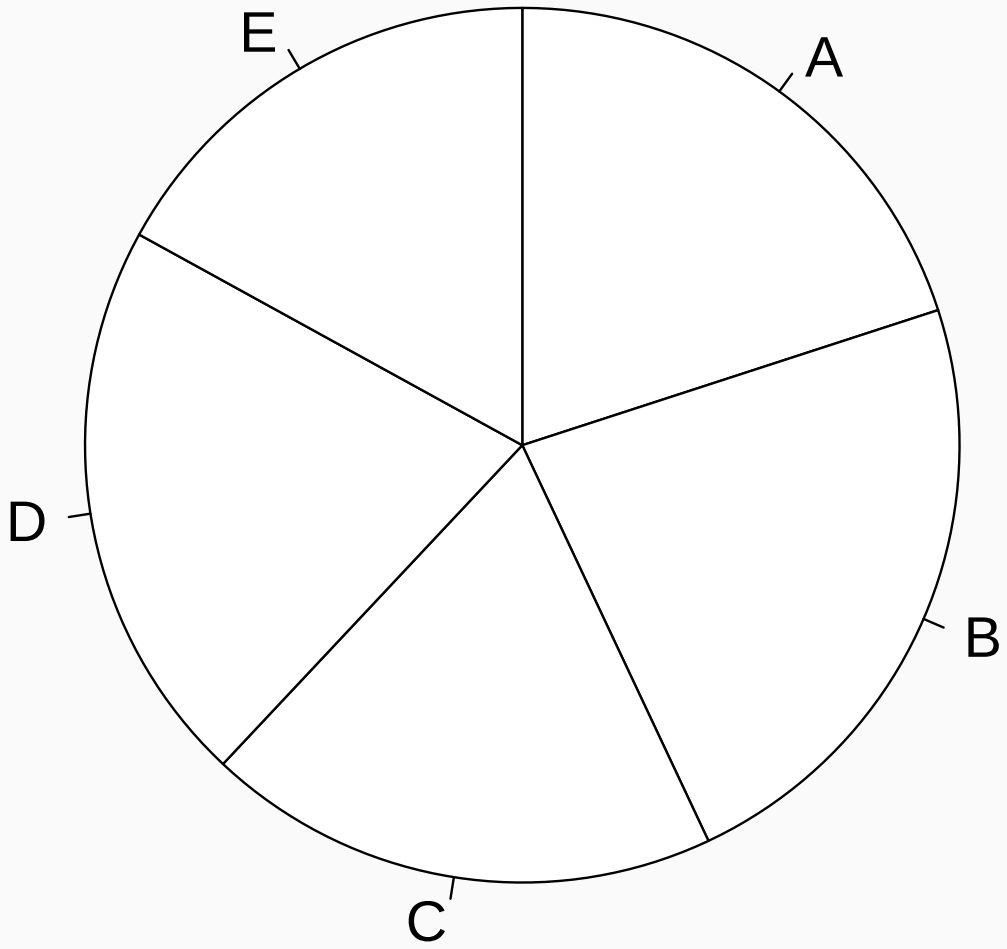
Elements of effective graphs

2019 MIDFIELD Institute

Richard Layton

June 2019

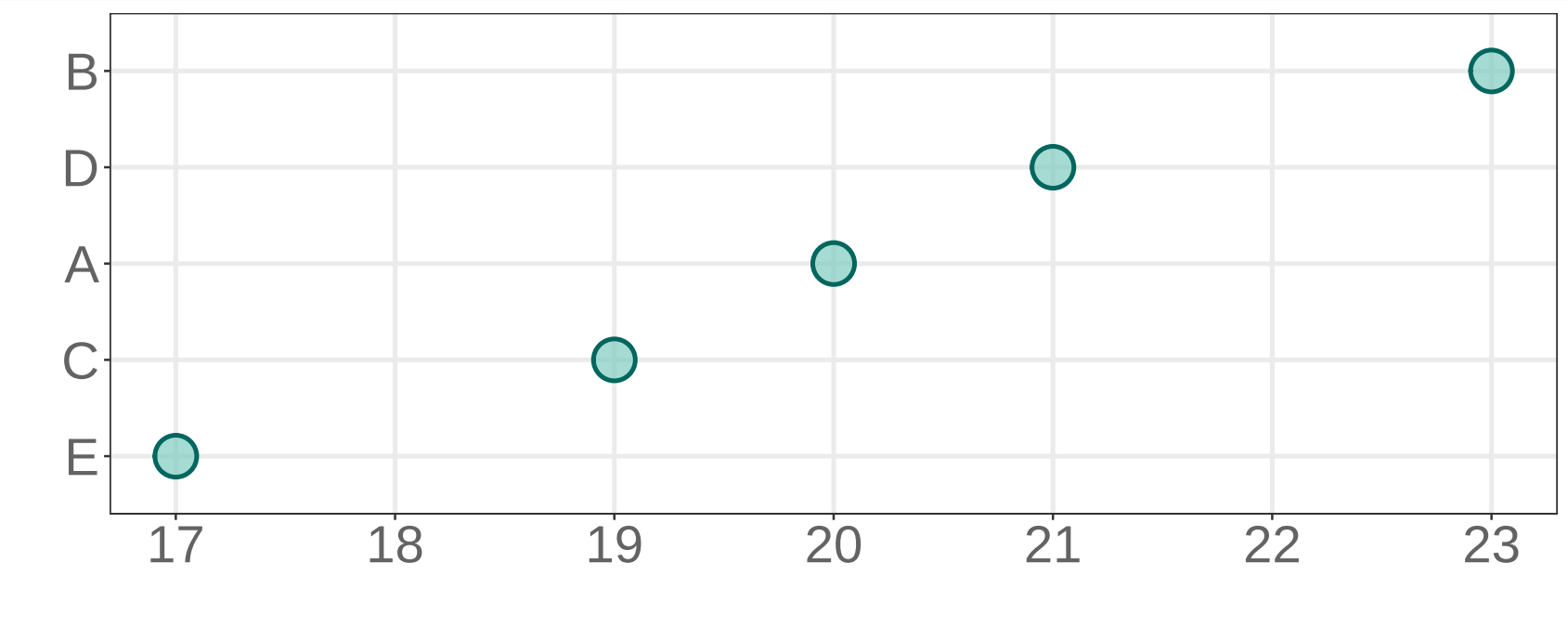
In your handout, list the slices A thru E from largest to smallest



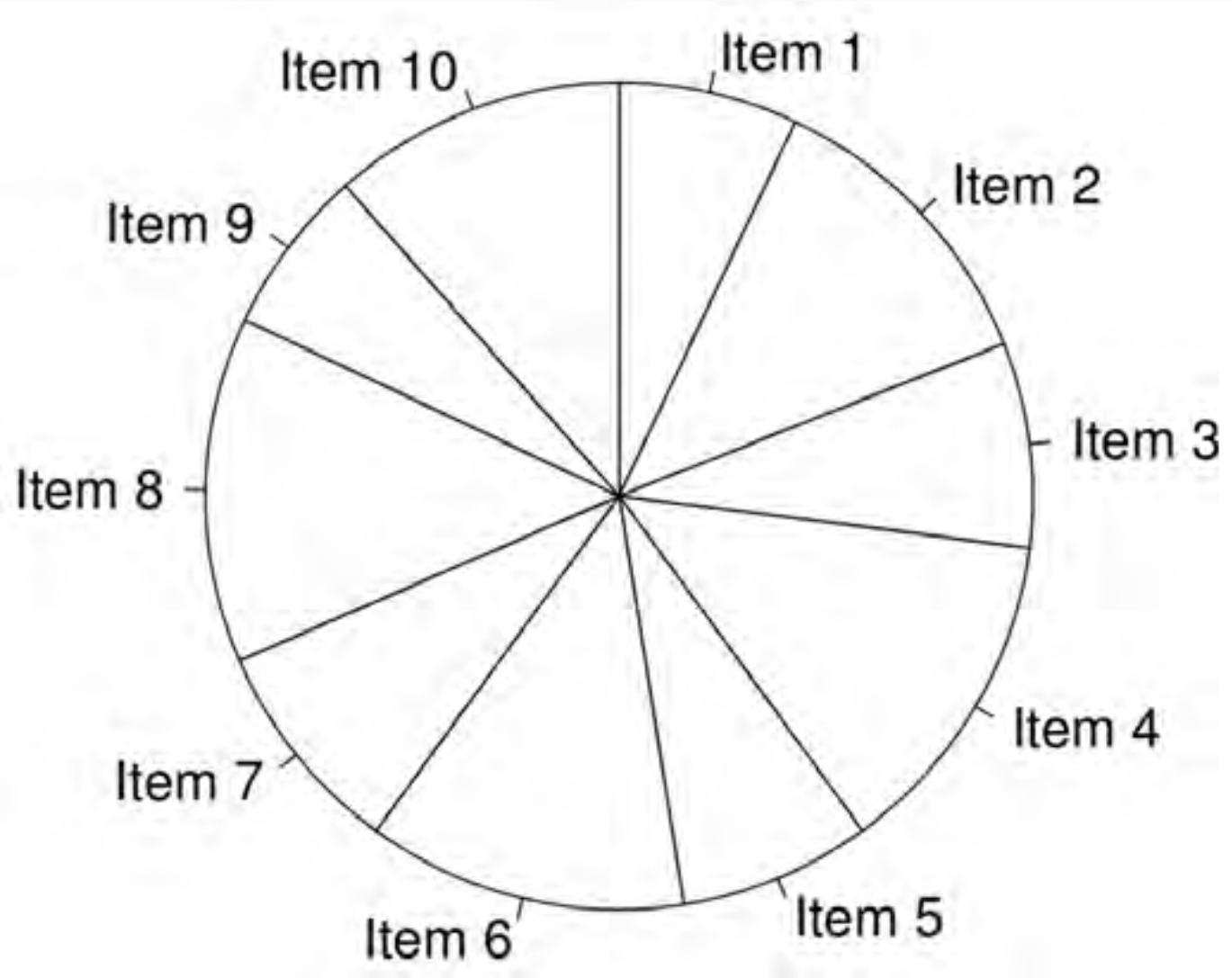
Adapted from (Robbins, 2013, Ch 2)

The same data arranged along a common axis

Comparing values along a common axis is a high-accuracy visual task.

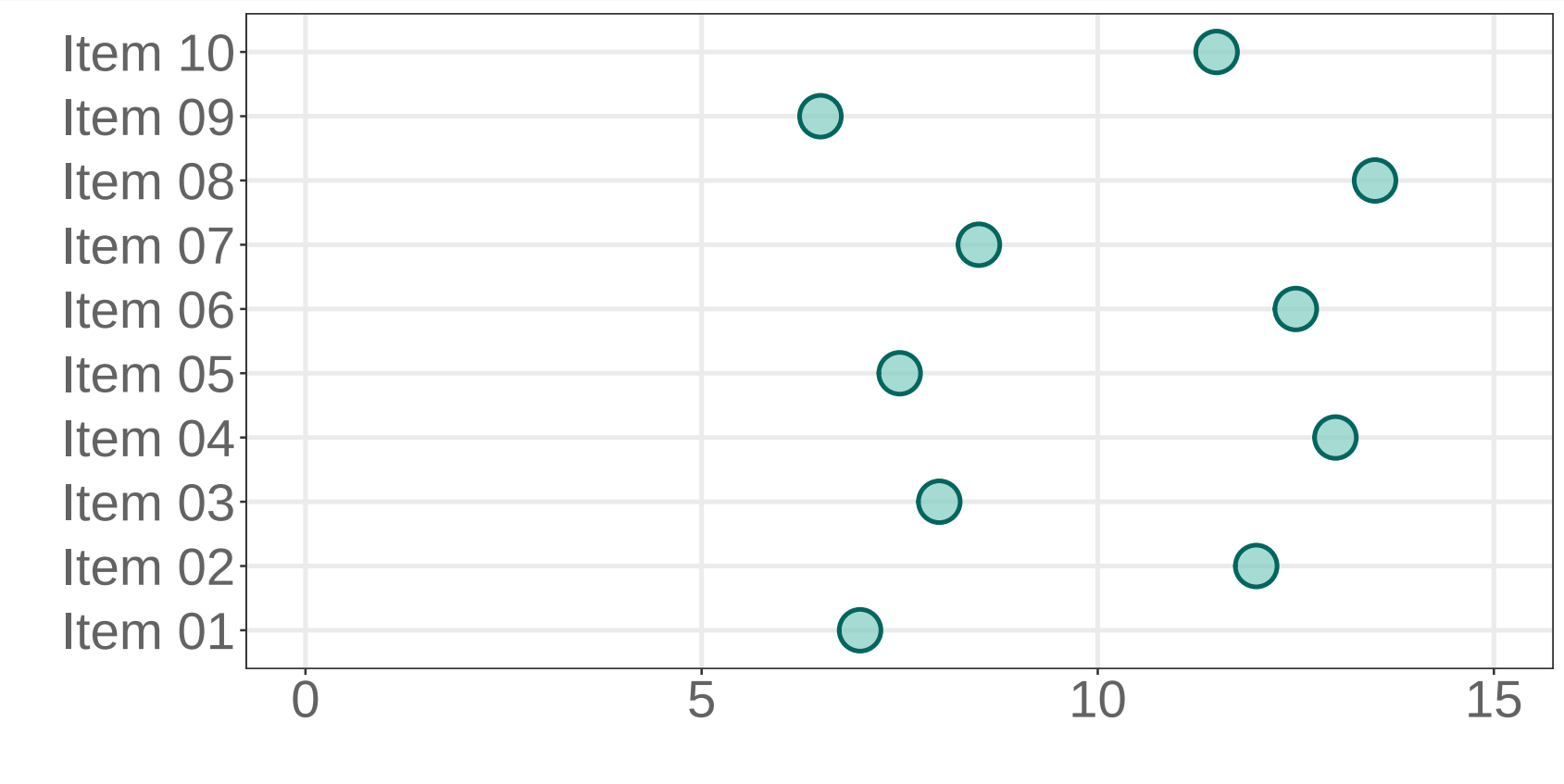


Structured data. What patterns do you see in these data?

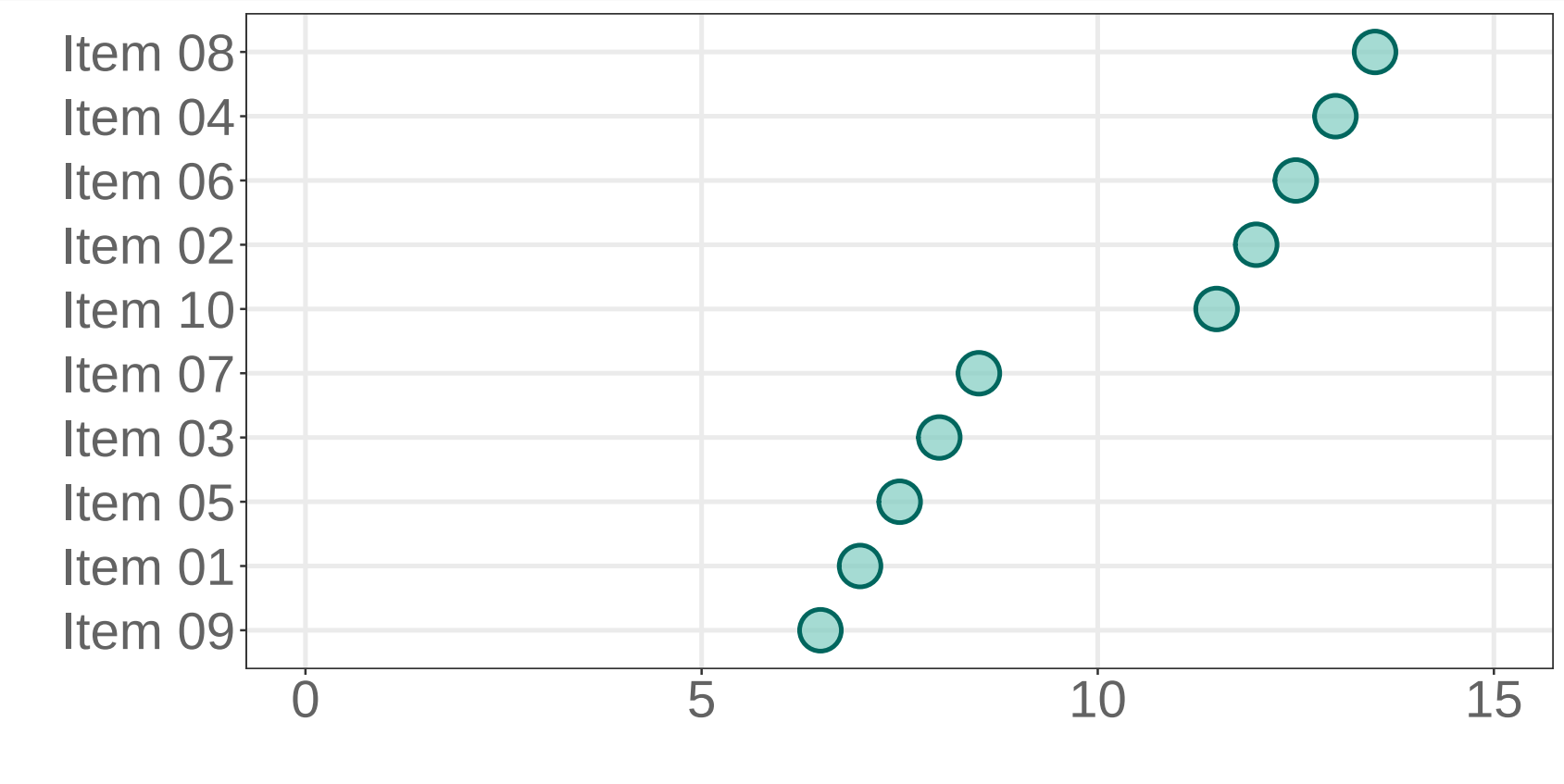


Adapted from (Robbins, 2013, Ch 2)

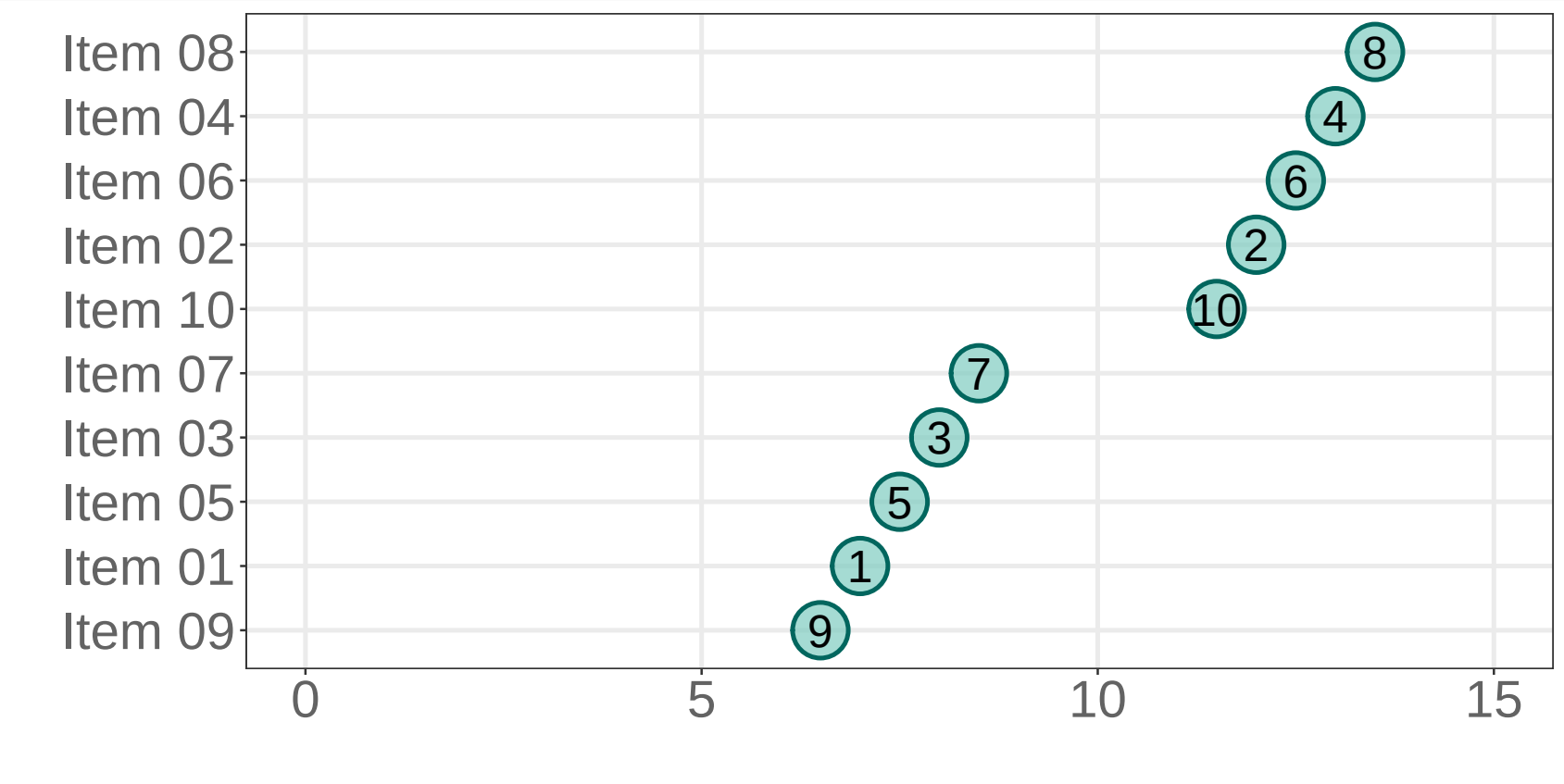
Structured data on a common scale. Any new observations?



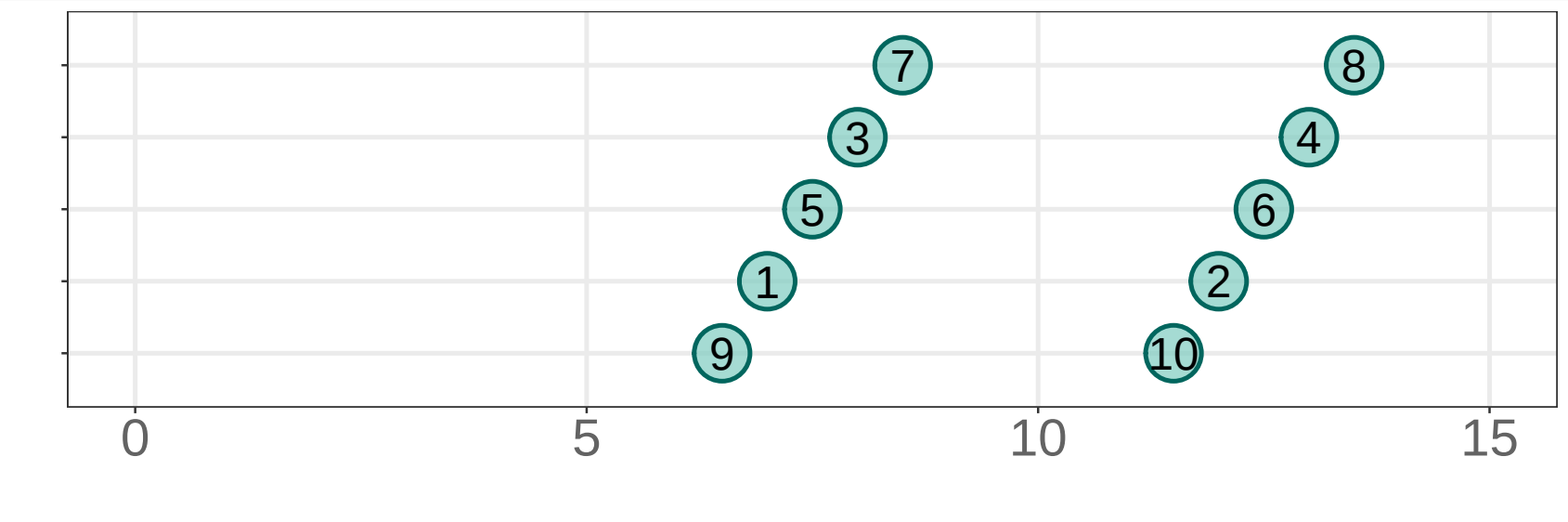
Reorder the rows by the data values. New observations?



Suppose we move the item number to the data marker.



Even-odd pairs emerge

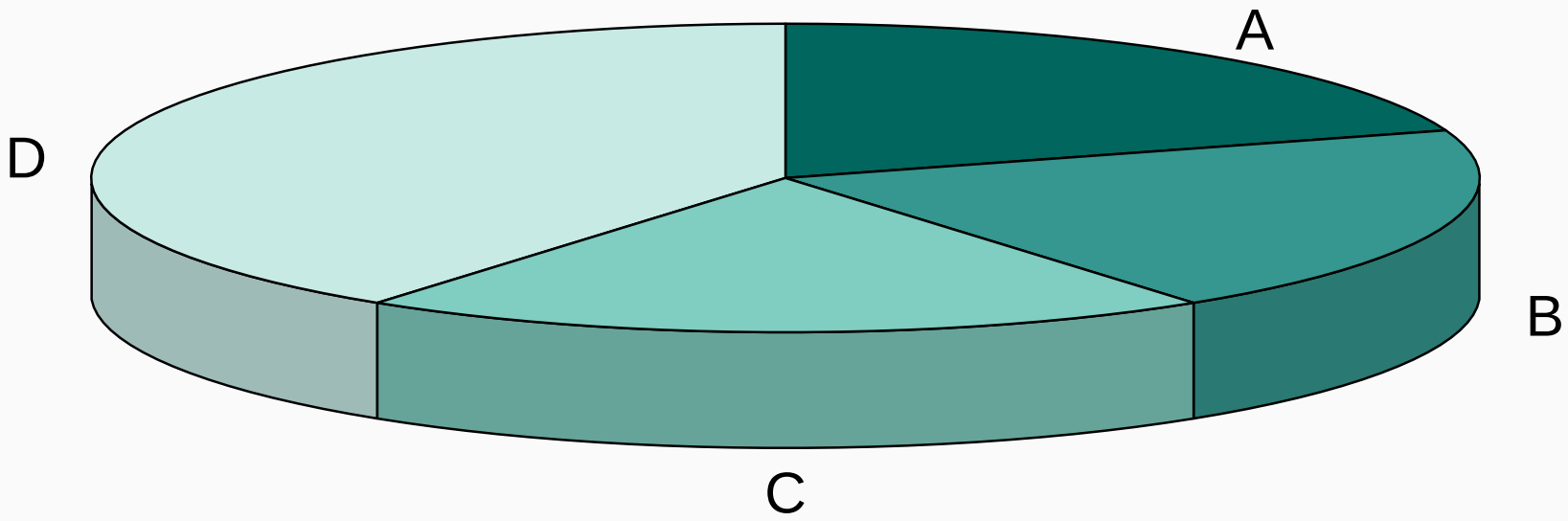


The greatest value of data visualization is when

it forces us to notice what we never expected to see.

— John Tukey (1915–2000)

Slices are what percentage of the whole?



Fill in the blanks

The total should be 100%

A. _____

B. _____

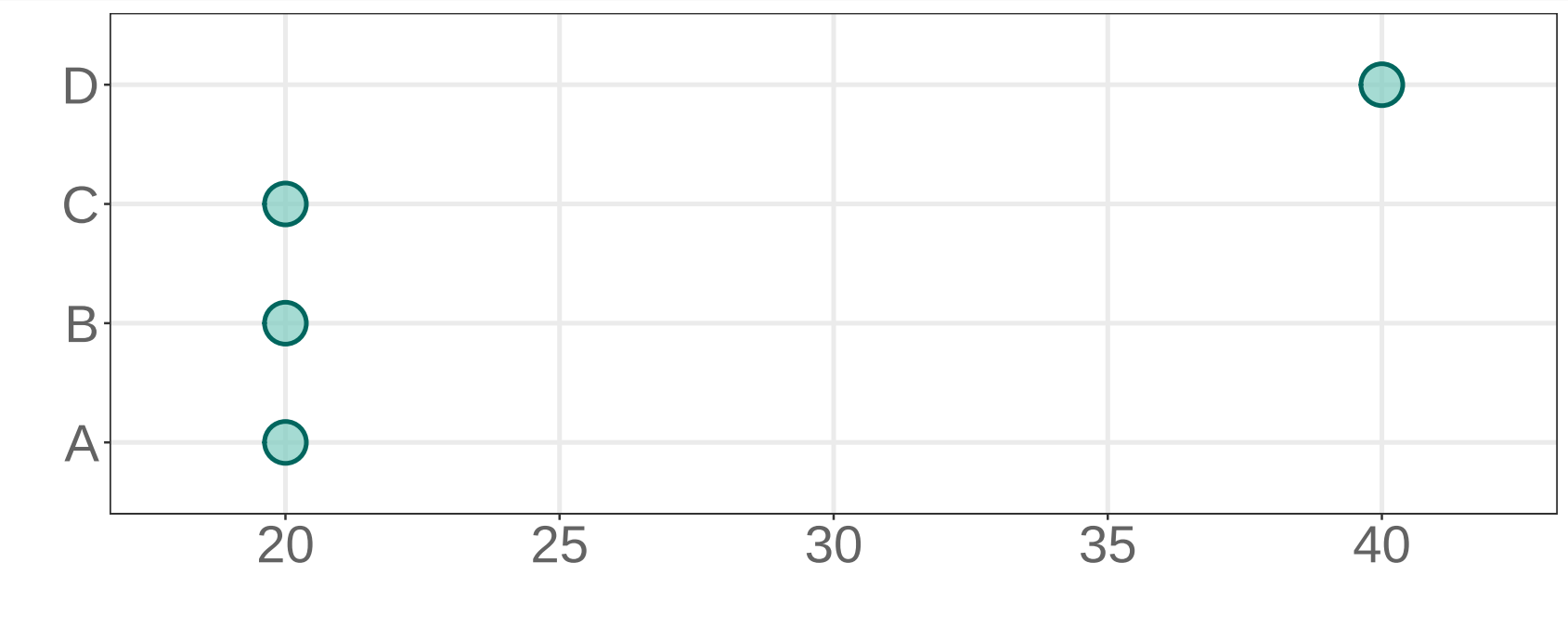
C. _____

D. _____

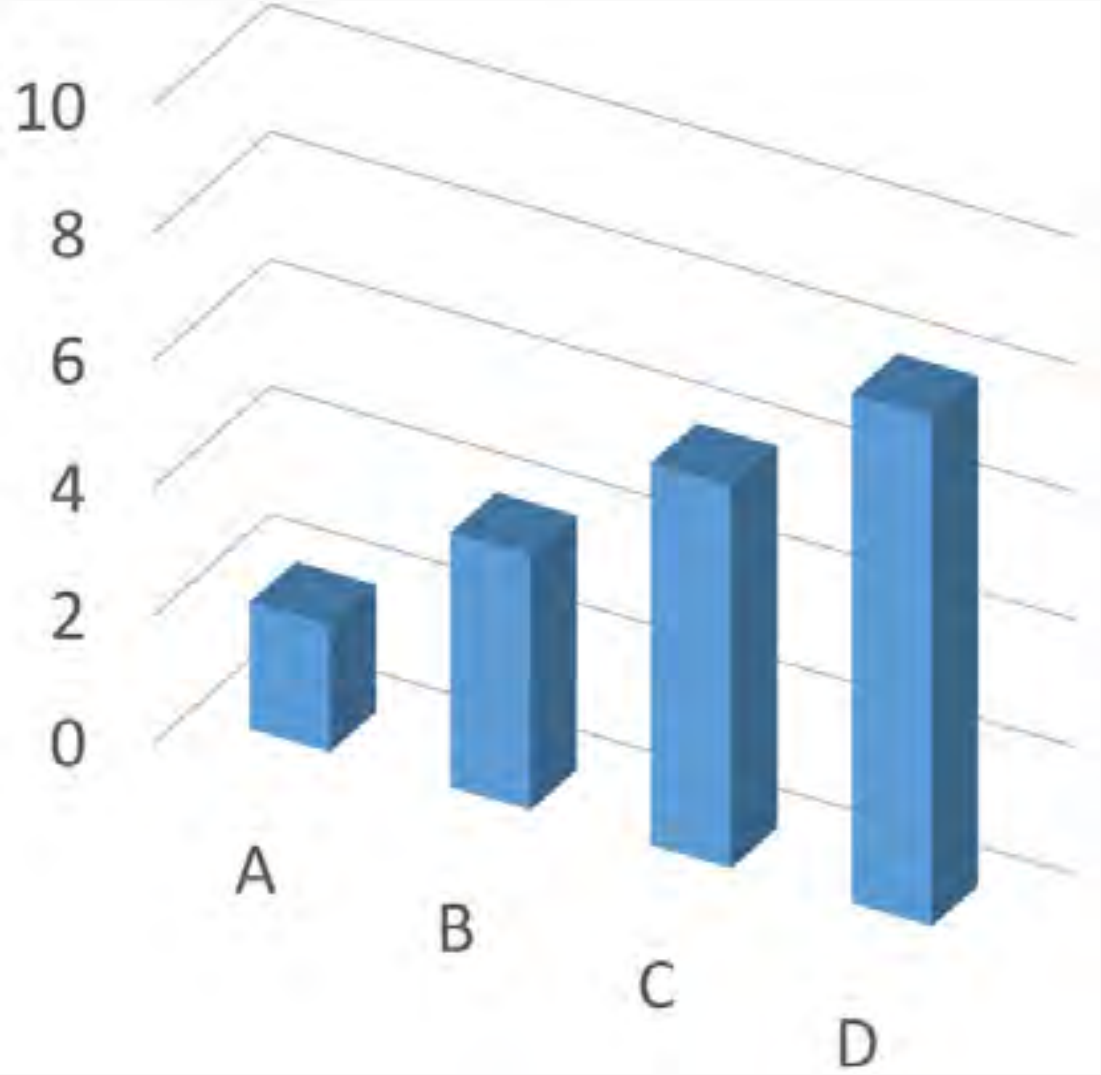
Adapted from (Robbins, 2013, Ch 2)

Again, the same data arranged along a common axis

A high-accuracy visual task.



Write down the heights of the bars



This is a visual inspection only.

Fill in the blanks

A. _____

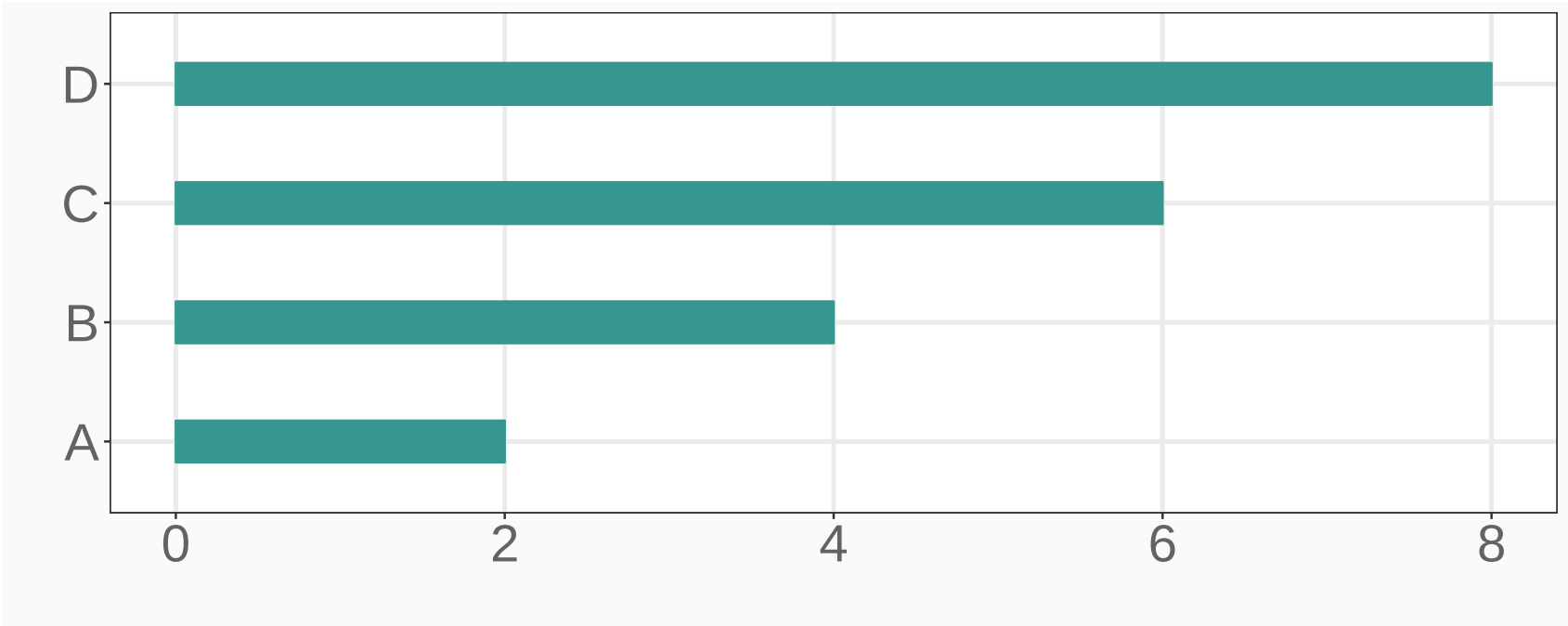
B. _____

C. _____

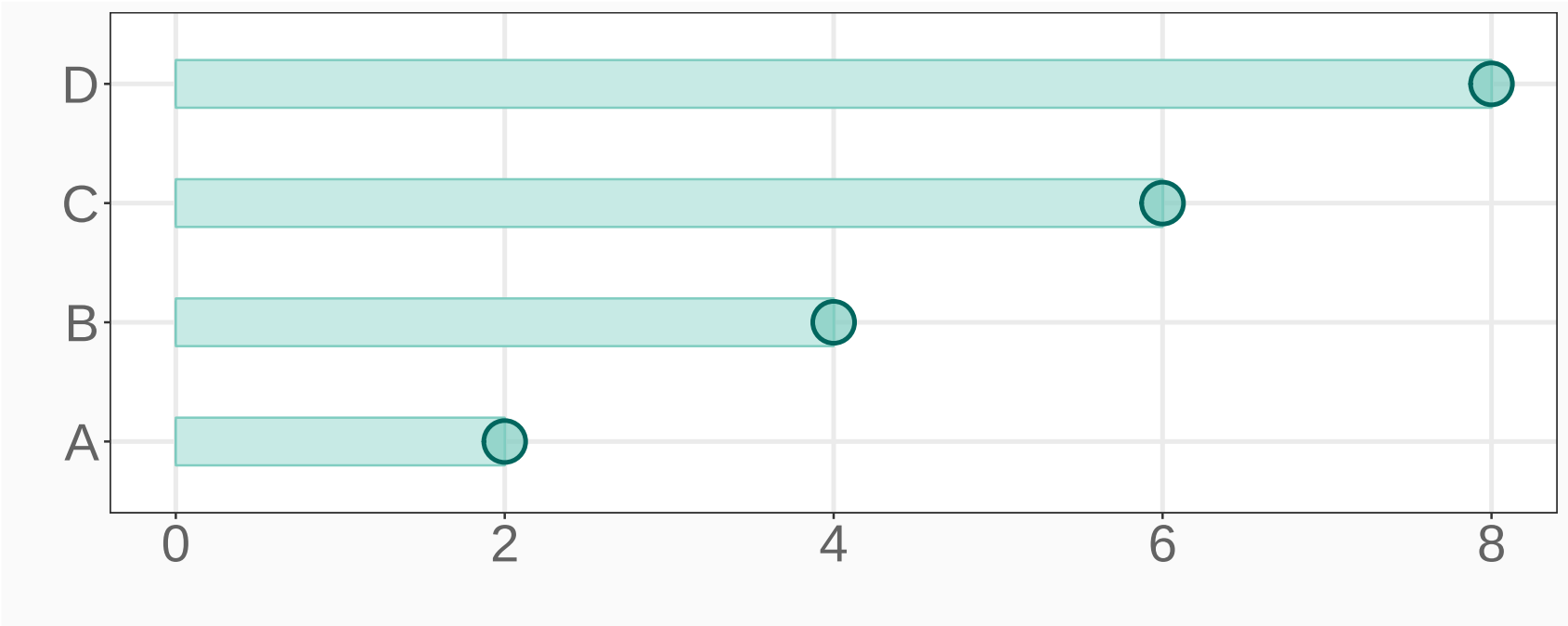
D. _____

Adapted from (Robbins, 2013, 22)

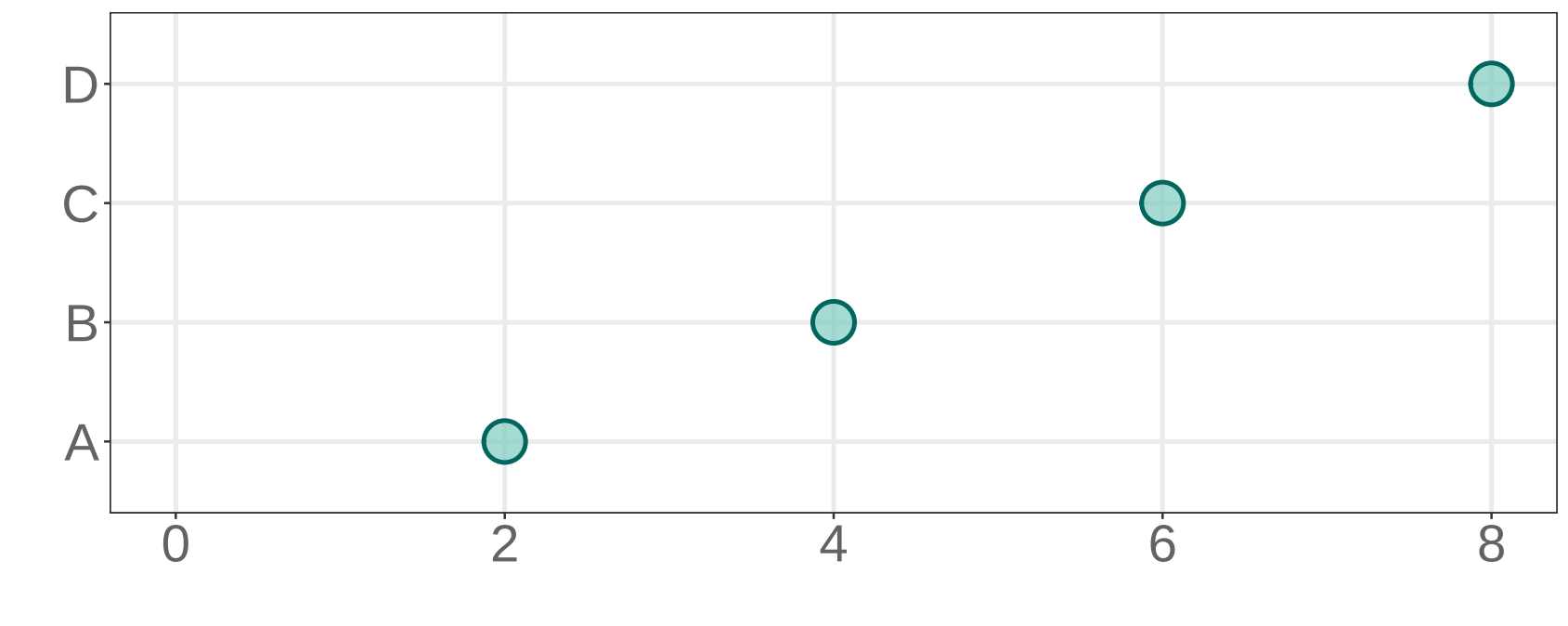
You can use bars, but must include zero



If you mark the endpoints, you can omit the bar



Producing a “dot plot” with rows ordered per the data



1st attempt: Visually estimate the state areas

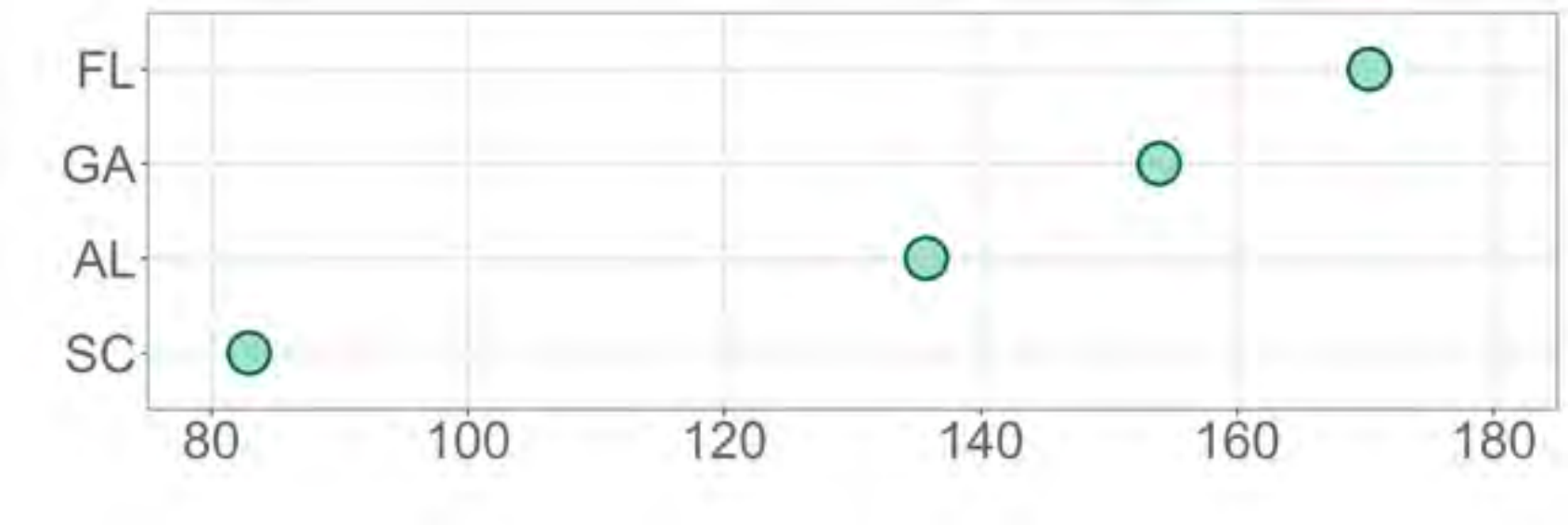
Visual estimation of area is a low-accuracy task.

South Carolina (SC) ≈ 83,000 sq km.

FL	_____	x 1000 sq. km
GA	_____	x 1000 sq. km
AL	_____	x 1000 sq. km
SC	83	x 1000 sq. km

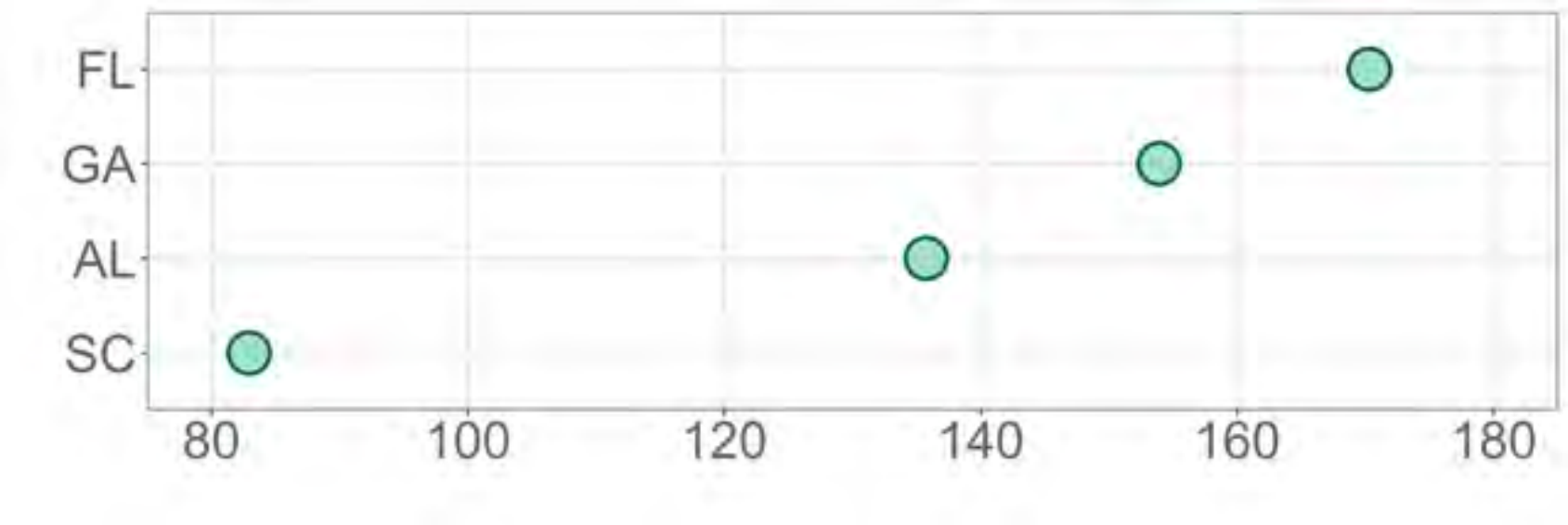
Adapted from (Ihaka, 2007)

2nd attempt: Visually estimate the state areas



FL	_____	x 1000 sq. km
GA	_____	x 1000 sq. km
AL	_____	x 1000 sq. km
SC	83	x 1000 sq. km

Your estimates have probably improved



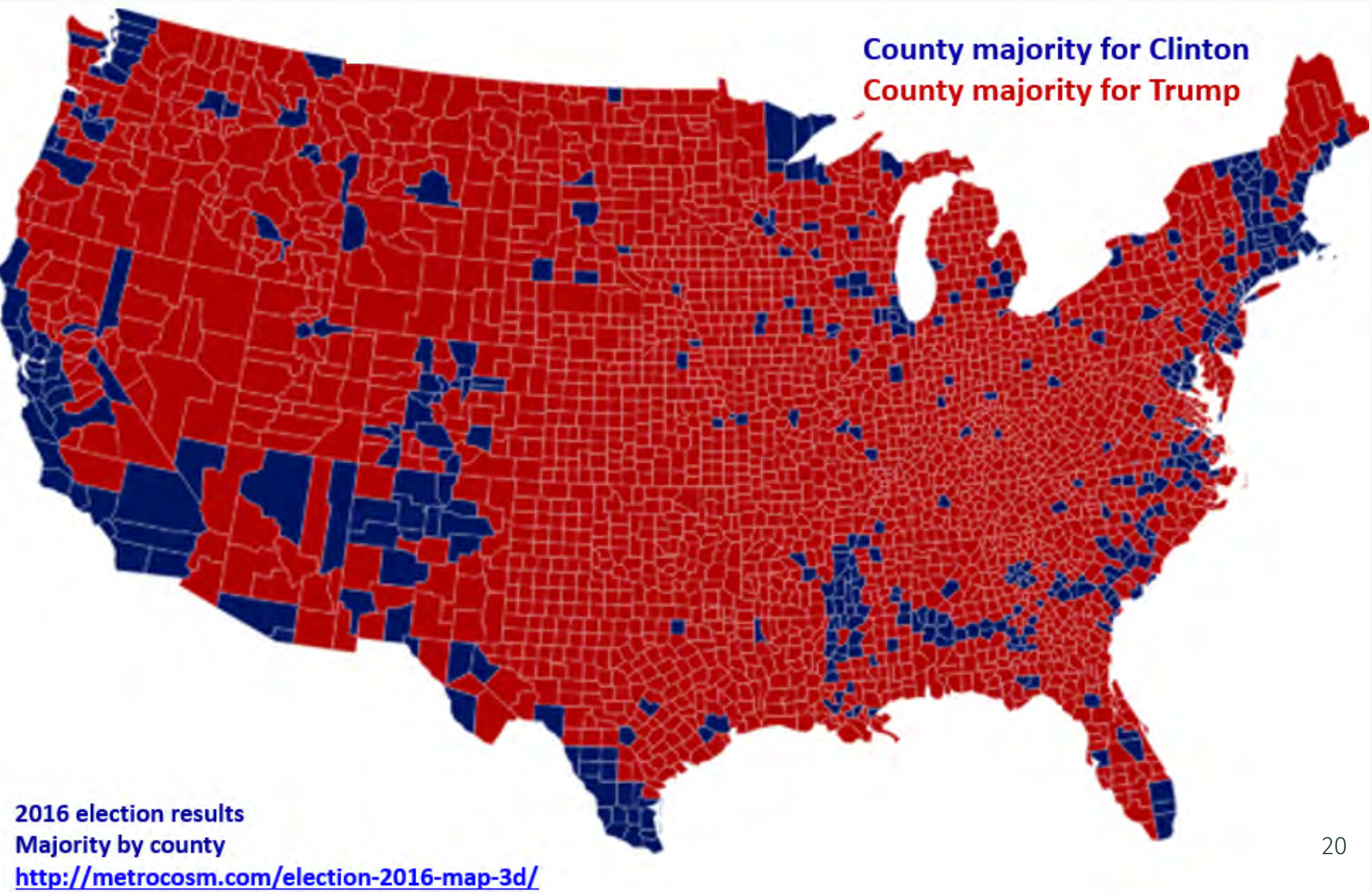
FL 170 x 1000 sq. km

GA 154 x 1000 sq. km

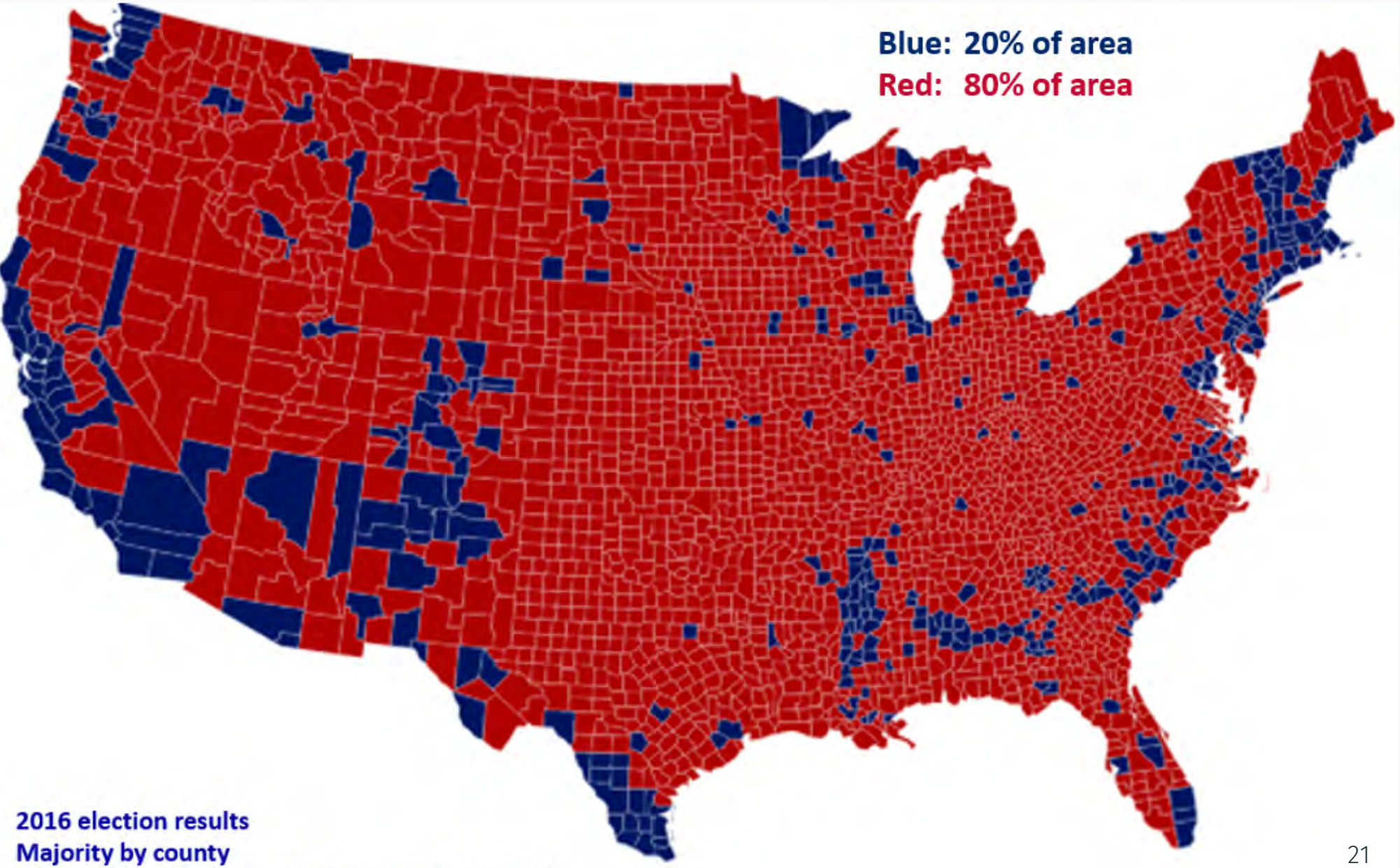
AL 136 x 1000 sq. km

SC 83 x 1000 sq. km

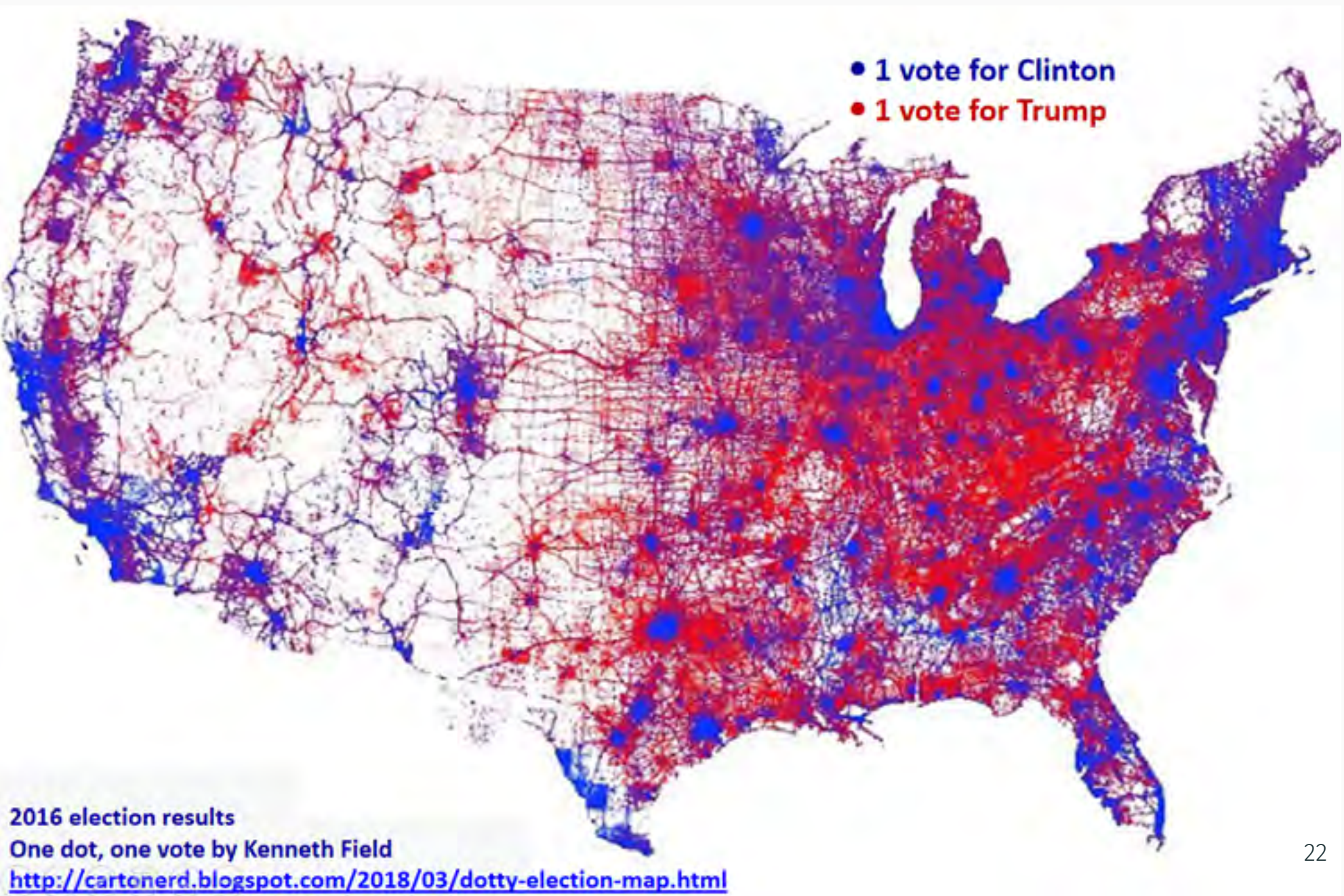
Color represents surface area. What is the visual story?



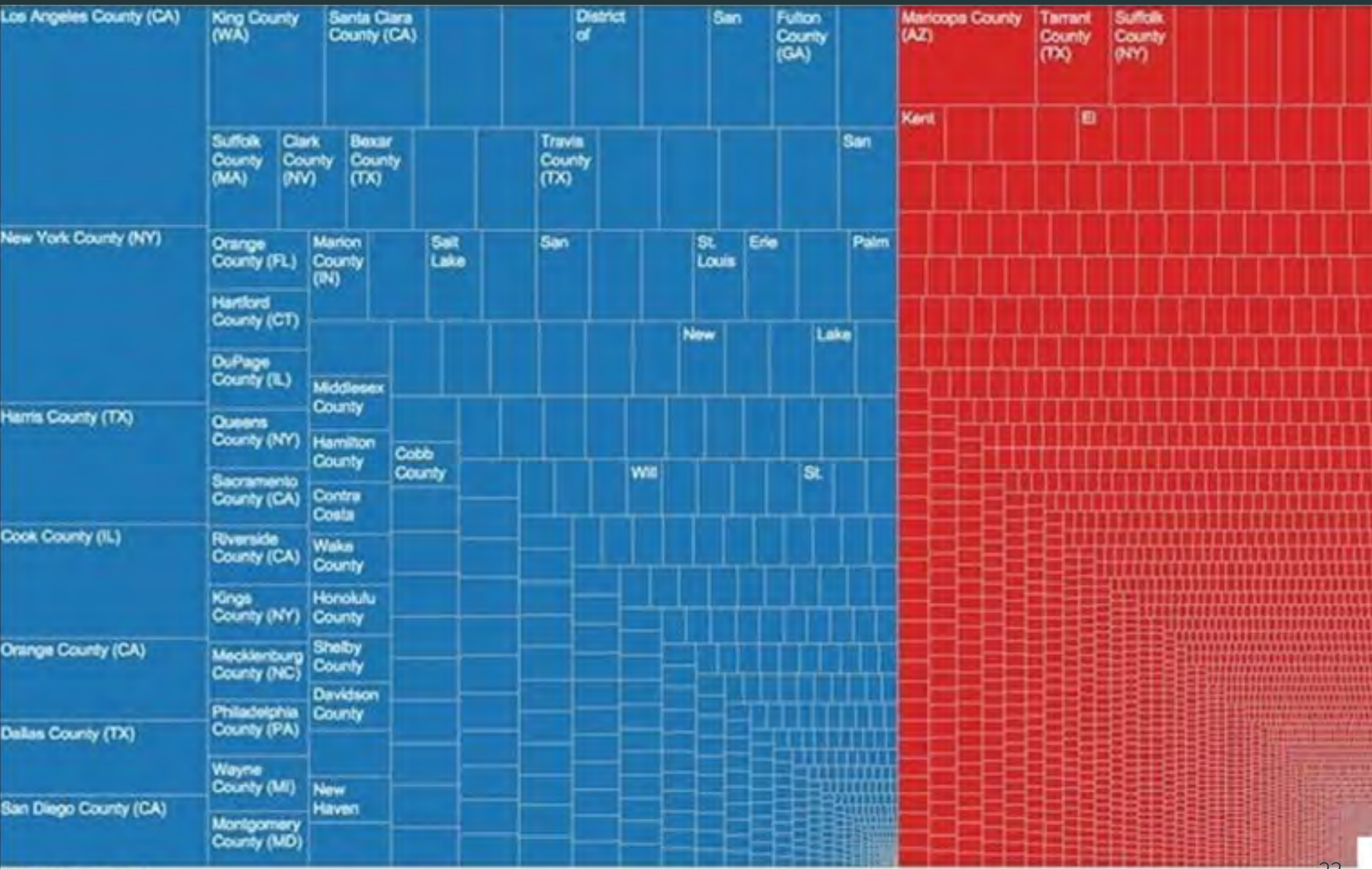
The visual ratio of surface areas $\approx 1:4$



The visual ratio of votes cast ≈ 1.05 : 1 (65.9 M to 63.0 M)

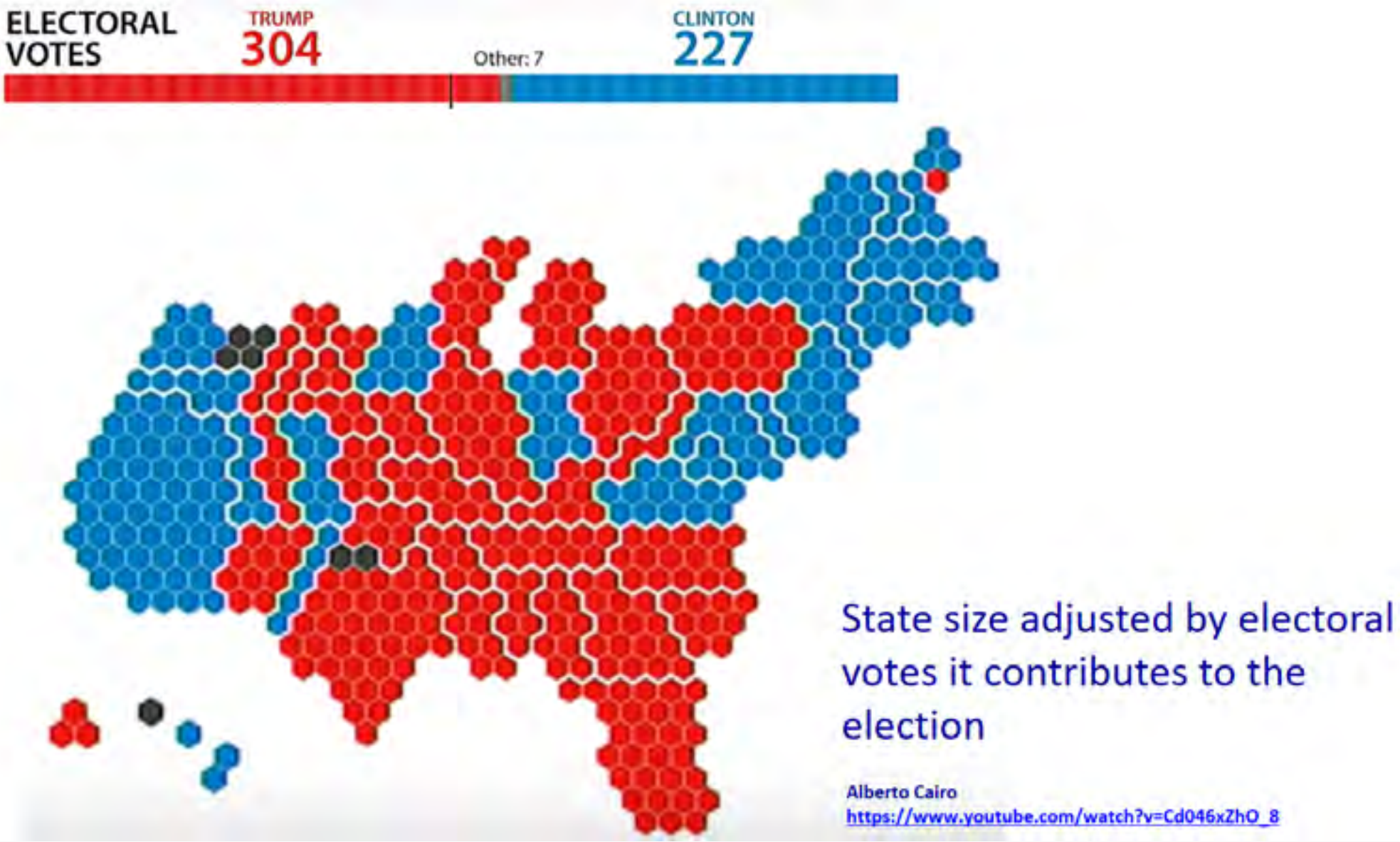


Area is county GDP. Color by party. What is the visual story?

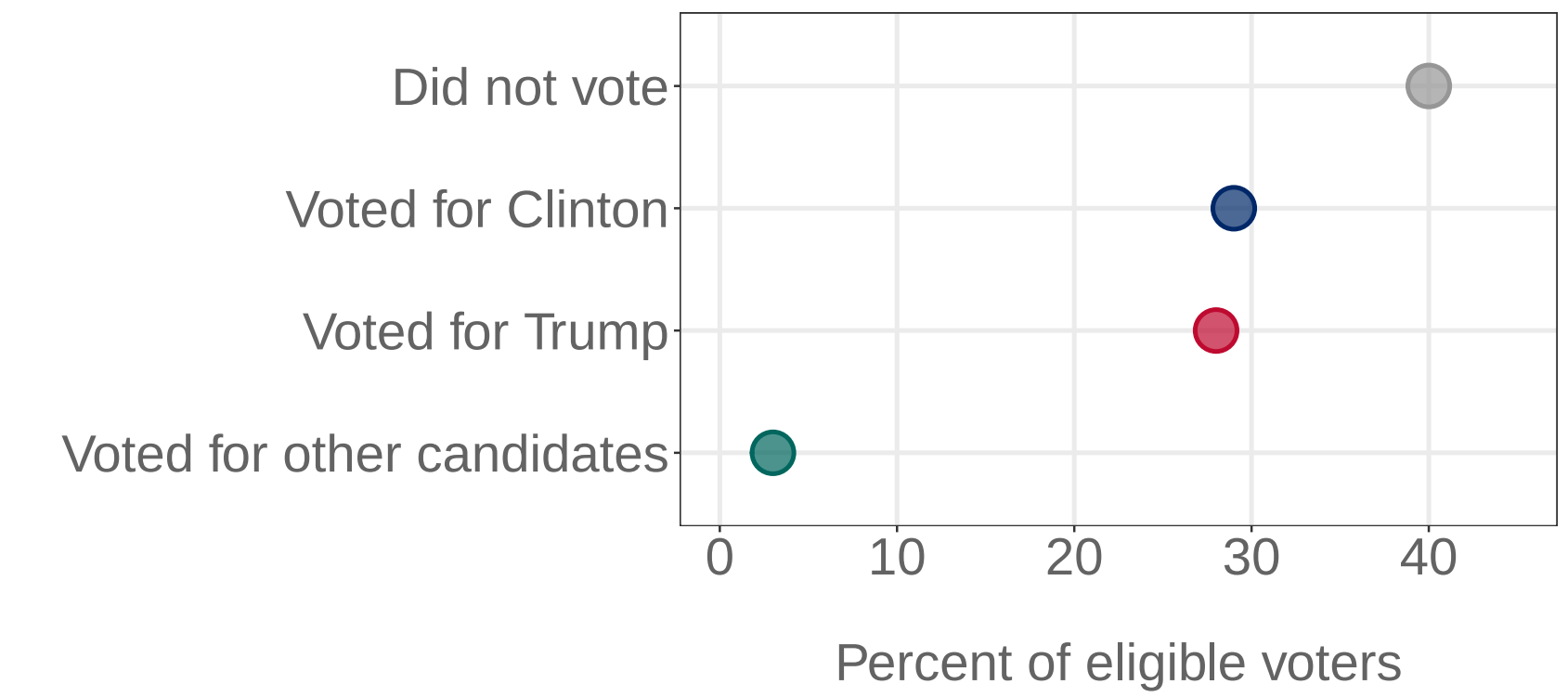


Washington Post
https://www.washingtonpost.com/news/wonk/wp/2016/11/22/donald-trump-lost-most-of-the-american-economy-in-this-election/?utm_term=.c4759565da83

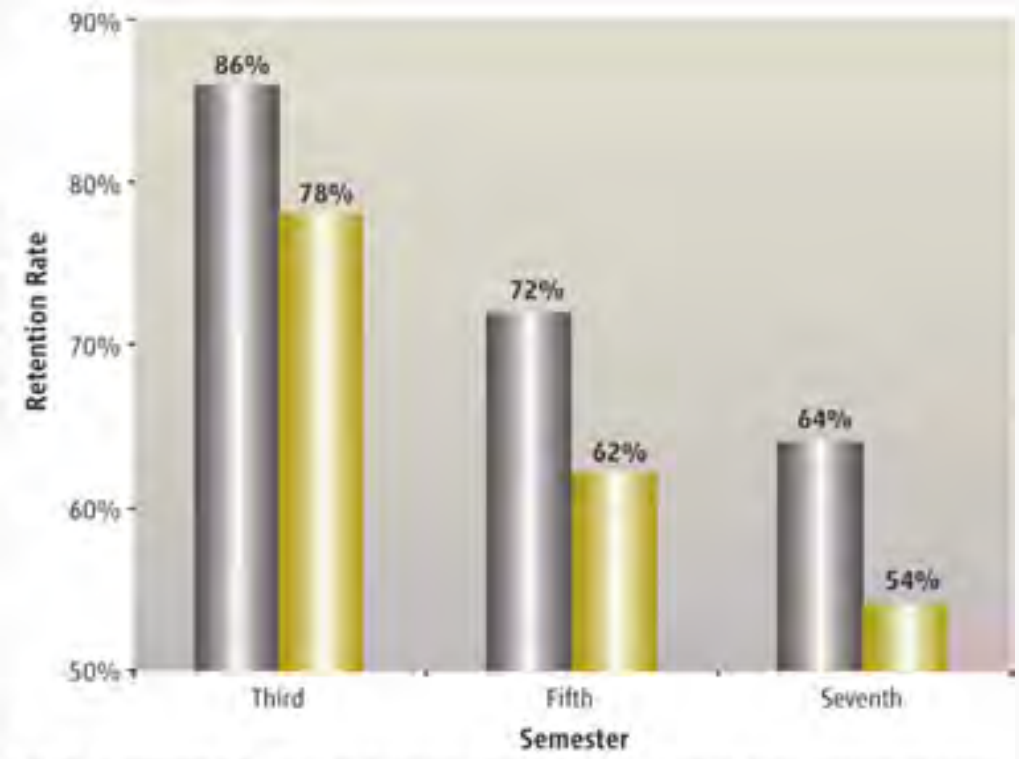
The metric that matters. What is the visual story?



The previous graphs conceal what story ?



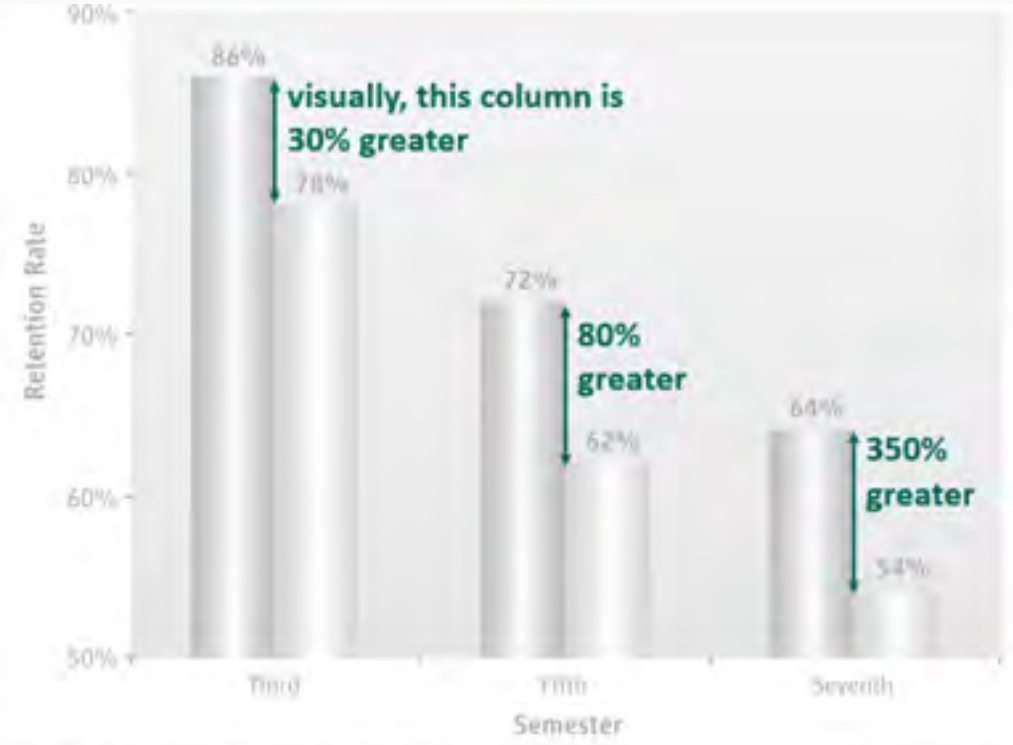
The trouble with bars. What's wrong with this graph?



Gains in retention. The FYEP course improved retention of engineering students into the third, fifth, and seventh semester. There were 2128 students who took the FYEP course (gray) and 2942 students who did not (gold). All retention gains over expected retention rates shown are significant ($P < 0.05$).

Norman L. Fortenberry, Jacquelyn F. Sullivan, Peter N. Jordan, and Daniel W. Knight (2007), Engineering education research aids instruction, *Science*, 31:1175–1176.

A nearly constant difference seems to increase in significance

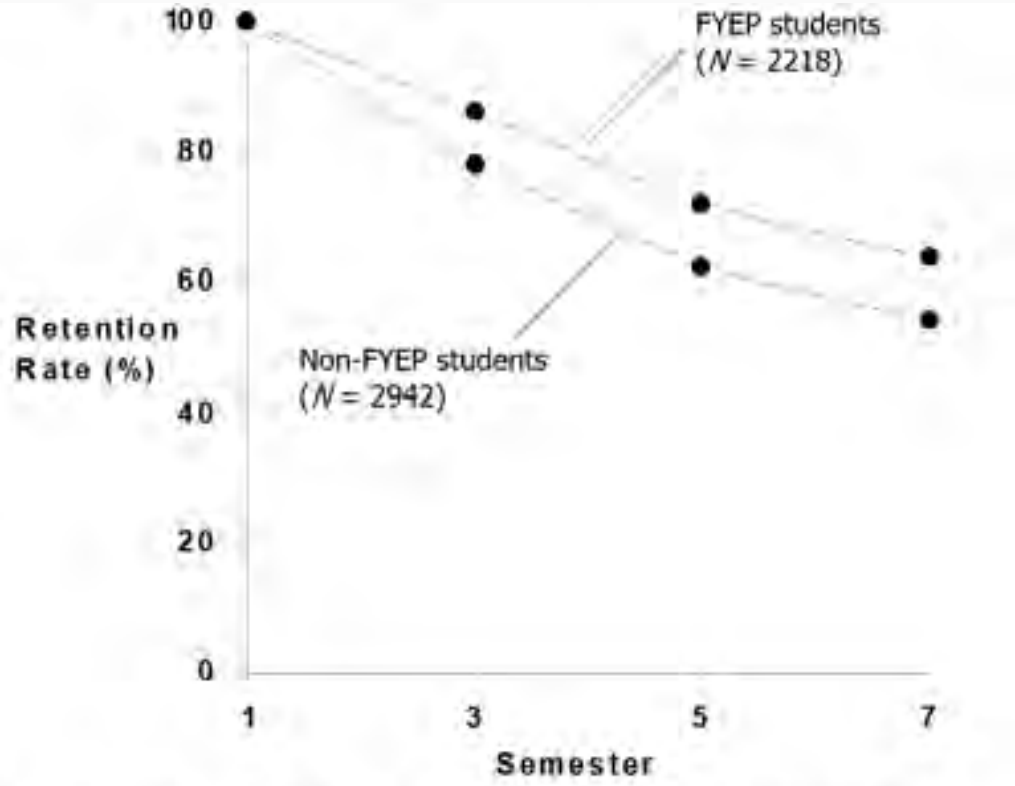


Gains in retention. The FYEP course improved retention of engineering students into the third, fifth, and seventh semester. There were 2128 students who took the FYEP course (gray) and 2942 students who did not (gold). All retention gains over expected retention rates shown are significant ($P < 0.05$).

- Can you identify the missing information?

Norman L. Fortenberry, Jacquelyn F. Sullivan, Peter N. Jordan, and Daniel W. Knight (2007), Engineering education research aids instruction, *Science*, 31:1175–1176.

Redesigned, with full scales, a different story emerges



First-year gains in retention. The primary impact of the first-year engineering projects (FYEP) course is in the higher retention rate in the third semester. Subsequently, both groups lose students at about the same rate with a persistent 10% difference between FYEP and non-FYEP students.

- FYEP impact is in the first year.
- Attrition rate afterwards is about the same for both groups.

The experts tell us

(Cairo, 2018)

Image from https://tinyurl.com/y5g7jbzt

"Graphicacy" is as important as numeracy to the modern educated citizen

Be aware that we all like charts that pander to our expectations or biases

The experts tell us

(Doumont, 2009)

Image from http://www.principiae.be/pdfs/Principiae-2014.pdf

Optimal design primarily depends on

- The message to be conveyed
- The variables to be shown

The experts tell us

(Tufte, 1983)

Image from https://en.wikipedia.org/wiki/Edward_Tufte

The task of the designer is to give visual access to the subtle and the difficult — that is, reveal the complex.

The experts tell us

What's your point?

Seriously, that's the most important question.

(Evergreen, 2017)

Image from https://tei.cgu.edu/people/stephanie-evergreen-phd/

References

Cairo A (2018) Ihaka lectures—Visual trumpery: How charts lie and how they make us smarter. https://tinyurl.com/y5g7jbzt

Doumont J-l (2009) *Trees, maps, and theorems: Effective communication for rational minds.* Principiae, Kraainem, Belgium http://www.treesmapsandtheorems.com/

Evergreen S (2017) *Effective Data Visualization: The Right Chart for the Right Data.*, 1st edn. Sage, Thousand Oaks, CA

Ihaka R (2007) Statistics 787 Lecture slides. https://tinyurl.com/yywbpuas

Kostelnick C and Hassett M (2003) *Shaping Information: The Rhetoric of Visual Conventions.* Southern Illinois University

Robbins N (2013) *Creating More Effective Graphs.* Chart House, Wayne, NJ

Tufte E (1983) *The visual display of quantitative information.* Graphics Press, Cheshire, CT https://www.edwardtufte.com/tufte/books_vdqi

Wainer H (2000) *Visual revelations: Graphical tales of fate and deception from Napoleon Bonaparte To Ross Perot.* Psychology Press, Mahwah, NJ