

Statistique descriptive avec R

TP 8

Nous allons travailler sur les données `mtcars` déjà présentes dans **R**.

1. Affichez les données et regardez la description dans l'aide des données. Pour les 11 colonnes, décrivez pour chacune s'il s'agit de variables

- **Quantitatives continues** : des variables numériques qui correspondent à une valeur mesurée et distribuées continument. Exemple : la surface en m² d'un appartement.
- **Quantitatives discrètes** : des variables numériques qui correspondent à une réalité et distribuée sur des valeurs discrètes précises et souvent en petit nombre : le nombre de pièces d'un appartement.
- **Qualitatives ordinaires** : des variables qualitatives où les modalités ont un ordre logique, parfois représentées numériquement mais dont le nombre n'a plus de sens numérique : le diagnostic de performance énergétique (DPE) d'un appartement.
- **Qualitatives nominales** : des variables qualitatives pour lesquelles il n'y a aucun ordre, parfois représentées numériquement. Exemple : le type de chauffage de l'appartement.

2. Avec les fonctions `table` et `prop.table` (usage : `table(x)` et `prop.table(table(x))`), affichez les fréquences et les proportions des variables `cyl` et `am`, puis des combinaires `cyl, am` en ajoutant le second en second argument de `table`.

3. Affichez l'histogramme de la variable quantitative continue `hp`.

4. Avec la fonction `pie` qui prend en premier argument le résultat d'un `table`, affichez le camembert de répartition des `am`. On essayera de mettre en `labels` la signification "**Automatique**" et "**Manuelle**" plutôt que `0` ou `1`, et le pourcentage de répartition.

5. Pour afficher les fréquences des modalités d'une variable qualitative, on peut utiliser un diagramme en barres avec `barplot` qui prend en premier argument le résultat d'un `table`. Affichez la répartition des `vs`.

6. Pour comparer des données quantitatives par groupes qualitatifs, on utilise un `boxplot`. Affichez `mpg` en fonction de `am`, la fonction `boxplot` fonctionne comme la fonction `lm` pour la formule et la mise des données.

7. Grâce à la fonction `cut`, ajouter une colonne `classehp` qui transforme la variable quantitative continue en une variable qualitative ordinale :

hp	classehp
[0, 100["Faible"
[100, 200["Moyenne"
[200, +∞["Forte"

8. Affichez les `mpg` en fonction de votre nouvelle classe avec un `boxplot`