

**Winter Term 2022/23****Introduction to Quantitative Methods****Instructor:** Leonce Röth*Time:* Friday 10:00 – 11:30*Room:* **WiSo Gebäude, Hörsaal XXV***First Session:* 25 November 2022**Contact**[leonce.roeth@uni-koeln.de](mailto:leonce.roeth@uni-koeln.de)

Office hours by appointment

Office: IBW Building, Herbert-Lewin-Str. 2

**Note on ILIAS:** There are separate platforms for the lecture and the lab on ILIAS.

- Lecture: Platform for all other material, including the assignments.
- Lab: We only use this for all material strictly related to the lab sessions, meaning the datasets, R scripts and all related files used during the labs.

**Registration**

Registration for exam in KLIPS2 (for Master students). PhD researchers should send an email to [cccp-sekretariat@wiso.uni-koeln.de](mailto:cccp-sekretariat@wiso.uni-koeln.de) instead.

Please also regularly check the CCCP information on teaching on the internet:

<http://www.cccp.uni-koeln.de/en/public/teaching/>

**Course description**

This course is open to MA and PhD students and provides an introduction to quantitative analysis in political science using R. Topics include the essential elements of sampling, statistical inference, the estimation and interpretation of the linear regression model and data visualization. In the last weeks students will have brief introductions to more advanced topics, including time-series and panel data analysis, multilevel models and interaction effects.

Data and empirical applications that we will use in class primarily focuses on political science. Throughout the course, students will become familiar with the statistical programming language R as a means to realize a quantitative analysis. The goal is to prepare students to critically interpret and conduct quantitative studies in political science.

The course will be taught on a weekly basis (first session and Lab: 25<sup>th</sup> of November). Registration takes place via KLIPS.

**Prerequisites**

Students should have a basic understanding of descriptive statistics (i.e., can you calculate an average?). No prior knowledge of R is required.

## Software

The open-source statistical programming language R is particularly suited for carrying out state-of-the-art computer-based simulations and data exercises. It also generates really nice publication-quality graphics. The software runs under a wide array of operating systems. R can be downloaded for free at <http://www.r-project.org/>. Learning R might seem a bit challenging at first, but you will realize that it is incredibly powerful. The lab sessions will be devoted to learning data management and analysis techniques using R.

Students are also strongly encouraged to install RStudio. This is a more user-friendly interface for R with integration to other packages we will use throughout the course. RStudio is available for free at <https://www.rstudio.com/>. Installing R can be problematic sometimes. Students must have installed R (and RStudio) before the first lab session (25<sup>th</sup> of November). If you are running into errors when installing, contact the instructor before the first session.

## Course Materials

The course is structured as an applied introduction to quantitative analysis. Students can choose many of the mandatory readings from one of two sources provided by the instructor:

Gravetter, F. et al. (2021): Essentials of Statistics for the Behavioral Sciences. Australia; United States: Cengage Learning, Inc.

Diez, David, Christopher Barr, and Mine Cetinkaya-Rundel (2015). OpenIntro Statistics, 3rd edition. Available at <https://www.openintro.org>.

***The following materials provide alternative coverage for topics in this course:***

Fox, John (2016). Applied Regression Analysis and Generalized Linear Models, 3rd edition. Los Angeles: Sage.

Fox, John and Harvey S. Weisberg (2019). An R Companion to Applied Regression, 4th edition. Thousand Oaks: Sage.

## Course Requirements

This course will have a portfolio evaluation. It will be comprised of six data analysis projects during the semester. The first five papers should be around 3 pages, while the last around 5 pages. It is mandatory to deliver all assignments to pass the course. For the first five, students will have a week to complete the assignment, and two weeks for the last. You can discuss it with other students, but each one must write and submit their own individual answers. Two (nearly) identical assignments will receive a 0. If you work with someone, please indicate their names in your homework. Assignments must be handed in through Ilias, with the following deadlines and weights towards the final grade:

**Total number of points and final grade**

| Points  | Grade |
|---------|-------|
| 100-95  | 1     |
| 94.5-90 | 1.3   |
| 89.5-85 | 1.7   |
| 84.5-80 | 2     |
| 79.5-75 | 2.3   |
| 74.5-70 | 2.7   |
| 69.5-65 | 3     |
| 64.5-60 | 3.3   |
| 59.5-55 | 3.7   |
| 54.5-50 | 4     |
| 0-49    | 5     |

| Homework     | Posted     | Deadline   | Weight    |
|--------------|------------|------------|-----------|
| Assignment 1 | 25.11.2022 | 02.12.2022 | 15 points |
| Assignment 2 | 02.12.2022 | 09.12.2022 | 15 points |
| Assignment 3 | 09.12.2022 | 13.01.2023 | 15 points |
| Assignment 4 | 13.01.2023 | 20.01.2023 | 15 points |
| Assignment 5 | 20.01.2023 | 27.01.2023 | 15 points |
| Assignment 6 | 27.01.2023 | 06.02.2023 | 25 points |

**Bonus Points**

It is possible to increase your grade by up to 12 points in this course, with two bonus point assignments. Each one will have a maximum grade of six. They will be posted on 9<sup>th</sup> of December and the 13<sup>th</sup> of January, and you will have one week to do each. Bonus points exercises cannot be used a replacement for one of the mandatory assignments. It means that you can only receive bonus points if you submit all six mandatory assignments.

Last, the projects submitted must be written with R Markdown (see below).

**R Markdown | A Note on Professional Presentation**

RStudio comes with a powerful authoring format called R Markdown. R Markdown documents look like a mix of a text document and R code. They enable easy creation of data analysis reports directly from R. Rather than copying and pasting into Word, your report is created automatically. R Markdown combines the core syntax of markdown (an easy-to-write plain text format) with embedded R code chunks that are run so their output can be included in the final document. R Markdown documents are fully reproducible (they can be automatically regenerated whenever underlying R code or data changes). Markdown is simple to use as it enables the use of a syntax like plain-text.

1. To use R Markdown, simply create a new R Markdown document in RStudio. This will load a sample document. Select "Knit PDF" to produce a PDF output file with the write up and the code output.

2. More information on the R Markdown syntax is available here: <http://rmarkdown.rstudio.com/>.

## Schedule

### **Week 1 (25.11.2022): Fundamentals of quantitative analysis and intro to the course**

Gravetter, F. et al. (2021): Essentials of Statistics for the Behavioral Sciences. Australia; United States: Cengage Learning, Inc: chapters 1 - 4.

Or OpenIntro Statistics, chapter 1 and 2.

### **Week 2 (02.12.2022): Distributions, sampling and p-values**

Gravetter, F. et al. (2021): Essentials of Statistics for the Behavioral Sciences. Australia; United States: Cengage Learning, Inc: chapters 7 and 8.

Or OpenIntro Statistics, p. 127-137.

### **Week 3 (09.12.2022): Statistical significance, testing, and confidence intervals**

Gravetter, F. et al. (2021): Essentials of Statistics for the Behavioral Sciences. Australia; United States: Cengage Learning, Inc: chapters 8 (again), 9.

### **Week 4 (17.12.2022): Errors, power, substantive significance, group differences and missings**

Gravetter, F. et al. (2021): Essentials of Statistics for the Behavioral Sciences. Australia; United States: Cengage Learning, Inc:: chapter 9 (again), 10, 15.

Schafer, J. L., and Graham, J. W. (2002). Missing Data: Our View of the State of the Art, *Psychological Methods*, 7(2):147-177.

### **Week 5 (13.01.2023): Correlation, bivariate and multivariable regression**

Gravetter, F. et al. (2021): Essentials of Statistics for the Behavioral Sciences. Australia; United States: Cengage Learning, Inc: chapter 16.

### **Week 6 (20.01.2023): Analysis of categorical data and dichotomous dependent variables**

Gravetter and Wallnau chapter 16 AND OpenIntro Statistics, chapters 6 and 8 (from p. 386).

### **Week 7 (27.01.2023): Nested observations – cross-sectional time-series and multi-level models**

Luke, D. A. (2004). *Multilevel Modeling*. Thousand Oaks, CA: Sage. Chapter 1 (pp. 1-7).

Cowpertwait, P. S. P., and Metcalfe, A. V. (2009). *Introductory Time Series with R*. New York: Springer. Chapter 1 (pp. 1-25).

Finkel, S. E. (1995). *Causal Analysis with Panel Data*. Thousand Oaks, CA: Sage. Chapter 1 (Modeling Change with Panel Data).

### **Week 8 (03.02.2023) | Interactions, Polynomials, and Assumptions**

Brambor, T., Clark, W. R., & Golder, M. (2006). Understanding interaction models: Improving empirical analyses. *Political analysis*, 14(1), 63-82.