

Agenda

- EDA Refresher
- Effective Visualization
 - Graphical Integrity
 - Scope
 - Displays
 - Sensible Design
- Communication
 - Motivation
 - Key Considerations

Agenda

- EDA Refresher
- Effective Visualization
 - Graphical Integrity
 - Scope
 - Displays
 - Sensible Design
- Communication
 - Motivation
 - Key Considerations

Key Considerations

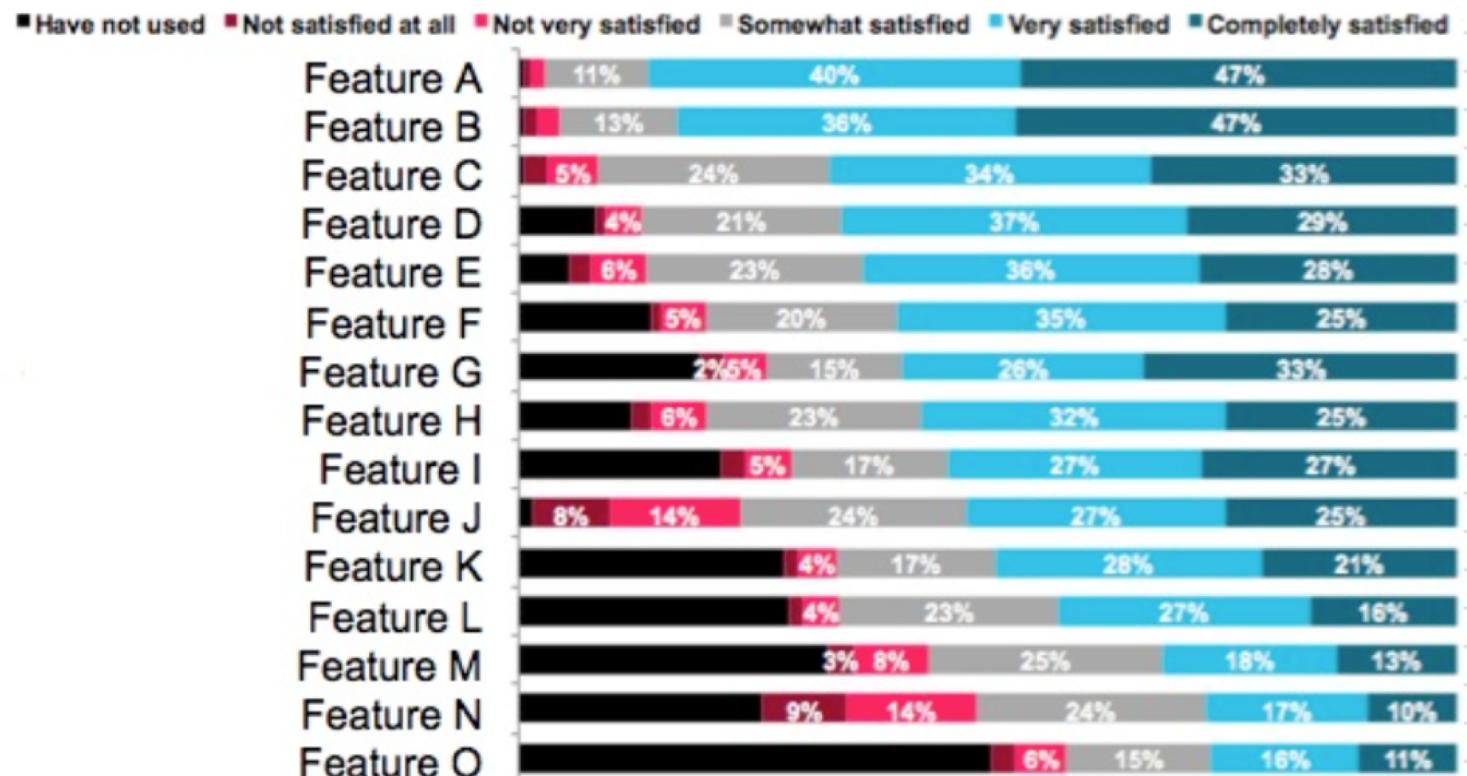
- Who is your **audience**
- What questions are you answering?
- Why should the audience care?
- What are your major insights and surprises?
- What change do you want to affect?

Don't make them think!

- The audience does not want to spend cognitive energy on dissecting and decoding your intended message.
- Lead them through the **major** steps of your story.
- Point out interesting key facts and insights using **captions** and **annotations**

Don't Bury the Lead

How satisfied have you been with each of these features?



User satisfaction varies greatly by feature

Product X User Satisfaction: Features

* Completely satisfied * Very satisfied * Somewhat satisfied * Not very satisfied * Not satisfied at all * Have not used

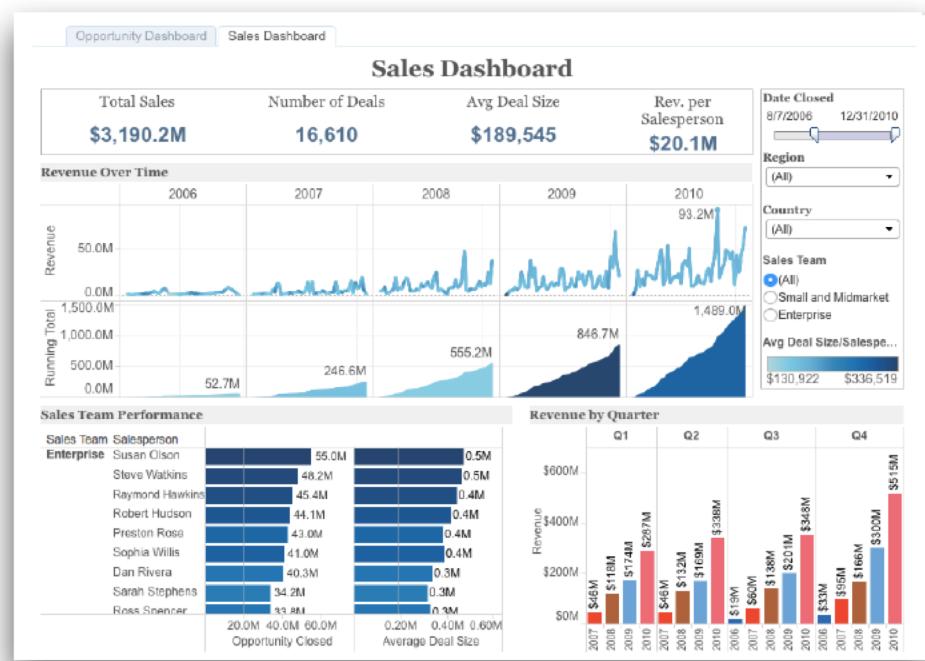


Feature O is least-used feature; what steps can we proactively take with existing users to increase use?

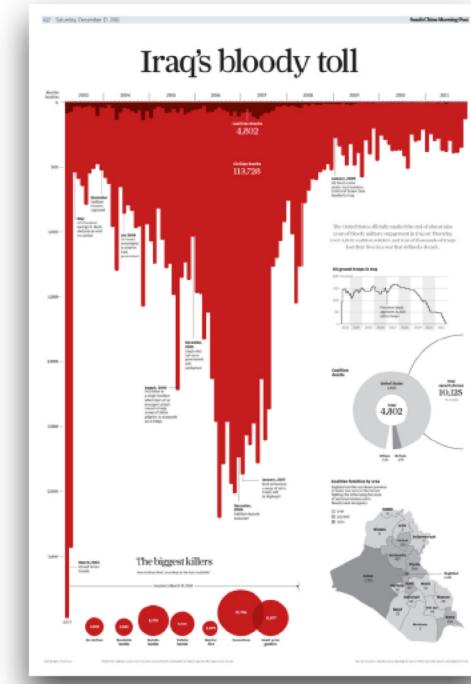
Communication

What is the message?

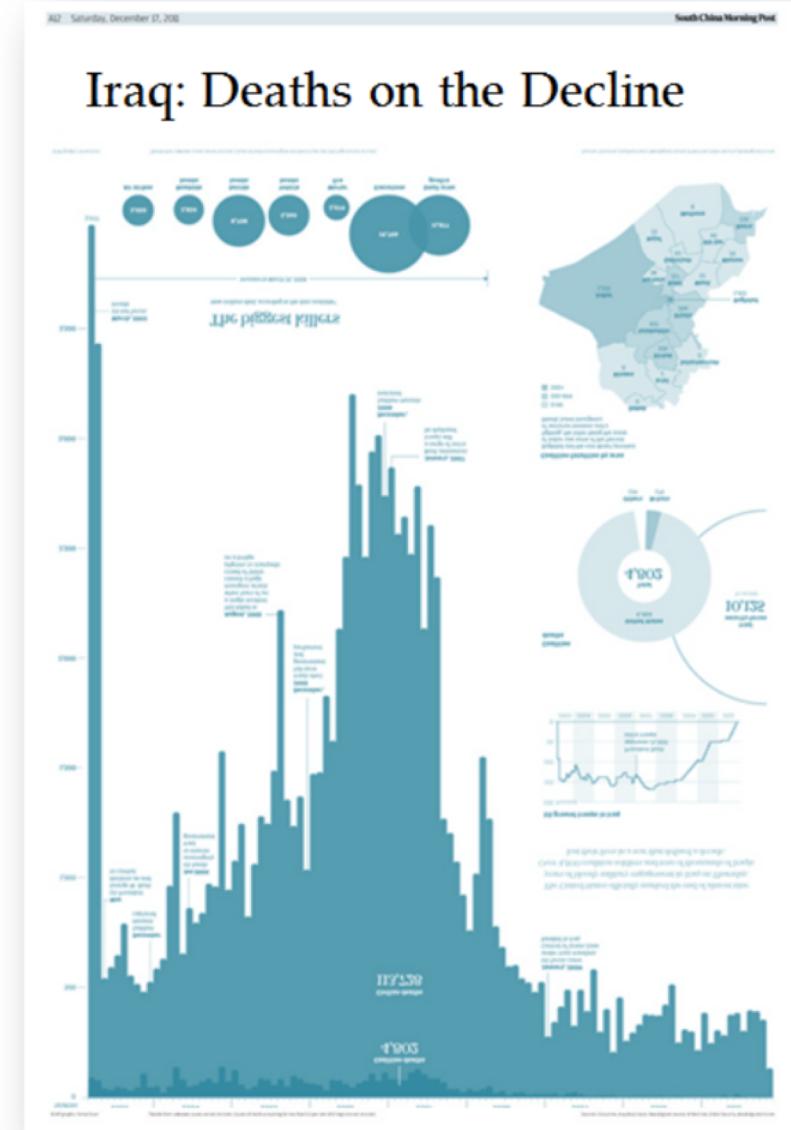
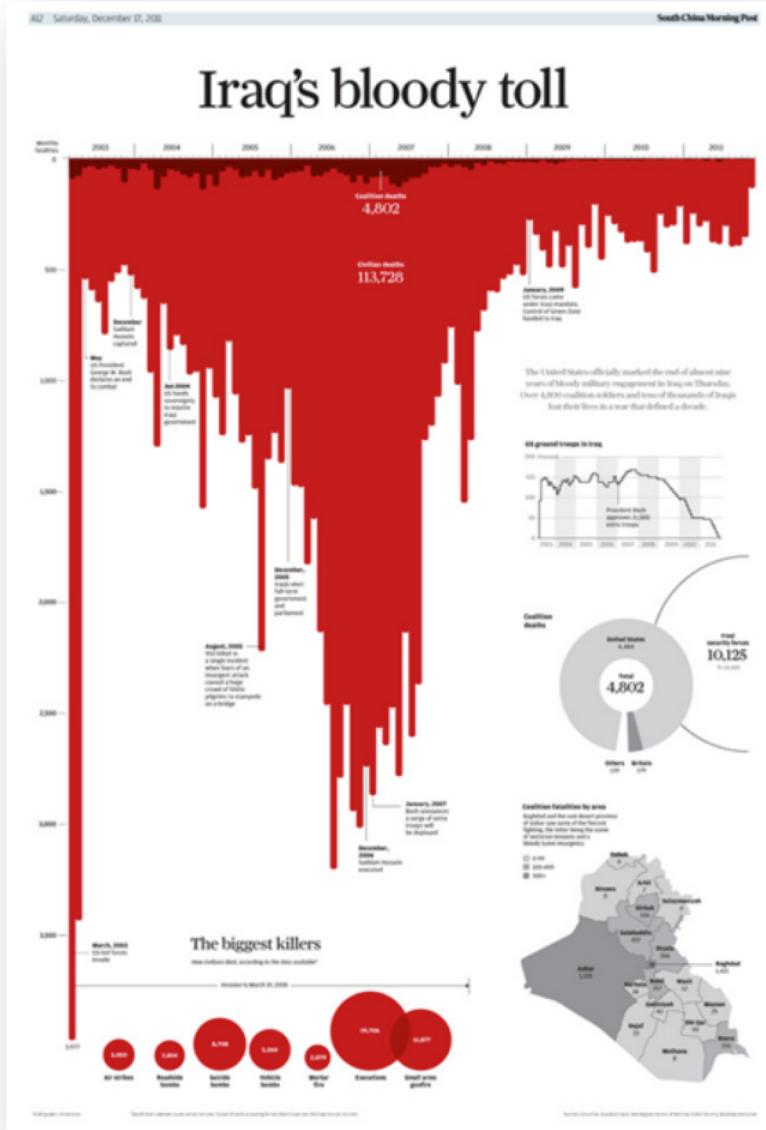
Exploratory
Neutral



Explanatory
Opinionated



Communication



Andy Cotgreave, Tableau

Final Takeaways

- How you choose to display your data greatly influences how people interpret the data
- Humans are visual, *emotional* creations; make graphs that don't make others **feel** confused, insulted, etc.
- Your graphs should illicit good feelings and effectively convey your narrative

Suggested Python Packages

- Matplotlib
- Seaborn
- plotly
- ggplot

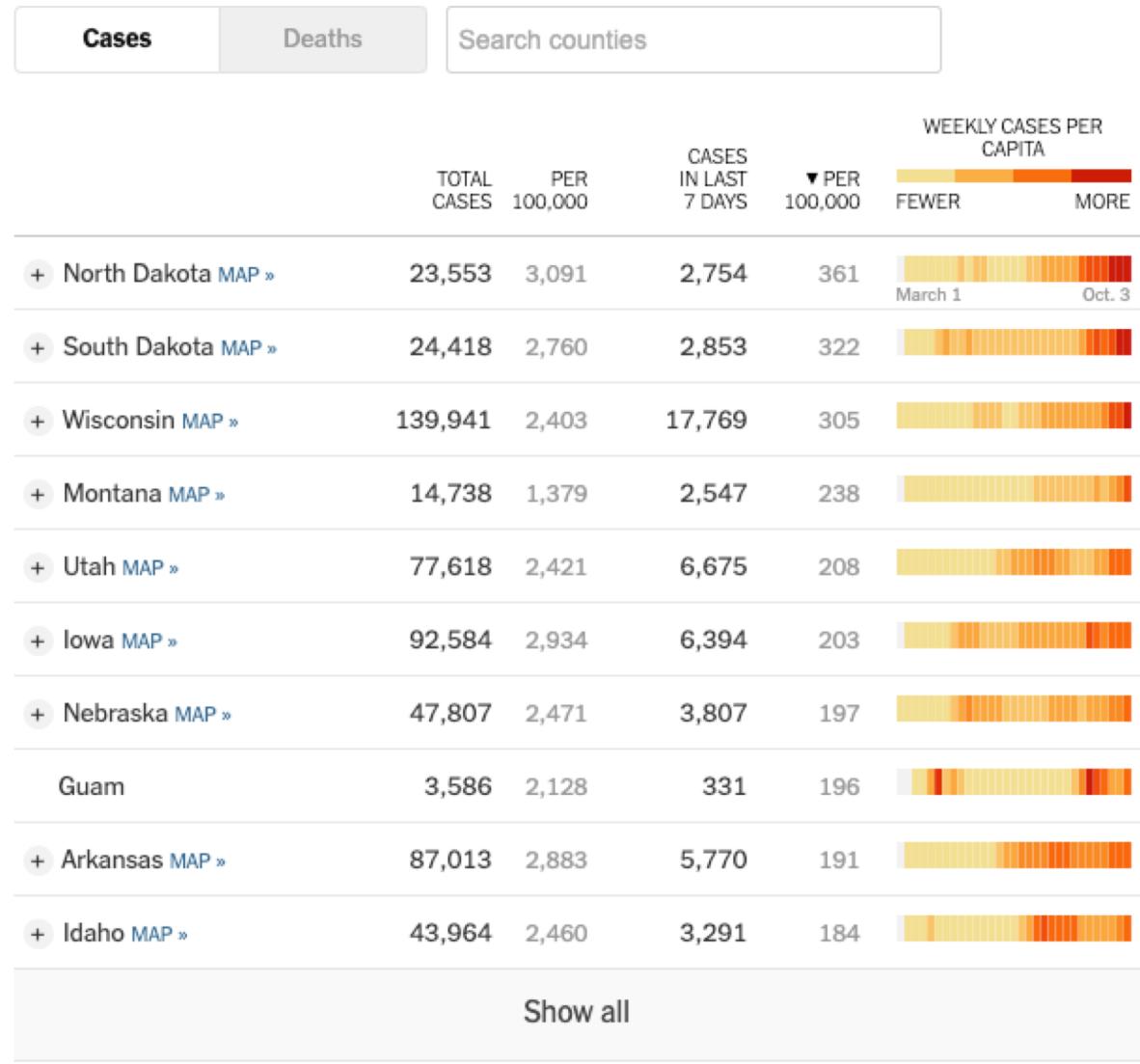


Further Good Examples

- <https://www.nytimes.com/> tends to have incredibly high-quality visualizations that convey information seamlessly
- <https://www.reddit.com/r/dataisbeautiful/>
- fivethirtyeight.com

Cases and deaths by state and county

This table is sorted by places with the most cases per 100,000 residents in the last seven days. Charts are colored to reveal when outbreaks emerged.



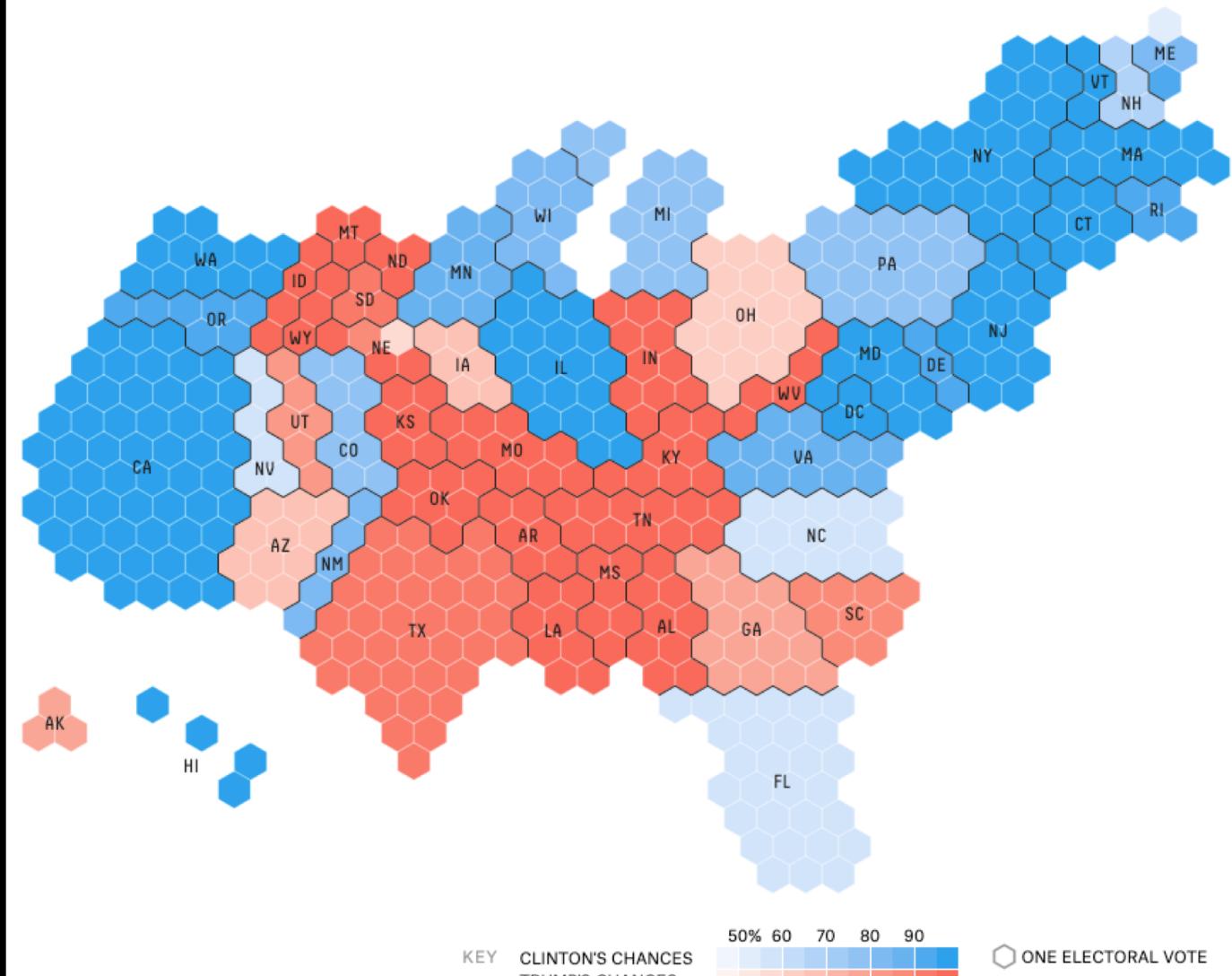
Where new cases are higher and staying high

States where new cases are higher had a daily average of at least 15 new cases per 100,000 people over the past week. Charts show daily cases per capita and are on the same scale. Tap a state to see detailed map page.



It's all about the 538 Electoral College votes

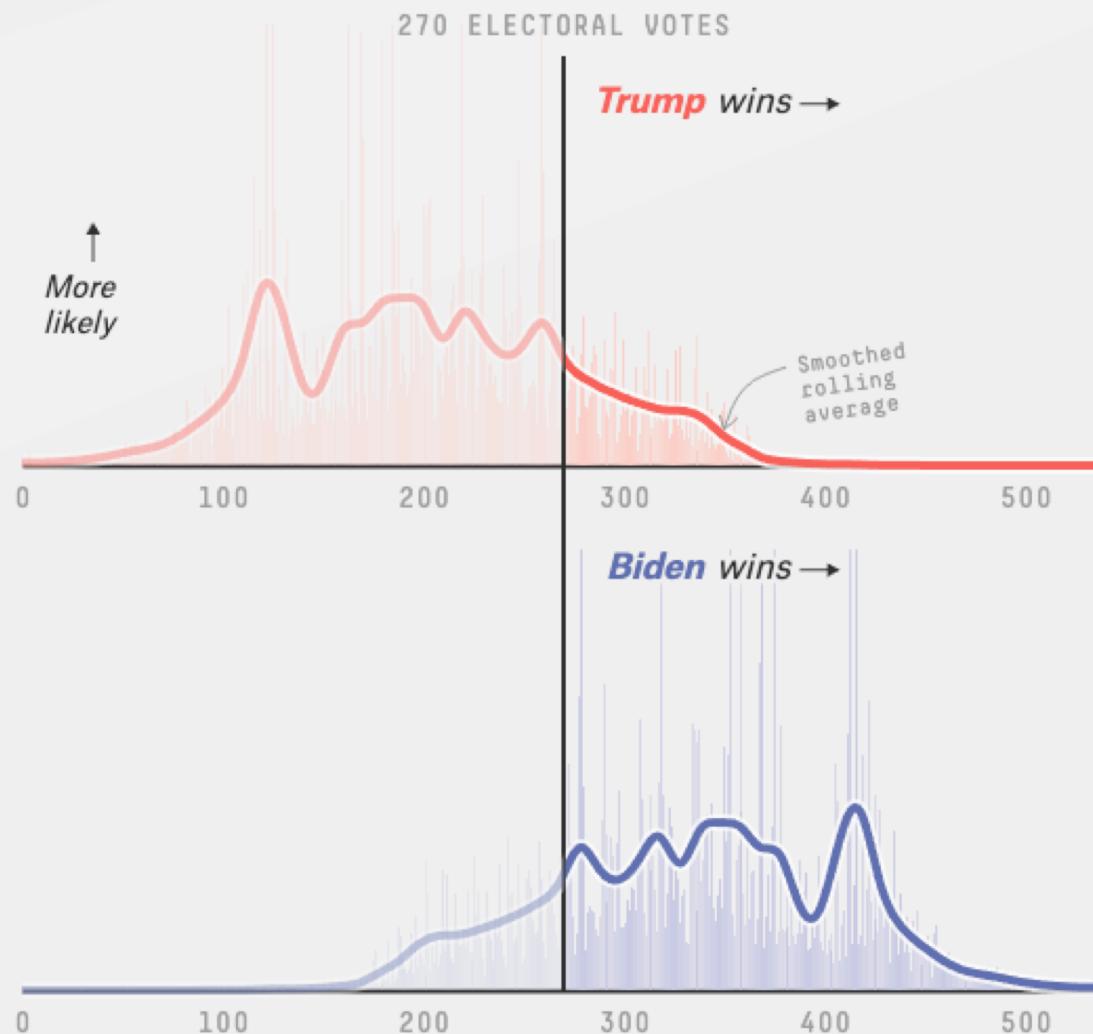
Here's a map of the country, with each state sized by its number of electoral votes and shaded by the leading candidate's chance of winning it.



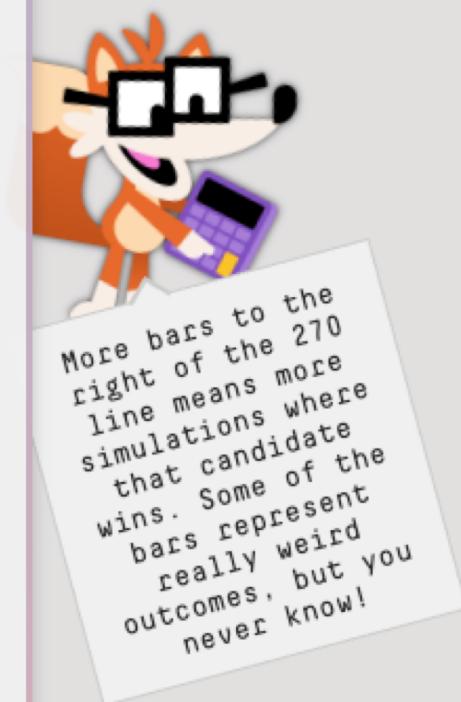
<https://projects.fivethirtyeight.com/2016-election-forecast/>

Every outcome in our simulations

All possible Electoral College outcomes for each candidate, with higher bars showing outcomes that appeared more often in our 40,000 simulations

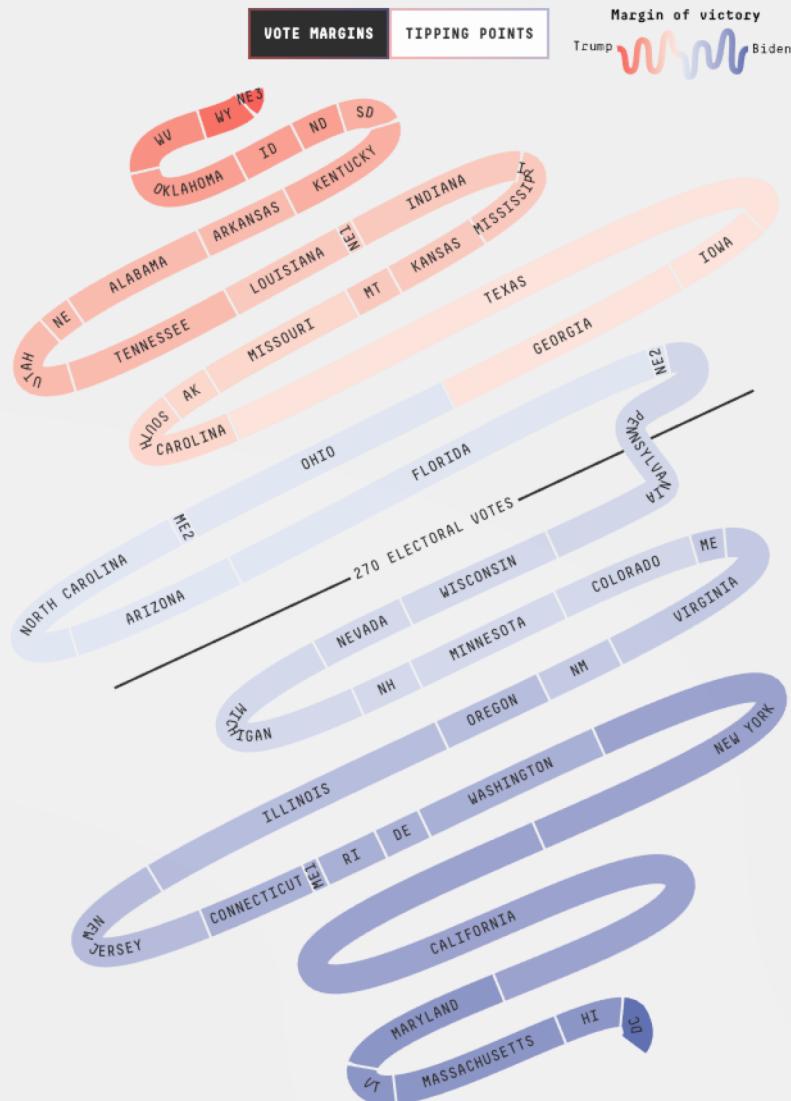


<https://projects.fivethirtyeight.com/2020-election-forecast/>



The winding path to victory

States that are forecasted to vote for one candidate by a big margin are at the ends of the path, while tighter races are in the middle. Bigger segments mean more Electoral College votes. Trace the path from either end to see which state could put one candidate over the top.



Maine and Nebraska's congressional districts are shown separately because those states split their Electoral College votes, allotting some to the statewide winner and some to the winner of each district.

<https://projects.fivethirtyeight.com/2020-election-forecast/>

Exercise time!