

Exercice MLL-28-6-2025

Minosoa Barphyl ANDRIAMIHARISOA

2025-07-12

Questions associées aux graphiques (base de données : datasets) Avec fonctions graphiques de base en R

1. Histogramme — `mtcars$mpg`

Fonction : `hist(mtcars$mpg)`

- Quelle est la consommation de carburant la plus fréquente (en miles per gallon) ?
 - La distribution est-elle symétrique, asymétrique à gauche ou à droite ?
 - Quelle est l'étendue approximative de la variable mpg ?

2. Diagramme en barres — `table(iris$Species)`

Fonction : `barplot(table(iris$Species))`

- Combien d'espèces différentes sont représentées dans le jeu de données iris ?
 - Quelle espèce est la plus fréquente ?
 - Quelle espèce est la moins représentée ?

3. Boxplot — `mpg ~ cyl`

Fonction : `boxplot(mpg ~ cyl, data = mtcars)`

- Quelle catégorie de cylindres a le plus grand mpg médian ?
 - Y a-t-il des valeurs aberrantes ? Dans quel groupe ?
 - Que peut-on dire de la variabilité de la consommation selon le nombre de cylindres ?

4. Nuage de points — `mtcarswtvsmtcarsmpg`

Fonction : `plot(mtcarswt, mtcarsmpg)`

- Quelle est la relation entre le poids de la voiture et la consommation (mpg) ?
 - Est-ce une relation linéaire ? Inverse ?
 - Peut-on identifier un groupe de voitures particulièrement économes ?

5. Courbe — `pressure` Fonction : `plot(pressuretemperature, pressurepressure, type = "l")`

- Comment la pression change-t-elle avec la température ?
 - À partir de quelle température observe-t-on une forte augmentation de la pression ?
 - Quelle est la forme générale de la courbe ?

6. Diagramme circulaire — `irisSpecies` Fonction : `pie(table(irisSpecies))`

- Quelle est la proportion de chaque espèce dans le jeu de données iris ?
 - Quelle espèce est majoritaire ?
 - Comment ce graphique se compare-t-il avec le diagramme en barres précédent ?

7. Matrice de nuages de points — `iris[1:4]`

Fonction : `pairs(iris[1:4])`

- Quelles paires de variables semblent fortement corrélées ?
 - Peut-on distinguer visuellement les espèces ?
 - Quels nuages montrent une séparation nette entre espèces ?

8. Densité — `airquality$Temp`

Fonction : `plot(density(airquality$Temp, na.rm = TRUE))`

- Quelle est la température la plus probable ?
 - La courbe présente-t-elle un ou plusieurs pics ?
 - Est-ce que la distribution semble normale ?

9. Graphique en escalier — Nile Fonction : Graphique en escalier — Nile

Fonction : `plot(Nile, type = "s")`

- En quelle année observe-t-on le débit maximum du Nil ?
 - Y a-t-il une tendance générale dans les débits annuels ?
 - Y a-t-il des années particulièrement faibles ou élevées ?

10. Diagramme en barres groupé — `mtcars` `cyl` vs `mtcars` `gear`

Fonction : `barplot(table(mtcars$cyl, mtcars$gear), beside = TRUE)`

- Quel nombre de cylindres est le plus associé aux voitures à 4 vitesses ?
 - Y a-t-il des combinaisons de cylindres et de vitesses très rares ?
 - Quelle combinaison est la plus fréquente ?

BONUS — `ggplot2`

a) Histogramme `ggplot2` — `mpg`

Fonction : `ggplot(mtcars, aes(x = mpg)) + geom_histogram()`

- Quelle est la consommation la plus fréquente ?
 - Comment le choix de `binwidth` influence-t-il la lisibilité du graphique ?

b) Nuage de points coloré par `cyl`

Fonction : `ggplot(mtcars, aes(x = wt, y = mpg, color = factor(cyl))) + geom_point()`

- Quel groupe de cylindres présente la meilleure efficacité (haut `mpg`) ?
 - Les groupes de cylindres sont-ils bien séparés visuellement ?

c) Boxplot ggplot2 — iris

Fonction : `ggplot(iris, aes(x = Species, y = Sepal.Length)) + geom_boxplot()`

- Quelle espèce a les sépales les plus longs en moyenne ?
 - Quelle espèce présente la plus grande variabilité de longueur de sépales ?
-

(c) MLL-2025