

## Statistiques descriptives

### Série statistique à une variable (qualitative et quantitative) et à deux variables (nuage de points)

## EXERCICES

### Table des matières

EXERCICE n°1: Série statistique à une variable : série qualitative .....	2
Partie A. Exemple : Opérateur de téléphonie mobile.....	2
Partie B. Autre exemple : Marque du véhicule.....	3
EXERCICE n°2: Série statistique à une variable : série quantitative discrète.....	4
Partie A. Exemple : Pointure.....	4
Partie B. Comparaison des séries statistiques discrètes.....	5
EXERCICE n°3: Série statistique à une variable : série quantitative continue.....	6
Partie A. Exemple : Taille (en cm).....	6
Partie B. Influence de la largeur des classes.....	7
EXERCICE n°4: Série statistique à deux variables : ajustement affine.....	8
Partie A. Exemple : Nuage mois de naissance vs jour de Naissance.....	8
Partie B. Exemple : Nuage masse (en kg) vs taille (en cm).....	9
Partie C. Exemple : Nuage pointure (sans unité) vs taille (en cm).....	10

*On utilisera les données fournies dans la feuille d'enquête jointe à ce support.*

**EXERCICE n°1: Série statistique à une variable : série qualitative**

*Partie A. Exemple : Opérateur de téléphonie mobile*

**1. Tableau de données**

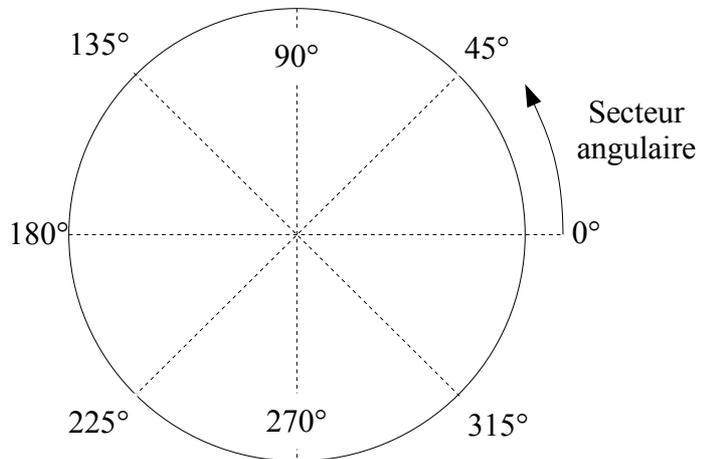
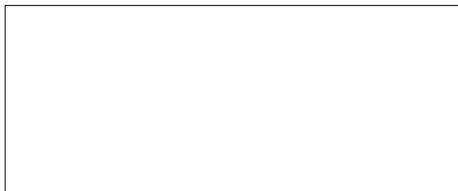
Consigne(s) : Compléter le tableau de données ci-dessous en classant les principaux **opérateurs** par **effectif décroissant**.  $N$  est l'effectif total de la série. Les **fréquences** seront données sous forme décimale arrondies à  $10^{-2}$  près. Compléter également la ligne angle  $\alpha$  (en degré) et somme cumulée des angles pour la création du diagramme camembert.

Opérateur								Sans	Total
Effectif $n$									
Fréquence $f$									1
Angle $\alpha$ (°)									360
Somme cumulée des angles (°)									

**2. Représentations graphiques**

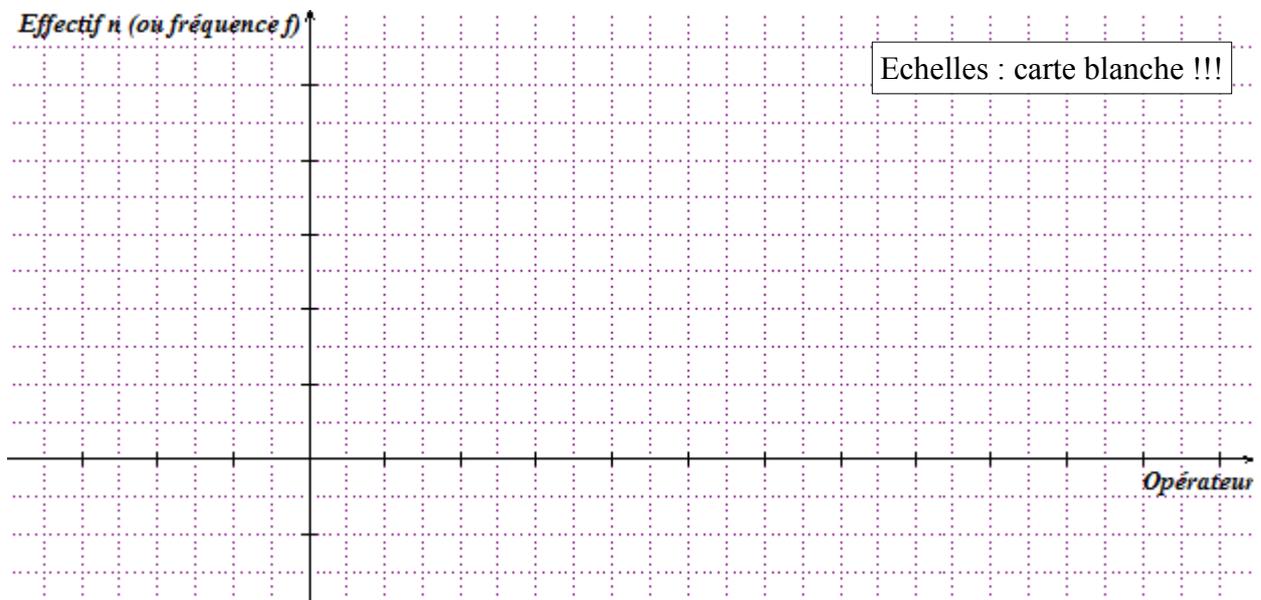
**a) Diagramme camembert**

Consignes : Utiliser la ligne somme cumulée des angles afin de pouvoir construire le diagramme camembert associé.



**b) Diagramme en tuyau d'orgue**

Consigne(s) : Construire de la même manière qu'un diagramme en bâtons (ou qu'un histogramme) le diagramme en tuyau d'orgues. L'axe horizontal contient les différentes qualités du caractère étudié, **l'axe vertical l'effectif associé**.



Partie B. Autre exemple : Marque du véhicule

1. Tableau de données

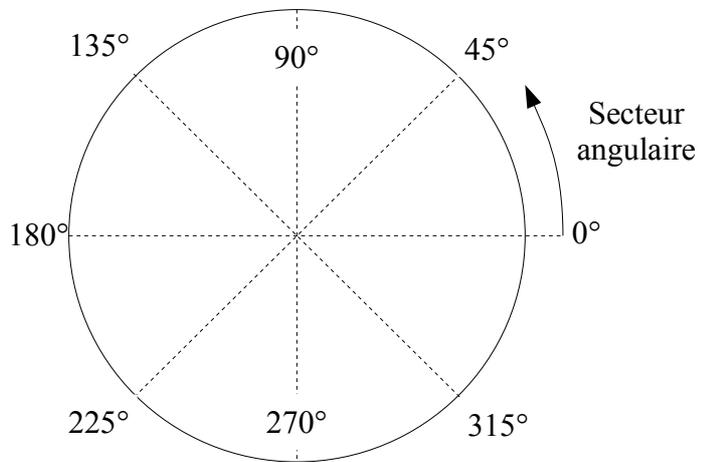
Consigne(s) : Compléter le tableau de données ci-dessous en classant les principales **marques** par **effectif décroissant**.  $N$  est l'effectif total de la série. Les **fréquences** seront données sous forme décimale arrondies à  $10^{-2}$  près. Compléter également la ligne angle  $\alpha$  (en degré) et somme cumulée des angles pour la création du diagramme camembert.

<b>Véhicule</b>								<b>Sans</b>	<b>Total</b>
<b>Effectif <math>n</math></b>									
<b>Fréquence <math>f</math></b>									<b>1</b>
<b>Angle <math>\alpha</math> (°)</b>									<b>360</b>
<b>Somme cumulée des angles (°)</b>									

2. Représentations graphiques

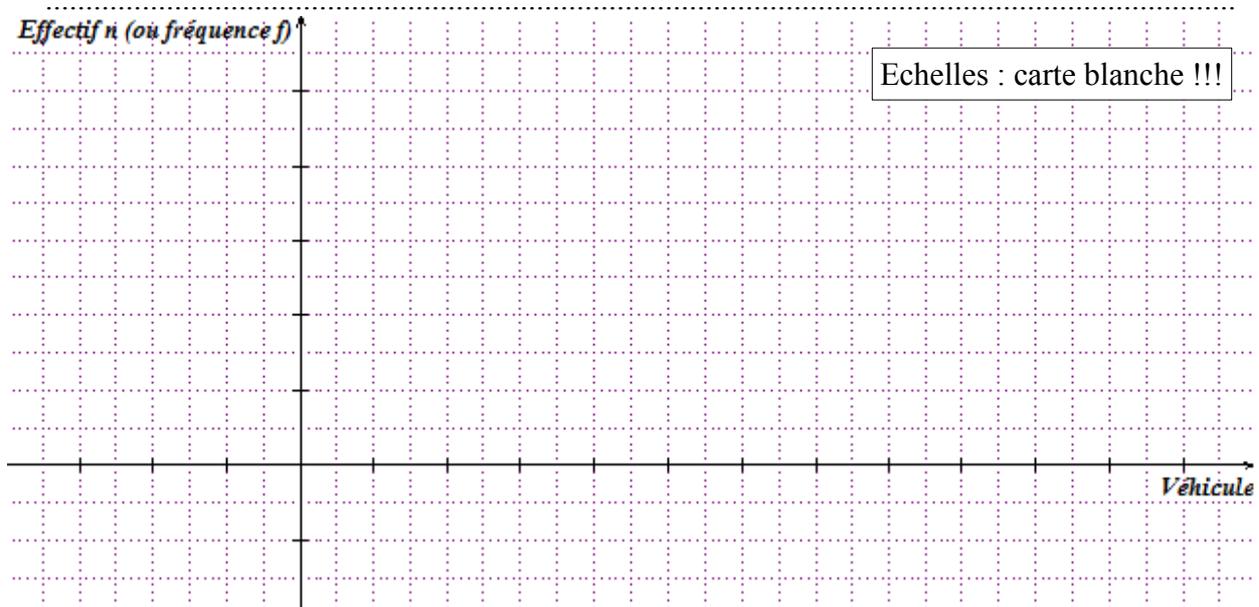
a) Diagramme camembert

Consignes : Utiliser la ligne somme cumulée des angles afin de pouvoir construire le diagramme camembert associé.



b) Diagramme en tuyau d'orgue

Consigne(s) : Construire de la même manière qu'un diagramme en bâtons (ou qu'un histogramme) le diagramme en tuyau d'orgues. L'axe horizontal contient les différentes qualités du caractère étudié, l'axe vertical l'effectif associé. Quel est le pourcentage de véhicule de marque « française » (Renault, Citroën et Peugeot) ?



## EXERCICE n°2: Série statistique à une variable : série quantitative discrète

Les deux parties de cet exercice sont indépendantes.

### Partie A. Exemple : Pointure

#### 1. Tableau de données

Consigne(s) : Compléter le tableau de données ci-dessous. La ligne **ECC** signifie **Effectifs cumulés croissants**. Les fréquences  $f$  seront données sous forme décimale arrondies à  $10^{-2}$  près. La ligne **FCC** signifie Fréquences cumulées croissantes, **exprimées en pourcentage**.

Pointure										Total
Effectif $n$										
ECC										
$f$										1
FCC (%)										

#### 2. Caractéristiques de la série statistique

Consigne(s) : Compléter le tableau ci-dessous donnant les différentes caractéristiques de la série statistique, **obtenues à la calculatrice**.

Série	Effectif $N$	Moyenne	Écart-type	Q1	Médiane	Q3	$x_{min}$	$x_{max}$	Étendue
Pointure									

Remarque(s) : Les **dernières versions des définitions** des quartiles sont les suivantes :

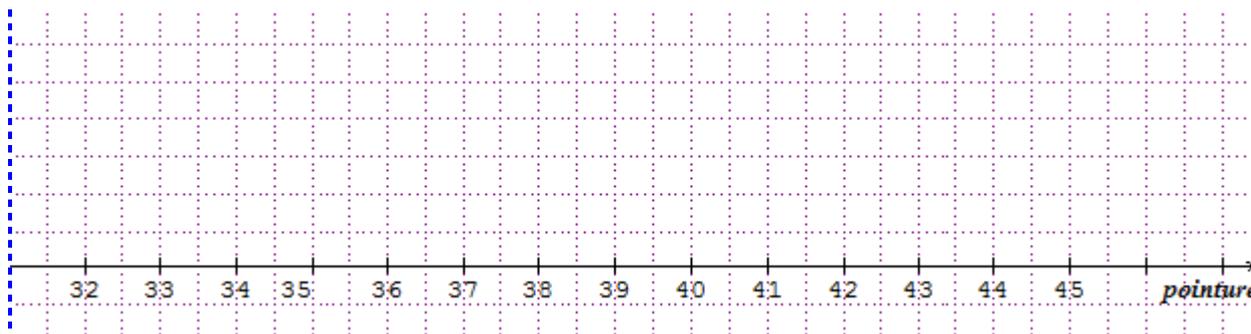
- La valeur de **Q1** est la **première valeur  $x$  du caractère étudié** telle que la valeur FCC associée est supérieure ou égale à 25 % (rappel : 25 % = 0,25 =  $\frac{1}{4}$ ).
- La valeur de la **médiane** (ou **Q2**) est la **première valeur  $x$  du caractère étudié** telle que la valeur FCC associée est supérieure ou égale à 50 % (rappel : 50 % = 0,5 =  $\frac{2}{4}$  =  $\frac{1}{2}$ ).
- La valeur de **Q3** est la **première valeur  $x$  du caractère étudié** telle que la valeur FCC associée est supérieure ou égale à 75 % (rappel : 75 % = 0,75 =  $\frac{3}{4}$ ).

#### 3. Représentations graphiques

a) Diagramme boîte à moustache 

Consigne(s) : Construire le diagramme boîte à moustache en respectant les consignes suivantes :

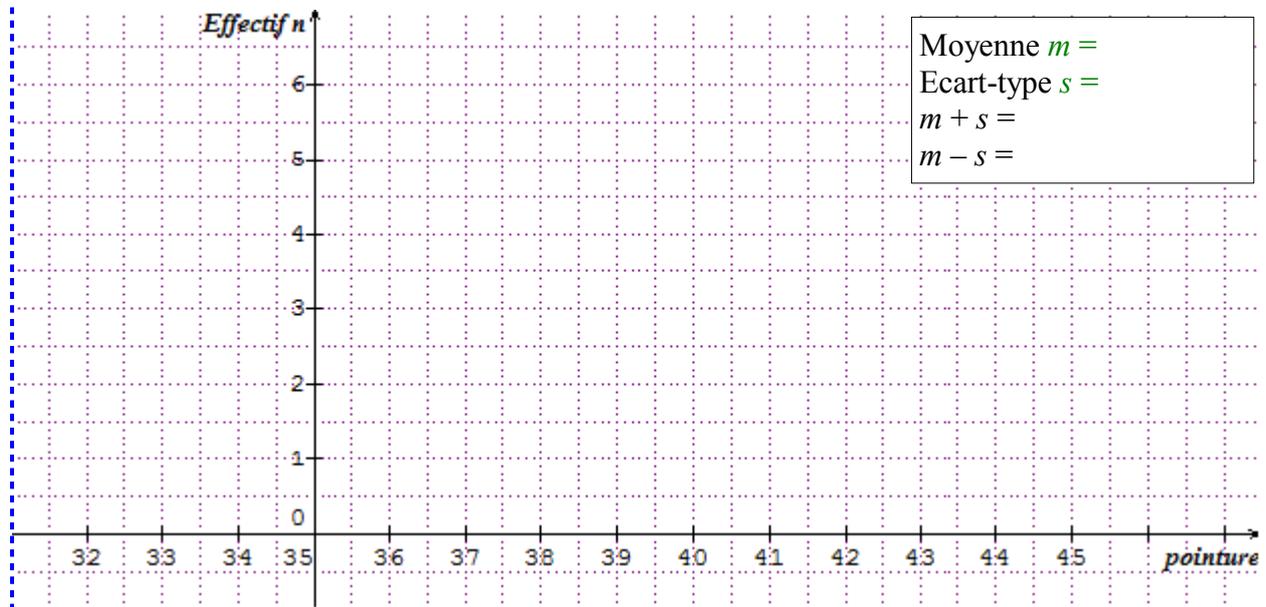
- Sur l'axe horizontal, positionner les valeurs  $x_{min}$ , le premier quartile **Q1**, la Médiane ou **Q2**, le troisième quartile **Q3**, et  $x_{max}$ .
- En démarrant de la valeur **Q1**, tracer un rectangle de largeur **Q3 – Q1** et de hauteur deux cm.
- En partant à mi-hauteur du rectangle, prolonger de part et d'autre le rectangle par deux traits horizontaux jusqu'aux valeurs  $x_{min}$  et  $x_{max}$ .



**b) Diagramme en bâtons**

Consigne(s) : Construire le **diagramme en bâtons** puis représenter **graphiquement** la **moyenne**  $\bar{x}$  et l'**écart-type**  $\sigma$  de la série en respectant les points suivants :

- Repérer sur l'axe horizontal les trois valeurs suivantes  $\bar{x}$ ,  $\bar{x} - \sigma$  et  $\bar{x} + \sigma$ .
- Tracer une **flèche verticale** partant de la valeur de  $\bar{x}$  et de hauteur l'**effectif maximal**,
- A mi-hauteur de cette flèche, tracer de part et d'autre deux flèches horizontales, chacune de longueur  $\sigma$ . (on pourra utiliser les valeurs  $\bar{x} - \sigma$  et  $\bar{x} + \sigma$  repérées précédemment).



*Partie B. Comparaison des séries statistiques discrètes*

Consigne(s) : Compléter le tableau ci-dessous donnant les résultats **obtenus à la calculatrice** des séries statistiques associées à la pointure (sans unité), la taille (en cm) et la masse (en kg), puis répondre aux questions suivantes : Quelle série possède le plus grand écart-type ? la plus grande étendue ? le plus grand intervalle inter-quartile  $Q3 - Q1$  ?

Série	Effectif $N$	Moyenne	Écart-type	$Q1$	Médiane	$Q3$	$x_{min}$	$x_{max}$	Étendue
Pointure									
Taille									
Masse									
Série	Écart-type $\div$ Moyenne			( $Q3 - Q1$ ) $\div$ Médiane			Étendue $\div$ Moyenne		
Pointure									
Taille									
Masse									

.....

.....

.....

.....

.....

.....

**EXERCICE n°3: Série statistique à une variable : série quantitative continue**

Les deux parties de cet exercice **ne sont pas** indépendantes.

Partie A. Exemple : Taille (en cm)

1. Tableau de données

Consigne(s) : Compléter le tableau de données ci-dessous. La ligne **ECC** signifie **Effectifs cumulés croissants (ECD pour Effectifs cumulés décroissants)**. Les fréquences  $f$  seront données sous forme décimale arrondies à  $10^{-2}$  près. La ligne **FCC** signifie **Fréquences cumulées croissantes (FCD pour Fréquences cumulées décroissantes)**, exprimées en pourcentage.

Taille5	[ 155 ; 160 [	[ 160 ; 165 [	[ 165 ; 170 [	[ 170 ; 175 [	[ 175 ; 180 [	[ 180 ; 185 [	[ 185 ; 190 [	[ 190 ; 195 [	[ 195 ; 200 [	[ 200 ; 205 [
Centre $x$	157,5	162,5	167,5	172,5	177,5	182,5	187,5	192,5	197,5	202,5
Effectif $n$										
ECC										
ECD										
$f$										
FCC (%)										
FCD (%)										

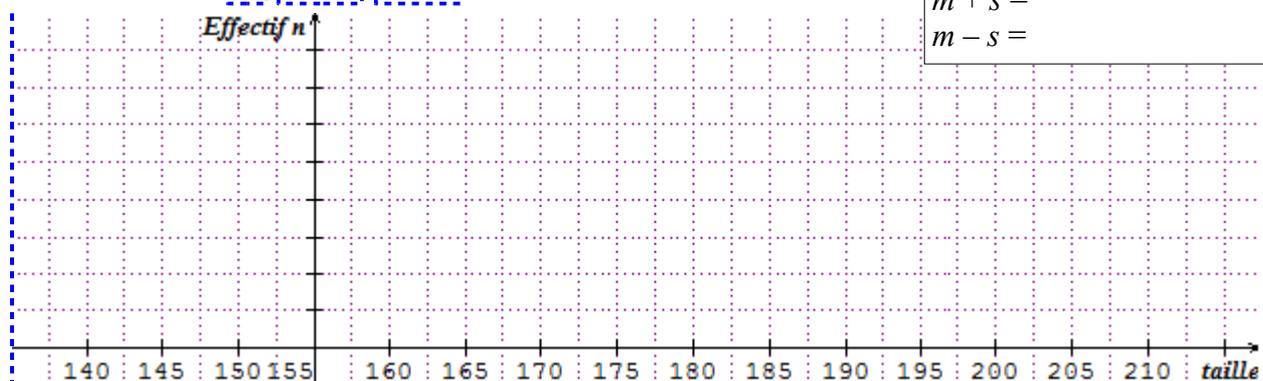
2. Caractéristiques de la série statistique

Consigne(s) : Compléter le tableau ci-dessous donnant les différentes caractéristiques de la série statistique, **obtenues à la calculatrice**, en utilisant les **centres de classes  $x$** . Seule **l'hypothèse** de l'utilisation des **centres de classes** permet d'accéder à la valeur de la **moyenne** et de **l'écart-type**. Pour les **quartiles**, leur détermination doit se faire à partir du tracé du **polygone des fréquences cumulées croissantes (FCC) ou décroissantes (FCD)**.

Série	Effectif $N$	Moyenne	Écart-type	Q1	Médiane	Q3	$x_{min}$	$x_{max}$	Étendue
Taille5									

3. Représentations graphiques

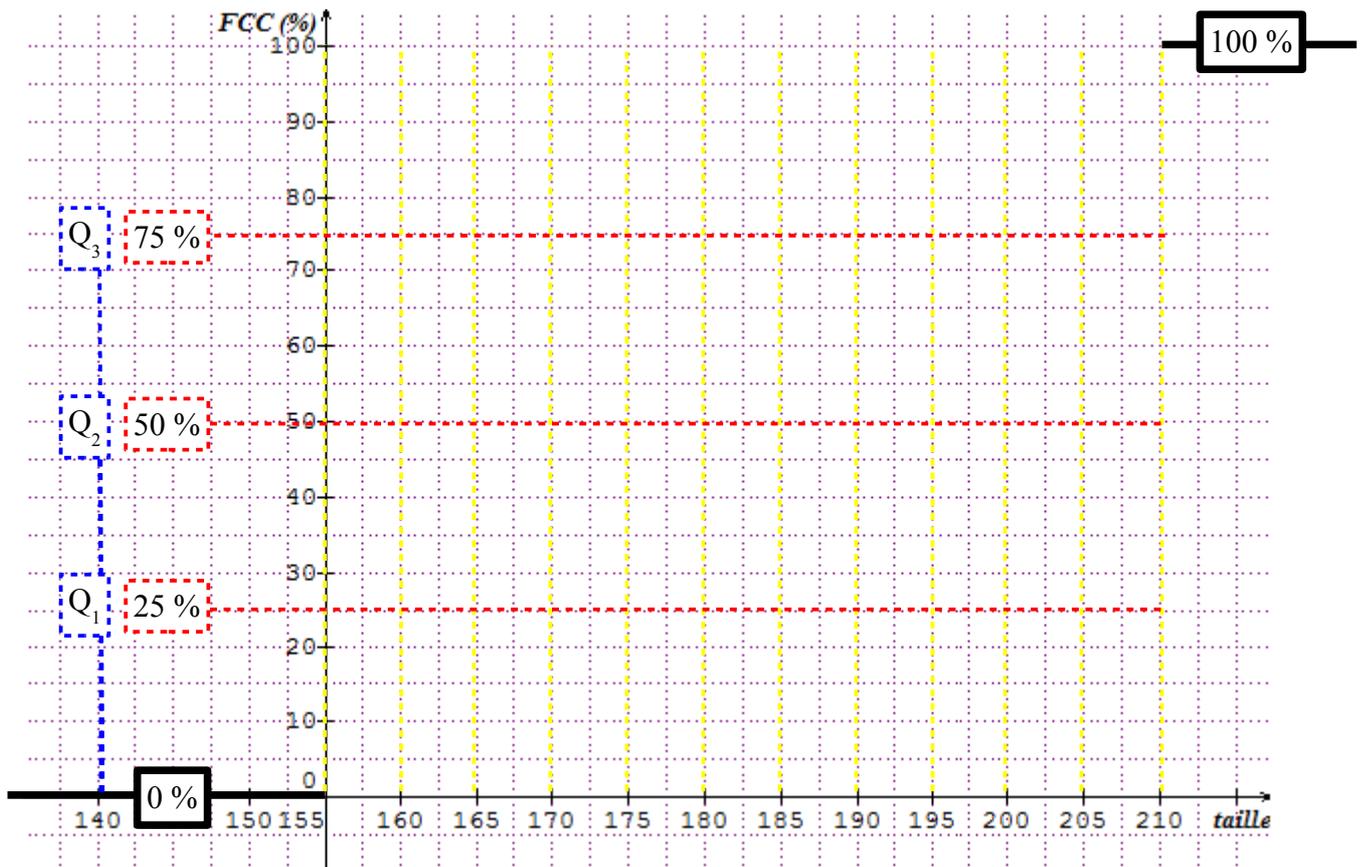
a) Histogramme  $m$   $m - s$   $m + s$



b) Polygone des fréquences cumulées.

Consigne(s) : Construire le polygone des fréquences cumulées croissantes (FCC). Repérer les intersections du polygone avec les valeurs 25 %, 50% et 75%. En déduire graphiquement les valeurs des quartiles **Q1, Q2 (Médiane) et Q3**. Comparer aux valeurs obtenues à la calculatrice.

Résultats calculatrice	<i>Q1</i>	<i>Médiane</i>	<i>Q3</i>	Résultats graphique	<i>Q1</i>	<i>Médiane</i>	<i>Q3</i>



*Partie B. Influence de la largeur des classes*

Consigne(s) : Reprendre le même tableau de données que dans la *Partie A*. mais cette fois ci avec des classes ou des intervalles de largeur 10 cm.

<b>Taille10</b>	[ 155 ; 165 [	[ 165 ; 175 [	[ 175 ; 185 [	[ 185 ; 195 [	[ 195 ; 205 [
<b>Centre <math>x</math></b>	160	170	180	190	200
<b>Effectif <math>n</math></b>					

Compléter le tableau ci-dessous donnant les différentes caractéristiques de la **même** série statistique, **obtenues à la calculatrice**, en **utilisant les centres de classes  $x$** .

<i>Série</i>	<i>Moyenne</i>	<i>Écart-type</i>	<i>Q1</i>	<i>Médiane</i>	<i>Q3</i>	<i><math>x_{min}</math></i>	<i><math>x_{max}</math></i>	<i>Étendue</i>
Taille (discrète)								
Taille5 (continu)								
Taille10 (continu)								

De quelle manière le choix de la largeur de la classe influe sur les caractéristiques de la série ?

.....

.....

.....

**EXERCICE n°4:** Série statistique à **deux** variables : ajustement affine.

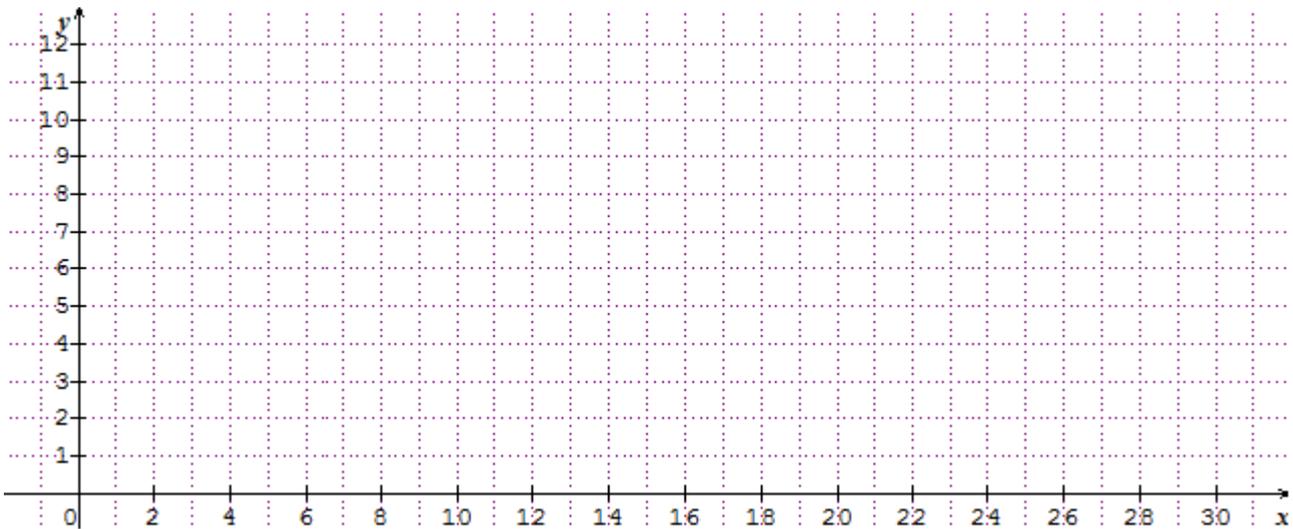
La question 3. de la partie C. dépend des résultats de la question 2. des parties A., B. et C.

Partie A. Exemple : Nuage **mois de naissance** vs **jour de Naissance**

Soit **jj / mm / aaaa** le format d'une date de naissance. Quelles sont les différentes valeurs possibles pour le **jour jj de naissance** ? pour le **mois mm de naissance** ?

1. Représentation graphique :

Consigne(s) : Représenter graphiquement le nuage de points  $(x, y)$  avec **x le jour jj de naissance** et **y le mois mm de naissance** de chaque individu de la série statistique.



a) Y a-t-il des personnes de même **jour de naissance** ? de même **mois de naissance** ? de même date d'anniversaire ?

b) Définir une qualité d'alignement du nuage de points (parfaitement, très, moyennement, peu ou pas du tout aligné). Comment justifier votre choix ?

c) Tracer une droite qui ajuste au mieux l'évolution du nuage de points **puis** déterminer son **coefficient directeur (ou pente) a**, son ordonnée à l'origine **b** et **donc** son équation  **$y = a x + b$** .

2. Caractéristiques de la série statistique

Consigne(s) : Compléter le tableau ci-dessous en **calculant** les coordonnées du **point moyen G** du nuage de point, les coefficients **a** et **b** de l'équation de la droite d'ajustement affine  **$y = a x + b$**  de ce nuage, le coefficient de corrélation **r** puis **r<sup>2</sup>** de ce nuage de points. Arrondir à **10<sup>-3</sup>** près.

Grandeur	$x_G$	$y_G$	$a$	$b$	$r$	$r^2$
Valeur						

Partie B. Exemple : Nuage masse (en kg) vs taille (en cm)

1. Représentation graphique :

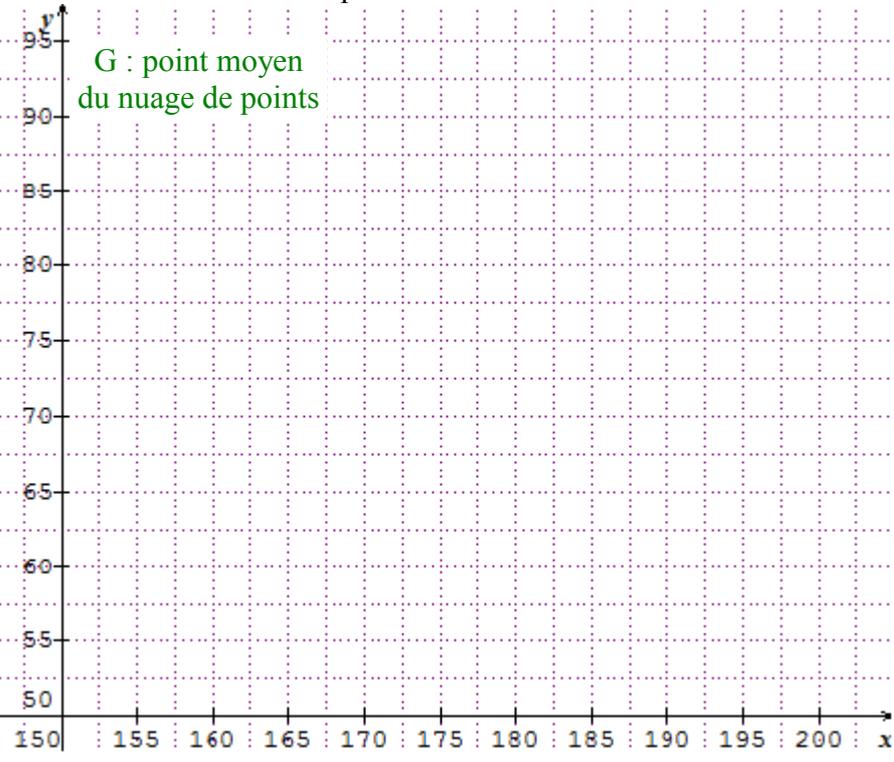
Consigne(s) : Représenter graphiquement le nuage de points  $(x, y)$  avec  $x$  la taille (en cm) et  $y$  la masse (en kg) de chaque individu de la série statistique.

a) Y a-t-il des personnes de même taille ?

Y a-t-il des personnes de même masse ?

Y a-t-il des personnes de même taille et de même masse ?

b) Définir une qualité d'alignement du nuage de points (parfaitement, très, moyennement, peu ou pas du tout aligné). Comment justifier votre choix ?



c) Tracer une droite qui ajuste au mieux l'évolution du nuage de points puis déterminer son coefficient directeur (ou pente)  $a$ , son ordonnée à l'origine  $b$  et donc son équation  $y = ax + b$ .

2. Caractéristiques de la série statistique

Consigne(s) : Compléter le tableau ci-dessous en **calculant** les coordonnées du point moyen G du nuage de point, les coefficients  $a$  et  $b$  de l'équation de la droite d'ajustement affine  $y = ax + b$  de ce nuage, le coefficient de corrélation  $r$  puis  $r^2$  de ce nuage de points. Arrondir à  $10^{-3}$  près.

Grandeur	$x_G$	$y_G$	$a$	$b$	$r$	$r^2$
Valeur						

3. Caractéristiques d'une autre série : Imaginons un nuage de points parfaitement alignés.

Consigne(s) : Refaire la question 2. avec la série suivante :

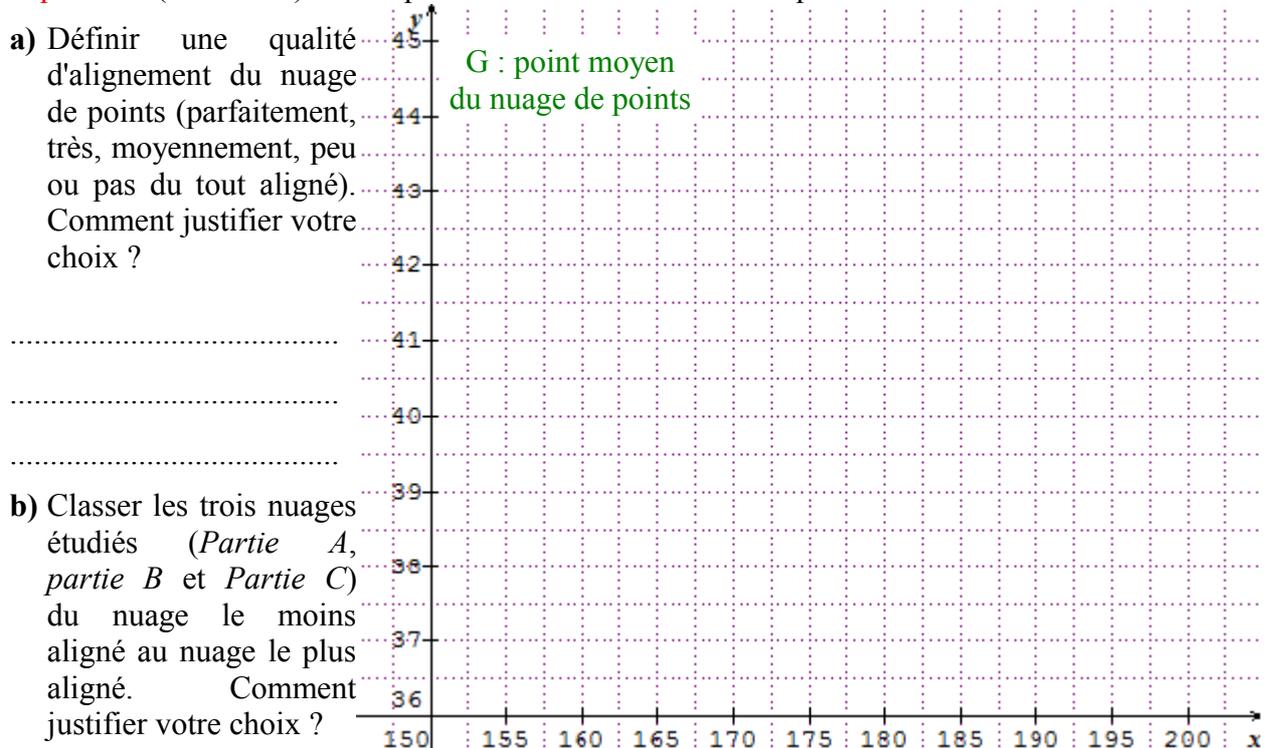
$x$ (cm)	150	160	170	180	190	200
$y$ (kg)	50	60	70	80	90	100

Grandeur	$x_G$	$y_G$	$a$	$b$	$r$	$r^2$
Valeur						

Partie C. Exemple : Nuage pointure (sans unité) vs taille (en cm)

1. Représentation graphique :

Consigne(s) : Représenter graphiquement le nuage de points  $(x, y)$  avec  $x$  la taille (en cm) et  $y$  la pointure (sans unité) de chaque individu de la série statistique.



a) Définir une qualité d'alignement du nuage de points (parfaitement, très, moyennement, peu ou pas du tout aligné). Comment justifier votre choix ?

.....  
 .....  
 .....

b) Classer les trois nuages étudiés (Partie A, partie B et Partie C) du nuage le moins aligné au nuage le plus aligné. Comment justifier votre choix ?

.....  
 .....

c) Tracer une droite qui ajuste au mieux l'évolution du nuage de points puis déterminer son coefficient directeur (ou pente)  $a$ , son ordonnée à l'origine  $b$  et donc son équation  $y = ax + b$ .

.....  
 .....

2. Caractéristiques de la série statistique

Consigne(s) : Compléter le tableau ci-dessous en **calculant** les coordonnées du point moyen G du nuage de point, les coefficients  $a$  et  $b$  de l'équation de la droite d'ajustement affine  $y = ax + b$  de ce nuage, le coefficient de corrélation  $r$  puis  $r^2$  de ce nuage de points. Arrondir à  $10^{-3}$  près.

Grandeur	$x_G$	$y_G$	$a$	$b$	$r$	$r^2$
Valeur						

3. Comparaison des trois nuages de points.

Classer par ordre croissant les valeurs des trois coefficients  $r^2$  des trois nuages de points. En déduire un classement du nuage le moins aligné jusqu'au nuage le plus aligné. Votre résultat est-il conforme au résultat de la question 1.b. De cette page ?

.....  
 .....