On peut simplifier cette formule à l'aide de la notation sigma, symbolisée par « Σ » (S dans l'alphabet grec). Ce symbole indique que l'on doit faire la somme de tous les termes de forme x_i, l'indice i variant de 1 à *N*.

$$\mu = \frac{\Sigma x_i}{N}$$

Dans le cadre d'un sondage, on emploie le symbole $\overline{x}$ pour désigner la moyenne des données de l'échantillon et μ pour la moyenne des données de la population. De même, on utilise le symbole n pour désigner le nombre de données de l'échantillon et N pour le nombre de données de la population. La formule pour obtenir la moyenne des données de l'échantillon s'écrit donc :

$$\overline{x} = \frac{\Sigma x_i}{n}$$

Pour la moyenne des données de la mise en situation, on a :

$$\mu = \frac{\Sigma x_i}{N} = \frac{4 + 3 + 5 + \cdots + 3 + 5}{48} = \frac{223}{48} \approx 4{,}6 \text{ programmes préuniversitaires par collège}$$

<u>Interprétation :</u> En 2006, si tous les collèges publics du Québec avaient offert le même nombre de cours, il y aurait eu 4,6 programmes préuniversitaires dans chaque collège.

Attention!!: Le résultat du calcul d'une moyenne brute ne doit pas être arrondi à l'entier sous prétexte que les données brutes sont entières : la moyenne est un nombre théorique. <u>Nous conviendrons de conserver une décimale après la virgule.</u>

ii. Calcul de la moyenne avec les effectifs du tableau de distribution

Répartition des 48 collèges publics du Québec selon le nombre de programmes préuniversitaires offerts, 2006

Nombre de programmes préuniversitaires	Nombre de collèges	Pourcentage de collèges
3	10	20,8 %
4	14	29,2 %
5	12	25,0 %
6	7	14,6 %
7	5	10,4 %
Total	48	100,0 %

Avec les effectifs du tableau de distribution, on calcule ainsi la moyenne du nombre de programmes préuniversitaires offerts par les collèges publics du Québec :

Moyenne avec des effectifs
$\text{Moyenne} = \dfrac{\text{somme des produits de chaque valeur de la variable par son effectif}}{\text{nombre total de données}}$

Écriture symbolique :

En notant par la lettre n_i l'effectif correspondant à la valeur x_i, on obtient la formule suivante : $\mu = \frac{x_1n_1+x_2n_2+x_3n_3+\cdots+x_kn_k}{N}$, où k représente le nombre de valeurs de la variable.

En utilisant la notation sigma, on obtient :

$$\mu = \frac{\Sigma x_i n_i}{N}, \text{ pour } i \text{ variant de 1 à } k.$$

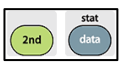
Dans l'exemple, on obtient donc :

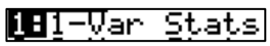
$$\mu = \frac{\Sigma x_i n_i}{N} = \frac{3 \times 10 + 4 \times 14 + 5 \times 12 + 6 \times 7 + 7 \times 5}{48} = \frac{223}{48} \approx 4{,}6$$

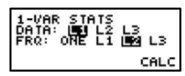
Entrer les données dans la calculatrice TI-30XS MultiView

1- 

: Entrer les modalités dans la première colonne et les effectifs dans la deuxième colonne

2- 

→ 

→ 

iii. Calcul de la moyenne avec les pourcentages du tableau de distribution

Pour trouver une formule qui permettrait de calculer une moyenne avec les pourcentages, reprenons le calcul de la moyenne avec les effectifs et apportons les modifications suivantes :

$$\mu = \frac{3\times10+4\times14+5\times12+6\times7+7\times5}{48} = 4{,}6$$

$$\mu = \frac{3\times10}{48}+\frac{4\times14}{48}+\frac{5\times12}{48}+\frac{6\times7}{48}+\frac{7\times5}{48} = 4{,}6$$

$$\mu = 3\times\frac{10}{48}+4\times\frac{14}{48}+5\times\frac{12}{48}+6\times\frac{7}{48}+7\times\frac{5}{48} = 4{,}6$$

$$\mu = 3\times20{,}8\,\% + 4\times29{,}2\,\% + 5\times25\,\% + 6\times14{,}6\,\% + 7\times10{,}4\,\% = 4{,}6$$

Moyenne avec des pourcentages
$\text{Moyenne} = \begin{matrix}\text{somme des produits de chaque valeur} \\ \text{de la variable par son pourcentage}\end{matrix}$

<u>Attention!!:</u> Il est à remarquer que nous n'avons pas à diviser, dans ce cas-ci, par le total des données, car cette division a déjà été faite dans le calcul du pourcentage.

Écriture symbolique
En notant f_i la fréquence relative correspondant à la valeur x_i, on obtient la formule suivante :

$$\mu = x_1 f_1 + x_2 f_2 + x_3 f_3 + \cdots + x_k f_k = \Sigma x_i f_i$$

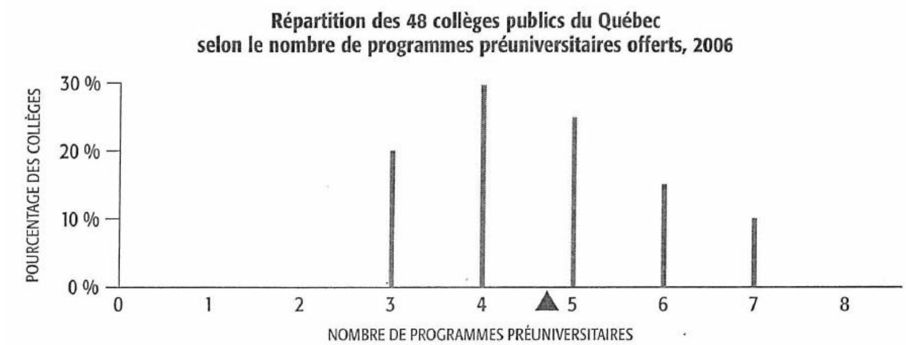
pour i variant de 1 à k, où k est le nombre de valeurs de la variable.

Dans l'exemple, on obtient donc :

$$\mu = \Sigma x_i f_i = 3 \times 20{,}8\ \% + 4 \times 29{,}2\ \% + 5 \times 25\ \% + 6 \times 14{,}6\ \% + 7 \times 10{,}4\ \% \approx 4{,}6$$

iv. Représentation graphique de la moyenne

Voici le diagramme en bâtons de la distribution du nombre de programmes préuniversitaires :



En plaçant un pivot sous l'axe horizontal du diagramme en bâtons à l'endroit où se situe la moyenne, on remarque que la moyenne correspond graphiquement au centre d'Équilibre du diagramme. On applique ici le principe des balançoires à bascule : on doit s'imaginer que l'axe horizontal est une planche sous laquelle il faut placer un pivot pour que les bâtons du diagramme se trouvent en position d'équilibre sur cette planche.

Exemple 1 :

Répartition des ménages selon le nombre de voitures, Québec, 2006

Nombre de voitures	Pourcentage
0	25 %
1	38 %
2	32 %
3 et plus	5 %
Total	100 %

Source : *L'actualité*, mai 2007.

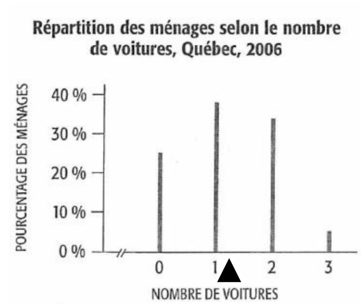
a) Calculez et interprétez la moyenne en remplaçant la catégorie « 3 et plus » par le nombre 3.

$$\mu = 0 \times 25\ \% + 1 \times 38\ \% + 2 \times 32\ \% + 3 \times 5\ \% = 1{,}17$$

Si tous les ménages québécois avaient le même nb de voitures, ils en auraient 1,2.

b) Représentez graphiquement la moyenne en plaçant le pivot au bon endroit.

(on peut procéder par essai-erreur...)

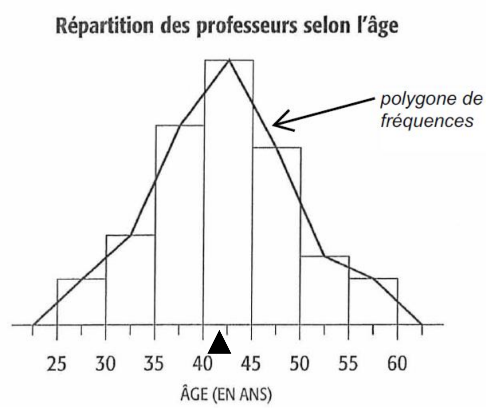


Exemple 2 : À la suite de l'étude menée auprès des 40 professeurs d'une école, nous avons construit le tableau de distribution et l'histogramme suivant.

Répartition des professeurs selon l'âge

Âge (en ans)	Nombre de professeurs
$25 \leq X < 30$	2
$30 \leq X < 35$	4
$35 \leq X < 40$	9
$40 \leq X < 45$	12
$45 \leq X < 50$	8
$50 \leq X < 55$	3
$55 \leq X < 60$	2
Total	40

(25+30)/2 = 27,5



a) Calculer et interpréter la moyenne de la distribution.

$$\mu = \frac{27,5 \times 2 + 32,5 \times 4 + \cdots + 57,5 \times 2}{40} = \frac{1685}{40} = 42,1$$

Si tous les profs avaient le même âge, ils auraient 42,1 ans.

b) Placer un pivot sous l'axe horizontal du graphique afin d'estimer la moyenne de la distribution.

Attention!!: Avec un histogramme ou un tableau de données groupées, la moyenne n'est pas exacte, mais c'est le mieux que l'on peut faire.

v. Moyenne pondérée

Lorsque le calcul d'une moyenne se fait en multipliant chaque valeur d'une série par sa pondération, on dit que l'on calcule la moyenne pondérée de la série. La pondération est déterminée selon l'importance qui est accordée à chaque valeur par rapport aux autres valeurs de la série. La somme des pondérations doit toujours être égale à 100 %.

Exemple 1 : Supposons qu'un étudiant ait obtenu les notes suivantes en français :
Examen 1 : 65 % Examen 2 :70 % Travail : 75 %

Trouvez sa moyenne pour la session si la pondération de chaque évaluation est :
Examen 1 : 25 % Examen 2 : 35 % Travail : 40 %

$$\text{Moyenne} = \boxed{65 \times 25\,\% + 70 \times 35\,\% + 75 \times 40\,\% = 70{,}8\,\%}$$

Exemple 2 :

Revenu moyen des familles monoparentales par province, Canada, 2005

Terre-Neuve-et-Labrador	22 300 $	Ontario	40 300 $
Île-du-Prince-Édouard	23 100 $	Manitoba	30 300 $
Nouvelle-Écosse	28 300 $	Saskatchewan	22 800 $
Nouveau-Brunswick	22 100 $	Alberta	41 000 $
Québec	35 300 $	Colombie-Britannique	32 400 $

Source : Statistique Canada, CANSIM, tableau 202-0202.

a) Calcule la moyenne de ces 10 revenus.

$$\mu = \frac{22\,300 + 23\,100 + \ldots + 32\,400}{10} = 29\,790\ \$$$

b) Peut-on dire que la moyenne trouvée ci-haut est le revenu moyen des familles monoparentales vivant dans les provinces canadiennes? Justifier.

Non, car le nombre de familles monoparentales par province n'est pas le même ; il faut donc pondérer ces revenus.

c) Utilisons le tableau suivant pour pondérer le calcul de la moyenne.

Répartition des familles monoparentales par province, Canada, 2005

Terre-Neuve-et-Labrador	1,8 %	Ontario	39,3 %
Île-du-Prince-Édouard	0,5 %	Manitoba	3,9 %
Nouvelle-Écosse	3,5 %	Saskatchewan	3,5 %
Nouveau-Brunswick	2,6 %	Alberta	9,7 %
Québec	23,0 %	Colombie-Britannique	12,2 %

Source : Statistique Canada, CANSIM, tableau 111-0009.

$$\mu = 22\,300\,\$ \times 1{,}8\,\% + 23\,100\,\$ \times 0{,}5\,\% + \ldots + 32\,400\,\$ \times 12{,}2\,\% = \mathit{35\ 948\$}$$

Bref, le fait que 39,3 % des familles monoparentales vivent en Ontario et que ces dernières ont, après l'Alberta, le revenu moyen le plus élevé (40 300 \$) fera augmenter la valeur trouvée pour μ à la question a).

d) Interpréter la moyenne pondérée.

Si toutes les familles monoparentales avaient le même revenu, elles gagneraient 35 948\$.

Exemple 3 : Calculer le salaire moyen des 200 employés d'une usine sachant que les 10 cadres gagnent en moyenne 45 200 \$, les 50 techniciens 35 250 \$ et que les 140 ouvriers 28 400 \$. Interpréter la moyenne et dire si elle est représentative.

$$\begin{aligned}\text{Salaire moyen } &= 45\,200 \times \frac{10}{200} + 35\,250 \times \frac{50}{200} + 28\,400 \times \frac{140}{200} \\ &= 45\,200 \times 5\,\% + 35\,250 \times 25\,\% + 28\,400 \times 70\,\% \\ &= 30\,952{,}50\,\$\end{aligned}$$

Si tous les employés de l'usine gagnaient le même salaire, ils gagneraient 30 952,50\$.

Non, la moyenne n'est pas représentative, car seulement 30% des employés gagnent plus que la moyenne.