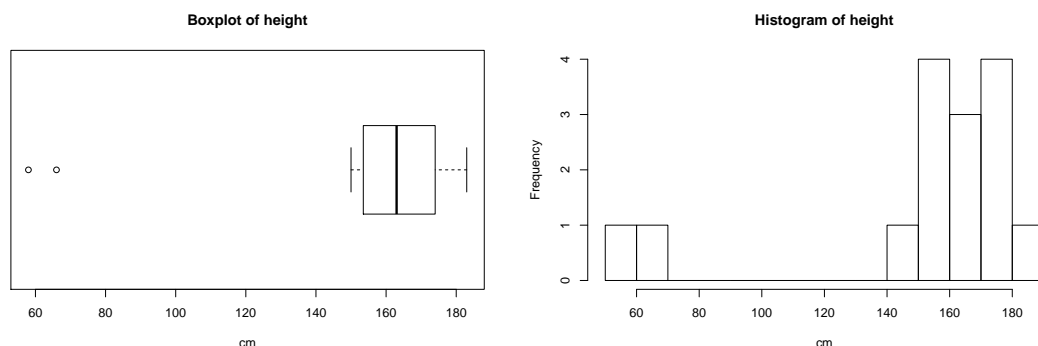


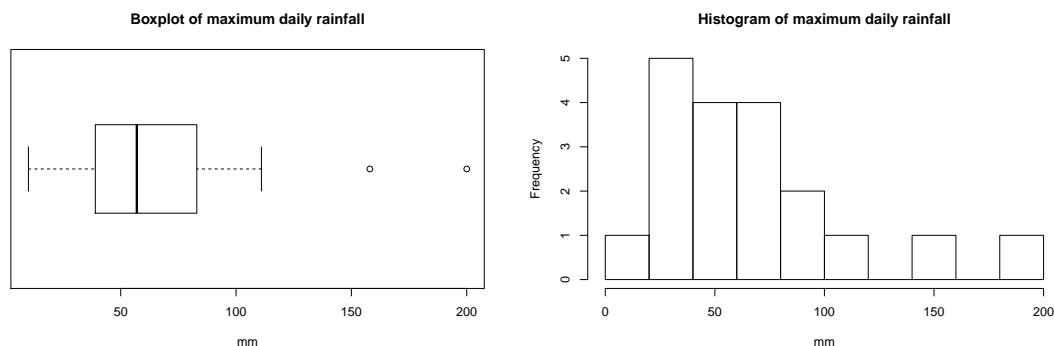
SÉRIE 2

Exercice 1. On a observé la taille de 15 femmes. Le boxplot et l'histogramme ci-dessous ont été construits avec ces données¹.



- La moyenne est égale à 152.3 cm et la médiane à 163 cm. Tracez la moyenne et la médiane dans les deux graphiques. Laquelle de ces deux caractéristiques représente le mieux la tendance centrale de la distribution des tailles ? Justifiez votre réponse.
- Si l'on considère seulement les tailles supérieures à 140 cm, on obtient une moyenne de 166.2 cm et une médiane de 165 cm. Commentez la phrase suivante : *“La médiane est plus robuste que la moyenne.”*
- L'écart-type des tailles observées est de 38.1 cm et l'écart inter-quartile est de 20.5 cm. Si l'on considère seulement les tailles supérieures à 140 cm, on obtient un écart-type de 11.1 cm et un écart inter-quartile de 18 cm. Laquelle de ces deux caractéristiques vous semble la plus robuste ?

Exercice 2. On a observé le maxima des pluies journalières du mois de novembre pour 19 années consécutives. Le boxplot et l'histogramme construits avec ces données² sont :



- La moyenne est égal à 69.32 mm et la médiane à 57 mm. Tracez-les dans les deux graphiques. Laquelle de ces deux caractéristiques représente le mieux la tendance centrale de la distribution ? Justifiez votre réponse.
- Pensez-vous qu'il s'agit d'une situation similaire à celle de l'exercice précédent ? (Cela n'est pas une question de statistique, mais plutôt une question d'interprétation des données.)

1. Les données sont tirées de *The World Almanac and Book of Facts*, 1975. Voir le corrigé pour plus de détails.

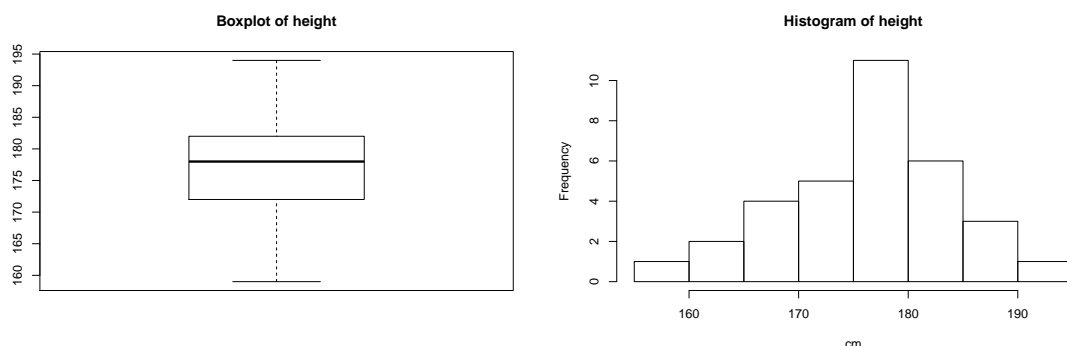
2. Les données sont tirées de <http://www.waterlog.info/>

Exercice 3. On a demandé à 1000 femmes âgées entre 25 et 29 ans quel était leur état civil. Les données recueillies sont présentées dans le tableau ci-dessous.

Etat civil	jamais mariée	mariée	veuve	divorcée
Proportion	0.353	0.574	0.002	0.071

- Vérifiez qu'il s'agit bien de fréquences relatives.
- Calculez le pourcentage de femmes divorcées ou jamais mariées.
- Calculez le pourcentage de femmes qui ne sont pas mariées.

Exercice 4. Les graphiques suivants représentent la taille, en centimètres, de 30 hommes :

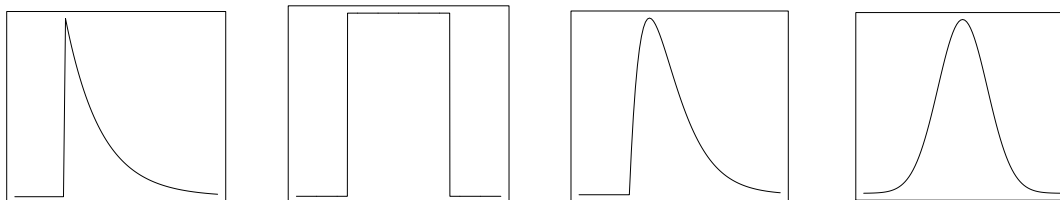


La moyenne et l'écart-type sont égales à $\bar{x} = 176.6$ cm et $s_x = 7.99$ cm respectivement. On va approximer la distribution des données par une distribution normale avec paramètres $\mu = \bar{x}$ et $\sigma = s_x$. En utilisant cette approximation,

- donnez le pourcentage d'observations attendues dans l'intervalle $[\bar{x} - s_x, \bar{x} + s_x]$. Comparez avec la fréquence relative des données observées dans cet intervalle, qui est de 63.3 %.
- calculez le pourcentage attendu d'hommes ayant une taille inférieure ou égale à 170 cm. Comparez avec la fréquence relative dans cet intervalle, qui est de 23.3 %.

Indication : vous pouvez centrer-réduire (standardiser) la distribution et ensuite utiliser la table de la loi $N(0, 1)$ (donnée dans le cours).

Exercice 5. Les graphiques suivants représentent les fonctions de densité de 4 distributions.



- Décrivez brièvement les caractéristiques attendues d'un échantillon de données tiré de chacune de ces distributions.
- Laquelle de ces densités choisiriez-vous pour approximer la distribution des données de l'exercice 2? Et pour les données de l'exercice 4?
- Est-ce qu'il y a une densité normale parmi ces quatre densités?