

Utilisation de R pour l'analyse de données

Prise en main de R et R Studio (3 jours – 21 h)

A qui s'adresse cette formation ?

Cette formation s'adresse aux personnes souhaitant **découvrir R et RStudio**, comprendre les bases du langage et acquérir de bonnes pratiques pour analyser des données et structurer efficacement leur travail sous R.

Elle est particulièrement adaptée aux :

professionnel(le)s de tous domaines amené(e)s à analyser des données dans le cadre de leur activité

étudiant(e)s, doctorant(e)s ou personnels de l'enseignement supérieur souhaitant utiliser R dans leurs travaux

personnes débutantes sur R, ou ayant déjà tenté d'utiliser R sans réussir à être autonomes

personnes souhaitant mettre en place un environnement de travail structuré et reproductible sous RStudio

Cette formation constitue une **première étape** pour prendre en main R et RStudio.

Elle peut ensuite être complétée par une formation de niveau avancé, notamment *R pour la data analyse*, ou par des **ateliers thématiques d'une journée** selon les besoins.

Ce que cette formation n'est pas

Cette formation n'a pas vocation à :

- enseigner l'ensemble du langage R en 3 jours
- aborder des développements avancés (programmation orientée objet, optimisation de performances, développement de packages)
- transformer les participant(e)s en développeurs ou développeuses R à l'issue de la formation

Elle vise au contraire à :

- poser des bases solides pour utiliser R et RStudio de manière efficace
- acquérir de bons réflexes pour structurer son travail et ses analyses
- développer une approche rigoureuse et progressive de l'analyse de données avec R

Prérequis

Seules **des connaissances de base de l'outil informatique** sont nécessaires.

Les participant(e)s doivent notamment être à l'aise avec :

- la création et l'organisation de dossiers
- la navigation dans l'arborescence de fichiers
- l'utilisation du copier-coller
- l'usage courant des outils bureautiques (type pack Office)

Aucune connaissance préalable du langage R ni de la programmation n'est requise.

La formation est conçue pour permettre une prise en main progressive de R et RStudio, en partant des bases.

Objectifs et compétences visées

Cette formation vise à permettre aux participant(e)s de **prendre en main R et RStudio**, de comprendre les bases du langage et d'acquérir les bonnes pratiques nécessaires pour analyser des données de manière autonome, structurée et reproductible.

À l'issue de la formation, les participant(e)s seront capables de :

utiliser l'interface de RStudio et organiser efficacement leur travail sous forme de projets R

comprendre les principes de base du langage R et manipuler les principales structures de données (vecteurs, data frames)

importer des données tabulaires (CSV, Excel), effectuer des vérifications simples et exporter des résultats

réaliser des premières manipulations de données (sélection, filtrage, combinaison de tables)

produire des analyses descriptives à l'aide d'indicateurs statistiques et de graphiques

réaliser des visualisations simples pour explorer, décrire et synthétiser des données
s'initier aux bases de la programmation avec R (fonctions apply, boucles, premières fonctions)

acquérir de bons réflexes pour écrire, lire et réutiliser du code R de manière fiable

Cette formation constitue une **base solide** pour poursuivre vers des formations plus spécialisées, notamment *R pour la data analyse*, ou vers des ateliers thématiques d'une journée.

Programme de la formation

Module 1 – Introduction à R et RStudio via un projet R

Objectif : Prendre en main R et RStudio en s'appuyant sur un projet R et comprendre l'organisation de l'environnement de travail.

- Présentation de R et RStudio
- Création et organisation d'un projet R
- Architecture du projet R
- Découverte des fenêtres et outils de RStudio
- Installation de packages et utilisation de l'aide

Module 2 – Les bases du langage R

Objectif : Comprendre les bases du langage R et manipuler les principales structures de données.

- Éléments à connaître pour bien démarrer
- Principaux types de données et principales structures (vecteurs, data frames)
- Opérateurs et fonctions de base

Module 3 – Importation, vérification et exportation de données

Objectif : Importer, vérifier et exporter des données tabulaires courantes de manière fiable.

- Importation de données tabulaires (.csv, .xlsx)
- Vérifications et modifications simples
- Exportations

Module 4 – Premières manipulations de données avec tidyverse / dplyr

Objectif : Réaliser des manipulations simples de tables de données à l'aide du tidyverse.

- Introduction au méta-package tidyverse
- Sélection de colonnes et filtrage de lignes
- Combinaison de tables par jointures

Module 5 – Analyses descriptives des données

Objectif : Décrire des données à l'aide d'indicateurs statistiques et de tables de synthèse.

- Indicateurs statistiques descriptifs
- Tables de synthèse

Module 6 – Réalisation de graphiques assistés avec {esquisse}

Objectif : Produire des graphiques exploratoires et descriptifs à l'aide d'une interface graphique.

- Graphiques à des fins exploratoires, descriptives et de synthèse
- Utilisation d'une interface graphique avec le package esquisse

Module 7 – Fonctions R pour les approches statistiques de base

Objectif : Mettre en œuvre les principales approches statistiques usuelles sous R en ligne de commande.

- Tests d'hypothèse
- Analyse de la variance
- Corrélation et régression

Module 8 – Initiation à la programmation avec R

Objectif : Automatiser des tâches simples grâce aux bases de la programmation avec R.

- Fonctions apply
- Boucles for
- Premières fonctions personnelles


Module 9 – Utiliser l'intelligence artificielle pour travailler avec R

Objectif : Utiliser l'intelligence artificielle pour générer, comprendre et adapter du code R de manière raisonnée.

- Obtenir du code R à partir d'un besoin
- Se faire expliquer un script existant
- Corriger et adapter du code proposé

Modalités pratiques

 **Durée** : 3 jours (21 heures)

 **Format** : distanciel (visioconférence), présentiel possible selon les besoins


 **Effectif** : groupes à taille limitée (environ 5 participant(e)s)

 **Pédagogie et organisation** :

Alternance d'apports théoriques et de mises en pratique

Supports remis aux participant(e)s. Temps dédiés aux questions et à l'interprétation

 **Logiciels** : R et R Studio

 **Évaluation des acquis** : Exercices pratiques, mises en situation et échanges oraux tout au long de la formation.

 **Tarif**

1 500 € HT par participant(e) pour la formation de 3 jours (21 h) - Supports pédagogiques inclus.

Les tarifs en intra-entreprise sont établis sur devis, en fonction du format et des modalités de la formation.

Organisme exonéré de TVA en vertu de l'article 261-4-4°a du CGI.

✓ **Suivi et assiduité** : Émargement par demi-journée.

 **Adaptation et accessibilité** :

Les exemples peuvent être adaptés au contexte des participants en intra-entreprise.

Toute situation de handicap ou besoin spécifique peut être signalé en amont afin d'étudier les adaptations possibles

Informations de contact

Organisme de formation : **DellaDATA**

NDA : 93040100204 auprès du préfet de région PACA

SIRET : 90838795400013

Organisme exonéré de TVA en vertu de l'article 261-4-4°a du CGI

Adresse : 1087 CH de Ste Roustagne - 04100 Manosque

Formatrice et interlocutrice : Claire Della Vedova (claire@delladata.fr / 06.64.98.21.42)