

# Data Science Ethics

Dr. Maria Tackett

12.03.19



[Click for PDF of slides](#)



STA 199

[datasciencebox.org](http://datasciencebox.org)

# Announcements

- Project Data Analysis due **TODAY** at 11:59p
- HW 06 - due **Friday 12/6** at 11:59p
- Project final write up & presentation - due **Friday, 12/13** at 11:59p
- Presentations Saturday, 12/14
  - Lab 01: 7p - 8p
  - Lab 02: 8p - 9p
  - Lab 03: 9p - 10p
- Exam 02 Extra Credit
  - 90% response rate on course eval: +3 pts on Exam 02 grades

# Agenda

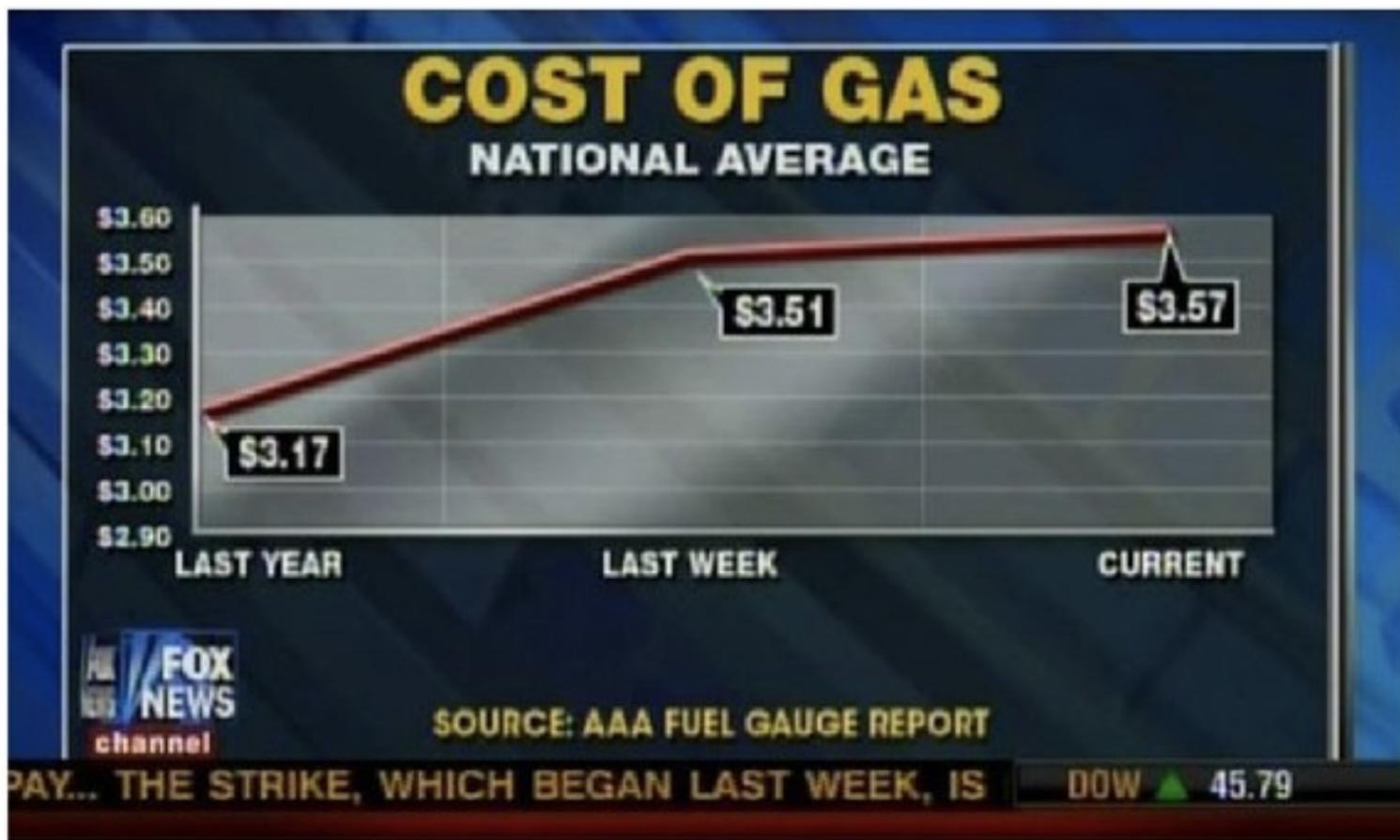
1. Misrepresenting data
2. Misusing p-values
3. Privacy
4. Algorithmic bias



# Misrepresenting data



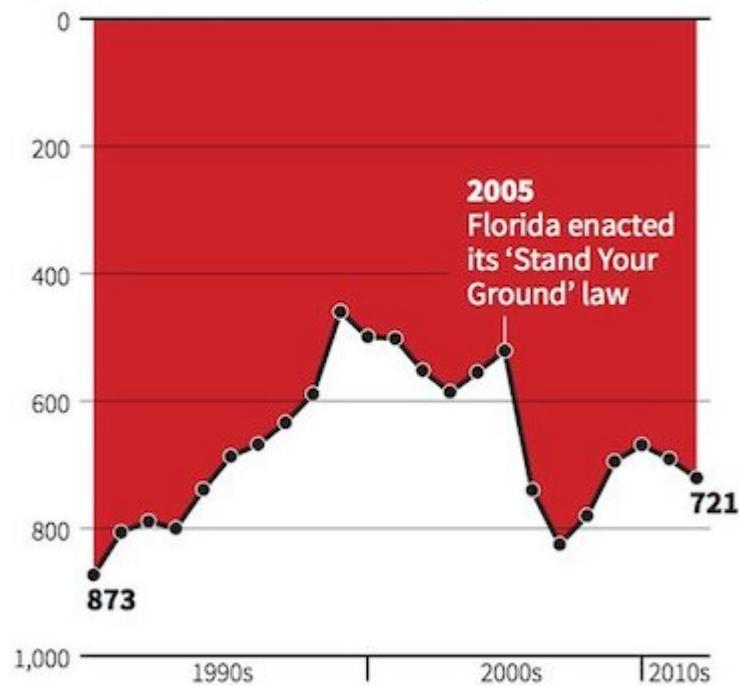
What is wrong with this graph? How would you fix it?



What is wrong with this graph? How would you fix it?

## Gun deaths in Florida

Number of murders committed using firearms

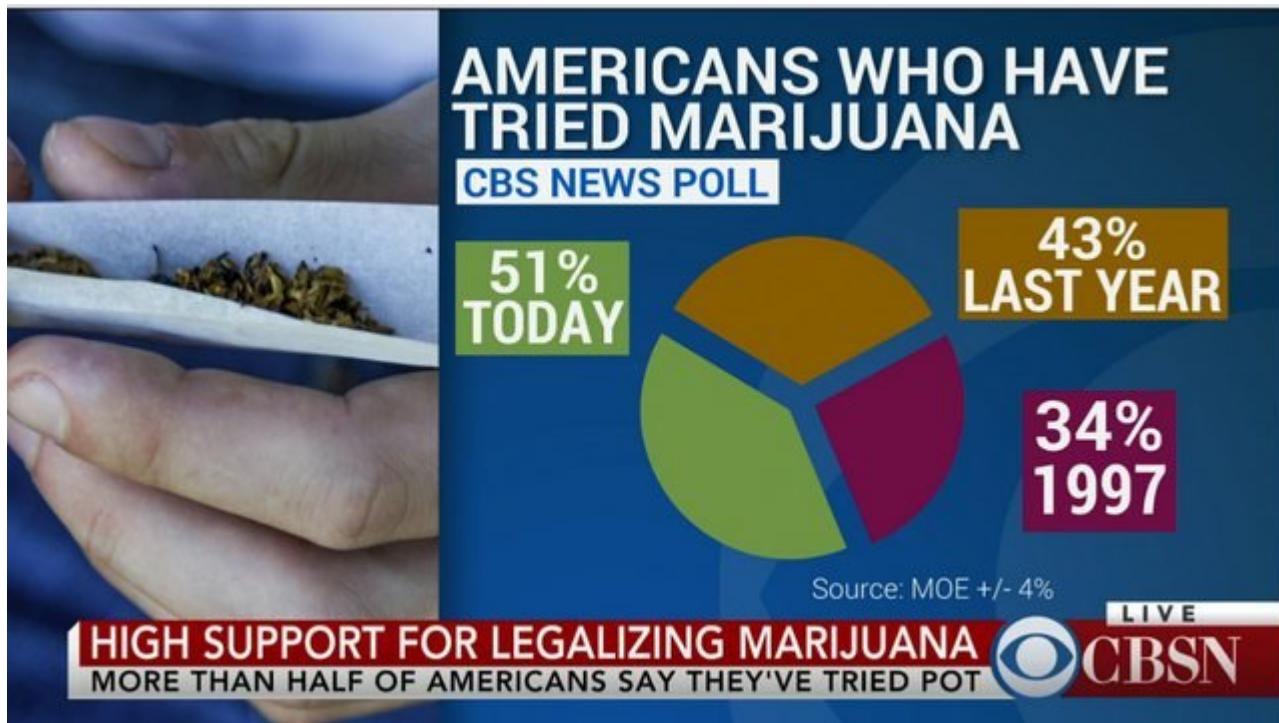


Source: Florida Department of Law Enforcement

C. Chan 16/02/2014

REUTERS

What is wrong with this graph? How would you fix it?



# Misusing p-values



STA 199

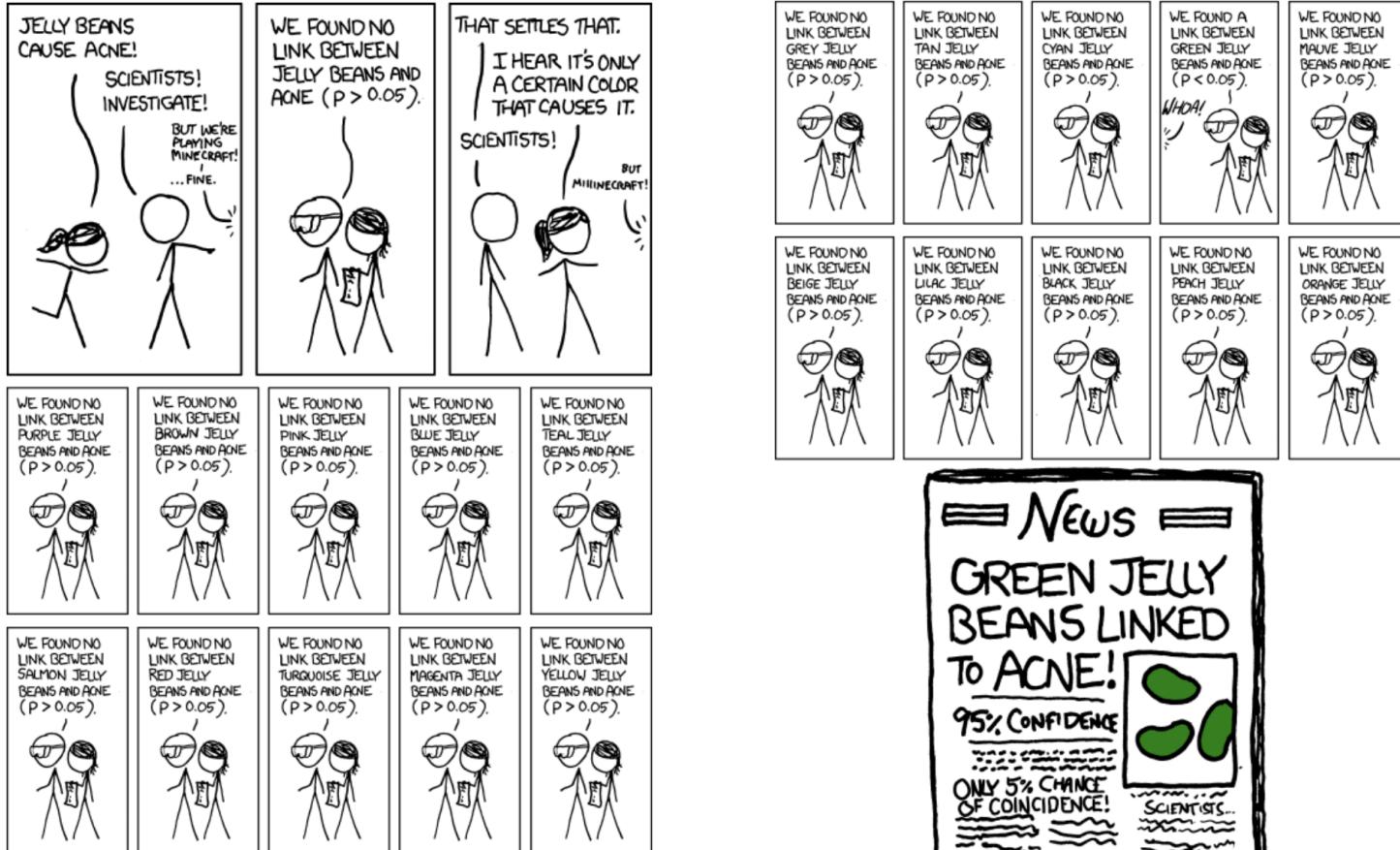
[datasciencebox.org](http://datasciencebox.org)

# What is a p-value?



Source: "[Not Even Scientists Can Easily Explain p-values](#)"

# Statistically significant?



Source: <https://xkcd.com/882/>

*Let's repeat it: P-values don't necessarily tell you if an experiment "worked" or not*

from [800 scientists say it's time to abandon "statistical significance"](#)

# Alternate ways to evaluate evidence

- Concentrate on effect sizes
  - How big of a difference does an intervention make
  - Is it practically meaningful?
- Use confidence intervals to estimate effect size
- Ask whether the result is from a novel study or a replication (put some more weight into a theory many labs have looked into)
- Ask whether underlying data is freely accessible (so anyone can check the math)



Source: [800 scientists say it's time to abandon "statistical significance"](#)

# Alternate ways to evaluate evidence

- Use alternative statistical techniques like likelihood ratios and Bayes factors
  - P-values ask the question “how rare are my results?”
  - Likelihood ratios and Bayes factors ask the question “what is the probability my hypothesis is the best explanation for the results we found?”
- You'll learn more about these in future statistics classes 😊

# Privacy

# OK Cupid 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.



[OKCupid Study Reveals the Perils of Big-Data Science](#)

When collecting and analyzing social media data, how do you make sure you don't violate reasonable expectations of privacy?



**Ethan Jewett** @esjewett · May 11, 2016



Replies to @KirkegaardEmil

@KirkegaardEmil This data set is highly re-identifiable. Even includes usernames? Was any work at all done to anonymize it?



**Emil O W Kirkegaard**

@KirkegaardEmil

@esjewett No. Data is already public.

3 12:30 PM - May 11, 2016



[See Emil O W Kirkegaard's other Tweets](#)



# Facebook & Cambridge Analytica

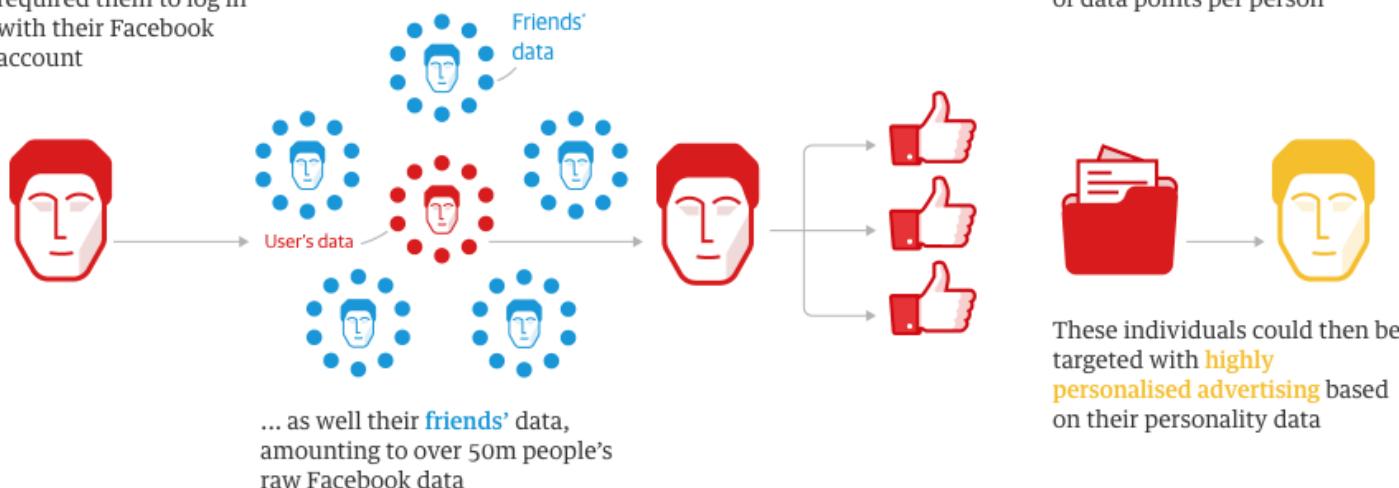
## Cambridge Analytica: how 50m Facebook records were hijacked

**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



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

# Algorithmic bias



# The Hatahway Effect

TECHNOLOGY

# Does Anne Hathaway News Drive Berkshire Hathaway's Stock?

ALEXIS C. MADRIGAL MAR 18, 2011

*Given the awesome correlating powers of today's stock trading computers, the idea may not be as far-fetched as you think.*



"Does Anne Hathaway News Drive Berkshire Hathaway's Stock?"

# The Hathaway Effect

- Oct. 3, 2008 - Rachel Getting Married opens: BRK.A up .44%
- Jan. 5, 2009 - Bride Wars opens: BRK.A up 2.61%
- Feb. 8, 2010 - Valentine's Day opens: BRK.A up 1.01%
- March 5, 2010 - Alice in Wonderland opens: BRK.A up .74%
- Nov. 24, 2010 - Love and Other Drugs opens: BRK.A up 1.62%
- Nov. 29, 2010 - Anne announced as co-host of the Oscars: BRK.A up .25%



The Hathaway Effect: How Anne Gives Warren Buffet a Rise

# Sentencing algorithms

# Machine Bias



There's software used across the country to predict future criminals. And it's biased...

“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.”

Then U.S. Attorney General Eric Holder (2014)

# ProPublica analysis

**Data:** Risk scores assigned to more than 7,000 people arrested in Broward County, Florida, in 2013 and 2014 + whether they were charged with new crimes over the next two years

## Results:

- 20% of those predicted to commit violent crimes actually did
- Algorithm had higher accuracy (61%) when full range of crimes taken into account (e.g. misdemeanors)

	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

Read more at

[propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing](https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing).

# Amazon



STA 199

[datasciencebox.org](http://datasciencebox.org)

# Amazon's Experimental Hiring Algorithm

- Amazon used AI to give job candidates scores ranging from one to five stars - much like shoppers rate products on Amazon, some of the people said
- 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 male candidates were preferable

"Gender bias was not the only issue. Problems with the data that underpinned the models' judgments meant that unqualified candidates were often recommended for all manner of jobs, the people said."

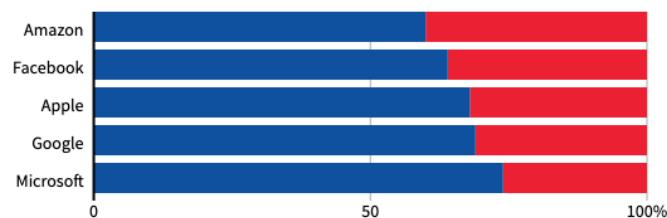
# Amazon's Experimental Hiring Algorithm

## Dominated by men

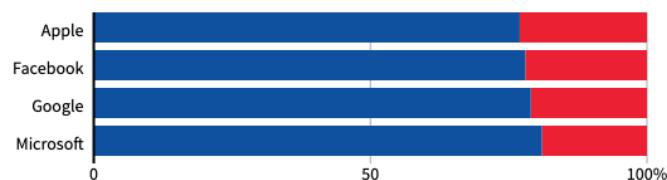
Top U.S. tech companies have yet to close the gender gap in hiring, a disparity most pronounced among technical staff such as software developers where men far outnumber women. Amazon's experimental recruiting engine followed the same pattern, learning to penalize resumes including the word "women's"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

### GLOBAL HEADCOUNT

■ Male ■ Female



### EMPLOYEES IN TECHNICAL ROLES



Note: Amazon does not disclose the gender breakdown of its technical workforce.

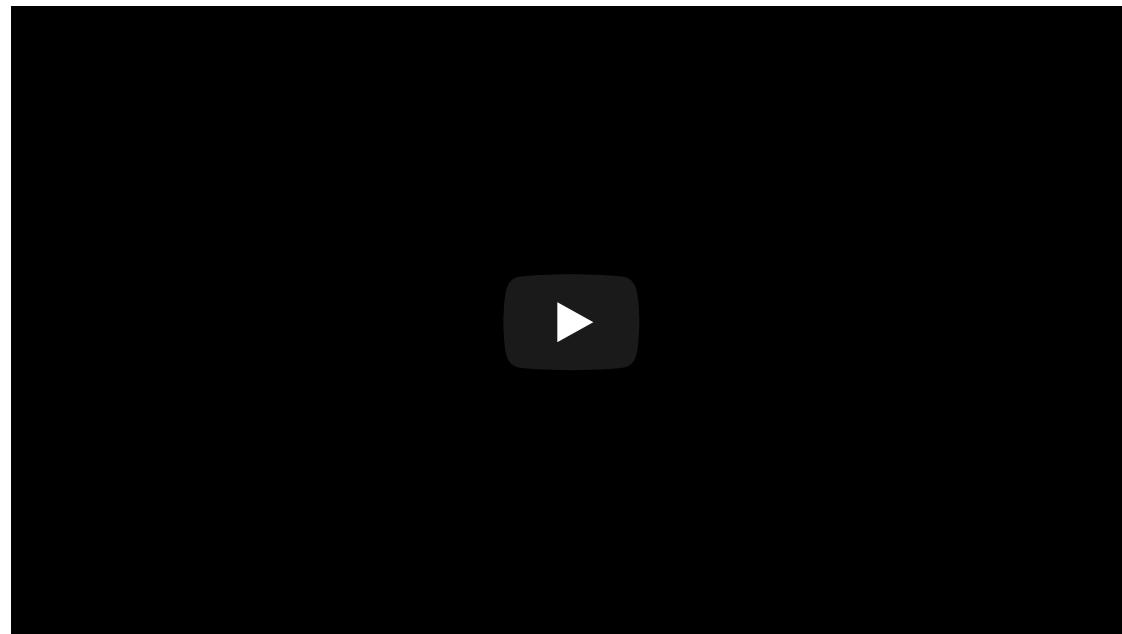
Source: Latest data available from the companies, since 2017.

By Han Huang | REUTERS GRAPHICS

# Algorithmic Justice



# Algorithmic Justice



Interview with Joy Buolamwini: 'A White mask worked better': why algorithms are not colour blind



# Review

A company uses a machine learning algorithm to determine which job advertisement to display for users searching for technology jobs. Based on past results, the algorithm tends to display lower paying jobs for women than for men (after controlling for other characteristics).

What ethical considerations might be considered when reviewing this algorithm?

Source: *Modern Data Science with R*, by Baumer, Kaplan, and Horton

# Further study on data science ethics



# Further Reading

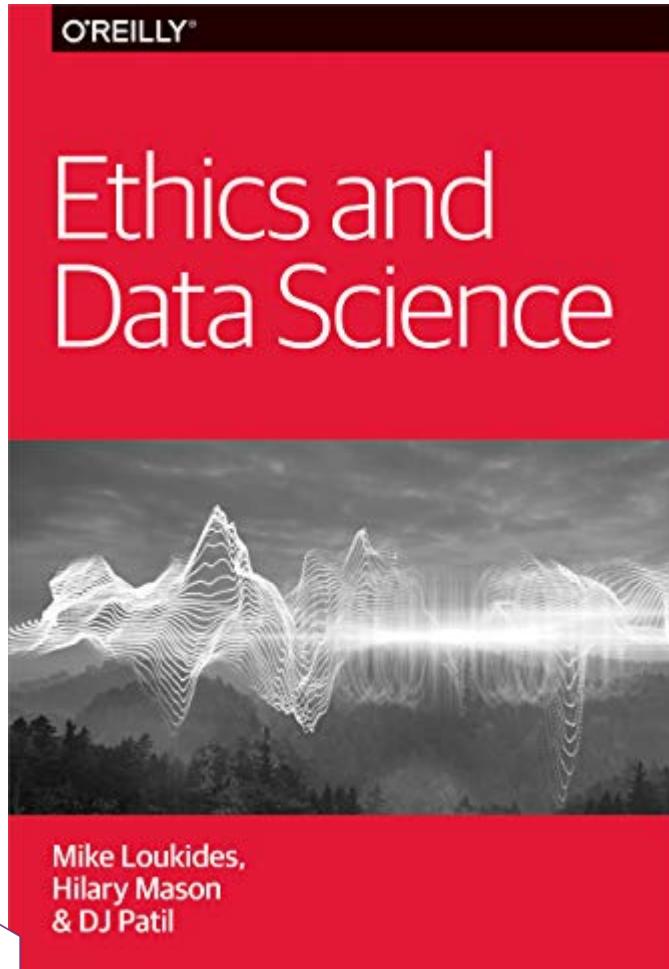


*Statistical Inference in the 21st  
Century: A World Beyond  $p < 0.05$*

March 2019 issue of American  
Statistician

(free access as student through  
Duke library)

# Further reading

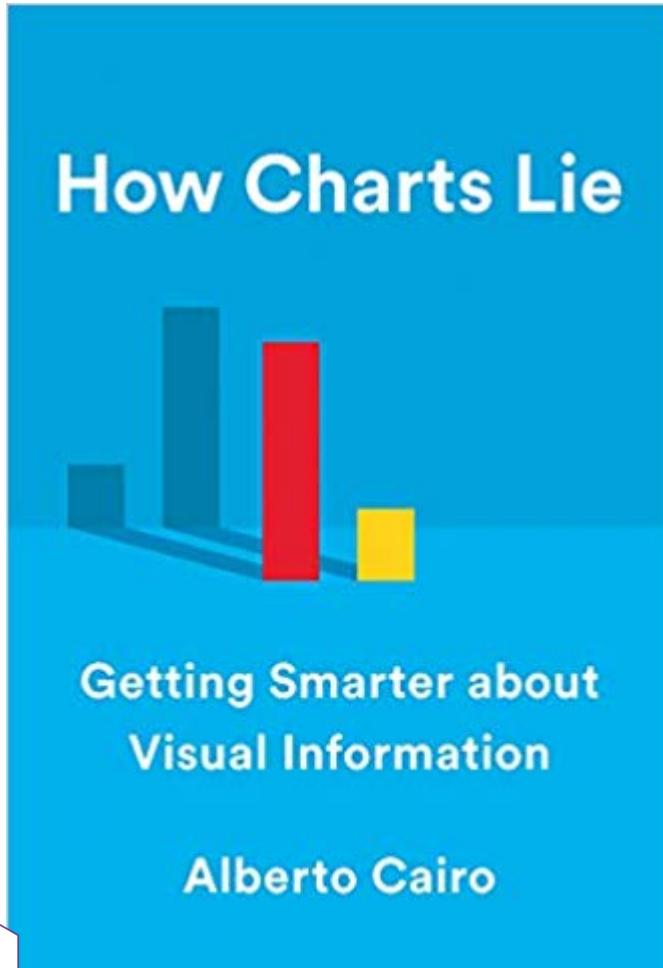


## Ethics and Data Science

by Mike Loukides, Hilary Mason, DJ Patil

(free Kindle download)

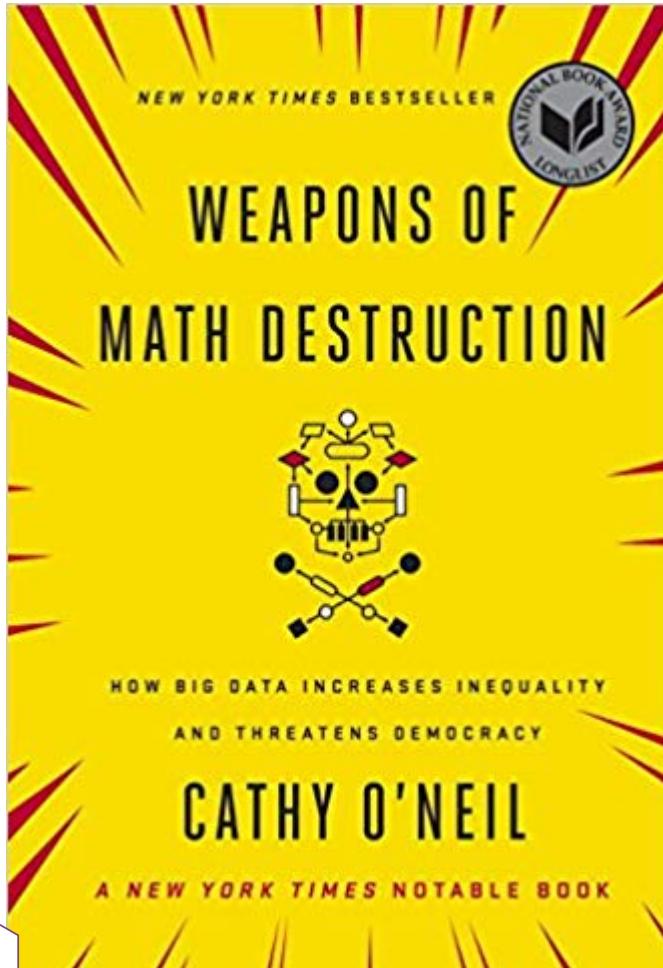
# Further reading



[How Charts Lie: Getting Smarter About Visual Information](#)

by Alberto Cairo

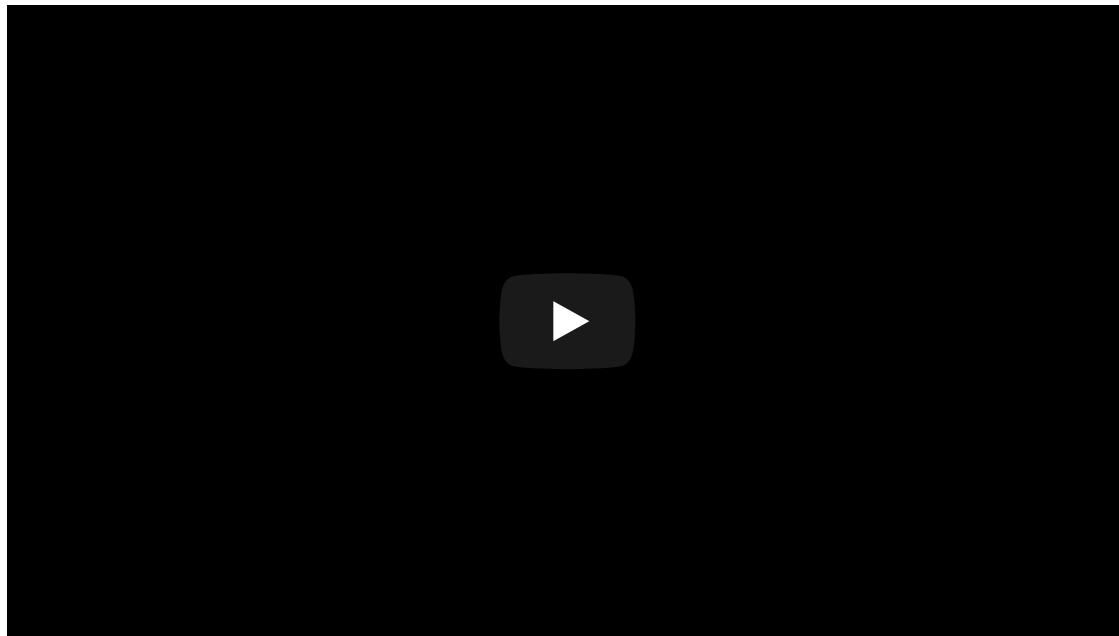
# Further reading



Weapons of Math Destruction:  
How Big Data Increases Inequality,  
and Threatens Democracy.

by Cathy O'Neil

# Further watching



Predictive Policing: Bias In, Bias Out  
by Kristian Lum

# Parting thoughts

- At some point during your data science journey you will learn tools that can be used unethically
- You might also be tempted to use your knowledge in a way that is ethically questionable either because of business goals or for the pursuit of further knowledge (or because your boss told you to do so)

How do you train yourself to make the right decisions (or reduce the likelihood of accidentally making the wrong decisions) at those points?

# Do good with data

- Data Science for Social Good: <http://www.dssgfellowship.org/>
- DataKind: <https://www.datakind.org/>
- Data Values & Practices: <https://datapractices.org/manifesto/>



# Acknowledgement

These slides were largely based on *Data Science Ethics*

