Business Data Analysis

ECTS : 6

Enseignant responsable : Alexandra DIMOU

Langue du cours : Français

Volume horaire : 36

Description du contenu de l'enseignement :

This course introduces students to fundamental methods in data analysis for business using the statistical software package R. It aims to teach students how to effectively present, analyse, interpret data, and make business decisions. Students will become familiar with the R ecosystem and learn to use R for various data analysis projects. Key topics include numerical and graphical summaries of data, univariate and bivariate
analysis, analysis of variance, regression, principal components, factor analysis, and cluster analysis. Practical application is emphasized through hands-on sessions in a computer lab.
The course prepares the next generation of managers to collect, manage, and analyse data to solve the most pressing challenges facing businesses and institutions. This course is on Moodle.
1 Introduction to R-software - Descriptive statistics
2 Sampling and statistical inference (review)
3 Analysing relationships among two categorical variables
4 t.test and ANOVA
5 Correlation and Simple regression
6 Multiple regression
7 ANOVA and ANCOVA models
8 Logistic regression
9 Principal components analysis
10 Correspondence analysis
11 Clustering
12 **Final session assessment**

Compétences à acquérir :

By the end of the course, students will be able to:

- Describe and present data using appropriate statistical and graphical methods;
- Summarize different types of variables and their distributions;
- Analyze relationships between variables through bivariate analysis;
- Conduct regression analysis and make predictions;
- Perform multivariate analyses, including principal component analysis and correspondence analysis;
- Execute cluster analysis to identify patterns within data;
- Confidently navigate and utilize the R environment for data analysis tasks;
- Apply advanced R features and packages for data visualization and manipulation;
- Work effectively in teams and communicate data analysis results clearly.

Pré-requis obligatoires

Basic knowledge of probability and statistics (sampling, estimation, testing) is required. Familiarity with any programming language is beneficial but not mandatory.

Mode de contrôle des connaissances :

Active participation is crucial for this course. Students will engage in individual and group work on real datasets using R. Assessment consists of:
- Class Participation and Homework (40%): Regular attendance, active engagement in lab sessions, and completion of homework assignments.
- Final Exam (60%): A comprehensive, practical exam conducted on the computer, assessing students' ability to apply course concepts using R. Closed book exam.
The numerical grade distribution will dictate the final grade. The passing grade for a course is 10/20.

Attendance

Attendance is mandatory. Students are expected to attend all classes, arrive on time, and stay for the entire session. Repeated absences or lateness may affect the final grade.

Class Participation

Active participation is encouraged, as it contributes to making classes more engaging and instructive. Students are expected to come prepared and contribute thoughtfully to discussions. When participation is part of the course assessment, it is evaluated based on the quality of contributions rather than their quantity.

Exam Policy

Students are not allowed to bring any materials into exams, except those explicitly authorized by the instructor. Unexcused absences from exams or failure to submit assigned cases will result in a grade of zero when calculating final averages. All exams must be submitted at the end of the examination period.

Communication and Grading

All questions or concerns regarding grading or course policies must follow the official procedures. No direct negotiation with instructors about grades or assessments is permitted. Be aware of the rules in Université Paris Dauphine about plagiarism and cheating during exams. All work turned in for this course must be your own work, or that of your own group. Working as part of a group implies that you are an active participant and fully contributed to the output produced by that group.

Bibliographie, lectures recommandées :

- Wickham, H., & Grolemund, G. (2016). R for Data Science: Import, Tidy, Transform, Visualize, and Model Data.O'Reilly Media.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning: With Applications in R. Springer.
- Field, A., Miles, J., & Field, Z. (2012). Discovering Statistics Using R. Sage Publications.
- Kabacoff, R. I. (2015). R in Action: Data Analysis and Graphics with R. Manning Publications.
- Peng, R. D. (2016). R Programming for Data Science. Leanpub.
- Healy, K. (2018). Data Visualization: A Practical Introduction. Princeton University Press.
- Nina Zumel and John Mount (2019), Practical Data Science with R.
- Heumann (2016), Introduction to Statistics and Data Analysis, Springer, 455 pages.
- J.L. Devore (2011), Introduction to Statistics and Data Analysis, 4th Edition, 944 pages.
- C. Judd (2017), Data Analysis, New Edition, 366 pages.
- Daniel J. Denis (2020), Univariate, Bivariate, and Multivariate Statistics Using R, Wiley, 384 pages.
- Mustapha Abiodun Akinkunmi (2019), Business Statistics with Solutions in R, De Gruyter, 276 pages.
- Christian Heumann, Michael Schomaker, et al. (2017), Introduction to Statistics and Data Analysis: With Exercises, Solutions and Applications in R, Springer.

Alexandra Schaffar is a Professor of Economics at the University of Toulon, with an agregation in mathematics. Her research focuses on econometrics and urban and regional economics. She currently serves as Head of the French Institute for Economic Development, is a former Vice-President of the French speaking section of the International Regional Science Association and is Editor of the journal Région et
Développement. Since 2018, Alexandra Schaffar has been teaching Business Data Analysis at the University Paris-Dauphine.