

Formation dplyr, tidyr et lubridate - Manipuler efficacement des tableaux de données

Objectifs : Exploiter les nombreuses fonctionnalités des packages dplyr, tidyr et lubridate pour manipuler efficacement des tableaux de données dans R

Compétences visées : - Apprendre à créer des subsets (par sélection de colonnes et ou de lignes) avec le package dplyr

- Apprendre à assembler deux tables de données avec le package dplyr

- Apprendre à pivoter un tableau de données, du format wide vers le format long, et inversement avec le package tidyr

- S'initier à la gestion des données de dates et d'heures avec le package lubridate

Durée : 1 jour(s) (7 heures)

Public : Toute personne amenée à travailler avec des données, à l'aide du logiciel R, et souhaitant les manipuler aisément et efficacement

Pré-requis : Pour suivre ce stage dans de bonnes conditions, il est recommandé d'avoir suivi en amont la formation [R - Prise en main, analyses statistiques et graphiques](#)

Méthode pédagogique : Pédagogie active mêlant exposés, exercices et applications pratiques dans le logiciel R.

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.

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

Tarifs :

- Présentiel : 550 € HT

- Distanciel : 500 € HT

(-10% pour 2 inscrits, -20% dès 3 inscrits)

Option(s) :

- Forfait déjeuners : 20 € HT

Nos prochaines sessions

Distance

le 18 avril 2025

le 13 octobre 2025

Lyon

le 12 mai 2025

le 7 novembre 2025

Paris

le 6 juin 2025

le 8 décembre 2025

Toulouse

le 10 mars 2025

le 8 septembre 2025

Programme :

- Manipulation de données avec le package dplyr

- Introduction au package tidyverse et à la notion de pipe
- Sélection des colonnes (variable) avec la fonction select()
- Sélection de lignes avec la fonction filter()
- Création de nouvelles variables avec la fonction mutate()
- Création de tables résumées avec les fonctions group_by() et summarise()

- Combinaison de deux tableaux de données

- Fonctions d'assemblage de lignes
- Notion de clé primaire
- Fonctions de fusion de tables

- Transformation du format de tables avec le package tidyr

- Format wide et format long
- Passage d'un format wide à un format long
- Passage d'un format long à un format wide

- Gestion des données de dates et d'heure avec le package lubridate

- Spécificité des données de date et d'heure et norme
- Les fonctions de parsing du package lubridate
- Les fonctions d'extraction du package lubridate

Date de dernière modification : 5 novembre 2024