

Reproducible Data Analysis with R

— Welcome —

Cédric Scherer // R Course TU Dresden // Feb 27-Mar 3, 2023



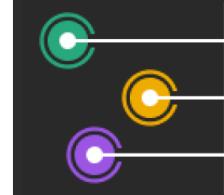
Welcome!



HELLO
my name is

Cédric





CÉDRIC SCHERER

Data Visualization & Information Design



Consulting

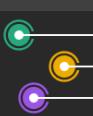
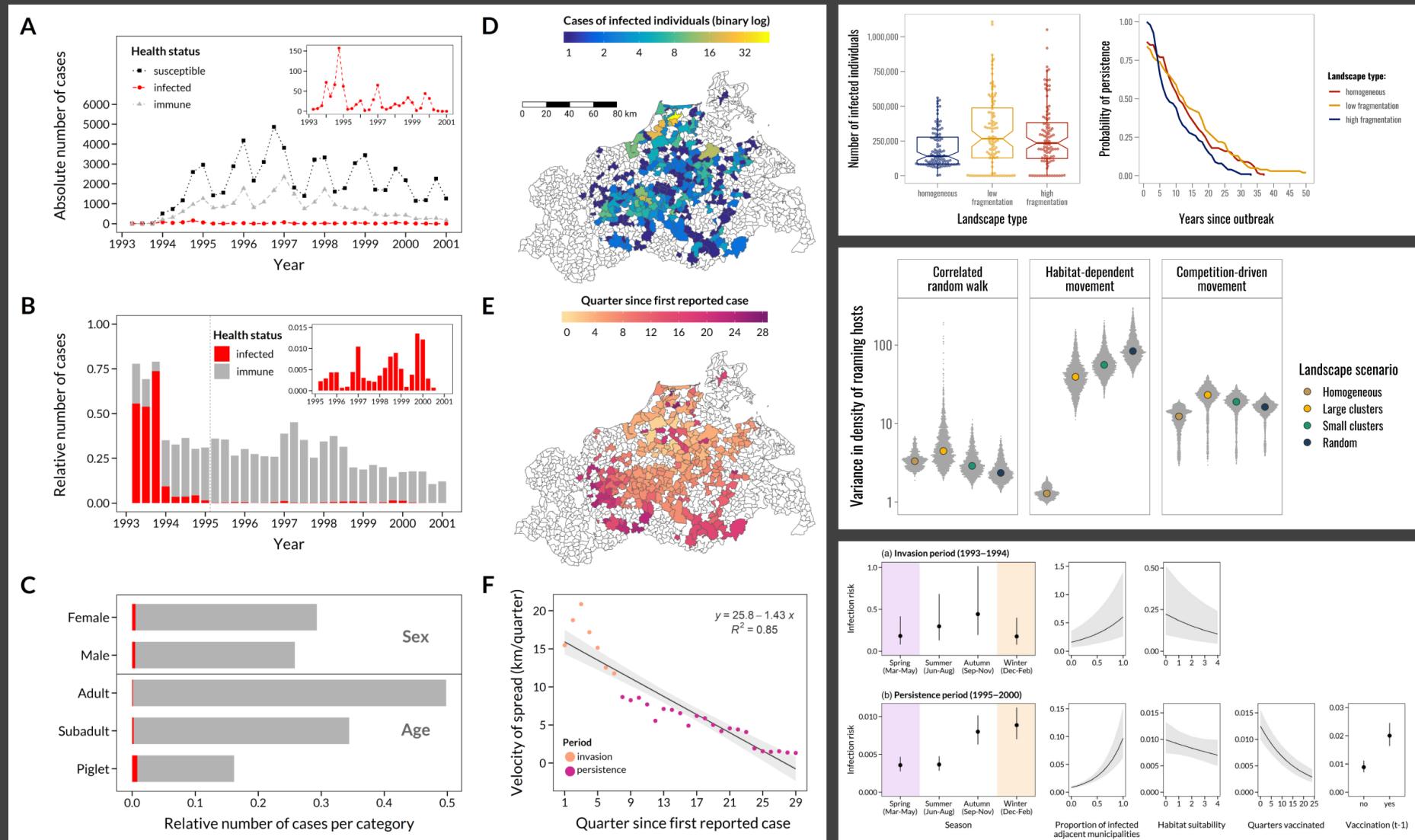


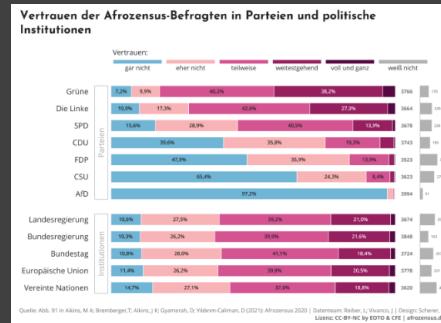
Coaching



Coding







Politiker rechnen bald mit einer Fortsetzung der Fußball-Bundesliga.
Wenn auch nicht im Stadion, so ist es voraussichtlich bald wieder möglich Fußball im Fernsehen zu sehen.

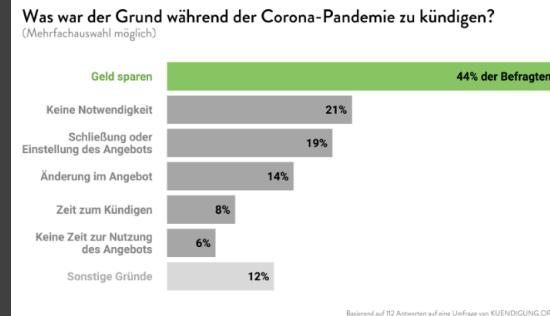
Ich habe **einen** Vertrag mit einem Anbieter für Sportübertragungen.

50% finden das richtig
16% haben dazu keine Meinung
34% finden das falsch

Ich habe **keinen** Vertrag mit einem Anbieter für Sportübertragungen.

31% finden das richtig
35% haben dazu keine Meinung
33% finden das falsch

Basiert auf 1546 Antworten auf eine Umfrage von KUENDIGUNG.ORG



Häufigkeit von Diskriminierungserfahrungen entlang ausgewählter Vielfaltdimensionen im Bereich „Arbeitsleben“



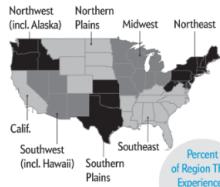
GRAPHIC SCIENCE

Text by Clara Moskowitz | Graphic by Cédric Scherer and Georgios Karamanis

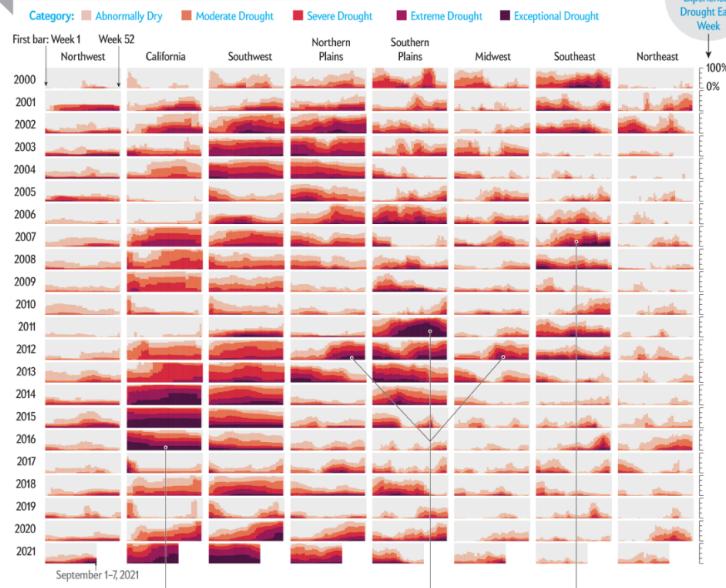
Escalating Drought

Climate change is intensifying periods of extreme dryness, particularly in the U.S. West

For more than 20 years the National Drought Mitigation Center (NDMC) has been monitoring dozens of indices of drought around the country, including satellite measurements of evaporation and color in vegetation, soil-moisture sensors, rainfall estimates, and river and streamflow levels. Although the agency's weekly assessments have identified periods of exceptional drought before, lately dryness has been ramping up. "The changing climate is definitely contributing to more natural disasters, drought being one of them," says Brian Fuchs, a climatologist who oversees the weekly report at the NDMC. "We're seeing more frequent and high-intensity episodes. This year some of these areas in the West have been in drought more than they have been without drought."



Drought Extent and Intensity by Region over Time



Source: U.S. Drought Monitor, created by the National Drought Mitigation Center at the University of Nebraska-Lincoln, U.S. Department of Agriculture, and National Oceanic and Atmospheric Administration (data).

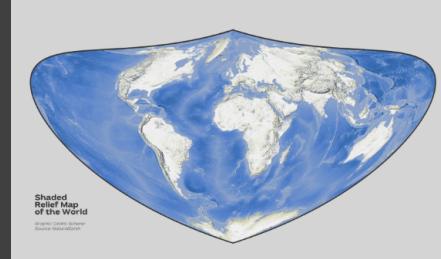
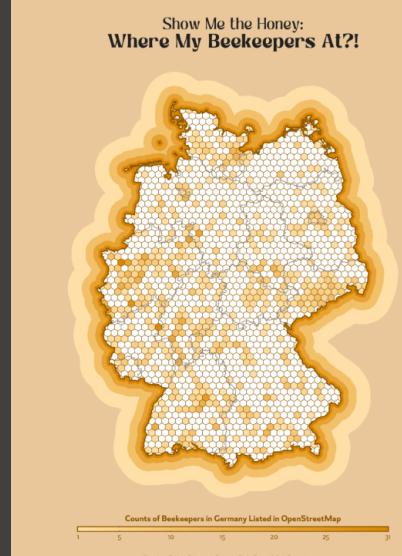
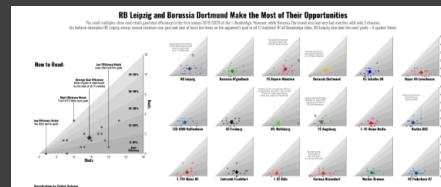
California experienced its hottest drought in recorded history from 2012 to 2016. A warming climate makes the atmosphere thirstier, which increases evaporation and boosts drought.

A drought that originated in the Southern Plains in 2011 eventually spread to the Midwest and Northern Plains when the moisture coming in from the Gulf of Mexico was absorbed by the parched South before it could reach the North.

The Southeast's driest year to date was 2007, when only 31.85 inches of rain fell in Atlanta, 62 percent of its average yearly rainfall.

74 Scientific American, November 2021





Not my cup of coffee...

Each dot on every coffee bean rated by Quality Initiative-trained roasters, in addition, the multiple vertical stripes show where beans are from. The red dots indicate coffee grown from one to six varieties. The green coffee beans come from coffee farms that have been certified organic. The blue dots indicate coffee beans with 20 or more tested beans are diverse. The black dots indicate each country's median score.

Illustration by Cédric Scherer

10 POINTS

30 POINTS

60 POINTS

90 POINTS

120 POINTS

150 POINTS

180 POINTS

210 POINTS

240 POINTS

270 POINTS

300 POINTS

330 POINTS

360 POINTS

420 POINTS

480 POINTS

540 POINTS

600 POINTS

660 POINTS

720 POINTS

780 POINTS

840 POINTS

900 POINTS

960 POINTS

1020 POINTS

1080 POINTS

1140 POINTS

1200 POINTS

1260 POINTS

1320 POINTS

1380 POINTS

1440 POINTS

1500 POINTS

1560 POINTS

1620 POINTS

1680 POINTS

1740 POINTS

1800 POINTS

1860 POINTS

1920 POINTS

1980 POINTS

2040 POINTS

2100 POINTS

2160 POINTS

2220 POINTS

2280 POINTS

2340 POINTS

2400 POINTS

2460 POINTS

2520 POINTS

2580 POINTS

2640 POINTS

2700 POINTS

2760 POINTS

2820 POINTS

2880 POINTS

2940 POINTS

3000 POINTS

3060 POINTS

3120 POINTS

3180 POINTS

3240 POINTS

3300 POINTS

3360 POINTS

3420 POINTS

3480 POINTS

3540 POINTS

3600 POINTS

3660 POINTS

3720 POINTS

3780 POINTS

3840 POINTS

3900 POINTS

3960 POINTS

4020 POINTS

4080 POINTS

4140 POINTS

4200 POINTS

4260 POINTS

4320 POINTS

4380 POINTS

4440 POINTS

4500 POINTS

4560 POINTS

4620 POINTS

4680 POINTS

4740 POINTS

4800 POINTS

4860 POINTS

4920 POINTS

4980 POINTS

5040 POINTS

5100 POINTS

5160 POINTS

5220 POINTS

5280 POINTS

5340 POINTS

5400 POINTS

5460 POINTS

5520 POINTS

5580 POINTS

5640 POINTS

5700 POINTS

5760 POINTS

5820 POINTS

5880 POINTS

5940 POINTS

6000 POINTS

6060 POINTS

6120 POINTS

6180 POINTS

6240 POINTS

6300 POINTS

6360 POINTS

6420 POINTS

6480 POINTS

6540 POINTS

6600 POINTS

6660 POINTS

6720 POINTS

6780 POINTS

6840 POINTS

6900 POINTS

6960 POINTS

7020 POINTS

7080 POINTS

7140 POINTS

7200 POINTS

7260 POINTS

7320 POINTS

7380 POINTS

7440 POINTS

7500 POINTS

7560 POINTS

7620 POINTS

7680 POINTS

7740 POINTS

7800 POINTS

7860 POINTS

7920 POINTS

7980 POINTS

8040 POINTS

8100 POINTS

8160 POINTS

8220 POINTS

8280 POINTS

8340 POINTS

8400 POINTS

8460 POINTS

8520 POINTS

8580 POINTS

8640 POINTS

8700 POINTS

8760 POINTS

8820 POINTS

8880 POINTS

8940 POINTS

9000 POINTS

9060 POINTS

9120 POINTS

9180 POINTS

9240 POINTS

9300 POINTS

9360 POINTS

9420 POINTS

9480 POINTS

9540 POINTS

9600 POINTS

9660 POINTS

9720 POINTS

9780 POINTS

9840 POINTS

9900 POINTS

9960 POINTS

10020 POINTS

10080 POINTS

10140 POINTS

10200 POINTS

10260 POINTS

10320 POINTS

10380 POINTS

10440 POINTS

10500 POINTS

10560 POINTS

10620 POINTS

10680 POINTS

10740 POINTS

10800 POINTS

10860 POINTS

10920 POINTS

10980 POINTS

11040 POINTS

11100 POINTS

11160 POINTS

11220 POINTS

11280 POINTS

11340 POINTS

11400 POINTS

11460 POINTS

11520 POINTS

11580 POINTS

11640 POINTS

11700 POINTS

11760 POINTS

11820 POINTS

11880 POINTS

11940 POINTS

12000 POINTS

12060 POINTS

12120 POINTS

12180 POINTS

12240 POINTS

12300 POINTS

12360 POINTS

12420 POINTS

12480 POINTS

12540 POINTS

12600 POINTS

12660 POINTS

12720 POINTS

12780 POINTS

12840 POINTS

12900 POINTS

12960 POINTS

13020 POINTS

13080 POINTS

13140 POINTS

13200 POINTS

13260 POINTS

13320 POINTS

13380 POINTS

13440 POINTS

13500 POINTS

13560 POINTS

13620 POINTS

13680 POINTS

13740 POINTS

13800 POINTS

13860 POINTS

13920 POINTS

13980 POINTS

14040 POINTS

14100 POINTS

14160 POINTS

14220 POINTS

14280 POINTS

14340 POINTS

14400 POINTS

14460 POINTS

14520 POINTS

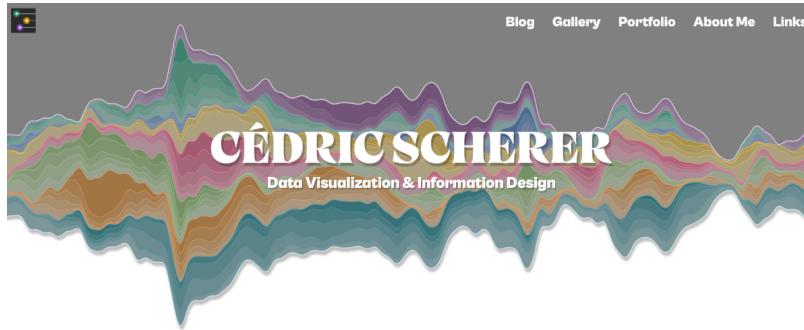
14580 POINTS

14640 POINTS

14700 POINTS

14760 POINTS

14820 POINTS



Hi, I am Cédric 🙌

Data Visualization Designer, Consultant and Instructor
for Engaging and Effective Graphical Storytelling.

[← Gallery](#) [← Portfolio](#) [← About Me](#)



Always coding. Passionate about
design. Worried about nature and
inequality. Proud dad.



[Schedule a discovery call!](#)

Quick Links

[ggplot2 Tutorial](#)
[Evolution of a ggplot](#)

Featured Tags



Friends

DataVisSociety · From Data to Vis
Corneel · Monica Sotiri
Georgios Karanikas · Julie Brunet
Matthias Stach · Inbal Rief · Wili Chose
Anne-Marie Dufour · Heureka Latte

2-Day Workshop on "Graphic Design with ggplot2" at rstudio::conf 2022

End of July, I had the honor to teach an in-person ggplot2 workshop at the rstudio::conf in Washington DC. All course resources are available on the course webpage featuring slides, hands-on R codes, recap notes, and exercises including prepared scripts and step-by-step solutions.

Tuesday • August 9, 2022

International Women's Day 2022: The Pay Gap in Europe

Sadly, inequality between women and men is still an issue. In 2020, women's gross hourly earnings were on average 13 percentage points below those of male paid employees in the European Union. Only five countries in the EU report a gap smaller than 5 percentage points with Luxembourg being closest to zero.

Tuesday • March 8, 2022

The World's Countries Colored by Their First Letter

While preparing the mapping section for a Pearson-O'Reilly training, I got the idea to visualize the first letter of each country. And got especially curious about how much landmass each letter covers. Turns out: A, C and R are covering the largest areas!

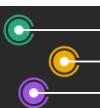
Friday • August 27, 2021

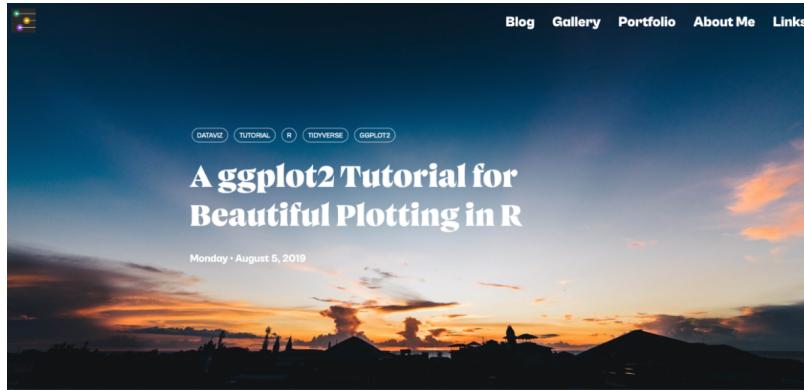
A Quick How-to on Labelling Bar Graphs in ggplot2

Bar charts are likely the most common chart type out there and come in several varieties. Most notably, direct labels can increase accessibility of a bar graph. I got a request how one can add percentage labels inside the bars and how to highlight specific bars with ggplot2. This short tutorial shows you multiple ways how to do so.

Monday • July 5, 2021

Cédric Scherer // R Course TU Dresden // Feb 27-Mar 3, 2023



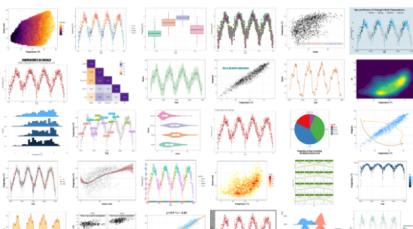


Introductory Words

I don't care, just show me the content!

Back in 2016, I had to prepare my PhD introductory talk and I started using [{ggplot2}](#) to visualize my data. I never liked the syntax and style of base plots in R, so I was quickly in love with ggplot. Especially useful was its faceting utility. But because I was short on time, I plotted these figures by trial and error and with the help of lots of googling. The resource I came always back to was a blog entry called [Beautiful plotting in R: A ggplot2 cheatsheet](#) by Zev Ross, updated last in January 2016. After giving the talk which contained some decent plots thanks to the blog post, I decided to go through this tutorial step-by-step. I learned so much from it and directly started modifying the codes and over the time I added additional code snippets, chart types and resources.

Since the blog entry by Zev Ross was not updated for some years and step by step this became a unique version of a tutorial, I decided to host the updated version on my GitHub. Now it finds its proper place on this homepage! (Plus I added a ton of other updates—just to name a few: The fantastic [{patchwork}](#), [{eggtext}](#) and [{eggforce}](#) packages. How to deal with custom fonts and colors. A collection of R packages tailored to create interactive charts. And several other chart types including pie charts because everyone looooves pie charts!)





End of July, I had the honor to teach a [2-day, in-person](#) workshop on "Graphic Design with ggplot2" at the [rstudio::conf\(2022\)](#) in Washington DC. Invited by RStudio (now named [Posit](#)), I developed a new course that covers the most important steps and helpful tips to create visually appealing, engaging and complex graphics with ggplot2. The course focused on the main concepts of the grammar of graphics and used hands-on examples to explore ggplot2's utility to create multi-layered, more complex graphs. [All course resources are available as open-source material on the course page.](#)



The course webpage as well as the slide decks and the exercises and solutions were developed with the new open-source scientific and technical publishing system [Quarto](#). The new workshop development was a perfect opportunity to give it a try and the experience was overall wonderful—the [reveal.js](#) integration for the slides works perfect and allows for a lot of customization. Thanks to Marco Sciolini for helping me setting up the course webpage which was, after learning about a few quirks, a smooth experience as well.

Graphic Design with ggplot2

How to Create Engaging and Complex Visualizations in R

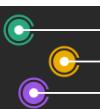
An [rstudio::conf\(2022\) Workshop](#) by Cédric Scherer

July 25 and 26, 2022
10am - 11am
National Harbor 4/F
[rstat.jsconf](#)

Course Overview

Creating engaging and accessible data visualizations of high quality in an efficient and preferably reproducible way is an essential tool in a data scientist's toolbox. After completing the course, participants will have a solid understanding of data visualization principles and the functionality of the ggplot2 package and its helpful extensions to create highly customized graphics without the need for post-processing.

We will discuss the main concepts of the grammar of graphics and use hands-on examples to explore ggplot2's utility to create multi-layered, more complex graphs. The workshop covers a short overview of the course objectives and links to all session and the GitHub repository.



Search or jump to... Pull requests Issues Codespaces Marketplace Explore

Cédric Scherer
z3tt

Data Visualization Specialist • Computational Ecologist

Edit profile

1.4k followers • 119 following

Self-Employed • IZW Berlin
Berlin
cedricscherer.com
@CedScherer

Achievements

My GitHub Activity

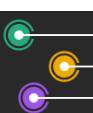
What I Actually Do

Pinned

Customize your pins

822 contributions in the last year Contribution settings

Cédric Scherer // R Course TU Dresden // Feb 27-Mar 3, 2023



HELLO
my name is

Meet each other!

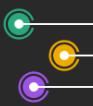


Conversation Starters

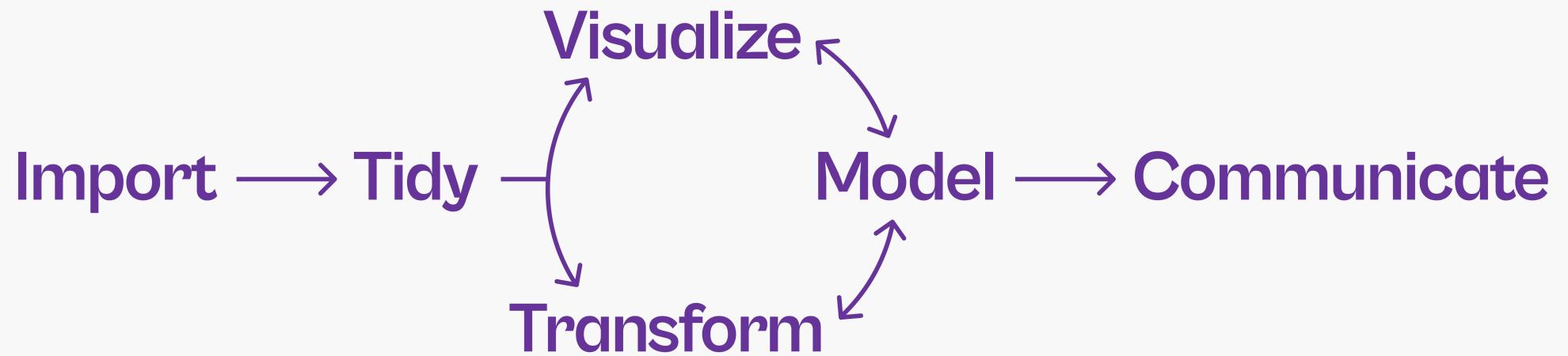
- What's your name?
- Where do you feel home?
- What is your research topic?
- When did you use R for the first time?
- What's your favorite animal / plant / color / typeface?
- Whom would you like to meet once in lifetime?
- Which R package are you looking forward to try?



What is This Course About?



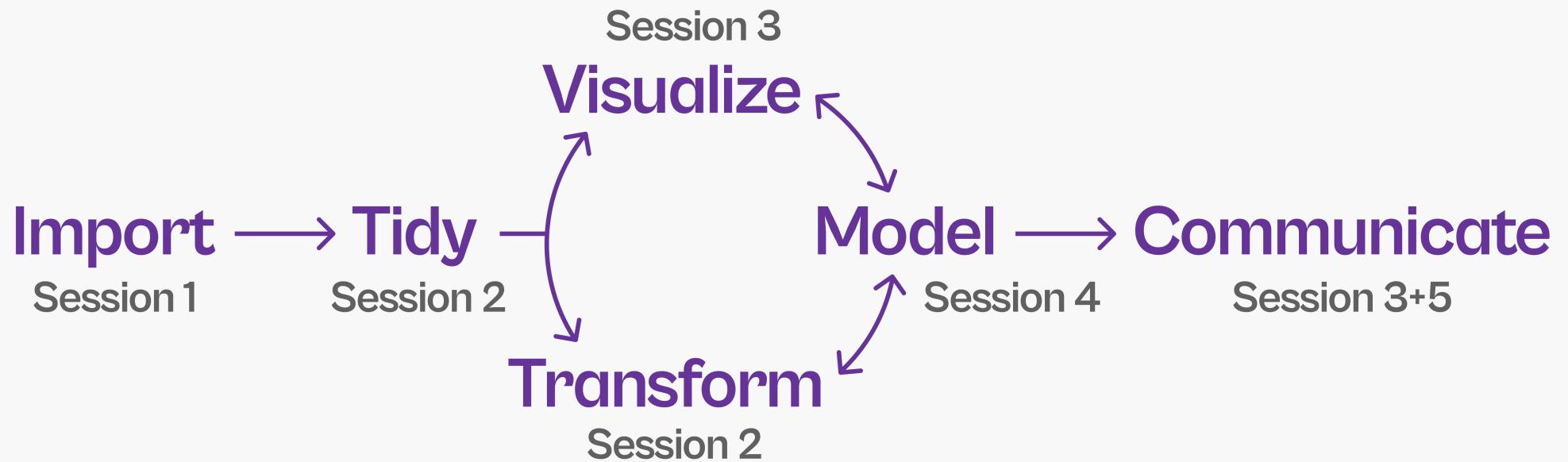
The Data Science Workflow



The data science workflow, modified from ["R for Data Science"]low, modified from "R for Data Science"



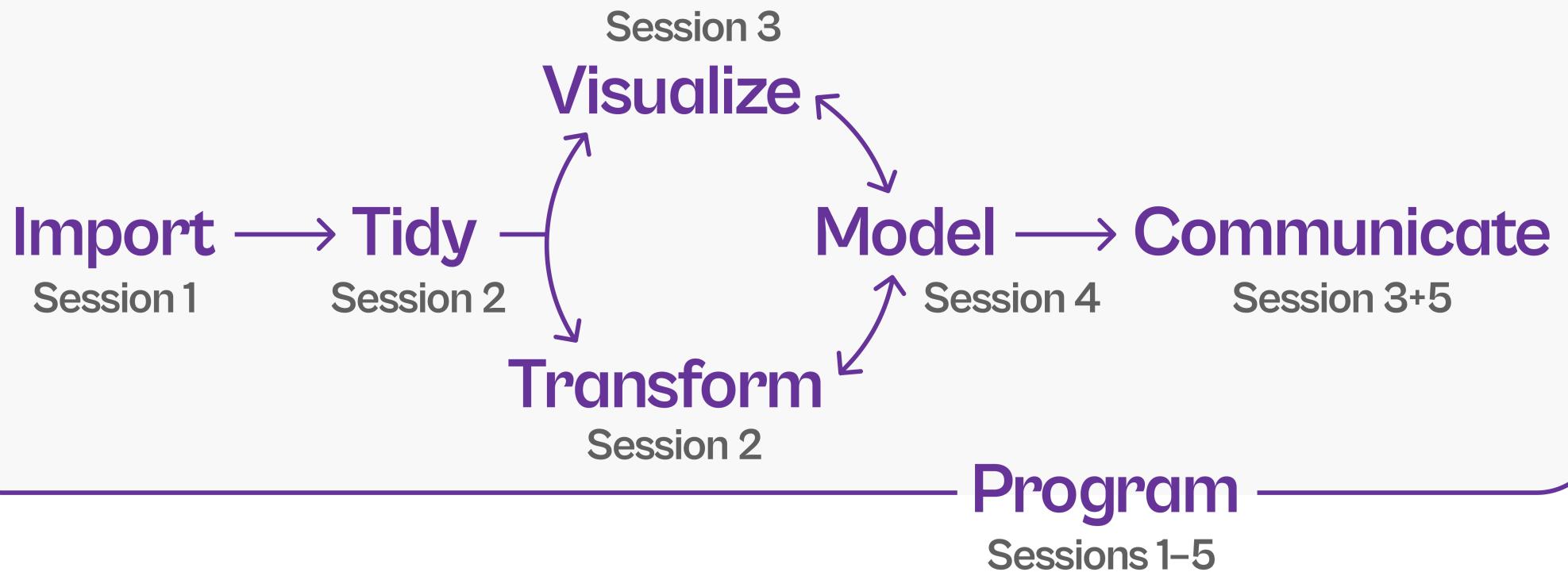
Course Structure



The data science workflow, modified from ["R for Data Science"]low, modified from "R for Data Science"



Course Structure



The data science workflow, modified from ["R for Data Science"]ow, modified from "R for Data Science"







Illustration by [Allison Horst](#)



Announcements

Course Materials

- Download all materials here: tinyurl.com/rcourse-tud
 - R Scripts: `./scripts`
 - Slides: `./slides`
 - Cheat Sheets: `./cheatsheets`
- HTML slides: z3tt.github.io/TUD_R/session-name.html

Course Preparation

- Software: **R (v4+)** and a suitable IDE, e.g. **RStudio**
- R Packages: You will learn about packages within the first session.
 - Just make sure that you have sufficient rights to install packages on your machine by running the following in the R console: `install.packages("palmerpenguins")`



Let's Get Started!

