

Hi! We're Todd Hartman, Aneta Piekut and Mark Taylor from the Sheffield Methods Institute and we look at how the media uses (and misuses) data and statistics. Ask us anything!

smi_tkhartman¹andr/ScienceAMAs¹

¹Affiliation not available

April 17, 2023

Abstract

Hi everyone! We are lecturers in quantitative social science at the Sheffield Methods Institute. Increasingly, the media bombards us with all sorts of data about how society is changing: opinion poll trends; migration data; economic results; government debt levels; and politicians' expenses claims. We look at where those numbers come from, can they be trusted and how they can be manipulated visually and in written form to support a contentious claim. Todd Hartman: I'm a political psychologist by training, and I've got extensive experience conducting surveys and experiments. My current research focuses on political attitudes and intergroup relations. Before I came to Sheffield, I was Director of Survey Research for the Centre for Economic Research and Policy Analysis as well as Assistant Professor of Political Science at Appalachian State University. I've been in Sheffield for about a year and a half, and in that time I've got heavily into rugby and real ale. Aneta Piekut: I was trained as a sociologist, but have been working in a different subdisciplines of social science, mixing various research methods. In my research I am interested in such topics as social diversity, social inclusion, integration of ethnic minority groups and socio-spatial segregation, working with surveys and secondary data. I spend my spare time in a gym or swimming, and walking Czarek, a rescue dog, whose adventures you can follow on Instagram. Mark Taylor: I'm a sociologist who's interested in culture, broadly defined - so music, video games, TV, books, and so on. I mainly work with survey data, but also work with data from schools, the labour market, and other more-or-less official sources. For graphics I'm a total evangelist for ggplot2, and I'm in the process of getting my head round Tableau as well. I also spend an inordinate amount of my time playing the Binding of Isaac. We also developed this course to help people brush up their social statistics skills and help combat the rising trend of misleading data visualizations. Here's proof that it's us! We'll be back at 11am ET/4pm GMT to answer your questions. Ask us anything! EDIT: We're ready to go, and we've been joined by our colleague Andrew Bell who's also a lecturer in quantitative social science! EDIT: We're signing off for now. Thanks everyone for some great questions and insightful discussion! We'll keep an eye on the AMA if you think there's any big questions we've missed and try to get round to them! Also if you want to freshen up your social statistics skills then check out our course on data in the media.

[REDDIT](#)

Hi! We're Todd Hartman, Aneta Piekut and Mark Taylor from the Sheffield Methods Institute and we look at how the media uses (and misuses) data and statistics. Ask us anything!

SMI_TKHARTMAN [R/SCIENCE](#)

Hi everyone! We are lecturers in quantitative social science at the Sheffield Methods Institute.

Increasingly, the media bombards us with all sorts of data about how society is changing: opinion poll trends; migration data; economic results; government debt levels; and politicians' expenses claims.

We look at where those numbers come from, can they be trusted and how they can be manipulated visually and in written form to support a contentious claim.

Todd Hartman: I'm a political psychologist by training, and I've got extensive experience conducting surveys and experiments. My current research focuses on political attitudes and intergroup relations. Before I came to Sheffield, I was Director of Survey Research for the Centre for Economic Research and Policy Analysis as well as Assistant Professor of Political Science at Appalachian State University. I've been in Sheffield for about a year and a half, and in that time I've got heavily into rugby and real ale.

Aneta Piekut: I was trained as a sociologist, but have been working in a different subdisciplines of social science, mixing various research methods. In my research I am interested in such topics as social diversity, social inclusion, integration of ethnic minority groups and socio-spatial segregation, working with surveys and secondary data. I spend my spare time in a gym or swimming, and walking Czarek, a rescue dog, whose adventures you can [follow on Instagram](#).

Mark Taylor: I'm a sociologist who's interested in culture, broadly defined - so music, video games, TV, books, and so on. I mainly work with survey data, but also work with data from schools, the labour market, and other more-or-less official sources. For graphics I'm a total evangelist for ggplot2, and I'm in the process of getting my head round Tableau as well. I also spend an inordinate amount of my time playing the Binding of Isaac.

We also developed [this course](#) to help people brush up their social statistics skills and help combat the rising trend of misleading data visualizations.

[Here's proof that it's us!](#)

We'll be back at 11am ET/4pm GMT to answer your questions.

Ask us anything!

EDIT: We're ready to go, and we've been joined by our colleague [Andrew Bell](#) who's also a lecturer in quantitative social science!

EDIT: **We're signing off for now. Thanks everyone for some great questions and insightful discussion!**

We'll keep an eye on the AMA if you think there's any big questions we've missed and try to get round to them! Also if you want to freshen up your social statistics skills then check out our course on [data in the media](#).

[READ REVIEWS](#)

Can we trust any data the media uses? There seems to be contradictory information published constantly.

[WRITE A REVIEW](#)

CORRESPONDENCE:

[georgiauv](#)

DATE RECEIVED:

February 05, 2016

Data isn't really the problem; poor use of data is. I think the best approach is to become a good

DOI:
10.15200/winn.145459.92397

ARCHIVED:
February 04, 2016

CITATION:
smi_tk hartman , r/Science , Hi!
We're Todd Hartman, Aneta
Piekut and Mark Taylor from
the Sheffield Methods Institute
and we look at how the media
uses (and misuses) data and
statistics. Ask us anything!, *The
Winnower* 3:e145459.92397 ,
2016 , DOI:
[10.15200/winn.145459.92397](https://doi.org/10.15200/winn.145459.92397)

© et al. This article is
distributed under the terms of
the [Creative Commons
Attribution 4.0 International
License](https://creativecommons.org/licenses/by/4.0/), which permits
unrestricted use, distribution,
and redistribution in any
medium, provided that the
original author and source are
credited.



consumer of data so that you can recognize what makes "good" and "bad" evidence. But I agree, I often find the contradictory evidence (e.g., health research) problematic.

Hi, I am a secondary school teacher - what would you like 16-18 year olds to know about this topic?

[NemeReddit](#)

One of the problems for this age group is that there's a lot of focus on mathematics but little on statistics. So, first, I'd like there to be coverage of things like the Law of Large Numbers, the Central Limit Theorem, Probability Theory, and hypothesis testing. I'd probably like to see one step further-- applied statistics taught with specific, compelling examples. When I first learned stats, I spent a lot of time going through proofs and other less practical statistics, which can turn people off.

Find really cool, fun examples that motivate people to explore data!

I have a broad question. Where does the media tend to get its data, polls, statistics from? Is there a go to source for each?

[Thegermanway](#)

Some commission their own polls or work with regular survey research firms. In these cases, the data collected is usually high quality because the organizations with whom they're working are reputable. For instance, the Washington Post is pretty good about being transparent re: they're data sources.

538 has a list of ranking of pollsters, which is helpful: <http://fivethirtyeight.com/interactives/pollster-ratings/>

Thanks for doing this!

You no doubt know of the *Literary Digest* poll of 1936 declaring Alf Landon the probable victor in the U.S. presidential election — of course, he was swamped by FDR. This failure has been analyzed and the faulty polling practices described. I'm wondering what sorts of trends are potentially sabotaging poll results today.

Many people no longer use land line telephones and aren't listed in directories. Caller ID has surely affected the rate of response for those people who do still have traditional phone service. Increased cynicism has led to calls for people to give misleading answers to pollsters.

Even more unsettling are the number of online polls. These allow anyone with a browser to answer — which would seem to qualify as a "self-selected survey". How are these sorts of result analyzed?

[gargle_ground_glass](#)

The short answer to your question is that they're often analyzed just as we've always done without really thinking about the problems that these issues create for data quality.

The biggest concern for survey research is that it has become increasingly difficult to recruit representative samples--as you note, fewer people have landlines and cell phone frames are inconsistent. Internet samples can be ok depending on how their recruited, and eventually someone will figure out how to have truly representative Internet panels (without contacting people via landlines).

In the end, I like to think of data as evidence, and although it's easier than ever to collect data these days, the quality can certainly vary significantly from project to project.

PhD Political Science student here, thanks for doing your work and this AMA!

1. What statistical factor do you think the media leaves out of data the most and that would be the most telling?
2. Even if they included it, do you think a big issue is that non-science majors or non-academic wouldn't understand these results (like when assumptions are violated or estimators or any other statistical jargon)?
3. What would you tell the average person to look for the most and why is it not R-squared?

Thanks!

[mapxvc](#)

1. That correlation is not the same as causation; sample sizes and resultant margin of errors for estimates.
2. I think journalists have a responsibility to help make sense of the information that they're reporting, so it's incumbent on them to be knowledgeable enough to properly interpret the results of data they're using in their stories.
3. R-squared just tells us how much variation our model explains; often the goal of a social scientist is not explaining the most variation but testing hypotheses--that is, finding effects in the real world. I tend to look at the source of the data, sample size, effect size, etc. Good luck in your program!

What are the most common mistakes made in survey design leading to misleading results? How does one go about designing a good survey?

[Emeson](#)

This is a tough question because there are many ways to go wrong. The big issues are with question wording problems (e.g., loaded questions, double-barrelled questions, etc.), question ordering, and sampling issues. Don Dillman has a really good how-to book worth checking out:
http://www.amazon.com/Internet-Phone-Mail-Mixed-Mode-Surveys/dp/1118456149/ref=sr_1_fkmr0_1?ie=UTF8&qid=1454603388&sr=8-1-fkmr0&keywords=the+tailor+design+method+dillman

Hello! My question is how many people (roughly) do you think actually believe false or biased data? If the number is large is there any way to minimize the effect that media has?

One such example would be the gun control debate. One party believes that gun violence is running rampant throughout the country, while the opposition believes that gun violence is on a massive decline. How can it be both? (it can't) More importantly how do people come to conclusions like this that could not be more different?

(Sorry I know a lot of questions haha)

[MarianasTrench](#)

There's a theory in social psychology called motivated reasoning that helps answer your question. Basically, people have different goals, one of which is accuracy and the other is partisan. Partisan goals mean that people will be biased based on their strong prior beliefs, which means that they'll be quick to accept information consistent with their own viewpoints and spend considerable energy attacking contrary information. So, I think people use data to support their own views without trying to be objective.

Any plans to have a regular newspaper column like Ben Goldacre used to have in the Guardian? Feels like a really useful public service to me.

[iff_true](#)

Would love to, but that's a lot of work! We all have our regular jobs to do....

Do you think a form of accreditation by an independent statistics body where media have to prove their journalists use data/statistics responsibly be of benefit to the media? Are you aware of this existing anywhere?

[butterflykiller](#)

I think that could be one way to help signify to readers that the information is being reported in a fair and transparent manner. I know some media outlets sign on to transparency initiatives from academic associations like AAPOR: <https://www.aapor.org/>

There's a good discussion of this here: http://www.huffingtonpost.com/2014/10/03/2014-election-poll-transparency_n_5921860.html

And, these are they types of issues that we discuss in our MOOC:

<https://www.futurelearn.com/courses/media-data>

Hello, and thank you for doing this AMA!

I currently teach an introductory level class on spreadsheets and databases. Quantitative analysis is a small component of the class currently, yet there is a need to increase this type of analysis. From your perspective, where should one begin in teaching quantitative analysis? What are the three to four biggest ideas/concepts that every student should walk out of my class knowing? Any suggestions and resources would be most helpful!

[Mouthoy](#)

Start with interesting examples and work backwards. Too often people try to teach statistics within arbitrary, uninteresting contexts and turn students off. In my experience, the single determining factor in whether someone does well or poorly in a stats class is motivation (at least in my classes).

The Central Limit Theorem and the Law of Large Numbers; probability theory; hypothesis testing are good concepts to cover. But really the key is getting students to appreciate the value of statistics--I often here "what can statistics teach me about X (e.g., politics)?" Statistics are simply a way to make sense of information; to find reliable patterns.

If your students can get used to seeing numbers and thinking about interesting relationships in the data, then you've done well.

What drove you all, individually and collectively, to research something like this? What trends are most surprising to you?

[validusrex](#)

My doctoral program was heavily quantitative. Like many students, I didn't really appreciate the value

of stats until I had been forced to learn them. Now, 15 years later, I really enjoy stats to help make sense of the world around me!

What do you think about the rise of "think tanks" and how they curate the data to support their policy objectives?

[loercase](#)

Some think tanks do some really good work, but others have political agendas first and work the data to fit their goals. So, I like the ones that do honest/fair data analysis and ditch the rest.

To an extent, is it fantasy to expect a lack of bias?

[alexgorale](#)

Yes,. But if we're aware of the potential problems and try our best to be objective/fair, I think that's better than just accepting that bias is the norm.

I'm a math teacher and we have a class at our school that talks about this very subject.

I don't really have a question anymore, as I just saw your website and was looking for cool/useful things to introduce to the students. This topic is never publicly talked about yet is around us at all times. Thank you for coming here and I hope others realize how big of an issue it really can be.

If you had something very cool for me to show the students though, by all means, post it here!

[MEuRaH](#)

There's some great data visualizations (see above examples). Our own Alasdair Rae does some pretty cool spatial data visualizations: <http://www.statsmapsnpix.com/>

What's the most effective methods of combatting bad data reporting? Given that the headline is usually so hyperbolic and effective and the analysis is (by necessity) long and in-depth, can we ever win in this era of tabloid shock and easy linkbating?

[DaveMedlo](#)

Read the methodology section or look more deeply into what's actually being reported. Often lazy or uninformed reports sensationalize things based upon pretty flimsy evidence.

How do you feel about Fox News obscuring facts go portray their image of the situation? What do you have to say if you could say anything to Bill O'Reilly?

[DatOneGuy-69](#)

I understand people like Bil O'Reilly do what they do because there's a market for it, but I really wish we'd get back to the days of high quality journalism, where journalists did more than simply reporting the news (of course, this is a massive oversimplification: I recognize that there's some really great journalism going on, but it often gets squeezed out by "talking heads").

Bill: Stop making people so angry (or at least feeding the anger).

Does reality have a liberal bias?

[CunningAllusionment](#)

It depends on your ideology. Liberals see reality one way; conservatives the other.

- What have you found in regards to meteorologists misrepresenting data or detracting the truth about the global warming?
- Do you think the appropriate development and survival of our civilization through science stands against the tactics of major oil, gas and motor industries?
- How can rectifying scientific data hope to make a difference if it is not a concern for these industries?

[lamgmg90](#)

I don't think the real issue is science or data. The real problem standing in the way of our ability to address key issues facing humanity is political.