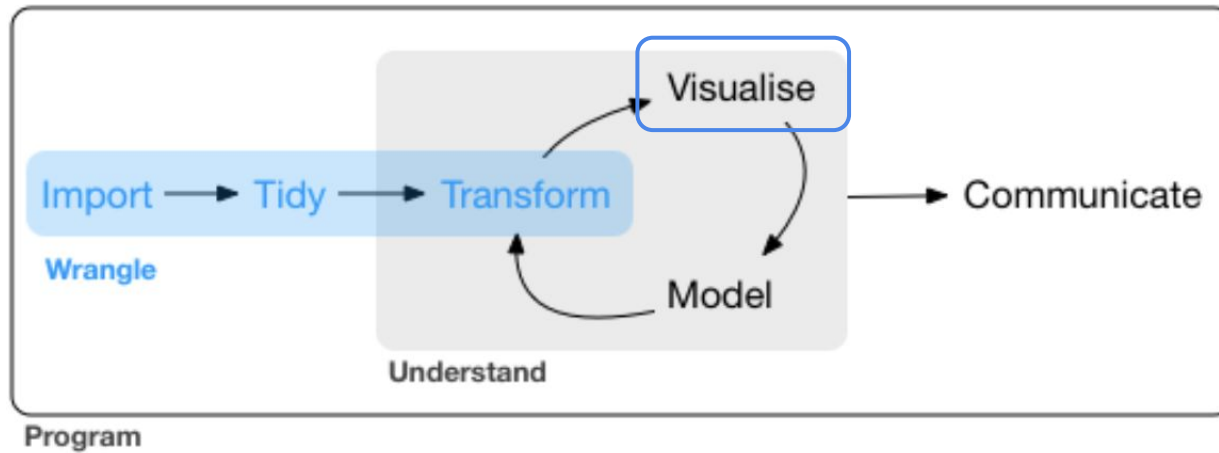


Data Visualization

Aurélien Berra & Matteo Romanello

Sunoikisis Digital Classics 2020



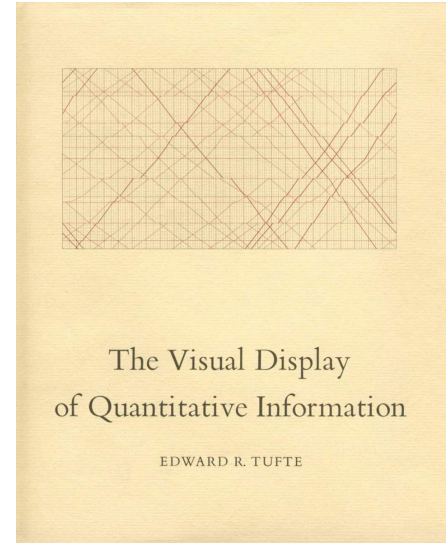
Tufte lore

“Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space.”

“What is to be sought in designs for the visual display of information is the clear portrayal of complexity.”

Edward Tufte, *The Visual Display of Quantitative Information*, 2nd ed., 2001

Aim at **data density**, avoid **chartjunk**!



Tufte 1983. *The visual display of quantitative information.*

Purpose(s)

Data visualization aims to make **quantitative** data **legible** and **intelligible**:

- for oneself: data exploration (know your data!)
- for others: communication of results

Data visualizations tell (potentially fake) stories.

All parts of the process, from the creation of quantified information to producing visualizations, are acts of interpretation: ***data are capta***.

Purpose(s)

Visualisation can and should be weaved into the data analysis process: figures are steps and tools in an iterative **exploratory data analysis**.

The automated production of figures is important for the sake of **re-use** and **reproducibility** (strong ethical and deontological issues).

Basic principles

The graphic qualities of visualizations have a strong rhetorical force, and different types of viz are different modes of expression.

Some visualizations are more suitable to a given dataset than others. In some cases a table works better than a diagram.

Keep it simple. Break down complexity. Small multiples are useful when comparison matters.

Keep it true. Use surfaces and lengths proportional to the quantity represented. Scales can lie.

A grammar of graphics

Visualization is about mapping data values onto aesthetics, i.e. dealing efficiently with variables (things for which you have data) and aesthetics (the visible dimensions of your representation).

Aesthetics can be: position, shape, size, color, line width, line type

Variables can be: quantitative/numerical continuous, quantitative/numerical discrete, qualitative/categorical unordered, qualitative/categorical ordered, date or time, text

Main data and visualisation types

- Amounts: bar plots, grouped and stacked bars, histograms, heatmaps
- Distributions (one or several): cumulative distribution plots, boxplots
- Proportions: pie charts, side-by-side bars, stacked bars
- X-Y relationships: scatterplots, principal components analysis (involving dimension reduction), time series
- Geospatial data: maps (with projections and layers), [choropleth maps](#)

Representing uncertainty: error bars, confidence bands

Multipanel figures: small multiples

(See the galleries mentioned in [our references](#).)

Breadth of text visualization types

Text Visualization Browser

A Visual Survey of Text Visualization Techniques ([IEEE PacificVis 2015 short paper](#))

Provided by ISOVIS group

[About](#) [Summary](#) [Add entry](#) [Other surveys](#) ▾

Techniques displayed:

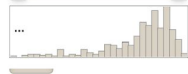
439

Search:

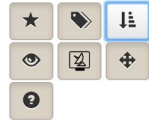
Time filter:

1976

2019



Visualization Tasks ☐



Data

Source



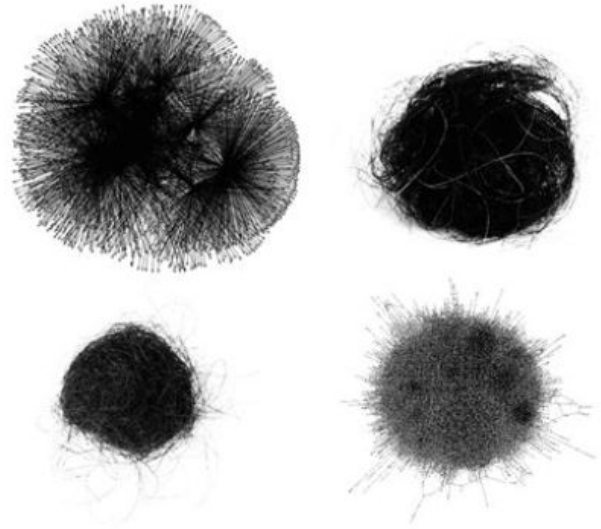
<http://textvis.lnu.se/>

Network data and hairballs

Graph visualization notoriously yields examples of illegible data representation.

To overcome this:

- use alternative representations (adjacency matrix)
- remove some data to show meaningful patterns

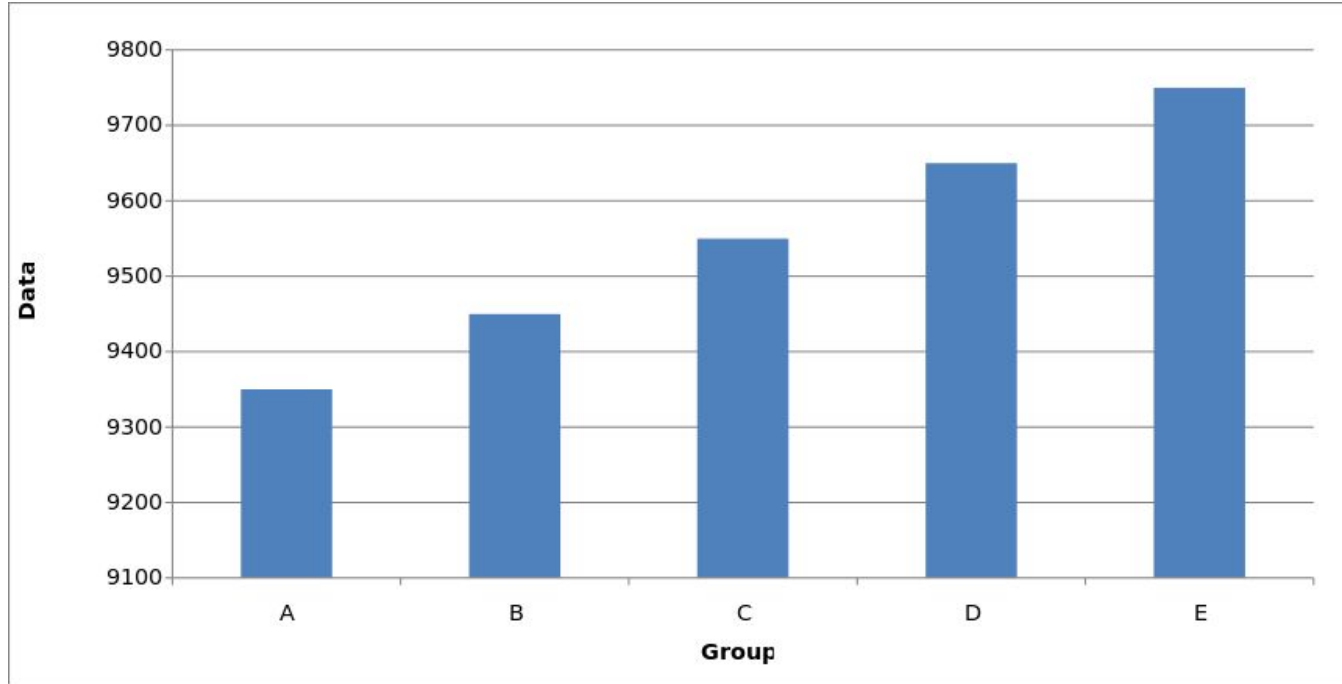


<https://dh2017.adho.org/abstracts/428/428.pdf>

Data visualization done wrong

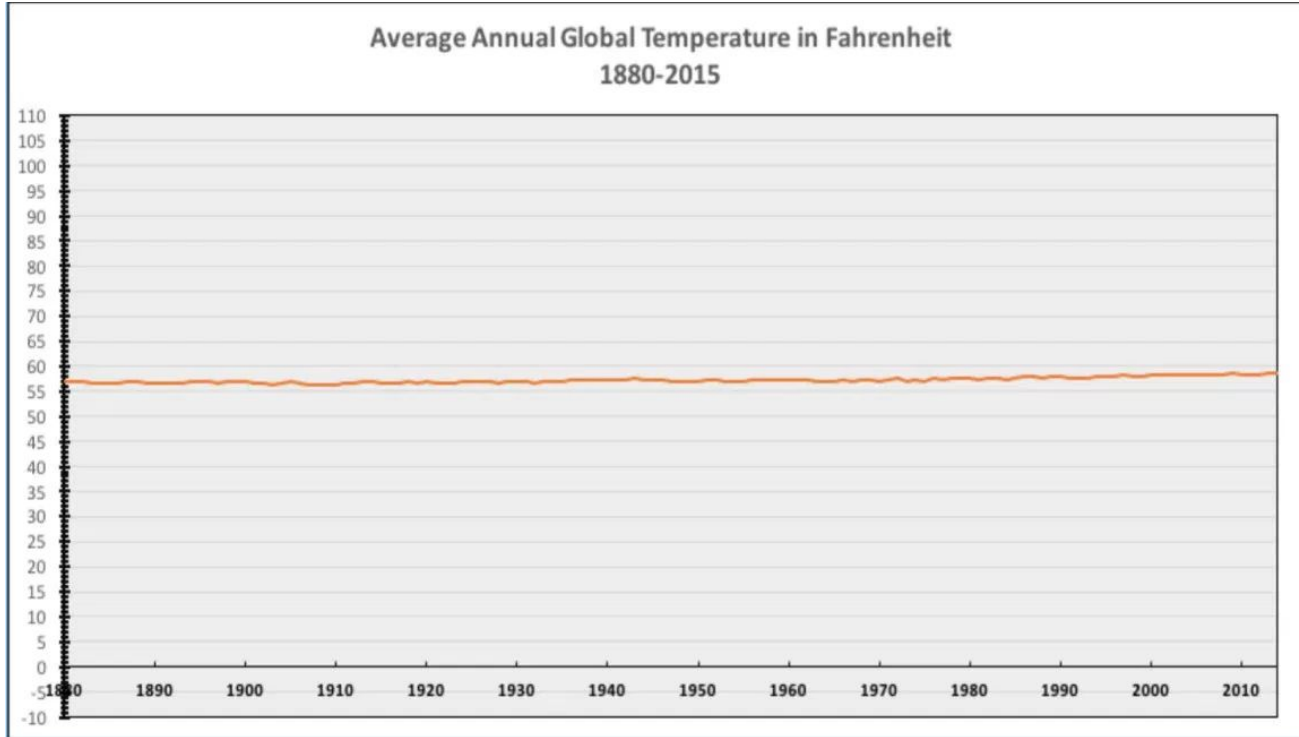
1. The y-axis
2. Binning
3. Unreadability
4. Missing proportionality
5. Improper scaling
6. Aberrations

See also <http://callingbullshit.org/tools.html>.



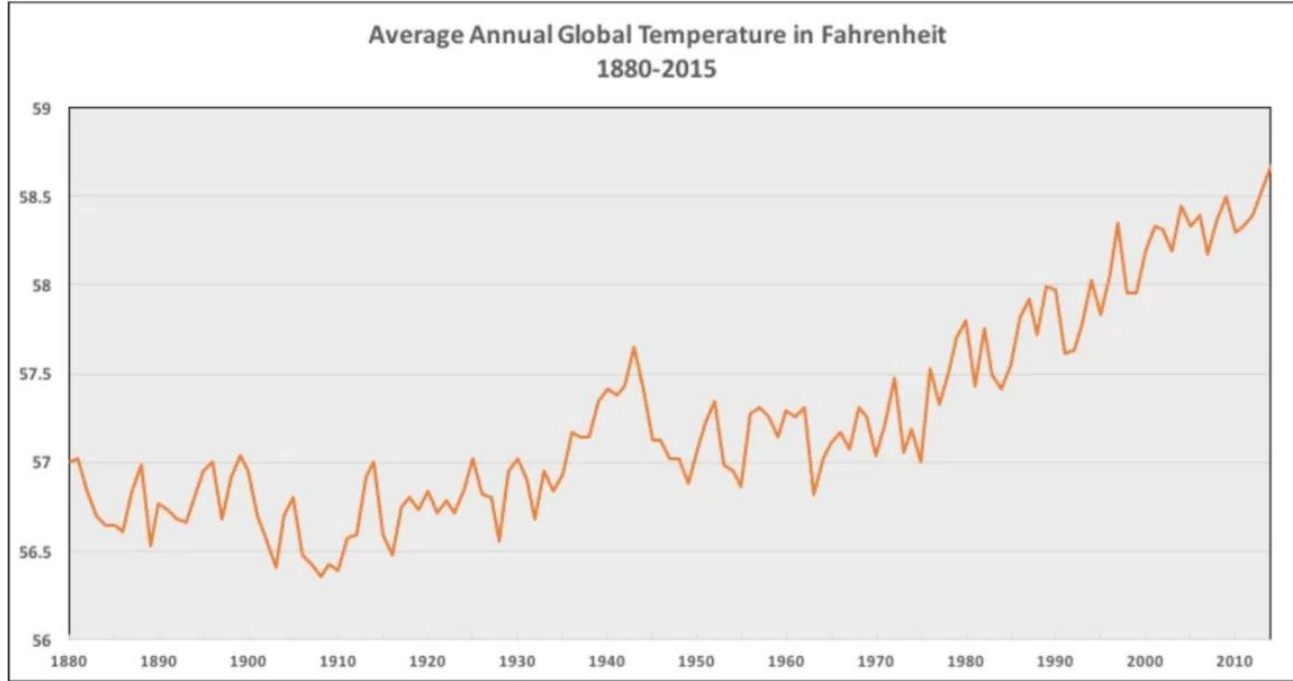
Misleading graph
https://en.wikipedia.org/wiki/Misleading_graph

The y-axis



The only global warming chart you need from now on
<http://www.powerlineblog.com/archives/2015/10/the-only-global-warming-chart-you-need-from-now-on.php>

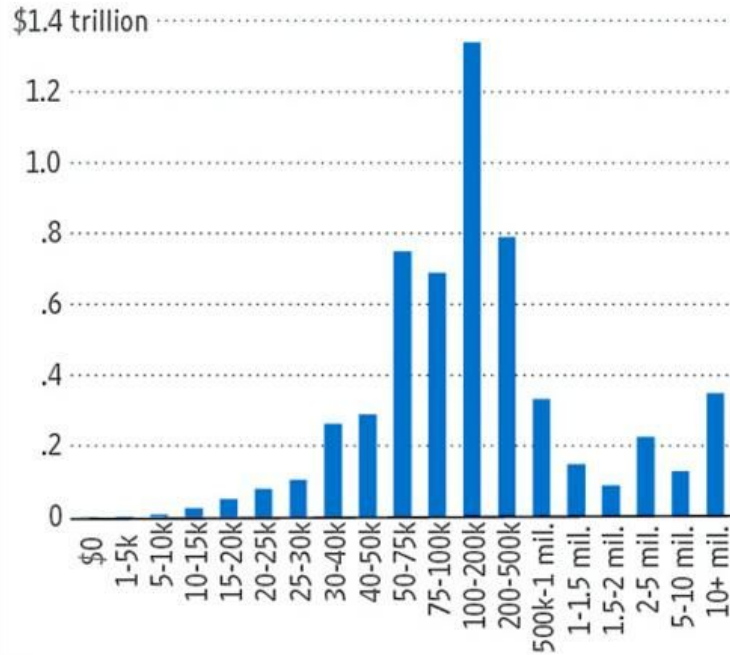
The y-axis



The only global warming chart you need from now on
<http://www.powerlineblog.com/archives/2015/10/the-only-global-warming-chart-you-need-from-now-on.php>

The Middle Class Tax Target

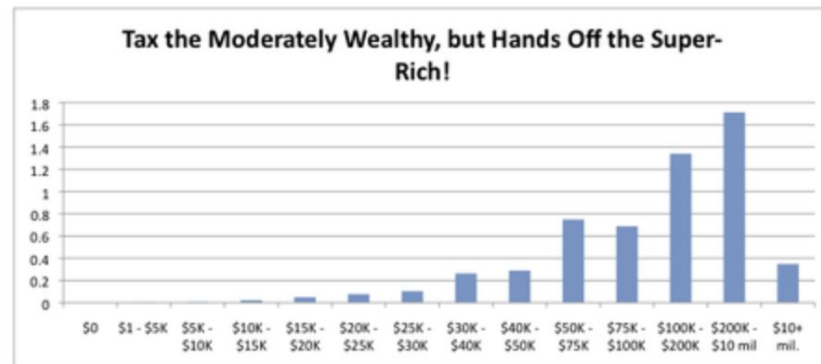
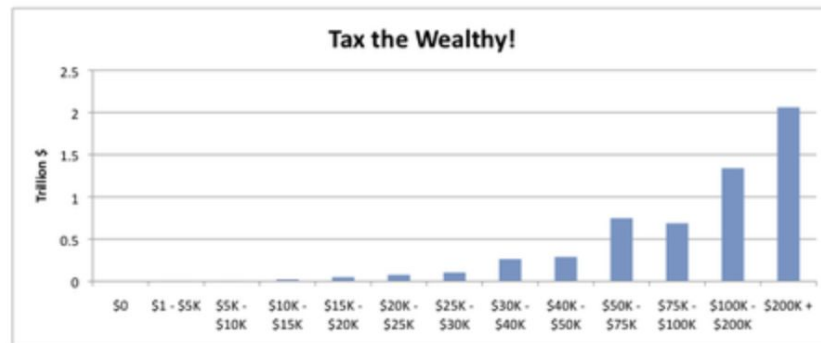
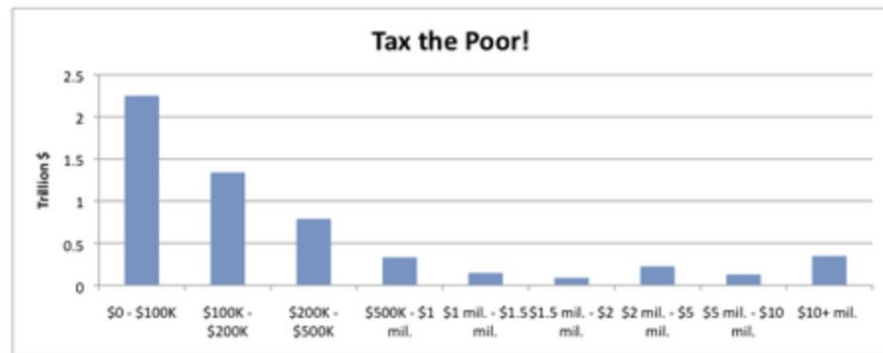
The amount of total taxable income (left scale) for all filers by adjusted gross income level for 2008



Source: IRS

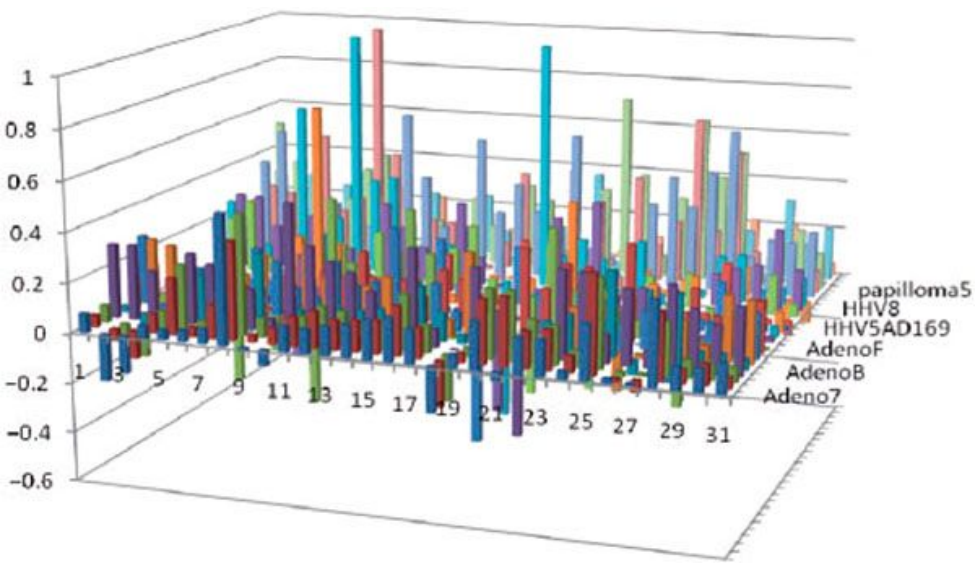
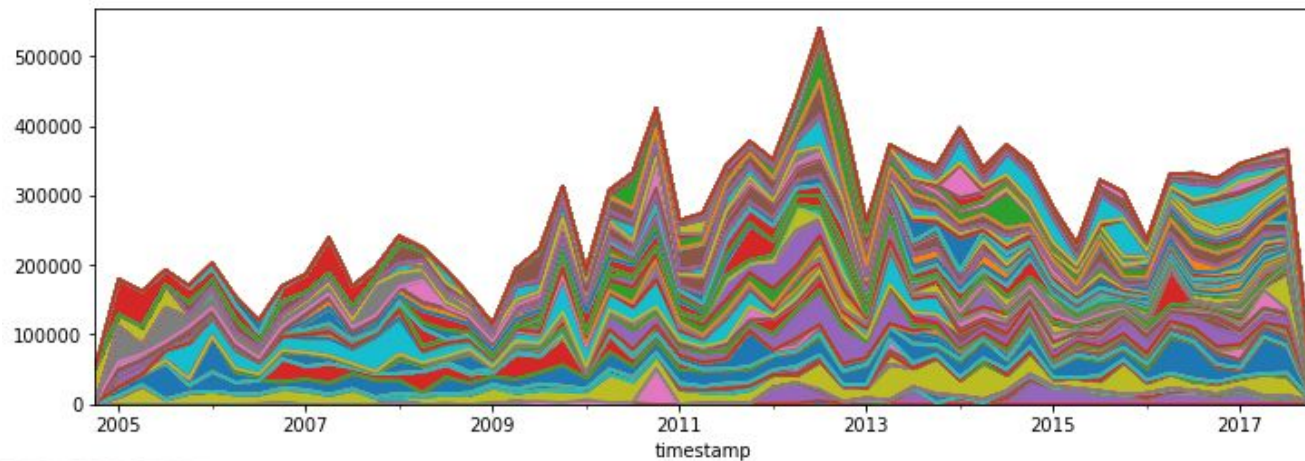
Wall Street Journal,
April 17th 2011

Binning

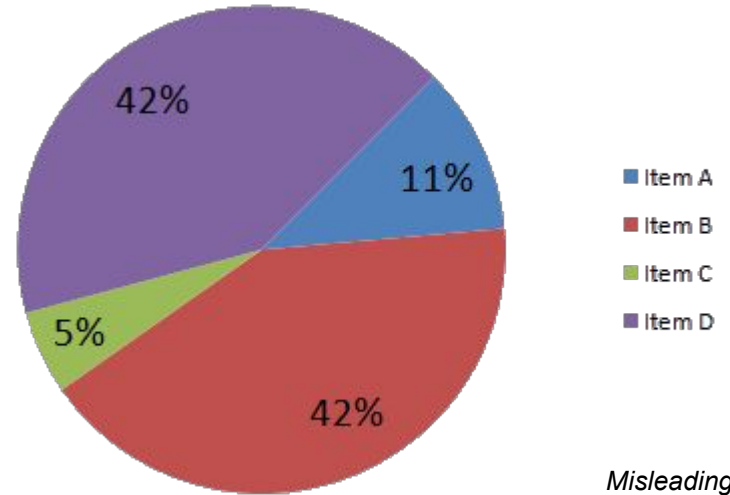
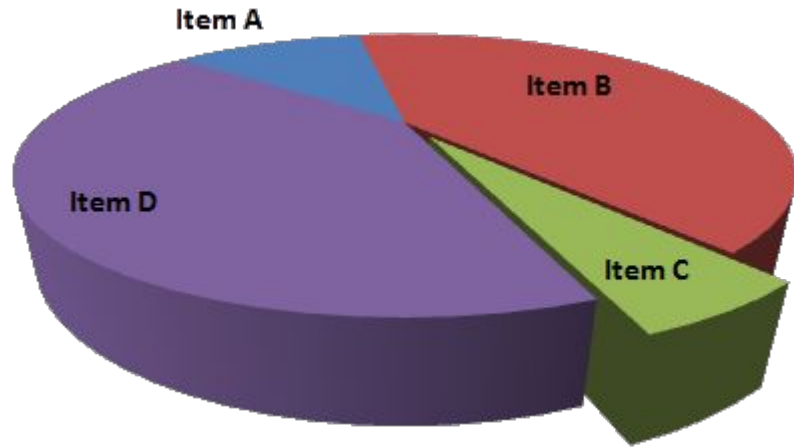


<http://www.brendan-nyhan.com/blog/2011/05/the-use-and-abuse-of-bar-graphs.html>

Unreadability

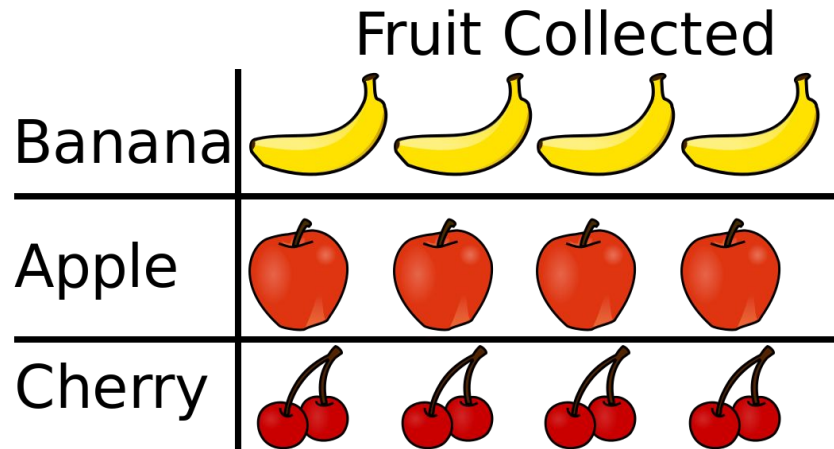
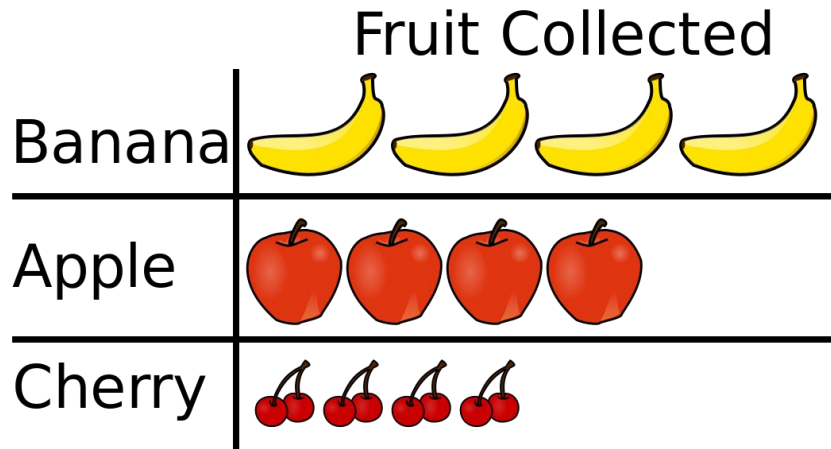


Missing proportionality



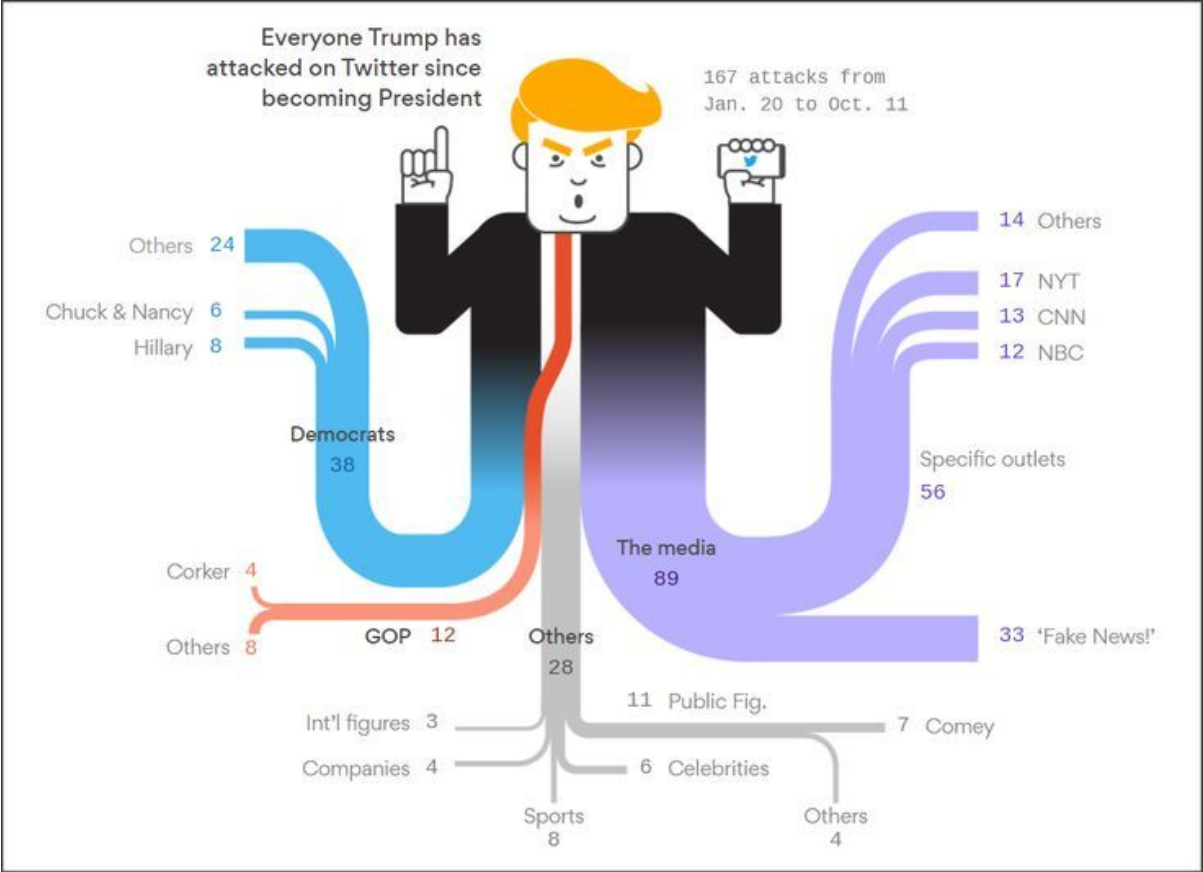
Misleading graph
https://en.wikipedia.org/wiki/Misleading_graph

Improper scaling



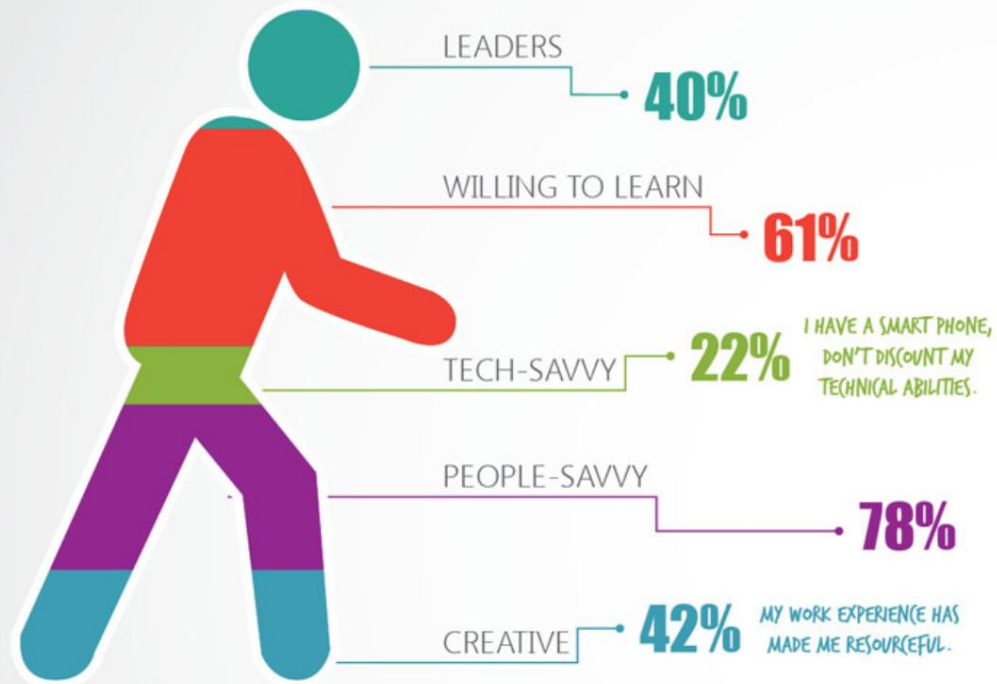
Misleading graph
https://en.wikipedia.org/wiki/Misleading_graph

Aberrations



Aberrations

HOW **BABY BOOMERS** DESCRIBE THEMSELVES



Tools, libraries, frameworks

GUI: [Voyant Tools](#), [Tableau public](#), [Palladio](#)

R: ggplot2 (static), shiny (interactive)

Python: matplotlib, seaborn (static), bokeh, plotly, dash (interactive)

JavaScript: D3.js, etc.

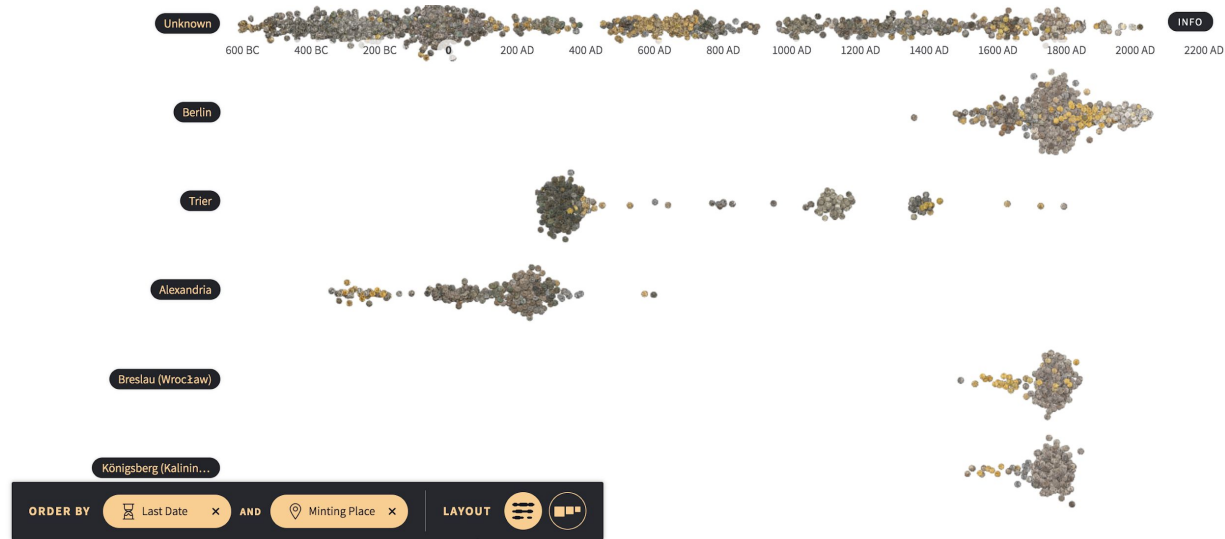
From exploration to
scientific communication

Static or dynamic, embedded in
documents or Web applications

DH Awards – best visualization winners

- (2019) [Witches – Mapping the Scottish Survey of Witchcraft Database](#)
- (2018) [Coins](#)
- (2017) [Mapping Islamophobia: Visualizing Islamophobia and Its Effects](#)
- (2016) [Peripleo – A search prototype by Pelagios Commons](#)
- (2015) [Mapping Metaphor with the Historical Thesaurus: Metaphor Map of English](#)
- (2014) [The Virtual Paul's Cross Project](#)
- (2013) [Infographic – the Humanities Matter](#)
- (2012) [A Thousand Words: Advanced Visualization for the Humanities](#)

Visualizing coin collections



Scalable reading interfaces

Cited Loci :: Aeneid

Home

Explore

About

Display: all

References: 291

Quotations: 314

I II III IV V VI VII VIII IX X XI XII

1-50

51-100

101-150

151-200

201-250

251-300

301-350

351-400

401-450

451-500

501-550

551-600

601-650

651-700

701-750

751-800

801-850

851-900

901-950

951-1000

In Focus: Book 1, lines 1-50

Arma virumque cano, Troiae qui primus ab oris

Italiam, fato profugus, Laviniaque venit

litora, multum ille et terris iactatus et alto

vi superum saevae memorem Iunonis ob iram;

5 multa quoque et bello passus, dum conderet urbem,

inferretque deos Latio, genus unde Latinum,

Albanique patres, atque altae moenia Romae.

Musa, mihi causas memora, quo numine laeso,

quidve dolens, regina deum tot volvere casus

10 insignem pietate virum, tot adire labores

impulerit. Tantaene animis caelestibus irae?

Urbs antiqua fuit, Tyrii tenuere coloni,

Karthago, Italiam contra Tiberinaque longe

ostia, dives opum studiisque asperrima belli;

15 quam Iuno fertur terris magis omnibus unam

posthabita coluisse Samo; hic illius arma,

hic currus fuit; hoc regnum dea gentibus esse,

si qua fata sinant, iam tum tenditque fovetque.

Results: quotations references

Aen. 1.1

112 52

Menin Virumque: Translating Homer with Virgil in "Epigrammata Bobiensia" 46_ 47 and 64

SCOTT MCGILL

The Classical Journal 2006

DOI: [10.2307/30038019](https://doi.org/10.2307/30038019)

By rendering Iliad 1.1 with Aeneid 1.1 rather than translating Homer with an eye to word-for-word correspondence, the EB poets offer a striking perspective on that aspect of ancient education.

— (reference: 571a5798b634c1c68fef6343)

Silius Punica 6. 1-53

EDWARD L. BASSETT

Classical Philology 1959

DOI: [10.2307/266203](https://doi.org/10.2307/266203)

(The Silian phrase will help to endorse foeda over

<http://aeneid.citedloci.org>