

Data Visualization Best Practices

March 29, 2018



Course Outline

- **Why visualize**
- **Good practice guidelines**
- **Basic visualizations**
- **Composition rules**
- **Interactivity**

Why Visualize?

- Data visualization shifts the balance from cognition (thinking) to perception (seeing)
- Seeing is handled by the visual cortex located in the rear of the brain, is extremely fast and efficient
- We see immediately, with little effort
- Thinking is handled primarily by the cerebral cortex in the front of the brain, is much slower and requires bigger effort
- Images are processed faster than audio or text
- Our brains have powerful image processing which works unconsciously

Gestalt Principles of Grouping

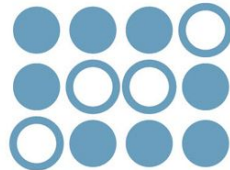
- a philosophy of mind developed by the Berlin School of experimental psychology, “Gestalt” is German for shape, form

- **Proximity:** close together objects are perceived as a group
- **Similarity:** object with same attributes (color, size) are perceived as a group
- **Enclosure:** objects within a boundary are perceived as a group
- **Closure:** open structures are perceived as closed if there is a way to interpret as such
- **Continuity:** aligned objects are perceived as a group
- **Connection:** connected objects are perceived as a group

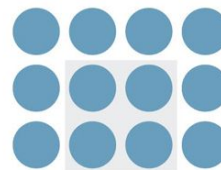
Proximity



Similarity



Enclosure



Connection



Continuity



Symmetry



Figure & Ground



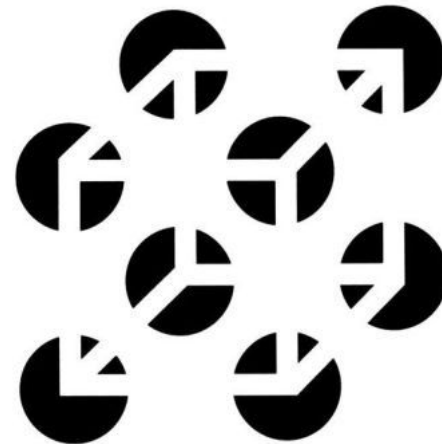
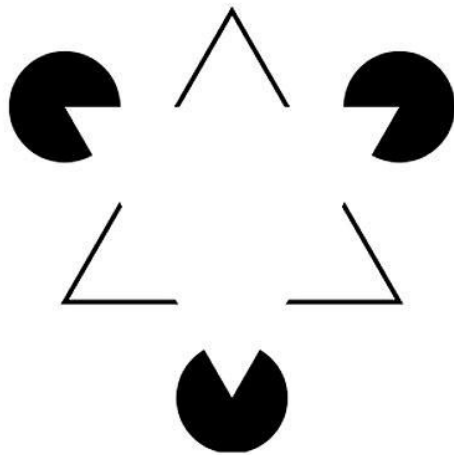
Closure



Common Fate

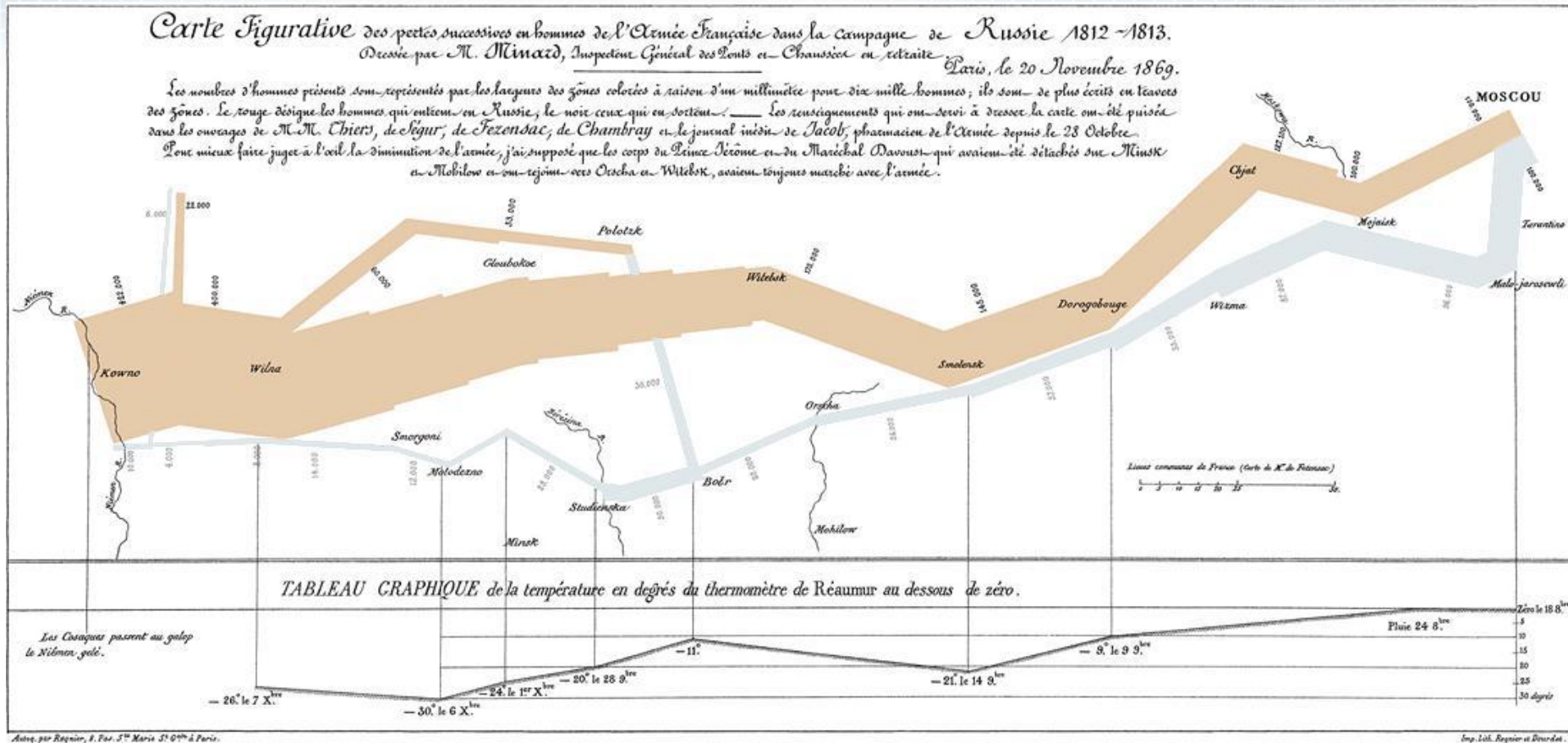


Exploit Ability of the Human Brain to Infer Simplest Solution for Incomplete Visual Information



Classic Visualization

Charles Minard: Napoleon's Russian Campaign of 1812

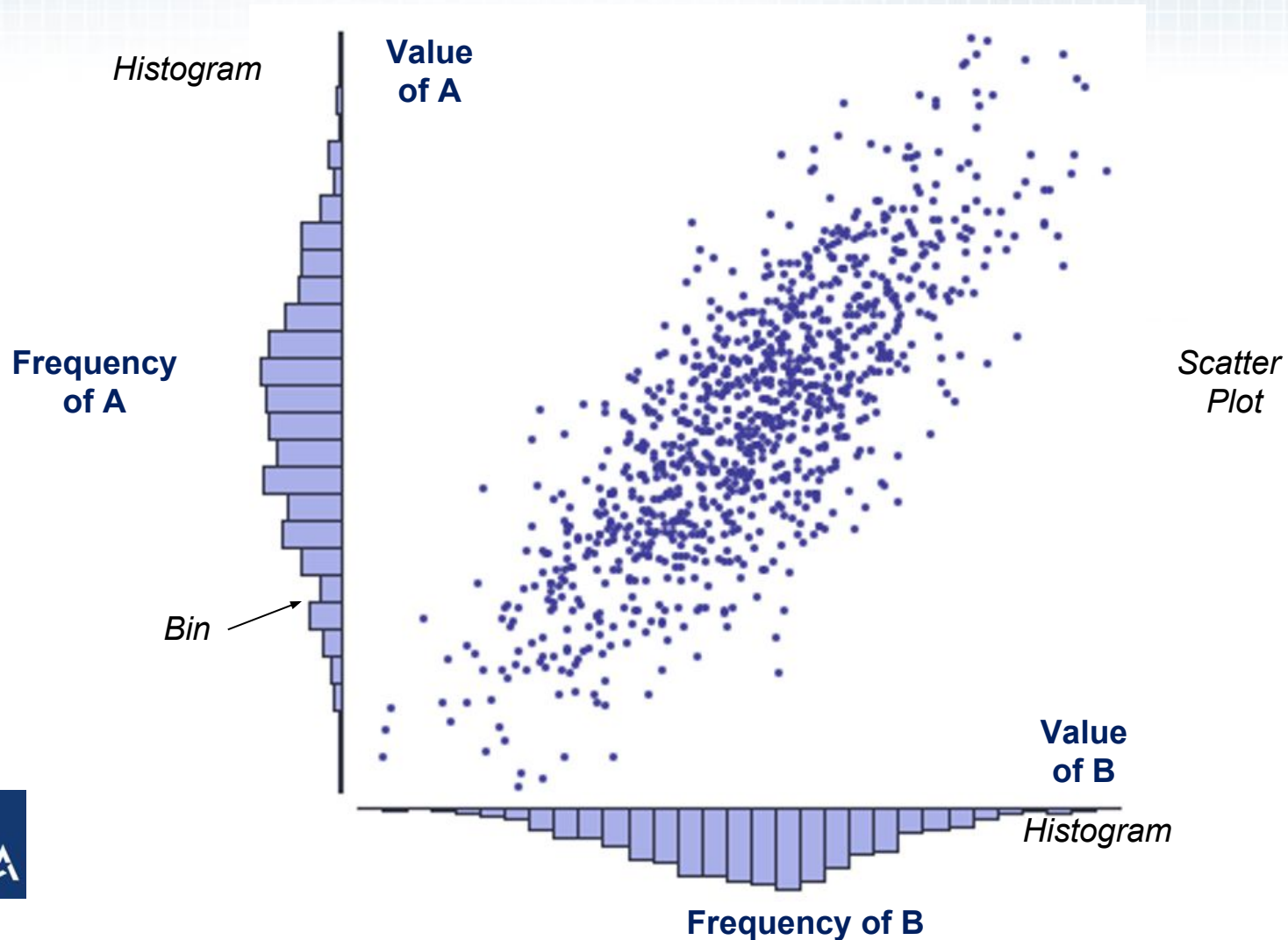


Edward Tufte's Visualization Goals

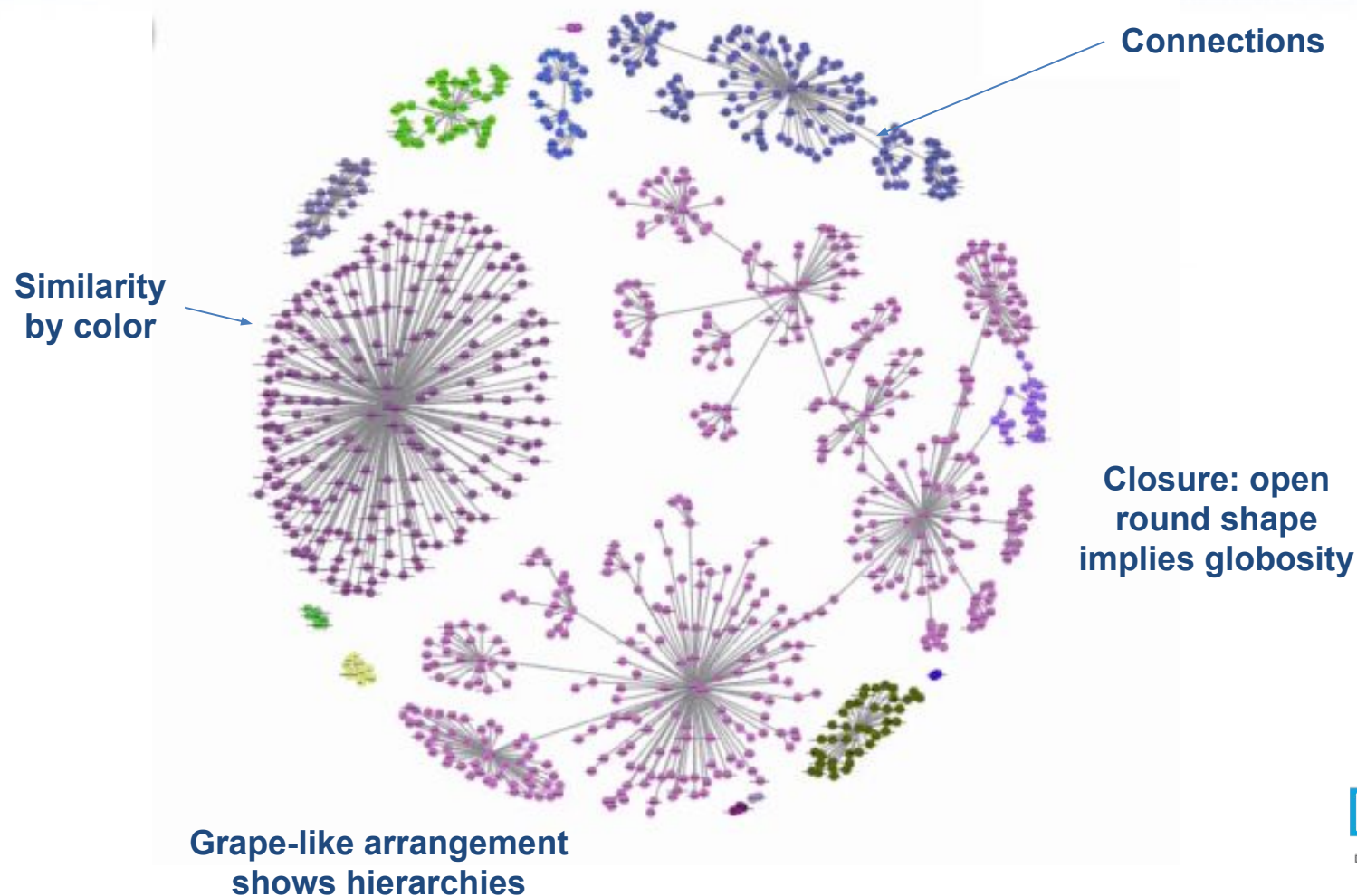
- **Explore**
- **Compare**
- **See trends**
- **Relate part to whole**
- **Rank**
- **Correlate**
- **Show distribution**

Data Exploration

Scatter Plot and Histogram

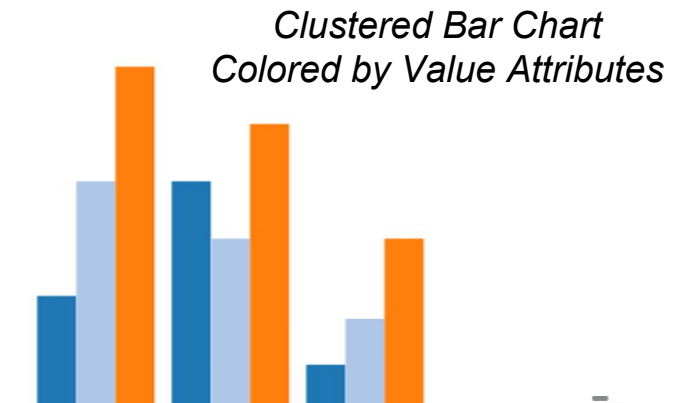
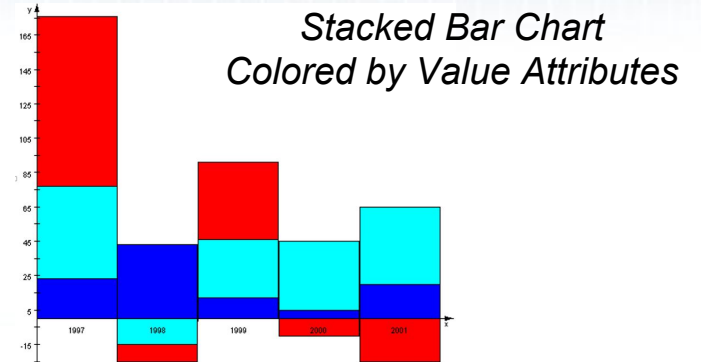
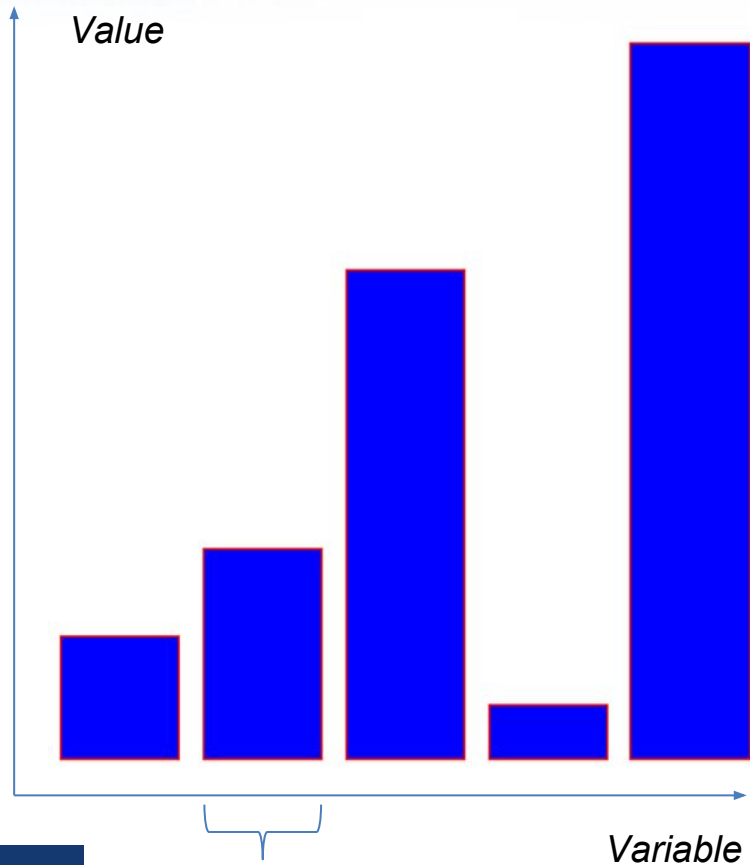


Cluster Analysis



Data Comparison

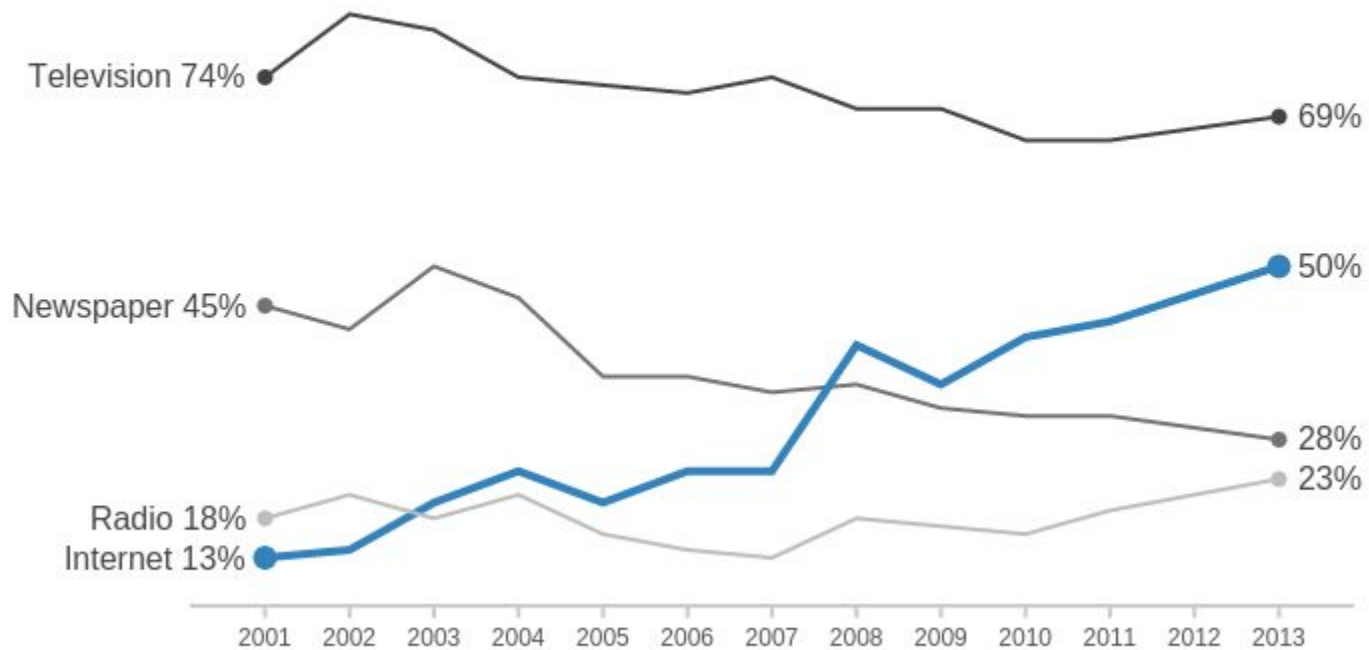
Bar Charts



Data Trends

Line Charts

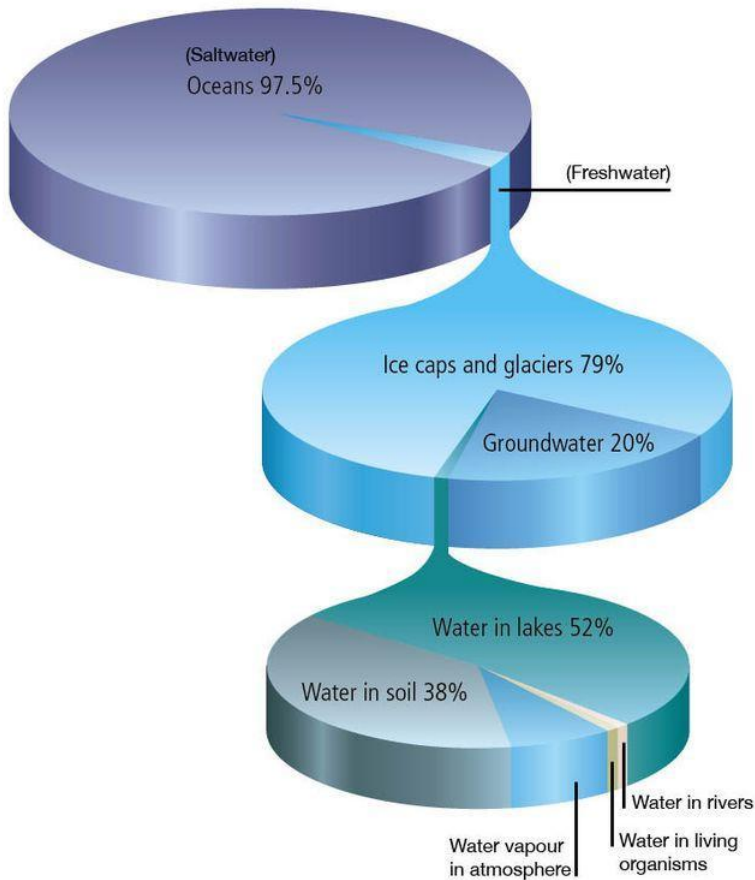
Main Source for News



Source: PewResearch Center & Storytelling with data

Relate Part to Whole

Pie Charts



2017 State Solar Power Rankings



How to read this chart:

This chart ranks the 50 states and the District of Columbia, from best (green) to worst (red), based on their solar-friendliness. For example, Massachusetts receives the best score, while Mississippi receives the worst.

The outermost ring (closest to each state label) reflects overall state rankings.

The inner rings represent factors contributing to each state's grade.

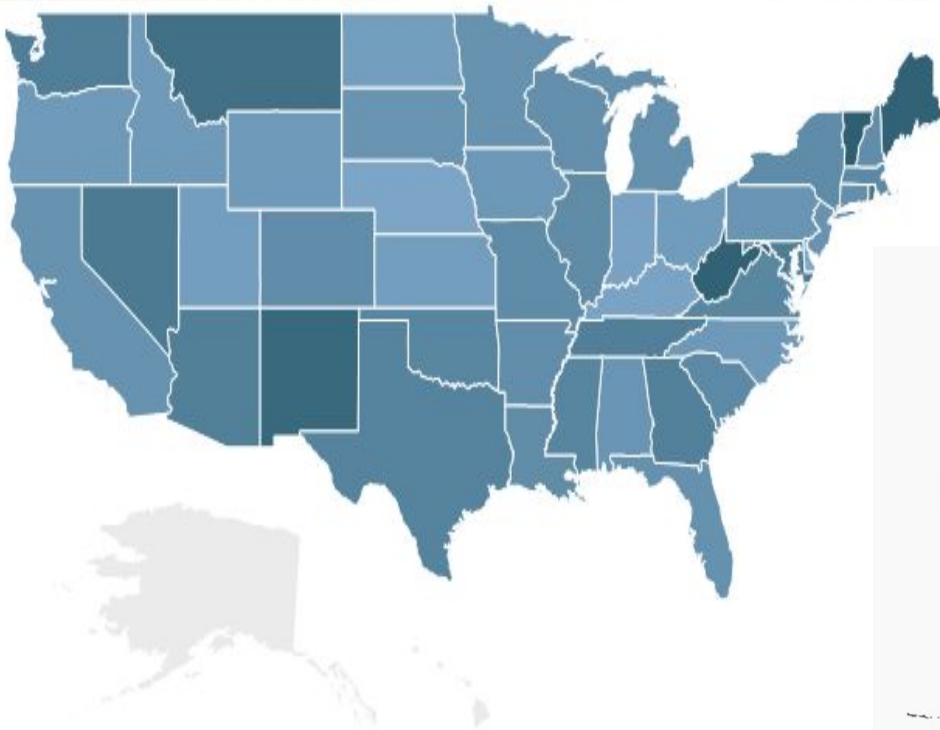
Factors:

1. Overall Grade
2. Renewable Portfolio Standard (RPS)
3. RPS Solar Carve-out
4. Electricity Cost
5. Net Metering
6. Interconnection
7. State Solar Tax Credit
8. State Solar Rebates
9. Performance Payments
10. Property Tax Exemption
11. Sales Tax Exemption

Grading Scale:



Maps



Heat Map
State borders polygons
Color density
proportional to value

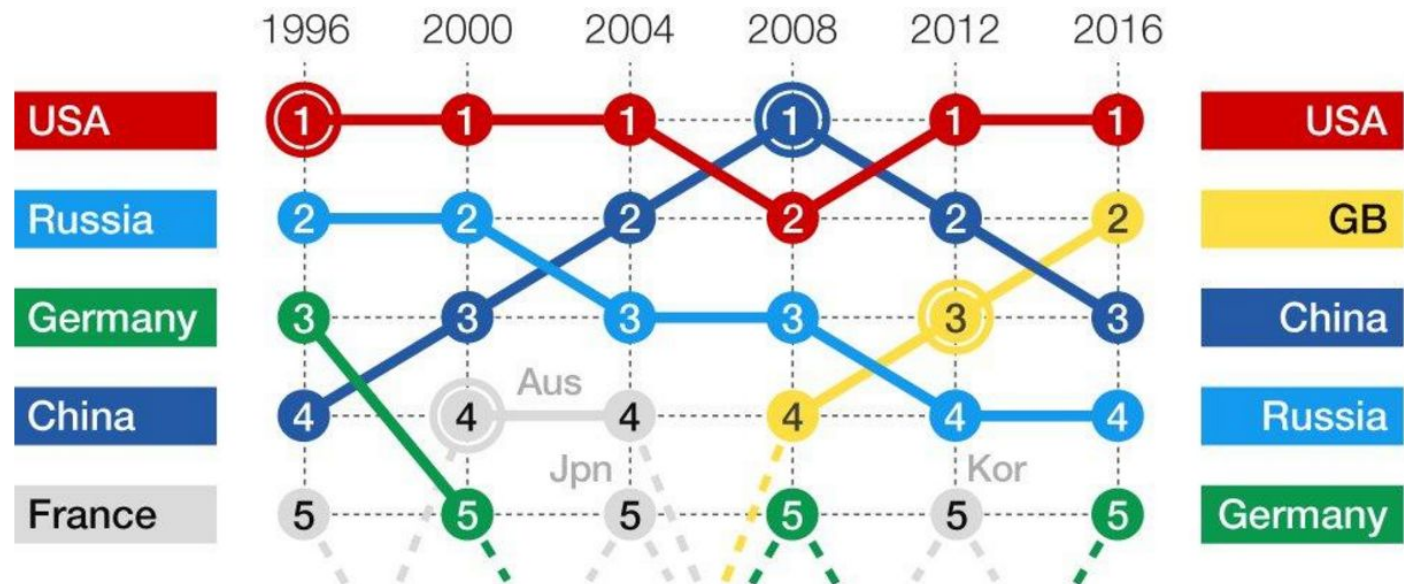
Bubble Markers
Pinpointed to coordinates
Size and density
proportional to values



Ranking Charts

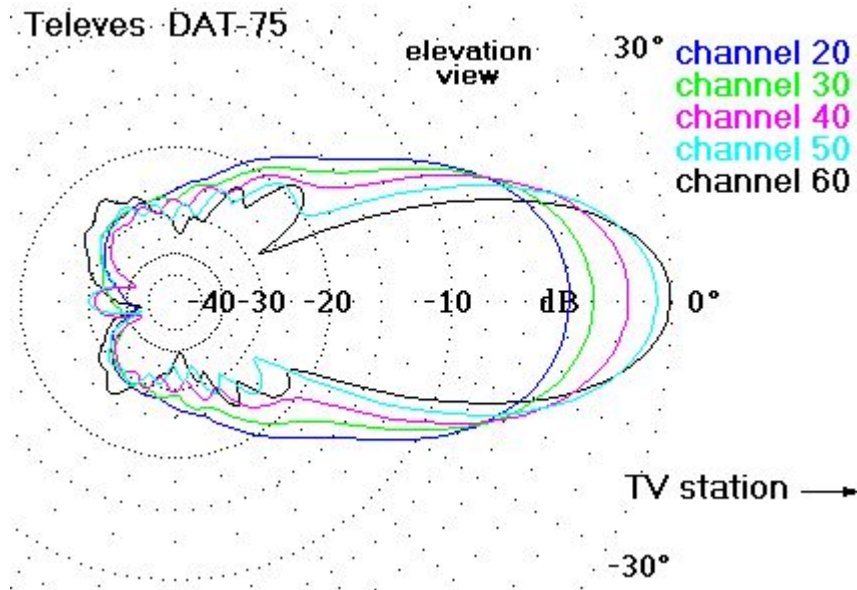
Olympic medal table: Top 5 rankings 1996–2016

○ = Host nation

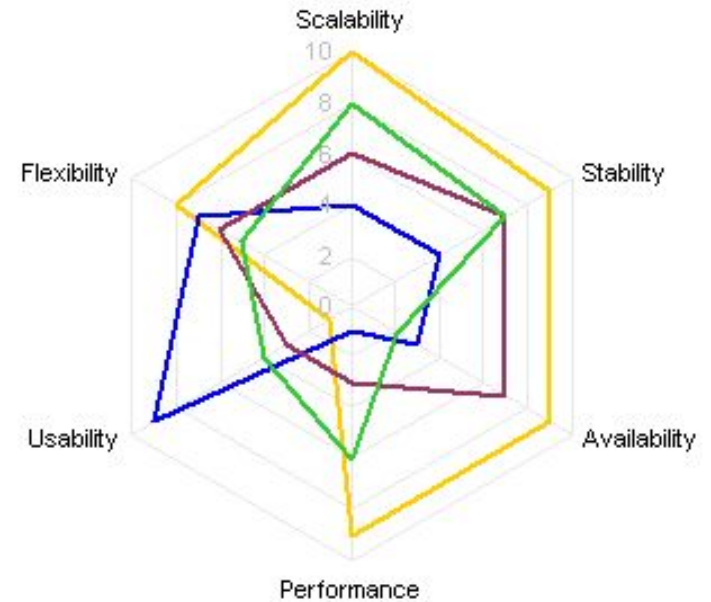


Radar Charts

*Antenna Diagram
Attenuation
Vs. Direction*

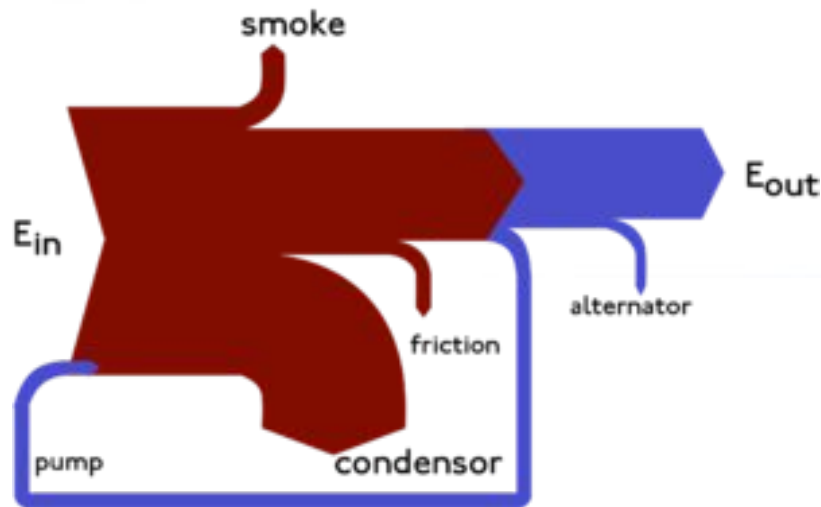


*Spider Diagram
Quick way to assess on
multiple parameters*

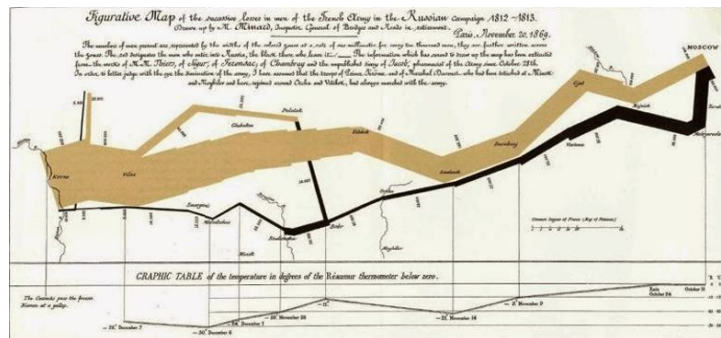
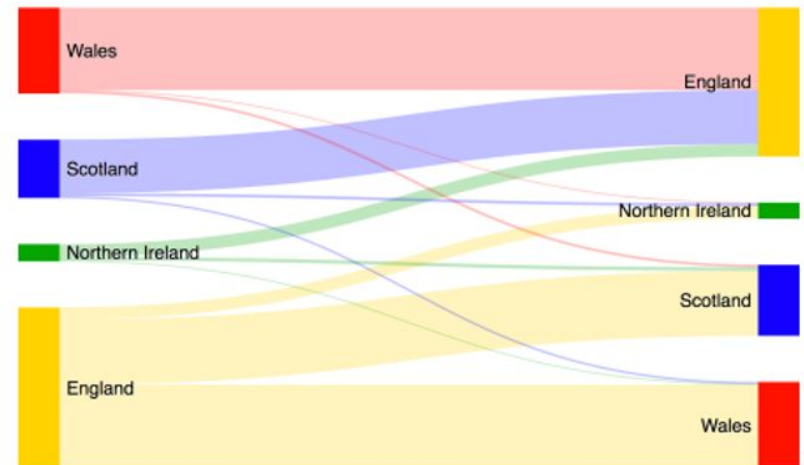


Sankey Diagram

Steam Engine
Diagram



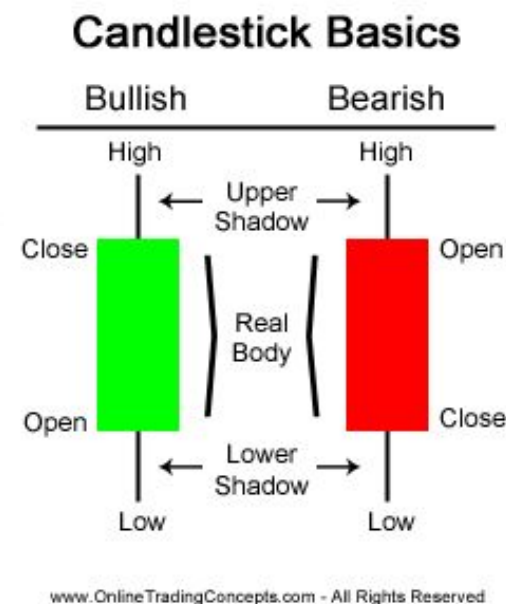
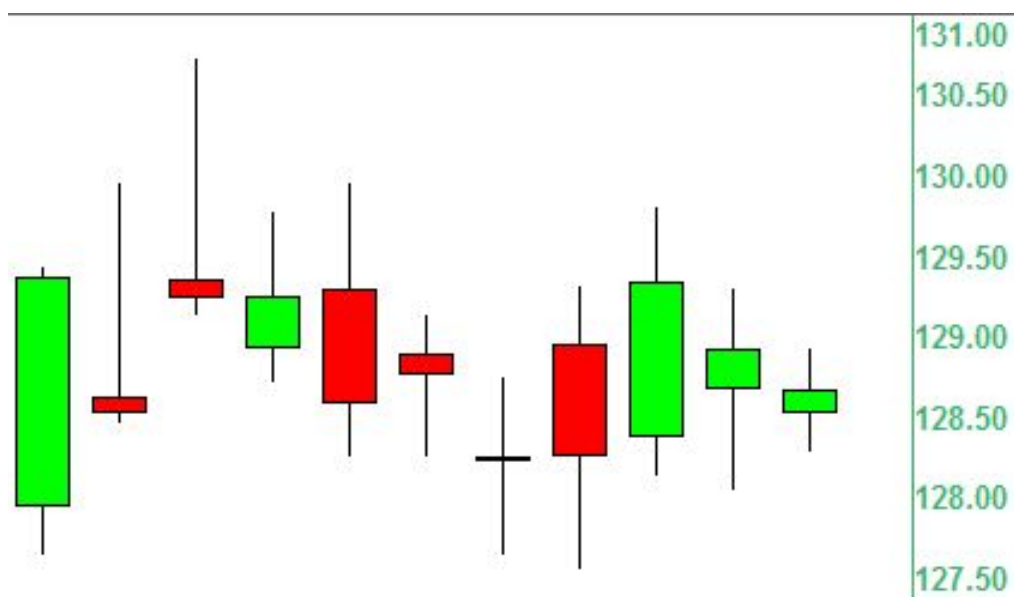
Migration Between
UK Countries



Remember
Charles Minard?

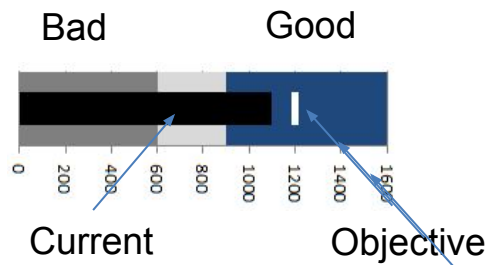
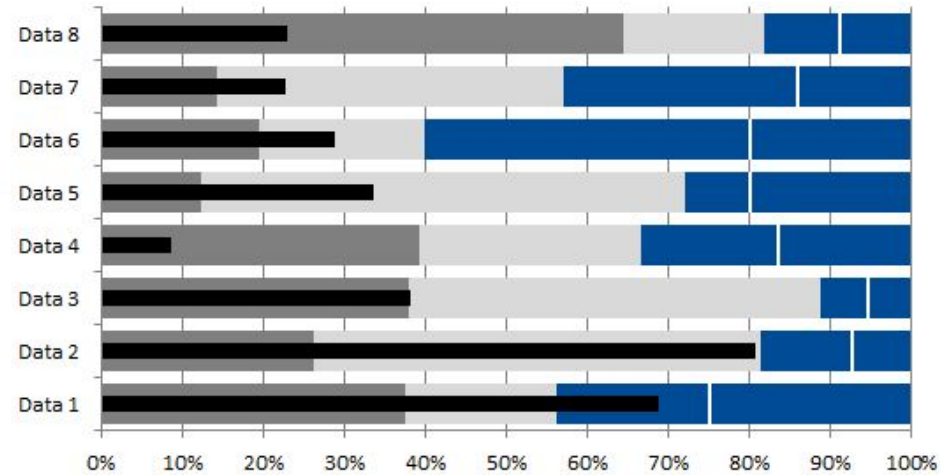
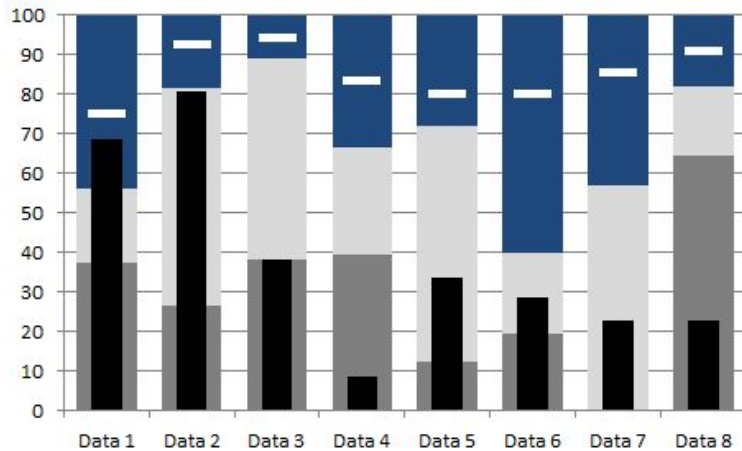
Special: Candlestick Diagram

(Shows Statistics: Range 1, Range 2, Gradient)

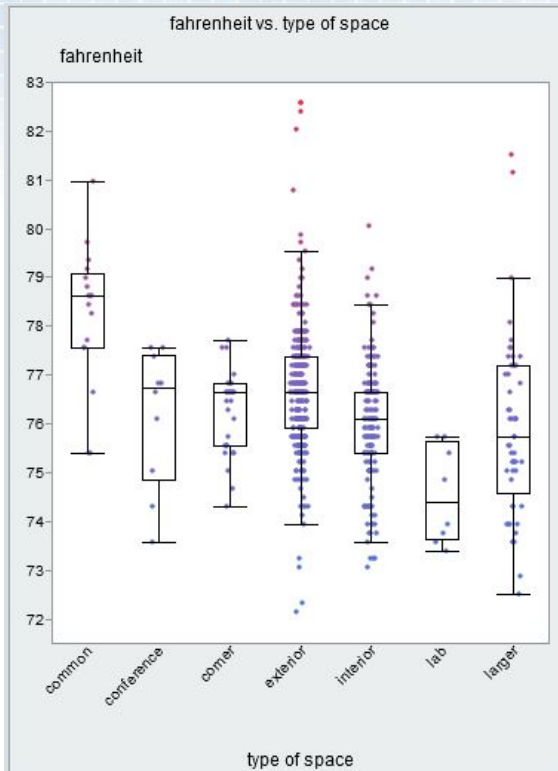


Special: Bullet Chart

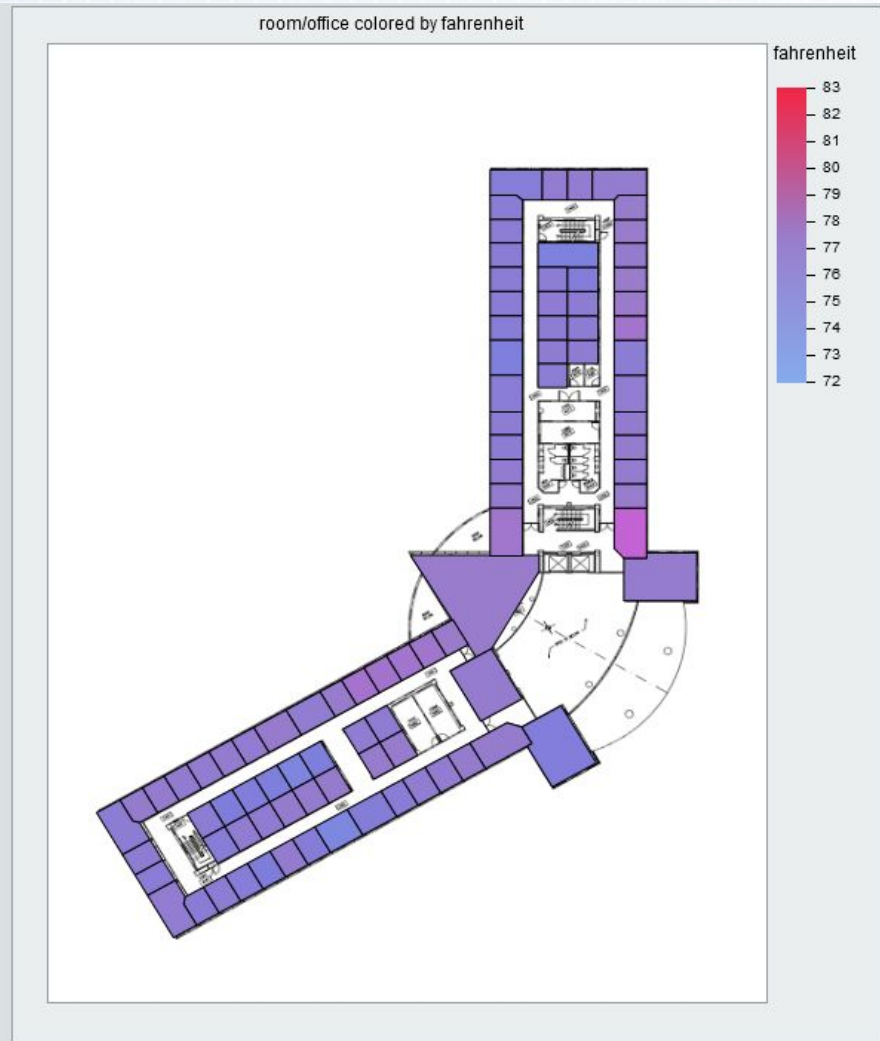
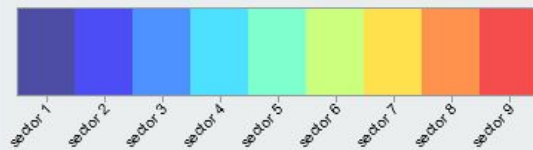
Compares current value to target, previous, and objective



Building Air Temperature

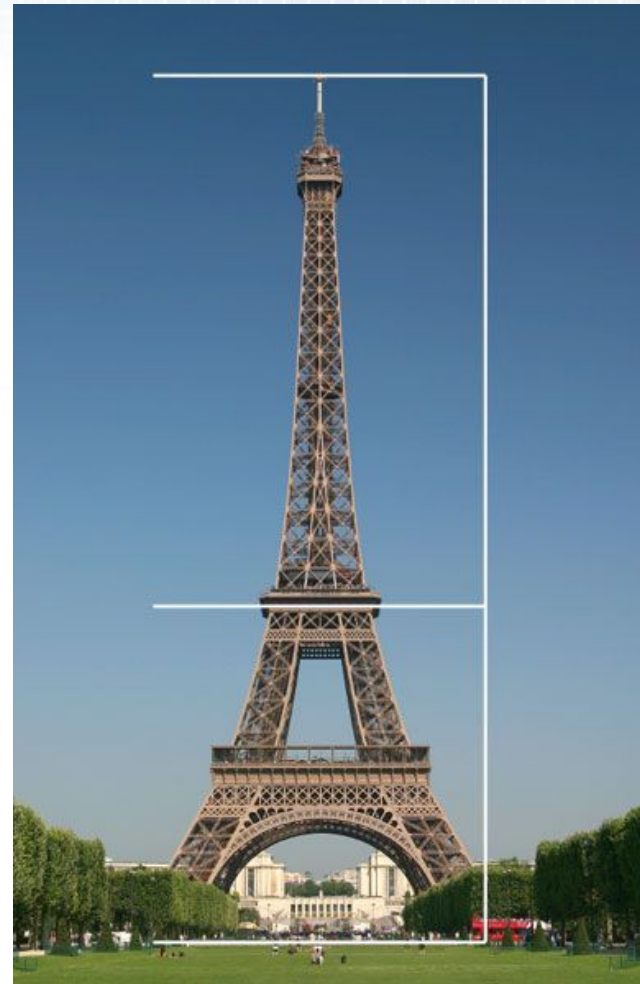


▼ Select Sector

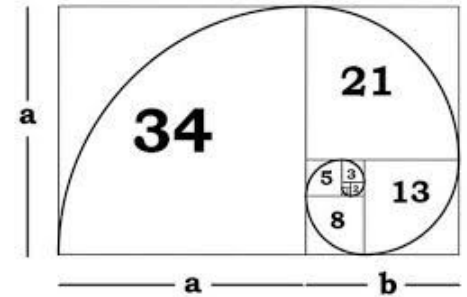


Composition Rules

- The visualization shall please the eye
- People spend more time with something they like
- People are more forgiving for what they like
- The eye likes
 - Symmetries, balance
 - Golden ratio ($\Phi = 1.618\dots$)

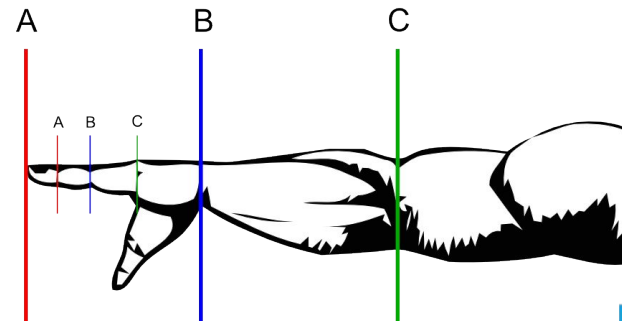


Golden Ratio (Fibonacci) Spiral (1,2,3,5,8,13,21...)



$$\text{Phi} = \frac{a+b}{a} = \frac{a}{b} = 1.618$$

Golden Ratio in Nature



Brain Wave Frequencies

The classical frequency bands of the EEG can indeed be described as a geometric series with a ratio between neighboring frequencies approximating $F = 1.618$

Frequency band		Frequency subband		Peak	Period
Name	[Hz]	name	[Hz]	[Hz]	[ms]
delta	1.5–4	delta1	1–2	1.5	667
		delta2	2–3	2.5	400
theta	4–10	theta1	3–5	4	250
		theta2	5–8	6.5	154
alpha	8–12	alpha	8–12	10	100
beta	10–30	beta1	12–20	16	62.5
		beta2	20–30	25	40
gamma	30–80	gamma1	30–50	40	25
		gamma2	50–80	65	15
fast	80–200	ripples1	80–120	100	10
ripples		ripples2	120–200	160	6.25

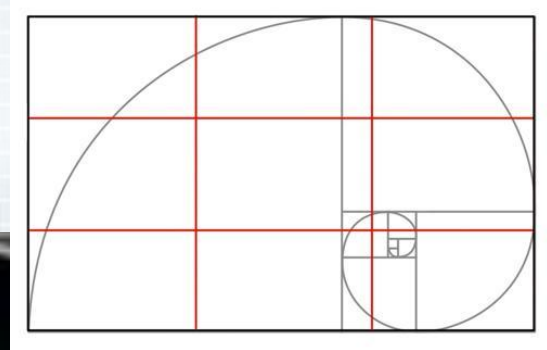
1,2,3,5,8 - Fibonacci Sequence => Golden Ratio = $n/(n-1) = 1.618$

Phi Grid



Rule of Thirds

Phi Grid for “Dummies”

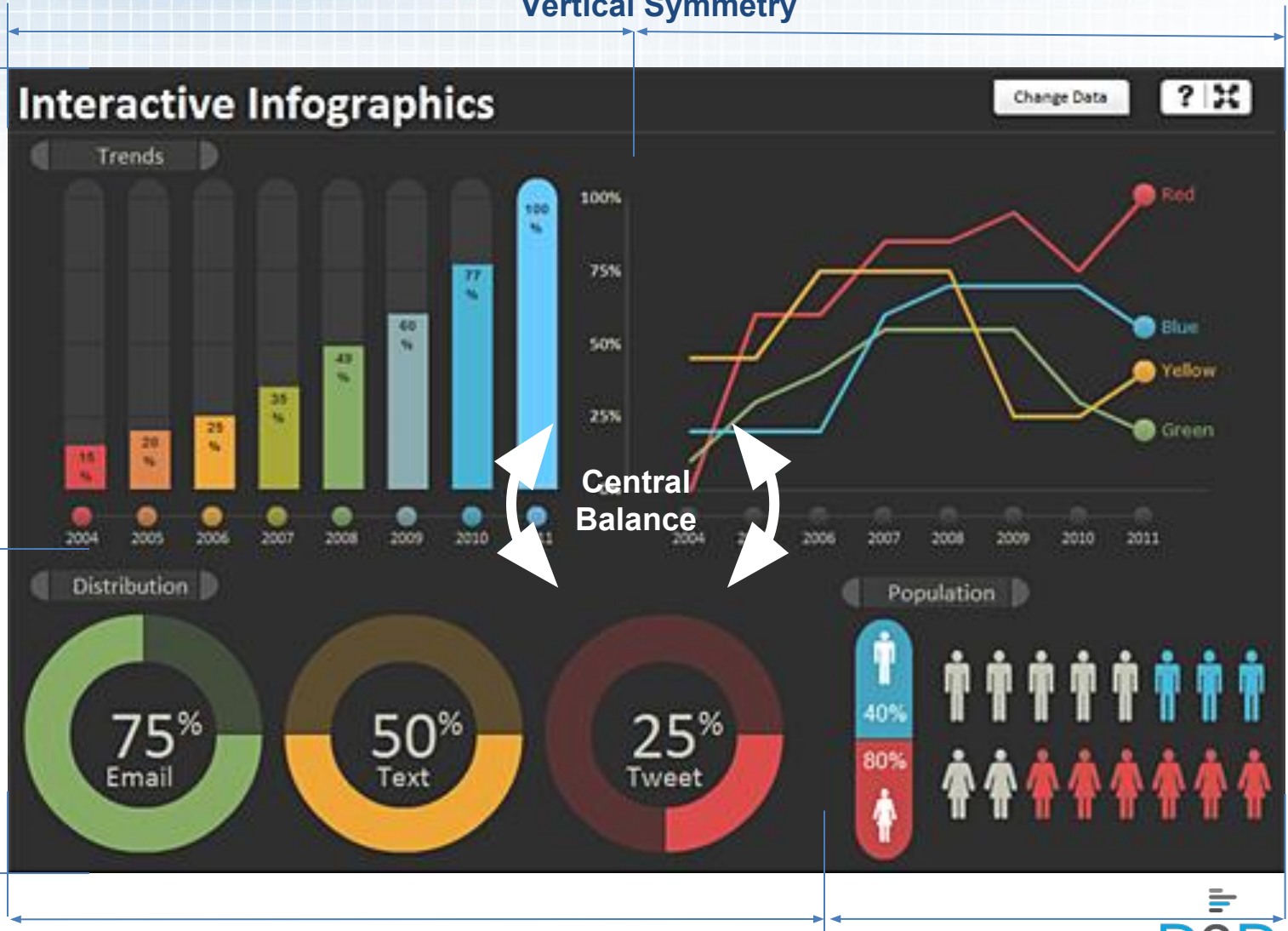


Dashboard Composition Example

Vertical Symmetry

Golden Ratio

Central Balance



GSA

Golden Ratio

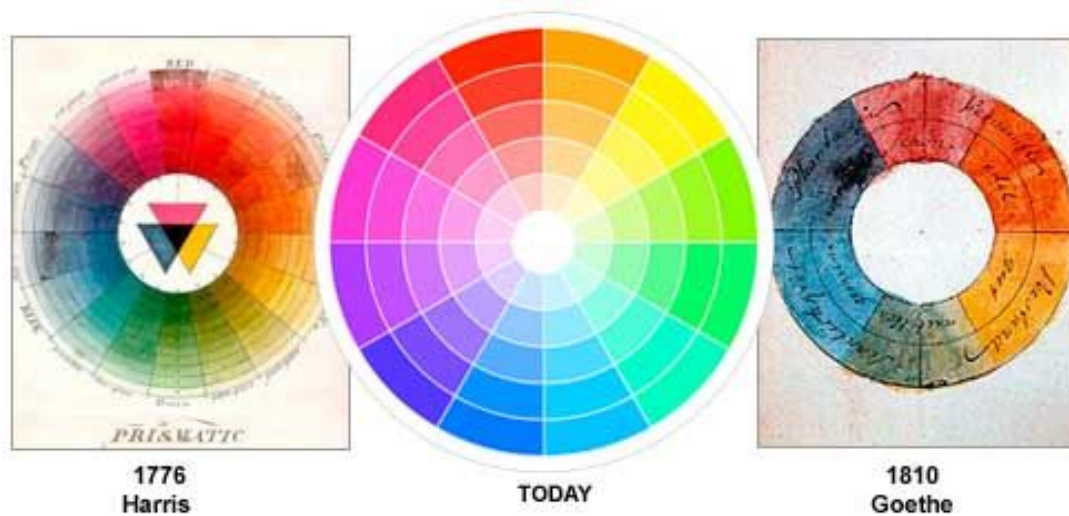
D2D
DATA TO DECISIONS

Colors

- **“...the first principle in bringing color to information: above all, do no harm.”**
(Envisioning Information, Edward Tufte, Graphics Press, 1990)
- **Poor choice of colors will obscure, muddle, confuse or irritate**
- **If possible, use color palettes provided by visualization tool**
- **If not, create customer-matching palettes**
 - **Use customer logos or webpages**
 - **Extract color codes**
 - **Use online tools**

Color Theory

- First color circle was developed by Isaac Newton in 1666
- Analogous colors are adjacent
- Complementary colors are opposite



Essential Color Terminology



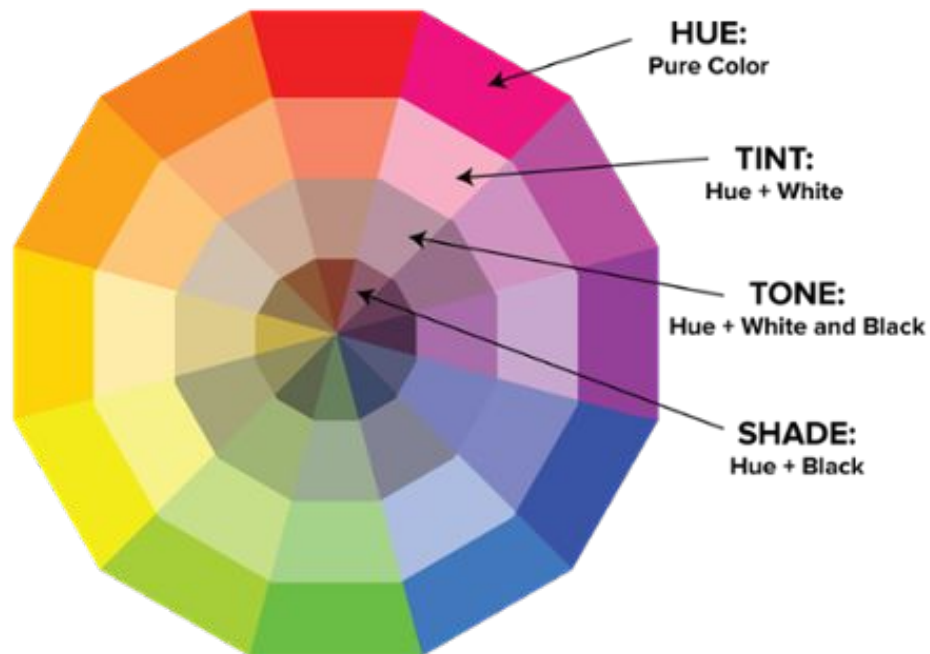
Primary Colors



Secondary Colors



Tertiary Colors



Why Printers and Monitors Use Different Primary Colors?

- You can get (almost) any color from either RGB or RYB (or other primary set)

Note: magenta, yellow, cyan cannot make black, hence there is black cartridge in printers

- You can mix colors by adding or subtracting (light)
- When you paint, you subtract light by adding another paint (reduce the amount of light going through – light is “subtracted”)
- Subtractive color wheel gives black when all colors are added
- Additive color wheel gives white when all colors are added

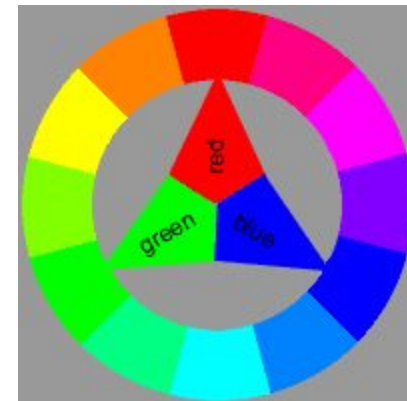
Color wheel
for painting



Color wheel
for printing



Color wheel
for projecting



Color Codes

Hex Color Codes - Websites

#FFFFFF

#000000

#**XXXXXX**

RGB Color Codes - Screens

255,255,255

0,0,0,

XXX,XXX,XXX

#9400D3

RGB 148, 0, 211

#4B0082

RGB 75, 0, 130

#0000FF

RGB 0, 0, 255

#00FF00

RGB 0, 255, 0

#FFFF00

RGB 255, 255, 0

#FF7F00

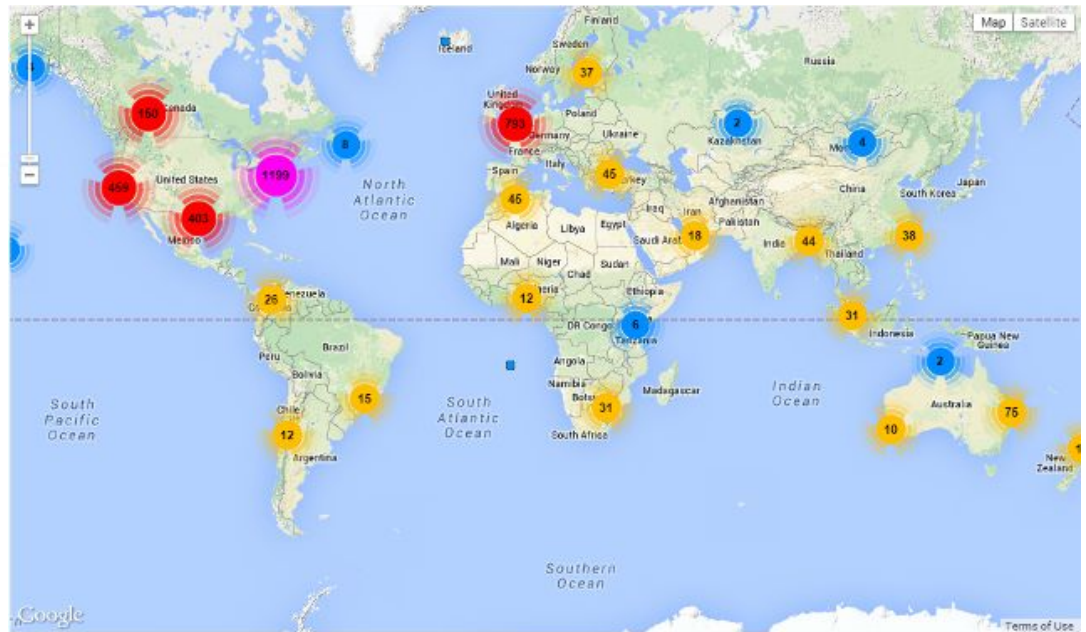
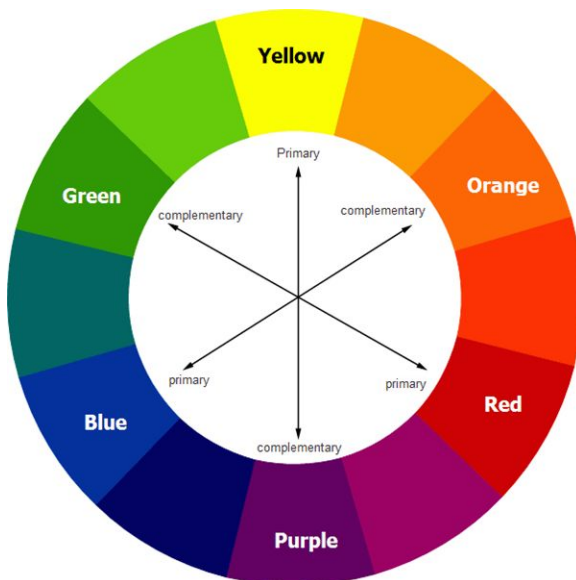
RGB 255, 127, 0

#FF0000

RGB 255, 0, 0

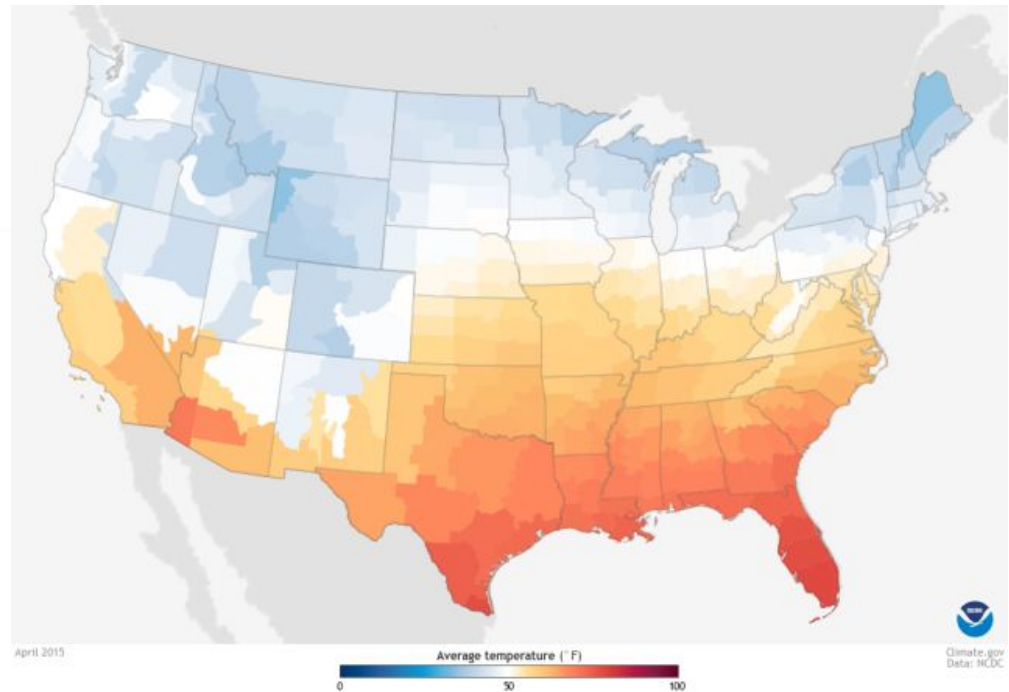
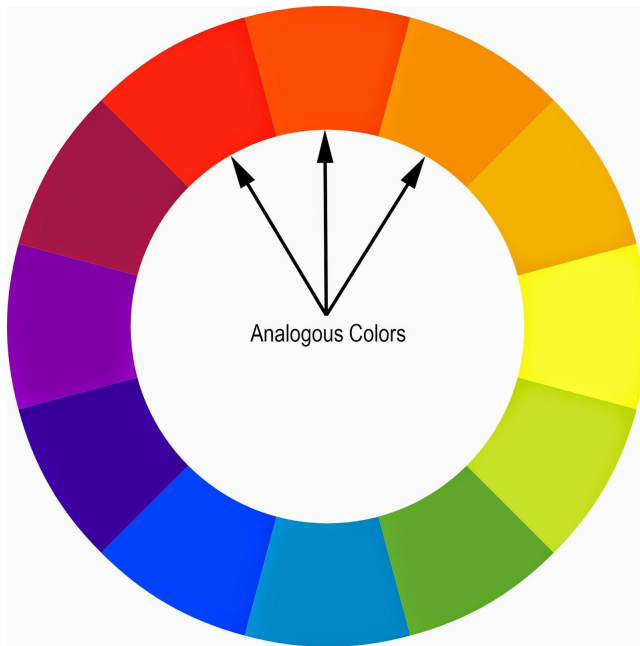
Contrast Colors

- Contrast colors draw attention
- Make things distinct
- Use color wheels to pick complementary contrast color



Analogous Colors

- Analogous colors suggest grouping
- Use them to show gradual changes




Match Customer Colors

- Use online tools to create customer-matching palettes
e.g.: <https://www.impalette.com/>

Sign Up

Sign In

HOMEEXAMPLESCONTACT

Enter a link to an image in this form to scan to find the most dominant Pantone colors (and their related colors) in your image.

Palette

Pantone Solid C

Link

or (max 4MB upload)

Choose File


No file chosen

Paired Color Offset Amount (1-160)

Brightness Range

Saturation Range

Scan



Elapsed Time: 1.519

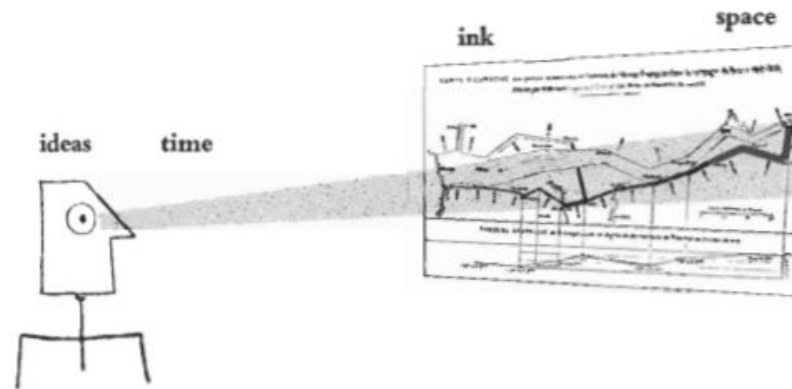
Wt.	Primary	Secondary	Tertiary
28	7527 C #d6d2c4	5807 C #d0d1ab	4745 C #cbb5a7
17	482 C #dbc8b6	616 C #d0c883	500 C #c6858f
13	4715 C #946c58	7497 C #7a7256	437 C #7b6469
8	7604 C #e4d5d3	468 C #ddcb4	7430 C #dca9bf
6	663 C #e5e1e6	677 C #e5cedb	2706 C #cbd3eb
6	Cool Gray 1 C #d9d9d6	Cool Gray 1 C #d9d9d6	Warm Gray 1 C #d7d2cb
6	479 C #a98066	7497 C #7a7256	7639 C #936d73
5	7615 C #866761	404 C #776e64	437 C #7b6469
4	4735 C #c9a78a	452 C #d1caa7e	5005 C #b07c83
2	621 C #d1e0d7	566 C #b9dcd2	344 C #a0dab3

Interactivity

- **Interactivity improves usability when**
 - **Intuitive**
 - **Forgiving (e.g., can get back to where you were)**
- **Most common interactivity features**
 - **Sort**
 - E.g.: descending, ascending, alphabetically, top 10, etc.
 - **Filter**
 - Exclude variables based on a parameter
 - **Drill down**
 - Path values from one visualization to another
 - Works best with established hierarchies

Edward Tufte's Principles of Graphical Excellence*

- **Complex ideas communicated with**
 - **Clarity**
 - **Precision**
 - **Efficiency**
- **Greatest number of ideas in shortest time with least ink in the smallest space**



Edward Tufte's Visualization Guidelines*

- **Show the data**
- **Induce the viewer to think over the substance, not graphic design, technology, etc.**
- **Do not distort what the data have to say**
- **Present many numbers in small space**
- **Make large data sets coherent**
- **Encourage the eye to compare different pieces of data**
- **Reveal the data at several layers of detail, from broad to fine**
- **Serve a clear purpose: description, explorations, tabulation, or decoration**
- **Be closely integrated with statistical and verbal descriptions of a data set**

MicroStrategy Examples

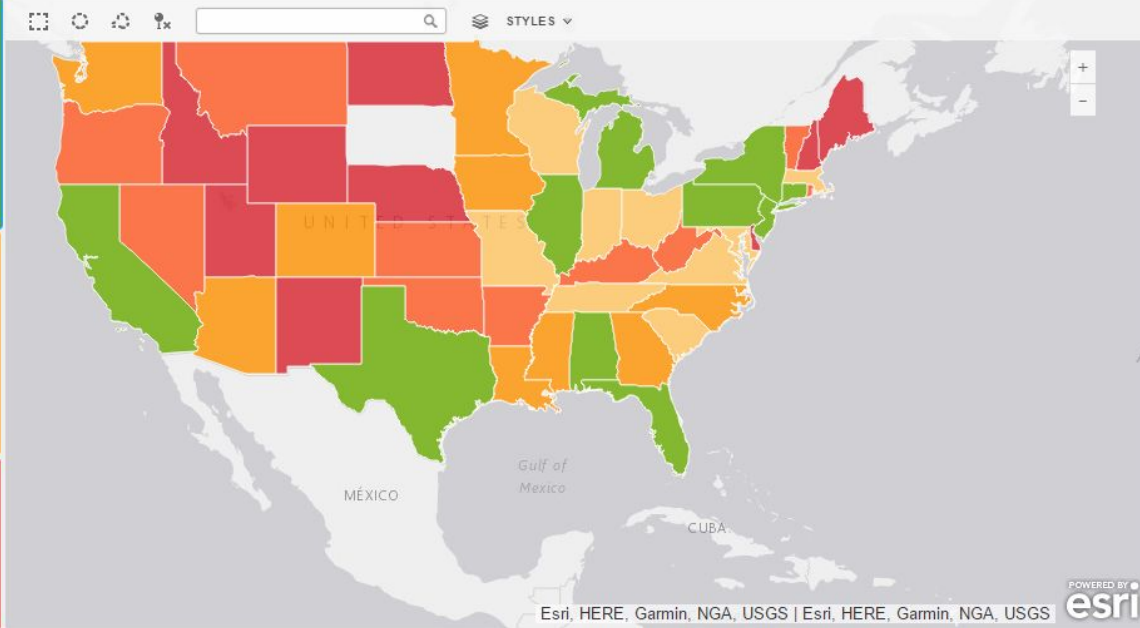
- **MicroStrategy Desktop is free**
- **Create data extracts in R**
- **Add data to MSTR desktop**
- **Create dashboards**
- **Save, and export as PDF**
- **<https://www.microstrategy.com/us/desktop>**

Sales Overview

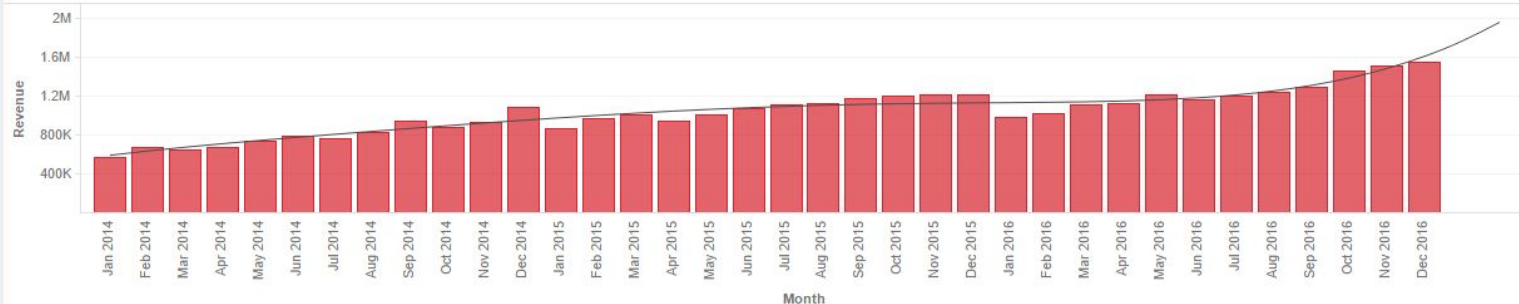
Revenue
\$37,253,703

Units Sold
870233

Cost
\$30,806,716



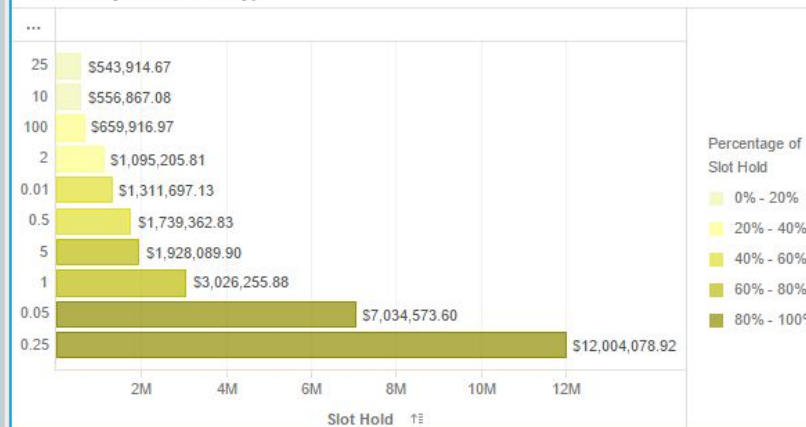
Revenue Over Time



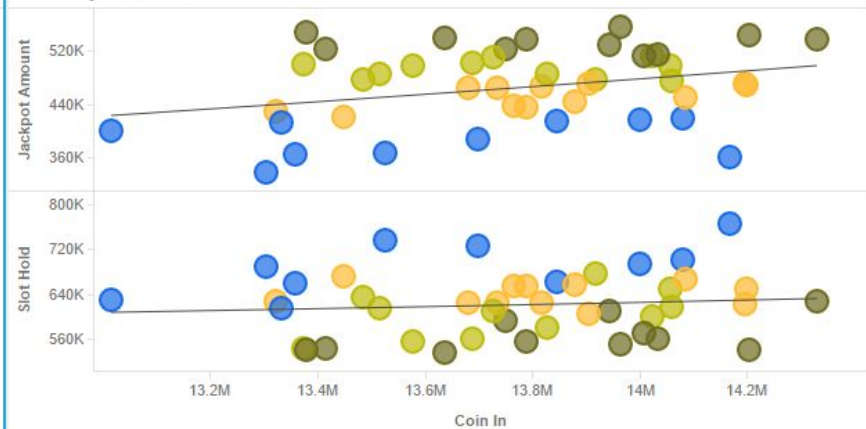
Sales ▾

+

Slot Hold by Slot Machine Type



Key Performance Indicator Review



Slot Machine Daily Activity



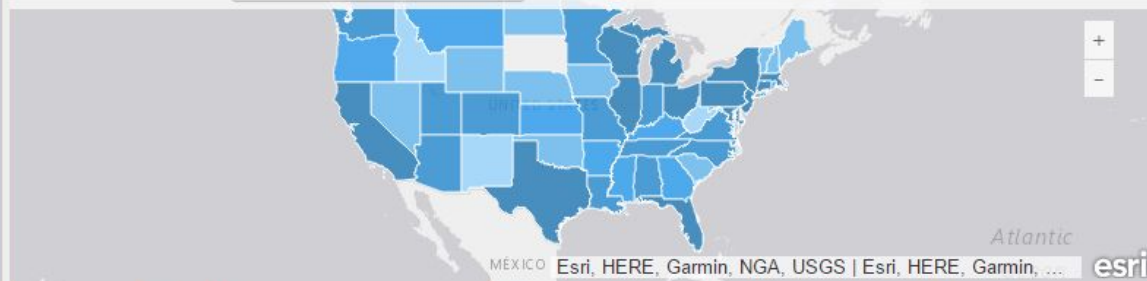
Product Sales Analysis

(All)



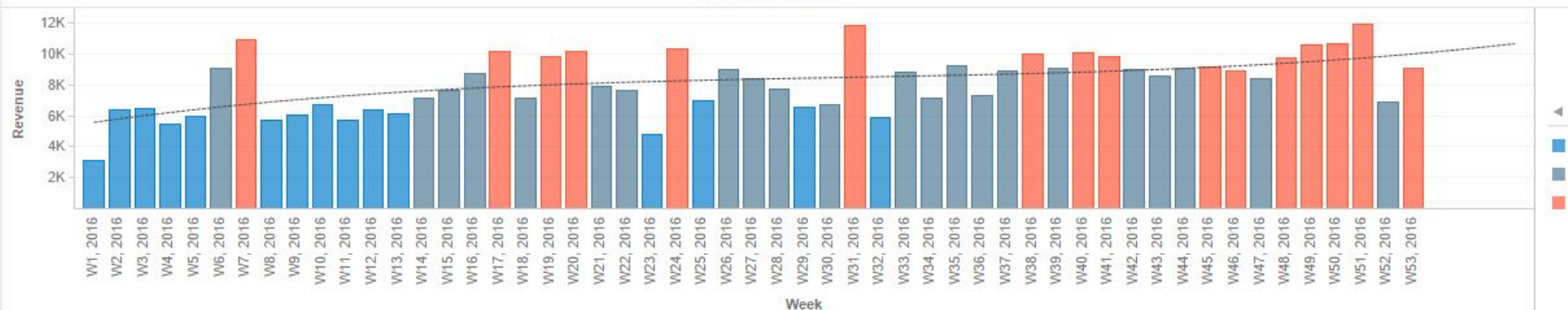
Search

STYLES

**Profit**
\$67,619**Revenue**
\$430,469**Cost**
\$362,850

Category	Books	Electronics	Movies	Music
Day	Revenue	Revenue	Revenue	Revenue
12/31/2016	\$69	\$1,338	\$284	\$178
12/30/2016	\$131	\$360	\$214	\$195
12/29/2016	\$142	\$50	\$132	\$229
12/28/2016	\$112	\$1,012	\$67	\$173
12/27/2016	\$126	\$2,144	\$179	\$183
12/26/2016	\$60	\$938	\$176	\$93
12/25/2016	\$145		\$125	\$193
12/24/2016	\$127	\$1,665	\$76	\$141
12/23/2016	\$49	\$80	\$167	\$200
12/22/2016	\$34	\$68	\$80	\$28
12/21/2016	\$34	\$453	\$27	\$130
12/20/2016	\$69	\$725	\$158	\$121
12/19/2016	\$102	\$1,046	\$270	\$174
12/18/2016	\$85	\$296	\$240	\$202
12/17/2016	\$89	\$1,580	\$207	\$123
12/16/2016	\$176	\$1,000	\$189	\$177

Revenue over Time



Sales

Future Sales Projections

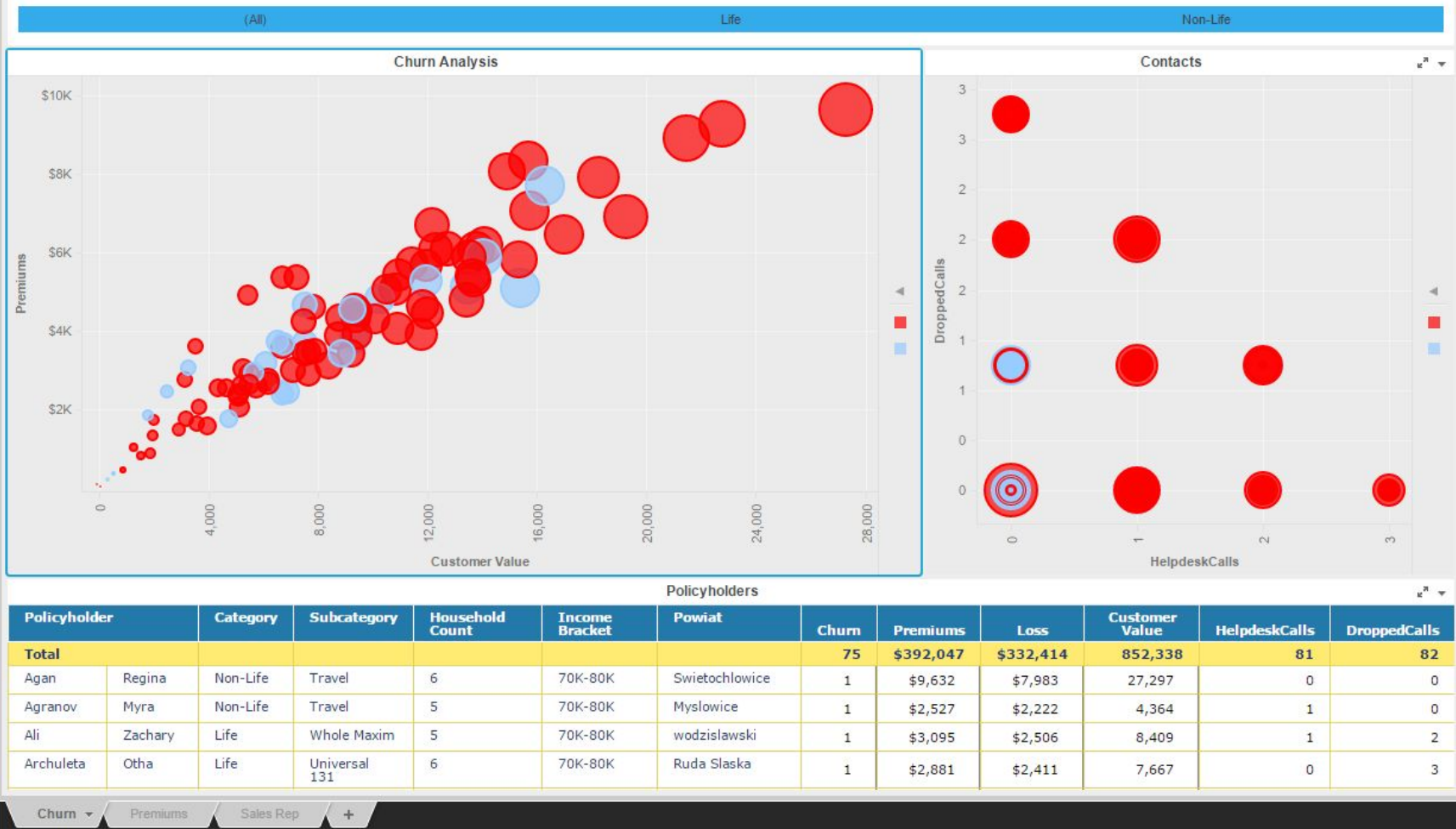
Customer Segmentation

Category and Product Popularity

Media Center

+

GSA



Q & A