

b) Poids à la naissance de 10 nouveau-nés :

2 350 g	3 150 g	3 252 g	3 334 g	3 552 g
3 843 g	3 926 g	4 125 g	4 650 g	3 684 g

### Solution

Attention, il faut mettre les données en ordre.

La médiane est égale à la moyenne de la 5<sup>e</sup> et de la 6<sup>e</sup> donnée :

$$\text{Médiane} = \frac{3\,552 + 3\,684}{2} = 3\,618 \text{ g}$$

### Interprétation

50 % des nouveau-nés pèsent moins de 3 618 grammes.

## Données groupées par valeurs

On peut déterminer la médiane soit avec les effectifs, soit avec les pourcentages associés aux valeurs.

- Pour déterminer la médiane avec les effectifs :  
On applique la même procédure que pour des données brutes.
- Pour déterminer la médiane avec les pourcentages :
  1. On repère la première valeur associée à un pourcentage cumulé égal ou supérieur à 50 %.
  2. Si le cumul des données est supérieur à 50 %, la médiane est égale à la valeur repérée.  
Si le cumul des données est égal à 50 %, la médiane est la moyenne de la valeur repérée et de la valeur suivante.

### EXEMPLE

a) Déterminer et interpréter la médiane de la distribution suivante.

#### Répartition des Centres de la petite enfance (CPE) d'une ville selon le nombre d'éducatrices

Nombre d'éducatrices	4	5	6	7	Total
Nombre de CPE	3	6	4	2	15
Pourcentage de CPE	20,0 %	40,0 %	26,7 %	13,3 %	100,0 %

### Solution

- Calcul de la médiane avec les effectifs :  
Avec un total de 15 données, un nombre impair, la médiane est la valeur de la 8<sup>e</sup> donnée.  
Médiane = 5 éducatrices
- Calcul de la médiane avec les pourcentages :  
Pour obtenir un pourcentage cumulé d'au moins 50 %, il faut additionner les pourcentages associés aux valeurs 4 et 5, soit 20 % + 40 %, ce qui donne 60 %, une valeur supérieure à 50 %.  
Médiane = 5 éducatrices

### Interprétation

Au moins 50 % des CPE de cette ville ont 5 éducatrices ou moins.

b) Déterminer et interpréter la médiane de la distribution suivante.

Répartition de 160 entreprises selon le nombre de cadres

Nombre de cadres	2	3	4	5	6	Total
Nombre d'entreprises	16	24	40	48	32	160
Pourcentage d'entreprises	10 %	15 %	25 %	30 %	20 %	100 %

**Solution**

- Calcul de la médiane avec les effectifs :

Avec un total de 160 données, un nombre pair, la médiane est égale à la moyenne de la 80<sup>e</sup> et de la 81<sup>e</sup> donnée.

$$\text{Médiane} = \frac{4 + 5}{2} = 4,5 \text{ cadres}$$

- Calcul de la médiane avec les pourcentages :

Pour obtenir un pourcentage cumulé d'au moins 50 %, il faut additionner les pourcentages associés aux valeurs 2, 3 et 4, soit 10 % + 15 % + 25 %, ce qui donne exactement 50 %. La médiane est la moyenne des valeurs 4 et 5.

$$\text{Médiane} = \frac{4 + 5}{2} = 4,5 \text{ cadres}$$

En raison du contexte, on utilisera 4 (ou 5) au lieu de 4,5 dans l'interprétation.

*Interprétation*

50 % des entreprises ont 4 cadres ou moins (ou 5 cadres ou plus).

**EXERCICES DE COMPRÉHENSION | 1.8**

1. Vrai ou faux ? Si la médiane d'un examen sur 100 points est 68, on peut alors dire qu'au moins 50 % des étudiants ont une note de 68. Faux, une note de 68 ou moins.
2. Calculer et interpréter la médiane des distributions suivantes.

- a) **Répartition des employés qui ont pris un congé de maladie dans la dernière année selon le nombre de jours d'absence**

Nombre de jours	1	2	3	4	5 et plus	Total
Nombre d'employés	9	6	7	3	5	30

$$\text{Médiane} = \frac{2 + 3}{2} = 2,5 \text{ jours}$$

*Interprétation*

50 % des employés qui ont pris un congé de maladie dans la dernière année se sont absentés 2 jours ou moins (ou 3 jours ou plus).

- b) Un sondage effectué auprès des étudiants d'un cégep a permis de construire le tableau suivant.

**Répartition des répondants selon le nombre de repas pris à la cafétéria durant la dernière semaine**

Nombre de repas	0	1	2	3	4	5	Total
Pourcentage de répondants	30 %	15 %	15 %	18 %	12 %	10 %	100 %

Médiane = 2 repas

*Interprétation*

Au moins 50 % des répondants ont pris 2 repas ou moins à la cafétéria au cours de la dernière semaine.

3. Donner et interpréter la meilleure mesure de tendance centrale pour la distribution suivante.

**Répartition des 50 albums<sup>1</sup> les plus vendus selon la langue d'enregistrement, Québec, 2012**

Langue d'enregistrement	Anglais	Français	Autres	Total
Nombre d'albums	35	14	1	50

1. Inclut les albums sur support physique et les albums numériques.

Source: Observatoire de la culture et des communications du Québec, Mai 2013.

À titre d'information, les 5 premiers albums de la liste des 50 albums les plus vendus au Québec en 2012 sont: *Sans attendre* (Céline Dion), *Star Académie 2012* (Artistes variés), *21* (Adele), *Star Académie Noël* (Artistes variés), *Mes amours mes amis* (Paul Daraïche).

Meilleure mesure de tendance centrale :

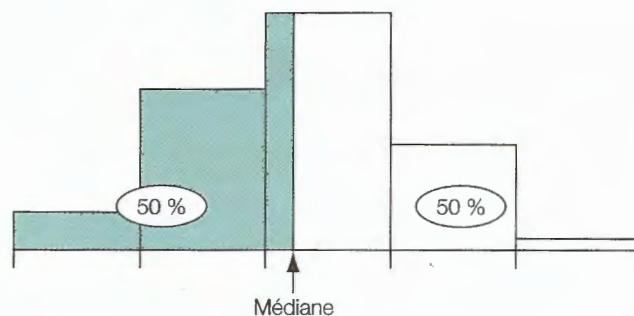
Mode = Anglais

*Interprétation*

En 2012, une majorité (70 %) des 50 albums les plus vendus au Québec sont enregistrés en anglais.

## Données groupées en classes

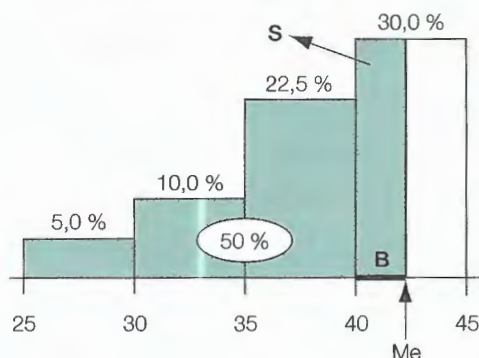
Lorsque les données d'une série statistique sont groupées en classes, la médiane est égale à la valeur sur l'axe horizontal qui divise la surface de l'histogramme (donc les données, en vertu du principe de la proportionnalité) en deux parties égales.





Cherchons la médiane de la distribution suivante.

Esquisse de l'histogramme



Répartition des 40 clients de l'échantillon selon la durée du branchement au réseau Wi-Fi

Durée du branchement (en minutes)	Pourcentage de clients
$25 \leq X < 30$	5,0 %
$30 \leq X < 35$	10,0 %
$35 \leq X < 40$	22,5 %
$40 \leq X < 45$	30,0 %
$45 \leq X < 50$	20,0 %
$50 \leq X < 55$	7,5 %
$55 \leq X < 60$	5,0 %
<b>Total</b>	<b>100,0 %</b>

Notre objectif est de trouver une durée de branchement sur l'axe horizontal tel que 50 % de la surface de l'histogramme (donc 50 % des données) se situe à gauche de ce nombre. Pour obtenir 50 % de la surface, il faut additionner la surface du 1<sup>er</sup> rectangle, du 2<sup>e</sup> rectangle, du 3<sup>e</sup> rectangle et une partie S, de la surface du 4<sup>e</sup> rectangle.

- Surface :  $50 \% = (5 \% + 10 \% + 22,5 \%) + S$

$$50 \% = 37,5 \% + S$$

$$S = 50 \% - 37,5 \%$$

$$S = 12,5 \%$$

- Médiane = 40 min + B min, où B est la longueur de la base du rectangle de surface S. On trouve la valeur B par une règle de trois établissant un rapport entre la surface et la base du 4<sup>e</sup> rectangle et la surface S et la base B du rectangle construit à l'intérieur de ce rectangle.

Surface

Base

30,0 % → 5 min

12,5 % → B min

$$\text{D'où } B = \frac{12,5 \times 5}{30} = 2,1 \text{ min}$$

$$\text{Médiane} = 40 \text{ min} + 2,1 \text{ min} = 42,1 \text{ min}$$

### Interprétation

On peut estimer que, pour 50 % des clients de l'échantillon, le temps de branchement au réseau Wi-Fi a duré moins de 42,1 minutes.

### EXEMPLE

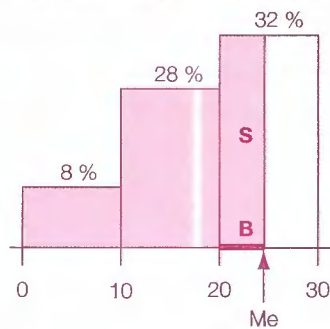
Le tableau suivant donne la distribution de l'âge des arbres recensés sur un terrain boisé. Trouver et interpréter la médiane de cette distribution.

Répartition des arbres selon l'âge

Âge (en ans)	Pourcentage
[0 ; 10[	8 %
[10 ; 20[	28 %
[20 ; 30[	32 %
[30 ; 40[	20 %
[40 ; 50[	12 %
<b>Total</b>	<b>100,0 %</b>

## Solution

### Esquisse de l'histogramme



- Surface:  $50 \% = (8 \% + 28 \%) + S$   
 $50 \% = 36 \% + S$   
 $S = 50 \% - 36 \%$   
 $S = 14 \%$

- Médiane = 20 ans + B ans

On a: 

<u>Surface</u>		<u>Base</u>
32 %	→	10 ans
14 %	→	B ans

D'où  $B = \frac{14 \times 10}{32} = 4,4$  ans

Médiane = 20 ans + 4,4 ans = 24,4 ans

### Interprétation

On peut estimer que 50 % des arbres ont moins de 24,4 ans.

## Quelle mesure de tendance centrale faut-il utiliser ?

Chaque mesure de tendance centrale présente des avantages et des inconvénients. La moyenne est certainement la mesure la plus couramment utilisée pour représenter une série de données, mais on ne peut pas l'employer si la variable est qualitative. Son principal inconvénient est qu'elle peut être influencée par quelques valeurs extrêmes de la série statistique. Dans ce cas, on choisit la médiane comme mesure de tendance centrale.

La médiane est une mesure intéressante, car elle donne le centre de la distribution. Une différence importante entre la moyenne et la médiane indique que certaines données de la série statistique sont beaucoup plus grandes ou beaucoup plus petites que les autres.

Le mode est la mesure de tendance centrale qu'il faut utiliser si la variable est qualitative. Dans le cas d'une variable quantitative, le mode est significatif uniquement si sa fréquence est élevée.

Voici un tableau dans lequel on associe mesures de tendance centrale et types de variable.

Type de variable	Mesure possible
Variable qualitative	Mode
Variable quantitative	Moyenne Mode Médiane

## 1.4 Les mesures de position (quantiles)

Les mesures de position que sont les quantiles et la cote  $z$  servent à situer une donnée par rapport aux autres données d'une série statistique. Dans la présente section, nous étudions les quantiles. Par contre, l'étude de la cote  $z$  sera présentée à la section 1.6, après l'élaboration des outils mathématiques nécessaires à sa compréhension. La mise en situation suivante sert à illustrer le concept de quantile pour des données groupées en classes.

### MISE EN SITUATION

Reprenons la distribution de l'âge des arbres recensés dans un boisé.

Répartition des arbres selon l'âge

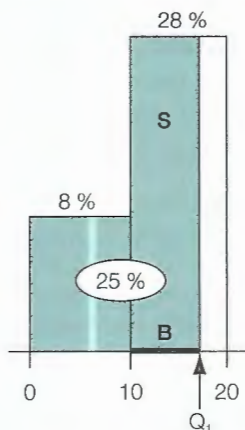
Âge (en ans)	Pourcentage
[0; 10[	8 %
[10; 20[	28 %
[20; 30[	32 %
[30; 40[	20 %
[40; 50[	12 %
Total	100 %

Nous avons déjà estimé que 50 % des arbres avaient moins de 24,4 ans. Supposons qu'un voisin curieux pose la question suivante au propriétaire du boisé : « Le quart de vos arbres ont moins de quel âge ? »

Pour répondre à cette question, il faut trouver un âge, disons  $Q_1$ , tel que 25 % des arbres aient un âge inférieur à  $Q_1$ . La marche à suivre pour déterminer l'âge  $Q_1$  est analogue à celle qui a servi à calculer la médiane.

### Solution

Esquisse de l'histogramme



- Surface :  $25 \% = 8 \% + S$

$$S = 25 \% - 8 \%$$

$$S = 17 \%$$

- $Q_1 : 10 + B$

On a :	<u>Surface</u>		<u>Base</u>
	28 %	→	10 ans
	17 %	→	B ans

$$\text{D'où } B = \frac{17 \times 10}{28} = 6,1 \text{ ans}$$

$$Q_1 = 10 + 6,1 = 16,1 \text{ ans}$$

### Interprétation

On peut estimer que 25 % des arbres ont moins de 16,1 ans.

La valeur  $Q_1$  est le 1<sup>er</sup> quartile de la distribution. Les quartiles partagent une distribution en quatre parties égales comprenant 25 % des données. La médiane correspond au 2<sup>e</sup> quartile :  $Me = Q_2$ .



Les **quantiles** sont des valeurs qui partagent une distribution en un certain nombre de parties égales. Les plus utilisés sont :

- les **quartiles** ( $Q_1, Q_2, Q_3$ ), qui partagent une distribution en quatre parties comprenant 25 % des données;
- les **quintiles** ( $V_1, V_2, V_3, V_4$ ), qui partagent une distribution en cinq parties comprenant 20 % des données;
- les **déciles** ( $D_1, D_2, \dots, D_9$ ), qui partagent une distribution en dix parties comprenant 10 % des données;
- les **centiles** ( $C_1, C_2, \dots, C_{99}$ ), qui partagent une distribution en cent parties comprenant 1 % des données.

## EXERCICES DE COMPRÉHENSION | 1.9

1. Quel âge ont les propriétaires de PME (petites et moyennes entreprises)? Le tableau suivant donne des statistiques à ce sujet.

**Répartition des propriétaires de PME selon l'âge, Canada, 2011**

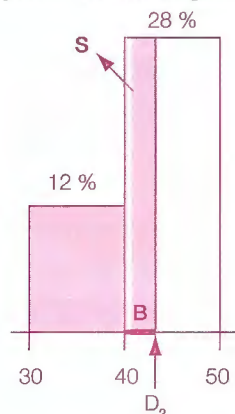
Âge	Pourcentage
Moins de 40 ans	12 %
[40 ans; 50 ans[	28 %
[50 ans; 60 ans[	40 %
60 ans et plus	20 %
<b>Total</b>	<b>100 %</b>

Source: Industrie Canada. *Principales statistiques relatives aux petites entreprises* – Août 2013.

- a) Calculer le 2<sup>e</sup> décile de la distribution. (Fermer la 1<sup>re</sup> classe en lui attribuant une amplitude de 10 ans.)

**Solution**

**Esquisse de l'histogramme**



- Surface:  $20 \% = 12 \% + S$

$$S = 20 \% - 12 \%$$

$$S = 8 \%$$

- $D_2 = 40 \text{ ans} + B \text{ ans}$

On a: Surface

Base

$$28 \% \rightarrow 10 \text{ ans}$$

$$8 \% \rightarrow B \text{ ans}$$

$$\text{d'où } B = \frac{10 \times 8}{28} = 2,9 \text{ ans}$$

$$D_2 = 40 \text{ ans} + 2,9 \text{ ans} = 42,9 \text{ ans}$$

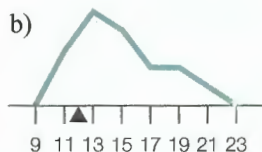
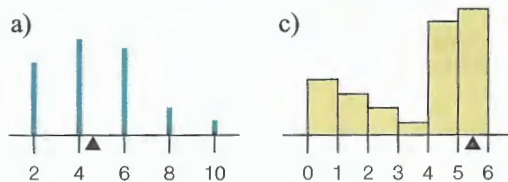
- b) Vrai ou faux. On interprète le décile trouvé en a) ainsi :

On peut estimer que 20 % des propriétaires de PME canadiennes ont 42,9 ans en 2011. Faux

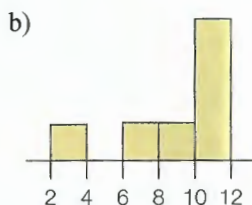
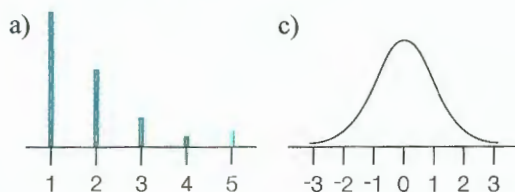
2. 50 ans correspond à quel quintile? Au 2<sup>e</sup> quintile ( $V_2$ ).  
 3. Compléter. 60 ans correspond au 80<sup>e</sup> ( $C_{80}$ ) centile de la distribution.

### EXERCICES 1.3

1. Le pivot placé sous l'axe horizontal de chacun des graphiques suivants permet d'estimer la moyenne de la distribution représentée. Y a-t-il des cas où cette estimation est de toute évidence erronée?



2. Placer un pivot permettant d'estimer la moyenne des distributions suivantes.



3. Pour les séries statistiques suivantes, donner et interpréter chacune des mesures de tendance centrale, puis indiquer laquelle serait la plus représentative de la série statistique.

- a) Le nombre de calendriers vendus en une journée par sept personnes :

7      8      6      9      6      36      10

- b) Le nombre de spectateurs à chacune des six représentations d'une pièce de théâtre :

724    802    715    825    650    790

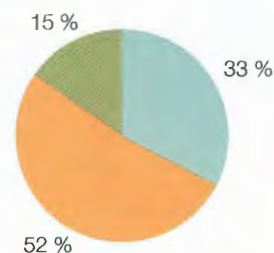
4. Quel est le rythme de production des écrivains professionnels<sup>5</sup>? Se cantonnent-ils dans un seul genre littéraire (roman, poésie, théâtre, etc.)? En 2010, une enquête effectuée auprès des 1 510 écrivains professionnels du Québec donne une réponse à ces questions.

Source: Institut de la statistique du Québec. *Les écrivains québécois. Portrait des conditions de pratique de la profession littéraire au Québec, 2010*, septembre 2011.

- a) Donner et interpréter la meilleure mesure de tendance centrale de la distribution ci-dessous.

#### Répartition des écrivains professionnels selon le rythme de publication, Québec, 2010

- Un livre ou moins tous les trois ans
- Un livre tous les deux ans
- Au moins un livre par année



- b) On compte 680 femmes parmi les 1 510 écrivains professionnels québécois. La moyenne de livres édités en carrière par celles-ci est de 10 livres alors qu'elle est de 11,8 livres pour les hommes. Utiliser ces informations pour calculer le nombre moyen de livres édités en carrière pour l'ensemble des auteurs québécois.

5. Écrivains qui ont publié au moins deux livres au cours de leur carrière.



- c) Déterminer et interpréter la moyenne, le mode et la médiane de la distribution suivante. Remplacer la catégorie «4 et plus» par le nombre «4» dans les calculs.

**Répartition des écrivains professionnels selon le nombre de genres littéraires auxquels appartiennent leurs publications, Québec, 2010**

Nombre de genres littéraires	Nombre d'écrivains	Pourcentage
1	512	33,9 %
2	450	29,8 %
3	285	18,9 %
4 et plus	263	17,4 %
<b>Total</b>	<b>1 510</b>	<b>100,0 %</b>

5. Peut-on vivre de sa plume au Québec? Voici des statistiques sur le revenu des écrivains.

**Répartition des écrivains professionnels selon le revenu tiré des activités de création littéraire, Québec, 2010**

Revenu (en \$)	Pourcentage
Moins de 5 000	64,9 %
[5 000; 20 000[	22,1 %
[20 000; 40 000[	8,8 %
[40 000; 60 000[	2,2 %
60 000 et plus	2,0 %
<b>Total</b>	<b>100,0 %</b>

Source: Institut de la statistique du Québec. *Les écrivains québécois. Portrait des conditions de pratique de la profession littéraire au Québec, 2010*, septembre 2011.

- a) Parmi les 1 510 écrivains professionnels, combien ont tiré un revenu d'au moins 40 000 \$ de leurs activités de création littéraire?
- b) Donner et interpréter la classe modale de la distribution.
- c) Donner et interpréter la médiane de la distribution.
- d) Donner et interpréter la moyenne de la distribution, puis expliquer l'écart que l'on observe entre la moyenne et la médiane. Attribuez une amplitude de 20 000 \$ à la dernière classe.
- e) Quelle mesure de tendance centrale représente le mieux le revenu que les auteurs professionnels tirent de leur création?
6. En 2011-2012, 26 316 cégépiens de la formation technique ont reçu une aide financière pour leurs études. À combien s'élève cette aide? Voici de l'information sur les montants attribués.

**Répartition des bénéficiaires du Programme de prêts et bourses de la formation technique selon le montant<sup>1</sup> d'aide attribué, 2011-2012**

Montant attribué	Pourcentage
Moins de 2 000 \$	26,3 %
[2 000 \$; 4 000 \$[	20,1 %
[4 000 \$; 6 000 \$[	17,7 %
[6 000 \$; 8 000 \$[	16,3 %
[8 000 \$; 10 000 \$[	9,7 %
10 000 \$ et plus	9,9 %
<b>Total</b>	<b>100,0 %</b>

1. Prêt + bourse.

Source: Ministère de l'Enseignement supérieur. *Statistiques. Rapport 2011-2012*, 2014.

- a) Quel pourcentage de bénéficiaires de la formation technique ont reçu 6 000 \$ ou plus du Programme de prêts et bourses?
- b) Calculer et interpréter la moyenne de la distribution.
- c) Donner et interpréter la classe modale.
- d) Donner et interpréter la médiane.
- e) Donner et interpréter le 1<sup>er</sup> quartile.
7. On mise sur l'immigration pour contrer le vieillissement de la population québécoise. Les nouveaux immigrants sont-ils plus jeunes que les Québécois? Voici des statistiques à ce sujet.

**Répartition des immigrants accueillis au Québec en 2012 selon le groupe d'âge**

Âge	Pourcentage
Moins de 15 ans	21,4 %
[15 ans; 30 ans[	29,4 %
[30 ans; 45 ans[	39,2 %
45 ans et plus	10,0 %
<b>Total</b>	<b>100,0 %</b>

Source: Institut de la statistique du Québec. Août 2013.

- a) En 2012, l'âge médian des Québécois est de 41,5 ans. Qu'en est-il de l'âge médian des immigrants que le Québec a accueillis cette année-là? Interpréter cette mesure.
- b) Donner et interpréter la classe modale de la distribution.
- c) Calculer et interpréter la moyenne de la distribution.
- d) Déterminer et interpréter le 2<sup>e</sup> quintile ( $V_2$ ) de la distribution.
- e) À quel décile correspond un âge de 45 ans? Interpréter cette mesure.

8. Calculer la note moyenne des cours suivants en tenant compte du nombre de crédits pour chaque cours. Quel nom donne-t-on à ce type de moyenne?

Cours	Nombre de crédits	Note (sur 100 points)
Mathématiques	3	60
Histoire	2	70
Géographie	2	65
Français	3	80

9. Un ménage est constitué d'une personne ou d'un groupe de personnes qui occupent le même logement. De combien de personnes sont composés les ménages québécois en 2011?

**Répartition des ménages selon le nombre de personnes par ménage, Québec, 2011**

Nombre de personnes par ménage	Pourcentage
1	32,2 %
2	34,8 %
3	14,6 %
4	12,4 %
5	4,2 %
6 et plus	1,8 %
<b>Total</b>	<b>100,0 %</b>

Source: Statistique Canada. Recensement 2011.

- a) Donner le type de la variable étudiée.
- b) Analyser la distribution du nombre de personnes par ménage.
- c) Calculer et interpréter la moyenne en remplaçant la catégorie « 6 et plus » par « 6 ».
- d) Calculer et interpréter la médiane.
10. Le poids moyen des personnes du groupe A est de 66,3 kg et celui du groupe B, de 60,2 kg. Supposons que l'on réunisse les deux groupes. Peut-on, dans les conditions suivantes, calculer le poids moyen du nouveau groupe? Si oui, donner la moyenne.
- a) Les deux groupes contiennent le même nombre de personnes.
- b) On ne connaît pas le nombre de personnes de chaque groupe.
- c) Le groupe A comprend 10 personnes et le groupe B, 40 personnes.
11. Si l'on compare les naissances de 2011 à celles de 1991, observe-t-on une différence en ce qui concerne la distribution de l'âge de la mère? Le tableau suivant permet d'effectuer cette analyse.

**Répartition des naissances selon l'âge de la mère, Québec, 1991 et 2011**

Âge de la mère (en ans)	Pourcentage des naissances	
	1991	2011
Moins de 20	4,1 %	2,5 %
[20; 25[	20,4 %	13,7 %
[25; 30[	41,0 %	32,7 %
[30; 35[	26,4 %	33,8 %
[35; 40[	7,2 %	14,4 %
40 et plus	0,9 %	2,9 %
<b>Total</b>	<b>100,0 %</b>	<b>100,0 %</b>

Source: Institut de la statistique du Québec. Naissances et fécondité, 2012.

- a) Compléter l'énoncé.

On peut estimer que la moyenne d'âge des mères qui ont donné naissance à un enfant était de 28,2 ans en 1991 et de \_\_\_\_\_ ans en 2011, une augmentation de \_\_\_\_\_ an(s) sur une période de 20 ans.

- b) Compléter l'énoncé.

L'âge médian des mères qui ont donné naissance à un enfant peut être estimé à 28,1 ans en 1991 et à \_\_\_\_\_ ans en 2011, une augmentation de \_\_\_\_\_ an(s) sur une période de 20 ans.

- c) Calculer et interpréter le 1<sup>er</sup> décile de l'âge des mères qui ont donné naissance à un enfant en 2011.

- d) Le 4<sup>e</sup> quintile de la distribution de l'âge de la mère pour l'année 2011 est 34,6 ans. Interpréter cette mesure.

12. Selon une étude de Statistique Canada, dans les villes canadiennes de moins de 250 000 habitants, les travailleurs mettent 19 minutes en moyenne pour se rendre au travail. Qu'en est-il des régions métropolitaines comme Montréal où la congestion routière est particulièrement fréquente?

**Répartition des travailleurs selon la durée des déplacements pour se rendre au travail, région de Montréal, 2010**

Durée (en min)	[0; 15[	[15; 30[	[30; 45[	45 et plus	Total
Pourcentage	20 %	27 %	26 %	27 %	100 %

Source: Statistique Canada. Se rendre au travail: résultats de l'enquête sociale générale de 2010, mai 2013.

- a) Calculer la moyenne de la distribution ci-dessus, l'interpréter et la comparer avec le temps moyen que prennent les travailleurs des villes de moins de 250 000 habitants pour se rendre au travail.