

Formation **Logiciel R**

Analyse Statistique & Data Science

Formation progressive combinant apprentissage du langage, manipulation des données, analyses statistiques et visualisation avancée.

Programme de la formation



Module 1 – Prise en Main & Manipulation des Données

- Installation, environnement RStudio, scripts, projets
- Importation de données : readr, readxl, haven
- Manipulation avec dplyr : select, filter, mutate, summarise, group_by

Module 2 – Nettoyage et Préparation des Données

- Gestion des valeurs manquantes, doublons et outliers
- Recodage des variables, transformations numériques et catégorielles
- Reshape de données (long vs wide) avec tidyr

Module 3 – Statistiques Descriptives & Visualisation

- Mesures centrales et dispersion (mean, sd, median, quantiles)
- Tableaux de fréquences et statistiques groupées
- Visualisations avec ggplot2 (barres, boxplots, densités, corrélation)

Module 4 – Tests Statistiques & Modèle

- Tests paramétriques : t-test, ANOVA, corrélation de Pearson
- Tests non-paramétriques : Mann-Whitney, Wilcoxon, Khi-2
- Modèles de régression : linéaire, multiple, logistique

Module 5 : Exportation et Reporting

- Export de graphiques et tableaux (png, csv, xlsx)
- R Markdown : création d'un rapport automatique (PDF/Word/HTML)
- Reproductibilité et traçabilité du code