Syllabus

INTRODUCTION TO R

Instructor	Gergely Daroczi; daroczig@ceu.edu
Office	by appointment
Credits	1 US credit (2 ECTS credits)
Term	Winter 2024-2024
Course level	Master's
Prerequisites	

Course drop 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

This is a hands-on and tech-focused course to get you up to speed using the R ecosystem (installing R, installing packages from CRAN and GitHub, manuals and vignettes, RStudio IDE, running R remotely) and R programming language for data transformations, data visualization, modeling, exploratory data analysis, and publishing HTML and PDF reports.

2. LEARNING OUTCOMES

Key outcomes.

By the end of the course, you will be able to use R to read and write data in different formats, do data transformations (e.g. filtering, aggregation, reshaping between long and wide tables, creating new variables), data visualizations (e.g. scatterplots, line plots, histograms, barcharts, boxplots), and generating reports in various formats (e.g. HTML and PDF).

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

Learning Area	Learning Outcome
Critical Thinking	Finding the appropriate tool for data processing.
Quantitative Reasoning	
Technology Skills	Using R.
Interpersonal Communication Skills	
Management Knowledge and Skills	
Cultural Sensitivity and Diversity	
Ethics and Social Responsibility	

3. READING LIST

Class materials will be available on GitHub.

Databases. The CEU Library boasts a range of databases covering financial and company data, market and industry reports, global news and more. For a full list of databases visit the [CEU Library](#).
- Refinitiv (Thomson Reuters) Eikon for Students + Datastream/Thomson ONE
  – Eikon: Platform used by finance practitioners including market traders to monitor and analyze financial information. Information, analytics and news on all major financial markets including real-time pricing data, financial research, global financial news and commentary, financial estimates, fundamentals analysis, visual analysis through charting. Import/export from Excel.
  – Datastream: Range of economic, securities and company financial data. Excel add-in.
  – Thomson ONE: Global overviews on 55,000 public companies, one million private companies. Reuters News, ownership, deals, private equity, key ratios, company filings, officers and directors. Investext analyst reports, active and historical research from 1,600 independent research firms, brokerages, investment banks.
- Standard & Poor's Capital IQ
  – Web and Excel-based platform combining deep global company information, credit ratings and research, and market research with powerful tools for risk assessments. Real-time and historical information on markets, industries, companies, transactions and people. Tearsheet data.
- Lexis Nexis Academic
  – Global database of news, business, legal and other sources. Full text of 350 newspapers, 300 magazines and journals, 600 newsletters. Wire services including Associated Press, Business Wire and PR Newswire. Company financial information, market research, industry reports.

4. TEACHING METHOD AND LEARNING ACTIVITIES

The course will involve coding sessions with demos and exercises.
Learning objectives will be achieved through actively taking part in the in-class exercises and solving homework and the final take-home assignment.

5. ASSESSMENT

40% homework and 60% final project.

Grading Policy

Students shall not miss more than 1 day of classes, failing to do so will yield an administrative fail grade. To pass, students will need to get at least 50% of the homework AND at least 50% of the final project.

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

6. TECHNICAL REQUIREMENTS

Laptop with R, RStudio and git installed is required in the class and for the take-home assignments. Installation steps will be shared on GitHub.

7. TOPIC OUTLINE AND SCHEDULE

Session	Topics	Readings
1	R ecosystem (base R, packages, RStudio), data frames, basic plots, reading CSV files, intro to modeling	Shared on GitHub.
2	Advanced plotting with ggplot2, data transformations using data.table	Shared on GitHub.
3	Publication-ready reports using R markdown, modeling	Shared on GitHub.

8. SHORT BIO OF THE INSTRUCTOR

Gergely Daroczi has a PhD in Sociology, 15 years of experience with R, founder of the Hungarian R meetup and main organizer of R conferences, authored a book on R and maintains a dozen of R packages, lived and worked in Hungary and USA at market research, fintech, adtech and healthtech companies as a data scientist and engineer both in individual contributor and management roles.