

Premiers pas avec R et les statistiques descriptives

Ahmed Fouad EL HADDAD

UPEC

September 24, 2025

Qu'est-ce que R et RStudio ?

- **R** : Un langage de programmation dédié au calcul statistique et à la visualisation des données.
- **RStudio** : Un environnement de développement intégré (IDE) qui rend l'usage de R plus simple.
 - Organisation en fenêtres : scripts, console, graphiques, environnement.
 - Interface pensée pour la recherche et l'enseignement.
- **Pourquoi utiliser R en science politique ?**
 - Analyser de grands jeux de données (enquêtes électorales, recensements, bases parlementaires).
 - Produire des visualisations (cartes, graphiques de tendances).
 - Réaliser des tests statistiques (comparaison de groupes, régressions).

Comment installer R et RStudio ?

Étape 1 : Installer R

- Téléchargement depuis <https://cran.r-project.org/>.
- Choisir la version selon votre système (Windows, Mac, Linux).

Étape 2 : Installer RStudio

- Télécharger RStudio Desktop (version gratuite) depuis <https://posit.co/download/rstudio-desktop/>.

Premier lancement :

- 1 **Éditeur de script** : écrire et sauvegarder du code (en haut à gauche).
- 2 **Console** : exécuter le code ligne par ligne (en bas à gauche).
- 3 **Environnement** : voir les objets créés (en haut à droite).
- 4 **Graphiques/Aide/Fichiers** : visualiser les sorties graphiques (en bas à droite).

Interface de RStudio

1. SOURCE

Click "Run" to send your code to the console

This is where you write your code!

Your code will not be evaluated until you "Run" them to the console.

2. CONSOLE

This is where your code from the Source is evaluated by R.

You can also use the console to perform quick calculations that you don't need to save

3. Environment / History

Here you can see what objects are in your working space (Environment) or view your command history (History)

4. Files / Plots / Packages / Help

Here you can see file directories, view plots, see your packages, and access R Help

Premières commandes

Exécuter des commandes dans la console :

Code R

```
print("Bonjour, R !")
```

Sortie attendue :

Résultat

```
[1] "Bonjour, R !"
```

Intuition : la console est comme une calculatrice améliorée : chaque commande saisie produit un résultat immédiat.

Variables et assignments

Créer et utiliser des variables :

Code R

```
x <- 5
```

```
y <- 10
```

```
z <- x + y
```

```
print(z)    # Affiche 15
```

Intuition : une variable est un espace mémoire nommé où une valeur est stockée. Ici, x, y, z sont comme des boîtes contenant des nombres.

Comprendre les fonctions

Définir et utiliser une fonction :

Code R

```
addition <- function(a, b) {  
  return(a + b)  
}
```

```
resultat <- addition(3, 5)  
print(resultat)    # Donne 8
```

Intuition : une fonction est une machine qui prend des **arguments**, applique une opération et retourne un **résultat**. Beaucoup de fonctions utiles sont déjà intégrées à R : `mean()`, `median()`, `summary()`, etc.

Types de données en R

Principaux types de données :

Type	Exemple	Description
Numérique	<code>x <- 3.14</code>	Nombres décimaux
Entier	<code>y <- 10L</code>	Nombres entiers
Caractère	<code>nom <- "Paris"</code>	Chaînes de texte
Logique	<code>actif <- TRUE</code>	Valeurs booléennes

Intuition : les objets R se distinguent d'abord par leur type : numérique, texte, booléen. On peut tester le type avec des fonctions natives comme `class()` ou `is.numeric()`.

Vecteurs

Créer un vecteur :

Code R

```
partis <- c("PJD", "USFP", "PAM")  
sieges <- c(125, 20, 102)
```

Accéder aux éléments :

Code R

```
partis[1]      # Retourne "PJD"  
sieges[2]     # Retourne 20
```

Intuition : un vecteur est une séquence **homogène**. Il peut contenir uniquement des nombres, ou uniquement du texte, mais pas un mélange.

Data Frames

Créer un data frame :

Code R

```
deputes <- data.frame(
  Nom = c("Ahmed", "Fatima", "Omar"),
  Parti = c("PJD", "USFP", "PAM"),
  Age = c(45, 50, 38)
)
```

Accéder aux données :

Code R

```
head(deputes)           # Premières lignes
deputes$Nom             # Colonne "Nom"
deputes[1, 2]           # 1ère ligne, 2e colonne
```

Intuition : un data frame est une **table à deux dimensions**. Chaque colonne est un vecteur homogène, mais les colonnes peuvent être de types différents (texte, nombre, booléen). C'est l'objet central de l'analyse en sciences sociales.