CAUSALITY
& CORRELATION

How are continuous variables related
and are these relations causal?

LEUPHANA
UNIVERSITÄT LÜNEBURG

CAUSALITY OR CORRELATION

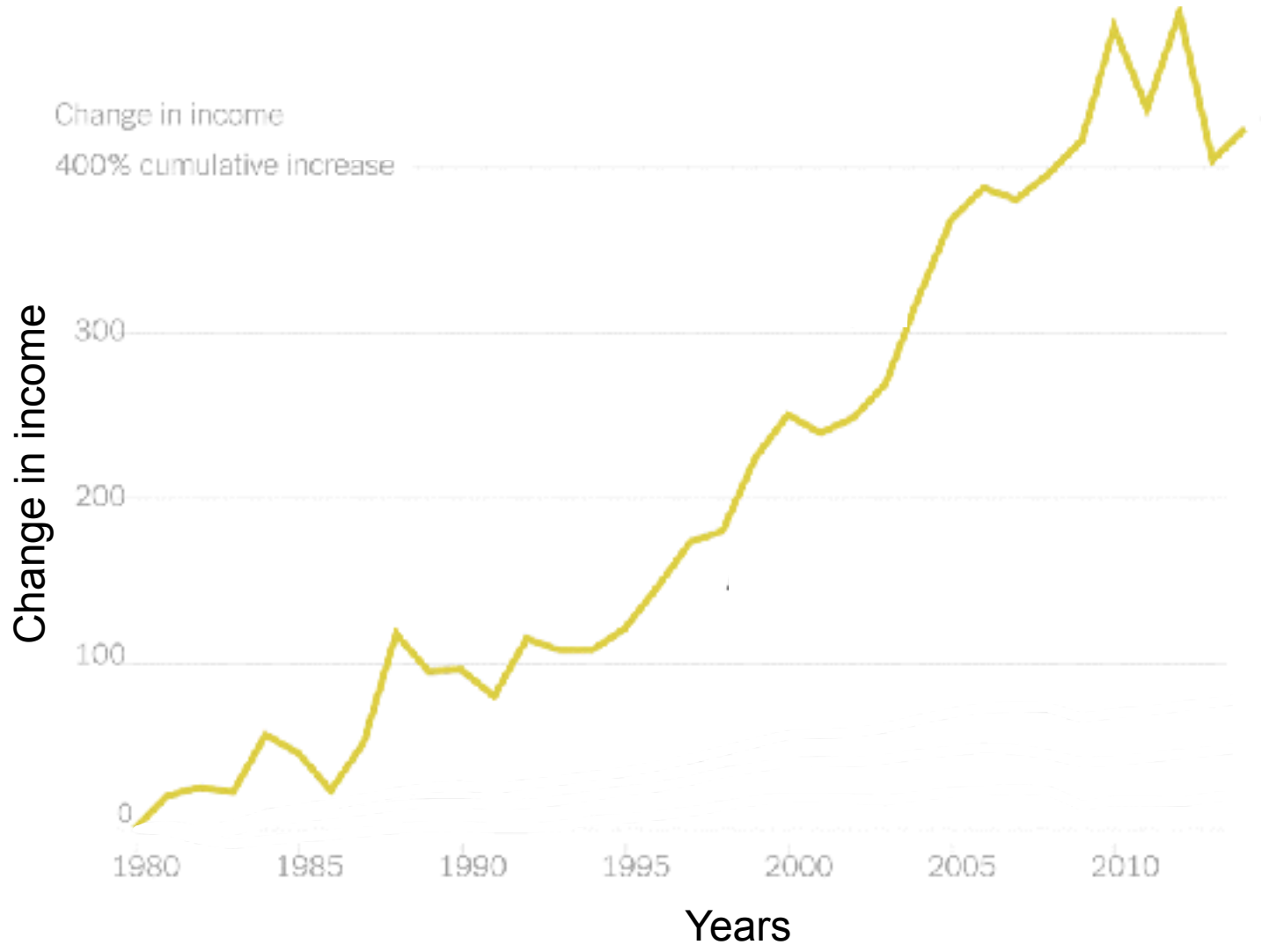
Are relations really causal?
And does this matter?

→ **Henrik von Wehrden - Methods I - V**

LEUPHANA
UNIVERSITÄT LÜNEBURG

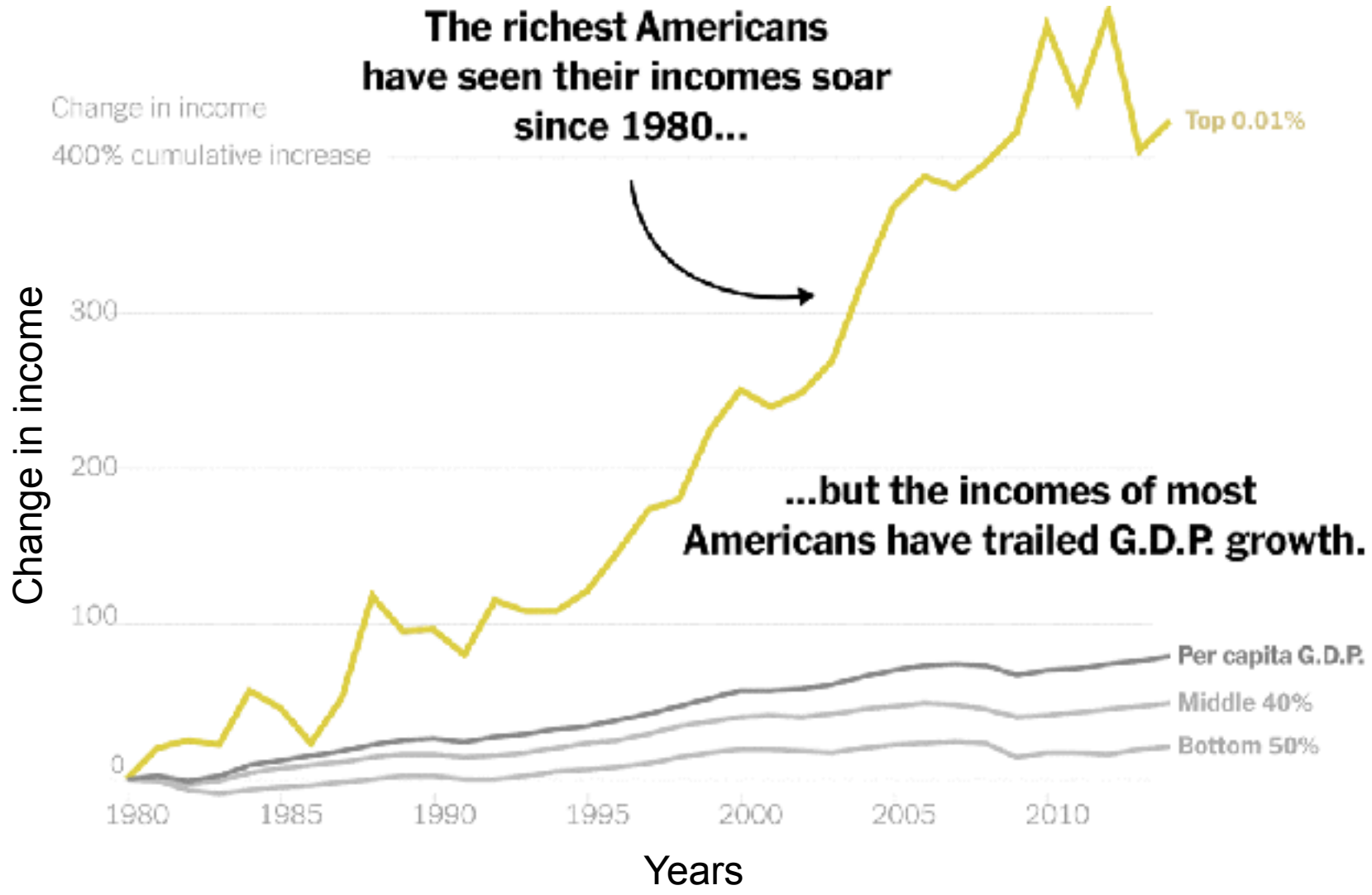
LET US START SIMPLE: STATISTICAL CORRELATIONS

How are two continuous variables in relation with each other?



LET US START SIMPLE: STATISTICAL CORRELATIONS

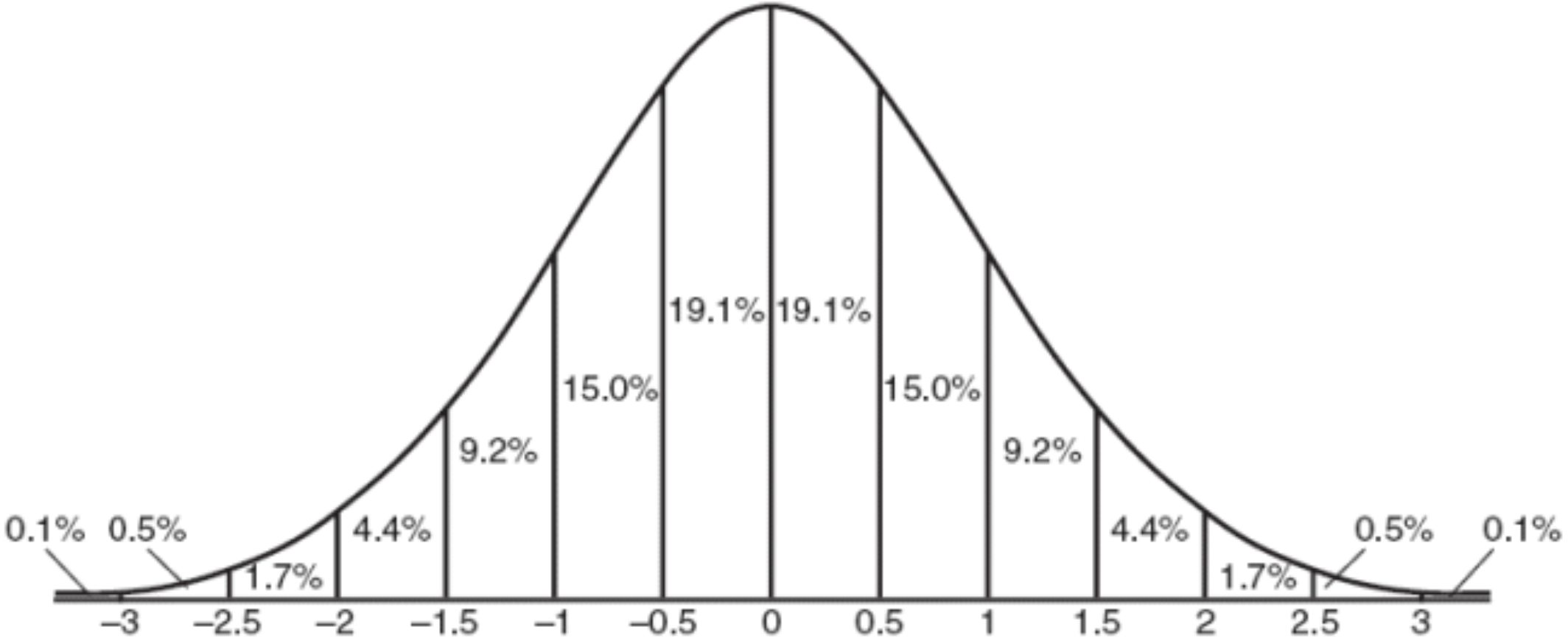
How are two continuous variables in relation with each other?



HISTORY OF CORRELATIONS

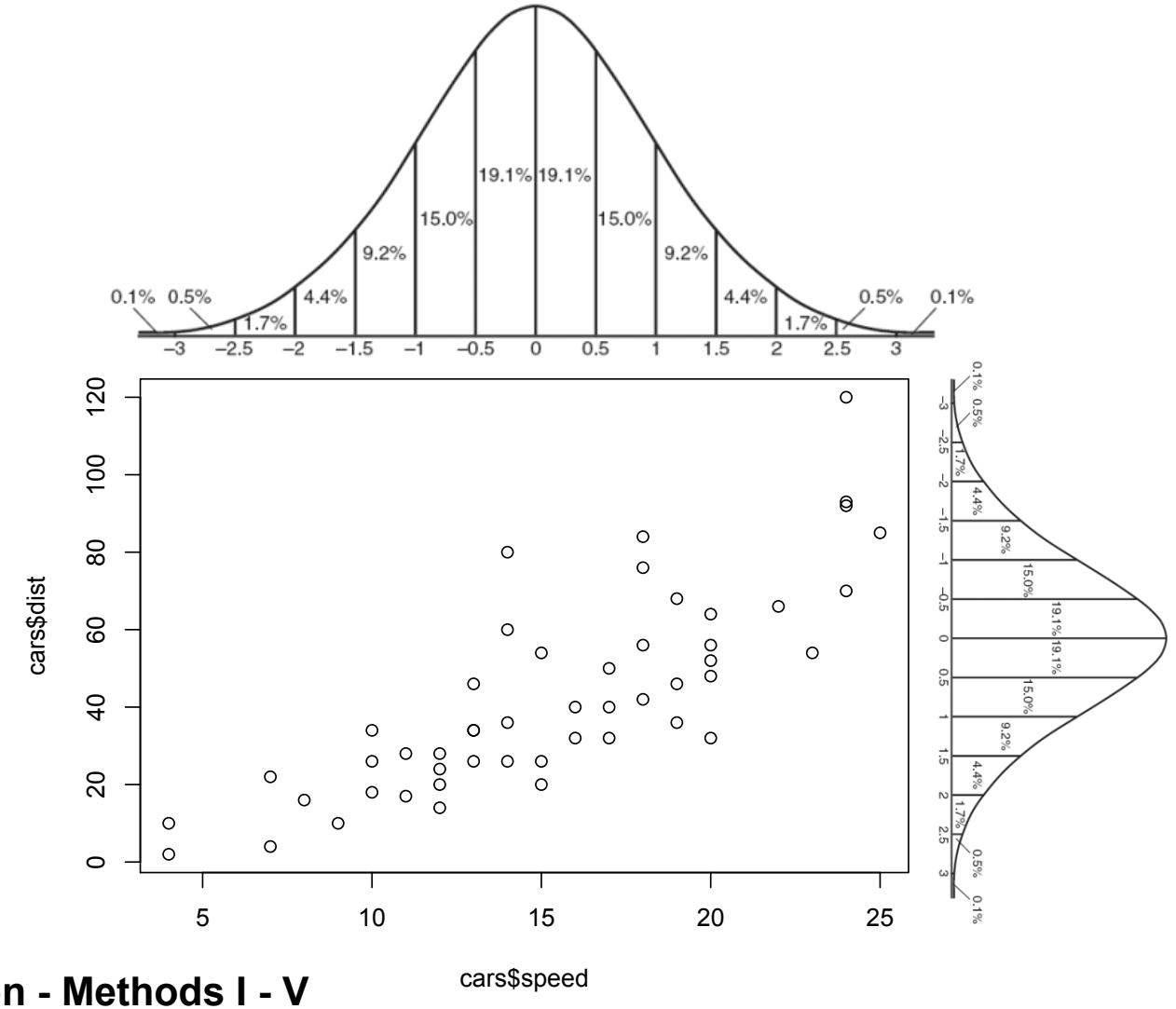
- Early focus by Pearson widely theoretical
- Fisher unraveled linear relations
- Explanatory measures have no real threshold
- Statistical power is highly relevant
- Shift from hard to soft science

THE NORMAL DISTRIBUTION

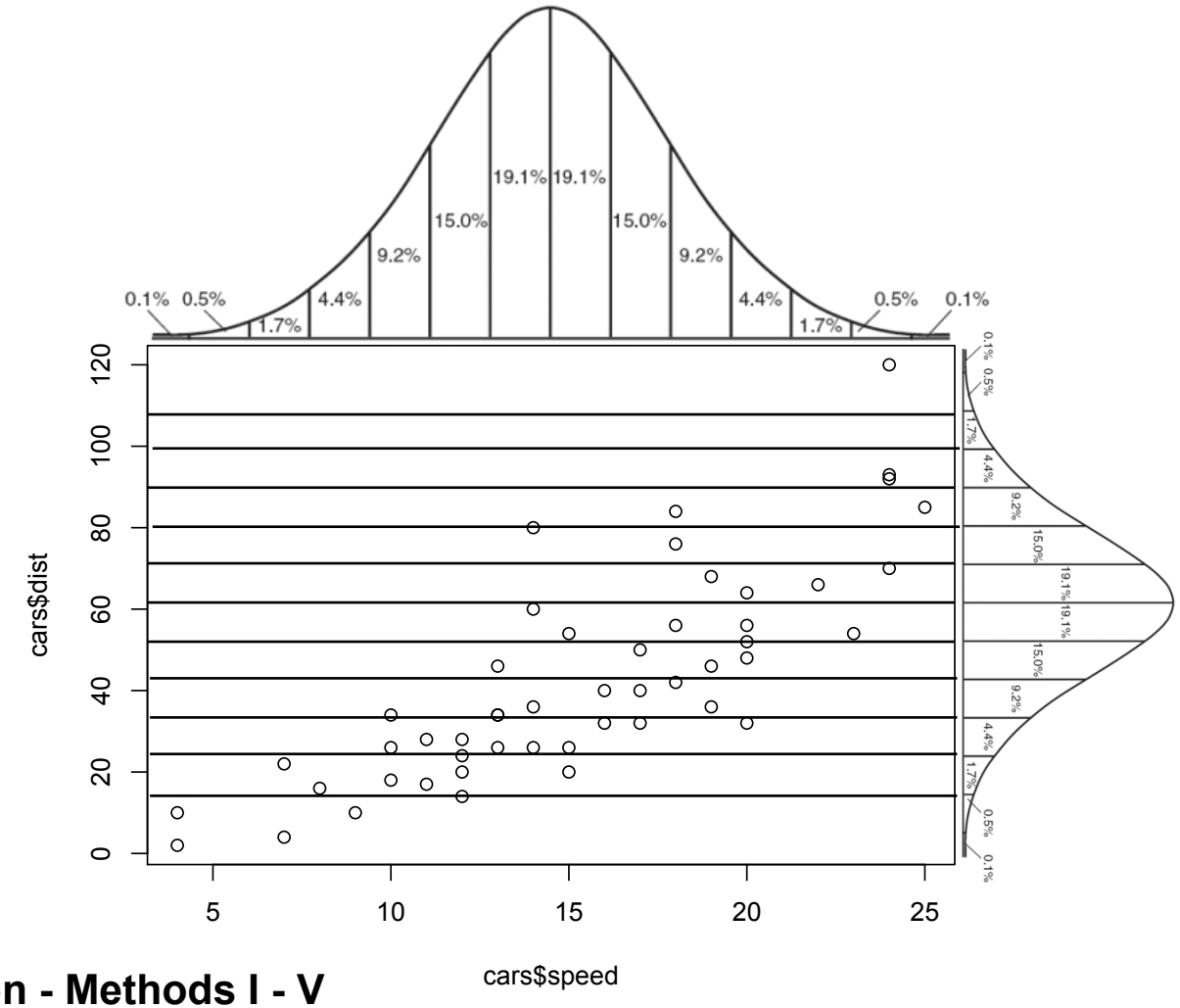


LET US START SIMPLE: STATISTICAL CORRELATIONS

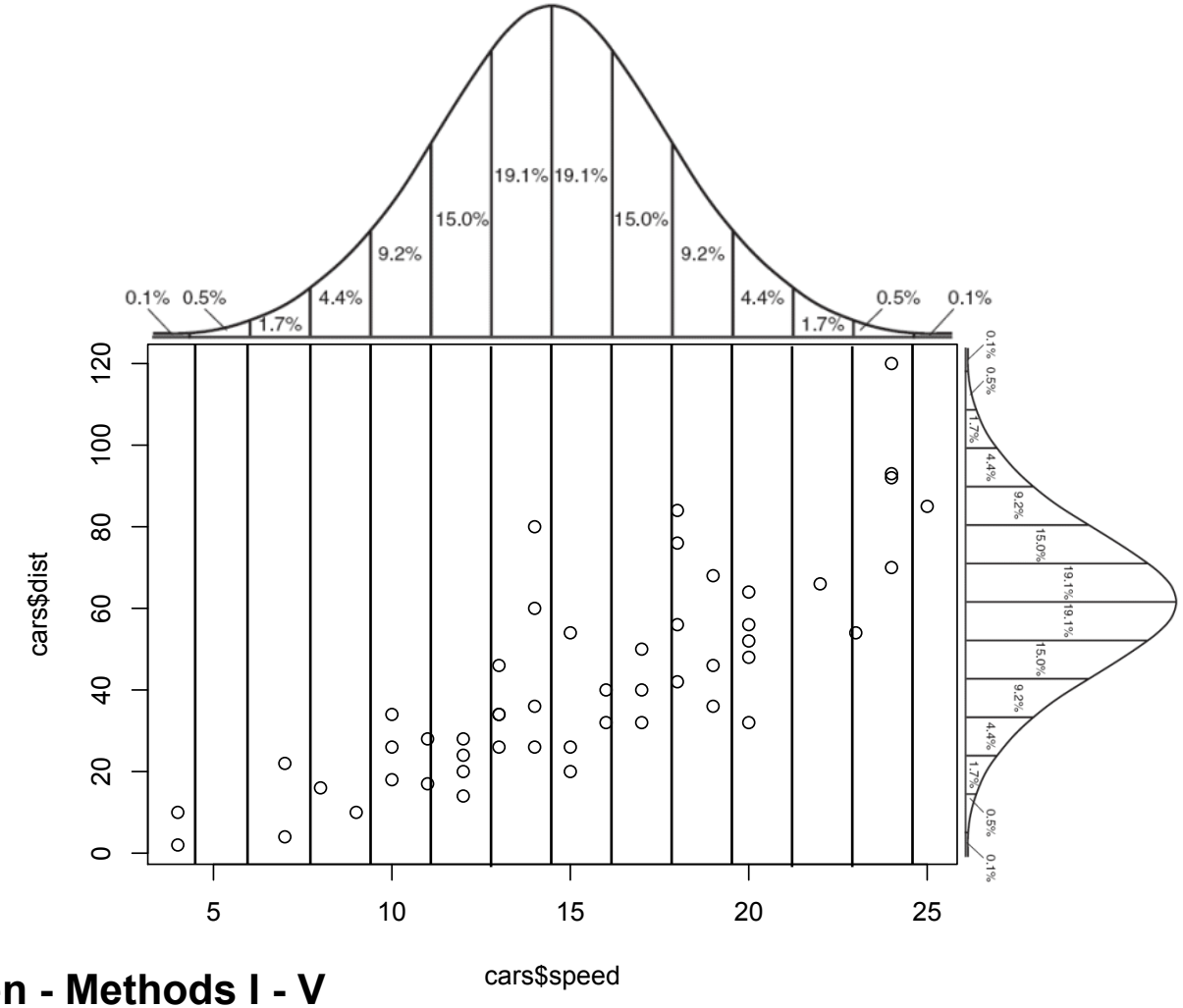
$$y \sim a + b * x$$



LET US START SIMPLE: STATISTICAL CORRELATIONS



LET US START SIMPLE: STATISTICAL CORRELATIONS



PROBABILITY - STATISTICS

The lady tasting tea

What is the probability that the lady guesses correctly if first the milk of first the tea was put into the cup?

What is the probability?

What is the probability?

What is the probability?

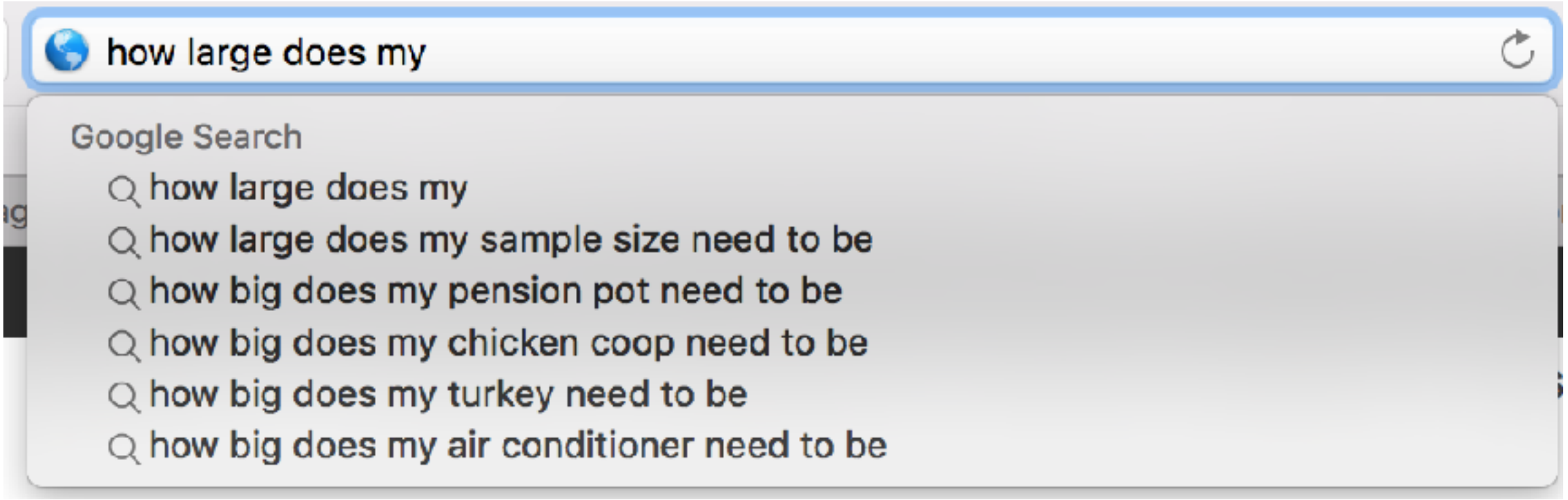
What is the probability?

What is the probability?

What is the probability?

What is the probability?

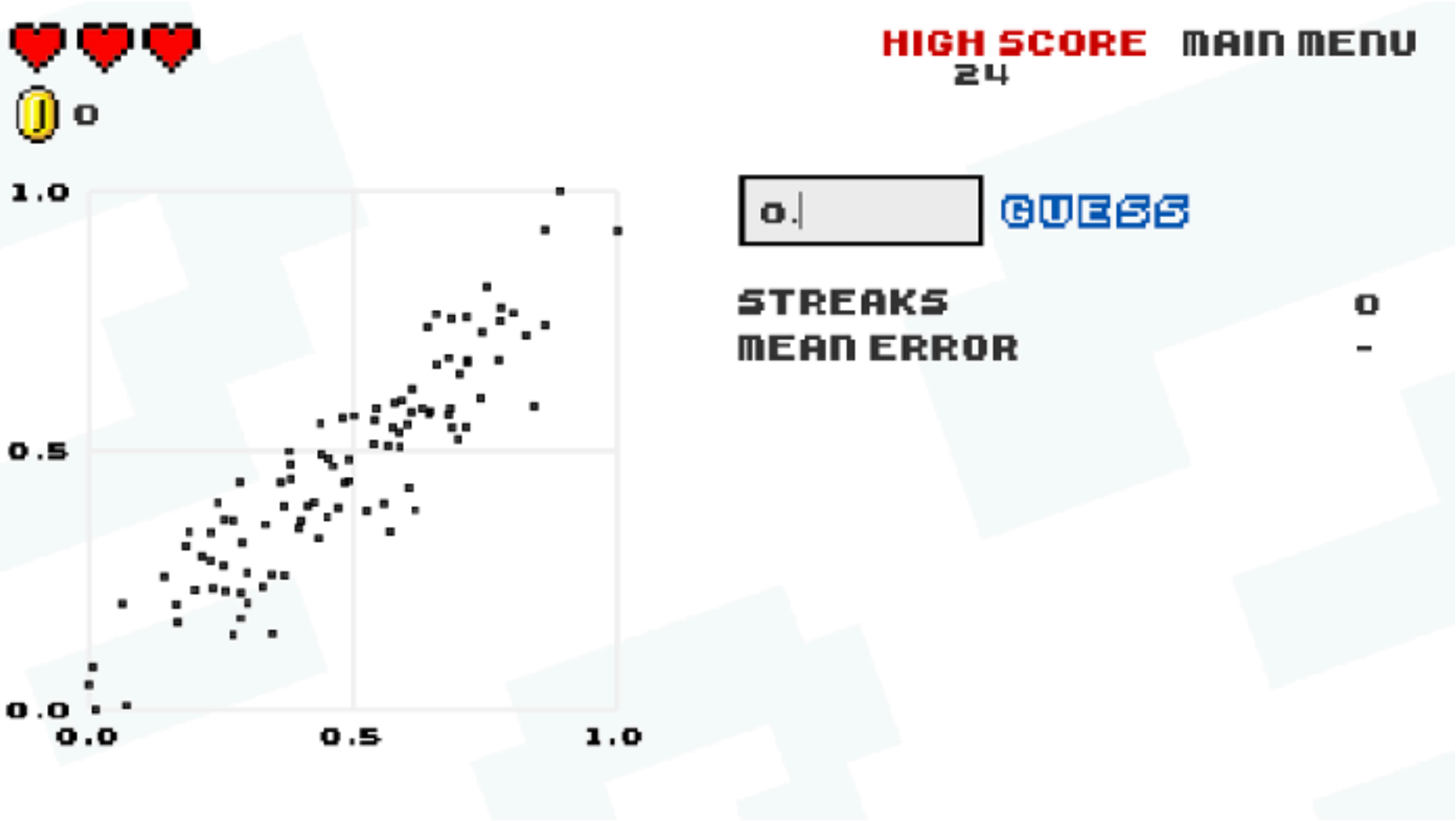
COUNTING AND SIZES MATTERS



HOW MUCH DOES A MODEL EXPLAIN

http://guessthecorrelation.com

HOW MUCH DOES A MODEL EXPLAIN?



http://guessthecorrelation.com

STATISTICAL CORRELATIONS: ONE VARIABLE RELATES ON ONE VARIABLE.

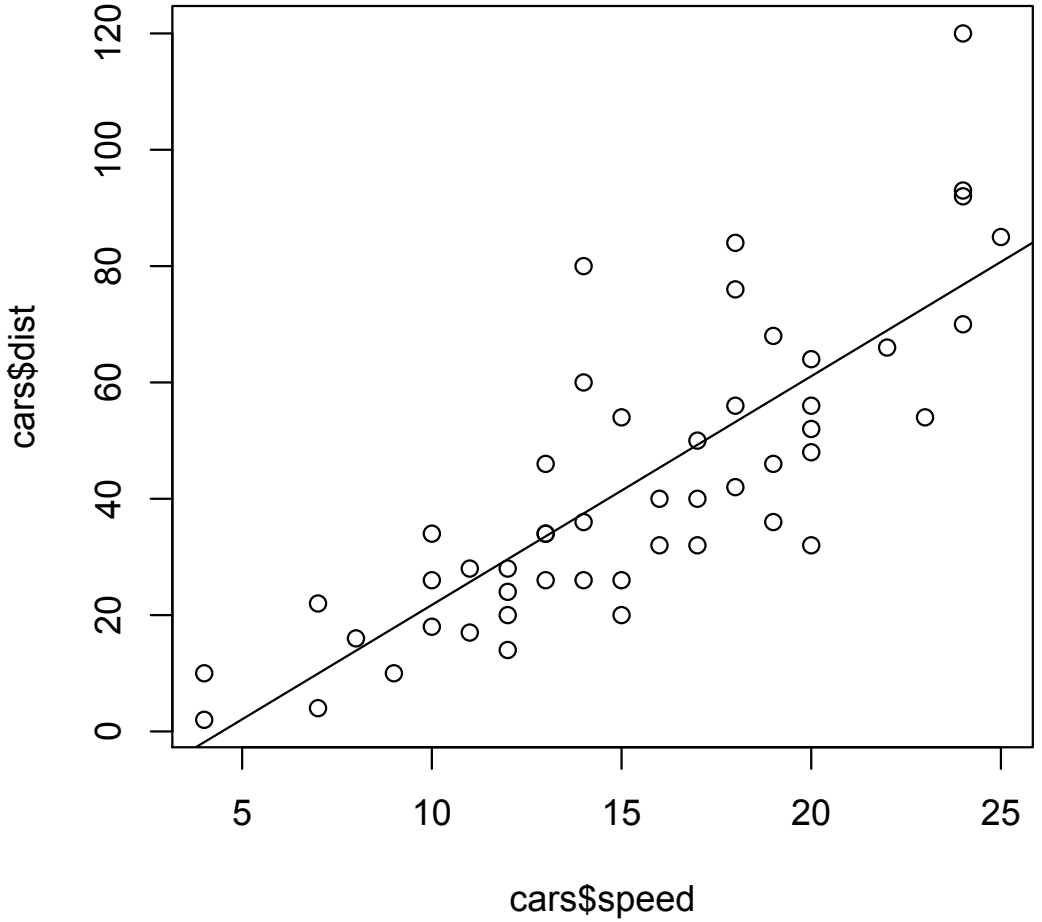
```
> cor.test(cars$dist,cars$speed)

        Pearson's product-moment correlation

data:  cars$dist and cars$speed
t = 9.464, df = 48, p-value = 1.49e-12
alternative hypothesis: true correlation is not equal
0
95 percent confidence interval:
 0.6816422 0.8862036
sample estimates:
      cor
0.8068949
```

P-VALUE:
0.000000000000149

CORRELATION COEFFICIENT
RANGES FROM -1 TO 1

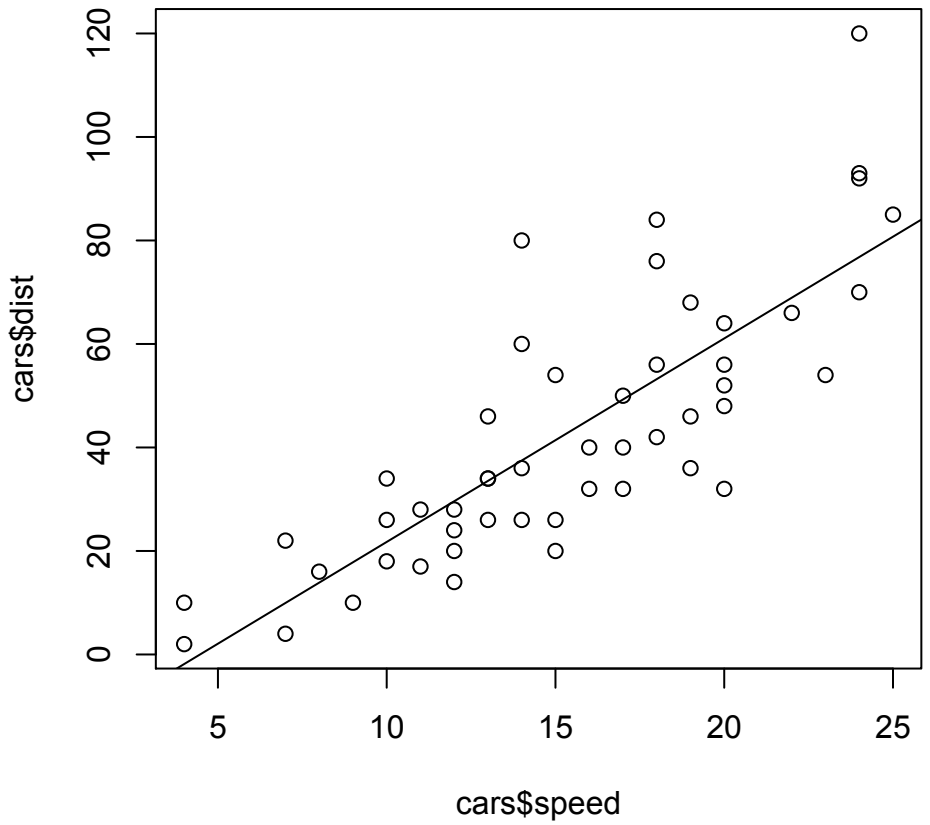
AND NOW THE REGRESSION: ONE VARIABLE IS EXPLAINED BY ANOTHER VARIABLE, AND HENCE DEPENDS ON IT

$y \sim a + b*x$



Breaking distance depends on the speed!

PREDICTIONS: SOMETIMES OUTSIDE OF THE DATA

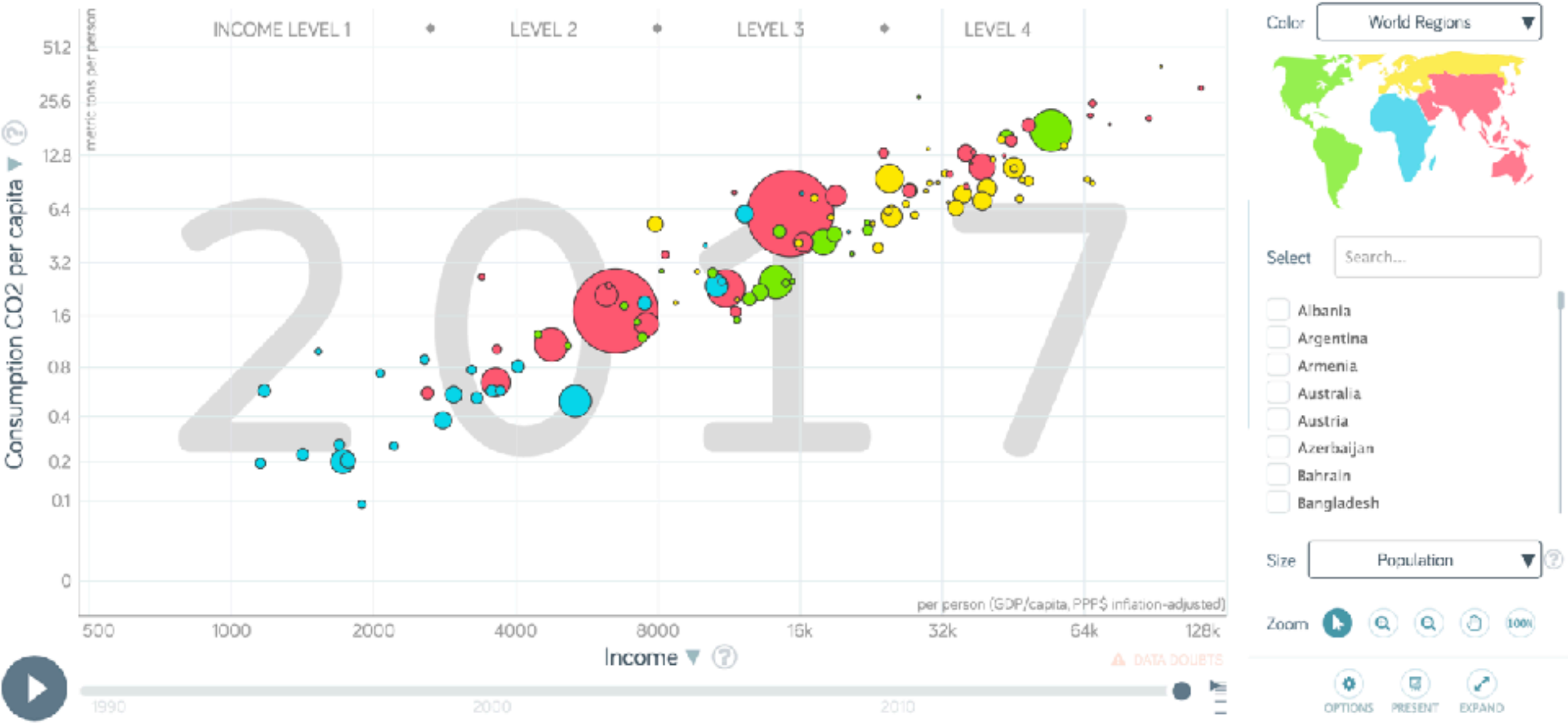


?

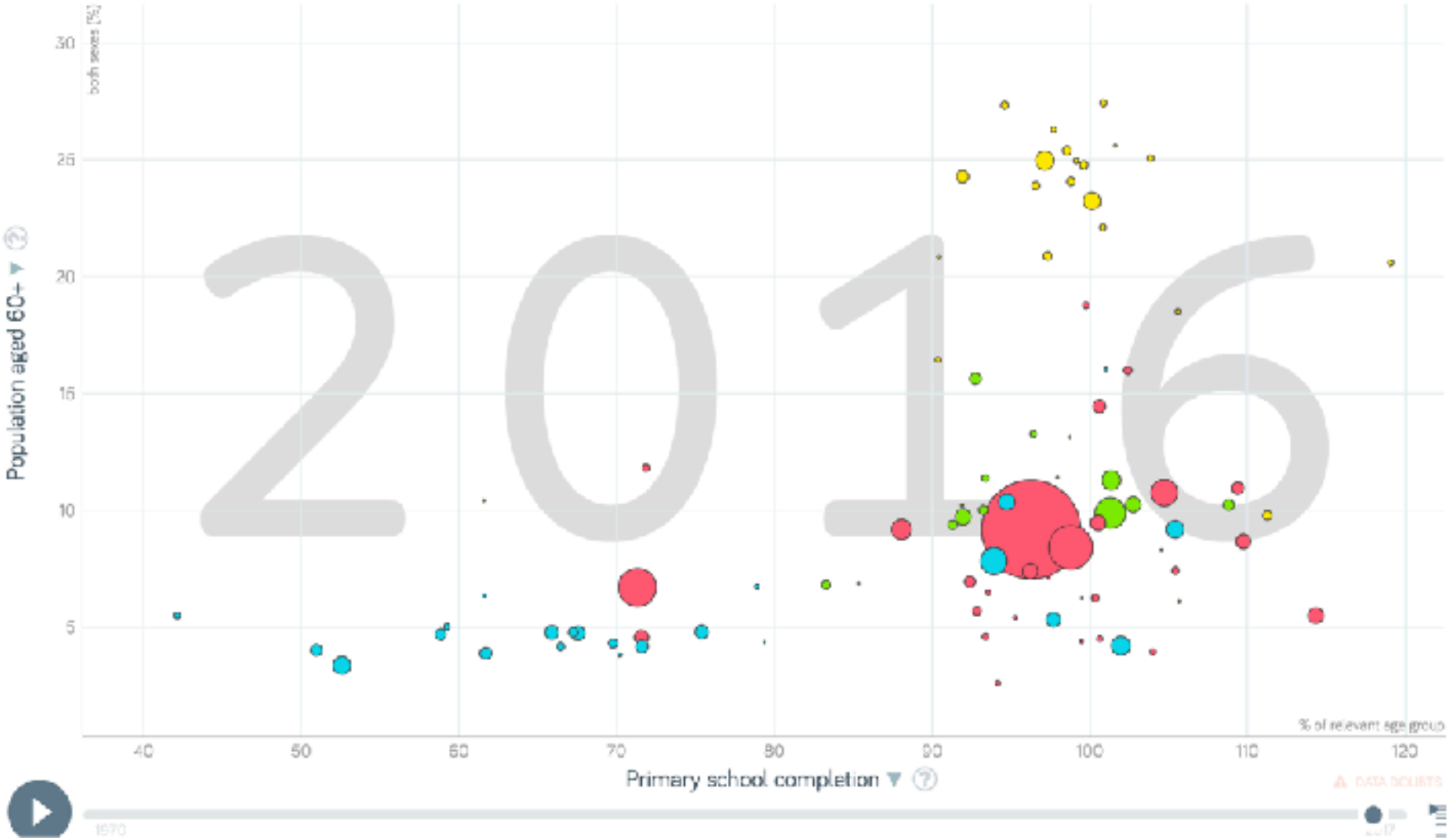
UNCERTAINTY - HOW CERTAIN ARE WE?

There are measures of uncertainty
Context matters
Everything is relative
No model is perfect

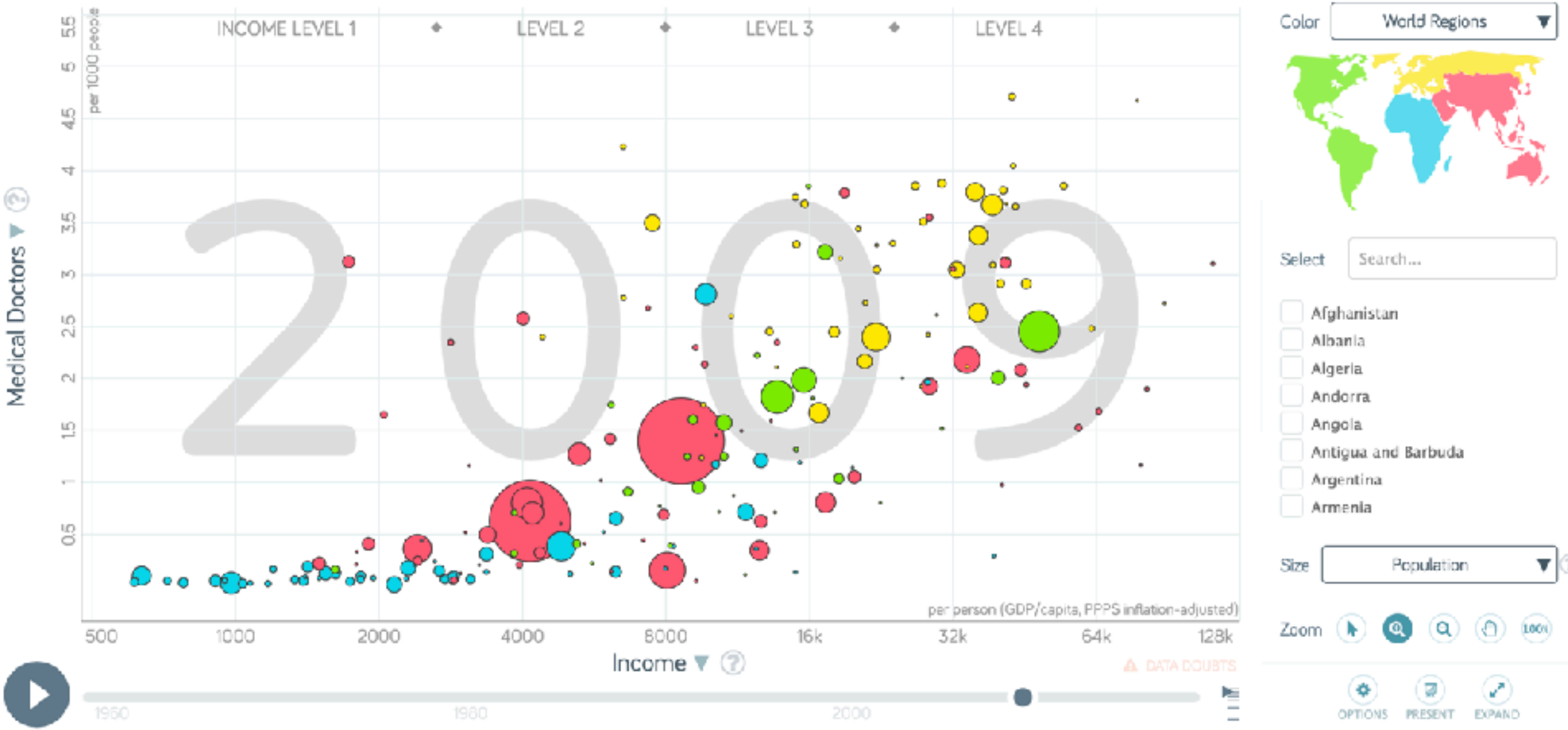
LEARN TO READ CORRELATION PLOTS



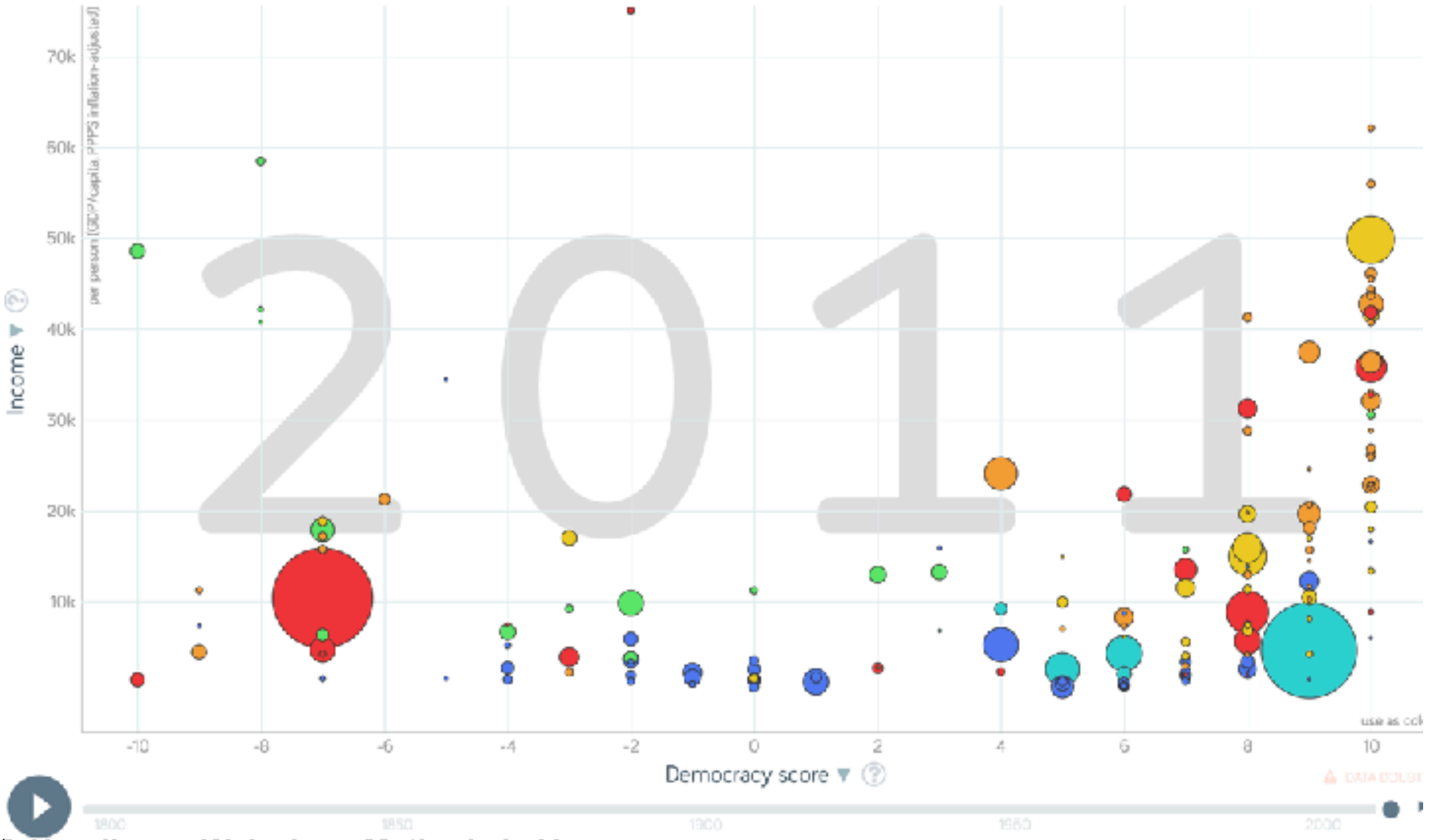
LEARN TO READ CORRELATION PLOTS



LEARN TO READ CORRELATION PLOTS



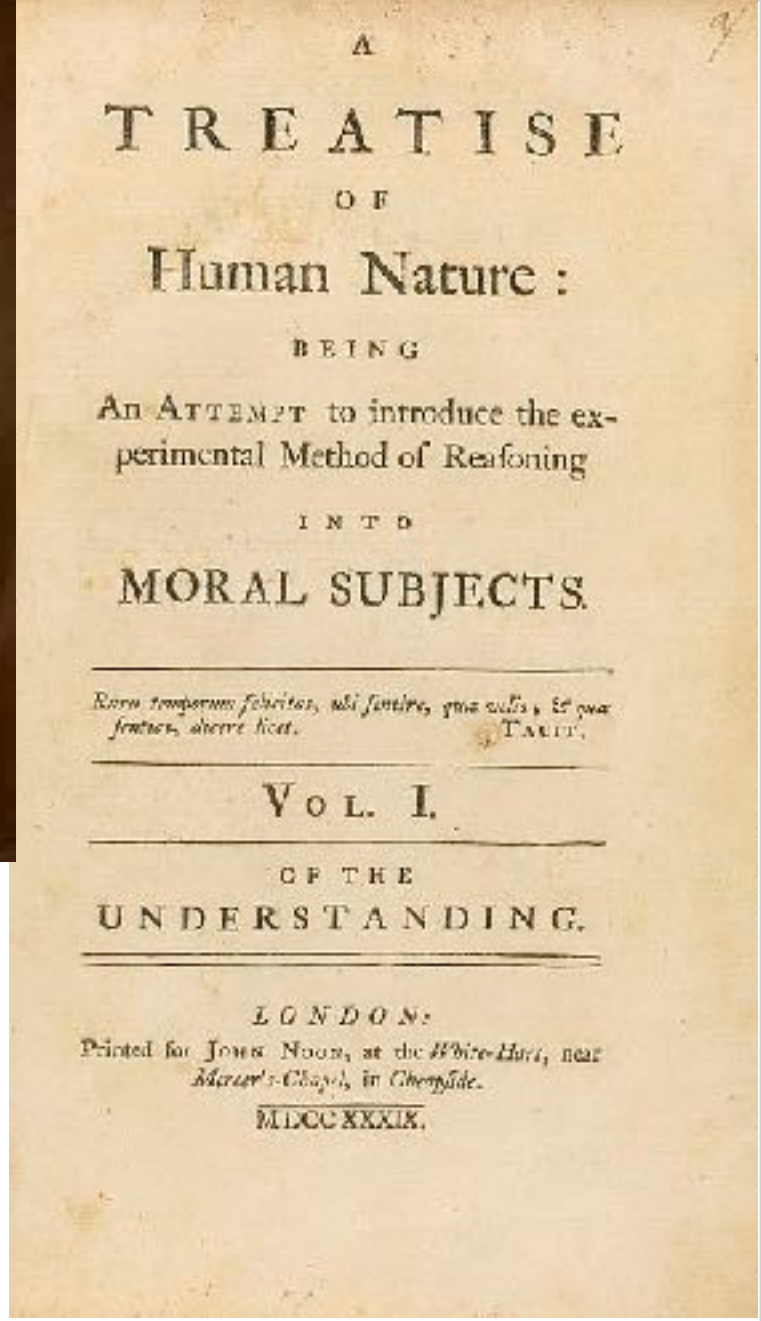
LET US START SIMPLE: STATISTICAL CORRELATIONS



CAUSALITY

?

Hume



A

TREATISE

OF

Human Nature:

BEING

An ATTEMPT to introduce the experimental Method of Reasoning

INTO

MORAL SUBJECTS.

Rara temporum felicitas, ubi sentire, quæ velis, & quæ sentias, dicere licet. TACIT.

VOL. I.

OF THE

UNDERSTANDING.

LONDON:

Printed for JOHN NOON, at the *White-Hart*, near *Mercer's-Chapel*, in *Cheapside*.

MDCCXXXIX.

Knowledge and probability

	Immediate	Inferential
Relations of ideas	intuition	demonstrative reasoning
Matters of fact	perception	probable reasoning

A GLIMPSE AT HUME'S CAUSALITY CRITERIA

1. The same cause produces the same effect.

2. If several objects create the same effect, then there must be a uniting criterion among them causing the effect.

3. If two objects have a different effect, there must be a reason that explains the difference.

1) THE SAME CAUSE PRODUCES THE SAME EFFECT.

- You fall from the 15th floor vs. you fall from the 40th floor.
- Heat and ice cream sales
- Fertilizer and crop yield

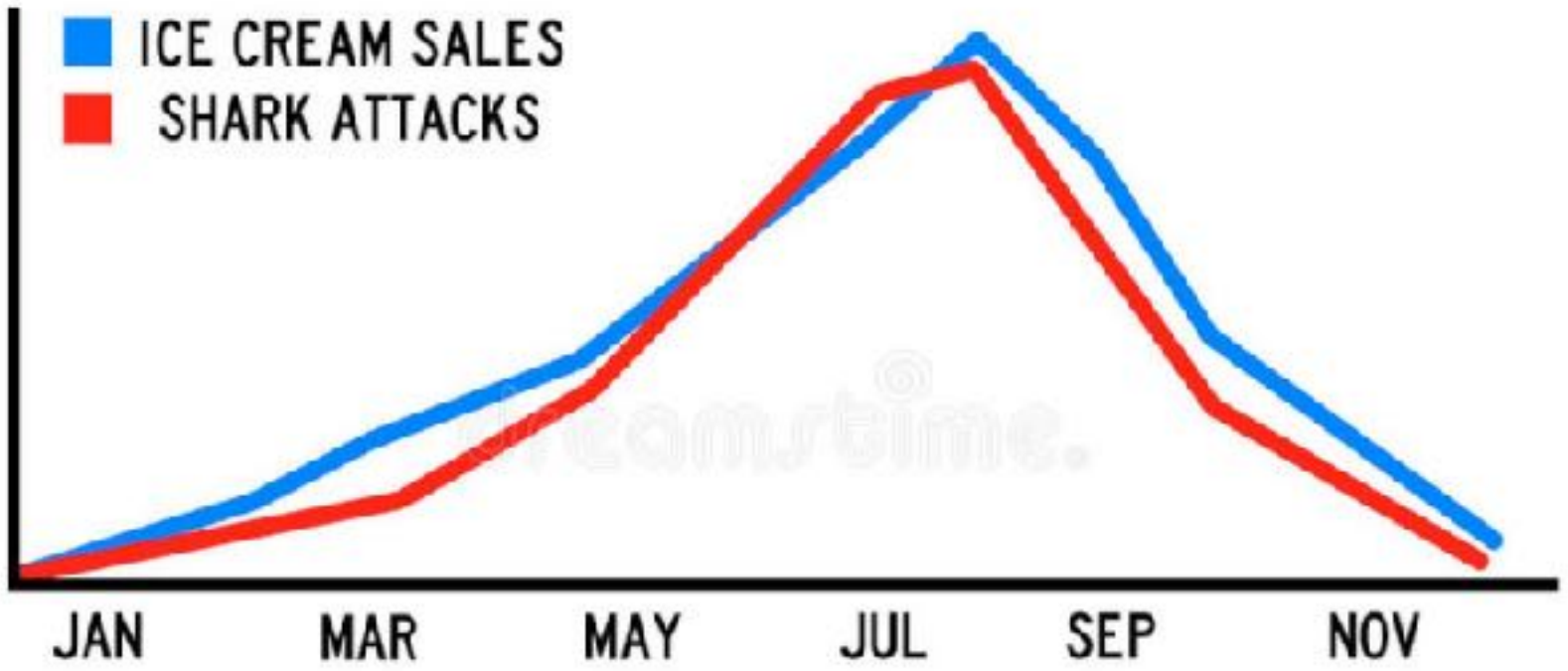
2) IF SEVERAL OBJECTS CREATE THE SAME EFFECT, THEN THERE MUST BE A UNITING CRITERION AMONG THEM CAUSING THE EFFECT.

- Different fertilisers
- Any types of soap against the Corona-Virus
- Ice cream or madeleines

3) IF TWO OBJECTS HAVE A DIFFERENT EFFECT, THERE MUST BE A REASON THAT EXPLAINS THE DIFFERENCE.

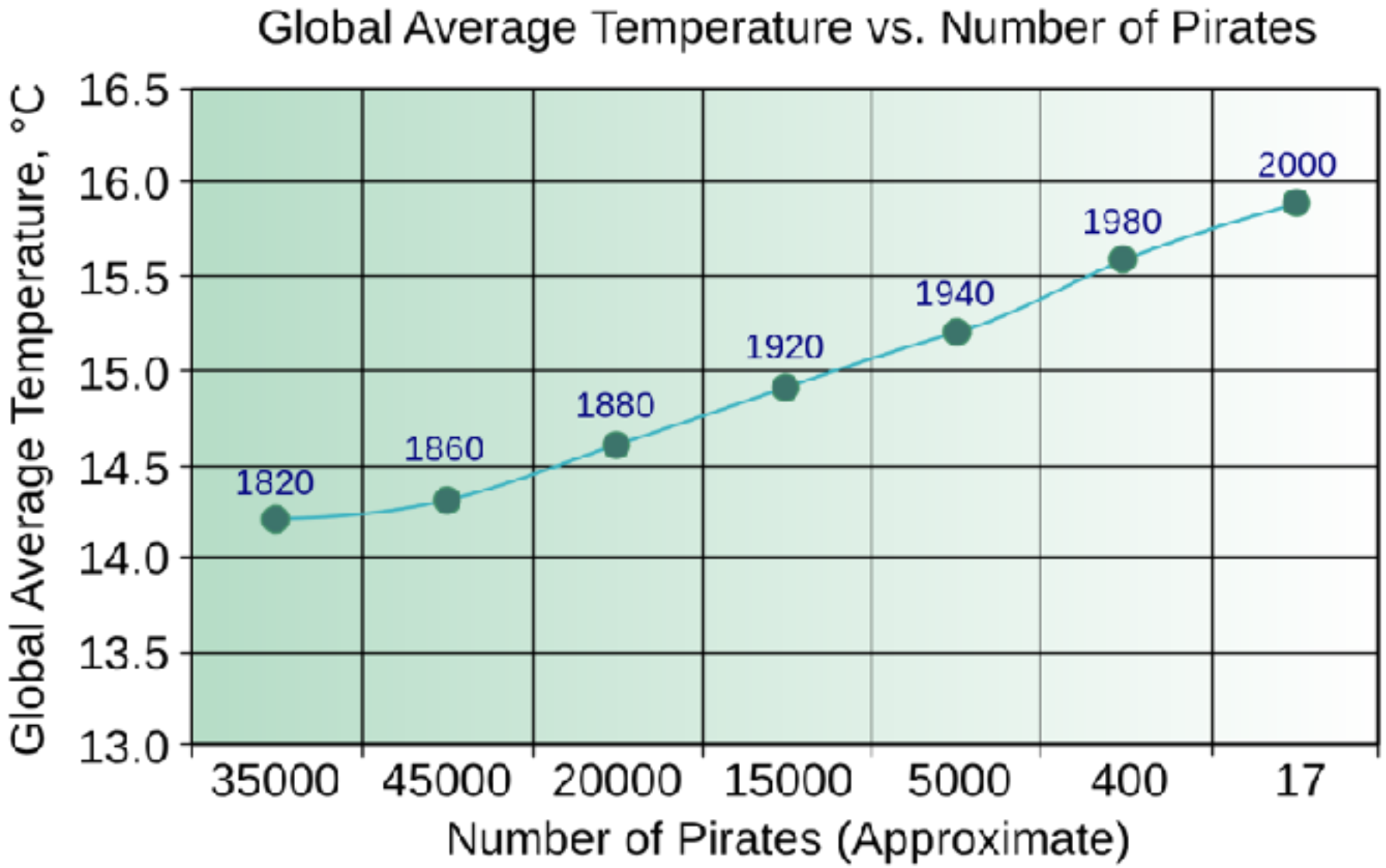
- Different drinks work in the same way against thirst
- Aspirin vs. Paracetamol
- Cow manure vs. horse manure

DO YOU WANT TO PREDICT OR TO EXPLAIN?



Both ice cream sales and shark attacks increase when the weather is hot and sunny, but they are not caused by each other (they are caused by good weather, with lots of people at the beach, both eating ice cream and having a swim in the sea)

AND CAN YOU EXPLAIN ANYTHING AT ALL?



THE HIGH ROAD AND THE LOW ROAD OF CAUSALITY

THE HIGH ROAD AND THE LOW ROAD OF CAUSALITY

The high road allows us to explain everything on how two things or phenomena are linked.

The low road of causality: May one thing or phenomena be causally linked to another thing or phenomena.

THE HIGH ROAD AND THE LOW ROAD OF CAUSALITY

The high road allows us to explain everything on how two things or phenomena are linked.

Positivism

The low road of causality: May one thing or phenomena be causally linked to another thing or phenomena.

Critical Realism

WHY A HIGH AND A LOW ROAD, THEN?

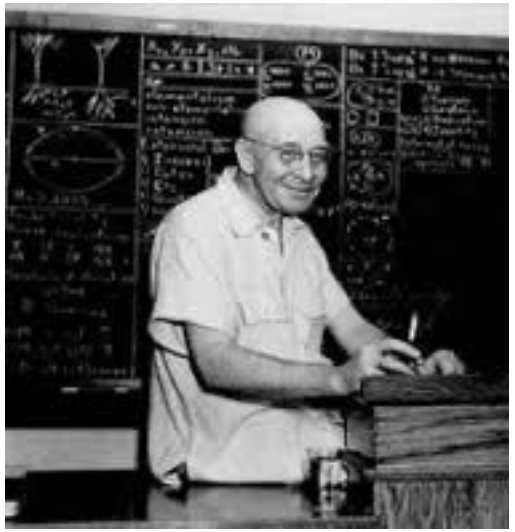
Some people assume that there are only two ways of doing Ethics, or arguing about morality. One is the *Low Road,* that merely appeals to our intuitions. The other is the *High Road,* Meta-Ethics. If we can give the best account of the nature of moral reasoning, we can hope that this will imply particular claims about morality. We can hope that our Meta-Ethics will imply conclusions in Ethics.

I believe that these are not the only ways in which we can argue about morality. I have not taken the High Road, except when I assumed that an acceptable moral theory cannot be directly collectively self-defeating. I have often taken the Low Road, appealing to our intuitions. But one of my main aims has been to explore a variety of different kinds of argument, that are between the Low and High extremes.

SOME GENERAL LAWS THAT MAY BE VALUABLE

All models are wrong
Some models are useful.

Everything needs to be as simple as possible, and as complex as necessary.

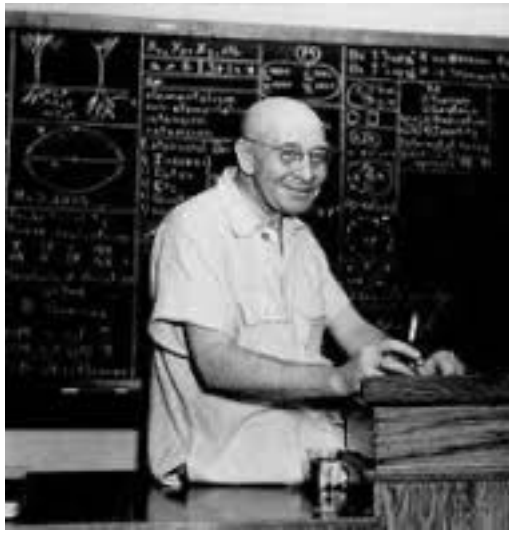


A map is not the territory it represents, but, if correct, it has a similar structure to the territory, which accounts for its usefulness.

WHAT SHOULD MODELS BE?

Georg Box

William of Ockham



Alfred Korzybski

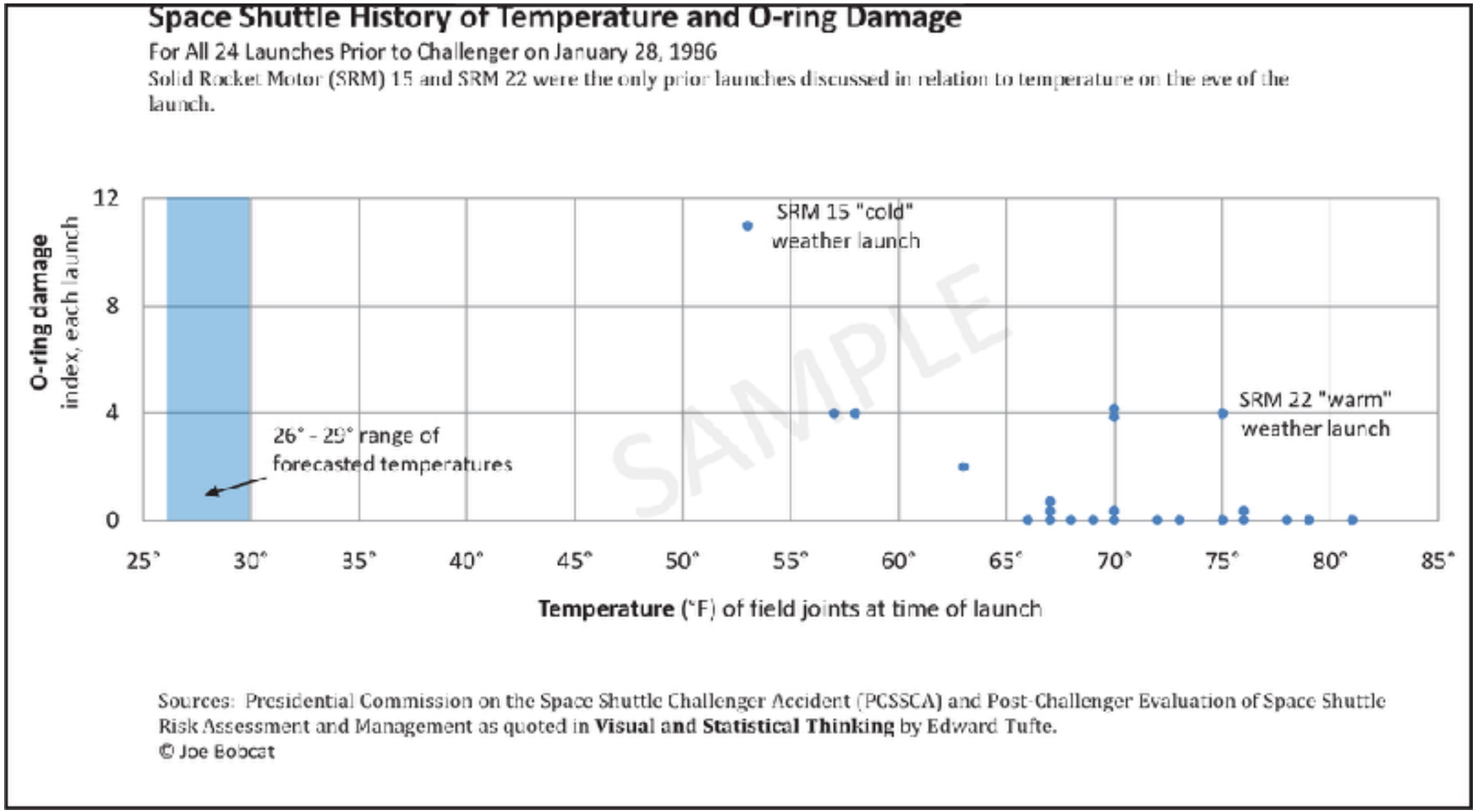
Approximations!

Parsimonious!

Generalisation!

AN EXAMPLE OF COMPLEXITY: NON-LINEAR STATISTICS

NON LINEAR RELATIONS



 https://www.youtube.com/watch?v=raMmRKGkGD4

NON LINEAR RELATIONS: ECONOMIC GROWTH SLOWS DRAMATICALLY WHEN THE SIZE OF A COUNTRY'S DEBT RISES ABOVE 90% OF GROSS DOMESTIC PRODUCT

Reinhart and Rogoff's work showed average real economic growth slows (a 0.1% decline) when a country's debt rises to more than 90% of gross domestic product (GDP) – and this 90% figure was employed repeatedly in political arguments over high-profile austerity measures.

Data omission

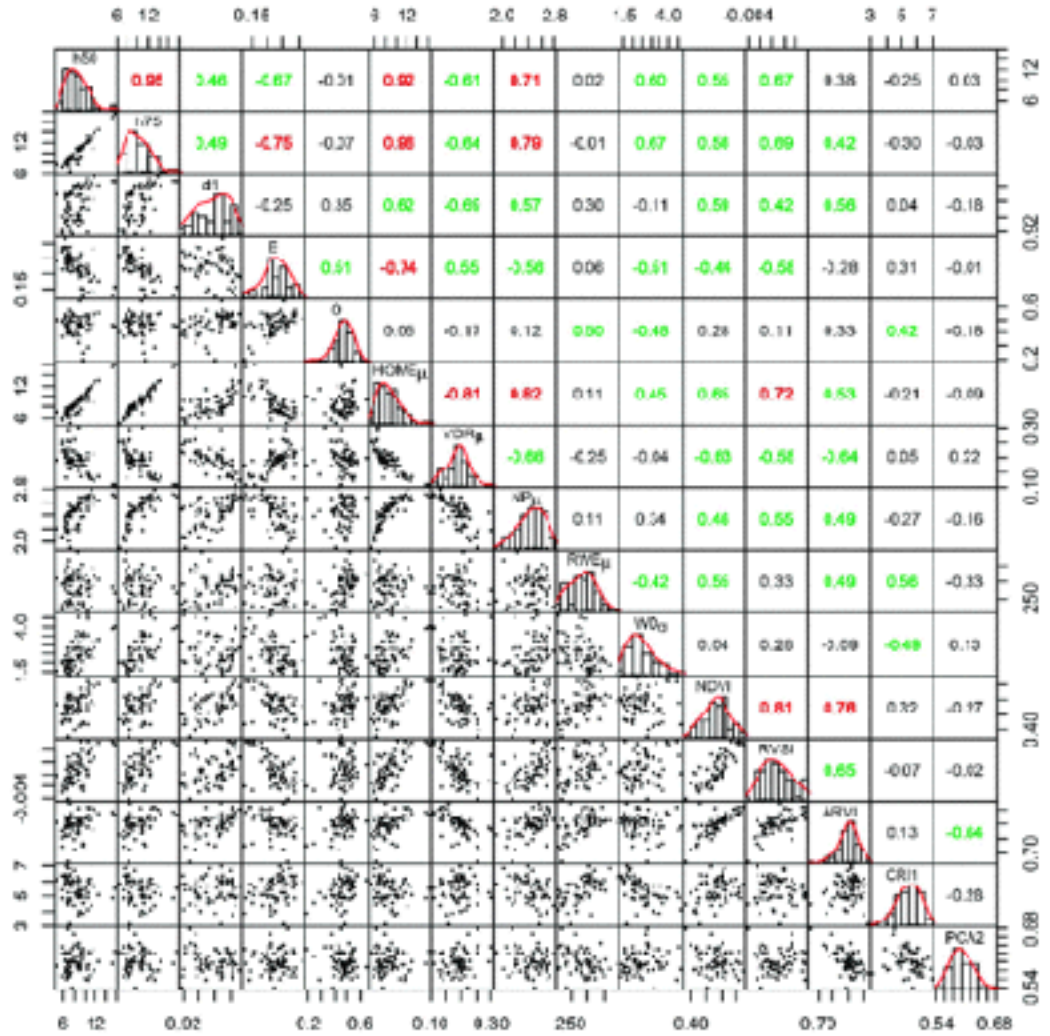
Spreadsheet error

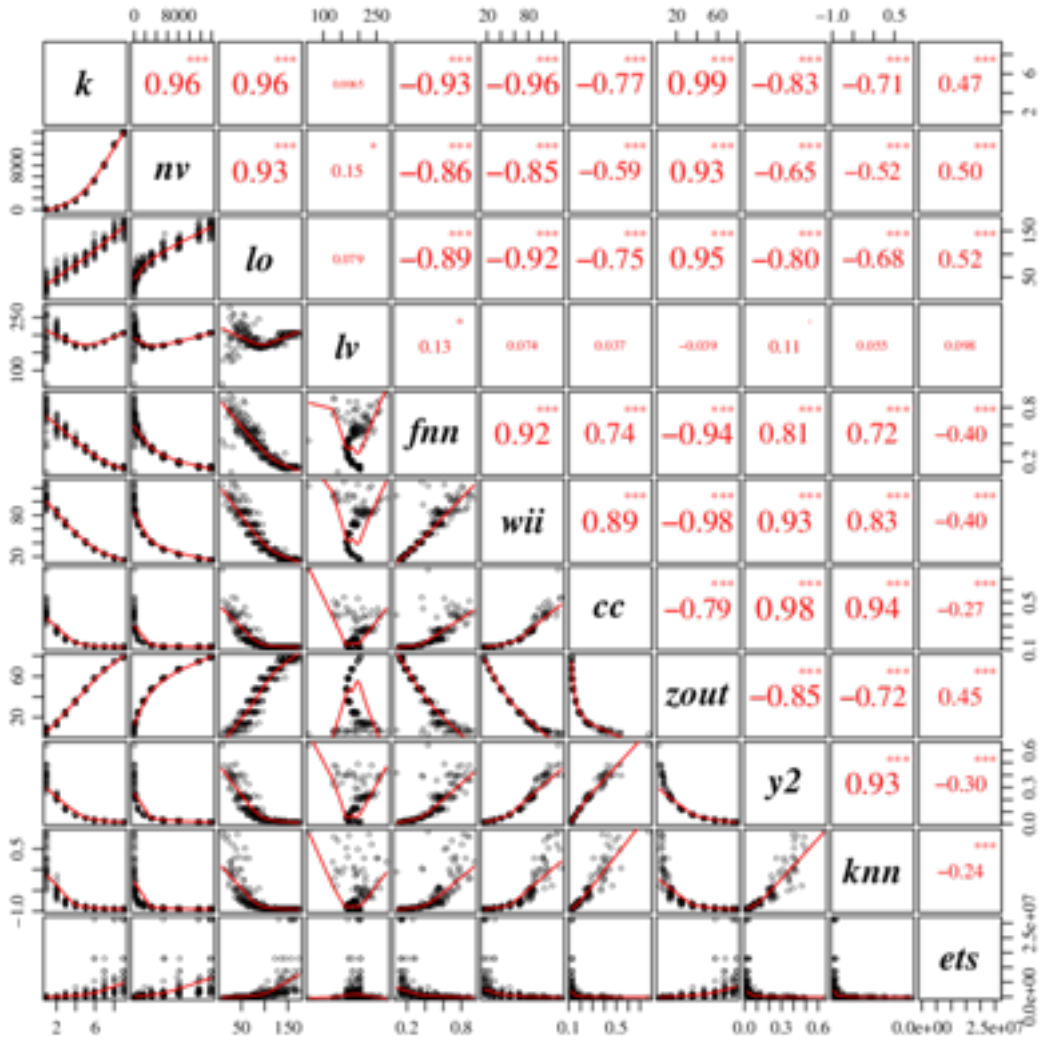
Misleading summary statistics

http://www.peri.umass.edu/fileadmin/pdf/working_papers/working_papers_301-350/WP322.pdf
https://theconversation.com/the-reinhart-rogoff-error-or-how-not-to-excel-at-economics-13646

NON LINEAR RELATIONS CAN BE CATASTROPHIC

AFTER TEN THOUSANDS OF PLOTS AND MODELS, INTUITION





CAUSALITY OR CORRELATION? ONE SUMMARY.

- In a century of numbers, correlations mattered
- Yet, correlation can predict, and may even help to explain
- Correlation models are generalisation, approximations, and (ideally) parsimonious
- Causality can be altogether different thing
- Suggestions of causality are rooted in logic
- Whether they matter is a question of philosophy of science
- Whether you may ultimately understand relations between to continuous variables is a matter of practice

CONTACT

Henrik von Wehrden | Faculty of Sustainability
Universitätsallee 1 | 21335 Lüneburg
Fon 04131.677-1671 | henrik.von_wehrden@leuphana.de
https://henrikvonwehrden.web.leuphana.de
Use Rocket chat!