

Thanks a lot for participating!



Artwork by [Allison Horst](#)

Bring your own data

Day 4 - Introduction to Data Analysis with R

Selina Baldauf

Freie Universität Berlin - Theoretical Ecology

October 6, 2025

Organization

Schedule of today

- **Now - 14.30:** Work on the data set(s)
 - Take break(s) as best fits your needs
- **14.30 - 15:** Short feedback round
 - What did you find out about your data set? Plots, summaries, ...
 - Which methods did you use?
 - Did you learn something new?
 - Was there something you struggled with?
 - ...
- **15-16:** Feedback, conclusion

Data set 1: What makes a good wine?

see [here](#)



Frederik Vandaele - originally posted to Flickr as Château Pétus, CC BY 2.0,
<https://commons.wikimedia.org/w/index.php?curid=5145286>

Data set 2: Paralympic games from 1980-2016

[see here](#)

Data set 4: Ice cover and temperature

see [here](#)




Image by LTER under CC BY-SA 4.0

Source ice data: [Magnuson, J.J., S.R. Carpenter, and E.H. Stanley. 2021.](#) North Temperate Lakes LTER: Ice Duration - Madison Lakes Area 1853 - current ver 35. Environmental Data Initiative.

Data set 5: RNAseq data

see [here](#)

 Download example script

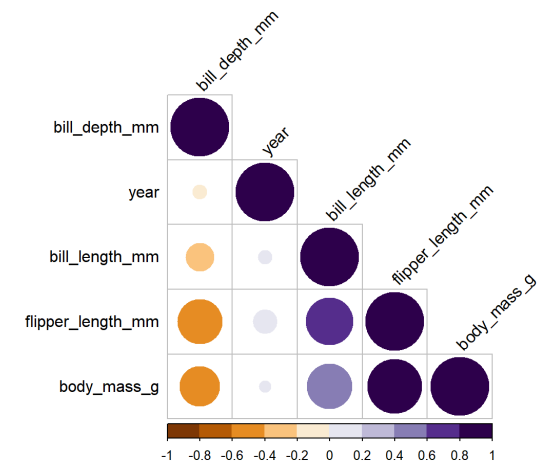
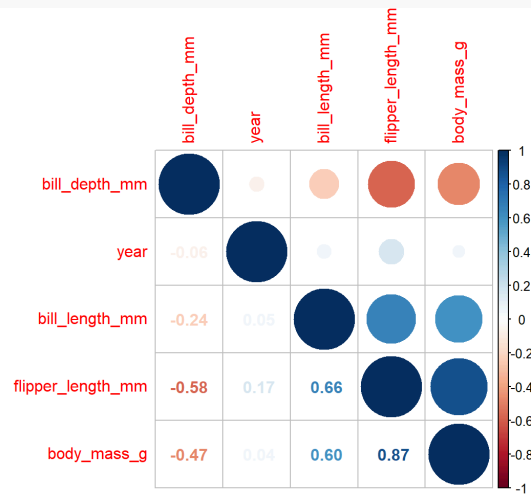
New methods: Read multiple files at once

- If you have multiple files with data you can
 - Read them all in at the same time
 - Combine them into one tibble

 Download example script

New methods: Correlation plots

 Download example script



New methods: PCAS

 Download example script

```
Error in `select()`:  
! Can't select columns that don't exist.  
✗ Column `bill_len` doesn't exist.
```

```
Error: object 'peng_pca' not found
```

```
Error: object 'peng.pca' not found
```

Some general tips

- First make a plan:
 - What do you want to achieve and what are the steps?
 - Start with something small, e.g. reading in the data and bringing it into the right format.
 - If you want, stop by in general to discuss your plan or write me in the chat
- Google, ask AI tools if you want, look at the function help (?)
 - Hint: Set your RStudio error messages to English: `Sys.setenv(LANGUAGE='en')` in the console
- If you get stuck, ask in the chat or stop by in General

Now you

Working with real data

- **Meet** in your group (if you want)
- **Work** on your data set
- **Take breaks** as you need and **be back** at 2.30 p.m.

Sharing

In 1-2 mins:

- What was the highlight of your analysis?
 - Your favorite plot
 - Some cool code
 - A problem that you finally solved
 - Something new you learned
- What was difficult?
- If you want: Share a screenshot in the chat or share your screen

Feedback

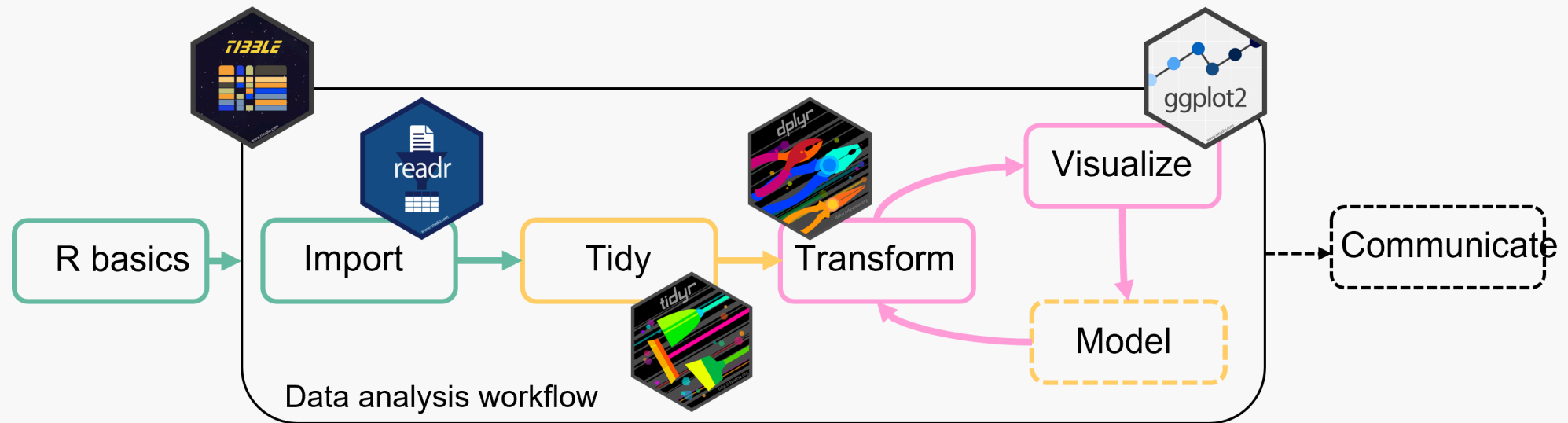
Please take 10 mins to complete the feedback survey for the Graduate center (don't use Internet Explorer)

<https://votingo.cedis.fu-berlin.de/PCNLP3>

Feedback

- Any other feedback or comments from your side?

Conclusion

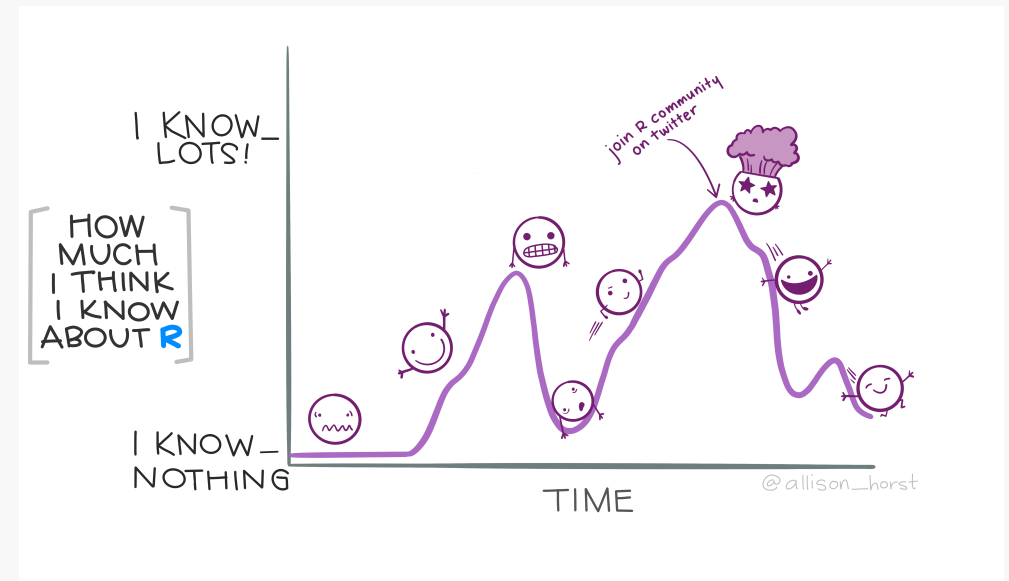


We learned a lot of stuff!

Conclusion

How to continue from here?

- Learning by doing!
- Have a look at some [online resources](#)
- [FU statistical consulting](#) for questions regarding statistical methods
- [R Consulting by me](#)
- [Tools and Tips lecture](#)



Artwork by [Allison Horst](#)

The End