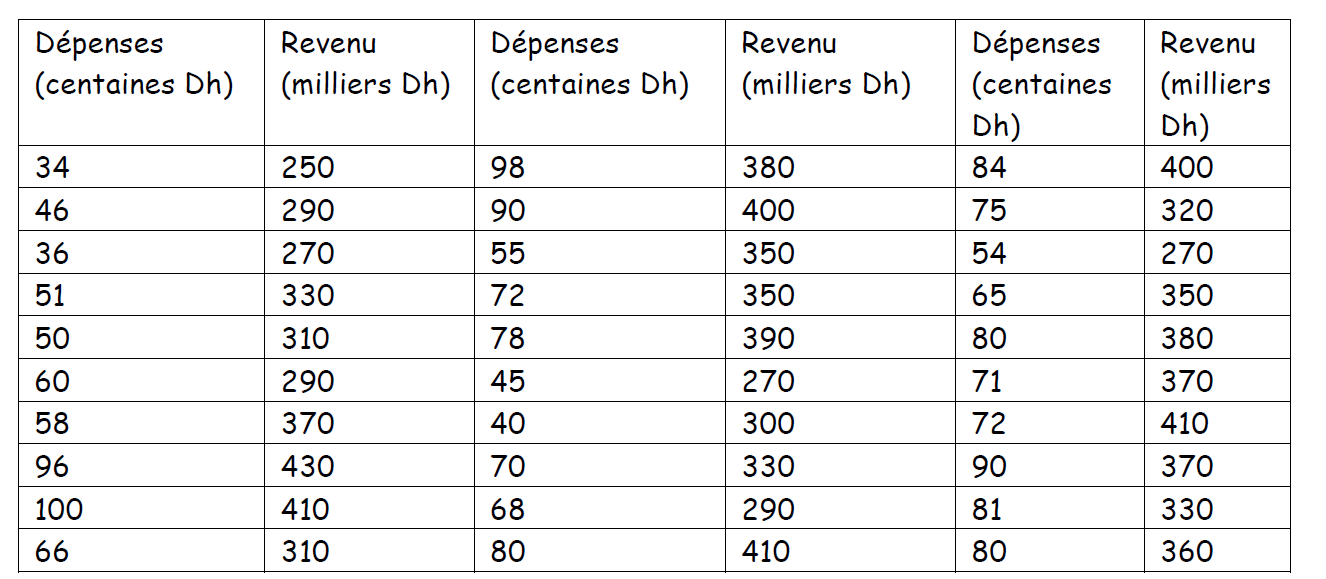
**Statistique Descriptive Année universitaire : 2022/2023**

**Série N°2**

**Exercice1:**

Dans une pré-enquête sur les habitudes de consommation, on a obtenu, auprès d’un échantillon de 30 familles, les données suivantes concernant les dépenses (en centaines de dirhams) pour les loisirs et le revenu annuel (en milliers de dirhams).



1-Tracer le nuage de Points.

2- Est-ce que le diagramme de dispersion permet d’affirmer que les deux variables varient dans le même sens ?

3- Le diagramme suggère-t-il que la liaison entre x et y est de forme linéaire ?

4- Calculer le coefficient de corrélation linéaire.

5- D’après cette étude de corrélation, comment peut-on qualifier les dépenses pour les loisirs des familles de la population échantillonnée qui ont un revenu annuel de 300000 dhs par rapport aux dépenses des familles qui ont un revenu annuel de 400000 dhs ? Sont-elles inférieurs, égales ou supérieurs ?

**Solution :**

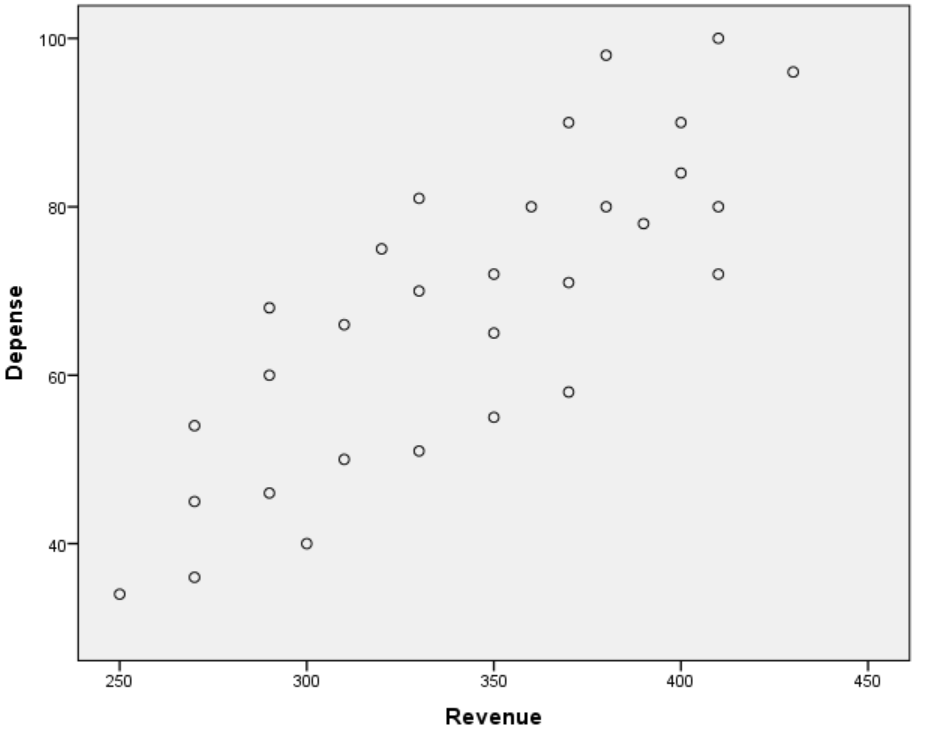
Dans cette étude nous voulons exprimer la relation qui existe entre la variable statistique Y = Dépenses exprimé en centaines Dh et la variable explicative X= Revenu présenté en milliers Dh sous forme d’une expression linéaire. Les deux variables sont de type quantitatif continu.

1-Le nuage de Points.

Il s’agit de tracer les couples points (,), i=1,2,…….,n. ou les et sont les réalisations c-à-d les données collectées pour les deux variables X et Y et n =30.

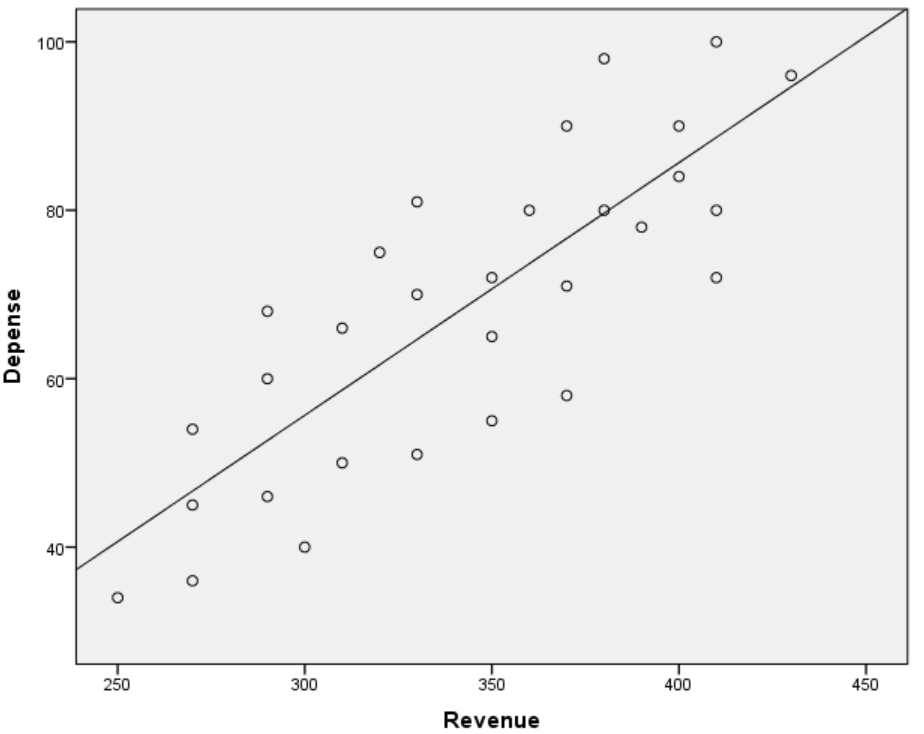
Le graphe de nuage de point obtenue sur SPSS.

Diagramme de nuage de points



2- Le diagramme de dispersion affirme que les deux variables varient dans le même sens suivant le sens de présentation des points.

3- Le diagramme suggère une liaison entre x et y sous forme linéaire puisque le tube de nuage de point prend une forme linéaire. Donc nous pouvons tracer une droite linéaire qui permet de liés des données de X par rapport à Y.



4- Le coefficient de corrélation linéaire est donné par l’expression suivante :

ρ(X,Y) = = 0.815 avec  ***COV(X,Y) =***  est la covariance de X et Y, les deux moyennes :  **=**  et  **=**  et les deux écart-type :  **=**  et  **=** .

La valeur calculée du coefficient de corrélation est supérieure de 0.8 ce qui montre qu’il existe une forte corrélation (liaison ou dépendance) entre les deux variables X= Revenu et Y= Dépenses.

5- Deux familles sont présentées par leurs revenus annuels de 300000 dhs et 400000 dhs et nous voulons comparer leurs éventuelles dépenses. La relation linéaire entre Y= Dépenses et X= Revenu est donnée par la formule suivante :

Y = -34.34 +0.23\*X.

En première temps, nous transformons les deux valeurs de revenus annuels en même unité de mesure présentées dans le tableau de données.

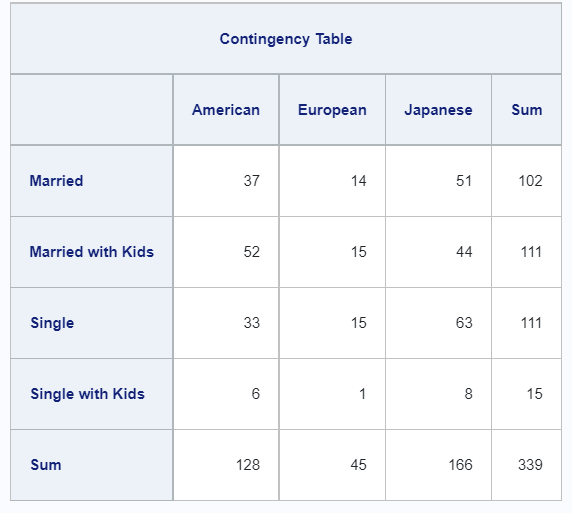
Pour la première famille, le revenu est exprimé par = 300, ce qui donne une estimation du dépense : = -34.34 +0.3\* 300 = 56. La valeur est présenté en unité utilisé qui est

Pour la deuxième famille, le revenu est exprimé par = 400, ce qui donne une estimation du dépense : = -34.34 +0.3\* 400 = 86.

Les dépenses de la deuxième famille est supérieure de celle de la première famille

**Exercice2:**

Les données proviennent d'un échantillon d'individus à qui on a demandé de fournir des informations sur leurs situations familiales (célibataire, marié, célibataire ayant des enfants ou marié ayant des enfants) et sur leur type de l’automobile (américaine, japonaise ou européenne). Le nombre total des individus faisant partie de l’étude est 339. Le tableau de contingence est donnée par :



1. Préciser le type des deux variables.
2. Que représente le tableau ci-dessus.
3. Donner la proportion des individus qui sont Married et ayant type de voiture European.
4. Calculer la proportion de personnes qui ont choisis le type Japanese parmi les Single.
5. Donner le critère et la formule pour mesurer la liaison entre les deux variables.
6. Calculer le coefficient de mesure de la dépendance des deux variables et interpréter la valeur trouvée.

**Solution:**

1. Les deux variables statistiques utilisées dans cette étude sont X = situations familiales et Y = leur type de l’automobile et sont deux variables qualitatives.
2. Le tableau croisé ci-dessus présente le tableau de contingence entre les deux variables X et Y.
3. La proportion des individus qui sont Married et ayant type de voiture European.

P((X = Married)Ո(Y= European)) = = = 0.04129.

1. La proportion de personnes qui ont choisis le type Japanese parmi les Single.

P(Y= Japanese / X= Single) = = = 0.5675.

1. Le critère qui permet de mesurer la liaison entre deux variables qualitatives est le coefficient de Khi-deux . La formule est donné par l’expression suivante :

1. Dans cette étude on a K = 4 et L = 3. La valeur du coefficient de Khi-deux qui permet de mesurer la dépendance entre les deux variables à partir la formule de et le tableau de contingence est 8.34. Dès que la valeur de

> 3 on peut déduire que les deux sont dépendants et plus cette valeur est grande plus la dépendance est significatif. Pour cet exemple nous pouvons conclure qu’il existe une dépendance entre la variable X = Situations familiales et Y = type de l’automobile.