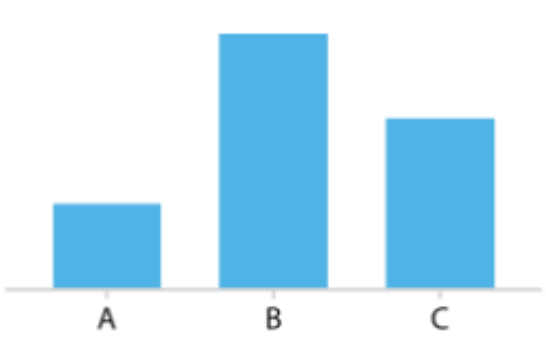
CSCI 491: Data Visualization

3- From Data to Visualization

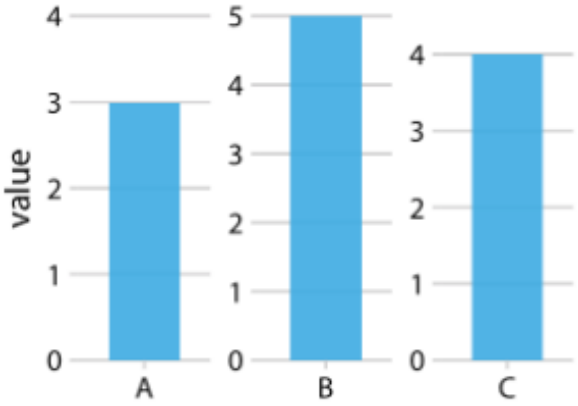
Ugly, Bad, Wrong Visualization

- Ugly: A figure that has aesthetic problems but otherwise is clear and informative (more subjective)
- Bad: A figure that has problems related to perception; it may be unclear, confusing, overly complicated, or deceiving (sometimes Deliberately)
- Wrong: A figure that has problems related to mathematics; it is objectively incorrect

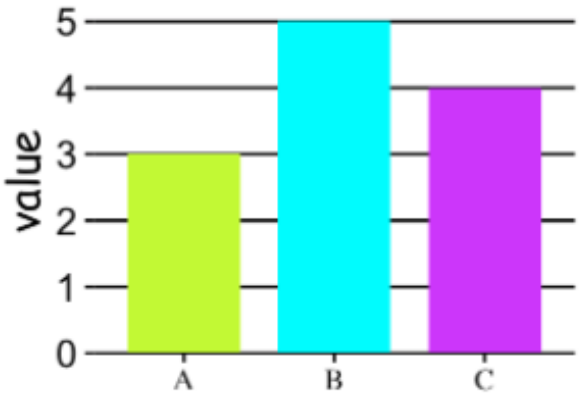
Ugly, Bad, Wrong Visualization



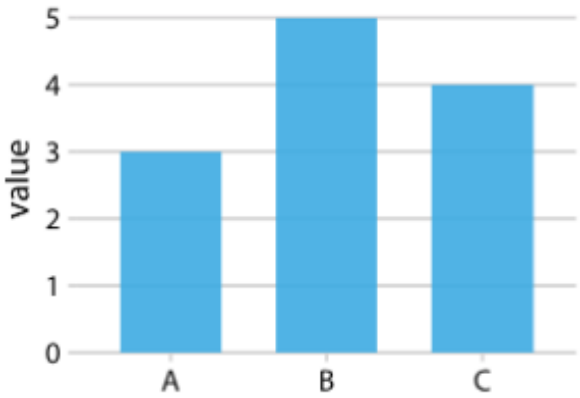
WRONG!!!



Bad

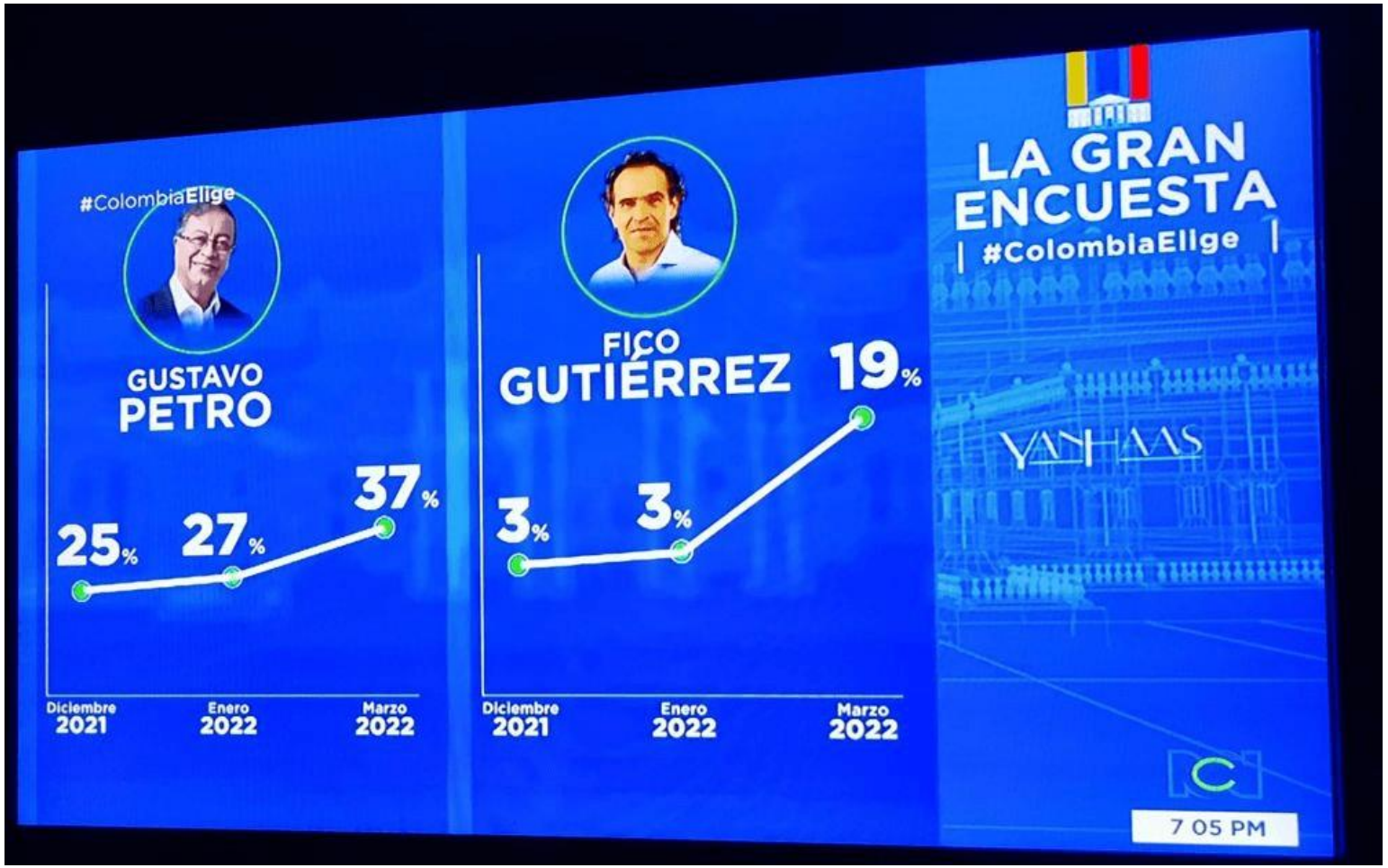


Ugly

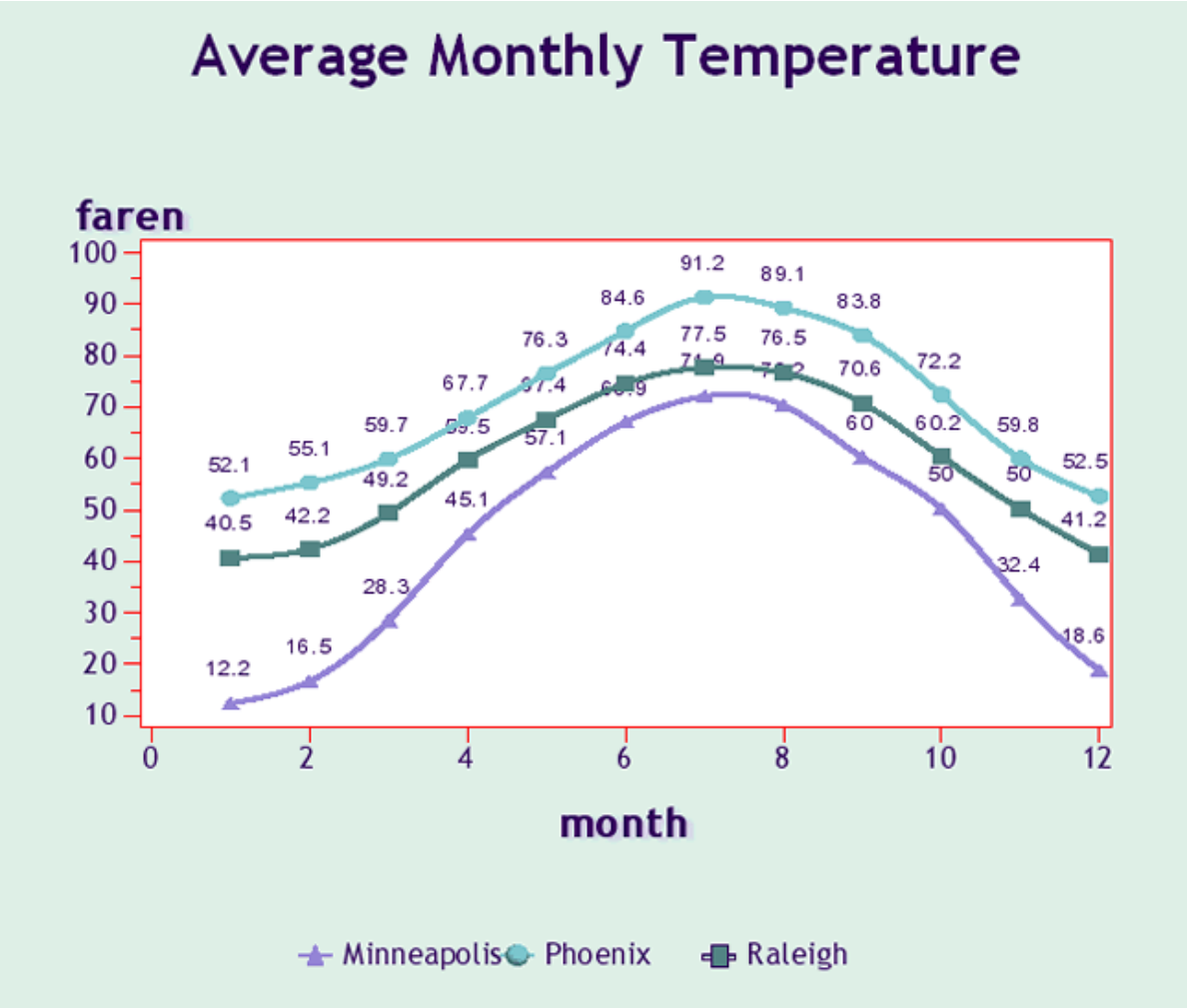


Acceptable

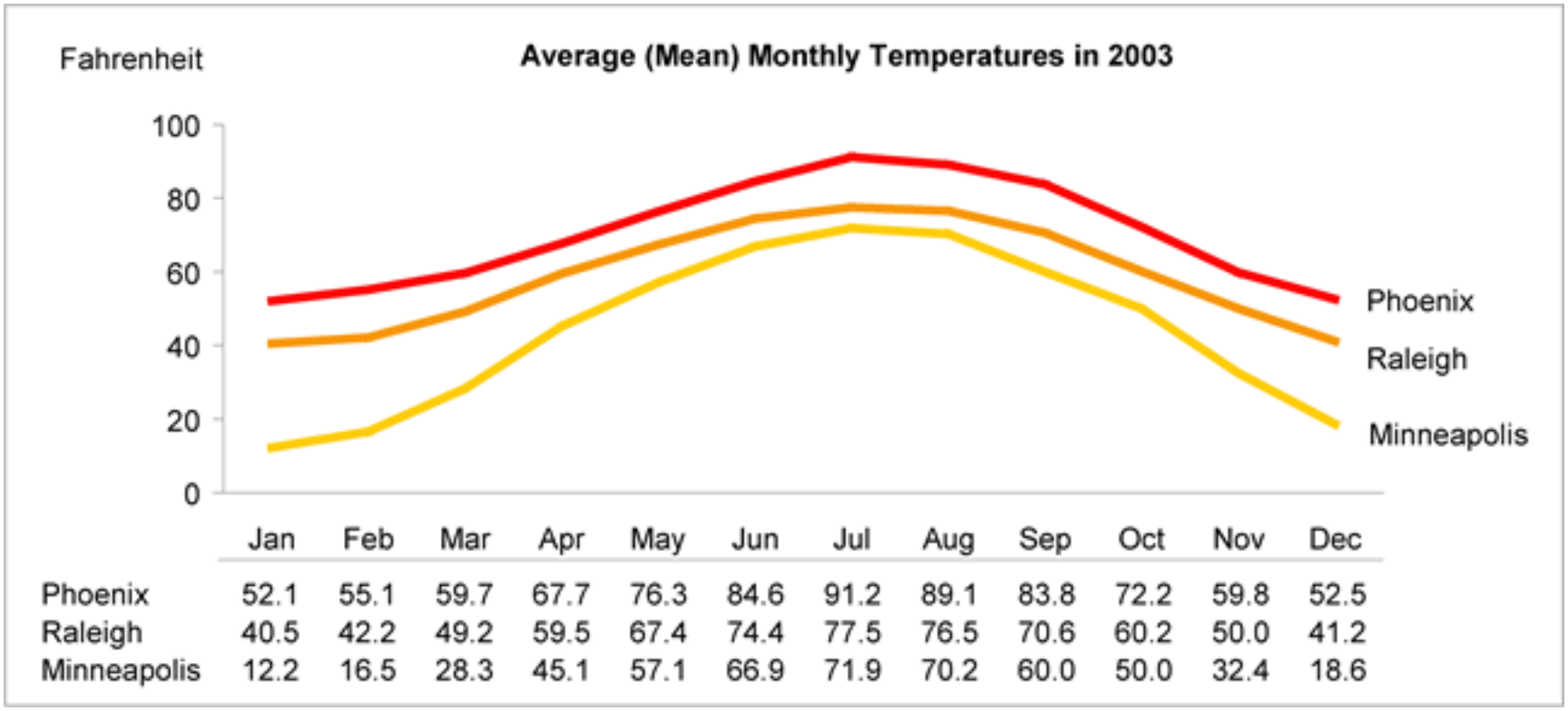
Good, Bad, Ugly, or Wrong?



Good, Bad, Ugly, or Wrong?

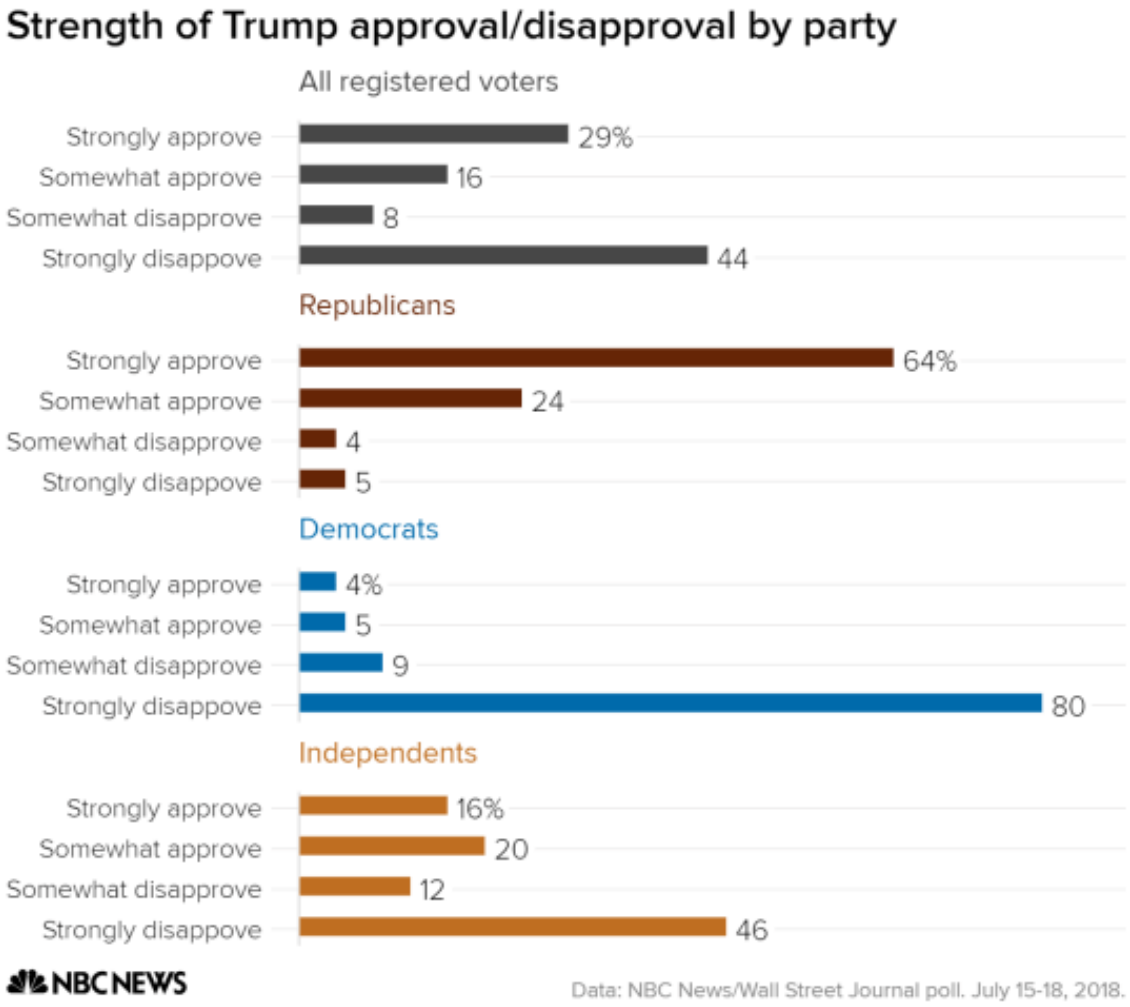


Good, Bad, Ugly, or Wrong?



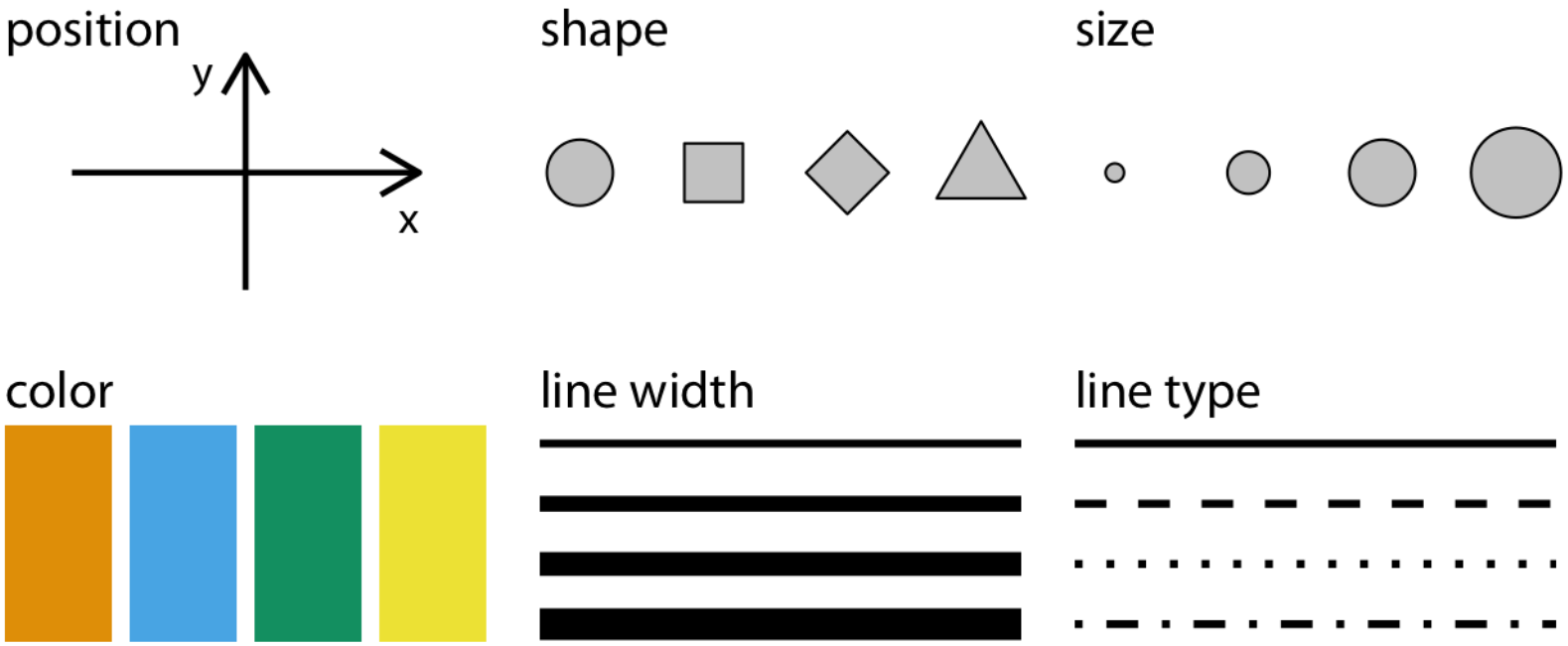
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Phoenix	52.1	55.1	59.7	67.7	76.3	84.6	91.2	89.1	83.8	72.2	59.8	52.5
Raleigh	40.5	42.2	49.2	59.5	67.4	74.4	77.5	76.5	70.6	60.2	50.0	41.2
Minneapolis	12.2	16.5	28.3	45.1	57.1	66.9	71.9	70.2	60.0	50.0	32.4	18.6

Good, Bad, Ugly, or Wrong?



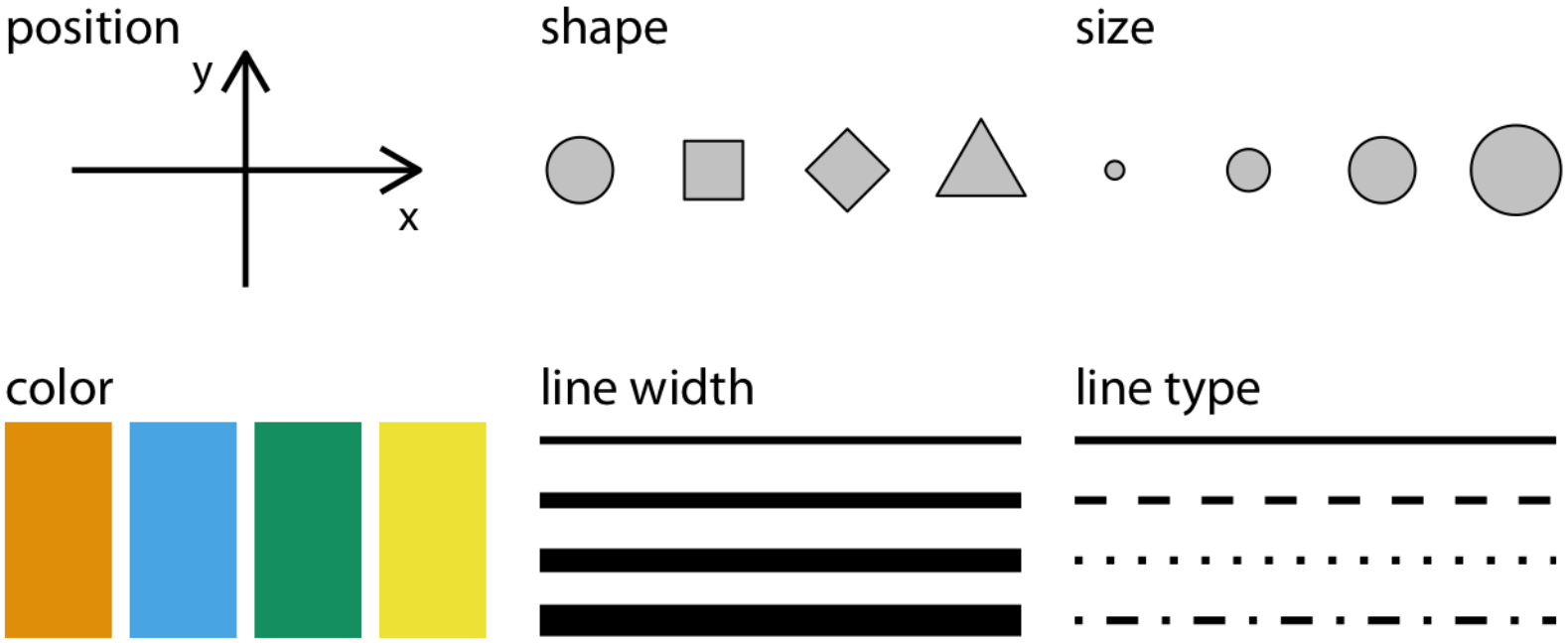
All data visualizations map data values into quantifiable features of the resulting graphic. We refer to these features as aesthetics.

Commonly Used Aesthetics



Continuous Vs. Discreet Data

- Which aesthetics can represent continuous data?
- Which aesthetics can represent Discreet data?



Data Types

- Quantitative/numerical Continuous
- Quantitative/numerical discrete
- Qualitative/categorical unordered (factors)
- Qualitative/categorical ordered
- Date or time
- Text

Example

Data source: National Oceanic and Atmospheric Administration (NOAA).

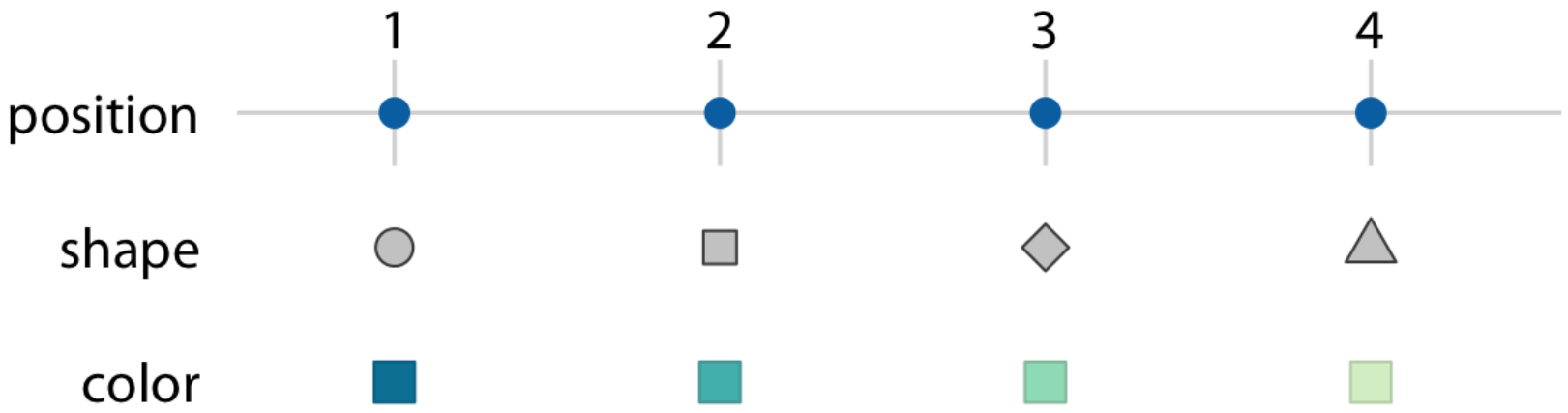
Month	Day	Location	Station ID	Temperature (°F)
Jan	1	Chicago	USW00014819	25.6
Jan	1	San Diego	USW00093107	55.2
Jan	1	Houston	USW00012918	53.9
Jan	1	Death Valley	USC00042319	51.0
Jan	2	Chicago	USW00014819	25.5
Jan	2	San Diego	USW00093107	55.3
Jan	2	Houston	USW00012918	53.8
Jan	2	Death Valley	USC00042319	51.2

iClicker

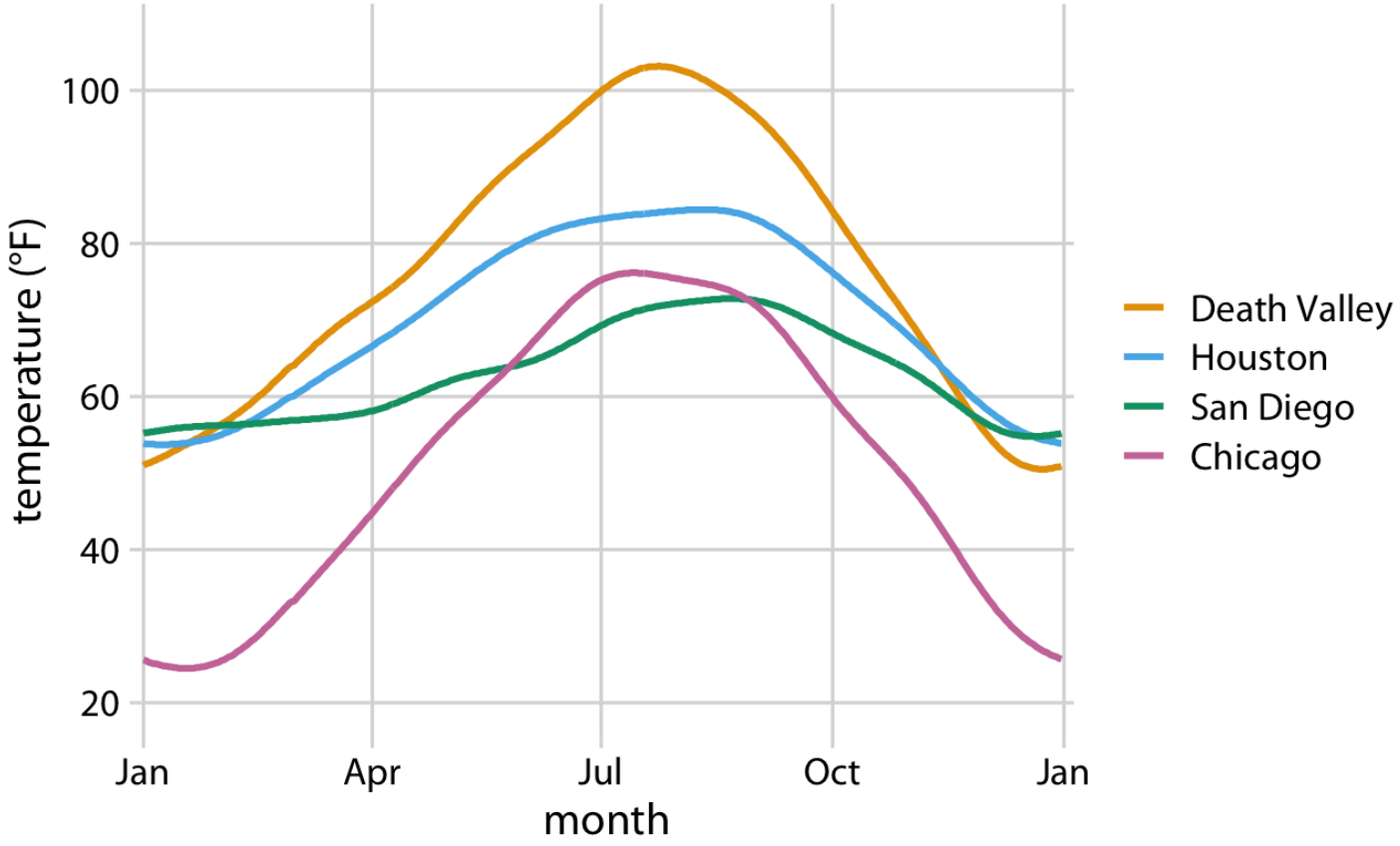
A: Quantitative/numerical Continuous
B: Quantitative/numerical discrete
C: Qualitative/categorical unordered (factors)
D: Qualitative/categorical ordered
E: Date or time

Scales Map Data Values onto Aesthetics

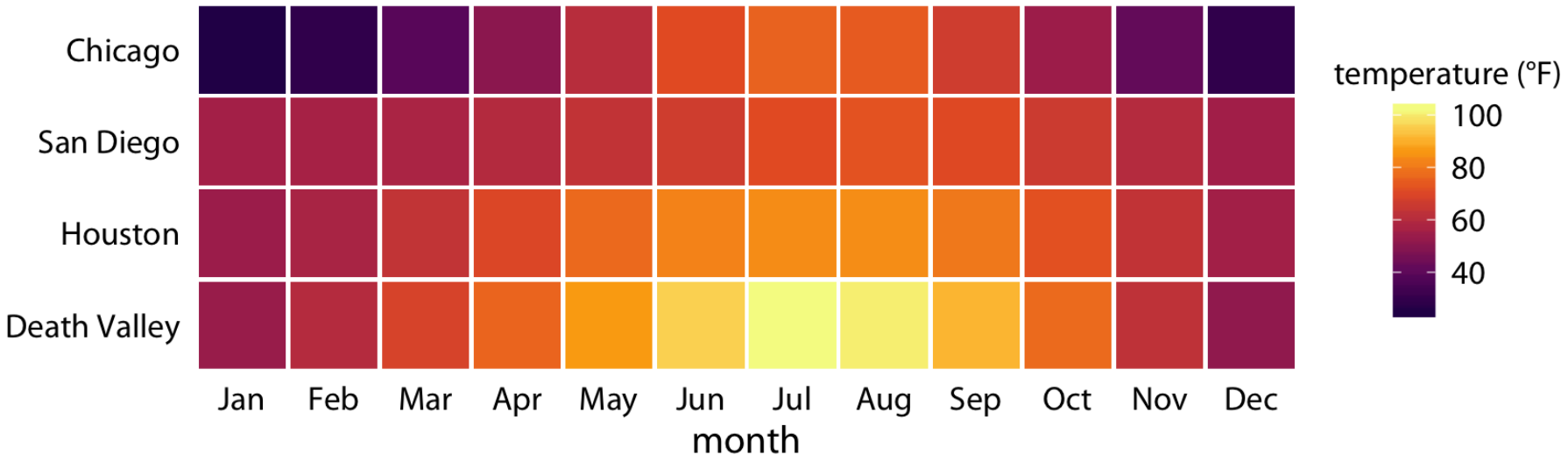
- The same data values can be mapped to different aesthetics.
- Importantly, a scale must be one-to-one, such that for each specific data value there is exactly one aesthetics value and vice versa.



Temperature Data Example



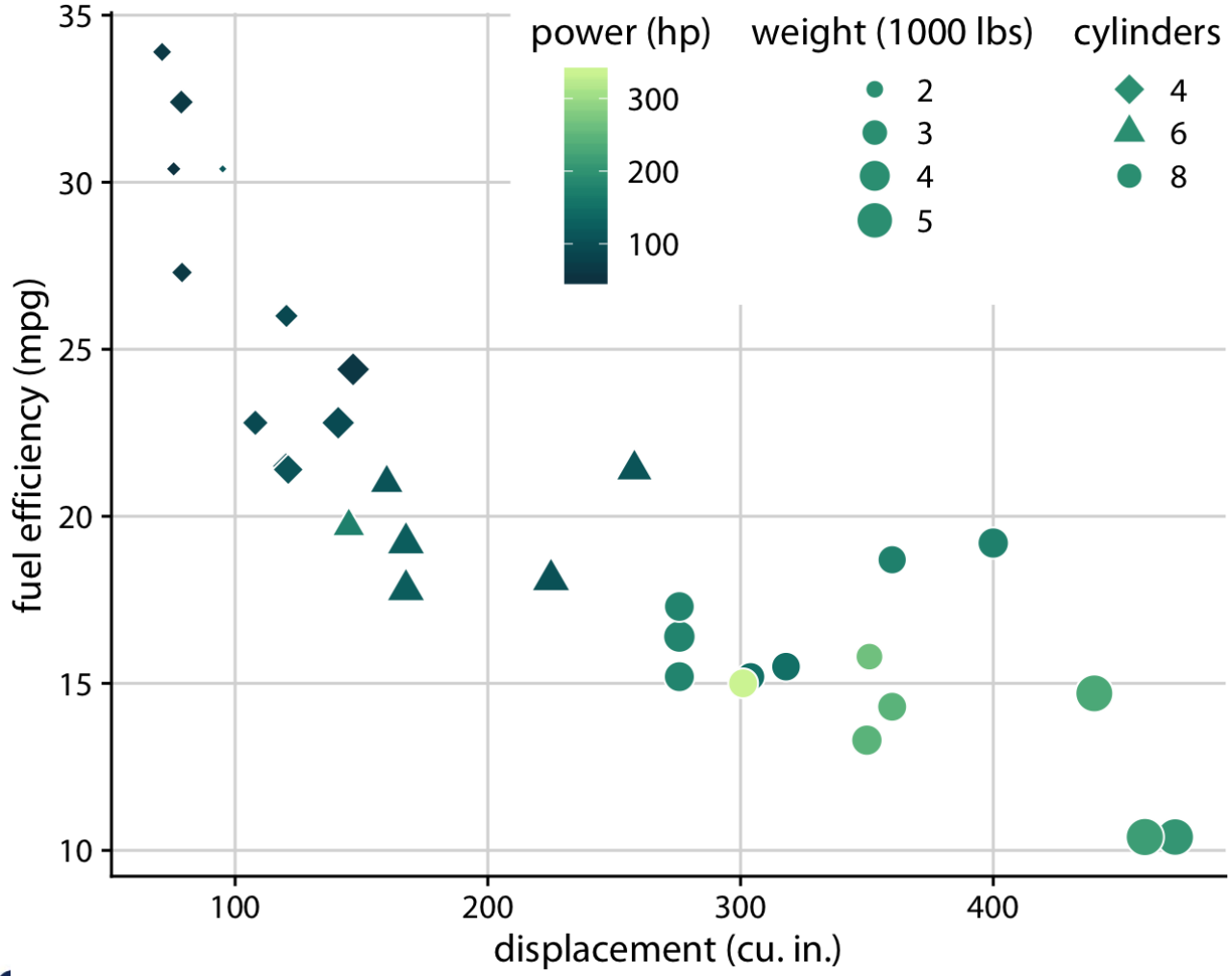
Temperature Data Example 2



iClicker

- How many scales were used in this plot?
- A: 1 B: 2 C: 3 D:4 E: 5

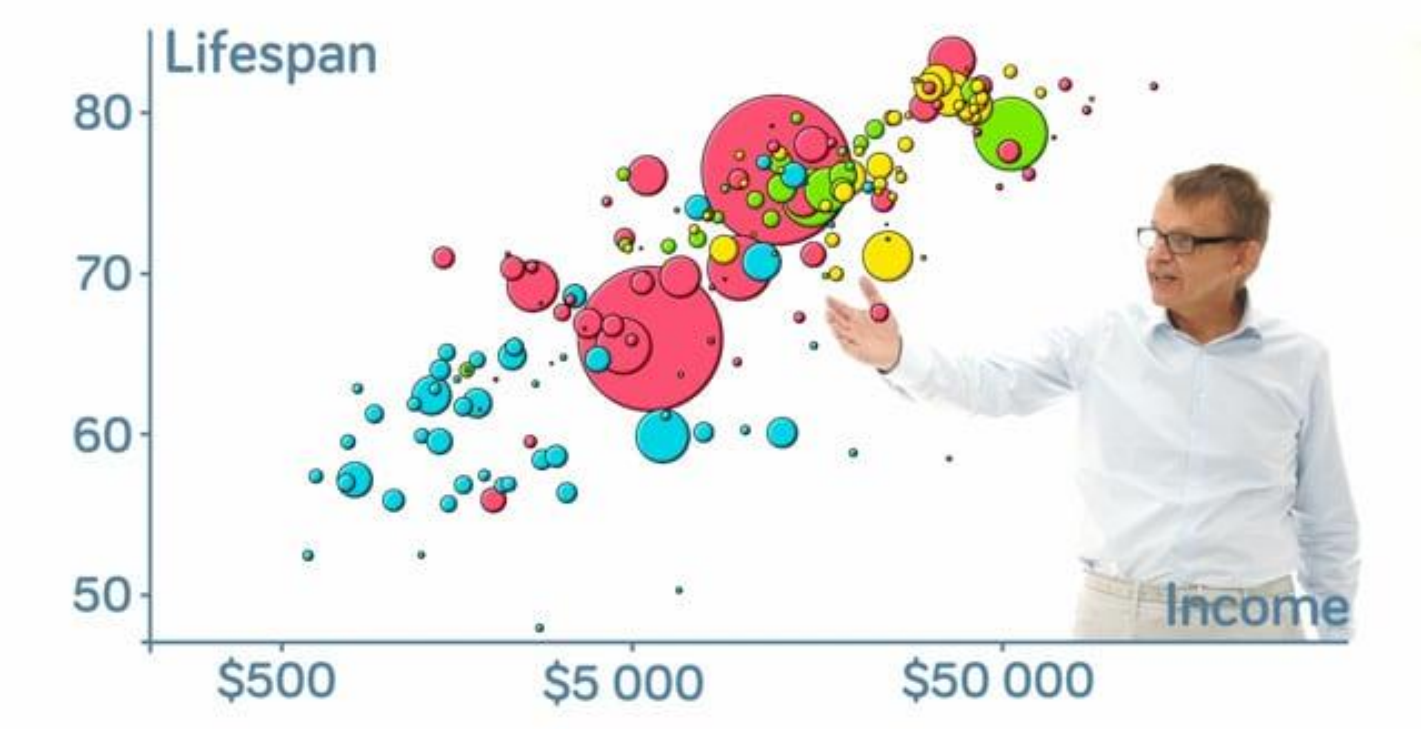
Using more than three scales



Hans Rosling

iClicker

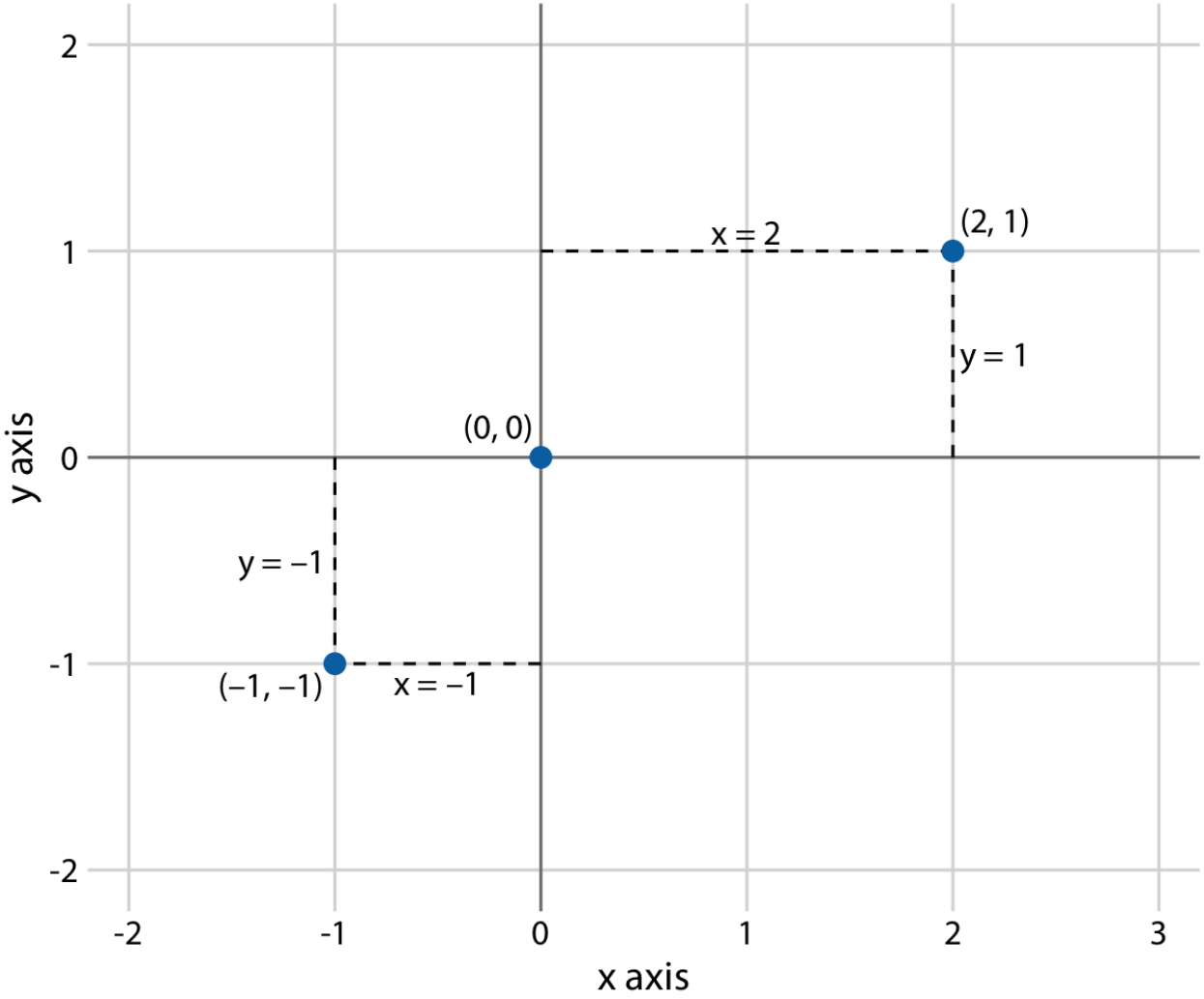
- How many scales were used in this plot?
- A: 1 B: 2 C: 3 D:4 E: 5



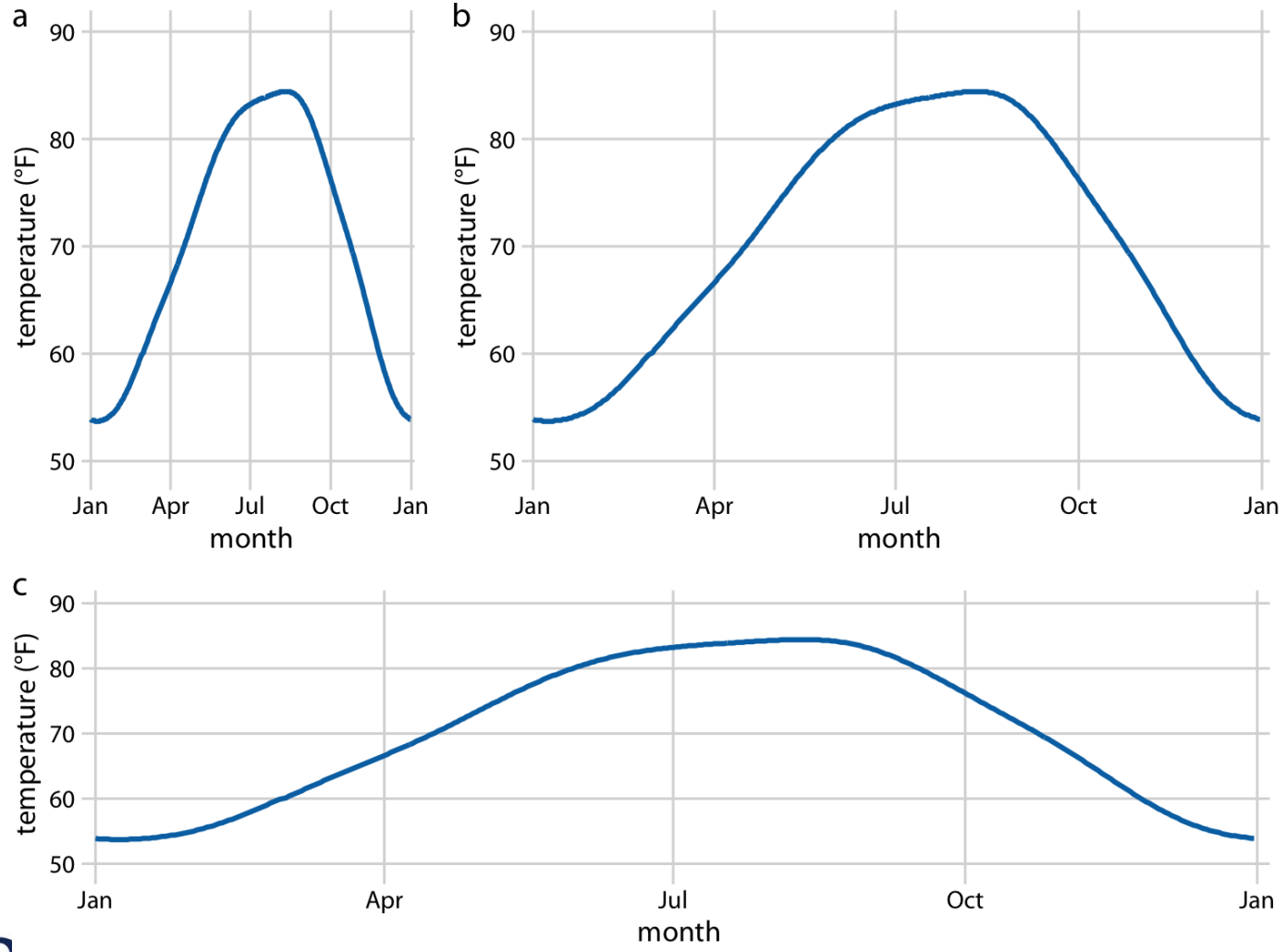
Coordinate Systems and Axes

- For regular 2D visualizations, two numbers are required to uniquely specify a point, and therefore we need two position scales.
- These two scales are usually but not necessarily the x and y axes of the plot.
- We also have to specify the relative geometric arrangement of these scales. Conventionally, the x axis runs horizontally and the y axis vertically, but we could choose other arrangements.
- The combination of a set of position scales and their relative geometric arrangement is called a coordinate system.

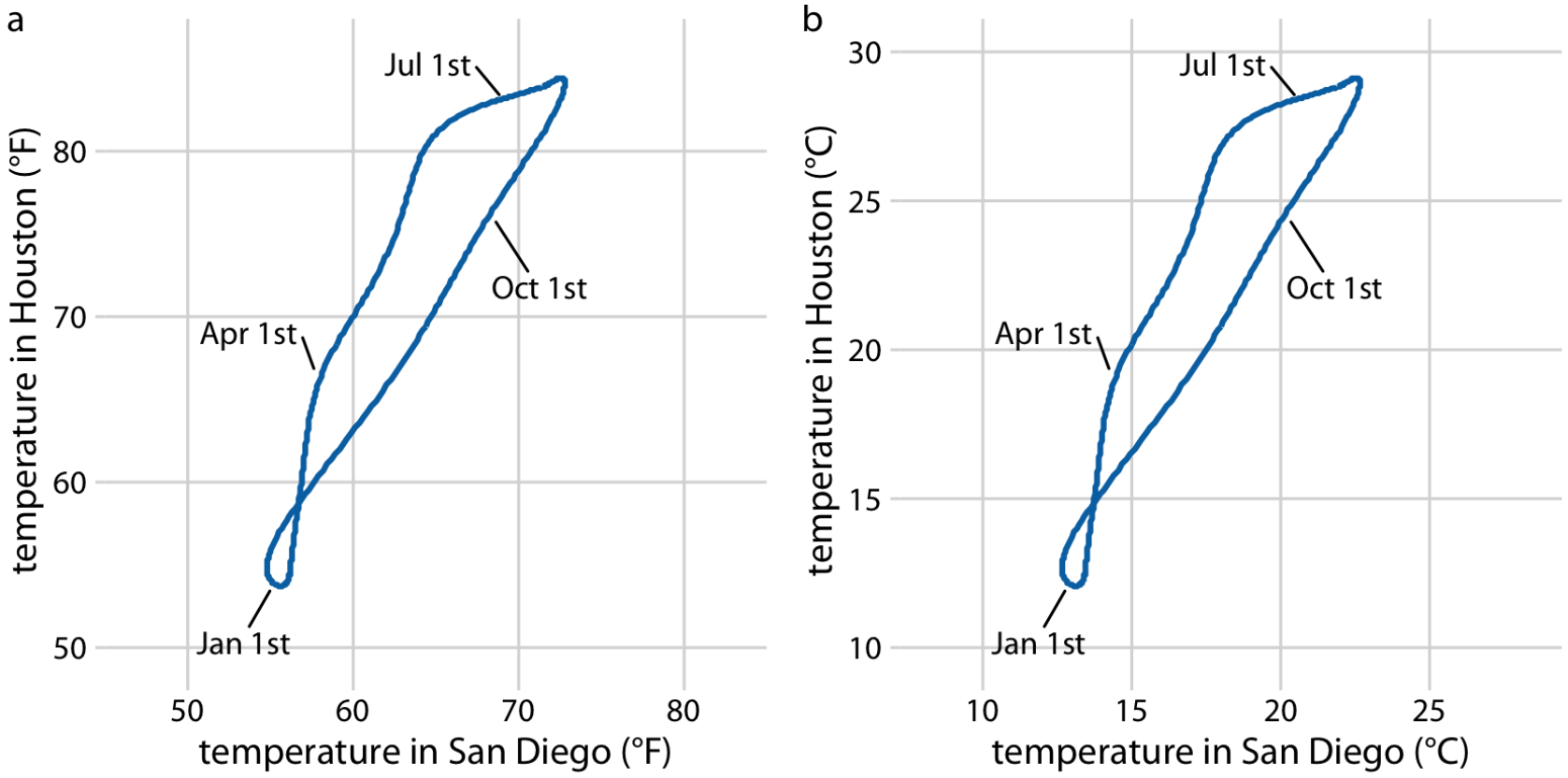
Cartesian Coordinates



Cartesian coordinate representing two different units

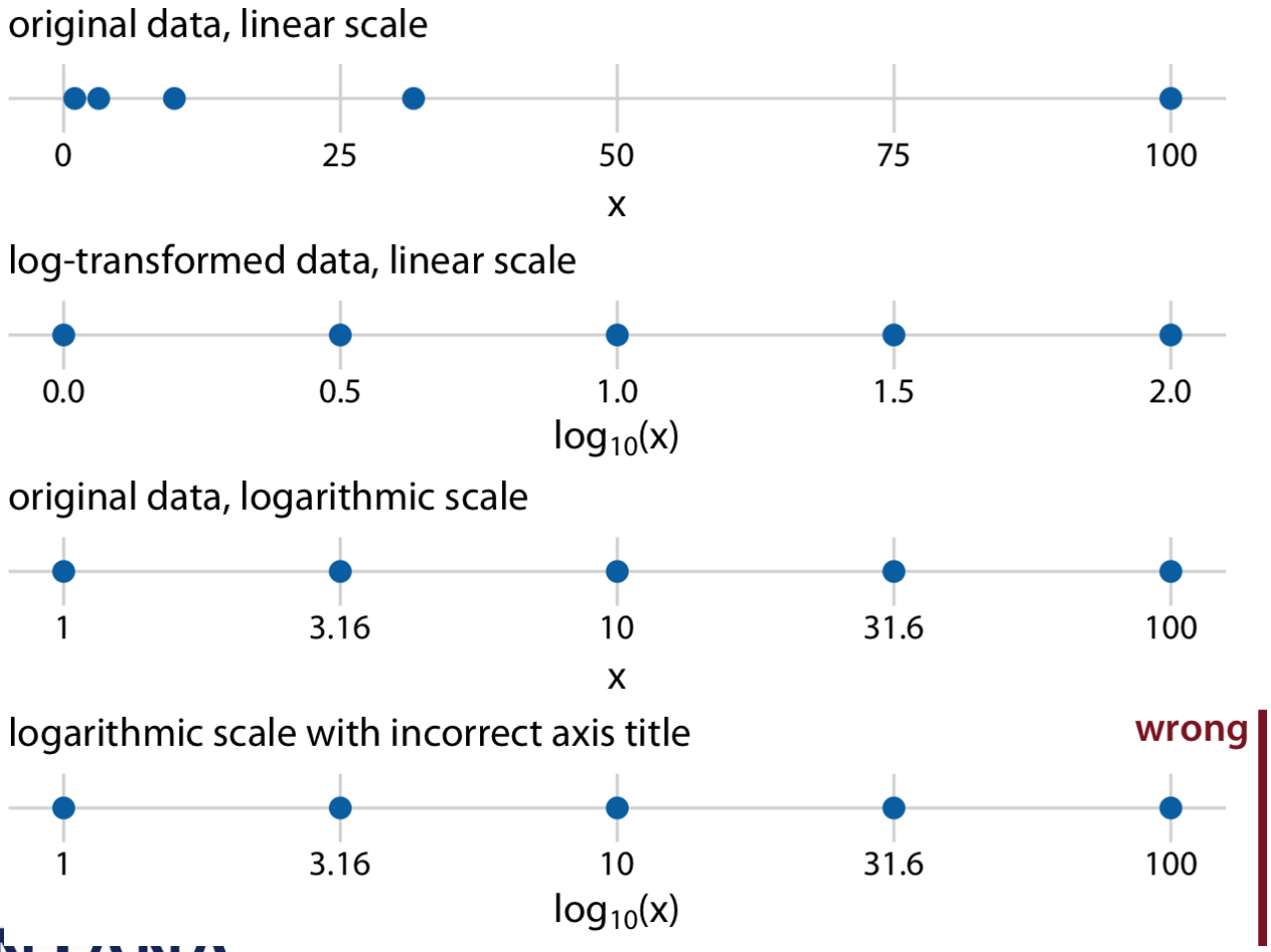


Cartesian coordinate representing similar units

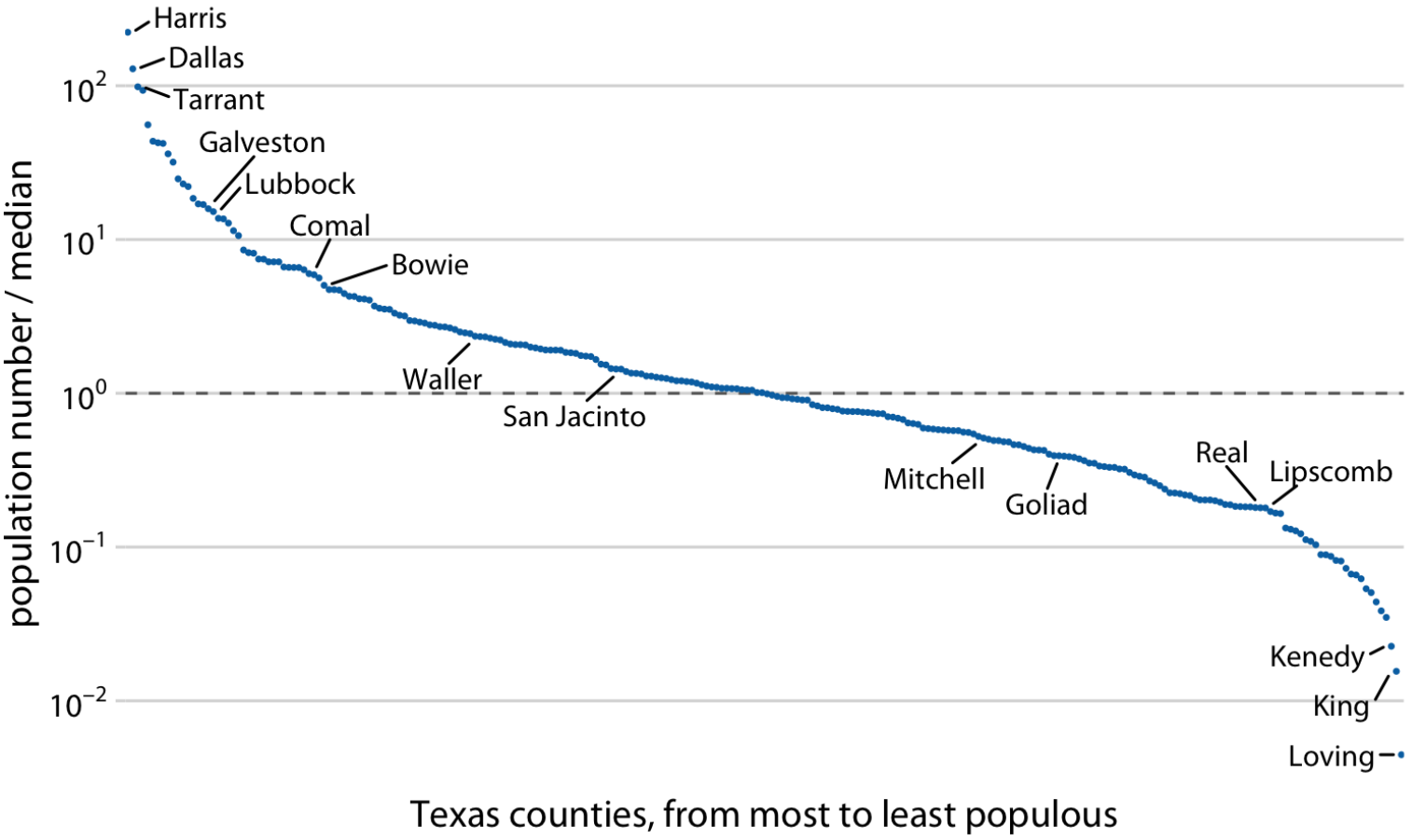


Nonlinear Axes

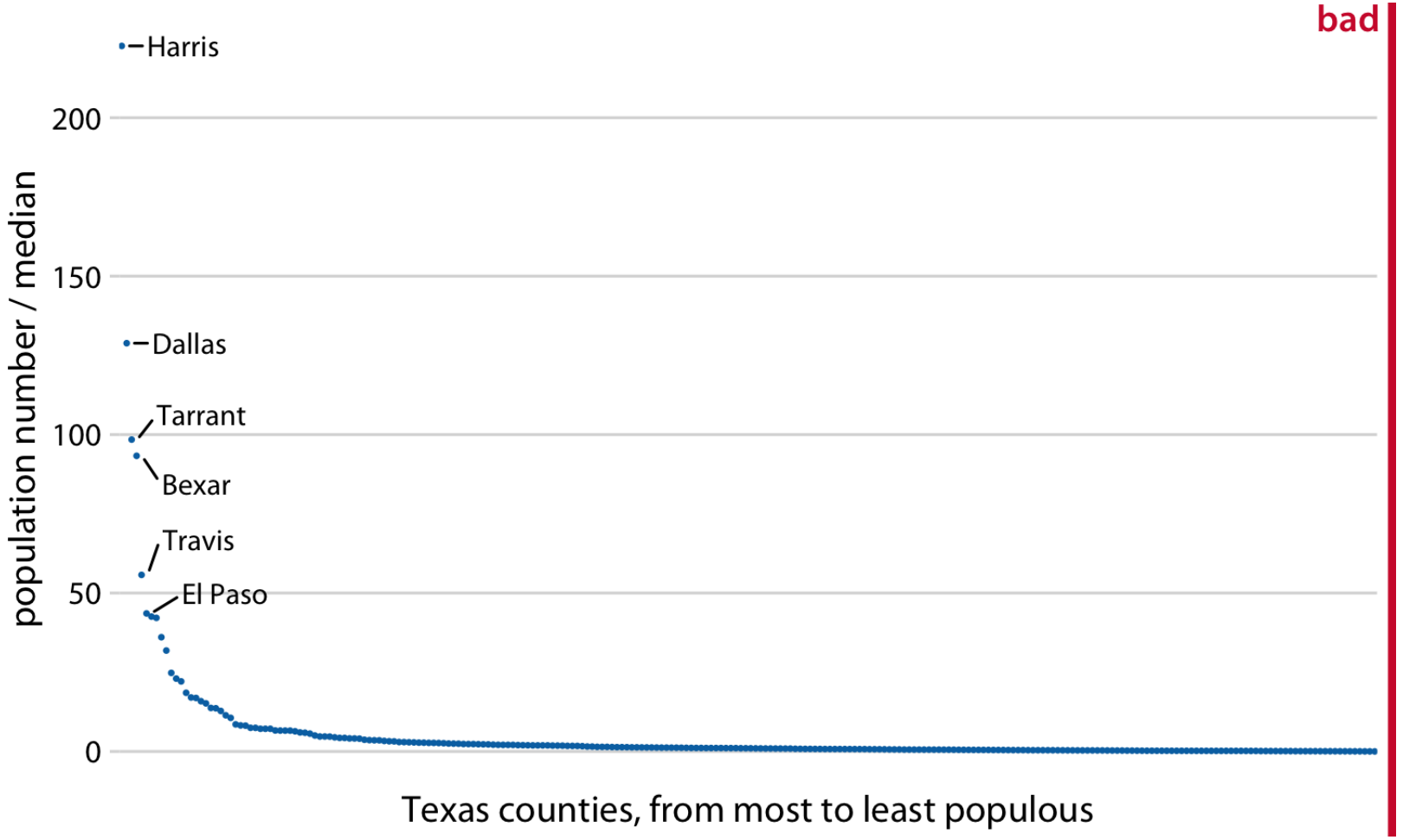
Values: 1, 3.16, 10, 31.6, and 100



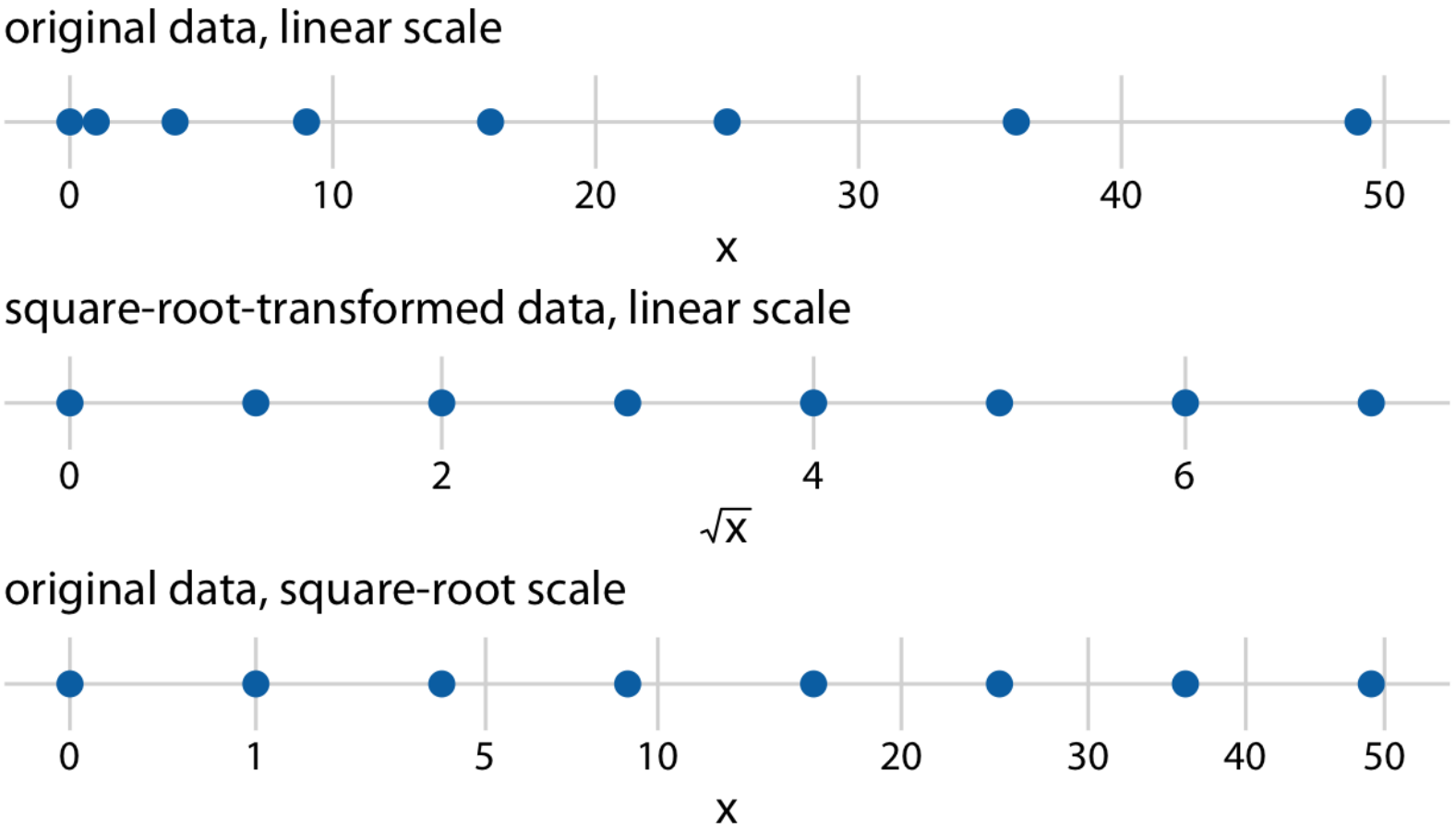
Example


Harris
Dallas
Tarrant
Galveston
Lubbock
Comal
Bowie
Waller
San Jacinto
Mitchell
Goliad
Real
Lipscomb
Kenedy
King
Loving
10^2
10^1
10^0
10^{-1}
10^{-2}
population number / median
Texas counties, from most to least populous

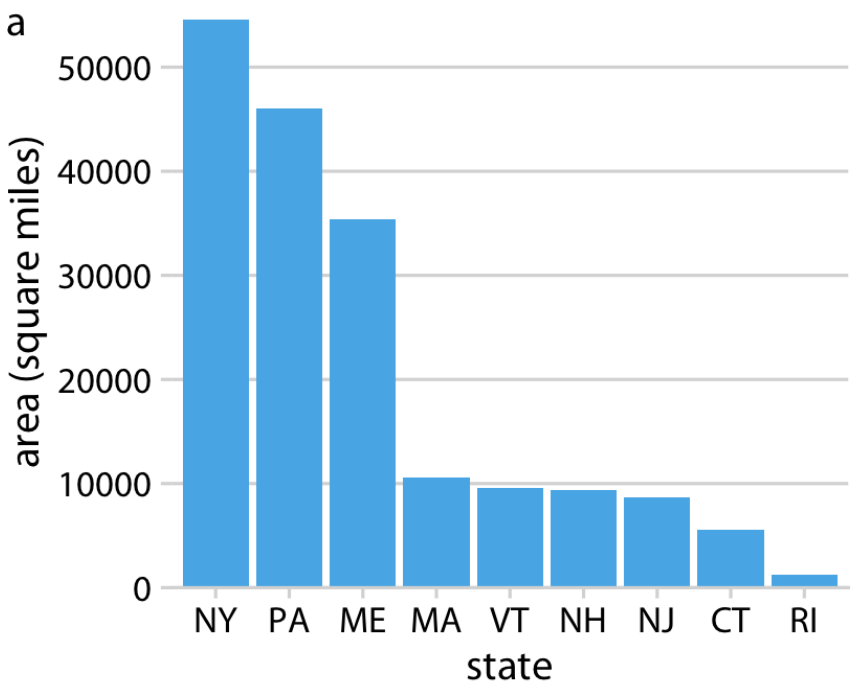
Example

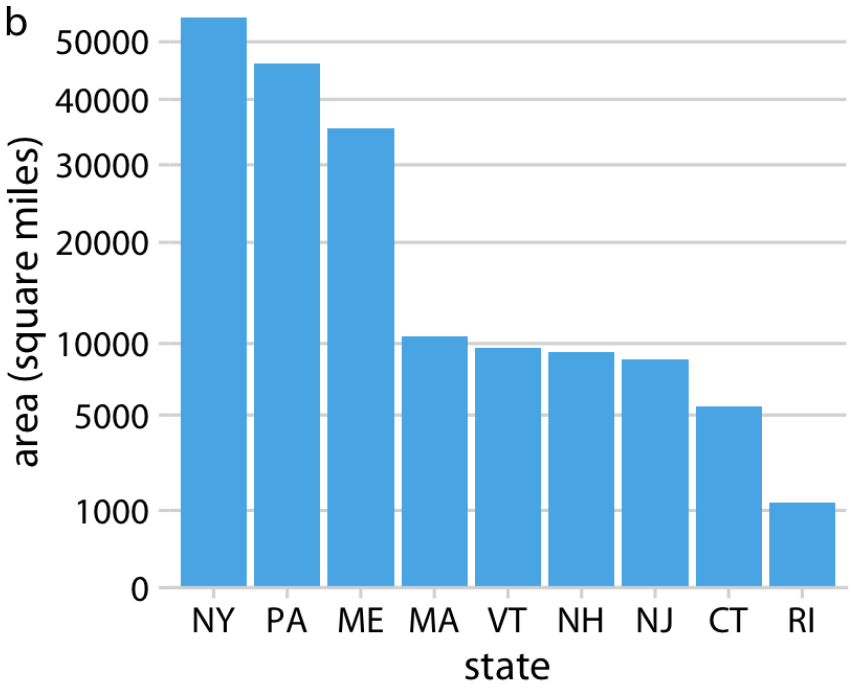

bad
Harris
200
150
Dallas
Tarrant
100
Bexar
Travis
50
El Paso
0
population number / median
Texas counties, from most to least populous

Square-root scale

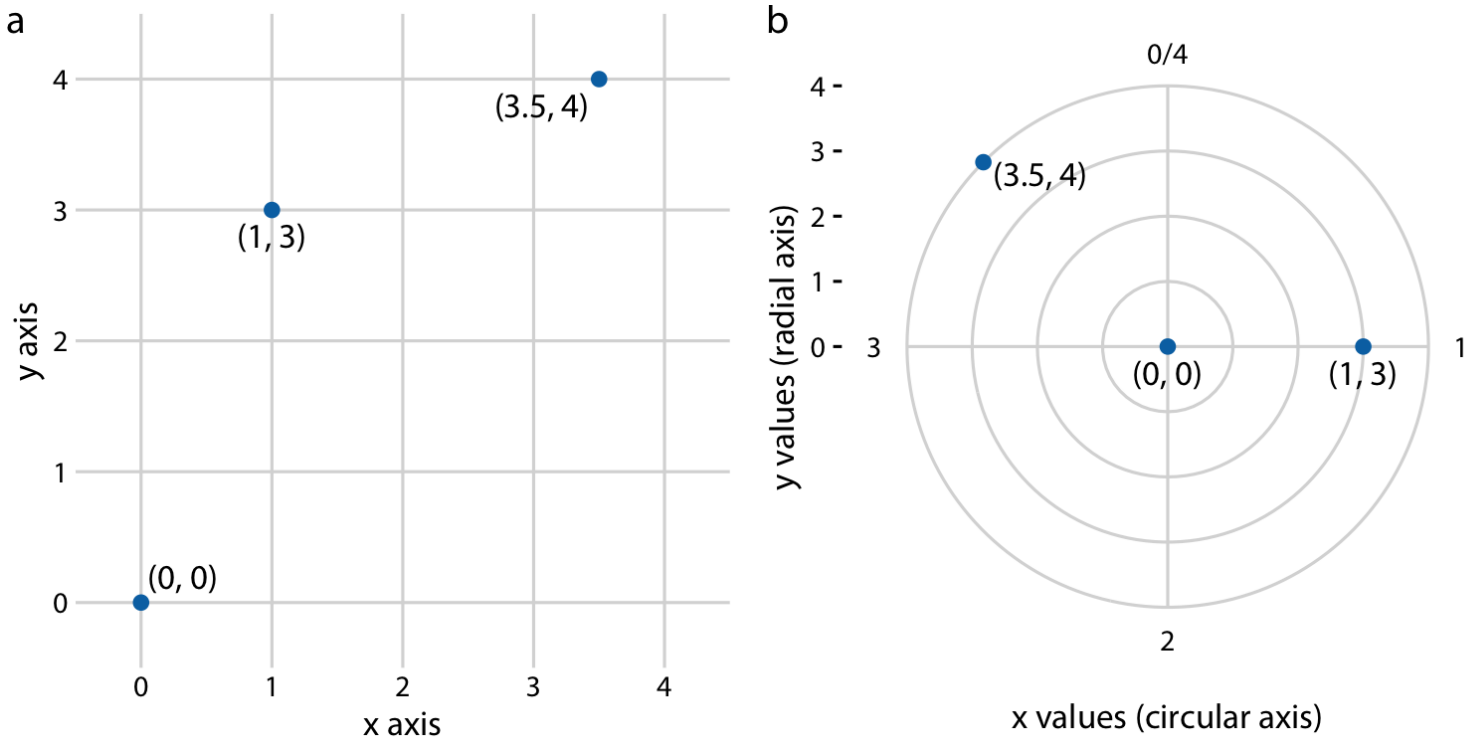


Example

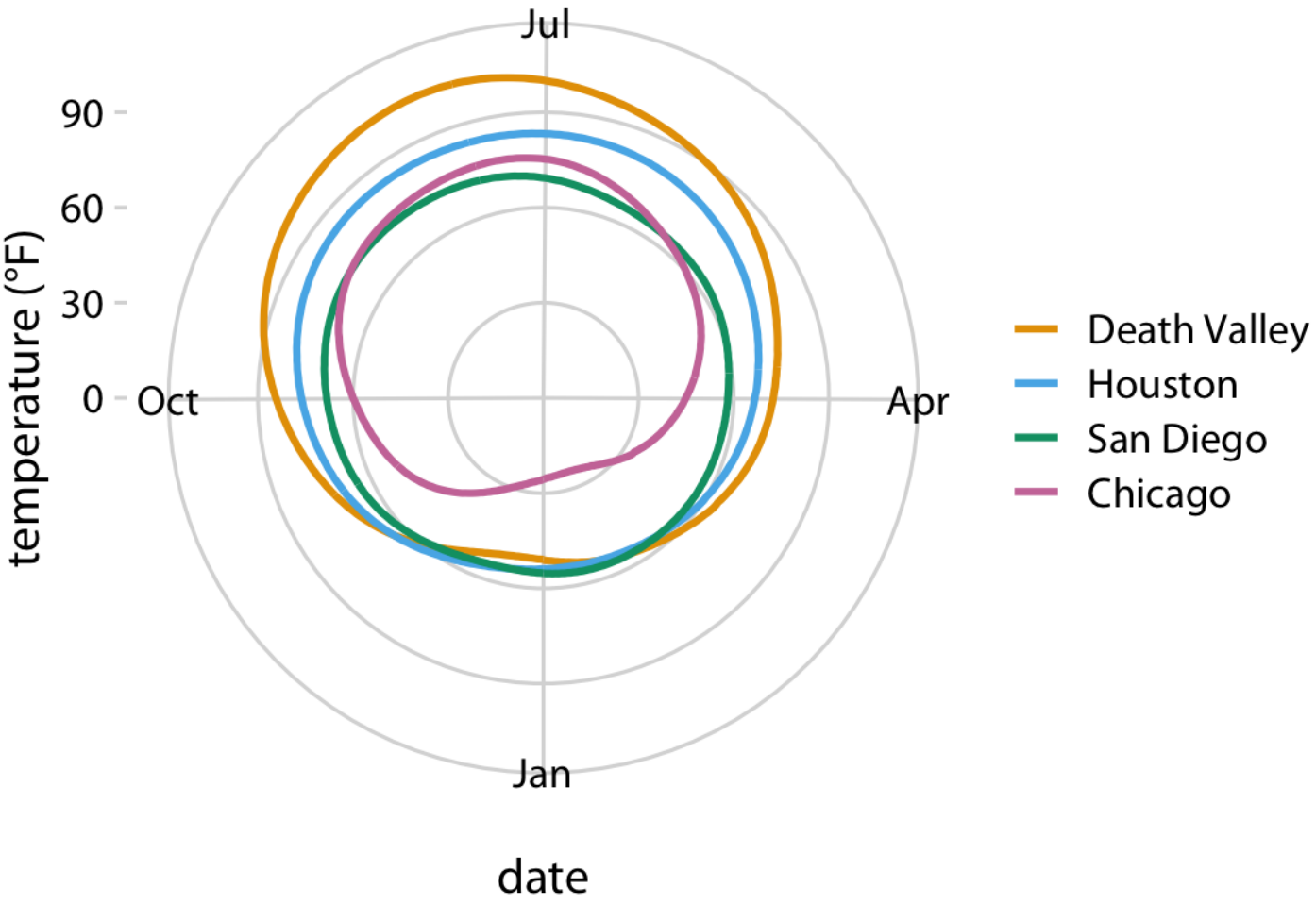

a
area (square miles)
50000
40000
30000
20000
10000
0
NY
PA
ME
MA
VT
NH
NJ
CT
RI
state


b
area (square miles)
50000
40000
30000
20000
10000
5000
1000
0
NY
PA
ME
MA
VT
NH
NJ
CT
RI
state

Coordinate Systems with Curved Axes

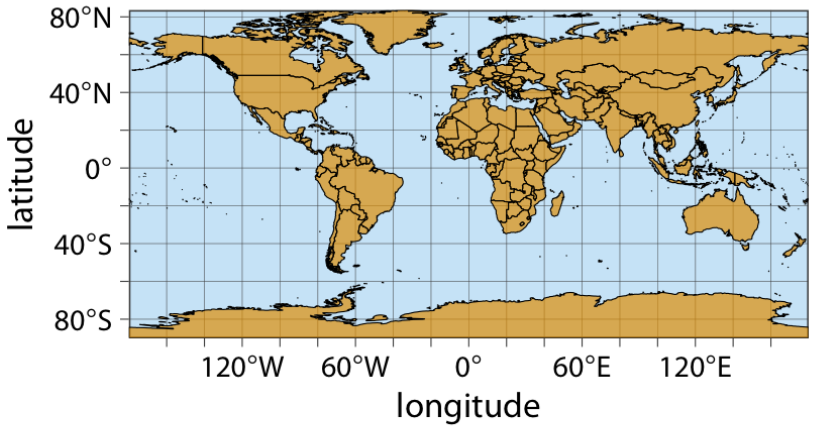


Avg. Temp Example

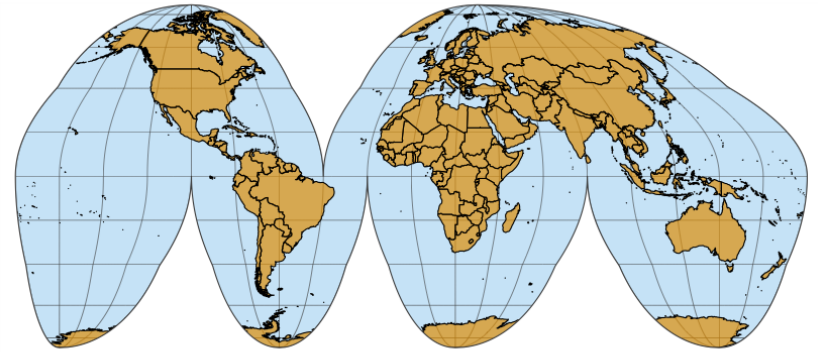


Map of the world, shown in four different projections

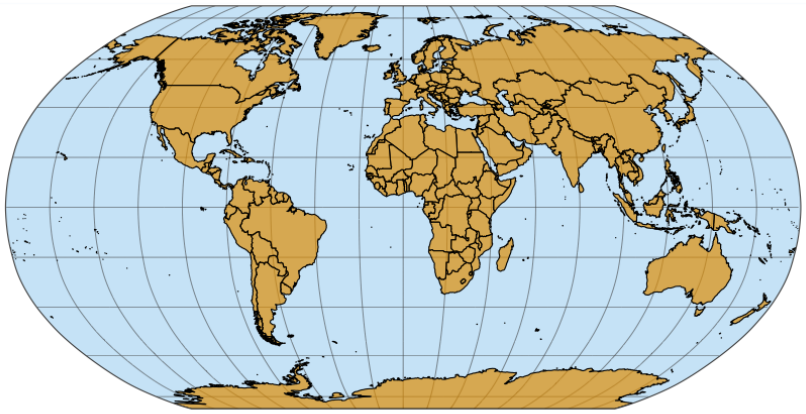
Cartesian longitude and latitude



Interrupted Goode homolosine



Robinson



Winkel tripel