
TD statistique-économétrie

Gabrielle GAMBULI

gabrielle.gambuli@gmail.com

COURS I

27 janvier 2026



Plan

- Présentation du cours
- Modalités d'évaluation
- Introduction à R
- Exercice

Organisation du TD

- 12 séances de 2h
- Mardi 15h30-17h30, SRB : salle G303
- Du 27/01/2026 au 24/04/2026
- 7-8 séances d'application sur R
- 4-5 séances de présentation de projets en binôme

OBJECTIFS

- Saisir l'intérêt de l'approche économétrique pour répondre à des questions centrales
- Être capable de traiter de manière autonome des questions économétriques simples avec R
- Compétences en économétrie très valorisées sur le marché du travail

EVALUATION

- Note sur 20 : 25% code et écrit + 75% présentation
- Rendre la liste des binômes pour le 03/02
- Choix du sujet et des données pour le 17/02
- Calendrier de passage tiré au sort
- M'envoyer la base de données, le script R (ou Rmd ou quarto) et les diapositives de la présentation par mail

CONSIGNES POUR LE PROJET

- Concevoir une **question de recherche** à laquelle une étude économétrique peut répondre
- Construire une **base de données** pour répondre à la question à partir d'une ou plusieurs sources
- Proposer une **stratégie économétrique** adaptée pour répondre à la question (appliquée sous R)
- Interpréter les **résultats**, leur validité statistique et proposer une réponse à la question soulevée
- Copier-coller les commandes R vues en cours ne suffit pas

CONSIGNES POUR L'EXPOSÉ

- **Objectif :** Présenter la problématique, les données, la méthode, les résultats et leur interprétation
- **Durée de l'exposé :** environ 30 minutes par groupe
- **Exemples de bases de données :** datasets dans les packages de R, INSEE, data.gouv.fr, Eurostat, OCDE, FMI, Banque de France, BCE, World Bank, data.gouv.fr...
- **Méthodes vues en cours:** MCO et tests associés, event studies, CAPM (Capital Asset Pricing Model), endogénéité et variables instrumentales, modélisation sur données de panel, variables qualitatives (modèles logit/probit)
- **Montrer les résultats :** tableaux/graphiques de statistiques descriptives, régressions et tests

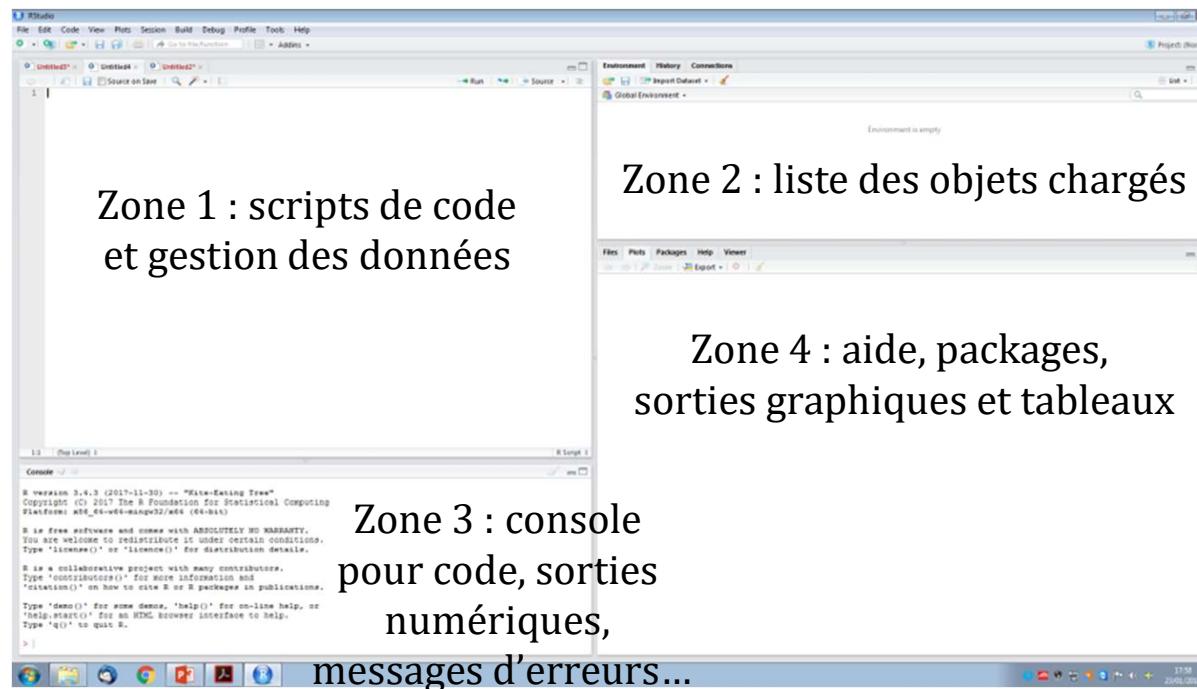
RÉFÉRENCES BIBLIOGRAPHIQUES

- « Séries temporelles avec R » d'Yves Aragon, Editions Springer
- « Régression avec R » de P-A Cornillon et E. Matzner-Lober, Editions Springer
- « Applied econometrics » with R de Kleiber, C., & Zeileis, A. Editions Springer

LE LOGICIEL R VIA L'ENVIRONNEMENT RSTUDIO

- **R** : langage et logiciel libres et gratuits pour l'extraction et le traitement de données, l'analyse statistique et économétrique, et les graphiques
 - Téléchargement : <https://cran.r-project.org/>
- **RStudio** : environnement gratuit avec des outils pour R et Python
 - Téléchargement : <https://posit.co/download/RStudio-desktop/>

LE LOGICIEL R VIA L'ENVIRONNEMENT RSTUDIO



FICHIERS ET OBJETS DE R

- **Fichiers** de différents formats : Rdata, csv, txt, SAS...
- Différents **objets** :
 - Vecteurs **c()** composés de valeurs appelées éléments
 - Facteurs **factor()** comme des vecteurs, mais pour des variables qualitatives ordinaires et nominales
 - Matrices **matrix()** pour 2 dimensions, **data.frames()** rangent des valeurs dans des tableaux de données
 - Listes **list()** rangent des éléments hétérogènes
 - Fonctions **nom_fonction()**

PACKAGES DE R

- Modules supplémentaires
- Installer un package : `install.packages(nom du package)`
- Charger un package : `library(nom du package)`
- Packages disponibles : <https://posit.co/products/open-source/rpackages/>
- Documentation : <https://cran.r-project.org/>
- Aide sur une fonction du package : `help(nom de la fonction, package=nom du package)` ou `?nom de la fonction`
- Principaux packages : dplyr (manipulation de data frames), ggplot2 (graphiques), MASS (régression multiple), lmtest (tests statistiques)

RÉPERTOIRE DE TRAVAIL

- Pour charger/importer des fichiers dans R ou les enregistrer
- **Connaître où se situe le répertoire de travail : `getwd()`**
- **Indiquer à R où travailler : `setwd("chemin d'accès au fichier")`**
- **Attention :** R ne reconnaît pas les backslashes (\), à remplacer par des slashes (/) dans le chemin d'accès
- **Exemple :** `setwd("C:/Users/Documents/TD_statistique_econometrie/")`

CRÉER UN SCRIPT.R

- Fichier externe pour enregistrer et réutiliser une série de commandes de R (format .R)
- **Exécuter le script :** `source("nomduscript.R")` exécute toutes les commandes enregistrées une par une (Ctrl Shift Entrée ou Ctrl Shift S)
- **Exécuter un code :** bouton « Run » (Ctrl Entrée)
- **Attention :** le répertoire de travail doit être le dossier où est enregistré le script
- **Importance de structurer le script :**
 - en-tête (information sur le contenu, date, auteur),
 - chargement des objets (packages, données),
 - titres, paragraphes, commentaires avec `#` (Ctrl Shift C),
 - sauvegarde du programme

IMPORTER/CHARGER DES OBJETS ET FICHIERS

- Charger un objet : `load()`
- Charger un `data.frame` : `read.table(nom de l'objet)`
- Charger un fichier `.csv` : `read.csv(file = "chemin/nom du fichier", sep = "séparateur", header = TRUE/FALSE)`
- Exemples :

```
data <- read.csv(file = "R:/test_ic.csv", sep = ";", header = TRUE)
data <- read.csv2(file = " R:/test_ic.csv ", sep=";", header = TRUE)
```
- Lire un fichier excel : `read_excel("chemin/nom du fichier", cheet = "nom de la feuille de calcul")`
avec le package `readxl`
- Information : `<-` est la flèche d'affectation

VISUALISER LES DONNÉES

- Visualiser les données : **View(nom de la base)**
- Attention : R est sensible aux majuscules et minuscules
- Voir les données après un **read** : **nom de la base** ou **nom de la base\$nom de la variable**

Exemple : mesdata <- read.csv(file = "R:/test_ic.csv", sep = ",", header = TRUE)

mesdata

mesdata\$nomvar1

- Visualisation graphique simple : **plot(nom de la variable)**

Exemple : plot (nomvar1 ~ nomvar2, data = mesdata, xlab = "titre abscisse", ylab = "titre ordonnée")

EDITER LES DONNÉES

- **Editer des données existantes à l'aide d'un mini-tableur : `edit(nom de la base)`**
- **Convertir en numérique : `as.numeric(nom de la variable)`**
Exemple : `newvar <- as.numeric(mesdata$nomvar1)`
- **Créer une nouvelle variable** avec la flèche d'affectation
Exemple : `newvar <- mesdata$nomvar1`

SAUVEGARDER/EXPORTER

- Sauvegarder des objets : `save (un ou plusieurs objets, file="nom du fichier")`
- Créer un fichier `.rdata` qui enregistre l'espace de travail : `save.image()`
- Enregistrer dans un fichier texte, un fichier csv ou fichier excel : `write.table()`, `write.csv2()`, `write_xlsx()` avec le package `openxlsx`

Exemple : `write.table(nom du data.frame à exporter, file = "nom nouveau fichier.txt" , sep = ',', row.names = FALSE)`

EXERCICE I – LES VECTEURS

ANALYSE DE DONNÉES DE GAIN