

# Data visualization

---

# Principles

---

# Data visualization principles

- Data-ink ratio:
  - ◆ Maximize the amount of data represented while minimizing unnecessary elements
- Graphical integrity:
  - ◆ Ensure that visual representations accurately convey the data without distortion.
- Simplicity:
  - ◆ Present information simply and clearly, avoiding unnecessary complexity.
- “Above all else, show the data”

(From Edward Tufte: “The Visual Display of Quantitative Information”)

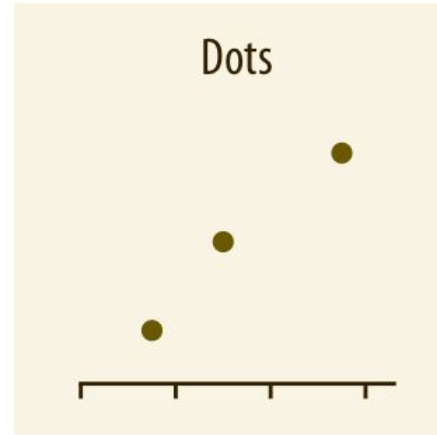
# Data visualization principles

- Design for your audience
  - ◆ Yourself or external viewers
- In exploratory visualizations, don't obscure data from yourself.
  - ◆ Explore multiple views.
- In presentation visualizations, tell a story/communicate a key finding.

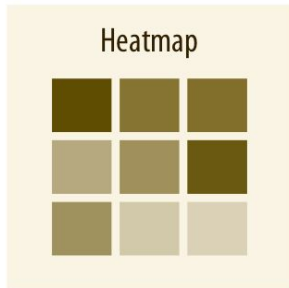
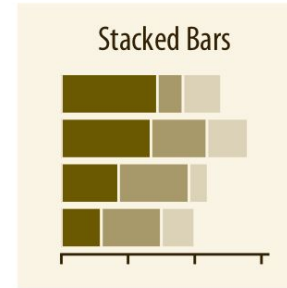
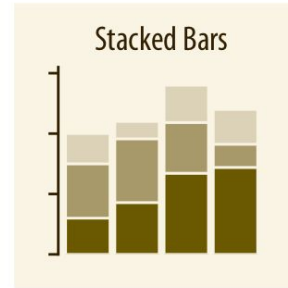
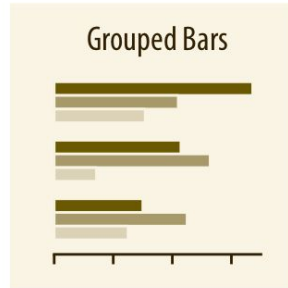
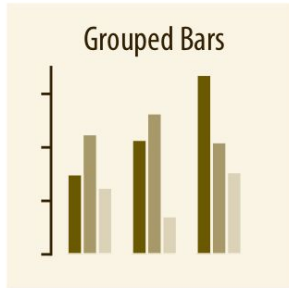
Use the right visualization type for the data

---

# Amounts

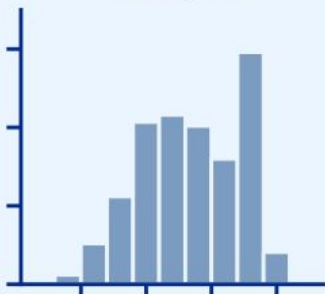


# Amounts (groups)

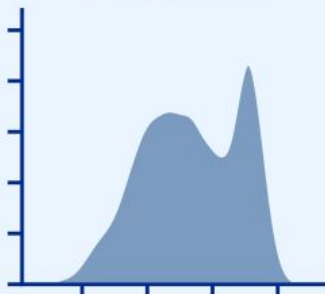


# Distributions

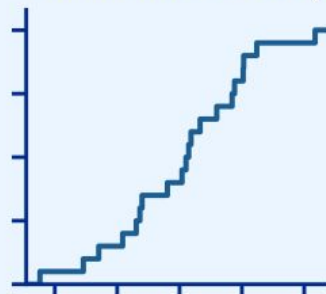
Histogram



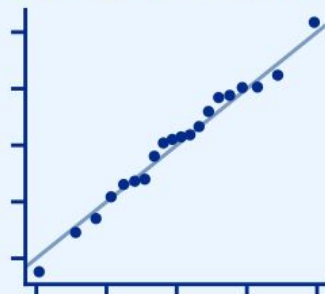
Density Plot



Cumulative Density

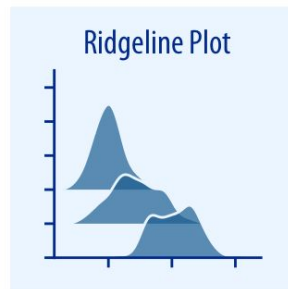
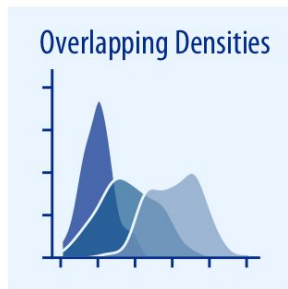
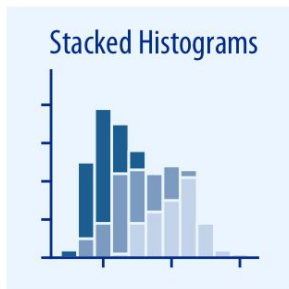
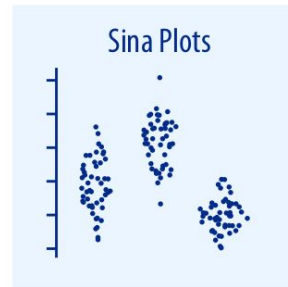
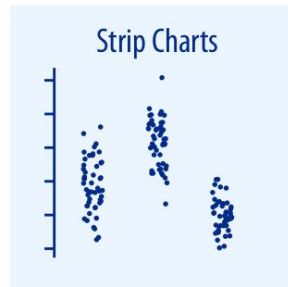
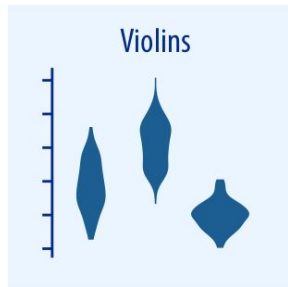
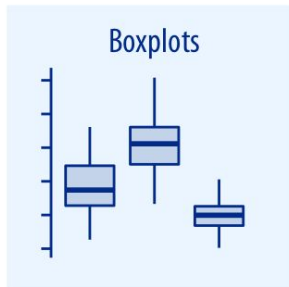


Quantile-Quantile Plot





# Distributions



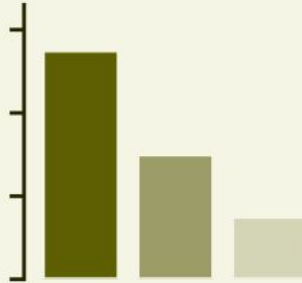
<https://clauswilke.com/dataviz/directory-of-visualizations.html>

# Proportions

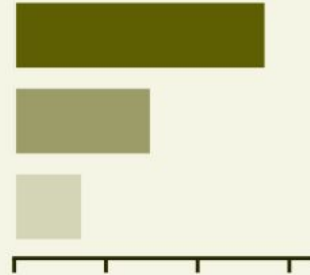
Pie Chart



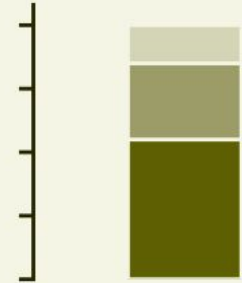
Bars



Bars



Stacked Bars

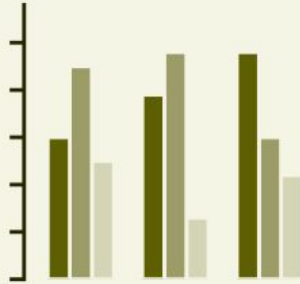


# Proportions (groups)

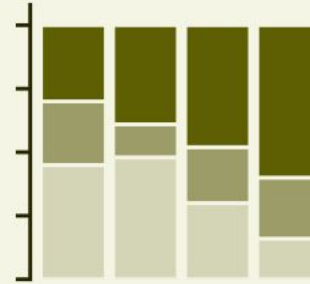
Multiple Pie Charts



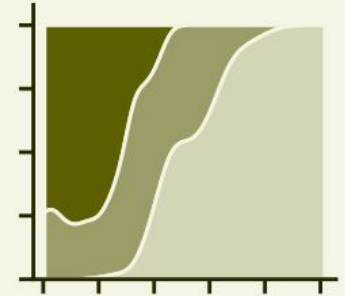
Grouped Bars



Stacked Bars

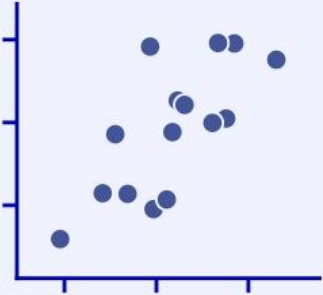


Stacked Densities

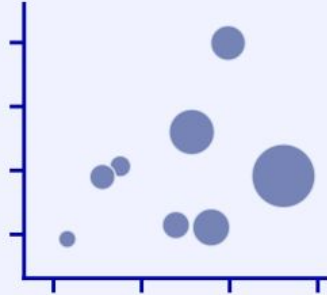


# Relationships

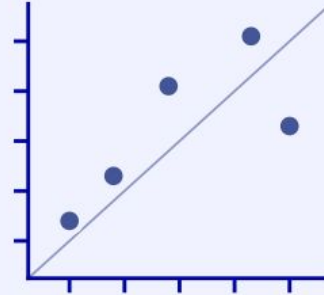
Scatterplot



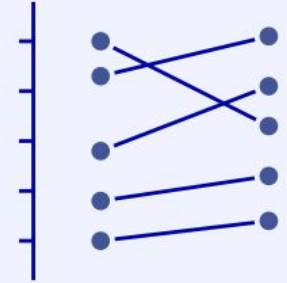
Bubble Chart



Paired Scatterplot

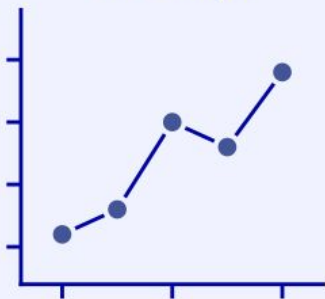


Slopegraph

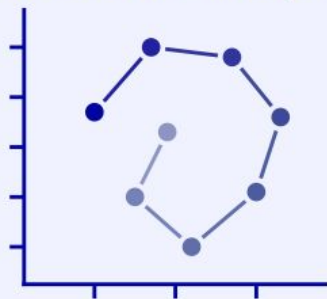


# Relationships (directional)

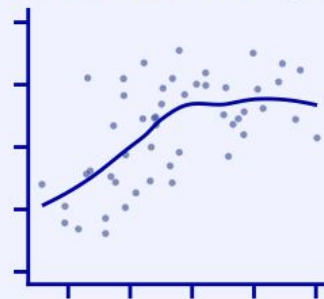
Line Graph



Connected Scatterplot



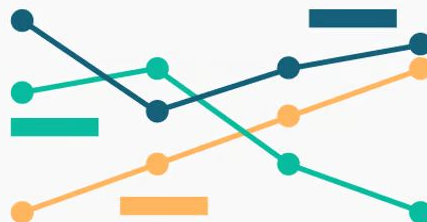
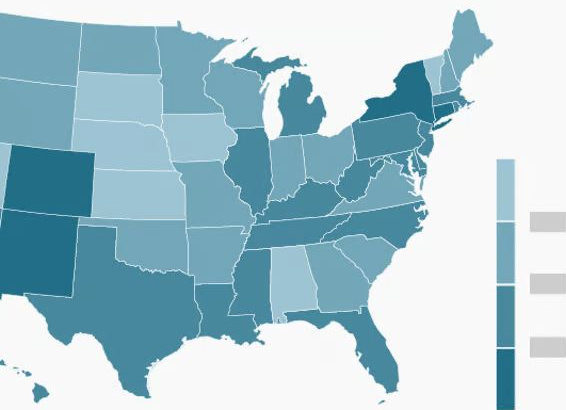
Smooth Line Graph



Use the right colors for the data

---

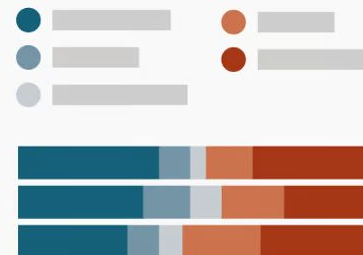
## SEQUENTIAL



## CATEGORICAL



## DIVERGING



### CATEGORICAL SCALE



Windows MacOS iOS All others

### ORDINAL SCALE



Very Low Low Neutral All others

### INTERVAL SCALE



1995 to 1999 2000 to 2004 2005 to 2009  
2010 to 2014 2015 to 2020

### RATIO SCALE



-10 to -5 -5 to 0 0 to 5 5 to 10  
10 to 15

### CATEGORICAL SCALE



Windows MacOS iOS Android All others

### ORDINAL SCALE



Low High

### INTERVAL SCALE



1995 2000 2005 2010 2015 2020

### RATIO SCALE



-10 -5 0 5 10 15

### CATEGORICAL SCALE



Windows MacOS iOS Android All others

### ORDINAL SCALE



Low Neutral High

### INTERVAL SCALE



1995 2000 2005 2010 2015 2020

### RATIO SCALE



-25 -15 -5 5 15 25

Categorical colors are not ordered. Use these for categorical scales. Do not use these for ordinal, interval, or ratio scales.

Sequential colors are ordered. Use these for ordinal and interval scales. It's also acceptable to use these for ratio scales. Do not use these for categorical scales.

Diverging colors are ordered. Use these for ordinal and ratio scales, especially when there is a meaningful middle value. These may also be used for interval scales. Do not use these for categorical scales.



Number of data classes: 3



[how to use](#) [updates](#) [downloads](#) [credits](#)

# COLORBREWER 2.0

color advice for cartography

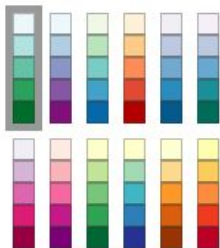
Nature of your data:



☒ sequential ☐ diverging ☐ qualitative

Pick a color scheme:

Multi-hue:



Single hue:



Only show:



- ☐ colorblind safe
- ☐ print friendly
- ☐ photocopy safe

Context:



- ☐ roads
- ☐ cities
- ☒ borders



Background:

- ☒ solid color
- ☐ terrain



color transparency

3-class BuGn



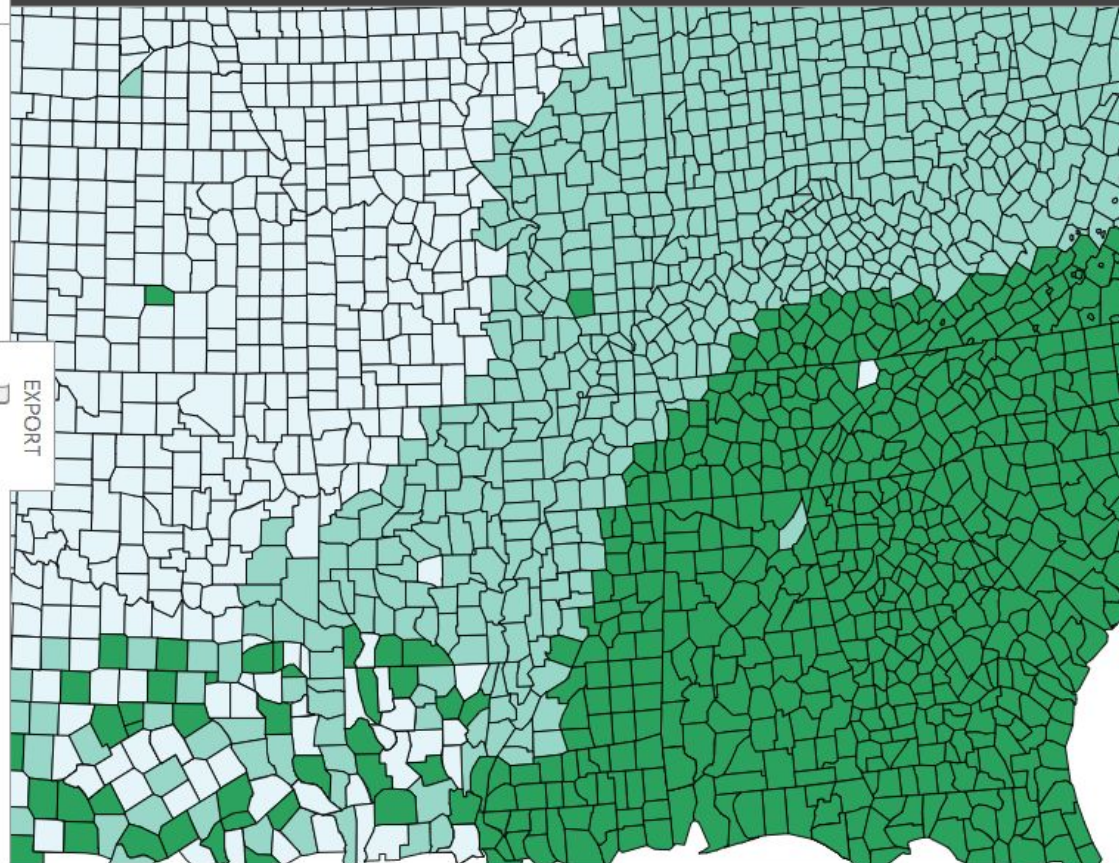
HEX

#e5f5f9

#99d8c9

#2ca25f

EXPORT





# Accessible color palette generator

Discover beautiful color combinations your whole audience can appreciate and follow Web Content Accessibility Guidelines (WCAG) with ease.

Random palette ↺

● #0073E6

Generate

The color pairings follow WCAG 2.1 AA based on a contrast ratio of 4.5:1. The pairings have sufficient contrast for use with normal text, large text and graphics.

## Vibrant palette

Download ↓

#00BF7D	#00B4C5	#0073E6	#2546F0	#5928ED
Contrast Black Text 8.75:1 ✓	Contrast Black Text 8.33:1 ✓	Contrast White Text 4.57:1 ✓	Contrast White Text 6.54:1 ✓	Contrast White Text 7.12:1 ✓

## Monochromatic palette

Download ↓

#B3C7F7	#8BABF1	#0073E6	#0461CF	#054FB9
Contrast Black Text 12.43:1 ✓	Contrast Black Text 9.19:1 ✓	Contrast White Text 4.57:1 ✓	Contrast White Text 5.81:1 ✓	Contrast White Text 7.43:1 ✓

## Contrasting palette 1

Download ↓

#C44601	#F57600	#8BABF1	#0073E6	#054FB9
Contrast White Text 4.97:1 ✓	Contrast Black Text 7.47:1 ✓	Contrast Black Text 9.19:1 ✓	Contrast White Text 4.57:1 ✓	Contrast White Text 7.43:1 ✓

## Contrasting palette 2

<https://venngage.com/tools/accessible-color-palette-generator>

Download ↓

# Declarative visualization

---

# Data visualization

Don't think about drawing the picture, think about what you want to be drawn.

*Encode* data values to visual elements

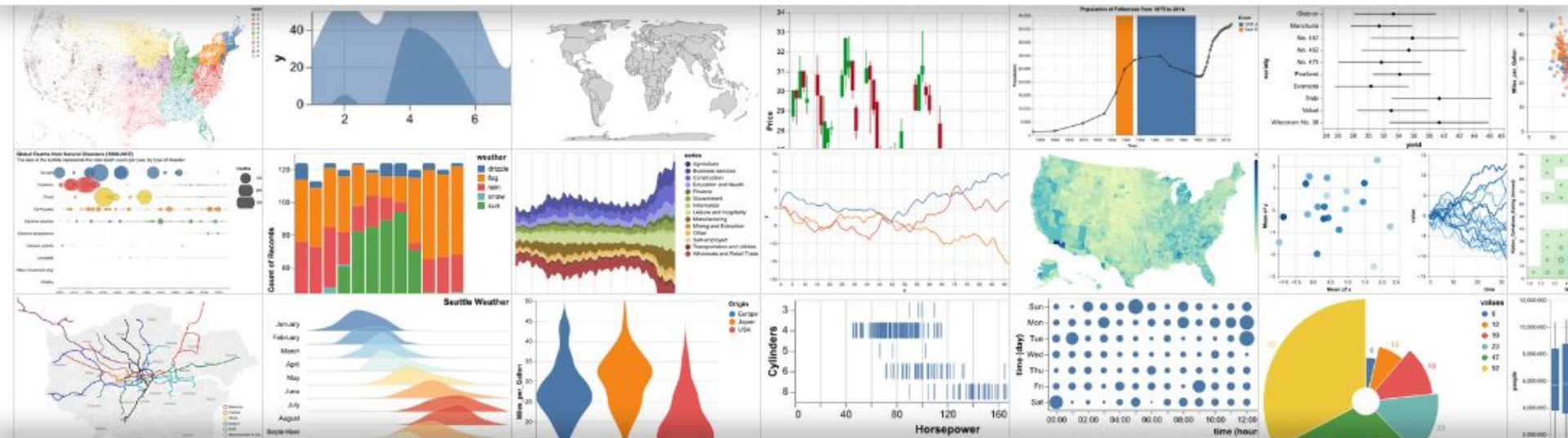
- $x$ -,  $y$ -axes
- shapes
- color
- location

# Altair

---



# Vega-Altair: Declarative Visualization in Python



Data

+ marks

+ encoding



# Wide form vs. long form data

---



# Wide-form vs long-form (*tidy*) data

*Wide* data:

→ Each column is a different property

*Long* or *tidy* data:

→ Each row is an single observation

→ Each column is a property or value

Some visualization packages (like altair) expect long-form data;  
some (like Datawrapper) expect wide-form

# Wide-form vs long-form (*tidy*) data

## Wide

	Date	Tmax	Tmin	Tmean
0	1948-01-01	8.9	3.3	6.10
1	1948-02-01	7.9	2.2	5.05
2	1948-03-01	14.2	3.8	9.00
3	1948-04-01	15.4	5.1	10.25
4	1948-05-01	18.1	6.9	12.50
...	...	...	...	...
908	2023-09-01	24.4	14.7	19.55
909	2023-10-01	18.1	10.2	14.15
910	2023-11-01	11.8	5.4	8.60
911	2023-12-01	10.9	5.9	8.40
912	2024-01-01	8.4	2.5	5.45

## Long

	Date	Category	Temperature
0	1948-01-01	Tmax	8.90
1	1948-01-01	Tmin	3.30
2	1948-01-01	Tmean	6.10
3	1948-02-01	Tmax	7.90
4	1948-02-01	Tmin	2.20
...	...	...	...
2734	2023-12-01	Tmax	10.90
2735	2023-12-01	Tmin	5.90
2736	2024-01-01	Tmin	2.50
2737	2024-01-01	Tmax	8.40
2738	2024-01-01	Tmean	5.45