R, TRAITEMENT DE DONNÉES ET ANALYSE STATISTIQUE

Formation proposée par BIZOÏ RAZVAN

R – Analyse statistique

R est un environnement logiciel Open Source spécialisé dans le calcul et l'analyse statistique. Ce stage vous présentera ses concepts et ses fonctionnalités. Il vous montrera comment manipuler les données et leur appliquer des modèles statistiques dans le cadre de cet environnement.

◷DURÉE
5 Jour(s)
▣MODALITÉS
Présentiel / Distanciel
♙PARTICIPANTS
Jusqu'à 10 participants
★LANGUE
Français
LIEU
France entière
♙CERTIFICATION
Attestation de formation
▦MISE À JOUR
12/04/2023

VOTRE FORMATEUR

BR **BIZOÏ RAZVAN**
Formateur professionnel

TARIFS INDICATIFS

Inter-entreprise	**Pas de dates planifiées**
Intra-entreprise	**Sur devis**
Individuelle	**Sur devis**
Durée	**5 Jour(s)**

Tarifs indicatifs, sous réserve d'analyse du besoin et de devis.

◎Objectifs pédagogiques

- ✓Installer l'environnement d'analyse R
- ✓Créer et regrouper plusieurs types d'objets de R
- ✓Créer des programmes d'analyse avec R
- ✓Elaborer des graphiques avec R
- ✓Utiliser les packages de R pour mettre en œuvre des modélisations statistiques (régression, ACP...)

⌘Compétences visées

R – Analyse statistique

Public cible et prérequis

Public cible

- Statisticiens, analystes, ingénieurs, développeurs.
- Toute personne ayant à manipuler des données, à traiter numériquement et représenter graphiquement des données, à réaliser des statistiques sous R.

Prérequis

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

Programme détaillé

Introduction
- Présentation du logiciel R et de ses fonctionnalités.
- Création et sauvegarde d'un script.
- Le répertoire sous R Installation.
- Les autres éditeurs R Studio et Jupyter Notebook.

Travaux pratiques
Installation de l'environnement R et R-Studio et Jupyter Notebook.

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, condition, switch.

Travaux pratiques
Manipulation des types d'objets.

Création et utilisation de fonctions
- Structure d'une fonction.
- Opérations ensemblistes.
- Les tables de contingences.

Travaux pratiques
Création et utilisation de fonctions.

Importation-exportation et production de données
- Lire un fichier texte ASCII, Excel, SPSS, Minitab, SAS ou Matlab.
- Lire des données au clavier et utiliser le copier-coller.
- Lecture/écriture des fichiers, bases de données.

Travaux pratiques
Lecture/Ecriture des données à partir des fichiers, base de données, DataLake et dans le format R.

Manipulation de données, fonctions
- Opérations sur les matrices ou les DataFrames.
- Les fonctions outer, apply, lapply, sapply et mapply.
- Opérations logiques et relationnelles.
- Manipulation de chaînes de caractères. Manipulation de dates et d'unités de temps.

Travaux pratiques
Traîter les matrices et DataFrames. Utiliser les fonctions lapply ou sapply pour remplacer les boucles for.

La manipulation de données via les packages du tidyverse
- Import de données avec readr.
- Manipulation de données avec dplyr, tidyr, tibble.
- Manipulation des variables catégorielles avec forcats.
- Manipulation des dates avec lubridate.
- Manipulation des chaînes de caractères avec stringr.

Travaux pratiques
Mise en œuvre des classifications supervisées sur plusieurs jeux de données.

Techniques pour tracer des courbes et des graphiques
- Les fenêtres graphiques : manipulation, sauvegarde.
- Les fonctions de tracé de bas niveau.
- La gestion des couleurs et paramètres graphiques.

- L'ajout de texte, titres, axes et légendes.
- Diagrammes en croix, tuyaux d'orgue, empilé ou circulaire. Boîte à moustaches.
- Graphe de la fonction de répartition empirique. Histogramme en densité à amplitudes de classes égales ou inégales.
- Polygone des fréquences. Représentations graphiques dans un cadre bivarié.

Travaux pratiques

Mise en œuvre des techniques pour tracer des courbes et des graphiques.

Mathématiques et statistiques élémentaires
- Structuration des variables suivant leur type.
- Résumés numériques.
- Mesures d'association.
- Notions sur la génération de nombres au hasard et de variable aléatoire.
- Loi des grands nombres et théorème de la limite centrale.
- Intervalles de confiance.
- Tests d'hypothèses usuels. Autres tests d'hypothèses.
- Analyse de la variance à un facteur, deux facteurs ou à mesures répétées.

Travaux pratiques

Mise en œuvre de l'algorithme de détermination du type d'une variable. L'interprétation d'un intervalle de confiance.

▣Méthodes pédagogiques & évaluation

Méthodes pédagogiques

- ✓ Apports théoriques et retours d'expérience
- ✓ Exercices pratiques et mises en situation
- ✓ Travaux de groupe et échanges entre pairs
- ✓ Études de cas concrets

Modalités d'évaluation

- ✓ Quiz de validation des acquis
- ✓ Mises en situation évaluées
- ✓ Projet fil rouge
- ✓ Feuille de route post-formation

♿Accessibilité

Une formation accessible à tous

Ce programme peut être adapté pour les personnes en situation de handicap. Contactez-nous pour étudier vos besoins spécifiques.