

Durée : 3 jours soit 21 heures

Référence : IF-LANG-R



Public visé :

- Utilisateurs finaux, statisticiens, analystes type **Data Miner**, acteurs impliqués dans l'analyse/fouille des données.



Pré-requis :

- Connaissances de base en statistiques (régressions, analyse en composantes principales, classification) et des concepts de programmation.



Objectifs pédagogiques :

- Installer l'environnement d'analyse **R**
- Utiliser la console de **R**, et les environnements de travail **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**, **ACP**...)



Compétences acquises à l'issue de la formation :

- Installer l'environnement d'analyse **R**
- Utiliser la console de **R**, et les environnements de travail **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**, **ACP**...)



Modalités pédagogiques :

Session dispensée en présentiel ou téléprésentiel, selon la modalité inter-entreprises ou intra-entreprises sur mesure. La formation est animée par un(e) formateur(trice) durant toute la durée de la session et présentant une suite de modules théoriques clôturés par des ateliers pratiques validant l'acquisition des connaissances. Les ateliers peuvent être accompagnés de Quizz. L'animateur(trice) présente la partie théorique à l'aide de support de présentation, d'animation réalisée sur un environnement de démonstration. En présentiel comme en téléprésentiel, l'animateur(trice) accompagne les participants durant la réalisation des ateliers.

Note relative au programme : Les ateliers, démonstrations, travaux pratiques et scénarios présentés dans ce programme sont fournis à titre d'exemples illustratifs et ne sont pas exhaustifs. Afin de garantir une adéquation permanente avec les évolutions technologiques, le contenu pédagogique ainsi que les environnements de travaux pratiques (TP) sont susceptibles d'être adaptés, ajustés ou mis à jour au cours du temps.



Moyens et supports pédagogiques :

Cadre présentiel

Salles de formation équipées et accessibles aux personnes à mobilité réduite.

- Un poste de travail par participant
- Un support de cours numérique ou papier (au choix)
- Un bloc-notes + stylo
- Vidéoprojection
- Connexion Internet
- Accès extranet pour partage de documents et émargement électronique

Cadre téléprésentiel

Session dispensée via notre solution iClassRoom s'appuyant sur Microsoft Teams.



- Un compte Office 365 par participant
- Un poste virtuel par participant
- Un support numérique (PDF ou Web)
- Accès extranet pour partage de documents et émargement électronique



Modalités d'évaluation et de suivi :

Avant

Afin de valider le choix d'un programme de formation, une évaluation des prérequis est réalisée à l'aide d'un questionnaire en ligne ou lors d'un échange avec le formateur(trice), qui valide la base de connaissances nécessaires.

Les stagiaires disposent également, depuis leur extranet, d'un **outil d'auto-positionnement** leur permettant d'évaluer leur niveau de maîtrise des compétences visées par la formation. Cette auto-évaluation est réalisée avant le démarrage de la formation à l'aide d'un curseur de positionnement allant du niveau le plus faible au niveau le plus élevé.

Pendant

Après chaque module théorique, un ou plusieurs ateliers pratiques permettent de valider progressivement l'acquisition des connaissances. Un quiz peut accompagner les ateliers pratiques afin de mesurer la compréhension des notions abordées.

Après

À l'issue de la formation, les stagiaires réalisent une nouvelle **auto-évaluation** sur les mêmes compétences que celles évaluées avant la formation. Cette comparaison permet de mesurer leur perception de la progression réalisée et de visualiser les compétences développées au cours de la formation.

Lorsque le programme de formation le prévoit, un examen de certification est organisé conformément aux modalités définies par l'éditeur ou le centre de certification (TOSA, Pearson VUE, ENI, PeopleCert...).

Le formateur réalise également une **évaluation individuelle des compétences** de chaque stagiaire sur l'ensemble des compétences visées par la formation. Pour chacune d'elles, il indique si la compétence est :

- Acquis ;
- En cours d'acquisition ;
- Non acquise.

Lorsque la compétence est évaluée comme **en cours d'acquisition** ou **non acquise**, le formateur formule des observations précisant les difficultés rencontrées ainsi que les actions ou recommandations permettant au stagiaire de poursuivre sa montée en compétences.

Enfin

Un questionnaire de satisfaction permet au participant d'évaluer la qualité de la prestation de formation.



Informations sur l'accessibilité :

Nos formations sont, dans la mesure du possible, conçues pour être accessibles à toutes et à tous. Afin de garantir les meilleures conditions d'accueil et d'apprentissage pour les personnes en situation de handicap, nous vous invitons à contacter notre référente handicap certifiée :

Céline SOLATGES – 05 61 34 39 80 – csolatges@iform.fr

Nous vous remercions de bien vouloir nous communiquer toute information utile à ce sujet en amont de la formation, afin de mettre en place les adaptations nécessaires et d'assurer un accompagnement optimal.

Pour en savoir plus sur les dispositifs d'accompagnement existants, vous pouvez consulter les sites suivants :

- [AGEFIPH](#)
- [FIPHFP](#)
- MDPH de votre département

Description / Contenu

Formation Environnement R, traitement de données et analyse statistique.

R est un **langage de programmation** dont le but est de pouvoir traiter et organiser des jeux de données en y

appliquant des tests **statistiques** plus ou moins complexes. R permet de se représenter ces données traitées graphiquement à l'aide d'une grande variété de graphique disponible.

Développé initialement au début des années 1990 par Ross Ihaka et Robert Gentleman, ce langage basé sur un autre **langage statistique appelé S** (oui, les informaticiens ont beaucoup d'imagination), est désormais maintenu et mis à jour par une équipe de développeurs au sein du **R Project**. Cette structure garantit des mises à jour fréquentes et une communauté importante d'utilisateurs apporte son aide pour notamment développer de nouvelles fonctionnalités (des librairies, nous en parlerons plus tard) au projet.

Le langage R est en effet très implanté dans la communauté scientifique et universitaire et ses nombreux utilisateurs organisent même des conférences pour discuter des nouveautés disponibles sous **R**.

Module 1 : Introduction

- Présentation du logiciel R et de ses fonctionnalités.
- Avantages et inconvénients.
- Accès au site de téléchargement de l'outil et installation. Exercice Installation de l'environnement d'analyse.

Module 2 : Premiers pas

- Environnement de base (console, script).
- Utilisation de la console.
- Création et sauvegarde d'un script.
- Le répertoire sous R Installation.
- Aide et commentaires.
- Les autres éditeurs Tinn-R et R Studio.
- Exercice *Création de scripts*.

Module 3 : 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.
- Exercice *Manipulation des types d'objets*.

Module 4 : Création et utilisation de fonctions

- Structure d'une fonction.
- Fonctions de type mathématique.
- Fonctions de type chaîne de caractères.
- Fonctions liées au temps/date.
- Opérations ensemblistes.
- Les tables de contingences.
- Exercice *Création et utilisation de fonctions*.

Module 5 : 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.
- Exercice *Exercice d'application sur les données*.

Module 6 : Analyses statistiques

- Présentation de la notion de package (librairie).
- Télécharger/charger des packages.
- Quelques packages utiles.
- Cas de la régression non linéaire multiple.
- Cas de l'analyse en composantes principales ACP.
- Cas de la classification CAH.
- Exercice *Exercice d'application sur les packages*.