

Statistique de poche

Introduction

I. Décrire :

par graphiques et indicateurs.

II. Estimer :

fréquences, moyennes, variances.

III. Tester :

conformité, homogénéité, indépendance.

Introduction.Définition

Afin d'apporter une réponse

à une question concernant un collectif

----- on fait : -----

un état quantitatif = *une statistique*

----- que l'on traite par : -----

des méthodes = *la (science) Statistique*

pour : réaliser l'état

le présenter (*Stat. descriptive*)

l'analyser (*estimations, tests*)

Introduction.Terminologie

Le collectif = Population

Un élément = Individu

Une partie = Echantillon

Variable Qualitative (Nominale/Ordinale)

Quantitative Discrète

Continue

Une valeur de variable = Modalité

Introduction.exemple

AGE	POIDS	TAILLE	SPONTANE	CESARI	POIDS_NAISS
30	78	158	oui	non	2700
22	55	154	oui	non	2940
32	87	179	non	non	2950
37	80	163	oui	non	2360
29	75	162	oui	oui	4380
35	110	164	non	non	3590
22	89	171	oui	non	3470
28	78	173	oui	non	3325
38	72	165	non	non	3470
35	64	160	oui	non	3360
43	50	164	oui	non	3240
35	135	186	non	oui	3290
28	82	171	non	oui	4030
32	87	160	non	non	3010
24	116	163	oui	non	4000
35	79	172	non	oui	2400
36	55	167	non	oui	2770
36	84	165	oui	non	2920
34	70	163	non	non	2790
36	77	165	oui	non	3230

Introduction.Notations

Variable	V
Nombre de modalités	M
Modalité de V indicée	j (1 à M)
Taille d'échantillon	N
Individu indicé i (1 à N)	
Valeur de V pour l'individu i	V_i
<hr/>	
Variable <u>continue</u> discrétisée en classes $[v_{inf_j} ; v_{sup_j}]$	
Une classe = une modalité d'expression	$V_j = (v_{sup_j} + v_{inf_j}) / 2$

Décrire.Univariée_Qualitative.Effectif, fréquence

Nombre d'individus à modalité v_j n_j
= **Effectif** (de la modalité v_j de V)

Fraction d'individus $f_j = n_j / N$
= **Fréquence** (de la modalité v_j de V)

Pourcentage d'individus $\%_j = 100 n_j / N$
= **Fréquence x 100** (de la modalité v_j de V)

Variable qualitative ou quantitative

Décrire.Univariée_Qualitative.Exemple

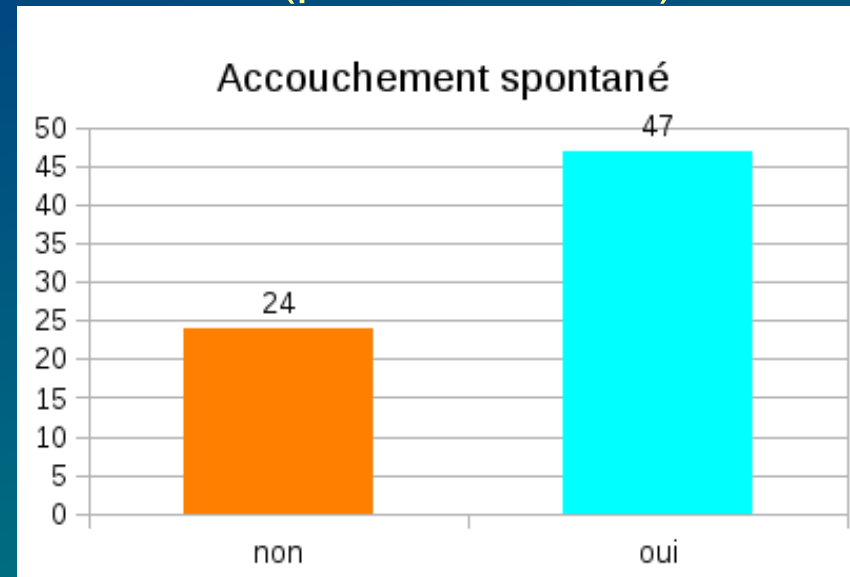
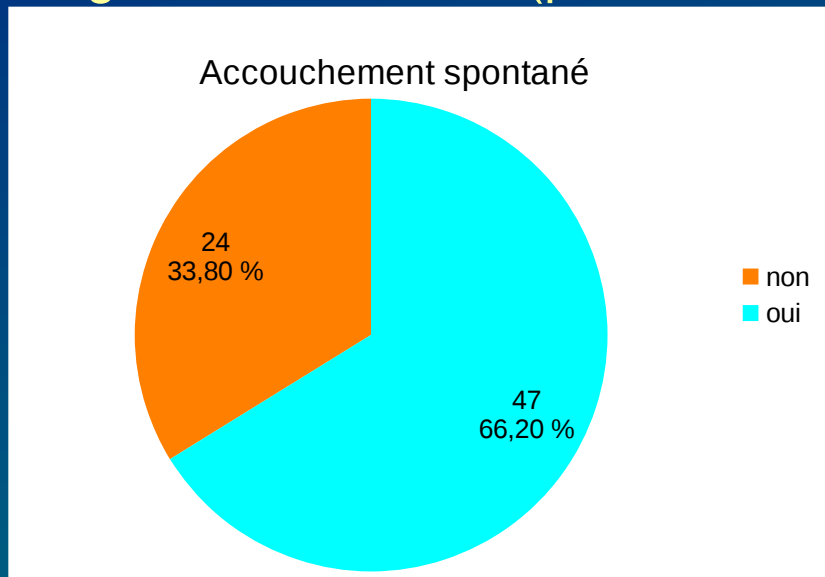
1 variable **qualitative** : accouchement **SPONTANE**

2 modalités : **oui** , **non**

Table d'effectifs, fréquences, pourcentages

	SPONTANE	n _j	f _j	% _j
	non	24	0,338	33,80 %
	oui	47	0,662	66,20 %

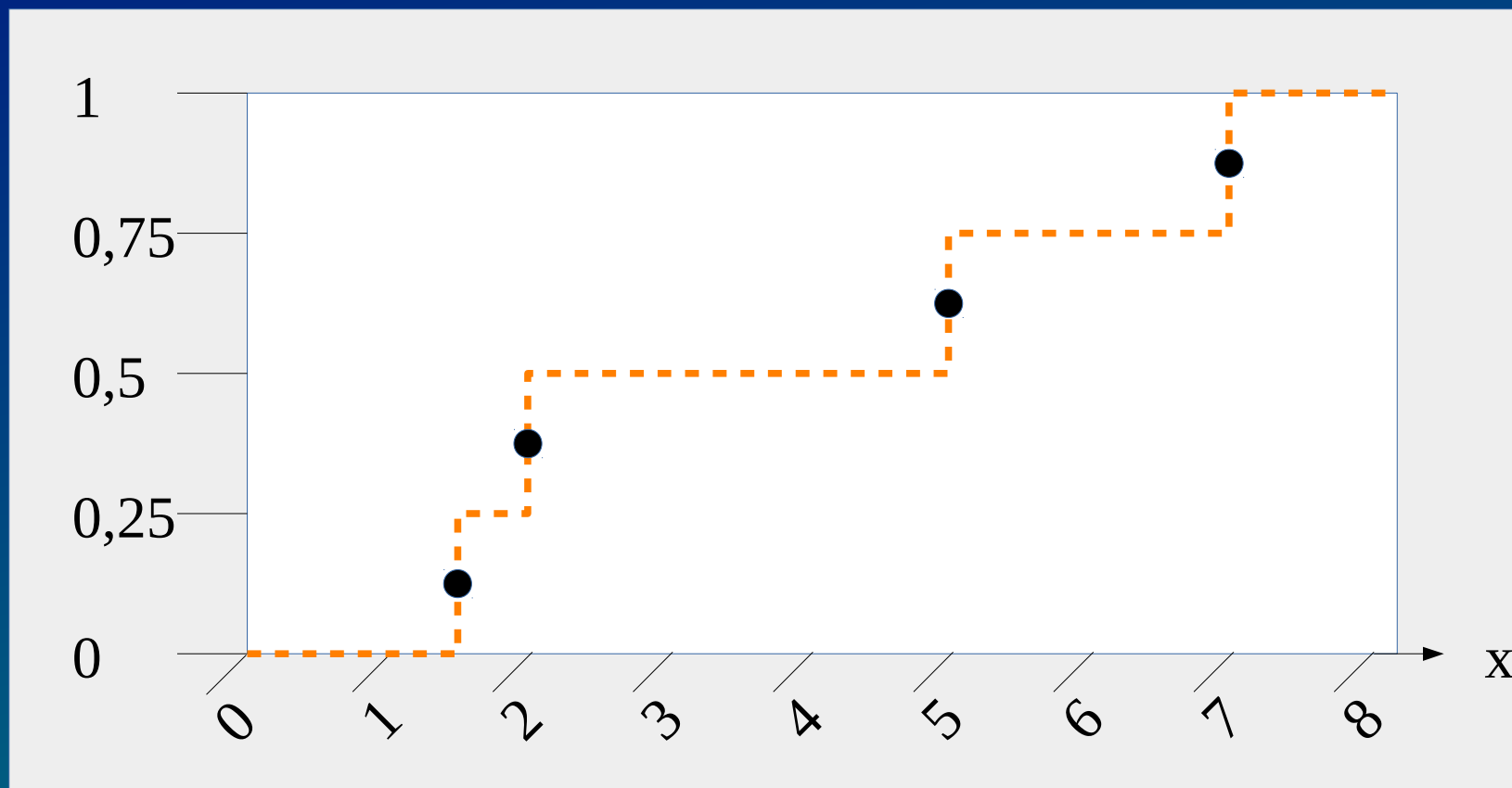
diagramme circulaire (priorité fractions) ou en barres (priorité effectifs)



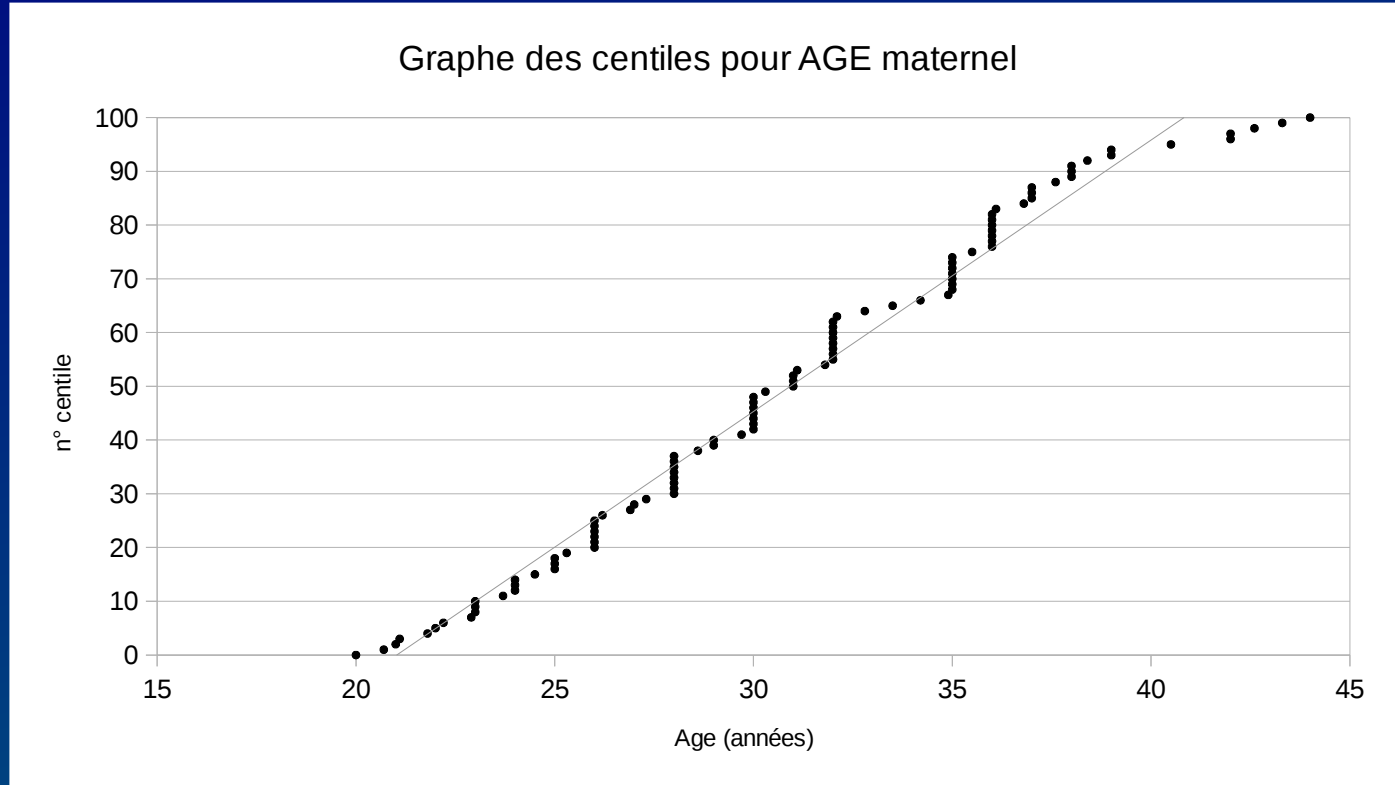
Décrire.Univariée_Quantitative.Graphe des centiles

Exemple de construction sur $N = 4$ points

$$(v_i) = (1,5 ; 2 ; 5 ; 7)$$



Décrire.Univariée_Quantitative.Exemple quasi uniforme



rampe quasi linéaire : distribution quasi uniforme.

Indicateurs appropriés :

min = 20 ans max = 44 max-min = étendue = 24 ans

Décrire.Univariée_Quantitative.Moyenne

Mesure de tendance centrale

$$m_e = \frac{1}{N} \sum_{i=1}^N v_i$$

Changement d'unité

v_i devient $a v_i + b$

m_e devient $a m_e + b$

Décrire.Univariée_Quantitative.Variance, écart-type

Mesure de dispersion

$$s_e^2 = \frac{1}{N} \sum_{i=1}^N (v_i - m_e)^2$$

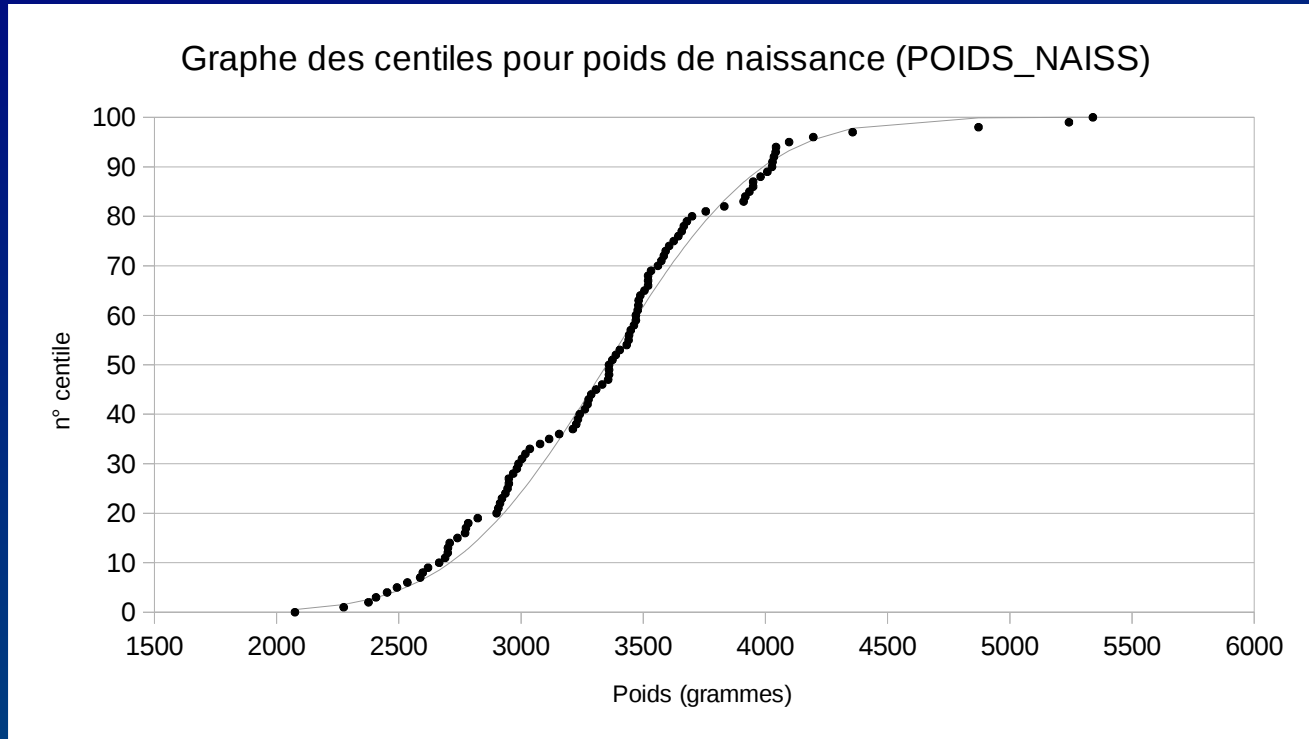
Changement d'unité

v_i	devient	$a v_i + b$
s_e^2	devient	$a^2 s_e^2$

Ecart-type

$$s_e = \sqrt{s_e^2}$$

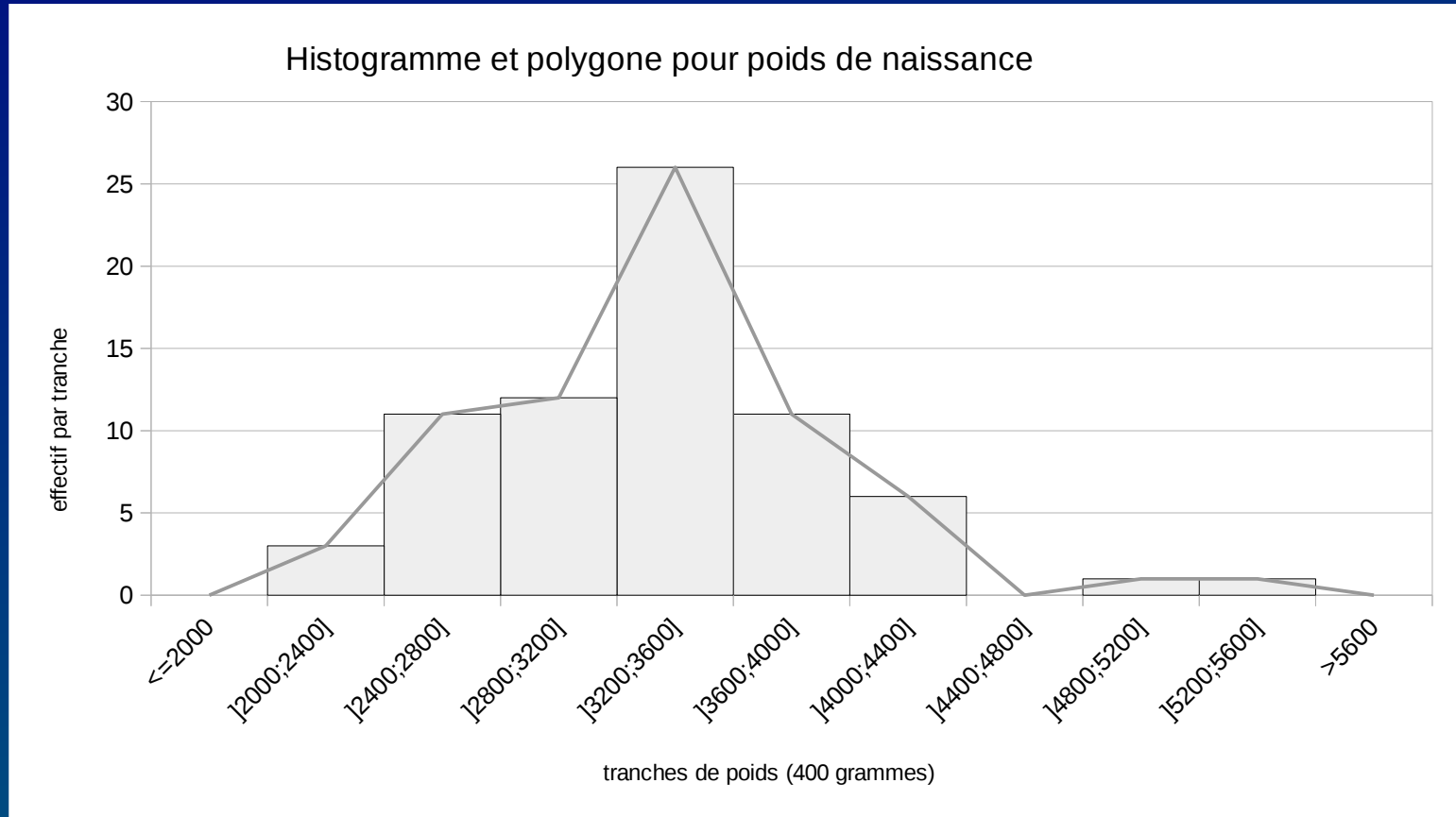
Décrire.Univariée_Quantitative.Exemple gaussien



graphe des centiles en sigmoïde quasi symétrique :
distribution quasi gaussienne.

Indicateurs appropriés : moyenne , écart-type :
 $me = 3347g$ $se = 495g$ (hors deux derniers points)

Décrire.Univariée_Quantitative.Exemple gaussien



le choix des classes est arbitraire.

le polygone améliore la forme en respectant la règle des surfaces

Décrire.Univariée_Quantitative.Médiane, quartiles

Ordonner les v_i

N pair $k=N/2$ $C_{50} = (v_k + v_{k+1}) / 2$

N impair $k=(N-1) / 2$ $C_{50} = v_{k+1}$

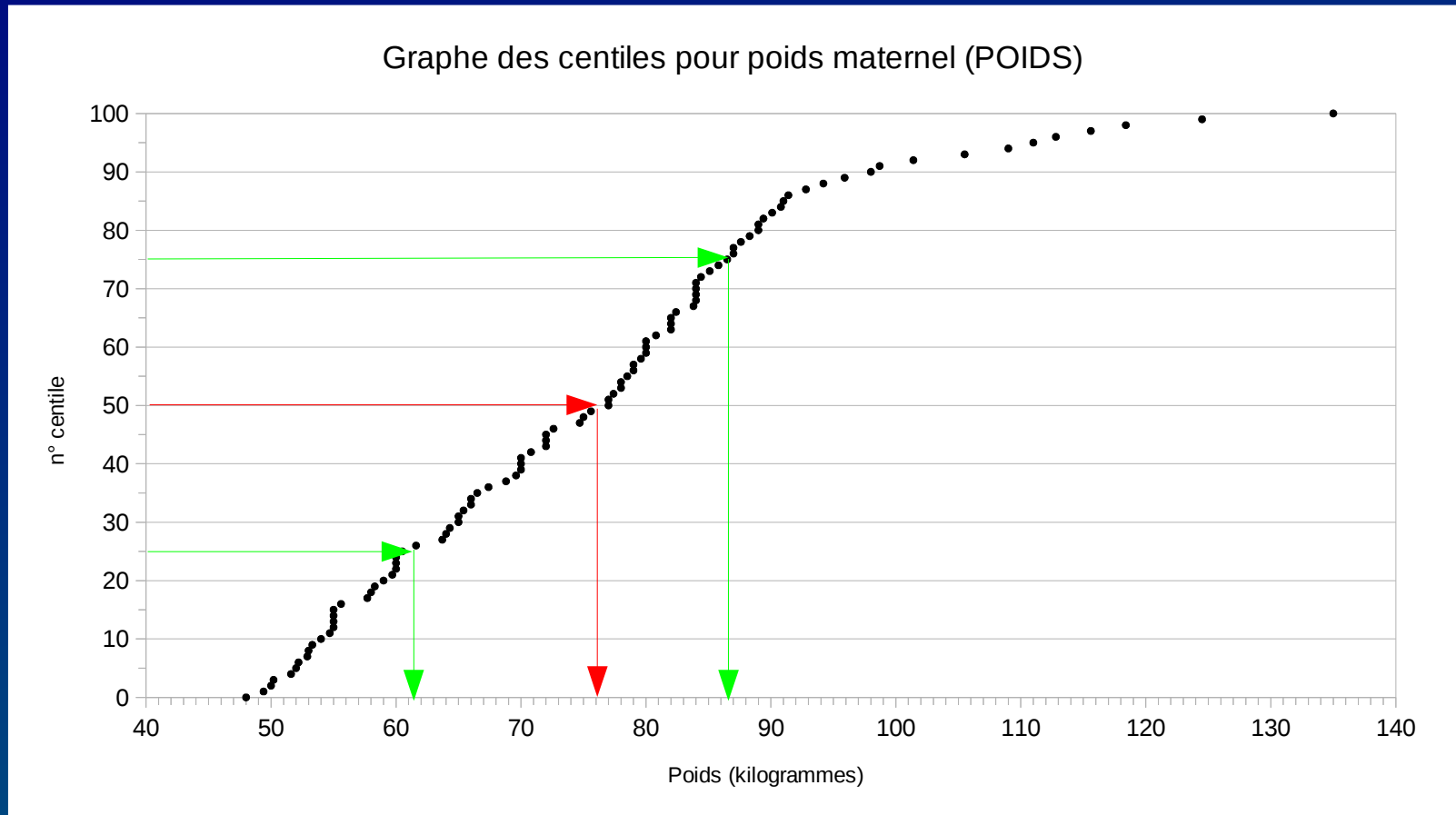
$$C_{50} = \text{« centile 50 »} = \text{médiane}$$

Insensible aux valeurs extrêmes

Quartiles : $Q1 = C_{25}$ $Q2 = C_{50}$ $Q3 = C_{75}$

Interquartile = $Q3 - Q1$

Décrire.Univariée_Quantitative.Exemple asymétrique

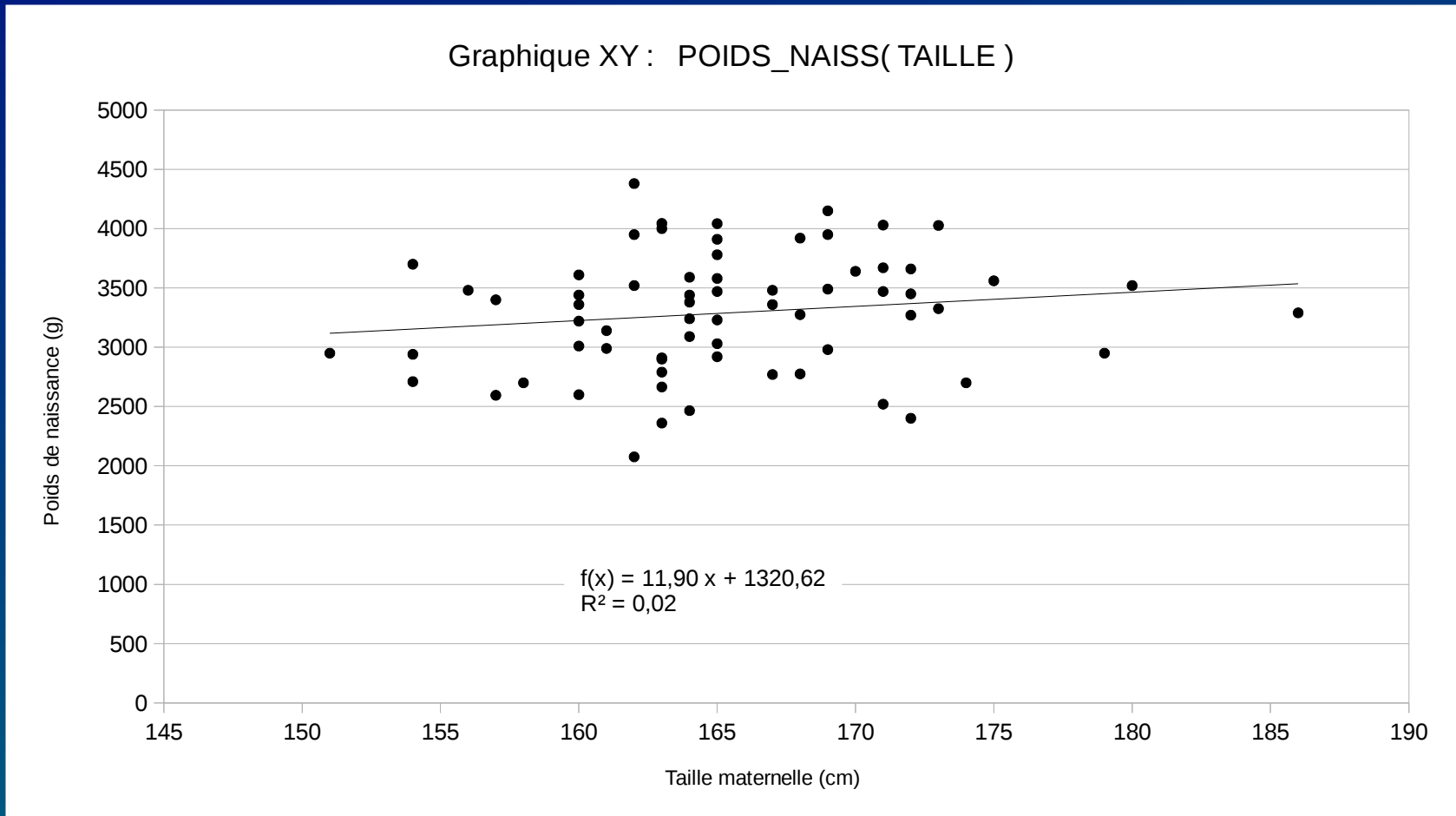


Q1 = 61,5kg médiane = 76kg Q3 = 86,5kg

Interquartile = 25kg

Décrire.Bivariée_2_Quantitatives.Exemple

visualiser une dépendance quantitative versus quantitative :
faire un graphique XY.



Décrire.Bivariée_2_Quantitatives.covariance corrélation

Mesure de liaison entre 2 variables X et Y

$$\text{COV}_{XYe} = \frac{1}{N} \sum_{i=1}^N (x_i - m_{xe})(y_i - m_{ye})$$

Pour s'affranchir des unités :

$$\text{corr}_{XYe} = \frac{\text{COV}_{XYe}}{S_{Xe} S_{Ye}}$$

Décrire.Bivariée_2_Qualitatives.exemple

visualiser une dépendance qualitative versus qualitative :
faire un tableau croisé (table de contingence).

	CESARI		
	non	oui	
SPONTANE			
non	10	14	24
oui	41	6	47
	51	20	71

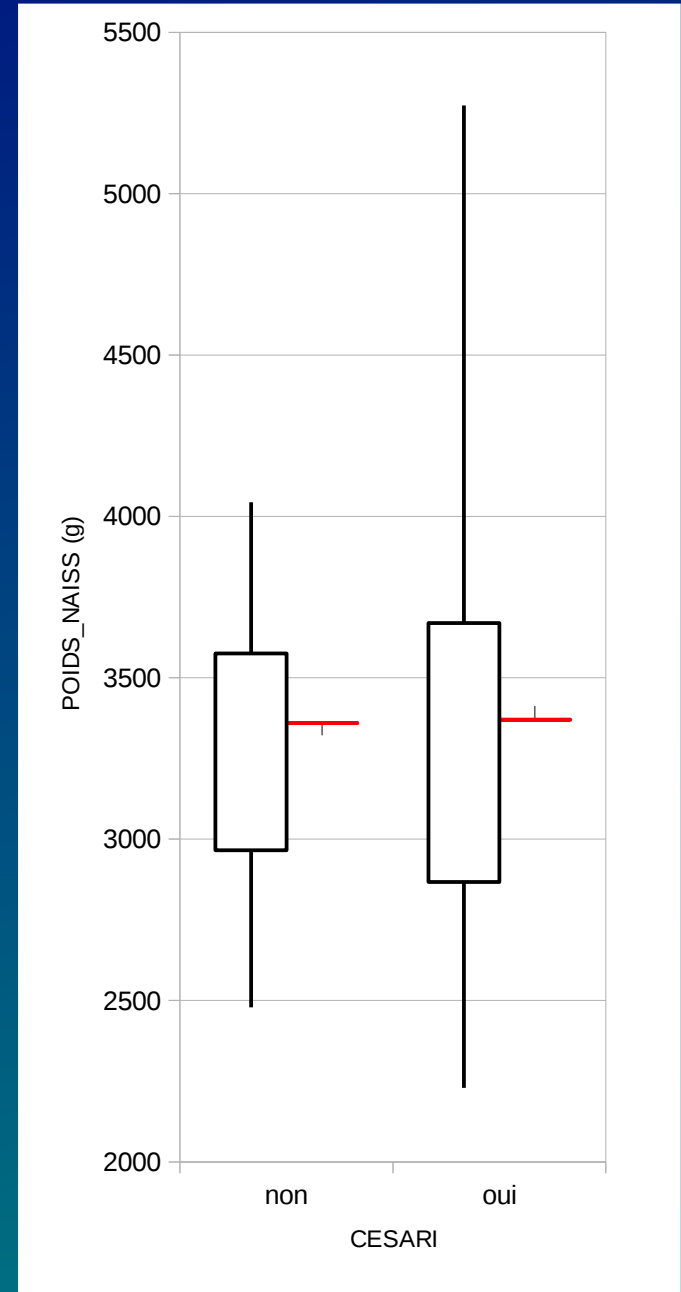
profils de lignes différents en % (dépendance).

	CESARI		
	non	oui	
SPONTANE			
non	41,67 %	58,33 %	100,00 %
oui	87,23 %	12,77 %	100,00 %

Décrire.Bivariée_Qualitative_Quantitative.exemple

visualiser une dépendance
quantitative versus qualitative :

Diagramme boîte et moustaches,
(adapté sur tableur en détournant
le graphique boursier)



Décrire.Bivariée_Qualitative_Quantitative.exemple

mettre en évidence une dépendance
quantitative versus qualitative :

Calculer moyennes et écart-types par groupes pour comparer

	Données		
CESARI	NB	Moyenne - POIDS_NAISS	Ecartype - POIDS_NAISS
non	51	3322	459
oui	20	3413	848
tout confondu	71	3347	589

Utiliser la fonctionnalité "tableau croisé dynamique" sur tableur