

# FORMATION AU LOGICIEL R

## 1. INITIATION AU LANGAGE

### Description

**Difficulté** : 1/5

**Durée** : 3 jours (21 heures)

**Public ciblé** : Toute personne souhaitant prendre en main le logiciel R, maîtriser les bases de la programmation et la mise en œuvre d'analyses statistiques de base.

**Formateur** : Olivier Delaigue (INRAE – U.R. HYCAR)

### Avant-propos

Cette formation s'adresse à des personnes souhaitant prendre en main le logiciel R : son interface, sa syntaxe, ses objets, afin que le stagiaire devienne autonome dans le traitement de ses données. Il ne s'agit donc pas d'une formation sur la méthodologie statistique, mais bien d'une formation sur le logiciel R. De même, il ne s'agit pas d'une formation sur les techniques de programmation avancée sous R.

### Prérequis

- Connaissance de l'environnement Windows
- Il est souhaitable que le stagiaire ait des connaissances sur les différentes analyses statistiques mises en œuvre au cours de cette formation

### Objectifs pédagogiques

À l'issue de cette formation, le stagiaire sera capable de :

- Installer et configurer R
- Installer un package
- Utiliser R en mode console
- Exécuter et exploiter des scripts simples
- Obtenir de l'aide en ligne, rechercher des commandes permettant de réaliser l'action souhaitée
- Créer et manipuler les objets sous R
- Importer des données contenues dans un fichier texte ou dans une feuille de calcul Excel
- Créer et modifier un graphique
- Mettre en œuvre des analyses statistiques de base sous R

### Objectifs opérationnels et contenu de la formation

#### Comprendre les principes généraux du langage R

- Bref historique
- Fonctionnement général de R
- Fonctionnalités, avantages, inconvénients
- Installation de R
- Site du projet R et documentation

### Utiliser un EDI

- Avantages et inconvénients
- Présentation de différents EDI
- Installation et prise en main de RStudio

### Se repérer dans l'interface du logiciel R et maîtriser les notions élémentaires

- Interface et gestion des fenêtres
- Utiliser la console R
- Organiser un projet R
- Saisir et rappeler des lignes de commandes
- Obtenir de l'aide en ligne
- Mode console, mode script
- Choisir le répertoire de travail
- Affecter des objets (créer des variables)
- Lister et effacer les objets en mémoire
- Sauver et charger l'environnement de travail et l'historique des commandes

### Créer et manipuler des objets sous R

- Attributs intrinsèques des objets
- Création et indexation des vecteurs (variables quantitatives)
- Opérations sur les vecteurs
- Extraction des éléments d'un vecteur
- Création et manipulation des facteurs (variables qualitatives)
- Création d'un tableau de données (`data.frame`)
- Extraction des éléments d'un tableau de données
- Extraction des éléments d'une sortie R

### Réaliser un calcul arithmétique et utiliser les fonctions simples

- Opérateurs arithmétiques (+, -, \*, /)
- Fonctions mathématiques de base sur les données (`log`, `exp`, `sum`, `mean`, etc.)

### Connaître et utiliser les opérateurs de comparaison

- Opérateurs de comparaison
- Extraction des éléments d'un vecteur par condition
- Filtrage des données

### Importer et exporter des données

- Contenues dans un fichier texte
- Contenues dans une feuille de calcul Excel
- Contenues dans le presse-papiers

### Créer et exécuter des scripts

- Lecture
- Écriture
- Exécution et débogage
- Ouverture sur la création de fonctions utilisateur simples



### Créer et manipuler des graphiques

- Les fonctions graphiques (`plot`, `hist`, `barplot`, `boxplot`, etc.)
- Gestion des graphiques
  - Ouverture de plusieurs fenêtres graphiques
  - Partitionnement des fenêtres graphiques
- Les fonctions graphiques secondaires (`points`, `lines`, `abline`, `title`, `legend`, etc.)
- Les commandes de mise en forme des graphiques
- Le paramétrage des graphiques

### Mettre en œuvre des analyses statistiques de base sous R

- Analyse descriptive des données
  - La description par le chiffre (`summary`, `sd`, `table`, etc.)
  - La description par le graphique
- Identification des valeurs suspectes
- Tests d'hypothèses paramétriques et non paramétriques
  - Comparaison des moyennes (Student)
  - Comparaison des médianes (Wilcoxon, Mann-Whitney)
  - Comparaison des variances (Fisher)
  - Comparaison des proportions ( $\chi^2$ , Fisher exact)
  - Test de Normalité (Shapiro-Wilk)
- Analyse de la variance
- Régression linéaire simple
- Analyses multivariées (A.C.P., A.F.C., etc.)

#### Moyens pédagogiques

##### Méthodes et moyens :

- Explications théoriques suivies de pratiques guidées, puis mise en autonomie
- 1 vidéoprojecteur par salle
- 1 ordinateur par stagiaire

##### Méthodes d'évaluation des acquis :

- Exercices d'applications

##### Support du stagiaire :

- Diaporama du cours
- Exercices et jeux de données d'accompagnement
- Corrigés des exercices

