

GRAPHIQUES ET VISUALISATION DES DONNÉES AVEC R

DATE: 6 octobre 2014

HEURES: 9h30 à 12h30 et 13h30 à 16h30

DESCRIPTION GÉNÉRALE:

Les capacités graphiques du progiciel R constituent un des ses atouts majeurs. Dans cet atelier, nous explorerons un éventail de fonctions graphiques visant trois types d'utilisation: exploration visuelle des données, aide visuelle aux diagnostics, et représentation graphique des résultats d'analyse. Plus précisément, cet atelier s'appuiera principalement sur l'utilisation de fonctions graphiques de base ainsi que de fonctions contenues dans deux modules à orientation graphique populaires: lattice et ggplot2. Nous aborderons aussi brièvement la préparation des données pour fins de visualisation.

Le contenu du séminaire est pensé principalement dans l'optique d'une application en sciences sociales, mais serait utile à quiconque intéressé par l'utilisation de R pour représenter visuellement des données.

Une connaissance élémentaire de R est un atout car le séminaire n'abordera pas l'utilisation générale du progiciel.

FORMATRICE:

Cet atelier sera animé par Marina Doucerain, étudiante au doctorat en psychologie culturelle et linguistique appliquée à l'Université Concordia. Fervente utilisatrice de Linux et partisane du mouvement "Open Source", son réflexe est de chercher un logiciel à code source libre pour toute nouvelle tâche informatique à laquelle elle est confrontée. C'est ainsi qu'elle a commencé à utiliser R il y a quelques années. Aujourd'hui, elle utilise R pour toutes ses analyses statistiques et elle est convaincue que R est un choix gagnant pour l'analyse de données en sciences sociales.

Elle peut être jointe à marinadoucerain@gmail.com

DESCRIPTION DES ACTIVITÉS:

Interface graphiques

- Brève présentation de Rstudio

Visualisation à fins de diagnostics

- Fonctions du module 'car'
- Identification interactive des points

Exercice 1

Module 'base'

- Exploration visuelle des données
- Visualisation de modèles

Exercice 2

Module 'lattice'

- Exploration visuelle des données
- Prédiction
- Visualisation de modèles
- Couleurs et thèmes

Exercice 3

Module 'ggplot2'

- La logique de la 'grammar of graphics'
- Exploration visuelle des données
- Préparer les données pour les visualiser (plyr, + dplyr si le temps le permet)
- Visualisation de modèles
- Couleurs et thèmes
- Stratégies pour visualiser les grands jeux de données

Exercice 4

Si le temps le permet: applications web, shiny, ggvis

RESSOURCES ADDITIONNELLES:

Une liste de références et de ressources éducatives sera fournie.