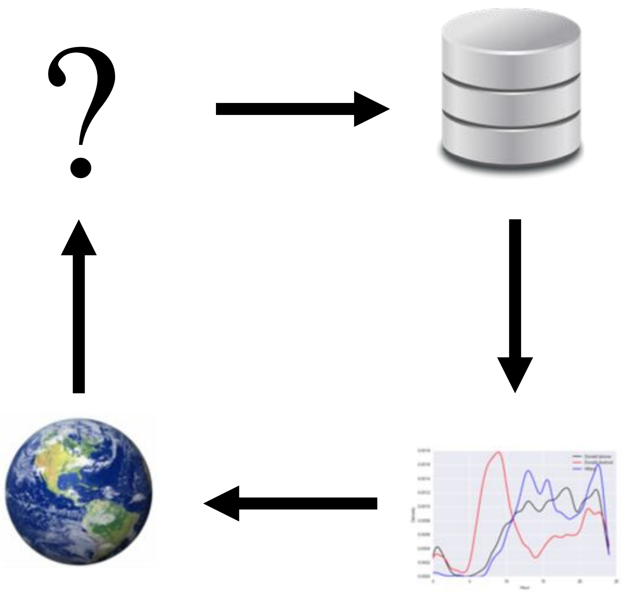
Data Science 100

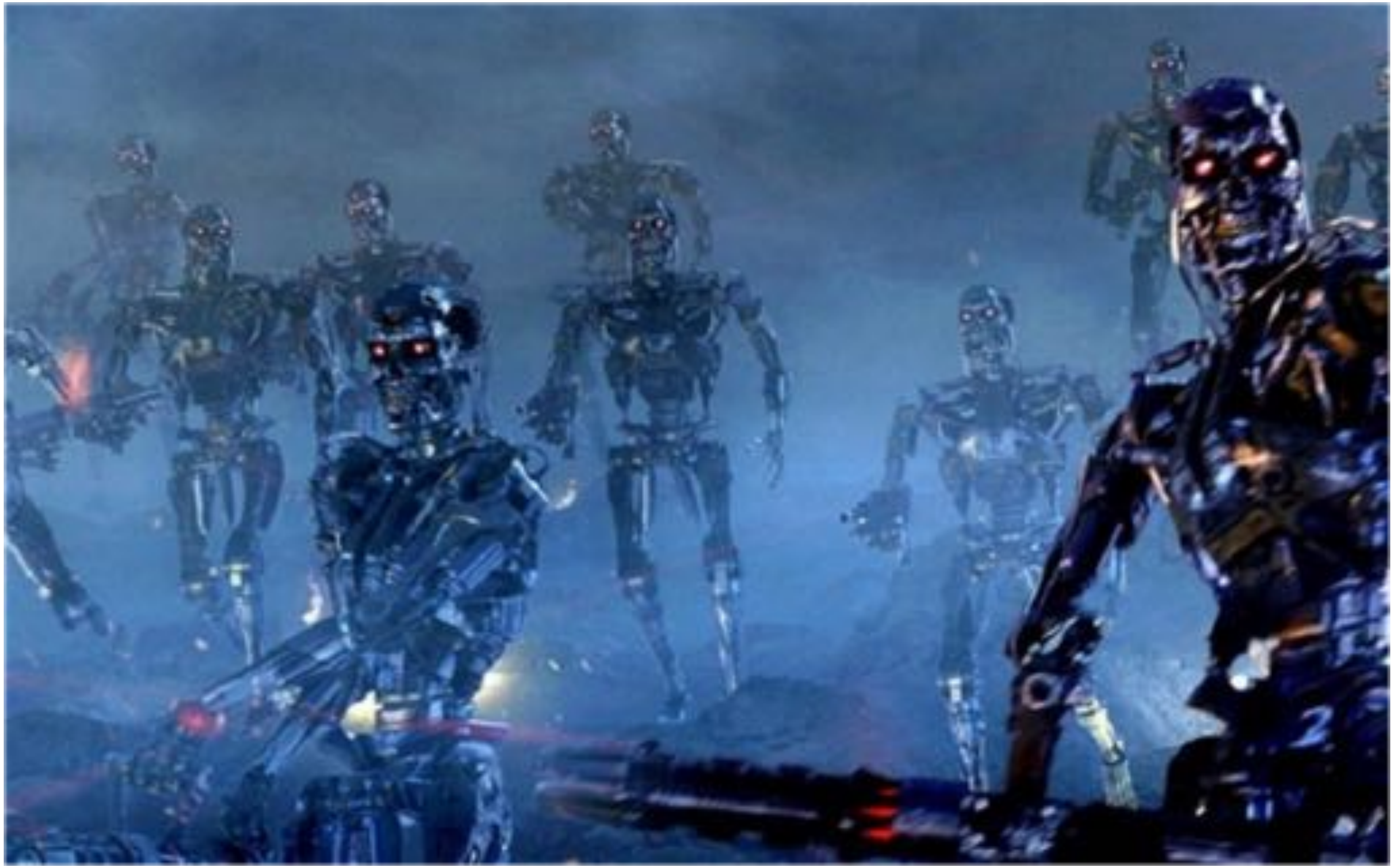
Lecture 27: Ethics in Data Science

Slides by:

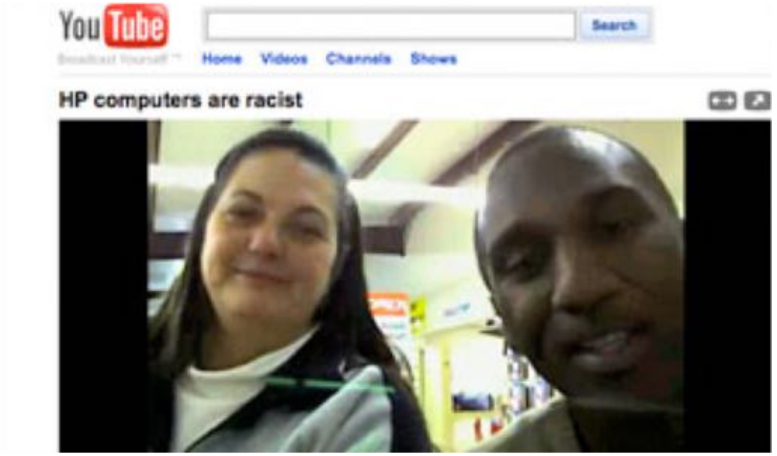
Joshua Kroll

kroll@berkeley.edu






4:00 05/06/2015 UTC
14:01 05/06/2015 UTC



HP Face-Tracking Webcams Don't Recognize Black People

Adam Frucci
12/21/09 10:00am • Filed to: WEBCAMS

154.6K 291

'A white mask worked better': why algorithms are not colour blind

When Joy Buolamwini found that a robot recognised her face better when she wore a white mask, she knew a problem needed fixing

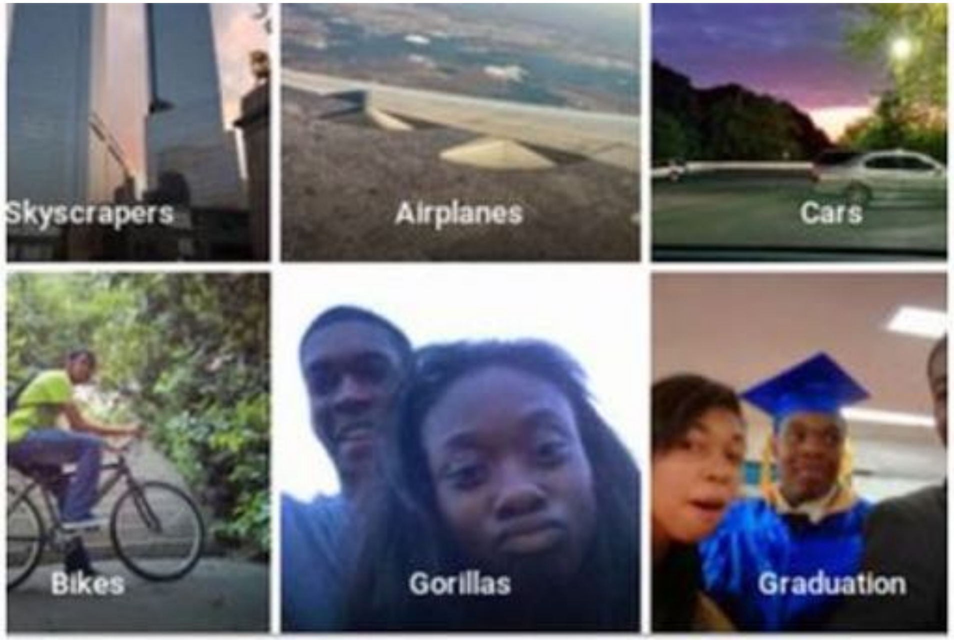
Gender Shades

Gender Classifier	Overall Accuracy on all Subjects in Pilot Parlaiments Benchmark (2017)
Microsoft	93.7%
FACE++	90.0%
IBM	87.9%

Gender Classifier	Female Subjects Accuracy	Male Subjects Accuracy	Error Rate Diff.
Microsoft	89.3%	97.4%	8.1%
FACE++	78.7%	99.3%	20.6%
IBM	79.7%	94.4%	14.7%

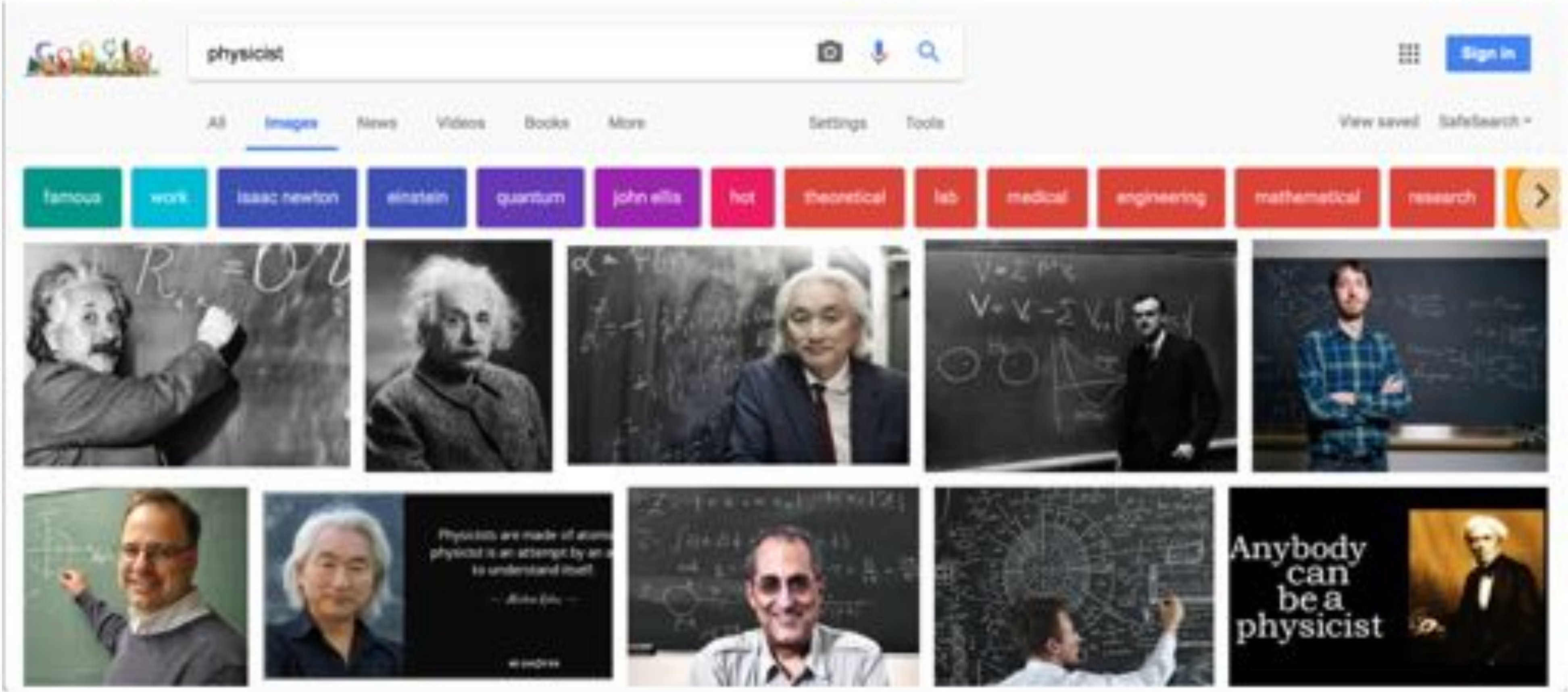
Gender Classifier	Darker Subjects Accuracy	Lighter Subjects Accuracy	Error Rate Diff.
Microsoft	87.1%	99.3%	12.2%
FACE++	83.5%	95.3%	11.8%
IBM	77.6%	96.8%	19.2%

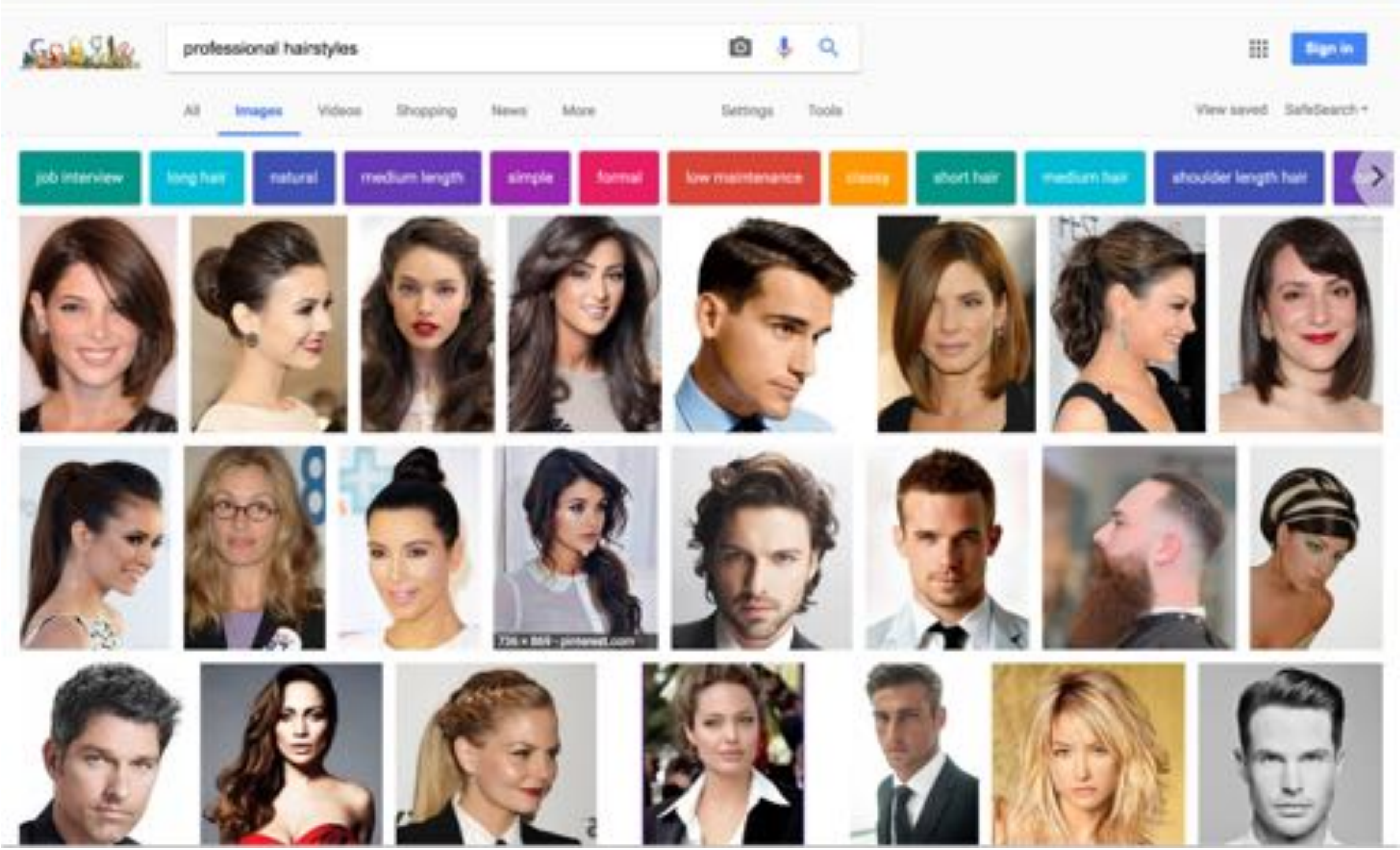
Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%



Google Photos Mistakenly Labels Black People 'Gorillas'

BY CONOR DOUGHERTY JULY 1, 2015 7:01 PM 41


physicist
Sign in
All
Images
News
Videos
Books
More
Settings
Tools
View saved
SafeSearch
famous
work
isaac newton
einstein
quantum
john ellis
hot
theoretical
lab
medical
engineering
mathematical
research
Anybody can be a physicist



professional hairstyles
Sign in
All
Images
Videos
Shopping
News
More
Settings
Tools
View saved
SafeSearch
job interview
long hair
natural
medium length
simple
formal
low maintenance
classy
short hair
medium hair
shoulder length hair

Data science is everywhere…

Scoring Systems

FICO

ChexSystems

Lotteries

U.S. Department of Commerce · Bureau of the Census

Criminal Justice & National Security

TSA Pre✓

Transportation Security Administration

National Security Agency · United States of America

PREDPOL

News

AUTOMATED INSIGHTS

NarrativeScience

Lending

zest finance

earnest

Kabbage

affirm

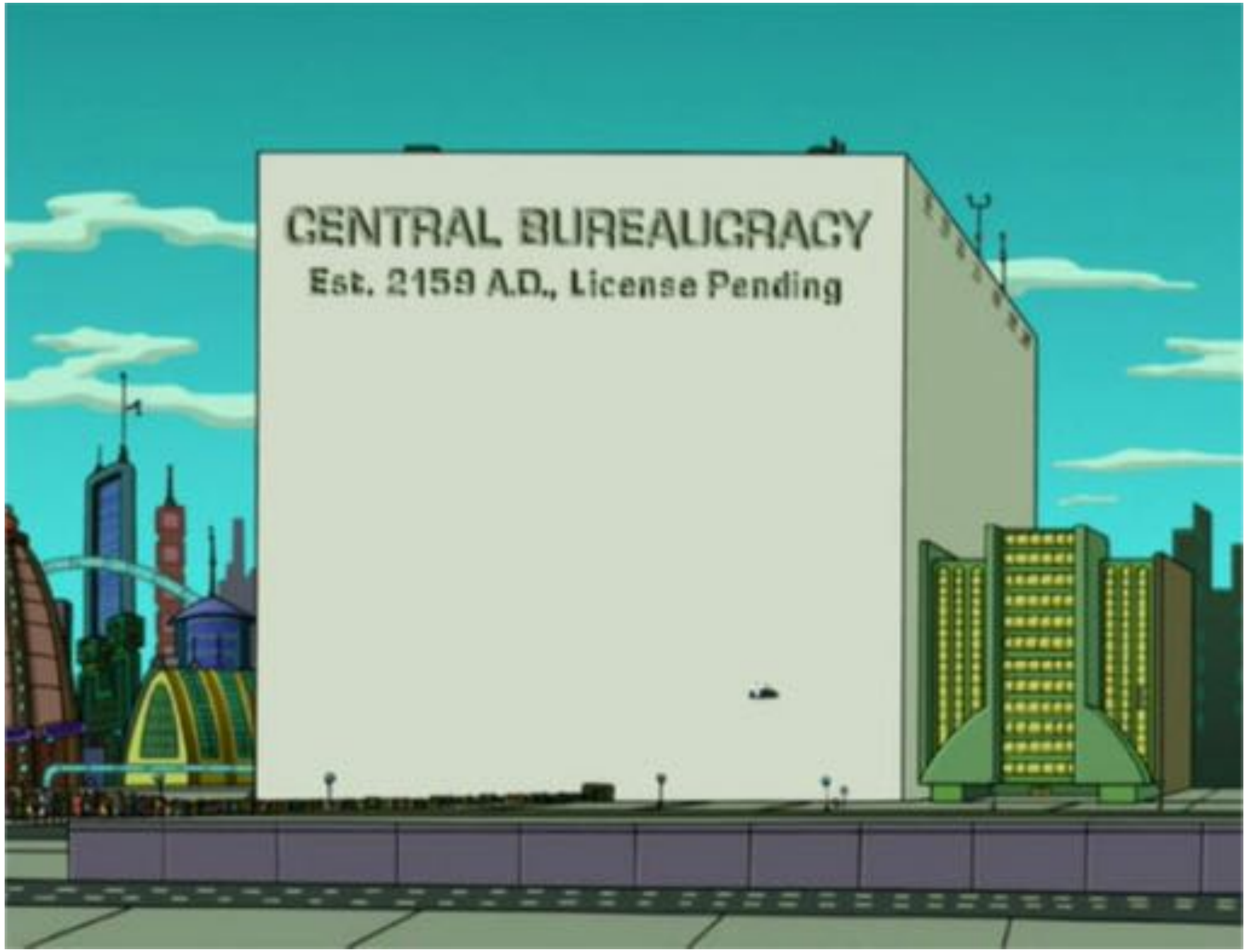
Hiring

gild

TalentBin

GJ

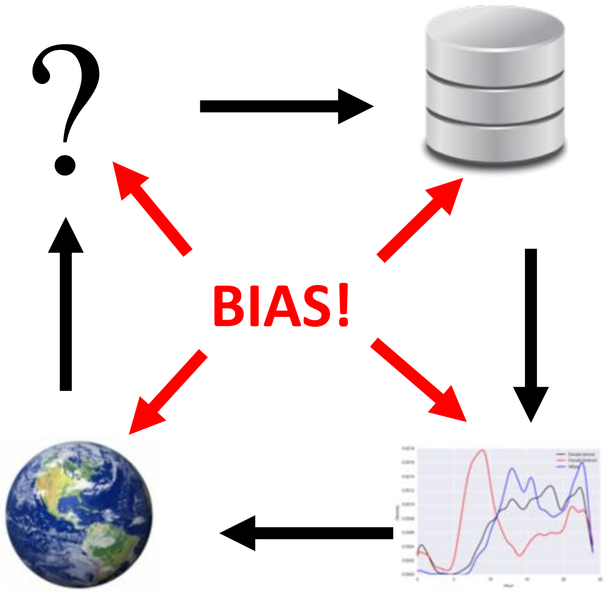
entelo


CENTRAL BUREAUCRACY
Est. 2159 A.D., License Pending


WOPR
War Operation Plan Response

Data Science Life Cycle

Context
Question
Refine Question to an one answerable with data



Design
Data Collection
Data Cleaning

Modeling
Test-train split
Loss function choice
Feature engineering
- Transformations,
- Dummy Variables

Model selection
- Best subset regression
- Cross-Validation

Model evaluation
Prediction error

Context & Refinement

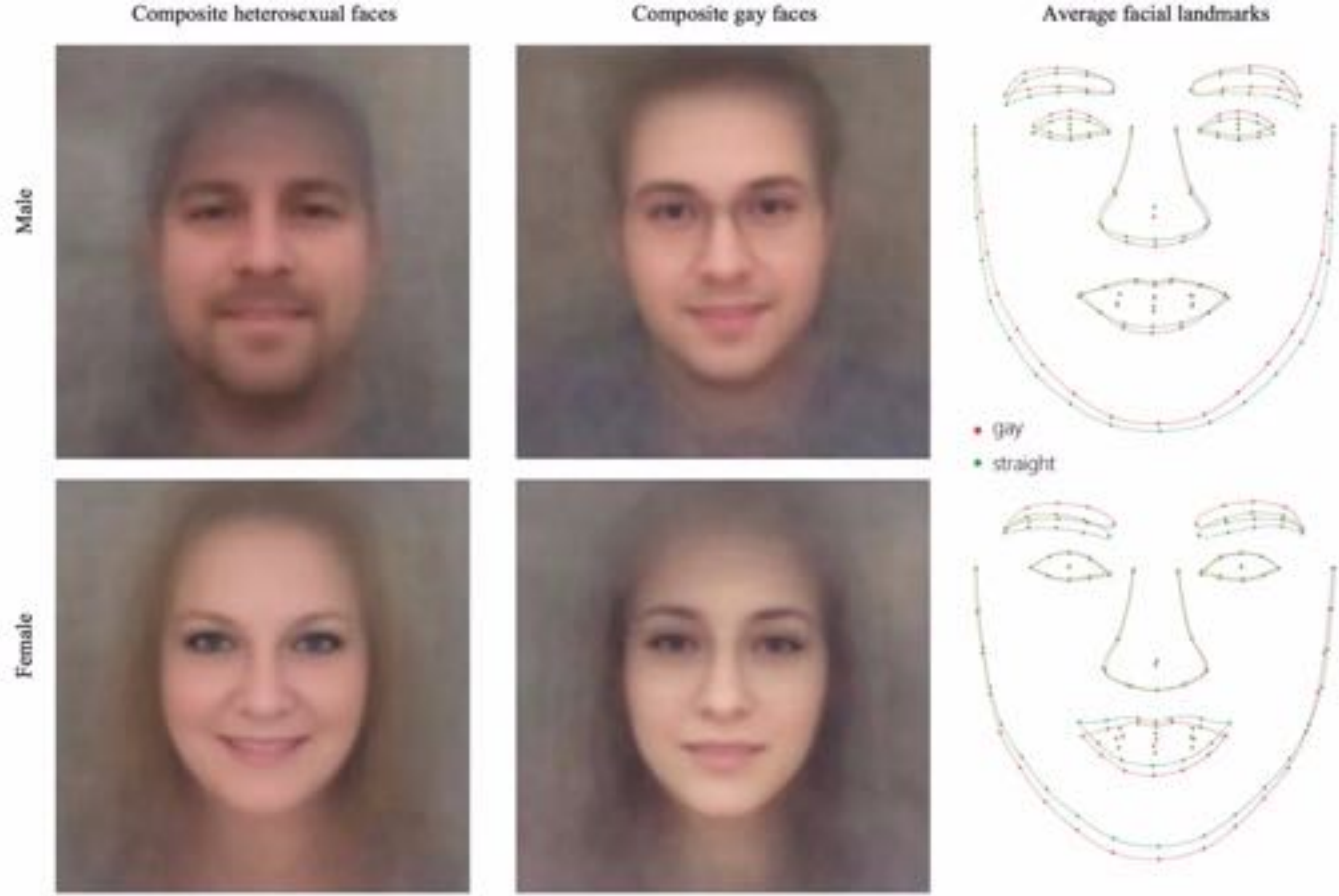
Extreme Digital Vetting of Visitors to the U.S. Moves Forward Under a New Name

ICE officials have invited tech companies, including Microsoft, to develop algorithms that will track visa holders' social media activity.

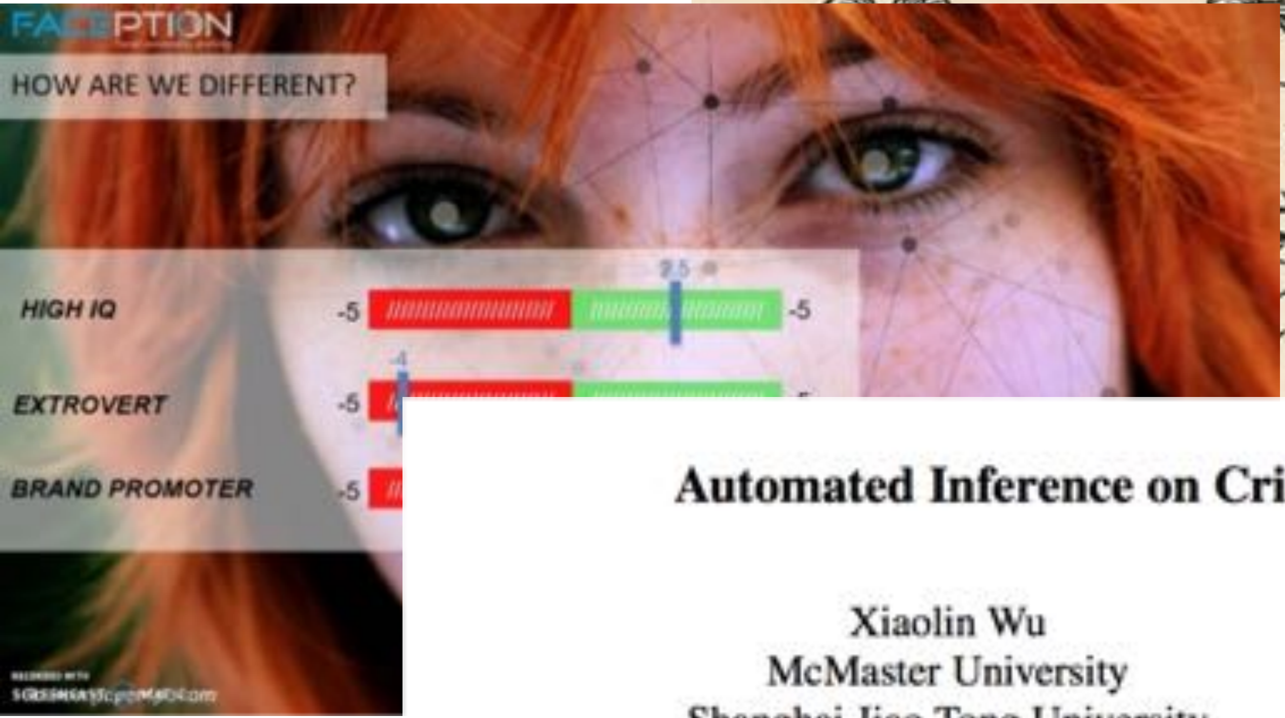
by **George Joseph**, Nov. 22, 8 a.m. EST

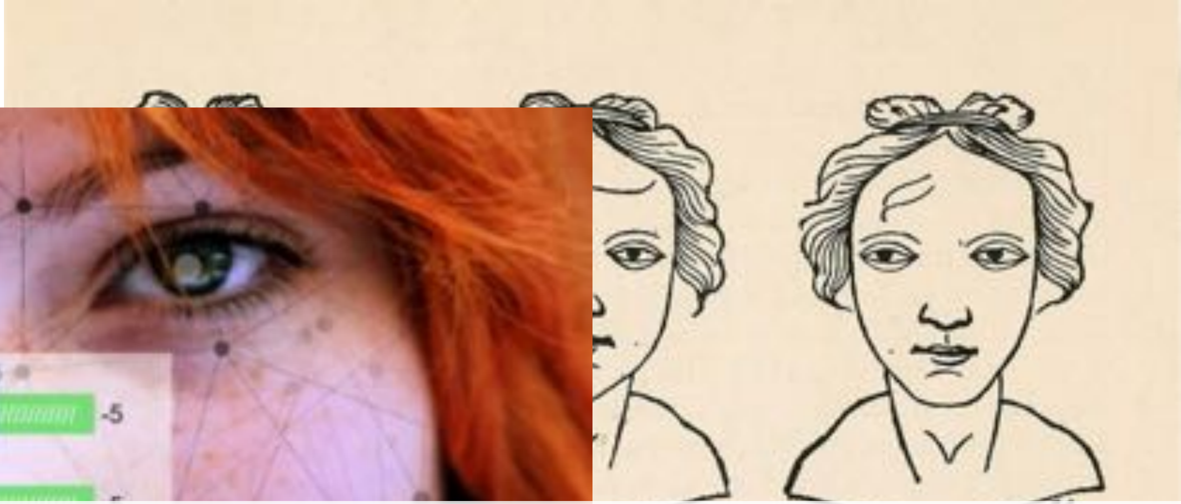
Context & Refinement

A Stanford scientist says he built a gaydar using “the lamest” AI to prove a point


Composite heterosexual faces
Composite gay faces
Average facial landmarks
Male
Female
gay
straight

Proxies



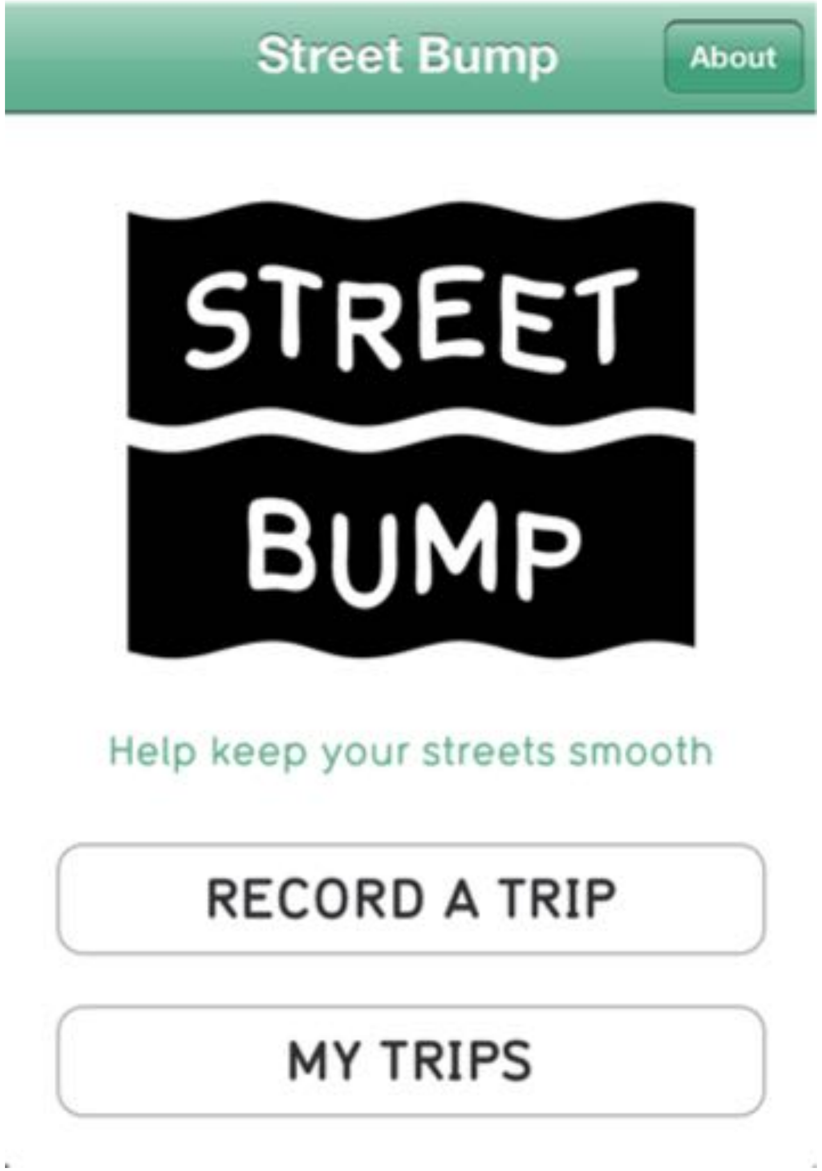


Automated Inference on Criminality using Face Images

Xiaolin Wu
McMaster University
Shanghai Jiao Tong University
xwu510@gmail.com

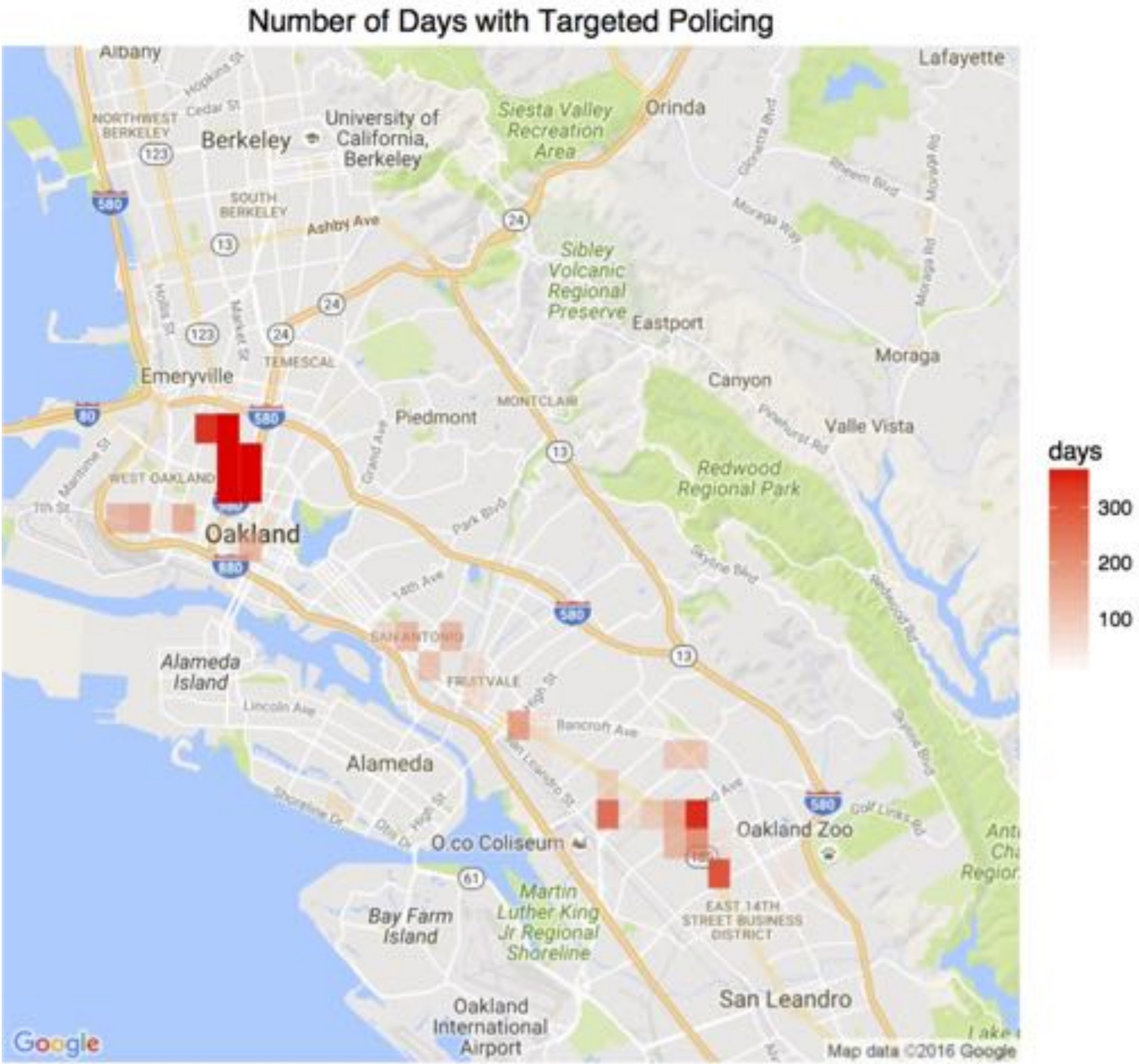
Xi Zhang
Shanghai Jiao Tong University
zhangxi.19930818@sjtu.edu.cn

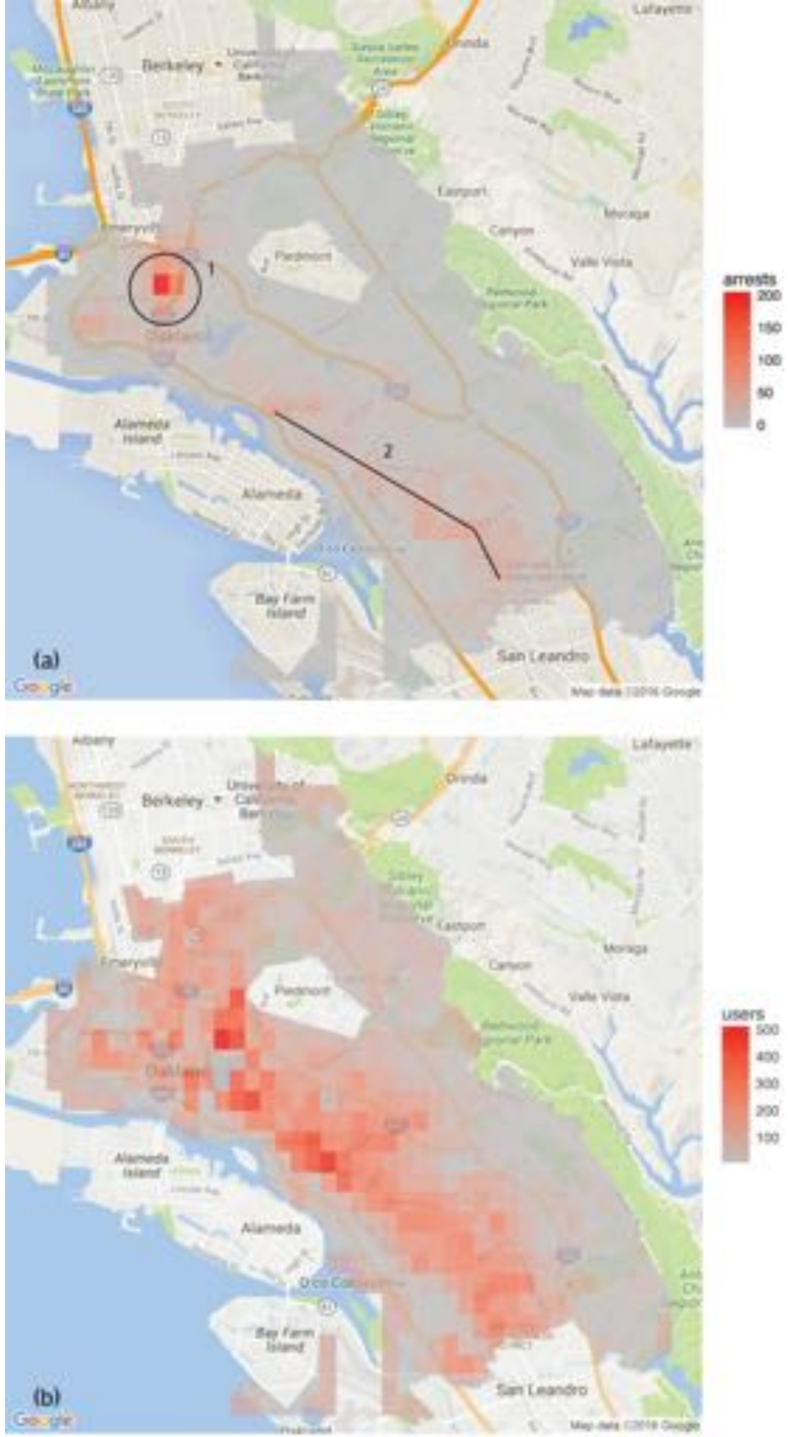
Data Collection



Data Collection

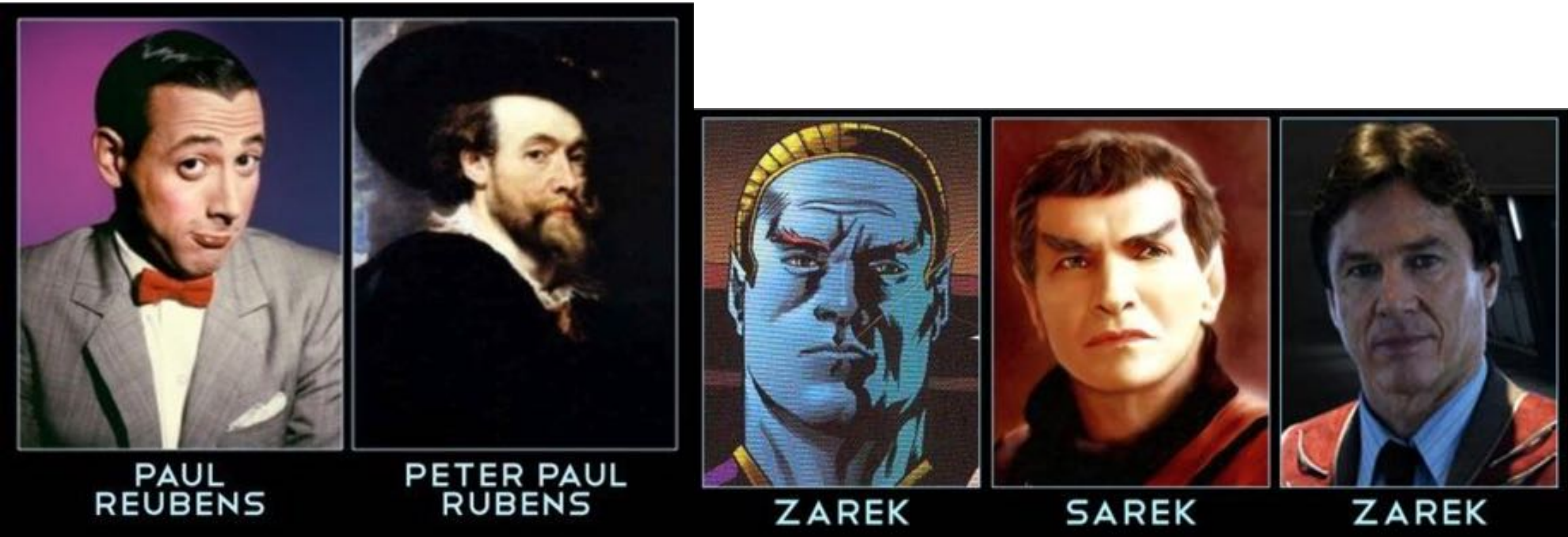
http://bit.ly/ds100-sp18-eth





From Lum, Kristian, and William Isaac. "To predict and serve?." Significance 13, no. 5 (2016): 14-19.

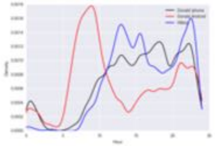
Data Cleaning



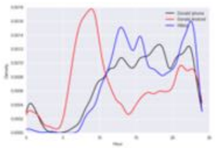
Data Cleaning

Data Cleaning

Test-Train Split

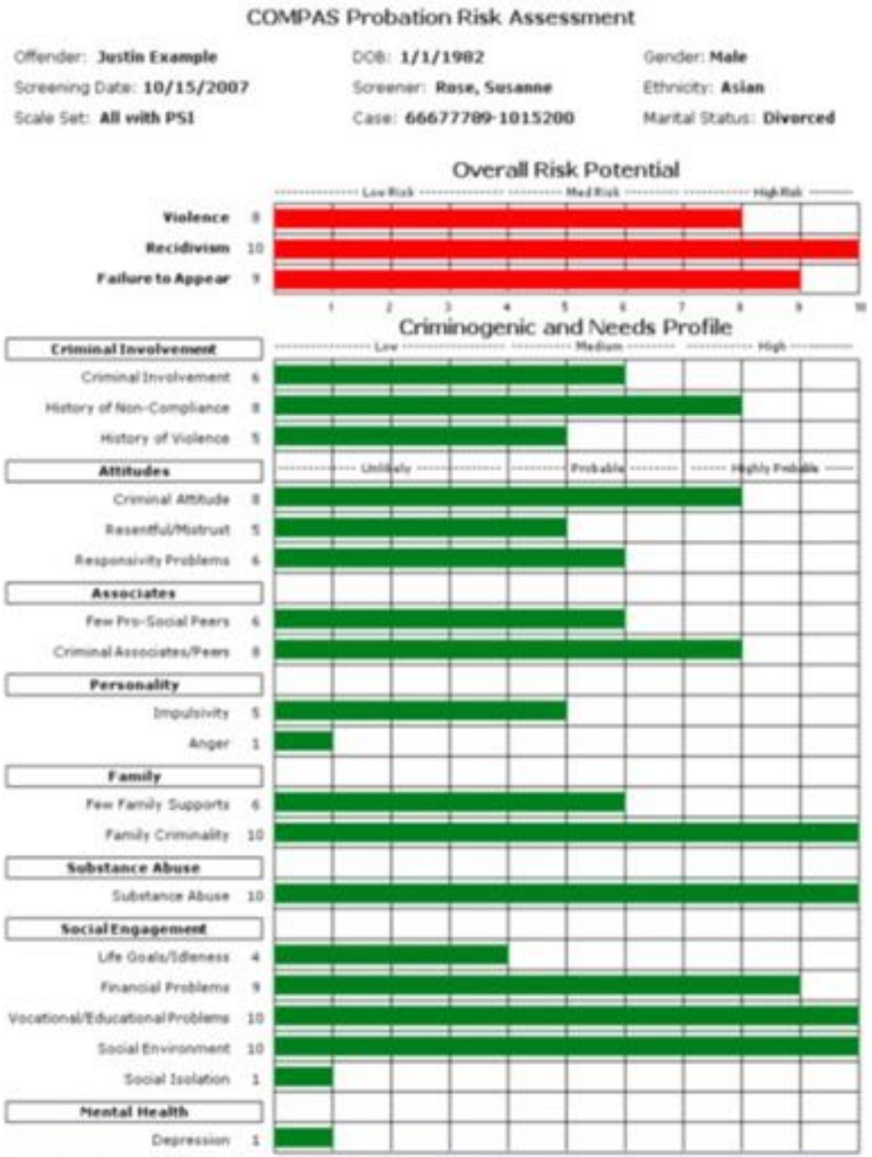


Loss Function Choice



DOES UNCONSCIOUS RACIAL BIAS AFFECT TRIAL JUDGES?

Jeffrey J. Rachlinski, Sheri Lynn Johnson,† Andrew J. Wistrich,‡ & Chris Guthrie††*

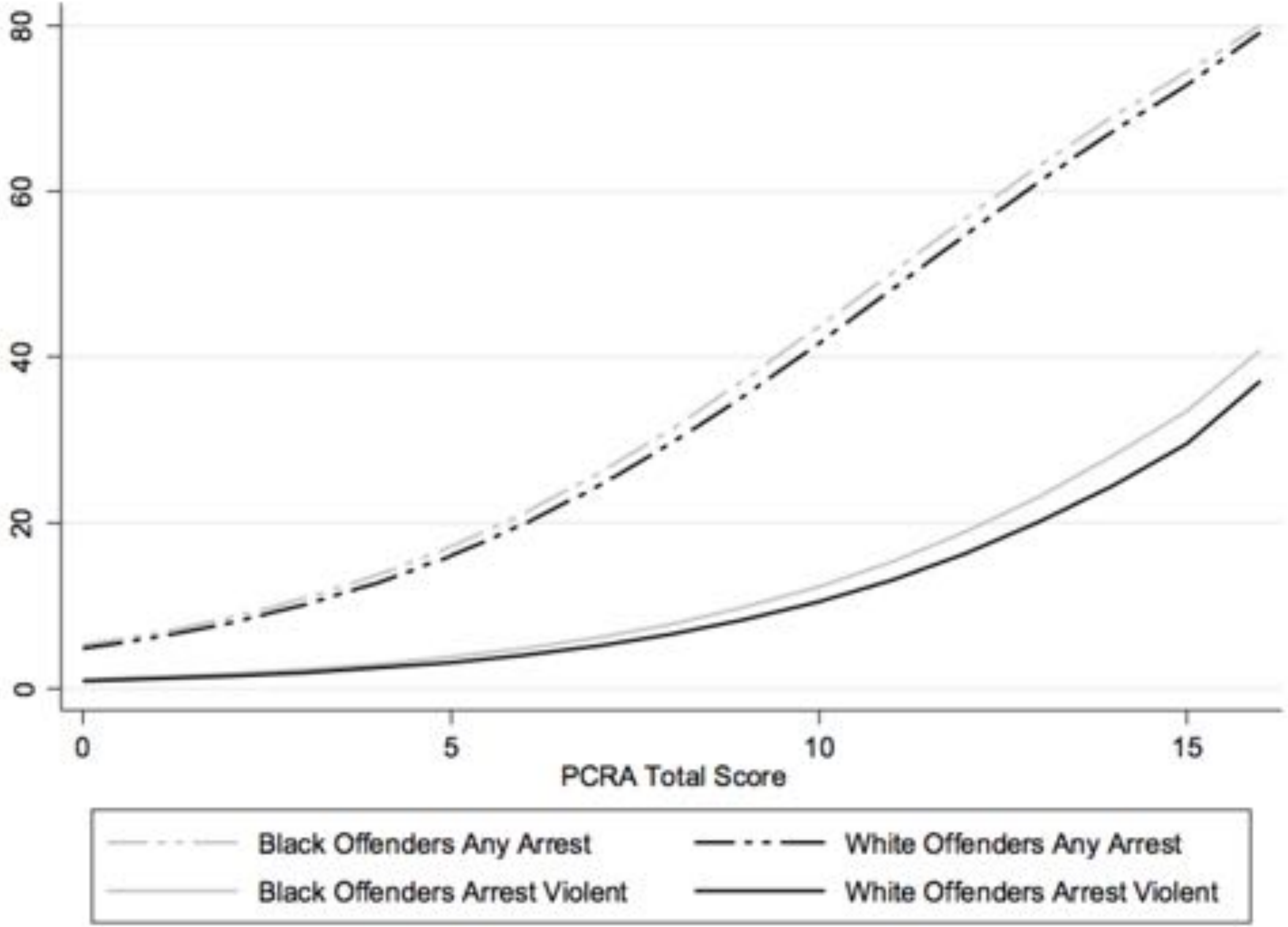


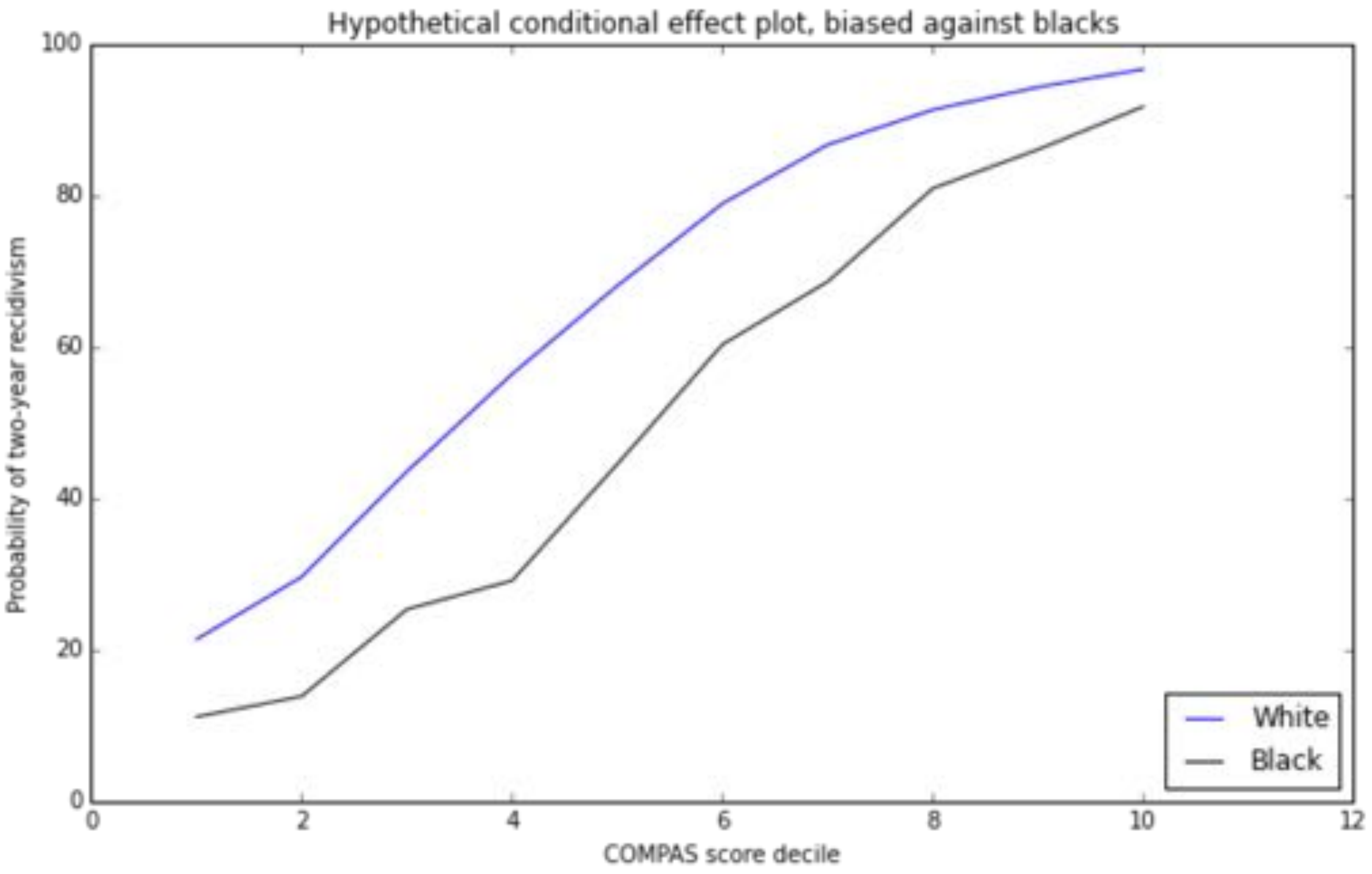
Machine Bias

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

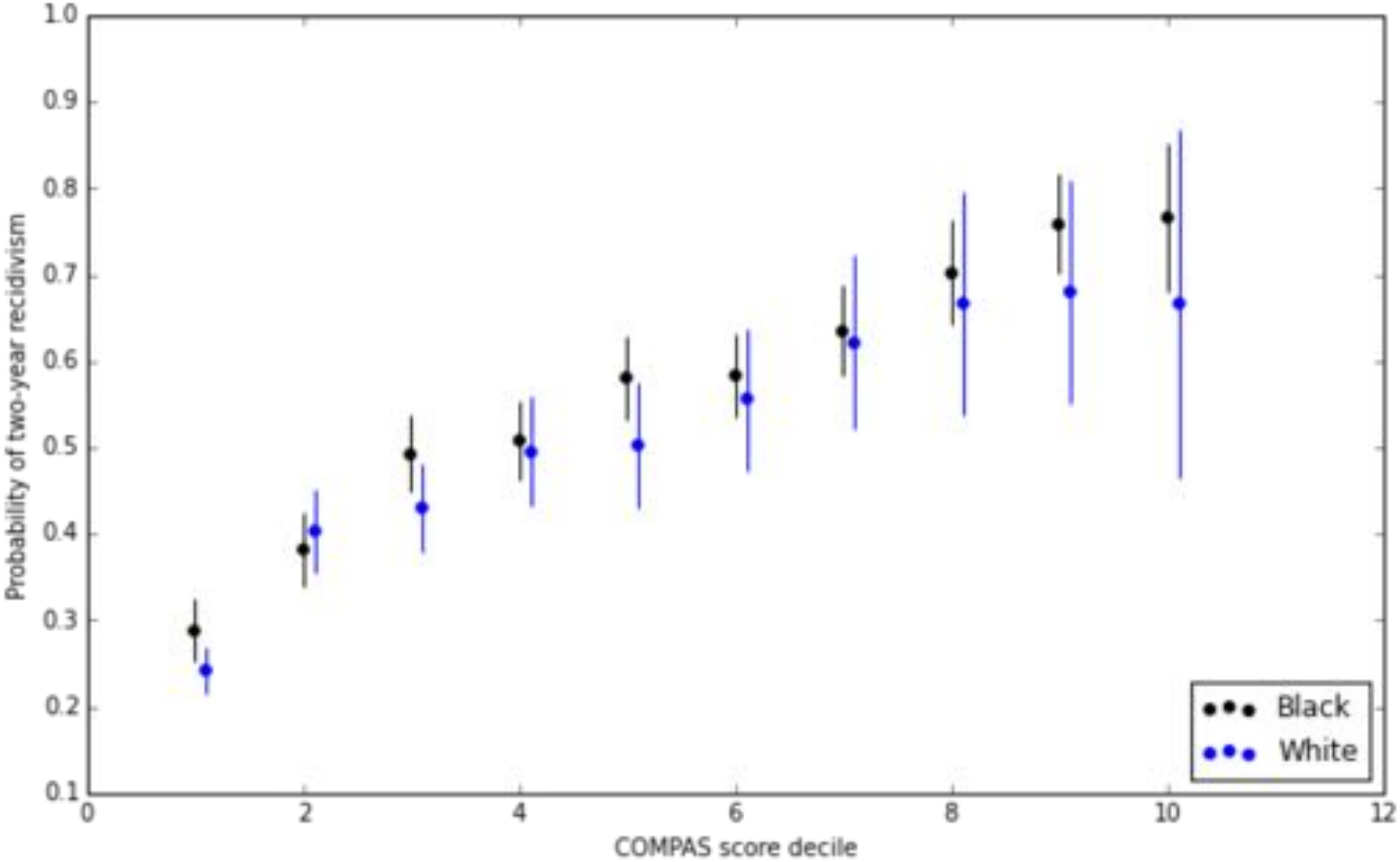
by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica
May 23, 2016

ON A SPRING AFTERNOON IN 2014, Brisha Borden was running late to pick up her god-sister from school when she spotted an unlocked kid's blue Huffy bicycle and a silver Razor scooter. Borden and a friend grabbed the bike and scooter and tried to ride them down the street in the Fort Lauderdale suburb of Coral Springs.

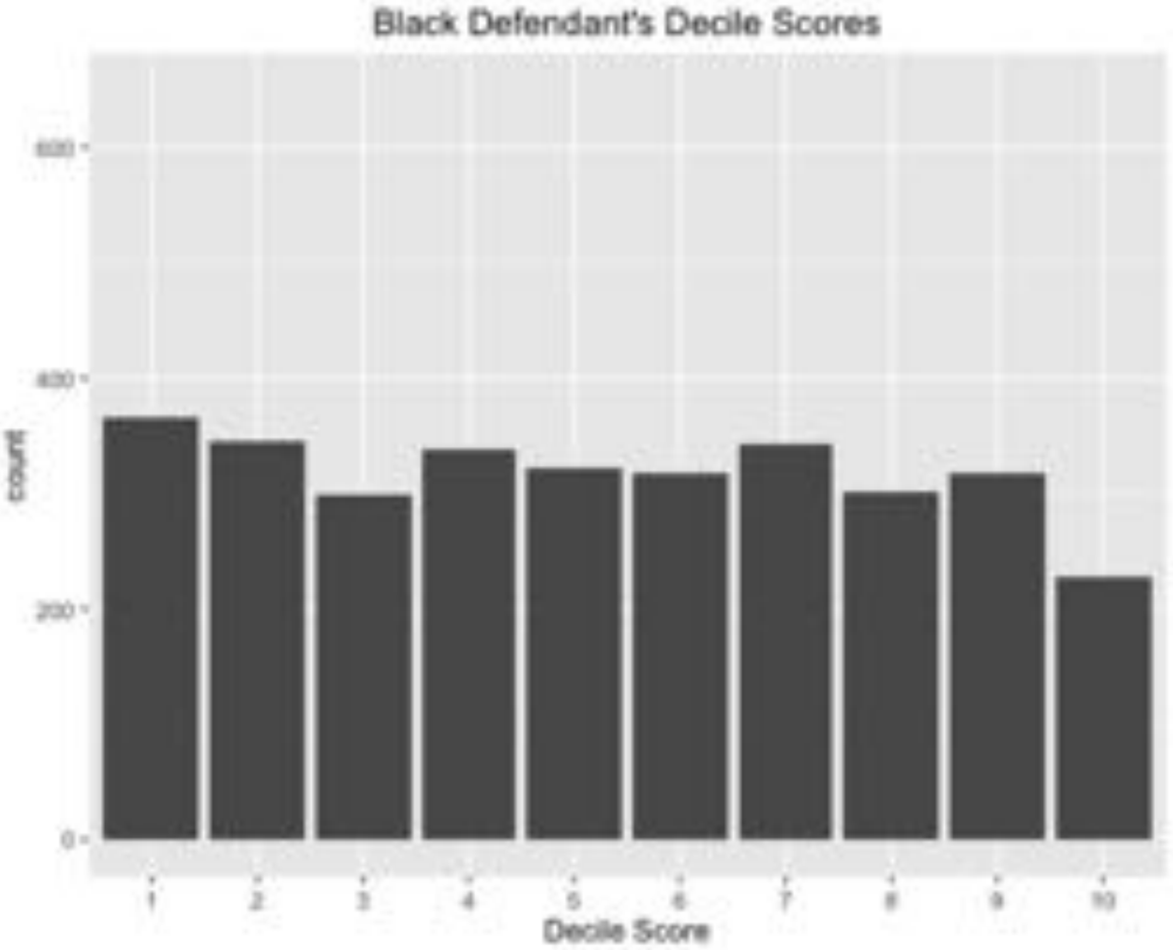
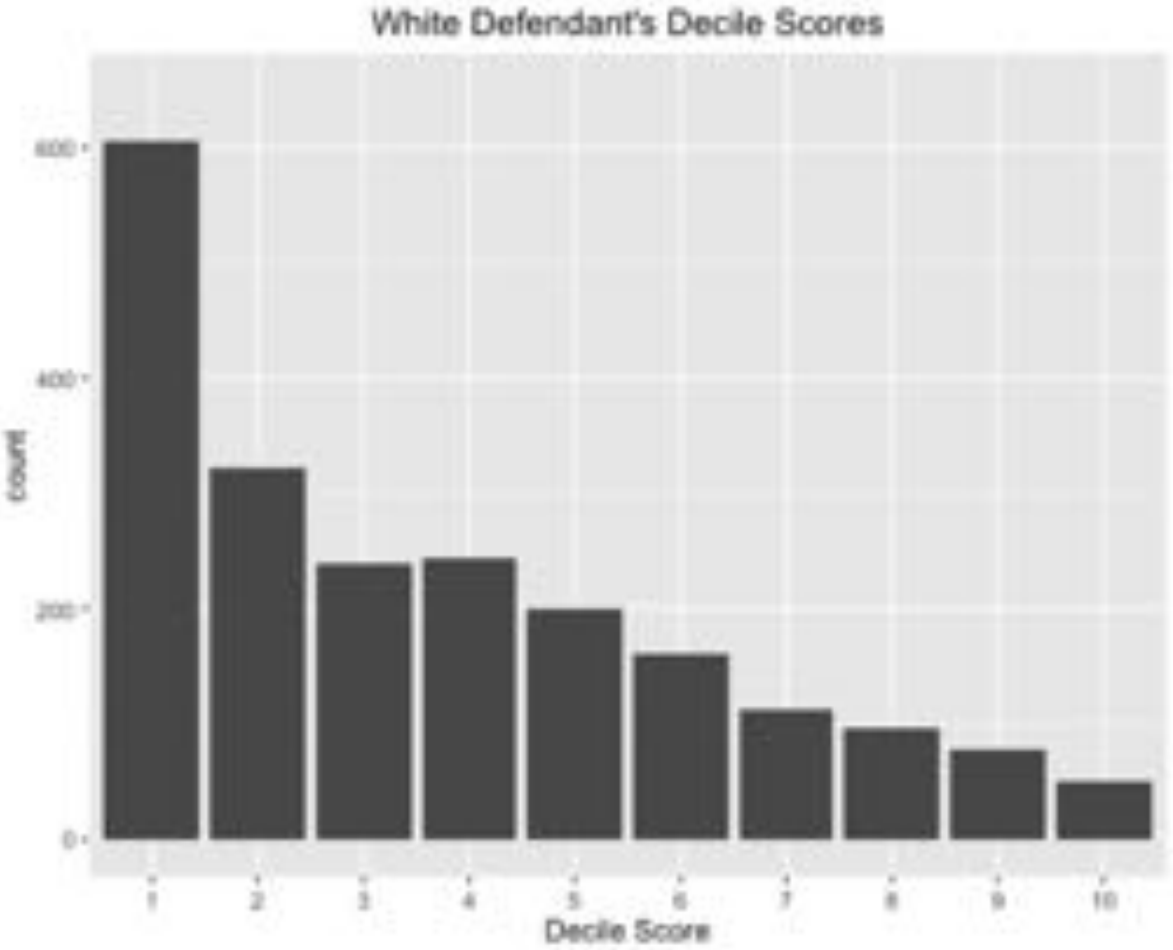

80
60
40
20
0
0
5
10
15
PCRA Total Score
Black Offenders Any Arrest
White Offenders Any Arrest
Black Offenders Arrest Violent
White Offenders Arrest Violent



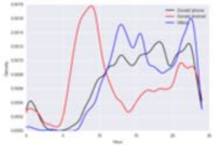
Hypothetical conditional effect plot, biased against blacks
Probability of two-year recidivism
COMPAS score decile
White
Black
0
20
40
60
80
100
0
2
4
6
8
10
12

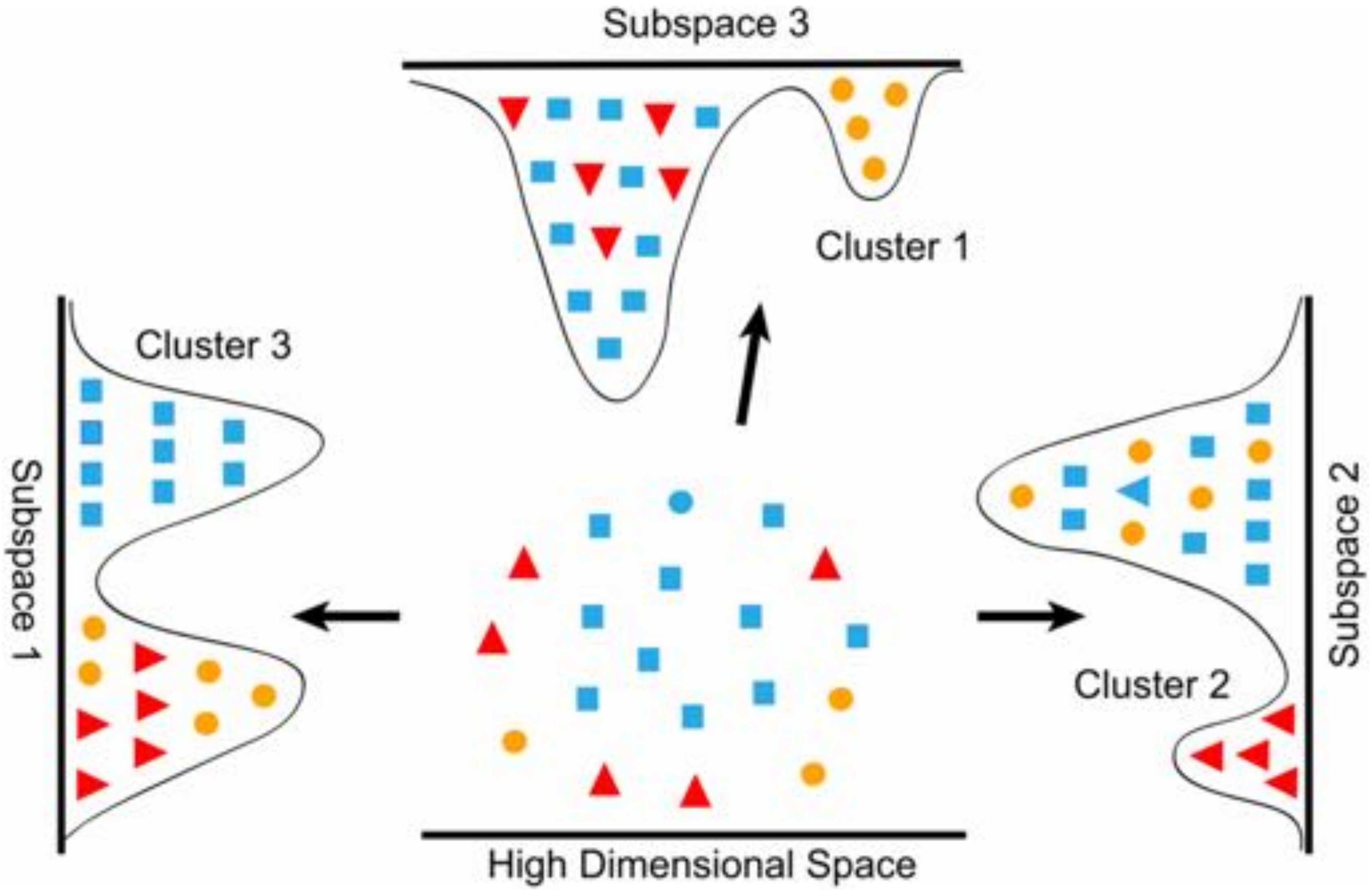

1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
Probability of two-year recidivism
0
2
4
6
8
10
12
COMPAS score decile
Black
White

So why does ProPublica think COMPAS is biased?

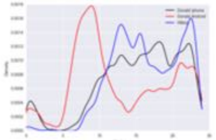

Black Defendant's Decile Scores
count
Decile Score

White Defendant's Decile Scores
count
Decile Score

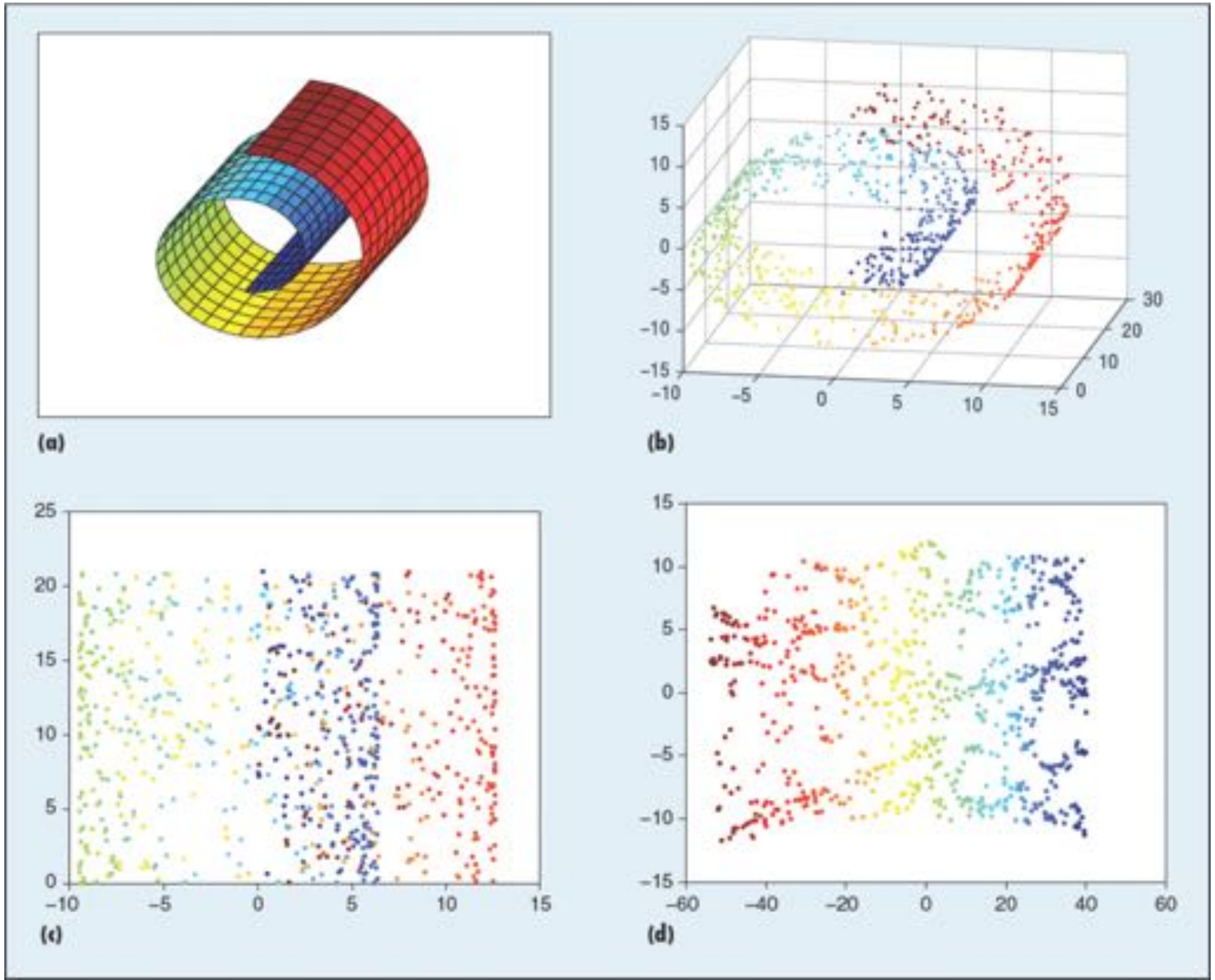
Feature Engineering



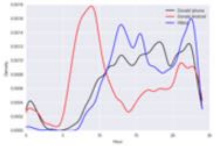


Feature Engineering

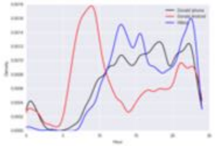


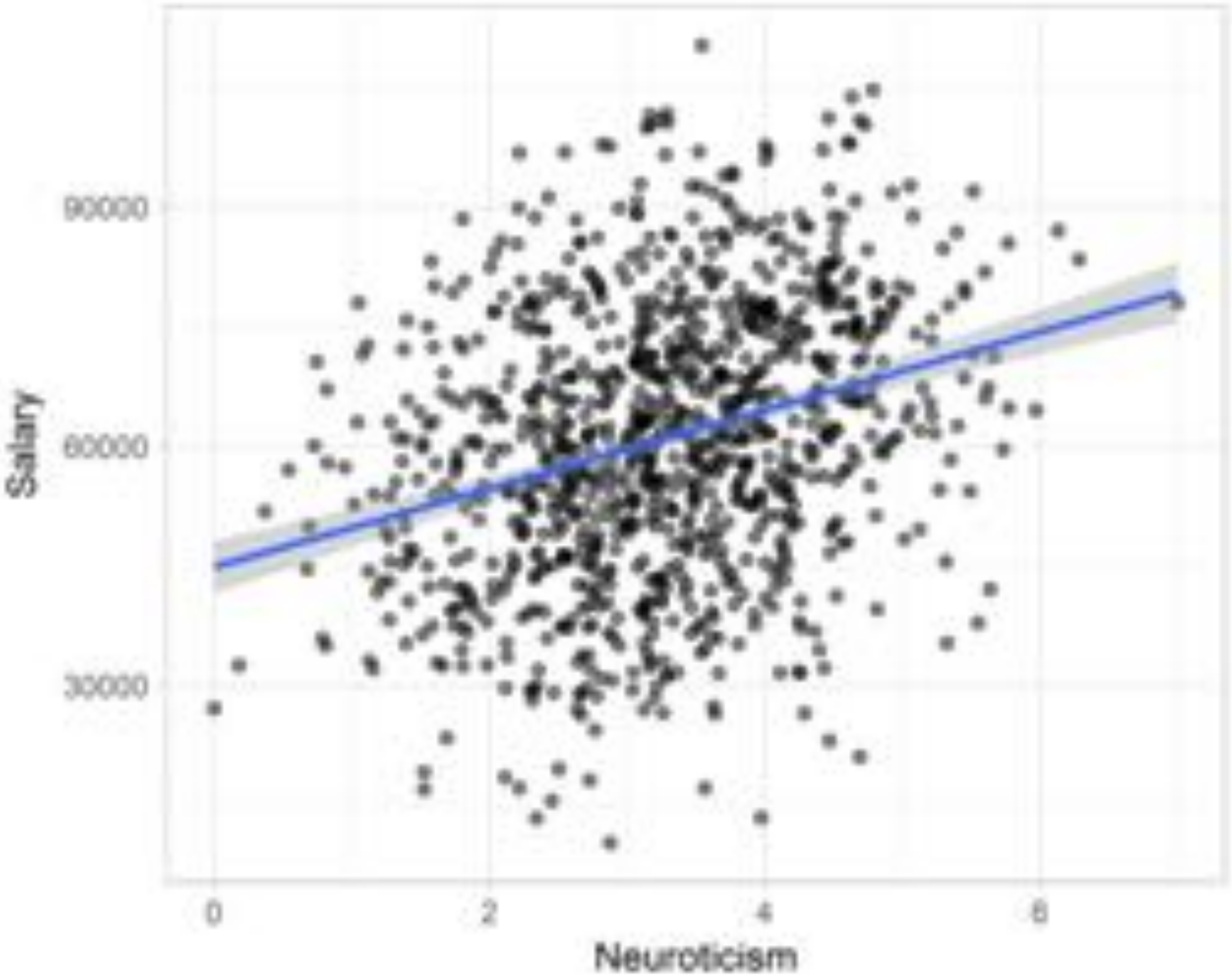


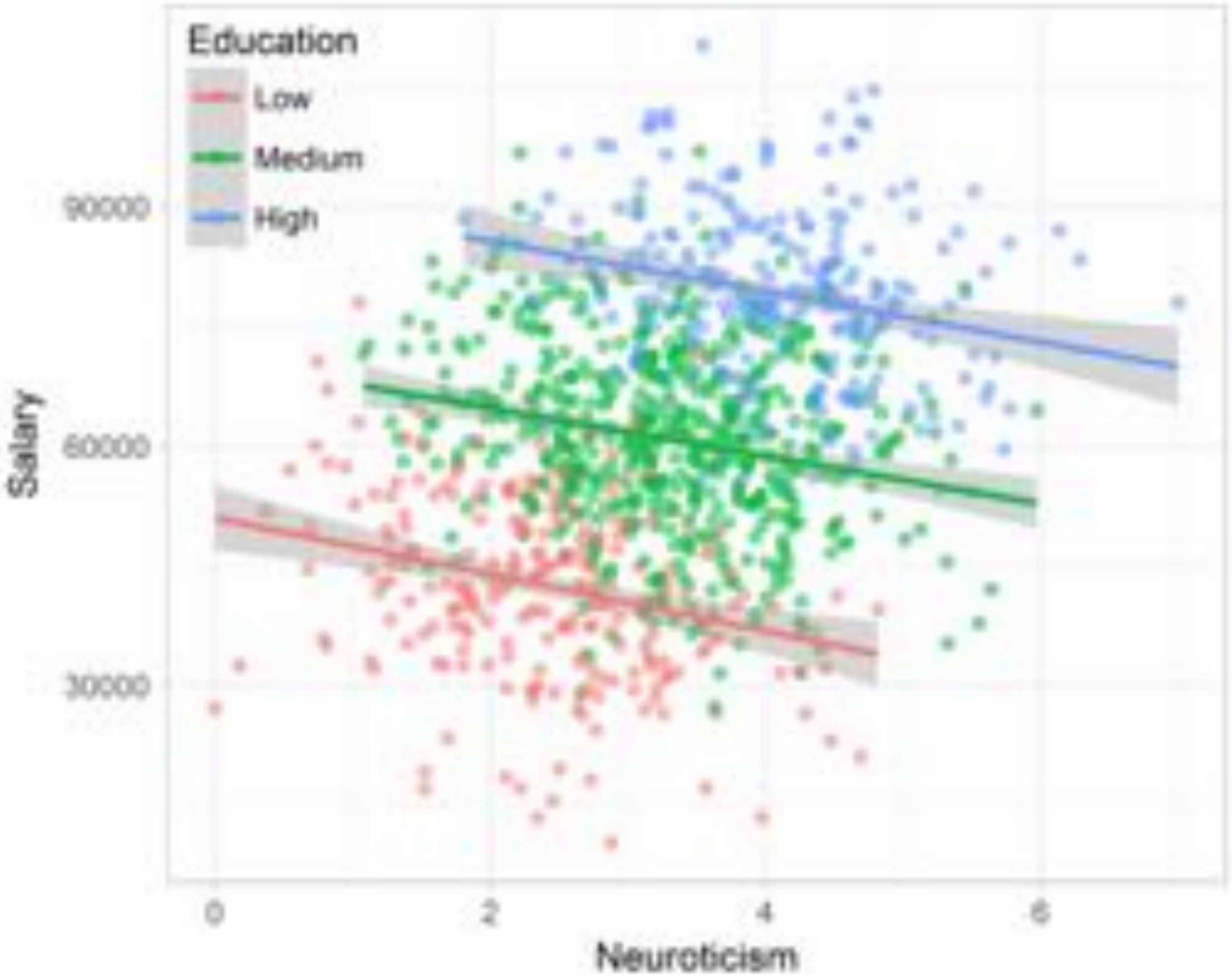
Feature Engineering



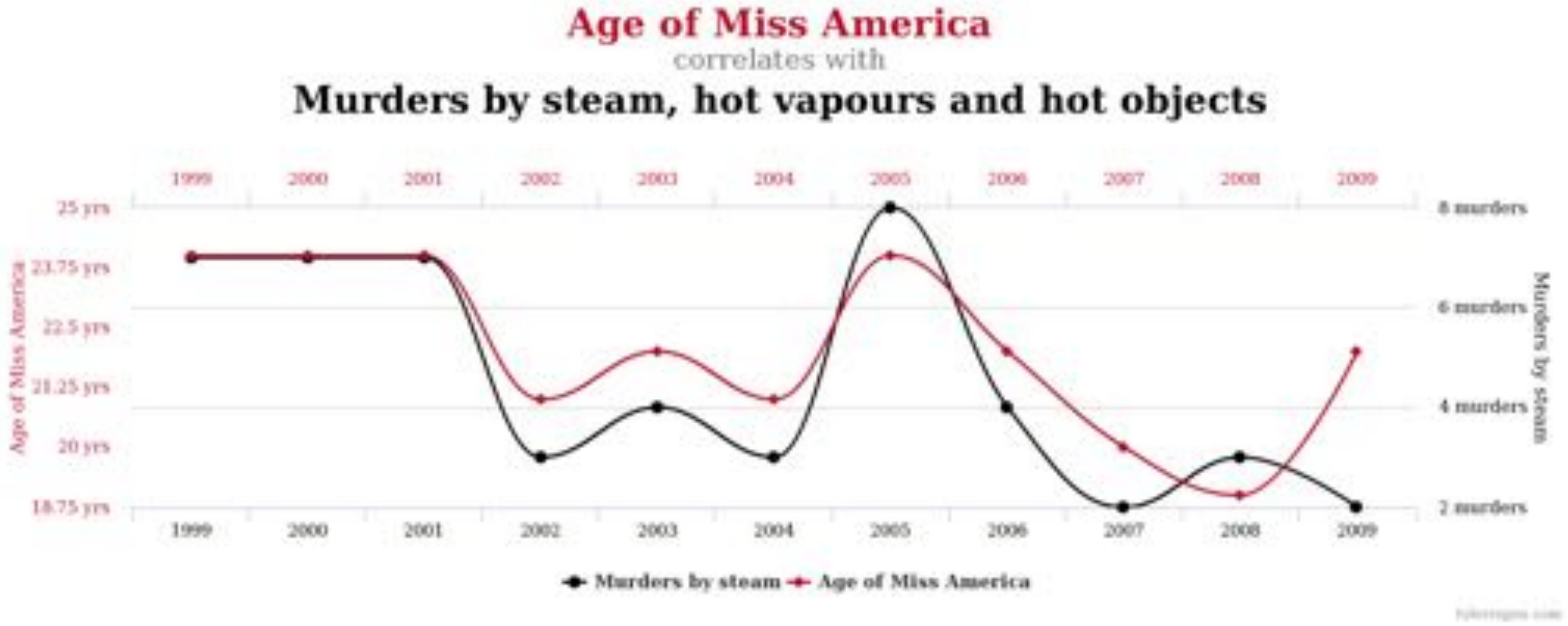
Model Selection



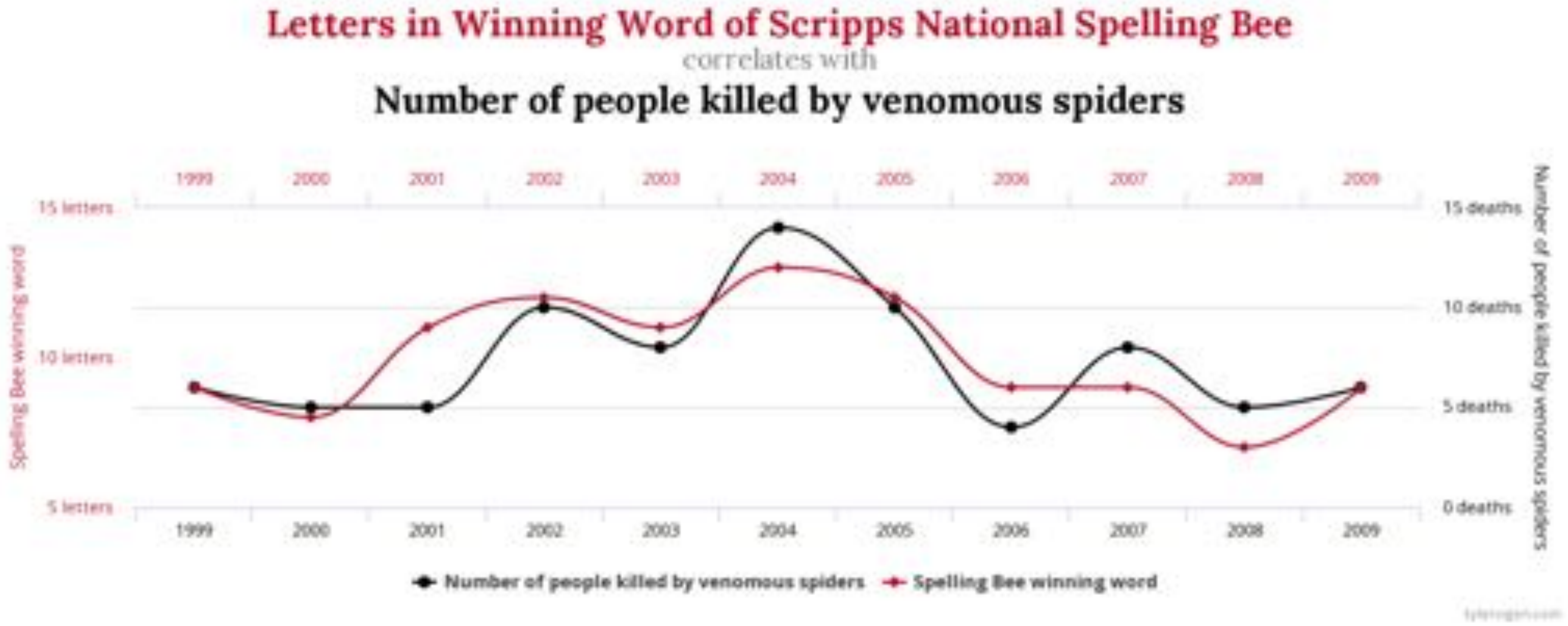




Model Evaluation



Model Evaluation



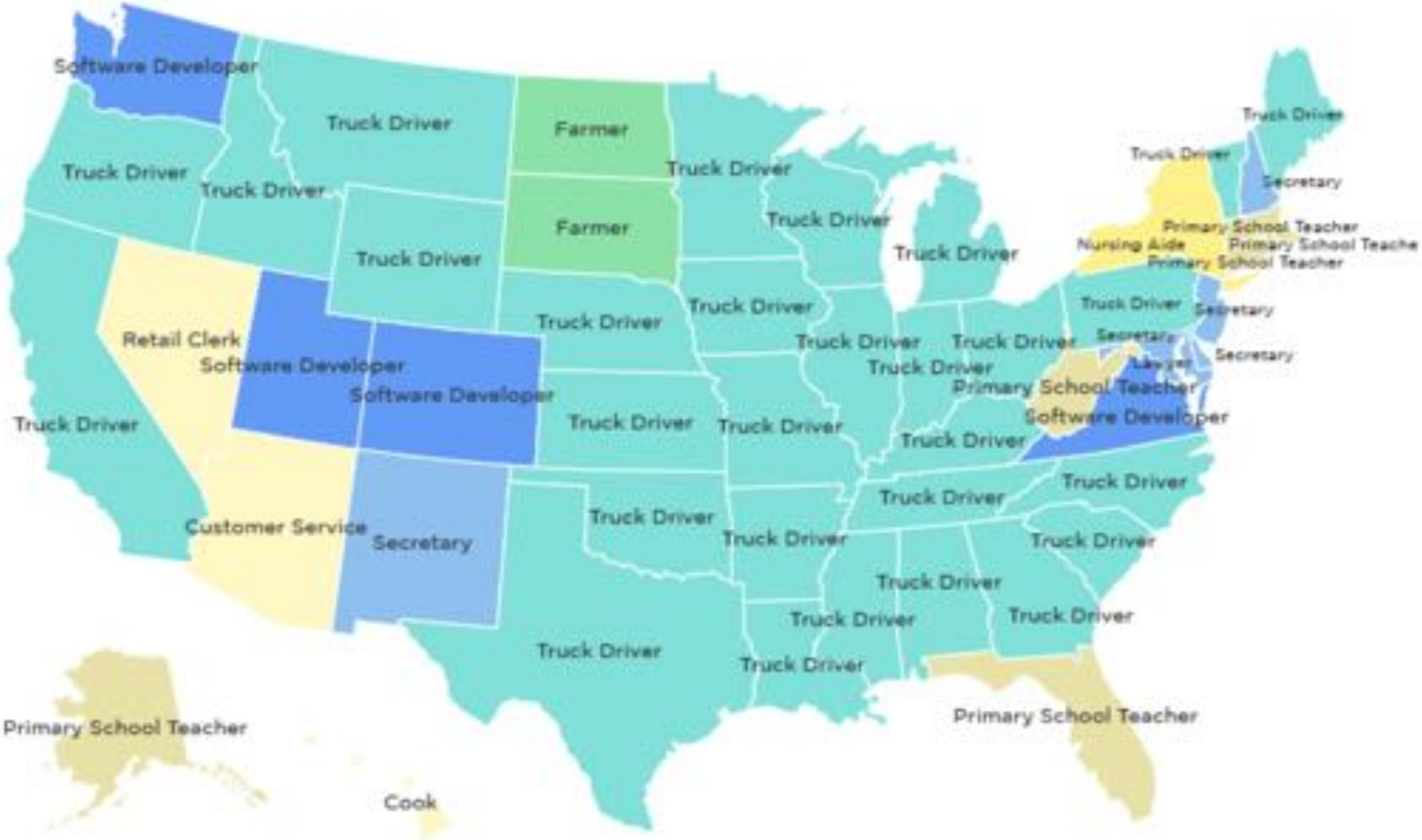
So what do we do about this?

Ethical Lessons

1. Know what the data are telling you
   - Understand your model, and validate it
   - Involve domain experts; “raw data” is an oxymoron
   - Communicate not just model *capabilities*, but also *limits* and *assumptions*
2. Data about people are about people
   - The data do not exist on their own, someone made them
3. *Always* consider why your model might be wrong in systematic ways
   - Don’t just **trust** the technology, make it **trustworthy**
   - **Close** the data science lifecycle: the world changes, and so should your models

Ethics beyond the Data Science Lifecycle

Should we deploy technology at all?



When is it ethical to collect data?


RYAN SINGEL SECURITY 02.17.12 03:02 PM

GOOGLE BUSTED WITH HAND IN SAFARI-BROWSER COOKIE JAR

When is it ethical to collect *sensitive* data?

- Without knowing sensitive attributes, it's not possible to know how to make good decisions.
  - E.g., in a credit-granting context, there's no reason to believe that qualified minority applicants will be similar to qualified members of the majority
- But having sensitive data means it can be repurposed for other uses

Inclusion and Representation

Resources

- Take a course at Berkeley on people and technology
- Visit https://fatml.org and https://fatconference.org
  - Reading list
  - Principles for Accountable Algorithms
- Books
  - *Weapons of Math Destruction,* by Cathy O'Neil
  - *How to Lie With Statistics*, by Darrell Huff
- Courses around the world: tinyurl.com/ethics-classes

"Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful."

-George E. P. Box