

# How Data Mining Discriminates

Solon Barocas

Princeton University



CENTER FOR  
INFORMATION  
TECHNOLOGY  
POLICY

[CITP.PRINCETON.EDU](http://CITP.PRINCETON.EDU)

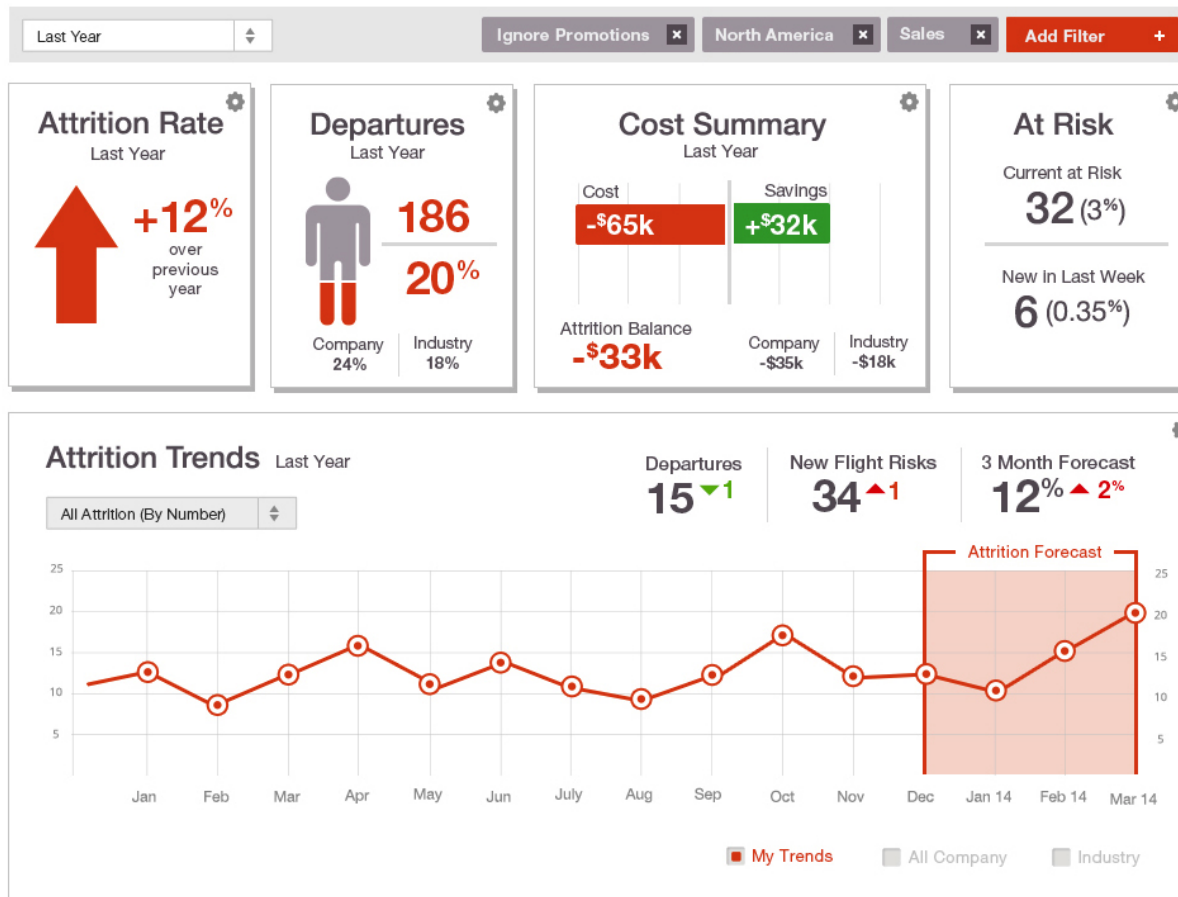
# Defining the Problem



john@acme.com | logout

Home Explore Track Manage Settings

## Attrition Management Console Detail



# Collecting and Labeling Examples

## A Skewed Set of Examples StreetBump



Crawford, K., 2013. The Hidden Biases in Big Data. *Harvard Business Review*.

## Setting a Bad Example Instant Checkmate

Ads by Google

[Latanya Sweeney, Arrested?](#)  
1) Enter Name and State. 2) Access Full Background Checks Instantly.  
[www.instantcheckmate.com/](http://www.instantcheckmate.com/)

[Latanya Sweeney](#)  
Public Records Found For: **Latanya Sweeney**. View Now.  
[www.publicrecords.com/](http://www.publicrecords.com/)

[La Tanya](#)  
Search for La Tanya Look Up Fast Results now!  
[www.ask.com/La+Tanya](http://www.ask.com/La+Tanya)

Sweeney, L., 2013. Discrimination in Online Ad Delivery. *Communications of the ACM*, 56(5), pp.44–54.

# Carving Up the Population

- Assessing people as individuals, not as members of a group or a type of person
  - At issue, really, is the coarseness and comprehensiveness of the criteria that permit distinction-making
- The uneven rates at which different groups happen to be subject to erroneous determinations
  - Should minority groups bear the disproportionate burden of faulty assessment, even if this means that the majority enjoys greater accuracy in decision-making?
- Errors can be harmful or beneficial...

Granularity of the Data

High

- Discovering attractive customers and candidates in populations previously dismissed out of hand → Financial inclusion
- Evidence-based and formalized decision-making

- Less favorable treatment in the marketplace → Finding specific customers not worth servicing (e.g., firing the customer)
- Individualization of risk

Low

- Equal treatment in the marketplace → Common level of service and uniform price
- Socialization of risk

- Underserving large swaths of the market → Redlining
- Informal decision heuristics plagued by prejudice and implicit bias

Benefit

Harm

Effects on historically disadvantaged communities

# Dealing with “Redundant Encodings”

- In many instances, making accurate determinations will mean considering factors that are somehow correlated with legally proscribed features
  - There is no obvious way to determine how correlated a relevant attribute or set of attributes must be with proscribed features to be worrisome
  - Nor is there a self-evident way to determine when an attribute or set of attributes is sufficiently relevant to justify its consideration, despite the fact that it is highly correlated with these features

# Masking

Race	<i>African-American</i>	<i>White American</i>
	I Support My President Fantasia Jill Scott Next Friday Erykah Badu Maxwell Taraji P Henson Madea Tyga Love And Basketball	Just Because You Can Reproduce Doesn't Mean You Should I Come From A Town Where A Traffic Jam Is 4 Cars Behind A Tractor Harley Davidson Halloween Bret Michaels David Bowie Official ASPCA Fly The American Flag Road Trips Bonfires