

## TD 02 : **Statistiques descriptives univariées**

### Exercice 1

Une enquête réalisée auprès de 538 étudiants de l'institut a permis d'obtenir les résultats suivants : Pratique de sport

Une enquête Pratique de sport	Effectif
Jamais	163
Rarement	246
Occasionnellement	108
Souvent	21

Quel est le caractère étudié ?  
 Représenter graphiquement ces données.

### Exercice 2

Soit la série des âges des salariés d'une entreprise :

33, 39, 25, 54, 25, 39, 21, 29, 21, 46, 56, 33, 23, 25, 39, 41, 60, 46, 41, 25,  
 21, 54, 46, 25, 30, 60, 21, 21, 26, 21, 33, 21, 37, 25, 55, 35, 33, 46, 33, 30

1. De quel type est le caractère étudié ?
2. Quel est l'effectif de l'échantillon ?
3. Représenter ce caractère par un graphique adéquat.
4. Déterminer les paramètres de position du caractère étudié.

Si on transforme ces données en classes d'âge (21-30 ; 31-40 ; 41-50 ; 51-60).

5. Combien y'a-t-il de modalités pour cette nouvelle série ?
6. Déterminer l'effectif de chaque modalité.
7. Représenter ces nouvelles données par un graphique adéquat. Quelle est la classe modale ?
8. Déterminer les paramètres de position de la nouvelle distribution.
9. Que constatez-vous ?

## TD 02 : Statistiques descriptives univariées

### TD Statistiques / Série N°2 / Corrigé-type Exo 1 et 2

#### Exercice 1

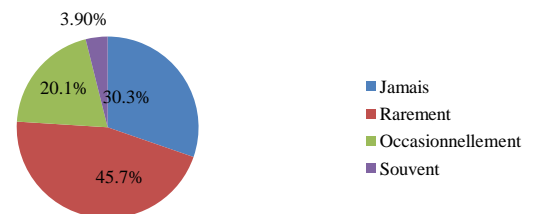
1) Quel est le caractère étudié ? Fréquence de pratique de sport : **caractère qualitatif ordinal**.

2) Représentation graphique ces données : Plusieurs types de représentations sont envisageables :

➤ **Le diagramme circulaire (camembert)**. Il n'y a qu'une réponse possible par affirmation, donc le total n'excède pas 100 %. Pour calculer l'angle de chaque secteur ( $\alpha$ ) :  $\alpha = [ni / n] \times 360^\circ$ .

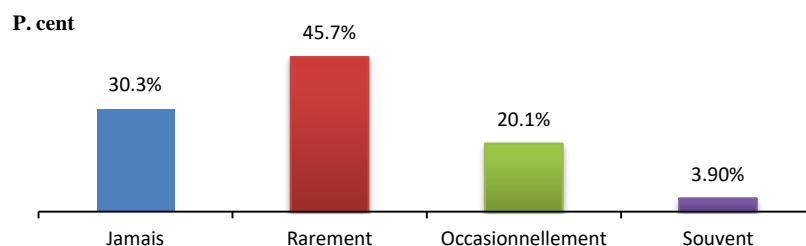
$$n = 538$$

Modalités	$ni$	$Pi$ (%)	$\alpha$
Jamais	163	30,3	109.07
Rarement	246	45,7	164.61
Occasionnellement	108	20,1	72.27
Souvent	21	3,9	14.05



**Figure.** Répartition des étudiants selon la fréquence de pratique de sport

➤ **Le diagramme en bâton**, soit un seul bâton par réponse, les hauteurs des bâtons correspondent aux effectifs (ou pourcentages) des modalités du caractère (variable).



**Figure.** Répartition des étudiants selon la fréquence de pratique de sport

#### Exercice 2

1) De quel type est le caractère étudié ?

Le caractère étudié (l'âge des salariés d'une entreprise) est de type **quantitatif discret**.

2. Quel est l'effectif de l'échantillon ?

L'effectif total est la somme des salariés concernés par cette étude, soit **40 salariés**.

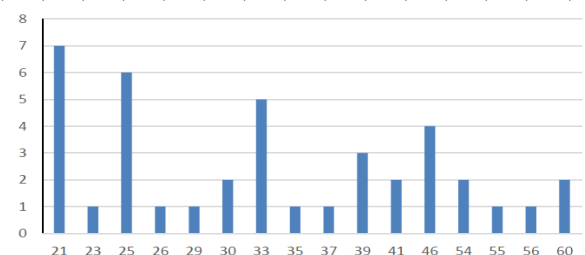
3. Représenter ce caractère par un graphique adéquat.

En premier, il faut classer par ordre l'âge des salariés et déterminer l'effectif de chaque âge.

21, 21, 21, 21, 21, 21, 21, 21, 23, 25, 25, 25, 25, 25, 25, 25, 26, 29, 30, 30, 33, 33, 33, 33, 33, 33, 35, 37, 39, 39, 39, 39, 41, 41, 46, 46, 46, 46, 54, 54, 55, 56, 60, 60. On aura le tableau suivant :

Age	21	23	25	26	29	30	33	35
$ni$	7	1	6	1	1	2	5	1
Age	37	39	41	46	54	55	56	60
$ni$	1	3	2	4	2	1	1	2

Le diagramme en bâton, soit un seul bâton par âge.



**Figure.** Répartition des salariés en fonction de l'âge

## TD 02 : Statistiques descriptives univariées

### 4. Déterminer les paramètres de position du caractère étudié.

a) Moyenne arithmétique :  $\bar{x} = \frac{1}{n} \sum_{i=0}^k x_i$  ou  $\bar{x} = \frac{1}{n} \sum_{i=0}^k x_i \cdot n_i$

$$\text{Moy} = 1/40 \times 1349 = 34,85 \text{ ans} \quad \text{---} \quad \text{Moy} = 1/40 \times (21 \times 7 + 23 \times 1 + \dots + 60 \times 2) = 34,85 \text{ ans}$$

b) Médiane on organise les données initiales du plus jeune au plus âgé :

21, 21, 21, 21, 21, 21, 21, 21, 23, 25, 25, 25, 25, 25, 25, 26, 29, 30, 30, 33, **33**,

**33**, 33, 33, 35, 37, 39, 39, 39, 41, 41, 46, 46, 46, 46, 54, 54, 55, 56, 60, 60

L'effectif total étant un nombre pair (40 salariés), on aura deux individus médians (21<sup>e</sup> et 22<sup>e</sup> individus) pour lesquels on détermine la médiane : **Médiane** =  $(33 + 33)/2 = 33 \text{ ans}$

Soit on utilise les effectifs cumulés

Age	21	23	25	26	29	30	<b>33</b>	35	37	39	41	46	54	55	56	60
$n_i$	7	1	6	1	1	2	5	1	1	3	2	4	2	1	1	2
$N_{\uparrow}$	7	8	14	15	16	18	<b>23</b>	24	25	28	30	34	36	37	38	40

c) Mode = **21 ans**

d) Autres paramètres de position : **Q1** = **25 ans** ; **Q3** =  $(41+46)/2 = 43,5 \text{ ans}$

Si on transforme ces données en classes d'âge (21-30 ; 31-40 ; 41-50 ; 51-60).

### 5. Combien y'a-t-il de modalités pour cette nouvelle série ?

Il y a 4 modalités.

### 6. Déterminer l'effectif de chaque modalité.

Modalités	$n_i$	$f_i$	$F_{i\uparrow}$	$C_i$
[21 – 30]	18	0,45	0,45	25,5
[31 – 40]	10	0,25	0,70	35,5
[41 – 50]	6	0,15	0,85	45,5
[51 – 60]	6	0,15	1	55,5

#### 4) Quelle est la classe modale ?

La classe modale correspond à l'intervalle qui présente l'effectif le plus important, soit l'intervalle [21 – 30].

### 5) Déterminer les paramètres de position de la nouvelle distribution

- La moyenne :  $\bar{x} = \sum_{i=0}^k f_i \cdot C_i$

$$\text{Moy} = (0,45 \times 25,5) + (0,25 \times 35,5) + (0,15 \times 45,5) + (0,15 \times 55,5) = 35,5 \text{ ans}$$

- La médiane : la classe médiane est [31 – 40] ; **M** =  $a + (b - a) \times [0,5 - F_a / F_b - F_a]$

$$M = 31 + (40 - 31) \times [0,5 - 0,45 / 0,70 - 0,45] = 32,8 \text{ ans}$$

- Le mode : classe modale est [21 – 30] le mode doit être le centre de la classe modale : **25,5 ans**

- Autres paramètres : **Q1** =  $21 + (30 - 21) \times [0,25 - 0 / 0,45 - 0] = 25,99 \text{ ans}$  ;

$$\text{Q3} = 41 + (50 - 41) \times [0,75 - 0,70 / 0,85 - 0,70] = 44 \text{ ans}$$

### 6) Que constatez-vous ? Débat ouvert.

### 7. Représentation graphique adéquat.

