

The Human Side of Data (Visualization)

CSC544

Announcements

- **TCEs! Please fill them out**
- **All following lectures - from now on, final project hands-on**
 - **Come prepared to give brief presentations (think no more than 1-2 minutes!)**
 - **I'll come prepared for impromptu Q&A sessions.**

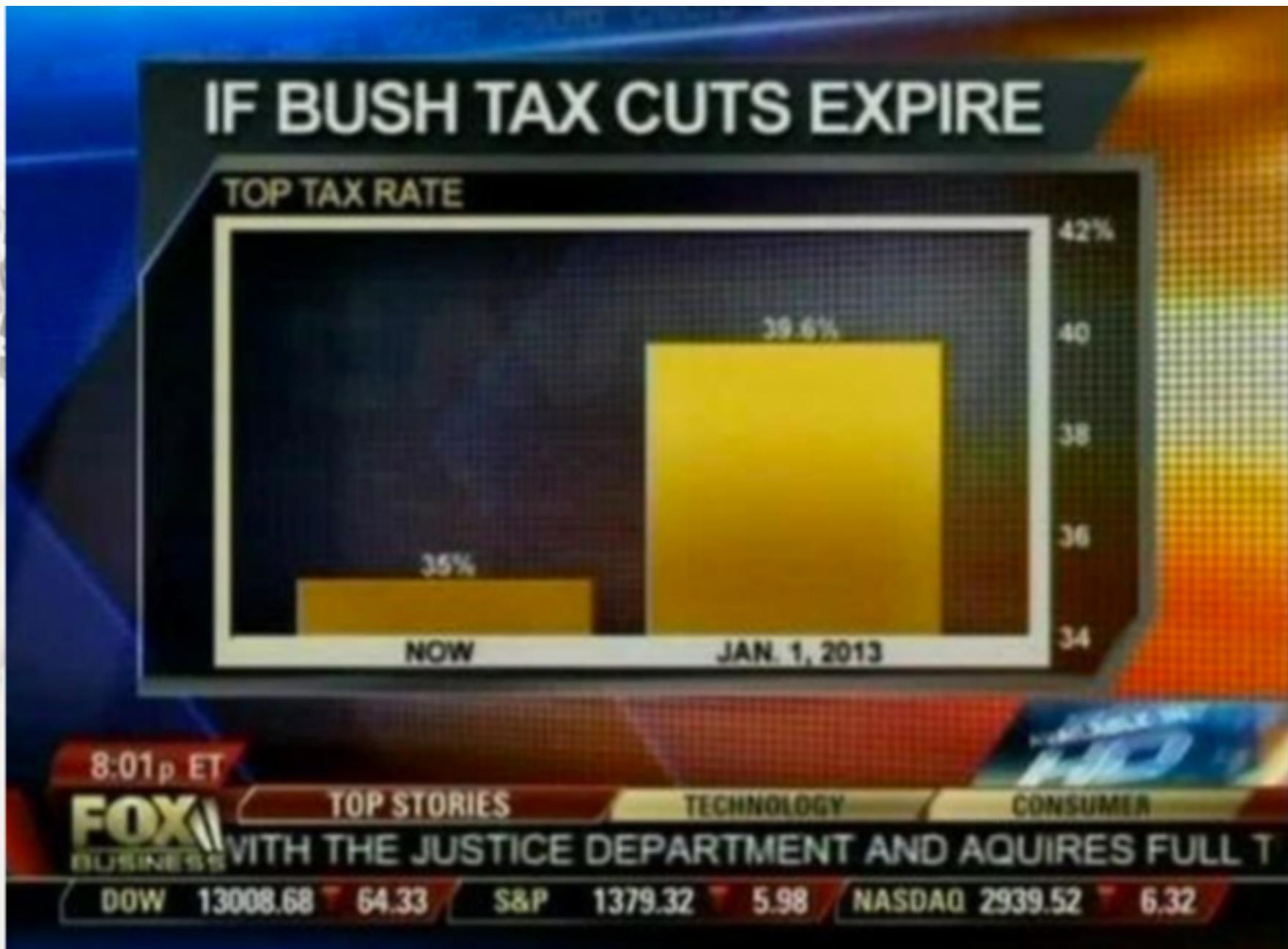
What's the political goal of data visualization?

Truth?

Objectivity?

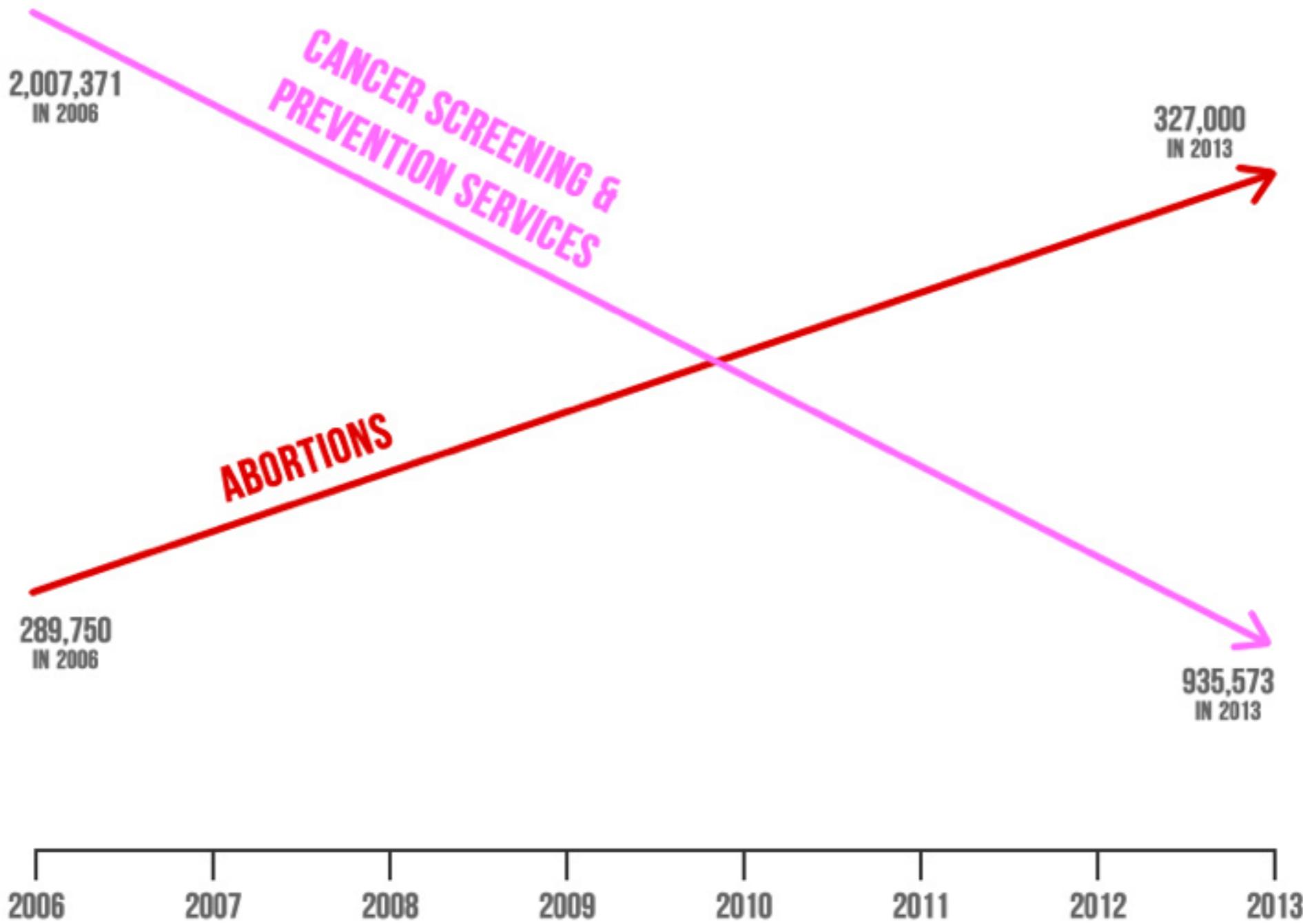
Transparency?

Can you build a visualization without politics?



<https://prezi.com/8dlwl4615t6b/ethics-of-data-visualization/>

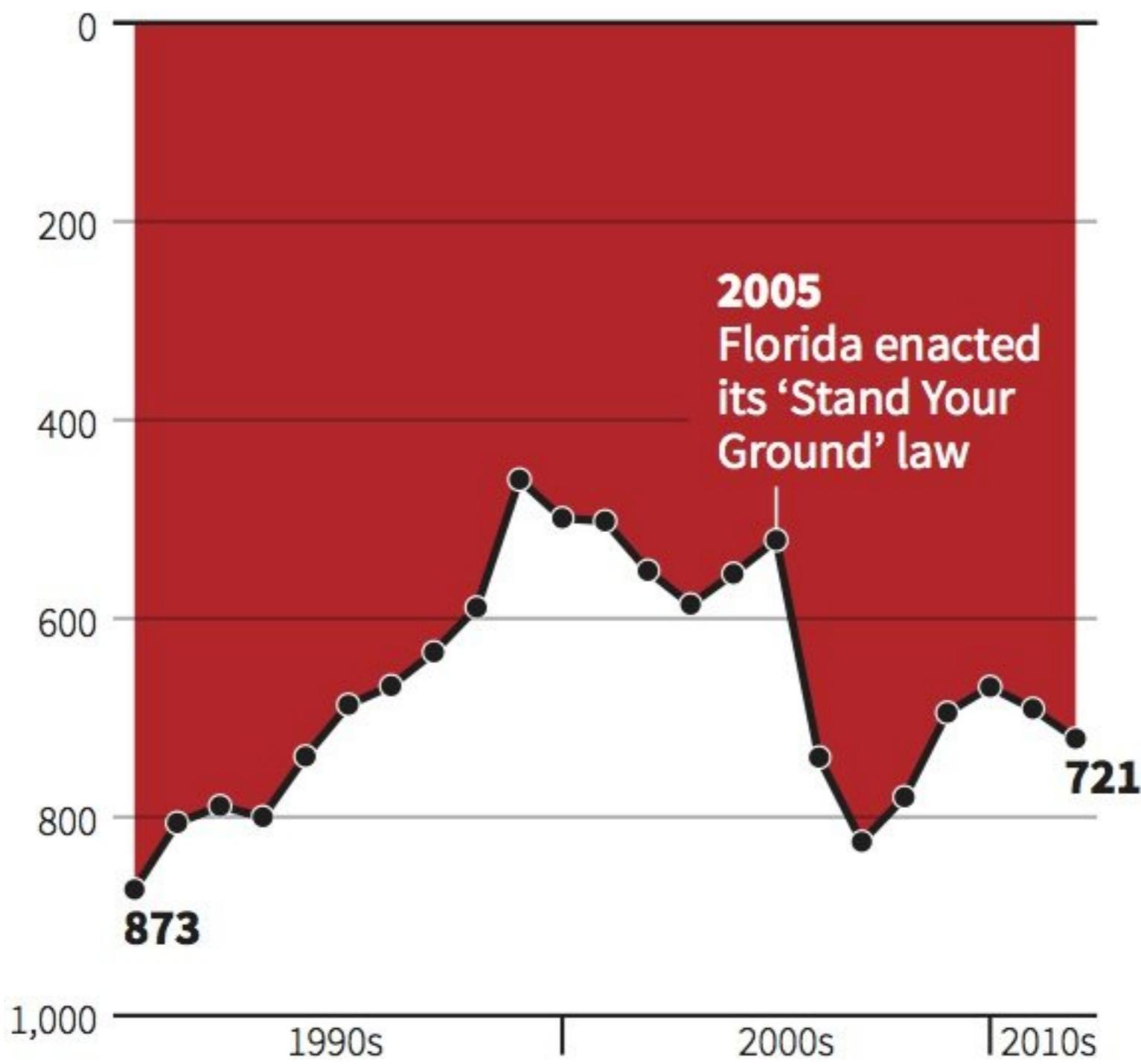
PLANNED PARENTHOOD FEDERATION OF AMERICA: ABORTIONS UP – LIFE-SAVING PROCEDURES DOWN



SOURCE: AMERICANS UNITED FOR LIFE

Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement



George Joseph

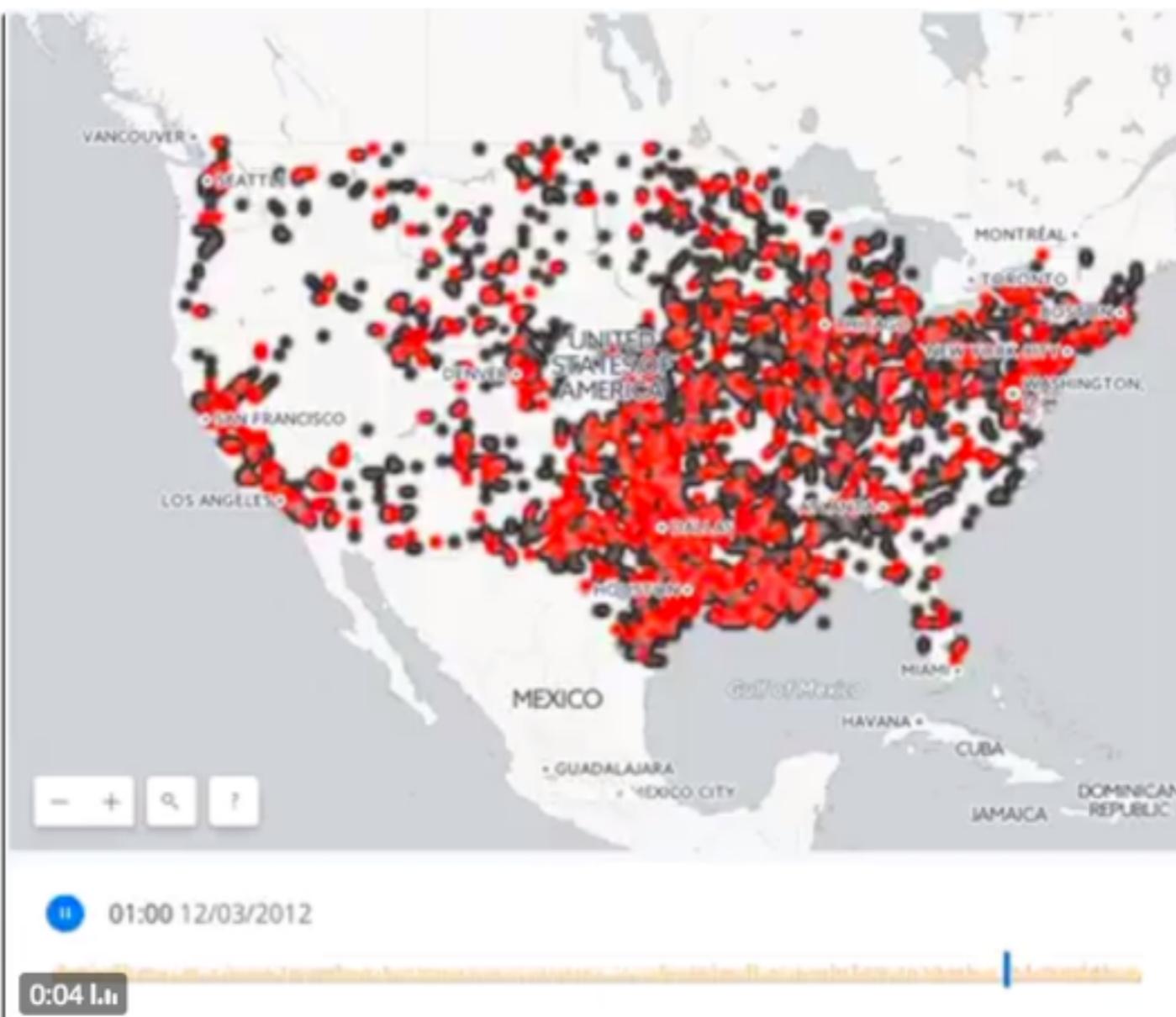
@georgejoseph94



Follow

Mapped the last 20 years of big pipeline related accidents. Red dots indicate those that led to fatalities. [#NoDAPL](#)

citylab.com/weather/2016/1/ ...



<https://twitter.com/georgejoseph94/status/804027897275154436>

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	
<input type="text" value="- Any -"/>	<input type="text"/>	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

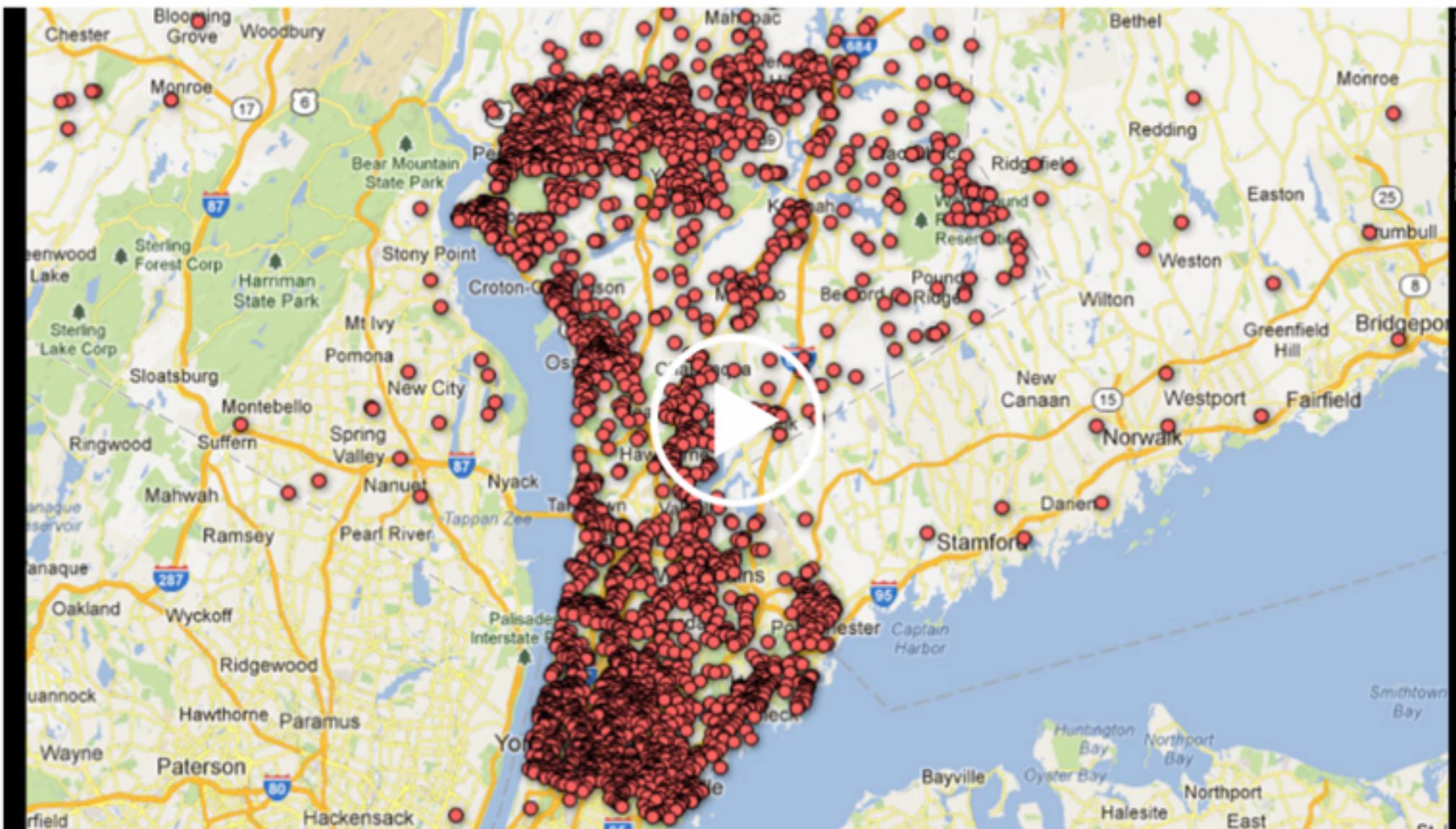
The table below shows individual contributions made by New York state residents to the Hillary Clinton and Donald Trump presidential campaigns. It contains information on more than 170,000 contributions made between January 1, 2015 and August...

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

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 Iaboni and Laura Ly, CNN

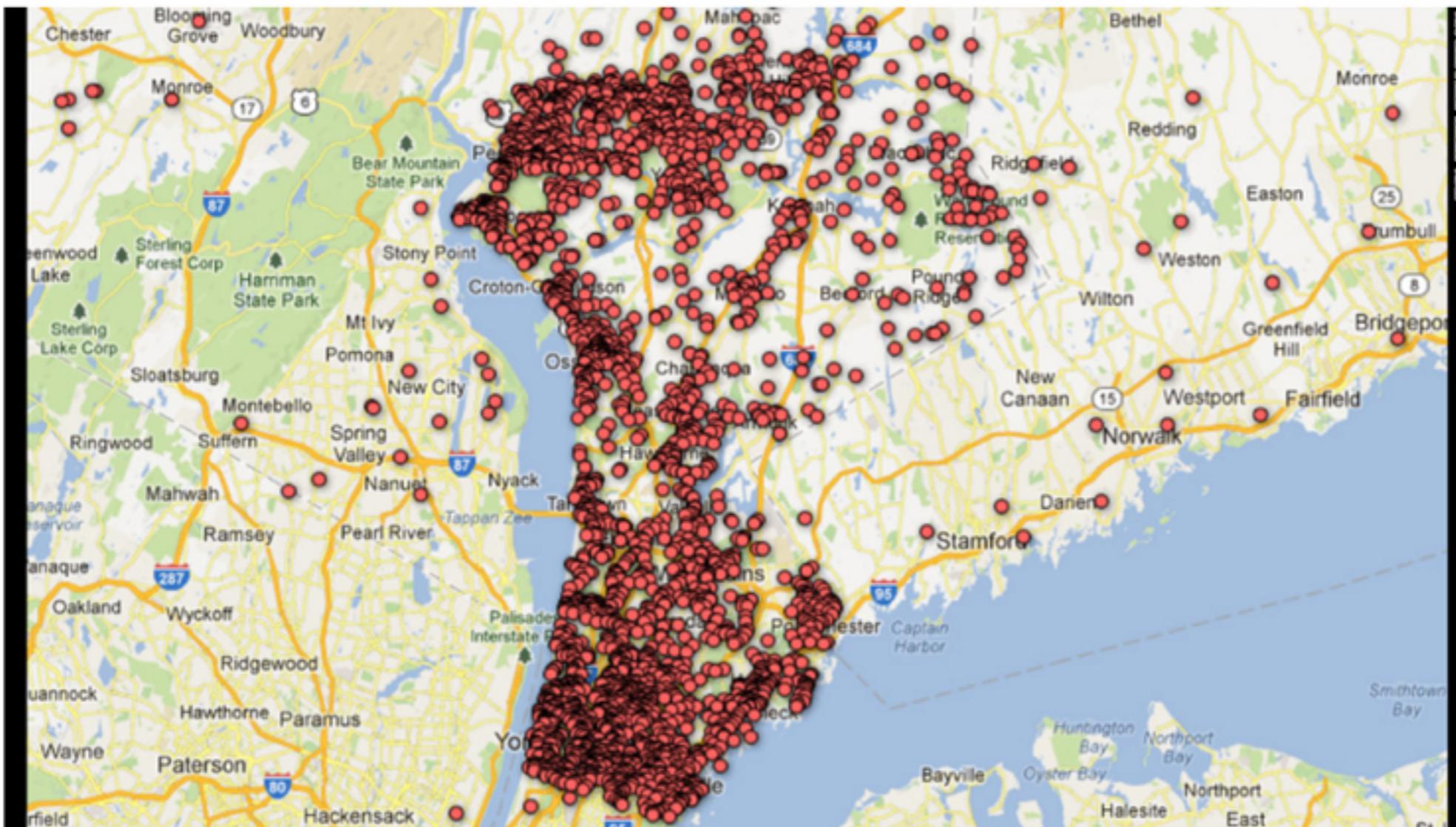
Updated 5:46 AM ET, Fri December 28, 2012



Newspaper removes controversial online database of gun permit holders

By Michael Martinez, CNN

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



The Persuasive Power of Data Visualization

Anshul Vikram Pandey, Anjali Manivannan,
Oded Nov, Margaret Satterthwaite, and Enrico Bertini, *Member, IEEE*

Abstract—Data visualization has been used extensively to inform users. However, little research has been done to examine the effects of data visualization in influencing users or in making a message more persuasive. In this study, we present experimental research to fill this gap and present an evidence-based analysis of persuasive visualization. We built on persuasion research from psychology and user interfaces literature in order to explore the persuasive effects of visualization. In this experimental study we define the circumstances under which data visualization can make a message more persuasive, propose hypotheses, and perform quantitative and qualitative analyses on studies conducted to test these hypotheses. We compare visual treatments with data presented through barcharts and linecharts on the one hand, treatments with data presented through tables on the other, and then evaluate their persuasiveness. The findings represent a first step in exploring the effectiveness of persuasive visualization.

Index Terms—Persuasive visualization, elaboration likelihood model, evaluation

1 INTRODUCTION

Recent years have witnessed a remarkable increase in the adoption of visualization as a means to convey messages through data. Popular and respected venues such as *The New York Times* and *The Guardian* have already popularized the idea of using data visualization to convey a powerful message, and an increasing number of scientists, journalists, activists, and businesses are following a similar path.

However, as visualization as a communication tool gains popularity, it is necessary to better understand how it impacts and influences people. Does graphical presentation of data make a message more persuasive? This and other similar questions are crucial as data visualization is often employed with the implicit or explicit assumption that graphical representation has a powerful persuasive effect on a target population.

Human rights advocates and activists are increasingly tapping into the power of data visualization as a way to influence and persuade their audience. Tactical Technology Collective, an organization that helps campaigners and activists use technology in their work, has extensively used and researched visualization as an instrument of influence in activism. Their recently published book "Visualizing Information for Advocacy" discusses and proposes strategies to transform mes-

We also deem important to clarify that by studying the effect of visualization on persuasion we by no means intend to imply, through our studies and results, that it is always ethically acceptable or even desirable to manipulate people's opinion through visualization. Rather, we believe further advancements in this area need to go hand-to-hand with further research on visualization misuse and its well-known powerful effects in terms of misrepresentation, disinformation, and even deception [23, 15, 18]. Research like that carried out in this study can help us shed light on how persuasion through visualization works and hopefully prevent or mitigate malevolent use.

In our work, we proceed from the basic research question: "*Does graphical depiction of data have a more persuasive effect than textual or tabular information?*". More precisely, following the tradition of the Elaboration Likelihood Model (ELM) of persuasion (explained in more detail in Section 2.1), we sought effects of visualization on *attitude change*, that is, the change of attitude on a proposed topic before and after being shown a persuasive message.

We ran several crowd-sourced randomized controlled experiments

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-*

Is “Data” Objective?

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

@amuellerml



Follow

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

The screenshot shows two separate instances of the Google Translate interface. In both cases, the source language is set to Turkish and the target language is English.

Top Translation:

- Source: "o bir doktor" (Turkish)
- Target: "he is a doctor" (English)

Bottom Translation:

- Source: "o bir hemşire" (Turkish)
- Target: "she is a nurse" (English)

In both examples, the gender neutral pronoun "o" is translated as a gendered pronoun ("he" or "she"), demonstrating a clear bias in the machine translation system.

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



In preparing this talk I decided to check out the data landscape, since I hadn't seen it for a while.

The terminology around Big Data is surprisingly bucolic. Data flows through streams into the data lake, or else it's captured in logs. A data silo stands down by the old data warehouse, where granddaddy used to sling bits.

And high above it all floats the Cloud. Then this stuff presumably flows into the digital ocean.



I would like to challenge this picture, and ask you to imagine data not as a pristine resource, but as a waste product, a bunch of radioactive, toxic sludge that we don't know how to handle.

Maciej Ceglowski

http://idlewords.com/talks/haunted_by_data.htm



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.

Resources, People to Follow

- Data & Society, <https://datasociety.net/>
- Zeynep Tufekci, Kate Crawford, danah boyd
- Mimi Onuoha
 - “To collect, record, and archive aspects of the world is an intentional act.”