



R EXPERTISE

Objectifs pédagogiques :

- Savoir structurer un traitement complexe sous forme de fonctions
- Connaître les outils adaptés pour manipuler des données volumineuses
- Savoir mesurer la performance d'un code et l'améliorer
- Savoir développer et maintenir un package

Prérequis

- Appétence pour l'approche analytique et quantitative
- Connaissances de base en programmation (concepts de boucles, conditions ...)
- Connaissances des notions de statistiques, d'analyse de données, d'économétrie et de Machine Learning

Modalités pédagogiques :

Public :

- ▶ Data scientists avec connaissances en R couvrant la formation « R consolidation »

Déroulé de la formation :

- ▶ Durée : 10 demi-journées

Modalités d'évaluation des participants :

- ▶ L'évaluation se fera sous forme d'exercices avec fil directeur lors d'une séance dédiée.

Moyens spécifiques de la formation :

- ▶ Formateur avec une expérience significative en R et une formation de data scientist (> 2 ans)

Moyens pédagogiques :

- Support de présentation
- Exercices pratiques
- Bases de données pour faire les exercices
- ▶ Dispositif de suivi de la formation : feuille de présence

Panorama des séquences :

Temps 1 - Fonctions

Séquence 1 : Maîtriser les fonctions

- Environnements (notamment les variable locales / globales)
- Notions de programmation fonctionnelle
-





Temps 2 - Données volumineuses

Séquence 2 : Les outils pour traiter des données volumineuses (1/2)

- Principes
- Traitement hors mémoire
- Calcul distribué

Séquence 3 : Les outils pour traiter des données volumineuses (2/2)

Temps 3 - Optimisation de code

Séquence 4 : Reconnaître la qualité d'un code

- Principaux outils et bonnes pratiques
- Temps de code vs. temps de calcul

Séquence 5 : Améliorer son code

- Optimisation de l'utilisation de mémoire
- Optimisation du temps de calcul
- Debugging de code
- Manipulation des chaînes de caractères avec les expressions régulières
-

Temps 4 - Packages

Séquence 6 : Les bases d'un package

- Quand et pourquoi faire un package
- Dépendances et documentation (roxygen2)

Séquence 7 : Créer et maintenir un package de qualité

- Gestion des tests
- Publication sur le CRAN
- Récapitulatif, approfondissement et évaluation

Séquence 8 : Synthèse

- Zoom sur renv.lock
- Rappels et exercice de mise en pratique (données DVF)

Séquence 9 : Approfondissement - Utilisation avancée des fonctions

- Fonctions avec fonctions en argument ou en retour
- Fonctions emboîtées, récursives, etc.
- Décorateurs

Séquence 10 : Evaluation

- Exercice basé sur les données SIRENE

► **Contact :** formations@veltys.com

► **01 87 44 69 70**

