A Random Rambling Rant

Mike Pearson

When Hans Rosling started to use his amazing data driven graphics to tell the world that it was not the world it thought it was, studying statistics suddenly became interesting. Even better, it became accessible. As Rosling demonstrates with his deft use of large international databases, animations, and internet video (27), technology now makes it possible to use data to tell exciting stories. In days gone by those stories would have lain undiscovered, neatly tabulated, filed and locked away.

There has never been a shortage of people telling stories, but now we can look at those stories with a degree of scepticism, find the underlying data, and make our own judgements. And if this all seems too much, we can rely on data savvy journalists (9) and bloggers (3) to pick up the trail. Still, a degree of numeracy together with common sense and an ability to spot a dodgy argument are statistical skills we all, as members of a democracy, should possess.

But what is statistics? In mathematics we often say 'probability and statistics' in the same breath, and perhaps too often we convince ourselves that they are just different aspects of the same thing. It's perhaps a little simplistic, but in a sense I believe them to be opposites. Probability starts with a model, a theory, and predicts possible future outcomes; whereas statistics starts with the outcomes, the hard, unchangeable data, and attempts to validate a model or theory.

This inversion is really important to keep in mind because the possibilities for confusion are endless. A good model should predict the data we obtained with a high probability. It's easy to confuse that probability with what we'd really like to know – the chance that the model is correct. In general,

$$P(\text{data obtained}|\text{our model is correct}) \neq P(\text{our model is correct}|\text{data obtained}), \quad (1)$$

where $P(A|B)$ means the probability of A happening given that B has already happened.

A big part of the over simplification here is that statistical thinking has to include a lot more than just analysing numbers and data. I'm going to leave the humour to Randall Munroe, the creator of the xkcd comic (see figure 1), who makes the point so much better than I could.

Data has provenance which cannot be ignored. Questions about funding, motivation, and methodology must be answered before attempting to make sense of any numbers. This is the vital non mathematical part of statistics and it's crucial if we want to arrive at anything close to the truth. And once the numbers are understood there remains the non-trivial job of communicating them to others.

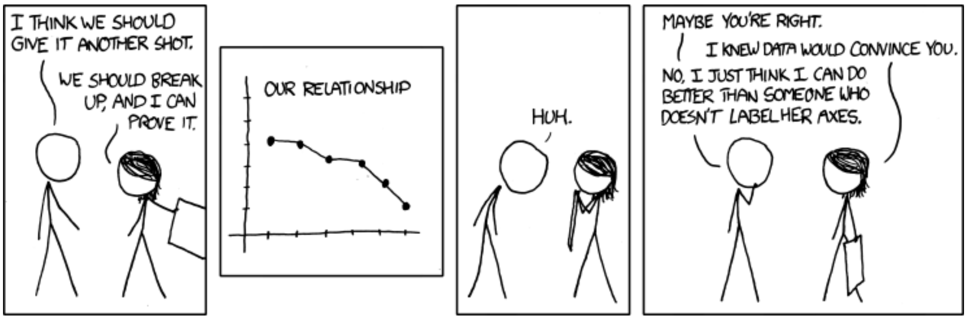
I THINK WE SHOULD GIVE IT ANOTHER SHOT.
WE SHOULD BREAK UP, AND I CAN PROVE IT.
OUR RELATIONSHIP
HUH.
MAYBE YOU'RE RIGHT.
I KNEW DATA WOULD CONVINCE YOU.
NO, I JUST THINK I CAN DO BETTER THAN SOMEONE WHO DOESN'T LABEL HER AXES.

Figure 1: *Convincing* http://xkcd.com/833/

I'm sure that much of our confused thinking related to inequality 1 comes from statistical techniques taught in school, and reinforced in media reports. In hypothesis testing we have alternate complementary models, usually called the null hypothesis H_0, and the alternate hypothesis H_1. We then proceed to calculate the probability of our data, given H_0. If this probability is below some arbitrary threshold (often 0.05), we reject the null hypothesis, saying that the alternative H_1 was found to be statistically significant at the 5% level. Notice that the technique calculates the probability of the data given H_0 rather than the probability of H_0 given the data. The logic is back to front and extremely difficult to justify at school level even though the calculations are mechanical. Of course, if we were to attempt to find what we really want to know – the probability of H_0 given the data – the calculations would probably become very difficult even though the logic would then be clear. Modern statisticians often overcome this complexity of calculation problem using computer simulations, known as Monte Carlo methods. They can then get the logic right.

One of the pitfalls of of using the back to front hypothesis testing technique is beautifully illustrated by Randall again in 'Significant' (see figure 2). Hint: count the colours.

In simple situations it is possible to get things the right way round without too much calculation. Take a good look at the carefully constructed school activity 'Bayes Ice-Breaker' by Alan Jessop (12).

'Screening for disease and dishonesty (33) provides another easy to digest example of this Bayesian reasoning which is very necessary when you only have the results of one test to go on.

1 Telling stories

There is no question that Rosling's popularity is mostly due to his ability to tell stories with data. His excitement and enthusiasm is infectious and has led to a flowering of data led journalism, most successfully in the New York Times and in the Guardian.

The Guardian Datablog (5) is a brilliant resource for teachers wanting to base lessons

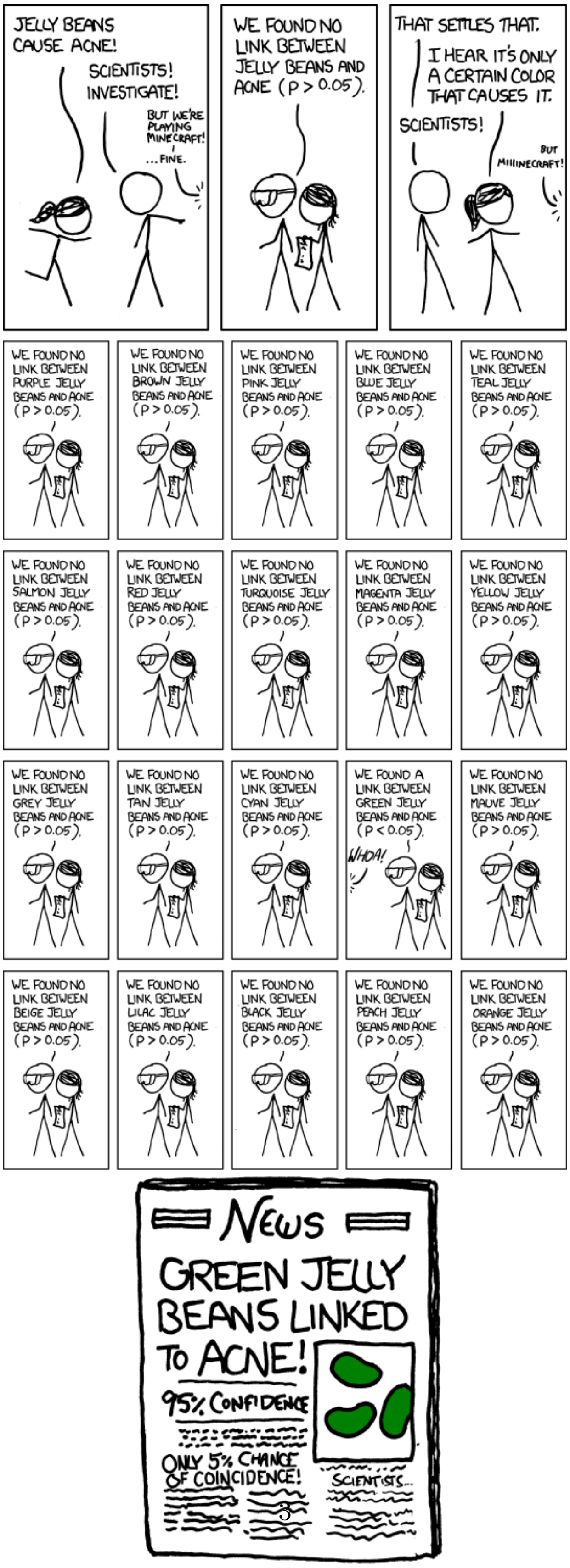
JELLY BEANS CAUSE ACNE!
SCIENTISTS! INVESTIGATE!
BUT WE'RE PLAYING MINECRAFT!
... FINE.
WE FOUND NO LINK BETWEEN JELLY BEANS AND ACNE (P > 0.05).
THAT SETTLES THAT.
I HEAR IT'S ONLY A CERTAIN COLOR THAT CAUSES IT.
SCIENTISTS!
BUT MIIIINECRAFT!
WE FOUND NO LINK BETWEEN PURPLE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BROWN JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN PINK JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BLUE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN TEAL JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN SALMON JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN RED JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN TURQUOISE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN MAGENTA JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN YELLOW JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN GREY JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN TAN JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN CYAN JELLY BEANS AND ACNE (P > 0.05).
WE FOUND A LINK BETWEEN GREEN JELLY BEANS AND ACNE (P < 0.05).
WHOA!
WE FOUND NO LINK BETWEEN MAUVE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BEIGE JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN LILAC JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN BLACK JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN PEACH JELLY BEANS AND ACNE (P > 0.05).
WE FOUND NO LINK BETWEEN ORANGE JELLY BEANS AND ACNE (P > 0.05).
NEWS
GREEN JELLY BEANS LINKED TO ACNE!
95% CONFIDENCE
ONLY 5% CHANCE OF COINCIDENCE!
SCIENTISTS...

Figure 2: *Significant* http://xkcd.com/882/

on topical data. The datablog publishes the data behind the current headlines. For example, with drought and famine raging in the Horn of Africa, you can download a spreadsheet showing the humanitarian aid to the region collected by the United Nations. PJ Harvey is in the newspapers as I write this for winning the Mercury Prize again. If that's of interest to your students, go to the data blog and download her sales figures.

It's not easy to find stories in data. Numbers in tables rarely leap out at you. One of the best ways to find the stories is to visualise the data, and it's really exciting to find one. One that I remember in particular appeared while working on a survival animation for the Understanding Uncertainty website. I was using the Human Life Database (11) maintained by the Max Plank Institute to extend the animation to all the countries it contains, going back as far as the records allowed. You can see the resulting animation here (25). Take a look at France, graph male survival, switch the scale to logarithmic, and then slowly wind the clock back to 1806. We have to thank Napoleon that we can take this dataset back so far, and we probably have to thank him too for some of its interest. The main story here is about war and its effect on survival rates of young men, but there are other gems in this dataset. Switch to women, and look at their survival rates from 1917 to 1919. Switch to the United States of America in the same period. You are probably looking at the post war flu epidemic which hit both young women and men. Now switch to Russia and look at what happened to male death rates after the fall of communism. What could explain that?

Wars have featured in statistics ever since Florence Nightingale drew her famous rose diagrams of British deaths during the Crimean campaign (32). She needed to tell a story about unnecessary deaths in order to gain support for better sanitary conditions.

Modern journalists have taken up the challenge – often with different motivations – using technology to tell more personal stories of loss alongside the overall numbers. See for example the New York Times visualisation 'Faces of the Dead' (37) for Iraq. In Britain, the war in Afghanistan has attracted similar attention. See the Guardian datablog for data and visualisations (7).

From war to sex. Always of interest to the sixteen year old and now with no shortage of data. The Guardian datablog once more obliges with a geographical breakdown of abortion statistics (6). If you are brave enough to go there – the site may offend – the American online dating agency OKCupid now has a statistics blog called OKTrends (23). This plots data from OKCupid's large database of users – so expect bias here! It's clear that Rosling's influence has even reached the online dating scene and is now providing an unexpected career path for budding infographics designers. The less adventurous amongst you will have to make do with another xkcd gem (see figure 3).

2 Statistics in School

So far I've neatly avoided talking coherently about what might happen in the classroom.

We certainly have a good selection of problems for you to use in this November 2011 issue of NRICH. I love the idea behind 'If the World were a Village' (18) – you might

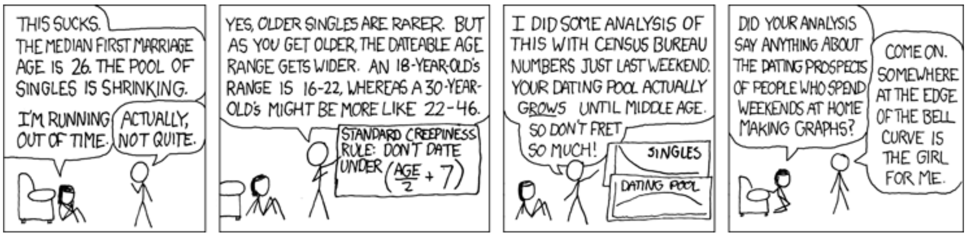
THIS SUCKS.
THE MEDIAN FIRST MARRIAGE AGE IS 26. THE POOL OF SINGLES IS SHRINKING.
I'M RUNNING OUT OF TIME.
ACTUALLY, NOT QUITE.
YES, OLDER SINGLES ARE RARER. BUT AS YOU GET OLDER, THE DATEABLE AGE RANGE GETS WIDER. AN 18-YEAR-OLD'S RANGE IS 16-22, WHEREAS A 30-YEAR-OLD'S MIGHT BE MORE LIKE 22-46.
STANDARD CREEPINESS RULE: DON'T DATE UNDER $(\frac{AGE}{2}+7)$
I DID SOME ANALYSIS OF THIS WITH CENSUS BUREAU NUMBERS JUST LAST WEEKEND. YOUR DATING POOL ACTUALLY GROWS UNTIL MIDDLE AGE.
SO DON'T FRET SO MUCH!
SINGLES
DATING POOL
DID YOUR ANALYSIS SAY ANYTHING ABOUT THE DATING PROSPECTS OF PEOPLE WHO SPEND WEEKENDS AT HOME MAKING GRAPHS?
COME ON. SOMEWHERE AT THE EDGE OF THE BELL CURVE IS THE GIRL FOR ME.

Figure 3: *Dating Pools* http://xkcd.com/314/

compare this to David McCandless’s ‘if Twitter was 100 People’ (14).

‘Charting Success’ (19) is a great stepping off point if you want to introduce your class to the world of infographics – McCandless (15) will get you started again but follow links to other practitioners as he is possibly more interested in aesthetics than communication.

The Domesday Project (NRICH (20)) invites you to examine data on pet keeping gleaned from the project that the BBC ran in 1986. There are useful skills to be picked up here - see Pereira-Mendoza (26) for a very relevant article on graphing in primary mathematics.

I have not seen a better introduction to work on averages than ‘Wisdom of the Crowds’ (4) a video by James Grime and David Spiegelhalter (YouTube’s Singing-Banana). Listen carefully to the discussion on the various averages that could be used to determine the crowd’s answer. This could be a good experiment to conduct in a school – especially if you can involve the whole school, and offer some prize for nearest guess to minimise silly responses.

The Royal Statistical Society Centre for Statistical Education (RSSCSE) publishes a Teaching Statistics journal and has even surfaced many of the best articles in a ‘Best of’ series (31). I’ve already mentioned a couple of these and they are well worth a read. They have also published Census At School and Sports at School, but I must admit that I’d classify the Census at School data tool (30) as a nice try rather than an unqualified success. I tried hard to find an interesting question to ask about the offered data set, and failed. Tools that I would have liked to use such as a 2D scatter graph and maps were unavailable. There was no possibility of time series analysis and the countries on offer were too similar. The exception – South Africa – was missing interesting data.

To provide interest we want to find our own data. Where should we start?

3 Finding stuff

We have to start at the top so let’s not forget the internet search engine. My favourite is unsurprisingly, Google. After all It’s used by xkcd too (see figure 4). Search engines will lead you to facts and figures sites like Wikipedia.

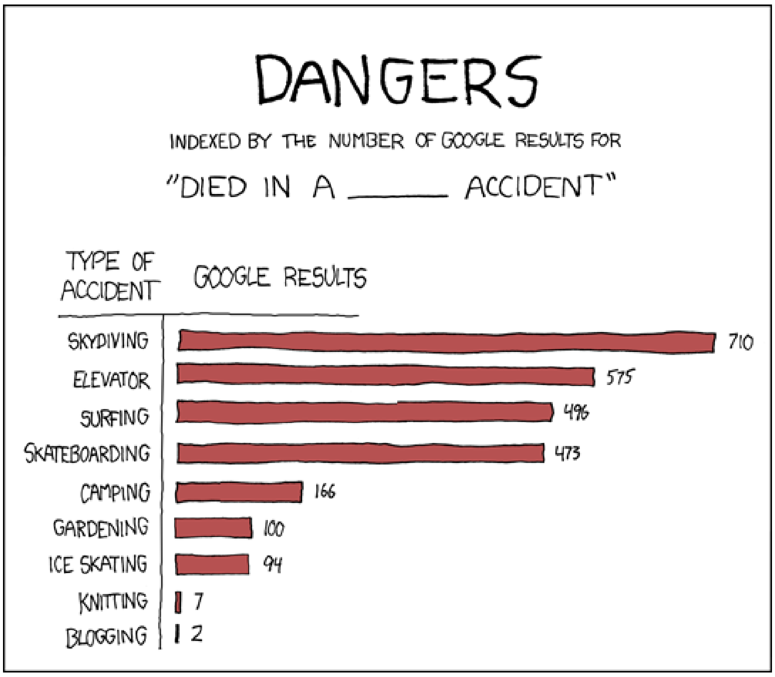

Figure 4: *Dangers* http://xkcd.com/369/

Google and Wikipedia are incredibly useful, as I'm sure you are aware. However, they don't often lead you to structured, tabulated data. Harvesting numbers from web pages can be a time consuming and painful process. It's a shame that Google shut down their Google Squared service which was cleverly designed to assist us in creating tabular data by scraping web pages.

That leaves Wolfram Alpha. Alpha is the result of a huge effort to build a *computational* search engine. It features natural language input, access to linked data sources, and the power of Mathematica running on the server, and it is very very clever. It does blurt out what it knows rather than what you asked it so you don't often get fresh answers by asking your questions in other interesting ways. It also tends to do a lot of analysis and graphing for you, which may well preempt your lesson plan. So use with care. Here (40), for example, is the result of a query on the UK population. Alpha is extremely conscientious about quoting its data sources.

4 Cautionary tales

This is probably a good moment for a word of caution. As we all know, you can't believe everything you see on the internet.

I've been working on the NRICH and Understanding Uncertainty web sites for some years. Over that time I've wasted many hours chasing problems that only show up in

Internet Explorer. So like many, I was primed to believe a recent report (16) that a study had shown that Internet Explorer users were dumber than most. The report was of course an easily discoverable hoax dressed up to look like it was backed by data. There is an excellent, if dated, text that tells you all you need to know about 'How to Lie with Statistics' (10), reviewed by Helen Joyce on our sister Plus website (13), but in this case the hoax was easily unmasked using sound statistical thinking (35).

5 Must Read Books

My one concern about mentioning texts such as 'How to lie with Statistics' (10) is that the title reinforces the oft quoted 'Lies, damned lies, and statistics' (Disraeli?, Twain) public perception. It definitely falls in the must read category however.

Edward Tufte's now classic text 'The Visual Display of Quantitative Information' (38) has a more positive approach with many examples of excellence in data visualisation discussed before he mentions less successful attempts in a chapter on graphic integrity. A useful rule of thumb promoted by Tufte is that a data visualisation should dedicate most of its ink to data. Even grids and axes are to be minimised or eliminated if they don't carry useful data. It's a minimalist approach that has its own powerful aesthetic, first developed by Otto Neurath and Gerd Arntz (1) working in pre war Vienna.

A good practical text with guidance on creating data visualisations is 'Visualize this' (41) by the author of the 'Flowing Data' blog, Nathan Yau.

6 Institutional Data

The earliest use of the term *statistics* recorded by the Oxford English Dictionary is in 1770. It helps to explain the origin of both the word - it's to do with nation states - and its early acceptance as a science.

> *The science, that is called statistics, teaches us what is the political arrangement of all the modern states of the known world.* [W. Hooper tr. Bielfeld Elem. Universal Educ. III. xiii. 269]

The UK Office for National Statistics (24) trumps the OED by tracing its history back to the Domesday book. That may be rather tongue in cheek, but the point is well made. The earliest statistics were data collected by governments about the states they administered.

We should expect to find a lot of useful data in national databases.

As I write this, the ONS is about to launch its new website so there's not much point in me explaining how to navigate the old one. One thing to watch though is that much of the data is in the form of downloadable reports published as PDF files. If you are looking for datasets do not despair - there's a good chance that these reports link through to spreadsheets, and the reports are there to supply the provenance needed to interpret the data. All very useful!

The ONS is responsible for the national census. The 2001 census feeds data to a Neighbourhood Statistics (36) web site. Here you will find interesting data for your local area. I expect this will all change once we have the 2011 census results, but at the moment the Neighbourhood summary is a good starting point, providing a number of summaries and visuals on health, education, crime etc.

For worldwide statistics, the OECD (22) is a fantastic resource. Look for the page called *Statistics from A to Z* for a useful set of links to spreadsheet downloads and to web pages displaying extracts from the real-time OECD database. For more detailed searches, try OECD StatExtracts (21) where you can make queries on the database and export the results in Excel spreadsheets or in other useful formats. If you delve deeper into the OECD iLibrary you will begin to hit subscription content.

The United Nations has a statistics division (39). You'll find there a home page leading you to many topical reports such as progress reports and charts on the Millennium Development Goals. Useful poster material here.

The UN site has a public *data search facility* http://data.un.org, but it is surprisingly patchy. 'Coffee' produces 12 documents from 4 sources, whereas 'Cocaine' produces none at all. On the plus side, this is one portal where you can access public data from a wide variety of institutions such as the World Heath Organisation, the International Monetary Fund, National Governments, and the UN itself.

The US government places a lot of information in the public domain. Even the CIA publishes its CIA World Factbook online (2). NASA especially has a wealth of interesting data together with support for schools access. Try for example 'My NASA Data' (17).

One goal of open government is that decisions should be transparent, and there has been a trend recently for governments to publish the data on which their decisions are based. The Guardian has a useful index of these sites called 'Government data sites around the world' (8). Our own government is present and correct at http://data.gov.uk/.

The Royal Statistical Society http://www.rss.org.uk has for some time been publishing its excellent public-facing Significance magazine (29) and has more recently launched its GetStats Campaign http://www.getstats.org.uk which promises more to come. Don't miss our own work in Cambridge, supporting Professor David Spiegelhalter's Understanding Uncertainty http://understandinguncertainty.org website. Lots of excellent articles, animations and videos there together with an active blog.

I'm just dipping my toe in the water here. Think of an institution, (e.g. the WHO), then type its name into Google along with the word 'statistics' and you are away. Happy hunting.

7 What have I left out?

I'm going to end by pointing you at a couple of must-see videos. The first is 'Professor Risk' (34) by David Spiegelhalter, and the last takes us back to the beginning – another Rosling production – The Joy of Stats (28).

And on a final sadder note, take a look at xkcd ‘Lanes’ http://xkcd.com/931/. Communicating this stuff matters.

References

[1] Gerd Arntz. Gerd arntz web archive. http://www.gerdarntz.org/isotype. Available from World Wide Web: http://www.gerdarntz.org/isotype.

[2] CIA. CIA - the world factbook. https://www.cia.gov/library/publications/the-world-factbook/, 2011. Available from World Wide Web: https://www.cia.gov/library/publications/the-world-factbook/.

[3] Ben Goldacre. Bad science. http://www.badscience.net/, 2011. Available from World Wide Web: http://www.badscience.net/.

[4] James Grime and David Spiegelhalter. Wisdom of the crowds by Singing-Banana | understanding uncertainty. http://understandinguncertainty.org/wisdom-crowds-singingbanana, 2011. Available from World Wide Web: http://understandinguncertainty.org/wisdom-crowds-singingbanana.

[5] The Guardian. Data journalism and data visualization from the datablog | news | guardian.co.uk. http://www.guardian.co.uk/news/datablog. Available from World Wide Web: http://www.guardian.co.uk/news/datablog.

[6] The Guardian. Abortion statistics for england and wales: see the latest breakdown | news | guardian.co.uk. http://www.guardian.co.uk/news/datablog/2011/may/24/abortion-statistics-england-wales, May 2011. Available from World Wide Web: http://www.guardian.co.uk/news/datablog/2011/may/24/abortion-statistics-england-wales.

[7] The Guardian. British dead and wounded in afghanistan, month by month | news | guardian.co.uk. http://www.guardian.co.uk/news/datablog/2009/sep/17/afghanistan-casualties-dead-wounded-british-data, 2011. Available from World Wide Web: http://www.guardian.co.uk/news/datablog/2009/sep/17/afghanistan-casualties-dead-wounded-british-data.

[8] The Guardian. Official government data sites around the world | news | guardian.co.uk. http://bit.ly/7l6a9M, 2011. Available from World Wide Web: http://bit.ly/7l6a9M.

[9] Tim Harford. BBC NEWS | programmes | more or less, 2011. Available from World Wide Web: http://news.bbc.co.uk/1/hi/programmes/more_or_less/default.stm.

[10] Darrell Huff. *How to Lie with Statistics*. Penguin, new ed edition, December 1991.

[11] Max Plank Institute. Human Life-Table database. http://www.lifetable.de/, 2007. Available from World Wide Web: http://www.lifetable.de/.

[12] Alan Jessop. Bayes Ice-Breaker - jessop - 2010 - teaching statistics - wiley online library. http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9639.2009.00374.x/pdf, 2010. Available from World Wide Web: http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9639.2009.00374.x/pdf.

[13] Helen Joyce. 'How to lie with statistics' | plus.maths.org. http://plus.maths.org/content/how-lie-statistics-0, 2009. Available from World Wide Web: http://plus.maths.org/content/how-lie-statistics-0.

[14] David McCandless. If twitter was 100 people.... http://www.informationisbeautiful.net/2009/if-twitter-was-100-people/, 2009. Available from World Wide Web: http://www.informationisbeautiful.net/2009/if-twitter-was-100-people/.

[15] David McCandless. Information is beautiful | ideas, issues, knowledge, data - visualized! http://www.informationisbeautiful.net/, 2011. Available from World Wide Web: http://www.informationisbeautiful.net/.

[16] Rik Myslewski. It's official: IE users are dumb as a bag of hammers • the register. http://www.theregister.co.uk/2011/07/29/aptiquant_iq_survey/, 2011. Available from World Wide Web: http://www.theregister.co.uk/2011/07/29/aptiquant_iq_survey/.

[17] NASA. My NASA data. http://mynasadata.larc.nasa.gov/, 2011. Available from World Wide Web: http://mynasadata.larc.nasa.gov/.

[18] NRICH. If the world were a village. http://nrich.maths.org/7725, October 2011. Available from World Wide Web: http://nrich.maths.org/7725.

[19] NRICH. NRICH: charting success. http://nrich.maths.org/7735, October 2011. Available from World Wide Web: http://nrich.maths.org/7735.

[20] NRICH. NRICH:The domesday project. http://nrich.maths.org/7554, October 2011. Available from World Wide Web: http://nrich.maths.org/7554.

[21] OECD. OECD statistics (GDP, unemployment, income, population, labour, education, trade, finance, prices,health,debt...). http://stats.oecd.org/Index.aspx, 2011. Available from World Wide Web: http://stats.oecd.org/Index.aspx.

[22] OECD. Organisation for economic co-operation and development. http://www.oecd.org/home/0,2987,en_2649_201185_1_1_1_1_1,00.html, 2011. Available from World Wide Web: http://www.oecd.org/home/0,2987,en_2649_201185_1_1_1_1_1,00.html.

[23] OKCupid. OkTrends. http://blog.okcupid.com/. Available from World Wide Web: http://blog.okcupid.com/.

[24] ONS. ONS home. http://www.ons.gov.uk/ons/index.html, 2011. Available from World Wide Web: http://www.ons.gov.uk/ons/index.html. ONS Website homepage.

[25] Mike Pearson. SurvivalWorldwide. http://bit.ly/nocrSB, 2009. Available from World Wide Web: http://bit.ly/nocrSB.

[26] Lionel Pereira Mendoza. Graphing in the primary school. *Teaching Statistics*, 17(1):2–6, March 1995. Available from World Wide Web: http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9639.1995.tb00704.x/abstract.

[27] Hans Rosling. Hans rosling shows the best stats you've ever seen | video on TED.com. http://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen.html, 2006. Available from World Wide Web: http://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen.html.

[28] Hans Rosling. The joy of the joy of stats!, 2010. Available from World Wide Web: http://www.rsscse.org.uk/newsandfeatures/rsscse-news/335-joyofstats.

[29] RSS. Significance magazine. http://www.significancemagazine.org/view/index.html, 2011. Available from World Wide Web: http://www.significancemagazine.org/view/index.html.

[30] RSSCSE. CensusAtSchool. http://www.censusatschool.org.uk/, 2011. Available from World Wide Web: http://www.censusatschool.org.uk/.

[31] RSSCSE. Getting the best from teaching statistics | teaching statistics. http://www.rsscse-edu.org.uk/tsj/?page_id=357#one, 2011. Available from World Wide Web: http://www.rsscse-edu.org.uk/tsj/?page_id=357#one.

[32] Ian Short. Nightingale's 'Coxcombs' | understanding uncertainty. http://understandinguncertainty.org/coxcombs, 2009. Available from World Wide Web: http://understandinguncertainty.org/coxcombs.

[33] David Spiegelhalter. Screening for disease and dishonesty | understanding uncertainty. http://understandinguncertainty.org/node/238, 2009. Available from World Wide Web: http://understandinguncertainty.org/node/238.

[34] David Spiegelhalter. Professor risk | understanding uncertainty. http://understandinguncertainty.org/node/604, 2010. Available from World Wide Web: http://understandinguncertainty.org/node/604.

[35] David Spiegelhalter. Spotting a hoax using statistics | understanding uncertainty. http://understandinguncertainty.org/spotting-hoax-using-statistics, 2011. Available from World Wide Web: http://understandinguncertainty.org/spotting-hoax-using-statistics.

[36] Neighbourhood Statistics. Neighbourhood statistics. http://bit.ly/pVAUmK, January 2007. Available from World Wide Web: http://bit.ly/pVAUmK. Neighbourhood Statistics - free to access government website providing range of local area information.

[37] New York Times. Faces of the dead: Service members killed in iraq and afghanistan - interactive feature - NYTimes.com. http://www.nytimes.com/interactive/us/faces-of-the-dead.html#/wood_zarian, 2010. Available from World Wide Web: http://www.nytimes.com/interactive/us/faces-of-the-dead.html#/wood_zarian.

[38] Edward R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press USA, 2nd edition edition, January 2001.

[39] UN. United nations statistics division. http://unstats.un.org, 2011. Available from World Wide Web: http://unstats.un.org.

[40] Wolfram. uk population - Wolfram|Alpha. http://bit.ly/iIvVT2, 2011. Available from World Wide Web: http://bit.ly/iIvVT2.

[41] Nathan Yau. Visualize this: The FlowingData guide to design, visualization, and statistics. http://book.flowingdata.com/, 2011. Available from World Wide Web: http://book.flowingdata.com/.