

Méthodes quantitatives

Licence 3 – Semestre 5

20 Heures ($10 \times 2h$)

Enseignant : Ahmed Fouad EL HADDAD

Syllabus

24 septembre 2025

Table des matières

1	Modalités d'évaluation	3
2	Ressources et manuels de référence	3
3	Programme du cours	3

Introduction

Ce cours propose une introduction critique à la collecte et au traitement des données statistiques en sciences sociales. En suivant les étapes du raisonnement empirique, il poursuit trois objectifs complémentaires : *(i)* classer les données selon leurs modes de production (enquêtes, sources administratives, traces numériques) ; *(ii)* problématiser les opérations de construction, de codification et de transformation des données ; *(iii)* initier aux principes élémentaires de l'inférence statistique.

Chaque séance articule un enjeu théorique à une étude de cas concrète, tirée d'épisodes historiques, d'enquêtes emblématiques ou de controverses contemporaines.

1 Modalités d'évaluation

- **Régime général** : 100% écrit (épreuve terminale).
- **Régime dérogatoire** : 100% écrit.
- **Session 2** : 100% oral.

2 Ressources et manuels de référence

- Chanvril-Ligneel, F., & Le Hay, V. (2014). *Méthodes statistiques pour les sciences sociales*. Paris : Ellipses.
- Fox, W. (1999). *Les statistiques sociales*. Québec : Presses de l'Université Laval.
- Pétry, F. (2003). *Guide pratique d'introduction à la régression en sciences sociales*. Québec : PUL.
- Arel-Bundock, V. (2021). *Analyse causale et méthodes quantitatives*. Montréal : PUM.
- Salganik, M. J. (2019). *Bit by Bit : Social Research in the Digital Age*. Princeton : Princeton UP.

3 Programme du cours

Séance 1 — Qu'est-ce qu'une donnée ? Genèse, matérialité, autorité

Objectif : Comprendre que la donnée n'est pas un donné.

Contenu : Définitions (trace, abstraction, artefact) ; origines des dispositifs statistiques (recensement, fiscalité, mortalité) ; autorité de la donnée.

Cas pratique : La circulaire du 16 floréal : la statistique de population sous le Consulat et l'Empire.

Séance 2 — La mise en données : problématiser la mesure

Objectif : Concevoir la donnée comme approximation structurée du réel.

Contenu : De la conceptualisation à l'opérationnalisation ; dimensions, indicateurs, indices ; typologies de variables (nominale, ordinale, continue).

Cas pratique : Mesurer un fait social : le suicide chez Durkheim (*Le Suicide*, 1897).

Séance 3 — Catégoriser, codifier, classer

Objectif : Comprendre le recodage comme compromis entre perte d'information et gain d'intelligibilité.

Contenu : Réduction (agrégation, discrétisation, codification) ; enrichissement (création de variables par combinaison).

Cas pratique : Recoder les CSP selon Bourdieu, Marx et Weber.

Séance 4 — Cartographie des sources

Objectif : Classifier et évaluer les grandes familles de données.

Contenu : Données digitales vs non digitales ; enquêtes vs administratives ; données trouvées vs construites.

Cas pratique : Extraction automatisée de fiches Wikipédia sur les lauréats du prix Nobel.

Séance 5 — Construire une enquête

Objectif : Concevoir un questionnaire *ad hoc*.

Contenu : Échantillonnage (aléatoire, raisonné, boule de neige) ; maturation et saturation empirique ; biais et effets de cadrage.

Cas pratique : L'élection présidentielle américaine de 1936 (*Literary Digest* vs *Gallup*).

Séance 6 — Exploiter un jeu de données secondaire

Objectif : Développer une lecture critique des bases existantes.

Contenu : Infrastructures de recherche ; dictionnaire de variables ; tendance centrale et dispersion ; pondération, imputation.

Cas pratique : European Social Survey (ESS).

Séance 7 — Structurer et nettoyer les données

Objectif : Comprendre le nettoyage comme alignement entre empirie et théorie.

Contenu : Structuration tabulaire ; correction des incohérences (valeurs manquantes, doublons).

Cas pratique : Données sur les candidats aux législatives en Seine-et-Marne (1988–2012).

Séance 8 — Décrire ses données

Objectif : Résumer la distribution des variables en vue de l'analyse statistique.

Contenu : Analyse de la distribution (asymétrie, aplatissement, outliers) ; indicateurs selon type de variable ; représentations graphiques.

Cas pratique : Comparative Manifesto Project (CMP).

Séance 9 — Tester une hypothèse

Objectif : Comprendre le test comme procédure d'évaluation de la vraisemblance.

Contenu : Logique des tests et interprétation de la p-value ; choix du test selon type de

variables ; erreurs statistiques.

Cas pratique : Données de l'enquête Trajectoires et Origines (TeO).

Séance 10 — Éthique, responsabilité et transparence

Objectif : Réfléchir aux responsabilités morales du traitement de données.

Contenu : Consentement, anonymisation, biais implicites ; reproductibilité et transparence scientifique.

Cas pratique : Affaire hiQ Labs vs LinkedIn (2017–2022).