

Acquérir les fondamentaux de l'analyse de données avec le langage R

Date et durée
Code formation : DEV023FR Durée : 2 jours Nombre d'heures : 14 heures
Description
<p>Le langage R est un outil incontournable dans le monde de la recherche et de la science des données. Il permet de réaliser des analyses statistiques robustes et reproductibles. Cette formation de 2 jours est conçue pour vous rendre autonome sur cet outil, en partant des bases de la syntaxe jusqu'à l'automatisation de vos rapports. Le parcours privilégie une approche pratique pour démysterifier la programmation en R. Vous apprendrez à configurer votre environnement avec RStudio, à structurer et nettoyer vos jeux de données, puis à les exploiter graphiquement grâce au package de référence ggplot2. Loin de la théorie aride, vous manipulerez des données concrètes pour produire des résultats exploitables immédiatement. La formation se conclut sur les bonnes pratiques d'automatisation. Vous découvrirez comment transformer des tâches répétitives en scripts efficaces et comment générer des documents dynamiques avec RMarkdown, garantissant ainsi la traçabilité de vos analyses.</p>
Objectifs
<p>À l'issue de cette formation, vous atteindrez les objectifs de compétences suivants :</p> <ul style="list-style-type: none">• utiliser l'environnement R et RStudio pour exécuter des scripts et gérer des projets ;• manipuler et préparer les structures de données (vecteurs, data frames, listes) ;• réaliser des visualisations de données professionnelles avec la librairie ggplot2 ;• appliquer les fonctions statistiques fondamentales (tests, corrélations, régressions) ;• automatiser la production de rapports d'analyse via RMarkdown.
Points forts
<ul style="list-style-type: none">• Standards du marché : vous apprenez directement sur RStudio et ggplot2, les outils les plus utilisés par la communauté R.• Approche reproductible : vous intégrez dès le début la logique de script et de reporting automatisé (RMarkdown), essentielle pour la traçabilité.• Pratique métier : les ateliers s'appuient sur des cas réalistes (données de vente, de santé) pour une transposition immédiate dans votre quotidien.• Montée en compétence rapide : en deux jours, vous passez de l'installation à la production d'un rapport complet.
Modalités d'évaluation
Travaux Pratiques
Pré-requis

Suivre cette formation nécessite les prérequis suivants :

- **Connaissance de base :** une aisance avec des outils de développement informatique et des notions de base en statistiques sont recommandées pour profiter pleinement des exercices.

Public

Cette formation s'adresse aux profils souhaitant professionnaliser leur traitement de données. Le public inclut notamment :

- **les data analysts et chargés d'études** qui souhaitent dépasser les limites d'Excel pour industrialiser leurs analyses ;
- **les chercheurs, doctorants et étudiants** (sciences, économie, santé) ayant besoin d'un outil standard pour leurs publications scientifiques ;
- **les développeurs** en reconversion vers la Data, désireux d'ajouter une compétence statistique à leur profil technique ;
- **Toute personne voulant automatiser le traitement, la visualisation et le reporting de ses données.**

Programme

Jour 1 : Maîtriser la syntaxe et la préparation des données

- L'installation et la prise en main de l'interface RStudio.
- La gestion des structures de données : vecteurs, matrices et Data frames.
- L'importation de fichiers (CSV, Excel) et le nettoyage des données (gestion des valeurs manquantes, filtrage).

Travaux pratiques

- Découvrir RStudio, créer un jeu de données et nettoyer un fichier réel (ex. données de santé).

Jour 2 : Produire des analyses visuelles et automatisées

- La création de graphiques avancés avec la grammaire ggplot2 (barres, courbes, nuages de points).
- L'application des statistiques descriptives et des tests simples (T-test, Chi²).
- L'automatisation des traitements : création de fonctions, boucles et génération de rapports RMarkdown.

Travaux pratiques

- Concevoir un tableau de bord graphique, mener une analyse statistique complète et générer le rapport final automatisé.

Les noms de marques et logos éventuellement cités dans cette fiche (ex. R, RStudio, ggplot2, RMarkdown) sont la propriété de leurs détenteurs respectifs. Leur mention à des fins pédagogiques ne constitue ni un engagement ni un partenariat.