Les fondamentaux de l'analyse statistique avec R

Première mise en oeuvre pratique

Présentiel ou en classe à distance

3 jours (21 h)
+ activité à distance

Prix inter : 2.390,00 € HT
Forfait intra : 7.650,00 € HT

Réf.: BI103

Considéré depuis quelques années comme le logiciel le plus complet et puissant pour l'analyse statistique avancée de données, R dispose aujourd'hui d'une large communauté d'utilisateurs. Fidèles à la philosophie Open Source, ils ont peu à peu créé et mis à disposition des extensions réutilisables qui permettent par exemple de se connecter aisément à des bases de données ou encore qui proposent des modèles d'analyses spécifiquement adaptés à des domaines particuliers (ingénierie, finance quantitative, sociologie, économie, marketing, ...). Cette formation de premier niveau sur R permettra aux participants de créer leurs premières analyses statistiques et d'apprendre à les restituer sous forme graphique.

A qui s'adresse cette formation ?

- Ingénieurs
- Analystes
- Data analysts
- Toute personne intéressée par l'analyse statistique avec R

- Avoir suivi la formation "Les fondamentaux de la statistique appliquée" (BI090) ou connaissances mathématiques équivalentes
- Être familier avec l'environnement Microsoft Windows
- Disposez-vous des connaissances nécessaires pour suivre cette formation ? Testez-vous !

Programme

Pendant

1 - Introduction

- Qu'est-ce que R ?
- Avantages et inconvénients
- Solutions concurrentes gratuites ou payantes

2 - Installation

- Installation de R ou Microsoft R Open sur MS Windows ou Scientific Linux
- Découverte de l'environnement
- Utiliser l'historique des commandes
- Exemple d'environnement superposé (R-Studio)
- Comment citer le logiciel dans une publication scientifique
- Comment citer des packages dans une publication scientifique

3 - Utilisation

- Vider la console de commande
- Utilisation de l'aide
- Changer la langue de l'interface

- Quitter en ligne de commande
- Changer le dossier de travail par défaut temporairement ou à chaque session (*.Rprofile)
- Changer définitivement le dossier par défaut des packages (*.Rprofile)
- Sauver/Charger l'espace de travail (*.Rdata)
- Sauver/Charger/Exécuter un historique des commandes (*.Rhistory)
- Sauver les commandes et sorties dans un fichier (*.txt)
- Sauver/Charger un script (*.R)

4 - Manipulation de packages

- Installer/Désinstaller/Mettre à jour des packages
- Obtenir des informations systèmes sur les packages
- Écrire plusieurs commandes sur une ligne
- Ajouter des commentaires

5 - Types de données

- Manipulations de scalaires (réels)
- Manipulations de nombres complexes
- Manipulations de variables
- Manipulations de vecteurs
- Manipulations de matrices
- Manipulations de textes
- Manipulations de dates et de durées
- Création/Édition de données en ligne de commande
- Gérer les variables dans la mémoire

6 - Import et export de données

- Importer/Exporter des données d'Excel
- Importer/Exporter des données en *.csv et gestion du passage d'encodage Linux/Windows
- Importer/Exporter des fichiers SAS
- Importer des données *.csv du web
- Fusionner des fichiers *.csv
- Importer de données de MS Access
- Importer des données *.xml du web
- Importer des données de MySQL
- Importer des données d'Oracle (Express)
- Importer des fichiers *.json

7 - Manipulation de données

- Utiliser le SQL
- Quelques Data set
- Manipuler les data frames
- Accélérer l'accès aux colonnes
- Renommer les colonnes
- Appliquer des rangs
- Trier des données
- Filtrer des données
- Réaliser des sous-sélections
- Fusionner des données
- Supprimer les doublons
- Échantillonnage
- Empiler/Désempiler des données

8 - Analyse de données

- Synthétiser des données (tables de contingence)
- Travailler avec des valeurs absentes
- Définir le nombre de décimales de chaque sortie
- Générer des variables pseudo-aléatoires
- Statistiques descriptives simples (comptage, éléments uniques, moyenne, max, min, centiles, somme, écart-type biaisé/non biaisé, cv, médiane, etc.)
- Plotter (tracer) des fonctions algébriques
- Racines d'équations univariées
- Intégration numérique
- Intégration algébrique
- Dérivation algébrique/numérique
- Optimisation linéaire uni ou multidimensionnelle
- Optimisation sous contrainte
- Programmation linéaire

Après

- Des vidéocasts pour revenir sur les points clés de la formation



Les objectifs de la formation

- Savoir installer R
- Comprendre comment manipuler des données avec R
- Savoir importer et exporter des données
- Être en mesure de réaliser des analyses statistiques basiques avec R
- Savoir restituer des résultats à l'aide de graphiques



Evaluation

Pendant la formation, le formateur évalue la progression pédagogique des participants via des QCM, des mises en situation et des travaux pratiques. Les participants passent un test de positionnement avant et après la formation pour valider leurs compétences acquises.



Les points forts de la formation

- Une formation très pratique : 90% du temps de la formation est dédié à la mise en pratique pour une meilleure assimilation de notions de base.
- Des consultants expérimentés partagent leur savoir-faire avec les participants.
- 95% des participants à cette formation se sont déclarés satisfaits ou très satisfaits au cours des 12 derniers mois.

Evaluation

Pendant la formation, le formateur évalue la progression pédagogique des participants via des QCM, des mises en situation et des travaux pratiques. Les participants passent un test de positionnement avant et après la formation pour valider leurs compétences acquises.

Dates et villes 2027 - Référence BI103

Dernières places disponibles Session garantie

A distance

du 8 févr. au 10 févr.	du 6 sept. au 8 sept.
du 3 mai au 5 mai	du 2 nov. au 4 nov.

Paris

du 8 févr. au 10 févr.	du 6 sept. au 8 sept.
du 3 mai au 5 mai	du 2 nov. au 4 nov.