

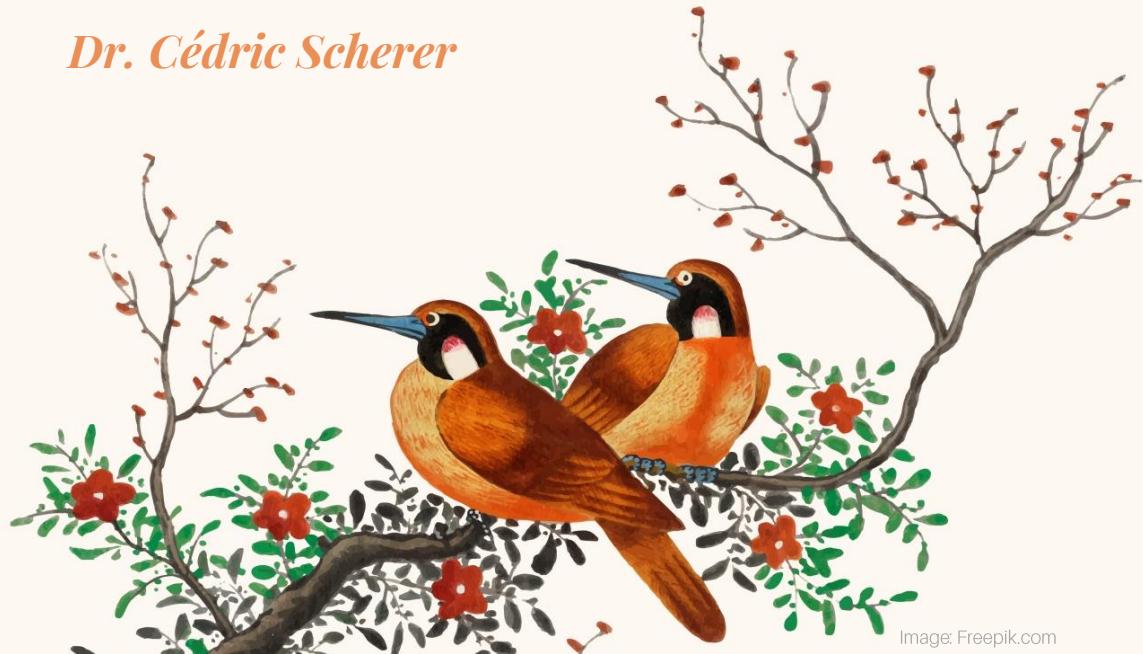


ggplot Wizardry

Extended Version

My Favorite Tricks and Secrets for Beautiful Plots in R

Dr. Cédric Scherer



MeetUp March 25, 2021

Image: Freepik.com

Scientist by Training



Computational Ecology at the Leibniz Institute for Zoo and Wildlife Research

DataVizard by Heart



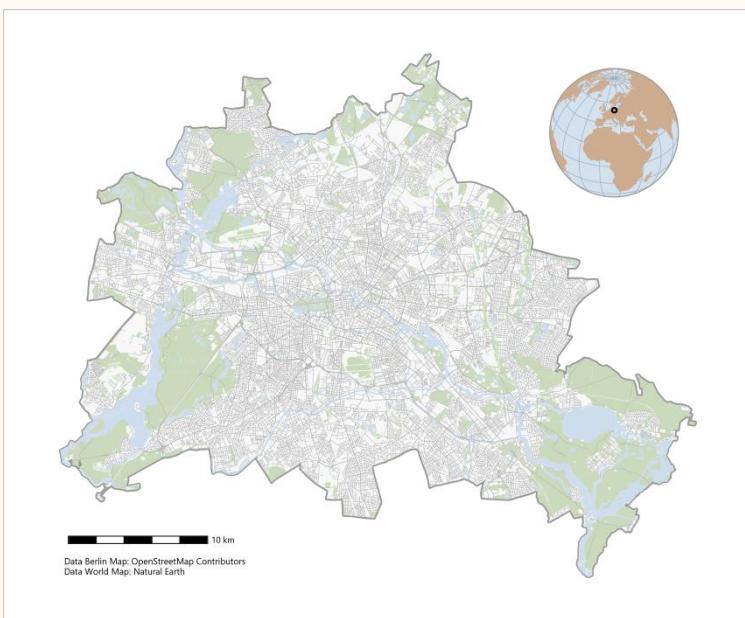
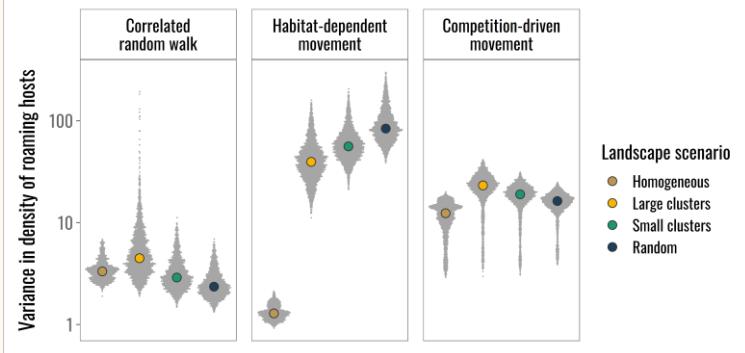
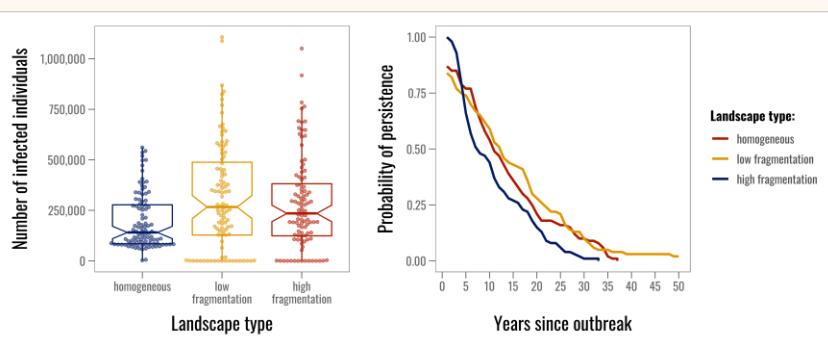
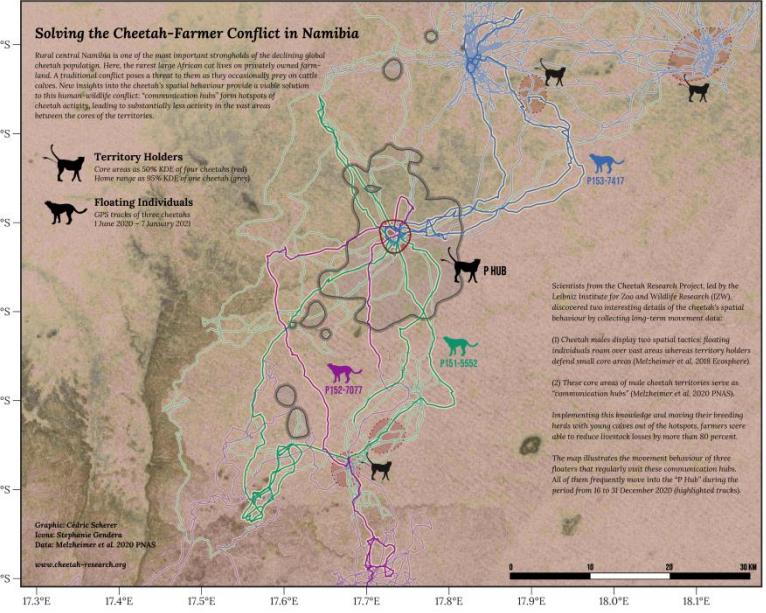
Self-employed Data Visualization Designer & Workshop Instructor

99% R | 95% ggplot2



- 2008 – my first time using R
- 2011 – frequent use of base R
- 2016 – discovered ggplot2 + rest of the tidyverse
- 2019 – using ggplot2 every day
- 2020 – freelancing with ggplot2 designs + courses





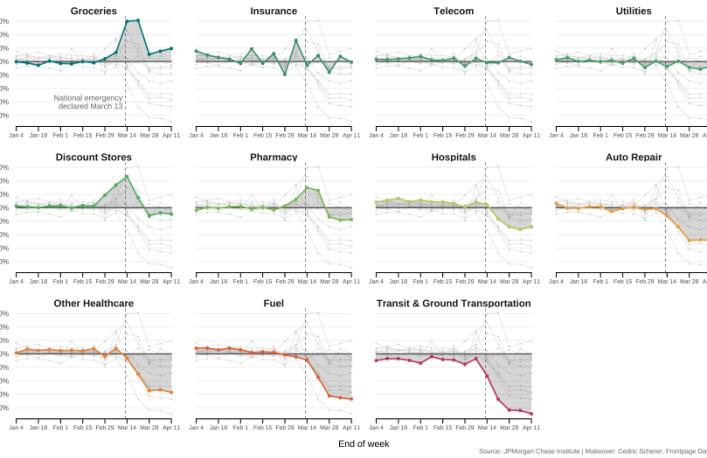
Top left:
BES MoveMap Contest –
Winner “Pretty Map” and
Runner-Up “Rmap”

Top right:
Scherer, Radchuk, Franz,
Thulke, Lange, Grimm &
Kramer-Schadt (2020)
Oikos 129 (5):651–667
(modified)

Bottom left:
Sciaini, Fritsch, Scherer
& Simpkins (2018)
*Methods in Ecology and
Evolution* 9 (11):2240–
2248 (modified)

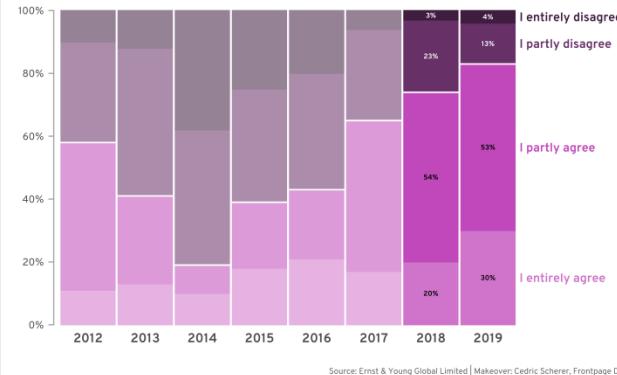
Bottom right:
Package to create
template maps of Berlin

Year-over-year percent change in spending by essential category

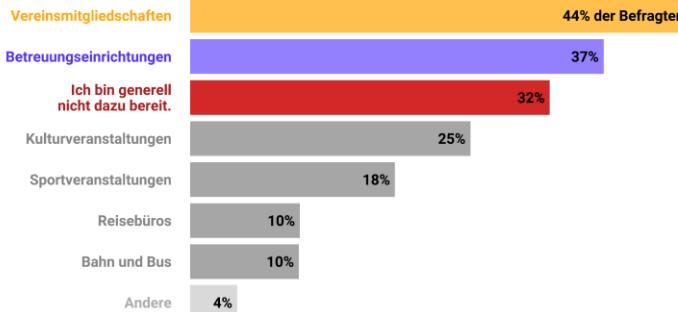


Do you agree with the following statement?

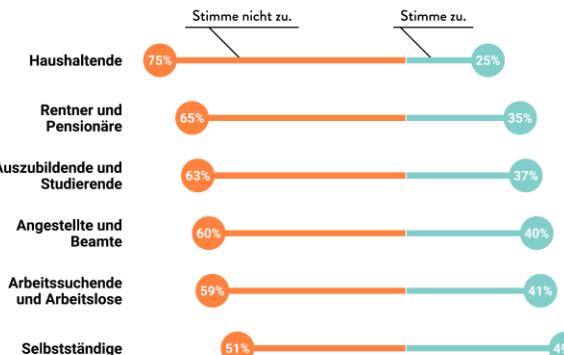
"The price of banking services will fall."



Ich bin bei diesen Betrieben bereit auf eine Rückerstattung zu verzichten.
(Mehrfachauswahl möglich)



Der Eventim-Chef Klaus-Peter Schulenberg sagt, dass die Pflicht zur Ticketerstattung der Kultur in Deutschland dauerhaft schadet.

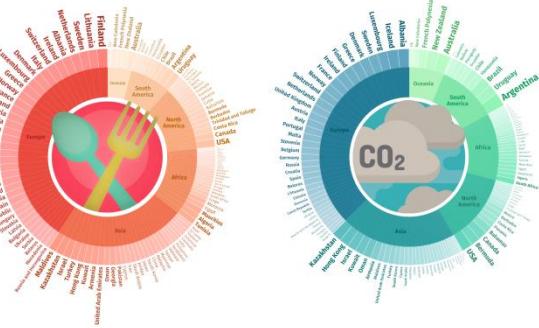


Upper row:
[Makeover examples for FrontPage Data](#)

Lower row:
[Survey on contract termination during the COVID-19 pandemic for kuendigung.org](#)

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 Cedric Scherer • Data by Food and Agriculture Organization of the United Nations (FAO) via mca • Icons by Freepik

THE WHOLE WORLD IS GETTING WARMER

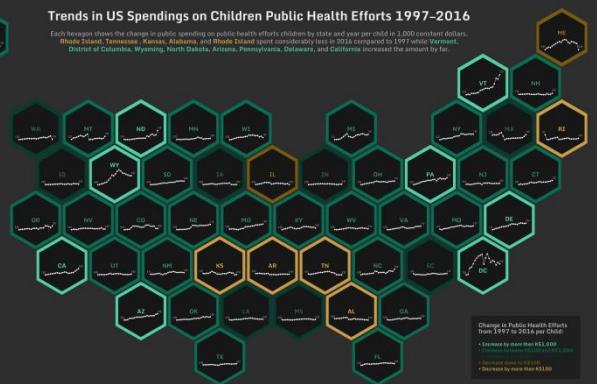
TEMPERATURE ANOMALIES BY COUNTRY 1880-2018 COMPARED TO 1850-1900 BASELINE



Visualization: Cedric Scherer • Data: GISTEMP (country-level data provided by Kevin Lipps)

Trends in US Spending on Children Public Health Efforts 1997–2016

Each hexagon shows the change in public spending on public health efforts children by the last year our child in 2010 completed school.
Rhode Island, Tennessee, & Kentucky All three Rhode Island spent considerably less in 2016 compared to 1997 while Vermont, District of Columbia, Wyoming, North Dakota, Arizona, Pennsylvania, Delaware, and California increased the amount by far.



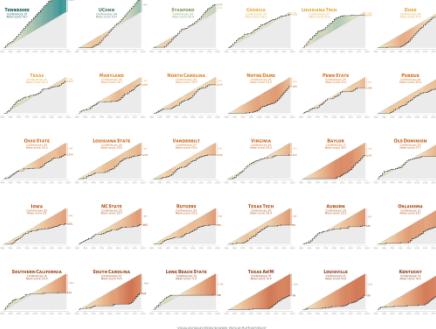
Visualization by Cedric Scherer • Data: Census Bureau's Annual State and Local Government Finance Survey (Expenditure Variable E22)

THE RISE & FALL OF WOMEN'S COLLEGE BASKETBALL DYNASTIES

A women's college basketball dynasty is the rise of the NCAA basketball program that has emerged as one of the most dominant programs in the history of college basketball. The term "dynasty" typically refers to a team that has won at least four national championships in a row.

The visualization shows the cumulative rise and fall of women's college basketball dynasties over time. It includes data from the early 1980s to the present day.

The visualization shows the cumulative rise and fall of women's college basketball dynasties over time. It includes data from the early 1980s to the present day.



Wind Turbines in Canada

Each vertical stripe contains the same number of wind turbines.



Visualization by Cedric Scherer • Data: Government of Canada

ggplot2

[gee-gee-plot-two]

an R package

is a system for declaratively creating graphics
based on “The Grammar of Graphics”.



Related searches



map

scatter plots

bubble chart

histogram

font

gganimate

thematic maps

line

switzerland

graphics

text

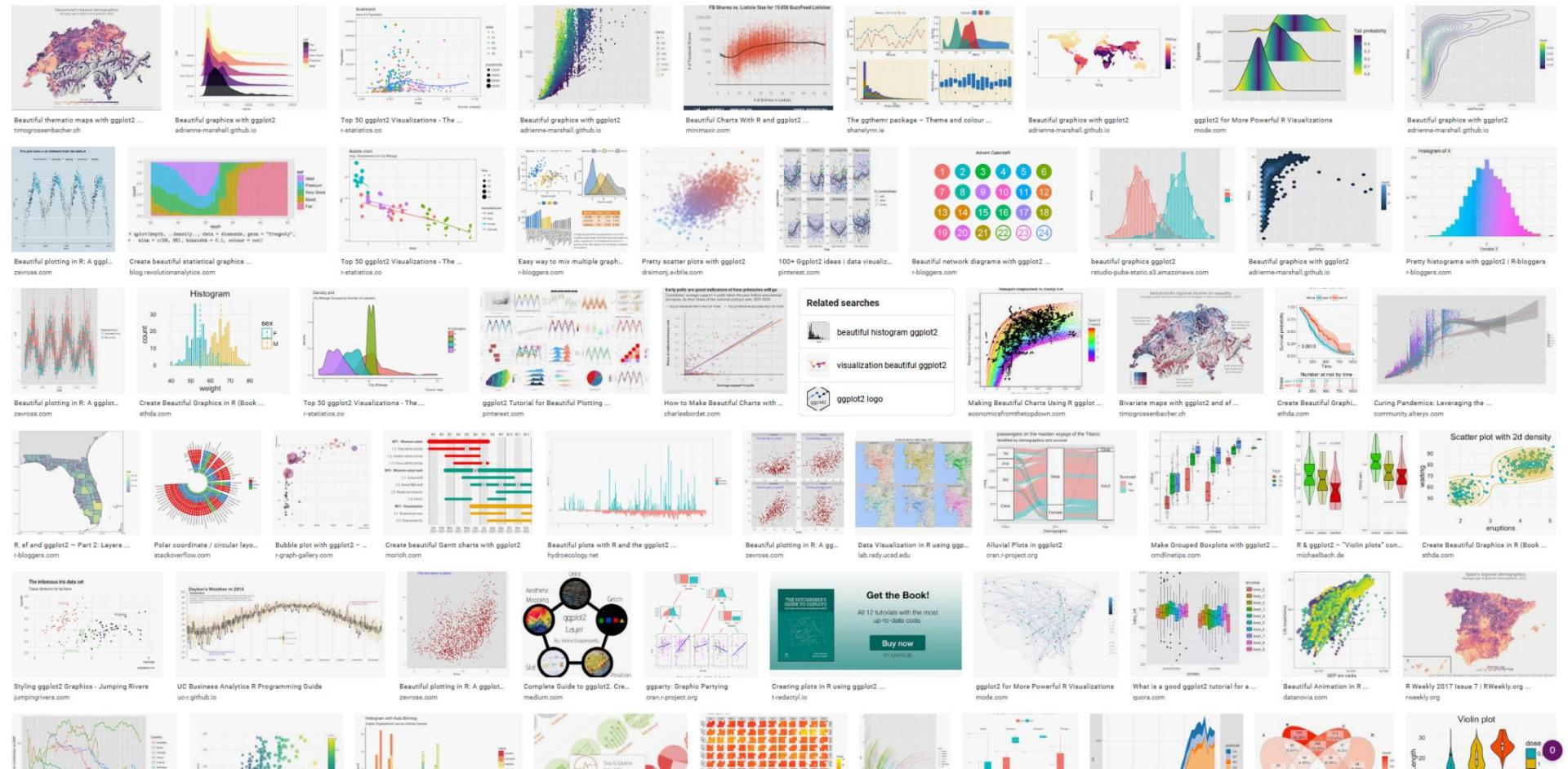
rstudio

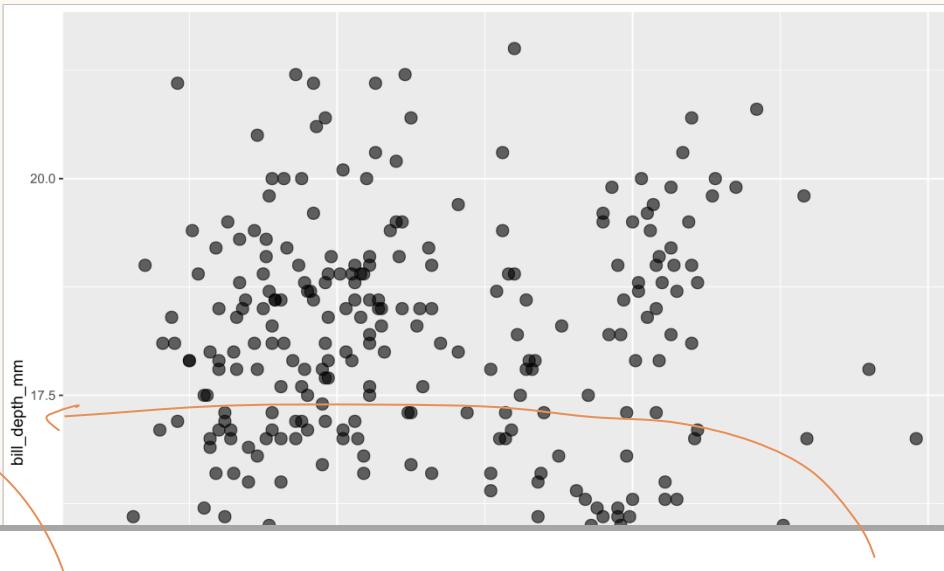
swiss

visualization

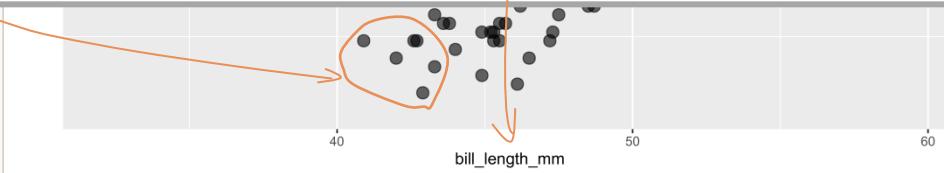
choropleth

geom





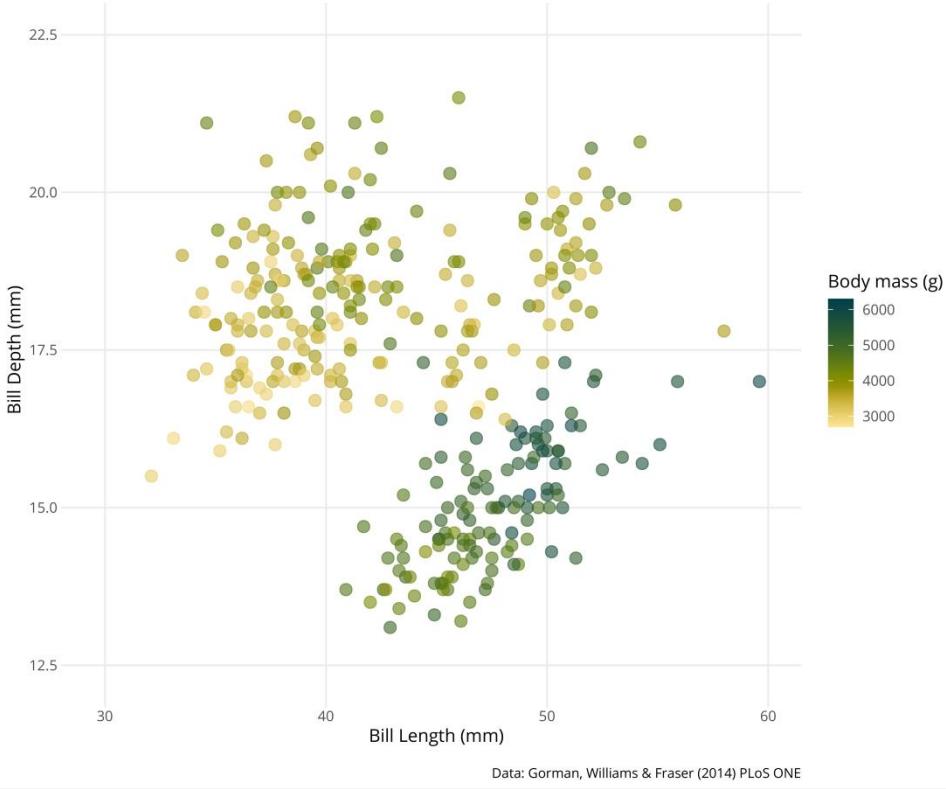
```
ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +  
  geom_point(alpha = .6)
```



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

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.



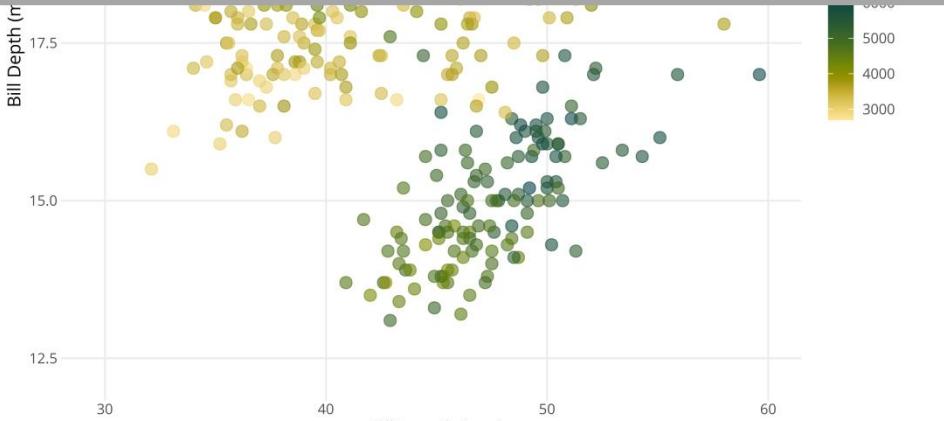
You provide the data, tell ggplot2 how to map variables to aesthetics, what graphical primitives to use, and it takes care of the details.

Bill Dimensions of Brush-Tailed Penguins (Pygoscelis)

A scatter plot of bill depth versus bill length.

22.5

```
ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +  
  geom_point(aes(color = body_mass_g), alpha = .6) +  
  ## custom colors  
  scico::scale_color_scico(palette = "bamako", direction = -1)
```



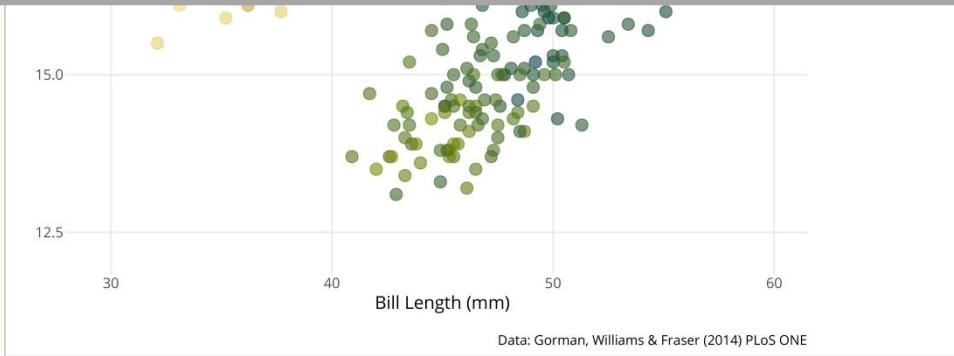
Data: Gorman, Williams & Fraser (2014) PLoS ONE

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

Bill Dimensions of Brush-Tailed Penguins (Pygoscelis)

A scatter plot of bill depth versus bill length.

```
ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +  
  geom_point(aes(color = body_mass_g), alpha = .6) +  
  ## custom colors  
  scico::scale_color_scico(palette = "bamako", direction = -1) +  
  ## custom axes scaling  
  scale_x_continuous(breaks = 3:6 * 10, limits = c(30, 60)) +  
  scale_y_continuous(breaks = seq(12.5, 22.5, by = 2.5), limits = c(12.5, 22.5))
```



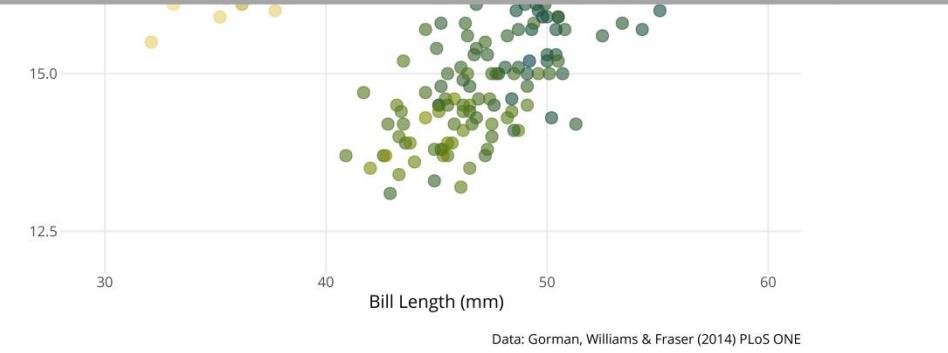
You provide the data, tell ggplot2 how to map variables to aesthetics, what graphical primitives to use, and it takes care of the details.

Bill Dimensions of Brush-Tailed Penguins (Pygoscelis)

A scatter plot of bill depth versus bill length.

22.5

```
ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +  
  geom_point(aes(color = body_mass_g), alpha = .6) +  
  ## custom colors  
  scico::scale_color_scico(palette = "bamako", direction = -1) +  
  ## custom axes scaling  
  scale_x_continuous(breaks = 3:6 * 10, limits = c(30, 60)) +  
  scale_y_continuous(breaks = seq(12.5, 22.5, by = 2.5), limits = c(12.5, 22.5))
```

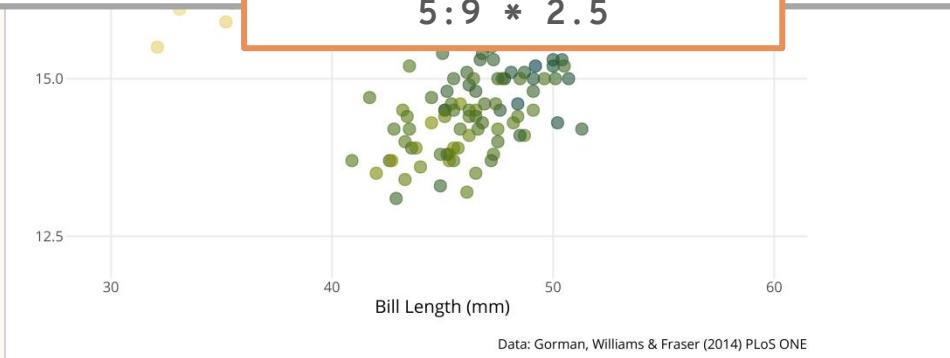


You provide the data, tell ggplot2 how to map variables to aesthetics, what graphical primitives to use, and it takes care of the details.

Bill Dimensions of Brush-Tailed Penguins (Pygoscelis)

A scatter plot of bill depth versus bill length.

```
ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +  
  geom_point(aes(color = body_mass_g), alpha = .6) +  
  ## custom colors  
  scico::scale_color_scico(palette = "bamako", direction = -1) +  
  ## custom axes scaling  
  scale_x_continuous(breaks = 3:6 * 10, limits = c(30, 60)) +  
  scale_y_continuous(breaks = seq(12.5, 22.5, by = 2.5), limits = c(12.5, 22.5))
```



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

Bill Dimensions of Brush-Tailed Penguins (Pygoscelis)

A scatter plot of bill depth versus bill length.

22.5

```
ggplot(data = penguins, aes(x = bill_length_mm, y = bill_depth_mm)) +  
  geom_point(aes(color = body_mass_g), alpha = .6) +  
  ## custom colors  
  scico::scale_color_scico(palette = "bamako", direction = -1) +  
  ## custom axes scaling  
  scale_x_continuous(breaks = 3:6 * 10, limits = c(30, 60)) +  
  scale_y_continuous(breaks = seq(12.5, 22.5, by = 2.5), limits = c(12.5, 22.5)) +  
  ## custom labels  
  labs(  
    title = 'Bill Dimensions of Brush-Tailed Penguins (Pygoscelis)',  
    subtitle = 'A scatter plot of bill depth versus bill length.',  
    caption = 'Data: Gorman, Williams & Fraser (2014) PLoS ONE',  
    x = 'Bill Length (mm)',  
    y = 'Bill Depth (mm)',  
    color = 'Body mass (g)'  
)
```

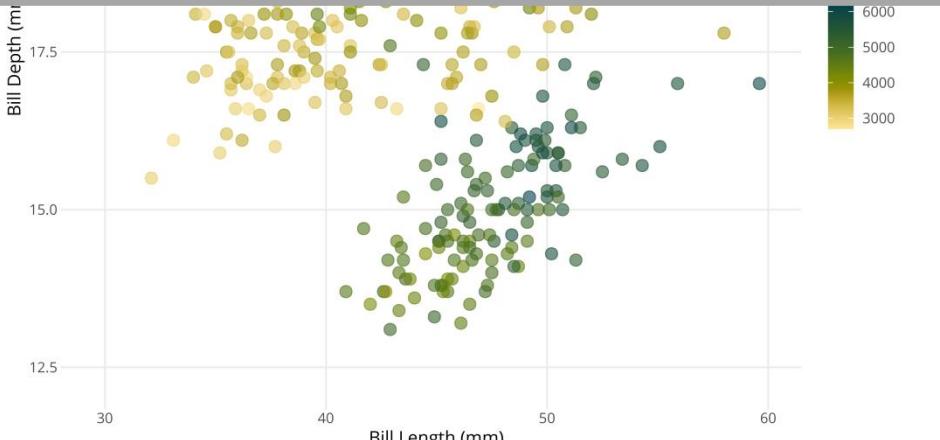
You provide the data, tell ggplot2 how to map variables to aesthetics, what graphical primitives to use, and it takes care of the details.

Bill Dimensions of Brush-Tailed Penguins (Pygoscelis)

A scatter plot of bill depth versus bill length.

Run before plotting!
(or add via +)

```
## change global theme settings (for all following plots)
theme_set(theme_minimal(base_size = 12, base_family = "Open Sans"))
```



Data: Gorman, Williams & Fraser (2014) PLoS ONE

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

Bill Dimensions of Brush-Tailed Penguins (Pygoscelis)

A scatter plot of bill depth versus bill length.



Run before plotting!
(or add via +)

```
## change global theme settings (for all following plots)
theme_set(theme_minimal(base_size = 12, base_family = "Open Sans"))

## modify plot elements globally (for all following plots)
theme_update(
  axis.ticks = element_line(color = "grey92"),
  axis.ticks.length = unit(.5, "lines"),
  panel.grid.minor = element_blank(),
  legend.title = element_text(size = 12),
  legend.text = element_text(color = "grey30"),
  plot.title = element_text(size = 18, face = "bold"),
  plot.subtitle = element_text(size = 12, color = "grey30"),
  plot.caption = element_text(size = 9, margin = margin(t = 15))
)
```

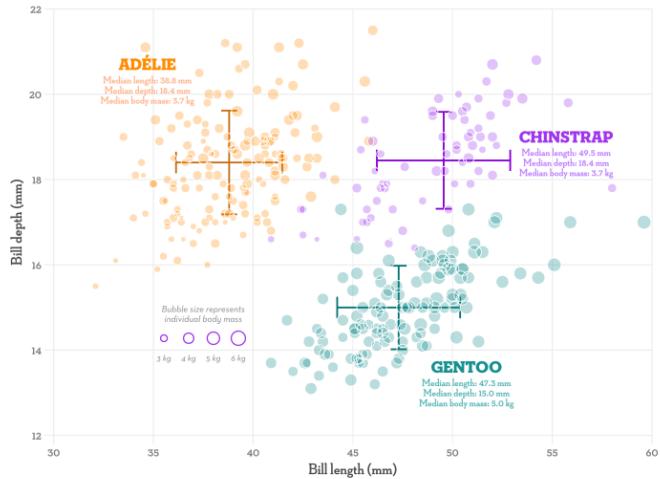
You provide the data, tell ggplot2 how to map variables to aesthetics, what graphical primitives to use, and it takes care of the details.

BILL DIMENSIONS OF BRUSH-TAILED PENGUINS

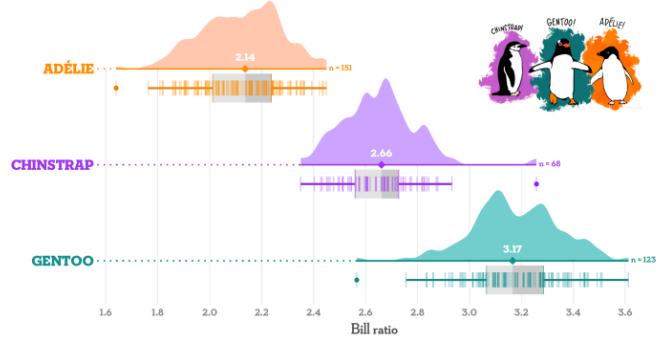
Pygoscelis adeliae (Adélie penguin) • *P. antarctica* (Chinstrap penguin) • *P. papua* (Gentoo penguin)



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



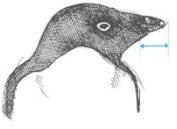
B. Distribution of the bill ratio, estimated as bill length divided by bill depth



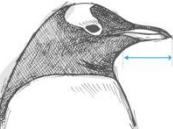
Palmer Penguins Bill Length

Palmer Archipelago is a group of islands off the northwestern coast of the Antarctic Peninsula. The histograms show that females have shorter bills than males in every species.

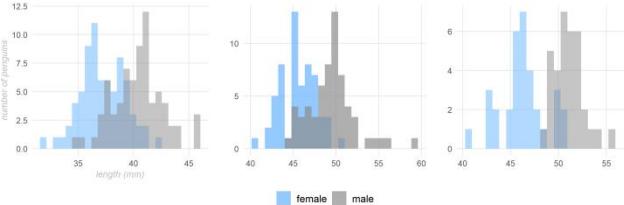
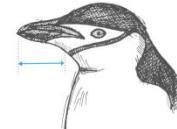
ADELIE



GENTOO



CHINSTRAP

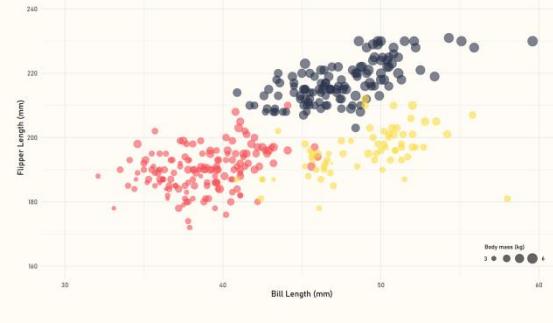


Visualization: Laura Navarro Soler | Data: Gorman, Williams & Fraser (2014)

Measurements of body mass, bill length, bill depth, and flipper length in **Adélie**, **Gentoo**, and **Chinstrap** penguins, collected from 3 islands in the Palmer Archipelago, Antarctica.

Data: Dr. K. Gorman, Dr. A. Horst, & Dr. A. Hill. allisonhorst/palmerpenguins

Visualization: Joseph Shaw / @JosephShaw.



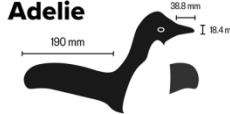
#TidyTuesday Contributions by

- ← myself
- ↖ Laura Navarro
- ↙ Joe Shaw
- ↓ Georgios Karamanis

Palmer Penguins

Median length of flipper, length and depth of bill, of 342 penguins recorded between 2007 and 2009

Adelie



Recorded penguins by species



Chinstrap



Recorded penguins by species



Gentoo



Source: Dr. Kristen Gorman and the Palmer Station, Antarctica LTER | Graphic: Georgios Karamanis

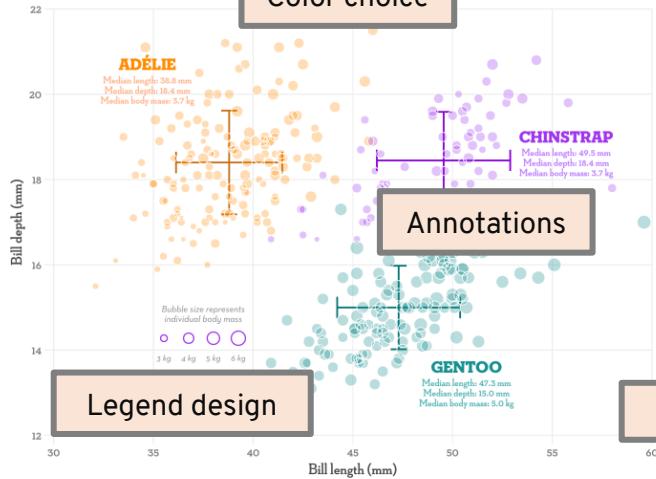
BILL DIMENSIONS OF BRUSH-TAILED PENGUINS

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

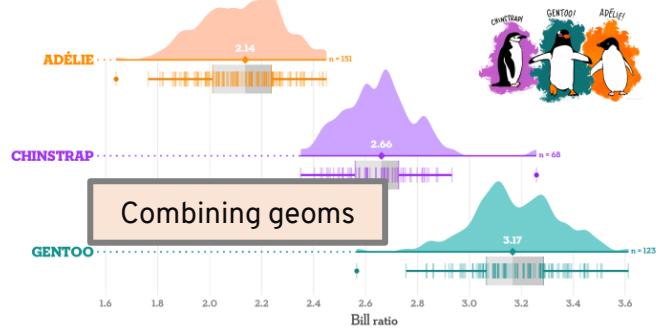


A. Scatterplot of bill length versus bill depth.

Color choice



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.

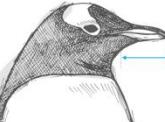
Palmer Penguins Bill Length

Palmer Archipelago is a group of islands off the northwestern coast of the Antarctic Peninsula. The histograms show that females have shorter bills than males in every species.

ADELIE



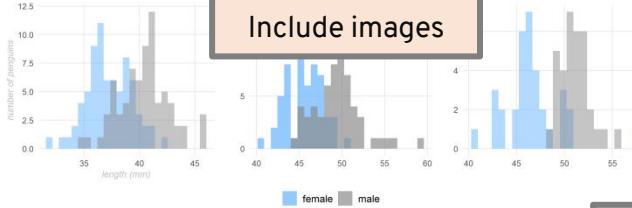
GENTOO



CHINSTRAP



Include images



#TidyTuesday Contributions by

- ← myself
- ↖ Laura Navarro
- ↙ Joe Shaw
- ↓ Georgios Karamanis

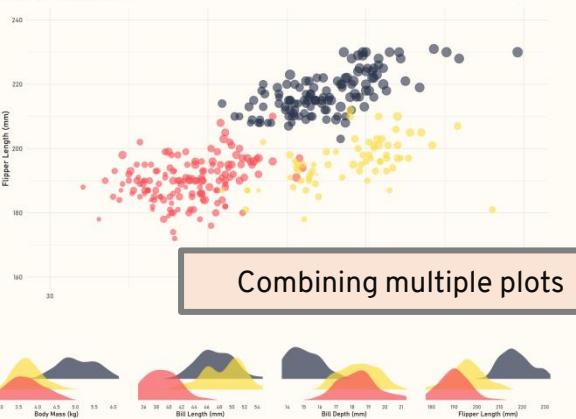
Spacing

Text styling

Measurements of body mass, bill length, bill depth, and flipper length in **Adelie**, **Gentoo**, and **Chinstrap** penguins, collected from 3 islands in the Palmer Archipelago, Antarctica.

Data: Dr. K. Gorman, Dr. A. Horst, & Dr. A. Hill. [allisonhorst/palmerpenguins](#)

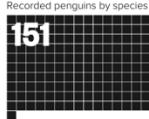
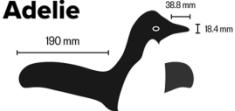
Visualization: Joseph Shaw / [@JosephShaw](#).



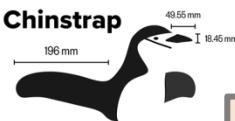
Palmer Penguins

Median length of flipper, length and depth of bill, of 342 penguins recorded between 2007 and 2009

Adelie



Chinstrap



Gentoo



Unusual geoms

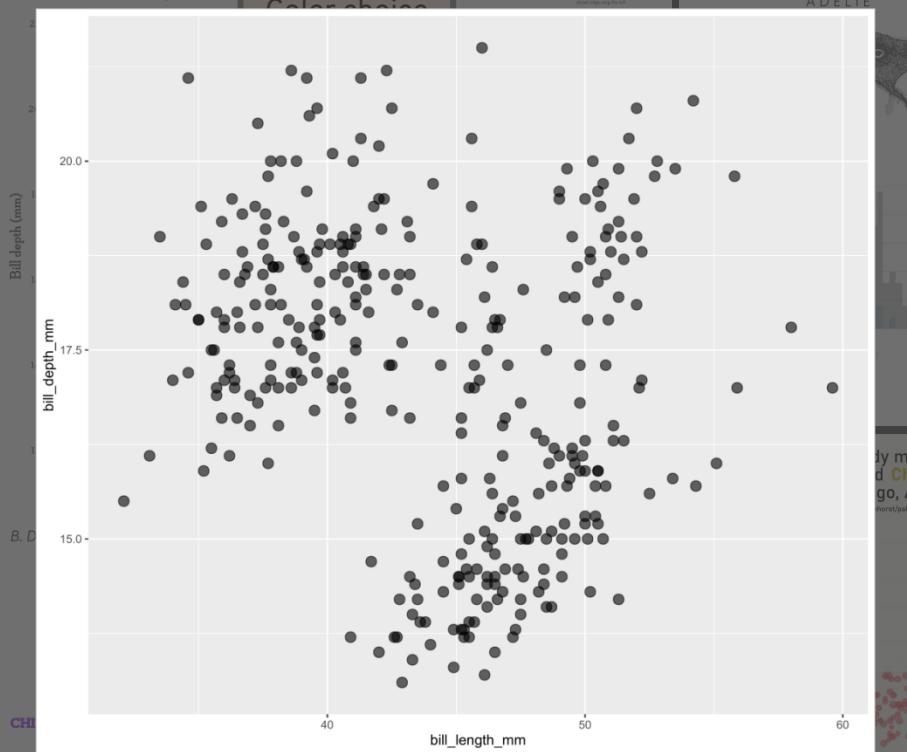
Source: Dr. Kristen Gorman and the Palmer Station, Antarctica LTER | Graphic: Georgios Karamanis

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



Combining geoms

GENTOO



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 Cedric Scherer • Data Cormac Williams & Fraser (2014) DOI: 10.1371/journal.pone.0090081 • Illustrations Alison Horst

Palmer Penguins Bill Length

Palmer Archipelago is a group of islands off the northwestern coast of the Antarctic Peninsula.
The histograms show that females have shorter bills than males in every species

ADELIE

GENTOO

CHINSTRAP

How to draw an owl

1.



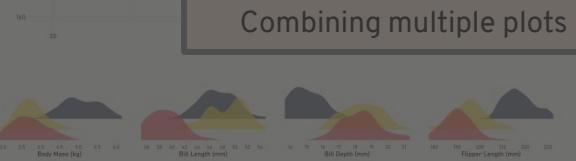
2.



1. Draw some circles

2. Draw the rest of the fowl owl

Combining multiple plots



Unusual geoms



123

Source: Dr. Kristen German and the Palmer Station, Antarctica LTER | Graphics: Georgios Karanikas

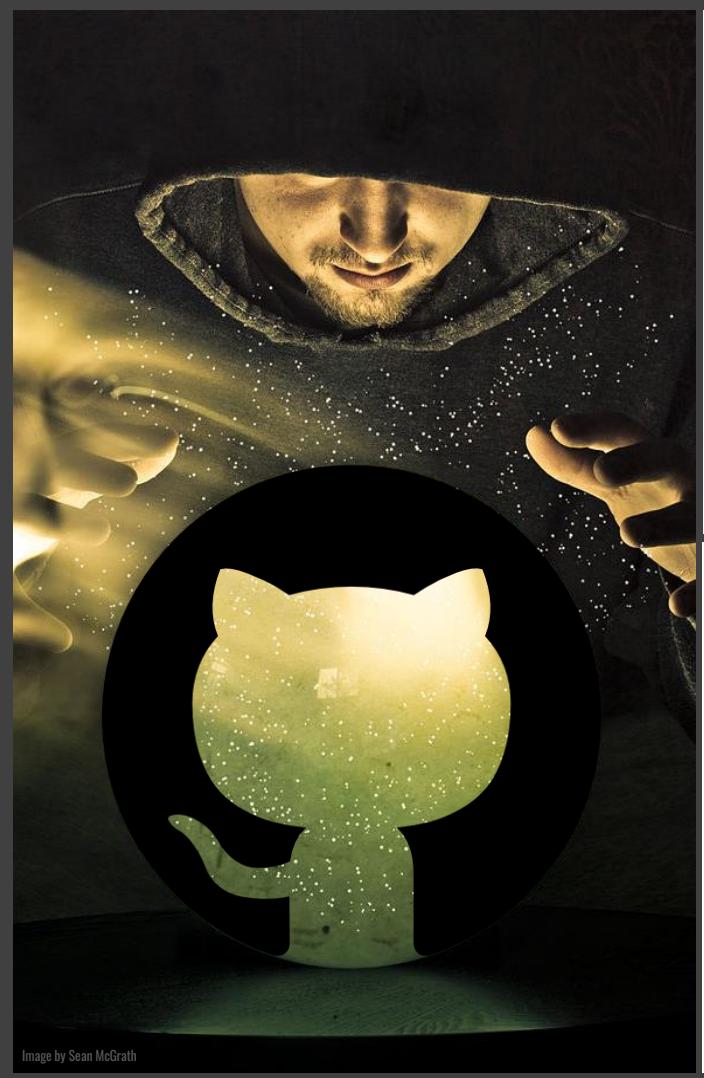
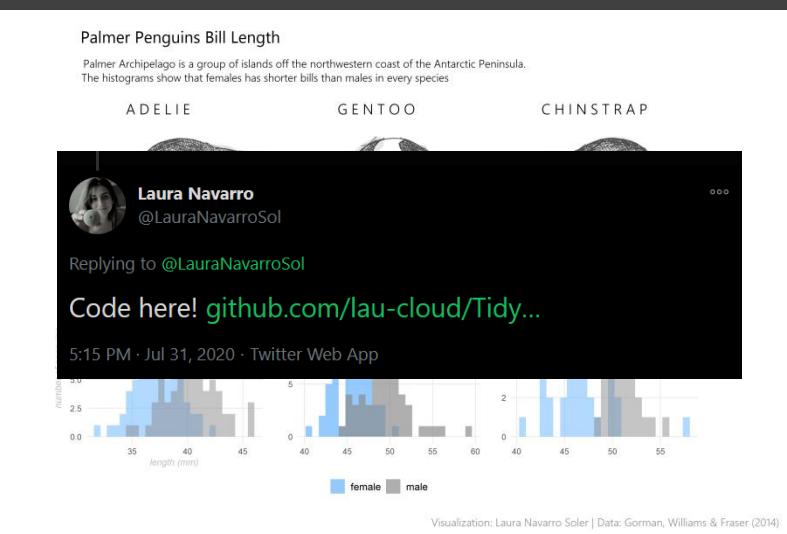
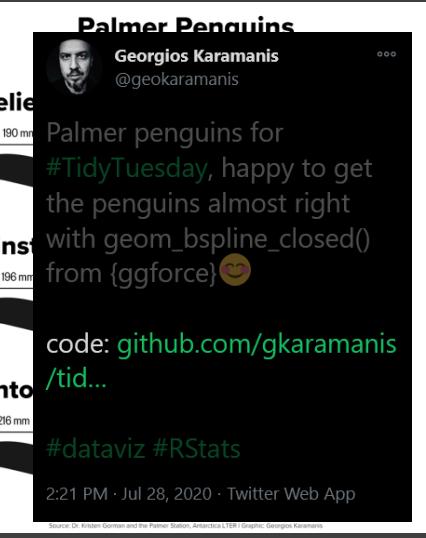


Image by Sean McGrath



#TidyTuesday Contributions by

- ← myself
- ↖ Laura Navarro
- ↙ Joe Shaw
- ↓ Georgios Karamanis



TidyTuesday

A weekly data project in R from the
R4DS online learning community

#TIDYTUESDAY on Twitter • [rfordatascience/tidytuesday](https://github.com/rfordatascience/tidytuesday) on GitHub • [R4DS](#) on Slack

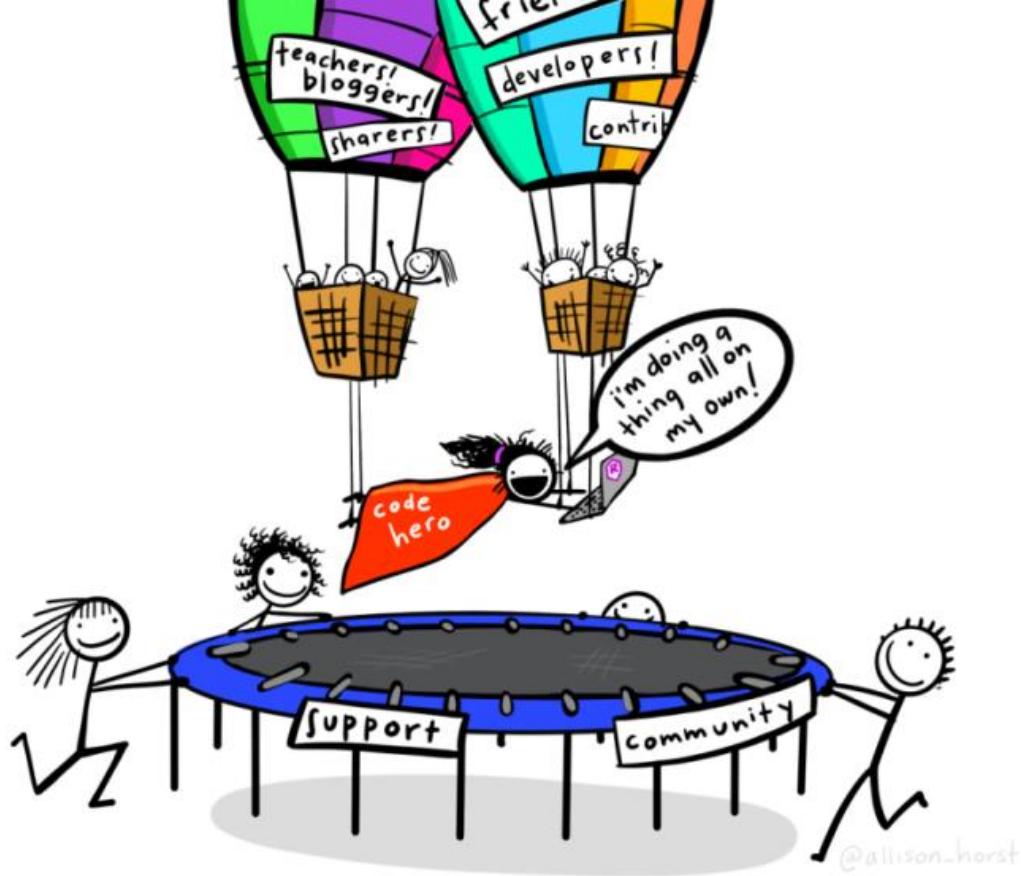


“I’m not looking to necessarily practice my skills as much as I am **to be inspired and know what I can do** based on what other people share.”

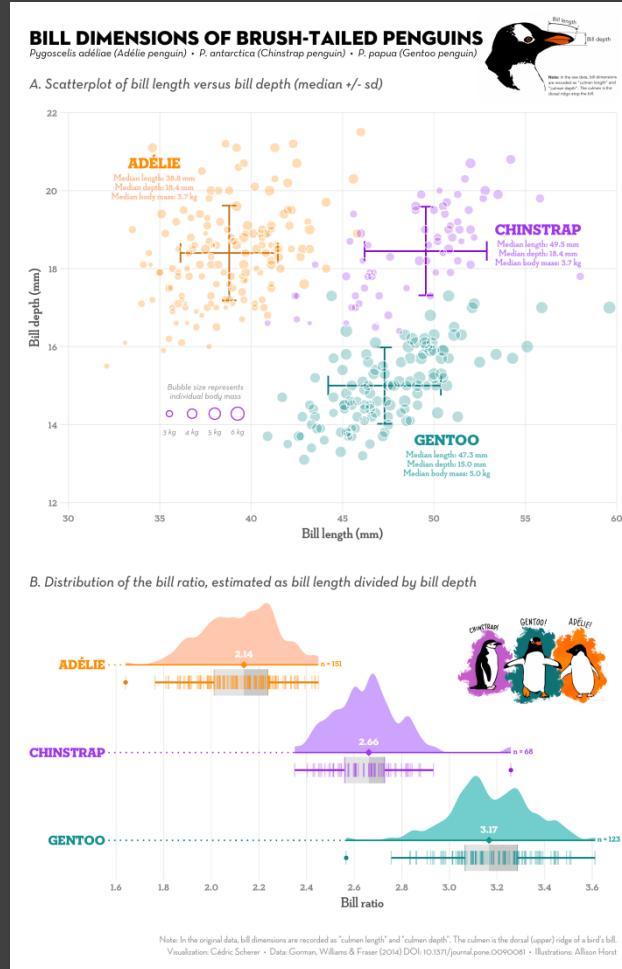
P3 in *Proc. ACM Hum.-Comput. Interact.* 37, 4, Article 111

“#TidyTuesday became a ‘choose your own adventure game’, which allowed participants **to ‘pursue something really weird’ beyond traditional visualizations.**”

I17 and P16 in *Proc. ACM Hum.-Comput. Interact.* 37, 4, Article 111



The Showcase



with the help of
{ggdist} & {ggttext} & {patchwork}

mjskay.github.io/ggdist

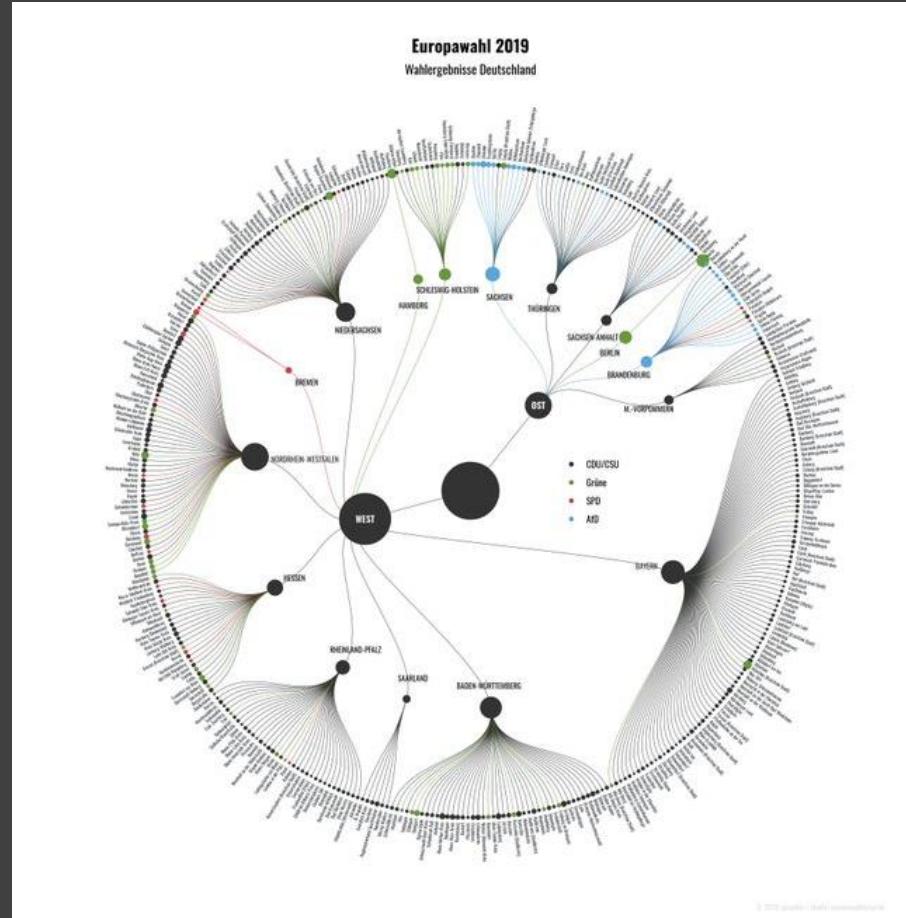
wilkelab.org/ggttext/

patchwork.data-imaginist.com

My Contribution to #TidyTuesday 2020/31

The Showcase

Tree Diagrams + Networks



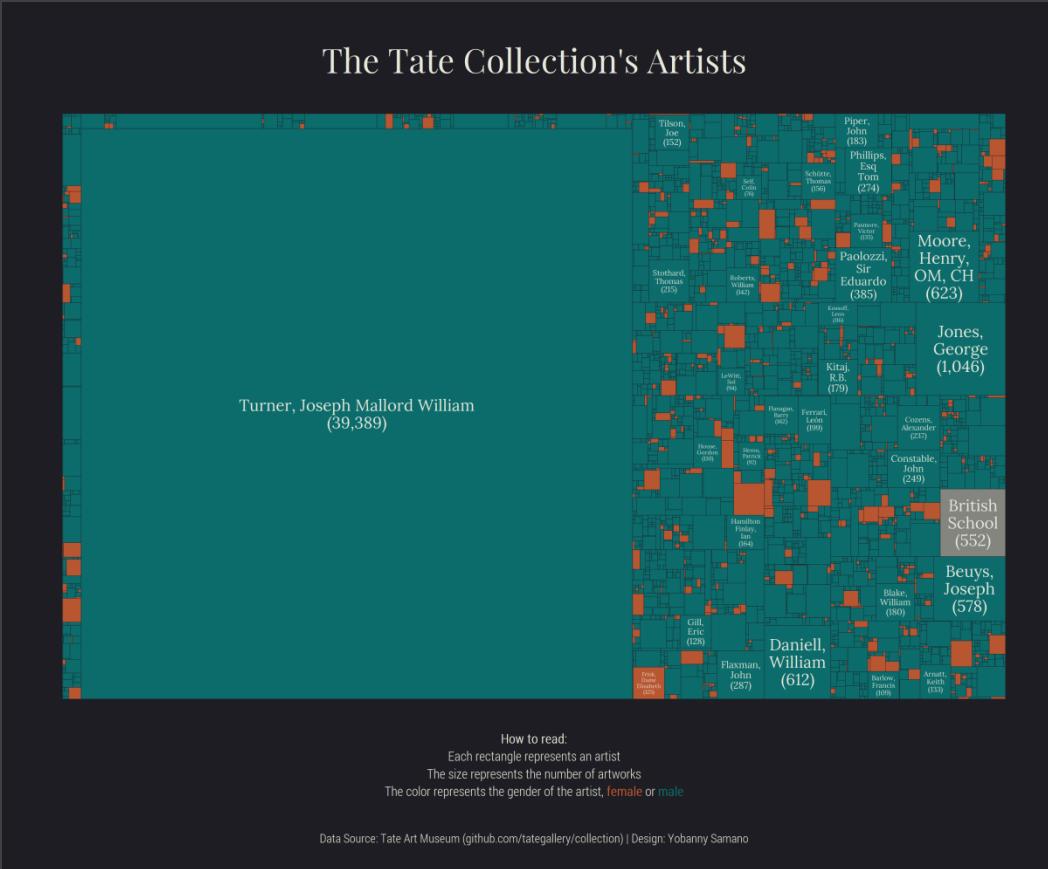
with the help of
{ggraph}

graph.data-imaginist.com

Thorsten Sprenger

The Showcase

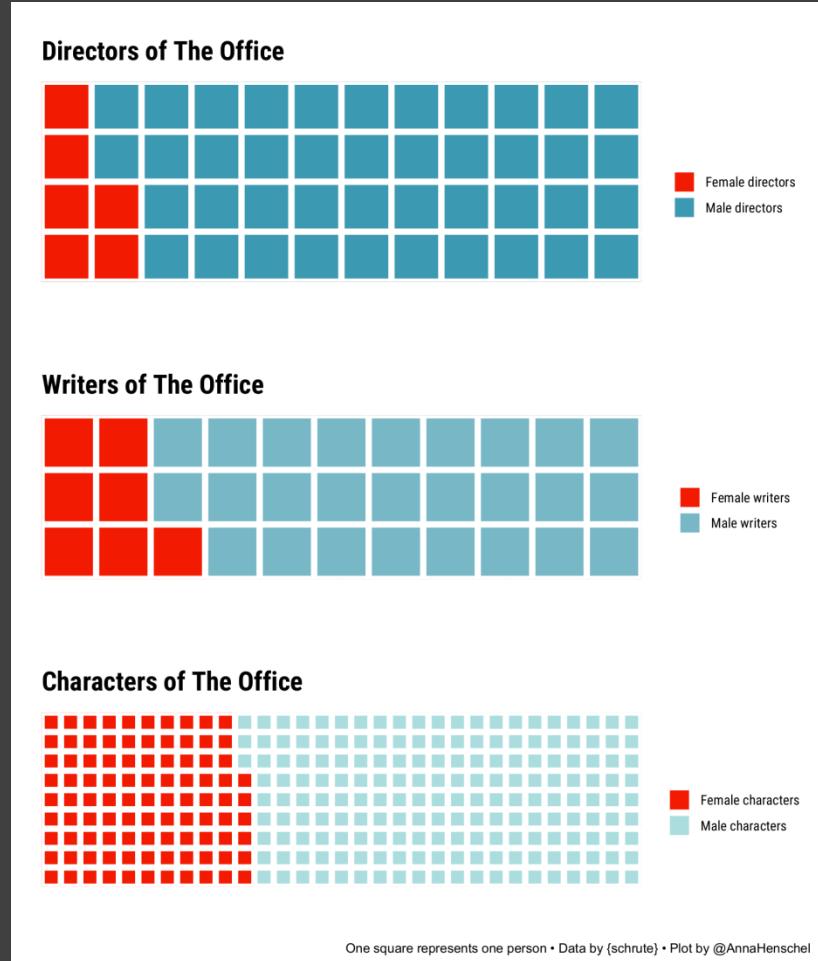
Treemaps



with the help of
{treemapify}
github.com/gwilkox/treemapify

The Showcase

Waffle Charts



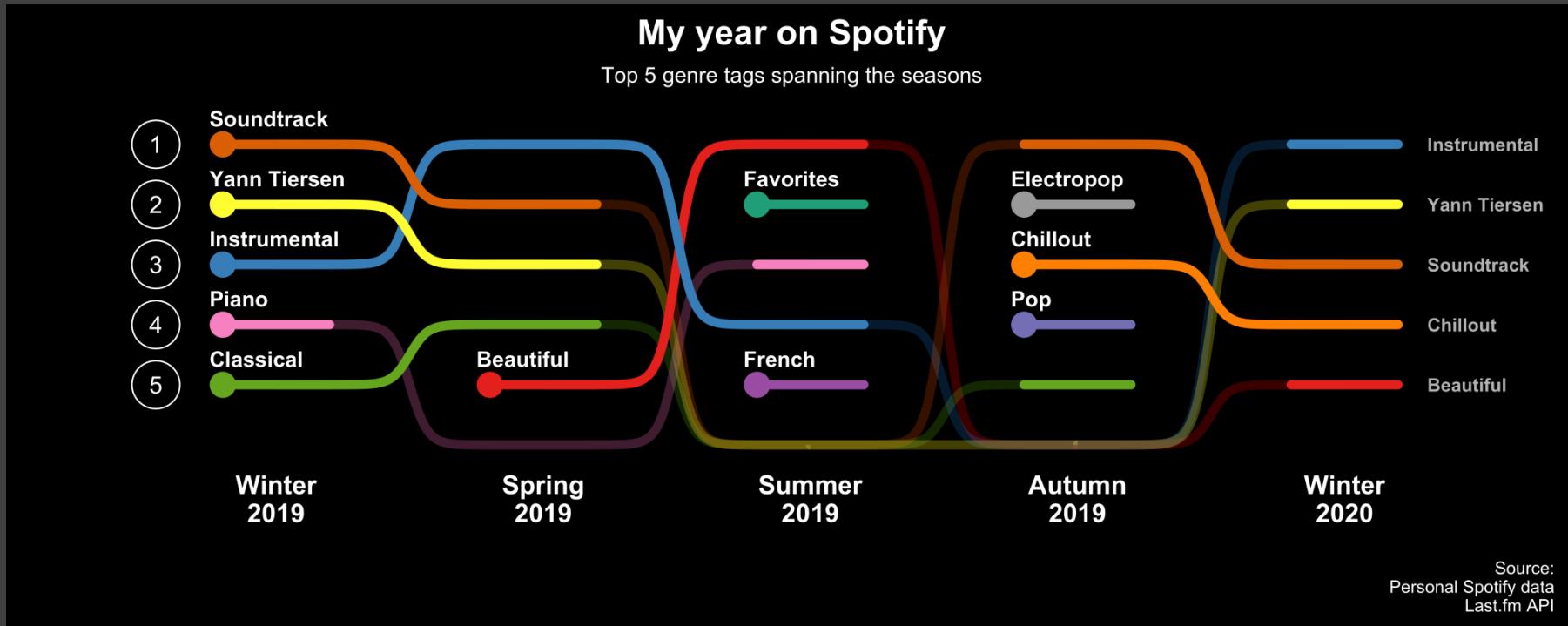
with the help of
{waffle}

github.com/hrbrmstr/waffle

Anna Henschel, #TidyTuesday 2020/12

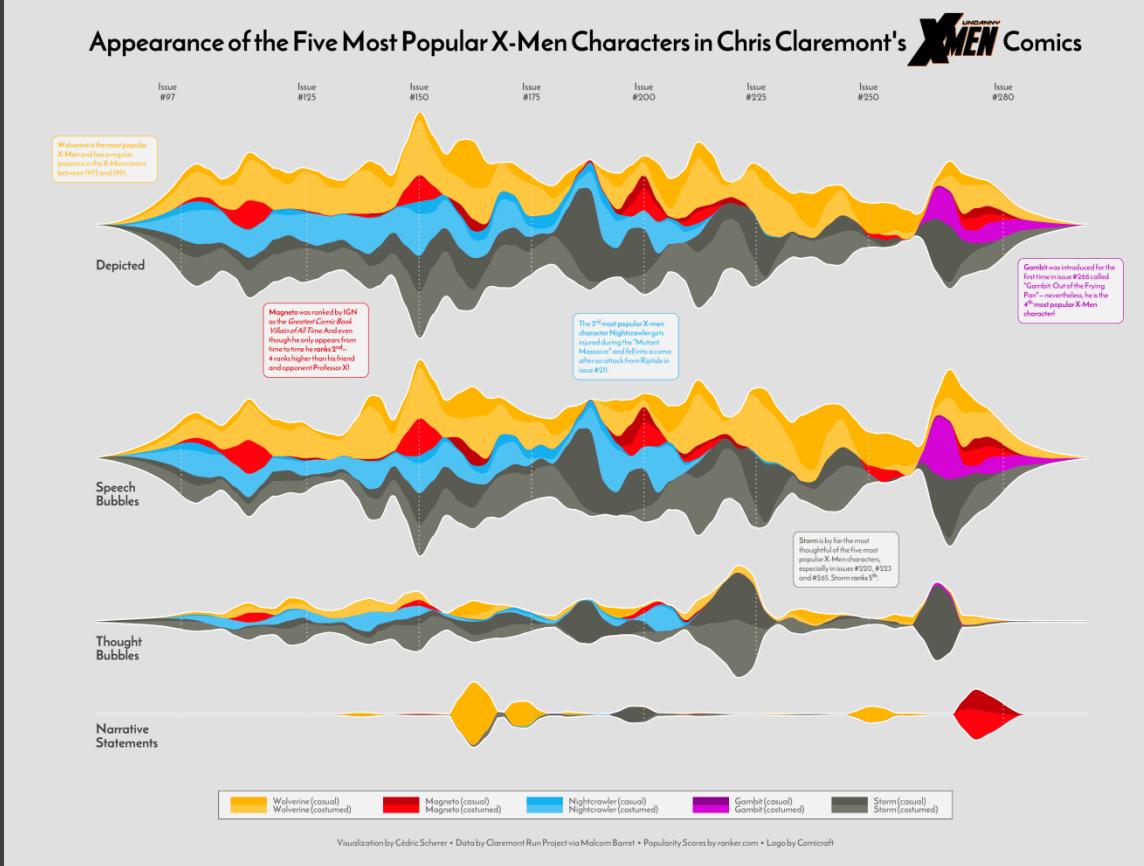
The Showcase

Bump Charts



The Showcase

Streamgraphs

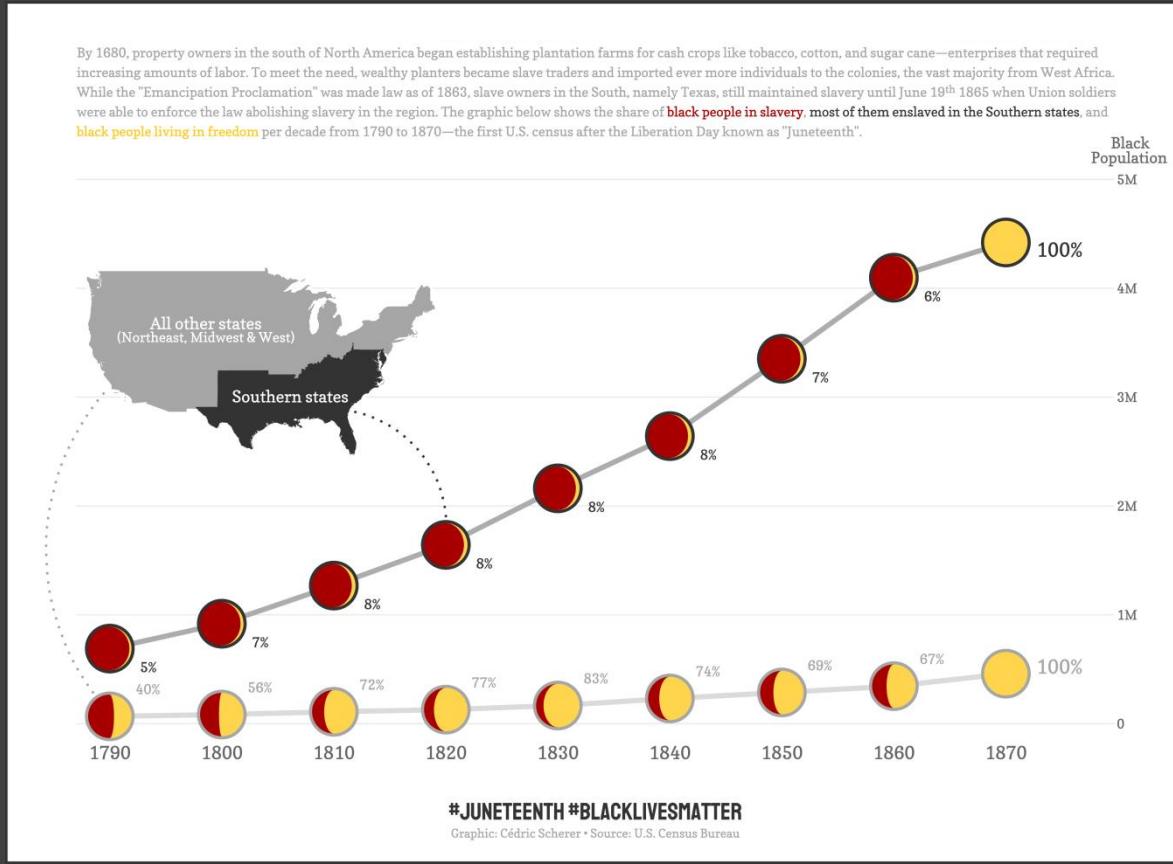


with the help of
{ggstream}
github.com/davidsjoberg/ggstream

My Contribution to #TidyTuesday 2020/27

The Showcase

Streamgraphs

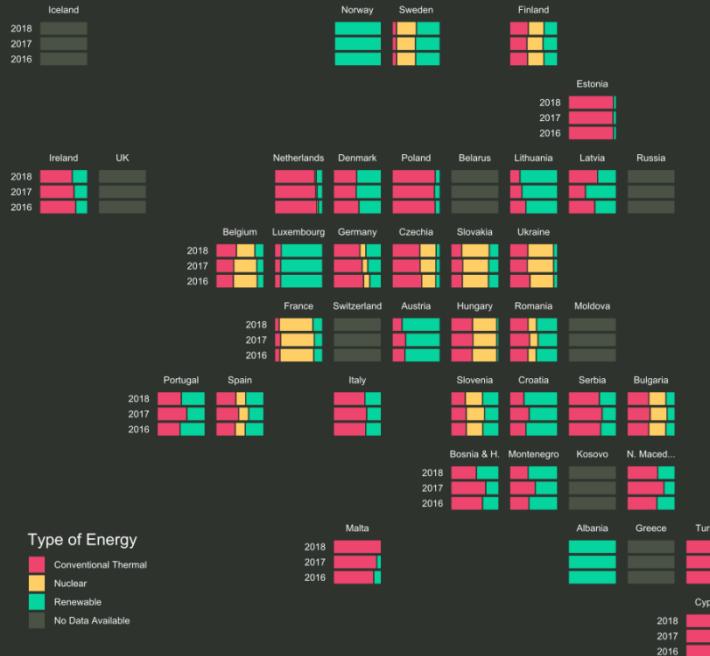


The Showcase

Geofacets

EUROPEAN ENERGY GENERATION

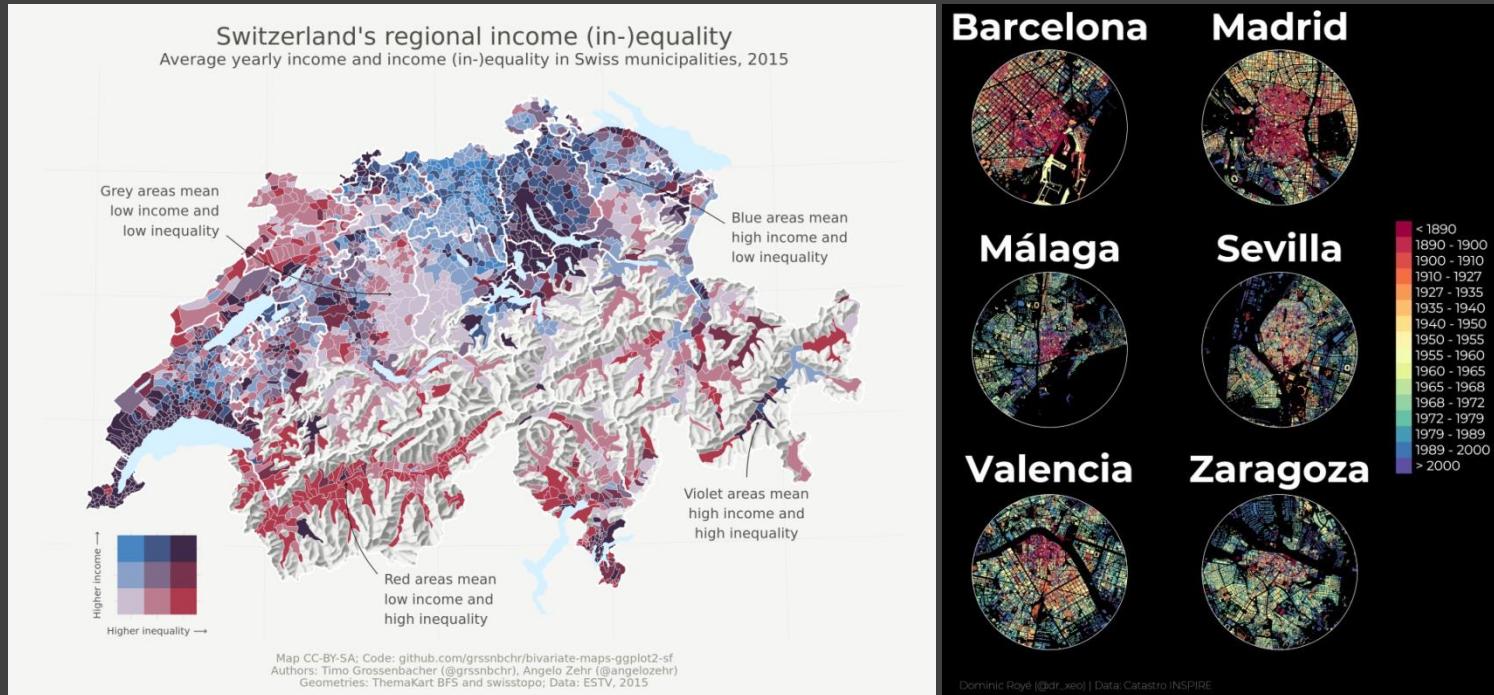
Each bar represents the **total energy generation** for each country per year.
The colours represent the proportion of energy generated a) using **conventional thermal power plants**, which is to say those that use coal, oil or natural gas,
b) using **nuclear power stations**, and c) using other **renewable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



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

The Showcase

Maps



The Showcase

Cartograms

What do most people die from?

Cardiovascular Diseases



Cancers



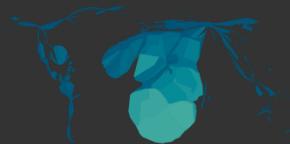
Diabetes



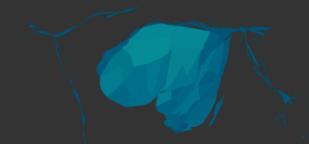
The leading causes of death across the world still vary significantly.

These cartograms show causes of deaths in 2016 that exceeded 20 percent of total deaths in at least 1 country.

HIV Infections & Aids



Malaria Infections



Wars & Conflicts



The data refers to the specific cause of death, which is distinguished from risk factors for death, such as air pollution, diet and other lifestyle factors.

with the help of
{cartogram}

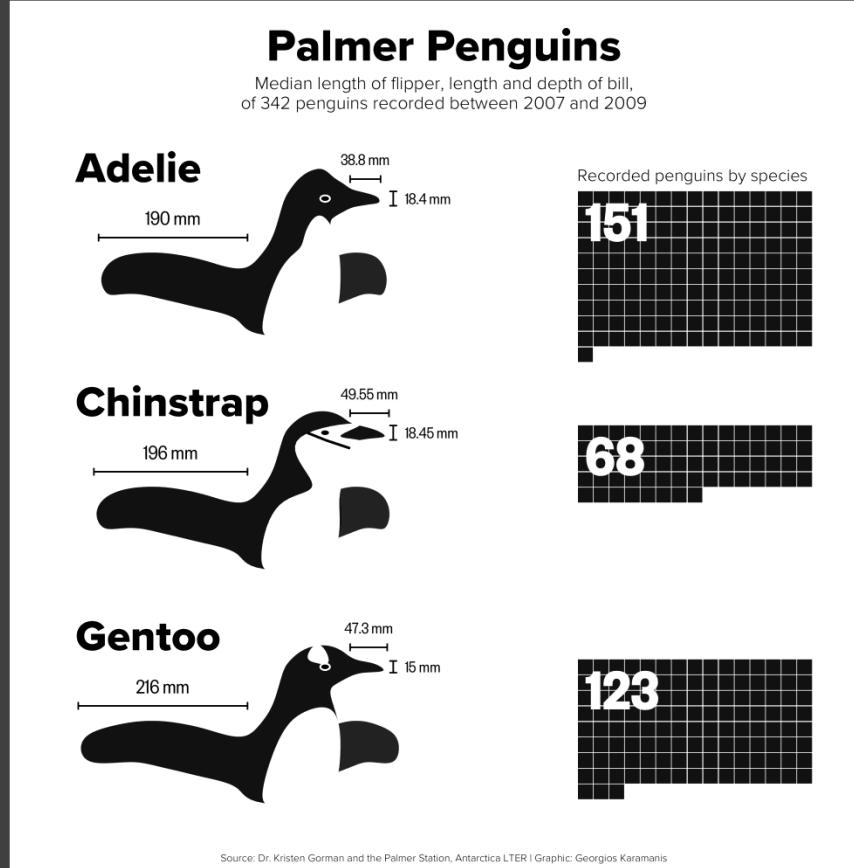
github.com/sjewo/cartogram

My Contribution to the #30DayMapChallenge 2019, Topic *Polygons*

Visualization by Cédric Schefer • Data by OurWorldInData.org

The Showcase

Drawings



The Showcase

Tables

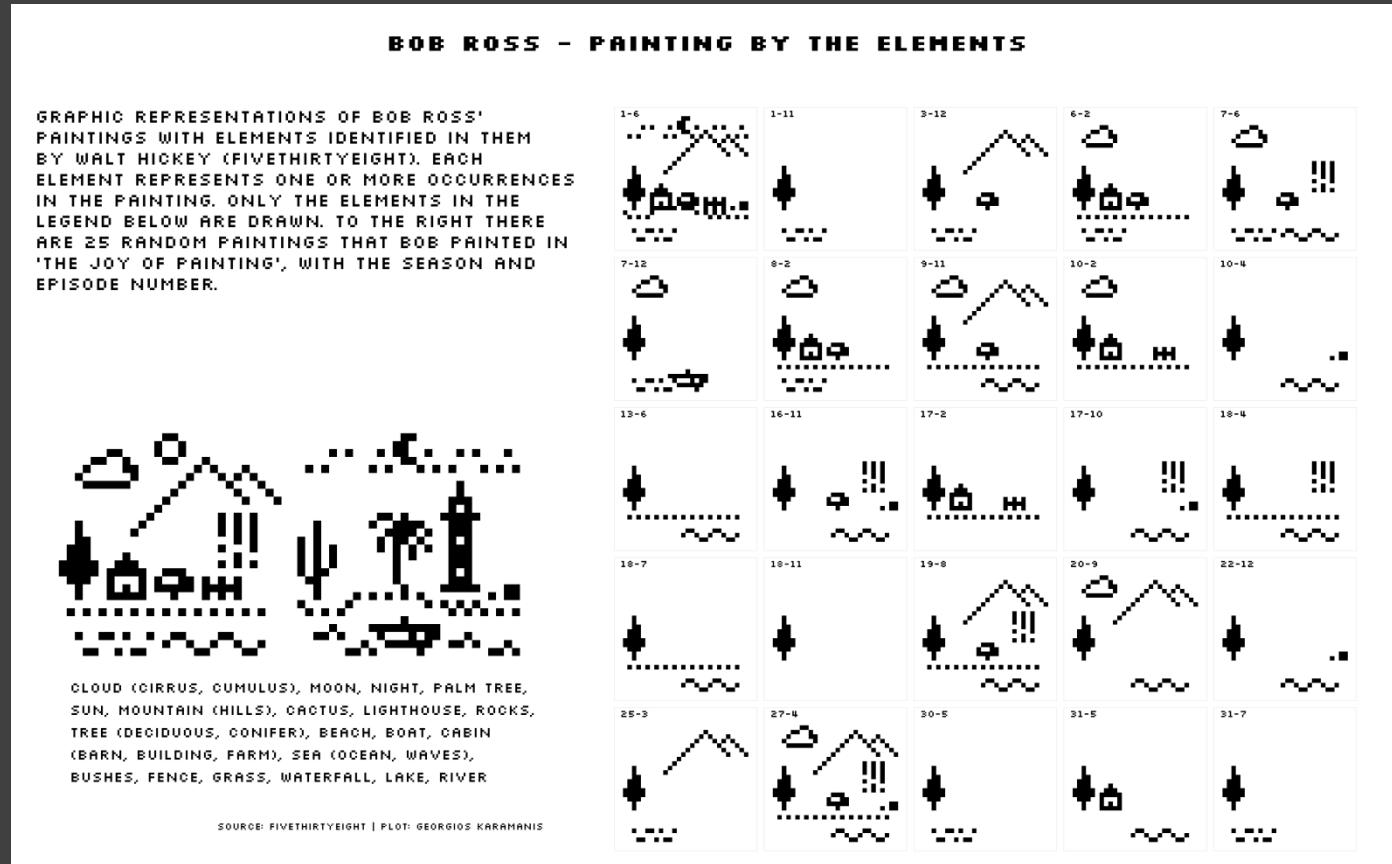
Tour de France Winners

Source: alastairrushworth/tdf & kaggle.com/jaminliu | Graphic: Georgios Karamanis

| YEAR | DISTANCE | WINNER | TEAM | AVERAGE SPEED | TOTAL TIME | YEAR |
|--|----------|---------------------------|---------------------------------|---------------|------------|------|
| 1985 | 0 | 8000 km | | | | |
| 1986 | 2428 | René Gavet (FRA) | La Française | 0 | 60 km/h | 1986 |
| 1987 | 4488 | Henri Cornet (FRA) | Gemini | 25.3 | + 96.1 | 1987 |
| 1988 | 4754 | Louis Trousselier (FRA) | Peugeot-Motul | | | 1988 |
| 1989 | 4754 | René Gavet (FRA) | Peugeot-Motul | | | 1989 |
| 1990 | 4754 | Lucien Petit-Breton (FRA) | Peugeot-Motul | | | 1990 |
| 1991 | 4754 | René Gavet (FRA) | Peugeot-Motul | | | 1991 |
| 1992 | 5287 | François Faber (LUX) | Acyon-Bulcup | | | 1992 |
| 1993 | 5287 | Gustave Serrigno (FRA) | Acyon-Bulcup | | | 1993 |
| 1994 | 5287 | Olivie Decroix (FRA) | Acyon-Bulcup | | | 1994 |
| 1995 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | 26.7 | + 197.9 | 1995 |
| 1996 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 1996 |
| 1997 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 1997 |
| 1998 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 1998 |
| 1999 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 1999 |
| 2000 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2000 |
| 2001 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2001 |
| 2002 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2002 |
| 2003 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2003 |
| 2004 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2004 |
| 2005 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2005 |
| 2006 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2006 |
| 2007 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2007 |
| 2008 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2008 |
| 2009 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2009 |
| 2010 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2010 |
| 2011 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2011 |
| 2012 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2012 |
| 2013 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2013 |
| 2014 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2014 |
| 2015 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2015 |
| 2016 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2016 |
| 2017 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2017 |
| 2018 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2018 |
| 2019 | 5287 | Philippe Thys (BEL) | Peugeot-Motul | | | 2019 |
| 1915-1918 Tour suspended because of World War I | | | | | | |
| 1919 | 5583 | Pierre Lefèvre (FRA) | La Sportive | | | 1919 |
| 1920 | 5583 | Philippe Thys (BEL) | La Sportive | 24.1 | + 228.6 | 1920 |
| 1921 | 5583 | León Scieur (BEL) | La Sportive | | | 1921 |
| 1922 | 5386 | Émile Georget (FRA) | Peugeot | 24.2 | + 222.3 | 1922 |
| 1923 | 5386 | Henri Pélissier (FRA) | Autosport-Hutchinson | 24.2 | + 224.4 | 1923 |
| 1924 | 5745 | Ottavia Bottecchia (ITA) | Autosport-Hutchinson | 24.1 | + 238.7 | 1924 |
| 1925 | 5745 | Lucien Buysse (BEL) | Autosport-Hutchinson | 24.1 | + 237.7 | 1925 |
| 1926 | 5745 | Nicolas Frantz (LUX) | Acyon-Bulcup | | | 1926 |
| 1927 | 5745 | Nicolas Frantz (LUX) | Acyon-Bulcup | | | 1927 |
| 1928 | 5745 | Nicolas Frantz (LUX) | Acyon-Bulcup | | | 1928 |
| 1929 | 5287 | André Leducq (FRA) | Acyon-Bulcup | 28.3 | + 184.7 | 1929 |
| 1930 | 5287 | André Leducq (FRA) | Acyon-Bulcup | | | 1930 |
| 1931 | 5287 | André Leducq (FRA) | Acyon-Bulcup | | | 1931 |
| 1932 | 5287 | André Leducq (FRA) | France | 29 | + 154.2 | 1932 |
| 1933 | 5287 | André Leducq (FRA) | France | | | 1933 |
| 1934 | 5287 | Antonio Rigo (ITA) | France | | | 1934 |
| 1935 | 5287 | Rosein Haes (BEL) | Belgium | 38.7 | + 141.4 | 1935 |
| 1936 | 5287 | Roger Lapebie (FRA) | France | | | 1936 |
| 1937 | 5287 | Sir Alfredo Binda (ITA) | Italy | | | 1937 |
| 1938 | 5287 | Désiré Deraf (BEL) | Belgium | 31.6 | + 148.1 | 1938 |
| 1939 | 5287 | Désiré Deraf (BEL) | Belgium | | | 1939 |
| 1940-1946 Tour suspended because of World War II | | | | | | |
| 1947 | 4922 | Jean Robic (FRA) | France | | | 1947 |
| 1948 | 4922 | Giaco Bartali (ITA) | Italy | 33.4 | + 167.2 | 1948 |
| 1949 | 4922 | Fausto Coppi (ITA) | Italy | | | 1949 |
| 1950 | 4922 | René Vietto (FRA) | France | | | 1950 |
| 1951 | 4499 | Hugo Koblet (SUI) | Switzerland | 32.9 | + 142.5 | 1951 |
| 1952 | 4499 | Paul Koechlin (FRA) | France | | | 1952 |
| 1953 | 4499 | Louis Bobet (FRA) | France | | | 1953 |
| 1954 | 4458 | Louis Bobet (FRA) | France | 33.2 | + 148.1 | 1954 |
| 1955 | 4458 | Roger Wileman (FRA) | France | | | 1955 |
| 1956 | 4458 | Jacques Anquetil (FRA) | France | 34.4 | + 135.7 | 1956 |
| 1957 | 4469 | Charly Gaul (LUX) | Luxembourg | | | 1957 |
| 1958 | 4469 | René Pottier (FRA) | France | | | 1958 |
| 1959 | 4469 | Gaston Menzini (ITA) | Italy | 37.2 | + 112.1 | 1959 |
| 1960 | 4473 | Jacques Anquetil (FRA) | France | | | 1960 |
| 1961 | 4473 | Jacques Anquetil (FRA) | France | | | 1961 |
| 1962 | 4138 | Jacques Anquetil (FRA) | France | | | 1962 |
| 1963 | 4138 | Jacques Anquetil (FRA) | France | 36.5 | + 113.5 | 1963 |
| 1964 | 4138 | Jacques Anquetil (FRA) | France | | | 1964 |
| 1965 | 4138 | François Géminiani (ITA) | Italy | | | 1965 |
| 1966 | 4138 | François Géminiani (ITA) | Italy | 36.0 | + 117.6 | 1966 |
| 1967 | 4138 | Roger Pingeon (FRA) | Peugeot-GP-Michelin | | | 1967 |
| 1968 | 4138 | Eddy Merckx (BEL) | Peugeot-GP-Michelin | 35.4 | + 116.3 | 1968 |
| 1969 | 4117 | Eddy Merckx (BEL) | Peugeot-GP-Michelin | | | 1969 |
| 1970 | 4117 | Eddy Merckx (BEL) | Peugeot-GP-Michelin | | | 1970 |
| 1971 | 3846 | Eddy Merckx (BEL) | Molteni | 35.5 | + 108.3 | 1971 |
| 1972 | 3846 | Eddy Merckx (BEL) | Molteni | | | 1972 |
| 1973 | 3846 | Eddy Merckx (BEL) | Molteni | | | 1973 |
| 1974 | 3846 | Eddy Merckx (BEL) | Molteni | | | 1974 |
| 1975 | 3846 | Eddy Merckx (BEL) | Molteni | | | 1975 |
| 1976 | 3846 | Eddy Merckx (BEL) | Ullman-Campagnolo | 34.9 | + 114.6 | 1976 |
| 1977 | 3908 | Bernard Thévenet (FRA) | Peugeot-Esses-Michelin | | | 1977 |
| 1978 | 3908 | Bernard Thévenet (FRA) | Peugeot-Esses-Michelin | 34.1 | + 108.3 | 1978 |
| 1979 | 3908 | Bernard Thévenet (FRA) | Peugeot-Esses-Michelin | | | 1979 |
| 1980 | 3753 | Bernard Hinault (FRA) | Renault-Gitane | | | 1980 |
| 1981 | 3753 | Bernard Hinault (FRA) | Renault-Gitane | 39 | + 96.3 | 1981 |
| 1982 | 3753 | Bernard Hinault (FRA) | Renault-Gitane | | | 1982 |
| 1983 | 3753 | Laurent Fignon (FRA) | Renault-Gitane | | | 1983 |
| 1984 | 4821 | Laurent Fignon (FRA) | Renault-Gitane | 35.9 | + 112.1 | 1984 |
| 1985 | 4821 | Laurent Fignon (FRA) | Renault-Gitane | | | 1985 |
| 1986 | 4231 | Greg LeMond (USA) | La Vie Claire | | | 1986 |
| 1987 | 4231 | Pedro Delgado (ESP) | Ag Axa-San-Servando-Vayphased | 36.6 | + 115.5 | 1987 |
| 1988 | 4231 | Pedro Delgado (ESP) | Z-Torosso | | | 1988 |
| 1989 | 3564 | Greg LeMond (USA) | Ag Axa-San-Servando-Vayphased | 38.6 | + 99.7 | 1989 |
| 1990 | 3714 | Miguel Indurain (ESP) | Banesto | | | 1990 |
| 1991 | 3714 | Miguel Indurain (ESP) | Banesto | 58.7 | + 96 | 1991 |
| 1992 | 3714 | Miguel Indurain (ESP) | Banesto | | | 1992 |
| 1993 | 3714 | Miguel Indurain (ESP) | Banesto | | | 1993 |
| 1994 | 3714 | Miguel Indurain (ESP) | Banesto | | | 1994 |
| 1995 | 3746 | Jan Ulrich (DEU) | Team Telekom | 39.2 | + 96 | 1995 |
| 1996 | 3746 | Lance Armstrong (USA) | US Postal Service | 48.3 | + 91.5 | 1996 |
| 1997 | 3746 | Lance Armstrong (USA) | US Postal Service | | | 1997 |
| 1998 | 3487 | Lance Armstrong (USA) | US Postal Service | 39.9 | + 82.1 | 1998 |
| 1999 | 3487 | Cadel Evans (AUS) | Team Telekom | 41.7 | + 86.3 | 1999 |
| 2000 | 3487 | Alberto Contador (ESP) | Discovery Channel | 40.5 | + 87.9 | 2000 |
| 2001 | 3487 | Alberto Contador (ESP) | Astana | | | 2001 |
| 2002 | 5272 | Alberto Contador (ESP) | Discovery Channel | 39.9 | + 86.2 | 2002 |
| 2003 | 5272 | Alberto Contador (ESP) | Discovery Channel | | | 2003 |
| 2004 | 5272 | Alberto Contador (ESP) | Discovery Channel | | | 2004 |
| 2005 | 3595 | Oscar Pereiro (ESP) | Caisse d'Épargne-Illles Balears | | | 2005 |
| 2006 | 3595 | Oscar Pereiro (ESP) | Caisse d'Épargne-Illles Balears | | | 2006 |
| 2007 | 3559 | Alberto Contador (ESP) | Discovery Channel | | | 2007 |
| 2008 | 3559 | Alberto Contador (ESP) | Discovery Channel | | | 2008 |
| 2009 | 3559 | Alberto Contador (ESP) | Astana | | | 2009 |
| 2010 | 3438 | Cadel Evans (AUS) | Team BMC Racing Team | | | 2010 |
| 2011 | 3438 | Alberto Contador (ESP) | Team BMC Racing Team | 39.8 | + 86.2 | 2011 |
| 2012 | 3438 | Alberto Contador (ESP) | Team BMC Racing Team | | | 2012 |
| 2013 | 3438 | Chris Froome (GBR) | Team Sky | | | 2013 |
| 2014 | 3668.5 | Chris Froome (GBR) | Team Sky | 40.7 | + 99 | 2014 |
| 2015 | 3668.5 | Chris Froome (GBR) | Team Sky | | | 2015 |
| 2016 | 3548 | Chris Froome (GBR) | Team Sky | | | 2016 |
| 2017 | 3548 | Chris Froome (GBR) | Team Sky | 41 | + 86.3 | 2017 |
| 2018 | 3548 | Graeme Obree (GBR) | Team Sky | | | 2018 |
| 2019 | 3548 | Graeme Obree (GBR) | Team Sky | | | 2019 |

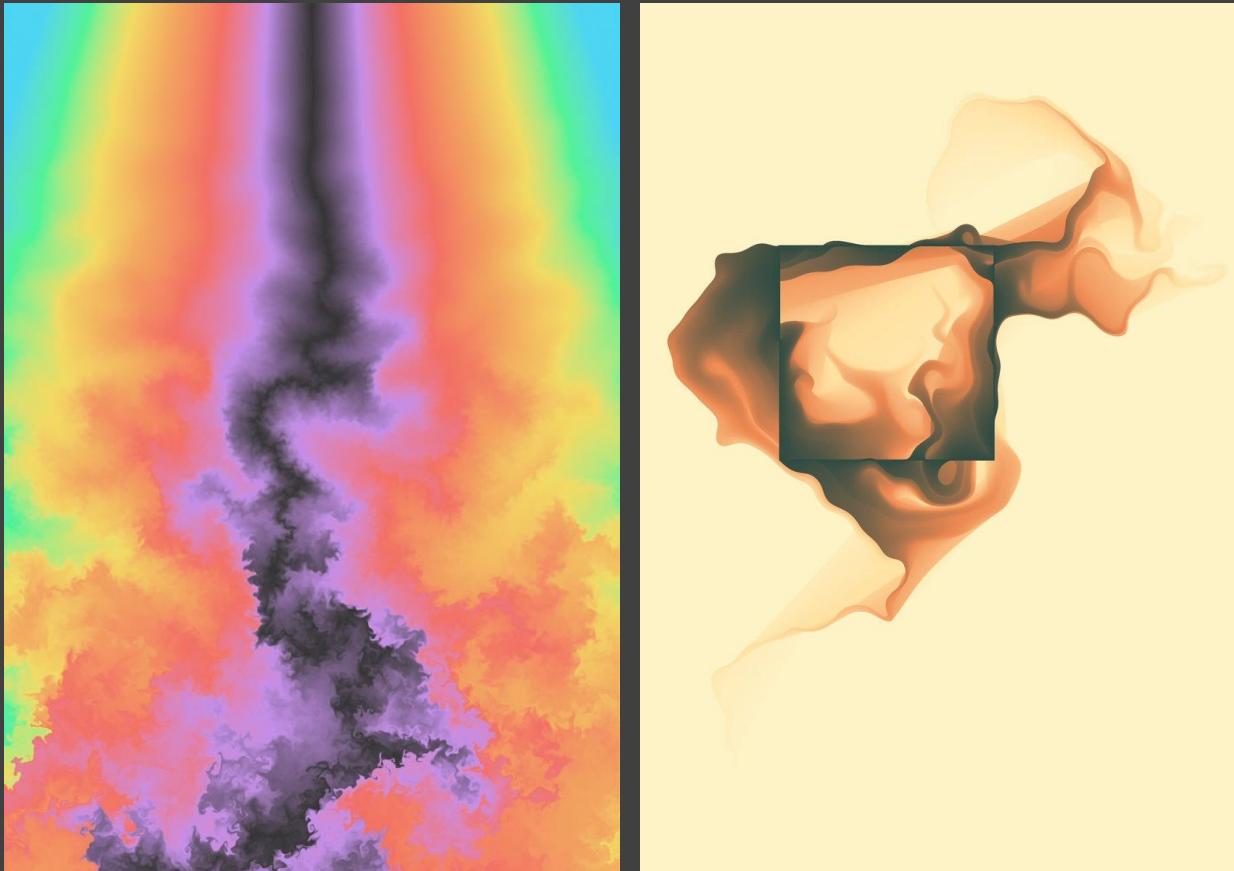
The Showcase

Pixel Art



The Showcase

Generative Art



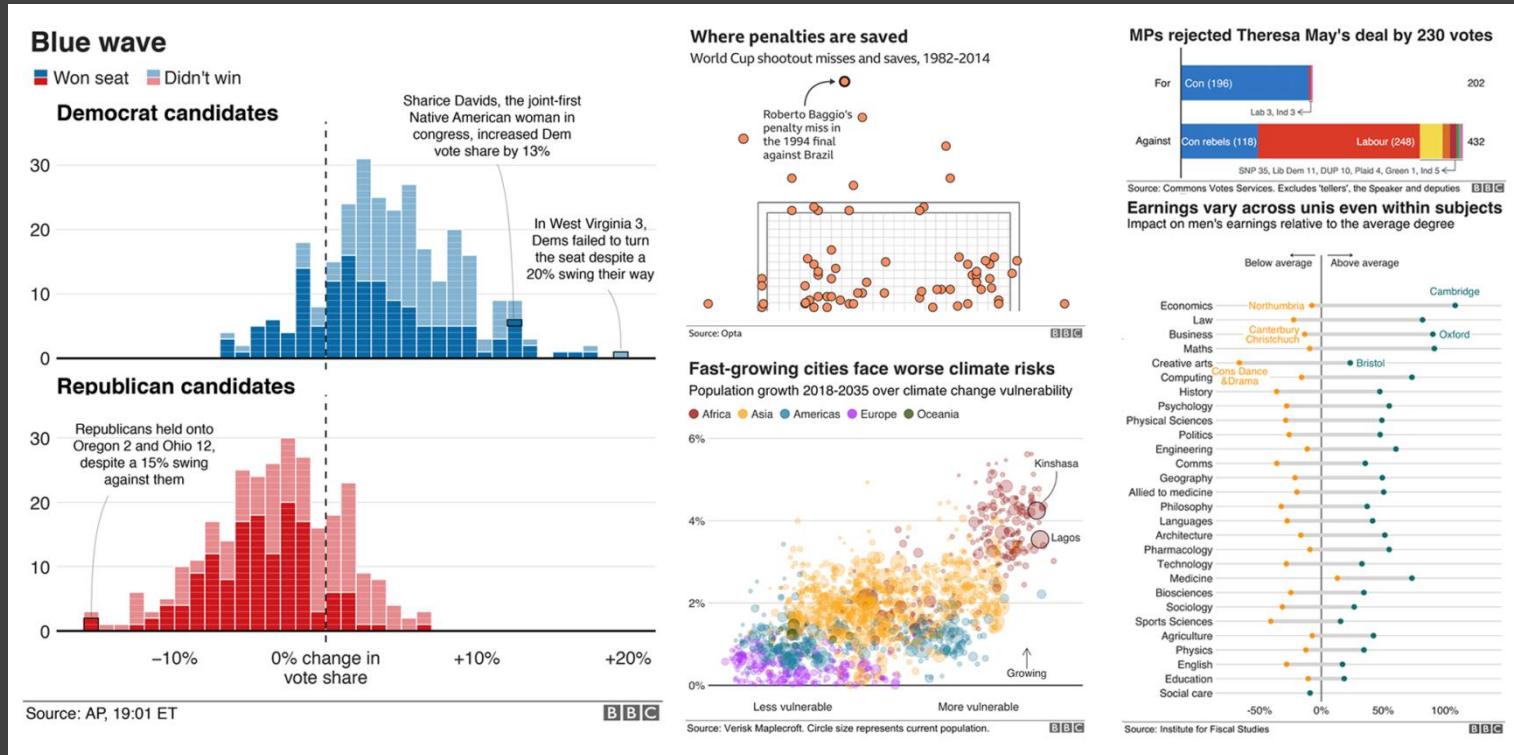
with the help of



Thomas Lin Pedersen

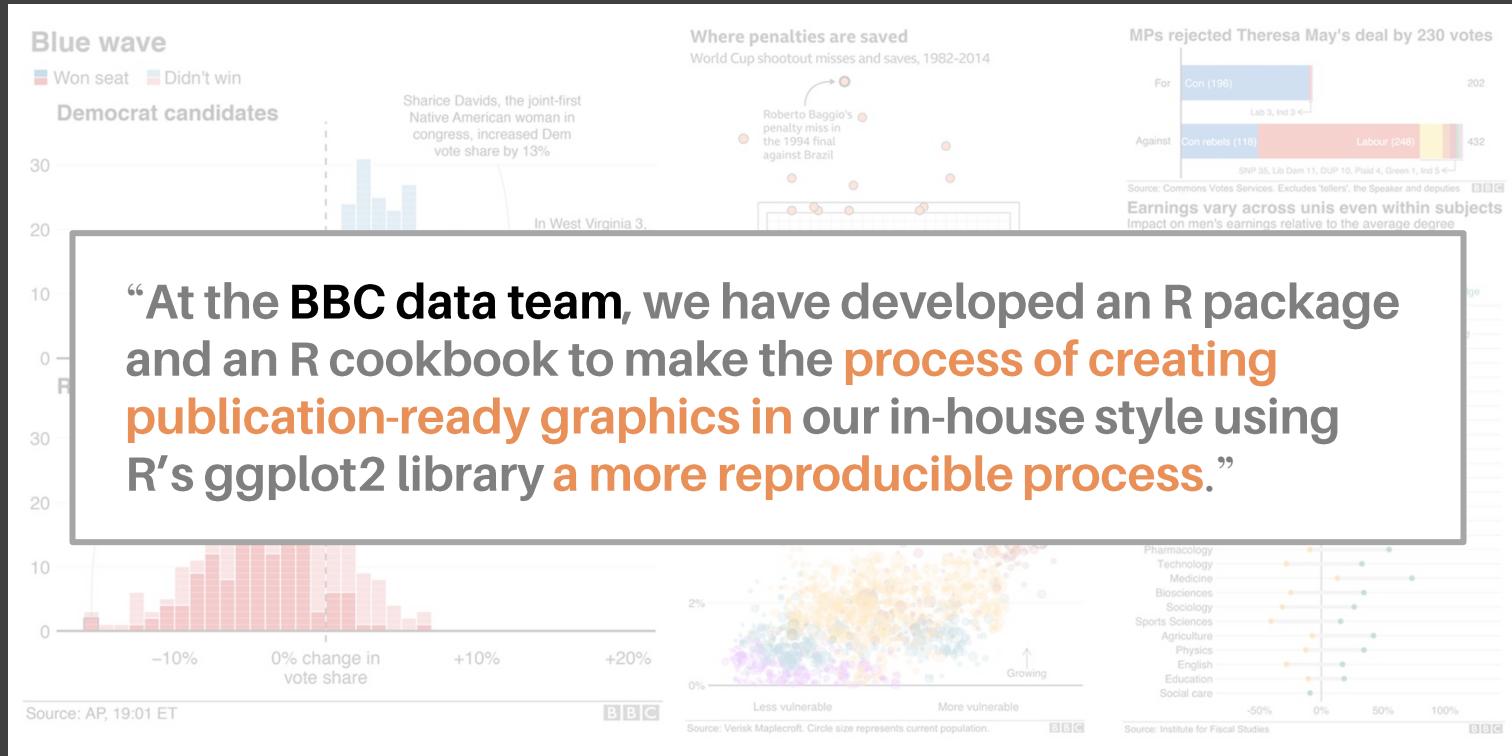
The Showcase

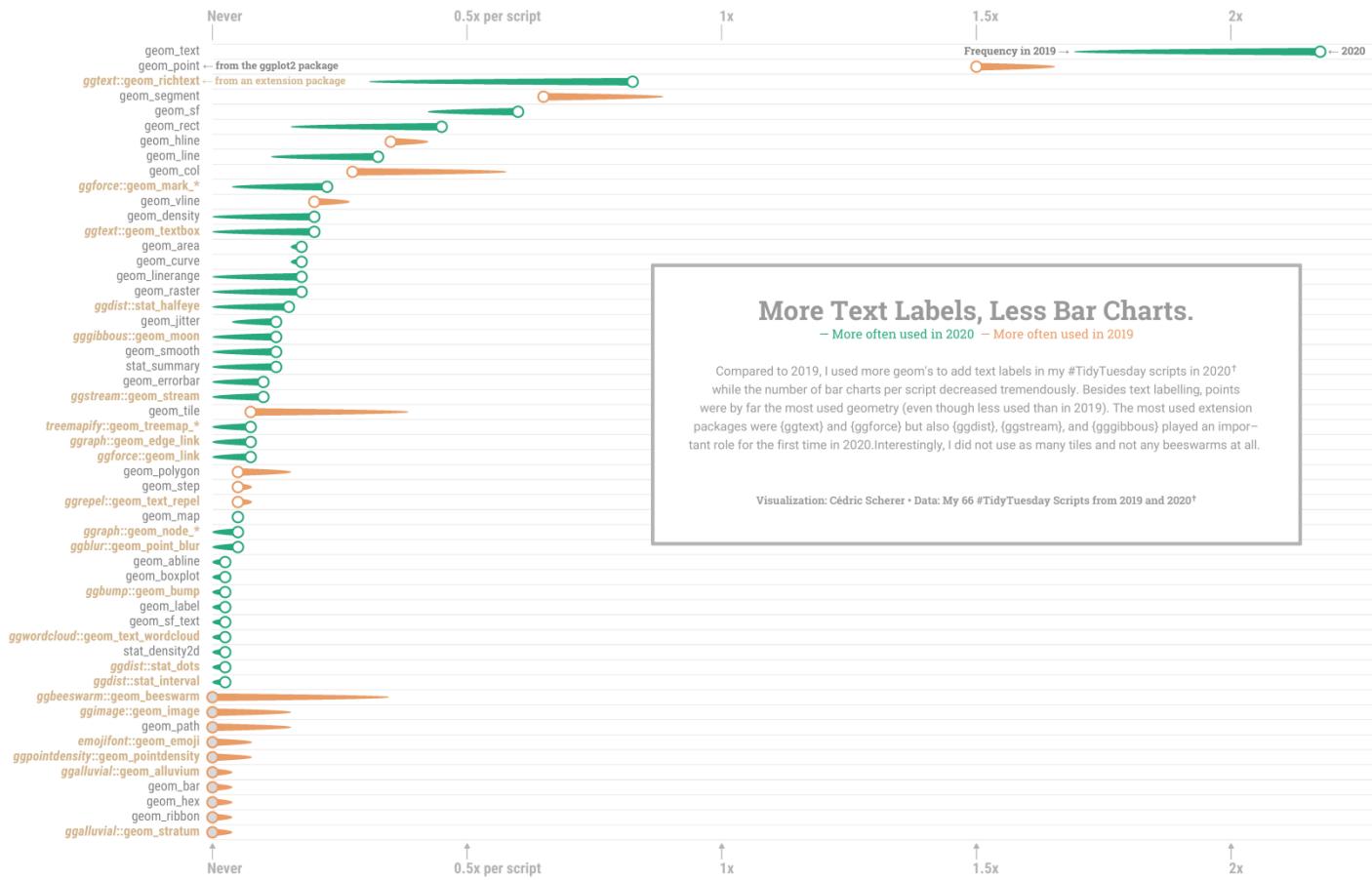
Consistent Theming



The Showcase

Consistent Theming + Reproducibility





[†] I extracted all functions starting with geom or stat from my Rmd files containing the code for all my #TidyTuesday contributions (thanks Georgios for the idea and script). For the contributions from 2019 ($n = 26$) and 2020 ($n = 40$) I calculated the frequency of usage per year for each geom/stat as times used divided by the number of contributions. Note that some geom's which usually appear together (e.g. treemapify::geom_treemap functions) or behave very similarly (e.g. ggforce::geommark functions) were grouped together.



More Text Labels, Less Bar Charts.

— More often used in 2020 — More often used in 2019

Compared to 2019, I used more geom's to add text labels in my #TidyTuesday scripts in 2020[†] while the number of bar charts per script decreased tremendously. Besides text labelling, points were by far the most used geometry (even though less used than in 2019). The most used extension packages were (ggtext) and (ggforce) but also (ggdist), (ggstream), and (gggibbous) played an important role for the first time in 2020. Interestingly, I did not use as many tiles and not any beeswarms at all.

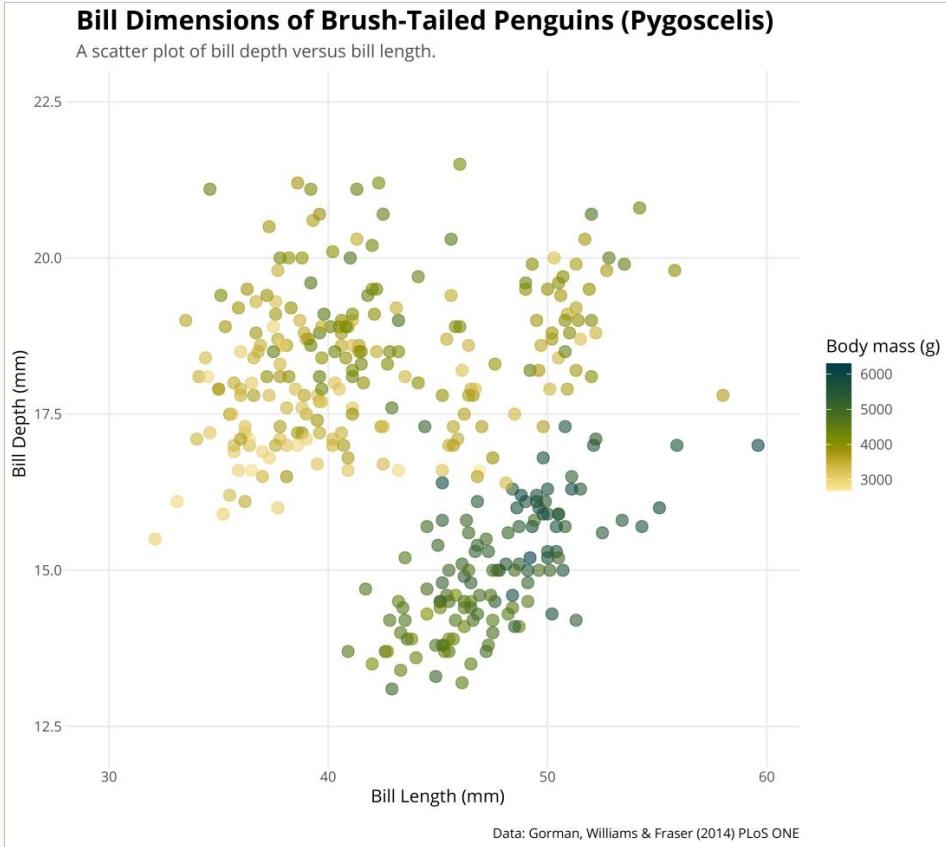
Visualization: Cédric Scherer • Data: My 66 #TidyTuesday Scripts from 2019 and 2020[†]

[†] I extracted all functions starting with `geom` or `stat` from my Rmd files containing the code for all my #TidyTuesday contributions (thanks Georgios for the idea and script). For the contributions from 2019 ($n = 26$) and 2020 ($n = 40$) I calculated the frequency of usage per year for each geomist as times used divided by the number of contributions. Note that some geom's which usually appear together (e.g. `treemapify::geom_treemap` functions) or behave very similarly (e.g. `ggforce::geommark` functions) were grouped together.

{ggtext}

Improved Text Rendering Support

{ggtext} Improved Text Rendering Support



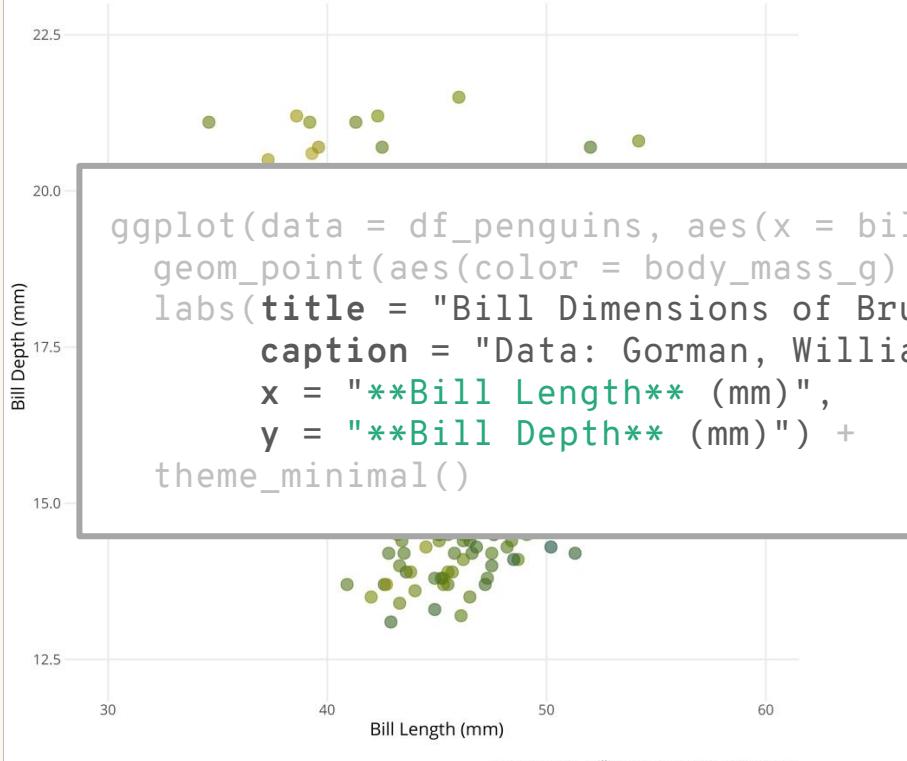
`element_markdown()`

- formatted text elements,
e.g. titles, caption, axis text, striptext

{ggtext} Improved Text Rendering Support

Bill Dimensions of Brush-Tailed Penguins (Pygoscelis)

A scatter plot of bill depth versus bill length.



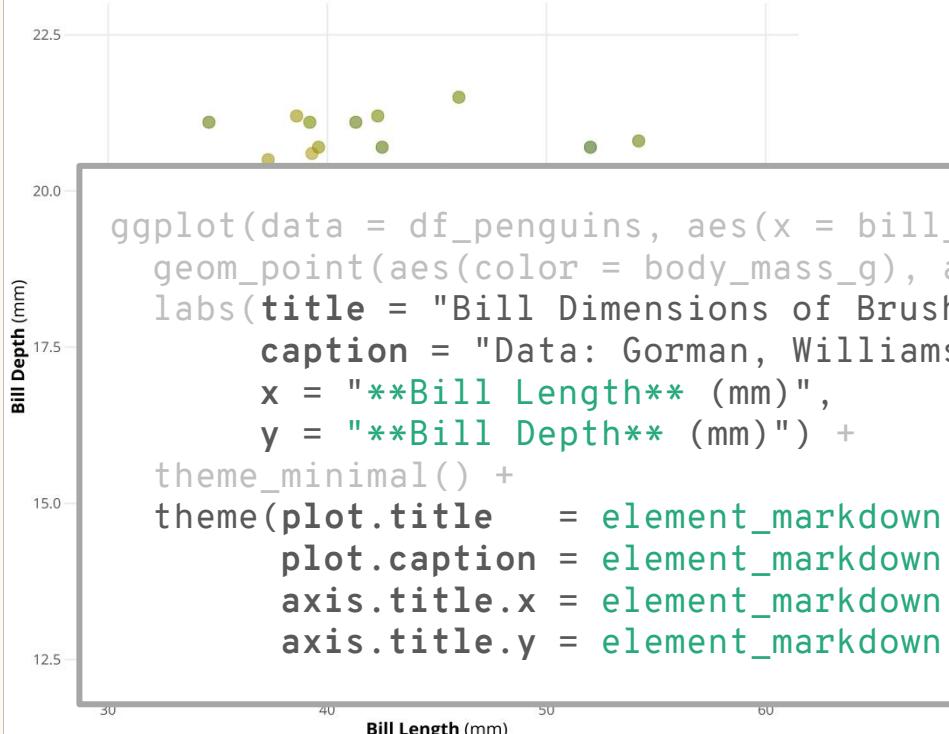
element_markdown()

- formatted text elements,
e.g. titles, caption, axis text, striptext

{ggtext} Improved Text Rendering Support

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.



element_markdown()

→ formatted text elements,
e.g. titles, caption, axis text, striptext

{ggtext} Improved Text Rendering Support

```
<i style='color:#28A87D;'>Pygoscelis</i>
```

Bill Dimensions of Brush-Tailed Penguins *Pygoscelis*

```
<b style='font-size:32pt;font-family:blacksword;'>Pygoscelis</b>
```

Bill Dimensions of Brush-Tailed Penguins *Pygoscelis*

```
<img src='https://cedricscherer.com/img/pygoscelis.jpg', width='10' />
```

Bill Dimensions of Brush-Tailed Penguins



Chats about Friends and their Past, Present, and Future Partners

Mentions of the main characters and their most popular partners in dialogues* during the ten seasons of Friends.



Font Color
Font Face

Ross & Rachel
Rachel & Barry

*For each of the 67,373 dialogues in 236 episodes it was determined whether the two names occur in the same text.
The area and luminance of the squares is mapped to the number of overall mentions of the two names per season.

Visualization by Cédric Scherer • Data by Emil Hvitfeldt via the *(friends)* R package

Contribution to #TidyTuesday 2020/37

Chats about Friends and their Past, Present, and Future Partners

```
tibble(  
  key = c("Chandler", "Joey", "Monica", "Monica &  
         Chandler", "Phoebe", "Rachel", "Rachel &  
         Joey", "Ross", "Ross & Rachel"),  
  color = c("#48508c", "#55331d", "#a64d64",  
           "#774f78", "#5b7233", "#ba2a22",  
           "#882f20", "#f6ab18", "#d86b1d"))  
)
```

Ross & Carol
Ross & Julie
Ross & Bonnie
Ross & Emily
Ross & Elizabeth
Ross & Mona
Ross & Charlie
Ross & Rachel
Rachel & Barry
Rachel & Paolo
Rachel & Tag
Rachel & Joshua
Rachel & Paul
Rachel & Joey
Joey & Kathy
Joey & Janine
Joey & Charlie
Monica & Richard
Monica & Pete
Monica & Chandler
Chandler & Janice
Chandler & Kathy
Phoebe & David
Phoebe & Gary
Phoebe & Mike

Font Color
Font Face

Chats about Friends and their Past, Present, and Future Partners

```
tibble(  
  key = c("Chandler", "Joey", "Monica", "Monica &  
          Chandler", "Phoebe", "Rachel", "Rachel &  
          Joey", "Ross", "Ross & Rachel"),  
  color = c("#48508c", "#55331d", "#a64d64",  
           "#774f78", "#5b7233", "#ba2a22",  
           "#882f20", "#f6ab18", "#d86b1d"))  
)
```

| key | color |
|---------------|---------|
| Ross | #f6ab18 |
| Chandler | #48508c |
| Ross & Rachel | #d86b1d |

Ross & Carol
Ross & Julie
Ross & Bonnie
Ross & Emily
Ross & Elizabeth
Ross & Mona
Ross & Charlie
Ross & Rachel
Rachel & Barry
Rachel & Paolo
Rachel & Tag
Rachel & Joshua
Rachel & Paul
Rachel & Joey
Joey & Kathy
Joey & Janine
Joey & Charlie
Monica & Richard
Monica & Pete
Monica & Chandler
Chandler & Janice
Chandler & Kathy
Phoebe & David
Phoebe & Gary
Phoebe & Mike

Font Color
Font Face



Chats about Friends and their Past, Present, and Future Partners

```
tibble(  
  key = c("Chandler", "Joey", "Monica", "Monica &  
    Chandler", "Phoebe", "Rachel", "Rachel &  
    Joey", "Ross", "Ross & Rachel"),  
  color = c("#48508c", "#55331d", "#a64d64",  
    "#774f78", "#5b7233", "#ba2a22",  
    "#882f20", "#f6ab18", "#d86b1d"))  
) %>%  
right_join(df_friends, by = "key")
```

| key | color | partners | episode |
|---------------|---------|-------------------|---------|
| Ross | #f6ab18 | Ross & Carol | S01E01 |
| Chandler | #48508c | Chandler & Janice | S01E03 |
| Ross & Rachel | #d86b1d | Ross & Rachel | S01E04 |

Ross & Carol
Ross & Julie
Ross & Bonnie
Ross & Emily
Ross & Elizabeth
Ross & Mona
Ross & Charlie
Ross & Rachel
Rachel & Barry
Rachel & Paolo
Rachel & Tag
Rachel & Joshua
Rachel & Paul
Rachel & Joey
Joey & Kathy
Joey & Janine
Joey & Charlie
Monica & Richard
Monica & Pete
Monica & Chandler
Chandler & Janice
Chandler & Kathy
Phoebe & David
Phoebe & Gary
Phoebe & Mike

Font Color
Font Face

Chats about Friends and their Past, Present, and Future Partners

```
tibble(
  key = c("Chandler", "Joey", "Monica", "Monica &
         Chandler", "Phoebe", "Rachel", "Rachel &
         Joey", "Ross", "Ross & Rachel"),
  color = c("#48508c", "#55331d", "#a64d64",
            "#774f78", "#5b7233", "#ba2a22",
            "#882f20", "#f6ab18", "#d86b1d")
) %>%
right_join(df_friends, by = "key") %>%
mutate(partners = if_else(
  key == partners,
  glue::glue("<b style='color:{color};'>
              {partners}</b>"),
  str_replace(partners, key,
             glue::glue("<b style='color:{color};'>
              {key}</b>")))
))
```

Ross & Carol
Ross & Julie
Ross & Bonnie
Ross & Emily
Ross & Elizabeth
Ross & Mona
Ross & Charlie
Ross & Rachel
Rachel & Barry
Rachel & Paolo
Rachel & Tag
Rachel & Joshua
Rachel & Paul
Rachel & Joey
Joey & Kathy
Joey & Janine
Joey & Charlie
Monica & Richard
Monica & Pete
Monica & Chandler
Chandler & Janice
Chandler & Kathy
Phoebe & David
Phoebe & Gary
Phoebe & Mike

Font Color
Font Face

Chats about Friends and their Past, Present, and Future Partners

| key | color | partners |
|---------------|---------|-------------------|
| Ross | #f6ab18 | Ross & Carol |
| Chandler | #48508c | Chandler & Janice |
| Ross & Rachel | #d86b1d | Ross & Rachel |

Ross & Carol
Ross & Julie
Ross & Bonnie
Ross & Emily
Ross & Elizabeth
Ross & Mona
Ross & Charlie
Ross & Rachel
Rachel & Barry
Rachel & Paolo
Rachel & Tag
Rachel & Joshua
Rachel & Paul
Rachel & Joey
Joey & Kathy
Joey & Janine
Joey & Charlie
Monica & Richard
Monica & Pete
Monica & Chandler
Chandler & Janice
Chandler & Kathy
Phoebe & David
Phoebe & Gary
Phoebe & Mike

Font Color
Font Face

```
right_join(df_friends, by = "key") %>%  
  mutate(partners = if_else(  
    key == partners,  
    glue::glue("<b style='color:{color};'>  
              {partners}</b>"),  
    str_replace(partners, key,  
               glue::glue("<b style='color:{color};'>  
              {key}</b>"))  
)
```

Visualization by Cédric Scherer • Data by Emil Hvitfeldt via the (friends) R package

Chats about Friends and their Past, Present, and Future Partners

| key | color | partners | |
|---------------|---------|-------------------|-------|
| Ross | #f6ab18 | Ross & Carol | FALSE |
| Chandler | #48508c | Chandler & Janice | FALSE |
| Ross & Rachel | #d86b1d | Ross & Rachel | TRUE |

Ross & Carol
Ross & Julie
Ross & Bonnie
Ross & Emily
Ross & Elizabeth
Ross & Mona
Ross & Charlie
Ross & Rachel
Rachel & Barry
Rachel & Paolo
Rachel & Tag
Rachel & Joshua
Rachel & Paul
Rachel & Joey
Joey & Kathy
Joey & Janine
Joey & Charlie
Monica & Richard
Monica & Pete
Monica & Chandler
Chandler & Janice
Chandler & Kathy
Phoebe & David
Phoebe & Gary
Phoebe & Mike

Font Color
Font Face

```
right_join(df_friends, by = "key") %>%  
  mutate(partners = if_else(  
    key == partners,  
    glue::glue("<b style='color:{color};'>  
              {partners}</b>"),  
    str_replace(partners, key,  
               glue::glue("<b style='color:{color};'>  
              {key}</b>"))  
)
```

Visualization by Cédric Scherer • Data by Emil Hvitfeldt via the (friends) R package

Chats about Friends and their Past, Present, and Future Partners

| key | color | partners | |
|--------------------------|----------------|---|-------|
| Ross | #f6ab18 | Ross & Carol | FALSE |
| Chandler | #48508c | Chandler & Janice | FALSE |
| Ross & Rachel | #d86b1d | < b style='color:#d86b1d;'>Ross & Rachel | |

```
right_join(df_friends, by = "key") %>%  
  mutate(partners = if_else(  
    key == partners,  
    glue::glue("<b style='color:{color};'>  
               {partners}</b>"),  
    str_replace(partners, key,  
               glue::glue("<b style='color:{color};'>  
               {key}</b>"))  
)
```

Ross & Carol
Ross & Julie
Ross & Bonnie
Ross & Emily
Ross & Elizabeth
Ross & Mona
Ross & Charlie
Ross & Rachel
Rachel & Barry
Rachel & Paolo
Rachel & Tag
Rachel & Joshua
Rachel & Paul
Rachel & Joey
Joey & Kathy
Joey & Janine
Joey & Charlie
Monica & Richard
Monica & Pete
Monica & Chandler
Chandler & Janice
Chandler & Kathy
Phoebe & David
Phoebe & Gary
Phoebe & Mike

Font Color
Font Face

Visualization by Cédric Scherer • Data by Emil Hvitfeldt via the (friends) R package

Chats about Friends and their Past, Present, and Future Partners

| key | color | partners |
|---------------|---------|---|
| Ross | #f6ab18 | <b style='color:#f6ab18;'>Ross & Carol |
| Chandler | #48508c | <b style='color:#48508c;'>Chandler & Janice |
| Ross & Rachel | #d86b1d | <b style='color:#d86b1d;'>Ross & Rachel |

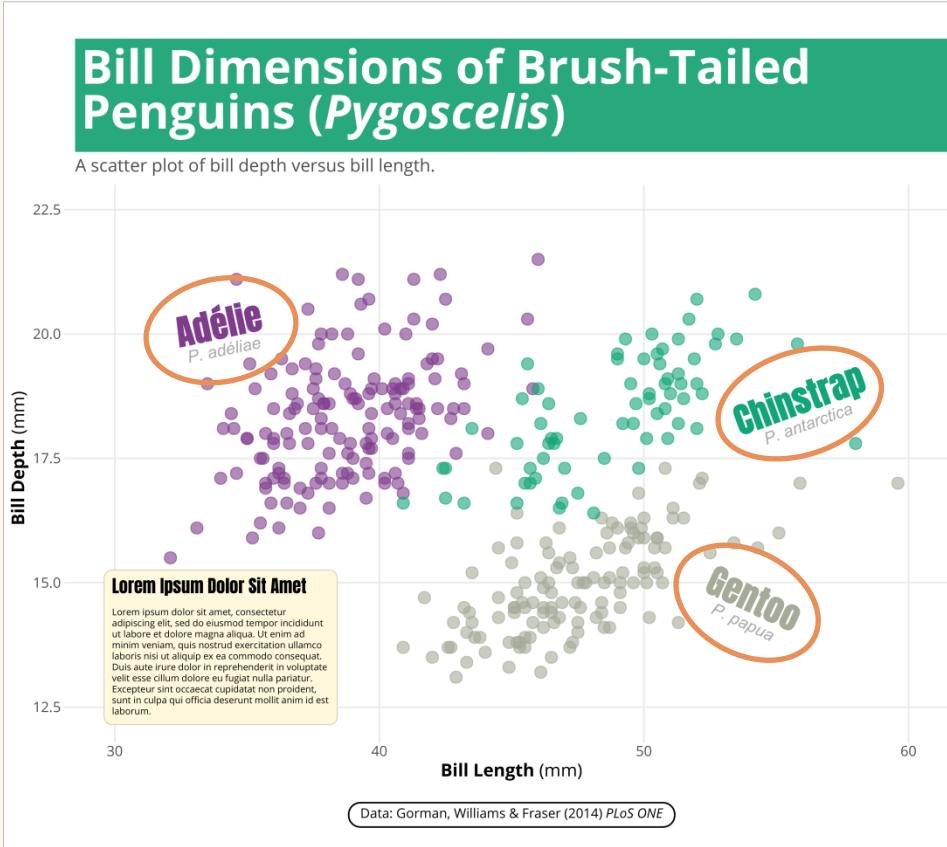
```
right_join(df_friends, by = "key") %>%  
  mutate(partners = if_else(  
    key == partners,  
    glue::glue("<b style='color:{color};'>  
               {partners}</b>"),  
    str_replace(partners, key,  
               glue::glue("<b style='color:{color};'>  
               {key}</b>"))  
)
```

Ross & Carol
Ross & Julie
Ross & Bonnie
Ross & Emily
Ross & Elizabeth
Ross & Mona
Ross & Charlie
Ross & Rachel
Rachel & Barry
Rachel & Paolo
Rachel & Tag
Rachel & Joshua
Rachel & Paul
Rachel & Joey
Joey & Kathy
Joey & Janine
Joey & Charlie
Monica & Richard
Monica & Pete
Monica & Chandler
Chandler & Janice
Chandler & Kathy
Phoebe & David
Phoebe & Gary
Phoebe & Mike

Font Color
Font Face

Visualization by Cédric Scherer • Data by Emil Hvitfeldt via the (friends) R package

{ggtext} Improved Text Rendering Support



`element_markdown()`

- formatted text elements,
e.g. titles, caption, axis text, striptext

`geom_richtext()`

- formatted text labels with 360° rotation

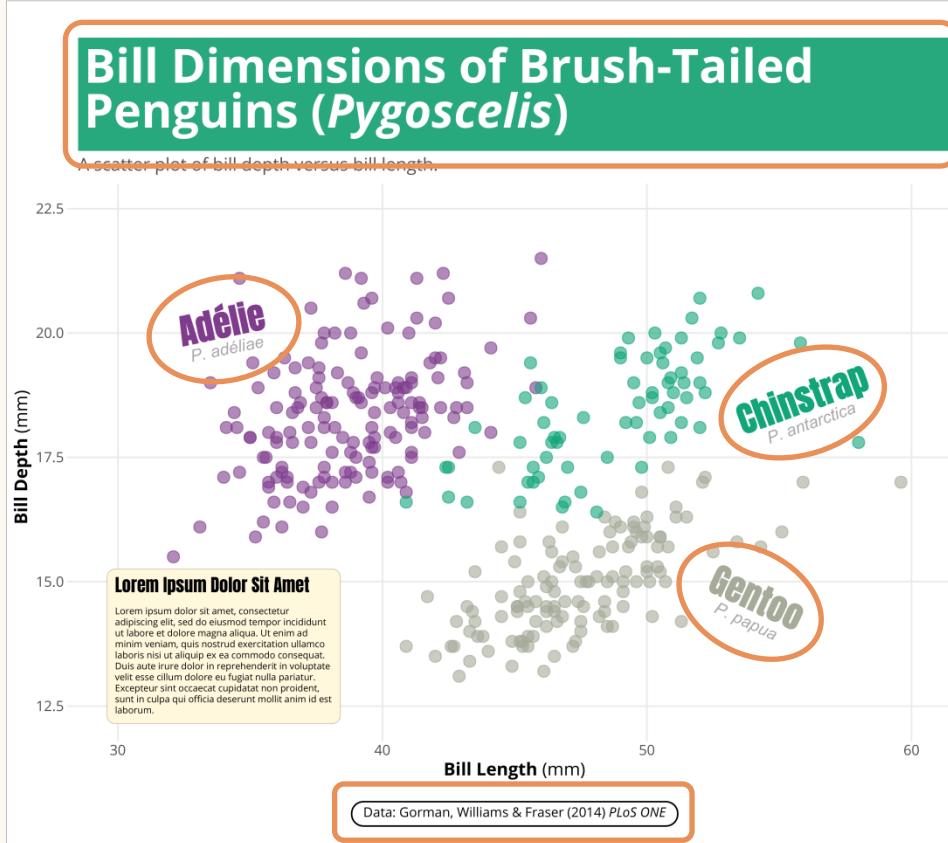
`element_textbox()` and
`element_textbox_simple()`

- formatted text boxes with word wrapping

`geom_textbox()`

- formatted text boxes with word wrapping

{ggtext} Improved Text Rendering Support



`element_markdown()`

- formatted text elements,
e.g. titles, caption, axis text, striptext

`geom_richtext()`

- formatted text labels with 360° rotation

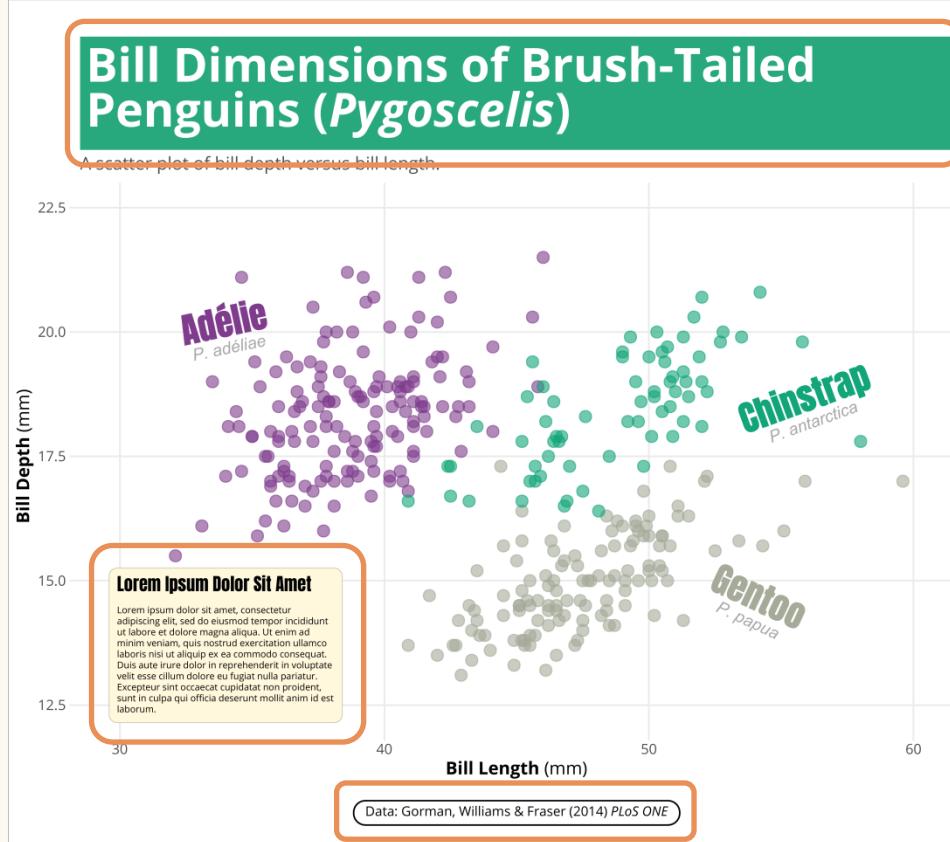
`element_textbox()` and
`element_textbox_simple()`

- formatted text boxes with word wrapping

`geom_textbox()`

- formatted text boxes with word wrapping

{ggtext} Improved Text Rendering Support



`element_markdown()`

→ formatted text elements,
e.g. titles, caption, axis text, striptext

`geom_richtext()`

→ formatted text labels with 360° rotation

`element_textbox()` and
`element_textbox_simple()`

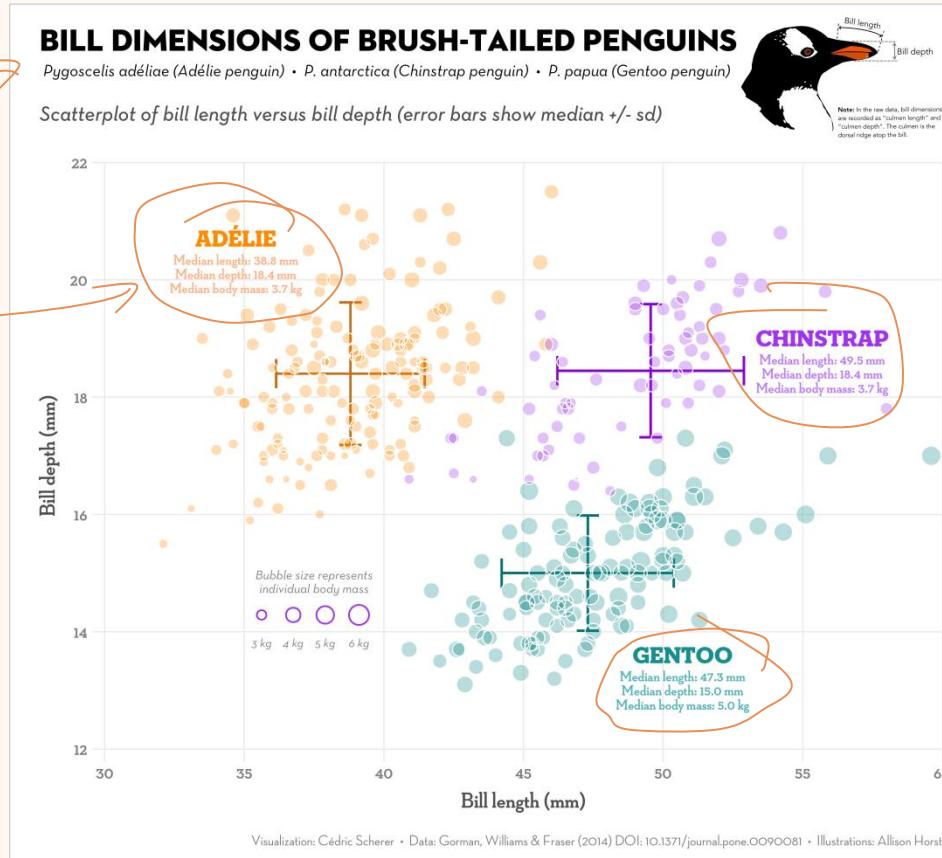
→ formatted text boxes with word wrapping

`geom_textbox()`

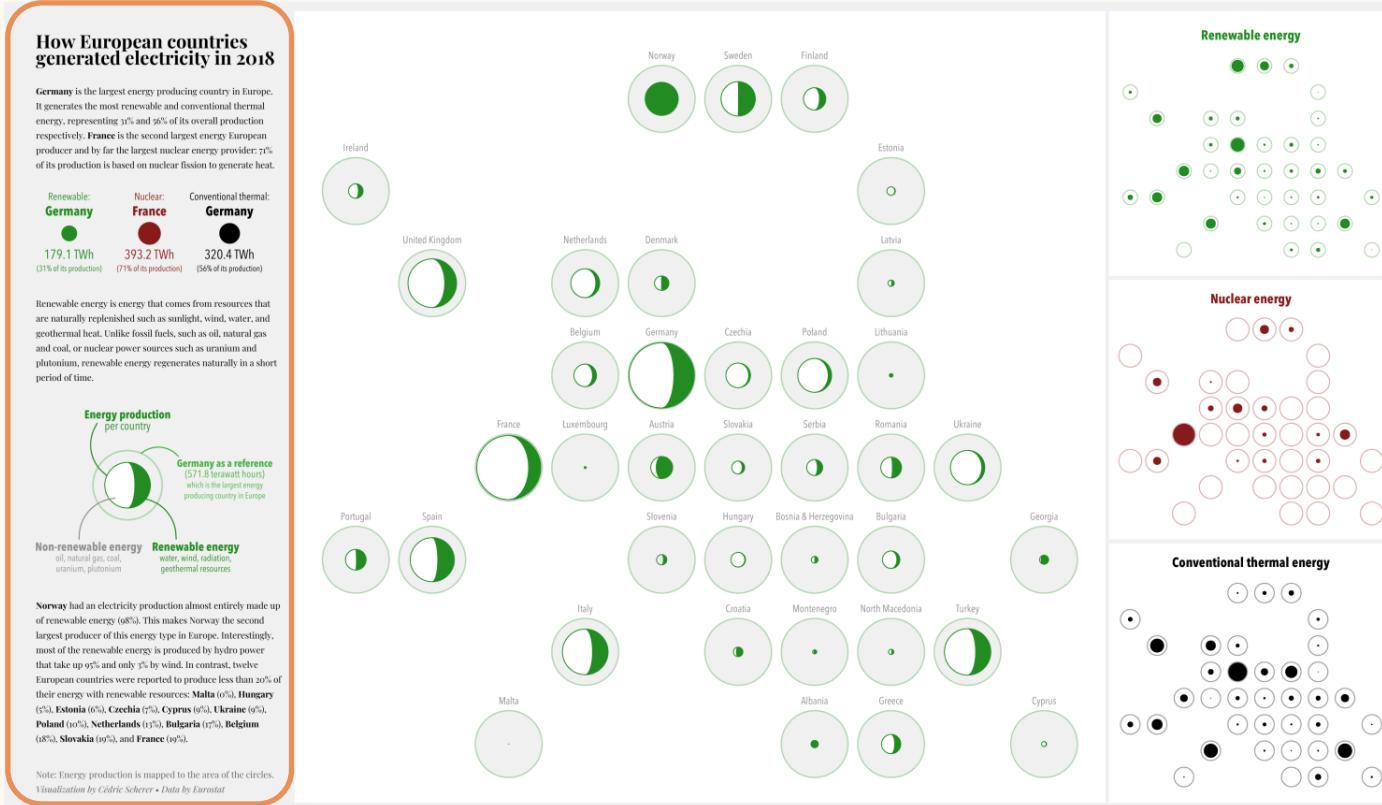
→ formatted text boxes with word wrapping

{ggtext} Improved Text Rendering Support

Font Family
Font Size
Font Face
Lineheight

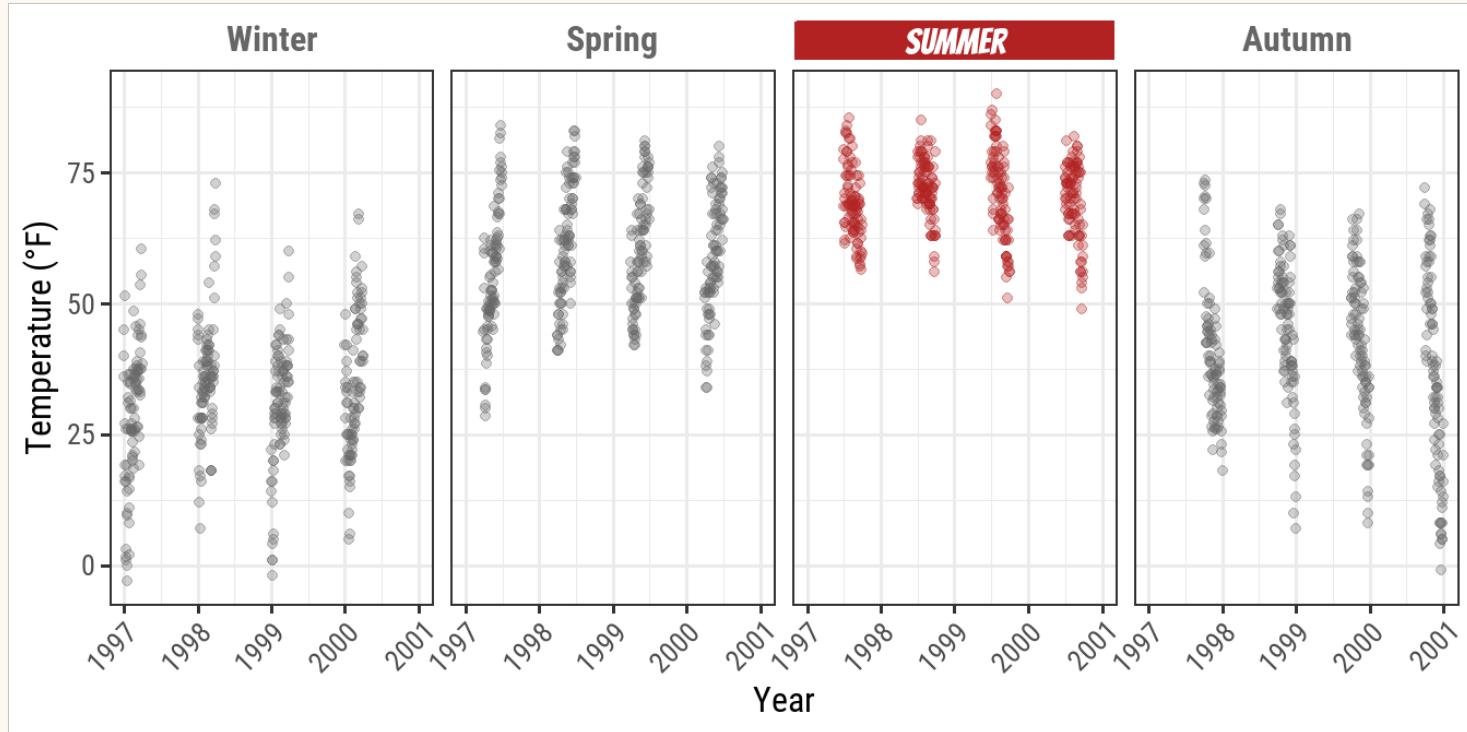


{ggtext} Improved Text Rendering Support



`geom_textbox() +
geom_richtext()`

{ggtext} Improved Text Rendering Support



→ stackoverflow.com/questions/60332202/conditionally-fill-ggtext-text-boxes-in-facet-wrap

→ cedricscherer.com/2019/08/05/a-ggplot2-tutorial-for-beautiful-plotting-in-r/#panels

custom_element_textbox_highlight()

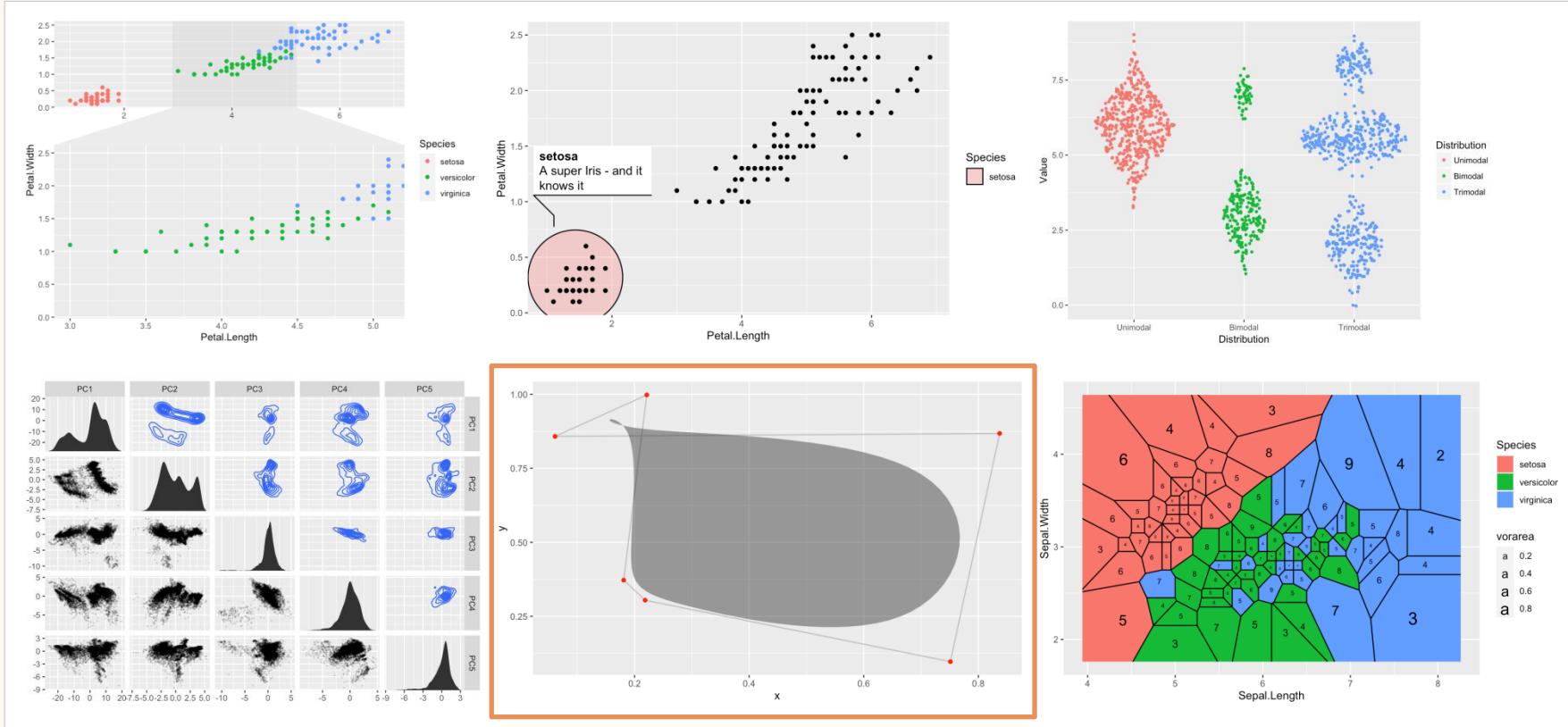
{ggforce}

Providing Missing Functionality to `ggplot2`



ggforce.data-imaginist.com

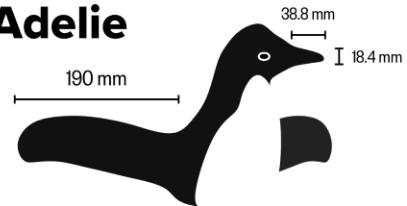
{ggforce} Providing Missing Functionality



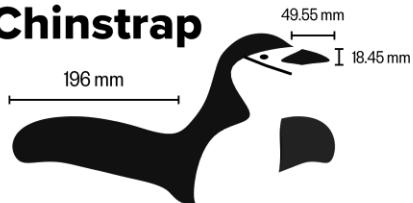
Palmer Penguins

Median length of flipper, length and depth of bill,
of 342 penguins recorded between 2007 and 2009

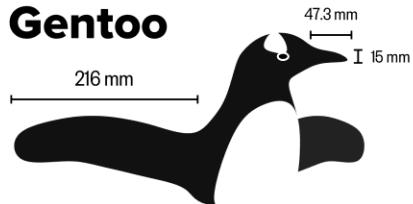
Adelie



Chinstrap

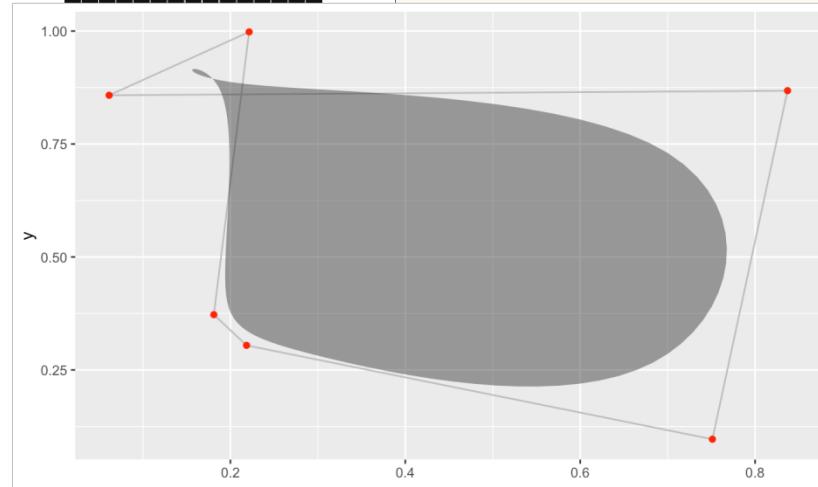


Gentoo



Recorded penguins by species

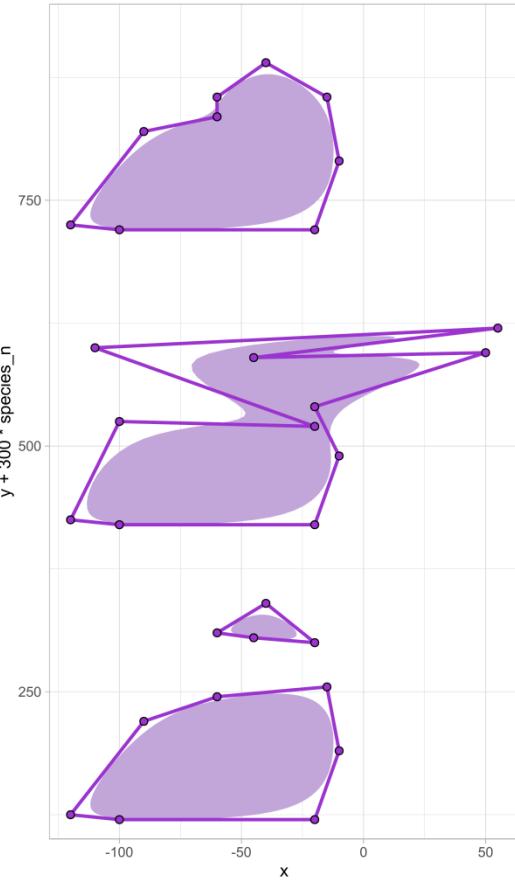
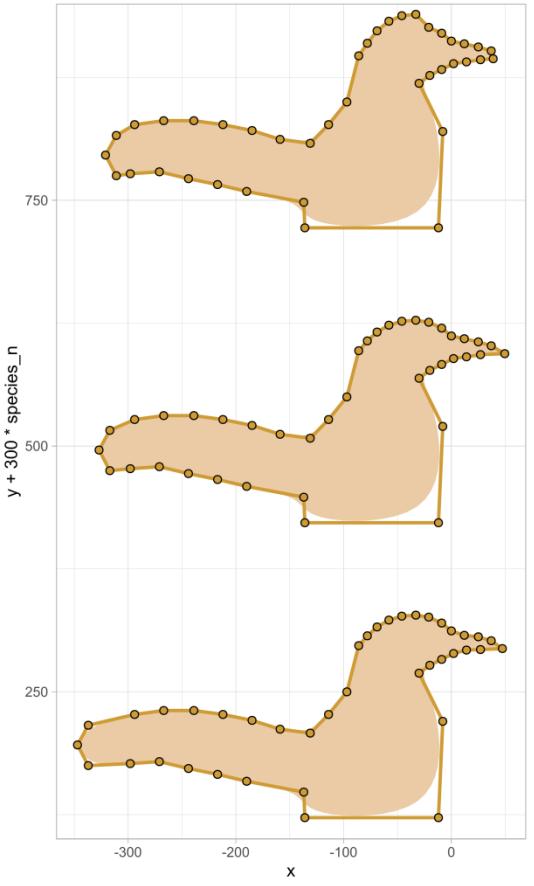
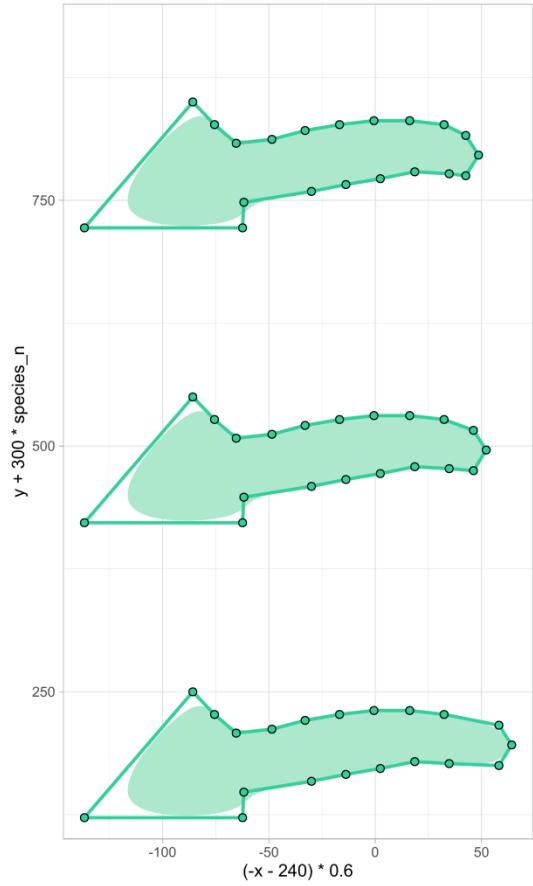
151



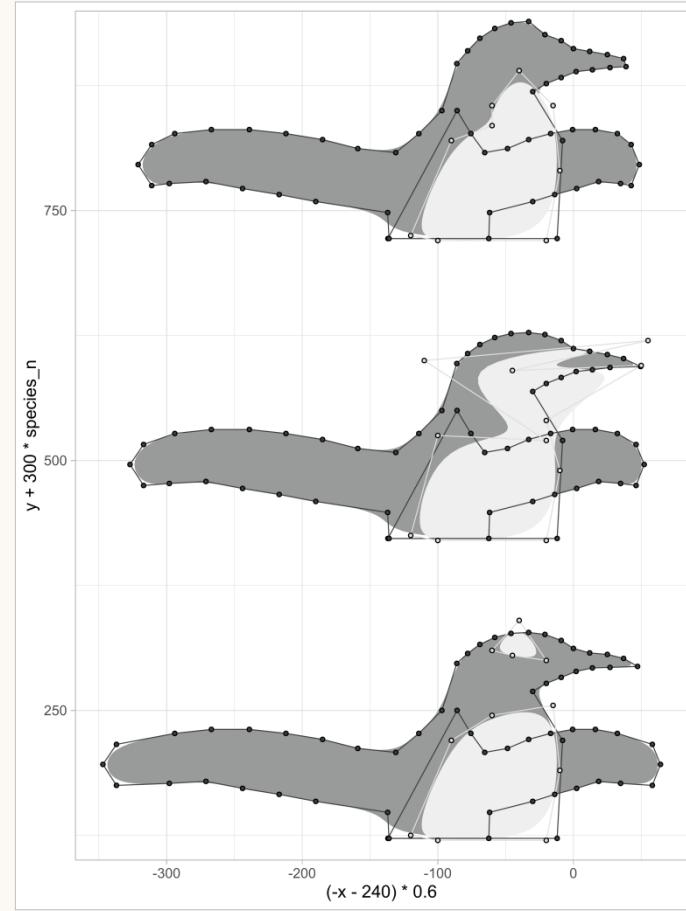
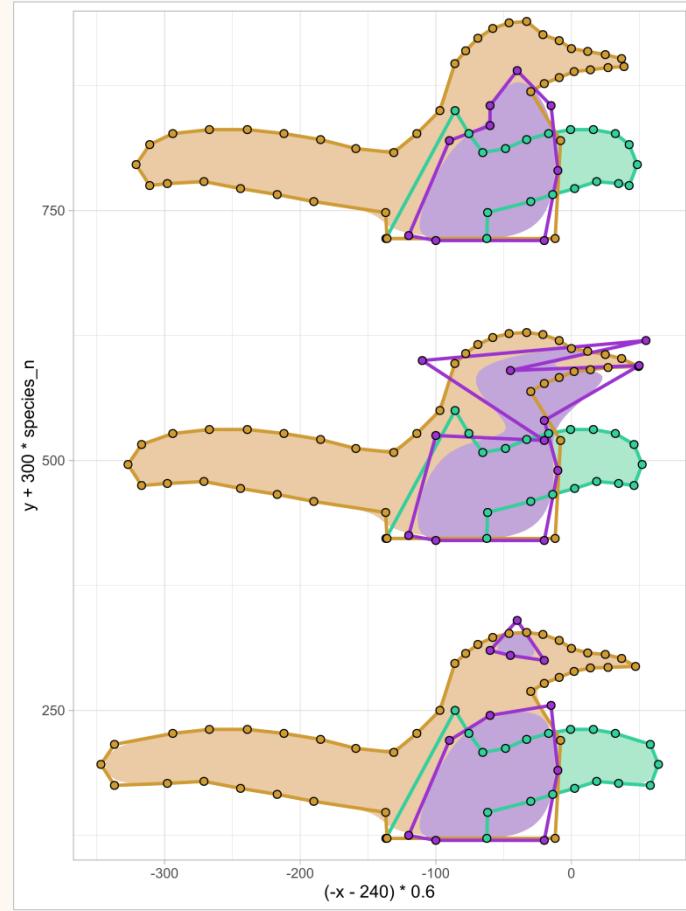
geom_bspline_closed()

Source: Dr. Kristen Gorman and the Palmer Station, Antarctica LTER | Graphic: Georgios Karamanis

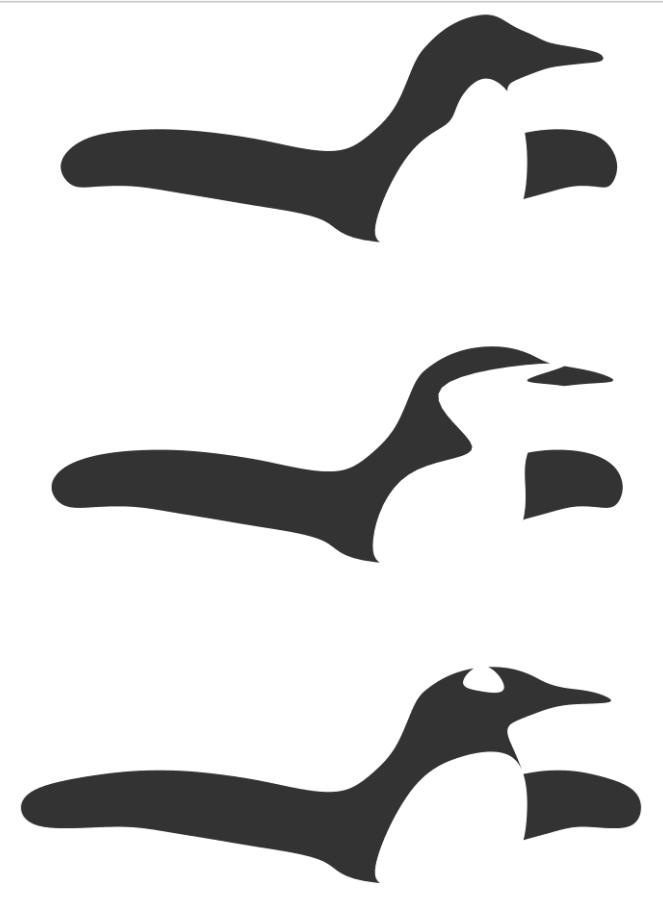
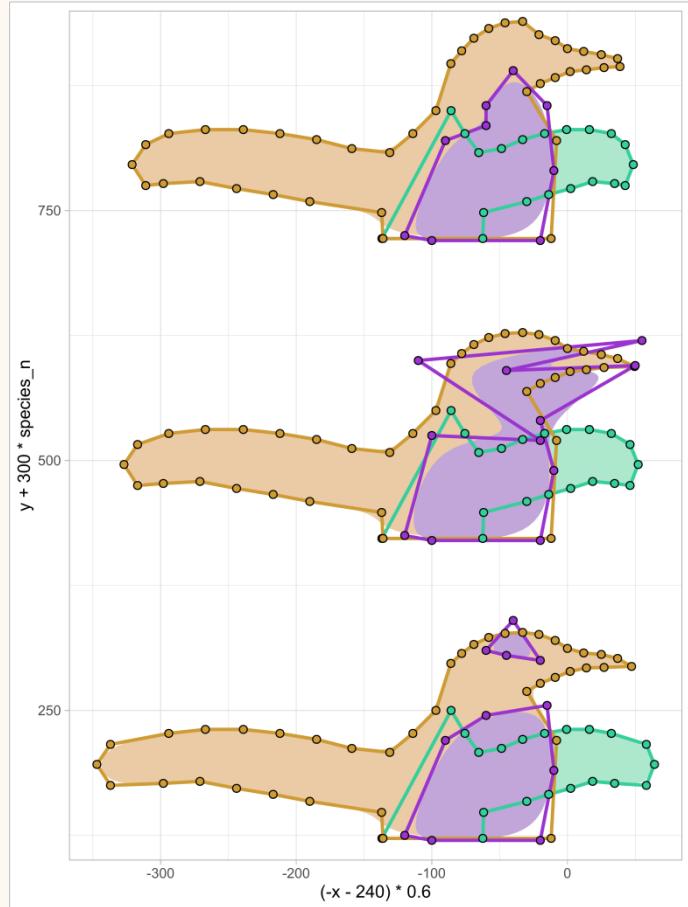
Georgios Karamanis, Contribution to #TidyTuesday 2020/31



Modified from Georgios Karamanis' Contribution to #TidyTuesday 2020/31

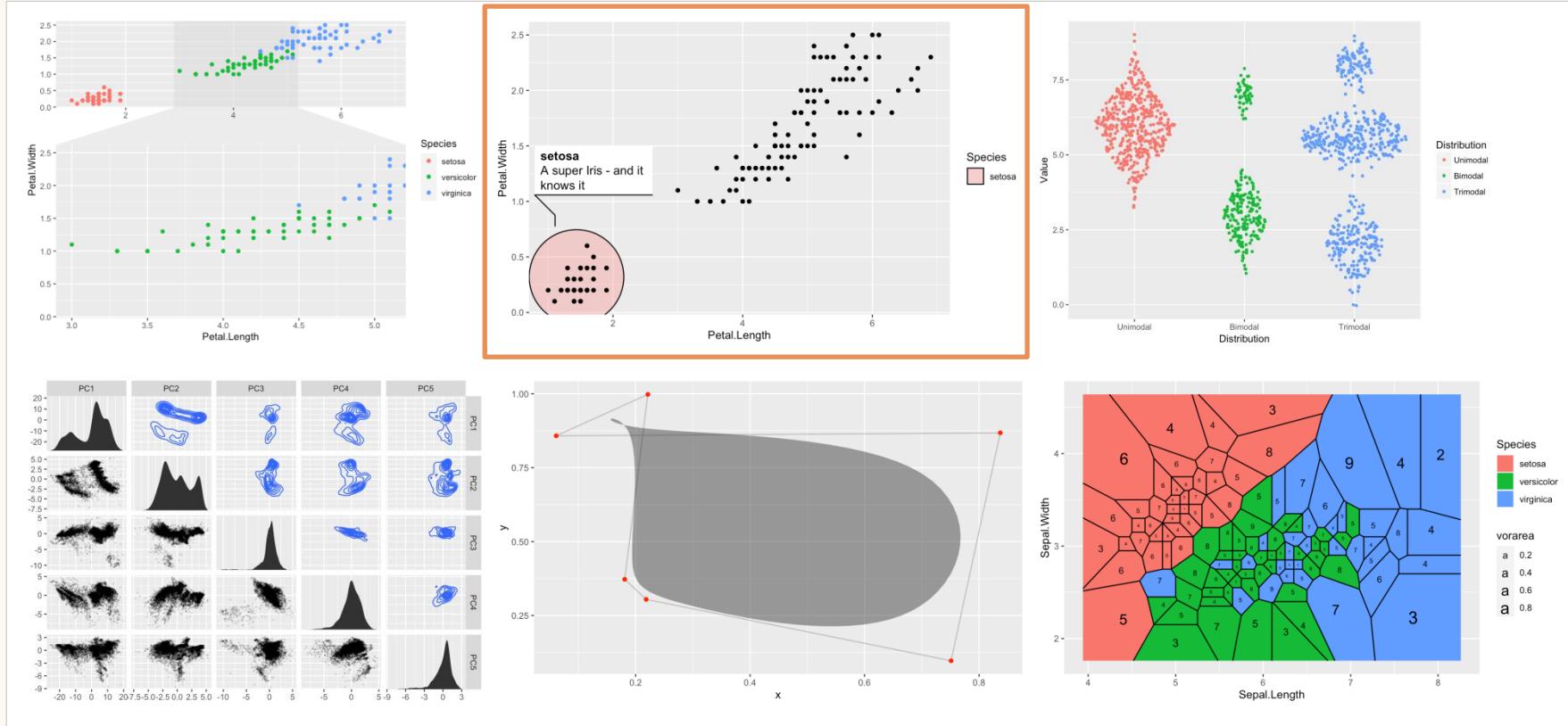


Modified from Georgios Karamanis' Contribution to #TidyTuesday 2020/31



Modified from Georgios Karamanis' Contribution to #TidyTuesday 2020/31

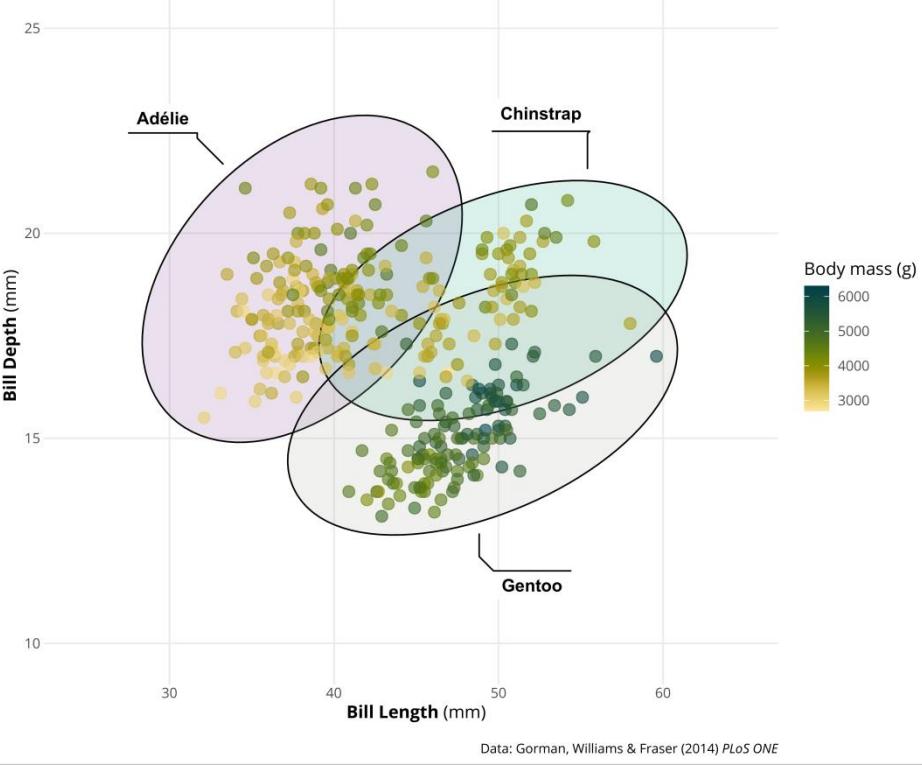
{ggforce} Providing Missing Functionality



{ggforce} Fancy Annotations

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.



`geom_mark_*`

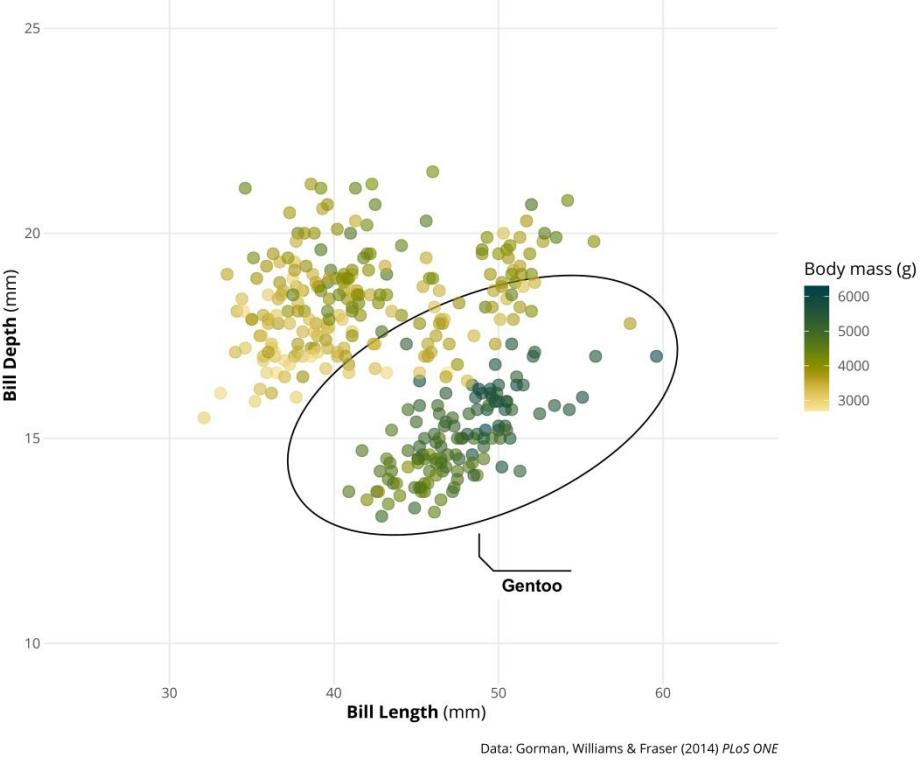
→ advanced labels for single or multiple points

```
geom_mark_ellipsoid(aes(fill = species, label = species))
```

{ggforce} Fancy Annotations

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.



```
geom_mark_ellipsoid(aes(fill = species, label = species,  
filter = species == 'Gentoo'))
```

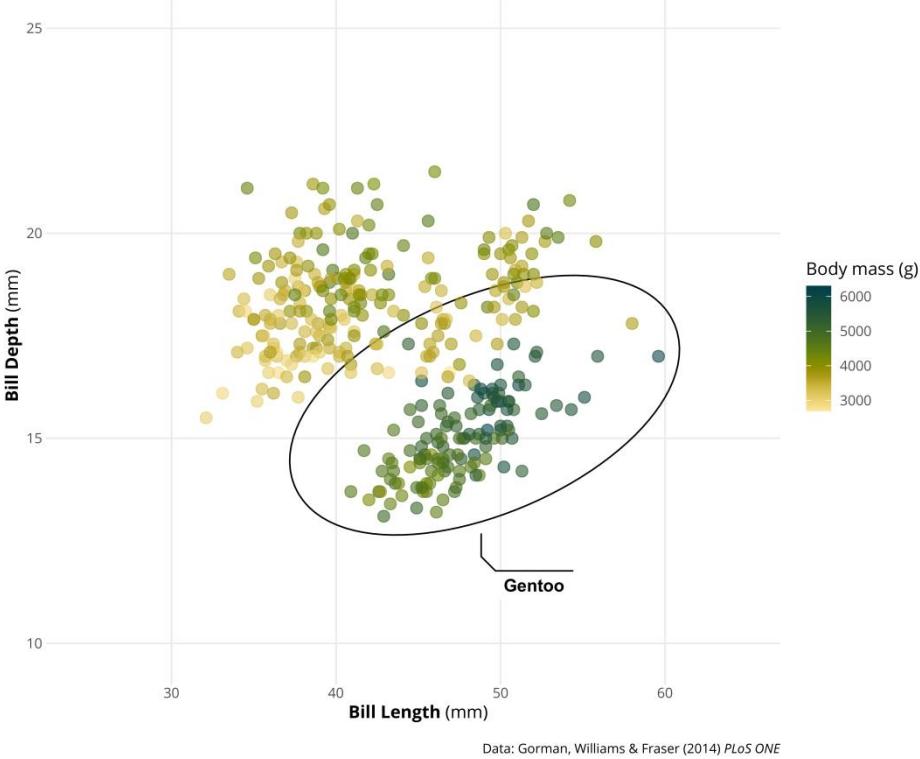
geom_mark_*

- advanced labels for single or multiple points
- show all groups or highlight interesting parts

{ggforce} Fancy Annotations

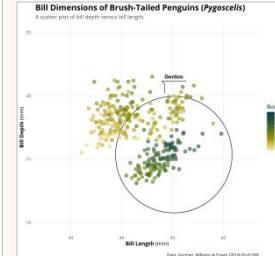
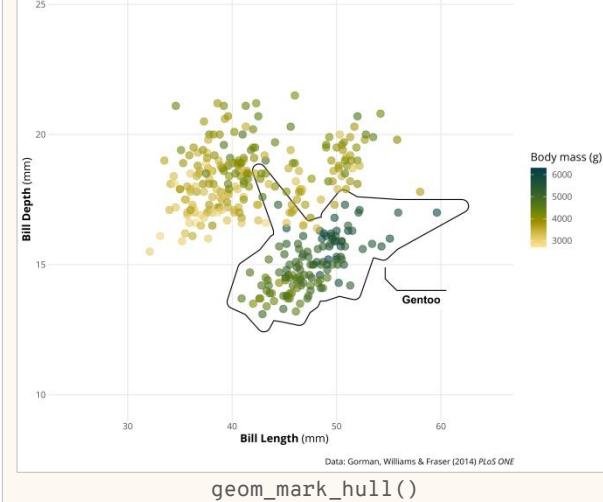
Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.

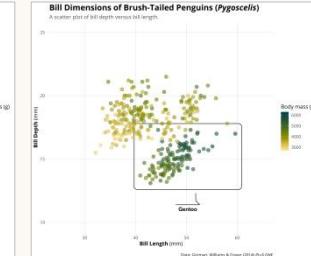


Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.



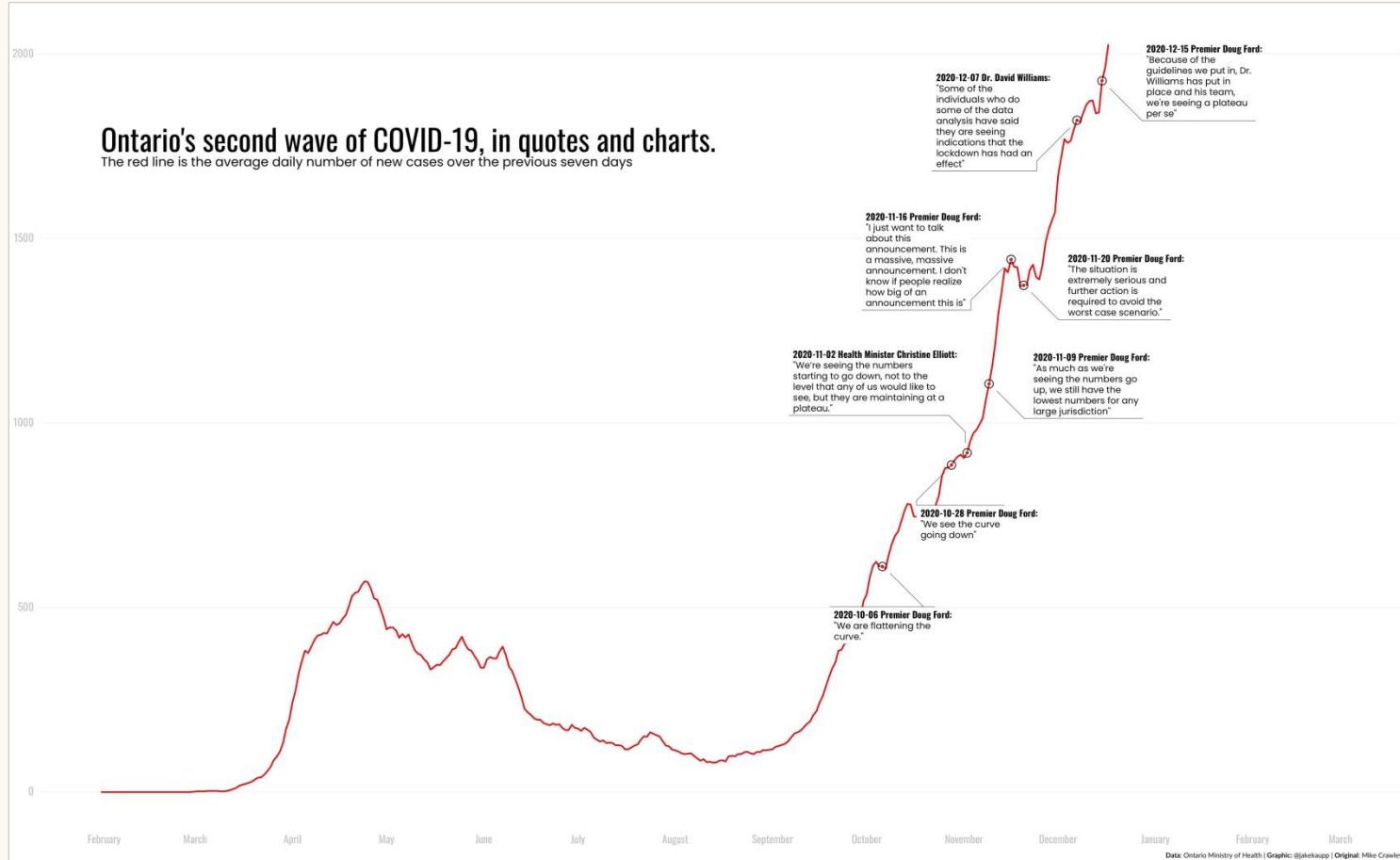
`geom_mark_circle()`



`geom_mark_rect()`

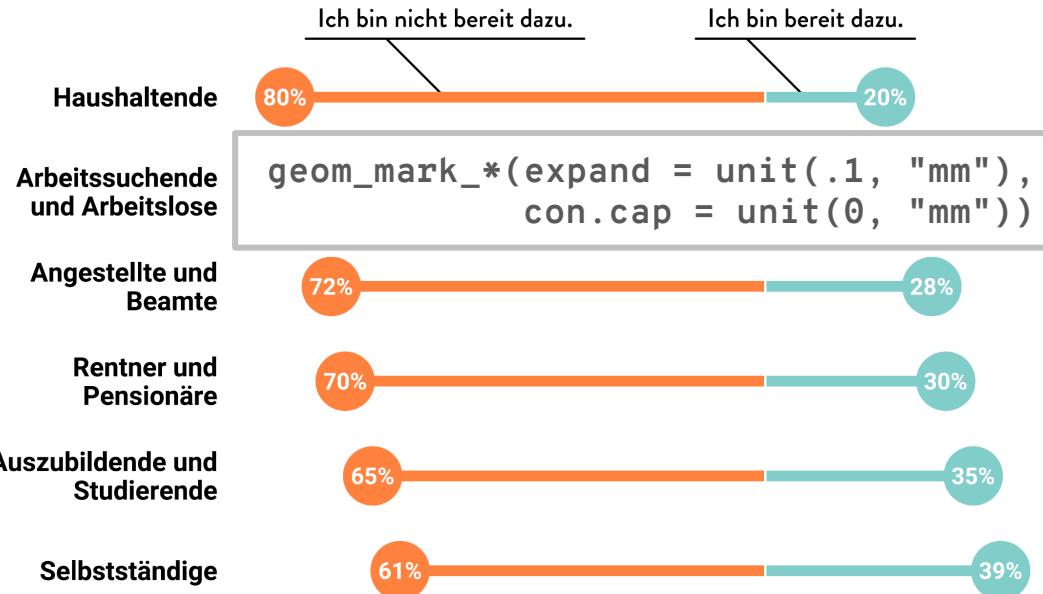
the points
ing parts

`geom_mark_ellipsoid()`



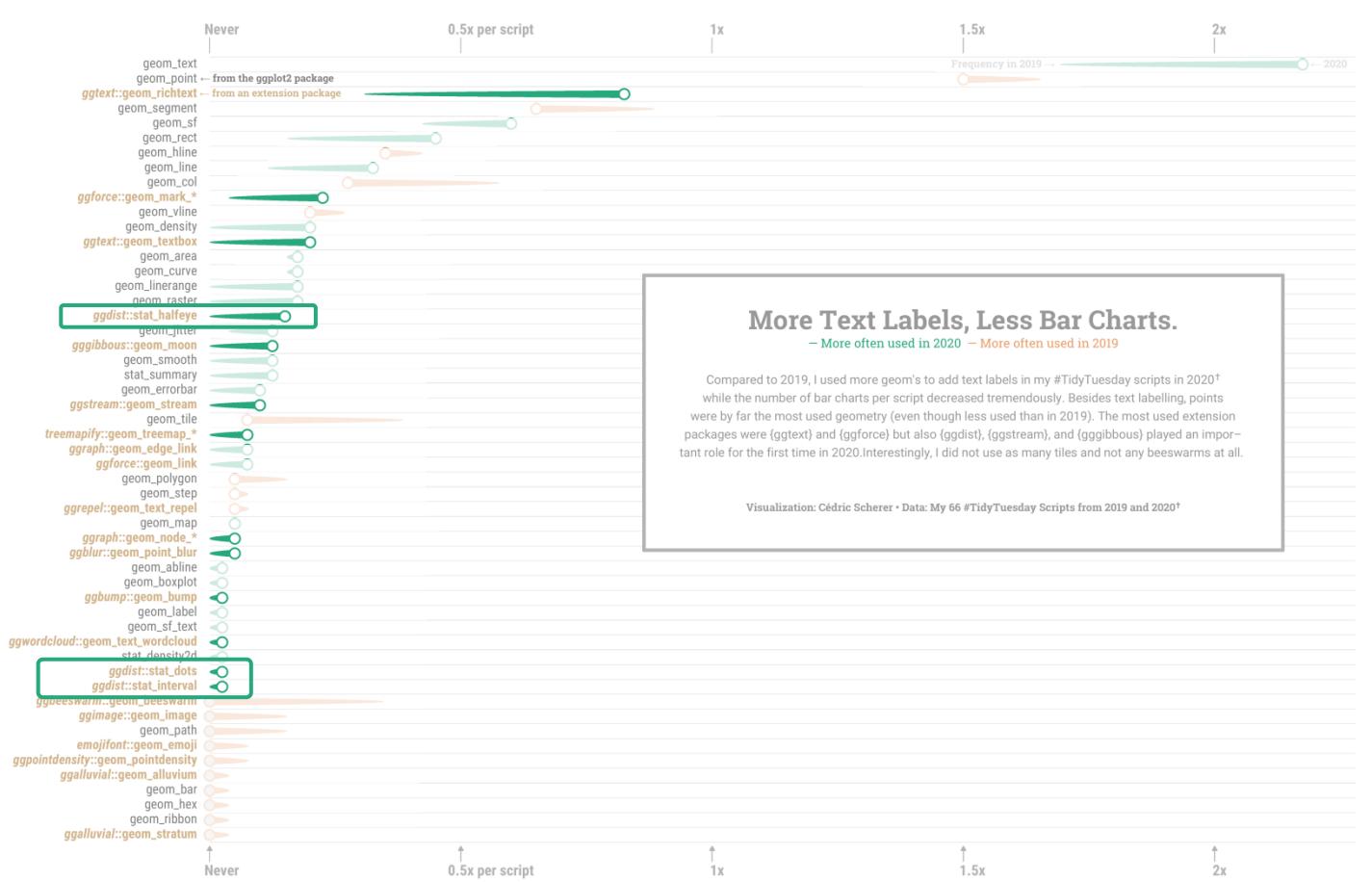
Jake Kaupp

Die Präsidentin der EU Ursula von der Leyen bittet "finanziell nicht notleidende Kunden" ihr Recht auf Rückerstattung aus Solidarität nicht in Anspruch zu nehmen.



Basierend auf 1057 Antworten auf eine Umfrage von KUENDIGUNG.ORG

Customer survey Kuendigung.org
(kuendigung.org/studien/verbraucherumfrage-zur-zukunft-nach-der-krise)



More Text Labels, Less Bar Charts.

— More often used in 2020 — More often used in 2019

Compared to 2019, I used more geom's to add text labels in my #TidyTuesday scripts in 2020[†] while the number of bar charts per script decreased tremendously. Besides text labelling, points were by far the most used geometry (even though less used than in 2019). The most used extension packages were (ggtext) and (ggforce) but also (ggdist), (ggstream), and (ggbibous) played an important role for the first time in 2020. Interestingly, I did not use as many tiles and not any beeswarms at all.

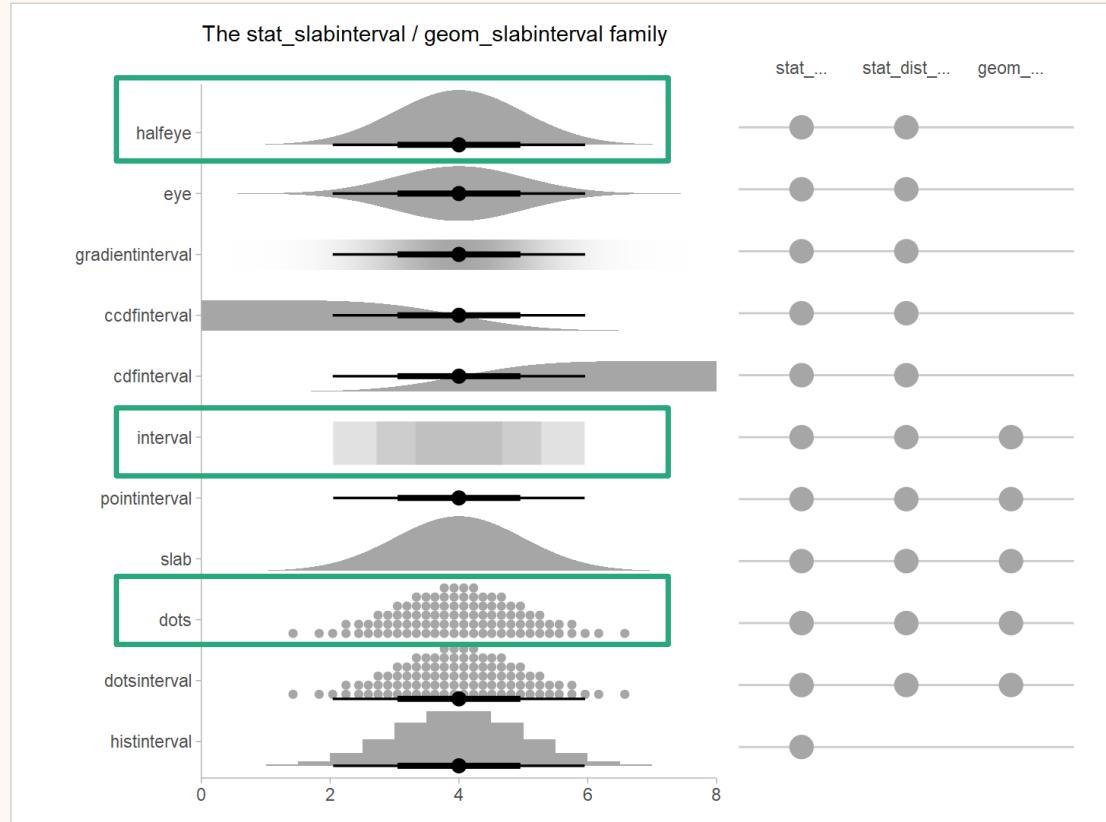
Visualization: Cédric Scherer • Data: My 66 #TidyTuesday Scripts from 2019 and 2020[†]

[†] I extracted all functions starting with `geom` or `stat` from my Rmd files containing the code for all my #TidyTuesday contributions (thanks Georgios for the idea and script). For the contributions from 2019 ($n = 26$) and 2020 ($n = 40$) I calculated the frequency of usage per year for each geom/stat as times used divided by the number of contributions. Note that some geom's which usually appear together (e.g. `treemapify::geom_treemap` functions) or behave very similarly (e.g. `ggforce::geommark` functions) were grouped together.

{ggdist}

Visualizations of Distributions and Uncertainty

{ggdist} Visualizations of Distributions and Uncertainty



Not my cup of coffee...

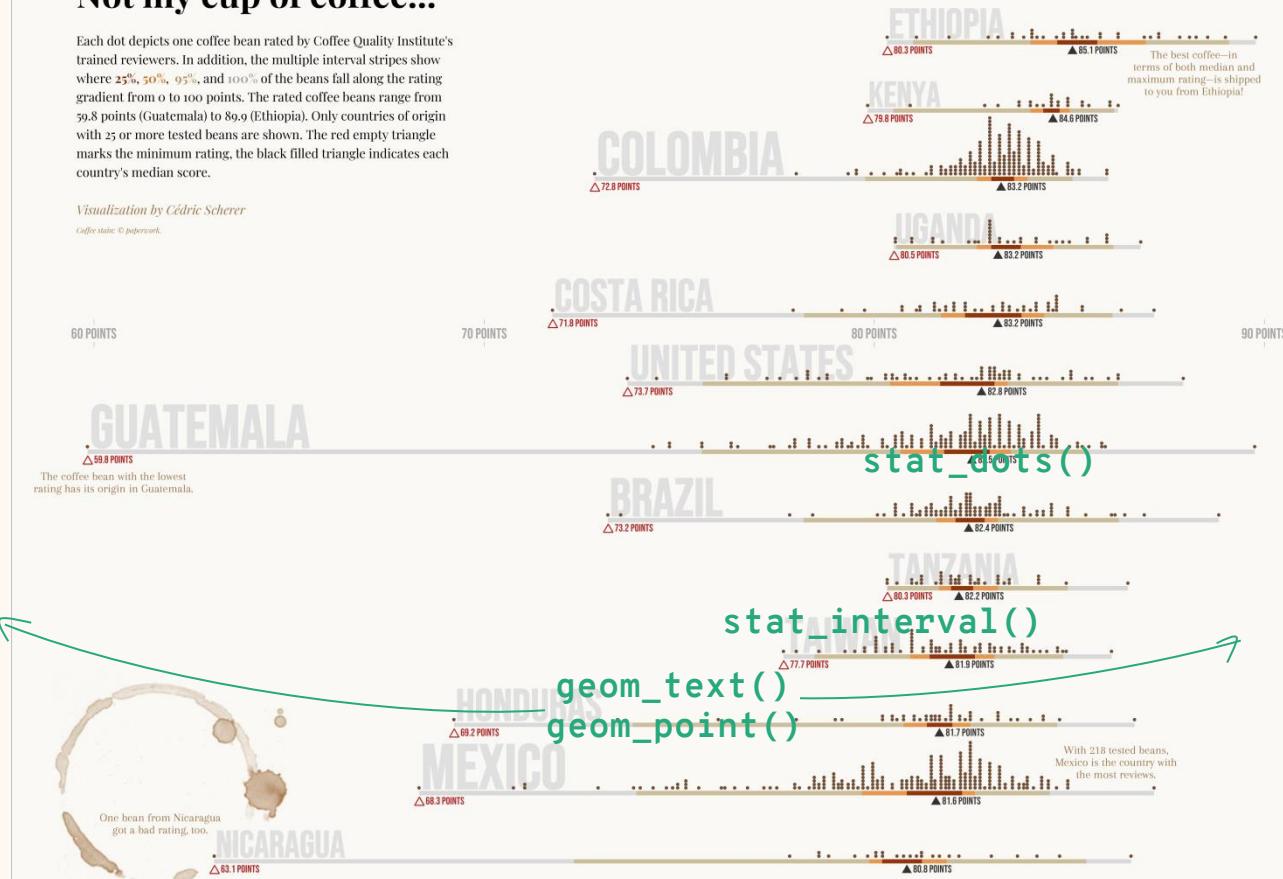
Each dot depicts one coffee bean rated by Coffee Quality Institute's trained reviewers. In addition, the multiple interval stripes show where 25%, 50%, 95%, and 100% of the beans fall along the rating gradient from 0 to 100 points. The rated coffee beans range from 59.8 points (Guatemala) to 89.9 (Ethiopia). Only countries of origin with 25 or more tested beans are shown. The red empty triangle marks the minimum rating, the black filled triangle indicates each country's median score.

Visualization by Cédric Scherer

Coffee stain: © paperwerk.



△ 59.8 POINTS
The coffee bean with the lowest rating has its origin in Guatemala.



Contribution to #TidyTuesday 2020/28

COLOMBIA

△ 72.8 POINTS

KENYA

△ 79.8 POINTS

stat_dots()

stat_interval()

▲ 84.6 POINTS

geom_text()
geom_point()

△ 80.5 POINTS

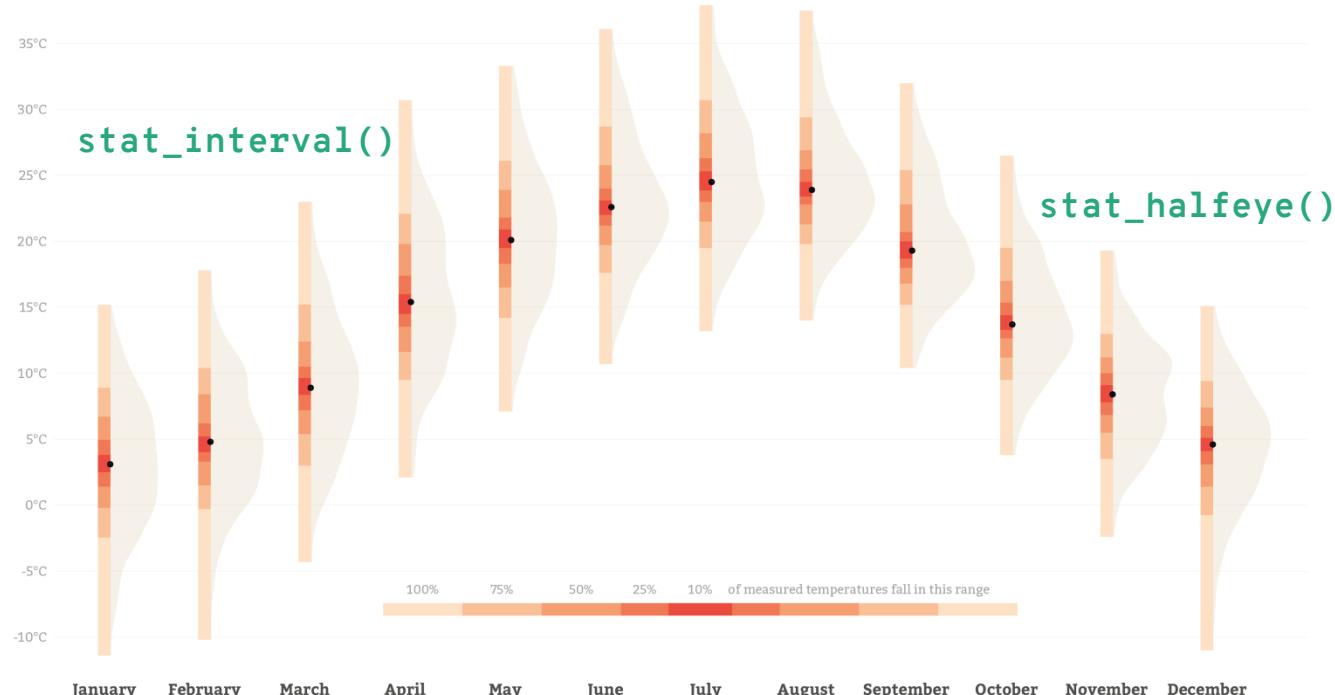
UGANDA

▲ 83.2 POINTS

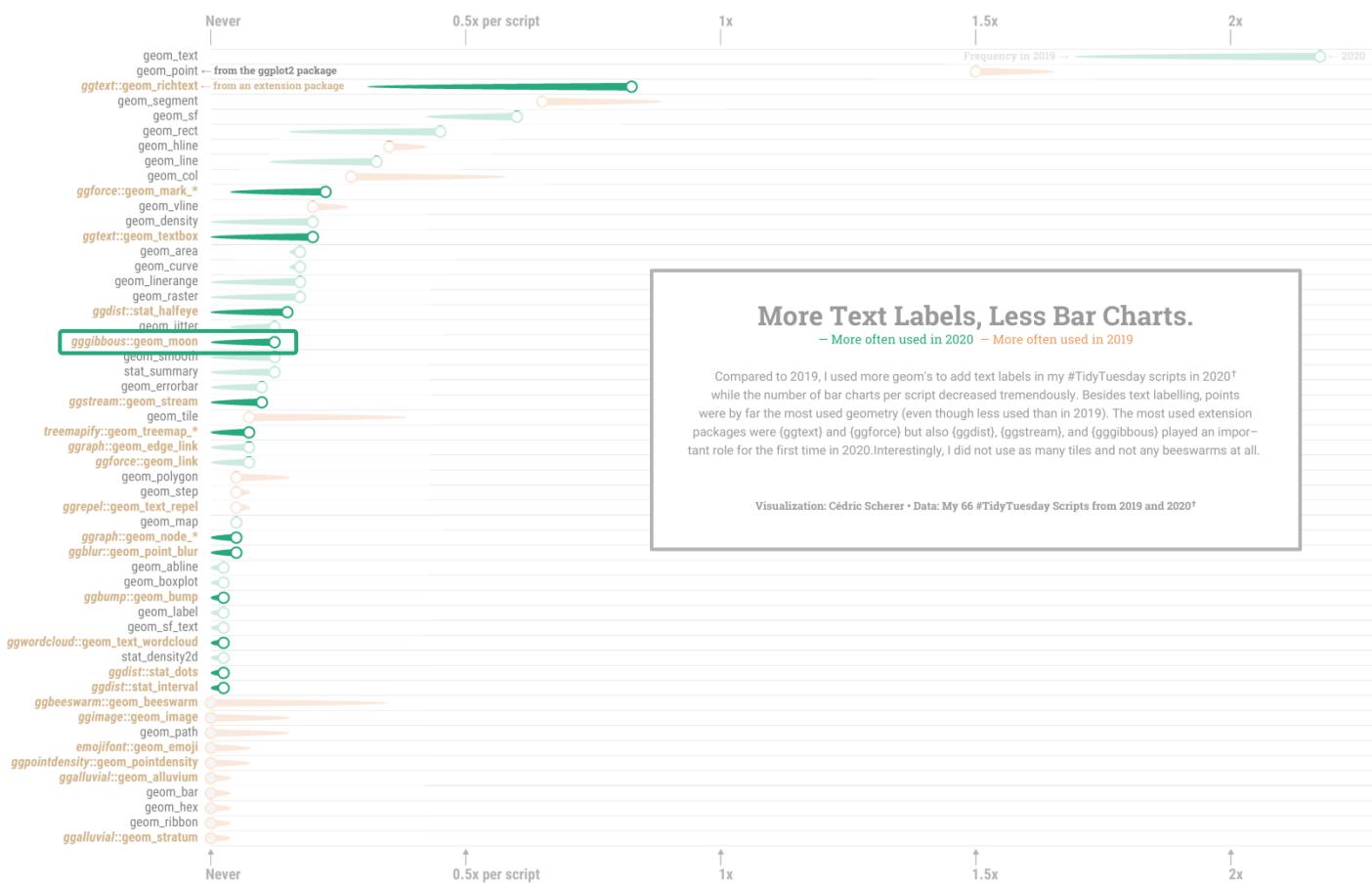
Contribution to #TidyTuesday 2020/28

Daily Temperatures in Berlin, Germany

Range and distribution of maximum daily temperatures in Celsius per month from 2000 to 2018 measured in Berlin-Dahlem, Germany



Contribution to the SWD Challenge September 2019



More Text Labels, Less Bar Charts.

— More often used in 2020 — More often used in 2019

Compared to 2019, I used more geom's to add text labels in my #TidyTuesday scripts in 2020[†] while the number of bar charts per script decreased tremendously. Besides text labelling, points were by far the most used geometry (even though less used than in 2019). The most used extension packages were (`ggtext`) and (`ggforce`) but also (`ggdist`), (`ggstream`), and (`gggibbous`) played an important role for the first time in 2020. Interestingly, I did not use as many tiles and not any beeswarms at all.

Visualization: Cédric Scherer • Data: My 66 #TidyTuesday Scripts from 2019 and 2020[†]

[†] I extracted all functions starting with `geom` or `stat` from my Rmd files containing the code for all my #TidyTuesday contributions (thanks Georgios for the idea and script). For the contributions from 2019 (n = 26) and 2020 (n = 40) I calculated the frequency of usage per year for each geomist as times used divided by the number of contributions. Note that some geom's which usually appear together (e.g. `treemapify::geom_treemap` functions) or behave very similarly (e.g. `ggforce::geommark` functions) were grouped together.

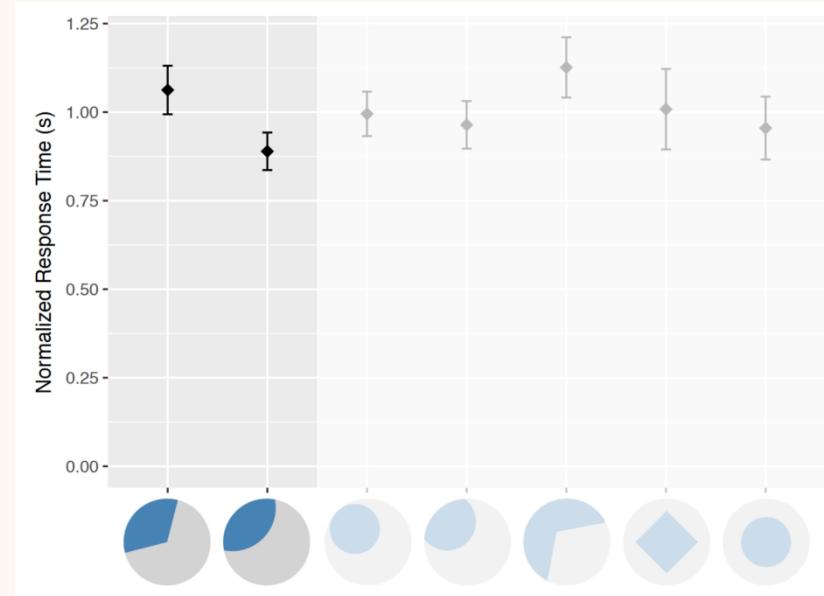
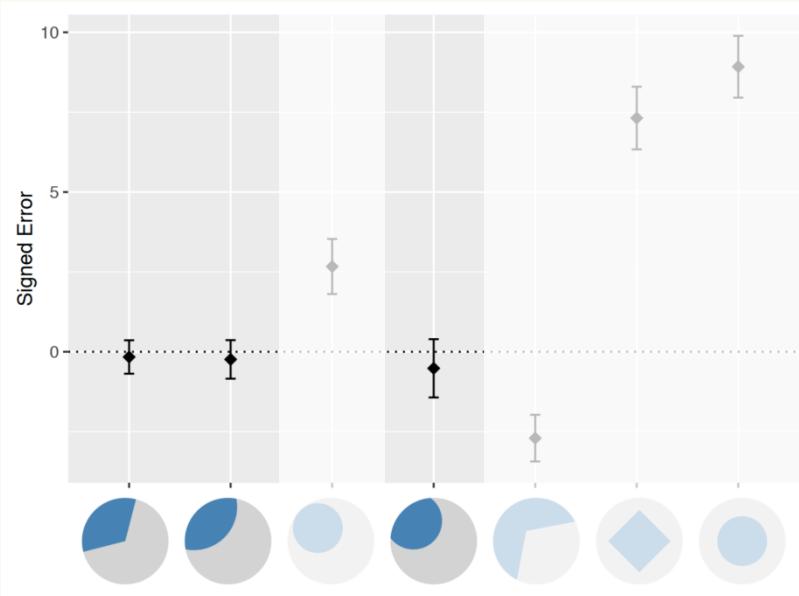
{gggibbous}

Moon charts for ggplot2



Circular Part-to-Whole Charts Using the Area Visual Cue

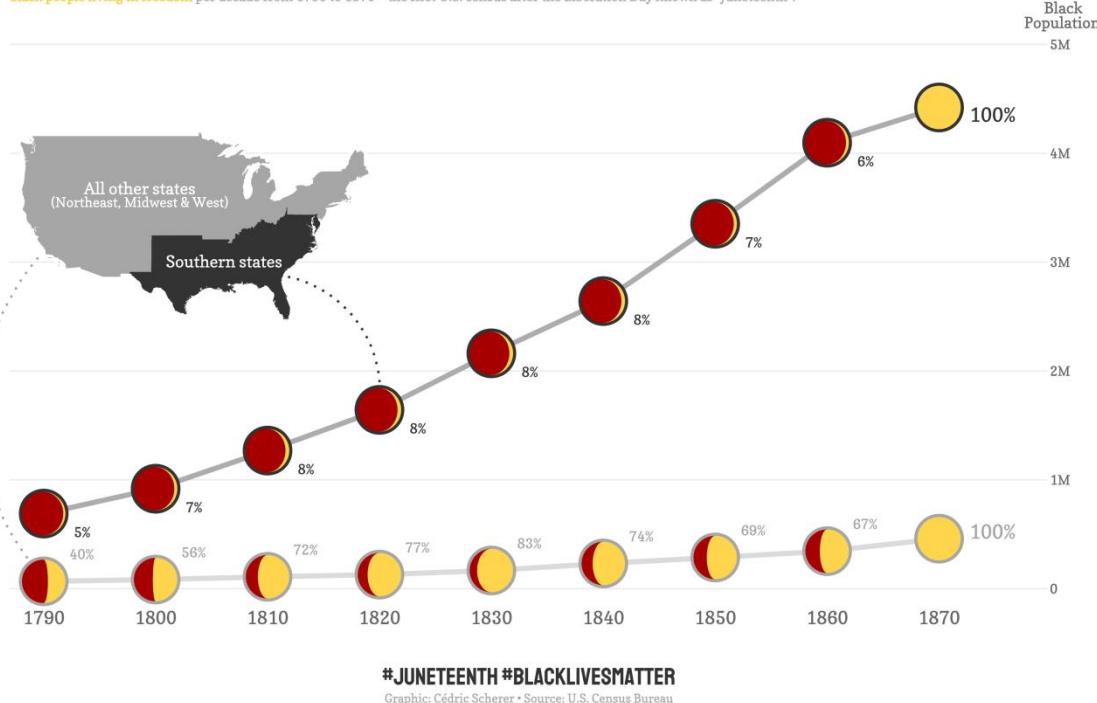
Robert Kosara
Tableau Research



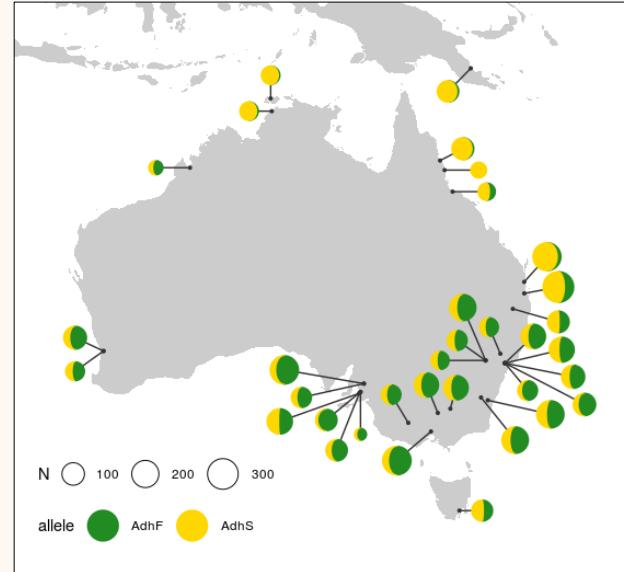
doi.org/10.2312/evs.20191163

{gggibbous} Moon charts

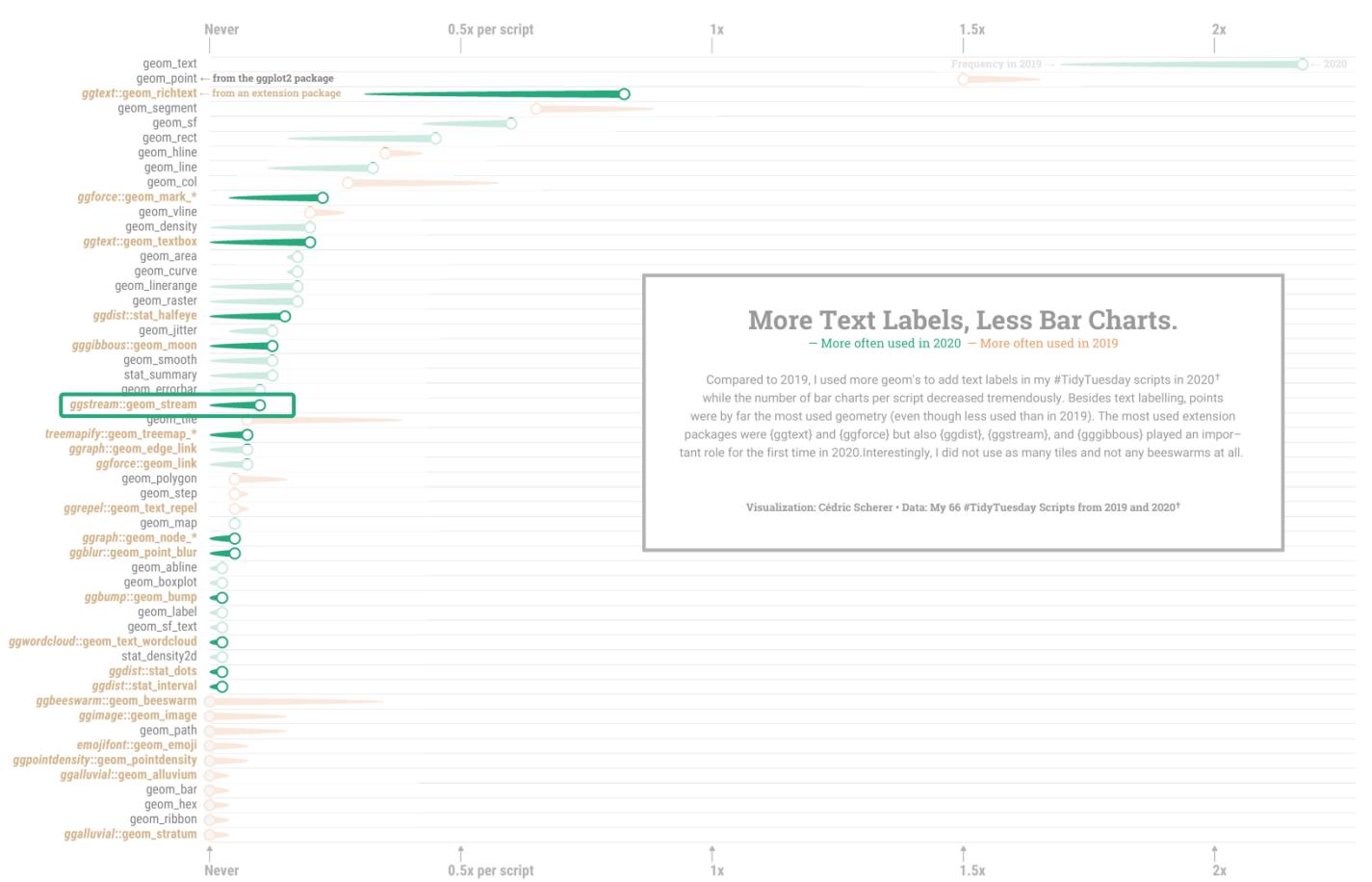
By 1680, property owners in the south of North America began establishing plantation farms for cash crops like tobacco, cotton, and sugar cane—enterprises that required increasing amounts of labor. To meet the need, wealthy planters became slave traders and imported ever more individuals to the colonies, the vast majority from West Africa. While the “Emancipation Proclamation” was made law as of 1863, slave owners in the South, namely Texas, still maintained slavery until June 19th 1865 when Union soldiers were able to enforce the law abolishing slavery in the region. The graphic below shows the share of **black people in slavery**, most of them enslaved in the Southern states, and **black people living in freedom** per decade from 1790 to 1870—the first U.S. census after the Liberation Day known as “Juneteenth”.



Contribution to #TidyTuesday 2020/25



Package example
(github.com/mnbrahm/gggibbous)



More Text Labels, Less Bar Charts.

— More often used in 2020 — More often used in 2019

Compared to 2019, I used more geom's to add text labels in my #TidyTuesday scripts in 2020[†] while the number of bar charts per script decreased tremendously. Besides text labelling, points were by far the most used geometry (even though less used than in 2019). The most used extension packages were (ggtext) and (ggforce) but also (ggdist), (ggsstream), and (gggibbous) played an important role for the first time in 2020. Interestingly, I did not use as many tiles and not any beeswarms at all.

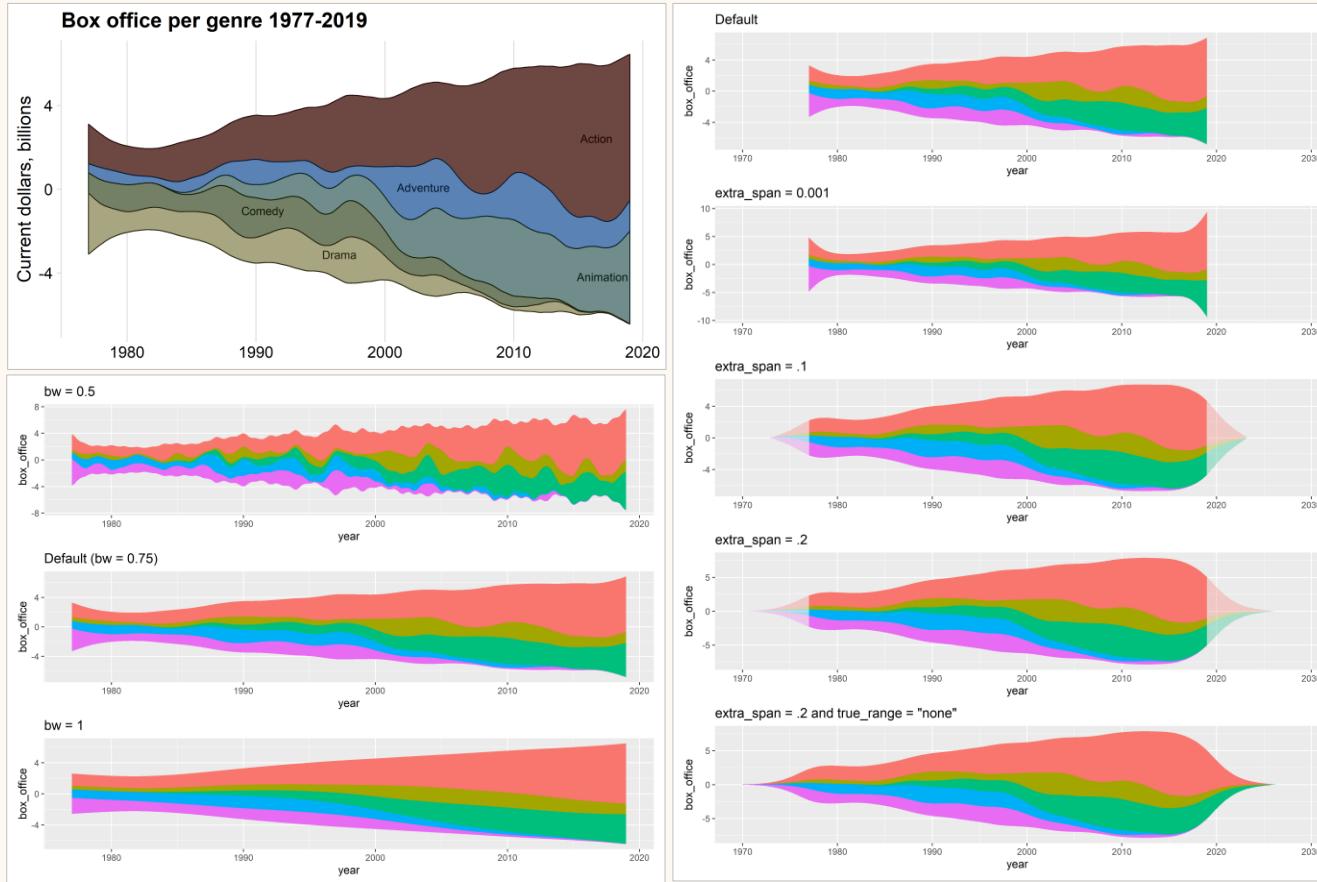
Visualization: Cédric Scherer • Data: My 66 #TidyTuesday Scripts from 2019 and 2020[†]

[†] I extracted all functions starting with `geom` or `stat` from my Rmd files containing the code for all my #TidyTuesday contributions (thanks Georgios for the idea and script). For the contributions from 2019 ($n = 26$) and 2020 ($n = 40$) I calculated the frequency of usage per year for each geom/stat as times used divided by the number of contributions. Note that some geom's which usually appear together (e.g. `treemapify::geomtreemap` functions) or behave very similarly (e.g. `ggforce::geommark` functions) were grouped together.

{ggstream}

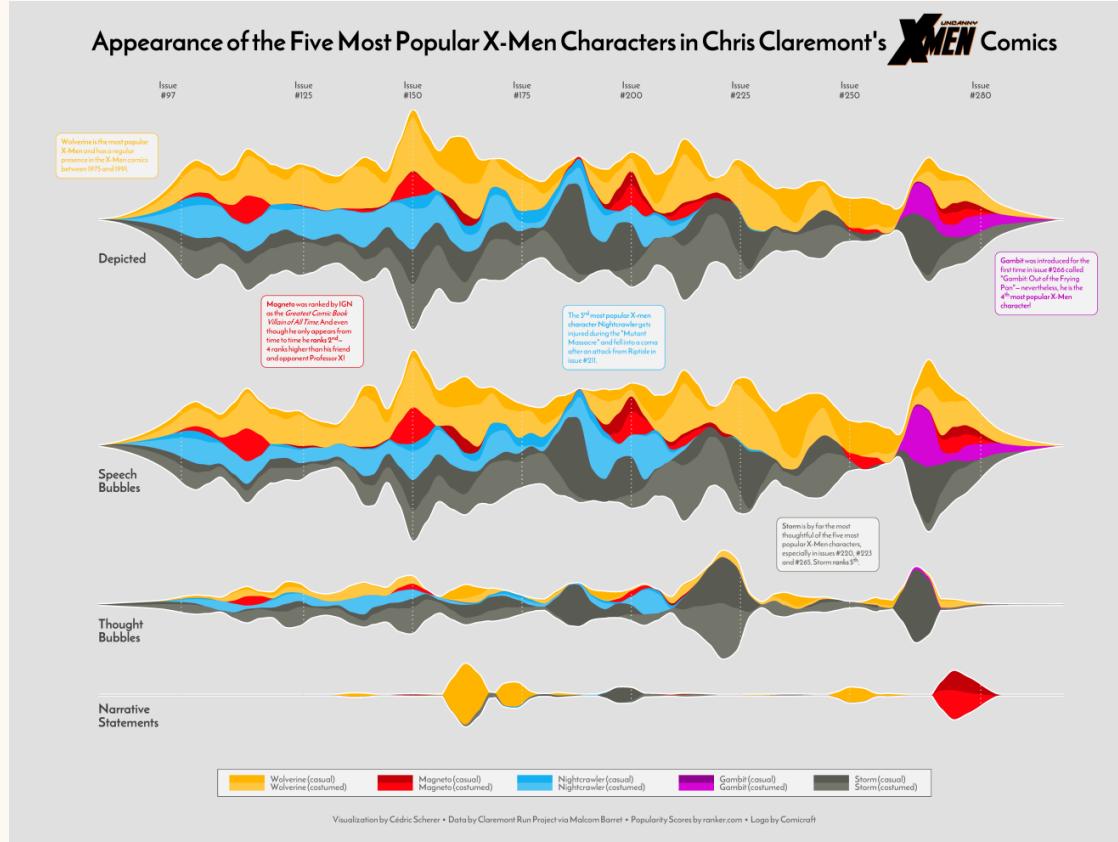
Make streamplots with ggplot2

{ggstream} Streamgraphs



Package examples
(github.com/davidsjoberg/ggstream)

{ggstream} Streamgraphs



My Contribution to #TidyTuesday 2020/27

{ggplot2}

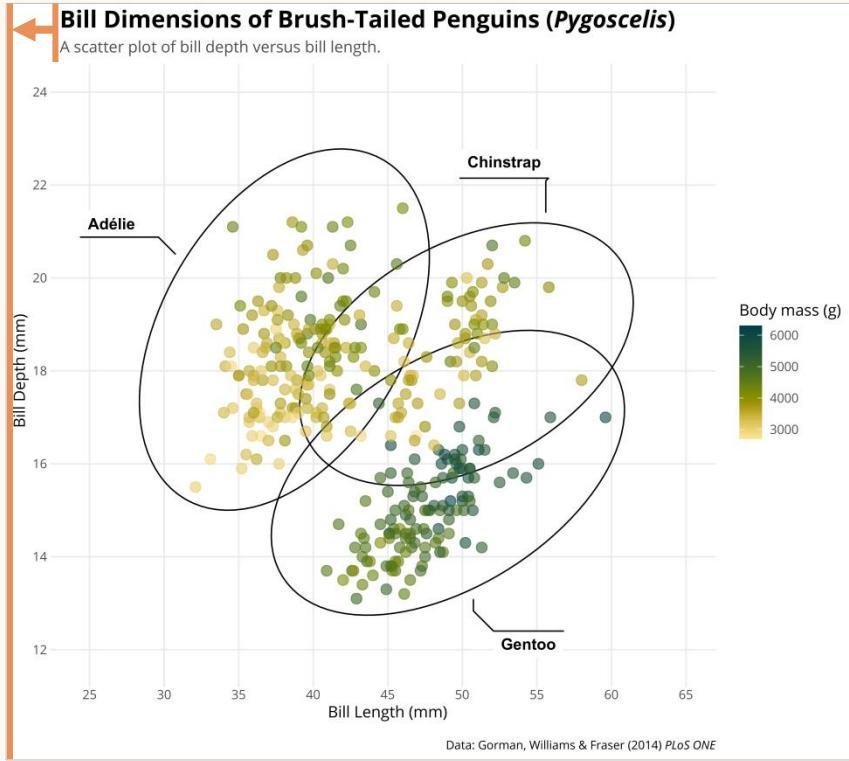
Details You May Not Know Yet
(or have never thought about)



ggplot2.tidyverse.org

Left-Aligned Title?

```
theme(plot.title = element_text(hjust = 0))
```



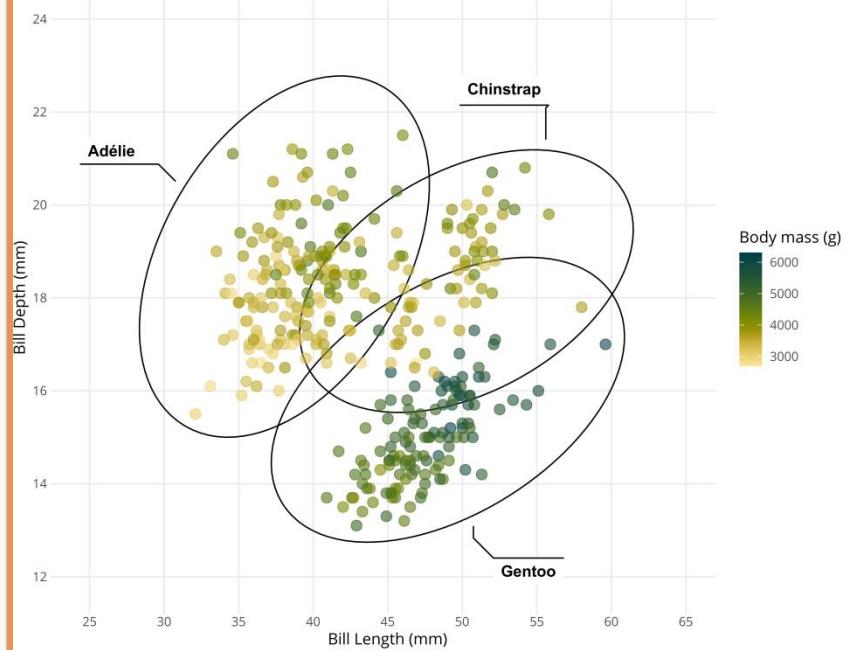
Left-Aligned Title

```
theme(plot.title.position = 'panel')
```

Default

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.

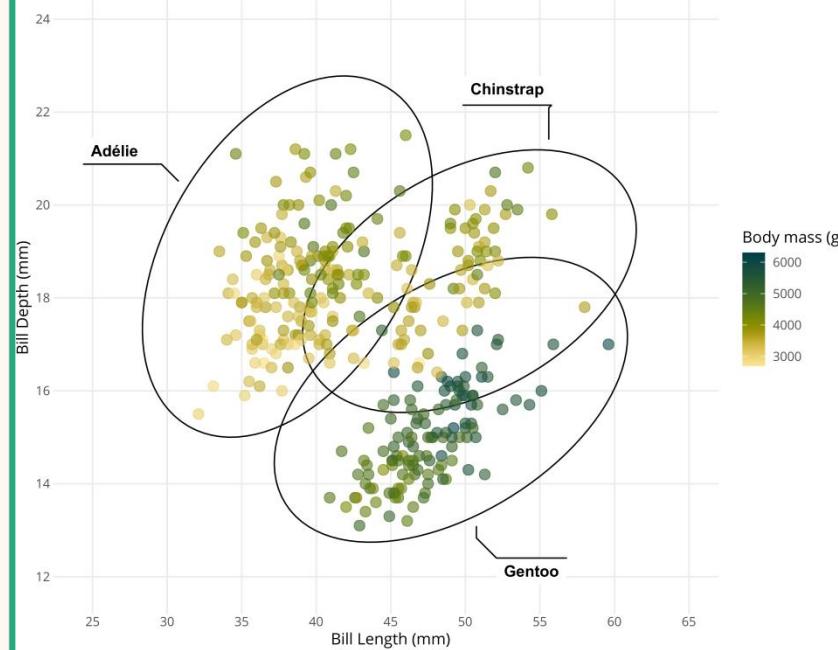


```
theme(plot.title.position = 'plot')
```

Better

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.



Data: Gorman, Williams & Fraser (2014) *PLoS ONE*

Data: Gorman, Williams & Fraser (2014) *PLoS ONE*

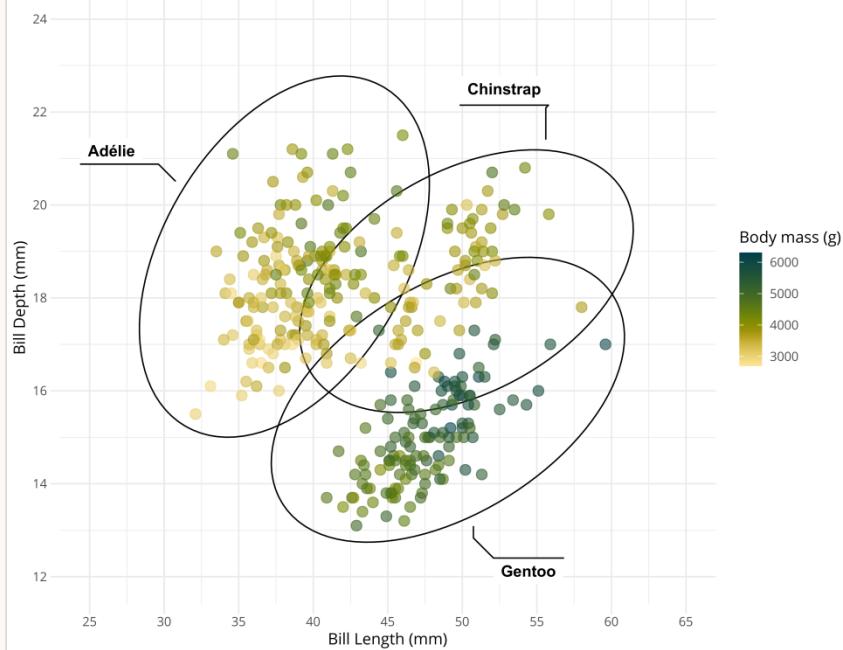
Right-Aligned Caption

```
theme(plot.caption.position = 'panel')
```

Default

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.



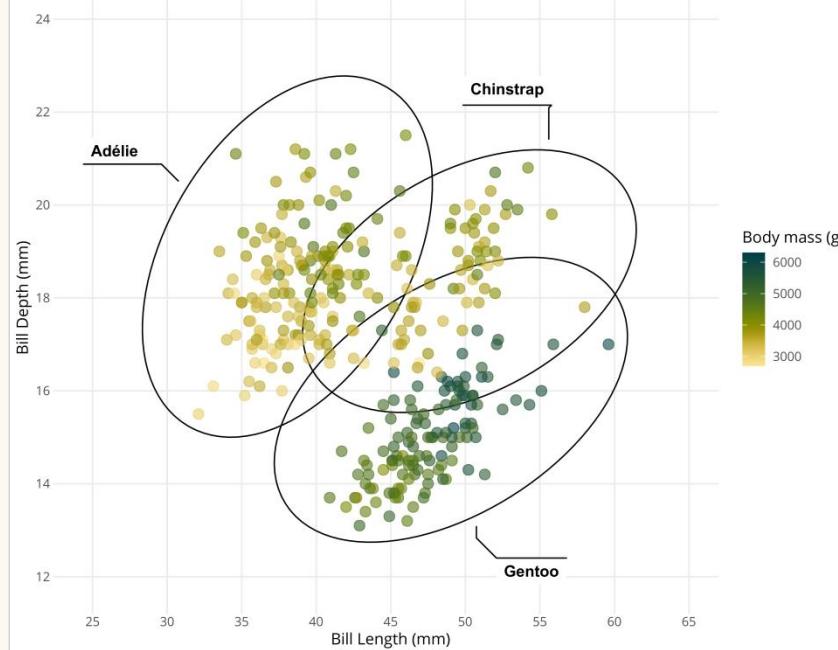
Data: Gorman, Williams & Fraser (2014) *PLoS ONE*

Better (?)

```
theme(plot.caption.position = 'plot')
```

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.



Data: Gorman, Williams & Fraser (2014) *PLoS ONE*

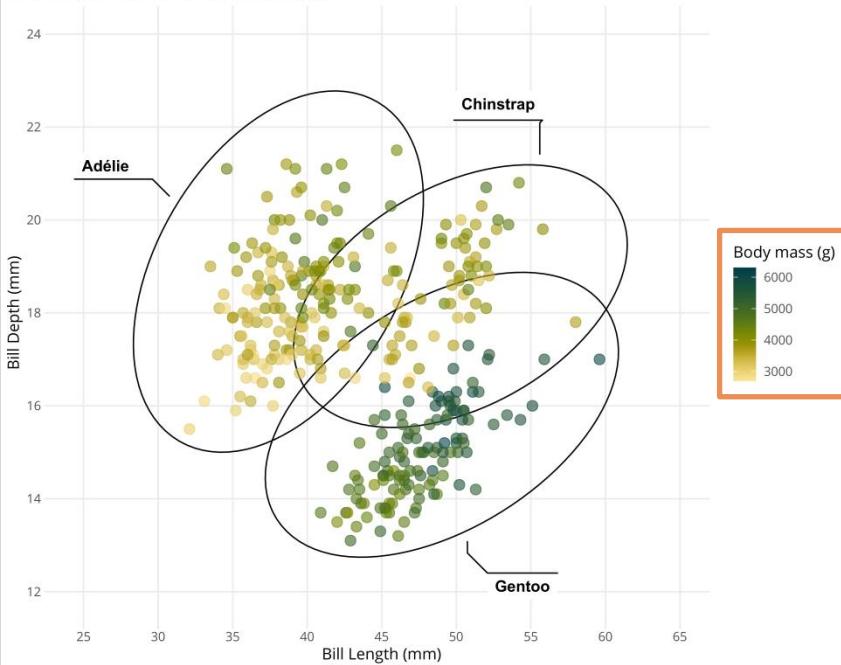
Legend Design

theme(legend.position = 'right')

Default

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.

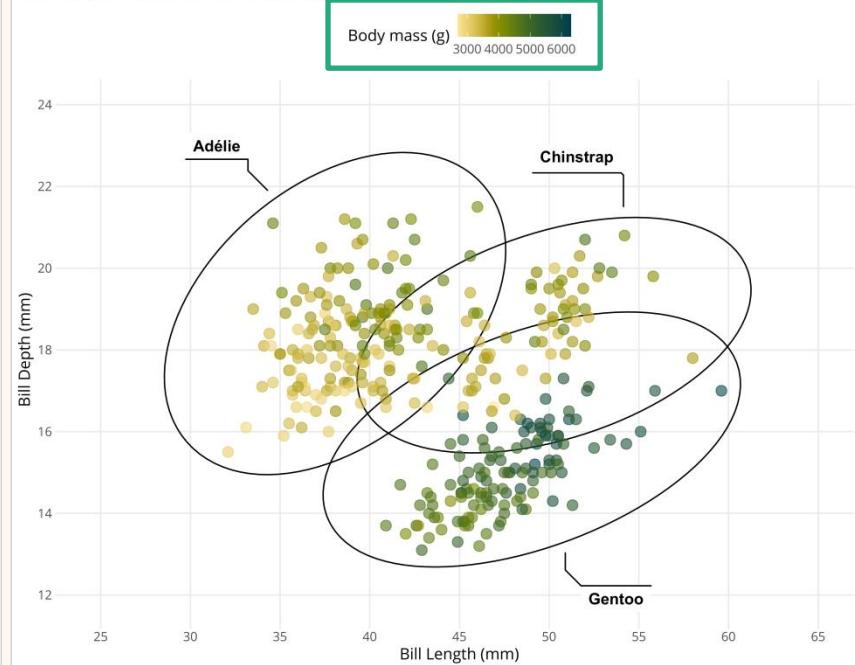


Well...

theme(legend.position = 'top')

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.



Legend Design

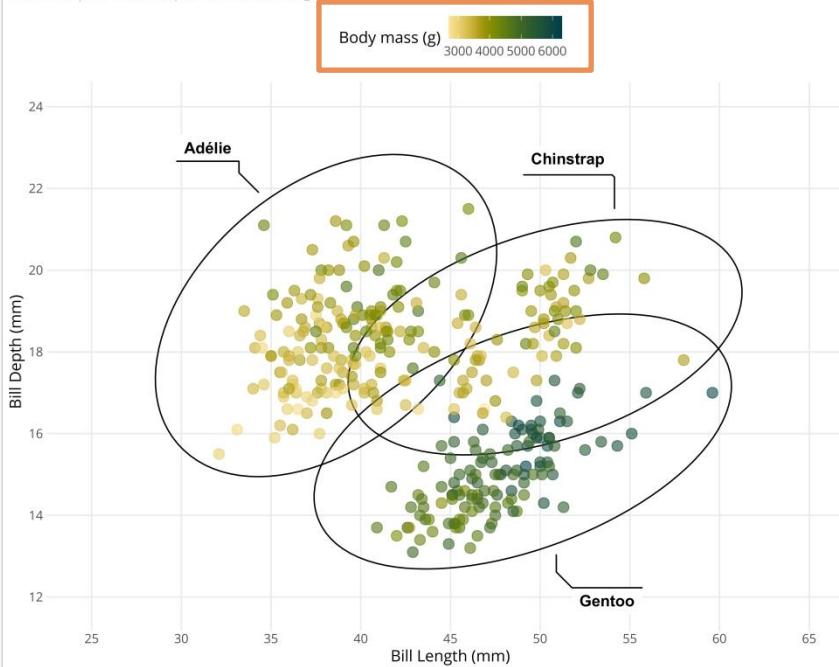
Default

```
guides(color = guide_colorbar(title.position = 'top', title.hjust = .5,
                                barwidth = unit(20, 'lines'), barheight = unit(.5, 'lines')))
```

Better!

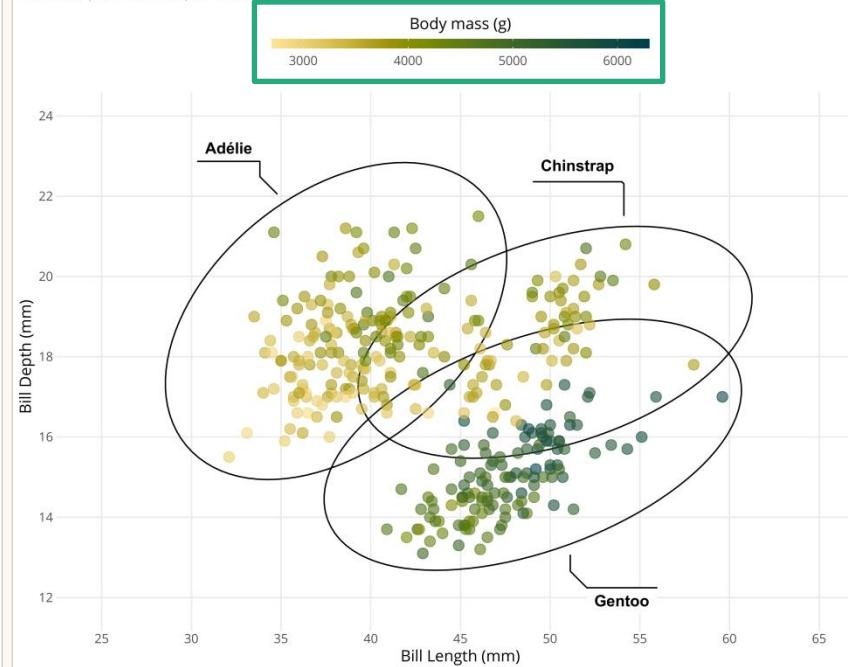
Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.



Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.



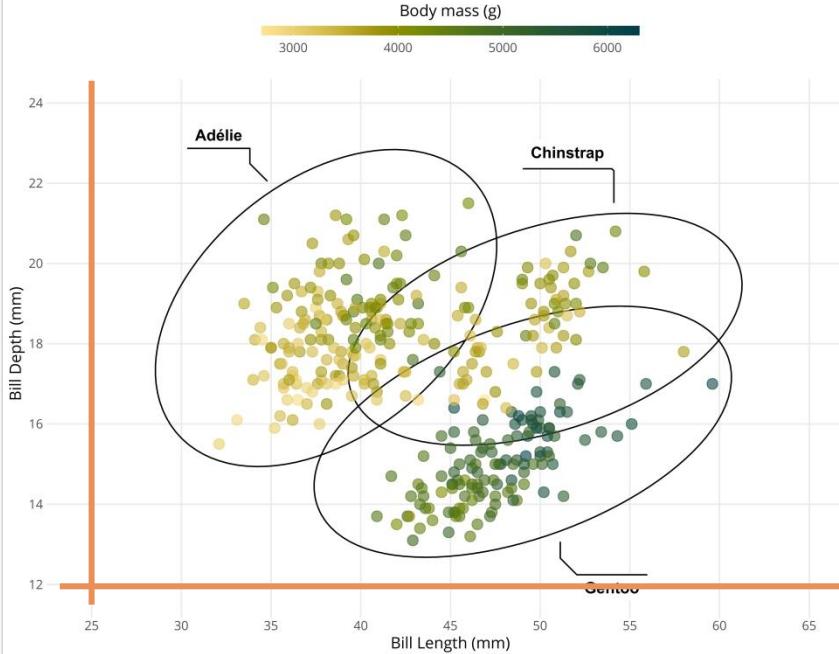
Limit Expansion

`coord_cartesian(expand = TRUE)`

Default

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.

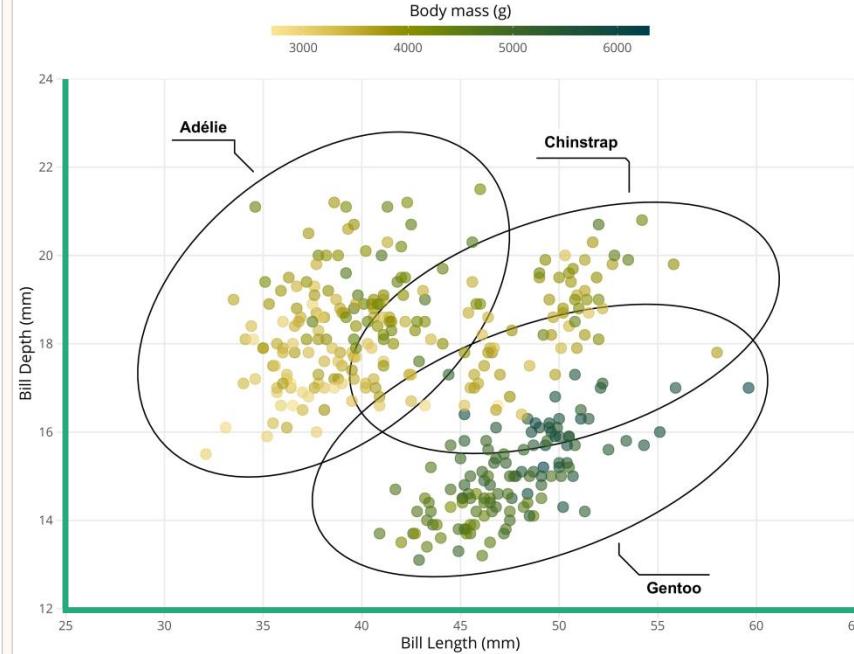


`coord_cartesian(expand = FALSE)`

Without

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.



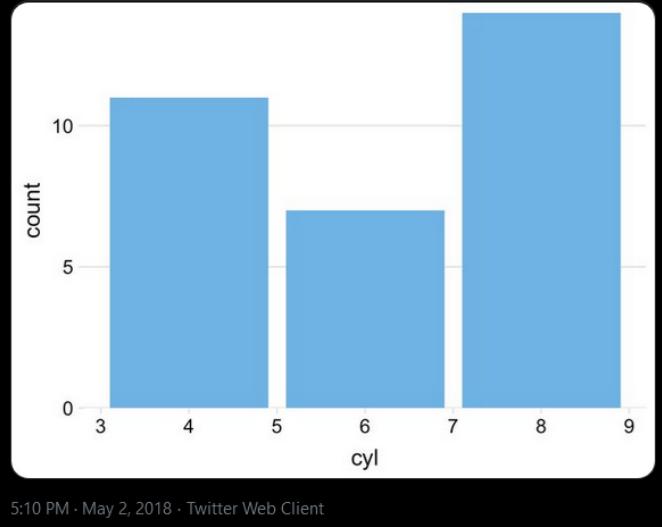
Geeky Details



Claus Wilke
@ClausWilke

Replies to @thomas85 and @hrbrmstr

This kind of stuff just really gets me. One of the main reasons I was motivated to add clip = "off" is plots like the attached, which look terrible in my opinion.



5:10 PM · May 2, 2018 · Twitter Web Client



gvdr @ipnosimmia · May 2, 2018

Replies to @ClausWilke @thomas85 and @hrbrmstr

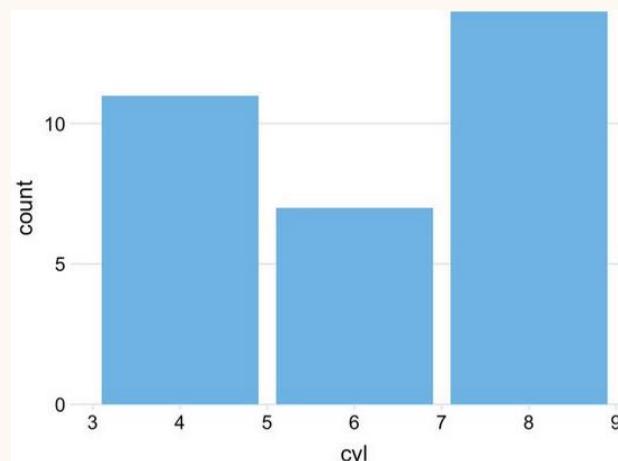
I must admit my limits. I can't see any difference between incorrect and correct. May I ask you to spell it out? (here to learn!)



1



2



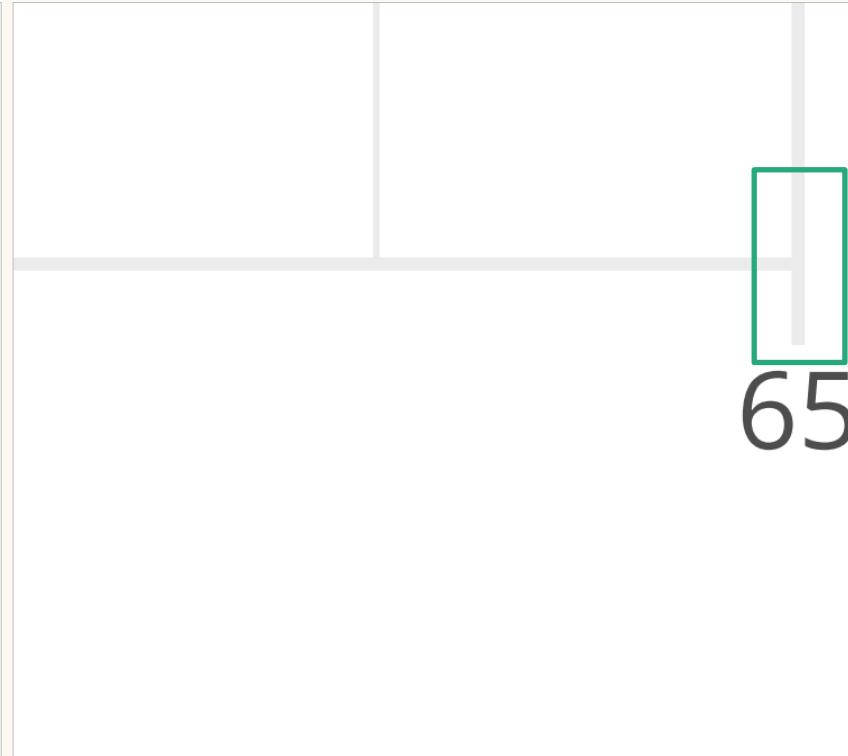
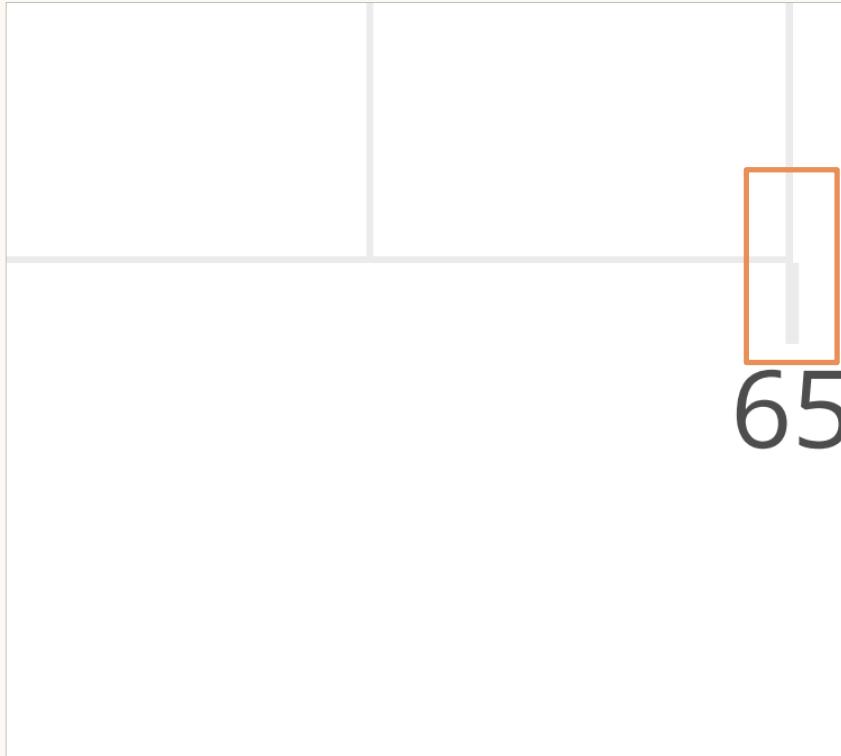
Geeky Details

`coord_cartesian(clip = 'on')`

Default

Without

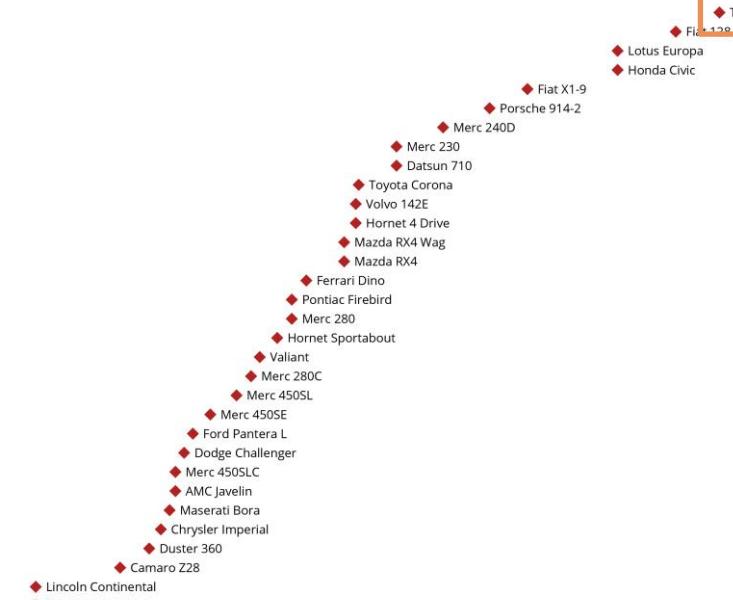
`coord_cartesian(clip = 'off')`



Geeky Details

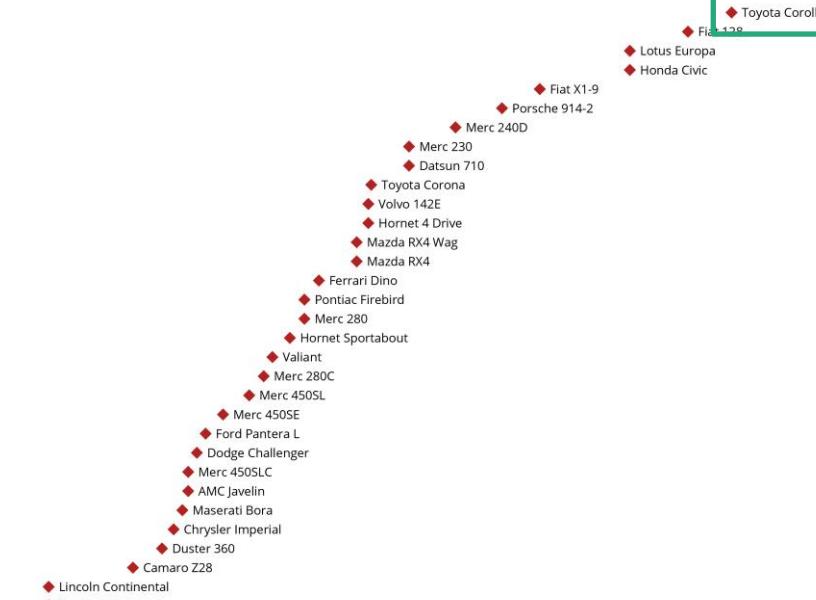
```
coord_cartesian(clip = 'on')
```

Default



```
coord_cartesian(clip = 'off')
```

Without



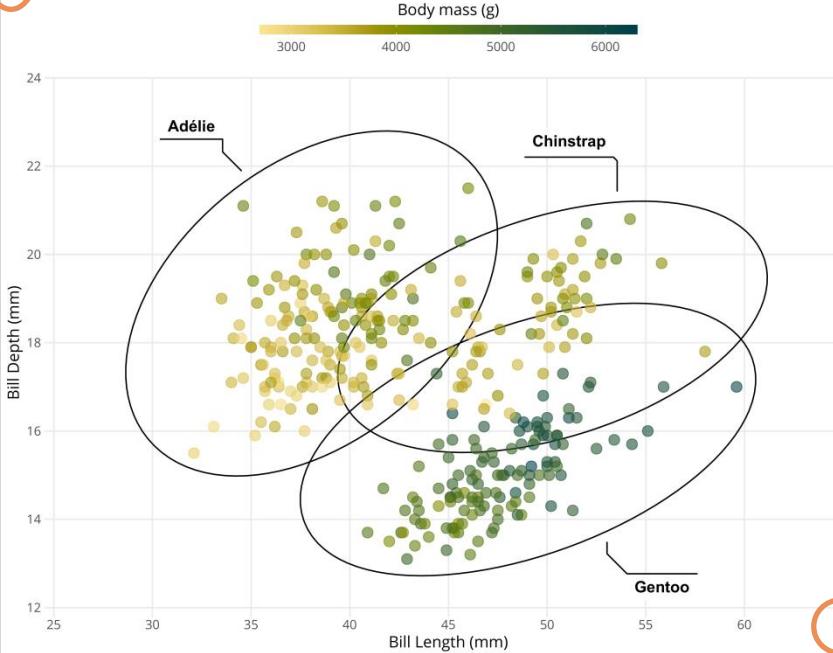
White Space

```
theme(plot.margin = margin(rep(base_size/2, 4)))
```

Default

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.

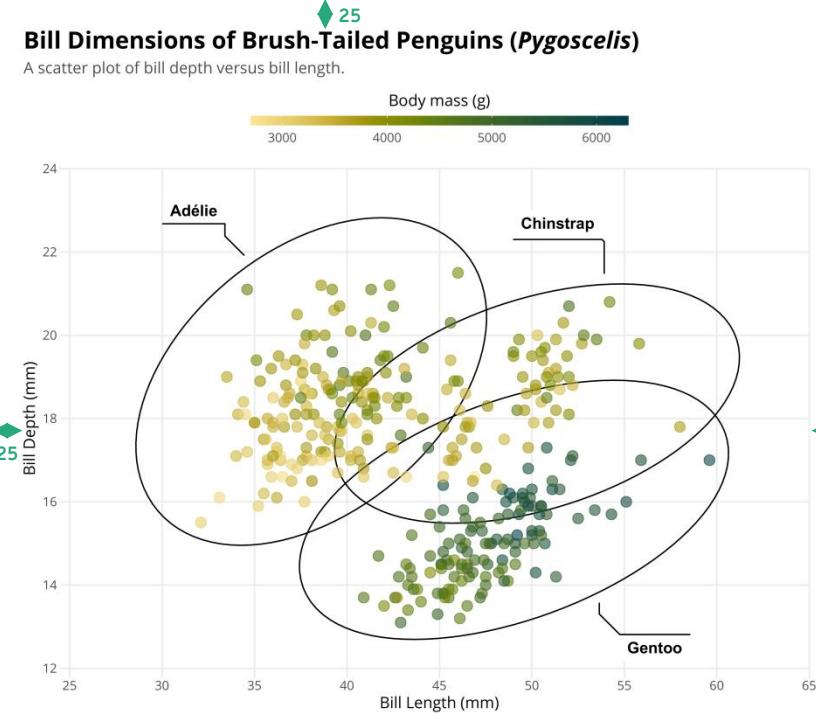


```
theme(plot.margin = margin(25, 25, 10, 25))
```

Better

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.

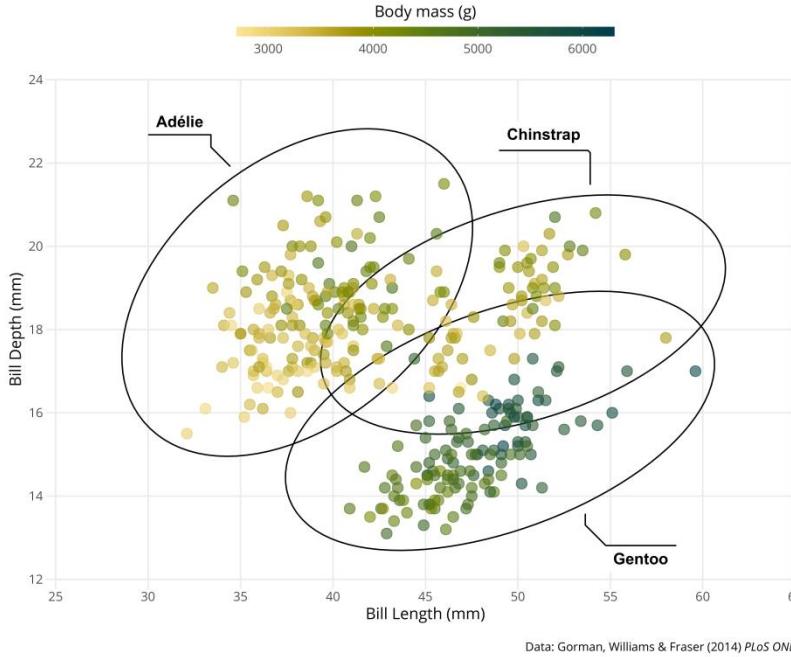


Adding Images

```
png <- magick::image_read("https://raw.githubusercontent.com/allisonhorst/.../culmen_depth.png")
img <- grid::rasterGrob(png, interpolate = TRUE)
annotation_custom(img, ymin = 22, ymax = 31, xmin = 55, xmax = 65.5)
```

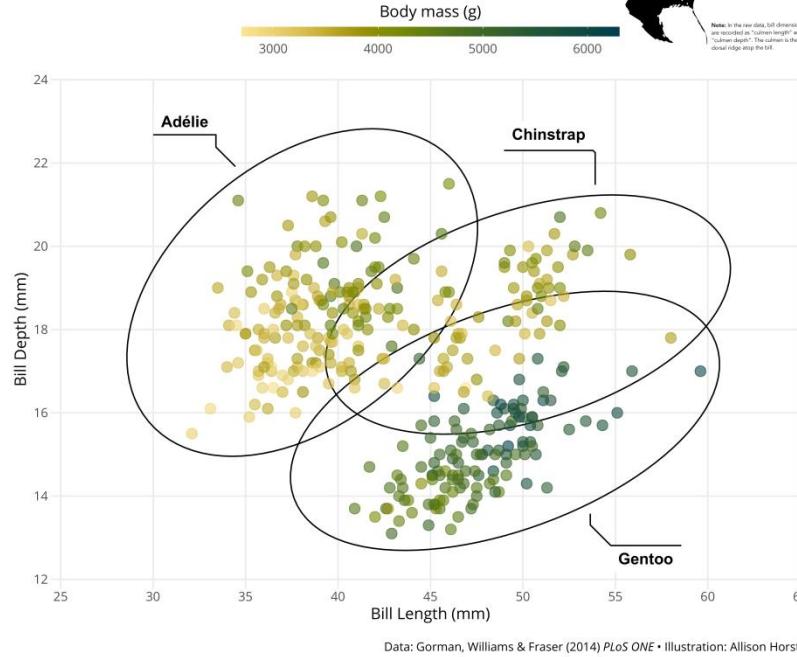
Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.



Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A scatter plot of bill depth versus bill length.



{patchwork}

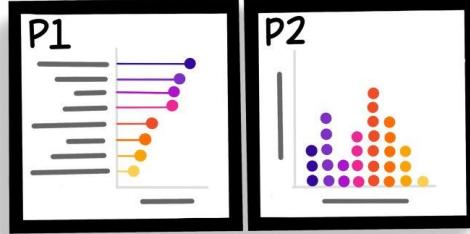
The Composer of ggplots



patchwork.data-imaginist.com

patchwork

Combine + arrange
your ggplots!

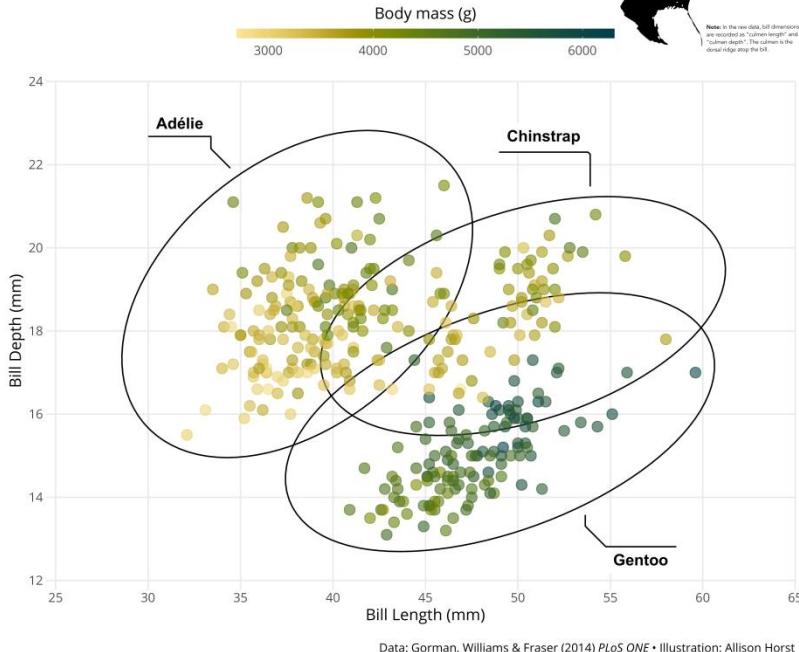


Artwork by Allison Horst

{patchwork} The Composer of ggplots

Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

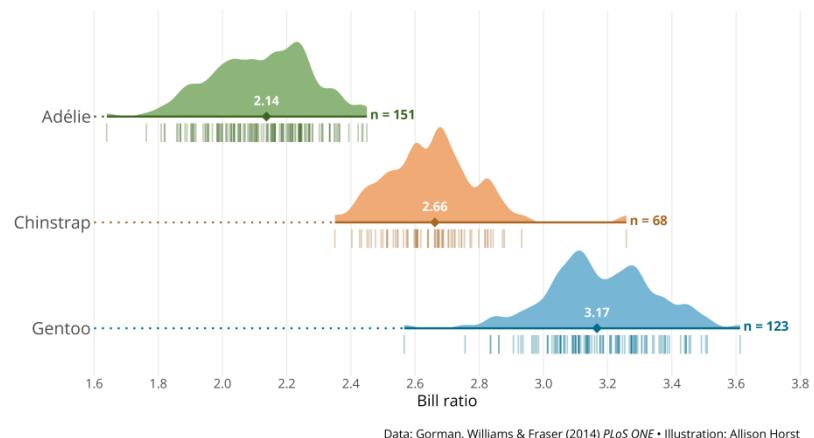
A scatter plot of bill depth versus bill length.



p1

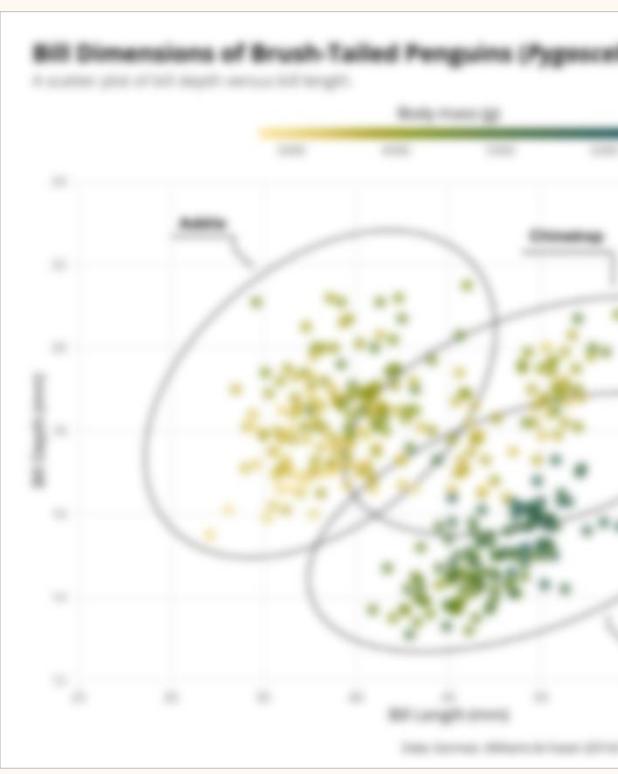
```
(p1 / p2) + plot_layout(heights = c(1, .65))
```

B. Raincloud plot showing the distribution of bill ratios, estimated as bill length divided by bill depth.

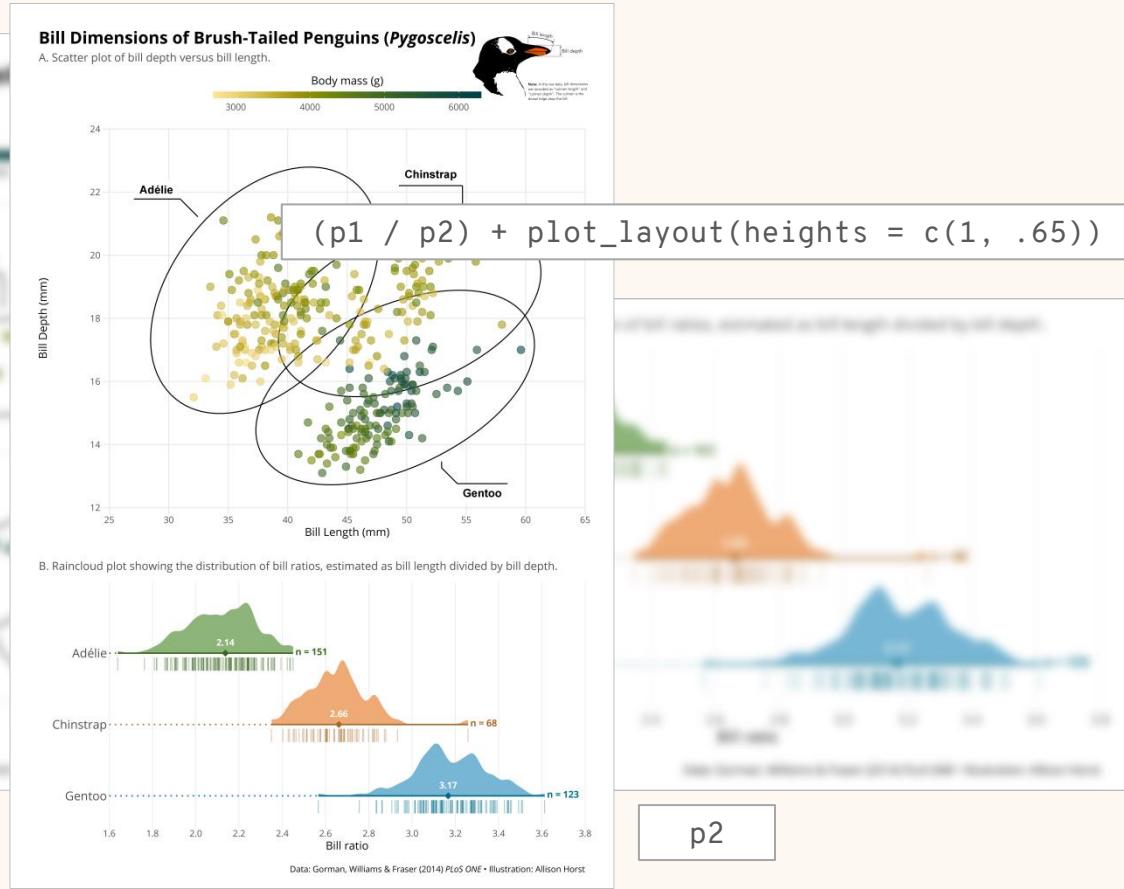


p2

{patchwork} The Composer of ggplots



p1



p2

```
ggtitle('*Pygoscelis*') + theme(plot.title = element_markdown())
```

```
theme(plot.position = 'plot')
```

```
theme(legend.position = 'top') +  
  guide(color = guide_colorbar())
```

```
theme(plot.margin = margin(t, r, b, 1))
```

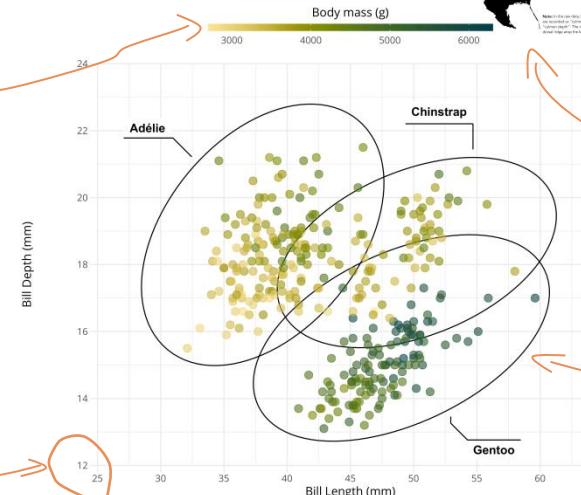
```
coord_cartesian(expand = c(0, 0),  
  clip = 'off')
```

```
ggdist::stat_halfeye()
```

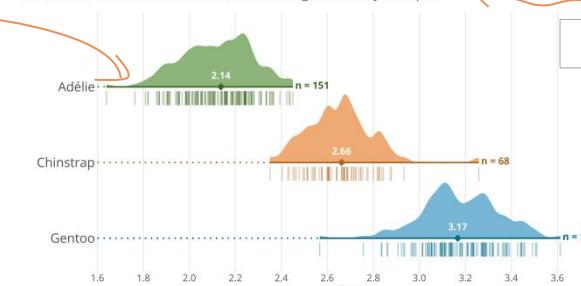


Bill Dimensions of Brush-Tailed Penguins (*Pygoscelis*)

A. Scatter plot of bill depth versus bill length.



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



Data: Gorman, Williams & Fraser (2014) PLoS ONE • Illustration: Allison Horst

ggplot2
ggplot2.tidyverse.org

Palmer Penguins

github.com/allisonhorst/palmerpenguins

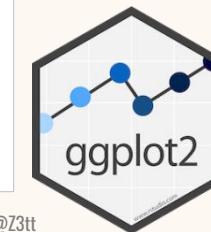


cedricscherer.com

cedricphilipscherer@gmail.com

@CedScherer

@Z3tt



```
annotation_custom(grid::rasterGrob(img))
```

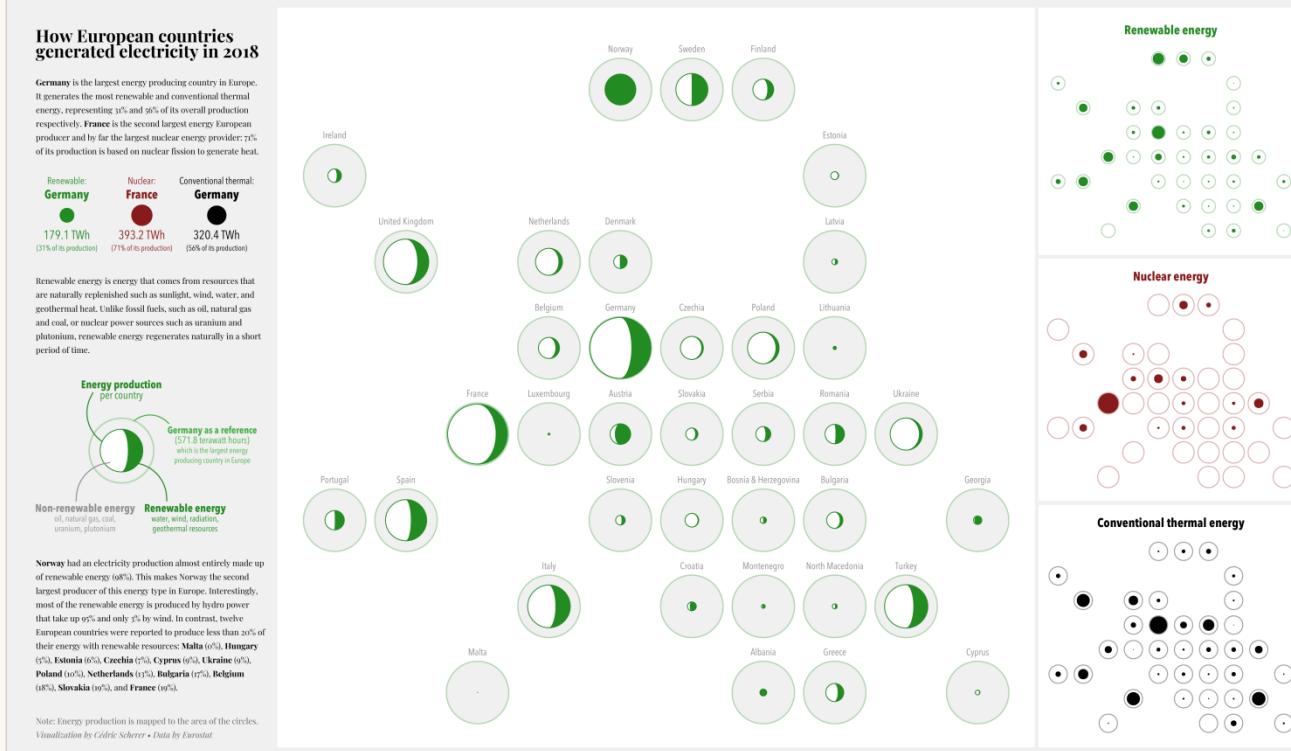


```
ggforce::geom_mark_*
```



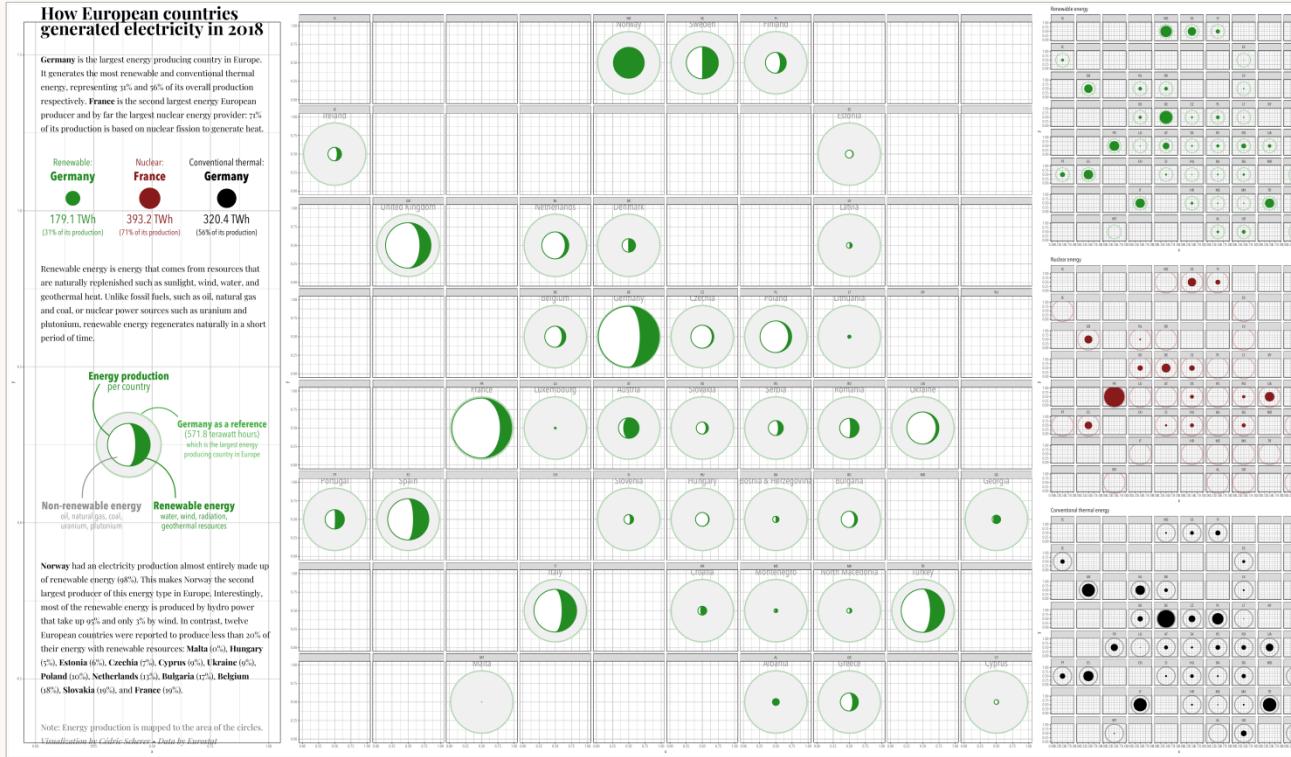
```
(p1 / p2) + plot_layout(heights = c(1, .65))
```

{patchwork} The Composer of ggplots



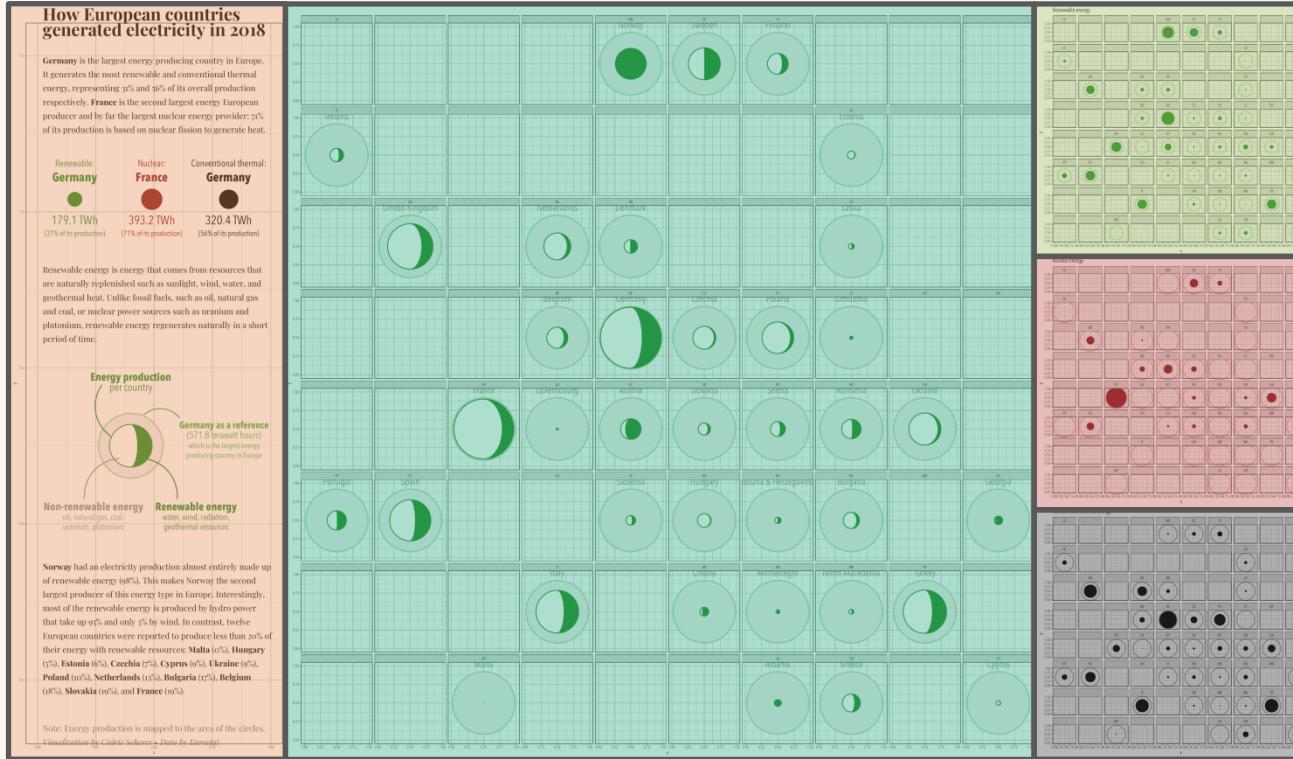
{patchwork} The Composer of ggplots

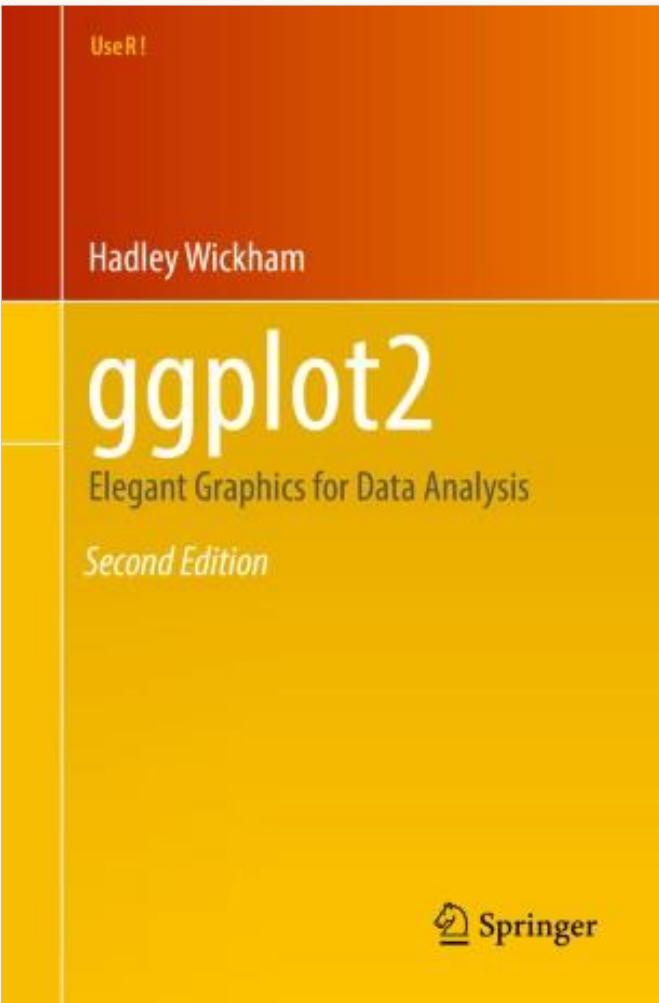
legend | main | (renewable / nuclear / thermal) + plot_layout(widths = c(.35, 1, .35))



{patchwork} The Composer of ggplots

legend | main | (renewable / nuclear / thermal) + plot_layout(widths = c(.35, 1, .35))





The screenshot of the R Graph Gallery website features a navigation bar at the top with links for CHART TYPES, QUICK, TOOLS, ALL, D3.JS, PYTHON, DATA TO VIZ, and ABOUT. A search icon is also present. Below the header, the text 'The R Graph Gallery' is centered, followed by social media sharing icons for Twitter, GitHub, LinkedIn, and others. A welcome message from the site's creator, Hadley Wickham, is displayed, encouraging users to explore the gallery and provide feedback. The main content area is organized into several sections, each showing a grid of small chart thumbnails:

- Distribution:** Violin, Density, Histogram, Boxplot, Ridgeline
- Correlation:** Scatter, Heatmap, Correlogram, Bubble, Connected scatter, Density 2d
- Ranking:** Barplot, Spider / Radar, Wordcloud, Parallel, Lollipop, Circular Barplot
- Part of a whole:** Treemap, Chord, Sunburst, Donut, Treemap, Chord

CÉDRIC SCHERER

Data Visualization & Computational Ecology

MY PERSONAL DATA VISUALIZATION YEAR 2020

Even though it was a crazy and exhausting year, there was also some good and exiting things happening. Therefore I've decided to take a short break on New Year's Day and look back at some of the positive moments of my personal data visualization journey during 2020.

POSTED BY CÉDRIC FRIDAY, JANUARY 1, 2020

WHAT DO I BINGE NEXT? A DETAILED OVERVIEW OF THE TOP 250 TV SHOWS

My contribution to the RStudio table contests visualizing relevant details of the top 250 TV shows as rated by IMDB users. I focused on displaying all the details I and my friends care about including in-line visualizations of rating trends and average runtime.

POSTED BY CÉDRIC SUNDAY, NOVEMBER 1, 2020

THE WORST DAYS OF THE CORONAVIRUS PANDEMIC SO FAR

Coronavirus SARS-CoV-2, COVID-19 or simply Corona—what started as an epidemic in China has become a global pandemic. I created an animated timeseries of daily deaths relative to each country's worst day so far to visualize the first wave of COVID-19.

POSTED BY CÉDRIC TUESDAY, MARCH 31, 2020

COMPARING THE EXTENT OF THE AUSTRALIAN BUSHFIRES 2019/20

The massive bushfires in Australia are in the news worldwide. The incredible extent of burnt land and plume of smoke is hard to imagine so I have compared the areas to countries in Europe and worldwide.

POSTED BY CÉDRIC THURSDAY, JANUARY 9, 2020

BEST TIDYTUESDAY 2019

Here are my favorite visualizations of the #TidyTuesday challenge in 2019 (from those I've seen and which I remember). I present my personal top 3 in terms of design and storytelling.

POSTED BY CÉDRIC MONDAY, DECEMBER 30, 2019

MERRY (WHITE?) CHRISTMAS!

At the end of the year, I explore the history of snow cover and white Christmas in Berlin. I wish you a merry Christmas and wonderful holidays 2019!

POSTED BY CÉDRIC TUESDAY, DECEMBER 24, 2019

OLDER POSTS →



ABOUT ME

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



Support me

FEATURED TAGS



FRIENDS

DataVizSociety
RDS Community CorrelAid
Will Chase
Georgios Karmanis
Marco Scisini Matthias Stahl
Heureka Labs

cedricscherer.com
cedricphilippscherer@gmail.com
twitter.com/CedScherer
github.com/z3tt

Thank you!

[Search or jump to...](#)

[Pull requests](#) [Issues](#) [Marketplace](#) [Explore](#)

[Unwatch](#) • 53 [Star](#) 415 [Fork](#) 77

[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Security](#) [Insights](#) [Settings](#)

[master](#) • [1 branch](#) • [0 tags](#)

[Go to file](#) [Add file](#) [Create](#)

About

[My contributions to the #TidyTuesday challenge](#)

[ggplot2](#) [dplyr](#) [tidyverse](#) [rlang](#)
[Hadley](#) [tidyverse](#) [tidyTuesday](#) [tidyTuesdayChallenge](#)

[Readme](#)

Releases

No releases published [Create a new release](#)

Packages

No packages published [Publish your first package](#)

Contributors

Z3tt Cédric Scherer
 gkaramanis Georgios Karamanis

Languages

R 96.3% CSS 1.7%

DISCLAIMER:
 ☣ No Illustrator or Photoshop was harmed during the making of these visualizations.
 ☣ I certify ggplot quality.

Feedback is very welcome on [Twitter](#) or [mail](#).

• Challenges in chronological order (click to expand)

- Challenges 2018
 - 2018/01 Global Mortality
 - 2018/03 Malaria
- Challenges 2019
 - 2019/1 Seattle Bike Traffic
 - 2019/2 The Economist's Data Viz Mistakes
 - 2019/3 Anime Data
 - 2019/4 Chicago Bird Collisions
 - 2019/5 Global Student to Teachers Ratios
 - 2019/6 Nobel Prize Winners
 - 2019/7 Global Plastic Waste
 - 2019/8 Wine Ratings
 - 2019/9 Meteorites
 - 2019/10 Christmas Bird Counts
 - 2019/11 Media Franchise Revenues
 - 2019/12 FIFA Women's World Cup
 - 2019/13 Roman Emperors
 - 2019/14 Nuclear Explosions
 - 2019/15 Simpsons Guest Stars
 - 2019/16 Moore's Law
 - 2019/17 US National Park Visits
 - 2019/18 US School Diversity
 - 2019/19 NYC Pizza Ratings
 - 2019/20 Big Cars
 - 2019/21 NYC Squirrels
 - 2019/22 Replicating Plots (Rap Edition)
 - 2019/23 Adoptable Dogs
 - 2019/24 Christmas Songs
- Challenges 2020
 - 2020/01 Bring Your Own Data: White Christmas in Berlin
 - 2020/02 Australian Fires
 - 2020/03 Passwords
 - 2020/04 Spotify Songs: The Golden Age of Hip Hop
 - 2020/05 San Francisco Trees
 - 2020/06 Hotel Bookings
 - 2020/07 Foot Carbon Footprint
 - 2020/08 The Office on IMDb
 - 2020/09 Beer Production
 - 2020/10 Tour de France History
 - 2020/11 Best Hip-Hop Songs of All Time
 - 2020/12 Broadway Musicals (RTFM Edition)
 - 2020/13 Animal Crossing
 - 2020/14 Volcano Eruptions
 - 2020/15 Beach Volleyball

[Search or jump to...](#)

[Pull requests](#) [Issues](#) [Marketplace](#) [Explore](#)

[Unwatch](#) • 6 [Star](#) 133 [Fork](#) 23

[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#) [Security](#) [Insights](#) [Settings](#)

[main](#) • [OutlierConf2021 / README.md](#)

[Go to file](#) ...

Z3tt [update readme](#)

Latest commit [f422c](#) of now [History](#)

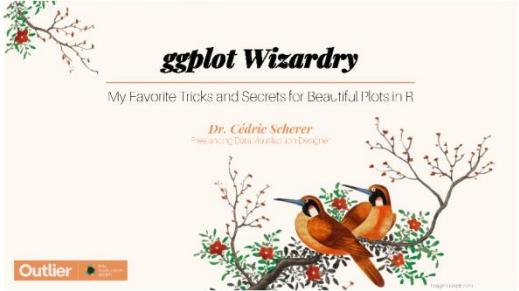
1 contributor

77 lines (84 loc) 4.96 kB

[Raw](#) [Name](#) [Edit](#)

ggplot Wizardry: My Favorite Tricks and Secrets for Beautiful Plots in R

[Slides](#) and [hands-on codes](#) for my talk at the 1st OutlierConf, February 4–7 2021.



ggplot Wizardry

My Favorite Tricks and Secrets for Beautiful Plots in R

Dr. Cédric Scherer
 Playing with Data Visuals, Design & Code

Outlier

TUTORIAL 

SLIDES FROM TODAY 

Links:

- Slides
- Recording
- [Handson Tutorial](#) ←
- Codes

Bonus: Extended version!

- Slides from the useR! Oslo meeting on March 24, 2021

About the talk:

In this talk, I present my favorite tips and tricks with regard to the ggplot2 package, a library for plotting in the programming language R. I will cover functions that are helpful but many may not be aware of as well as a collection of interesting functions from a large range of extension packages. The talk is intended for people who already know how to code in R and ggplot2. However, I am going to cover a diverse collection of tips so I hope everyone can pick something helpful independent from their level of expertise.



cedricscherer.com

cedricphilippscherer@gmail.com

twitter.com/CedScherer

github.com/z3tt

Thank you!

[DATAVIS](#) [TUTORIAL](#) [TOYVERSE](#) [GGPLOT2](#)

A GGPLOT2 TUTORIAL FOR BEAUTIFUL PLOTTING IN R

POSTED BY CÉDRIC ON MONDAY, AUGUST 5, 2013

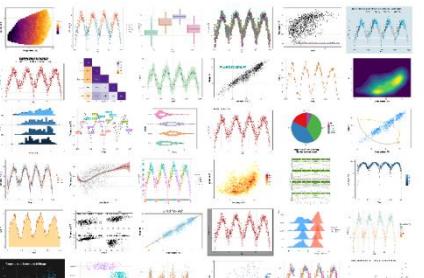
Last update: 2020-12-07

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.

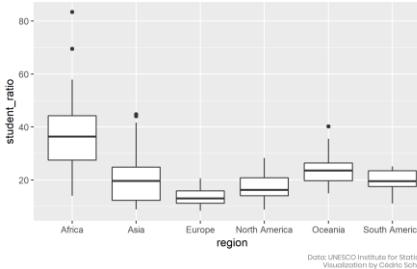
Since the blog entry by Zev Ross was not updated for some years and step by step this became a unique version of a tutorial, I decided to host the updated version on my GitHub. Now it finds its proper place on this homepage! (Plus I added a ton of other updates—just to name a few: The fantastic [\(patchwork\)](#), [\(ggttext\)](#) and [\(ggeforce\)](#) 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!)


[DATAVIS](#) [TUTORIAL](#) [ANIMATIONS](#) [GGPLOT EVOLUTION](#) [GSPLOT2](#) [TOYVERSE](#) [TOYTUESDAY](#)

THE EVOLUTION OF A GGPLOT (EP. I)

POSTED BY CÉDRIC ON FRIDAY, MAY 17, 2019

The Evolution of a ggplot



- Aim of this Tutorial
- Data Preparation
- The Default Boxplot
- Sort Your Data!
- Let Your Plot Shine—Get Rid of the Default Settings
- The Choice of the Chart Type
- More Geoms, More Fun, More Info!
- Add Text Boxes to Let The Plot Speak for Itself
- Bonus: Add a Tile Map as Legend
- The Final Evolved Visualization
- Complete Code for Final Plot
- Post Scriptum: Mean versus Median

AIM OF THIS TUTORIAL

In this series of blog posts, I aim to show you how to turn a default ggplot into a plot that visualizes information in an appealing and easily understandable way. The goal of each blog post is to provide a step-by-step tutorial explaining how my visualization have evolved from a typical basic ggplot. All plots are going to be created with 100% [\(ggplot2\)](#) and 0% Inkscape.

In the first episode, I transform a basic boxplot into a colorful and self-explanatory combination of a jittered dot strip plot and a lollipop plot. I am going to use [data](#) provided by the UNESCO on global student to teacher ratios that was selected as data for the #TidyTuesday challenge 19 of 2019.

- cedricscherer.com
 - cedricphilippscherer@gmail.com
 - twitter.com/CedScherer
 - github.com/z3tt

Thank you!