

Introduction à la sociologie quantitative. Plan de cours

Samuel Coavoux

février-mai 2019

Niveau 1 : Lire les statistiques

Objectifs théoriques :

- être capable de lire de façon critique un article de sociologie quantitative ;
- être capable de concevoir un projet d'analyse statistique.

Objectif pratique :

Concevoir un projet d'analyse statistique en rapport avec votre thématique de recherche :

- trouver une base de données ou développer un plan de production d'une base de donnée ;
- sélectionner les outils statistiques pertinents

À rendre à la fin du cours : une page de projet d'analyse

Séance 1 : Les bases de données

- Vocabulaire de la statistique
- Sources de données
- Construire une base de donnée
- Théorie de l'échantillonnage

Séance 2 : Statistique descriptive

- Recodage, nettoyage de données
- Statistique descriptive : univariée
 - numérique : tendance centrale, dispersion, représentations graphiques (diagramme de densité, histogramme, boîte à moustache)
 - catégorielle : tri à plat, représentation graphique (diagramme en barre)
- Statistique descriptive : bivariée
 - num/num : diagramme en point, diagramme en ligne, coefficient de corrélation
 - num/cat : boîte à moustache, comparaison de moyennes
 - cat/cat : tableau croisée

Séance 3 : Inférence et test d'hypothèse

- Théorie des tests d'hypothèse
- Test du khi-2
- Test de student
- Crise la répliquabilité et limites de la p-value

Séance 4 : Analyse géométrique de données

- Introduction au raisonnement géométrique
- Analyse en composantes principales
- Analyse des correspondances multiples

Séance 5 : Modèles de régression

- Modèle OLS
- Modèle linéaire généralisé : régression logit

Niveau 2 : Produire les statistiques

Objectif théorique : être capable de mettre en pratique de façon autonome l'analyse statistique d'une base de donnée simple sous R.

Objectif pratique : écrire le code et l'interprétation d'une analyse statistique complète (mais qui peut avoir des dimensions mesurées). Scripts d'import et de recodage, manipulation de données, analyses, graphiques, et interprétation.

À rendre à la fin du cours : un document html/pdf écrit en Rmarkdown

Séance 6 : introduction à R, manipulation de données

- Prise en main de R et de Rstudio
- Scripter : conseils, méthodes
- Principe de R : objets et fonctions
- Import et manipulation de données : introduction au tidyverse
- Indexation
- Recodage/ transformation

Séance 7 : Analyses statistiques

- Tri à plat, indicateurs de tendance centrale et de dispersion
- Tris croisés
- Test d'hypothèses
- Manipulation d'objet complexes

- AGD avec FactoMineR : ACM, ACP
- Régression (lm, glm)

Séance 8 : Produire des graphiques

- Introduction à la grammaire des graphiques
- Syntaxe de ggplot2
- Principaux types de graphiques :
 - diagramme en point
 - diagramme en ligne
 - diagramme de densité
 - histogramme
 - boîtes à moustache
 - barplot
- Comparaison : facettes, couleurs, etc.

Séance 9 : Écriture de rapports/rmarkdown

- Introduction à Markdown
- Écrire en Rmarkdwon
 - des rapports
 - des slides
- Savoir chercher de l'aide pour R
- Résoudre une erreur. Fabriquer un exemple reproductible minimal

Bibliographie

Couto, Fanny, Marie-Paule et Bugeja-Bloch. 2015. *Les méthodes quantitatives*. Paris : Presses Universitaires de France.

Maillochon, Marion, Florence et Selz. 2009. *Le raisonnement statistique en sociologie*. Paris : Presses Universitaires de France.

Martin, Olivier. 2009. *L'analyse de données quantitatives*. Paris : Armand Colin.

Reinhart, Alex. 2015. *Statistics done wrong. The Woefully Complete Guide*. San Francisco : No Starch Press. <https://www.statisticsonewrong.com/index.html>.

Wickham, Hadley. 2015. *Advanced R*. Boca Raton : CRC Press.