Formation R pour la Data Analyse avancée

Objectifs : Perfectionner son utilisation du langage R dans l'optique de structurer, automatiser et restituer efficacement des analyses de données reproductibles.

Compétences visées : - Structurer un projet R reproductible en organisant scripts, données, gestion d'environnement et bonnes pratiques de codage
- Automatiser des traitements analytiques en concevant des fonctions R robustes
- Mettre en œuvre des traitements de données avancés avec le tidyverse (pivots, traitements multicolonnes, gestion de facteurs)
- Analyser et optimiser les performances de traitements de données volumineuses à l'aide de {data.table} et {duckDB}
- Produire des visualisations et tableaux de synthèse professionnels, statiques ou interactifs, adaptés à la restitution analytique
- Générer, automatiser et mettre en forme des rapports d'analyse reproductibles avec Quarto

Durée : 3 jour(s) (21 heures)

Public : Toute personne utilisant déjà le langage R pour la réalisation d'analyses de données, de reporting ou d'études statistiques.

Pré-requis : Une pratique minimale de R et des packages du tidyverse (dplyr, ggplot2) est requise.

Tarifs :
- Présentiel : 1950 € HT
- Distanciel : 1800 € HT
(-10% pour 2 inscrits, -20% dès 3 inscrits)

Option(s) :
- Forfait déjeuners : 75 € HT

Nos prochaines sessions

Distance
du 23 au 25 novembre 2026
du 12 au 14 mai 2027

Lyon
du 12 au 14 octobre 2026
du 7 au 9 juin 2027

Paris
du 21 au 23 septembre 2026
du 3 au 5 mai 2027

Toulouse
du 14 au 16 décembre 2026

du 21 au 23 juin 2027

Programme :

- Bonnes pratiques générales avec R

- Projets R Studio : principes, organisation et chemins relatifs
- L'approche tidyverse : rappels des fondamentaux, différences avec le R base et principaux packages
- Bonnes pratiques de codage : conventions et outils d'aide
- Structurer son code et différencier les traitements préparatoires de l'analyse
- Gestion d'environnement avec renv

Pratique : passer en revue un script R « fonctionnel mais perfectible » et l'améliorer en mettant en place des bonnes pratiques de codage et d'organisation

- Automatisation et fonctions

- Rappels sur la création et l'appel de fonctions
- Intégration de messages et warning dans une fonction
- Approche pratique du tidyeval
- Bonnes pratiques dans l'écriture des fonctions (nommage, organisation, imbrication, débuggage)

Pratique : refactoriser un script R en encapsulant dans des fonctions diverses étapes du traitement de données

- Traitements de données avancés avec le tidyverse

- Rappels sur les principales fonctions et packages du tidyverse
- Outils transverses au tidyverse : {tidyselect}, {glue}, regex/li>
- Traitements sur de multiples colonnes (across et dérivés)
- Modifier la structure d'un dataframe avec les pivots
- Gestion des facteurs : package {forcats} et bonnes pratiques

Pratique : réaliser divers traitements analytiques courants impliquant notamment gestion de facteurs, pivots de données et traitements sur de multiples colonnes d'un dataframe

- Gestion de données massives

- Evaluer les temps de traitements de manipulations de données
- Fondamentaux du package {data.table}

- Traitements de données courants avec {data.table}
- {data.table} vs {dplyr} : avantages et inconvénients
- Connexion à {duckDB} pour des traitements accélérés

Pratique : mettre en place des traitements analytiques à partir d'une base de données conséquente, en utilisant différentes approches et en comparant leurs performances

- Dataviz & tableaux

- Aller plus loin avec {ggplot2} : thèmes, combinaisons de graphiques, gestion des couleurs, formatage de texte
- Création de visualisations interactives avec {ggiraph}
- Création et personnalisation de tableaux de synthèse avec {gt}
- Autres packages pour la dataviz et les tableaux

Pratique : créer des graphiques interactifs et des tableaux de synthèse, avec un rendu propre et professionnel

- Reporting avec Quarto

- Fondamentaux de Quarto : objectifs, principes, différences avec R markdown
- Options de code des chunks
- Personnalisation du YAML
- Intégration de tableaux et graphiques dans un rapport Quarto
- Trucs et astuces Quarto
- Rapports paramétrés

Pratique : automatiser la création de plusieurs rapports d'analyse avec un rendu propre et professionnel

Modalités d'évaluation : Un formulaire d'auto-évaluation proposé en amont de la formation nous permettra d'évaluer votre niveau et de recueillir vos attentes. Ce même formulaire soumis en aval de la formation fournira une appréciation de votre progression.
Des exercices pratiques seront proposés à la fin de chaque séquence pédagogique pour l'évaluation des acquis.
En fin de formation, vous serez amené(e) à renseigner un questionnaire d'évaluation à chaud.
Une attestation de formation vous sera adressée à l'issue de la session.
Trois mois après votre formation, vous recevrez par email un formulaire d'évaluation à froid sur l'utilisation des acquis de la formation.

Qualification du formateur : Les formations DATA VALUE sont animées par des intervenants experts, alliant expérience opérationnelle et compétences pédagogiques reconnues. Ils disposent de plus de 10 ans d'expérience professionnelle et assurent une veille constante sur les pratiques, outils et normes de leur domaine afin de garantir des contenus actualisés et adaptés aux besoins des stagiaires.

Délais d'accès et modalités de formation : Les inscriptions peuvent être effectuées jusqu'à 48 heures avant le début de la formation.
Vous pourrez trouver des informations plus détaillées concernant les modalités de nos formations à cette adresse : https://www.datavalue.fr/modalites-acces-formation.

Accessibilité : Vous souhaitez suivre notre formation Formation par ville et êtes en situation de handicap ? Merci de nous contacter afin que nous puissions envisager les adaptations nécessaires et vous garantir de bonnes conditions d'apprentissage

Date de dernière modification : 25 août 2026