

Méthodes quantitatives pour les sciences sociales 2019-2020

Syllabus

Enseignant-e-s :

Augustin Guillot (histoire) – Solenn Huitric (histoire) – Camille Martin (sociologie) – Emmanuelle Santoire (géographie) – Sinan Sarpca (économie)

Présentation

L'objectif général est de proposer une offre unifiée et cohérente de formation aux méthodes quantitatives, accessible aussi bien aux étudiants du Master de sciences sociales qu'aux étudiants des autres mentions dans lesquelles le département s'implique. Cette formation s'appuie sur un socle de remise à niveau en statistiques simples et d'initiation aux méthodes quantitatives pour les sciences sociales, ensuite prolongé et approfondi dans le cadre d'un atelier de recherche s'appuyant sur une formation avancée aux méthodes quantitatives, dans le cadre de la mise en œuvre d'une recherche collective sur les données d'une grande enquête quantitative réalisée par des instituts de statistique et de recherche. Cet atelier de recherche est également le cadre dans lequel les étudiant-e-s qui le souhaitent peuvent bénéficier d'un accompagnement pour la mobilisation de méthodes quantitatives dans leurs propres travaux, du Master au doctorat.

Prérequis

La formation aux méthodes quantitatives en sciences sociales ne comporte pas de difficulté mathématique particulière. Nous vous proposons tout de même un niveau 1 de formation, qui, au début de l'année, vous permettra de voir ou de revoir les bases des méthodes quantitatives, et d'arriver plus assuré-e aux autres étapes de la formation. En dehors de cette initiation, il n'y a aucun prérequis particulier pour le premier semestre du niveau 2 : tou-te-s les étudiant-e-s de toutes les disciplines, de la première année de scolarité à l'ENS jusqu'au doctorat, peuvent suivre cette formation, et cela quel que soit leur niveau initial en mathématiques, statistiques, informatique : en fonction de ce niveau initial, justement, vous serez orienté-e vers le niveau de la formation qui vous convient le mieux.

Matériel

Le suivi des cours et ateliers de méthodes quantitatives requiert l'utilisation en séance d'un ordinateur portable personnel, équipé d'un « tableur » (Excel recommandé ou logiciel libre équivalent) et permettant de se connecter à internet en Wifi. Les étudiant-e-s ne disposant pas d'ordinateur pourront le signaler aux enseignant-e-s, quelques ordinateurs de prêt pouvant être mis à disposition pendant les séances.

Bibliographie indicative

BARNIER Julien, *Introduction à R. Version 2019* (en ligne),
<https://juba.github.io/tidyverse/>

BUGEJA-BLOCH Fanny, COUTO Marie-Paule, 2015, *Les méthodes quantitatives*, Paris, Puf, coll. « Que sais-je ? ». Voir en ligne <http://www.cairn.info/acces.bibliotheque-diderot.fr/les-methodes-quantitatives--9782130631613.htm>.

CELLIER Jacques COCAUD Martine, *Le traitement de données en histoire et sciences sociales*, Rennes, Presses universitaires de Rennes.

CIBOIS Philippe 2015 1ère éd 2007 *Les méthodes d'analyse d'enquête*, Lyon, ENS Editions, coll. « Bibliothèque idéale des sciences sociales », préf. Pierre Mercklé. Voir en ligne : <http://books.openedition.org/enseditions/1443>.

COMBESSIE Jean-Claude, 2001, *La méthode en sociologie*, Paris, La Découverte, coll. « Repères » Voir en ligne <http://www.cairn.info/la-methode-en-sociologie--9782707152411.htm>.

CORNILLON Pierre-André (dir.), 2010, *Statistiques avec R*, Rennes, Presses universitaires de Rennes, 2ème éd . revue et augmentée.

DUMOLARD Perre, DBUS Nahalie, CHARLEUX Laure, *Les statistiques en géographie*, Paris, Belin.

GROUPE CHADULE, 1997, *Initiation aux pratiques statistiques en géographie*, 4e éd., Paris, ArmandColin, 203 p

LEMERCIER Claire, ZALC Claire, 2008, *Méthodes quantitatives pour l'historien*, Paris, La Découverte, coll. « Repères ». Voir en ligne <http://www.cairn.info/methodes-quantitatives-pour-l-historien--9782707153401.htm>.

MAILLOCHON Forence, SELZ Marion, 2009, *Le raisonnement statistique en sociologie*, Puf, coll. « Licence Socio »

MARTIN Olivier, 2005, *L'analyse de données quantitatives*, Paris, ArmandColin, coll. « 128 Sociologie ».

PAUGAM Serge (dir.), 2010, *L'enquête sociologique*, Paris, Puf, coll. « Quadrige Manuels »

QUANTI Tutti, 2016, *Introduction à l'analyse des données pour les sciences sociales avec R*, Lyon, Département de sciences sociales,

<https://docs.google.com/document/d/1Y8jt6cz3EWg8qABJH3jzURBgiS9BVJsnUxxCETZoagE/edit>

Niveau 1 : Les principes fondamentaux de l'analyse des données

Enseignants : Emmanuelle Santoire et Augustin Guillot

Présentation

Le niveau 1 de l'offre de formation aux méthodes quantitatives propose, à celles et ceux qui n'ont fait que peu de mathématiques et de statistiques au cours de leur cursus, une remise à niveau en matière d'analyse des données. Cette remise à niveau, qui a lieu au tout début du semestre 1, permet aux étudiant-e-s qui la suivent de redécouvrir les grands principes élémentaires de l'analyse des données : structuration des données, types de variables, calcul d'indicateurs de position (moyenne, médiane...) et de dispersion (écart-type, quantiles...), analyse des distributions, analyse des relations entre deux variables, représentations graphiques simples... Ces « révisions » font l'objet de courts rappels théoriques et d'exercices d'appropriation du tableur. En cours, les présentations seront proposées à partir d'Excel.

Ce module est conçu comme une sorte de première formation dans le domaine des méthodes quantitatives visant à donner les tout premiers outils d'appréhension des statistiques et de leurs usages en sciences sociales. Ce premier ensemble de cours s'adresse à tous les étudiants au cursus littéraire (CPGE série A/L, cursus universitaires en sciences sociales ou littéraires) qui n'ont fait jusque-là que peu de mathématiques et de statistiques, pour leur permettre d'aborder dans de bonnes conditions les outils et les logiciels utilisés ensuite en méthodes quantitatives.

Conçu à l'origine pour les géographes et historiens en première année, le cours est désormais ouvert dans le cadre du Diplôme à tous ceux qui souhaitent s'initier aux méthodes quantitatives, y compris dans la perspective d'études textométriques ou lexicographiques. Ce cours ne se substitue pas aux cours d'initiation au tableur dispensé par le CID mais le complète. Organisé en 5 séances hebdomadaires en tout début d'année, il est conçu également pour les étudiants arrivés en deuxième année du diplôme de l'ENS et qui suivent le cours de méthodes quantitatives de niveau 2 lors du même semestre. Cet enseignement vise à :

- présenter de manière critique la place des méthodes quantitatives en sciences sociales et d'en présenter les principaux usages;
- rappeler les principes de base de la collecte et de la mise en forme de données ;
- poser les principes de la statistique descriptive, univariée et bivariée

Le cours se propose d'utiliser comme exemple filé les données d'une station météo et d'une station de mesure de la pollution lyonnaises.

Les dates, heures et salles du cours seront précisées au dernier moment en fonction des disponibilités d'emploi du temps. Si vous souhaitez vous inscrire, outre le questionnaire en ligne, écrivez directement sur l'adresse suivante : camille.martin@ens-lyon.fr

Modalités d'évaluation

L'évaluation repose sur le contrôle de l'assiduité et des travaux écrits (sous la forme d'exercices).

Bibliographie

Cf. Bibliographie indicative présentée plus haut.

Séances

Séance 1 : Les données : où ? quoi ? comment ?

Séance 2 : Appropriation du tableau élémentaire

Séance 3 : Analyse univariée : Valeurs centrales et valeurs de dispersion

Séance 4 : Discrétisation et introduction aux principes des analyses bivariées

Séance 5 : Régression linéaire et initiation au Khi^2

Niveau 2 : niveau intermédiaire

Résumé du cours :

Le cours de niveau 2 se divise en deux semestres (2 x 21h), et a pour objectif de rendre les étudiant-e-s autonomes dans la réalisation d'une étude statistique simple. Au cours du premier semestre, il s'agit de se former à l'utilisation du logiciel de traitement de données R et à son langage de programmation, pour parvenir petit à petit à maîtriser l'ouverture des bases de données, leur nettoyage, la réalisation de travaux statistiques simples (analyses uni- et bivariées), et la présentation des résultats sous forme de tableaux ou de graphiques. Il s'agit également de maintenir un usage réflexif du traitement des données, et de s'interroger sur le mode de raisonnement statistique.

Au cours du second semestre, les étudiant-e-s sont formé-e-s à des techniques statistiques plus avancées (analyse géométrique, régressions, traitement spatial, analyse textuelle) et sont chargé-e-s de réaliser en autonomie une petite étude quantitative sur une base de données de la statistique publique. Les travaux seront présentés par groupe à l'occasion d'une journée d'études *ad hoc* à la fin de l'année.

Semestre 1 : SCS04105

Planning et salles pour le semestre :

[Groupe 1 : le lundi 14-17h // groupe 2 : 9h30-12h30 // Groupe 3 : Mardi 15-18h]

- Séance 1 : 23-24 septembre // D2 115 // D2 018 // D2 121

- Séance 2 : 7-8 octobre // D2 115 // D2 018 // D2 121

- Séance 3 : 21-22 octobre // D2 115 // D2 018 // D2 121

- Séance 4 : 4-5 novembre // D2 115 // D2 018 // D2 121

- Séance 5 : 18-19 novembre // D2 115 // D2 018 // D2 121

- Séance 6 : 2-3 décembre // D2 115 // D2 109 // D2 121

- Séance 7 : 16-17 décembre // D2 115 // D2 002 // D2 002

Les données utilisées durant ce semestre seront les données de l'enquête ESS (European Social Survey) de 2016.

Séance 1 : Connaître les sources des données et les obtenir

Augustin Guillot, Camille Martin et Solenn Huitric, lundi 23 et mardi 24 septembre 2019

Salles :

Lundi AM : D2 115

Mardi matin : D2 018

Mardi AM : D2 121

Bibliographie indicative

Caveng Remy, « La production des enquêtes quantitatives », *Revue d'anthropologie des connaissances* 1/2012 (Vol. 6, n° 1), p 65-88. Voir en ligne www.cairn.info/revue-anthropologie-des-connaissances-2012-1-page-65.htm.

Cayouette-Remblière Joanie « Reconstituer une cohorte d'élèves à partir de dossiers scolaires. La construction d'une statistique ethnographique », *Genèses* 4/2011 (n° 85), p. 115-133. Voir en ligne www.cairn.info/revue-geneses-2011-4-page-115.htm.

Chenu Alain et Lesnard Laurent, 2011, *La France dans les comparaisons internationales. Guide d'accès aux grandes enquêtes statistiques en sciences sociales*, Paris, Les Presses de Sciences Po, 221 p.

Grémy Jean-Pierre, 1987, « Les expériences françaises sur la formulation des questions d'enquête. Résultats dun premier inventaire », *Revue française de sociologie*, vol. XXVIII n° 4, octobre-décembre, pp. 567-599. Voir en ligne

http://www.persee.fr/articleAsPDF/rf_soc0035-29691987num2842446/articlerfsoc0035-29691987num2842446.pdf.

Hobeika Alexandre, Ollion Etienne, 2012, "Vers une "data-sociologie" ? Tour d'horizon des techniques de collecte automatisée de données sur le web", *Quanti*, 3 octobre 2012, en ligne : <http://quanti.hypotheses.org/647/>.

Ollion Etienne (2011), « Collecte automatique de données, l'exemple des évaluateurs anonymes (referees) de l'ASR », *Sociologie*, n° 3, vol. 2. Voir en ligne <http://sociologie.revues.org/1126>.

Séance 2 : Introduction à R

Augustin Guillot, Camille Martin et Solenn Huitric, lundi 7 octobre et mardi 8 octobre 2019

Salles :

Lundi AM : D2 115

Mardi matin : D2 018

Mardi AM : D2 121

Cette séance vise à présenter le logiciel, à définir la différence entre commande par "bouton" et par programmation et à se familiariser avec l'utilisation de commandes dans la console ou dans des scripts. Dans un deuxième temps, il s'agit d'apprendre des opérations simples : importer des données, les inspecter, acquérir les commandes pour des traitements de statistique univariée, produire un graphique.

Bibliographie indicative

BARNIER Julien, *Introduction à R. Version 2.2*, 2016 (en ligne),

<https://alea.fr.eu.org/pages/intro-R>

CORNILLON Pierre-André (dir.), 2010, *Statistiques avec R*, Rennes, Presses universitaires de Rennes, 2ème éd. revue et augmentée.

GROUPE ELEMENTR 2014, *R et espace. Traitement de l'information géographique*, Framasoft, 244 p., <http://framabook.org/r-et-espace/>.

QUANTI Tutti, 2016, *Introduction à l'analyse des données pour les sciences sociales avec R*, Lyon, Département de sciences sociales,

<https://docs.google.com/document/d/1Y8jt6cz3EWg8qABJH3jzURBgiS9BVjsnUxxCETZoagE/edit>

Séance 3 : Explorer et manipuler les données **Emmanuelle Santoire, Camille Martin et Solenn Huitric, lundi 21 et** **mardi 22 octobre 2019**

Salles :

Lundi AM : D2 115

Mardi matin : D2 018

Mardi AM : D2 121

Bibliographie indicative

Caporali Arianna, Morisset Amandine, Legleye Stéphane, 2015, « La mise à disposition des enquêtes quantitatives en sciences sociales: l'exemple de l'Ined », *Population* 3/ (Vol. 70), p 567-597. Voir en ligne www.cairn.info/revue-population-2015-3-page-567.htm.

QUANTI Tutti, 2016, *Introduction à l'analyse des données pour les sciences sociales avec R*, Lyon, Département de sciences sociales,

<https://docs.google.com/document/d/1Y8jt6cz3EWg8qABJH3jzUR BgiS9BVjsnUxxCEtZoagE/edit>

Gollac Michel, 1997, « Des chiffres insensés ? Pourquoi et comment on donne un sens aux données statistiques ». In *Revue française de sociologie*, 38-1. pp. 5-36. Voir en ligne www.persee.fr/doc/rfsoc0035-29691997num3814570

Séance 4 : Recoder les données : principes fondamentaux (1) **Emmanuelle Santoire, Camille Martin et Solenn Huitric, lundi 4 et mardi 5** **novembre 2019**

Salles :

Lundi AM : D2 115

Mardi matin : D2 018

Mardi AM : D2 121

Bibliographie indicative

Amossé T, Chardon O., 2006, « Les travailleurs non qualifiés une nouvelle classe sociale ? », *Économie et Statistique*, n° 393-394

Henrotin, J., Vaissière, M., Etaix, M., Malard, S. Dziurla, M. & Lafon, D. (2016). Codage d'emplois (PCS 2003) : retour d'expérience d'une étude menée en service de santé au travail. *Santé Publique*, vol. 28,(4), 471-480.

Chauvel, Louis. « Les données et le codage des groupes socio-professionnels », *Le destin des générations. Structure sociale et cohortes en France du XX^e siècle aux années 2010*. Presses Universitaires de France, 2010, pp . 374 -381.

Séance 5 : Recoder les données : principes fondamentaux (2) **Augustin Guillot, Camille Martin et Solenn Huitric, lundi 18 et mardi 19 novembre 2019**

Salles :

Lundi AM : D2 115
Mardi matin : D2 018
Mardi AM : D2 121

Bibliographie indicative

Jean-Claude Passeron, « L'espace mental de l'enquête (II) », *Enquête* [En ligne], 3 | 1996, mis en ligne le 11 juillet 2013, URL <http://enquete.revues.org/393>.

Thébaud-Mony, A. (2006) Histoires professionnelles et cancer . *Actes de la recherche en sciences sociales*, n. 163,(3), 18-31

Béliard, A. & Eideliman, J (2014) Mots pour maux Théories diagnostiques et problèmes de santé. *Revue française de sociologie*, vol. 55,(3), 507-536.

Séance 6 : Analyser les relations entre variables **Augustin Guillot, Camille Martin et Solenn Huitric, lundi 3 et mardi 3 décembre 2019**

Salles :

Lundi AM : D2 115
Mardi matin : D2 109
Mardi AM : D2 121

Bibliographie indicative

Herpin Nicolas, 2003, « La taille des hommes son incidence sur la vie en couple et la carrière professionnelle », *Économie et statistique*, n°361, p. 71-90. Voir en ligne <http://www.insee.fr/fr/ffc/docsffc/ES361D.pdf>.

Mercklé Pierre, 2008, « On peut faire des tableaux croisés avec Excel ! », *Quanti*, <http://quanti.hypotheses.org/33/>.

Merle Pierre, « Démocratisation ou accroissement des inégalités scolaires? L'exemple de l'évolution de la durée des études en France (1988-1998)», *Population* 4/2002 (Vol. 57) p. 633-659. Voir en ligne

www.cairn.info/revue-population-2002-4-page-633.htm.

Passeron Jean-Claude, « Ce qu'un tableau dit et ce qu'on en dit », in *Le raisonnement sociologique. Un espace non-poppérien de l'argumentation*, Paris, Albin Michel, coll. « « Bibliothèque de l'évolution de l'Humanité », 2006 [1991].

Zarca Bernard, 2006, « Mathématicien : une profession élitare et masculine », *Sociétés Contemporaines*, n°64, p. 41-65. Voir en ligne : <http://www.cairn.info/revue-societes-contemporaines-2006-4-p-41.htm>.

Séance 7 : Visualiser les données : tableaux, graphiques, figures Augustin Guillot, Camille Martin et Solenn Huitric, lundi 3 et mardi 3 décembre 2019

Salles :

Lundi AM : D2 115

Mardi matin : D2 002

Mardi AM : D2 002

Bibliographie indicative

Benveniste Corinne, 2009, « Des graphiques et des tableaux », *Courrier des statistiques*, n° hors-série, pp. 13-17,

<http://www.insee.fr/fr/themes/document.asp?regid=0&id=2759>.

Bertin Jacques 2013 1ère éd 1967 *Sémiologie graphique. Les diagrammes, les réseaux, les cartes*, Paris, EHESS éditions, réimpr. de la quatrième édition, 2005.

Grémy Jean-Paul, 2003, « Introduction à la lecture de tableaux statistiques » *Méthodes quantitatives pour les sciences sociales*,

<http://www.scienceshumaines.com/textesInedits/Gremy.pdf>.

Henry Nathalie et Fekete Jean-Daniel, 2008, « Représentations visuelles alternatives pour les réseaux sociaux », *Réseaux*, n° 152, pp. 59-92, <http://www.cairn.info/revue-reseaux-2008-6-page-59.htm>.

Monso Olivier et De Saint Pol Thibaut, 2009, « La perception graphique : mieux construire et interpréter les graphiques », *Courrier des statistiques*, n° 129, pp. 67-74, <http://www.insee.fr/fr/themes/document.asp?regid=0&refid=COU126L> .

Passeron Jean-Claude, 1991, « Ce qu'un tableau dit et ce qu'on en dit », *Le raisonnement sociologique*, Paris, Nathan, pp. 111-133.

Relieu Marc, 1999, « Du tableau statistique à l'image audiovisuelle. Lieux et pratiques de la représentation en sciences sociales », *Réseaux*, vol. 17, n° 94, pp. 49-86, <http://www.persee.fr/web/revues/home/prescript/article/reso0751-79711999num17942140>.

Tufte Edward R., 1983, *The Visual Display of Quantitative Information*, Cheshire (Connecticut), Graphics Press, 2nd ed. 2001.

Niveau 2, semestre 2 : SCSO4205

Les cours ont lieu le mardi de 9h30 à 12h30.

Le cours sera assuré par Augustin Guillot, doctorant en histoire. Le contenu des données est encore à définir. Les séances permettront de se former aux regressions logistiques, à la réalisation d'analyses factorielles. Ce qui aura été vu au premier semestre sera approfondi.

Dates des séances :

- Séance 1 : 14 janvier
- Séance 2 : 21 janvier
- Séance 3 : 28 janvier
- Séance 4 : 04 février
- Séance 5 : 11 février
- Séance 6 : 25 février
- Séance 7 : Datathon + Journée d'étude : début mai, certainement

Niveau 3 : Séminaire de recherche sur la mesure de la mobilité sociale

Les cours ont lieu le jeudi matin, de 9h à 13h.

Le niveau 3 est un atelier de recherche dont le but est de produire une recherche collective originale.

Cet atelier est d'abord pensé pour ceux qui ont déjà suivi le niveau 2 de la formation quantitative, mais il est possible de suivre les deux niveaux en parallèle.

La première partie de l'atelier sera généralement l'occasion de présenter des techniques statistiques plus complexes que celles vues lors du deuxième niveau (économétrie, modèles log linéaire, techniques de webscraping, etc.). La seconde partie se fera sous forme d'un travail collectif sur les bases de données que nous fouillerons..

En début d'année, un travail de bibliographie et de formulation d'une question de recherche sera également fait, en accord avec ce qui nous semble à la fois exploitable et sous-exploité dans ces bases de données.

Puis il s'agira de réaliser des traitements statistiques simples et nécessaires à la suite de la recherche : recodage, tableaux croisés, etc. Un premier rendu contenant les statistiques descriptives de notre objet de recherche sera à rédiger à la fin du premier semestre. La seconde partie de l'année sera consacrée à l'application de méthodes statistiques avancées et explicatives sur notre objet de recherche (régressions, modèles log linéaire, etc.).

Des chercheurs seront invités à certaines séances pour nous présenter certains résultats et techniques statistiques.

C'est un séminaire annuel : le cours ne peut être validé qu'en suivant les deux semestres.

Semestre 1 : SCS05105

Les séances de l'atelier auront lieu de 9h à 13h (avec, rassurez vous, une bonne pause au milieu !), à un rythme mensuel :

- - Séance 1 : 3 octobre // D2 108
- - Séance 2 : 24 octobre // D2 108
- - Séance 3 : 14 novembre // D2 108
- - Séance 4 : 19 décembre // D2 108
- - Séance 5 : 16 janvier // D2 020

Semestre 2 : SCS05205

Les dates du second semestre sont les suivantes (les salles ne sont pas encore réservées) :

- - Séance 1 : 30 janvier
- - Séance 2 : 20 février
- - Séance 3 : 12 mars
- - Séance 4 : 2 avril
- - Séance 5 : 23 avril