

# **CORH 203 Data Visualization**

March, 2024

*Data! data! data! [...] I can't make bricks  
without clay.*

— Sherlock Holmes, The Adventure of the Copper Beeches

<http://tinyurl.com/corh-203-slides>

# **Data Literacies**

**What is (are) Research Data?**

*[R]esearch data represent any material  
derived from a source of potential  
information, both measurable and abstract,  
**gathered for analysis** or as part of  
developing findings or insights on a  
subject.*

- Unpublished RDM manuscript.

**What are Possible Sources of Data?**

people, films, photographs, recordings,  
artifacts, specimens, music score sheets,  
art, an existing data set...

**How are data classed?**



# **Scales**

Nominal (No Order)

Ordinal (Rank Ordering)

Interval (Arbitrary Zero, Quantifiable  
Difference)

Ratio (Proportional)

Stevens, S. (1946). On the theory of scales of Measurement. Science.  
103 (2684).

# **Collection**

Counted (Discrete, Observerd)

Measured (Continuous)

Set (Parameter, Categorical)

Reported (Survey, Nominal, Ordinal, Free  
Text)

# **Computational Representations**

Numeric (integers, rationals)

Character (strings)

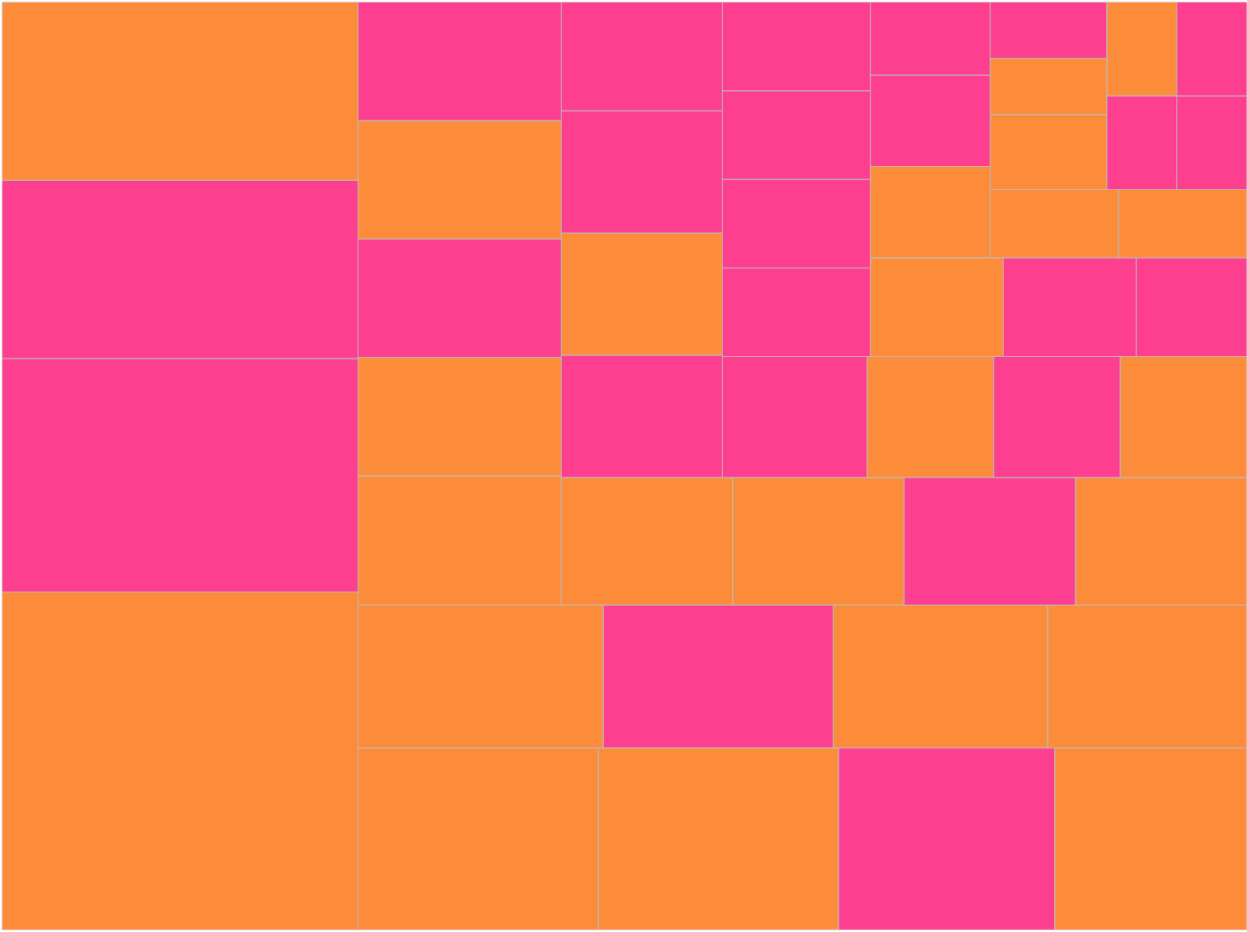
Boolean (true, false)...

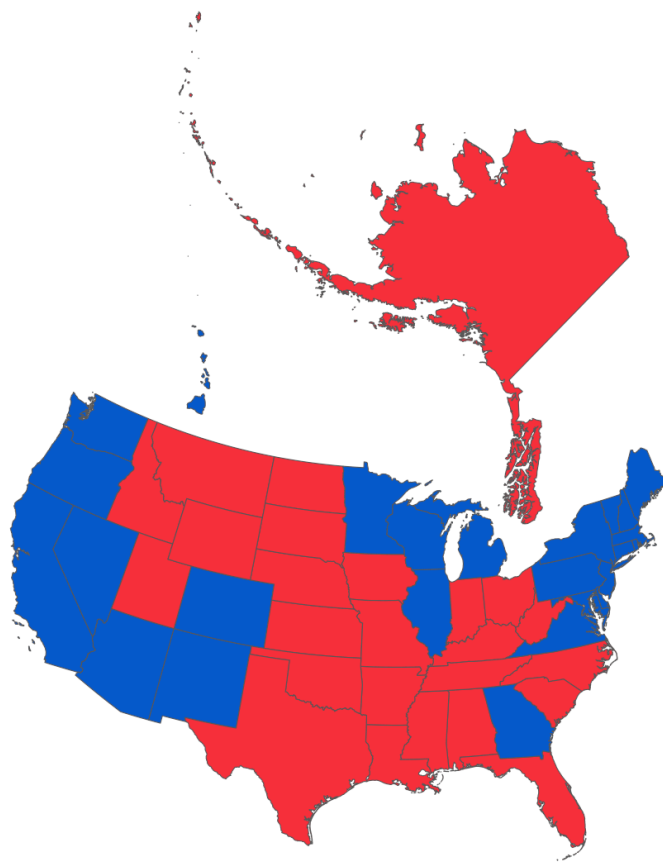
# **Attributes**

Dates, Currencies, Coordinates...

# **What is Data Visualization?**

The **abstraction** of data, using shapes,  
colour, and position.







306

Joseph R. Biden Jr. ✓

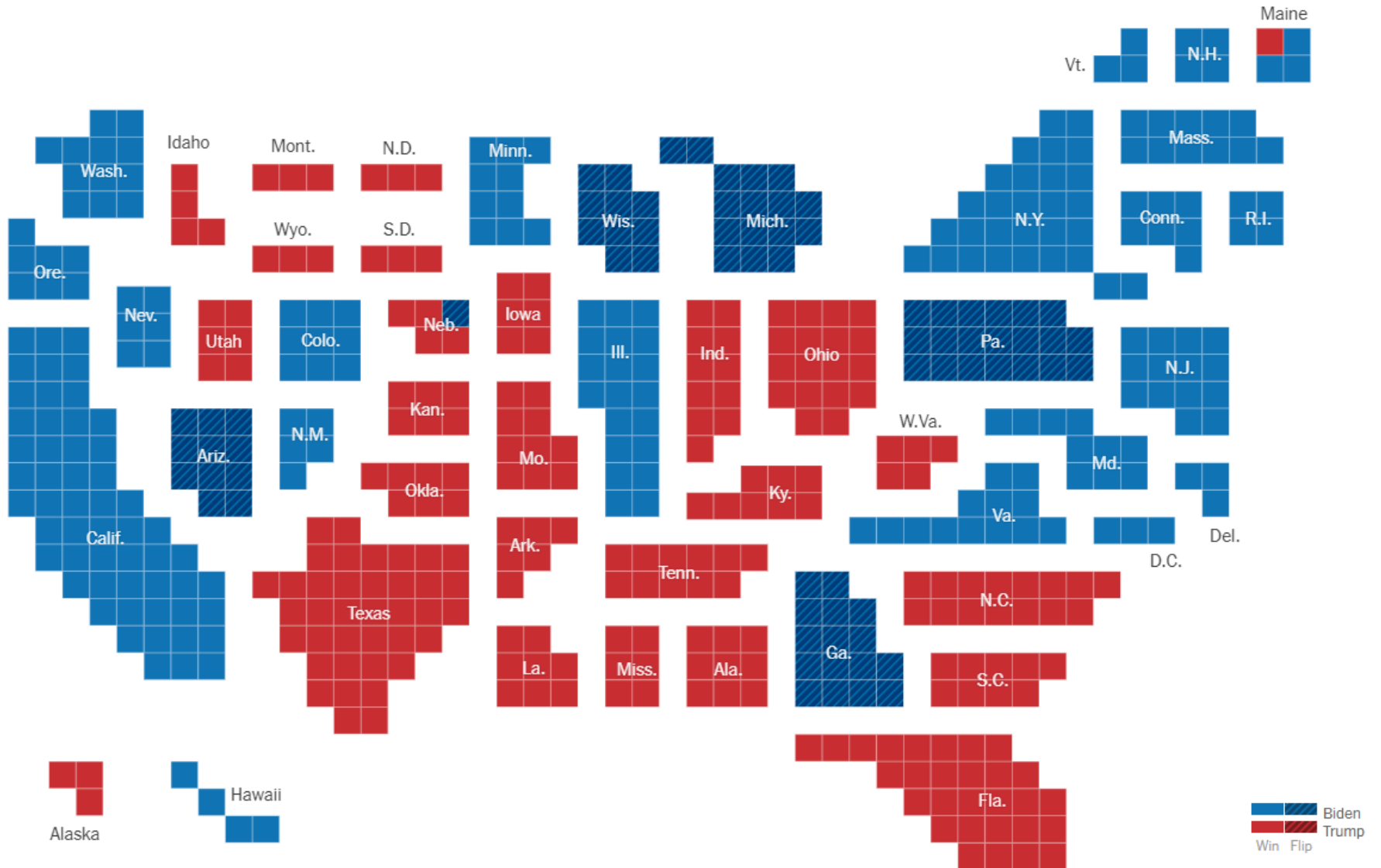
79,554,207 votes (51.0%)

232

Donald J. Trump

73,611,180 votes (47.2%)

270  
TO WIN



Source: <https://static01.nyt.com/images/2020/11/20/learning/2020electoralvotesmapLN/2020electoralvotesmapLN-superJumbo.png?quality=75&auto=webp>

# **Why do we Visualize Data?**

*Data visualization is the graphical display of abstract information for two purposes: sense-making (also called data analysis) and communication.*

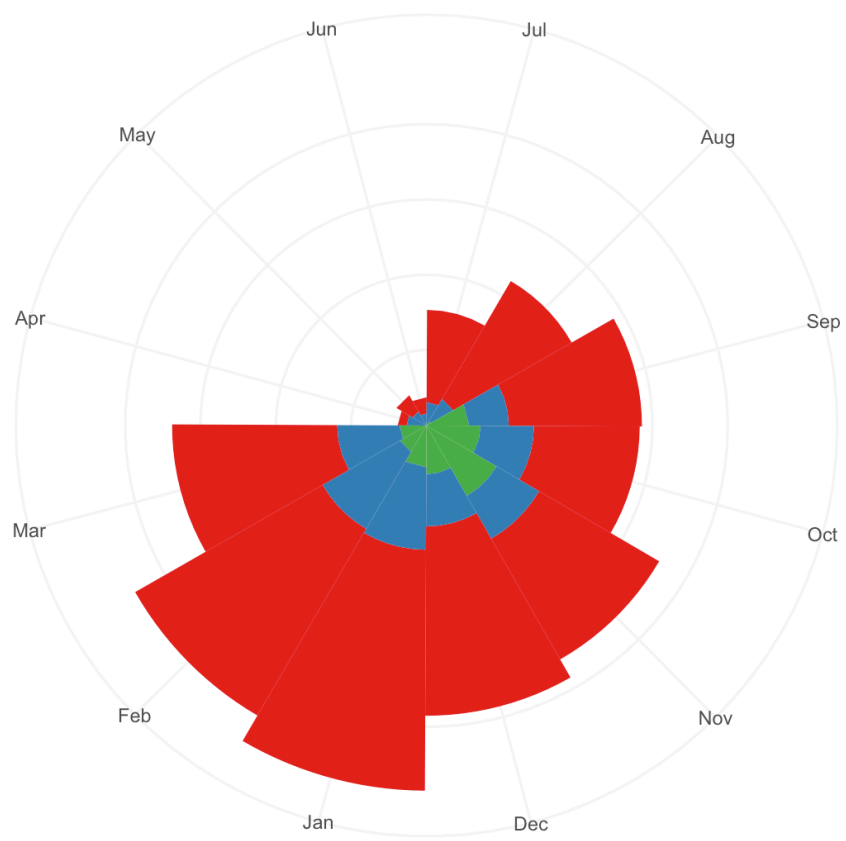
Stephen Few. [Data Visualization for Human Perception](#).

*The greatest value of a picture is when it  
forces us to notice what we never expected to  
see.*

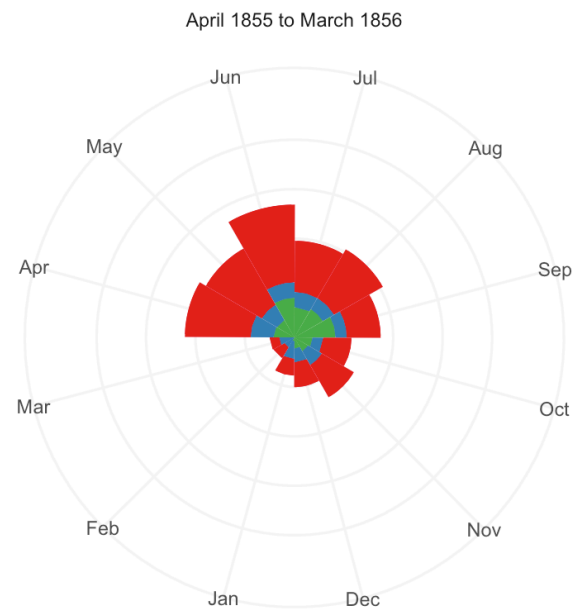
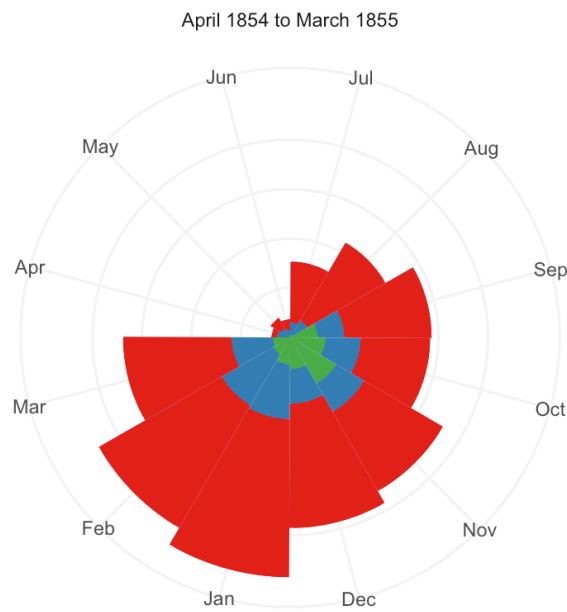
John W. Tukey (1977). "Exploratory Data Analysis"



*[T]he [...] optic nerves are sending what we now know are 20 megabits a second of information back to the brain [...] [It] is being transformed into information, into thinking, right as that step from the retina to the brain. And the brain is really busy, and it likes to economize. And so it's quick to be active and jump to conclusions. So if you're told what to look for, you can't see anything else.*

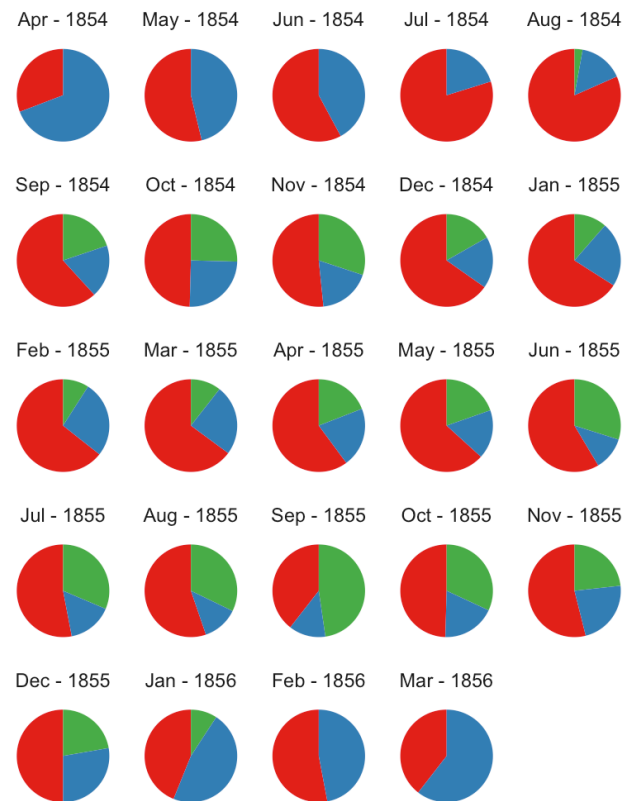






Cause of Mortality

Disease Other Wounds

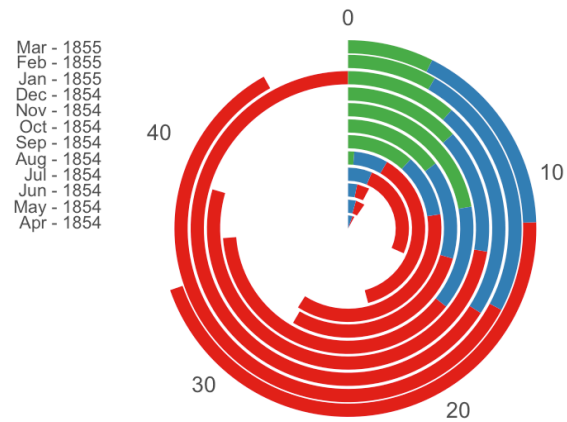


Cause of Mortality ■ Disease ■ Other ■ Wounds

*There is no data that can be displayed in a pie chart, that cannot be displayed better in some other type of chart.*

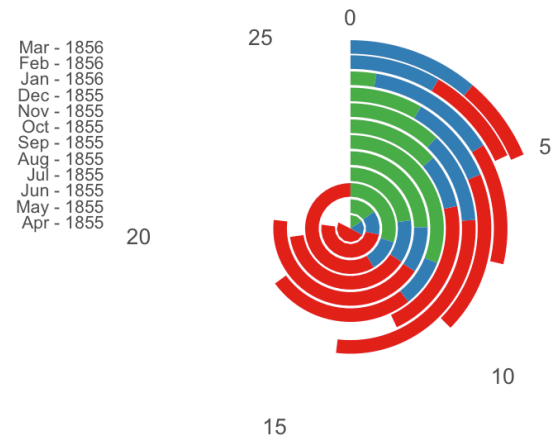
- John W. Tukey. Attributed quote.

April 1854 to March 1855



Mortality (rate / 1000)

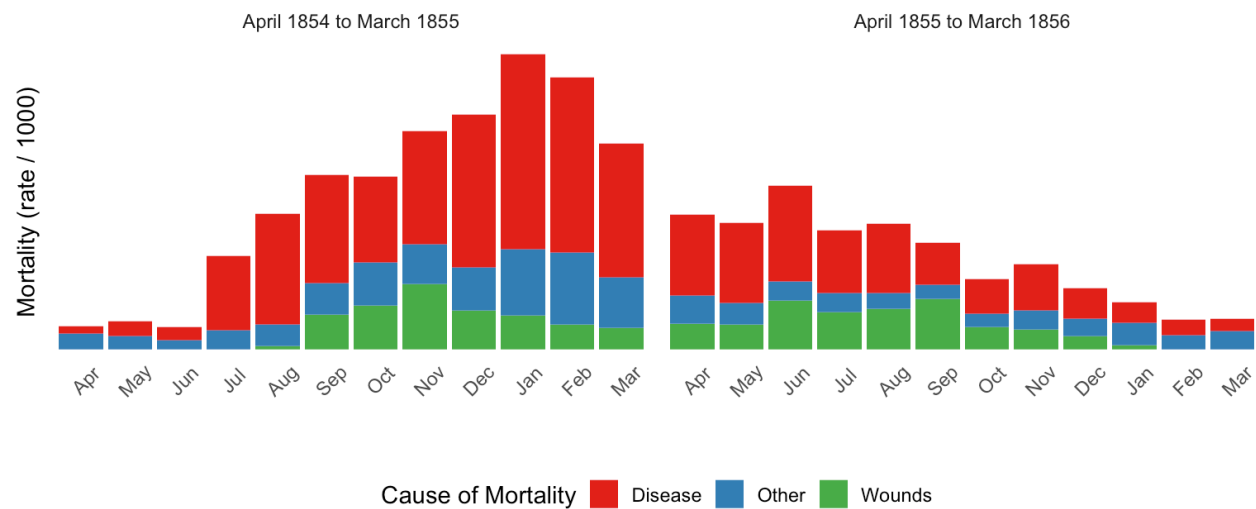
April 1855 to March 1856

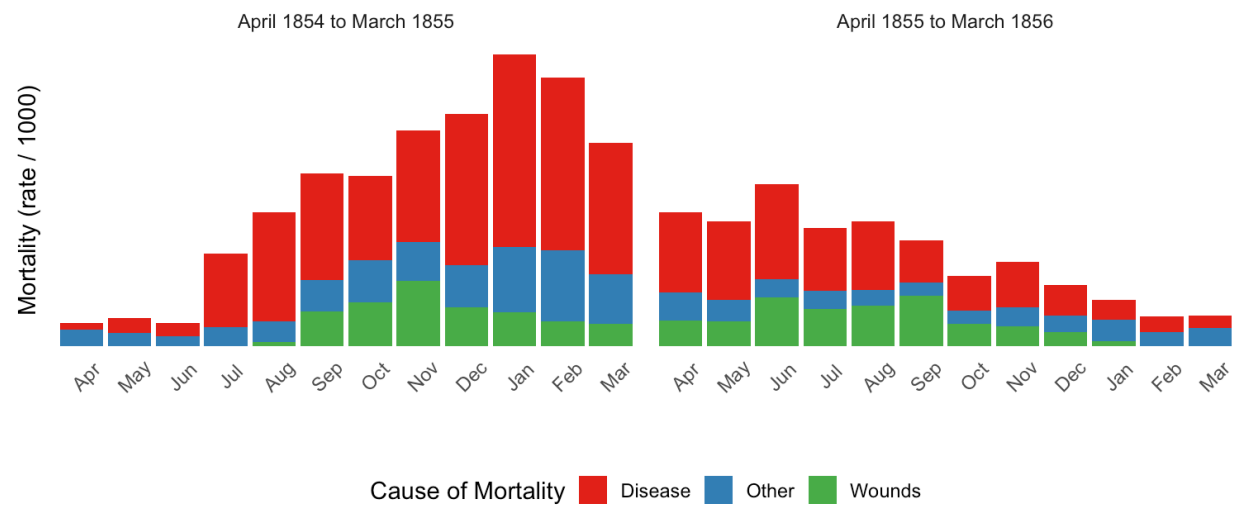
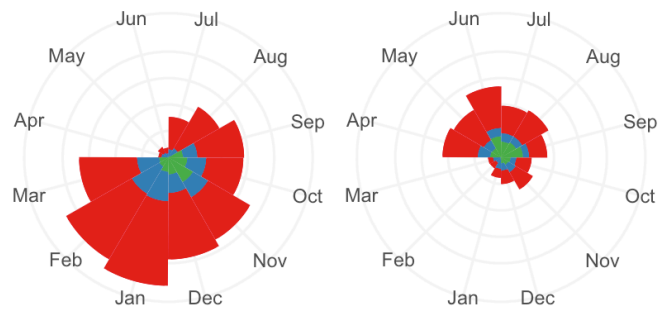


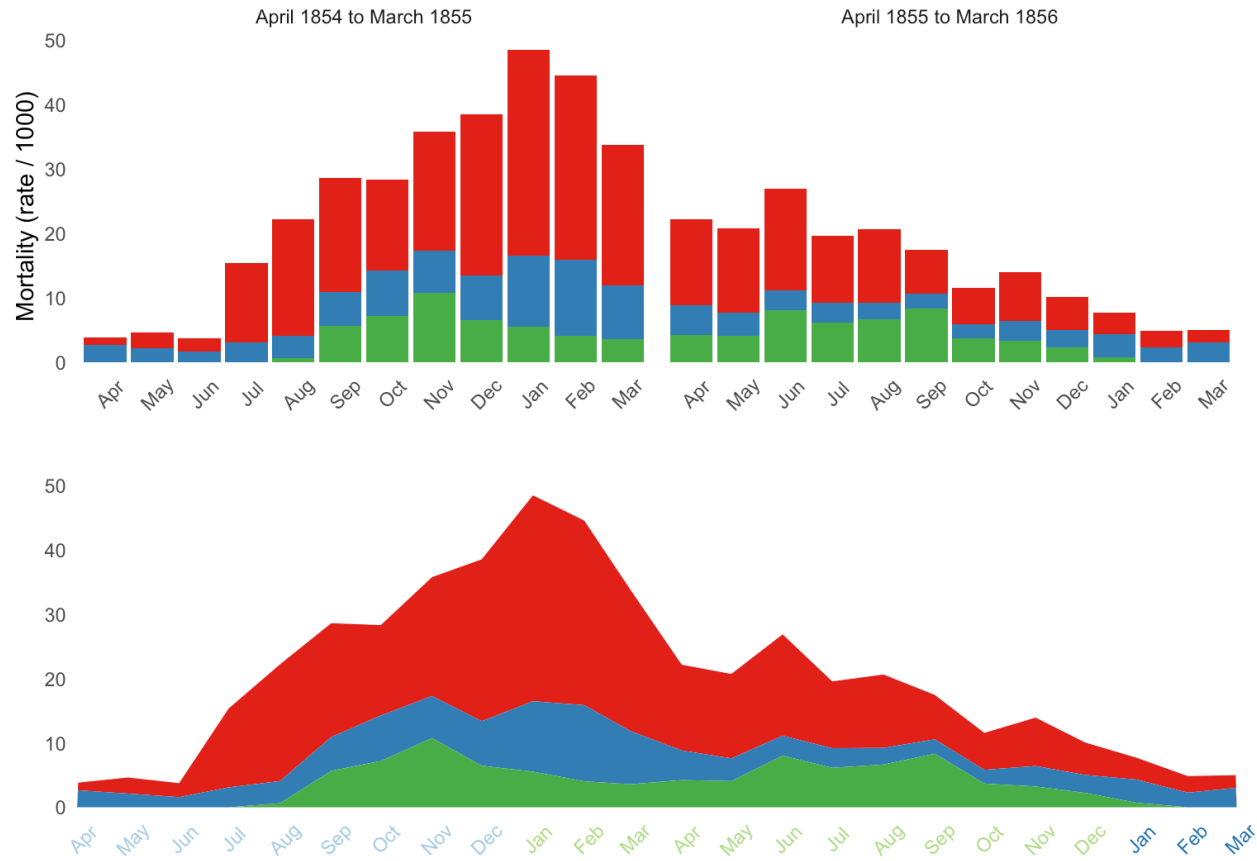
Mortality (rate / 1000)

*There is no such thing as information  
overload, just bad design. If something is  
cluttered and/or confusing, fix your design.*

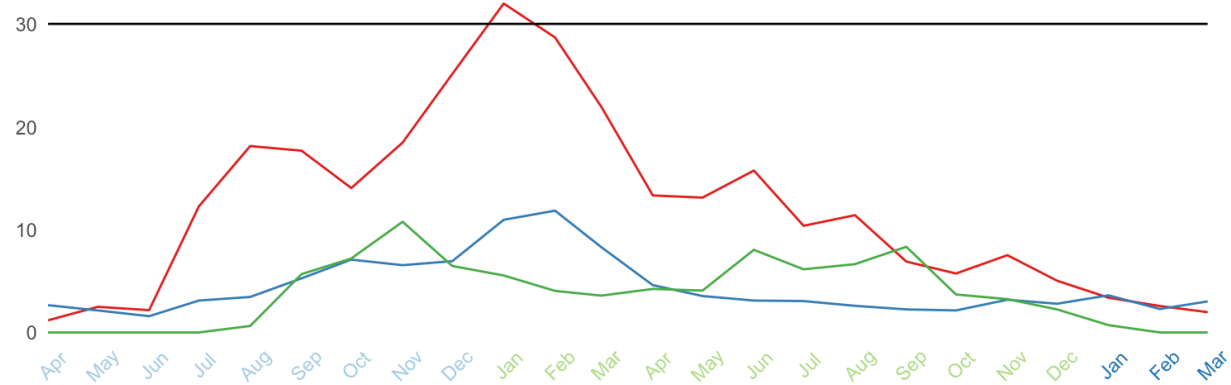
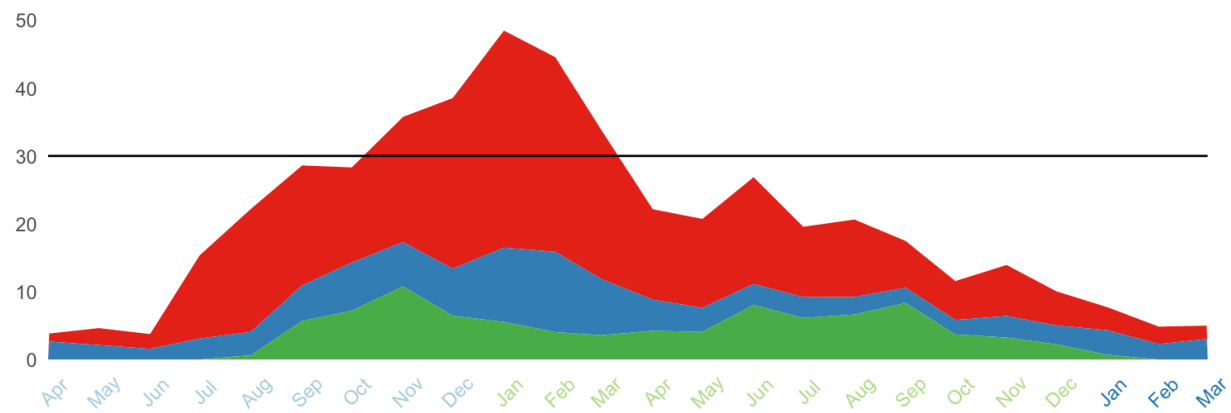
- Edward Tufte. Attributed quote.









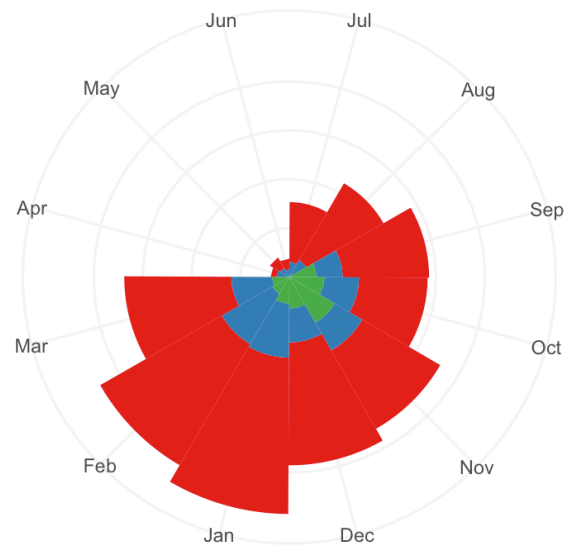


*A beautiful visualization has a clear goal, a message, or a particular perspective on the information that it is designed to convey.*

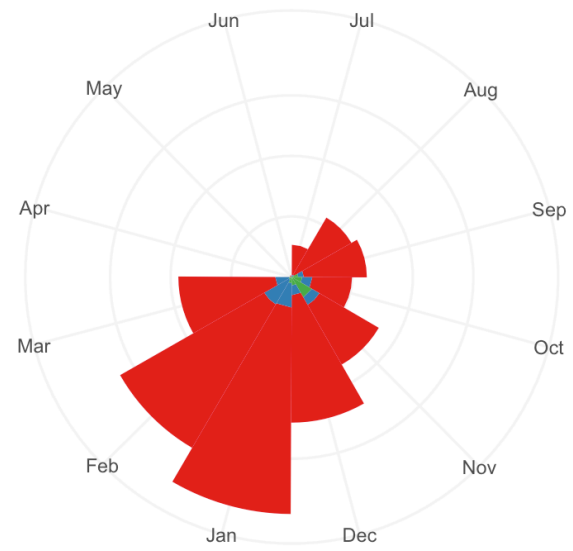
*Access to this information should be as straightforward as possible, without sacrificing any necessary, relevant complexity.*

Noah Iliinsky (2010). [On Beauty](#). In *Beautiful visualization: looking at data through the eyes of experts*.

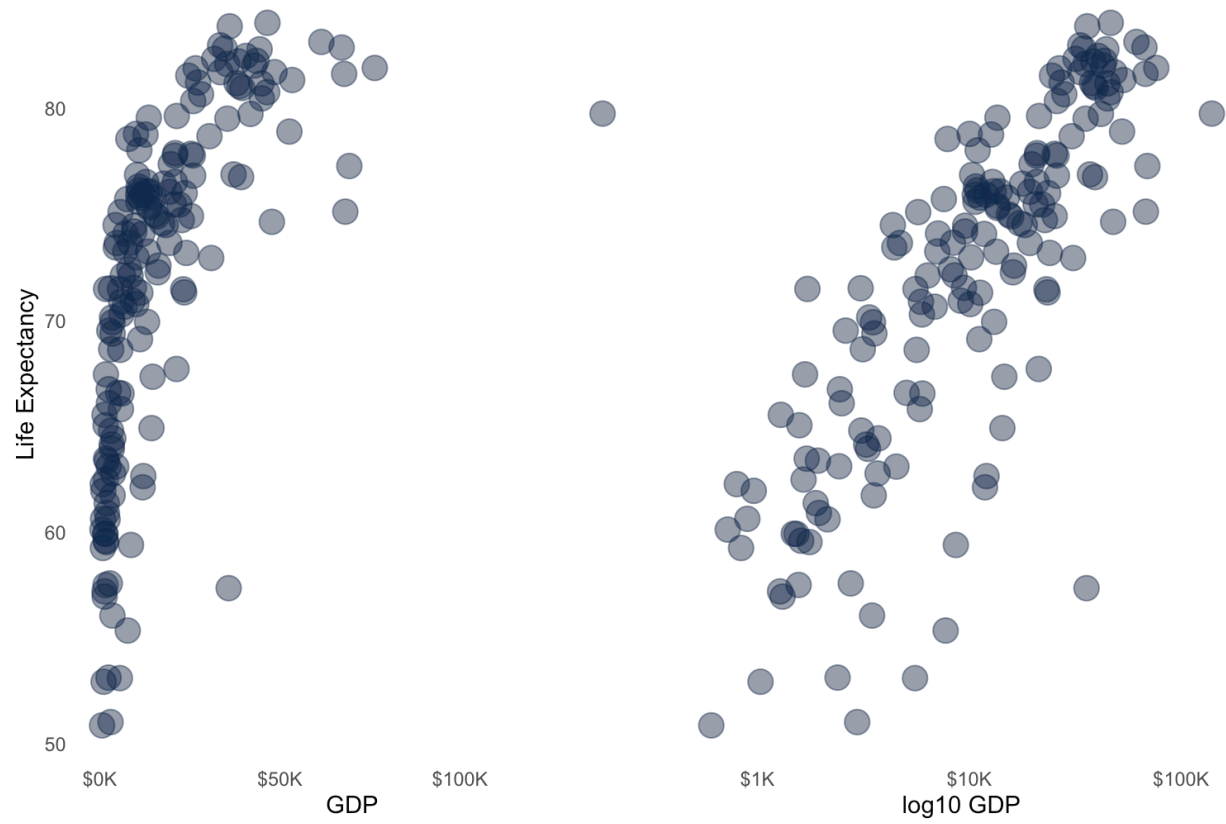
1854 - 1855, square root of rates / 1000



1854 - 1855, rates / 1000



Global Life Expectancy Against GDP, 2015  
Comparing non vs log transformed gdp.





0

250000

500000

750000

1000000



3



4



5



6

# **Colour and Your Data**

Continuous (Ratio)



Discrete (Ordinal)



Diverging (Integer)

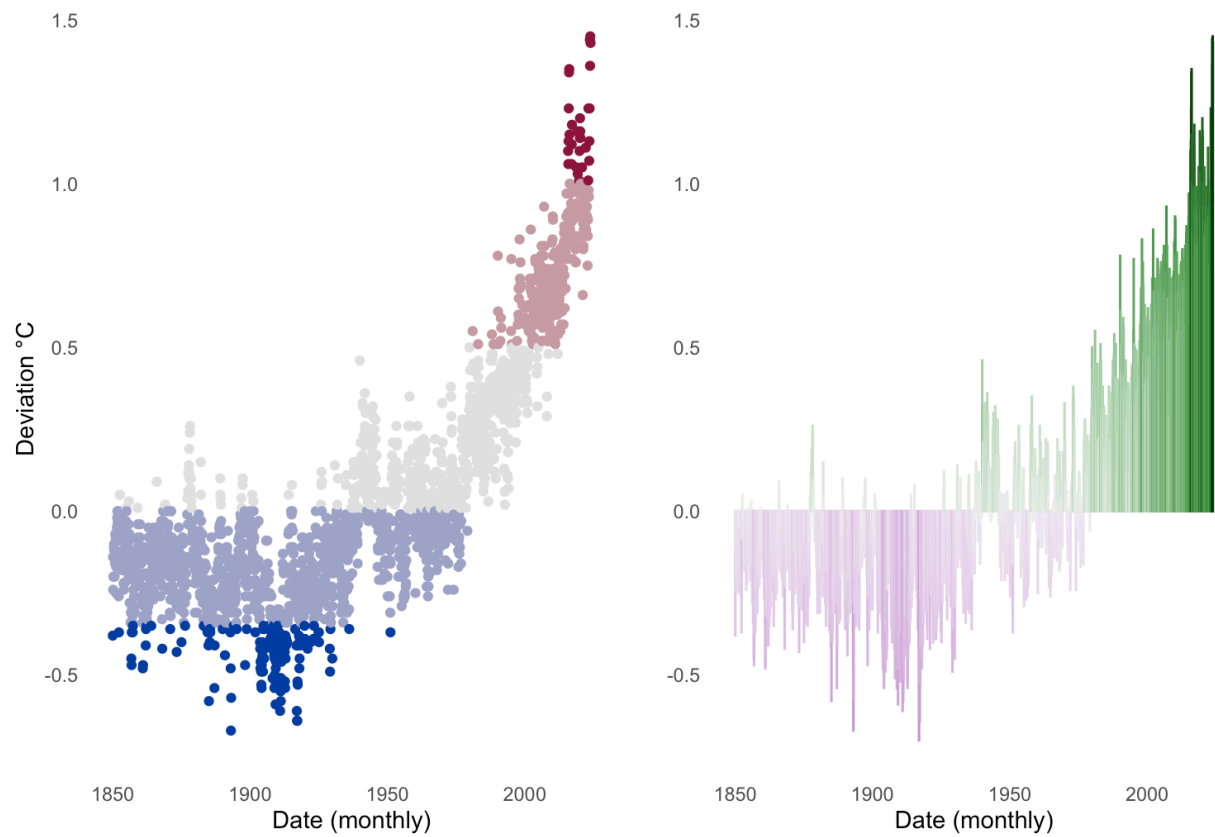


Diverging (continuous)



Categorical (Nominal)





Mean global land and ocean temperature data.  
Divergence calculated from 1901 - 2000 mean.  
Source: <https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/global/time-series>

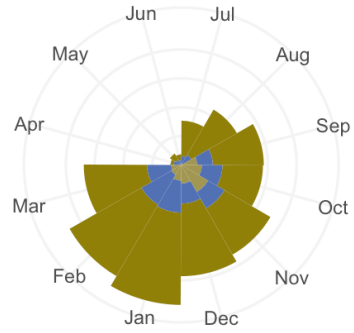


# **Colour and Your Audience**



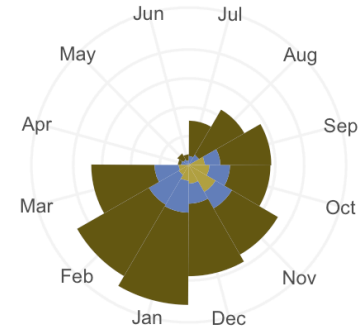
### Deutanomaly

1854 - 1855, square root of rates / 1000



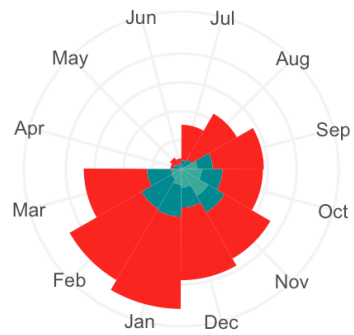
### Protanomaly

1854 - 1855, square root of rates / 100



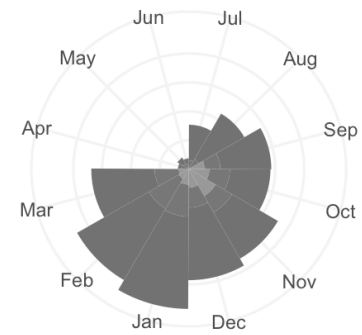
### Tritanomaly

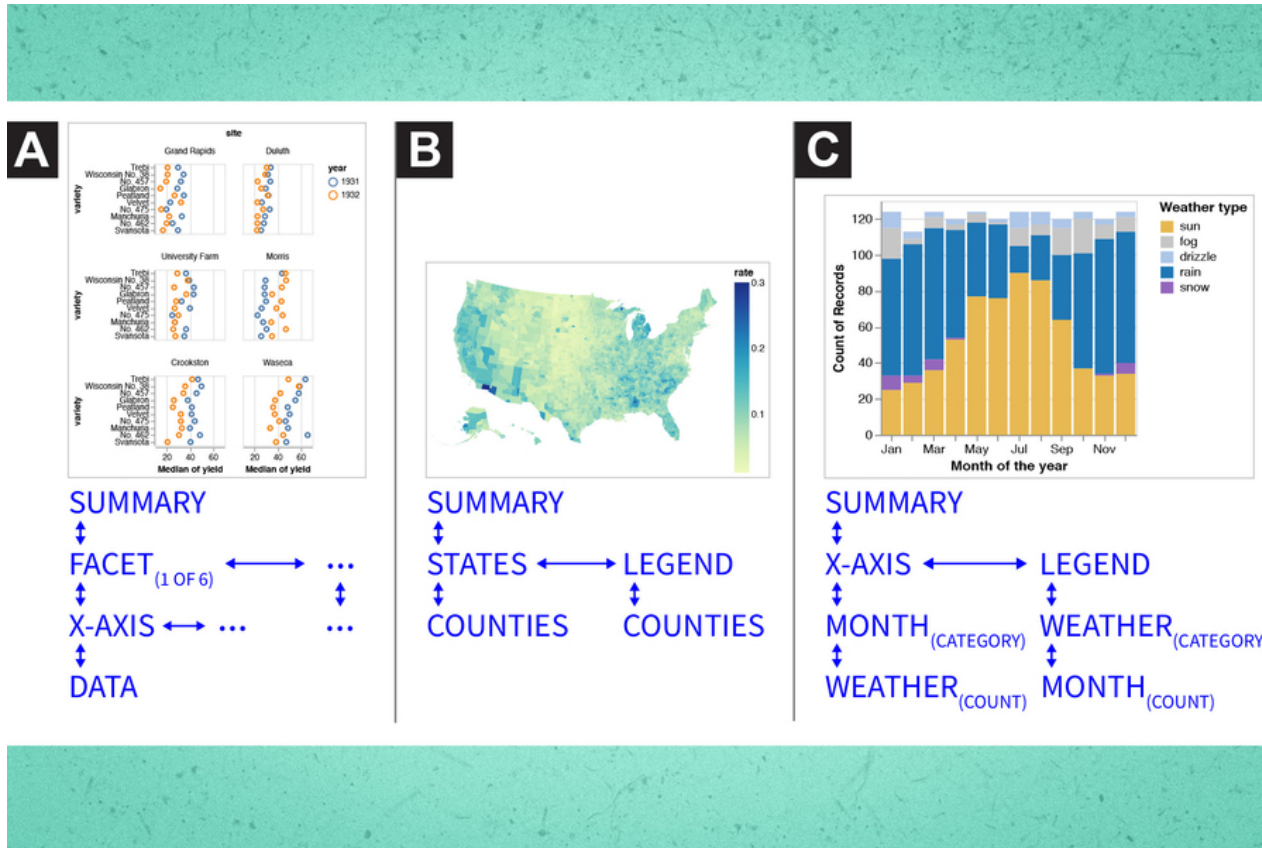
1854 - 1855, square root of rates / 1000



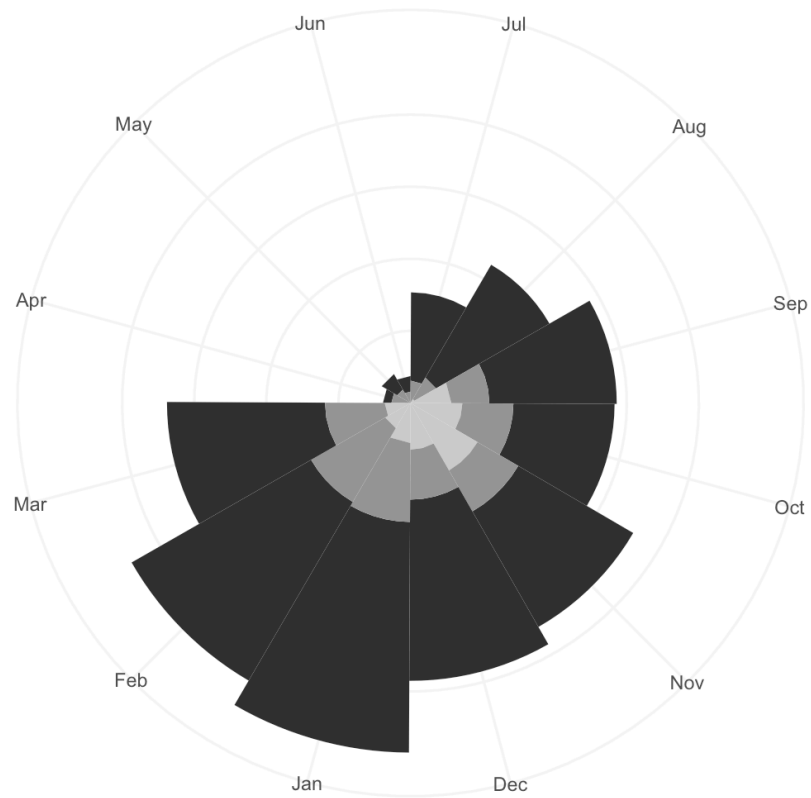
### Desaturated

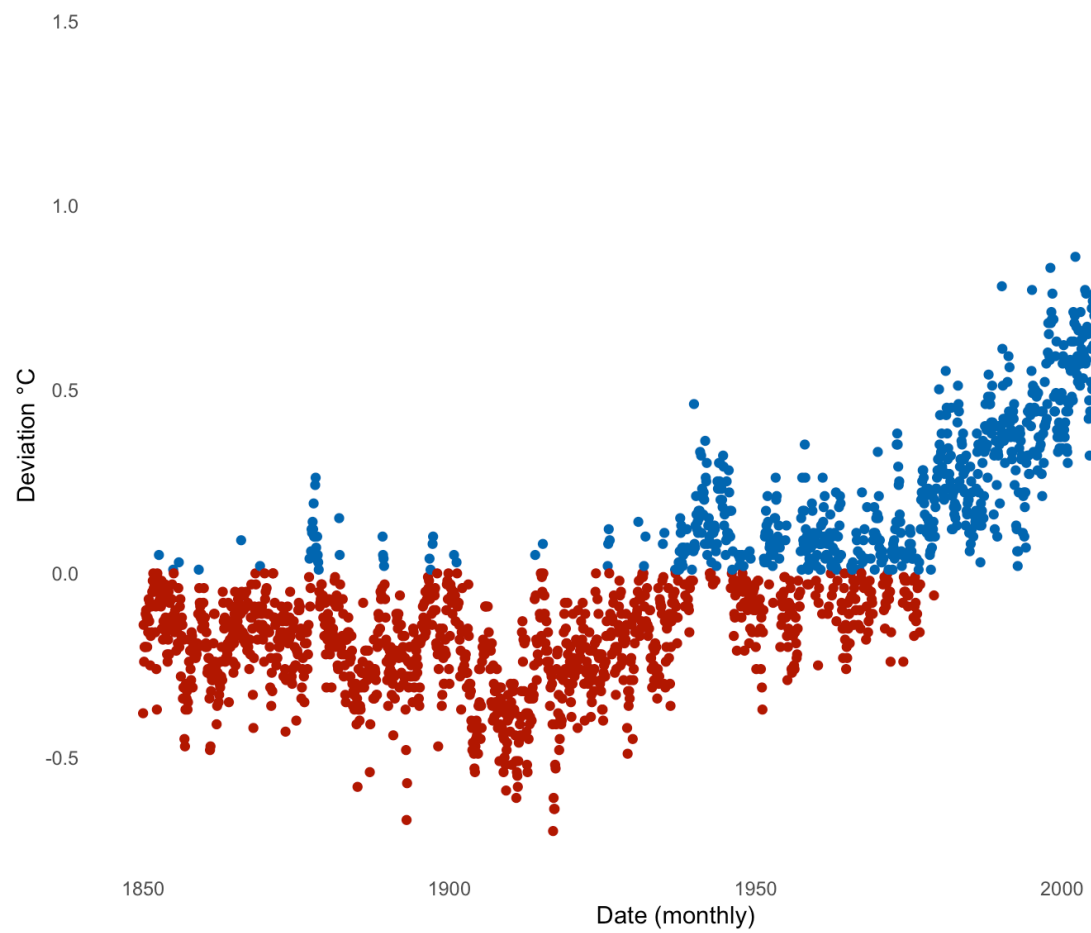
1854 - 1855, square root of rates / 100





Zewe, A. (June 2, 2022). Making data visualization more accessible for blind and low-vision individuals. *MIT News*. <https://news.mit.edu/2022/data-visualization-accessible-blind-0602>







A diagram consisting of a 2x2 grid of colored squares. The top-left square is blue and contains the word 'Trust'. The top-right square is magenta and contains the word 'Love'. The bottom-left square is red and contains the word 'Passion'. The bottom-right square is orange and contains the word 'Joy'. The squares are separated by thin white lines.

Trust

Love

Passion

Joy

Medium (print, digital, distance)



# **Creating a Colour Scheme**

**Hue**

Colour

**Saturation**

Amount of Grey

**Luminosity**

Amount of White and Black

Try building palettes by adjusting a single  
element at a time.

**ColorBrewer**

Qualitative colour schemes

**Data Color Picker**

Input a colour, request a palette

**Picular**

Word associated colour selections

**Color Palette Generator**

Derive a palette from an image

**Colour Maps**

**Generator & Overview**

**Contrast Checker**

Evaluate Contrast Ratio

**Coblis**

Colour Blindness Checker

