EduGroupe
Accompagner pour réussir

Les fondamentaux de l'analyse statistique avec R

Référence : **PYCB052B** Durée : **2 jours (14 heures)** Certification : **Aucune**

Connaissances préalables

- Avoir des connaissances de base en statistiques

Profil des stagiaires

- Ingénieurs, Data analysts, statisticiens, développeurs en environnement statistique ou toute personne intéressée par l'analyse statistique avec R

Objectifs

- Être capable de réaliser une analyse statistique avec le logiciel R
- En restituer les résultats sous forme graphique

Certification préparée

- Aucune

Méthodes pédagogiques

- 6 à 12 personnes maximum par cours, 1 poste de travail par stagiaire
- Remise d'une documentation pédagogique papier ou numérique pendant le stage
- La formation est constituée d'apports théoriques, d'exercices pratiques et de réflexions

Formateur

- Consultant-Formateur expert Bigdata

Méthodes d'évaluation des acquis

- Auto-évaluation des acquis par le stagiaire via un questionnaire
- Attestation des compétences acquises envoyée au stagiaire
- Attestation de fin de stage adressée avec la facture

Contenu du cours

1. Savoir installer et utiliser l'environnement d'analyse R

- Le projet R Programming. R Project
- Utilisation de R en mode commande
- Commandes de base. Syntaxe de base. Interfaces et environnements : RStudio, Jupyter
- Opérations de base. Expressions, variables, types de données, nombres, chaînes de caractères
- Exemple de Travaux Pratiques : installation et tests d'une plateforme R

2. Comprendre comment manipuler des données avec R

- Base de R : tests, boucles, fonctions
- Transformations de données, conversions de types
- Fonctions mathématiques de base
- Exemple de Travaux Pratiques : conversion de données numériques en texte. conversion de données texte en nombres

3. Savoir reconnaître les différents types d'objets de R

- Manipulations de nombres, vecteurs, tableaux, matrices, listes, factors
- Liste et DataFrames : Définitions, cas d'utilisation
- Attachement, détachement. Chargement d'un dataframe. La fonction scan
- Tableaux et matrices : Déclaration, dimensionnement, indexation
- Opérations de base : produit de tableaux, transposition, produits de matrices
- Matrices : équations linéaires, inversion, valeur propre, vecteur propre, déterminant, moindre carré, ...
- Exemple de Travaux Pratiques : exercices sur les objets R (matrice, vecteur)

4. Savoir importer et exporter des données

- Import/export : formats texte, csv, xml, binaire, largeur fixe, images (jpeg, png)
- Encodage. Filtrage. Bibliothèques : rjson, readr, xml2
- Interfaçage avec le BigData : SparkR
- Transformation d'un dataframe R en un dataframe Spark
- Importation SQL. Importation depuis un socket réseau
- Exemple de Travaux Pratiques : importation de données géodésiques et export au format Json

5. Créer des programmes d'analyse avec R

- Structure générale d'un programme d'analyse. Découpage en sections métiers. Notion d'ETL
- Fonctions spécifiques : définition de nouvelles fonctions. Appels. Passage d'argument
- Construction d'une bibliothèque
- Diffusion, installation avec R CMD INSTALL
- Exemple de Travaux Pratiques : construction d'un programme de calcul de moyennes en trois parties : chargement, calcul, restitution

6. Être en mesure de réaliser des analyses statistiques basiques avec R

- Statistiques : Distributions embarquées : uniforme, normale, poisson, exponentielle, ...
- Calculs statistiques. Modèles statistiques
- Exemple de Travuax Pratiques : analyse statistique d'une population

7. Savoir restituer des résultats à l'aide de graphiques

- Affichage en graphes, histogrammes : Plot, Line, Pie Charts, Scatterplot, Bars
- Exemple de Travaux Pratiques : affichage des données de l'atelier précédent sous forme de nuage de points

Notre référent handicap se tient à votre disposition au 01.71.19.70.30 ou par mail à referent.handicap@edugroupe.com pour recueillir vos éventuels besoins d'aménagements, afin de vous offrir la meilleure expérience possible.