

THE ANNUAL DEATH RATE AMONG PEOPLE
WHO KNOW THAT STATISTIC IS ONE IN SIX.

TYPE I ERROR: FALSE POSITIVE

TYPE II ERROR: FALSE NEGATIVE

TYPE III ERROR: TRUE POSITIVE FOR
INCORRECT REASONS

TYPE IV ERROR: TRUE NEGATIVE FOR
INCORRECT REASONS

TYPE V ERROR: INCORRECT RESULT WHICH
LEADS YOU TO A CORRECT
CONCLUSION DUE TO
UNRELATED ERRORS

TYPE VI ERROR: CORRECT RESULT WHICH
YOU INTERPRET WRONG

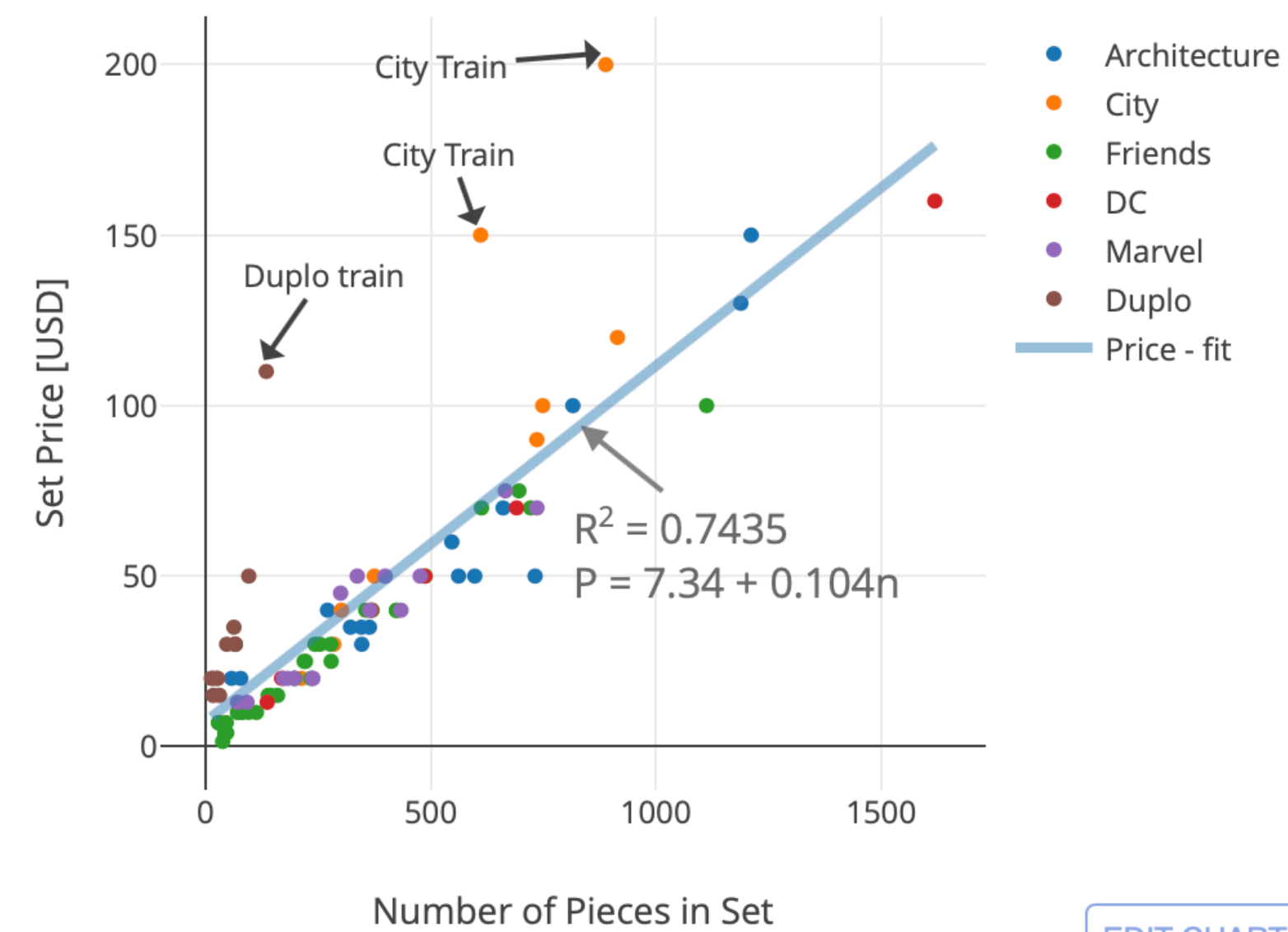
TYPE VII ERROR: INCORRECT RESULT WHICH
PRODUCES A COOL GRAPH

TYPE VIII ERROR: INCORRECT RESULT WHICH
SPARKS FURTHER RESEARCH
AND THE DEVELOPMENT OF
NEW TOOLS WHICH REVEAL
THE FLAW IN THE ORIGINAL
RESULT WHILE PRODUCING
NOVEL CORRECT RESULTS

TYPE IX ERROR: THE RISE OF SKYWALKER

■ <https://www.wired.com/2014/08/lego-cost/>

Price of Lego Sets vs. Number of Pieces in Set



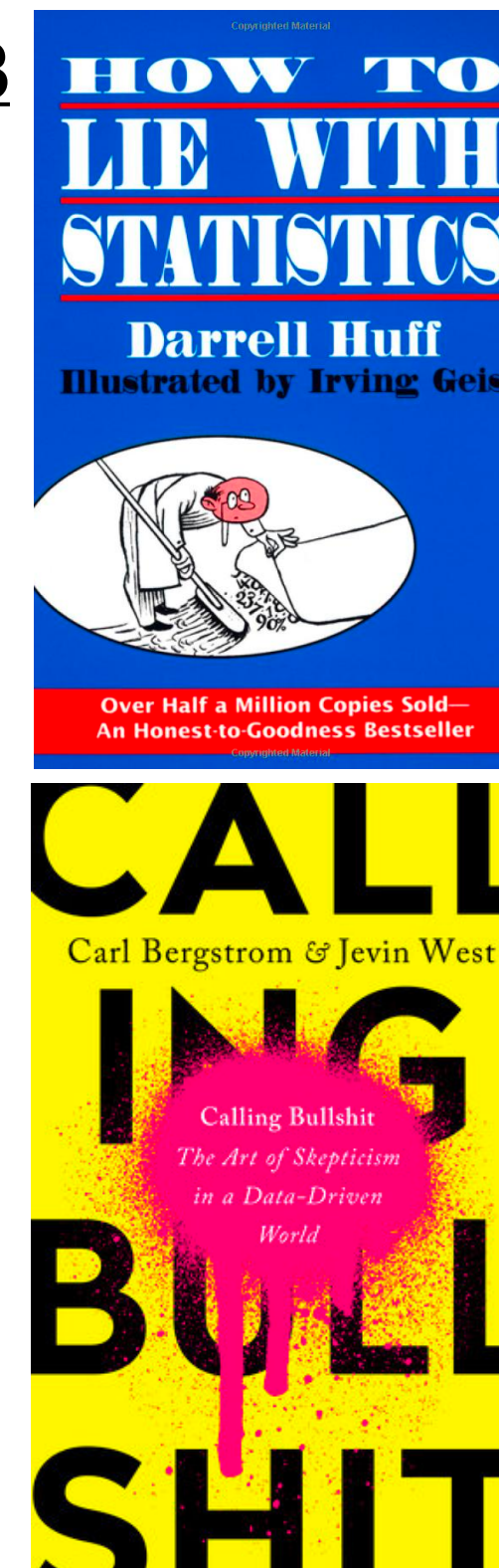
EDIT CHART

Lies, Damned Lies, and Statistics + Logical Fallacies

Frédo Durand
MIT EECS 6.00

Good resources

- <http://www.amazon.com/How-Lie-Statistics-Darrell-Huff/dp/0393310728>
- <http://www.nature.com/news/policy-twenty-tips-for-interpreting-scientific-claims-1.14183>
- <http://www.ncl.ac.uk/ceg/assets/documents/seminars/StatisticalFallacies.pdf>
- http://en.wikipedia.org/wiki/Misleading_graph
- http://en.wikipedia.org/wiki/Misuse_of_statistics
- <http://www.thenakedscientists.com/forum/index.php?topic=18520.0>
- <https://yourlogicalfallacyis.com/>
- <http://www.theskepticsguide.org/resources/logical-fallacies>
- <https://arxiv.org/pdf/1811.07271.pdf>
- <https://callingbullshit.org/syllabus.html>



So far

- advocated statistics and computational data analysis
- today: beware of misleading data
and misleading data analysis

Types of problems

- Bad data (bias, garbage)
 - Bad visualization
 - Bad logic (fallacies)
-
- And a bunch of my hobby horses

Bad data

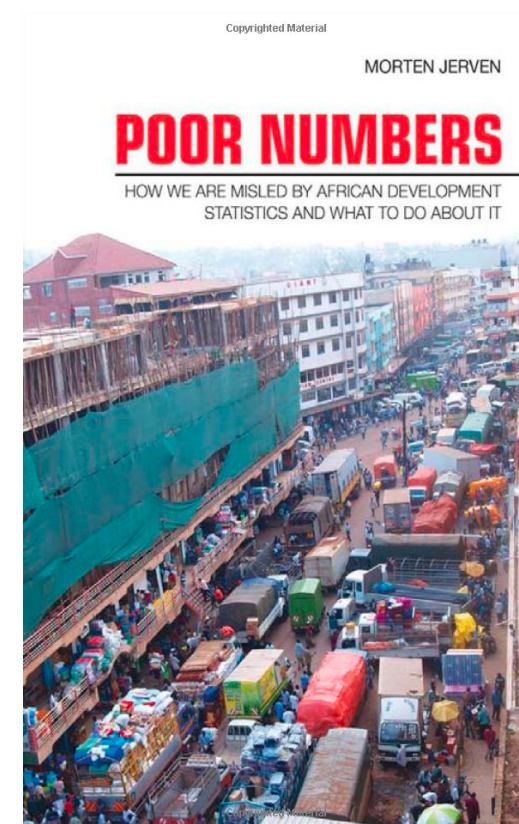
**Garbage in,
garbage out**

Garbage In Garbage Out (GIGO)

- On two occasions I have been asked,
"Pray, Mr. Babbage, if you put into the machine wrong figures, will the right answers come out?" ...
I am not able rightly to apprehend the kind of confusion of ideas that could provoke such a question.
- — Charles Babbage, Passages from the Life of a Philosopher[3]

GIGO

- World bank statistics such as GDP about African countries are deeply flawed
- http://www.econtalk.org/archives/2013/01/jerven_on_measu.html
- <http://www.cornellpress.cornell.edu/book/?gcoi=80140100939320>
- <http://afraf.oxfordjournals.org/content/early/2009/11/11/afraf.adp064.abstract?etoc>



poor

Table 2. Relative income ranking in Africa according to three data sources

	Maddison	WDI	PWT	
1	Congo, Dem. Rep.*	217	Congo, Dem. Rep.*	92
2	Sierra Leone	410	Ethiopia	115
3	Chad	429	Burundi	139
4	Niger	486	Sierra Leone	153
5	Burundi	496	Malawi	169
6	Tanzania	535	Tanzania	190
7	Guinea	572	Liberia	191
8	Central African Rep.	576	Mozambique	191
9	Comoro Islands	581	Niger	200
10	Ethiopia*	605	Guinea-Bissau	210
11	Togo	614	Chad	218
12	Zambia	645	Rwanda	242
13	Malawi	656	Burkina Faso	243
14	Guinea-Bissau	681	Madagascar	246
15	Madagascar	706	Nigeria	254
16	Angola	765	Mali	294
17	Uganda	797	Sudan	313
18	Rwanda	819	Togo	323
19	Mali	892	Kenya	328
20	Gambia	895	Central African Rep.	339
21	Burkina Faso	921	São Tomé & Príncipe	341
22	Liberia	990	Uganda	348
23	Sudan	991	Gambia, The	370
24	Mauritania	1017	Zambia	394
25	Kenya	1031	Ghana	413
26	Cameroon	1082	Benin	414
27	São Tomé & Príncipe	1226	Comoros	436
28	Nigeria	1251	Mauritania	495
29	Ghana	1270	Angola	524
30	Benin	1283	Lesotho	548
31	Zimbabwe	1328	Guinea	605
32	Côte d'Ivoire	1352	Senegal	609
33	Senegal	1358	Zimbabwe	620
34	Mozambique	1365	Cameroon	675
35	Lesotho	1490	Côte d'Ivoire	739
36	Cape Verde	1777	Congo-Brazzaville	791
37	Congo-Brazzaville	2005	Swaziland	1538
38	Swaziland	2630	Cape Verde	1541
39	Namibia	3637	Equatorial Guinea	1599
40	Gabon	3847	Namibia	2366
41	South Africa	3978	Botswana	3931
42	Botswana	4269	South Africa	4020
43	Seychelles	6354	Mauritius	4104
44	Equatorial Guinea	7973	Gabon	4378
45	Mauritius	10652	Seychelles	6557

richer

• <http://afraf.oxfordjournals.org/content/early/2009/11/11/afraf.adp064.abstract?etoc>

Tanzania

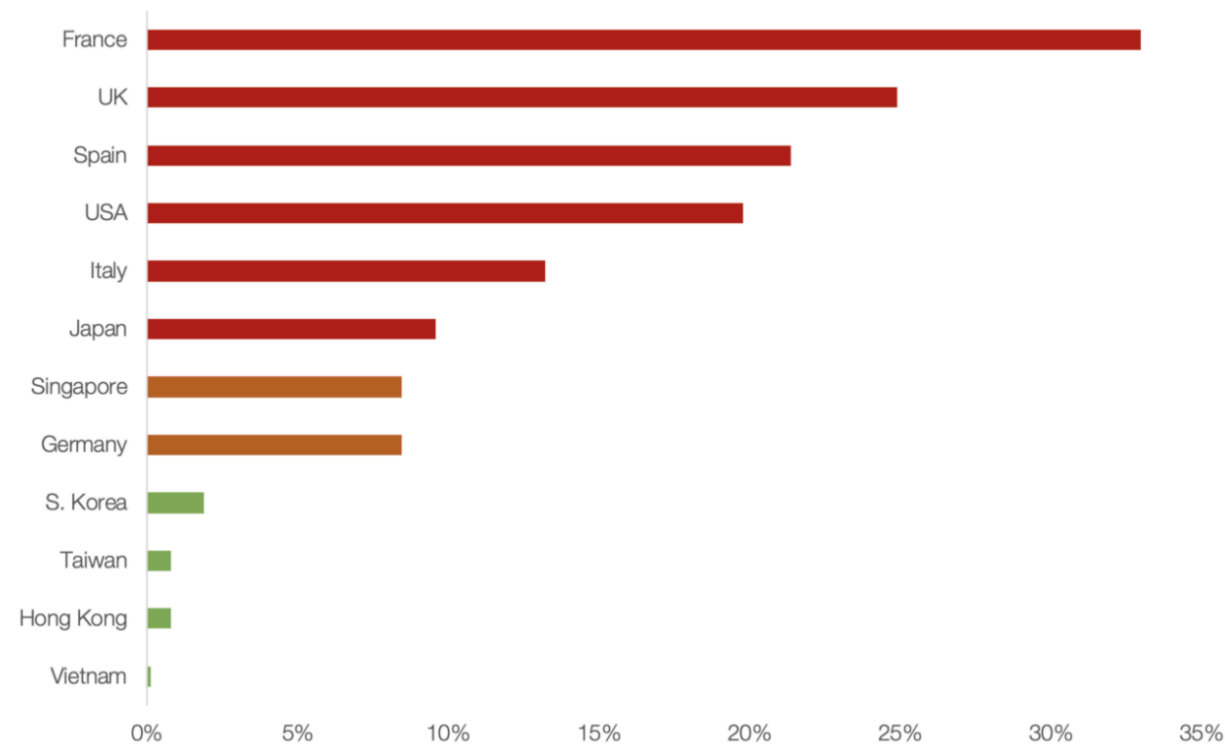
- until the 60s, does not take into account unrecorded economy
- 67, adds unrecorded construction and rent (+25%)
- 90s, adds more unrecorded stuff, +67%

• <http://afraf.oxfordjournals.org/content/early/2009/11/11/afraf.adp064.abstract?etoc>

Enormous flaws in COVID data

- Number of cases probably wrong by a factor of 2-10
- Not enough tests !
- Number of tests (and quality) varies in time and space

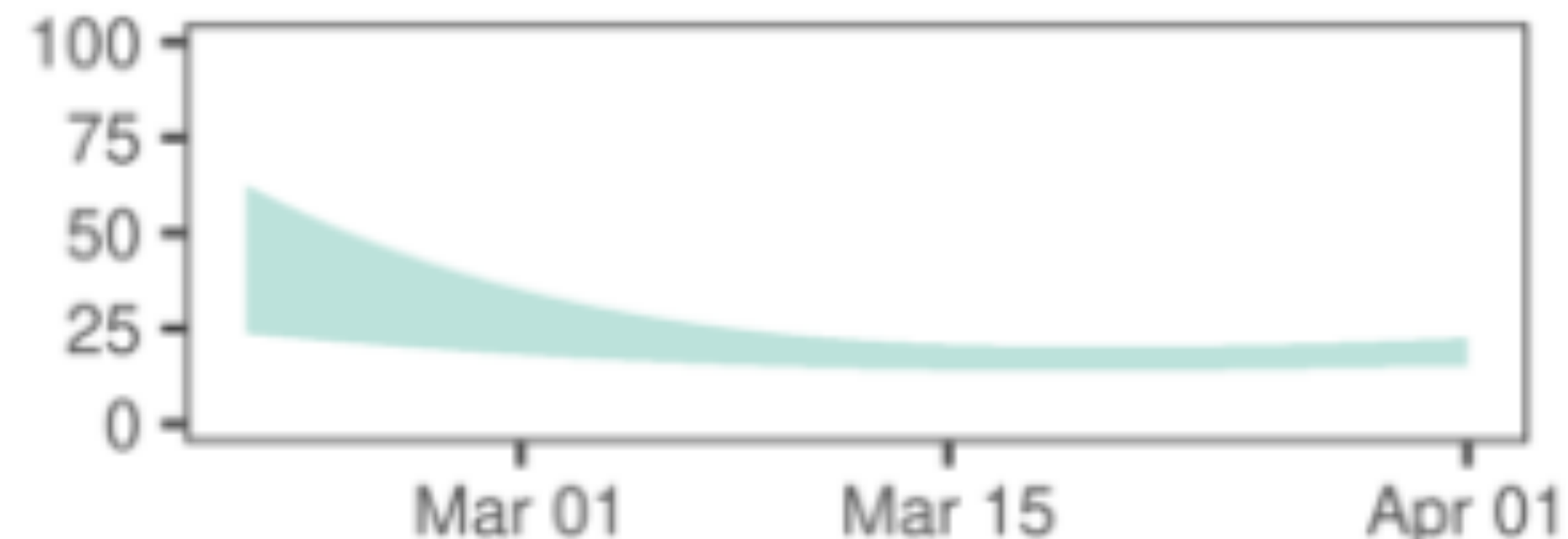
Chart 7: Share of Positive Coronavirus Tests
As of April 20th 2020



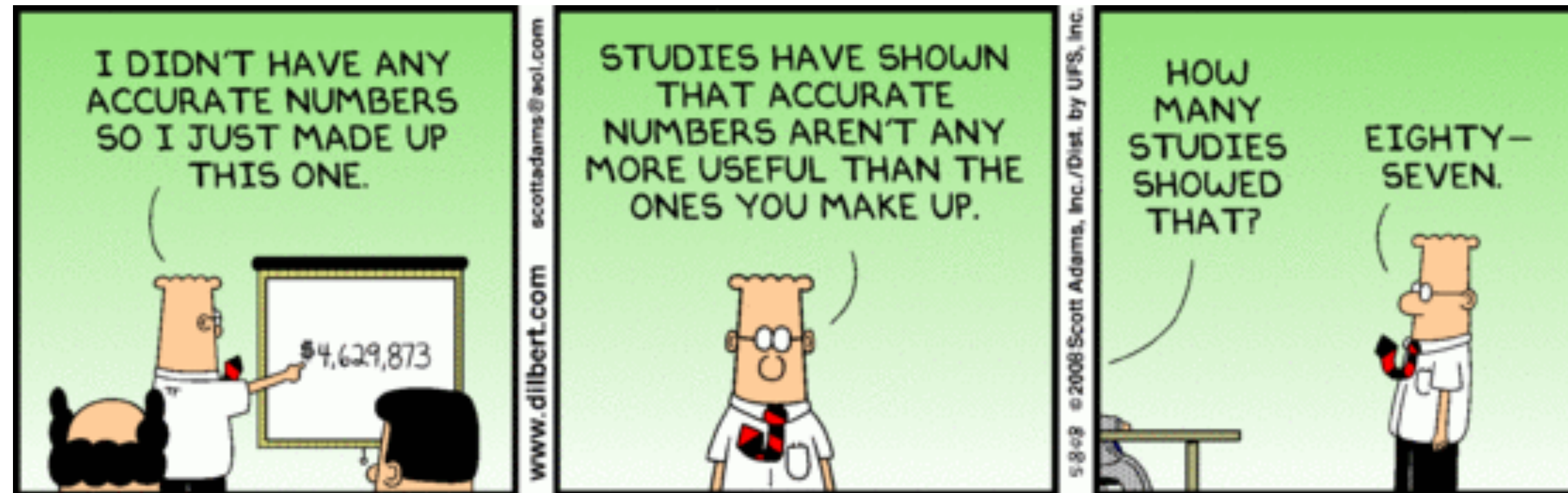
<https://medium.com/@tomaspueyo/coronavirus-learning-how-to-dance-b8420170203e>

Estimate of % of cases reported over time

United States of America



https://cmmid.github.io/topics/covid19/severity/global_cfr_estimates.html



Selection bias

Non-response bias, or the Non-representative Sample

- Statistical techniques are based upon the assumption that a by sampling a subset of a population we can infer things about the population as a whole
- Unfortunately, most studies involve convenience (or accidental) sampling
- e.g. in WWII, analysis of hits in returning planes

Poll



Non-response bias, or the Non-representative Sample

- Statistical techniques are based upon the assumption that a by sampling a subset of a population we can infer things about the population as a whole
- Unfortunately, most studies involve convenience (or accidental) sampling
- e.g. in WWII, analysis of hits in returning planes
- But couldn't look at planes that were shot down!



Hedge funds beat the average

Poll

2 1 7

Non-response bias

- Surveys that only call land lines
 - or in 1936, Literary digest predicted that Roosevelt would lose because they only surveyed people rich enough to have a phone.
- Course evaluations done at the end of the term
- Grading courses on a curve when many students have dropped

US employment data

- Includes only households
- Excludes:
 - inmates
 - homeless people
 - military
- See e.g. http://www.econtalk.org/archives/2012/12/pettit_on_the_p.html

At MIT

- Program XYZ to help students
 - assesses effectiveness by how these students then perform
 - but students must apply to these programs.
 - Selection bias!

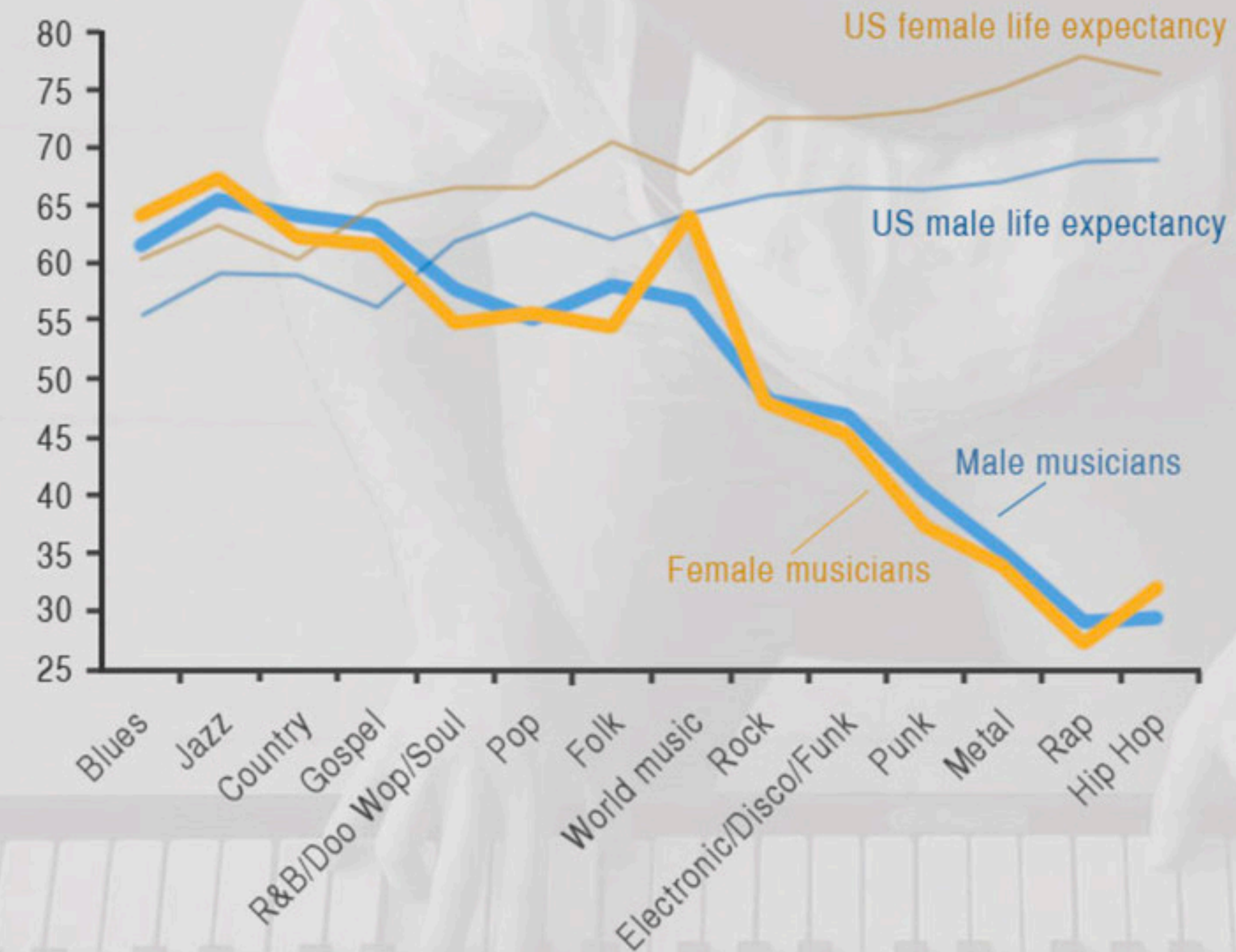
- https://callingbullshit.org/case_studies/case_study_musician_mortality.html

-

Age of death and musical genre

Average age of death for popular musicians by genre and sex

Average age at death



Questions?

Bad presentation and visualization

Context, order of magnitude

Data Enhancement by lack of context.

- Most auto accidents happen within 10 miles of home.

Poll

Data Enhancement by lack of context.

- **Most auto accidents happen within 10 miles of home.**
- **Also most driving!**

- **Context matters!**
- **Numbers should be provided with a baseline**

Budget numbers

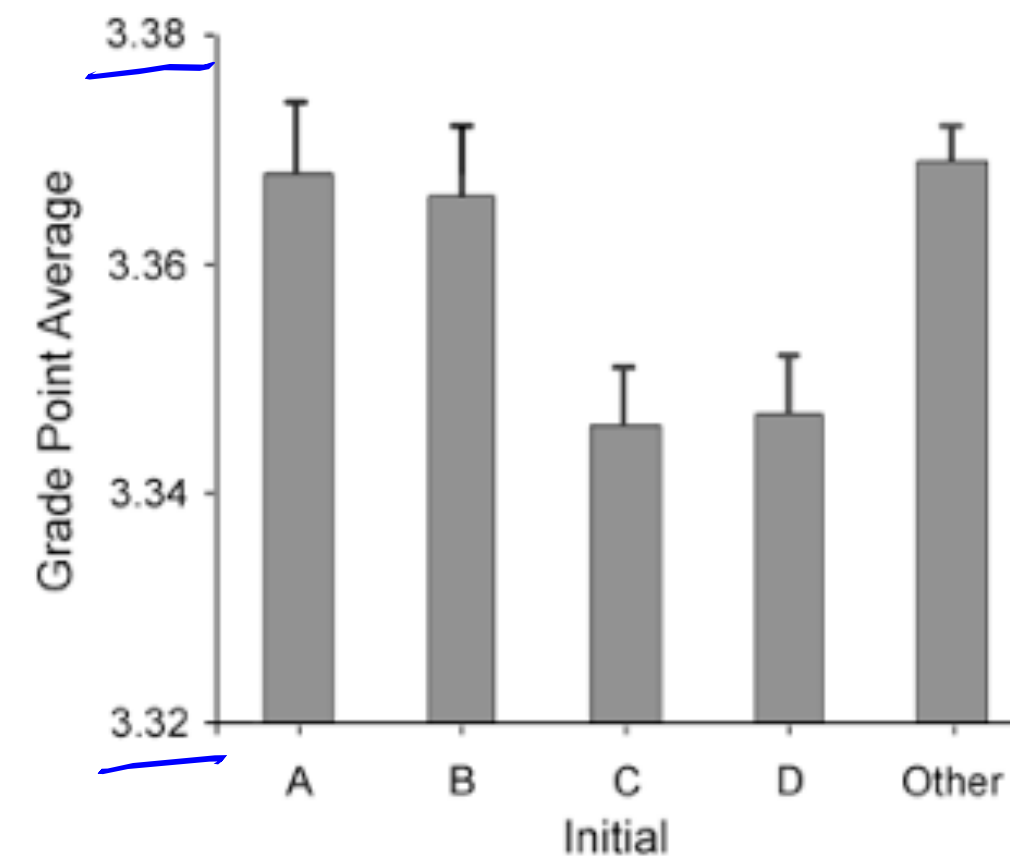
- Should be in context
- <http://www.cepr.net/index.php/blogs/beat-the-press/numbers-in-context-big-congrats-to-the-new-york-times-and-margaret-sullivan>
- <http://www.cepr.net/index.php/blogs/beat-the-press/when-is-the-nyt-going-to-start-putting-budget-numbers-in-context>
- <http://www.cepr.net/index.php/responsible-budget-reporting>

Dollar Amount of Program/Cut/Etc*		Over	Starting		Percent of Budget
\$ 0	thousand ▾	1 ▾ years	2013 ▾	=	0%

Plots

What is your last name initial

- Nelson and Simmons, “Monkier Maladies,” Psychological Science, Vol. 18, No. 12, 2007.
- (Mgmt professors at UCSD and Yale)

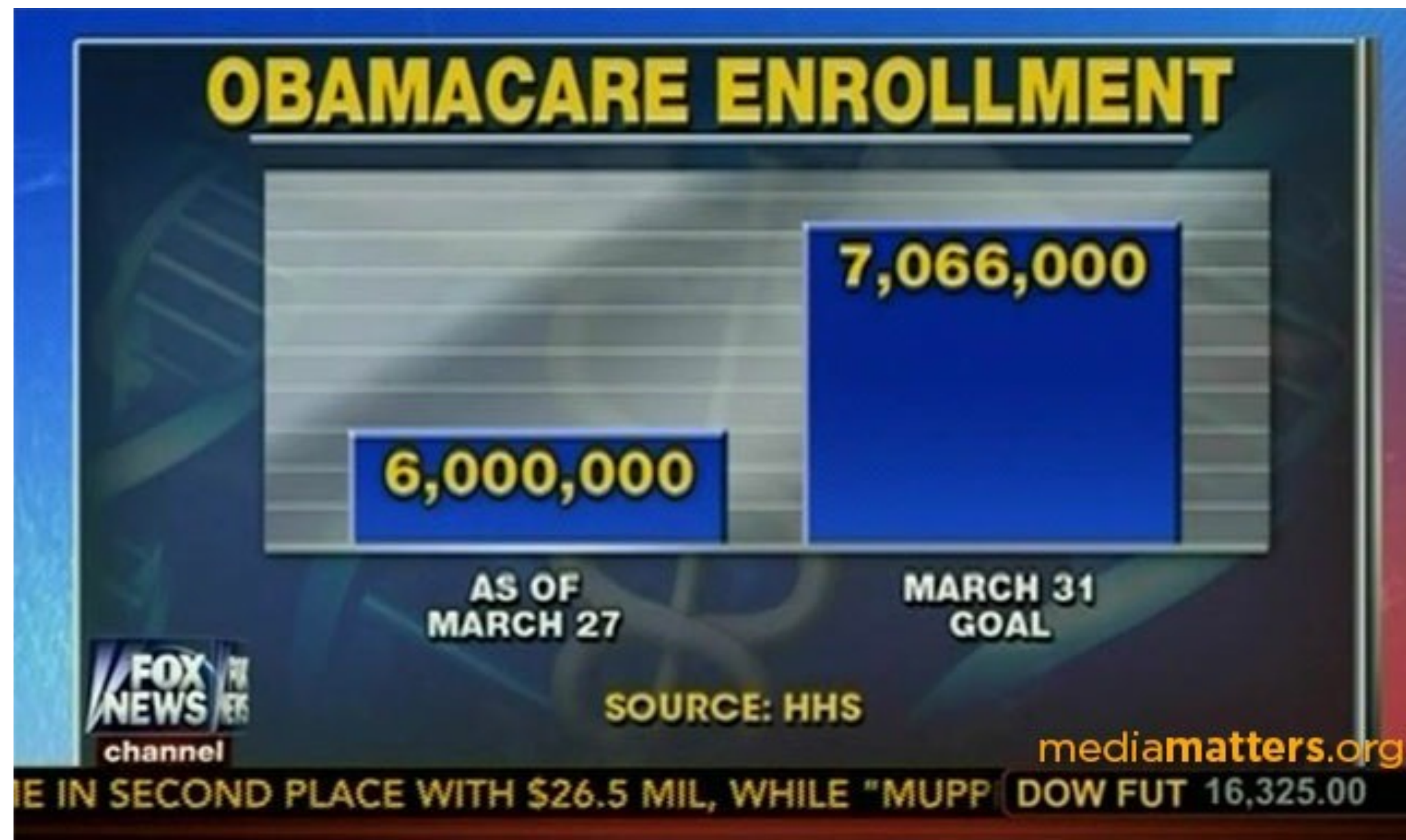


- Notice the suspicious y scale
- other reasons to be wary....

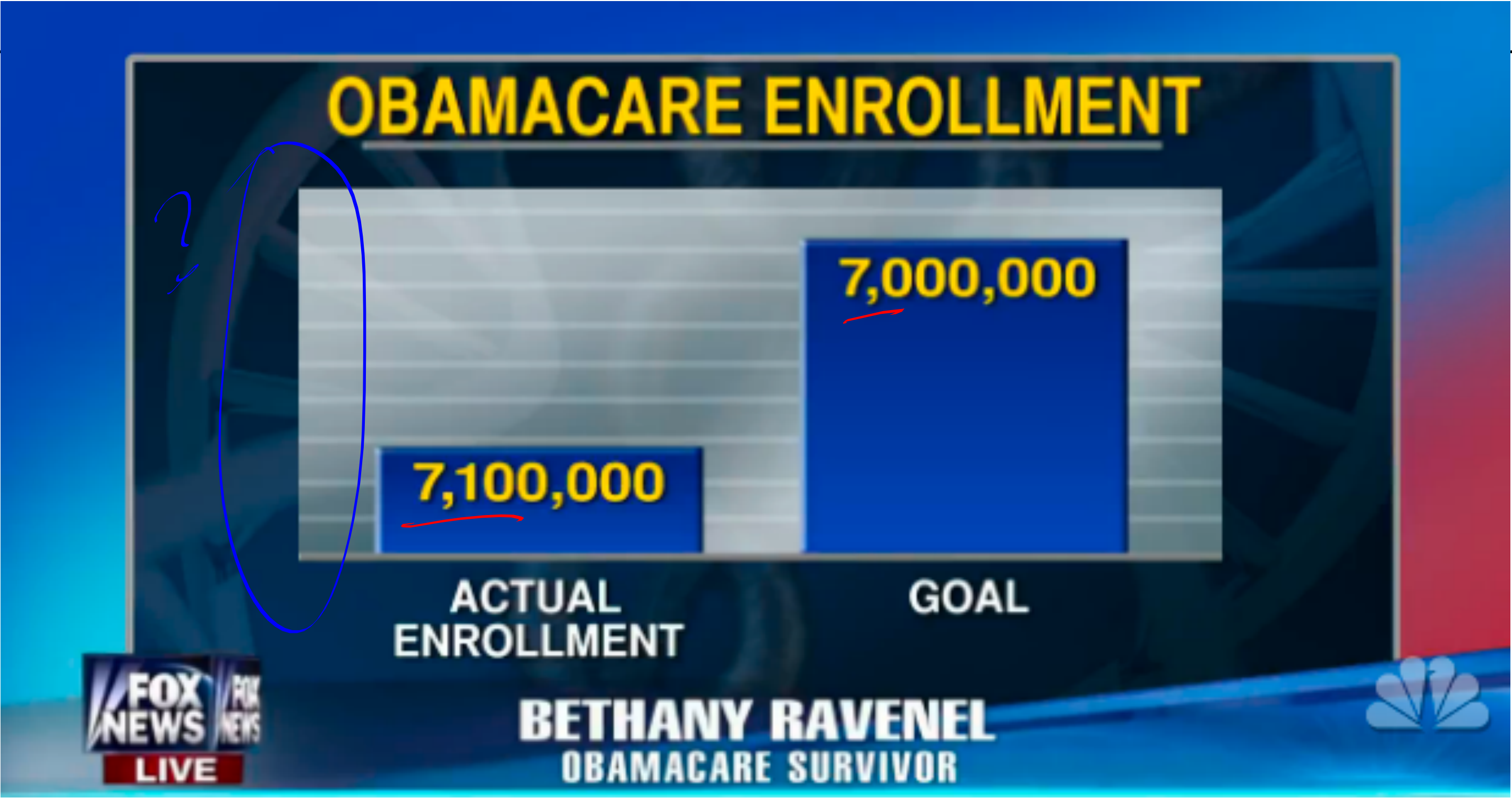
http://www.stat.columbia.edu/~gelman/stuff_for_blog/Nelson&Simmons,2007.pdf

Fox news





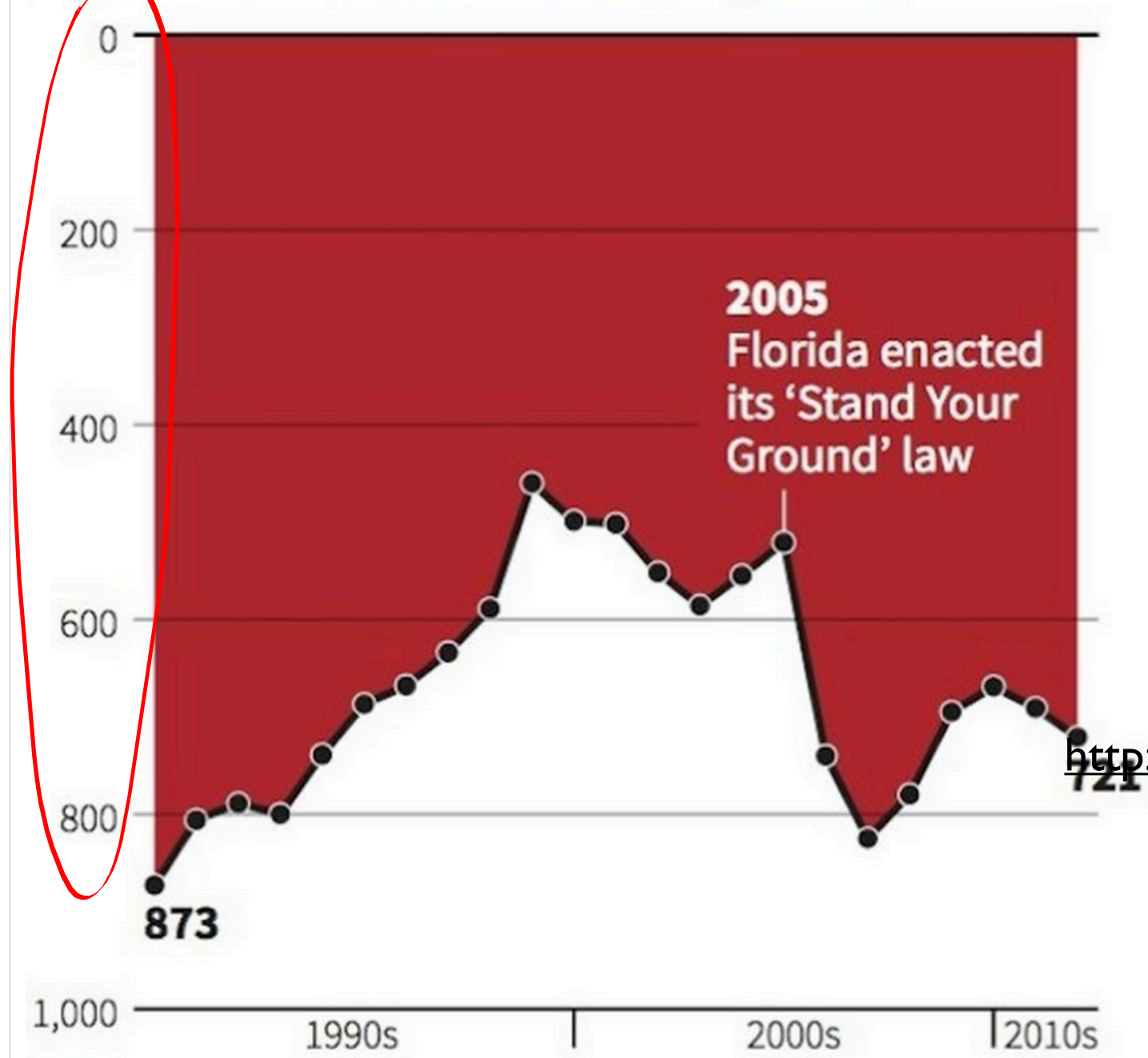
<http://mediamatters.org/blog/2014/03/31/dishonest-fox-charts-obamacare-enrollment-edit/198679>



SNL version

Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

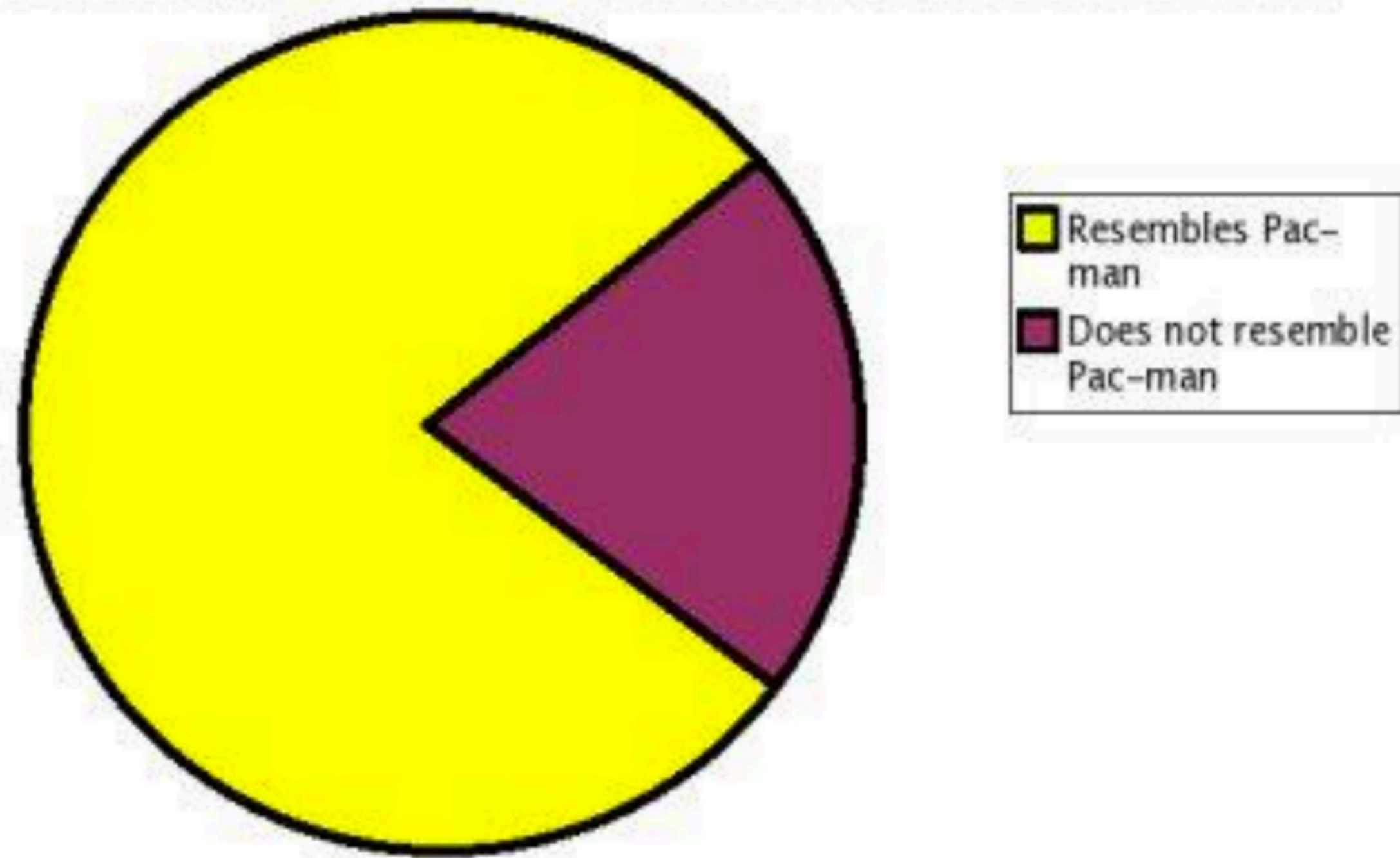
• <http://gizmodo.com/how-to-lie-with-data-visualization-1563576606>

<http://usvsth3m.com/post/82779802419/creator-defends-graph-that-appears-to-erron>

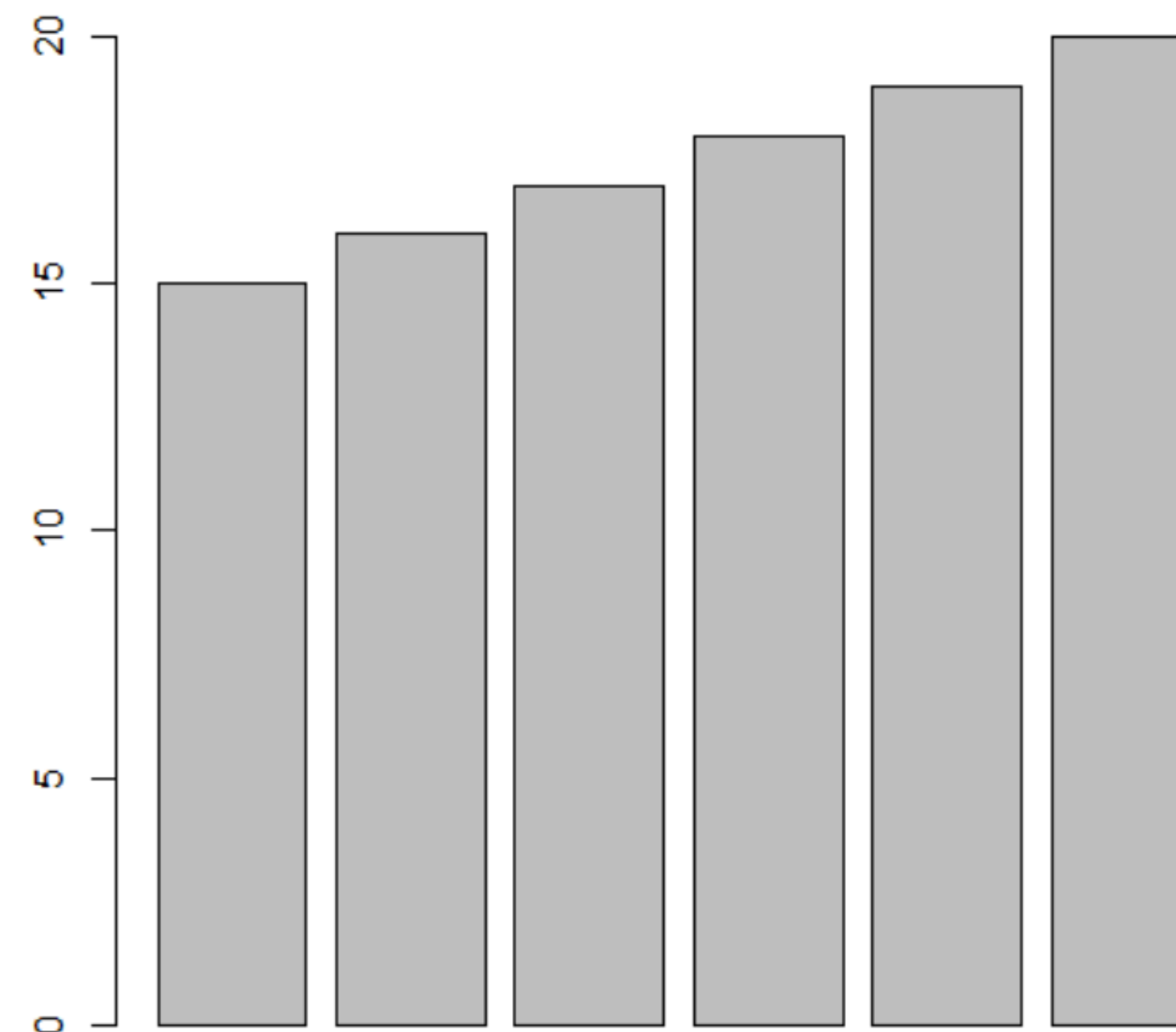
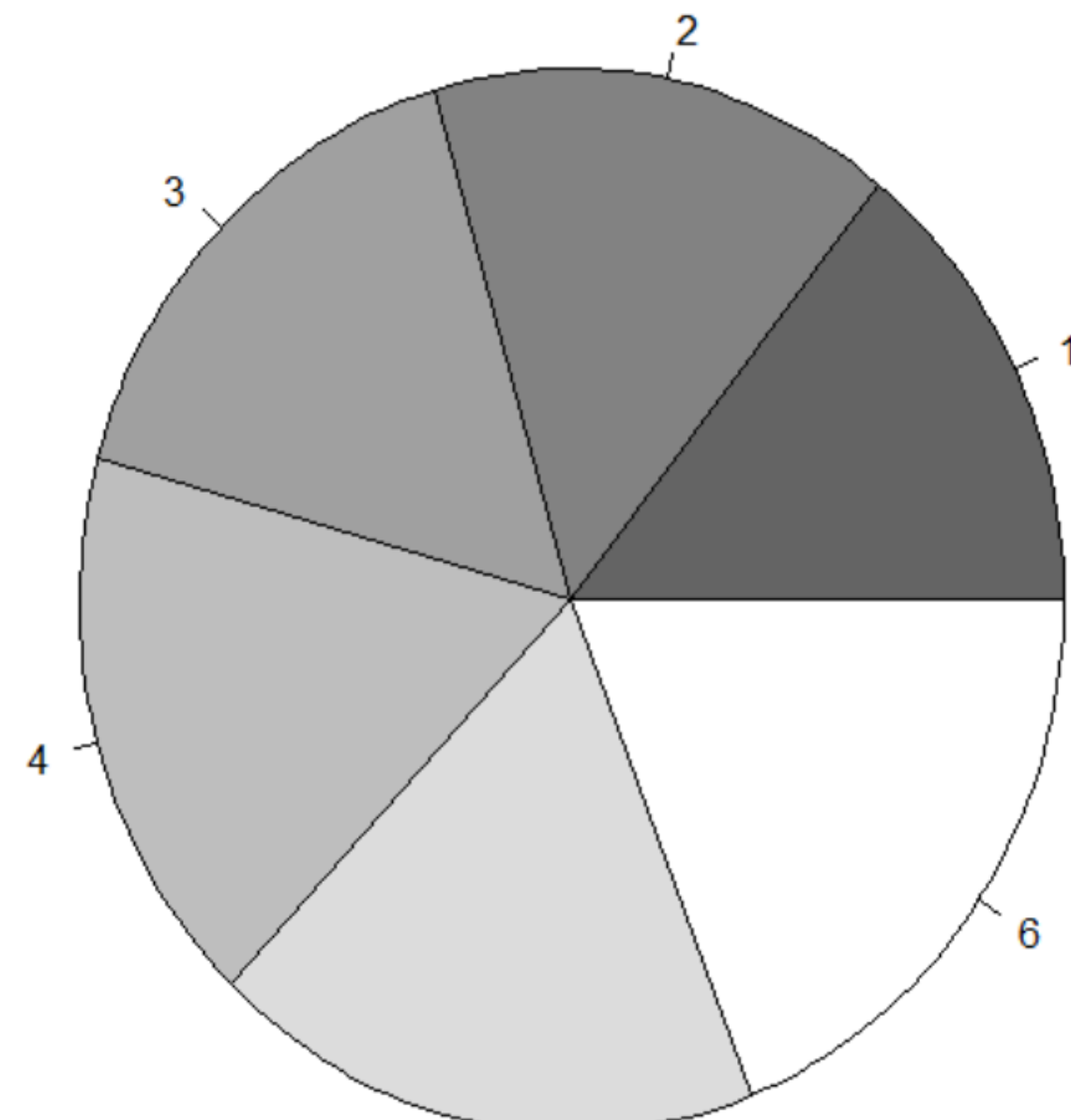
<https://www.visualisingdata.com/2014/04/the-fine-line-between-cor>

Pie charts ARE THE WORST

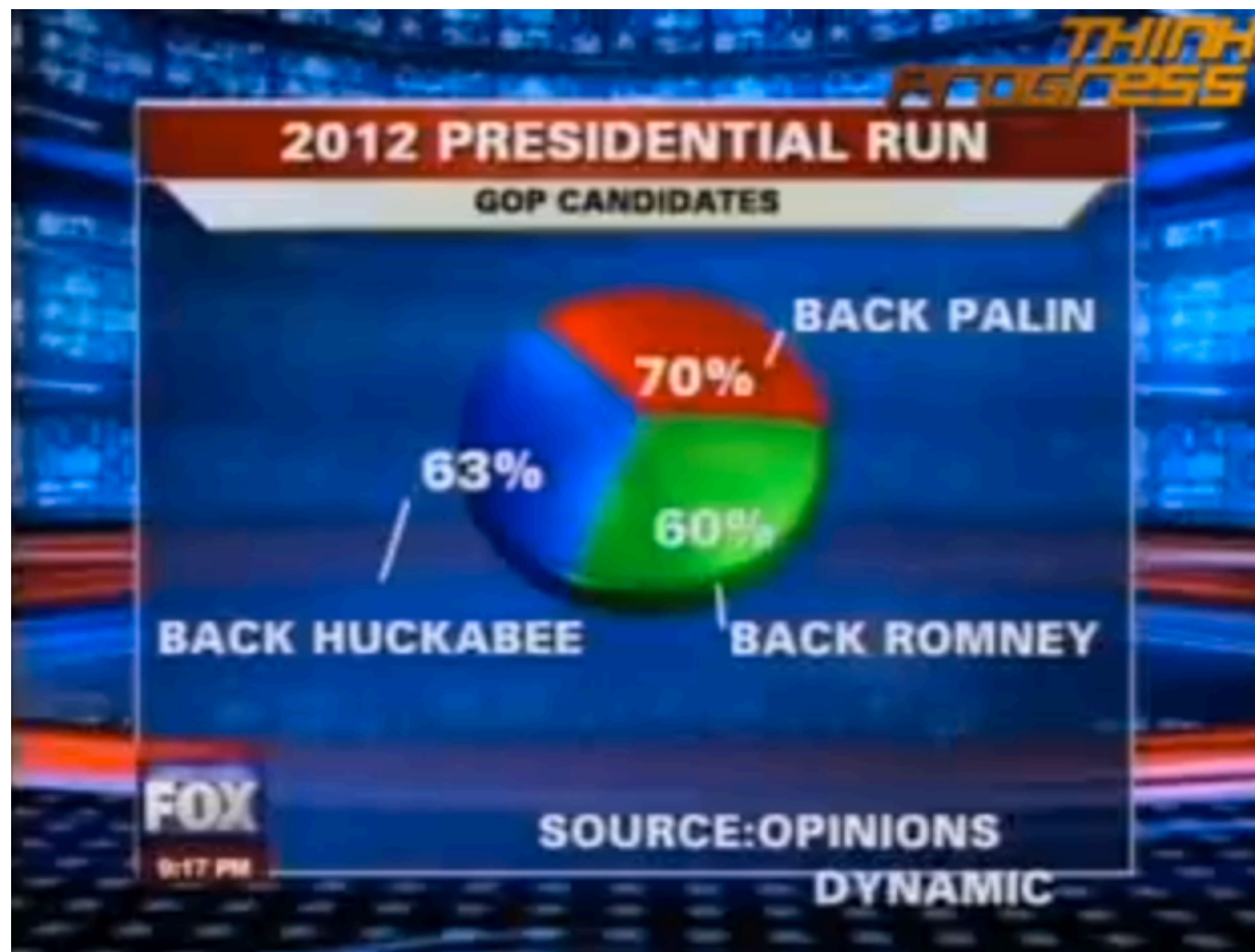
Percentage of Chart Which Resembles Pac-man



Pie charts minimize differences

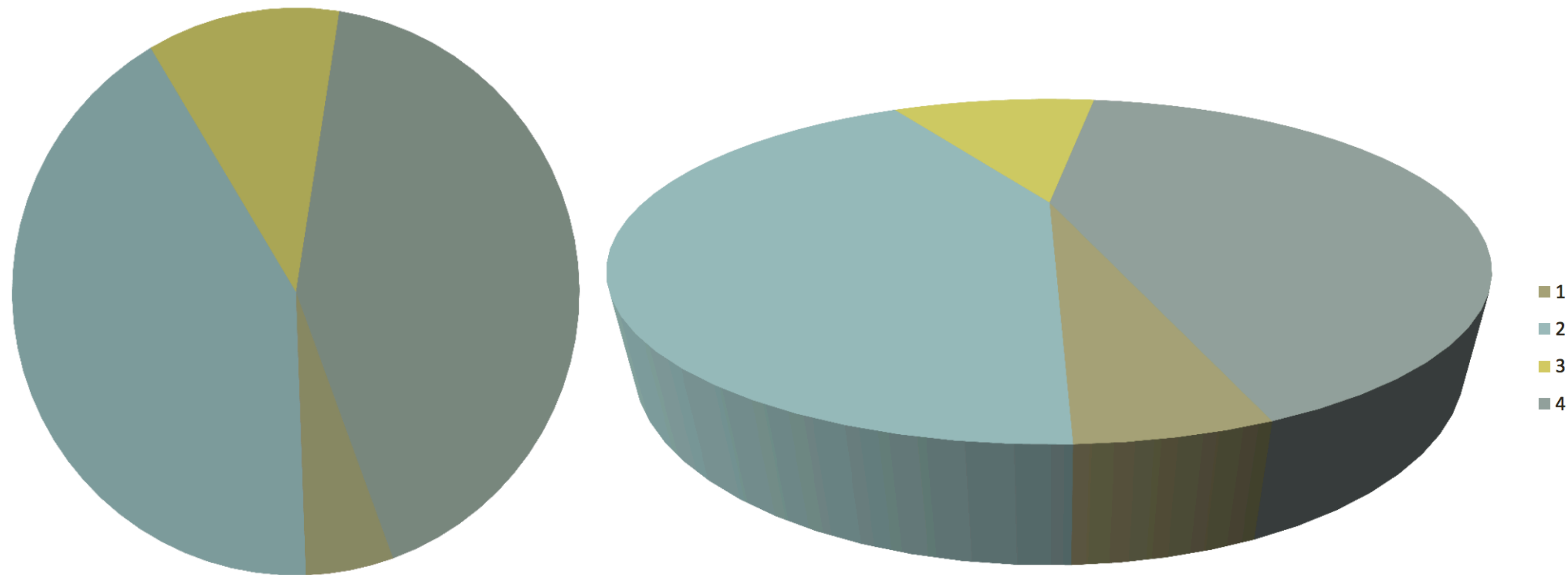


Making pie charts worse



<http://data.heapanalytics.com/how-to-lie-with-data-visualization/>

3D is even worse



Questions?

Bad logic (fallacies)



strawman

Misrepresenting someone's argument to make it easier to attack.

By exaggerating, misrepresenting, or just completely fabricating someone's argument, it's much easier to present your own position as being reasonable, but this kind of dishonesty serves to undermine rational debate.

After Will said that we should put more money into health and education, Warren responded by saying that he was surprised that Will hates our country so much that he wants to leave it defenceless by cutting military spending.



slippery slope

Asserting that if we allow A to happen, then Z will consequently happen too, therefore A should not happen.

The problem with this reasoning is that it avoids engaging with the issue at hand, and instead shifts attention to baseless extreme hypotheticals. The merits of the original argument are then tainted by unsubstantiated conjecture.

Colin Closet asserts that if we allow same-sex couples to marry, then the next thing we know we'll be allowing people to marry their parents, their cars and even monkeys.



special pleading

Moving the goalposts or making up exceptions when a claim is shown to be false.

Humans are funny creatures and have a foolish aversion to being wrong. Rather than appreciate the benefits of being able to change one's mind through better understanding, many will invent ways to cling to old beliefs.

Edward Johns claimed to be psychic, but when his 'abilities' were tested under proper scientific conditions, they magically disappeared. Edward explained this saying that one had to have faith in his abilities for them to work.



the gambler's fallacy

Believing that 'runs' occur to statistically independent phenomena such as roulette wheel spins.

This commonly believed fallacy can be said to have helped create a city in the desert of Nevada USA. Though the overall odds of a big run happening may be low, each spin of the wheel is itself entirely independent from the last.

Red had come up six times in a row on the roulette wheel, so Greg knew that it was close to certain that black would be next up. Suffering an economic form of natural selection with this thinking, he soon lost all of his savings.



black-or-white

Where two alternative states are presented as the only possibilities, when in fact more possibilities exist.

Also known as the false dilemma, this insidious tactic has the appearance of forming a logical argument, but under closer scrutiny it becomes evident that there are more possibilities than the either/or choice that is presented.

Whilst rallying support for his plan to fundamentally undermine citizens' rights, the Supreme Leader told the people they were either on his side, or on the side of the enemy.



false cause

Presuming that a real or perceived relationship between things means that one is the cause of the other.

Many people confuse correlation (things happening together or in sequence) for causation (that one thing actually causes the other to happen). Sometimes correlation is coincidental, or it may be attributable to a common cause.

Pointing to a fancy chart, Roger shows how temperatures have been rising over the past few centuries, whilst at the same time the numbers of pirates have been decreasing; thus pirates cool the world and global warming is a hoax.



ad hominem

Attacking your opponent's character or personal traits in an attempt to undermine their argument.

Ad hominem attacks can take the form of overtly attacking somebody, or casting doubt on their character. The result of an ad hom attack can be to undermine someone without actually engaging with the substance of their argument.

After Sally presents an eloquent and compelling case for a more equitable taxation system, Sam asks the audience whether we should believe anything from a woman who isn't married, was once arrested, and smells a bit weird.



loaded question

Asking a question that has an assumption built into it so that it can't be answered without appearing guilty.

Loaded question fallacies are particularly effective at derailing rational debates because of their inflammatory nature - the recipient of the loaded question is compelled to defend themselves and may appear flustered or on the back foot.

Grace and Helen were both romantically interested in Brad. One day, with Brad sitting within earshot, Grace asked in an inquisitive tone whether Helen was having any problems with a fungal infection.



appeal to emotion

Manipulating an emotional response in place of a valid or compelling argument.

Appeals to emotion include appeals to fear, envy, hatred, pity, guilt, and more. Though a valid, and reasoned, argument may sometimes have an emotional aspect, one must be careful that emotion doesn't obscure or replace reason.

Luke didn't want to eat his sheep's brains with chopped liver and brussels sprouts, but his father told him to think about the poor, starving children in a third world country who weren't fortunate enough to have any food at all.



tu quoque

Avoiding having to engage with criticism by turning it back on the accuser - answering criticism with criticism.

Literally translating as 'you too' this fallacy is commonly employed as an effective red herring because it takes the heat off the accused having to defend themselves and shifts the focus back onto the accuser themselves.

Nicole identified that Hannah had committed a logical fallacy, but instead of addressing the substance of her claim, Hannah accused Nicole of committing a fallacy earlier on in the conversation.



burden of proof

Saying that the burden of proof lies not with the person making the claim, but with someone else to disprove.

The burden of proof lies with someone who is making a claim, and is not upon anyone else to disprove. The inability, or disinclination, to disprove a claim does not make it valid (however we must always go by the best available evidence).

Bertrand declares that a teapot is, at this very moment, in orbit around the Sun between the Earth and Mars, and that because no one can prove him wrong his claim is therefore a valid one.



the fallacy fallacy

Presuming a claim to be necessarily wrong because a fallacy has been committed.

It is entirely possible to make a claim that is false yet argue with logical coherency for that claim, just as it is possible to make a claim that is true and justify it with various fallacies and poor arguments.

Recognising that Amanda had committed a fallacy in arguing that we should eat healthy food because a nutritionist said it was popular, Alyse said we should therefore eat bacon double cheeseburgers every day.



personal incredulity

Saying that because one finds something difficult to understand, it's therefore not true.

Subjects such as biological evolution via the process of natural selection require a good amount of understanding before one is able to properly grasp them; this fallacy is usually used in place of that understanding.

Kirk drew a picture of a fish and a human and with effusive disdain asked Richard if he really thought we were stupid enough to believe that a fish somehow turned into a human through just, like, random things happening over time.



ambiguity

Using double meanings or ambiguities of language to mislead or misrepresent the truth.

Politicians are often guilty of using ambiguity to mislead and will later point to how they were technically not outright lying if they come under scrutiny. It's a particularly tricky and premeditated fallacy to commit.

When the judge asked the defendant why he hadn't paid his parking fines, he said that he shouldn't have to pay them because the sign said 'Fine for parking here' and so he naturally presumed that it would be fine to park there.



bandwagon

Appealing to popularity or the fact that many people do something as an attempted form of validation.

The flaw in this argument is that the popularity of an idea has absolutely no bearing on its validity. If it did, then the Earth would have made itself flat for most of history to accommodate this popular belief.

Shamus pointed a drunken finger at Sean and asked him to explain how so many people could believe in leprechauns if they're only a silly old superstition. Sean, however, had had a few too many Guinness himself and fell off his chair.



appeal to authority

Saying that because an authority thinks something, it must therefore be true.

It's important to note that this fallacy should not be used to dismiss the claims of experts or scientific consensus. Appeals to authority are not valid arguments, but nor is it reasonable to disregard the claims of experts who have a demonstrated depth of knowledge unless one has a similar level of understanding.

Not able to defend his position that evolution 'isn't true' Bob says that he knows a scientist who also questions evolution (and presumably isn't herself a primate).



composition /division

Assuming that what's true about one part of something has to be applied to all, or other, parts of it.

Often when something is true for the part it does also apply to the whole, but because this isn't always the case it can't be presumed to be true. We must show evidence for why a consistency will exist.

Daniel was a precocious child and had a liking for logic. He reasoned that atoms are invisible, and that he was made of atoms and therefore invisible too. Unfortunately, despite his thinking skills, he lost the game of hide and go seek.



no true scotsman

Making what could be called an appeal to purity as a way to dismiss relevant criticisms or flaws of an argument.

This fallacy is often employed as a measure of last resort when a point has been lost. Seeing that a criticism is valid, yet not wanting to admit it, new criteria are invented to disassociate oneself or one's argument.

Angus declares that Scotsmen do not put sugar on their porridge, to which Lachlan points out that he is a Scotsman and puts sugar on his porridge. Furious, like a true Scot, Angus yells that no true Scotsman sugars his porridge.



genetic

Judging something good or bad on the basis of where it comes from, or from whom it comes.

To appeal to prejudices surrounding something's origin is another red herring fallacy. This fallacy has the same function as an ad hominem, but applies instead to perceptions surrounding something's source or context.

Accused on the 6 o'clock news of corruption and taking bribes, the senator said that we should all be very wary of the things we hear in the media, because we all know how very unreliable the media can be.



begging the question

A circular argument in which the conclusion is included in the premise.

This logically incoherent argument often arises in situations where people have an assumption that is very ingrained, and therefore taken in their minds as a given. Circular reasoning is bad mostly because it's not very good.

The word of Zorbo the Great is flawless and perfect. We know this because it says so in The Great and Infallible Book of Zorbo's Best and Most Trust Things that are Definitely True and Should Not Ever Be Questioned.



appeal to nature

Making the argument that because something is 'natural' it is therefore valid, justified, inevitable, good, or ideal.

Many 'natural' things are also considered 'good', and this can bias our thinking; but naturalness itself doesn't make something good or bad. For instance, murder could be seen as very natural, but that doesn't mean it's justifiable.

The medicine man rolled into town on his bandwagon offering various natural remedies, such as very special plain water. He said that it was only natural that people should be wary of 'artificial' medicines like antibiotics.



anecdotal

Using personal experience or an isolated example instead of a valid argument, especially to dismiss statistics.

It's often much easier for people to believe someone's testimony as opposed to understanding variation across a continuum. Scientific and statistical measures are almost always more accurate than individual perceptions and experiences.

Jason said that that was all cool and everything, but his grandfather smoked, like, 30 cigarettes a day and lived until 97 - so don't believe everything you read about meta analyses of sound studies showing proven causal relationships.



the texas sharpshooter

Cherry-picking data clusters to suit an argument, or finding a pattern to fit a presumption.

This 'false cause' fallacy is coined after a marksman shooting at barns and then painting a bullseye target around the spot where the most bullet holes appear. Clusters naturally appear by chance, and don't necessarily indicate causation.

The makers of Sugarette Candy Drinks point to research showing that of the five countries where Sugarette drinks sell the most units, three of them are in the top ten healthiest countries on Earth, therefore Sugarette drinks are healthy.



middle ground

Saying that a compromise, or middle point, between two extremes must be the truth.

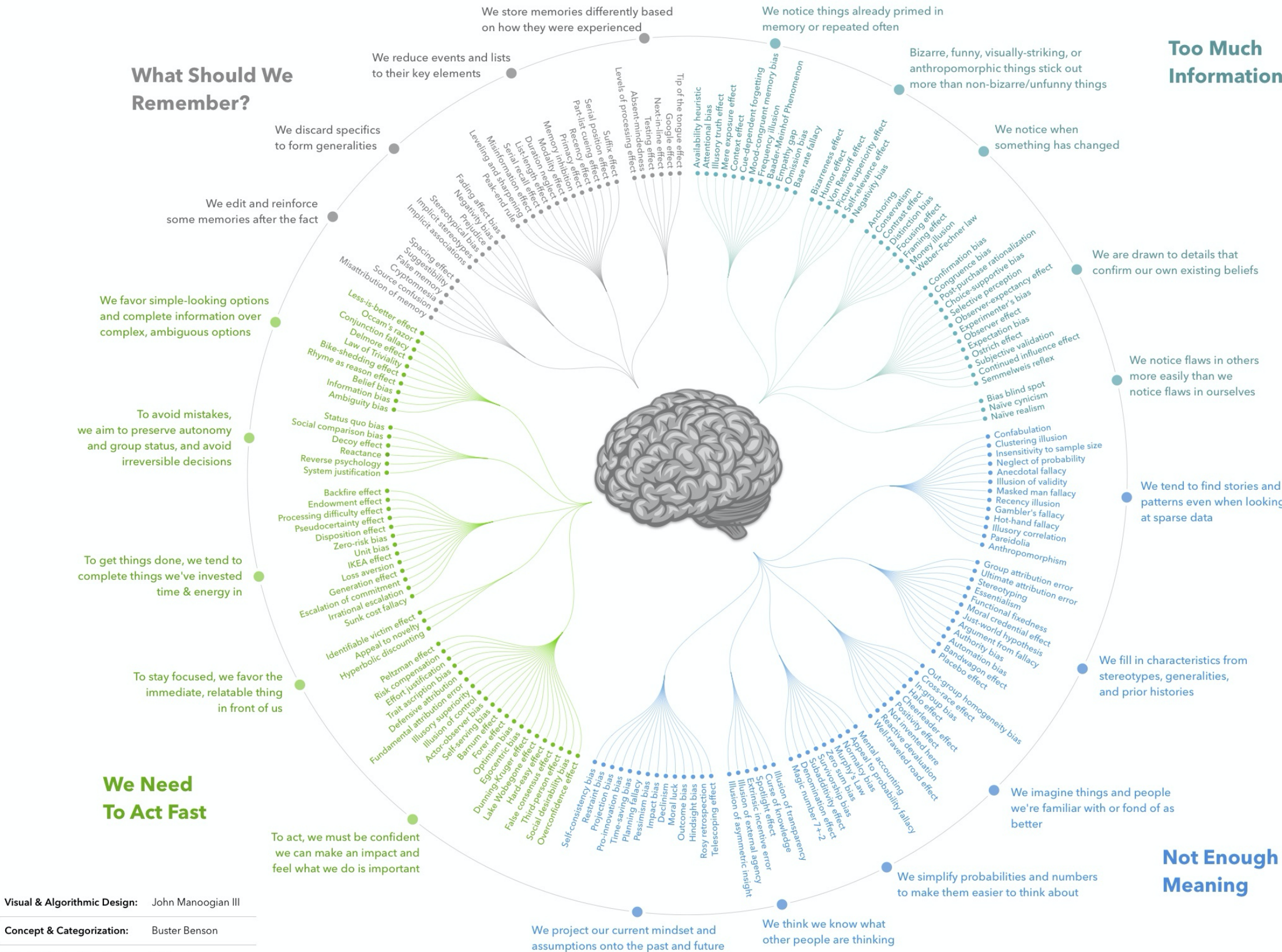
Much of the time the truth does indeed lie between two extreme points, but this can bias our thinking; sometimes a thing is simply untrue and a compromise of it is also untrue. Half way between truth and a lie, is still a lie.

Holly said that vaccinations caused autism in children, but her scientifically well-read friend Caleb said that this claim had been debunked and proven false. Their friend Alice offered a compromise that vaccinations cause some autism.

• [https://cogn](https://cogn.designhacks.co)

/09/

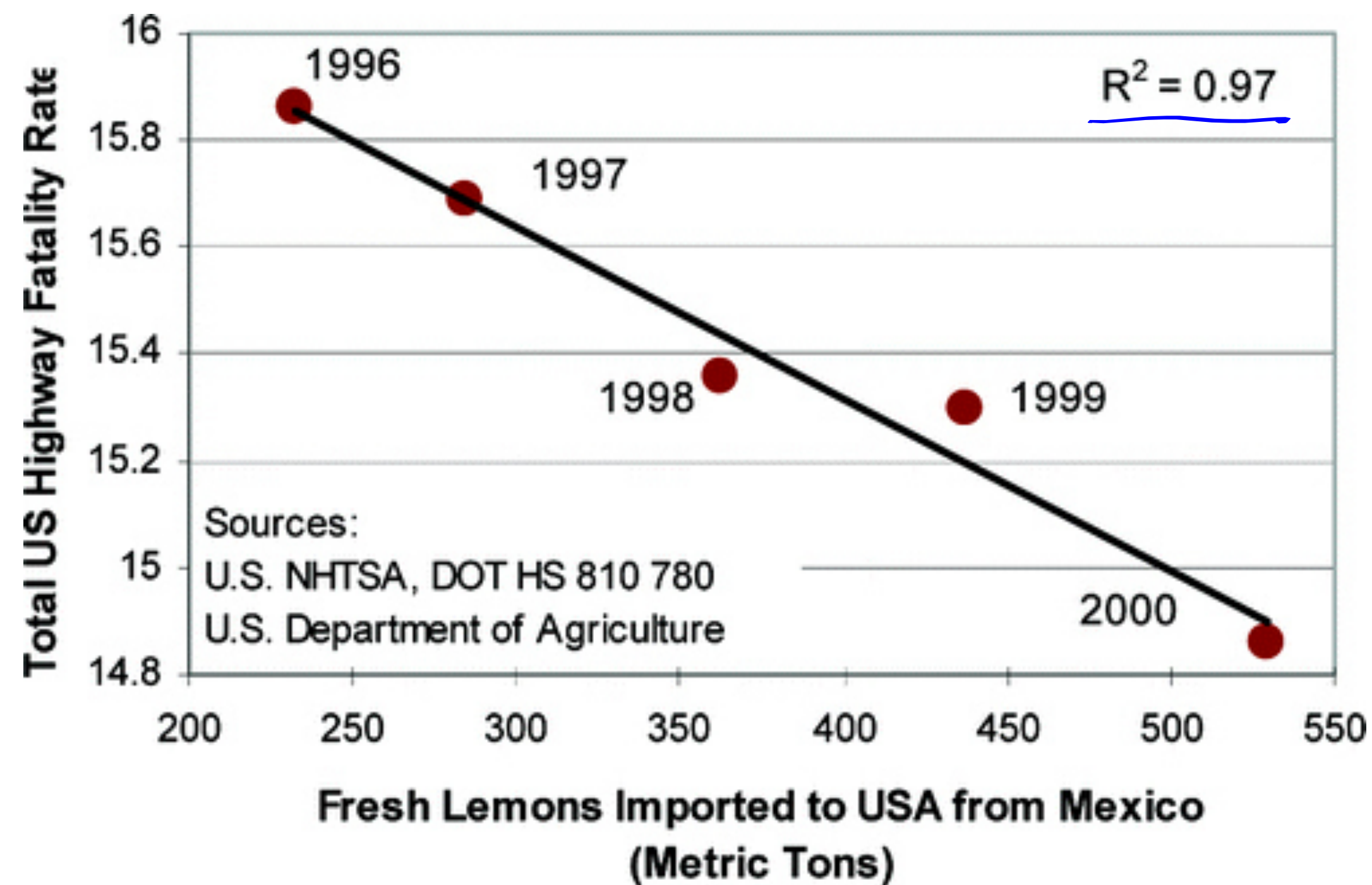
COGNITIVE BIAS CODEX



designhacks.co

Correlation & Causation

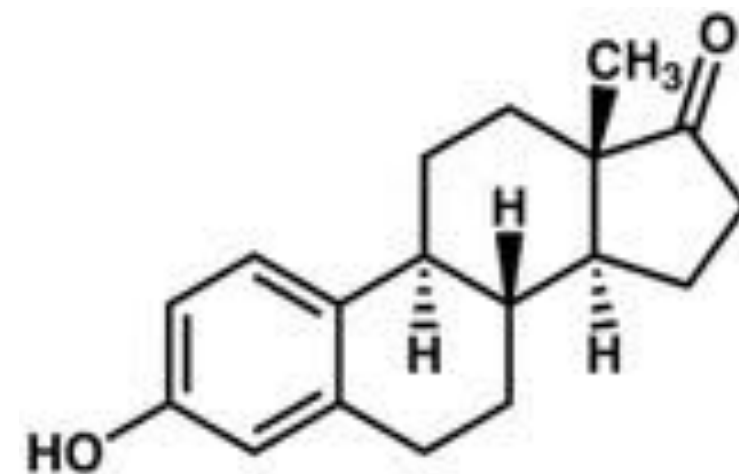
Mexican lemons cause highway deaths



- <http://www.buzzfeed.com/kjh2110/the-10-most-bizarre-correlations>

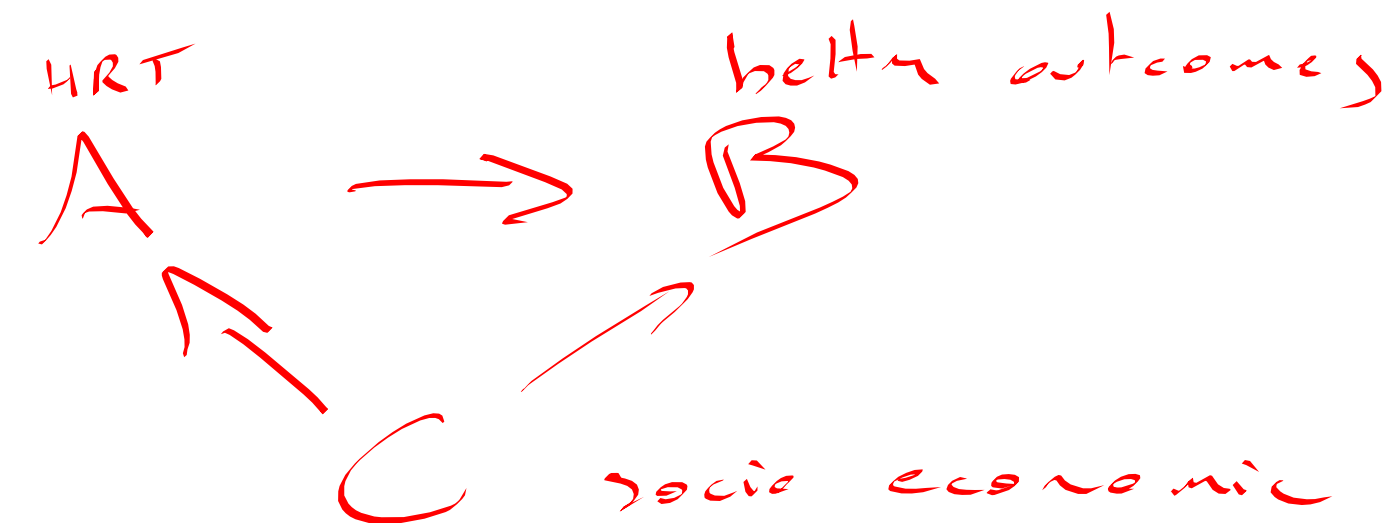
Estrogen

- Hormone replacement therapy was supposed to substantially lower risks of cardiovascular disease
- Supported by several studies that demonstrated reduced incidence of cardiovascular death among women using it.

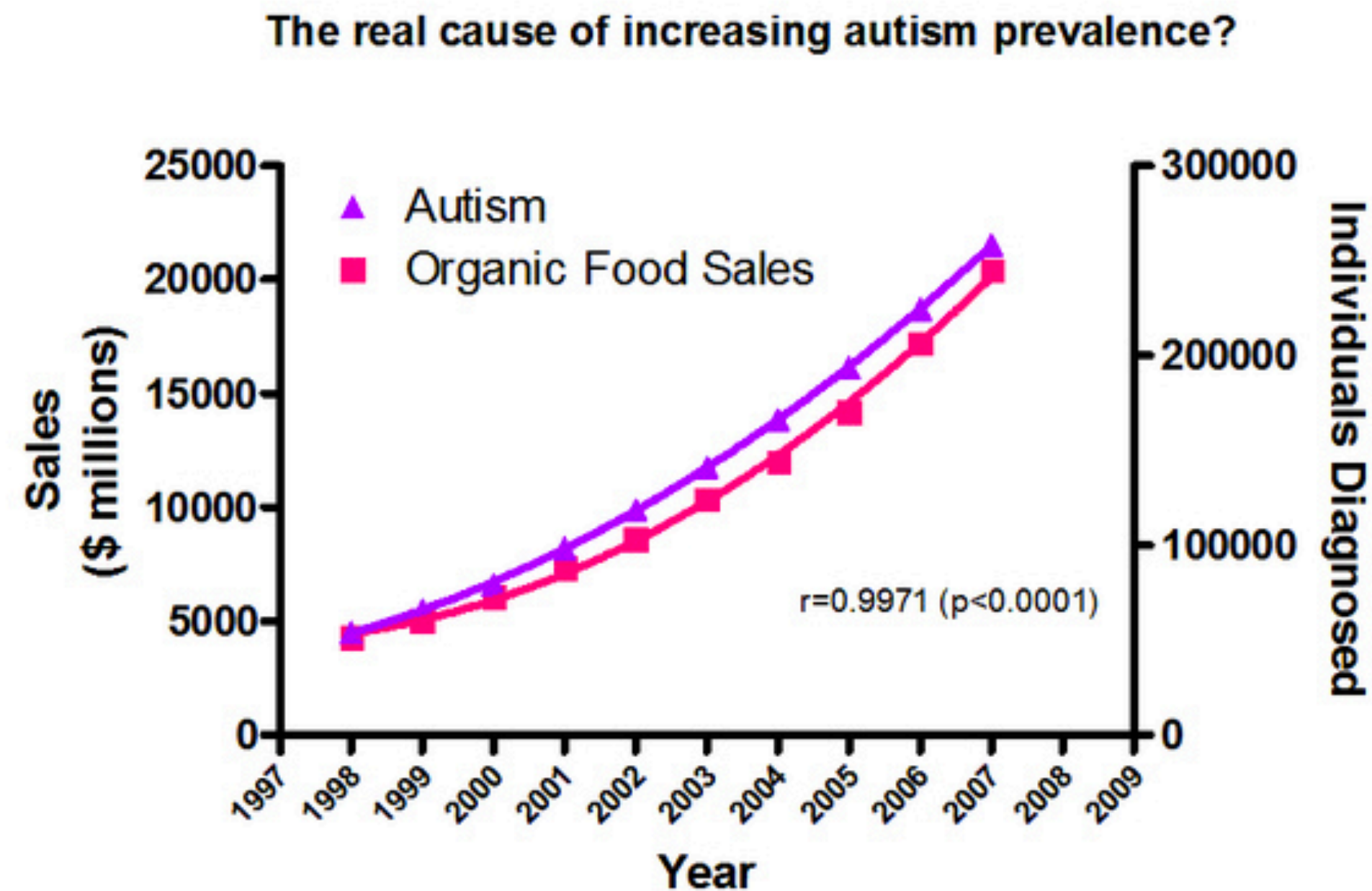


Hormone replacement therapy

- Later that year when the Journal of the American Medical Society published an article asserting that HRT in fact increased the risk of cardiovascular disease.
- Women undertaking HRT were more likely to be from groups with better than average diet and exercise regimes,
- i.e., there was a lurking variable. It seems that taking HRT and improved cardiac health were coincident effects of a common cause.



Autism & organic food

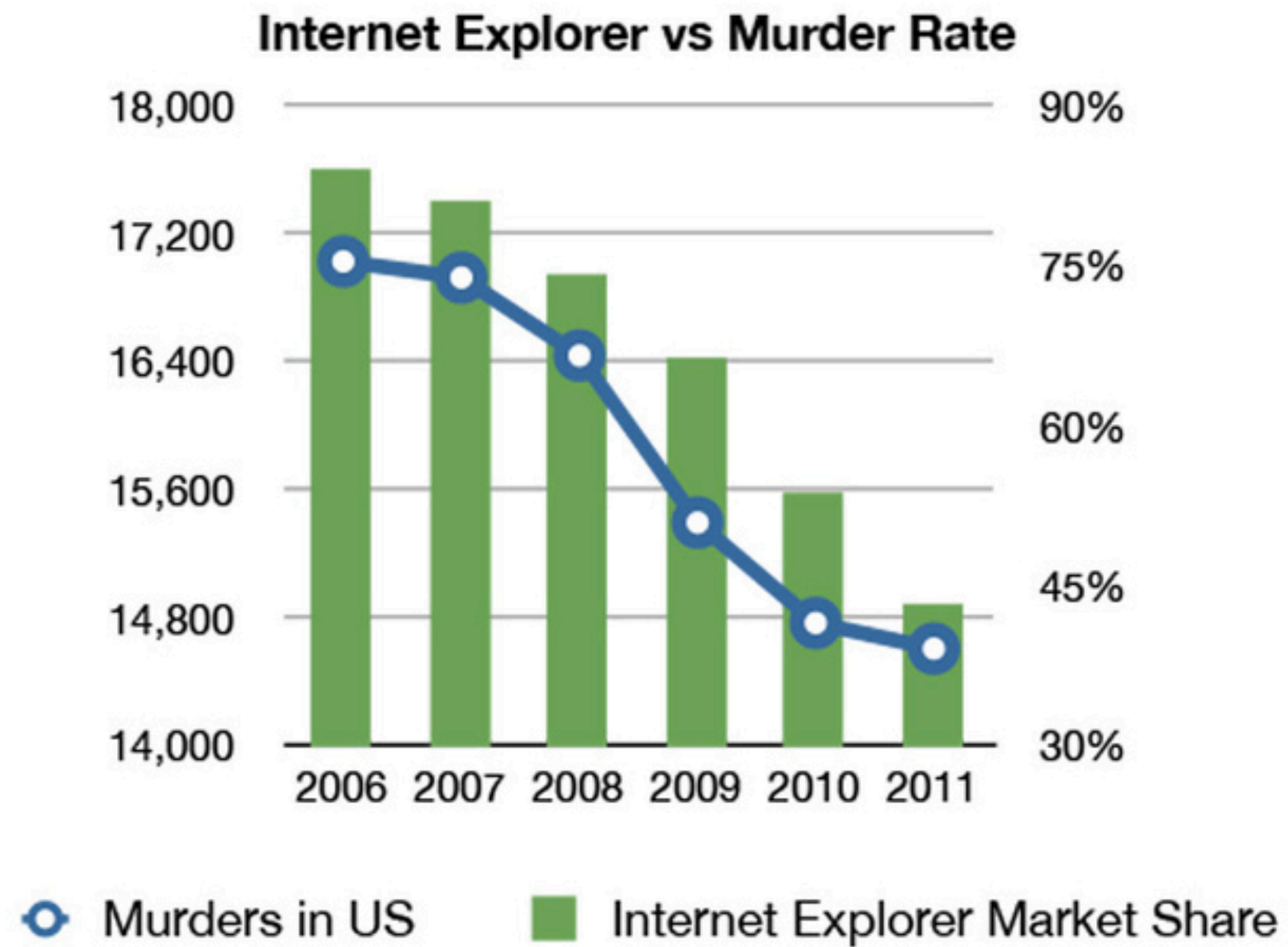


Sources: Organic Trade Association, 2011 Organic Industry Survey; U.S. Department of Education, Office of Special Education Programs, Data Analysis System (DANS), OMB# 1820-0043: "Children with Disabilities Receiving Special Education Under Part B of the Individuals with Disabilities Education Act"

- <http://www.buzzfeed.com/kjh2110/the-10-most-bizarre-correlations>

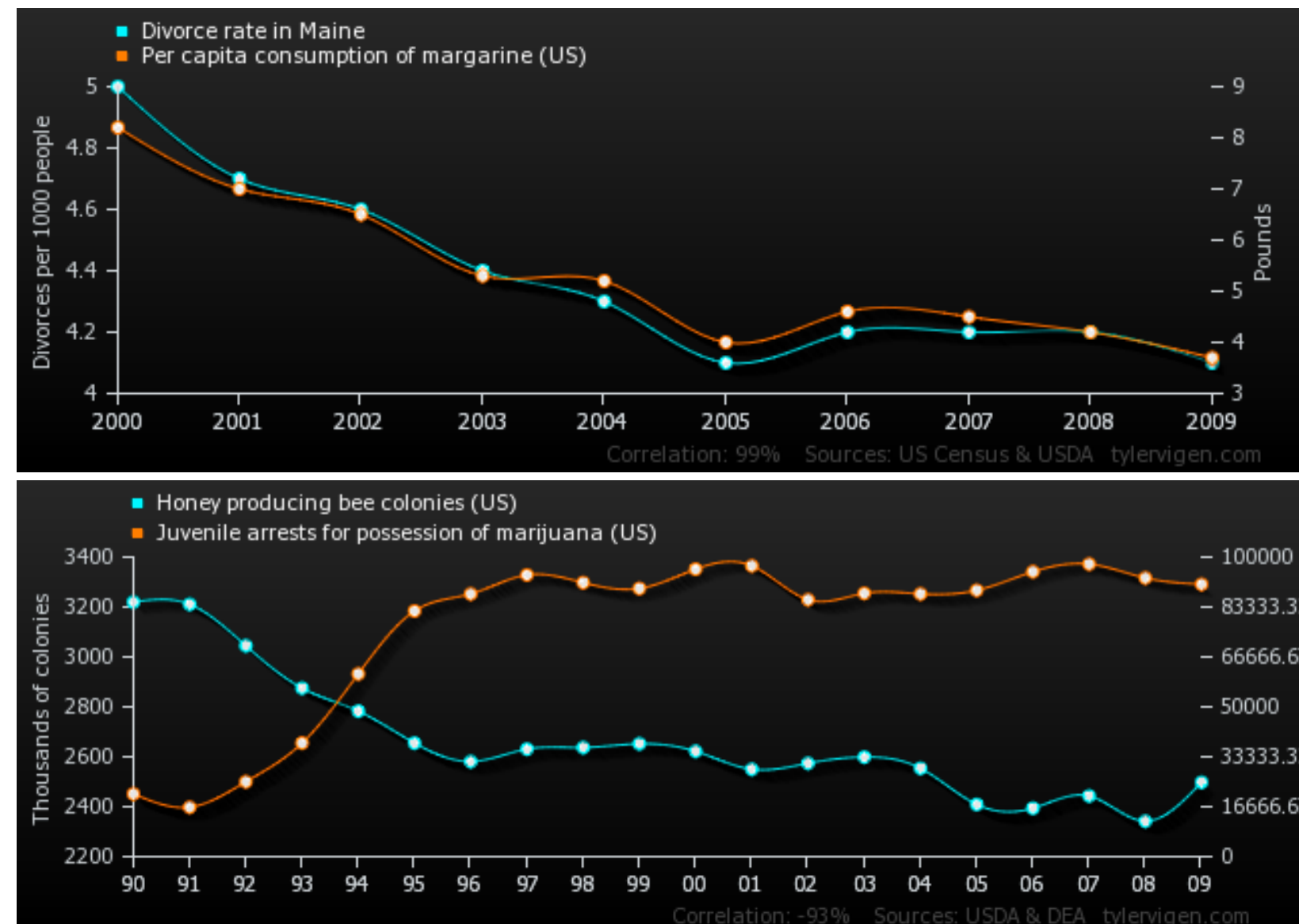
Internet explorer causes murder

- <http://www.buzzfeed.com/kjh2110/the-10-most-bizarre-correlations>



Correlation finders

- http://tylervigen.com/view_correlation?id=3267
- <http://www.google.com/trends/correlate/draw>



Recap: correlation without causation

- Can be a coincidence
- Or a lurking variable
 - common cause
- e.g. http://en.wikipedia.org/wiki/Correlation_does_not_imply_causation
- http://rationalwiki.org/wiki/Correlation_does_not_imply_causation

Best solution

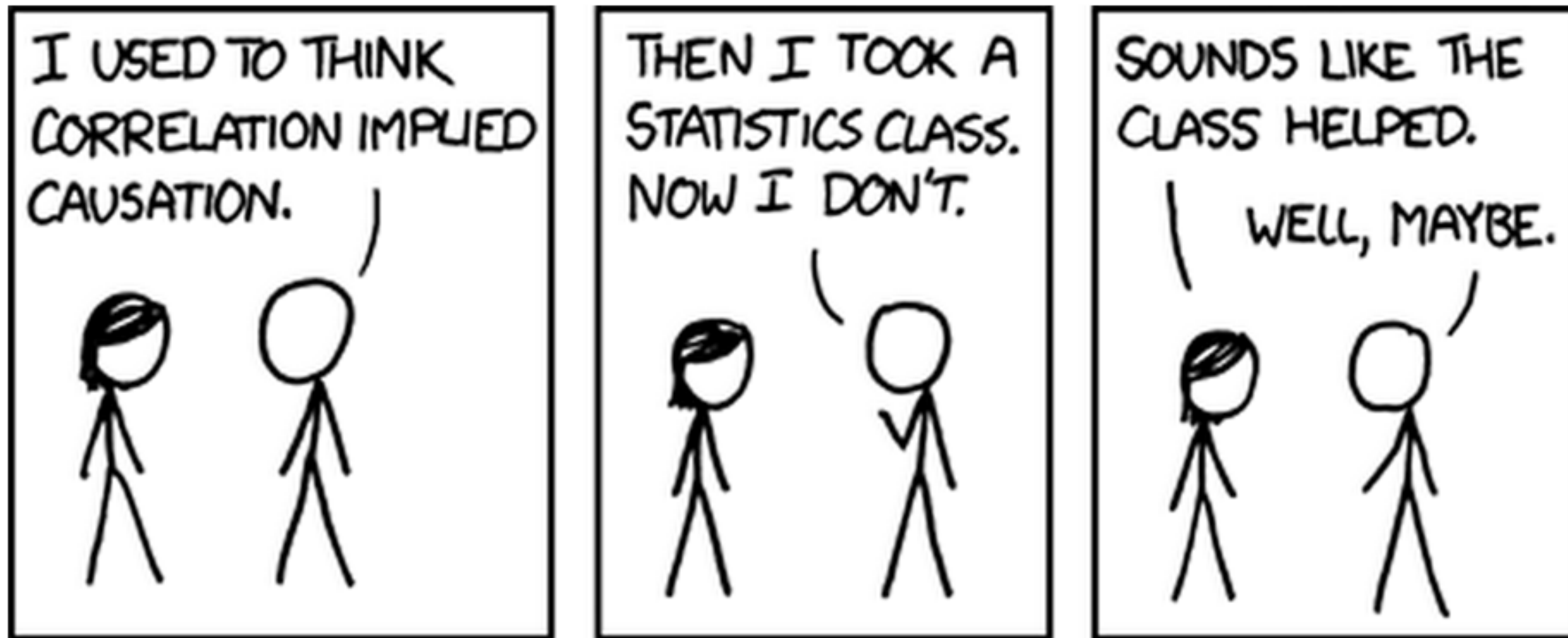
- Randomized trial
- Subjects are assigned randomly to test vs. control group
- Correlation analysis is good for *exploration study*.
 - Points to interesting directions
- But then requires clean randomized *confirmation study*

Milton Friedman's Thermostat Analogy

- <https://justinhohn.typepad.com/blog/2013/01/milton-friedmans-thermostat-analogy.html>
- Imagine a house has a reasonably good thermostat. We would observe negative correlation between outside temperature (O) and energy consumption (E) (more heat needed when colder). We would observe no correlation between inside temperature (I--it's constant, remember) and energy consumption. Also, the constancy of inside temperature shows no correlation to outside temperature, either. In mathematical terms, we see a negative correlation of O to E, but no correlation at all of O to I, or I to E.

Milton Friedman's Thermostat Analogy

- This causes a problem for analyzing the data. One economist might look at the data and conclude that the amount of energy consumption had no effect on indoor temperatures (no correlation). Likewise, the temperature outside has no effect on the temperature inside (again, no correlation). He concludes that the only effect of using more energy is that it appears to reduce outdoor temperatures.
- Another economist thinks the causation runs the opposite way-- that warmer temperatures cause a decrease in the amount of energy used. Convinced that energy consumption and outdoor temperatures are irrelevant as factors that affect indoor temperatures, they turn off the furnace to save energy.



Correlation doesn't imply causation, but it does waggle its eyebrows suggestively and gesture furtively while mouthing 'look over there'.

<http://xkcd.com/552/>









Sharpshooter fallacy & Data dredging

Texas sharp shooter fallacy

- Barn with six targets painted on it, and a bullet hole at the very center of each. “Yes sir,” says the owner of the barn, “I never miss.”

Poll



Texas sharp shooter fallacy

- Barn with six targets painted on it, and a bullet hole at the very center of each. “Yes sir,” says the owner of the barn, “I never miss.”
- “That’s right,” says his spouse, “there ain’t a man in the state of Texas who’s more accurate with a paint gun.”



Texas sharp shooter fallacy

- A million participant raffle was drawn, and Joe was found to be the winner. Afterwards, someone points out that the odds of Joe winning are a million to one, and thus, he couldn't have won randomly and must have cheated. Of course, the chances of anyone else winning was also a million to one, and this person could've accused everyone of cheating. However, the chances of somebody winning is 100% guaranteed. In this case, Joe lucked out. Somebody had to have lucked out.

https://rationalwiki.org/wiki/Texas_sharpshooter_fallacy

Prosecutor's fallacy

- The prosecutor's fallacy[12] (pp 203-205 and Appendix C) has led, in the UK, to the false imprisonment of women for murder when the courts were given the prior statistical likelihood of a woman's 3 children dying from Sudden Infant Death Syndrome as being the chances that their already dead children died from the syndrome. This led to statements from Roy Meadow that the chance they had died of Sudden Infant Death Syndrome were extremely small (one in millions). The courts then handed down convictions in spite of the statistical inevitability that a few women would suffer this tragedy. The convictions were eventually overturned (and Meadow was subsequently struck off the U.K. Medical Register for giving “erroneous” and “misleading” evidence, although this was later reversed by the courts).[13] Meadow's calculations were irrelevant to these cases, but even if they were, using the same methods of calculation would have shown that the odds against two cases of infanticide were even smaller (one in billions).[13]
- http://en.wikipedia.org/wiki/Misuse_of_statistics
- 2 big problems: texas sharshooter + non-independent probabilities

Randomness?

- August 2001, New Scientist, John Eagles of the Royal Cornhill hospital in Aberdeen:
“anorexic women are most likely to have been born in the spring or early summer... Between March and June there were 13% more anorexics born than average, and 30% more in June itself.”

http://www.numberwatch.co.uk/of_birthdays_and_clusters.htm

Data

- total 446 anorexic women
- average of 37 per month

Monte Carlo to the rescue

```
def juneProb(numTrials):  
    june48 = 0.0  
    for trial in range(numTrials):  
        june = 0.0  
        for i in range(446):  
            if random.choice([0,1,2,3,4,5,6,7,8,9,10,11]) == 5:  
                june += 1.0  
        if june >= 48:  
            june48 += 1  
    juneProb = str(june48/numTrials)  
    print 'Probability of at least 48 births in June = ' + juneProb  
  
juneProb(10000)
```

null hypothesis

Probability of at least 48 births in June = 0.04255

< 0.05

Monte Carlo to the rescue

```
def anyProb(numTrials):  
    anyMonth = 0.0  
    for trial in range(numTrials):  
        months = [0.0]*12  
        for i in range(446):  
            months[random.choice([0,1,2,3,4,5,6,7,8,9,10,11])] += 1  
        if max(months) >= 48:  
            anyMonth += 1  
    aProb = str(anyMonth/numTrials)  
    print('Probability of at least 48 births in a Month = ' + aProb )  
  
anyProb(10000)
```

Probability of at least 48 births in a Month = 0.43574

Problems

- 1/ small data size
- 2/ multiple possible hypotheses (one per month)

Sharpshooter fallacy

- A Swedish study in 1992 tried to determine whether power lines caused some kind of poor health effects. The researchers surveyed everyone living within 300 meters of high-voltage power lines over a 25-year period and looked for statistically significant increases in rates of over 800 ailments. The study found that the incidence of childhood leukemia was four times higher among those who lived closest to the power lines, and it spurred calls to action by the Swedish government.[6]

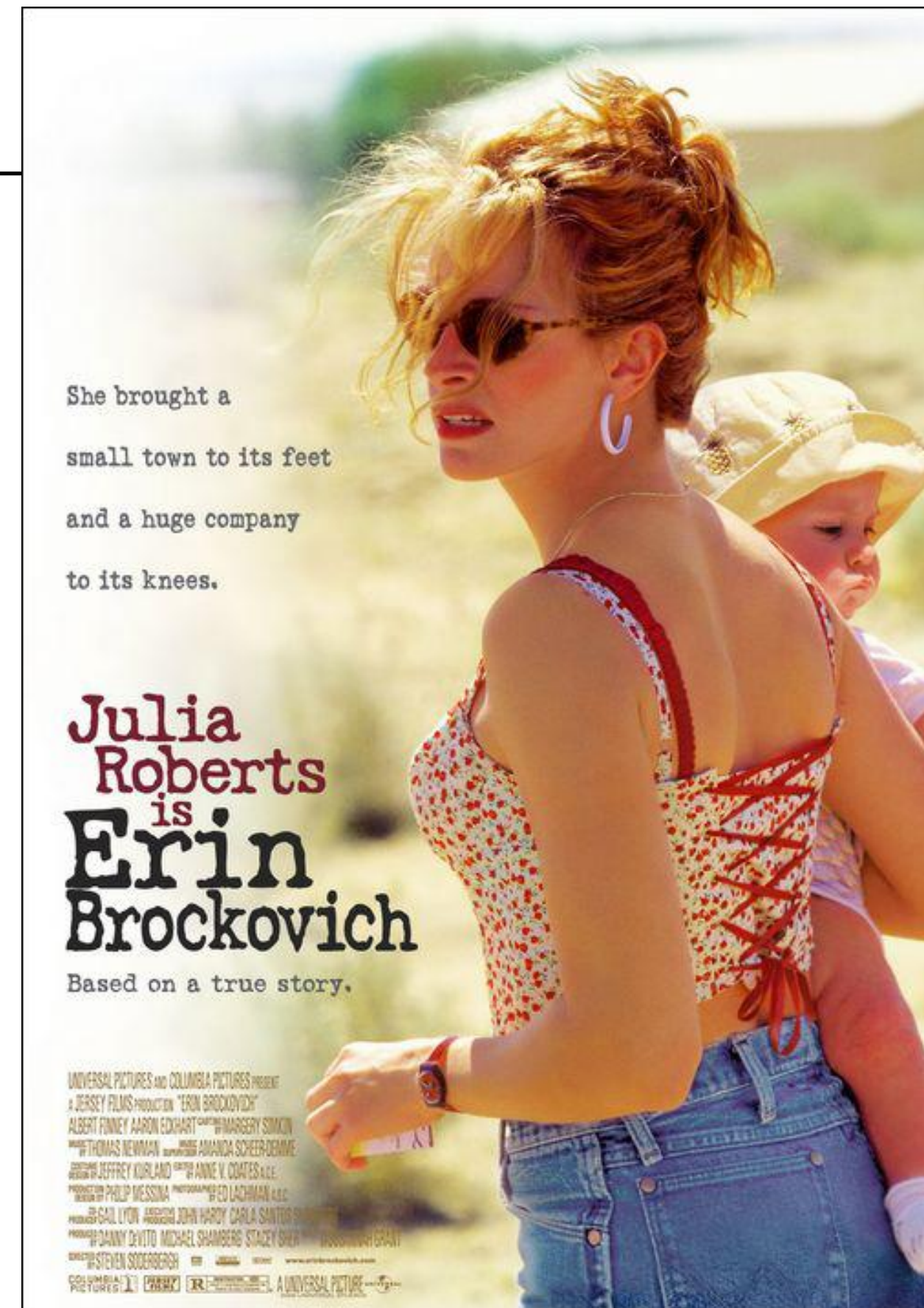
https://en.wikipedia.org/wiki/Texas_sharpshooter_fallacy

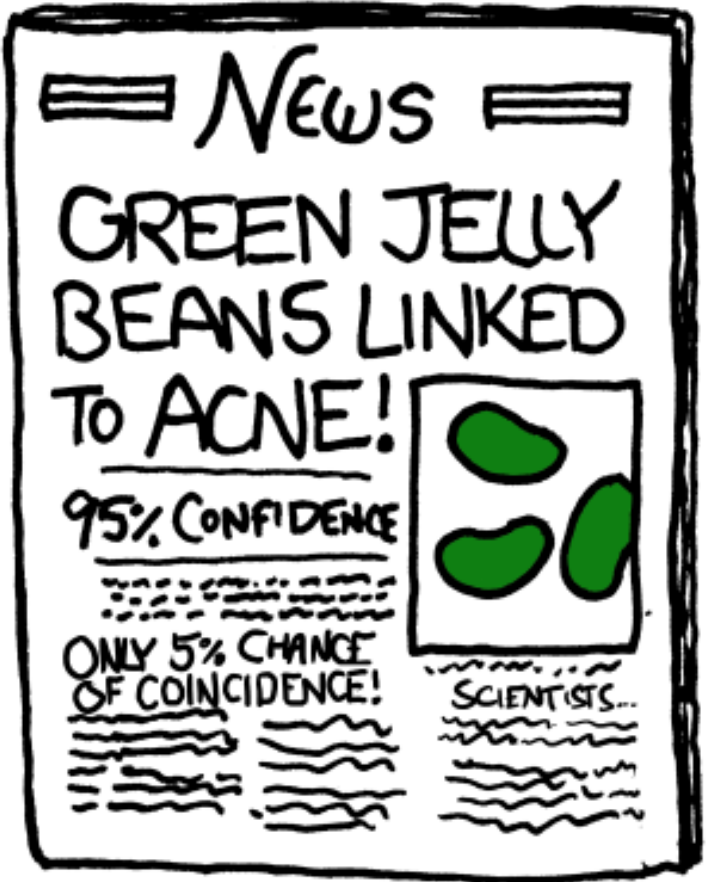
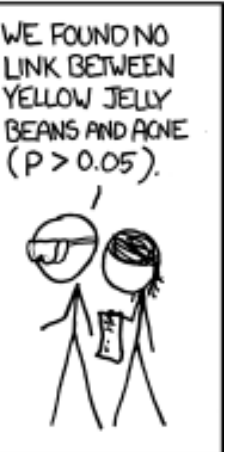
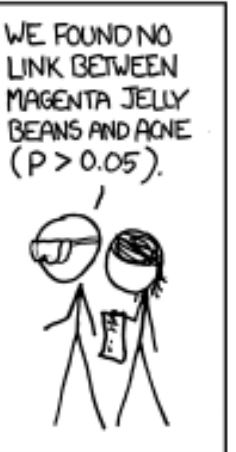
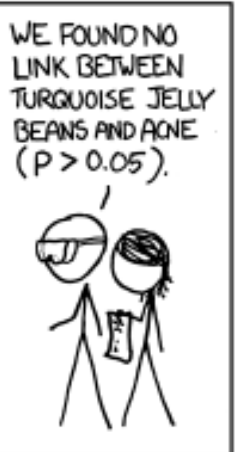
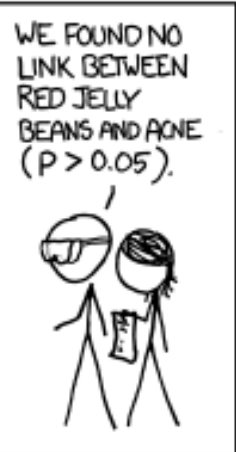
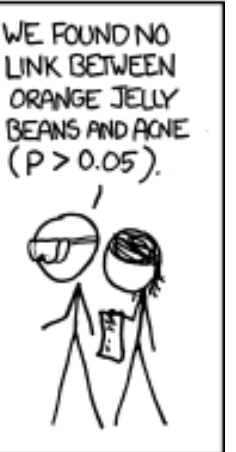
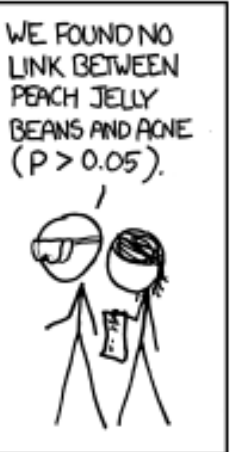
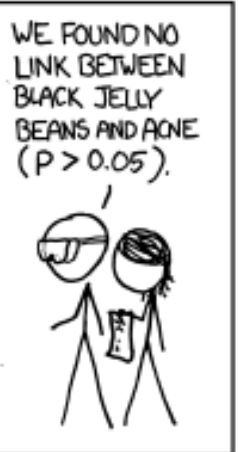
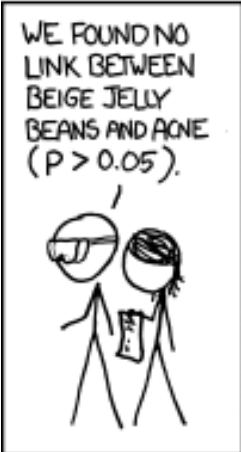
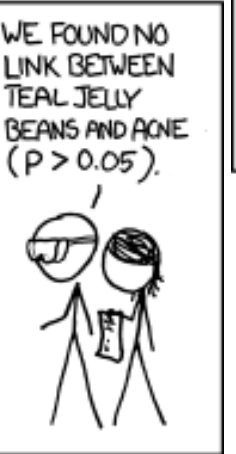
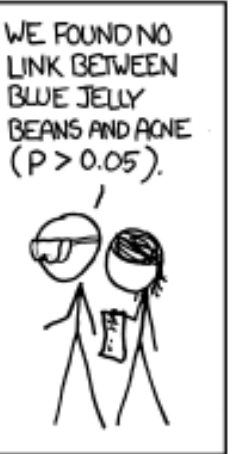
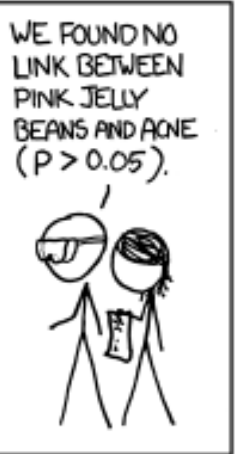
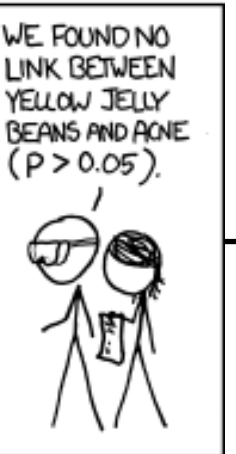
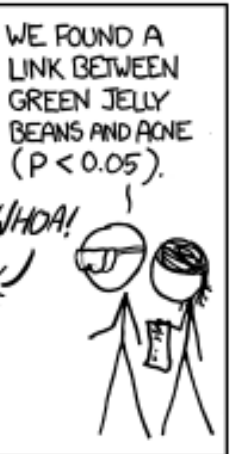
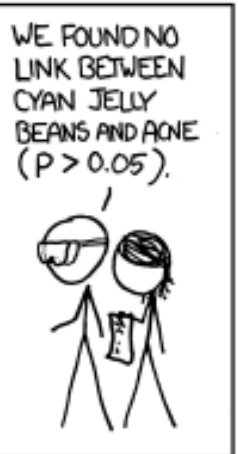
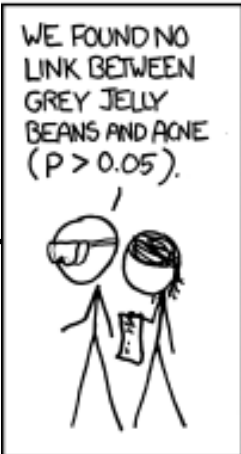
Sharpshooter fallacy

- A Swedish study in 1992 tried to determine whether power lines caused some kind of poor health effects. The researchers surveyed everyone living within 300 meters of high-voltage power lines over a 25-year period and looked for statistically significant increases in rates of over 800 ailments. The study found that the incidence of childhood leukemia was four times higher among those who lived closest to the power lines, and it spurred calls to action by the Swedish government.[6]
- Subsequent studies failed to show any links between power lines and childhood leukemia, neither in causation nor even in correlation.

https://en.wikipedia.org/wiki/Texas_sharpshooter_fallacy

- http://www.slate.com/articles/health_and_science/medical_examiner/2013/03/cancer_cluster_in_toms_river_new_jersey_the_link_to_a_superfund_site_is.single.html





Base rate fallacy

Base rate fallacy

- Test accurate 95% of the time
- Base rate of disease: 0.1 %
- Test on 100,000 people

Base rate fallacy

- Test accurate 95% of the time
- Base rate of disease: 0.1 %
- Test on 100,000 people

		Have Disease	
		Yes	No
Test Results	Positive	95	~ 5000
	Negative	5	~ 99400

Handwritten notes: 100 (above Yes), 99,900 (above No), 5000 (next to Positive No), only 95 (next to Positive Yes)

- Excellent discussion at <http://www.dangriller.com/jumping-to-conclusions-base-rate-neglect/>

Base rate fallacy

- Test accurate 95% of the time
- Base rate of disease: 0.1 %
- Test on 100,000 people

		Have Disease		Probability	
		Yes	No		
Test Results	Positive	95	4,995	1.90%	of having disease if test is positive
	Negative	5	94,905	0.01%	of having disease if test is negative

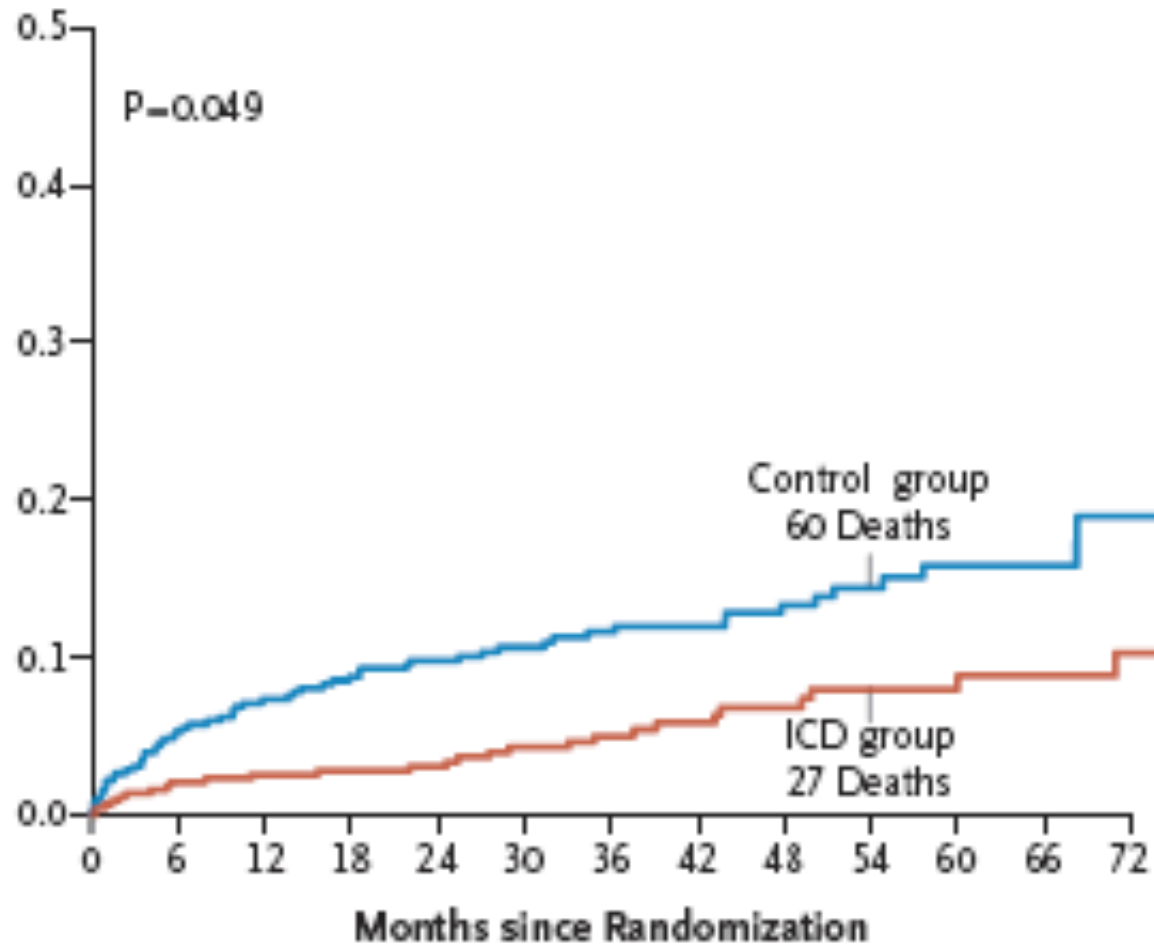
- Excellent discussion at <http://www.dangreller.com/jumping-to-conclusions-base-rate-neglect/>

Implantable cardiac defibrillator

EECS

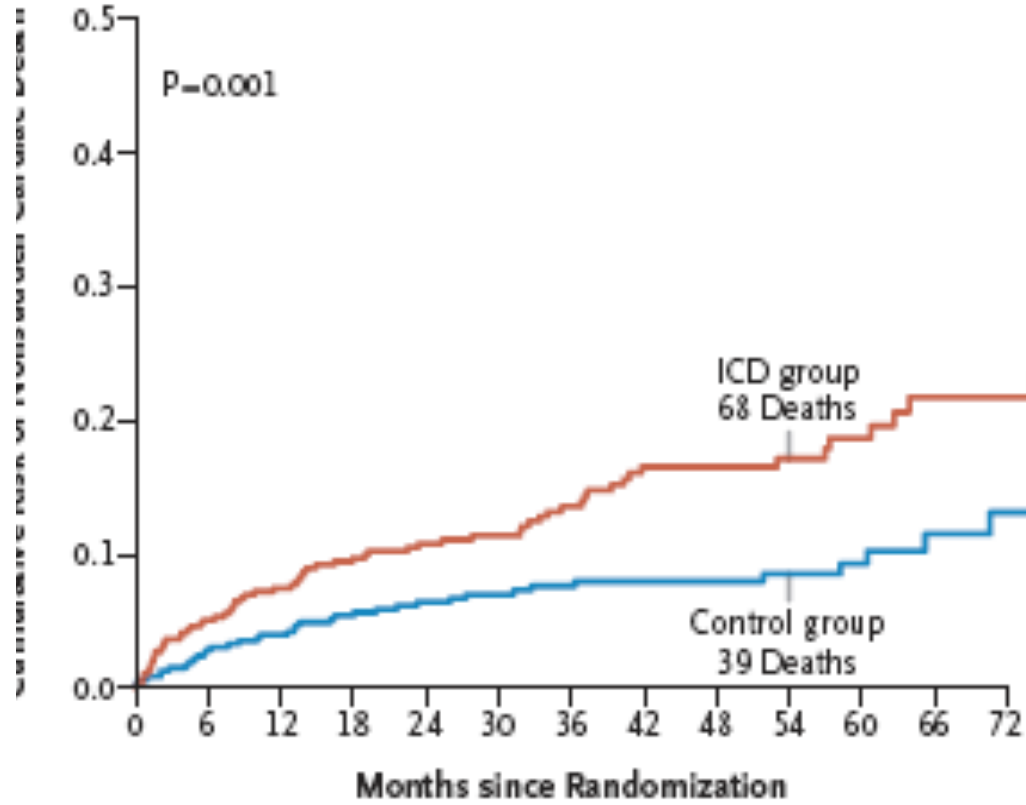


Cumulative Risk of Sudden Cardiac Death

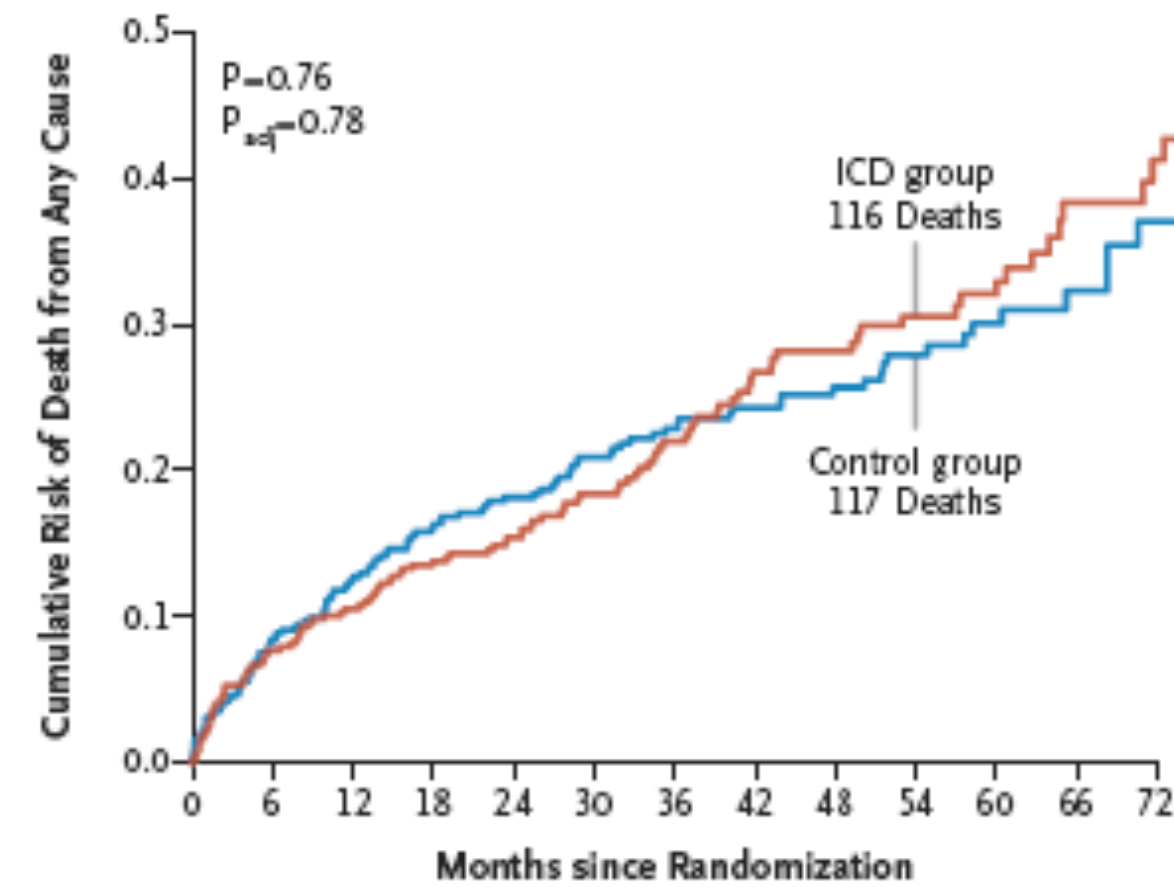


December 11, 2012

Cumulative Risk of Non-Sudden Cardiac Death



©John Guttag



NEJM Oct. 2009

Too many: Potentially risky, always expensive (~\$50k)

90% of recipients received < 0 medical benefit

Too few: 100's of deaths/day potentially avoidable

Base rate fallacy

- False positive error rate is X%
- Does not mean that when you get a positive, the probability that it's false is X%
- Depends on the base rate of positives vs negatives

- See also http://en.wikipedia.org/wiki/Representativeness_heuristic for cognitive basis

Medicine and base rate fallacy

- Medical tests are never 100% accurate
- Treatments can have (negative) side effects
- Think carefully about base rate before treating

Reporting bias

Reporting bias can means 2 things

- **Bad data:**
 - The tendency of subjects to omit vs. report some information might correlate with outcomes
(e.g. subjects may lie about smoking, sexual experience, pain)
- **Bad science/logic**
 - Negative studies don't get reported

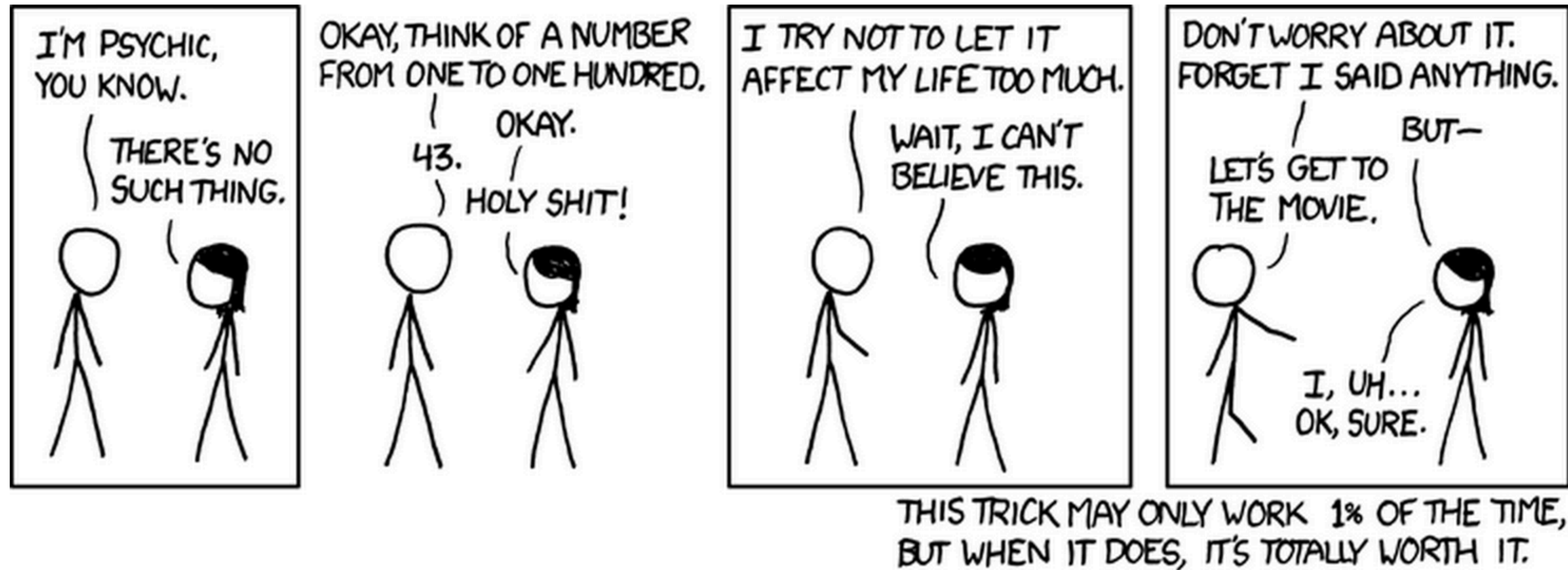
Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy

Erick H. Turner, M.D., Annette M. Matthews, M.D., Eftihia Linardatos, B.S.,
Robert A. Tell, L.C.S.W., and Robert Rosenthal, Ph.D.

- <https://www.nejm.org/doi/pdf/10.1056/NEJMsa065779>

- **RESULTS**

Among 74 FDA-registered studies, 31%, accounting for 3449 study participants, were not published. Whether and how the studies were published were associated with the study outcome. A total of 37 studies viewed by the FDA as having positive results were published; 1 study viewed as positive was not published. Studies viewed by the FDA as having negative or questionable results were, with 3 exceptions, either not published (22 studies) or published in a way that, in our opinion, conveyed a positive outcome (11 studies). According to the published literature, it appeared that 94% of the trials conducted were positive. By contrast, the FDA analysis showed that 51% were positive. Separate meta-analyses of the FDA and journal data sets showed that the increase in effect size ranged from 11 to 69% for individual drugs and was 32% overall.



- See also elaborate scam described by Alfred Hitchcock
<http://www.imdb.com/title/tt0508320/>
- <http://www.hulu.com/watch/48285>



All

Movies, TV
& Showtimes

Celebs, Events
& Photos

News &
Community

Watchlist



Alfred Hitchcock Presents (TV Series)

Mail Order Prophet (1957)

[Edit](#)

Plot

Showing all 1 items

Jump to: [Summaries](#) (1)

Summaries

One day, Ronald Grimes receives a letter from a Mr. Christianai who says he can predict the future. The letter correctly predicts the outcome of an upcoming election. More letters follow and through gambling, Grimes acquires a large amount of money. A final letter from Christianai asks for a contribution. Grimes gives it quite willingly. Later Grimes finds out that Christiani was a fraud. He was really a con man who sent out thousands of letters, half of them predicting one kind of outcome, and the other half predicting another. Grimes was lucky; he got the right predictions time after time.

Science issue

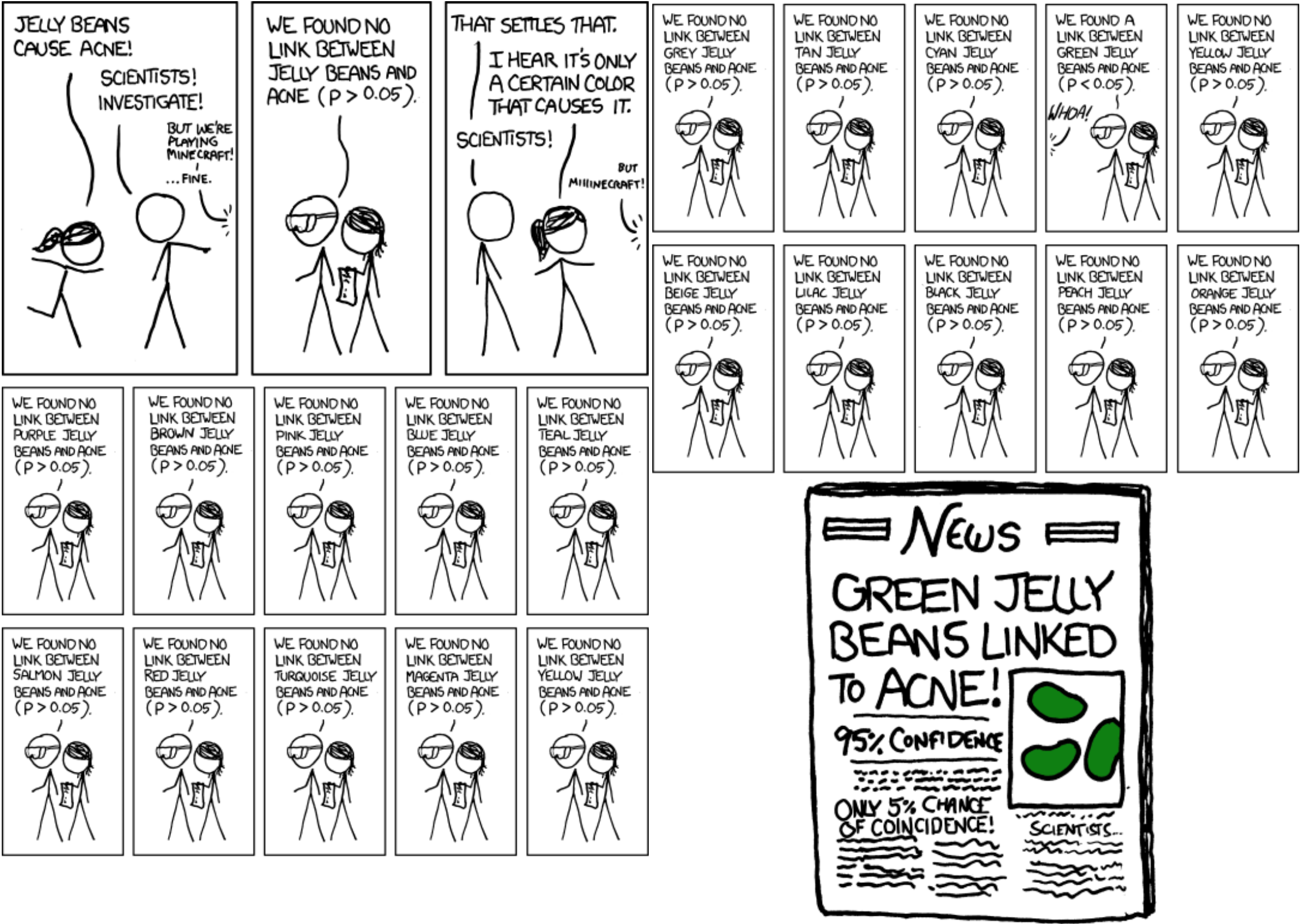
- Sharpshooter + base rate + reporting bias
=> lots of false positives
- <http://www.statisticsonewrong.com/p-value.html>

How science goes wrong (John Ioannidis)

- <https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0020124>
- <http://www.economist.com/news/leaders/21588069-scientific-research-has-changed-world-now-it-needs-change-itself-how-science-goes-wrong>
- <http://www.economist.com/blogs/graphicdetail/2013/10/daily-chart-2>



Essentiall the jelly bean problem



Regression to the mean

Sports illustrated cover jinx?

- http://en.wikipedia.org/wiki/Sports_Illustrated_cover_jinx



Regression to the means (Galton)

- If a variable is extreme on first measurement
- It will likely become less extreme

Coin tosses

- After you've been very lucky...
- You'll probably get less lucky.
- Not a causation. Just because on average you're only 50% lucky
 - Beware of gambler's fallacy

Children of exceptional people

- Are usually less exceptional
- Coined by Galton
 - Darwin's half cousin
 - Unfortunately eugenicist

After you're sick

- You usually get better
- Very important to take into account
 - hence control groups

Common cold study (hypothetical)

- 50% subjects take treatment, 50% take placebo
- Test group: 81% get better
- Placebo group: 79% get better
- Should we conclude that mind controls matter and the psychological effect of the placebo made people better?

Placebo

Placebo

- “simulated or otherwise medically ineffectual treatment for a disease or other medical condition”



Placebo effect in popular culture

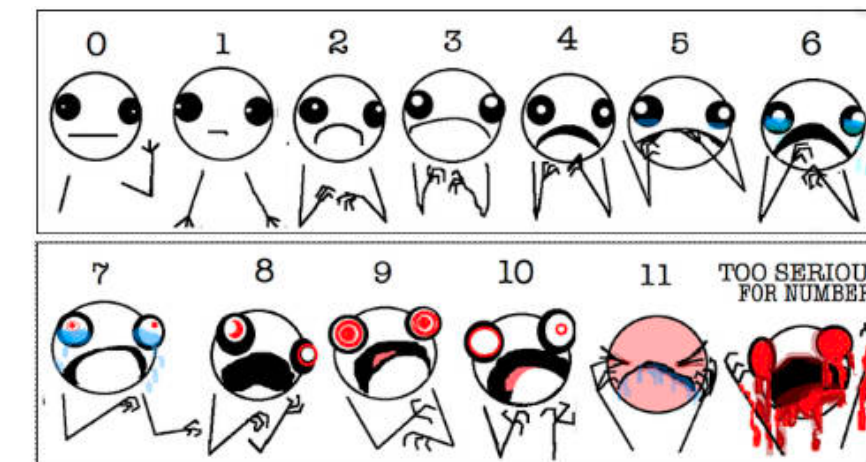
- Power of mind over matter
- Psychosomatic
- “Placebo effect, the tendency of any medication or treatment, even an inert or ineffective one, to exhibit results simply because the recipient believes that it will work”

Actual main role of placebo

- So that the control groups has exactly the same conditions as the test group except one variable: the presence of treatment
- i.e. to compare apples to apples

Placebo non-effects

- Two main factor why patients with placebo (in control group) show improvement:
- Regression to the mean
- For subjective metrics
(e.g. pain level as opposed to objective ones like blood pressure)
Reporting bias:
 - people **say** they feel better to make the doctor happy, or because pain is hard to self-assess
(Subject-expectancy effect)



<http://hyperboleandahalf.blogspot.com/2010/02/boyfriend-doesnt-have-ebola-probably.html>

Is the Placebo Powerless?

- Third, more questionable factor:
psychosomatic “placebo effect”
has not been clearly established for objective measures
 - <http://www.nejm.org/doi/full/10.1056/NEJM200105243442106>
 - <http://onlinelibrary.wiley.com/doi/10.1111/j.1365-2796.2004.01355.x/abstract>
 - <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD003974.pub3/abstract>
 -



<http://comicskingdom.com/pardon-my-planet/2014-05-14>

Gold standard

Standard of evidence

- 1- Methodologically rigorous, properly blinded, and sufficiently powered studies that adequately define and control for all relevant variables (confirmed by surviving peer-review and post-publication analysis).
- 2- Positive results that are statistically significant.
- 3- A reasonable signal to noise ratio (clinically significant for medical studies, or generally well within our ability to confidently detect).
- 4- Independently reproducible. No matter who repeats the experiment, the effect is reliably detected.

<http://www.sciencebasedmedicine.org/evidence-thresholds/>

-
- <http://www.nature.com/news/policy-twenty-tips-for-interpreting-scientific-claims-1.14183>

Recap

Fallacies, statistics and science

- Correlation & causation
- Statistically vs. practically significant
- Data dredging (Texas sharpshooter fallacy)
- Base rate fallacy
- Reporting bias (file drawer fallacy)
- Regression to the mean
- Golden standard: replicated, well-powered randomized blind trial with controls
 - Placebo: mostly for control, psychosomatic effect debatable

Nature's 20 tips

<http://www.nature.com/news/policy-twenty-tips-for-interpreting-scientific-claims-1.14183>

- Differences and chance cause variation
- No measurement is exact
- Bias is rife
- Bigger is usually better for sample size
- Correlation does not imply causation
- Regression to the mean can mislead
- Extrapolating beyond the data is risky
- Beware the base-rate fallacy
- Controls are important
- Randomization avoids bias
- Seek replication, not pseudoreplication
- Scientists are human
- Significance is significant
- Separate no effect from non-significance
- Effect size matters
- Study relevance limits generalizations
- Feelings influence risk perception
- Dependencies change the risks
- Data can be dredged or cherry picked.
- Extreme measurements may mislead

Computational thinking

- Formulate solution to a problem as a computation
- Identify or invent useful abstractions
- Design and construct a sufficiently efficient implementation of experiment
- Validate experimental setup (i.e., debug it)
- Run experiment
- Evaluate results of experiment
- Repeat as needed

The Two A's of Computational Thinking

- **Abstraction**

- Choosing the right abstractions
- Operating in terms of multiple layers of abstraction simultaneously
- Defining the relationships the between layers

- **Automation**

- Think in terms of mechanizing our abstractions
- Mechanization is possible
 - Because we have precise notations and models
 - There is some “machine” that can interpret our notations

What next?

- **Many of you have worked very hard**
 - staff and I appreciate it
- **Only you know your return on investment**
 - Take a look at early problem sets
 - Think about what you'd be willing tackle now
- **Remember that you can write programs to get answers**
- **There are other EECS courses you are prepared to take**
 - 6.01, 6.009, 6.004, 6.006, 6.034, 6.036, 6.031
- **You could major in Course VI**
- **You are qualified for interesting UROP's involving computation**

Aside

- Best thing you can do at MIT is a urop
- Also critical to get into graduate school

- Wash hands, wear a mask, try to stay 6 feet away