

Multivariate Analyses 2021:

Welcome to the course!

Thomas Gschwend, Oliver Rittmann, Viktoriia Semenova & David Grundmanns

September 6, 2021

Table of contents

1. About Us
2. About You
3. About This Course
4. Learning R
5. Our Support Infrastructure

About Us



Thomas Gschwend

Professor for Quantitative Methods in the Social Sciences

📍 A5, 6, B 220
📞 +49 621 181 2413
✉️ gschwend@uni-mannheim.de
🌐 methods.sowi.uni-mannheim.de

Substantive Research Interests

- Comparative Political Behavior
- Judicial Politics
- Electoral Systems

Methodological Expertise

- Statistical Modeling
- Experimental Design
- Simulations



Oliver Rittmann

PhD Student, CDSS

- 📍 A5, 6, B 226
- 📞 +49 621 181 2542
- ✉️ orittman@mail.uni-mannheim.de
- 🌐 methods.sowi.uni-mannheim.de

Substantive Research Interests

- Parliamentary Research
- Legislative Studies
- Political Representation

Methodological Expertise

- Statistical Modeling
- Automated Video- and Text Analysis
- Sample Selection and Causal Inference



Viktoriia Semenova

PhD Student, CDSS

📍 B6, 30-32, 250
✉️ semenova@uni-mannheim.de
🌐 methods.sowi.uni-mannheim.de

Substantive Research Interests

- Authoritarian Politics
- Information Manipulation
- Election Fraud

Methodological Expertise

- Statistical Modeling & Bayesian Stats
- Automated Text Analysis
- Data Visualization



David Grundmanns

PhD Student, CDSS

- 📍 A5, 6, A 217
- ✉ dgrundma@mail.uni-mannheim.de
- 🌐 methods.sowi.uni-mannheim.de

Substantive Research Interests

- Judicial Politics
- Separation of Powers
- Public Opinion

Methodological Expertise

- Statistical Modeling & Bayesian Stats
- Automated Text Analysis
- Formal Modeling

About You

About You

1. Even though we are slowly getting back onto the campus, the circumstances make it harder for us to get to know **you**.
 - Please fill out the survey on ILIAS (you will receive a mail later today).
 - Attend our virtual office hours.
2. The situation also makes it harder for you to get to know **each other**.
 - Not everyone can be in Mannheim or come to the campus these days. Please look out for one another! Integrate those of you who cannot meet in person, we all know what isolation feels like.
 - Help each other in the student forum on ILIAS and to collaborate and chat use the course Slack channel.
 - Support each other while collaboratively working on our weekly assignments.
3. If the current situation and the way we handle it makes it hard for you to study effectively:
 - Talk to us about your concerns and let us know if/how we can help.

About This Course

Methods sequence in our political science graduate program:

- M.A.: Quantitative Methods (this class), required
- M.A.: Advanced Quantitative Methods (in spring), optional
- CDSS: Additional advanced classes focusing on specific techniques

Goals and Objectives

- Passive
 - Understand logic of statistical inference
 - Recognize and understand basic linear models
 - Sound critical judgment of quantitative studies of political problems
 - Interpret quantitative analyses in published work
 - Be prepared for an advanced quantitative methods course (AQM) in spring
- Active
 - Use basic linear models to analyze political problems
 - Understand statistical results and interpret them correctly
 - Present results in an understandable way, preferably using effective visualizations
 - Use a powerful and free software to do all this: **R**

Roadmap

- Understand and model stochastic processes
- Understand statistical inference
- Implement it mathematically and learn how to estimate it
 - OLS
 - Maximum Likelihood
- Implement it using statistical software
 - Learn how to use R
 - Basic programming skills

- Quantitative Methods = Lecture + Lab
 - Lecture: “Multivariate Analyses”
 - Lab session: “Tutorial Multivariate Analyses”
- The lecture introduces the methods and models, the labs apply them to data.
- You can then practice and deepen your understanding in the weekly homework assignments.

Expectations

- This is a time-intensive course with a high work load. Really!
- There are no prerequisites. We start from the beginning, but will cover substantial ground rather quickly; both in statistical theory and programming.
- Practice, practice, practice: Work together on weekly homework assignments, ask your instructors.
- Good news: Mastering quantitative methods is essential for the academic and non-academic job market.
- Even if you had a quantitative methods course during your undergraduate study, you will learn more advanced material here.

You will receive a grade for the lecture, and a pass/fail for the tutorial

- Weekly Homework Assignments: pass/fail
 - Short problems sets: Apply methods to data and gain proficiency with statistical software.
 - Homework for Weeks 1 and 2 must be completed individually. After that, you will work in groups of 2-3 students (preferably 3).
 - On each group submission, please indicate about how much each group member contributed towards it (in %).
 - **Deadline: Tuesdays, 23h59.** Late submissions will not be accepted.
- Midterm Exam: 1/2
 - 90min open-book take-home exam on **3 November 2021**. No collaboration allowed.
 - We will provide sample questions beforehand.
- Data Analysis Essay: 1/2
 - Final paper (around 2,000 words). No collaboration allowed.
 - Application of statistical techniques to a substantive problem.
 - **Deadline: 15 December 2021, 10h00.** Late submissions will not be accepted.

Sessions

1. 8 September 2021: Introduction. Visualizing Data.
2. 15 September 2021: Fundamentals of Probability.
3. 22 September 2021: Sampling & Statistical Inference.
4. 29 September 2021: Linear Regression: Basics & Hypothesis Testing.
5. 6 October 2021: Linear Regression: Statistical Control & Causality.
6. 13 October 2021: Linear Regression: Dummies & Interactions.
7. 20 October 2021: Linear Regression: Interpreting Substantive Effects via the Simulation Method.
8. 27 October 2021: Linear Regression: Diagnostics.
9. 3 November 2021: Mid-Term Exam. No homework assignment due.
10. 10 November 2021: Non-linear Probability Models - The Likelihood Theory of Statistical Inference.
11. 17 November 2021: Binary Data.
12. 24 November 2021: Count Data.
13. 1 December 2021: Data Essay Q&A. Semester Wrap-Up.

Basics:

- Lecture
 - The **first lecture** on Wednesday (Sep 8) will be **online** on Zoom.
 - **From week 2 on** (starting on Sep 15), the lecture is planned to be **in person**. The lectures will be either streamed live or recordings will be made available online.
- Lab Sessions
 - The **first sequence of lab sessions** (Sep 9, Sep 13, Sep 14) will be **online** on Zoom.
 - **From week 2 on** (starting on Sep 15), **two of the three lab sessions will be in person**, one lab session will be online on Zoom. We will determine which of the three lab sessions will be online after surveying students in week 1.
- Office Hours
 - All office hours will be **online** on Zoom.

Learning R

What is R?

- R is not only a statistical software package but also a statistical programming language.
- R is probably the most versatile statistical tool out there.
- It can do all the things other programs (like STATA, SPSS etc.) can do, but way more (and better).
- It is very flexible, and can also be used beyond traditional statistics, for example for webscraping, machine learning or deep learning.

- Learning R is like learning a new language.
 - Vocabulary (=commands)
 - Grammar (=structure of commands)
 - Practice, Practice, Practice!
- **Very steep learning curve**, may be overwhelming at first.
- Has its quirks you have to get used to.
- It will be hard work, and frustrating at times.
- We are here to help you to navigate around those cliffs.

Why R?

High investment means high returns

- Learning R will already pay off during your Master studies, in term papers and for your thesis.
- Used e.g. in
 - all fields of academia, from biology to economics
 - Wallstreet
 - NY Times data journalists [<https://www.nytimes.com/2019/11/13/technology/personaltech/data-journalism-economics.html>]
- Knowing R will not only **help you in academia**, but will **pay off on the job market!**
- At one point in your studies/academic work you probably have to learn it anyways. Do it now when you have the time and support!



Companies that use R for Analytics



Figure 1: List of Companies that use R for Analytics

Why R?



Figure 2: Work by a former QM student.

<https://projekte.sueddeutsche.de/artikel/politik/europas-parteien-in-der-analyse-e894574> & <https://yannikbuhl.netlify.com>

Why R?

R is open-source.

- Developed from the commercial S language by two statisticians from New Zealand in the early 1990s.
- Today, the R development core team maintains “the engine”.
- R has a vibrant community, that is still growing exponentially.
- Thousands of user/expert-generated, community-reviewed packages.
- Works very well together with other open-source programs like \LaTeX .
- You get everything for free!

Why R?

R is more than just a statistical software package.

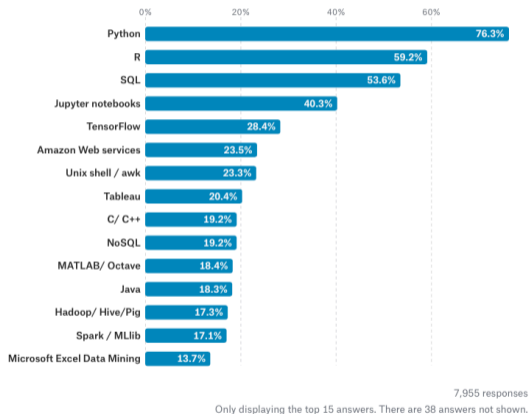


Figure 3: What tools are used at work? <https://www.kaggle.com/surveys/2017>

Our Support Infrastructure

Our Support Infrastructure

The learning curve... Looks frightening, but we will help you to get up there!

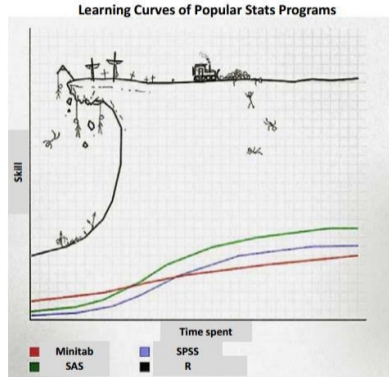


Figure 4: Source: Twitter/@rogierK

Our Support Infrastructure

Seek help from your colleagues, us, and the R community:

- Use the Forum for this course on ILIAS for your specific problems.
- **Help your colleagues.** You will learn a lot in the process of helping.
- Support each other when working on collaborative assignments. We have created a Slack Workspace to facilitate communication among group members.
- Use the R community's online resources, use **Stack Overflow**, use **Google**.
- We are always happy to help. Come to our virtual **office hours**.
 - From our experience we can best help if you come with a specific question in mind.
 - So if you want to attend our office hours, **please send us an email by 17h00 on Sunday**, with **at least one specific question** you want to have answered.
 - Given the number of people in the class, we might need to restrict the time we can offer for each student.

We set up a homepage where you can find all the information and material that you need

- `qm2021.netlify.app`

Our Support Infrastructure

- Thomas Gschwend's email: `gschwend@uni-mannheim.de`
- Oliver Rittmann's email: `orittman@mail.uni-mannheim.de`
- Viktoriia Semenova's email: `vsemenov@mail.uni-mannheim.de`
- David M. Grundmanns's email: `dgrundma@mail.uni-mannheim.de`
- Thomas Gschwend's office hours (Lecture): Tue, 13h30-14h30
- Oliver Rittmann's, Viktoriia Semenova's & David M. Grundmanns's office hours (Labs):
Mon, 15h30-17h00