Introduction to Data Science

Session 12: Monitoring and Communication

Simon Munzert
Hertie School | GRAD-C11/E1339

Table of contents

Communicating data science

The final piece of the pipeline

Preparatory work

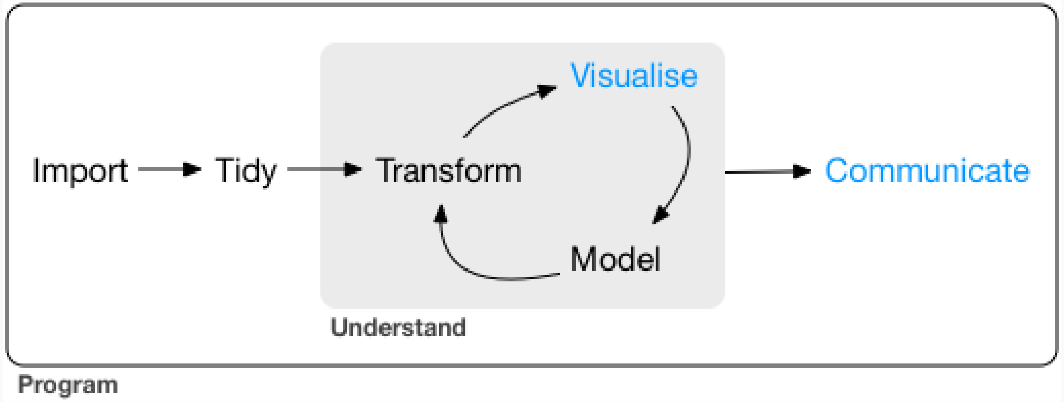
- **Problem definition** predict, infer, describe
- **Design** conceptualize, build data collection device
- **Data collection** recruit, collect, monitor

Data operation

- **Wrangle**: import, tidy, manipulate
- **Explore**: visualize, describe, discover
- **Model**: build, test, infer, predict

Dissemination

- 👉👉 **Communicate**: to the public, media, policymakers
- 👉👉 **Publish**: articles, blogs, software
- **Productize**: make usable, robust, scalable



Communicating data science

"[I]t doesn’t matter how great your analysis is unless you can explain it to others: you need to communicate your results."

Hadley Wickham & Garrett Grolemund, *R for Data Science*

Lasswell model of communication **for data scientists**

Laswell's framework of communication[1] dissects the task of communication along the following dimensions: (1) Who communicates (2) what (3) in what form (4) to whom (5) to what effect?

Let's apply this to us. Data scientists communicate…

[1]HD Lasswell. 1948. The structure and function of communication in society. In *The communication of ideas* (ed. Bryson L), 37-51.

Lasswell model of communication **for data scientists**

Laswell's framework of communication[1] dissects the task of communication along the following dimensions: (1) Who communicates (2) what (3) in what form (4) to whom (5) to what effect?

Let's apply this to us. Data scientists communicate...

What

- Estimates
- Uncertainty
- Model implications
- Substantive knowledge
- Product
- Themselves

[1] HD Lasswell. 1948. The structure and function of communication in society. In *The communication of ideas* (ed. Bryson L), 37-51.

Lasswell model of communication **for data scientists**

Laswell's framework of communication[1] dissects the task of communication along the following dimensions: (1) Who communicates (2) what (3) in what form (4) to whom (5) to what effect?

Let's apply this to us. Data scientists communicate...

What

- Estimates
- Uncertainty
- Model implications
- Substantive knowledge
- Product
- Themselves

How

- Spoken word
- Technical reports
- Academic papers
- Web applications
- Policy briefs

[1] HD Lasswell. 1948. The structure and function of communication in society. In *The communication of ideas* (ed. Bryson L), 37-51.

Lasswell model of communication **for data scientists**

Laswell's framework of communication[1] dissects the task of communication along the following dimensions: (1) Who communicates (2) what (3) in what form (4) to whom (5) to what effect?

Let's apply this to us. Data scientists communicate...

What

- Estimates
- Uncertainty
- Model implications
- Substantive knowledge
- Product
- Themselves

How

- Spoken word
- Technical reports
- Academic papers
- Web applications
- Policy briefs

To whom

- The public
- The media
- Policymakers
- Other scientists
- Managers / co-workers

[1] HD Lasswell. 1948. The structure and function of communication in society. In *The communication of ideas* (ed. Bryson L), 37-51.

Lasswell model of communication **for data scientists**

Laswell's framework of communication[1] dissects the task of communication along the following dimensions: (1) Who communicates (2) what (3) in what form (4) to whom (5) to what effect?

Let's apply this to us. Data scientists communicate...

What

- Estimates
- Uncertainty
- Model implications
- Substantive knowledge
- Product
- Themselves

How

- Spoken word
- Technical reports
- Academic papers
- Web applications
- Policy briefs

To whom

- The public
- The media
- Policymakers
- Other scientists
- Managers / co-workers

To what end

- Inform
- Influence
- Instruct
- Motivate
- Monitor
- Document

[1] HD Lasswell. 1948. The structure and function of communication in society. In *The communication of ideas* (ed. Bryson L), 37-51.

Lasswell model of communication **for data scientists**

Laswell's framework of communication[1] dissects the task of communication along the following dimensions: (1) Who communicates (2) what (3) in what form (4) to whom (5) to what effect?

Let's apply this to us. Data scientists communicate...

What

- Estimates
- Uncertainty
- Model implications
- Substantive knowledge
- Product
- Themselves

How

- Spoken word
- Technical reports
- Academic papers
- Web applications
- Policy briefs

To whom

- The public
- The media
- Policymakers
- Other scientists
- Managers / co-workers

To what end

- Inform
- Influence
- Instruct
- Motivate
- Monitor
- Document

What, how, and to what end you communicate depends on your **audience/stakeholders** because they will differ in interest, contextual knowledge, data literacy, and motives.

[1] HD Lasswell. 1948. The structure and function of communication in society. In *The communication of ideas* (ed. Bryson L), 37-51.

Statistical communication

Statistical communication

What we communicate

- The quantity of interest
- The selection/generation of data
- The empirical setup
- The model mechanics and results (estimates/predictions/uncertainty)

Common challenges

- There's epistemological and statistical uncertainty.
- Effect sizes have implications that are often not easy to grasp.
- Conclusions about data science output crucially hinge on the validity of design aspects, which are extremely difficult to communicate.

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 DECEMBER 31, 2020 VOL. 383 NO. 27

Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine

Fernando P. Polack, M.D., Stephen J. Thomas, M.D., Nicholas Kitchin, M.D., Judith Absalon, M.D., Alejandra Gurtman, M.D., Stephen Lockhart, D.M., John L. Perez, M.D., Gonzalo Pérez Marc, M.D., Edson D. Moreira, M.D., Cristiano Zerbini, M.D., Ruth Bailey, B.Sc., Kena A. Swanson, Ph.D., Satrajit Roychoudhury, Ph.D., Kenneth Koury, Ph.D., Ping Li, Ph.D., Warren V. Kalina, Ph.D., David Cooper, Ph.D., Robert W. Frenck, Jr., M.D., Laura L. Hammitt, M.D., Özlem Türeci, M.D., Haylene Nell, M.D., Axel Schaefer, M.D., Serhat Ünal, M.D., Dina B. Tresnan, D.V.M., Ph.D., Susan Mather, M.D., Philip R. Dormitzer, M.D., Ph.D., Uğur Şahin, M.D., Kathrin U. Jansen, Ph.D., and William C. Gruber, M.D., for the C4591001 Clinical Trial Group*

ABSTRACT

BACKGROUND
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection and the resulting coronavirus disease 2019 (Covid-19) have afflicted tens of millions of people in a worldwide pandemic. Safe and effective vaccines are needed urgently.

METHODS
In an ongoing multinational, placebo-controlled, observer-blinded, pivotal efficacy trial, we randomly assigned persons 16 years of age or older in a 1:1 ratio to receive two doses, 21 days apart, of either placebo or the BNT162b2 vaccine candidate (30 μg per dose). BNT162b2 is a lipid nanoparticle–formulated, nucleoside-modified RNA vaccine that encodes a prefusion stabilized, membrane-anchored SARS-CoV-2 full-length spike protein. The primary end points were efficacy of the vaccine against laboratory-confirmed Covid-19 and safety.

RESULTS
A total of 43,548 participants underwent randomization, of whom 43,448 received injections: 21,720 with BNT162b2 and 21,728 with placebo. There were 8 cases of Covid-19 with onset at least 7 days after the second dose among participants assigned to receive BNT162b2 and 162 cases among those assigned to placebo; BNT162b2 was 95% effective in preventing Covid-19 (95% credible interval, 90.3 to 97.6). Similar vaccine efficacy (generally 90 to 100%) was observed across subgroups defined by age, sex, race, ethnicity, baseline body-mass index, and the presence of coexisting conditions. Among 10 cases of severe Covid-19 with onset after the first dose, 9 occurred in placebo recipients and 1 in a BNT162b2 recipient. The safety profile of BNT162b2 was characterized by short-term, mild-to-moderate pain at the injection site, fatigue, and headache. The incidence of serious adverse events was low and was similar in the vaccine and placebo groups.

CONCLUSIONS
A two-dose regimen of BNT162b2 conferred 95% protection against Covid-19 in persons 16 years of age or older. Safety over a median of 2 months was similar to that of other viral vaccines. (Funded by BioNTech and Pfizer; ClinicalTrials.gov number, NCT04368728.)

The authors' affiliations are listed in the Appendix. Address reprint requests to Dr. Absalon at Pfizer, 401 N. Middletown Rd., Pearl River, NY 10965, or at judith.absalon@pfizer.com.

*A complete list of investigators in the C4591001 Clinical Trial Group is provided in the Supplementary Appendix, available at NEJM.org.

Drs. Polack and Thomas contributed equally to this article.

This article was published on December 10, 2020, and updated on December 16, 2020, at NEJM.org.

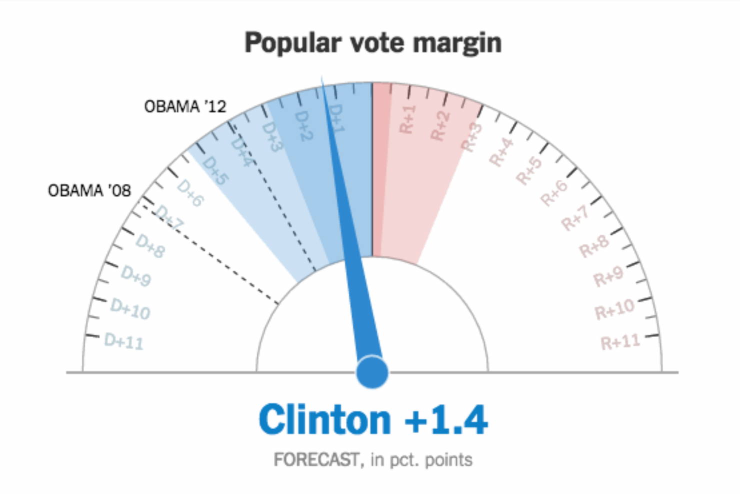
N Engl J Med 2020;383:2603-15.
DOI: 10.1056/NEJMoa2034577

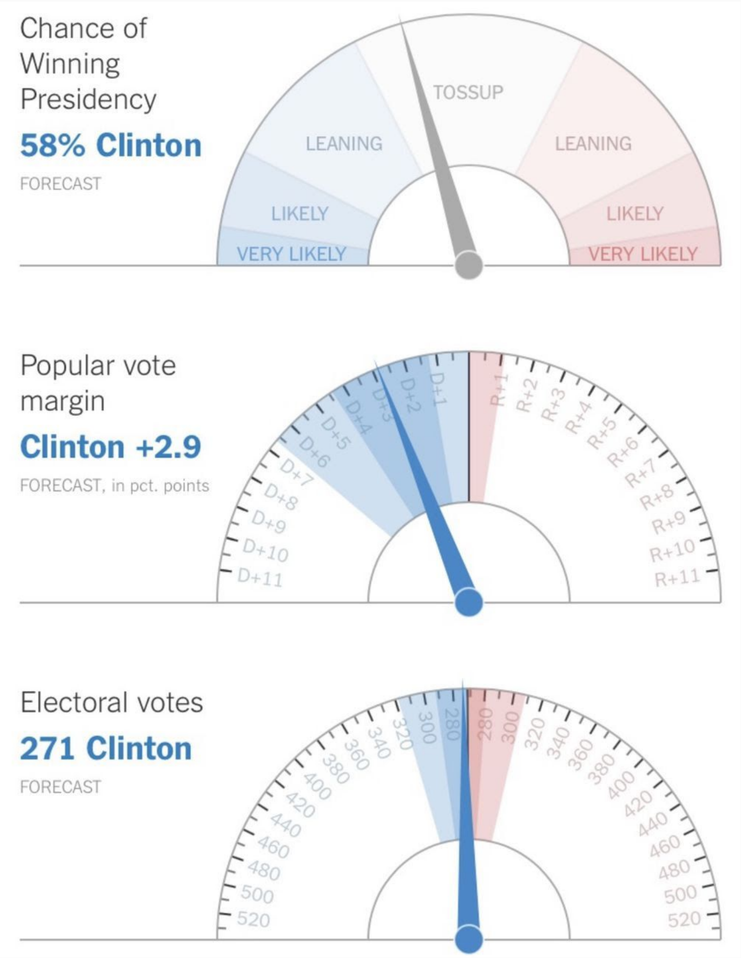
`Credit` Polack et al. 2020, NEJM

Uncertainty

Question to reflect on

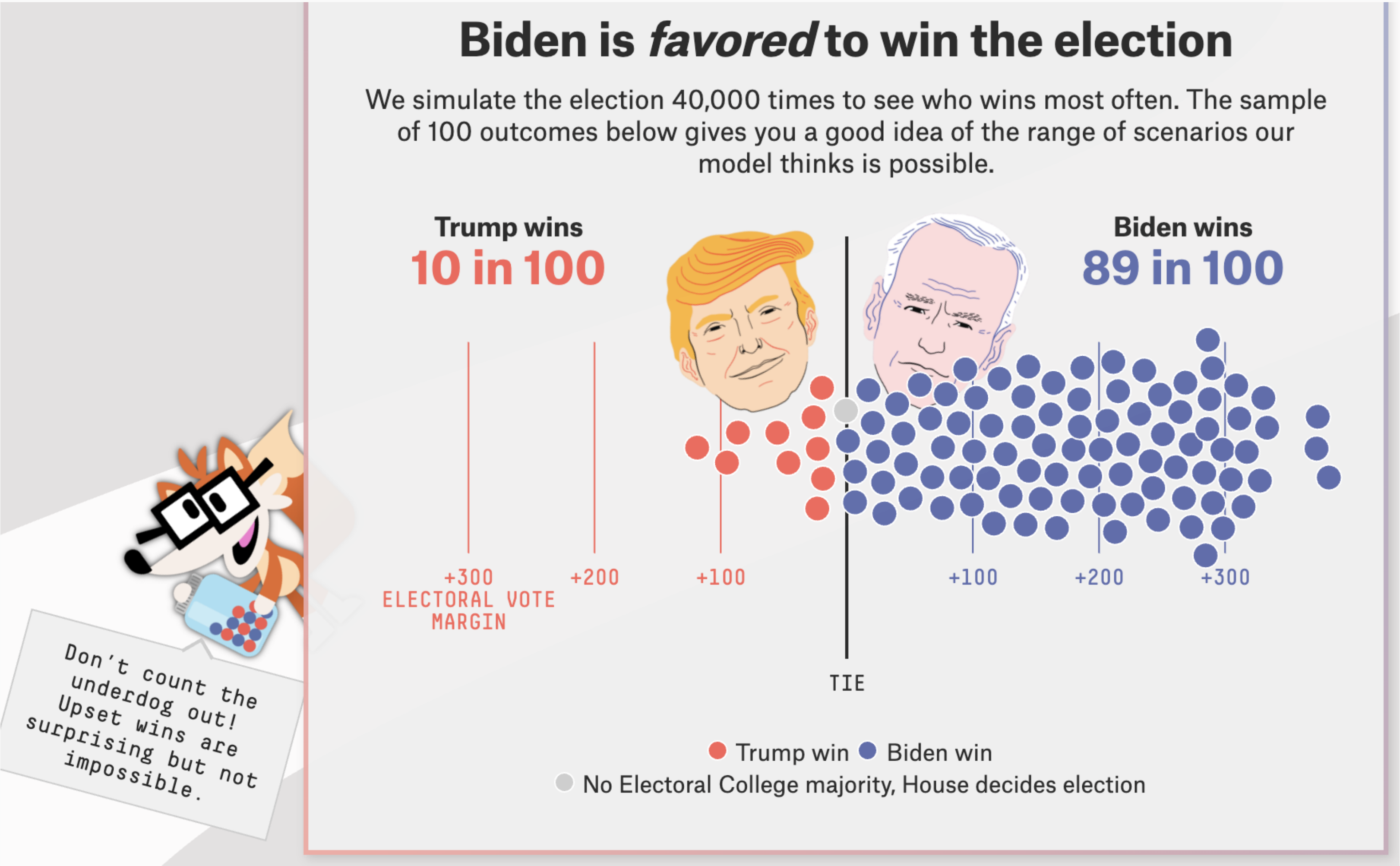
Thinking about the task of publicly forecasting the U.S. presidential election using polling data, what kinds of uncertainty would you have to deal with as a forecaster?





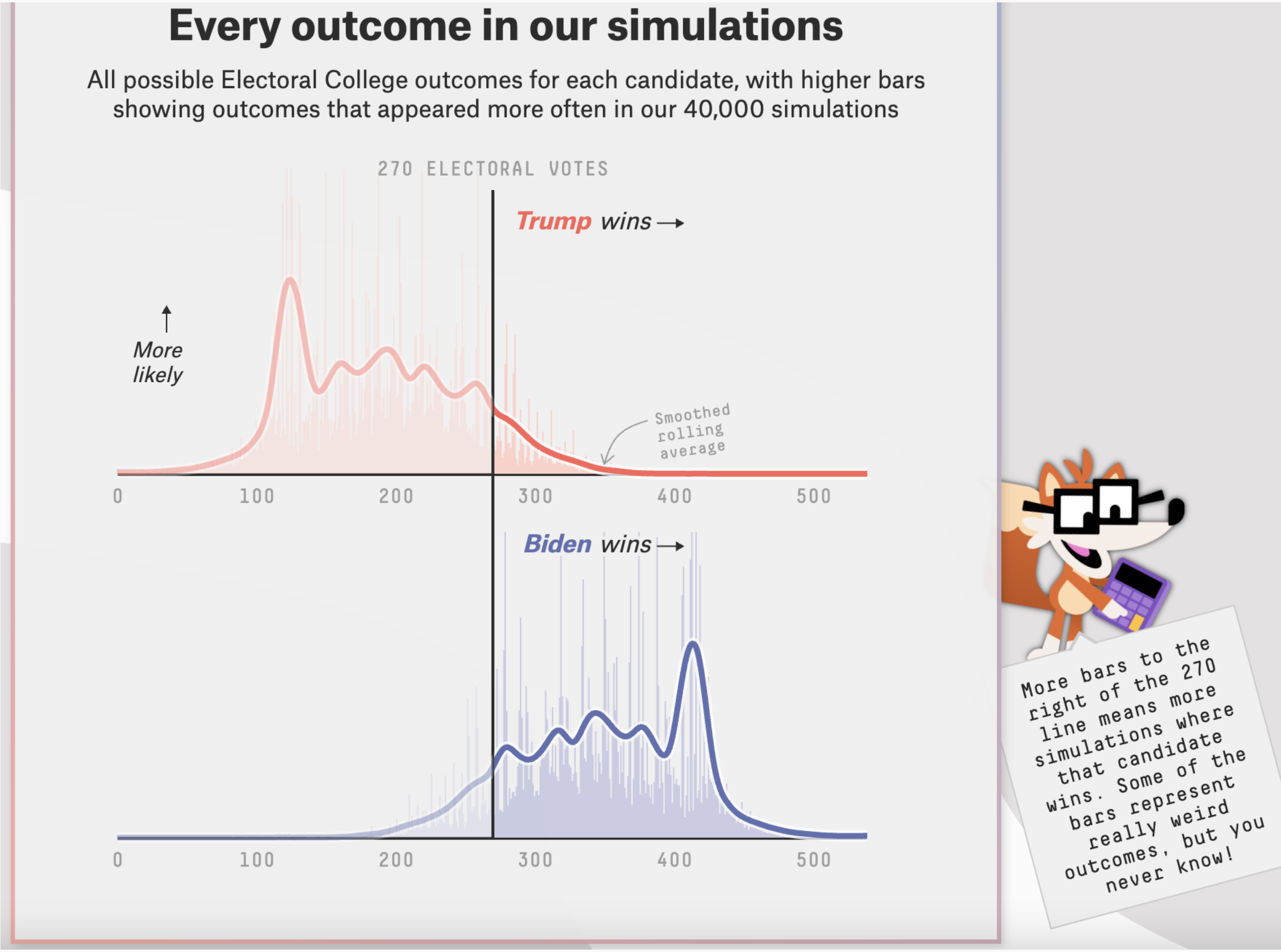
`Credit` NYTimes.com at 9:20 p.m. Nov. 8, 2016

Example: FiveThirtyEight 2020 election forecast



Source FiveThirtyEight

Example: FiveThirtyEight 2020 election forecast



Source FiveThirtyEight

Example: FiveThirtyEight 2020 election forecast

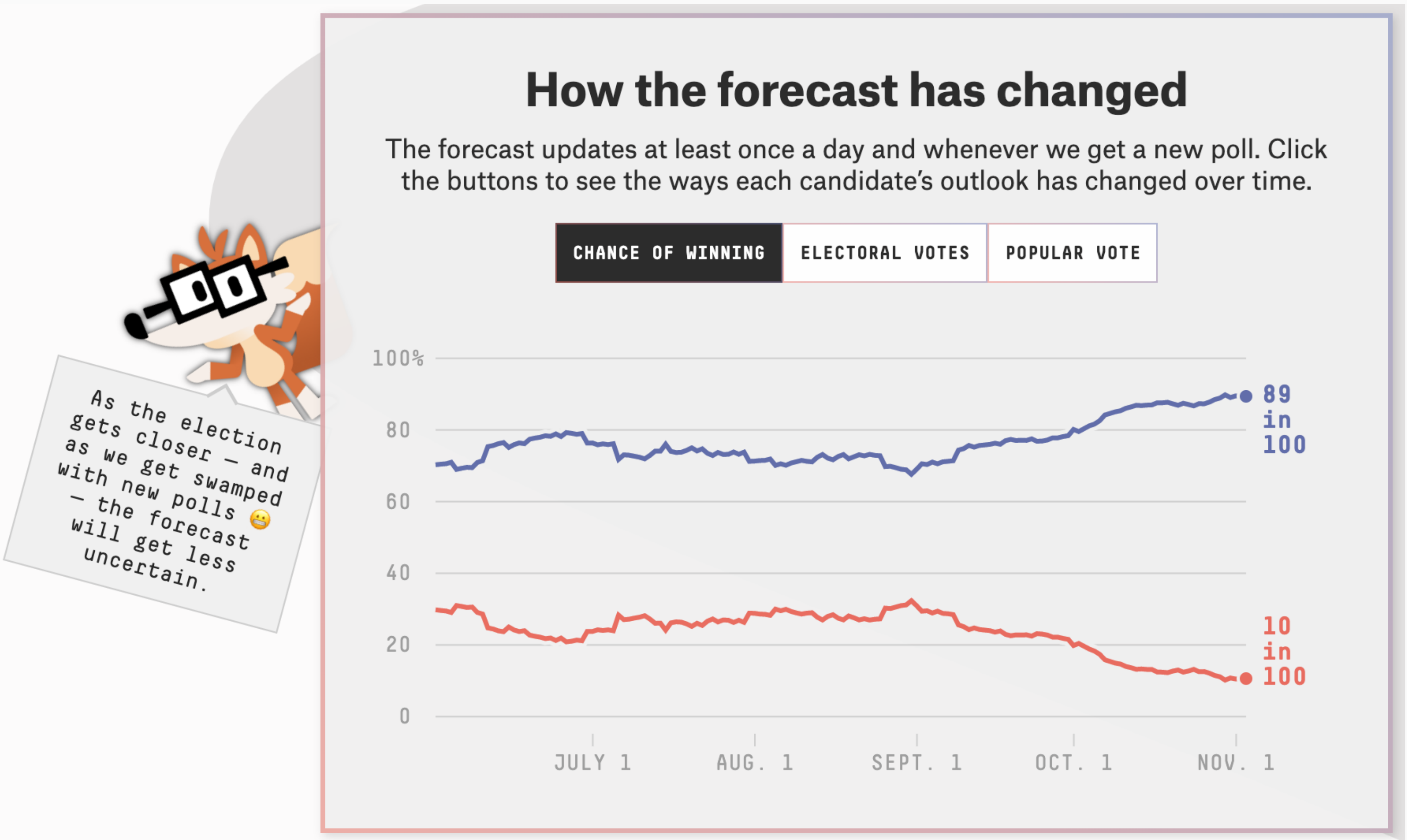
Weird and not-so-weird possibilities

The chances that these situations will crop up

Trump wins the popular vote Regardless of whether he wins the Electoral College	3 in 100
Biden wins the popular vote Regardless of whether he wins the Electoral College	97 in 100
Trump wins more than 50% of the popular vote Regardless of whether he wins the Electoral College	1 in 100
Biden wins more than 50% of the popular vote Regardless of whether he wins the Electoral College	95 in 100
Trump wins in a landslide Defined as winning the popular vote by a double-digit margin	<1 in 100
Biden wins in a landslide Defined as winning the popular vote by a double-digit margin	29 in 100
Trump wins the popular vote but loses the Electoral College	<1 in 100
Biden wins the popular vote but loses the Electoral College	8 in 100
No one wins the Electoral College No candidate gets 270 electoral votes and Congress decides the election	<1 in 100

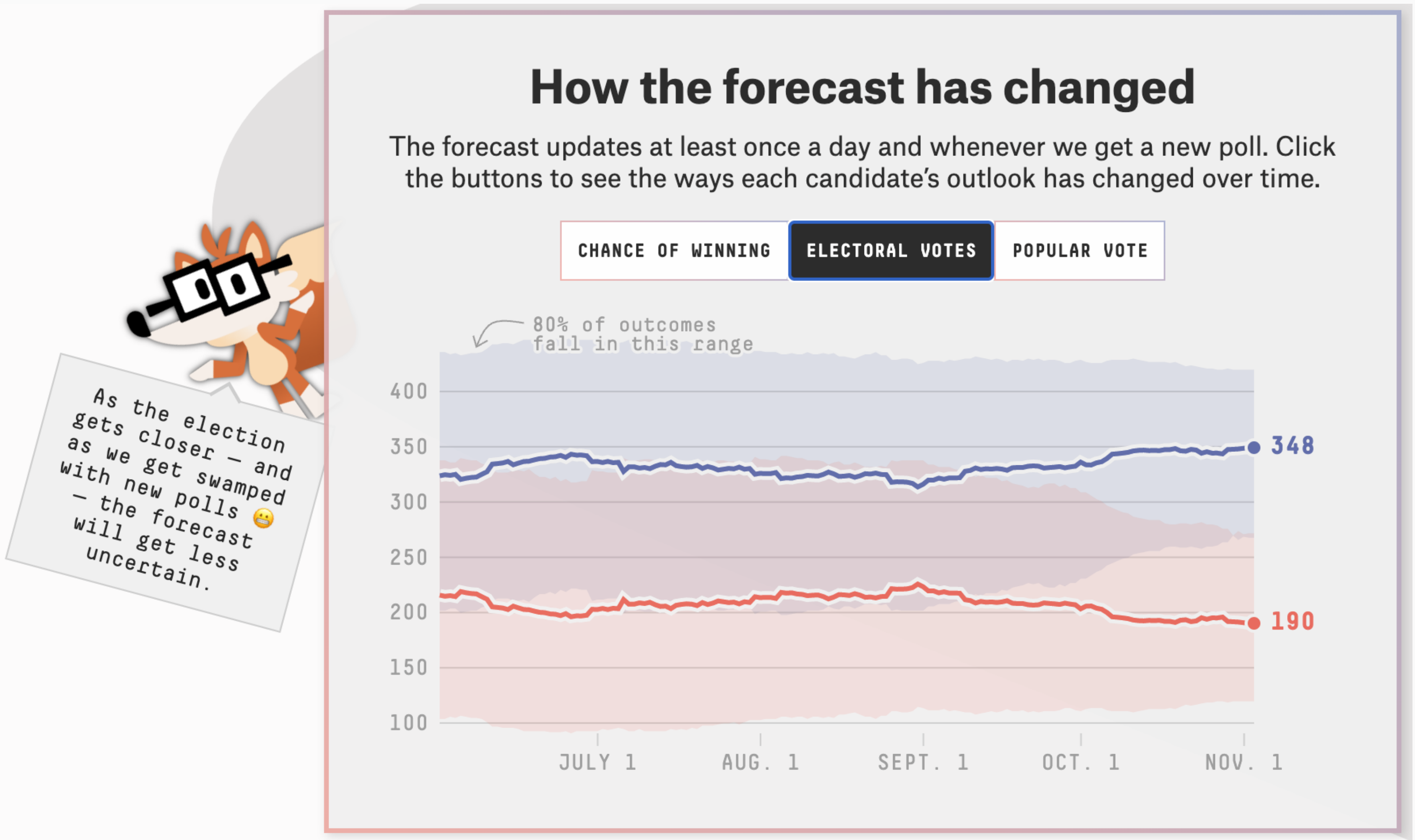
Source FiveThirtyEight

Example: FiveThirtyEight 2020 election forecast



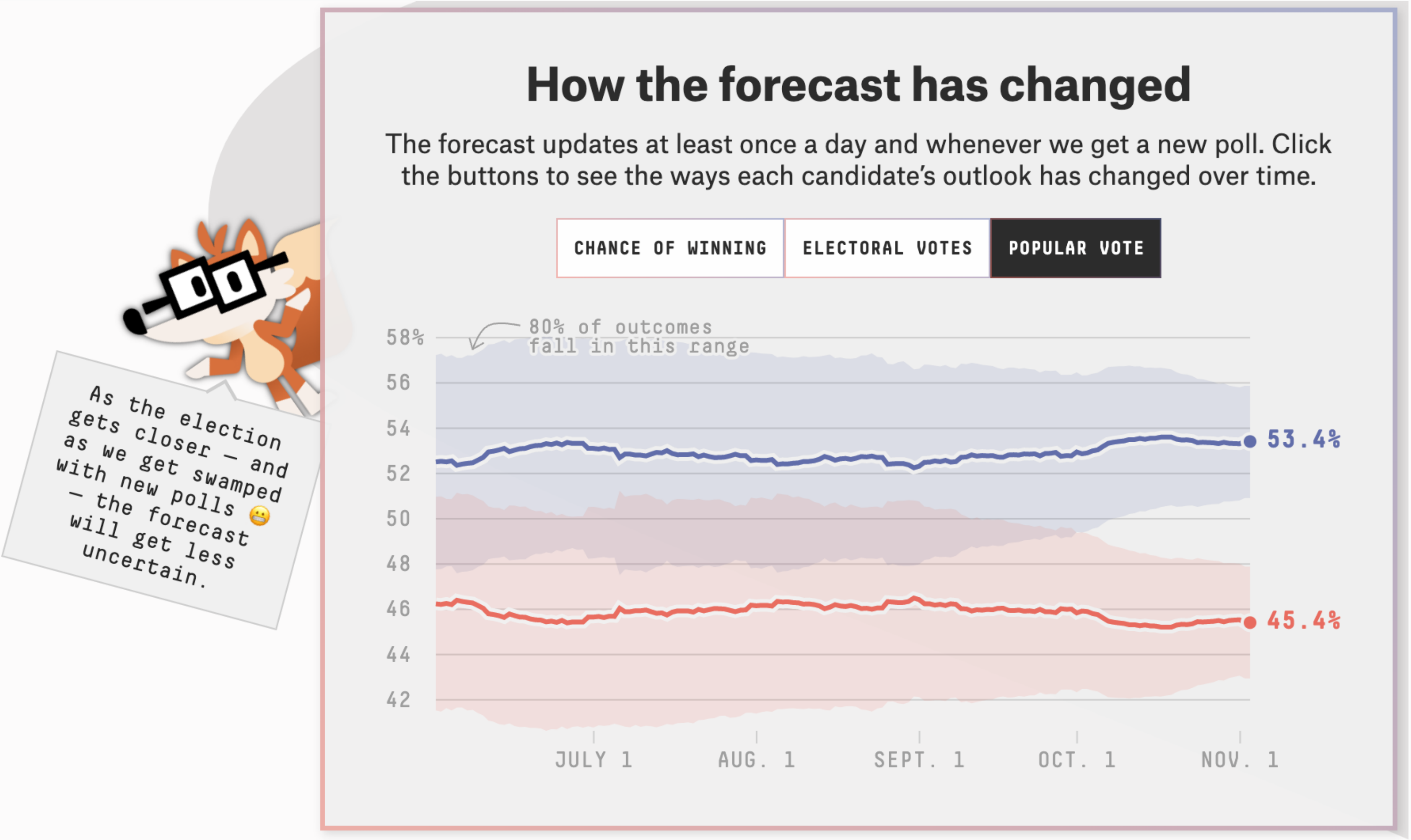
Source FiveThirtyEight

Example: FiveThirtyEight 2020 election forecast



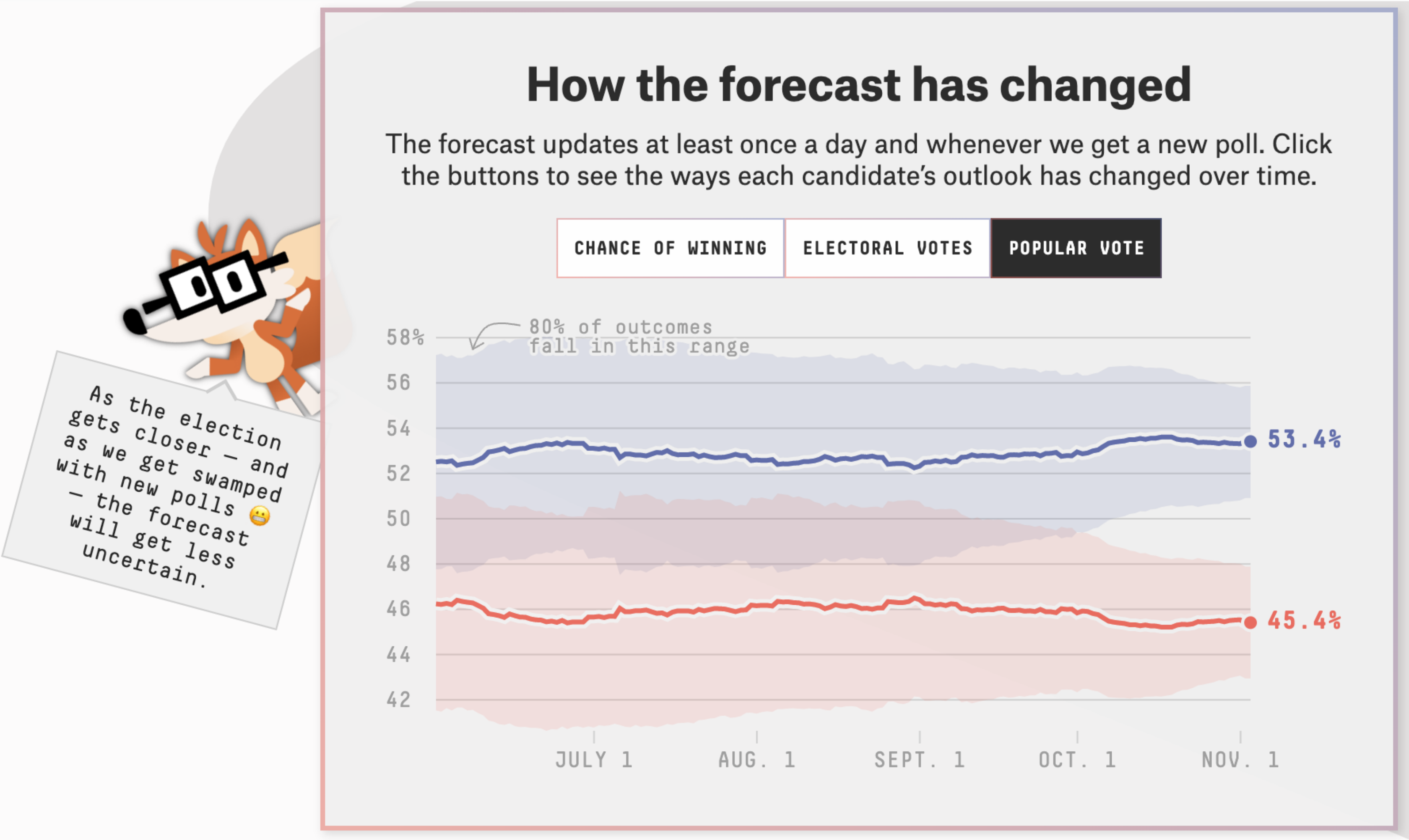
Source FiveThirtyEight

Example: FiveThirtyEight 2020 election forecast



Source FiveThirtyEight

Example: FiveThirtyEight 2020 election forecast



Source FiveThirtyEight

Example: How the sausage is made

AUG. 12, 2020, AT 6:30 AM

How FiveThirtyEight's 2020 Presidential Forecast Works — And What's Different Because Of COVID-19

By Nate Silver

Filed under 2020 Election

Our presidential forecast, which launched today, is not the first election forecast that FiveThirtyEight has published since 2016. There was our midterms forecast in 2018, which was pretty accurate in predicting the makeup of the House and the Senate. And there was our presidential primaries model earlier this year, which was a bit of an adventure but mostly notable for being bullish (correctly) on Joe Biden and (incorrectly) on Bernie Sanders. But we're aware that the publication of our first presidential forecast since 2016 is liable to be fraught.

`Source` FiveThirtyEight

Example: How the sausage is made

AUG. 12, 2020, AT 6:30 AM

uncertain 1/26

How FiveThirty
Presidential Forecast Works — And What's Different Because Of COVID-19

By Nate Silver

Filed under 2020 Election

Our presidential forecast, which launched today, is not the first election forecast that FiveThirtyEight has published since 2016. There was our midterms forecast in 2018, which was pretty accurate in predicting the makeup of the House and the Senate. And there was our presidential primaries model earlier this year, which was a bit of an adventure but mostly notable for being bullish (correctly) on Joe Biden and (incorrectly) on Bernie Sanders. But we're aware that the publication of our first presidential forecast since 2016 is liable to be fraught.

`Source` FiveThirtyEight

Example: How the sausage is made

Step 3: Account for uncertainty, and simulate the election thousands of times

As complicated though it may seem, everything I've described up until this point is, in some sense, the easy part of developing our model. There's no doubt that Biden is comfortably ahead as of the forecast launch in mid-August, for example, and the choices one makes in using different methods to average polls or combine them with other data isn't likely to change that conclusion.

What's trickier is figuring out how that translates into a probability of Biden or Trump winning the election. That's what this section is about.

`Source` FiveThirtyEight

Example: How the sausage is made

With that disclaimer out of the way, here are the four types of uncertainty that the model tries to account for:

1. **National drift**, or how much the overall national forecast could change between now and Election Day.
2. **National Election Day error**, or how much our final forecast of the national popular vote could be off on Election Day itself.
3. **Correlated state error**, which reflects errors that could occur across multiple states along geographic or regional lines — for instance, as was relevant in 2016, a systematic underperformance relative to polls for the Democratic candidate in the Midwest.
4. **State-specific error**, an error relative to our forecast that affects only one state.

`Source` FiveThirtyEight

Example: How the sausage is made

The components of our uncertainty index are as follows:

1. The number of undecided voters in national polls. More undecided voters means more uncertainty.
2. The number of undecided plus third-party voters in national polls. More third-party voters means more uncertainty.
3. Polarization, as measured elsewhere in the model, is based on how far apart the parties are in roll call votes cast in the U.S. House. More polarization means less uncertainty since there are fewer swing voters.
4. The volatility of the national polling average. Volatility tends to predict itself, so a stable polling average tends to remain stable.
5. The overall volume of national polling. More polling means less uncertainty.
6. The magnitude of the difference between the polling-based national snapshot and the fundamentals forecast. A wider gap means more uncertainty.
7. The standard deviation of the component variables used in the FiveThirtyEight economic index. More economic volatility means more overall uncertainty in the forecast.
8. The volume of major news, as measured by the number of full-width New York Times headlines in the past 500 days, with more recent days weighted more heavily. More news means more uncertainty.

`Source` FiveThirtyEight

Example: How the sausage is made

The components of our uncertainty index are as follows:

1. The number of undecided voters in national polls. More undecided voters means more uncertainty.
2. The number of undecided plus third-party voters in national polls. More third-party voters means more uncertainty.
3. Polarization, as measured elsewhere in the model, is based on how far apart the parties are in roll call votes cast in the U.S. House. More polarization means less uncertainty since there are fewer swing voters.
4. The volatility of the national polling average. Volatility tends to predict itself, so a stable polling average tends to remain stable.
5. The overall volume of national polling. More polling means less uncertainty.
6. The magnitude of the difference between the polling-based national snapshot and the fundamentals forecast. A wider gap means more uncertainty.
7. The standard deviation of the component variables used in the FiveThirtyEight economic index. More economic volatility means more overall uncertainty in the forecast.
8. The volume of major news, as measured by the number of full-width New York Times headlines in the past 500 days, with more recent days weighted more heavily. More news means more uncertainty.

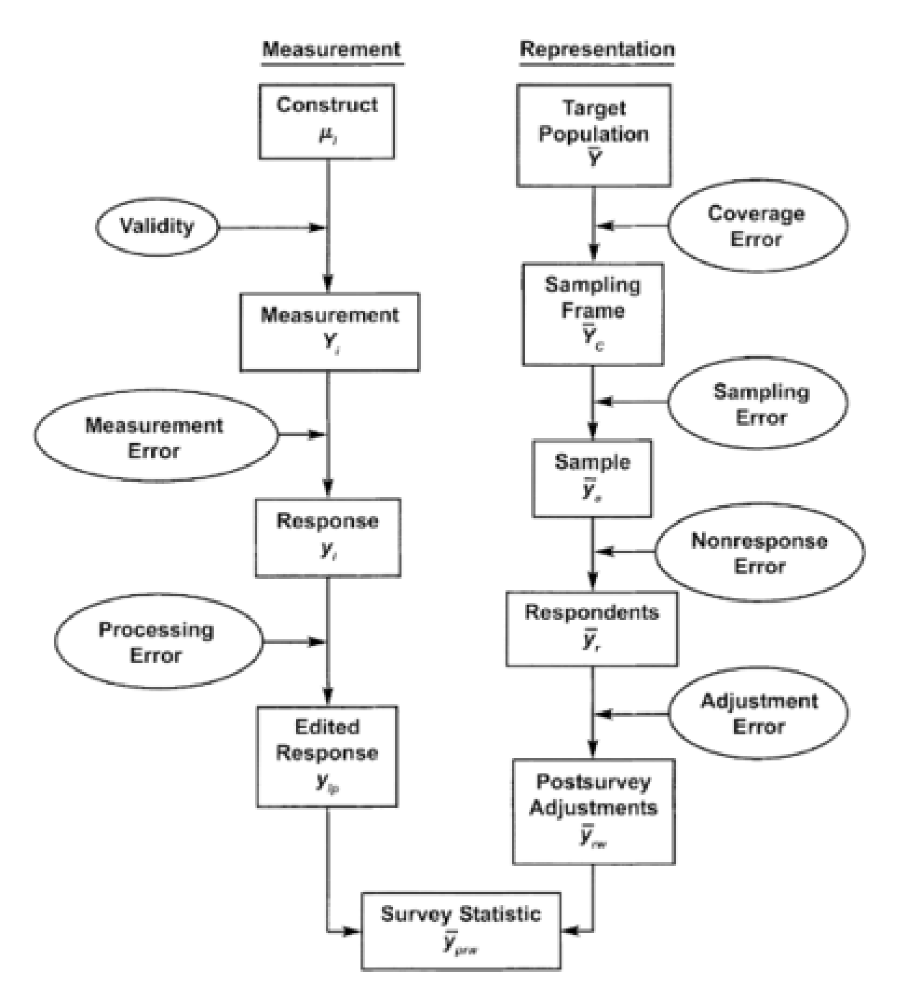
`Source` FiveThirtyEight

Uncertainty

What we are uncertain about

- **Measurement** → uncertainty in single variables
- **Model specification** → uncertainty across multiple variables and how they connect
- **Parameter estimates** → uncertainty about bias and precision
- **Model outcomes** → uncertainty about (out-of-sample) fit
- **Generalizability** to other samples, the future

Depending on the empirical setup, various specific **sources of error** might enter (e.g., survey data, digital trace data).



`Credit` Robert Groves, Total Survey Error

Communicating uncertainty

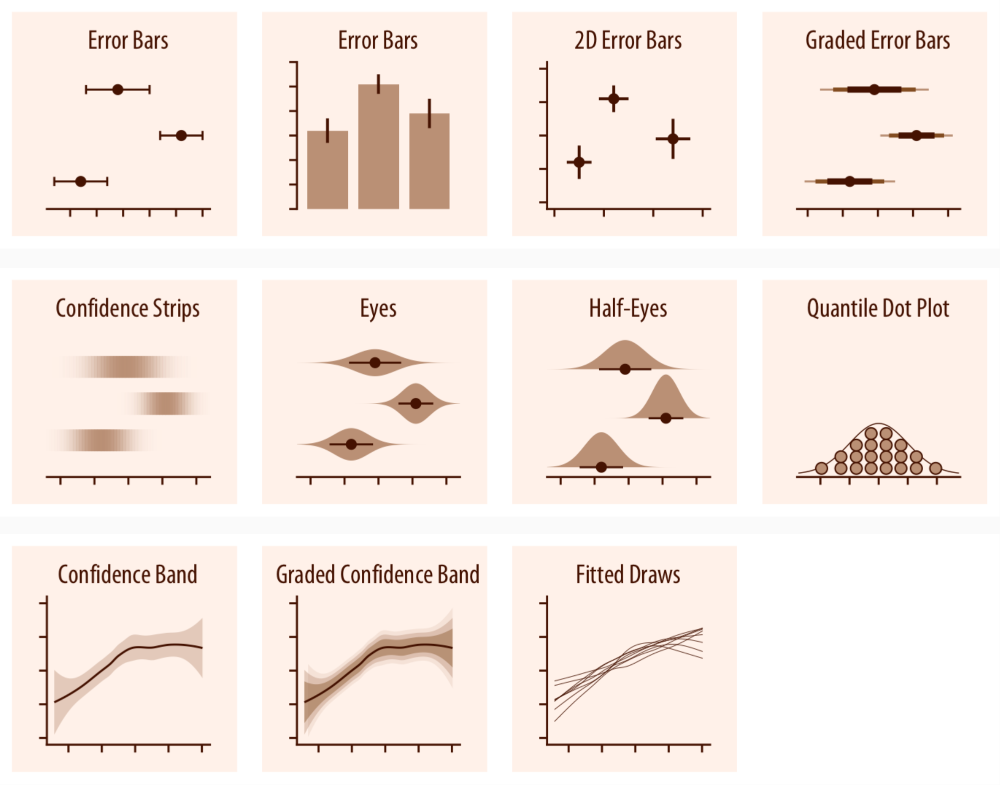
The difficulty of communicating uncertainty

- The concept is complex. Not all people think in probabilistic terms.
- Many humans are bad at understanding (conditional and unconditional) probabilities.
- Adding information about uncertainty might distract, confuse, and undermine trust.

`Credit` Richard McElreath

Communicating uncertainty (cont.)

Visualizing uncertainty



`Source` Claus Wilke

Uncertainty by numbers

```
> broom::tidy(model_out, conf.int = TRUE, conf.level = 0.95)
# A tibble: 4 × 7
  term        estimate std.error statistic   p.value conf.low conf.high
  <chr>          <dbl>     <dbl>     <dbl>     <dbl>    <dbl>     <dbl>
1 (Intercept) 13.4      0.175         76.7 0         13.1      13.8
2 distance    -0.00405  0.000110     -36.9 5.53e-297 -0.00426  -0.00383
3 originJFK   -2.70     0.189        -14.3 1.46e- 46 -3.07     -2.33
4 originLGA   -4.46     0.194        -23.0 3.04e-117 -4.84     -4.08
```

Strategies by precision

Decreasing precision

i. A full explicit probability distribution
ii. A summary of a distribution
iii. A rounded number, range or an order-of-magnitude assessment
iv. A predefined categorisation of uncertainty
v. A qualifying verbal statement
vi. A list of possibilities or scenarios
vii. Informally mentioning the existence of uncertainty
viii. No mention of uncertainty
ix. Explicit denial that uncertainty exists

`Credit` van der Bles et al. 2019

Communicating probabilities with verbal expressions

JCOM

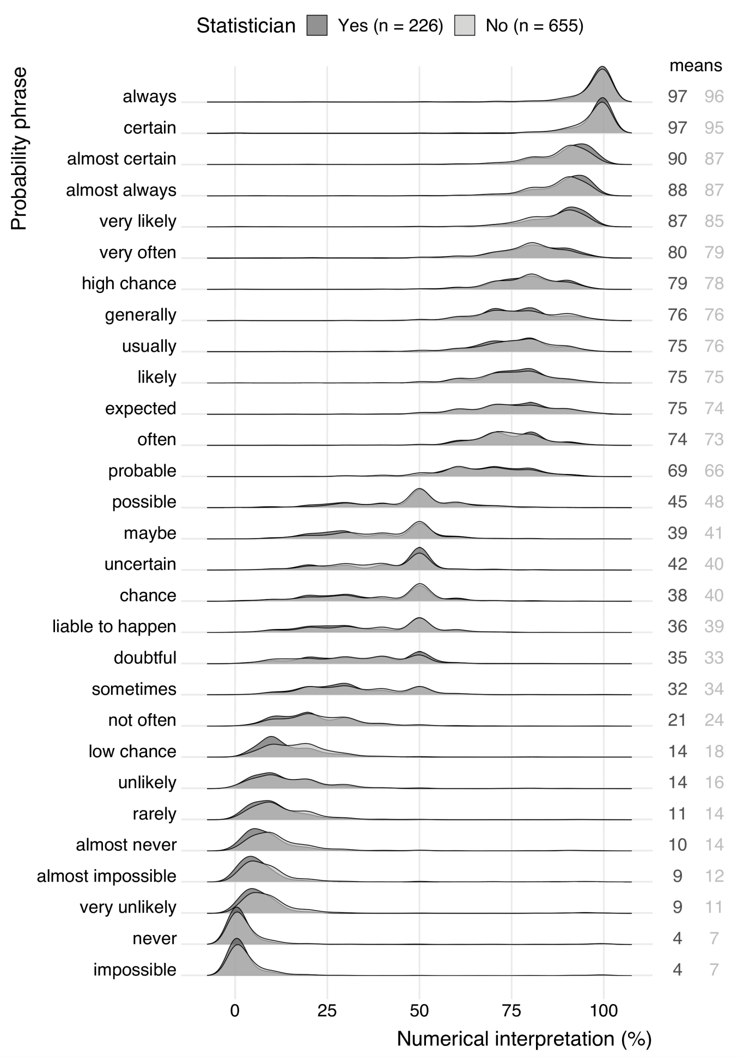
Variability in the interpretation of probability phrases used in Dutch news articles — a risk for miscommunication

Sanne Willems, Casper Albers and Ionica Smeets

Abstract Verbal probability phrases are often used in science communication to express estimated risks in words instead of numbers. In this study we look at how laypeople and statisticians interpret Dutch probability phrases that are regularly used in news articles. We found that there is a large variability in interpretations, even if the phrases are given in a neutral context. Also, statisticians do not agree on the interpretation of the phrases. We conclude that science communicators should be careful in using verbal probability expressions.

Keywords Risk communication; Science and media; Science writing

`Source` Willems et al. 2020



Probabilities: confusing vote share with p(win)

Projecting Confidence: How the Probabilistic Horse Race Confuses and Demobilizes the Public

Sean Jeremy Westwood, Dartmouth College
Solomon Messing, Acronym
Yphtach Lelkes, University of Pennsylvania

Recent years have seen a dramatic change in horse-race coverage of elections in the United States—shifting focus from late-breaking poll numbers to sophisticated meta-analytic forecasts that emphasize candidates' chance of victory. Could this shift in the political information environment affect election outcomes? We use experiments to show that forecasting increases certainty about an election's outcome, confuses many, and decreases turnout. Furthermore, we show that election forecasting has become prominent in the media, particularly in outlets with liberal audiences, and show that such coverage tends to more strongly affect the candidate who is ahead—raising questions about whether they contributed to Trump's victory over Clinton in 2016. We bring empirical evidence to this question, using American National Election Studies data to show that Democrats and Independents expressed unusual confidence in a decisive 2016 election outcome—and that the same measure of confidence is associated with lower reported turnout.

I don't know how we'll ever calculate how many people thought it was in the bag, because the percentages kept being thrown at people—"Oh, she has an 88% chance to win!"
—Hillary Clinton quoted in Traister (2017)

`Source` Westwood et al. 2020

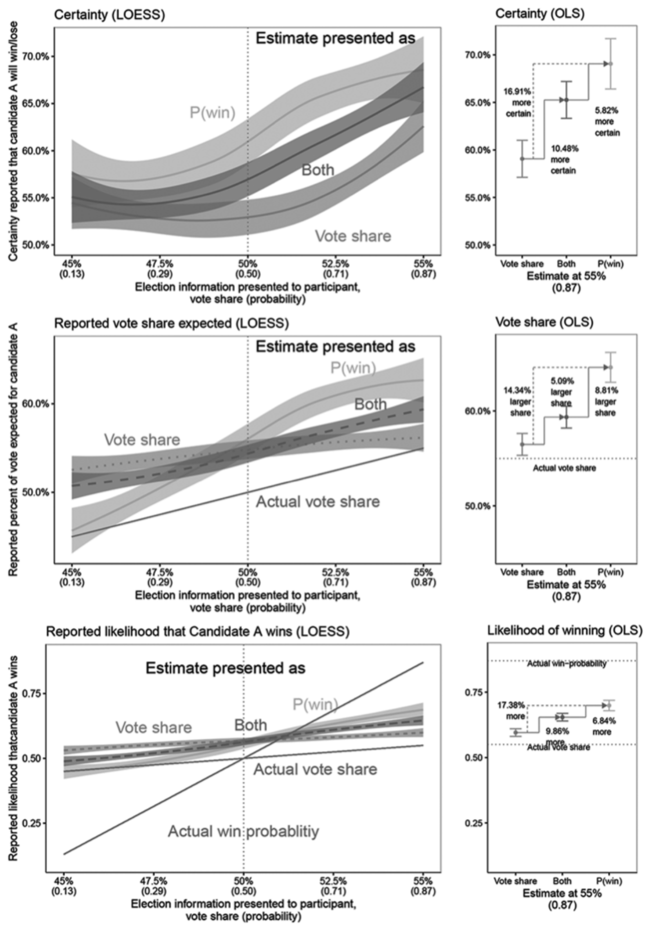


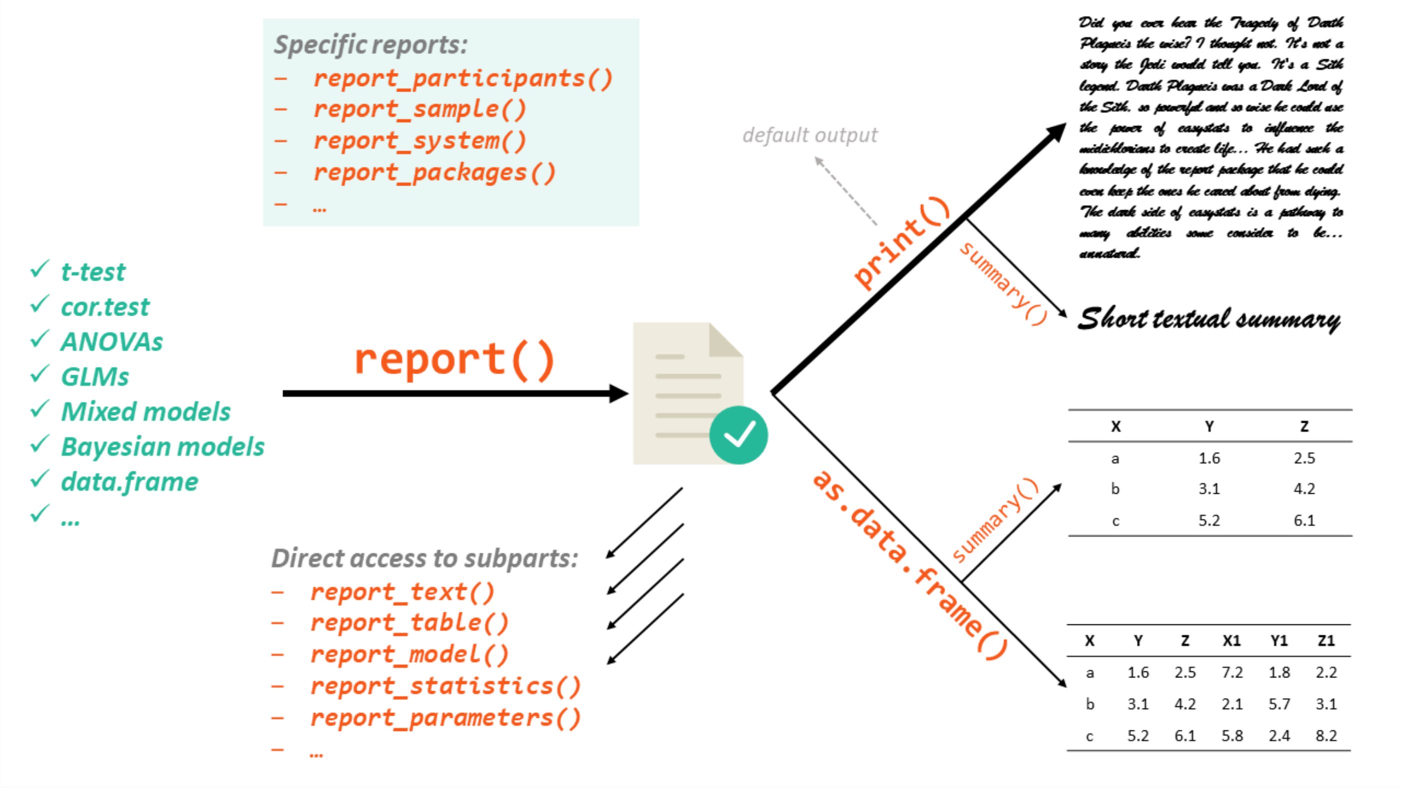
Figure 5. Effects of probabilistic forecasts on perceptions of an election. Probabilistic forecasts create the impression that the leading candidate will win more decisively, with higher certainty in judgments about which candidate will win, particularly for the leading candidate (*top*) and more extreme judgments of anticipated vote share (*bottom*), even when accompanied by vote share projections ("both" condition). Participants are less accurate when attempting to judge the likelihood of winning (*middle*) than vote share (*top*). Plots on the right show differences when vote share is fixed at 55% (.87 probability). Lines fit using LOESS in plots on the left; results based on OLS regression in plots on the right, 95% confidence bands/intervals shown. Color version available as an online enhancement.

Communicating data and analyses with report

The package

- The `report` package (part of the easystats project) provides verbal reports of models, tests, and data frames.
- In doing so, it helps ensure standardization in reporting.

The workflow



Communicating data and analyses with report

The package

- The `report` package (part of the easystats project) provides verbal reports of models, tests, and data frames.
- In doing so, it helps ensure standardization in reporting.

Use with care

- Fully automating this part of communication is probably not a good idea (do you enjoy talking to a bot?).
- The output is pseudo-objective (reporting some quantities but not others).

Example

```
R> library(report)
R> model <- lm(Sepal.Length ~ Species, data = iris)
R> report(model)
```

```
We fitted a linear model (estimated using OLS) to predict Sepal.Length
with Species (formula: Sepal.Length ~ Species). The model explains a
statistically significant and substantial proportion of variance
(R2 = 0.62, F(2, 147) = 119.26, p < .001, adj. R2 = 0.61). The model's
intercept, corresponding to Species = setosa, is at 5.01 (95% CI
[4.86, 5.15], t(147) = 68.76, p < .001). Within this model:

  - The effect of Species [versicolor] is statistically significant
  and positive (beta = 0.93, 95% CI [0.73, 1.13], t(147) = 9.03,
  p < .001; Std. beta = 1.12, 95% CI [0.88, 1.37])
  - The effect of Species [virginica] is statistically significant
  and positive (beta = 1.58, 95% CI [1.38, 1.79], t(147) = 15.37,
  p < .001; Std. beta = 1.91, 95% CI [1.66, 2.16])

Standardized parameters were obtained by fitting the model on a
standardized version of the dataset.
```

Written communication with R Markdown

Written communication

When to communicate in writing

- For communicating to **the public and decision makers**, who want to focus on the conclusions, not the code behind the analysis.

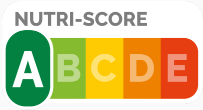
- For collaborating with **other data scientists**, who are interested in both your conclusions and how you reached them (i.e. the code).

Authoring as part of the workflow

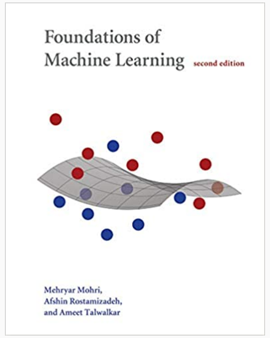
- Many different formats, including reports, briefs, blog posts, books, presentations, …
- Form follows function: the write-up tool should talk to the analytic toolset.

`Source` Kurt Newman

The continuum of written data science communication

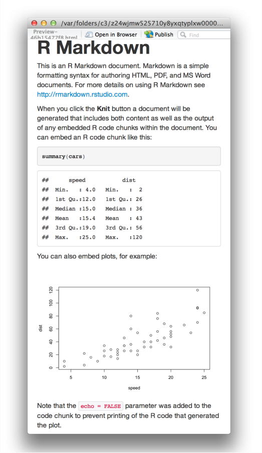


Book



Conference proceeding, journal article

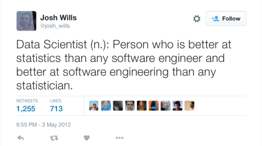
Technical report



Executive summary

Dashboard

Tweet



Authoring with R Markdown

What you already know

- R Markdown (and the `rmarkdown` package) helps you create dynamic analysis documents that combine code, rendered output (such as figures), and prose.
- You can use it to
  - Do data science interactively with notebooks.
  - Modify the layout of your report.
  - Communicate your results with others.
- You take care of content, R Markdown of format.

More resources

- The official website
- The R Markdown Cookbook
- R Markdown - The Definitive Guide

Authoring with R Markdown

What you already know

- R Markdown (and the `rmarkdown` package) helps you create dynamic analysis documents that combine code, rendered output (such as figures), and prose.
- You can use it to
  - Do data science interactively with notebooks.
  - Modify the layout of your report.
  - Communicate your results with others.
- You take care of content, R Markdown of format.

More resources

- The official website
- The R Markdown Cookbook
- R Markdown - The Definitive Guide

What you probably don't know (yet)

- R Markdown can do much more than reports. You can use it to author
  - Reports (in PDF, HTML, Word, etc.)
  - Interactive documents
  - Dashboards
  - Slideshows
  - Books
  - Websites
- It stands on the shoulders of Pandoc, a program that converts markup files into virtually any other format.

RMarkdown formats

bookdown

- A package that facilitates writing books and long-form articles/reports with R Markdown.
- See here for an overview of books written with `bookdown`.

blogdown

- A package that lets you create websites (not only blogs!) using R Markdown.
- It integrates Hugo (or other site generators).

pagedown

- A package that lets you paginate the HTML output of R Markdown with CSS for print (PDF).
- Lots of different templates available.

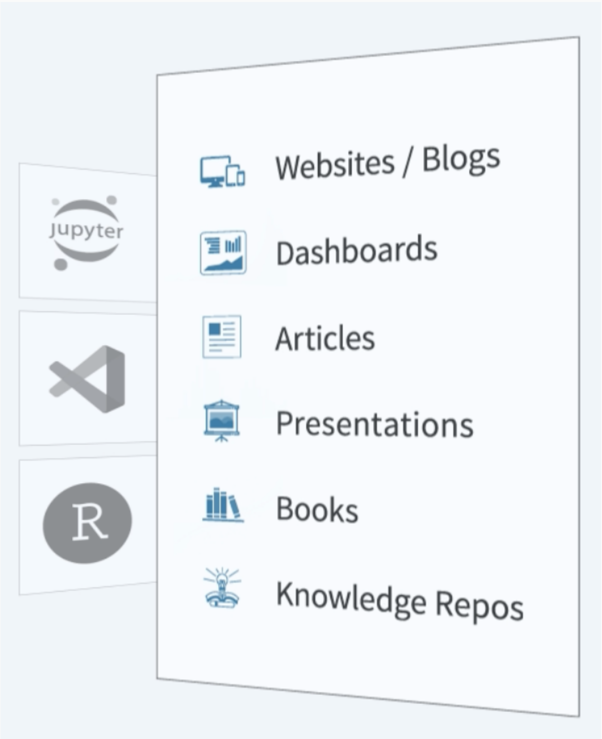
xaringan

- A package that lets you crate slideshows with remark.js through R Markdown.
- These slides have been created using this package.

The next generation of technical publishing: Quarto

What's Quarto?

- Quarto is software developed at Posit, the company behind RStudio.
- It's the "next generation of R Markdown" and also built on Pandoc. If you know R Markdown well, you already know Quarto well.
- It facilitates embedding code and output from R, Python, Julia, and other languages.
- It combines the functionality of R Markdown and all the other tools (bookdown, xaringan, etc.) into one single consistent system.
- Quarto is still fairly new and under active development. More and more extensions are coming out that increase the flexibility of the suite.
- Check out the comprehensive guide to learn more.

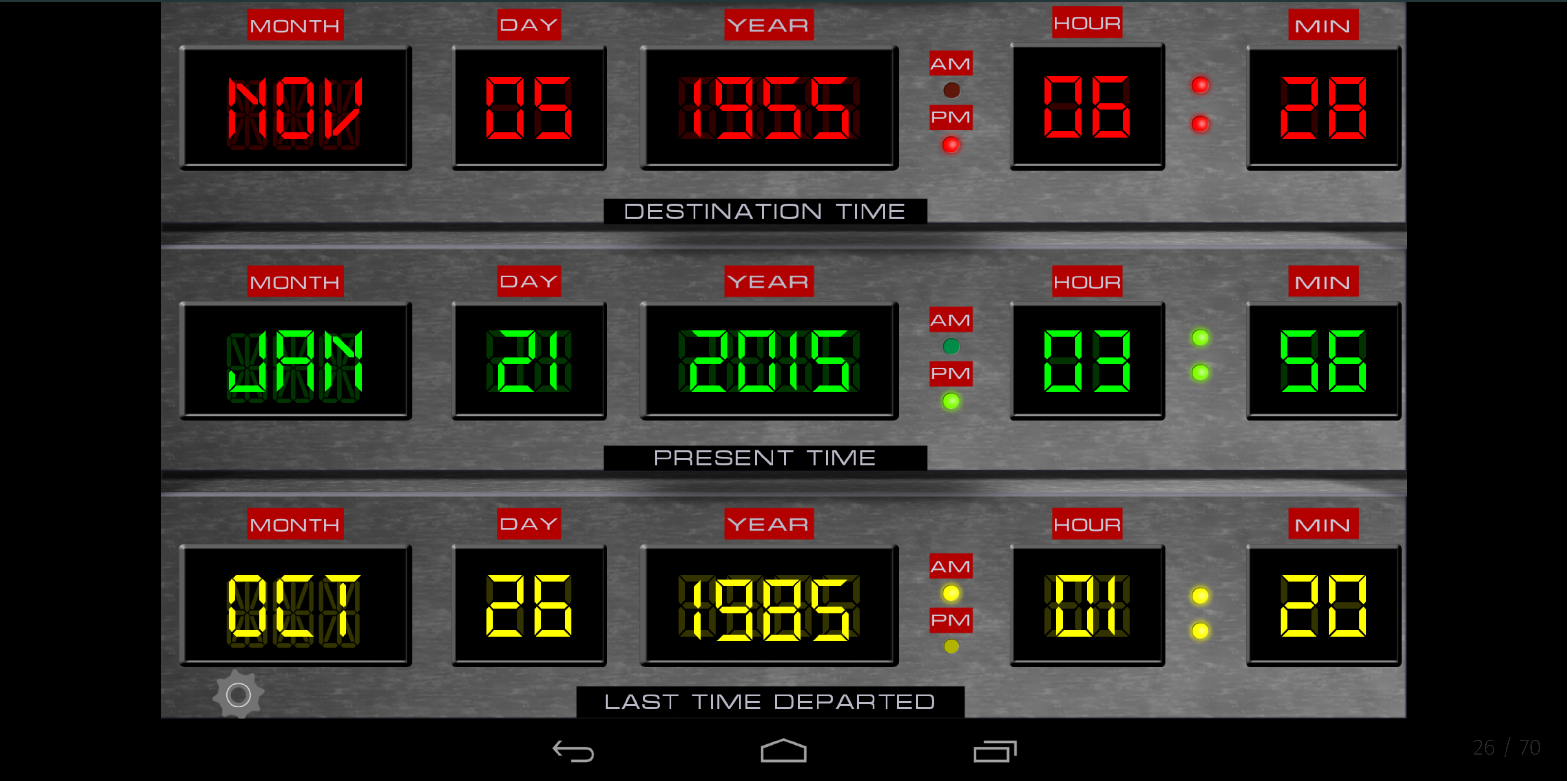


Quarto and Jupyter

- Quarto's support for both Knitr and Jupyter means that you can use it to create documentation projects that contain content from both systems.
- Something to keep in mind next semester and going forward! :)

Interactive communication with dashboards

What are dashboards?

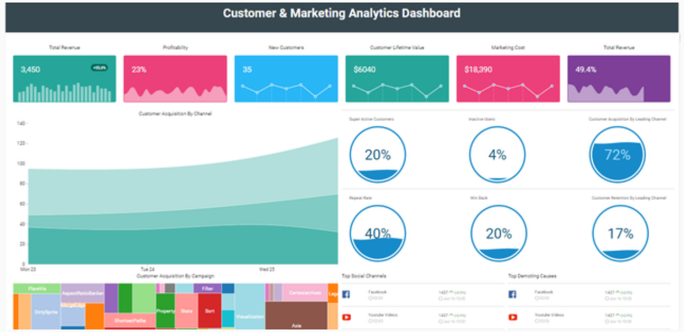


What are dashboards?

Yeah, what are they really?

- A (business or data) dashboard is a GUI that provides high-level overviews of performance indicators or other quantities of interest.
- It's a **monitoring** (and not so much analysis) **tool**.
- Think of dashboards as a **mash-up of data visualization and report**.
- Dashboards are increasingly popular in businesses and organizations to **synthesize data points** from operative units (for strategic and analytical purposes).
- Data journalism has started to embrace dashboards in the context of elections, the COVID-19 pandemic, and sports.
- **Common features** are:
  - Accessibility via web browser
  - Featuring of interactives
  - Heavy focus on comparative visualization
  - Provision of trends on key performance indicators (KPIs)

`Credit` Tim Green

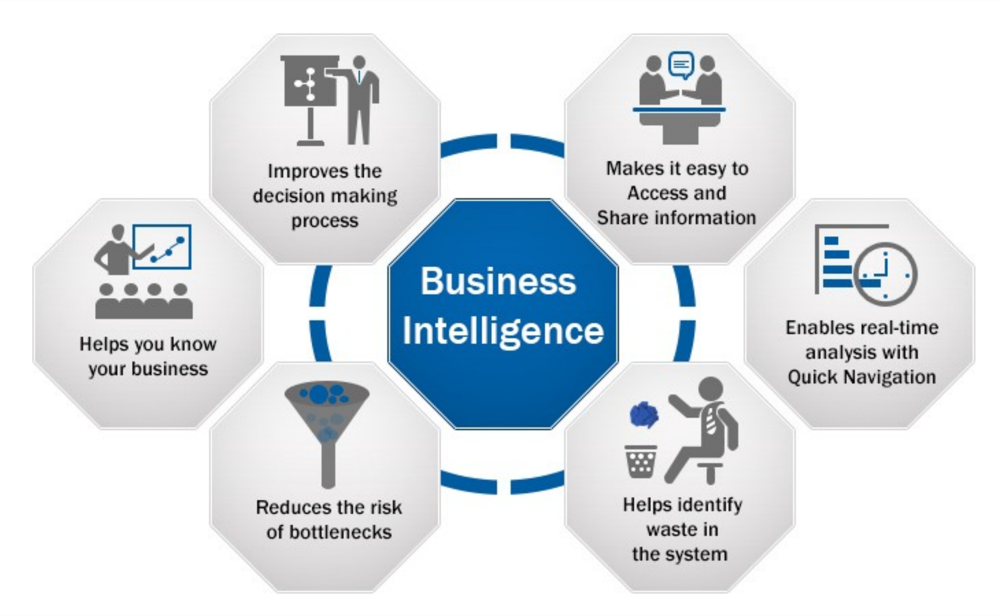


`Credit` HelicalInsight OpenSourceBI

Why are dashboards?

Why are they a thing?

- There is increasing **abundance of data** (often process-generated) that cannot speak for itself.
- If used wisely, these data can provide an **important part of business intelligence** and a basis for high-level evidence-based decision-making.
- Provide continuous quantification of indicators of interest (→ **monitoring**).
- **Reduce information differential** between analysts and stakeholders.
- Also, **measuring the health of organizations** can help stay in control (if only as a performative act) and satisfy managers' need for micromanaging.



`Credit` towardsdatascience.com

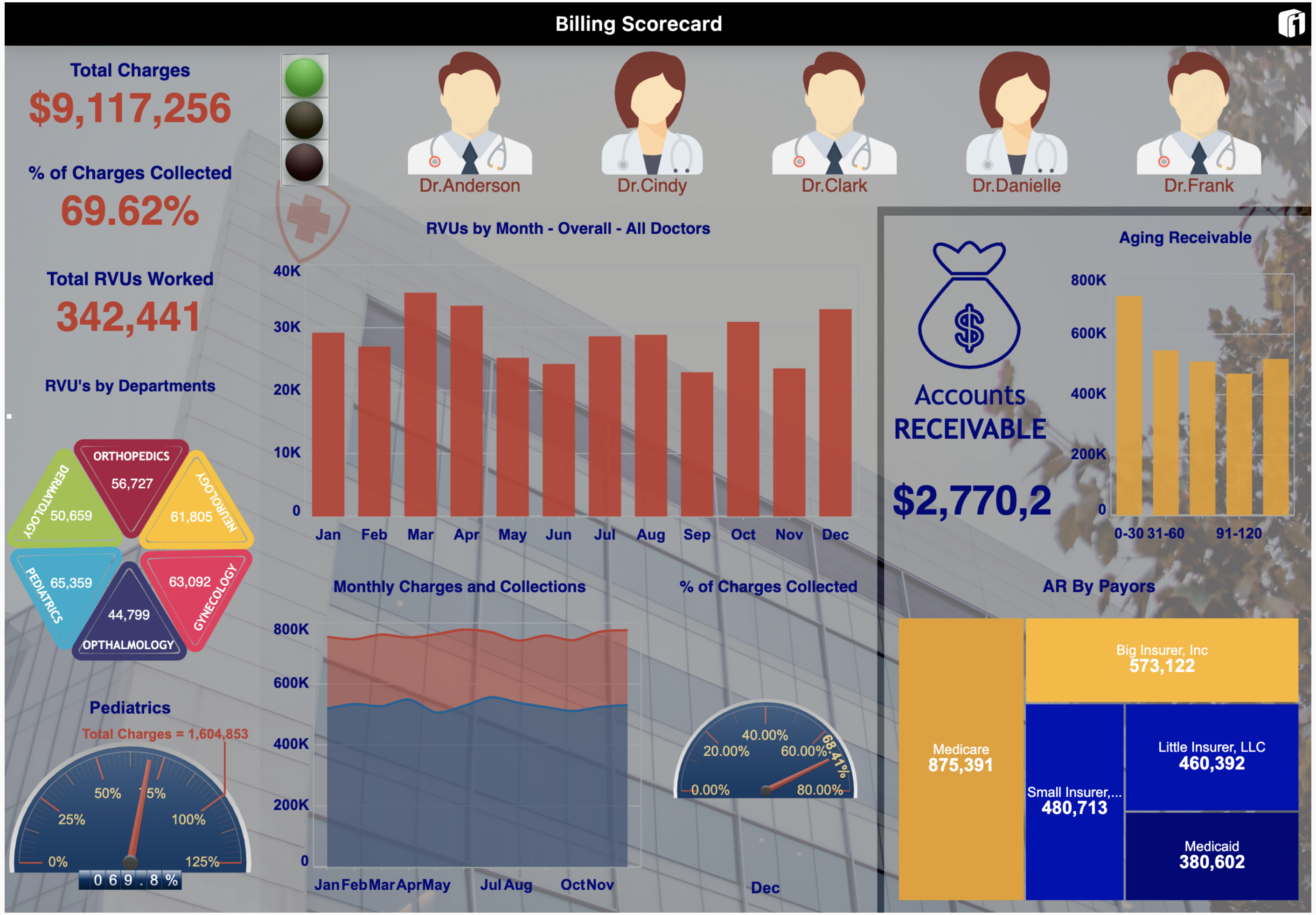
Dashboards in the wild

Credit geckoboard.com

Dashboards in the wild

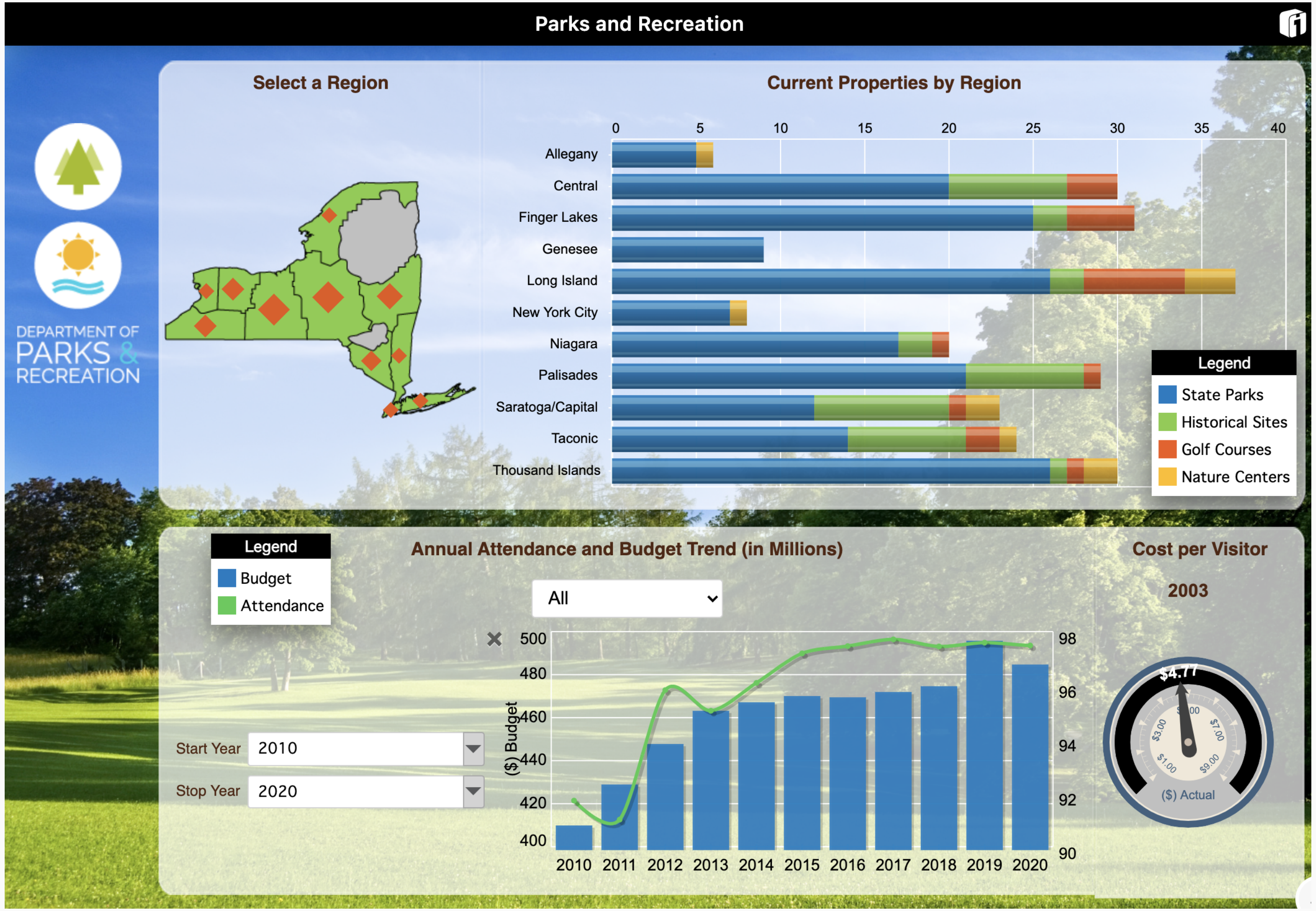
`Credit` geckoboard.com

Dashboards in the wild



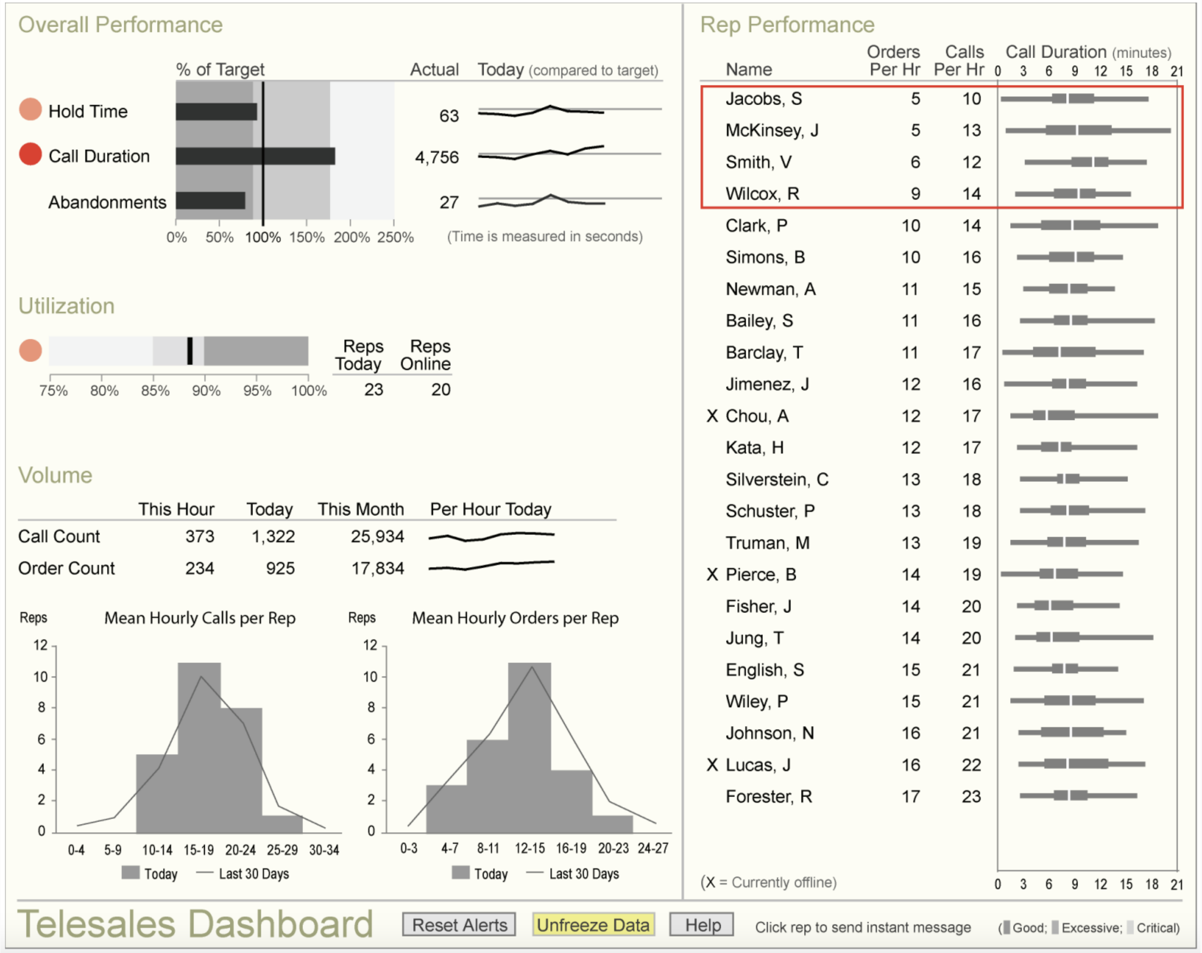
Credit idashboards.com

Dashboards in the wild



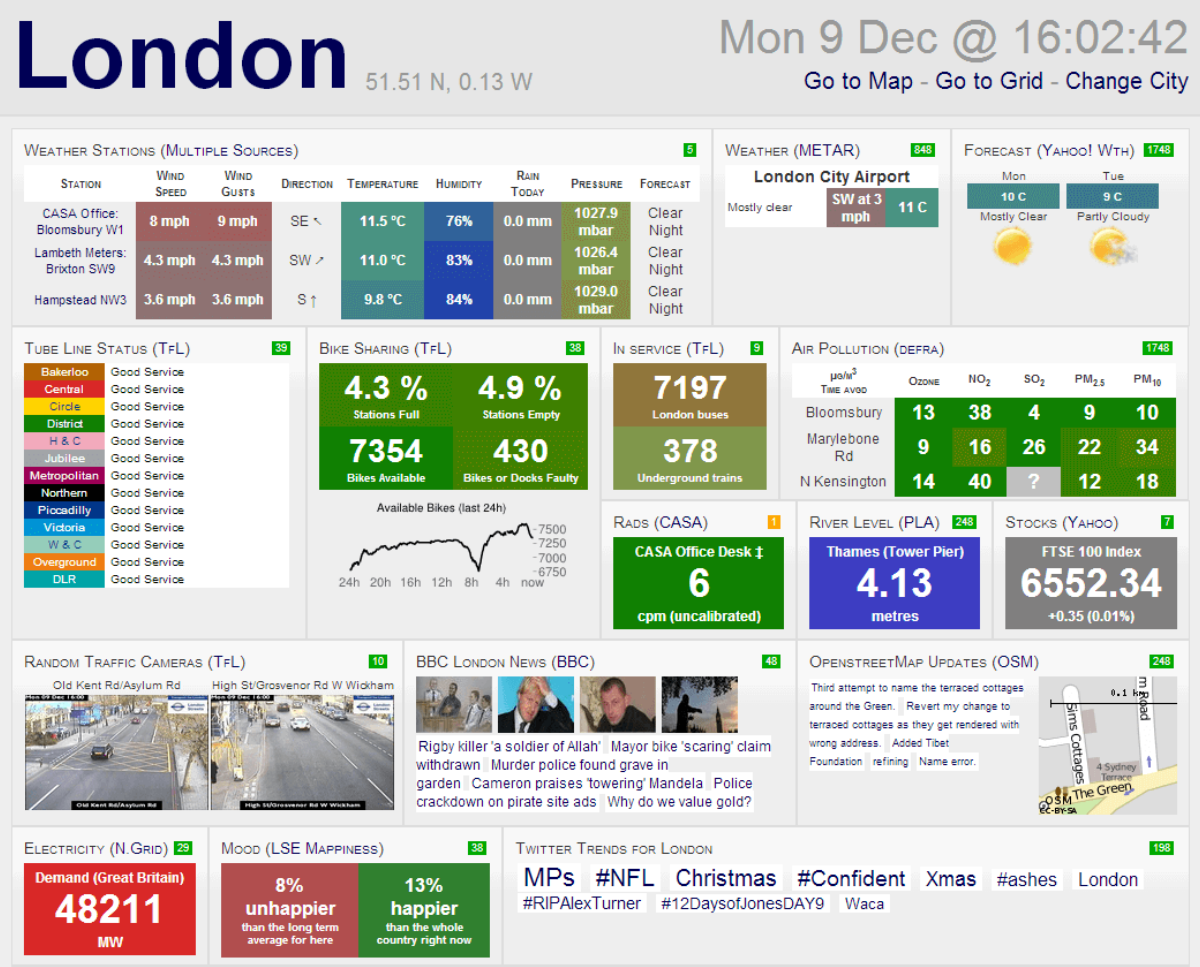
Credit idashboards.com

Dashboards in the wild



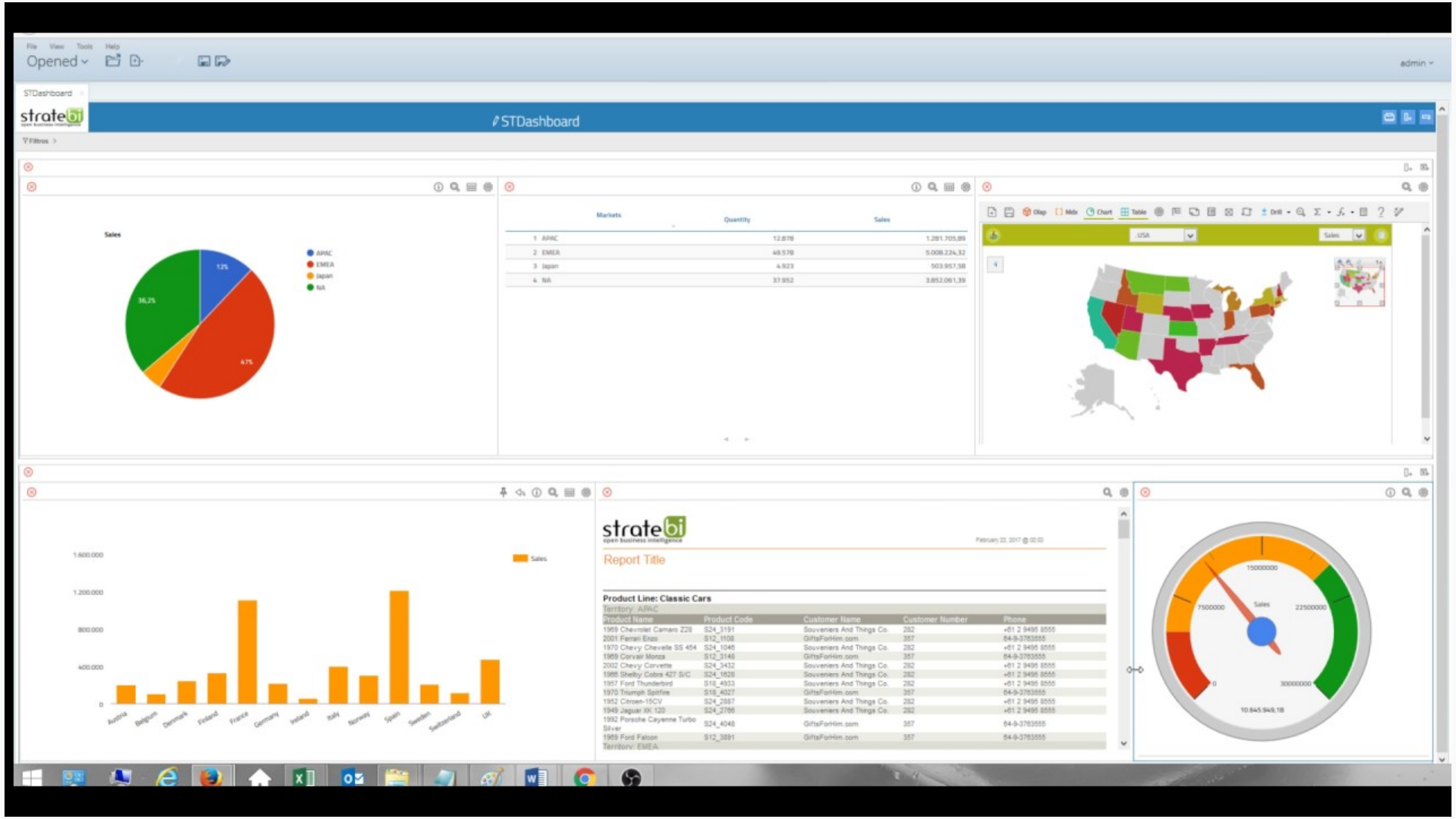
`Credit` Stephen Few

Dashboards in the wild



Credit idashboards.com

Dashboards in the wild



`Credit` carmel.es

The problem(s) with dashboards

Design challenges

- **They say too little.** Loss of information is fatal for good decision-making when aggregating results into few KPIs.
- **They say too much** (irrelevant things).
- Dashboards often fail not in technology but in communication (rooted in poor design).
- "Dashboards are not for show. No amount of cuteness and technical wizardry can substitute for clear communication." Stephen Few, Perceptual Edge
- Dashboards are a subgenre of data viz, so **all rules of good/bad viz apply**.
- So, there is hope since we do know a bit about how to design good visuals. (See here for a nice case study on improving the design of a dashboard.)

The problem(s) with dashboards

Design challenges

- **They say too little.** Loss of information is fatal for good decision-making when aggregating results into few KPIs.
- **They say too much** (irrelevant things).
- Dashboards often fail not in technology but in communication (rooted in poor design).
- "Dashboards are not for show. No amount of cuteness and technical wizardry can substitute for clear communication." Stephen Few, Perceptual Edge
- Dashboards are a subgenre of data viz, so **all rules of good/bad viz apply**.
- So, there is hope since we do know a bit about how to design good visuals. (See here for a nice case study on improving the design of a dashboard.)

Analytic challenges

- Dashboards cater to the desire to be able to make good decisions on the basis of few selected metrics.
- This logic reflects a **gross simplification of reality**.
- All challenges that pop up in careful analytic work - issues of selection, measurement, causality, predictiveness - are still valid but will be obscured when aggregating data.
- Simple metrics can still be useful, but often **you need contextual knowledge** (which is difficult to communicate in dashboards).
- Another consequence of "dashboarding" business intelligence can be that by making decisions a function of metrics, they stop working well because they will be gamed.

Thoughtful dashboard design and usage

Checklist before you start[1]

1. Are you tackling a monitoring task that needs your data/metrics to be updated frequently?
2. Who will use the dashboard and to what end? What questions will they use it to answer? What actions will they take in response to these answers?
3. What specific information should be displayed, and is it meaningful without much context?
4. What could lead to the metrics being wrong/misleading?

[1] Source: Stephen Few/Perceptual Edge

Thoughtful dashboard design and usage

Checklist before you start[1]

1. Are you tackling a monitoring task that needs your data/metrics to be updated frequently?

2. Who will use the dashboard and to what end? What questions will they use it to answer? What actions will they take in response to these answers?

3. What specific information should be displayed, and is it meaningful without much context?

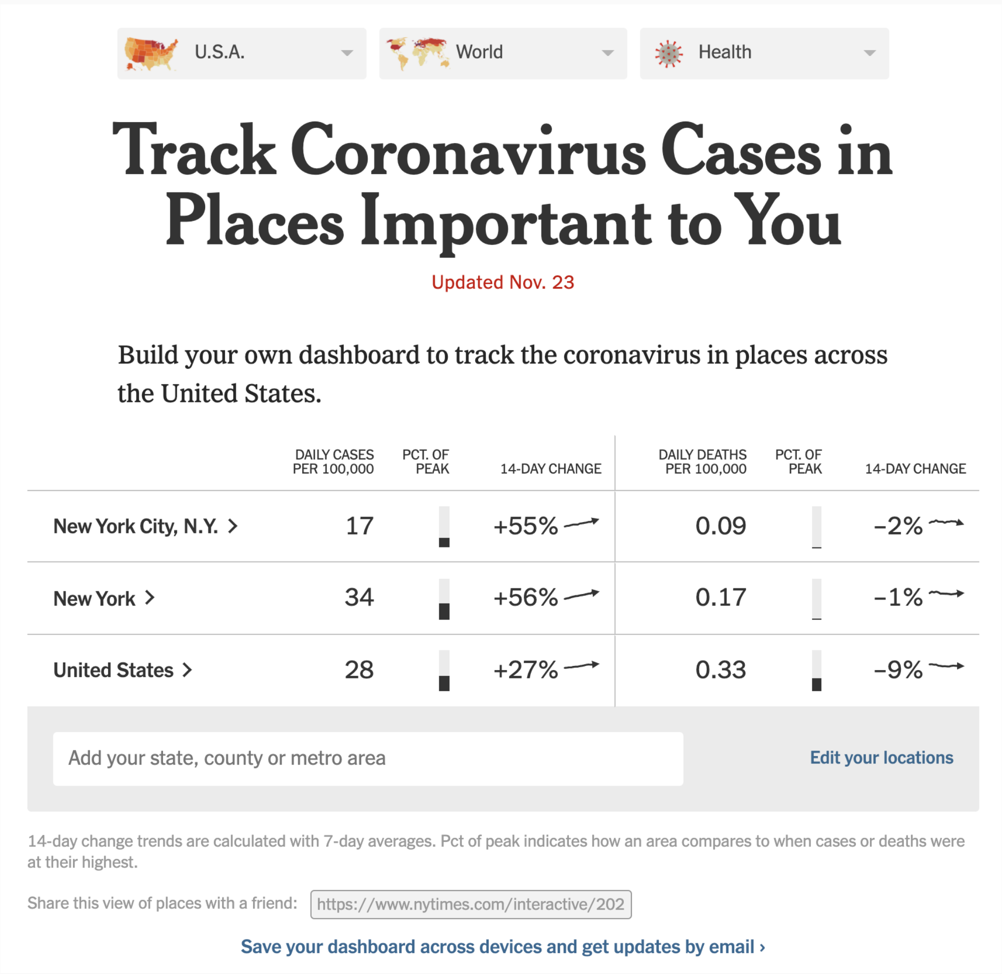
4. What could lead to the metrics being wrong/misleading?

Design advice

- Minimize distractions.
- Focus on meaningful quantities of interest, not the ones that look cool.
- Don't overload with information.
- Apply all rules of good data viz.
- Use interactives with care (e.g., to make optional content conditionally visible)
- Try not to exceed the boundaries of a single screen.
- Ensure desktop/mobile screen responsiveness.

[1]Source: Stephen Few/Perceptual Edge

Dashboards in the wild: COVID-19 edition


U.S.A. | World | Health

Track Coronavirus Cases in Places Important to You

Updated Nov. 23

Build your own dashboard to track the coronavirus in places across the United States.

	DAILY CASES PER 100,000	PCT. OF PEAK	14-DAY CHANGE	DAILY DEATHS PER 100,000	PCT. OF PEAK	14-DAY CHANGE
New York City, N.Y. ›	17		+55%	0.09		-2%
New York ›	34		+56%	0.17		-1%
United States ›	28		+27%	0.33		-9%

Add your state, county or metro area | Edit your locations

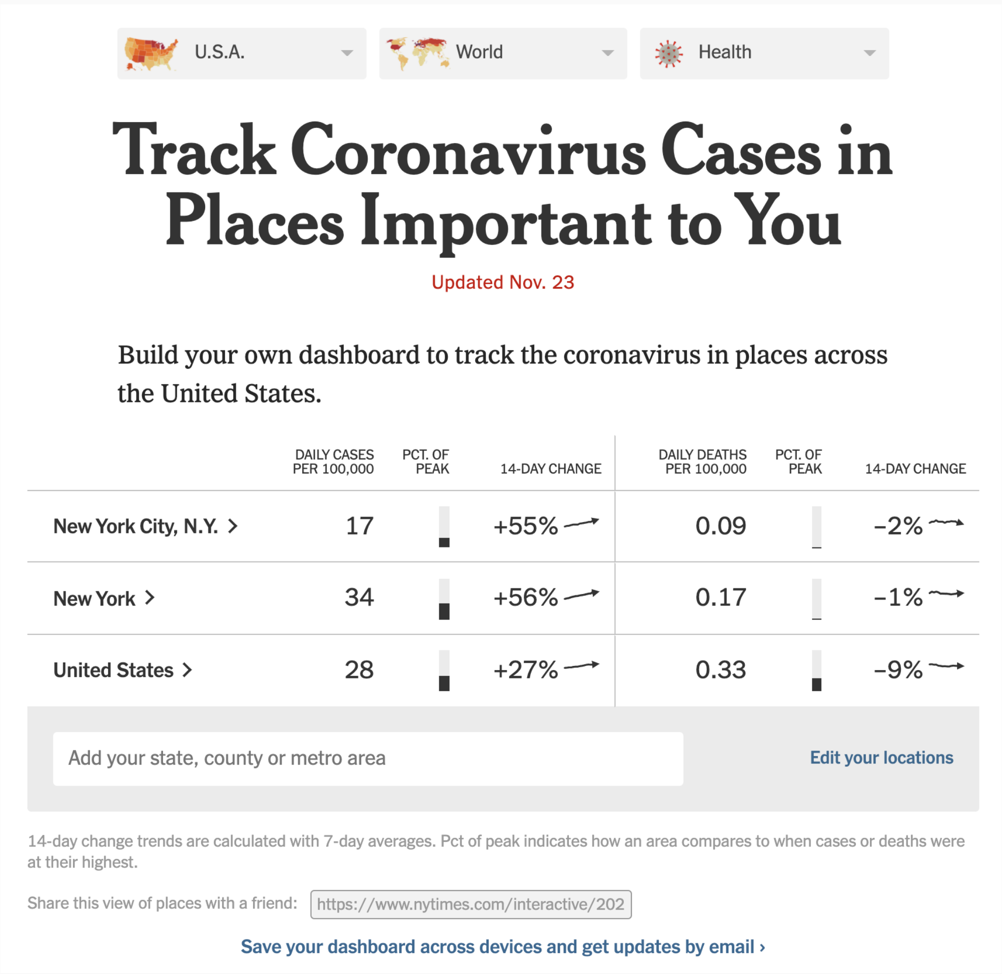
14-day change trends are calculated with 7-day averages. Pct of peak indicates how an area compares to when cases or deaths were at their highest.

Share this view of places with a friend: https://www.nytimes.com/interactive/202

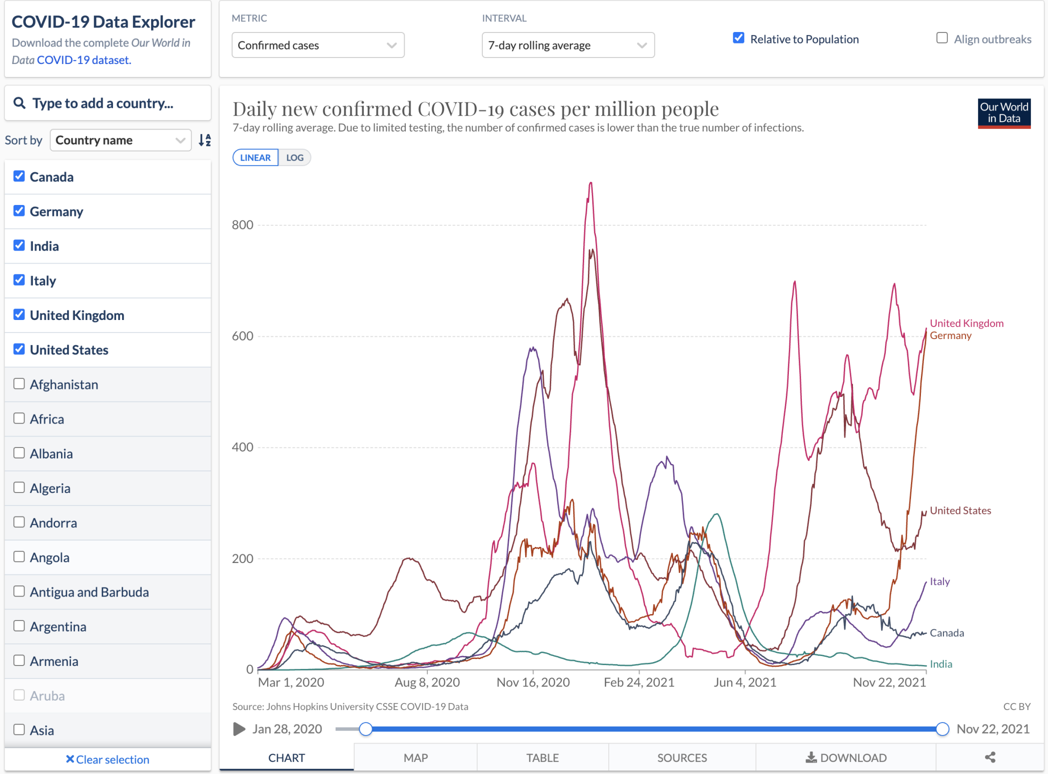
Save your dashboard across devices and get updates by email ›

`Credit` NY Times

Dashboards in the wild: COVID-19 edition

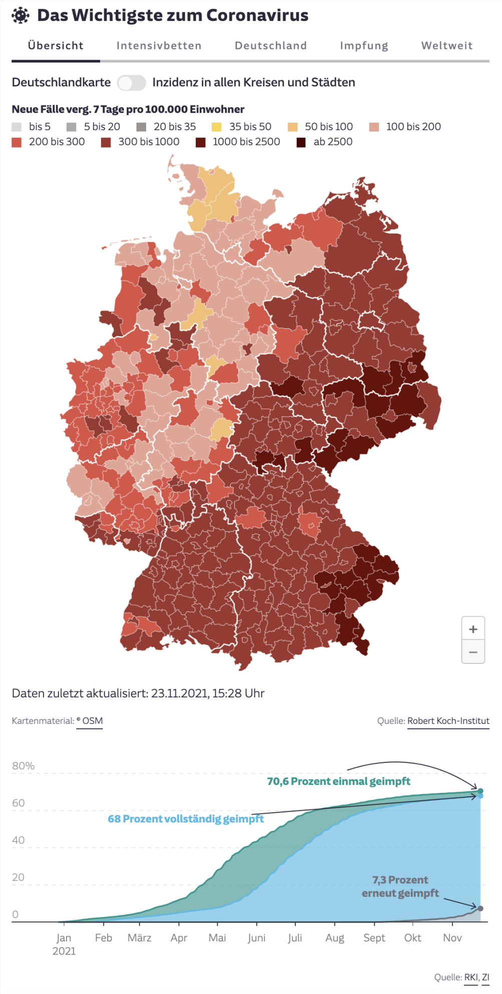


Credit NY Times



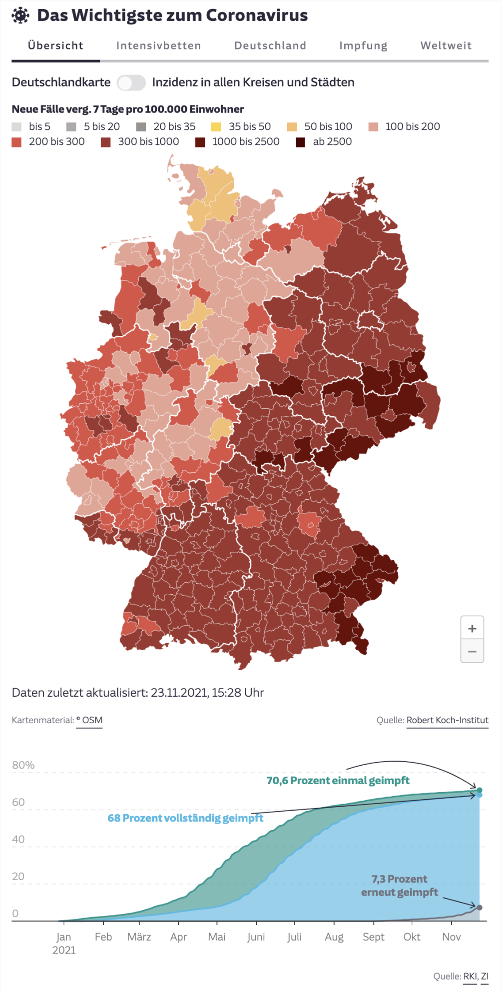
Credit Our World in Data

Dashboards in the wild: COVID-19 edition

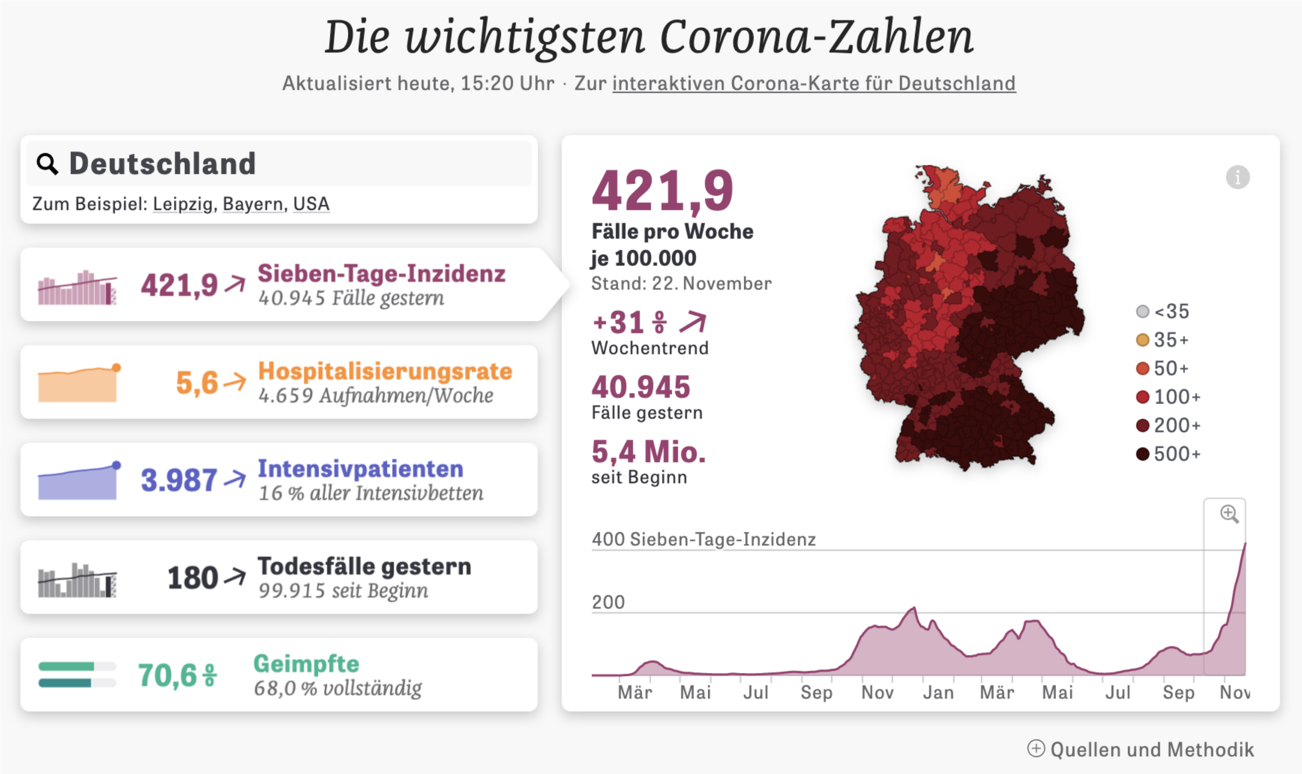


`Credit` SZ Online

Dashboards in the wild: COVID-19 edition

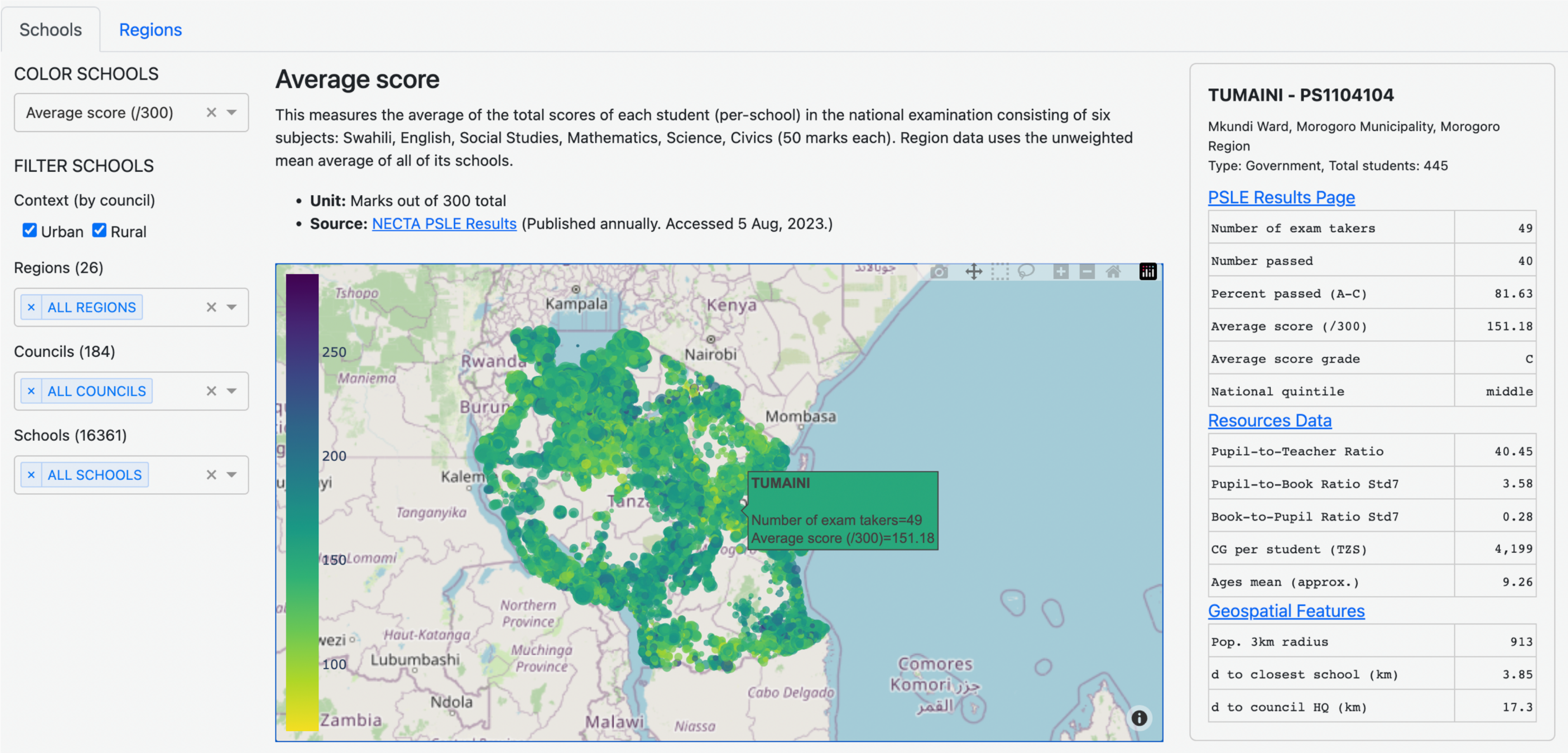


Credit SZ Online



Credit ZEIT Online

Dashboards: another example

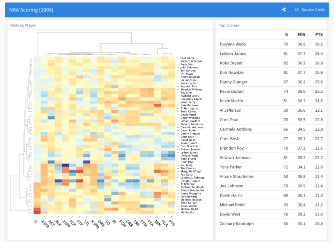


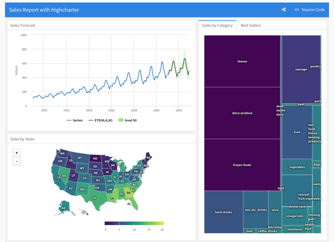
`Source` Lonny Chen, NECTA PSLE Dashboard 2022 | https://bit.ly/psle2022mvp

Dashboards with R

`flexdashboard` package

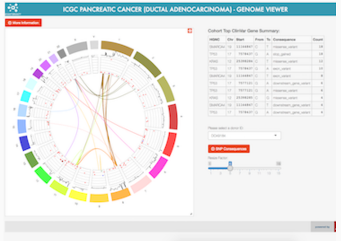
- Overview here.
- Good for easy dashboard building
- Just a document that looks like a dashboard
- Can be compiled into a static file (just like regular Markdown)
- Can only run interactive code client-fleside (in embedded JavaScript)
- Shiny and `htmlwidgets` (`leaflet`, `plotly`, `highcharter`, etc.) can be integrated (with all the up- and downsides)





`shiny` package

- Overview here.
- More complex to program, but the best option for complex apps.
- Can implement any layout.
- Needs a server behind it to execute R code on user input.
- Can run interactive code either by processing serverside (in R) or clientside (in embedded JavaScript).
- The shinydashboard package provides another way to create dashboards with Shiny.



Dashboards with `flexdashboard`

Functionality

- Use simple R Markdown to build a dashboard.

```
---
title: "NBA Scoring (2008)"
output:
  flexdashboard::flex_dashboard:
    orientation: rows
    social: menu
    source_code: embed
---

```{r setup, include=FALSE}
library(knitr)
library(d3heatmap)
library(flexdashboard)

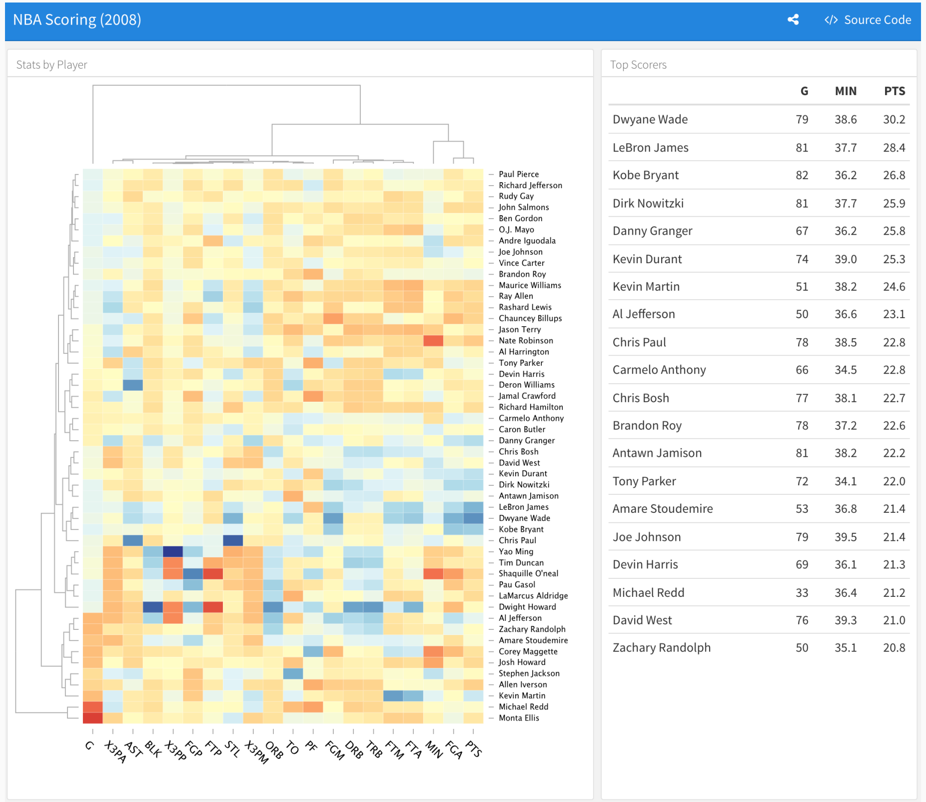
url <- "http://datasets.flowingdata.com/ppg2008.csv"
nba_players <- read.csv(url, row.names = 1)
```

Stats by Player {data-width=650}

```{r}
d3heatmap(nba_players, scale = "column")
```

Top Scorers {data-width=350}

```{r}
knitr::kable(nba_players[1:20,c("G", "MIN", "PTS")])
```
```



Source: jjallaire

Dashboards with flexdashboard

Functionality

- Use simple R Markdown to build a dashboard.
- Arrange panels as blocks with flexible syntax.

Layout by Column

By default, level 2 markdown headers (`------------------`) within dashboards define columns, with individual charts stacked vertically within each column. Here's the definition of a two column dashboard with one chart on the left and two on the right:

```
---
title: "Column Orientation"
output: flexdashboard::flex_dashboard
---

Column
-------------------------------------

### Chart 1

```{r}
```

Column

Chart 2

```{r}
```

Chart 3

```{r}
```

```

# Dashboards with `flexdashboard`

## Functionality

- Use simple R Markdown to build a dashboard.
- Arrange panels as blocks with flexible syntax.

### SCROLLING LAYOUT

By default flexdashboard charts are laid out to automatically fill the height of the browser. This works well for a small number of vertically stacked charts, however if you have lots of charts you'll probably want to scroll rather than fit them all onto the page. You can control this behavior using the `vertical_layout` option. Specify `fill` to vertically re-size charts so they completely fill the page and `scroll` to layout charts at their natural height, scrolling the page if necessary.

For example, the following layout includes 3 charts and requests that the page scroll as necessary to accommodate their natural height:

````
---
title: "Chart Stack (Scrolling)"
output:
  flexdashboard::flex_dashboard:
    vertical_layout: scroll

Chart 1

```{r}
```

Chart 2

```{r}
```

Chart 3

```{r}
```
````

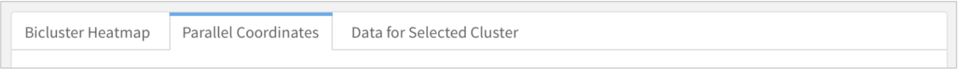
Dashboards with `flexdashboard`

Functionality

- Use simple R Markdown to build a dashboard.
- Arrange panels as blocks with flexible syntax.

TABSETS

If you have several components you'd like to display within a row or column then rather than attempting to fit them all on screen at the same time you can lay them out as a tabset. This is especially appropriate when one component is primary (i.e. should be seen by all readers) and the others provide secondary information that might be of interest to only some readers.



In many cases tabsets are a better solution than `vertical_layout: scroll` for displaying large numbers of components since they are so straightforward to navigate.

To layout a row or column as a tabset you simply add the `{.tabset}` attribute to the section heading. For example, the following code lays out the second column in tabset:

```
---
title: "Tabset Column"
output: flexdashboard::flex_dashboard
---

Column
-------------------------------------

### Chart 1

```{r}
```

Column {.tabset}

Chart 2

```{r}
```

Chart 3

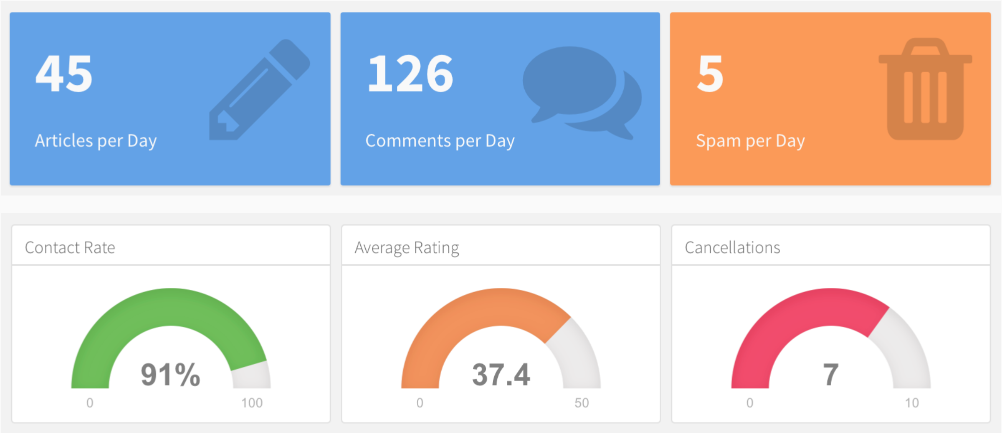
```{r}
```

```

# Dashboards with `flexdashboard`

## Functionality

- Use simple R Markdown to build a dashboard.
- Arrange panels as blocks with flexible syntax.
- Add elements like gauges and value boxes.



# Dashboards with `flexdashboard`

## Functionality

- Use simple R Markdown to build a dashboard.
- Arrange panels as blocks with flexible syntax.
- Add elements like gauges and value boxes.
- Couple it with `shiny`.

```
---
title: "Old Faithful Eruptions"
output: flexdashboard::flex_dashboard
runtime: shiny

```{r global, include=FALSE}
# load data in 'global' chunk so it can be shared by all users of the dashboard
library(datasets)
data(faithful)
```

Column {.sidebar}

Waiting time between eruptions and the duration of the eruption for the
Old Faithful geyser in Yellowstone National Park, Wyoming, USA.

```{r}
selectInput("n_breaks", label = "Number of bins:",
            choices = c(10, 20, 35, 50), selected = 20)

sliderInput("bw_adjust", label = "Bandwidth adjustment:",
            min = 0.2, max = 2, value = 1, step = 0.2)
```

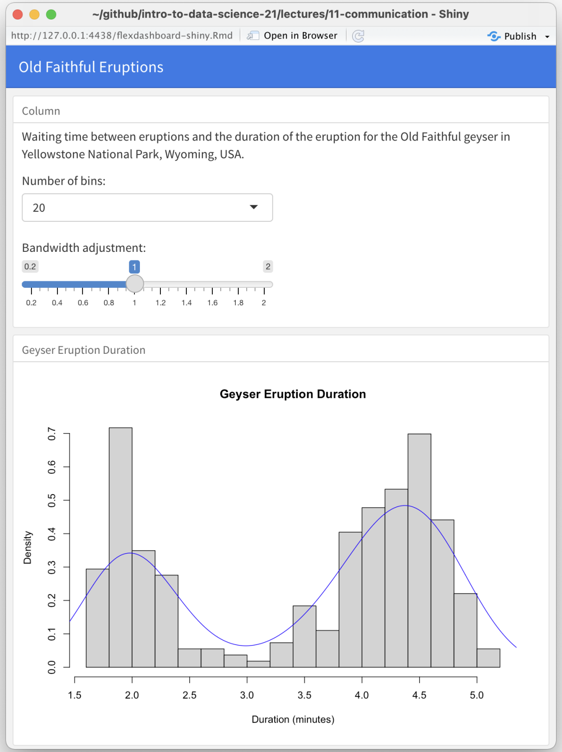
Column

Geyser Eruption Duration

```{r}
renderPlot({
  hist(faithful$eruptions, probability = TRUE, breaks = as.numeric(input$n_breaks),
       xlab = "Duration (minutes)", main = "Geyser Eruption Duration")

  dens <- density(faithful$eruptions, adjust = input$bw_adjust)
  lines(dens, col = "blue")
})
```

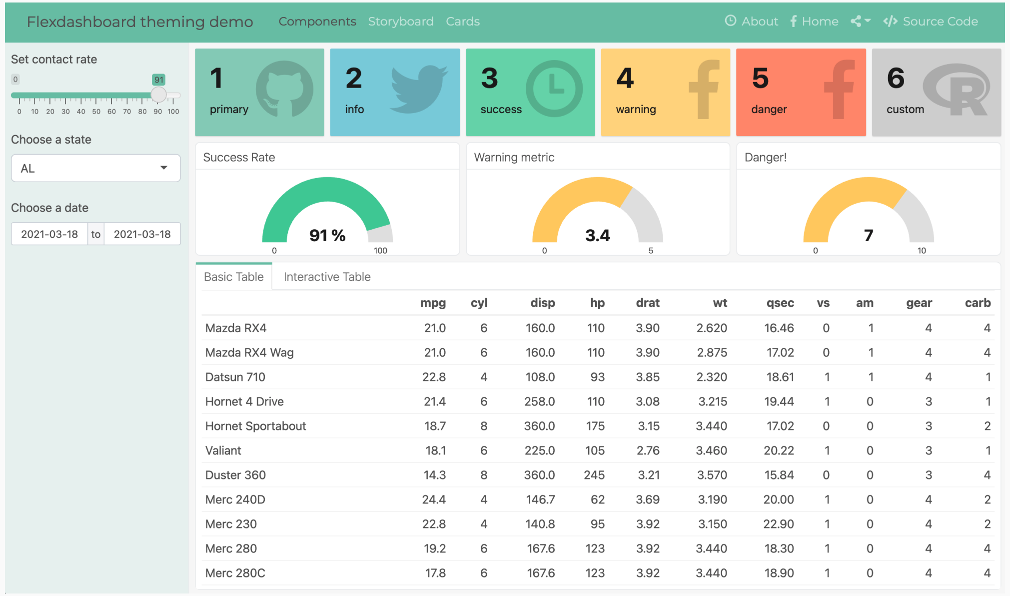
```



# Dashboards with `flexdashboard`

## Functionality

- Use simple R Markdown to build a dashboard.
- Arrange panels as blocks with flexible syntax.
- Add elements like gauges and value boxes.
- Couple it with `shiny`.
- Customize themes.



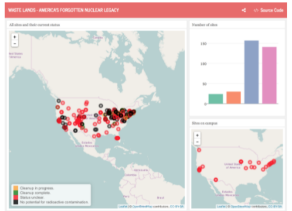
# Dashboards with `flexdashboard`

## Functionality

- Use simple R Markdown to build a dashboard.
- Arrange panels as blocks with flexible syntax.
- Add elements like gauges and value boxes.
- Couple it with `shiny`.
- Customize themes.
- Explore more examples here.

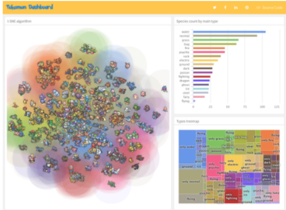

MetricsGraphics: Tor Project

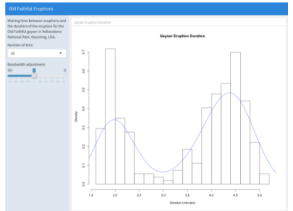

Shiny: kmeans clustering

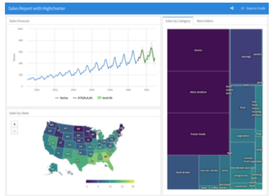

leaflet: nuclear waste sites

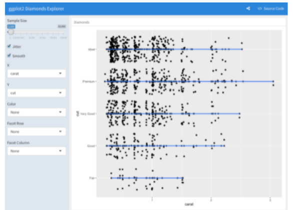

Shiny: biclust example

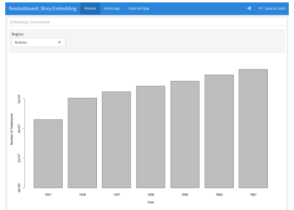

ggplotly: various examples

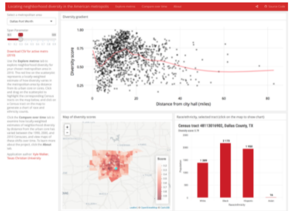

Pokemon characters with highcharter

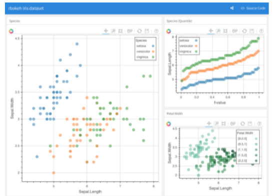

Shiny: Old faithful eruptions


Sales report with highcharter


Shiny: Diamonds explorer


Shiny: Embedding

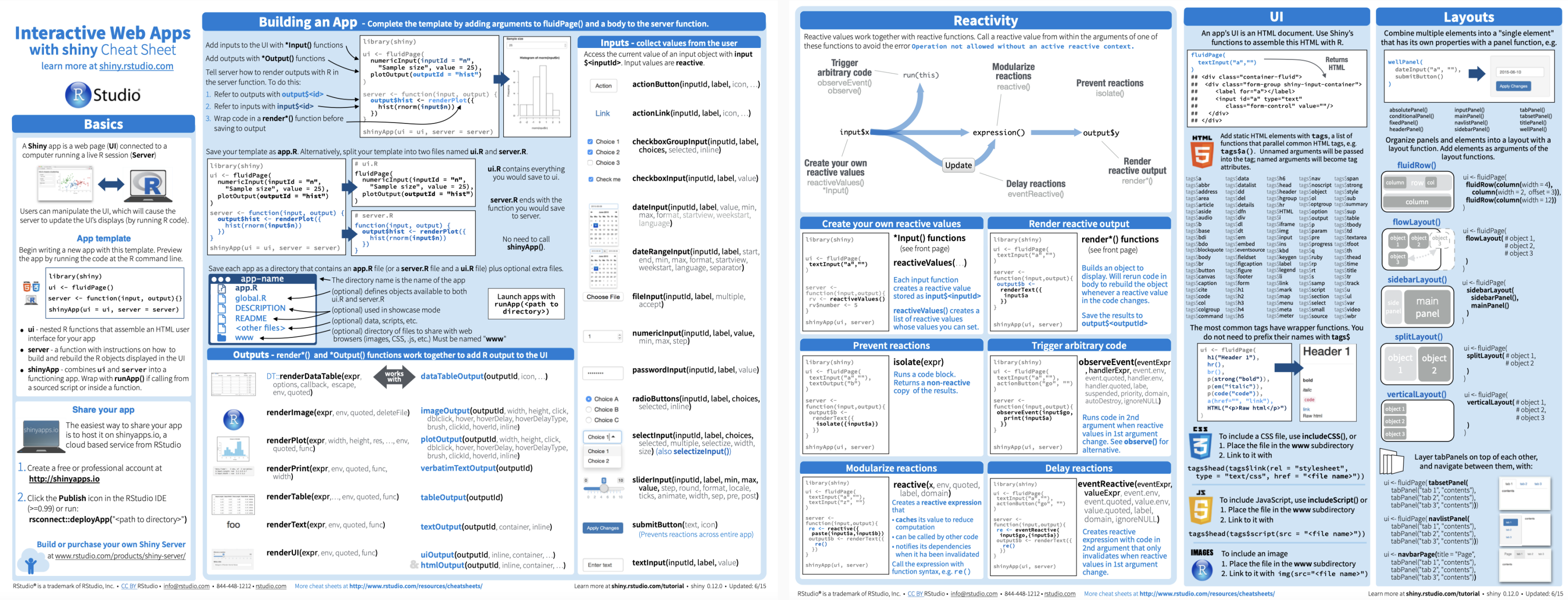

Shiny: Neighborhood diversity (Source)


rbokeh: iris dataset

# Web apps with shiny

## Functionality

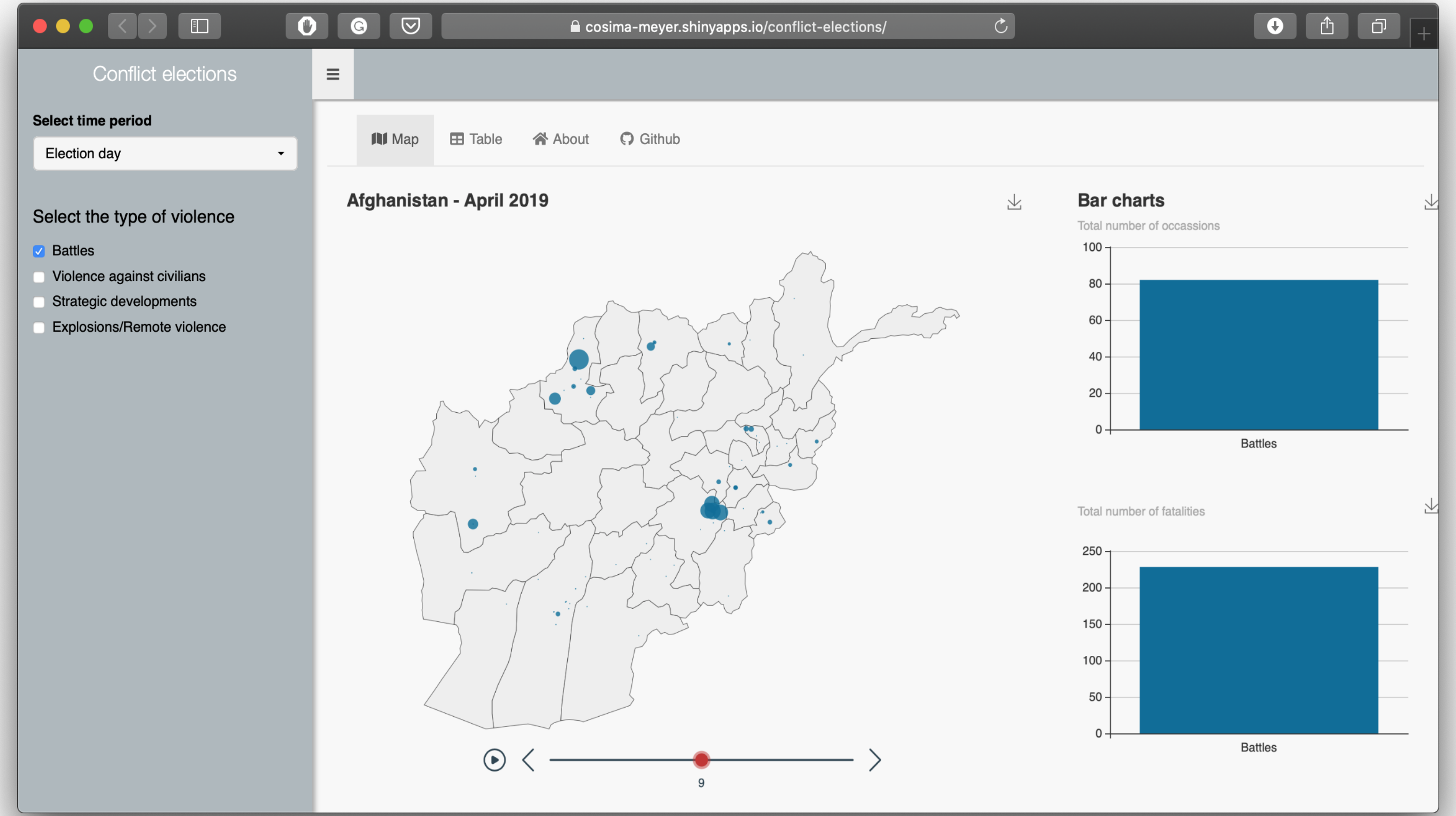
- Shiny's functionality is too complex and rich to introduce it on a couple of slides. Wait for the labs!
- It certainly can do much more than dashboards.
- Think of it as a tool to create **web apps** that allow interaction with raw and cooked data.



# Web apps with shiny

## Example applications

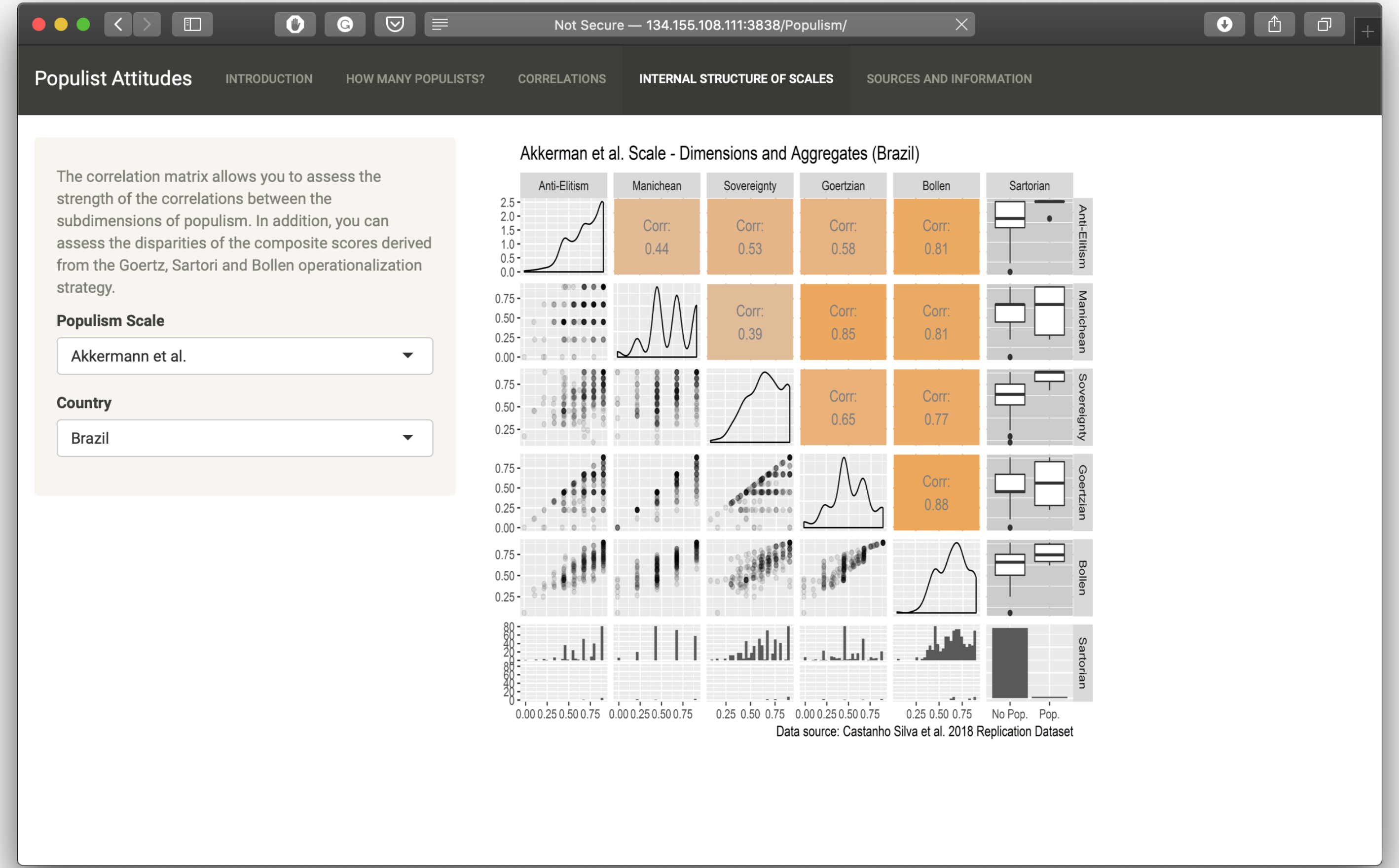
- Data explorer



# Web apps with shiny

## Example applications

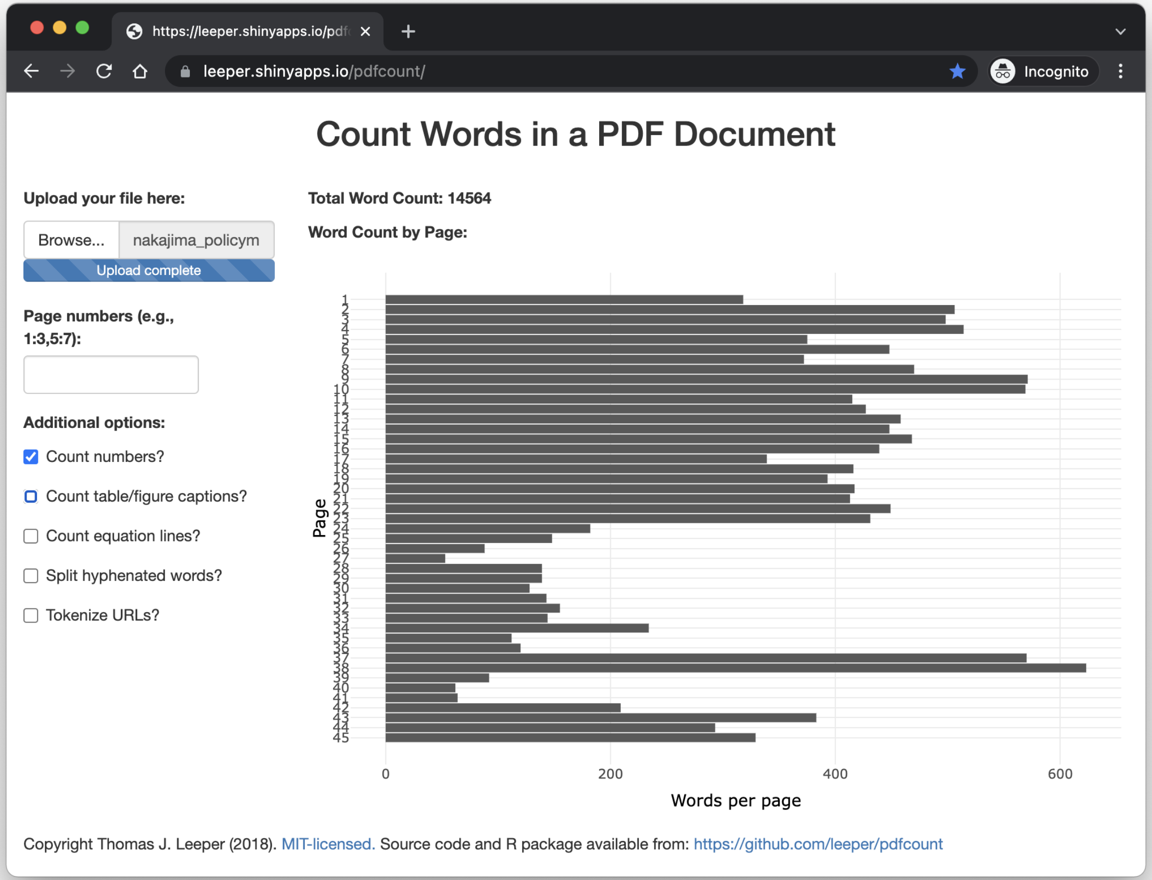
- Data explorer
- Interactive appendix



# Web apps with shiny

## Example applications

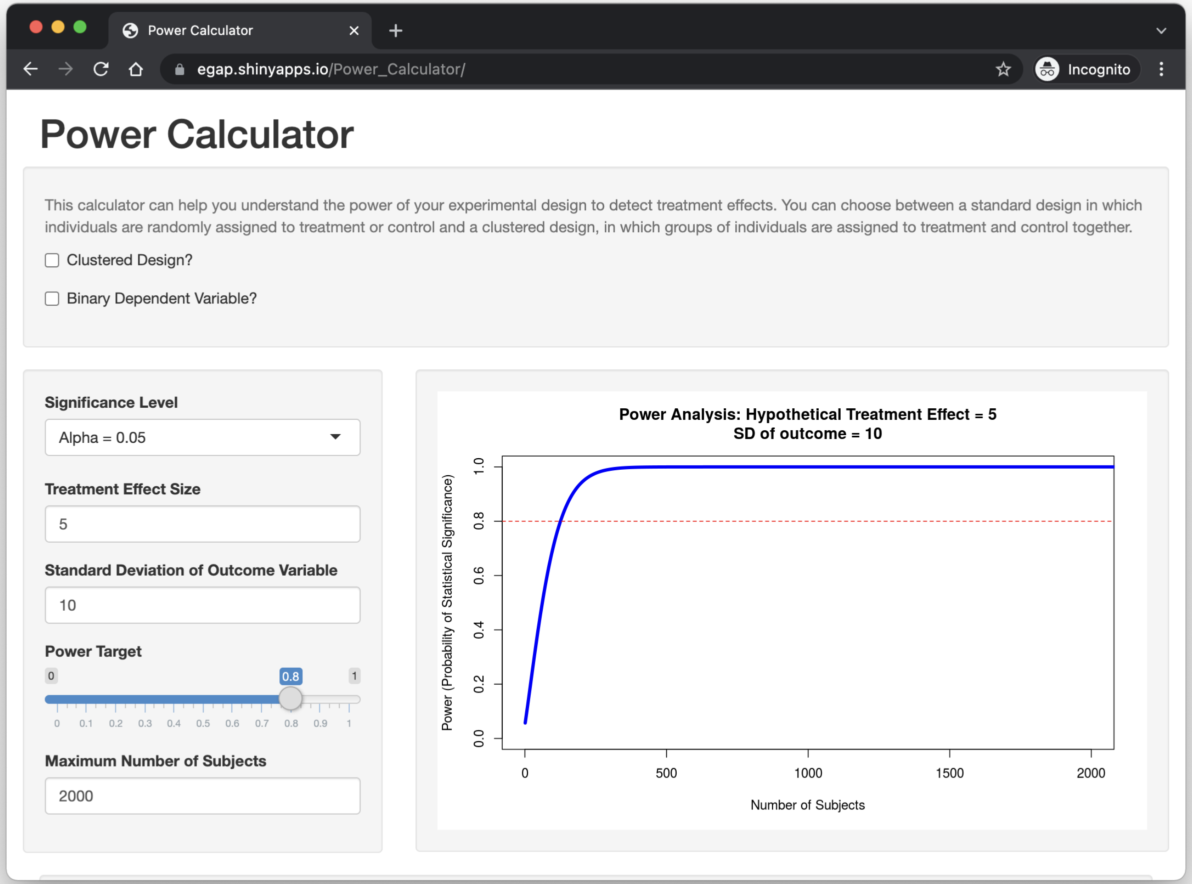
- Data explorer
- Interactive appendix
- Workflow apps



# Web apps with shiny

## Example applications

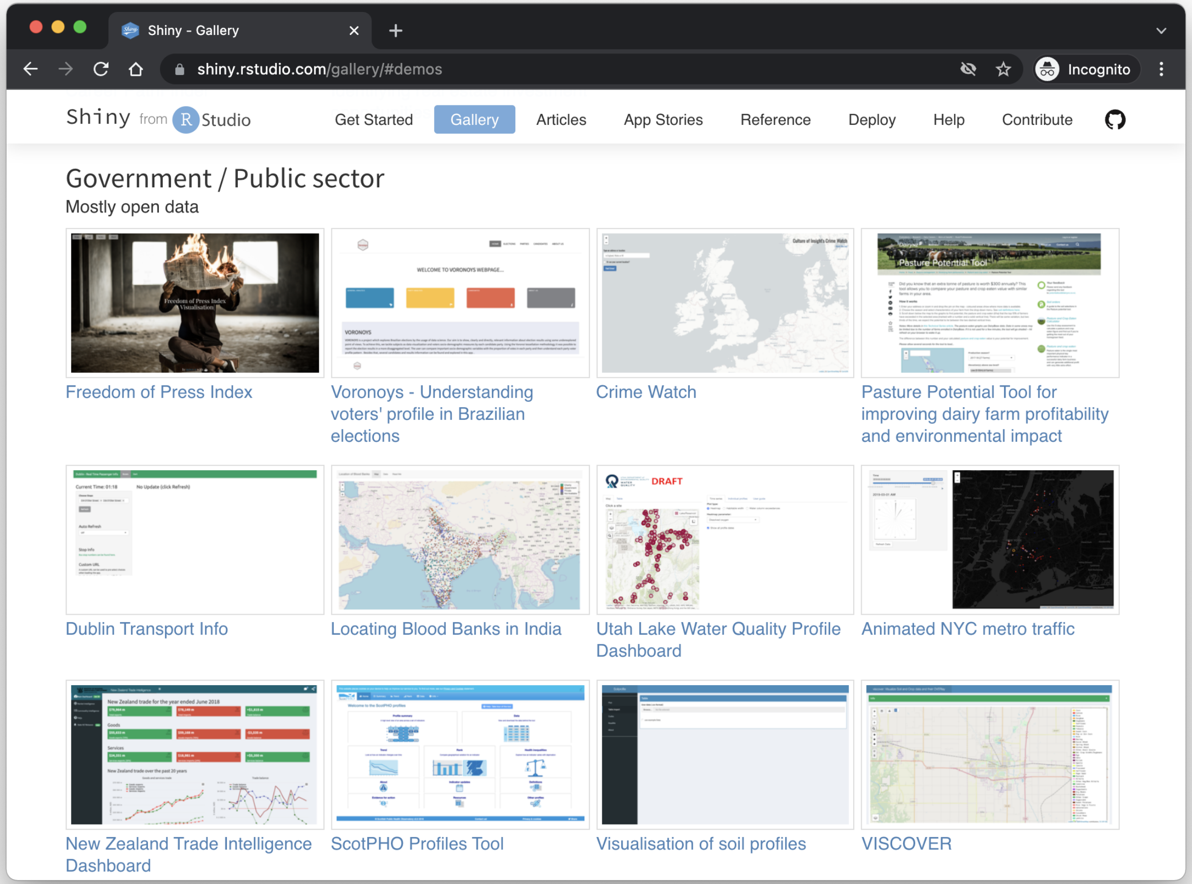
- Data explorer
- Interactive appendix
- Workflow apps
- Learning tools



# Web apps with shiny

## Example applications

- Data explorer
- Interactive appendix
- Workflow apps
- Learning tools
- … and so much more!



# More Shiny resources

## Online resources

- Shiny official website
- Shiny official tutorial
- Shiny cheatsheet
- Mastering Shiny, book by Hadley Wickham
- Many useful articles about different topics
- Publishing own Shiny apps for free with shinapps.io
- Hosting your Shiny app on your own server
- Debugging Shiny

## An overview of Shiny extensions

- awesome-shiny-extensions

## Some highlights

- shinythemes: Altering the overall appearance of Shiny apps
- shinyjs: Enrich apps with JavaScript operations
- leaflet: Interactive maps
- ggvis: Similar to ggplot2 but with focus on web and interaction
- shinydashboard: Tools to create visual dashboards

# The science of science communication

# The science of science communication

## Motivation

- You have learned the basic rules of good visualization and reporting, but what is the external validity of this advice?
- How are facts and figures perceived by the public, policymakers, or even other scientists?
- How does science communication affect attitudes and behaviors?

## Show us the data

- It turns out there's increasing evidence on how science communication is consumed by stakeholders, and to what effect.
- However, much more research is needed to better understand how what we do and communicate travels to stakeholders.
- The following slides report some selected findings.

# Lawmakers, scientists, and evidence

## Lawmakers' use of scientific evidence can be improved

**D. Max Crowley[a,1], J. Taylor Scott[a], Elizabeth C. Long[a], Lawrie Green[a], Azaliah Israel[a], Lauren Supplee[b], Elizabeth Jordan[b], Kathryn Oliver[c], Shannon Guillot-Wright[d,e], Brittany Gay[f], Rachel Storace[a], Naomi Torres-Mackie[g], Yolanda Murphy[a], Sandra Donnay[h], Jenna Reardanz[i], Rebecca Smith[j], Kristina McGuire[j], Elizabeth Baker[k], Ana Antonopoulos[l], Mary McCauley[a], and Cagla Giray[a]**

[a]Evidence-to-Impact Collaborative, Pennsylvania State University, University Park, PA 16802; [b]Child Trends, Bethesda, MD 20814; [c]Transforming Evidence, London School of Hygiene & Tropical Medicine, London, WC1H 9SR, United Kingdom; [d]Department of Obstetrics & Gynecology, University of Texas Medical Branch, Galveston, TX 77550; [e]Center for Violence Prevention, University of Texas Medical Branch, Galveston, TX 77550; [f]Department of Psychology, University of Maryland, Baltimore County, Baltimore, MD21250; [g]Teachers College, Columbia University, New York, NY 10027; [h]The Racial Equity Initiative, Skillman, NJ 08558; [i]Department of Psychology, University of Alabama, Tuscaloosa, AL 35487-0348; [j]Department of Psychology, Virginia Commonwealth University, Richmond, VA 23238; [k]Department of Psychology, University of Calgary, Calgary, Alberta, Canada T2N 1N4; and [l]Georgetown University Medical School, Washington, DC20007

Edited by Douglas S. Massey, Princeton University, Princeton, NJ, and approved December 12, 2020 (received for review July 6, 2020)

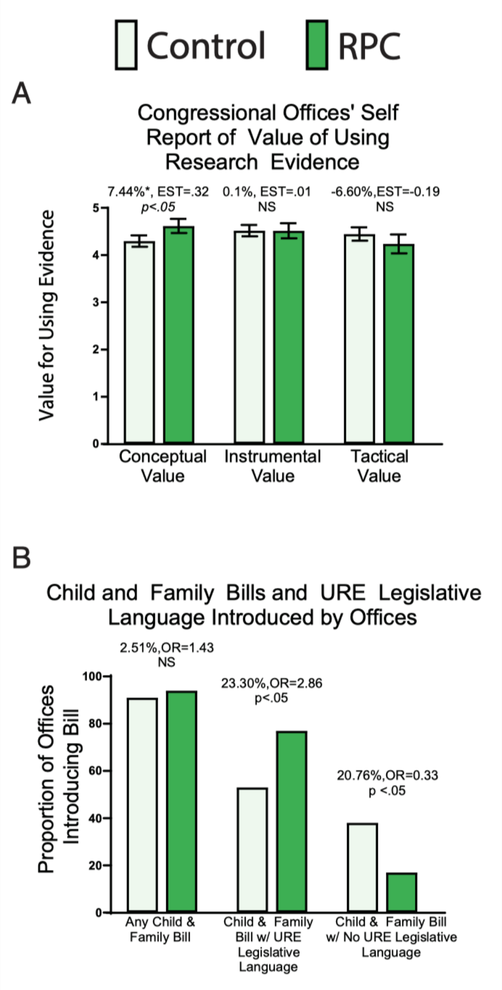
**Core to the goal of scientific exploration is the opportunity to guide future decision-making. Yet, elected officials often miss opportunities to use science in their policymaking. This work reports on an experiment with the US Congress—evaluating the effects of a randomized, dual-population (i.e., researchers and congressional offices) outreach model for supporting legislative use of research evidence regarding child and family policy issues. In this experiment, we found that congressional offices randomized to the intervention reported greater value of research for understanding issues than the control group following implementation. More research use was also observed in legislation introduced by the intervention group. Further, we found that researchers randomized to the intervention advanced their own policy knowledge and engagement as well as reported benefits for their research following implementation.**

evidence-based policymaking | randomized controlled trial | Congress

response to opportunities or crises (7, 8, 12, 13). Timeliness of researcher engagement is particularly challenging since public policy goals often shift suddenly in response to socio-political factors (9, 14). Thus, there is a need for engaging researchers in real-time during discrete, time-limited opportunities for policy change (10, 15).

Policymakers can decide to use research evidence for varied purposes or intentions. A widely used typology in URE investigations is informed by foundational work of multiple scholars (16, 17). While researchers often deplore political uses of research for persuading others, justifying, or challenging existing policy proposals (i.e., tactical use), research evidence can also be used to guide policy development itself. This includes instances in which research is used to directly inform policy decisions (i.e., instrumental use) as well as instances in which research is indirectly used by changing the way policymakers think about problems or solutions (i.e., conceptual use). While instrumental uses may be relatively observable in specific policy efforts, conceptual use may influence

Source Crowley et al. 2021, *PNAS*



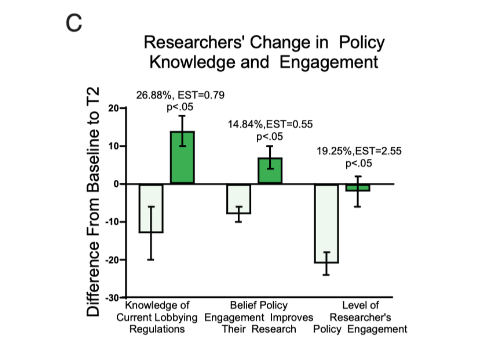




**Fig. 3.** Effects of the RPC Intervention Model. (*A*) Congressional offices participating in the RPC reported greater value of conceptual research evidence than controls, but there were no significant differences in instrumental or tactical uses. (*B*) Offices participating in the RPC were 1) not more likely to write more child and family bills, but were 2) more likely to include research evidence terms in legislative language, and (3) were less likely to introduce bills that did not include URE research evidence terms. (*C*) Researchers in the RPC reported improved (1) knowledge of current lobbying restrictions and (2) belief that engaging with policymakers would improve their own research. Their level of policy engagement was sustained, whereas control group researchers declined in policy engagement.

# Study characteristics and appreciation

Weighing the Evidence: Which Studies Count?

Eva Vivalt* Aidan Coville† Sampada KC‡

April 30, 2021

**Abstract**

We present results from two experiments run at World Bank and Inter-American Development Bank workshops on how policy-makers, practitioners and researchers weigh evidence and seek information from impact evaluations. We find that policy-makers care more about attributes of studies associated with external validity than internal validity, while for researchers the reverse is true. These preferences can yield large differences in the estimated effects of pursued policies: policy-makers indicated a willingness to accept a program that had a 6.3 percentage point smaller effect on enrollment rates if it were recommended by a local expert, larger than the effects of most programs.

Table 5: Attributes and Levels used for IDB 2016 & 17, Nairobi, and Mexico City Sample

| Attributes | Levels |
|---|---|
| Method | Experimental, Quasi-experimental, Observational |
| Location | Different country, Same country, Different country in the same region |
| Impact | −5, 0, +5, +10 percentage points |
| Organization | Government, NGO |
| Sample Size | 50, 3,000, 15,000 |

Table 6: Seeking Research Results by Type of Respondent

| | *World Bank* | | | *IDB* | |
|---|---|---|---|---|---|
| | Policy-maker (1) | Practitioner (2) | Researcher (3) | Policy-maker (4) | Practitioner (5) |
| Impact | 1.053*** | 1.035* | 1.014 | 1.023* | 1.011 |
| | (0.017) | (0.018) | (0.021) | (0.012) | (0.019) |
| Quasi-Experimental | 1.625** | 2.180*** | 4.267*** | 1.331* | 1.527* |
| | (0.341) | (0.469) | (1.294) | (0.224) | (0.383) |
| Experimental | 2.473*** | 2.728*** | 8.869*** | 1.371** | 2.327*** |
| | (0.592) | (0.677) | (3.431) | (0.218) | (0.595) |
| Different country, same region | 1.563** | 1.492* | 1.077 | 1.556*** | 2.118*** |
| | (0.328) | (0.325) | (0.344) | (0.236) | (0.491) |
| Same country | 1.728** | 2.011*** | 1.386 | 2.363*** | 2.537*** |
| | (0.369) | (0.453) | (0.346) | (0.391) | (0.674) |
| Sample size: 3000 | 1.455* | 1.607** | 6.413*** | 2.007*** | 3.095*** |
| | (0.313) | (0.380) | (2.481) | (0.325) | (0.723) |
| Sample size: 15000 | 1.656** | 1.372 | 6.946*** | 1.974*** | 4.680*** |
| | (0.358) | (0.309) | (2.578) | (0.321) | (1.280) |
| Government | 1.338* | 1.015 | 0.951 | 0.948 | 1.434** |
| | (0.208) | (0.167) | (0.209) | (0.106) | (0.243) |
| Observations | 209 | 206 | 180 | 394 | 233 |

This table reports the results of conditional logit regressions on which impact evaluation was selected, using odds ratios. The omitted categories are "Observational", "Different region", "Sample size: 50", and "NGO". The number of observations represents the total number of choices made across individuals. The IDB results use only the pre-workshop sample. Standard errors are provided in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Source Vivalt et al. 2022, *working paper*

# Reported uncertainty and public trust

## The effects of communicating uncertainty on public trust in facts and numbers

Anne Marthe van der Bles[a,b,c,1], Sander van der Linden[a,b,d,1], Alexandra L. J. Freeman[a,b], and David J. Spiegelhalter[a,b]

[a]Winton Centre for Risk and Evidence Communication, University of Cambridge, Cambridge CB3 0WA, United Kingdom; [b]Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge CB3 0WA, United Kingdom; [c]Department of Social Psychology, University of Groningen, 19712 TS Groningen, The Netherlands; and [d]Cambridge Social Decision-Making Lab, Department of Psychology, University of Cambridge, Cambridge CB2 3RQ, United Kingdom

Edited by Arild Underdal, University of Oslo, Oslo, Norway, and approved February 20, 2020 (received for review August 7, 2019)

**Uncertainty is inherent to our knowledge about the state of the world yet often not communicated alongside scientific facts and numbers. In the "posttruth" era where facts are increasingly contested, a common assumption is that communicating uncertainty will reduce public trust. However, a lack of systematic research makes it difficult to evaluate such claims. We conducted five experiments—including one preregistered replication with a national sample and one field experiment on the *BBC News* website (total $n$ = 5,780)—to examine whether communicating epistemic uncertainty about facts across different topics (e.g., global warming, immigration), formats (verbal vs. numeric), and magnitudes (high vs. low) influences public trust. Results show that whereas people do perceive greater uncertainty when it is communicated, we observed only a small decrease in trust in numbers and trustworthiness of the source, and mostly for verbal uncertainty communication. These results could help reassure all communicators of facts and science that they can be more open and transparent about the limits of human knowledge.**

communication | uncertainty | trust | posttruth | contested

the general sense of honesty evoked [by uncertainty] . . . this did not seem to offset concerns about the agency's competence" (p. 491). In fact, partly for these reasons, Fischhoff (1) notes that scientists may be reluctant to discuss the uncertainties of their work. This hesitation spans across domains: For example, journalists find it difficult to communicate scientific uncertainty and regularly choose to ignore it altogether (10, 11). Physicians are reluctant to communicate uncertainty about evidence to patients (12), fearing that the complexity of uncertainty may overwhelm and confuse patients (13, 14). Osman et al. (15) even go as far as to argue explicitly that "the drive to increase transparency on uncertainty of the scientific process specifically does more harm than good" (p. 131).

At the same time, many organizations that produce and communicate evidence to the public, such as the European Food Safety Authority, have committed themselves to openness and transparency about their (scientific) work, which includes communicating uncertainties around evidence (16–19). These attempts have not gone without criticism and discussion about the potential impacts on public opinion (15, 20). What exactly do we know about the effects of communicating uncertainty around

**Table 1. Overview of the conditions and manipulation texts of experiment 3 and 4**

| Format | Experiment 3 |
|---|---|
| Control (no uncertainty) | "Official figures from the first quarter of 2018 show that UK unemployment fell by 116,000 compared with the same period last year. [. . .]" |
| Numerical range with point estimate | . . .by 116,000 (range between 17,000 and 215,000). . . |
| Numerical range without point estimate | . . .by between 17,000 and 215,000. . . |
| Numerical point estimate ±2 SEs | . . .by 116,000 (±99,000). . . |
| Verbal explicit uncertainty statement | . . .by 116,000 compared with the same period last year, although there is some uncertainty around this figure: It could be somewhat higher or lower. [. . .] |
| Verbal implicit uncertainty statement | . . .by 116,000 compared with the same period last year, although there is a range around this figure: could be somewhat higher or lower. [. . .] |
| Verbal uncertainty word | . . .by an estimated 116,000. . . |
| Mixed numerical and verbal phrase | . . .by an estimated 116,000 (±99,000). . . |

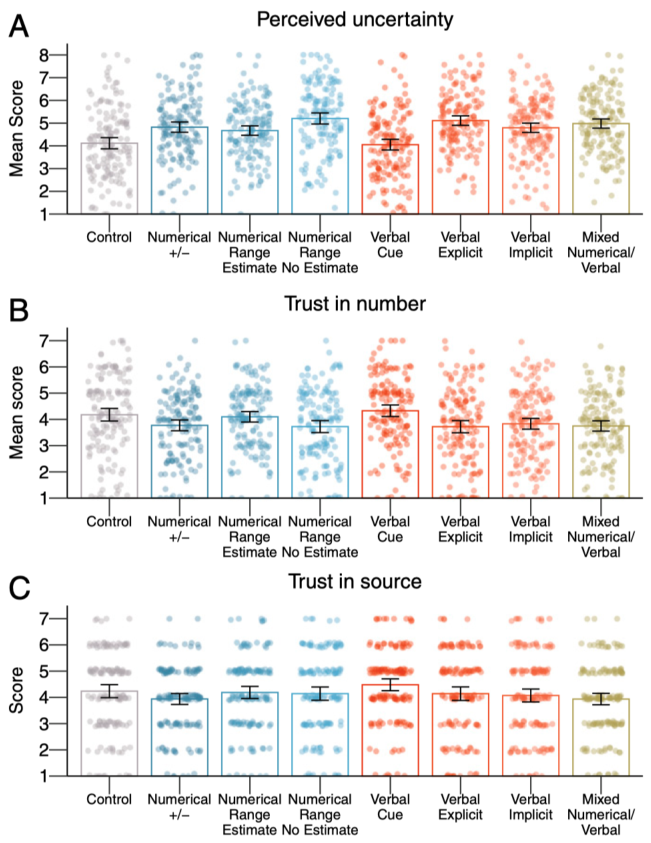
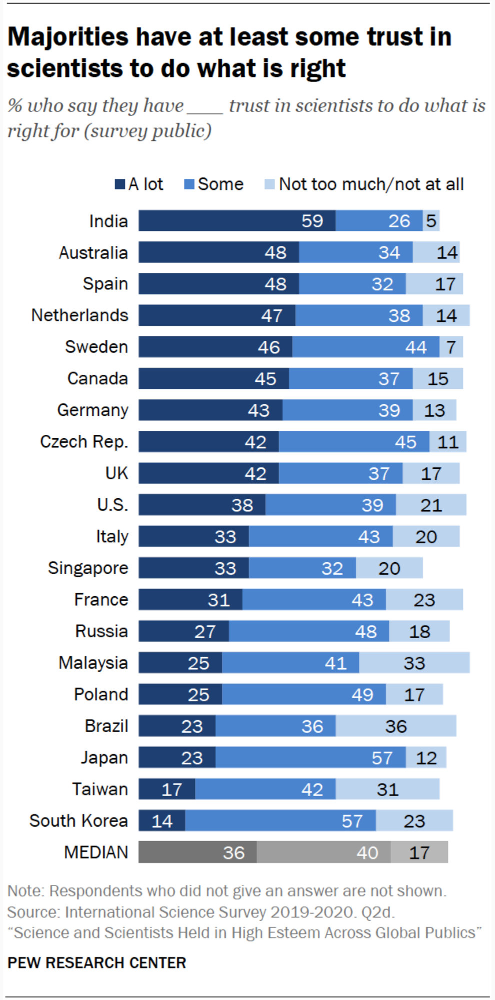


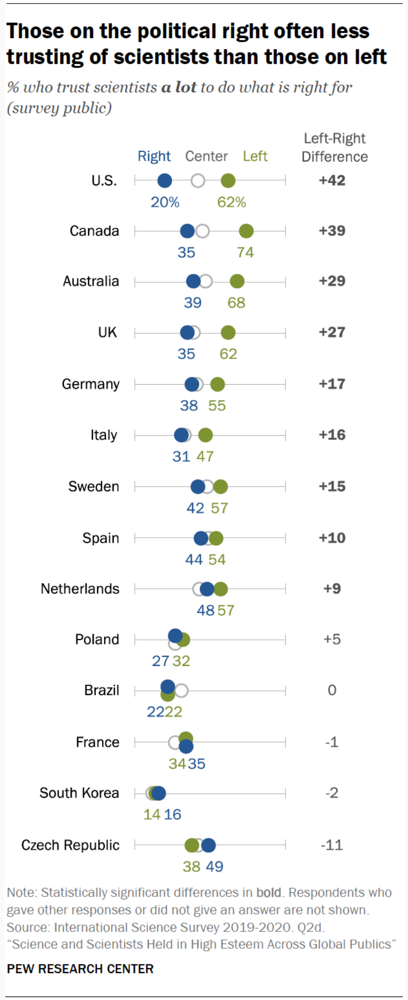

**Fig. 3.** The results of experiment 3: Means per condition for perceived uncertainty (*A*), trust in numbers (*B*), and trust in the source (*C*). The error bars represent 95% CIs around the means, and jitter represents the distribution of the underlying data.

`Source` Van der Bles et al. 2020, *PNAS*

# Towards open data science

# Trust in science




`Source` Pew 2020

# Trust in data science

## Data scientists have the potential to help save the world

By Leo Borrett May 17, 2017

With an untold number of crises emerging every year, big data is becoming increasingly important for helping aid organisations respond quickly to chaotic and evolving situations.

## HOW DATA SCIENCE IS SAVING LIVES

AVINASH N Sep 29 · 2 min read

For all the people first priority is about their life. Life is one of the most precious thing in the world. Can Data Science techniques save life, is it possible? Yes, using Data Science techniques to analyze large data sets today has a huge impact on saving lives.

**Health**

## Artificial intelligence and covid-19: Can the machines save us?

**Analytics And Data Science**

## Data Scientist: The Sexiest Job of the 21st Century

Meet the people who can coax treasure out of messy, unstructured data. by Thomas H. Davenport and D.J. Patil

From the Magazine (October 2012)

## How AI Will Save Thousands of Lives

Sepsis is the problem; data are the cure

Drew Smith, PhD Jan 10, 2020 · 5 min read ★

**STUDENTS**

## Data Science: Why It Matters and How It Can Make You Rich

# Trust in data science?

**The Cambridge Analytica case: What's a data scientist to do?**

The Cambridge Analytica controversy has highlighted data ethics issues especially dear to early career stage data scientists

**Researchers just released profile data on 70,000 OkCupid users without permission**

By Brian Resnick | @B_resnick | brian@vox.com | May 12, 2016, 6:00pm EDT

**An Algorithm That 'Predicts' Criminality Based on a Face Sparks a Furor**

Its creators said they could use facial analysis to determine if someone would become a criminal. Critics said the work recalled debunked "race science."

**Data Failed the Election, But There's Still Hope for Business** Everyone is blaming data for failing to predict Trump's win. But it's the data handlers who need the real reexamination.

# The replication crisis

## What the crisis is about

- The finding that many scientific studies are difficult or impossible to reproduce.
- Reproducibility is a cornerstone of science as an enterprise of knowledge generation → bad.

## Factors fueling the replication crisis

- Solo, silo-ed investigators limited to small sample sizes
- Wrong incentives in science
- No pre-registration of hypotheses being tested
- Post-hoc cherry picking of hypotheses with best P values
- Only requiring $P < .05$
- No replication
- No data sharing

*Open access, freely available online*

Essay

**Why Most Published Research Findings Are False**

John P. A. Ioannidis

**Summary**

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance. Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias. In this essay, I discuss the implications of these problems for the conduct and interpretation of research.

Published research findings are sometimes refuted by subsequent evidence, with ensuing confusion and disappointment. Refutation and controversy is seen across the range of research designs, from clinical trials and traditional epidemiological studies [1–3] to the most modern molecular research [4,5]. There is increasing concern that in modern research, false findings may be the majority or even the vast majority of published research claims [6–8]. However, this should not be surprising. It can be proven that most claimed research findings are false. Here I will examine the key factors that influence this problem and some corollaries thereof.

The Essay section contains opinion pieces on topics of broad interest to a general medical audience.

**Modeling the Framework for False Positive Findings**

Several methodologists have pointed out [9–11] that the high rate of nonreplication (lack of confirmation) of research discoveries is a consequence of the convenient, yet ill-founded strategy of claiming conclusive research findings solely on the basis of a single study assessed by formal statistical significance, typically for a *p*-value less than 0.05. Research is not most appropriately represented and summarized by *p*-values, but, unfortunately, there is a widespread notion that medical research articles should be interpreted based only on *p*-values. Research findings are defined here as any relationship reaching formal statistical significance, e.g., effective interventions, informative predictors, risk factors, or associations. "Negative" research is also very useful. "Negative" is actually a misnomer, and the misinterpretation is widespread. However, here we will target relationships that investigators claim exist, rather than null findings.

**It can be proven that most claimed research findings are false.**

As has been shown previously, the probability that a research finding is indeed true depends on the prior probability of it being true (before doing the study), the statistical power of the study, and the level of statistical significance [10,11]. Consider a $2 \times 2$ table in which research findings are compared against the gold standard of true relationships in a scientific field. In a research field both true and false hypotheses can be made about the presence of relationships. Let $R$ be the ratio of the number of "true relationships" to "no relationships" among those tested in the field. $R$ is characteristic of the field and can vary a lot depending on whether the field targets highly likely relationships or searches for only one or a few true relationships among thousands and millions of hypotheses that may be postulated. Let us also consider, for computational simplicity, circumscribed fields where either there is only one true relationship (among many that can be hypothesized) or the power is similar to find any of the several existing true relationships. The pre-study probability of a relationship being true is $R/(R+1)$. The probability of a study finding a true relationship reflects the power $1 - \beta$ (one minus the Type II error rate). The probability of claiming a relationship when none truly exists reflects the Type I error rate, $\alpha$. Assuming that $c$ relationships are being probed in the field, the expected values of the $2 \times 2$ table are given in Table 1. After a research finding has been claimed based on achieving formal statistical significance, the post-study probability that it is true is the positive predictive value, PPV. The PPV is also the complementary probability of what Wacholder et al. have called the false positive report probability [10]. According to the $2 \times 2$ table, one gets $PPV = (1 - \beta)R/(R - \beta R + \alpha)$. A research finding is thus

**Citation:** Ioannidis JPA (2005) Why most published research findings are false. PLoS Med 2(8): e124.



**Abbreviation:** PPV, positive predictive value

John P. A. Ioannidis is in the Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece, and Institute for Clinical Research and Health Policy Studies, Department of Medicine, Tufts-New England Medical Center, Tufts University School of Medicine, Boston, Massachusetts, United States of America. E-mail: jioannid@cc.uoi.gr

**Competing Interests:** The author has declared that no competing interests exist.

**DOI:** 10.1371/journal.pmed.0020124

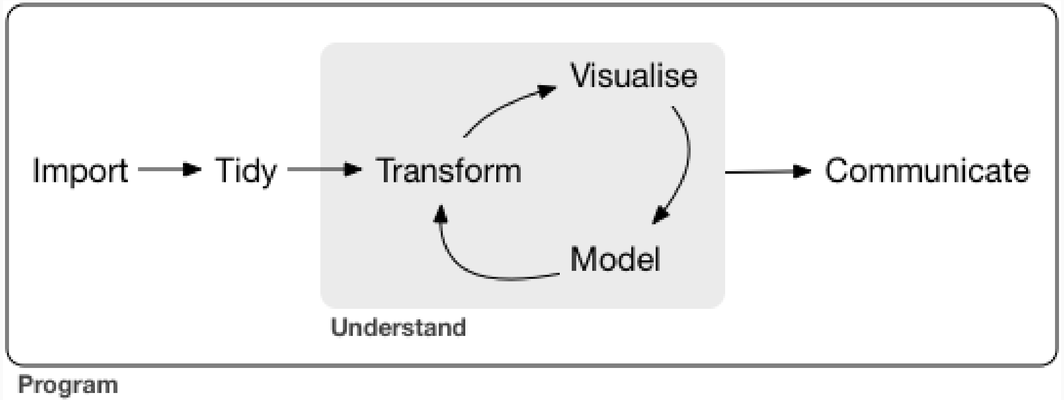


Source Ioannidis 2005/PLOS Medicine

# Fostering trust in science through open science

## Communicating the entire workflow

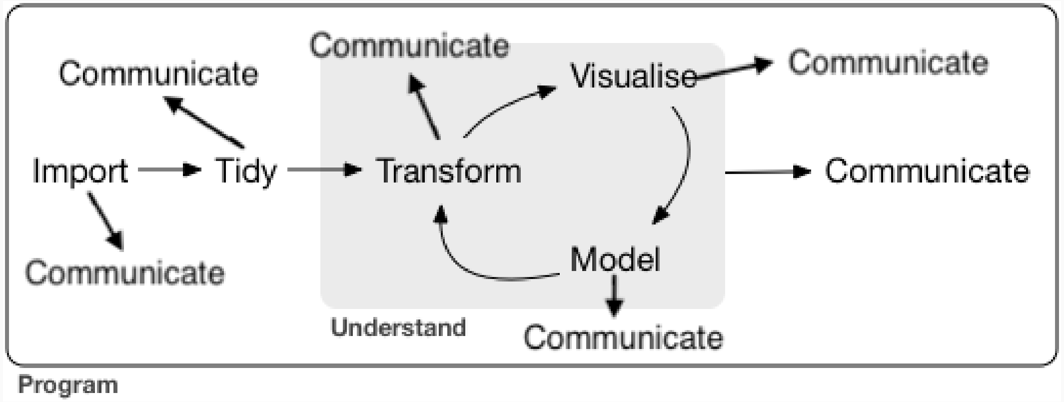
- Any decision - from conceptualizing measures to formatting tables - is meaningful for your output.
    - Tiny mistakes can have massive technical consequences (→ debugging).
    - Various decisions can have ethical implications (→ next session).
    - For others to follow (and potentially invalidate) your analyses, all the details are important.
- What makes the scientific endeavor unique is that it has self-correcting mechanisms.
- But in order for them to work effectively, you have to **be transparent about every step in your workflow**.



# Fostering trust in science through open science

## Communicating the entire workflow

- Any decision - from conceptualizing measures to formatting tables - is meaningful for your output.
  - Tiny mistakes can have massive technical consequences (→ debugging).
  - Various decisions can have ethical implications (→ next session).
  - For others to follow (and potentially invalidate) your analyses, all the details are important.
- What makes the scientific endeavor unique is that it has self-correcting mechanisms.
- But in order for them to work effectively, you have to **be transparent about every step in your workflow**.
- That's a meta meaning of communication: tell others, by publishing everything from input to code to output, what you've done.



# Towards open data science (cont.)

## Good practice

- Pre-register designs `osf.io`, `aspredicted.org`
- Do version control `GitHub`
- Publish all research outputs (and inputs if possible) `GitHub`, `plain-text formats`
- Disclose and document software pipeline `targets`, `make`
- Make analysis reproducible `renv`, `Docker`
- Make preprints accessible `arXiv`, `osf.io`
- Public in open access journals

# Towards open data science (cont.)

## Good practice

- Pre-register designs `osf.io`, `aspredicted.org`
- Do version control `GitHub`
- Publish all research outputs (and inputs if possible) `GitHub`, `plain-text formats`
- Disclose and document software pipeline `targets`, `make`
- Make analysis reproducible `renv`, `Docker`
- Make preprints accessible `arXiv`, `osf.io`
- Public in open access journals


**Notice something?**

After the last 10 sessions, you have already become open science practitioners. You are welcome.