u^b

_b_
UNIVERSITÄT BERN

Doctoral Program Brain and Behavioral Sciences

University of Bern
Department of Psychology
Fabrikstrasse 8
3012 Bern

Fall Courses 2024

„Data Wrangling and Project Management"
(1-day workshop – 1 ECTS)

Friday, 1st of November 2024 | Room 002, vonRoll, Fabrikstrasse 2e

09:00 – 17:00h	Dr. Ian Hussey, University of Bern

„Simulated Data and Monte-Carlo Simulation Studies"
(2-day workshop – 2 ECTS)

Friday, 8th of November 2024 | Room 004, vonRoll, Fabrikstrasse 2e

09:00 – 17:00h	Dr. Ian Hussey, University of Bern

Friday, 15th of November 2024 | Room 003, vonRoll, Fabrikstrasse 2e

09:00-17:00h	Dr. Ian Hussey, University of Bern

Both courses can be attended independently. These courses are organized for PhD students in the Doctoral Program Brain and Behavioral Sciences. We also welcome other Ph.D. students and Postdocs outside the Doctoral Program, yet students enrolled in the Doctoral Program Brain and Behavioral Sciences have priority for registration. **The deadline for registration is the 16th of October, 2024.**

More information

Registration

Data Wrangling & Project Management

Effective data analysis goes far beyond simply running statistical tests. This one-day workshop is designed to equip participants with foundational skills in key areas of data science: data wrangling, data visualization, reproducible reporting, and digital project management. Using the R and the tidyverse packages, participants will learn how to organise their data management process, ensuring that their data is well-organized, their analyses are transparent, and their outputs are reproducible, sharable, and robust to errors. By learning these skills early on in their graduate training, participants can avoid many common pitfalls that lead to wasted time, errors, and poor reproducibility.

Key Takeaways:
- Data Wrangling: Introduction to the skills to clean, organize, and prepare your datasets for analysis with R and the tidyverse (dplyr, tidyr).
- Reproducible Reports: Discover how to generate reproducible reports with R Markdown and Quarto, integrating analysis and documentation in one seamless workflow, crucial for collaboration and publication.
- Digital Project Management: Learn how to manage and share your research projects using GitHub, ensuring version control and collaborative efficiency.

Who is this course for?
This workshop is particularly relevant to early-career graduate students who are working with data who want to improve the reproducibility and efficiency of their projects but who do not yet have experience with the above tools. If you have taken longer courses on these topics, this course may be too introductory for you. While this one-day workshop provides an introduction to these tools, it is important to remember that mastering them will require extensive subsequent self-study and practice.

Simulated Data and Monte-Carlo Simulation Studies

Simulated data is incredibly useful. Before a study design has been finalised, you can use stimulated data to refine the design and analysis to ensure they address your research question. Before the real data has been collected, you can use simulated data to write and debug code for data processing, analysis, and visualisation. Without overfitting on your real data, you can use simulated data to learn how to use a statistical method. And using Monte Carlo simulation studies, you can better understand statistical methods, from simply how to implement and debug them to understanding their statistical power, false discovery rate, or what happens when you violate their assumptions.

This two-day workshop introduces participants to data simulation using R and the tidyverse. On the first day, participants will learn how to simulate datasets ranging from extremely simple to complex (e.g., non-normal distributions, truncated, multi-level). On the second day, participants will be introduced to simulation studies, where data is simulated and then analysed many times under different known conditions in order to estimate properties of interest (e.g., statistical power, false positive rates).

Key Takeaways:
- Simulating Unrealistic Data: In its simplest form, learning how to simulate even totally unrealistic data can be a useful starting point to help you write and debug code.
- Data Simulation for Workflow Development: Learn how to simulate realistic data to write and test your data wrangling, visualization, and analysis code, even before real data is available.
- Data Simulation for Methodological Exploration: Use realistic simulated data to practice new analysis techniques or methods in a controlled environment, helping you build confidence before applying them to real data.
- Monte Carlo Simulation Studies: Conduct simulation studies in R and purrr to improve your statistical inferences, e.g. by estimating statistical power, false positive rates, or robustness.

This workshop is particularly relevant to early-career graduate students who are working with data who want to improve their data analysis skills and confidence, but who do not yet have experience with the above tools and methods. If you have taken longer courses on these topics, this course may be too introductory for you. This course assumes some degree of experience with R and data wrangling (dplyr, tidyr) –the previous one-day course on data wrangling plus some practice will be sufficient. While this one-day workshop provides an introduction to these tools, it is important to remember that mastering them will require extensive subsequent self-study and practice.