Learning from Data: the art of statistics

#LSEStats

David Spiegelhalter

Winton Professor of the Public Understanding of Risk, University of Cambridge

Chair: Fiona Steele

Professor in Statistics and Deputy Head of Department

Hosted by The Department of Statistics, LSE

Learning from data: the art of statistics

David Spiegelhalter

Chair of the Winton Centre for Risk & Evidence Communication, University of Cambridge

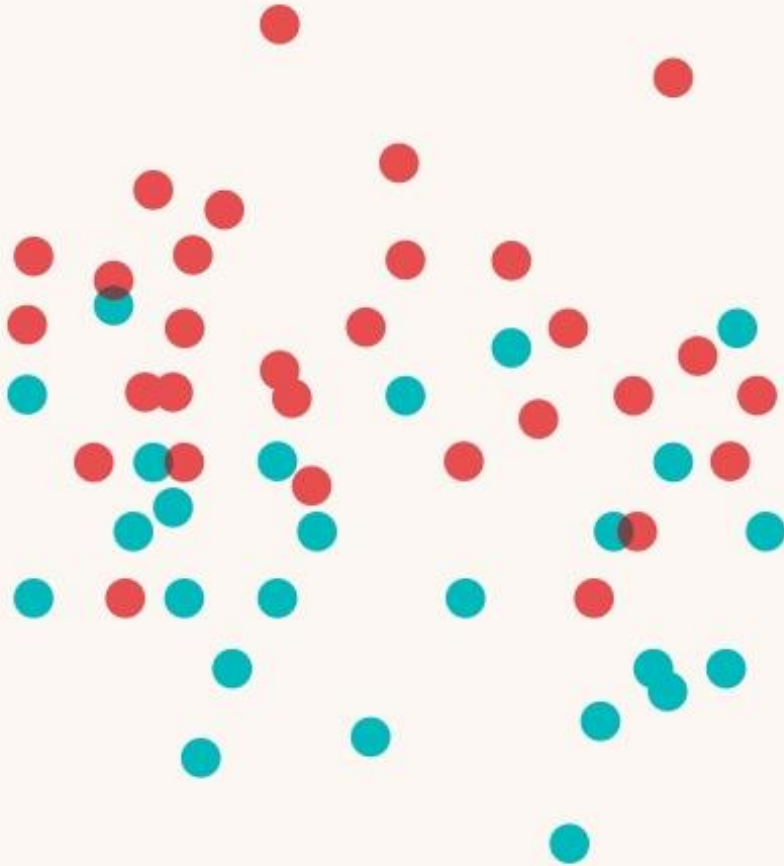*ex-President, Royal Statistical Society (2017-2018)*

LSE 2019

out March 28th!

Information Science and Statistics

Robert G. Cowell · A. Philip Dawid
Steffen L. Lauritzen · David J. Spiegelhalter

Probabilistic Networks and Expert Systems

Exact Computational Methods
for Bayesian Networks

Springer

WILEY

Bayesian Approaches to Clinical Trials and Health-Care Evaluation

David J. Spiegelhalter
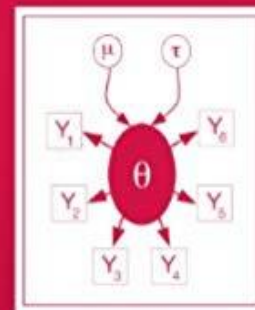Keith R. Abrams
Jonathan P. Myles

STATISTICS IN PRACTICE

Texts in Statistical Science

The BUGS Book

A Practical Introduction to Bayesian Analysis

David Lunn
Christopher Jackson
Nicky Best
Andrew Thomas
David Spiegelhalter

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

FOUR Climate Change by Numbers

Home   Clips

This programme is not currently available on BBC iPlayer

At the heart of the climate change debate is a paradox - we've never had more information about our changing climate, yet surveys show that the public are, if anything, getting less sure they understand what's…

Last on

BBC FOUR   Thu 5 Mar 2015
22:00
BBC FOUR

🕐 1 hour, 15 minutes

FOUR Tails You Win: The Science of Chance

Home   Clips

Next on

DURATION: 1 HOUR
Smart and witty, jam-packed with augmented-reality graphics and fascinating history, this film, presented by Professor David Spiegelhalter, tries to pin down what chance is and how it works in the real world. For…
> SHOW MORE

78 Share f ✆ ▣

BBC FOUR   **Next Thursday**
21:00
BBC Four

See all upcoming broadcasts of Tails You Win: The Science of Chance (3)

Numbers are often used to persuade rather than inform

Data does not speak for itself

INTRODUCTION

The numbers have no way of speaking for themselves. We speak for them. We imbue them with meaning.
— Nate Silver, *The Signal and the Noise*[1]

The traditional statistics course

- Describing data with summary statistics
  - *dull*
- Probability theory for drawing random observation from a population distribution
  - *difficult and mathematical*
- Probability theory for distributions of summary statistics
  - *mathematical and incomprehensible*
- Formulae for statistical tests
  - *mathematical, unmotivated, just a bag of tools*
- (If lucky) Examples of using statistical models in real life.

A 'modern' statistical course

- Motivate by problem solving
- Start with visualisation and exploring data
- Focus on what can be reasonably learned from data, biases in data, concluding causation, etc
- Models and algorithms
- Assessing uncertainty through re-sampling data ('bootstrap')
- Probability theory as neat way of turning random variation into uncertainty about what is true
- Hypothesis testing and its potential problems
- Bayesian methods

All these rather abstract, challenging, ideas are there to help answer real questions

- The 'data cycle'

- eg PPDAC (promoted in New Zealand)

Are You a Data Detective?

PROBLEM
* understanding and defining the problem
* how do we go about answering this question

PLAN
* what to measure & how?
* study design?
* recording?
* collecting?

DATA
* collection
* management
* cleaning

ANALYSIS
* sort data
* construct table, graphs
* look for patterns
* hypothesis generation

CONCLUSION
* interpretation
* conclusions
* new ideas
* communication

Data detectives use PPDAC

www.censusatschool.org.nz

Looking at data
What was the pattern of Harold Shipman's murders?
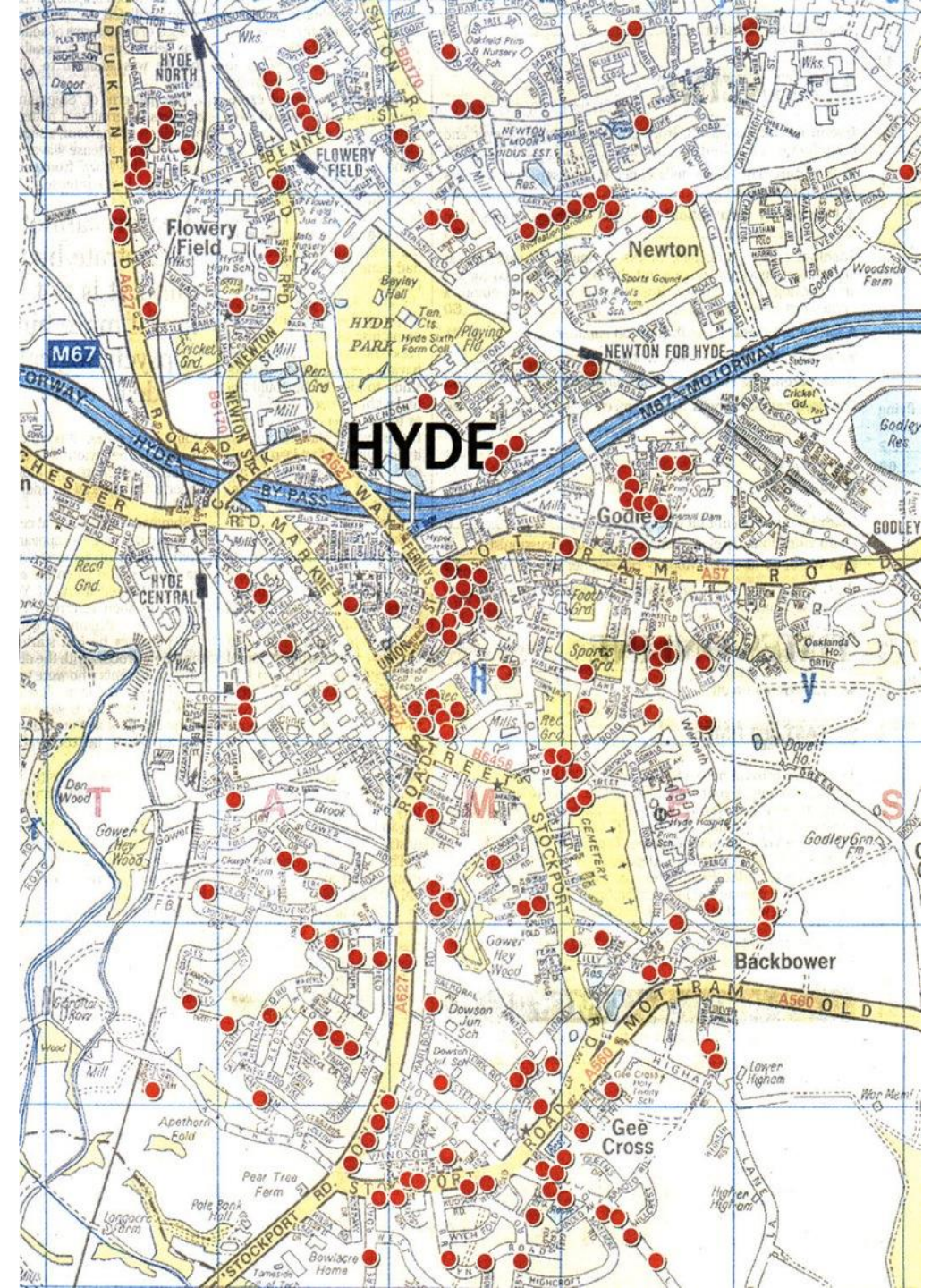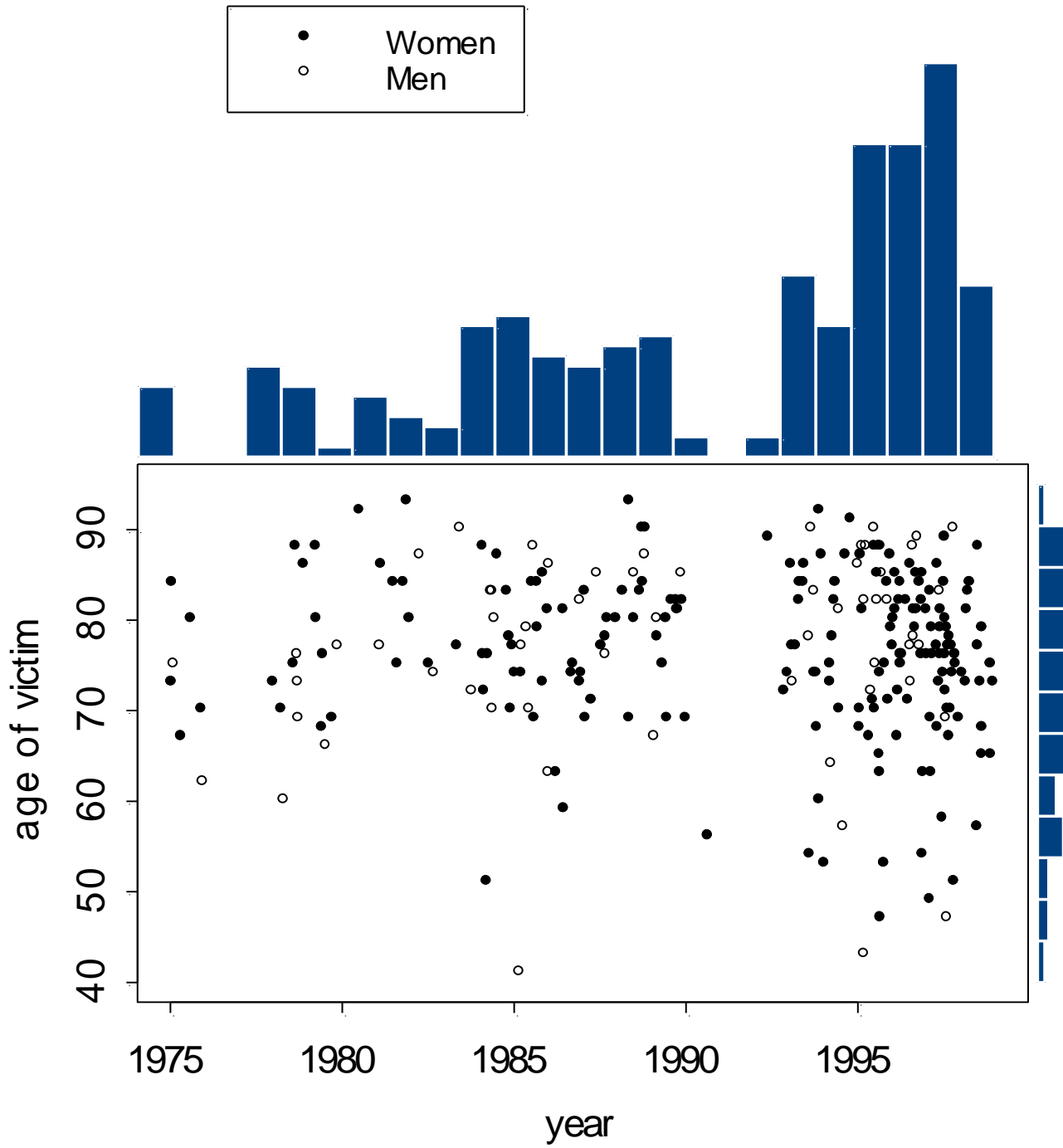
'I have nothing to hide'

Dr Harold Shipman, general practitioner, on his arrest in September 1998

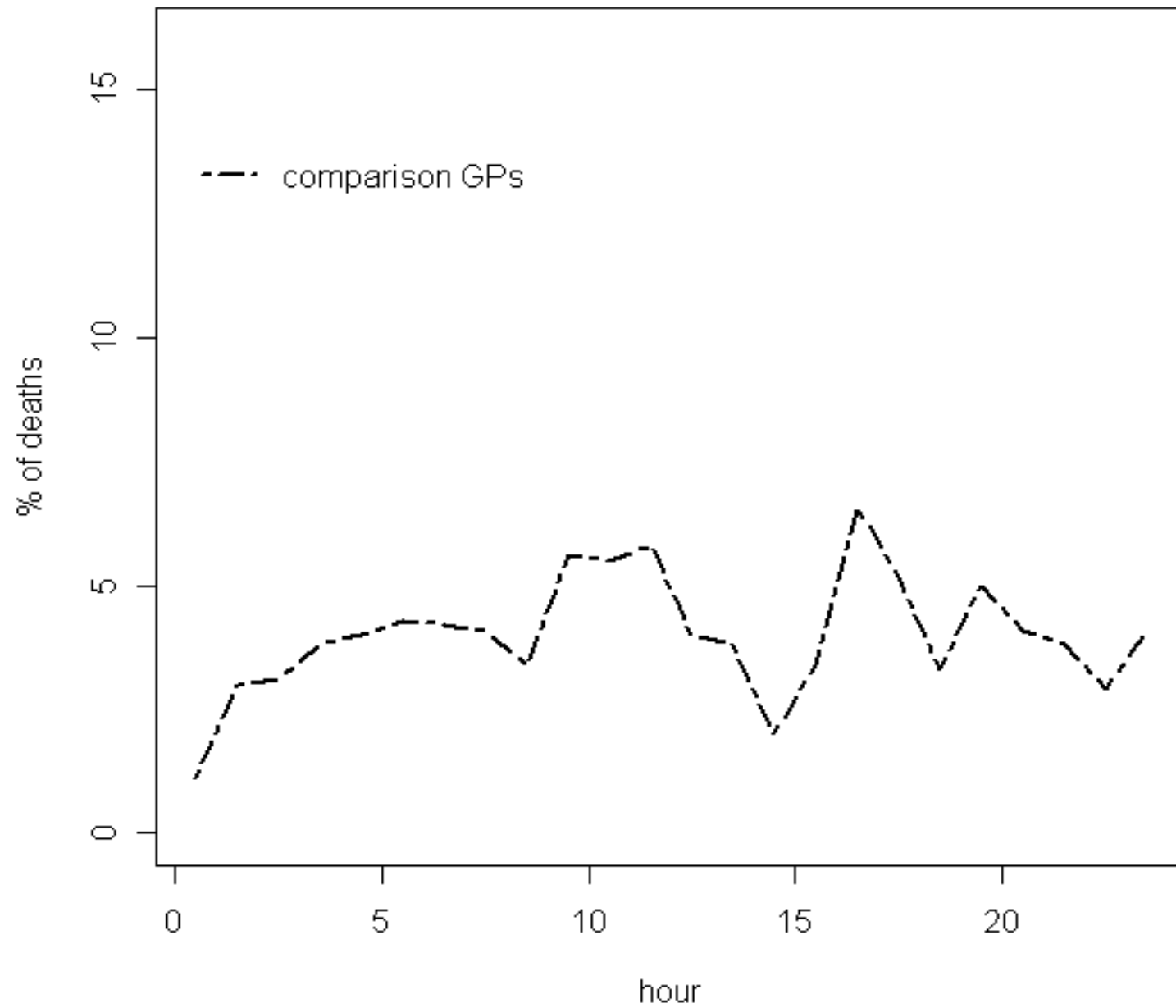Shipman Inquiry July 2002:

215 definite victims,

45 probable

Looking at data
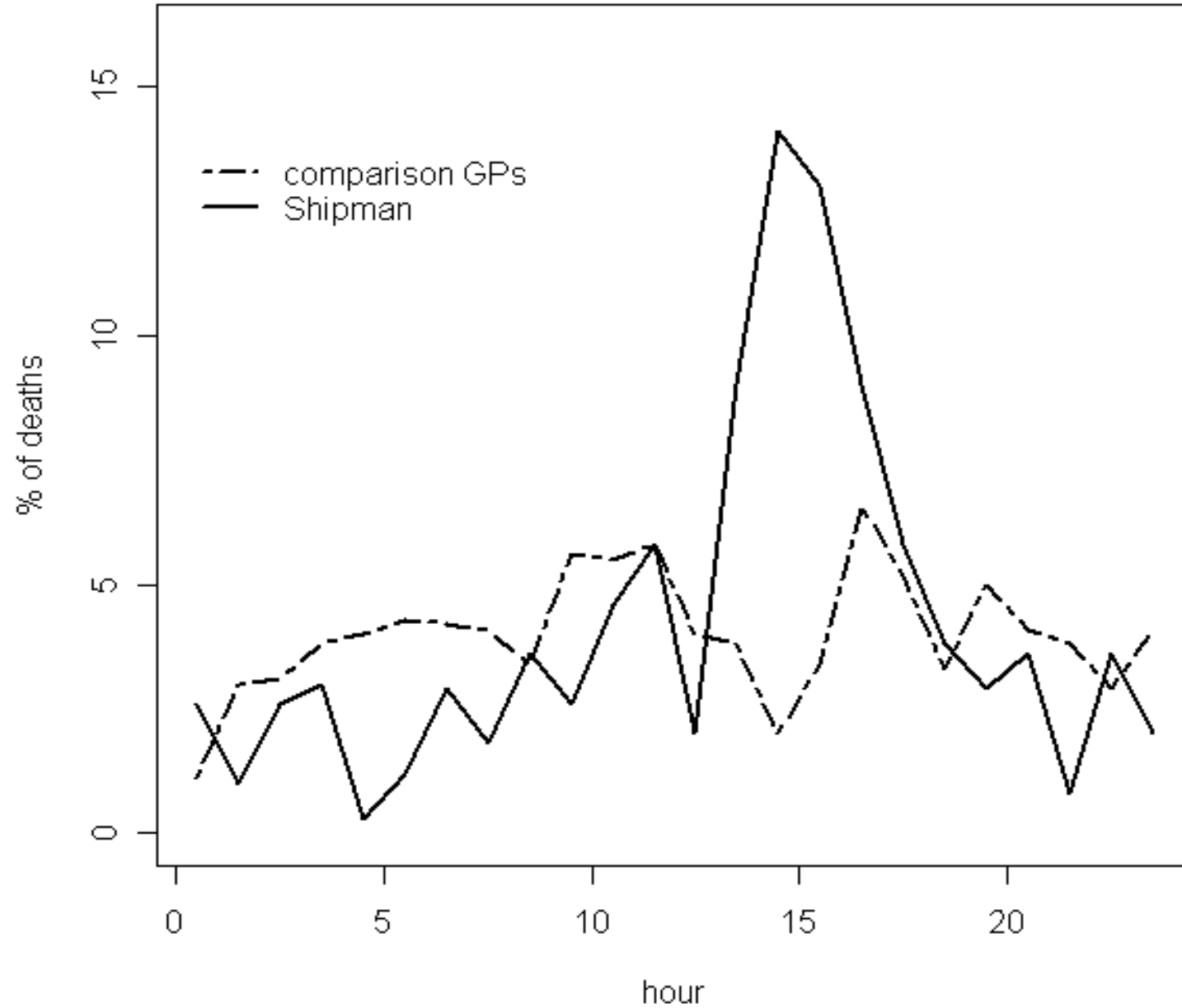## *What was the pattern of Harold Shipman's murders?*

- **Problem**: can more detail tell us more about what Shipman did?

- **Plan**: compare actual times at which his patients died with the times of deaths recorded by other local GPs

- **Data**: a huge exercise requiring examination of death certificates

- **Analysis**: simple plotting.....

# % of deaths in each hour of the day

People
die at all
hours

Legend: — · — comparison GPs

Y-axis: % of deaths (0, 5, 10, 15)

X-axis: hour (0, 5, 10, 15, 20)

% of deaths in each hour of the day

People die at all hours

- but not Shipman's victims

comparison GPs
Shipman

Inference and bias
How many sexual partners have people in Britain had in their lifetime?

- **Problem**: cannot know this as a fact

- **Plan**: survey in which people are carefully asked about the sexual activity (Natsal)

- **Data**: reports of numbers of partners

- **Analysis**: plotting and summary statistics

# How many sexual partners do people report?
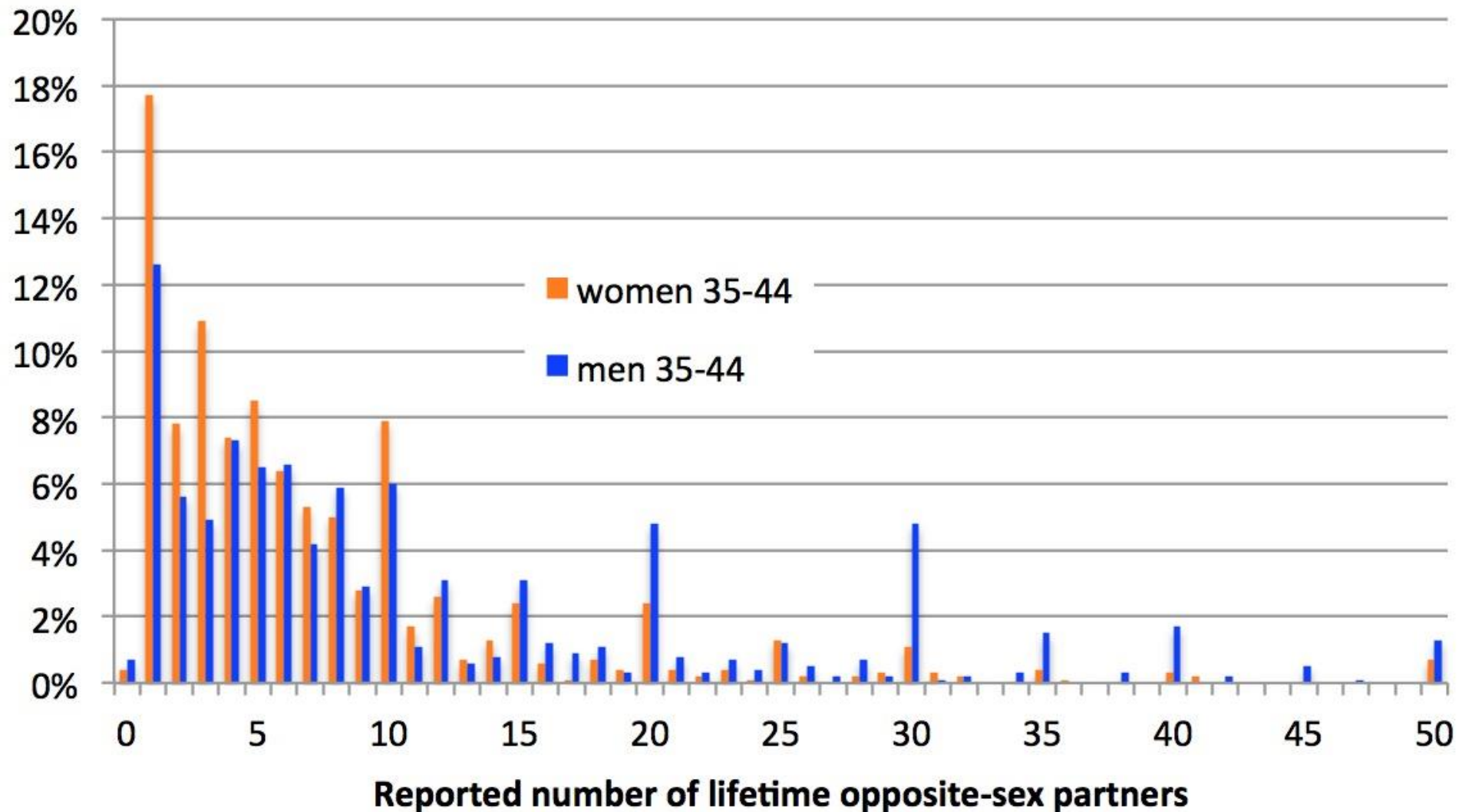
Inference and bias
*How many sexual partners have people in Britain **really** had in their lifetime?*

Reported number of sexual partners in lifetime	Men aged 35–44	Women aged 35–44
Mean	14.3	8.5
Median	8	5
Mode	1	1
Range	0 to 500	0 to 550
Inter-quartile range	4 to 18	3 to 10
Standard deviation	24.2	19.7

- **Conclusions**: can we generalise this to the whole population?????

Induction: the stages in generalising from data
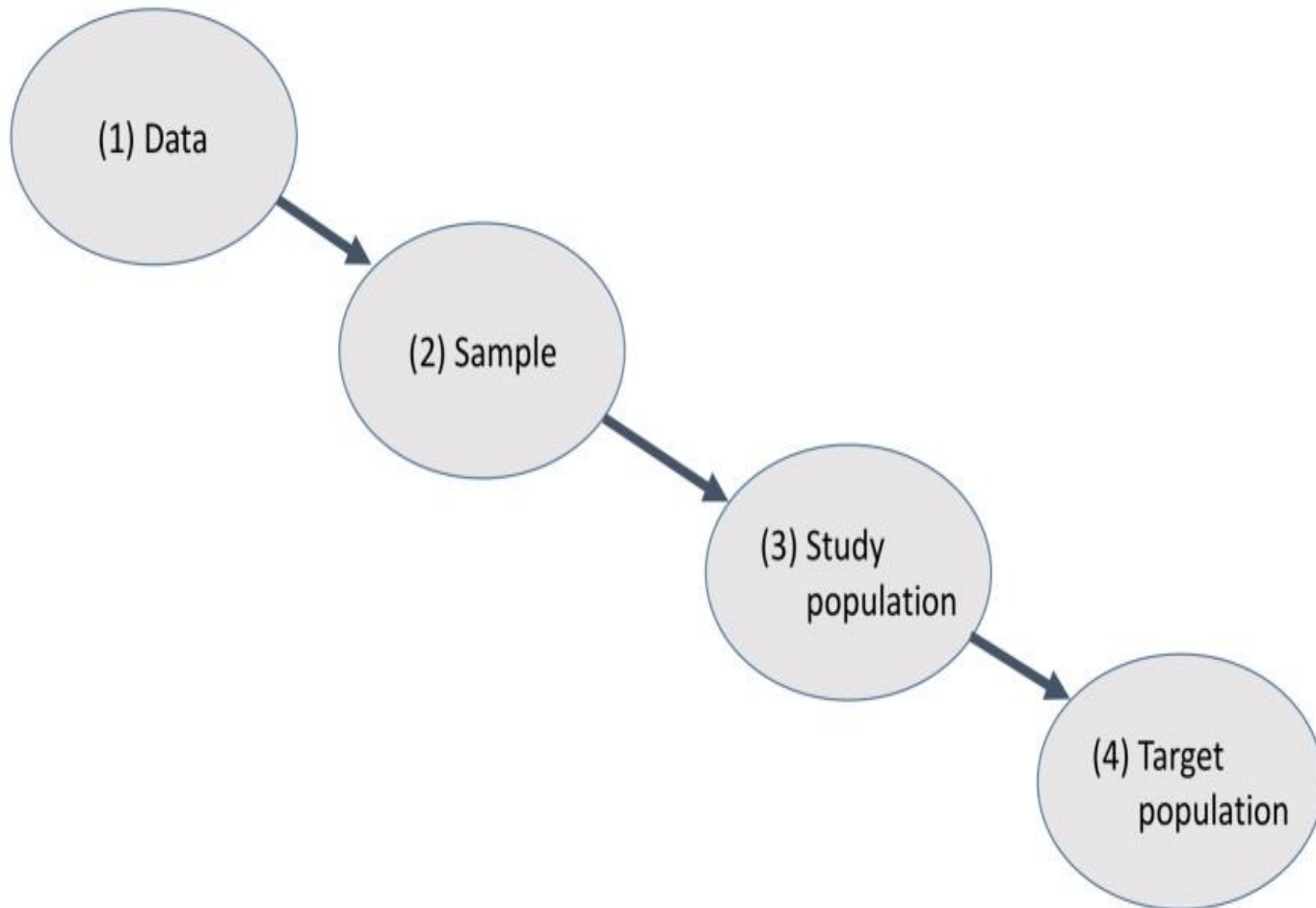
- ***1 to 2.*** How reliable are the reports?
- *Poor memory, social acceptability bias etc*

- ***2 to 3.*** How representative is the sample of those eligible for the study?
- *Random sampling of families (soup), 66% response*

- ***3 to 4.*** How close does the study population match the target population?
- *No people in institutions, etc*

Causation (or correlation)
The power of the press release....

Socioeconomic position and the risk of brain tumour: a Swedish national population-based cohort study

Amal R Khanolkar,[1,2] Rickard Ljung,[2] Mats Talbäck,[2] Hannah L Brooke,[2] Sofia Carlsson,[2] Tiit Mathiesen,[3] Maria Feychting[2]

- abstract:
  - *We observed consistent associations between higher socio-economic position and higher risk of glioma*

- press release
  - *High levels of education linked to heightened brain tumour risk*

- Daily Mirror…

EU REFERENDUM
Latest news, opinion, polls & analysis >

Mirror

Our new FREE apps are here

News · Politics · Football · Sport · Celebs · TV & Film · Weird News

Most read · Videos

TRENDING · IPHONE SE · APPLE · GOOGLE · FACEBOOK · WHATSAPP · VIDEO GAMES

Technology · Money · Travel · Fashion

M › Science › tumour

Why going to university increases risk of getting a brain tumour

23:30, 20 JUN 2016 · BY ANDREW GREGORY

Highly educated people are more likely to suffer from brain tumours than those who do not progress as far in their education

13 SHARES · 1 COMMENT

Scientists might even have an agenda...

APOCALYPSE WARNING: Insect population must be saved or 'LIFE WILL DISAPPEAR'

A DECLINE in insect populations across the globe is a major "CRISIS" and it could spell the "disappearance of the foundation of life on Earth", researchers have warned.

By SEAN MARTIN

Insects

Plummeting insect numbers 'threaten collapse of nature'

Damian Carrington
Environment editor

🐦 @dpcarrington
Sun 10 Feb 2019 18.00 GMT

The world's insects are hurtling down the path to extinction, threatening a "catastrophic collapse of nature's ecosystems", according to the first global scientific review.

More than 40% of insect species are declining and a third are endangered, the analysis found. The rate of extinction is eight times faster than that of mammals, birds and reptiles. The total mass of insects is falling by a precipitous 2.5% a year, according to the best data available, suggesting they could vanish within a century.

Review

Worldwide decline of the entomofauna: A review of its drivers

Francisco Sánchez-Bayo [a] ✉, Kris A.G. Wyckhuys [b, c, d]

⊞ Show more

Get rights and content

- *we performed a search on the online Web of Science database using the keywords [insect*] AND [declin*] AND [survey], which resulted in a total of 653 publications.*

Predictive analytics

Regression, prediction and algorithms
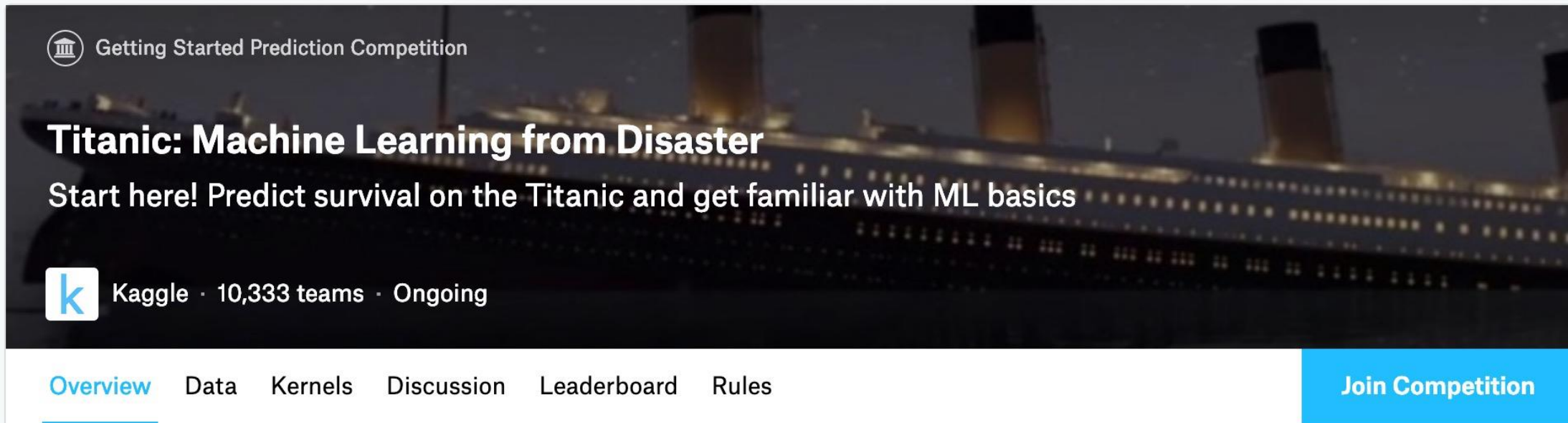*Who was the luckiest person on the Titanic?*

Ilfracombe, North Devon

ALSO OF
FRANCIS WILLIAM
SON OF THE ABOVE
WHO PERISHED
IN THE "TITANIC" DISASTER
APRIL 14, 1912, AGED 30 YEARS.
ALSO OF
HANNAH

William Somerton's entry in a public database of 1309 passengers (39% survive)

pclass	survived	name	sex	age	sibsp	parch	ticket	fare	cabin	embarked	boat	body
3	0	Somerton, Mr. Francis William	male	30	0	0	A.5. 18509	8.0500		S		
3	0	Spector, Mr. Woolf	male		0	0	A.5. 3236	8.0500		S		
3	0	Spinner, Mr. Henry John	male	32	0	0	STON/OQ. 369943	8.0500		S		
3	0	Staneff, Mr. Ivan	male		0	0	349208	7.8958		S		
3	0	Stankovic, Mr. Ivan	male	33	0	0	349239	8.6625		C		
3	1	Stanley, Miss. Amy Zillah Elsie	female	23	0	0	CA. 2314	7.5500		S	C	
3	0	Stanley, Mr. Edward Roland	male	21	0	0	A/4 45380	8.0500		S		
3	0			60.5	0	0	2701			S		261

- Can we construct an algorithm to predict who survives?

Copy structure of Kaggle competition (currently over 59,000 entries)

Split data-base of 1309 passengers at random into
training set (70%)
test set (30%)

Unsurprising factors predict survival
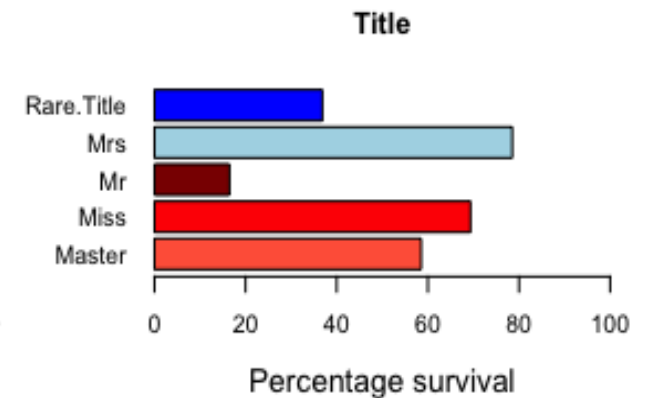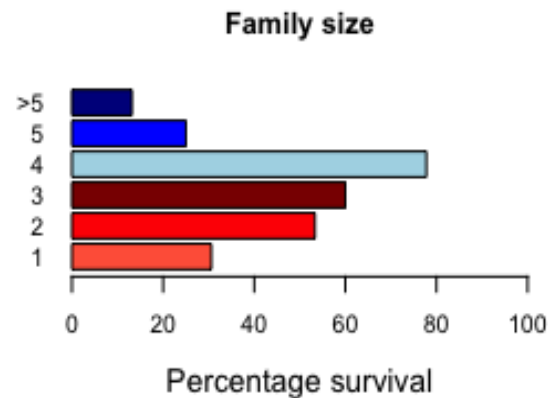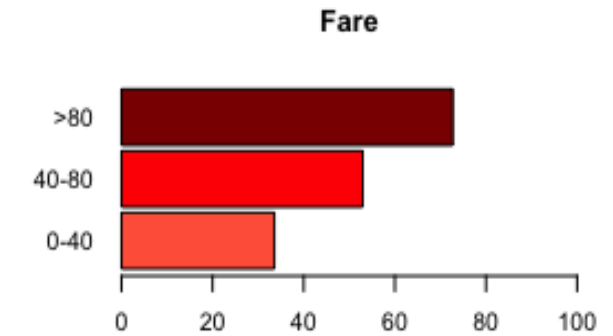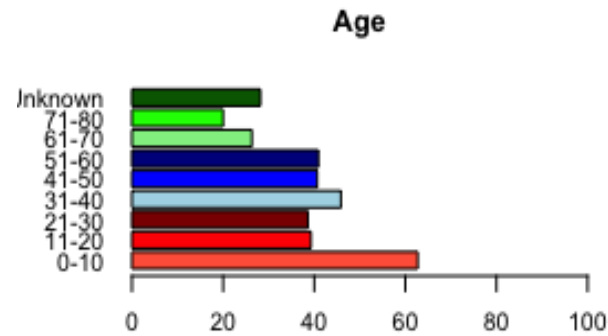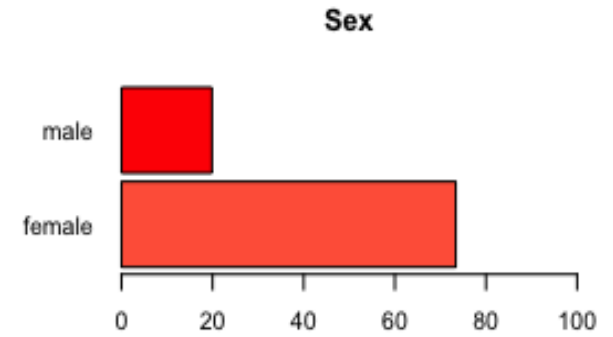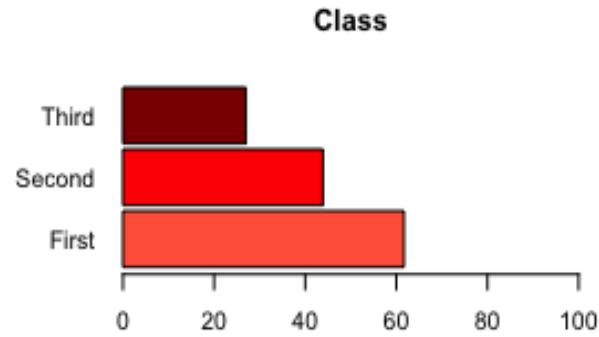
A simple classification tree
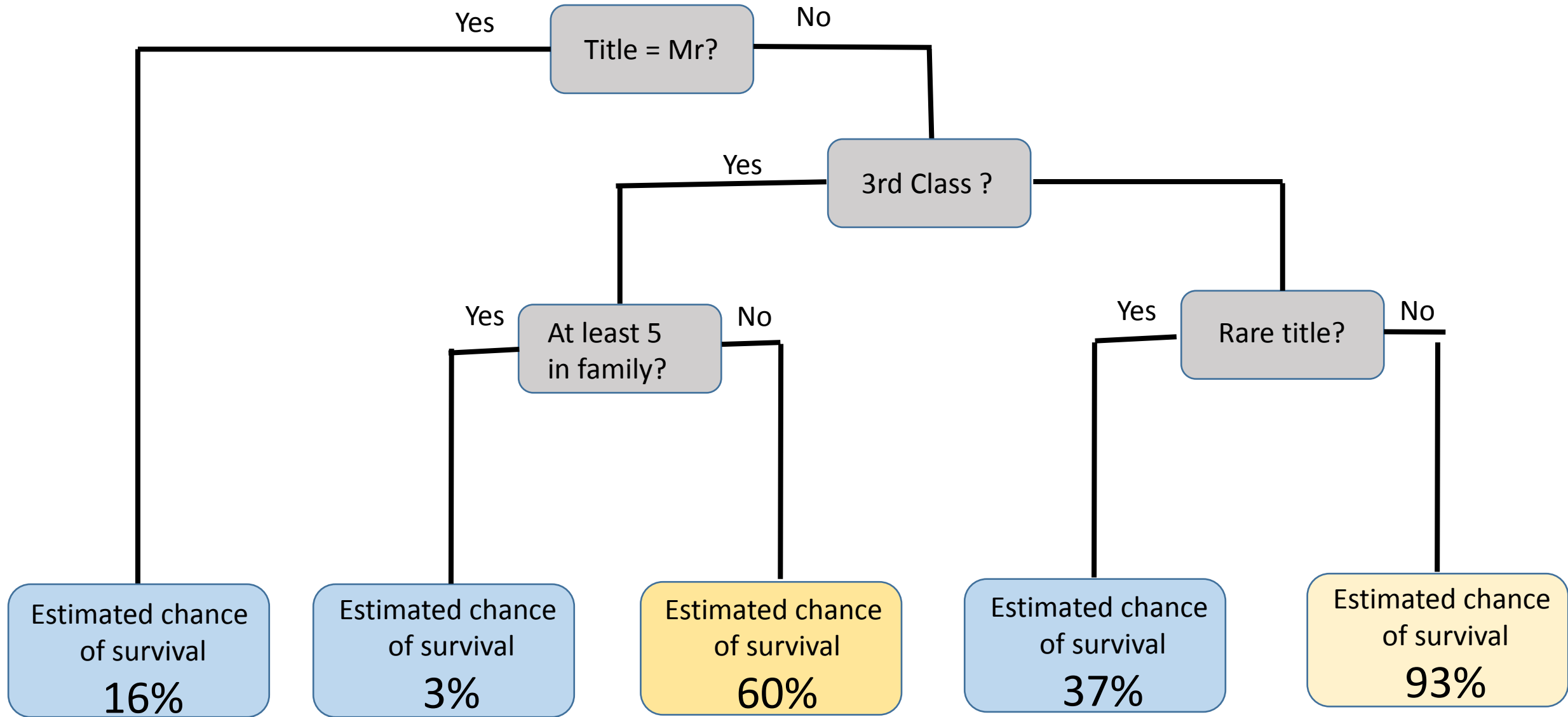
How good is my algorithm?

- 'Accuracy' is a very crude way of judging an algorithmic prediction

- Better to use the probabilities provided

- If probability p is given to an event X (0,1), then the Brier score is $(X - p)^2$

Performance of a range of methods on the test set

Method	Accuracy (high is good)	Brier score (low is good)
Everyone has a 39% chance of surviving	0.639	0.232
All females survive, all males do not	0.786	0.214
Simple classification tree	**0.806**	**0.139**
Classification tree (over-fitted)	0.806	0.150
Logistic regression	0.789	0.146
Random forest	0.799	0.148
Support Vector Machine (SVM)	0.782	0.153
Neural network	0.794	0.146
Averaged neural network	0.794	0.142
K-nearest-neighbour	0.774	0.180

Who was the luckiest person on the Titanic?

- Karl Dahl, a 45-year-old Norwegian/Australian joiner travelling on his own in third class, paid the same fare as Francis Somerton

- Had the lowest average Brier score among survivors – a very surprising survivor

- He apparently dived into the freezing water and clambered into Lifeboat 15, in spite of some on the lifeboat trying to push him back.

- Hannah Somerton was left just £5, less than Francis spent on his ticket.

Statistical methods are not always used well..

February 17, 2019

Effect of a Resuscitation Strategy Targeting Peripheral Perfusion Status vs Serum Lactate Levels on 28-Day Mortality Among Patients With Septic Shock

The ANDROMEDA-SHOCK Randomized Clinical Trial

Glenn Hernández, MD, PhD[1]; Gustavo A. Ospina-Tascón, MD, PhD[2]; Lucas Petri Damiani, MSc[3];
et al

Figure 2. Kaplan-Meier Estimates of Cumulative Mortality Within 28 Days Among Patients Treated With Peripheral Perfusion–Targeted Resuscitation vs Lactate Level–Targeted Resuscitation

Hazard ratio, 0.75 (95% CI, 0.55-1.02); $P = .06$
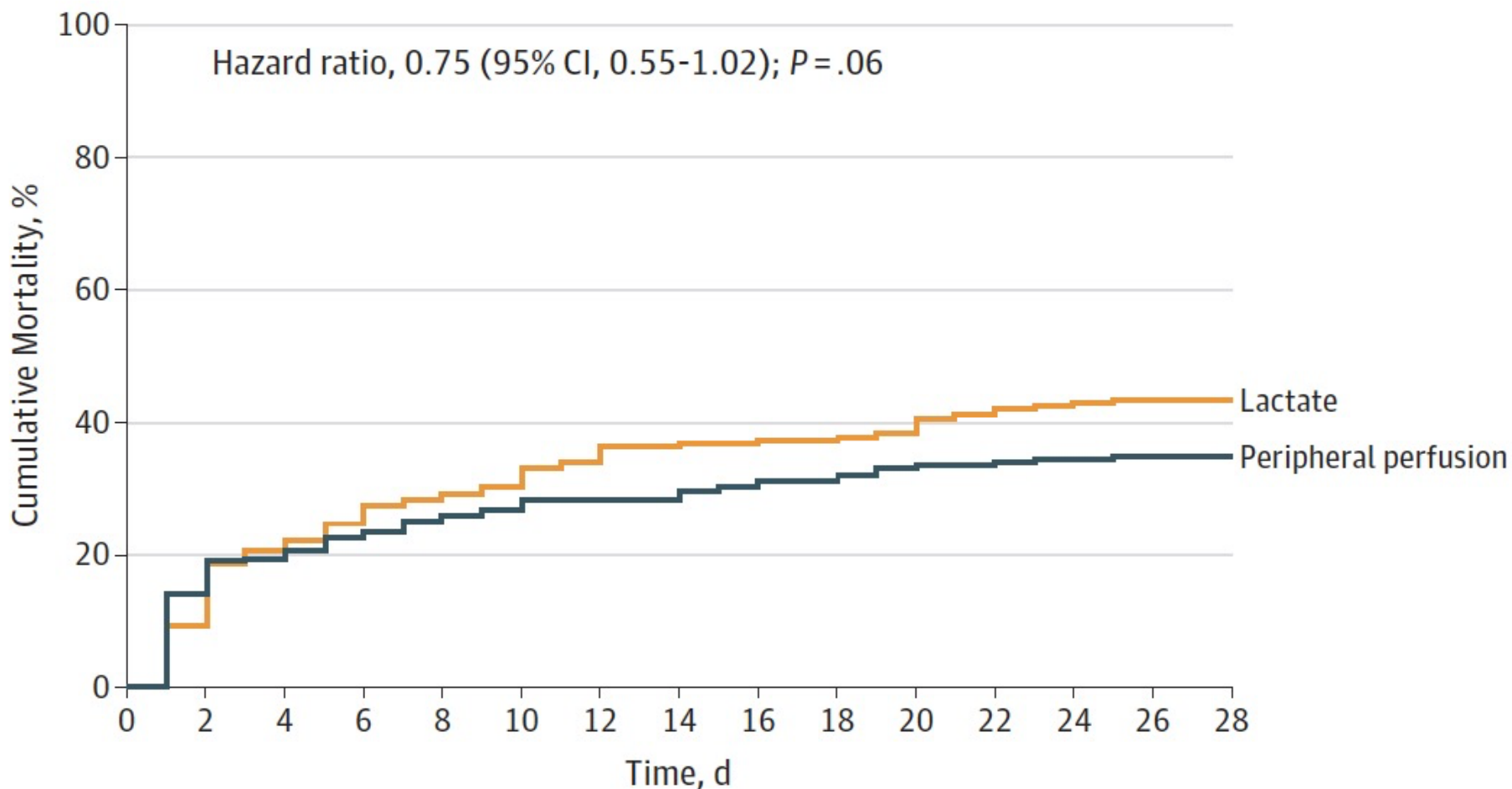
No. at risk

Lactate	212	192	168	160	152	148	140	135	134	133	130	124	122	120	120
Peripheral perfusion	212	182	171	164	159	155	152	152	148	146	142	141	139	138	138

The mysteries of the P-value...

- P-value: a measure of the conflict between the data and a 'null hypothesis' of no effect
- Specifically, P = probability of getting such an extreme result, were the null hypothesis true.
- *Not* the probability of the null hypothesis
- Traditional threshold of 5%, to declare 'statistically significant'
- Not significant does *not* mean 'no effect'
- If many tests, or crucial decision, use more stringent threshold

Rare example of accurate reporting of meaning of P-value

Science & Environment

Higgs boson-like particle discovery claimed at LHC

By Paul Rincon
Science editor, BBC News website, Geneva

4 July 2012 | Science & Environment | 1665

They claimed that by combining two data sets, they had attained a confidence level just at the "five-sigma" point - about a one-in-3.5 million chance that the signal they see would appear if there were no Higgs particle.

So what did Andromeda find?

- *'Two-sided P = 0.06'*
- i.e. the probability of observing such a big *improvement,* were there no effect, is 0.03
- Could say there is 97% confidence of improvement.
- So what is the authors' conclusion?

CONCLUSIONS AND RELEVANCE Among patients with septic shock, a resuscitation strategy targeting normalization of capillary refill time, compared with a strategy targeting serum lactate levels, did not reduce all-cause 28-day mortality.

Peripheral Perfusion Fails to Cut Septic Shock Mortality

— But showed improvements over lactate-guided resuscitation for some outcomes

by Elizabeth Hlavinka, Staff Writer, Med

February 19, 2019

2 minute medicine®
Physician Press ™

All Specialties   Cardiology   Emergency

Peripheral perfusion-targeted resuscitation does not lower mortality in septic shock patients

February 20, 2019 | Jason Nam, MD and Daniel Fisher

But just last week….

COMMENT · 20 MARCH 2019

Scientists rise up against statistical significance

Valentin Amrhein, Sander Greenland, Blake McShane and more than 800 signatories call for an end to hyped claims and the dismissal of possibly crucial effects.

- Not against P-values
- Just their dichotomisation

When might a split into 'significant' / 'not-significant' be more reasonable?

- Where a **decision** has to be made, e.g.

    - Drug regulation
    - Monitoring the performance of a list of centres/hospitals/doctors – when to intervene?

Hypothesis testing
Could Harold Shipman have been caught earlier?

- Using mortality rates from local GPs, calculate how many deaths he would have been **expected** to observe each year, under the **null hypothesis** that his mortality rates were normal.

- Subtract **expected** from **observed** number to get **excess mortality**
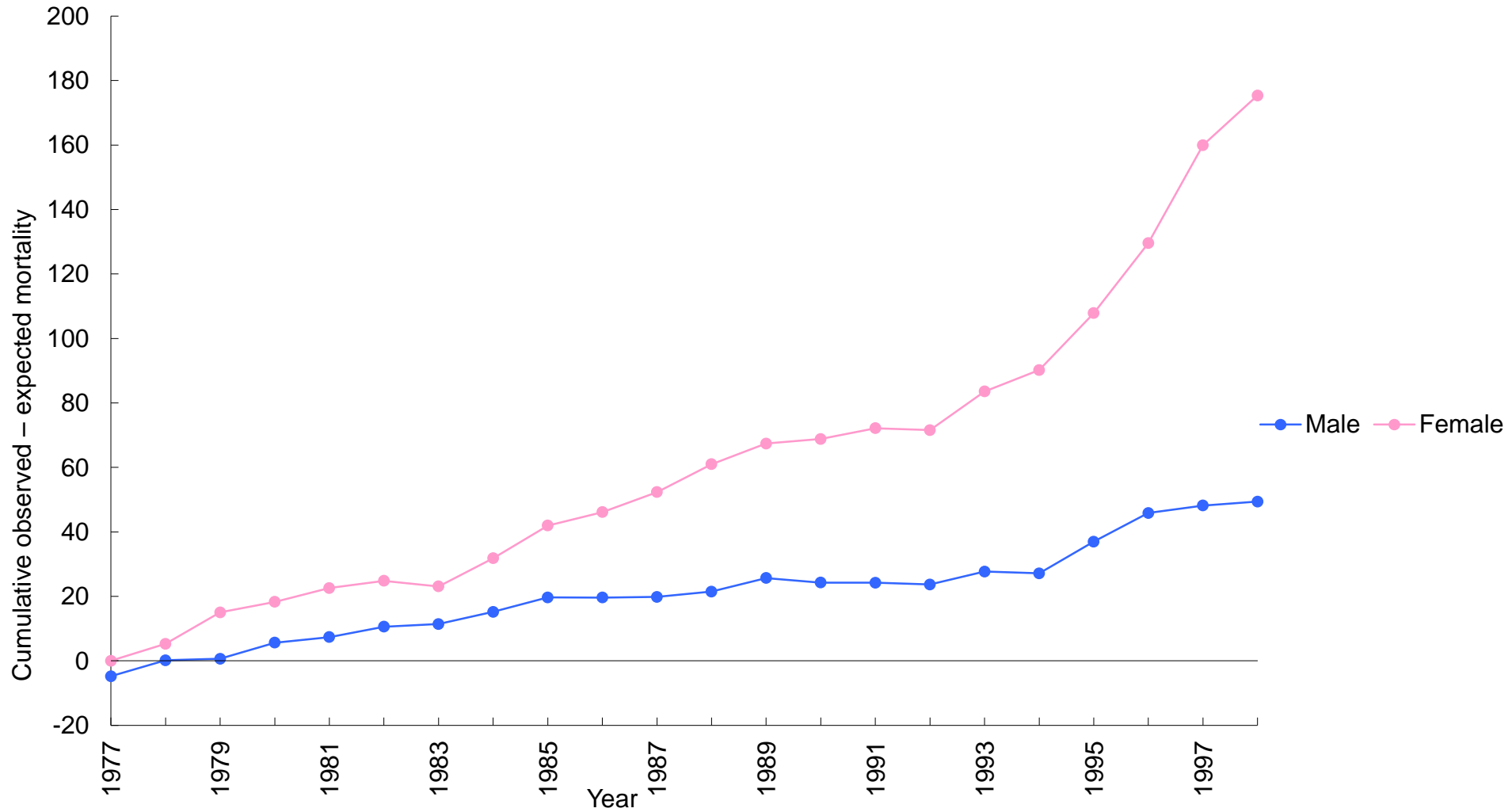
(NB: Shipman Inquiry total of definite or probable victims:
189 female > 65, **55** male over 65)

Hypothesis testing
Could Harold Shipman have been caught earlier?

- But when to 'blow the whistle'?
- This are two possible types of error -
  - **Type I error:** falsely accuse an innocent person (the null hypothesis)
  - **Type II error:** miss someone with true increased risk
- Generally, we want to
  - control the probability of a Type I error at a low value (α)
  - collect enough data to make Type II errors rare (β)

Shipman: "Sequential probability ratio test" (SPRT)
older females would have set off 'alarm' in 1985, after only 40 deaths
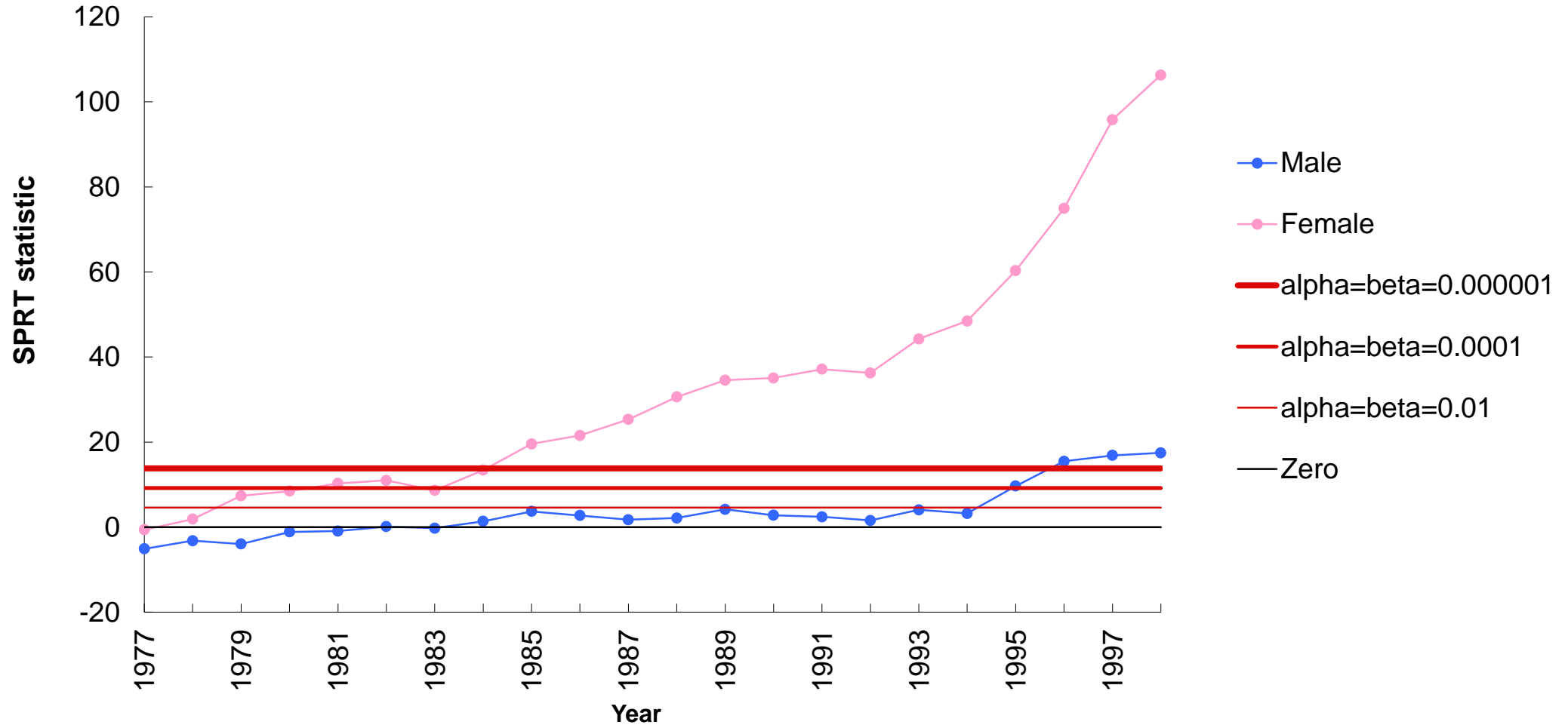
Probability and Bayes

Bayes theorem

the initial odds for a hypothesis

x   the *likelihood ratio*

=   the final odds for a hypothesis

Bayes theorem

- Suppose 1,000 possible perpetrators of a crime, plus one suspect

- The initial odds that a suspect is guilty = 1 / 1,000

- The 'likelihood ratio'

- $= \dfrac{\text{Pr(evidence | suspect guilty)}}{\text{Pr(evidence| someone else did it)}} = 1,000,000$

- After evidence is considered, final odds that a suspect is guilty

$$\dfrac{1,000}{i} \times \dfrac{1,000,000}{1} = \dfrac{1}{1,000}$$

Probability and Bayes
What is the probability that the skeleton in a Leicester car park was really Richard III?

A recent case

- On Saturday 25 August 2012, archeologists started digging in a car park in Leicester – the site of Grey Friars friary

- In a few hours they found their first skeleton

- This was later claimed to be Richard III

Identification of the remains of King Richard III

Turi E. King[1,2], Gloria Gonzalez Fortes[3,4,*], Patricia Balaresque[5,*], Mark G. Thomas[6], David Balding[6], Pierpaolo Maisano Delser[1], Rita Neumann[1], Walther Parson[7,8], Michael Knapp[9], Susan Walsh[10,11], Laure Tonasso[5], John Holt[12], Manfred Kayser[11], Jo Appleby[2], Peter Forster[13,14], David Ekserdjian[15], Michael Hofreiter[3,4] & Kevin Schürer[16]

$$\text{Likelihood ratio} = \frac{\text{probability of evidence, if skeleton were Richard III}}{\text{probability of evidence, if someone else}}$$

Suggested 'verbal equivalents' for bands of likelihood ratios

Value of likelihood ratio	Verbal equivalent
>1–10	Weak support for proposition
10–100	Moderate support
100–1000	Moderately strong support
1000–10,000	Strong support
10,000–1,000,000	Very strong
>1,000,000	Extremely strong

Standards for the formulation of evaluative forensic science expert opinion

Association of Forensic Science Providers

Evidence	Likelihood ratio (conservative estimate)	Verbal equivalent
Radiocarbon dating AD 1456–1530	2	Weak support
Age and sex of skeleton	5	Weak support
Scoliosis	212	Moderately strong support
Post-mortem wounds	42	Moderate support
mtDNA match	478	Moderately strong support
Y chromosome not matching	0.2	Weak evidence against
Combined evidence	6.5 million	More than extremely strong support

Researchers claimed at least 0.999994 probability that they had found Richard III

When communication goes wrong..

SEX

BY
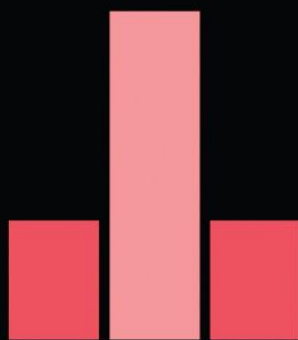## NUMBERS

The Statistics of Sexual Behaviour

DAVID SPIEGELHALTER

How often do (opposite sex) couples report having sex?

Median (middle) number of occasions of sex in the past 4 weeks (people aged 16–44)

Natsal–1	2 — 5 — 9
Natsal–2	2 — 4 — 9
Natsal–3	1 — 3 — 7

1990
2000
2010

The bars represent the interquartile range: 50% of the population were in this range, 25% were below the lower value, and 25% were above the higher value.

Britons having sex less often

National Survey of Sexual Attitudes and Lifestyles finds frequency of sex has fallen by 20% since last survey in 2000

When I said all this in a talk….

Britons are having less sex, and Game of Thrones could be to blame, warns Cambridge professor

David Spiegelhalter, the Winton Professor of the Public Understanding of Risk, said the trend in declining sex rates over the last 30 years was 'very worrying.'

In 1990 couples had sex around five times a month, but now it is just three times, a 40 per cent decrease in just 20 years. If currents trends continue couples would not be having sex at all by 2030.

CULTURE

IS 'GAME OF THRONES' RUINING OUR SEX LIVES?

BY **TUFAYEL AHMED** ON 6/6/16 AT 1:42 PM

Most read | **What's on** ▾ | **News** ▾ | **Sport** ▾ | **Traffic & Travel** | **Property** | **Fun Stuff** | **Lifestyle**

TRENDING ▸ CRIME | LUAS | THINGS TO DO DUBLIN   Soccer   GAA   Music and Nightlife   Things to Do

D » News » Dublin News » Television

Couples 'will stop having sex by 2030' due to the large rise in TV ratings

11:55, 6 JUN 2016 | BY JOHN PATRICK-KIERANS

Sex Will Be Obsolete by 2030 Because of Netflix, According to One Lone Scientist

Apparently, the future holds no chill.

BY KORIN MILLER, *June 7, 2016*

stern

SPIELE . SCHLAGZEILEN . THEMEN . ABO & SHOP . APPS & EMAGS . TOOLS . FOTOGRAFIE . TV

Home . Panorama . Politik . Kultur . Digital . Lifestyle . Wirtschaft . Sport . Familie . Genuss . Gesundheit . Reise . Auto . Video .

Home > Digital > Online > Game of Thrones: Warum wir wegen der Serie weniger Sex haben sollen

Von wegen heiß

02. Juni 2016 15:59 Uhr

Warum wir wegen "Game of Thrones" alle weniger Sex haben

Wo man auch hinschaut, findet man heute Sex, nur in den Schlafzimmern wird er immer weniger. Ein Statistikprofessor gibt daran nun ausgerechnet der heißen Serie "Game of Thrones" die Schuld.

Drucken

Di·Lei

Digitale, singolare, femminile

Il sesso sparirà, nel 2030 tutti in bianco

Lo dice la scienza e la colpa è tutta della tecnologia, lo studio arriva da Cambridge

f Condividi 25   Tweet   P   G+

Dai mitici anni settanta tutti sesso, droga e rock and roll a un futuro senza sesso, una generazione triste quella futura. E le dinamiche di coppia e i giochi di potere e il toy boy e il vechio con la Lolita? Sarà un mondo allo sfascio quello futuro, senza punti di riferimento, senza trasgressioni, senza tradimenti e fughe d'amore, perché il sesso aiuta a vivere.

L'allarme arriva dall'Univeristà Cambridge dove un esperto di statistica il professor **David Spiegelhalter**, ha realizzato che gli inglesi fanno sesso il 40% in meno rispetto a quanto ne facevano 20 anni fa. Secondo

EN 2030, ON NE FERA PLUS L'AMOUR

▶ 🔊 0:08 ━━━━━●━━━━━━━━━━━━━━━ 1:02 ⚙ ⛶

Sexe : d'après une étude scientifique, nous ne ferons plus l'amour en 2030

Le chercheur britannique David Spiegelhalter l'affirme dans une nouvelle étude : selon lui, nous ne ferons plus l'amour en 2030. Voici pourquoi.

NoticiasDeAcá
@TwDeACA

👨‍👩‍👧 👨‍👩‍👧‍👦 👨‍👩‍👧 Un trabajo de la Universidad de Cambridge, liderado por el experto en parejas David Spiegelhalter, alertó de que si las estadísticas continúan a ese ritmo, para el 2030 las parejas podrían no tener sexo en absoluto.

#NoMoreSex #VideosDeACÁ
facebook.com/noticiasdeaca/...

Translate Tweet

> **Noticias de Aca**
> facebook.com

9:30pm · 26 Mar 2019 · Facebook

@MaîtreMuabila G.
@MMGlody

D'après le chercheur britannique David Spiegelhalter, d'ici à 2030, nous ne ferons plus l'amour. ''Profitons-en tant qu'il est encore temps'' dit-il.
#ScienceTwitter #science #santépublique #OMS

Translate Tweet
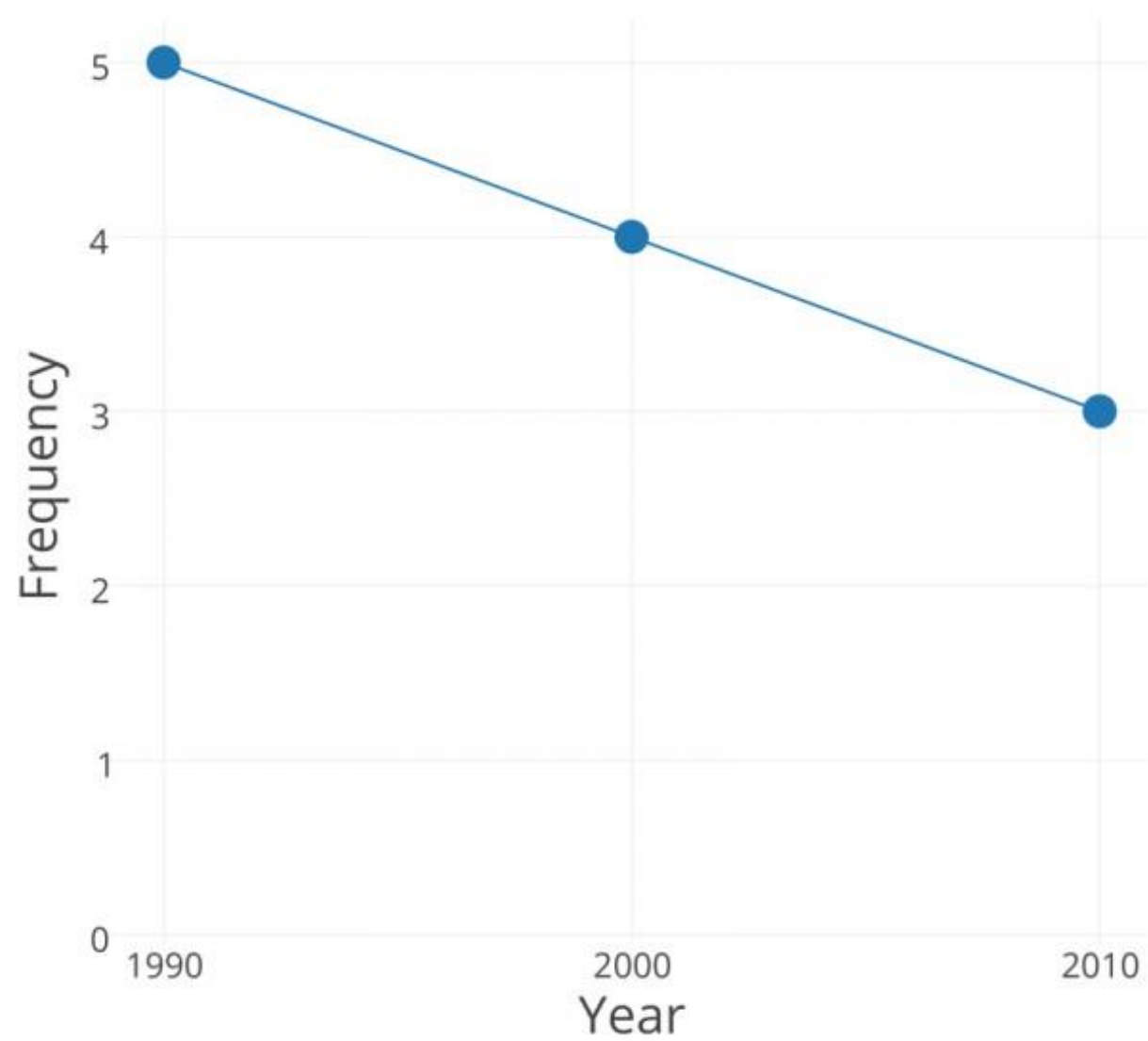
12:13pm · 26 Mar 2019 · Twitter for Android

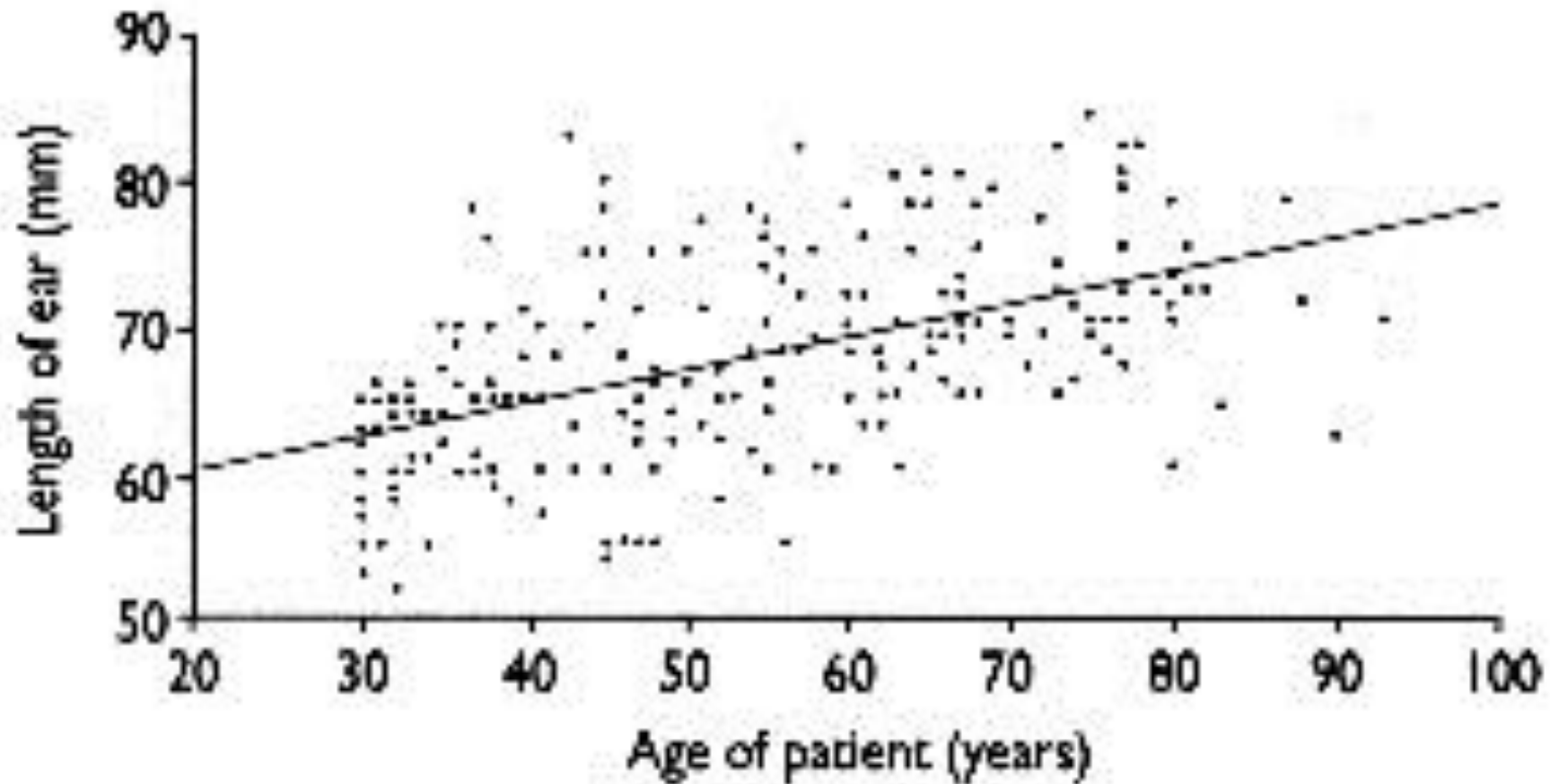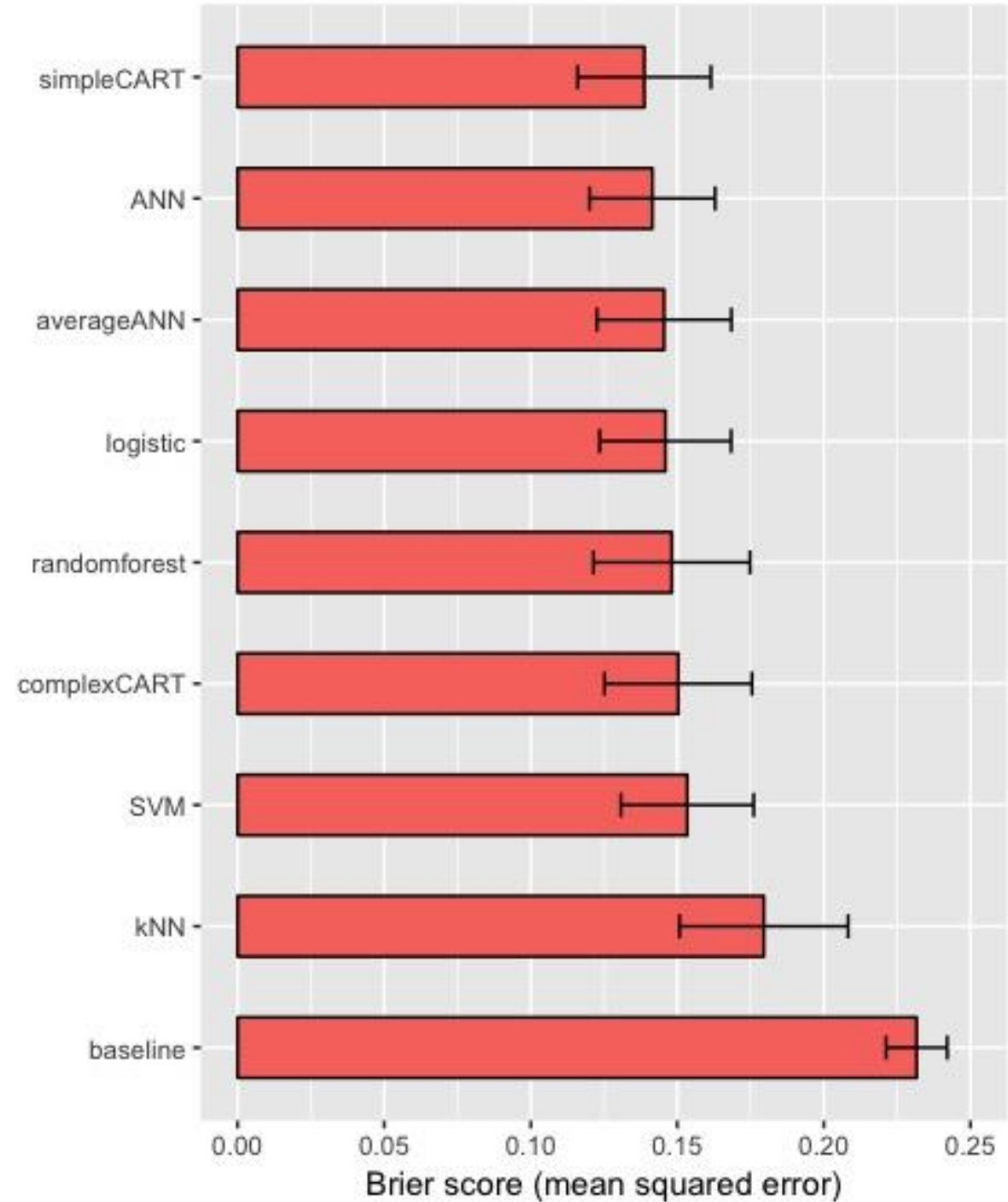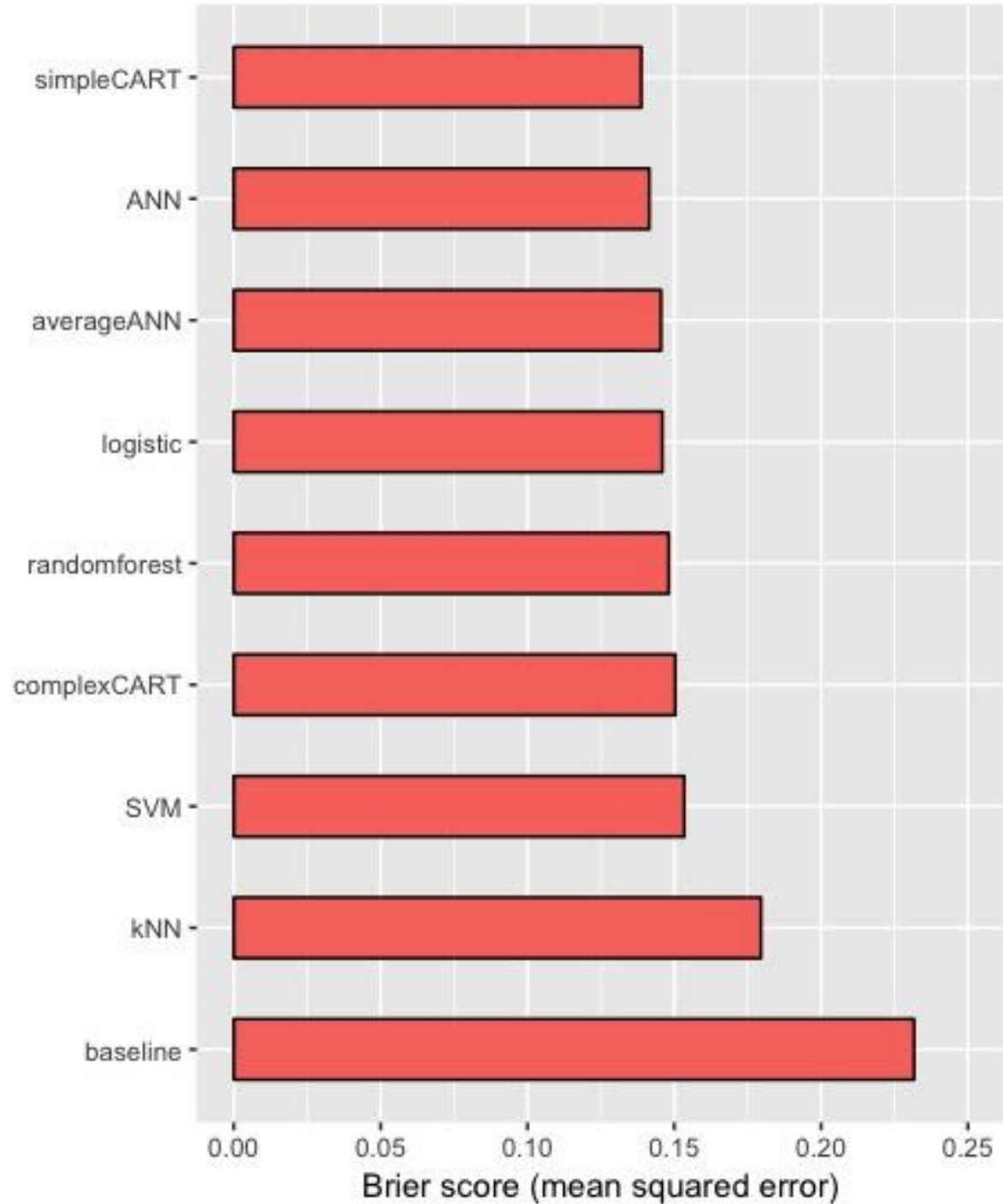# Is 'Game Of Thrones' Killing Your Sex Life?

JV Chamary, CONTRIBUTOR

I cover the science of life. **FULL BIO** ⌄

Number of times the average person had sex in the past 4 weeks

Why do old men have big ears?

Why do old men have big ears? BMJ 1995; 311 : 1668 (Published 23 December 1995)

- Potentially a very misleading graphic!
- When comparing, need to acknowledge that tested on same cases
- Calculate differences and their standard error
- How confident can we be that simple CART is best algorithm?

Excess Brier score (mean squared error)