

Data Visualization

**DATA VISUALIZATION IS
COMMUNICATION**

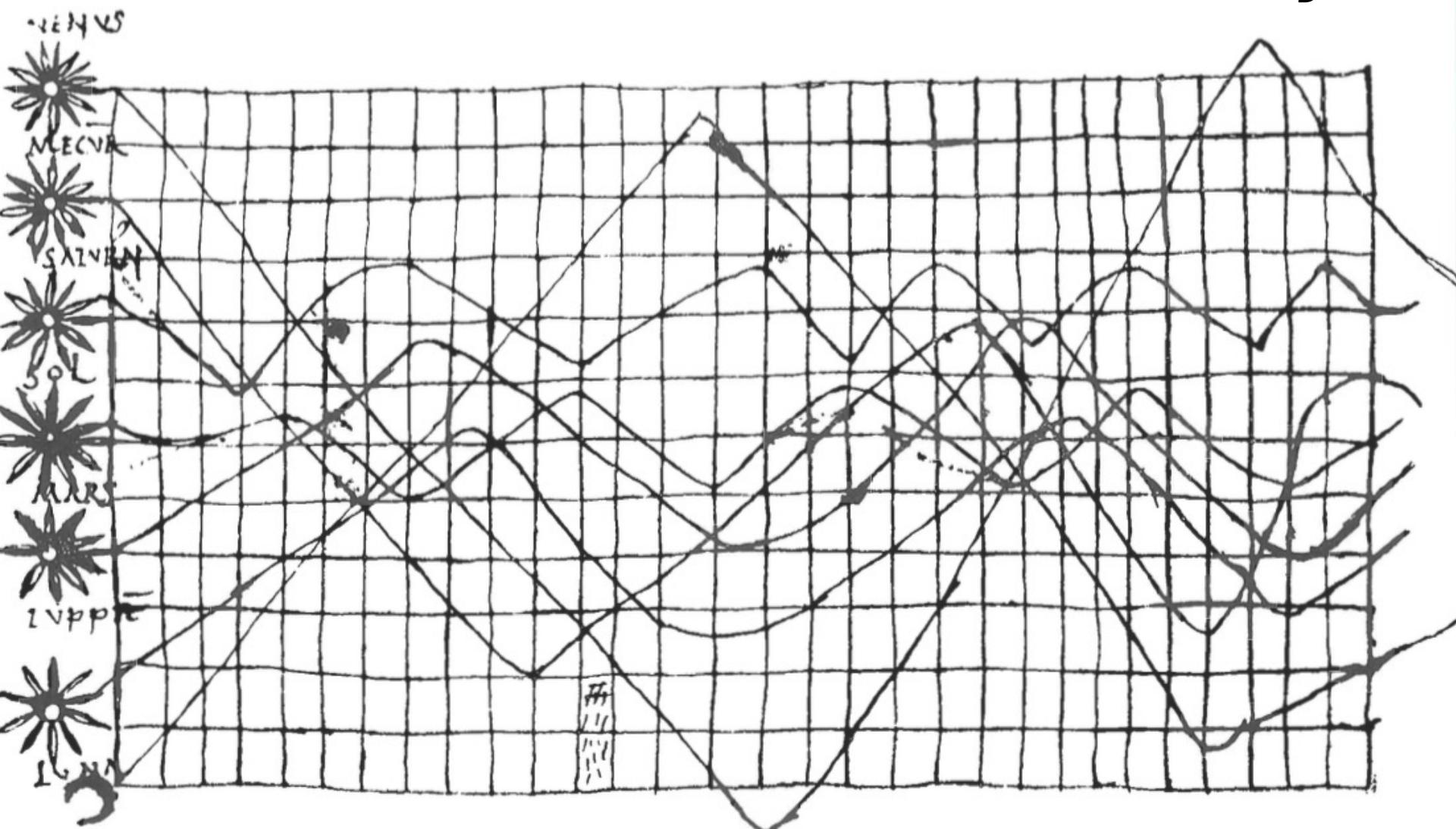
Principles of Graphical Excellence

- Clarity
- Precision
- Efficiency
- Maximize ideas, minimize ink

Data Viz in a Nutshell

1. History
2. Graphical Basics
3. Minimalist principles

Data Viz in the 10th Century



Ecce formulam, vsum, atque

structuram Tabularum Ptolomati, cum quibusdam locis, in
quibus studiosus Geographia se satis exercere potest.

SEPTENTRIO. pars superior.



OCCIDENTIS.
Sinistra manus.

	28	29	30	31	32	33	34	35	36	37	
52						Lyptzim				52	
51						Erfordre	Leyflich			51	
50						Nurenberch	Praga			50	
49						Ingolstade				49	
48						Monachii	Munichia	Vienne et Austrice		48	
47						Bauari				47	
46								Venize		46	
	28	29	30	31	32	33	34	35	36	37	

CIRCE.
Dextra manus.

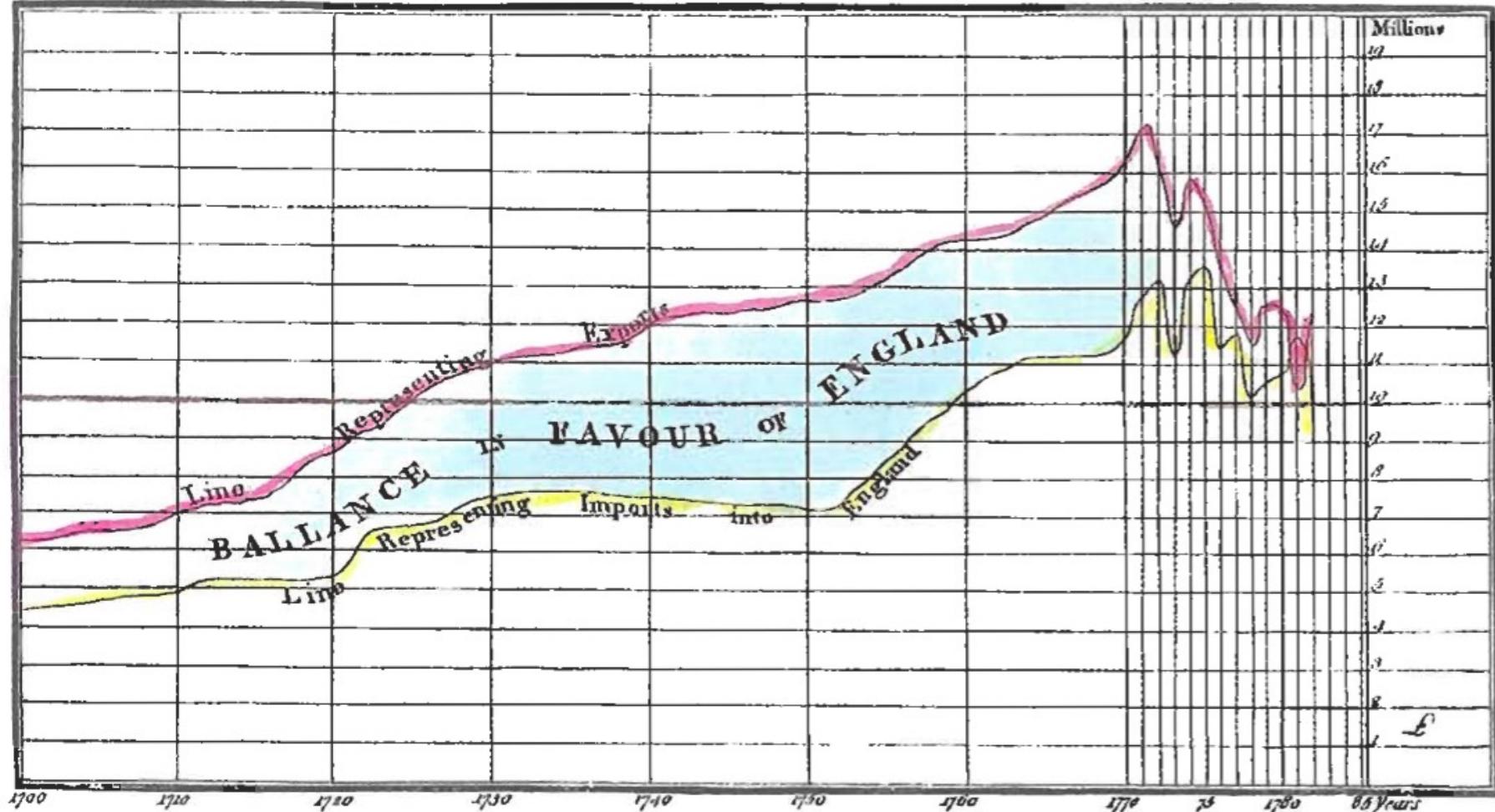
pars inferior.
MERIDIES.

Maps as Rich Information Source





CHART of all the IMPORTS and EXPORTS to and from ENGLAND
From the Year 1700 to 1782 by W. Playfair

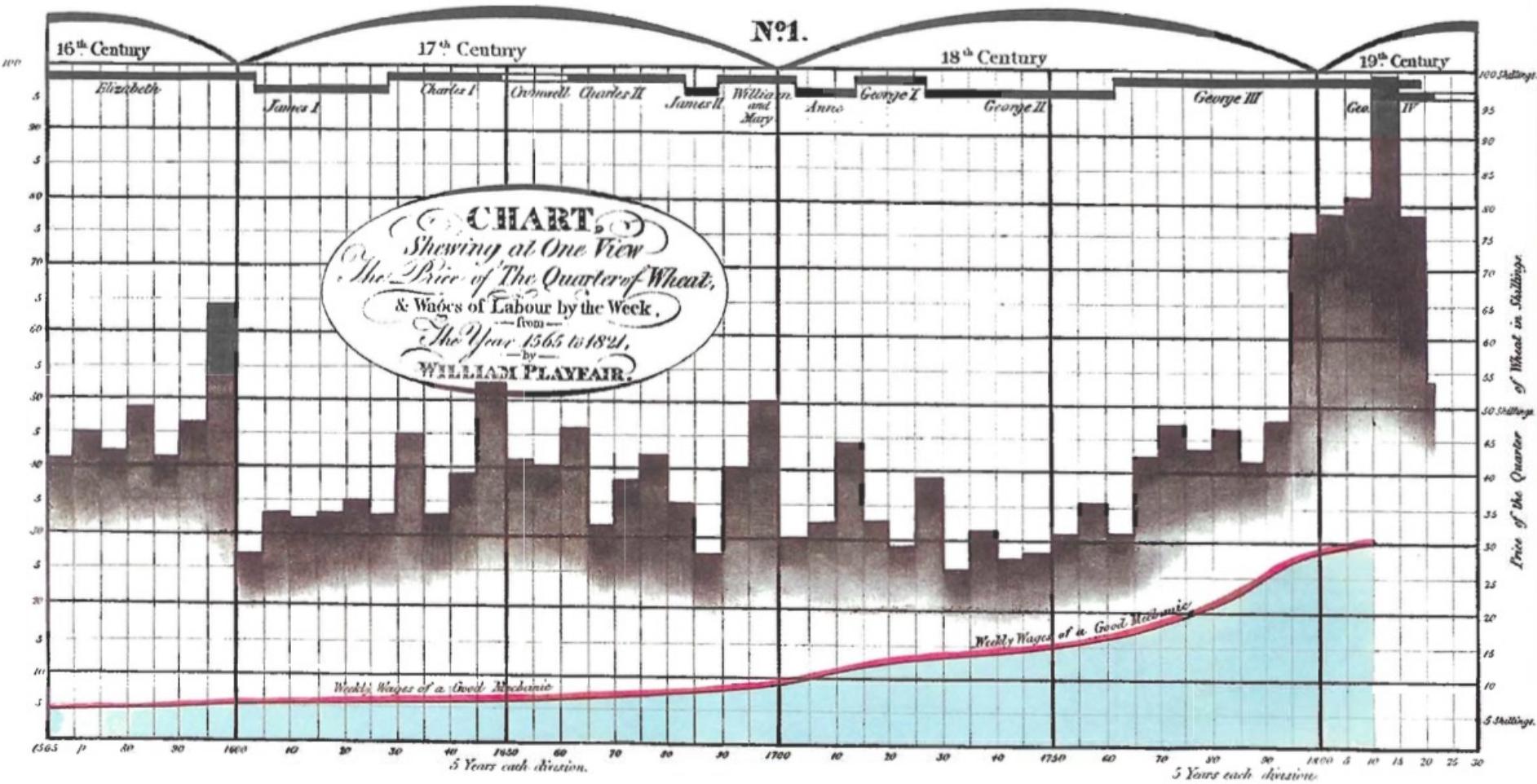


The Divisions at the Bottom, express Years, & those on the Right-hand Millions of Pounds

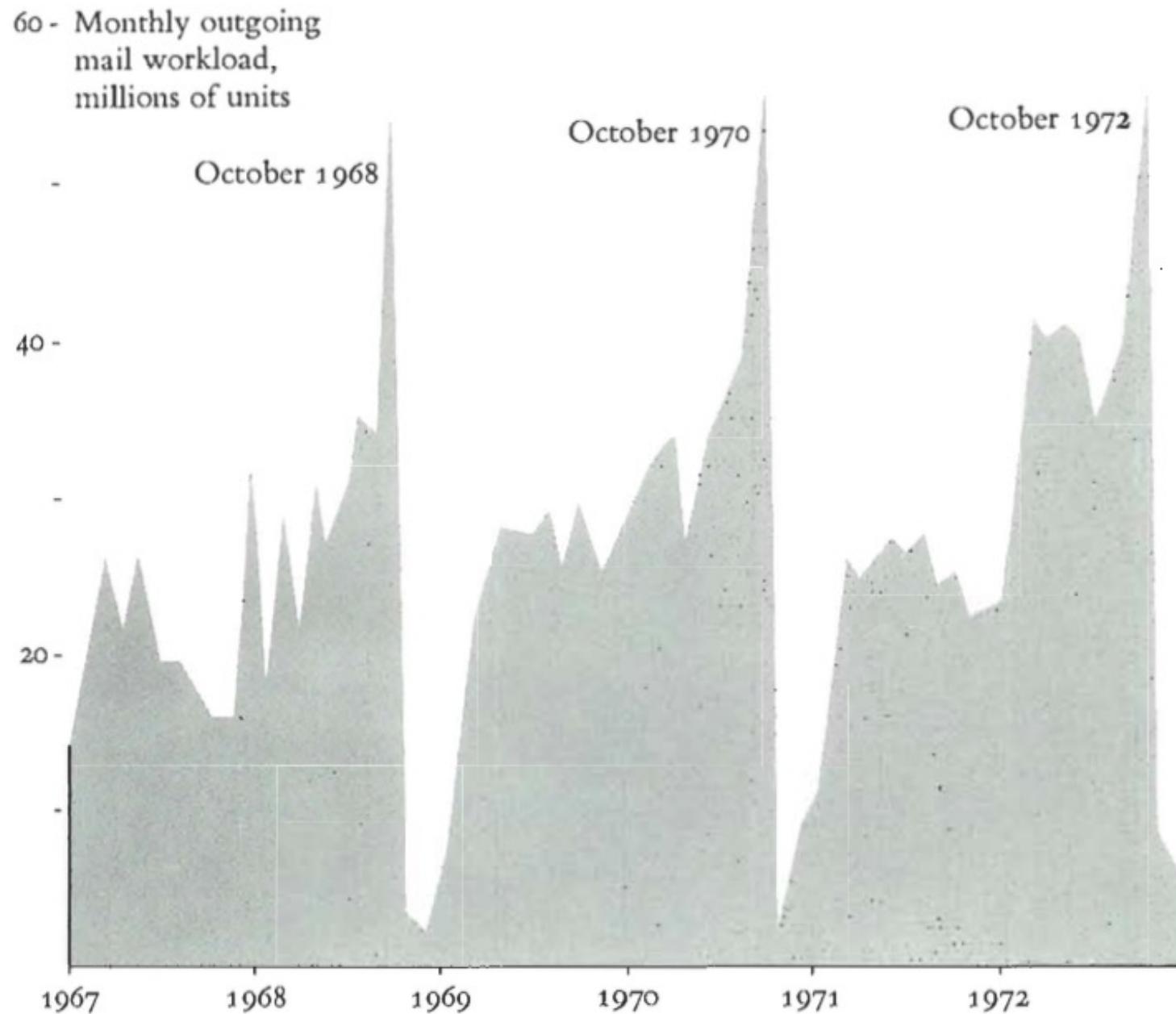
J. André Sculp't.

Published as the Act directs, 20th Aug^r. 1785

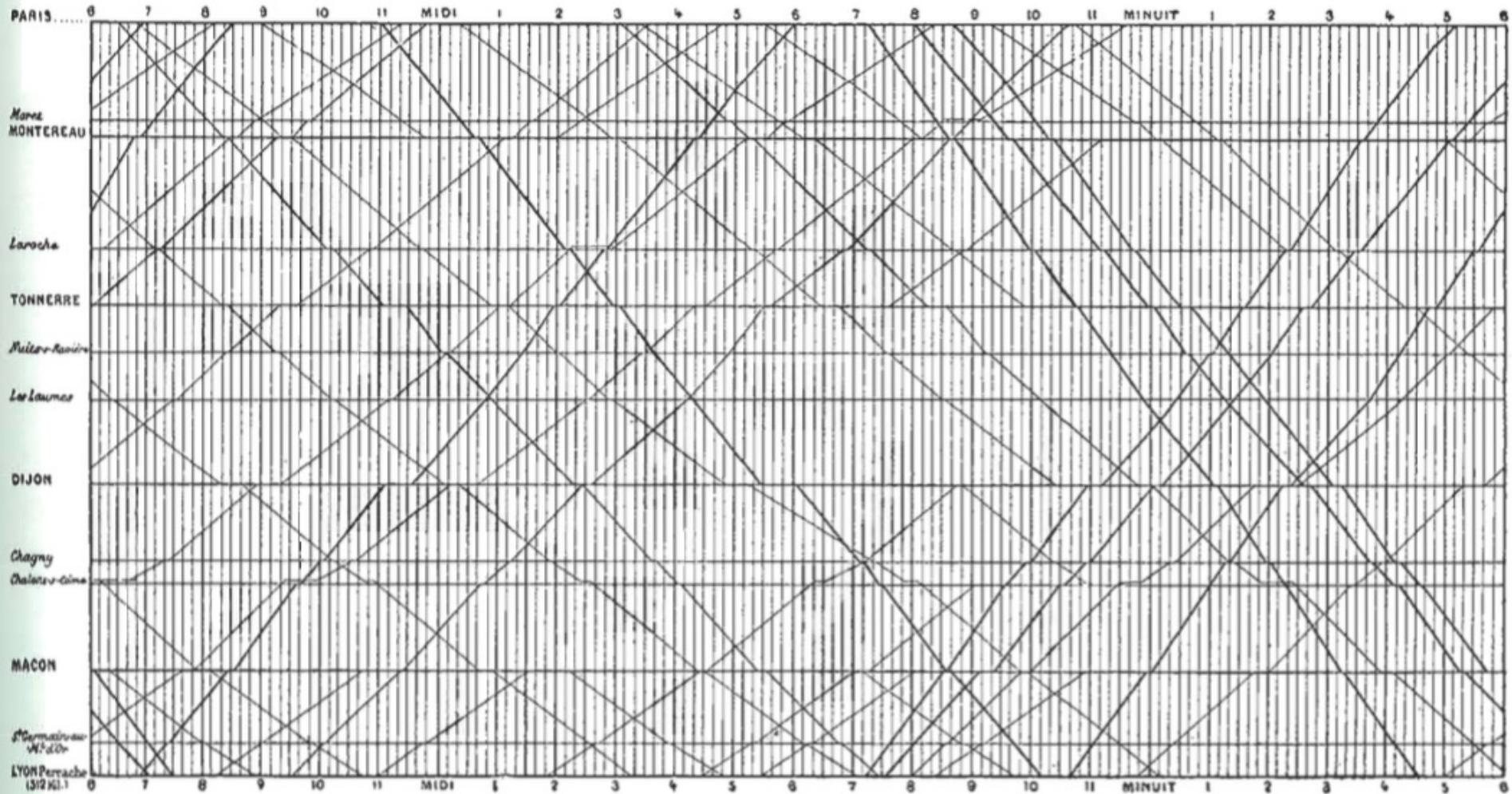
Multiple Data Sources to Make a Point



Sometimes One Variable is Enough



Adding Space to Time



Unifying Maps and Timeseries

CARTE FIGURATIVE des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite.

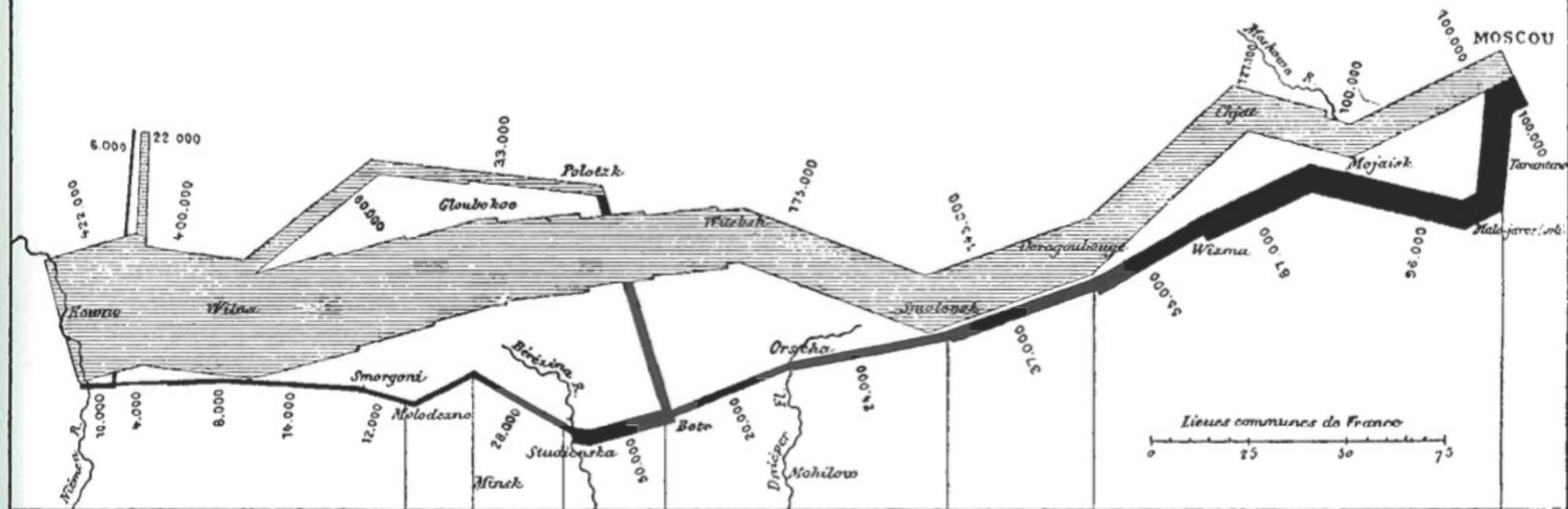
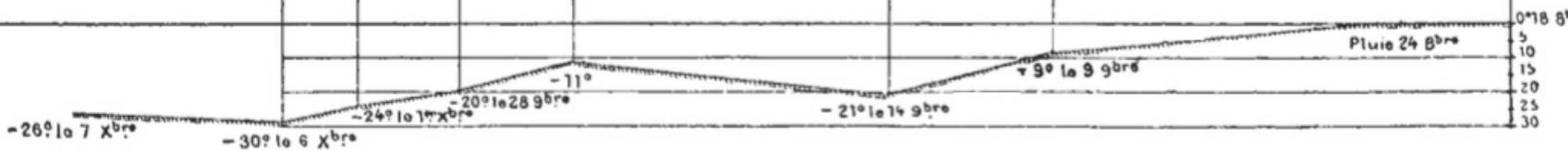


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro

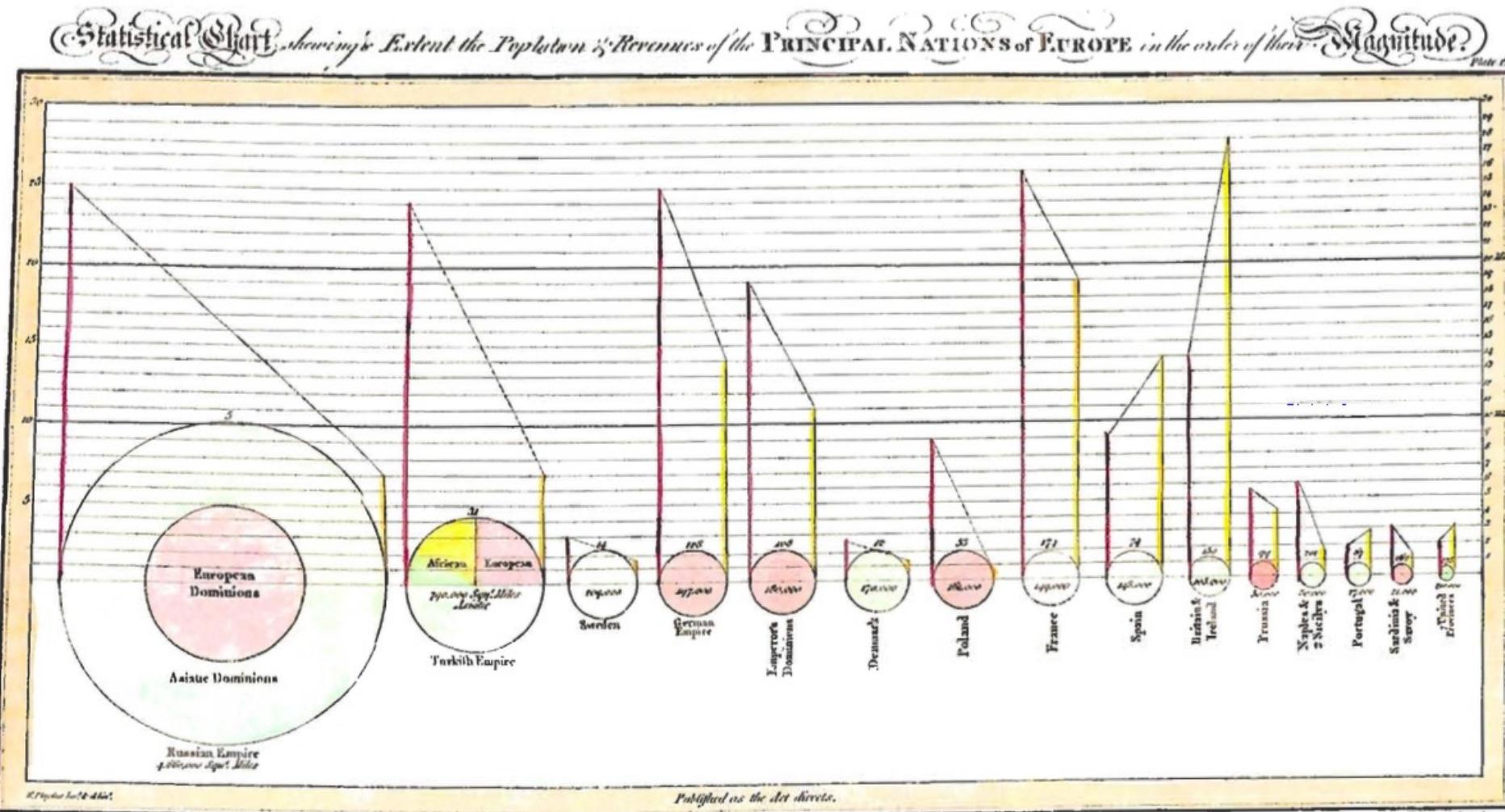


X^{bre} = December

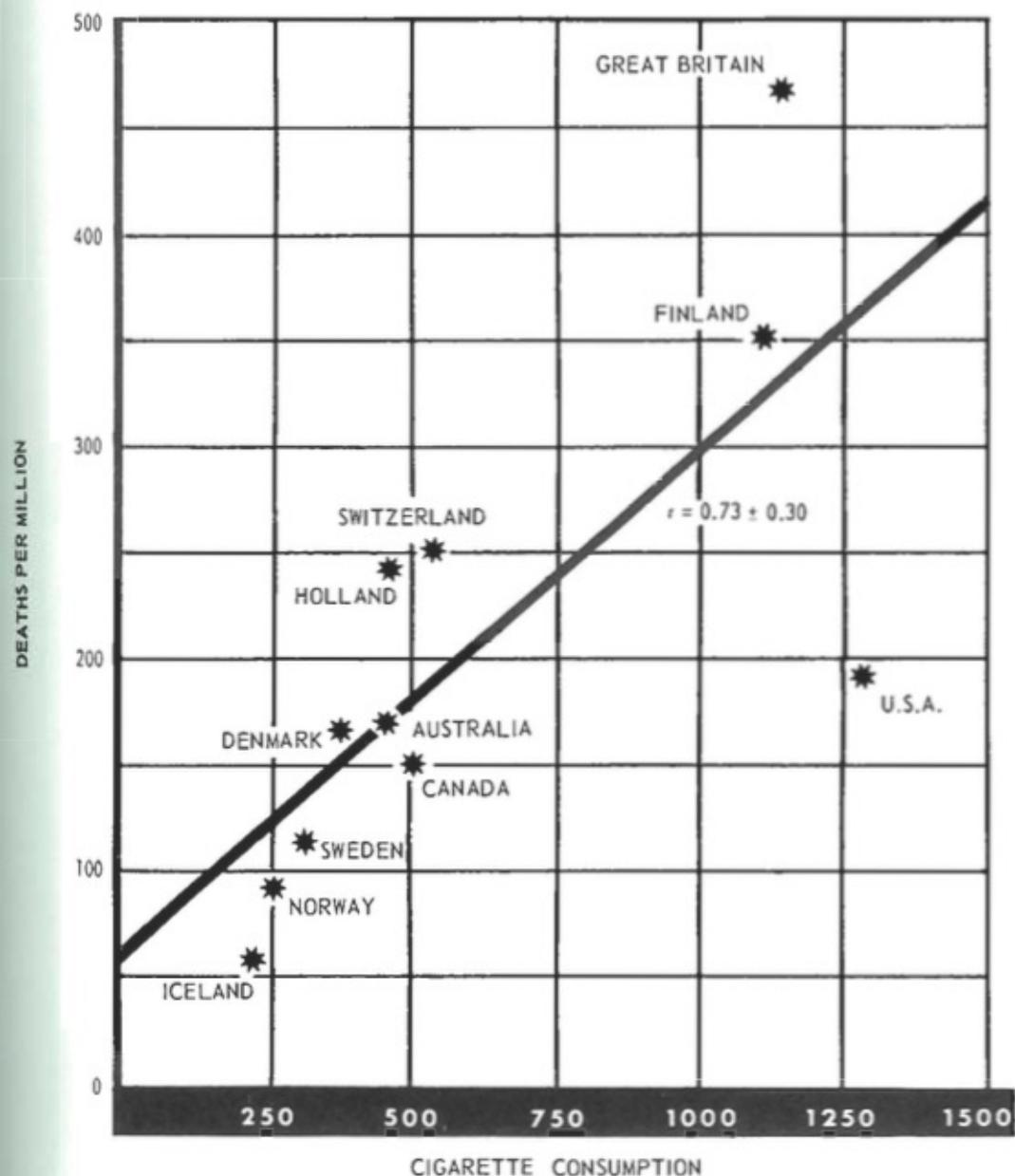
9bre = November

gbre = October

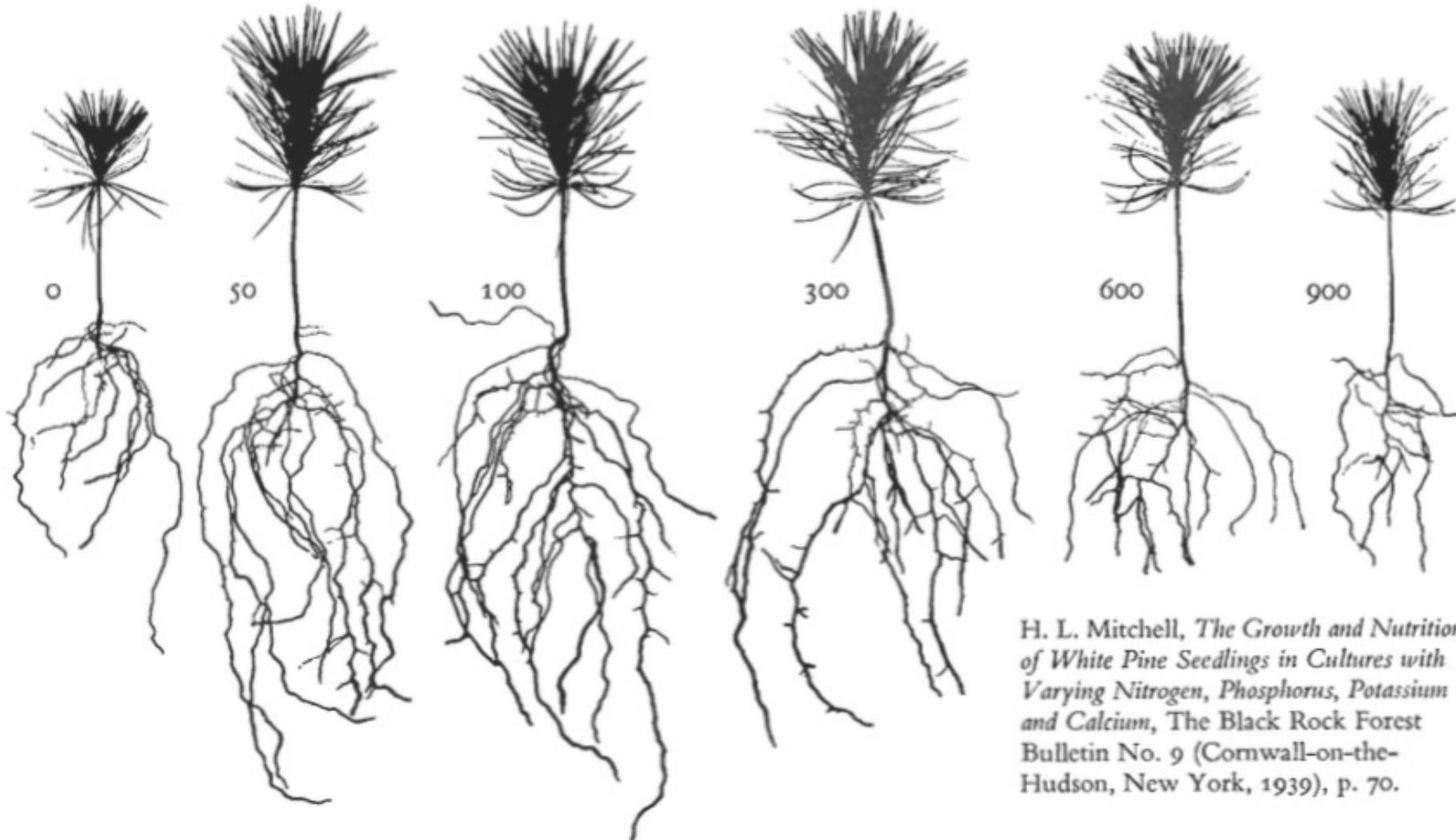
Abstracting Variables to Make a Point



CRUDE MALE DEATH RATE FOR LUNG CANCER
IN 1950 AND PER CAPITA CONSUMPTION OF
CIGARETTES IN 1930 IN VARIOUS COUNTRIES.



Abstraction Can Tell Wonderful Stories



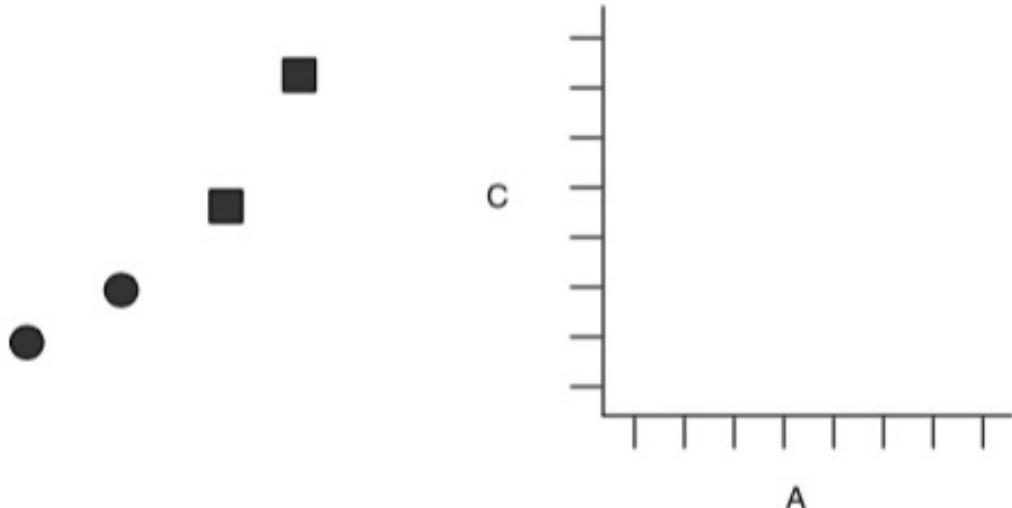
H. L. Mitchell, *The Growth and Nutrition of White Pine Seedlings in Cultures with Varying Nitrogen, Phosphorus, Potassium and Calcium*, The Black Rock Forest Bulletin No. 9 (Cornwall-on-the-Hudson, New York, 1939), p. 70.

Data Viz in a Nutshell

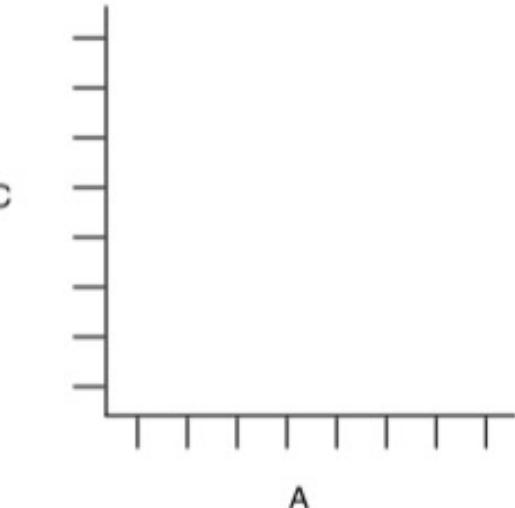
1. History
2. Graphical Basics
3. Minimalist principles

Elements of a Plot

Geometric Objects



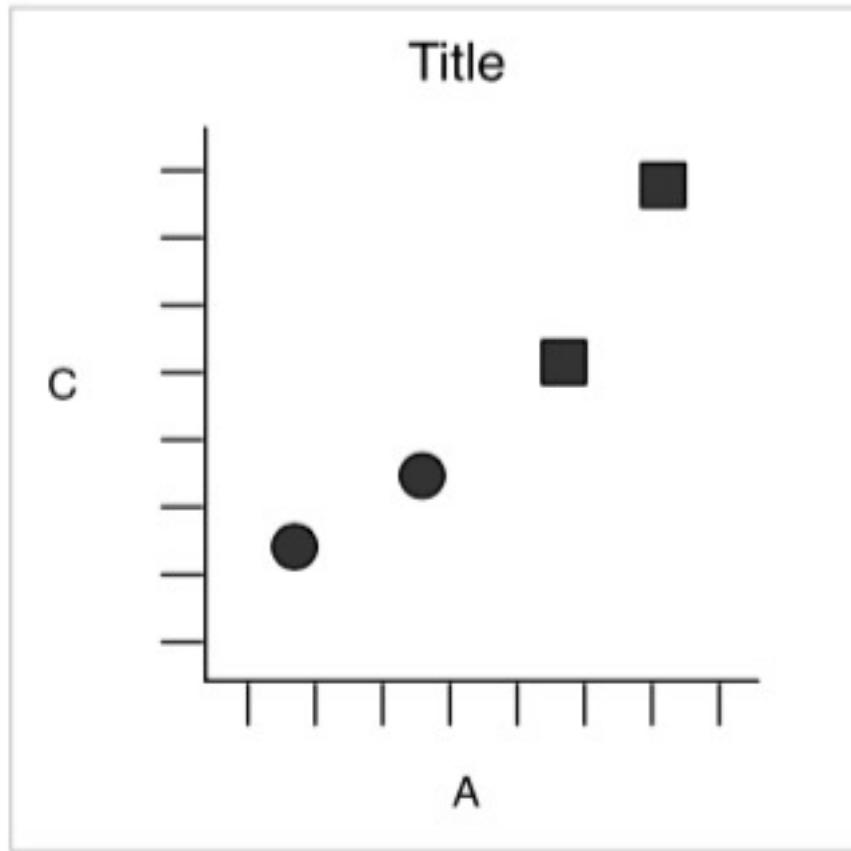
Scales & Coordinates



Annotations

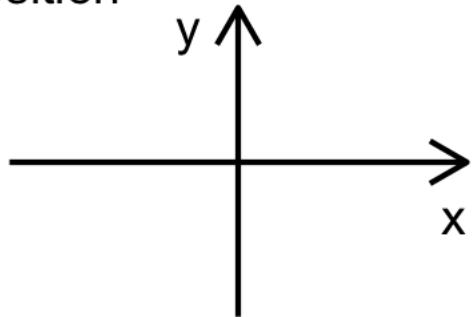


Elements of a Plot

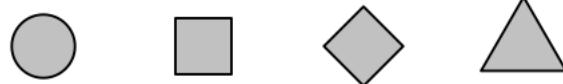


Aesthetics of a Plot

position



shape



size



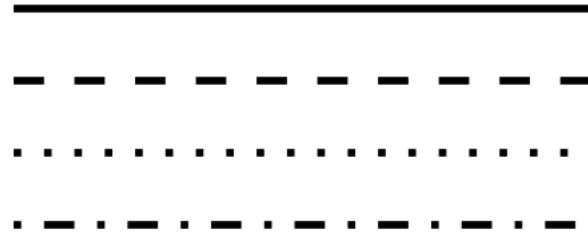
color



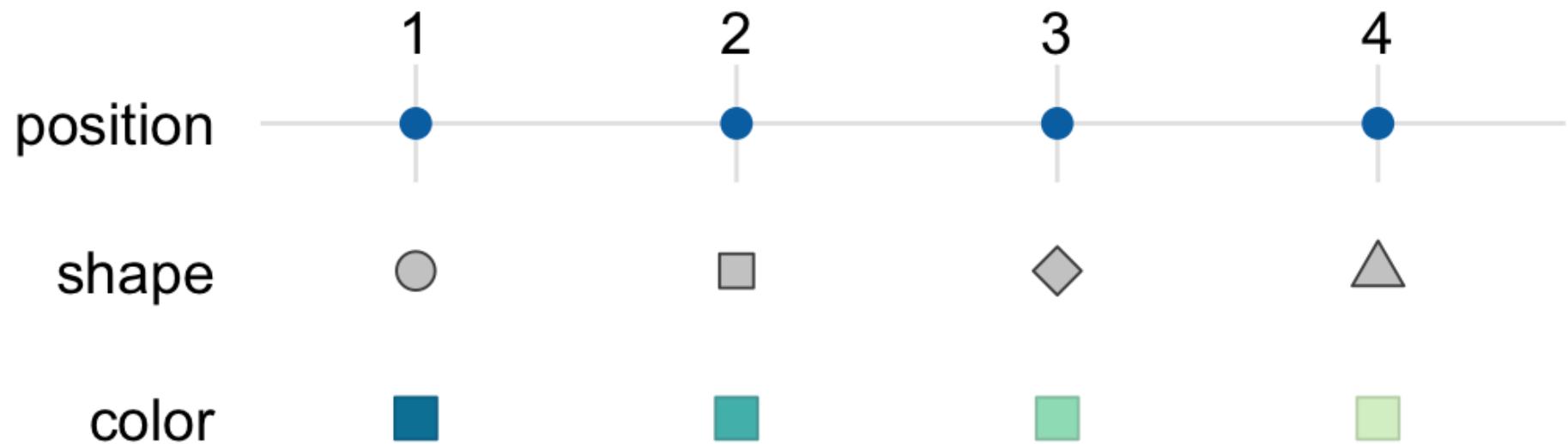
line width



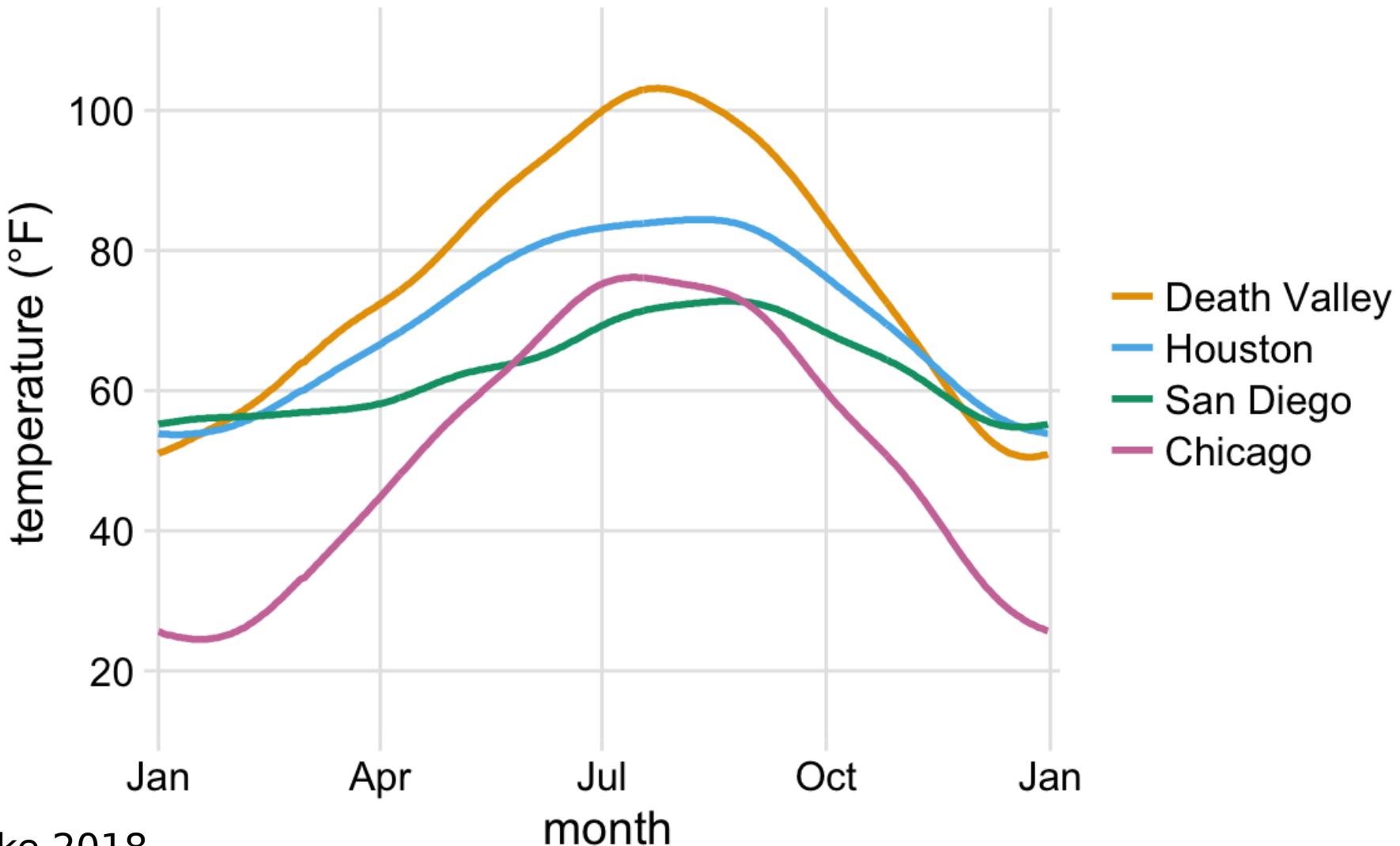
line type



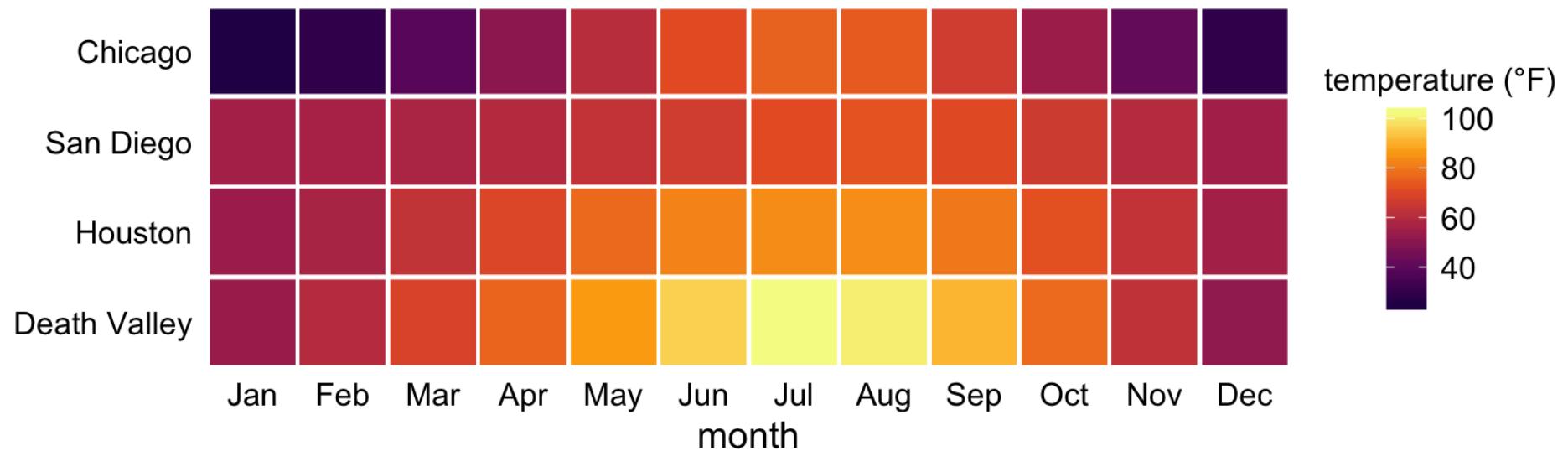
Aesthetics Map Data to Visual Representation



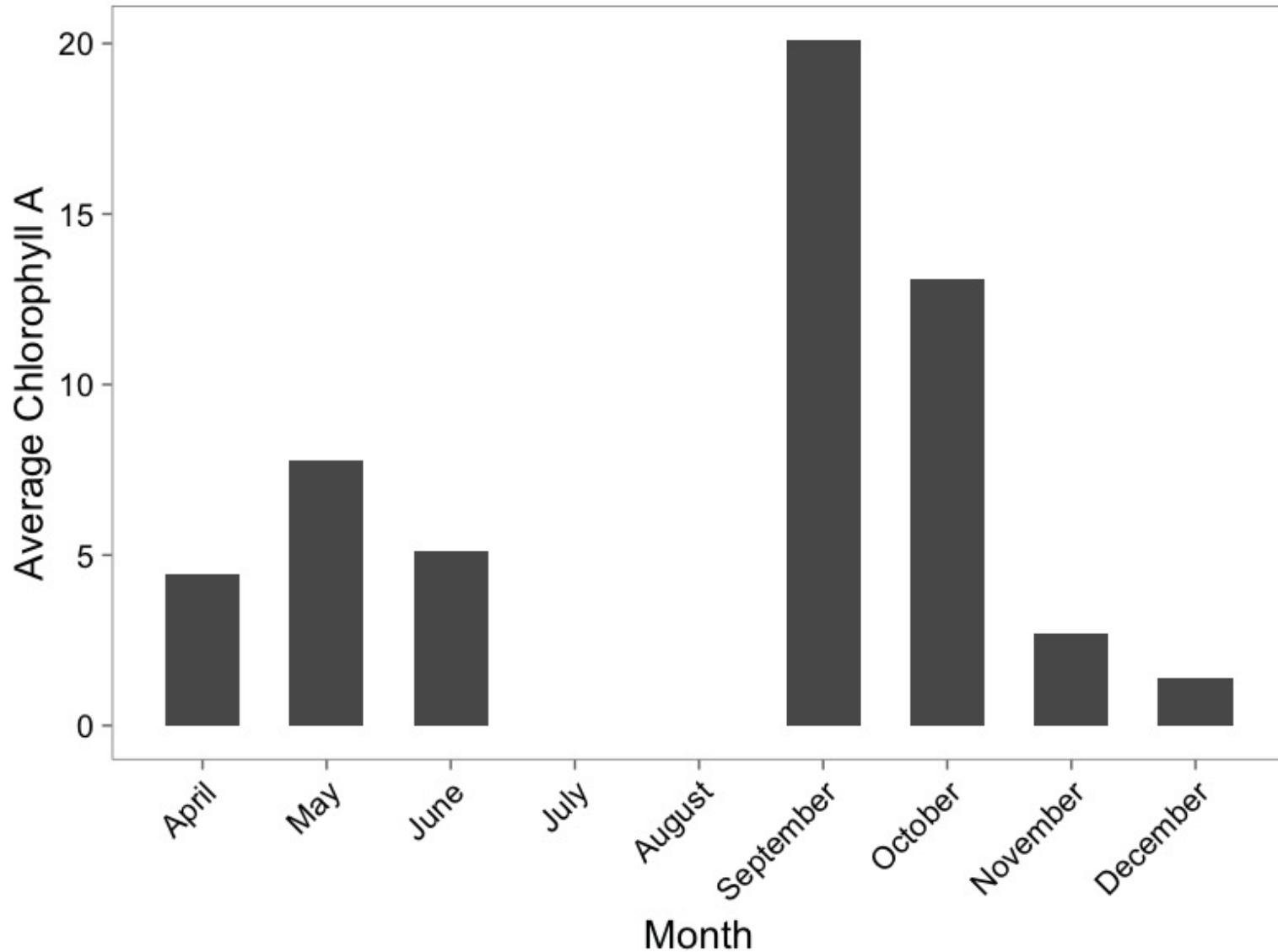
Aesthetics Map Data to Visual Representation



Aesthetics Map Data to Visual Representation

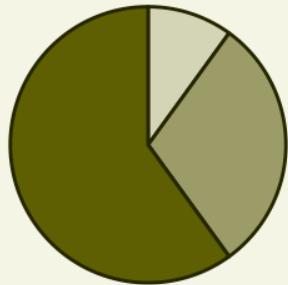


Barplots

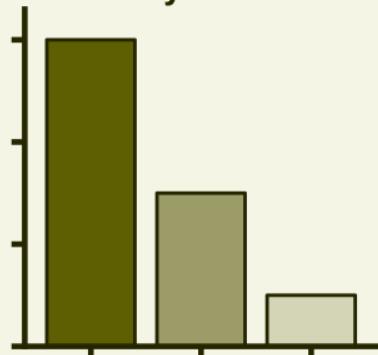


Types of Visualizations: Proportions

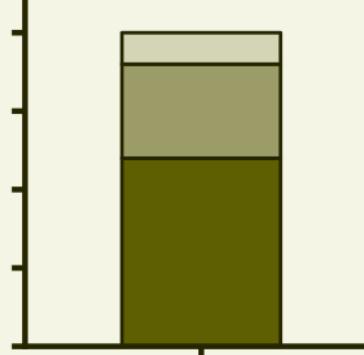
Pie chart



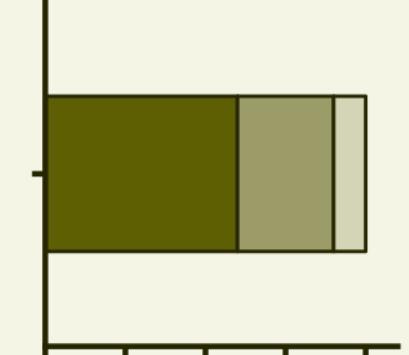
Side-by-side bars



Stacked bars



Stacked bars

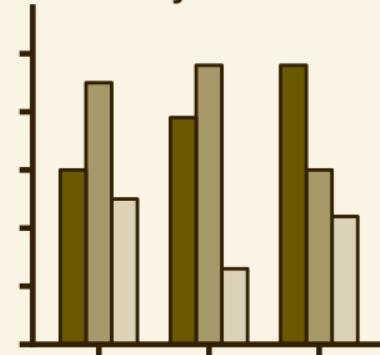


Types of Visualizations: Multiple Proportions

Multiple pie charts



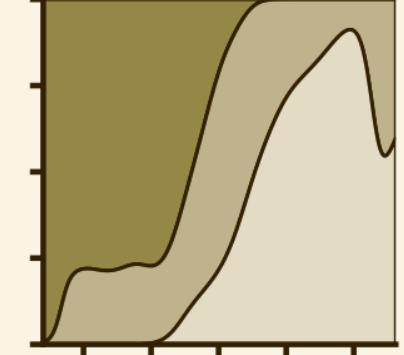
Side-by-side bars



Stacked bars

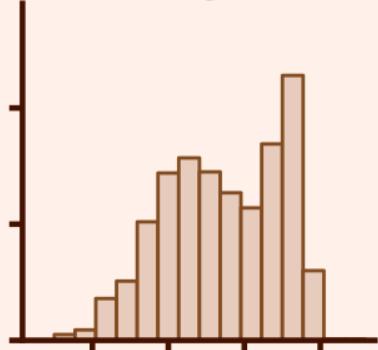


Stacked densities

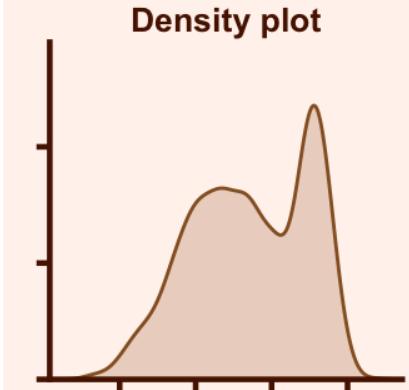


Types of Visualizations: Distributions

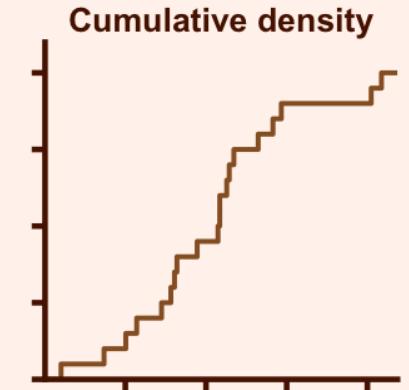
Histogram



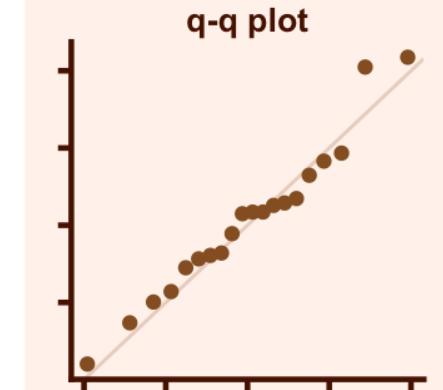
Density plot



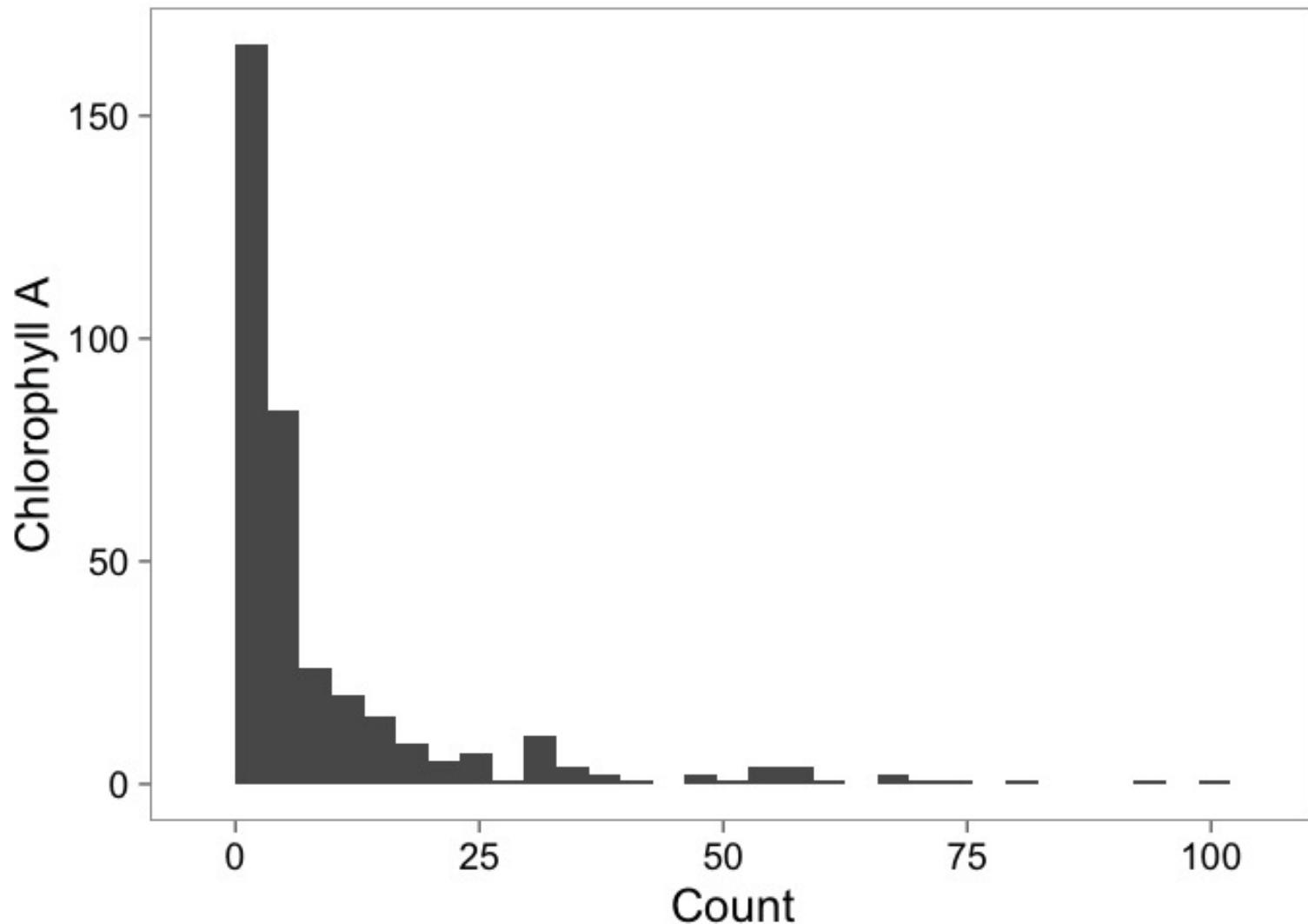
Cumulative density



q-q plot

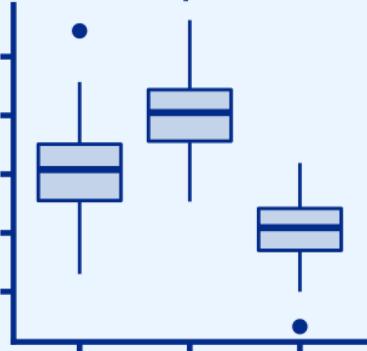


Histograms Show Frequency

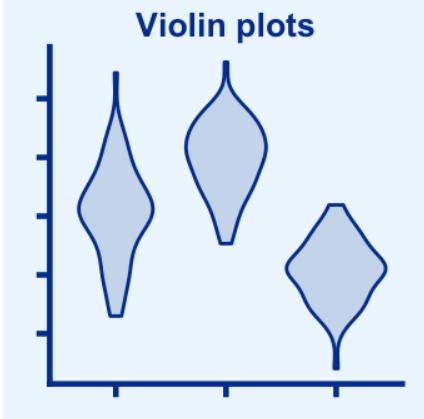


Types of Visualizations: Multiple Distributions

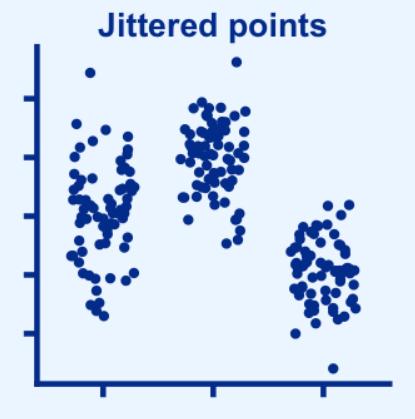
Boxplots



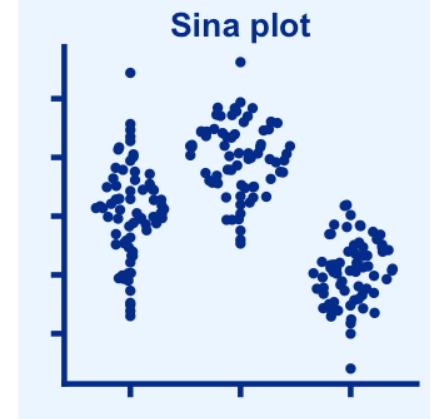
Violin plots



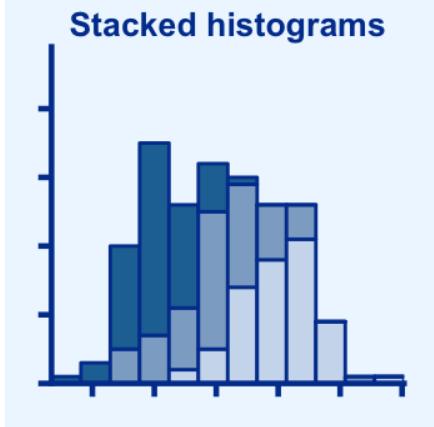
Jittered points



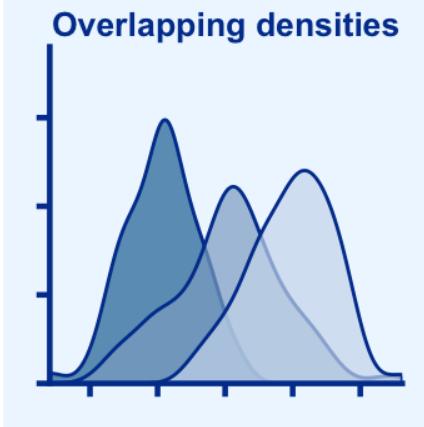
Sina plot



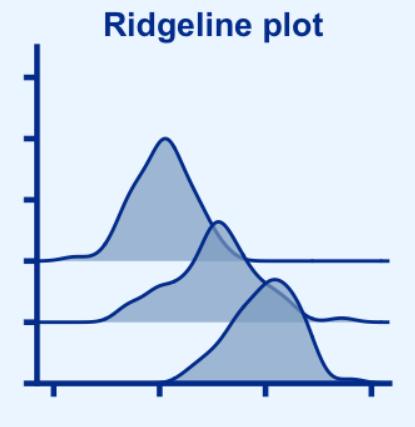
Stacked histograms



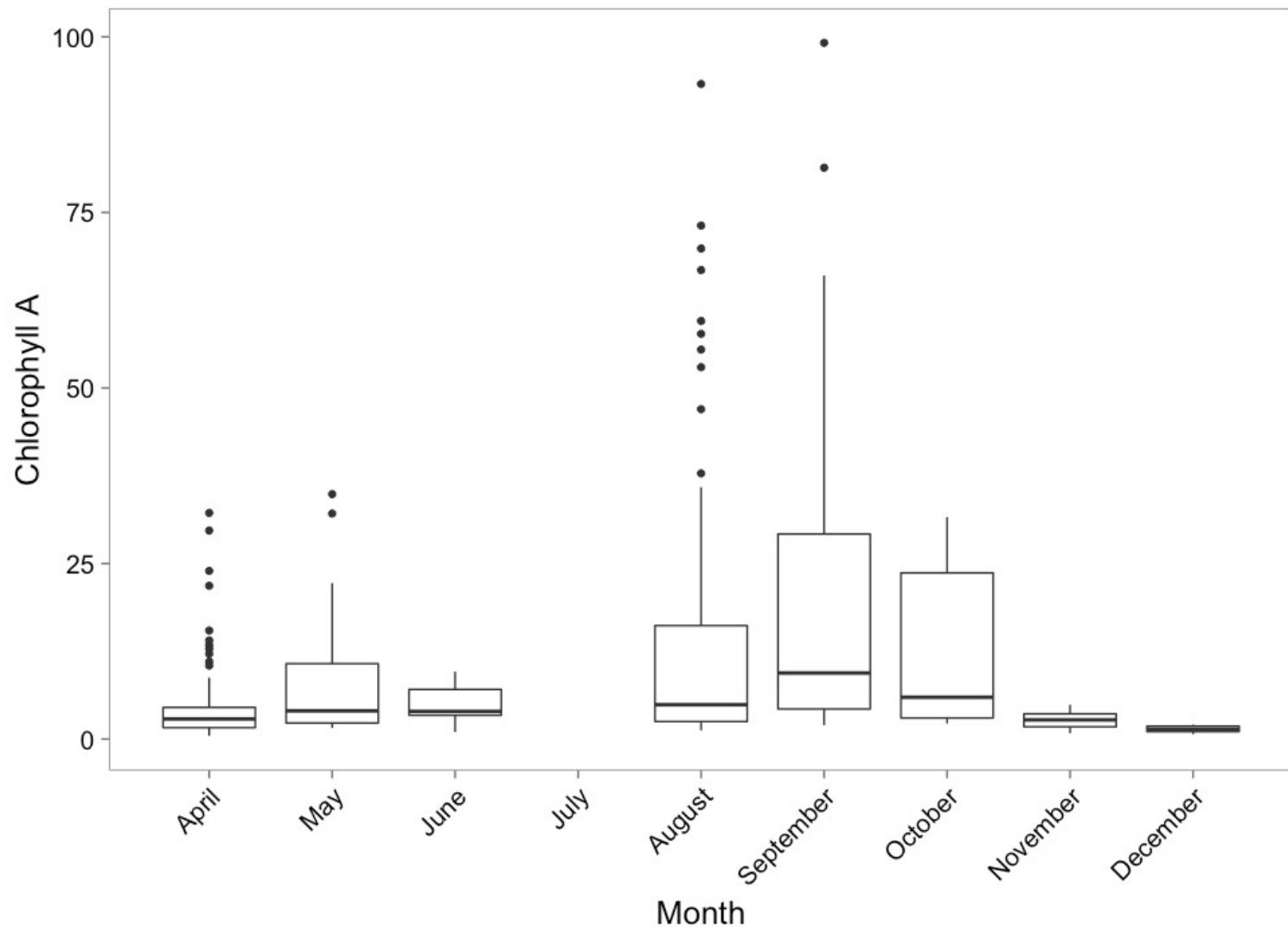
Overlapping densities



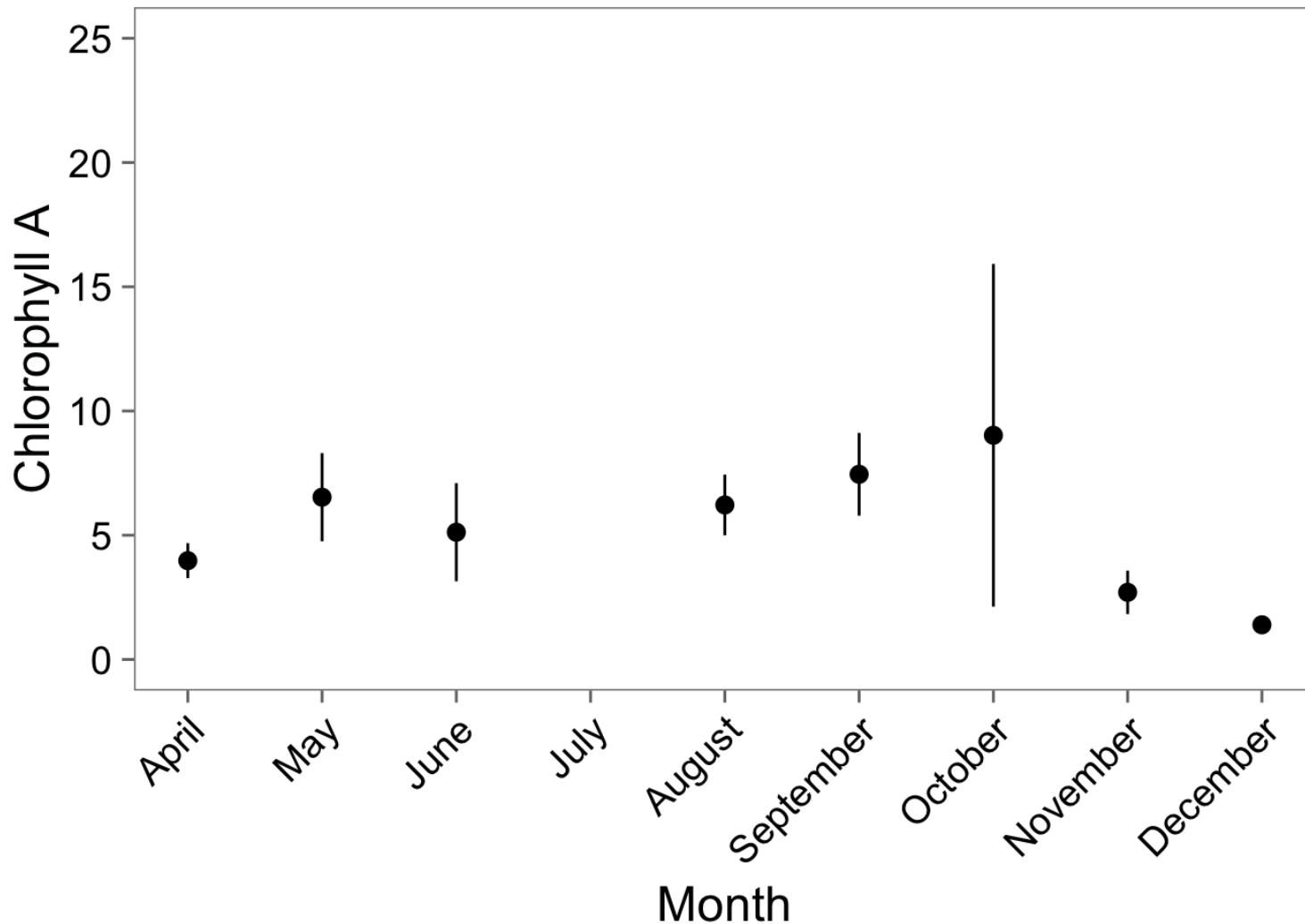
Ridgeline plot



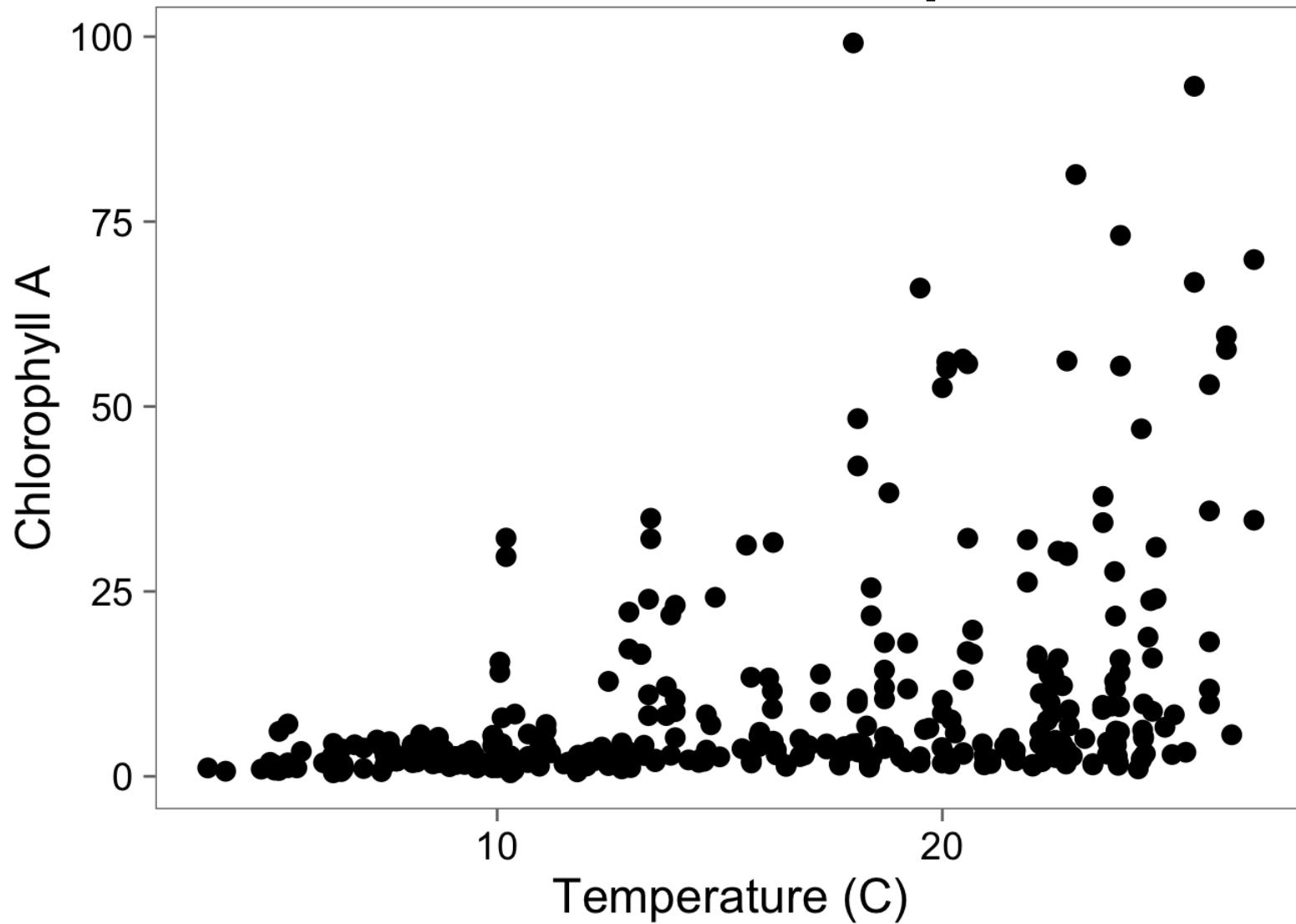
Boxplots to Show Variation



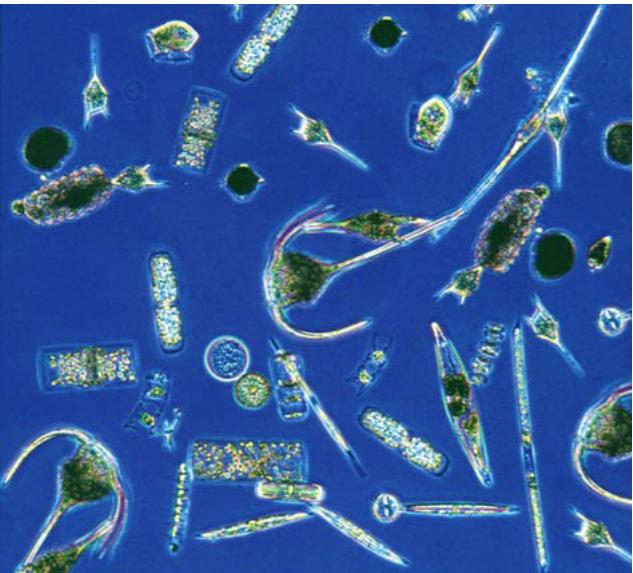
...Or Point-Ranges



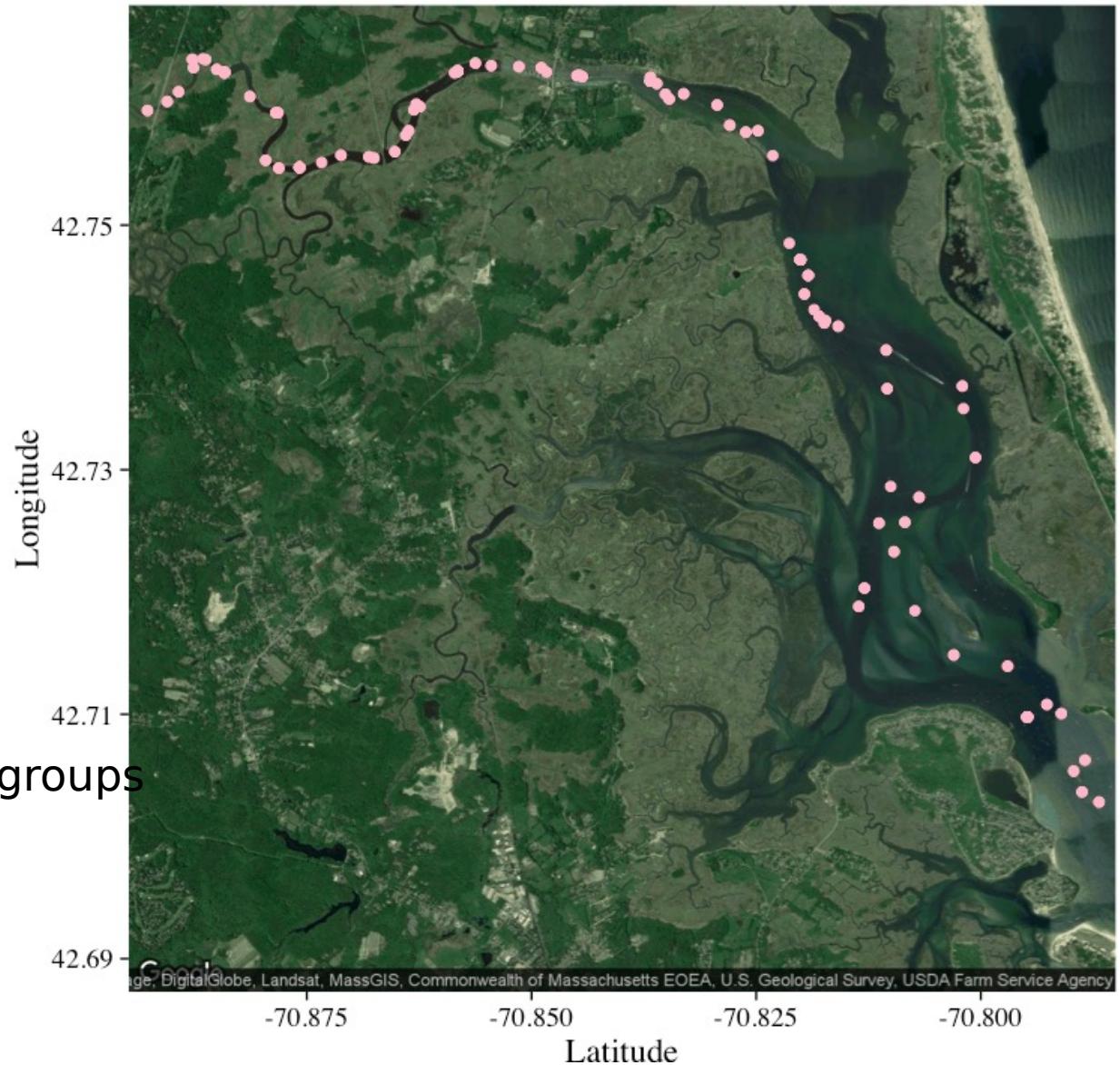
Scatterplots Show Relationships



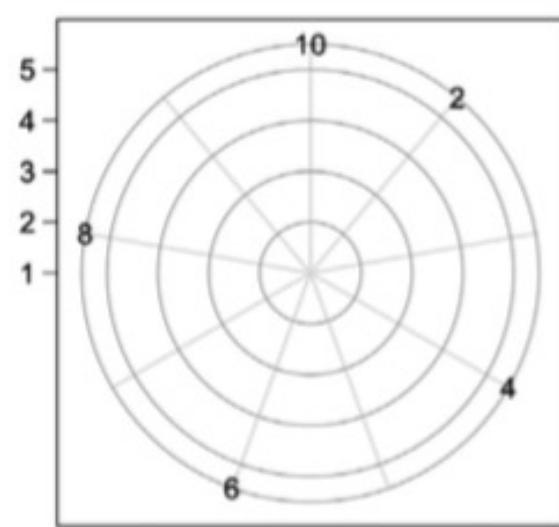
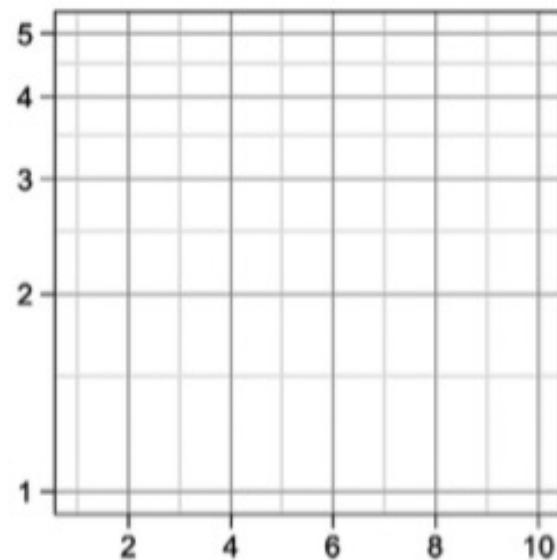
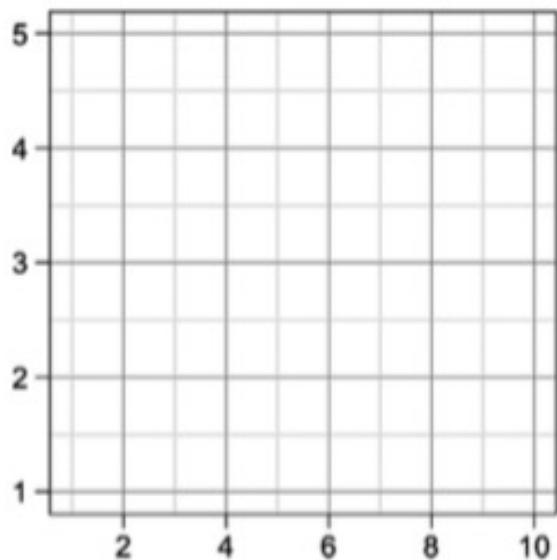
Combining Data Sources and Maps



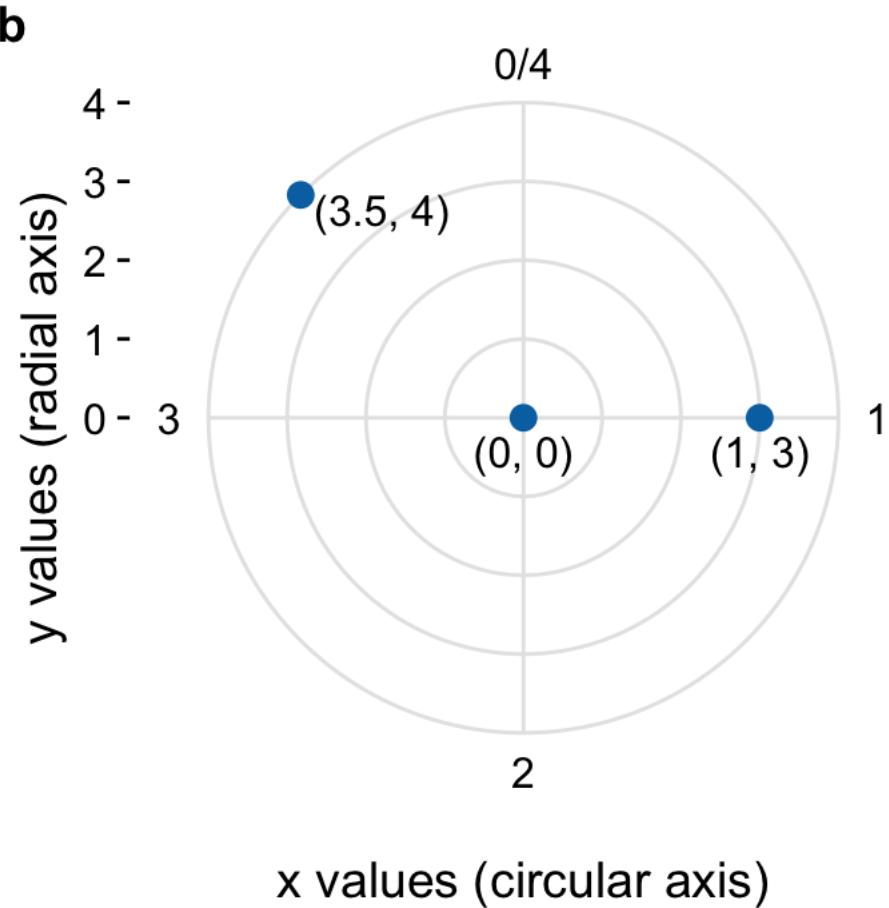
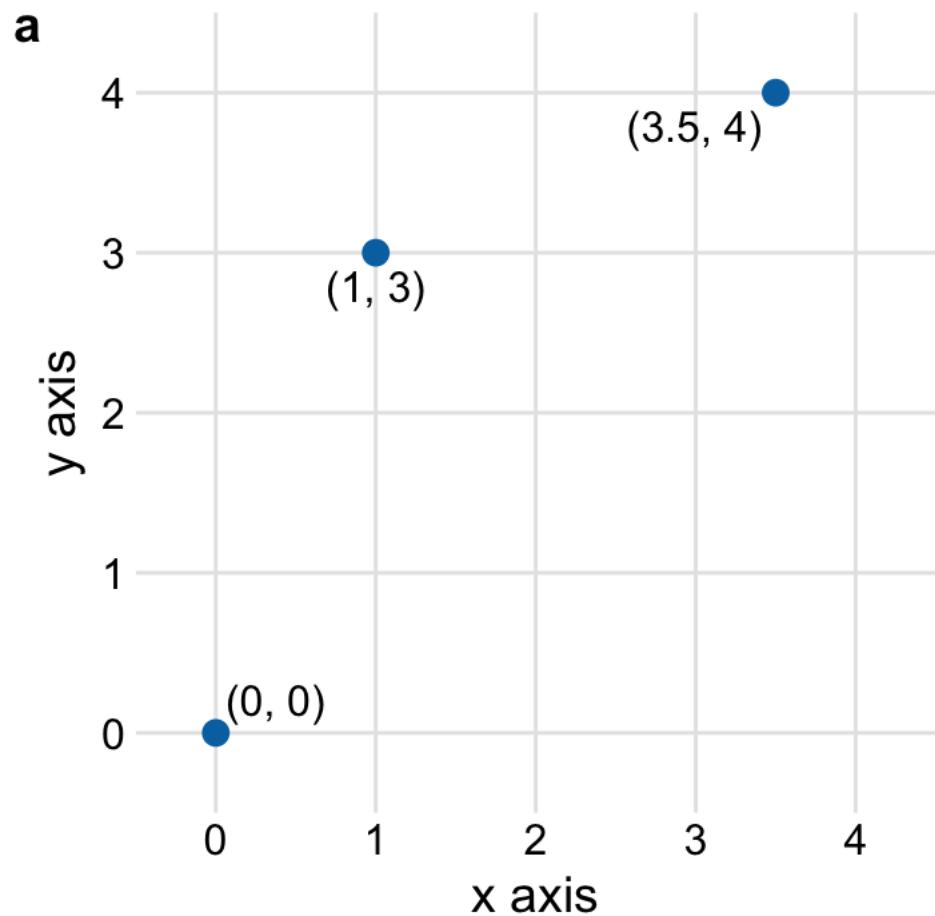
Chlorophyll a
Abundance of taxonomic groups
Temperature
Salinity



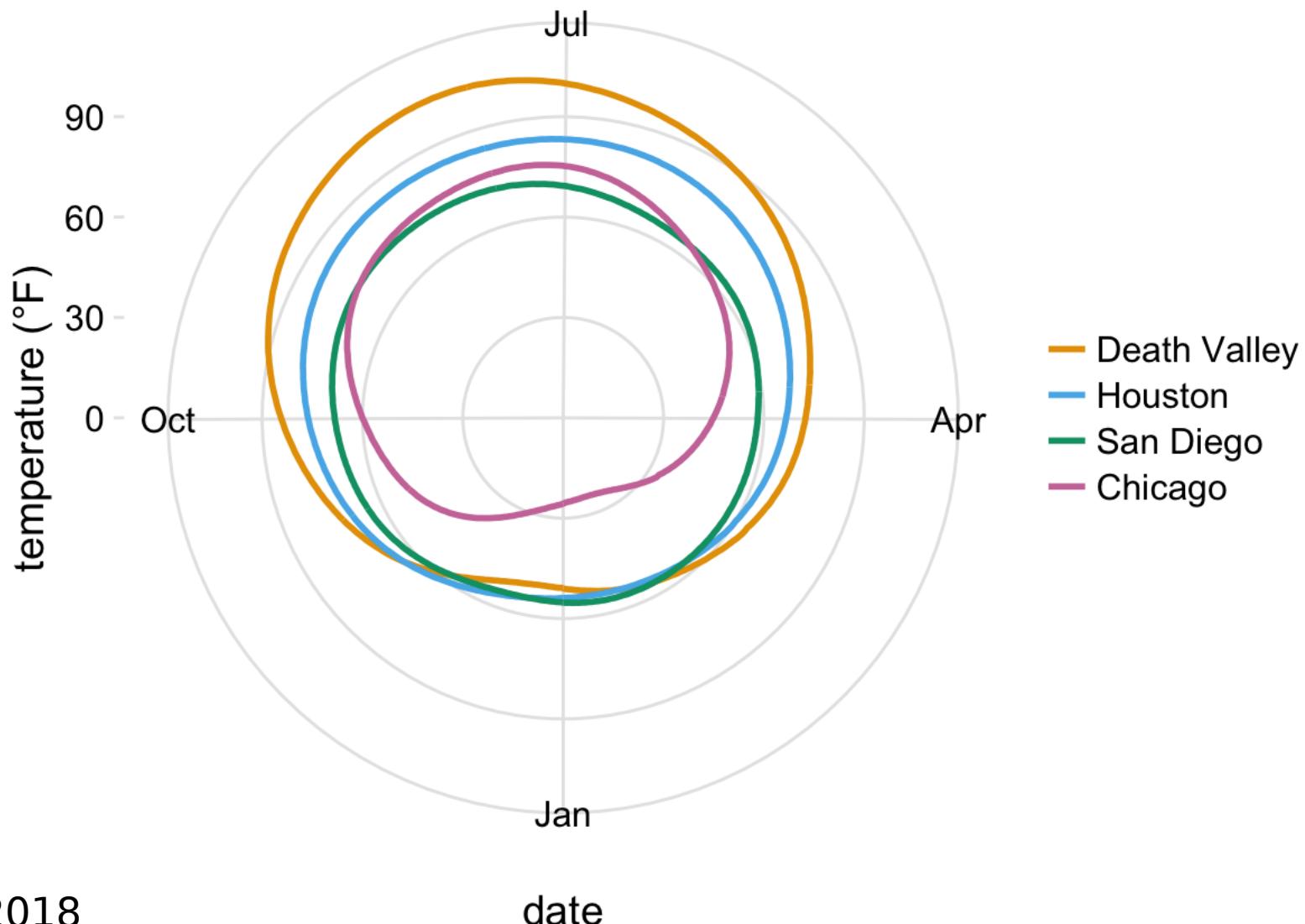
Coordinate Systems Transform Relationships



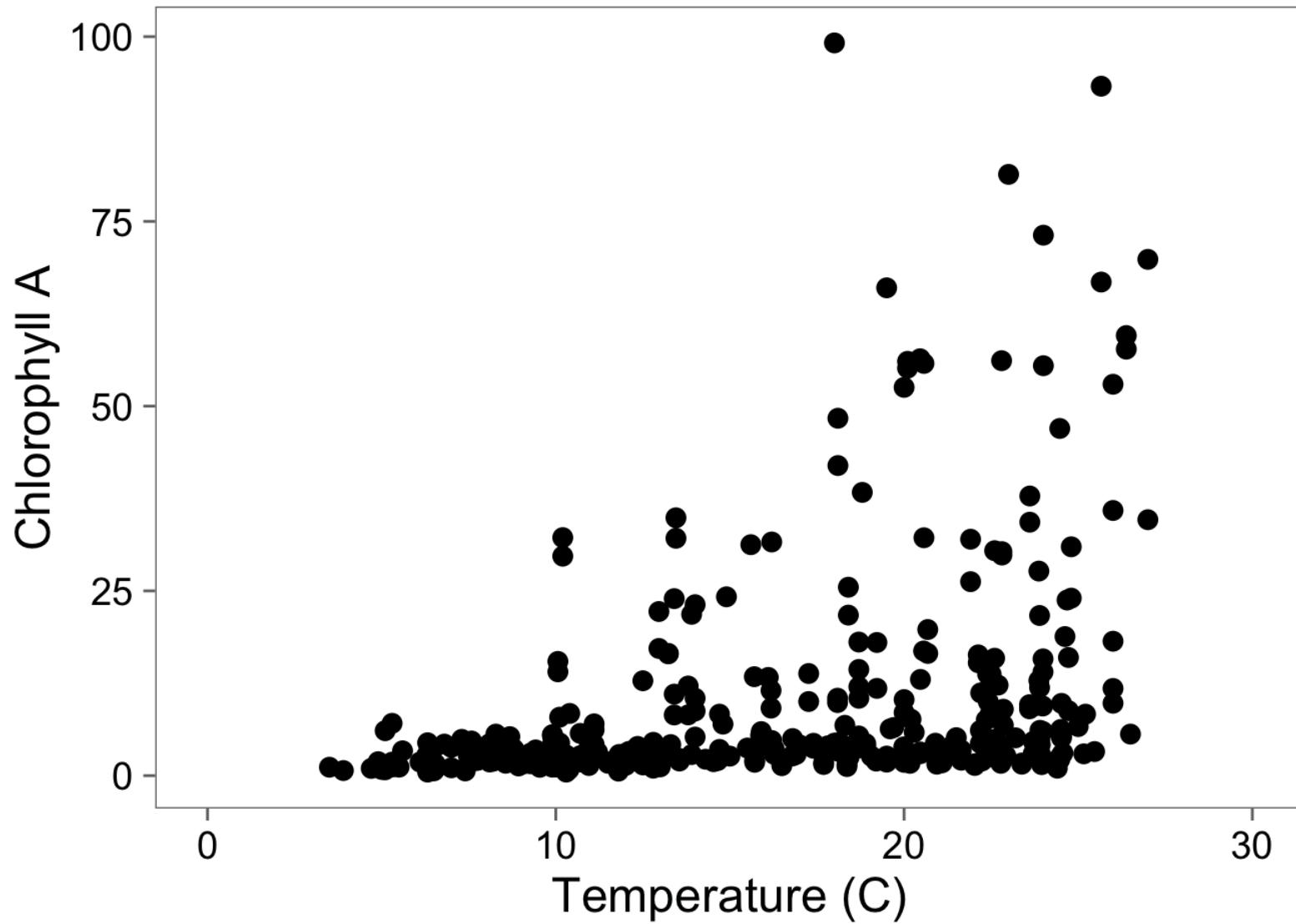
The Polar Coordinate System is Useful!



The Polar Coordinate System is Useful!

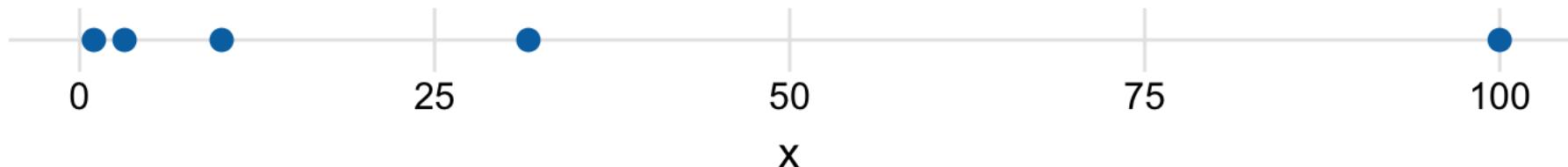


Adding Full Scale to 0

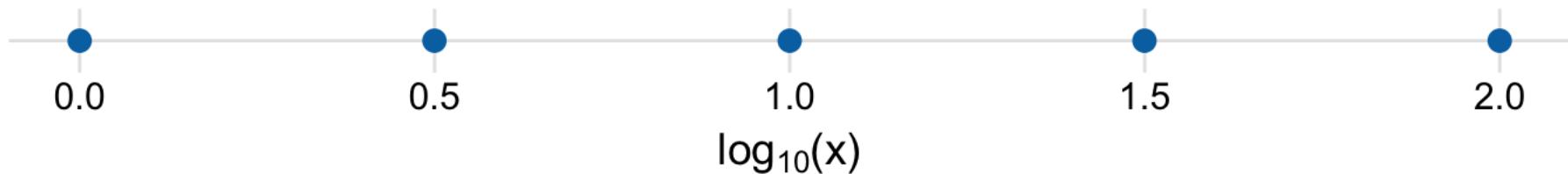


Why Transform?

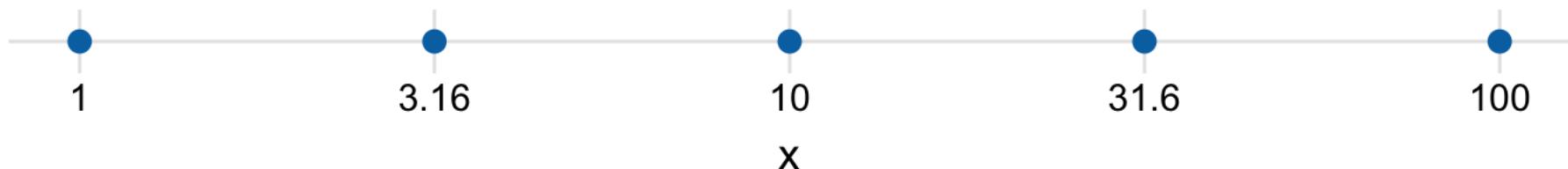
original data, linear scale



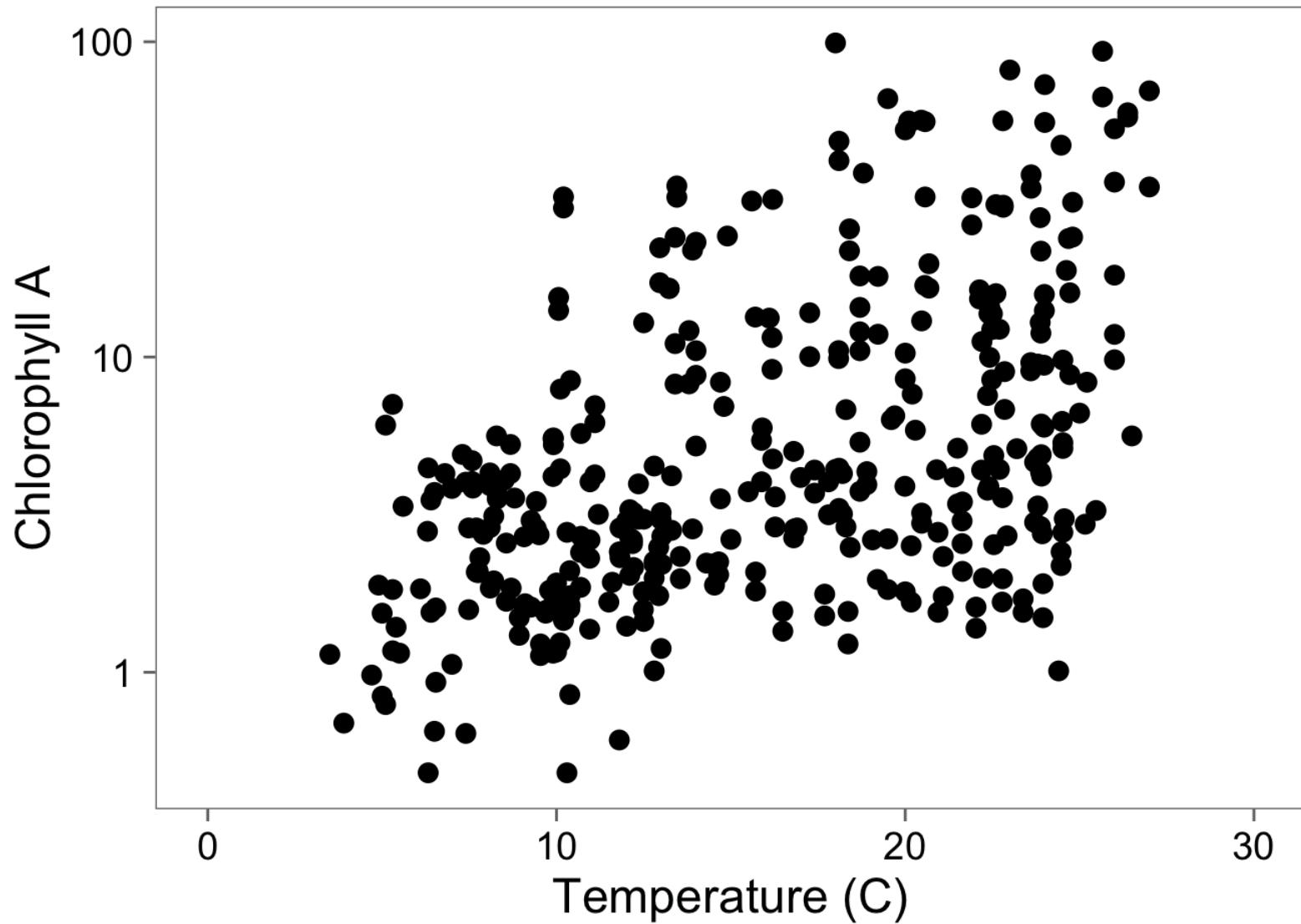
log-transformed data, linear scale



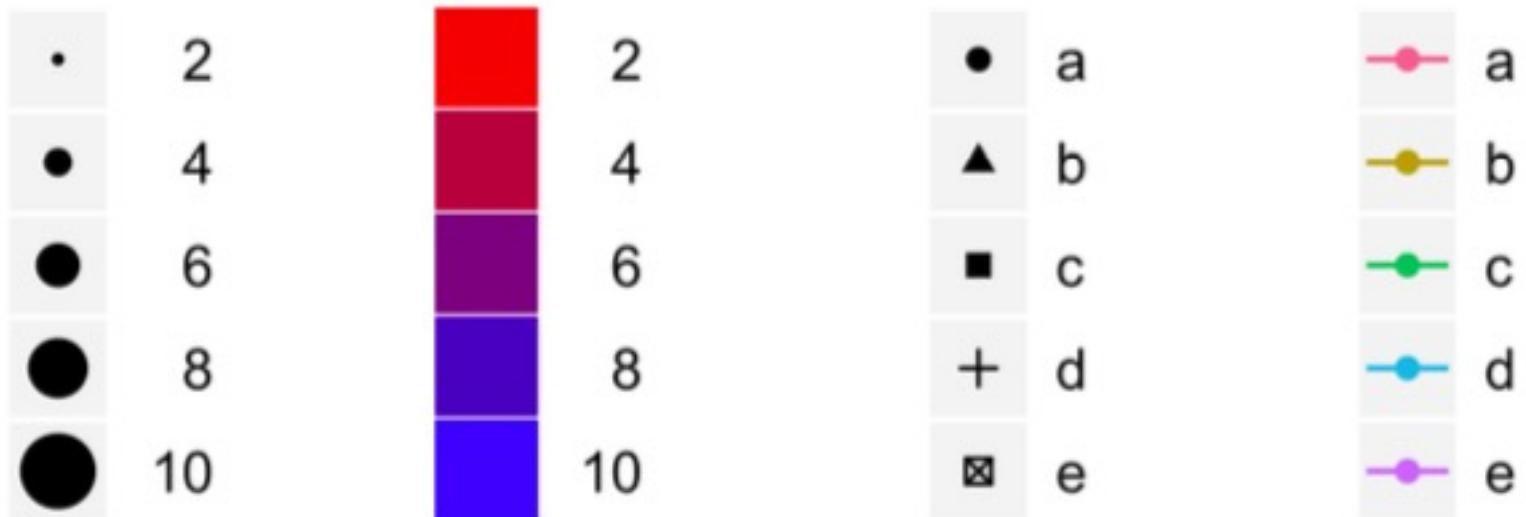
original data, logarithmic scale



Log-Transformation To See Relationship



Scales to Add Dimensions of Data



Colors Can Distinguish Groups

Okabe Ito



ColorBrewer Dark2



ggplot2 hue



Colors Can Show Data

ColorBrewer Blues



Heat



Viridis



Beware Not Thinking About Color Blindness

original



deuteranomaly



protanomaly



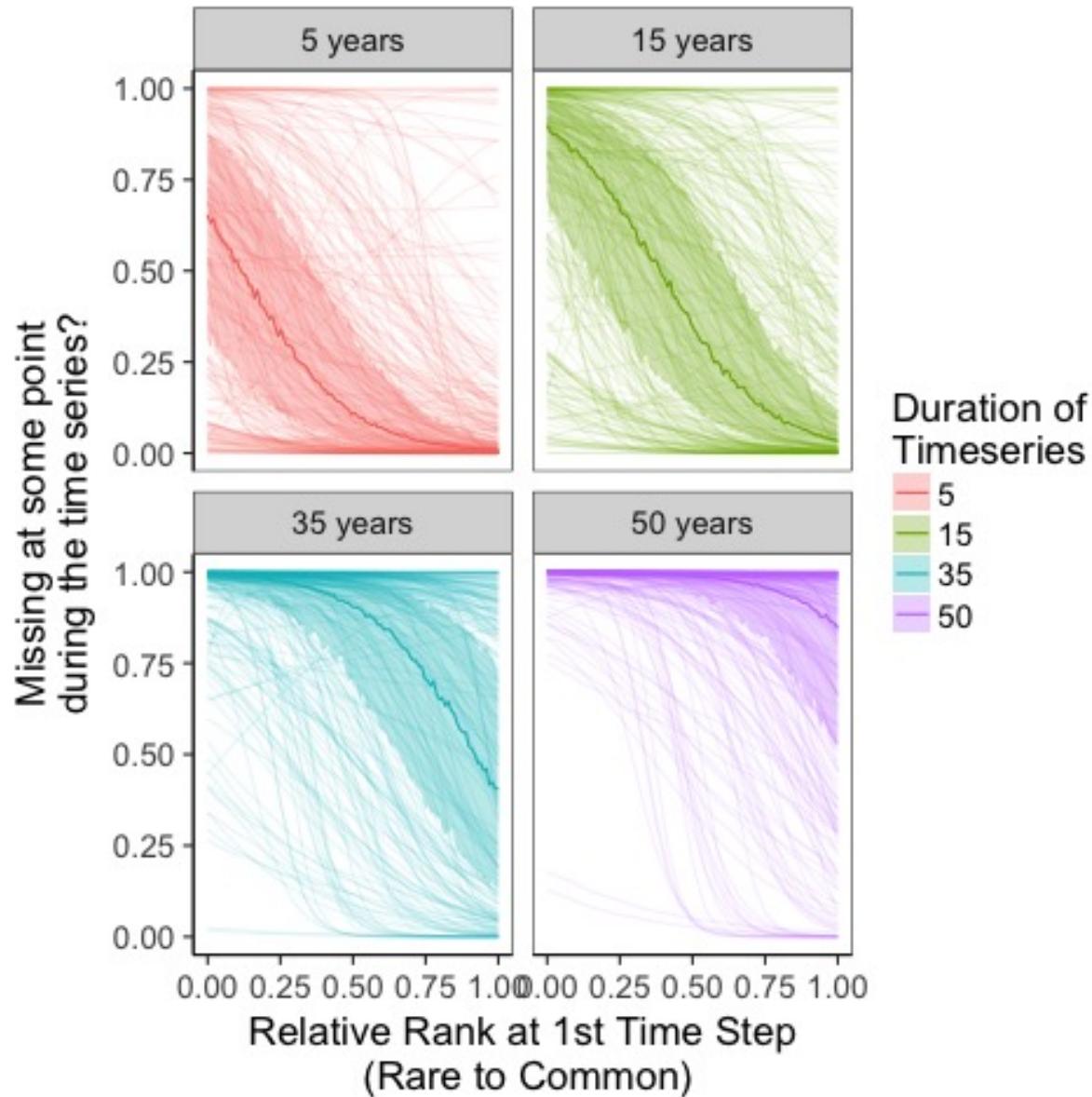
tritanomaly



Never use Red-Green!

Use redundant coding: shapes, sizes, etc.

Redundant Coding

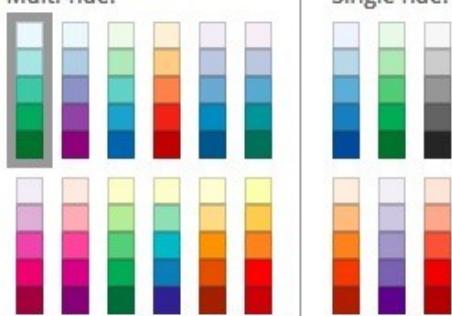


Colorbrewer.org

Number of data classes: 3

Nature of your data:
 sequential diverging qualitative

Pick a color scheme:

Multi-hue:


Single hue:

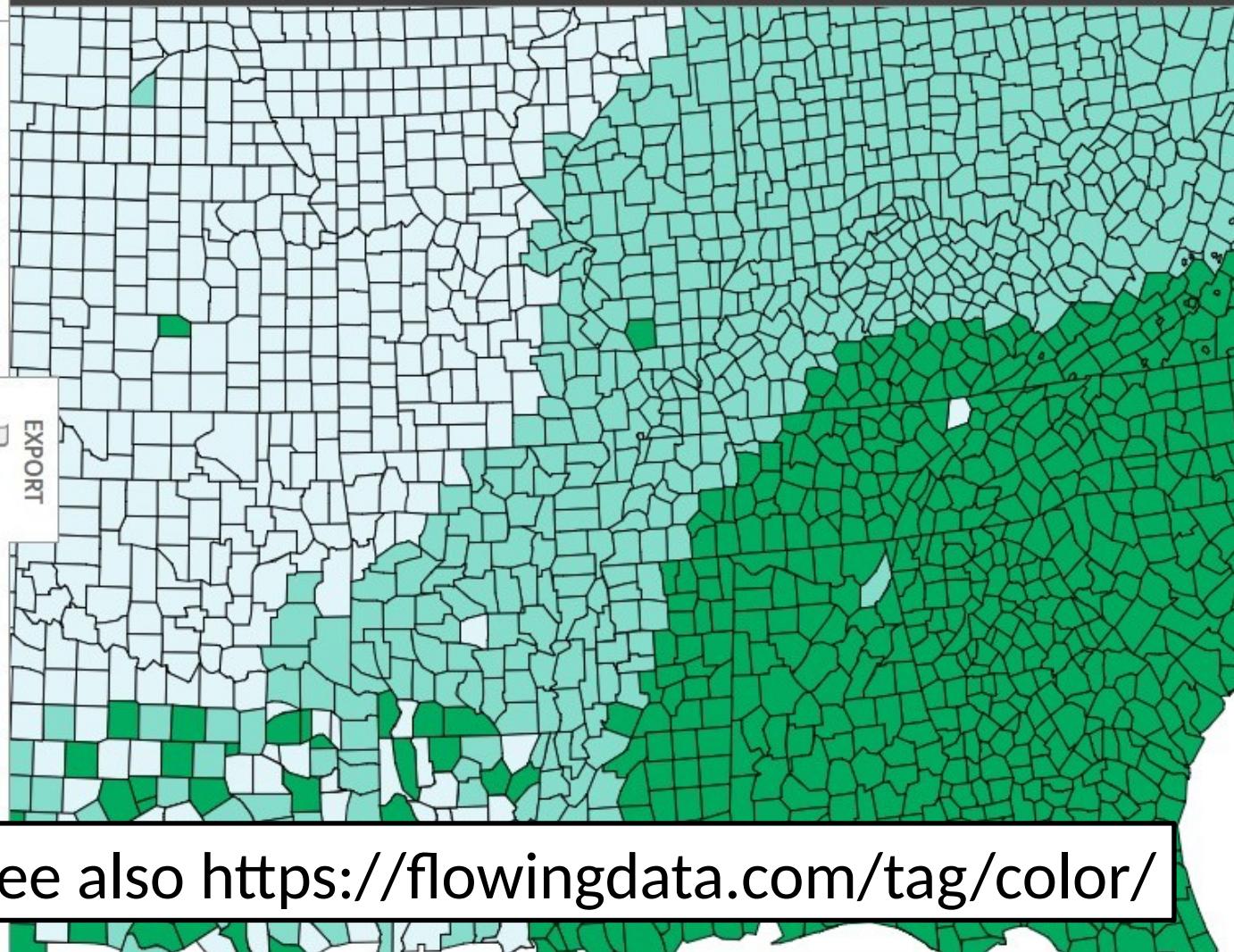

Only show:
 colorblind safe
 print friendly
 photocopy safe

Context:
 roads
 cities
 borders

Background:
 solid color terrain
 color transparency

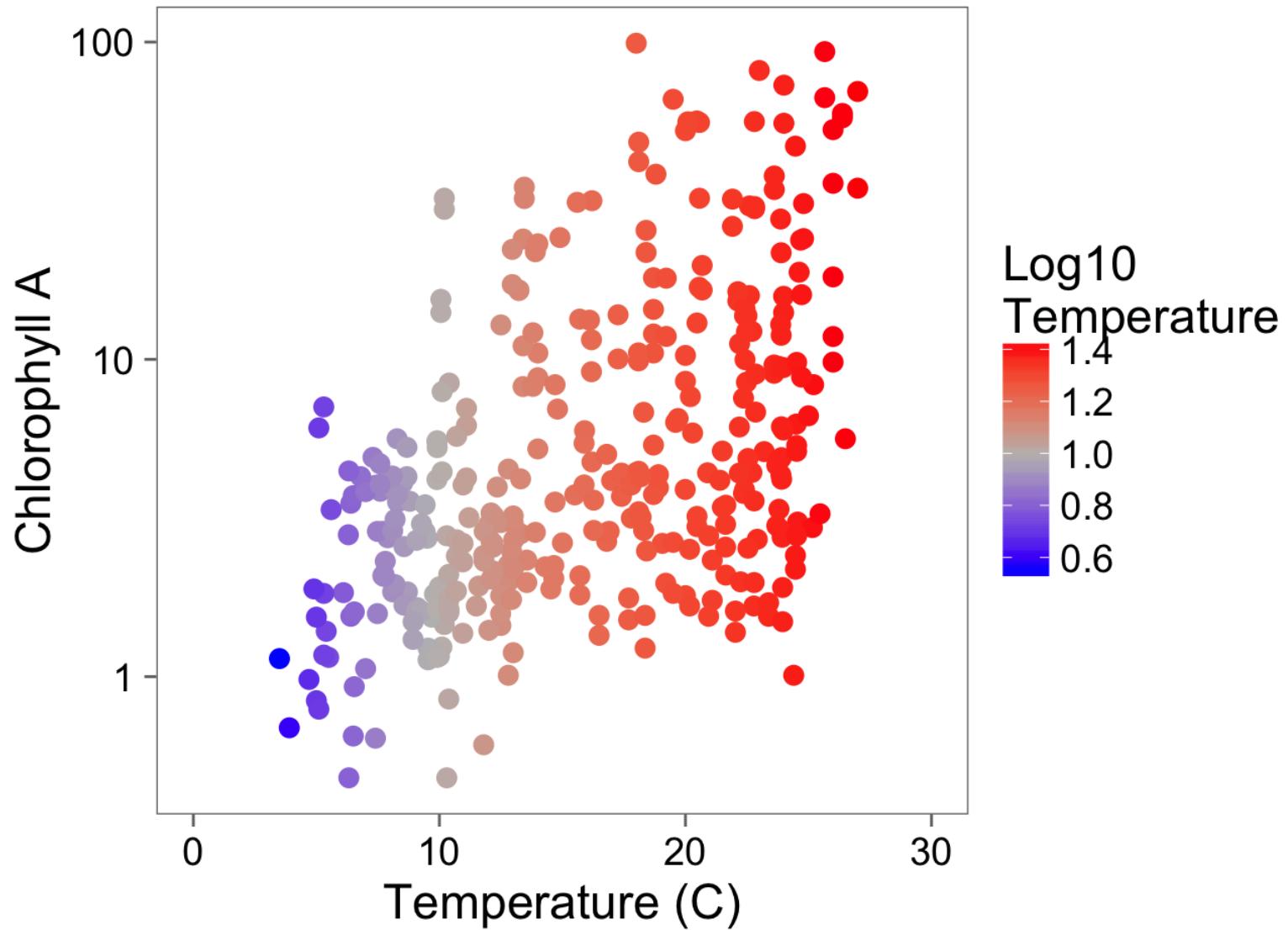
how to use | updates | downloads | credits

COLORBREWER 2.0
color advice for cartography

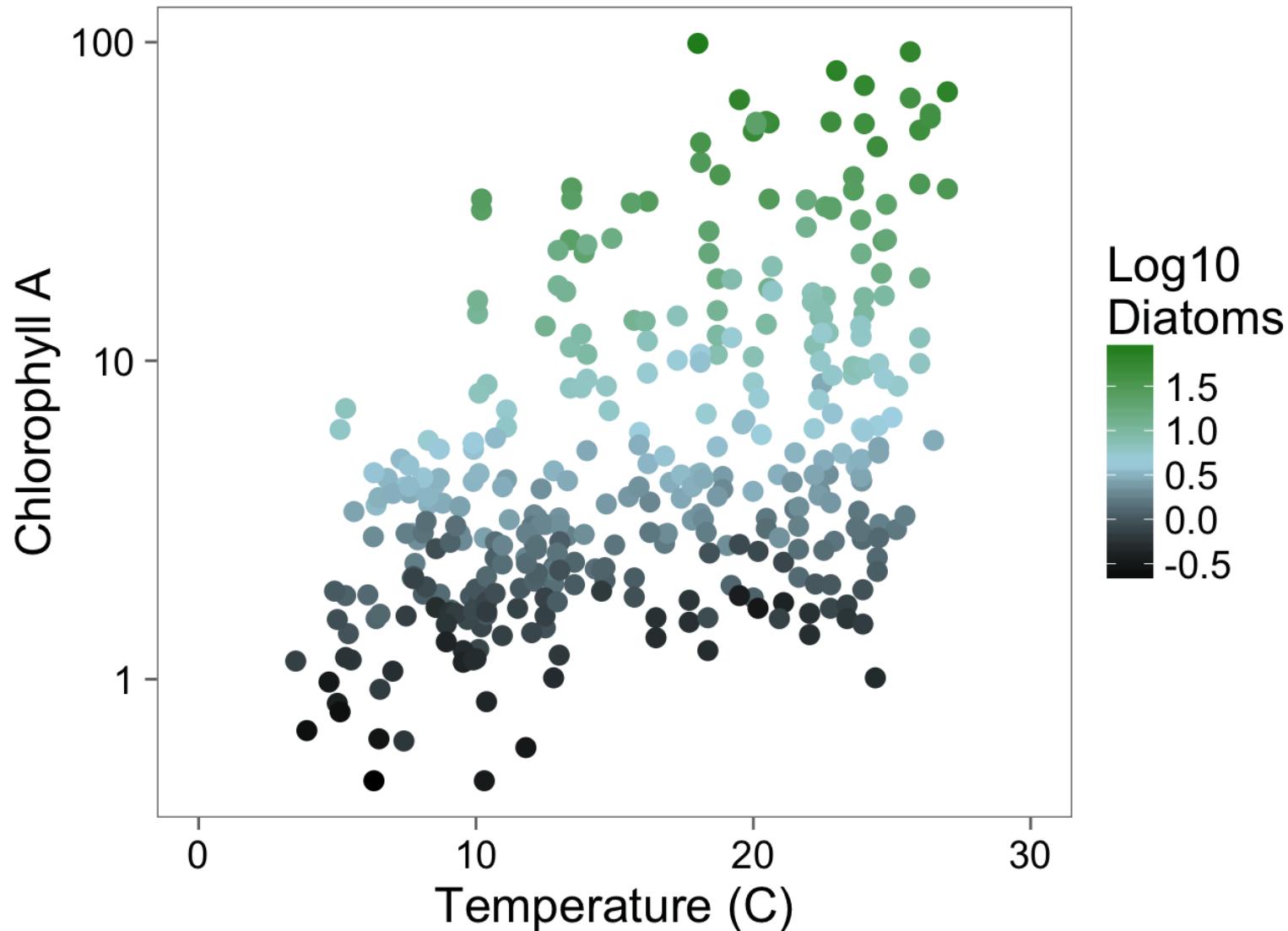


See also <https://flowingdata.com/tag/color/>

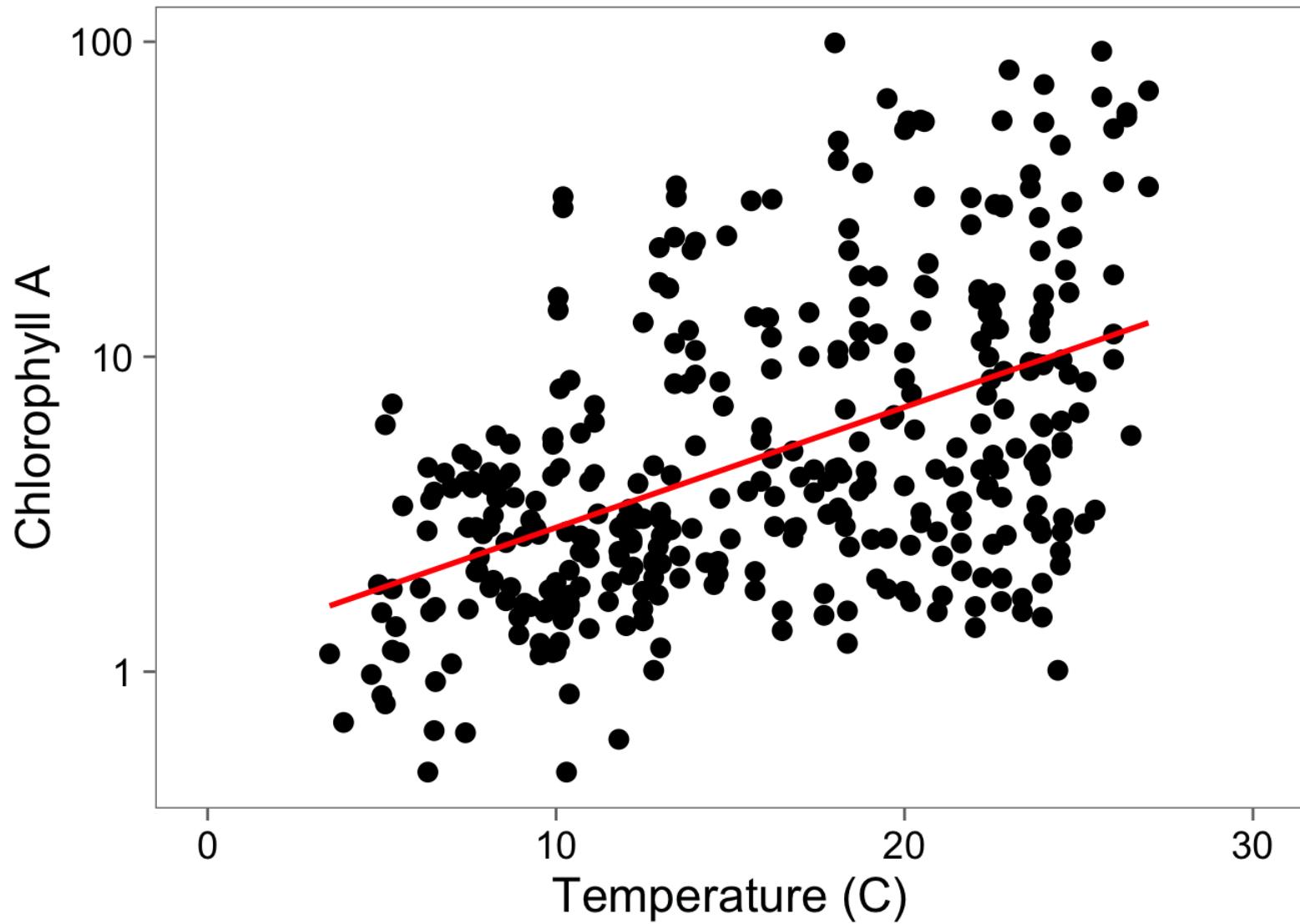
A Touch of Color!



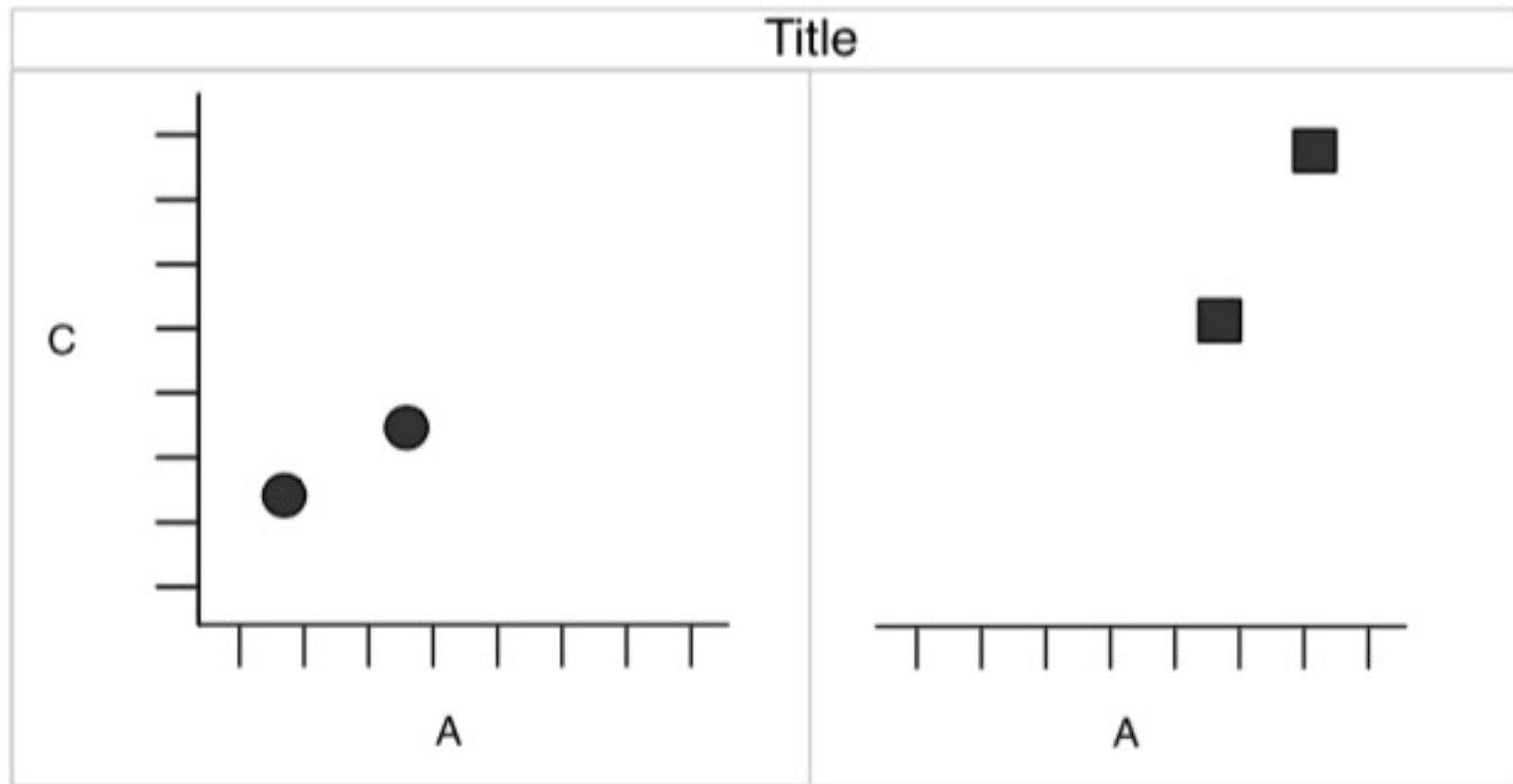
Color Can Bring in Another Dimension



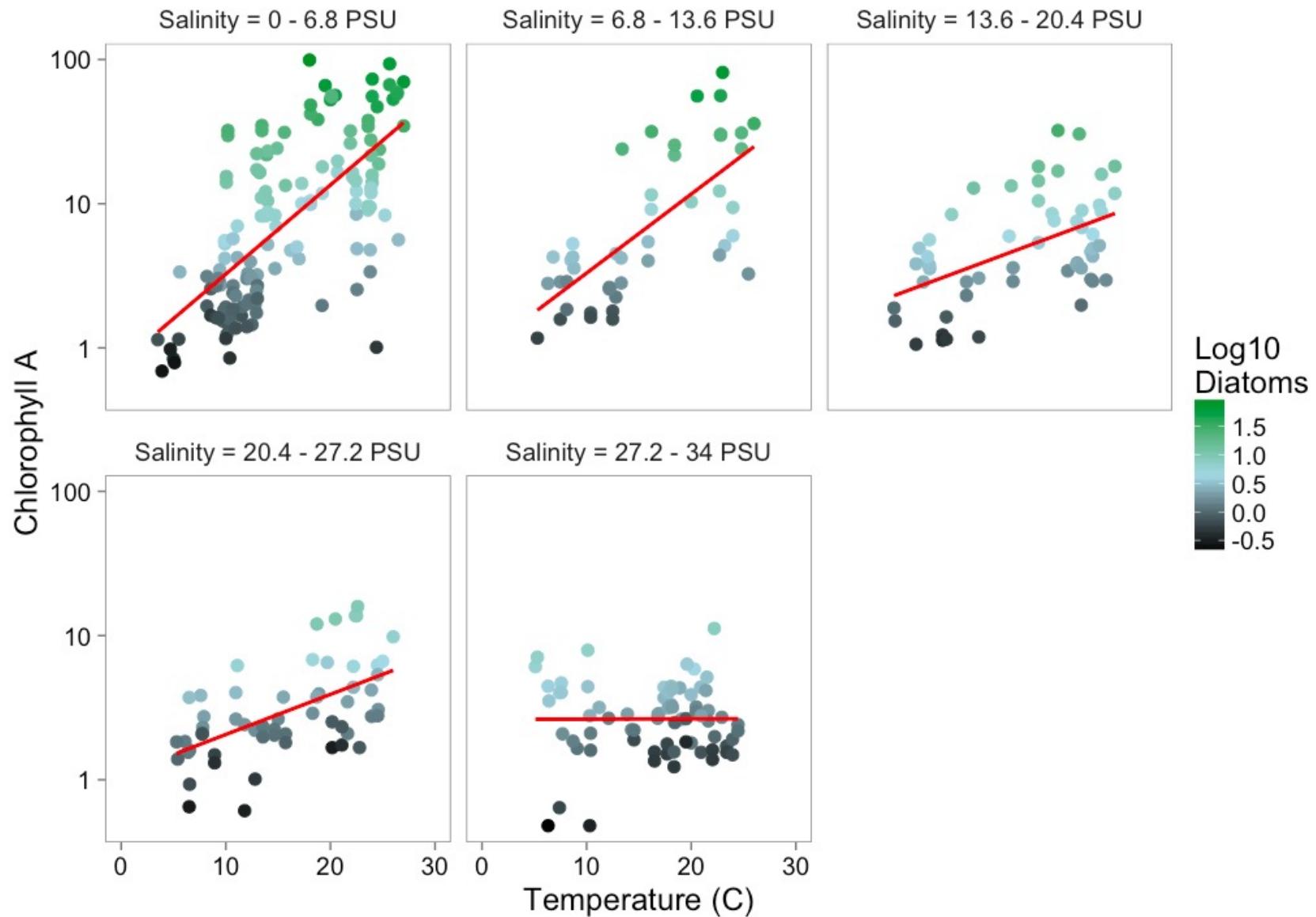
Statistical Fit to Aid Understanding



Facets to Add Fine-Grained Information or New Dimensions



Facets Add Information



Data Viz in a Nutshell

1. History
2. Graphical Basics
3. Minimalist principles

Minimalist Presentation

1. Above all else show data.
2. Maximize the data-ink ratio
3. Erase non-data-ink.

number of ideas in the shortest time with the least ink in the smallest space.

4. Erase redundant data-ink.

