



FORMATION R INITIATION

La formation « R initiation » débute par une demi-journée d'acculturation aux nouveaux environnements de la data (logiciels, sources, algorithmes, nouvelles pratiques, etc.). Les objectifs de la formation R initiation sont ensuite de prendre en main et maîtriser l'environnement de travail de R avec Rstudio, d'apprendre à manipuler et gérer des données avec R et les packages de data manipulation les plus récents (dplyr, data.table...). Seront abordés les différences et les points communs entre SAS et R, traduire des codes SAS en codes R, connaître les packages de gestion de données avec R et leur équivalent SAS

Durée : 5 demi-journées

Objectifs pédagogiques :

- ▶ Être à l'aise avec l'environnement de R
- ▶ Maîtriser les principales fonctionnalités de R et ses objets
- ▶ Importer des données tabulaires de fichiers divers (texte, excel, SAS)
- ▶ Explorer le contenu d'une table via des statistiques descriptives et des visualisations graphiques
- ▶ Manipuler des données de sorte à créer de nouvelles variables, transformer des tables et associer plusieurs tables entre elles
- ▶ Sauvegarder son travail sous R ou d'autres formats plus universels (excel, jpg, txt...)

Modalités d'évaluation des participants

- ▶ L'évaluation se fondera à la fois sur un questionnaire de type coding-game ...
- ▶ ... et sur une application des compétences enseignées sous forme d'un projet.

Exemples de notions enseignées dans le module

- ▶ Organiser son travail sous R :
 - Importer et exporter des données
 - Installer et utiliser des bibliothèques
 - Organiser des scripts au sein d'un projet
 - Connaître les principaux objets et fonctionnalités (list, vector, array, matrix, data.frame, factor...)
 - Bonnes pratiques (noms de variables, formatage du code, commentaires, hard coding, reproductibilité...)
- ▶ Vérifier le contenu de données :
 - Inspecter les lignes et les colonnes d'une table
 - Décrire le contenu d'une variable
 - Afficher des distributions
 - Réaliser des tests
 - Savoir lire un log
- ▶ Associer plusieurs tables :
 - Empiler / concaténer / appairer des datasets
 - Transposer





- ▶ Identifier les passerelles et les différences avec SAS :
 - différentes façons de faire des sélections de variables et d'observations
 - transformations, renommages, recodage,
 - équivalents des traitements séquentiels, first, last, in, etc., agrégations

- ▶ Contact : formations@veltys.com