

## Formations Intensives de l'École HED

### Formations R 2020

**Formateur :** Joseph Larmarange, Chargé de Recherche au CEPED - Université de Paris

En mars et mai 2020, 3 sessions de formation R, du niveau Débutant au niveau avancé seront organisées dans le cadre des *Formations intensives de l'École des Hautes Études en Démographie* (EUR HED) (anciennement Résodemo) :

- Introduction à l'analyse d'enquêtes avec R et RStudio • niveau débutant, du **11 au 13 mars 2020** ;
- Manipulation de données et statistiques avancées avec R et RStudio, **25 au 27 mai 2020** ;
- Introduction à l'analyse de séquences avec R et RStudio, du **28 au 29 mai 2020**.

Les formations peuvent être suivies de manière indépendante. Soyez attentifs à la progression dans le niveau des formations proposées.

**Il ne s'agit pas de cours de statistiques :** les différents modules présupposent donc que vous avez déjà une connaissance des différentes techniques présentées.

#### Lieu de formation :

Institut National d'Études Démographiques,

Campus Condorcet

9 cours des Humanités,

93322 AUBERVILLIERS CEDEX, **Salle 1. 080 (pour les 3 formations)**

**Information :** Pour ces 3 ateliers, vous pouvez apporter votre ordinateur portable mais cela n'est pas indispensable, les formations étant dispensées comme des cours magistraux.

**NB :** *Les formations intensives de l'École HED* (ex *Résodemo*) peuvent délivrer des crédits ECTS. Une attestation de présence pourra vous être délivrée à l'issue de la formation ou sur demande.

Nous invitons chaque doctorant à se renseigner auprès de la direction de son école doctorale.

Merci de retrouver le descriptif de chacune des formations ci-dessous.

Information et inscriptions: [contact@resodemo.fr](mailto:contact@resodemo.fr)

## Introduction à l'analyse d'enquêtes avec R et RStudio • niveau débutant

11 au 13 mars 2020

### Objectif :

Présenter comment réaliser des analyses statistiques et diverses opérations courantes (comme la manipulation de données ou la production de graphiques) avec R.

Il ne s'agit pas d'un cours de statistiques : les différents modules présupposent donc que vous avez déjà une connaissance des différentes techniques présentées.

### Thématiques abordées :

- Prise en main de R et RStudio
- Import de données
- Gestion des données labellisées
- Manipulation des données et recodages
- Statistique descriptive uni- et bivariée (incluant intervalles de confiance et tests)
- Plan d'échantillonnage complexe et pondération
- Introduction à ggplot2 (graphiques)
- Régression logistique
- Sélection pas à pas d'un modèle
- Présentation d'autres modèles : modèle de poisson, régression multinomiale, modèles mixtes

### Prérequis :

- Pas de connaissance de R préalable
- Connaître les statistiques uni- et bivariées et la régression logistique
- Une connaissance d'un autre logiciel de statistique (tel que SPSS, SAS ou Stata) est préférable

### Acquis de la formation :

Être en capacité de réaliser sous R et RStudio des analyses statistiques de base.

## Manipulation de données et statistiques avancées avec R et RStudio • niveau avancé

25 au 27 mai 2020

### Objectif :

L'objectif de ce second atelier est de perfectionner son usage de R et RStudio en abordant des techniques statistiques plus avancées (analyse des correspondances, analyse multivariée...) qui nécessitent le plus souvent une manipulation plus complexe des données en amont. Sera également abordé la réalisation de rapports automatisés.

Il ne s'agit pas d'un cours de statistiques : les différents modules présupposent donc que vous avez déjà une connaissance des différentes techniques présentées.

### Thématiques abordées :

- Manipulation avancée des données : les packages data.table et dplyr
- Gestion des dates : le package lubridate
- Analyse des correspondances multiples et classification hiérarchique
- Modèles multivariés (approches stepwise, effets d'interaction, multicollinéarité...)
- Rapports automatisés avec knitr et rmarkdown

### Prérequis :

- Connaissance du logiciel R et de sa syntaxe (cf. stage d'introduction)

### Acquis de la formation :

- Connaître et utiliser des outils avancés de manipulation de données
- Mettre en œuvre sous R des analyses de correspondance
- Réaliser des analyses multivariées
- Générer des rapports automatisés en Rmarkdown

**Approches longitudinales avec R et RStudio • niveau avancé**

28 et 29 mai 2020

**Objectif :**

Introduction aux approches longitudinales dont analyse de survie, analyse de séquences et *optimal matching* dans le champ des sciences sociales et à son implémentation sous R.

**Thématiques abordées :**

- Analyse de survie (Kaplan-Meier et modèle de Cox)
- Introduction à l'analyse de séquences
- Introduction aux modèles temporels

**Prérequis :**

- Connaissance du logiciel R et de sa syntaxe (cf. stage d'introduction)
- Manipulation avancée de données et gestion des dates (cf. stage avancé)
- Savoir réaliser un modèle multivarié sur R et une classification hiérarchique (cf. stage avancé)

**Acquis de la formation :**

- Mettre en œuvre sous R des analyses longitudinales