Logiciel R

La formation « Logiciel R » vous permettra de comprendre les bases du langage R et manipuler des données avec le tidyverse, de réaliser des analyses statistiques et des visualisations claires, et de produire des documents reproductibles avec RMarkdown.. Le programme est donné à titre indicatif et sera adapté à vos besoins et votre niveau après audit. N'hésitez pas à nous contacter pour toute demande spécifique.

Pré-requis

Des connaissances de l'environnement Windows et des connaissances de base en statistiques sont souhaitées. Toutefois, tous les éléments nécessaires sont donnés au fur et à mesure de la formation.

Public concerné

Développeur en environnement statistique

Durée et tarif de la formation

La durée de la formation varie en fonction des besoins et des objectifs déterminés après audit. Les tarifs sont disponibles sur devis.

Contenu de la formation

Introduction à R et RStudio

Présentation de R, RStudio et de l'écosystème R.

Philosophie du langage R vs Python vs Excel.

Installation de packages : install.packages(), library().

Présentation de RProjects pour une organisation de travail propre.

Bases du langage R

Types de données : numériques, chaînes, logiques.

Structures de base : vecteurs, matrices, listes, data frames, facteurs.

Opérateurs de base : arithmétiques, logiques, assignation.

Fonctions de base : length(), class(), typeof(), str(), etc

Premières manipulations de données

Import de données avec readr::read_csv(), readxl::read_excel().

Aperçu des données : head(), glimpse(), summary().

Accès aux éléments d’un data frame

Gestion des valeurs manquantes : is.na(), na.omit(), naniar.

Le tidyverse : manipulation moderne de données

Introduction au tidyverse et à la syntaxe %>% (ou |>).

Filtrer (filter()), sélectionner (select()), créer/modifier (mutate()), trier (arrange())

Regrouper et résumer (group_by(), summarise()).

Restructurer les données avec pivot_longer() et pivot_wider() (tidyr).

Jointures avec left_join(), inner_join() (dplyr).

Contrôle de flux et programmation de base

Structures conditionnelles : if, else.

Boucles : for, while, fonctions apply() et purrr::map().

Écriture de fonctions personnalisées : function(), arguments, retour de valeurs.

Statistiques de base

Statistiques descriptives : mean(), median(), var(), sd(), quantile().

Tests statistiques de base : test de normalité (shapiro.test()), t de Student (t.test()), test de variance (var.test()).

Visualisation de données

Introduction à ggplot2.

Créer des graphiques de base : barres, histogrammes, boxplots, courbes.

Esthétique et thèmes : labs(), theme_minimal(), scale_*.

Sauvegarde de graphiques.

Analyse en Composantes Principales (ACP)

Rappels théoriques sur l’ACP.

ACP avec FactoMineR::PCA() et factoextra::fviz_pca_*().

Visualisation avancée et interprétation des axes.