

STATISTIQUE DESCRIPTIVE

1. DEFINITION

Branche de la statistique qui regroupe les nombreuses techniques utilisées pour décrire un ensemble relativement important de données.

Analyse numérique

- ✓ **Description** = Résumé et représentation par des statistiques (indices) de données nombreuses
- ✓ Fournir toute l'information disponible sur le phénomène en **moins de chiffres et de mots possibles.**

2. MISE EN ORDRE DES DONNEES

2^{ème} étape après le recueil

Construction **d'un tableau de fréquences** ou
distribution des fréquences

Faire correspondre chaque modalité de la variable à l'effectif correspondant

Distribution des fréquences

Modalité de la variable (X_i)	Effectif (N_i)

3. GROUPEMENT OU CLASSEMENT DES DONNEES

3.1. Variable qualitative

- ❖ Bien définir avant l'étude
- ❖ **Collectivement exhaustives**
- ❖ **Mutuellement exclusives**

Classification internationale des maladies (OMS)

Distribution de 50 malades selon le sexe

Sexe	Effectif	%
Masculin	15	30
Féminin	35	70
Total	50	100

3.2. Variable quantitative

3.2.1. Variable quantitative discrète

Modalités peu nombreuses : Même procédé que les variables qualitatives

*Distribution du nombre annuel d'épisodes de syndrome grippal
chez une population de 77 sujets*

X_i	N_i
0	14
1	16
2	18
3	17
4	12
$\sum N_i$	77

3.2.2. Variable quantitative continue

Regroupement des valeurs de la variable en classes

Classe = Intervalle contenant un certain nombre de valeurs successives

- *Successives*

- *Contigües*

- *Ne se recouvrent pas*

Caractéristiques des classes :

➤ Limites :

- Limite inférieure
- Limite supérieure
- **Limite inférieure toujours incluse**
- **Limite supérieure toujours exclue**
- Limite supérieure = Limite inférieure de la classe suivante
- Limite inférieure = Limite supérieure de la classe précédente
- **Une observation ne doit appartenir qu'à une seule classe à la fois**

➤ **Amplitude de classe :**

Largeur de la classe

Amplitude = Limite supérieure - Limite inférieure

En général : Classes d'amplitude égale

➤ **Centre de classe :**

Demi-somme des 2 limites

$$\text{Centre de classe} = \frac{\text{Limite inférieure} + \text{limite supérieure}}{2}$$

Calcul du nombre de classes :

Nombre de classes C : toujours compris entre **5 et 20**

- $C = \sqrt{N}$

- $C = 1 + 3,3 \log_{10} N$

- **Décomposition de l'étendue en produits de deux facteurs :**

Etendue

= Marge

= Domaine de variation

= Différence entre les valeurs extrêmes

= Valeur supérieure – valeur inférieure

$$\begin{array}{r|l} 50 & 2 \\ 25 & 5 \\ 5 & 5 \\ 1 & \end{array}$$

$$50 = 25 \times 2$$

$$50 = 5 \times (2 \times 5) = 5 \times 10$$

$$50 = 2 \times 25$$

$$50 = (2 \times 5) \times 5 = 10 \times 5$$

Taille de 307 footballeurs algériens

Minimum 159,5 cm

Maximum 191,5 cm

Marge = $191,5 - 159,5 = 32$ cm

32		2
16		2
8		2
4		2
2		2
1		

$$32 = 16 \times 2$$

$$32 = 8 \times (2 \times 2) = 8 \times 4$$

$$32 = 4 \times (2 \times 2 \times 2) = 4 \times 8$$

$$32 = 2 \times (2 \times 2 \times 2 \times 2) = 2 \times 16$$

4. PRESENTATION TABULAIRE

Résumé des données

Toutes les indications utiles pour sa compréhension

- Titre** : Explicite
 - Définition du contenu
 - Objet de l'étude
 - Lieu de l'étude
 - Période d'étude
- Modalités des caractères étudiés**
- Unités de mesure**

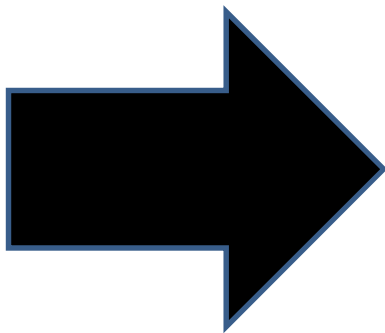
4.1. Etude d'une seule variable

Tableau à **simple entrée**

Tableau à **une dimension**

Tableau **unidimensionnel**

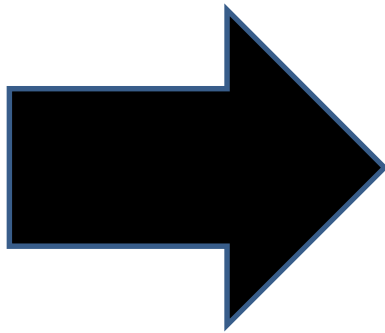
*Distribution du nombre annuel d'épisodes de syndrome grippal
chez une population de 77 sujets*



X_i	N_i
0	14
1	16
2	18
3	17
4	12
$\sum N_i$	77

Taille de 307 footballeurs algériens

Indice de classe	Classe	Contre de classe X_i	Effectif N_i
1	159,5 – 161,5	160,5	7
2	161,5 – 163,5	162,5	4
3	163,5 – 165,5	164,5	10
4	165,5 – 167,5	166,5	23
5	167,5 – 169,5	168,5	19
6	169,5 – 171,5	170,5	39
7	171,5 – 173,5	172,5	55
8	173,5 – 175,5	174,5	48
9	175,5 – 177,5	176,5	35
10	177,5 – 179,5	178,5	31
11	179,5 – 181,5	180,5	16
12	181,5 – 183,5	182,5	9
13	183,5 – 185,5	184,5	5
14	185,5 – 187,5	186,5	3
15	187,5 – 189,5	188,5	1
16	189,5 – 191,5	190,5	2
			$\sum N_i = 307$



4.2. Etude de deux variables

Tableau à **double entrée**

Tableau à **deux dimensions**

Tableau **bidimensionnel**

✓ **Deux variables qualitatives : Tableau de contingence**

Comparaison de répartitions



*Répartition de 826 élèves scolarisés
selon la survenue d'accident et le sexe – Alger, 1985.*



Accident / sexe	Masculin	Féminin	Total
Oui	102	105	207
Non	209	410	619
Total	311	515	826

✓ Deux variables quantitatives : **Tableau de corrélation**

Test de corrélation

Age

*Répartition de 247 élèves scolarisés
selon l'âge et le poids – Alger, 1985.*

Poids (Kg)/Age (Années)	5 - 7	7 - 9	9 - 11	11 - 13	13 - 15
11 - 15	1				
15 - 19	4	1			
19 - 23	8	6		1	
23 - 27	4	16	7		
27 - 31	2	8	13	5	
31 - 35		3	16	14	5
35 - 39		1	10	13	16
39 - 43			2	15	7
43 - 47			1	8	21
47 - 51				5	15
51 - 55				1	9
55 - 59				1	3
59 - 63					4
63 - 67					1

Poids

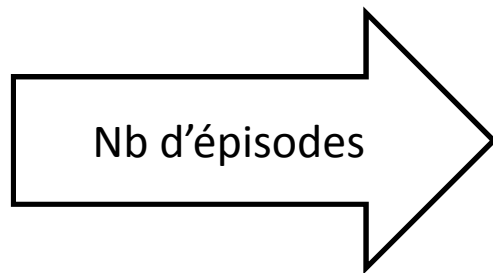
- ✓ Une variables qualitative et un variable quantitative :

Tableau de comparaison

Comparaison de moyennes



Distribution du nombre annuel d'épisodes de syndrome grippal chez une population de 77 sujets selon le sexe



Nb d'épisodes	Hommes	Femmes	Total
0	8	6	14
1	9	7	16
2	10	8	18
3	11	6	17
4	7	5	12
Total	45	32	77
<i>Moyenne</i>	<i>2,2</i>	<i>1,9</i>	

5. PRESENTATION GRAPHIQUE

Moyen suggestif

Vue synoptique

Allure générale

Impression durable

- **Simple**
- **Clair**
- **Explicite**

Moins de précision que le tableau

Tableaux et graphiques sont complémentaires

5.1. Généralités

Graphique = 3 parties

- ❖ **Titre** : Même principes que le titre du tableau
- ❖ **Coordonnées, axes et échelles** :
 - Axe des abscisses : Modalités de la variable (*Axe des x*)
 - Axe des ordonnées : Effectifs – Absolus, relatifs ou cumulés –
(*Axe des y*)
 - Echelle arithmétique ou logarithmique
- ❖ **Tracé** :
 - Surfaces rectangulaires
 - Bâtonnets
 - Polygones (lignes brisées)
 - Courbes
 - Nuages de points

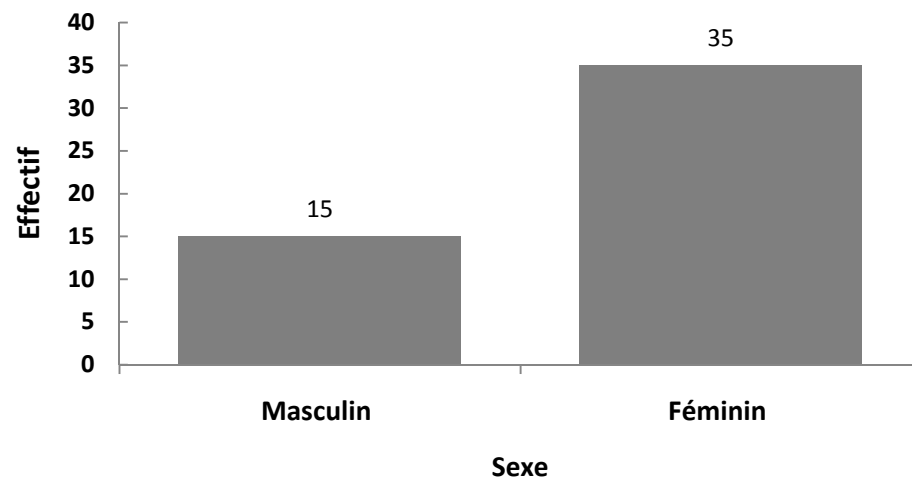
5.2. Variable qualitative

Diagramme à bâtonnets

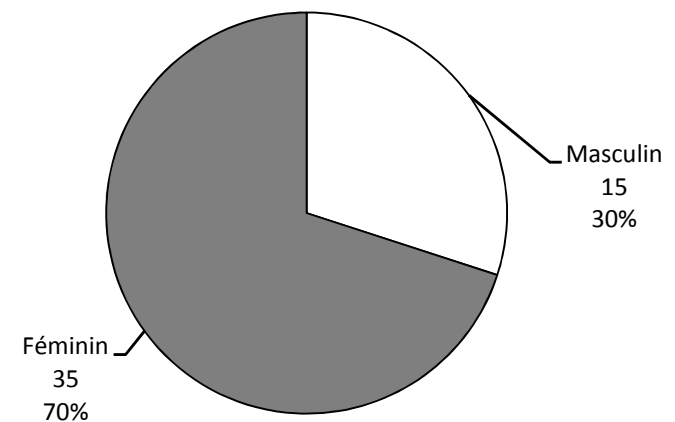
Chaque modalité est schématisée par un bâtonnet

Bâtonnet :

- Largeur a : Constante et arbitraire
- Hauteur h : Proportionnelle à l'effectif
- Surface s (a x h) : Proportionnelle à l'effectif



Distribution de 50 malades selon le sexe

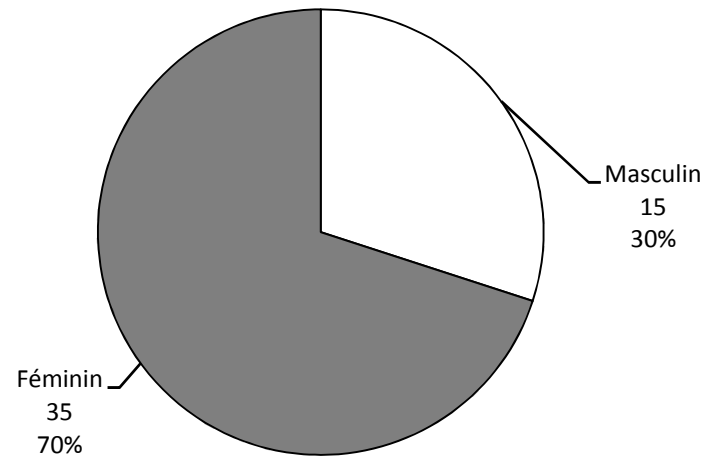


Distribution de 50 malades selon le sexe

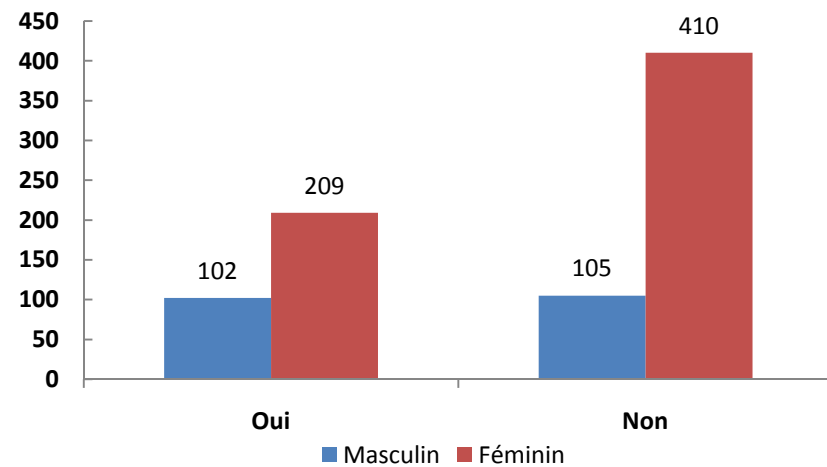
Graphique circulaire

Chaque modalité est schématisée par une portion proportionnelle à l'effectif

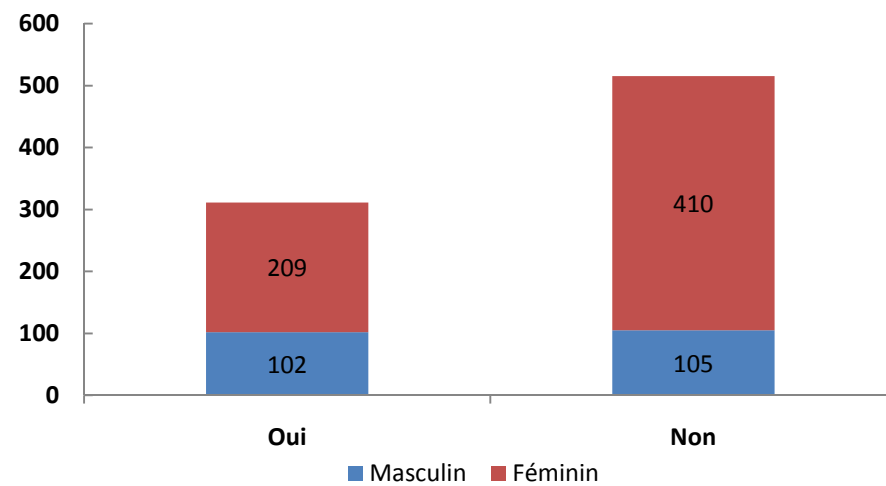
100 % de l'effectif = 360°



Distribution de 50 malades selon le sexe



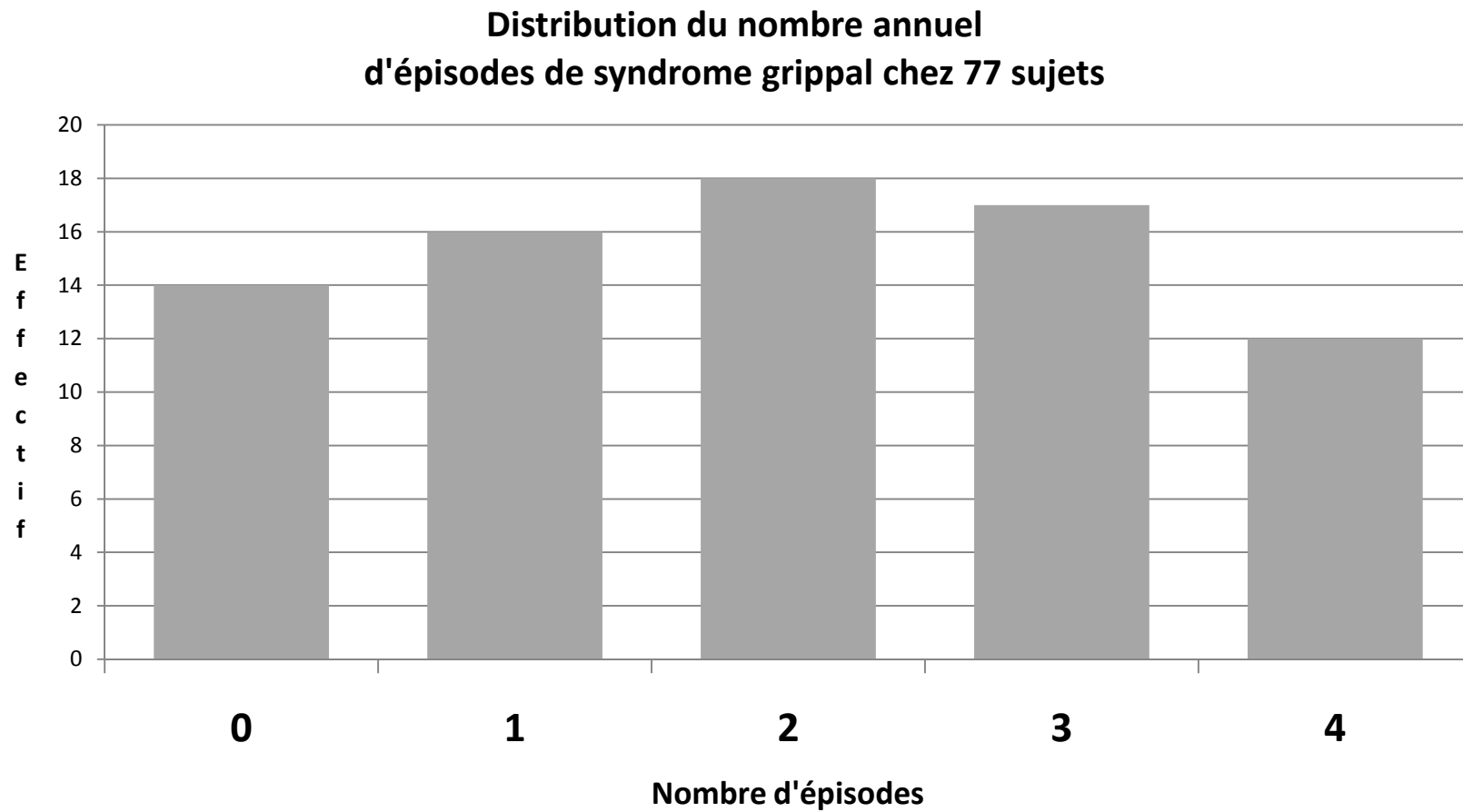
**Répartition de 826 élèves scolarisés
selon la survenue d'accident et le sexe – Alger, 1985.**



**Répartition de 826 élèves scolarisés
selon la survenue d'accident et le sexe – Alger, 1985.**

5.3. Variable quantitative discontinue

Diagramme à bâtonnets



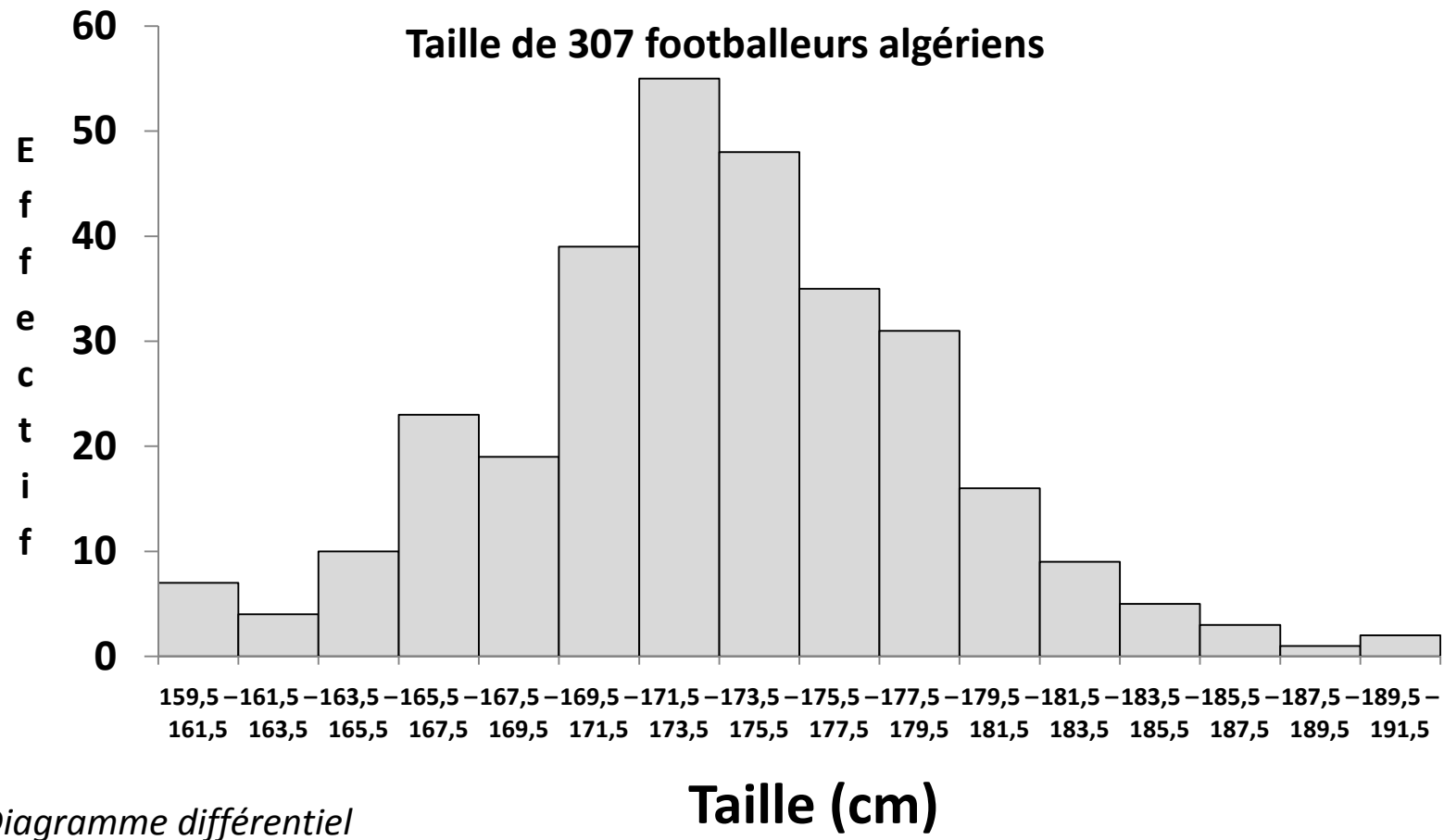
5.4. Variable quantitative continue

Histogramme :

Même principe que le diagramme à bâtonnets

Avec des surfaces juxtaposées

Largeur a proportionnelle à l'amplitude de classe



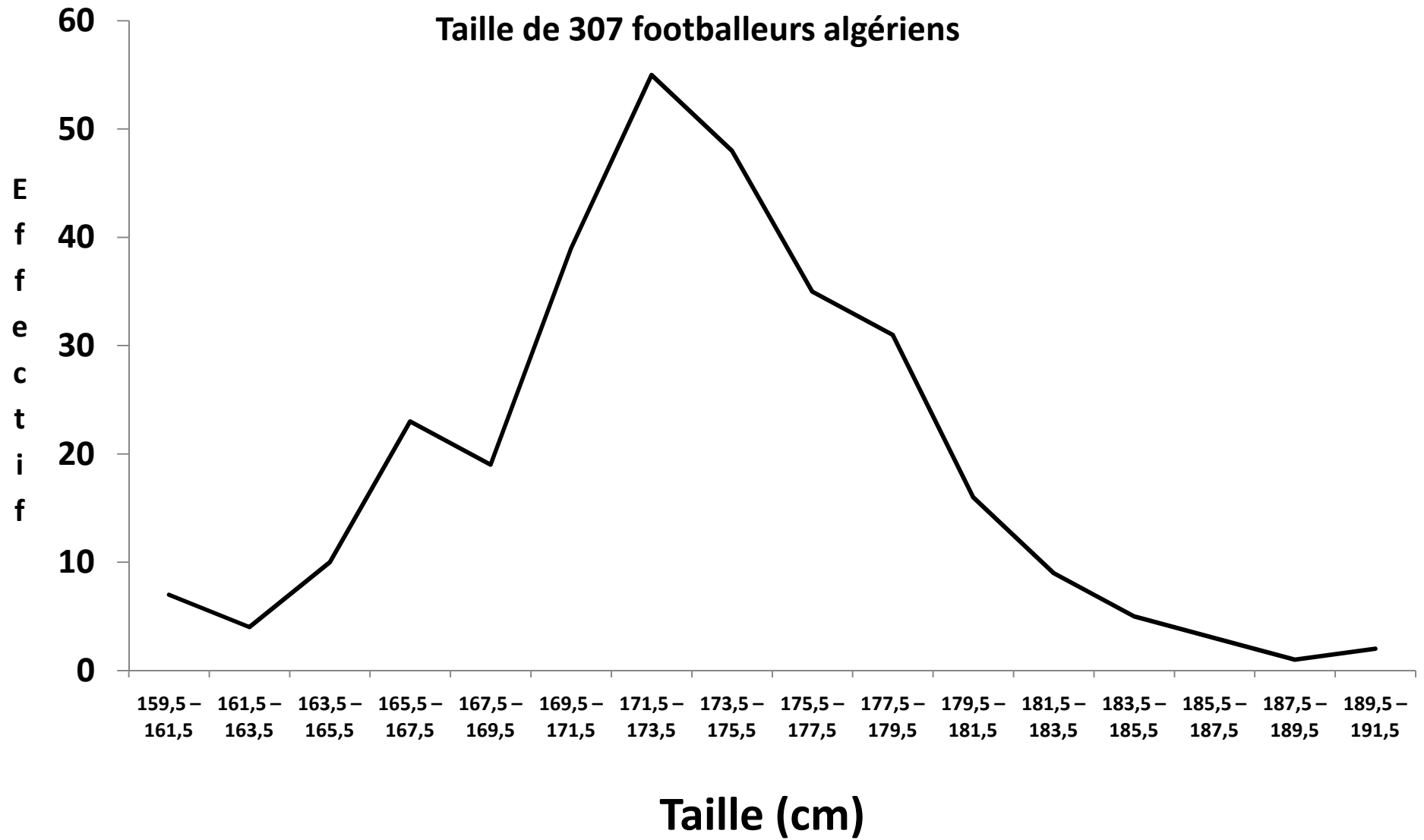


Diagramme différentiel

Taille de 307 footballeurs algériens

Indice de classe	Classe	Contre de classe X_i	Effectif N_i	Effectif Cumulé « moins de » $N_i (-)$	Effectif Cumulé « plus de » $N_i (+)$
1	159,5 – 161,5	160,5	7	7	307
2	161,5 – 163,5	162,5	4	11	300
3	163,5 – 165,5	164,5	10	21	296
4	165,5 – 167,5	166,5	23	44	286
5	167,5 – 169,5	168,5	19	63	263
6	169,5 – 171,5	170,5	39	102	244
7	171,5 – 173,5	172,5	55	157	205
8	173,5 – 175,5	174,5	48	205	150
9	175,5 – 177,5	176,5	35	240	102
10	177,5 – 179,5	178,5	31	271	67
11	179,5 – 181,5	180,5	16	287	36
12	181,5 – 183,5	182,5	9	296	20
13	183,5 – 185,5	184,5	5	301	11
14	185,5 – 187,5	186,5	3	304	6
15	187,5 – 189,5	188,5	1	305	3
16	189,5 – 191,5	190,5	2	307	2

$\sum N_i = 307$

Taille de 307 footballeurs algériens

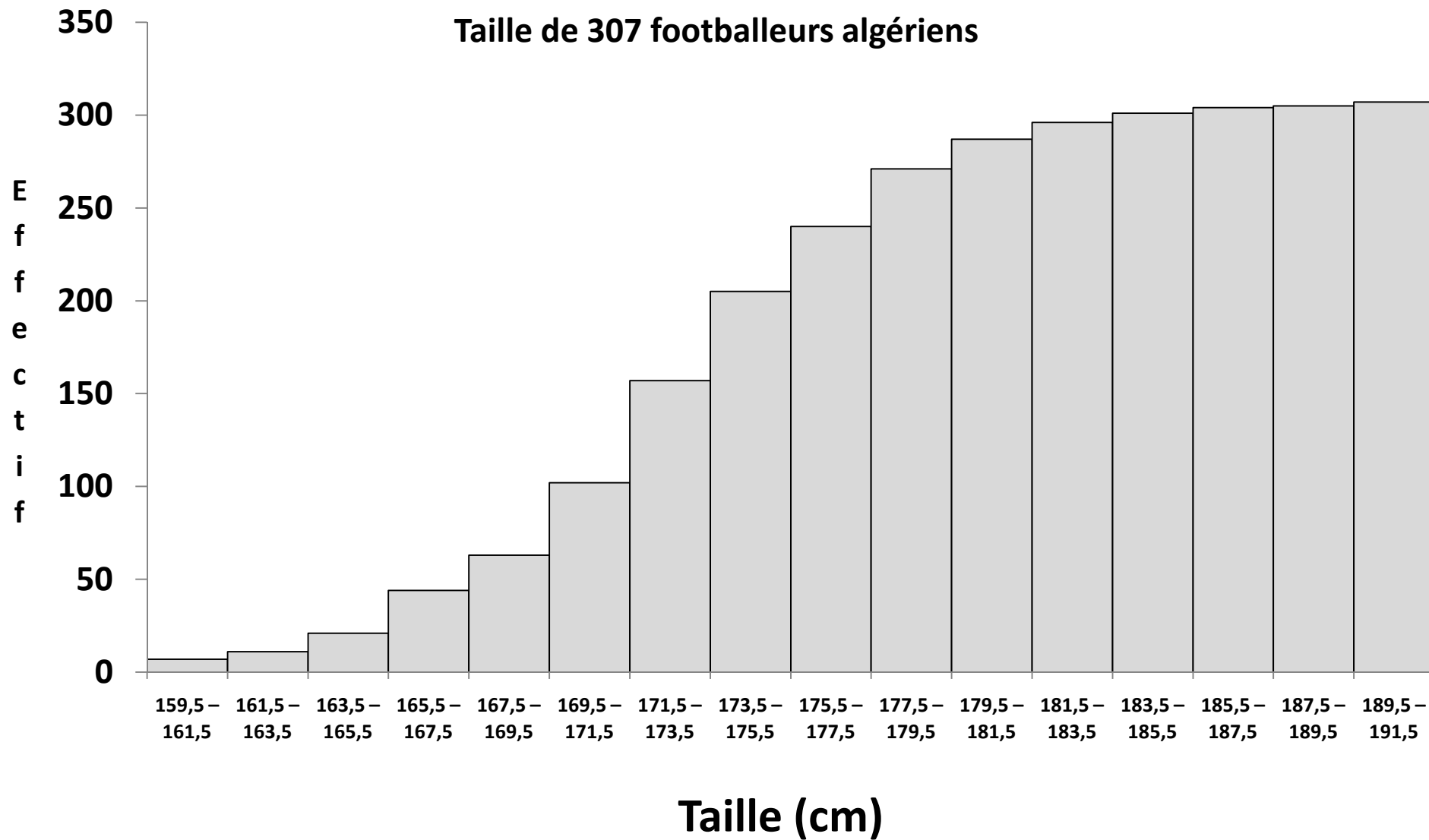


Diagramme Intégral

Taille de 307 footballeurs algériens

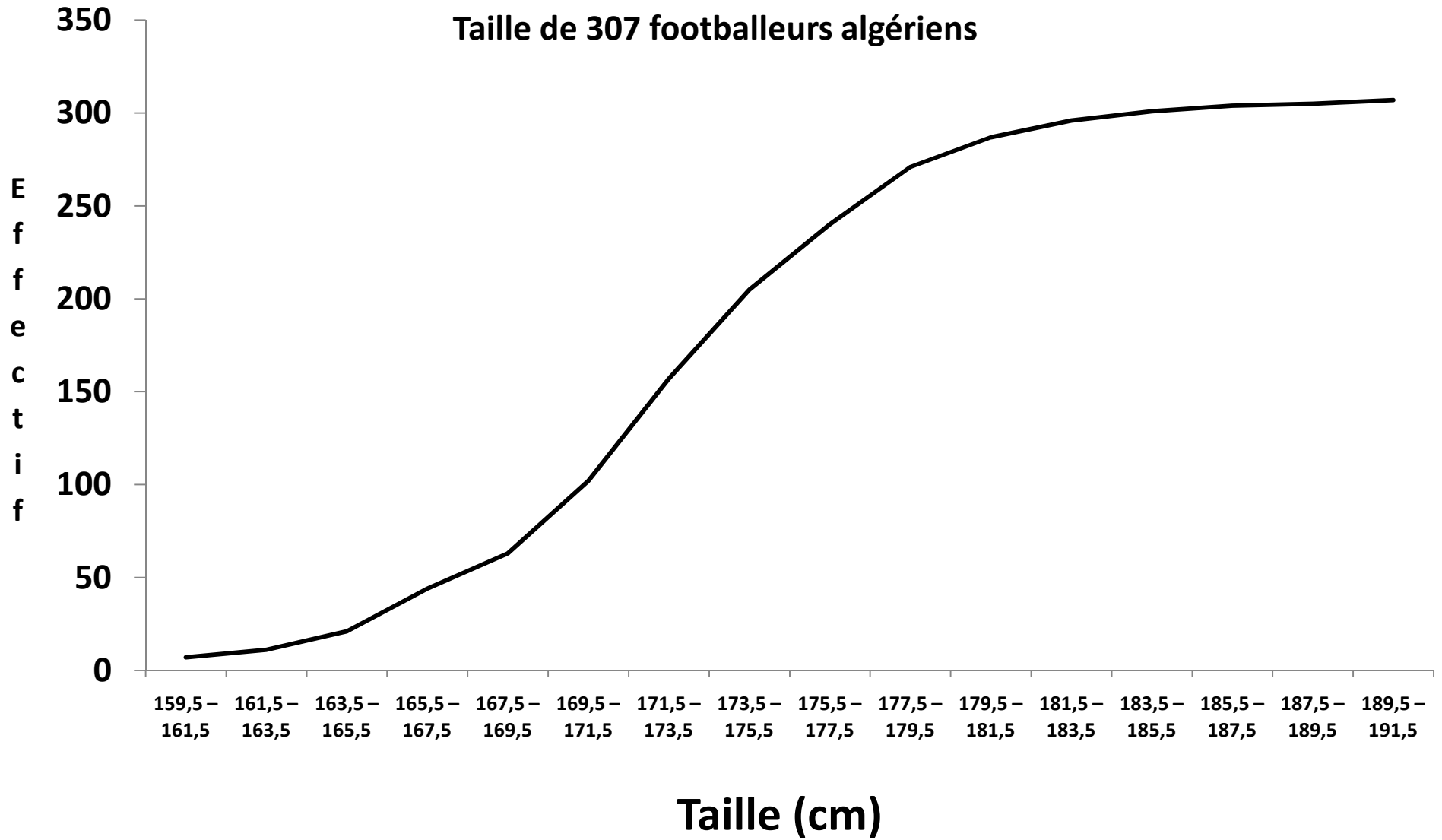


Diagramme Intégral

Taille de 307 footballeurs algériens

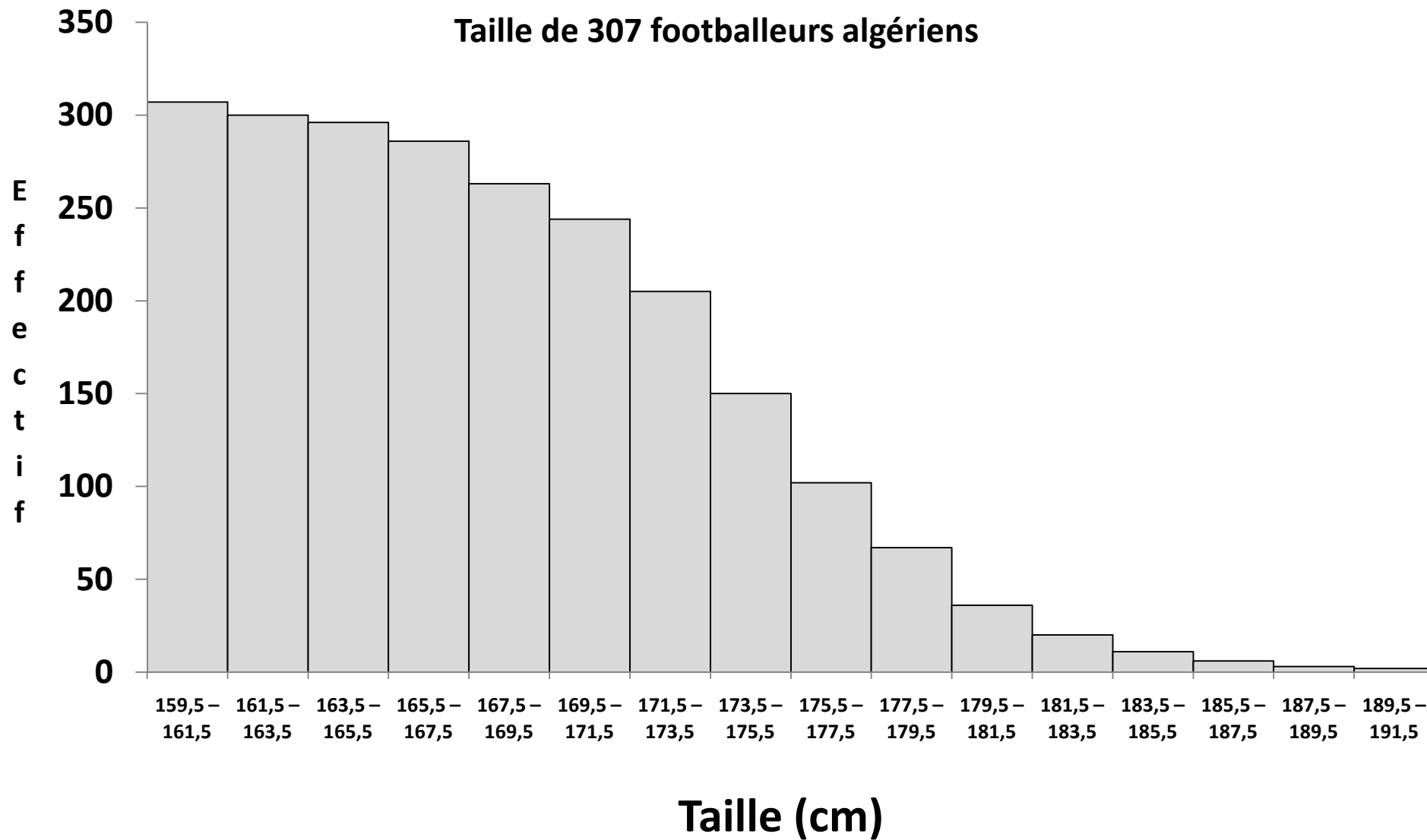


Diagramme Intégral

Taille de 307 footballeurs algériens

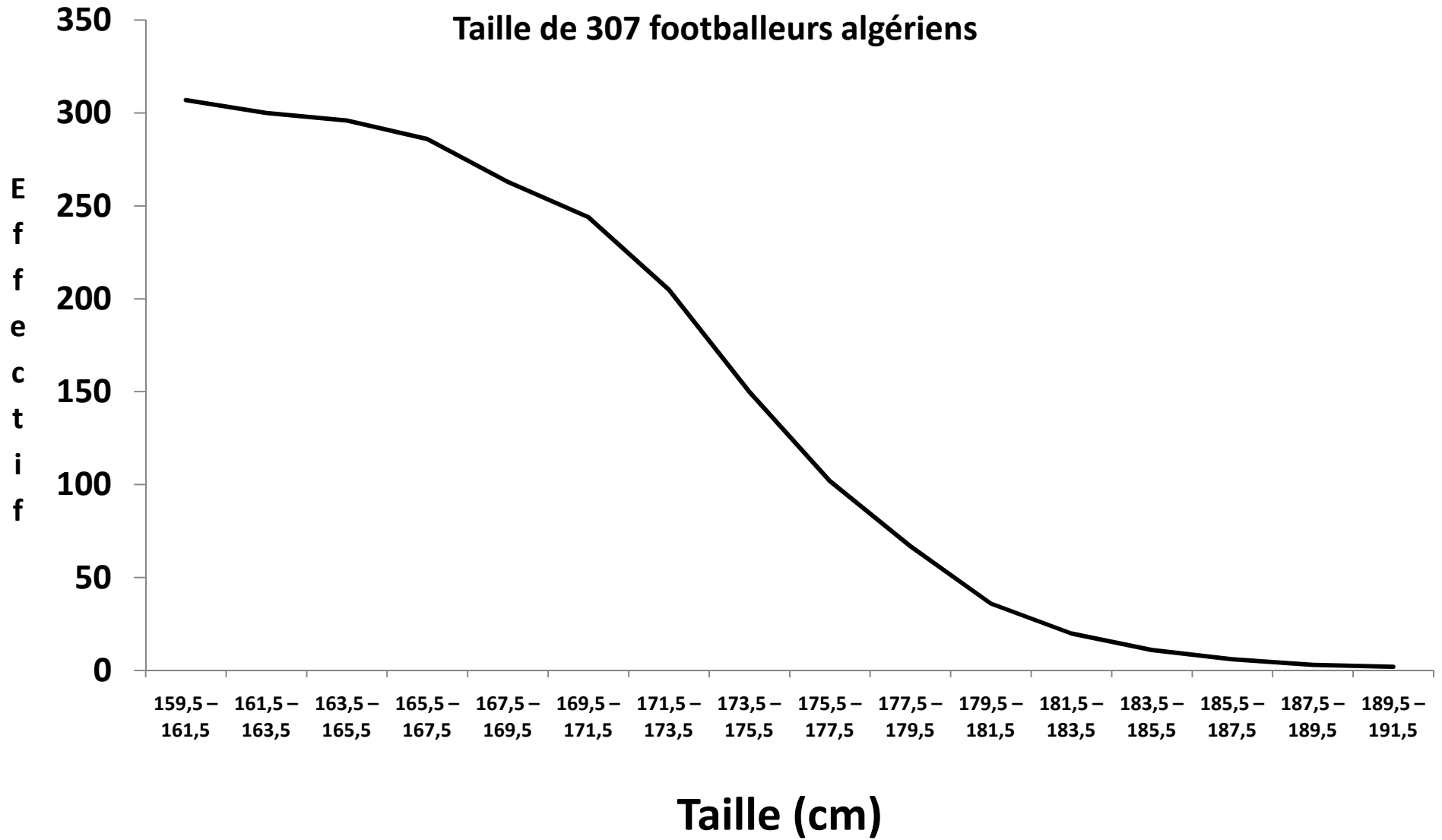


Diagramme Intégral