

SÉMINAIRE MÉTHODOLOGIQUE

proposé par le service méthodes statistiques de l'Ined

Les rencontres de
statistique appliquée

**VISUALISATION DE DONNÉES :
GRAMMAIRE, PRATIQUES, OUTILS**

Mardi 18 juin 2019

14h-17h

Ined

Salle sauvy

statapp.site.ined.fr/

La « data vizualisation » (dataviz) connaît un essor important depuis quelques années sous l'effet conjoint de l'explosion des données numériques, du développement de technologies performantes et de l'augmentation des puissances de calculs. La maîtrise de l'élaboration des visualisations de données, la capacité de les comprendre sont des enjeux majeurs à l'ère du « big data ».

L'objet de cette séance est de faire un point sur les possibilités actuellement offertes à tout chercheur souhaitant utiliser ou décrypter des représentations graphiques de données. Il s'agira d'évoquer les aspects théoriques (sémiologie graphique et communication visuelle) et les pratiques associées (explorer, synthétiser, communiquer) à la dataviz. Seront ensuite présentés quelques outils performants (intégrés dans les logiciels standards ou applications web) permettant d'élaborer des graphiques « classiques » en analyse statistique et des visualisations de données spatiales ou relationnelles.

Programme et résumés des présentations

14h00	Arnaud Bringé (Ined) et Pascal Cristofoli (EHESS ; Ined) Introduction de la séance
14h15	Timothée Giraud (CNRS, Université de Paris) Outils de cartographie pour la recherche - Magrit et cartography L' UMS RIATE propose deux solutions de cartographie qui visent deux types de besoins de production différents : L'application en ligne MAGRIT a été pensée comme un support à l'enseignement de la cartographie thématique. Son interface graphique est calquée sur le raisonnement suivi lors de la conception d'une carte. Cette application permet à un public non spécialiste de s'initier à la production de cartes respectant les règles de la discipline. Le package « cartography » est une extension du logiciel de statistique et de visualisation libre R. Dans une démarche de recherche reproductible, le package répond au besoin d'intégrer la cartographie dans des chaînes de traitement complètes, comprenant notamment l'acquisition de données ou la statistique.
15h00	Jean-Daniel Fekete (Inria, responsable du projet AVIZ) et Paola Valdivia (Inria) Exploration visuelle d'hypergraphes dynamiques La technique PAOH (Parallel Aggregated Ordered Hypergraph) permet de visualiser des hypergraphes dynamiques. Un hypergraphe est une généralisation de graphe où chaque arrête peut connecter un ou plusieurs sommets. Les hypergraphes permettent de modéliser des réseaux de co-auteurs avec plusieurs auteurs par article, ou des réseaux commerciaux où plusieurs partenaires sont impliqués par contrat. Notre visualisation représente ce type de réseaux évoluant dans le temps, elle est conçue pour des réseaux de taille moyenne, soit 200-1000 sommets, quelques dizaines ou centaines de périodes temporelles. Elle est facile à lire, et le système est très interactif et permet de réaliser des analyses visuellement et interactivement.



	Nous montrerons comment il est utilisé dans le cas d'études en histoire sociale.
16h00	<p>Joseph Larmarange (IRD / Ceped (IRD, Université de Paris, Inserm))</p> <p>ggplot2 ou la grammaire des graphiques sous R</p> <p>R possède un puissant moteur graphique interne. Toutefois, produire un graphique complet avec les fonctions basiques de R est peu aisée. L'extension ggplot2, développée par Hadley Wickham et mettant en œuvre la grammaire des graphiques théorisée par Leland Wilkinson, est devenue en quelques années l'une des extensions R les plus utilisées. La grammaire des graphiques invite l'utilisateur à repenser sa manière de concevoir un graphique en associant variables (données), paramètres graphiques (esthétiques) et formes géométriques, puis en complexifiant progressivement son graphique (coordonnées, échelles, facettes, étiquettes, légendes, thèmes...).</p>
16h45	Discussion et conclusion