Statistique Descriptive
Chapitre 03 : Introduction à RStudio

Présentation de RStudio comme un environnement de développement intégré (IDE) pour le langage R.

Importance de RStudio pour l'analyse de données, la visualisation et le développement de modèles statistiques.

Le langage R



langage de programmation + environnement logiciel



- analyse statistique
- visualisation de données
- science des données



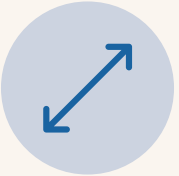
Open Source



communauté d'utilisateurs et de développeurs très active



vaste écosystème de bibliothèques étendant ses fonctionnalités.



capacités de visualisation de données



intégration avec d'autres langages et outils

Le langage R : exemple

```r
# paramètres généraux de l'exercice
exo_fn <- "01-syllabus-exoTemp.csv"
class_min <- 11
class_max <- 21

# Charger les données depuis le fichier CSV
input <- read.csv(exo_fn)

# Vérifier les premières lignes des données
head(input)
head(input$inputData)

# 13. Calculer et afficher la moyenne
moyenne <- mean(input$inputData, na.rm = TRUE)

# 14. Calculer et afficher la variance
variance <- var_p(input$inputData)
```

```r
# 22. Afficher la courbe cumulative
plot(ecdf(input$inputData),
     main = "fréquences cumulées",
     ylab = "Proportion Cumulative",
     col = chart_color )

# 23. Afficher la boîte à moustaches
boxplot(input$inputData,
        main = "Boîte à Moustaches",
        col = chart_color,
        border = chart_border)
```

L'IDE RStudio



Environnement de développement intégré (IDE) pour R



Interface utilisateur graphique (GUI)



Gestion de projet



Aperçu de l'environnement de travail



Outils de développement (débogage, gestion des paquets)



Assistance intégrée



Visualisation et exportation simples des graphiques



Collaboration et partage

RStudio

File Edit Code View Plots Session Project Build Tools Help

Untitled1*

Source on Save | Run | Source

```
1
2  rm(list = ls())
3  N <- 1000
4  u <- rnorm(N)
5  x1 <- -2 + rnorm(N)
6  x2 <- 1 + x1 + rnorm(N)
7  y <- 1 + x1 + x2 + u
8  r1 <- lm(y ~ x1 + x2)
9
10
```

10:1 (Top Level) R Script

Console ~/

```
Tapez <Entrée> pour voir le graphique suivant :
Tapez <Entrée> pour voir le graphique suivant :
Tapez <Entrée> pour voir le graphique suivant :
>
> ?lm
> rm(list = ls())
> N <- 1000
> u <- rnorm(N)
> x1 <- -2 + rnorm(N)
> x2 <- 1 + x1 + rnorm(N)
> y <- 1 + x1 + x2 + u
> r1 <- lm(y ~ x1 + x2)
>
```

Workspace History

Import Dataset

Values	
N	1000
r1	lm[12]
u	numeric[1000]
x1	numeric[1000]
x2	numeric[1000]
y	numeric[1000]

Files Plots Packages Help

R: Fitting Linear Models — Find in Topic

lm {stats} R Documentation

Fitting Linear Models

Description

`lm` is used to fit linear models. It can be used to carry out regression, single stratum analysis of variance and analysis of covariance (although `aov` may provide a more convenient interface for these).

Usage

```
lm(formula, data, subset, weights,
   method = "qr", model = TRUE, x =
   singular.ok = TRUE, contrasts =
```

Arguments

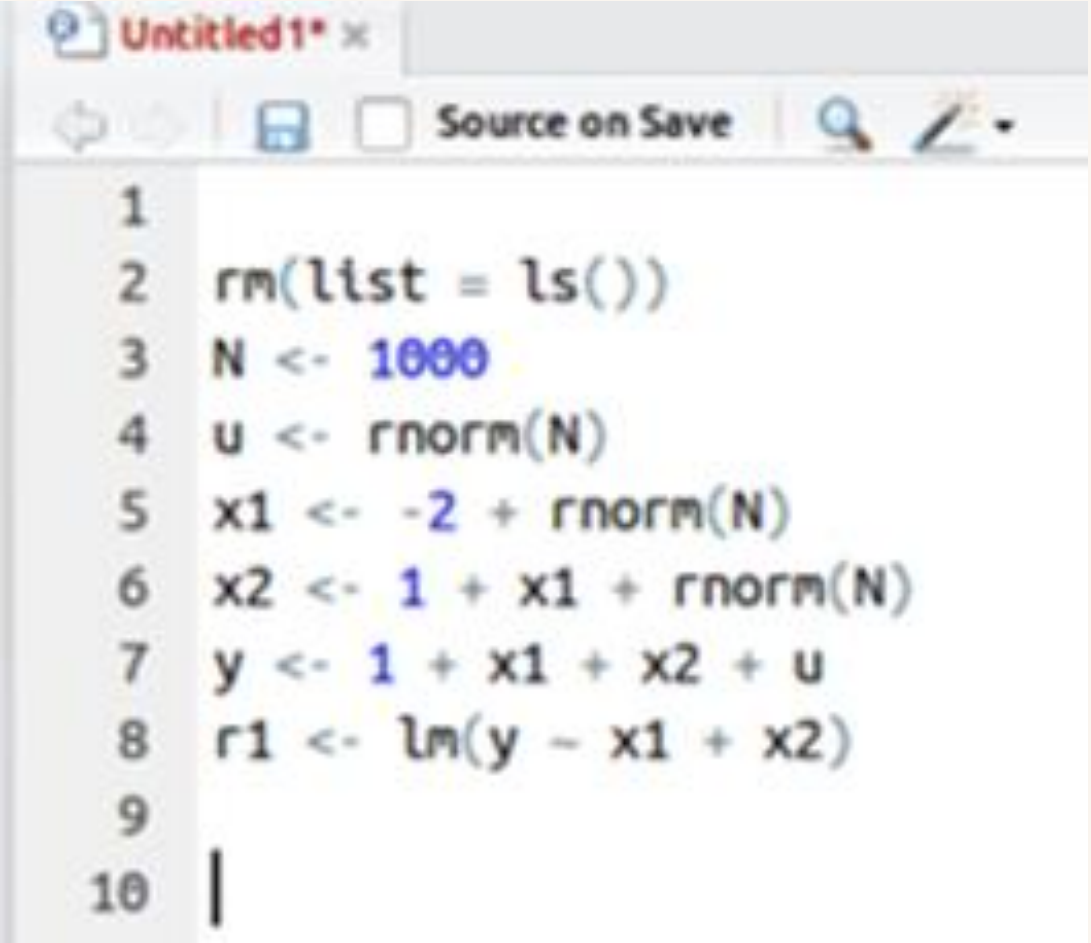
Pourquoi Utiliser RStudio ?

- Interface conviviale (?) pour travailler avec R.
- Outils intégrés pour
  - l'édition de code
  - la visualisation de données
  - la gestion des projets
- Large communauté d'utilisateurs et de ressources disponibles pour l'apprentissage et le dépannage.

Installation de RStudio

- Étape 1 : Télécharger R
  - depuis le site CRAN
  - depuis le site Burotix
- Étape 2 : Télécharger Rstudio
  - depuis le site Rstudio
  - depuis le site Burotix
- Étape 3 : Installer R et RStudio
  - en suivant les instructions à l'écran
- Étape 4 : Lancer RStudio

Interface de RStudio



Éditeur de Script

Emplacement

- panneau supérieur gauche

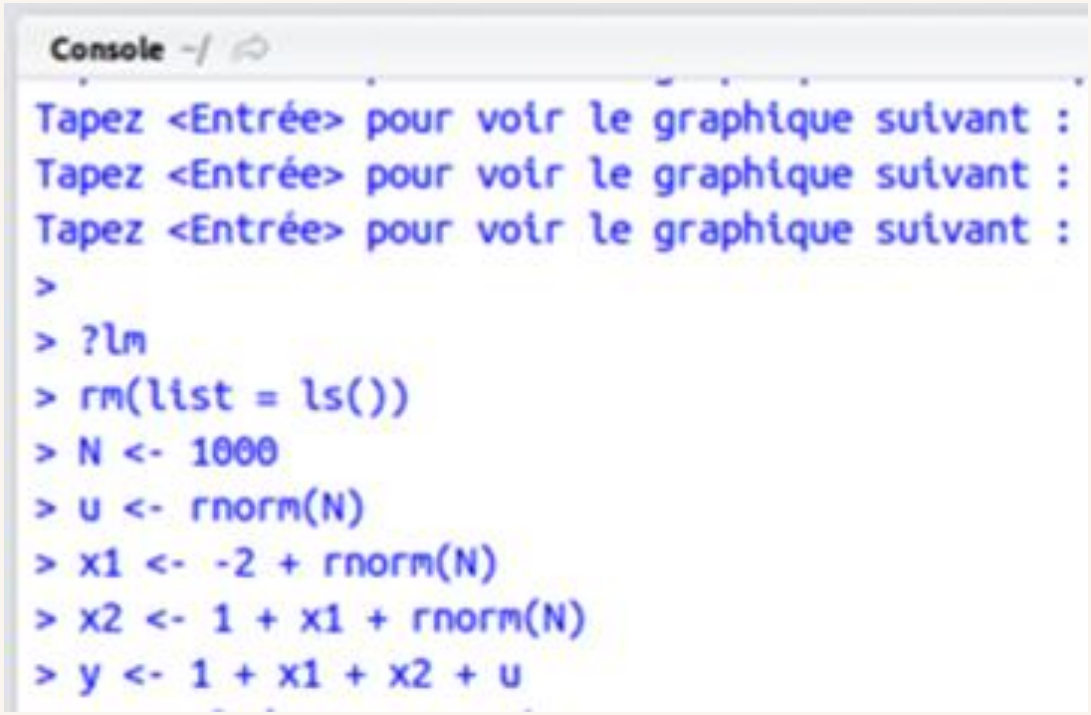
Fonction

- écrire et éditer des scripts R
- coloration syntaxique
- complétion de code
- indentation automatique

Utilisation

- écrire des fonctions, des analyses de données et des scripts longs

Interface de RStudio



Console

Emplacement

- panneau inférieur gauche

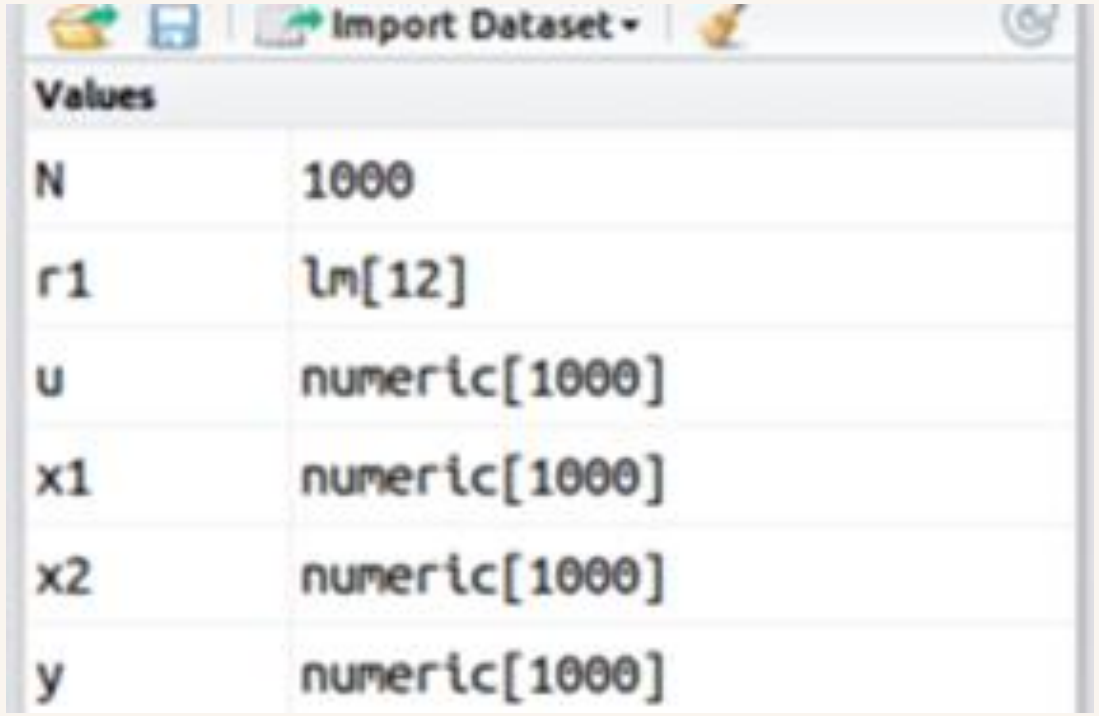
Fonction

- exécuter des commandes R directement
- affichage du résultat des commandes exécutées

Utilisation

- tester des commandes rapidement
- voir les résultats en temps réel

Interface de RStudio



Environnement ; Histoire

Emplacement

- panneau supérieur droit

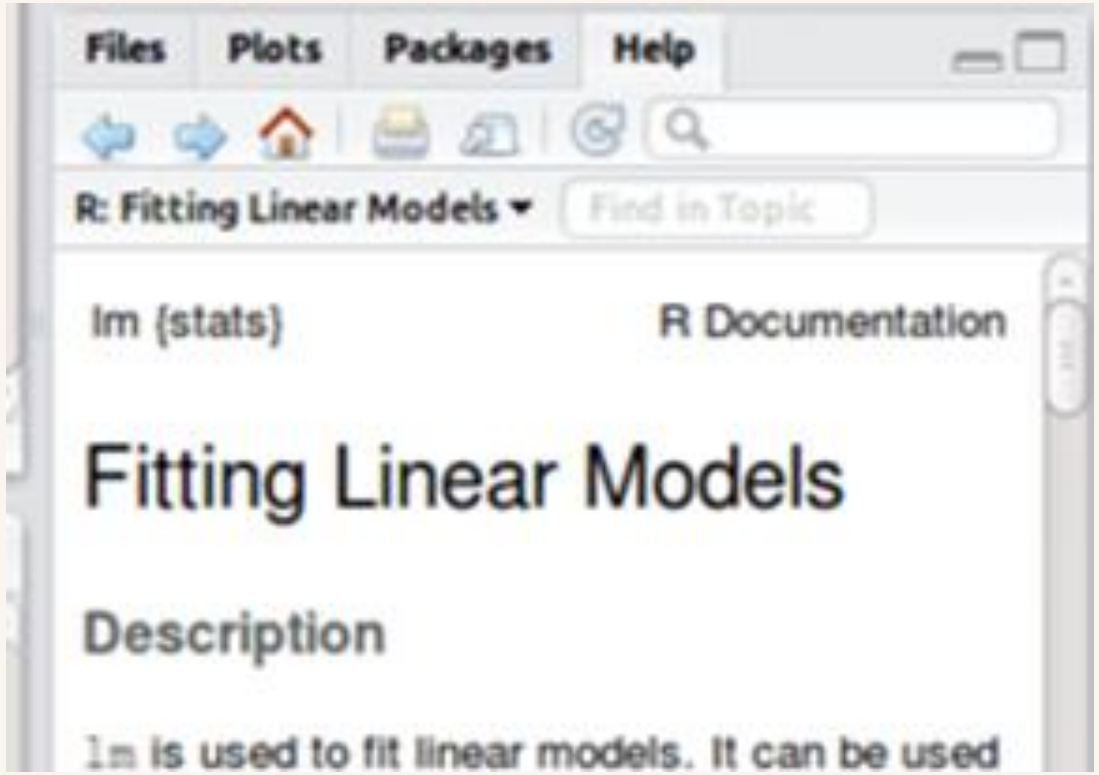
Fonction

- affichage des objets (variables, fonctions, etc.) de l'environnement de travail
- historique des commandes exécutées

Utilisation

- gérer les objets en mémoire
- revoir les commandes précédentes

Interface de RStudio



Fichiers ; Graphiques ; Packages ; Aide

Emplacement

- panneau inférieur droit

Fonction

- Fichiers : gestion des fichiers dans le répertoire de travail actuel
- Graphiques : affichage des graphiques générés par le code R
- Packages : gestion des packages installés et installation de nouveaux
- Aide : documentation et aide sur les fonctions et packages R.

Utilisation

- facilitation de la navigation dans les fichiers
- visualisation des graphiques
- gestion des packages
- accès à la documentation

Statistiques Univariées avec RStudio

- Exemple du syllabus : les températures
  - données d'entrée : `01-syllabus-exoTemp.csv`
  - traitement : `01-syllabus-exoTemp.R`
  - (données de sorties : `01-syllabus-solTemp.csv`)
- Lancer Rstudio
- Charger le fichier de traitement
- Exécuter ce fichier ligne par ligne
- Enjoy !

Visualisation de données avec RStudio

Ressources pour apprendre R

- Tutoriels et cours en ligne
  - R for Data Science : livre en ligne gratuit, Hadley Wickham et Garrett Grolemund, R for Data Science.
  - Coursera : Cours en ligne sur R, tels que "Data Science: Foundations using R", Johns Hopkins University.
  - DataCamp : Plateforme interactive pour apprendre R avec des cours comme "Introduction to R".
- Communautés et forums
  - Stack Overflow
  - RStudio Community : forum dédié aux utilisateurs de RStudio
  - R-bloggers : blog communautaire
- Documentation officielle
  - CRAN (Comprehensive R Archive Network) : site officiel, CRAN
  - RStudio Documentation : Documentation et tutoriels, RStudio.