

Environnement R, traitement de données et analyse statistique

Cours Pratique de 2 jours - 14h
Réf : TDA - Prix 2024 : 1 620€ HT

Environnement logiciel Open Source et langage, R est spécialisé dans les traitements statistiques. Ce cours vous apprendra à programmer en R, à utiliser l'éditeur R studio, à exploiter les possibilités de visualisation des datas et à appliquer des algorithmes statistiques fondamentaux : régressions, ACP, CAH.

OBJECTIFS PÉDAGOGIQUES

À l'issue de la formation l'apprenant sera en mesure de :

- Savoir installer et utiliser l'environnement d'analyse 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
- Comprendre comment manipuler des données avec R
- Savoir restituer des résultats à l'aide de graphiques

EXERCICE

Mise en pratique des connaissances théoriques sur la base de jeux de données variés.

LE PROGRAMME

dernière mise à jour : 11/2022

1) Introduction

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

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.
- Travaux pratiques : Manipulations de la console. Ecriture de scripts.*

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.

FINANCEMENT

Ce cours fait partie des actions collectives Atlas.

PARTICIPANTS

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

PRÉREQUIS

Être familier avec l'environnement Microsoft Windows et avoir des connaissances de base en statistiques.

COMPÉTENCES DU FORMATEUR

Les experts qui animent la formation sont des spécialistes des matières abordées. Ils ont été validés par nos équipes pédagogiques tant sur le plan des connaissances métiers que sur celui de la pédagogie, et ce pour chaque cours qu'ils enseignent. Ils ont au minimum cinq à dix années d'expérience dans leur domaine et occupent ou ont occupé des postes à responsabilité en entreprise.

MODALITÉS D'ÉVALUATION

Le formateur évalue la progression pédagogique du participant tout au long de la formation au moyen de QCM, mises en situation, travaux pratiques...

Le participant complète également un test de positionnement en amont et en aval pour valider les compétences acquises.

MOYENS PÉDAGOGIQUES ET TECHNIQUES

- Les moyens pédagogiques et les méthodes d'enseignement utilisés sont principalement : aides audiovisuelles, documentation et support de cours, exercices pratiques d'application et corrigés des exercices pour les stages pratiques, études de cas ou présentation de cas réels pour les séminaires de formation.
- À l'issue de chaque stage ou séminaire, ORSYS fournit aux participants un questionnaire d'évaluation du cours qui est ensuite analysé par nos équipes pédagogiques.
- Une feuille d'émargement par demi-journée de présence est fournie en fin de formation ainsi qu'une attestation de fin de formation si le stagiaire a bien assisté à la totalité de la session.

MODALITÉS ET DÉLAIS D'ACCÈS

L'inscription doit être finalisée 24 heures avant le début de la formation.

ACCESSIBILITÉ AUX PERSONNES HANDICAPÉES

Vous avez un besoin spécifique d'accessibilité ? Contactez Mme FOSSE, référente handicap, à l'adresse suivante psh-accueil@orsys.fr pour étudier au mieux votre demande et sa faisabilité.

- Notions de boucle (for et while), condition (if), switch.

Travaux pratiques : Ecriture de programmes R manipulant des types d'objets.

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.

Travaux pratiques : Création de fonctions et utilisation de celles-ci dans les programmes R.

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.

Travaux pratiques : Exercices d'application sur les données, restitution des résultats avec des graphiques.

6) Analyses statistiques

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

Travaux pratiques : Poursuite d'écriture de programmes statistiques, intégration des packages.

LES DATES

CLASSE À DISTANCE

2024 : 23 mai, 08 juil., 07 oct.

PARIS

2024 : 16 mai, 01 juil., 30 sept.

LYON

2024 : 23 mai, 08 juil., 07 oct.