

Hi everyone, we're Matt Jukes and Rob Fry from the UK Office for National Statistics. We create data visualizations and data explorers based on UK official data. Ask us anything!

UK_{ONS}¹*andr/ScienceAMAs*¹

¹Affiliation not available

April 17, 2023

Abstract

Matt Jukes: I have been called the 'Ian Holloway of Digital Government.' You may have heard of the ONS website's reputation, or experienced it for yourself - it's my job to deliver a new site that changes those perceptions. I promote open source, open data and am a little militant with it comes to user research. We named our publishing system Florence, as in Nightingale, but over time Magic Roundabout characters ruled the roost in our app naming strategy. We even have a 'Chris Giles test' for the site. Rob Fry: I am a self proclaimed data vis geek working at the UK's Office for National Statistics. I'm a trained statistician and love all things D3. I look for ways to make government data useful or relevant to people's everyday lives. With my team I've created a pension calculator to see how long your pension pot will need to last, a quiz to challenge people perceptions of their own area, and lots of maps. I post all of this on Visual.ONS which is the ONS sister website aimed at non-experts just interested in data stories. Here's proof that it's us. We're here to talk with you about UK government open data, open source, data visualization, or anything else. Ask us anything! That's the end of our AMA! Thanks guys for all your questions, really interesting discussions and hope it was useful!

[REDDIT](#)

Hi everyone, we're Matt Jukes and Rob Fry from the UK Office for National Statistics. We create data visualizations and data explorers based on UK official data. Ask us anything!

UK_ONS [R/SCIENCE](#)

ABSTRACT

Matt Jukes: I have been called the 'Ian Holloway of Digital Government.' You may have heard of the ONS website's reputation, or experienced it for yourself - it's my job to deliver a [new site](#) that changes those perceptions. I promote open source, open data and am a little militant with it comes to user research. We named our publishing system Florence, as in Nightingale, but over time Magic Roundabout characters ruled the roost in our app naming strategy. We even have a 'Chris Giles test' for the site.

Rob Fry: I am a self proclaimed data vis geek working at the UK's Office for National Statistics. I'm a trained statistician and love all things D3. I look for ways to make government data useful or relevant to people's everyday lives. With my team I've created a [pension calculator](#) to see how long your pension pot will need to last, a [quiz](#) to challenge people perceptions of their own area, and [lots of maps](#). I post all of this on Visual.ONS which is the ONS sister website aimed at non-experts just interested in data stories. [Here's proof that it's us.](#)

We're here to talk with you about UK government open data, open source, data visualization, or anything else. Ask us anything!
That's the end of our AMA! Thanks guys for all your questions, really interesting discussions and hope it was useful!

[READ REVIEWS](#)

[WRITE A REVIEW](#)

CORRESPONDENCE:

DATE RECEIVED:
September 24, 2015

DOI:
10.15200/winn.144301.12432

ARCHIVED:
September 23, 2015

CITATION:
UK_ONS , r/Science , Hi everyone, we're Matt Jukes and Rob Fry from the UK Office for National Statistics. We create data visualizations and data explorers based on UK official data. Ask us anything!, *The Winnower* 2:e144301.12432 , 2015 , DOI: [10.15200/winn.144301.12432](#)

© et al. This article is

Whats the worst misinterpretation of data that you have come across?

How much influence does the government of the day have on the operations of the ONS? Do you come under pressure for example to add or remove certain items from the RPI or CPI baskets in order to tell the inflation story the government wants to tell and support its policy aims?

[confuzzledfather](#)

First part is difficult for us to answer, sometimes politicians use our stats to suit their needs, but the UK Statistics Authority acts as a watchdog to challenge misuse of official statistics. The answer to the second part is no influence. ONS stats are independently produced in line with the Code of Practice for Official Statistics.

Can you remember a time where the use of statistics dramatically changed your opinion on something? A scenario where the stats disproved many of your preconceived notions about a topic?

[frostickle](#)

I'm not sure that statistics have ever dramatically changed my opinion on something, but I'm continually surprised by statistics.

distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and redistribution in any medium, provided that the original author and source are credited.



One instance that immediately springs to mind is a short piece we recently worked on looking at how the UK's welfare budget was spent. Whilst I was putting this together I was shocked at how wrong I was at understanding what Government spends the welfare budget on. I won't give you the answers incase you want to have a guess yourselves (<http://visual.ons.gov.uk/welfare-spending/>) but I was surprised at the amount that we spend on pensions and how little we spend on unemployment benefits. Whilst looking at this we also saw how much that has share increased over the past years... presumably as a result of an ageing population. Rob

To what extent do you think the ONS should be adding value to releases by interpreting the data, especially when the data could be interpreted multiple ways? Or should the ONS primarily be releasing the raw data (in an open format) for the wider statistical world to interpret?

And as a independent organisation, set-up and funded by the government how difficult do you find it to release interpretations of data that are critical of the present day government?

[beetex](#)

I think we should be doing both. We know from user research that we have a wide and varied audience and have been developing a strategy to ensure we're not just providing raw data to expert users. We're also using our statisticians who know the data intimately to provide unique insight which can be valuable to interpret data correctly. Our stats are used to inform debate and its important for us to remember that being neutral is not the same as saying nothing. We should be explaining the intricacies of our data. Open data is something we think is important but getting people to view raw data sets and then use them is hard – we've described data vis as the pretty sister of open data and we feel that a lot of the work we're doing on <http://visual.ons.gov.uk> helps open up our data to a much wider audience compared to those willing to download the data and explore it themselves.

We inform the debate not create the debate so I would say it's unlikely we'd be critical of government, rather we'd aim to provide context and remain neutral, highlighting other factors to be considered so people can form their own opinion.

What is your favorite statistical anomaly?

[rhiever](#)

I'm not sure I have a favourite, but one anomaly that I always have trouble believing is the fact that for every one hundred females born there are 105 male (on average). This doesn't make sense to me, but apparently mother-nature has made an adjustment <http://www.livescience.com/33491-male-female-sex-ratio.html>

We also teach an introduction to data visualisation course across ONS and UK government and one of my favourite statistical anomalies we cover as part of that is Anscombes Quartet. For those of you who aren't familiar with this they are 4 different data tables with 11 pairs of x & y values, they're all different, but remarkably they have the same mean, median, variance, linear regression line. The shock comes when you visualise these data sets, they all look very different. It's a great example that helps convey importance of visualisation.

<http://bl.ocks.org/alansmithy/07218980d8bac3689104>

I am a great fan of the new ONS stuff. IS there any plan for you to take on more data presentation functions from the rest of govt? Some specialist data - like FE - is really tricky and (post-GDS) is no longer presented with context (you have to compare first releases to other first releases, not the most recent data - stuff like that).

p.s. yes, it's me. <http://digitalbydefault.com/2014/12/21/dont-look-back-in-anger/> I stand by 2009.

2010 saw a load of pointless UX nobbling,

p.p.s. edited for clarity. to make it clear that, yes, the GDS is still giving me routine coronaries.

[xtophercook](#)

Rob here, and thanks for your kind comments on the recent work. There are no plans at the moment to extend our work to cover the data presentation functions from the rest of government. But we do run a basics of data visualisations course which we run across UK government...introducing them to the basics of chart design, the importance of context, the importance of succinct messaging. We've run this for some time, but appreciate it's having a limited impact!

Which one of you is Gareth and which is Tim?

[crayish](#)

I am.

If you could do one thing to improve the capturing of statistics in the UK what would it be?

[TheRealDave24](#)

To capture data more naturally rather than having to go out and seek it. For instance there's some great work in using more admin data at the moment. It's important to maintain confidentiality in this area but if we get it right there's so much potential.

What is the best/most interesting statistic about the UK you can give us?

[FelonyT1](#)

I couldn't limit it to one statistic, but if I'm ever at a house party here are some of the statistics I'd roll out... A female born today has a 1 in 4 chance of reaching 104. We spend more on lottery tickets, than newspapers. And pilots and flight engineers is the top earning occupation. I don't get invited to many house parties... Rob

Is there any plan to provide a top-down entry point to up-to-date visualisations with current data sets? I like the look of the visual.ONS entries, but they're couched like a blog with interesting tidbits - I (and I'm sure others) would love to see all relevant data presented in such a way that you can easily drill down to a relevant topic (or better yet, search for datasets and have visualisations brought back for you).

Think something along the lines of wheredoesmymoneygo but for a selection of disparate datasets (in terms of presentation / drilldown) and WolframAlpha in terms of putting in search terms and getting relevant graphed data back.

[Fingerbob](#)

Interesting question. There's work lots of work going on the new Beta ONS website – you can have a look at it here - <http://beta.ons.gov.uk/>. Here you can drill down and find the data by topic and for each topic we start with a data visualisation and a key message for each dataset.

One of the reasons for separating out Visual.ONS from the main website was the wide range of users we have. From expert users who just want the data, to members of the public who may be interested in how statistics influence their lives and the economy. There's an interesting blogpost here which attempts to explain some of the differences - <http://digitalpublishing.ons.gov.uk/2015/01/15/official-data-new-light-introducing-visual-ons/>

First, I'm a huge fan of your work! I'm a Brit living in the USA and it's a great way to get an "objective" view of what's happening back home: A few questions if I may, no need to answer them all if you don't want to!

1) Within the ONS as an institution, how seriously is data visualization and communicating with data taken? Is it viewed as simply ancillary to the core data collection and compilation work or as a "first class activity" in and of itself?

2) How do you chose which projects to work on / issues to tackle? Are they driven by public interest, policy needs, current affairs, your own curiosities or something else?

3) Do you get much management or "political interference" when it comes to portraying information that doesn't show either the ONS or UK government in a flattering light? (e.g. missing or poor data or poor/variable service delivery or quality of life across the country)

4) How do you feel about incorporating official data sources from outside the ONS into your work? E.g. data from other parts of government or international agencies like The World Bank?

5) How do you feel about incorporating "unofficial" data sources into your work? For example, there are now several crowdsourced CPI series available (I believe you guys outsource CPI collection now?) and new methods for estimating socio-economic activity e.g. satellite imagery, sensor data, anonymized cell phone metadata, internet / social media data.

Thanks!

[bantamug](#)

Thanks, nice to hear our work helps on the other side of the pond. Here are a few answers, hope they are helpful: 1) we are very fortunate at the ONS, there is a lot of support across the business when publishing statistics and we work closely with lots of different teams to communicate the data as best we can, teams are always engaged in visualising their statistics 2) yes, all of the above. we are still working out what works and what doesn't, each project is a learning curve. We rate on strategic interest, topicality, user persona as well as a list of other criteria. 3) we are independent so don't have political influence. 4) We work closely with the GSS (government statistical society) and our focus is mainly on our stats. We do use statistics from across Europe when relevant for projects. 5) we are interested in "unofficial" data but would be really careful about using it, we would need to understand the samples, methodology and trustworthiness of the data before using it. One of the really interesting projects happening in ONS at the moment is the big data project which is looking into the use of some of the sources you've mentioned. We don't outsource our CPI collection

As a relatively fresh-faced reporter trying to get into data journalism, what resources would you recommend starting with?

[Newbeams](#)

There's obviously loads of resources available, but the two "bibles" we've used extensively in our team are Alberto Cairo's the Functional Art – this gives a great overview of the basics of data visualisation, and Scott Murray's Interactive Data Visualisation for the Web (O'Reilly) which is great introduction for getting into data visualisation programming (primarily d3).

If you're a trained reporter then you've already got the majority of skills you need for data journalism: an understanding of news values, objectivity, balance, and a nose for a story. To become a more refined data journo I would highly advise becoming more proficient on Excel and practicing interrogating data on Excel. Learn to ask the right questions, always question the data at an international level, regional level and personal level. Look at how it has changed over time, and look for anomalies. Most of all, get experience. Newsrooms aren't the only places crying out for data journos, PR firms need them, as do charities and the public sector. Also, look at other people's work so

you grow to learn what great data journalism looks like, you can learn a lot from emulating others.

Statisticians have taken such a bad beating in the United States lately. People seem to generally feel that statistics are just made up "14% of all stats are made up, 90% of all people know that", or are viewed as being misleading (538 vs unskewed from 2012).

What do you think could be done to change this perception?

[e6c](#)

We are independent, meaning we aren't under control of any ministers and separate from the rest of government, this was partly to gain trust. We also publish our methods and how we collect data. Not sure this is the answer but may help to gain trust.

How do I get a job like yours?

[SobanSa](#)

My advice would be to do a project in your own time, build some skills and show people what you can do. Then visit this page - <http://digitalpublishing.ons.gov.uk/jobs/> ;-)

Do you have any preferred real-time open data source? Every time I look for one I come up empty-handed

[ANTIVAX JUGGALETTE](#)

We don't have any real time data sources at ONS but this is a really good example
<https://www.gov.uk/performance>

How long does it take to complete a study, between the government asking for your analysis, to gathering the dataset, to running your analysis, to your final presentation and upload to the Internet? Do you have control over the gathering of the dataset?

[Meetthekingofwhales](#)

It really depends on the piece of analysis being carried out. Our data comes from lots of different sources e.g. surveys, Census and admin data sources. The analysis time depends on the release, some intend to just present key points, others may have more cross-cutting/in-depth analysis. As for uploading, some are uploaded as datasets or statistical bulletins whereas others have accompanying data visualisations or infographics created by us. It varies greatly depending on lots of different factors but we have a release calendar for our regular pre-announced publications

<http://www.ons.gov.uk/ons/release-calendar/index.html>

Hi guys, I agree, open data and open source is great in government matters and its important that the general public are enabled to analyse the performance of the government of the day. However, there are few that take the time to dig through such data to draw impartial conclusions. What do you think can be (or is being) done by ONS to engage the public with its output? And what is the downside of such engagement?

[lagori](#)

A lot of the work we're doing on <http://visual.ons.gov.uk/> is precisely to engage the public with ONS statistics. Have a look. We focus on pieces that are topical, where ONS statistics can help shed light on an issue, or a personal aspect where people can see themselves in the data. In addition to this we encourage syndication - we're really happy if the media picks up our work and reuses it and gets our message out to a wider audience than we could achieve on our own. I don't think there are many

downsides that we've encountered so far although every media outlet will put their own spin on these messages.

Hi! Statistics student here! Could you talk about some of the more challenging/ interesting problems you have seen? Also, Bayesian or frequentist? :)

[datrandomgal](#)

Some of the memorable challenges I've had working with data stem from the fact that we have such a wealth of data in ONS. Some of our releases publish hundreds of tables and extracting key messages can be tricky.

How much does it get annoying that some data is for England, some for England and Wales, some for GB, some for the UK, and so on..? Do you have to constantly double check that you're talking about the right geographical/political entity?

[grogipher](#)

It's like any editorial check. Before we post a tweet or an article, we're always like, 'oh god, is it UK or England and Wales?!' and have to ring the statisticians to check. It's fine though, keeps up on our toes.

You have my dream job. How do I apply, and what are the qualifications?

[the_original_kiki](#)

Thanks, it is a good job! We currently have a number of vacancies in the wider team. Have a look here for latest jobs in our team - <http://digitalpublishing.ons.gov.uk/jobs/> Rob

I'm a normal voter without any real knowledge of statistics or maths. What is the best way for me to be informed about the state of the country from a non-partisan viewpoint and how much does your work at the ONS feed into this?

[SamKeeley](#)

I'd recommend reading content on our Visual.ONS site which is all aimed at the non-statistical experts. We keep our copy simple and our content topical. Also our ONS Twitter feed and our tweeting statisticians will keep you up to date on current issues in a plain and objective way. Externally, I highly recommend reading Fullfact.

What is the most unintuitive correlation you've ever noticed?

[XSplain](#)

We don't often report on correlation as it can be tricky to explain and we're not in a position where we would want to imply causation without good cause
https://en.wikipedia.org/wiki/Correlation_does_not_imply_causation. There are many amusing pages I've noticed being shared on social media such as <http://www.buzzfeed.com/kjh2110/the-10-most-bizarre-correlations#.pejRO0ykVX> which have a bit of fun considering what causation could be inferred from different correlations

What data do you which were collected but aren't?

[brainburger](#)

Cats and stats! #icanhazstatistics

Hi guys, i am a 27 yo from UK with a statistics degree and 6 years exp in waste management, local govt and waste water industry/environmental engineering... How do you go about getting a job at ONS?

[jimmy011087](#)

Best place to start is our current vacancies page where you can sign up for email alerts when new vacancies are added <http://www.ons.gov.uk/ons/about-ons/careers/vacancies/index.html> Find out more about careers at ONS here <http://www.ons.gov.uk/ons/about-ons/careers/index.html>

Why is visualisations spelt with a Z?

[AdamWrigley](#)

It's for our american friends ;-) Rob

I was thinking about grad school for biostatistics; can you tell me about your experiences in academia? Is the field of statistics hard to get into?

[spacebabes](#)

Sorry, I don't know anything about biostatistics, and although I studied statistics at university I haven't spent anytime in academia. Rob

Hi both - Matt, love the title: "Ian Holloway of Digital Government", that's fine praise!

Anyway my question is about D3 - I'm coming to the end of an introductory course on D3.js (lectured by the excellent Scott Murray and Alberto Cairo) and am keen to take it further, therefore I'm interested to know what else D3 is capable of. Given any data visualisation you might plan, is it possible (with sufficient skill/experience) to create using D3? What are the best resources to learn from in order to take the next steps into making some seriously impressive visualisations?

[kusasi2000](#)

Our interactive developers started in the same way...by reading tutorials and getting to grips with worked examples. Eventually we gained confidence and started to create our own products. The majority of our interactive/responsive content on <http://visual.ons.gov.uk/> uses d3 and it has become one of our essentials when creating an interactive. When developing your skills I'd recommend working on a project which really interests you and look around the internet for inspiration. Utilise sites like <http://stackoverflow.com/> where you run into problems and gradually you'll gain confidence and create some cool things. There's a fairly steep learning curve but it's rewarding when you've finally got a product working that you're proud of.

What is your education background?

[Dawsy](#)

We all have different backgrounds in the team, some of us come from a mathematical/statistical background, journalism, designers, coders etc. We all work together on various projects drawing on the skills needed for each project. I'm personally have a statistical background. Rob

How do you collect the data, and how do you do you test the integrity of your data?

I do like your stats, have had to quote them a number of times.

[trowawayatwork](#)

Fill yer boots! <http://www.ons.gov.uk/ons/guide-method/index.html>

I'm interested in vegan related stats but I'm having a difficult time finding UK statistics. I can only ever manage to find statistics based on the USA, could you point me in the right direction or do you know if any statistics for this are coming up?

It'd be great to know - percentage of vegans in the UK, across a period of time? How much water is saved and how many animals are spared by going vegan? The rate in which vegan/vegan friendly businesses are growing?

[vegan kong](#)

We don't produce vegan statistics at ONS. Although recently we held a census consultation where we asked what topics could be included in the next census. You may have a opportunity like this in the U.S. I'd highly recommend expressing your opinions. Also, pressure groups (if you're part of one) can lobby or pay for questions to be added to social surveys.

We have an annual survey of businesses, you may have something similar in the U.S. where you could look through the dataset for vegan businesses.

what are the main best practices that you follow when creating data visualisations?

could you recommend any resources or documentation?

Marc

[Marc_BT](#)

Our website <http://style.ons.gov.uk> aims to cover all the more common mistakes which can be made when creating data visualisations, it's based on our own experience, but also on historical good practice and the principles of human perception (https://en.wikipedia.org/wiki/Gestalt_psychology). We're hoping to add more content soon on maps and other topics... I think there are several other posts where we've recommended resources, Alberto Cairo and Edward Tufte are a good starting point.

What do you think of Canada getting rid of the long form census? Is that something the UK would consider doing or is Canada out of step with its fellow OECD countries in doing so?

[Suivoh](#)

We've just consulted publicly about the topics for our 2021 census. I think the office is moving towards more online forms and using admin data

Can you recommend any background reading material to learn more about visualization and information representation?

I'm currently reading the books by Edward R. Tufte. Are there any other books like those that you'd recommend?

[thread-safe](#)

Alberto Cairo <http://www.thefunctionalart.com/> is also a good start. I'd recommend following members of the data vis community on social media too as there are some great discussions on good/bad vis you can learn a lot from.

Why is it so hard to get Scottish and Northern Irish data? Sometimes it's easy to get the data for all regional levels across the whole UK, sometimes its only England and Wales, sometimes

it's just England. Don't even get me started about how hard it is to get unified spatial data. Is this a vast conspiracy against Scottish researchers?

[bad_jew](#)

ONS is responsible for collecting data for England and Wales only in a number of areas. Scottish and Northern Irish data are collected by the Scottish Government and the Northern Ireland Statistical Research Agency. Often we will then publish full UK data using figures from them, e.g. Census, but that's why you might find it more difficult to find Scotland and NI-specific figures on our site.

I read most of the title and here's my question. Do you prefer David Brent or Michael Scott from the US version? Thanks.

[whilweaton](#)

Michael Scott of course. Everyone should follow his wise words in life

<http://www.buzzfeed.com/lyapalater/the-wisest-things-michael-scott-ever-said#.paRpGWzjN>

How do I get started with D3 without any knowledge of JS?

[datavizzy](#)

Incredibly difficult. You need some basic web developer skills (HTML, CSS & JS) to make the most of D3. We'd recommend sites such as <https://www.codecademy.com/> or <http://www.w3schools.com/> as a starting point on this. Scott Murray's book on Interactive Data Visualization for the Web - An Introduction to Designing with D3 is a also a great starting point -

<http://shop.oreilly.com/product/0636920026938.do>

Why isn't there ever anyone at the ONS desks in airports?

[willz1208](#)

If you're talking about the International Passenger Survey it's because they work on shifts and you've just been unlucky!

Missed the ama:

What advice would you give to someone who wants a career working with data, such as yourselves?

[lunchtimeredit](#)

Best place to start is our current vacancies page where you can sign up for email alerts when new vacancies are added <http://www.ons.gov.uk/ons/about-ons/careers/vacancies/index.html> Find out more about careers at ONS here <http://www.ons.gov.uk/ons/about-ons/careers/index.html>

What are the most interesting trends in data visualization these days?

[printculture](#)

One of the most interesting approaches I've seen recently was from the New York Times where they've used geolocation to push readers deeper into the data. Ie giving them personalised maps for their area, and personalised analysis. It's something I'd love to explore at some point -

<http://www.niemanlab.org/2015/05/the-upshot-uses-geolocation-to-push-readers-deeper-into-data/> Rob

Have you got any plans to try and produce combined data sets with your Scottish and Northern Irish colleagues? For example, there's no logical reason why the Small Area Population

statistics should be presented as different releases (understand that internally they may need to be worked separately, but having the end presentation as one would be very useful!)

[pauljefferson](#)

Can't comment on specific outputs but as we noted above, ONS is responsible for collecting data for England and Wales only in a number of areas. Often we will publish full UK data using figures from Scottish Government and NISRA. For updates on any new developments head here

<http://www.ons.gov.uk/ons/guide-method/index.html>

Your economic data is generally split by deciles or quintiles. This hides the realities at the extreme ends of the distribution. e.g. income of the top and bottom 0.5% and 0.1%. Same for taxes, benefits, and others. Do you plan to provide separate data for these extremes?

[LogJamAxe](#)

Many of these data are already available (albeit sometimes only on request) and we would always seek to provide the most useful analysis in our publications. Of course, in any sample-based survey, the robustness of figures reduces as the sample gets smaller, so we have to perform that balancing act when we produce data. Email info@ons.gsi.gov.uk if you can't find specific data you're after.

How much of the NHS budget goes to the private sector?

[MarvinFrohike](#)

You'd need to contact the Treasury on that one, sorry.

Should I move to the UK?

[Dear_Fuck_WHY](#)

Depends where you are now :)

How often are you surprised by what you discover? Is it every day, rarely, never? Also, do you ever "tweak" the statistics so they work better in the representation?

[vesparider](#)

We're often surprised by what we see and we like other people to challenge their perceptions too. Our "How well do you know your local area?" quiz was developed to do just this.....

<http://www.theguardian.com/news/datablog/ng-interactive/2014/jul/21/seven-questions-that-will-tell-you-how-well-you-know-your-local-area>. Have a go and let me know your score. "Tweaking" is not

acceptable in order to make things a better fit!