

15/12/2023

Master 1 Data Analyst, 2023-2024

GOUPIL Victor

Le milieu social influence-t-il la consommation des jeux d'argents ?



Sous la direction de Yannick L'Horty

CAMPUS MARNE-LA-VALLEE

Université Gustave Eiffel
Cité Descartes - Champs-Sur-Marne
Bâtiment Bois de l'étang
5 rue Galilée
77454 Marne-La-Vallée cedex 2
01 60 95 70 48

www.iae-eiffel.fr



Table des matières

Introduction	2
Présentation de l'enquête et de l'échantillon	3
a. Administration du questionnaire et création de la base de données.....	3
b. Statistiques descriptives.....	4
Les tests d'indépendance (Khi2).....	9
a. Test simple et détaillé	9
b. Test d'indépendance généralisé	11
c. Test d'indépendance sur des sous-échantillons	12
Régressions linéaires	14
a. La régression simple.....	14
b. La régression multiple.....	15
c. Les régressions sur des sous échantillons	17
d. Régression multiple avec une autre variable à expliquer.....	18
Conclusion.....	19
Annexe	21
Questionnaire	21
Bibliographie	25
Régressions : Résultats complémentaires	26

Introduction

Selon l'autorité nationale des jeux (ANJ), le marché des jeux en France représentait un chiffre d'affaires de 12,9 milliards d'euros en 2022. Toujours selon cette autorité, près d'un français sur deux est consommateur de jeux d'argents. Ces derniers, en tant que phénomène socioculturel, occupent une place significative dans nos sociétés contemporaines. Au-delà de l'aspect récréatif, ils soulèvent des questions cruciales sur les influences environnementales et sociales qui guident les comportements de consommation, ce qui suscite un intérêt croissant du point de vue des recherches académiques.

Alors que le jeu peut être perçu comme une activité de loisir accessible à tous, il est devenu essentiel d'explorer les liens intrinsèques entre le milieu social et la consommation de jeux d'argent. Plusieurs études, comme notamment celles de JM Costes dirigées pour l'observatoire des jeux (ODJ) en 2012, 2017 et 2019, ou bien celle de T. Amadiou en 2015 convergent vers la conclusion que certaines caractéristiques sociodémographiques, telles que le sexe, l'âge, le niveau de diplôme et le niveau de revenu, autrement dit le milieu social, exercent une influence significative sur le niveau de jeu problématique. Bien qu'elles montrent de petites différences entre elles, ces recherches suggèrent que toutes choses égales par ailleurs, être un homme, jeune et disposant d'un niveau d'éducation modeste accroît sensiblement la probabilité d'être un joueur problématique. Cependant, le niveau de revenu n'interfère pas significativement sur le comportement de jeu.

En 2017, l'ODJ s'est concentré uniquement sur les joueurs en ligne et a trouvé un profil un peu différent des joueurs traditionnels. En effet, le profil type du joueur en ligne montrait un homme plus jeune et plus diplômé que l'analyse sur les jeux traditionnels. Dans les deux cas, les hommes semblent accorder une importance bien plus grande aux jeux d'argent que les femmes.

Ces résultats vont aussi dans le sens de notre première intuition, avant même d'avoir lu les études. En effet, notre hypothèse est que les individus qui évoluent au sein d'un milieu social plus modeste consomment plus de jeux d'argent, en budget ou en fréquence de jeu, que ceux qui évoluent dans un milieu social plus élevé. Le milieu social se définit principalement par le niveau d'étude, le revenu et la catégorie socio-professionnelle. La consommation de jeux d'argents quant à elle sera analysée principalement par la fréquence de jeu par mois, puis le budget de jeu par mois.

Ainsi, cette étude vise à répondre à des questions précises que l'on se pose comme : Qui consomme les jeux d'argent ? Est-ce que ce sont les personnes à revenu modeste qui cherchent à en avoir davantage au risque d'en perdre ? Ou bien à l'inverse des individus à revenu plus élevé qui cherchent à se divertir ? De manière plus générale, nous tenterons de répondre à la problématique suivante : Le milieu social influence-t-il la consommation des jeux d'argents ? Pour ce faire, nous examinerons de plus près les caractéristiques sociodémographiques spécifiques qui émergent comme des facteurs influents, puis nous chercherons à comprendre comment ces éléments contribuent à la création d'un profil de joueur problématique. En analysant ces relations, nous espérons éclairer les déterminants individuels du jeu problématique.

Pour faire réaliser l'étude, nous construirons dans un premier temps une base de données grâce à l'administration d'un questionnaire en ligne. Ensuite, nous la décortiquerons à l'aide de graphiques et de statistiques descriptives avant de réaliser des tests d'indépendances pour avoir une première idée

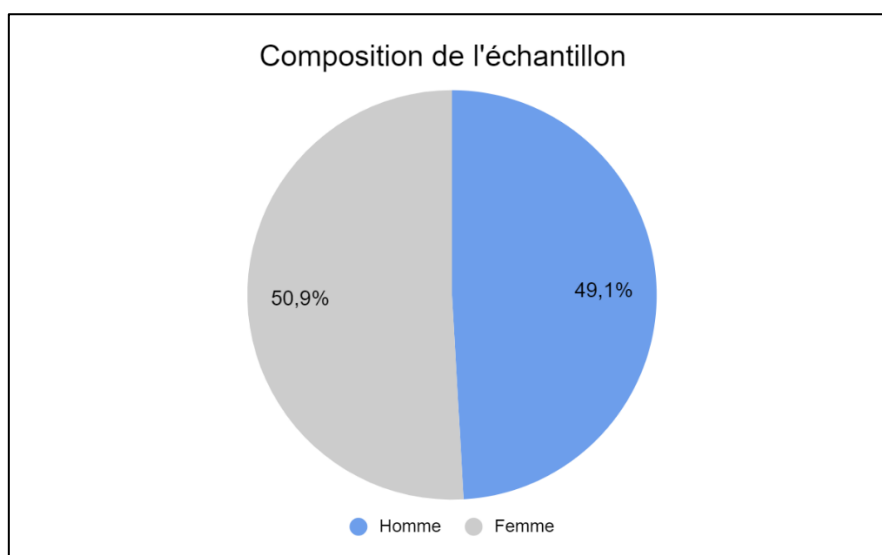
sur la significativité de nos variables. Enfin, nous chercherons à obtenir des résultats grâce aux régressions linéaires qui répondront à toutes nos questions.

Présentation de l'enquête et de l'échantillon

a. Administration du questionnaire et création de la base de données

Dans un premier temps, nous avons dû constituer une base de données afin de réaliser notre étude et répondre à notre problématique sur l'influence du milieu social et la consommation des jeux d'argents. Pour ce faire, nous avons construit un questionnaire à l'aide de Google Form avec des questions qui nous renseignent sur le milieu social de l'individu et ses rapports avec les jeux d'argents. Nous avons aussi quelques questions plus générales comme le genre, l'âge, le type de logement, le paiement du loyer ou encore le numéro de département qui a servi à créer une variable « région » que nous avons ensuite séparées en 5 modalités qui coupe la France de la sorte : « Nord, Sud, Est, Ouest, Centre ». Ces variables nous serviront de variables de contrôle pour notre régression. Ensuite, nous avons partagé ce questionnaire à notre entourage (camarades de classe, amis, famille proche ou éloignée) afin de récolter un maximum de réponse. Au bout de 14 jours, nous avons atteints notre objectif de 100 réponses et nous avons pu commencer le nettoyage de la base. Pour ce faire, nous nous sommes servis du logiciel R. Nous avons donc dans un premier temps transformé toutes nos variables binaires « Oui/Non » en variables binaires numériques prenant les valeurs 0 et 1. Pour finir, nous avons supprimé les valeurs aberrantes puis nous avons créé des variables catégorielles à partir de nos variables numériques pour réaliser les tests du Khi2. Ce questionnaire s'adresse à tout le monde et ne cible personne en particulier. Au total, notre base de données contient 110 observations avec un nombre d'hommes et de femmes quasiment égaux (50.9% vs 49.1%).

Figure 1 : Composition de l'échantillon



A l'aide des réponses que nous avons obtenu, nous avons pu former clairement des variables qualitatives et quantitatives, pour les X comme pour les Y. Nous avons notamment défini 3 variables X qui représentent le milieu social. Il s'agit du niveau de diplôme, du revenu et de la catégorie socio-professionnelle. De plus, nous avons un nombre conséquent de variable Y qui nous permettront de jauger la consommation des jeux d'argents des individus. Parmi celles-ci, deux variables semblent plus pertinentes que les autres. Il s'agit de la fréquence de jeu par mois, et du budget de jeu par mois. Les personnes jouant moins d'une fois par mois prendront la valeur « 0 », tandis que ceux qui ne jouent jamais recevront la valeur « -1 » afin de pouvoir les inclure dans nos régressions. Pour finir, nous avons aussi créé nos variables de contrôle dites « variables Z » à partir des questions portant sur les renseignements généraux.

Après avoir expliqué comment notre base de données a été constituée, nous allons passer l'analyse descriptive de cette dernière afin d'avoir une vision plus éclairée sur les différents types d'observations dont nous disposons.

b. Statistiques descriptives

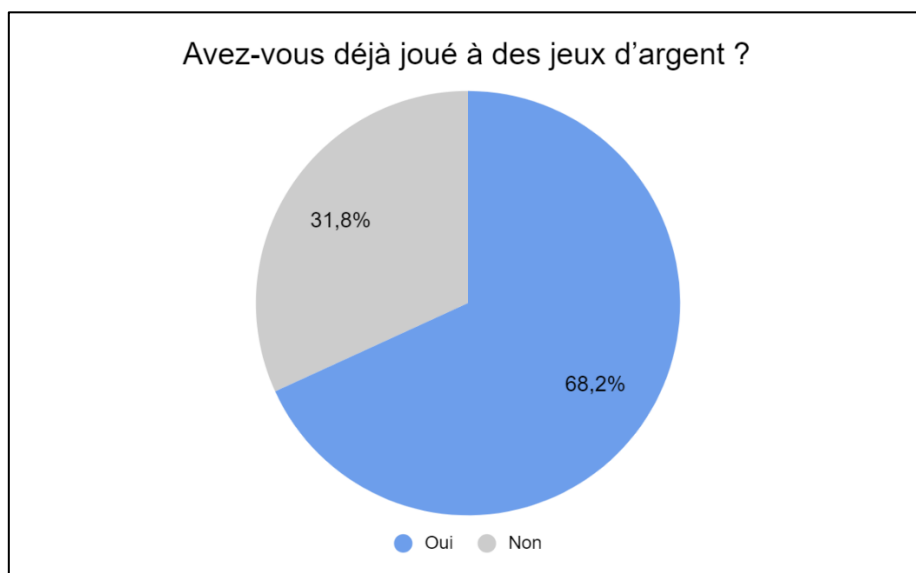
Figure 2 : Tableau des statistiques descriptives pour les variables quantitatives

	Min	1 ^{er} Qu	Médiane	Moyenne	3 ^{ème} Qu	Max
Age	15	21	21	28	36	72
Niveau diplôme	0	4	5,5	4,9	6	7
Fréquence/mois	0	0	0	4,7	2	100
Mise	0	3,5	6	9,6	10	50
Budget jeu	0	0	5	44,3	20	750
Plus grosse perte	2	5	20	38,8	29	600
Plus gros gain	0	20	50,5	317,7	150	5000
% du gain réinvesti	0	17,5	50	48,7	100	100

Voici un tableau descriptif avec nos variables numériques intéressantes à analyser. Tout d'abord, nous pouvons remarquer que nous avons une grande dispersion de l'âge avec des observations allant de 15 à 72 ans. On constate aussi une forte représentation des personnes âgées de 21 ans, la médiane et le 1^{er} quartile étant tous les deux égaux à cet âge. Ensuite, nous travaillons sur une base de données qui a une moyenne de niveau de diplôme de niveau 5, soit l'équivalent d'un bac +2. Les six variables suivantes concernent les jeux d'argent et prennent en compte seulement les personnes ayant répondu avoir déjà jouer à des jeux d'argent. A première vue, une déduction semble évidente ; la base de données contient des consommateurs de jeux d'argents très différents voire opposés, comme le montre la grande dispersion de ces données. En effet, certains sont de « petits joueurs » et jouent occasionnellement dans l'année pendant que d'autres jouent 100 fois par mois. Le budget par mois le montre aussi, avec des montants allant de 0 à 750€ par mois. Cependant, on constate qu'il y a une majorité de « petits joueurs » en se fiant aux différentes médiane. Il va donc être intéressant de voir

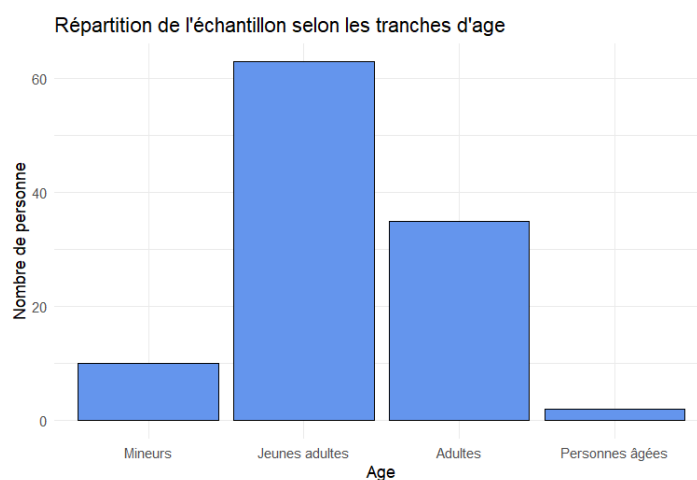
quel est le profil de ces différents joueurs, et quelles variables, si il y en a, semblent être corrélées avec le fait d’être un petit ou un gros consommateur de jeux d’argent.

Figure 3 : Part de personnes ayant déjà joué aux jeux d’argent.



Parmis les 110 observations, 68.2% ont déjà joué aux jeux d’argent. C’est donc sur ce groupe que l’on concentrera notre analyse, tout en essayant de comprendre pourquoi les 31.8% restants ne jouent pas.

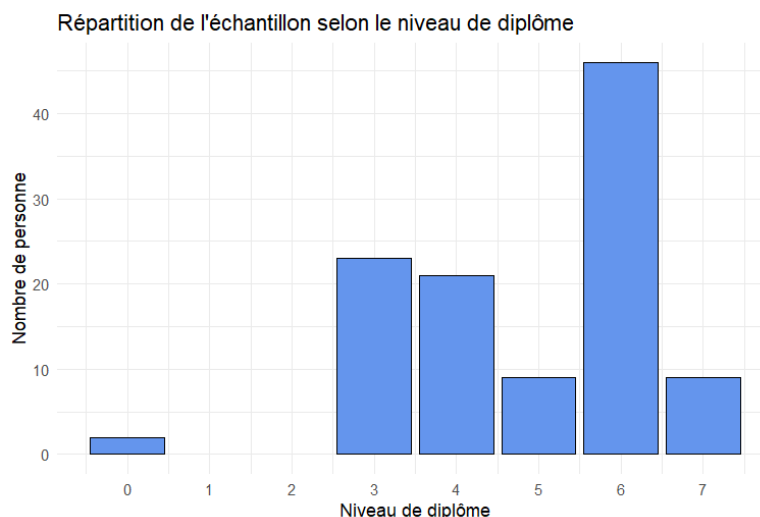
Figure 4 : Répartition de l’échantillon selon les tranches d’âge



Ce second graphique montre la répartition de l’échantillon selon les tranches d’âges. Sans surprise, nous avons une grande majorité de jeunes adultes (18-30 ans) qui représente plus de 60% de notre échantillon. De plus, l’échantillon compte environ 35% d’adultes (31-64 ans). Il n’est pas dérangeant

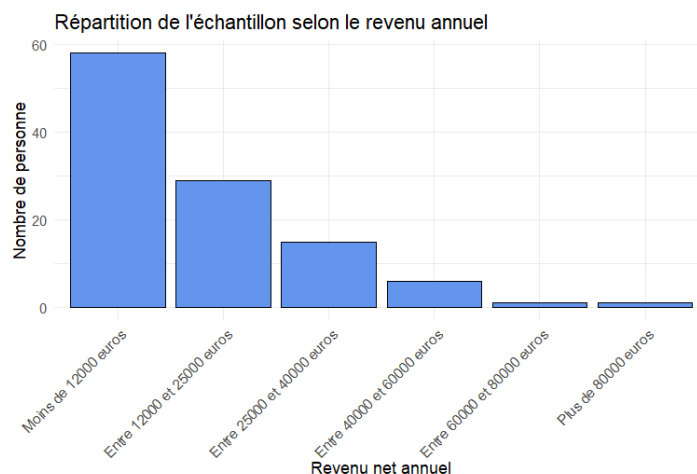
de n'avoir que 10% de mineurs dans cette étude puisque ces derniers sont interdits de jeux d'argents. Il aurait cependant été intéressant d'avoir plus de personnes âgées (65 ans et plus) pour notre analyse.

Figure 5 : Répartition selon le dernier diplôme obtenu



Ce troisième graphique nous permet de mieux visualiser comment se répartit le niveau de diplôme au sein de notre échantillon. Ainsi, plus de 50% des répondants possèdent un diplôme de niveau 6, soit l'équivalent d'un bac +3. Nous savons que le milieu social se définit à partir des caractéristiques socio-économiques (profession, niveau de diplôme, revenus). Cette variable est donc importante pour notre analyse et nous explorerons s'il existe un lien avec le niveau d'éducation et la consommation de jeux d'argents. Elle sera cependant recodée en « nivétude2 » de la manière suivante : pour chaque niveau de diplôme supérieur ou égal à 3, $\text{nivétude2} = \text{nivétude} - 2$. Ceci permet de ne pas avoir de « trou » entre les personnes sans diplôme (0) et les personnes avec un diplôme (3 ou plus).

Figure 6 : Répartition selon le revenu



Comme la variable précédente, le revenu net annuel est en partie un indicateur du milieu social d'une personne. Cette variable est donc tout aussi importante pour nous que la précédente. Comme nous pouvons le voir, presque 60% des répondants gagnent moins de 12000€ par an. Cela s'explique par le fait qu'il y a un grand nombre d'étudiant au sein de l'échantillon. En effet, le graphique en secteur ci-dessous révèle que 57,3% des répondants prétendent ne pas avoir terminé leurs études. Il sera donc intéressant de séparer les étudiants de ceux qui n'étudient plus lors des tests d'indépendances de Khi2 sur des sous populations. Enfin, selon les autres catégories de revenu annuel net, nous voyons que la classe populaire et la classe moyenne sont les classes les plus représentées.

Figure 7 : Avez-vous terminé vos études ?

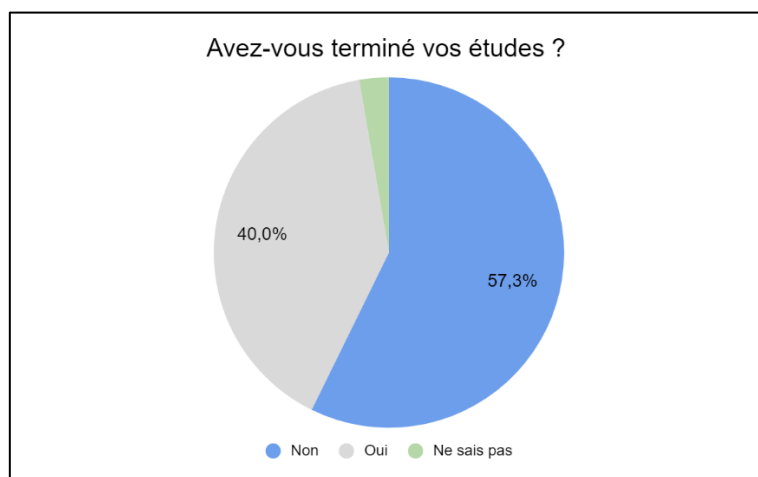
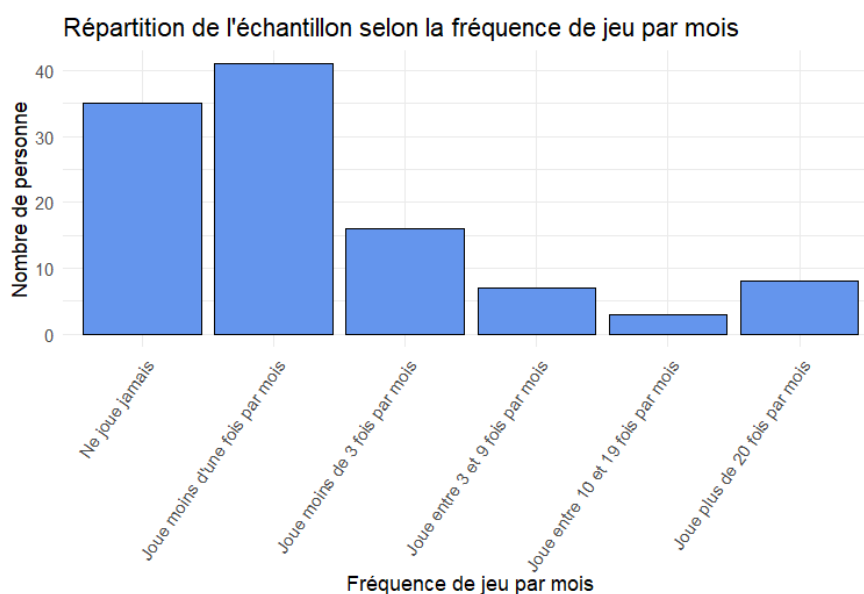


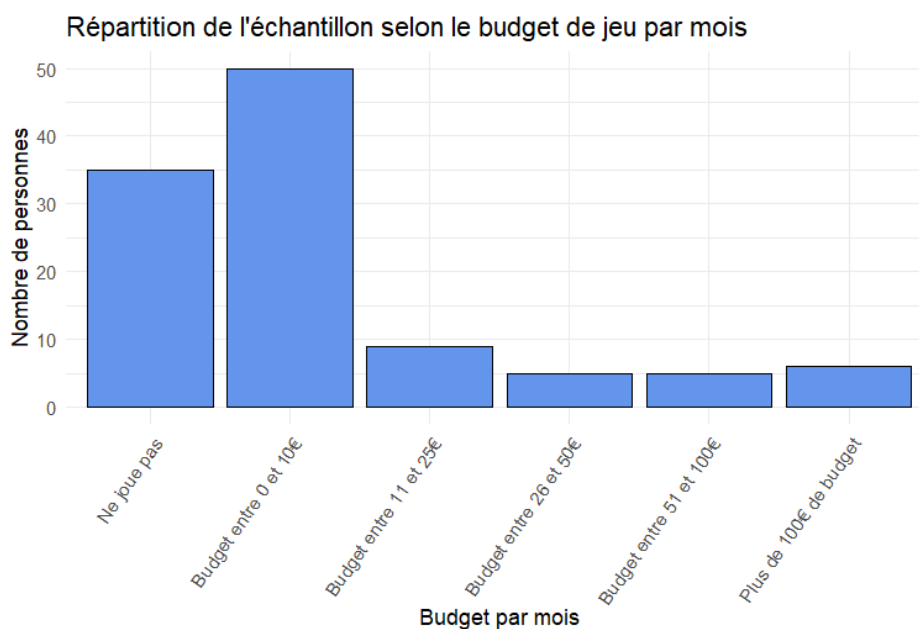
Figure 8 : Répartition selon la fréquence de jeu par mois



Il est temps de s'intéresser aux graphiques qui nous informent sur la partie à expliquer de notre étude, c'est-à-dire la consommation des jeux d'argent. Pour cela, nous avons tout d'abord choisi d'afficher un graphique montrant la répartition de l'échantillon selon la fréquence de jeu par mois. Nous savons

déjà que 31.8% des répondants n'ont jamais joué aux jeux d'argents. Désormais, nous voyons comment se répartit la base de données pour ceux qui ont déjà joué. Tout d'abord, il y a plus de 40% de personnes qui jouent très peu, soit moins d'une fois par mois. On imagine qu'ils considèrent les jeux d'argent comme un divertissement très occasionnel. Ils sont conscients des risques et n'y voient aucun but lucratif. A l'inverse, un peu plus de 10% de personnes jouent plus de 10 fois par mois. Il peut s'agir ici d'addiction et il sera intéressant de voir de quels types de profils il s'agit.

Figure 9 : Répartition selon le budget de jeu par mois



Pour finir sur la partie descriptive, ce graphique montre la répartition de l'échantillon selon le budget de jeu par mois. Dans la même lignée que le précédent, il y a une importante différence de consommation entre les répondants au questionnaire avec une majorité de joueurs occasionnels qui prévoient entre 0 et 10€ pour les jeux d'argents dans le mois. A l'opposé, environ 10% prévoient plus de 50€ par mois pour jouer. Est-ce parce qu'ils ont plus de revenu que les autres ? Ou bien à l'inverse, ont-ils moins de revenu et cherchent à en avoir d'avantage au risque de perdre de l'argent ? Nous tenterons d'éclairer ce point pour le moins intéressant au cours de notre étude.

Ainsi, après avoir examiné en détail la composition de notre échantillon et la répartition selon certaines variables importantes à notre étude, nous allons procéder à des tests d'indépendance pour vérifier l'existence d'une corrélation entre nos variables à expliquer et nos variables explicatives.

Les tests d'indépendance (Khi2)

a. Test simple et détaillé

Après avoir passé au peigne fin notre base de données, nous allons procéder à des tests d'indépendance afin de vérifier si notre hypothèse nulle H0 (hypothèse d'indépendance) est acceptée face à l'hypothèse alternative H1 (hypothèse de non-indépendance). Dans un premier temps, nous effectuerons manuellement et de manière détaillée ce test entre notre variable à expliquer préférée (fréquence de jeu par mois) et notre variable explicative préférée (niveau de diplôme). Ensuite, nous ferons ce test pour afficher toutes les P-Values entre nos variables Y et X dans un même tableau.

Voici donc ci-dessous le détail du test d'indépendance du Khi2 entre la fréquence de jeu par mois des individus et leur niveau d'études. Pour la réalisation de ces tests, et dans le but d'obtenir une analyse pertinente, nous avons décidé de retirer les 10 mineurs de notre base de données, après avoir pris soin de vérifier qu'ils n'avaient jamais joués aux jeux d'argent. Cela nous permet aussi d'avoir un échantillon non biaisé par rapport au niveau de diplôme. Nous obtenons donc une base de données contenant 100 observations.

Figure 10 : Tableau des effectifs observés

Effectifs observés					
Niveau de diplôme des individus					
		Diplôme inférieur au bac	Diplôme entre bac et bac +2	Diplôme entre bac +3 et bac+5	Total
Fréquence de jeu par mois	Ne joue jamais	1	4	20	25
	Joue moins d'une fois par mois	5	17	19	41
	Joue entre 1 et 3 fois par mois	3	2	11	16
	Joue entre 3 et 9 fois par mois	2	3	2	7
	Joue entre 10 et 19 fois par mois	0	3	0	3
	Joue plus de 20 fois par mois	4	1	3	8
	Total	15	30	55	100

Tout d'abord, nous avons construit un tableau des effectifs observés obtenus grâce à un tri croisé entre nos variables. C'est dans ce but que nous avons recodés nos 2 variables afin de les rendre catégorielles. Nous avons donc 3 modalités pour la variable portant sur le niveau de diplôme, et 6 modalités pour la variable de la fréquence de jeu par mois. On remarque immédiatement que 66 joueurs jouent moins d'une fois par mois, ou bien n'ont tout simplement jamais joués. Les observations sont ainsi concentrées sur les 2 premières modalités de la variable fréquence. Bien que ceci soit intéressant, ce n'est pas important pour la réalisation du test du Khi2. Passons donc à la seconde étape qui consiste à créer des effectifs théoriques à partir des effectifs observés à l'aide de la formule suivante :

$$x_1 = \frac{\text{Total ligne 1} * \text{Total colonne 1}}{\text{Total}}$$

Figure 11 : Effectifs théoriques sous hypothèse d'indépendance

Effectifs théoriques					
Niveau de diplôme des individus					
		Diplôme inférieur au bac	Diplôme bac ou bac +2	Diplôme bac +3 ou bac+5	Total
Fréquence de jeu par mois	Ne joue jamais	3,75	7,5	13,75	25
	Joue moins d'une fois par mois	6,15	12,3	22,55	41
	Joue entre 1 et 3 fois par mois	2,4	4,8	8,8	16
	Joue entre 3 et 9 fois par mois	1,05	2,1	3,85	7
	Joue entre 10 et 19 fois par mois	0,45	0,9	1,65	3
	Joue plus de 20 fois par mois	1,2	2,4	4,4	8
	Total	15	30	55	100

Après avoir construit nos deux tableaux, nous cherchons calculer la distance entre les effectifs observés et théoriques grâce à la formule suivante :

$$\frac{(\text{effectif observé} - \text{effectif théorique})^2}{\text{effectif théorique}}$$

Figure 12 : Distance entre effectifs observés et théoriques

Distance entre effectifs observés et théoriques					
Niveau de diplôme des individus					
		Diplôme inférieur au bac	Diplôme bac ou bac +2	Diplôme bac +3 ou bac+5	Total
Fréquence de jeu par mois	Ne joue jamais	2,02	1,63	2,84	6,49
	Joue moins d'une fois par mois	0,22	1,80	0,56	2,57
	Joue entre 1 et 3 fois par mois	0,15	1,63	0,55	2,33
	Joue entre 3 et 9 fois par mois	0,86	0,39	0,89	2,13
	Joue entre 10 et 19 fois par mois	0,45	4,90	1,65	7,00
	Joue plus de 20 fois par mois	6,53	0,82	0,45	7,80
	Total			Statistique calculée	28,32

Ainsi, nous obtenons ce dernier tableau. La somme de la dernière colonne nous donne la statistique calculée qui est égale à 28,32. Nous devons maintenant la comparer avec la valeur critique qui est égale à 18,31 au seuil de 5% et 23,20 au seuil de 1%. Dans les deux cas, notre statistique calculée est supérieure à la valeur critique, ce qui signifie que nous pouvons rejeter l'hypothèse nulle. De plus, notre P-Value est de 0.0016, soit significative au seuil de 1%. Cela suggère qu'il existe une association significative entre le niveau de diplôme obtenu par les individus et la fréquence de jeu par mois.

b. Test d'indépendance généralisé

Maintenant que nous avons vu en détail comment réaliser et interpréter un test d'indépendance, nous allons regrouper dans un tableau toutes les P-Values obtenues grâce à un code sur le logiciel R permettant d'effectuer les tests entre nos différentes variables. Chaque variable a été recodée au préalable afin de toutes les rendre catégorielles. Voici les détails des différentes catégories pour chaque variable (hormis freq2 et nivdip qui correspondent aux variables utilisées pour le test détaillé).

Les variables à expliquer :

- La variable temps2 qui renseigne sur le nombre d'années depuis lequel l'individu a commencé à jouer aux jeux d'argents est composée de 4 modalités qui sont les suivantes : "Ne joue jamais", "Joue depuis moins de 4 ans", "Joue depuis moins de 12 ans", "Joue depuis plus de 12 ans".
- La variable nbtype2 nous indique le nombre de différents types de jeux d'argents auxquels joue l'individu. Elle se constitue des 4 modalités suivantes : "Ne joue jamais", "Moins de deux types de jeux", "3 ou 4 types de jeux", "Plus de 5 types de jeux".
- La variable mise2 concerne la mise moyenne des individus à chaque fois qu'ils jouent. Elle compte 6 modalités : "Ne joue jamais", "Mise 5€ ou moins", "Mise entre 6 et 10€", "Mise entre 11 et 15€", "Mise entre 16 et 20€", "Mise plus de 20€".
- La variable budgetjeu2 renseigne le budget par mois des individus qui est utilisé pour les jeux d'argents. Elle contient les 6 modalités suivantes : "Ne joue jamais", "Budget entre 0 et 10€", "Budget entre 11 et 25€", "Budget entre 26 et 50€", "Budget entre 51 et 100€", "Plus de 100€ de budget".
- La variable perte2 nous indique la plus grosse perte d'un individu en une seule mise. Elle compte 6 modalités qui sont les suivantes : "Ne joue jamais", "Plus grosse perte = 10€ ou moins", "Plus grosse perte entre 11 et 25€", "Plus grosse perte entre 26 et 50€", "Plus grosse perte entre 51 et 100€", "Plus grosse perte = plus de 100€".
- La variable bestgain2 porte sur le meilleur gain d'un individu en une seule mise. Elle se forme des 7 modalités suivantes : "Ne joue jamais", "N'a jamais gagné", "Entre 1 et 20€ en une seule mise", "Entre 21 et 75€ en une seule mise", "Entre 76 et 200€ en une seule mise", "Entre 201 et 1000€ en une seule mise", "Plus de 1000€ en une seule mise".
- La variable reinvesti2 nous donne le pourcentage du gain que l'individu réinvesti dans les jeux d'argents lorsqu'il gagne. Elle contient les 6 modalités suivantes : "Ne joue jamais", "Ne réinvestit pas", "Réinvestit entre 1 et 25% du gain", "Réinvestit entre 26 et 50% du gain", "Réinvestit entre 51 et 75% du gain", "Réinvestit entre 75 et 100% du gain".

Les variables explicatives :

- La variable catsocio concerne la catégorie socio-professionnelle de l'individu. Elle ne contient que les 7 modalités suivantes car la base de données n'a pas d'observations sur les autres catégories socio-professionnelles : "Sans activité", "Etudiant", "Cadre et profession intellectuelle supérieure", "Profession intermédiaire", "Employé", "Ouvrier", "Retraité".
- La variable revenu nous indique la classe de revenu annuel dans laquelle l'individu se situe. Elle compte les 6 modalités suivantes : "Moins de 12000 euros", "Entre 12000 et 25000 euros", "Entre 25000 et 40000 euros", "Entre 40000 et 60000 euros", "Entre 60000 et 80000 euros", "Plus de 80000 euros".

Figure 13 : Test Khi2 global

	nivdip	catsocio	revenu
temps2	0,01	0,00	0,00
freq2	0,00	0,01	0,64
nbtype2	0,00	0,66	0,02
mise2	0,16	0,12	0,41
budgetjeu2	0,01	0,12	0,54
perte2	0,04	0,27	0,47
bestgain2	0,12	0,55	0,38
reinvesti2	0,00	0,02	0,03

Voici donc ci-dessus le tableau contenant l'intégralité des P-Values entre nos variables à expliquer et nos variables explicatives avec une case de couleur verte lorsqu'elles sont significatives au seuil de 5%. Ainsi, presque la moitié des tests sont significatifs (20/48). De plus, notre variable explicative préférée (nivdip) semble être significative avec la quasi-totalité de nos variables concernant les jeux d'argent, et notamment les deux principales qui sont la fréquence de jeu par mois et le budget de jeu par mois, ce qui rejoint notre hypothèse émise précédemment. La variable catsocio, également importante dans la définition du milieu social, est-elle aussi significative avec la variable freq2. En revanche, elle ne l'est pas avec la variable budgetjeu2. Ensuite notre variable revenu est malheureusement très peu significative, si ce n'est avec des variables peu importante pour notre étude.

c. Test d'indépendance sur des sous-échantillons

Figure 14 : Test khi2 hommes

	nivdip	catsocio	revenu
temps2	0,25	0,01	0,01
freq2	0,06	0,11	0,46
nbtype2	0,09	0,47	0,33
mise2	0,18	0,09	0,05
budgetjeu2	0,14	0,71	0,58
perte2	0,32	0,54	0,48
bestgain2	0,23	0,90	0,85
reinvesti2	0,20	0,40	0,06

Ce tableau montre les P-values des tests d'indépendances réalisés sur une sous population rassemblant uniquement les hommes. Notre intuition et ce qu'on observe parmi notre entourage dans la vie de tous les jours semble montrer que les hommes ont tendances à plus consommer les jeux d'argents que les femmes. Ainsi, nous avons voulu séparer les deux sexes pour chercher un lien significatif entre les variables. Cependant, le tableau des hommes ne montre aucune significativité intéressante pour notre étude.

Figure 15 : Test Khi2 femmes

	nivdip	catsocio	revenu
temps2	0,04	0,00	0,03
freq2	0,47	0,51	0,63
nbtype2	0,30	0,86	0,01
mise2	0,35	0,31	0,78
budgetjeu2	0,03	0,64	0,64
perte2	0,04	0,69	0,81
bestgain2	0,01	0,29	0,26
reinvesti2	0,07	0,10	0,36

A l'inverse du tableau des tests sur les hommes qui montrait peu de significativité, celui sur les femmes en montre bien plus. Nous remarquons notamment une significativité intéressante pour notre étude entre la variable nivdip et la variable budgetjeu2.

Figure 16 : Test Khi2 étudiant

	nivdip	catsocio	revenu
temps2	0,11	0,00	0,00
freq2	0,02	0,96	0,89
nbtype2	0,04	0,59	0,83
mise2	0,15	0,04	0,48
budgetjeu2	0,03	0,99	0,95
perte2	0,02	0,95	0,64
bestgain2	0,24	0,96	0,75
reinvesti2	0,29	0,58	0,53

Notre base de données étant malheureusement peu hétérogène en termes de catégorie socio-professionnelle, nous avons jugé pertinent de séparer les étudiants des non étudiants afin de comprendre ce qui impact réellement la consommation de jeux d'argent au sein de ces deux sous-populations. Ici, notre variable nivdip est une fois de plus significative avec freq2 et budgetjeu2, il sera très intéressant de voir si ça se confirme lors de la régression.

Figure 17 : Test Khi2 non étudiant.

	nivdip	catsocio	revenu
temps2	0,57	0,14	0,09
freq2	0,14	0,00	0,80
nbtype2	0,33	0,90	0,28
mise2	0,77	0,04	0,20
budgetjeu2	0,17	0,38	0,68
perte2	0,13	0,09	0,78
bestgain2	0,38	0,83	0,81
reinvesti2	0,05	0,55	0,52

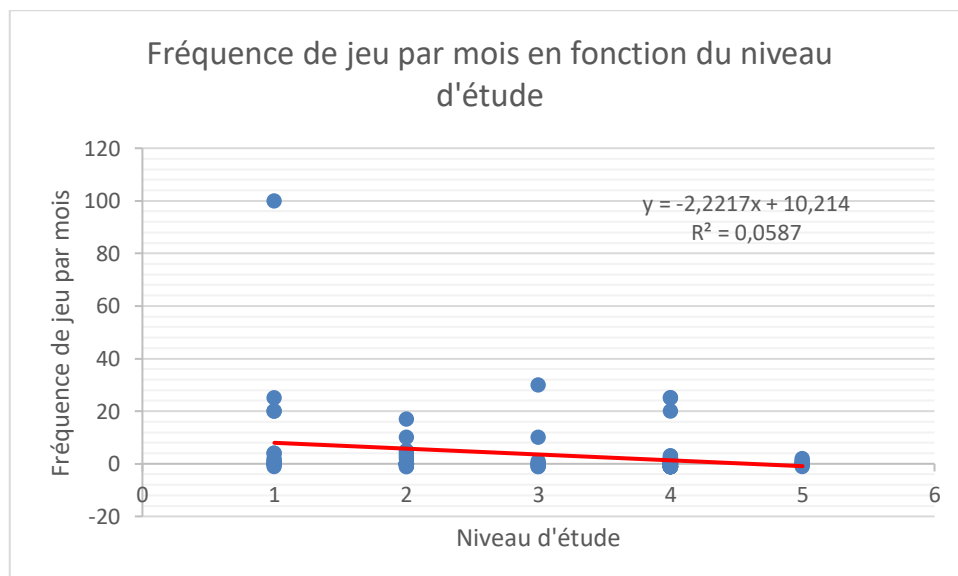
Enfin, le tableau des non étudiants nous permet d'y voir plus clair parmi les adultes qui sont censés être rentrés dans la vie professionnelle. Ici, le niveau de diplôme n'est plus significatif avec freq2 mais la catégorie socioprofessionnelle l'est quant à elle. Il serait très intéressant de voir quelle catégorie à le plus de rapports avec les jeux d'argents.

Régressions linéaires

a. La régression simple

Il est temps de passer à la partie des régressions linéaires, afin d'obtenir des résultats et d'en tirer des conclusions qui répondront ou non à nos hypothèses de départ. Tout d'abord, commençons par nos principales variables qui sont la fréquence de jeu par moi et le niveau de diplôme des individus.

Figure 18 : Fréquence de jeu par mois en fonction du niveau d'étude



Ce nuage de points entre nos 2 variables préférées semble nous montrer qu'il existe une corrélation négative entre le niveau d'étude et la fréquence de jeu par mois. En effet, la droite décroissante signifie que lorsqu'on augmente le niveau de diplôme, cela donne lieu à une réduction de la fréquence de jeu par mois. Nous allons maintenant procéder aux régressions afin de voir si ce lien est significatif ou non.

Dans un premier temps, regardons le résultat d'une régression simple entre la fréquence de jeu par mois et le niveau d'étude.

Figure 19 : Régression simple entre freq et nivétude2

Régression simple entre freq et nivétude2	
	nivétude2
β de nivétude2 sur freq	-2,22
Significativité de nivétude2	**
R2 ajusté du modèle	5,90%
Significativité globale du modèle	**
*p<0,1; **p<0,05; ***p<0,01	

Cette régression indique que notre coefficient est significatif au seuil de 5%. Ainsi, il s'interprète de la manière suivante : Toutes choses égales par ailleurs, une augmentation de 1 niveau de diplôme entraîne une baisse de 2,22 unités en moyenne par mois dans la fréquence de jeu des individus.

Pour donner suite à ce tableau, nous avons effectué toutes les régressions simples possibles avec nos variables indépendantes « revenu2 » et « nivétude2 » (voir annexe). Cependant, comme pour les tests du Khi2, les résultats de la variable explicative « revenu2 » ne sont pas très concluants. A l'inverse, le niveau d'étude affiche une significativité pour la majorité des variables à expliquer. Il sera intéressant de procéder à une régression multiple avec les plus pertinentes d'entre-elles.

b. La régression multiple

Revenons maintenant sur notre régression entre freq et nivétude2, et observons l'évolution des résultats en ajoutant cette fois-ci les variables de contrôle.

Figure 20 : Tableau des « blocs » de variables de contrôle.

Bloc 1 : Renseignement généraux	age, genre2, region, finétude2
Bloc 2 : Précisions sur le logement	habitat, loyer, apl_loyer

Figure 21 : Régression entre freq et nivétude2 avec variables de contrôle

De la régression simple au modèle final
--

	nivetude2	nivetude2 + bloc 1	nivetude2 + bloc 1:2
β de nivetude2 sur freq	-2,22	-2,81	-2,78
Significativité de nivetude2	**	***	**
R2 ajusté du modèle	5,90%	19,40%	20,30%
Significativité globale du modèle	**	***	***
*p<0,1; **p<0,05; ***p<0,01			

On remarque dans un premier temps que notre variable nivetude2 reste significative au seuil de 5% malgré l'ajout au fur et à mesure des blocs de variables, et est même significative au seuil de 1% après l'ajout du premier bloc. De plus, notre coefficient varie peu, puisqu'il reste entre -2,22 et -2,81. Enfin, notre modèle final nous donne un R^2 de 20,30%, ce qui signifie que la fréquence de jeu par mois est expliquée à 20,30% par notre modèle, ce qui est assez peu.

Figure 22 : Modèle final et interprétation des coefficients significatifs

Dependent variable:	
freq	
nivetude2	-2.782** (1.065)
age	-0.419*** (0.123)
genre2	4.745** (2.249)
regionEst	-6.940 (5.952)
regionNord	-10.359 (8.384)
regionOuest	-3.007 (3.596)
regionSud	0.616 (6.248)
finetude2	9.006*** (3.291)
habitatEn collocation	-3.338 (4.403)
habitatSeul	-0.318 (3.304)
loyerJe paye seul(e) mon loyer	-6.186 (3.731)
loyerMes parents payent mon loyer	-8.795** (4.232)
loyerPas de loyer	-8.252* (4.605)
apl_loyer	1.990 (3.849)
Constant	26.751*** (7.746)
Observations	100
R2	0.316
Adjusted R2	0.203
Residual Std. Error	10.449 (df = 85)
F Statistic	2.801*** (df = 14; 85)
Note: *p<0.1; **p<0.05; ***p<0.01	

- Toutes choses égales par ailleurs, une augmentation de 1 niveau de diplôme entraine une diminution de la fréquence de jeu par mois de 2,78 unités
- Toutes choses égales par ailleurs, lorsque l'âge augmente d'une année, la fréquence de jeu diminue de 0,41 unités par mois.
- Genre : Homme = 1. Toutes choses égales par ailleurs, les hommes jouent en moyenne 4,74 unités par mois de plus que les femmes.
- Toutes choses égales par ailleurs, les personnes ayant fini leurs études ont une fréquence de jeu par mois de 9 unités supplémentaires que les étudiants.
- Toutes choses égales par ailleurs, les individus qui se voient payer leur loyer par leurs parents jouent 8,79 unités par mois de moins que ceux qui partagent leur loyer avec leur famille ou leurs colocataires.
- Toutes choses égales par ailleurs, les individus n'ayant pas de loyer à payer ont une fréquence de jeu par mois inférieure de 8,25 unités comparé à ceux qui partagent leur loyer avec leur famille ou leurs colocataires.

Ainsi, notre modèle final est constitué d'une variable X qui est significative, et de 13 autres variables de contrôle dont 5 qui affichent une significativité plus ou moins forte. De plus, ce modèle explique à 20,30% la variance de notre variable dépendante.

c. Les régressions sur des sous échantillons

Lors des tests du Khi2 nous avons divisé la population entre les hommes et les femmes puis entre les étudiants et les non étudiants afin de voir s'il y avait une différence entre les groupes. Pour aller jusqu'au bout de notre analyse, nous avons réalisé des régressions sur ces quatre groupes distinctivement. Voici ce que les résultats ont donnés :

Figure 23 : Régression entre freq et nivetude2 avec variables contrôle sur la sous population « Etudiants »

De la régression simple au modèle final (Etudiants)			
	nivetude2	nivetude2 + bloc 1	nivetude2 + bloc 1:2
β de revenu sur freq	-0,07	0,35	0,63
Significativité de nivetude2			
R2 ajusté du modèle	0,01%	-2,20%	0,30%
Significativité globale du modèle			
*p<0,1; **p<0,05; ***p<0,01			

Commençons par le groupe des étudiants. Le niveau de diplôme qui semblait significatif avec la fréquence de jeu par mois lors du test du Khi2 ne l'est plus du tout lors de la régression simple et multiple. Ainsi, avec les 66 observations que l'on a et selon nos résultats, le niveau de diplôme des personnes qui n'ont pas terminé leurs études n'impact pas significativement sur la fréquence de jeu d'argent par mois.

Voyons si il en est de même pour les non étudiants :

Figure 24 : Régression entre freq et nivetude2 avec variables de contrôle sur la sous population « Non étudiants »

De la régression simple au modèle final (Non étudiants)			
	nivetude2	nivetude2 + bloc 1	nivetude2 + bloc 1:2
β de revenu sur freq	-3,37	-4,71	-3,67
Significativité de nivetude2	**	**	
R2 ajusté du modèle	9,90%	26,60%	26,30%
Significativité globale du modèle	**	***	**
*p<0,1; **p<0,05; ***p<0,01			

Les résultats indiquent le contraire de ce qu'on avait pu constater lors des tests du Khi2 et les variables affichent une significativité entre elles, à l'inverse du groupe étudiant. Cependant, avec l'ajout des variables de contrôle sur le logement, on remarque que le niveau d'étude perd sa significativité. Si on se fie seulement à la deuxième colonne, on peut en déduire que le niveau de diplôme pour les personnes ayant terminé leurs études impact significativement la fréquence de jeu par mois. En effet,

toutes choses égales par ailleurs, une augmentation d'un niveau de diplôme entraîne une diminution de la fréquence de jeu par mois de 4,71 unités. Ce résultat est cependant à prendre avec des pincettes étant donné que cette sous population ne contient que 44 observations.

Ensuite, nous allons comparer les hommes et les femmes par rapport à l'impact de leur niveau d'étude sur leur fréquence de jeu par mois.

Figure 25 : Régression entre freq et nivétude2 avec variables de contrôle sur la sous population « Femmes »

De la régression simple au modèle final (Femmes)			
	nivétude2	nivétude2 + bloc 1	nivétude2 + bloc 1:2
β de revenu sur freq	-0,21	-0,09	-0,07
Significativité de nivétude2	*		
R2 ajusté du modèle	5,80%	17,00%	8,40%
Significativité globale du modèle	*	**	
*p<0,1; **p<0,05; ***p<0,01			

La régression multiple sur le groupe des femmes indique que l'impact du niveau de diplôme n'est pas significatif, malgré une régression simple qui semblait montrer l'inverse au seuil de 10%. La variable freq pour cette sous population s'étend de -1 à 4, avec seulement 10 valeurs au dessus de 0. Autrement dit, les 51 observations ne nous permettent pas de dire que le niveau d'étude des femmes influe sur leur fréquence de jeu. Qu'en est-il du groupe des hommes ?

Figure 26 : Régression entre freq et nivétude2 avec variables de contrôle sur la sous population « Hommes »

De la régression simple au modèle final (Hommes)			
	nivétude2	nivétude2 + bloc 1	nivétude2 + bloc 1:2
β de revenu sur freq	-3,7	-4,61	-4,83
Significativité de nivétude2	**	**	**
R2 ajusté du modèle	8,50%	21,10%	28,70%
Significativité globale du modèle	**	**	**
*p<0,1; **p<0,05; ***p<0,01			

Pour finir avec les sous-populations, intéressons-nous au groupe masculin composé de 49 observations. Ici, on peut constater une significativité constante tout au long du modèle entre nos 2 variables. Ainsi, malgré un faible nombre d'observation, on peut tirer la conclusion que chez les hommes, toutes choses égales par ailleurs, lorsque le niveau d'étude augmente de 1 unité, la fréquence de jeu diminue de 4,83 unités.

d. Régression multiple avec une autre variable à expliquer

Après avoir construit un modèle de régression multiple entre la variable dépendante « freq » et la variable explicative « nivétude2 » accompagnée de quelques variables de contrôles, puis avoir effectué l'analyse sur une sous population, nous allons réaliser une seconde régression en remplaçant cette fois

la fréquence de jeu par mois par le budget de jeu par mois, tout en gardant les mêmes variables explicatives.

Figure 27 : Régression entre budgetjeu et nivetude2 avec variables de contrôle

De la régression simple au modèle final			
	nivetude2	nivetude2 + bloc 1	nivetude2 + bloc 1:2
β de nivetude2 sur budgetjeu	-16,36	-14,48	-10,21
Significativité de nivetude2	*		
R2 ajusté du modèle	3,60%	11,00%	9,40%
Significativité globale du modèle	*	**	*
*p<0,1; **p<0,05; ***p<0,01			

Malgré une régression simple significative au seuil de 10%, le budget de jeu par mois ne semble pas être expliqué significativement par le niveau d'étude lorsqu'on ajoute les blocs de variables de contrôle.

Conclusion

Pour résumer, nous avons réalisé une étude sur la base d'un questionnaire comptant 110 réponses au total. Après avoir analysé des graphiques de répartition pour mieux comprendre la composition de notre base de données, nous avons réalisé des tests d'indépendance du Khi2 afin de se faire une idée sur la signification de nos variables entre elles. Pour finir, nous avons procédé à des régressions simples et multiples sur notre population et nos sous populations.

Nous avons trouvé des résultats qui convergent globalement tous et tendent à dire que plus les individus sont diplômés et moins ils consomment de jeux d'argent. Ainsi, ces résultats rejoignent notre hypothèse de départ, ainsi que les résultats de l'ODJ en 2012, 2017 et 2019. Ils rejoignent aussi l'étude de T. Amadiou en 2015. De plus, nous avons trouvé une forte significativité pour les hommes que nous n'avons pas retrouvé chez les femmes. Cela s'explique par le manque d'observation de femmes « joueuses ». En effet, bien que nous ayons un échantillon qui respecte la parité homme-femme, avec respectivement 49 et 51 observations, nous avons constaté une plus large majorité de joueurs chez les hommes. Cette constatation rejoint une nouvelle fois les travaux cités précédemment, ainsi que notre intuition de départ. La question ouverte posée aux personnes n'ayant jamais joué aux jeux d'argent n'a pas permis d'expliquer la raison de cette différence entre les hommes et les femmes, les réponses étant globalement orientées autour de « l'inutilité de jouer » ou du « manque d'argent ». On peut tout de même en déduire que les femmes ont une aversion au risque bien plus élevée que les hommes en matière de jeux d'argent.

Bien que les résultats de notre étude rejoignent ceux des études officielles de l'ODJ par exemple, il est important de les interpréter avec des pincettes. En effet, notre étude contient plusieurs limites qui peuvent biaiser le coefficient et la significativité des variables. La principale limite est le nombre

d'observations récoltées, qui est bien trop faible pour produire des résultats fiables. La majorité des répondants sont des étudiants de 21 ans, ce qui est peu représentatif de la société. Ensuite, nous estimons après réalisation de l'étude que le questionnaire est imparfait dans de nombreux points. Tout d'abord, la question du revenu aurait été bien plus pertinente si elle demandait aux répondants de donner leur salaire mensuel net par mois, et non pas de le donner par classe comme il a été fait. De ce fait, nous avons basé notre interprétation du milieu social uniquement sur le niveau d'étude, ce qui n'est pas tout à fait exact. Ensuite, dans la continuité de la limite précédente, nous trouvons qu'il a manqué de questions plus précises sur le milieu social ou les caractéristiques socio-démographiques. Il aurait été intéressant de demander si les individus habitent en ville ou en campagne, s'ils pratiquent un sport, s'ils jouent sur internet ou pas, la catégorie socioprofessionnelle des parents, etc... Enfin, les variables de fréquence de jeu et de budget de jeu auraient pu être demandées par année plutôt que par mois, afin d'avoir une réponse plus précise sur les personnes jouant moins d'une fois par mois. Pour finir sur les limites, notre étude ne prend pas en compte les variables de contrôles basées sur les jeux d'argent directement. Par exemple, l'étude de l'ODJ en 2019 montre que le genre est significatif dans un premier modèle, mais qu'en ajoutant le type de jeux d'argent (paris sportifs, poker, casino...), il ne l'est plus pour certains de ces jeux. Nous cherchions à nous focaliser uniquement sur la consommation de jeux d'argent en elle-même, sans rentrer dans les détails de cette dernière.

Si l'étude était à refaire, tout en prenant en compte les limites citées précédemment qu'il faut améliorer, nous avons pensé à la création d'un algorithme très précis qui donne un « score » de milieu social sur 100 sur la base de critères très précis. Cet algorithme permettrait de regrouper les différentes variables qui représentent le milieu social en une seule, à l'inverse de l'étude actuel qui finalement ne s'est basé que sur le niveau de diplôme. Ainsi, nous pourrions évaluer le réel impact du milieu social sur la consommation des jeux d'argents.

Annexe

Questionnaire

Partie A : Renseignements généraux

- 1- Vous êtes
 - a. Homme
 - b. Femme

- 2- Quel âge avez-vous ?
...

- 3- Dans quel type de logement habitez-vous ?
 - a. Chez vos parents
 - b. Seuls
 - c. En collocation
 - d. Autres

- 4- Quel est votre numéro de département ?

Partie B : Evaluation du milieu social

- 5- Quel est votre dernier diplôme obtenu ?
 - a. Bac général
 - b. Bac technologique ou professionnel, brevet professionnel
 - c. Licence
 - d. Master
 - e. BTS, IUT
 - f. Pas de diplôme
 - g. Autres

- 6- Avez-vous terminé vos études ?
 - a. Oui
 - b. Non
 - c. Ne sais pas

- 7- Quelle est votre catégorie socioprofessionnelle ?
 - a. Etudiants
 - b. Agriculteurs exploitants

- c. Artisans, commerçants et chefs d'entreprise
- d. Cadres et professions intellectuelles supérieures
- e. Professions intermédiaires
- f. Employés
- g. Ouvriers
- h. Retraités
- i. Sans activité
- j. Autres

8- Quel est le revenu net annuel de votre ménage ?

- a. Moins de 12000 euros
- b. Entre 12000 et 25000 euros
- c. Entre 25000 et 40000 euros
- d. Entre 40000 et 60000 euros
- e. Entre 60000 et 80000 euros
- f. Plus de 80000 euros

9- Votre salaire représente combien de pourcents de vos revenus (environ) ?

...

10- Avez-vous bénéficié de bourses, d'aides sociales ou d'autres formes de soutien financier ?

- a. Oui
- b. Non

11- Avez-vous voyagé à l'étranger ou dans d'autres régions du pays récemment ?

- a. Oui
- b. Non

Partie C : Evaluation de la consommation des jeux d'argent

12- Avez-vous déjà joué à des jeux d'argent ?

- a. Oui
- b. Non

12(bis)- Si non, pourquoi ?

...

13- A quel âge avez-vous commencé à jouer ?

...

14- Depuis combien de temps jouez-vous (en année) ?

...

15- A quelle fréquence jouez-vous ?

- a. Jamais
- b. De temps en temps
- c. Fréquemment
- d. Tous les jours

16- Plus précisément, environ combien de fois par mois jouez-vous ?

...

17- A quels jeux jouez-vous ?

- a. Jeu de grattage
- b. Jeux de Casino
- c. Paris sportifs
- d. Paris hippiques
- e. Loto, Euro millions
- f. Autres

18- A combien de jeux différents jouez-vous ?

...

19- Quelle somme dépensez-vous pour chaque mise jouée (en euros) ?

...

20- Quel budget consacrez-vous environ aux jeux mensuellement (en euros) ?

...

21- Avez-vous déjà gagné ?

- a. Oui
- b. Non

22- Quelle est votre plus grosse perte en une seule mise (en euros) ?

...

23- Quel est votre plus gros gain en une seule mise (en euros) ?

...

24- Quelle est votre somme totale de gains depuis que vous jouez ?

- a. 1 à 20 euros
- b. 21 à 60 euros
- c. 61 à 100 euros
- d. 101 à 200 euros
- e. 201 à 500 euros
- f. 501 à 1000 euros
- g. Plus de 1000 euros

25- Quelle est votre technique de jeu ?

- a. Hasard
- b. Chiffre symbolique et signe astrologique
- c. Logique statistique
- d. Connaissances hippiques/sportives
- e. Autres

26- Pour quelles raisons jouez-vous ?

- a. Gains d'argent
- b. Plaisir du jeu
- c. Habitudes
- d. Influence de l'entourage
- e. Goût du risque
- f. Autres

27- Que faites-vous lorsque vous gagnez ?

- a. Réinvestissement des gains (rejouer la somme gagnée)
- b. Economie, placement
- c. Achats divers
- d. Autre réponse

28- Lorsque vous gagnez, quel pourcentage réinvestissez-vous dans les jeux ?

...

29- Pensez-vous prendre des risques lorsque vous jouez de l'argent ?

- a. Oui
- b. Non

30- Le fait de jouer vous apporte-t-il une satisfaction ?

- a. Oui
- b. Non

31- Vous sentez vous obligé de jouer de l'argent ?

- a. Oui
- b. Non

32- Avez-vous déjà misé plus que ce que vous pouviez vous permettre de perdre ?

- a. Oui
- b. Non

33- Avez-vous déjà emprunté de l'argent à quelqu'un ou vendu quelque chose pour jouer ?

- a. Oui
- b. Non

34- Avez-vous déjà rejoué dans la même journée pour récupérer ce que vous avez perdu ?

- a. Oui
- b. Non

Bibliographie

AMADIEU Thomas, (2015). Prises de risques délibérées avec l'argent : les modalités de consommation des jeux de hasard.

CHEVALIER Serge, MARTIN Isabelle, GUPTA Rina, DEREVENSKY Jeffrey, (2005). Jeux de hasard et d'argent (chapitre 5, page 131 à 147), extrait de « Enquête québécoise sur le tabac, l'alcool, la drogue et le jeu chez les élèves du secondaire, 2004 »

COSTES Jean-Michel, EROUKMANOFF Vincent, (2018). Les pratiques de jeux d'argent sur Internet en France en 2017.

COSTES Jean-Michel, RICHARD Jean-Baptiste, EROUKMANOFF Vincent, (2020). Les problèmes liés aux jeux d'argent en France, en 2019.

TOVAR Marie-Line, COSTES Jean-Michel, EROUKMANOFF Vincent, (2013). Les jeux d'argent et de hasard sur Internet en France en 2012.

Régressions : Résultats complémentaires

Figure 28 : Toutes les régressions simples possibles avec nivetude2 et revenu2

nivetude2	Variable_X	Variable_Y	Coefficient_Beta	P_Value
nivetude2 & temps	nivetude2	temps	-1,49	0,11
nivetude2 & freq	nivetude2	freq	-2,22	0,02
nivetude2 & nbtype	nivetude2	nbtype	-0,25	0,07
nivetude2 & mise	nivetude2	mise	-1,16	0,11
nivetude2 & budgetjeu	nivetude2	budgetjeu	-16,36	0,06
nivetude2 & perte	nivetude2	perte	-11,21	0,05
nivetude2 & bestgain	nivetude2	bestgain	-102,16	0,06
nivetude2 & reinvesti	nivetude2	reinvesti	-5,17	0,11
revenu2	Variable_X	Variable_Y	Coefficient_Beta	P_Value
revenu2 & temps	revenu2	temps	0,00	0,00
revenu2 & freq	revenu2	freq	0,00	0,28
revenu2 & nbtype	revenu2	nbtype	0,00	0,05
revenu2 & mise	revenu2	mise	0,00	0,66
revenu2 & budgetjeu	revenu2	budgetjeu	0,00	0,42
revenu2 & perte	revenu2	perte	0,00	0,56
revenu2 & bestgain	revenu2	bestgain	0,00	0,85
revenu2 & reinvesti	revenu2	reinvesti	0,00	0,34