



## Programme des séances 2014-2015

### Séance 1: Présentation générale de R - Générations de rapports dynamiques

**Judi 18 décembre 2014 de 9h à 11h**, salle 015, RdC, bât. Le France, 190 av de France 75013 Paris

**Christophe Genolini** (Université Paris Ouest Nanterre La Défense), *Présentation générale de R - Générations de rapports dynamiques* A la fois langage de programmation et environnement d'analyse, R s'est progressivement imposé sur le devant de la scène des statistiques. Les raisons de ce succès sont diverses: gratuité, simplicité de la prise en main, mais également son côté open source qui permet à un chercheur inventant une nouvelle méthode de la programmer et de la mettre lui-même à la disposition de la communauté. On trouve sur R nombre des méthodes les plus récentes dans des domaines aussi variés que la biologie, la psychologie, l'économie, la médecine, l'histoire, la linguistique et bien sûr les sciences sociales.

Lors de ce séminaire, nous essayerons dans un premier temps de donner une vue d'ensemble du logiciel. Puis nous nous attarderons un peu plus précisément sur l'«automatic reporting», la possibilité d'intégrer automatiquement une analyse statistique dans un rapport.

### Séance 2: Text mining et recherche de thèmes dans un ensemble de documents

**Judi 19 Mars 2015 de 9h00 à 11h00** à l'EHESS en salle 015, RdC, bât. Le France, 190 av de France 75013 Paris.

**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.*

Au cours de la séance sera abordé 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.

### Séance 3: Représentation et traitement de l'information géographique avec R

**Judi 21 mai 2015 de 9h00 à 11h00** à l'EHESS en salle 015, RdC, bât. Le France, 190 av de France 75013 Paris.

**Timothée Giraud et Nicolas Lambert** (Ingénieurs, CNRS UMS Riate), *Représentation et traitement de l'information géographique avec R*

R permet de simplifier les chaînes de traitements d'analyse spatiale et de cartographie qui nécessitaient auparavant l'utilisation de nombreux logiciels. A partir d'exemples seront abordés notamment l'utilisation d'API de cartographie, des opérations classiques de géotraitement ainsi que des traitements plus poussés d'analyse spatiale faisant appel aux méthodes de lissage.