

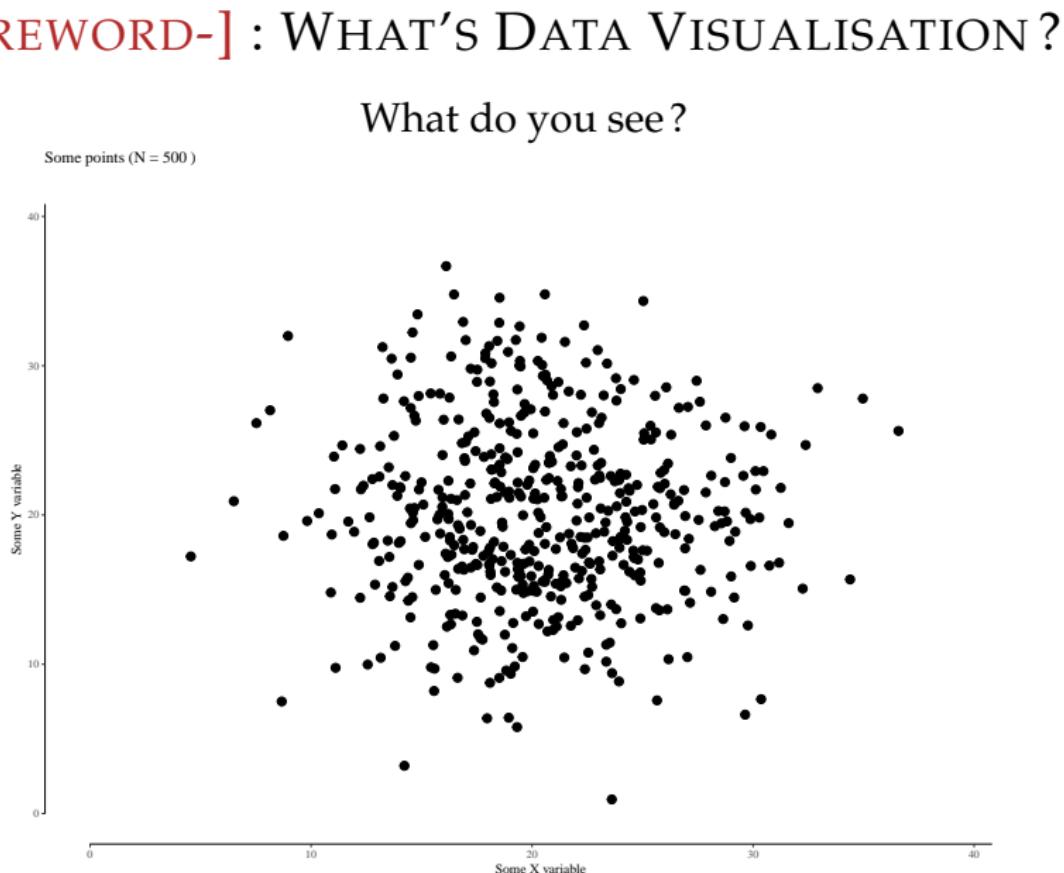
How to Lie with Graphics?

Christophe Bontemps

Statistical Institute for Asia and the Pacific (United Nations)

Toulouse School of Economics (INRAE)

BPLIM DataViz Workshop, Porto
December, 2020



Foreword Visual lies!

Classics

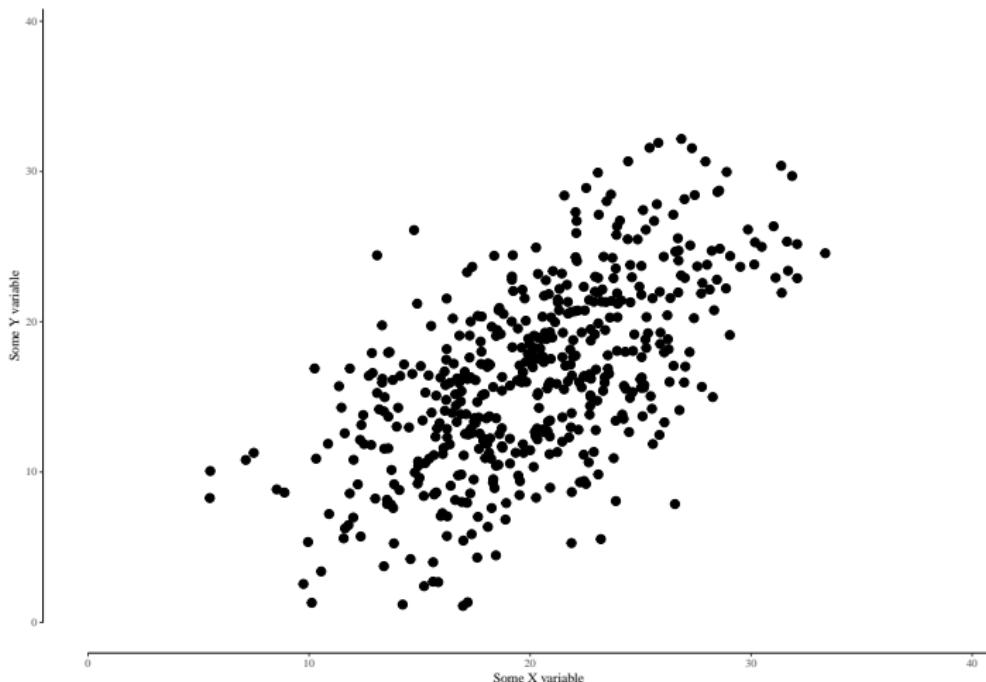
Tricky stuff Serious lies



[-FOREWORD-] : WHAT'S DATA VISUALISATION ?

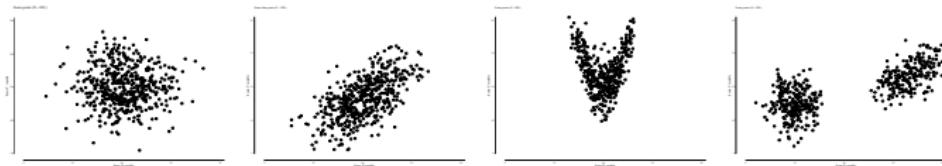
And here, what do you see?

Some other points ($N = 500$)

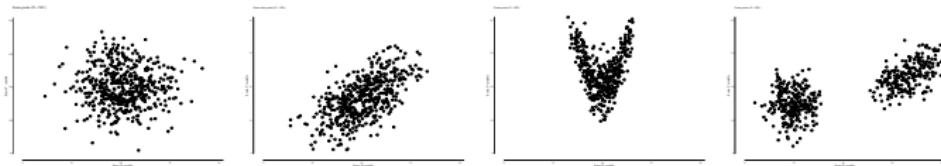




“DATA VISUALISATION” AS A STATISTICAL TEST



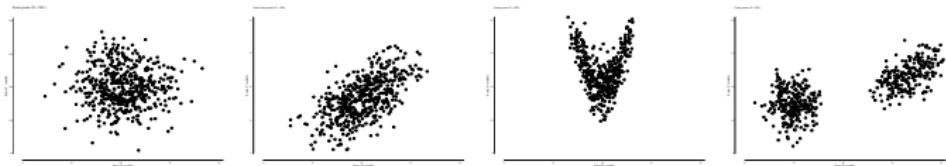
“DATA VISUALISATION” AS A STATISTICAL TEST



“The human eye acts as a broad feature detector and general statistical test”. Buja et al. (2009)



“DATA VISUALISATION” AS A STATISTICAL TEST

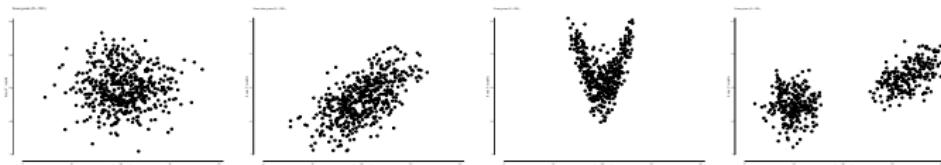


“The human eye acts as a broad feature detector and general statistical test”. Buja et al. (2009)

Test : H_0 : {There is "nothing"} = {No relation}



“DATA VISUALISATION” AS A STATISTICAL TEST



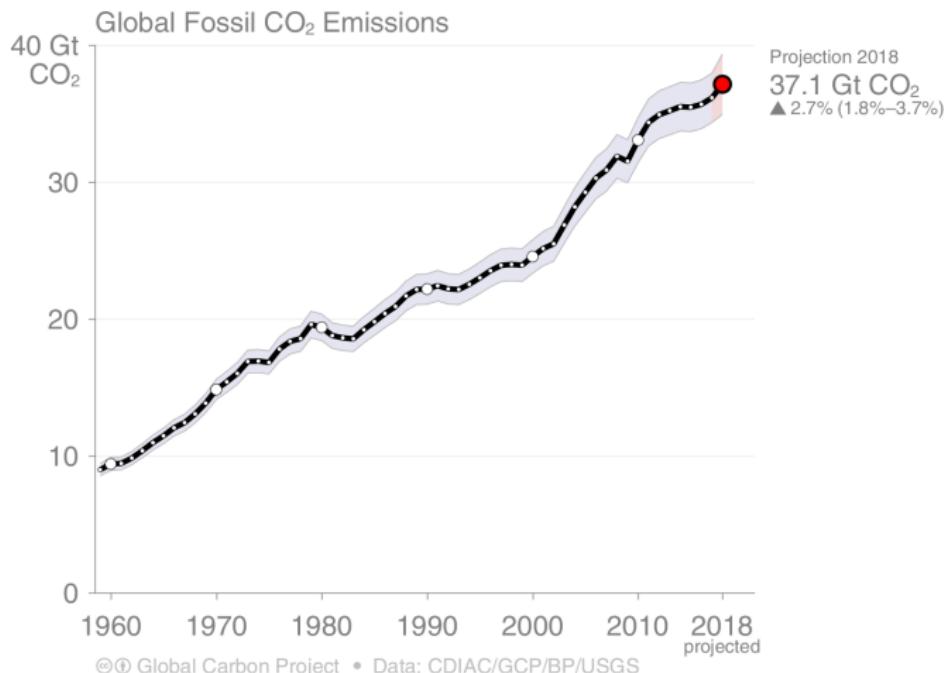
“The human eye acts as a broad feature detector and general statistical test”. Buja et al. (2009)

Test : H_0 : {There is "nothing"} = {No relation}

$H_1 : \{ \text{There is "something"} \} = \{ \text{There is some relation (Correlation, linearity, heterogeneity, groups..)} \}$

[- WHAT WE DO : -] : IMPLICIT COMPARISONS

What does this curves tells you?



[- WHAT WE DO :-] : EXPLICIT COMPARISONS

We compare : **surfaces...**



[- WHAT WE DO : -] : EXPLICIT COMPARISONS

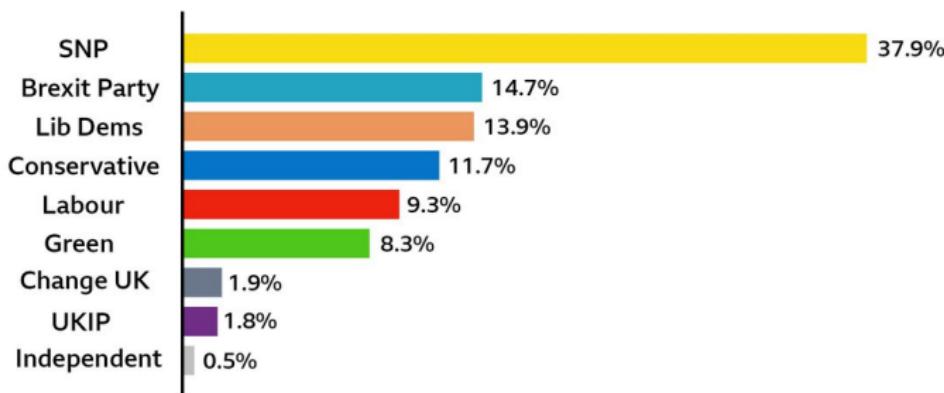
lines...



[- WHAT WE DO :-] : EXPLICIT COMPARISONS

lengths...

European Parliament election results 2019
Percentage of votes won in Scotland



(After 31 of 32 council results)

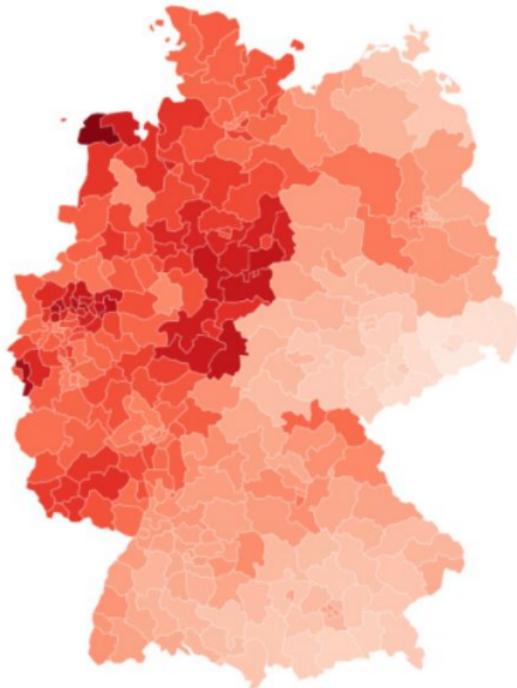
[- WHAT WE DO :-] : EXPLICIT COMPARISONS

colors...

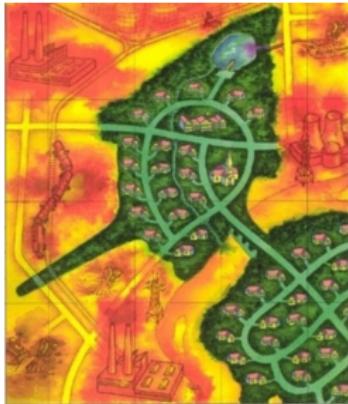
SPD

5%

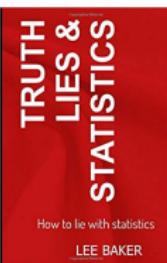
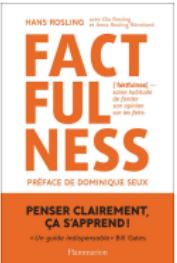
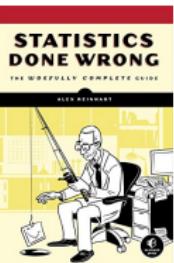
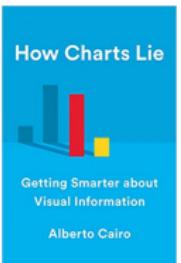
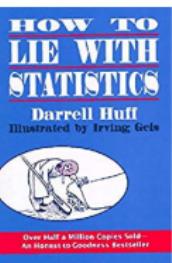
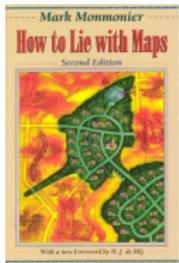
40%



What if all of this was a lie?



Should we learn to lie?!



[- THIS TALK -]

Definition :

[- THIS TALK -]

Definition :

"Lie" : What You See (on a graphic)

Is **Not**

What You Have (in the data)

[- THIS TALK -]

Definition :

“Lie” : What You See (on a graphic)

Is **Not**

What You Have (in the data)

Fact :

Many (conflicting) “rules” on dataviz

[- THIS TALK -]

Definition :

“Lie” : What You See (on a graphic)

Is **Not**

What You Have (in the data)

Fact :

Many (conflicting) “rules” on dataviz

My goal :

Decipher graphics and identify visual lies

[- THIS TALK -]

Definition :

“Lie” : What You See (on a graphic)

Is **Not**

What You Have (in the data)

Fact :

Many (conflicting) “rules” on dataviz

My goal :

Decipher graphics and identify visual lies

Outcome :

10+ rules for “lying” in a future communication ?

[- THIS TALK -]

Definition :

"Lie" : What You See (on a graphic)

Is **Not**

What You Have (in the data)

Fact :

Many (conflicting) "rules" on dataviz

My goal :

Decipher graphics and identify visual lies

Outcome :

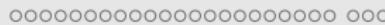
10+ rules for "lying" in a future communication ?

At least, it is **fun !**

Foreword Visual lies!



Classics



Tricky stuff Serious lies



VISUAL LIES CAN BE ANYWHERE [PACKAGING!]

VISUAL LIES CAN BE ANYWHERE [PACKAGING!]



VISUAL LIES CAN BE ANYWHERE [PACKAGING!]



Foreword Visual lies!



Classics



Tricky stuff Serious lies



VISUAL LIES CAN BE ANYWHERE [PICTURES!]

VISUAL LIES CAN BE ANYWHERE [PICTURES!]



Source : Astrid Helmer Mørck

VISUAL LIES CAN BE ANYWHERE [PICTURES!]



Source : Astrid Helmer Mørck

Foreword Visual lies!



Classics



Tricky stuff Serious lies



VISUAL LIES CAN BE ANYWHERE [POLITICS]

VISUAL LIES CAN BE ANYWHERE [POLITICS]

Covid-19 > Tests completed through April 5



VISUAL LIES CAN BE ANYWHERE [POLITICS]



Source : David Yanofsky (Quartz)

VISUAL LIES CAN BE ANYWHERE [POLITICS]

Covid-19 > Tests completed through April 5

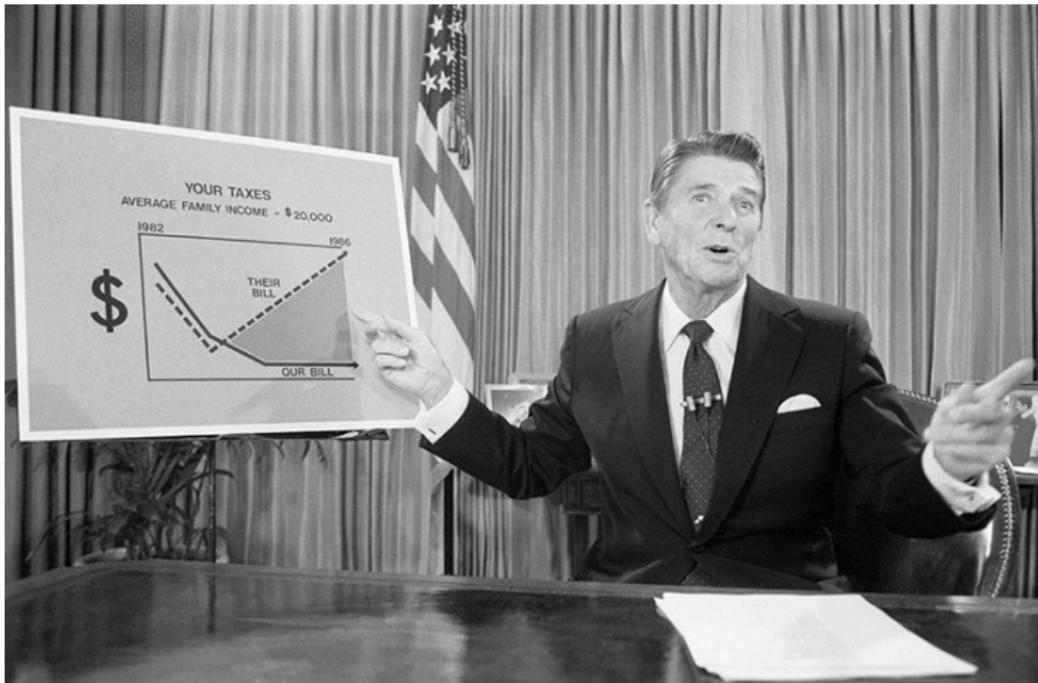


VISUAL LIES CAN BE ANYWHERE [POLITICS]

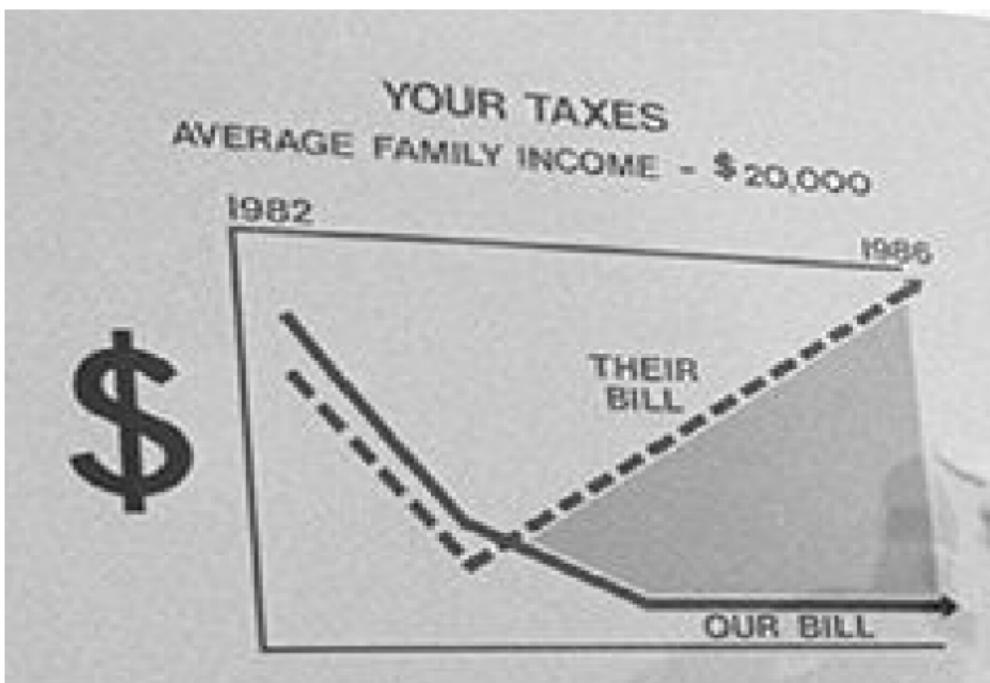


Source : Dave Gilson (Mother Jones)

VISUAL LIES CAN BE ANYWHERE [POLITICS]



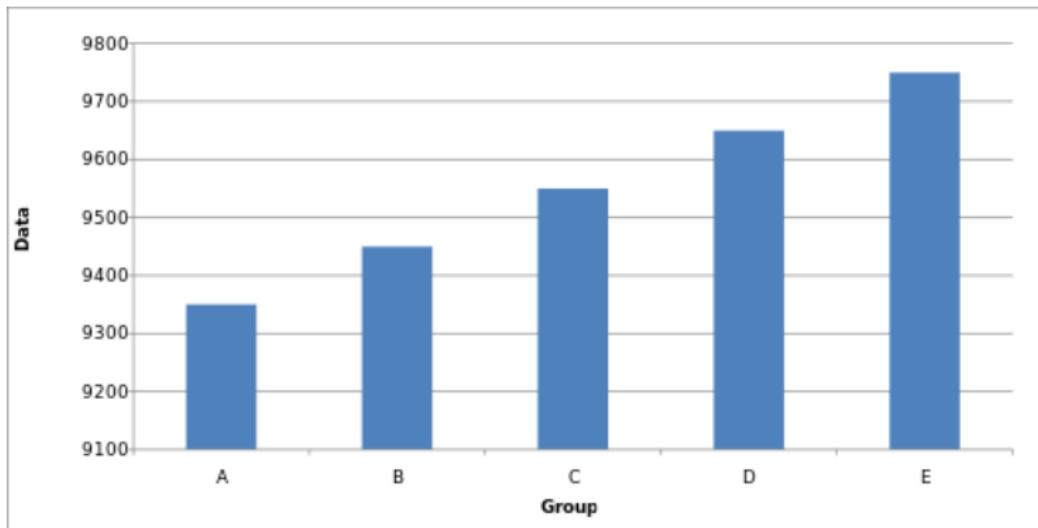
VISUAL LIES CAN BE ANYWHERE [POLITICS]



[- RULE #1 : USE DEVILISH AXIS! -]

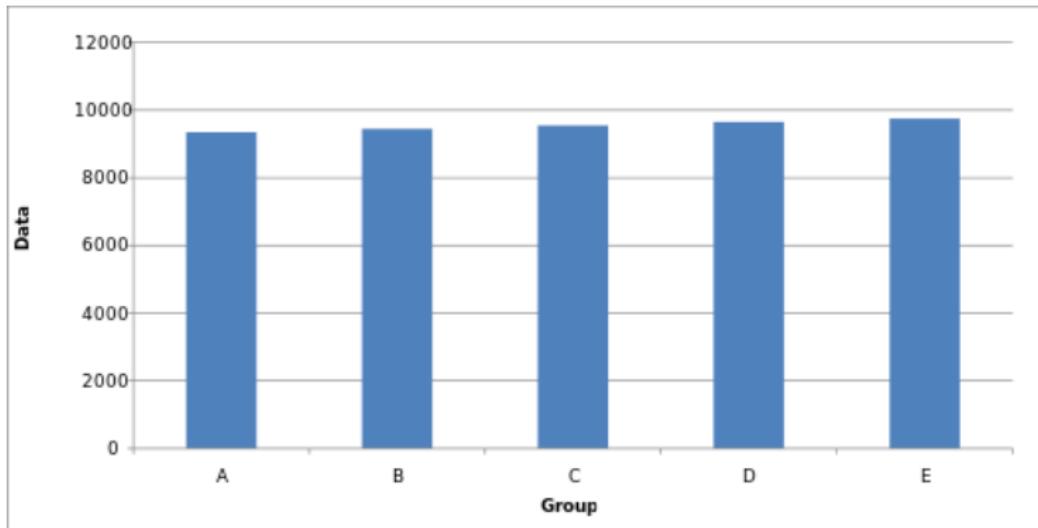


[-RULE #1 : USE DEVILISH (TRUNCATED) AXIS ! -]



Source : Wikipedia: Misleading graph

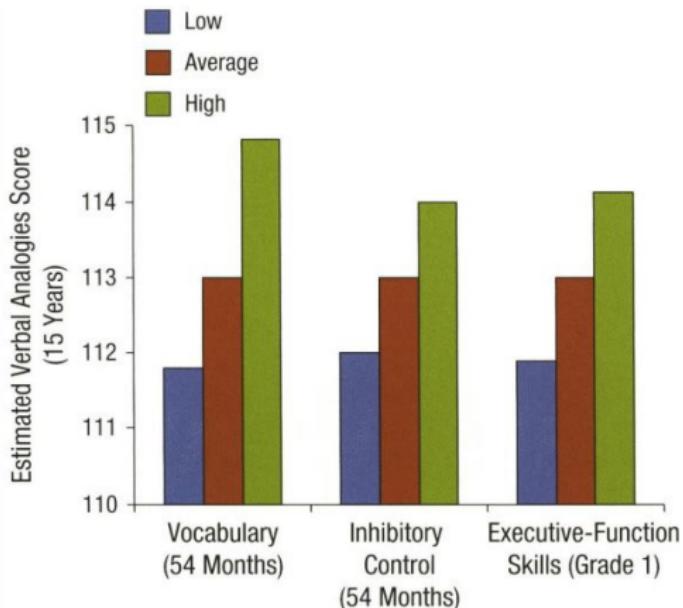
[-RULE #1 : USE DEVILISH (TRUNCATED) AXIS!-]



Source : Wikipedia: Misleading graph

[-RULE #1 : USE TRUNCATED Y-AXIS!-]

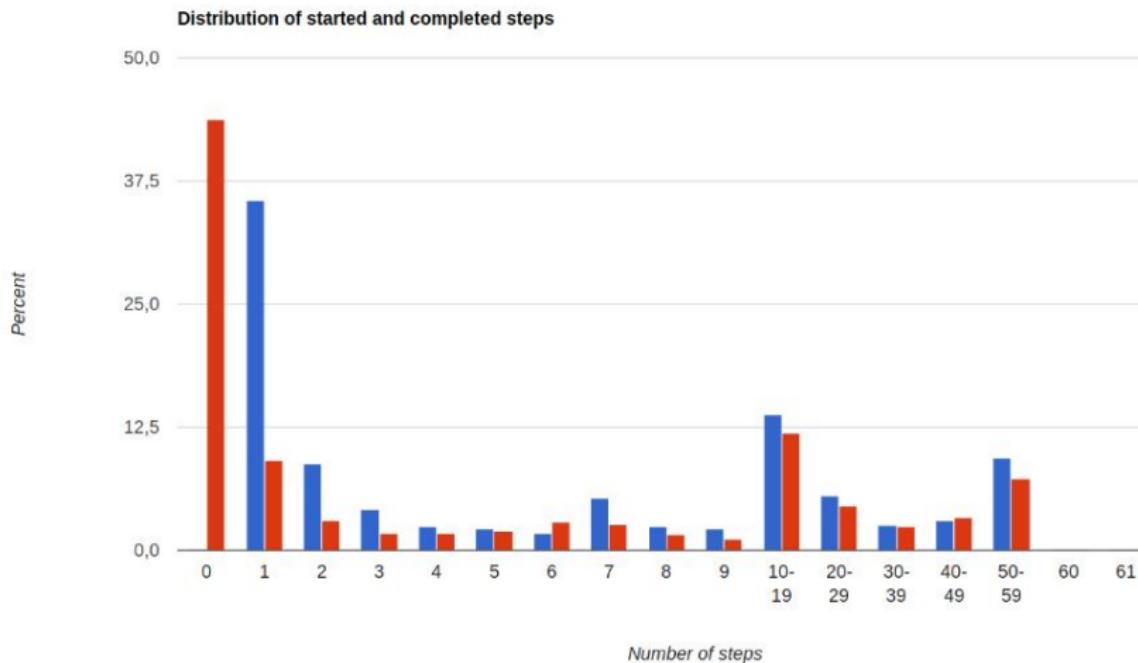
Example : Published paper



Source : Richland and Burchinal (2013)

[-RULE #1-BIS : USE TRUNCATED X-AXIS!-]

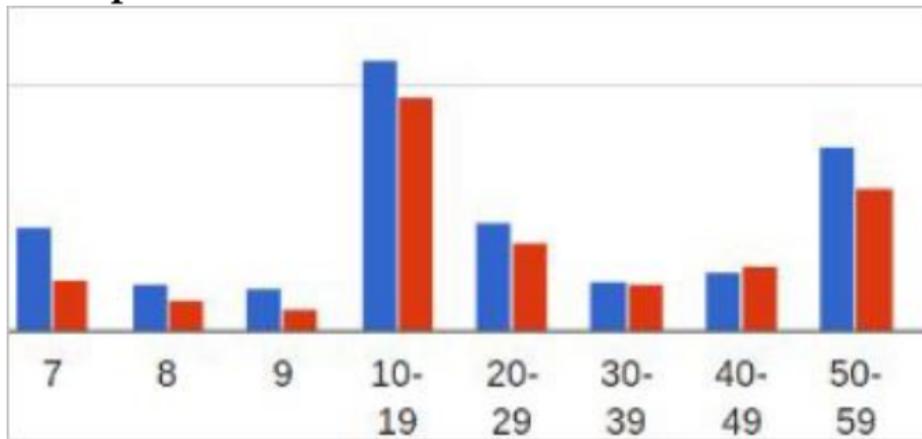
Examples are common !



Source : My 2017' students (Jan & Mohamed), but also recent researchers' presentations

[-RULE #1-BIS : USE TRUNCATED X-AXIS!-]

Examples are common!



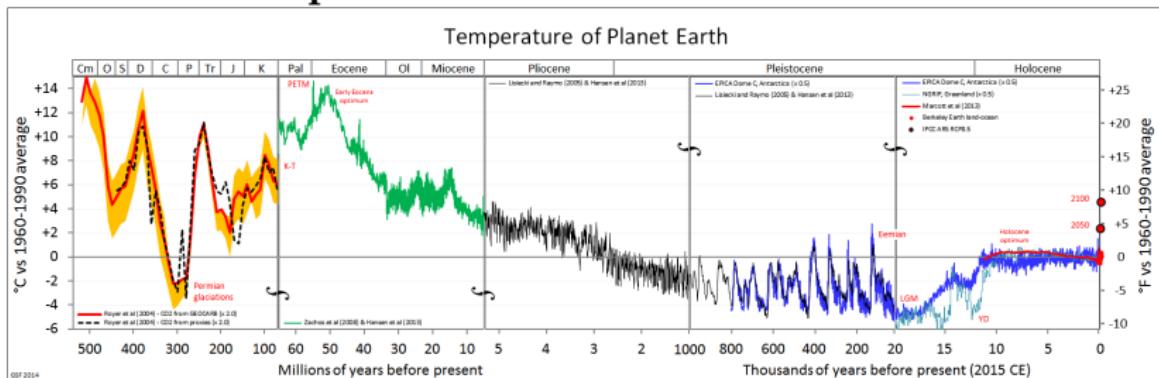
Source :

My 2017' students (Jan & Mohamed), but also recent researchers' presentations



[-RULE #1-BIS : USE TRUNCATED X-AXIS !-]

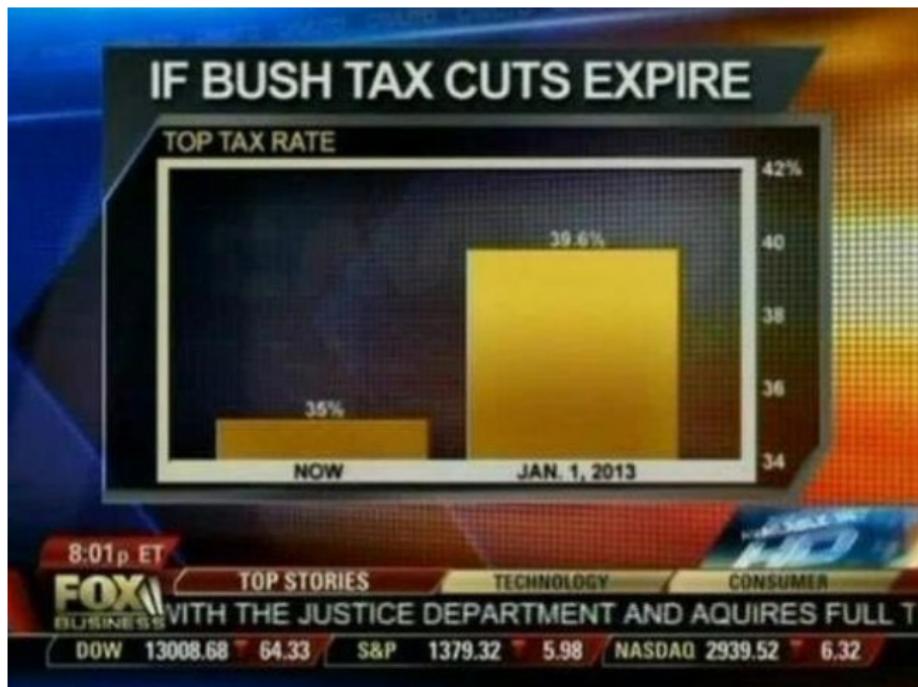
Published example are common!



Source : Wikipedia

[-ENTERING COMPLEXITY -]

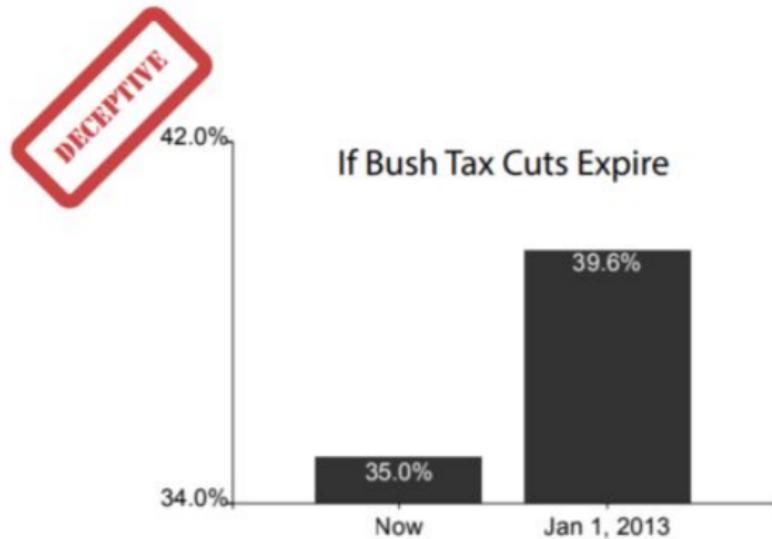
Example : Fox news



Source : Techna Verba Scripta

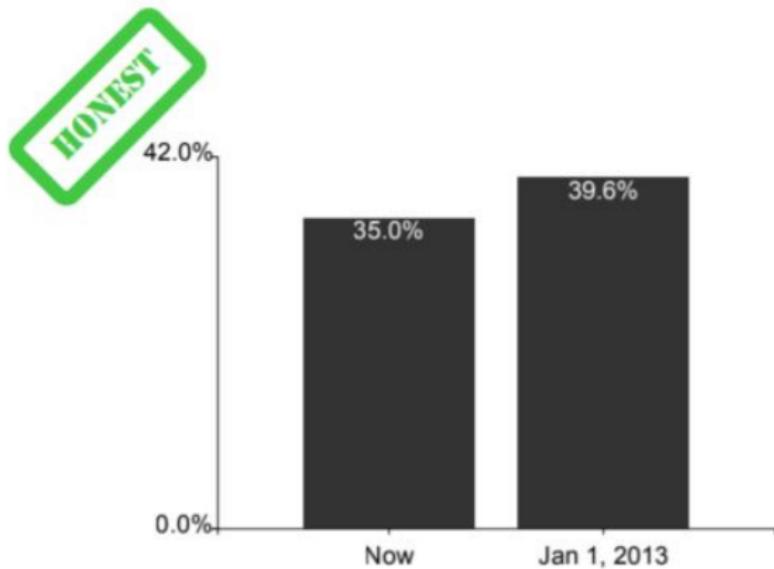
[-ENTERING COMPLEXITY-]

So this is bad :



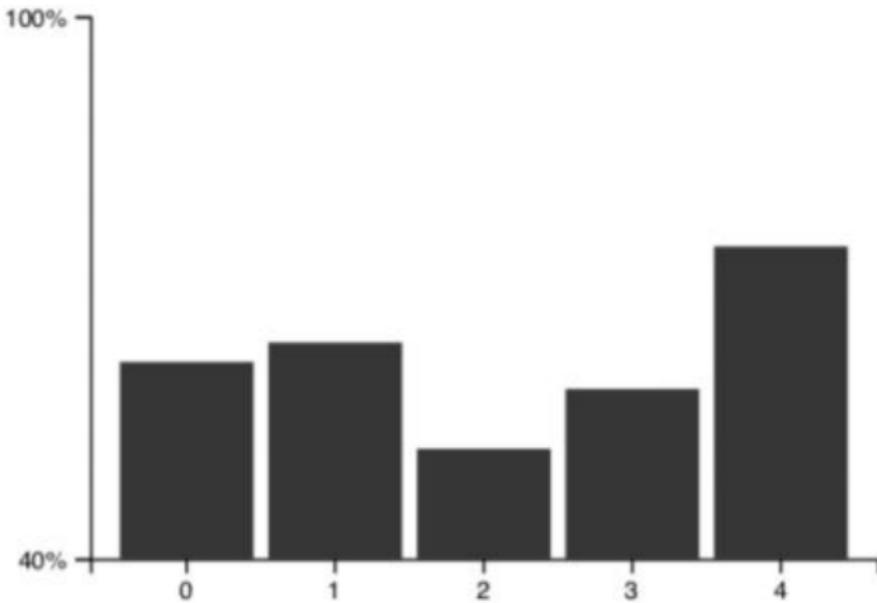
[-ENTERING COMPLEXITY-]

This is good :



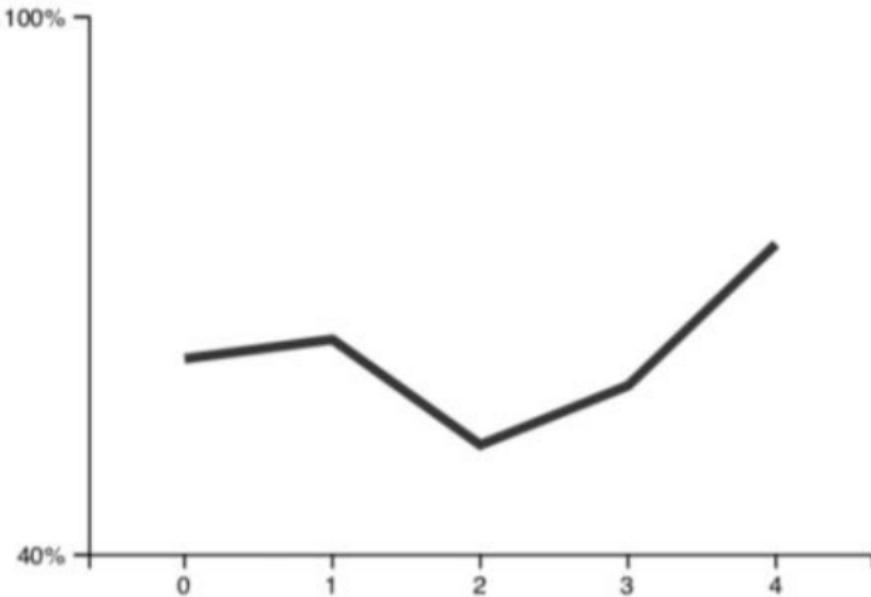
[-ENTERING COMPLEXITY-]

This is also bad :



[-ENTERING COMPLEXITY-]

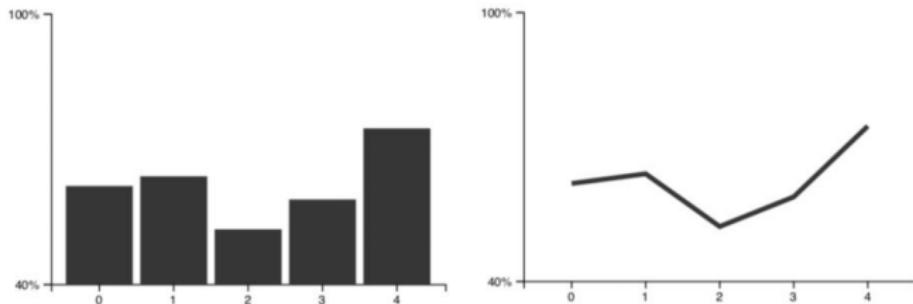
Bad or good?



[-ENTERING COMPLEXITY-]

Correll et al. (2019) show empirically that there is :

“ no robust difference (...) : truncation serves to exaggerate effect sizes in both types of graphs.”



[-ENTERING COMPLEXITY-]

Bad?

Average global temperature by year

Data from NASA/GISS.

60°

59

58

57

56

55

1900

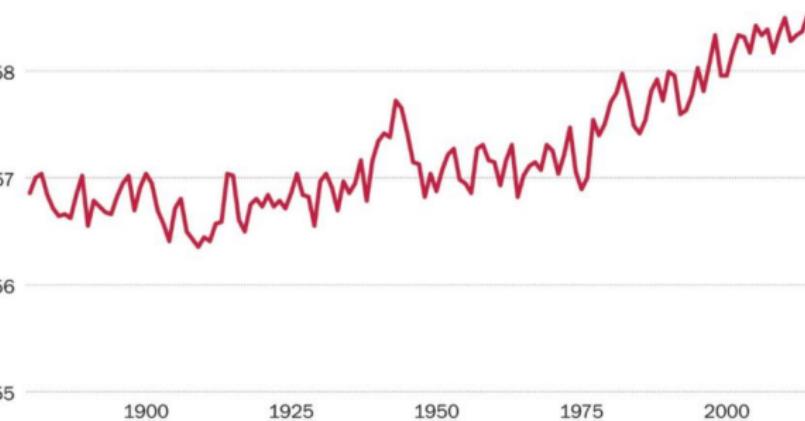
1925

1950

1975

2000

Washington Post (December, 2015)

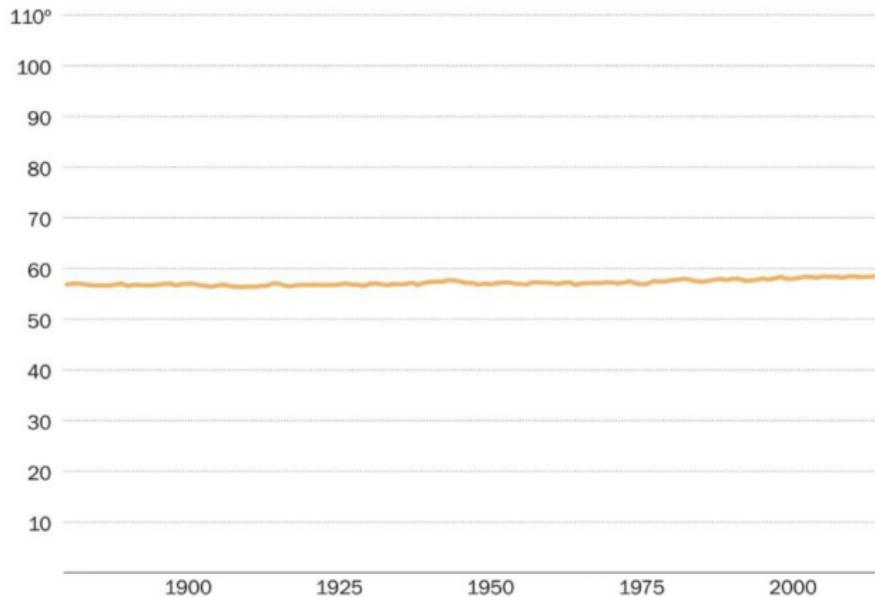


[-ENTERING COMPLEXITY-]

Good?

Average global temperature by year, 50x scale

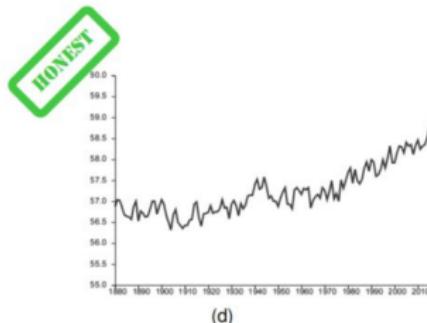
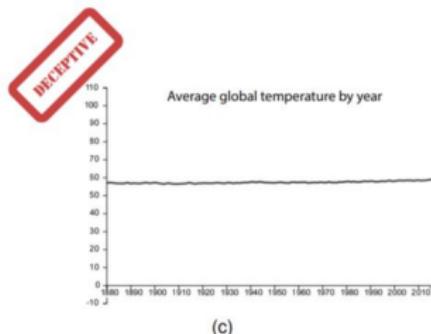
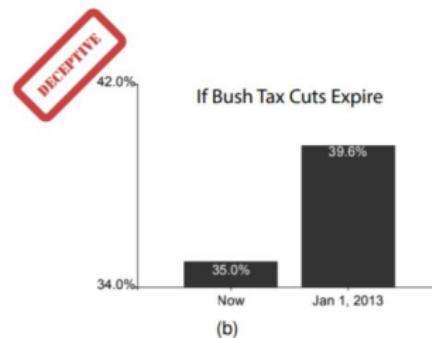
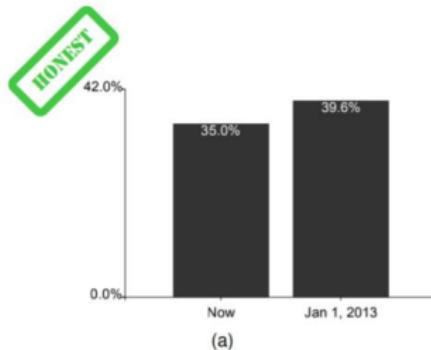
Data from NASA/GISS.



Washington Post, reproducing National Review (December, 2015)

[-HOW (NOT) TO BE MISLEADING ?-]

Ruling is hard!



[-HOW (NOT) TO BE MISLEADING ?-]

Discussed by many :

Cairo (2019),

Correll et al. (2019),

Allen and Erhardt (2016),

Isenberg et al. (2011),

Kelleher and Wagener (2011),

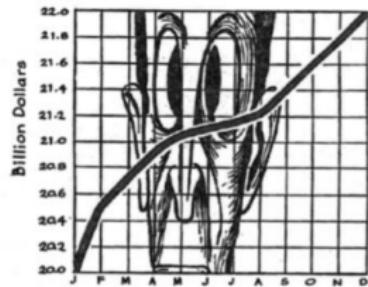
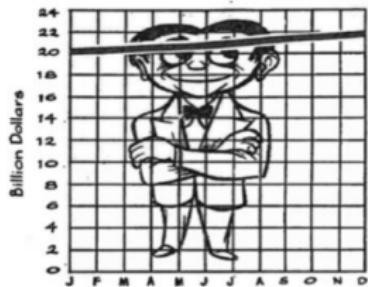
Tufte (2001),

Cleveland et al. (1982)

...

[-HOW (NOT) TO BE MISLEADING ?-]

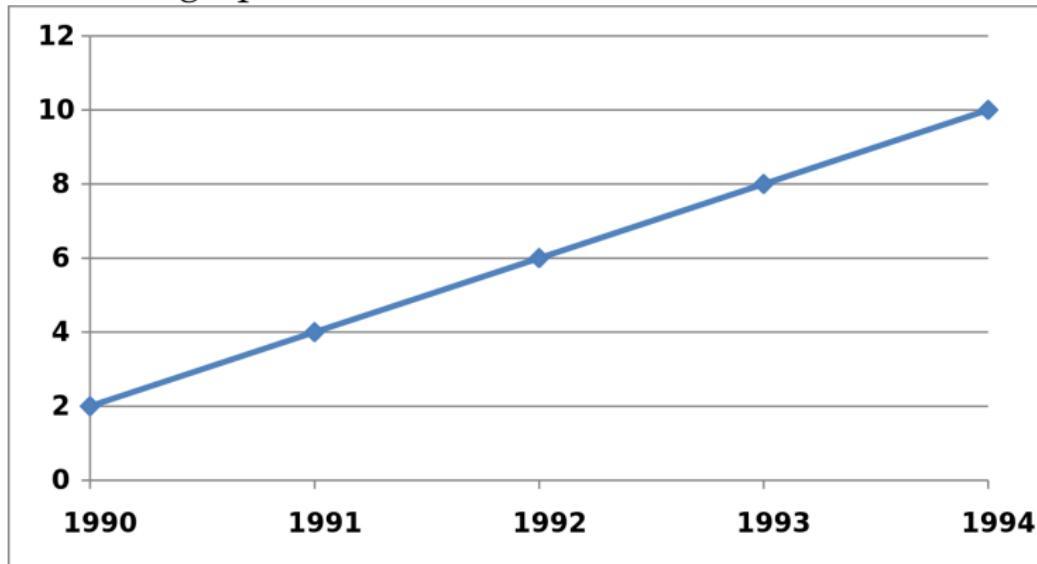
So **lying** is easy :



From Huff (1993)

[-RULE #2 : PLAY WITH SCALES!-]

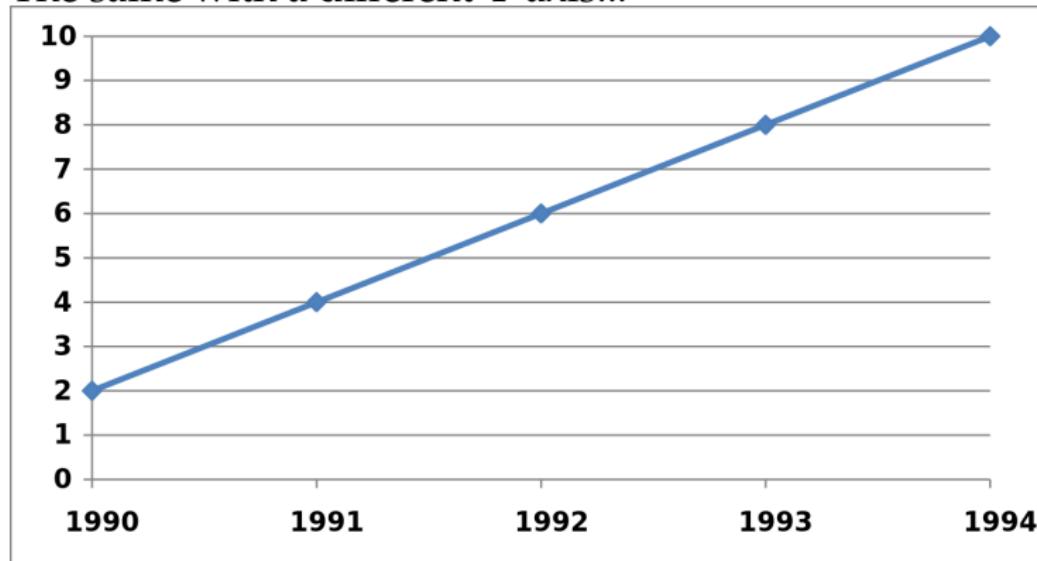
Here is a graphic



Source : Wikipedia: Misleading graph

[-RULE #2 : PLAY WITH SCALES!-]

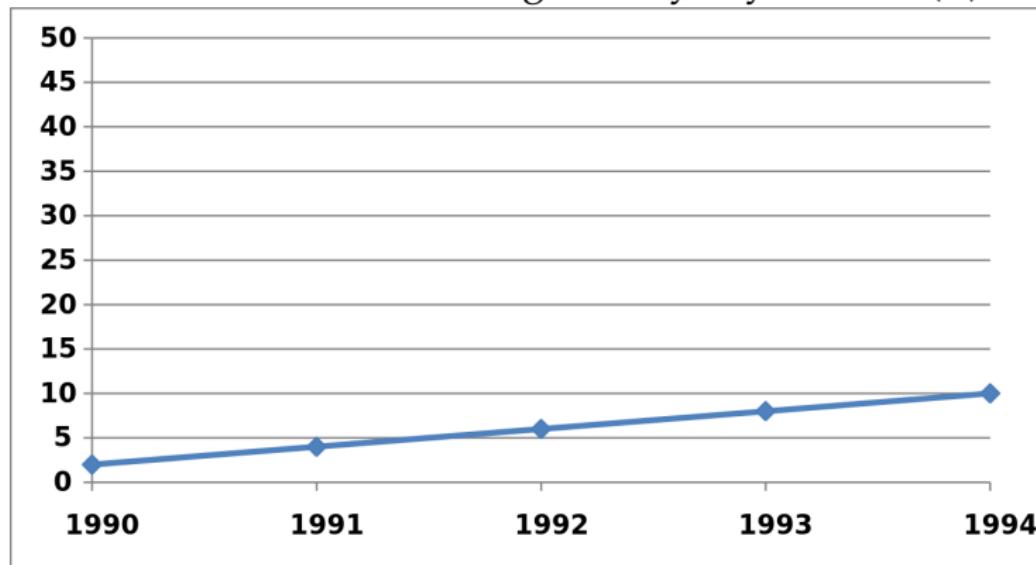
The same with a different Y-axis...



Source : Wikipedia: Misleading graph

[-RULE #2 : PLAY WITH SCALES !-]

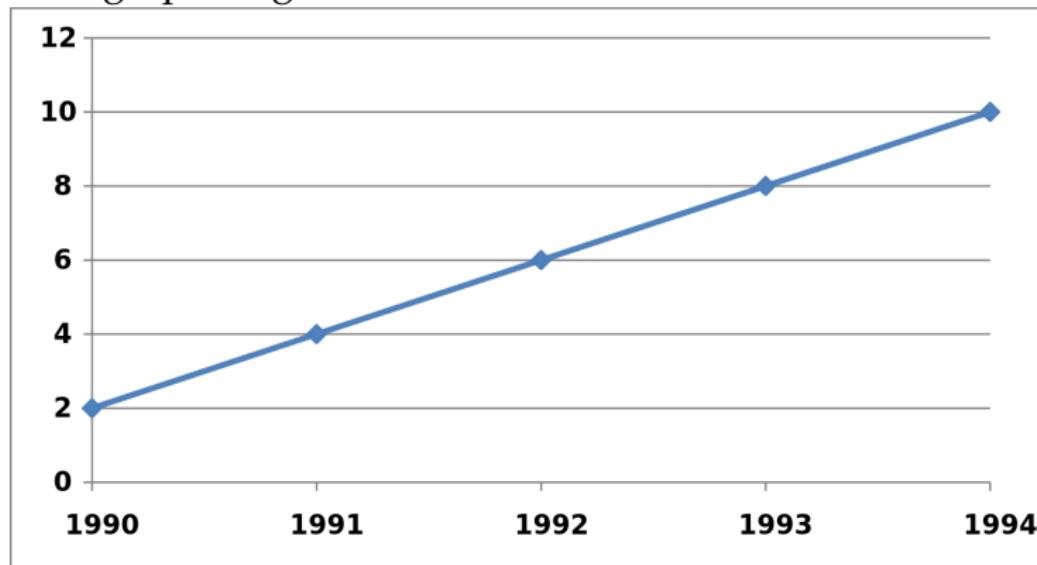
The same with a Y-axis that goes way beyond max(Y)...



Source : Wikipedia: Misleading graph

[-RULE #2-BIS : PLAY WITH SHAPE!-]

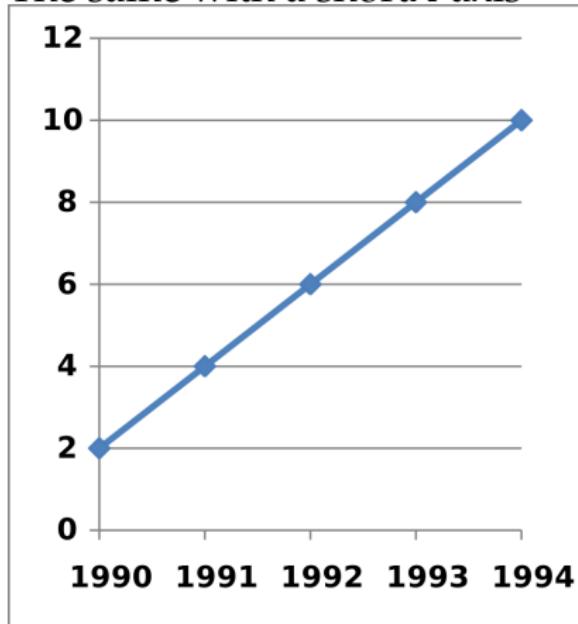
First graphic again



Source : Wikipedia: Misleading graph

[-RULE #2-BIS : PLAY WITH SHAPE!-]

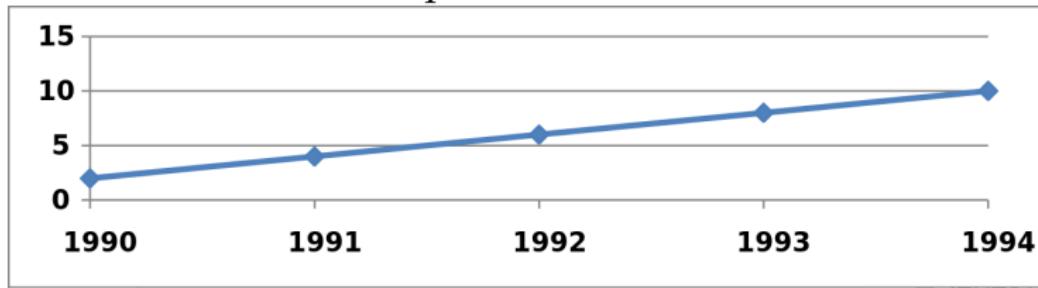
The same with a short X-axis



Source : Wikipedia: Misleading graph

[-RULE #2-BIS : PLAY WITH SHAPE!-]

The same with a widespread X-axis and a short Y-axis



Source : Wikipedia: Misleading graph

[-RULE #2 : EXAMPLE-]

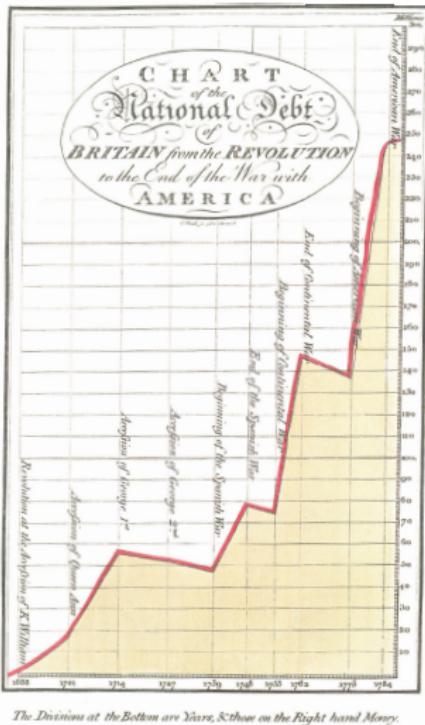


FIGURE – Source : Tufte (2001) from Playfair(1786).

[-RULE #2-BIS : EXAMPLE-]

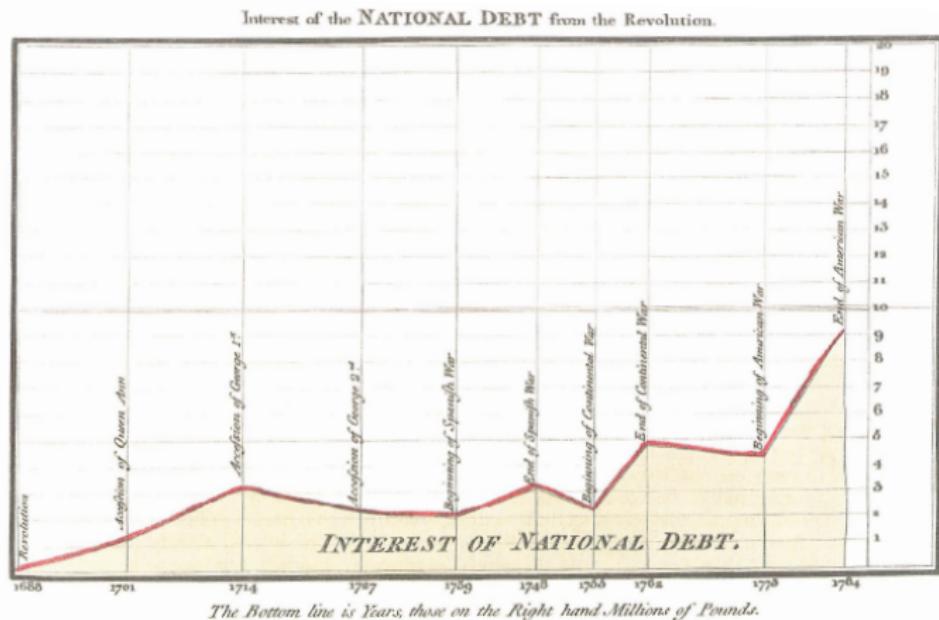
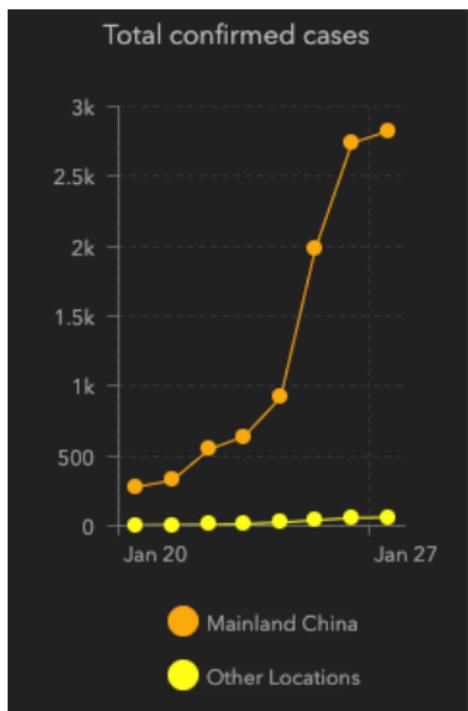


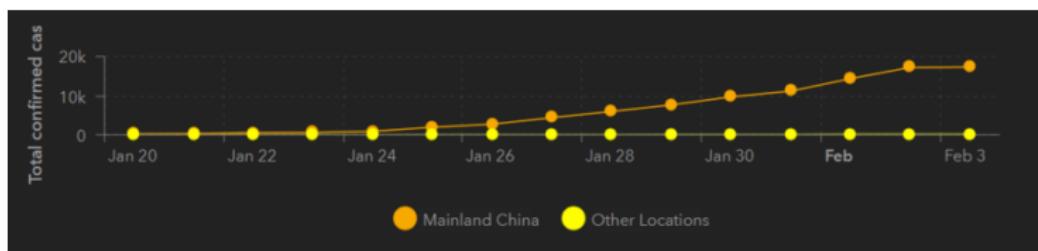
FIGURE – Source : Tufte (2001) from Playfair(1786).

[-APPLYING RULE #2-BIS : CORONAVIRUS-]



Source : Science Alert (2020)

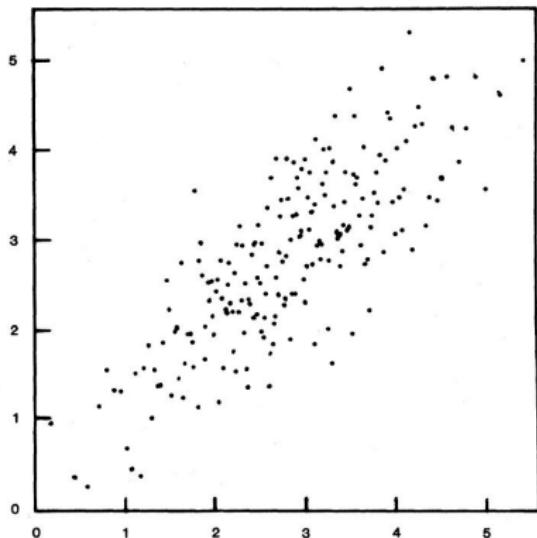
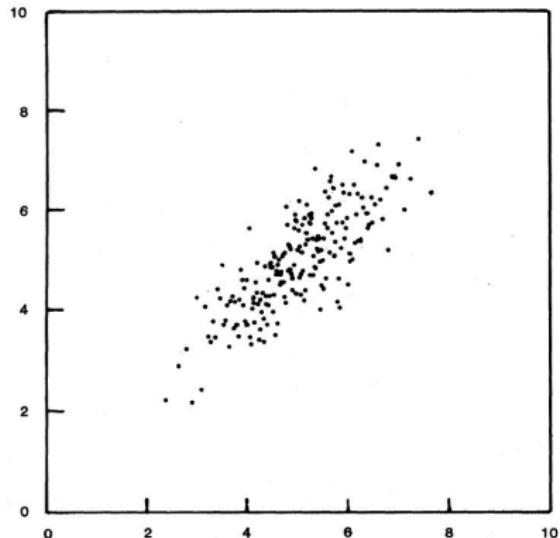
[-APPLYING RULE #2-BIS : CORONAVIRUS-]



Source : John Hopkins CSSE (2020)

[-RULE #2-BIS : SUBTLE THINGS...-]

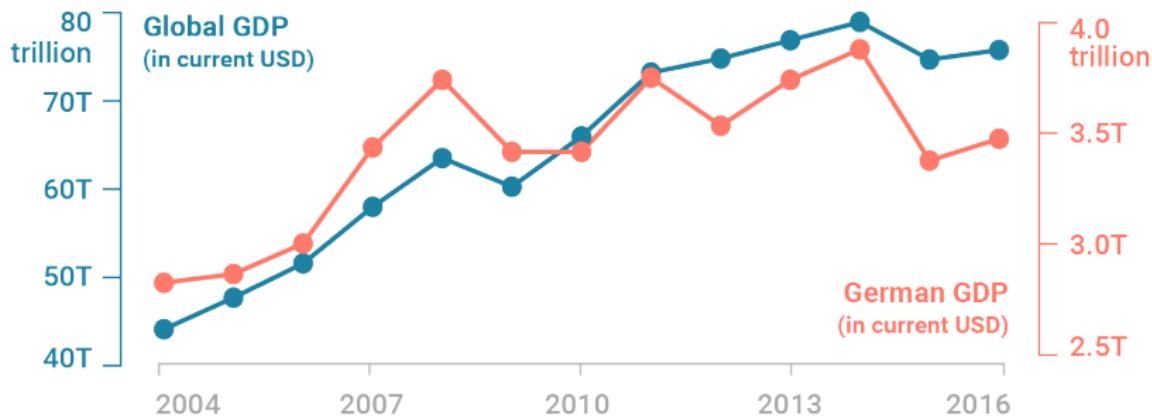
Which sample has the highest correlation?



From Cleveland et al. (1982)

[- RULE #3 : USE DOUBLE AXES! -]

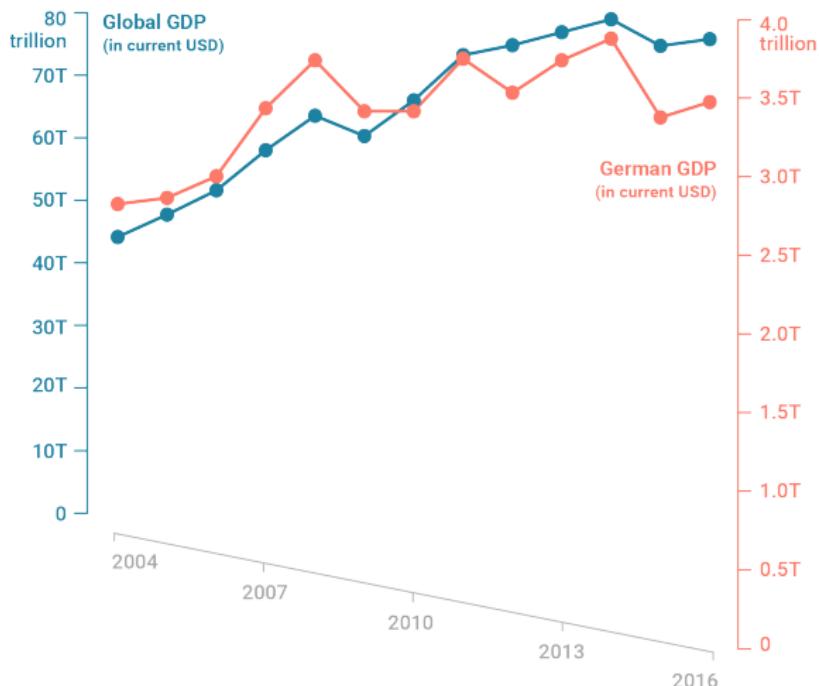
An apparently harmless example :



Source : Lisa Charlotte Rost, data from World bank

[- RULE #3 : USE DOUBLE AXES! -]

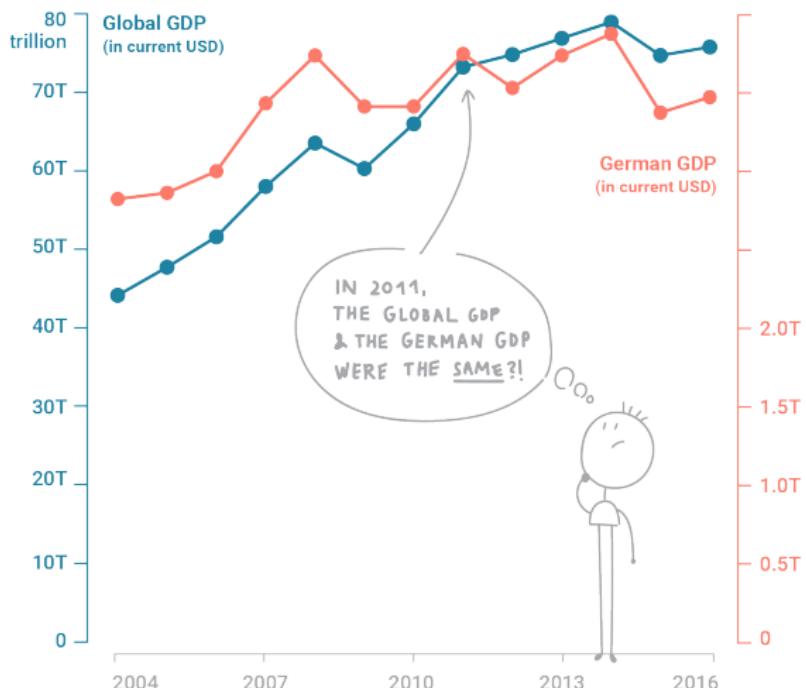
Zero baseline are not aligned!



Source : Lisa Charlotte Rost, data from World bank

[- RULE #3 : USE DOUBLE AXES! -]

Does aligned zero baselines solve the problem?



Source : Lisa Charlotte Rost, data from World bank

[- RULE #3 : USE DOUBLE AXES! -]

When playing with scales, anything may happen...



Orange steady,
Blue massively increasing.



Blue steady,
Orange increasing.



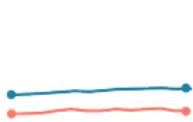
Both started at the same
level, but Orange increased
far more than Blue.



Both started at the same
level, but Blue increased far
more than Orange.



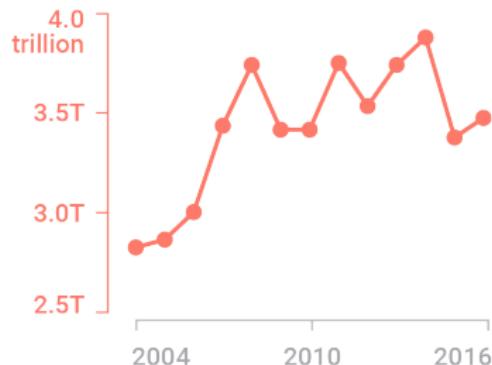
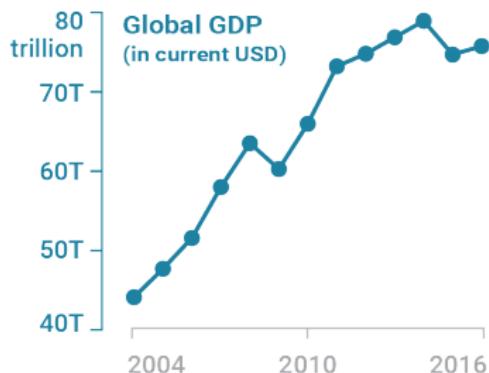
Both started with the
same increase, then Blue
raced to the top.



Both steady.

[- RULE #3 : USE DOUBLE AXES! -]

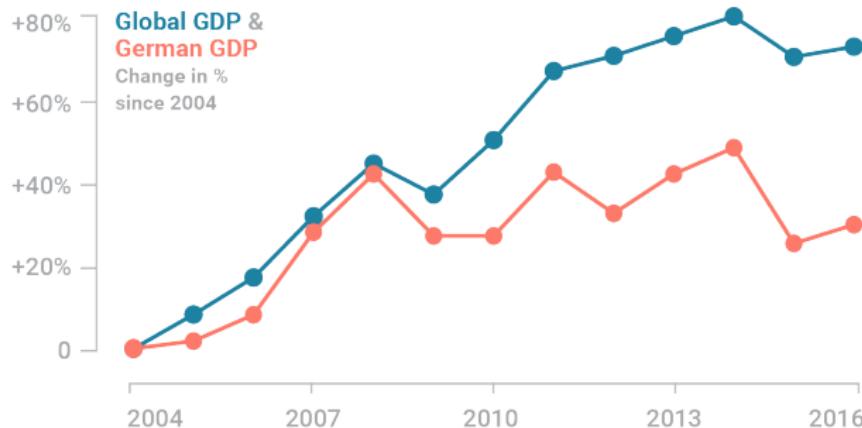
Solution 1 : Two graphs



Source : Lisa Charlotte Rost, data from World bank

[- RULE #3 : USE DOUBLE AXES! -]

Solution : Indexes



The German GDP and the global GDP **are not** growing at the same rate since 2008!

Source : Lisa Charlotte Rost, data from World bank

[- RULE #3 : EXAMPLES -]

Glyphosate vs Senile dementia

Age Adjusted Deaths from Senile Dementia
(ICD F01, F03 & 290)

Plotted against glyphosate use on corn & soy

($R = 0.9942$, $p \leq 1.822e-09$)

Sources: USDA:NASS; CDC

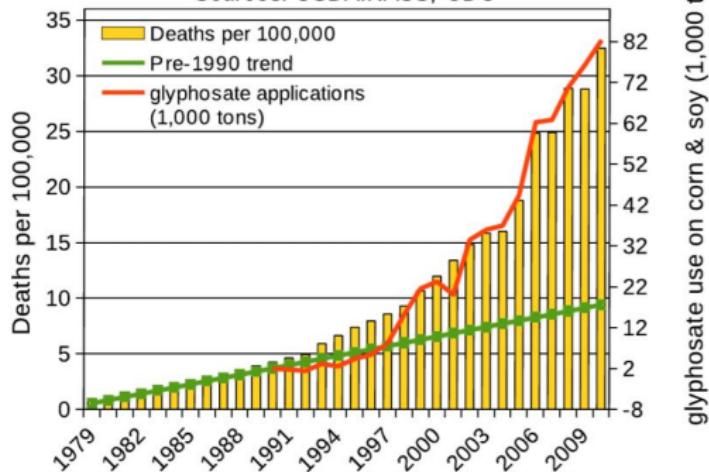


Figure 24. Correlation between age-adjusted dementia deaths and glyphosate applications.

Source : Calling Bullshit- Original paper Swanson et al. Journal of Organic System (2014)

[- RULE #3 : EXAMPLES -]

Glyphosate vs Tyroid

Thyroid Cancer Incidence Rate (age adjusted)

plotted against glyphosate applied to U.S. corn & soy ($R = 0.988$, $p \leq 7.612e-09$)
along with %GE corn & soy crops $R = 0.9377$, $p \leq 2.152e-05$
sources: USDA:NASS; SEER

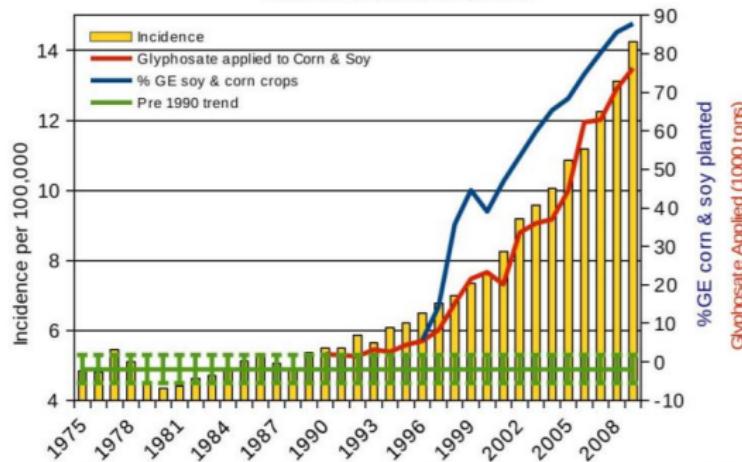


Figure 10. Correlation between age-adjusted thyroid cancer incidence and glyphosate applications and percentage of US corn and soy crops that are GE.

Source : Calling Bullshit- Original paper Swanson et al. Journal of Organic System (2014)

[- RULE #3 : EXAMPLES -]

Glyphosate vs Parkinson

Age Adjusted Deaths from Parkinson's disease
(ICD G20 & 332.0)

plotted against glyphosate use on corn & soy ($R = 0.8754$, $p \leq 1.631e-06$)
and percent GE corn & soy planted ($R = 0.9516$, $p \leq 5.398e-06$)
sources: USDA:NASS; CDC

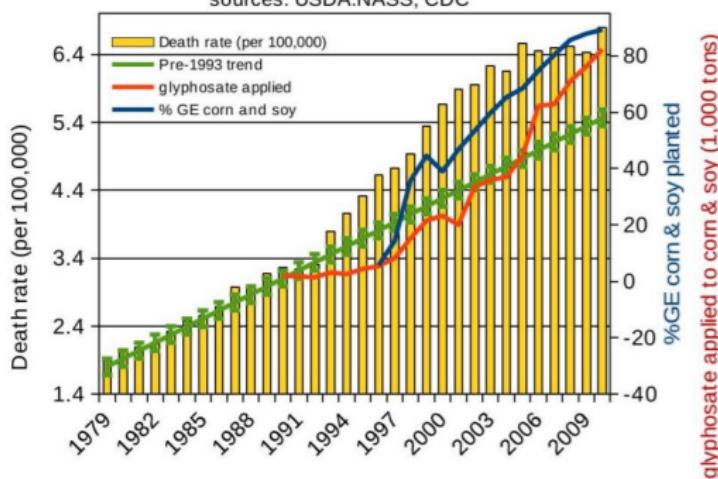


Figure 26. Correlation between age-adjusted Parkinson's disease deaths and glyphosate applications and percentage of US corn and soy crops that are GE.

Source : Calling Bullshit- Original paper Swanson et al. Journal of Organic System (2014)

[- RULE #3 : EXAMPLES -]

Glyphosate vs Diabetes

Annual Incidence of Diabetes (age adjusted)

plotted against %GE corn & soy crops planted ($R = 0.9547$, $p \leq 1.978e-06$)
along with glyphosate applied to corn & soy in US ($R = 0.935$, $p \leq 8.303e-08$)
sources: USDA:NASS; CDC

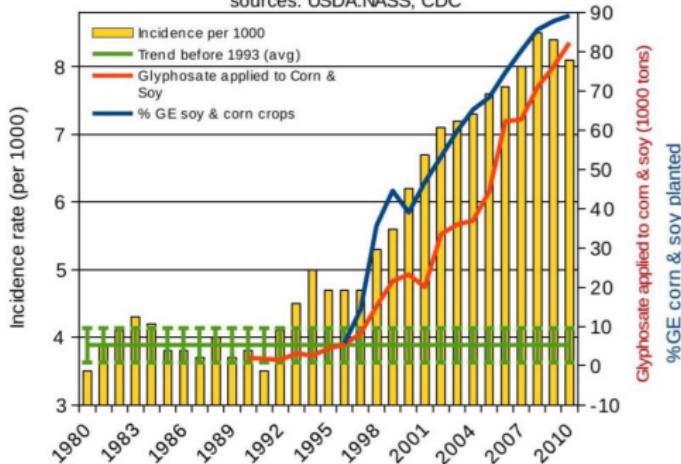


Figure 14. Correlation between age-adjusted diabetes incidence and glyphosate applications and percentage of US corn and soy crops that are GE.

Source : Calling Bullshit- Original paper Swanson et al. Journal of Organic System (2014)

[- RULE #3 : EXAMPLES -]

Glyphosate vs Strokes

Age Adjusted Deaths due to Stroke
(ICD I62.9 & 432.9 hemorrhage, non embolic)

Plotted against %GE corn and soy ($R = 0.9827$, $p \leq 1.354e-06$)
& glyphosate applied to corn and soy ($R = 0.9246$, $p \leq 1.471e-07$)

Sources: USDA:NASS; CDC

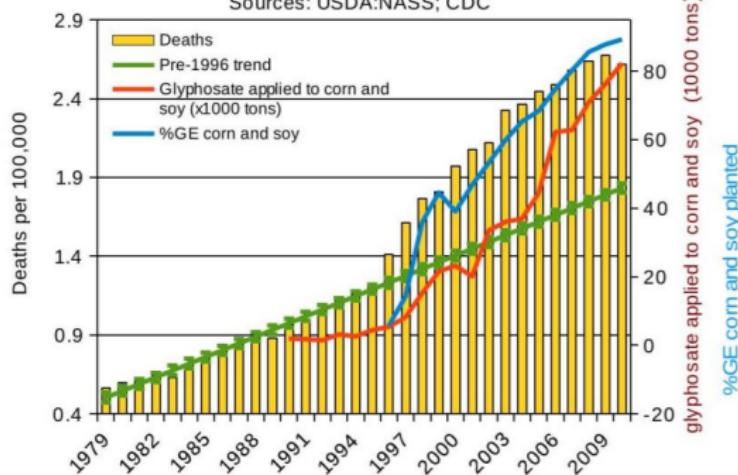


Figure 12. Correlation between age-adjusted hemorrhagic stroke deaths and glyphosate applications and percentage of US corn and soy crops that are GE.

Source : Calling Bullshit- Original paper Swanson et al. Journal of Organic System (2014)

[- RULE #3 : EXAMPLES -]

Glyphosate vs Hypertension

Age Adjusted Deaths due to Hypertension (ICD I10 & 401)

plotted against %GE corn and soy ($R = 0.9607$, $p \leq 3.675e-06$)
& glyphosate applied to corn and soy ($R = 0.923$, $p \leq 1.603e-07$)

Sources: USDA:NASS; CDC

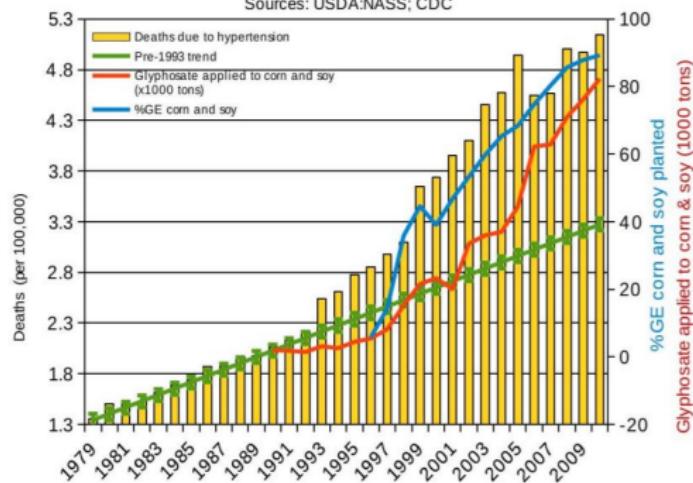


Figure 11. Correlation between age-adjusted hypertension deaths and glyphosate applications and percentage of US corn and soy crops that are GE.

Source : Calling Bullshit- Original paper Swanson et al. Journal of Organic System (2014)

[- RULE #3 : EXAMPLES IN ECONOMICS -]

Assets vs Working Time

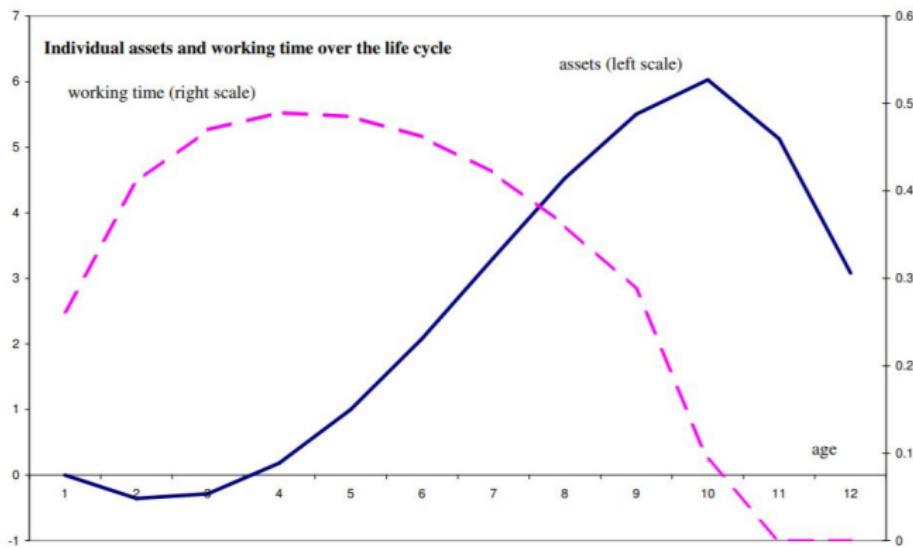
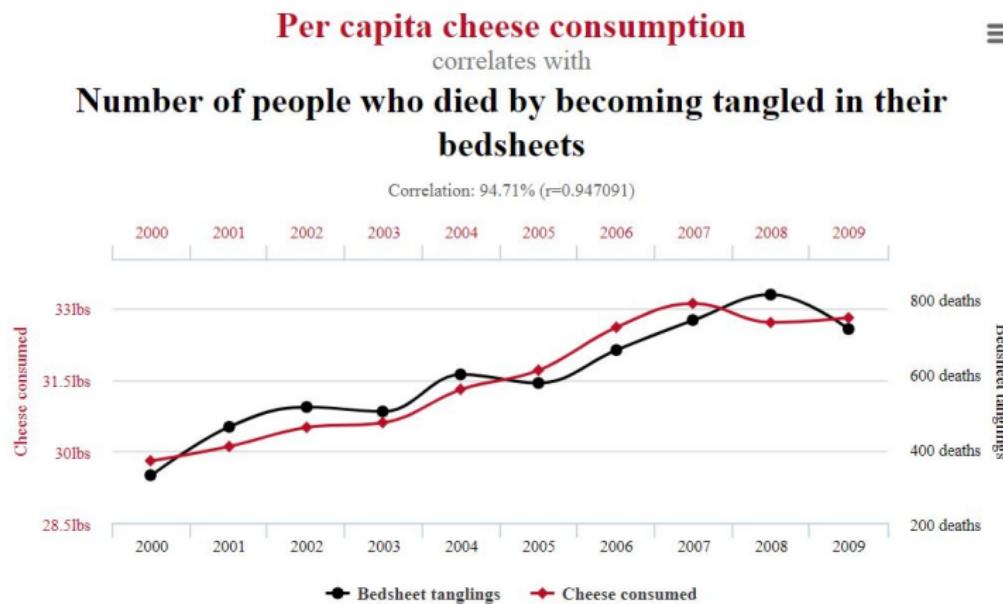


Figure 8. Life-cycle in the benchmark equilibrium.

Source : Mateos-Planas (2010) - AER!

[- RULE #3 : EXAMPLES IN ECONOMICS -]

Nearly as funny :



Data sources: U.S. Department of Agriculture and Centers for Disease Control & Prevention

tylervigen.com

Source : Vigen (2015) - Spurious Correlations

[-RULE #4 : SELECT YOUR SCOPE -]

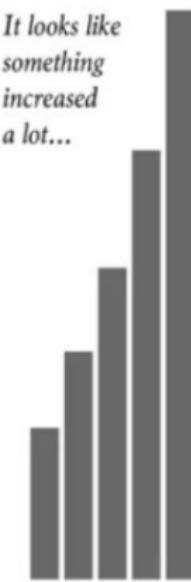


FIGURE – Are you looking at the right thing?

[-RULE #4 : SELECT YOUR SCOPE-]

LIMITED SCOPE

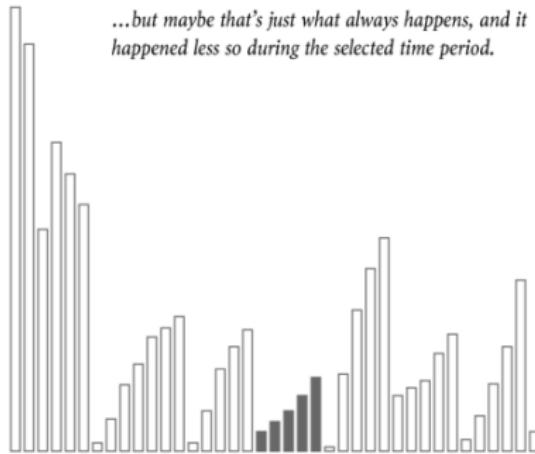
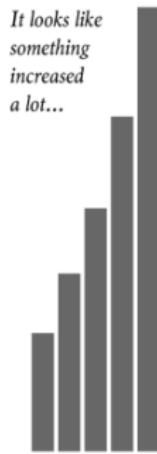


FIGURE – Cherry picking ?

Source : Flowing data

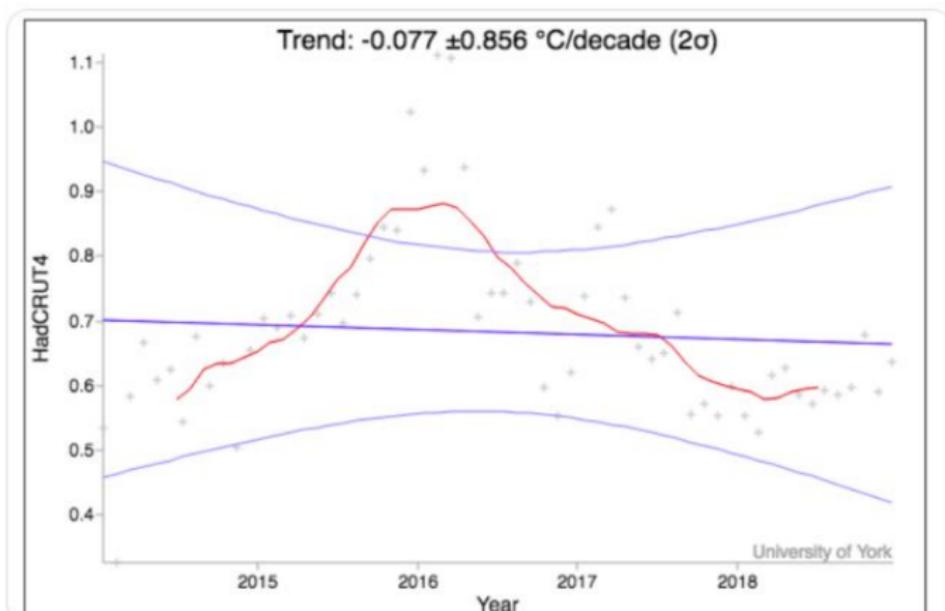
[-APPLYING RULE #4 : GLOBAL WARMING-]



GWPF

@thegwpfcom

Ignore climate hysteria & look at the facts: Global temperature has been falling for the last 3 years.



[-APPLYING RULE #4 : GLOBAL WARMING-]

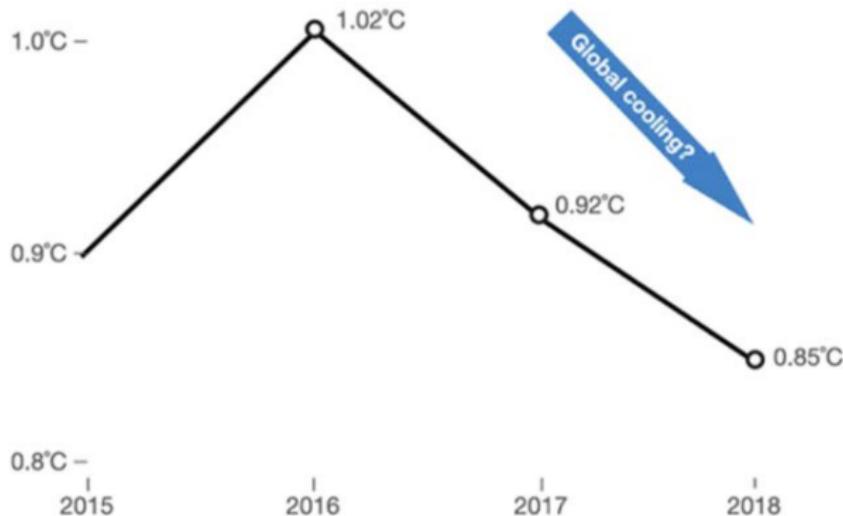


FIGURE – Cherry picking!

Source : "Nope, Earth Isn't Cooling". NASA, 2019.

[-APPLYING RULE #4 : GLOBAL WARMING-]

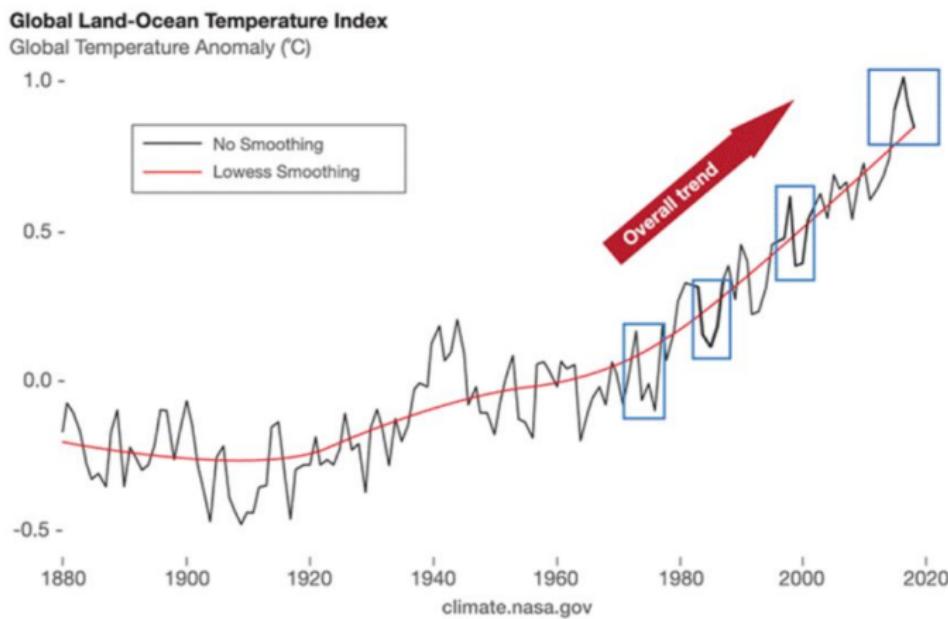


FIGURE – Cherry picking!

Source : "Nope, Earth Isn't Cooling". NASA, 2019.

[-APPLYING RULE #4 : BREXIT!-]

POUND RISES +9% THIS YEAR

Brexit currency confounds doomsayers

Occasionally we look at general economic indicators to keep you informed. Below is the progress of the pound against the dollar this year.



Source : Pro-Brexit site(Facts4eu.org)

[-APPLYING RULE #4 : BREXIT!-]

GBP to USD Chart

3 Feb 2015 00:00 UTC - 2 Feb 2020 22:22 UTC **GBP/USD** close:1.31794 low:1.20224 high:1.58788



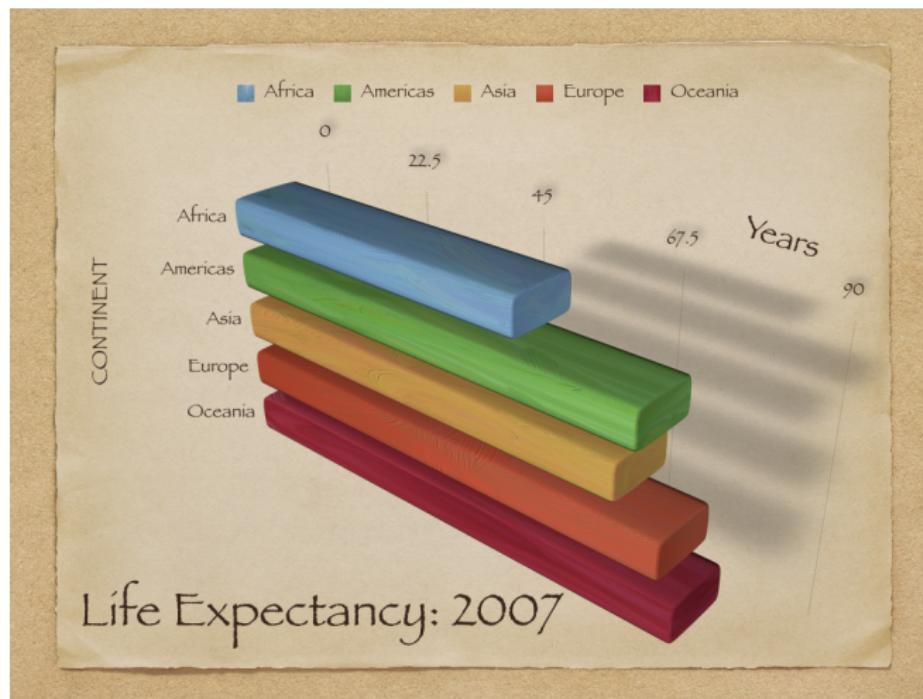
[-APPLYING RULE #4 : BREXIT!-]

GBP to USD Chart

5 Feb 2010 00:00 UTC - 2 Feb 2020 22:22 UTC GBP/USD close:1.31794 low:1.20224 high:1.71602



[- RULE #5A : USE 3D BARS -]



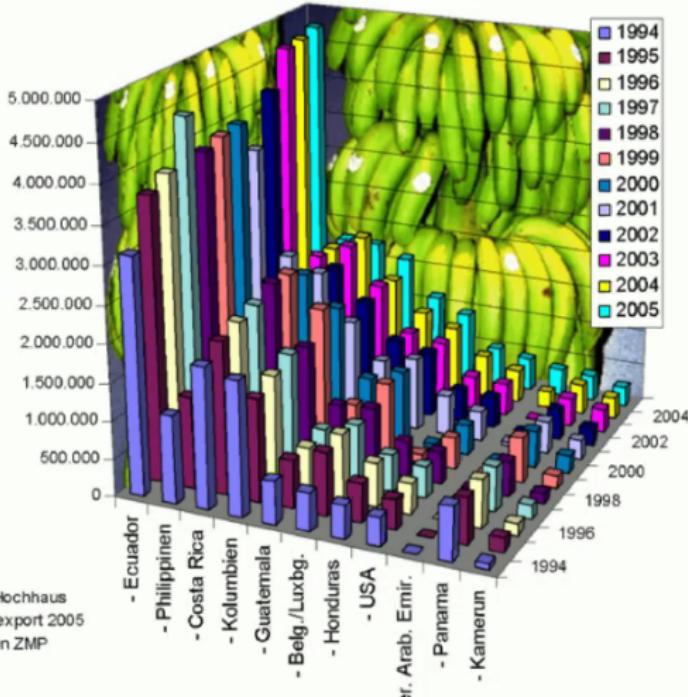
Source Healy

(2019)

[- RULE #5A : USE 3D -]

This is Bananas

Export von Bananen in Tonnen von 1994-2005



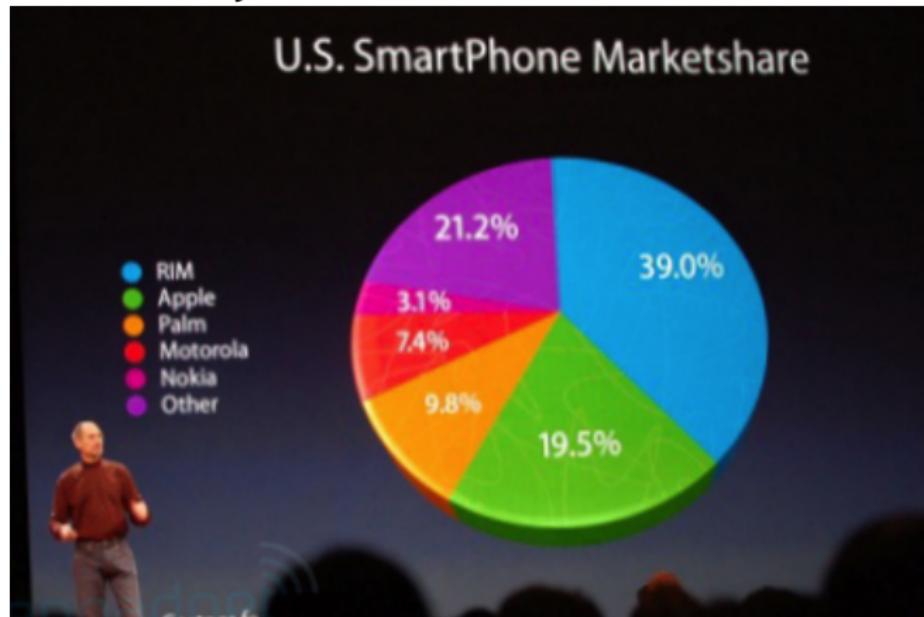
[- RULE #5A : USE 3D -]

Monthly expenses...



[- RULE #5B : USE 3D PIE CHARTS -]

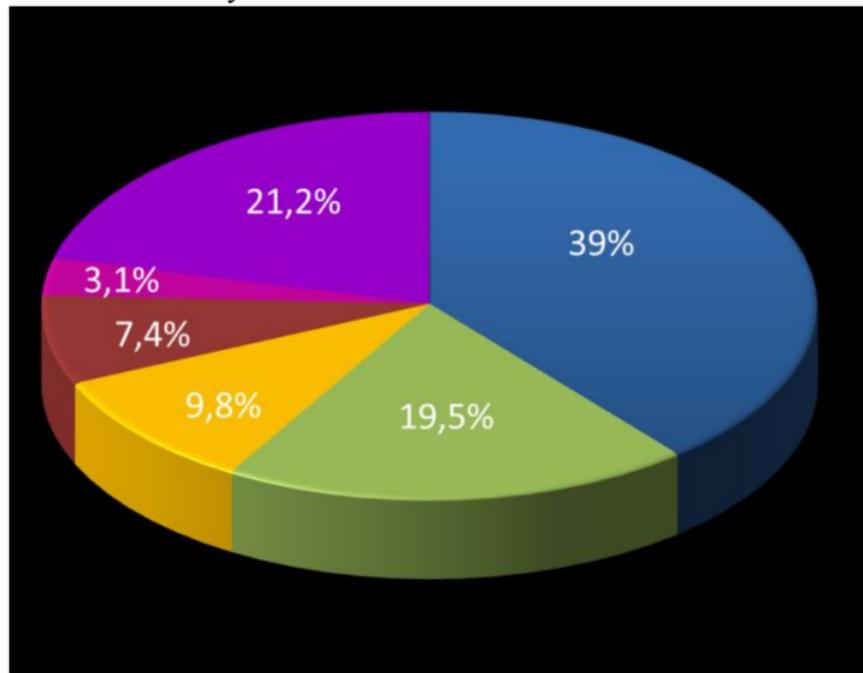
Do as Steve Jobs!



Source (Macworld 2008 keynote lecture)

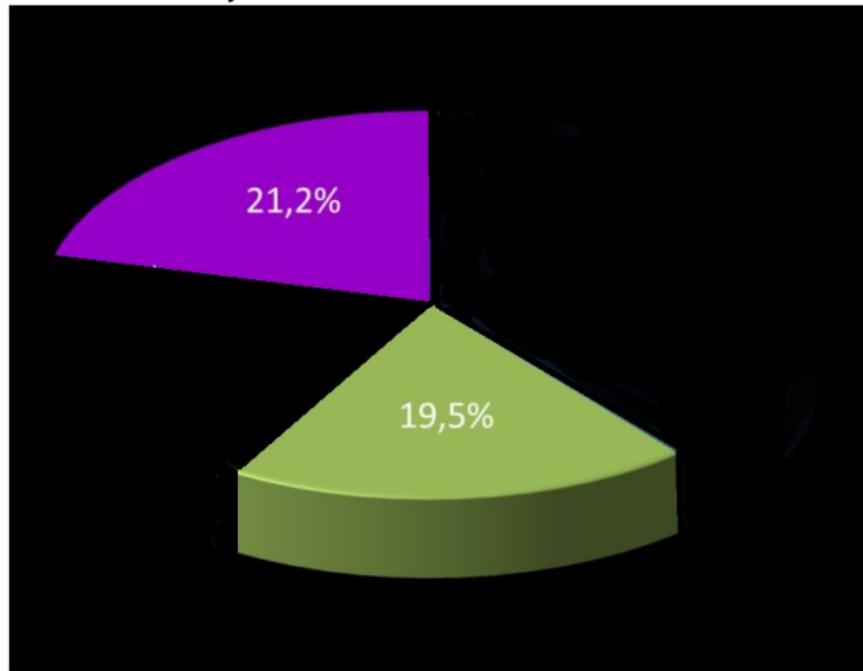
[- RULE #5B : USE 3D PIE CHARTS -]

Do as Steve Jobs !



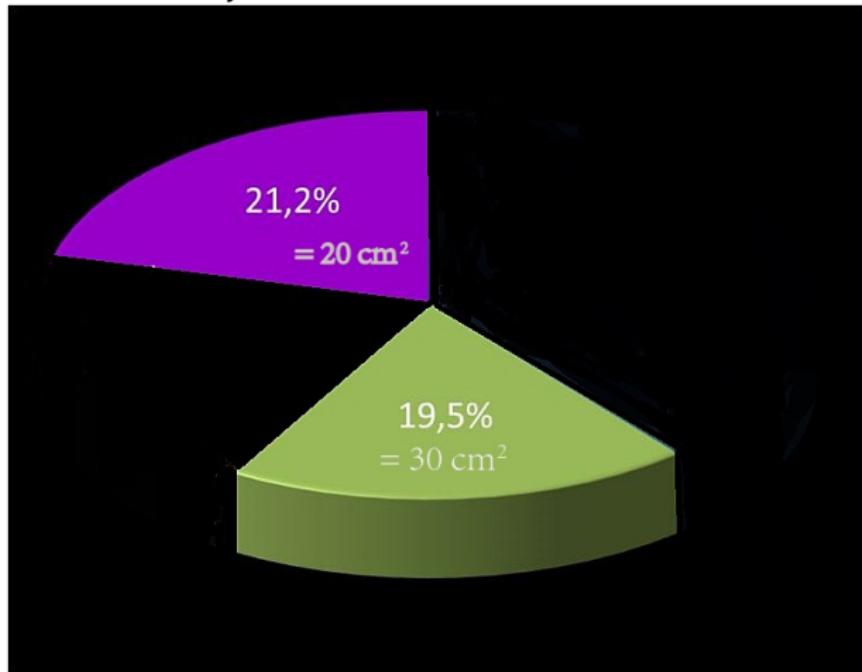
[- RULE #5B : USE 3D PIE CHARTS -]

Do as Steve Jobs !

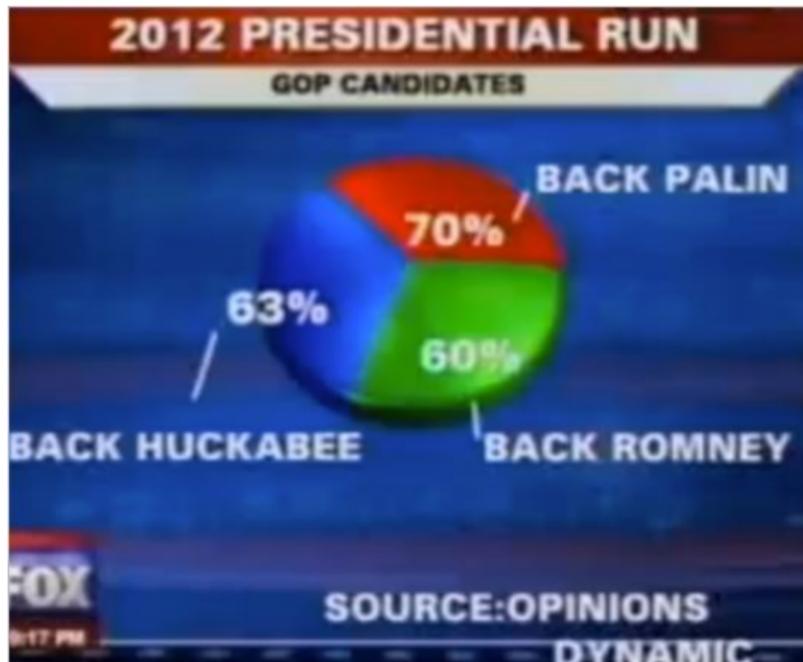


[- RULE #5B : USE 3D PIE CHARTS -]

Do as Steve Jobs : Lie!



[- A NOTE ON PIE CHARTS -]



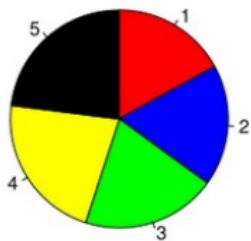
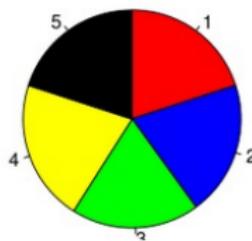
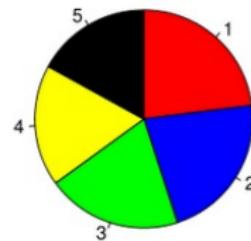
From FOX News

[- A NOTE ON PIE CHARTS -]

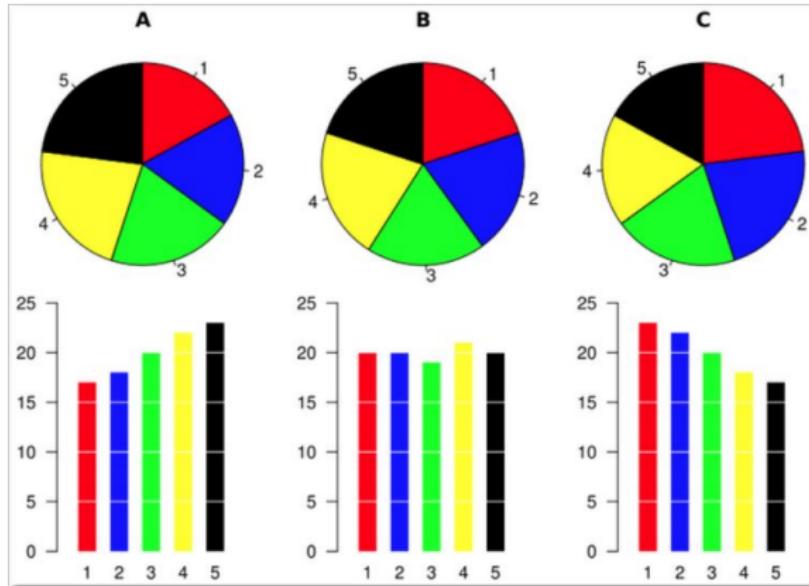


From WCVB (via Twitter)

[- RULE #6 : USE SEVERAL PIE CHARTS -]

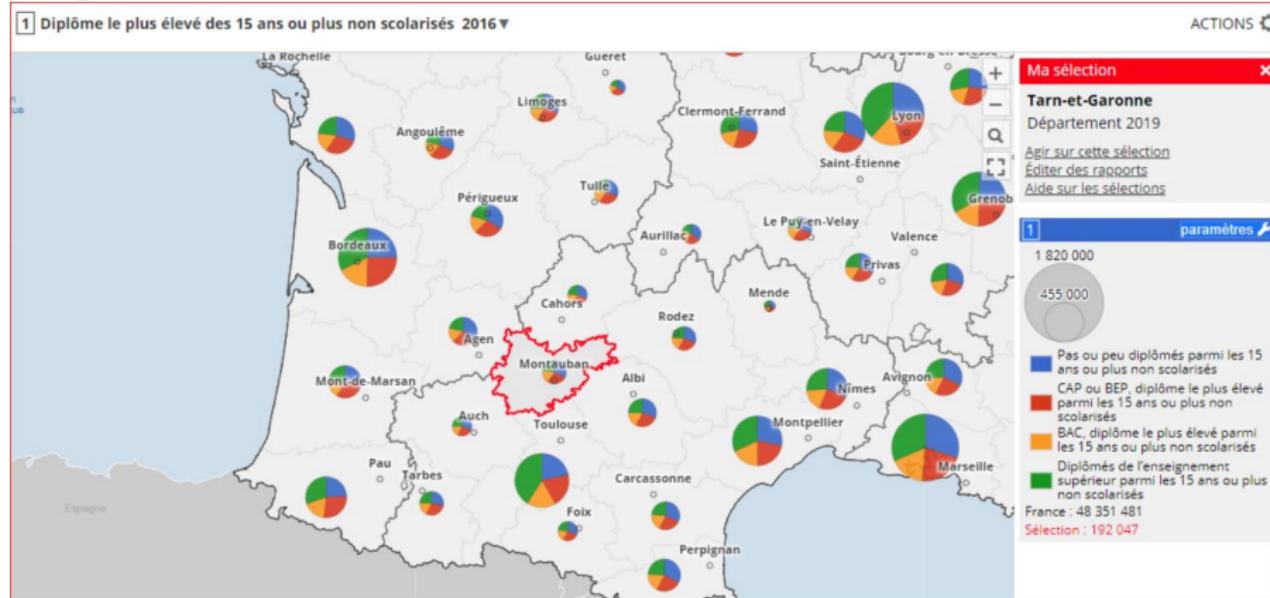
A**B****C**

[- RULE #6 : USE SEVERAL PIE CHARTS -]



[- RULE #6 : USE SEVERAL PIE CHARTS -]

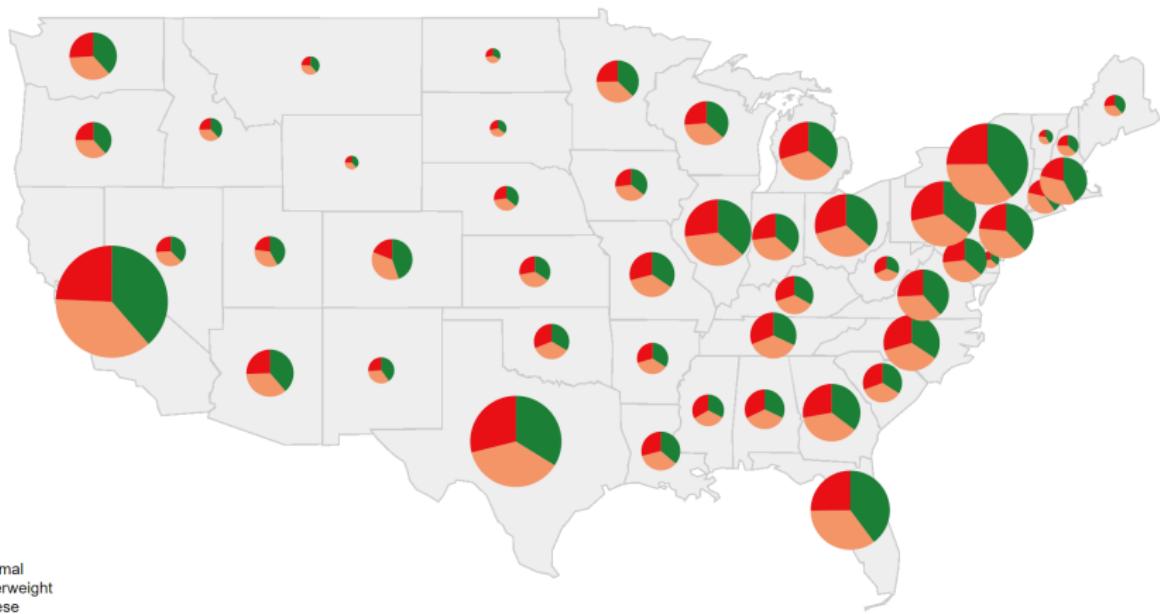
People use it.



Source : INSEE

[- RULE #6 : USE SEVERAL PIE CHARTS -]

People use it.



Source : Anita Graser

[- RULE #7 : USE AREAS -]

Maximize your intakes : One big or two smaller pizzas ?



Source : Fermat's library

[- RULE #7 : USE AREAS -]

One big > two smaller pizzas!



$$\text{Area} = \pi(18/2)^2 = 254 \text{ in}^2$$

$$\text{Area} = 2\pi(12/2)^2 = 226 \text{ in}^2$$

Source : Fermat's library



[- RULE #7 : USE AREAS -]

How to compare circles on maps?

Source : Statista, el pais