Introduction to Data Analysis with R

Day 1: Welcome to the workshop

Selina Baldauf
Freie Universität Berlin - Theoretical Ecology

March 2, 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?

What happens when cyanobacteria cells acclimate to far-red light conditions for photosynthesis?

What is the neural basis of individual variability in visual and non-visual behavior?

Biophysical characterization of membrane proteins

Microbiology and Immunology

How does fasting affect the inflammation in Multiple Sclerosis

How are affective systems and their inter-individual variability associated with variability in the processing of emotionally and motivationally relevant stimuli in the brain?

dissecting the innate lymphoid cells in skin

colloidal chemistry

psychology

Is there a plant immunity receptor network at the endoplasmatic reticulum?

What is the phylogenetic relationship between Central and South American species?

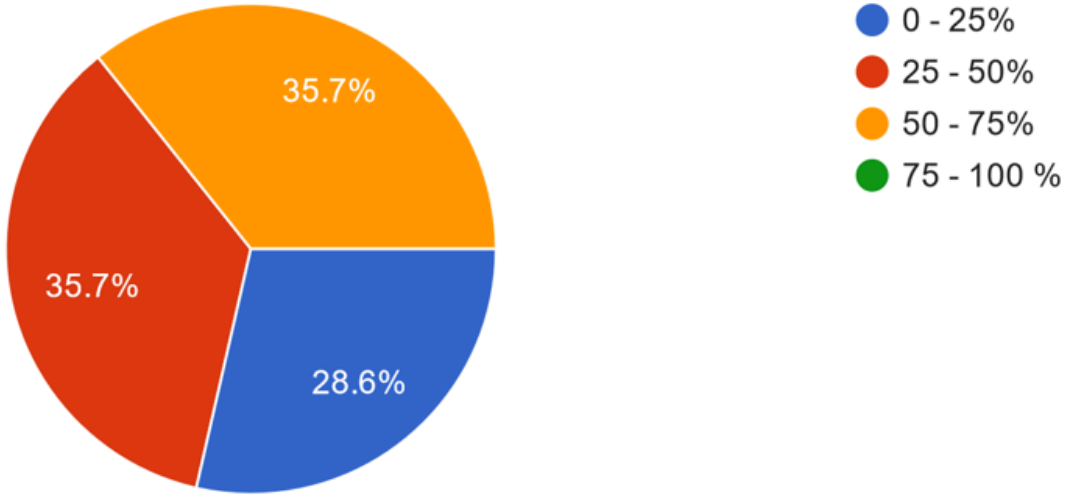
Identification of immune regulators with altered nuclear retention in Arabidopsis thaliana

How do raccoons adapt to the urban structure?

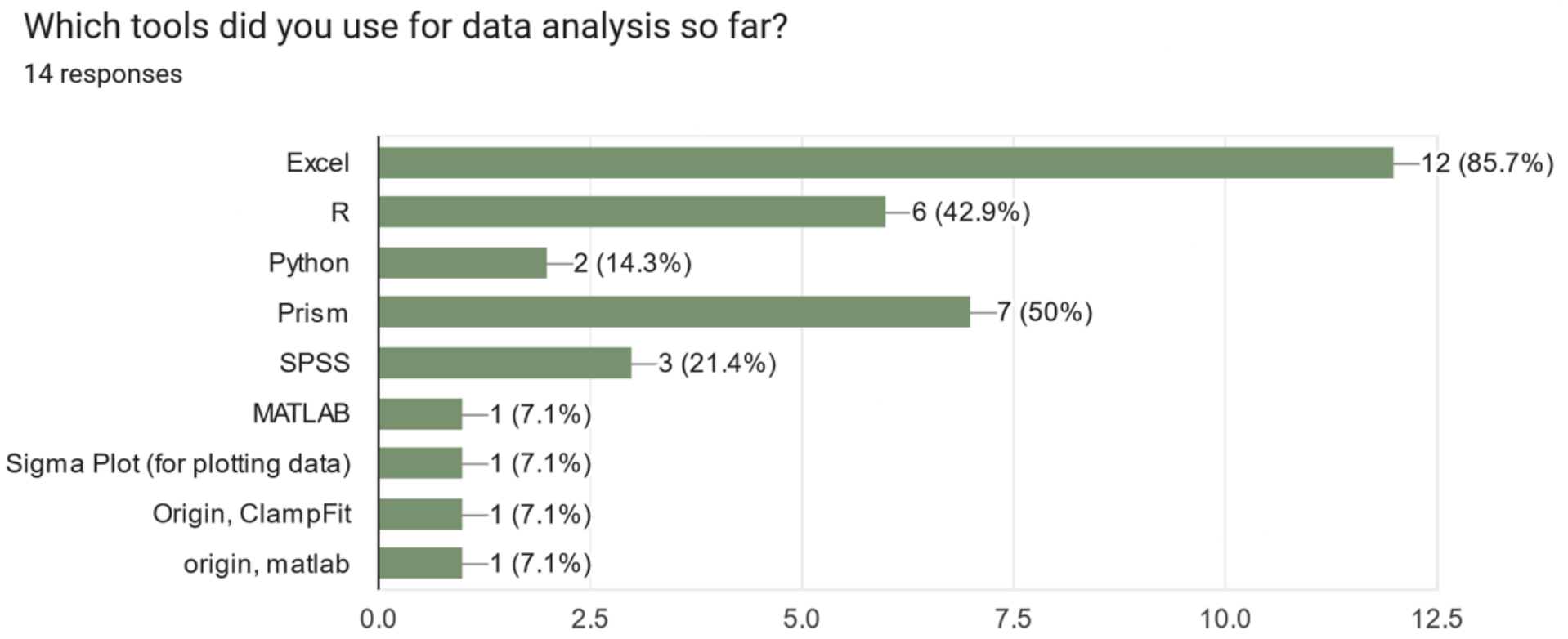
Who are you?

How much of your PhD time do you roughly spend with data analysis (including data preparation, plotting, statistics, ...)?

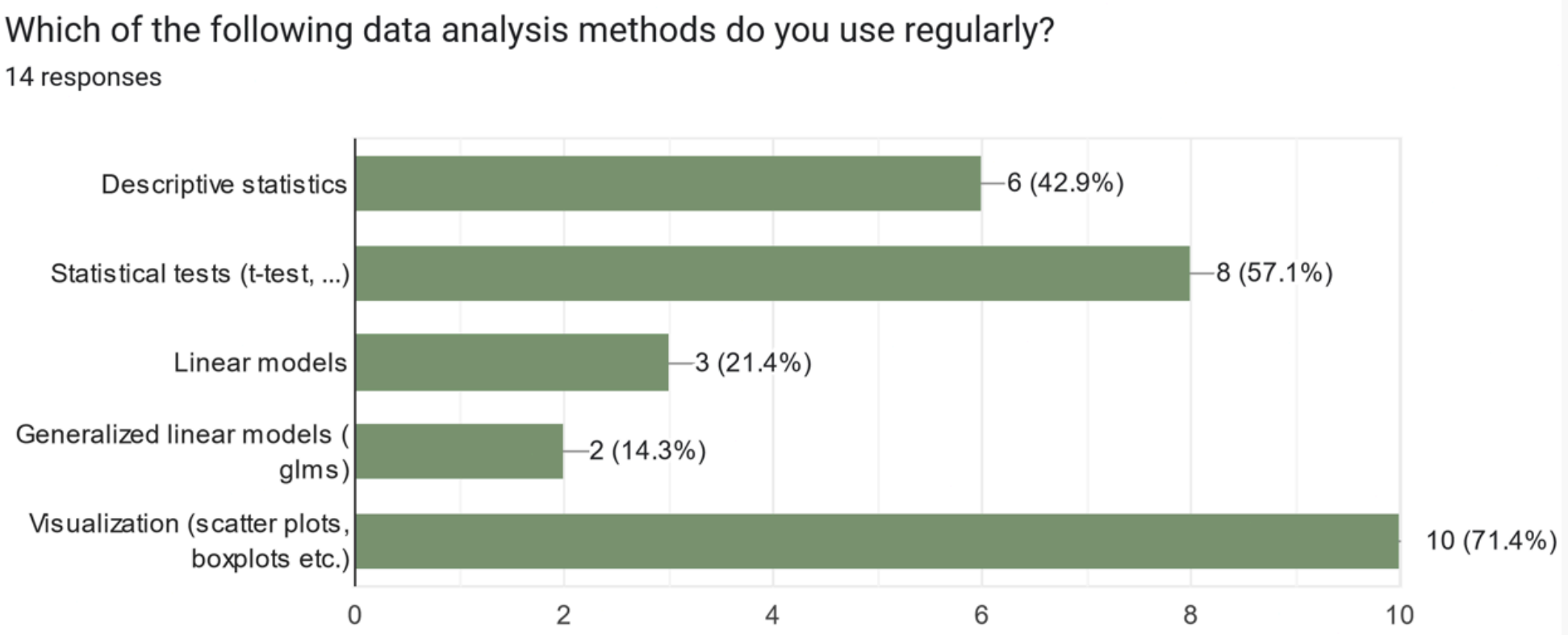
14 responses



Who are you?


Which tools did you use for data analysis so far?
14 responses
Excel
R
Python
Prism
SPSS
MATLAB
Sigma Plot (for plotting data)
Origin, ClampFit
origin, matlab
12 (85.7%)
6 (42.9%)
2 (14.3%)
7 (50%)
3 (21.4%)
1 (7.1%)
1 (7.1%)
1 (7.1%)
1 (7.1%)
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?
14 responses
Descriptive statistics
6 (42.9%)
Statistical tests (t-test, ...)
8 (57.1%)
Linear models
3 (21.4%)
Generalized linear models (glms)
2 (14.3%)
Visualization (scatter plots, boxplots etc.)
10 (71.4%)
0
2
4
6
8
10

Workshop topics

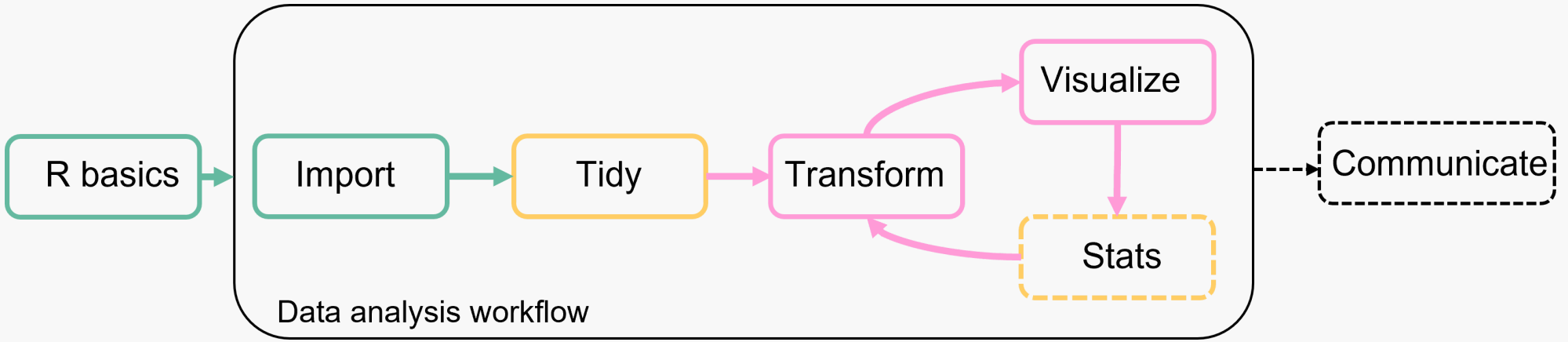


Image adapted from Wickham & Grolemund: R for Data Science

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

Schedule and Organization

16.10.2023 - 17.10.2023 from 9 a.m. - 4 p.m.
30.10.2023 - 31.10.2023 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

Bring your own data

- Keep the last workshop day in mind during the next days
  - Take a mental note if you learn something that is applicable to your data
  - Think of questions that you would like to answer for your data set
- 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
- 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