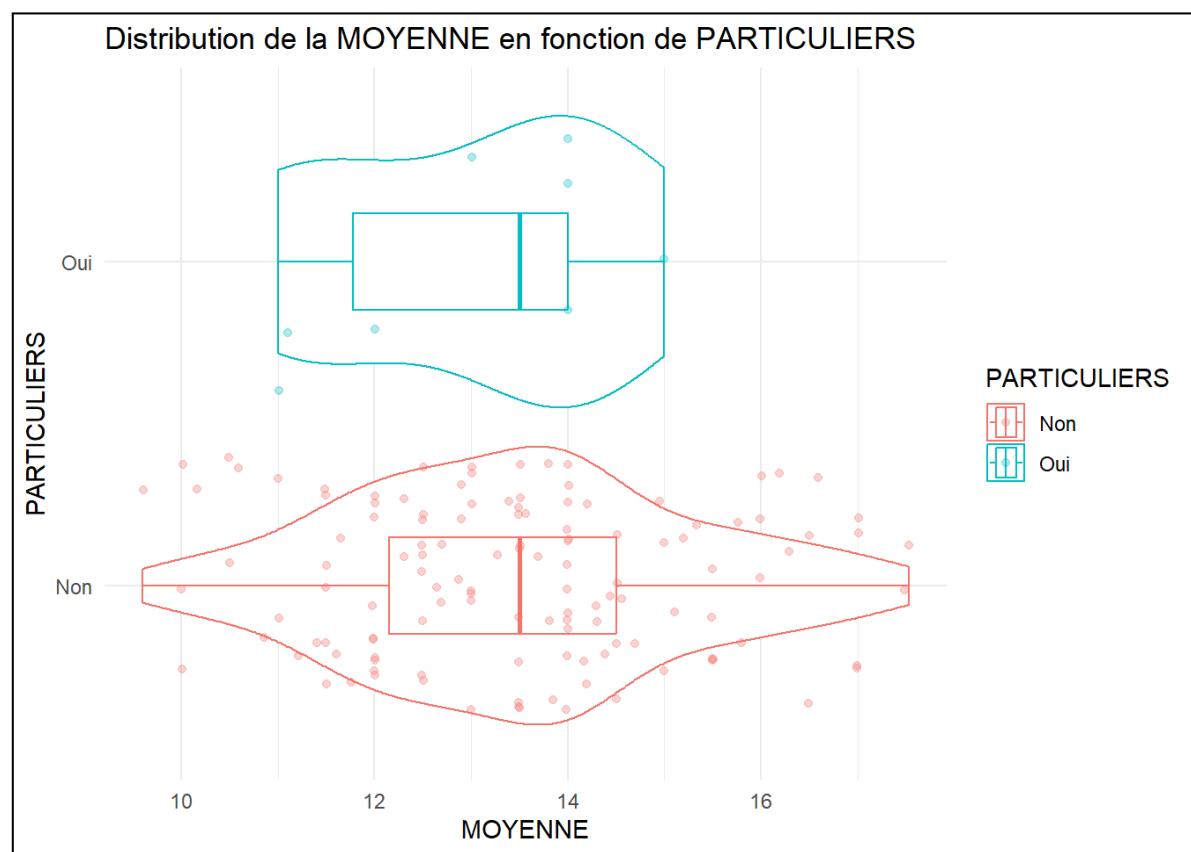
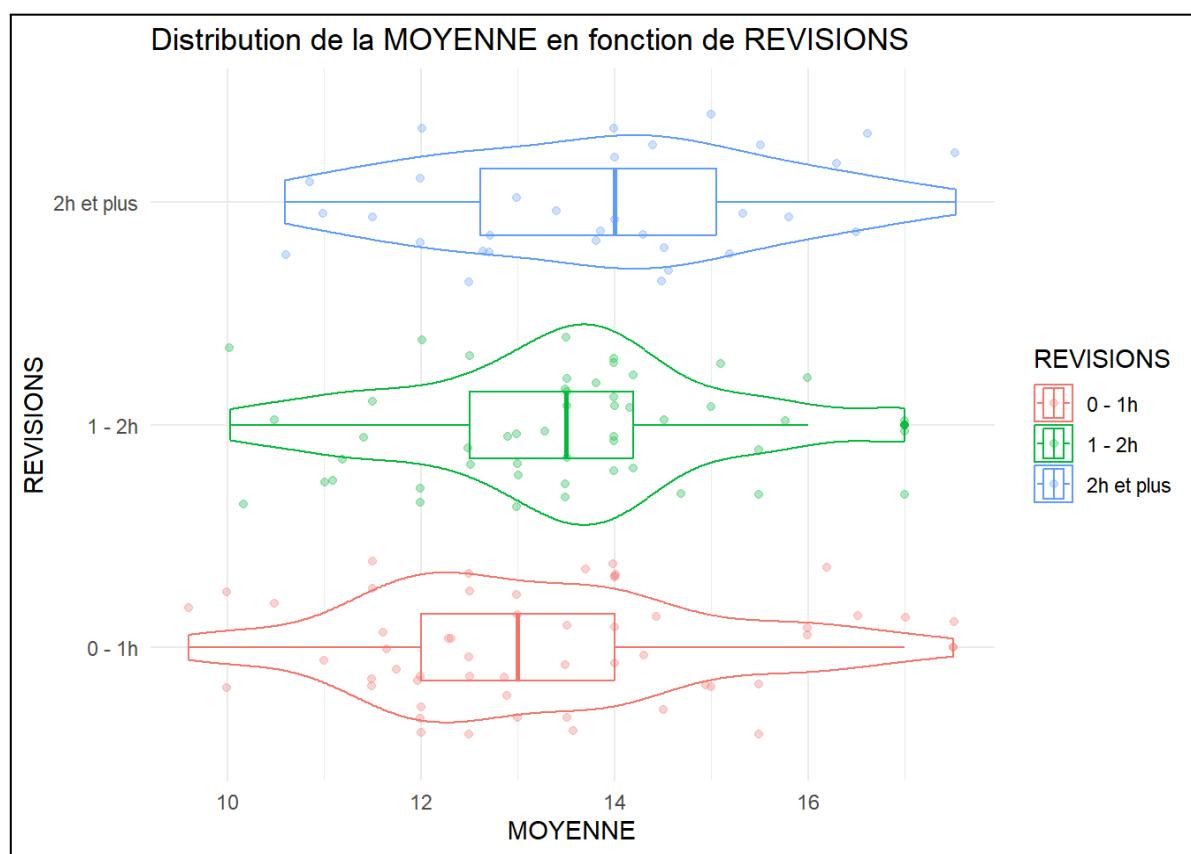


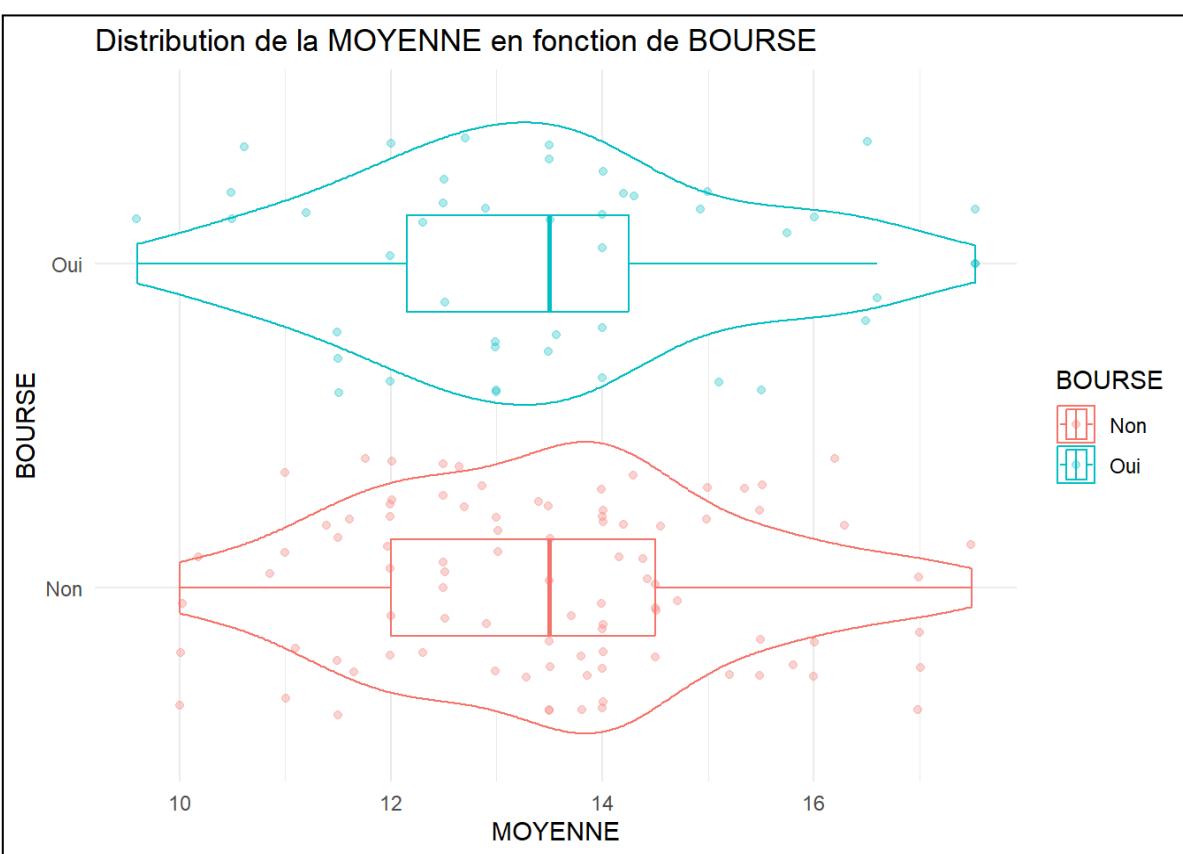
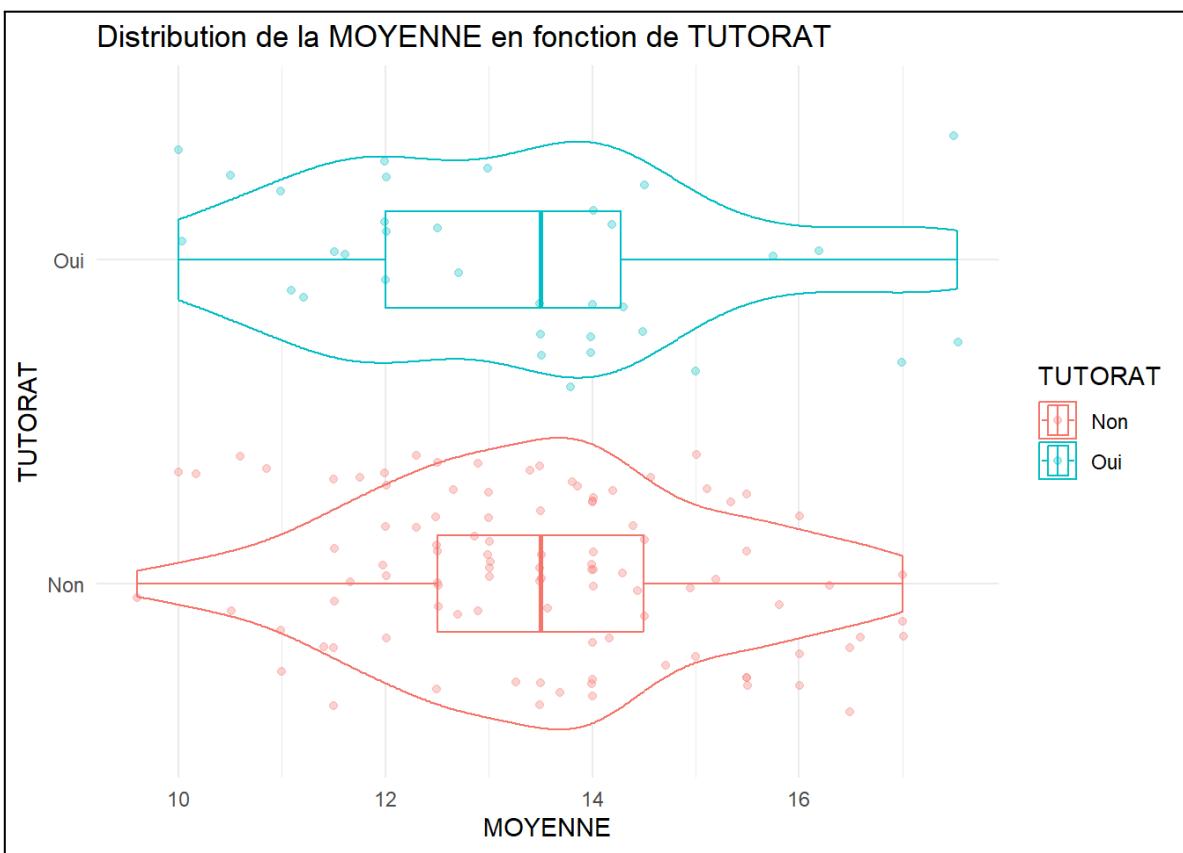
Annexe 30 : Test du khi-deux d'indépendance entre deux variables qualitatives.

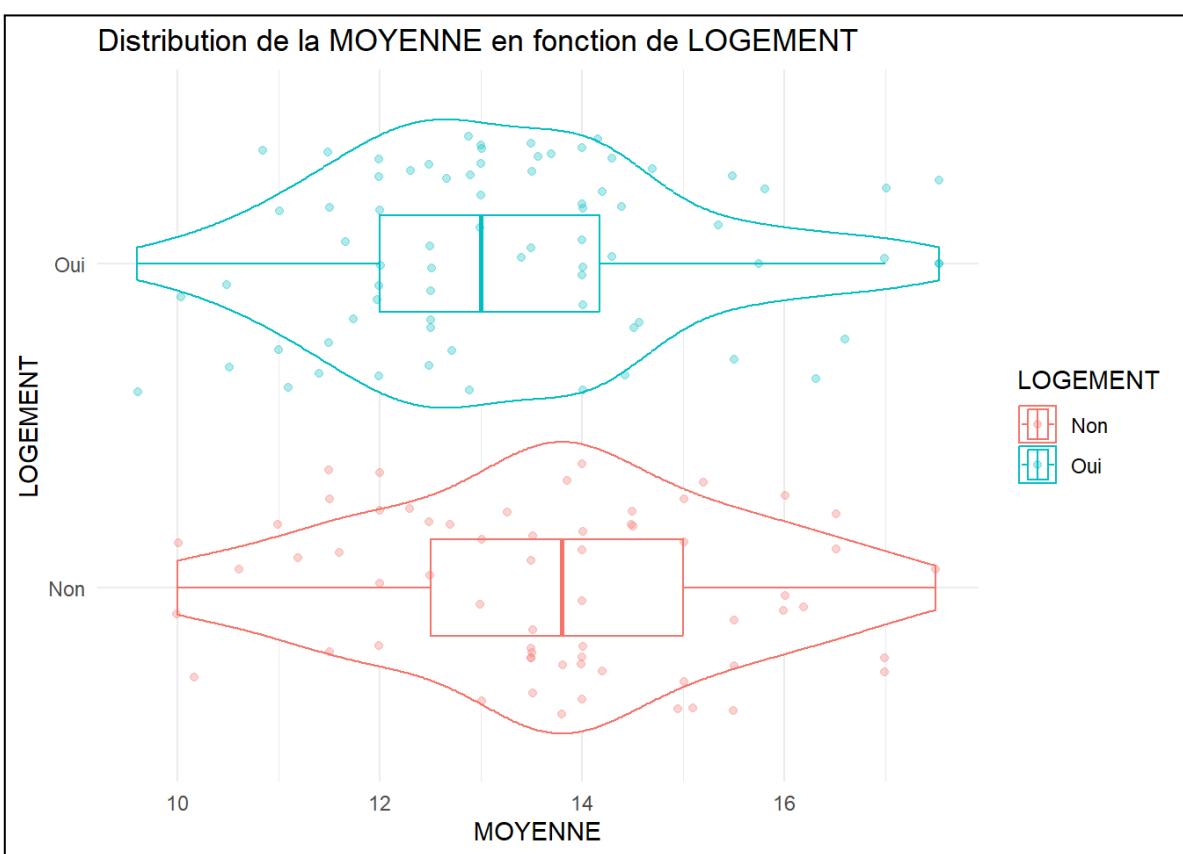
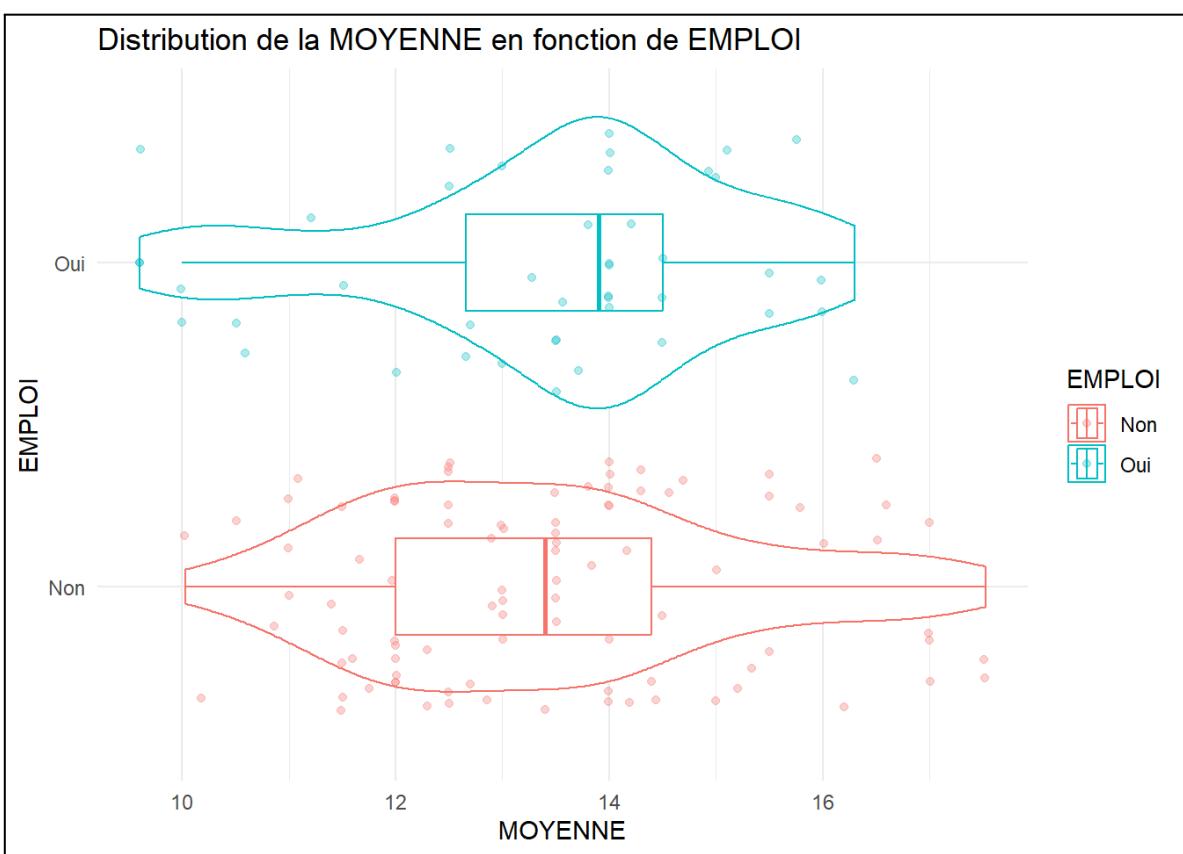
```
data: PARTICULARS_TUTORAT  
X-squared = 8.5656, df = 1, p-value = 0.003426  
  
data: BOURSE_EMPLOI  
X-squared = 4.1776, df = 1, p-value = 0.04096  
  
data: LOGEMENT_CAF  
X-squared = 75.302, df = 1, p-value < 2.2e-16  
  
data: STATUT_ARGENT  
X-squared = 9.6782, df = 2, p-value = 0.007914
```

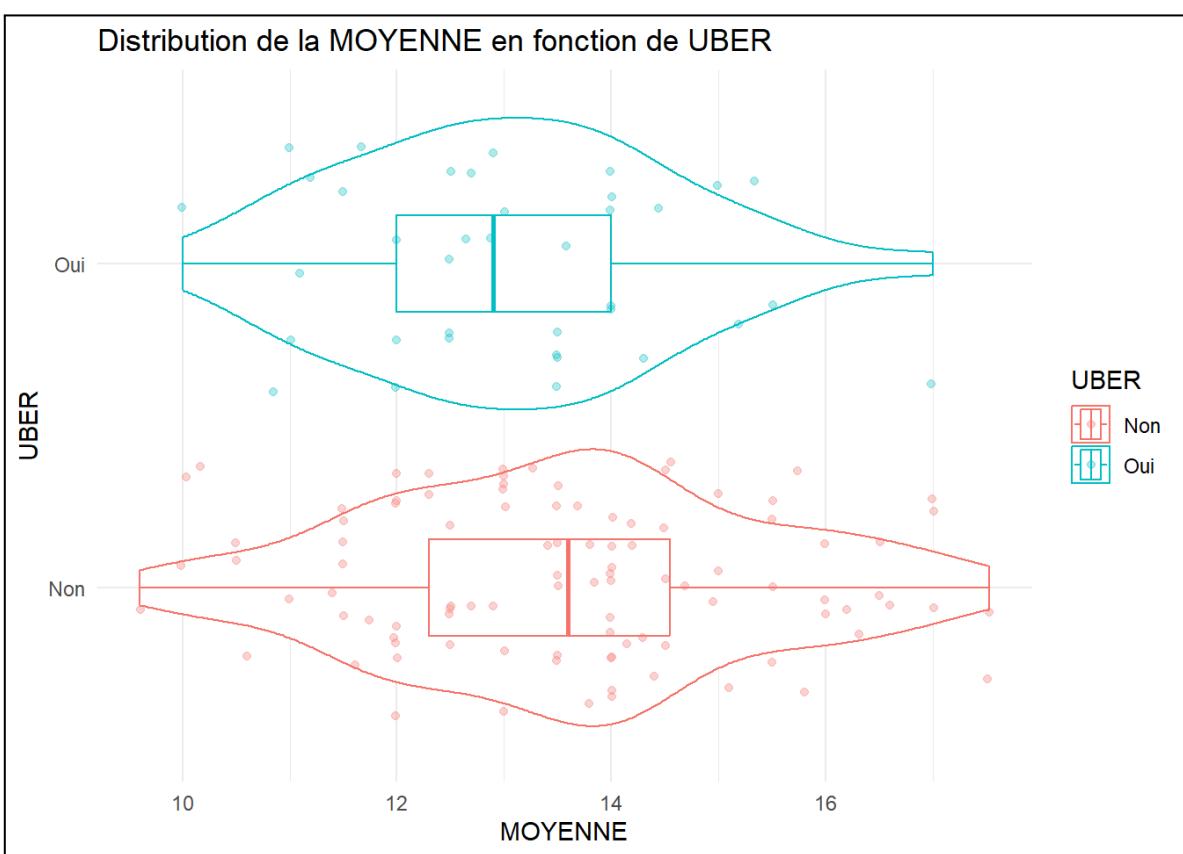
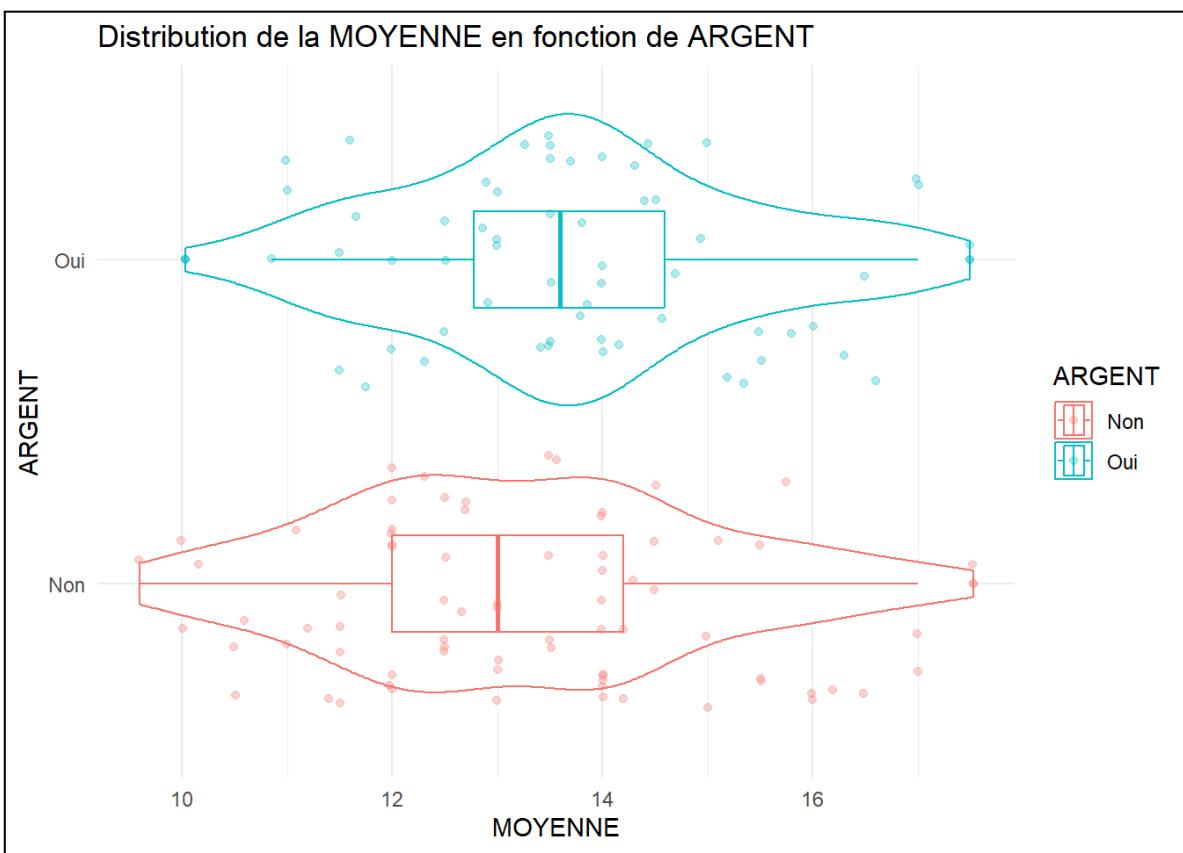
Annexe 31 : Distribution de MOYENNE en fonction des modalités de chaque variable qualitative.

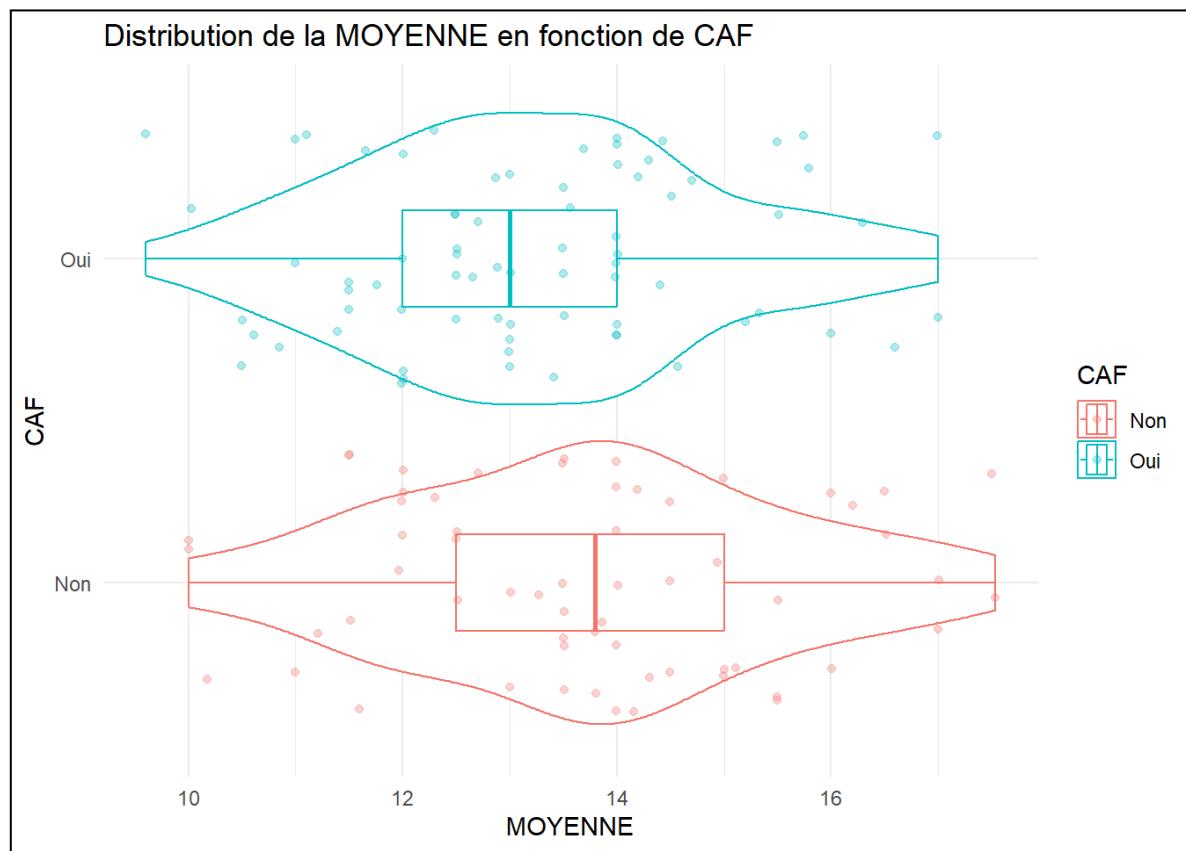
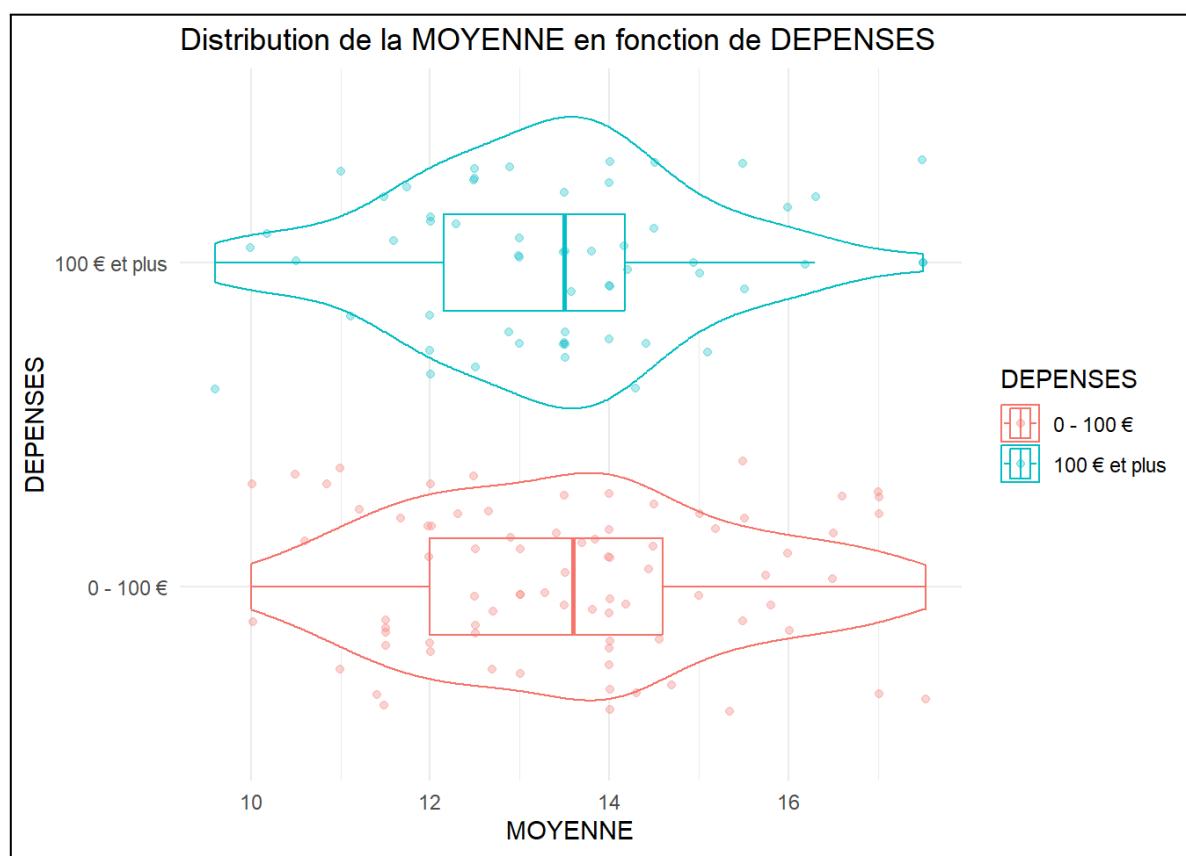


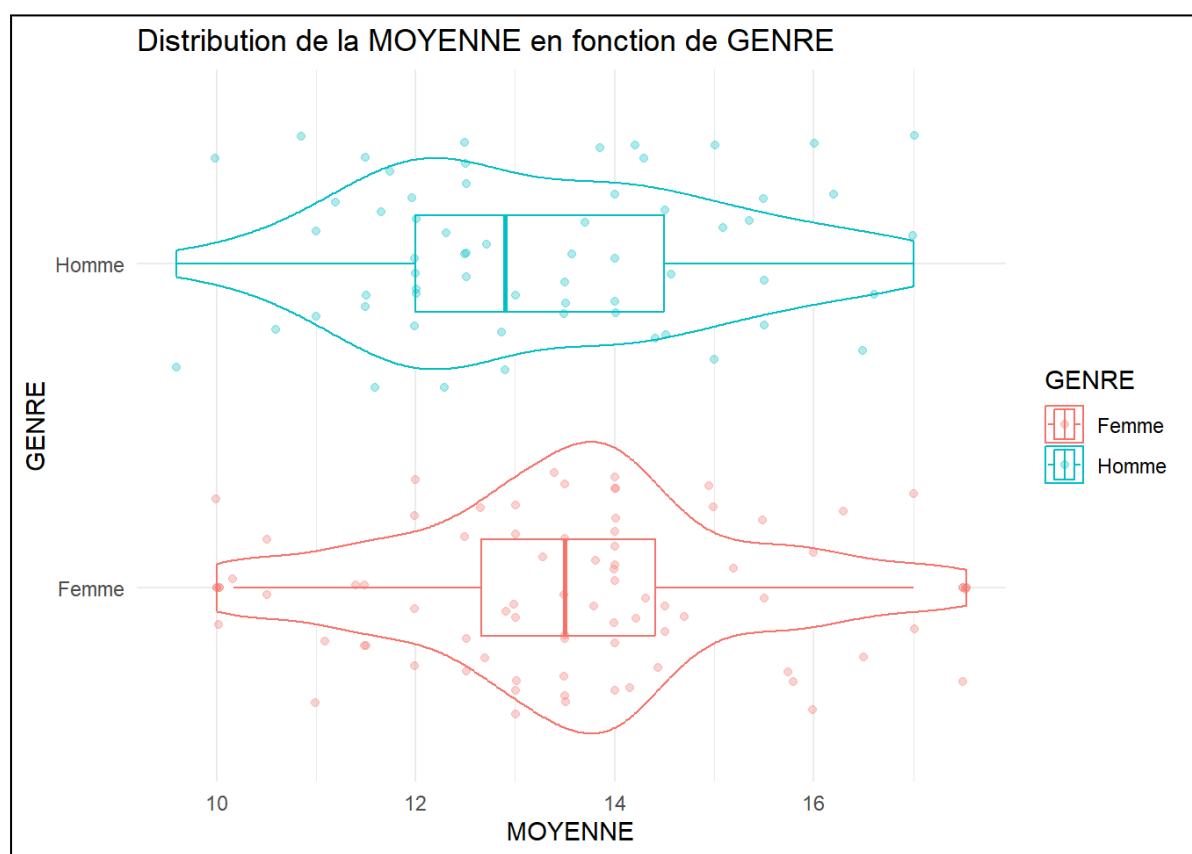
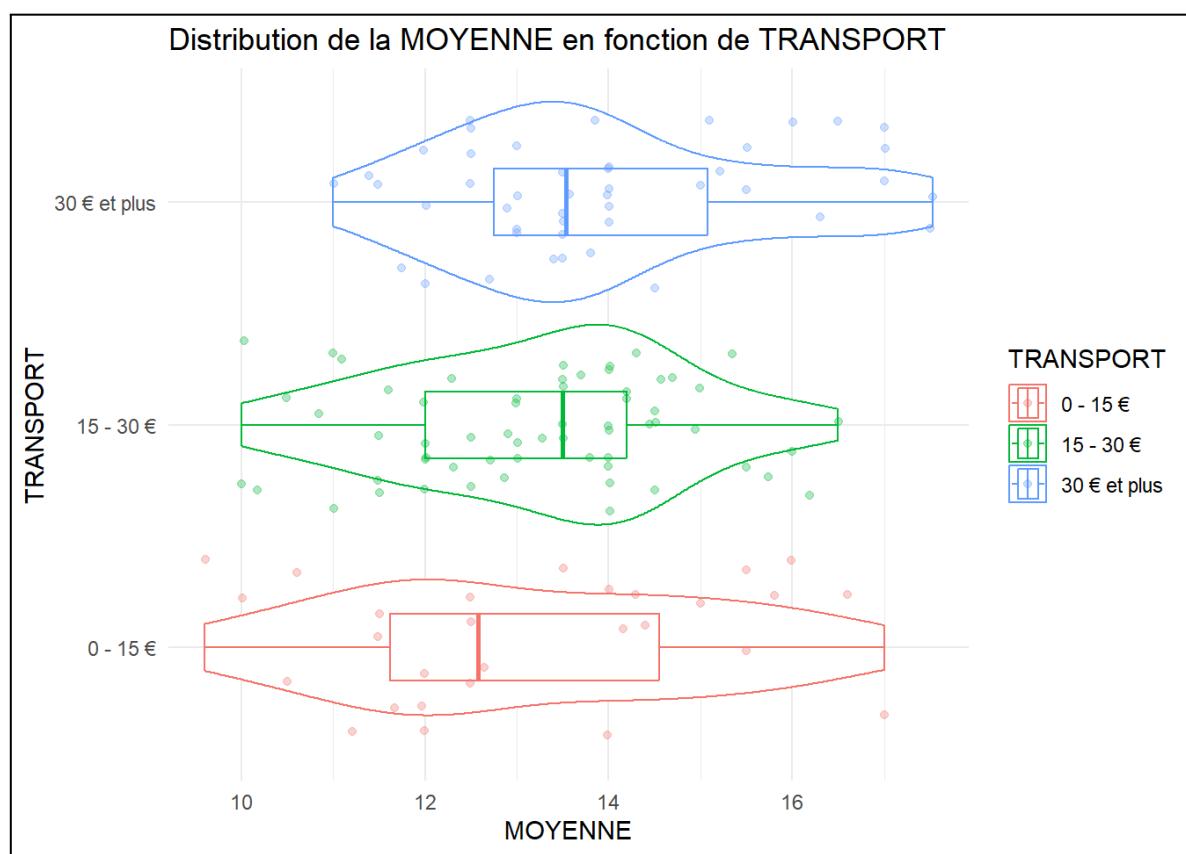


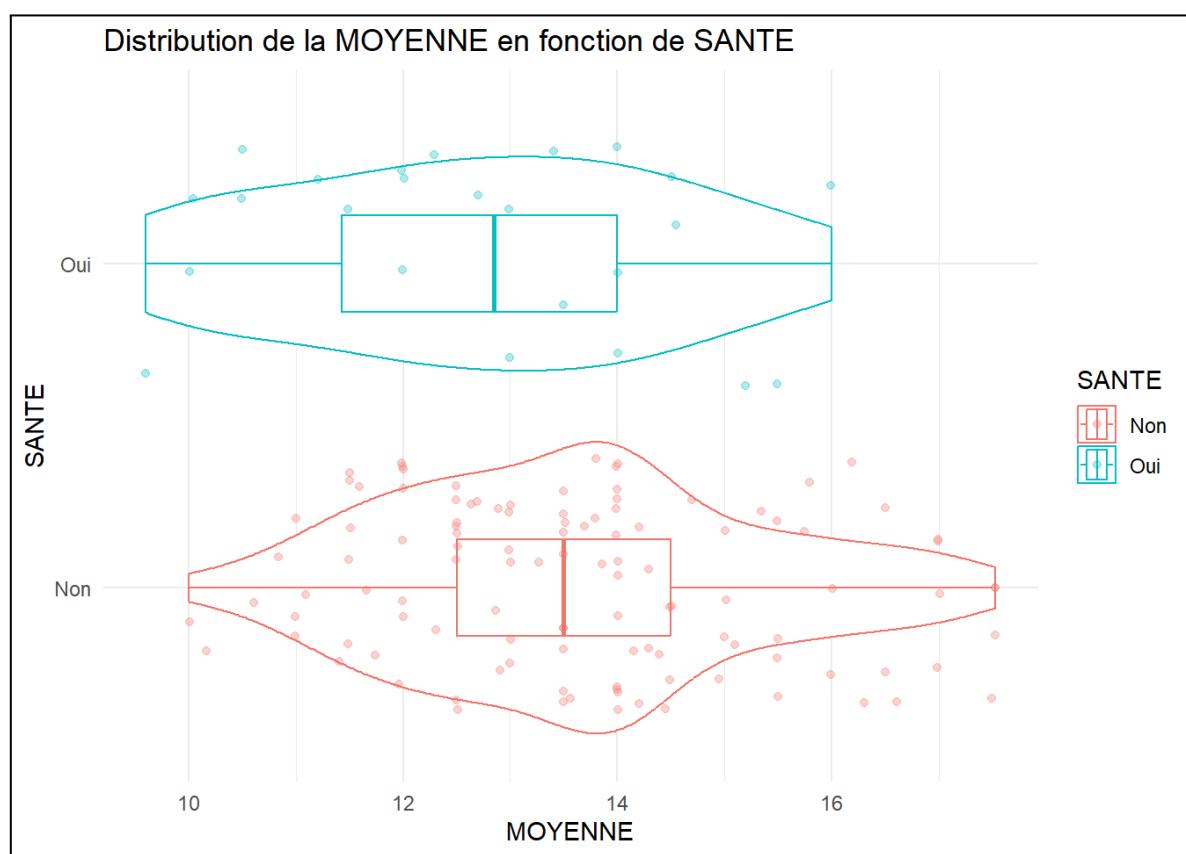
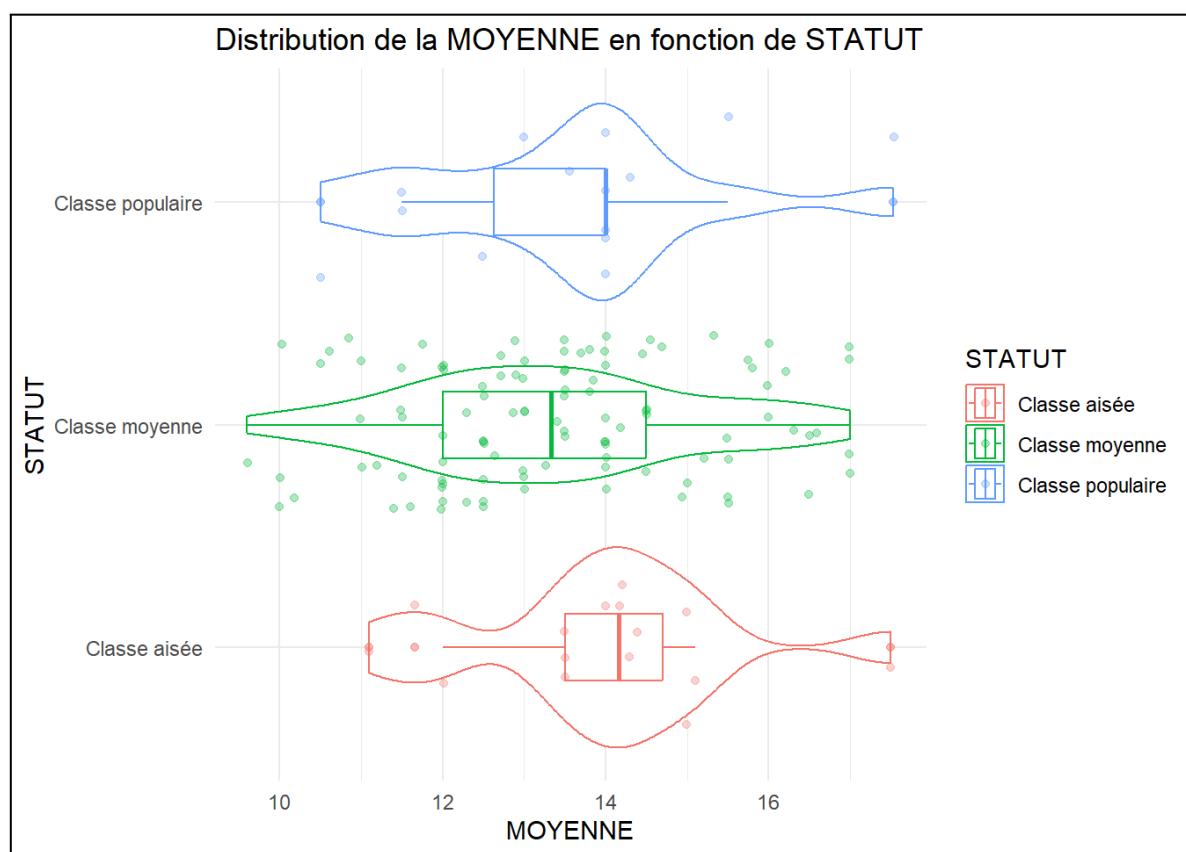


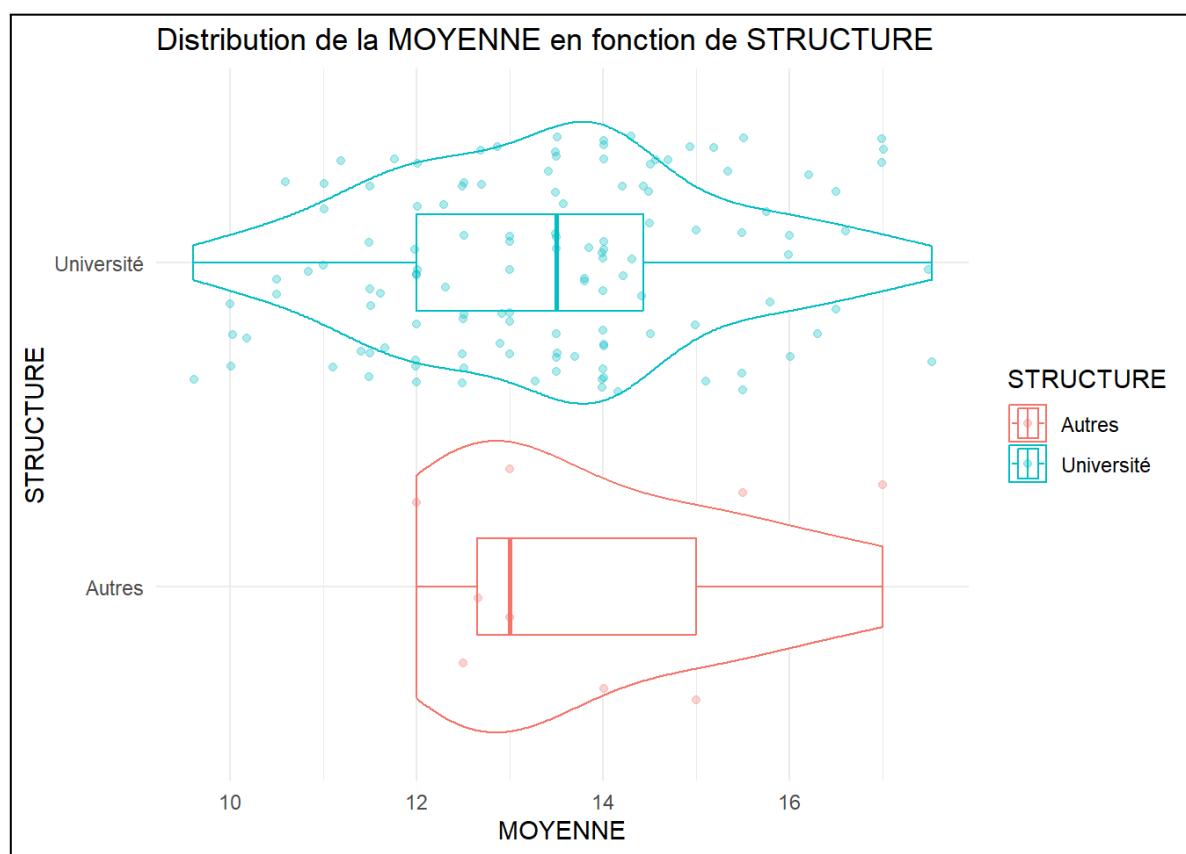












Annexe 32 : fonction step, méthode ascendante

```
Step: AIC=136.19  
MOYENNE ~ ASSIDUITE + RESTAURANT + STRESS + SANTE + TRANSPORT
```

	Df	Sum of Sq	RSS	AIC
<none>		333.76	136.19	
+ ARGENT	1	4.7217	329.04	136.27
+ PARTICULIERS	1	3.1885	330.57	136.90
+ GENRE	1	2.4111	331.35	137.22
+ EMPLOI	1	1.7176	332.04	137.50
+ DEPENSES	1	1.1451	332.61	137.73
+ TUTORAT	1	0.9180	332.84	137.82
+ AGE	1	0.3876	333.37	138.04
+ STRUCTURE	1	0.1092	333.65	138.15
+ BOURSE	1	0.0421	333.72	138.18
+ FORMATION	1	0.0296	333.73	138.18
+ REVISIONS	2	4.7405	329.02	138.26

Call:

```
lm(formula = MOYENNE ~ ASSIDUITE + RESTAURANT + STRESS + SANTE +  
TRANSPORT, data = Budget3)
```

Coefficients:

(Intercept)	ASSIDUITE	RESTAURANT
13.13260	0.15724	-0.26005
STRESS	SANTEOui	TRANSPORT15 - 30 €
-0.14693	-0.77632	-0.08956
TRANSPORT30 € et plus		
0.63734		

Annexe 33 : fonction step, méthode descendante

Step: AIC=136.04

MOYENNE ~ ARGENT + RESTAURANT + TRANSPORT + GENRE + STRESS + SANTE

	Df	Sum of Sq	RSS	AIC
<none>		328.47	136.04	
- GENRE	1	6.0237	334.49	136.49
- ARGENT	1	6.1063	334.58	136.53
- TRANSPORT	2	11.3161	339.79	136.61
- SANTE	1	13.5003	341.97	139.47
- STRESS	1	19.0549	347.52	141.65
- RESTAURANT	1	23.4979	351.97	143.37

Call:

```
lm(formula = MOYENNE ~ ARGENT + RESTAURANT + TRANSPORT + GENRE +  
    STRESS + SANTE, data = Budget3)
```

Coefficients:

(Intercept)	ARGENTOui	RESTAURANT
14.67747	0.45055	-0.32039
TRANSPORT15 - 30 €	TRANSPORT30 € et plus	GENREHomme
-0.07023	0.58878	-0.43077
STRESS	SANTEOui	
-0.16269	-0.84895	

Annexe 34 : fonction step, méthode dans les deux sens

Step: AIC=136.19

MOYENNE ~ ASSIDUITE + RESTAURANT + STRESS + SANTE + TRANSPORT

	Df	Sum of Sq	RSS	AIC
<none>		333.76	136.19	
+ ARGENT	1	4.7217	329.04	136.27
- ASSIDUITE	1	5.6749	339.43	136.47
+ PARTICULIERS	1	3.1885	330.57	136.90
+ GENRE	1	2.4111	331.35	137.22
+ EMPLOI	1	1.7176	332.04	137.50
+ DEPENSES	1	1.1451	332.61	137.73
- TRANSPORT	2	14.1191	347.88	137.79
+ TUTORAT	1	0.9180	332.84	137.82
+ AGE	1	0.3876	333.37	138.04
+ STRUCTURE	1	0.1092	333.65	138.15
+ BOURSE	1	0.0421	333.72	138.18
+ FORMATION	1	0.0296	333.73	138.18
+ REVISIONS	2	4.7405	329.02	138.26
- SANTE	1	11.2250	344.98	138.66
- RESTAURANT	1	14.7882	348.55	140.05
- STRESS	1	15.1666	348.93	140.19

Call:

```
lm(formula = MOYENNE ~ ASSIDUITE + RESTAURANT + STRESS + SANTE +
TRANSPORT, data = Budget3)
```

Coefficients:

(Intercept)	ASSIDUITE	RESTAURANT
13.13260	0.15724	-0.26005
STRESS	SANTEOui	TRANSPORT15 - 30 €
-0.14693	-0.77632	-0.08956
TRANSPORT30 € et plus		
0.63734		

Annexe 35 : Estimation de notre modèle par la RLM en appliquant la méthode des MCO.

```
Call:
lm(formula = MOYENNE ~ ARGENT + RESTAURANT + TRANSPORT + GENRE +
    STRESS + SANTE, data = Budget3)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.9542 -1.1593 -0.0678  1.0925  3.6724 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) 14.67747   0.47591 30.841 < 2e-16 ***
ARGENTOui    0.45055   0.29322  1.537  0.12689  
RESTAURANT   -0.32039   0.10629 -3.014  0.00311 ** 
TRANSPORT15 - 30 € -0.07023   0.37167 -0.189  0.85042  
TRANSPORT30 € et plus 0.58878   0.39852  1.477  0.14204  
GENREHomme   -0.43077   0.28227 -1.526  0.12947  
STRESS       -0.16269   0.05994 -2.714  0.00757 ** 
SANTEOui     -0.84895   0.37158 -2.285  0.02399 *  
---
Signif. codes:  0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.608 on 127 degrees of freedom
Multiple R-squared:  0.203, Adjusted R-squared:  0.1591 
F-statistic: 4.621 on 7 and 127 DF,  p-value: 0.0001237
```

Annexe 36 : Test de Shapiro-Wilk pour vérifier la normalité des résidus

```
Shapiro-Wilk normality test
```

```
data: residus  
W = 0.99103, p-value = 0.5421
```

Annexe 37 : Test de Kolmogorov-Smirnov pour vérifier la normalité des résidus

```
Asymptotic one-sample Kolmogorov-Smirnov test
```

```
data: residus  
D = 0.060712, p-value = 0.7022  
alternative hypothesis: two-sided
```

Annexe 38 : Test de Ramsey afin de vérifier la forme fonctionnelle du modèle

```
RESET test
```

```
data: regf  
RESET = 1.0884, df1 = 2, df2 = 125, p-value = 0.3399
```

Annexe 39 : Test du VIF pour vérifier l'hypothèse d'absence de multicolinéarité

	GVIF	Df	GVIF^(1/(2*Df))
ARGENT	1.108109	1	1.052668
RESTAURANT	1.066001	1	1.032473
TRANSPORT	1.124012	2	1.029657
GENRE	1.030033	1	1.014905
STRESS	1.101970	1	1.049748
SANTE	1.053461	1	1.026383

Annexe 40 : Test de Breusch-Pagan afin de vérifier l'homoscédasticité des résidus

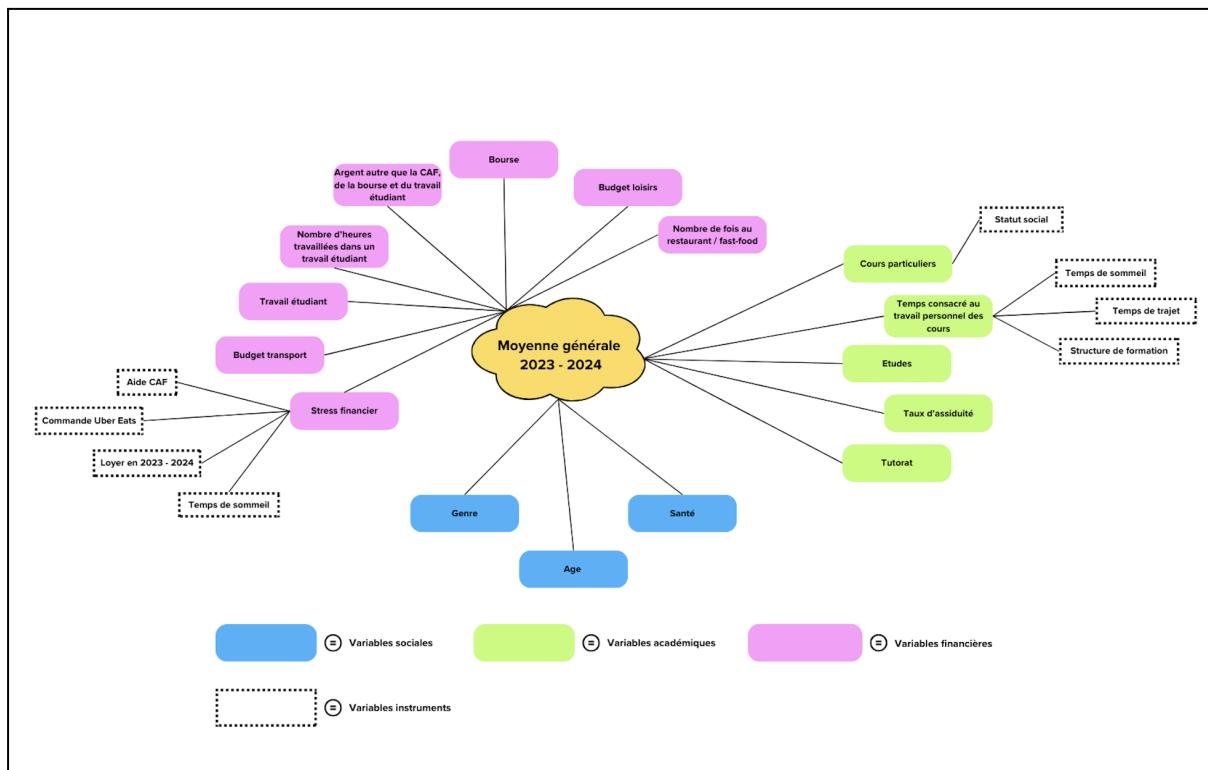
```
studentized Breusch-Pagan test

data: regf
BP = 3.6683, df = 7, p-value = 0.8171
```

Annexe 41 : Estimation de notre modèle par la méthode des DMC et tests d'endogénéité

```
Call:  
ivreg(formula = MOYENNE ~ ARGENT + RESTAURANT + TRANSPORT + GENRE +  
STRESS + SANTE | ARGENT + RESTAURANT + TRANSPORT + GENRE +  
SANTE + CAF + UBER + LOGEMENT + SOMMEIL, data = Budget3)  
  
Residuals:  
    Min         1Q     Median        3Q       Max  
-4.680474 -1.193129  0.005118  1.113470  4.505118  
  
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)  
(Intercept) 15.66727  0.83301 18.808 < 2e-16 ***  
ARGENTOui    0.25105  0.34053  0.737  0.46234  
RESTAURANT   -0.36706  0.11783 -3.115  0.00227 **  
TRANSPORT15 - 30 € -0.03144  0.39823 -0.079  0.93721  
TRANSPORT30 € et plus 0.77345  0.44351  1.744  0.08359 .  
GENREHomme   -0.53462  0.30963 -1.727  0.08667 .  
STRESS        -0.41831  0.18198 -2.299  0.02316 *  
SANTEOui     -0.58740  0.43384 -1.354  0.17816  
  
Diagnostic tests:  
                 df1 df2 statistic p-value  
Weak instruments 4 124      4.389 0.00236 **  
Wu-Hausman      1 126      2.607 0.10886  
Sargan          3  NA      0.865 0.83384  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 1.72 on 127 degrees of freedom  
Multiple R-Squared: 0.08888, Adjusted R-squared: 0.03866  
Wald test: 3.877 on 7 and 127 DF, p-value: 0.0007381
```

Annexe 42: Schéma présentant le modèle de la moyenne générale de 2023 - 2024



XI - Table des matières

I - Introduction	2
II - Analyse détaillée de la gestion budgétaire dans la réussite académique et présentation des variables	4
A. Pertinence de la variable à expliquer : la moyenne générale lors de l'année 2023-2024	4
B. Pertinence des variables explicatives	5
1. Variables sociales	5
1.1. Genre	5
1.2. Age	7
1.3. Santé	7
2. Variables académiques	8
2.1. Tutorat	8
2.2. Taux d'assiduité	9
2.3. Filière d'étude	9
2.4. Temps dédié au travail personnel pour les cours	10
2.5. Cours particulier	12
3. Variables financières	13
3.1. Stress Financier	13
3.2. Budget transport	14
3.3. Travail étudiant	15
3.4. Nombre d'heures travaillées	16
3.5. Budget loisirs	17
3.6. Argent autre que la bourse, la CAF et le travail étudiant	18
3.7. Bourse	19
3.8. Nombre de restaurants par mois	19
C. Récapitulatif du modèle	20
III - Méthode	22
A. Méthode de recueil de données	22
B. Méthode économétrique utilisée	26
IV - Analyse de la base de donnée	27
A. Représentativité de l'échantillon	27
B. Statistiques descriptives univariées	28
1. Variables quantitatives	29
2. Atypicité des variables quantitatives	31
3. Variables qualitatives	35
C. Statistiques descriptives bivariées	38
1. Analyse des variables Quantitatives - Quantitatives	38
2. Analyse des variables Qualitatives - Qualitatives	40
3. Analyse des variables Qualitatives - Quantitatives	41
V - Estimations économétriques	43

A. Estimations préliminaires	43
1. Première estimation du modèle par la méthode des STEP	43
2. Choix du meilleur modèle	44
B. Tests Statistiques	45
1. Normalité des résidus	46
2. Forme fonctionnelle	47
3. Absence de multicolinéarité	48
4. Homoscédasticité des erreurs	49
5. Distance de Cook	50
C. Vérification de l'endogénéité	51
1. Choix des instruments de notre modèle	51
2. Tests Statistiques	52
D. Interprétation des résultats	54
VI - Prévisions	56
VII - Discussion	58
A. Comparaison de nos résultats avec la littérature	58
B. Limites	59
VIII - Conclusion	61
IX - Bibliographie	63
X - Annexes	70
XI - Table des matières	107