

**DURÉE : 2 A 3 JOURS**

7 heures par jour - Présentiel

**PRÉ-REQUIS**

Connaissances de base en statistiques (centrage et dispersion, tests statistiques, corrélation) et des concepts de programmation

**OBJECTIFS**

Installer l'environnement d'analyse R  
Utiliser la console de R, et les environnements de travail Tinn-R et R Studio  
Créer et regrouper plusieurs types d'objets de R  
Créer des programmes d'analyse avec R  
Élaborer des graphiques avec R  
Utiliser les packages de R pour mettre en œuvre des modélisations statistiques (régression, classification)

**PUBLIC VISÉ**

Utilisateurs finaux, data scientists, statisticiens, analystes type Data Miner, acteurs impliqués dans l'analyse/fouille des données

**MOYENS PÉDAGOGIQUES**

- Diagnostic pédagogique amont permettant de regrouper les apprenants par niveau homogène et d'assurer la parfaite adéquation entre vos besoins et le programme
- Organisation en petit groupe de 1 à 6 apprenants maximum garantissant une meilleure interactivité
- 1 poste informatique par apprenant
- Environnement confortable, calme et propice à la concentration (dans nos salles)
- Supports de cours et documentation individuels de qualité (livrets papier)
- Enchaînement de : ① théorie  
② démonstration par l'exemple  
③ mise en pratique sur exercices
- Visualisation et partage des connaissances transmises par projection audiovisuelle en appui
- A l'issue de la formation, tous les exercices et leurs corrigés vous sont remis

**MOYENS D'ENCADREMENT**

Consultant formateur spécialisé, validé par notre équipe tant sur la capacité pédagogique que la connaissance technique métier

**MOYENS D'ÉVALUATION**

Diagnostic préalable des connaissances individuelles à partir d'un questionnaire de positionnement  
Évaluation de l'atteinte des objectifs par l'apprenant  
Évaluation du transfert des acquis par le formateur

**ORGANISATION - INTER OU INTRA****INTER-ENTREPRISES**

Prix et dates sur calendrier si programmé

**PRIX INTRA-ENTREPRISE**

Sur devis - Programme adaptable en intra

**CONTENU PEDAGOGIQUE****Introduction**

Présentation du logiciel R et de ses fonctionnalités  
Avantages et inconvénients  
Accès au site de téléchargement de l'outil et installation

**Premiers pas**

Environnement de base (console, script)  
Utilisation de la console  
Création et sauvegarde d'un script  
Aide et commentaires  
L'éditeur R Studio

**Objets et notions de programmation R**

Les objets de type vecteurs, matrix, array, factor, data.frame, list  
Manipulation des objets, classe d'objet, fonctions spécifiques, jointure  
Sauvegarde, suppression de mémoire  
Notions de boucle (for et while), condition (if), switch

**Création et utilisation de fonctions**

Structure d'une fonction  
Fonctions de type mathématique  
Fonctions de type chaîne de caractères  
Fonctions liées au temps/date  
Les tables de contingences

**Génération, gestion et visualisation des données**

Les données : séquences régulières et aléatoires  
Données exemples de R  
Importation et exportation de données  
Modifier les données d'un objet  
Exemples de graphiques construits avec R  
Création des graphiques de base  
Les options graphiques, partager une fenêtre graphique, sauvegarder un graphique

**Analyses statistiques**

Présentation de la notion de package (librairie)  
Télécharger/charger des packages  
Quelques packages utiles  
Cas de la régression linéaire multiple  
Cas de la régression logistique  
Cas de l'analyse en composantes principales ACP  
Cas de la classification CAH  
Cas du clustering KMeans