# Annexe 2. LOGICIEL R - INITIATION

**Durée :** 3 jours

**Nombre de participants maximum :** 10

**Objectifs pédagogiques :**

* Exécuter et exploiter des scripts simples
* Créer et manipuler des objets sous R (vecteurs, facteurs, dataframe)
* Importer des données
* Créer et modifier un graphique simple
* Mettre en œuvre des analyses statistiques de base sous R

**Public cible :** Agents CNRS souhaitant prendre en main le logiciel R, s'initier aux bases de l’utilisation

et à la mise en œuvre d’analyses statistiques de bases.

**Pré-requis :** Avoir suivi la formation « Notions fondamentales de statistiques » ou avoir un niveau équivalent

**Programme :**

1. Introduction au logiciel R

* Configuration du logiciel R/R Studio
* Utiliser l’aide en ligne (recherche de commandes R)

1. Principes généraux du langage R, principales structures de données, objets et fonction

* Vecteurs
* Types de données
* Dataframe : structure de base des traitements statistiques
* Listes
* Packages

1. Principaux éléments de R Studio

* Object browser, packages, fenêtre de scripts
* Sauvegarder l’espace de travail, charger un fichier sauvegardé
* Travailler en mode projet
* Importer dans une dataframe
* Exporter un fichier CSV/Excel

1. Langage R : prise en mains des objets R

* Extraire des parties d’un vecteur
* Concaténer, supprimer, insérer
* Manipulation de différents types de données
* Regrouper/ réordonner les niveaux d’un facteur
* Extraire des parties d’une matrice
* Extraire des parties d’une liste
* Dataframe en tant que liste et en tant que matrice
* Gérer les données manquantes

1. Statistiques descriptives : principaux résumés statistiques sur des données numériques ou nominales

* Médiane, moyenne, quantiles
* Distribution statistique
* Corrélations et covariances
* Tableaux de contingence

1. Graphiques : visualiser des données multi-variées sous R

* Histogrammes
* Box-plots, dot plots, scatter plot
* Graphiques en barre
* Graphiques en mosaïques
* Fonction qplot de ggplot2

1. Exercice pratique : réalisation d’un mini-projet (analyse de données et production du rapport)

* Description de chaque variable
* Corrélation et corrélations partielles
* Représentation graphique multidimensionnelle