Data in public communications

•••

Giuseppe Sollazzo
Battistini Lecture
Institute of Advanced Studies, University of Bologna

1 December 2020

Some definitions

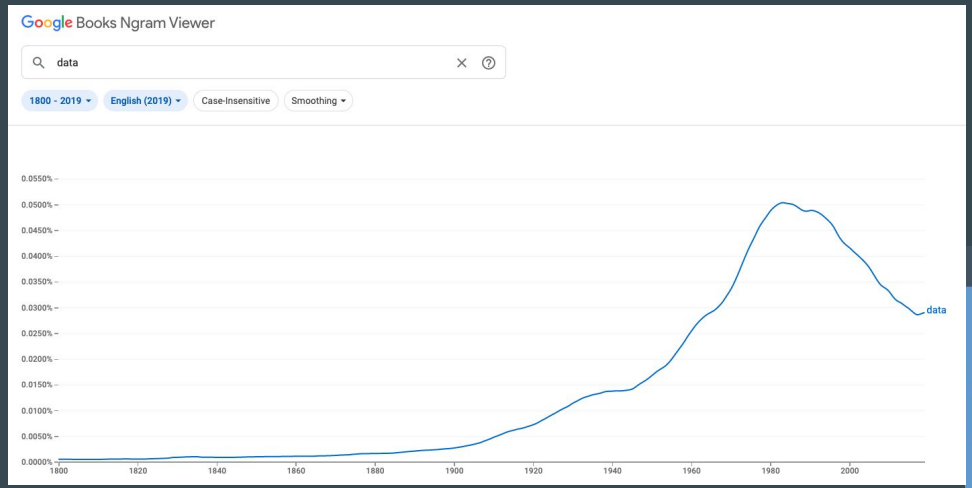
Data can be stored in a digital form processed on a computer

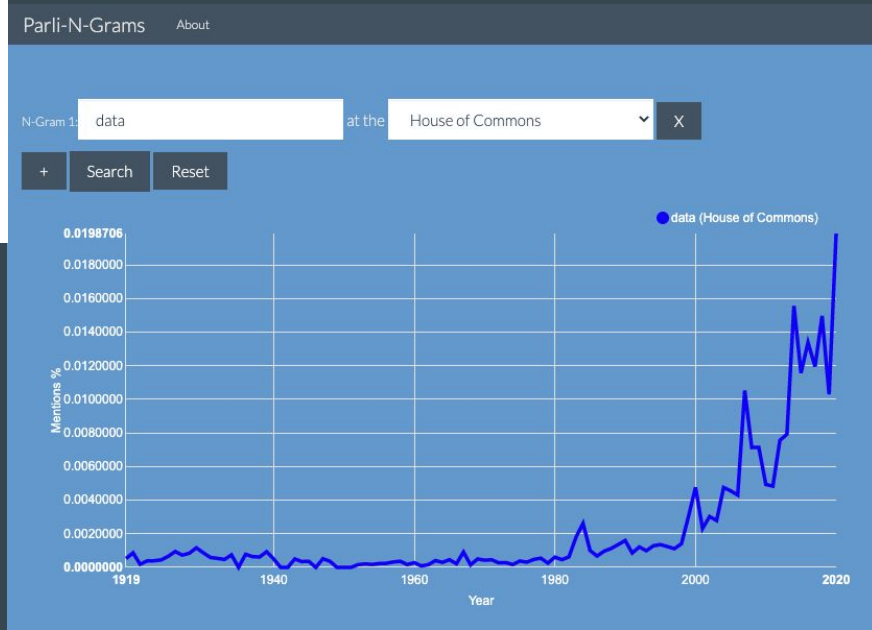
“Public” means intended for the broadest audience possible

Aims

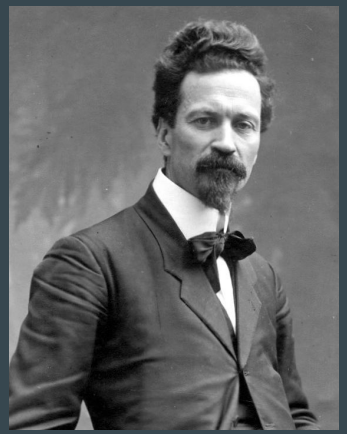
1. Historical and technical elements of data in public communications
2. The advent of official statistics in public communications
3. Data in the recent pandemic
4. Common problems of data communications

“Data”


Google Books Ngram Viewer
data
1800 - 2019
English (2019)
Case-Insensitive
Smoothing
0.0550%
0.0500%
0.0450%
0.0400%
0.0350%
0.0300%
0.0250%
0.0200%
0.0150%
0.0100%
0.0050%
0.0000%
1800
1820
1840
1860
1880
1900
1920
1940
1960
1980
2000
data


Parli-N-Grams
About
N-Gram 1:
data
at the
House of Commons
X
+
Search
Reset
data (House of Commons)
0.0198706
0.0180000
0.0160000
0.0140000
0.0120000
0.0100000
0.0080000
0.0060000
0.0040000
0.0020000
0.0000000
Mentions %
1919
1940
1960
1980
2000
2020
Year

“Di più avrei potuto fare, specialmente nel campo statistico, se non ci fosse nel nostro paese, e nei privati e negli enti morali, una tal quale ritrosia a confidare al dominio del pubblico dati, fatti e notizie.”



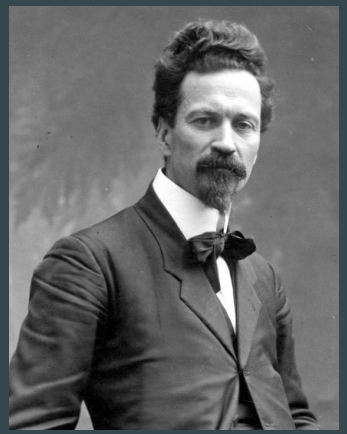
Cesare Battisti

Wikipedia, Public Domain

“I could have done more, especially in the statistical field, if there was not in our country, and in private individuals and moral bodies, such a reluctance to confide in the public domain data, facts and news.”

(Il Trentino, 1898)

“Non tutta la regione compresa fra i limiti dell’Italia fisica coincide coi confini dell’Italia linguistica.”

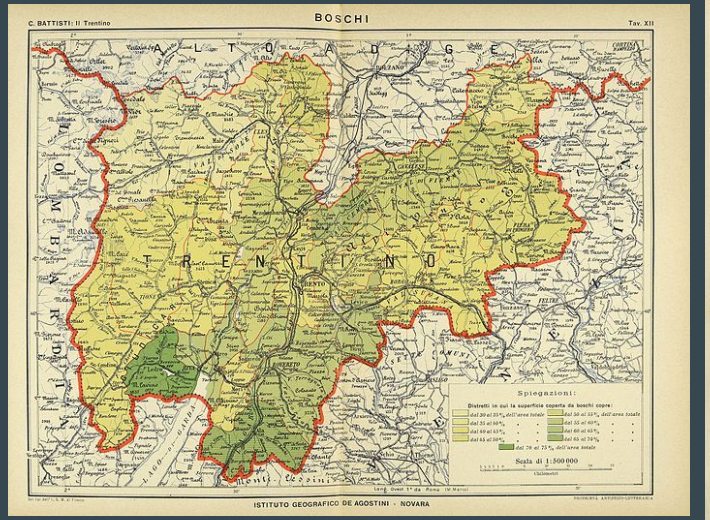


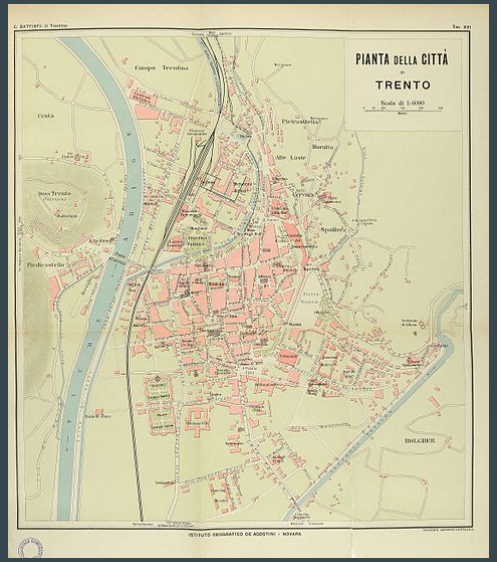
“Only part of the region within the geographic borders of Italy coincides with the borders of linguistic Italy”

(Il Trentino, 1898)

Cesare Battisti

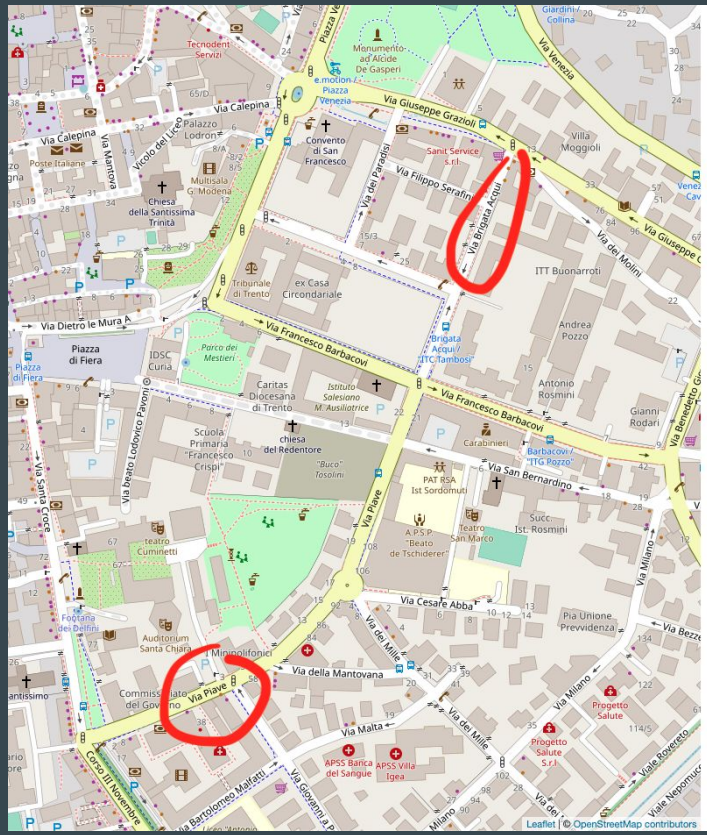
Wikipedia, Public Domain

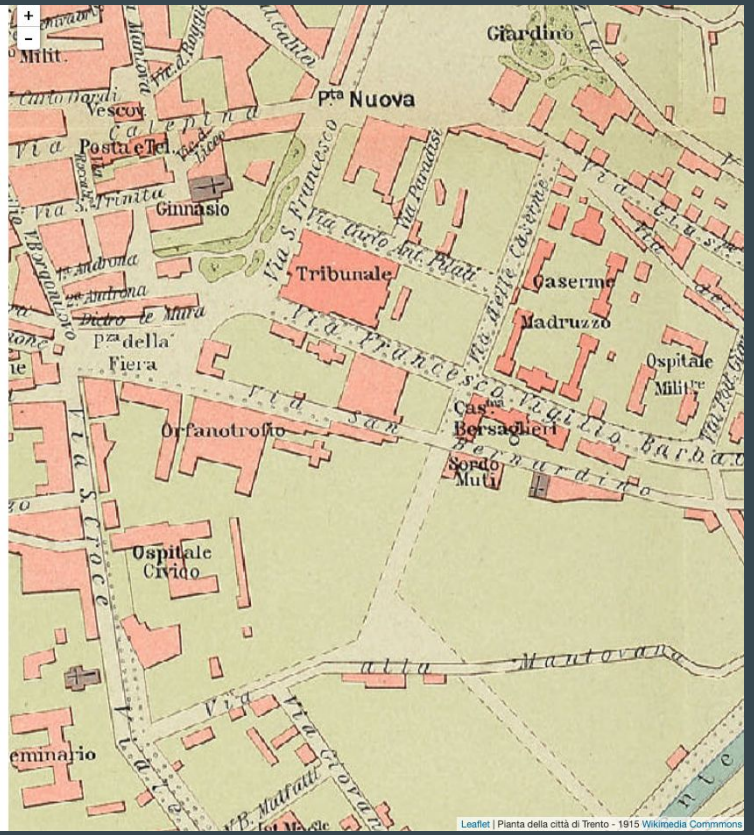




Cesare Battisti, Il Trentino (1915)

Wikipedia Commons





Trento, today and in 1915

[“Trento nel 1915”, Labmod]

Number of provinces [hide]

Year	Provinces
1861	59
1866	68
1870	69
1923	75
1924	76
1927	92
1934	93
1935	94
1941	95
1944	94
1945	93
1947	91
1954	92
1968	93
1970	94
1974	95
1992	103
2001	107
2004	110
2016	107

Provinces of Italy 1942 vs 2020

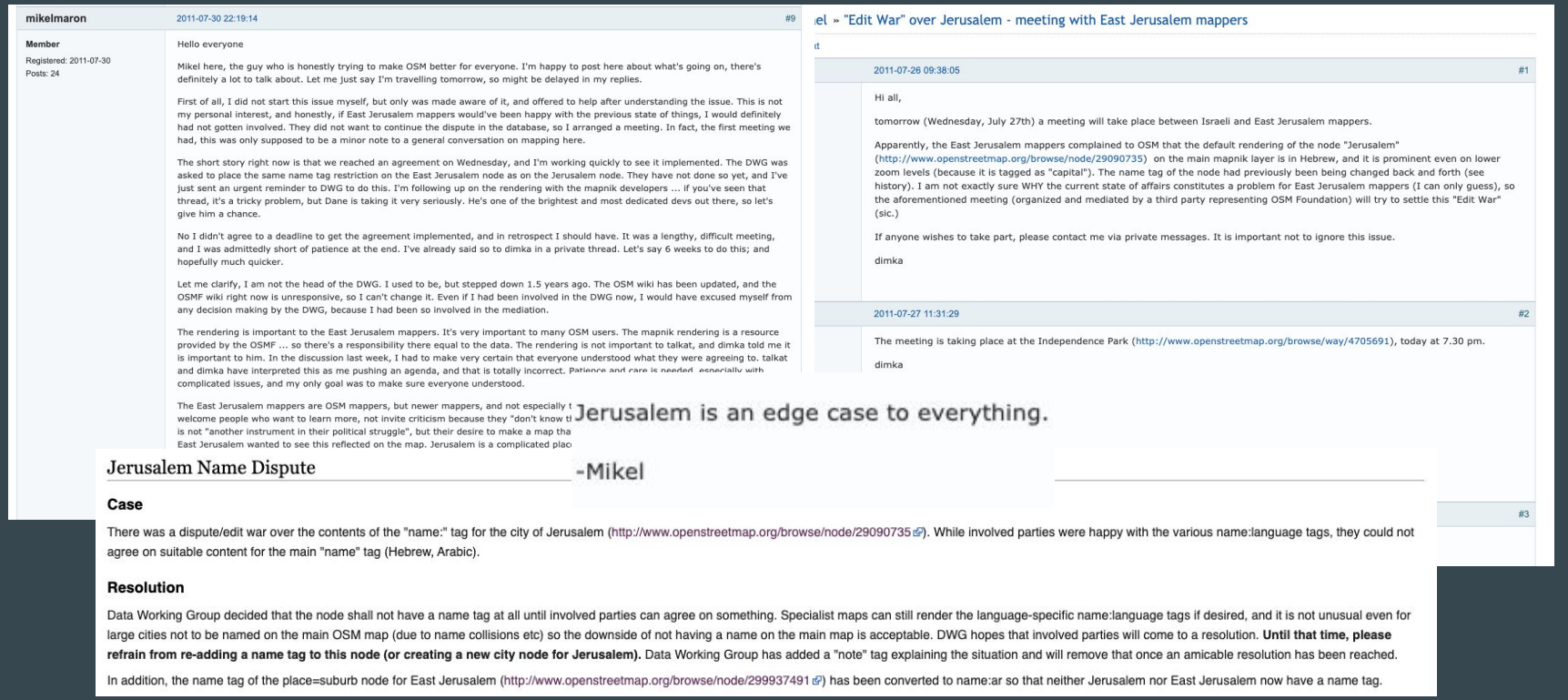
[1942, Wikipedia]

[2020, Wikipedia]

Crimea	Russia Ukraine	45.159/34.634	In 2014, Russian forces occupied and annexed the Crimean peninsula, seizing it from Ukrainian territories. Russian claims to Crimea are not internationally recognized. The United Nations, through General Assembly Resolution 68/262, affirmed that Crimea was part of Ukraine and that Russia's annexation is invalid. Several nations imposed international sanctions on Russia in reaction to its action.
Aegean dispute, Imia/Kardak	Greece Turkey		Broad number of delimitation disputes about a.o. national airspace, territorial waters and exclusive economic zones. Includes Imia/Kardak dispute.
Mont Blanc summit dispute	France Italy		France asserts that the principal peaks on the Mont Blanc massif—Dôme du Goûter, Punta Helbronner, and Mont Blanc lie in French territory, while Italy asserts that the summits are shared.
Carlingford Lough and Lough Foyle boundary dispute	Ireland United Kingdom		Lough Foyle divides County Donegal, Republic of Ireland, and County Londonderry, Northern Ireland. Carlingford Lough divides County Louth, Republic of Ireland, and County Down, Northern Ireland.
Gibraltar	United Kingdom Spain		Dispute over the interpretation of the Treaty of Utrecht and the location of the border.
Dollart/Dollard bay	Germany Netherlands		The exact course of the border through this bay is disputed, yet the countries have agreed to disagree by signing a treaty in 1960.
Lake Constance	Germany Austria Switzerland		Switzerland holds the view that the border runs through the middle of the lake. Austria is of the opinion that the contentious area belongs to all the states on its banks. Germany holds an ambiguous opinion.
Olivenza and Vila Real(including the municipality of Táliga)	Spain Portugal		In 1801, during the War of the Oranges, Spain, with French military support, occupied the territory of Olivenza (in Portuguese *Olivença*). During the Treaty of Vienna (1815), the signatory powers (including Spain) agreed with the Portuguese arguments concerning its claim on Olivença but Spain never fulfilled its duty of giving the city of Olivença and its territory back to Portugal.
Croatia-Serbia border dispute	Croatia Serbia		Limited areas along the Danube Parts of Osijek-Baranja and Vukovar-Syrmia Counties and West and South Bačka Districts
Gulf of Piran	Slovenia Croatia		An agreement was signed (and ratified by Croatia's parliament on 20 November 2009) to pursue binding arbitration to both the land and maritime portions of this continuing dispute. In 2015 collusion between the Slovenian judge on the arbitration panel and a representative from the Slovenian government was uncovered. The Croatian Sabor voted to withdraw from the arbitration, citing allegations of significant breaches of arbitration rules by Slovenia as the reason. Despite this the arbitration tribunal continued its work, issuing a ruling in 2017.
Prevlaka	Croatia Montenegro		

Disputed territories

OpenStreetMap



mikelmaron — 2011-07-30 22:19:14 #9

Member
Registered: 2011-07-30
Posts: 24

Hello everyone

Mikel here, the guy who is honestly trying to make OSM better for everyone. I'm happy to post here about what's going on, there's definitely a lot to talk about. Let me just say I'm travelling tomorrow, so might be delayed in my replies.

First of all, I did not start this issue myself, but only was made aware of it, and offered to help after understanding the issue. This is not my personal interest, and honestly, if East Jerusalem mappers would've been happy with the previous state of things, I would definitely had not gotten involved. They did not want to continue the dispute in the database, so I arranged a meeting. In fact, the first meeting we had, this was only supposed to be a minor note to a general conversation on mapping here.

The short story right now is that we reached an agreement on Wednesday, and I'm working quickly to see it implemented. The DWG was asked to place the same name tag restriction on the East Jerusalem node as on the Jerusalem node. They have not done so yet, and I've just sent an urgent reminder to DWG to do this. I'm following up on the rendering with the mapnik developers ... if you've seen that thread, it's a tricky problem, but Dane is taking it very seriously. He's one of the brightest and most dedicated devs out there, so let's give him a chance.

No I didn't agree to a deadline to get the agreement implemented, and in retrospect I should have. It was a lengthy, difficult meeting, and I was admittedly short of patience at the end. I've already said so to dimka in a private thread. Let's say 6 weeks to do this; and hopefully much quicker.

Let me clarify, I am not the head of the DWG. I used to be, but stepped down 1.5 years ago. The OSM wiki has been updated, and the OSMF wiki right now is unresponsive, so I can't change it. Even if I had been involved in the DWG now, I would have excused myself from any decision making by the DWG, because I had been so involved in the mediation.

The rendering is important to the East Jerusalem mappers. It's very important to many OSM users. The mapnik rendering is a resource provided by the OSMF ... so there's a responsibility there equal to the data. The rendering is not important to talkat, and dimka told me it is important to him. In the discussion last week, I had to make very certain that everyone understood what they were agreeing to. talkat and dimka have interpreted this as me pushing an agenda, and that is totally incorrect. Patience and care is needed, especially with complicated issues, and my only goal was to make sure everyone understood.

The East Jerusalem mappers are OSM mappers, but newer mappers, and not especially t
welcome people who want to learn more, not invite criticism because they "don't know tl
is not "another instrument in their political struggle", but their desire to make a map tha
East Jerusalem wanted to see this reflected on the map. Jerusalem is a complicated plac

el » "Edit War" over Jerusalem - meeting with East Jerusalem mappers

2011-07-26 09:38:05 #1

Hi all,

tomorrow (Wednesday, July 27th) a meeting will take place between Israeli and East Jerusalem mappers.

Apparently, the East Jerusalem mappers complained to OSM that the default rendering of the node "Jerusalem" (http://www.openstreetmap.org/browse/node/29090735) on the main mapnik layer is in Hebrew, and it is prominent even on lower zoom levels (because it is tagged as "capital"). The name tag of the node had previously been being changed back and forth (see history). I am not exactly sure WHY the current state of affairs constitutes a problem for East Jerusalem mappers (I can only guess), so the aforementioned meeting (organized and mediated by a third party representing OSM Foundation) will try to settle this "Edit War" (sic.)

If anyone wishes to take part, please contact me via private messages. It is important not to ignore this issue.

dimka

2011-07-27 11:31:29 #2

The meeting is taking place at the Independence Park (http://www.openstreetmap.org/browse/way/4705691), today at 7.30 pm.

dimka

#3

Jerusalem is an edge case to everything.

-Mikel

Jerusalem Name Dispute

Case

There was a dispute/edit war over the contents of the "name:" tag for the city of Jerusalem (http://www.openstreetmap.org/browse/node/29090735). While involved parties were happy with the various name:language tags, they could not agree on suitable content for the main "name" tag (Hebrew, Arabic).

Resolution

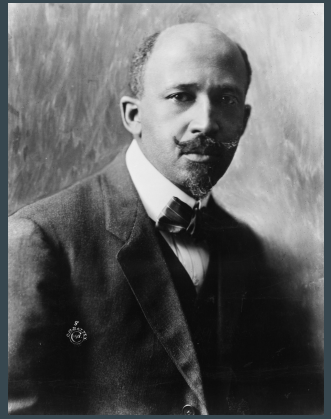
Data Working Group decided that the node shall not have a name tag at all until involved parties can agree on something. Specialist maps can still render the language-specific name:language tags if desired, and it is not unusual even for large cities not to be named on the main OSM map (due to name collisions etc) so the downside of not having a name on the main map is acceptable. DWG hopes that involved parties will come to a resolution. **Until that time, please refrain from re-adding a name tag to this node (or creating a new city node for Jerusalem).** Data Working Group has added a "note" tag explaining the situation and will remove that once an amicable resolution has been reached.

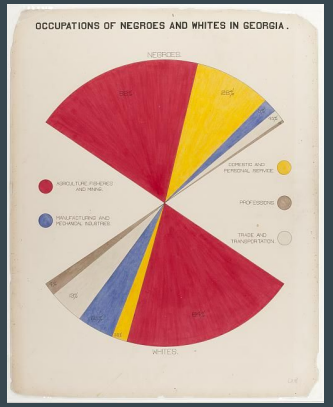
In addition, the name tag of the place=suburb node for East Jerusalem (http://www.openstreetmap.org/browse/node/299937491) has been converted to name:ar so that neither Jerusalem nor East Jerusalem now have a name tag.

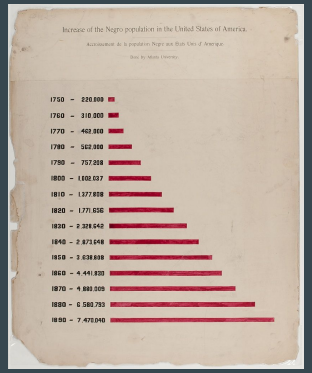
Edit war over Jerusalem

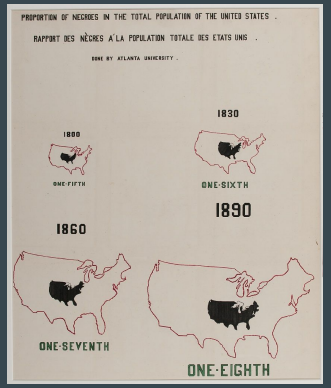
OpenStreetMap Forum
OpenStreetMap Data Working Group Disputes

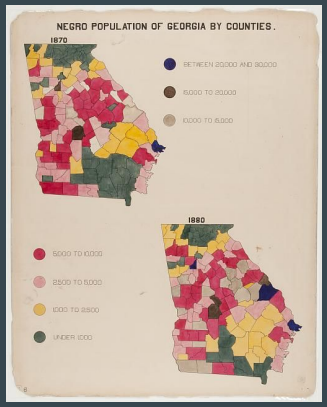
Historical elements

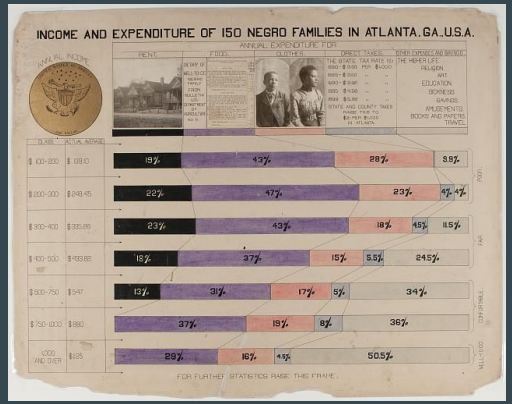


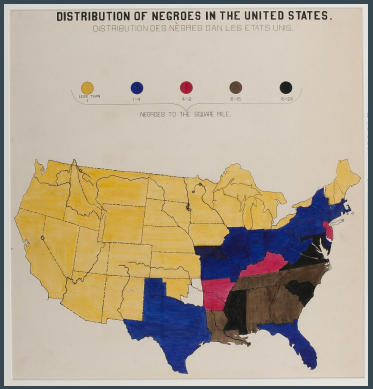








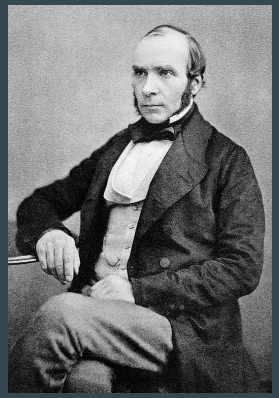


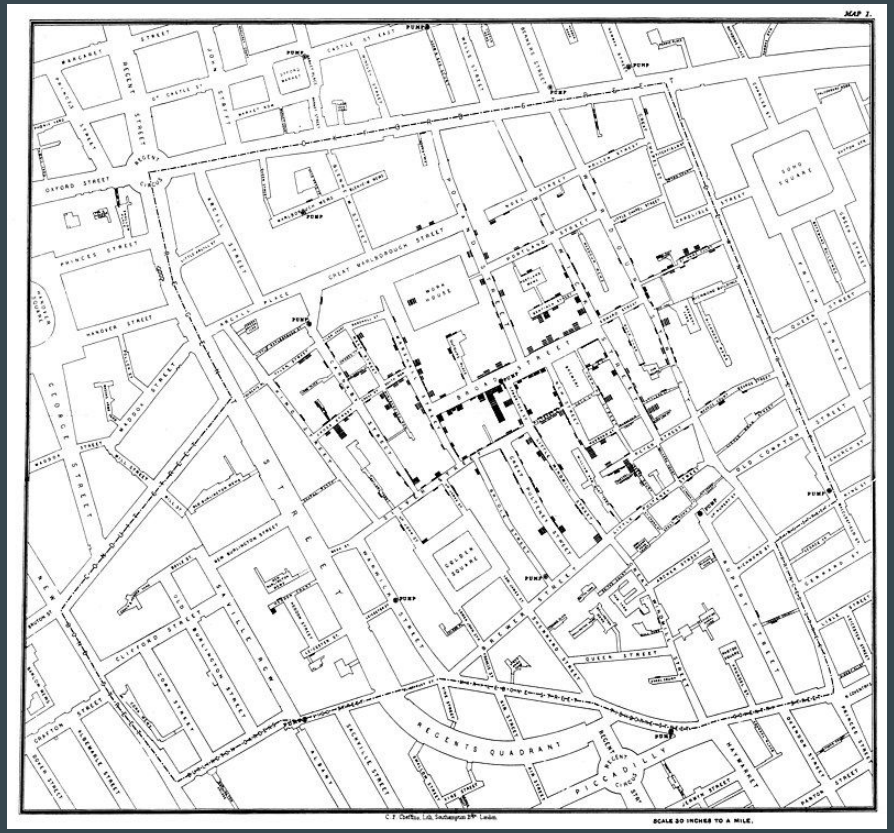


William Edward Burghardt Du Bois

[Wikipedia]

[Library of Congress]





John Snow

[Wikipedia (portrait)]
[Wikipedia (map)]

Official statistics

4000 BC
Babylonians

2500 BC
Egyptians

~500 BC
Romans

2 AD
Han Dynasty

1790
US Census

1801
UK Census

1926
ISTAT

1941
UK Central Statistics Office

Timeline of census history

Process

How is the data collected?

Definitions

How are common sense concepts translated into data?

BBC | BBC Account | Home | News | Sport | Weather | iPlayer | Sounds

NEWS

Home | Coronavirus | US Election | UK | World | Business | Politics | Tech | Science | Health | Family & Education

Business | Your Money | Market Data | Companies | Economy | Global Car Industry | Business of Sport

UK unemployment falls to 1.44 million

24 January 2018

UK unemployment fell by 3,000 to 1.44 million in the three months to November, official figures show.

The number of those in work increased sharply and wages rose at their fastest rate in almost a year, the Office for National Statistics said.

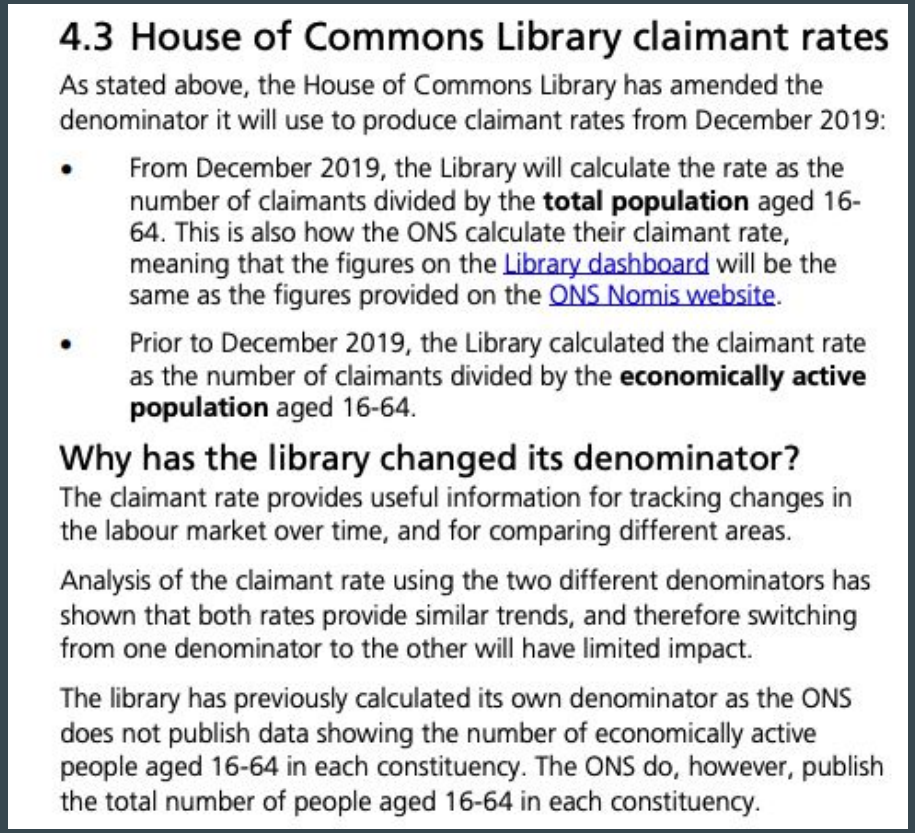
As well as calculating precision measures around the numbers and rates obtained from the survey, we can also calculate them for changes in the numbers. For example, for September to November 2017, the estimated change in the number of unemployed people since June to August 2017 was a small fall of 3,000, with a 95% confidence interval of plus or minus 77,000. This means that we are 95% confident the actual change in unemployment was somewhere between an increase of 74,000 and a fall of 80,000, with the best estimate being a small fall of 3,000. As the estimated fall in unemployment of 3,000 is smaller than 77,000, the estimated fall in unemployment is said to be “not statistically significant”.

Misreporting of survey data

BBC
ONS

“Anyone
aged 16, or over,
who has completed
at least one hour of work per week”

Definition of employment used in official UK statistics
ONS



4.3 House of Commons Library claimant rates

As stated above, the House of Commons Library has amended the denominator it will use to produce claimant rates from December 2019:

- From December 2019, the Library will calculate the rate as the number of claimants divided by the **total population** aged 16-64. This is also how the ONS calculate their claimant rate, meaning that the figures on the Library dashboard will be the same as the figures provided on the ONS Nomis website.
- Prior to December 2019, the Library calculated the claimant rate as the number of claimants divided by the **economically active population** aged 16-64.

Why has the library changed its denominator?

The claimant rate provides useful information for tracking changes in the labour market over time, and for comparing different areas.

Analysis of the claimant rate using the two different denominators has shown that both rates provide similar trends, and therefore switching from one denominator to the other will have limited impact.

The library has previously calculated its own denominator as the ONS does not publish data showing the number of economically active people aged 16-64 in each constituency. The ONS do, however, publish the total number of people aged 16-64 in each constituency.

Example of changing definition

House of Commons Library

1. Does anyone in this household have a business or a farm?

2. **Last week**, did you do **any** work for (either) pay (or profit)?
   If the answer to question 1 is "yes" and the answer to question 2 is "no," the next question is:

3. **Last week**, did you do any unpaid work in the family business or farm?
   For those who reply "no" to both questions 2 and 3, the next key questions used to determine employment status are:

4. **Last week**, (in addition to the business) did you have a job, either full or part time? Include any job from which you were temporarily absent.

5. **Last week**, were you on layoff from a job?

6. What was the main reason you were absent from work **last week**?
   For those who respond "yes" to question 5 about being on layoff, the following questions are asked:

7. Has your employer given you a date to return to work?
   If "no," the next question is:

8. Have you been given any indication that you will be recalled to work within the next 6 months?
   If the responses to either question 7 or 8 indicate that the person expects to be recalled from layoff, he or she is counted as unemployed. For those who were reported as having no job or business from which they were absent or on layoff, the next question is:

9. Have you been doing anything to find work during the last 4 weeks?
   For those who say "yes," the next question is:

10. What are all of the things you have done to find work during the last 4 weeks?
    If an active method of looking for work, such as those listed at the beginning of this section, is mentioned, the following question is asked:

11. **Last week**, could you have started a job if one had been offered?
    If there is no reason, except temporary illness, that the person could not take a job, he or she is considered to be not only looking but also available for work and is counted as unemployed.

Questions used in interviews of the employment panel

US Bureau of Labor Statistics

HOUSEHOLD DATA

Table A-15. Alternative measures of labor underutilization

[Percent]

Measure	Not seasonally adjusted			Seasonally adjusted					
	Oct. 2019	Sept. 2020	Oct. 2020	Oct. 2019	June 2020	July 2020	Aug. 2020	Sept. 2020	Oct. 2020
U-1 Persons unemployed 15 weeks or longer, as a percent of the civilian labor force	1.3	4.5	3.8	1.3	2.1	5.0	5.1	4.6	3.8
U-2 Job losers and persons who completed temporary jobs, as a percent of the civilian labor force	1.4	5.5	4.5	1.6	8.9	8.1	6.4	5.7	4.8
U-3 Total unemployed, as a percent of the civilian labor force (official unemployment rate)	3.3	7.7	6.6	3.6	11.1	10.2	8.4	7.9	6.9
U-4 Total unemployed plus discouraged workers, as a percent of the civilian labor force plus discouraged workers	3.5	8.0	6.9	3.8	11.5	10.6	8.7	8.2	7.2
U-5 Total unemployed, plus discouraged workers, plus all other persons marginally attached to the labor force, as a percent of the civilian labor force plus all persons marginally attached to the labor force	4.1	8.8	7.7	4.3	12.5	11.3	9.6	8.9	8.0
U-6 Total unemployed, plus all persons marginally attached to the labor force, plus total employed part time for economic reasons, as a percent of the civilian labor force plus all persons marginally attached to the labor force	6.5	12.4	11.6	6.9	18.0	16.5	14.2	12.8	12.1

NOTE: Persons marginally attached to the labor force are those who currently are neither working nor looking for work but indicate that they want and are available for a job and have looked for work sometime in the past 12 months. Discouraged workers, a subset of the marginally attached, have given a job-market related reason for not currently looking for work. Persons employed part time for economic reasons are those who want and are available for full-time work but have had to settle for a part-time schedule. Updated population controls are introduced annually with the release of January data.

Alternative measures of labor underutilization

US Bureau of Labor Statistics

Average

Examples of mathematical concepts of “average”

Mean
Median
Mode
Mid-range
...

If using a mean, which mean?

$$\mathrm{AM}(x_1, \ldots, x_n) = \frac{1}{n}(x_1 + \cdots + x_n)$$

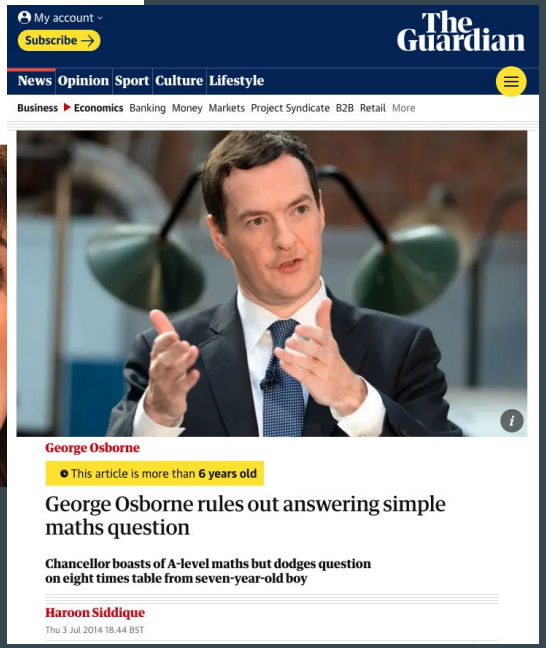
$$\mathrm{GM}(x_1, \ldots, x_n) = \sqrt[n]{|x_1 \times \cdots \times x_n|}$$

$$\mathrm{HM}(x_1, \ldots, x_n) = \frac{n}{\frac{1}{x_1} + \cdots + \frac{1}{x_n}}$$

Chair	*[...] if "good" requires pupil performance to exceed the national average, and if all schools must be good, how is this mathematically possible?*
Gove	*By getting better all the time.*
Chair	*So it is possible, is it?*
Gove	*It is possible to get better all the time.*
Chair	*Were you better at literacy than numeracy, Secretary of State?*
Gove	*I cannot remember.*

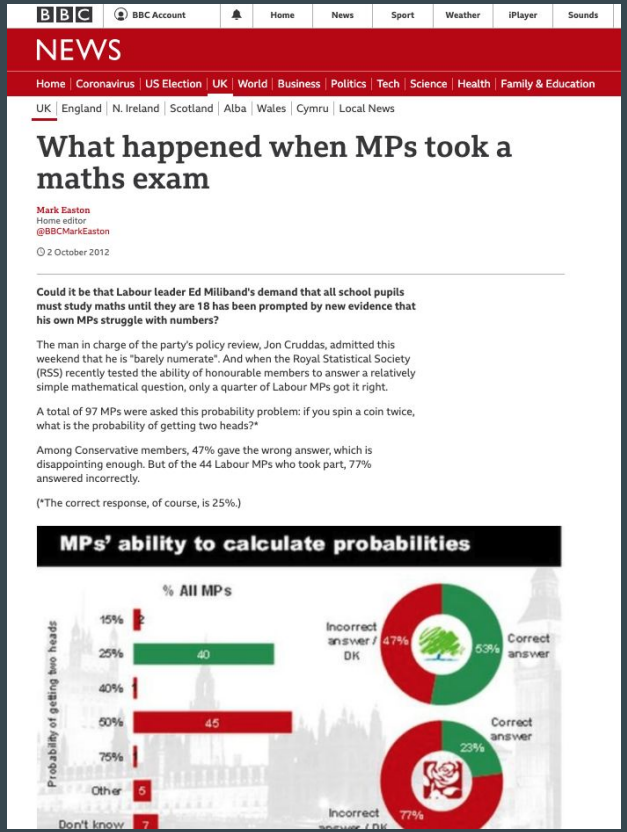
Michael Gove and the “average”

Hansard



Politicians refusing to answer maths questions

[Independent]
[The Sun]
[Evening Standard]
[The Guardian]



BBC | BBC Account | Home | News | Sport | Weather | iPlayer | Sounds

NEWS

Home | Coronavirus | US Election | UK | World | Business | Politics | Tech | Science | Health | Family & Education

UK | England | N. Ireland | Scotland | Alba | Wales | Cymru | Local News

What happened when MPs took a maths exam

Mark Easton
Home editor
@BBCMarkEaston

2 October 2012

Could it be that Labour leader Ed Miliband's demand that all school pupils must study maths until they are 18 has been prompted by new evidence that his own MPs struggle with numbers?

The man in charge of the party's policy review, Jon Cruddas, admitted this weekend that he is "barely numerate". And when the Royal Statistical Society (RSS) recently tested the ability of honourable members to answer a relatively simple mathematical question, only a quarter of Labour MPs got it right.

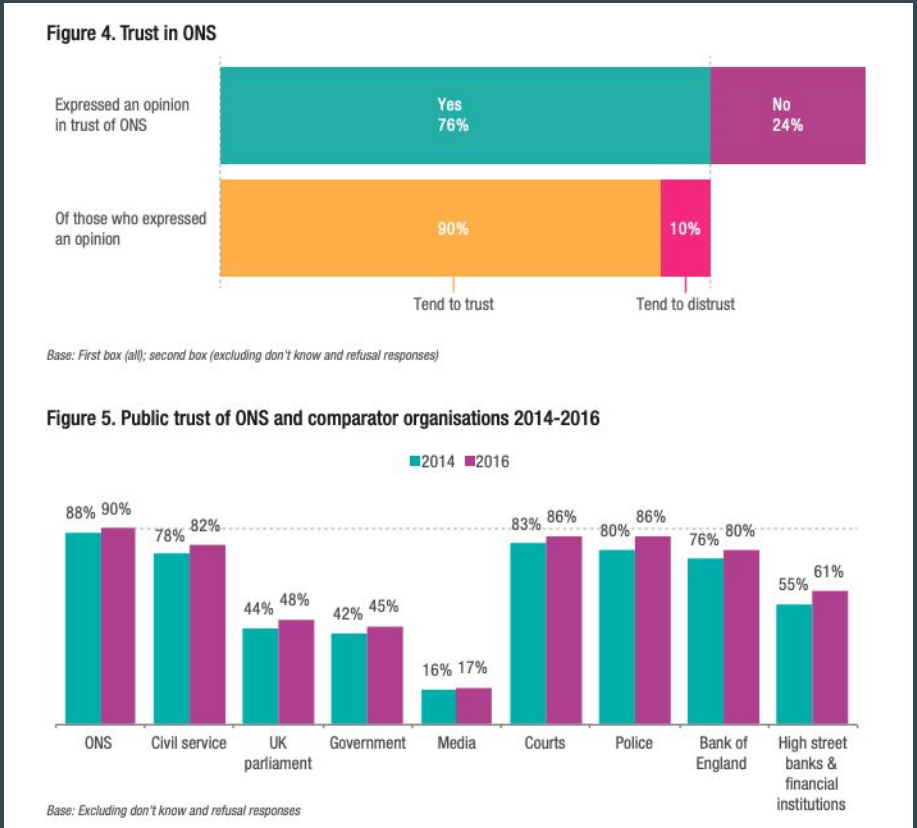
A total of 97 MPs were asked this probability problem: if you spin a coin twice, what is the probability of getting two heads?*

Among Conservative members, 47% gave the wrong answer, which is disappointing enough. But of the 44 Labour MPs who took part, 77% answered incorrectly.

(*The correct response, of course, is 25%.)

MPs and probability

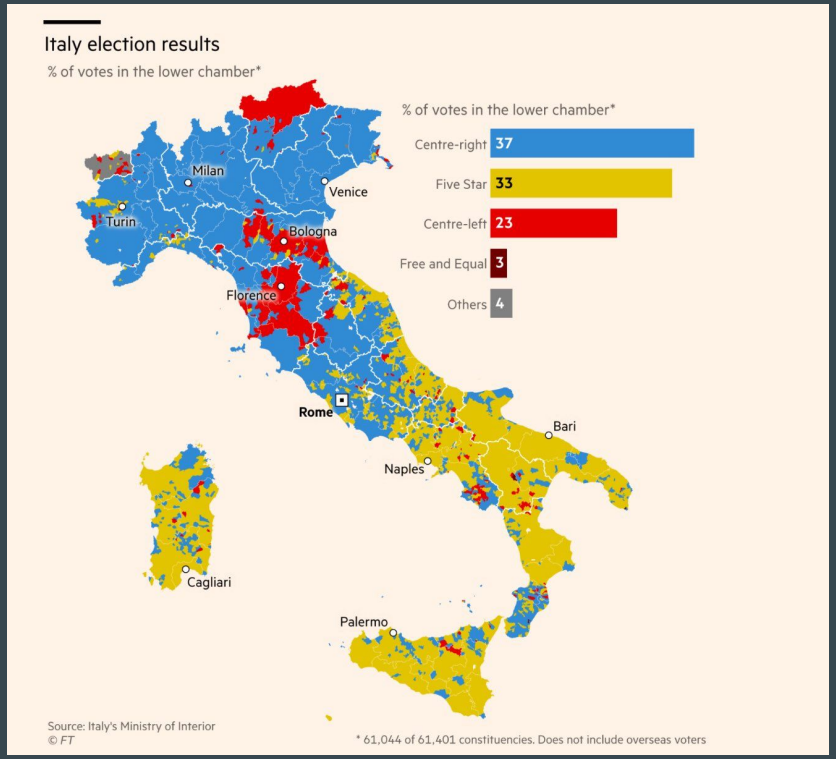
BBC



Trust in UK National Statistics (2016)

[NatCen]

Data journalism



Electoral maps tell a story

[Financial Times]
[Giuseppe Sollazzo]

“The map is not the territory“

(Alfred Korzybski, Science and Sanity, Wikipedia)

How The Washington Post Estimates Outstanding Votes for the 2020 Presidential Election

John Cherian, Lenny Bronner

October 2020

1 Introduction

On November 3rd, 2020, Americans will be closely watching The Washington Post as results pour in from across the country.[1] Making sense of election results as they arrive is no simple task. Careful observers of previous elections recognize that early leads do not necessarily translate to ultimate victories. The same patterns often seem to recur: In some states, a Democratic-leaning early vote is buried by a tidal wave of Republican election day votes. In others, a Republican lead is slowly whittled down as slower-counting and Democratic-leaning metropolitan counties report their tallies. To address this, The Washington Post, in collaboration with our friends at Decision Desk HQ/0ptimus Analytics, has developed a election night model that will make sense of live vote totals by delivering *reliable* inferences regarding what those early votes imply about the final results. For this election in particular, we think that our model is vitally important. It bridges the gap between what the results are showing at any given moment and the likely true underlying result.

Based on the votes reported, our model will estimate three quantities for each state: The overall turnout, the number of Democratic votes, and the number of Republican votes. For each of these outputs, we want to display the uncertainty in the prediction via a prediction interval. Accurate estimation of these intervals is critically important. An overly confident forecast could mislead readers relying on these figures, while being too conservative runs the risk of not being honest about what we know.

We'd like to provide a technical overview of how and why this model can provide forecasts and prediction intervals for unseen votes in each state.

Replicable journalism

Washington Post

Predictive journalism?

Independent

“Since when is it our job to predict?”

Mona Chalabi, Guardian US Data Editor

Today in Focus

Formerly solid Republican states

Usually solid Republican territory, these reddish states look like they could be in play for Biden in 2020. If he can win one of the three, it could be curtains for Trump.

All to Biden | All to Trump

Arizona
11 electoral votes

2000 | 2016 margin + 3.5 | 2020 polling + 3.7

Biden's best chance to invade the Sun Belt. Biden's relative strength with older voters could help him in a state full of politically active retirees and the Democrats have fielded a strong Democratic Senate candidate this year in former astronaut Mark Kelly.

Biden | Trump

Georgia
16 electoral votes

2000 | 2016 margin + 5.1 | 2020 polling + 1.5

The state voted for southern favourite Bill Clinton in 1992 but since then has gone reliably Republican in national elections. Trump will hope for solid rural turnout to help him over the line here, away from the Democratic votes in Atlanta.

Biden | Trump

Texas
38 electoral votes

2000

The last time Democrats wo making this a solidly red stat race is looking much closer i 2016 vote share was the Repu presidential performance he a combination of coronaviru national disasters hasn't help

Biden

Important swing states

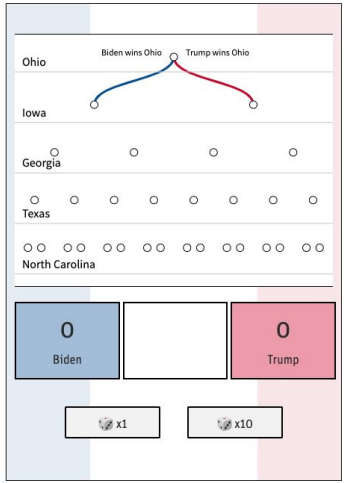
These states have been trending Republican in recent presidential elections, but Biden could put them back in play for t Democrats in 2020.

All to Biden | All to Trump

Florida
29 electoral votes

Iowa
6 electoral votes

North Carolina
15 electoral votes



US presidential election 2020

Biden vs Trump: live results 2020

LAST UPDATED YESTERDAY

522 out of 538 electoral college votes called

270 to win

Donald Trump REPUBLICAN 232

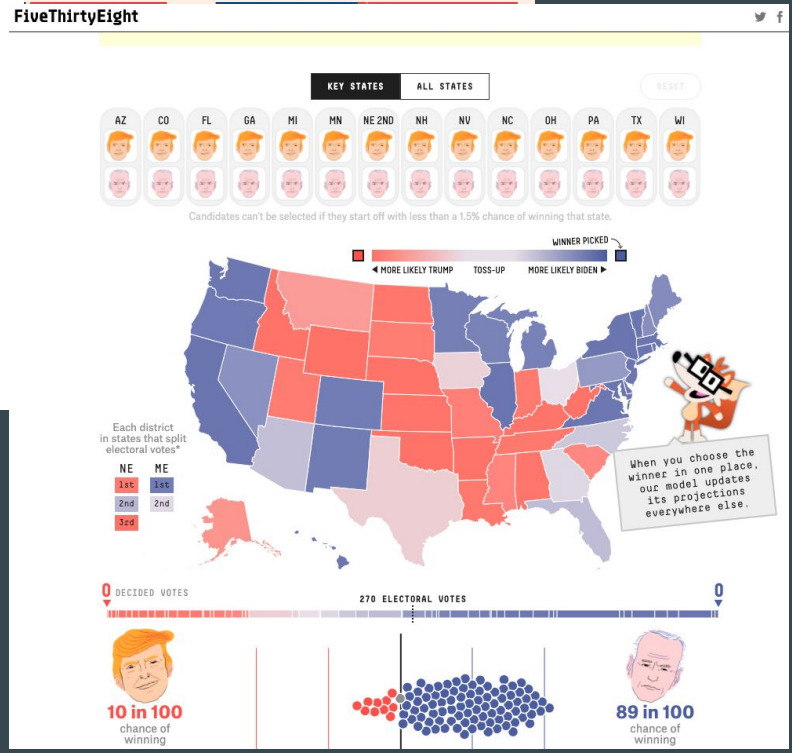
423 out of 435 races called

House 197

Senate

33 out of 35 races called

47 Outgoing Senate 53



Uncertainty in data visualization

[FiveThirtyEight]
[Financial Times]
[New Statesman]
[The Guardian]

Letting readers explore their (lack of) knowledge

Berliner Morgenpost (left)
Berliner Morgenpost (right)

Data in the pandemic

What data?

Number of COVID-19 positives

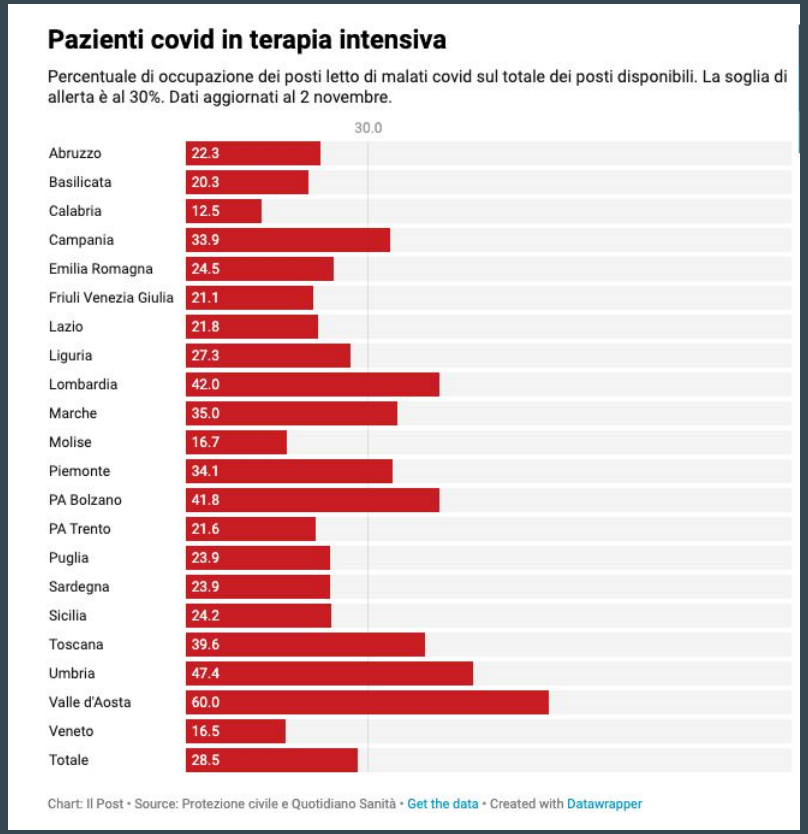
Number of tests

Number of admissions

Number of cases in intensive therapy units

Number of hospital beds available

Number of deaths



Anche questo dato, seppur molto affidabile, presenta dei problemi, soprattutto perché non è sempre chiaro quanti posti di terapia intensiva siano disponibili per ciascuna regione. Dall'inizio dell'emergenza, infatti, molti ospedali hanno aumentato il numero di posti letto in rianimazione, spesso inglobando reparti finora considerati di terapia "sub intensiva", altre volte riconvertendo reparti con ventilatori idonei. Sono i posti chiamati "attivabili".

Counting intensive therapy unit beds

Il Post

Public Health England has changed its definition of deaths: here's what it means

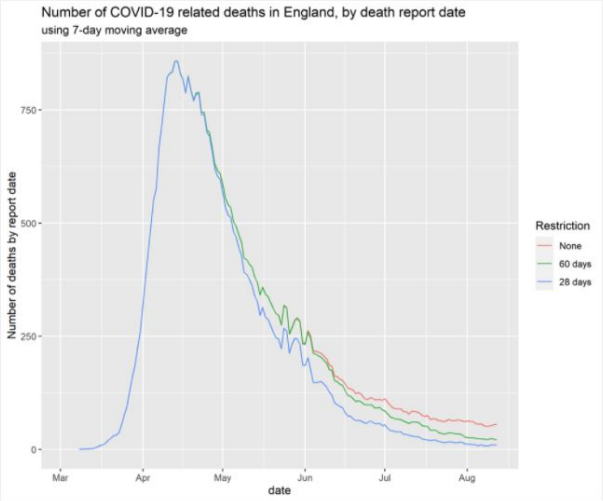
August 12, 2020

Carl Heneghan, Jason Oke

Public Health England has changed its definition of deaths. The new definition is now death in a person with a laboratory-confirmed positive COVID-19 test and died within (equal to or less than) 28 days of the first positive specimen date will now be reported

The new data can be accessed here.

What this means is that England has had 36,695 deaths using this definition as opposed to the previous reported 42,072 – a difference of 5,377.

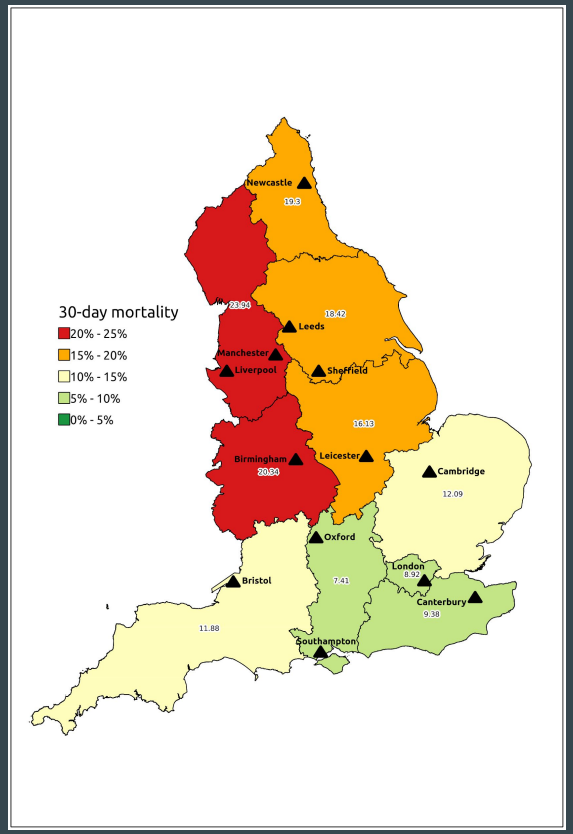


Changing the definition of death

Centre for Evidence Based Medicine, University of Oxford

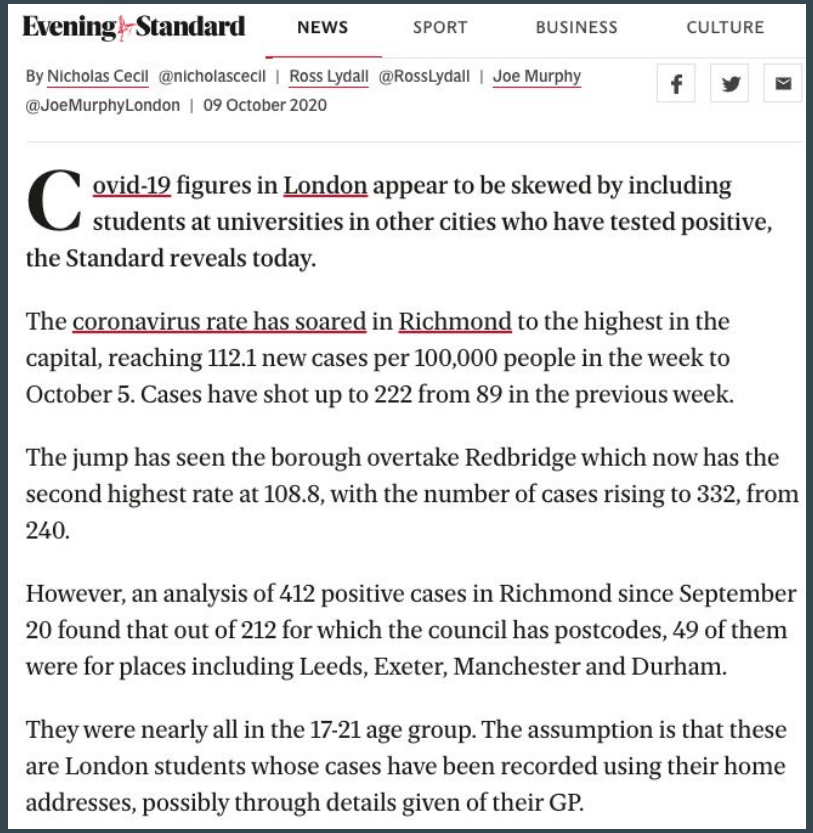
“Miscarriage” or “death”?

Independent
The Lancet



Surgery mortality rates

PLOS One



Evening Standard NEWS SPORT BUSINESS CULTURE

By Nicholas Cecil @nicholascecil | Ross Lydall @RossLydall | Joe Murphy @JoeMurphyLondon | 09 October 2020

Covid-19 figures in London appear to be skewed by including students at universities in other cities who have tested positive, the Standard reveals today.

The coronavirus rate has soared in Richmond to the highest in the capital, reaching 112.1 new cases per 100,000 people in the week to October 5. Cases have shot up to 222 from 89 in the previous week.

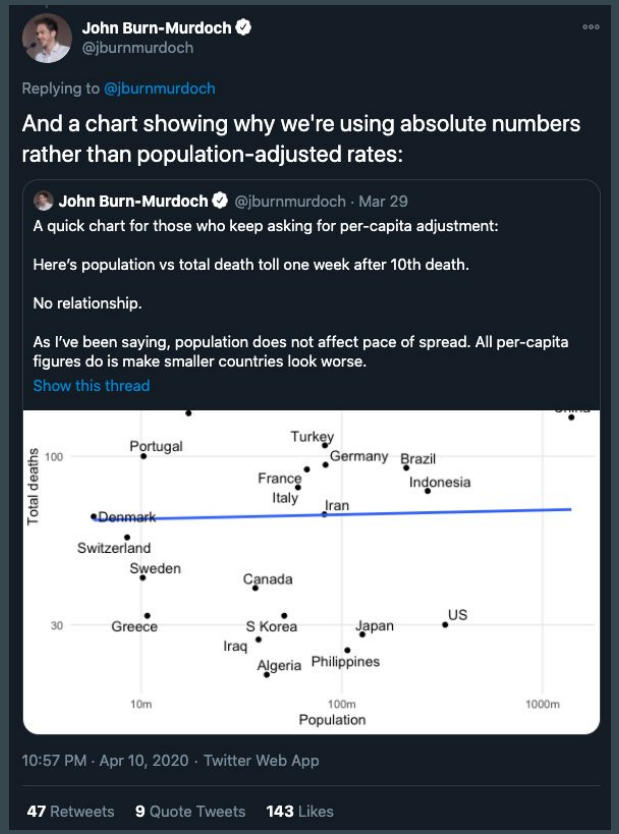
The jump has seen the borough overtake Redbridge which now has the second highest rate at 108.8, with the number of cases rising to 332, from 240.

However, an analysis of 412 positive cases in Richmond since September 20 found that out of 212 for which the council has postcodes, 49 of them were for places including Leeds, Exeter, Manchester and Durham.

They were nearly all in the 17-21 age group. The assumption is that these are London students whose cases have been recorded using their home addresses, possibly through details given of their GP.

Attributing cases to the right context

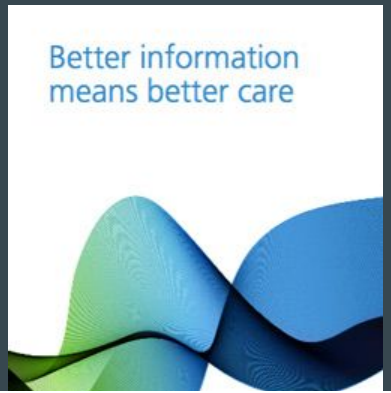
Evening Standard



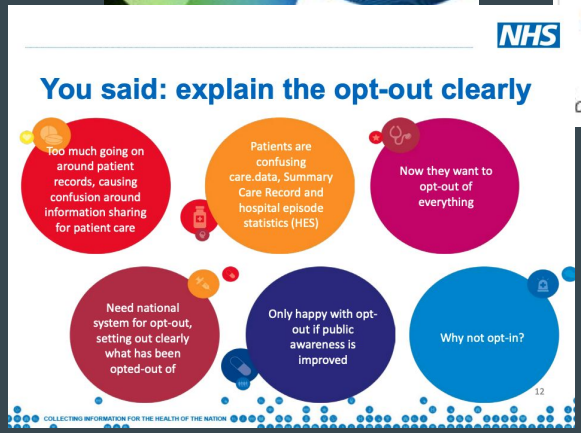
Data presentation is a conscious choice

[John Burn-Murdoch]

Trust and engagement


Better information
means better care

thebmj
covid-19
Research
Education
News & Views
Editor's Choice
What can we salvage from care.data?
BMJ 2016 ; 354 doi: https://doi.org/10.1136/bmj.i3907 (Published 14 July 2016)
Cite this as: BMJ 2016;354:i3907


NHS
You said: explain the opt-out clearly
Too much going on around patient records, causing confusion around information sharing for patient care
Patients are confusing care.data, Summary Care Record and hospital episode statistics (HES)
Now they want to opt-out of everything
Need national system for opt-out, setting out clearly what has been opted-out of
Only happy with opt-out if public awareness is improved
Why not opt-in?
12
COLLECTING INFORMATION FOR THE HEALTH OF THE NATION

Article
Related
ona Godlee, editor in chie
News
NHS England sets out the next steps of public awareness about care.data
16 October 2013
Digital
Patient safety
The Care.Data Programme closed in 2016

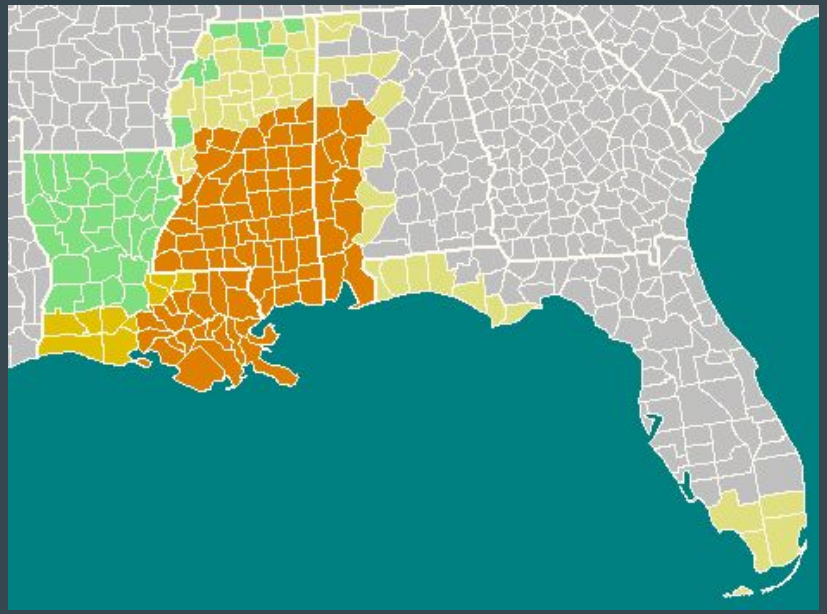
Care.data

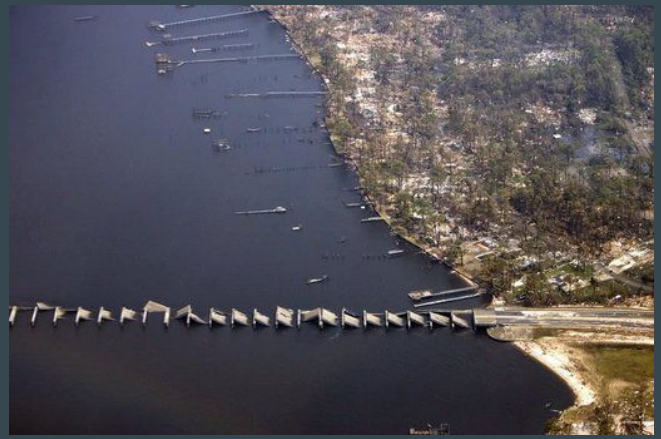
Open data

Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike

Open Data Definition

Open Knowledge Foundation





Hurricane Katrina

[Federal distaster areas, FEMA]
[EPA]
[US Coast Guard]
[Executive Office of the USA President]

“Just imagine the effect that an army of armchair auditors is going to have on those expense claims”

David Cameron (2009)

“Armchair auditors”

BBC News

Open data—public information and shared data from private sources—can help create $3 trillion a year of value in seven areas of the global economy.

The Open Data Economy?

McKinsey

Vast data reserves from Defra are set to transform the world of food and farming in the single biggest government data giveaway the UK has ever seen, Environment Secretary Elizabeth Truss will announce today.

Outlining her vision for the future of British food, farming and the natural environment, the Environment Secretary will say that over the next year, virtually all the data Defra holds—at least 8,000 sets—will be made freely available to the public, putting Britain at the forefront of the data revolution.

Harnessing this open data will help food and farming achieve its full potential and enable the UK to become a one-nation economy, where the productivity of the countryside will be brought up to the level of our towns and cities.

Environment Secretary Liz Truss announces a dataset release target (2015)

Gov.Uk

“Opening our data was at most half the issue. More important was the effect it had on open working among colleagues across many organisations in the DEFRA group, and with a range of external partners from universities to charities to commercial organisations.

That led to interesting discussions about how open government relates to accountability, and also to user-centred design. It took us into how we can be more open as individuals, sharing a bit more about ourselves as a way of making connections with others: building respect and trust.

Open has all sorts of dimensions. It encompasses being open to challenge, open to new experiences, open to new ideas and to difference”.

Permanent Secretary Clare Moriarty on Open Data (2019)

Wikipedia
FDA

A few final points

Communicating with data

1. ...is difficult
2. ...requires users to challenge assumptions
3. ...requires understanding that common sense concepts may differ from their definition in data
4. ...requires a multidisciplinary approach
5. ...is badly impacted by lack of scientific education in the public
6. ...needs data stewardship/leadership

Thank you

Giuseppe Sollazzo
@puntofisso

References, slides, and transcript (soon) at

http://www.puntofisso.net/battistini

#datibenecomune

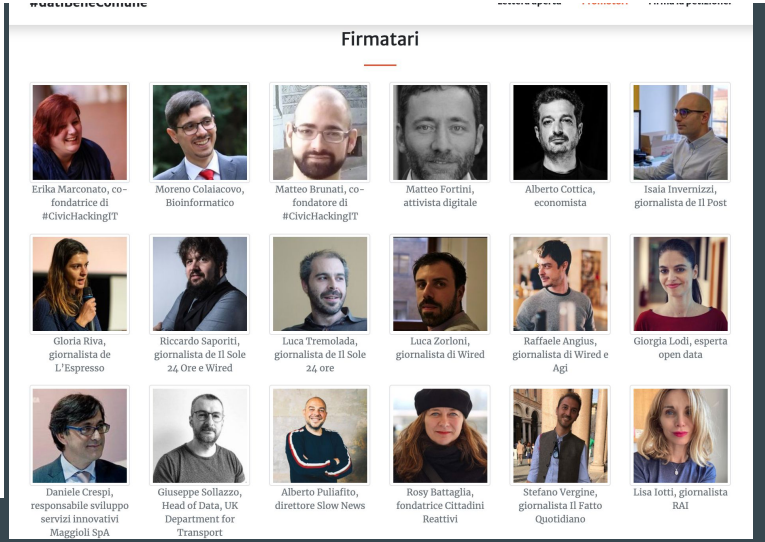
AMA 12 January, 5pm, Rome time

Chiediamo al Governo dati aperti e machine readable sull'emergenza Covid-19 per monitorare realmente la situazione e poterla gestire al meglio. Siamo 38238 firmatari e 154 organizzazioni promotrici.

La nostra lettera aperta

Viviamo una grave crisi. La società civile italiana, una delle più mature e competenti del mondo, è pronta a supportare le Istituzioni nel farvi fronte.

Promotori

nData · Transparency International Italia · actionaid · scienzainrete · info.nodes · ALTROCONSUMO · ALC Associazione Luca Coscioni · cittadini reattivi · COVID19Italia.help · parlamentwatch.it · openpolis Sicilia · The good Lobby · Legambiente · Medici Senza Frontiere · GIMBE Evidence for Health · Slow News. · openpolis · Unione degli Studenti · Priorità alla Scuola · LINK Coordinamento Universitario · Rete della Conoscenza · ANBI · REPORT · F.I.Bio.

Firmatari