Introduction to Data Analysis with R

Day 1: Welcome to the workshop

Selina Baldauf
Freie Universität Berlin - Theoretical Ecology

October 14, 2024

Who am I?

💻 **Scientific programmer** @theoretical ecology group
🎓 PhD in **dryland ecology**
👩‍🏫 **Teaching** R, Git, good scientific practice, …

I'm using R a lot for ...

... data analysis

... ecological modelling

... writing documents

.... workflow automation

...

Who are you?

How Osr1 controls the mechanical response of Fibro-adipogenic Progenitors?

osteopontin in colitis associated cancer

How do diverse institutions shape land-use dynamics in agricultural frontiers in the Peruvian Amazon?

Chemistry: Materials for water purification and some coatings

How can the recombinant toxin cCPE help to modulate and investigate the properties of the tight junction?

Epigenome Biology

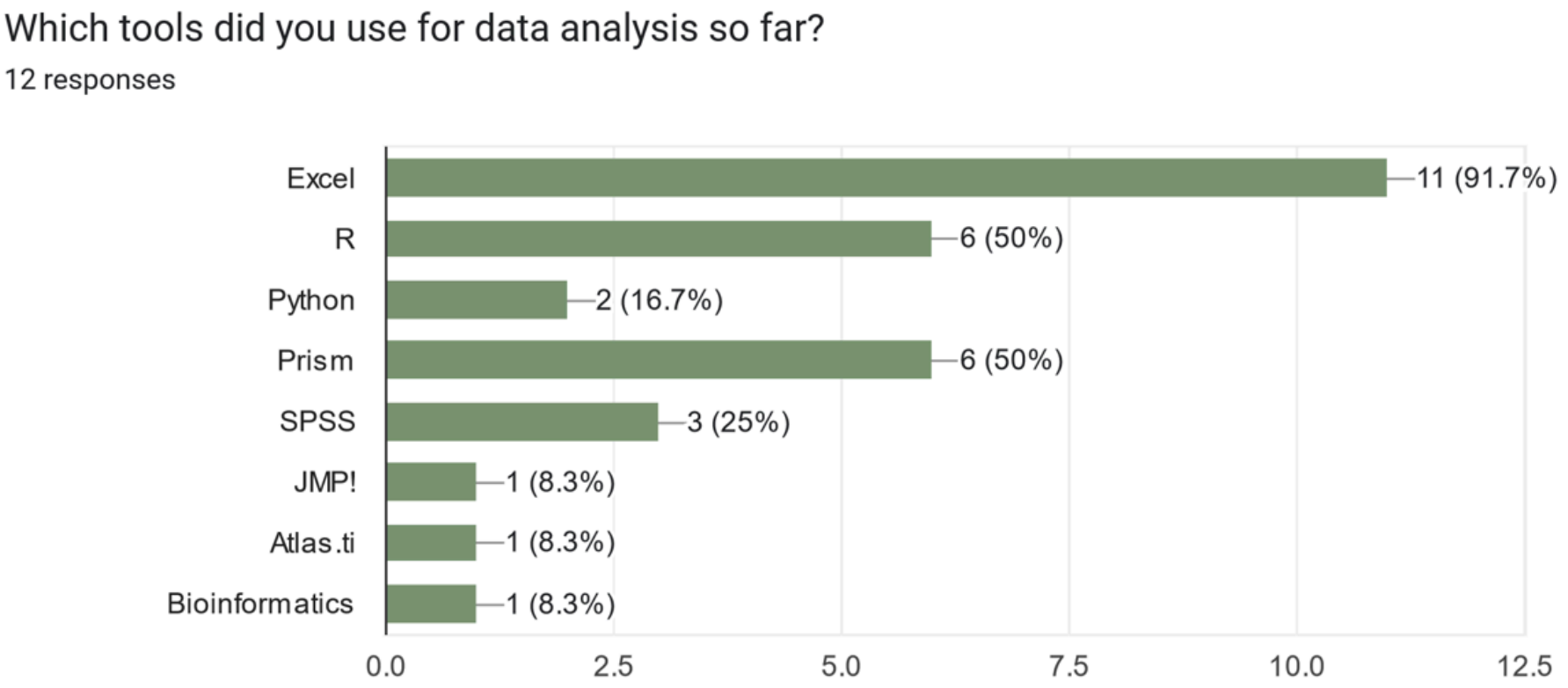
How do healthy adults recognize visual objects?

What is the influence of the gut microbiota on either IBD or MS?

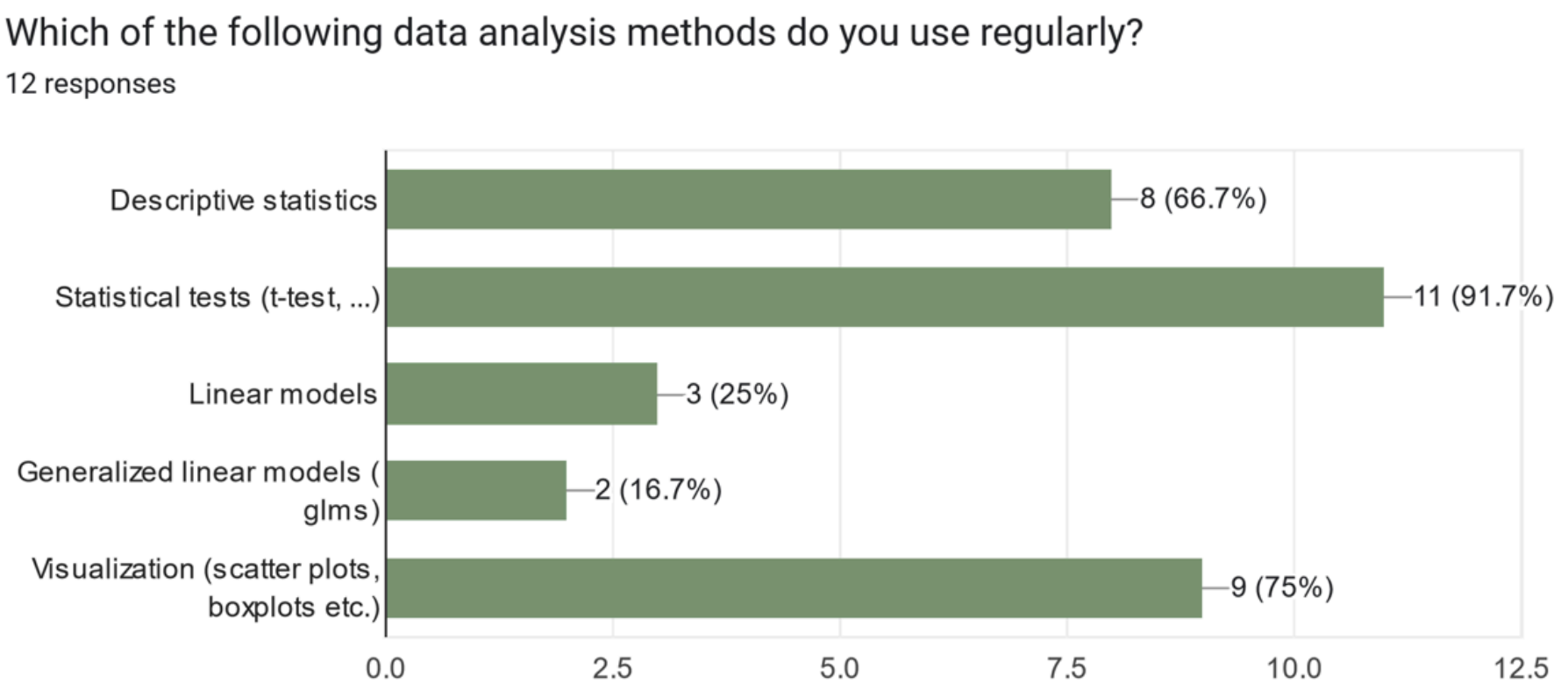
How can climate adaptation efforts become more conflict-sensitive, that is, avoid exacerbating the underlying drivers of conflict and contribute to a sustainable peace?

Disentangling the land-use effects on different dimensions of multitrophic ecosystem structure and functioning

Who are you?


Which tools did you use for data analysis so far?
12 responses
Excel
R
Python
Prism
SPSS
JMP!
Atlas.ti
Bioinformatics
11 (91.7%)
6 (50%)
2 (16.7%)
6 (50%)
3 (25%)
1 (8.3%)
1 (8.3%)
1 (8.3%)
0.0
2.5
5.0
7.5
10.0
12.5

Who are you?


Which of the following data analysis methods do you use regularly?
12 responses
Descriptive statistics
8 (66.7%)
Statistical tests (t-test, ...)
11 (91.7%)
Linear models
3 (25%)
Generalized linear models (glms)
2 (16.7%)
Visualization (scatter plots, boxplots etc.)
9 (75%)
0.0
2.5
5.0
7.5
10.0
12.5

Workshop topics

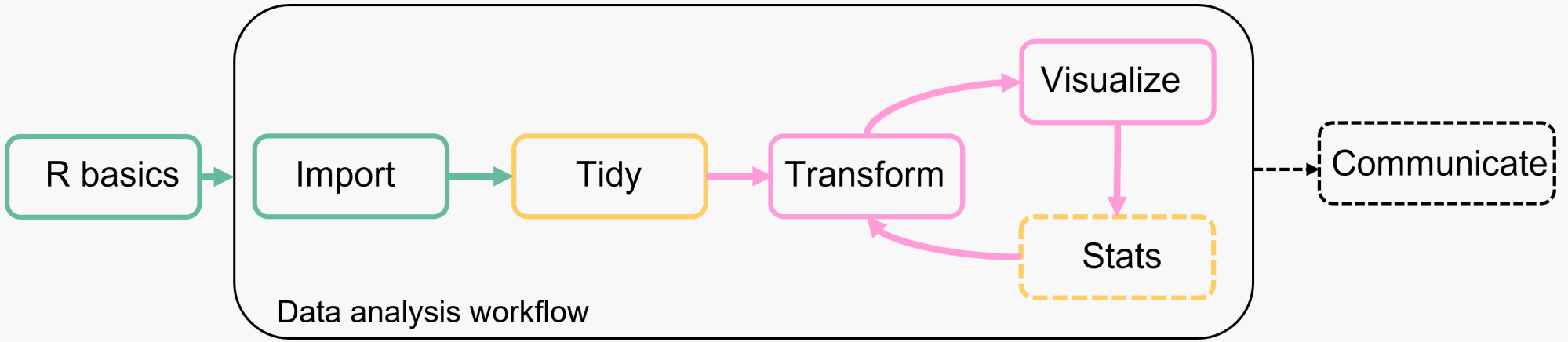


Image adapted from Wickham & Grolemund: R for Data Science

Day1 Introduction to R and RStudio and data import

Day2 Data transformation and visualization with the tidyverse

Day3 Cleaning data and statistical tests

Day 4 Bring your own data

- good practice (project organization, code structure, ...) + AI Tools

Schedule and Organization

📅 07.10.2023 - 08.10.2023 from 🕘 9 a.m. - 4 p.m.
📅 14.10.2023 - 15.10.2023 from 🕘 9 a.m. - 4 p.m.
🍽️ ~ 12 a.m. - 1 p.m., ☕ in between
📌 We will meet in the `General` meeting on Webex

Input sessions

- Presentation and demonstration of a topic
- Some examples

Tasks

- Solve them in small groups

Joint discussion of tasks & questions

Material

- All material can be found on the workshop's website
  - Presentations, Tasks, Solutions, Additional resources
- You can download slides and R Scripts from there
- Website will stay online after the workshop

Bring your own data

On the last workshop day, you can **work with your own research data**. I will also provide some **real life data sets** from different topics.

Learning by doing

- Get started using R for your own analyses
- Use any of the methods from the course or try new things, ...
- Present/discuss your results, questions and problems at the end of the day

- Add your name and some details on what you plan to do in this joint table

Before we get started I

- **Help each other** if possible
  - Have an eye on the chat
- All questions and comments are welcome
- If possible, please turn on your camera
- Feedback is welcome (Evaluation at the end of the workshop)

Before we get started II

How to use Webex teams

- `General` channel for our joint meetings and chat
- `Groups 1-4` for group work
  - Group spaces have their own chat and meetings
  - Use the chat in the groups for questions
  - Meet in your groups and solve the tasks jointly
- Please always have a look at the general chat during the tasks

Now you (5 min)

> Let's try the Webex groups

- **Go into your groups** and **start a meeting** there
- **Say hello** to your team briefly
  - What do you study and what types of data analysis do you do?
- **Come back to the general meeting** when I write in the general chat

Before we get started III

Did anyone have problems installing R and RStudio?

Download and install R from https://cran.r-project.org

Download and install RStudio from https://www.posit.co