

Que dire d'un histogramme ?

Maïeutique en statistique

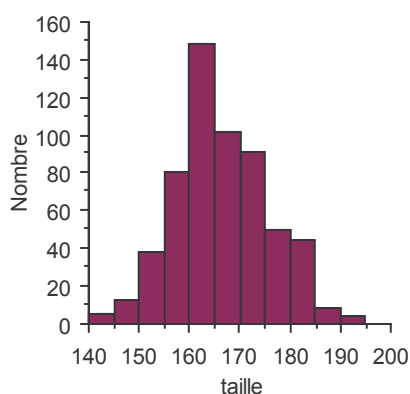
Léo Sila

Pièce en trois actes

Les données dont on parle dans ce texte ont été recueillies en 2005 dans des classes de l'Isère (voir le fichier des données).

Acte I Un ou deux histogrammes

M. : Voilà, j'ai obtenu ça (*il montre le dessin ci-dessous*). Qu'est-ce que tu en dis ?



S. : Je dis : c'est un histogramme. En violet, comme ça, c'est joli.

M. : Des goûts et des couleurs... Je ne suis pas statisticien, je n'ai pas la tête dans des nuages de données, je suis face à des étudiants. Alors qu'est-ce que je leur dis ?

S. : Tu peux leur demander s'ils trouvent que c'est joli.

M. : Oh, ça va !

S. : Pourquoi veux-tu leur dire quelque chose ? Ou l'as-tu trouvé, cet histogramme ? Au supermarché, au rayon des *data*, en promotion : "*data* de premier choix", "pas cher" ?

M. : Justement, pas du tout ! Ce sont de vraies données.

S. : De vraies données ?

M. : Mes étudiants, *prof-stagiaires-en situation*, ont demandé leur taille à leurs élèves.

S. : Des données qu'on n'a pas recueillies soi-même ne sont pas de vraies données ?

M. : Euh, enfin, si. Mais créer ses données soi-même, c'est plus ludique...

S. : Donc tu voulais jouer avec des données ?

M. : Disons plus simplement, donner du sens à ma séquence d'enseignement.

S. : Quelle générosité : donner du sens, dans un monde où tout se vend ! Je suis ému. Et à quoi tu veux donner du sens ?

M. : Aux histogrammes.

S. : Aux histogrammes en général, à la notion d'*histogramme en soi* comme disent les philosophes, ou à cet histogramme là ?

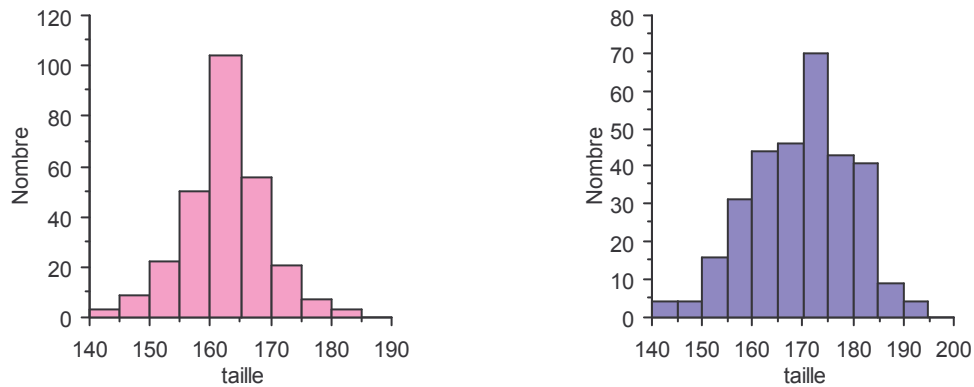
M. : Commençons par ce qui est sous nos yeux ! Ensuite, je ferai un camembert pour la répartition filles garçons, et un autre pour la répartition des années de naissance des élèves.

S. : Quel festin ! Les anglais ne parlent pas de camemberts mais de « pie chart », c'est-à-dire « diagrammes en tarte », à chacun ses plats favoris. On verra s'il y a des directives européennes. Fais voir tes données.

M. : Tu veux voir quoi dans les données ?

S. : Ce que tu veux.

M. : Voilà l'histogramme des filles et celui des garçons. Grosso modo, les garçons sont plus grands.



S. : Tu vois, quand tu veux... Choix classique de couleurs. Maintenant on a deux jolis dessins.

M. : C'est tout ce qu'il y a à dire ?

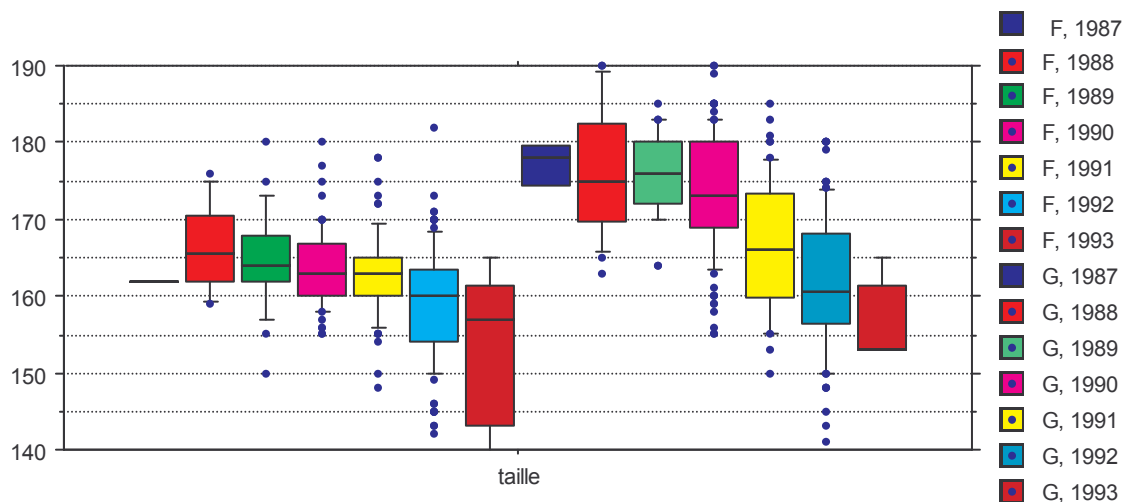
S. : As-tu bien regardé tes données ?

M. part en claquant la porte.

Acte II Diagrammes en boîtes

M. : J'ai bien regardé les données (*il montre le dessin ci-dessous*). Là, c'est vraiment joli !

S. : C'est toi qui dis que c'est joli maintenant. Que vas-tu dire à tes étudiants ?



M. : Je ne vais pas parler d'histogrammes mais de diagrammes en boîtes, comme ça ici je tiens compte à la fois du sexe et de l'année de naissance pour observer les tailles.

S. : Donc tu choisis d'adapter le traitement aux données ?

M. : Evidemment !

S. : Que vas-tu dire sur ces diagrammes ?

M. : Les filles sont plus petites que les garçons, et pour chaque sexe, on grandit entre 12 et 18 ans.

S. : Comme tu n'avais pas de questions, retrouver dans un mode d'expression (graphique) nouveau ce qu'on sait déjà est plein de bon sens. Et tes deux histogrammes d'hier, tu les as jetés ?

M. : Les voici, on dirait que c'est toi qui y tiens.

S. : Pourquoi un pas de 5cm ?

M. : Pourquoi pas ? Comme ça, il n'y a pas de « trous », c'est assez régulier, ça a une forme.

S. : Et les pics là, 160-165cm pour les filles, 170-175cm pour les garçons, c'est quoi ?

M. : Les tailles sont ici en nombre entier de centimètres ; mais de fait, la valeur la plus fréquente, le *mode*, est 160 cm à la fois pour les filles et pour les garçons avec 36 données sur 275 pour les filles et 27 données sur 312 pour les garçons !

S. : Juste 160 ? Ce *mode* me semble particulièrement élégant...

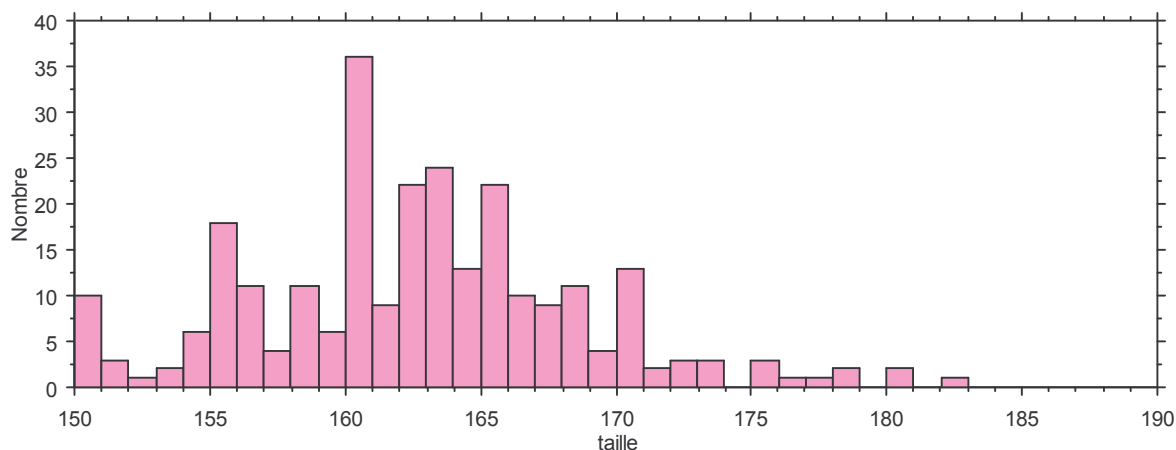
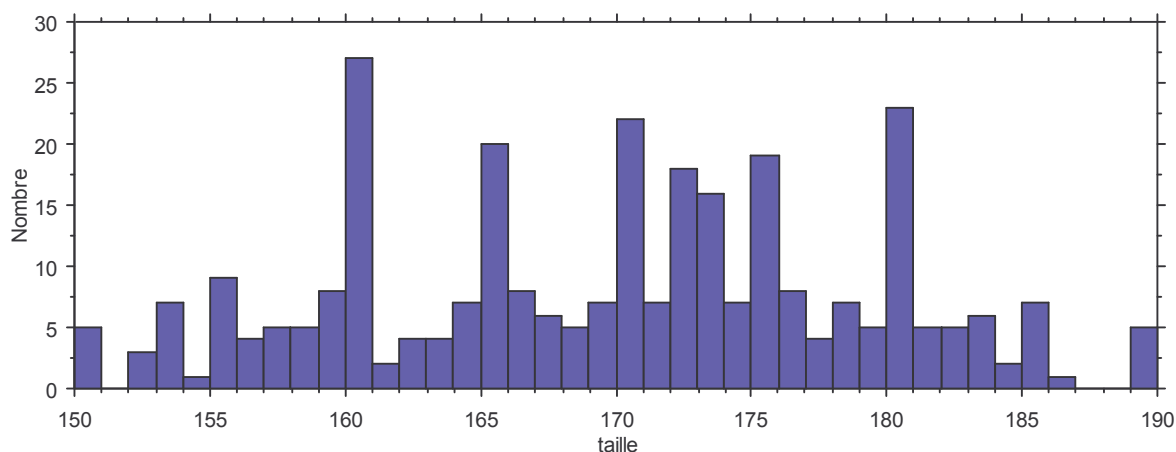
M. : ...mais, en fait, ce n'est pas normal.

S. : Normal, normal, encore des grands mots.

(M. *part en claquant la porte*).

Acte III Des nombres ronds

M. : Je suis déçu, complètement déçu ! Regarde (*il montre les dessins ci-dessous*) si on fait les histogrammes avec un pas de 1cm, ... il y a des pics presque tous les 5 cm. On ne grandit pas par bond de 5 cm ! Moi qui croyait avoir de vraies données ! Les élèves trichent, filles comme garçons ! Mes données, poubelle.



S. : N'as-tu pas demandé aux élèves : quelle est ta taille ? Pourquoi parles-tu de tricherie ? Tu as posé une question, tu as eu de vraies réponses. Tu n'as pas fait des mesures avec un protocole de mesure bien défini. D'ailleurs, si tu mesurais 179 ou 181cm, que dirais-tu ?

M. : Je pourrais dire que je mesure un mètre quatre vingt...

S. : J'ai fait plusieurs fois l'expérience de demander à des étudiants leur taille. J'ai aussi eu des pics sur des *nombres ronds*... *Le tropisme des nombres ronds*.

M. : C'est comme ça pour tous les histogrammes ?

S. : Devine...

M. : Si on pèse des objets avec une balance automatique qui inscrit directement les résultats dans un fichier, on n'a pas de raison d'observer ce tropisme.

S. : Sûrement. Essaie... Et les camemberts ?

M. : Eh bien, j'ai de quoi dire sur ces histogrammes et ces diagrammes en boîtes. Je pense qu'on n'aura plus faim pour du fromage ou un « pie ».

S. : Un jour, on pourra parler de « la forme » des histogrammes de taille.....

Fin