

Formation Logiciel R, Perfectionner l'analyse de données

Présentation

Cette formation vous permettra d'acquérir des compétences en développement R de manière à pouvoir mener à bien des projets et élaborer des programmes efficaces et intuitifs.

Durée : 21,00 heures (3 jours)

Tarif INTRA : Nous consulter

Objectifs de la formation

- Savoir organiser son environnement de travail pour fluidifier son processus d'analyse
- Connaître les packages utilisés dans la plupart des projets professionnels
- Savoir utiliser les bonnes structures de données pour les manipuler efficacement
- Savoir visualiser de façon élégante et rapide les données
- Maîtriser la manipulation des données avec les package dplyr, forecast, stringr, lubridate, etc.
- Savoir assembler les données (jointure)
- Savoir représenter graphiquement les données avec le package ggplot2
- Savoir générer dynamiquement son rapport d'analyse avec rmarkdown
- Comprendre la programmation fonctionnelle (purrr)

Prérequis

- Avoir des connaissances de base sur le logiciel R

Public

- Pour les utilisateurs du logiciel R

Programme de la formation

Organiser son travail sous R



- Travailler en projet R : notion de working directory, workspace, history
- Architecture de son projet R : data, plots, images, scripts, etc..
- Bonnes pratiques pour la création de fichiers de données
- Importation et exportation de fichiers avec le package here
- Mettre à jour ses packages
- Mettre à jour R et R Studio

Manipuler facilement ses données avec le package dplyr

- Introduction au package tidyverse et à la notion de pipe
- Filtrer des lignes avec la fonction filter()
- Sélectionner des colonnes (variable) avec la fonction select()
- Création de nouvelles variables avec la fonction mutate()
- Renommer ses variables avec la fonction rename()
- Calcul de paramètres par sous-groupes : fonctions group_by() et summarise()
- Passage du format wide au format long

Manipulation des variables catégorielles avec le package forecats

- Inspecter les variables catégorielles avec les fonctions levels(), fct_count et fct_unique
- Modifier l'ordre des modalités
- Modifier le nom des modalités

Manipuler les chaînes de caractères avec le package stringr

- Détection de patterns
- Découpage
- Gestion des longueurs
- Remplacement

Manipuler des données de date : utilisation du package lubridate

- Convertir les données au format YYYY-MM-DD et HH:MM:SS
- Décomposer les éléments d'année, de mois et de jour
- Calculer des différences de dates et les exprimer en jours, ou heures

Assemblage de tables

- Les différentes jointures (par colonne): left join, right join, inner join et full join
- Assemblage par lignes
- Exemple d'applications pour l'analyse de données

Réaliser des représentations graphiques performantes avec le package ggplot2

- Le principe des couches successives de ggplot2
- Réalisation des graphiques de base : scatterplot, barplots, line plot, boxplots
- Gestion des couleurs, titres, axes et légendes
- Représentation des séries temporelles
- Utilisation du format long et facetting
- Ajouter du texte sur un graphique (ex : équation)
- Exporter son graphique : format et résolution
- Utilisation des addins esquisse et Colour Picker

Générer dynamiquement son rapport d'analyse avec R Markdown

- Principe, formats de sorties (html, docx, pdf)
- Les différents éléments d'un fichier Rmd : en-tête, chunk, etc...
- Gestion des éléments de texte : gras, italique, titre
- Gestion des tables
- Gestion des images
- Gestion des graphiques
- Gestion des options des éléments de code
- Gestion de la table des matières et numérotation
- Les rapports paramétrés : automatisation des rapports d'analyse par sous groupe

Introduction à la programmation fonctionnelle avec le package purrr

- Les list
- Les fonctions map()
- Nested data

Organisation

Formateur

Les formateurs de Docaposte Institute sont des experts de leur domaine, disposant d'une expérience terrain qu'ils enrichissent continuellement. Leurs connaissances techniques et pédagogiques sont rigoureusement validées en amont par nos référents internes.

Moyens pédagogiques et techniques

- Apports des connaissances communes.
- Mises en situation sur le thème de la formation et des cas concrets.
- Méthodologie d'apprentissage attractive, interactive et participative.
- Equilibre théorie / pratique : 60 % / 40 %.
- Supports de cours fournis au format papier et/ou numérique.
- Ressources documentaires en ligne et références mises à disposition par le formateur.
- Pour les formations en présentiel dans les locaux mis à disposition, les apprenants sont accueillis dans une salle de cours équipée d'un réseau Wi-Fi, d'un tableau blanc ou paperboard. Un ordinateur avec les logiciels appropriés est mis à disposition (le cas échéant).

Dispositif de suivi de l'exécution et de l'évaluation des résultats de la formation

En amont de la formation

- Recueil des besoins des apprenants afin de disposer des informations essentielles au bon déroulé de la formation (profil, niveau, attentes particulières...).
- Auto-positionnement des apprenants afin de mesurer le niveau de départ.

Tout au long de la formation

- Évaluation continue des acquis avec des questions orales, des exercices, des QCM, des cas pratiques ou mises en situation...

A la fin de la formation

- Auto-positionnement des apprenants afin de mesurer l'acquisition des compétences.
- Evaluation par le formateur des compétences acquises par les apprenants.
- Questionnaire de satisfaction à chaud afin de recueillir la satisfaction des apprenants à l'issue de la formation.
- Questionnaire de satisfaction à froid afin d'évaluer les apports ancrés de la formation et leurs mises en application au quotidien.

Accessibilité

Nos formations peuvent être adaptées à certaines conditions de handicap. Nous contacter pour toute information et demande spécifique.

