



Rapport

Sport et Drogue

Étude réalisée par les
L2 STAPS et Aline Laufer (L3 STAPS)

sous la direction de
Christophe Genolini

Table des matières

1	Introduction	7
1.1	EPO	7
1.2	Avertissement	7
1.3	Enquête “Drogue et pratique sportive”	7
2	Présentation de l’étude	7
3	Analyse univariée	8
3.1	Nettoyage des données	8
3.2	Détail	9
3.2.1	Vous êtes : Un homme/ Une femme ?	9
3.2.2	Quel est votre âge ?	9
3.2.3	Hors des heures obligatoires en fac, combien d’heures de sport par semaine pratiquez-vous ?	10
3.2.4	Etes-vous inscrit dans un club sportif ?	10
3.2.5	Pratiquez-vous ou avez-vous pratiqué un sport en compétition ?	10
3.2.6	Quelle est votre consommation de tabac ?	11
3.2.7	Combien de fois avez-vous consommé de l’alcool au cours du dernier mois ?	11
3.2.8	Combien de fois avez-vous consommé du cannabis au cours du dernier mois ?	11
3.3	Bilan de l’analyse univariée	12
4	Analyse bivariée de [CONSO_CANNABIS]	13
4.1	[CONSO_CANNABIS] vs [AGE]	13
4.1.1	Graphe	13
4.1.2	Diagnostic	13
4.1.3	Test	13
4.2	[CONSO_CANNABIS] vs [SEXE]	14
4.2.1	Graphe	14
4.2.2	Diagnostic	14
4.2.3	Test	14
4.3	[CONSO_CANNABIS] vs [SPORTCLUB]	15
4.3.1	Graphe	15
4.3.2	Diagnostic	15
4.3.3	Test	15
4.4	[CONSO_CANNABIS] vs [SPORTCOMPET]	16
4.4.1	Graphe	16
4.4.2	Diagnostic	16

4.4.3	Test	16
4.5	[CONSO CANNABIS] vs [CONSO TABAC]	17
4.5.1	Graphe	17
4.5.2	Diagnostic	17
4.5.3	Test	17
4.6	[CONSO CANNABIS] vs [CONSO ALCOOL]	18
4.6.1	Graphe	18
4.6.2	Diagnostic	18
4.6.3	Test	18
4.7	[CONSO CANNABIS] vs [SPORT HEURE]	19
4.7.1	Graphe	19
4.7.2	Diagnostic	19
4.7.3	Test	19
5	Construction d'un modèle pour la variable [CONSOCANNABIS]	20
5.1	Détail du modèle	20
5.2	Bilan	21
6	Analyse bivariée de [CONSOALCOOL]	21
6.1	[CONSO ALCOOL] vs [AGE]	21
6.1.1	Graphe	21
6.1.2	Diagnostic	22
6.1.3	Test	22
6.2	[CONSO ALCOOL] vs [SEXE]	22
6.2.1	Graphe	23
6.2.2	Diagnostic	23
6.2.3	Test	23
6.3	[CONSO ALCOOL] vs [SPORT CLUB]	23
6.3.1	Graphe	24
6.3.2	Diagnostic	24
6.3.3	Test	24
6.4	[CONSO ALCOOL] vs [SPORT COMPET]	24
6.4.1	Graphe	25
6.4.2	Diagnostic	25
6.4.3	Test	25
6.5	[CONSO ALCOOL] vs [CONSO TABAC]	25
6.5.1	Graphe	26
6.5.2	Diagnostic	26
6.5.3	Test	26
6.6	[CONSO ALCOOL] vs [CONSO CANNABIS]	26
6.6.1	Graphe	27

6.6.2	Diagnostic	27
6.6.3	Test	27
6.7	[CONSOALCOOL] vs [SPORTHEURE]	27
6.7.1	Graphe	28
6.7.2	Diagnostic	28
6.7.3	Test	28
6.8	Modèle	28
7	Conclusion générale	28
7.1	Réponses aux questions initiales	28
7.2	Limites de l'étude	29

1 Introduction

1.1 EPO

Le projet Enquête Paris Ouest est une collaboration inter UFR proposant aux étudiants de mener des enquêtes statistiques dans l'université. L'idée maîtresse est de proposer aux étudiants de travailler sur des sujets qu'ils choisissent eux-mêmes et de devenir ainsi acteurs de leur enseignement. Ils interviennent à tous les niveaux : conception du questionnaire, passation, saisie des données, analyse statistique et le cas échéant demande de financement.

L'enquête conduit à la construction d'une base de données. Cette base est mise à disposition de ceux qui le souhaitent via le site web de EPO. Selon les UFR, selon les niveaux, selon les programmes, les enseignants sont libres d'utiliser les données pour faire faire à leurs étudiants les analyses correspondant à leur domaine d'étude. Au final, une compilation des différentes analyses est réalisée puis diffusée sur l'ensemble de l'université, à la plus grande fierté des étudiants... C'est un de ces rapports que vous avez entre les mains.

1.2 Avertissement

Le présent rapport est une compilation des travaux effectués par les étudiants des différentes années. IL N'A PAS LA PRETENTION D'ÊTRE UN RAPPORT SCIENTIFIQUE (voir la section 7.2 page 29 : "Limites de l'enquête"). Il constitue avant tout le résumé d'un exercice fait par et pour les étudiants, pour que l'enseignement des statistiques soit un peu moins austère...

1.3 Enquête "Drogue et pratique sportive"

Cette enquête, réalisé en 2006, a été la première des enquêtes EPO. Elle a été réalisée au sein de l'UFR STAPS. Seul un groupe TD avait enquêté, le nombre de sujet inclus est assez faible (147). Les L2 ont fait l'analyse univariée, Aline Laufer, étudiante de L3, a fait l'analyse bivariée et le modèle.

- Conception, passation du questionnaire et analyse univarié : L2 STAPS
- Analyse univarié, analyse bivariée et modélisation : Aline Laufer

2 Présentation de l'étude

L'étude porte sur la consommation de drogue, d'alcool et la pratique sportive. 22 étudiants de L2 ont participé à la construction d'un questionnaire. Ils ont eu pour mission de faire remplir le questionnaire par des étudiants STAPS.

L'étude a deux objectifs :

1. Examiner la consommation de drogue légère (cannabis) selon le niveau de pratique sportive. Notre hypothèse sera la suivante : plus le niveau de pratique est élevé, plus la consommation de drogue est faible.
2. Examiner la consommation d'alcool selon le niveau de pratique sportive. Notre hypothèse sera la suivante : plus le niveau de pratique est élevé, plus la consommation d'alcool est faible.

L'enquête a été réalisée sous forme d'un questionnaire. Le niveau de pratique sportive a été mesuré par deux questions : le nombre d'heure de sport par semaine, et le niveau de compétition.

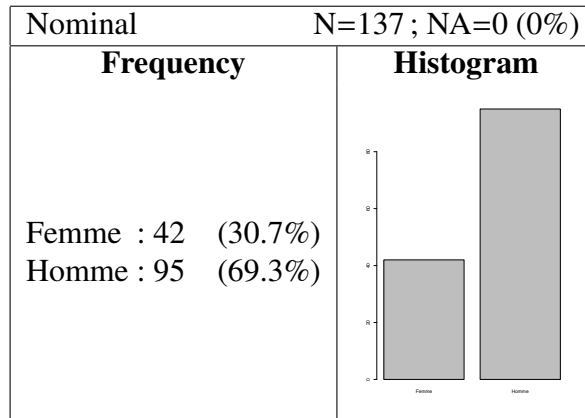
3 Analyse univariée

3.1 Nettoyage des données

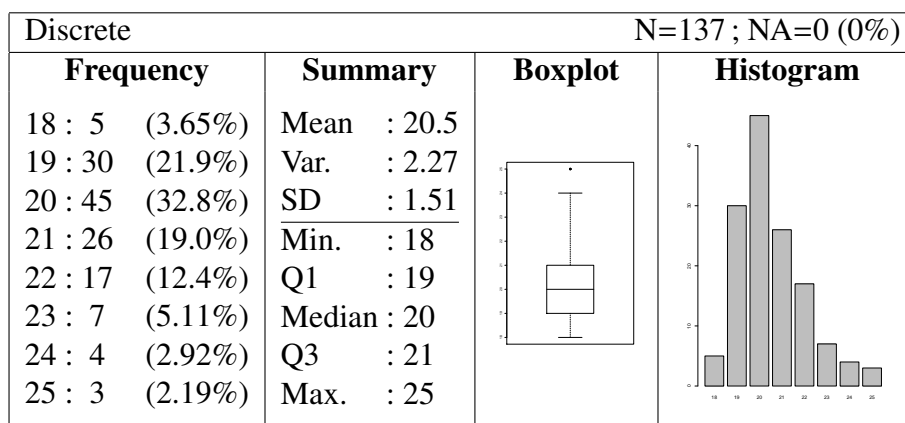
1. Dix questionnaires, saisis par une seule et même enquêtrice, présentent des valeurs extraimement élevées (abérantes). Les dix questionnaires sont exclus.
2. [SEXE] : les modalités sont uniformisées en (Homme) et (Femme).
3. [CONSOALCOOL] : un sujet a écrit (6litres). Cette valeur n'étant pas exploitable, elle est supprimée.
4. [CONSOALCOOL] et [CONSOCANNABIS] : un sujet a saisi un intervalle (15-20). Il est remplacé par sa valeur centrale (17.5).
5. [SPORTCOMPET] : seules trois personnes ont un niveau (International). Les modalités (National) et (International) sont fusionnées en (National et plus).
6. [SPORTPRINCIPAL] : avec 48 modalités différentes, cette variable est clairement inexploitable.
7. Pas de problème dans les autres variables.

3.2 Détail

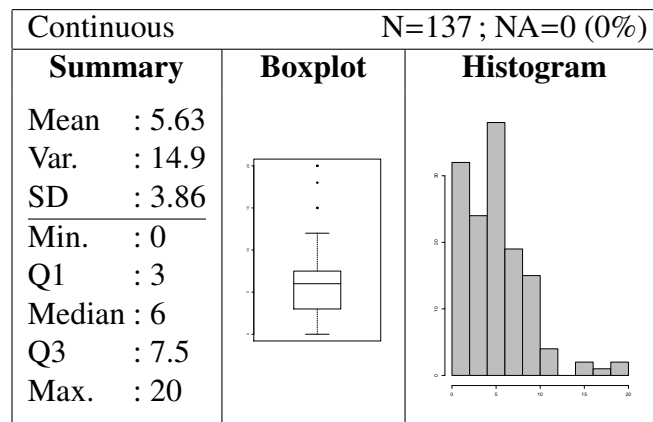
3.2.1 Vous êtes : Un homme/ Une femme ?



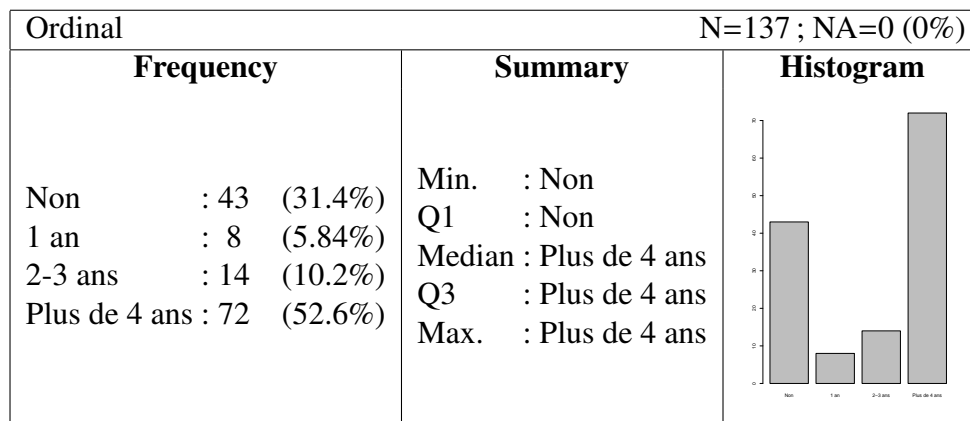
3.2.2 Quel est votre âge ?



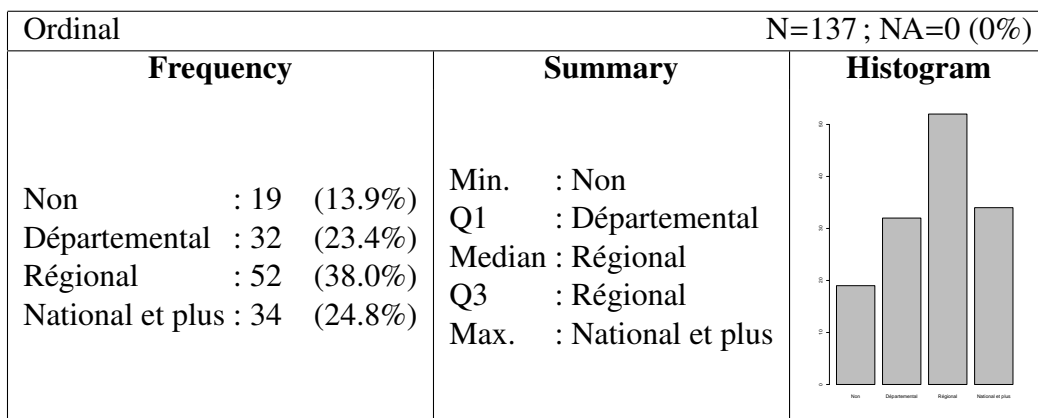
3.2.3 Hors des heures obligatoires en fac, combien d'heures de sport par semaine pratiquez-vous ?



3.2.4 Etes-vous inscrit dans un club sportif ?



3.2.5 Pratiquez-vous ou avez-vous pratiqué un sport en compétition ?



3.2.6 Quelle est votre consommation de tabac ?

Ordinal		N=137 ; NA=0 (0%)		
Frequency		Summary		Histogram
Jamais	: 87 (63.5%)	Min.	: Jamais	
Arrêté de fumer	: 21 (15.3%)	Q1	: Jamais	
Moins de 10	: 21 (15.3%)	Median	: Jamais	
Plus de 10	: 8 (5.84%)	Q3	: Arrêté de fumer	
		Max.	: Plus de 10	

3.2.7 Combien de fois avez-vous consommé de l'alcool au cours du dernier mois ?

Continuous		N=137 ; NA=1 (0.73%)	
Summary		Boxplot	Histogram
Mean	: 5.77		
Var.	: 63.4		
SD	: 7.96		
Min.	: 0		
Q1	: 0		
Median	: 3		
Q3	: 8.5		
Max.	: 55		

3.2.8 Combien de fois avez-vous consommé du cannabis au cours du dernier mois ?

Continuous		N=137 ; NA=0 (0%)	
Summary		Boxplot	Histogram
Mean	: 3.41		
Var.	: 83.3		
SD	: 9.12		
Min.	: 0		
Q1	: 0		
Median	: 0		
Q3	: 1		
Max.	: 69		

3.3 Bilan de l'analyse univariée

L'échantillon de notre étude est constitué de 137 étudiants

- [AGE] : Ils ont entre 18 et 25 ans (moyenne 20.5)
- [SEXE] : On trouve plus d'hommes, 95 (69.3%) contre 42 (30.7%) femmes.
- [SPORTHEURE] : Ils pratiquent de 0 à 20 heures de sport par semaine (moyenne 5.63), cette variable suit approximativement une loi de poisson.
- [SPORTCLUB] : 72 (52.6%) étudiants sont inscrits en club depuis plus de 4 ans, 43(31.4%) étudiants ne sont pas inscrits en club.
- [SPORTCOMPET] : La majorité des étudiants font des compétitions, seulement 19 (13.9%) étudiants n'en font pas.
- [CONSO TABAC] : Ils sont 108 (78.8%) à ne pas ou plus fumer contre 29 (21.2%) fumeurs.
- [CONSOALCOOL] : Cette variable suit une loi de poisson. Les étudiants ont consommé entre 0 et 55 fois de l'alcool au cours du dernier mois (moyenne 5.77). On note la présence une valeur extrême (55) mais la présence d'un 40 fait qu'elle n'est pas considéré comme abérante..
- [CONSOCANNABIS] : 101 (73.7%) étudiants n'ont pas consommé de cannabis au cours du dernier mois. La moyenne est de 3.41. La encore, on note une valeur extrême (69). Consommer du cannabis 69 fois dans le mois, cela fait une moyenne de 2,3 par jour. Cela semble possible, nous décidons de conserver cette valeur. Néanmoins, en cas de doute et cette valeur étant très isolée, nous ferons les tests une fois avec et une fois sans cette valeur, pour vérifier la robustesse des résultats.

4 Analyse bivariée de [CONSO_CANNABIS]

La variable qui nous intéresse est la variable [CONSO_CANNABIS]. D'une manière exploratoire, nous allons examiner les variables avec lesquelles elle est liée.

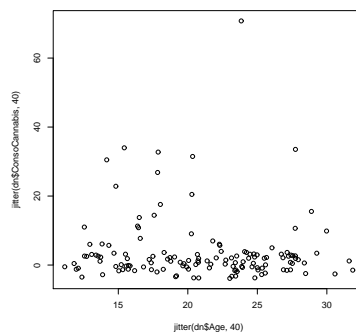
4.1 [CONSO_CANNABIS] VS [AGE]

La variable [AGE] est une variable discrète

- **Représentation graphique** : nuage de point
- **Diagnostic** : une des deux variables doit suivre une loi normale.
- **Test paramétrique** : r de Pearson
- **Test non paramétrique** : r de Spearman

4.1.1 Graphe

Pour éviter la superposition des points aux coordonnées identiques, une légère dispersion (instruction `jitter` sous R) est ajoutée à la variable [AGE].



4.1.2 Diagnostic

La variable [AGE] suit approximativement une loi normale, on applique le r de Pearson.

4.1.3 Test

- H_0 : Il n'y a pas de différence selon l'âge, les jeunes et les moins jeunes prennent du cannabis dans les mêmes proportions.
- r de Pearson : 0.142
- Degrés de liberté : $df = 135$
- p : 0.887

P est supérieur à 5%, on accepte H_0 :

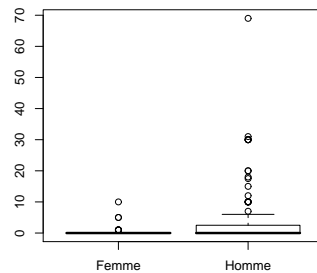
**Il n'y a pas de lien entre les variables [AGE] et [CONSO CANNABIS],
les jeunes et moins jeunes consomment autant de cannabis.**

4.2 [CONSO CANNABIS] VS [SEXE]

La variable [SEXE] est une variable nominale à deux classes.

- **Représentation graphique** : boîtes à moustaches.
- **Diagnostic** : la variable continue doit être normale ou d'effectif supérieur à 30.
De plus, les variances des groupes doivent être égales.
- **Test paramétrique** : T de Student
- **Test non paramétrique** : Test des rangs de Wilcoxon

4.2.1 Graphe



4.2.2 Diagnostic

- **Groupe Homme** moyenne : 4.66 ; variance 10.67
- **Groupe Femme** moyenne : 0.59 ; variance 1.85
- Les variances sont significativement différentes : on applique tout de même le test T de Student.

4.2.3 Test

- H_0 : Il n'y a pas de différence entre les sexes, les hommes et les femmes consomment autant de Cannabis.
- T : -3.59
- Degré de liberté, DDL : 106.08
- p : 0.00050

P est inférieur à 5%, on rejette H_0 :

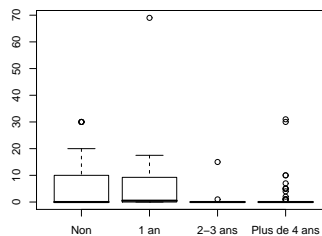
Il y a un lien entre les variables [SEXE] et [CONSO CANNABIS], les hommes consomment significativement plus de cannabis que les femmes.

4.3 [CONSO CANNABIS] VS [SPORT CLUB]

La variable [SPORT CLUB] est une nominale ordonnée. Nous choisissons de la traiter comme une variable discrète.

- **Représentation graphique** : boîtes à moustaches juxtaposées.
- **Diagnostic** : Au moins une des deux variables doit suivre une loi normale
- **Test paramétrique** : R de Pearson
- **Test non paramétrique** : R de Spearman

4.3.1 Graphe



4.3.2 Diagnostic

Les variables [CONSO CANNABIS] et [SPORT CLUB] ne suivent pas de loi normale, on applique le R de Spearman.

4.3.3 Test

- H_0 : il n'y a pas de différence selon la pratique ou non en club par rapport à la consommation de cannabis, les étudiants inscrits ou non en club consomment autant de cannabis.
- $\rho = -0.22$
- $p : 0.011$

P est inférieur à 5%, on rejette H_0 :

Il y a un lien entre les variables [SPORT CLUB] et [CONSO CANNABIS], les gens inscrits depuis plusieurs années dans un club consomment moins de cannabis.

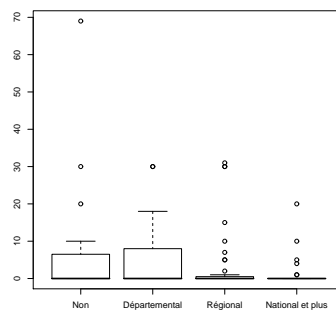
Le même test a été effectué sans la valeur (69) de la variable [CONSO CANNABIS], nous conservons un lien entre les deux variables.

4.4 [CONSO_CANNABIS] VS [SPORTCOMPET]

La variable [SPORTCOMPET] est une nominale ordonnée. Nous choisissons de la traiter comme une variable discrète.

- **Représentation graphique** : boîtes à moustaches juxtaposées.
- **Diagnostic** : Au moins une des deux variables doit suivre une loi normale
- **Test paramétrique** : R de Pearson
- **Test non paramétrique** : R de Spearman

4.4.1 Graphe



4.4.2 Diagnostic

La variable [SPORTCOMPET] suit approximativement une loi normale, on applique le R de Pearson.

4.4.3 Test

- H_0 : il n'y a pas de différence selon le niveau de pratique par rapport à la consommation de cannabis, les étudiants qui font des compétitions consomment autant de cannabis que les autres.
- r de Pearson : -2.59
- Degrés de liberté : $df = 135$
- p : 0.010

P est inférieur à 5%, on rejette H_0 :

Il y a un lien entre les variables [SPORTCOMPET] et [CONSO_CANNABIS], les étudiants qui font des compétitions consomment moins de cannabis

Nous avons effectué le test de Pearson en enlevant la valeur (69) de la variable [CONSO_CANNABIS] :

- r de Pearson : -1.87
- Degrés de liberté : $df = 134$
- p : 0.06 La valeur (69) de la variable [CONSO_CANNABIS] influence le résultat obtenu.

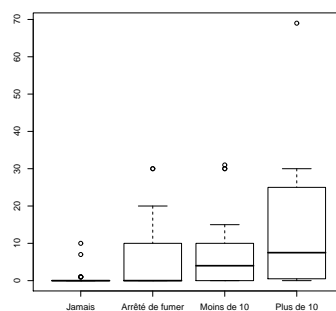
Conclusion : Il y a un lien entre les variables [CONSO_CANNABIS] et [SPORT_COMPET], cependant le résultat demande confirmation.

4.5 [CONSO_CANNABIS] vs [CONSO_TABAC]

La variable [CONSO_TABAC] est une nominale ordonnée. Nous choisissons de la traiter comme une variable discrète.

- **Représentation graphique** : boîtes à moustaches juxtaposées.
- **Diagnostic** : Au moins une des deux variables doit suivre une loi normale
- **Test paramétrique** : R de Pearson
- **Test non paramétrique** : R de Spearman

4.5.1 Graphe



4.5.2 Diagnostic

Les variables [CONSO_CANNABIS] et [CONSO_TABAC] ne suivent pas de loi normale, on applique le R de Spearman.

4.5.3 Test

- H_0 : il n'y a pas de lien entre la consommation de tabac et la consommation de cannabis, les fumeurs de tabac fument autant de cannabis que les autres.
- S : 163120.7
- $\rho = 0.62$
- p : $7.159e-16$

P est inférieure a 5%, on rejette H0 :

Il y a un lien entre les variables [CONSO TABAC] et [CONSO CANNABIS], les fumeurs de tabac fument plus de cannabis.

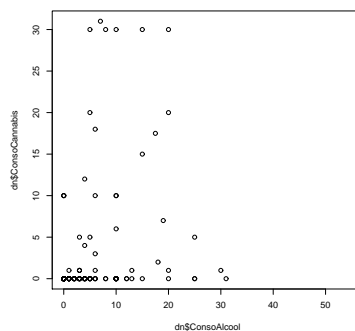
Le même test a été effectué sans la valeur (69) de la variable [CONSO CANNABIS], nous conservons un lien entre les deux variables.

4.6 [CONSO CANNABIS] VS [CONSO ALCOOL]

La variable [CONSO ALCOOL] est une variable continue.

- **Représentation graphique** : nuage de point
- **Diagnostic** : une des deux variables doit suivre une loi normale.
- **Test paramétrique** : r de Pearson
- **Test non paramétrique** : r de Spearman

4.6.1 Graphe



4.6.2 Diagnostic

Les variables [CONSO CANNABIS] et [CONSO ALCOOL] ne suivent pas de loi normale, on applique le r de Spearman.

4.6.3 Test

- H0 : Il n'y a pas de différence selon la consommation d'alcool, ceux qui prennent de l'alcool consomment du cannabis dans les mêmes proportions que les autres.
- r de Spearman : 246181.9
- p : 5.925e-07

P est inférieure a 5%, on rejette H0 :

Il y a un lien entre les variables [CONSO CANNABIS] et [CONSO ALCOOL], les étudiants qui consomment de l'alcool consomment plus de cannabis que les autres

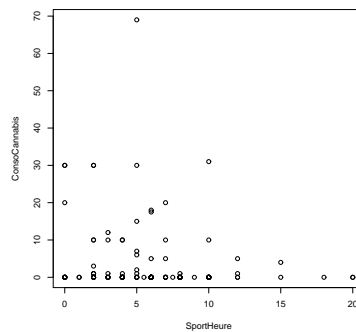
Le même test a été effectué sans la valeur (69) de la variable [CONSO CANNABIS], nous conservons un lien entre les deux variables.

4.7 [CONSO CANNABIS] VS [SPORT HEURE]

La variable [SPORT HEURE] est une variable continue.

- **Représentation graphique** : nuage de point
- **Diagnostic** : une des deux variables doit suivre une loi normale.
- **Test paramétrique** : r de Pearson
- **Test non paramétrique** : r de Spearman

4.7.1 Graphe



4.7.2 Diagnostic

Les variables [CONSO CANNABIS] et [SPORT HEURE] ne suivent pas de loi normale, on applique le r de Spearman.

4.7.3 Test

- H_0 : Il n'y a pas de différence selon le nombre d'heure de sport pratiqué, ceux qui font beaucoup de sport consomment autant de cannabis que les autres.
- r de Spearman : 487392.6
- p : 0.11

P est supérieur à 5%, on accepte H_0 :

Il n'y a pas de lien entre les variables [CONSO CANNABIS] et [SPORT HEURE], les étudiants qui font du sport consomment autant de cannabis que les autres

5 Construction d'un modèle pour la variable [CONSO CANNABIS]

5.1 Détail du modèle

Pour leur intégration a un modèle, les variables [SPORTCLUB], [SPORTCOMPET] et [CONSO TABAC] sont transformées en variables discrètes. Le modèle initial est :

```
ConsoCannabis~Sexe+Age+SportHeure+SportClub+SportCompet+ConsoTabac+ ConsoAlcool
```

Le résultat de la régression linéaire est le suivant :

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	7.26944	9.16172	0.793	0.42898	
Age	-0.56494	0.42842	-1.319	0.18964	
SexeHomme	4.01314	1.36766	2.934	0.00396	**
as.integer(SportClub)	-0.82378	0.59714	-1.380	0.17013	
as.integer(SportCompet)	-0.55638	0.75771	-0.734	0.46412	
as.integer(ConsoTabac)	3.55435	0.73049	4.866	3.29e-06	***
ConsoAlcool	0.38926	0.08452	4.605	9.80e-06	***
SportHeure	0.12743	0.19591	0.650	0.51656	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Le modèle ci-dessus nous montre qu'il existe un lien fort entre les variables [CONSO TABAC], [CONSOALCOOL] et [CONSO CANNABIS].

Les résultats sont des résultats bien connus et ne concernent pas notre étude (qui se concentre sur un lien entre les pratiques sportives et la consommation de drogue). Nous décidons donc d'enlever les variables [CONSO TABAC] et [CONSOALCOOL] et de construire un nouveau modèle. La sélection se fait sur le r^2 ajusté.

Le modèle final est :

```
ConsoCannabis~Sexe+SportCompet+Age
```

Le résultat de la régression linéaire est le suivant :

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	12.9406	11.1497	1.161	0.2479	
Age	-0.3727	0.5188	-0.718	0.4738	
SexeHomme	3.3645	1.6724	2.012	0.0463	*
as.integer(SportClub)	-1.4830	0.5845	-2.537	0.0123	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Note : Le même modèle après suppression de la valeur (69) de la variable [CONSO_CANNABIS] donne sensiblement les mêmes résultats.

5.2 Bilan

Le modèle présente les variables qui ont un impact sur la consommation de cannabis après ajustement sur les autres variables¹

- Le [SEXE] a une influence sur la consommation de cannabis. En moyenne, la consommation de cannabis des garçons est supérieur de 3.4 points à celui des filles.
- Nous avons classé la variable [SPORTCLUB] en quatre catégories ("non", "1an", "2-3ans" et "Plus de 4ans"); chaque fois que l'on augmente de catégorie, la consommation de cannabis baisse de 1.5 points. En conséquence notre hypothèse initiale est vérifiée.

1. En moyenne, la consommation de cannabis des garçons est plus élevée que celle des filles.
2. Plus l'engagement dans un club est fort, moins les étudiants consomment de cannabis.

6 Analyse bivariée de [CONSOALCOOL]

Nous nous intéressons maintenant à la variable [CONSOALCOOL]. D'une manière exploratoire, nous allons examiner les variables avec lesquelles elle est liée.

6.1 [CONSOALCOOL] VS [AGE]

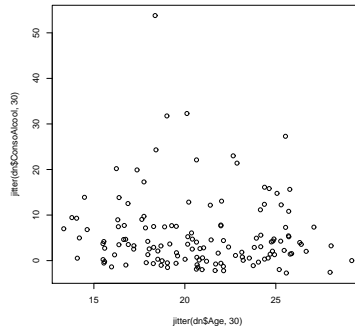
La variable [AGE] est une variable discrète

- **Représentation graphique** : nuage de point
- **Diagnostic** : une des deux variables doit suivre une loi normale.
- **Test paramétrique** : r de Pearson
- **Test non paramétrique** : r de Spearman

6.1.1 Graphe

Pour éviter la superposition des points aux coordonnées identiques, une légère dispersion (instruction `jitter` sous R) est ajoutée à la variable [AGE].

¹Ce type d'ajustement permet d'éviter les facteurs de confusion. Par exemple, s'il est avéré que les femmes consomment moins de cannabis, si celles-ci sont majoritaires au sein des compétitions de niveau régional, ce niveau présenterait l'apparence d'un niveau où l'on consomme peu. Cette illusion constituerait un biais. L'ajustement de la variable SportCompetition par la variable Sexe permet de comparer les niveaux de compétition à sexe constant et donc corrige ce biais potentiel.



6.1.2 Diagnostic

La variable [AGE] suit approximativement une loi normale, on applique le r de Pearson.

6.1.3 Test

- H0 : Il n'y a pas de différence selon l'âge, les jeunes et les moins jeunes consomment de l'alcool dans les mêmes proportions.
- r de Pearson : 0.53
- Degrés de liberté : $df = 134$
- p : 0.597

P est supérieur à 5%, on accepte H0 :

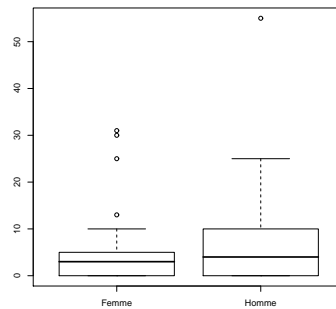
**Il n'y a pas de lien entre les variables [AGE] et [CONSOALCOOL],
les jeunes et moins jeunes consomment autant d'alcool.**

6.2 [CONSOALCOOL] vs [SEXE]

La variable [SEXE] est une variable nominale à deux classes.

- **Représentation graphique** : boîtes à moustaches.
- **Diagnostic** : la variable continue doit être normale ou d'effectif supérieur à 30. De plus, les variances des groupes doivent être égales.
- **Test paramétrique** : T de Student
- **Test non paramétrique** : Test des rangs de Wilcoxon

6.2.1 Graphe



6.2.2 Diagnostic

- **Groupe Homme** moyenne : 6.14 ; variance 66.63
- **Groupe Femme** moyenne : 4.93 ; variance 56.60
- Les variances ne sont pas significativement différentes et les effectifs des groupes sont supérieur à 30 : on applique le Test T de Student.

6.2.3 Test

- H0 : Il n'y a pas de différence entre les sexes, les hommes et les femmes consomment autant d'Alcool.
- T de Student : -0.82
- Degrés de liberté : df = 134
- p : 0.413

P est supérieur a 5%, on accepte H0 :

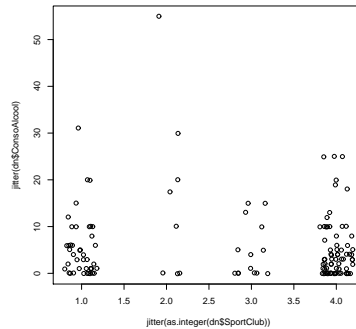
**Il n'y a pas de lien entre les variables [SEXE] et [CONSOALCOOL],
les hommes ne consomment pas significativement plus d'alcool que les femmes.**

6.3 [CONSOALCOOL] VS [SPORTCLUB]

La variable [SPORTCLUB] est une nominale ordonnée. Nous choisissons de la traiter comme une variable discrète.

- **Représentation graphique** : boîtes à moustaches juxtaposées.
- **Diagnostic** : Au moins une des deux variables doit suivre une loi normale
- **Test paramétrique** : R de Pearson
- **Test non paramétrique** : R de Spearman

6.3.1 Graphe



6.3.2 Diagnostic

Les variables [CONSOALCOOL] et [SPORTCLUB] ne suivent pas de loi normale, on applique le R de Spearman.

6.3.3 Test

- H_0 : il n'y a pas de différence selon la pratique ou non en club par rapport à la consommation d'alcool, les étudiants inscrits ou non en club consomment autant d'alcool.
- $\rho = -0.067$
- $p : 0.44$

P est supérieur à 5%, on accepte H_0 :

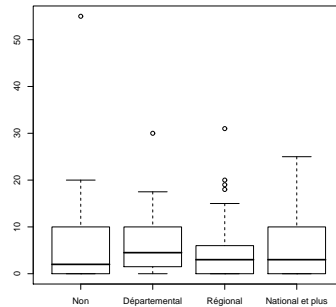
Il n'y a pas de lien entre les variables [SPORTCLUB] et [CONSOALCOOL], les gens inscrits depuis plusieurs années dans un club consomment autant d'alcool que les autres.

6.4 [CONSOALCOOL] VS [SPORTCOMPET]

La variable [SPORTCOMPET] est une nominale ordonnée. Nous choisissons de la traiter comme une variable discrète.

- **Représentation graphique** : boîtes à moustaches juxtaposées.
- **Diagnostic** : Au moins une des deux variables doit suivre une loi normale
- **Test paramétrique** : R de Pearson
- **Test non paramétrique** : R de Spearman

6.4.1 Graphe



6.4.2 Diagnostic

La variable [SPORTCOMPET] suit approximativement une loi normale, on applique le R de Pearson.

6.4.3 Test

- H_0 : il n'y a pas de différence selon le niveau de pratique par rapport à la consommation d'alcool, les étudiants qui font des compétitions consomment autant d'alcool que les autres.
- r de Pearson : -1.22
- Degrés de liberté : $df = 134$
- p : 0.22

P est supérieur à 5%, on accepte H_0 :

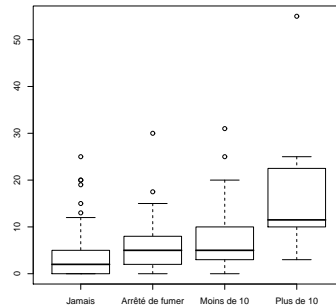
Il n'y a pas de lien entre les variables [SPORTCOMPET] et [CONSOALCOOL], les étudiants qui font des compétitions consomment autant d'alcool que les autres

6.5 [CONSOALCOOL] VS [CONSO TABAC]

La variable [CONSO TABAC] est une nominale ordonnée. Nous choisissons de la traiter comme une variable discrète.

- **Représentation graphique** : boîtes à moustaches juxtaposées.
- **Diagnostic** : Au moins une des deux variables doit suivre une loi normale
- **Test paramétrique** : R de Pearson
- **Test non paramétrique** : R de Spearman

6.5.1 Graphe



6.5.2 Diagnostic

Les variables [CONSOALCOOL] et [CONSO TABAC] ne suivent pas de loi normale, on applique le R de Spearman.

6.5.3 Test

- H_0 : il n'y a pas de lien entre la consommation de tabac et la consommation d'alcool, les fumeurs de tabac consomment autant d'alcool que les autres.
- $\rho = 0.37$
- $p : 8.709e-06$

P est inférieur a 5%, on rejette H_0 :

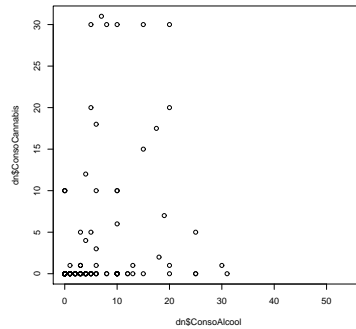
Il y a un lien entre les variables [CONSO TABAC] et [CONSOALCOOL], les fumeurs de tabac consomment plus d'alcool que les autres.

6.6 [CONSOALCOOL] vs [CONSO CANNABIS]

La variable [CONSOALCOOL] est une variable continue.

- **Représentation graphique** : nuage de point
- **Diagnostic** : une des deux variables doit suivre une loi normale.
- **Test paramétrique** : r de Pearson
- **Test non paramétrique** : r de Spearman

6.6.1 Graphe



6.6.2 Diagnostic

Les variables [CONSO CANNABIS] et [CONSO ALCOOL] ne suivent pas de loi normale, on applique le r de Spearman.

6.6.3 Test

- H_0 : Il n'y a pas de différence selon la consommation d'alcool, ceux qui prennent de l'alcool consomment du cannabis dans les mêmes proportions que les autres.
- r de Spearman : 246181.9
- p : 5.925e-07

P est inférieur à 5%, on rejette H_0 :

Il y a un lien entre les variables [CONSO CANNABIS] et [CONSO ALCOOL], les étudiants qui consomment de l'alcool consomment plus de cannabis que les autres

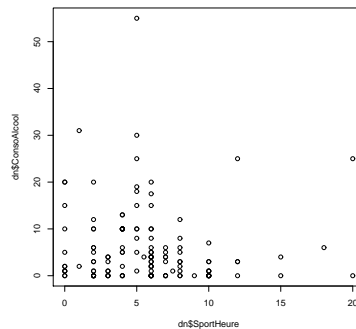
Le même test a été effectué sans la valeur (69) de la variable [CONSO CANNABIS], nous conservons un lien entre les deux variables.

6.7 [CONSO ALCOOL] VS [SPORT HEURE]

La variable [SPORT HEURE] est une variable continue.

- **Représentation graphique** : nuage de point
- **Diagnostic** : une des deux variables doit suivre une loi normale.
- **Test paramétrique** : r de Pearson
- **Test non paramétrique** : r de Spearman

6.7.1 Graphe



6.7.2 Diagnostic

Les variables [CONSOALCOOL] et [SPORTHEURE] ne suivent pas de loi normale, on applique le r de Spearman.

6.7.3 Test

- H0 : Il n'y a pas de différence selon le nombre d'heure de sport pratiqué, ceux qui font beaucoup de sport consomment autant d'alcool que les autres.
- r de Spearman : 477969.4
- p : 0.10

P est supérieur à 5%, on accepte H0 :

Il n'y a pas de lien entre les variables [CONSOALCOOL] et [SPORTHEURE], les étudiants qui font du sport consomment autant d'alcool que les autres

6.8 Modèle

Nous ne présenterons pas de modèle pour la variable [CONSOALCOOL] puisque l'analyse bivariée ne détecte aucun lien. Le manque de puissance de notre étude pourrait en être la cause.

7 Conclusion générale

7.1 Réponses aux questions initiales

En se basant sur les résultats apportés par le modèle pour la variable [CONSOCANNABIS] (résultats les plus fiables puisqu'ils prennent en compte les ajustements sur les autres variables) et sur les résultats de l'analyse bivariée pour la variable [CONSOALCOOL], on peut répondre à nos questions initiales :

- Les hommes consomment plus de cannabis que les femmes

- Plus la pratique au sein d'un club est forte, moins les étudiants consomment de cannabis
- La pratique sportive n'a pas d'impact sur la consommation d'alcool

7.2 Limites de l'étude

Même s'ils présentent une certaine cohérence, ces résultats sont toutefois à relativiser. En particulier :

- La consommation de drogue étant illégale, la désirabilité sociale peut avoir biaisé les résultats.
- Les étudiants enquêteurs ne sont pas des professionnels de l'enquête et les conditions de passation des questionnaires n'étaient parfois pas optimales.
- Aucun contrôle n'a été fait concernant l'échantillonnage ou le choix des sujets.
- Enfin, le choix de faire cette enquête a été une décision collective. Les étudiants de L2 ont eu à voter² pour choisir entre un projet "intéressant mais demandant un investissement" et utiliser une base de donnée "moins intéressante, mais déjà construite". Néanmoins, certains d'entre eux ont pu considérer cette enquête comme une contrainte et ont pu avoir la tentation de remplir eux-même les questionnaires ou de les faire remplir par leurs copains.

²Vote a main levé hors présence enseignante