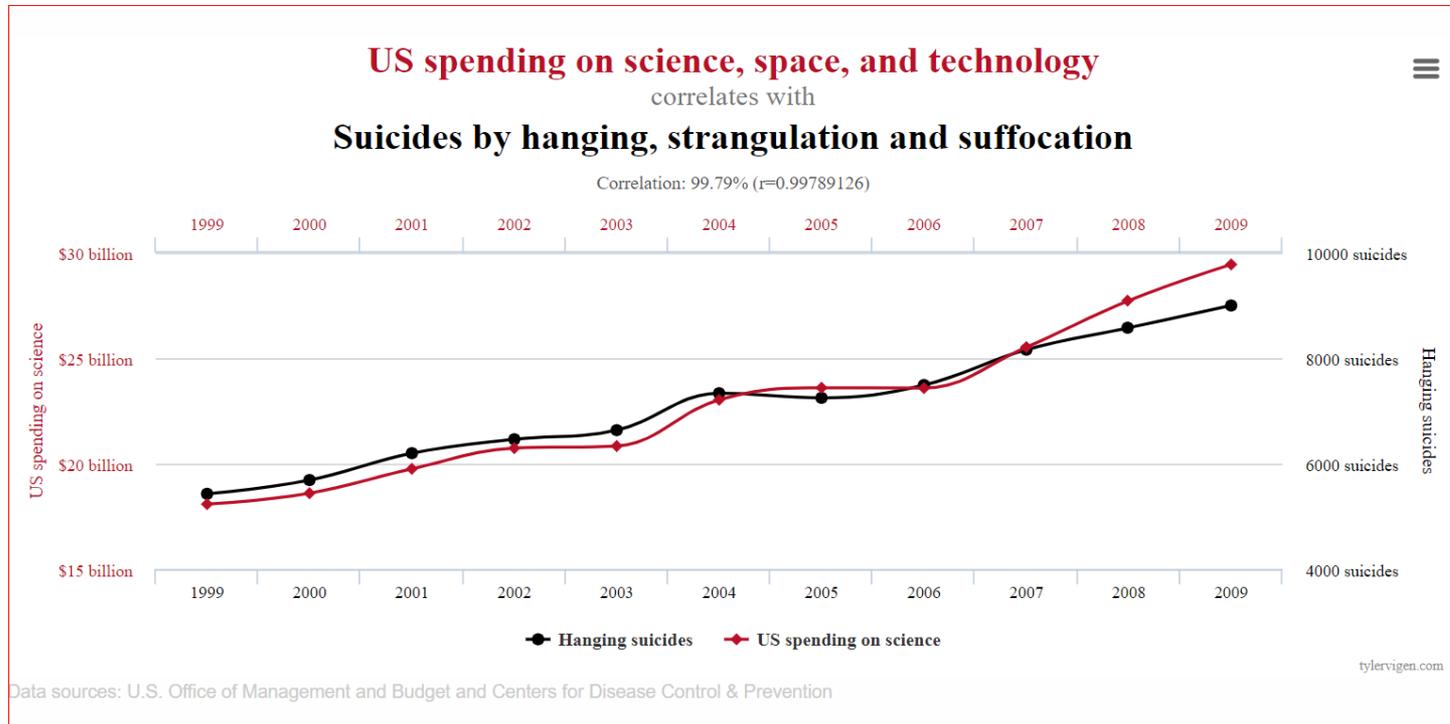


Value of Data – What if Physics could help?

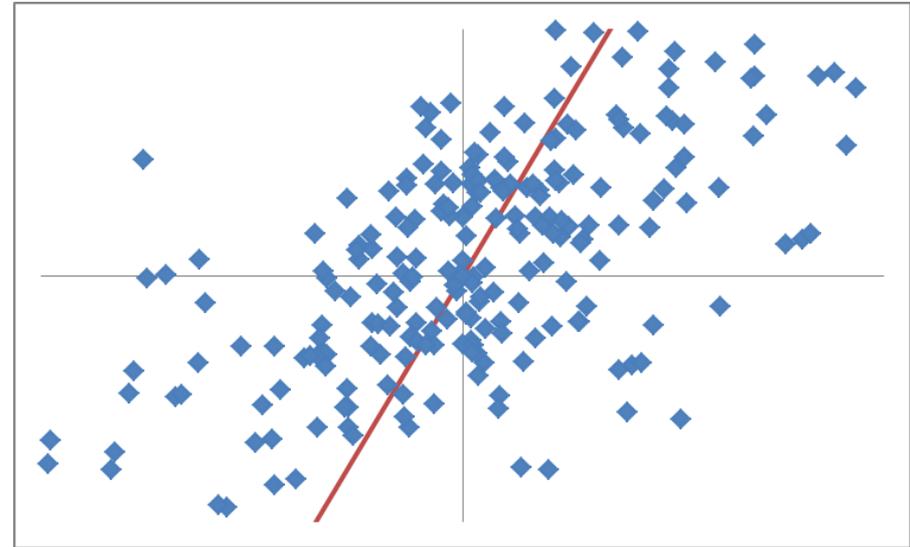
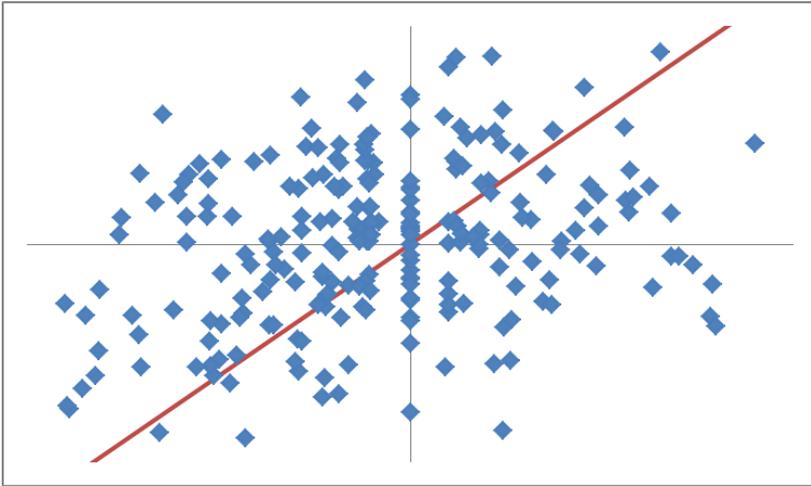
André Gorius, PhD

November 2019

The usual pitfall: Correlation does not mean Causality



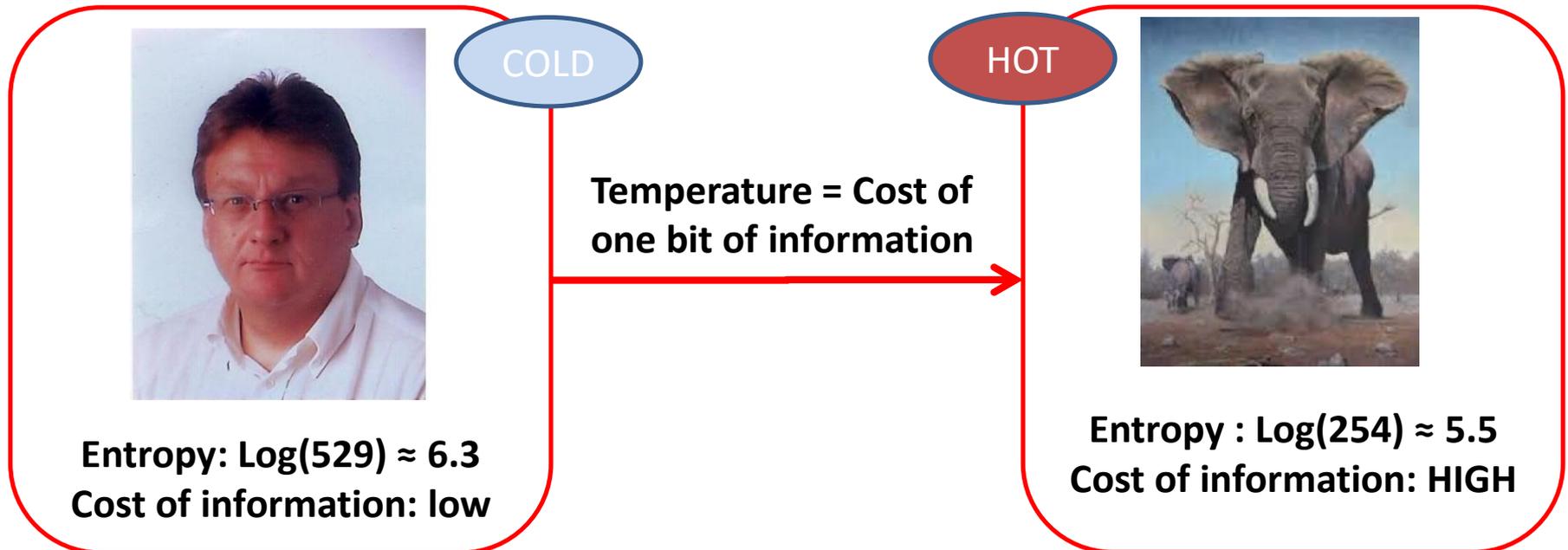
And when even correlations are highly questionable, many assume causality



Two (sadly) real examples of determination of « Beta »... for
CAPM determination of discount rate

Attempt to a Principled Approach: Value of datasets in Thermodynamics Terms

- Shannon's **Information** (bit) is the equivalent of Thermodynamic's **Entropy**
- **Value** (€) of a dataset is the equivalent of its **internal energy**
- If we have an equivalent of **temperature**, no need for a new **theoretical frame**: we have **Thermodynamics**
- **Take two jpg pictures with similar information content**



Our First and Second Laws

Thermodynamics

1. Work and Heat can be transformed into one another
Variation of **$U = \text{Work} + \text{Heat}$** depends only on the initial and final states of the system

$U = \text{Internal Energy}$

2. One can create Work from Heat only if Heat flows from a hot to a cold reservoir

State Function Entropy

Intangibles

1. Cost of Information and (Economic) Value can be transformed into one another

Variation of **$U (= \text{Information cost} + \text{Value})$** depends only on the initial and final states of the system

Exemple $U = \text{Ressources} = D + E$

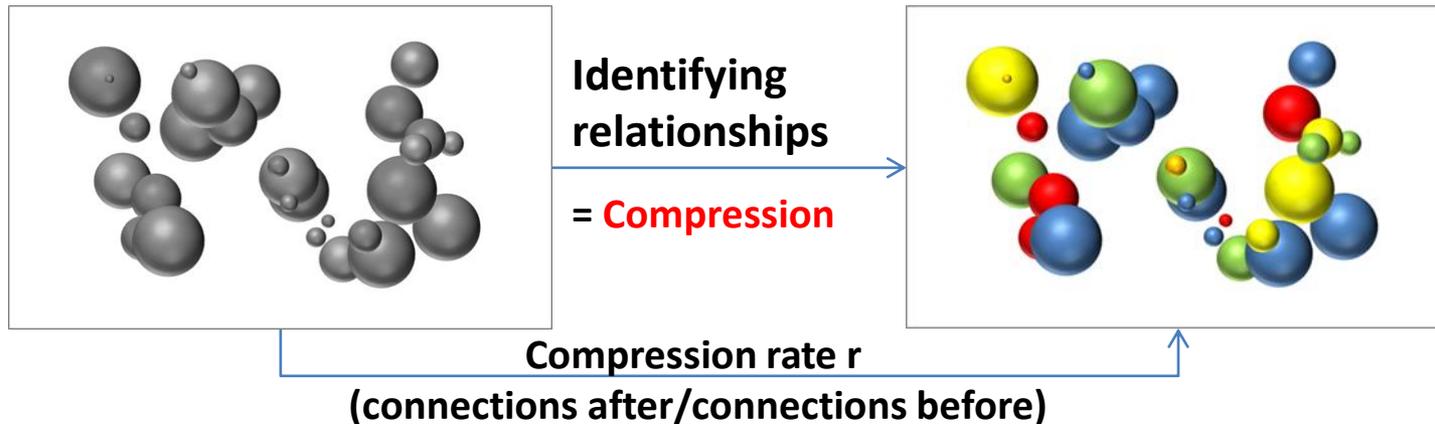
2. One can create sustainable Value from acquired information only if own information is more expensive to acquire than competition's Knowledge

Rigorous Analog to Entropy

With the right variables and functions, no need to re-invent the math!

A new way of looking to the value of datasets utilization

- Making business from data is a « heat engine »:
 - Input is the cash to create and use the model
 - Output is the cash one generates with the model (e.g. new business)
- What is data analytics?



- Applying the 2 principles determines the **maximum** value creation potential (attained only in an **ideal/reversible** world) :

$$\frac{\text{Cash out}}{\text{Cash in}} \leq \frac{r^{\gamma}-1}{r^{\gamma-1}-1}$$

γ related to value creation capacities of the organization

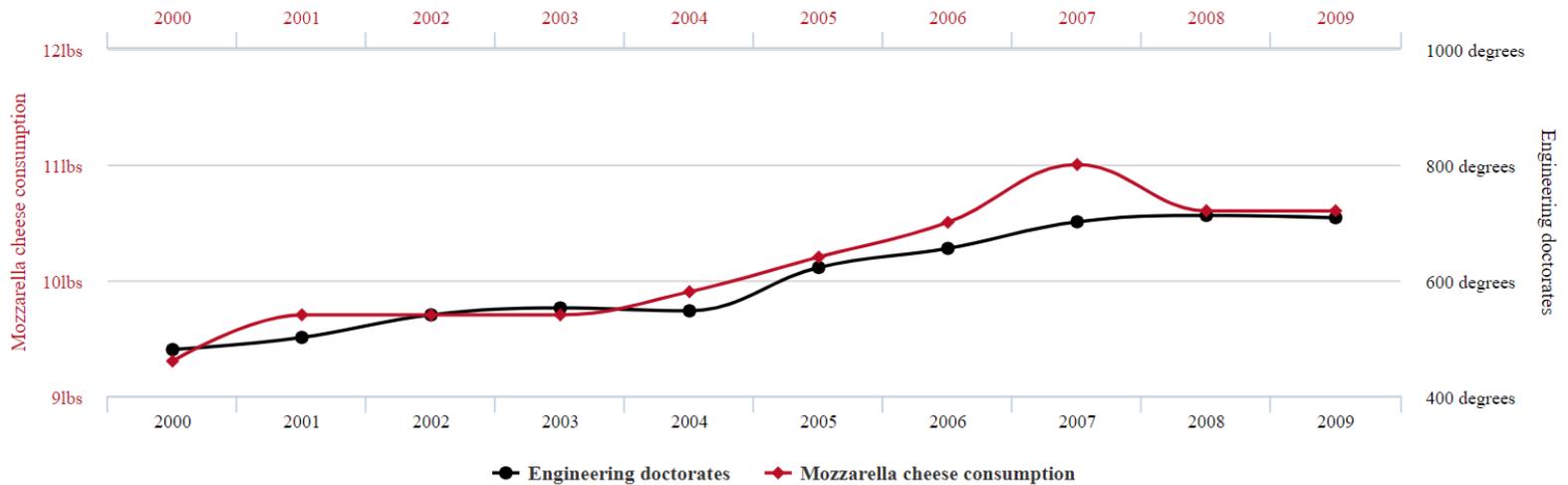
Conclusions

- Revisit value creation from simple first principles
 - In particular: why try to study separately capitals that are intrinsically complimentary?
- Link unit cost of information (« Temperature ») to value (« Work »)
 - WICI 2017 (Frankfurt): an « Insing model » links value creation capacity of a set of humans to their **achievement drive in front of incentives**
- Next step: capital evaluation in real situations, i.e. where irreversible phenomena are present

Per capita consumption of mozzarella cheese correlates with Civil engineering doctorates awarded



Correlation: 95.86% ($r=0.958648$)



Data sources: U.S. Department of Agriculture and National Science Foundation

tylervigen.com