

LES FONDAMENTAUX DE L'ANALYSE STATISTIQUE AVEC R

PROGRAMME DE FORMATION DIG16



PUBLIC CONCERNÉ

Professionnels du marketing, RH, finance, santé, logistique, ou R&D souhaitant exploiter des données.



PRÉ-REQUIS

Connaître les bases d'Excel ou d'un tableur (tri, filtres, formules simples).



DURÉE

21 heures



TARIF INTER

1 260,00 €



TARIF INTRA

Nous consulter



MOYENS TECHNIQUES & PÉDAGOGIQUES

Support de cours, salle et équipement informatique mis à disposition. Plateforme de visioconférence pour les cours à distance. Evaluation des besoins en amont de la formation et adaptation du programme aux attentes. Alternance d'exercices pratiques corrigés, de questionnaires.



QUALIFICATION DES INTERVENANTS

Formateurs disposant d'une d'expérience minimum de 5 ans sur la thématique enseignée.



MODALITÉ D'ÉVALUATION

Remise d'une attestation individuelle en fin de bilan. Cette attestation valide le niveau de progression et les acquis.



ACCESSIBILITÉ HANDICAP

Formation ouverte aux personnes en situation de handicap. Contactez-nous pour étudier les adaptations nécessaires ou les solutions alternatives.

Délai minimum d'accès à la formation de 15 jours à partir de confirmation du financement.



OBJECTIFS

- 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



PROGRAMME

MODULE 1 - INTRODUCTION

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

MODULE 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
- Comment citer des packages dans une publication scientifique

MODULE 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)

MODULE 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

MODULE 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

MODULE 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

atkGROUP
CONSTRUIRE DEMAIN

01 40 54 50 37

113



PROGRAMME (SUITE)

MODULE 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/Déempiler des données

MODULE 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

MODULE 9 - SYNTHÈSE

- Lister les principales bonnes pratiques à retenir

Evaluation : Un quizz de 40 questions est à renseigner par les participants afin d'évaluer le niveau des connaissances acquises (15 minutes). Les résultats seront fournis en séance.