

The Human Side of Data (Visualization)

CSC444

Announcements

- TCEs! Please fill them out
- Last lecture! Tuesday will be course review and retrospective

What's the political goal of data visualization?

Truth?

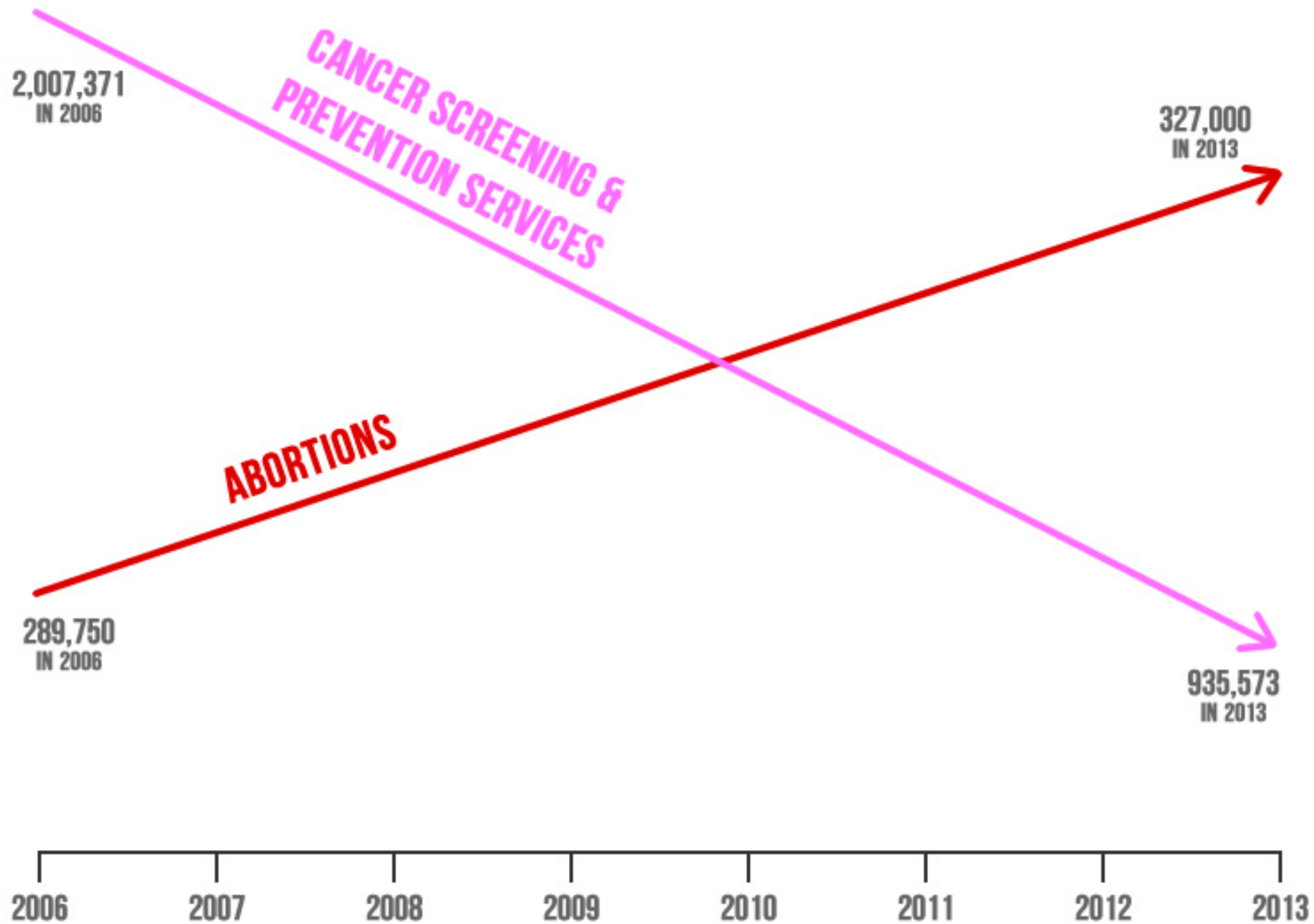
Objectivity?

Transparency?

Can you build a visualization without politics?

DECEPTION?

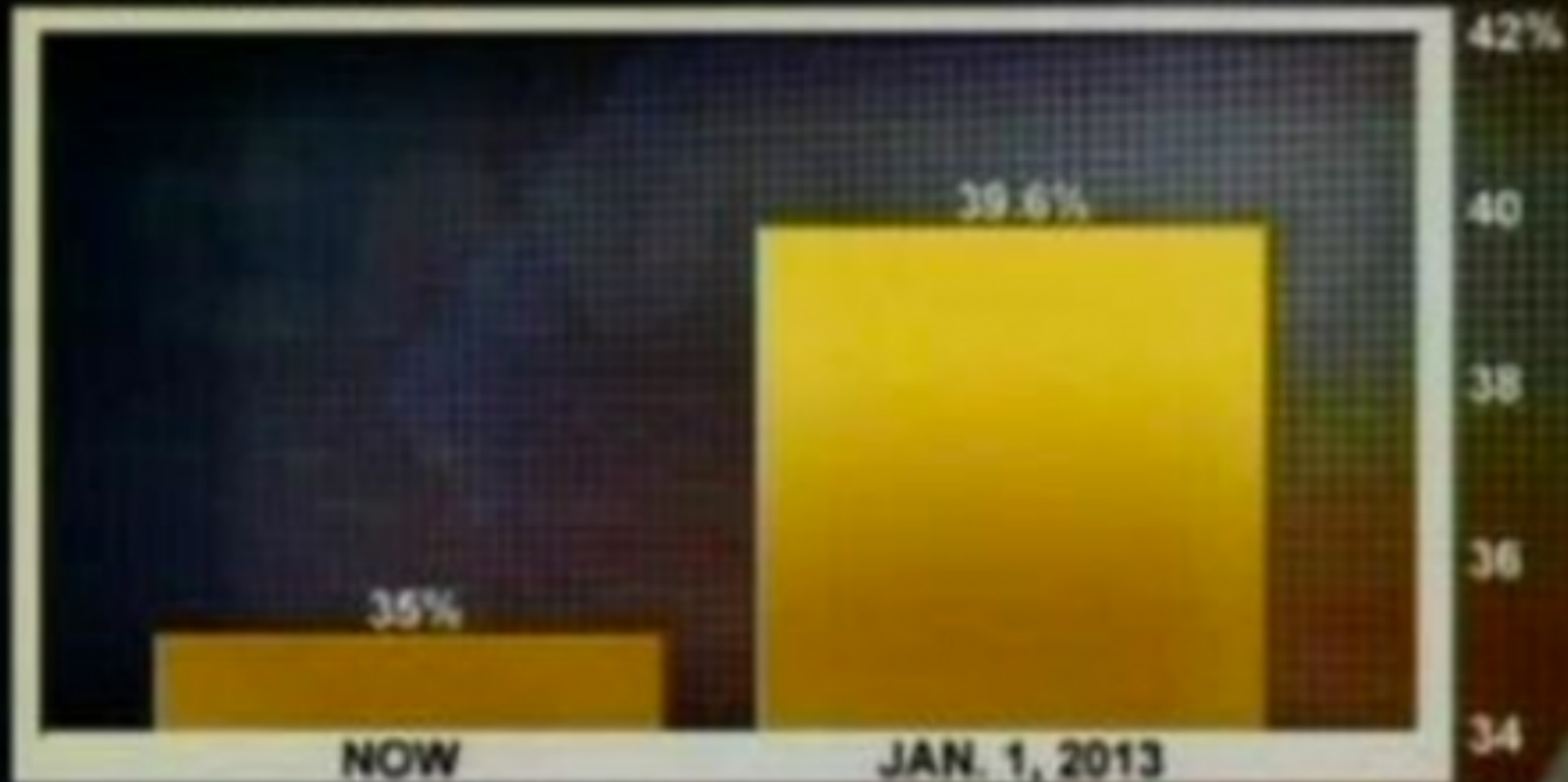
PLANNED PARENTHOOD FEDERATION OF AMERICA: ABORTIONS UP — LIFE-SAVING PROCEDURES DOWN



SOURCE: AMERICANS UNITED FOR LIFE

IF BUSH TAX CUTS EXPIRE

TOP TAX RATE



8:01p ET

FOX
EXTREMES

TOP STORIES

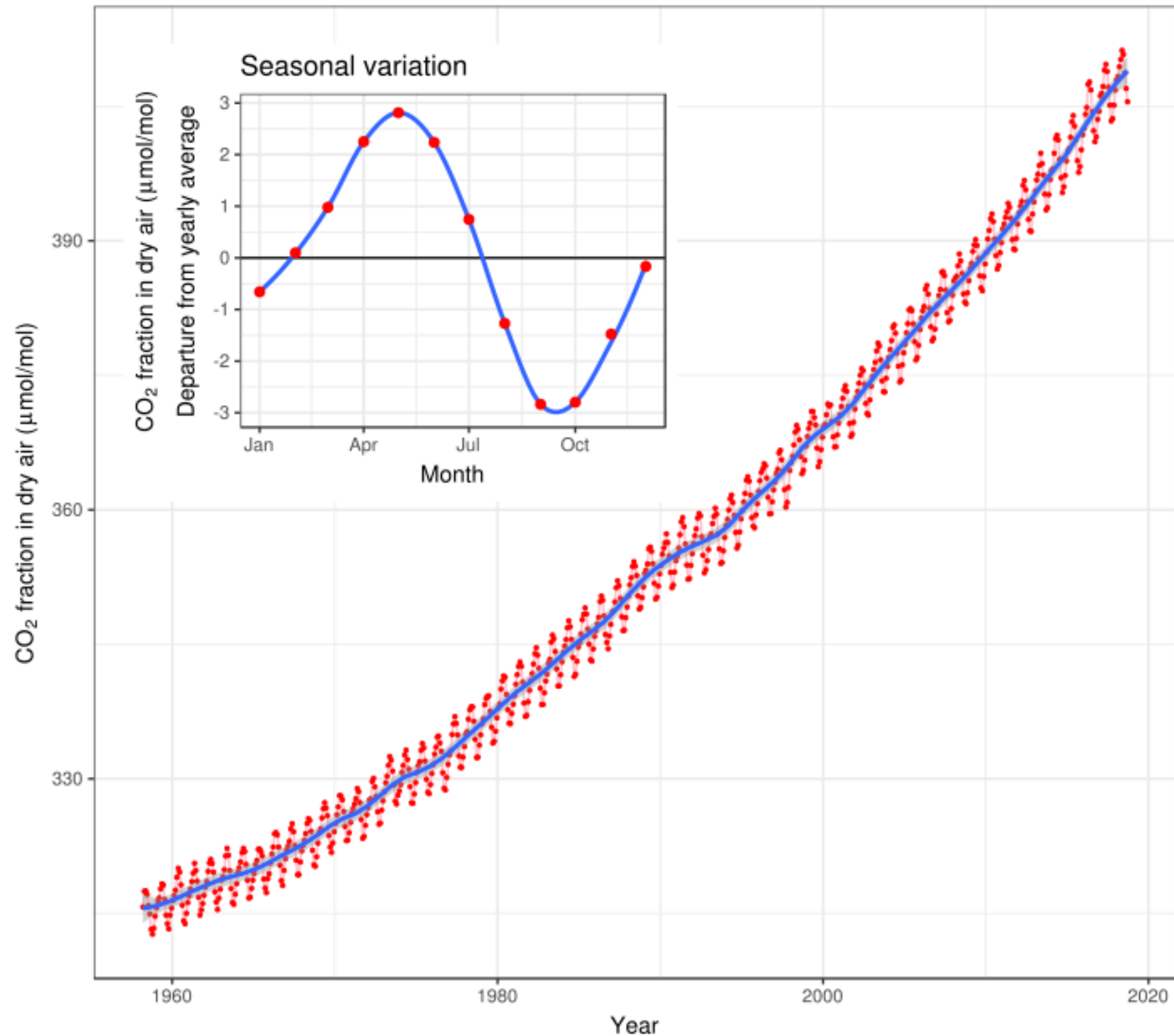
TECHNOLOGY

CONSUMER

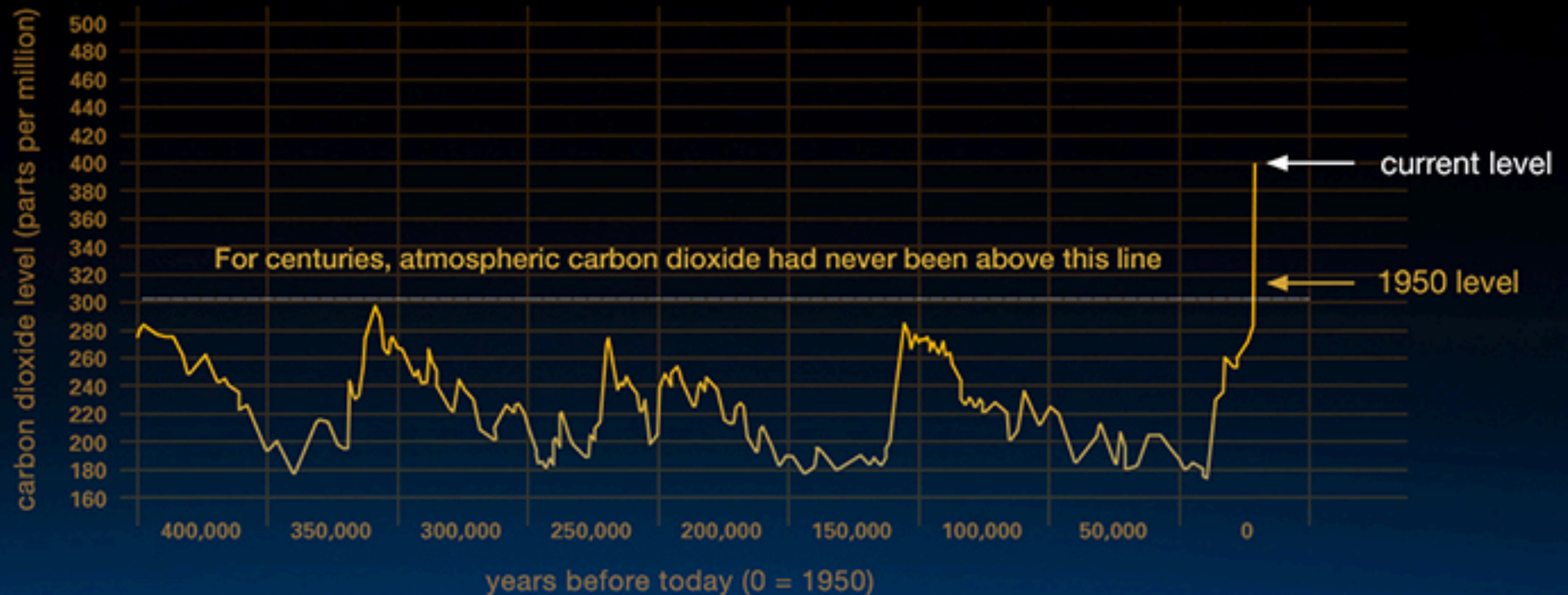
WITH THE JUSTICE DEPARTMENT AND ACQUIRES FULL T

Monthly mean CO₂ concentration

Mauna Loa 1958 - 2018



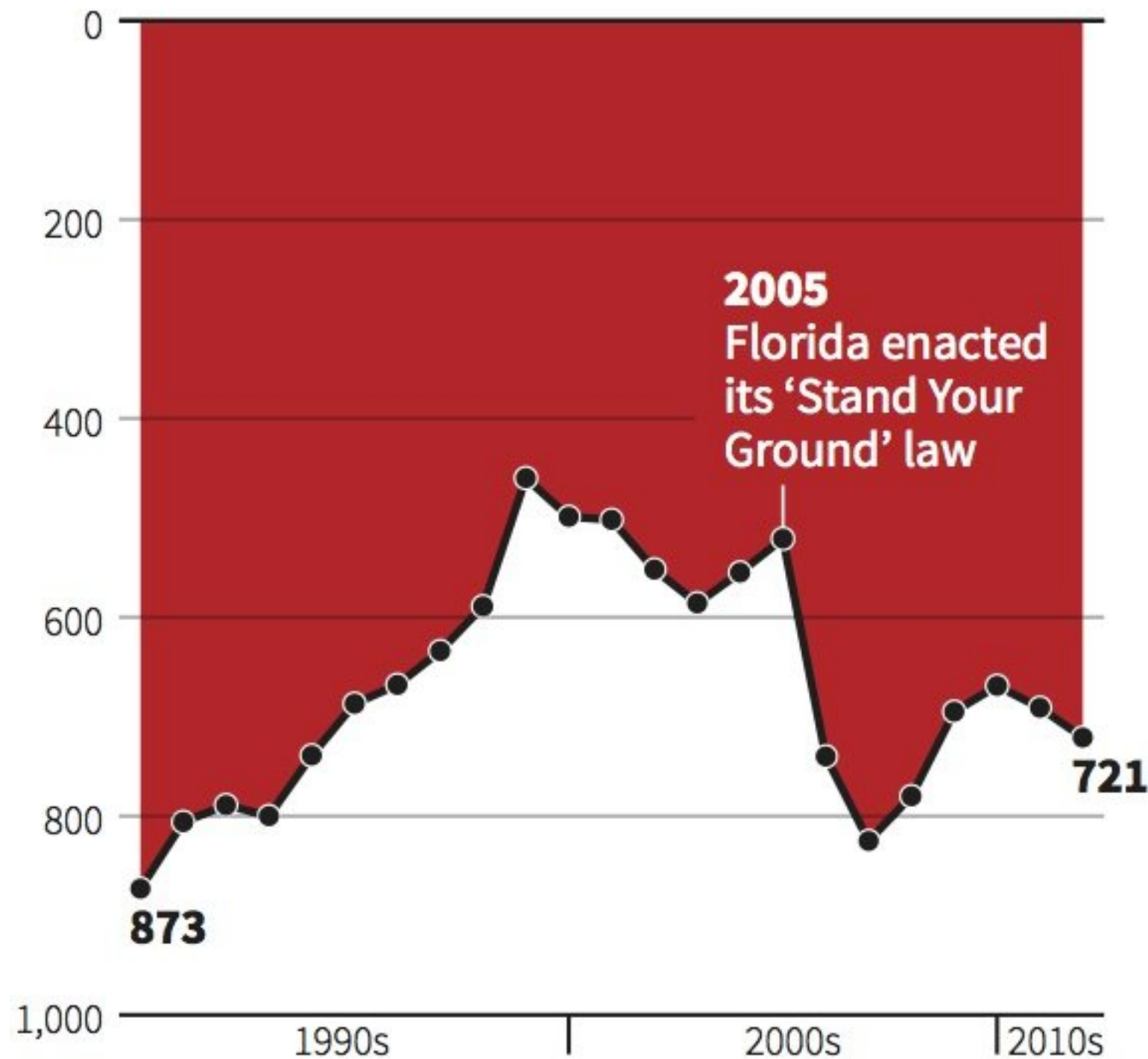
Data : Dr. Pieter Tans, NOAA/ESRL (www.esrl.noaa.gov/gmd/ccgg/trends/) and Dr. Ralph Keeling, Scripps Institution of Oceanography (scrippsco2.ucsd.edu/). Accessed 2018-10-21



https://climate.nasa.gov/climate_resources/24/graphic-the-relentless-rise-of-carbon-dioxide/

Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

How Deceptive are Deceptive Visualizations?: An Empirical Analysis of Common Distortion Techniques

Anshul Vikram Pandey
School of Engineering,
New York University
anshul.pandey@nyu.edu

Katharina Rall
School of Law,
New York University
kr1326@nyu.edu

Margaret L. Satterthwaite
School of Law,
New York University
satterth@exchange.law.nyu.edu

Oded Nov
School of Engineering,
New York University
onov@nyu.edu

Enrico Bertini
School of Engineering,
New York University
enrico.bertini@nyu.edu

ABSTRACT

In this paper, we present an empirical analysis of deceptive visualizations. We start with an in-depth analysis of what deception means in the context of data visualization, and categorize deceptive visualizations based on the type of deception they lead to. We identify popular distortion techniques and the type of visualizations those distortions can be applied to, and formalize why deception occurs with those distortions. We create four deceptive visualizations using the selected distortion techniques, and run a crowdsourced user study to identify the deceptiveness of those visualizations. We then present the findings of our study and show how deceptive each of these visual distortion techniques are, and for what kind of questions the misinterpretation occurs. We also analyze individual differences among participants and present the effect of some of those variables on participants' responses. This paper presents a first step in empirically studying deceptive visualizations, and will pave the way for more research in this direction.

journalism [14, 35], specialists and laypersons are using data to shape compelling, informative, and convincing narratives, conveyed through or supported by visualizations. While the use of such visual depictions as persuasion devices is not new, the popular use of visualizations has undoubtedly increased due in part to user-friendly software that allows non-experts to create visualizations. As such practices become more widespread and accessible, important new challenges and questions arise. If visualizations can make messages more accessible, comprehensible and persuasive [27, 37], visual representations can also be easily misused and misunderstood - even by their creators.

This problem has been known for a long time and it is not limited to visual representations but more to the general problem of communicating through numbers and statistics. Darrell Huff's "How to Lie with Statistics", published in 1954, popularized the problem and warned against the many traps of using statistics and charts in communication [15]. In the 1980s, Edward Tufte introduced the concept of *graphical in-*

TRANSPARENCY?

Search

Select a category and/or enter a text search and click "Search" to find what you're looking for.

Hint: If your search comes up with no results, try a broader search.

CATEGORY

TEXT SEARCH

Search

Educator salaries in New York

The database below shows the gross pay for roughly 250,000 educators at public schools in New York state during the 2015-16 school year, as reported by the New York State Teachers Retirement System. Their membership includes all full-time...

EDUCATION | Updated: November 16th 2016 | 10:15 PM

Unclaimed Funds

The New York state comptroller's office is overseeing \$14.5 billion dollars worth of unclaimed funds. Banks, insurance companies, and other institutions are required to turn over the contents of inactive or abandoned accounts. More than 35...

LOCAL BUSINESS | Updated: October 17th 2016 | 09:14 AM

Individual contributions to presidential candidates from New Yorkers

The table below shows individual contributions made by New York state residents to the Hillary Clinton and Donald Trump presidential campaigns. It contains information on more than 170,000 contributions made between January 1, 2015 and August...

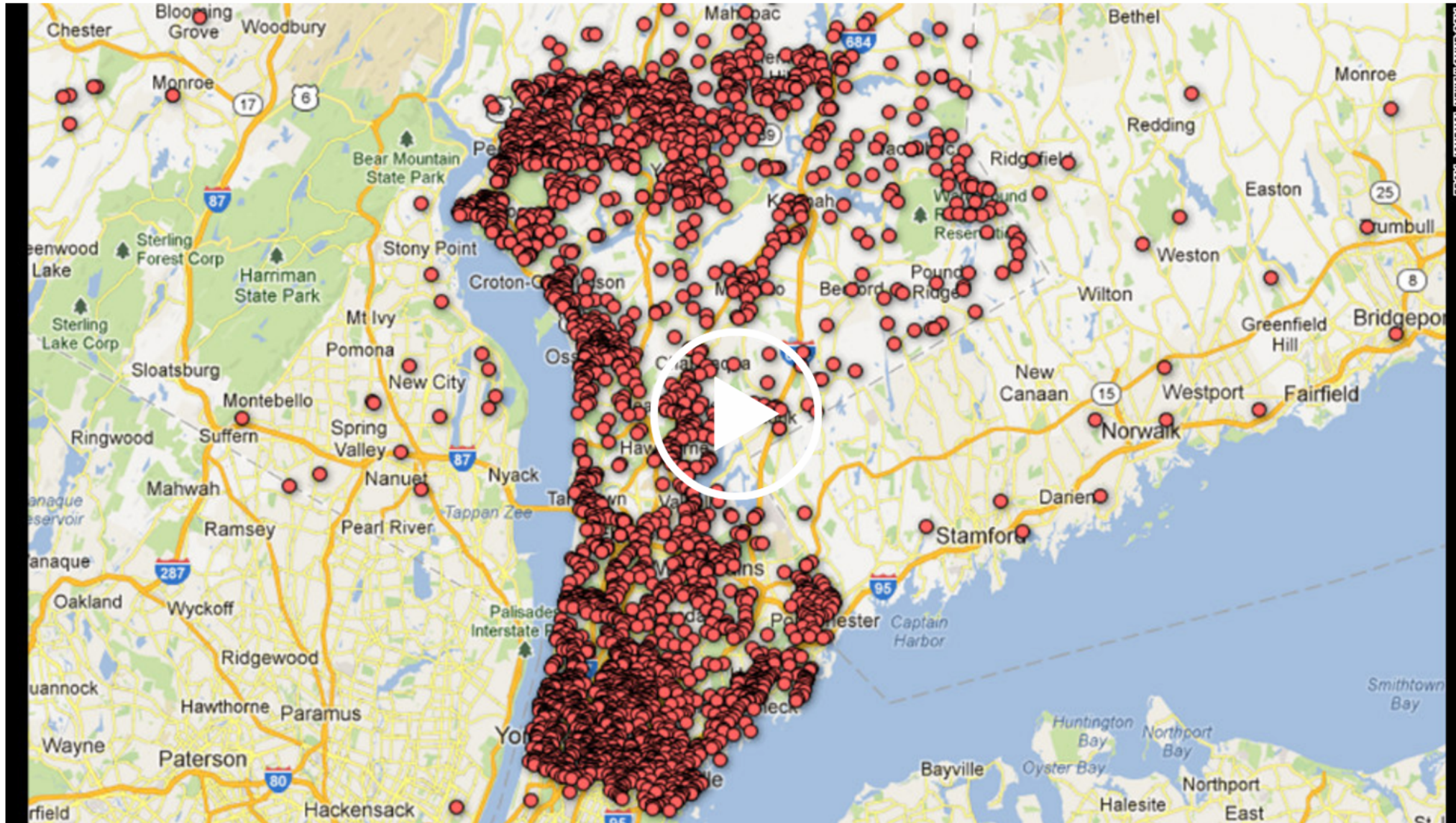
PUBLIC SPENDING | Updated: October 10th 2016 | 10:33 AM

<http://lohud.nydatabases.com/>

Newspaper sparks outrage for publishing names, addresses of gun permit holders

By KC Maas and Josh Levs, CNN

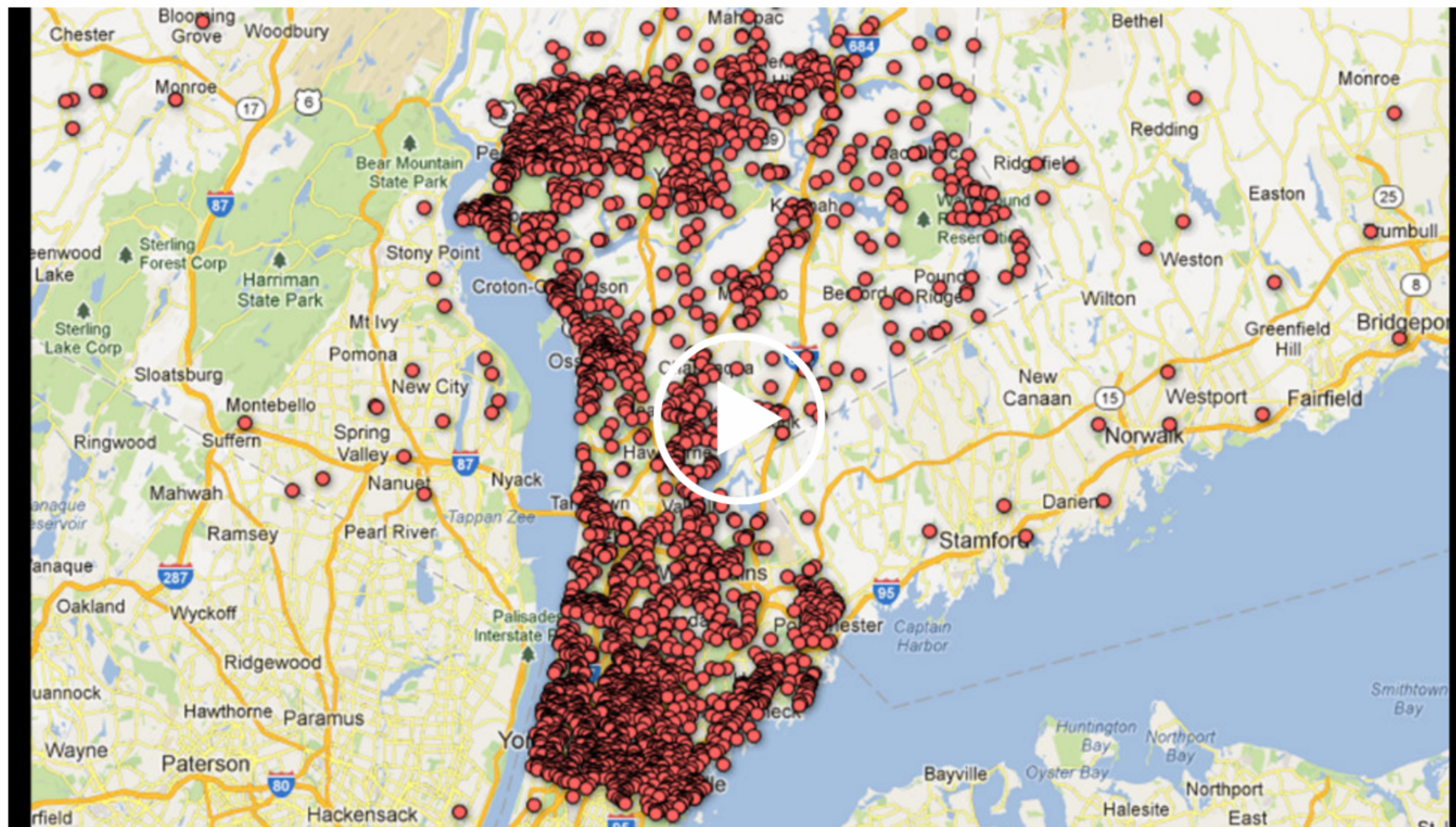
🕒 Updated 10:23 AM ET, Thu December 27, 2012



After newspaper's gun-permit map, blogger posts journalists' addresses

By Rande Jaconi and Laura Ly, CNN

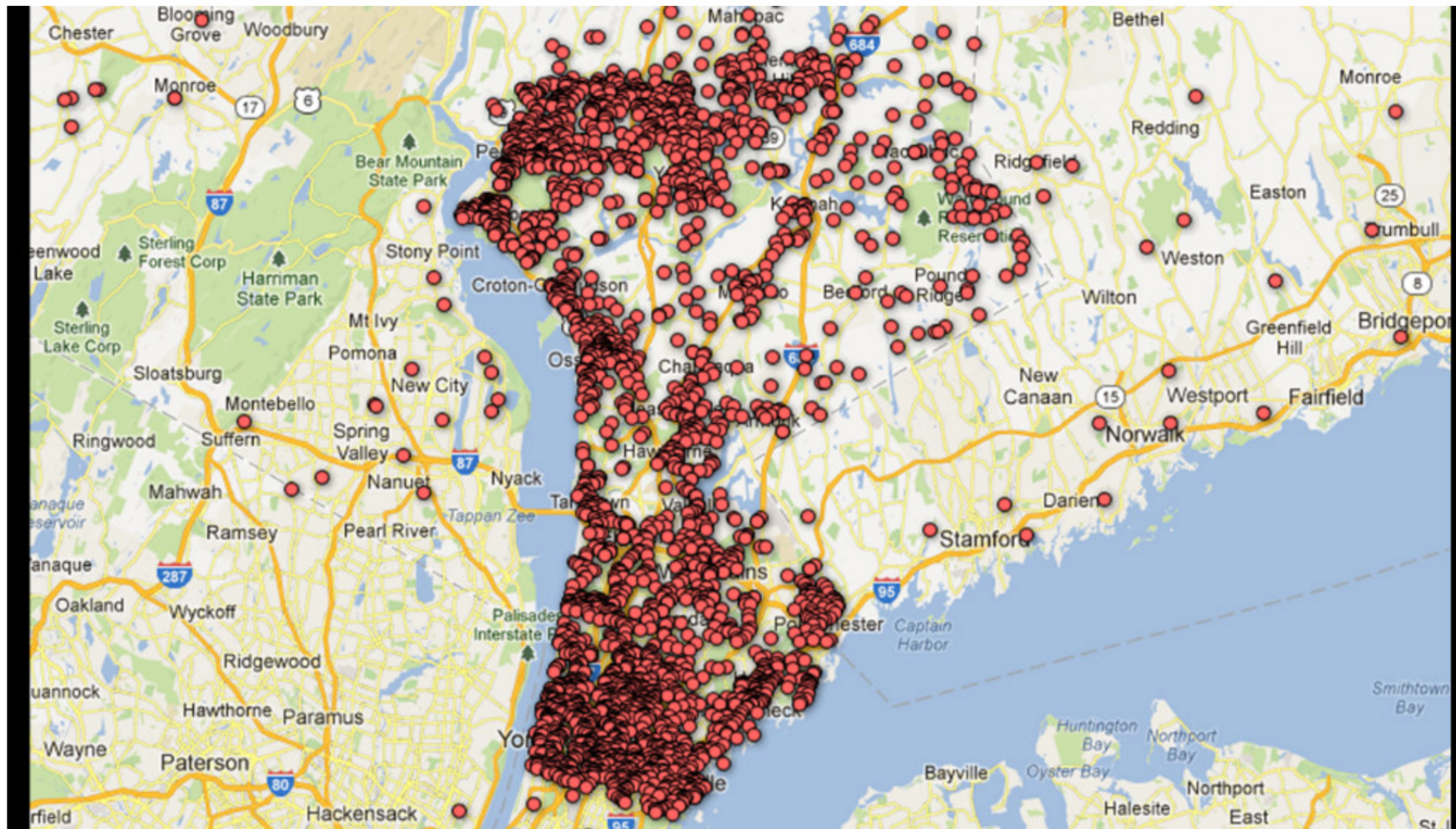
🕒 Updated 5:46 AM ET, Fri December 28, 2012



Newspaper removes controversial online database of gun permit holders

By Michael Martinez, CNN

🕒 Updated 1:50 PM ET, Sat January 19, 2013



I	J	K	L
College of Agric and Life Sci	Agricultural Resource Econ-Res	1.000	135,667.00
College of Medicine - Tucson	Medicine Administration	1.000	42,169.00
College of Agric and Life Sci	Sch of Family & Consum Sci-Ins	0.850	45,827.00
RDI Research Infrastructure	University Animal Care	1.000	28,080.00
College of Pharmacy	Pharmacy Practice and Science	1.000	48,495.00
Univ Info Tech Services Div	Research & Discovery Tech	1.000	167,000.00
Eller College of Management	Marketing	1.000	40,002.00
Enrollment Services	Office of the Registrar	1.000	32,502.00
College of Science	Astronomy	1.000	76,498.00
College of Medicine - Tucson	Medicine	1.000	49,909.00
College of Education	Disability Psychoeduc Studies	0.825	39,023.00
College of Pharmacy	Pharmacy Administration	0.490	16,307.00
College of Humanities	English as a Second Language	1.000	59,160.00
Libraries	UA Library Administration ADM	1.000	118,000.00
College of Medicine - Tucson	Dept of Emergency Medicine	0.250	8,320.00
College of Medicine - Tucson	Physiology	1.000	141,759.00
College of Social & Behav Sci	Social & Behavioral Sci Admin	1.000	39,520.00
College of Medicine - Tucson	Arthritis Center-Research	1.000	40,200.00
RDI Research Infrastructure	University Animal Care	1.000	28,080.00

https://docs.google.com/spreadsheets/d/1jFmxbDx6FP5qk5KKbFBJ5RvS0R2_HEoCkaw83P_DUG0/edit#gid=2006091355

OBJECTIVITY?

Not (only) a vis problem

Semantics derived automatically from language corpora necessarily contain human biases

Aylin Caliskan-Islam¹, Joanna J. Bryson^{1,2}, and Arvind Narayanan¹

¹Princeton University

²University of Bath

*Address correspondence to aylinc@princeton.edu, bryson@conjugateprior.org, arvindn@cs.princeton.edu.

[†]Draft date August 31, 2016.

ABSTRACT

Artificial intelligence and machine learning are in a period of astounding growth. However, there are concerns that these technologies may be used, either with or without intention, to perpetuate the prejudice and unfairness that unfortunately characterizes many human institutions. Here we show for the first time that human-like semantic biases result from the application of standard machine learning to ordinary language—the same sort of language humans are exposed to every day. We replicate a spectrum of standard human biases as exposed by the Implicit Association Test and other well-known psychological studies. We replicate these using a widely used, purely statistical machine-learning model—namely, the GloVe word embedding—trained on a corpus of text from the Web. Our results indicate that language itself contains recoverable and accurate imprints of our historic biases, whether these are morally neutral as towards insects or flowers, problematic as towards



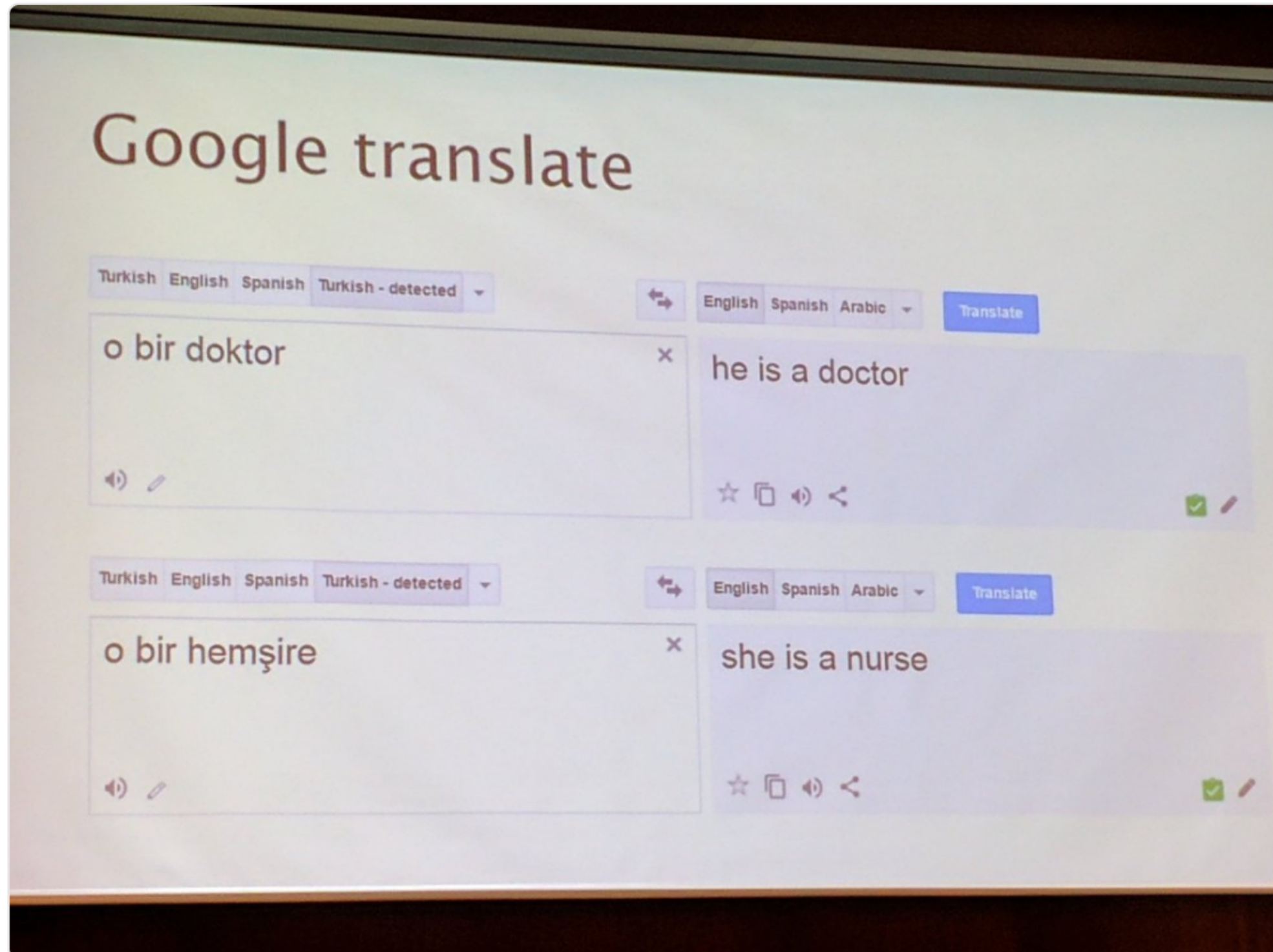
Andreas Mueller

@amuellerm1



Follow

Gender neutral pronouns get translated to gender stereotypes by Google translate #FATML



<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>

SAN FRANCISCO (Reuters) - Amazon.com Inc's ([AMZN.O](#)) machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.

. . .

In effect, Amazon's system taught itself that male candidates were preferable. It penalized resumes that included the word "women's," as in "women's chess club captain." And it downgraded graduates of two all-women's colleges, according to people familiar with the matter. They did not specify the names of the schools.

<https://www.cnbc.com/2018/05/30/silicon-valley-is-stumped-even-a-i-cannot-remove-bias-from-hiring.html>

HireVue - Hiring Intelligence | Assessment & Video Interview Software

<https://www.hirevue.com/> ▼

HireVue transforms the hiring & recruiting process by providing enterprises with an all-in-one video interview & pre-hire assessment solution. Hirevue's online ...

Unilever has improved the diversity of its talent pool by 10 percent since partnering with HireVue. "If the team does notice a skew in results, it can evaluate the algorithm to see what went wrong and remove the bad data," Zuloaga said.

"AI is not impartial or neutral," said Meredith Whittaker, co-founder of the AI Now Institute at New York University, and founder of Google's Open Research group. AI Now — which Whittaker co-founded with Kate Crawford, an NYU professor and principal researcher at Microsoft Research — aims to move beyond what it describe as "minimal oversight" of AI. Algorithmic bias is one of its core research areas.

IT'S ABOUT PEOPLE

Human Subjects

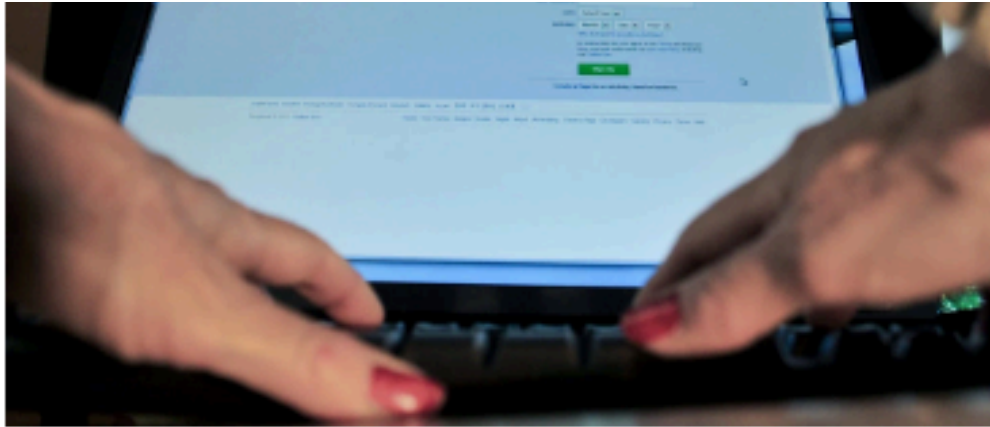
- If you work in the field, you will, at some point, be asked to collect data about people, and create a visualization of it
- What now?

INFORMED CONSENT

The Institutional Review Board (IRB)

- If you receive federal funds to to human-subjects research, you need approval from your institution's IRB
- But why?
 - History lesson: the horrendous Tuskegee experiment

<https://www.nytimes.com/2014/06/30/technology/facebook-tinkers-with-users-emotions-in-news-feed-experiment-stirring-outcry.html>



Facebook revealed that it had altered the news feeds of over half a million users in its study.
Karen Bleier/Agence France-Presse — Getty Images

To Facebook, we are all lab rats.

Facebook routinely adjusts its users' news feeds — testing out the number of ads they see or the size of photos that appear — often without their knowledge. It is all for the purpose, the company says, of creating a more alluring and useful product.

But last week, Facebook revealed that it had manipulated the news feeds of over half a million randomly selected users to change the number of positive and negative posts they saw. It was part of a psychological study to examine how emotions can be spread on social media.



Bill Sourour [Follow](#)

DevMastery.com | Consultant | Teacher

Nov 13 · 5 min read

The code I'm still ashamed of



If you write code for a living, there's a chance that at some point in your career, someone will ask you to code something a little deceitful – if not outright unethical.

Pearson Embedded a 'Social-Psychological' Experiment in Students' Educational Software



Sidney Fussell

Yesterday 5:50pm • Filed to: AI/ETHICS



52



3

The screenshot shows a web application interface with a sidebar and a main content area.

Table of Contents (Left Sidebar):

- prev Search the TOC... next
- Chapter 10 File IO
- Chapter 11 Recursion
- Chapter 12 [click for information]
- Chapter 13 Interfaces and Inner Classes
- Chapter 14 Generics and the ArrayList C
- Chapter 15 [click for information]
- Chapter 16 [click for information]
- Chapter 17 Swing I
- Chapter 18 Swing II
- Chapter 19 [click for information]
- Chapter 20 Applets and HTML (online at
- My assignment
- 70904
- Chapter 3 Flow of Control
 - 3.1 BRANCHING MECHANISM
 - 3.2 BOOLEAN EXPRESSIONS
 - 3.3 LOOPS
 - The while Statement
 - The do-while Statement
 - The for Statement
 - 20674
 - 20675
 - 20676
 - 20677
 - 20993**
 - 20994
 - 20995
 - 20996
 - 20997
 - 20998
 - 20688
 - 20686

Workbench (Main Content Area):

Exercise 20993

WORK AREA RESULTS SOLUTIONS ROSTER NEW

Print Results

CODELAB ANALYSIS: COMPILER ERROR(S)

More Hints:

- ⇒ You almost certainly should be using: **for**
- ⇒ You almost certainly should be using: **int**

YOUR SUBMISSION

Possible locations of errors or opportunities for improvement have been marked. [\(Details\)](#)

```
for( i =0; i<40; i++)  
{  
    System.out.print(i + " ")  
}
```

COMPILER ERROR MESSAGES

CTest.java:17: error: **'}' expected**

```
1 import java.util.Scanner;  
2  
3 class CTest {  
4     public static void main(String [] args) {  
5         Scanner stdin = new Scanner(System.in);  
6  
7  
8 For(i =0; i<40; i++)
```


What's our (your!)
responsibility on this?

The ACM Code of Ethics

- <https://ethics.acm.org/>
- (It's 2018, and it has existed for <6 months. That's telling.)