Faculté des sciences

1ère année Master Informatique (ISII)

Matière : Analyse de données



Année : 2022/2023

TD1

#### Série 1 d'exercices

## Exercice 1:

Dans une petite localité, on a relevé une série de nombre de pièces par appartement

Nombre de pièces	1	2	3	4	5	6	7
Nombre d'appartements	48	72	96	64	39	25	3

- 1) Définir la population, la variable (le caractère) étudiée et quelle est sa nature.
- 2) Calculer les paramètres de tendance centrale et les paramètres de dispersion, sachant que l'effectif total de l'échantillon est 347.
- 3) Représenter graphiquement dans la boite à moustache correspondante la médiane (Me) et les quartiles ( $Q_1$  et  $Q_3$ ).
- 4) Représenter graphiquement cette série par les graphiques correspondants.

#### Exercice 2:

Dans une ferme, on a pesé les œufs qui ont été produits (les masses des œufs sont exprimées en grammes) :

Masse de l'œuf	[28-38 [	[38-48 [	[48-58 [	[58-62 [	[62-72 [	[72-82 [
Nombre d'œufs	4	55	78	112	95	6

- 1. Déterminer :
- La population statistique.
- Le caractère étudié.
- La nature du caractère.
- 2) Calculer: étendue (E), mode (Mo), Centile 85 (C<sub>85</sub>), médiane (Me), écarts interquantiles ( $I_D$ ,  $I_Q$ ),  $\bar{X}$ ,  $\sigma_X$ .
- 3) Déterminer le pourcentage d'individus appartenant à l'intervalle  $[\bar{X}-\sigma_X, Q_3]$ .
- 4) Dessiner les graphiques correspondants.

# Exercice 3:

On s'intéresse à l'étude de l'âge et le niveau d'étude d'un groupe d'étudiants, ce qui nous a fourni le tableau suivant :

	Age (X)	[18-20[	[20-22[	[22-24[
Nombre d'années				
à l'université (Y)				
1		19	56	11
2		71	182	23
3		149	391	57
4		36	187	51
5		24	181	79

- 1. Déterminer les deux séries statistiques X et Y (ainsi que leur type).
- 2. Calculer la moyenne et l'écart-type de chaque série.
- 3. Calculer le point moyen du nuage de points.

## Exercice 4:

On cherche à étudier la relation entre le nombre d'enfants d'un couple et son salaire. On dispose de la série bidimensionnelle suivante :

Salaire en euros (Y)	Nombre d'enfants (X)
510	4
590	3
900	2
1420	1
2000	0
600	5
850	6
1300	7
2200	8

Calculer le coefficient de corrélation.

# Exercice 5:

Une expérience a été réalisée dans une entreprise sur 250 personnes pour étudier la relation qui existe entre l'âge (X) et le temps de sommeil (Y), le tableau ci-après a été obtenu.

X Y	[1-3[	[3-11[	[11-19[	[19-31[	[31-59[
[5-7[	0	0	2	9	29
[7-9[	0	3	8	26	15
[9-11[	2	12	35	22	6
[11-15[	36	26	16	3	0

- 1. Calculer les moyennes marginales  $\bar{X}$  et  $\bar{Y}$  et les écarts-types marginaux  $\sigma_x$  et  $\sigma_y$ .
- 2. Déterminer la covariance et le coefficient de corrélation linéaire.