Winning Ways for Your Visualization Plays

Mark Grundland
Functional Elegance
Mark Grundland

consultancy  entrepreneurship  research  analysis  design  development
innovation  commercialization  software  system  engineering
user experience  design  requirements  analysis
technology  research  business  strategy
market analysis  prototype  design
startup fundraising  business  development
markets  customers  strategy  insight  data  technology  opportunity

Strategy Consultant

Data Scientist

Product Manager
Visualization is old as art but it is just getting started

“I am here”
Hand cloud
invented 35,000 B.C.

http://en.wikipedia.org/wiki/Cueva_de_las_Manos
Visualization is old as art but it is just getting started

“I am here”
Hand cloud
invented 35,000 B.C.

“I blog here”
Word cloud
invented 1992 A.D.

http://en.wikipedia.org/wiki/Cueva_de_las_Manos
Seeing the pattern in the data can change how we view our world

Modern epidemiology started with plotting dots on a map.

http://en.wikipedia.org/wiki/1854_Broad_Street_cholera_outbreak
Modern epidemiology started with plotting dots on a map. In 1854, a cholera outbreak in Soho killed over 600 people. John Snow plotted the locations of the deaths to show that they were clustered around the neighborhood water pump.

http://en.wikipedia.org/wiki/1854_Broad_Street_cholera_outbreak
Communicating data effectively can change what we do with our world.

Hockey stick graph was arguably the most controversial chart in science: our future depends on how we read it.

Figure 1: Variations of the Earth’s surface temperature over the last 140 years and the last millennium.

(a) The Earth’s surface temperature is shown year by year (red bars) and approximately decade by decade (black line, a filtered annual curve suppressing fluctuations below near decadal time-scales). There are uncertainties in the annual data (thin black whisker bars represent the 95% confidence range) due to data gaps, random instrumental errors and uncertainties, uncertainties in bias corrections in the ocean surface temperature data and also in adjustments for urbanisation over the land. Over both the last 140 years and 100 years, the best estimate is that the global average surface temperature has increased by 0.6° ± 0.2°C.

(b) Additionally, the year by year (blue curve) and 50 year average (black curve) variations of the average surface temperature of the Northern Hemisphere for the past 1000 years have been reconstructed from “proxy” data calibrated against thermometer data (see list of the main proxy data in the diagram). The 95% confidence range in the annual data is represented by the grey region. These uncertainties increase in more distant times and are always much larger than in the instrumental record due to the use of relatively sparse proxy data. Nevertheless the rate and duration of warming of the 20th century has been much greater than in any of the previous nine centuries. Similarly, it is likely that the 1990s have been the warmest decade and 1998 the warmest year of the millennium.

[Based upon (a) Chapter 2, Figure 2.7c and (b) Chapter 2, Figure 2.20]
Communicating data effectively can change what we do with our world.

Hockey stick graph was arguably the most controversial chart in science: our future depends on how we read it.

Figure 1: Variations of the Earth's surface temperature over the last 140 years and the last millennium.

(a) The Earth's surface temperature is shown year by year (red bars) and approximately decade by decade (black line, a filtered annual curve suppressing fluctuations below near decadal time-scales). There are uncertainties in the annual data (thin black whisker bars represent the 95% confidence range) due to data gaps, random instrumental errors and uncertainties, uncertainties in bias corrections in the ocean surface temperature data and also in adjustments for urbanisation over the land. Over both the last 140 years and 100 years, the best estimate is that the global average surface temperature has increased by 0.6 ± 0.2°C.

(b) Additionally, the year by year (blue curve) and 50 year average (black curve) variations of the average surface temperature of the Northern Hemisphere for the past 1000 years have been reconstructed from “proxy” data calibrated against thermometer data (see list of the main proxy data in the diagram). The 95% confidence range in the annual data is represented by the grey region. These uncertainties increase in more distant times and are always much larger than in the instrumental record due to the use of relatively sparse proxy data. Nevertheless the rate and duration of warming of the 20th century has been much greater than in any of the previous nine centuries. Similarly, it is likely that the 1990s have been the warmest decade and 1998 the warmest year of the millennium.

[Based upon (a) Chapter 2, Figure 2.7c and (b) Chapter 2, Figure 2.20]
How can we display our data without distorting the truth?

Hockey stick graph was arguably the most controversial chart in science: our future depends on how we read it.
How can we display our data without distorting the truth?

The visualization design decisions we make affect which interpretations of the data are facilitated or impeded.

Departures in temperature ($°C$) is likely the previous nine centuries. Similarly, it has been much greater than in any of the duration of warming of the 20th century proxy data. Nevertheless the rate and record due to the use of relatively sparse grey region. These uncertainties increase the annual data is represented by the diagram). The 95% confidence range in list of the main proxy data in the calibrated against thermometer data (see reconstructed from "proxy" data for the past 1000 years have been temperature of the Northern Hemisphere curve) variations of the average surface curve) and 50 year average (black line, a filtered annual curve suppressing approximately decade by decade (black shown year by year (red bars) and (b) Additionally, the year by year (blue NORTHERN HEMISPHERE

Data from thermometers (red) and from tree rings, corals, ice cores and historical records (blue).

http://www.takepart.com/an-inconvenient-truth/film
How can we select the aspect ratio?

- Use 1:1 ... **Why?** It is fair and square.
- Use 3:2 ... **Why?** It is wider than taller, like a landscape photo.
- Use the golden ratio... **Why?** It is a most pleasing proportion found in nature and art.
- Make the average slope of all line segments 45°... **Why?** It is perceptually optimal for orientation discrimination.
- Minimize arc length, keeping area under the plot constant... **Why?** It is short, sweet, and mathematically optimal.
- Take the screen size or the widow size as given... **Why?** It fits, so obviously this must be what the user wants.
- Depends on the situation... **Why?** It depends on the story the user is meant to believe.

http://www.manorpress.co.uk/printing%20info.html
How can we select the map projection?

Representing the Earth on a flat map must in some way distort distances, directions, angles, shapes, and/or areas.

**Mercator Projection**
Preserves angles but not areas

How can we select the map projection?

Representing the Earth on a flat map must in some way distort distances, directions, angles, shapes, and/or areas.

Mercator Projection
Preserves angles but not areas

http://maps.google.com/
How can we select the map projection?

Representing the Earth on a flat map must in some way distort distances, directions, angles, shapes, and/or areas.

- Tissot indicatrix measures geometric distortion by showing how circles on the globe appear as ellipses on the map.

**Mercator Projection**
Preserves angles but not areas

**Robinson Projection**
Nearly preserves areas but not angles

http://wordpress.mrreid.org/2011/06/07/tissots-indicatrix/
http://www.progonos.com/furuti/MapProj/Normal/TOC/cartTOC.html
How can we select the map projection?

Representing the Earth on a flat map must in some way distort distances, directions, angles, shapes, and/or areas.

- Cartograms distort the size and shape of regions in order to make their area proportional to a given variable of interest.
- Computed using density diffusion or cellular automata.

[Image of equal area cartograms]
How can we select the map projection?

Representing the Earth on a flat map must in some way distort distances, directions, angles, shapes, and/or areas.

- Cartograms distort the size and shape of regions in order to make their area proportional to a given variable of interest
- Computed using density diffusion or cellular automata

Population
Equal area cartogram

GDP Wealth
Equal area cartogram

http://www.worldmapper.org/
Why do we visually represent values?

Though data is easily summarized by numbers, information is best communicated by patterns.

Same pattern ≠ Same statistics
Same means $\mu$, variances $\sigma^2$, correlation $R^2$, and regression

http://en.wikipedia.org/wiki/Anscombe%27s_quartet

Image Credit: Michael Taylor
Why do we visually represent values?

Though data is easily summarized by numbers, information is best communicated by patterns.

Same relationship ≠ Same correlation

http://en.wikipedia.org/wiki/File:Correlation_examples.png#filelinks
Why do we visually represent values?

Though data is easily summarized by numbers, information is best communicated by patterns.

Visualizing the data ≠ Visualizing the statistics

http://www.nature.com/nmeth/journal/v11/n2/full/nmeth.2811.html
http://flowingdata.com/2012/05/15/how-to-visualize-and-compare-distributions/
How can we visually represent values?

Numeric values express absolute magnitudes but visual perception makes relative judgments.

- Position
- Shape
- Length
- Orientation
- Area and volume
- Hue, saturation, brightness
- Texture and transparency
- Alignment and proximity
- Containment and connection
- Labels and glyphs
- Motion and flicker

How can we visually represent values?

Numeric values express absolute magnitudes but visual perception makes relative judgments, not very well.

Which is bigger?
And by how much?

http://michaelbach.de/ot/
How can we visually represent values?

Numeric values express absolute magnitudes but visual perception makes relative judgments, not very well.

Which is bigger?
And by how much?

http://michaelbach.de/ot/
How can we visually represent values?

Numeric values express absolute magnitudes but visual perception makes relative judgments, not very well.

**Equity Market Heat Map**

[Image of the equity market heat map]

http://www.marketwatch.com/tools/stockresearch/marketmap
How can we visually represent values?

Numeric values express absolute magnitudes but visual perception makes relative judgments, not very well.

Which is bigger?
And by how much?

http://en.wikipedia.org/wiki/Ebbinghaus_illusion
How can we visually represent values?

Numeric values express absolute magnitudes but visual perception makes relative judgments, not very well.

Which is bigger? And by how much?

http://en.wikipedia.org/wiki/Ebbinghaus_illusion
How can we visually represent values?

Numeric values express absolute magnitudes but visual perception makes relative judgments, not very well.

Which is bigger?
And by how much?

http://en.wikipedia.org/wiki/Ebbinghaus_illusion
How can we visually represent values?

Numeric values express absolute magnitudes but visual perception makes relative judgments, not very well.

UK Government Spending Bubble Chart

How can we visually represent values?

Numeric values express absolute magnitudes but visual perception makes relative judgments, not very well.

Which is brighter? And by how much?

How can we visually represent values?

Numeric values express absolute magnitudes but visual perception makes relative judgments, not very well.

Which is brighter?
And by how much?

How can we visually represent values?

Numeric values express absolute magnitudes but visual perception makes relative judgments, not very well.

US Presidential Election Map

http://bsumaps.blogspot.co.uk/2012/11/cartographic-election-resources.html
How can we visually represent values?

Numeric values express absolute magnitudes but visual perception makes relative judgments, not very well.

US Presidential Election Map

How can we visually represent values?

Help users by labeling data and adding trend indicators.

US Presidential Election Map

How can we visually represent values?

Help users by labeling data and adding trend indicators.

**Online Ad Campaign Performance**
**Bar Chart**

**Best Performing Content Categories**

Top English Content Categories:

- Europe (Regional news)
- Movies and video games (Arts)
- Issues (Society)
- North America (Regional news)
- Literature (Arts)
- Software (Computers)
- Asia (Regional news)
- Oceania (Regional news)
- Financial services (Business)
- Graphics (Computers)
- Music (Arts)
- Investing (Business)
- Technology (Science)
- Middle East (Regional news)
- History (Society)
- Business services (Business)
- Accounting (Business)
- Construction and maintenance (Business)
- Consumer goods and services (Business)
- Marketing and advertising (Business)
- Transportation and logistics (Business)
- Systems (Computers)
- Collecting (Recreation)
- Guns (Recreation)
- Recreation (Shopping)

**Legend:**
- Red: Impressions
- Green: Clicks

http://www.FunctionalElegance.com
How can we visually represent values?

Help users by labeling data and adding trend indicators.

Online Ad Campaign Performance
Traffic Light Bar Chart

Best Performing Content Categories:

- Europe (Regional news) [28.0% → 20.5%]
- Movies and video games (Arts) [9.6% → 10.7%]
- Issues (Society) [10.2% → 9.8%]
- North America (Regional news) [6.2% → 9.8%]
- Literature (Arts) [9.2% → 2.7%]
- Software (Computers) [0.6% → 6.3%]
- Asia (Regional news) [3.7% → 6.3%]
- Oceania (Regional news) [5.6% → 6.3%]
- Financial services (Business) [3.1% → 5.4%]
- Graphics (Computers) [0.0% → 5.4%]
- Music (Arts) [4.0% → 0.9%]
- Investing (Business) [1.7% → 3.6%]
- Technology (Science) [1.8% → 0.9%]
- Middle east (Regional news) [0.9% → 1.8%]
- History (Society) [1.4% → 0.9%]
- Business services (Business) [0.9% → 0.0%]
- Accounting (Business) [0.1% → 0.9%]
- Construction and maintenance (Business) [0.7% → 0.9%]
- Consumer goods and services (Business) [0.4% → 0.9%]
- Marketing and advertising (Business) [0.2% → 0.9%]
- Transportation and logistics (Business) [0.5% → 0.9%]
- Systems (Computers) [0.4% → 0.9%]
- Collecting (Recreation) [0.1% → 0.9%]
- Guns (Recreation) [0.1% → 0.9%]
- Recreation (Shopping) [0.0% → 0.9%]

Legend:
- Impressions
- Clicks
- Favorable
- Neutral
- Unfavorable

http://www.FunctionalElegance.com
How can we visually represent values?

Perceptually uniform color gradients for continuous values. Avoid rainbow color maps:

✦ Hue order is not obvious.
✦ Hue changes make edges.
✦ Yellows make highlights.
✦ Detail is harder to see.
✦ Eyes are more sensitive to brightness than hues.

http://blog.visual.ly/rainbow-color-scales/
http://researchweb.watson.ibm.com/people/l/lloydt/color/color.HTM
How can we visually represent values?

Use transparency to overlay information layers.

- Normally transparent layers are composited using linear interpolation, an averaging operation that reduces variation.
- Blending by linear interpolation can result in reduced contrast, dull colors, detail loss, and a lack of selective emphasis.

Satellite Map Overlay

http://lsntap.org/blogs/gis-mapping-part-i
How can we visually represent values?

Use transparency to overlay information layers.

- Apply image blending operators that are designed to produce composite images that preserve key visual characteristics of their components: contrast, color, detail, and salience.
How can we visually represent values?

Use transparency to overlay information layers.
- Apply image blending operators that are designed to produce composite images that preserve key visual characteristics of their components: contrast, color, detail, and salience.
How can we visually represent values?

Use transparency to overlay information layers.

- Render the arc of each node connection in order of decreasing length using a color gradient that emphasizes short links.

Facebook Friend Connection Map

http://paulbutler.org/archives/visualizing-facebook-friends/
How can we visually represent values?

Familiar visual metaphors make interpretation easier.

**SnapShot News Analysis Tool**

[Image of the SnapShot News Analysis Tool interface]

- Title news analysis: News Analysis
- Read news data file: [File name]
- Use news articles (%): + 100

- Publication date (%): + 50
- Publication date range (%): + 100

- Group news: Featured topics
- Subgroup news: None

- Rank news groups: Total news articles
- Order news groups: Descending order
- Show top news groups (%): + 42
- Skip small news groups: [Checkbox]

- Plot vertical axis: Average story influence
- Plot horizontal axis: Median query relevance
- Scale plot points: Total news articles
- Colour plot points: Average source prestige
- Label plot points: [Checkbox]
- Lock plot axes: [Checkbox]

[News Analysis: iPod Zune graph]

News Analysis: iPod Zune

12:06AM 12/02/2008 – 10:54PM 12/02/2008

- High Impact
- Low Impact
- News Article Relevance
- News
- Electronic Equipment
- Business
- Wires
- Information Technology
- Media
- Arts
- US Local News
- Miscellaneous

URLs:
- [Grapeshot.co.uk](http://www.grapeshot.co.uk/snapshot/snapshot.html)
- [FunctionalElegance.com](http://www.FunctionalElegance.com)
How can we visually represent values?

Familiar visual metaphors make interpretation easier.

**SnapShot News Radar**
How can we visually represent values?

Familiar visual metaphors make interpretation easier.

SnapShot News Radar

http://www.grapeshot.co.uk/snapshot/snapshot.html
http://www.FunctionalElegance.com
How can we visually represent values?

Familiar visual metaphors make interpretation easier.

SnapShot News Radar

Annotate
Measure
Group
Aggregate

News Annotations:
- Precedents
- Consequences

News Metrics:
- Influence
- Attention
- Novelty
- Effect

Influence = (Consequences - Precedents) / (Precedents + Consequences)
Attention = Median Rank of News Articles
Novelty = Median Rank of News Articles
Effect = (+ Consequences)

http://www.grapeshot.co.uk/snapshot/snapshot.html
http://www.FunctionalElegance.com
How can we visually represent connections?

Use proximity to find connections in a cloud of points.

- Take a sphere of influence around each point, with radius equal to its nearest neighbor distance, and connect every pair of points whose spheres of influence intersect.

**Sphere of Influence Graphs**

Work in $\mathbb{R}^n$ for any $L_p$ Metric

http://dx.doi.org/10.1016/S0166-218X(02)00246-9
How can we visually represent connections?

Use tree drawing to see patterns in hierarchical data.

Coloring each node according to its data type reveals the structure of expression trees, such as XML, JSON, and HTML.

Website Home Pages as HTML Trees

Website:
- Google.com
- Yahoo.com

Images from:
- http://www.generatorx.no/20060528/websites-as-graphs/
How can we visually represent connections?

Use word clouds that place their terms meaningfully.

- TextArc writes the sentences of a text along a circular arc and places each term according to its average position in the text.

“Alice in Wonderland” TextArc

http://www.textarc.org/
How can we visually represent connections?

Use word clouds that place their terms meaningfully.
- Self-organizing maps (SOM) are neural networks, which can create knowledge maps that cluster closely associated topics.

Self-organizing Map of a Mailing List

http://dx.doi.org/10.1371/journal.pone.0058779
http://wiki.sugarlabs.org/go/Sugar_Labs/SOM
How can we visually represent connections?

Use graph drawing tools to see how data is connected.

- Graph drawing algorithms, such as simulated annealing or spring systems, make it easier to follow how data is related.

Graph Drawing of Medical Knowledge

http://cs.brown.edu/~rt/gdhandbook/
http://visualization.geblogs.com/visualization/network/
How can we visually represent connections?

Use graph drawing tools to see how data is connected.

- Hierarchical edge bundles group connections belonging to related nodes, which can be placed radially along a circle.

**Graph Drawing of Medical Knowledge**

![Graph Drawing of Medical Knowledge](http://www.win.tue.nl/~dholten/papers/bundles_infovis.pdf)

http://visualization.geblogs.com/visualization/network/
What is effective visualization design?

Communicating the message that makes a difference means more than just plotting the data.

Global Income Inequality

Bottoms Up
% of world’s wealth owned by...

- richest 20% of the world: 82.7%
- poorest 20%: 1.4%

Source: UN

https://www.amazon.co.uk/Information-Beautiful-David-McCandless/dp/0007294662
What is effective visualization design?

Communicating the message that makes a difference means more than just plotting the data.

Global Income Inequality

Share of Global Income Going to the Richest 20% and Poorest 20% of the World's Population

https://www.office.com/
What is effective visualization design?

Communicating the message that makes a difference means more than just plotting the data.

**Global Income Inequality**

- Share of Global Income Going to the Richest 20% and Poorest 20% of the World's Population

- Richest 20%: 82.7%
- Poorest 20%: 1.4%

Sources:
- [Show Me the Numbers](https://www.amazon.co.uk/Show-Me-Numbers-Stephen-Few/dp/0970601972)
What is effective visualization design?

Communicating the message that makes a difference means more than just plotting the data.

Global Income Inequality

The richest 20% own 82.7% of the world’s income. The poorest 20% own 1.4% of the world’s income.

https://www.huffingtonpost.co.uk/entry/hunger-photo-series_n_6297410
What is effective visualization design?

Communicating the message that makes a difference means more than just plotting the data.

The richest 20% own more than 80% of the world’s income.

The poorest 20% own less than 2% of the world’s income.

Is that right?
What is the secret to great information visualization design?

Start with the user, not the data and not the graphic.

- What will this allow you to do that you can't do now?
- What difference can you observe in your business?
- What value do you expect that this will add to your business?
- What will this let your customers do that they can't do now?
- What difference can your customers observe in their business?
- What value can your customers expect that this will add?
- How does this fit in with your other strategic plans?
- How will you know that this has been a success?
- How will you build on this success?
- So what?
What does the future hold for information visualization?

In an information economy, there is no shortage of information; only understanding is in short supply.

- Will interactive charts become more common as letting people play with the data drives engagement?
- Will subtly animated infographics become more common as graphic designers compete for attention?
- Will augmented reality glasses superimpose an information visualization layer on our everyday lives?
- Will cheap displays make ambient visualization ubiquitous?
- Will virtual reality and gesture interfaces have an impact?

Let me know!

Mark@FunctionalElegance.com

http://www.FunctionalElegance.com
Online information visualization resources

Thank you for making this presentation possible!

Visualization Galleries:
- Tree Visualizations (Hans-Jörg Schulz): http://vcg.informatik.uni-rostock.de/~hs162/treeposter/poster.html
- Time Series Visualizations (Christian Tominski & Wolfgang Aigner): http://survey.timeviz.net/
- D3 JavaScript Visualization Library: https://github.com/mbostock/d3/wiki/Gallery

Visualization Courses:
- University of Utah (Miriah Meyer): http://www.sci.utah.edu/~miriah/cs6964/
- University of British Columbia (Tamara Munzner): http://www.cs.ubc.ca/~tmm/courses/533-09/
- University of California Berkeley (Michael Porath): http://blogs.ischool.berkeley.edu/i247s13/
- University of Washington (Jeffrey Heer): https://courses.cs.washington.edu/courses/cse512/14wi/
- Georgia Institute of Technology (John Stasko): http://www.cc.gatech.edu/~stasko/7450/syllabus.html

Visualization Tutorials:
- Storytelling with Data (Jonathan Corum): http://style.org/tapestry/
- Principles of Information Visualization (Jessie Kennedy): http://mkweb.bcgsc.ca/vizbi/2012/
- Information Visualization for Knowledge Discovery (Ben Shneiderman): http://bit.ly/1cw3oa2
- Data Visualization Best Practices (Jen Underwood): http://www.slideshare.net/idigdata/
- Data Visualization (Jan Willem Tulp): http://www.slideshare.net/janwillemtulp/

Visualization Sites:
- Perceptual Edge (Stephen Few): http://www.perceptualedge.com/
- Information is Beautiful (David McCandless): http://www.informationisbeautiful.net/
- Information Aesthetics (Andrew Vande Moere): http://infoaesthetics.com/
- Flowing Data (Nathan Yau): http://flowingdata.com/learning/