

# Introduction à R

---

Nom : **N'DOYE**

Prénom : **Abdoul-Aziz**

Année : **M1**

Semestre : **7**

Nature : **CM**

Volume horaire : **20 H**

ECTS / Coef : **2**

---

## Prérequis

Notions de base en statistique et économétrie

---

## Résumé

Au cours des dernières années, R a connu une ascension fulgurante qui s'est manifestée par une forte augmentation du nombre de ses utilisateurs due à la progression remarquable dans ses fonctionnalités et dans la variété de ses domaines d'application. Il dispose d'une très riche bibliothèque "package" de quasiment toutes les procédures et méthodes statistiques.

Dans sa structure, R est un langage de programmation d'une syntaxe voisine à celle du langage S et capable de manipuler des objets complexes sous forme d'objet.

Cet enseignement de 20h est un apprentissage de R en tant que logiciel statistique dans le traitement exploratoire de données statistiques et en tant que langage de programmation interprété pour tirer profit de sa particularité.

---

## Objectifs

- Découvrir le logiciel et apprendre à organiser son travail sous R
  - Comprendre les principaux objets de données et apprendre à les manipuler
  - Savoir appliquer les fonctions disponibles dans les packages
  - Décrire ses données et réaliser ses premières visualisations.
  - Programmer sous R : Ecrire un algorithme sous R - création de nouvelles fonctions de traitement de données sous R
  - Manipuler les fonctions des packages connectés du Tidyverse
- 

## Bibliographie

- R est un logiciel libre, complètement gratuit. Il dispose d'une documentation en ligne très complète avec des ressources externes sur les sites [cran.rproject.org](http://cran.rproject.org)

### Livres récents

- **Baumer B. & Horton, J.**, CRC Press (2021), Modern Data Science with R (2<sup>nd</sup> ed.),
  - **Chambers J.**, CRC Press (2016) : Extending R
  - Wickham H., CRC Press (2019), Advanced R (2<sup>nd</sup> edition),
  - **Wickham H. & Grolemund G.**, O'Reilly (2017), R for Data Science
-

---

# PLAN

---

- 1- **Utilisation de R comme un logiciel de traitement statistique des données :**
  - 1.1 Présentation générale du logiciel et de l'environnement R
    - 1.1.1 Description de R, commande de base, éditeurs de texte
    - 1.1.2 Utilisation de l'aide en ligne, documentation générale,
    - 1.1.3 Installation et utilisation des packages.
  - 1.2 Manipulation d'objet sous R
    - 1.2.1 Manipulation d'objet sous R : vecteurs, matrices, listes et data frame.
    - 1.2.2 Manipulation de données externes par les fonctions du package « base »
    - 1.2.3 Manipulation de fonctions disponibles dans les packages pour le traitement et l'analyse exploratoire de données
    - 1.2.4 Représentation graphique avec les packages de base
  
- 2- **Programmation modulaire sous R**
  - 2.1 Les structures de contrôle : Les alternatives et les boucles
  - 2.1 Création de nouvelles fonctions de traitement de données
  - 2.3 Ecriture d'algorithme en langage R
  
- 3- **Manipulation des fonctions des différents packages connectés de *Tidyverse***
  - 3.1 Présentation des cadres de données type *tibble*
  - 3.2 Transformation des données avec les fonctions de *dplyr*
  - 3.3 Importation et exportation des données avec les fonctions de *readr*
  - 3.4 Rangement des données avec les fonctions de *tidyr*
  - 3.5 Visualisation des données avec les fonctions du *ggplot2*
  - 3.6 Création de fonctions sous *Tidyverse*