Syllabus

CEU : Department
of Economics
and Business

CENTRAL
EUROPEAN
UNIVERSITY

Data Analysis 2: Finding Patterns with Regressions – Business Analytics track

Instructor:	Alice Kügler
Email:	KueglerA@ceu.edu
Office hours:	by appointment
Credits:	2 US credits (4 ECTS credits)
Term:	Fall 2023; Mondays Week 7-12, 13:30-15:10 and 15:40-17:20
Course level:	Master's
Prerequisites:	Mathematics and Informatics Pre-session for Business Analytics;
	Data Analysis 1: Exploration – Business Analytics track
Course drop:	The course can be dropped free of charge 24 hours after the first session. After this date drop is possible until the course is halfway over (late drop fee applies). No changes are allowed past that date.

1. COURSE DESCRIPTION

Content. The course starts with simple regression analysis, the method that compares expected y for different values of x to learn the patterns of association 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 time series data, binary dependent variables, as well as nonlinear models such as logit and probit.

The seminars will focus on selected case studies introduced in lectures. In classes, interpretation and coding solutions will both be discussed using Python.

Relevance. Uncovering patterns in the data can be an important goal in itself, and it is the prerequisite to establishing cause and effect, and for carrying out predictions.

2. LEARNING OUTCOMES

Key outcomes. By the end of the course, 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

Other outcomes. The course will also help develop skills in the following areas:

Learning Area	Learning Outcome
Critical Thinking	Ability to discuss and interpret results
Quantitative Reasoning	Ability to present empirical analysis and write short reports with data
Technology Skills	
Interpersonal Communication Skills	Ability to work on assignments in groups
Management Knowledge and Skills	
Cultural Sensitivity and Diversity	
Ethics and Social Responsibility	Learn about ethical and legal issues of data collection

3. READING LIST

Textbook
- **BK**: Békés, Gábor and Gábor Kézdi, "Data Analysis for Business, Economics, and Policy". Cambridge University Press, May 2021: hardcover, paperback and e-book.
  [textbook website, Cambridge UP, Amazon]
- Data Analysis 2 covers Chapters 7-12 including all under the hood sections

Optional reading
- Further reading suggestions are contained at the end of the textbook chapters.

Data and code

Data and code are already available, see the information on the textbook website: https://gabors-data-analysis.com/data-and-code/. In particular:

1. Data is available from an online data repository: https://osf.io/7epdj/
2. Code is available from the textbook github repository: https://github.com/gabors-data-analysis/da_case_studies

4. TEACHING METHOD AND LEARNING ACTIVITIES

The course will involve lectures and seminars.

- Lectures focus on theory, while seminars will focus more on the process of analysis via case studies.
- Learning objectives will be achieved by students working through the material.

Students may use R or Python to submit assignments. Codes for the textbook are available in R and Python. Support is guaranteed in Python.

5. ASSESSMENT

Grading will be based on the total score out of 100, in line with CEU Department of Economics and Business grading guidelines. In particular:

a. The median student can expect to get a B+.
b. Probably not more than 1/3 of the students can expect to get an A or A-.
c. To pass, students will need to get at least 50% of the overall grade AND at least 50% of the exam. Failure to do so, will yield a Fail grade.

The final grade is based on:
- start-of-the-class quizzes [10%]
- assignments [40%]
- a final closed-book exam [50%]

6. TECHNICAL REQUIREMENTS

A computer with R or Python.

7. TOPIC OUTLINE AND SCHEDULE

Session	Topics	Readings
1	Week 1: Simple regression analysis Topics: non-parametric regression, linear regression, OLS, predicted values and residuals, regression and causality	BK Chapter 7

2	Week 2: Complicated patterns and messy data Topics: taking log and other transformations of variables, piecewise linear splines and polynomials, measurement error in variables, influential observations, using weights	BK Chapter 8
3	Week 3: Generalizing results of regression analysis Topics: standard error, confidence interval, prediction interval, testing, external validity	BK Chapter 9
4	Week 4: Multiple linear regression Topics: linear regression mechanics, binary and other qualitative right-hand-side variables, interactions, ceteris paribus vs. conditioning in multiple regression	BK Chapter 10
5	Week 5: Modelling probabilities Topics: linear probability, logit and probit, marginal differences, goodness of fit, calibration	BK Chapter 11
6	Week 6: Regression with time series data Topics: trends, seasonality, leads and lags, serial correlation, cumulative association, appropriate standard errors	BK Chapter 12

8. SHORT BIO OF THE INSTRUCTOR

Alice Kügler joined the Department of Economics and Business at the Central European University as an Assistant Professor in 2021. Before joining CEU, Alice was a Postdoctoral Research Fellow at University College London. She completed her PhD in Economics at the University of Cambridge and holds an MSc in Economics from Pompeu Fabra University. Alice has previously worked at the World Bank and at the Ministry of Finance of Mozambique. Her research interests are Labor Economics, Public Economics, Inequality and Innovation.