

Analyse des données sur R

ECTS : 3

Description du contenu de l'enseignement :

À la suite du cours de « Théorie des statistiques », ce cours est conçu comme un enseignement pratique en salle informatique. Les étudiant.e.s appliqueront ainsi les méthodes statistiques étudiées à des bases de données secondaires (issues notamment de grandes enquêtes en sciences sociales) grâce au logiciel d'analyse de données quantitatives R.

L'objectif du cours de « Pratique des statistiques sur R » est d'amener les étudiant.e.s à produire eux/elles-mêmes des résultats intéressant les sciences sociales à partir de bases de données quantitatives secondaires grâce au logiciel R. La production de données quantitatives en sciences sociales est centrale pour les métiers de chargé.e.s d'enquête et d'études ou de chercheur.e.s auxquels prépare le Master 2 PPO.

Le cours se présente sous la forme de 6 séances de travail de pratique sur le logiciel R en salle informatique. Des exercices pratiques seront effectués avant de travailler sur des bases de données secondaires. La dernière séance sera consacrée au travail individuel de chaque étudiant.e sur la base de données qu'il/elle aura choisi, en fonction de la thématique qu'il/elle aura choisi, en vue du dossier « 8 pages » à rendre pour valider le cours.

Il est attendu des étudiant.e.s qu'ils/elles soient ponctuel.le.s, assidu.e.s d'une séance à l'autre et qu'ils/elles réalisent les exercices demandés d'une séance sur l'autre.

Compétence à acquérir :

À la fin de ce cours, les étudiant.e.s auront les compétences suivantes :

- Maîtrise des logiciels R et RStudio
- Data management de base de données quantitatives (nettoyage, recodage, etc.)
- Maîtrise et mise en œuvre des méthodes statistiques univariées, bivariées et d'inférence statistique
- Maîtrise et mise en œuvre des méthodes d'analyse géométrique des données (ACP, AFC, ACM) et de classification
- Compréhension de l'utilité des méthodes quantitatives étudiées pour les sciences sociales
- Production de résultats intéressant les sciences sociales à partir de données secondaires
- Conception de rapports types "4 pages".

Mode de contrôle des connaissances :

Les étudiant.e.s valideront les deux cours de « Pratique des statistiques sur R » et « Pratique des statistiques sur SAS » à partir d'un unique dossier de 8 pages. Ils/elles travailleront sur une même base de données secondaire, qu'ils/elles auront choisie, à partir des logiciels R et SAS.

Les exercices réguliers demandés entre chaque cours devront être envoyés à l'enseignante et pourront être notés.

La présence et la ponctualité étant requises, des points pourront être retirés sur la note finale si elles font défaut.

Bibliographie, lectures recommandées :

- La technique statistique

- Pour commencer

BUGEJA-BLOCH Fanny et COUTO Marie-Paule, 2015. Les Méthodes quantitatives, Que sais-je ?, PUF, Paris, 128 p.

CHANVRIL-LIGNEEL Flora et LE HAY Viviane, 2014. Méthodes statistiques pour les sciences sociales, Ellipses, Paris, 261 p.

MARTIN Olivier, 2009. L'Analyse de données quantitatives, Armand Colin, Paris, 128 p.

- Pour se former

BUSCA Didier et TOUTAIN Stéphanie (dir.), 2009. Analyse factorielle simple en sociologie. Méthodes d'interprétation et études de cas, De Boeck, Bruxelles, 259 p.

CIBOIS Philippe, 2009. Les Méthodes d'analyse d'enquêtes, PUF, Paris, 78 p.

Disponible sur internet : <http://cibois.pagesperso-orange.fr/PhilippeCiboisMethAnaEnq.pdf>

LEBARON Frédéric, 2005. L'Enquête quantitative en sciences sociales. Recueil et analyse de données, Dunod, Paris, 182 p.

- Pour approfondir

ESCOFFIER Brigitte et PAGES Jérôme, 2008. Analyses factorielles simples et multiples. Objectifs, méthodes et interprétation, Dunod, Paris, 284 p.

LE ROUX Brigitte et ROUANET Henry, 2010. Multiple Correspondance Analysis, Sage, Thousand Oaks, 128 p.

ROUANET Henry, LEBARON Frédéric, LE HAY Viviane, ACKERMANN Werner et LE ROUX Brigitte, 2002. « Régression et analyse géométrique des données : réflexions et suggestions », Mathématiques et sciences humaines, n°160, p. 13-45.

• Le logiciel

- Pour commencer

BARNIER Julien, Introduction à R et au tidyverse

<https://juba.github.io/tidyverse/>

PARADIS Emmanuel, R pour les débutants

https://cran.r-project.org/doc/contrib/Paradis-rdebuts_fr.pdf

ZORN Thierry, LETHROSNE Murielle, ROUSSEZ Vivien, IRZ Pascal & TORTEROTOT Nicolas, Découvrir R et RStudio

https://mtes-mct.github.io/parcours_r_socle_introduction/

- Pour approfondir

HUSSON François (dir.), R pour la statistique et la science des données, Rennes, PUR, 2018.

<https://r-stat-sc-donnees.github.io/>

LARMARANGE Joseph, analyse-R

<https://larmarange.github.io/analyse-R/>

- Pour effectuer des recherches

Rseek (moteur de recherche en anglais dédié à R)

<https://rseek.org/>

Rzine (recensement de ressources en livres d'accès sur la pratique de R en sciences humaines et sociales)

<https://rzine.fr/>

Document susceptible de mise à jour - 03/04/2026

Université Paris Dauphine - PSL - Place du Maréchal de Lattre de Tassigny - 75775 PARIS Cedex 16