



# PROJET RISE

(recherche, innovation, surveillance et évaluation)

---

## TERMES DE REFERENCE

### Formation : R pour la régression linéaire et la régression logistique

**Durée de la formation** : 5 jours en présentiel

**Dates de la formation** : Du 17 au 21 octobre 2022

**Lieu** : Institut Pasteur de Madagascar

**Date limite de dépôt de candidature** : Le 11 septembre 2022

## 1 Introduction

Le logiciel R, associé à l'outil Rstudio, permet l'analyse de grands jeux de données épidémiologiques, en permettant de réaliser des tests statistiques classiques (chi-2 ou test de Student) mais aussi en allant plus loin et en utilisant des modèles de régression linéaire ou logistique. Apprendre à programmer en R permet, en quelques lignes de code, de charger un jeu de données, faire des représentations graphiques, et appliquer un modèle.

## 2 Objectif de la formation

La formation a pour objectif de renforcer les compétences des participants à l'utilisation du logiciel R et R studio pour analyser les données épidémiologiques et de santé publique.

## 3 Participants ciblés

Cette formation est dédiée aux professionnels de santé qui sont impliqués dans la collecte, l'analyse des données de santé publique et la dissémination des résultats d'études, travaillant principalement dans les différentes directions du ministère de la santé publique au niveau national, régional et district ainsi que des institutions et laboratoires de recherche à Madagascar.

Au total, 15 participants seront sélectionnés.

Prérequis :

- Avoir suivi une formation pour débutant en programmation R.
- Savoir utiliser Rstudio.
- Être familier du concept de p-value

## 4 Organisation de la formation

En plus des cours théoriques, les participants auront l'occasion de pratiquer les différentes notions abordées sur des bases de données réelles.

Cette formation sera dispensée par 2 experts de l'Institut Pasteur à Paris

- Yoann MADEC
- Vincent GUILLEMOT

## 5 Résultats attendus

A la fin de cette formation, les participants seront capables de :

- Gérer les données avec R (importation, exportation etc...)
- Effectuer une exploration basique des données (description statistique de base, représentation graphique)
- Choisir un modèle adapté et l'appliquer
- Interpréter les résultats

---

## 6 Contenu du programme et agenda

### **Jour 1 : 8h30-16h30**

- Rappel sur R et R-Studio
- Introduction aux tests statistiques

### **Jour 2 : 8h30-16h30**

- Analyse des variables continues
  - Test de Student
  - Tests non paramétriques
  - Analyse de la variance (ANOVA)

### **Jour 3 : 8h30-16h30**

- Analyse des variables continues – suite
  - Principes de la régression linéaire
  - Analyse univariée
  - Analyse multivariée

### **Jour 4 : 8h30-16h30**

- Analyser les variables catégorielles
  - Test du chi-2
  - Généralisation du test du chi-2
  - Test de Fisher

### **Jour 5 : 8h30-16h30**

- Aller plus loin avec les variables binaires
  - Principe de la régression logistique
  - Interaction