

## Introduction - remise à niveau en langage R

*L'objectif de cette formation est de fournir les outils modernes de manipulation, de traitement de données et les connaissances nécessaires à la réalisation d'analyse en langage R. Un focus sera réalisé sur les usages modernes de R et sur l'usage de l'IDE Rstudio. A l'issue de cette formation l'apprenant sera autonome sur le logiciel R. Il sera capable d'importer, manipuler et traiter des données dans le « tidyverse » et de concevoir des analyses reproductibles. Un focus sera également réalisé sur les bonnes pratiques d'usage de R (travail en notebook, notion de « reproducible research »). Cette formation sera aussi l'occasion de présenter (rapidement) les autres fonctionnalités de R et Rstudio.*

**Durée:** 17.50 heures (2.50 jours)

### Profils des stagiaires

- Utilisateurs d'Excel pour l'analyse de données
- Utilisateurs de logiciels de statistiques propriétaires
- Utilisateur de R nécessitant une remise à niveau

### Prérequis

- Affinité pour la programmation
- Affinité pour les données
- Aisance informatique

### Objectifs pédagogiques

- Connaître et comprendre l'environnement R
- Importer et exporter des données
- Concevoir des opérations de manipulation de données (filtrer, trier, agréger, transformer...) comme un flux de traitement organisé
- Dessiner le graphique approprié à la représentation souhaitée
- Maîtriser le paramétrage de rapports mis en forme d'analyse de données

### Contenu de la formation

- Présentation de R, premiers pas dans la console
  - Démystifier la console, une «super-calculatrice» qui exécute du code
  - Prendre ses marques dans RStudio, créer son premier projet
  - Comprendre ce qu'est un vecteur
  - Se familiariser avec les packages et le moyen de les utiliser
  - Utiliser Rmarkdown pour la prise de notes
- Visualiser des données
  - Choisir le graphique adapté
  - Construire, modifier et exporter des graphiques avec le package {ggplot2}
- Importer et exporter des données
  - Importer des fichiers plats de type .csv et .xls/.xlsx avec les packages {readr} et {readxl}
  - Vérifier la qualité de l'import
  - Typage des données (déterminer si chaque colonne est une date, une valeur numérique, un texte...)
  - Exporter les données vers les formats .csv, .xlsx et .RDS

- Manipuler et nettoyer des données
  - Se familiariser avec les packages {dplyr} et {tidyr}
  - Concevoir les opérations de manipulation de données comme un flux de traitement organisé
  - Chainer les opérations avec l'opérateur %>%
  - Manipuler, transformer et résumer des données avec {dplyr}
  - Remodeler les données sous un format adéquat avec {tidyr}
- Produire des rapports d'analyse
  - Appréhender le concept de recherche reproductible
  - Compiler des rapports commentés imbriqués de code R avec Rmarkdown.
  - Exporter les rapports en docx, en html et en pdf
  - Paramétrer finement les rapports d'analyse réalisés avec Rmarkdown
- Bonnes pratiques
  - Trouver de l'aide face à un problème rencontré sur R
  - Concevoir un exemple reproductible
  - Respecter les conventions de style

## Organisation de la formation

### Equipe pédagogique

Une équipe de passionnés de R, formateurs, experts et contributeurs open-source

### Moyens pédagogiques et techniques

- Formation en distanciel synchrone (5 demi-journées)
- Mise à disposition en ligne de documents supports à la suite de la formation
- Plateforme d'apprentissage intégrée en ligne
- Exposés théoriques
- Travaux dirigés
- Mise à disposition d'une base de ressources documentaires

### Dispositif de suivi de l'exécution de l'évaluation des résultats de la formation

- Emargement en ligne
- Évaluations formatives régulières sous forme de QCM
- Évaluations sommatives régulières sous forme de travaux dirigés
- Formulaires d'évaluation de la formation, à chaud et à froid
- Alternance régulière entre théorie et pratique