



Les fondamentaux de l'analyse statistique avec R

CB052

Durée: 2 jours

1 670 €

30 au 31 janvier
5 au 6 mai

25 au 26 août
4 au 5 décembre

Public :

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

Objectifs :

A l'issue de la formation, le stagiaire sera capable de réaliser une analyse statistique avec le logiciel R et d'en restituer les résultats sous forme graphique.

Connaissances préalables nécessaires :

avoir des connaissances de base en statistiques

Objectifs pédagogiques :

Savoir installer et utiliser l'environnement d'analyse R
Comprendre comment manipuler des données avec R
Savoir importer et exporter des données
Savoir reconnaître les différents types d'objets de R
Créer des programmes d'analyse avec R
Être en mesure de réaliser des analyses statistiques basiques avec R
Savoir restituer des résultats à l'aide de graphiques

Programme :

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,

Atelier : installation et tests d'une plateforme R

Atelier : travail sur les chaînes de caractères et sur les types scalaires.

Comprendre comment manipuler des données avec R

Base de R : tests, boucles, fonctions.
Transformations de données, conversions de types.
Fonctions mathématiques de base.

Atelier : conversion de données numériques en texte. conversion de données texte en nombres



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, ...

Atelier : exercices sur les objets R : matrice, vecteur

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.

Atelier : importation de données géodésiques et export au format Json

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.

Atelier : construction d'un programme de calcul de moyennes en trois parties : chargement, calcul, restitution

Être en mesure de réaliser des analyses statistiques basiques avec R

Statistiques : Distributions embarquées : uniforme, normale, poisson, exponentielle, ...
Calculs statistiques. Modèles statistiques.

Atelier : analyse statistique d'une population

Savoir restituer des résultats à l'aide de graphiques

Affichage en graphes, histogrammes : Plot, Line, Pie Charts, Scatterplot, Bars.

Atelier : affichage des données de l'atelier précédent sous forme de nuage de points.