Coding 1: Data management and analysis with R

Ágoston Reguly

2021/22

This course

This course introduces to R statistical programming.

The aim of this course is to teach you how to carry out a complete, reproducible data analysis project.

Agenda

We are going to cover:

1. How R and coding in general works.
2. Importing and exporting data.
3. Data wrangling (management, cleaning, set ready to analyse).
4. Visualization and carry out explanatory data analysis.
5. Creating your own functions, loops and simulations.
6. How to implement Data Analysis theory in practice.
7. How to create nice reports and presentations with RMarkdown.
8. Some advanced analysis tool which might not covered in DA 2: spatial and financial time-series econometrics.

Office hours

Instructor: Ágoston Reguly (reguly_agoston@phd.ceu.edu)

- *Office hours*: Wednesday 11:00-12:00, online or in-person.
- Weekends - I am not checking my email.

Further help:

- Mihály Orsós (orsosm@ceu.edu) and
- Viktória Mészáros (meszaros_viktoria@alumni.ceu.edu)
- Instructors/mentors for R (Coding Practice with R)
- That course is suppose to help you if stuck with materials in this course.
- Later Mihály will teach Coding 2: Web scraping with R

Course Material

The core course material is what we cover during the classes.

- Codes are going to be uploaded to github:
  - https://github.com/CEU-Economics-and-Business/ECBS-5208-Coding-1-Business-Analytics
- We are going to cover the tidyverse approach.
  - Note that there are many possible solutions in coding!

There is a great book by Garrett Grolemund and Hadley Wickham:

- R for data science
  - You can look up and enrich your knowledge there!
  - It is highly recommended to practice through these exercises!

Assessment I

- Assignments (50%)
  - 6 short assignments through the semester (up to 35%)
    - After 1-5 and 10th session, each worth 7%, thus one can neglect one of them without loosing any points.
    - Form: upload to github/github classroom.
    - Deadline: before next class.
  - Team project (15%): write a descriptive report for your dataset.
    - Based on the created dataset from Data Analysis 1
    - Deadline: 31 October Sunday 23.55, github repo.
- Demonstrate useful resources/materials (up to 10%)
  - Reference your resource/material on slack channel
  - Demonstrate it in class (1-3 min)
  - Package with functions - 5%
  - Alternative solution to problems - 2.5%

Assessment II

- Take home examination (50%)
  - Carry out a complete data analysis project from scratch.
  - Joint project with Data Analysis 2.
  - Coding part will be evaluated only in this class.
  - Deadlines - TBA

The points sums to 110%, so you can lose 10% without it affecting your grade at all. There will be no paper based exam.

Grading policy

- To pass, you will need to get at least 50% of the overall grade AND at least 50% on the take-home exam.
- Lectures - 75% attendance is required (8/12)
  - In case of online participation: write me with the reason unless there is a departmental change.

Any questions?