Contenus

Code du Cours	Nom du Cours	Semestre du Cours	Cours Théoriques	Travaux Dirigés (TD)	Travaux Pratiques (TP)	Crédit du Cours	ECTS
RI203	Méthodes statistiques en sciences sociales	3	2	0	0	2	3

Cours Pré-Requis	
Conditions d'Admission au Cours	

Langue du Cours	Français
Type de Cours	Obligatoire
Niveau du Cours	Licence
Objectif du Cours	Ce cours vise à permettre aux étudiants menant des recherches en sciences sociales et en relations internationales d'acquérir l'état d'esprit statistique nécessaire pour gérer les incertitudes auxquelles ils sont confrontés. Ne nécessitant aucune connaissance préalable en programmation, le cours commence par l'acquisition de compétences en manipulation, nettoyage, transformation et visualisation de données avec des techniques graphiques modernes sous R ; la logique de l'analyse reproductible via R Markdown/Quarto est transmise aux étudiants dès les premières semaines. Au fil des semaines, la construction et l'interprétation de modèles de régression linéaire simple et multiple en adéquation avec les questions de recherche sont abordées, ainsi que l'analyse des différences entre groupes et des relations conditionnelles au moyen de variables muettes et de termes d'interaction. En fin de semestre, la logique de base de l'inférence causale et l'approche des doubles différences sont introduites ; la structure des données de panel et la manière dont l'hétérogénéité non observée des unités est traitée par les approches des effets fixes, des premières différences et des effets aléatoires sont présentées. Tout au long du cours, les étudiants sont amenés à interpréter de manière critique les outils d'inférence statistique tels que l'erreur standard, l'intervalle de confiance, la p-valeur et la taille d'effet, et à transformer leurs résultats en intuitions valides et en preuves empiriques. Dans ce cours, il n'est pas attendu des étudiants qu'ils mémorisent des formules ; il est attendu qu'ils sélectionnent des données appropriées à la question de recherche, les analysent dans l'environnement R, interprètent les résultats de manière critique, mènent un processus d'analyse commun au sein d'un groupe et rapportent ce processus de manière reproductible. Les étudiants peuvent recourir à des outils d'IA pour l'écriture du code, le débogage et le développement de l'interprétation, mais ils sont tenus d'expliquer et de défendre chaque étape d'analyse utilisée.
Contenus	Le cours commence par la mise en place d'un processus d'analyse de données reproductible dans l'environnement R/RStudio et Quarto/R Markdown. Il aborde ensuite la gestion des données, l'analyse descriptive et la visualisation avec ggplot2. La régression linéaire simple et multiple ainsi que l'inférence statistique sont ensuite traitées. En fin de semestre, l'approche des doubles différences et l'analyse de données de panel sont introduites.
Ressources	1. Wickham, H., Çetinkaya-Rundel, M., & Grolemund, G. (2023). R for Data Science (2nd ed.). Ressource fondamentale en accès libre pour la gestion, le nettoyage, la transformation et la visualisation des données avec R. https://r4ds.hadley.nz/ (anglais) et https://tr.r4ds.hadley.nz/ (turc) 2. Wooldridge, J. M. (2020). Introductory Econometrics: A Modern Approach (7th ed.). Cengage Learning. Pour ce cours, la lecture des 4 premiers chapitres et des chapitres 6, 7, 13 et 14 est suffisante. Pour les jeux de données et le code R utilisés dans l'ouvrage, voir https://justinmshea.github.io/wooldridge/ 3. Quelques base de données ouverts et packages R à utiliser tout au long du cours : • V-Dem Institute : https://www.v-dem.net/data/, package R : vdemdata • World Bank Open Data : https://data.worldbank.org/, packages R : WDI, worldbank. • UCDP/PRIO Armed Conflict Dataset : https://www.prio.org/data/4 • Gapminder : package R : gapminder • World Values Survey : https://www.worldvaluessurvey.org/ • European Social Survey : https://www.europeansocialsurvey.org/, package R : essurvey • Pour d'autres données facilement téléchargeables et explorables via des packages R, voir https://github.com/rOpenGov

Intitulés des Sujets Théoriques

Semaine	Intitulés des Sujets
1	Introduction à l'environnement R/RStudio ; documents dynamiques avec R Markdown/Quarto
2	Transformation/nettoyage des données avec dplyr et tidyr
3	Analyse exploratoire des données – résumés numériques
4	Analyse exploratoire des données – exploration visuelle avec ggplot2
5	Examen Partiel.
6	Régression linéaire simple : causalité, modèle de régression, interprétation des coefficients.
7	Régression linéaire multiple
8	Inférence statistique : erreurs standard, intervalles de confiance, valeurs de la statistique t, p-valeurs.
9	Variables muettes et variables explicatives catégorielles
10	Termes d'interaction (variables produit)
11	Approche de base des doubles différences
12	Analyse de données de panel
13	PRÉSENTATIONS DE POSTERS I
14	PRÉSENTATIONS DE POSTERS II