How to lie with Graphics?

Christophe Bontemps
Toulouse School of Economics, INRA

@Xtophe_Bontemps

5 mars 2019
[-FOREWORD-] : What’s Data Visualisation?

What do you see?
[-FOREWORD-]: **What’s Data Visualisation?**

And here, what do you see?
“DATA VISUALISATION” AS A STATISTICAL TEST

Buja et al. (2009)
“**DATA VISUALISATION**” AS A STATISTICAL TEST

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

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

Test: $H_0: \{\text{There is "nothing" } \} = \{\text{No relation}\}$
“Data visualisation” as a statistical test

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

**Test**: $H_0 : \{\text{There is "nothing" } \} = \{\text{No relation}\}$

$H_1 : \{\text{There is "something" } \} = \{\text{There is some relation (Correlation, linearity, heterogeneity, groups..) } \}$
"I’ll pause for a moment so you can let this information sink in."

Source: New Yorker
[- Rule #1: Use devilish scales! -]
[-Rule #1: Use devilish (truncated) scales! -]

[-Rule #1: Use devilish (truncated) scales!-]

[-Rule #1 : Use devilish (truncated) scales!-]

Example : Fox news

Source : Techna Verba Scripta
[-Rule #1: Use devilish (truncated) scales!-]

Example: SNCF (2018)

Source: Les Décodeurs (le Monde)
[**-Rule #1 : Use devilish (truncated) scales!-**]

**Example : SNCF - Corrected (2018)**

**COMMUNIQUÉ DE PRESSE**

**TAUX DE PARTICIPATION À LA GRÈVE DU 18 AVRIL 2018**

Pour la journée du mercredi 18 avril, sur les cheminots devant travailler aujourd'hui (1), le taux de grévistes en milieu de matinée s'établit à 19,84% : 4 cheminots sur 5 travaillent aujourd'hui.

![Graph showing participation rates](image-url)

(1) Les cheminots « devant travailler » ne sont pas forcément ceux qui devaient le faire hier, compte tenu des repos et des congés.

**Source :** Les Décodeurs (le Monde)
[RULE #1 BIS: CHANGE THE SCALE/UNITS -]

Distribution of started and completed steps

Source: My 2017’ students (Jan & Mohamed), but also recent researchers’ presentations
[- Rule #1 bis : Change the scale/units -]

**Zoom on X axis :**

Source: My 2017’ students (Jan & Mohamed), but also recent researchers’ presentations
[Rule #1 bis: Reverse the scale/units -]

Gun deaths in Florida

Number of murders committed using firearms

Source: Florida Department of Law Enforcement
C. Chan 16/02/2014
[- RULE #1 BIS : REVERSE THE SCALE/UNITS -]

Gun deaths in Florida

Number of murders committed using firearms

Source: Florida Department of Law Enforcement

C. Chan 16/02/2014 Inverted 11/01/2017 J. Sabin
[- RULE #1TER : USE DOUBLE AXES! -]

An apparently harmless example:

Source: Lisa Charlotte Rost, data from World bank
[- **RULE #1TER** : **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.

Source: Lisa Charlotte Rost, data from World bank
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 #1TER**: USE DOUBLE AXES! -]

Example: Americans United for Life

![Graph]

**PLANNED PARENTHOOD FEDERATION OF AMERICA:**

**ABORTIONS UP – LIFE-SAVING PROCEDURES DOWN**
[- **RULE #1TER**: USE DOUBLE AXES! -]

The “correct” curves comparing Abortion & Cancer Prevention Services

Source: Linda Qiu (Polитifact)

See also tons of funny examples on *Spurious Correlation Website*
[Rule #2: Select your scope -]

Figure – Are you looking at the right thing?
[**Rule #2 : Select your scope**-]

**LIMITED SCOPE**

*It looks like something increased a lot…*

...but maybe that’s just what always happens, and it happened less so during the selected time period.

**FIGURE – Cherry picking?**

Source: Flowing data
[- Rule #3 : Use 3D pie charts -]

Do as Steve Jobs!

(Source: Macworld 2008 keynote lecture)
Do as Steve Jobs!
[- **Rule #3 : Use 3D pie charts -**]

Do as Steve Jobs!
[RULE #3: USE 3D PIE CHARTS -]

Do as Steve Jobs: Lie!

- 21.2% = 20 cm²
- 19.5% = 30 cm²
[- **Rule #4: Use pie charts** -]

From Freakeconometrics
[- **Rule #4: Use pie charts** -]
[- RULE #4 : USE PIE CHARTS -]

From WTF Visualisations
[- Rule #4 : Use pie charts -]

From FOX News
[ Rule #4: Use Pie Charts - ]

How to make a pie chart if your percentages don’t add up to 100

From XKCD
Maximize your intakes: One big or two smaller pizzas?

Source: Fermat’s library
[- Rule #5 : Use areas -]

One big > two smaller pizzas!

\[
\text{Area} = \pi \left( \frac{18}{2} \right)^2 = 254 \text{ in}^2 \\
\text{Area} = 2\pi \left( \frac{12}{2} \right)^2 = 226 \text{ in}^2
\]

Source: Fermat’s library
If 100% of the US prisoners are represented by the big green square...what is the percentage for each group?

Figure – Ethnic composition of prisoners in Jail in 2008 in the USA. (Le Monde 5/12/2014)
[- RULE #5 BIS : WORKS ALSO WITH SQUARES-]

If 100% of the US prisoners are represented by the big green square...what is the percentage for each group?

**Figure** – Ethnic composition of prisoners in Jail in 2008 in the USA. *(Le Monde 5/12/2014)*
[- Rule #6 : Use [unaligned] bars -]

Source: Cleveland and McGill (1984)
[**Rule #6**: *Use [unaligned] bars*]

Source: Cleveland and McGill (1984)
[- Rule #6: Use Stacked BarCharts -]
[Solution #6: Align the bars! -]

FIGURE – From Dix and Ellis (1998) example

See also the dynamic version
[ - Rule #6 : Use Stacked barcharts - ]

Kind of stupid example:

Are you willing to pay...?

Number of people

<table>
<thead>
<tr>
<th>Euros</th>
<th>6000</th>
<th>5000</th>
<th>4000</th>
<th>3000</th>
<th>2000</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
<td>[20]</td>
<td>[40]</td>
<td>[60]</td>
<td>[80]</td>
<td>[100]</td>
<td>[120]</td>
</tr>
<tr>
<td>YES</td>
<td>[40]</td>
<td>[20]</td>
<td>[20]</td>
<td>[20]</td>
<td>[20]</td>
<td>[20]</td>
</tr>
</tbody>
</table>

Source: The TSEconomics Journal (TSE)
[- Rule #7: Use lines (lots of them!) -]

**Figure** – Major Cause of Worker Disability (1975-2010) (J. Schwabish, 2014).
LEGITIMATE QUESTIONS:

▶ Is cancer (red curve) increasing over the period?
LEGITIMATE QUESTIONS:

▶ Is cancer (red curve) increasing over the period?
▶ In the recent years, which causes have increased (decreased) the most?
**LEGITIMATE QUESTIONS:**

- Is cancer (red curve) increasing over the period?
- In the recent years, which causes have increased (decreased) the most?
- . . .
LEGITIMATE QUESTIONS:

- Is cancer (red curve) increasing over the period?
- In the recent years, which causes have increased (decreased) the most?
- ... 
- You do not remember a damn thing of this graph!
[- **Solution #7 : Use small multiples -]**

Initial DI Worker Awards by Major Cause of Disability—Calendar Years 1975–2010
(Percent)

- Circulatory
  - 32
  - 11

- Mental
  - 23

- Musculoskeletal
  - 17
  - 26

- Cancer
  - 10
  - 14

**Figure – Major Cause of Disability - 1975-2010 (J. Schwabish, 2014).**
[- Rule #7bis : Use lines (Compare) -]
[- Rule #7bis : Use lines (Compare) -]
[- Rule #7bis : Use lines (Compare) -]

From Cleveland and McGill (1984)
[- Rule #7: Use radar plots (or Web plots) -]
[- Rule #7 : USE RADAR PLOTS -]

![Radar plots for different scores](image)

Source: Xtophe’s blog: Why you should never use radar plots
[- **RULE #7 : USE RADAR PLOTS** -]

Source: Xtophe’s blog: Why you should never use radar plots
[- SUPER RULE #1 : USE MAPS -]

True size of countries

Source: http://metrocosm.com/mercator/, see also https://thetruesize.com/
"Welcome" version

REFUGEES WELCOME

Number of Syrian exiles welcomed in European countries

- 69199
- 31042
- 7978

Source: Françoise BAHOKEN & Nicolas LAMBERT
"Less welcome" version

THE BIG INVASION

Number of Syrian illegal immigrants

69199
31042
7978
6

Source: Françoise BAHOKEN & Nicolas LAMBERT
"Not welcome" version

Source: Françoise BAHOKEN & Nicolas LAMBERT
"Invasion" version

Natural Earth & United Nations, 2018
Lambert & Bahoken, 2018

Source: Françoise BAHOKEN & Nicolas LAMBERT
"Relative to population" version

Source: Françoise BAHOKEN & Nicolas LAMBERT
"All of this is a lie!" The picture is zoomed!

Source: Françoise BAHOKEN & Nicolas LAMBERT
Be suspicious of maps, not of migrants!

Source: Françoise BAHOKEN & Nicolas LAMBERT
[- **Super Rule #3 : Use Electoral Maps** -]

More on electoral maps

Source: Ken Field
[- SUPER RULE #4 : EXTRAPOLATE -]

An example: The gender gap in 100-meters Olympics

Adapted from: Calling Bullshit
[Super Rule #4: Extrapolate]

Let us reduce the y-axis scale

Adapted from: Calling Bullshit
Now, let us draw a regression line

Adapted from: Calling Bullshit
[- SUPER RULE #4 : EXTRAPOLATE -]

Extrapolate: Women will run faster in 2156!

Adapted from: Calling Bullshit
[- Super Rule #4 : Extrapolate -]

No journal will publish this!

No journal will publish this!

No journal will publish this!

[- Super Rule #4 : JUST FOR FUN -]

Reductio ad absurdum : Move backwards!

Adapted from : Calling Bullshit
[ - Super Rule #4: Just for Fun - ]

Reductio ad absurdum: Move forward!

Adapted from: Calling Bullshit
Curves and lines influence your senses

Source: Xtophe Bontemps
Curves and lines influence your senses

Source: Xtophe Bontemps
[ - **SUPER RULE #4 : ABOUT CURVE FITTING** - ]

It is just a 2D random point cloud...

Point plot ( 500 points)
(Point size = 2 , Alpha transparency = 1 )

Source : Xtophe Bontemps
There are many visualisation that transform the data for clarity: **Subway maps** for example.

Source The Guardian
There are many visualisation that transform the data for clarity: Subway maps for example.

Source: The Guardian
Lying for a good reason

Subway maps that match the physical reality are quite rare

Source Benjamin Schmidt
Ski resort maps

Source Pierre Novat
[- THE TRUTH... -]

please consider this before talking/typing
Keep in Touch!

Christophe
http://data.visualisation.free.fr
@Xtophe_Bontemps

Keep in Touch!

Christophe

http://data.visualisation.free.fr

@Xtophe_Bontemps
[- SUPER RULE #3 : USE ELECTORAL MAPS -]

USA election : 2016 results as a map : Washington Post
USA election: 2012 results as a map: New York Times

The blue states reflect a total of 332 electoral votes for Barack Obama.
[Super Rule #3: Use Electoral Maps]

USA election: 2008 results as a map: New York Times

The blue states reflect a total of 365 electoral votes for Barack Obama.
[\textbf{\textit{SUPER RULE #3 : USE ELECTORAL MAPS -}}]

Maybe a better map? From \textit{Financial Times blog}

From \textit{Financial Times}
[- Super Rule #3 : Use Electoral Maps -]

Maybe spatial information is not the most relevant!
[Super Rule #3: Use Electoral Maps -] Maybe spatial information is not the most relevant! (Back to maps)