



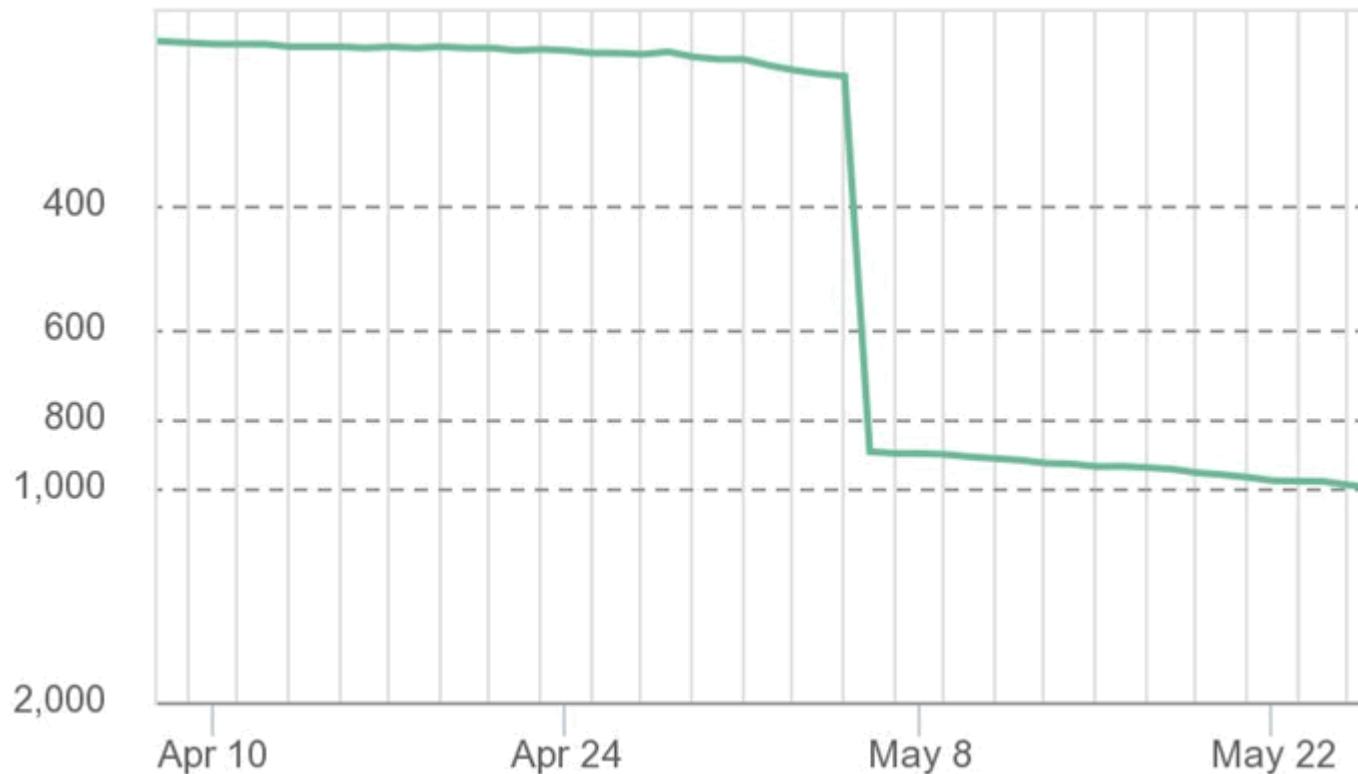
How to lie with statistics – Ein Kurs für die HAW in Hamburg

Prof. Dr. Katharina A. Zweig

**Eure persönliche
Matheberaterin
Nina**



What would you do with the following diagram?



Relative rank of Breitbart, according to Alexa as shown in Niggemeier's blog:
<http://uebermedien.de/16362/verlorene-breitbart-leser-wiedergefunden/>

What did others do?



- Most media titled: Breitbart loses acceptance of broad readership¹.
- Niggemeier found out that something happened in December 16 AND in May 17.



1. e.g.: <http://www.vanityfair.com/news/2017/05/breitbart-traffic-numbers-are-cratering>

What's the problem?



- According to Niggemeier, Breitbart registered as „certified“ website with Alexa in December 16.
- Certified websites are allowed to measure the traffic on their own servers – otherwise it is approximated by sampling.
- (Nonetheless, the interest in Breitbart has decreased, but not as drastically as stated beforehand).

Take Home Messages

Statistics is often
rather
unintuitive.



It's easy to be fooled



It's even easier to
fool others.
(but evil...)





Statistics

Exercise: Interpretation of data



- Based on the data, flight attendant's union says: average salary is \$ 26,700 – almost poverty level!
- Chief of airline says: average salary of full-time worker was over \$ 150,000!
- How would you describe it?

Job Title	Months Worked	Total Wages
President	12	\$920,000
Vice President	12	\$350,000
Vice President	12	\$350,000
Pilot	12	\$85,000
Pilot	12	\$72,000
Attendant	12	\$36,000
Attendant	12	\$36,000
Attendant	12	\$34,000
Attendant	12	\$32,000
Attendant	12	\$29,000
Attendant	12	\$29,000
Attendant	12	\$28,000
Attendant	12	\$28,000
Attendant	6	\$13,000
Attendant	1	\$2,000

Statistik des Tages



Skandal! Nur 1,4 % aller SpiegelOnline-Artikel arbeiten mit Statistiken!

US-Warenhandel mit ausgewählten Länder

Angaben in Milliarden Dollar



Mexiko

Quelle: US-Handelsministerium (Gesamtjahr 2015)

SPIEGEL ONLINE

Definition: Prozent



- Prozente beschreiben den Anteil einer Sache von Interesse an einer fiktiven Grundgesamtheit von 100.
- Wenn von 16532 jugendlichen Kriminellen 3306 wieder rückfällig werden, sind dies $(3306/16532)*100 = 20\%$.
- Von denen, die rückfällig werden, sind 2% Frauen.

Wieviele der jugendlichen Kriminellen sind Frauen, die wieder rückfällig werden?

- 1) 10%
- 2) 2%
- 3) 20%
- 4) 4%
- 5) 0,4%

Definition: Prozent



- Prozente beschreiben den Anteil einer Sache von Interesse an einer fiktiven Grundgesamtheit von 100.
- Wenn von 16532 jugendlichen Kriminellen 3306 wieder rückfällig werden, sind dies $(3306/16532)*100 = 20\%$.
- Von denen, die rückfällig werden, sind 2% Frauen.

Wieviele der jugendlichen Kriminellen sind Frauen, die wieder rückfällig werden?

- 1) 10%
- 2) 2%
- 3) 20%
- 4) 4%
- 5) **0,4%**

Natürliche Frequenzen



- Mathematisch gesehen kann man die Prozentzahlen einfach miteinander multiplizieren: $20/100 * 2/100 * 100\% = 40/100\% = 0.4\%$
- Der beste Trick besteht aber darin, die Prozentzahlen in „natürliche Frequenzen“ zu übersetzen.
- Dafür wählt man sich eine Grundgesamtheit, so dass der prozentuale Anteil sich als eine natürliche Zahl davon darstellen lässt:
 - 20% ist gleichbedeutend mit 20 von 100, 200 von 1.000, 2.000 von 10.000 und so weiter.
 - Davon 2% wären dann: 0.4 von 100, 4 von 1.000, 40 von 10.000 und so weiter.
 - So sieht man gleich, dass die Antwort 0,4% beträgt.
- Die Wahl der Grundgesamtheit ist eigentlich egal, aber es bietet sich an, sie so zu wählen, dass man während der Berechnung keine Zahlen < 1 erhält.

Alle 11 Minuten...



- Aus einer Unstatistik-Kolumne von Prof. Dr. Thomas Bauer¹.
- Wenn sich (der leichteren Rechnung zuliebe) alle 10 Minuten 2 Mitglieder verlieben,
- sind das pro Jahr 105.120 verliebte Mitglieder.
- Von geschätzten 5 Millionen Mitgliedern wären das nur ca. 2%, von 1 Million dementsprechend ca. 10%.
- Ist das erfolgsversprechend?



¹ <http://www.rwi-essen.de/unstatistik/50/>



Prozent von was?

Die richtige Basis wählen

Faires Urteil?



- Einem Arbeitnehmer wurde zu Unrecht das Gehalt um 25% gekürzt.
- Er zieht vor Gericht.
- Der Richter entscheidet, dass dem Arbeitnehmer auf sein jetziges Gehalt wieder eine 25%ige Gehaltserhöhung gezahlt werden muss, damit er wieder beim Ausgangsgehalt ist.

Ist das ein

- 1) faires Urteil?
- 2) unfaires Urteil?



Prozentrechnung richtig gemacht



- Das Urteil ist unfair, weil der Arbeitnehmer danach nicht wieder 100% seines Anfangsgehaltes hat.
- Immer bedenken, von was die Prozente berechnet werden.
- Wenn die Person vorher 2000 € brutto hatte, dann hat sie nach der Kürzung 1500 € brutto.
- Davon 25% sind nur 375 €, also insgesamt nur 1875 € brutto.
- Er hätte eine Erhöhung von $\frac{1}{3}$ benötigt, um wieder beim Ausgangsgehalt anzukommen.



Prozent von was?

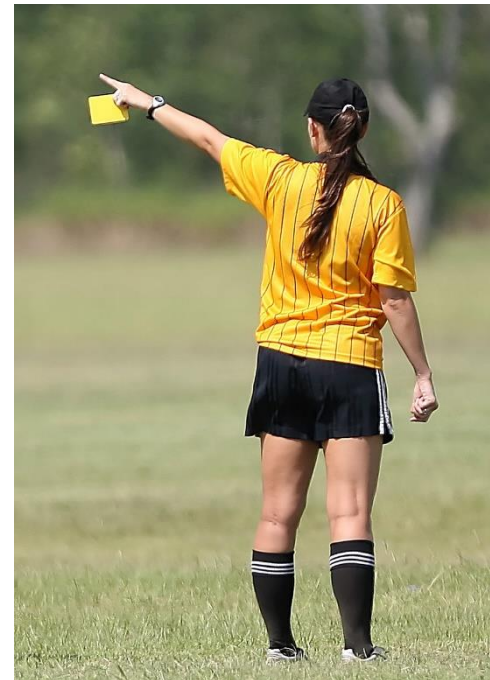


- Online-Zeitung des Schleswig-Holsteinischer Zeitungsverlages basierend auf einer Studie von XING (Durchsuchung von 45.000 Nutzerprofilen):
- „Die gute Nachricht zuerst: Mehr als die Hälfte (63,5 Prozent) der HSV-Fans hat einen Hochschulabschluss. Die schlechte Nachricht: Damit landen sie aber im Vergleich zu anderen Clubs auf dem drittletzten Platz.“ 24.8.2016)



Richtig ist:

- Von den Personen, die sich auf XING als Fußballfans outen, sind von den HSV-Fans 63% Personen mit Hochschulabschluss.
- Diese Teilmenge ist natürlich **nicht** repräsentativ für alle Fans von Fussballclubs¹.
- Ganz klar: Gelbe Karte!



1 <http://www.rwi-essen.de/unstatistik/58/>

Cola



- Laut der Unstatistik des Monats vom 30.6.2015 warb LIDL folgendermaßen für ihre Freeway-Cola:
 - Von 510 Befragten denken 80% der Personen, die Freeway-Cola am liebsten mochten, dass diese wie Pepsi oder Coca Cola schmeckte.
 - Laut Prof. Dr. Thomas Bauer fehlt die Angabe, dass nur 18% der Teilnehmer Freeway-Cola als ihre Lieblingscola angaben¹.
 - Von wieviel Prozent der Teilnehmer wissen wir somit, dass sie Freeway-Cola nicht von Pepsi und Coca Cola unterscheiden konnten?

1) 62%

2) 18%

3) 14%

4) 10%

1 <http://www.rwi-essen.de/unstatistik/44/>

Cola



- Laut der Unstatistik des Monats vom 30.6.2015 warb LIDL folgendermaßen für ihre Freeway-Cola:
 - Von 510 Befragten denken 80% der Personen, die Freeway-Cola am liebsten mochten, dass diese wie Pepsi oder Coca Cola schmeckte.
 - Laut Prof. Dr. Thomas Bauer fehlt die Angabe, dass nur 18% der Teilnehmer Freeway-Cola als ihre Lieblingscola angaben¹.
 - Von wieviel Prozent der Teilnehmer wissen wir somit, dass sie Freeway-Cola nicht von Pepsi und Coca Cola unterscheiden konnten?

1) 62%

2) 18%

3) 14%

4) 10%

1 <http://www.rwi-essen.de/unstatistik/44/>

Antwort



92 gaben Freeway als Lieblingscola an, davon sind 80% nur 74 Teilnehmer, also 14% von 510.

Es ist aber möglich, dass weitere Personen die Colas nicht unterscheiden konnten (und dann auf Pepsi oder Coca Cola zeigten).



Aufgaben



- Weitere Beispiele:
 - Arbeitsblatt, S. 5 und 6





Irreführende Grundgesamtheit

Faulty Basis



“That said, gimme another beer.”

Response of a driver to the fact that
70% of all accidents are caused by
non-drunk drivers.

Translated from Krämer, “So lügt man mit Statistik”,
Piper Verlag, München, 2000, S. 27

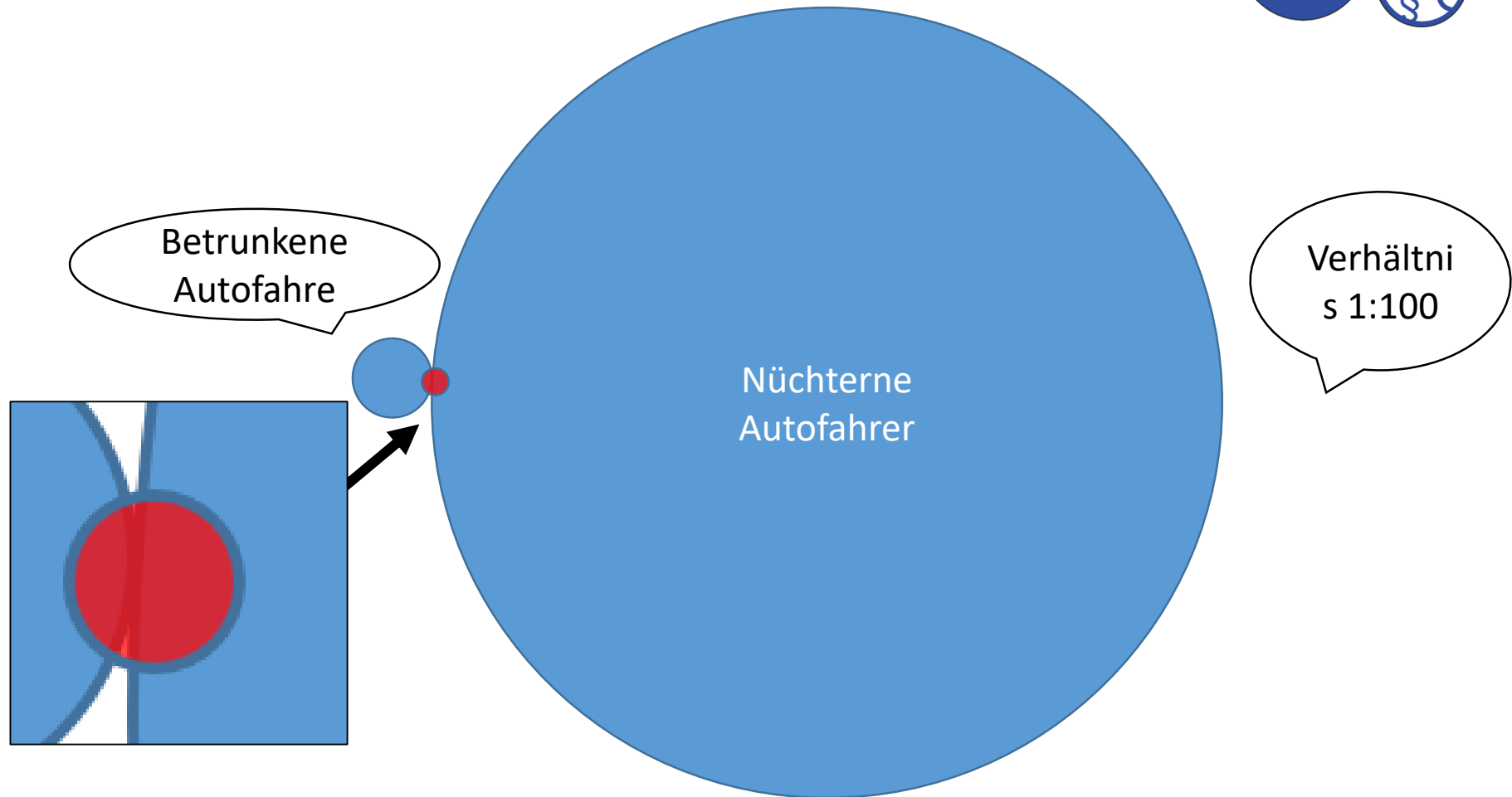
Betrunkene Autofahrer



- Nach dieser Statistik verursachen sie 30% der Unfälle.
- Zu keinem Zeitpunkt (Köln im Karneval mag eine Ausnahme sein?) machen Betrunkene 30% der Autofahrer aus.
- Daher ist Alkohol ein nicht zu unterschätzender Risikofaktor beim Autofahren.



Venn-Diagramme

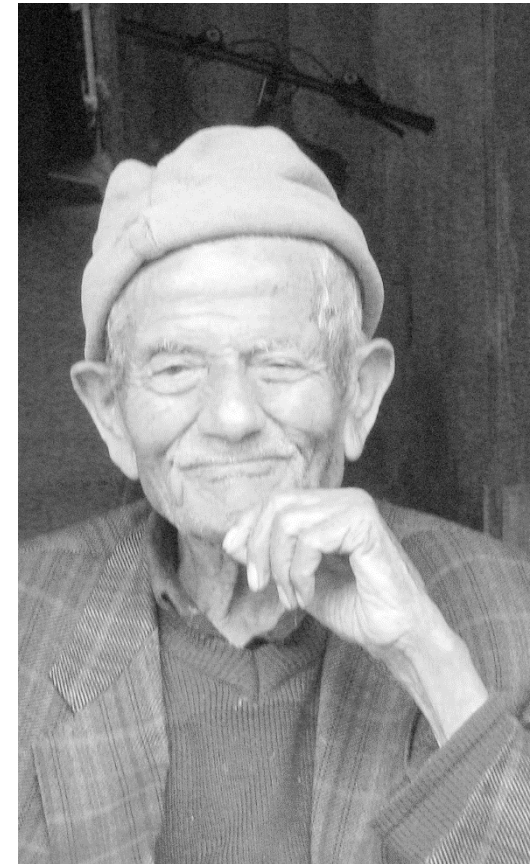


Happier when older?



- Der Tod durch Selbstmord unter allen Todesfällen bei Teenagern beträgt 25%.
 - Unter den 30-40jährigen sind es nur noch 10% aller Todesfälle.
 - Unter den 70jährigen sind es noch 2%.
→ Sind alte Leute weniger selbstmordgefährdet?
- 1) Sieht so aus, ja!
 - 2) Nein, das kann man aus diesen Daten nicht ableiten.

Aus Krämer, "So lügt man mit Statistik",
Piper Verlag, München, 2000, S. 33



Richtige Antwort



- Die Grundgesamtheit ändert sich, denn junge Menschen sterben glücklicherweise nur selten.
- Unfälle und Selbstmorde machen daher einen großen Anteil aus.
- Ältere Personen sterben vor allen Dingen aus natürlichen Gründen und an verschiedenen Krankheiten, da machen Unfälle und Selbstmorde nur noch einen kleinen Anteil aus.





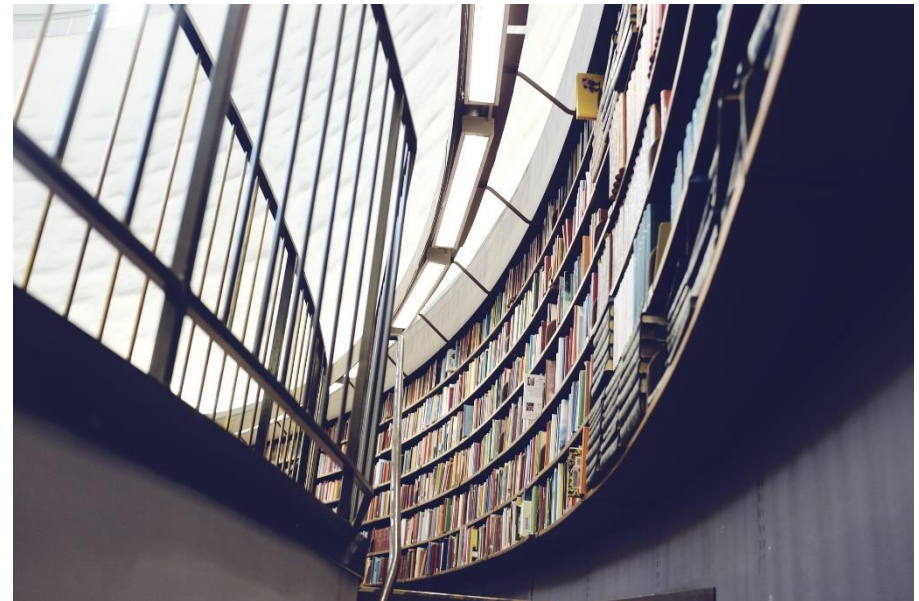
Yule-Simpson Effekt

Warum Wahrscheinlichkeiten manchmal so schwer zu interpretieren sind

Klage wegen Diskriminierung



- An einer Universität bewerben sich 1000 Frauen und 1000 Männer auf einen Studienplatz.
- Nur 180 Frauen, aber 820 Männer werden angenommen.
- Eine Frau klagt, aber ihrer Klage wird nicht stattgegeben, da das Verfahren fair ist. Wie kann das sein?



Simpson-Effekt



- Es gibt zwei Studiengänge (Vorsicht, Klischee!)
 - Pädagogik mit 100 Studienplätzen
 - Informatik mit 900 Studienplätzen.
 - 900 Frauen bewerben sich auf Pädagogik, 100 auf Informatik.
 - Bei den Männern ist es umgekehrt.
- Es wird jeweils proportional zum Geschlechtsverhältnis angenommen.
- D.h., je 90 Frauen in der Pädagogik und Informatik, damit insgesamt 180 Frauen.
- Von den Männern bekommen 10 eine Zulassung zur Pädagogik, und 810 eine zur Informatik.



Autoproduktion: Produktivität vs. Qualität

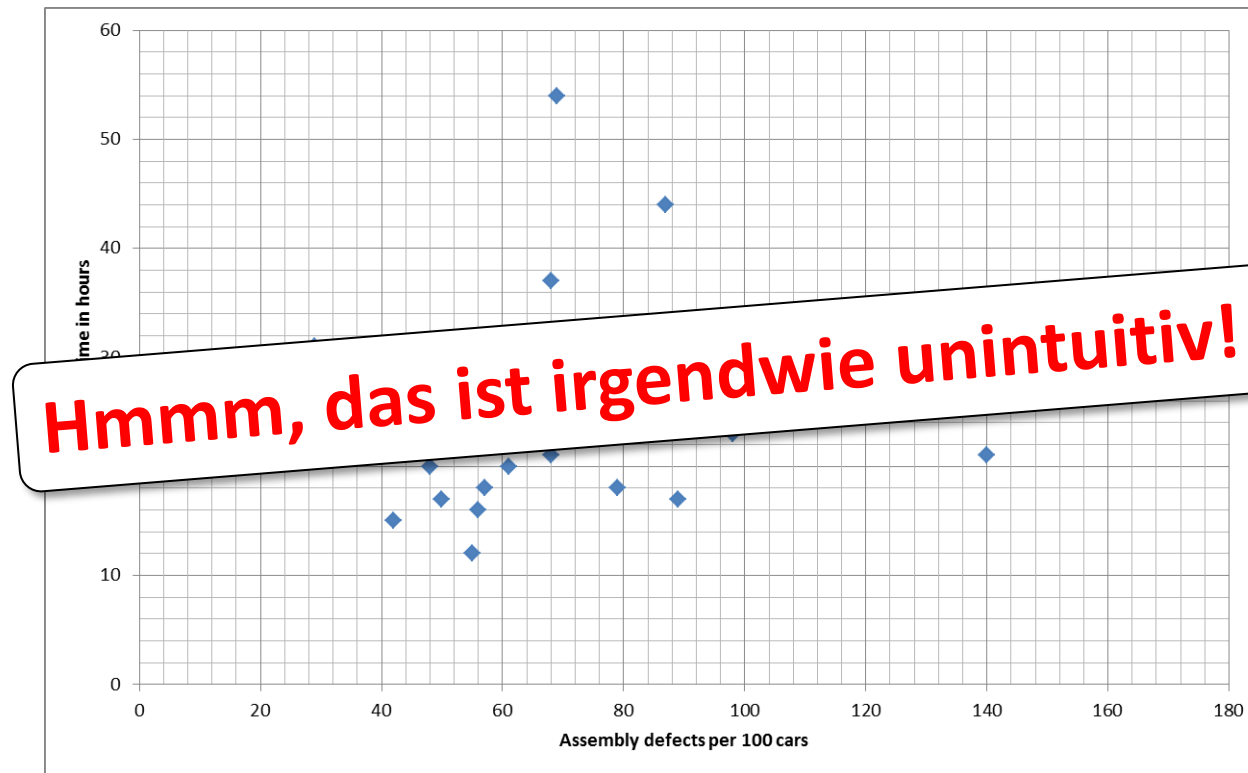


- Produktivität: gemessen in Stunden pro Auto [h/Auto]
- Qualität: Fehlerhaft zusammengebaute Autos pro 100 Autos
- Hypothese: Je länger der Zusammenbau, desto weniger Fehler



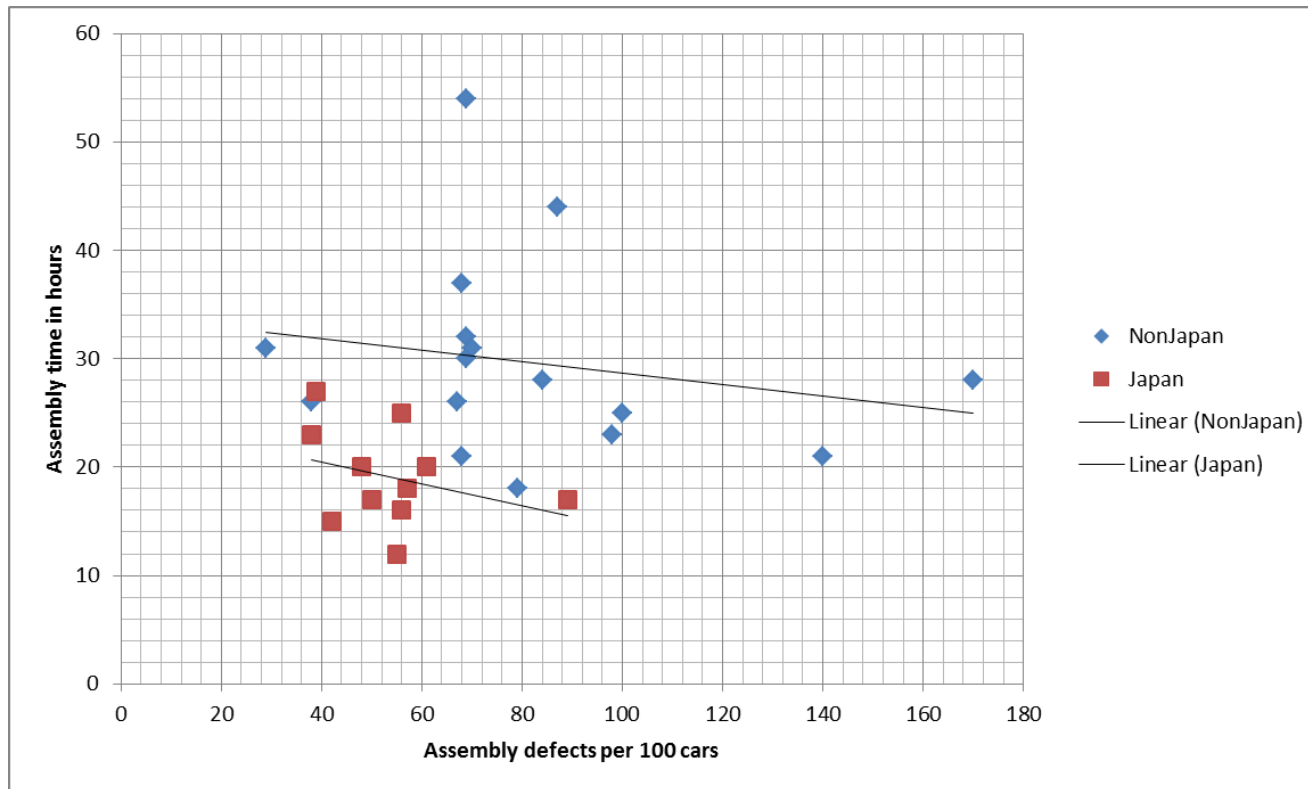
Data from: „A casebook for a first course in statistics and data analysis“, Samprit Chatterjee, Mark S. Handcock, Jeffrey S. Simonoff, Jon Wiley & Sons, Inc., USA, 1995. Data available from <http://www.stat.ucla.edu/~handcock/> under books -> A casebook ... -> Data files ...

Productivity vs. Quality



Reconstruction of figure on p. 26 in Chatterjee et al. 1995

Divided in Subgroups



Reconstruction of figure on p. 28 in Chatterjee et al. 1995

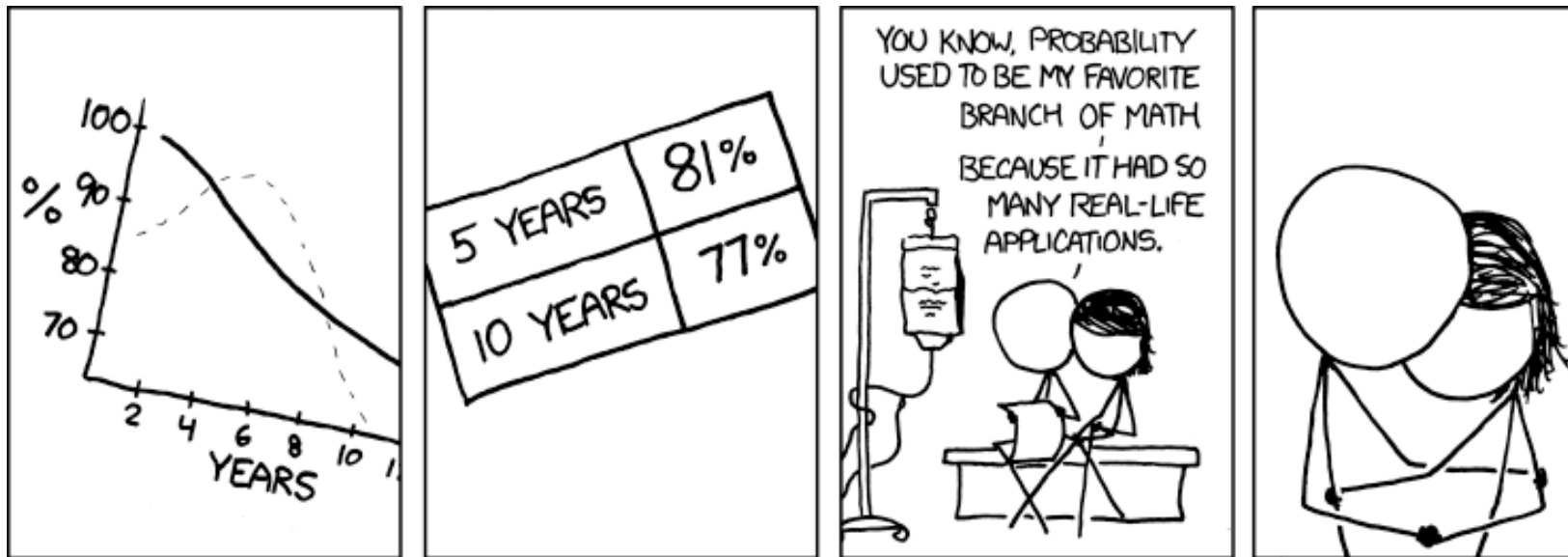


Anwendung: Predictive Policing

Predictive Policing



German Criminal Rate	<table border="1"><tr><td>A</td><td>B</td></tr><tr><td>C</td><td>D</td></tr></table>	A	B	C	D	<table border="1"><tr><td>A-B</td></tr><tr><td>C-D</td></tr></table>	A-B	C-D	<table border="1"><tr><td>A-C</td><td>B-D</td></tr></table>	A-C	B-D	<table border="1"><tr><td>A-B-C-D</td></tr></table>	A-B-C-D
	A	B											
C	D												
A-B													
C-D													
A-C	B-D												
A-B-C-D													
Foreigner Criminal Rate	<table border="1"><tr><td>A</td><td>B</td></tr><tr><td>C</td><td>D</td></tr></table>	A	B	C	D	<table border="1"><tr><td>A-B</td></tr><tr><td>C-D</td></tr></table>	A-B	C-D	<table border="1"><tr><td>A-C</td><td>B-D</td></tr></table>	A-C	B-D	<table border="1"><tr><td>A-B-C-D</td></tr></table>	A-B-C-D
A	B												
C	D												
A-B													
C-D													
A-C	B-D												
A-B-C-D													



Disclaimer: auf den nächsten Folien geht es u.a. um Krebs, Selbstmord, HIV, und den plötzlichen Kindstod.

<https://xkcd.com/881/>



5 Minuten PAUSE





Von Prozenten zu
Wahrscheinlichkeiten

Wahrscheinlichkeiten



- Zuerst muss es eine Menge von Ereignissen geben, die jeweils nur exklusiv auftreten können und zusammen die Menge aller möglichen Ereignisse beschreiben.
 - Beispiel: Jeder Zeitpunkt lässt sich eindeutig einem Jahr zuschreiben.
 - Gegenbeispiel: Nicht jeder Mensch lässt sich eindeutig einem Geschlecht zuschreiben.
- Eine Wahrscheinlichkeit für Ereignis A in % gibt an, in wievielen von 100 Fällen wir erwarten (oder bisher gesehen haben), dass Ereignis A eintritt.
 - Zukünftig: Vorhersage (predictive statistics)
 - Bisher: deskriptive Statistik
- Da keine zwei Ereignisse gleichzeitig eintreffen können und alle möglichen Ereignisse beschrieben wurden, gilt, dass die Summe aller Wahrscheinlichkeiten gleich 1 ist.

Wahrscheinlichkeit unabhängiger Ereignisse



- Wenn Ereignisse unabhängig voneinander eintreten, kann ihre Wahrscheinlichkeit durch Multiplikation der Einzelwahrscheinlichkeiten errechnet werden.
- Beispiel: Die Wahrscheinlichkeit, in einer beliebigen Woche des Jahres Geld zu finden sei 10%.
- Die Wahrscheinlichkeit, in dieser und der nächsten Woche Geld zu finden ist $10/100 * 10/100 * 100\% = 1\%$
- Die Wahrscheinlichkeit, in dieser Woche etwas zu finden und in der nächsten Woche nichts ist $10\% * 90\% = 9\%$.

Frage

Eine Studie besagt, dass 65% aller Kinder keinen Kohl mögen. Wenn 4 zufällig angerufen würden: wie hoch ist die Wahrscheinlichkeit, dass **alle** keinen Kohl mögen?

- 1) 260%
- 2) 17,9%
- 3) 1.5%
- 4) Nichts davon.

Vorsicht: was heißt Unabhängigkeit?



- Tragischer Fall von Sally Clark:
- Verlor ihre zwei Söhne innerhalb weniger Wochen nach der Geburt.
- Wahrscheinlichkeit dafür bei Unabhängigkeit wurde vom Gutachter auf 1:73 Millionen geschätzt.
- Genetische Komponente unterschätzt (kann zweiten Tod wahrscheinlicher machen), Risiken klein geredet, ...
- Andere Gutachter sahen wesentlich kleinere Wahrscheinlichkeiten dafür, dass Clark ihre Kinder getötet haben sollte.
- Ihr Freispruch kam zu spät – sie starb wenig später an einer Alkoholsucht.





Relative Wahrscheinlichkeiten

Vorsicht bei relativen Wahrscheinlichkeiten



- Brustkrebs-Frühvorsorge reduziert die Brustkrebstoten um 20%!
- Hier handelt es sich um die Reduktion des relativen Risikos!
 - Ohne Screening sterben 5 von 10.000 Frauen
 - Mit Screening nur 4 von 10.000
 - Daher handelt es sich um eine Reduktion von 5 auf 4, gleich 80% des Risikos ohne Screening.



Dramatische Entwicklung!



- Lukas Breitenbach von der Berliner Zeitung schreibt am 29.4.2015: „Es ist eine dramatische Entwicklung. Die Zahl der Verkehrstoten in Berlin ist um 40 Prozent angewachsen.“
- Wieder eine relative Betrachtung:
 - Insgesamt 52 Tote im Jahr 2014.
- Wieviele waren es dann 2013?
 - 1) Ca. 30
 - 2) Ca. 73
 - 3) Ca. 29
 - 4) Ca. 37
- Antwort: Wenn 52 Tote das Ergebnis einer Steigerung von 40% des Vorjahres sind, handelt es sich um 140% der Vorjahresanzahl.
- Damit müssen wir $(52/140) * 100$ rechnen.
- Wir kommen auf 37,12 Personen, also ca. 37.

Relative Risiken



- Bauer schreibt: „Jeder Verkehrstote ist einer zu viel. Nichtsdestotrotz müssen die genannten Meldungen aus mehreren Gründen mit Vorsicht interpretiert werden. Zum einen erhält man bei einer sehr kleinen Basis sehr schnell große Wachstumsraten.“
- Er betont auch, dass die absolute Anzahl von Toten immer in Bezug auf das Verkehrsaufkommen (regional oder lokal, z.B. durch Sperrungen und Umleitungen) bewertet werden muss.



Risikoerhöhung

- Babies von übergewichtigen Müttern haben ein 35% erhöhtes Risiko zu sterben, bevor sie 60 sind.
- Ist das eine wichtige Entdeckung?
- Hängt von der Basisrate ab: wenn normalerweise 10 von 10.000 Personen so früh sterben, wären es nun erwartet 13,5.



WE SHOULD GO TO THE NORTH BEACH.
SOMEONE SAID THE SOUTH BEACH HAS
A 20% HIGHER RISK OF SHARK ATTACKS.

YEAH, BUT STATISTICALLY, TAKING
THREE BEACH TRIPS INSTEAD OF TWO
INCREASES OUR ODDS OF GETTING
SHOT BY A SWIMMING DOG CARRYING
A HANDGUN IN ITS MOUTH BY **50%**!

OH NO! THIS IS
OUR THIRD TRIP!



REMINDER: A 50% INCREASE
IN A TINY RISK IS **STILL TINY.**



Bedingte Wahrscheinlichkeiten

Verwechslung vor Gericht



- Ein Dieb war in London gesehen worden, konnte aber nur sehr grob anhand fünf Eigenschaften beschrieben werden.
- Ein Mann mit diesen Eigenschaften wurde zwei Tage später gefasst.
- Die Wahrscheinlichkeit für das gemeinsame Auftreten der Eigenschaften sei 1:500.000.
- Wie hoch ist die Wahrscheinlichkeit, dass jemand der 5 unabhängige Eigenschaften besitzt, die auch ein bisher nicht gefasster Dieb hat, der Dieb ist?
- Basierend auf diesem Wissen:
 - 1) Der Mann ist sehr wahrscheinlich der Dieb.
 - 2) Der Mann ist wahrscheinlich nicht der Dieb.



Antwort



- Der Mann ist eher nicht der Dieb:
 - In London mit geschätzten 10 Millionen Einwohnern gibt es erwartet 20 Männer mit dieser Eigenschaftenkombination.
 - Basierend darauf hat er eine Chance von $1/20$ der Dieb zu sein.
 - Und eine Chance von $19/20$, nicht der Dieb zu sein.



Von: http://www.conceptstew.co.uk/pages/prosecutors_fallacy.html (sehr gut!)

https://en.wikipedia.org/wiki/Prosecutor's_fallacy

Typische Frage nach bedingten Wahrscheinlichkeiten



- $P(A|B)$ beschreibt die Wahrscheinlichkeit von Ereignis A, wenn wir schon wissen, dass Ereignis B eingetreten ist.
- Beispiel: wenn wir irgendeine Person in Deutschland anrufen, ist der Erwartungswert für ihre Körperlänge ca. 1.70 m
- Wenn wir wüssten, dass die Person eine Frau ist, liegt der Erwartungswert bei 1.65 m; wäre es ein Mann liegt der Erwartungswert bei 1.75
- Wie hoch wäre die Wahrscheinlichkeit $P(\text{Abitur} | \text{Hochschulabschluss})$?



Der AIDS Test

Eine Millionenstadt in Deutschland spendet geschlossen Blut – außer denen, die schon wissen, dass sie HIV-infiziert sind.

Natürlich werden alle Blutspenden einem HIV-Test unterzogen.

Welche Wahrscheinlichkeit wirklich infiziert zu sein hat eine Person, deren Test sich als positiv erweist?



<http://commons.wikimedia.org/wiki/File:ELISA.jpg>



Ein paar Informationen über HIV Tests



- Was müssen Sie dafür über HIV-Tests wissen?
- Sehr sorgfältig designte Tests, die mit hoher Wahrscheinlichkeit Infizierte ausfindig machen (**Sensitivität**) und nicht-Infizierte möglichst wenig oft fälschlich identifizieren (**Spezifizität**).
- Wir nehmen vereinfachend an:
 - Sensitivität ist 99,9%, d.h., von 1.000 Infizierten ist der Test erwartungsgemäß bei 999 positiv.
 - Spezifizität sei 99.8%, d.h., von 1.000 nicht-Infizierten ist der Test erwartungsgemäß bei 998 negativ.
 - Im Umkehrschluss heißt dies: einen von 1.000 Infizierten erkennt der Test nicht und pro 1.000 nicht-Infizierten bekommen zwei die beunruhigende Nachricht, dass ihr HIV-Test positiv sei.

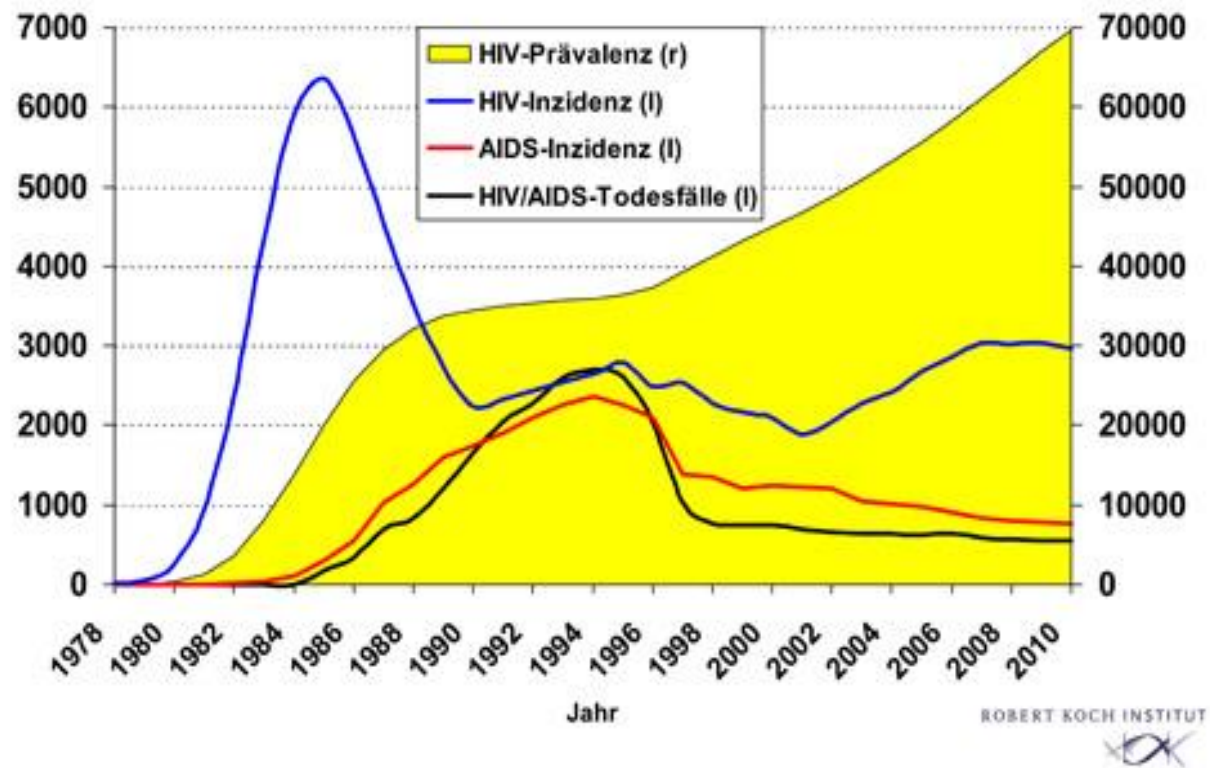
Ein paar Informationen über HIV in Deutschland



- Prävalenz: Wahrscheinlichkeit, dass eine bestimmte Gruppe infiziert ist.
- Bezogen auf HIV ist dieser Wert, für die meisten Bevölkerungsgruppen in Deutschland, sehr, sehr klein.
- Inzidenz: Wahrscheinlichkeit, dass sich eine Person innerhalb eines bestimmten Zeitraumes neu infiziert.
- In Deutschland liegt die Inzidenz bei 3.000 Personen/Jahr.

http://www.rki.de/DE/Content/InfAZ/H/HIVAIDS/hiv_node.html, 23.11.2010 ;
"Zum Welt-AIDS-Tag", Epidemiologisches Bulletin, 22.11.2010, Robert-Koch-Institut

Geschätzte HIV/AIDS Inzidenz, Prävalenz und Todesfälle in Deutschland, Ende 2010 (Modell)



http://www.rki.de/DE/Content/InfAZ/H/HIVAIDS/hiv_node.html, 23.11.2010 ;
 "Zum Welt-AIDS-Tag", Epidemiologisches Bulletin, 22.11.2010, Robert-Koch-Institut



Ist ein positiver HIV-Test ein sicheres Indiz für eine Infektion? Bedingte Wahrscheinlichkeiten

- Wahrscheinlichkeit der Stadtbewohner sich infiziert zu haben wird gleichgesetzt mit der Inzidenz in Deutschland (Modellannahme !)
- $\approx \frac{3000}{80000000} = 0,000375$
- Damit erwarten wir, dass von 1,000,000 Bürgerinnen und Bürgern, 37,5 (wir runden auf 38) infiziert sind.
- Wegen der hohen Sensitivität (Fähigkeit, Infizierte zu entdecken), sind erwartet 38 ihrer HIV-Tests positiv.
- Von den erwartet nicht-Infizierten, bekommen 0.002 von 100, entsprechend 2000 von 1000000 aber auch einen positiven Test.
- Jetzt fragen wir uns: wieviel Prozent der Personen mit positivem Test sind wirklich infiziert?
- Das sind nur 38 von $(2000+38) = 1.9\%$

Big Data und Bauchspeicheldrüsenkrebs



- 2016 zeigten Paparrizos, White und Horvitz, dass Suchanfragen bei Bing ihnen dabei helfen konnten, Leute mit Bauchspeicheldrüsenkrebs ausfindig zu machen.
 - Bauchspeicheldrüsenkrebs ist oft sehr schnell wachsend.
 - Erkrankte bekommen die Diagnose oft sehr spät (Stadium IV), so dass die Fünf-Jahres-Überlebenschance bei nur 5% liegt.
 - Können Suchbegriffe schon früher die Diagnose liefern?
 - Studie zeigt: ca. 5% bis 15% der Suchenden, die später angeben “I have been diagnosed with ...” können identifiziert werden, bei false positive rates von nur 0.00001 bis 0.0001.
- Übersetzen Sie in natürliche Frequenzen.
- Wie hoch ist die Wahrscheinlichkeit für eine Erkrankung, wenn der Algorithmus jemanden als „vermutlich“ krank markiert, wenn ca. 2 von 10.000 pro Jahr neu erkranken?

Antwort



- Wir nehmen 1.000.000 Bewohner (die die Suchmaschine Bing nutzen und nehmen an, dass diese dasselbe Risiko wie alle Amerikaner haben. Das könnte falsch sein. MODELLANNAHME!)
- Dann würden davon erwartet 200 in diesem Jahr neu erkranken.
- Davon könnte der Algorithmus von Bing 5% bis 15% erkennen, also zwischen 10 und 30.
- False positive rate war 0.00001 bis 0.0001, also zwischen „Einer von 100.000“ bis „Einer von 10.000“.
- Gleichzeitig würde er aber auch von den 999.800, die nicht erkranken werden, zwischen ca. 10 und ca. 100 fälschlicherweise als „vermutlich krank“ markieren.
- Damit liegt die Wahrscheinlichkeit, erkrankt zu sein, wenn man als solcher markiert wurde, zwischen 30/40 (75%) und 10/110 (9%).

Siehe auch: Gerd Gigerenzer:

Microsofts Suchmaschine erhöht die Überlebensrate bei Bauchspeicheldrüsenkrebs, Unstatistik vom 30.6.2016, <http://www.rwi-essen.de/unstatistik/56/>

DNA-Matching



- Gleicher Fall: DNA Match zeigt an, dass es sehr unwahrscheinlich ist, einen weiteren Menschen mit diesen Gen-Markern zu finden.
- Hängt aber wieder davon ab, wer grundsätzlich in Frage kommt.
- Bei einer Frau mit nur wenigen Geschlechtspartnern ist die Wahrscheinlichkeit gering, dass ein zweiter Mann dieselben Marker hat.
- Bei einem Bankraub, bei dem man nichts über die Täter weiß und zufällig einem Menschen mit denselben DNA-Markern begegnet, könnte die Menge der potentiellen Täter gleich der Hälfte der Menschheit sein und da gibt es dann wieder ein paar Dutzend Männer mit denselben genetischen Markern.

Oder ganz anders: das Heilbronner Phantom



- Nach dem Polizistenmord in Heilbronn wurde eine Frau gesucht.
- Ihre DNA konnte in 40 weiteren Kriminalfällen zugeordnet werden.
- Frau konnte nicht gefunden werden.
- Stellte sich als diagnostisches Artefakt heraus, eine Verpackungsmitarbeiterin des DNA-Abstrichmaterials war die Quelle.
- Wo gearbeitet wird, passieren Fehler!



Arbeitsblatt:



- Was noch so schief gehen kann:
 - Arbeitsblatt S. 7
 - Arbeitsblatt S. 9



5 Minuten PAUSE



Korrelationen und Ursachenbeziehungen



- Wenn ein Ereignis A ein anderes Ereignis B wahrscheinlicher macht, dann wird der Anstieg der Anzahl von Ereignis A einen Anstieg der Anzahl von Ereignis B bedingen.
- Diese Ereignisse sind also miteinander korreliert.
- Wenn B aber auch von anderen Ereignissen beeinflusst wird, kann die ursächliche Beziehung u.U. nicht entdeckt werden
 - Nicht jede ursächliche Beziehung kann in einer Statistik entdeckt werden.
 - Insbesondere bei zeitlichen Verzögerungen ist das oft der Fall.
- Es gibt verschiedene Gründe, warum die Anzahl zweier Ereignisse korrelieren kann, ohne dass eine ursächliche Beziehungen zwischen diesen Ereignissen besteht.



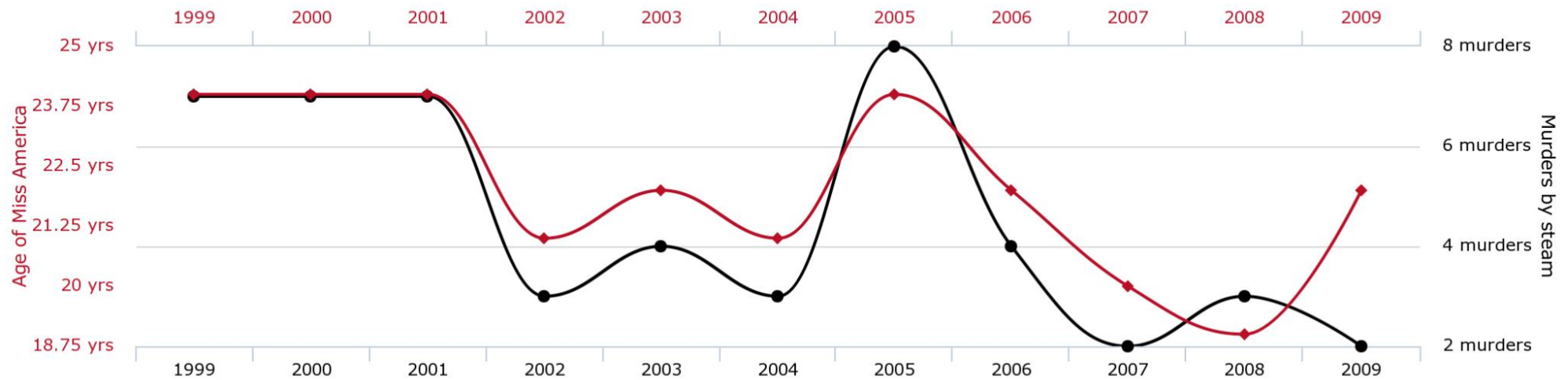
If the fossil record shows
more dinosaur footprints
in one period than another,
it does not necessarily mean
that there were more dinosaurs

—
it may be
that there was more mud.

Elise E. Morse-Gagné in her letter to
Science with the title „Culturomics –
Statistical Traps Muddy the Data“,
Science, 332, p. 35



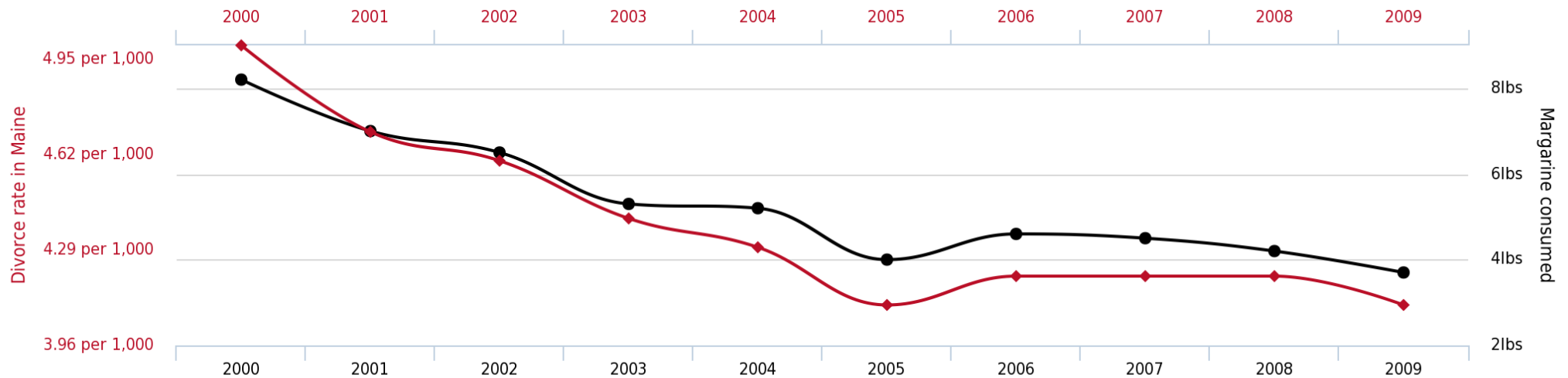
Tyler Vigen: Spurious Correlations



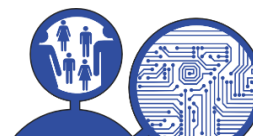
tylervigen.com

Donated to the public under a CC-licence by Tyler Vigen

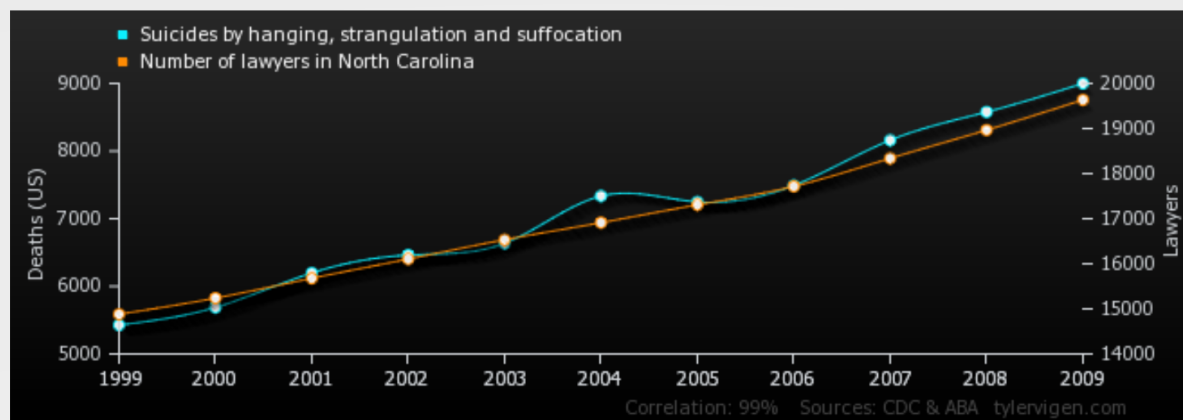
Tyler Vigen: Spurious Correlations



Donated to the public under a CC-licence by Tyler Vigen



Suicides by hanging, strangulation and suffocation correlates with Number of lawyers in North Carolina



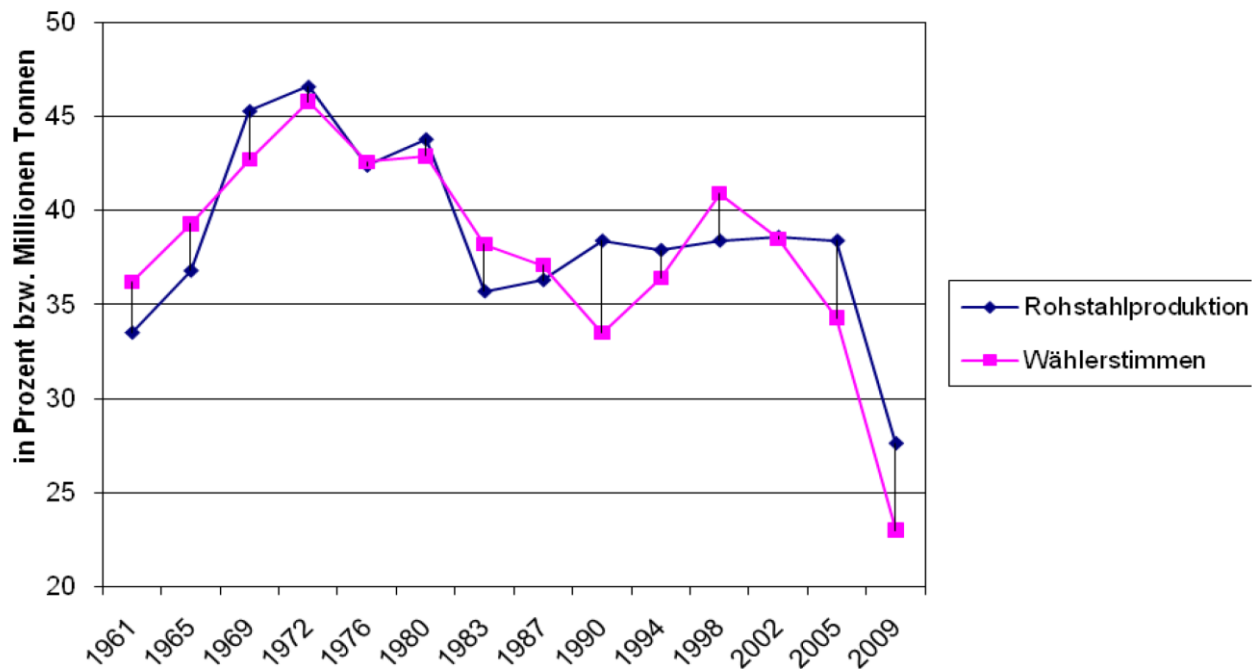
[Upload this chart to imgur](#)

	<u>1999</u>	<u>2000</u>	<u>2001</u>	<u>2002</u>	<u>2003</u>	<u>2004</u>	<u>2005</u>	<u>2006</u>	<u>2007</u>	<u>2008</u>	<u>2009</u>
<i>Suicides by hanging, strangulation and suffocation</i> <i>Deaths (US) (CDC)</i>	5,427	5,688	6,198	6,462	6,635	7,336	7,248	7,491	8,161	8,578	9,000
<i>Number of lawyers in North Carolina</i> <i>Lawyers (ABA)</i>	14,881	15,239	15,678	16,105	16,534	16,912	17,308	17,717	18,339	18,966	19,637

Correlation: 0.993796

[Permalink](#) - [Mark as interesting \(1,874\)](#) - [Not interesting \(1,151\)](#)

Das geht auch in Deutschland: Mierscheid-Gesetz

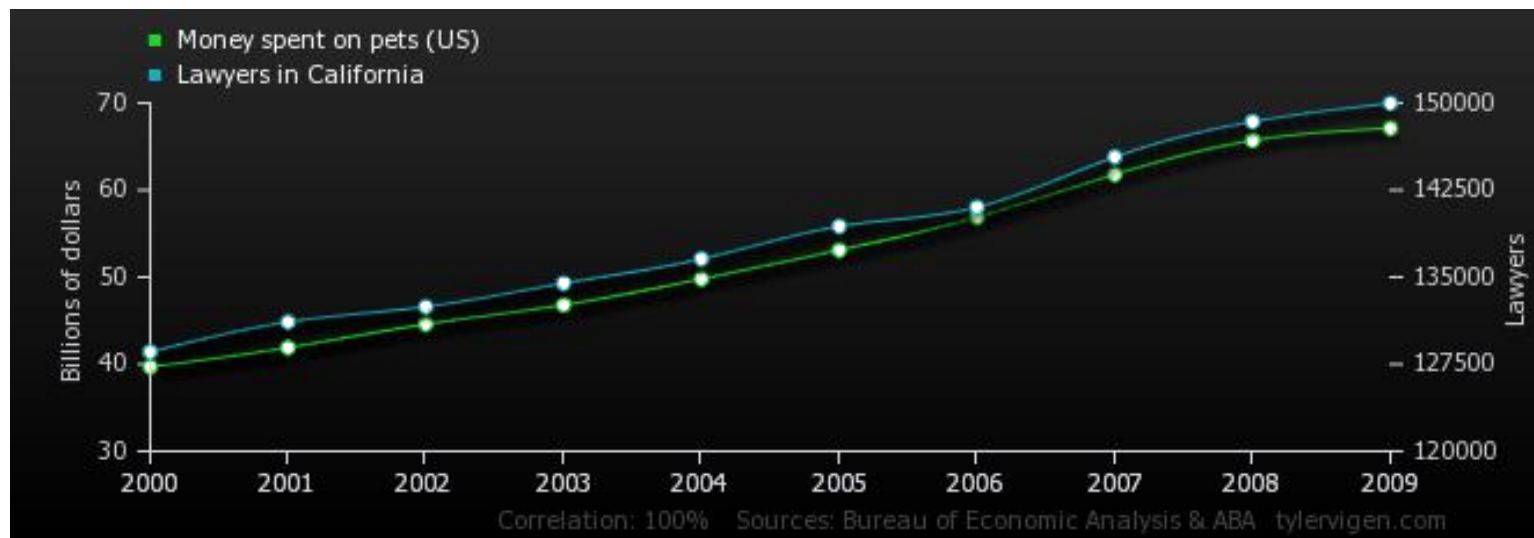


Aus einem Brief des fiktiven Bundestagsabgeordneten Jakob Maria Mierscheid vom 22.2.2010,
<http://www.wahlrecht.de/doku/doku/20100222.pdf>

Finden Sie Ihre eigene „spurious correlation“



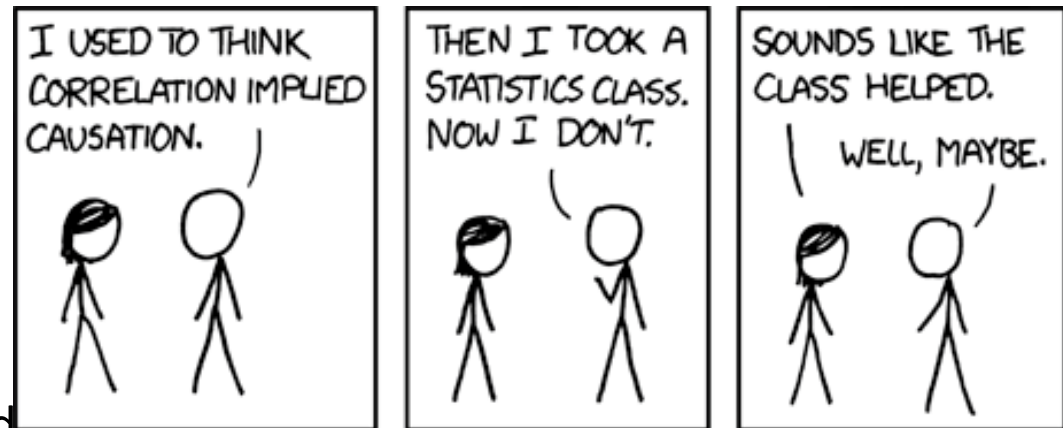
- Auf Vigen's Seite können Sie auch selber Korrelationen finden:
- http://tylervigen.com/discover?type_select=fun
- Wählen Sie eine erste Variable, z.B. "Ausgaben für Haustiere"
- Dann schlägt die Seite alle Variablen vor, die damit korrelieren.
- Hier mein Ergebnis.



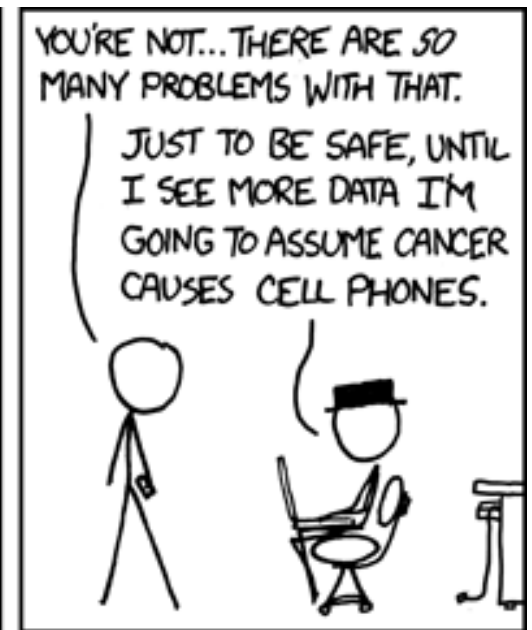
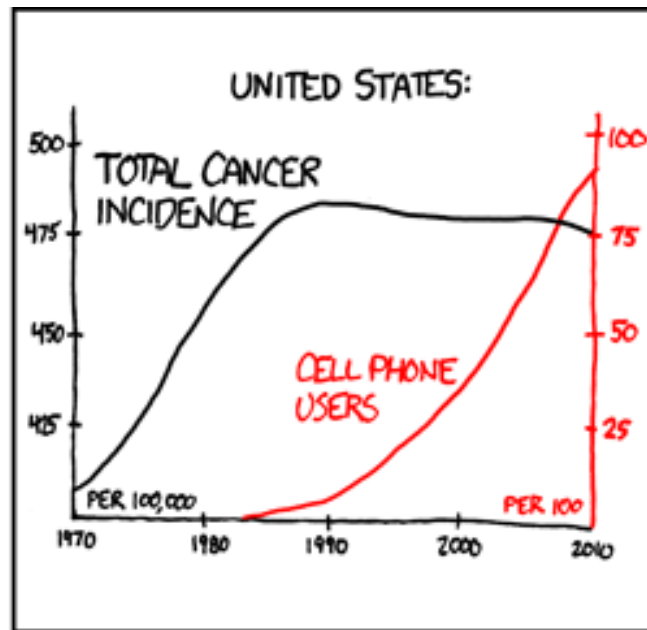
Kategorien von Korrelation



- Es gibt eine ursächliche Beziehung
 - A bedingt B
 - B bedingt A
- Es gibt eine dritte Variable, die A und B bedingt:
 - X bedingt A und X bedingt B oder
 - X bedingt Y, das A bedingt, und X bedingt Y', das B bedingt (etc.)
- Weder bedingt A Ereignis B noch B Ereignis A.



<https://xkcd.com/552/>
Thanks to Randall Munroe
for the right to use his
cartoons in lectures



Sind Vegetarier die besseren Menschen?



- Die Nordwest-Zeitung schreibt am 25.7.2015: „Vegetarier und Veganer haben weniger Vorurteile und widersetzen sich eher autoritären Strukturen. Dazu kommt zumindest eine am Freitag veröffentlichte Studie von Medizinerinnen und Psychologen der Universitäten Mainz und Wuppertal. Bei der Studie mit 1400 Personen sei ein eindeutiger Zusammenhang zwischen Ernährung und einem Hang zur Vorverurteilung von Minderheiten und „sozialer Dominanzorientierung“ aufgefallen.“ [1]
- Krämer merkt an, dass auch eine längere Lebenszeit bei Vegetariern nicht durch die vegetarische Lebensweise, sondern durch den Verzicht auf Zigaretten und mehr Sport und ähnliches erklärt werden könnte².
- Er erwartet, dass „unkonventionelle und antiautoritäre Charaktere auch bei der Ernährung zu eher unkonventionellen Methoden neigen.“



http://www.nwzonline.de/panorama/studie-vegetarier-vorurteilsfreier_a_30,0,1372869499.html

Walter Krämer: Macht uns eine vegetarische Lebensweise zu besseren Menschen?, Unstatistik vom 25.7.2015, <http://www.rwi-essen.de/unstatistik/45/>

Klicker



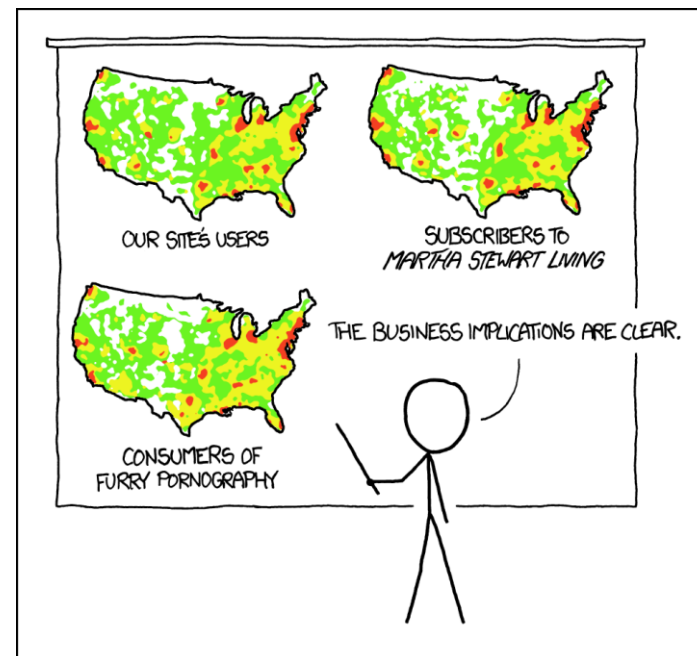
- Zu welcher Kategorie gehört diese Korrelation damit wahrscheinlich?
- Ereignis A: vegetarische Lebensweise
- Ereignis B: weniger Vorurteile, antiautoritärer?
 - 1) A bedingt B
 - 2) B bedingt A
 - 3) Weder noch
 - 4) Direkter, konfundierender Faktor
 - 5) Indirekter, konfundierender Faktor



Logikfehler: Cum hoc ergo propter hoc



- „Mit diesem, daher durch diesen“.
- Falsche Ableitung einer Kausalität aufgrund einer Korrelation.
- Sehr oft konfundierende Variablen
 - Also: X bedingt A und B direkt oder indirekt



PET PEEVE #208:
GEOGRAPHIC PROFILE MAPS WHICH ARE
BASICALLY JUST POPULATION MAPS

<https://xkcd.com/1138/>

Thanks to Randall Munroe for the right to use his cartoons in lectures

Beispiel: Störche und Babies



- Es gibt verschiedene Studien, die zeigen, dass die Anzahl von Störchen und Babies korreliert.
- Yule zeigte an diesem Beispiel, was eine konfundierende Variable ist:
 - Anzahl der Häuser in einem Dorf korreliert mit Babies und Möglichkeiten für Störche zu nisten¹.



¹ <https://www.psychologytoday.com/blog/how-we-do-it/201307/the-stork-and-baby-trap>

Babies und Muttermilch und das Gehirn



- Psychology Today nennt noch ein weiteres Beispiel:
 - Stillen wurde korreliert mit höheren kognitiven Fähigkeiten der Babies.
 - Mögliche konfundierende Variable: Mütter mit höherer Bildung tendieren eher dazu zu stillen.
 - Weitere Studien, die diese Variable „kontrollierten“ zeigten, dass ein Effekt bleibt.



Macht Googeln denkfaul?



- Aus einer Unstatistik-Kolumne von Prof. Dr. Gerd Gigerenzer¹:
- In einer Studie zeigen die Personen eine schlechtere Denkleistung (gemessen im Lösen von Rätseln), die längere Zeit auf Suchmaschinen verbringen.
- Gesundheitsstadt-berlin.de titelt: „Googeln mit dem Smartphone macht denkfaul“².
- Es gibt keinen Unterschied zwischen Nicht-Nutzern von Smartphones und Wenig-Nutzern von Suchmaschinen.
- Daher kann eine kausale Beziehung nahezu ausgeschlossen werden.



1 Prof. Dr. Gerd Gigerenzer: „Googeln mit Smartphone macht denkfaul!“ Unstatistik vom 30.03.2015, <http://www.rwi-essen.de/unstatistik/41/>

2 <http://www.gesundheitsstadt-berlin.de/googeln-mit-dem-smartphone-macht-denkfaul-5854/>



“Politicians use statistics in the same way that a drunk uses lamp-posts— for support rather than illumination.”

Attributed to Andrew Lang from a 1910 Speech, quoted in Alan L. Mackay *The Harvest of a Quiet Eye* (1977), as reported in *Chambers Dictionary of Quotations* (2005), p. 488. (Quote from Wikiquote, https://en.wikiquote.org/wiki/Andrew_Lang)

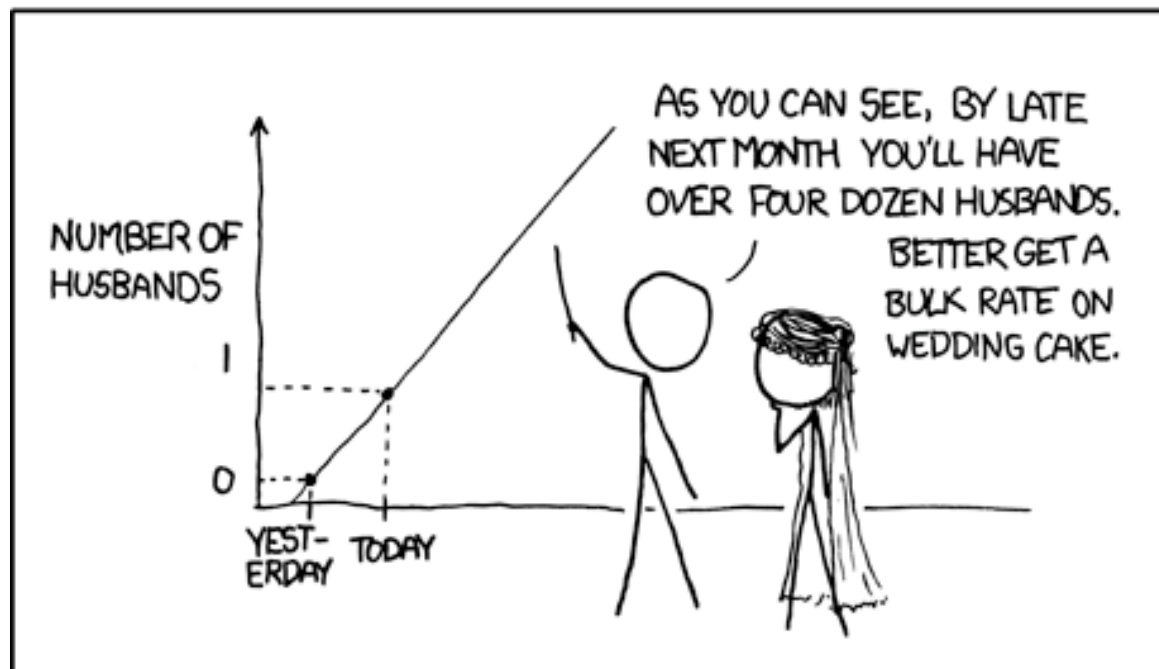


Vorhersagen

Extrapolation: almost always wrong



MY HOBBY: EXTRAPOLATING





Prognose und Wirklichkeit

Entwicklung der Endenergiebereitstellung aus Erneuerbaren Energien in Deutschland

Terawattstunden



Quelle: BMU / AGEE-Stat; Stand: 4/2009

www.unendlich-viel-energie.de





Prediction is very difficult, especially if it's about the future.

Assigned to too many people, among them Niels Bohr and Yogi Barra which results in the quote: „Correct attribution is hard, especially for the past.“ (Doug Arnold)



Philosophy of Science

How does knowledge come into being?

Francis Bacon (1561-1626)



Father of *empirism*

Physical laws need to be induced
from facts

Inductive theories cannot be
proven by more observations but
need to result in falsifiable
predictions.



<http://en.wikipedia.org/wiki/File:Francis-Bacon.jpg>

Basic Requirements

Observations
need to be repeatable.

The result must not be part of the
hypothesis.



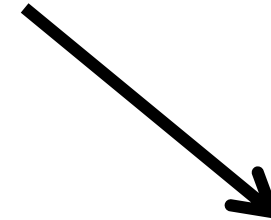
The scientific method



State the problem



Create hypothesis



Induce prediction from hypothesis

Prediction wrong



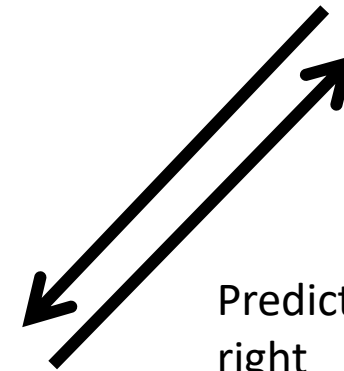
Affirming the consequent

~~**Hypothesis right?**~~



Test prediction

Prediction right



Prediction right

Swans



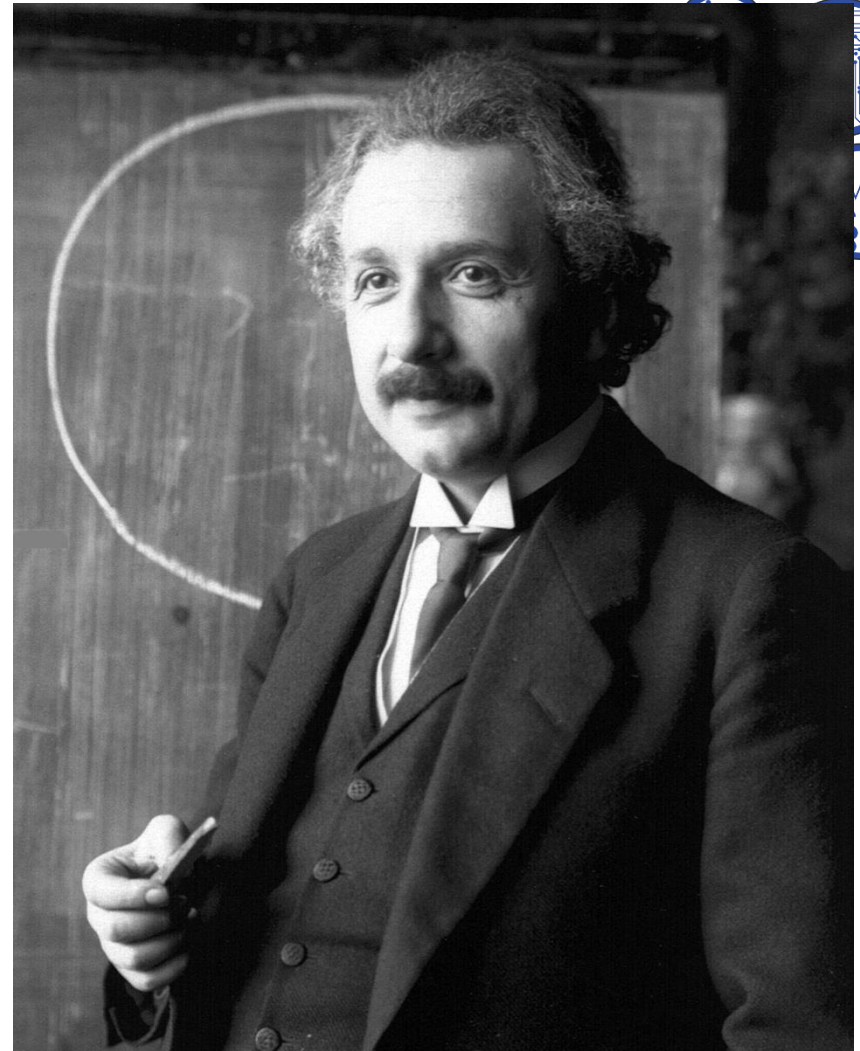
You have seen
1,000 white swans

You hypothesize:
All swans are white

How can you prove this
hypothesis?



- “No amount of experimentation can ever prove me right; a single experiment can prove me wrong”



As noted by A. Calaprice [editor of the *New Quotable Einstein*, University Press and Hebrew University of Jerusalem, 2005, p. 291] paraphrased from a translation of Einstein's *Induction and Deduction*, Document 28 in *Collected Papers of Albert Einstein, Volume 7, Editors M. Janssen et al., Princeton University Press, Princeton, 2002.*

What is wrong with this picture?





Animated sequence of a race horse galloping. Photos taken by [Eadweard Muybridge](#) (died 1904), first published in 1887 at Philadelphia (*Animal Locomotion*).

http://en.wikipedia.org/wiki/File:Muybridge_race_horse_animated.gif

We „see“ what we „know“



Confirmation bias



Self-fulfilling prophecy by observing what is already hypothesized.

Oedipus as a child with his stepfather



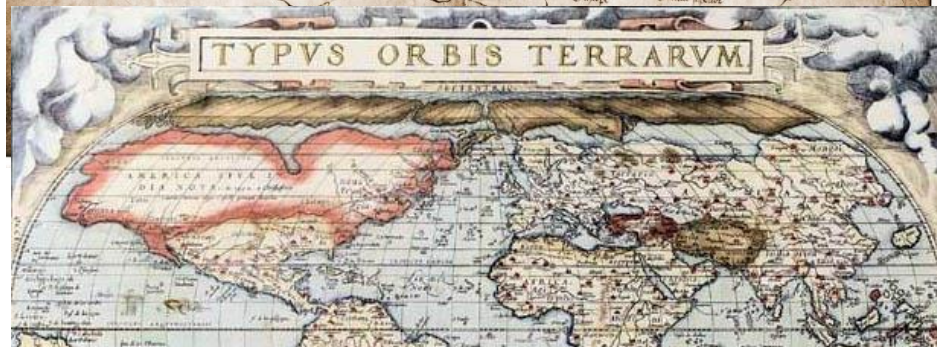
Model:

Simplified description of reality with only a few parameters.

Theory:

Theory combines the model and the assumption that the model is 'good enough'

Theory is valid while there is no contradicting observation.



Ockham's Razor



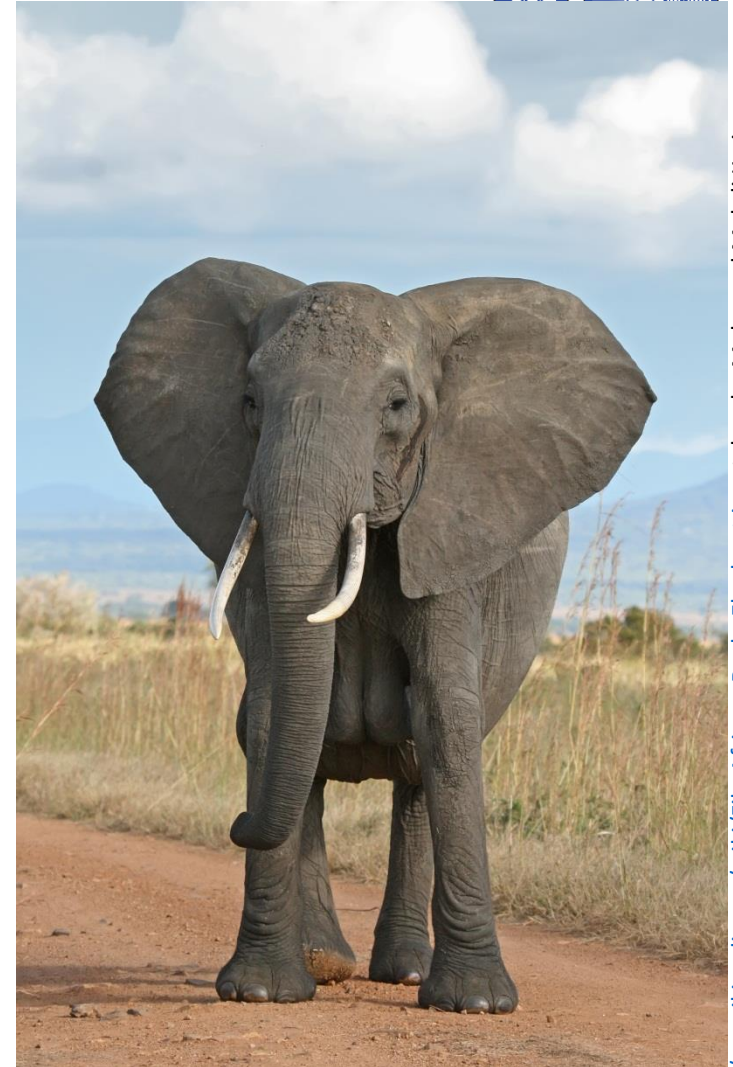
All other things being equal,
the simplest hypothesis
is the best.





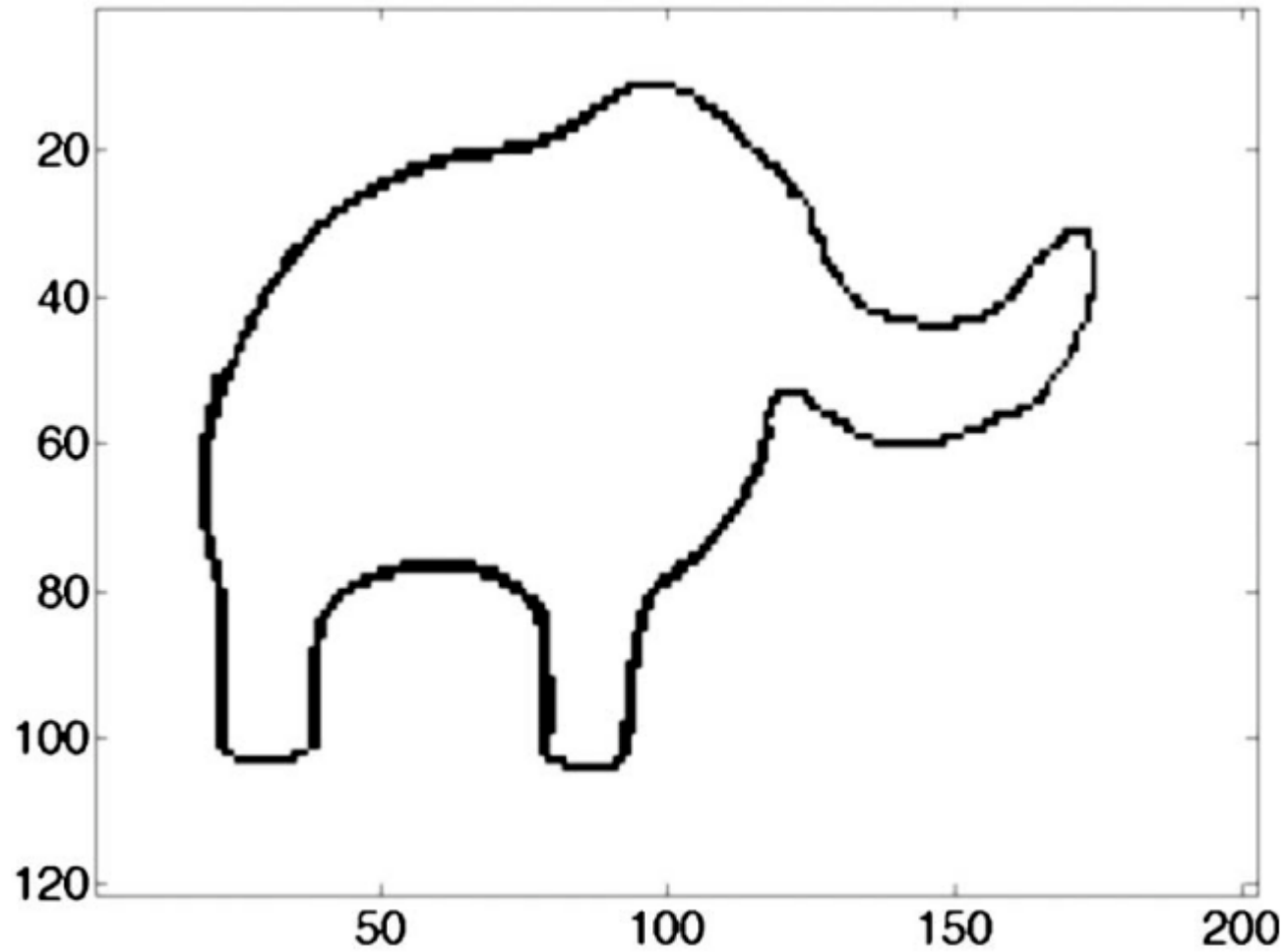
“With four parameters
I can fit an elephant,
and with five I can make him
wiggle his trunk.”

Attributed to von Neumann by Enrico Fermi,
as quoted by Freeman Dyson in
"A meeting with Enrico Fermi"
in Nature 427 (22 January 2004) p. 297

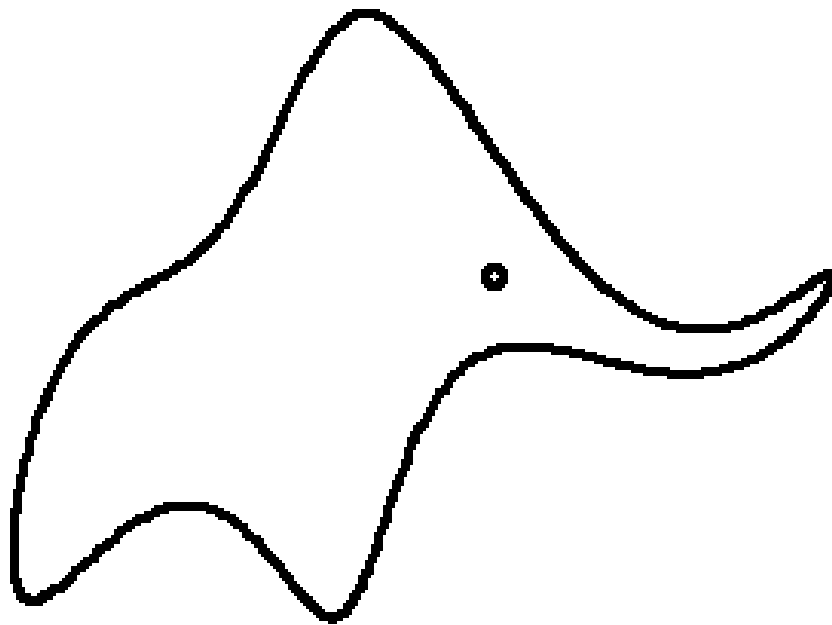


Realized (!!)

by Mayer et al., 2010



The Wiggling

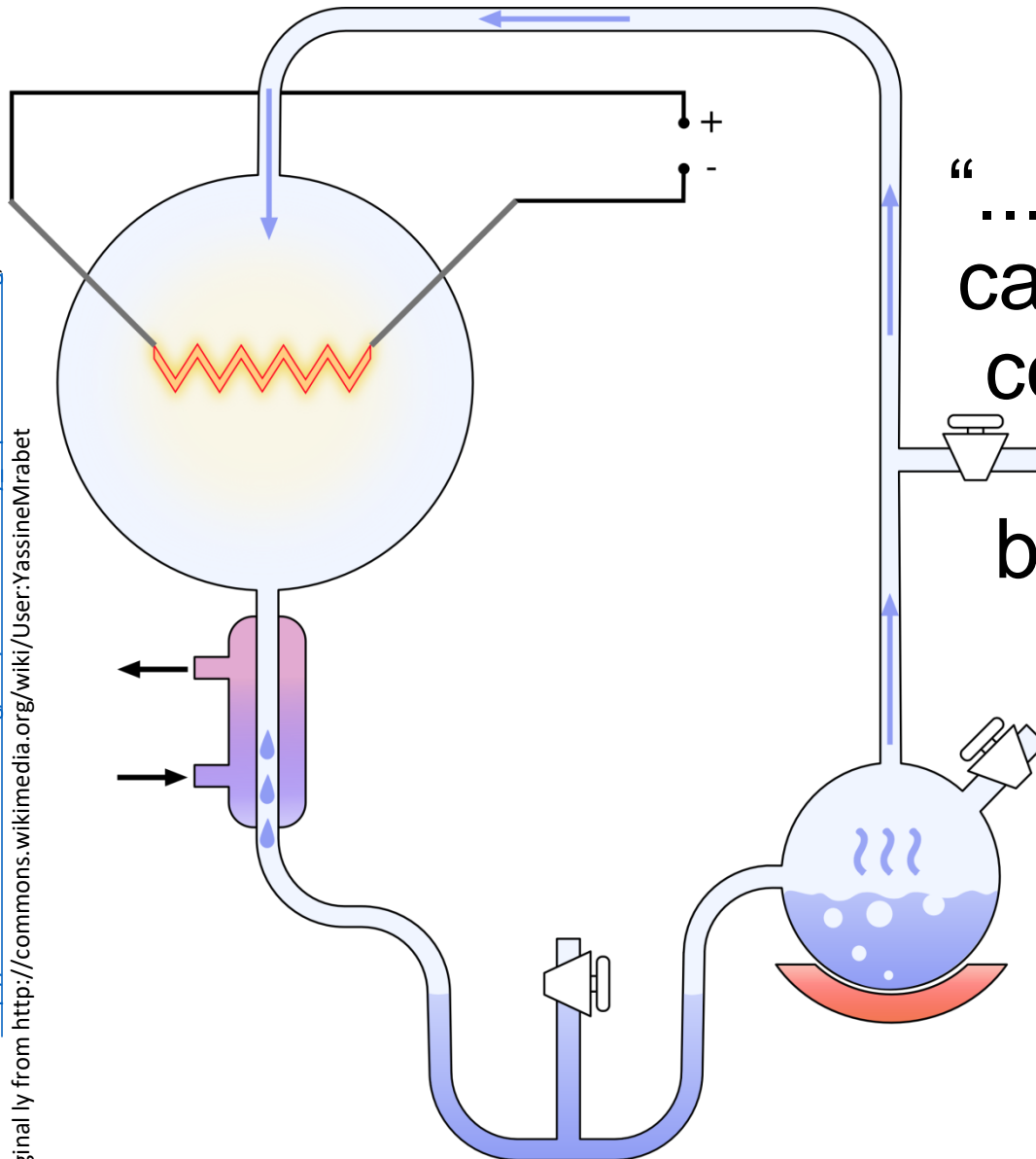


Falsification or Verification?



A scientific theory must be falsifiable.

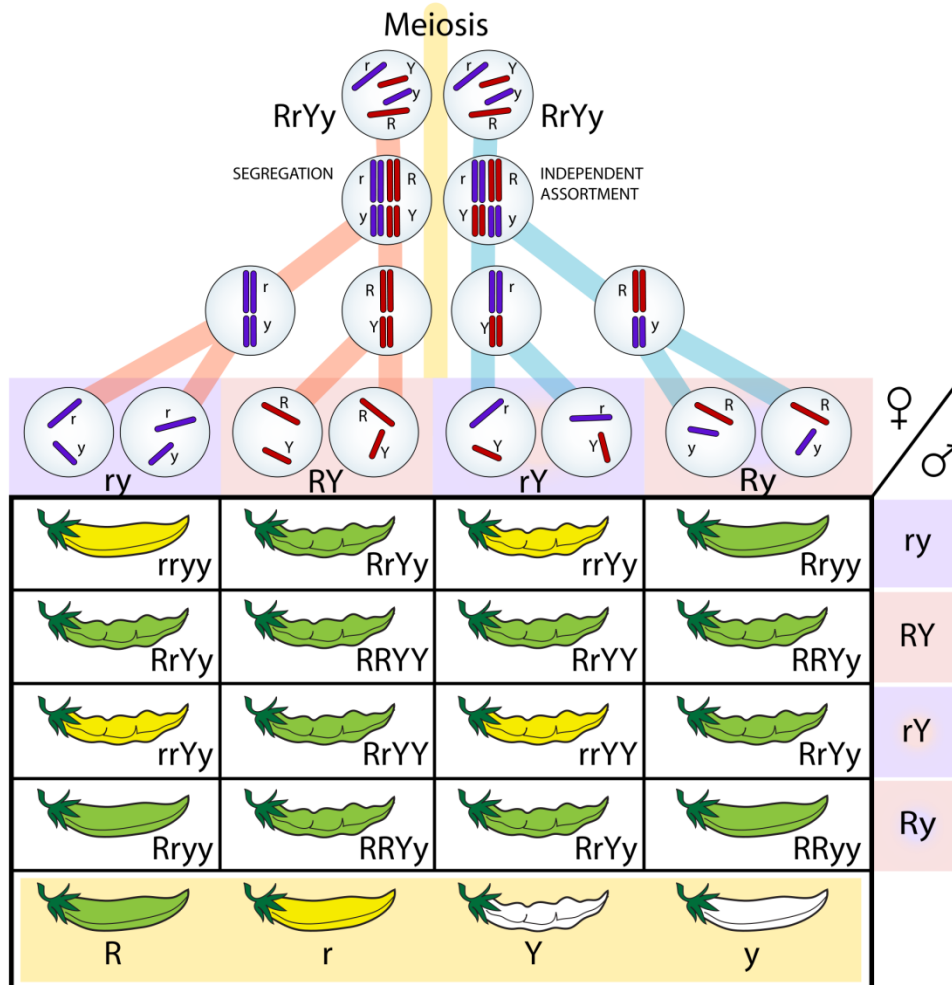




“...research, no matter how carefully and painstakingly conceived and executed, approaches proof, but never fully attains it.”

Douglas Futuyma,
Evolutionary Biology, 3rd ed.,
Sinauer Associates,
1998, pp. 9-12

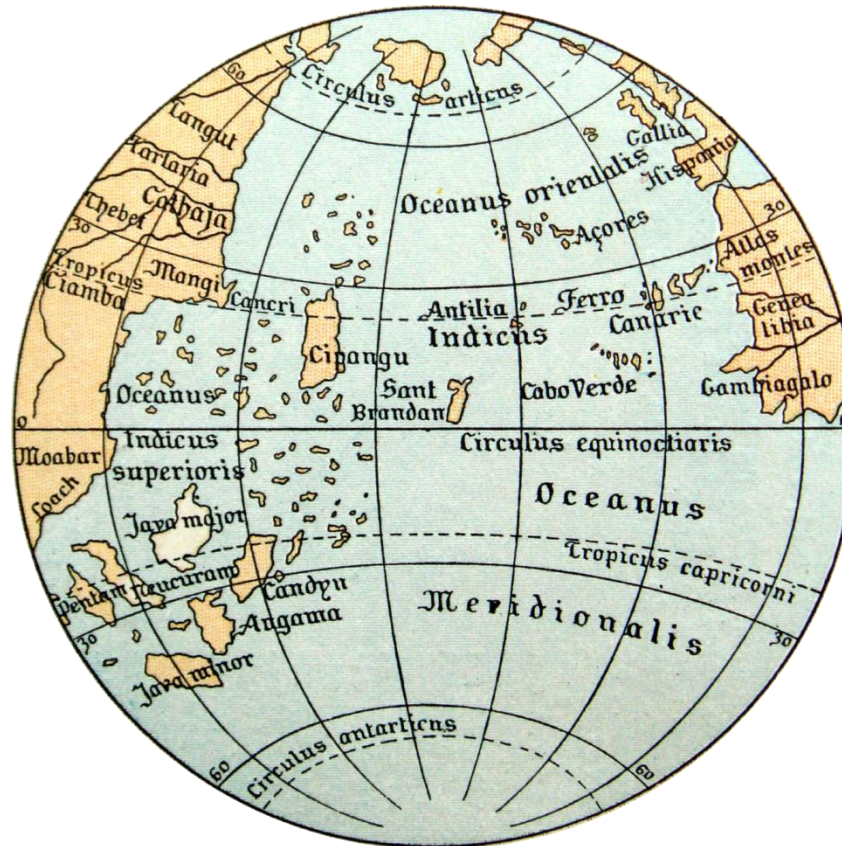
Verifiability



A fact is a hypothesis that is so firmly supported by evidence that we assume it is true, and act as if it were true.

Douglas Futuyma,
Evolutionary Biology, 3rd ed.,
Sinauer Associates, 1998, pp. 9-12

Essentially, all models are wrong,



but some are useful.

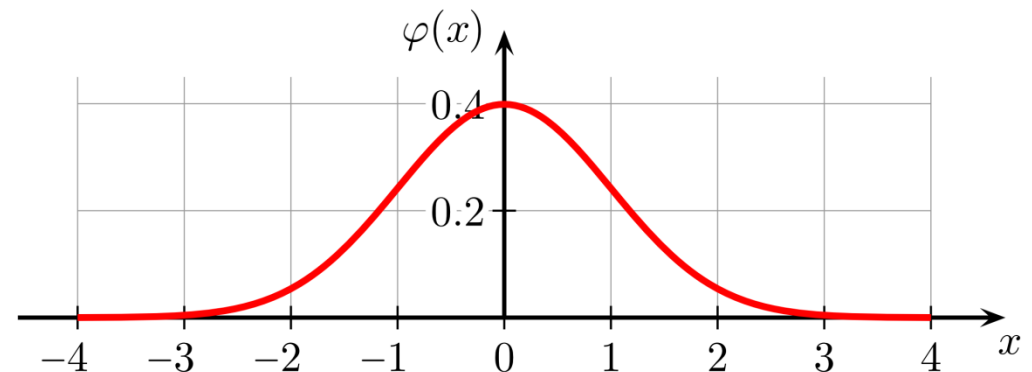
Attributed to statistician George E. P. Box, <http://en.wikiquote.org/wiki/Statistics>, 2.11.2010

Statistical models



Statistical models
make assumptions
about genesis and structure
of the data to be analyzed.

[SachsHedderich2006]





“I gather, young man, that you wish to be a Member of Parliament. The first lesson that you must learn is that, when I call for statistics about the rate of infant mortality, what I want is proof that fewer babies died when I was Prime Minister than when anyone else was Prime Minister. That is a political statistic.”

W. Churchill according to <http://en.wikiquote.org/wiki/Churchill>, 2.11.2010



Although we often hear that data speaks for themselves, their voices can be soft and sly.

Frederick Mosteller, Stephen E. Fienberg, and Robert E.K. Rourke, *Beginning Statistics with Data Analysis* (Reading, Massachusetts, 1983), p. 294; Cited by Edward R. Tufte in „Visual Explanations – Images and Quantities, Evidence and Narrative“, Graphics Press, Cheshire, Connecticut, USA, 1997, p. 26



Statistical Significance



Exercise:

Throw the dice 100 times and note down how often you saw exactly X 6's.

1		
2		
3		
4		
5		
6		



What is our model?



Assumption:
No die with more than one 6

$$P = \frac{\#6s}{\#die}$$

determines probability to throw X 6's.

Theory vs. Empiricism



Even if P is the same for two sets of die, the *empirical results* will look slightly different.



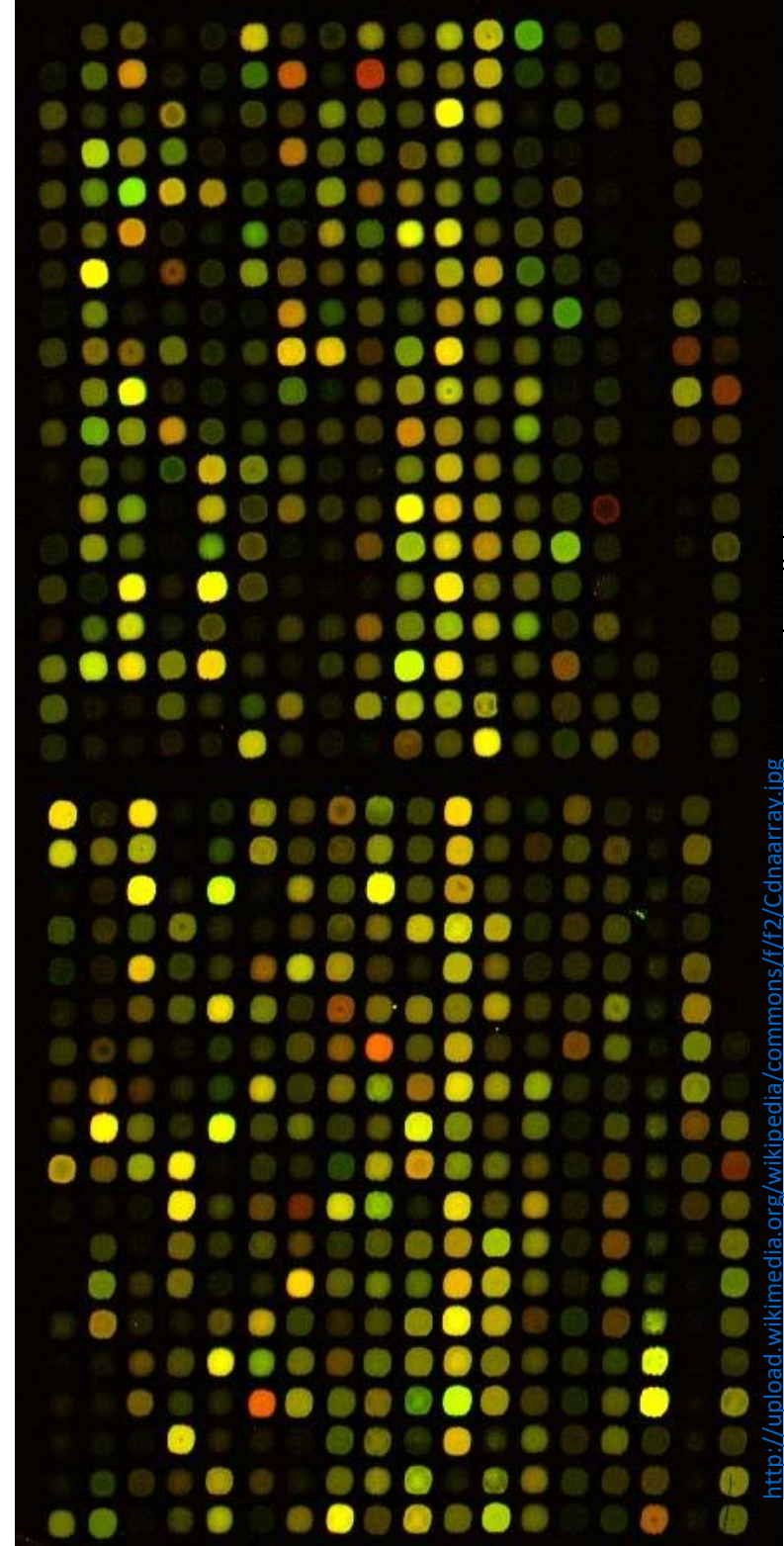
Tests of statistical significance

Expensive Observations

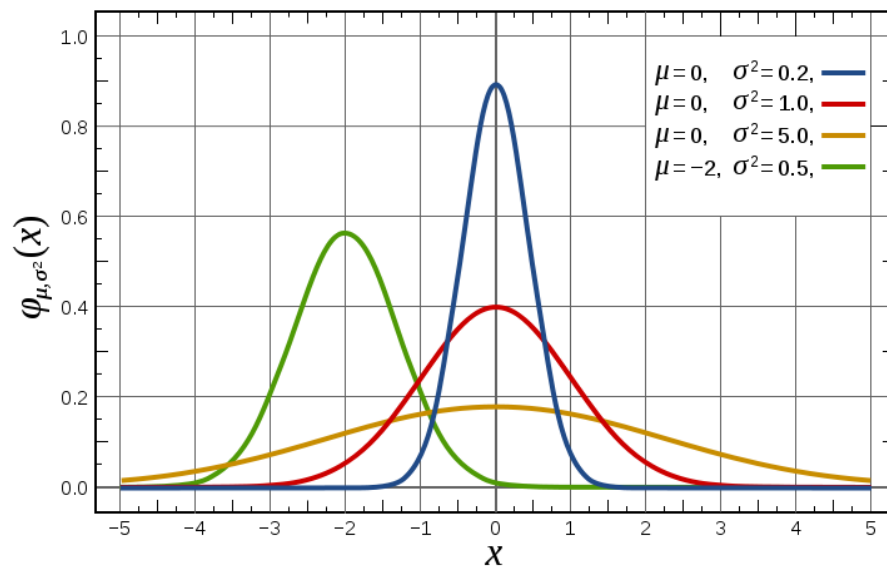
In biology, experiments are expensive:

Only a few samples for each drug/experimental setting but often many different drugs/settings.

Test: A distribution and the question whether one of the values is likely to be from this distribution.



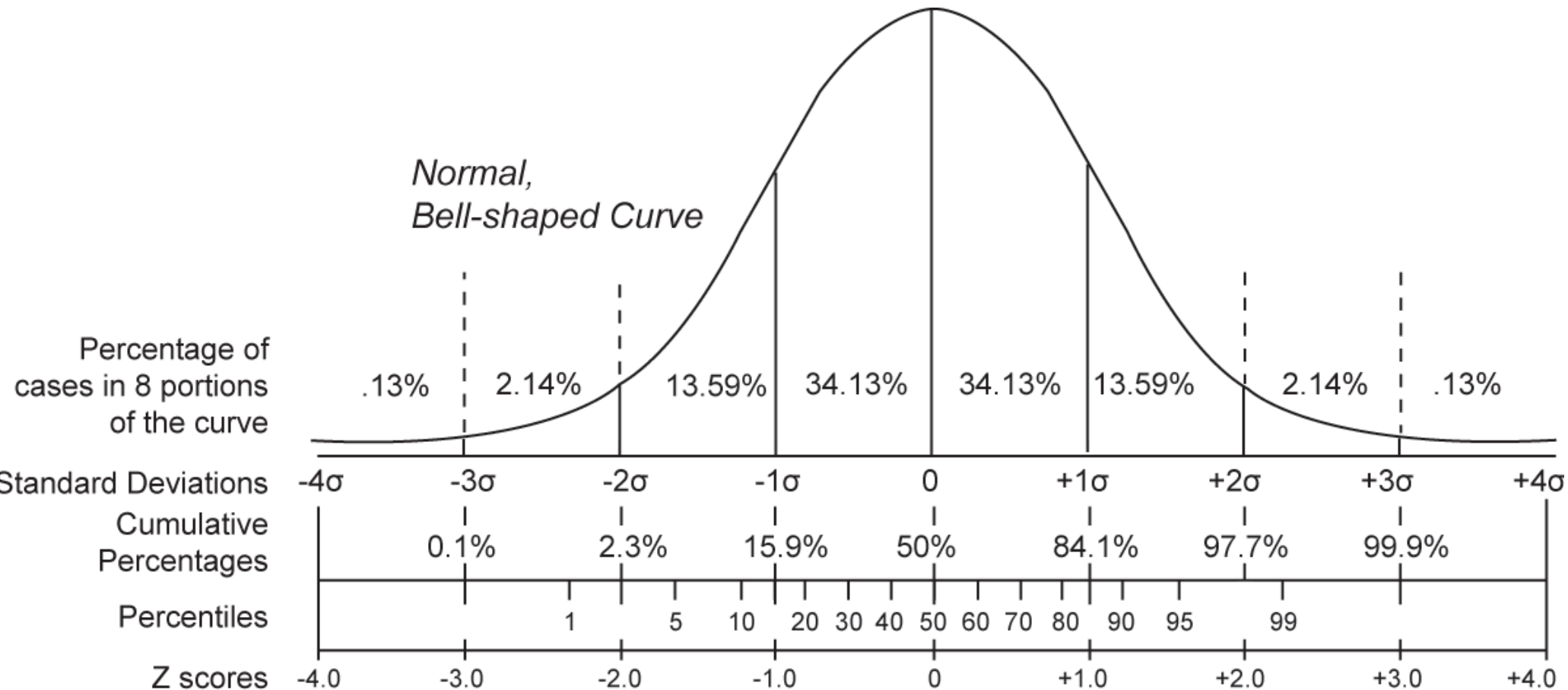
Model: normal distribution



- Typically, it is assumed that all observations are drawn from a normal distribution.
- Here we assume, we could only measure a single case and had a case against which this single value is to be tested.
- Null-Hypothesis: all values come from the **same** normal distribution.

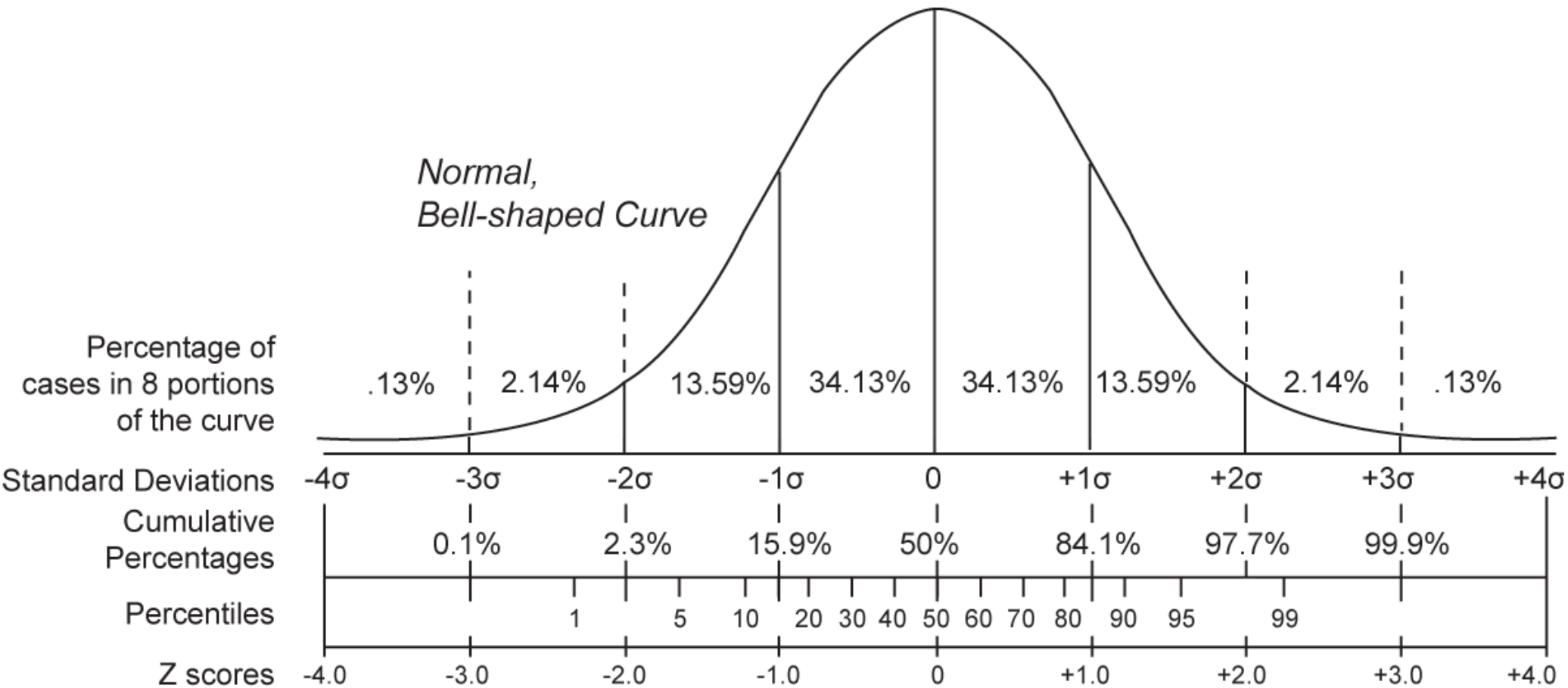
Z-Score and p-Value

Null-hypothesis: observed value is in the same distribution as the controls.



Z-Score and p-Value

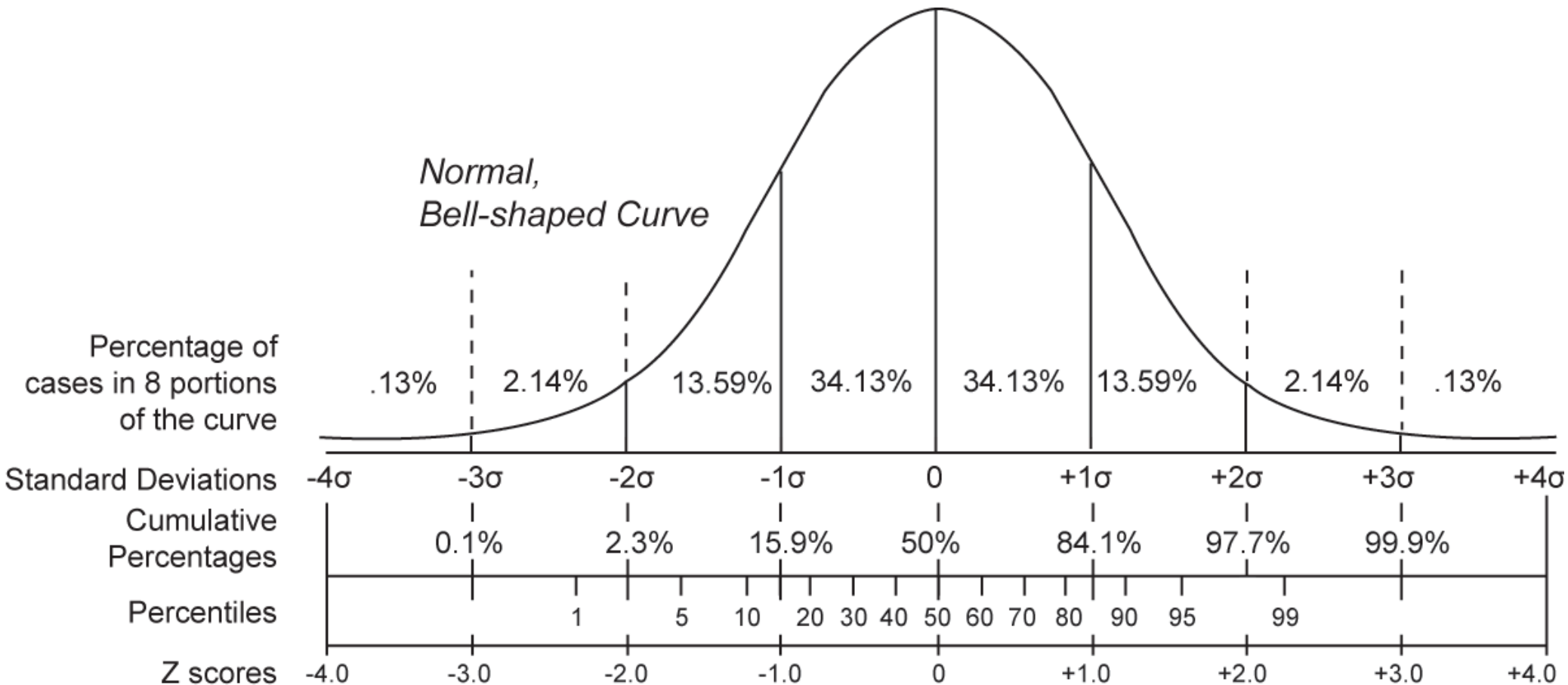
Z-score is a measure of how far away an observed value is from the mean.



Z-Score and p-Value

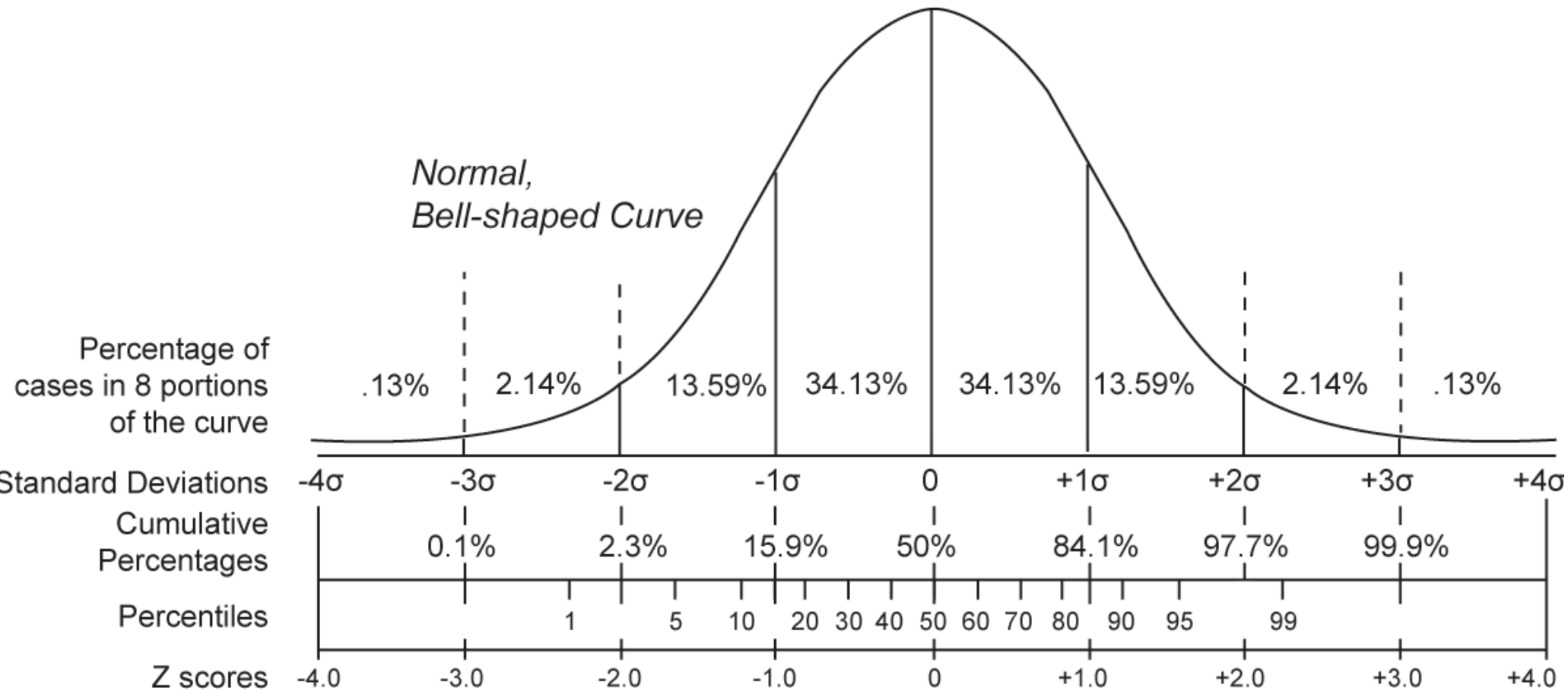


The z-score can be transformed into the probability to see a value as extreme as this in a normal distribution (p-value).



Z-Score and p-Value

If smaller than some predefined value α , the null-hypothesis is rejected.



Statistical Significance



Difference between observed values are said to be significant if the probability that the difference is pure chance is small.

$$\alpha = 0.05$$

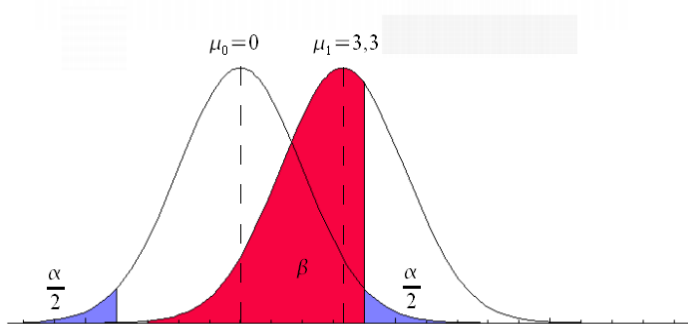
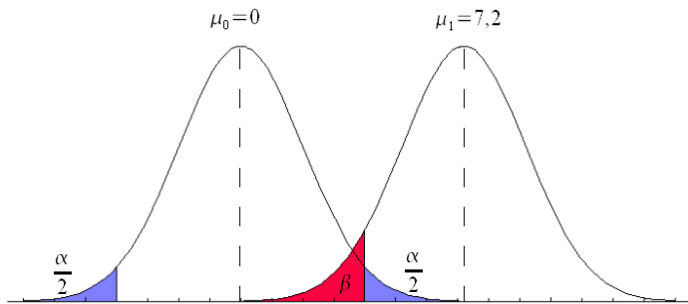
If the difference is significant we conclude that the values come from different distributions.

or

What is a low probability?

$$\alpha = 0.01$$

Cultural convention



**Null hypothesis
(H_0) is true**

**Null hypothesis
(H_0) is false**

Reject null hypothesis

Type I error (α)
False positive

Correct outcome
True Positive

Fail to reject null
hypothesis

Correct outcome
True Negative

Type II error (β)
False negative

Type II errors



The absence of evidence, however,
does not imply the evidence of absence.

Modified from E.R. Tufte, „Visual explanations – Images and Quantities, Evidence and Narrative“, Graphicis Press, Cheshire, Connecticut, USA, 1997, p. 28: „This absence of evidence, however, was not evidence of absence“, regarding the finding that the water of a certain pump well suspected to be the source of an epidemic seemed to be pure.

I CAN'T BELIEVE SCHOOLS
ARE STILL TEACHING KIDS
ABOUT THE NULL HYPOTHESIS.

I REMEMBER READING A BIG
STUDY THAT CONCLUSIVELY
DISPROVED IT *YEARS* AGO.



Relevance



Drug testing:

Type I error: a new drug is approved although it is not better than the old one (but more expensive → higher total costs)

Type II error: the new drug is better but it is not approved → higher costs through longer sickness periods



A valuable lesson (by G.E. Dallal)



- <http://www.jerrydallal.com/LHSP/multtest.htm>
- <http://www.jerrydallal.com/LHSP/jellybean.htm>
- <http://www.jerrydallal.com/LHSP/cellphone.htm>
- <http://www.jerrydallal.com/LHSP/coffee.htm>



What is significant?



<u>P-VALUE</u>	<u>INTERPRETATION</u>
0.001	HIGHLY SIGNIFICANT
0.01	
0.02	
0.03	
0.04	SIGNIFICANT
0.049	
0.050	OH CRAP. REDO CALCULATIONS.
0.051	ON THE EDGE OF SIGNIFICANCE
0.06	
0.07	HIGHLY SUGGESTIVE, SIGNIFICANT AT THE P<0.10 LEVEL
0.08	
0.09	
0.099	HEY, LOOK AT THIS INTERESTING SUBGROUP ANALYSIS
≥ 0.1	

Mean



„How many people do you know in prison?“

The average was 1.0

Isn't that quite high?

The solution



Here, we rather see two totally different populations.

According to the data, 70% of the respondents reported knowing 0 people in prison. However, the responses show a wide range of variation, with almost 3% reporting that they know at least 10 prisoners. [Zheng2006]



The Monty-Hall-Problem

Goat or Car?



- „Geh auf’s Ganze“/“Let’s make a deal“
- Candidate chooses a door
- Two doors reveal a goat, one an expensive car
- After the choice is made, the host opens a goat door and gives the option to switch to the other closed door.
- Better to switch or to stay?



5 year survival rate



Giuliani had prostate cancer and said that the survival rate for prostate cancer in the US was 82% compared to 44% in England.

(First of all, the numbers were wrong)

[http://www.themoneytimes.com/articles/20071103/giuliani_gets_his_numbers_wrong-id-1012641.html]

More importantly, he compared apples with pears.

Survival rate



After diagnosis how many people are still alive after 5 years?

USA: screening of men without symptoms

-> more diagnoses at earlier age

UK: no screening

-> diagnosed men show symptoms and are older on average



Increase of Risk

Increase of risk



Babies of obese mothers have an 35% increased risk of dying before 60

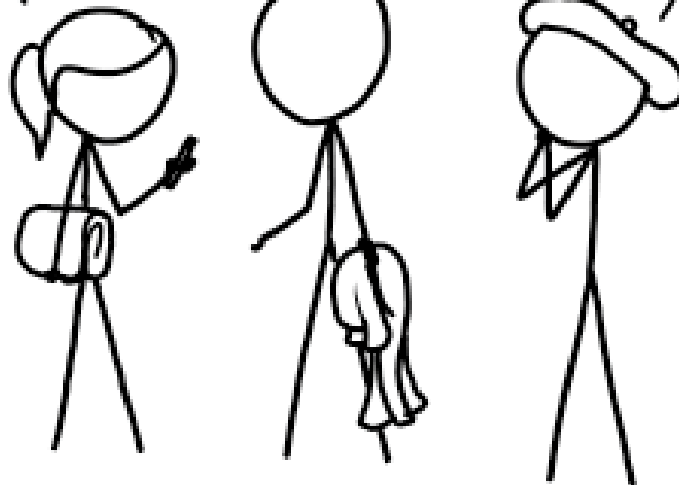
Is that an important information?

It depends: if normally 10 out of 10,000 die prematurely, then now expectedly 13.5 die.

WE SHOULD GO TO THE NORTH BEACH.
SOMEONE SAID THE SOUTH BEACH HAS
A 20% HIGHER RISK OF SHARK ATTACKS.

YEAH, BUT STATISTICALLY, TAKING
THREE BEACH TRIPS INSTEAD OF TWO
INCREASES OUR ODDS OF GETTING
SHOT BY A SWIMMING DOG CARRYING
A HANDGUN IN ITS MOUTH BY **50%**!

OH NO! THIS IS
OUR THIRD TRIP!



REMINDER: A 50% INCREASE
IN A TINY RISK IS **STILL TINY.**





Don't Fool Others

Statistics is hard

Scientific misconduct

- Made up/manipulated data
- Deletion of data
- Showing data in a different context
- Stealing of data/results
- Manipulate data or experiments of others
- Wrong reviews
- Hide financial interests



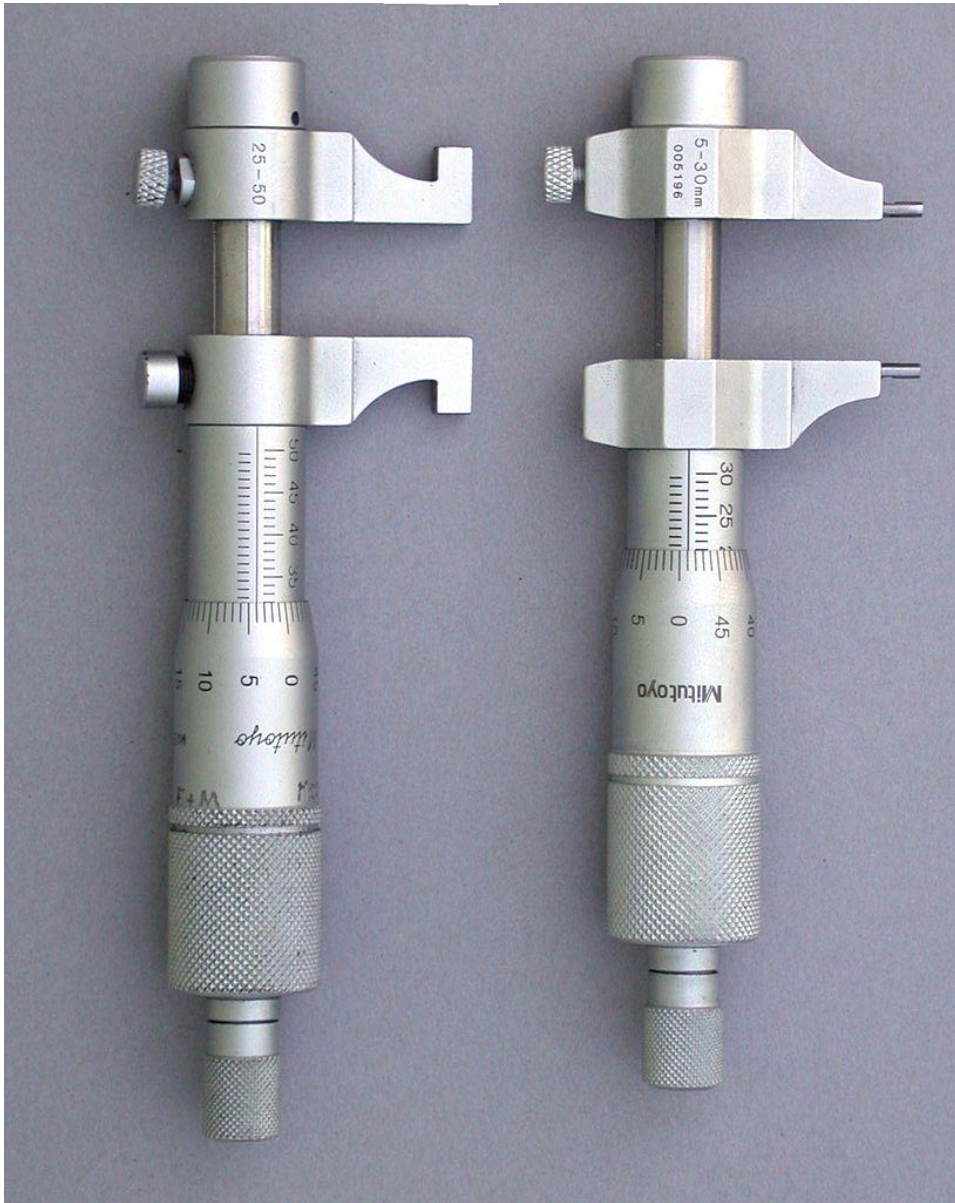
How to lie with statistics



- Precision
 - Wrong basis
 - Definition
 - Visualization
 - Usage of meaningless superlatives
 - Correlation vs. Causality
 - Ex post analysis
 - Reports of single quantities of distributions
 - Extrapolation
- Examples from
H.-H. Dubben, H.-P. Beck-Bornholdt: “Mit an Wahrscheinlichkeit grenzender Sicherheit”, rororo, 2005
A.K. Dewdney: “200 Prozent von nichts”, Birkhäuser, 1994
W. Krämer: “So lügt man mit Statistik”, Piper, 2000
H.-H. Dubben, H.-P. Beck-Bornholdt: “Der Hund, der Eier legt”, rororo 2000

Precision

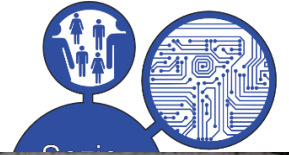
http://commons.wikimedia.org/wiki/File:MicrometerInside5-30_25-50.jpg, user graibeard



Higher precision invokes feelings of higher accuracy which might not have been achievable with the methods at hand.

Rule: presented precision should not be higher than the precision of the least accurate observed value.

Whom would you trust?



A bank offering an interest yield of 10.5% or one that offers 9-11%.

If you are asked to denote your property in order to get a credit, should you state "653.249€" or "approx. 650.000 €"?

In court, one expert reports the driver was breaking for around 150 Meter the other claims it to be 126,3 Meter.



Faulty precision

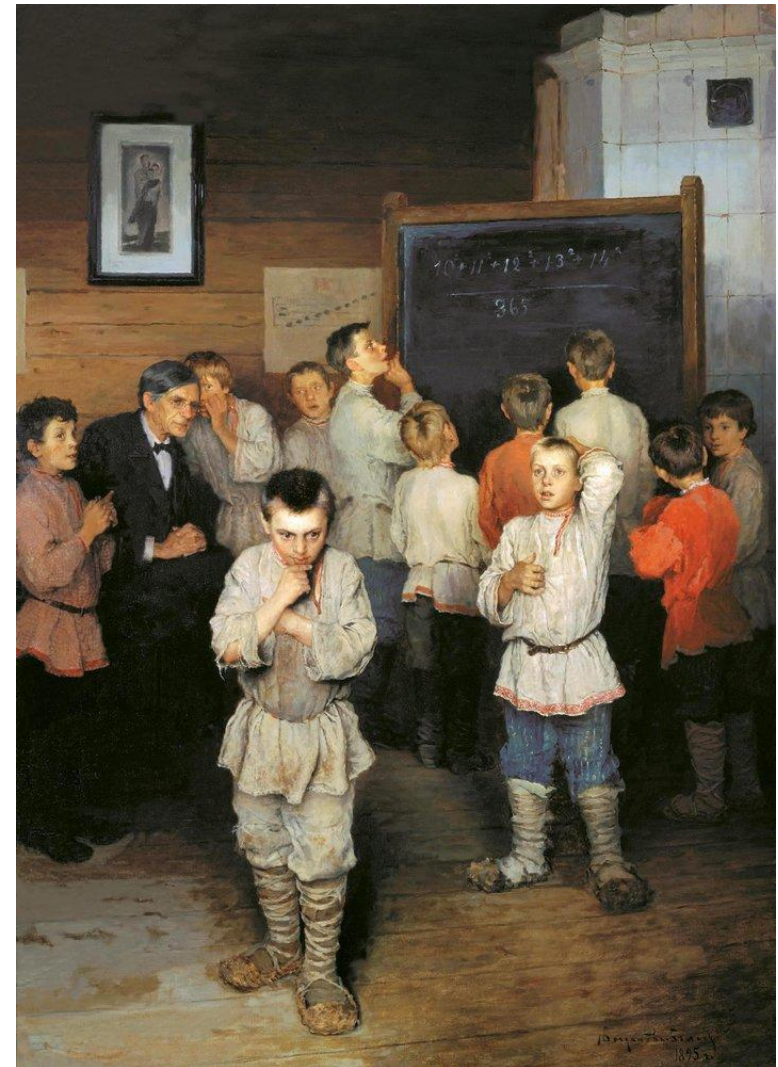


By addition of values
with different precision:

A house, ca. 600.000 €
plus 53.429 € cash

Note: In which interval is
the value of the house?
+/- 50.000

Thus, in sum, the total
property should be stated as
~650,000€



Faulty Precision II



Faulty precision very often arises by averaging:

Two engineers with a salary of around 30 k€ and one with 40k\$ have an average salary of 33.333 €. The result pretends to be very accurate.

It was also reported that Germans eat 3381 kcal (rank 19) per capita and day, just behind Greeks with 3400 kcal and in front of Canadians with 3374 kcal. Since these numbers cannot be that accurate, a ranking is meaningless. Krämer, S. 20-25



Number Psychology



What is the debt of Germany?

2 Trillion €

1.516 Billion €

1.515.608 Million €

1.515.608.069.526 €

Radioactivity can be measured in Ci or Bq

1 Ci = 1 decay

1 Bq = $37 * 10^9$ decays

What do you think, who uses which? Atom lobby? Ecologists?



4.9.2016: 2,26 Trillion €

(Im Deutschen sind es „nur“ 2 Billionen – leider immer noch dieselbe Zahl, aber unterschiedliche Benamungen...)



According to statistics in 1992, of all households in Frankfurt there are

49.2% with 1 person

28.3% with 2 persons

11.7% with 3 persons

8.2% with 4 persons

2.7% with 5 persons

A newspaper created the following headlines:
"Every second person lives alone!"

Presidential Candidates



Claim from Romney in 2012: „92.3% of job losses during the Obama presidency were from women“

At this time point there was a 19% point gap of women's preferences towards Obama.

Where does the number come from?

Compared to Jan 2009, there are still 740,000 jobs missing (due to the global crisis) and of these 683,000 were previously held by women, i.e., 92.3%.

Full story



- In Jan 2009, 4.5 million jobs were lost since the start of the crisis – predominantly among men. That was at the end of Bush’s presidency – but was he anti-male?
- The first phase of the job losses concerned construction and manufacturing – mostly occupied by men. The second phase those in industries occupied by women.
- If properly accounted for since the beginning of the crisis, all in all 5 million jobs were lost, of which 1.8 million were held by women (36%).

98% of all Catholic women have practiced contraception!

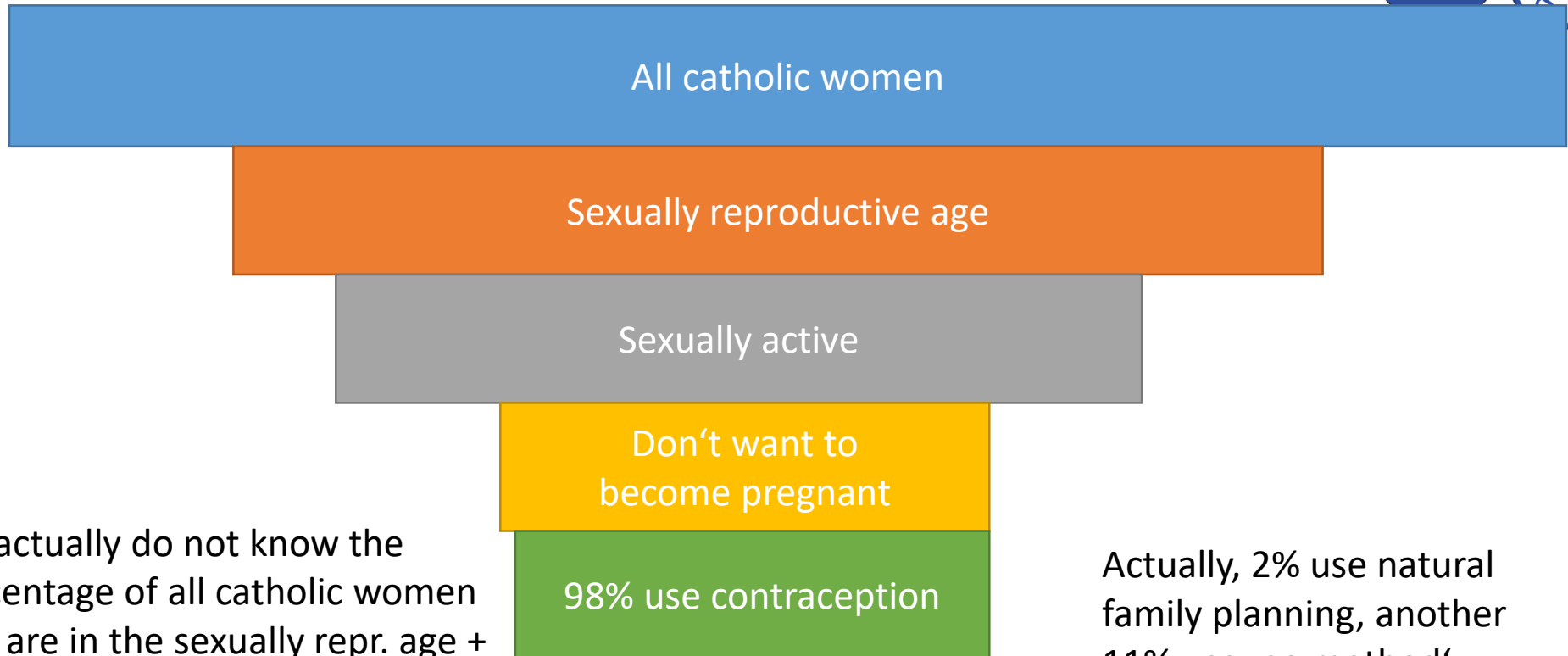


This finding of a research study was used as an argument for Obama's HHS mandate which requires all employer health plans to provide free contraceptives, sterilization, and abortion-inducing drugs.

The discussion of the above statement and the study from which it originated was given by Lydia Crew on a pro-catholic blog, found on the 29th of June, 2012:

http://www.whatswrongwiththeworld.net/2012/02/how_to_lie_with_statistics_exa_1.html

Full picture



We actually do not know the percentage of all catholic women that are in the sexually repr. age + and sexually active + do not want to be pregnant. But of those 98% (87% ?) practice contraception.

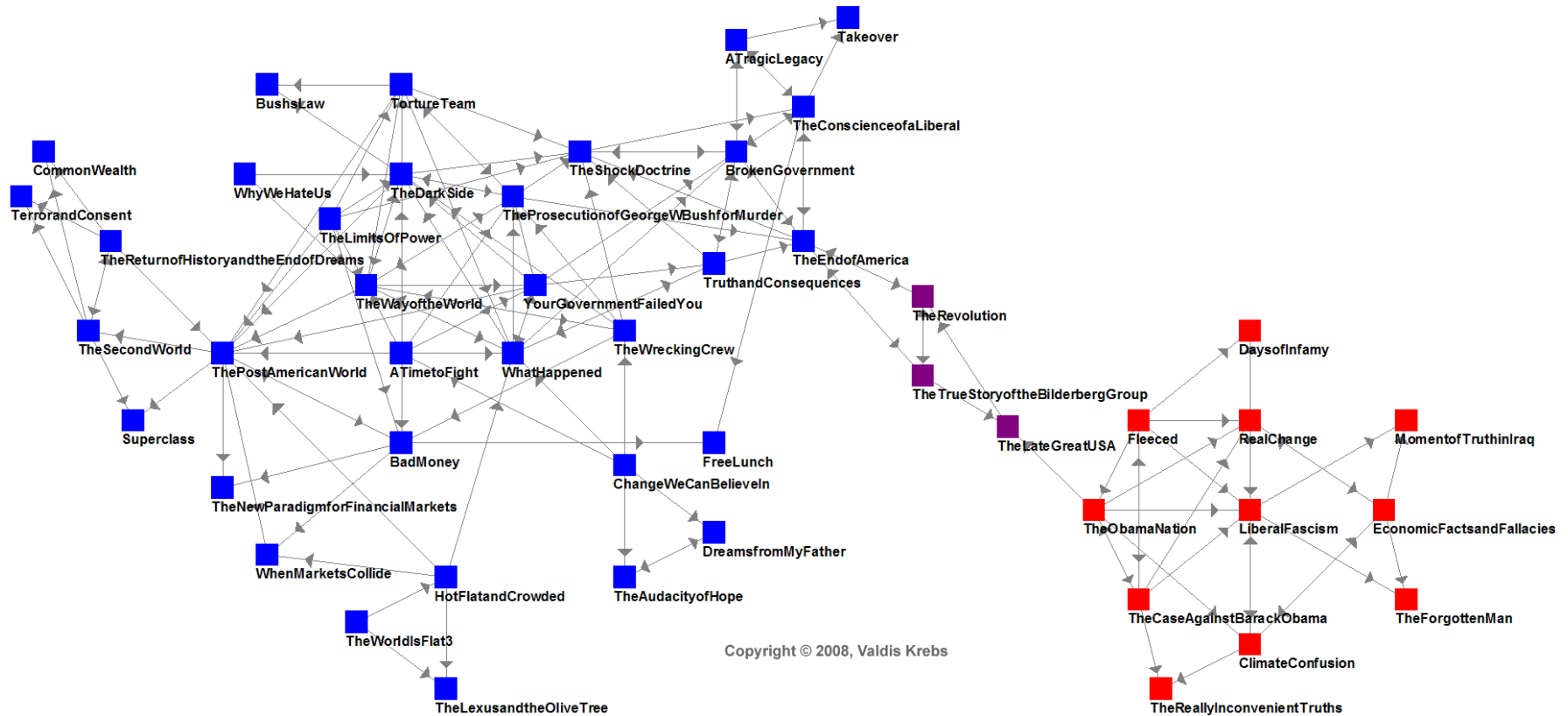
Actually, 2% use natural family planning, another 11% use ,no method'. So, maybe that is more like 87% practice contraception.

Disclaimer



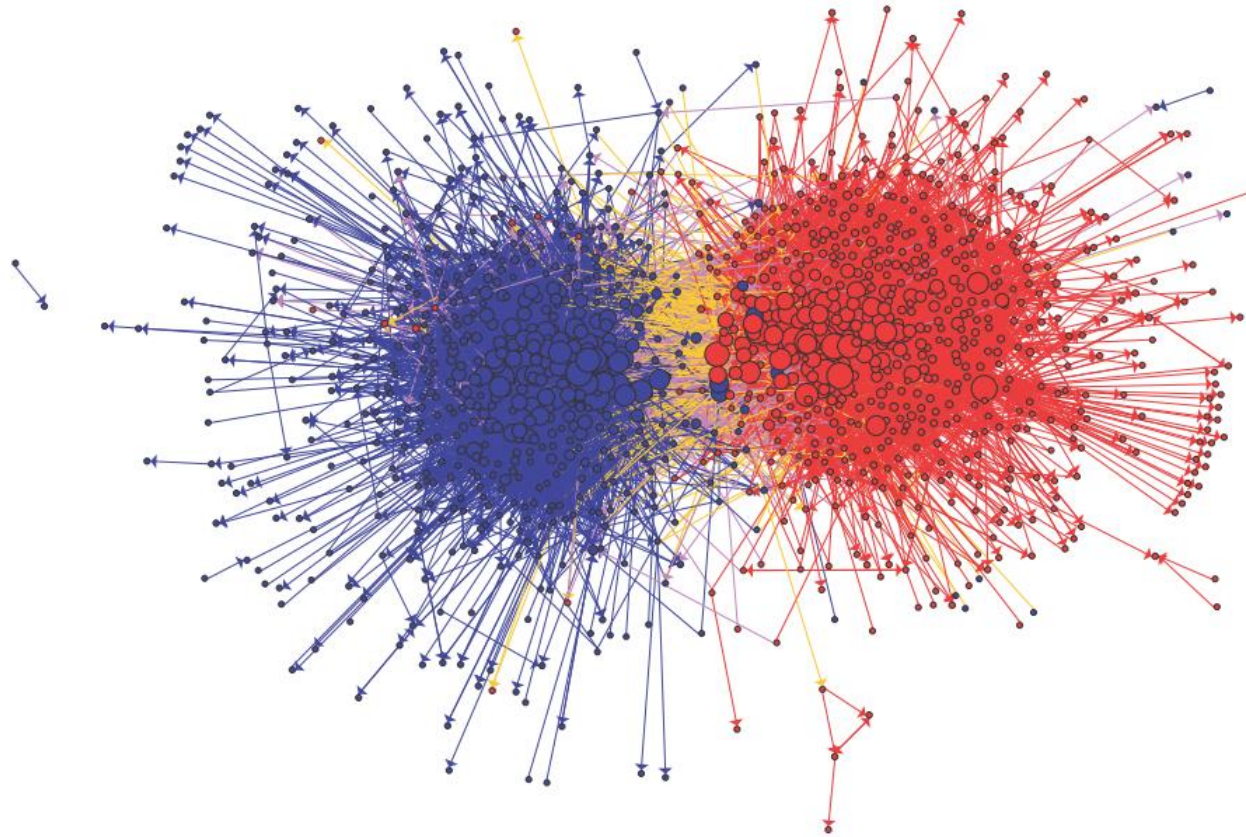
- The last two examples concern heavily discussed topics in US american politics. In both blog posts discussed here, a statistics used by one camp is attacked by someone from the other camp. It seems to me that the two blog posts do a good job on dissecting the wrong use of a statistics. I tried to pick one ill-used statistics from both camps for fairness. The topics chosen are by chance – just the ones that came up on a Google search on the 29th of June and do not express any opinion of mine.
- These examples should be taken as a warning sign that different political camps have very different incentives to use statistics and that it is always valuable to learn about the other camp's position.

Divided we fall...



Unfortunately, Valdis Krebs from orgnet.com showed in 2008 that people prefer to read books from their own camp, rather than sharpen their arguments by reading the other camp's books.

Divided they blog...



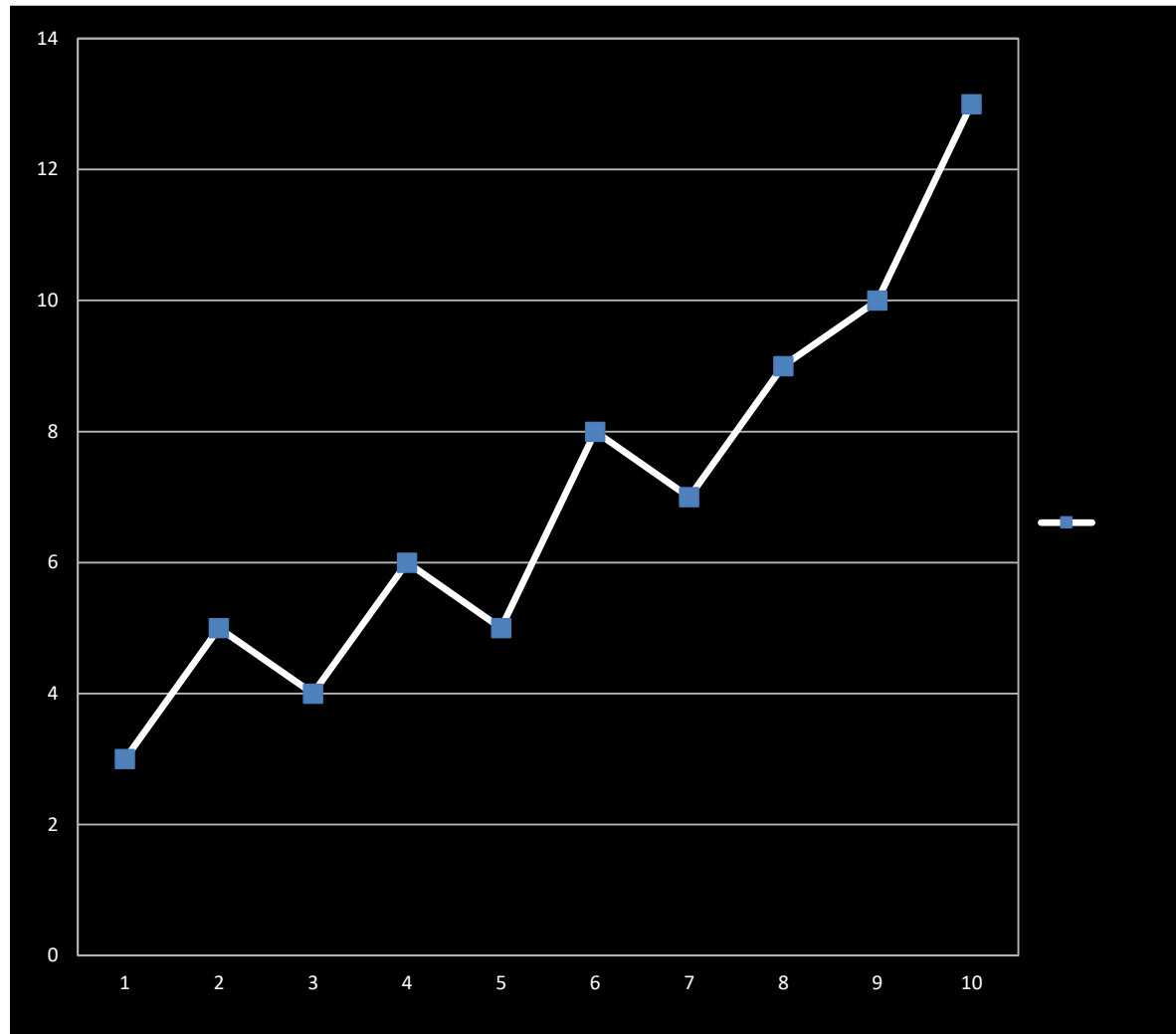
Similarly, Adamic and Glance showed that bloggers of the two camps also do not very often refer to each other.



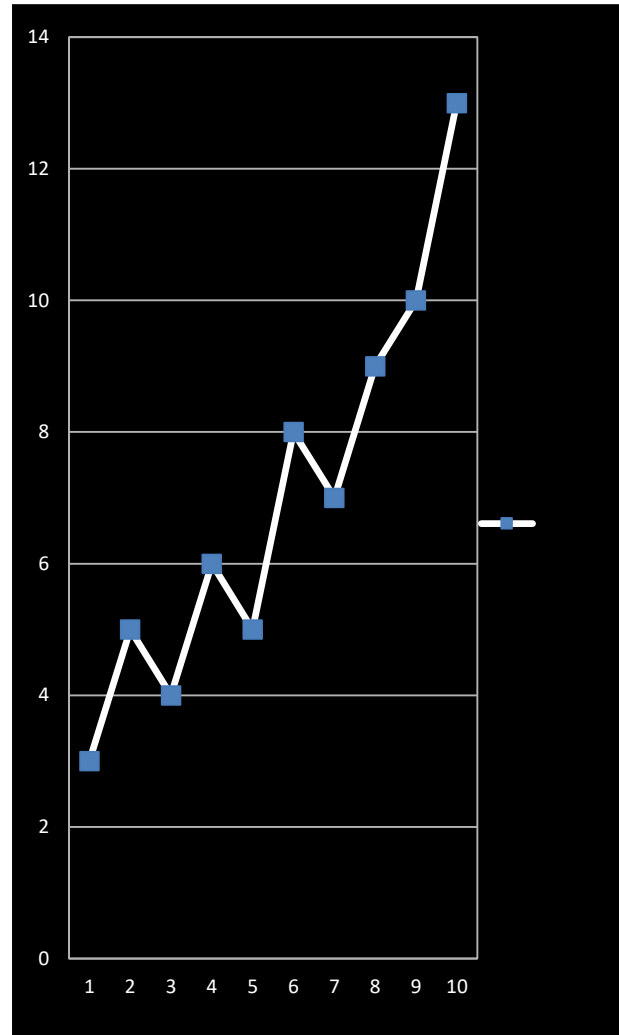


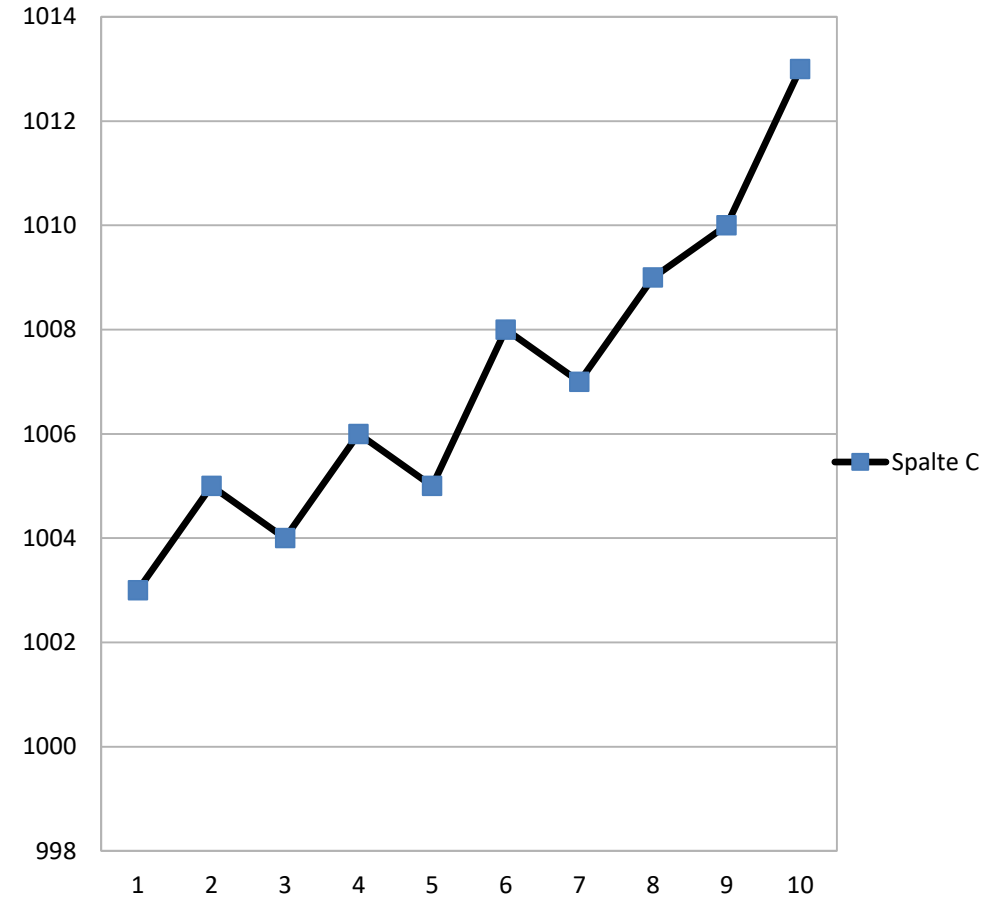
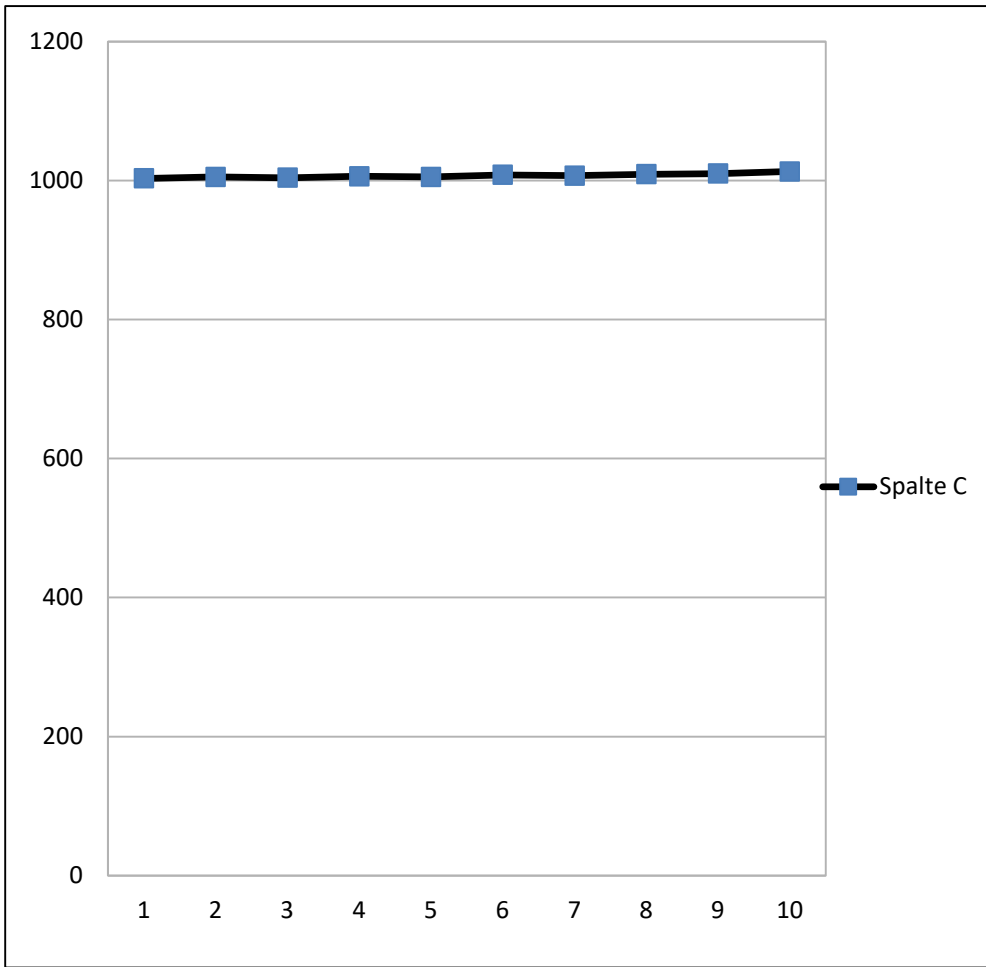
Back to Statistics

Beware of the curve!



Beware of the curve!

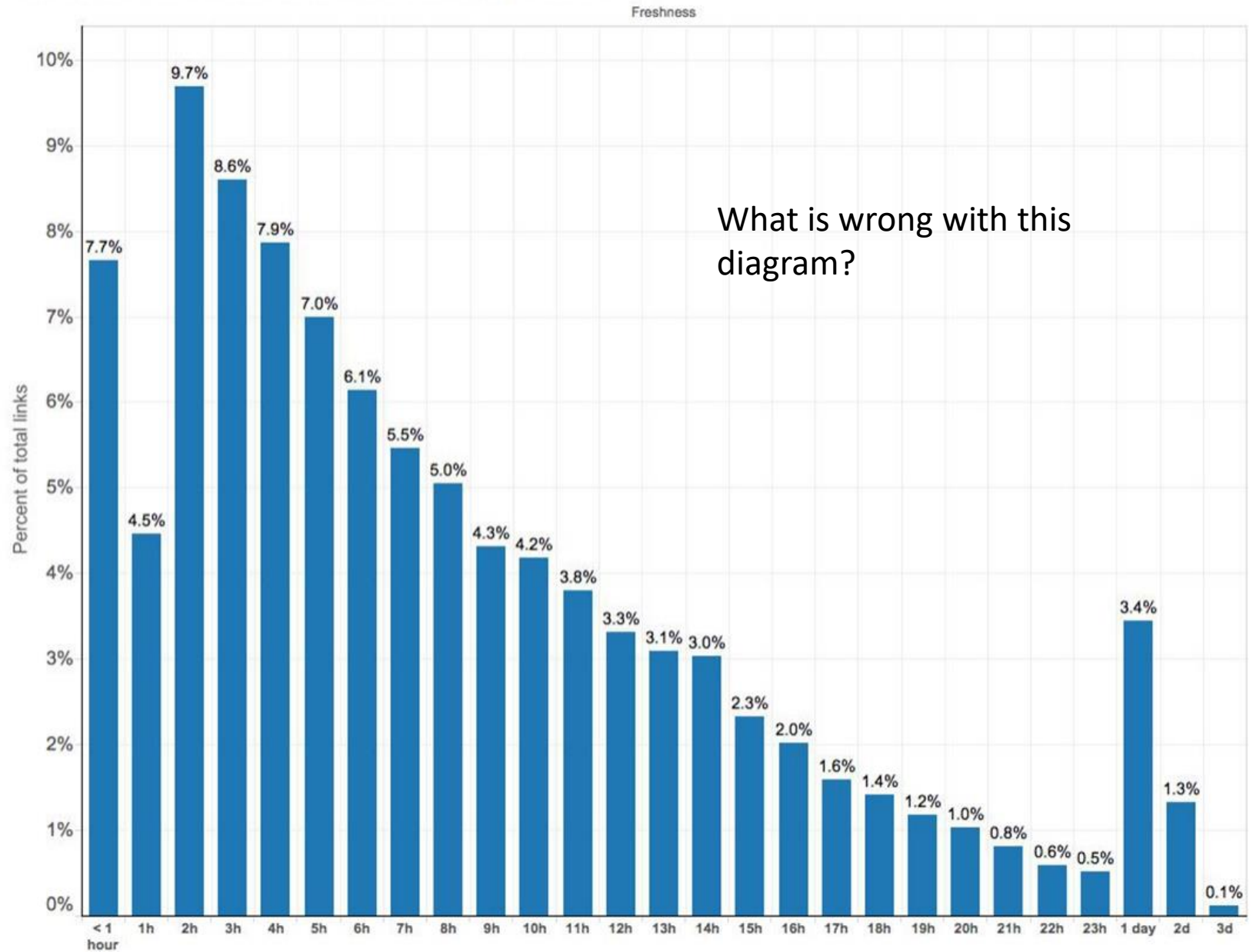






Inhomogeneous time axes

Distribution of the freshness of the links in the Google News box



From: Daniel Trielle and Nicholas Diakopoulos: „How Google Shapes the News You See About the Candidates - Not all links in the Google News box are equal.”, Slate, 3rd of November, 2016

http://www.slate.com/articles/technology/future_tense/2016/11/how_google_shapes_the_news_you_see_about_the_candidates.html

Inhomogeneous axes



- For any diagram, it is important that equally distant points on the axes have the same meaning.
- While for the first 24 intervals, each represents an hour, the last three each represent a day.
- Thus, the sudden „increase“ from 0.5% to 3.4% does not indicate that articles with some „staying power“ are favored, as stated by the authors:
 - „Overall, the fraction of articles with different freshness tapers off over the course of a day, though there is a small bump again for articles with some staying power.“ (see the same source)
- This ‘bump’ is actually the integral over the probability of 24 hours (from 24 hours of publication age to 47 hours).
- It seems as if the “real” curve (with respect to hours) is rather smooth.



Will Rogers Phenomenon

'Helpful' redefinitions– The Will-Rogers-Phenomenon



Assume that in 1990 patients are differentiated into two groups (with or without metastases).

Assume, patients with metastases have a life expectancy of 10 months, those without one of 10 years.

There are patients with small metastases that cannot be detected in 1990. Their LE is 30 months.

Will-Rogers-Phenomenon



In 2000 the method is reexamined.

Diagnosis has improved such that patients with small metastases are now put into the group 'with metastases'.

Both groups now have a 'better' life expectancy!

The method seems to work – but that is just because of a redefinition of the two groups of patients.

Actually, the method might make things worse and this would be disguised by the new definition of the two groups.



Exaggerations

Superlatives



Let's assume that a popular tennis player B has never played against a lefthander. The following head-lines are 'correct' but misleading:

Poor B! Never successful against lefthanders!

Bravo, B! Undefeated by lefthanders!

Translated from Krämer, S. 92





A Last insight

Laws of discovery



Arnold's law (implied by statements in his many letters disputing priority, usually in response to what he sees as neglect of Russian mathematicians)

Discoveries are rarely attributed to the correct person
(of course Arnold's law is self-referential).

Berry's law (prompted by the observation that the sequence of antecedents under law 1 is endless)

Nothing is ever discovered for the first time

Whitehead's law (quoted by Max Dresden at the beginning of his biography of Kramers)

Everything of importance has been said before by
someone who did not discover it.

Take Home Messages



Statistics is often rather **unintuitive.**



It's easy to be fooled



It's even easier to fool others.

(but evil...)





The difference between
theory and **practice**
is much higher
in **practice**
than in **theory**.

Translated from Ernst Ferstl, "Unter der Oberfläche", Wien-Klosterneuburg, EDITION VA BENE, Ausgabe 1996, ISBN 3-85167-026-4



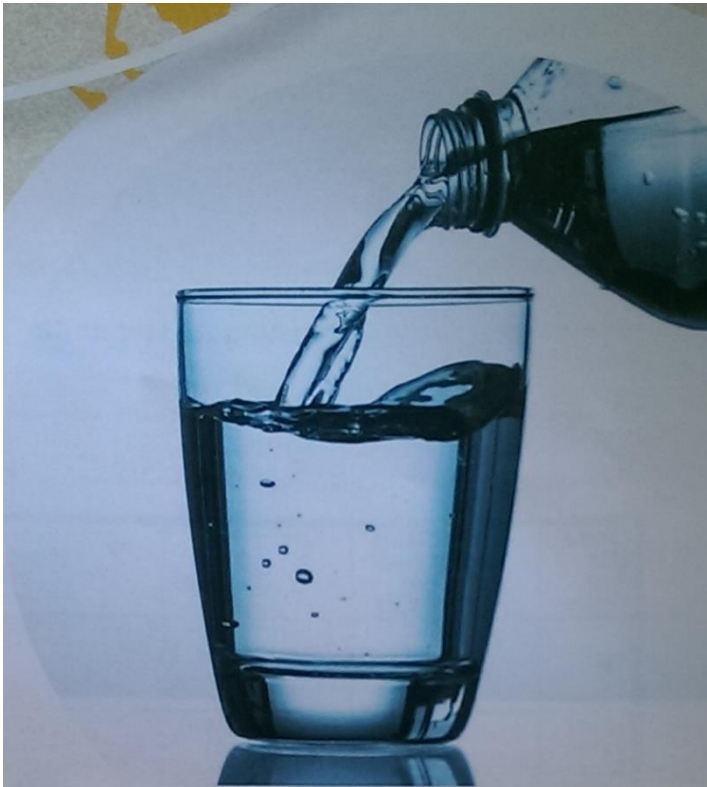
Some Real-World Statistics Fails

Fehlende Daten



Spekulation:

Die zugrunde liegende Statistik enthält höchstwahrscheinlich die Menge an **verkauftem Mineralwasser**. Diese war sicherlich 1970 kleiner als die Menge verkauften Bieres. Dass die Leute damals tatsächlich weniger Wasser als Bier



WASSER

Wer hätte das gedacht? Mineralwasser ist das beliebteste Getränk in Deutschland. Durchschnittlich 144 Liter konsumieren wir pro Kopf im Jahr. Egal, ob still oder sprudelnd, Wasser trinken ist in. Dabei gilt Deutschland als Land der Biertrinker. Hier muss aber bis ins Jahr 1970 zurückgeschaut werden, da wurde noch mehr Bier als Wasser getrunken. Aktuell liegt der Pro-Kopf-Verbrauch an Bier bei 107 Litern.



Quelle des Artikels

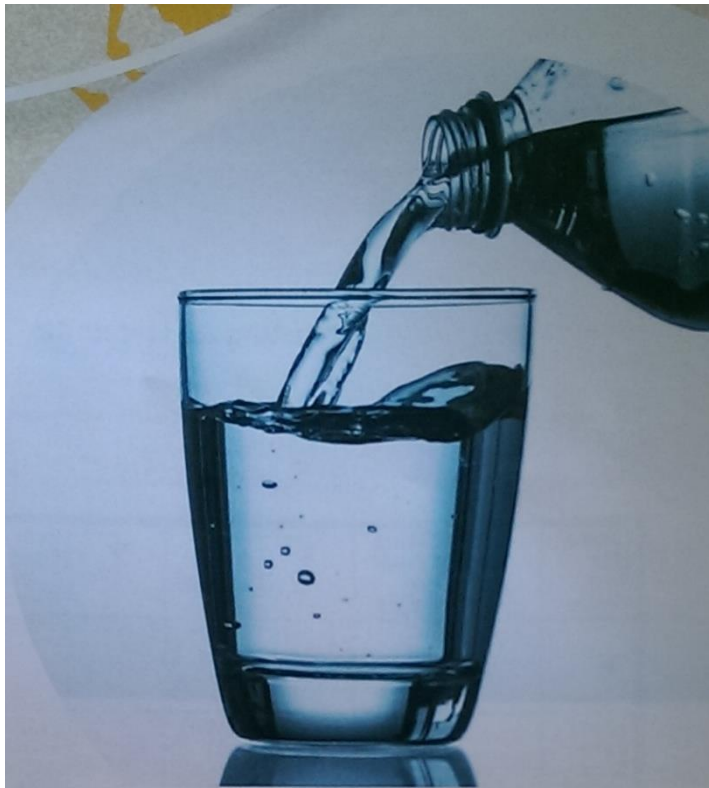
Fehlende Daten



Spekulation:

Die zugrunde liegende Statistik enthält höchstwahrscheinlich die Menge an **verkauftem Mineralwasser**. Diese war sicherlich 1970 kleiner als die Menge verkauften Bieres. Dass die Leute damals tatsächlich weniger Wasser als Bier **getrunken** haben, ist dagegen unwahrscheinlich!

Anfrage an Redaktion läuft (11.7.2015) – wurde nie beantwortet (04.09.2016)



WASSER

Wer hätte das gedacht? Mineralwasser ist das beliebteste Getränk in Deutschland. Durchschnittlich 144 Liter konsumieren wir pro Kopf im Jahr. Egal, ob still oder sprudelnd, Wasser trinken ist in. Dabei gilt Deutschland als Land der Biertrinker. Hier muss aber bis ins Jahr 1970 zurückgeschaut werden, da wurde noch mehr Bier als Wasser getrunken. Aktuell liegt der Pro-Kopf-Verbrauch an Bier bei 107 Litern.

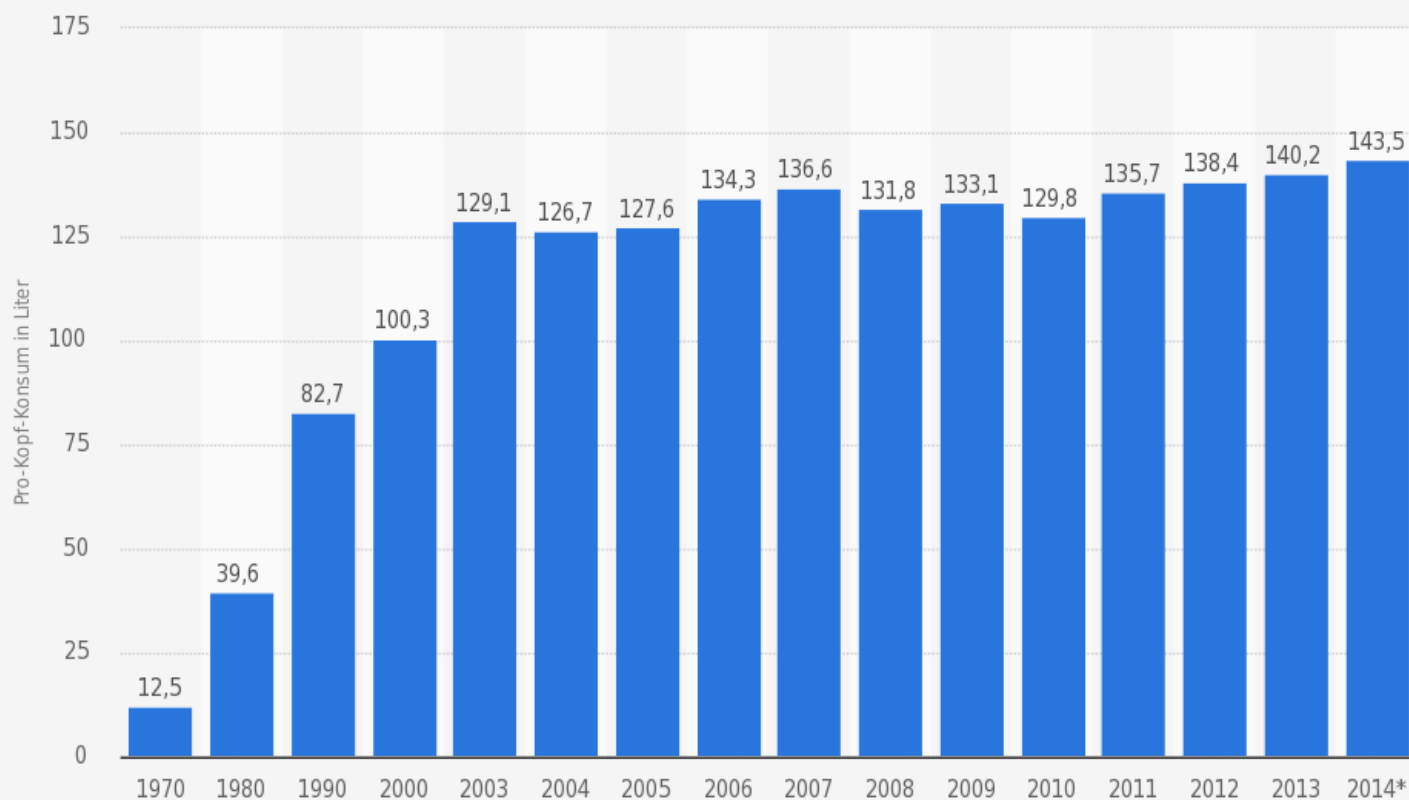


Quelle des Artikels

Vermutliche Quellen: Statista.de



Pro-Kopf-Konsum von Mineral- und Heilwasser in Deutschland in den Jahren 1970 bis 2014* (in Liter)



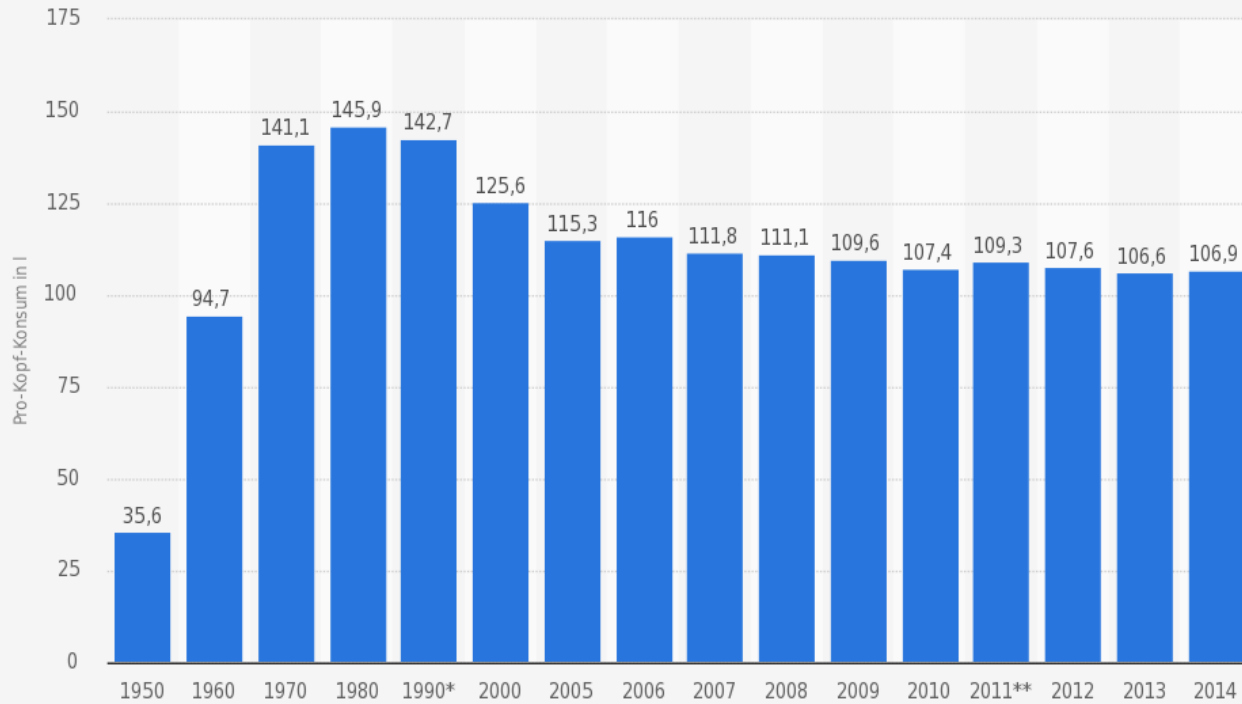
Quelle:
Verband Deutscher Mineralbrunnen
© Statista 2015

Weitere Informationen:
Deutschland

Bier



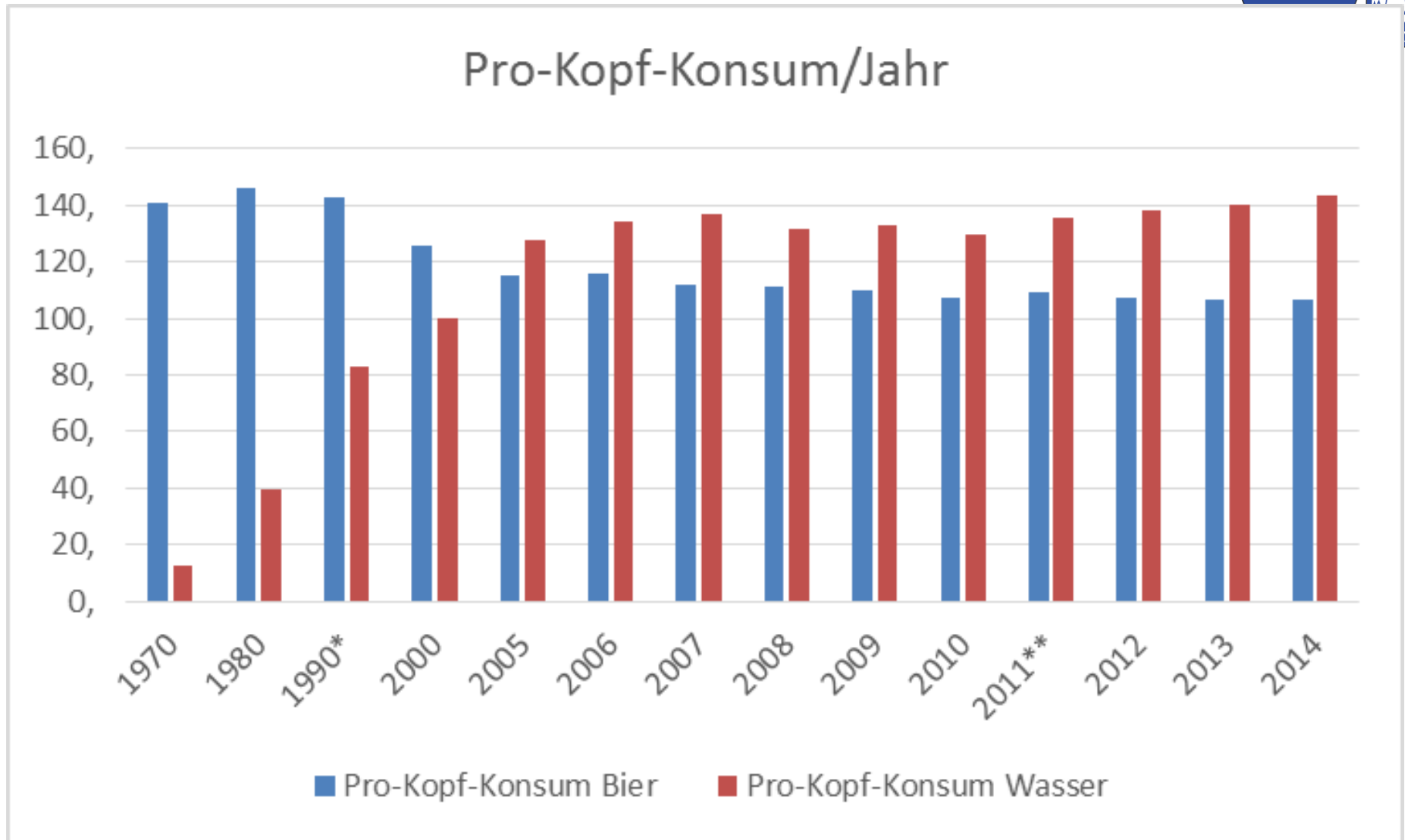
Pro-Kopf-Konsum von Bier in Deutschland in den Jahren 1950 bis 2014 (in Liter)



Quelle:
Deutscher Brauer-Bund
© Statista 2015

Weitere Informationen:
Deutschland; inkl. alkoholfreier Sorten

Pro-Kopf-Konsum



The ultimate test

Where does this
statistic suggest
something that
is not
truly
there?



Nährwerte	Pro 100 g Pulver	Pro Portion (15 g Pulver + 150 ml fettarme Milch)	% Tages- zufuhr*
Brennwert	1550 kJ = 367 kcal	533 kJ = 127 kcal	6
Eiweiß	7,7 g	6,1 g	12
Kohlenhydrate	67,2 g	17,3 g	6
- davon Zucker	63,6 g	16,7 g	19
Fett	7,5 g	3,5 g	5
- davon gesättigte Fettsäuren	4,5 g	2,1 g	10
Ballaststoffe	9,7 g	1,5 g	6
Natrium	0,04 g	0,07 g	3
Calcium	50 mg (6%)**	185 mg (23%)**	

* Richtwerte basierend auf einer Ernährung von 2000 kcal. Der persönliche Bedarf variiert nach Alter, Geschlecht, Gewicht und körperlicher Aktivität.
** der empfohlenen Tageszufuhr

Mit einer Tasse Feinste heiße Schokolade, zubereitet mit fettarmer Milch, nehmen Sie fast ein Viertel der empfohlenen Tageszufuhr an Calcium auf.

NUTRITIONAL SOURCE

References



Theory

<http://en.wikipedia.org/wiki/Theory>,
2.11.2010

Francis Bacon

http://de.wikipedia.org/wiki/Francis_Bacon,
2.11.2010

Douglas Futuyma on evolution theory

http://www.stephenjaygould.org/library/futuyma_theory.html

J. Mayer, K. Khairy, J. Howard: "How to draw an elephant with four complex parameters", Am. J. Phys. 78 (6), p. 648-9, 2010

L. Sachs, J. Hedderich: Angewandte Statistik, Springer Verlag, 12. Auflage, 2006

Walter Krämer, "So lügt man mit Statistik", Piper Verlag, München, 2000

A.K. Dewdney, "200 Prozent von nichts", Birkhäuser Verlag, Basel, 1994

Hans-Hermann Dubben und Hans-Peter Beck-Bornholdt:

Der Hund, der Eier legt, rororo, Reinbek, 2010 (5. Auflage)

Der Schein der Weisen, rororo, Reinbek, 2010 (7. Auflage)

Mit an Wahrscheinlichkeit grenzender Sicherheit, rororo, Reinbek, 2005

Gerd Gigerenzer, *Das Einmaleins der Skepsis: Über den richtigen Umgang mit Zahlen und Risiken*. Berlin: Berlin-Verlag, 2002

References



Essential:

Gerd Gigerenzer, Stefan Krauss and Oliver Vitouch: „The Null Ritual - What You Always Wanted to Know About Significance Testing but Were Afraid to Ask”, In: [The SAGE Handbook of Quantitative Methodology for the Social Sciences](#) (Editor: David Kaplan), (pp. 392-409). : SAGE Publications Ltd, 2004