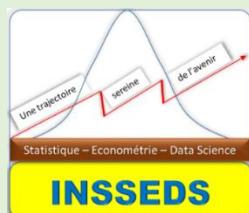


MINISTERE DE L'ENSEIGNEMENT SUPERIEUR  
ET DE RECHERCHE SCIENTIFIQUE



INSTITUT SUPERIEUR DE STATISTIQUE  
D'ECONOMETRIE ET DE DATA SCIENCE

REPUBLIQUE DE COTE D'IVOIRE



MASTER 2

**STATISTIQUE – ECONOMETRIE – DATA SCIENCE**

### **MINI PROJET**

### **STATISTIQUE DESCRIPTIVE**

**ETUDE DES FACTEURS INFLUENCANT LA DEPRESSION CHEZ  
LES ETUDIANTS**

**ANNEE ACADEMIQUE**

**2024-2025**

ETUDIANT

NOM : SEKA

PRENOM : ISRAEL ASSI JUNIOR  
CHRIST

ENSEIGNANT-ENCADREUR

**AKPOSSO DIDIER  
MARTIAL**

## Avant-propos

La santé mentale des étudiants est un enjeu majeur qui mérite une attention particulière, notamment en raison des nombreuses pressions auxquelles ils sont confrontés au cours de leur parcours académique. Ce mini-projet s'inscrit dans cette problématique en explorant les facteurs influençant la dépression chez les étudiants, à travers une approche statistique et analytique rigoureuse.

Ce travail a été réalisé dans le cadre de mon cursus en **Statistique, Économétrie et Data Science**, et vise à apporter une contribution significative à la compréhension des éléments déclencheurs de la dépression dans le milieu universitaire. En utilisant des méthodes statistiques avancées, l'objectif est d'identifier les principales variables en corrélation avec la dépression et de formuler des recommandations pour mieux prévenir ce phénomène.

Je tiens à exprimer ma gratitude à mon encadrant, **M. Akposso Didier Martial**, pour son accompagnement et ses conseils précieux tout au long de cette étude. Je remercie également **l'Institut Supérieur de Statistique, d'Économétrie et de Data Science** pour les ressources mises à disposition et le cadre propice à la réalisation de ce travail.

Que ce projet puisse servir d'outil de réflexion et d'action pour améliorer le bien-être des étudiants et encourager une prise de conscience accrue sur l'importance de la santé mentale dans le milieu académique.

## Table des matières

Avant-propos .....	1
Introduction .....	3
1 STRATEGIE DE TRAITEMENT DE DONNEE .....	4
1.1 Enonciation du dictionnaire de donnée .....	4
1.2 Enonciation du Jeu de donnée .....	4
1.3 Analyse de donnée .....	5
1.3.1 détection et apurement des valeurs manquantes .....	5
2 ANAYSE NAVIGATRICE DES DONNEES.....	8
2.1 Analyse univarié.....	8
2.1.1 Paramètres statistiques .....	8
2.1.2 Représentation de donnée sous forme de donnée numérique .....	11
2.2 Analyse bivariée.....	13
2.2.1 Analyse relationnel entre la variable qualitative et la variable quantitative.....	13
2.2.2 Analyse relationnelle entre deux variables qualitatives .....	16
3 Recommandations .....	27
Conclusion.....	27
ANNEXE .....	27

## Introduction

La dépression est un trouble psychologique majeur qui affecte des millions de personnes à travers le monde. Elle se manifeste par une tristesse persistante, une perte d'intérêt pour les activités quotidiennes, une diminution de l'énergie, des troubles du sommeil et une altération des fonctions cognitives. Chez les étudiants, ce phénomène prend une ampleur particulière en raison des multiples pressions auxquelles ils sont confrontés. En effet, le passage dans l'enseignement supérieur représente une période charnière marquée par de nombreux défis, tant académiques que personnels. L'adaptation à un nouvel environnement, la gestion du stress lié aux performances académiques, la pression sociale, les contraintes financières et parfois l'éloignement familial peuvent être des facteurs de risque majeurs favorisant l'apparition de troubles dépressifs.

Plusieurs études ont démontré que la prévalence de la dépression chez les étudiants est en constante augmentation, notamment en raison des exigences accrues du monde universitaire et des transformations socio-économiques contemporaines. La question de la santé mentale des étudiants est ainsi devenue un enjeu de santé publique majeur, nécessitant une analyse approfondie afin de mieux comprendre les facteurs de vulnérabilité et d'identifier les populations les plus à risque. Cette compréhension est essentielle pour élaborer des stratégies d'intervention efficaces, adaptées aux besoins spécifiques des étudiants et visant à prévenir la détérioration de leur bien-être mental.

Dans cette perspective, notre étude repose sur l'analyse d'un ensemble de données contenant des informations variées sur les étudiants, incluant leurs caractéristiques démographiques, leurs habitudes de vie, leur situation académique et professionnelle ainsi que leur état de santé mentale. L'objectif est d'examiner les corrélations entre ces différentes variables afin d'identifier les principaux facteurs contribuant au développement de la dépression. Plus spécifiquement, cette étude vise à répondre aux questions suivantes : Quels sont les déterminants majeurs de la dépression chez les étudiants ? Existe-t-il des profils à risque plus élevés ? Comment les facteurs environnementaux et personnels interagissent-ils pour influencer la santé mentale des étudiants ?

L'analyse de ces données permettra d'apporter une contribution significative aux recherches en psychologie, en éducation et en sciences sociales, tout en fournissant des recommandations pour améliorer la prise en charge et le soutien psychologique des étudiants en difficulté. En mettant en lumière les facteurs de vulnérabilité et en identifiant des indicateurs précoces, cette étude pourra également aider à la mise en place de dispositifs de prévention et d'accompagnement plus efficaces au sein des institutions académiques et des politiques publiques.

## 1 STRATEGIE DE TRAITEMENT DE DONNEE

### 1.1 Enonciation du dictionnaire de donnée

NOM DE LA VARIABLE	NATURE	DESCRIPTION	MODALITES
sexe	qualitative	sexe de l'Homme	male / femelle
age	qualitative	Age de l'étudiant	num
ville	qualitative	Ville de résidence	ex. : bangalor , pune
profession	qualitative	étudiant à temps plein ou autre activité	ex. : architect , chef
pression académique	qualitative	Pression académique	. Faible . Modéré . Elevé
pression liée au travail	qualitative	pression liée au travail	. Faible . Modéré . Forte
moyennes note	quantitative	Moyenne générale des notes	num
satisfaction etude	qualitative	Satisfaction par rapport aux études	. Faible . Modéré . Elevé
satisfaction travail	qualitative	Satisfaction par rapport au travail	. Faible . Modéré . Elevé
duree sommeil	qualitative	Durée du sommeil	5-6 heures , moins de 5 heures
habitude alimentaire	qualitative	Habitudes alimentaires	saine modéré
diplôme suivi	qualitative	Diplôme suivi ou obtenu	ex. : BSc, M.Tech
nombre heure travail etude	quantitative	Nombre d'heures de travail ou d'études par jour	num
stress financier	quantitative	Stress financier	num
antecedent familiaux maladie mentale	qualitative	Antécédents familiaux de maladies mentales	Oui / Non
depression	qualitative	depression	Oui / Non
pensée suicidaire	qualitative	pensée suicidaire	Oui / Non

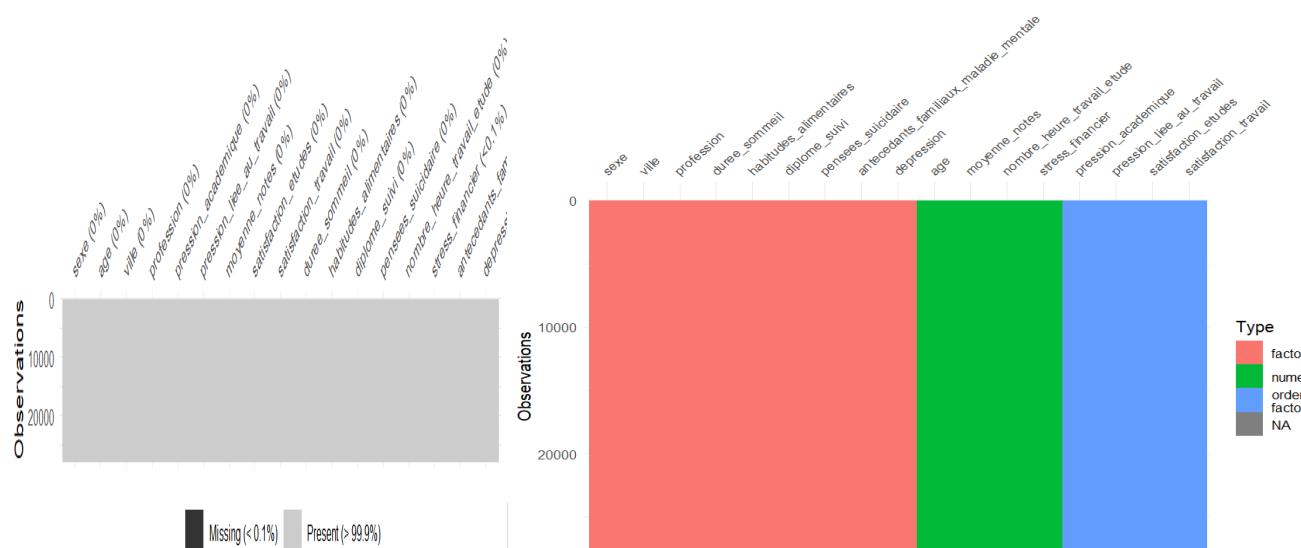
### 1.2 Enonciation du Jeu de donnée

La préparation des données englobe l'ensemble des étapes nécessaires pour transformer des données brutes en informations exploitables et de qualité, prêtes à être analysées. Ce processus inclut la collecte, le nettoyage, la transformation, l'intégration et l'organisation des données, afin de garantir leur cohérence, leur précision et leur pertinence pour l'analyse. Une préparation rigoureuse est essentielle pour minimiser les erreurs et maximiser la valeur des données, facilitant ainsi les analyses ultérieures et permettant des résultats plus fiables et pertinents.

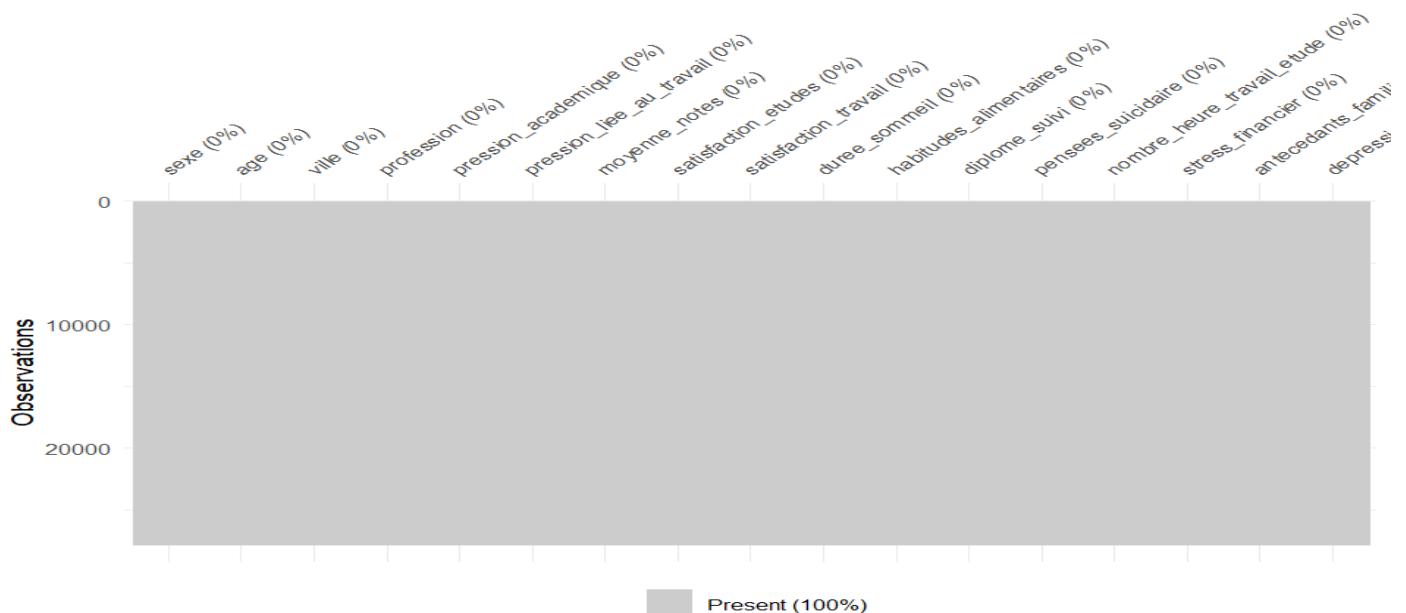
age	ville	profession	pression_academique	pression_liee_au_travail	moyenne_notes	satisfaction_etudes	satisfaction_travail	duree_sommeil
<dbl>	<fctr>	<fctr>	<ord>	<ord>	<dbl>	<ord>	<ord>	<fctr>
1	33	Visakhapatnam	Student	Élevée	Faible	8.97	Faible	Faible 5-6 hours
2	24	Bangalore	Student	Faible	Faible	5.90	Élevée	Faible 5-6 hours
3	31	Srinagar	Student	Modérée	Faible	7.03	Élevée	Faible Less than 5 hours
4	28	Varanasi	Student	Modérée	Faible	5.59	Faible	Faible 7-8 hours
5	25	Jaipur	Student	Élevée	Faible	8.13	Modérée	Faible 5-6 hours
6	29	Pune	Student	Faible	Faible	5.70	Modérée	Faible Less than 5 hours
7	30	Thane	Student	Modérée	Faible	9.54	Élevée	Faible 7-8 hours
8	30	Chennai	Student	Faible	Faible	8.04	Élevée	Faible Less than 5 hours
9	28	Nagpur	Student	Modérée	Faible	9.79	Faible	Faible 7-8 hours
10	31	Nashik	Student	Faible	Faible	8.38	Modérée	Faible Less than 5 hours

## 1.3 Analyse de donnée

### 1.3.1 détection et apurement des valeurs manquantes



Nous pouvons remarquer d'après le graphique qu'il existe une variable qui contient de valeur manquante. Raison pour laquelle nous allons supprimer la valeur manquante.



Nous pouvons remarquer d'après le graphique qu'aucune variable ne contient aucune de valeurs manquantes ni de case vide.

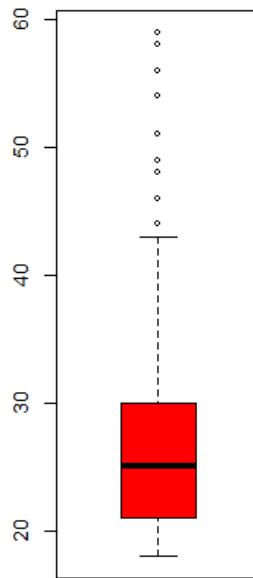
- **Traitement des doublons**

Ici nous constatons que nous n'avons pas de doublons

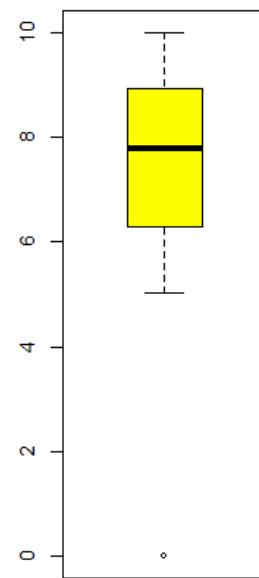
- **Visualisation des données extrêmes ou aberrantes**

Une donnée extrême ou aberrante est une valeur qui est très différente des autres dans un ensemble de données. Elle est soit beaucoup plus grande, soit beaucoup plus petite que les autres valeurs. la technique de boxplot est prépondérant pour la visualisation des données .

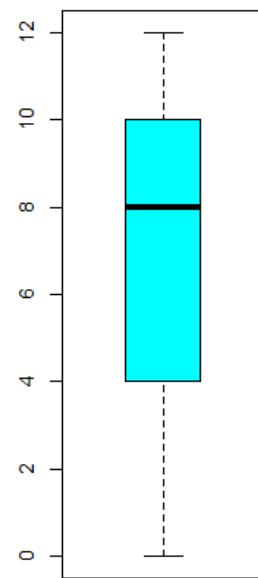
1- AGE



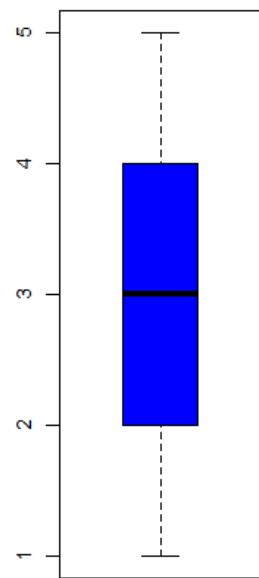
2-MOYENNE NOTE



3-NBRE HEURE



4-STRESS FINANC



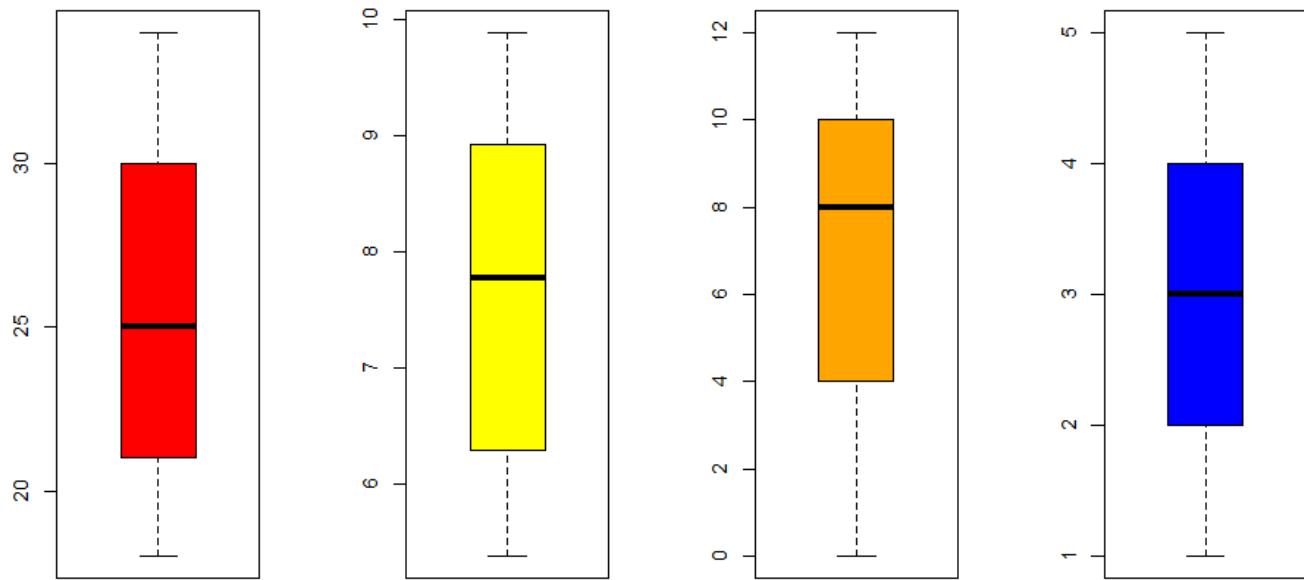
Source : auteur, sortie du logiciel R studio à partir des données secondaires auprès de l'INSSEDS

Notons la présence de donnée extrême dans notre data frame (jeu de donnée) lorsqu'on visualise.  
Nous passerons à la visualisation pour traiter ces valeurs aberrantes

- **Traitement de donnée extrême ou aberrante**

La technique de Winzorisation la technique été utilisée pour traiter ces valeurs aberrantes en les ramenant à la borne (inférieure et supérieure).

Comme nous avons la présence de valeur extrême nous allons procéder au traitement des données extrêmes puis revisualiser pour voir si les données ont été traité



Source : auteur, sortie du logiciel R studio à partir des données secondaires auprès de l'inssts

Les données ont été traité. Il n'y a plus de donnée extrême ou aberrantes

## 2 ANAYSE NAVIGATRICE DES DONNEES

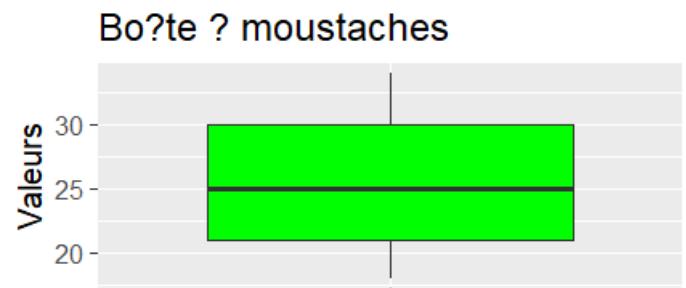
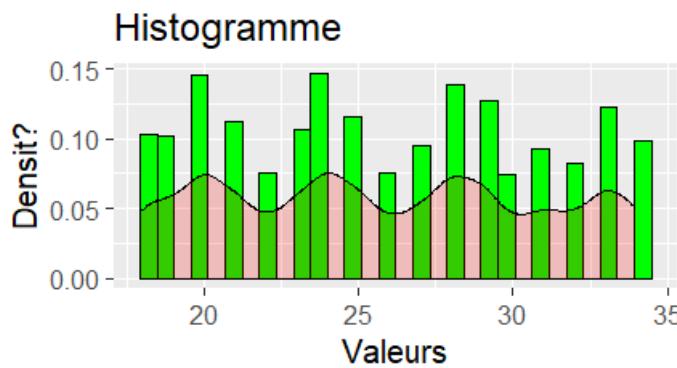
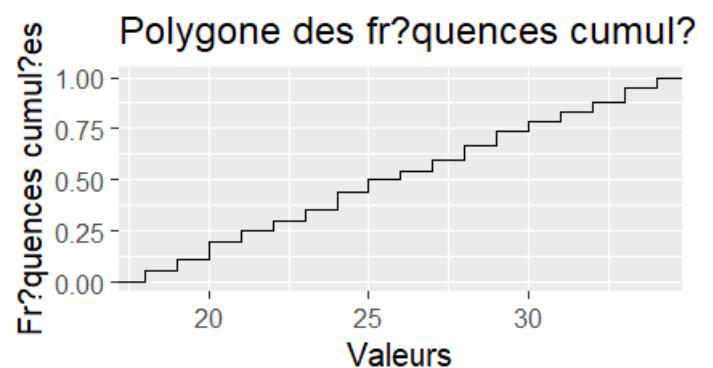
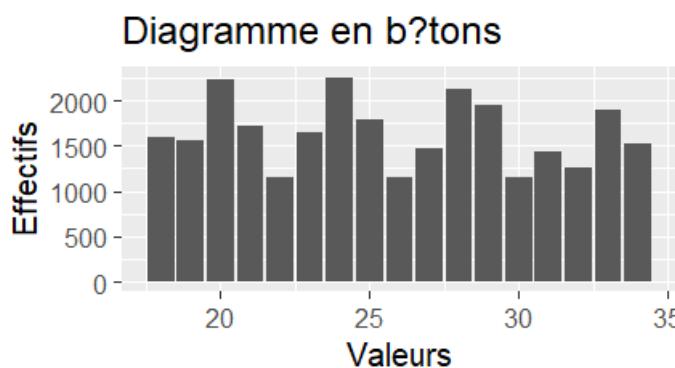
### 2.1 Analyse univarié

Cette partie nous permettra sans doute d'avoir certaines informations sur chaque variable et aussi de déterminer les tendances générales de chacune des variables.

#### 2.1.1 Paramètres statistiques

Dans cette partie, nous allons aborder une analyse totalement univariée des variables. Nous allons cependant décrire quelques principales caractéristiques de nos variables quantitatives. Ce qui nous permettra d'avoir des détails plus instructifs avant d'entamer la modélisation statistique. Tableau des résumés numériques de nos variables. Ainsi que des graphiques.

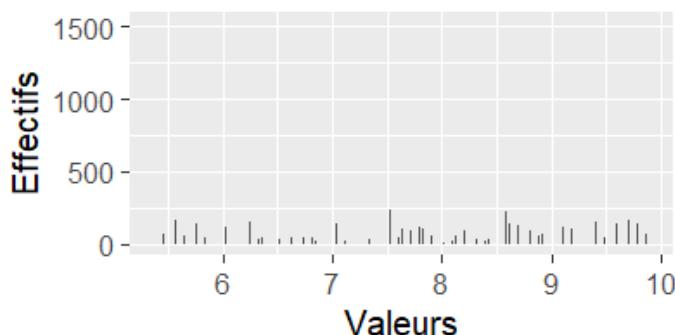
- LA VARIABLE AGE



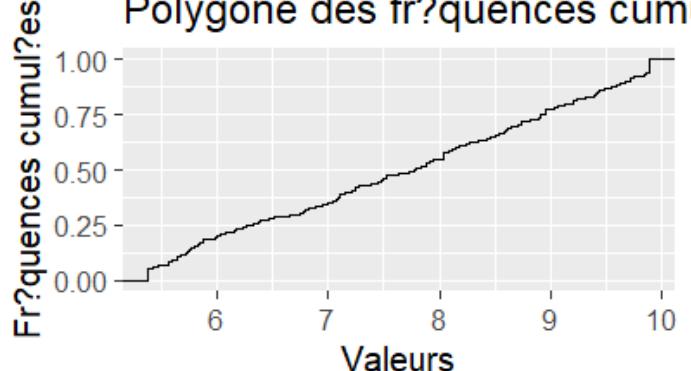
- La répartition de l'âge est **relativement équilibrée** avec une légère concentration autour de 24-25 ans.
- Il existe une certaine **hétérogénéité** dans les âges, mais la majorité reste dans la tranche **18-30 ans**.
- **Coefficient d'asymétrie (Skewness)** : **0.066** → **Légère asymétrie à droite**, ce qui signifie qu'il y a **quelques étudiants plus âgés**.
- **Coefficient d'aplatissement (Kurtosis)** : **1.81** → **Distribution platikurtique**, donc les âges sont **plus dispersés** autour de la moyenne.
- La distribution est **équilibrée**, sans forte concentration d'un côté.
- La présence de quelques étudiants **plus âgés** (vers 30-34 ans) **étend légèrement la distribution**.

- LA VARIABLE MOYENNE NOTE

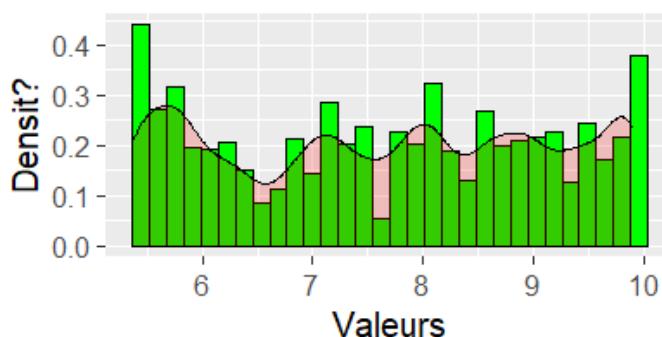
Diagramme en bâtons



Polygone des fréquences cumulées



Histogramme



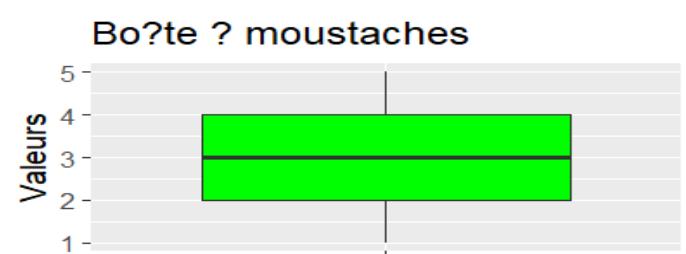
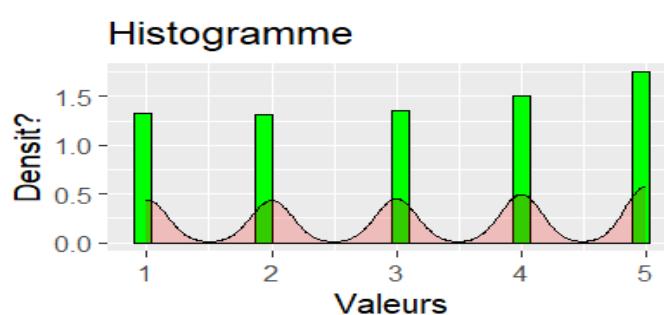
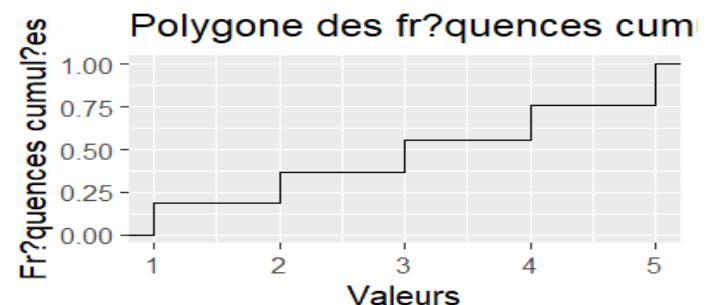
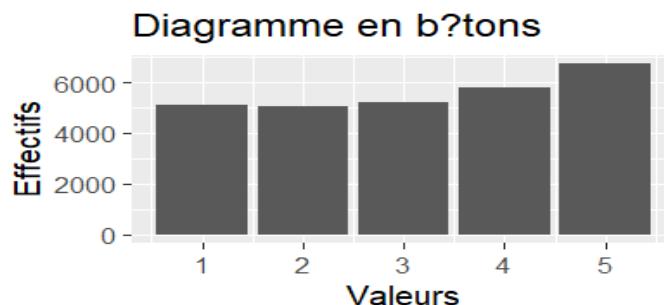
Boîte à moustaches



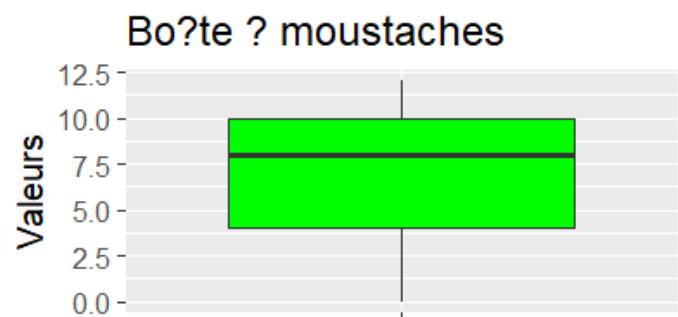
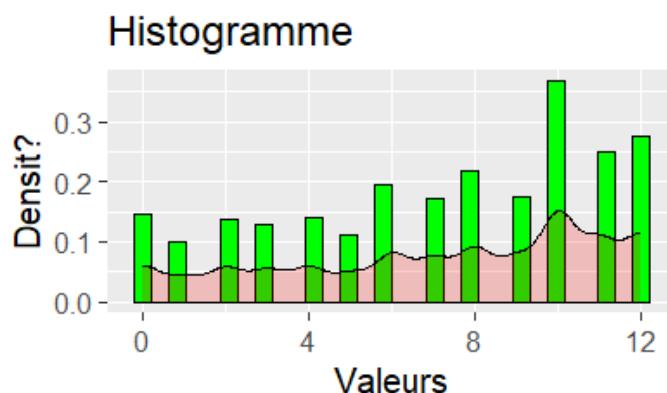
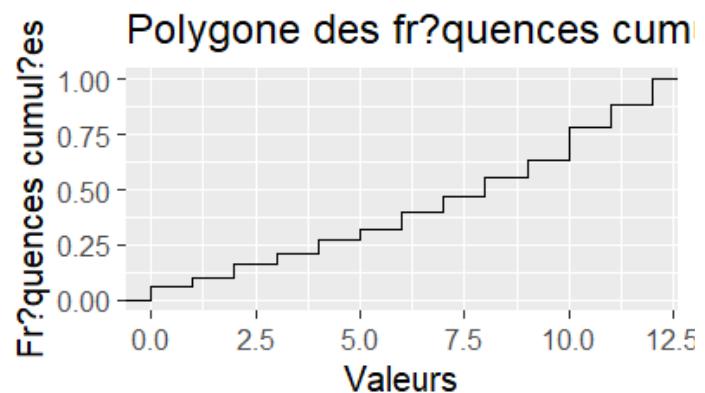
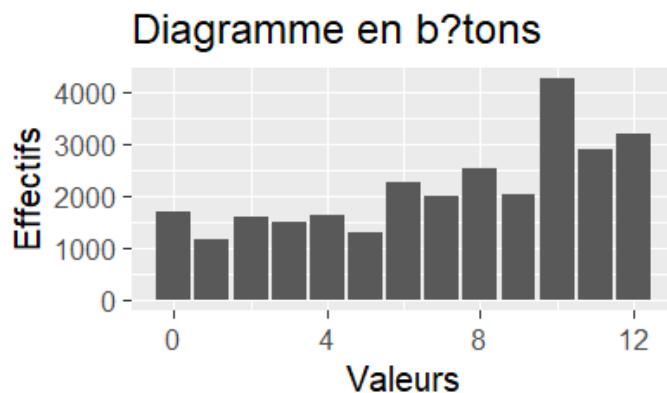
Source : auteur, sortie du logiciel R studio à partir des données secondaires auprès de l'INSSEDS

- La note la plus fréquente est autour de **5.69**, ce qui signifie que certains étudiants sont juste au-dessus de la moyenne minimale.
- La majorité des étudiants ont des **résultats satisfaisants**.
- **Légère concentration** des notes vers le bas (**mode autour de 5.69**), suggérant que certains étudiants rencontrent des difficultés académiques.
- La **dispersion est faible**, donc peu de différence entre les étudiants.
- Les étudiants sont **globalement performants**, avec **peu de notes extrêmes**.
- **-0.054** → **Légèrement asymétrique à gauche**, ce qui signifie que la majorité des étudiants ont des notes élevées avec quelques notes faibles.
- **Distribution platikurtique**, donc les notes sont **plus dispersées** autour de la moyenne, sans concentration excessive.
- La distribution étalée à gauche signifie que **beaucoup d'étudiants réussissent bien**, mais quelques-uns ont des **notes plus basses**.
- Il n'y a **pas de forte concentration** autour d'une seule note, donc la diversité des résultats est **modérée**.

- LA VARIABLE STRESS FINANCIER



- LA VARIABLE NOMBRE HEURE TRAVAIL ETUDE



Source : auteur, sortie du logiciel R studio à partir des données secondaires auprès de l'INSSEDS

- La plupart des étudiants travaillent ou étudient entre 7 et 10 heures par jour.
- Un groupe important travaille 10 heures, ce qui peut être un facteur de stress élevé.
- Certains étudiants ont un emploi du temps très chargé (10-12 heures).
- Une grande variabilité entre ceux qui travaillent peu et ceux qui travaillent énormément.
- Légèrement asymétrique à gauche, donc une majorité d'étudiants travaille beaucoup, mais quelques-uns travaillent très peu.
- Distribution platikurtique, indiquant une dispersion modérée autour de la moyenne.
- La distribution montre plus d'étudiants ayant un nombre d'heures élevé.
- Certains étudiants travaillent peu voire pas du tout (0 h), ce qui pourrait être lié à des différences de motivation ou de contraintes personnelles

## 2.1.2 Représentation de donnée sous forme de donnée numérique

- PARAMETRE DE TENDANCE CENTRALE**

VARIABLE	MINIMUM	MAXIMUM	MODE	MOYENNE	MEDIANE
AGE	18.00	34.00	24.00	25.81	25.00
MOYENNE NOTES	5.38	9.89	5.68	7.66	7.77
STRESS FINANCIER	1.00	5.00	5.00	3.14	3.00
NOMBRE HEURE TRAVAIL					
ETUDE	0.00	12.00	10.00	7.16	8.00

- PARAMETRE DE DISPERSION**

VARIABLE	VARIANCE	ECART TYPE	COEFFICIENT DE VARIATION
AGE	18.00	34.00	24.00
MOYENNE NOTES	5.38	9.89	5.68
STRESS FINANCIER	1.00	5.00	5.00
NOMBRE HEURE TRAVAIL	0.00	12.00	10.00
ETUDE	0.00	12.00	7.16

AGE	23.72	4.87	18.87
MOYENNE NOTES	2.09	1.45	18.88
STRESS FINANCIER	2.06	1.44	45.78
NOMBRE HEURE TRAVAIL ETUDE	13.75	3.71	51.81

- **PARAMETRE DE FORME**

VARIABLE	SKEWNESS	KURTOSIS
AGE	0.67	1.81
MOYENNE NOTES	-0.05	1.73
STRESS FINANCIER	-0.13	1.67
NOMBRE HEURE TRAVAIL ETUDE	-0.45	2.00

## INTERPRETATION DU PARAMETRE DU TABLEAU STATISTIQUE SIMPLE

- **PARAMETRE DE TENDANCE CENTRALE**

- **Âge** : La moyenne (25.81 ans) et la médiane (25 ans) sont proches, indiquant une distribution assez symétrique. Le mode est à 24 ans, ce qui signifie que cette valeur est la plus fréquente parmi les étudiants.
- **Moyenne des notes** : La moyenne (7.66) et la médiane (7.77) montrent que la performance académique tourne autour de ces valeurs. Un mode à 5.68 suggère qu'un certain nombre d'étudiants ont des notes relativement basses.
- **Stress financier** : La médiane de 3 et la moyenne de 3.14 montrent que le stress financier est modéré pour la plupart des étudiants, mais le mode à 5 indique qu'un grand nombre d'étudiants ressentent un fort stress financier.
- **Nombre d'heures d'étude/travail** : La médiane de 8 heures et la moyenne de 7.16 suggèrent que la majorité des étudiants travaillent ou étudient autour de 7 à 8 heures par jour. Le mode de 10 heures indique que beaucoup d'étudiants dépassent cette moyenne.

- **PARAMETRE DE DISPERSION**

- **Âge** : Un écart-type de 4.87 indique que l'âge des étudiants varie peu autour de la moyenne (25.81 ans). Le coefficient de variation (18.87%) montre une faible dispersion des âges.
- **Moyenne des notes** : L'écart-type de 1.45 indique que les notes sont assez regroupées autour de la moyenne (7.66). Le coefficient de variation (18.88%) confirme une dispersion modérée des performances.
- **Stress financier** : Avec un écart-type de 1.44 et un coefficient de variation de 45.78%, cette variable présente une dispersion assez forte, ce qui signifie que les étudiants ont des niveaux de stress financier très variés.

- **Nombre d'heures d'étude/travail** : L'écart-type de 3.71 et le coefficient de variation de 51.81% indiquent une forte variabilité dans le temps consacré aux études ou au travail. Certains étudiants étudient très peu, d'autres énormément.

- **PARAMETRE DE FORME**

- **Âge (Skewness = 0.67, Kurtosis = 1.81)**

- Une asymétrie légèrement positive (0.67) signifie que la distribution est légèrement décalée vers la droite (plus d'étudiants plus jeunes que la moyenne).
- Une kurtosis de 1.81 indique une distribution légèrement aplatie, donc sans valeurs extrêmes marquées.

- **Moyenne des notes (Skewness = -0.05, Kurtosis = 1.73)**

- L'asymétrie est très proche de zéro (-0.05), ce qui suggère une distribution quasi-symétrique.
- La kurtosis de 1.73 montre une distribution un peu aplatie, ce qui signifie qu'il y a peu de notes extrêmes.

- **Stress financier (Skewness = -0.13, Kurtosis = 1.67)**

- Une très légère asymétrie négative (-0.13) suggère que la majorité des étudiants ressentent un stress financier modéré, avec peu d'extrêmes.
- Une kurtosis de 1.67 indique une distribution légèrement aplatie, ce qui signifie que les valeurs extrêmes ne sont pas très fréquentes.

- **Nombre d'heures d'étude/travail (Skewness = -0.45, Kurtosis = 2.00)**

- Une asymétrie négative (-0.45) indique que la distribution est légèrement décalée vers la gauche, ce qui signifie qu'il y a plus d'étudiants qui étudient longtemps.
- Une kurtosis de 2.00 montre une distribution modérément pointue, suggérant quelques valeurs extrêmes mais pas en grand nombre.

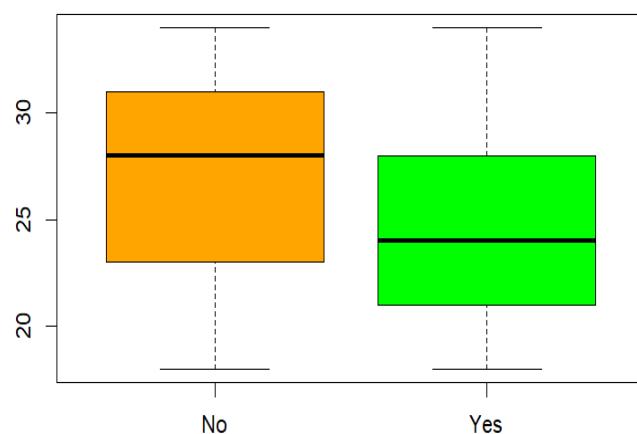
## 2.2 Analyse bivariée

Cette analyse aura pour but de mettre en exergue les relations entre les autres variables de \*\$données et la variable dépression afin de vérifier si elles sont ou non liées entre elles dans un premier temps et liés à la variable dépression dans un second temps.

### 2.2.1 Analyse relationnel entre la variable qualitative et la variable quantitative

#### 2.2.1.1 Représentation graphique : boxplot

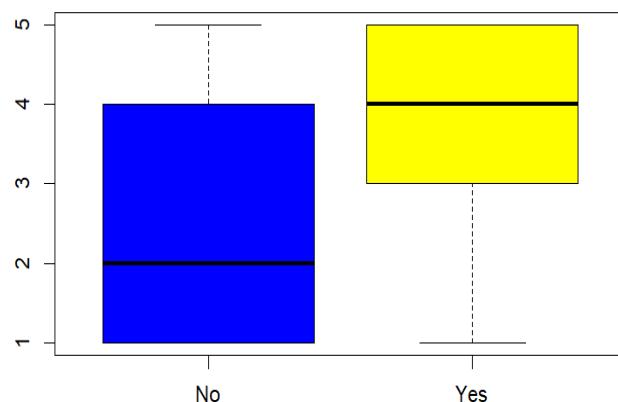
- La dépression en fonction de l'âge



Les étudiants compris entre 22 à 24 ans sont dépressifs alors que les étudiants compris entre 26 à 30 ne sont pas dépressifs

Source : auteur, sortie du logiciel R studio à partir des données secondaires auprès de l' INSSEDS

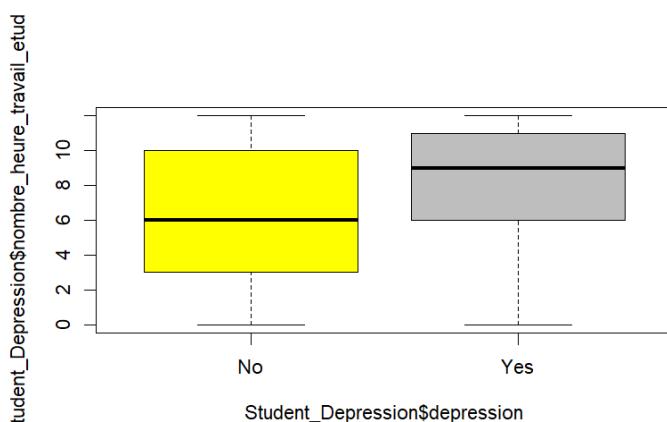
- La dépression en fonction du stress financier



La moitié des étudiants qui ont un stress financier sont dépressifs tandis que les étudiants qui ont un stress financier entre 1 et 2 ne sont pas dépressifs

Source : auteur, sortie du logiciel R studio à partir des données secondaires auprès de l' INSSEDS

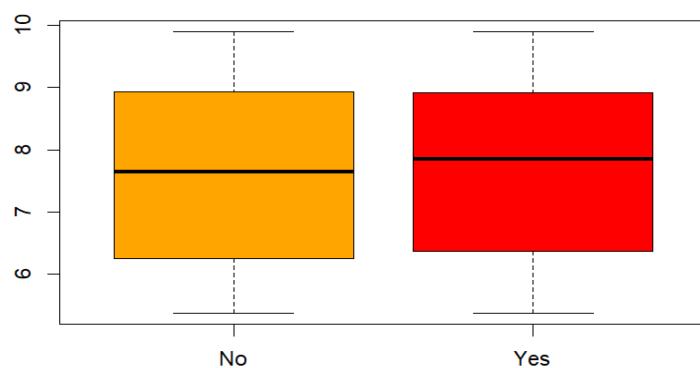
- La dépression en fonction du nombre d'heure de travail d'étude



Les étudiants qui ont un nombre d'heure de travail compris entre 6 et 9 sont dépressif alors que ceux qui sont compris entre 3 et 6 ne sont pas dépressif

Source : auteur , sortie du logiciel R studio à partir des données secondaires aupres de l'insseeds

- La dépression en fonction des moyennes notes



Ici on peut noter que les étudiants ayant des moyennes notes comprise entre 6 et 8 sont dépressif tandis que ceux qui sont compris entre 6,5 et 7,5 ne sont pas dépressif

Source : auteur, sortie du logiciel R studio à partir des données secondaires auprès de l' INSSEDS

#### 2.2.1.2 Mesure de liaison : rapport de corrélation

- La dépression en fonction de l'âge

$$\eta^2 Y|X = \frac{V_{inter}(Y)}{Var(Y)} = \frac{1 - V_{intra}(Y)}{Var(Y)}$$

Le rapport de corrélation est de 0.0512825.  $\eta^2 Y|X$  est proche de 0 ( $V_{inter} < < V_{intra}$ ), cela se traduit par des moyennes conditionnelles qui fluctuent peu entre elles, il y a une très faible liaison entre la dépression et l'âge.

- La dépression en fonction du stress financier

Le rapport de corrélation est de 0.1321987.  $\eta^2 Y|X$  est proche de 0 ( $V_{inter} << V_{intra}$ ), cela se traduit par des moyennes conditionnelles qui fluctuent peu entre elles, il y a une liaison entre deux variables mais une liaison qui n'est pas forte, elle est faible.

- La dépression en fonction des moyennes notes

Le rapport de corrélation est de 0.0004831228.  $\eta^2 Y|X$  est proche de 0 ( $V_{inter} << V_{intra}$ ), cela se traduit par des moyennes conditionnelles qui fluctuent peu entre elles, il y a une liaison entre deux variables mais une liaison qui n'est pas forte, elle est très faible.

- La dépression en fonction du nombre d'heure de travail d'étude

Le rapport de corrélation est de 0.04352529.  $\eta^2 Y|X$  . il n'y a pas de lien linéaire fort entre la dépression et le nombre d'heure travail d'étude

## 2.2.2 Analyse relationnelle entre deux variables qualitatives

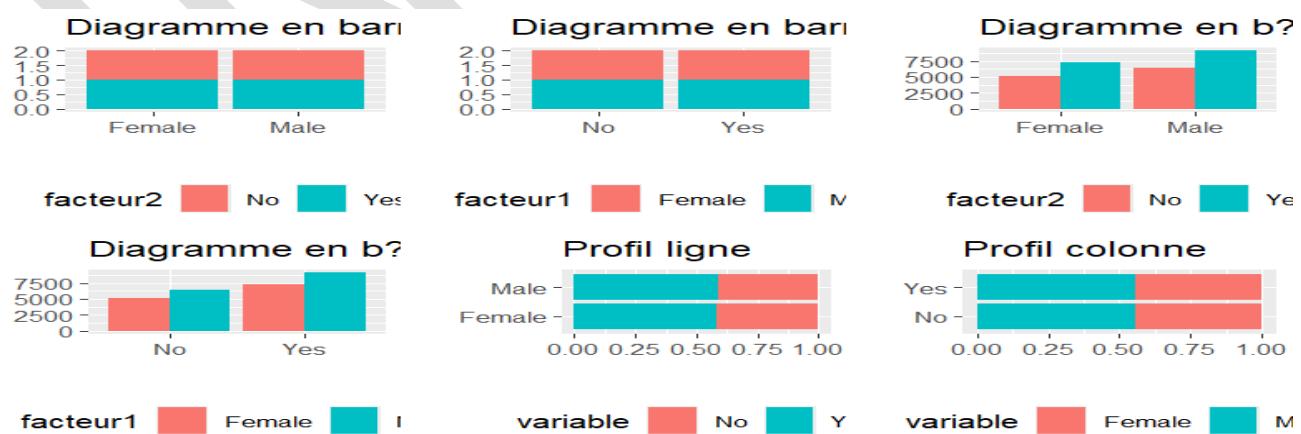
### 2.2.2.1 Tableau statistique de contingence

- Le sexe est-il facteur de dépression ?

TABLEAU DE CONTINGENCE	Colonne1	Colonne2
	NO	YES
MALE	5132	7220
FEMALE	6431	9115
TABLEAU DE FREQUENCE		
	NO	YES
MALE	0.18	0.26
FEMALE	0.23	0.33
TABLEAU PROFIL LIGNE		
	NO	YES
MALE	0.42	0.58
FEMALE	0.41	0.59
TABLEAU PROFIL COLONNE		
	NO	YES
MALE	0.44	0.44
FEMALE	0.56	0.56

### 2.2.2.2 Représentation graphique

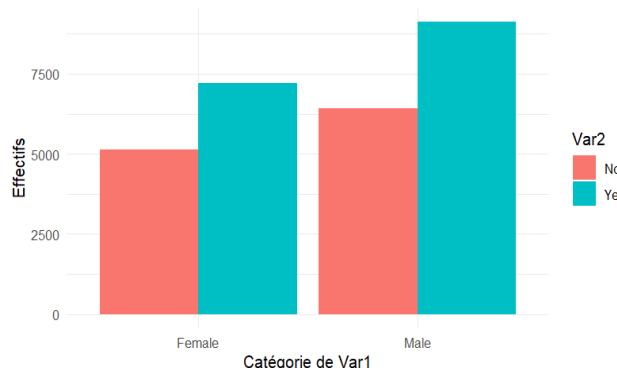
Graphique 1



Source : auteur, sortie du logiciel R studio à partir des données secondaires auprès de l' INSSEDS

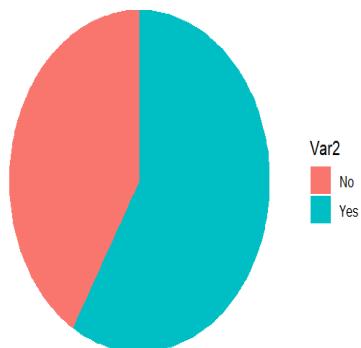
Graphique

Diagramme en barres de la répartition des catégories



Graphique

Diagramme en secteurs de la répartition des catégories



Graphique 2

Source : auteur, sortie du logiciel R studio à partir des données secondaires auprès de l'INSSEDS

### 2.2.2.3 Représentation de données sous forme de résumé numérique

graphique 3

TABLEAU DES EFFECTIFS THEORIQUES		
	NO	YES
MALE	5119.585	7232.415
FEMALE	6443.415	9102.585
ECART A L INDEPENDANCE		
	NO	YES
MALE	12.415	-12.415
FEMALE	-12.415	12.415

Le khi-deux est de 0.08.

Le v de cramer est de 0.002. On constate que le v de cramer est sensiblement égale à 0 donc les deux variables sont indépendantes d'où la liaison très faible.

- La pression académique est-elle facteur de dépression ?**
- Représentation de donnée sous forme de tableau**

```
$Tableau_Contingence
facteur2
facteur1   No  Yes
Faible    6484 2502
Modérée   2973 4488
Élevée    2106 9345
```

```
$Tableau_Frequence
facteur2
facteur1   No  Yes
Faible    0.23 0.09
Modérée   0.11 0.16
Élevée    0.08 0.33
```

```
$Tableau_Profil_Ligne
facteur2
facteur1   No  Yes
Faible    0.72 0.28
Modérée   0.40 0.60
Élevée    0.18 0.82
```

```
$Tableau_Profil_Colonne
facteur2
facteur1   No  Yes
Faible    0.56 0.15
Modérée   0.26 0.27
Élevée    0.18 0.57
```

Représentation de données sous forme de graphique

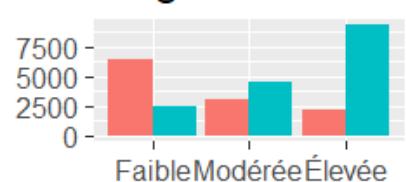
Diagramme en barri



Diagramme en barre



Diagramme en b?



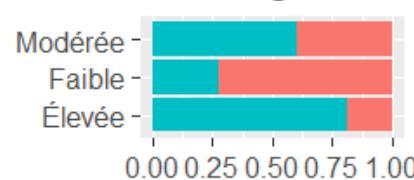
facteur2      No      facteur1      Faible      Modérée

facteur2      No      Ye

Diagramme en b?



Profil ligne



Profil colonne



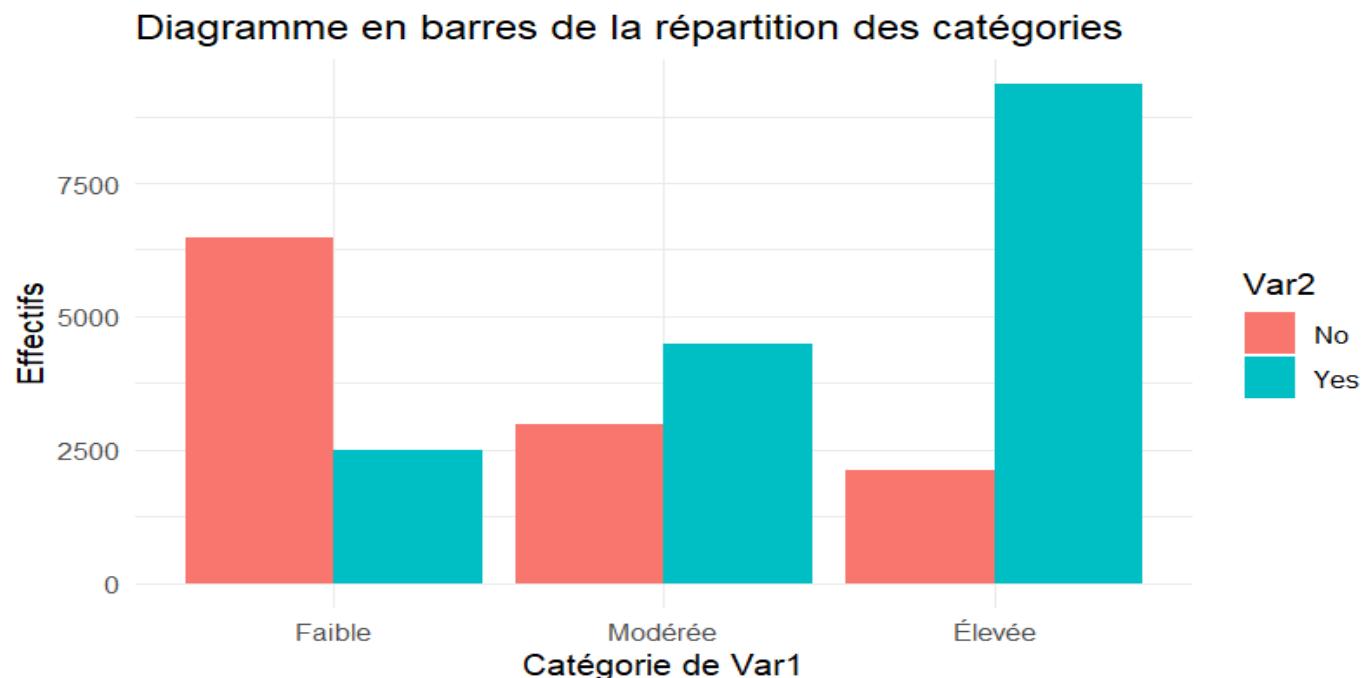
facteur1      Faible      Modérée

variable      No      variable      Faible      Modérée

Source : auteur, sortie du logiciel R studio à partir des données secondaires auprès de l' INSSEDS

L'image montre plusieurs types de diagrammes en barres et de profils (ligne et colonne), qui semblent analyser la relation entre deux variables catégorielles. Ces variables représentent la dépression et la pression académique, ces graphiques permettent d'examiner leur association.

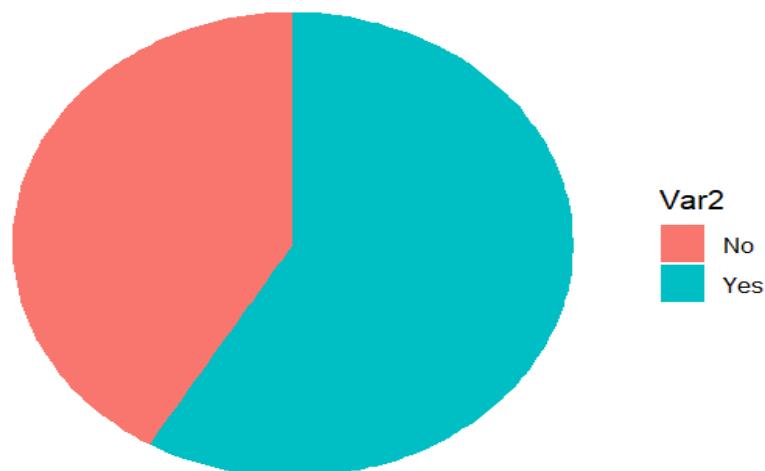
Graphique 2



Source : auteur, sortie du logiciel R studio à partir des données secondaires auprès de l'INSSEDS

Graphique 3

### Diagramme en secteurs de la répartition des catégories



Source : auteur, sortie du logiciel R studio à partir des données secondaires auprès de l'INSSEDS

### Représentation de donnée sous forme de résumé numérique

Tableau des effectifs théoriques

	No	Yes
Faible	1.156176e+04	1.633324e+04
Modérée	4.144742e-01	5.855258e-01
Forte	8.289483e-01	1.171052e+00

écart d'indépendance

	No	Yes
Faible	-0.7565775	0.7565775
Modérée	0.5855258	-0.5855258
Forte	0.1710517	-0.1710517

Le v de cramer est de 0.007 . 0.007 est une valeur proche de 0 ce qui signifie qu'il n'y a presque aucune liaison entre les deux variables qualitatives.

- La pensée suicidaire est-elle facteur de depression ?

Représentation de données sous forme de tableau

\$Tableau_Contingence		
facteur2		
facteur1	No	Yes
No	7866	2379
Yes	3699	13957

- Ligne "No" (absence de dépression) :

- 7 866 personnes sans dépression n'ont pas de pensées suicidaires.
- Ligne "Yes" (présence de dépression) :
- 3 699 personnes avec dépression n'ont pas de pensées suicidaires.

```
$Tableau_Fréquence  
facteur2  
facteur1 No Yes  
No 0.28 0.09  
Yes 0.13 0.50
```

```
$Tableau_Profil_Colonne  
facteur2  
facteur1 No Yes  
No 0.68 0.15  
Yes 0.32 0.85
```

```
$Tableau_Profil_Ligne  
facteur2  
facteur1 No Yes  
No 0.77 0.23  
Yes 0.21 0.79
```

Tableau de fréquence : 50% des étudiants ayant des pensées suicidaires sont dépressif tandis que 13% des étudiants ayant des pensées suicidaires ne sont pas dépressif

Ceci est pareil pour les profils lignes et colonnes

Representation de donnée sous forme de graphique

Diagramme en barre

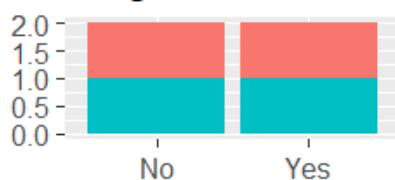
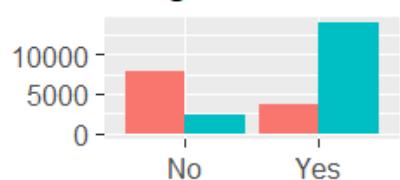


Diagramme en barre



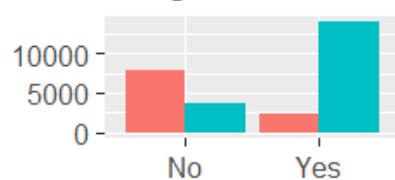
Diagramme en barre



facteur2

No Yes

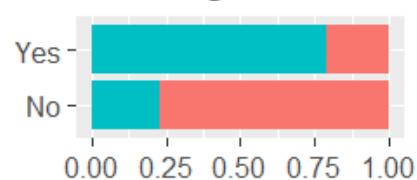
Diagramme en barre



facteur1

No Yes

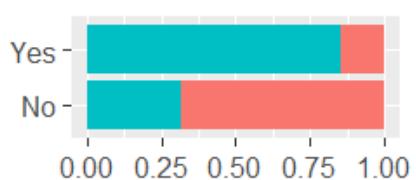
Profil ligne



facteur2

No Yes

Profil colonne



facteur1

No Yes

variable

No Yes

variable

No Yes

Source : auteur , sortie du logiciel R studio à partir des données secondaires auprès de l' INSSEDS

L'image montre plusieurs types de diagrammes en barres et de profils (ligne et colonne), qui semblent analyser la relation entre deux variables catégorielles . ces variables représentent la dépression et les pensées suicidaires, ces graphiques permettent d'examiner leur association.

Représentation graphique sous forme de résumé numérique

- ◆ Tableau des effectifs observés  

```
> print(table_contingence)
```

	No	Yes
No	7864	2378
Yes	3699	13957

-2378 personnes dépressif n'ont pas de pensée suicidaire et 7864 personnes non dépressif n'ont pas de pensée suicidaire.

-13957 personnes dépressif ont également des pensées suicidaires alors que 3699 personnes non dépressif ont des pensées suicidaires

- ◆ Tableau des effectifs théoriques  

```
> print(effectifs_theoriques)
```

	No	Yes
No	4245.044	5996.956
Yes	7317.956	10338.044

- ◆ Écarts d'indépendance (br)  

```
> print(ecarts_indep)
```

	No	Yes
No	3618.956	-3618.956
Yes	-3618.956	3618.956

Les tableaux ci-dessous montrent les résultats des effectifs théoriques et des à l'indépendance

Le v de cramer est de 0.546. 0.546 est une valeur proche de 1 ce qui signifie qu'il y a une forte liaison entre les deux variables qualitatives que sont la dépression et la pensée suicidaire. La pensée suicidaire est facteur de dépression

### L'habitude alimentaire est-elle facteur de dépression ?

- Représentation de données sous forme de tableau

**\$Tableau\_Contingence**  
facteur2

facteur1	No	Yes
Healthy	4177	3472
Moderate	4363	5558
Others	4	8
Unhealthy	3019	7297

**\$Tableau\_Frequence**  
facteur2

facteur1	No	Yes
Healthy	0.15	0.12
Moderate	0.16	0.20
Others	0.00	0.00
Unhealthy	0.11	0.26

**\$Tableau\_Profil\_Ligne**  
facteur2

facteur1	No	Yes
Healthy	0.55	0.45
Moderate	0.44	0.56
Others	0.33	0.67
Unhealthy	0.29	0.71

**\$Tableau\_Profil\_Colonne**  
facteur2

facteur1	No	Yes
Healthy	0.36	0.21
Moderate	0.38	0.34
Others	0.00	0.00
Unhealthy	0.26	0.45

Analyse des résultats

On constate qu'il n'y a plus de personne malade dépressif que de personne en bonne santé et modéré

Cela se constate par les tableaux des profils colonnes des profils ligne et le tableau des fréquences

-Représentation de donnée sous forme de graphique

graphique

Diagramme en barre

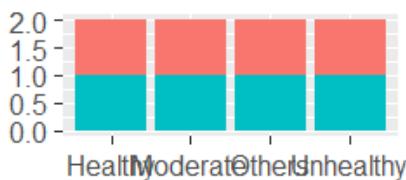


Diagramme en barre

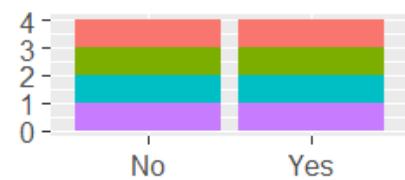
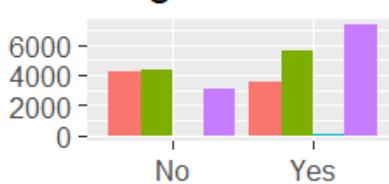


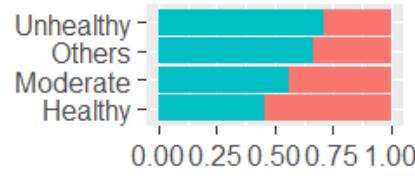
Diagramme en b?



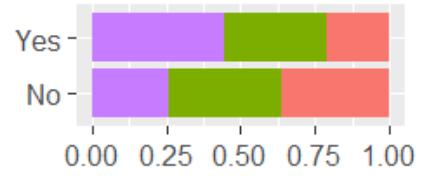
Diagramme en b?



Profil ligne



Profil colonne



Healthy      Moderate      Oth

variable      variable      Healthy      Moderate      Oth

Source : auteur, sortie du logiciel R studio à partir des données secondaires auprès de l'INSSEDS

-Représentation de donnée sous forme de résumé numérique

- ◆ Tableau des effectifs théoriques :
- > print(effectifs\_theoriques)
- ◆ Ecarts d'indépendance (Observé - Théorique) :
- > print(ecarts\_indep)

	No	Yes
Healthy	3170.31282	4478.68718
Moderate	4111.99810	5809.00190
Others	4.97369	7.02631
Unhealthy	4275.71539	6040.28461

	No	Yes
Healthy	1006.6871819	-1006.6871819
Moderate	251.0018998	-251.0018998
Others	-0.9736899	0.9736899
Unhealthy	-1256.7153918	1256.7153918

Le v de cramer est de 0.208. 0.208 est proche de 0 cela se traduit par une liaison faible entre l'habitude alimentaire et la dépression.

### **3 Recommandations**

À la lumière de votre étude sur la dépression chez les étudiants, voici quelques recommandations pour améliorer la prise en charge et la prévention de ce phénomène :

1. **Renforcement du soutien psychologique** : Les institutions académiques devraient mettre en place des services de soutien psychologique accessibles aux étudiants, notamment des consultations gratuites et anonymes.
2. **Sensibilisation et prévention** : Des campagnes de sensibilisation sur la santé mentale, le stress et les risques de la dépression pourraient être organisées pour aider les étudiants à identifier les signes précoces et à chercher de l'aide.
3. **Gestion du stress financier** : Étant donné que le stress financier a une influence sur la dépression, des aides financières, des bourses ou des emplois étudiants adaptés pourraient être mis en place pour alléger cette pression.
4. **Équilibre entre études et vie personnelle** : Encourager une bonne gestion du temps, promouvoir des activités extrascolaires et sportives, ainsi qu'un sommeil régulier, peut réduire le risque de dépression.
5. **Suivi des étudiants à risque** : Les étudiants présentant des signes de dépression ou ayant des antécédents de maladies mentales dans leur famille devraient être suivis de près par des professionnels de la santé et les établissements d'enseignement supérieur.

### **Conclusion**

Notre étude met en évidence les facteurs influençant la dépression chez les étudiants, notamment le stress académique, financier et les pensées suicidaires. Bien qu'il y a d'autres facteurs dans le jeu de donnée j'ai trouvé ces termes énumérés pertinents dans mon rapport. Si certains facteurs comme l'âge et les résultats académiques ont une faible corrélation avec la dépression, d'autres, tels que les pensées suicidaires et le stress financier, présentent une relation plus significative.

En comprenant mieux ces facteurs, il devient possible d'élaborer des stratégies ciblées pour améliorer le bien-être mental des étudiants. Il est essentiel que les institutions académiques, les pouvoirs publics et les étudiants eux-mêmes s'engagent ensemble dans la mise en place de mesures préventives et de soutien pour réduire l'impact de la dépression dans le milieu universitaire.

### **ANNEXE**

#### **Source du code R**

```
# importation des bibliotheques nécessaires  
library(israel)  
library(visdat)
```

```
library(ggplot2)
library(corrplot)
library(vcd)

# importation des données
Student_Depression <- read.csv("C:/Users/PCFRANCESERVICES/Desktop/inssts/jeudedonnee/Student_Depression.csv",
stringsAsFactors=TRUE)

View(Student_Depression)
head(Student_Depression)

# extraction de la variable id
Student_Depression <- Student_Depression[,-1]

#--- decodage de quelques variables----#
Student_Depression$depression <- factor(Student_Depression$depression, levels = c("0","1"), labels = c("No","Yes"))

Student_Depression$pession_academique <- cut(Student_Depression$pession_academique,
breaks = c(-Inf, 2, 3, 5), # Définir les intervalles
labels = c("Faible", "Modérée", "Élevée"),
right = TRUE, # Inclure la limite supérieure
ordered_result = TRUE) # Catégorie ordonnée

# Afficher un aperçu
table(Student_Depression$pession_academique)

# Catégorisation pession_liee_travail
Student_Depression$pession_liee_travail <- factor(Student_Depression$pession_liee_travail,
levels = c(0, 2, 5),
labels = c("Faible", "Modérée", "Forte"),
ordered = TRUE) # Facteur ordonné

# Vérifiez la table des catégories
table(Student_Depression$pession_liee_travail)

# Catégoriser "satisfaction etudes"
Student_Depression$satisfaction_etudes <- cut(Student_Depression$satisfaction_etudes,
breaks = c(-Inf, 2, 3, 5), # Définir les intervalles
```

```
labels = c("Faible", "Modérée", "Élevée"),  
right = TRUE, # Inclure la limite supérieure  
ordered_result = TRUE) # Catégorie ordonnée  
  
# Afficher un aperçu  
table(Student_Depression$satisfaction_etudes)  
  
# Catégoriser "pression_academique"  
Student_Depression$satisfaction_travail <- cut(Student_Depression$satisfaction_travail,  
breaks = c(-Inf, 1, 2, 4), # Définir les intervalles  
labels = c("Faible", "Modérée", "Élevée"),  
right = TRUE, # Inclure la limite supérieure  
ordered_result = TRUE) # Catégorie ordonnée  
  
# Afficher un aperçu  
table(Student_Depression$satisfaction_travail)  
print(Student_Depression)  
View(Student_Depression)  
  
# visualisation des données manquantes  
library(visdat)  
vis_miss(Student_Depression)  
vis_dat(Student_Depression)  
  
# suppression des valeurs manquantes  
Student_Depression= na.omit(Student_Depression)  
  
#---revisualisation des données manquantes----#  
vis_miss(Student_Depression)  
vis_dat(Student_Depression)  
  
#---traitement des doublons-----#  
traitement_doublons(Student_Depression)
```

```
#### visualisation des données extremes ou abberantes --#
par(mfrow=c(1,4),mar=c(3,3,3,3))
boxplot(Student_Depression$age, col = "red")
boxplot(Student_Depression$moyenne_notes, col = "yellow")
boxplot(Student_Depression$nombre_heure_travail_etude, col = "cyan")
boxplot(Student_Depression$stress_financier,col = "blue")
par(mfrow=c(1,1),mar=c(3,3,3,3))

#-----#
###traitement des données extremes ou abberantes-#
library(DescTools)
Student_Depression$moyenne_notes <- Winsorize(Student_Depression$moyenne_notes)
Student_Depression$age<- Winsorize(Student_Depression$age)
Student_Depression$nombre_heure_travail_etude<- Winsorize(Student_Depression$nombre_heure_travail_etude)
Student_Depression$stress_financier<- Winsorize(Student_Depression$stress_financier)

##-- revisualisation des données extremes ou abberantes
# visualisation des donnees extrêmes ou aberrantes
par(mfrow=c(1,4),mar=c(3,3,3,3))
boxplot(Student_Depression$age, col = "red")
boxplot(Student_Depression$moyenne_notes,col = "yellow")
boxplot(Student_Depression$nombre_heure_travail_etude,col = "orange")
boxplot(Student_Depression$stress_financier,col = "blue")
par(mfrow=c(1,1),mar=c(3,3,3,3))

#####ANALYSE UNIVARIE-----#
#####VARIABLE QUANTITATIVE-----#
summary(Student_Depression)

#tableau statistique
```

```
israel.qt.tableau(Student_Depression$age)
israel.qt.tableau(Student_Depression$moyenne_notes)
israel.qt.tableau(Student_Depression$stress_financier)
israel.qt.tableau(Student_Depression$nombre_heure_travail_etude)
```

```
# representation graphique-----#
israel.qt.graph(Student_Depression$age)
israel.qt.graph(Student_Depression$moyenne_notes)
israel.qt.graph(Student_Depression$stress_financier)
israel.qt.graph(Student_Depression$nombre_heure_travail_etude)
```

```
# resume numerique
israel.qt.resume(Student_Depression$age)
israel.qt.resume(Student_Depression$moyenne_notes)
israel.qt.resume(Student_Depression$stress_financier)
israel.qt.resume(Student_Depression$nombre_heure_travail_etude)
```

```
#####---VARIABLE QUALITATIVE-----#####
#-----TABLEAU STATISTIQUE
```

```
israel.ql.tableau(Student_Depression$sexe)
israel.ql.tableau(Student_Depression$profession)
israel.ql.tableau(Student_Depression$habitudes_alimentaires)
israel.ql.tableau(Student_Depression$antecedants_familiaux_maladie_mentale)
israel.ql.tableau(Student_Depression$pensees_suicidaire)
israel.ql.tableau(Student_Depression$depression)
israel.ql.tableau(Student_Depression$pression_academique)
israel.ql.tableau(Student_Depression$satisfaction_etudes)
israel.ql.tableau(Student_Depression$satisfaction_travail)
israel.ql.tableau(Student_Depression$duree_sommeil)
israel.ql.tableau(Student_Depression$diplome_suivi)
israel.ql.tableau(Student_Depression$ville)
israel.ql.tableau(Student_Depression$pression_liee_au_travail)
```

----- representation graphique---#

```
israel.ql.graph(Student_Depression$sexe)
israel.ql.graph(Student_Depression$profession)
israel.ql.graph(Student_Depression$habitudes_alimentaires)
israel.ql.graph(Student_Depression$antecedants_familiaux_maladie_mentale)
israel.ql.graph(Student_Depression$depression)
israel.ql.graph(Student_Depression$pensees_suicidaire)
israel.ql.graph(Student_Depression$pression_academique)
israel.ql.graph(Student_Depression$satisfaction_etudes)
israel.ql.graph(Student_Depression$satisfaction_travail)
israel.ql.graph(Student_Depression$duree_sommeil)
israel.ql.graph(Student_Depression$diplome_suivi)
israel.ql.graph(Student_Depression$ville)
israel.ql.graph(Student_Depression$pression_liee_au_travail)
```

----- ANALYSE BIVARIE -----#

####LA DEPRESSION DEPEND T 'ELLE DE L'AGE ?

#tableau statistique contingence

```
table(Student_Depression$age,Student_Depression$depression)
```

## Représentation des données sous forme de GRAPHIQUE

#Diagramme en boîte par niveau de facteur

```
boxplot(Student_Depression$age ~ Student_Depression$depression, col=c("orange","green"))
```

# liaison entre l'age et la depression

```
israel.qtql.liaison(Student_Depression$age,Student_Depression$depression)
```

#####

## LA DEPRESSION DEPEND T'ELLE DU STRESS FINANCIER ?

```

#tableau statistique contingence
table(Student_Depression$stress_financier,Student_Depression$depression)

## Représentation des données sous forme de GRAPHIQUE

#Diagramme en boîte par niveau de facteur
boxplot(Student_Depression$stress_financier ~ Student_Depression$depression, col=c("blue","yellow"))

# liaison entre l'age et la depression
israel.qtql.liaison(Student_Depression$stress_financier,Student_Depression$depression)

#####
#####

# LA DEPRESSION DEPEND T ELLE DU NOMBRE D'HEURE TRAVAIL ETUDE ?
#tableau statistique contingence
table(Student_Depression$nombre_heure_travail_etude,Student_Depression$depression)

## Représentation des données sous forme de GRAPHIQUE

#Diagramme en boîte par niveau de facte
boxplot(Student_Depression$nombre_heure_travail_etude ~ Student_Depression$depression,
col=c("yellow","gray"))

# liaison entre l'age et la depression
israel.qtql.liaison(Student_Depression$nombre_heure_travail_etude,Student_Depression$depression)

#####
#####

# LA DEPRESSION DEPEND T ELLE DES MOYENNES NOTES ?
# tableau statistique de contingence
table(Student_Depression$moyenne_notes,Student_Depression$depression)

## Représentation des données sous forme de GRAPHIQUE

#Diagramme en boîte par niveau de facte
boxplot(Student_Depression$moyenne_notes ~ Student_Depression$depression, col=c("orange","red"))

# liaison entre l'age et la depression
israel.qtql.liaison(Student_Depression$moyenne_notes,Student_Depression$depression)

#####
#####

#### LA DEPRESSION DEPEND T'ELLE DU SEXE ?

```

```
# tableau statistique de contingence
israel.2ql.tableau(Student_Depression$sexe,Student_Depression$depression)

# representation graphique
israel.2ql.graph(Student_Depression$sexe,Student_Depression$depression)

# indicateur de liaison
israel.2ql.liaison(Student_Depression$sexe,Student_Depression$depression)

# Construction du tableau de contingence
table_contingence <- table(Student_Depression$sexe,Student_Depression$depression)

# Affichage du tableau des effectifs observés
cat("\n ◆ Tableau des effectifs observés :\n")
print(table_contingence)

# Calcul des effectifs théoriques sous l'hypothèse d'indépendance
effectifs_theoriques <- chisq.test(table_contingence)$expected

# Affichage du tableau des effectifs théoriques
cat("\n ◆ Tableau des effectifs théoriques :\n")
print(effectifs_theoriques)

# Calcul des écarts d'indépendance (observé - théorique)
ecarts_indep <- table_contingence - effectifs_theoriques

# Affichage des écarts d'indépendance
cat("\n ◆ Écarts d'indépendance (Observé - Théorique) :\n")
print(ecarts_indep)

# Calcul du V de Cramer
v_cramer <- assocstats(table_contingence)$cramer
cat("\n ◆ V de Cramer =", round(v_cramer, 3), "\n")

# Interprétation de l'intensité de l'association
if (v_cramer < 0.1) {
  cat("➡ Association très faible (quasi inexistante).\n")
```

```

} else if (v_cramer < 0.3) {
  cat("➡ Association faible.\n")
} else if (v_cramer < 0.5) {
  cat("➡ Association modérée.\n")
} else {
  cat("➡ Association forte.\n")
}

# Visualisation : Diagramme en barres
ggplot(as.data.frame(table_contingence), aes(x = Var1, y = Freq, fill = Var2)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "Diagramme en barres de la répartition des catégories",
       x = "Catégorie de Var1", y = "Effectifs", fill = "Var2") +
  theme_minimal()

# Visualisation : Diagramme en secteurs
ggplot(as.data.frame(table_contingence), aes(x = "", y = Freq, fill = Var2)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y", start = 0) +
  labs(title = "Diagramme en secteurs de la répartition des catégories", fill = "Var2") +
  theme_void()

#####
### LA DEPRESSION DEPEND T'ELLE DE LA PRESSION ACADEMIQUE

# tableau statistique de contingence
israel.2ql.tableau(Student_Depression$pression_academique, Student_Depression$depression)
## representation graphique
israel.2ql.graph(Student_Depression$pression_academique, Student_Depression$depression)
# indicateur de liaison
# Construction du tableau de contingence
table_contingence <- table(Student_Depression$pression_academique, Student_Depression$depression)

```

```
# Affichage du tableau des effectifs observés  
cat("\n ◆ Tableau des effectifs observés :\n")  
print(table_contingence)  
  
# Calcul des effectifs théoriques sous l'hypothèse d'indépendance  
effectifs_theoriques <- chisq.test(table_contingence)$expected  
  
# Affichage du tableau des effectifs théoriques  
cat("\n ◆ Tableau des effectifs théoriques :\n")  
print(effectifs_theoriques)  
  
# Calcul des écarts d'indépendance (observé - théorique)  
ecarts_indep <- table_contingence - effectifs_theoriques  
  
# Affichage des écarts d'indépendance  
cat("\n ◆ Écarts d'indépendance (Observé - Théorique) :\n")  
print(ecarts_indep)  
  
# Calcul du V de Cramer  
v_cramer <- assocstats(table_contingence)$cramer  
cat("\n ◆ V de Cramer =", round(v_cramer, 3), "\n")  
  
# Interprétation de l'intensité de l'association  
if (v_cramer < 0.1) {  
  cat("➡ Association très faible (quasi inexistante).\n")  
} else if (v_cramer < 0.3) {  
  cat("➡ Association faible.\n")  
} else if (v_cramer < 0.5) {  
  cat("➡ Association modérée.\n")  
} else {  
  cat("➡ Association forte.\n")  
}
```

```
# Visualisation : Diagramme en barres
ggplot(as.data.frame(table_contingence), aes(x = Var1, y = Freq, fill = Var2)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "Diagramme en barres de la répartition des catégories",
       x = "Catégorie de Var1", y = "Effectifs", fill = "Var2") +
  theme_minimal()

# Visualisation : Diagramme en secteurs
ggplot(as.data.frame(table_contingence), aes(x = "", y = Freq, fill = Var2)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y", start = 0) +
  labs(title = "Diagramme en secteurs de la répartition des catégories", fill = "Var2") +
  theme_void()

#####
#####

### LA DEPRESSION DEPEND T'ELLE DES PENSES SUICIDAIRES ?

# tableau statistique de contingence
israel.2ql.tableau(Student_Depression$pensees_suicidaire,Student_Depression$depression)
#representation graphique
israel.2ql.graph(Student_Depression$pensees_suicidaire,Student_Depression$depression)
#indicateur de liaison
# Construction du tableau de contingence
table_contingence <- table(Student_Depression$pensees_suicidaire,Student_Depression$depression)

# Affichage du tableau des effectifs observés
cat("\n ♦ Tableau des effectifs observés :\n")
print(table_contingence)

# Calcul des effectifs théoriques sous l'hypothèse d'indépendance
```

```

effectifs_theoriques <- chisq.test(table_contingence)$expected

# Affichage du tableau des effectifs théoriques

cat("\n ◆ Tableau des effectifs théoriques :\n")
print(effectifs_theoriques)

# Calcul des écarts d'indépendance (observé - théorique)

ecarts_indep <- table_contingence - effectifs_theoriques

# Affichage des écarts d'indépendance

cat("\n ◆ Écarts d'indépendance (Observé - Théorique) :\n")
print(ecarts_indep)

# Calcul du V de Cramer

v_cramer <- assocstats(table_contingence)$cramer

cat("\n ◆ V de Cramer =", round(v_cramer, 3), "\n")

# Interprétation de l'intensité de l'association

if (v_cramer < 0.1) {
  cat("➡ Association très faible (quasi inexistante).\n")
} else if (v_cramer < 0.3) {
  cat("➡ Association faible.\n")
} else if (v_cramer < 0.5) {
  cat("➡ Association modérée.\n")
} else {
  cat("➡ Association forte.\n")
}

# Visualisation : Diagramme en barres

ggplot(as.data.frame(table_contingence), aes(x = Var1, y = Freq, fill = Var2)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "Diagramme en barres de la répartition des catégories",
       x = "Catégorie de Var1", y = "Effectifs", fill = "Var2") +
  theme_minimal()

# Visualisation : Diagramme en secteurs

ggplot(as.data.frame(table_contingence), aes(x = "", y = Freq, fill = Var2)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y", start = 0) +

```

```
labs(title = "Diagramme en secteurs de la répartition des catégories", fill = "Var2") +  
theme_void()  
#####  
#####  
### LA DEPRESSION DEPEND T'ELLE De L'HABITUDE ALIMENTAIRE ?  
  
# tableau statistique de contingence  
israel.2ql.tableau(Student_Depression$habitudes_alimentaires,Student_Depression$depression)  
#représentation graphique  
israel.2ql.graph(Student_Depression$habitudes_alimentaires,Student_Depression$depression)  
#indicateur de liaison  
# Construction du tableau de contingence  
table_contingence <- table(Student_Depression$habitudes_alimentaires,Student_Depression$depression)  
  
# Affichage du tableau des effectifs observés  
cat("\n ◆ Tableau des effectifs observés :\n")  
print(table_contingence)  
  
# Calcul des effectifs théoriques sous l'hypothèse d'indépendance  
effectifs_theoriques <- chisq.test(table_contingence)$expected  
  
# Affichage du tableau des effectifs théoriques  
cat("\n ◆ Tableau des effectifs théoriques :\n")  
print(effectifs_theoriques)  
  
# Calcul des écarts d'indépendance (observé - théorique)  
ecarts_indep <- table_contingence - effectifs_theoriques  
  
# Affichage des écarts d'indépendance  
cat("\n ◆ Écarts d'indépendance (Observé - Théorique) :\n")  
print(ecarts_indep)  
  
# Calcul du V de Cramer
```

```
v_cramer <- assocstats(table_contingence)$cramer
cat("\n ♦ V de Cramer =", round(v_cramer, 3), "\n")

# Interprétation de l'intensité de l'association
if (v_cramer < 0.1) {
  cat("➡ Association très faible (quasi inexiste).\\n")
} else if (v_cramer < 0.3) {
  cat("➡ Association faible.\\n")
} else if (v_cramer < 0.5) {
  cat("➡ Association modérée.\\n")
} else {
  cat("➡ Association forte.\\n")
}

# Visualisation : Diagramme en barres
ggplot(as.data.frame(table_contingence), aes(x = Var1, y = Freq, fill = Var2)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "Diagramme en barres de la répartition des catégories",
       x = "Catégorie de Var1", y = "Effectifs", fill = "Var2") +
  theme_minimal()

# Visualisation : Diagramme en secteurs
ggplot(as.data.frame(table_contingence), aes(x = "", y = Freq, fill = Var2)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y", start = 0) +
  labs(title = "Diagramme en secteurs de la répartition des catégories", fill = "Var2") +
  theme_void()
```