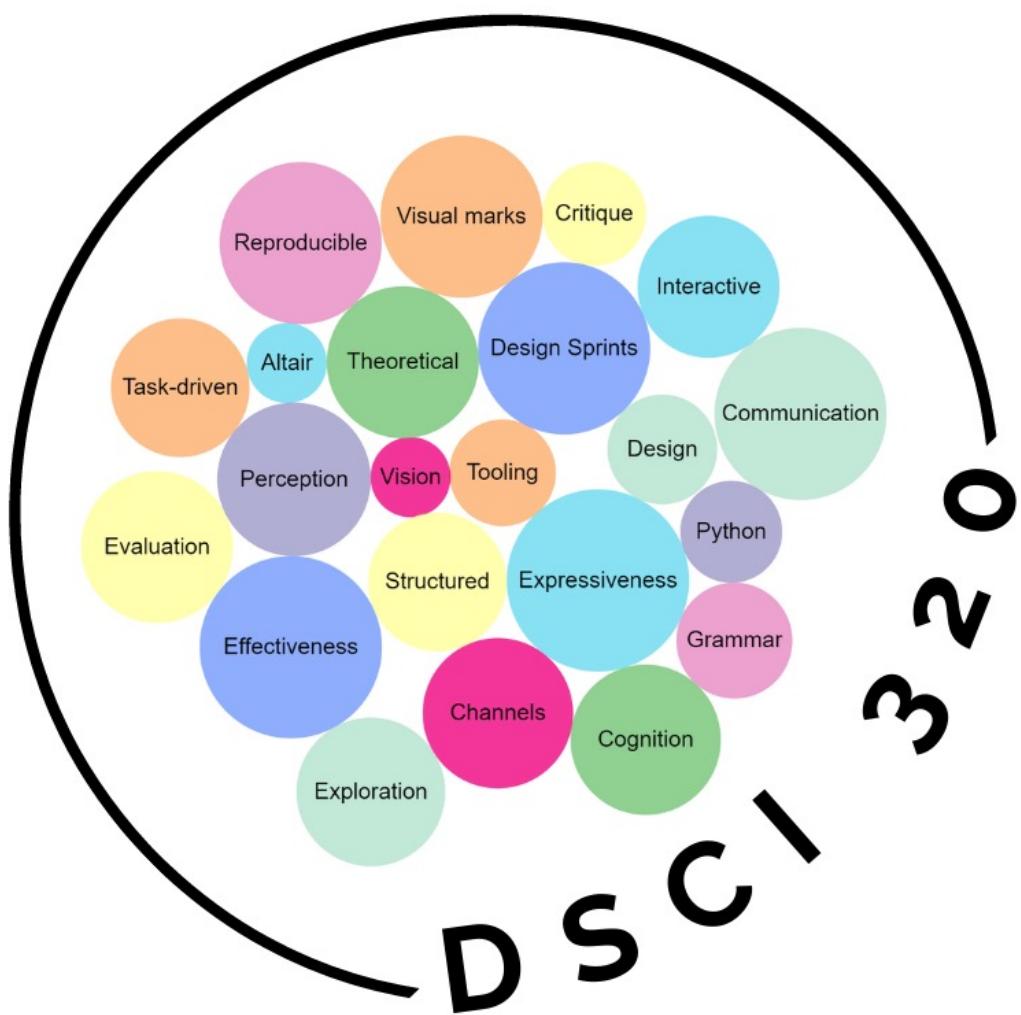


Visualization for Data Science

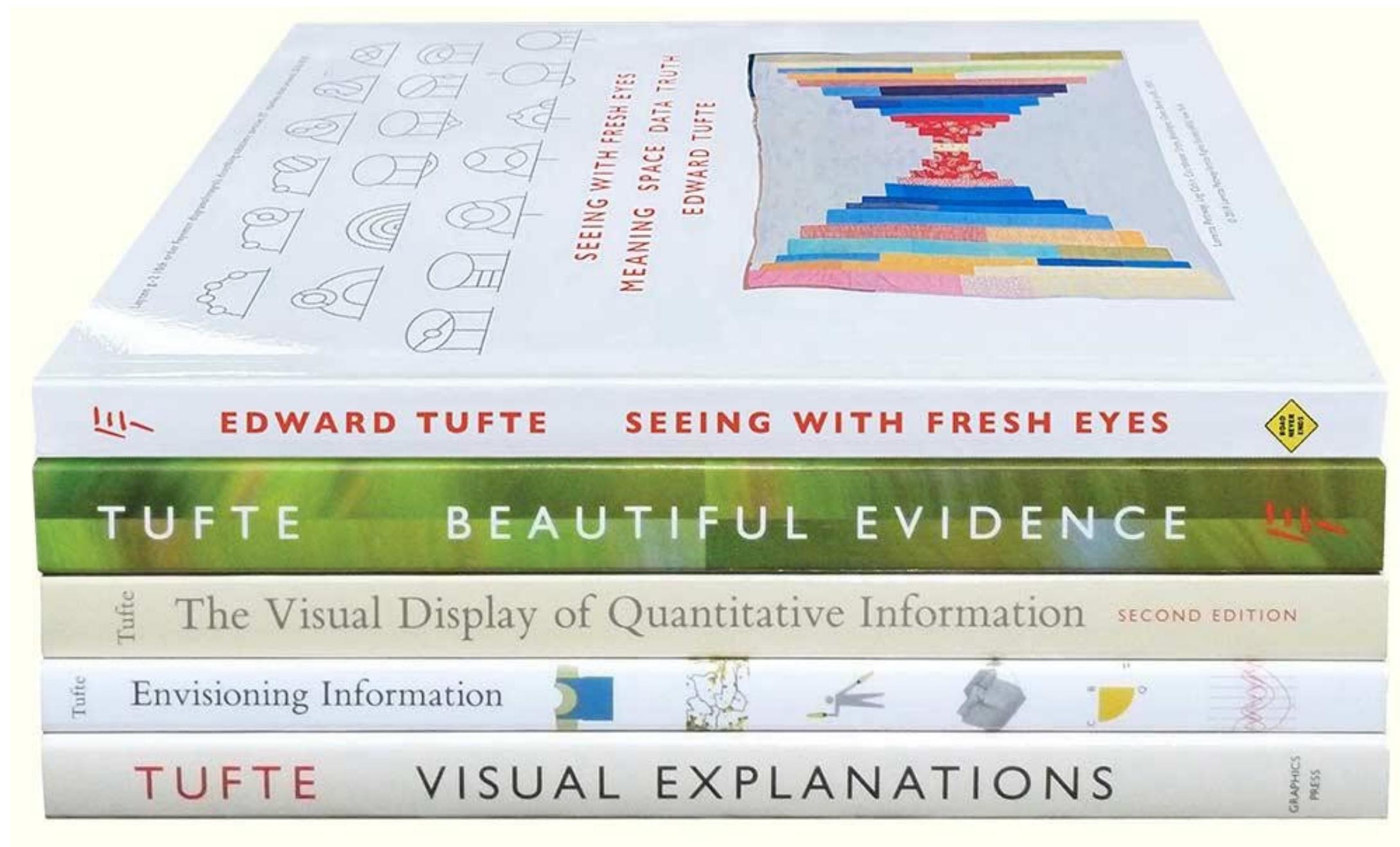
Design Rules of Toe



Adminstrivia

See ED Discussion

Edward Tufte



Principles of Effective Visualizations

Principle

Definition

Examples

Proportional Ink

The amount of ink used to indicate a value should be proportional to the value itself

Truncating a position channel on a bar chart to exaggerate the difference between bars violates the principle of proportional ink.

Data:ink ratio

Remove distracting visual elements to focus attention on the data

Lighten line weights, remove backgrounds, never use 3D or special effects, avoid unnecessary/redundant labels.

Labels & legends

Use axes labels and titles to highlight/communicate data

Never leave your data column names as axes labels! Generally good to add a title.

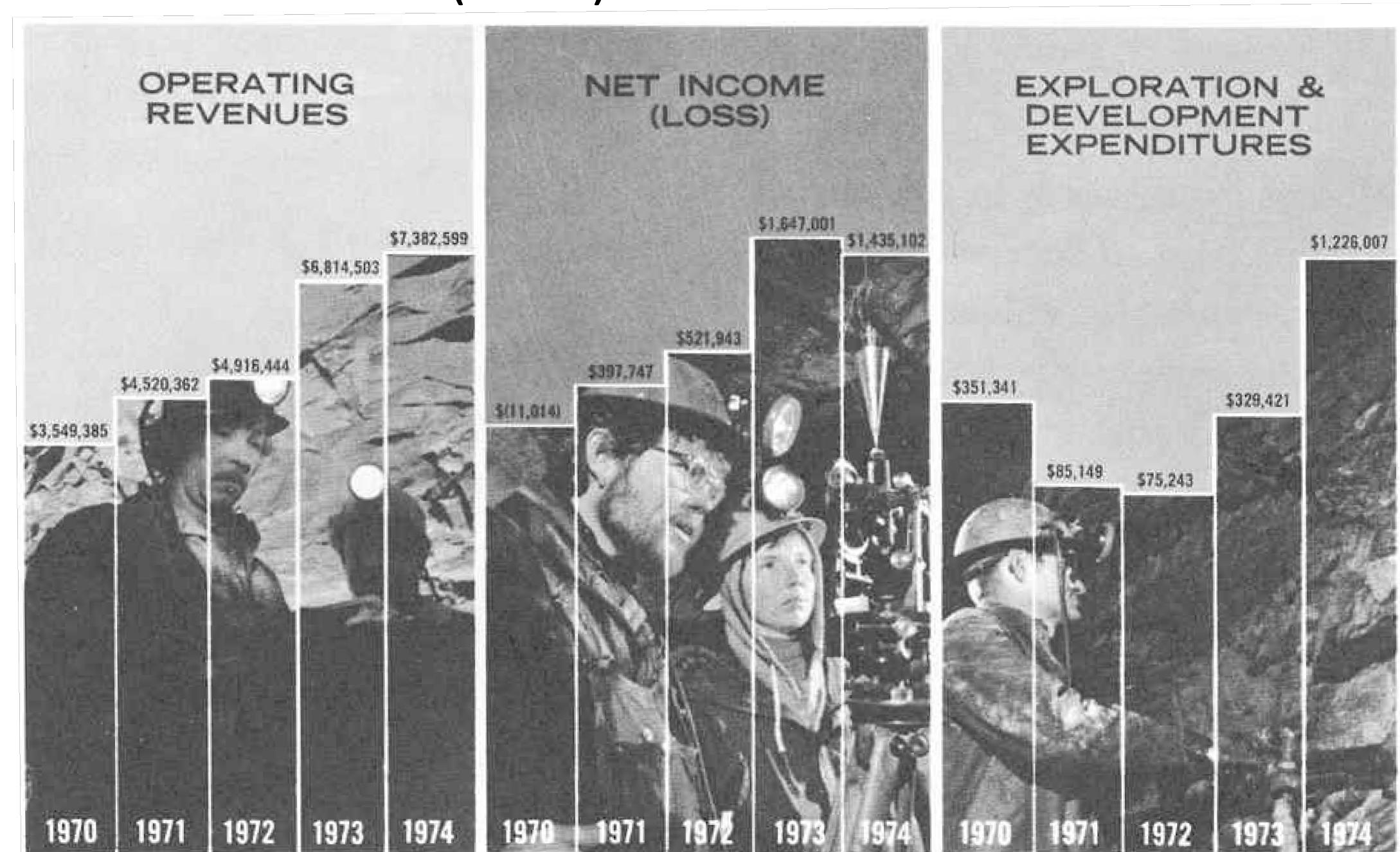
Colour & Accessibility

Colour can be used to encode information or for aesthetics/style/design. However, colour can also be distracting if used inappropriately or poorly.

Choose a perceptually uniform colour palette; can be sequential or diverging for quantitative data. Opt for colour-blind friendly palettes. Categorical data can use qualitative colour schemes.

Function first, form next

“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.” Tufte (1983)

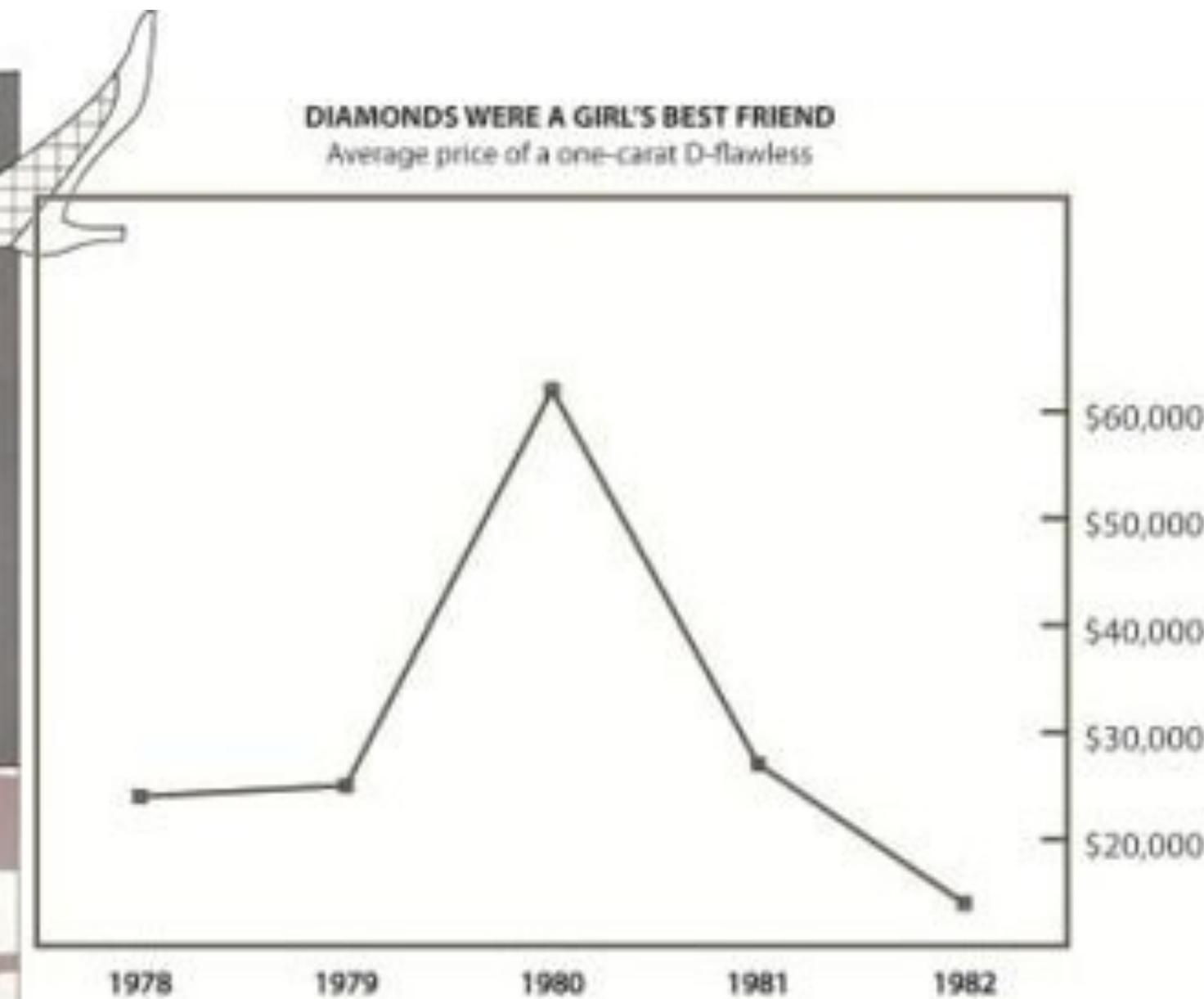


Which is better? Why?

A



B



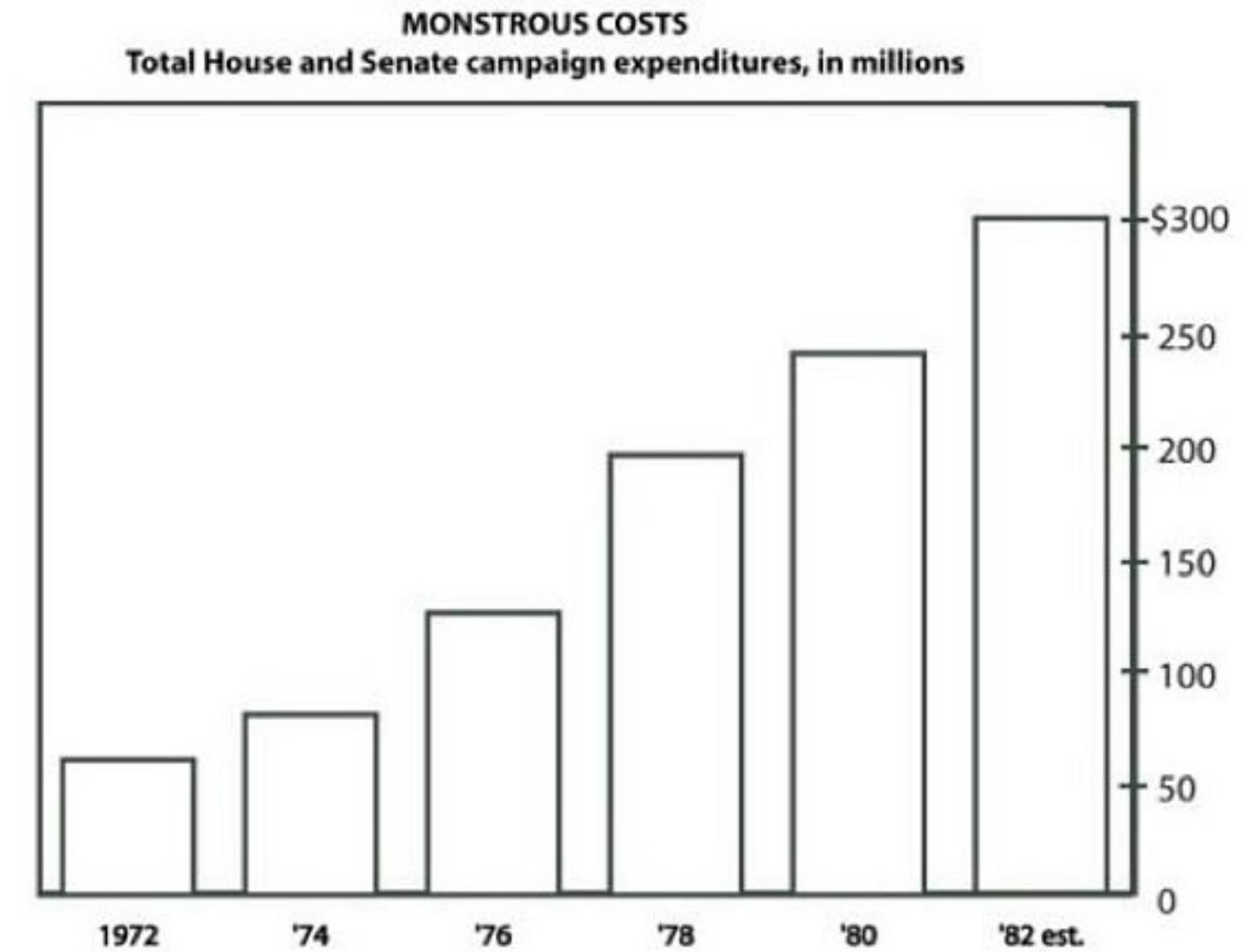
[Bateman et al. 2010]

Which is better? Why?

A



B



Use Embellished Charts? It Depends!

PROS

CONS

Use Embellished Charts? It Depends!

PROS

- persuasion
- memorability
- engagement

CONS

- biased analysis
- trustworthiness
- interpretability
- space efficiency
- effort

Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts

Scott Bateman, Regan L. Mandryk, Carl Gutwin,
Aaron Genest, David McDine, Christopher Brooks

Department of Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan, Canada
scott.bateman@usask.ca, regan@cs.usask.ca, gutwin@cs.usask.ca,
aaron.genest@usask.ca, dam085@mail.usask.ca, cab938@mail.usask.ca

ABSTRACT

Guidelines for designing information charts often state that the presentation should reduce ‘chart junk’ – visual embellishments that are not essential to understanding the data. In contrast, some popular chart designers wrap the presented data in detailed and elaborate imagery, raising the questions of whether this imagery is really as detrimental to understanding as has been proposed, and whether the visual embellishment may have other benefits. To investigate these issues, we conducted an experiment that compared embellished charts with plain ones, and measured both interpretation accuracy and long-term recall. We found that people’s accuracy in describing the embellished charts was no worse than for plain charts, and that their recall after a two-to-three-week gap was significantly better. Although we are cautious about recommending that all charts be produced in this style, our results question some of the premises of the minimalist approach to chart design.

Author Keywords

Charts, information visualization, imagery, memorability.

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI):
Miscellaneous.

General Terms

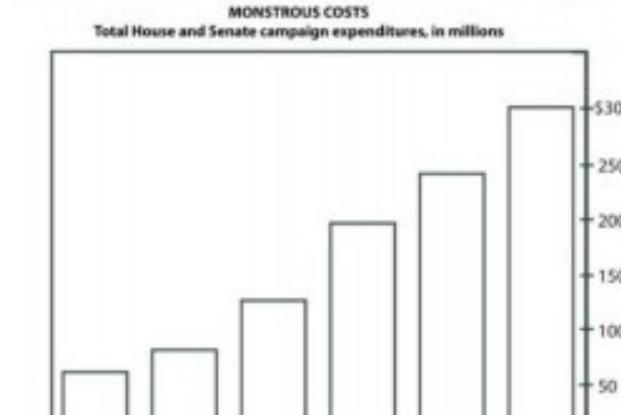
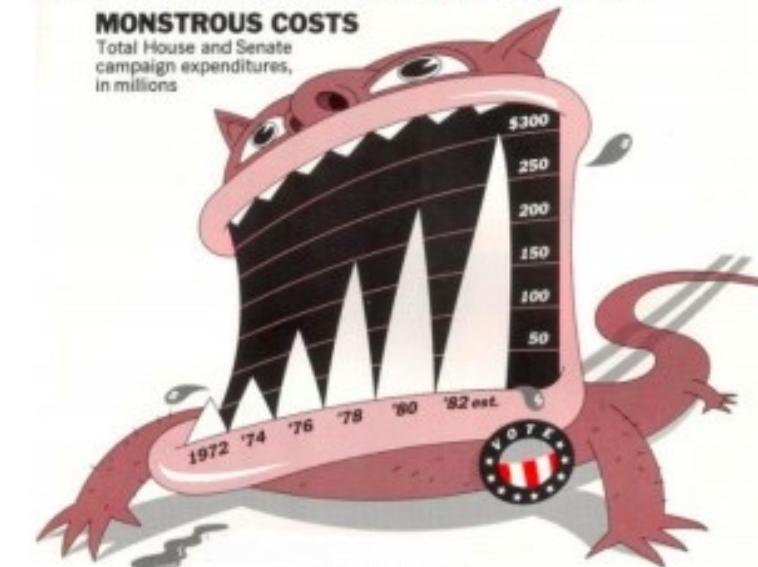
Design, Human Factors

INTRODUCTION

Many experts in the area of chart design, such as Edward Tufte, criticize the inclusion of visual embellishment in charts and graphs; their guidelines for good chart design often suggest that the addition of *chart junk*, decorations and other kinds of non-essential imagery, to a chart can make interpretation more difficult and can distract readers from the data [22]. This *minimalist* perspective advocates

data-ink – or the ink in the chart used to represent data.

Despite these minimalist guidelines, many designers include a wide variety of visual embellishments in their charts, from small decorations to large images and visual backgrounds. One well-known proponent of visual embellishment in charts is the graphic artist Nigel Holmes, whose work regularly incorporates strong visual imagery into the fabric of the chart [7] (e.g., Figure 1).

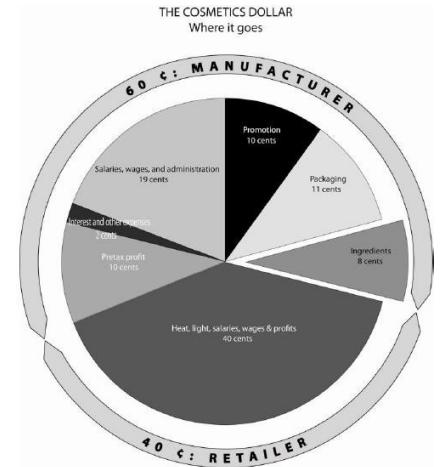
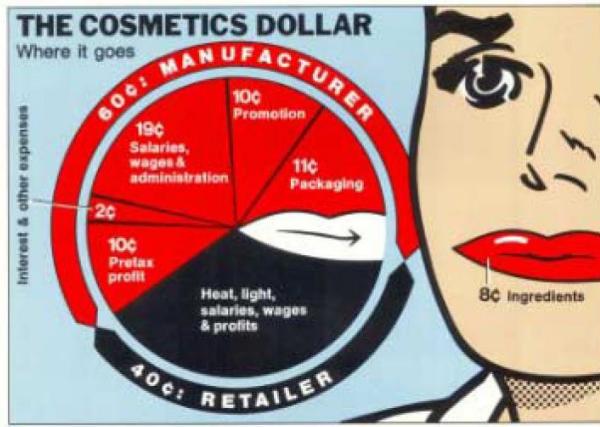


Experimental results

1. No difference for interpretation accuracy
2. No difference in recall accuracy after a five-minute gap
3. Significantly better recall for Holmes charts of both the chart topic and the details (categories and trend) after long-term gap (2-3 weeks).
4. Participants saw value messages in the Holmes charts significantly more often than in the plain charts.
5. Participants found the Holmes charts more attractive, most enjoyed them, and found that they were easiest and fastest to remember.

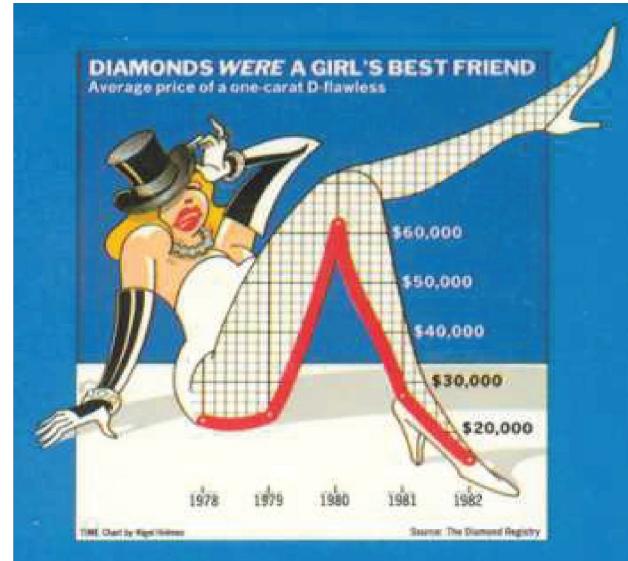
“Chart Embellishments Debate”

Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts



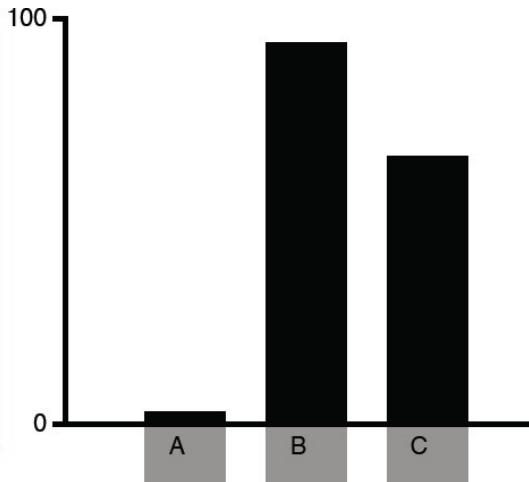
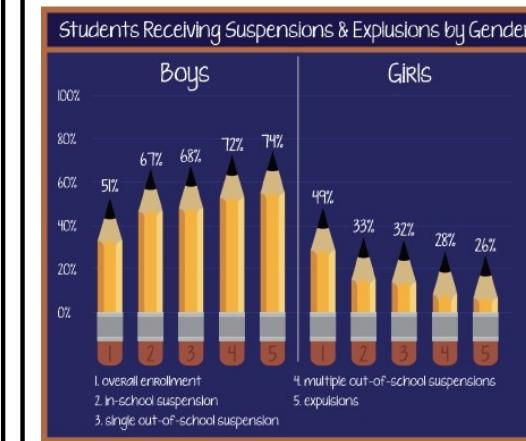
Bateman, et al. (2010)

Benefitting InfoVis with Visual Difficulties



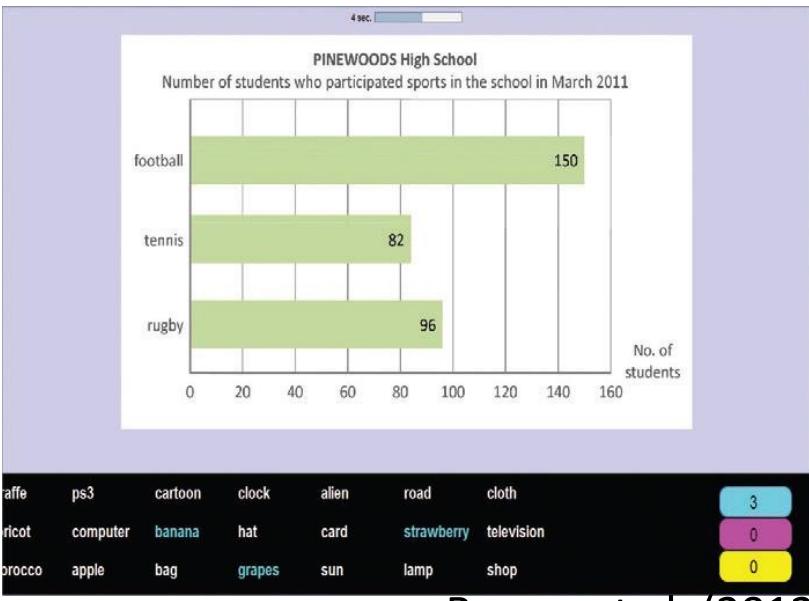
Hullman, et al. (2011)

An Evaluation of the Impact of Visual Embellishments in Bar Charts



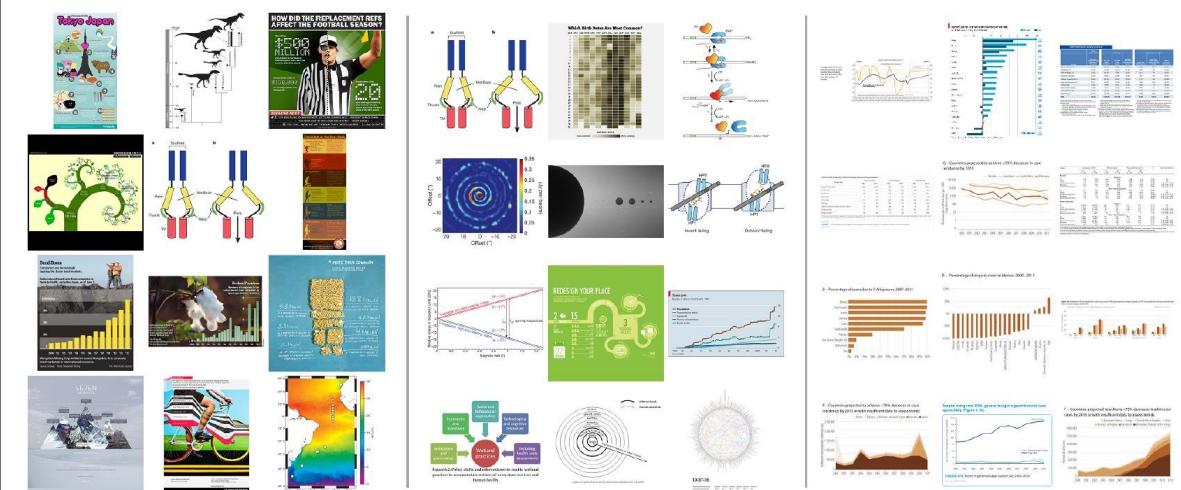
Skau, et al. (2015)

An Empirical Study on Using Visual Embellishments in Visualization



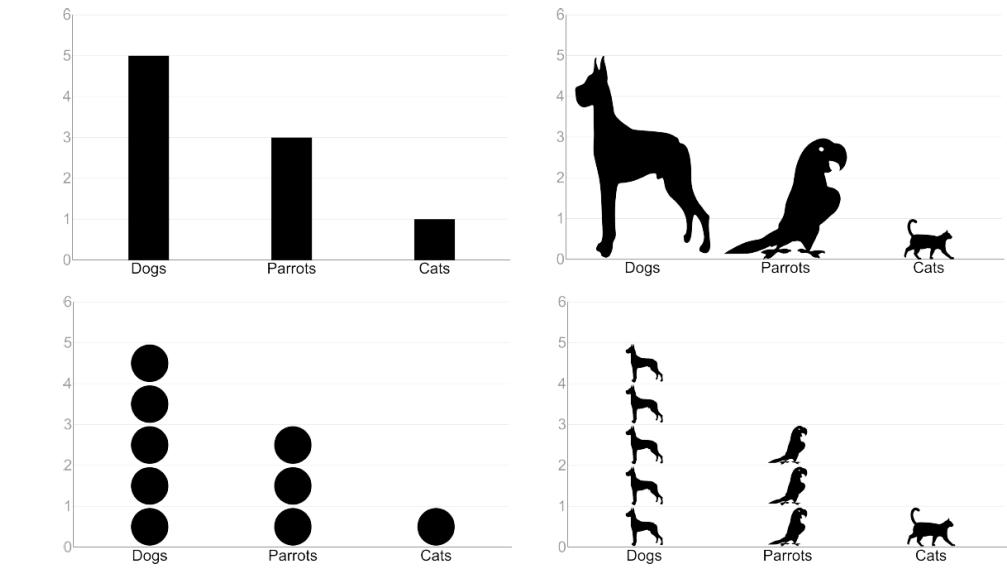
Borgo, et al. (2012)

What makes a visualization memorable?



Borkin, et al. (2013)
Borkin, et al. (2015)

ISOTYPE Visualization – Working Memory, Performance, and Engagement with Pictographs



Haroz, et al. (2015)

Function first, form next

- dangerous to start with aesthetics
 - usually impossible to add function retroactively
- start with focus on functionality
 - possible to improve aesthetics later on, as refinement
 - if no expertise in-house, find good graphic designer to work with
 - aesthetics do matter! another level of function
 - visual hierarchy, alignment, flow
 - Gestalt principles in action

Animation vs. multiple views

principle: external cognition vs. internal memory

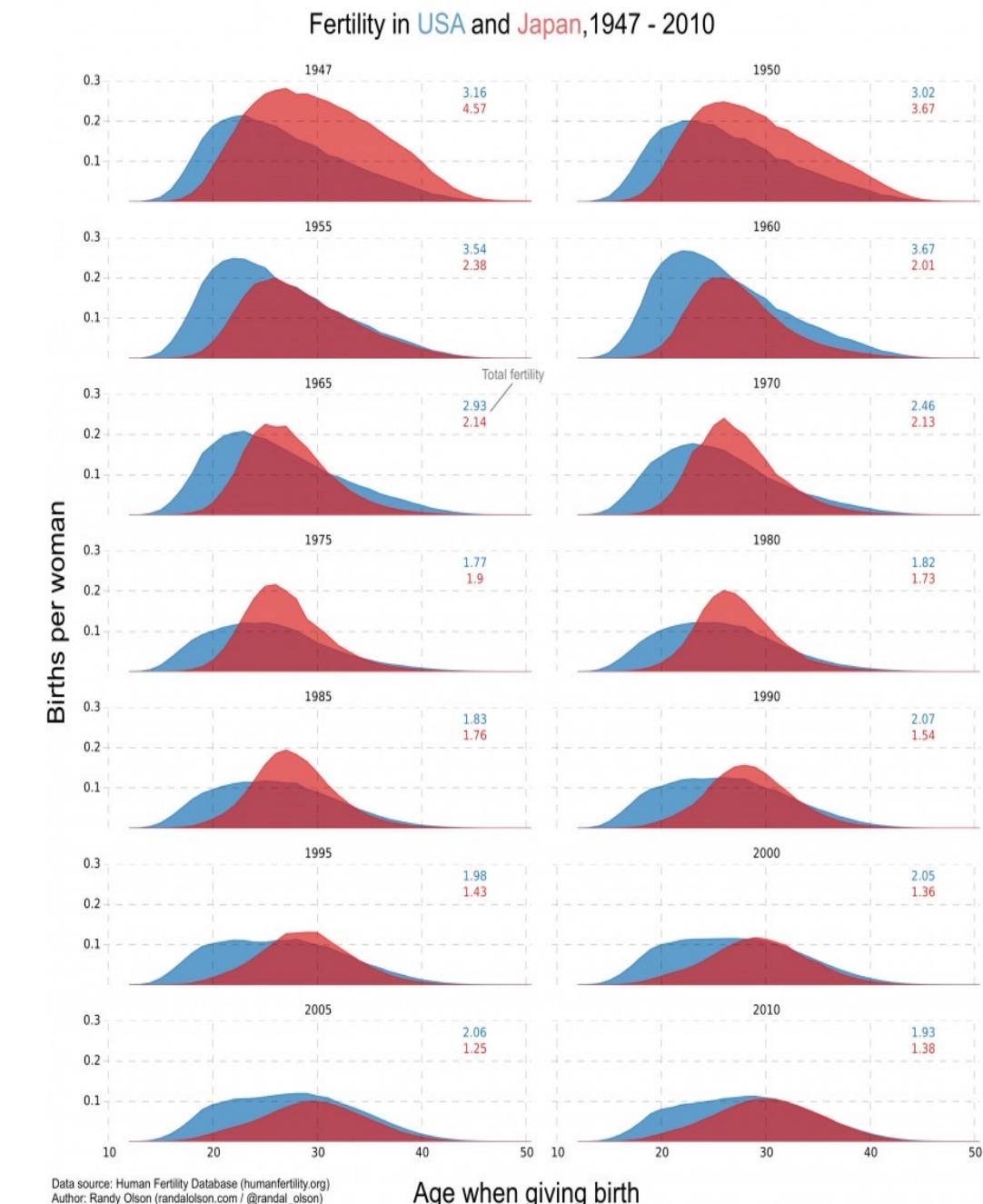
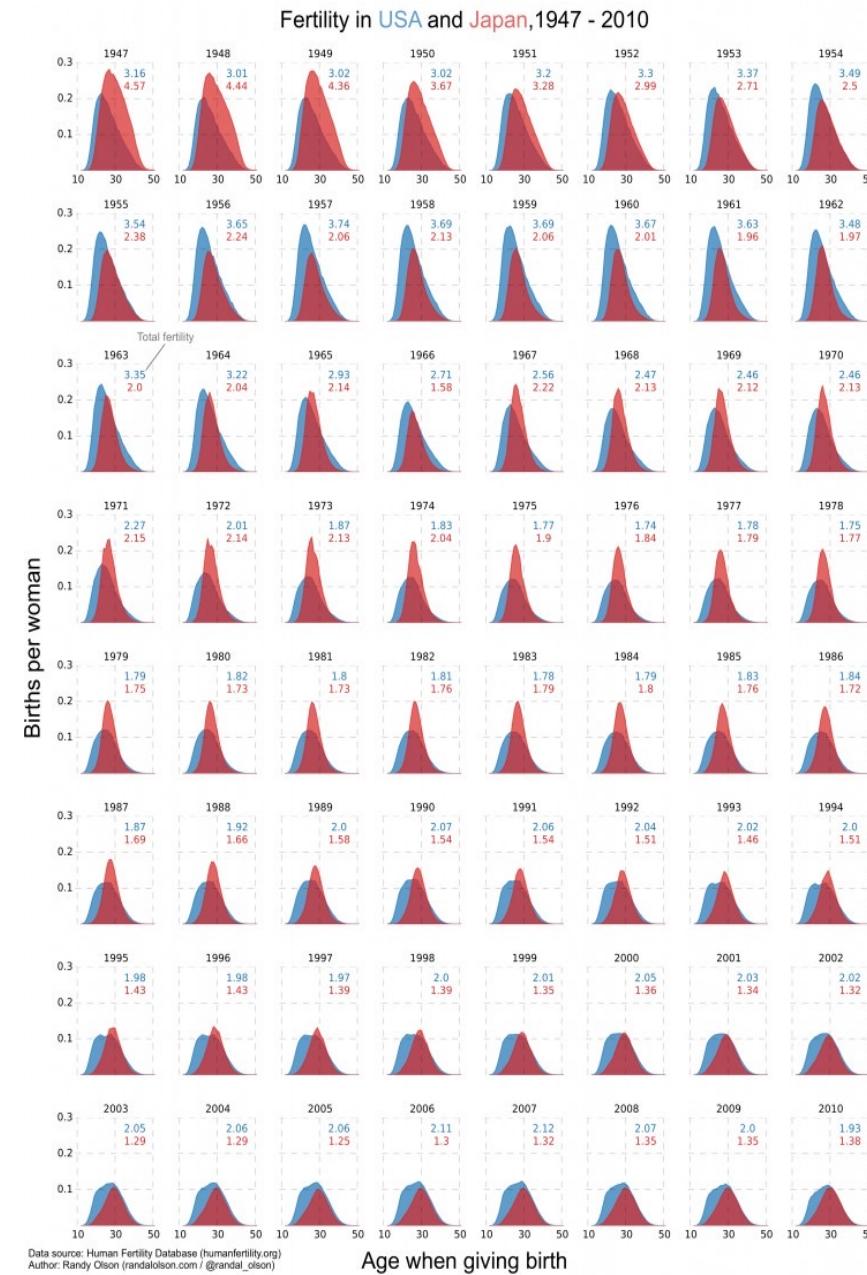
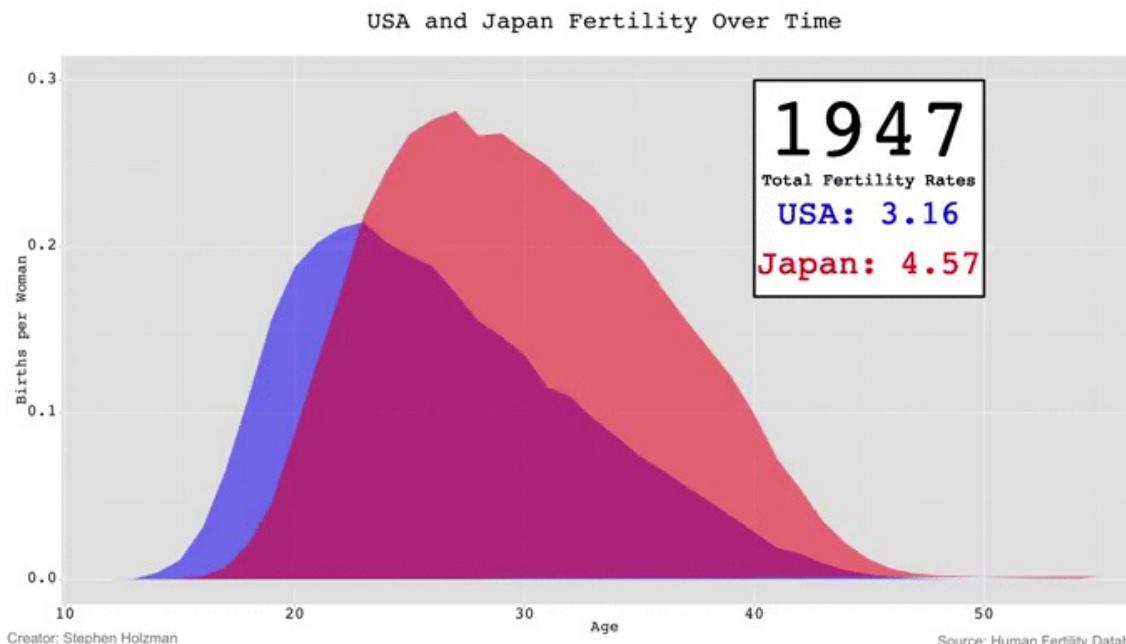
- easy to compare by moving eyes between side-by-side views
- harder to compare visible item to memory of what you saw



- Tufte advocates the use of small multiples – a series of similar graphs or charts using the same scale and axes, allowing for easy comparison across different categories or time periods. This technique is particularly effective in revealing patterns and changes in data over time.

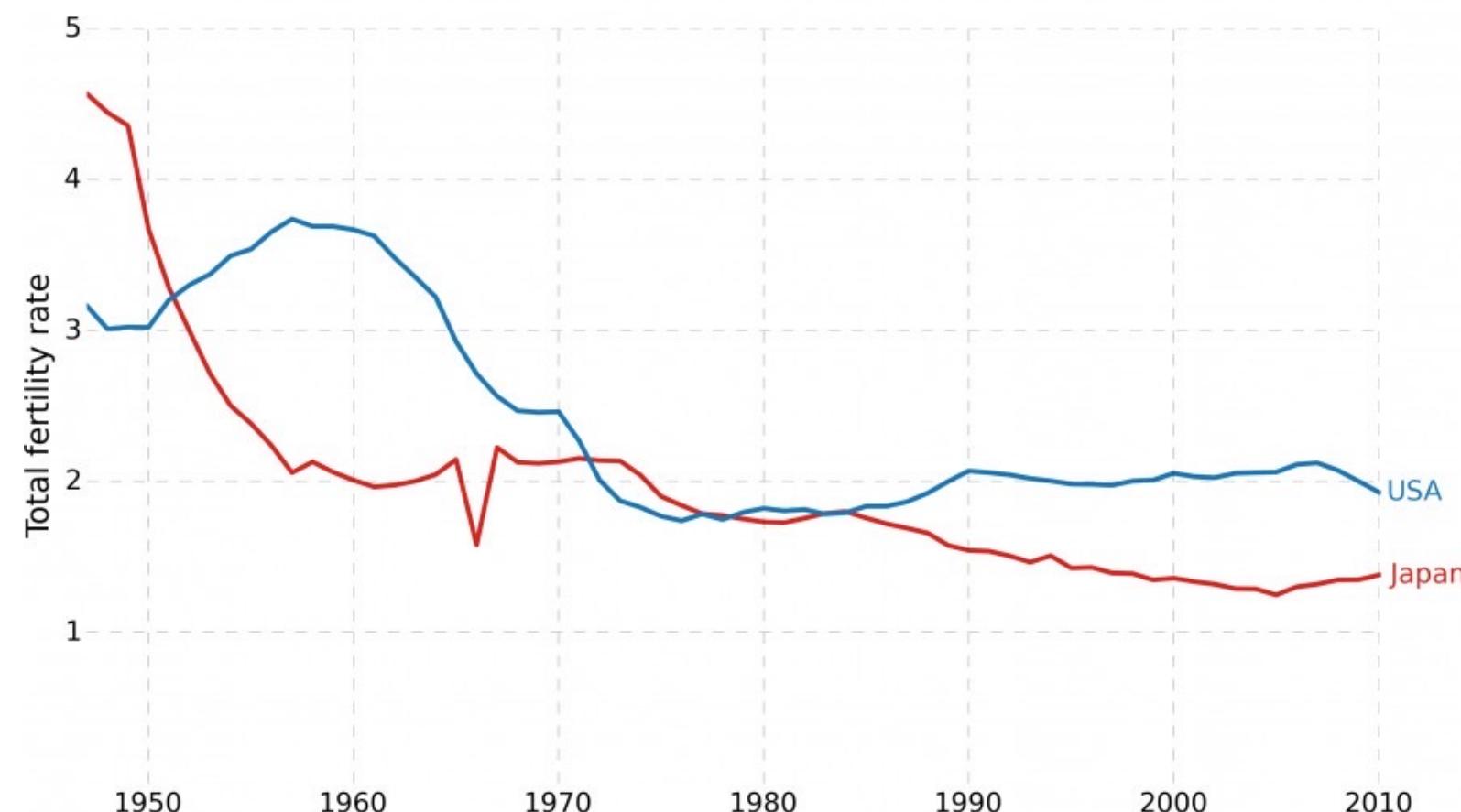
Animation vs. multiple views

DP: The task matters, if comparison is the goal, then don't make people memorize **Show them**



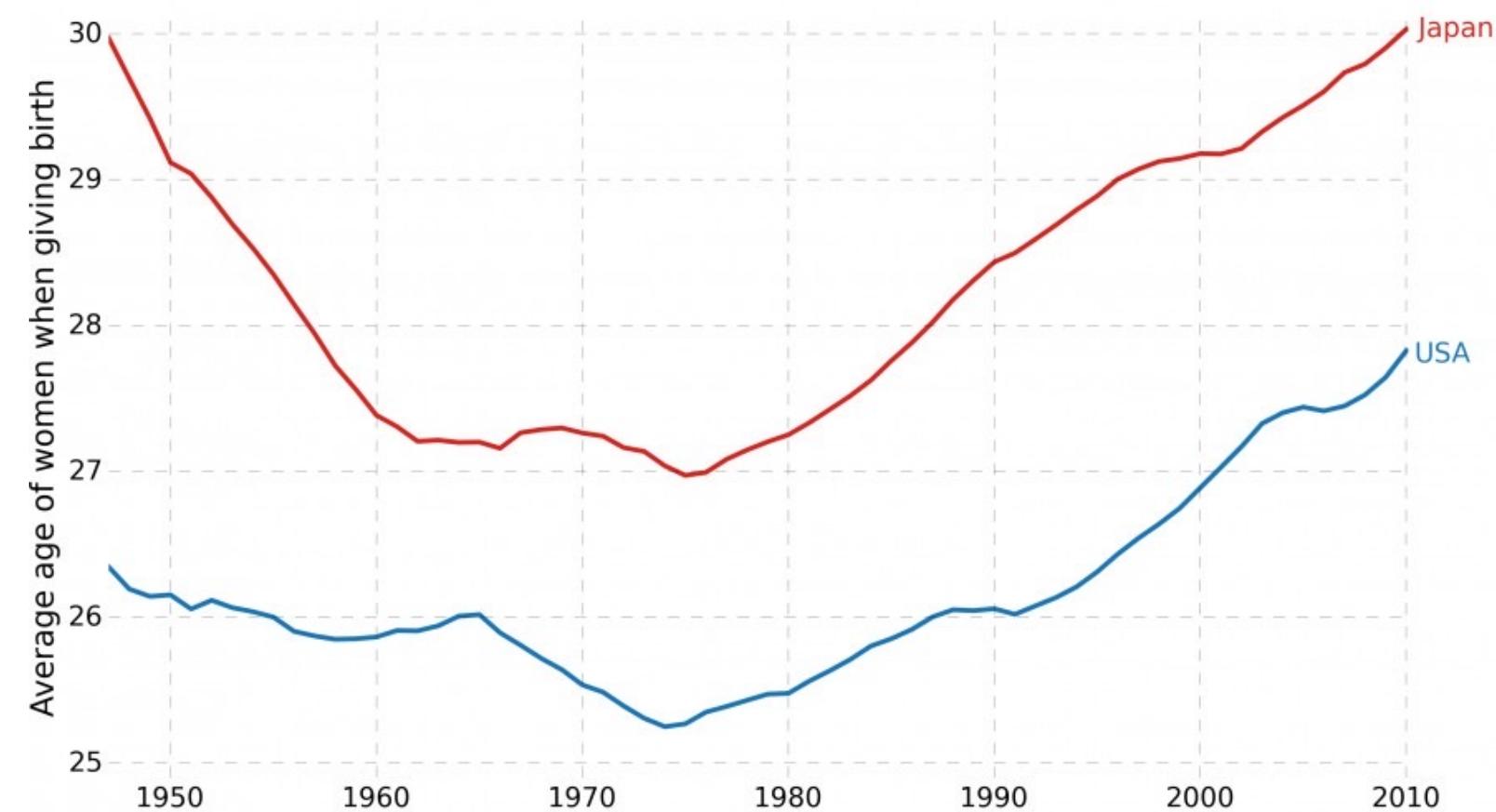
Simplify – Go back to the task and ask yourself what matters

Total fertility rate in USA and Japan, 1947 - 2010



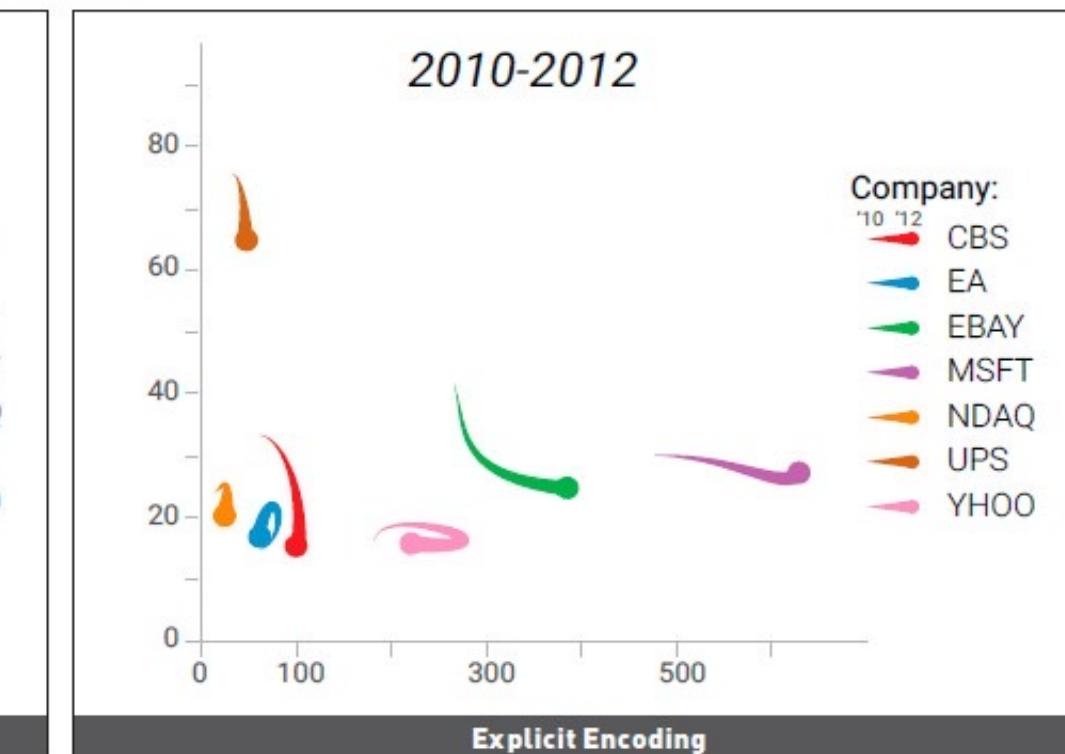
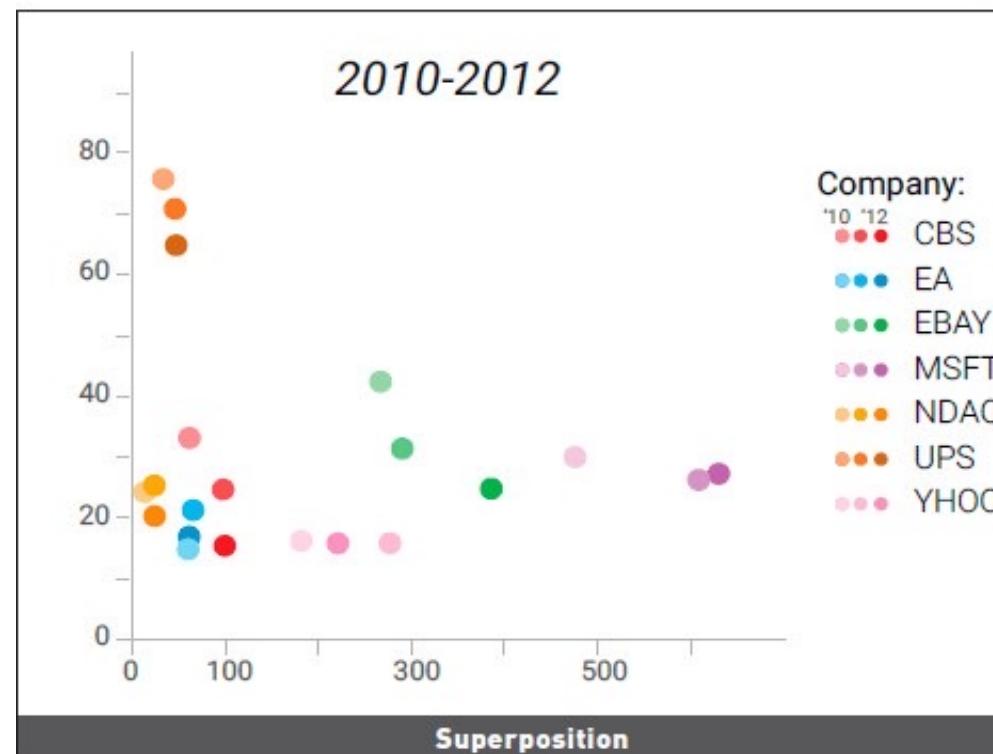
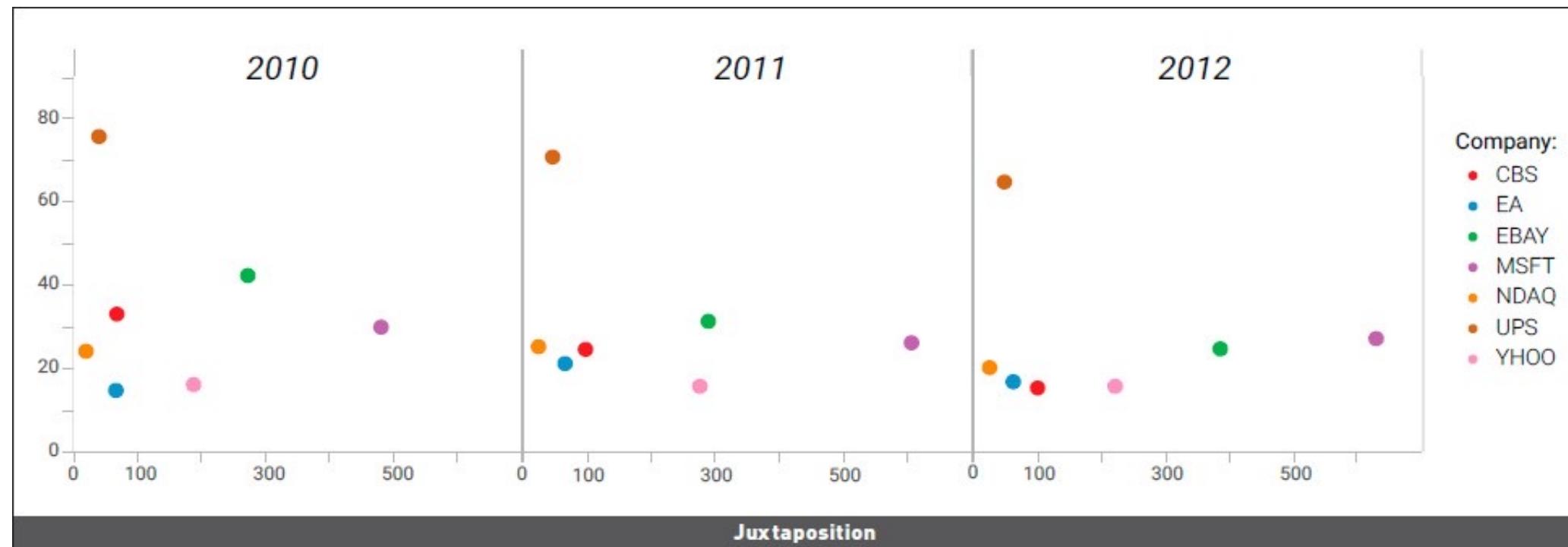
Data source: Human Fertility Database (humanfertility.org)
Author: Randy Olson (randalolson.com / [@randal_olson](https://twitter.com/randal_olson))

Average age when giving birth in USA and Japan, 1947 - 2010

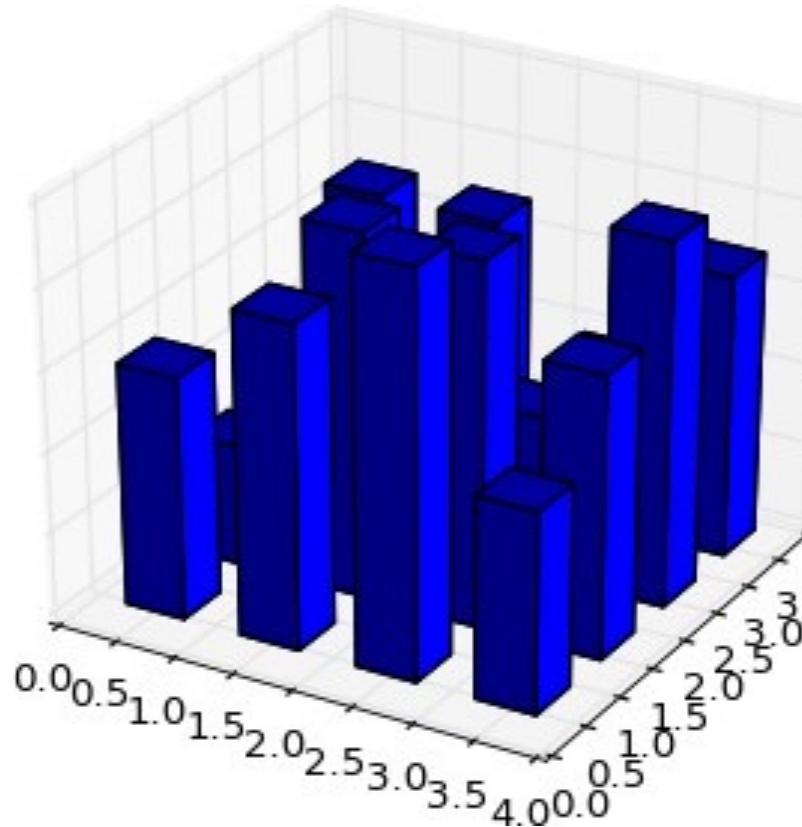


Data source: Human Fertility Database (humanfertility.org)
Author: Randy Olson (randalolson.com / [@randal_olson](https://twitter.com/randal_olson))

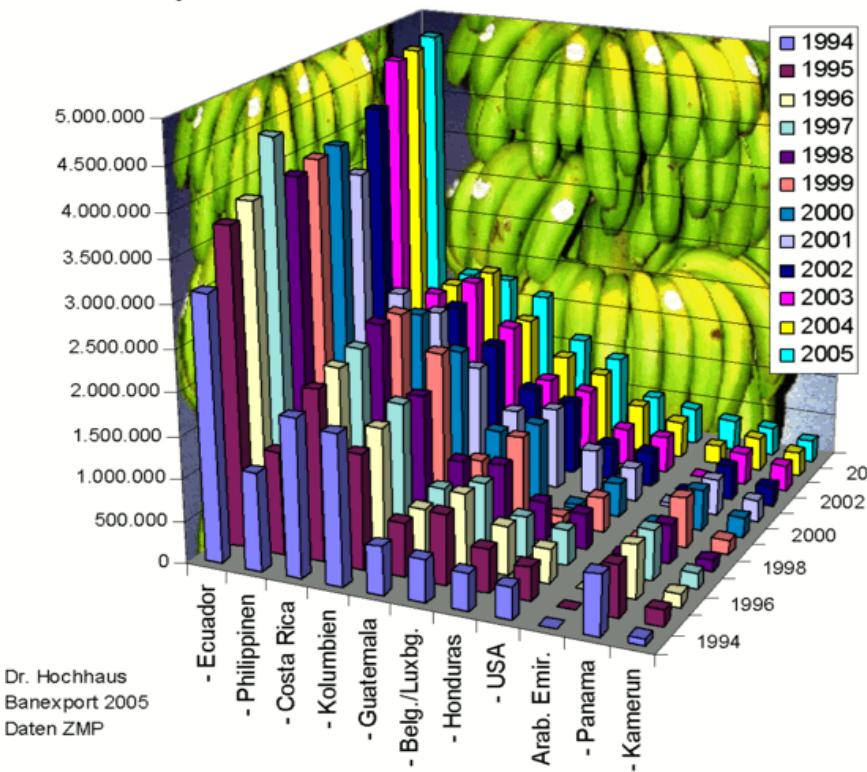
Small Multiple Design Alternatives



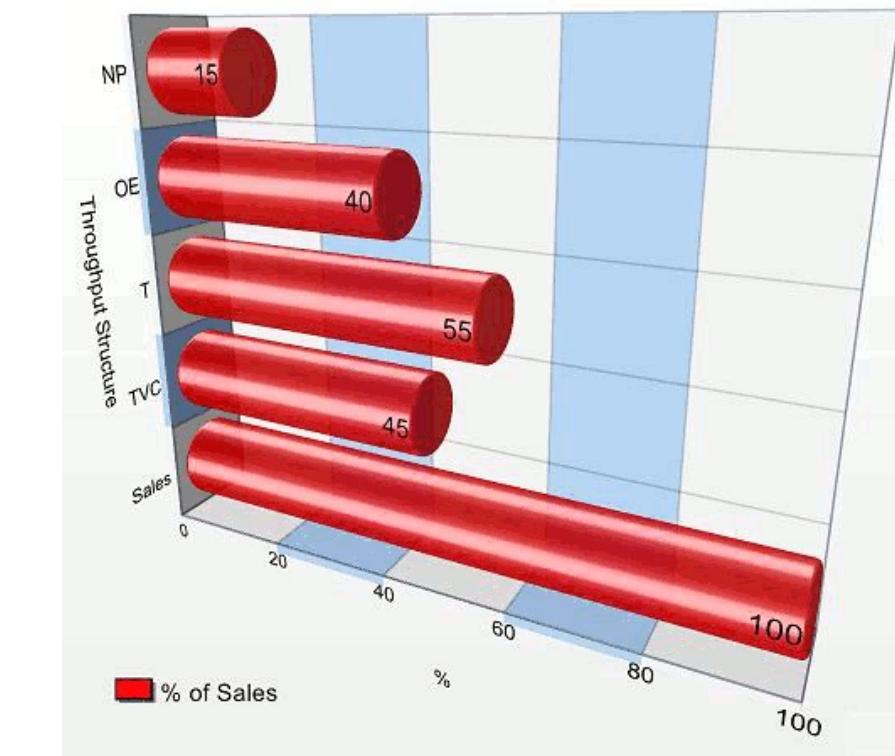
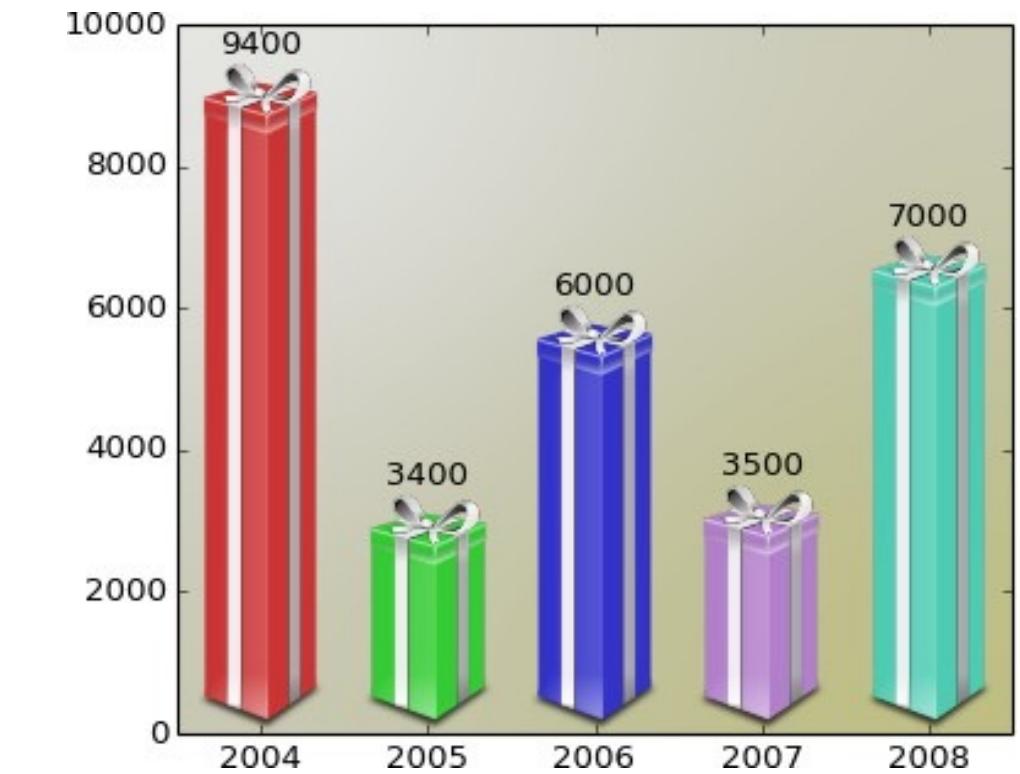
3D - Unjustified 3D all too common, in the news and elsewhere



Export von Bananen in Tonnen von 1994-2005



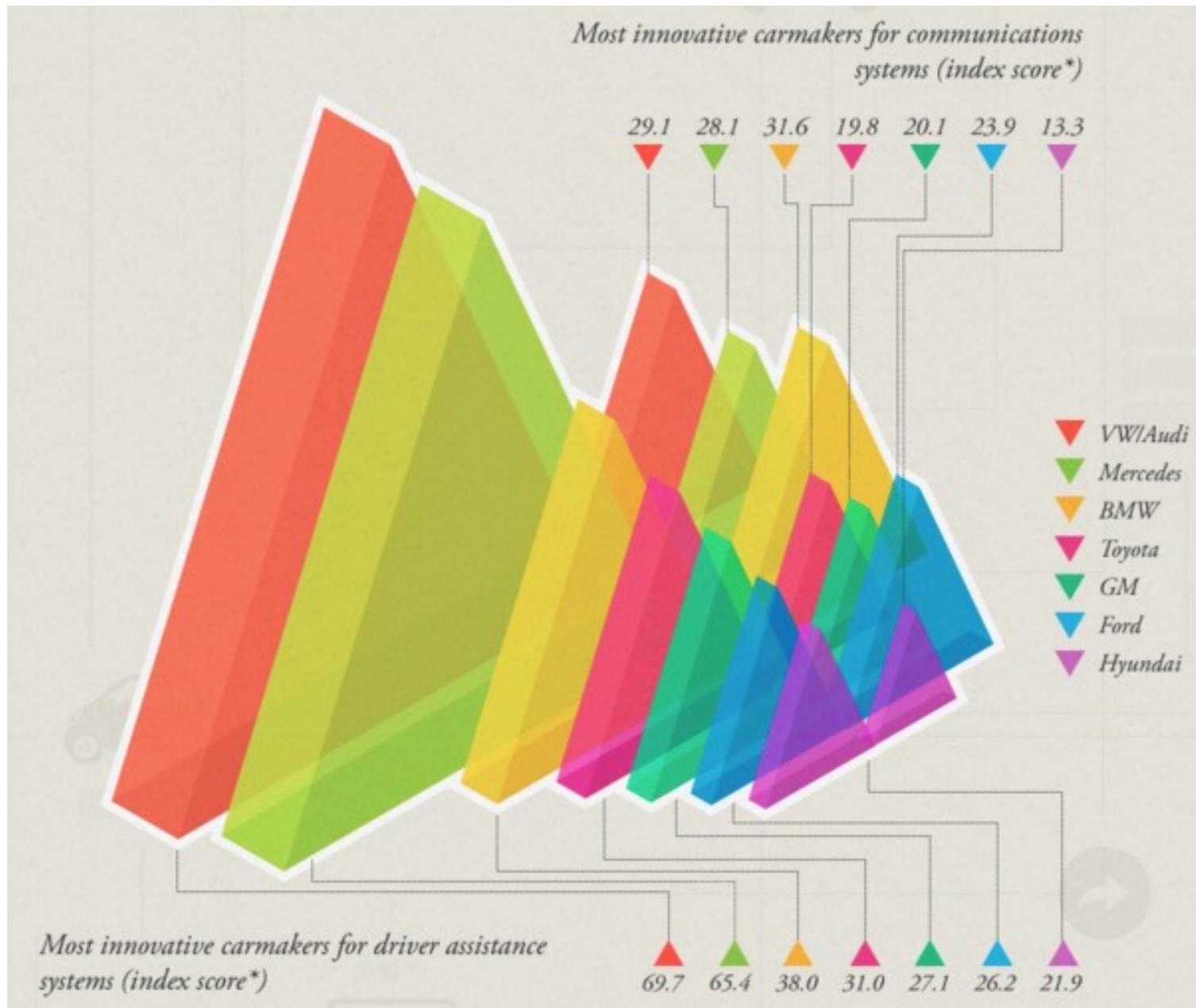
Dr. Hochhaus
Banlexport 2005
Daten ZMP



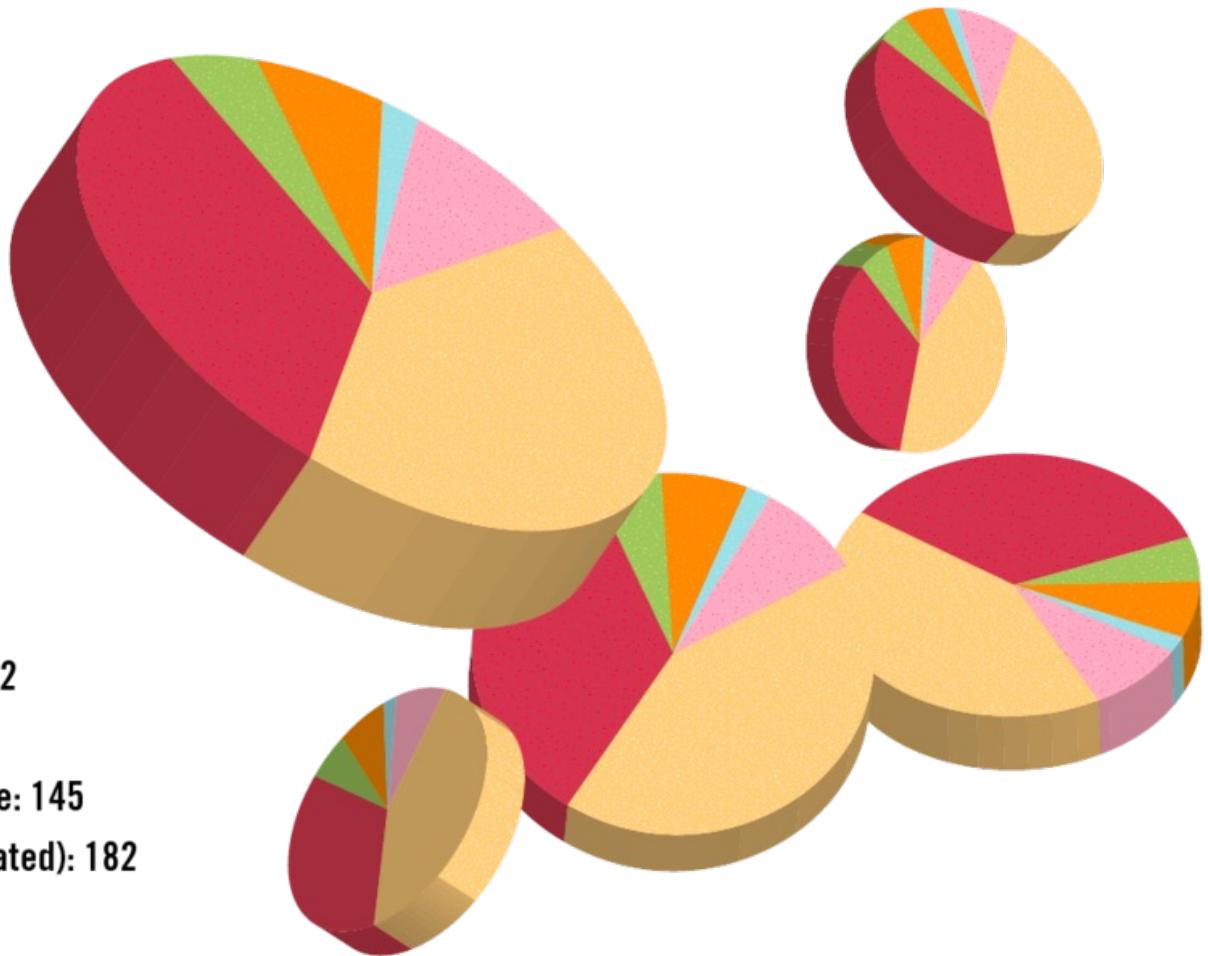
Excel Charts
Blog

matplotlib
gallery

3D - Unjustified 3D all too common, in the news and elsewhere



Convictions in London for class A drug supply.



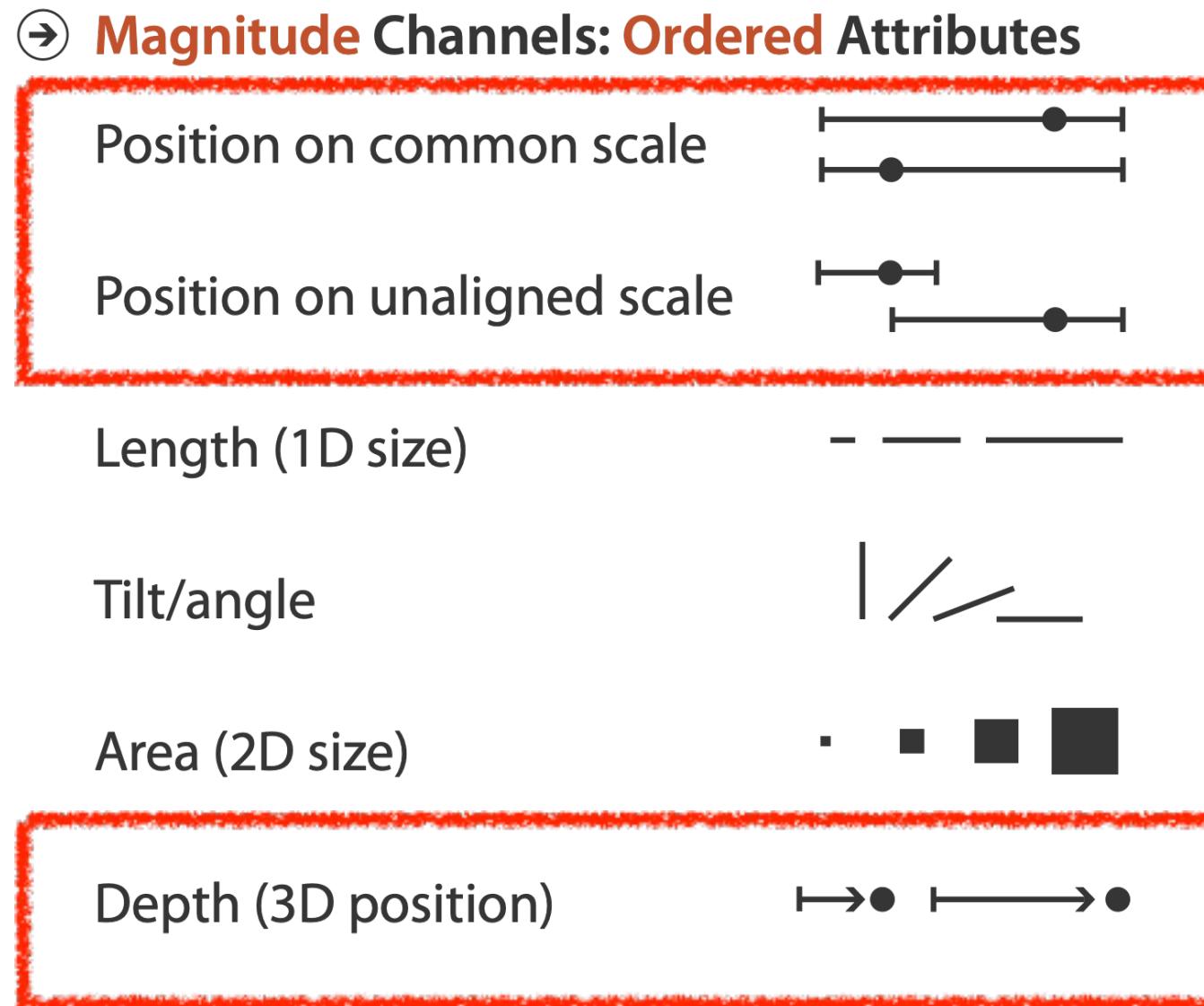
<http://viz.wtf/post/137826497077/eye-popping-3d-triangles>

<http://viz.wtf/post/139002022202/designer-drugs-ht-ducqn>

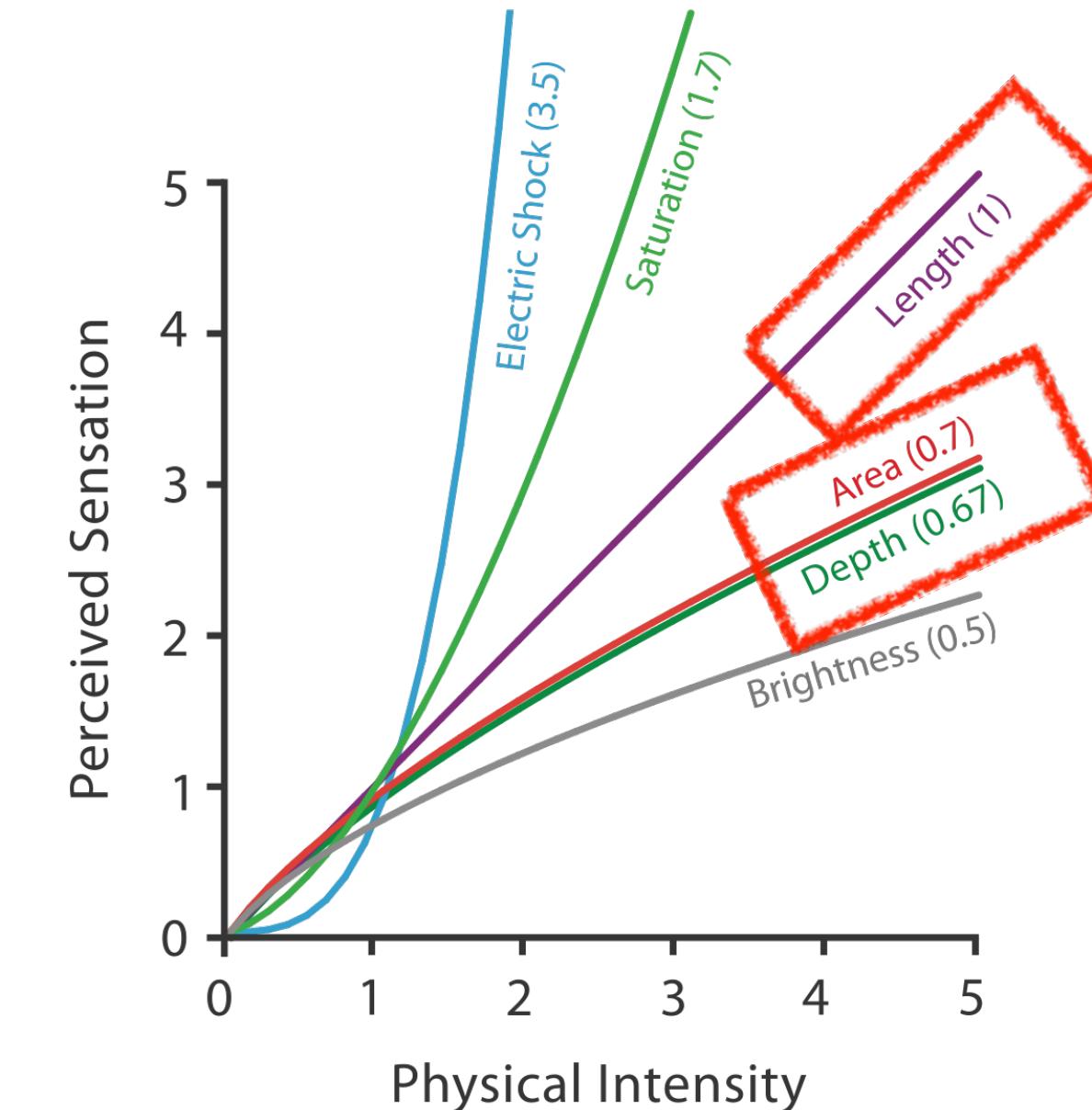
Reasons to Avoid 3D - Depth vs power of the plane

high-ranked spatial position channels: planar spatial position

– not depth!

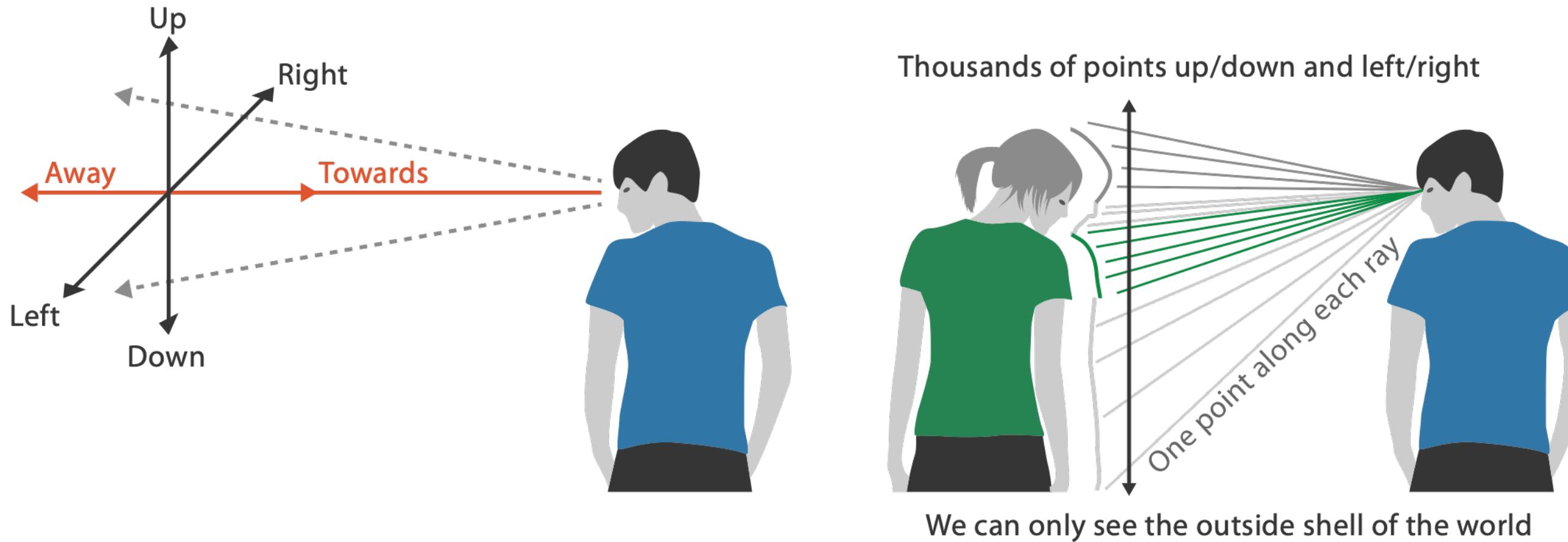


Steven's Psychophysical Power Law: $S = I^N$



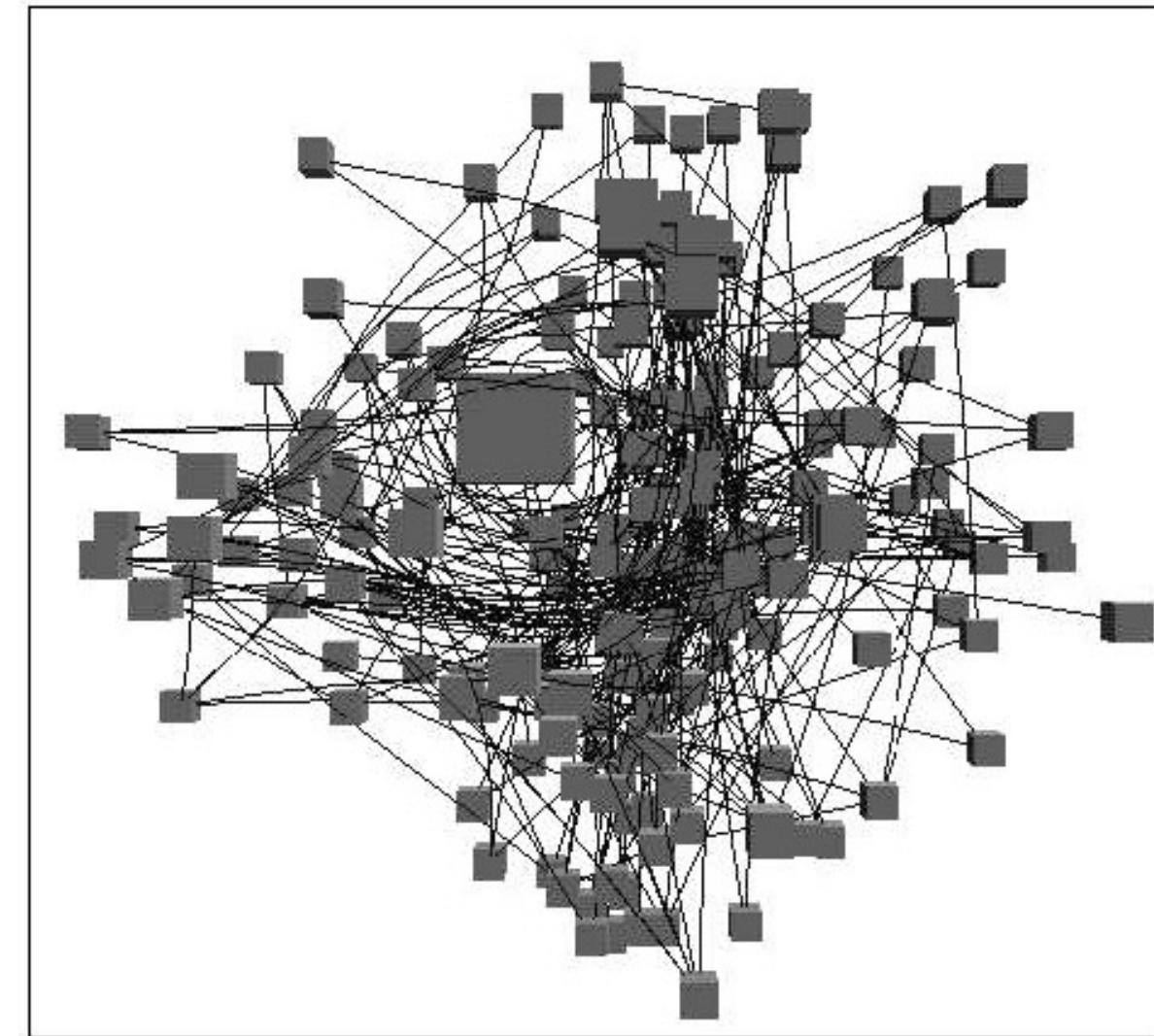
Reasons to Avoid 3D - Danger of depth

- we don't really live in 3D: we *see* in 2.05D
 - acquire more info on image plane quickly from eye movements
 - acquire more info for depth slower, from head/body motion



Reasons to Avoid 3D - Occlusion hides information

- occlusion
- interaction can resolve, but at cost of time and cognitive load

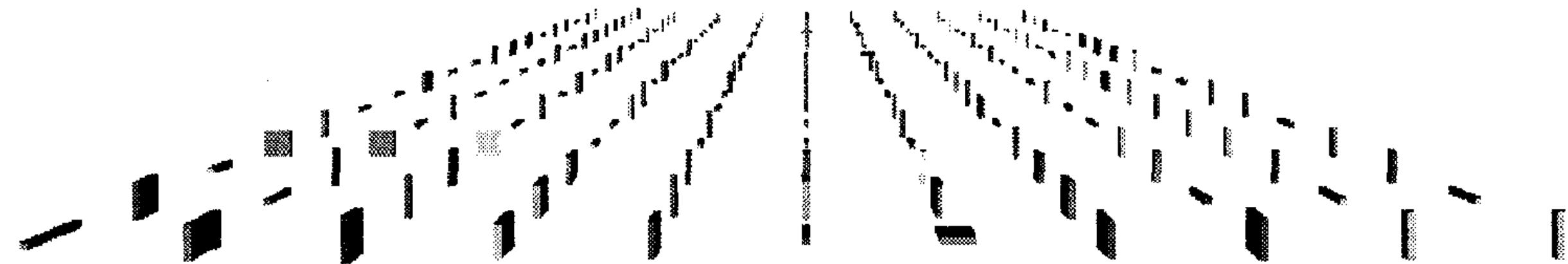


[*Distortion Viewing Techniques for 3D Data. Carpendale et al. InfoVis1996.*]

Reasons to Avoid 3D - Perspective distortion loses information

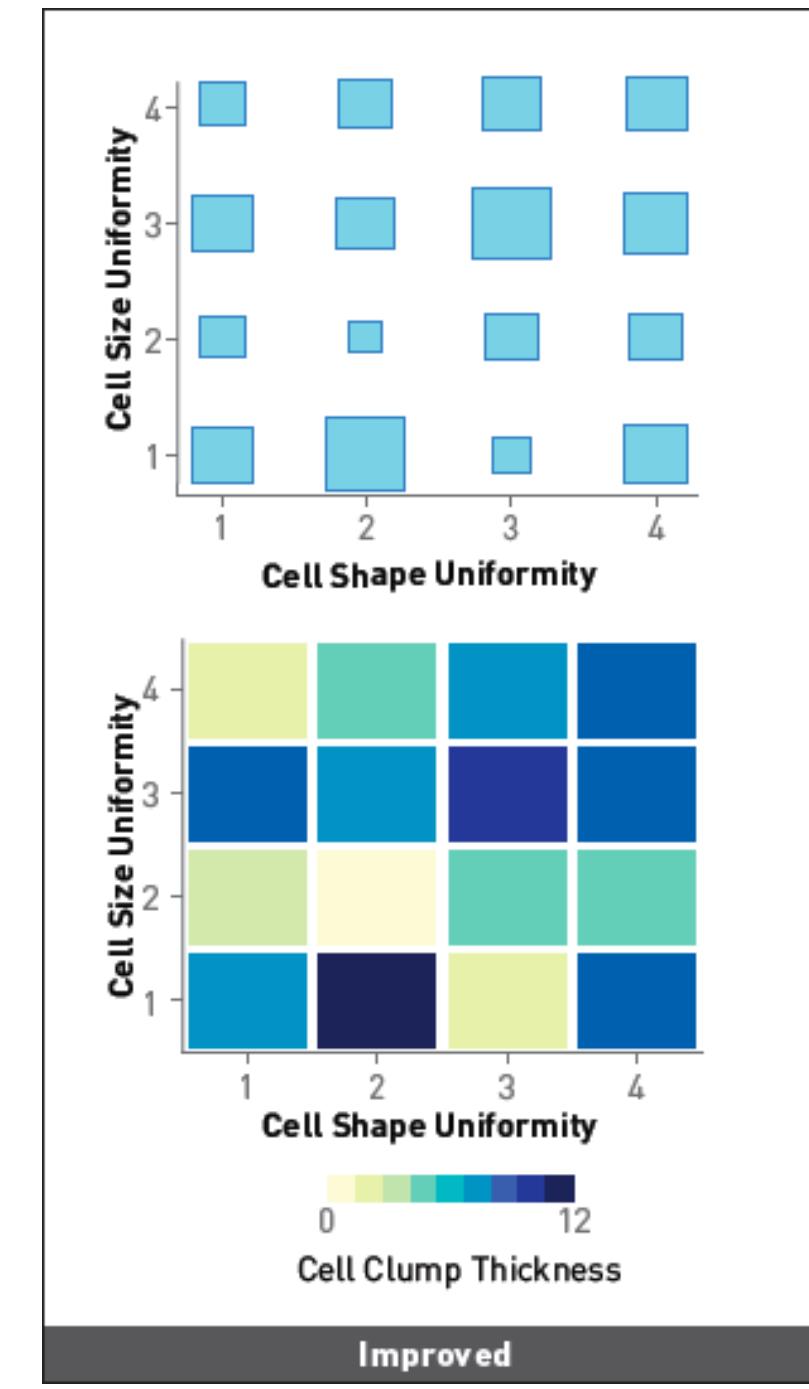
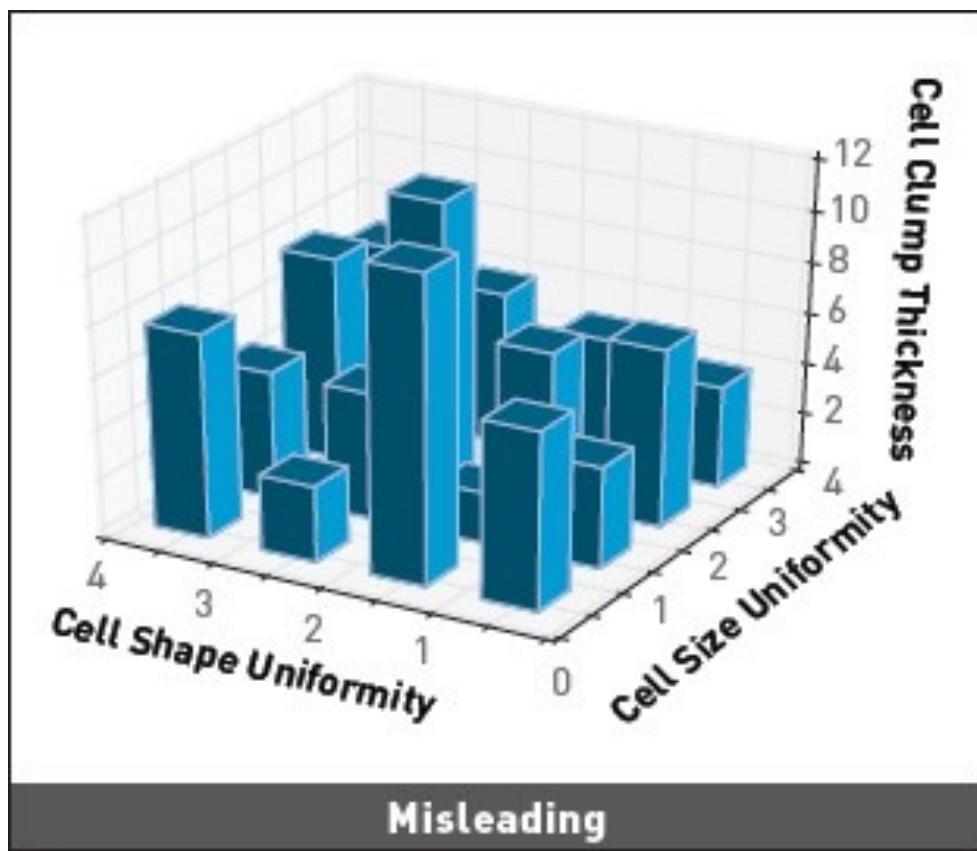
perspective distortion

- interferes with all size channel encodings
- power of the plane is lost!

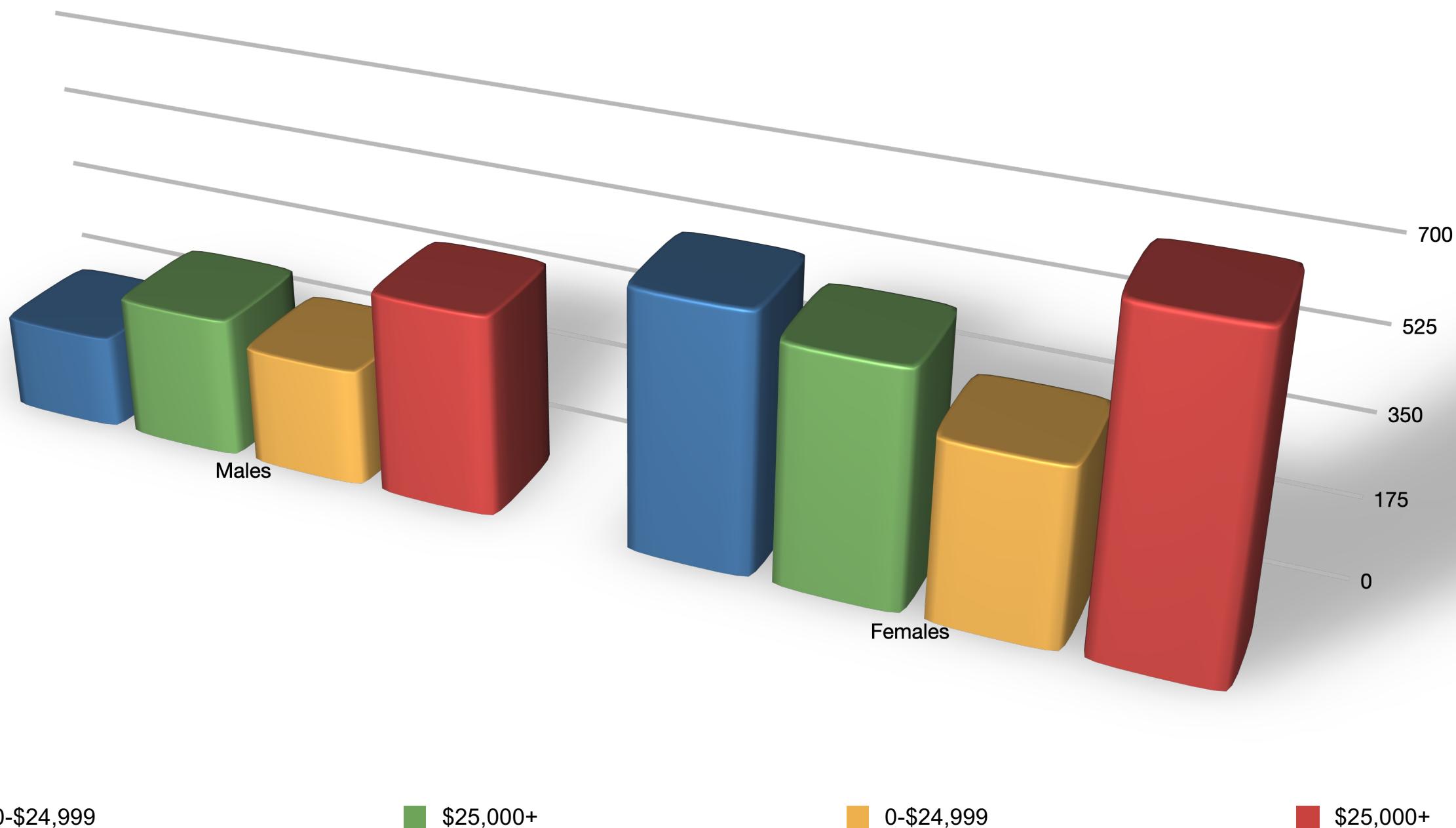


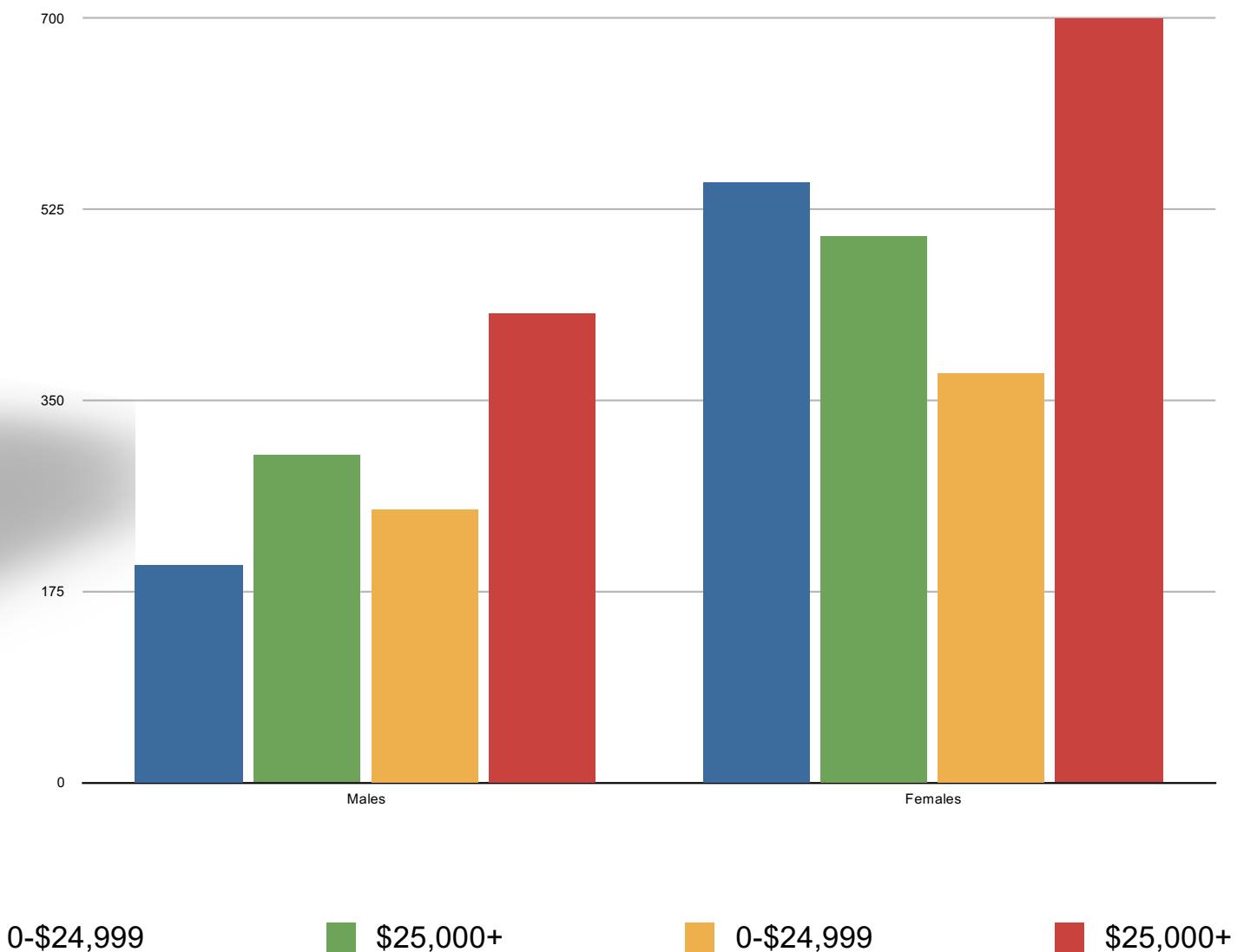
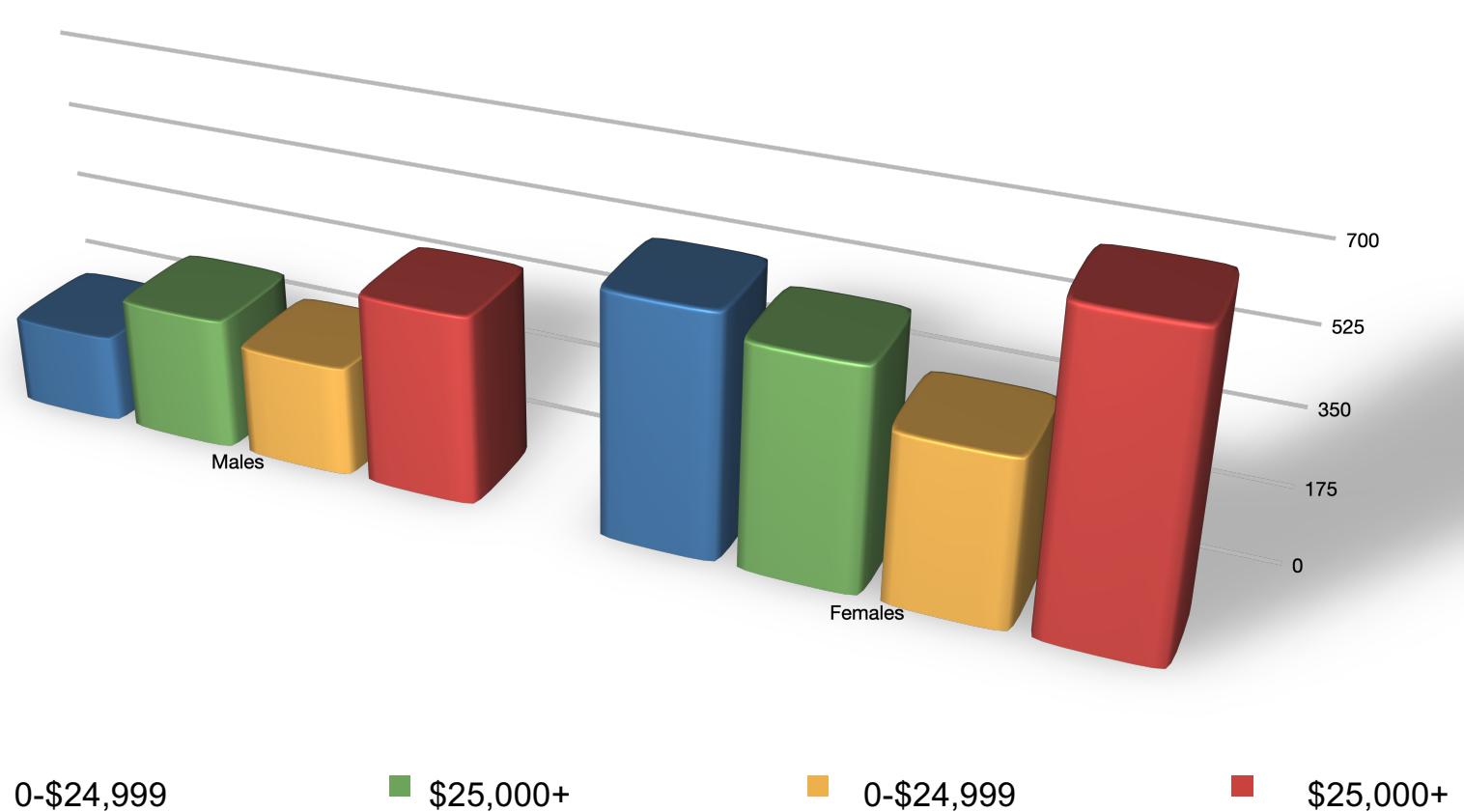
[Visualizing the Results of Multimedia Web Search Engines. Mukherjea, Hirata, and Hara. InfoVis 96]

3D Design Alternatives



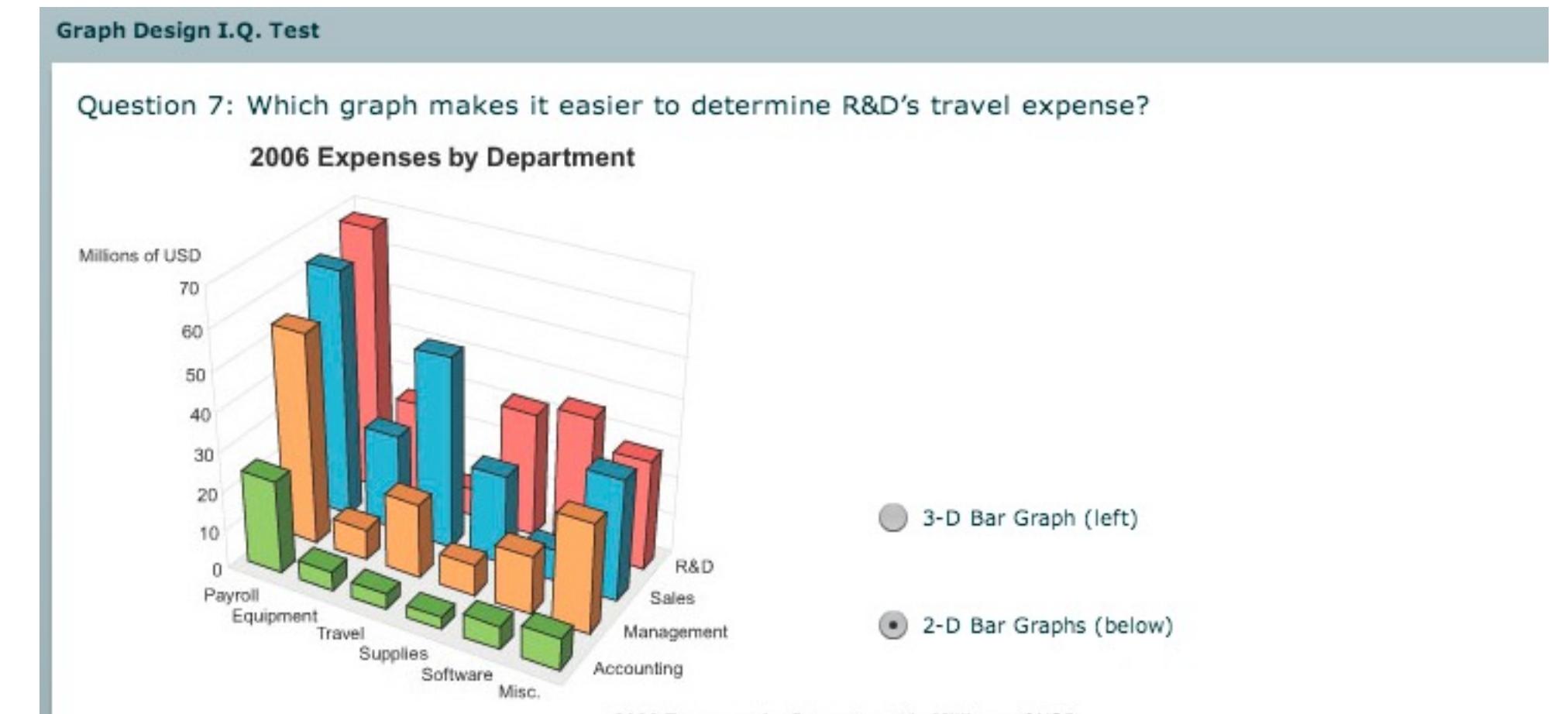
3D Depth Perception issues





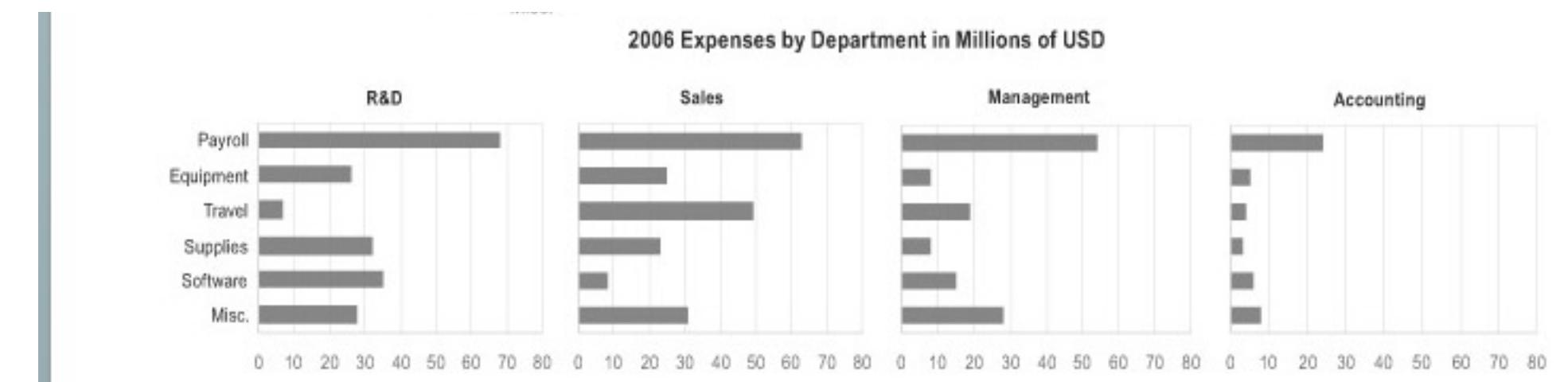
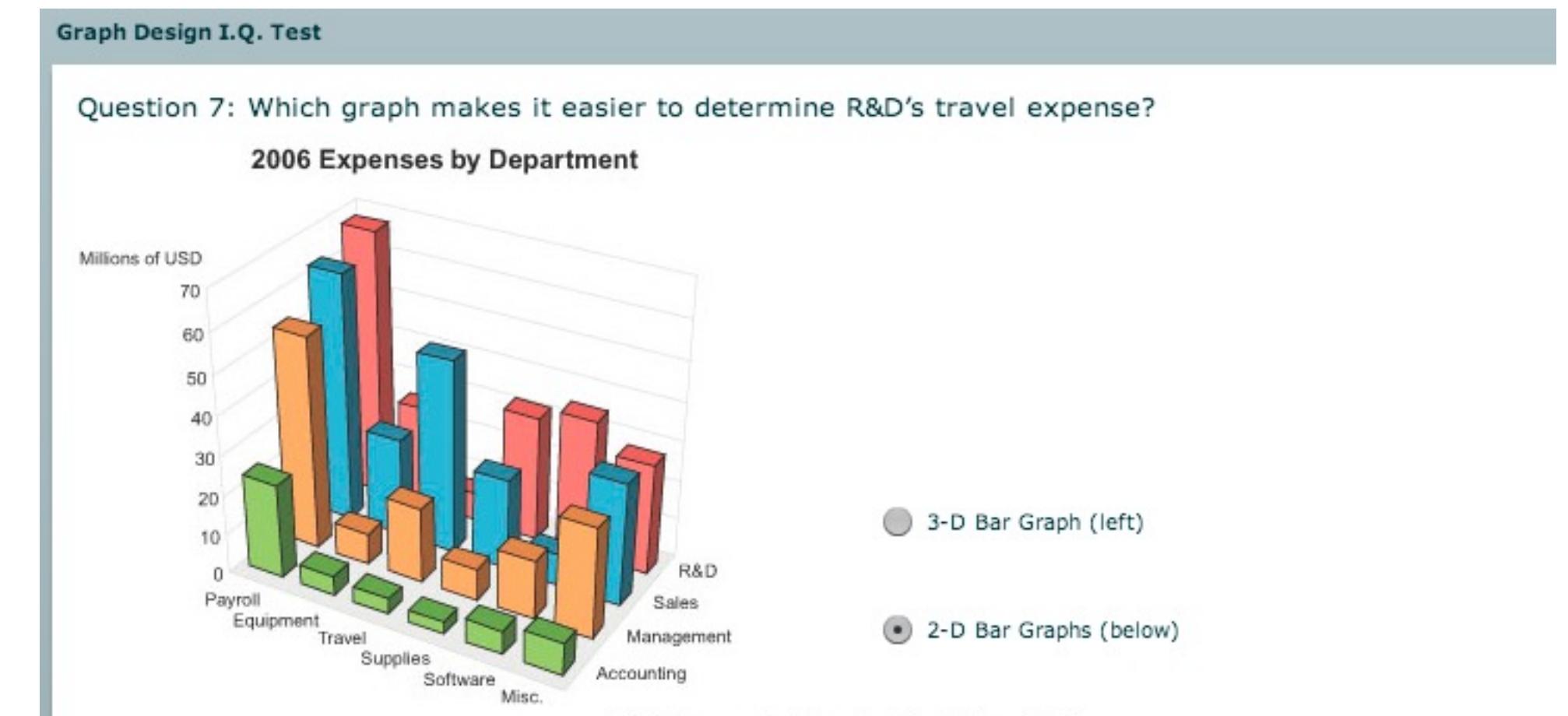
3D vs 2D bar charts

- 3D bars:
very difficult to justify!
 - perspective distortion
 - occlusion
- faceting into 2D almost
always better choice



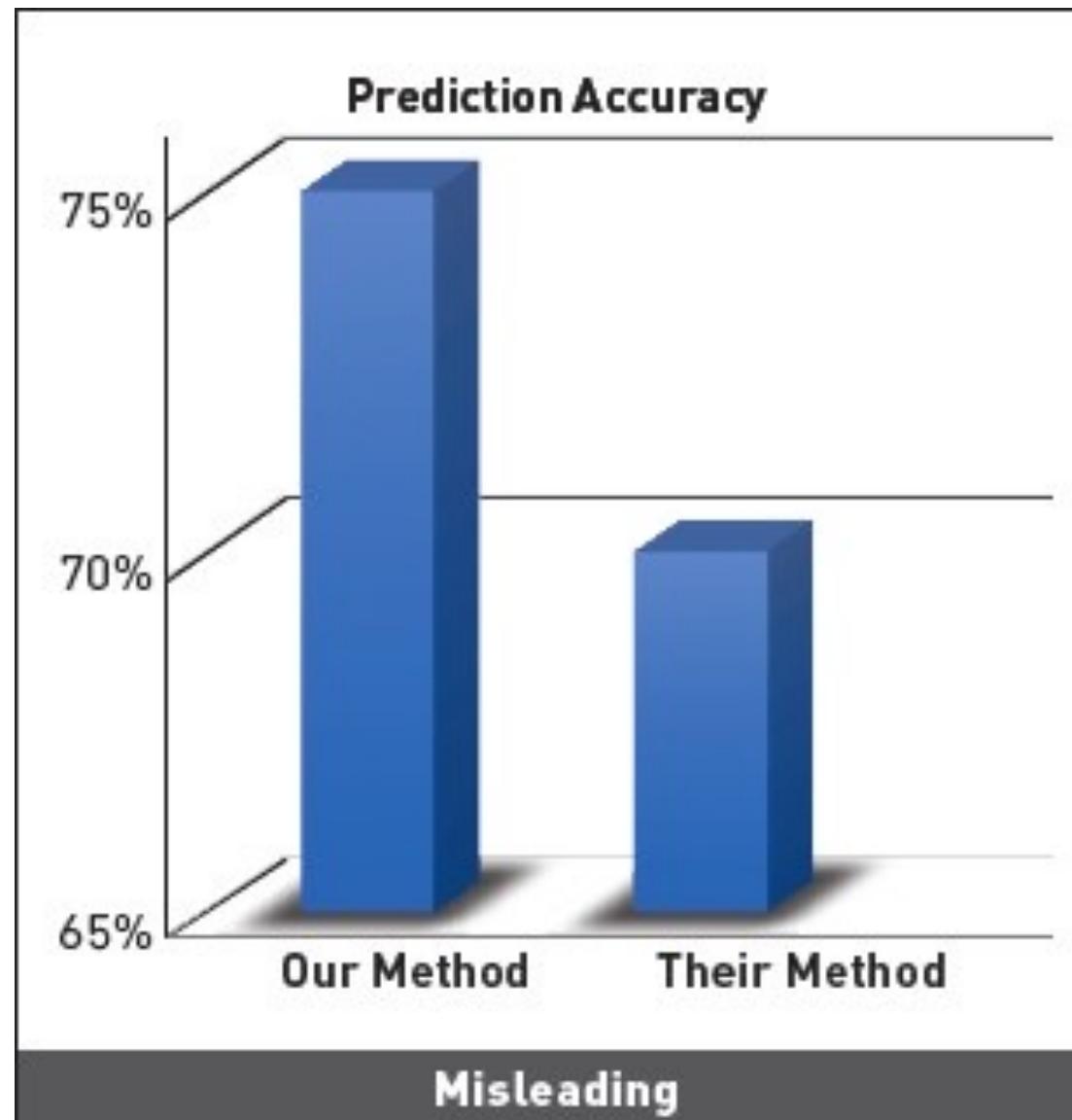
3D vs 2D bar charts

- 3D bars:
very difficult to justify!
 - perspective distortion
 - occlusion
- faceting into 2D almost
always better choice

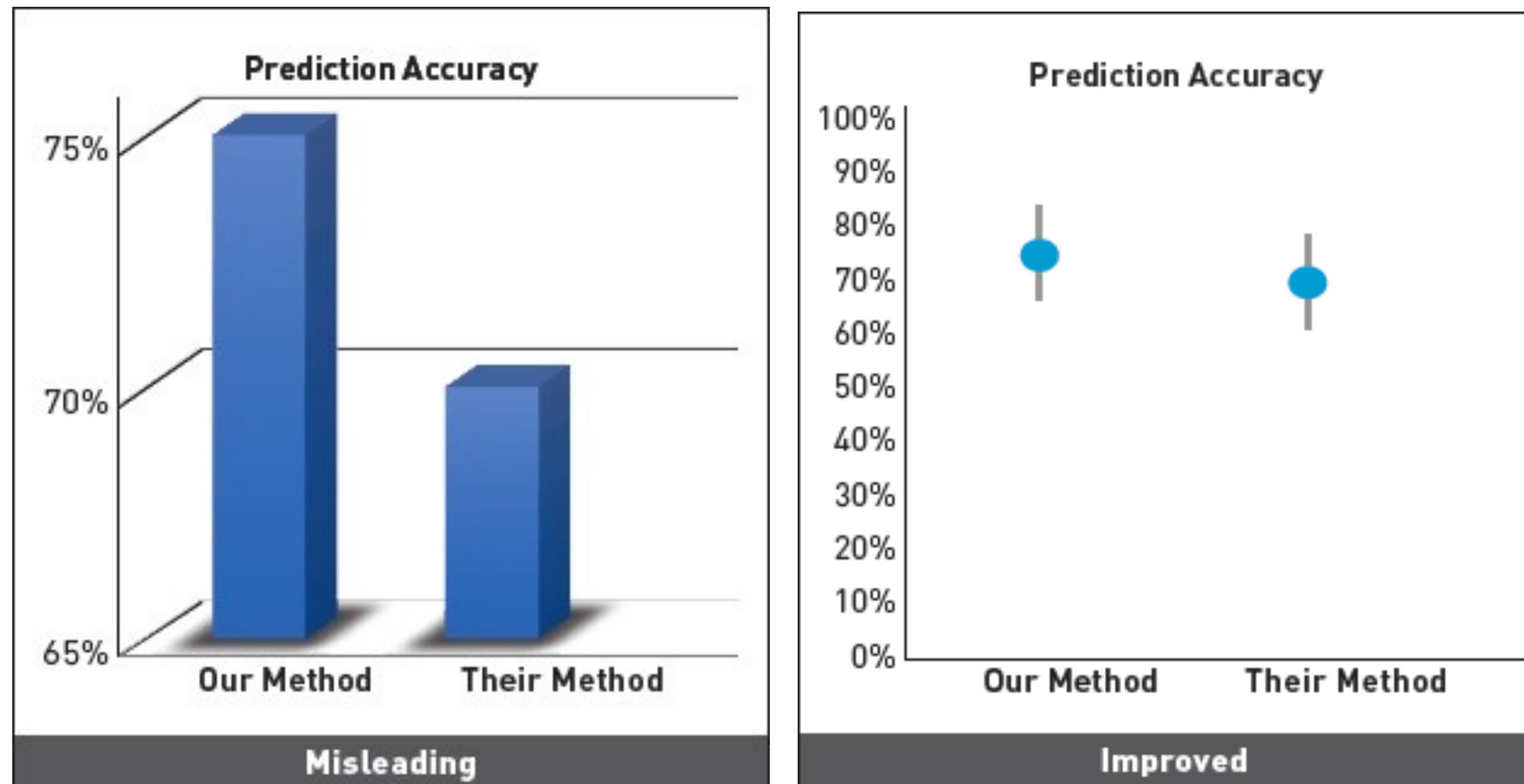


[<http://perceptualedge.com/files/GraphDesignIQ.html>]

3D Design Alternatives



3D Design Alternatives



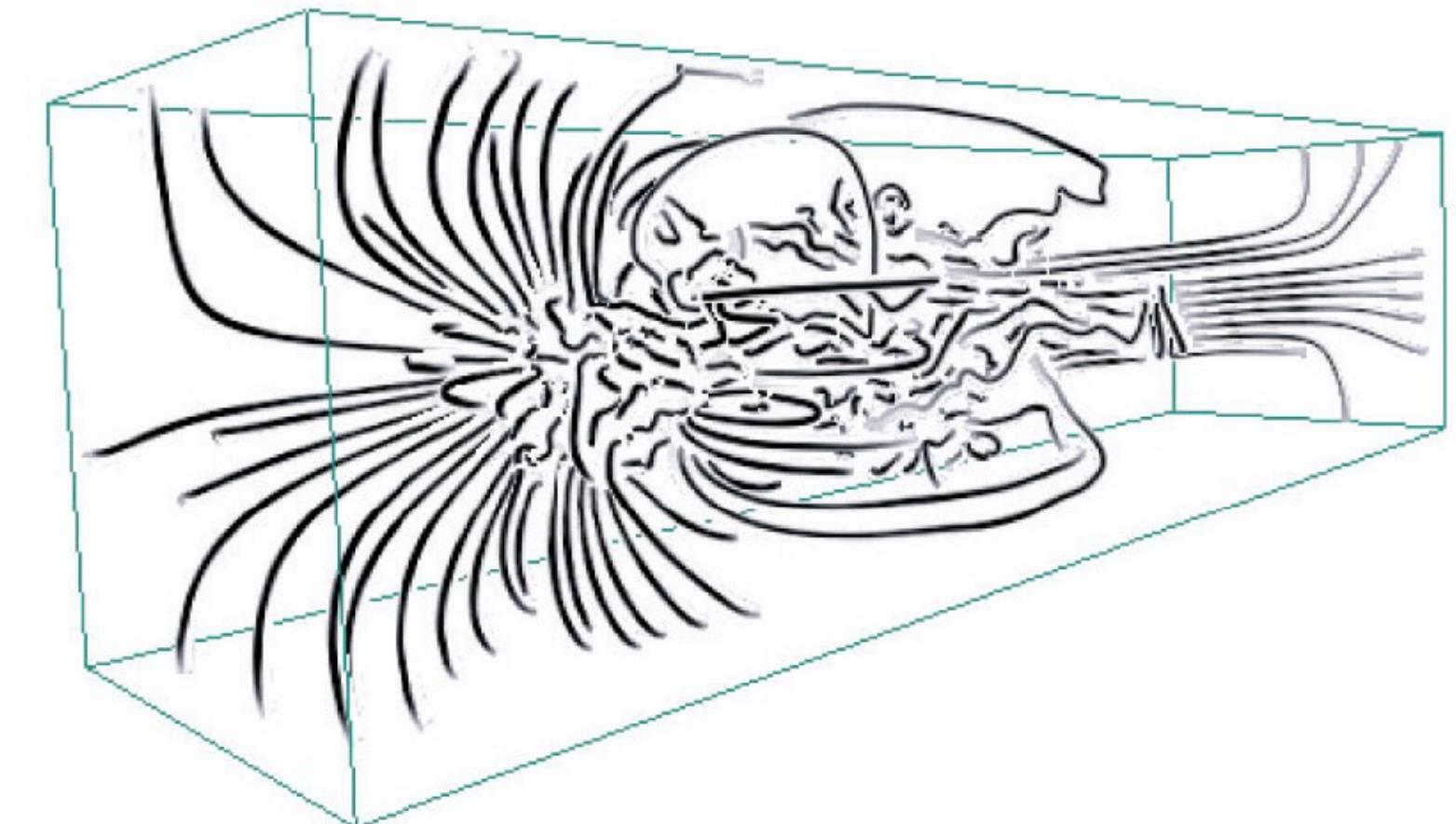
Justified 3D: shape perception

- benefits outweigh costs when task is shape perception for 3D spatial data
 - interactive navigation supports synthesis across many viewpoints

Targets

→ Spatial Data

→ Shape



[*Image-Based Streamline Generation and Rendering*. Li and Shen.

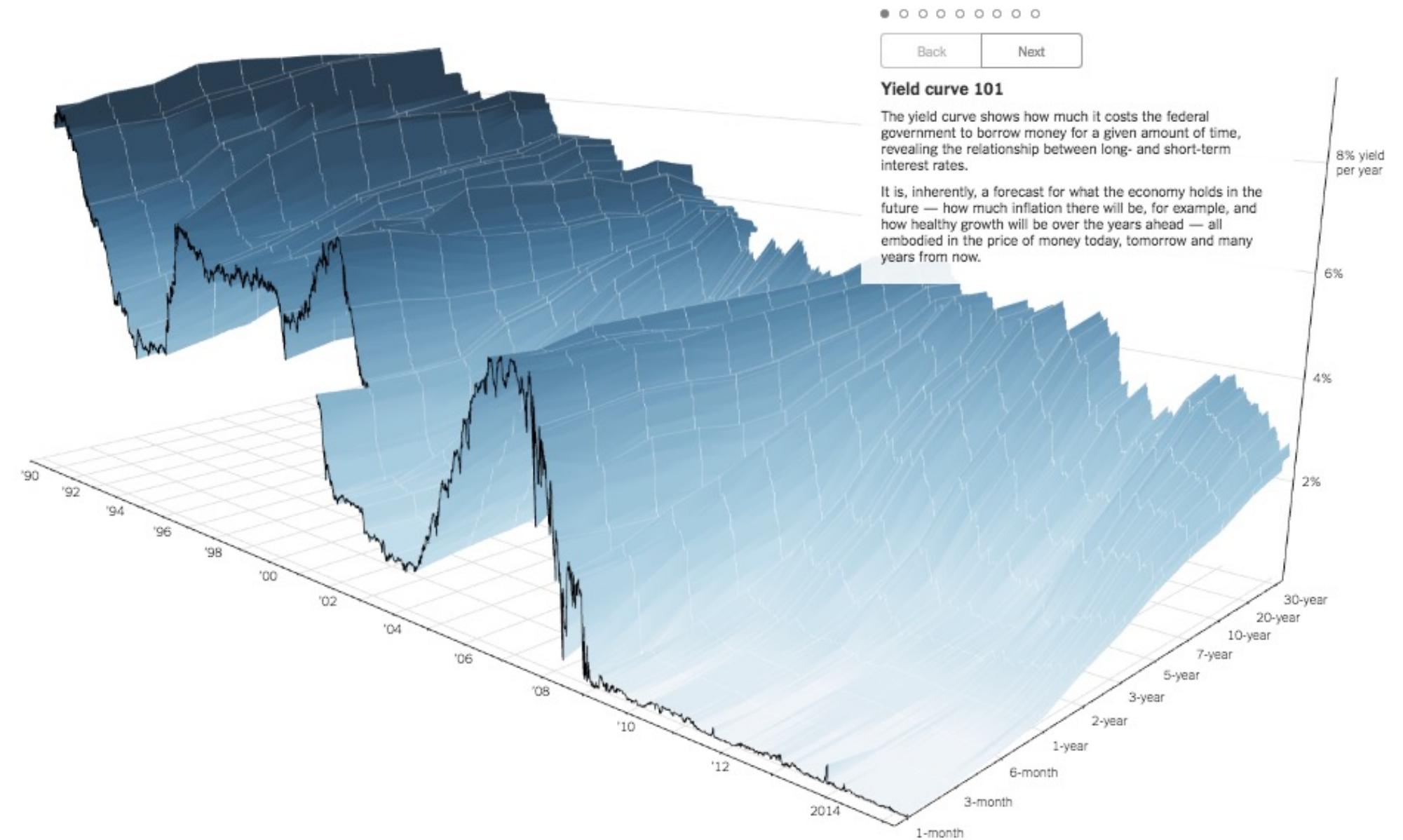
IEEE Trans. Visualization and Computer Graphics (TVCG) 13:3 (2007), 630–640.]

Justified 3D: Economic growth curve

- constrained navigation steps through carefully designed viewpoints

A 3-D View of a Chart That Predicts The Economic Future: The Yield Curve

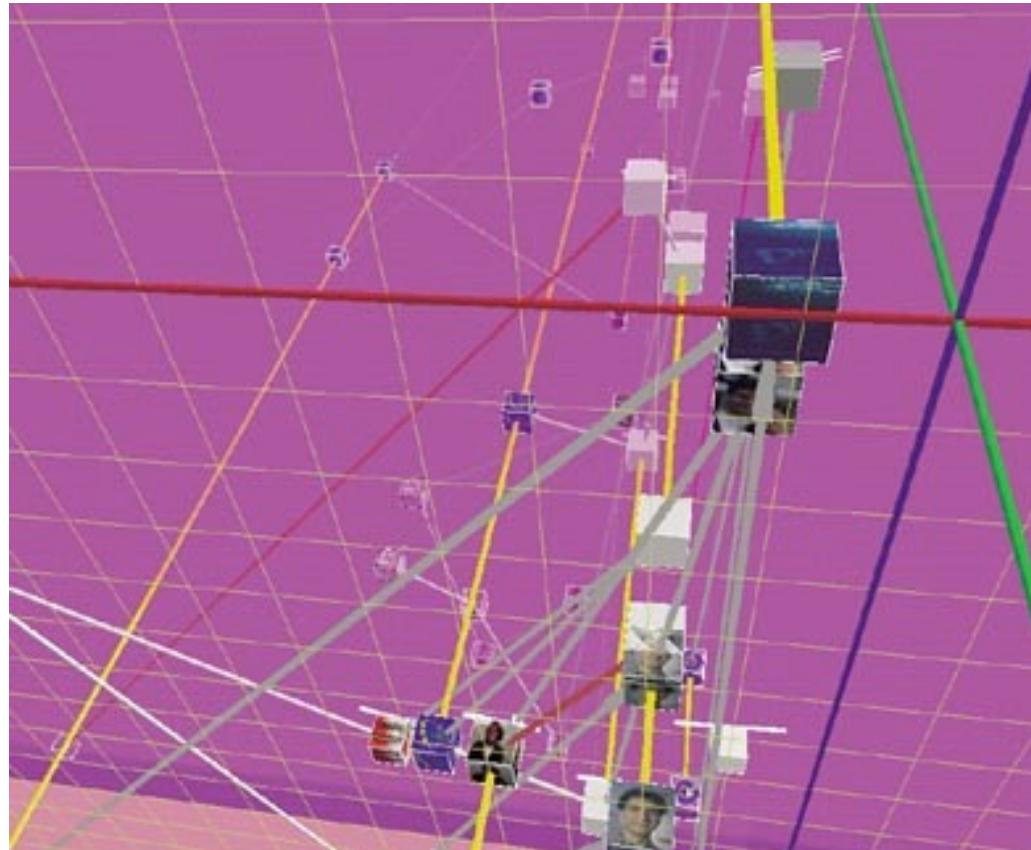
By GREGOR AISCH and AMANDA COX MARCH 18, 2015



<http://www.nytimes.com/interactive/2015/03/19/upshot/3d-yield-curve-economic-growth.html>

No unjustified 3D

- 3D legitimate for true 3D spatial data
- 3D needs very careful justification for abstract data
 - enthusiasm in 1990s, but now skepticism
 - be especially careful with 3D for point clouds or networks

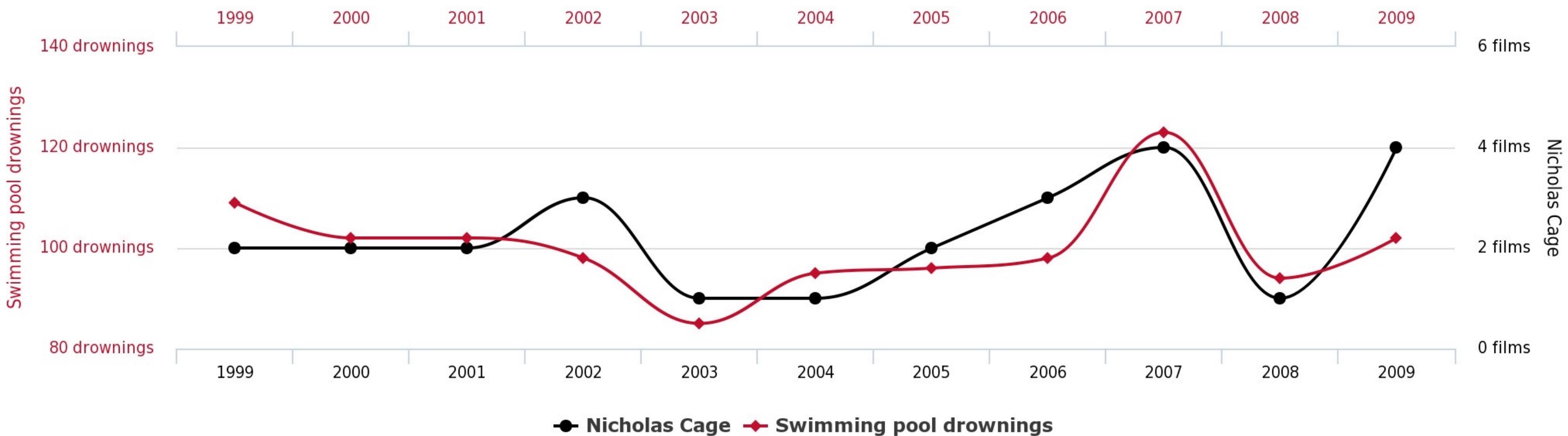


[WEBPATH-a three dimensional Web history. Frecon and Smith. Proc. InfoVis 1999]

Dual Axis

What is the message of this viz?

Number of people who drowned by falling into a pool
correlates with
Films Nicolas Cage appeared in



Dual Axis

DON'T TEAM

Misinterpretation Risk: They can easily lead to misunderstandings if viewers incorrectly assume a direct relationship between the two data sets.

Scale Issues: Different scales on the axes can misrepresent the relationship or magnitude of the data sets.

Complexity: They often result in cluttered and complicated visuals, hindering quick and clear comprehension.

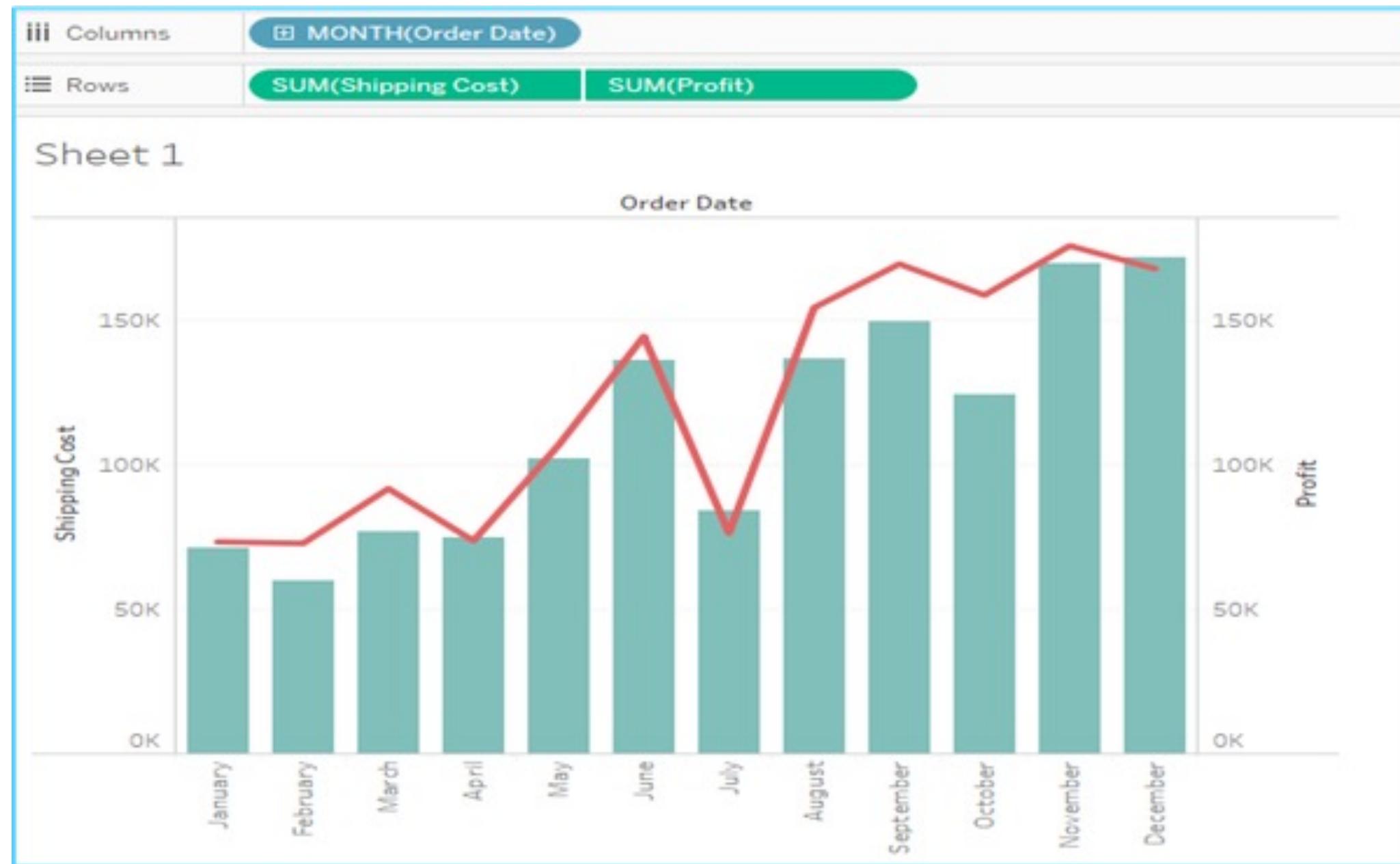
DO TEAM

Space Efficiency: They allow compact displaying of two related data sets in a single chart.

Correlation Highlighting: When used appropriately, they can effectively demonstrate correlations or interactions between variables.

Flexibility: They provide a versatile method for presenting and comparing diverse types of data that might not be easily combined otherwise.

Dual Axis



Dual Axis Design Principles

- **Use only IF YOU MUST**
- **Clear Labeling:** Ensure that both axes are clearly labeled, and it's evident which data set corresponds to which axis.
- **Consistent Scale:** If possible, use consistent scales for both axes to avoid misleading viewers.
- **Purposeful Use:** Only use dual-axis charts when there's a clear, logical reason to compare the two data sets.
- **Educate the Audience:** When presenting a dual-axis chart, it's often helpful to guide the audience through the data to avoid misinterpretation.

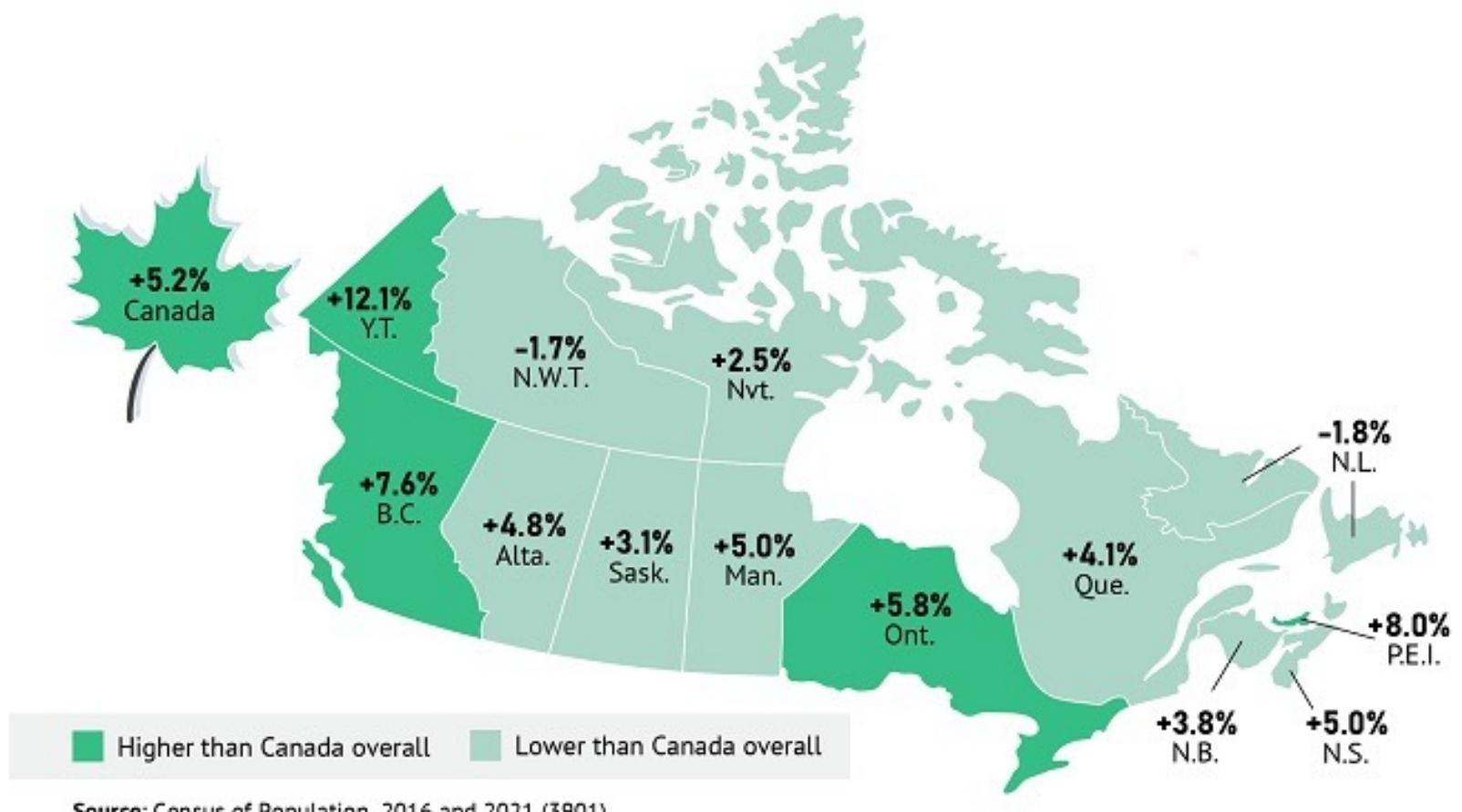
Rules of Thumb

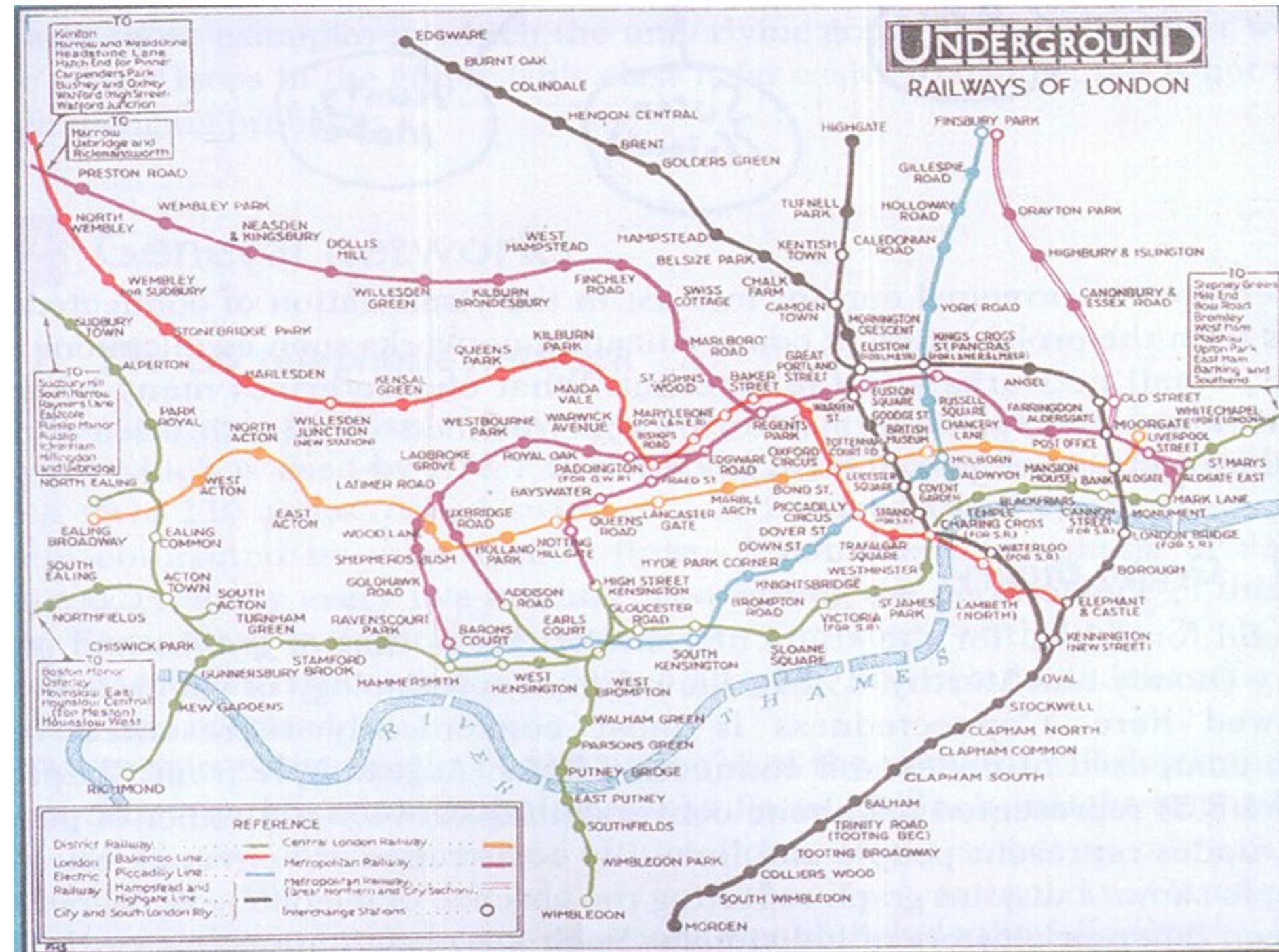
Guidelines and considerations, not absolute rules

- Animation vs. multiple views?
- when to use 3D? when to use 2D?
- which comes first, form or function?
- Labelling
- Maximize data ink ratio & Keep it simple
- If 1 is good, 2 must be better, the case for Dual Axis
- Vis, channel and data agreements

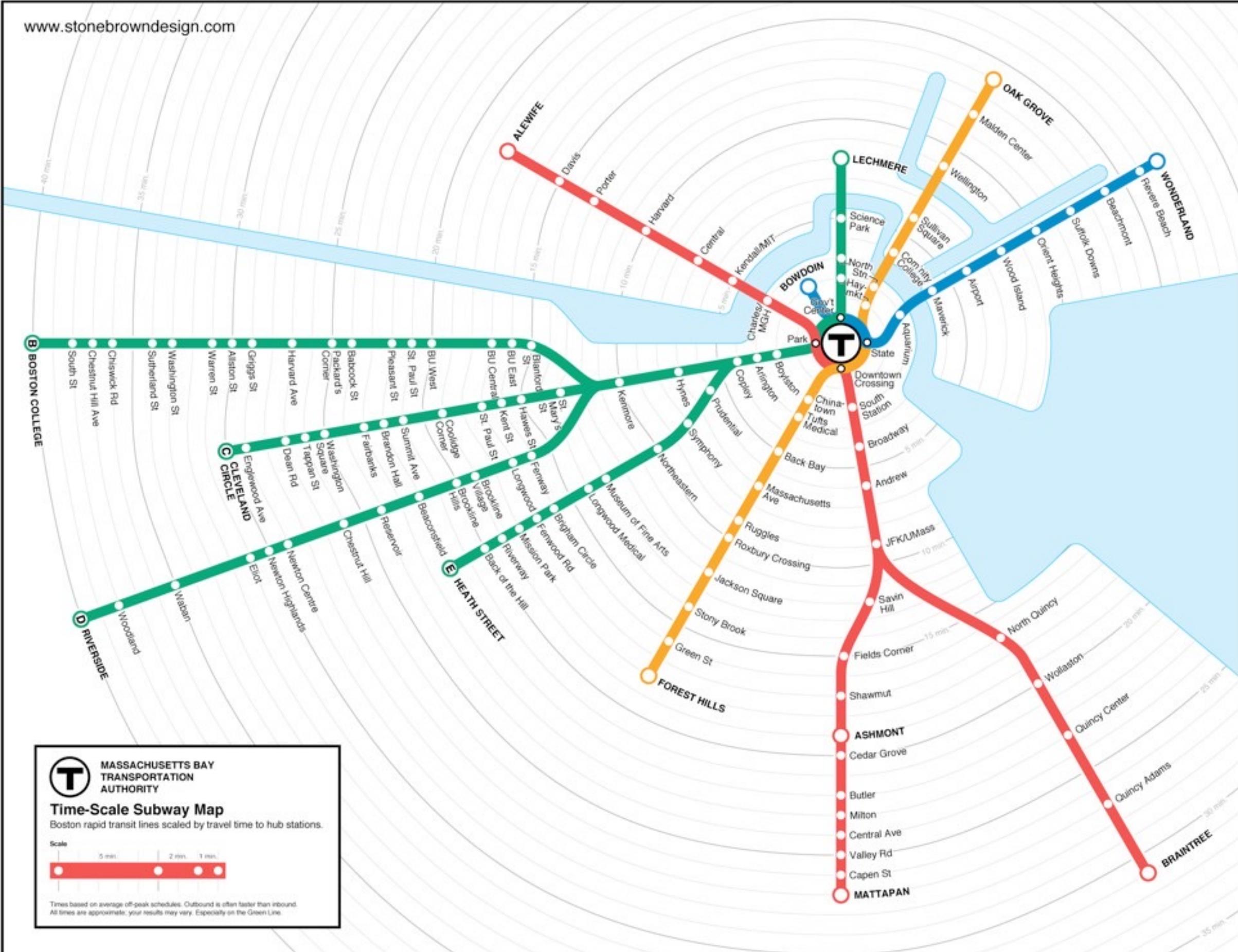
Vis, Channel and Data Agreement

- Channel/Data: Use the Best Visual Channel Available for the Most Important Aspect of your Data
- Vis/Data: The visualization should show all of the data, and only the data that is relative to the given task





London Subway Map,
1927



Rules of Thumb

Guidelines and considerations, not absolute rules

- Animation vs. multiple views?
- when to use 3D? when to use 2D?
- which comes first, form or function?
- Labelling
- Maximize data-ink ratio & Keep it simple
- If 1 is good, 2 must the better, the case for Dual Axis
- Vis, channel and data agreements