SYLLABUS

Course title	**DATA ANALYSIS 2** **- PATTERN DISCOVERY AND REGRESSION ANALYSIS**
Instructor	**Tímea Laura Molnár**
Email	MolnarTL@ceu.edu
Office	A509 on QS 3rd floor
Office Hours	Wednesdays, 5pm-5.45pm
Teaching Assistant	Costanza De Acutis
Email	De-Acutis_Costanza@phd.ceu.edu

(Optional) TA Discussion Group//Office Hours: 5 Mondays on Nov6 (C323),13,20,27,Dec11 (D108)

Credits	2 US credits (4 ECTS credits)
Module	***Mandatory in:*** Master of Arts in Economic Policy in Global Markets
Term and Time	Fall 2 term in AY 2023-2024, Tue and Thu 13.30-15.20 (with 10 mins break after 50 mins); venue either B505 or C322/C323, as indicated in your schedule
Course level	Master
Prerequisites	Data Analysis 1

1. COURSE DESCRIPTION

Uncovering patterns in the data can be an important goal in itself, and it is the prerequisite to establishing cause and effect and carrying out predictions. The course starts with simple regression analysis, the method that compares expected y for different values of x to learn the patterns of mean dependence between the two variables. It discusses nonparametric regressions and focuses on the linear regression. It builds on simple linear regression and goes on to enriching it with nonlinear functional forms, generalizing from a particular dataset to other data it represents, adding more explanatory variables, *etc.* We also cover regression analysis for binary dependent variables, as well as nonlinear models such as logit and probit. We will discuss selected case studies in lectures, and interpretation and coding solutions will both be discussed using STATA.

2. LEARNING OUTCOMES

By successfully completing the course the students will be able to:

- Successfully formulate research questions that are answerable by empirical analysis.
- Produce meaningful descriptive statistics and informative graphs.
- Carry out simple regression analysis.
- Discuss and interpret results, understand validity and constraints.
- Present empirical analysis and write short reports with data.

3. READING LIST

Lecture notes and slides will be provided by the instructor.

Optional Textbook (selected case studies will be covered from this textbook, along with the underlying code): Békés - Kézdi: Data Analysis for Business, Economics and Policy, Cambridge University Press, Chapters 7-11. Data and codes will be provided. Accessing the textbook: The CEU Library has 10 hard copies on site in Vienna. In addition, the library has paid for basic (read-only) online access, available through the publisher's website: https://www.cambridge.org/highereducation/books/data-analysis-for-business-economics-and-policy/D67A1B0B56176D6D6A92E27F3F82AA20#overview

IMPORTANT: Online access will work automatically only on the CEU campuses; if you are not on campus, you must connect through the CEU VPN.

Data sets and other supporting materials available through the authors' website: https://gabors-data-analysis.com/

4. ASSESSMENT, ATTENDANCE AND GRADING POLICY

Grading will be based on the total score out of 100, in line with CEU's standard grading guidelines, as in:

Homework (HW) Assignments (see schedule below) **50%** (5 in total, each worth 10%)
Final Exam: on Wednesday, December 13, 2023, 9am-noon: 50%
Students have to score at least 50% for *both* the HW Assignments and the Final Exam to pass the course.

- Regular class attendance is a precondition for course completion. Students who miss 3 or more classes, either excused or unexcused, cannot receive a passing grade. In line with the departmental no-phone-policy regulation, no cell phones are allowed to be used in class (and computers/tablets only for note-taking if needed).
- The Teaching Assistant (TA) for the course is PhD student Costanza De Acutis (email: De-Acutis_Costanza@phd.ceu.edu). The TA will hold TA Discussion Groups on selected Monday afternoons on which attendance is not compulsory but highly recommended, as the solution of the assignments will be presented and discussed, as well as students can ask the TA prior HW submission for hints and clarifications.
- Questions regarding the assignments, the assignment solutions and grading should be directed to the TA.
- All teaching materials (lecture notes, assignments, assignment solutions, more practice problems and other materials) will be uploaded to Moodle.
- Students are allowed to work in study groups (discussion groups) on homework assignments; but everyone is required to submit her/his/their own assignment spelling out the solution on her/his/their own; copying from other students will lead to zero grade for that HW assignment for all parties involved. Late assignments receive 0 credit.

	Tue Lecture	Thu Lecture		TA session
week1	31-Oct	2-Nov	post HW1 on Nov2, due on Nov6 midnight	6-Nov
week2	7-Nov	9-Nov	post HW2 on Nov9, due on Nov13 midnight	13-Nov
week3	14-Nov	16-Nov	post HW3 on Nov16, due on Nov20 midnight	20-Nov
week4	21-Nov	23-Nov	post HW4 on Nov23, due on Nov27 midnight	27-Nov
week5	28-Nov	30-Nov		
week6	5-Dec	7-Dec	post HW5 on Dec7, due on Dec10 midnight	11-Dec

5. TOPIC OUTLINE AND SCHEDULE

1. Simple regression analysis (conditional expectations, correlations, non-parametric regression, linear regression, OLS, predicted values and residuals, regression and causality, identification)
2. Complicated patterns (taking log of variables, piecewise linear splines, polynomials, measurement error)
3. Generalizing results of regression analysis (standard error, confidence interval, prediction interval, testing, external validity)
4. Multiple linear regression (linear regression mechanics, binary and other qualitative right-handside variables, interactions, *ceteris paribus* vs. conditioning in multiple regression)
5. Probability models (linear probability models, logit and probit, marginal differences)

6. SHORT BIO OF THE INSTRUCTOR

Tímea Laura Molnár is an Assistant Professor in Economics at CEU since 2020. She obtained her MA in Economics at CEU, and her PhD in Economics at the University of British Columbia. Her current research studies parents' intra-household time allocation decisions, focusing on parental quality time, and the implications for early childhood development, as well as the effect of academic redshirting on student and mental health outcomes, the effect of transitory health shocks on labor reallocation across firms, and the long-term effects of childcare availability on the gender wage gap, mothers' well-being and child development. In the era of COVID-19, she studied the impact of school reopenings and labor market risks on parental labor supply, focusing on the extent to which COVID-19 exacerbates gender inequalities via unequal obligations due to school closures and disproportionate employment losses in virus-exposed occupations. In the past she studied how closely private insurers' payment schedules follow that of Medicare's. She has published in the Journal of Health Economics and the Canadian Journal of Economics.