



Big Data - Data Science

Les fondamentaux de l'analyse statistique avec R

2 jours (14h00) | ★★★★★ 4,6/5 | BIG-STAT | Évaluation qualitative de fin de stage |
Formation délivrée en présentiel ou distanciel

Formations Informatique > Data > Big Data - Data Science

Contenu mis à jour le 18/10/2024. Document téléchargé le 08/12/2024.

Objectifs de formation

À l'issue de cette formation, vous serez capable de :

- Installer et utiliser l'environnement d'analyse R
- Manipuler des données avec R
- Importer et exporter des données
- Reconnaître les différents types d'objets de R
- Créer des programmes d'analyse avec R
- Réaliser des analyses statistiques basiques avec R
- Restituer des résultats à l'aide de graphiques.

Modalités, méthodes et moyens pédagogiques

Formation délivrée en présentiel ou distanciel* (blended-learning, e-learning, classe virtuelle, présentiel à distance).

Le formateur alterne entre méthode** démonstrative, interrogative et active (via des travaux pratiques et/ou des mises en situation).

Variables suivant les formations, les moyens pédagogiques mis en oeuvre sont :

- Ordinateurs Mac ou PC (sauf pour certains cours de l'offre Management), connexion internet fibre, tableau blanc ou paperboard, vidéoprojecteur ou écran tactile interactif (pour le distanciel)
- Environnements de formation installés sur les postes de travail ou en ligne
- Supports de cours et exercices

En cas de formation intra sur site externe à M2i, le client s'assure et s'engage également à avoir toutes les ressources matérielles pédagogiques nécessaires (équipements informatiques...) au bon déroulement de l'action de formation visée conformément aux prérequis indiqués dans le programme de formation communiqué.

* nous consulter pour la faisabilité en distanciel

** ratio variable selon le cours suivi

Prérequis

Etre familier avec l'environnement Microsoft Windows et avoir des connaissances de base en statistiques.

Public concerné

Ingénieurs, Data Analysts, statisticiens, développeurs en environnement statistique et/ou toute personne intéressée par l'analyse statistique avec R.

Cette formation :

- Est animée par un consultant-formateur dont les compétences techniques, professionnelles et pédagogiques ont été validées par des diplômes et/ou testées et approuvées par l'éditeur et/ou par M2i Formation
- Bénéficie d'un suivi de son exécution par une feuille de présence émargée par demi-journée par les stagiaires et le formateur.

Programme

Jour 1

Introduction

- Présentation du logiciel R et de ses fonctionnalités
- Avantages et inconvénients

Exemples de travaux pratiques (à titre indicatif)

- Accès au site de téléchargement de l'outil et installation

Premiers pas

- Utilisation de la console
- Création et sauvegarde d'un script
- Les autres éditeurs Tinn-R et RStudio
- Accès à l'aide en ligne et aide sur le CRAN (Comprehensive R Archive Network)

Exemples de travaux pratiques (à titre indicatif)

- Installation de RStudio, exécution d'ordres simples et sauvegarde de session

Objets et notions de programmation R

- Création, affichage et suppression d'un objet
- Les types de base
 - Logique
 - Numérique
 - Caractères
 - Complexe
 - Date
- Les objets de type vecteur, matrice, facteur, liste, DataFrame
- Manipulation des objets, classe d'objet, fonctions spécifiques, concaténation
- Fonctions prédéfinies
- Création et utilisation de fonctions
- Structures de contrôle

- Boucles (for, while, repeat)
- Conditions (if, else)
- Switch

Exemple de travaux pratiques (à titre indicatif)

- Programmer une fonction factorielle $f(n) = n \times (n - 1) \times \dots \times 2 \times 1$

Jour 2

Manipuler les données

- Séquences régulières et aléatoires

Exemples de travaux pratiques (à titre indicatif)

- Simuler une loi normale par addition de lois uniformes
- Données exemples de R
- Importation et exportation de données
- Modifier le type et les données d'un objet
- Import d'un jeu de données qualitatives et création d'un tableau de contingence, test d'indépendance du χ^2

Représenter les données

- Les fonctions graphiques conventionnelles : plot, barplot, hist, lines
- Amélioration et personnalisation des graphiques

Exemples de travaux pratiques (à titre indicatif)

- Représentations graphiques des données du TP précédent
- Représentation graphique d'une loi normale et de sa fonction de répartition

Packages R

- Installation, désinstallation, mise à jour d'un package
- Utilisation d'un package
- Exemples de packages populaires : dplyr, data.table, ggplot2, FactoMineR

Exemples de travaux pratiques (à titre indicatif)

- Analyse factorielle avec FactoMineR et représentation graphique du résultat

Le contenu de ce programme peut faire l'objet d'adaptation selon les niveaux, prérequis et besoins des apprenants.

Modalités d'évaluation des acquis

- En cours de formation, par des études de cas ou des travaux pratiques
- Et, en fin de formation, par un questionnaire d'auto-évaluation

Accessibilité de la formation

Le groupe M2i s'engage pour faciliter l'accessibilité de ses formations. Toutes nos formations sont accessibles aux personnes en situation de handicap : les détails de l'accueil des personnes sont consultables sur la page Accueil PSH.

Modalités et délais d'accès à la formation

Les formations M2i sont disponibles selon les modalités proposées sur la page programme.
Les inscriptions sont possibles jusqu'à 48 heures ouvrées avant le début de la formation.
Dans le cas d'une formation financée par le CPF, ce délai est porté à 11 jours ouvrés.