



practical
tips

OPEN ACCESS

EDITORIAL

Ten Simple Rules for Better Figures

Nicolas P. Rougier , Michael Droettboom, Philip E. Bourne

Published: September 11, 2008



OPEN ACCESS

PERSPECTIVE

Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm

Tracey L. Weissgerber , Natasha M. Milic, Stacey J. Winham, Vesna D. Garovic

Published: April 22, 2015 • <https://doi.org/10.1371/journal.pbio.1002128>



Practical Rules for Using Color in Charts

Stephen Few, Perceptual Edge
Visual Business Intelligence Newsletter

February 2008

BROWSE PUBL

Data Visualization: Chart Dos and Don'ts

This LibGuide collects resources and tutorials related to data visualization. It is a companion to the visualization services provided by Data and Visualization Services at Duke University Libraries.

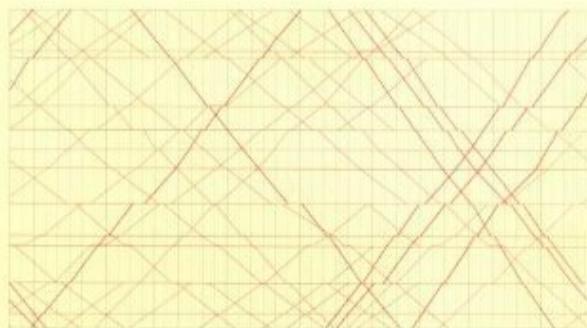
Tufte's Rules.

[THE RULES](#) | [ABOUT TUFTE](#)

About Tufte:

1. [Biography](#)
2. Contributions
3. Books

Copyrighted Material



SECOND EDITION

The Visual Display of Quantitative Information

EDWARD R. TUFTE

Copyrighted Material

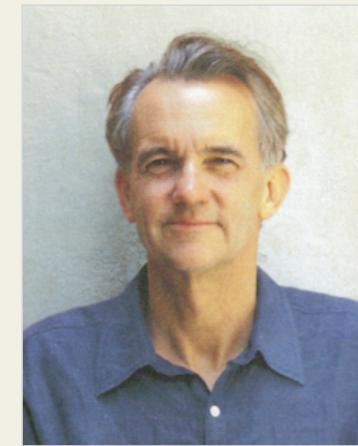
Biography:

EDWARD TUFTE (Tuff-tee) was born on 1942 in Kansas City, Missouri. He has been called "*The da Vinci of data*," "*Chart Master*," and "*The Data Artist*."

Tufte grew up in California, with his father and mother. His father was an engineer and his mother, a teacher. Together, they taught him the importance of a purposeful life and the value of contribution.

Upon graduating Beverly Hills High School in 1960, Tufte went on to Stanford University to earn his BA and MA in Statistics and then Yale University to earn his PhD in Political Science.

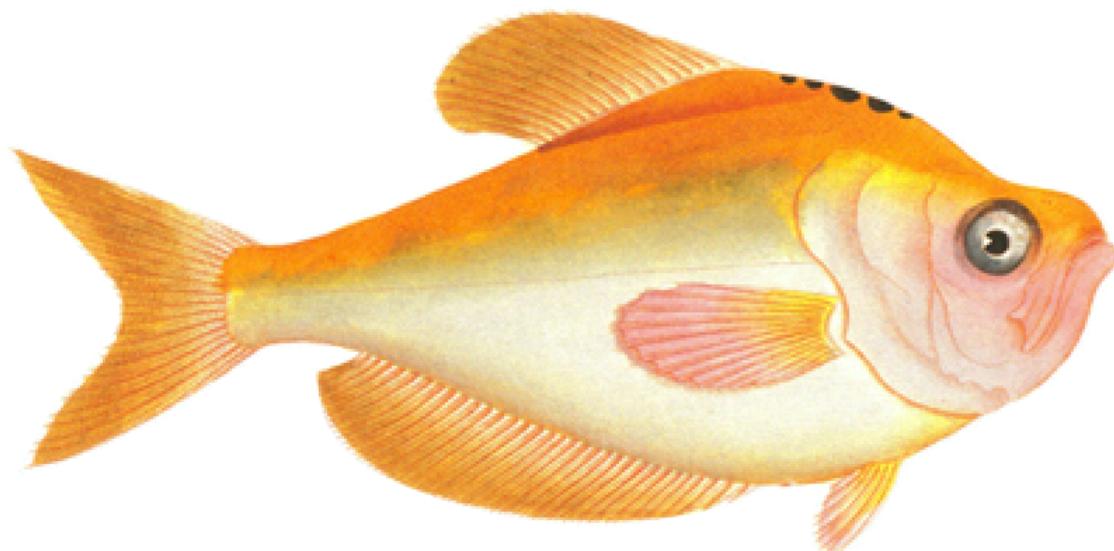
Tufte's education was the foundation for all of his later achievements and contributions, which includes becoming a highly respected professor.



Mr. Edward Rolf Tufte

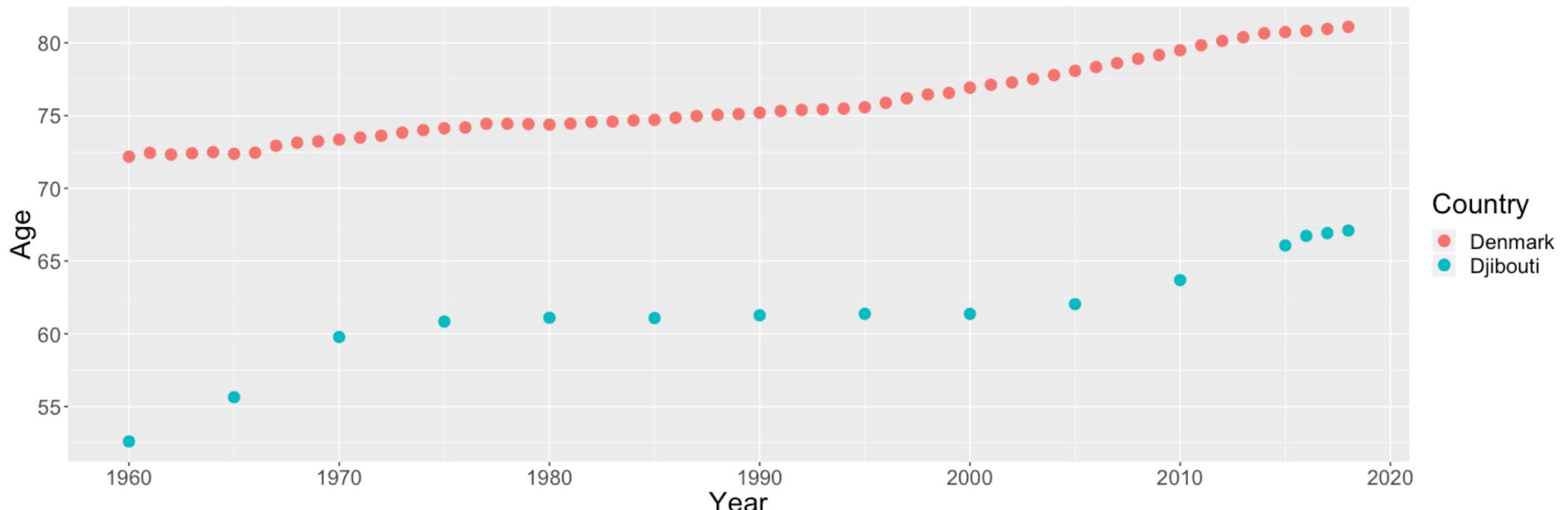
"If you're not doing something different, you're not doing anything at all."

i - show your data



i - show your data

- avoid summaries & aggregation if not required
- show where data is missing but do not distract the user
- rely on the deductive, inductive and adductive reasoning of the user

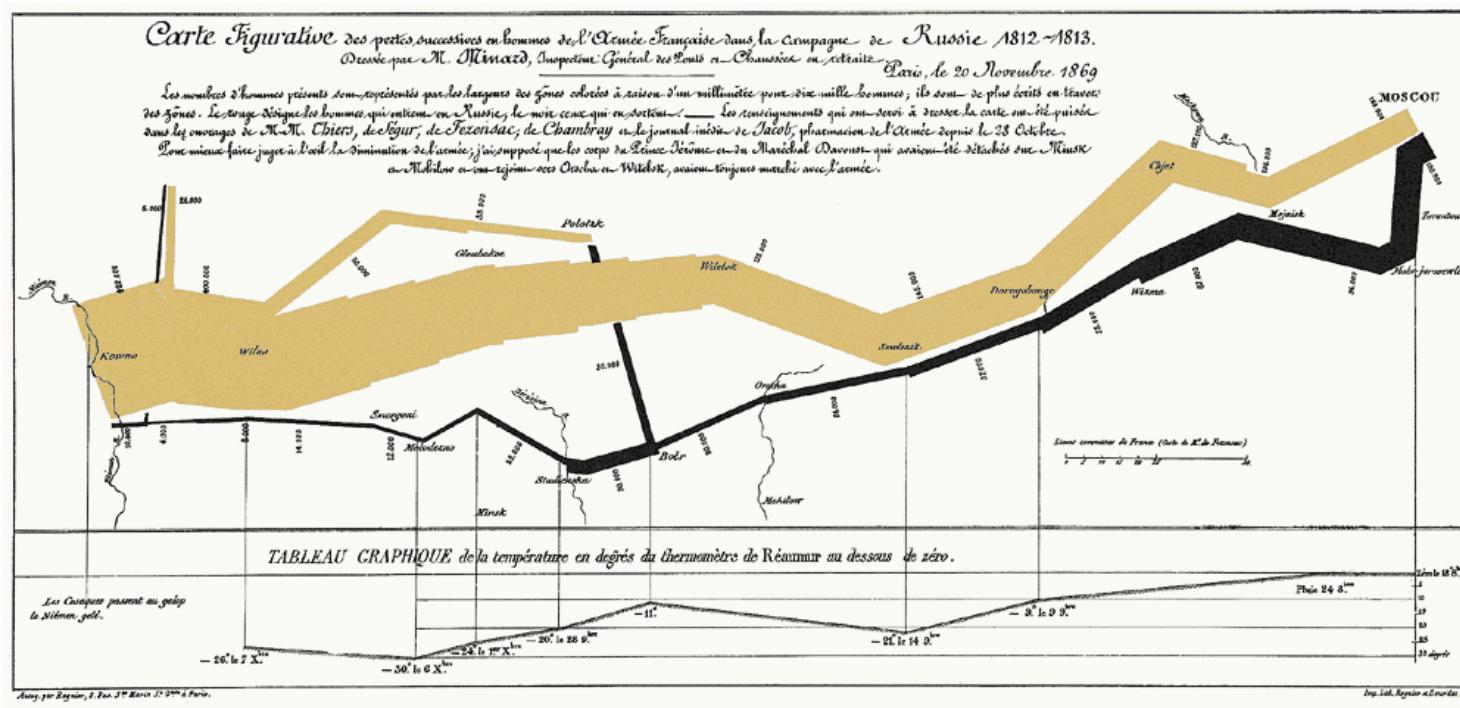


ii - use graphics

a picture is worth a thousand words

show with a diagram or a picture things that it would take a lot of text to describe

“only a picture can carry such volume of data in such a small space” [tufte]

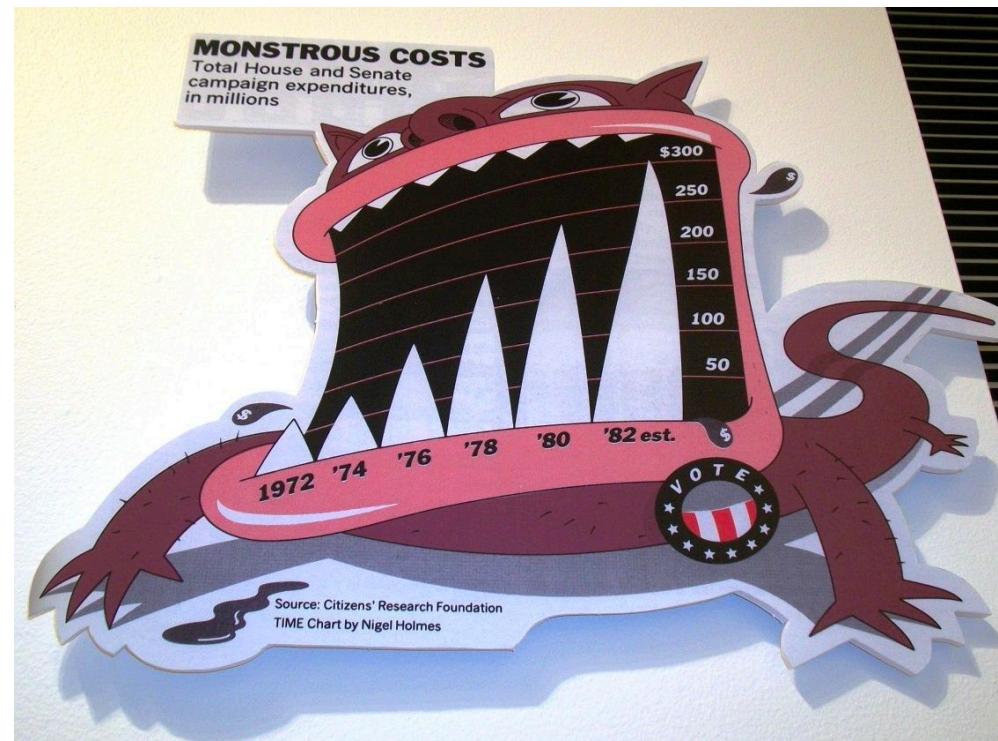


iii - avoid chartjunks

too many decorative elements added, detracting the user from the actual message

typical example: the unneeded 3d effects

“chartjunk can turn bores into disasters, but it can never rescue a thin data set” [tufte]

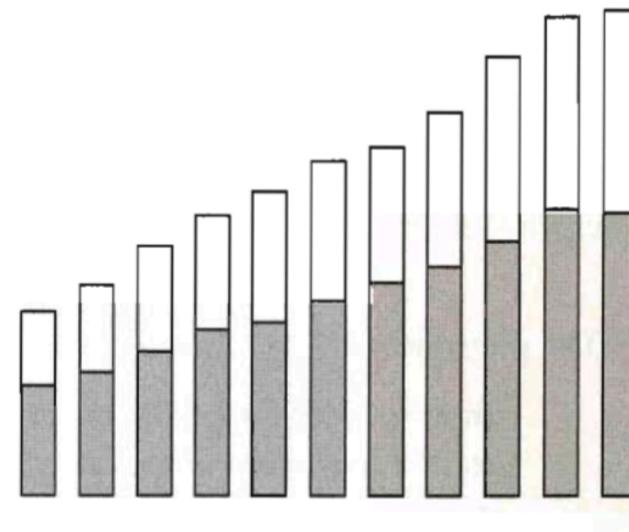
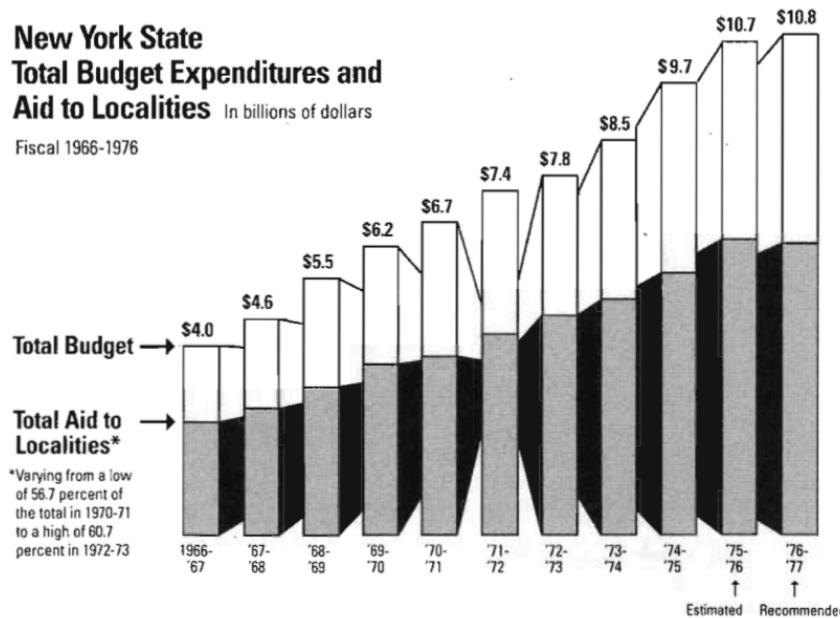


iii - avoid chartjunks

too many decorative elements added, detracting the user from the actual message

typical example: the unneeded 3d effects

“chartjunk can turn bores into disasters, but it can never rescue a thin data set” [tufte]



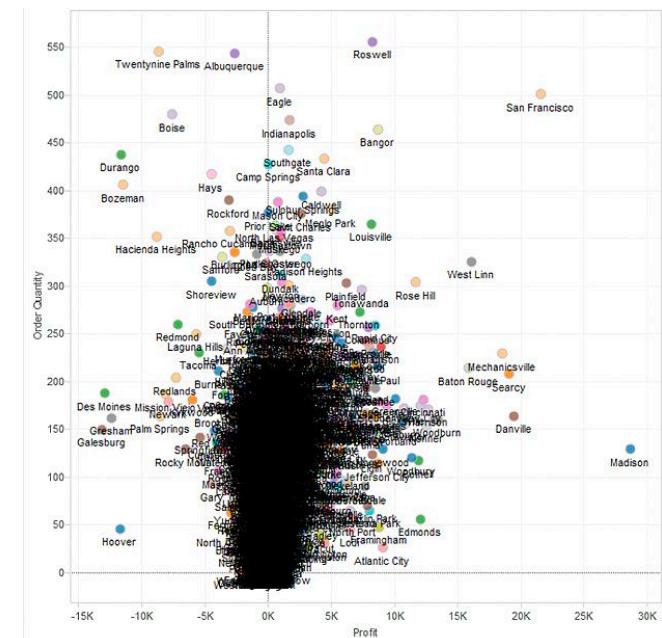
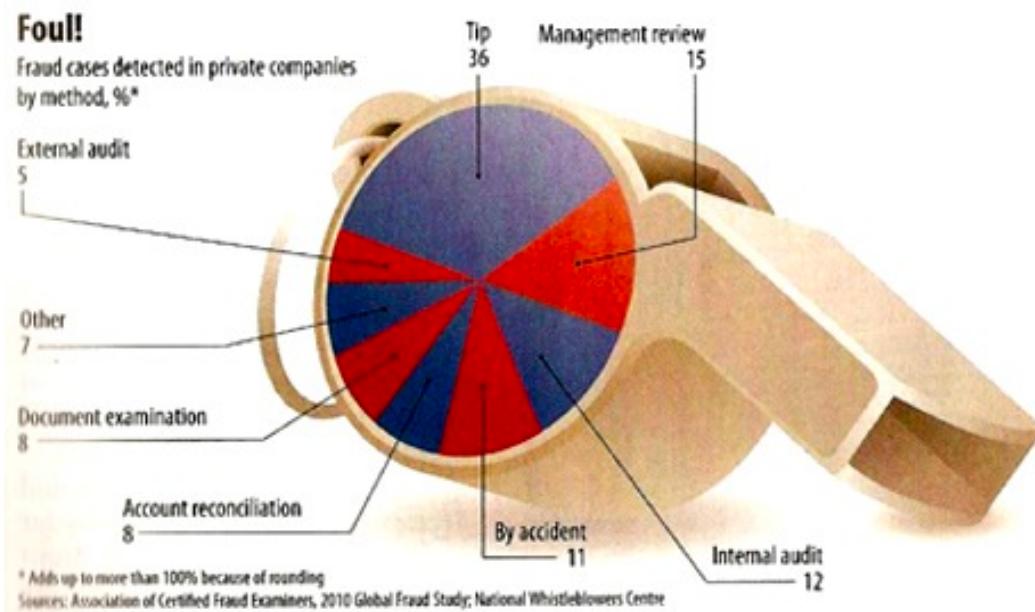
[new york times, 1976]

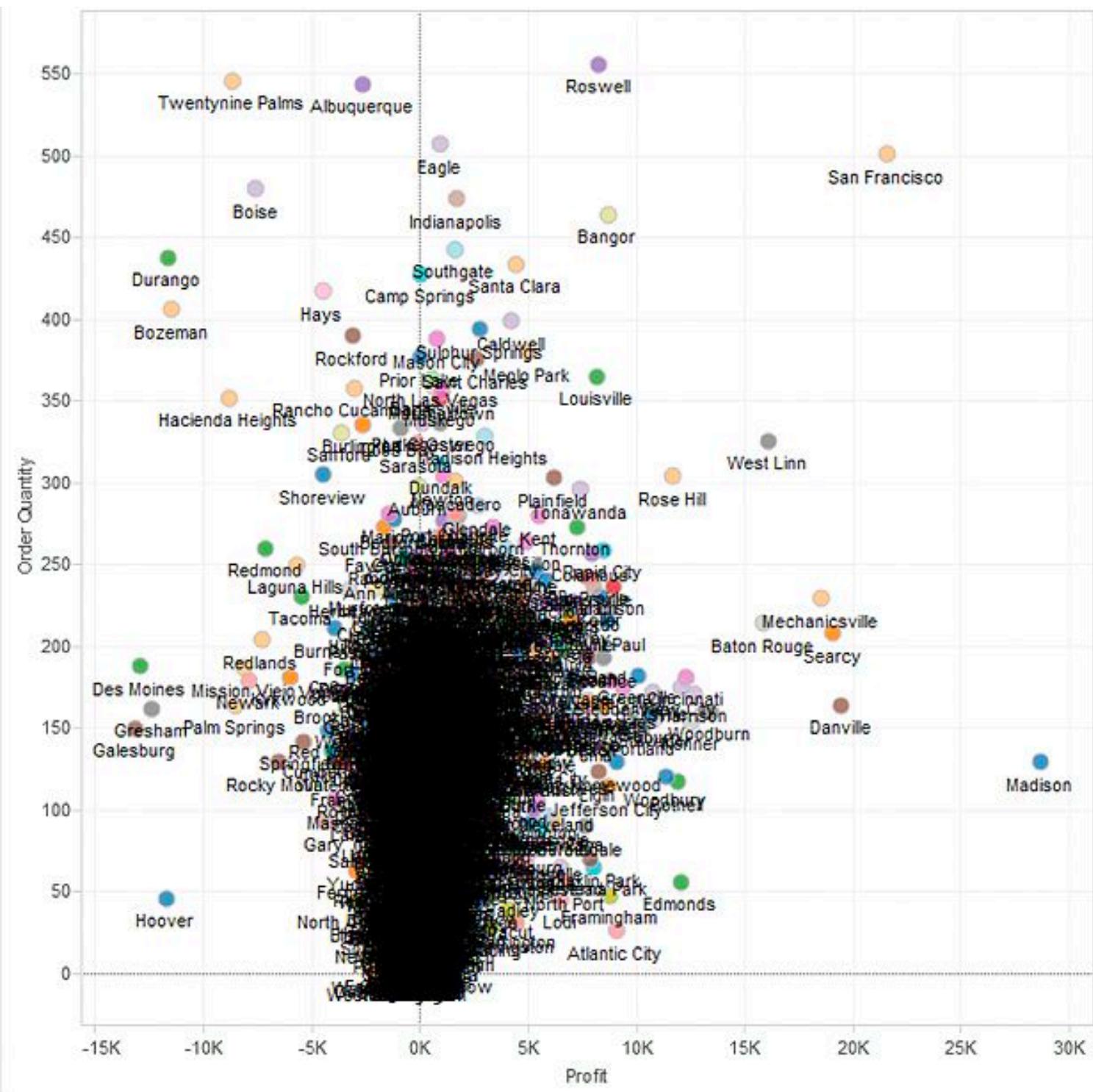
iii - avoid chartjunks

too many decorative elements added, detracting the user from the actual message

typical example: the unneeded 3d effects

“chartjunk can turn bores into disasters, but it can never rescue a thin data set” [tufte]



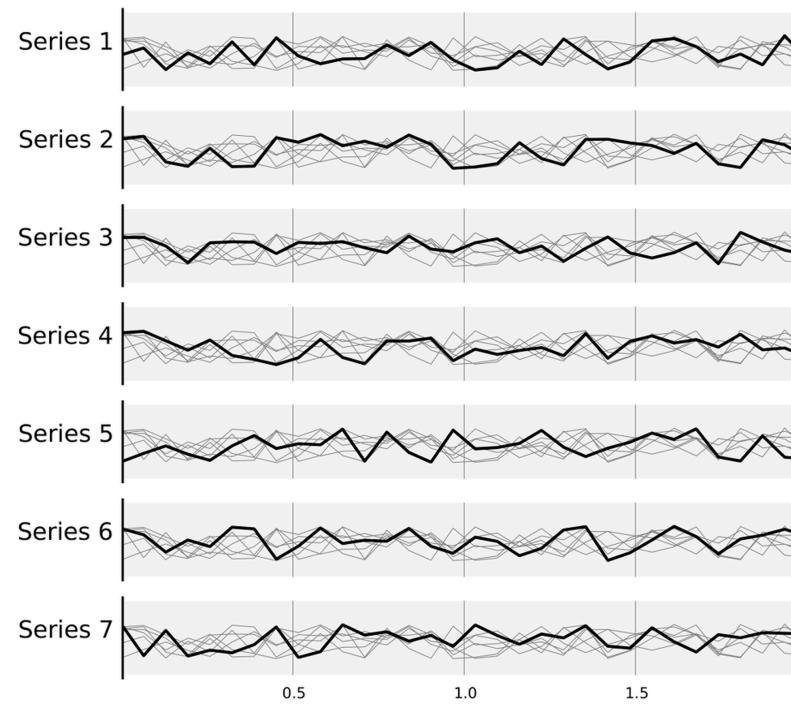
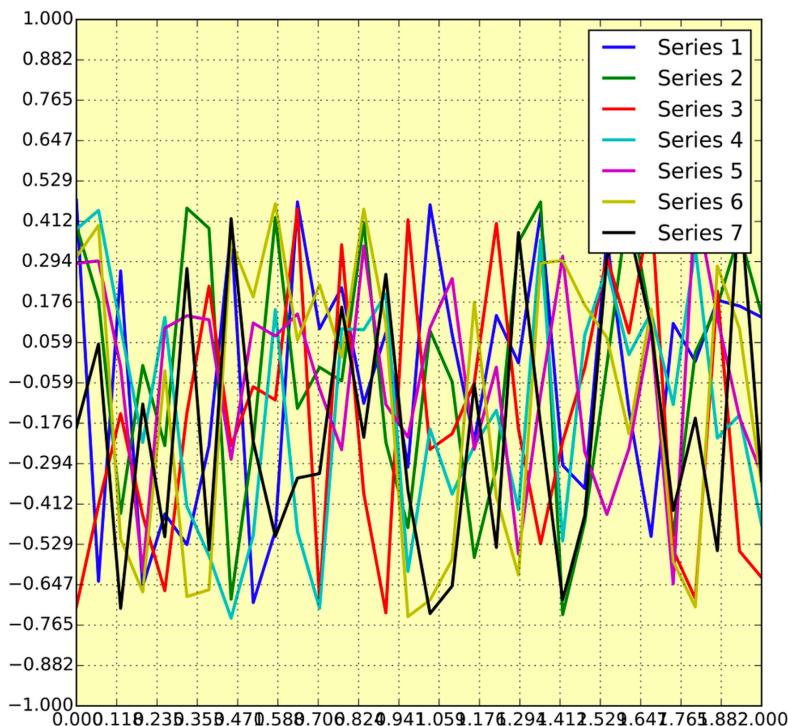


iii - avoid chartjunks

too many decorative elements added, detracting the user from the actual message

typical example: the unneeded 3d effects

“chartjunk can turn bores into disasters, but it can never rescue a thin data set” [tufte]



iv - the data-ink ratio

aka “remove2improve”

Data-ink ratio =
$$\frac{\text{data-ink}}{\text{total ink used to print the graphic}}$$

= proportion of a graphic's ink devoted to the non-redundant display of data-information

= $1.0 - \text{proportion of a graphic that can be erased without loss of data-information.}$

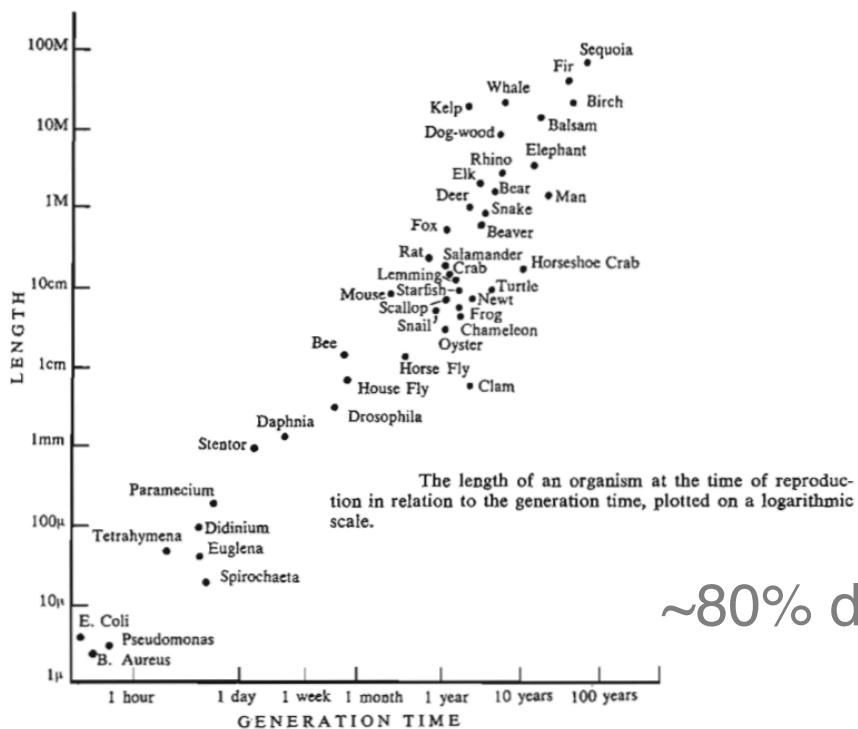
iv - the data-ink ratio

aka “remove2improve”

Data-ink ratio = $\frac{\text{data-ink}}{\text{total ink used to print the graphic}}$

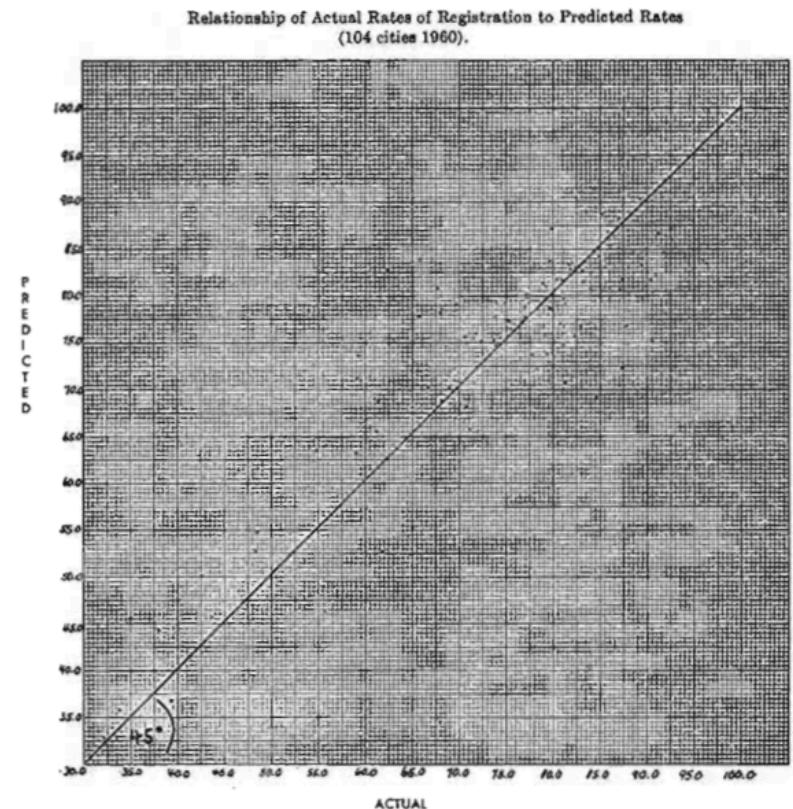
- = proportion of a graphic's ink devoted to the non-redundant display of data-information
- = $1.0 - \text{proportion of a graphic that can be erased without loss of data-information.}$

~90% non-data-ink



~80% data-ink

[j t bonner, 1965]



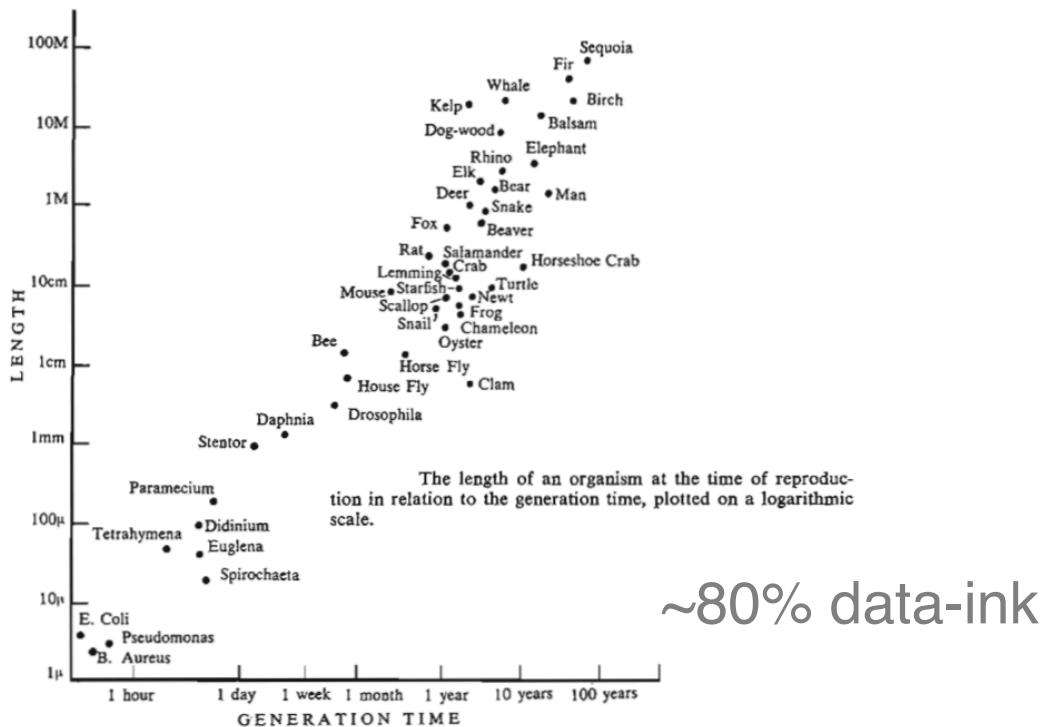
iv - the data-ink ratio

aka “remove2improve”

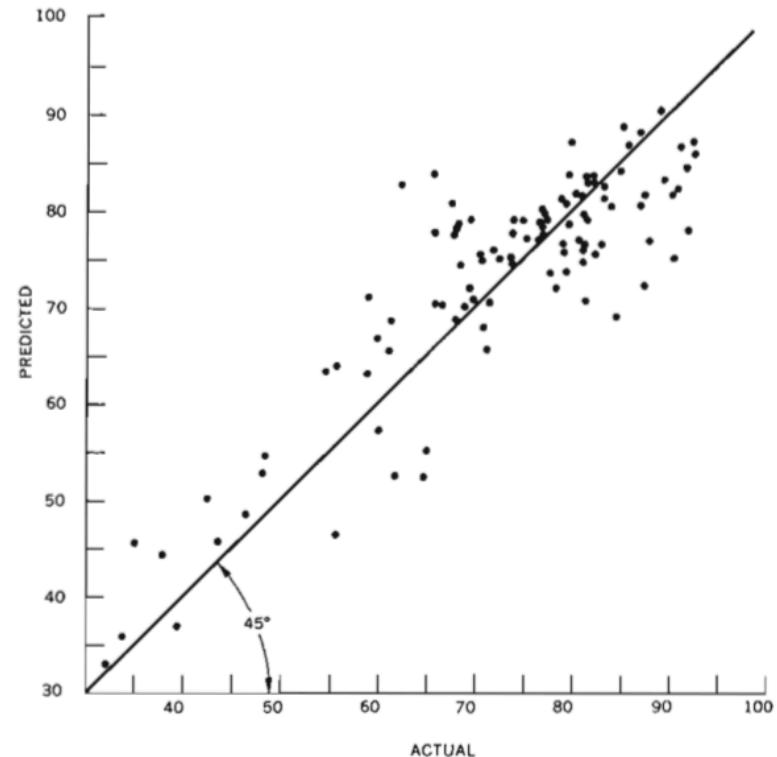
Data-ink ratio = $\frac{\text{data-ink}}{\text{total ink used to print the graphic}}$

- = proportion of a graphic's ink devoted to the non-redundant display of data-information
- = $1.0 - \text{proportion of a graphic that can be erased without loss of data-information.}$

~30% non-data-ink



[j t bonner, 1965]



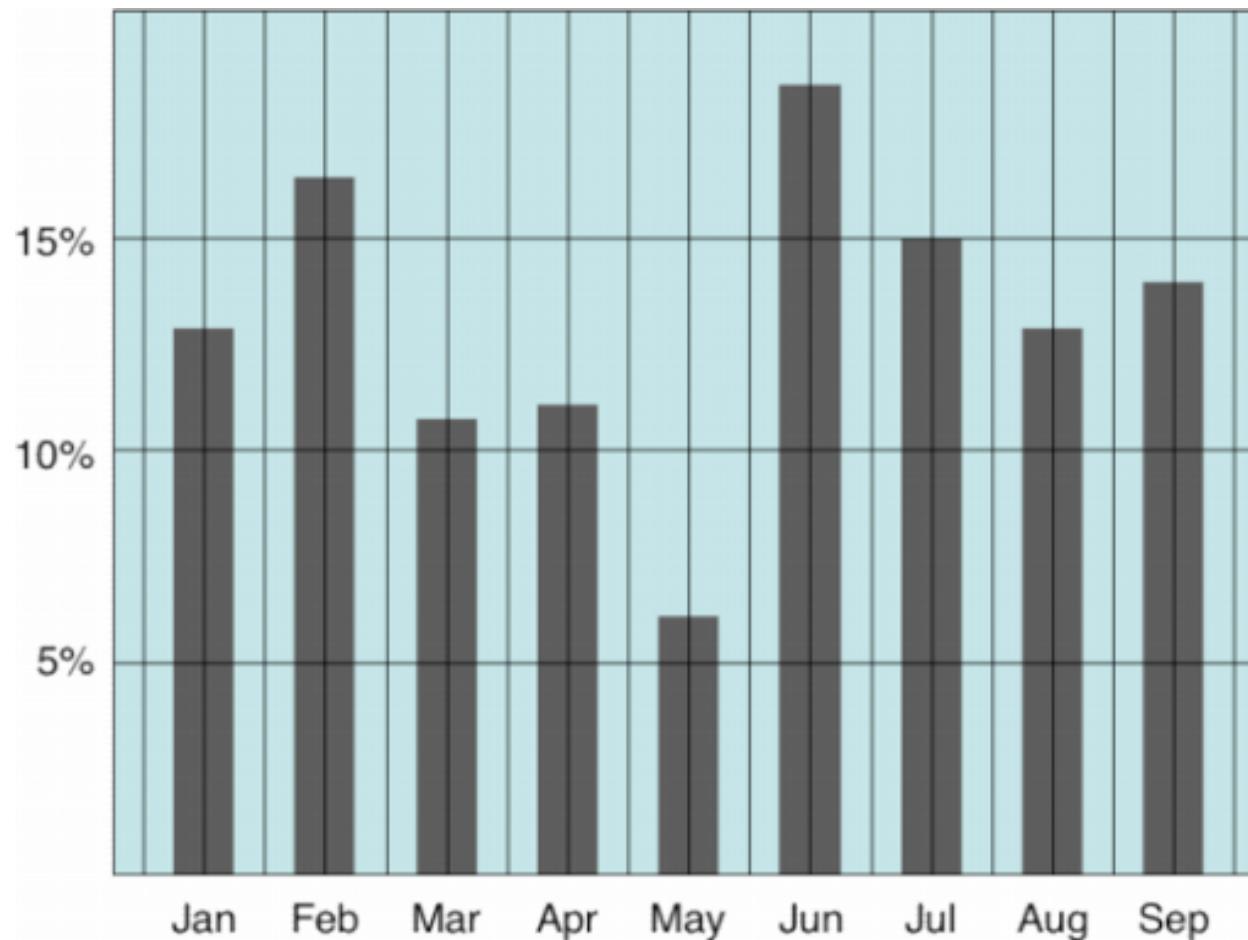
iv - the data-ink ratio

aka “remove2improve”

Data-ink ratio = $\frac{\text{data-ink}}{\text{total ink used to print the graphic}}$

= proportion of a graphic's ink devoted to the non-redundant display of data-information

= $1.0 - \text{proportion of a graphic that can be erased without loss of data-information}$.



Above all else show the data.
Maximize the data-ink ratio.
Erase non-data-ink.
Erase redundant data-ink.
Revise and edit.

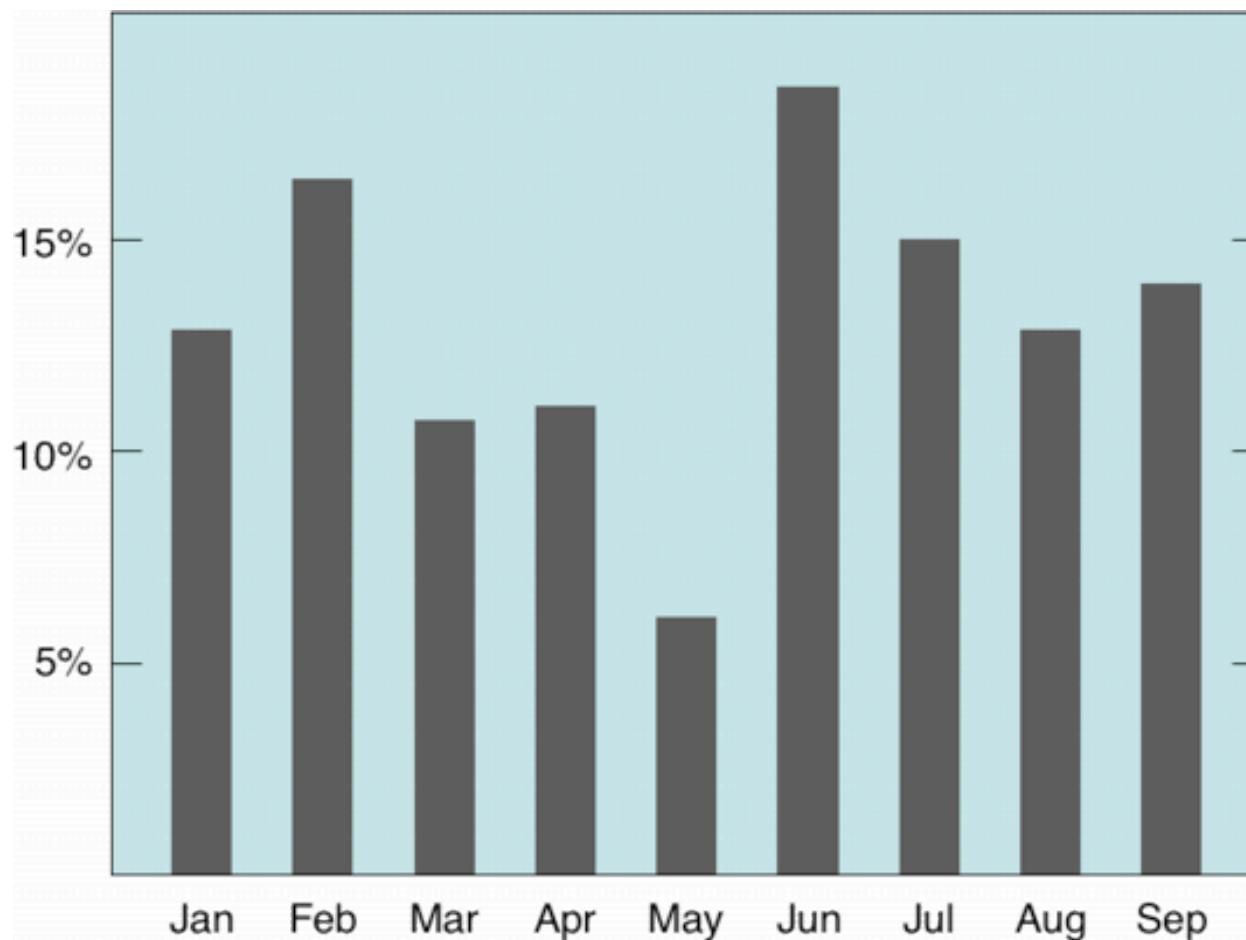
iv - the data-ink ratio

aka “remove2improve”

Data-ink ratio = $\frac{\text{data-ink}}{\text{total ink used to print the graphic}}$

= proportion of a graphic's ink devoted to the non-redundant display of data-information

= $1.0 - \text{proportion of a graphic that can be erased without loss of data-information}$.



Above all else show the data.
Maximize the data-ink ratio.
Erase non-data-ink.
Erase redundant data-ink.
Revise and edit.

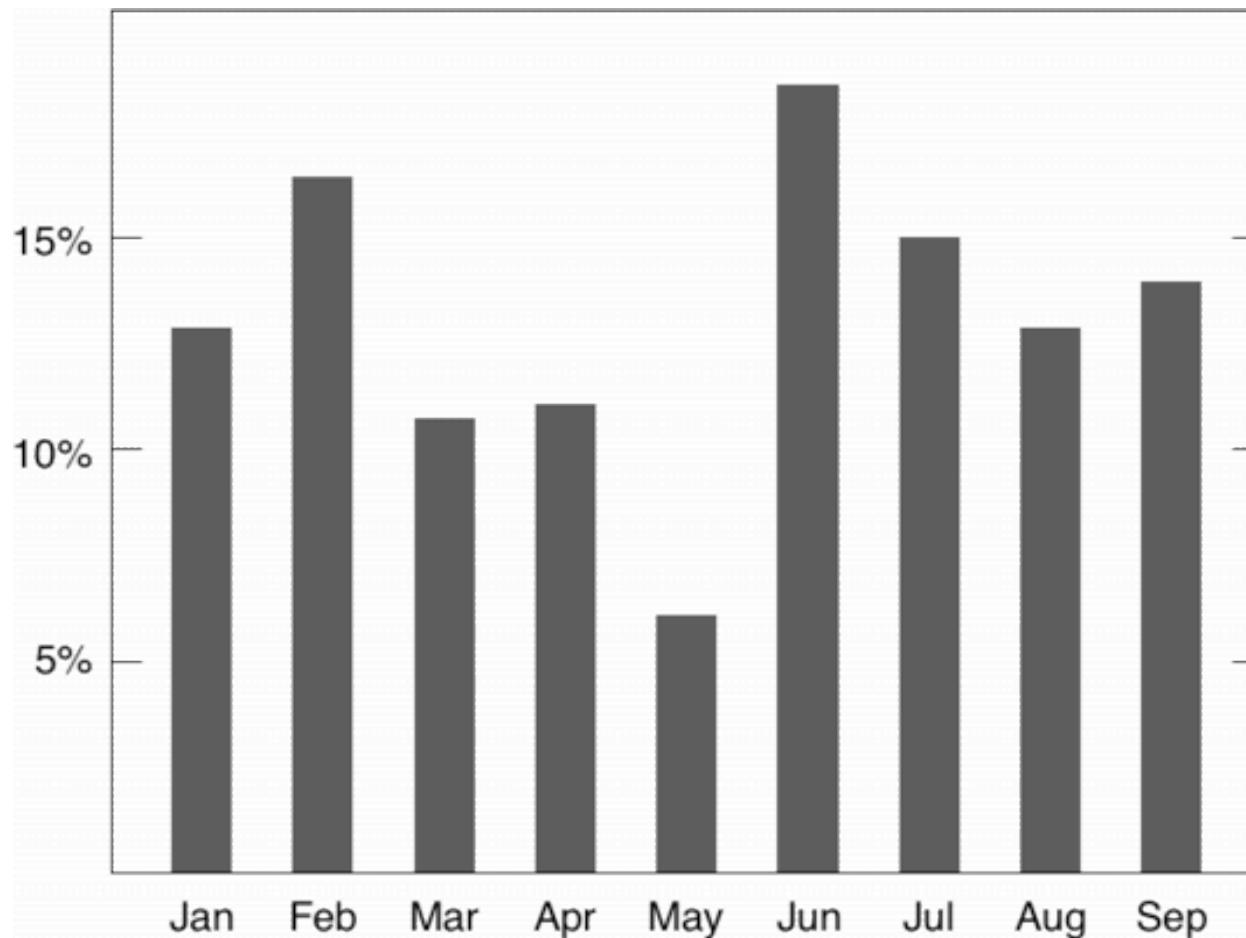
iv - the data-ink ratio

aka “remove2improve”

Data-ink ratio = $\frac{\text{data-ink}}{\text{total ink used to print the graphic}}$

= proportion of a graphic's ink devoted to the non-redundant display of data-information

= $1.0 - \text{proportion of a graphic that can be erased without loss of data-information}$.



Above all else show the data.
Maximize the data-ink ratio.
Erase non-data-ink.
Erase redundant data-ink.
Revise and edit.

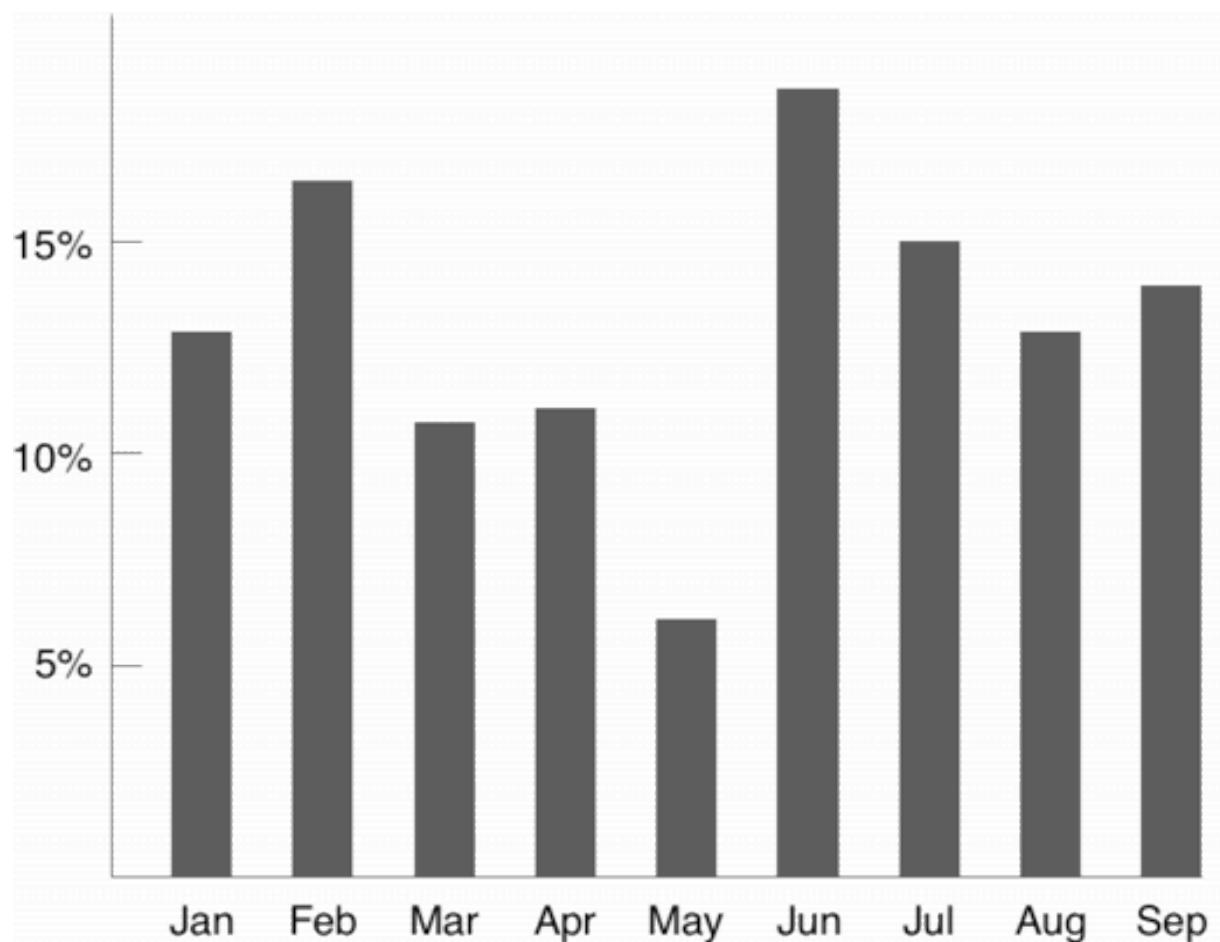
iv - the data-ink ratio

aka “remove2improve”

Data-ink ratio = $\frac{\text{data-ink}}{\text{total ink used to print the graphic}}$

= proportion of a graphic's ink devoted to the non-redundant display of data-information

= $1.0 - \text{proportion of a graphic that can be erased without loss of data-information}$.



Above all else show the data.
Maximize the data-ink ratio.
Erase non-data-ink.
Erase redundant data-ink.
Revise and edit.

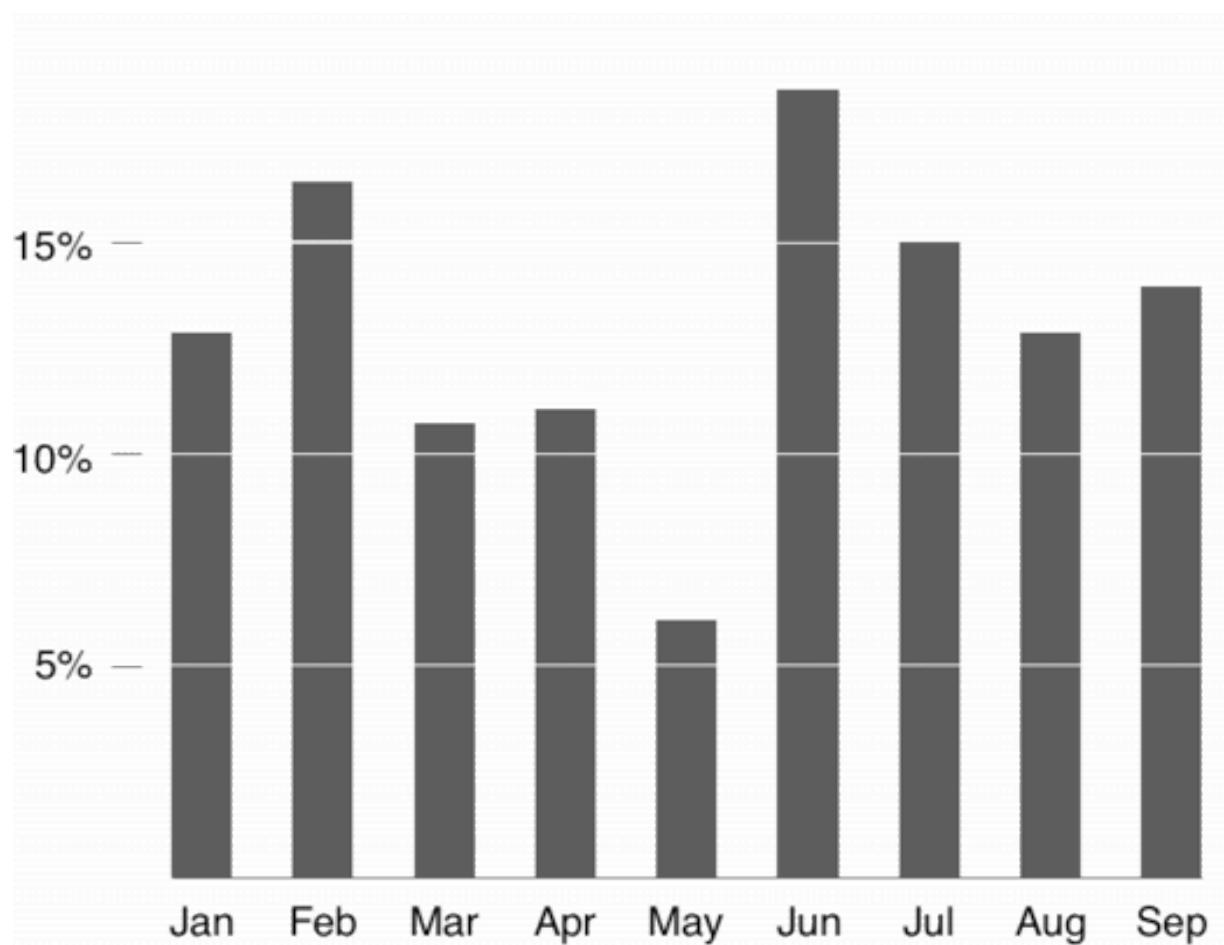
iv - the data-ink ratio

aka “remove2improve”

Data-ink ratio = $\frac{\text{data-ink}}{\text{total ink used to print the graphic}}$

= proportion of a graphic's ink devoted to the non-redundant display of data-information

= $1.0 - \text{proportion of a graphic that can be erased without loss of data-information}$.



Above all else show the data.
Maximize the data-ink ratio.
Erase non-data-ink.
Erase redundant data-ink.
Revise and edit.

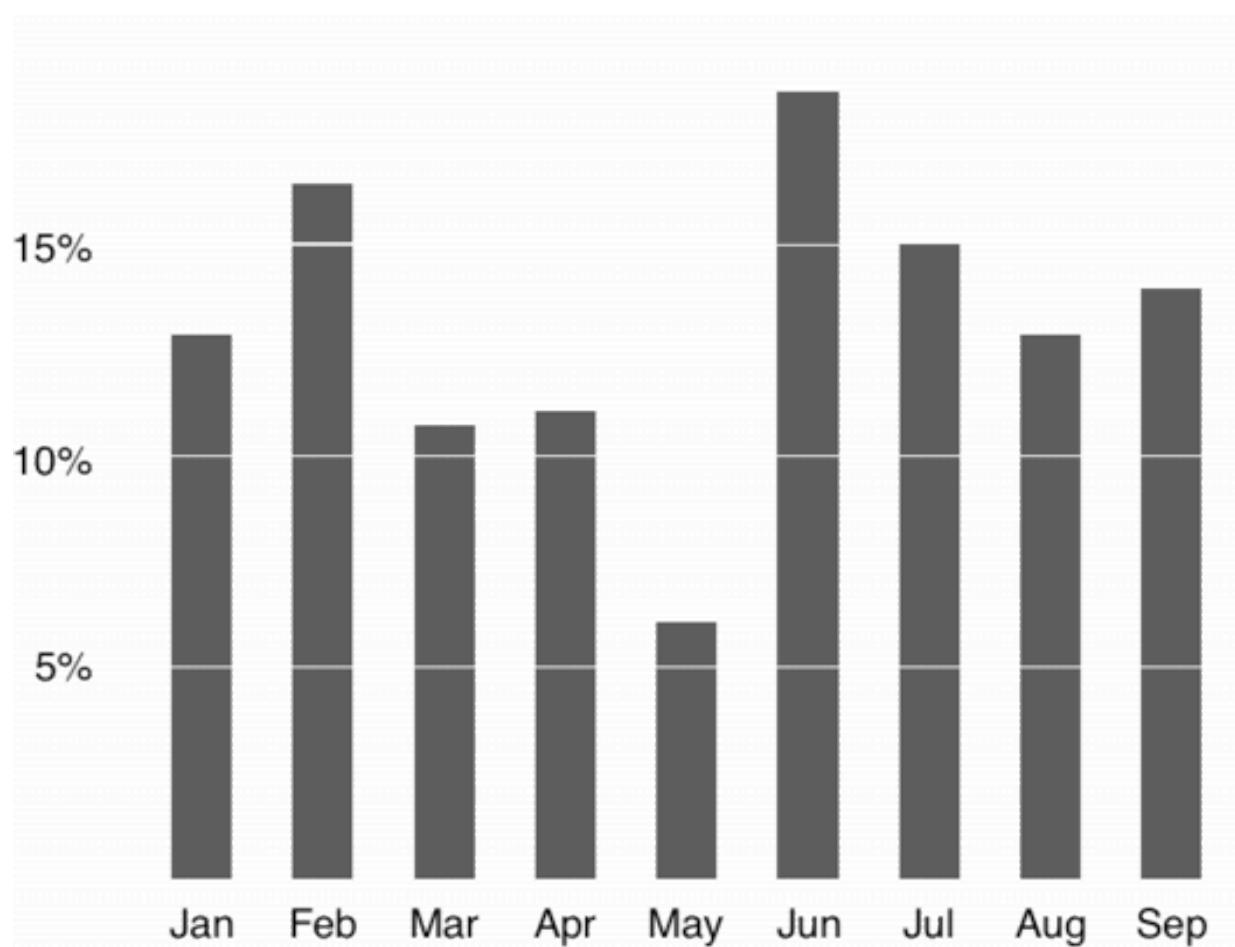
iv - the data-ink ratio

aka “remove2improve”

Data-ink ratio = $\frac{\text{data-ink}}{\text{total ink used to print the graphic}}$

= proportion of a graphic's ink devoted to the non-redundant display of data-information

= $1.0 - \text{proportion of a graphic that can be erased without loss of data-information}$.



Above all else show the data.
Maximize the data-ink ratio.
Erase non-data-ink.
Erase redundant data-ink.
Revise and edit.

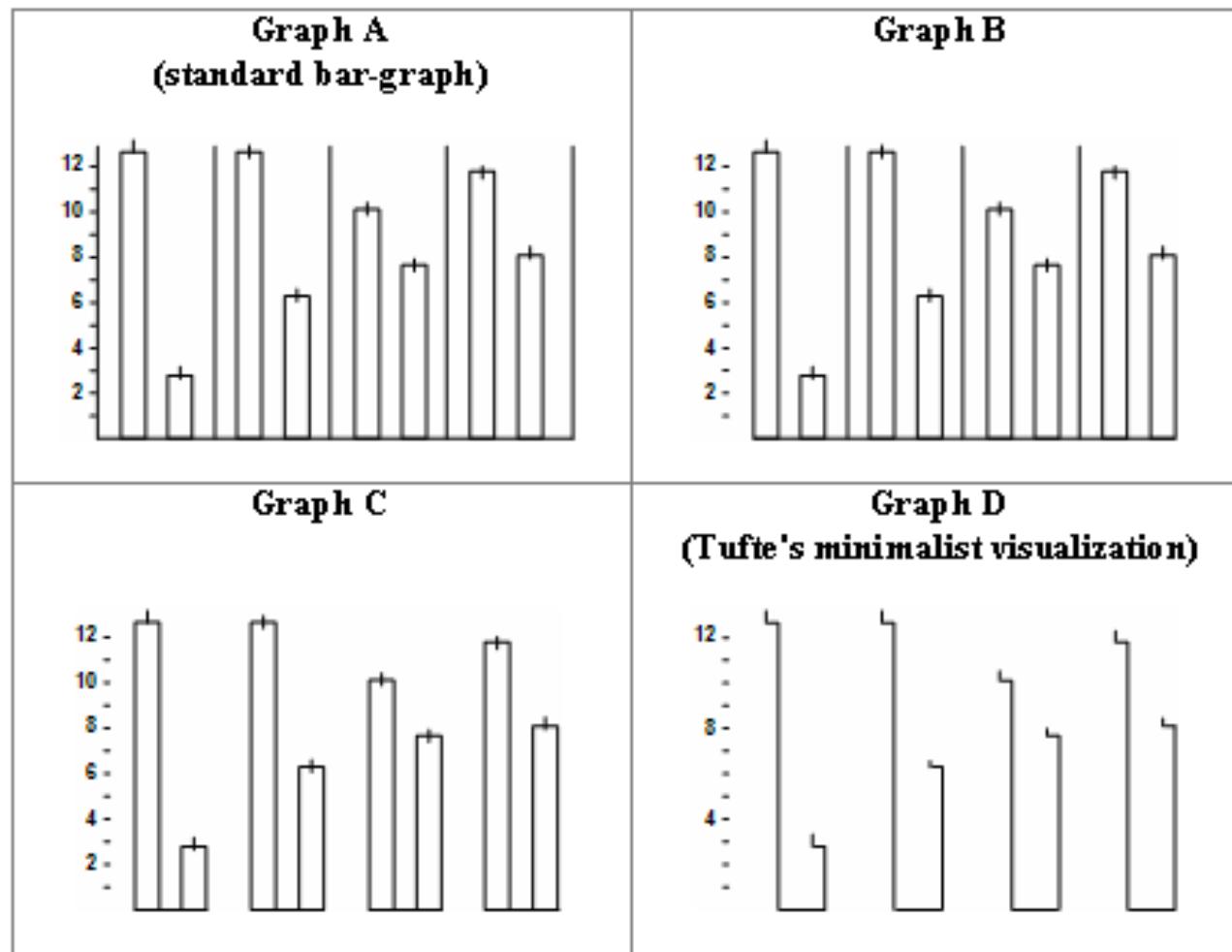
iv - the data-ink ratio

aka “remove2improve”

Data-ink ratio = $\frac{\text{data-ink}}{\text{total ink used to print the graphic}}$

= proportion of a graphic's ink devoted to the non-redundant display of data-information

= $1.0 - \text{proportion of a graphic that can be erased without loss of data-information}$.



Above all else show the data.
Maximize the data-ink ratio.
Erase non-data-ink.
Erase redundant data-ink.
Revise and edit.

tufte's minimalism

iv - the data-ink ratio

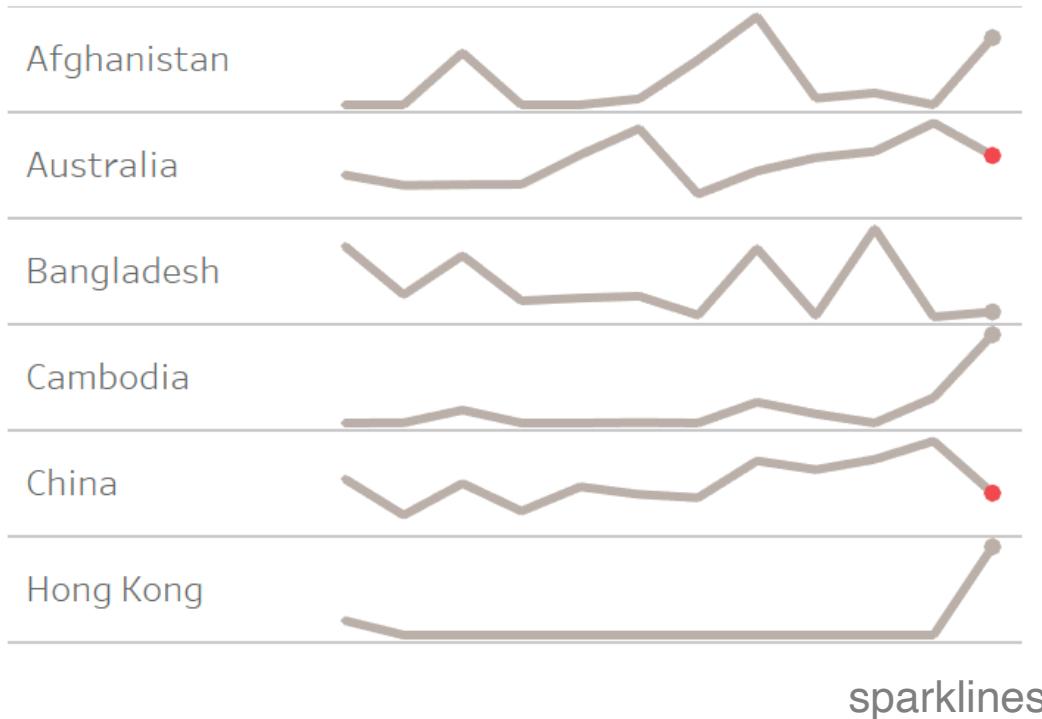
aka “remove2improve”

Data-ink ratio = $\frac{\text{data-ink}}{\text{total ink used to print the graphic}}$

= proportion of a graphic's ink devoted to the non-redundant display of data-information

= $1.0 - \text{proportion of a graphic that can be erased without loss of data-information}$.

what chart maximise data-ink ratio?



Above all else show the data.
Maximize the data-ink ratio.
Erase non-data-ink.
Erase redundant data-ink.
Revise and edit.

tufte's minimalism

adolf loos - 'ornament is a crime'

V - annotation

pictures are the focus, but words are important, too

labels are informative guides and they are essential in data design

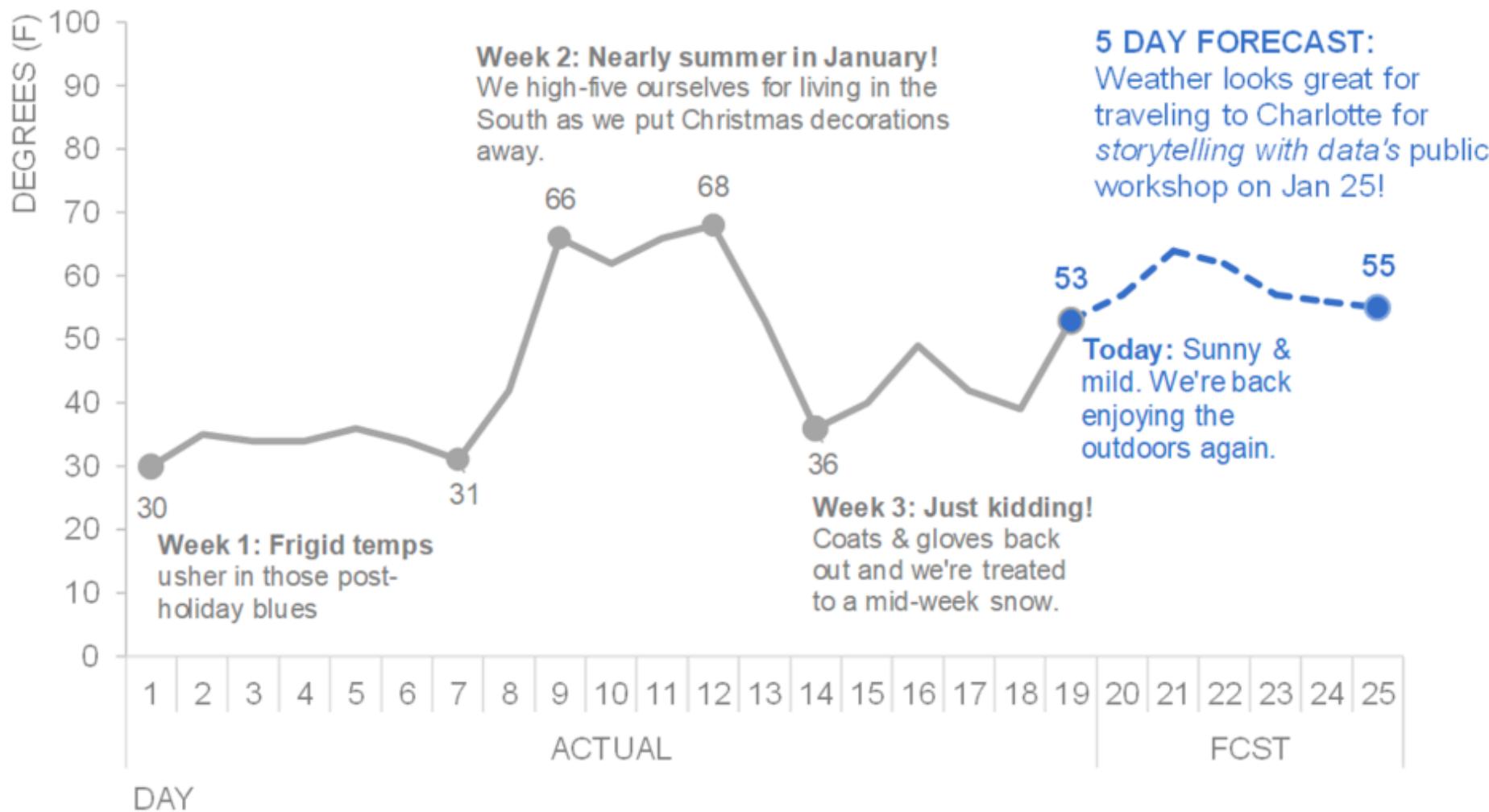
when labelling, focus on clarity, readability and differentiation

labels should stand out from data

V - annotation

January temperature trends in Charlotte

Daily high recorded in Jan 2018 (source: WeatherUnderground)



vi - micro/macro

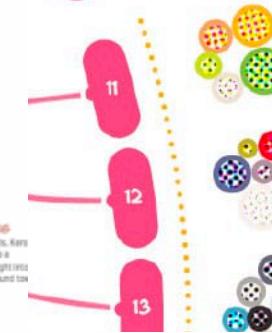
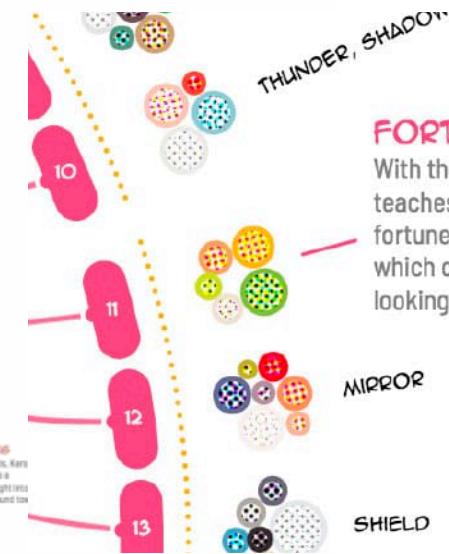
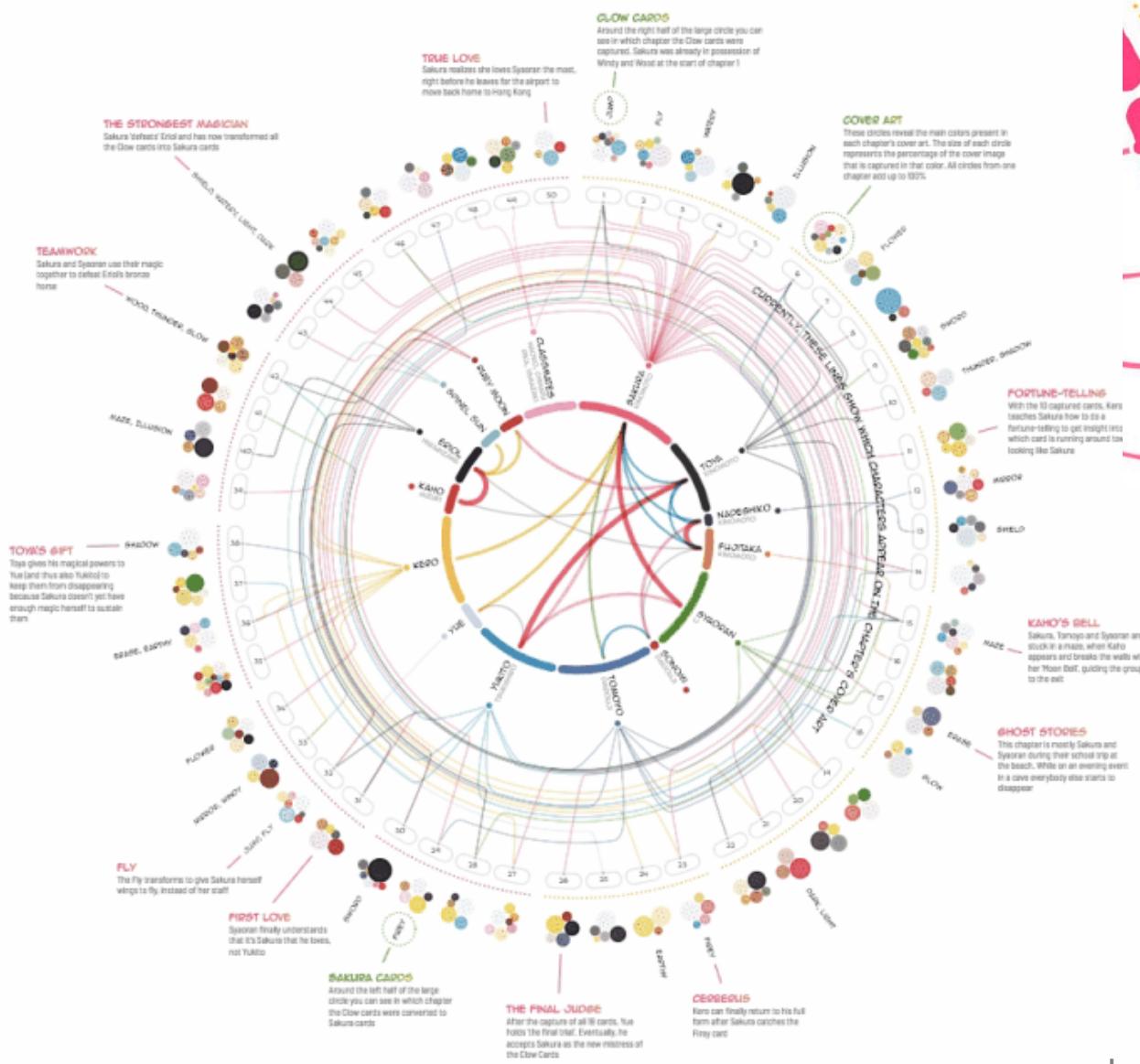
fine micro-level details become textures when viewed at the
macro level

the clarity of the macro is determined by the quality and the
quantity of the micro

ben schneiderman's mantra: overview first, then details on
demand

use interactive zooming when possible

vi - micro/macro



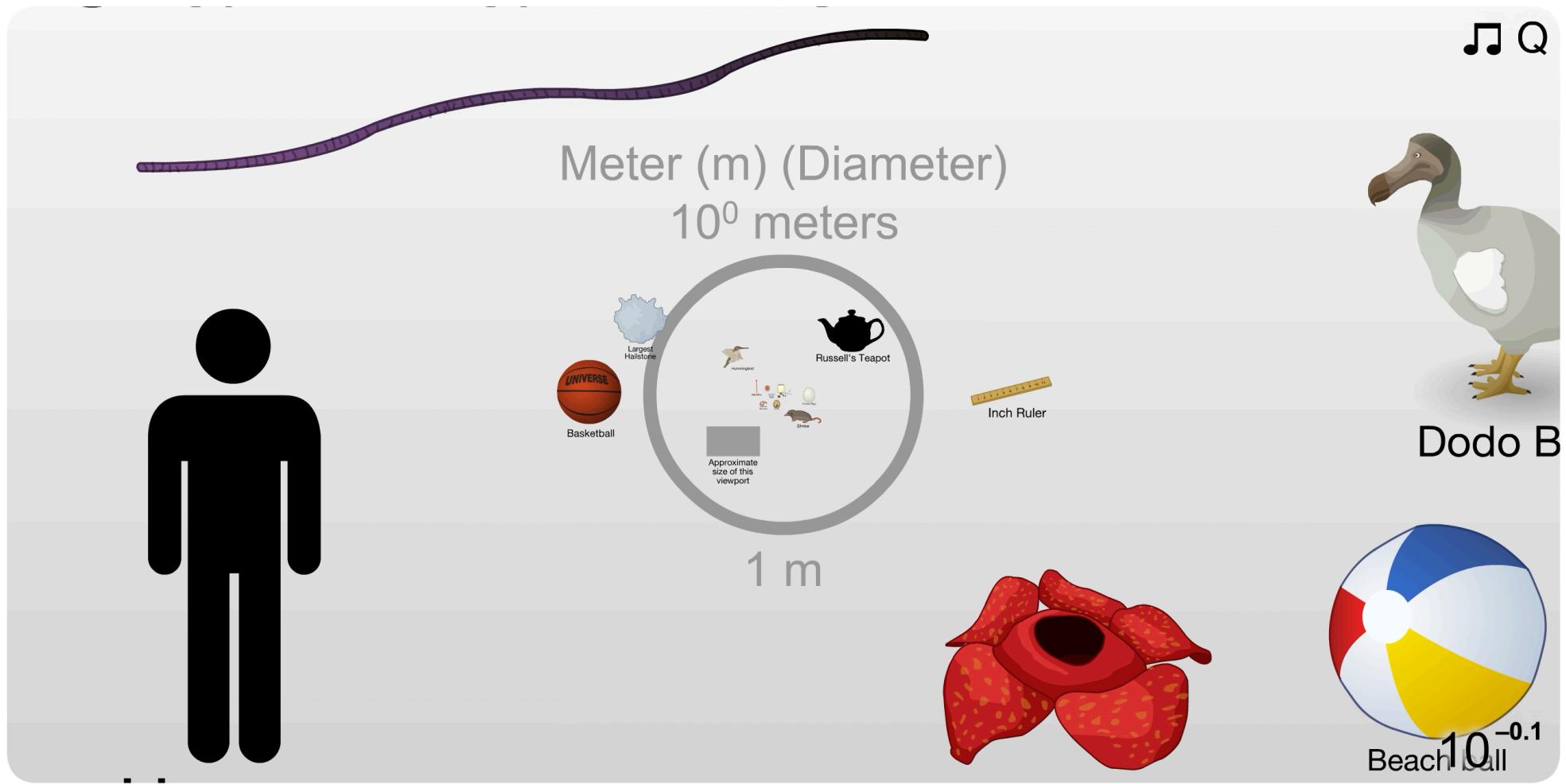
FORTUNE-TELLING

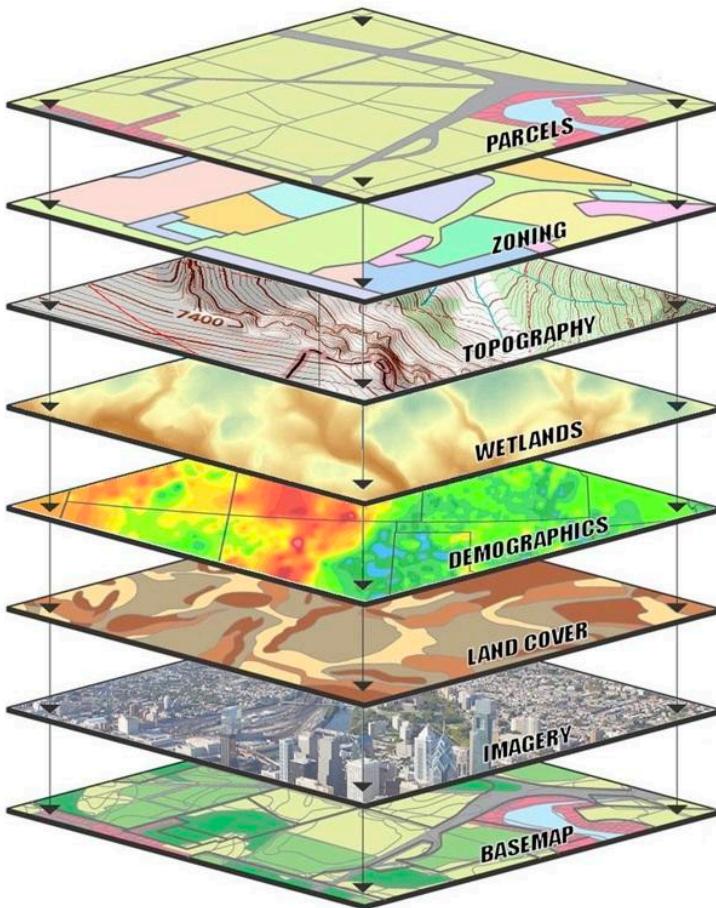
With the 10 captured cards, Kero teaches Sakura how to do a fortune-telling to get insight into which card is running around town looking like Sakura

an ode to cardcaptor sakura [n. bremer]

vi - micro/macro

The Scale of the Universe

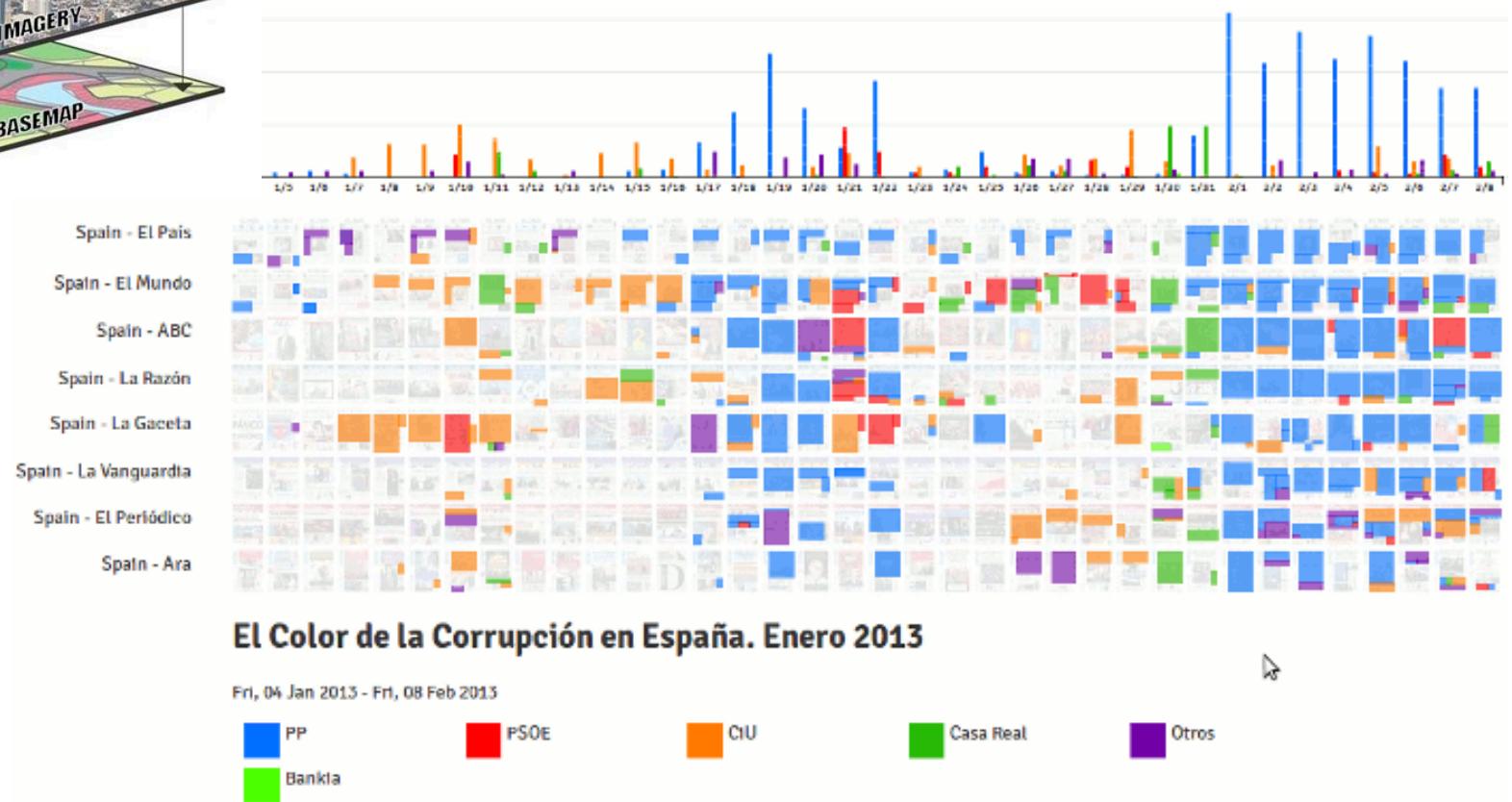




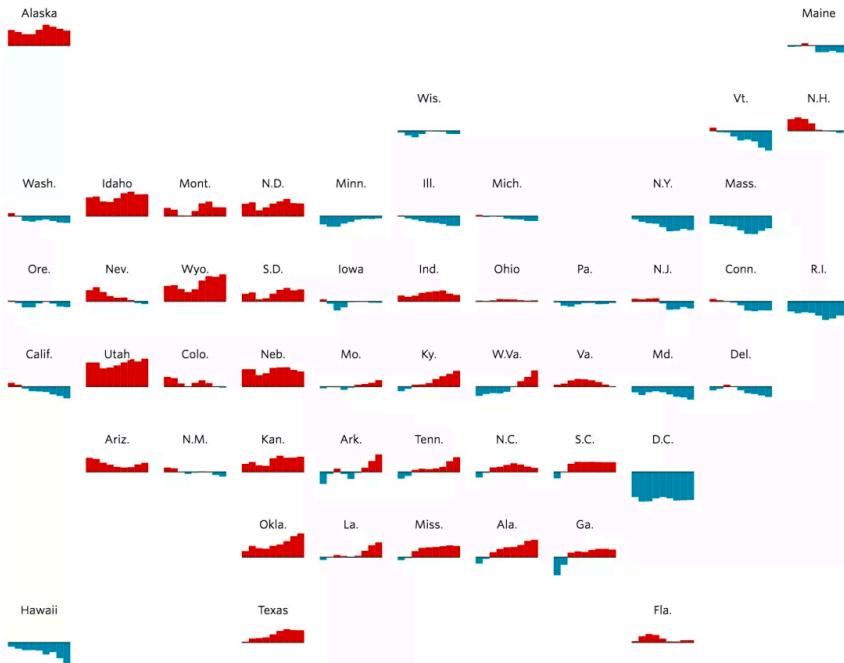
vii - layers

different elements of a visualisation
should have different appearance

use substantial difference in tones, colours
and weights



viii - small multiples

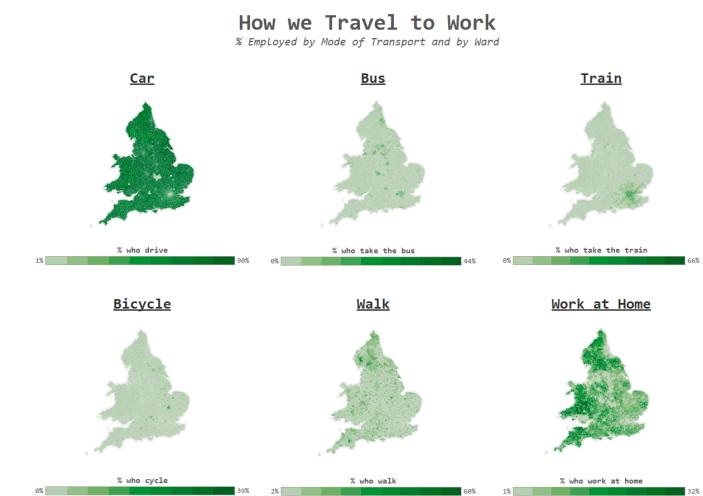
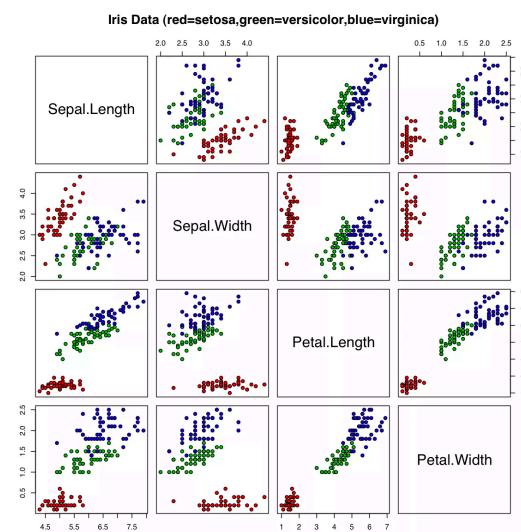


keep a consistent design, do not change appearance

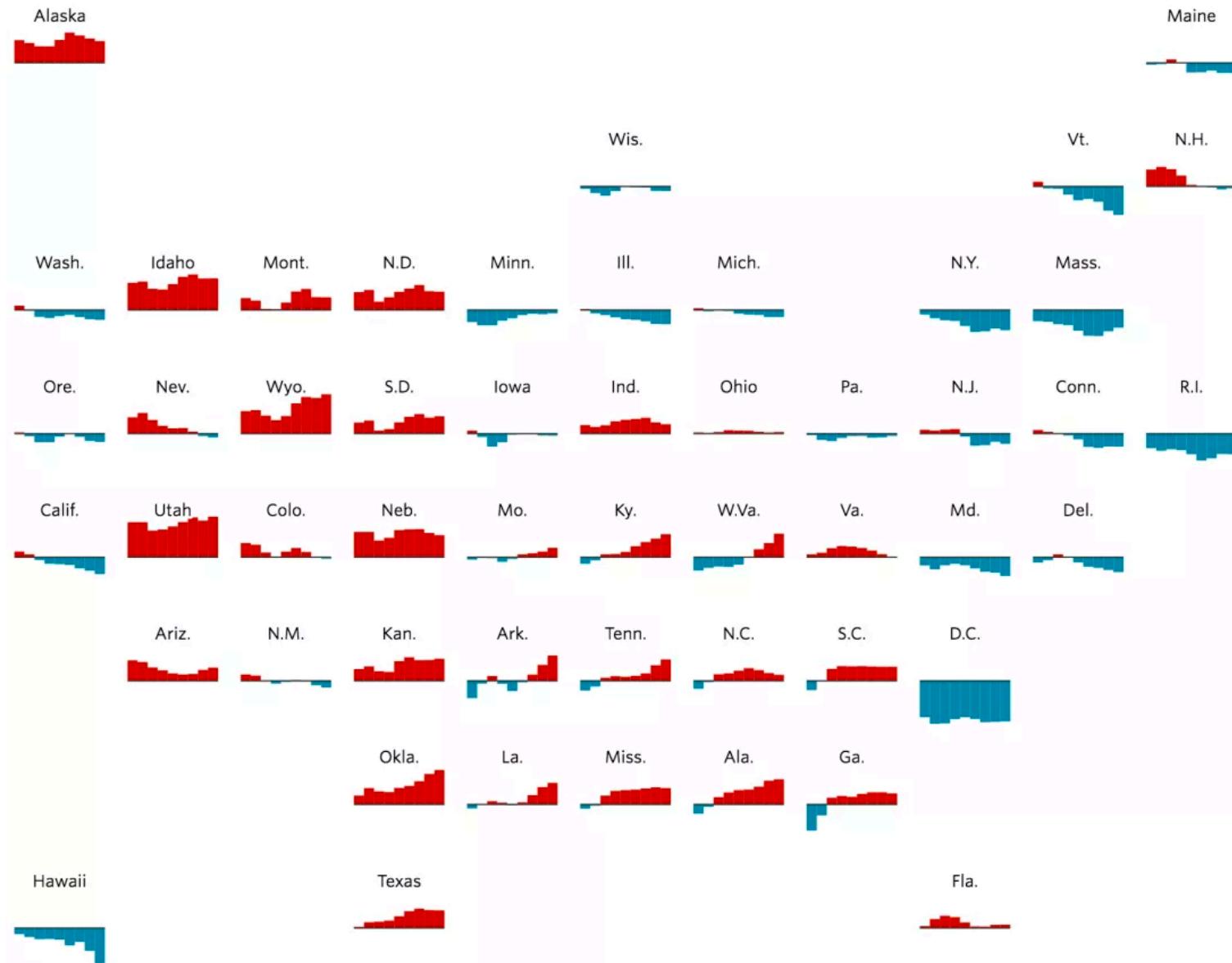
put emphasis on data, not on the visual design

changes in design may distract the viewer from irregularities in the data

use same annotations, same axes



viii - small multiples

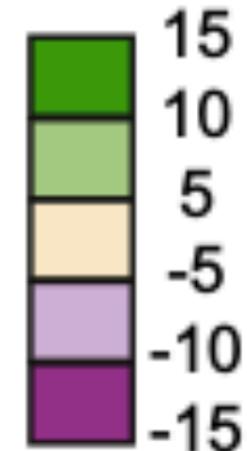
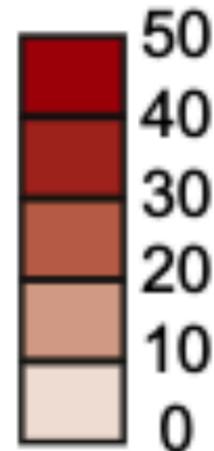
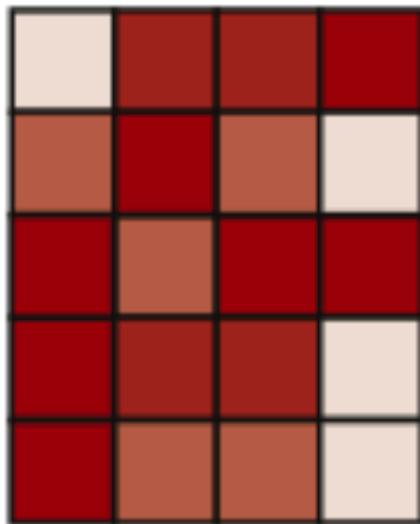


ix - colour

colour can be helpful if used properly

colour can be harmful if used naively

rainbows and gradients have different purposes



ix - colour

colour can be harmful

Atmos. Chem. Phys., 6, 5183-5191, 2006
https://doi.org/10.5194/acp-6-5183-2006
© Author(s) 2006. This work is licensed under
the Creative Commons Attribution-NonCommercial-ShareAlike 2.5 License.



Volume 6, issue 12

Article Peer review Metrics Related articles

15 Nov 2006

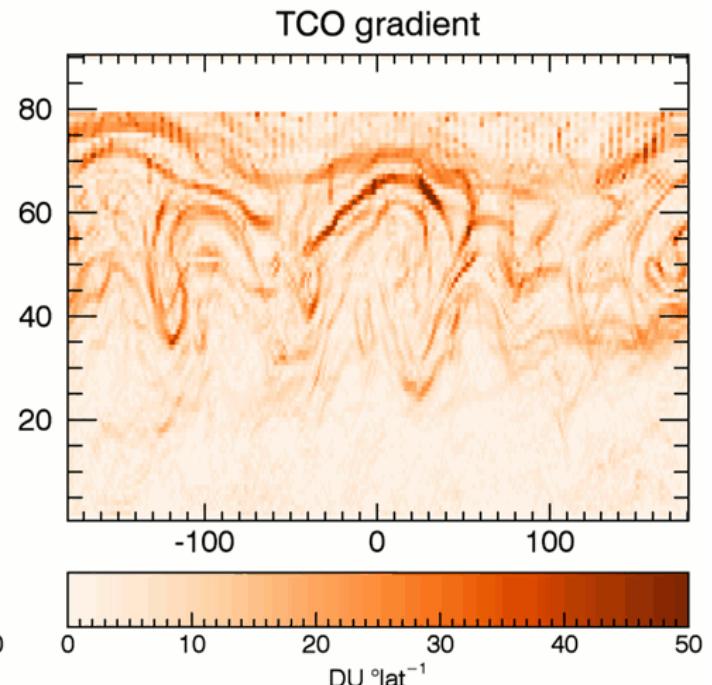
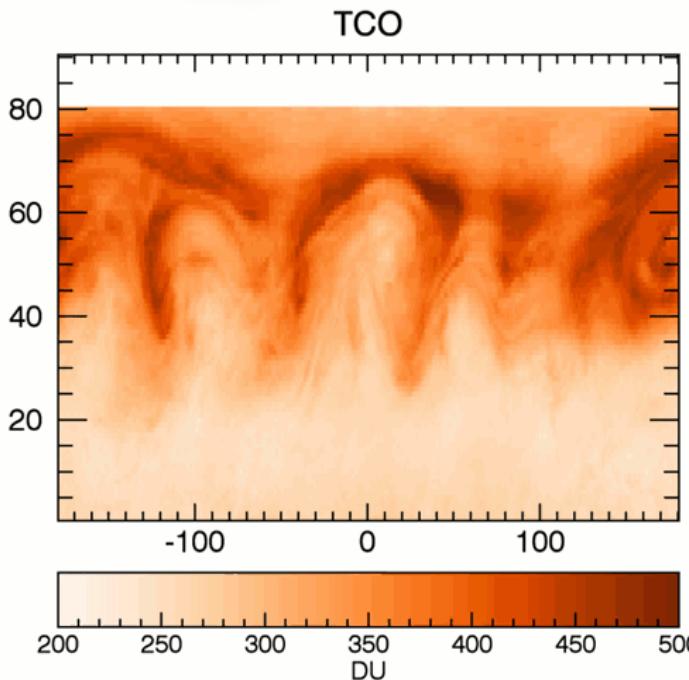
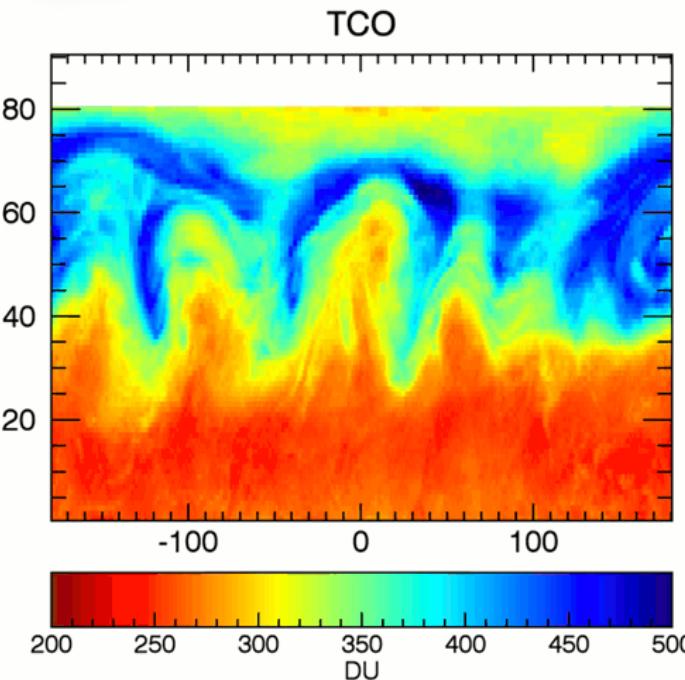
The total ozone field separated into meteorological regimes – Part II: Northern Hemisphere mid-latitude total ozone trends

R. D. Hudson¹, M. F. Andrade², M. B. Follette¹, and A. D. Frolov³

¹Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD 20742-2425, USA

²Atmospheric Physics Laboratory, Campus Universitario Cota-Cota, Calle 27, Universidad Mayor de San Andres, La Paz, Bolivia

³STG, Inc., 11710 Plaza America Drive, Reston, VA, USA



the blue front is an artefact, evident in rainbow but not present in gradient

IEEE Comput Graph Appl. 2007 Mar-Apr;27(2):14-7.

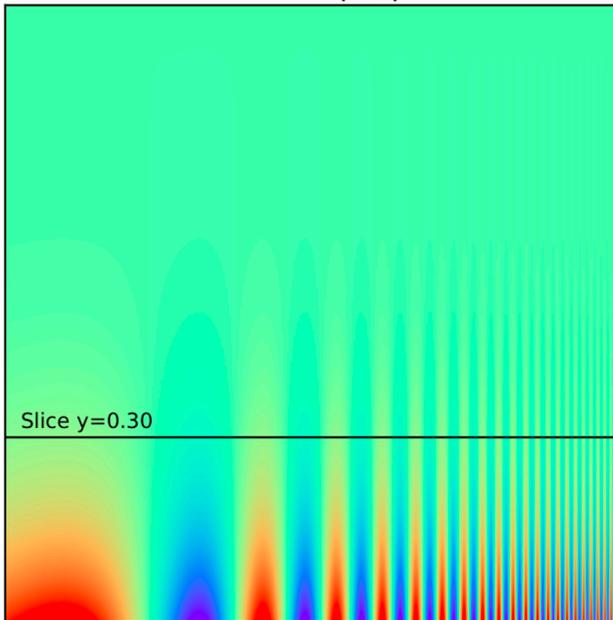
Rainbow color map (still) considered harmful.

Borland D¹, Taylor MR 2nd.

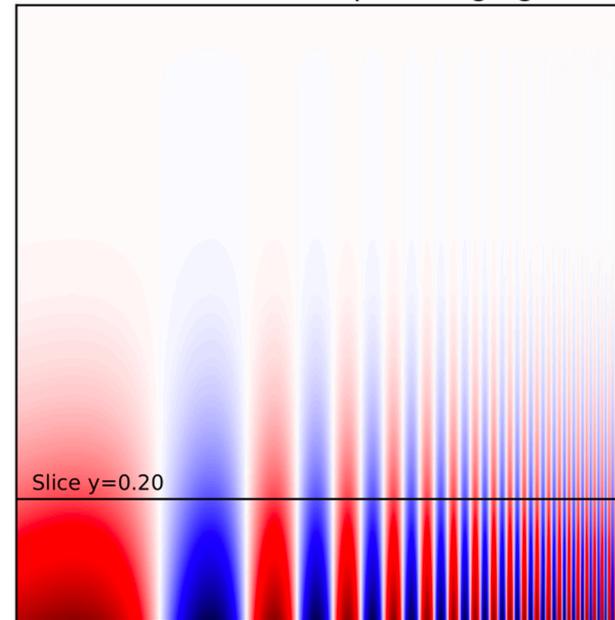
ix - colour

colour can be harmful

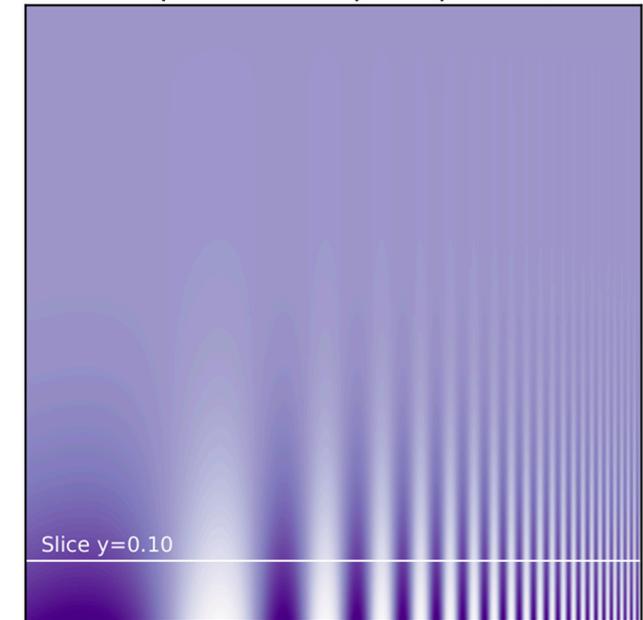
Rainbow colormap (qualitative)



Seismic colormap (diverging)



Purples colormap (sequential)

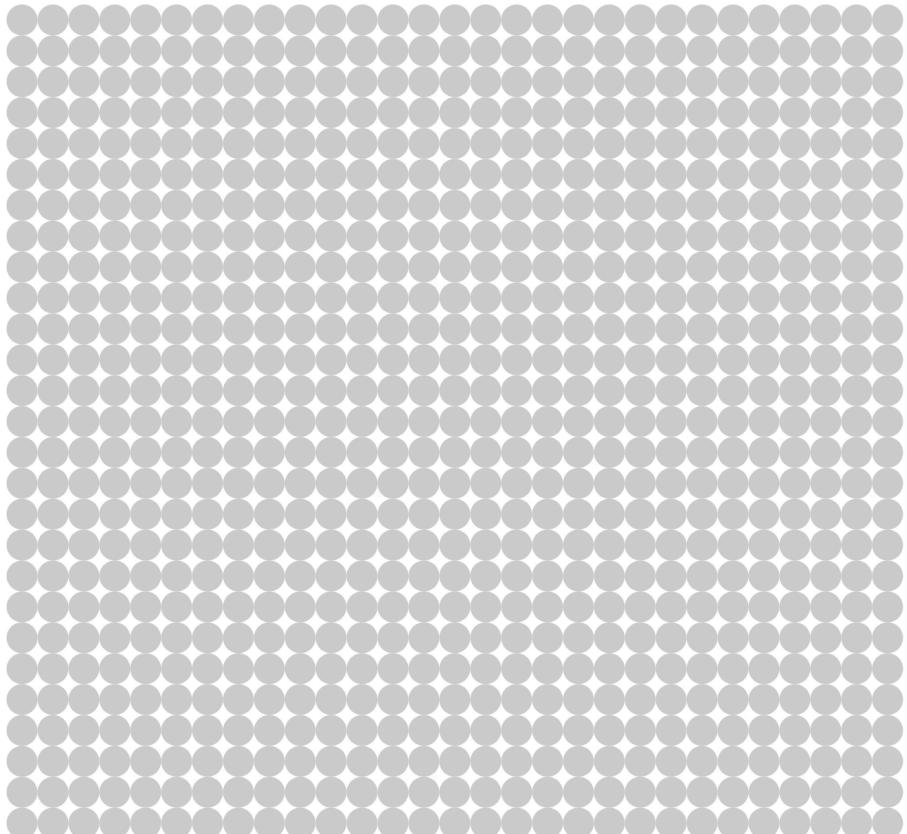


rainbow & seismic hide details in the bottom-right part

X - narrative

what story is your dataviz telling?
what is happening, over time and across space?
is it explicit, or left to the reader to explore?

I am **female** ▾ **White** ▾ and currently  years old. LIVE

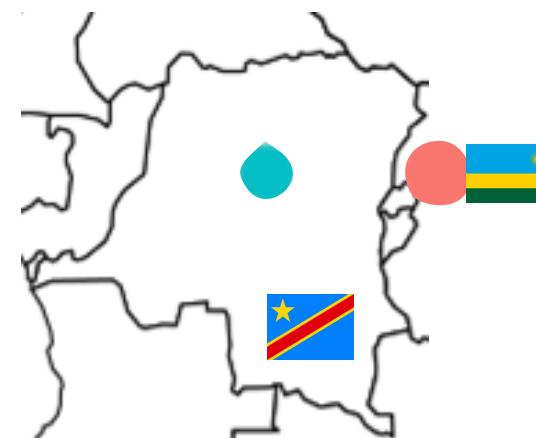
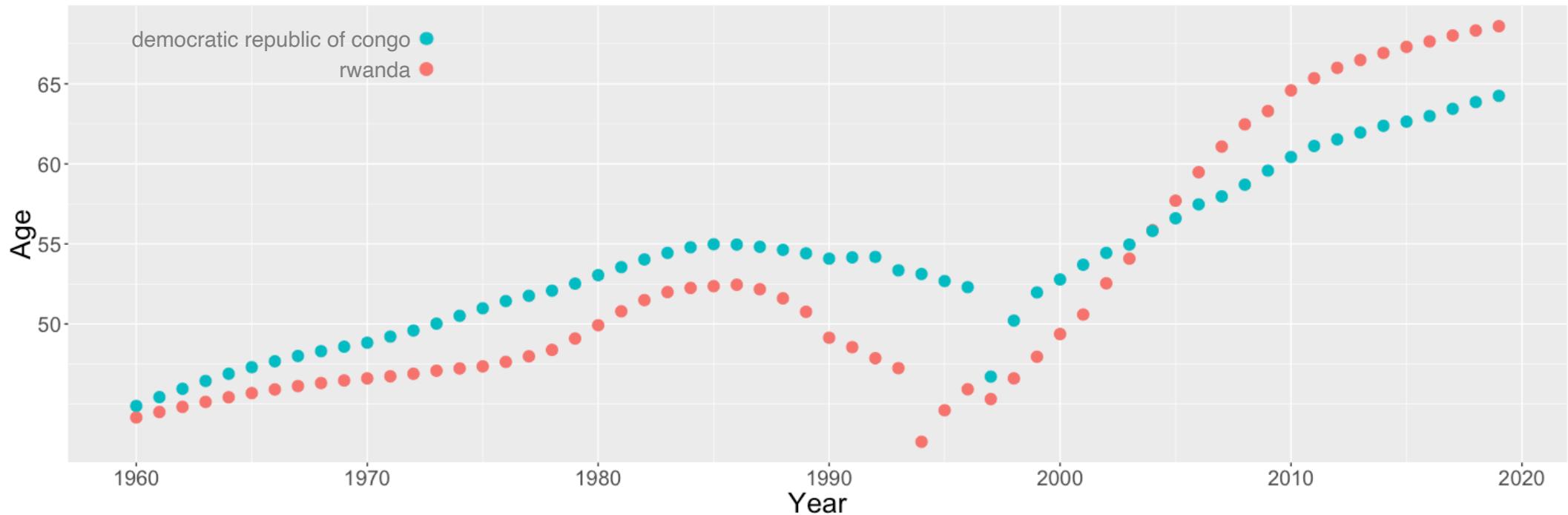


AGE OF DEATH	DIED SO FAR	STILL ALIVE
0 years	1%	99%

CAUSE OF DEATH SO FAR	
0%	Infection
0%	Cancer
0%	Blood
0%	Endocrine
0%	Mental
0%	Nervous
0%	Circulatory
0%	Respiratory
0%	Digestive
0%	Musculoskeletal
0%	Genitourinary
0%	Perinatal
0%	Congenital
0%	External Causes
0%	Other

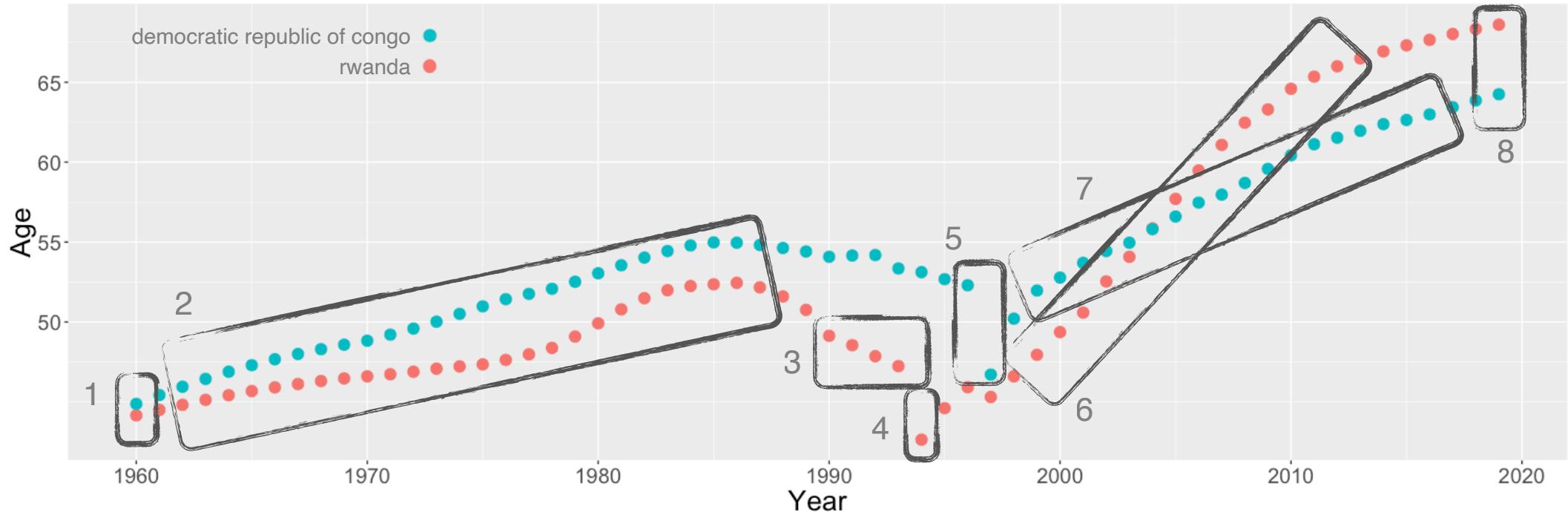
x - narrative

life expectancy



x - narrative

life expectancy



1. similar conditions out of the post-colonialism phase

2. similar trends from 60s to mid-80s, with congo slightly better

3. decline in rwanda due to internal turmoils

4. hutu/tutsi genocide 1994

5. civil war in congo (then zaire)

6. quick development of economy in rwanda, poverty drops from ~60% to ~40%

7. slow recovery in congo, due to several wars and government instability

8. increasing gap in favour of rwanda

modern version

for compbiol... but not so different after all

- i. know your audience
- ii. identify your message
- iii. adapt the figure to the support medium
- iv. captions are not optional
- v. do not trust the defaults
- vi. use colour effectively
- vii. do not mislead the reader
- viii. avoid chart junk
- ix. message trumps beauty
- x. get the right tool



OPEN ACCESS

EDITORIAL

Ten Simple Rules for Better Figures

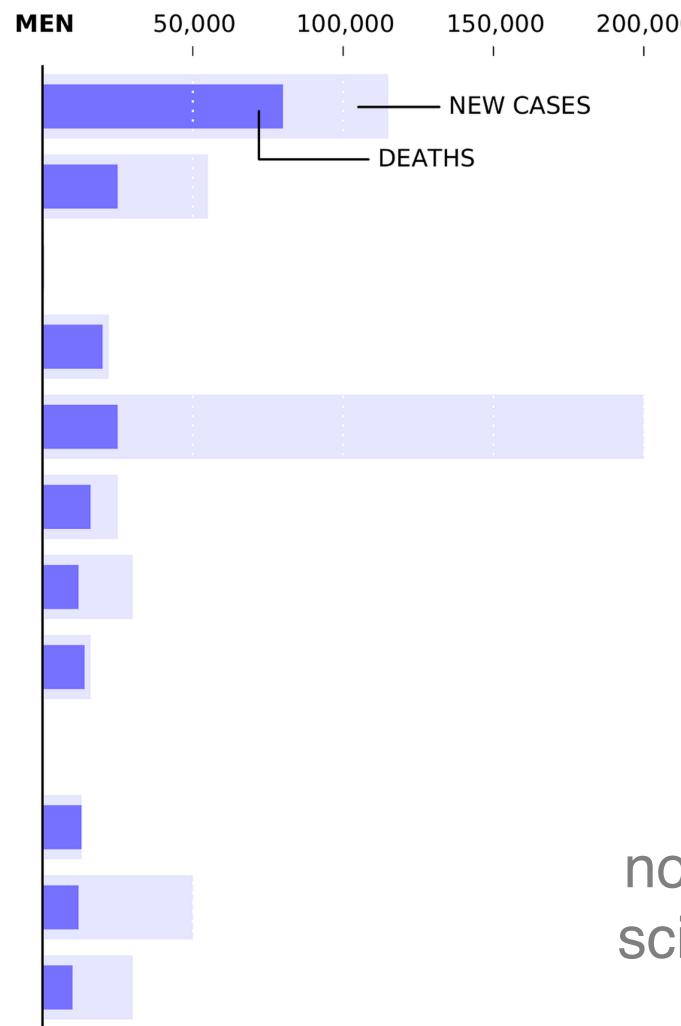
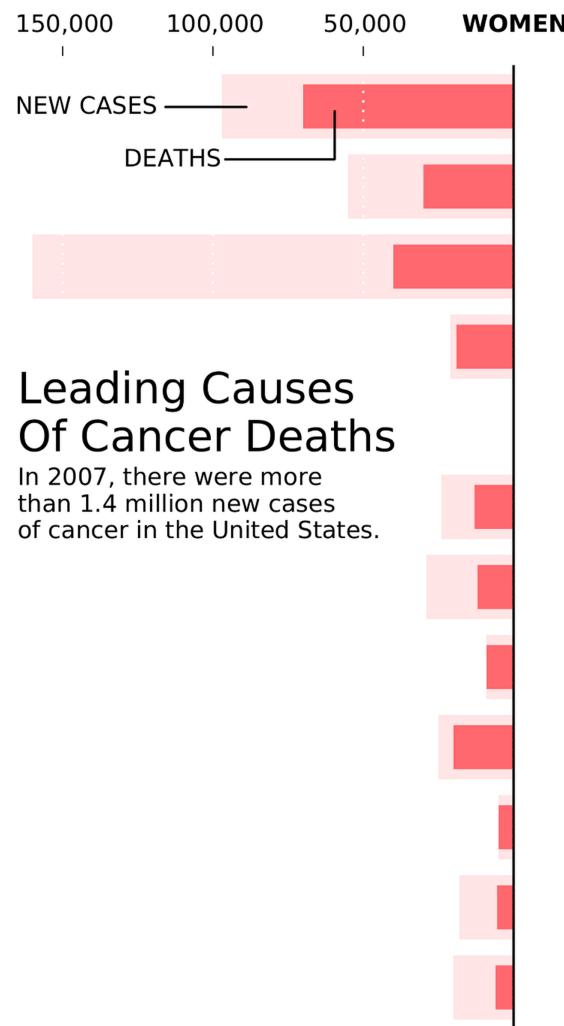
Nicolas P. Rougier , Michael Droettboom, Philip E. Bourne

Published: September 11, 2014 • <https://doi.org/10.1371/journal.pcbi.1003833>

modern version

for compbiol... but not so different after all

i. know your audience

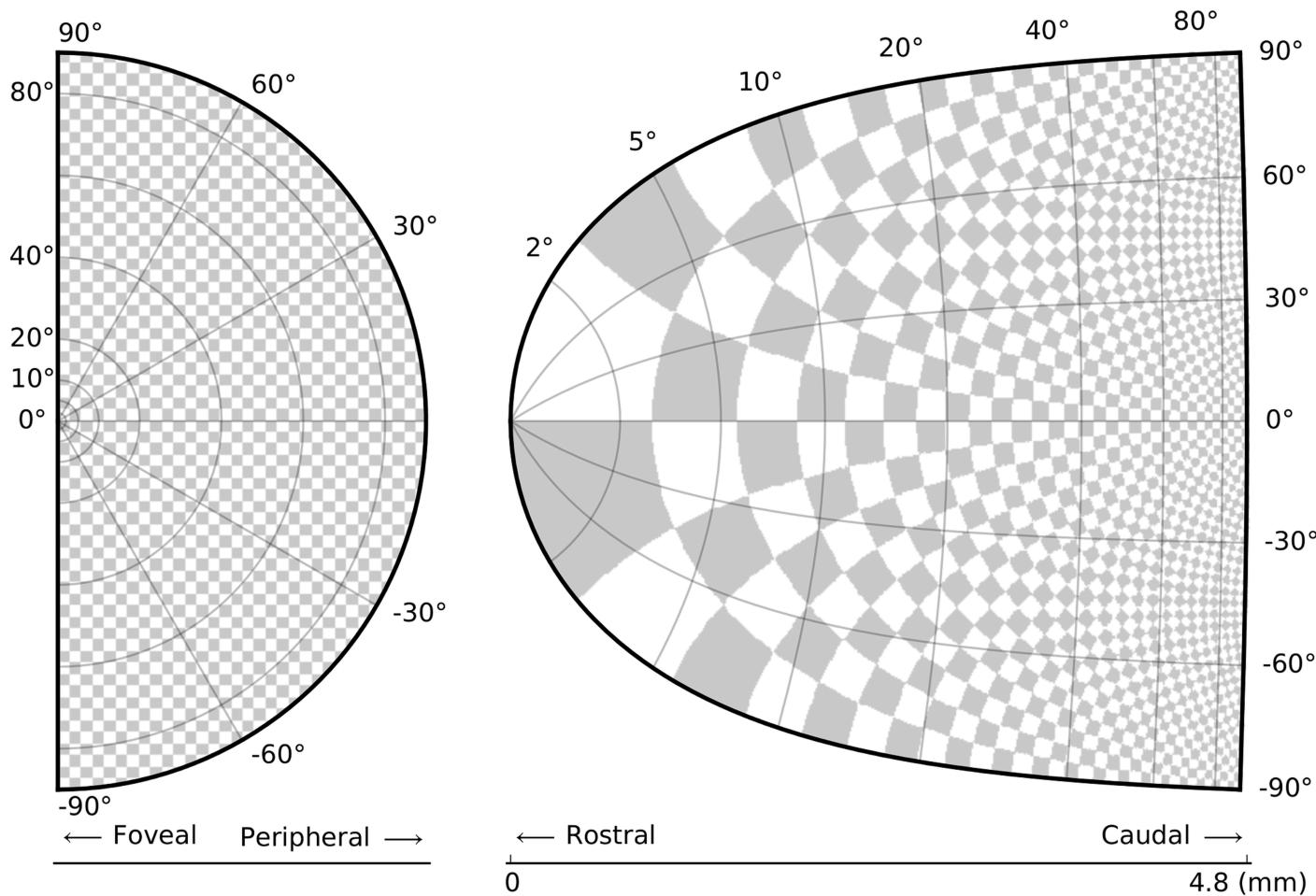


not suitable for scientific paper

modern version

for compbiol... but not so different after all

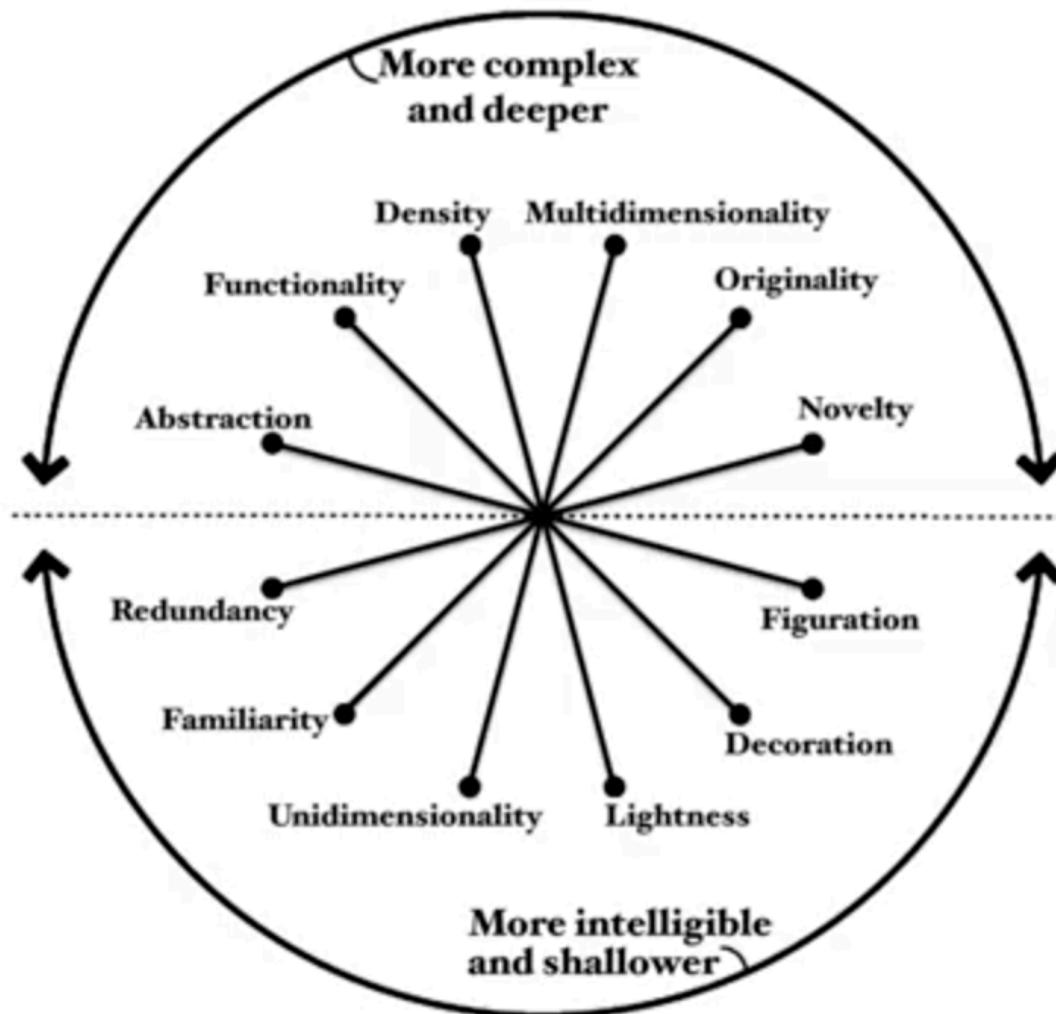
ii. identify your message



the checkerboard pattern highlights the logarithmic mapping involved in the model

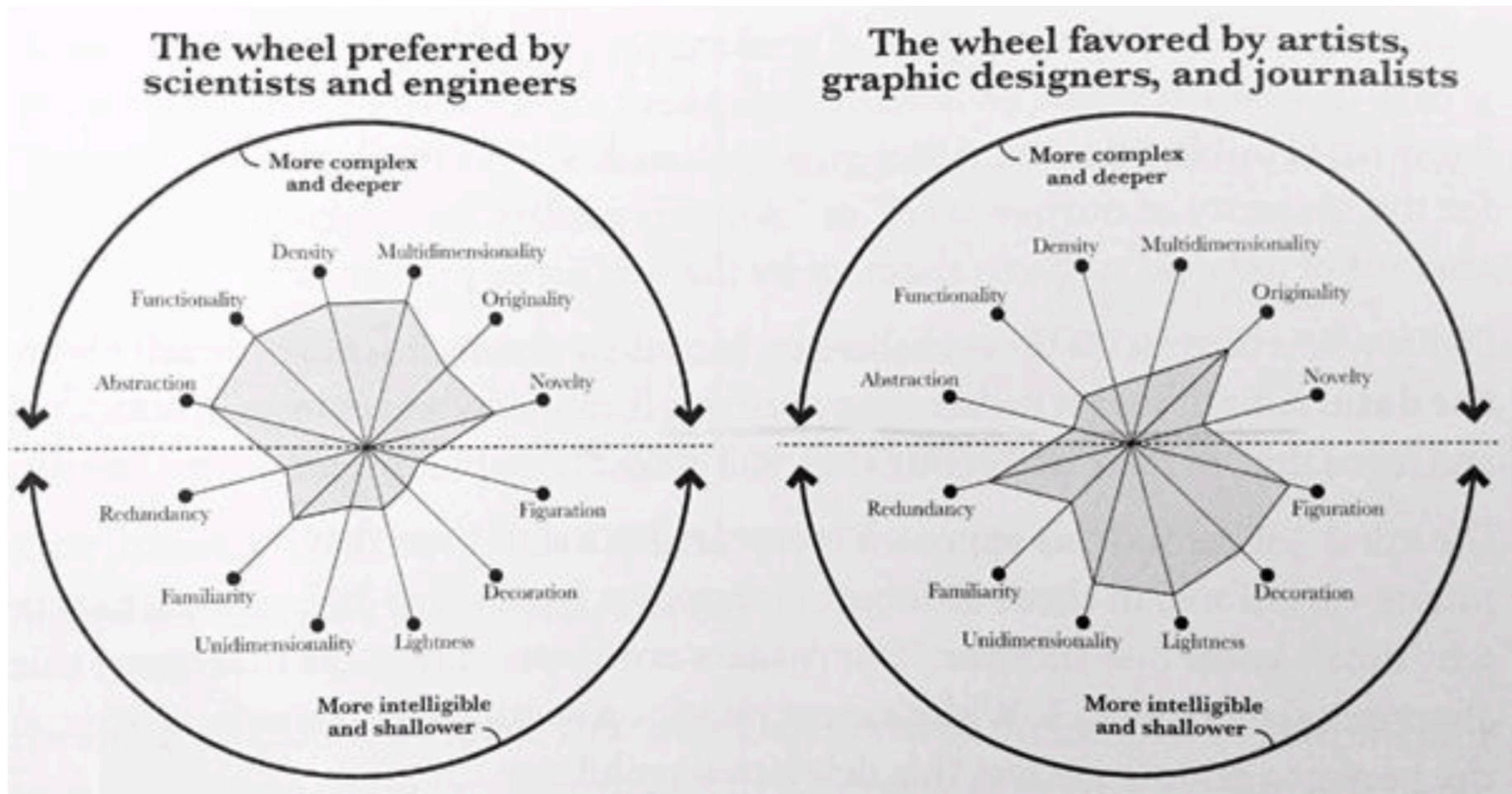
cairo's dataviz wheel

a metaviz to judge dataviz



cairo's dataviz wheel

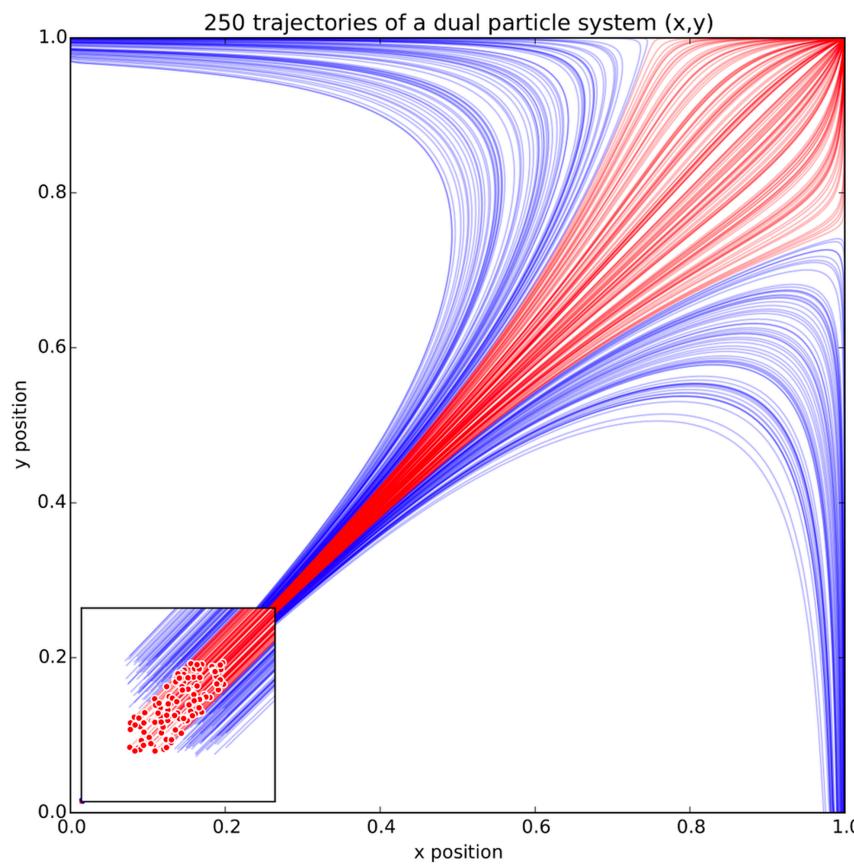
a metaviz to judge dataviz



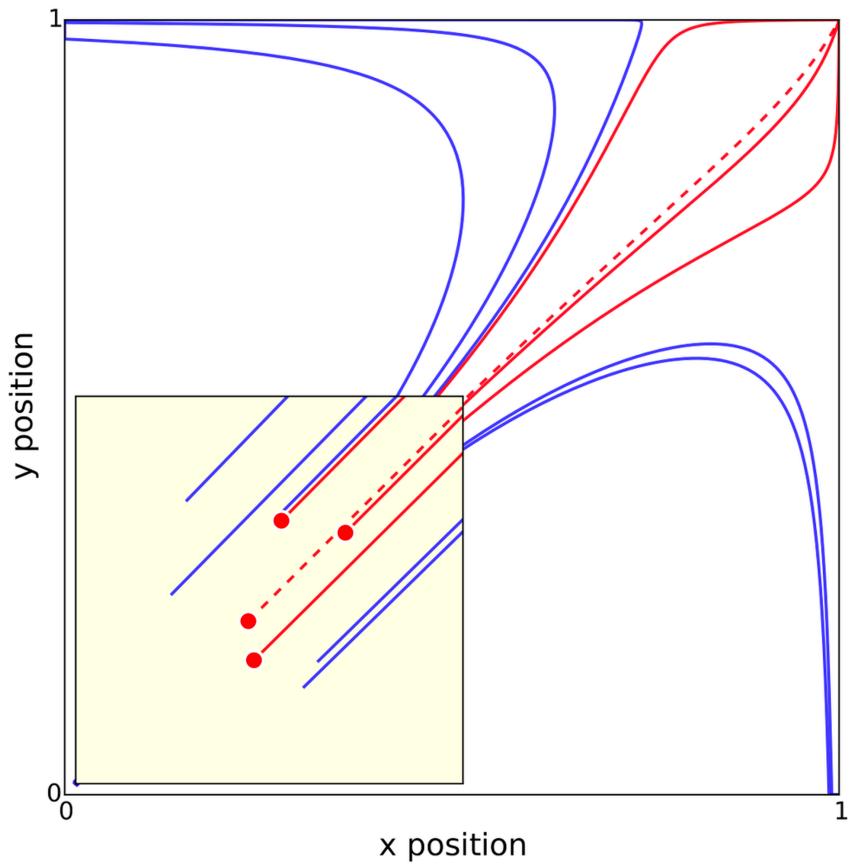
modern version

for compbiol... but not so different after all

iii. adapt the figure to the support medium



paper

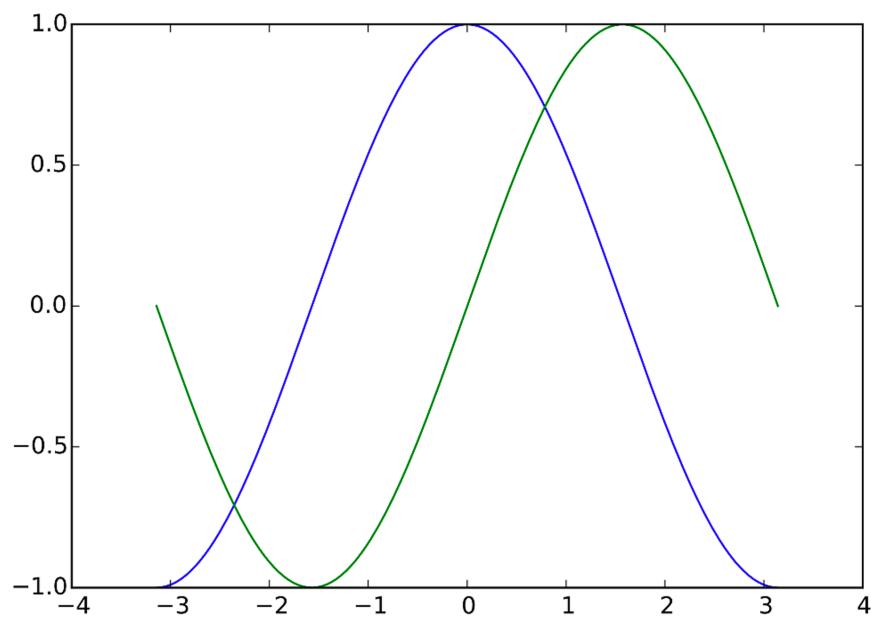


presentation

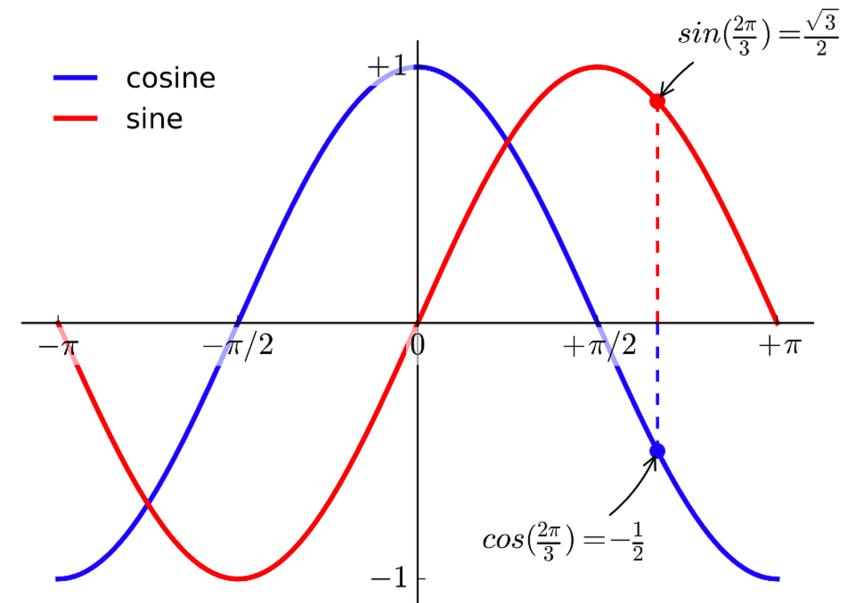
modern version

for compbiol... but not so different after all

v. do not trust the defaults



defaults

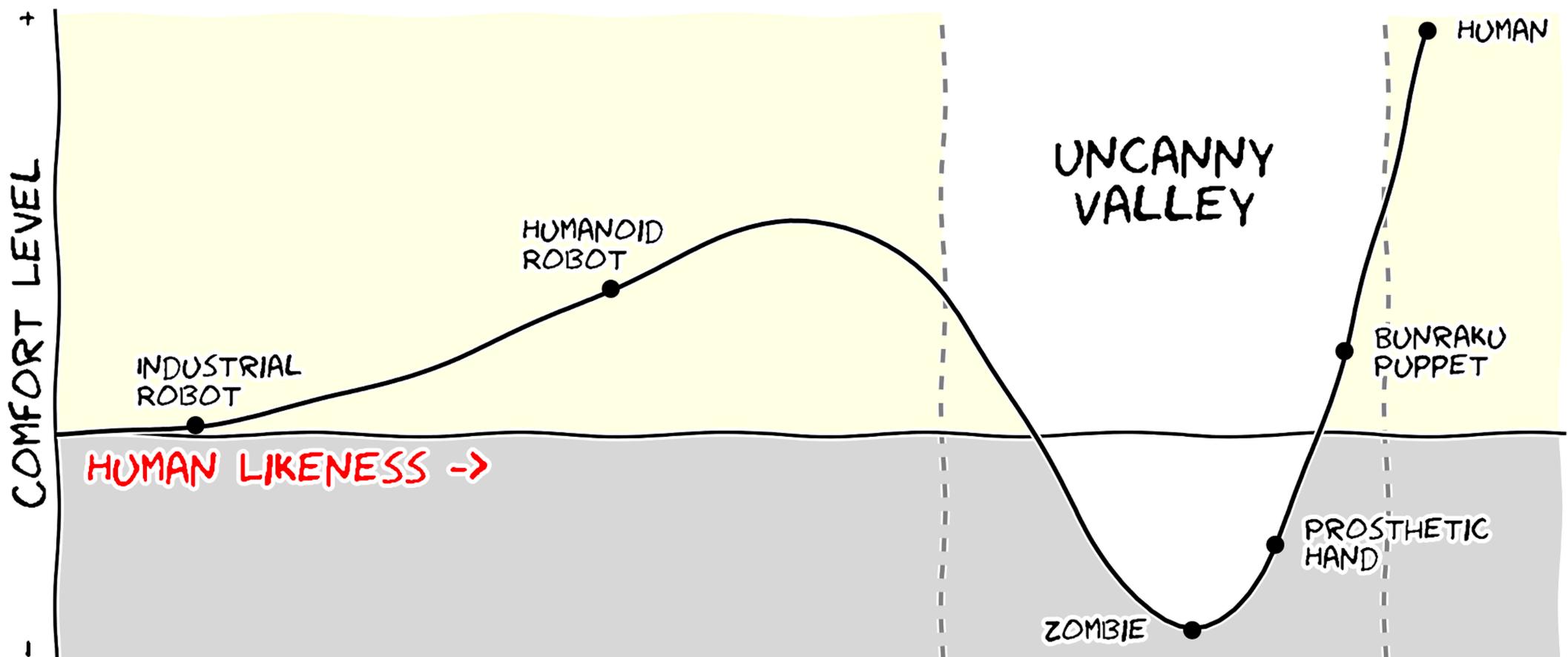


custom
(better)

modern version

for compbiol... but not so different after all

ix. message trumps beauty



questionable aesthetic but clear message: hypothetical data, no ticks, cartoonish font, sketched look

AI and data viz



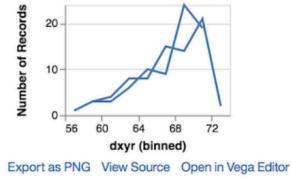
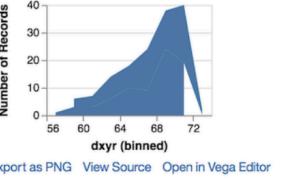
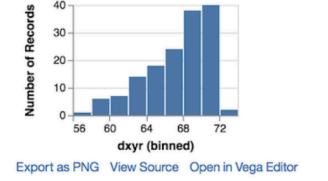
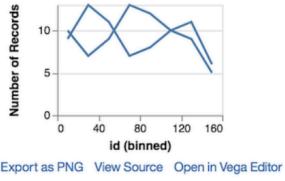
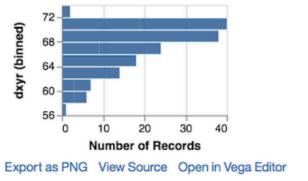
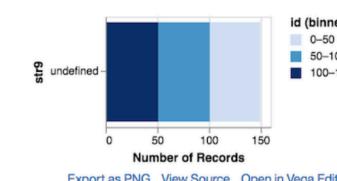
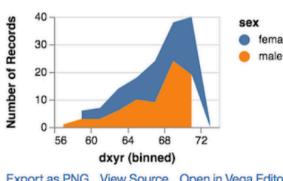
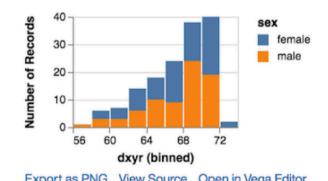
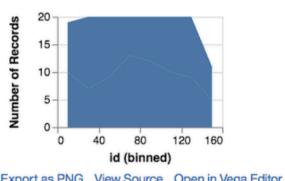
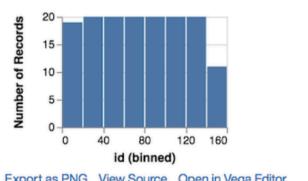
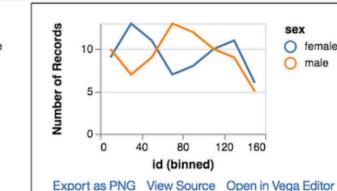
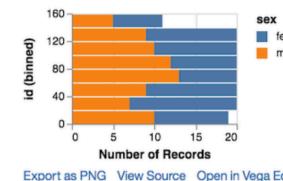
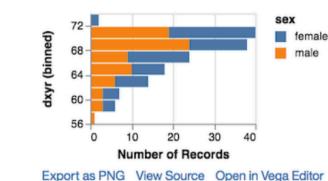
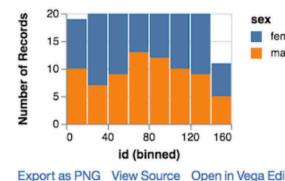
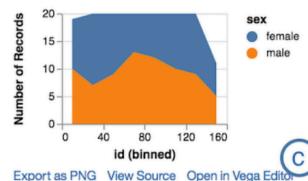
Automation

Generate Visualizations

Load random data or paste a valid JSON array.

```
graph TD; A[Load Dataset] -- a --> B[Generate]; B -- b --> A
```

```
"age":52,"abl":NA,"creat":0.6,"death":1,"dxy":65,"futime":5084,"hgb":12.6,"id":146,"mspike":2.1,"pcdx":NA,"pctime":NA,"sex":female  
"age":74,"abl":2.6,"creat":1,"death":1,"dxy":70,"futime":7246,"hgb":14.5,"id":147,"mspike":1.5,"pcdx":NA,"pctime":NA,"sex":female}  
"age":56,"abl":1.1,"death":1,"dxy":57,"futime":5930,"hgb":14.8,"id":148,"mspike":2.4,"pcdx":NA,"pctime":NA,"sex":female}  
"age":47,"abl":31,"creat":14,"death":1,"dxy":65,"futime":6147,"hgb":11.4,"id":149,"mspike":1.7,"pcdx":NA,"pctime":NA,"sex":female}  
"age":68,"abl":3.1,"creat":12,"death":1,"dxy":68,"futime":9607,"hgb":13.3,"id":150,"mspike":1.8,"pcdx":NA,"pctime":NA,"sex":male]]
```

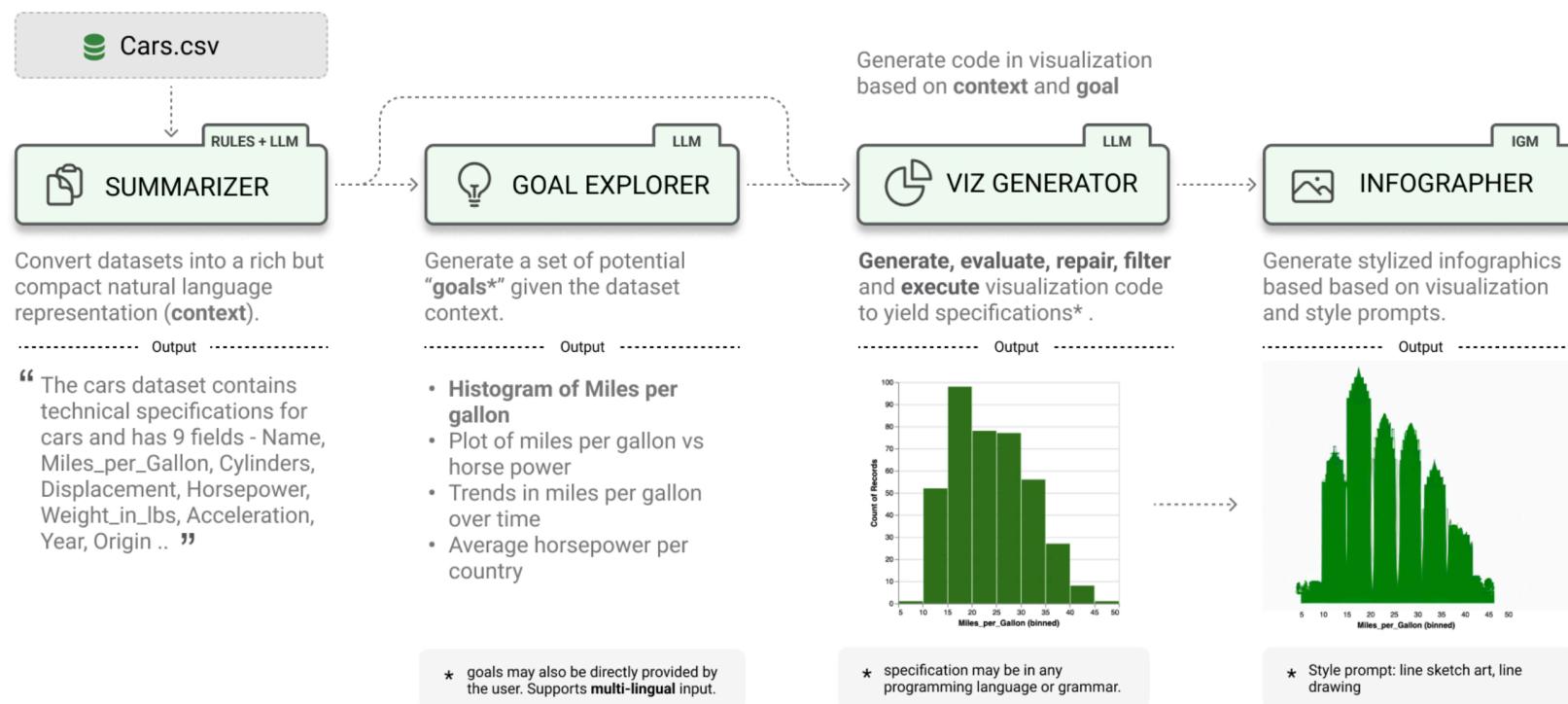


AI and data viz



System Architecture

System architecture for LIDA



Architecture for LIDA

LLM = Large Language Model | IGM = Image Generation Model

LIDA, Dibia, ACL 2023 Conference
<https://microsoft.github.io/lida/>