

Séries regroupées par classes.

La leçon est en deux parties : mode d'emploi puis application.

Trouver des exercices plus pertinent par exemple des interprétations de séries de résultats d'une expérience scientifique.

I Les classes.

Afin de rendre les très grandes séries plus aisément manipulables sans moyen informatique nous utiliserons les *classes* pour réduire le nombre de données différentes à traiter. Plutôt que de regarder les 36 706 nombres d'habitants nous allons regrouper celles qui sont à peu près la même population.

D'après le travail fait sur ordinateur, pour la série des *populations municipales* les minimum et maximum sont : $Min = 0$ et $Max = 453\,317$. Donc les population municipales appartiennent toutes à l'intervalle $[0 ; 453\,317]$.

Les *classes* sont de plus petits intervalles dans lesquels on regroupera les *populations municipales* qui sont proches.

Définition 1

En statistiques, les *classes* sont des intervalles (souvent semi-ouverts) dans lesquels sont regroupés des valeurs (modalités) de la série.

Remarques.

1. Lorsque la série comporte un très grand nombre de modalités (valeurs distinctes de la série), regrouper par modalités ne simplifie pas la situation. C'est dans ce cas que l'on regroupe la série par classe.
2. Nous manipulerons la classe comme une modalité.

3. Nous associerons à la série des classes la série des effectifs (ou des fréquences) ainsi que des effectifs cumulés croissants (ou des fréquences cumulées croissantes).
4. Nous ne pourrions pas trouver les quartiles et médianes nous nous contenterons des classes des quartiles et médianes. Pour la moyenne c'est différent (voir *infra*).

La *moyenne* d'une série regroupée par classes est la moyenne pondérée de la série des *centres des classes* et des effectifs (ou fréquences associées). Les centres des classes sont les moyennes des bornes des classes : par exemple 4,5 est le centre de $[2; 7[$.

Deux représentations graphiques des séries des classes sont couramment utilisées l'histogramme et le polygone des fréquences cumulées croissantes.

Histogramme : sur l'axes des abscisses sont indiquées les bornes des classes et les effectifs sont représentés par des rectangles dont l'aire est proportionnelle aux effectifs. Il n'y a pas d'unité sur l'axe des ordonnées : ce sont les carrés du quadrillage qui représentent les effectifs.

Polygone des fréquences cumulées croissantes : obtenu en reliant par un segment les points ayant comme abscisse la borne supérieure de la classe et comme ordonnée les fréquences cumulées croissantes de la classe.

Exemples.

1. Considérons la répartition des lycées selon le nombre d'élèves :

Nombre d'élèves classes	Nombre de lycée	Nombre d'élèves classes	Nombre de lycée
$[0; 100[$	4	$[600; 700[$	121
$[100; 200[$	12	$[700; 800[$	149
$[200; 300[$	37	$[800; 900[$	143
$[300; 400[$	71	$[900; 1200[$	393
$[400; 500[$	105	$[1200; 1500[$	231
$[500; 600[$	125	1500 et plus	176

Déterminons la classe médiane.

Étape 1 - La série des classes est déjà ordonnée.

Étape 2 - Il y a $N = 4 + 12 + 37 + 71 + 105 + 125 + 121 + 149 + 143 + 393 + 231 + 176 = 1567$ valeurs (nombre impair). La médiane est donc la 784-ième $\left(\frac{N}{2} + 0,5\right)$ valeur.

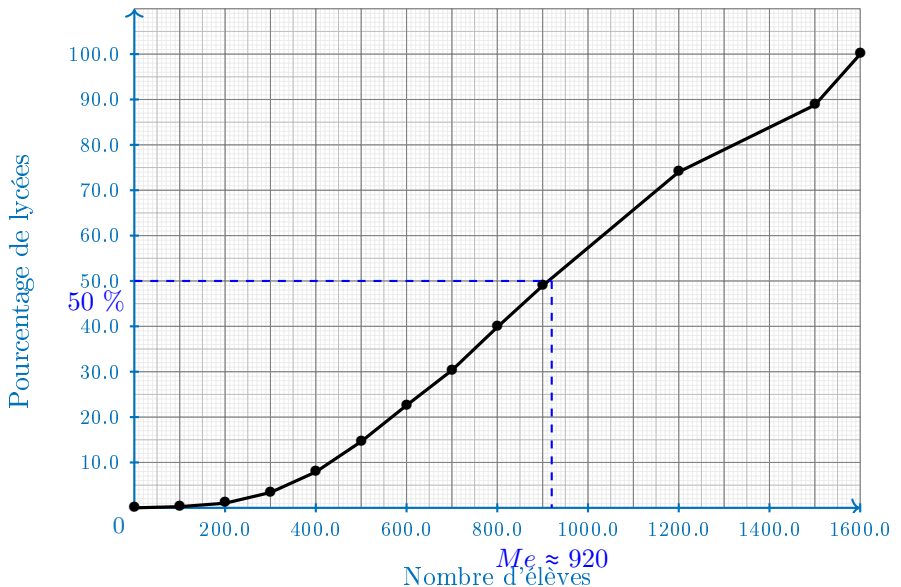
Étape 3 -

Nombre d'élèves classes	Effectif	E.C.C.	Nombre d'élèves classes	Effectif	E.C.C.
[0; 100[4	4	[600; 700[121	475
[100; 200[12	16	[700; 800[149	624
[200; 300[37	53	[800; 900[143	767
[300; 400[71	124	[900; 1200[393	1160
[400; 500[105	229	[1200; 1500[231	1391
[500; 600[125	354	1500 et plus	176	1567

La classe médiane est donc [900; 1200[.

2. Déterminons la médiane par lecture graphique sur le polygone des fréquences cumulées croissantes. Nous savons que la médiane correspond à 50% des effectifs.

Nombre d'élèves classes	Effectif	F.C.C (%)	Nombre d'élèves classes	Effectif	F.C.C. (%)
[0; 100[4	0,26	[600; 700[121	30,31
[100; 200[12	1,02	[700; 800[149	39,82
[200; 300[37	3,38	[800; 900[143	48,95
[300; 400[71	7,91	[900; 1200[393	74,03
[400; 500[105	14,61	[1200; 1500[231	88,77
[500; 600[125	22,59	1500 et plus	176	100



Séries regroupées par classes.

Exercice 1. ♥

hachette déclic 2014 page 178

Un hôpital cherche à tester l'efficacité d'un nouveau traitement oral pour des patients atteint de diabète de type 2. Pour cela , il étudie le taux de sucre dans le sang (glycémie en g/L) à jeun d'un groupe de 50 patients.

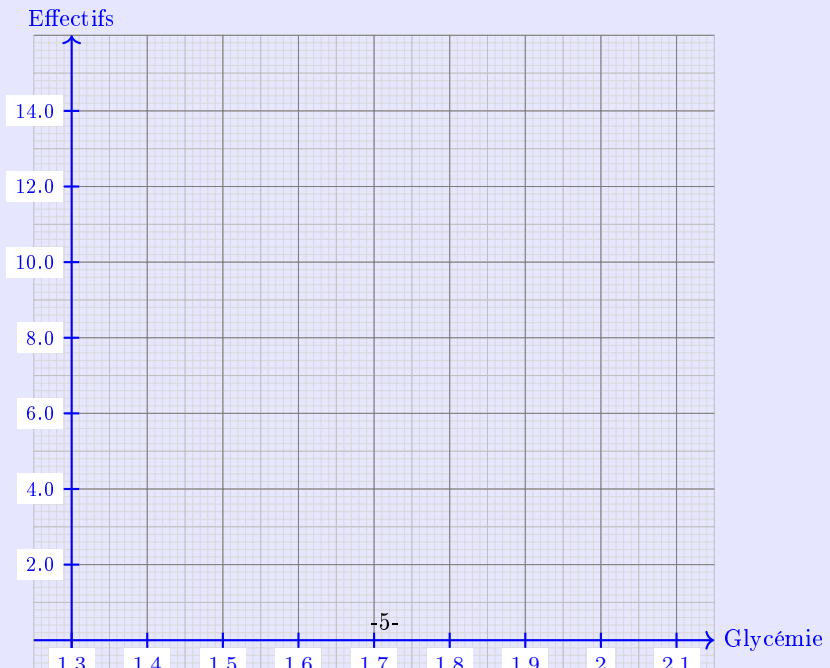
Les résultats sont répertoriés dans le tableau suivant.

1,73	2,07	1,53	1,47	2,05	1,41	1,84	1,97	1,82	1,8
1,53	1,6	2,05	1,53	1,69	1,61	1,92	1,45	1,68	1,94
1,62	1,73	1,47	1,34	1,53	1,61	1,78	1,64	1,61	1,61
1,82	2,02	2,02	1,73	1,77	1,99	1,68	1,6	1,9	1,39
1,7	1,45	1,68	1,82	1,96	1,91	1,7	1,6	1,48	1,56

1. Complétez le tableau d'effectifs du groupe en regroupant les résultats en classe d'amplitude 0,1.

Glycémie	[1,3; 1,4[[1,4; 1,5[[1,5; 1,6[[1,6; 1,7[
Effectifs				
Glycémie	[1,7; 1,8[[1,8; 1,9[[1,9; 2[[2; 2,1[
Effectifs				

2. Déduisez-en la moyenne de la série regroupée par classe.
3. Construisez l'histogramme correspondant.

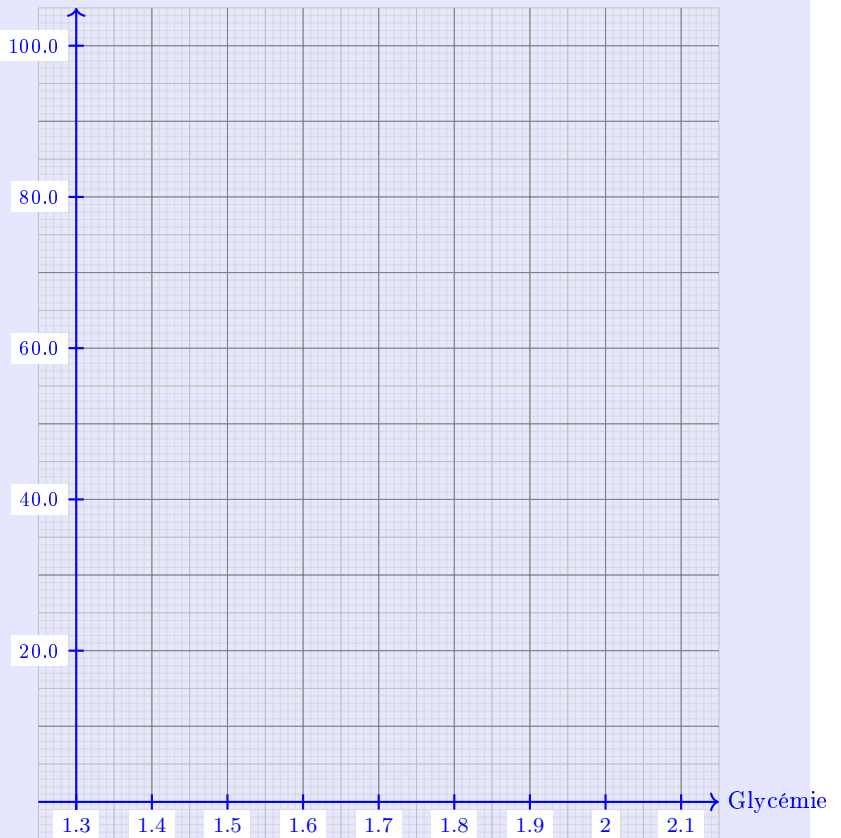


Exercice 1.

(SUITE)

4. (a) Dressez le tableau des fréquences et fréquences cumulées croissantes de la série des classes.
 (b) Déterminez la classe médiane.
 (c) Dessinez le polygone des fréquences cumulées croissantes.

F.C.C. (en %)



- (d) Par lecture graphique sur le polygone des fréquences cumulées croissantes déterminez la médiane de la série.

Correction exercice 1

1.

Glycémie	[1,3; 1,4[[1,4; 1,5[[1,5; 1,6[[1,6; 1,7[
Effectifs	2	6	5	13
Glycémie	[1,7; 1,8[[1,8; 1,9[[1,9; 2[[2; 2,1[
Effectifs	7	5	7	5

2. Calculons \bar{x} .

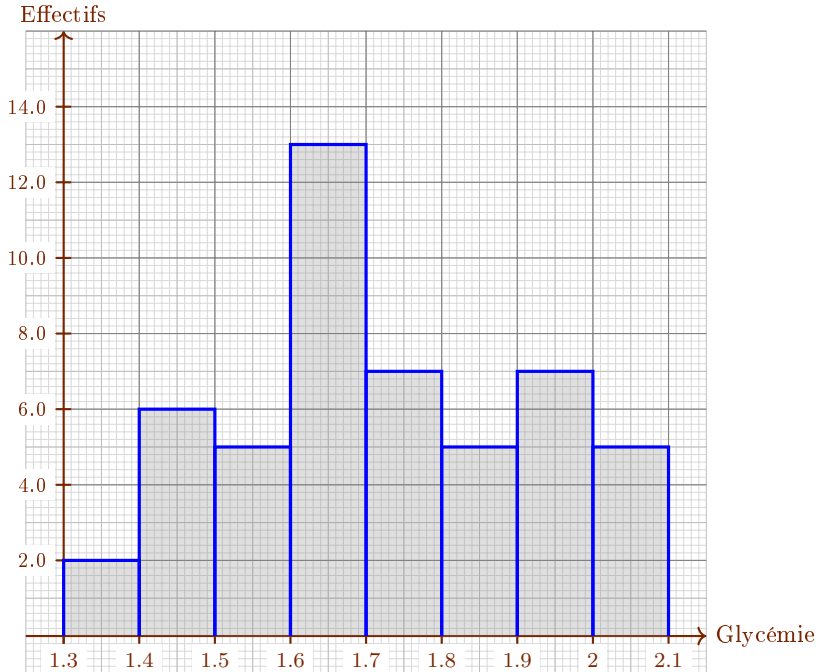
Les centres des classes sont, dans l'ordre, 1,35, 1,45, 1,55, 1,65, 1,75, 1,85, 1,95, 2,05. En utilisant la formule de la moyenne pondérée

$$\begin{aligned}\bar{x} &= \frac{n_1x_1 + n_2x_2 + \dots + n_r x_r}{n_1 + n_2 + \dots + n_r} \\ &= \frac{2 \times 1,35 + 6 \times 1,45 + \dots + 5 \times 2,05}{2 + 6 + \dots + 5} \\ &= 1,72\end{aligned}$$

Donc

$$\bar{x} = 1,72.$$

3. En règle générale un histogramme est pénible à construire car c'est la surface du rectangle qui représente l'effectif. Comme toutes les classes ont ici la même taille la construction est la même que celle d'un diagramme en barres.



4. (a) La fréquence est le quotient de l'effectif sur l'effectif total.

Glycémie	[1,3; 1,4[[1,4; 1,5[[1,5; 1,6[[1,6; 1,7[
Effectifs	2	6	5	13
Fréquences (%)	4	12	10	26
F.C.C	4	16	26	52

Glycémie	[1,7; 1,8[[1,8; 1,9[[1,9; 2[[2; 2,1[
Effectifs	7	5	7	5
Fréquences	14	10	14	10
F.C.C.	66	76	90	100

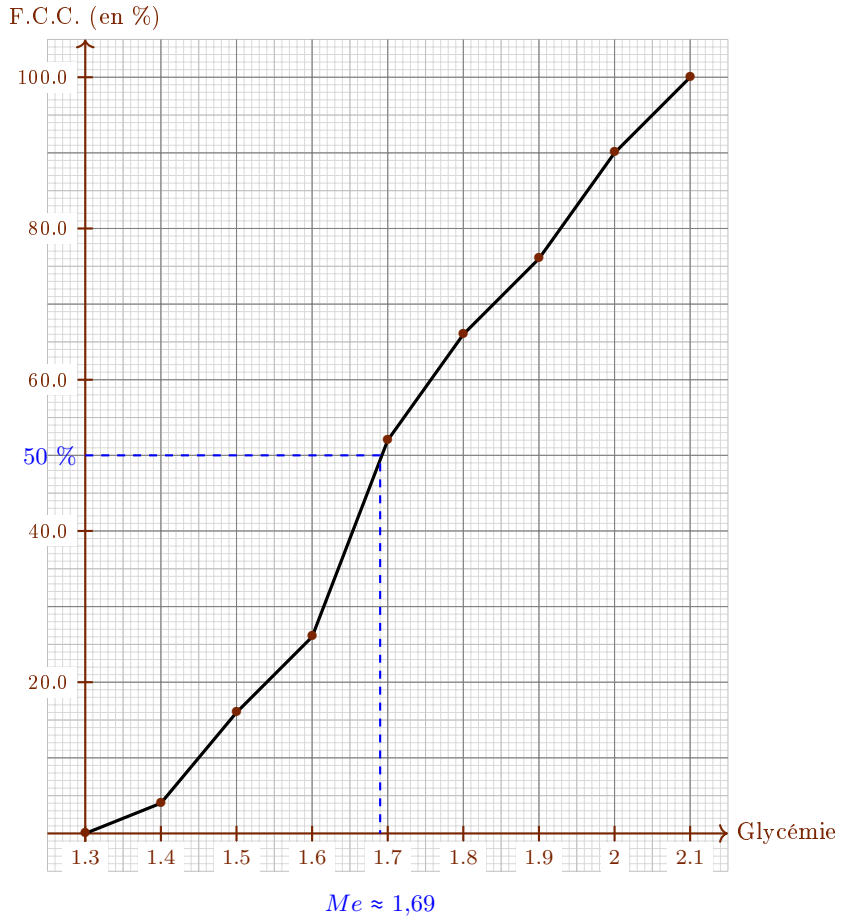
- (b) Déterminons la classe médiane.

Nous retrouvons les trois mêmes étapes que pour la recherche de la médiane.

- * Les classes sont déjà ordonnées.
- * La médiane correspond à 50 % des effectifs. Comme nous travaillons déjà avec des pourcentages il n'y a pas de calcul supplémentaire requis.
- * D'après les fréquences cumulées croissantes la médiane (obtenue pour 50 %) est dans [1,6; 1,7[.

La classe médiane est [1,6; 1,7[.

- (c) Traçons le polygone des fréquences cumulées croissantes.



(d) Par lecture graphique du polygone des fréquences cumulées croissantes :

$Me \approx 1,69.$

II Exercices.

Pour comparer des séries très différentes en effectifs on utilise les fréquences.
Exercice mettant en jeu cette situation.

Utilité : comparer des séries statistiques (double affichage, déterminer des valeurs approchées des médianes et quartiles.

Exercice là aussi.

Exercice 2.

Une étude menée auprès des consommateurs mesure la durée de vie d'ampoules de deux marques différentes que l'on souhaite comparer.

Les données recueillies sont les suivantes :

Durée de vie en milliers d'heures.	[0 ; 4[[4 ; 6[[6 ; 8[[8 ; 10[[10 ; 12]
Nombre d'ampoules Fiateloux	12	20	34	60	75
Nombre d'ampoules Ténébraé	234	432	309	456	607

1. Peut-on comparer les nombres d'ampoules des deux marques entre eux ?
2. Faites un tableau, pour chacune des marques, en remplaçant le nombre d'ampoules par la fréquence.
3. Dessinez dans un repère les deux polygones des fréquences cumulées croissantes.
4. Un chef d'entreprise souhaite acheter les ampoules qui dure le plus longtemps en se basant sur la médiane. Quelle marque va-t-il choisir ?

Correction exercice 2

1. Il n'est pas possible de comparer directement les données car le nombre total d'ampoules n'est pas le même pour les deux marques.
2. Les fréquences s'obtiennent comme le quotient de l'effectif par l'effectif total donc :

Durée de vie en milliers d'heures.	[0 ; 4[[4 ; 6[[6 ; 8[[8 ; 10[[10 ; 12]
Nombre d'ampoules Fiateloux	12	20	34	60	75
Fréquence Fiateloux	0,0597	0,0995	0,1692	0,2985	0,3731
Nombre d'ampoules Ténébraé	234	432	309	456	607
Fréquence Ténébraé	0,1148	0,212	0,1516	0,2237	0,2978

Pour gagner du temps il est possible de faire des formules de calcul directement avec les listes sur la calculatrice. Si les effectifs sont dans la liste L_1

L1	L2	L3	L4	L5	2
12					
20					
34					
60					
75					

$L_2(1)=$

nous pouvons faire apparaître les effectifs cumulés croissants dans la liste L_2 en faisant sur la page d'accueil de la calculatrice : `2nde`, `stats`, (OP), "6 : somCum", `2nde`, `1`, `entrer` :

NORMAL FLOTT AUTO RÉEL DEGRÉ MP 
`somCum(L1)→L2`

Nous obtenons alors dans l'éditeur de liste :

Séries regroupées par classes.

L1	L2	L3	L4	L5	2
12	12				
20	32				
34	66				
60	126				
75	201				

L2(1)=12

Nous voyons que l'effectif total est 201 et nous calculons alors les fréquences en L_3 en tapant sur la page d'accueil de la calculatrice : [2nde], [1], [÷], [2], [0], [1], [→], [2nde], [3], [entrer] :

NORMAL FLOTT AUTO RÉEL DEGRÉ MP	
somCum(L1)→L2	(12 32 66 126 201)
L1/201→L3	(0,0597014925 0,099502487)

Et en retournant dans l'éditeur de listes :

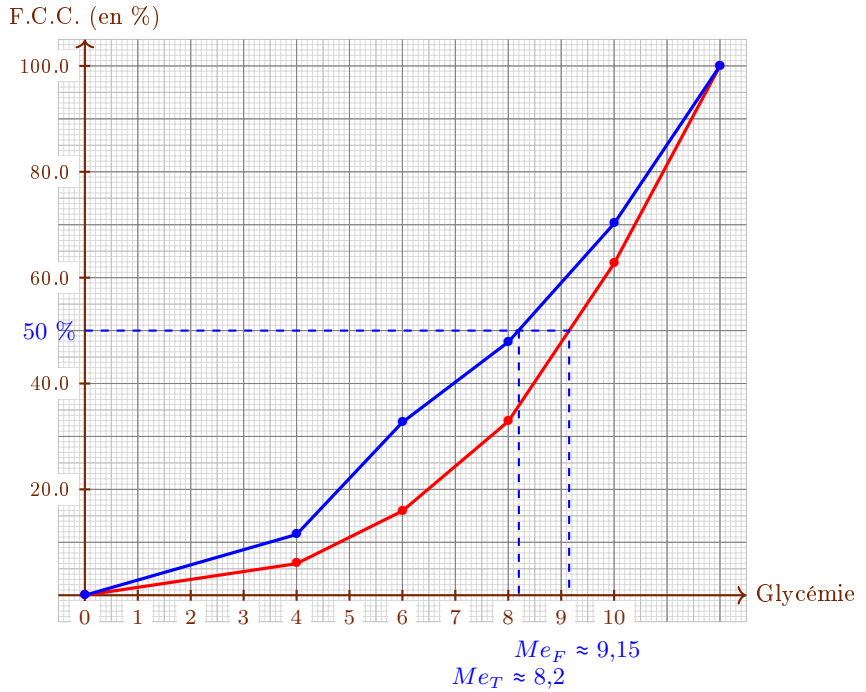
L1	L2	L3	L4	L5	3
12	12	0,0597			
20	32	0,0995			
34	66	0,1692			
60	126	0,2985			
75	201	0,3731			

L3(6)=

3. On peut là encore faire le travail à la calculatrice :

Durée de vie en milliers d'heures.	[0 ; 4[[4 ; 6[[6 ; 8[[8 ; 10[[10 ; 12]
Fréquence Fiateloux	0,0597	0,0995	0,1692	0,2985	0,3731
FCC Fiateloux	0,0597	0,1592	0,3284	0,6269	1
Fréquence Ténébraé	0,1148	0,212	0,1516	0,2237	0,2978
FCC Ténébraé	0,1148	0,3268	0,4784	0,7022	1

4.



5. Par lecture graphique la médiane de la marque Fiatlux est $Me_F = 9,15$ milliers d'heures et celle de la marque Ténébraé est $Me_T = 8,2$ milliers d'heures.

Il choisira la marque Fiatlux.

Exercice 3.

Exercice 37 page 209 du manuel déclic de première ES.

En 2010 un recensement de la population en Corse a permis de répertorier l'âge des femmes.

Age	[0 ; 10[[10 ; 20[[20 ; 30[[30 ; 40[
Effectif	16 115	17 476	17 698	19 574

Age	[40 ; 50[[50 ; 60[[60 ; 70[[70 ; 80[
Effectif	22 222	21 455	18 106	12 537

Age	[80 ; 90[[90 ; 100[[100 ; 110[
Effectif	5 692	768	22

Calculez la moyenne de cette série.

Correction exercice 3

Calculons \bar{x} .

Pour calculer le centre d'une classe $[a; b[$ (ou plus généralement le centre d'un intervalle) il faut la moyenne des bornes : $c = \frac{a+b}{2}$.

Les centres des classes sont, dans l'ordre, 5, 15, 25, 35, 45, 55, 65, 75, 85, 95, 105. En utilisant la formule de la moyenne pondérée

$$\begin{aligned}\bar{x} &= \frac{n_1 c_1 + n_2 c_2 + \dots + n_r c_r}{n_1 + n_2 + \dots + n_r} \\ &= \frac{16\,115 \times 5 + 17\,476 \times 15 + \dots + 22 \times 105}{16\,115 + 17\,476 + \dots + 22} \\ &= 1,72\end{aligned}$$

Donc

$$\bar{x} \approx 41,7151.$$

Exercice 4.

Exercice 29 page 19 du manuel Sésamath questions 1 et 2 : regrouper les données par classe (super long), moyenne pondérée d'une série regroupée par classes.

Exercice 30 page 19 du manuel Sésamath : moyenne pondérée d'une série regroupée par classes.

III Ce qu'il faut retenir.

1. Lorsqu'une série est présentée avec des classes penser à utiliser les effectifs et les effectifs cumulés croissants pour la médiane.
2. Déterminer les classes médianes et classe des quartiles, ainsi que la moyenne (avec le centre des classes).
3. Représentation graphiques : histogramme et polygone des fréquences cumulées croissantes.
4. Déterminer quartiles et médiane avec le polygone des fréquences cumulées croissantes.