STATISTIQUES DESCRIPTIVES

TRAVAUX DIRIGES

EXERCICE 1:

Indiquer de quel type sont les variables présentées ci-dessous :
Qualitatives nominales ou qualitatives ordinales
Quantitatives discrètes ou quantitatives continues

1. L'état-civil des habitants du Portugal
2. La taille des étudiants de l'Université de Harvard
3. Le nombre de pages d'un support de cours
4. Les professions reconnues en Suisse
5. Le nombre de ventes d'un appareil électro-ménager
6. Le nombre d'accidents non-professionnels
7. Le nombre d'enfants dans une famille
8. La nationalité des élèves d'une classe
9. Le poids d'un nouveau né
10. Le nombre de télévisions par famille
11. La couleur des yeux des étudiants de l'Université de Neuchâtel
12. Le nombre de jours de pluie pendant le mois d'août

EXERCICE 2:

Pour chaque ensemble de données ci-dessous :

Nombre de jours de chômage pour 40 personnes :

180	10	30	50	420	30	180	360
200	30	360	120	500	200	30	420
360	370	360	150	180	280	30	500
180	720	420	180	40	500	120	180
194	400	30	360	40	400	180	200

Qualité de production de 30 produits :

D : Défectueux
Q : De bonne qualité

Q	D	Q	D	Q	Q	Q	Q	Q	Q
D	Q	Q	D	Q	D	D	Q	Q	Q
D	D	D	Q	Q	Q	Q	Q	Q	D

1. Définir la population
2. Définir la variable
3. Préciser les modalités de cette variable
4. Déterminer de quel type de variables il s'agit

EXERCICE 3:

Le tableau ci-dessous donne la répartition d'une population d'employés selon leur qualification professionnelle

Variable : **Qualification professionnelle**	**Effectifs** **Ou fréquences absolues**
Qualifié	13
Semi-qualifié	10
Non-qualifié	17
Total	40

Donner la nature de la variable étudiée.
Construire le diagramme en bâtons et le diagramme circulaire (pie-chart) des différentes qualifications professionnelles de cette population.

EXERCICE 4:

Considérons un ensemble de 1250 ouvriers dans le cadre d'une étude sur la récurrence du chômage (personne se trouvant au chômage deux fois ou plus sur une période donnée).
Les unités statistiques de la population sont les ouvriers.
La variable, que nous désignons par X, est le nombre de fois qu'un ouvrier a été au chômage pendant une durée spécifiée, par exemple, une année.
Les modalités de la variables sont ainsi 0, 1, 2, 3, …

Nombre de périodes de chômage	**Effectifs ou** **Fréquences absolues**
0	1150
1	50
2	30
3	20
Total	1250

1. Donner la nature de la variable
2. Compléter le tableau en calculant les fréquences relatives et les pourcentages
3. Construire le diagramme en bâtons et le diagramme circulaire (pie-chart)
4. Compléter le tableau en calculant les fréquences absolues cumulées, les fréquences relatives cumulées et les pourcentages cumulés
5. Donner la représentation graphique de la distribution des fréquences relatives cumulées (cas d'une variable quantitative discrète)

EXERCICE 5:

Considérons la question suivante :
Comment la population du canton de Neuchâtel est-elle répartie selon le revenu ?
Chaque année, en mars, les contribuables sont invités à remplir une déclaration fiscale portant sur l'année précédente et sur la base de laquelle leur revenu imposable est déterminé.
Au total, 60528 contribuables ont été pris en compte pour l'année fiscale 1975/76.
Comme il est impossible d'appréhender efficacement un aussi grand nombre d'observations, il est essentiel de les organiser systématiquement, en les regroupant par classes de revenu.
Le tableau suivant présente un tel regroupement :

Répartition de la population du canton de Neuchâtel selon le revenu

Classe de revenu net en milliers de francs (suisses)	**Fréquences absolues Ou Effectifs**	**Pourcentages (%)**
[0-10[	284	0.47
[10-20[	13177	21.77
[20-50[	40318	66.61
[50-80[	5054	8.35
[80-120[	1028	1.70
[120 et plus[	667	1.10
Total	60 528	100.00

1. Quelle est la nature de la variable étudiée ?
2. Calculer les centres des classes
3. Calculer les amplitudes des classes
4. Construire l'histogramme des revenus du canton de Neuchâtel de la période 1975/76
5. Calculer la proportion de contribuables ayant un revenu strictement inférieur à 50 milliers de francs
6. Calculer la proportion de contribuables ayant un revenu supérieur ou égale à 20 milliers de francs
7. Représenter graphiquement la distribution de fréquences relatives cumulées (cas d'une variable quantitative continue)

EXERCICE 6:

Lors d'un cours de statistique, en 1989, 32 étudiants ont été invités à indiquer leur taille. Le tableau ci-dessous reproduit ces données.
On notera que, dans cet exemple, nous sommes en présence d'un tableau de données individuelles et non regroupées.

N° d'ordre	Taille en cm	N° d'ordre	Taille en cm
1	174	17	170
2	175	18	182
3	180	19	168
4	168	20	171
5	175	21	181
6	170	22	178
7	170	23	180
8	178	24	180
9	187	25	178
10	178	26	182
11	177	27	180
12	172	28	160
13	167	29	165
14	165	30	174
15	174	31	165
16	170	32	165

1. Quelle est la nature de la variable étudiée
2. Construire l'histogramme des tailles de la population d'étudiants

EXERCICE 7:

Considérons le nombre de personnes par ménage dans le canton de Neuchâtel en 1980

x_c	n_c
Ménage de 1 pers.	20 734
Ménage de 2 pers.	20 798
Ménage de 3 pers.	10 067
Ménage de 4 pers.	10 381
Ménage de 5 pers.	3 053
Ménage de 6 pers.	832
Total (nombre de ménages)	65 865

Calculer le nombre moyen de personnes par ménage

EXERCICE 8:

La distribution de fréquences du nombre de lettres par mot dans la langue française, telle qu'elle est obtenue à partir de 10 pages (228 mots) dans le Petit Robert, Edition 1973, est présentée ci-dessous :

Distribution de fréquences du nombre de lettres par mot

Nombre de lettres par mot	Fréquences relatives	Nombre de lettres par mot	Fréquences relatives
4	7/228	11	17/228
5	12/228	12	15/228
6	31/228	13	9/228
7	37/228	14	0/228
8	29/228	15	6/228
9	35/228	16	1/228
10	29/228	Total	1

Calculer le nombre moyen de lettres par mot dans ces 10 pages.

EXERCICE 9:

Soit la distribution de fréquences des revenus annuels des ménages en Suisse, en 1976, présentée dans le tableau ci dessous :

Distribution de fréquences des revenus

Revenu annuel (en francs suisse)	Nombre de ménages	Revenu moyen (en francs suisse)
[24 000 – 36 000[	41	31 953
[36 000 – 48 000[	151	42 596
[48 000 – 60 000[	153	53 916
[60 000 – 72 000[	82	65 562
[72 000 – 84 000[	39	78 064
[84 000 – 96 000[	29	89 573
[96 000 – 108 000[	9	101 018
Total	504	

1. Calculer le revenu moyen des ménages en Suisse en 1976, tout revenu confondu
2. Estimer le revenu moyen des ménages en supposant que l'on ne connaisse pas les revenus moyens par tranche de revenu (i.e : tableau précédent sans la troisième colonne !)

EXERCICE 10:

Le tableau suivant regroupe les données relatives à 5 classes d'étudiants fictifs ayant subi un examen d'anglais :

N° de classe	**Moyennes obtenues**	**Nombre d'élèves n_c par classe**
1	4.5	30
2	5.2	20
3	4.7	25
4	5.0	35
5	5.9	40
		$\sum_{c=1}^{5} n_c = 150$

Calculer la moyenne de l'ensemble des élèves

EXERCICE 11:

1. Calculer la médiane des cinq observations suivantes :
   4.9 –5.3 – 2.6 – 3.1 –5.5
2. Calculer la médiane des quatre observations suivantes :
   4.9 – 3.1 – 2.6 –5.3

EXERCICE 12:

Considérons les données du tableau suivant.

N° de classe	**Scores Intervalles de classe**	**Effectifs ou Fréquences absolues n_c**	**Fréquences Relatives f_c**
1	[16-21[	2	0.022
2	[21-26[	5	0.055
3	[26-31[	8	0.089
4	[31-36[	17	0.189
5	[36-41[	11	0.122
6	[41-46[	26	0.289
7	[46-51[	15	0.167
8	[51-56[	5	0.056
9	[56-60[	1	0.011
		$\sum n_c = 90$	$\sum f_c = 1.000$

1. Dans quel intervalle se trouve la médiane ?
2. Calculer sa valeur

EXERCICE 13:

Prenons ces dix observations :

1 2 4 4 5 5 5 6 7 9

1. Calculer les quartiles Q_1, Q_2 et Q_3 avec la technique proposée dans le cours
2. Calculer l'écart interquartile

EXERCICE 14:

Profits (en milliers de francs) de 100 épiceries

Profit (en milliers de francs)	Fréquences absolues	Fréquences absolues cumulées	Fréquences relatives cumulées
[100 – 200[	10	10	0.1
[200 – 300[	20	30	0.3
[300 – 400[	40	70	0.7
[400 – 500[	30	100	1.0
Total	100		

1. Calculer les quartiles Q_1, Q_2 et Q_3
2. Calculer l'écart interquartile

EXERCICE 15:

Lors d'un cours de statistique, en 1989, 32 étudiants ont été invités à indiquer leur poids. Le tableau ci dessous reproduit ces données.

N° d'ordre	Poids en kg	N° d'ordre	Poids en kg
1	59	17	55
2	67	18	75
3	60	19	49
4	61	20	60
5	82	21	88
6	76	22	52
7	60	23	54
8	61	24	69
9	67	25	66
10	50	26	61
11	71	27	67
12	72	28	57
13	68	29	78
14	55	30	69
15	60	31	60
16	54	32	46

1. Calculer le poids moyen
2. Calculer la variance empirique
3. Calculer l'écart-type empirique

EXERCICE 16:

Considérons le tableau suivant représentant les résultats obtenus par 71 élèves d'une école technique à un test d'habilité manuel noté de 1 à 9 :

Note	Fréquence
1	1
2	5
3	9
4	13
5	10
6	17
7	6
8	7
9	3

1. Calculer la note moyenne
2. Calculer la variance empirique
3. Calculer l'écart-type empirique

EXERCICE 17:

Considérons de nouveau les données du tableau suivant.

N° de classe	Scores Intervalles de classe	Effectifs ou Fréquences absolues n_c	Fréquences Relatives f_c
1	[16-21[	2	0.022
2	[21-26[	5	0.055
3	[26-31[	8	0.089
4	[31-36[	17	0.189
5	[36-41[	11	0.122
6	[41-46[	26	0.289
7	[46-51[	15	0.167
8	[51-56[	5	0.056
9	[56-60[	1	0.011
		$\sum n_c = 90$	$\sum f_c = 1.000$

1. Calculer le score moyen
2. Calculer la variance empirique

EXERCICE 18:

Les chiffres suivants donnent la température en centigrade (C) durant 7 jours consécutifs à Thèbes à 13h :
38 40 39 38 38 41 41
Les températures en degrés Farenheit (F = 32 + 9/5C) sont :
100.4 104.0 102.2 100.4 100.4 105.8 105.8

1. Calculer l'écart-type empirique de la première série
2. Calculer l'écart-type empirique de la deuxième série
3. Que remarque t-on ?

EXERCICE 19:

Soit la distribution suivante :
$\mathbf{x_c}$: les différentes valeurs de la variable étudiée X ($c \in 1;4$)

$\mathbf{f_c}$: la fréquence relative de la valeur $\mathbf{x_c}$

$\mathbf{x_c}$	$\mathbf{f_c}$
0	0.216
1	0.432
2	0.288
3	0.064

1. Calculer les coefficients de Pearson $\gamma 1$ (coefficient d'asymétrie) et $\gamma 2$ (coefficient d'aplatissement)
2. Quelle est la forme de la distribution ?

EXERCICE 20:

Un échantillon de 10 000 étudiants en 1975-76 se ventile selon le tableau de contingence suivant :

Etudes (Y): / **CSP père (X):**	**Droit**	**Sciences économiques**	**Lettres**	**Sciences**	**Médecine et dentaire**	**Pharmacie**	**Pluridis-ciplinaire**	**IUT**	**Total**
Exploitant agricole	80	36	134	99	65	28	11	58	511
Salarié agricole	6	2	15	6	4	1	1	4	39
Patron, Profession libérale, cadre sup.	168 470	74 191	312 806	137 400	208 876	53 164	21 45	62 79	1035 3031
Cadre moyen	236	99	493	264	281	56	36	87	1552
Employé	145	52	281	133	135	30	20	54	850
Ouvrier	166	64	401	193	127	23	28	129	1131
Personnel de service	16	6	27	11	8	2	2	8	80
Autres	305	115	624	247	301	47	42	90	1771
Total	1592	639	3093	1490	2005	404	206	571	10000

1. Construire le tableau des profils-lignes et commenter-le.
2. Construire le tableau des profils-colonnes et commenter-le.
3. Construire le tableau des effectifs sous l'hypothèse d'indépendance des deux variables

EXERCICE 21:

Pour estimer le nombre de poissons dans un lac, dans un premier temps, on saisit quelques poissons, on les compte, on les marque puis on les libère dans le lac. Un peu plus tard, on ressaisit quelques poissons, on compte combien ils sont, distinguant entre le nombre de poissons déjà saisis et le nombre de nouveaux poissons. Si les prises sont faites d'une façon aléatoire et indépendante, les résultats obtenus nous permettent d'obtenir une estimation du nombre total de poissons dans le lac.
Soit n_1, le nombre de poissons de la première prise, et n_2 celui de la deuxième prise. Soit a le nombre de poissons déjà marqués de la deuxième prise. Finalement n le nombre total de poissons du lac. Ces chiffres peuvent être disposés dans un tableau de contingence tel qu'indiqué ci-dessous :

Deuxième saisie / Première saisie	Poissons pêchés	Poissons non pêchés	Total
Poissons pêchés	a		n_1
Poissons non pêchés			
Total	n_2		n

1. Montrer que le nombre total de poissons doit être supérieur à $n_1 + n_2 - a$
2. Peut-on obtenir une estimation plus exacte tenant compte du fait que les deux prises étaient indépendantes ?
   Montrer que dans ce cas :
   $$\frac{a}{n_2} = \frac{n_1}{n}$$
   et donc
   $$n = \frac{n_1 * n_2}{a} \quad a \neq 0$$

3. Si $n_1 = 84$ et $n_2 = 207$ et $a = 2$, calculer la valeur estimée du nombre total de poissons

EXERCICE 22 (démonstration de cours):

Nous avons vu en cours que les paramètres b_0 et b_1 de la droite des moindres carrés (i.e : droite obtenue à l'aide de la méthode des moindres carrés) étaient égaux à :

$$b_1 = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

1. Démontrer que :

$$b_1 = \frac{S_{XY}}{S_X^2}$$

2. Vérifier les conditions suffisantes d'optimalité

EXERCICE 23 (démonstration de cours):

Démontrer analytiquement la formule de décomposition de la variance :

$$\sum_{i=1}^{n}(y_i-\bar{y})^2=\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2+\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$$

Avec $\sum_{i=1}^{n}(y_i-\bar{y})^2$: Variation totale

$\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2$: Variation expliquée

$\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$: Variation inexpliquée ou résiduelle

EXERCICE 24:

Le tableau ci-dessous contient la liste de 14 pays d'Amérique du Nord et d'Amérique centrale, dont la population dépassait le million d'habitants en 1985. Pour chaque pays, le tableau nous indique son taux de natalité (nombre de naissances par année pour 1000 personnes) ainsi que son taux d'urbanisation (pourcentage de la population vivant dans des villes de plus de 100 000 habitants) en 1980.
Nous désirons déterminer si le taux de natalité des pays pris en considération peut être expliqué uniquement par le taux d'urbanisation.

Pays	Taux de natalité	Taux d'urbanisation
Canada	16.2	55.0
Costa Rica	30.5	27.3
Cuba	16.9	33.3
Etats-Unis	16.0	56.5
El Salvador	40.2	11.5
Guatemala	38.4	14.2
Haïti	41.3	13.9
Honduras	43.9	19.0
Jamaïque	28.3	33.1
Mexique	33.9	43.2
Nicaragua	44.2	28.5
Trinidade/Tobago	24.6	6.8
Panama	28.0	37.7
Rep. Dominicaine	33.1	37.1

1. Quelle sont les variables Y et X ?
2. Représenter le nuage de points. Que peut-on en conclure ?
3. Calculer les coefficients b_0 et b_1 de la droite de régression
4. Tracer la droite sur le graphique et le point de coordonnées $(\bar{x}, \bar{y})$
5. Calculer le coefficient de détermination. Commentez.
6. Calculer le coefficient de corrélation. Commentez.

EXERCICE 25:

On a relevé, pour différentes années, numérotées de 1 à 8, le chiffre d'affaires d'une entreprise et les charges d'exploitation annuelles correspondantes :

Années	Chiffre d'affaires annuel en milliers d'euros	Charges en milliers d'euros
1	780	400
2	860	440
3	870	450
4	890	460
5	920	470
6	810	420
7	790	410
8	800	430

On cherche à expliquer le chiffre d'affaires en fonction des charges

1. Quelle sont les variables Y et X ?
2. Représenter le nuage de points. Que peut-on en conclure ?
3. Calculer les coefficients b_0 et b_1 de la droite de régression
4. Calculer le coefficient de détermination. Que peut-on en dire ?
5. Calculer le coefficient de corrélation. Que peut-on en dire ?

EXERCICE 26:

Les relevés de deux caractères quantitatifs relatifs aux éléments d'une même population donnent lieu au tableau suivant :

x_i	2	4	7	9	10	13	16	19
y_i	32	29	22	26	19	10	14	8

1. Représenter le nuage de points et déterminer les équations des deux droites de régression (droites des moindres carrés). Tracer ces deux droites sur le graphique
2. Quel est le point d'intersection de ces deux droites ?
3. Une corrélation linéaire est-elle possible ?

EXERCICE 27:

L'expert d'un journal a noté sur une échelle de 0 à 10 la tenue de route (variable X) et le confort (variable Y) des modèles automobiles proposées par 4 constructeurs dans la même catégorie (« familiale ») et à des prix voisins.

Marque	X	Y
Hyundai	7	4
Fiat	5	7
Peugeot	8	9
Renault	9	8

1. Calculer le coefficient de corrélation linéaire entre X et Y
2. On estime que l'expert a en fait utiliser une échelle numérique pour représenter les variables qui sont seulement ordinales : autrement dit, il pense que Peugeot tient mieux la route que la Hyundai et le Hyundai mieux que la Fiat mais il est incapable de donner un sens à une appréciation du type « l'écart de qualité entre les tenues de route de la Hyundai et de la Fiat est double de celui entre les tenues de route de la Peugeot et de la Renault.
   Déterminer le tableau des rangs qu'aurait dû donner l'expert.
3. Calculer le coefficient de corrélation des rangs de Spearman. Comparer le au coefficient de corrélation linéaire calculé à la première question.

EXERCICE 28:

On demande à 10 étudiants de donner une note de 0 à 100 pour deux caractéristiques : l'intêret qu'il porte à la presse écrite et la qualité d'information qu'ils estiment donnée par la télévision.
Ces données sont conciliées dans le tableau suivant :

Presse écrite 20 60 10 80 90 30 70 75 0 50
Télévision 50 30 60 10 20 70 40 5 0 90

Calculer le coefficient de corrélation des rangs de Spearman.

EXERCICE 29:

En 1973, F. J. Anscombe a publié dans le numéro 27 de American Statistician un jeu de données très intéressantes pour montrer les pièges du calcul « aveugle » du coefficient de corrélation linéaire.

```
> anscombe
   x1 x2 x3 x4    y1   y2    y3    y4
1  10 10 10  8  8.04 9.14  7.46  6.58
2   8  8  8  8  6.95 8.14  6.77  5.76
3  13 13 13  8  7.58 8.74 12.74  7.71
4   9  9  9  8  8.81 8.77  7.11  8.84
5  11 11 11  8  8.33 9.26  7.81  8.47
6  14 14 14  8  9.96 8.10  8.84  7.04
7   6  6  6  8  7.24 6.13  6.08  5.25
8   4  4  4 19  4.26 3.10  5.39 12.50
9  12 12 12  8 10.84 9.13  8.15  5.56
10  7  7  7  8  4.82 7.26  6.42  7.91
11  5  5  5  8  5.68 4.74  5.73  6.89
```

Question 3 *Calculer la moyenne et la variance de* `y1`, `y2`, `y3` *et* `y4`.

Question 4 *Calculer les coefficients de corrélation des couples* (`x1`,`y1`),...,(`x4`,`y4`). *Que constate-t-on ?*

Question 5 *Tracer la représentation des couples* (`x1`,`y1`),...,(`x4`,`y4`). *Commenter.*

EXERCICE 30 (démonstration de cours):

Soit X une variable statistique qualitative à k modalités et Y une variable statistique quantitative. Chaque modalité de X définit une sous-population : celle des individus ayant cette modalité. On note n_c l'effectif correspondant à la modalité c de X et $\overline{y}^c$ la moyenne des valeurs de la variable Y pour les individus de la modalité c de X.
Démontrer la formule d'analyse de variance :

$$\frac{1}{n}\sum_{c=1}^{k}\sum_{i=1}^{n_c}\left(y_i^c-\overline{y}\right)^2=\frac{1}{n}\sum_{c=1}^{k}\sum_{i=1}^{n_c}\left(y_i^c-\overline{y}^c\right)^2+\frac{1}{n}\sum_{c=1}^{k}n_c\left(\overline{y}^c-\overline{y}\right)^2$$

avec

$\frac{1}{n}\sum_{c=1}^{k}\sum_{i=1}^{n_c}\left(y_i^c-\overline{y}\right)^2$: Variance totale

$\frac{1}{n}\sum_{c=1}^{k}\sum_{i=1}^{n_c}\left(y_i^c-\overline{y}^c\right)^2$: Variance intra-groupe

$\frac{1}{n}\sum_{c=1}^{k}n_c\left(\overline{y}^c-\overline{y}\right)^2$: Variance inter-groupe

EXERCICE 31:

Soit x une série statistique. Démontrer la formule de Koenig pour la variance : $S_X^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - (\overline{x})^2$ avec $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

EXERCICE 32:

La distribution des demandeurs d'emploi selon le sexe et la classe d'âge dans une localité est la suivante :

âge	Hommes	Femmes
[16 ;26[	280	160
[26 ;40[	310	360
[40 ;50[	240	120
[50 ;60[	420	530
[60 ;65[	70	50

1. Tracer la courbes de fréquences relatives cumulées pour les demandeurs d'emploi masculins. Même question pour les demandeurs d'emploi de sexe féminin.
2. Déterminer les quartiles de la variable Age pour les demandeurs d'emploi masculins. Même question pour les demandeurs d'emploi de sexe féminin.
3. Conclusions.

EXERCICE 33 :

Soient x et y deux séries statistiques de taille n. On note rx et ry les séries des rangs correspondantes.

1. Montrer que $\overline{rx} = \frac{n+1}{2}$.
2. Montrer que $s_{rx}^2 = \frac{n^2-1}{12}$.
3. En posant $d_i = rx_i - ry_i$, montrer que $2S_{(rx,ry)} = S_{rx}^2 + S_{ry}^2 - \frac{1}{n}\sum_{i=1}^{n} d_i^2$.
4. En déduire l'expression du coefficient linéaire entre ces deux séries rx er ry, appelé coefficient de corrélation des rangs de Spearman : $r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}$.

EXERCICE 34 :

Un chercheur veut mesurer la résistance de l'homme (Y) en fonction de l'âge (X). Un test de résistance donne à chaque individu allant de 1 à 50. Les données sont les suivantes :

Age	**Résistance**		
15-20	7	7	15
20-25	12	17	12
25-30	14	18	18

1. Représenter graphiquement ces données.
2. Calculer les moyennes arithmétiques dans chaque groupe
3. Calculer la variance totale et les variances inter et intra-groupes.
4. Calculer et interpréter le rapport de corrélation empirique entre X et Y.Conclusion ?