

Graphic Design with ggplot2

How to Create Engaging and Complex Visualizations in R

Cédric Scherer // rstudio::conf // July 2022

Welcome!

HELLO
my name is

*Meet the
teaching team!*

THE Teaching TEAM



Cédric Scherer

Workshop Instructor (he/him)

Cédric Scherer

Independent Data Visualization Specialist
Computational Ecologist at IZW Berlin



Consulting

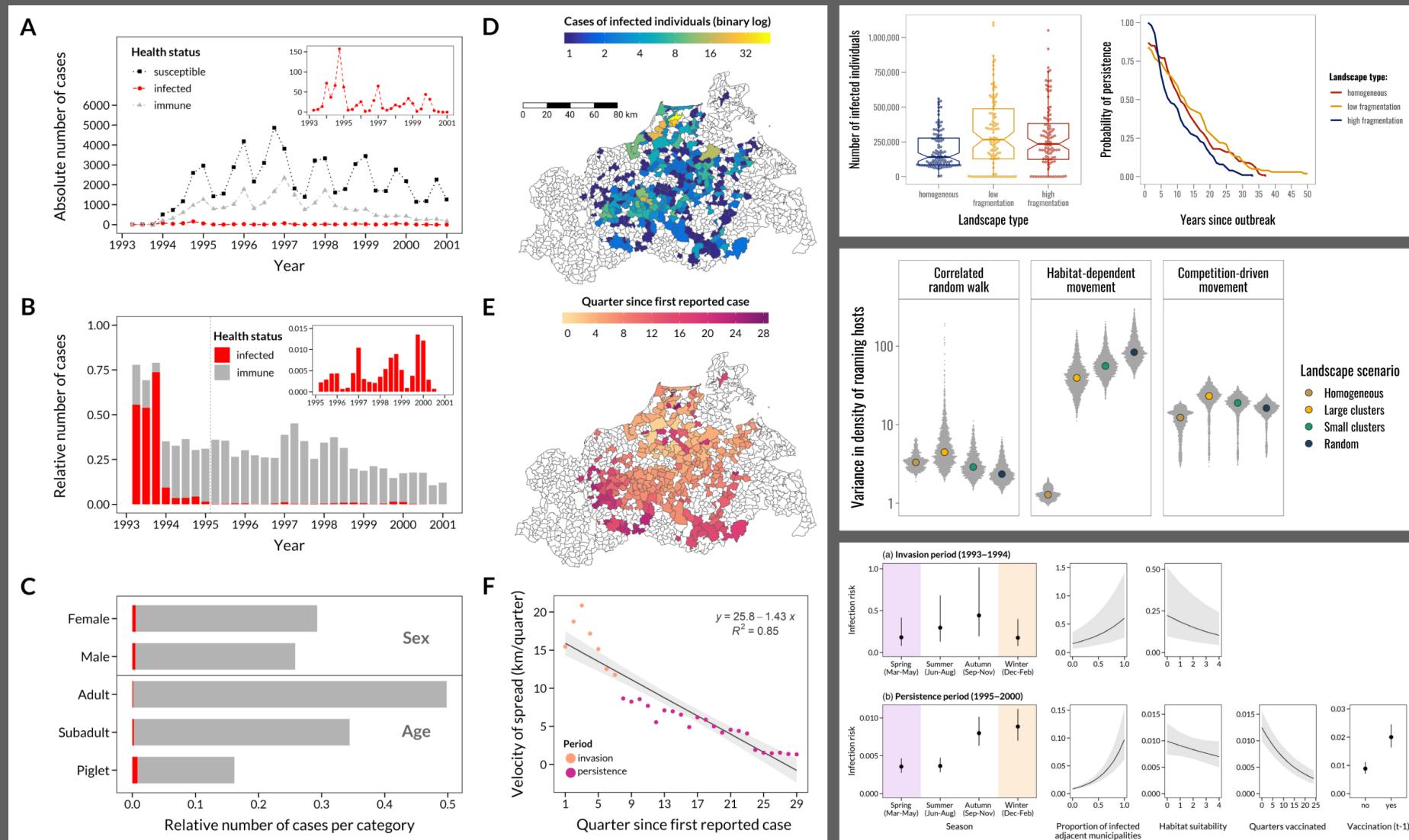
Coaching

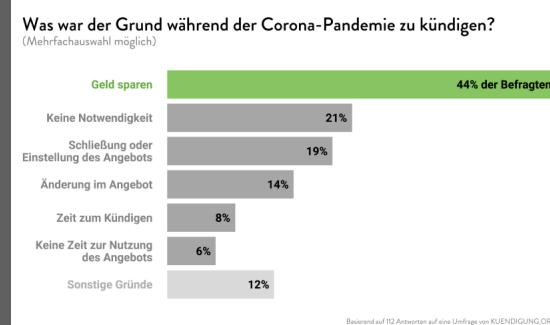
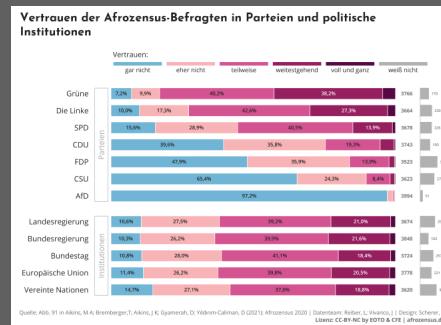
Coding

★ cedricscherer.com

Twitter Instagram Be @CedScherer

z3tt





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.



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



Basierend auf 1546 Antworten auf eine Umfrage von KUENDIGUNG.ORG

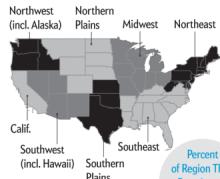
GRAPHIC SCIENCE

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

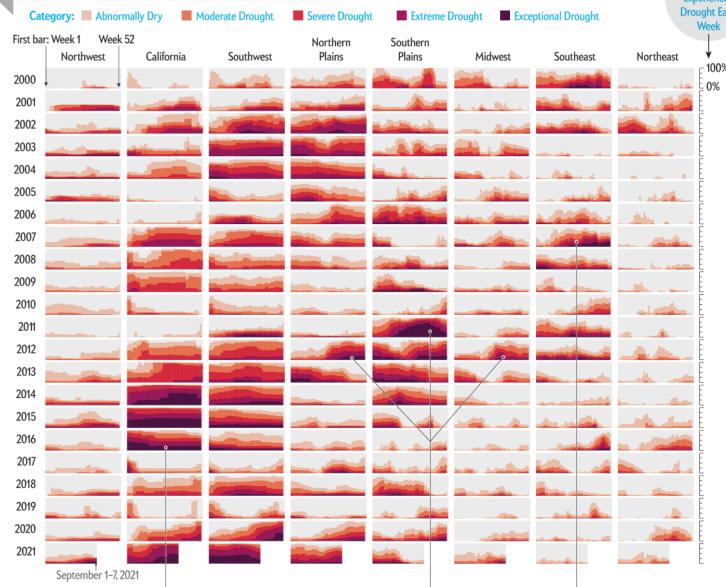
Escalating Drought

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

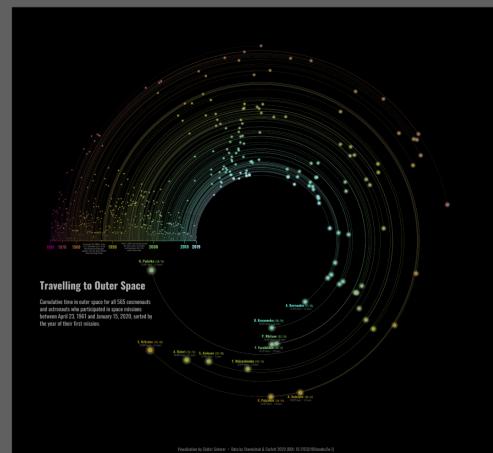
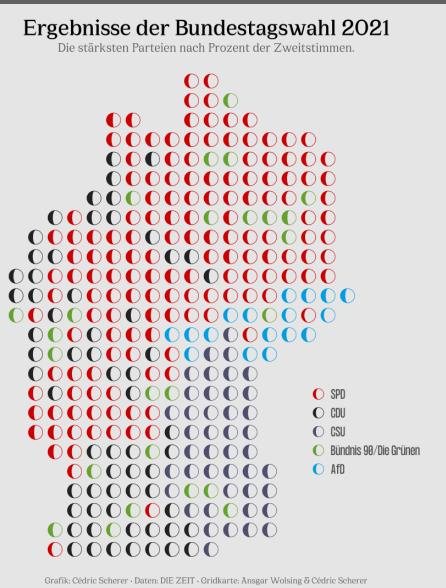
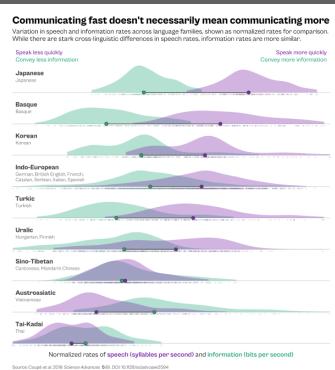
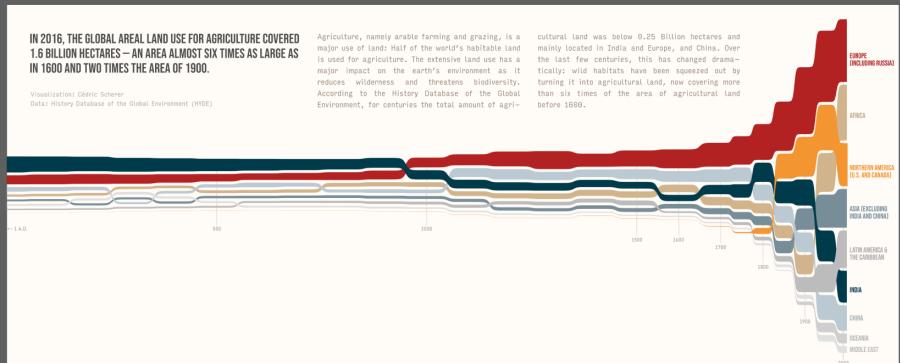
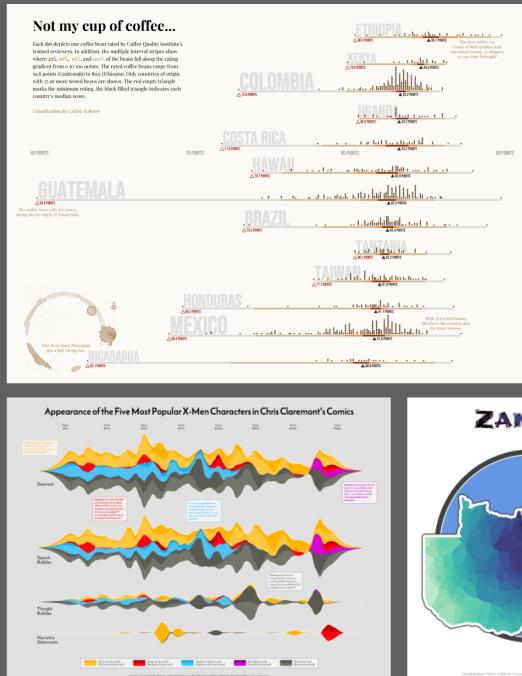
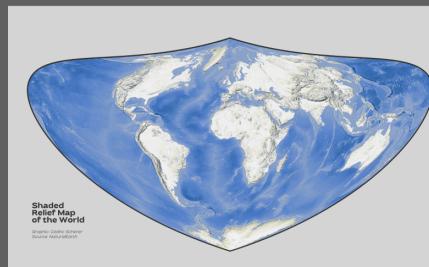
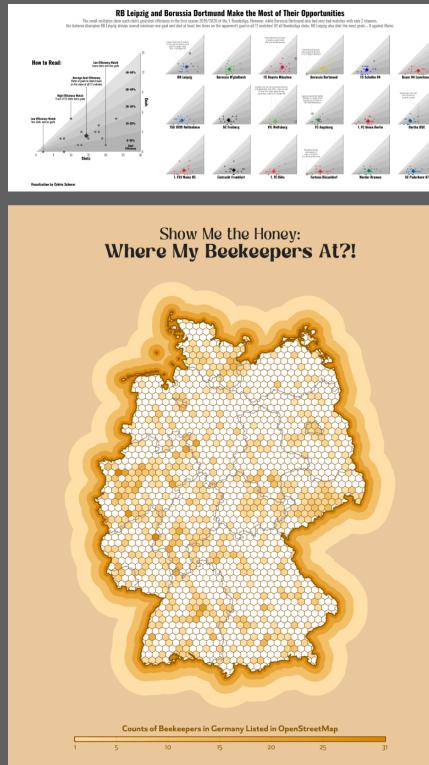


Source: U.S. Drought Monitor, created by the National Drought Mitigation Center at the University of Nebraska-Lincoln, U.S. Department of Agriculture, and National Oceanic and Atmospheric Administration (data).

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.



THE Teaching TEAM



Cédric Scherer

Workshop Instructor (he/him)



Christine Jin

Teaching Assistant & Logistics Coordinator (she/her)

THE Teaching TEAM



Cédric Scherer

Workshop Instructor (he/him)



Christine Jin

Teaching Assistant & Logistics Coordinator (she/her)



Thomas Lin Pedersen

Teaching Assistant (he/him)

THE Teaching TEAM



Cédric Scherer

Workshop Instructor (he/him)



Christine Jin

Teaching Assistant & Logistics Coordinator (she/her)



Thomas Lin Pedersen

Teaching Assistant (he/him)



Ahmaddou Dicko

Teaching Assistant (he/him)

THE Teaching TEAM



Cédric Scherer

Workshop Instructor (he/him)



Christine Jin

Teaching Assistant & Logistics Coordinator (she/her)



Thomas Lin Pedersen

Teaching Assistant (he/him)



Ahmadou Dicko

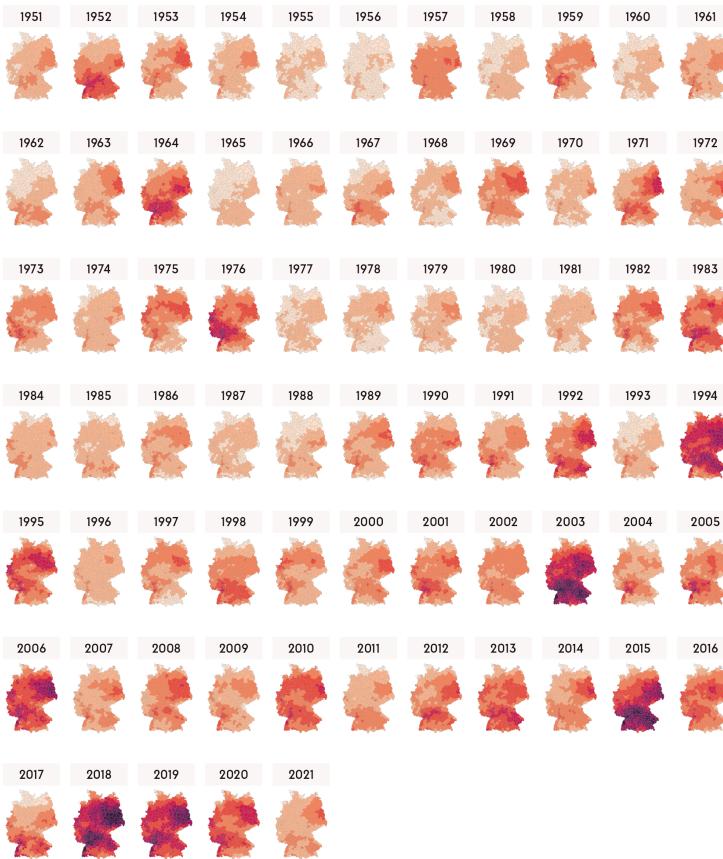
Teaching Assistant (he/him)



Marco Sciaiani

Teaching Assistant (he/him)

Germany's heating-up



Data: BKG & DWD | Visualisation: Marco Scialini

THE Teaching TEAM



Cédric Scherer

Workshop Instructor (he/him)



Christine Jin

Teaching Assistant & Logistics Coordinator (she/her)



Thomas Lin Pedersen

Teaching Assistant (he/him)



Ahmadou Dicko

Teaching Assistant (he/him)



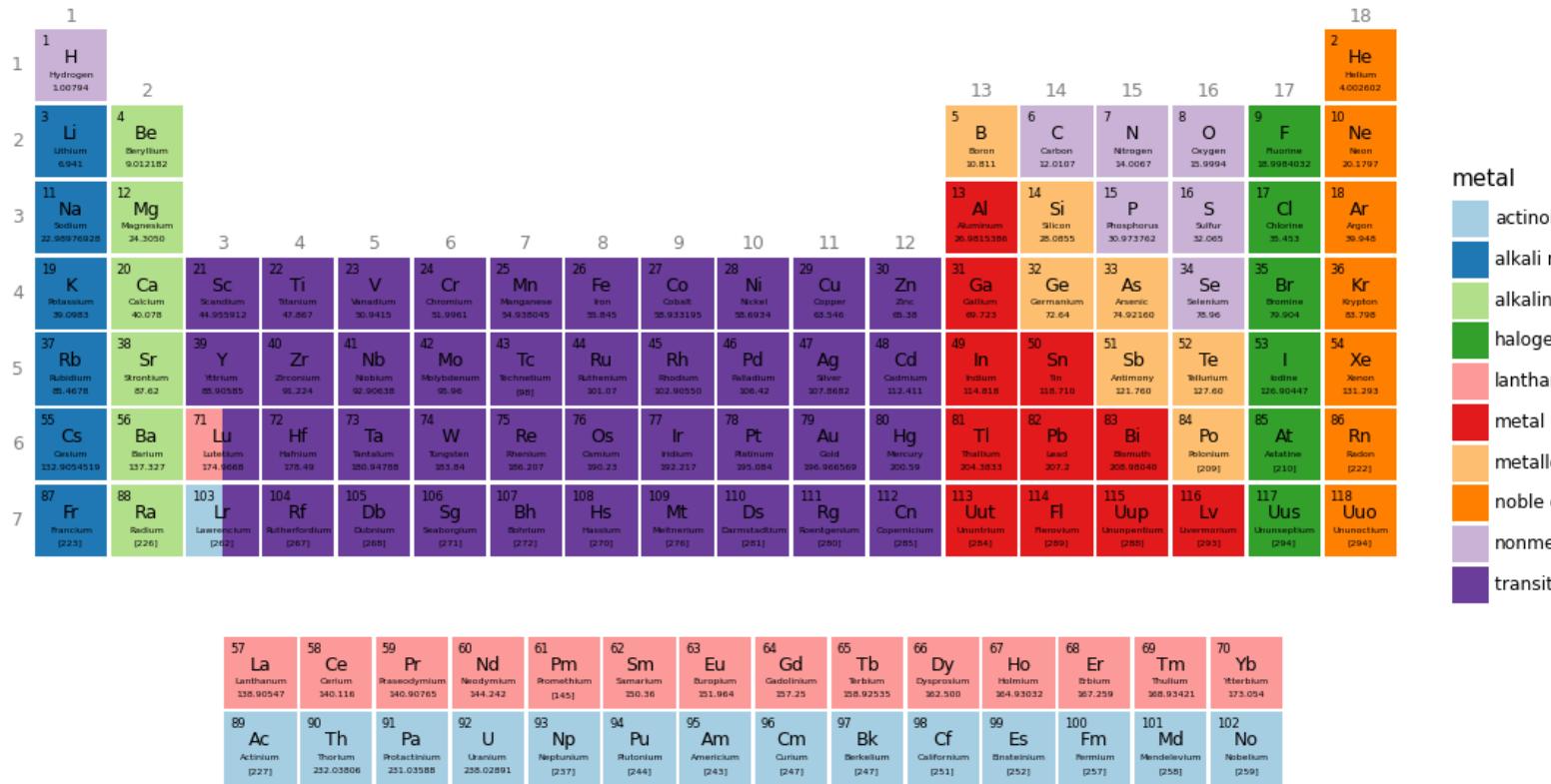
Marco Sciaiani

Teaching Assistant (he/him)



Hassan Kibirige

Teaching Assistant (he/him)



metal

- actinoid
- alkali metal
- alkaline earth metal
- halogen
- lanthanoid
- metal
- metalloid
- noble gas
- nonmetal
- transition metal

HELLO
my name is

Meet each other!

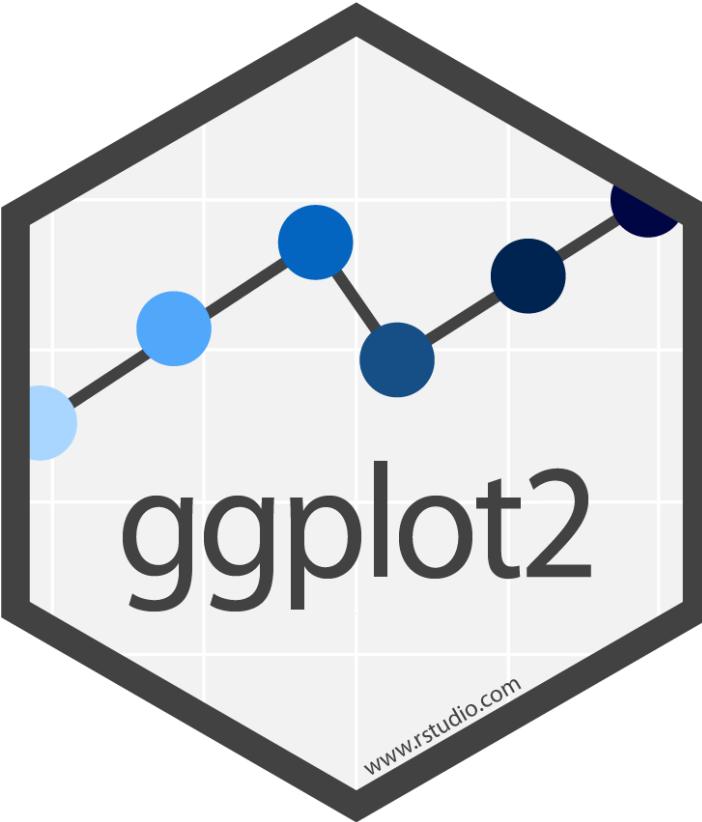
Conversation Starters

- What's your name?
- Where do you feel home?
- When did you use R for the first time?
- What's your favorite animal / plant / color / typeface?
- Whom would you like to meet during the rstudio::conf?
- Which R package are you looking forward to try?

Communication Strategy

- **Green sticky note** – I am done with the exercise
- **Pink sticky note** – I need help or support
- You can ask questions / comment at any time during the course.
- Please avoid questions during the breaks to give us a chance to recover and reflect.
- Team up with your neighbors for exercises – and lunch!
- We will collect feedback twice during the day (more later).

Introduction



{ggplot2} is a system for declaratively creating graphics,
based on “The Grammar of Graphics” (Wilkinson, 2005).

You provide the data, tell **{ggplot2}** how to map variables to aesthetics,
what graphical primitives to use, and it takes care of the details.

Advantages of {ggplot2}

- consistent underlying “grammar of graphics” (Wilkinson 2005)
- very flexible, layered plot specification
- theme system for polishing plot appearance
- lots of additional functionality thanks to extensions
- active and helpful community

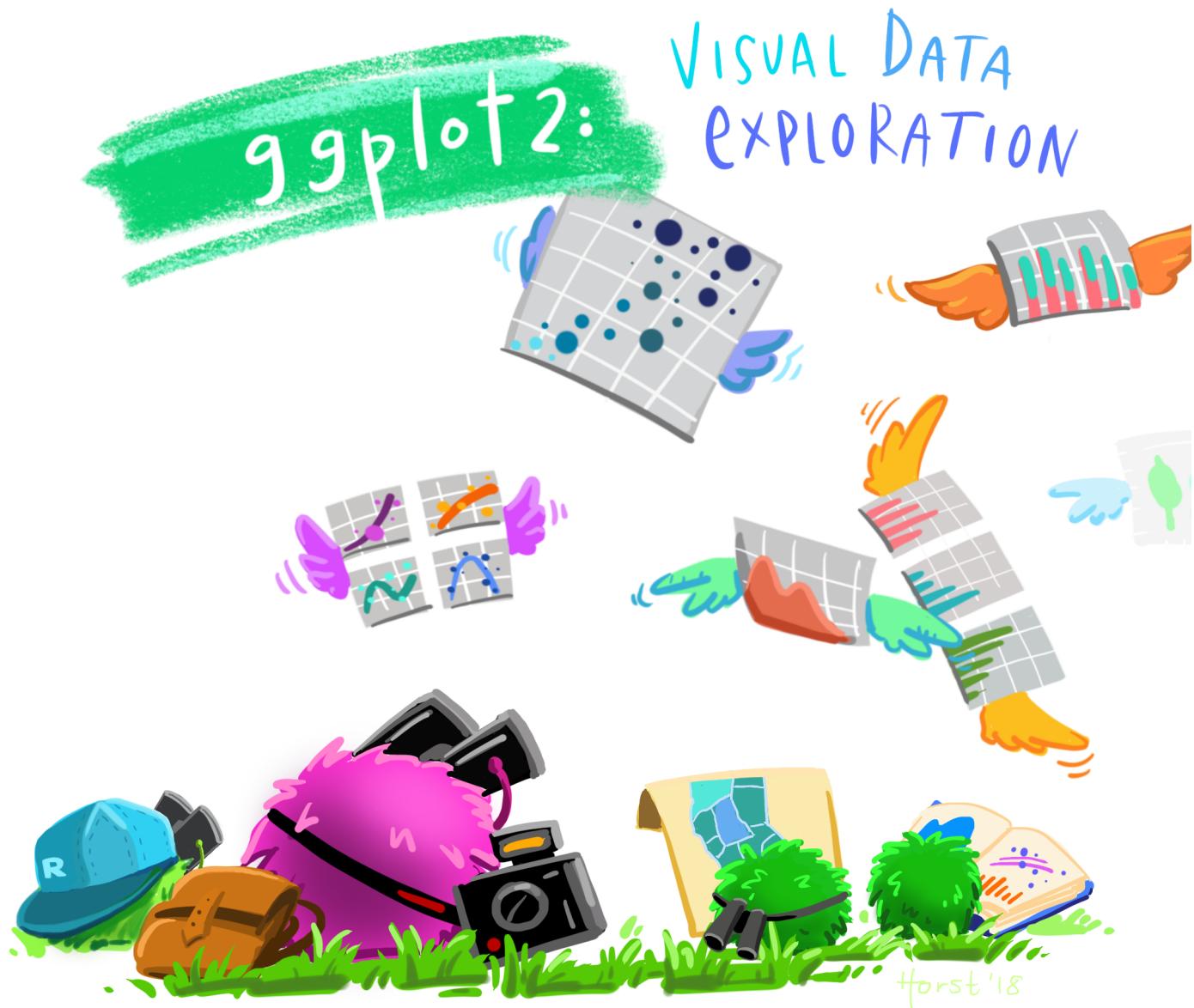
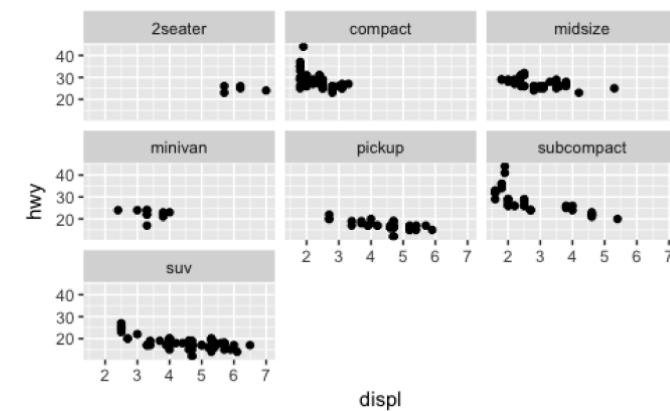
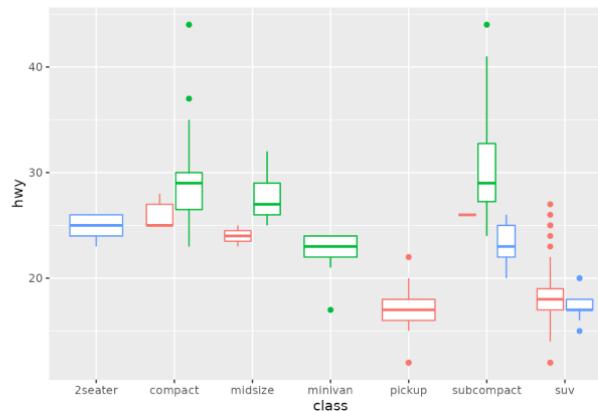
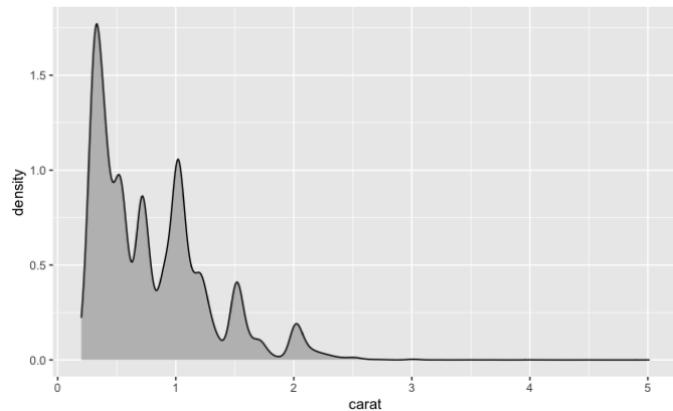
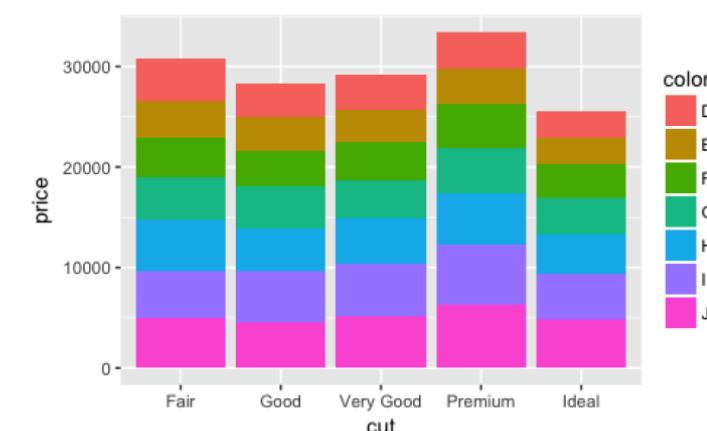
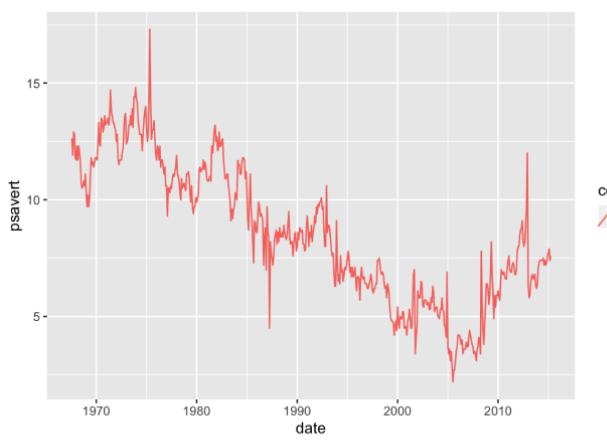
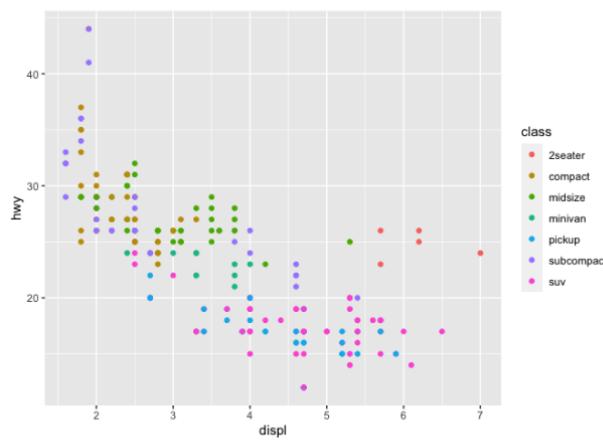
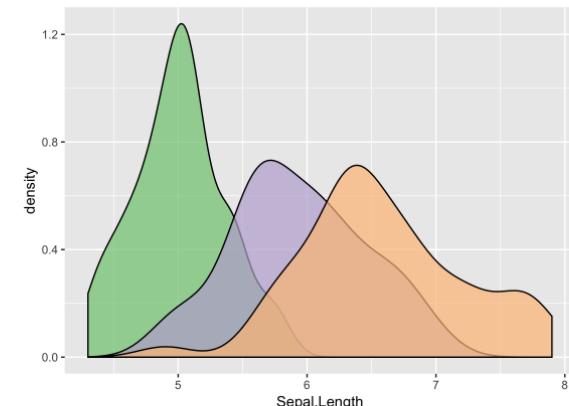
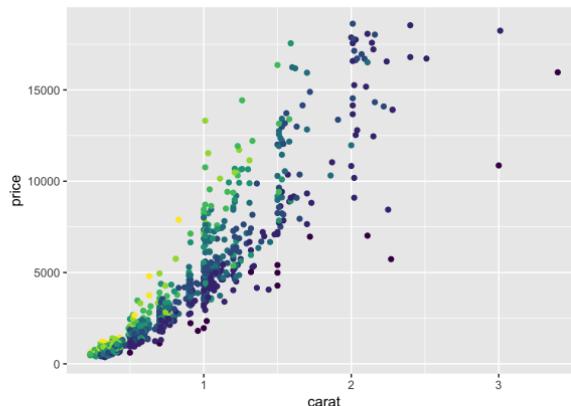
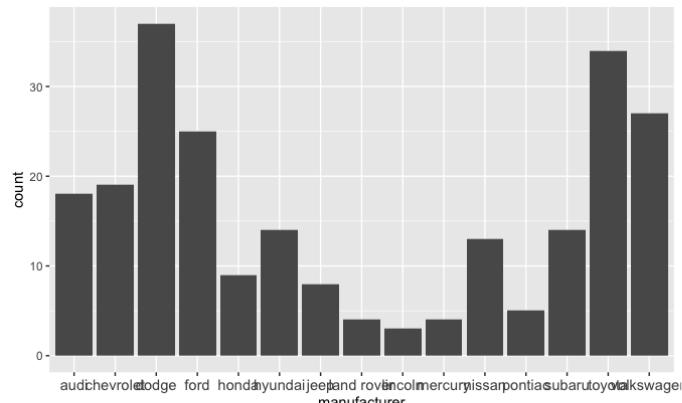


Illustration by Allison Horst

Cédric Scherer // rstudio::conf // July 2022



ggplot2 Examples featured on ggplot2.tidyverse.org

Cédric Scherer // rstudio::conf // July 2022

ggplot2:

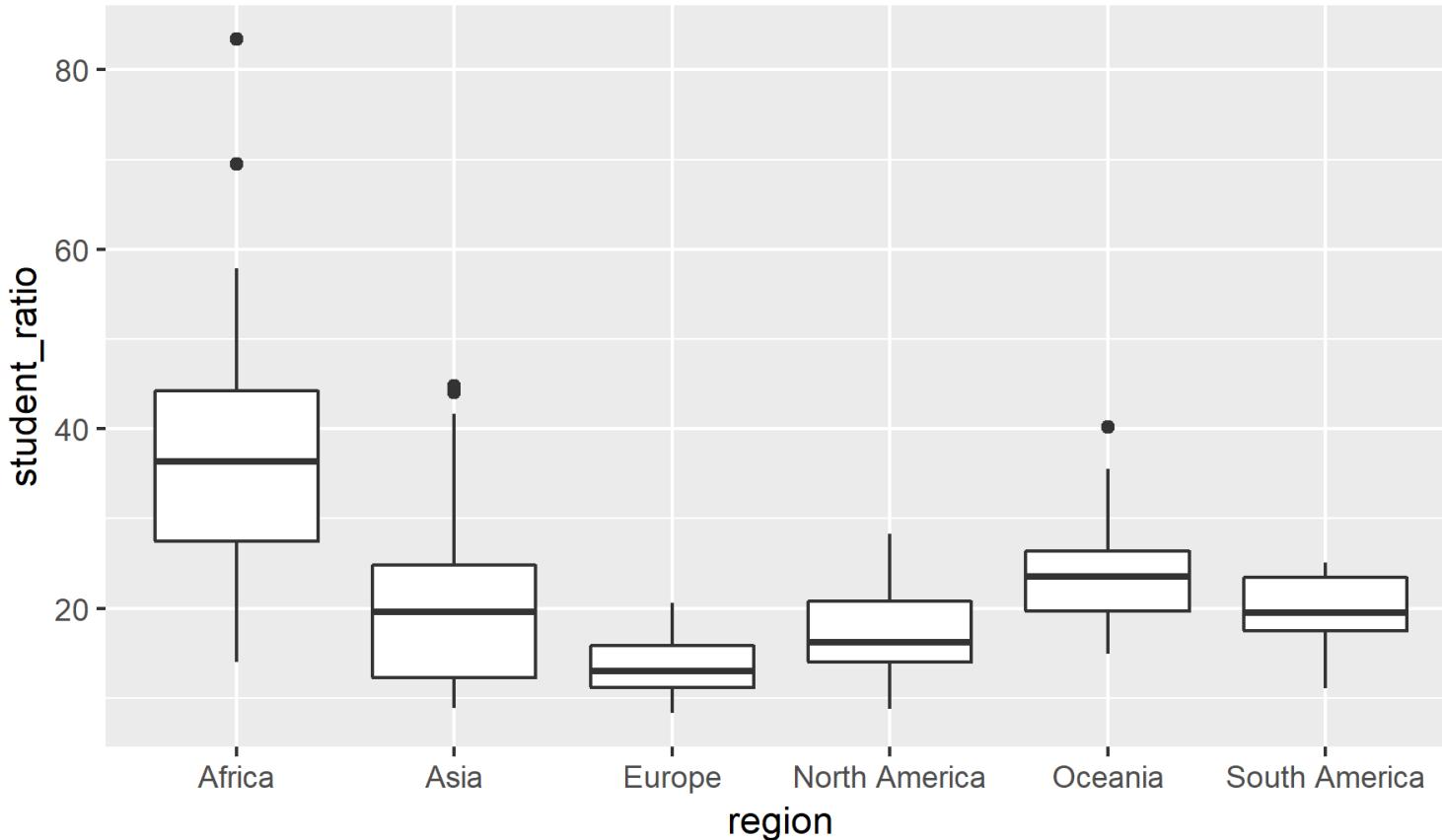
Build a data
MASTERpiece



Illustration by Allison Horst

Cédric Scherer // rstudio::conf // July 2022

The Evolution of a ggplot



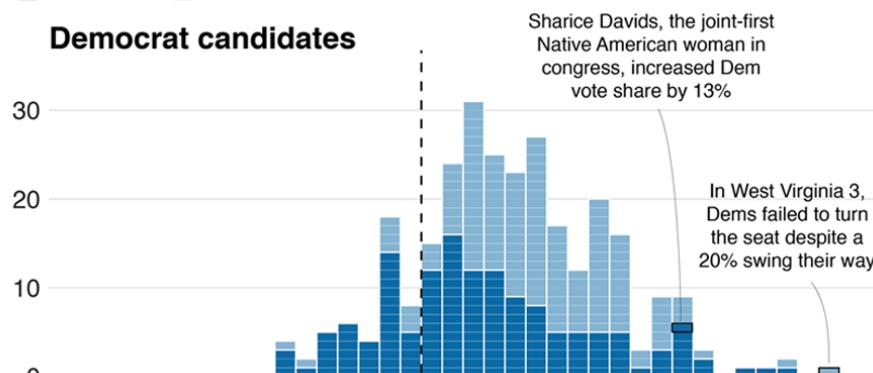
Data: UNESCO Institute for Statistics
Visualization by Cédric Scherer

The `ggplot2` Showcase

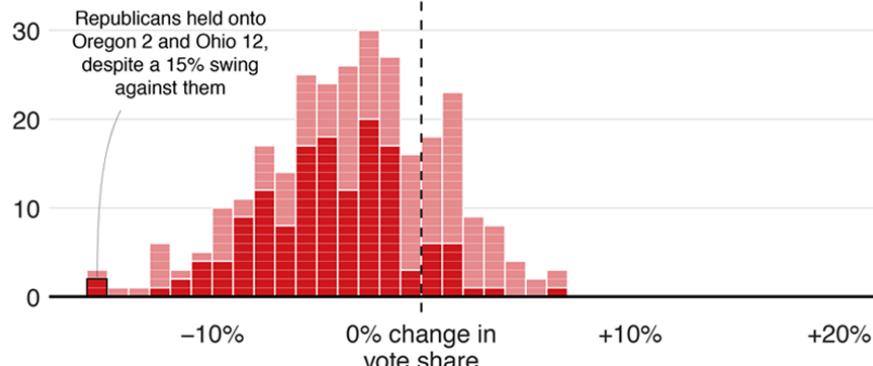
Blue wave

■ Won seat ■ Didn't win

Democrat candidates



Republican candidates

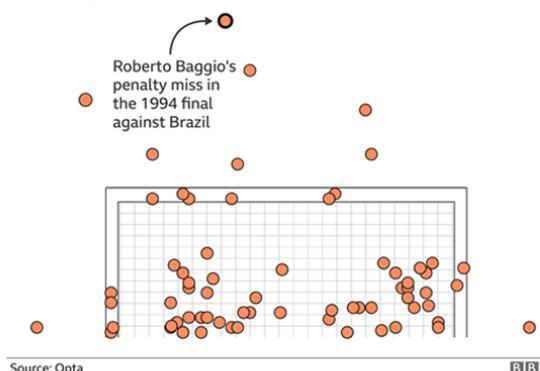


Source: AP, 19:01 ET

BBC

Where penalties are saved

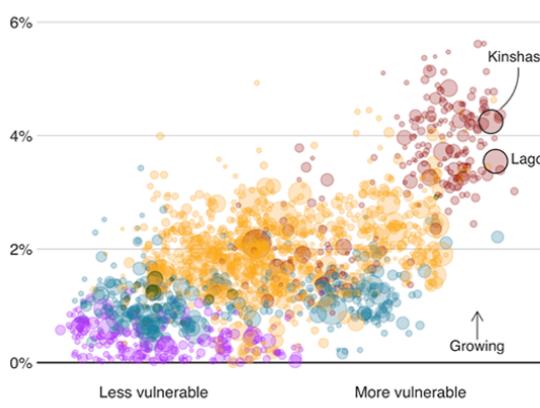
World Cup shootout misses and saves, 1982-2014



Fast-growing cities face worse climate risks

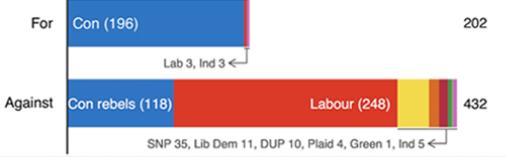
Population growth 2018-2035 over climate change vulnerability

■ Africa ■ Asia ■ Americas ■ Europe ■ Oceania



Source: Verisk Maplecroft. Circle size represents current population.

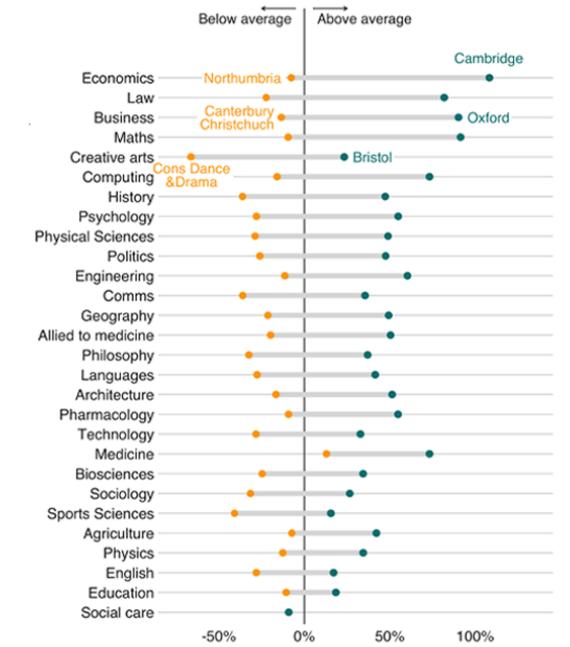
MPs rejected Theresa May's deal by 230 votes



Source: Commons Votes Services. Excludes 'tellers', the Speaker and deputies

Earnings vary across unis even within subjects

Impact on men's earnings relative to the average degree

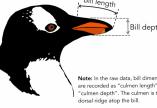


Source: Institute for Fiscal Studies

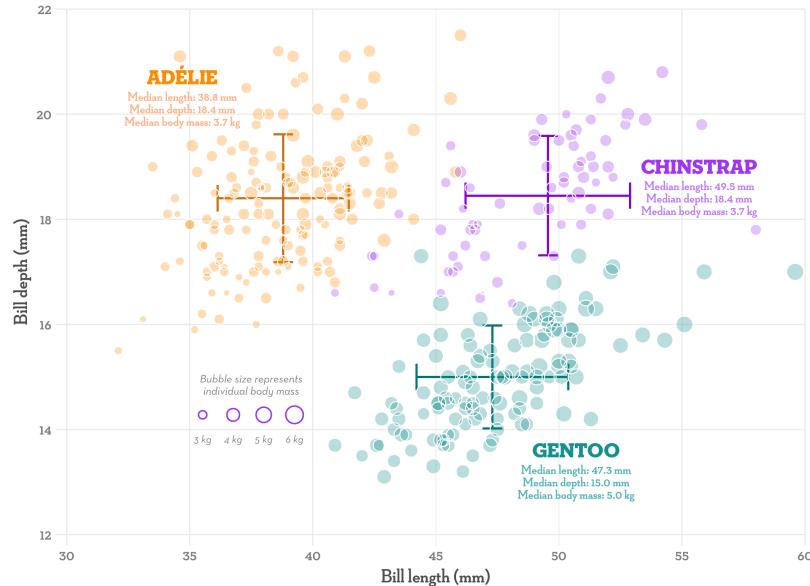
Collection of Graphics from the BBC R Cookbook

BILL DIMENSIONS OF BRUSH-TAILED PENGUINS

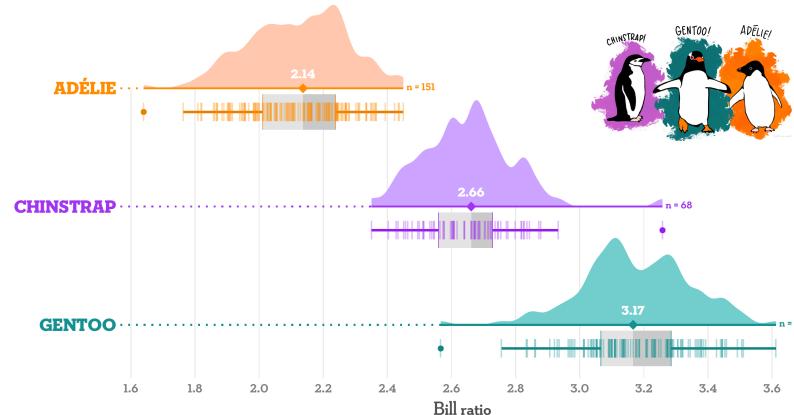
Pygoscelis adeliae (Adélie penguin) • *P. antarctica* (Chinstrap penguin) • *P. papua* (Gentoo penguin)



A. Scatterplot of bill length versus bill depth (median +/- sd)



B. Distribution of the bill ratio, estimated as bill length divided by bill depth



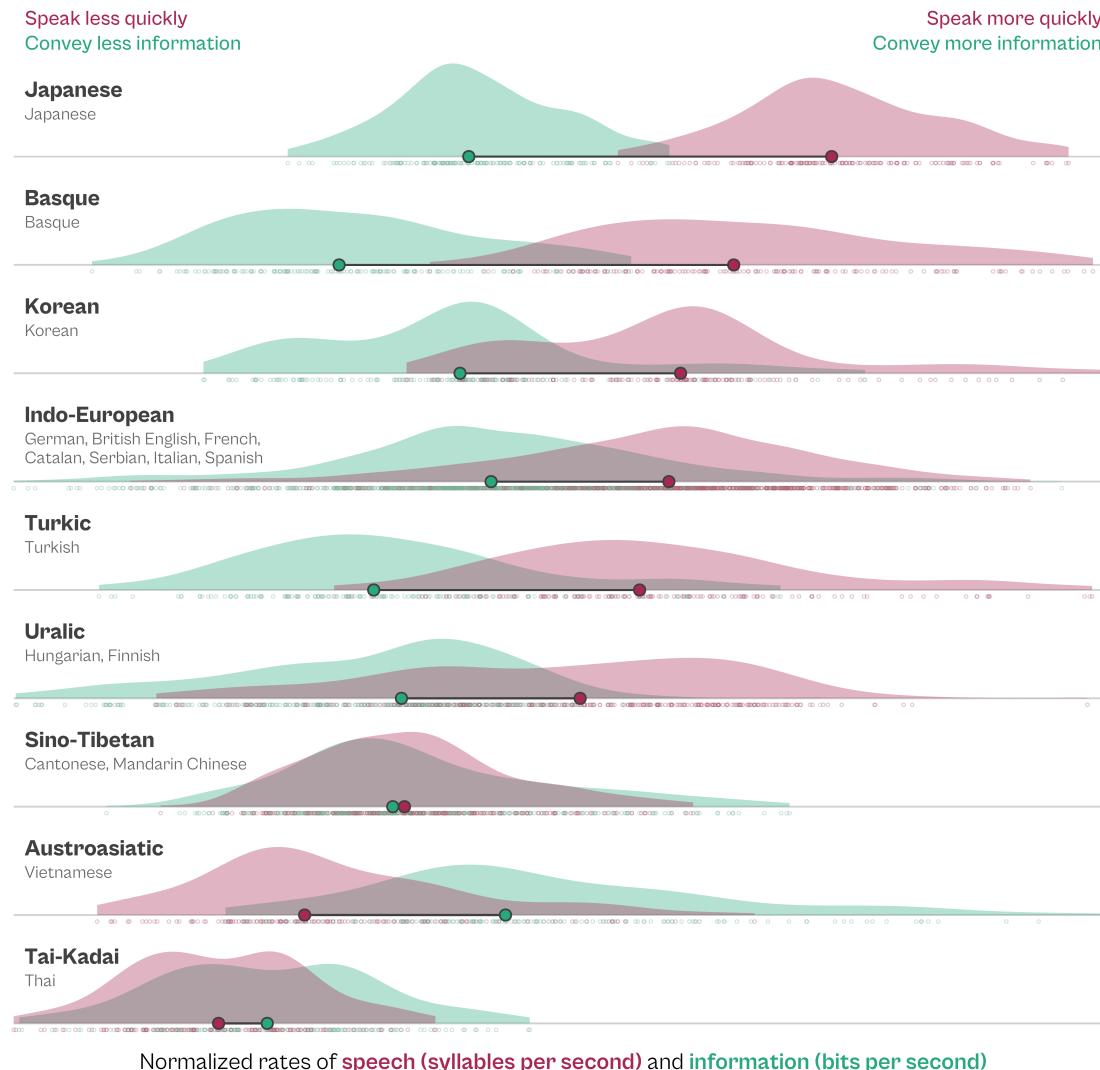
Note: In the original data, bill dimensions are recorded as "culmen length" and "culmen depth". The culmen is the dorsal (upper) ridge of a bird's bill.
Visualization: Cédric Scherer • Data: Gorman, Williams & Fraser (2014) DOI: 10.1371/journal.pone.0090081 • Illustrations: Allison Horst

Contribution to #TidyTuesday Week 2020/31

Cédric Scherer // rstudio::conf // July 2022

Communicating fast doesn't necessarily mean communicating more

Variation in speech and information rates across language families, shown as normalized rates for comparison. While there are stark cross-linguistic differences in speech rates, information rates are more similar.



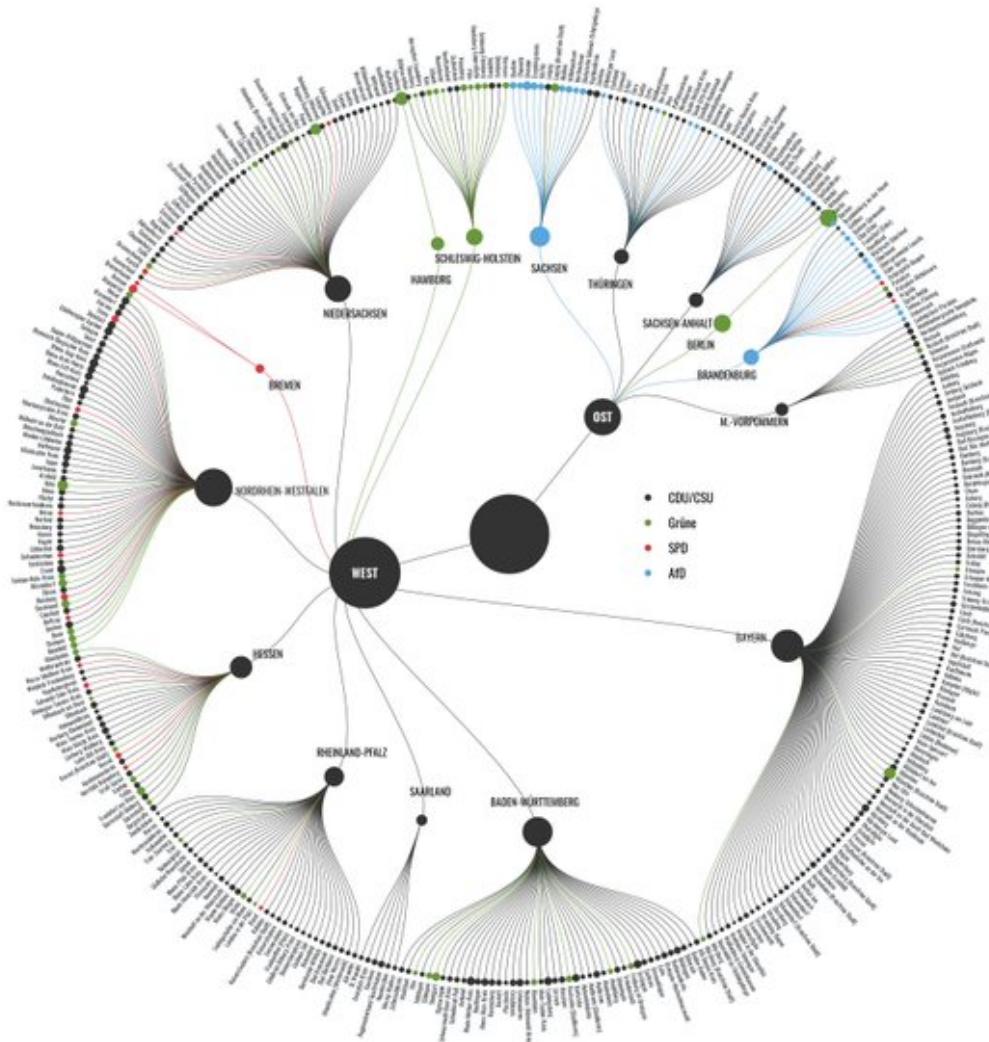
Source: Coupé et al. 2019 Science Advances 5(9). DOI: 10.1126/sciadv.aaw2594

Graphic: Cédric Scherer • Large dots show the median rates for each language family. Small dots show single estimates.

My reinterpreted The Economist graphic

Cédric Scherer // rstudio::conf // July 2022

Europawahl 2019
Wahlergebnisse Deutschland

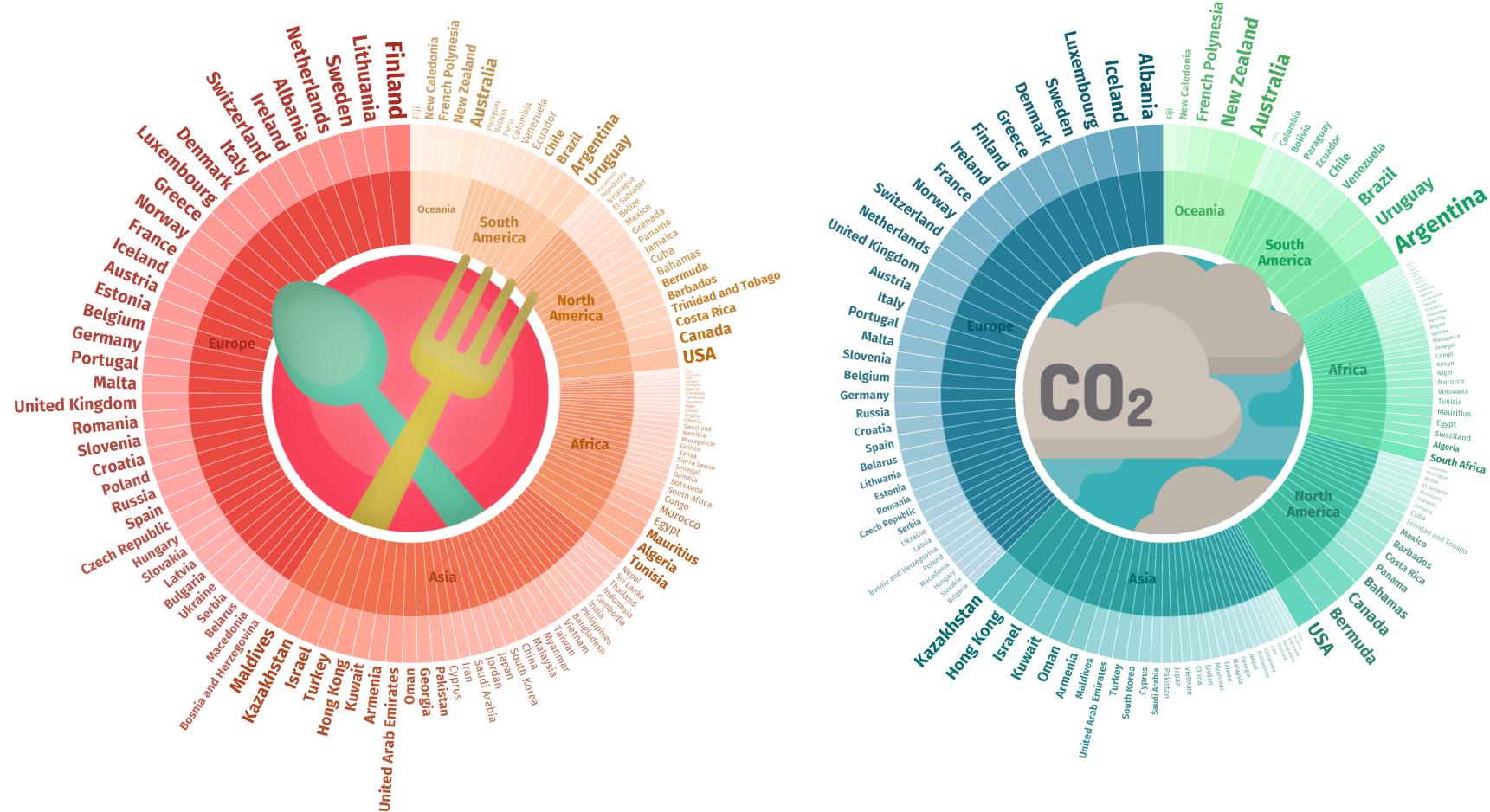


© 2019 sunburst | Statis kommunikation

“European Elections” by Torsten Sprenger
Cédric Scherer // rstudio::conf // July 2022

Food Carbon Footprint Index 2018

Global comparison of different diets in terms of **Average Consumption (kg/person/year)** of both animal and non-animal products as well as **Carbon Emissions (kg CO₂/person/year)** per continent and country. Font size and color intensity indicate each country's estimate with **countries printed in bold** belonging to the upper 50% of consumers and CO₂ emitters, respectively.



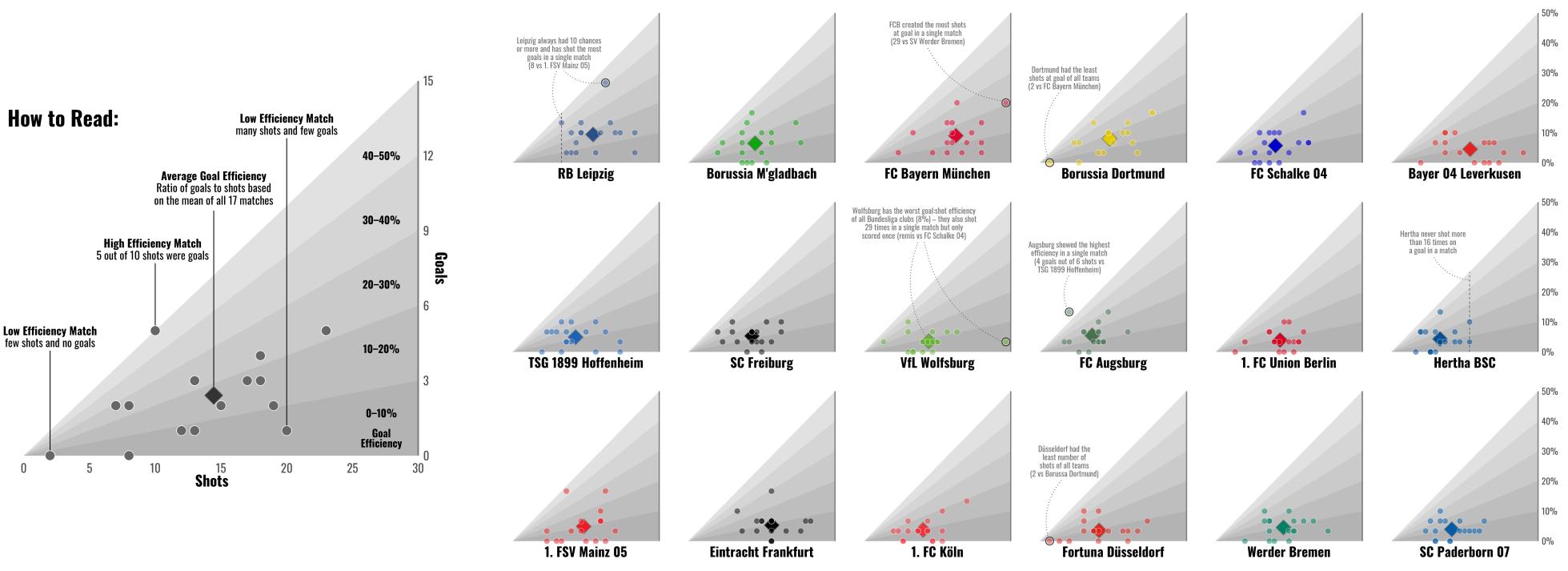
Visualization by Cédric Scherer • Data by Food and Agriculture Organization of the United Nations (FAO) via nu3 • Icons by FreePik

Contribution to #TidyTuesday Week 2020/08

Cédric Scherer // rstudio::conf // July 2022

RB Leipzig and Borussia Dortmund Make the Most of Their Opportunities

The small multiples show each club's goal:shot efficiency in the first season 2019/2020 of the 1. Bundesliga. However, while Borussia Dortmund also had very bad matches with only 2 chances, the Autumn champion RB Leipzig always scored minimum one goal and shot at least ten times on the opponent's goal in all 17 matches! Of all Bundesliga clubs, RB Leipzig also shot the most goals – 8 against Mainz.

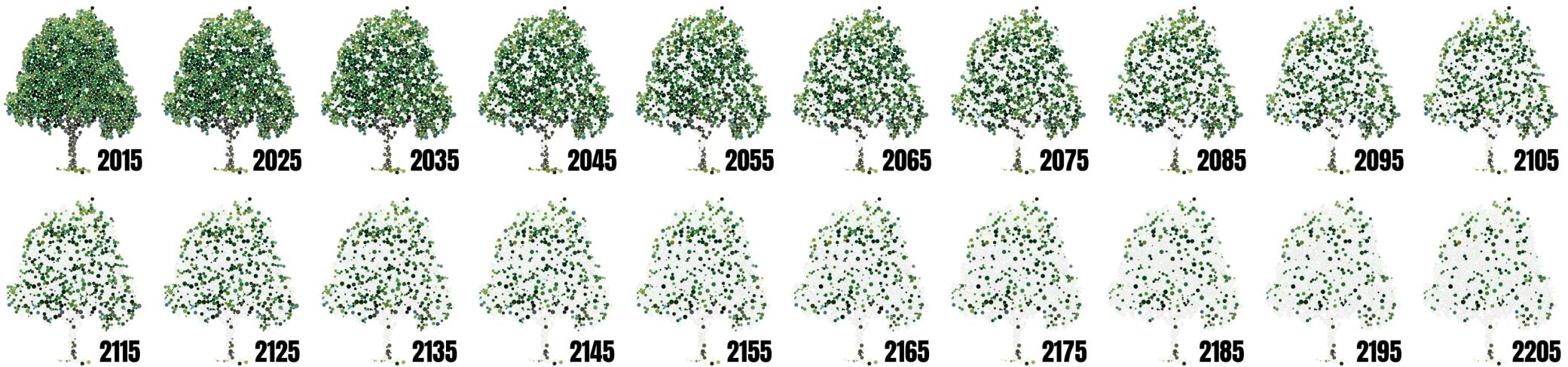


Visualization by Cédric Scherer

Contribution to the #SWDchallenge "Small Multiples"

Unless someone like you cares a whole awful lot, nothing is going to get better. It's not.

In a 2015 study, published in Nature, Thomas Crowther and colleagues mapped tree density across the world. They estimated that there were approximately 3.04 trillion trees in the world. The authors also estimated that over 15 billion trees are cut down each year, and the global number of trees has fallen by almost half (46%) since the start of human civilization. Each dot on the trees below represent one billion trees. Each year 15 dots are faded out to illustrate the progressive loss. At this rate, our forests will evaporate within 200 years.

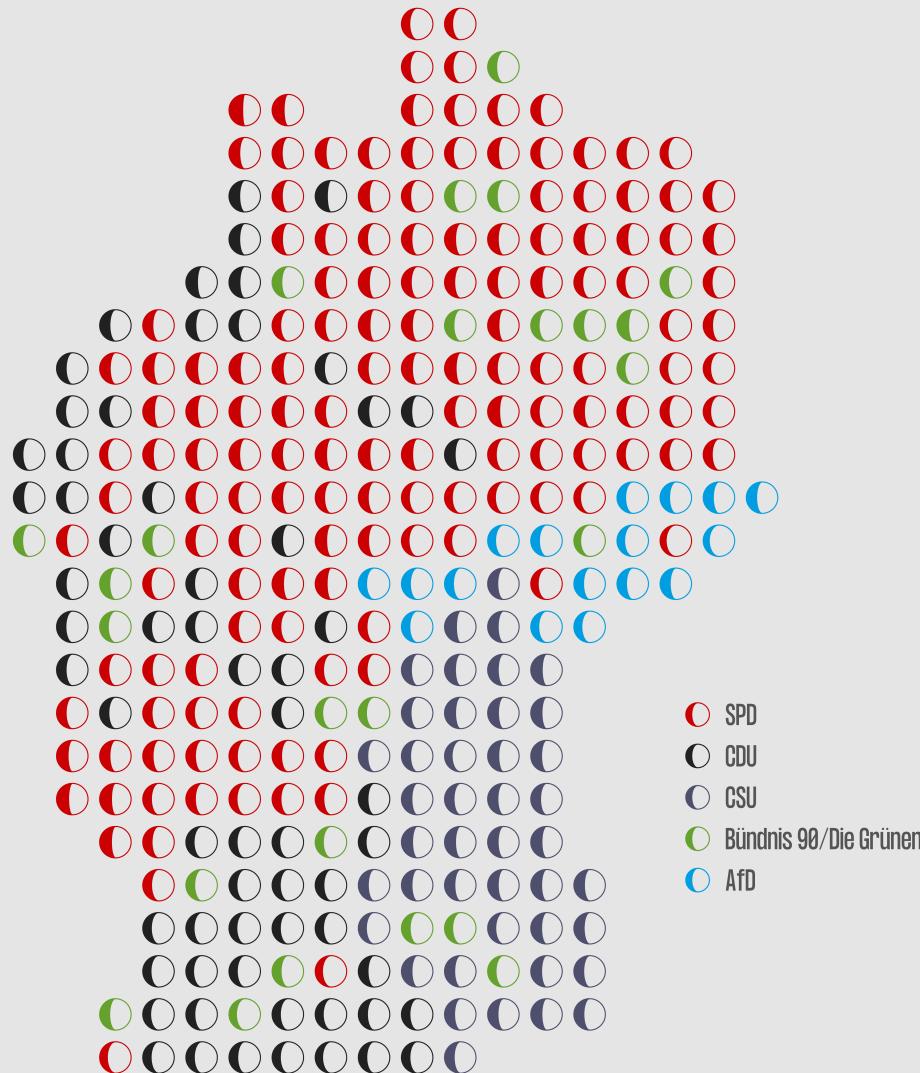


Data: ourworldindata.org/forests | Article: doi.org/10.1038/nature14967 | Graphic: @jakekaupp

Contribution to the #TidyTuesday Week 2021/15 x #30DayChartChallenge 2021 by Jake Kaupp

Ergebnisse der Bundestagswahl 2021

Die stärksten Parteien nach Prozent der Zweitstimmen.



- SPD
- CDU
- CSU
- Bündnis 90/Die Grünen
- AfD

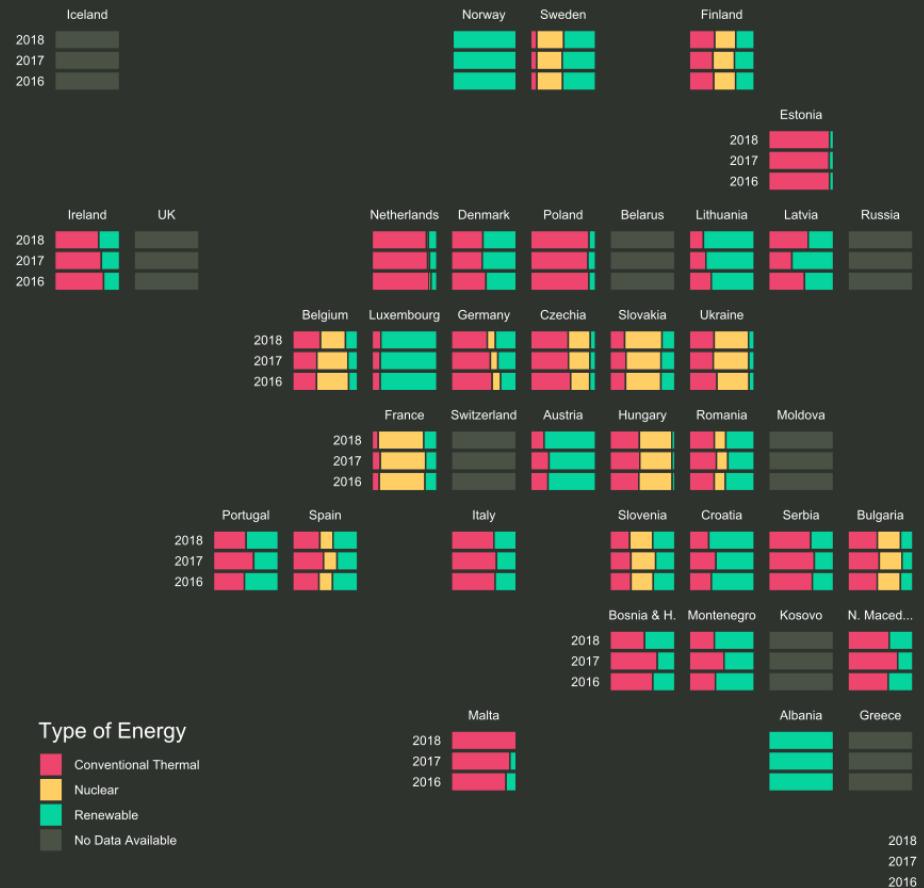
Grafik: Cédric Scherer · Daten: DIE ZEIT · Gridkarte: Ansgar Wolsing & Cédric Scherer

Moon Charts as a Tile Grid Map showing the 2nd Vote Results from the German Election 2021

Cédric Scherer // rstudio::conf // July 2022

EUROPEAN ENERGY GENERATION

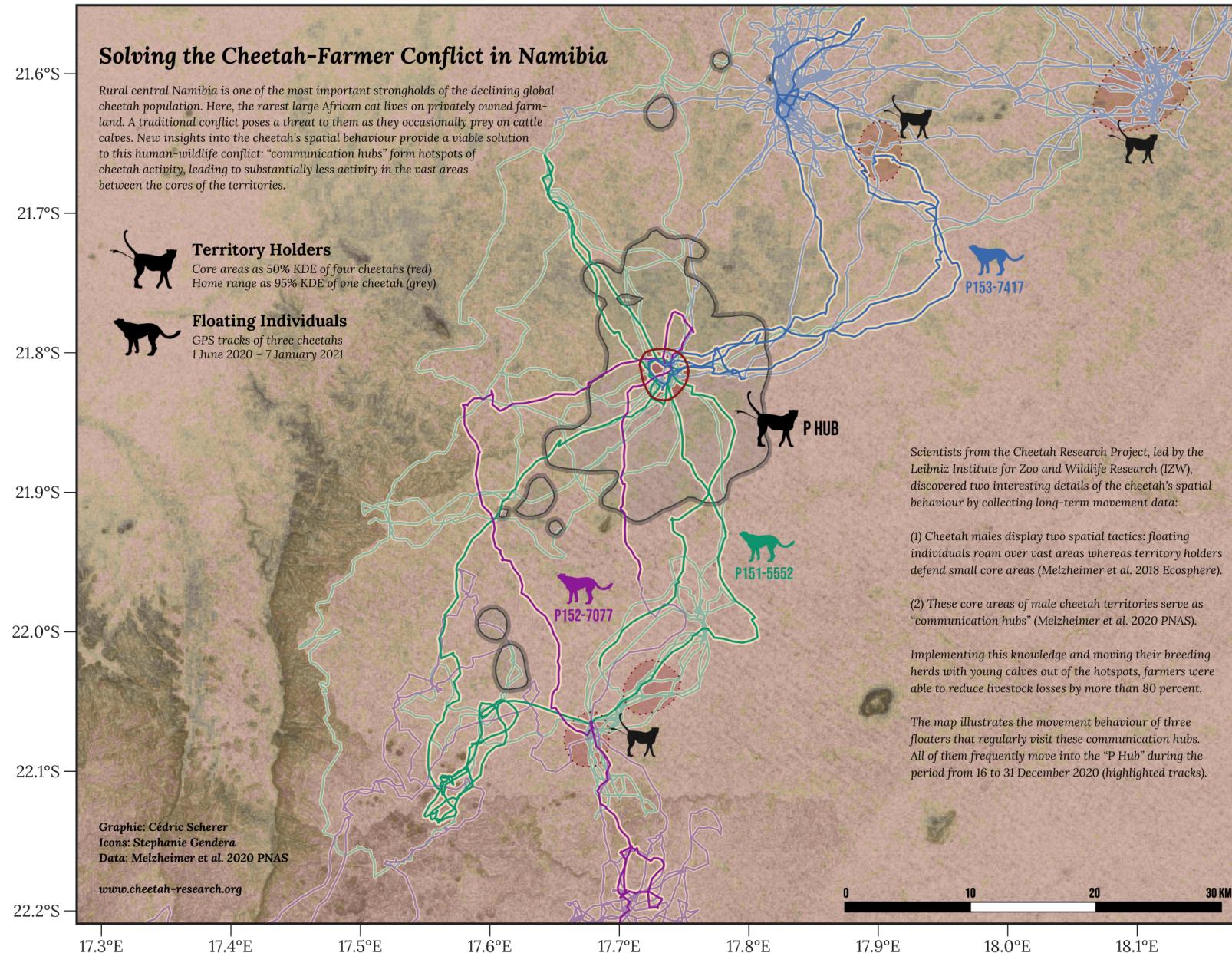
Each bar represents the **total energy generation** for each country per year.
The colours represent the proportion of energy generated
a) using **conventional thermal power plants**, which is to say those that use coal, oil or natural gas,
b) using **nuclear power stations**, and c) using other **renewable sources**.



Data from 'Electricity generation statistics - First Results' (ec.europa.eu/eurostat/statistics-explained)
Visualisation by [Jack Davison](#) ([@JDavison_](https://twitter.com/JDavison_))
Code found at github.com/jack-davison

#TidyTuesday Contribution by Jack Davison

Cédric Scherer // rstudio::conf // July 2022

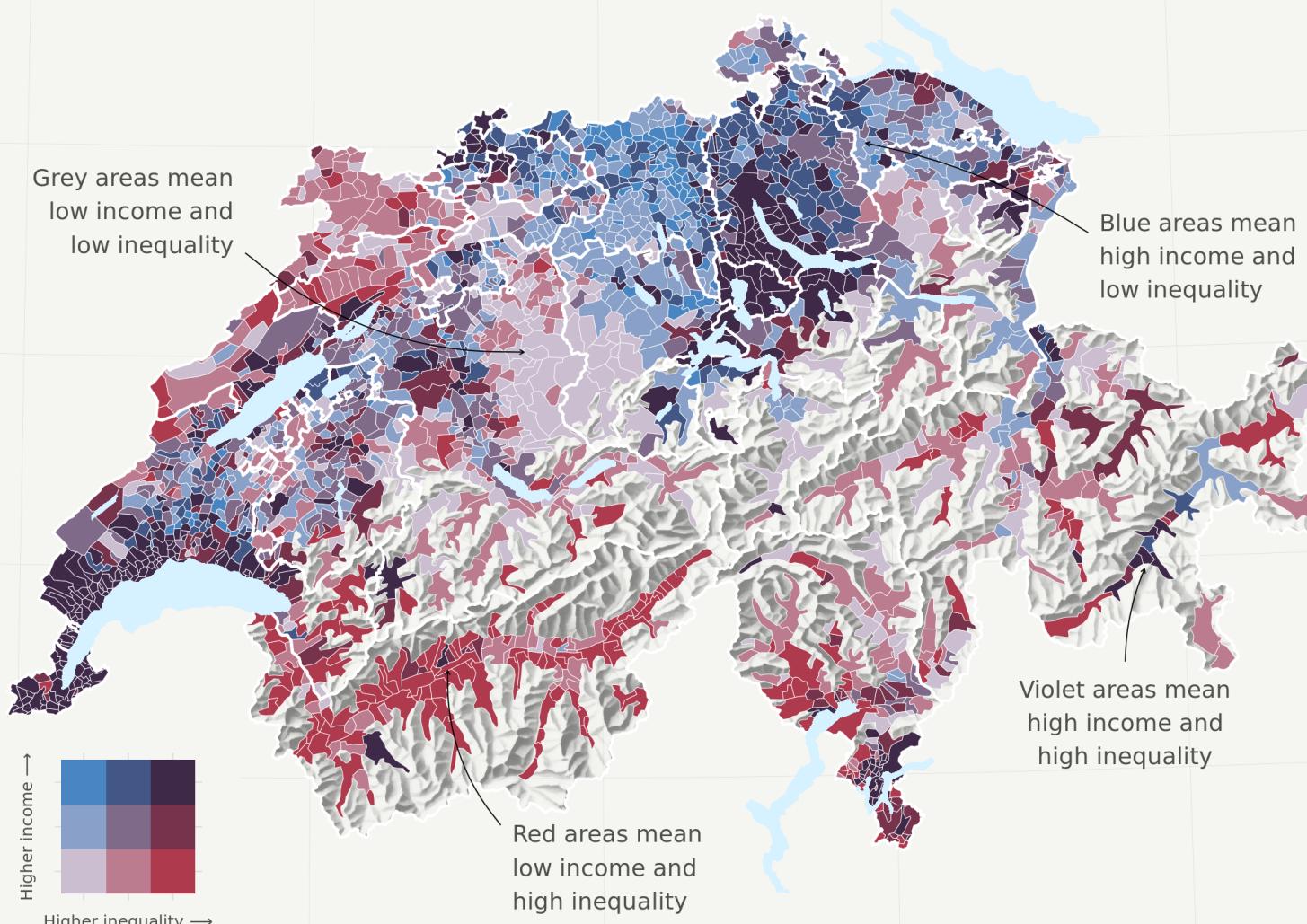


Our Winning Contribution to the BES MoveMap Contest

Cédric Scherer // rstudio::conf // July 2022

Switzerland's regional income (in-)equality

Average yearly income and income (in-)equality in Swiss municipalities, 2015



Map CC-BY-SA; Code: github.com/grssnbchr/bivariate-maps-ggplot2-sf
Authors: Timo Grossenbacher (@grssnbchr), Angelo Zehr (@angelozehr)
Geometries: ThemaKart BFS and swisstopo; Data: ESTV, 2015

Bivariate Choropleth x Hillshade Map by Timo Grossenbacher

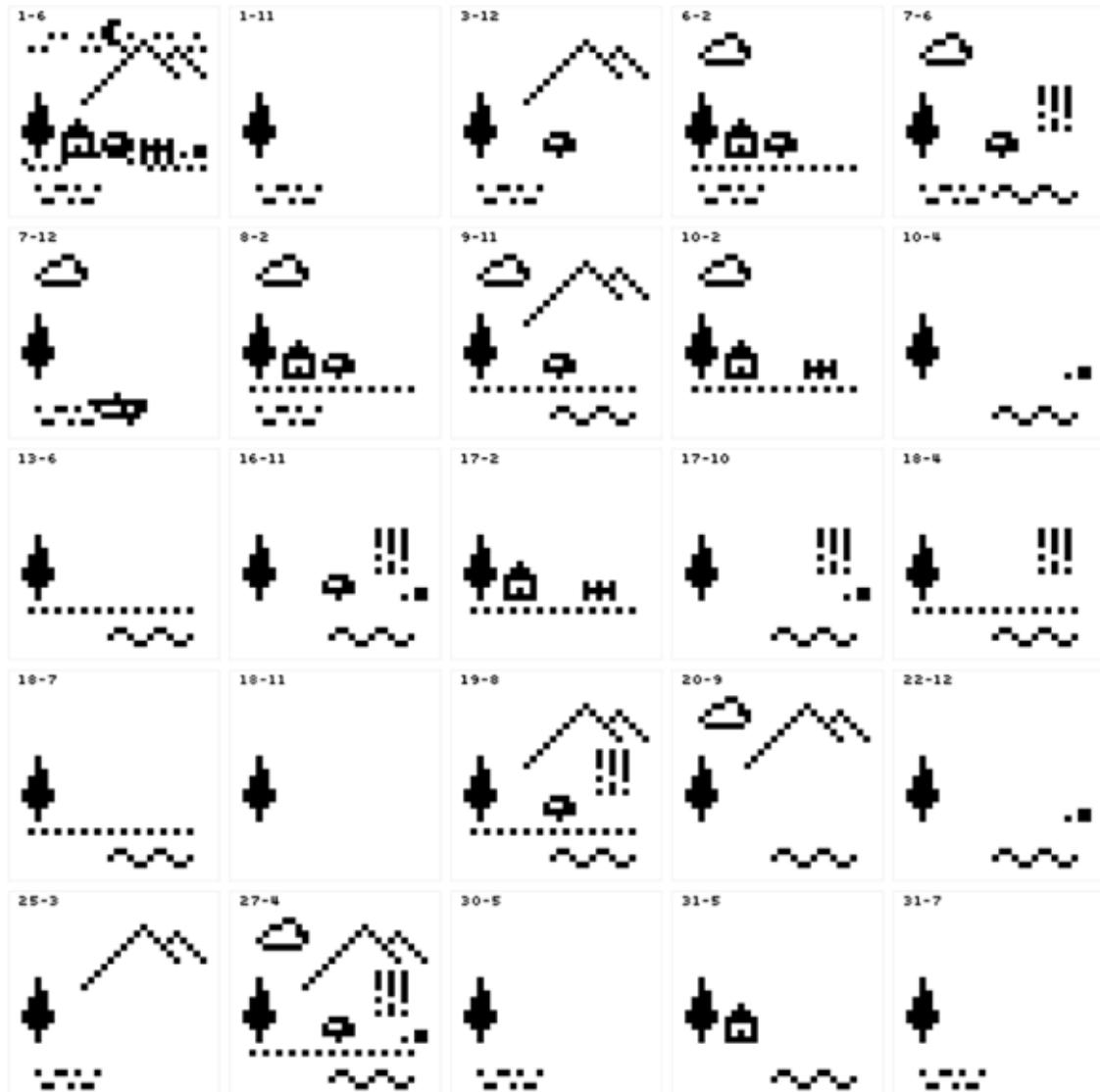
Cédric Scherer // rstudio::conf // July 2022

BOB ROSS - PAINTING BY THE ELEMENTS

GRAPHIC REPRESENTATIONS OF BOB ROSS' PAINTINGS WITH ELEMENTS IDENTIFIED IN THEM BY WALT HICKEY (FIVETHIRTYEIGHT). EACH ELEMENT REPRESENTS ONE OR MORE OCCURRENCES IN THE PAINTING. ONLY THE ELEMENTS IN THE LEGEND BELOW ARE DRAWN. TO THE RIGHT THERE ARE 25 RANDOM PAINTINGS THAT BOB PAINTED IN 'THE JOY OF PAINTING', WITH THE SEASON AND EPISODE NUMBER.



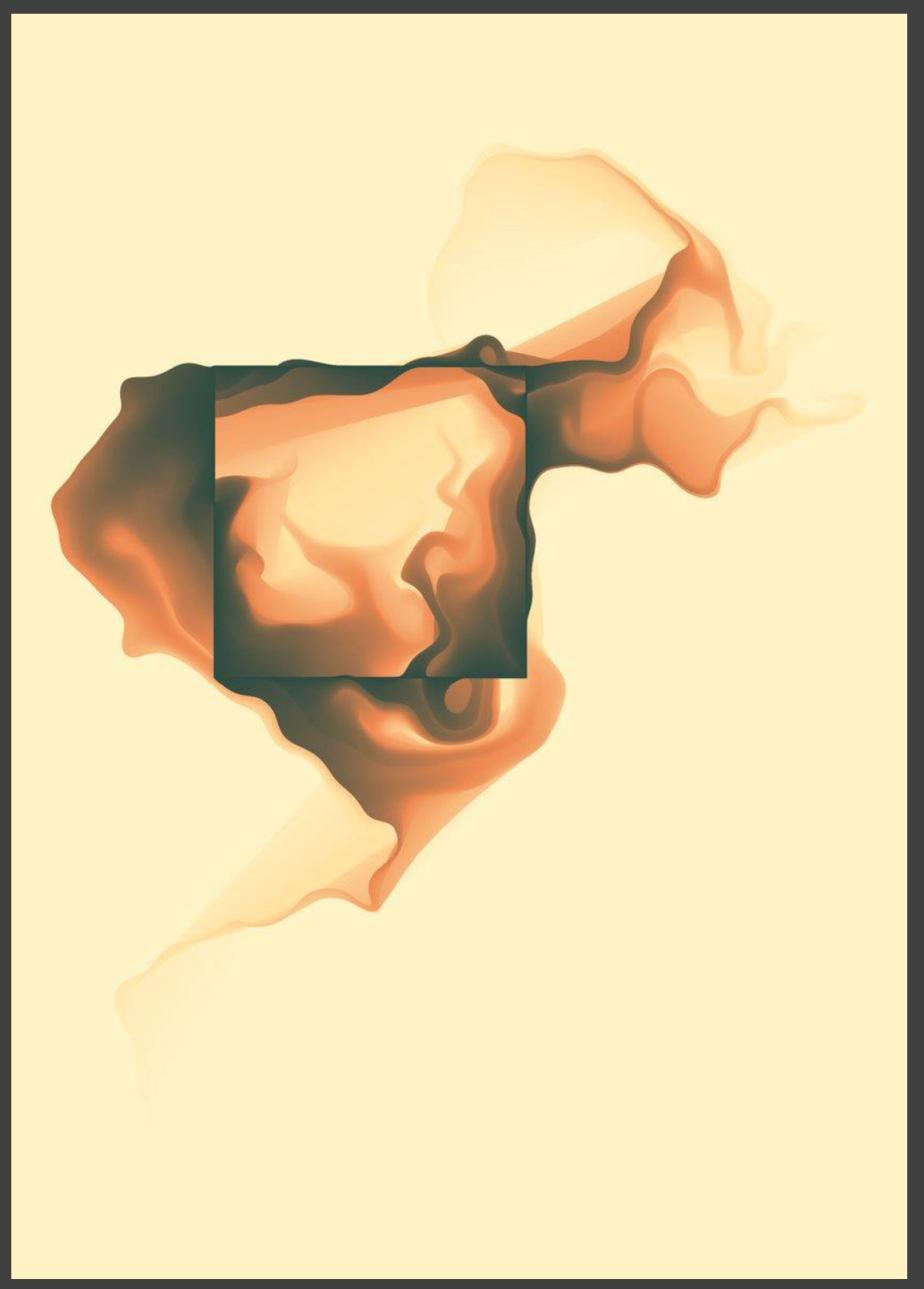
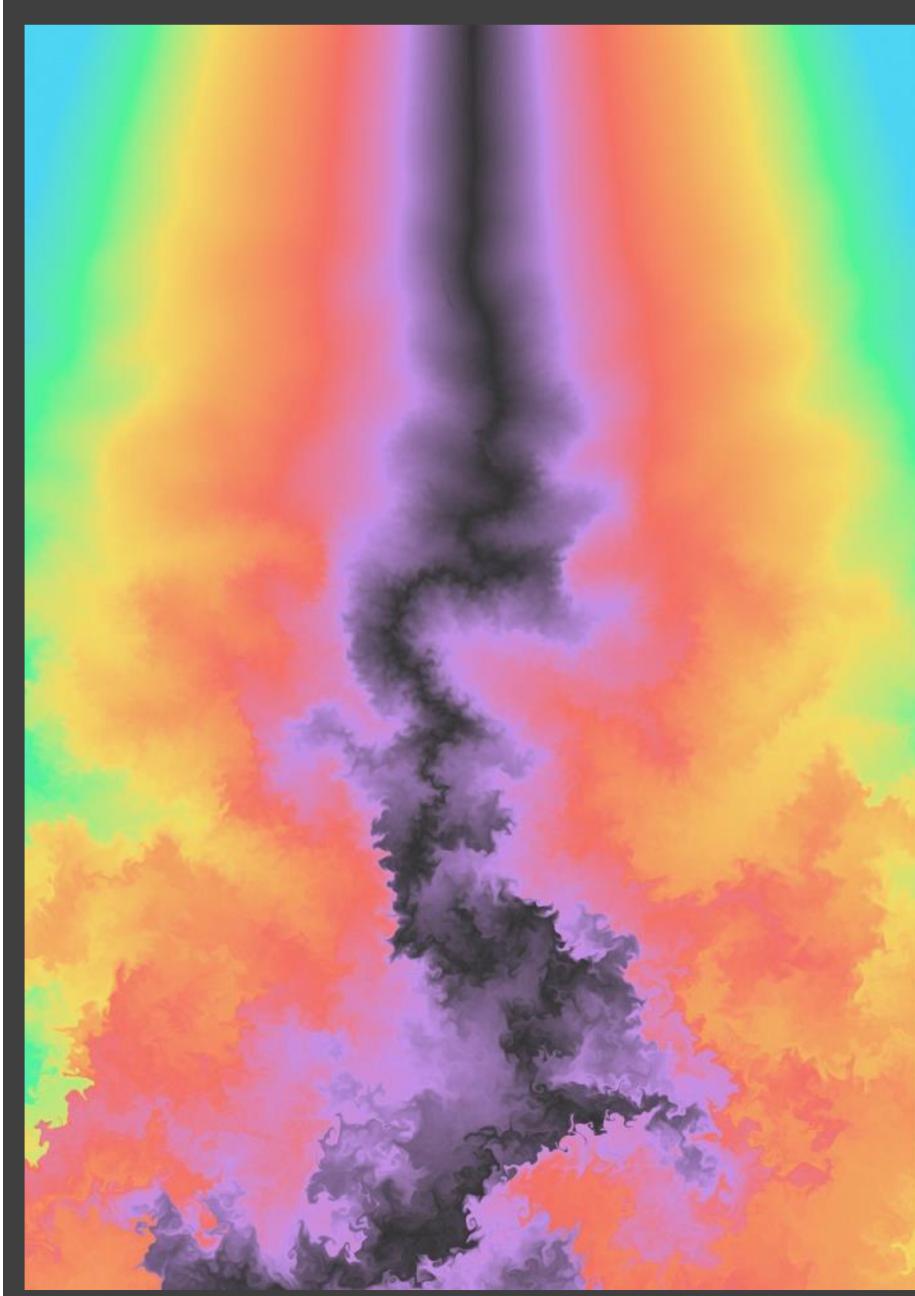
CLOUD (CIRRUS, CUMULUS), MOON, NIGHT, PALM TREE, SUN, MOUNTAIN (HILLS), CACTUS, LIGHTHOUSE, ROCKS, TREE (DECIDUOUS, CONIFER), BEACH, BOAT, CABIN (BARN, BUILDING, FARM), SEA (OCEAN, WAVES), BUSHES, FENCE, GRASS, WATERFALL, LAKE, RIVER



SOURCE: FIVETHIRTYEIGHT | PLOT: GEORGIOS KARAMANIS

Pixel art by [Georgios Karamanis](#)

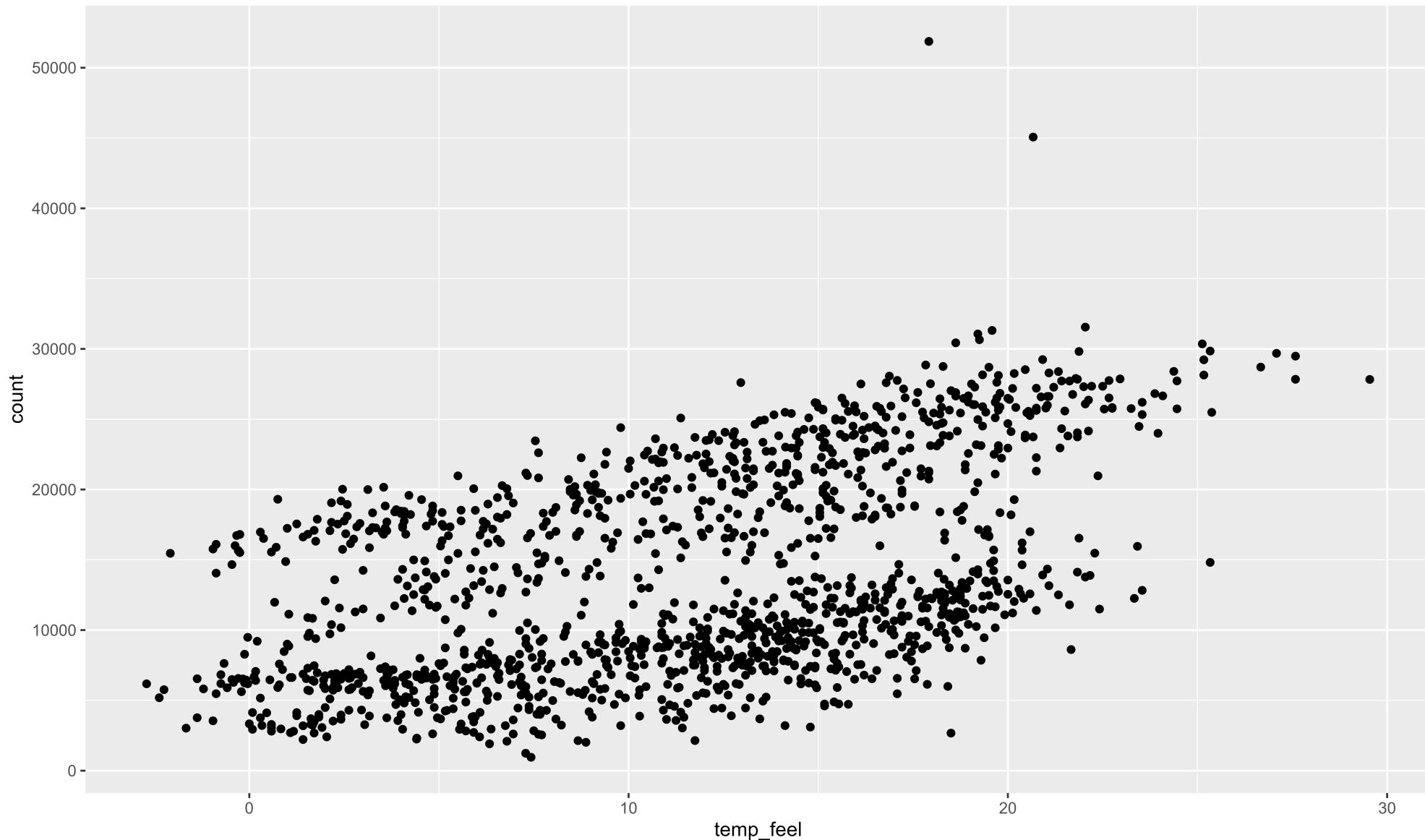
[Cédric Scherer](#) // rstudio::conf // July 2022

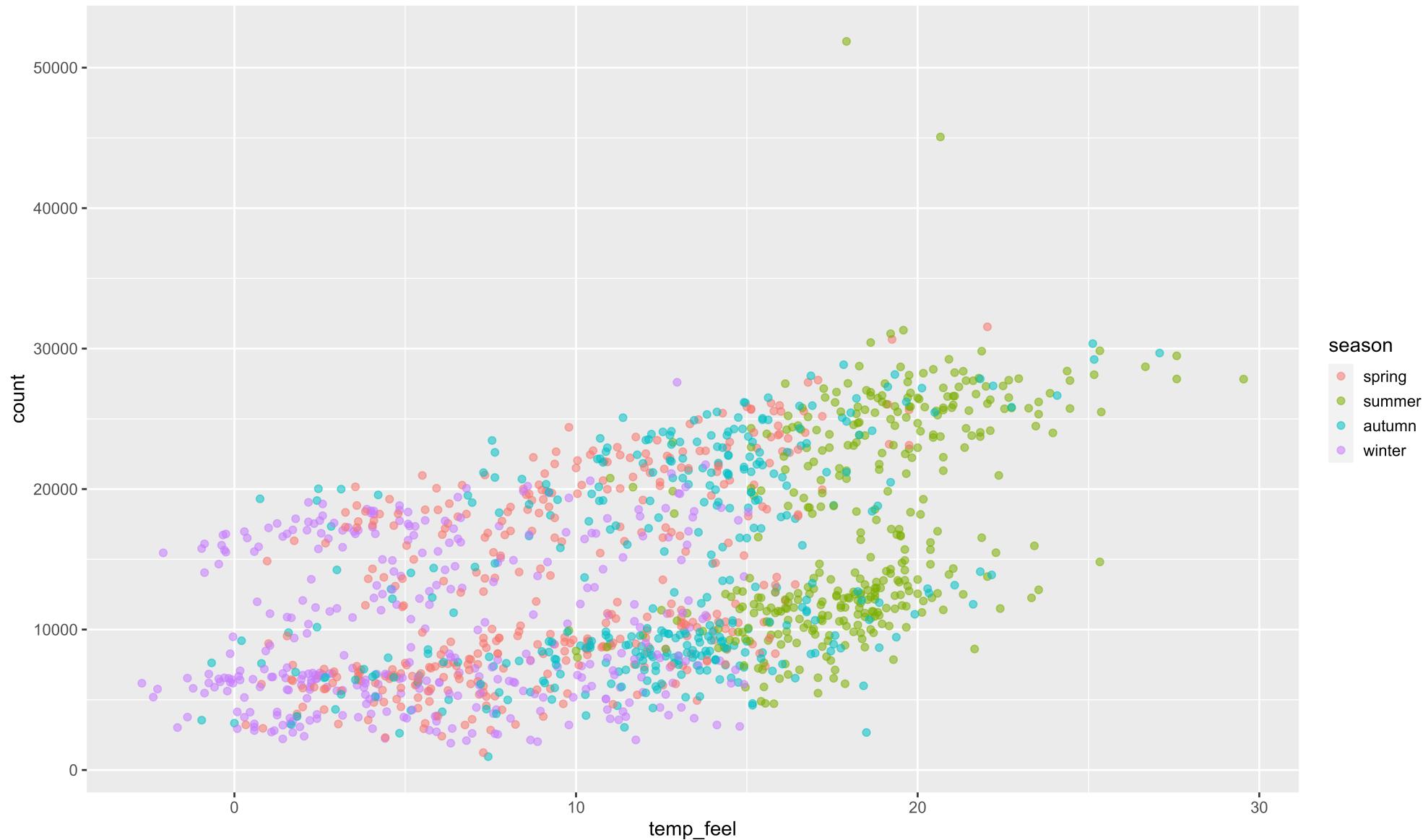


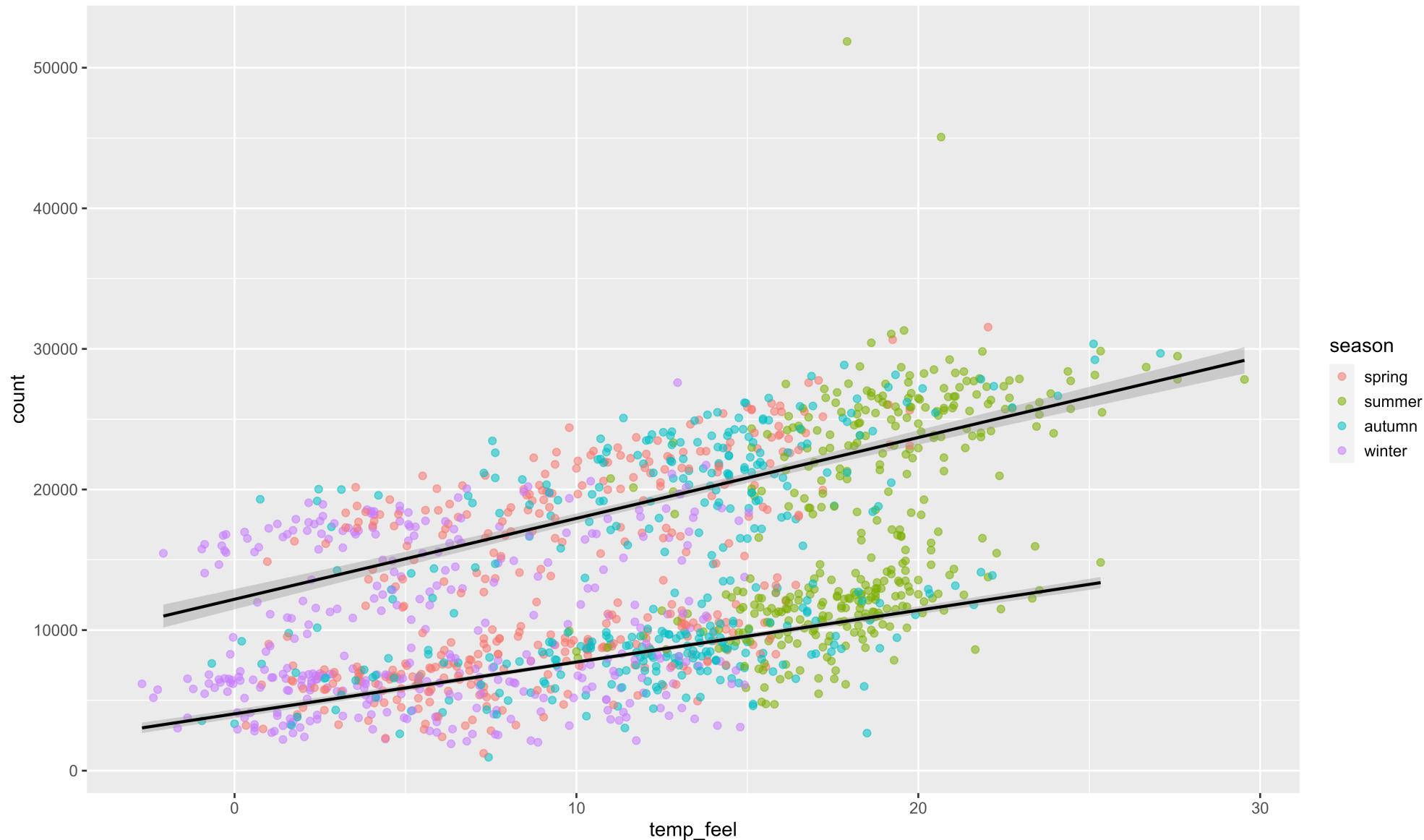
Generative art by Thomas Lin Pedersen

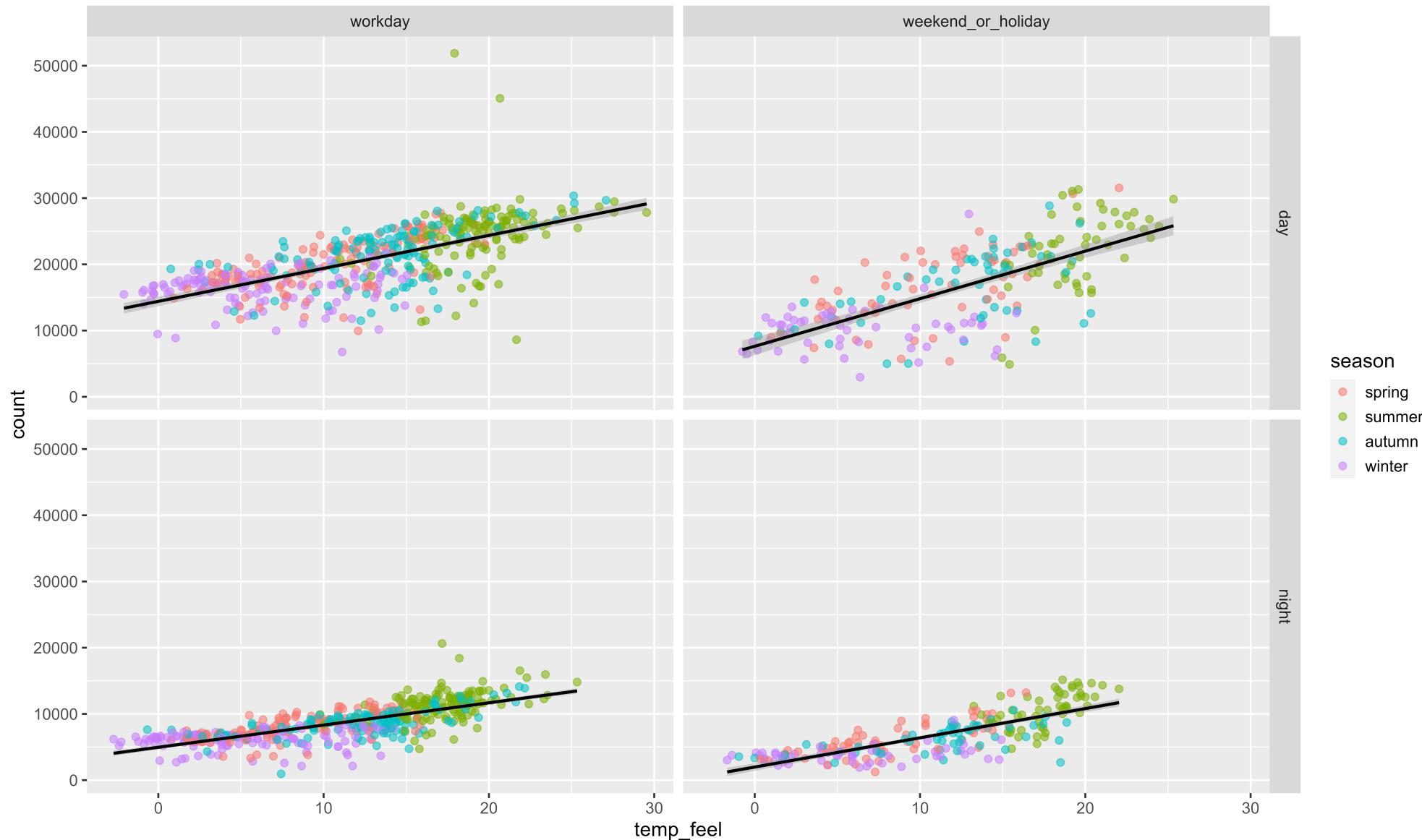
Cédric Scherer // rstudio::conf // July 2022

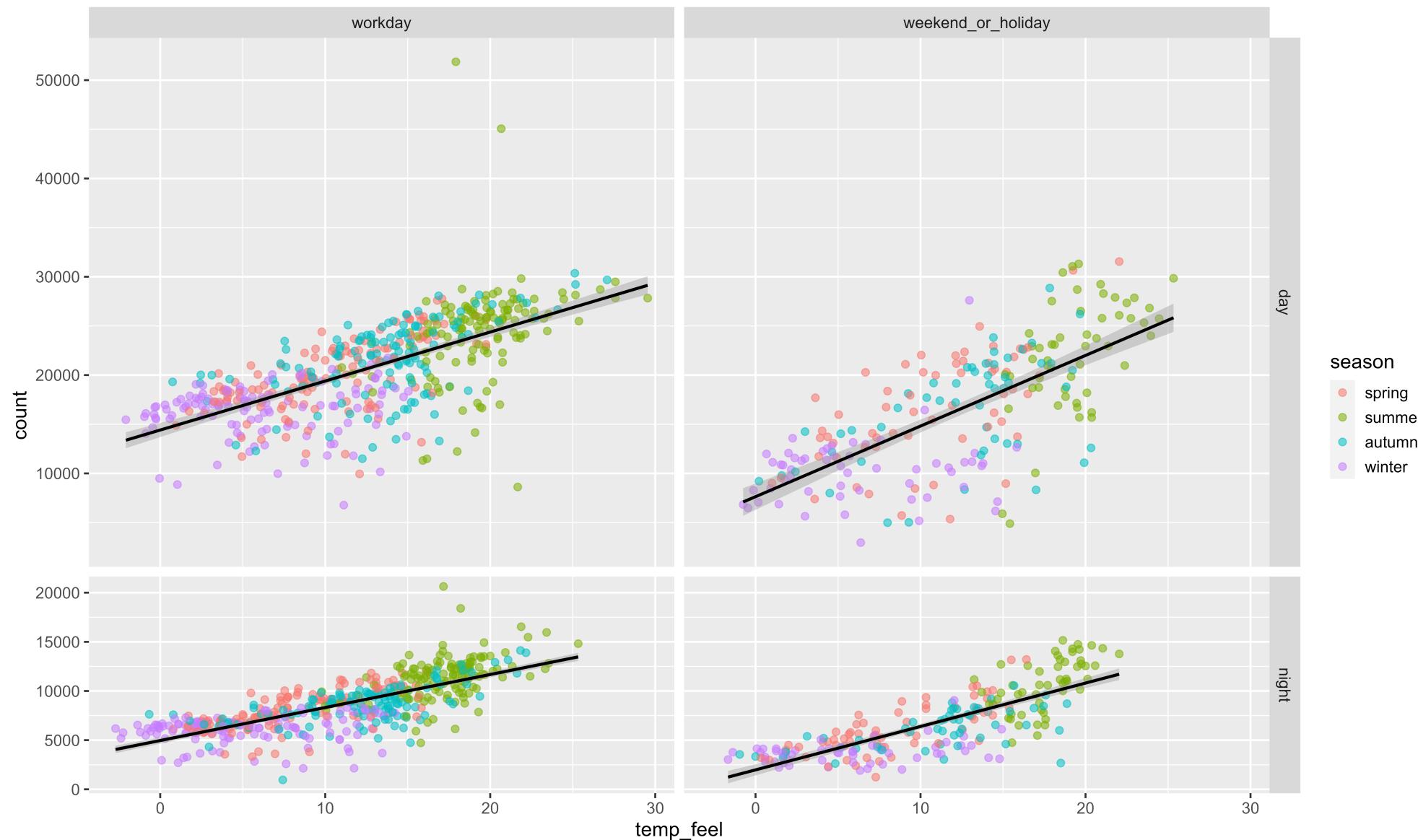
A Step-by-Step Example



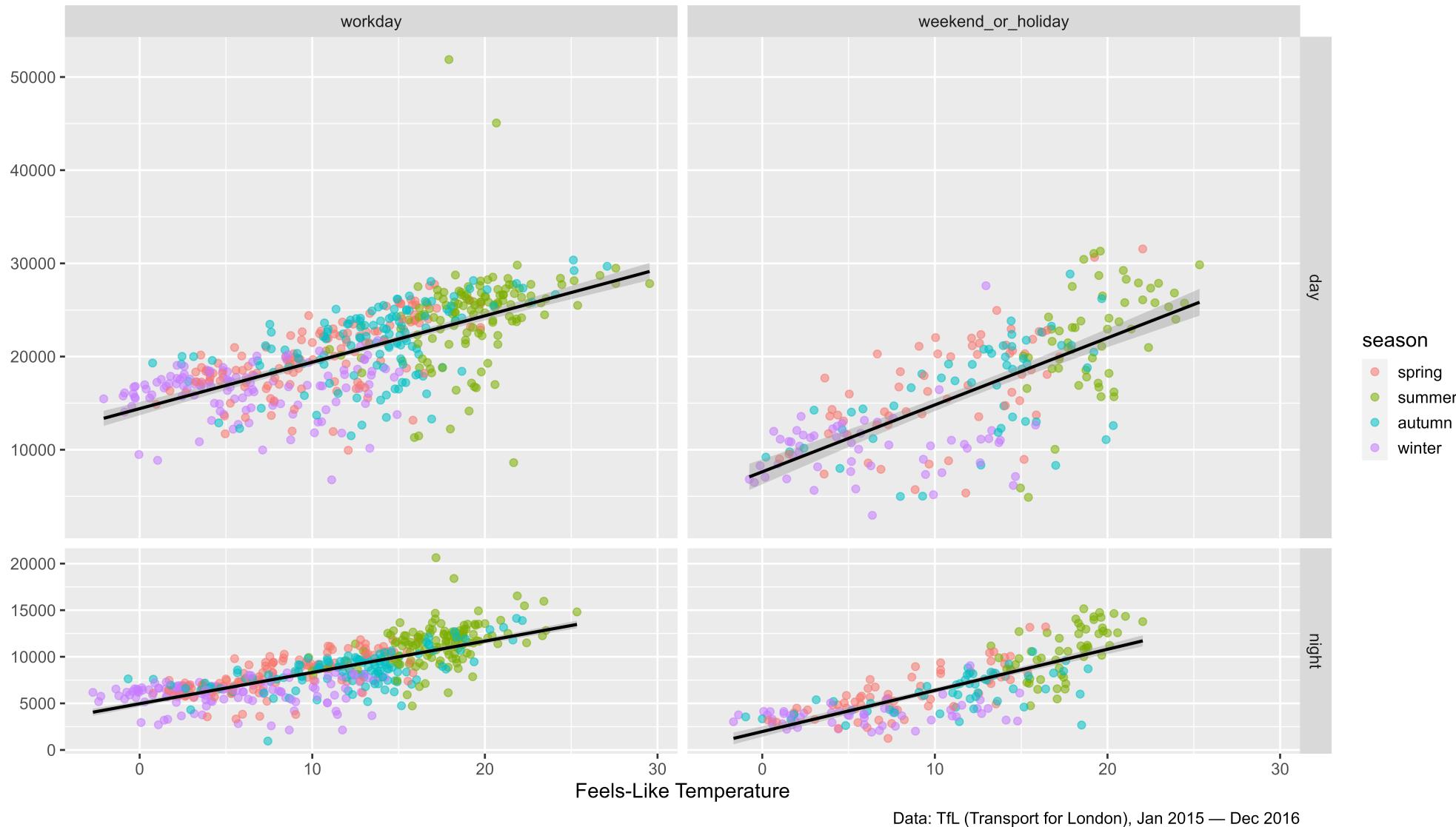




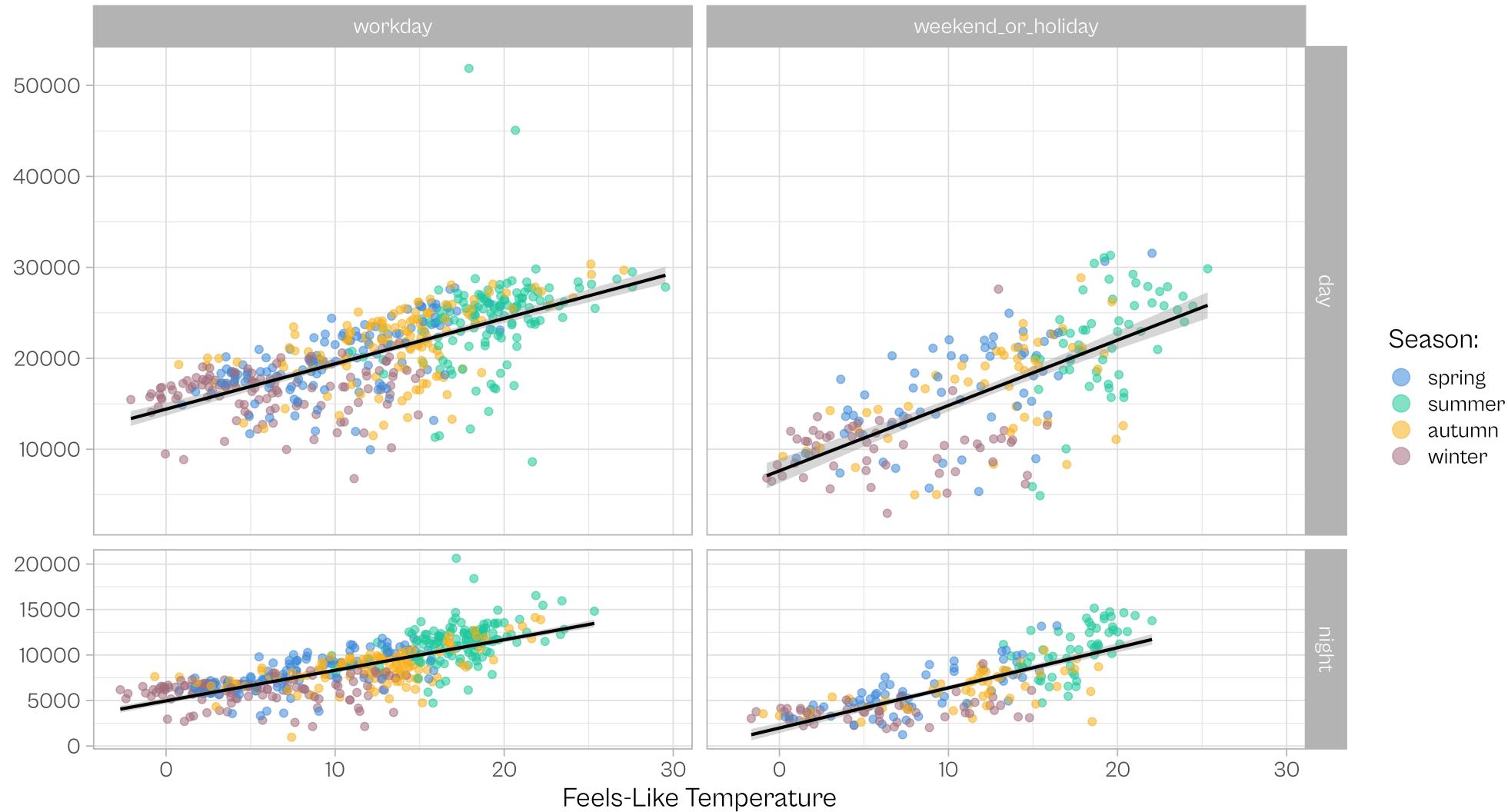




Reported bike rents versus feels-like temperature in London per time of day, period, and season.

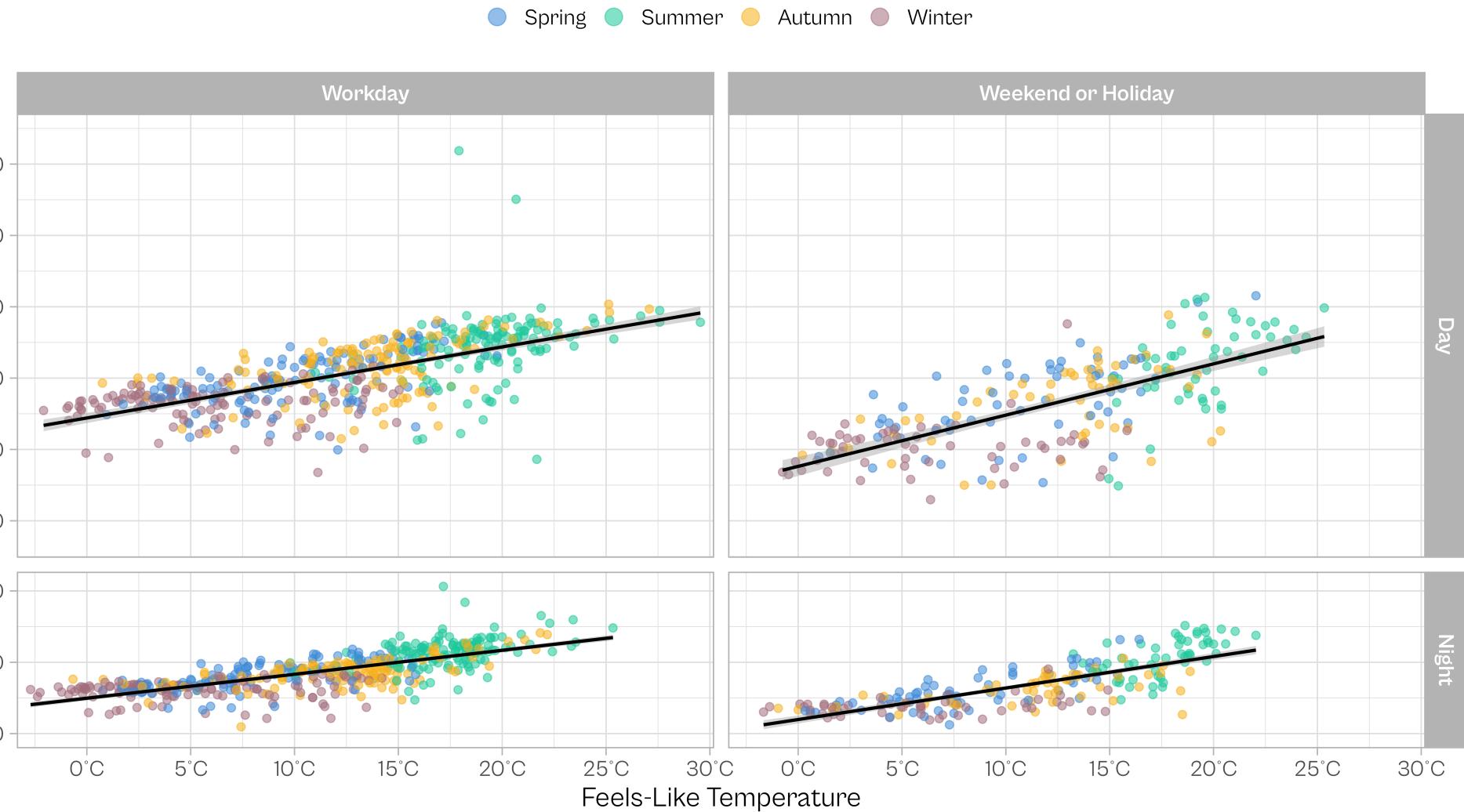


Reported bike rents versus feels-like temperature in London per time of day, period, and season.



Data: TfL (Transport for London), Jan 2015 – Dec 2016

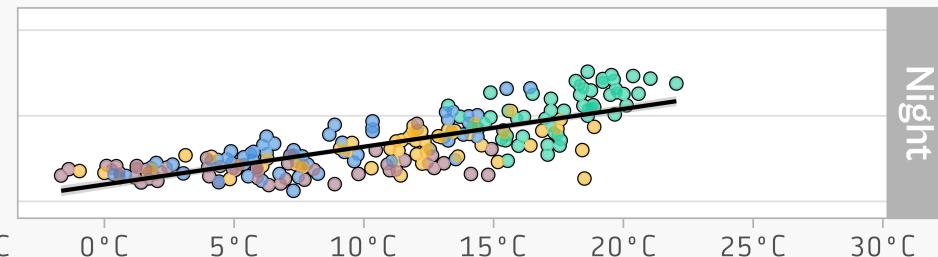
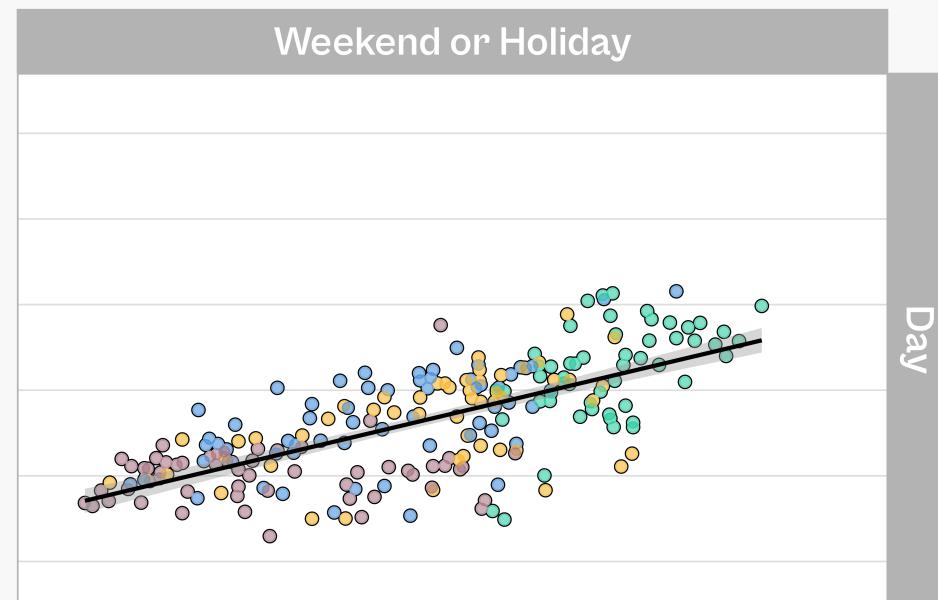
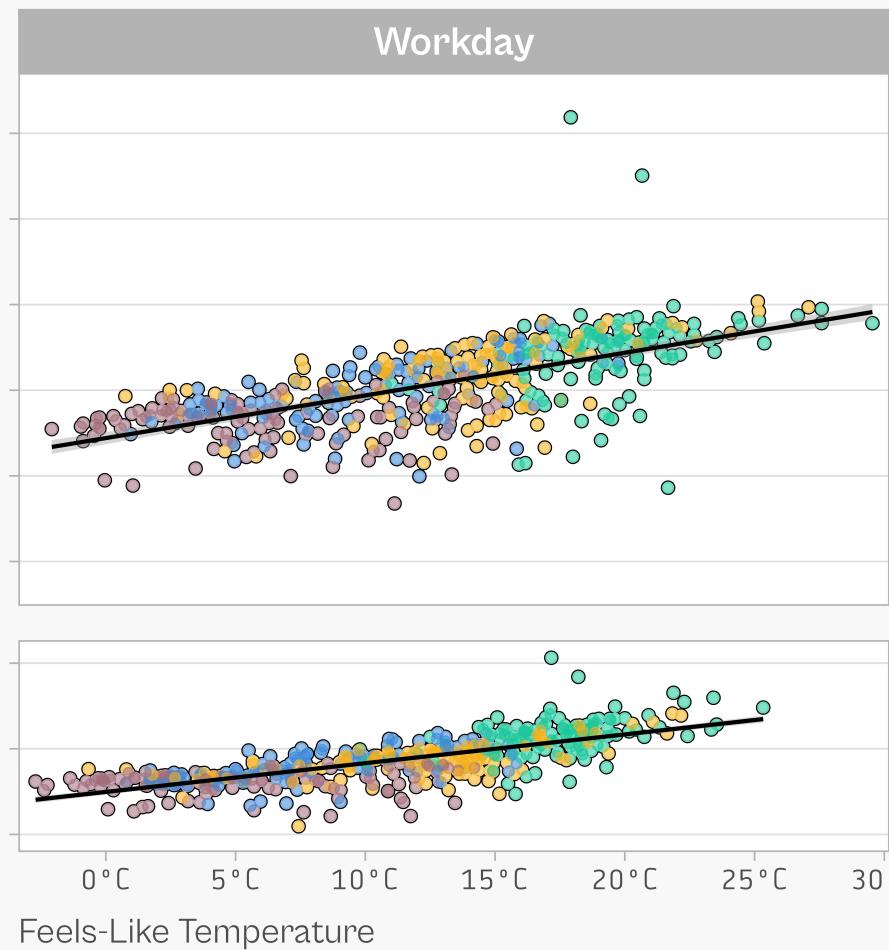
Reported bike rents versus feels-like temperature in London per time of day, period, and season.



Data: TfL (Transport for London), Jan 2015 – Dec 2016

Reported bike rents versus feels-like temperature in London per time of day, period, and season.

● Spring ● Summer ● Autumn ● Winter



Data: TfL (Transport for London), Jan 2015 – Dec 2016