Data Privacy and Biased Algorithms

Data Privacy

Example: It’s Easy to be Unsafe Online

Data Privacy

Example: Shiru Cafe

Located in Providence, Rhode Island near Brown University

Turns away customers who are not college students or faculty members

Faculty members pay in $$$, college students pay in their personal information!

To get free coffee, students give away information such as their name, date of birth, phone #, email address, major, and professional interests

Students open themselves up to receiving information from corporate sponsors who pay the cafe to reach its clientele through logos, apps, digital ads in the cafe, and even the baristas

Would you trade your information for free coffee?

Source: No Cash Needed At This Cafe. Students Pay The Tab With Their Personal Data

Data Privacy

Example: Your Data!

Think About It: If this data was obtained, how could it be used? What is the potential impact?

Google Maps Timeline

Google Ad Settings

Your Facebook Information

Google Activity History

Data Privacy

Example: OkCupid Data Breach

In 2016, researchers published data of 70,000 OKCupid users—including usernames, political leanings, drug usage, and intimate sexual details

"Some may object to the ethics of gathering and releasing this data. However, all the data found in the dataset are or were already publicly available, so releasing this dataset merely presents it in a more useful form." — Researchers Emil Kirkegaard and Julius Daugbjerg Bjerrekær

Although the researchers did not release the real names and pictures of the OKCupid users, critics noted that their identities could easily be uncovered from the details provided —such as from the usernames

Source: OKCupid Study Reveals the Perils of Big-Data Science

Data Privacy

Example: Facebook and Cambridge Analytica

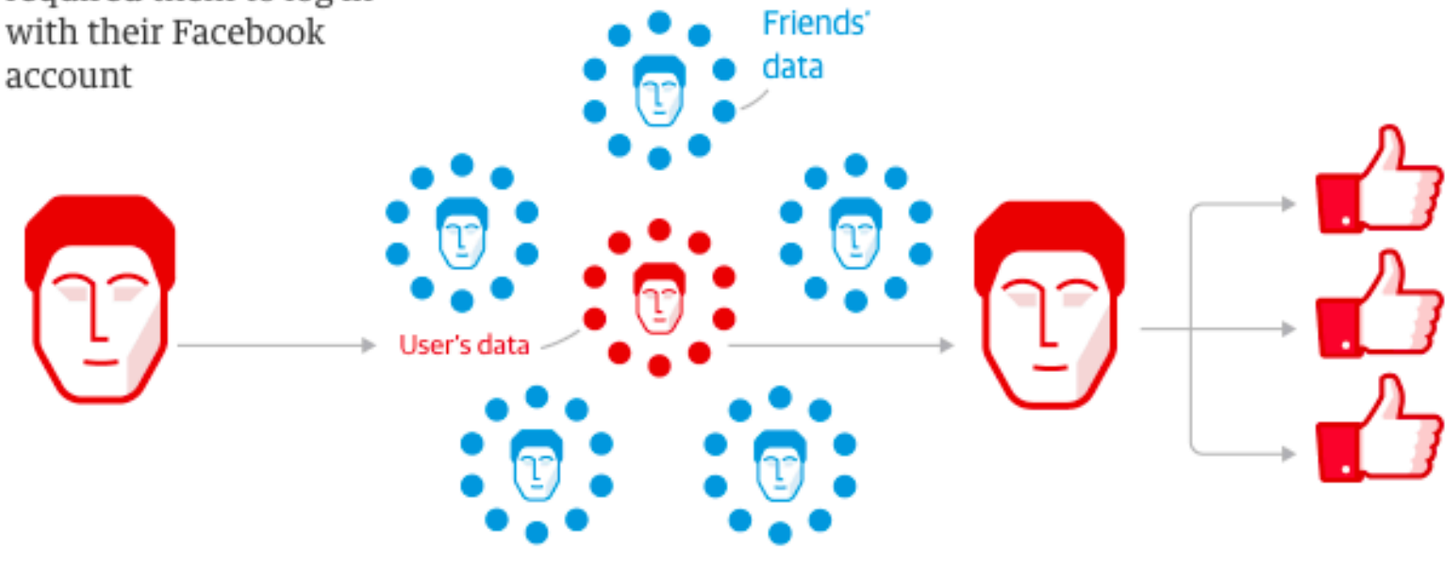
Cambridge Analytica: how 50m Facebook records were hijacked

1
Approx. 320,000 US voters ('seeders') were **paid $2-5 to take a detailed personality/ political test** that required them to log in with their Facebook account

2
The app also **collected data such as likes and personal information** from the test-taker's Facebook account ...

3
The **personality quiz results** were paired with their Facebook data - such as **likes** - to seek out psychological patterns

4
Algorithms combined the data with other sources such as voter records to **create a superior set of records (initially 2m people in 11 key states*)**, with hundreds of data points per person



... as well their **friends'** data, amounting to over 50m people's raw Facebook data

These individuals could then be targeted with **highly personalised advertising** based on their personality data

Guardian graphic. *Arkansas, Colorado, Florida, Iowa, Louisiana, Nevada, New Hampshire, North Carolina, Oregon, South Carolina, West Virginia

Data Privacy

“Like all security, privacy is hard.” — Cory Doctorow

What can we do?

Privacy by design

Considering privacy at the beginning of a project not near the end.

Anonymization generalization

Reducing the fallout if there is a breach.

Risk assessment

Identifying potential threats, external and internal users, potential use of data.

Secure yourself

Make sure your practices are up to snuff.

Data Privacy

Privacy by Design

Be proactive, not reactive; be preventative; not remedial

Privacy as the default setting (opt out)

End-to-end security: full lifecycle protection

Visibility and transparency: keep it open

Respect for user privacy

Data Privacy

Data Anonymization

Generalization

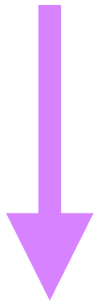
Encryption

Hashing

Pseudonymization

Perturbation

Name	Address	Birth Date	GPA
John Dorian	72 Sacred Heart Street Pittsburgh	06/18/1996	3.65
Perry Cox	856 Miller Rd Zelienople	12/06/1997	2.56
Elliot Reid	567 Street St Canton, Ohio	03/23/1996	3.89
Christopher	3734 McKnight Rd Pittsburgh	10/14/1997	2.97



ID	Location	Age Range	GPA
69546	Western Pennsylvania	19-24	High
45369	Western Pennsylvania	19-24	Low
49369	Western Pennsylvania	19-24	High
43695	Western Pennsylvania	19-24	Average

Data Privacy

Risk Assessment

External Threats: Hackers, social engineering, malware, spyware, privacy watch dogs, legislation

VS

Internal Threats: Customers/ employees scraping data, employees accidentally sharing data, lost devices

Impact to End Users: What is the risk to your users if the data you have is obtained? Is there threat to life, general privacy, or reputation? Is there risk of identity fraud? Is there a monetary risk?

Likelihood: How many devices is the data stored on? Are people likely to target you? How savvy are your employees?

High likelihood, high impact risks are your top priority!

Data Privacy

Check Yourself

Lock It Down:

Encrypt your data

Make sure your computer/the place the data are stored is secure

Make sure any software used is secure + reliable

Secure Sharing:

Make sure the method of sharing is secure (encryption, VPN, SFTP, etc)

Share Sparingly:

Share only what the end-user really needs

Misrepresenting Data

The best-selling book with “statistics” in the title is *How to Lie with Statistics* by Darrell Huff

Shows graphical ploys used to fool people even with accurate data

Let’s take a look at a few examples!

Misrepresenting Data

What is wrong with this plot?

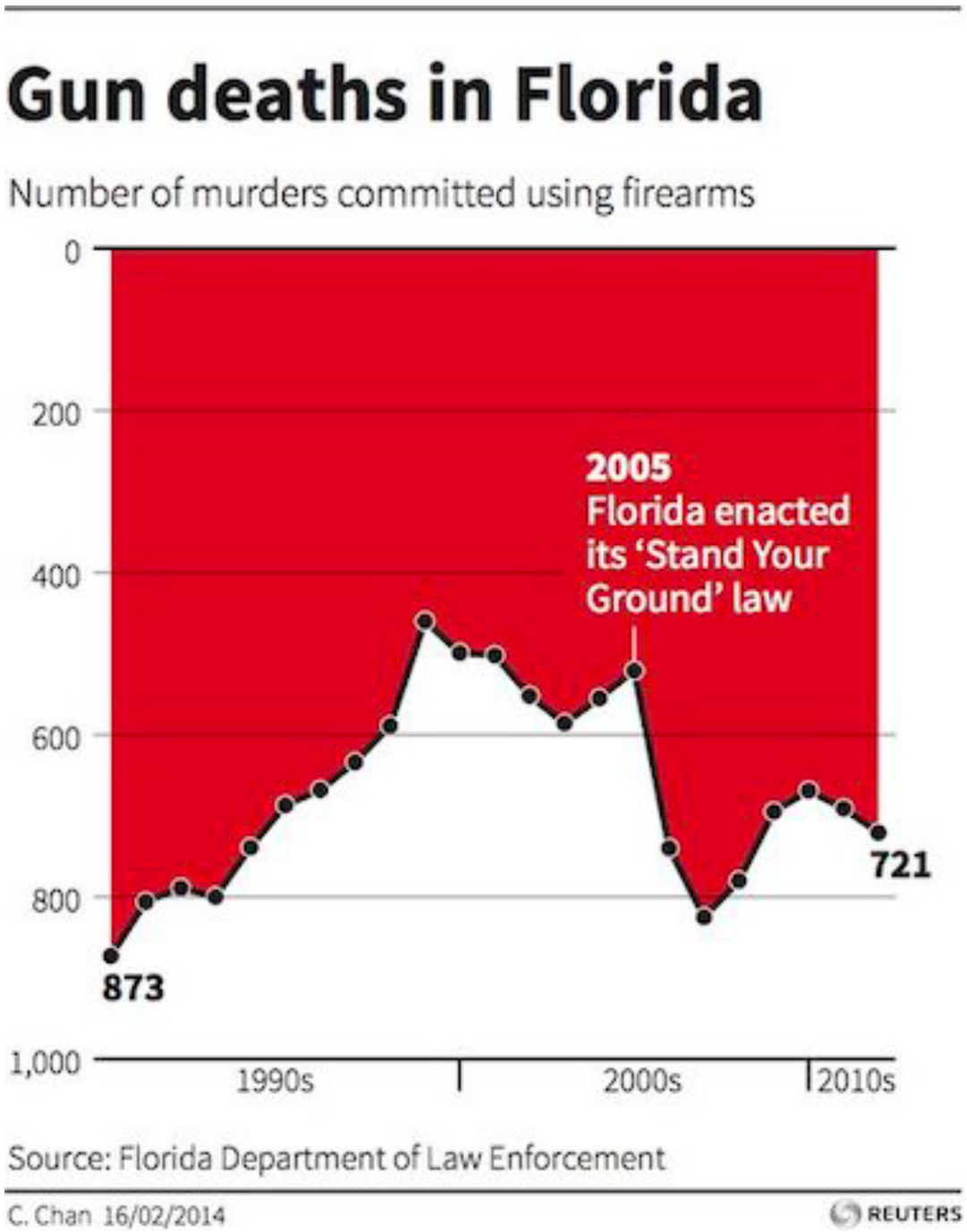
The *y*-axis is flipped!

This is misleading because it makes the increase that occurs after 2005 look like a decrease

This plot is misleading, but it wasn't meant to be!

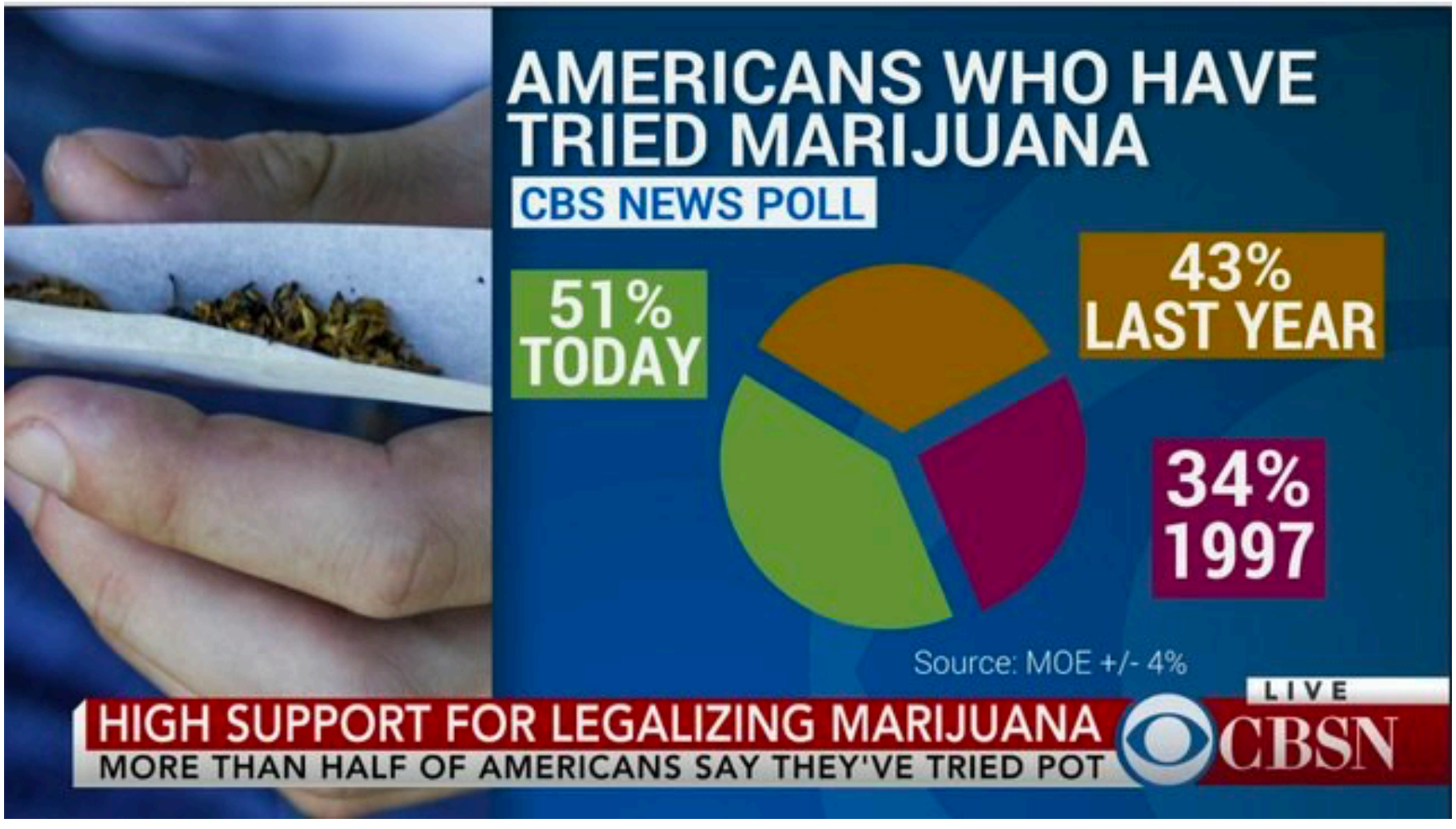
The designer of the chart, Christine Chan, stated "I prefer to show deaths in negative terms (inverted). It's a preference really, it can be shown either way"

The red shading was inspired by a graphic she had seen using red "dribble" lines that evoke blood running down a wall



Misrepresenting Data

What is wrong with this plot?

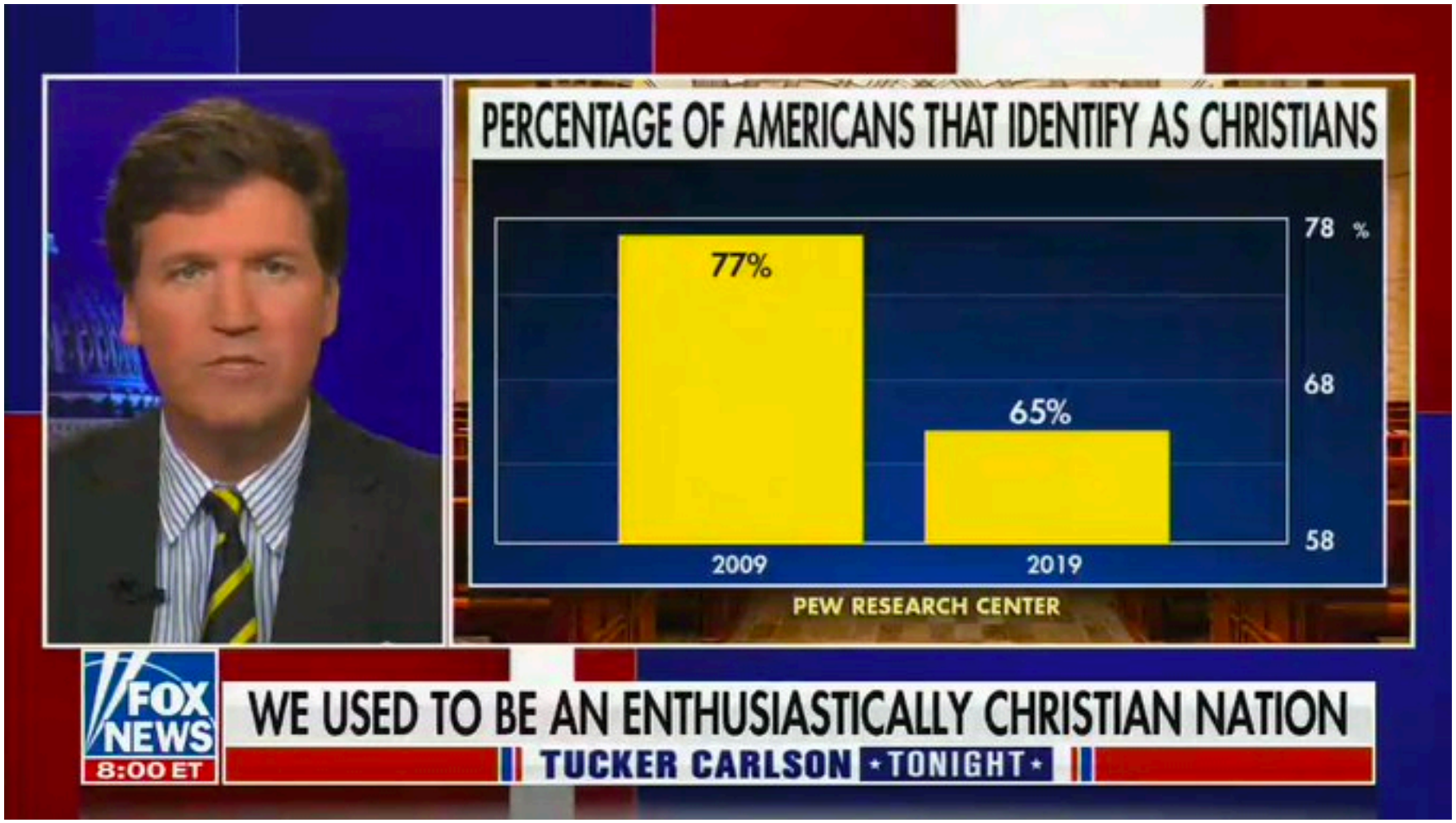


The percentages add to 128%!

A pie chart might not be the best way to represent this type of data (when people can choose multiple responses to a question)

Misrepresenting Data

What is wrong with this plot?



The *y*-axis begins at 58, instead of 0, making a 12% difference look much larger than it actually is

Biased Algorithms

There's software used across the country to predict future criminals, and it's biased…

Example: ProPublica Analysis

Data:

7,000 + arrested in Broward County, Florida in 2013 — 2014

Each person was assigned a risk score

Whether or not they were charged with new crimes in the next two years was recorded

Results:

20% of those predicted to commit violent crimes actually did

Algorithm had a higher accuracy (61%) when full range of crimes taken into account (e.g. misdemeanors)

Biased Algorithms

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Algorithm was more likely to falsely flag African American defendants as future criminals, at almost twice the rate as Caucasian defendants!

"Although these measures were crafted with the best of intentions, I am concerned that they inadvertently undermine our efforts to ensure individualized and equal justice," he said, adding, "they may exacerbate unwarranted and unjust disparities that are already far too common in our criminal justice system and in our society." — U.S. Attorney General Eric Holder (2014)

Source: Machine Bias

Biased Algorithms

Example: Amazon

Amazon used AI to give job candidates scores ranging from one to five stars (*like how shoppers rate products on Amazon)*

Company realized its new system was not rating candidates for software developer jobs and other technical posts in a gender-neutral way

Amazon's system taught itself that candidates identifying as male were preferable!

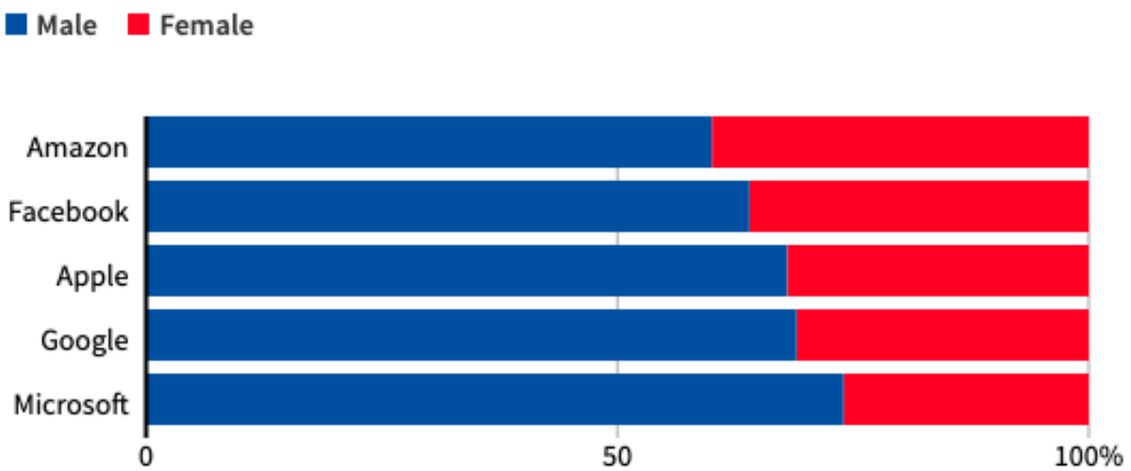
Source: Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women

Biased Algorithms

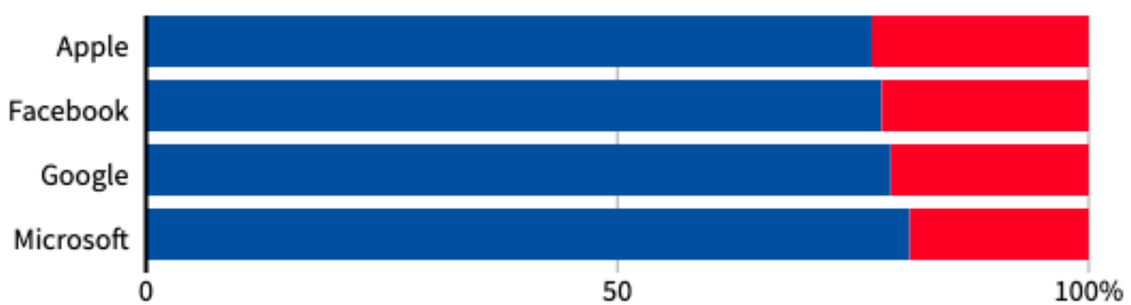
Dominated by men

Top U.S. tech companies have yet to close the gender gap in hiring, a disparity most pronounced among technical staff such as software developers where men far outnumber women. Amazon's experimental recruiting engine followed the same pattern, learning to penalize resumes including the word "women's" until the company discovered the problem.

GLOBAL HEADCOUNT

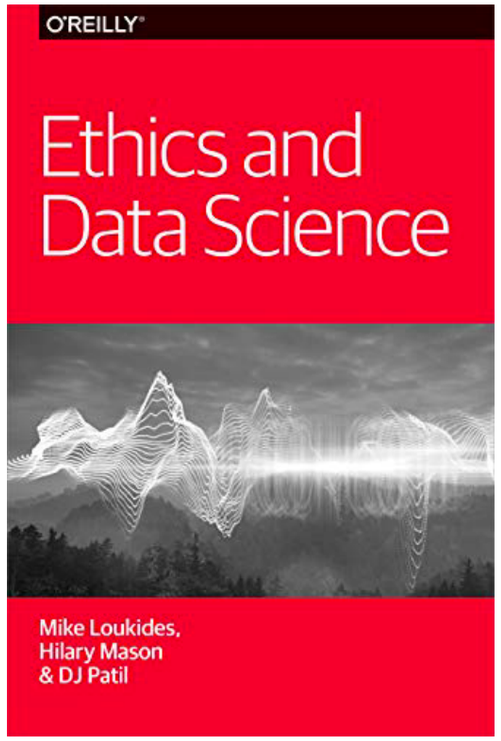


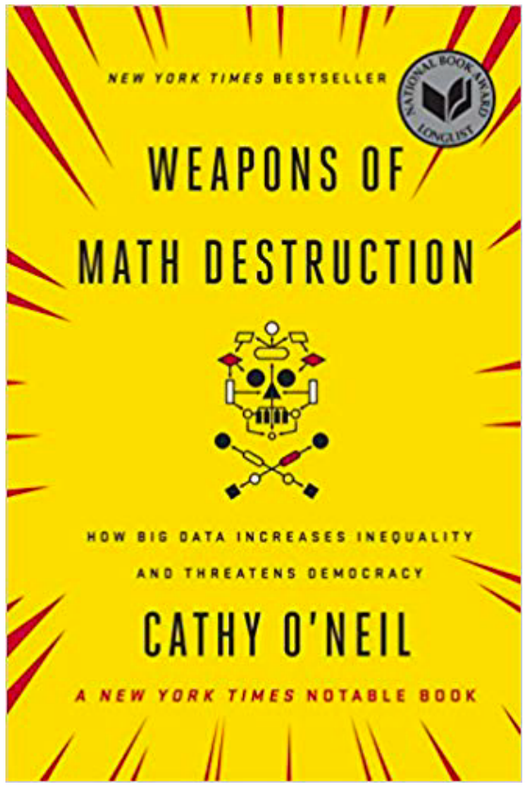
EMPLOYEES IN TECHNICAL ROLES

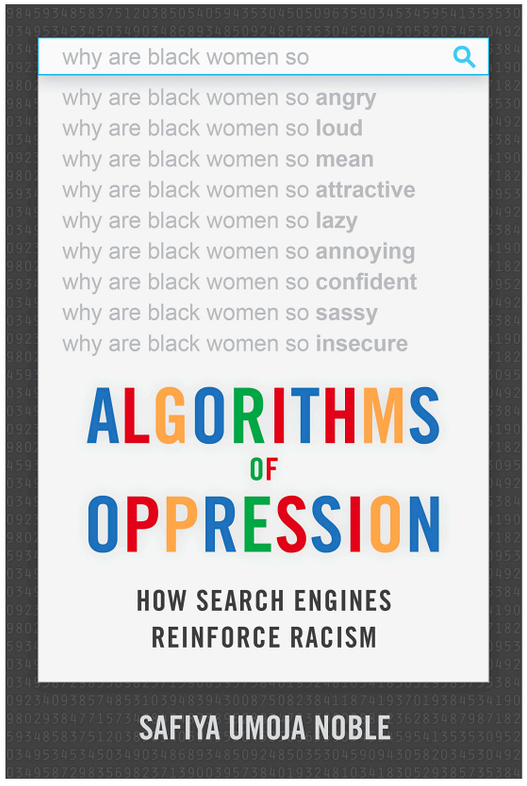


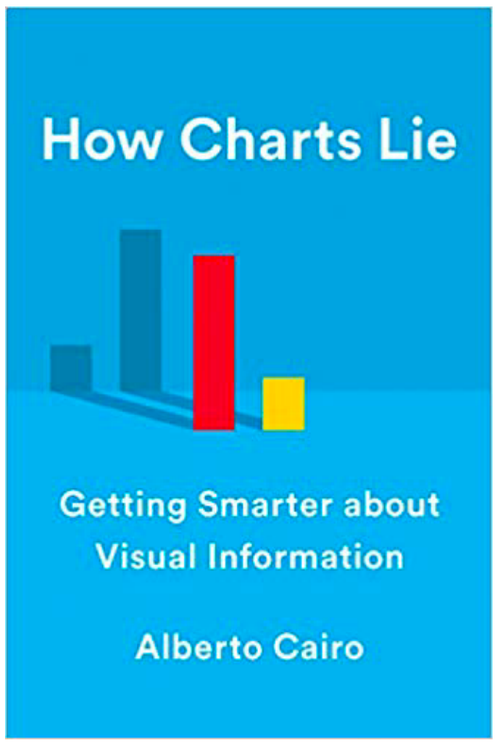
Note: Amazon does not disclose the gender breakdown of its technical workforce.
Source: Latest data available from the companies, since 2017.
By Han Huang | REUTERS GRAPHICS

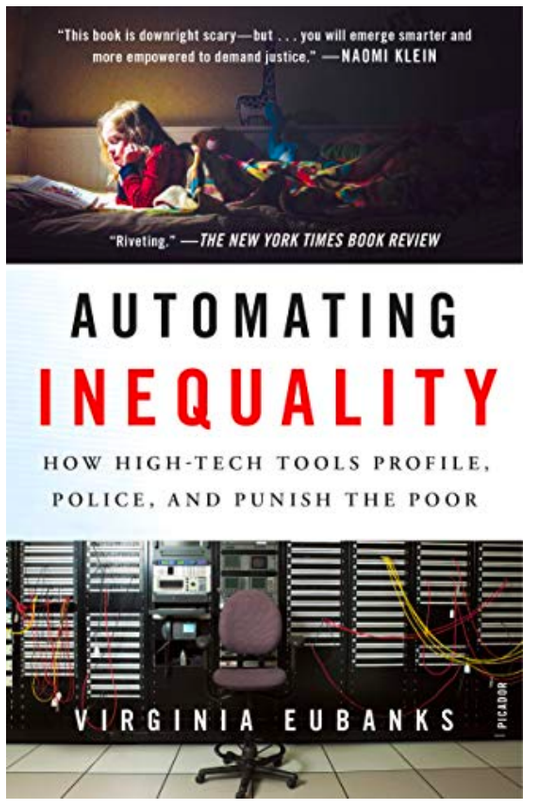
Some Resources











Do Good with Data

Data Science for Social Good: https://www.dssgfellowship.org/

DataKind: https://www.datakind.org/

Data Values & Practices: https://datapractices.org/manifesto/