

L'objectif de ce séminaire est d'échanger autour du logiciel de statistique libre, gratuit et multiplateforme *R* (<http://www.r-project.org/>). Il s'adresse aux praticiens impliqués dans le traitement quantitatif des données en sciences humaines et sociales (utilisateurs de données, chercheurs, ingénieurs ou étudiants) qu'ils aient ou non déjà utilisé le logiciel *R*.

Chaque séance est organisée autour de la présentation d'une expérience de traitement de données avec le logiciel (fonction spécifique et/ou packages). Le cadre de ces réunions est informel et les participants doivent se sentir libres d'intervenir afin de confronter leurs expériences.

Élisabeth Morand et Bénédicte Garnier (Ined),
Timothée Giraud (CNRS UMS Riate), Pascal Cristofoli (EHESS)
<https://enseignements-2017.ehess.fr/2017/ue/1010/>

Séance 2 : jeudi 1^{er} février 2018 (9h30-12h)

Présentation du tidyverse

Julien Barnier

(Centre Max Weber, ENS de Lyon)

Le tidyverse est un ensemble de packages pour *R* qui abordent les étapes les plus courantes d'un projet d'analyse de données : importation (*readr*, *readxl*, *haven*), manipulation et recodage (*dplyr*, *tidyr*, *forcats*), sorties graphiques (*ggplot2*), etc. Ces packages sont conçus pour fonctionner ensemble et partagent une philosophie commune, notamment le souci d'une syntaxe cohérente et un modèle d'organisation des données (baptisé tidy data).

L'objectif de cette présentation est de faire découvrir les principes du tidyverse, de donner un aperçu des principaux packages qui le composent, de leurs avantages mais aussi de certains inconvénients, et de permettre à chacun de juger s'il lui semble souhaitable d'apprendre ou d'enseigner l'utilisation de ces extensions.

Prochaines séances

5 avril

7 juin