

Syllabus

Mastering R Skills

Instructor: Gergely Daroczi

Email: daroczig@ceu.edu

Office hours: by appointment

Credits: 1 US credit (2 ECTS credits)

Term: Spring 2022-2023

Course level: Master's

Prerequisites: Data Engineering 3: Using R in Production

Course drop: Spring term course registration modification is not allowed after April 25th, 2023.
Course may not be dropped past that date.

Audit: A maximum nr of 2 auditing students are allowed to take this course. Prerequisites are mandatory for auditing students as well. Please consult the instructor for auditing requirements.

1. COURSE DESCRIPTION

This class focuses on actual programming questions and problems related to R -- that will help you writing more elegant, maintainable, reliable and performant R applications, covering topics like

- Writing and documenting R functions
- Debugging and profiling R scripts and functions
- Refactoring R scripts
- Developing R packages
- Writing unit tests and running those in an automated way
- Logging, database connections

2. LEARNING OUTCOMES

Key outcomes:

By the end of this course, you will have a good understanding on how to structure and deploy R scripts and packages for production.

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

Learning Area	Learning Outcome
Critical Thinking	The importance of security, logging and initial design of applications and scripts.
Quantitative Reasoning	
Technology Skills	Interacting with APIs, profiling and benchmarking R code, database connections, writing R packages.
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 METHODS 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

Weekly homeworks (50%) and final take-home assignment (50%)

To pass, students will need to get at least 50% of the overall grade AND at least 50% of the final take-home assignment. Failure to do so, will yield a Fail grade.

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

6. TECHNICAL/LAPTOP REQUIREMENT

Laptop with the most recent version of R and RStudio is required in the class and for the take-home assignments as well. Installation steps will be shared on GitHub.

7. TOPIC OUTLINE AND SCHEDULE

Session	Topics	Readings
1	API connections, data transformations.	Shared on GitHub.
2	Creating an R package. Profiling and benchmarks.	Shared on GitHub.
3	Reporting in Excel and Google Spreadsheets.	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.