



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),
Florence Maillachon (CMH-Cnrs), Pascal Cristofoli (EHESS)
<http://www.ehess.fr/fr/enseignement/enseignements/2014/ue/562/>

Séance 2 : jeudi 19 mars 2015 de 9h à 11h

Claire Tissot (INRS)

« **Text mining et recherche de thèmes dans un ensemble de documents : une application de R au traitement de données d'accidents de salariés nouvellement embauchés dans une entreprise** »

Nous aborderons la mise en œuvre complète de l'analyse, de la lecture des données aux traitements et sorties graphiques. Quatre packages seront utilisés : tm pour la création d'un corpus et le traitement du texte, topicmodels pour l'identification de thèmes et d'une structure intra-document, FactoMineR pour les analyses factorielles et une classification des accidents, worcloud pour un exemple d'illustration graphique.

Le nombre de places étant limité à 30, l'inscription est indispensable à : <http://www.ined.fr/fr/actualites/rencontres-scientifiques/seminaires-colloques-ined/russ-seance-2/>

Prochaine séance jeudi 21 mai 2015