Dcid CONSULTING data & strategy

Conseil - Formation - Expertise en Data Science

" Investissez dans votre capital humain, facteur clé de compétitivité "

IMPORTATION ET MANIPULATION DE DONNÉES AVEC R

Le traitement de données est une étape incontournable de tout projet de data science. Grâce à ses librairies 'dplyr' et 'data.table' (notamment), R est un des langages de référence des data scientists pour le traitement de données et requiert des compétences méthodologiques et informatiques. Maitriser R fera de vous un acteur incontournable pour le traitement et la valorisation des données de votre entreprise.

Durée

- 1 jour

Tarif

- Inter : 550€ HT / participant (6 participants maximum)
- Intra : *nous consulter*

Lieu

- Locaux de Dcid Consulting (Orléans)
- Locaux de votre entreprise (Région Centre et Paris)

Public

Tout public ayant déjà quelques notions en programmation et/ou analyse de données

- Analyse de données ★☆☆☆☆
- Statistique ★☆☆☆☆
- Programmation ★★☆☆☆
- Base de données ★☆☆☆☆
- Curiosité ★★★★★

Renseignements

- contact@dcid-consulting.fr
- 02 38 25 33 90

Prérequis

- Ordinateur individuel
- Connexion internet

Ressources utiles

- R, le site officiel
- R Studio, le site officiel
- Dplyr dans le tidyverse

Objectifs pédagogiques

- Maitriser l'importation et la manipulation de données de sources variées (numérique, catégorielle, textuelle, temporelle)
- Savoir préparer un jeu de données en vue d'une utilisation ultérieure (dataviz, machine learning, analyses statistiques)
- Savoir extraire, filtrer et résumer un jeu de données
- Savoir visualiser et explorer un jeu données

Programme

- Découvrir R, son environnement et son écosystème
  - Présentation et chargement de R, des librairies principales
  - Présentation de R Studio et son écosystème
- Importer des données de sources variées
  - Chargement de fichiers de données
  - Connection et requête sur une base de données
  - Extraction de données depuis un site web
- Manipuler les données de sources variées
  - Données numériques, catégorielles, textuelles et temporelles
- Manipuler et préparer un jeu de données
  - Extraction d'un sous ensemble du jeu de données
  - Sélection, filtrage, jointure des données
  - Résumé statistique des données
  - Modification du format du jeu de données
- Visualiser et explorer un jeu des données
  - Représentation graphique simple
  - Présentation d'outils et méthodes pour l'exploration
- Et après ? Perspectives et cas d'usage

Méthodes et moyens pédagogiques

- Alternance d'exposés et d'applications pratiques
- Possibilité de travailler sur les jeux de données des participants
- Codes source et supports de formation remis aux participants
- Formation dispensée par un data scientist expérimenté (pratique de R depuis 7 ans minimum)