This Syllabus Template is based on New York State Education Department and AQ Austria requirements

COURSE SYLLABUS

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

Instructor:
Ágoston Reguly
Department of Economics
Central European University
Fall, 2024
Master of Science
2 US Credits (4 ECTS Credits)
Mandatory for Business Analytics and, elective for Finance
Business Analytics and Finance
Mathematics and Informatics Pre-session, Data Analysis 1: Exploration
Course e-learning site: ceulearning.ceu.edu
Office hours: by appointment (regulya@ceu.edu)

Course Description
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.

Learning Outcomes

By the end of this 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
- ✓ Ability to work on assignments in groups

Learning activities and student learning

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 submit assignments with the help of Python. Codes for the textbook are available in R and Python.

Course Requirements

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

- *The median student can expect to get a B+.*
- *Probably not more than 1/3 of the students can expect to get an A or A-.*
- *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:

(1) Start of the class quizzes (10% of the final grade). Students will fill out a short quiz on ceulearning's site, which asks questions from last lectures

(2) Assignment 1 (10% of the final grade). Students has to submit a 1 page report with analyzing a specific regression task.

(3) Mid-term exam (20% of the final grade). Closed-book exam from single linear regression theory.

(4) Assignment 2 (20% of the final grade). Students submit a complete analysis in pairs on a specific topic.

(5) Final exam (40% of the final grade). Closed-book exam from all materials

COURSE SCHEDULE

Week 1: Simple regression analysis and non-linear transformations

Topics: non-parametric regression, linear regression, OLS, predicted values and residuals, regression and causality. Taking log and other transformations of variables, piecewise linear splines and polynomials. Readings: BK Chapter 7-8

Week 2: Messy data and generalizing regression results

Topics: Measurement error in variables, influential observations, using weights. Standard error, confidence interval, prediction interval, testing, external validity. Readings: BK Chapter 8-9

Week 3: Multiple linear regression

Topics: linear regression mechanics, binary and other qualitative right-hand-side variables, interactions, ceteris paribus vs. conditioning in multiple regression. Readings: BK Chapter 10

Week 4: Modelling probabilities and introduction to time series data

Topics: Probability models: linear probability, logit and probit, marginal differences, goodness of fit, calibration. Time series analysis: trends, seasonality, leads and lags, serial correlation, cumulative association, appropriate standard errors. Readings: BK Chapter 11-12

Reading:

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]
- We will cover Chapters 7-12, including the under the hood sections.

Optional:

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