

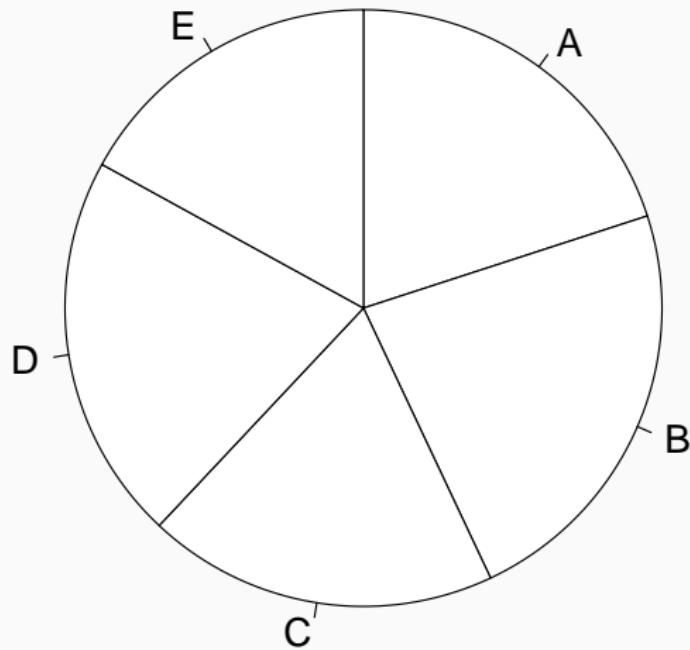
Elements of effective graphs

2019 MIDFIELD Institute

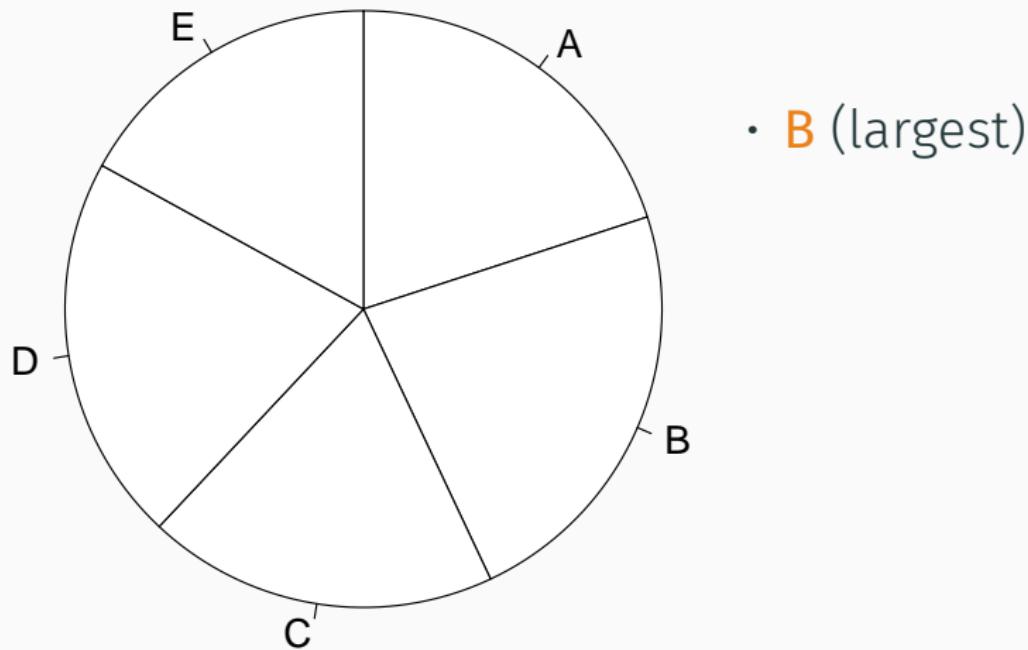
Richard Layton

June 2019

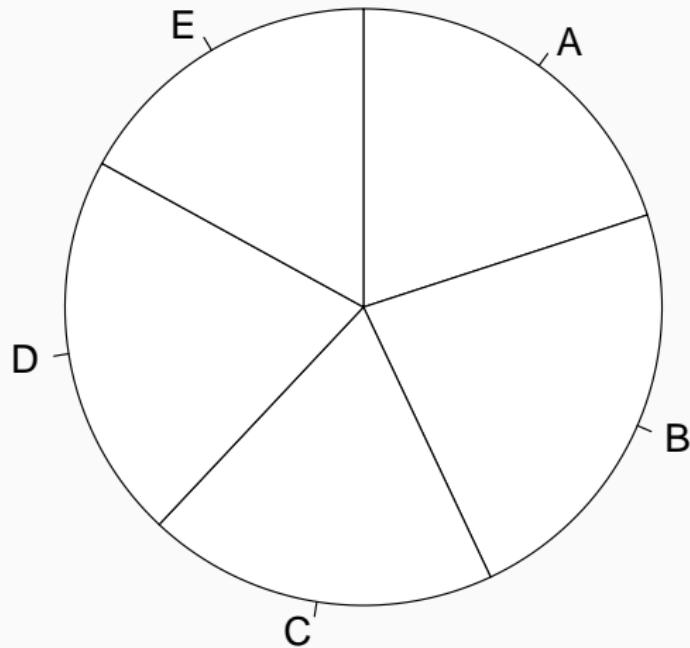
In your handout, list the slices A thru E from largest to smallest



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



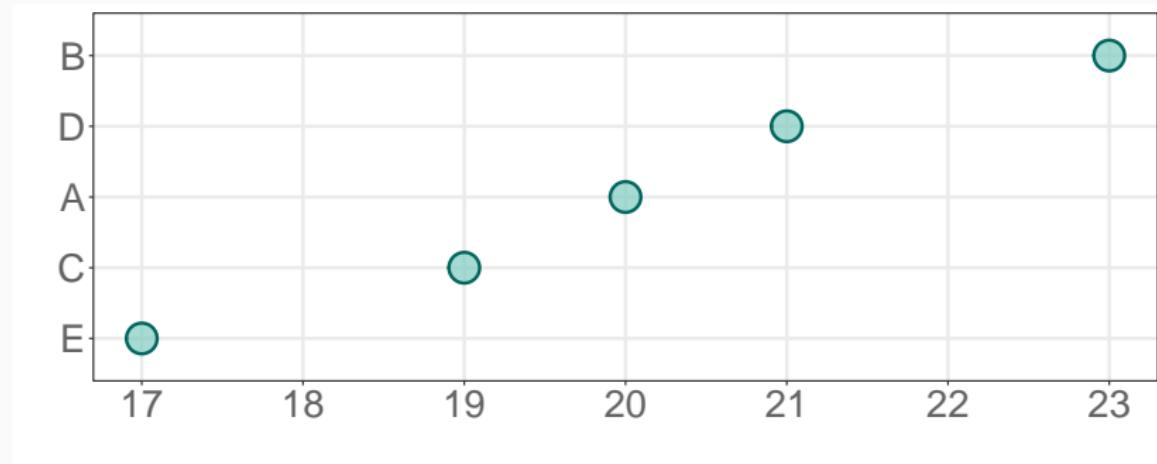
In your handout, list the slices A thru E from largest to smallest



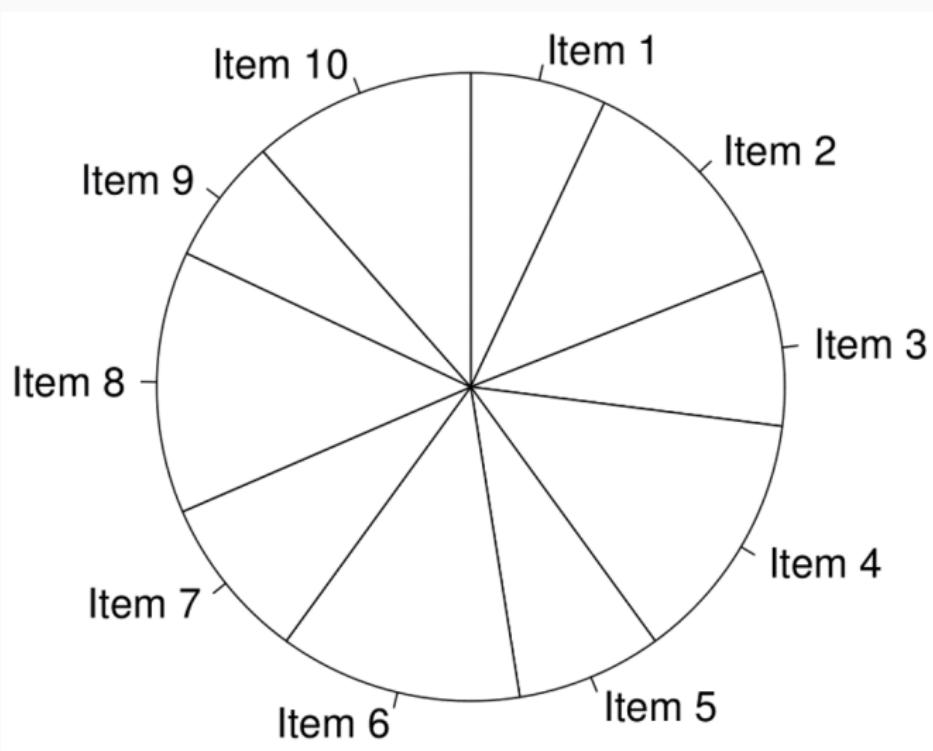
- B (largest)
- D
- A
- C
- E (smallest)

The same data arranged along a common axis

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

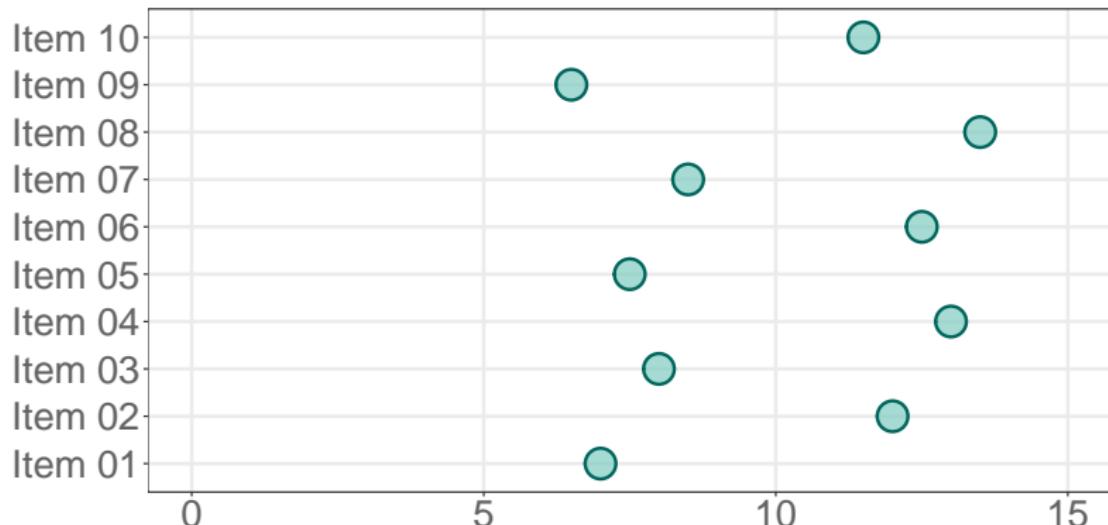


Structured data. What patterns do you see in these data?

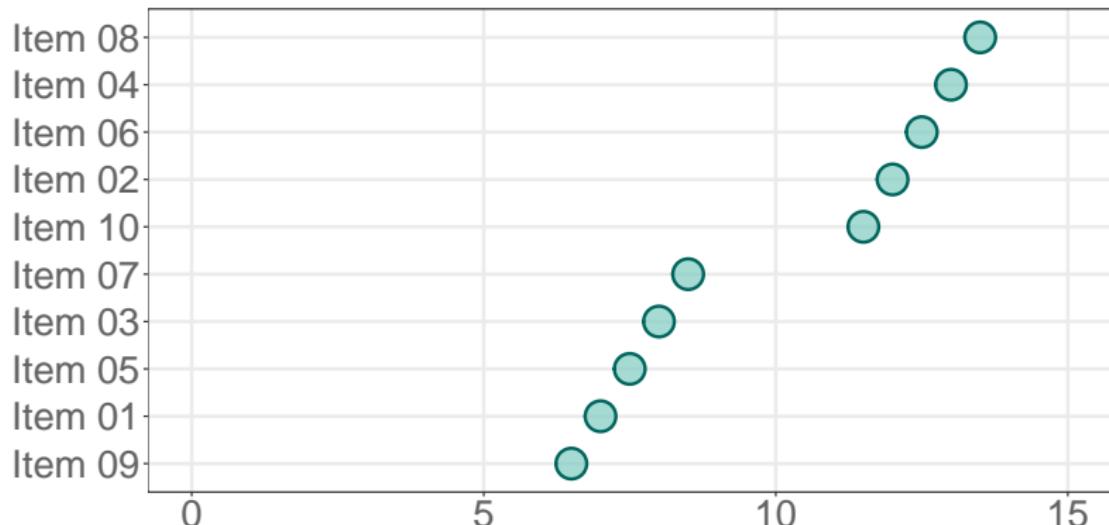


Adapted from (Robbins 2013, Ch 2)

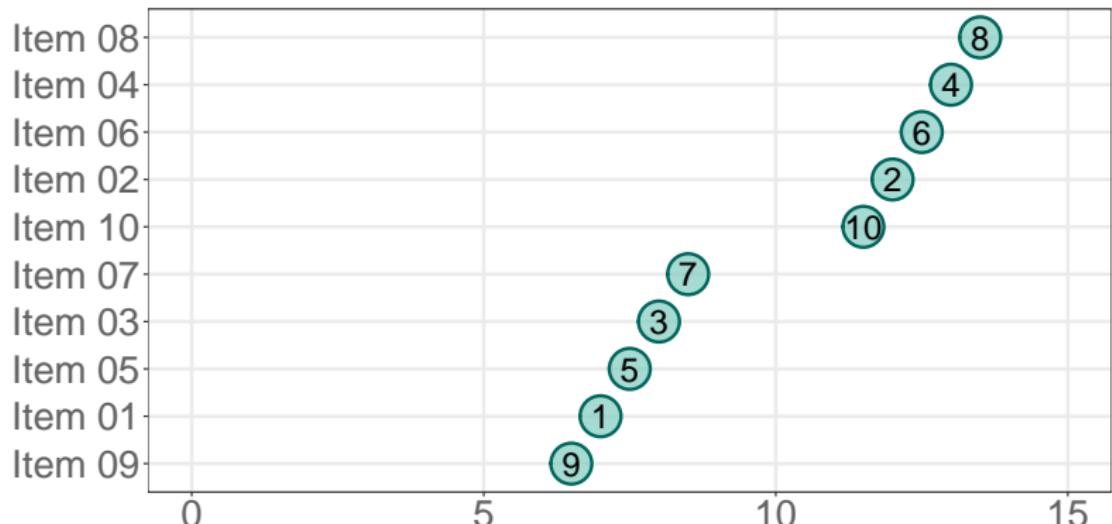
Structured data on a common scale. Any new observations?



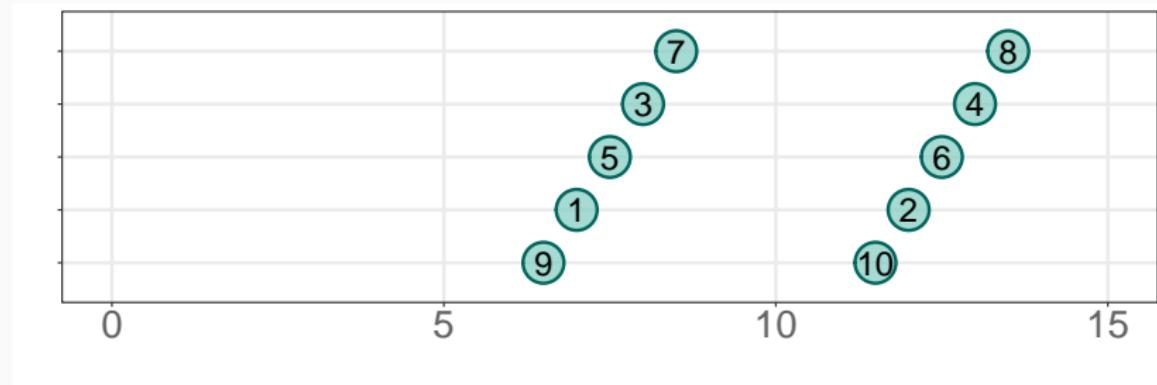
Reorder the rows by the data values. New observations?



Suppose we move the item number to the data marker.



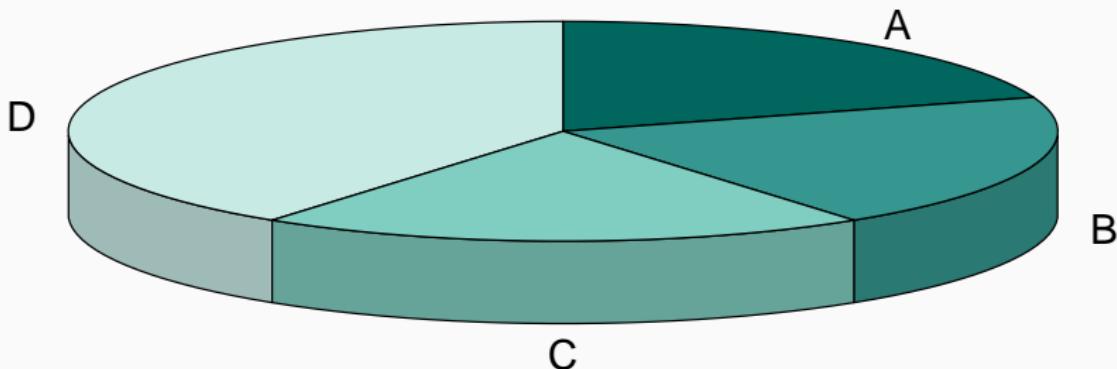
Even-odd pairs emerge



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

A. _____

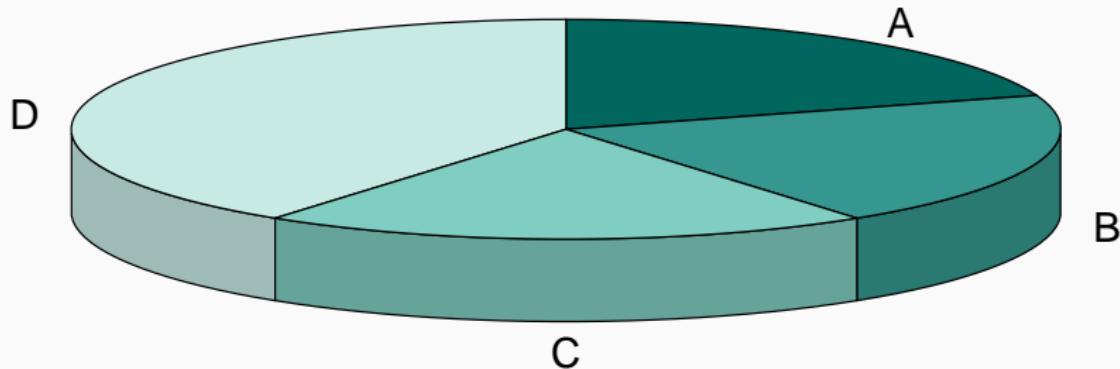
The total should be 100%

B. _____

C. _____

D. _____

3D-effects distort our judgment



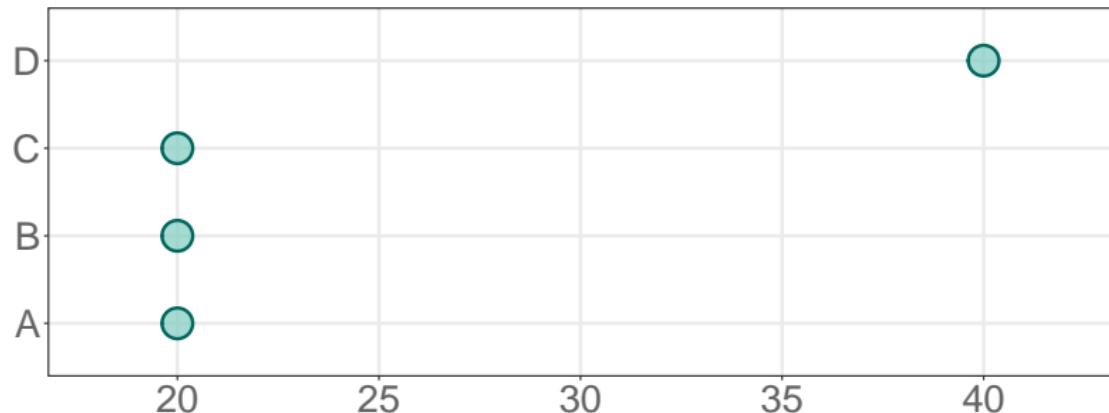
Fill in the blanks

The total should be 100%

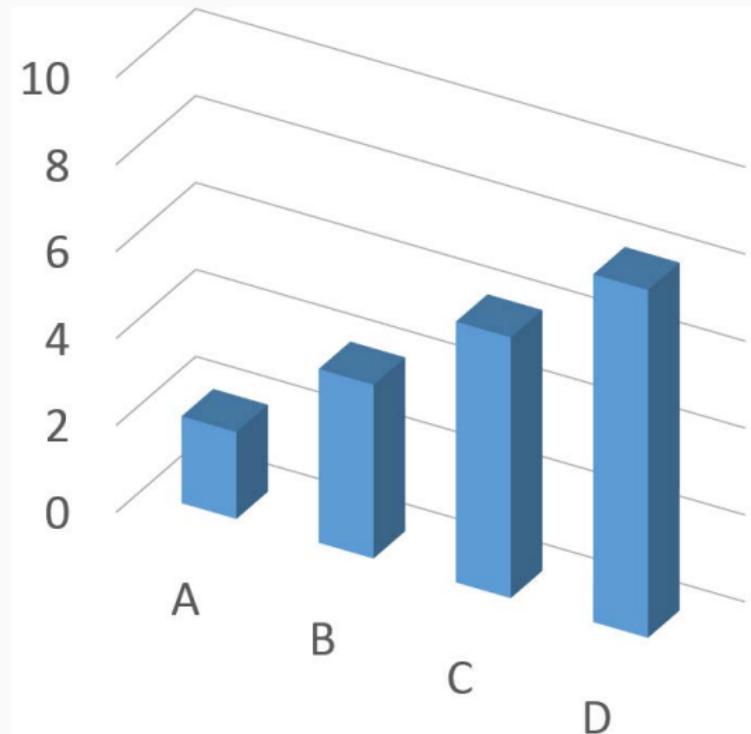
- A. 20%
- B. 20%
- C. 20%
- D. 40%

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

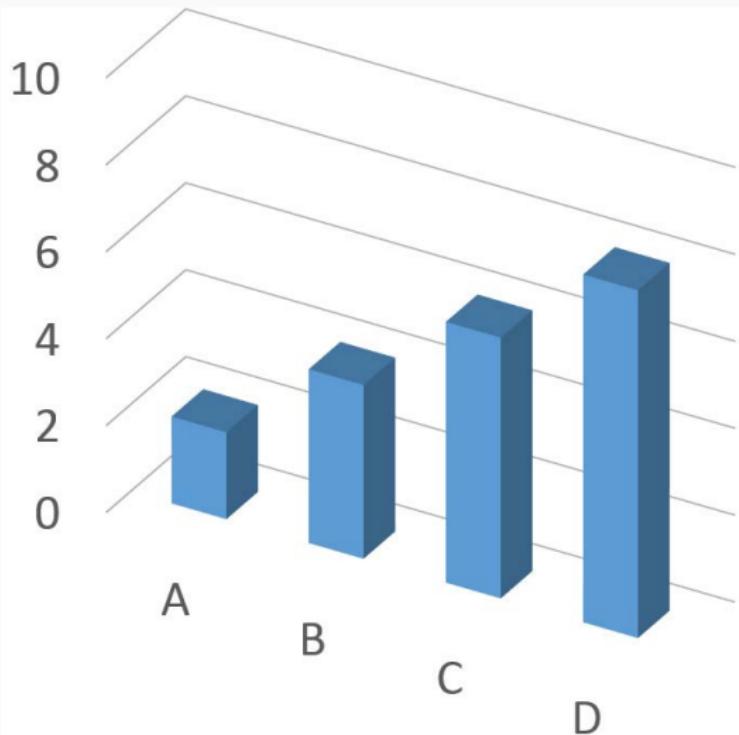
A. _____

B. _____

C. _____

D. _____

Again, 3D-effects distort our judgment

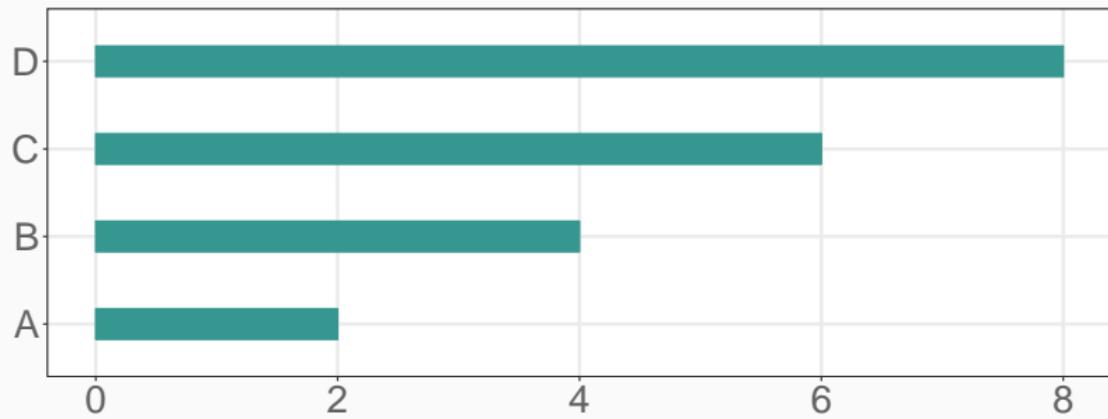


This is a visual inspection only.

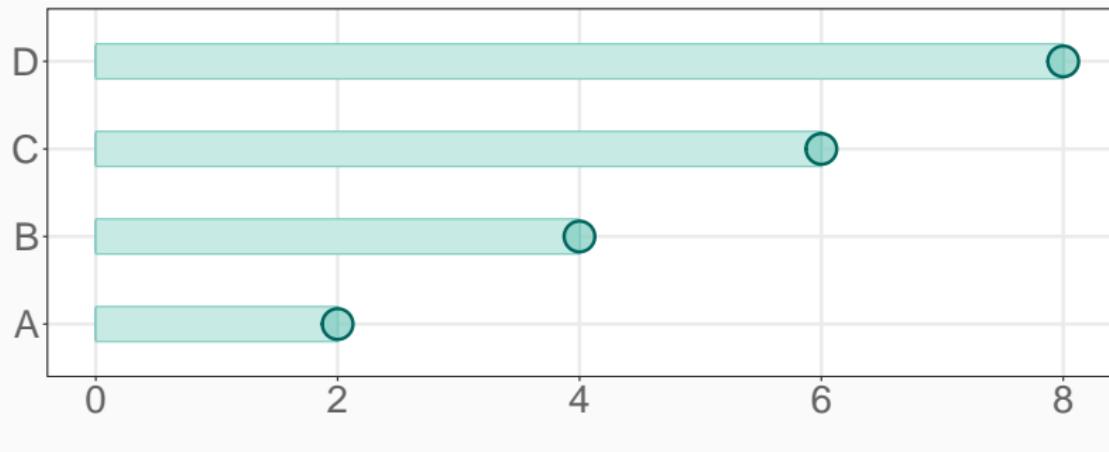
Fill in the blanks

- A. 2
- B. 4
- C. 6
- D. 8

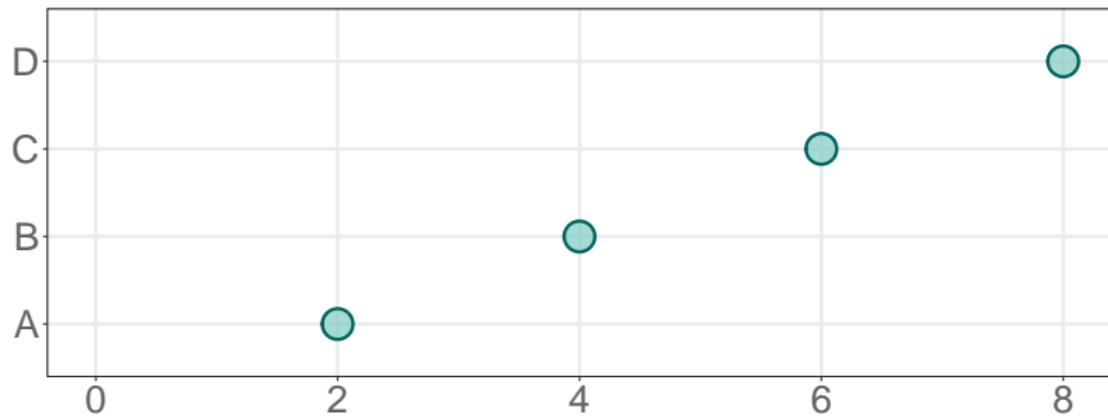
You can use bars, but must include zero



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) \approx 83,000 sq km.

FL _____ $\times 1000$ sq. km

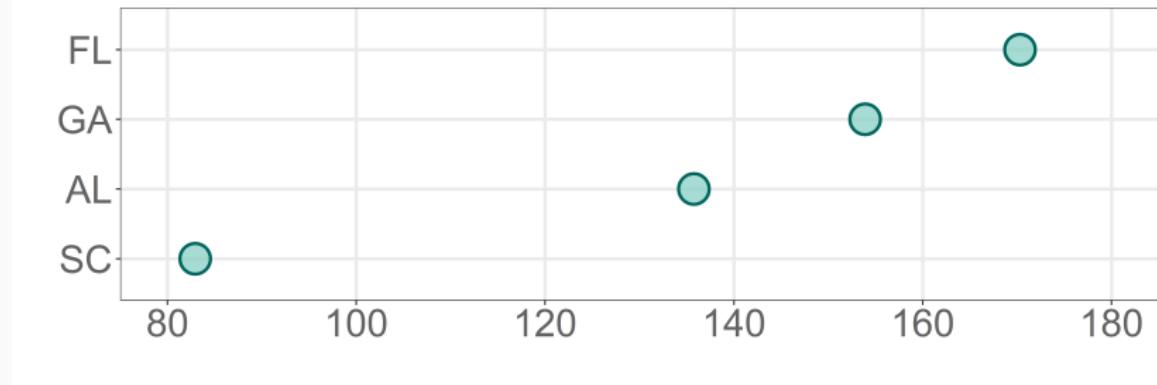
GA _____ $\times 1000$ sq. km

AL _____ $\times 1000$ sq. km

SC 83 $\times 1000$ sq. km

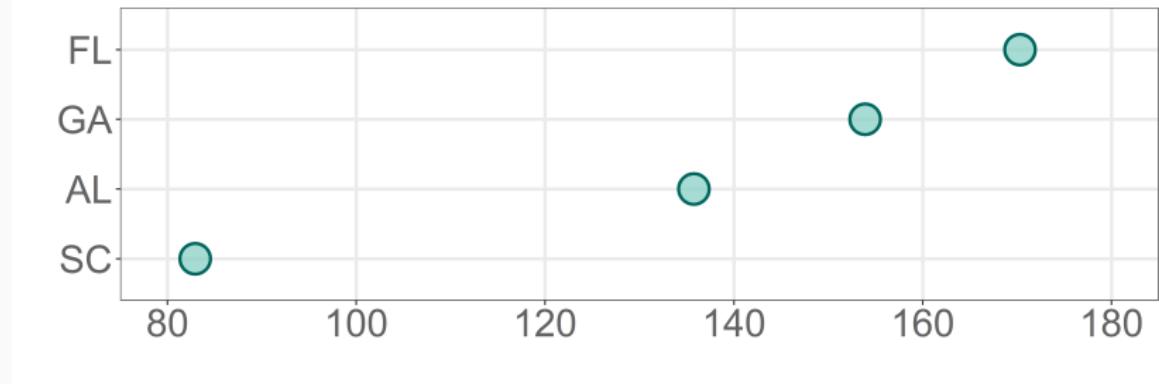
Adapted from (Ihaka 2007)

2nd attempt: Visually estimate the state areas



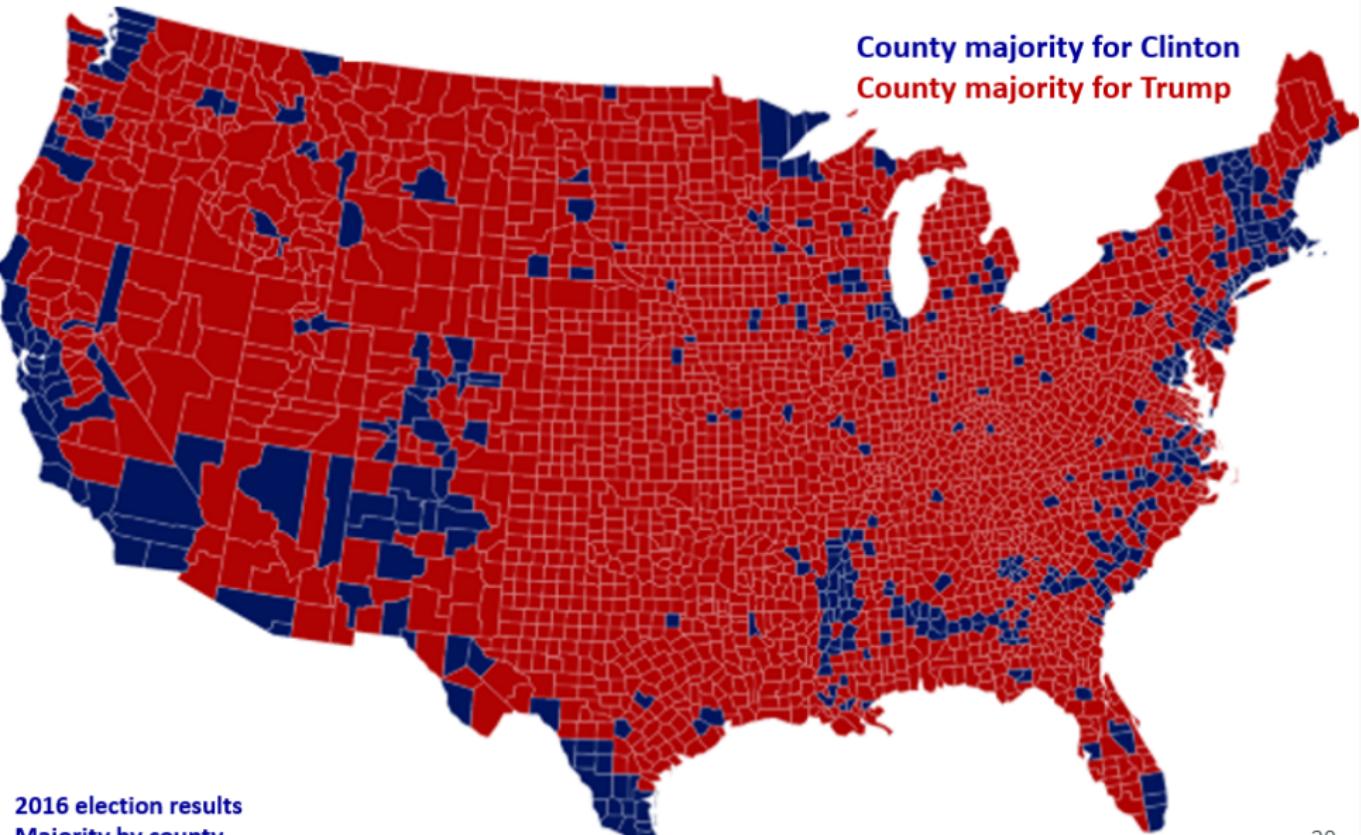
FL	_____	x 1000 sq. km
GA	_____	x 1000 sq. km
AL	_____	x 1000 sq. km
SC	83	x 1000 sq. km

Your estimates have probably improved

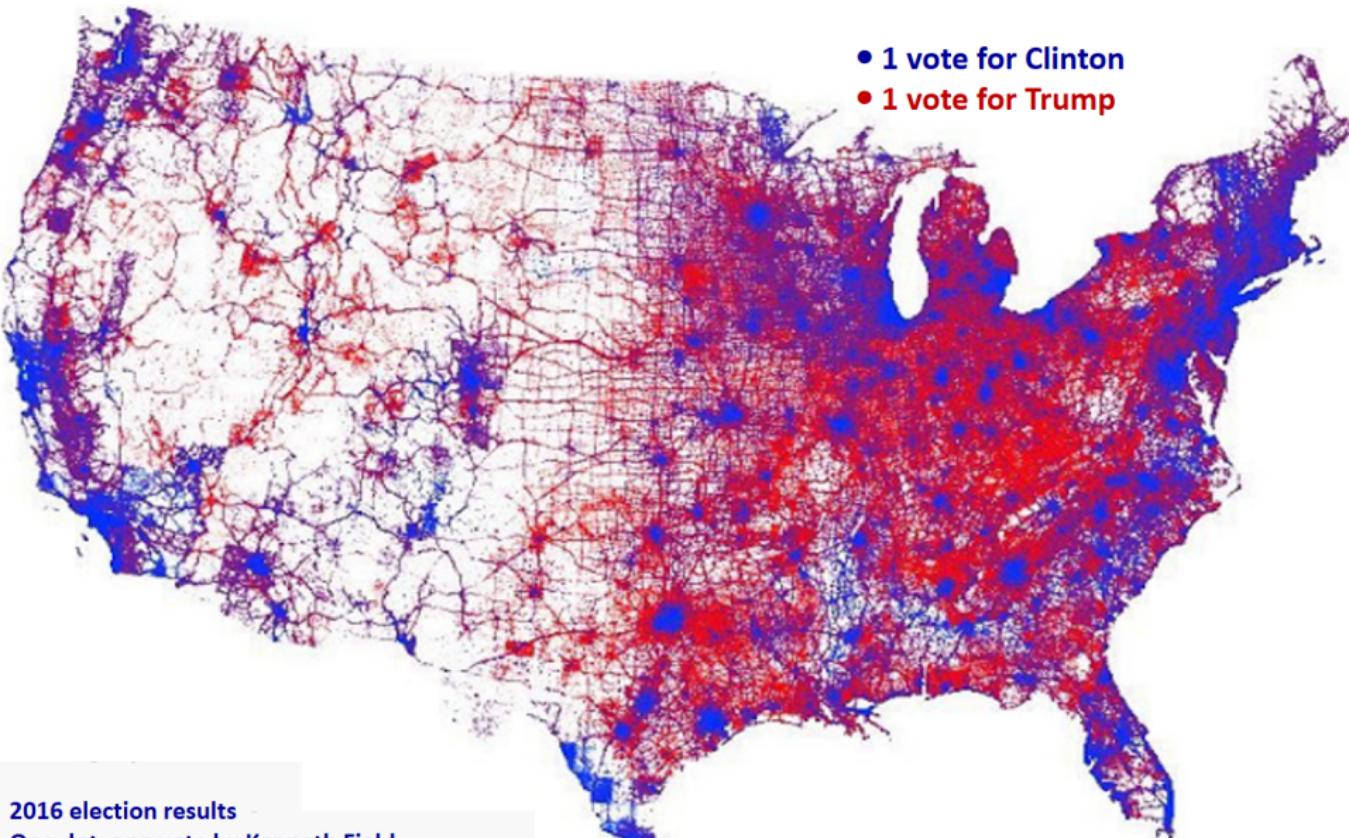


- FL **170** x 1000 sq. km
- GA **154** x 1000 sq. km
- AL **136** x 1000 sq. km
- SC **83** x 1000 sq. km

Color represents geographic area. What is the visual story?



One dot is one vote. Color by party. What is the visual story?



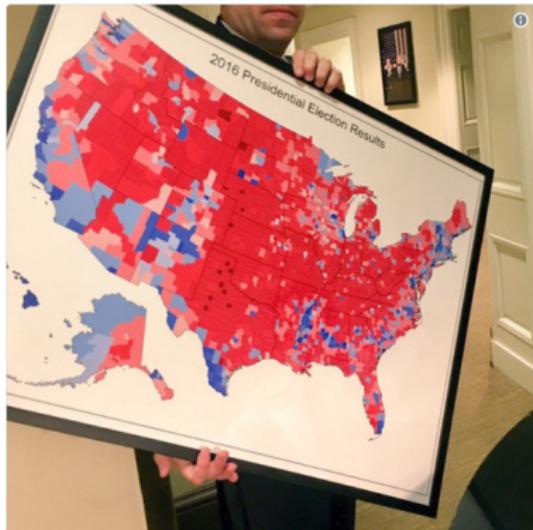
2016 election results

One dot, one vote by Kenneth Field

<http://cartonerd.blogspot.com/2018/03/dotty-election-map.html>

Which story does the President favor?

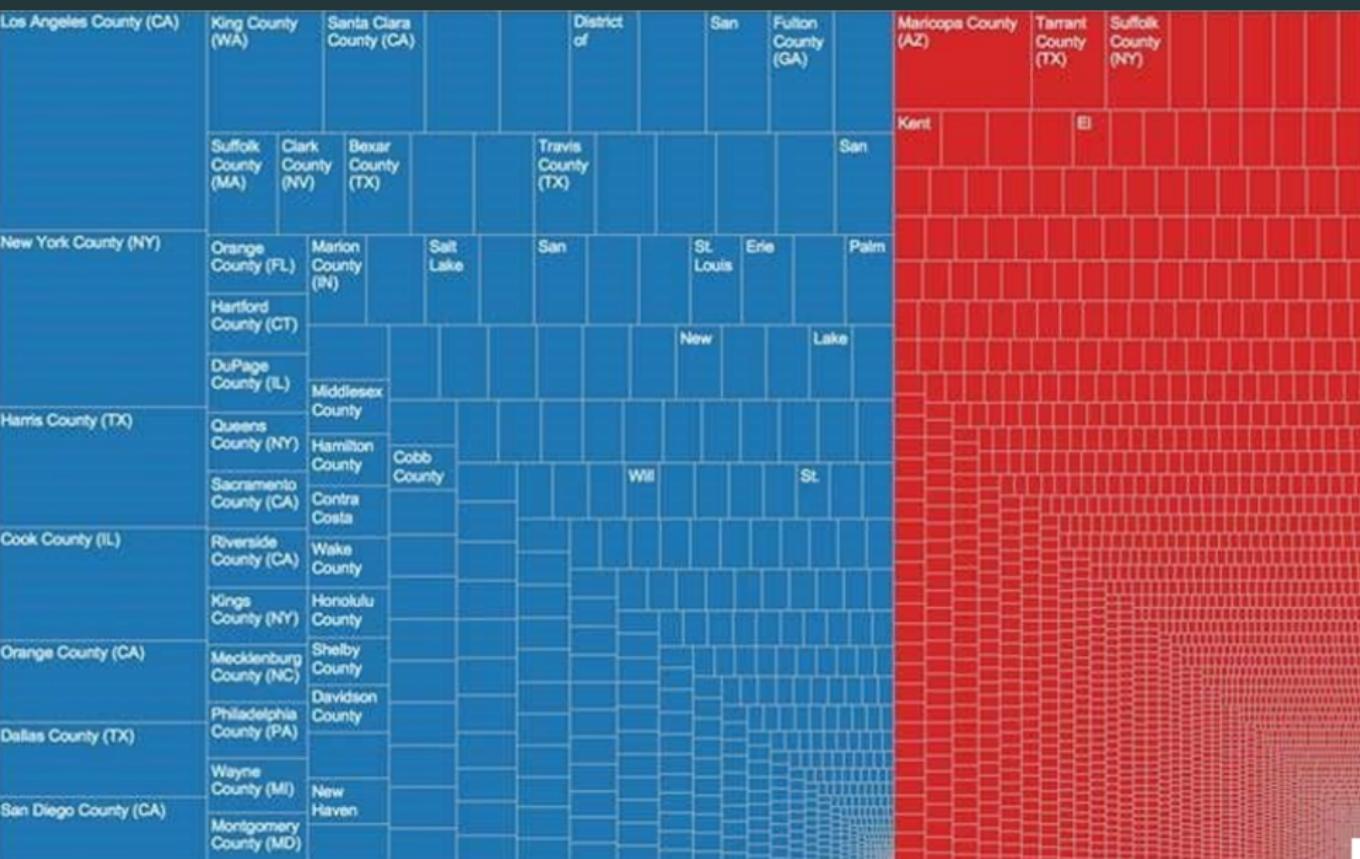
During a discussion about Chinese President Xi Jinping, Trump stopped to hand copies of the 2016 electoral map to each of the reporters.
Business Insider, April 2017



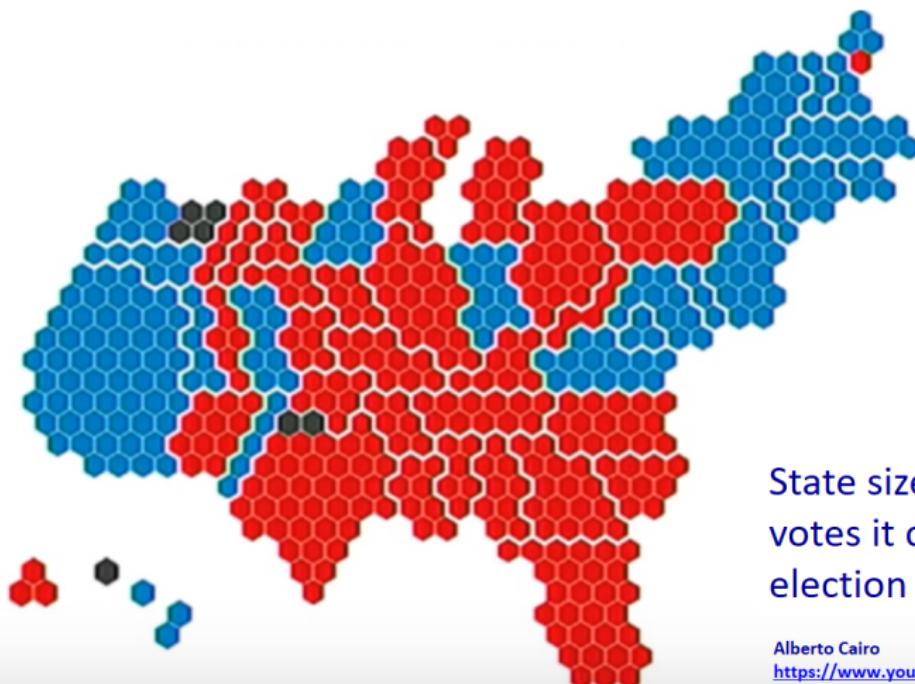
Trey Yingst @TreyYingst

Spotted: A map to be hung somewhere in the West Wing
8,329 10:03 AM - May 11, 2017

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



The numbers that decide the election. What is the visual story?

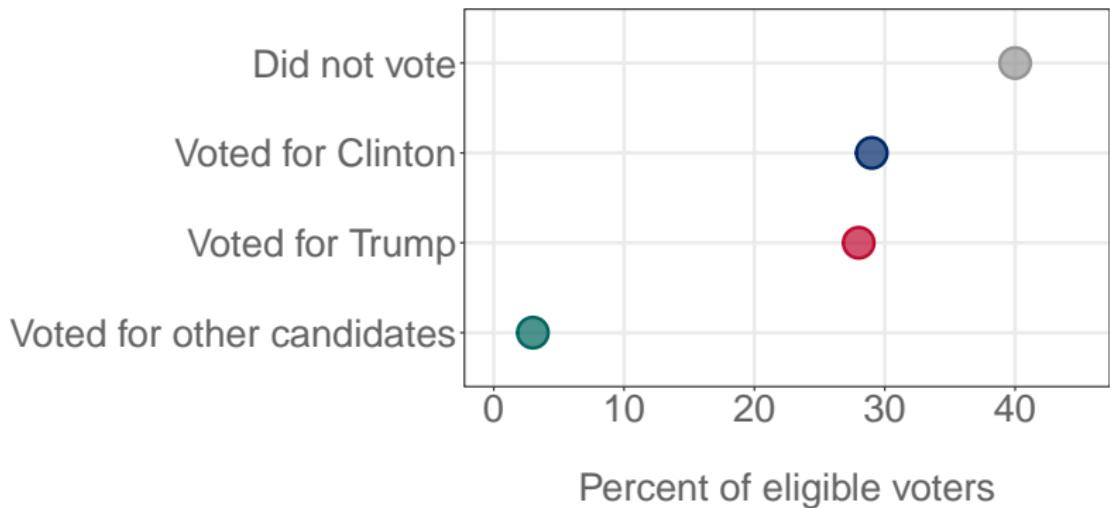


State size adjusted by electoral votes it contributes to the election

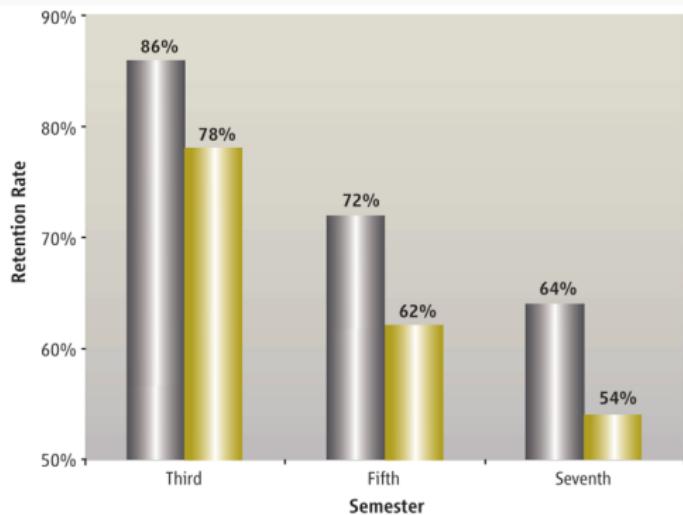
Alberto Cairo

https://www.youtube.com/watch?v=Cd046xZhO_8

The previous graphs conceal what story ?



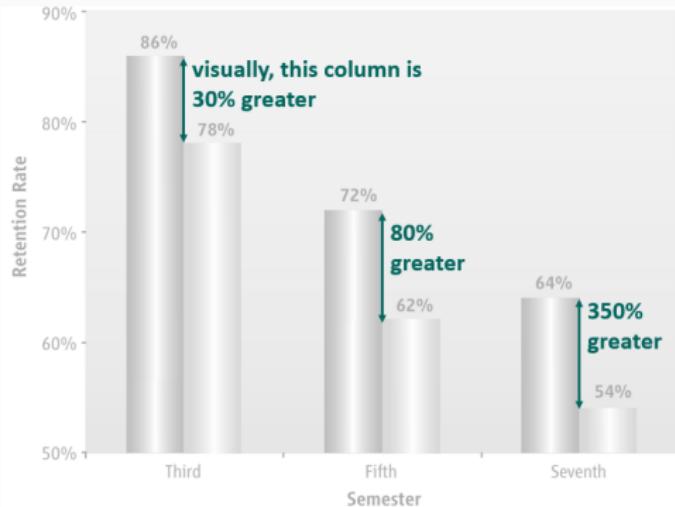
The trouble with bars. What's wrong with this graph?



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.

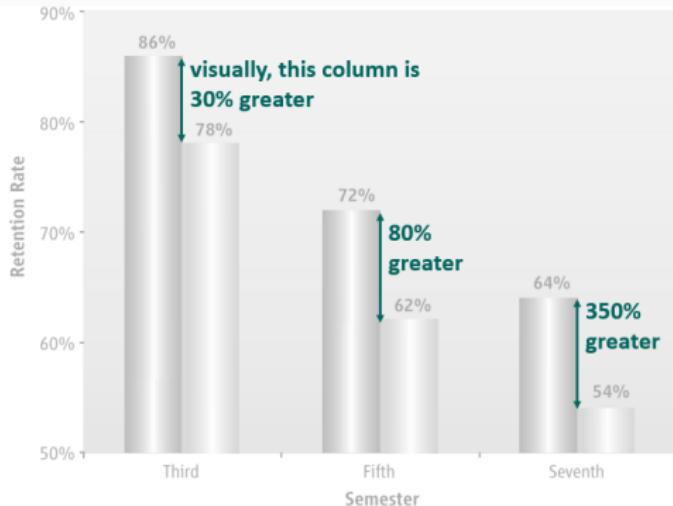
A nearly constant difference seems to increase in significance



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 ($P < 0.05$).

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

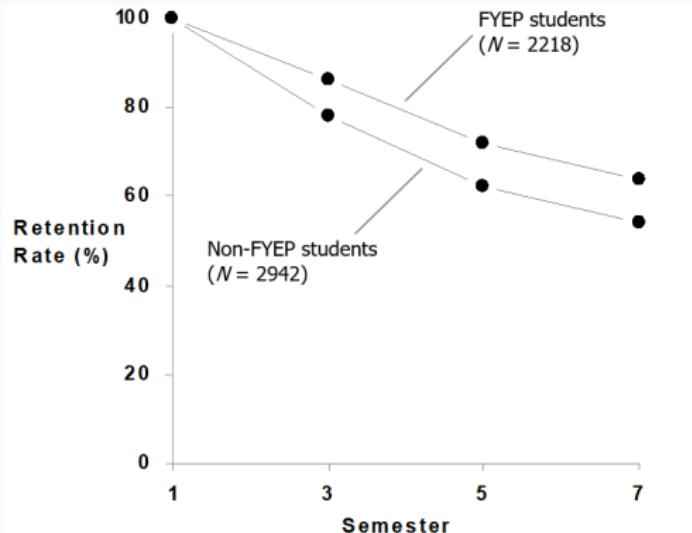


- Can you identify the missing information?

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 ($P < 0.05$).

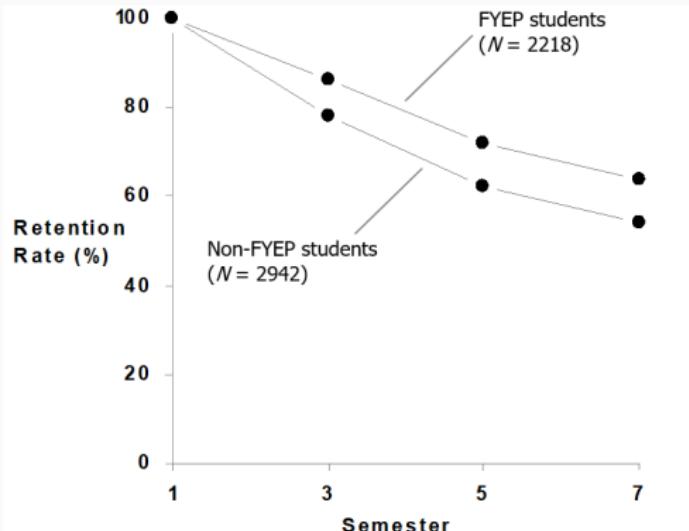
Norman L. Fortenberry, Jacquelyn F. Sullivan, Peter N. Jordan, and Daniel W. Knight (2007), Engineering education research aids instruction, *Science*, 31:1175–1176.

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.

Redesigned, with full scales, a different story emerges



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

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.

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



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

(Tufte 1983)

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

The experts tell us



(Evergreen 2017)

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

What's your point?

Seriously, that's the most important question.

References

- Doumont, Jean-luc. 2009. *Trees, maps, and theorems: Effective communication for rational minds*. Kraainem, Belgium: Principiae. <http://www.treessmapsandtheorems.com/>.
- Evergreen, Stephanie. 2017. *Effective Data Visualization: The Right Chart for the Right Data*. 1st ed. Thousand Oaks, CA: Sage.
- Ihaka, Ross. 2007. "Statistics 787 Lecture slides." <https://tinyurl.com/yywbpua>.
- Kostelnick, Charles, and Michael Hassett. 2003. *Shaping Information: The Rhetoric of Visual Conventions*. Southern Illinois University.
- Robbins, Naomi. 2013. *Creating More Effective Graphs*. Wayne, NJ: Chart House.
- Tufte, Edward. 1983. *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press. https://www.edwardtufte.com/tufte/books_vdqi.
- Wainer, Howard. 2000. *Visual revelations: Graphical tales of fate and deception from Napoleon Bonaparte To Ross Perot*. Mahwah, NJ: Psychology Press.