

DATA VISUALIZATION AND VISUAL ANALYTICS

S. Rinzivillo – rinzivillo@isti.cnr.it

TAXONOMY OF VISUAL VARIABLES

CLEVELAND MCGILL [1984]

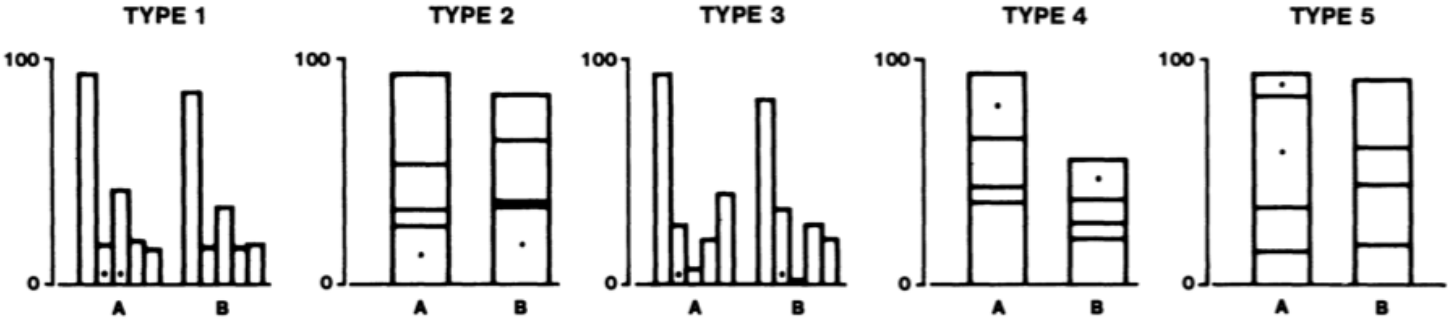


Figure 4. Graphs from position-length experiment.

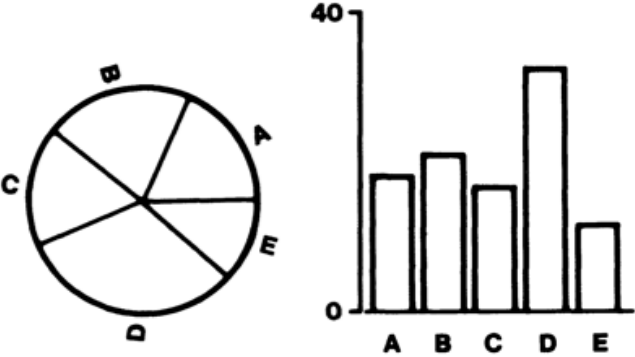
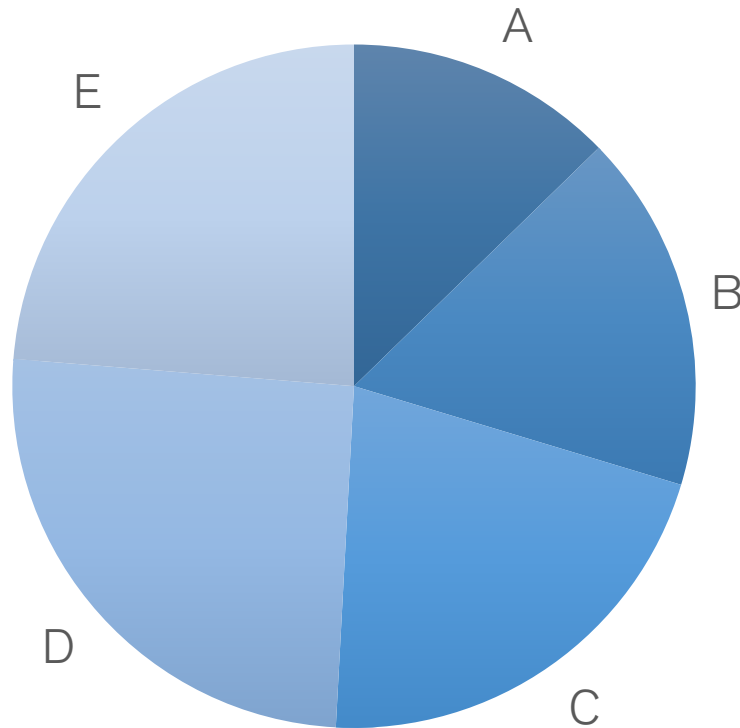


Figure 3. Graphs from position-angle experiment.

CLEVELAND & MCGILL: GRAPHICAL ENCODINGS

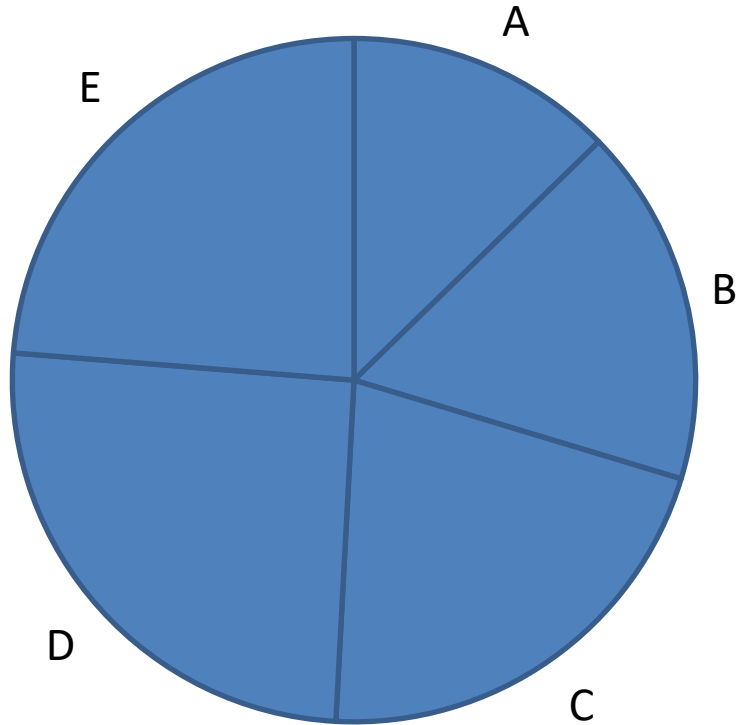
- Angle
- Area
- Color Hue
- Color Saturation
- Density
- Length
- Position on a common scale
- Position on non aligned scale
- Slope
- Volume

ANGLE DECODING



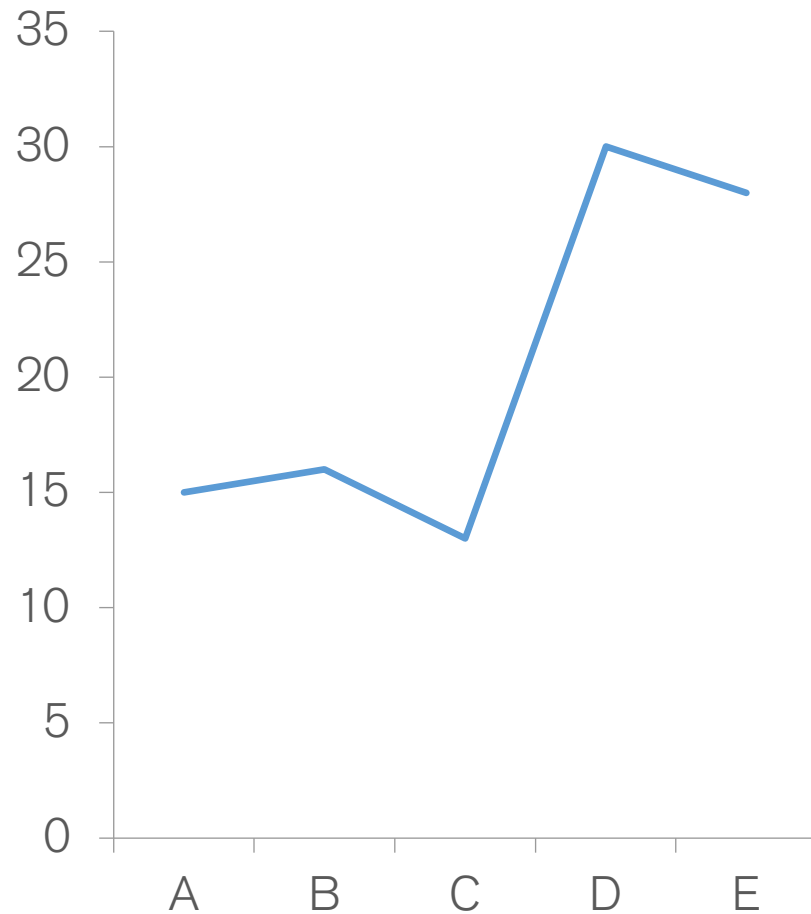
- It is difficult to compare angles
 - Underestimation of acute angles
 - Overestimation of obtuse angles
 - Easier if bisectors are aligned
- Area estimation helps

ANGLE DECODING



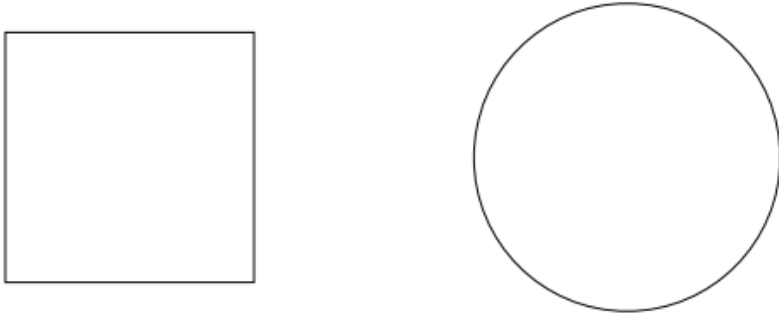
- It is difficult to compare angles
 - Underestimation of acute angles
 - Overestimation of obtuse angles
 - Easier if bisectors are aligned

SLOPES DECODING



- Same difficulties as angles
- Easier task since one branch is aligned with x-axis

AREA DECODING

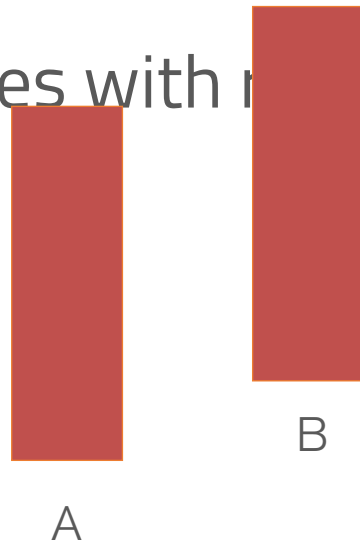


- Area is not well decoded
 - Different regular shapes
 - Irregular shapes
 - Context influences (thin area within compact thick area)

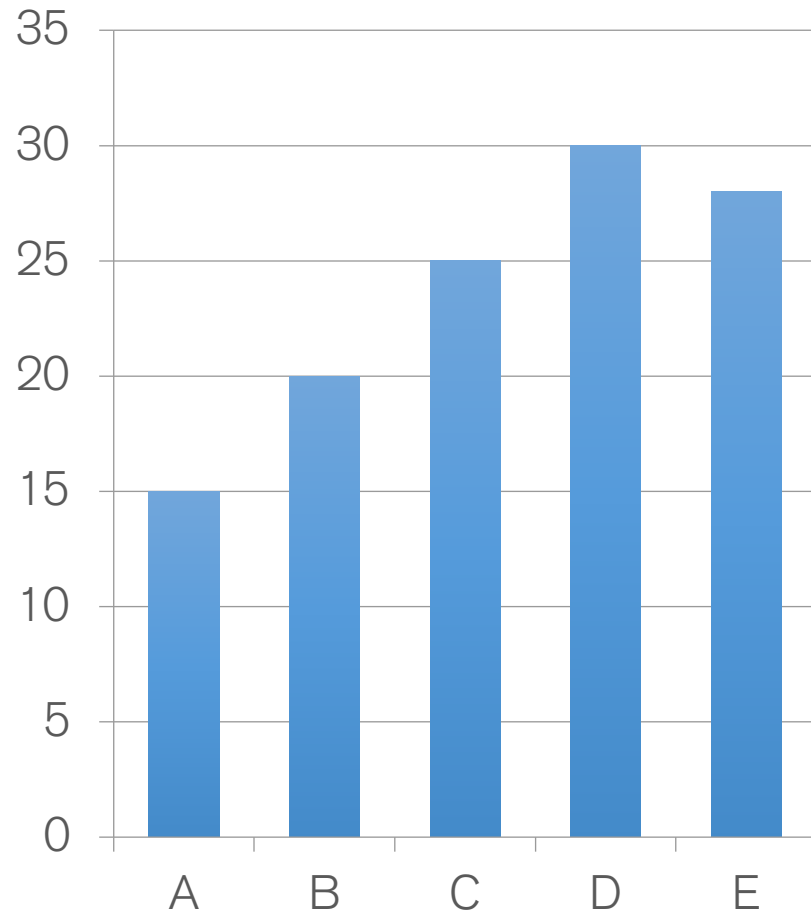


LENGTH DECODING

- Straight forward to encode numerical values
- Difficulties with relative lengths



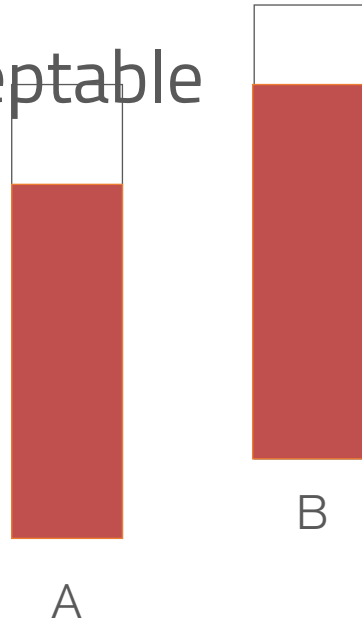
POSITION ON A COMMON SCALE



- Widely used in statistical charts

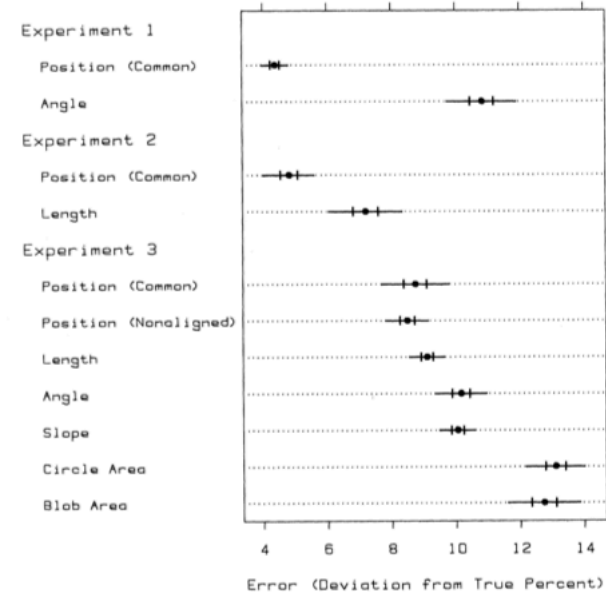
POSITION ON NON-ALIGNED SCALE

- Not as bas as common scale
- Still acceptable



DESIGNING EFFECTIVE VISUALIZATIONS

- If possible, use graphical encoding that are easily decoded
- Graphical Attributes ordered(Cleveland & McGill):
 - Position along a common scale
 - Position on non aligned scales
 - Length
 - Angle and Slope
 - Area
 - Volume, density, color saturation
 - Color Hue



Most Efficient

t



Least Efficient

t

Position



Length



Slope



Angle



Area



Intensity



Color



Shape



Quantitative

Ordinal

Nominal

© Microsoft

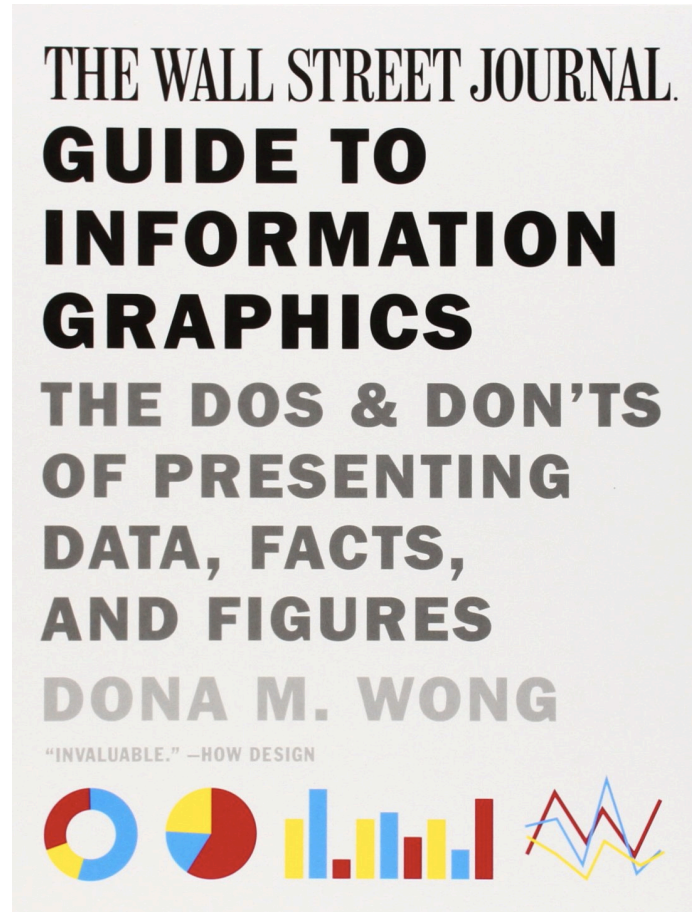
TAKEAWAY MESSAGES

- Data type for entities and relationships
- Visual variables for representation
- Mapping of types to VVs
- Some VVs are more appropriate for specific data types

6

VISUAL ANALYTICS DOS AND DON'TS FOR VISUAL CHARTS

CRASH COURSE ON EFFECTIVE CHARTING



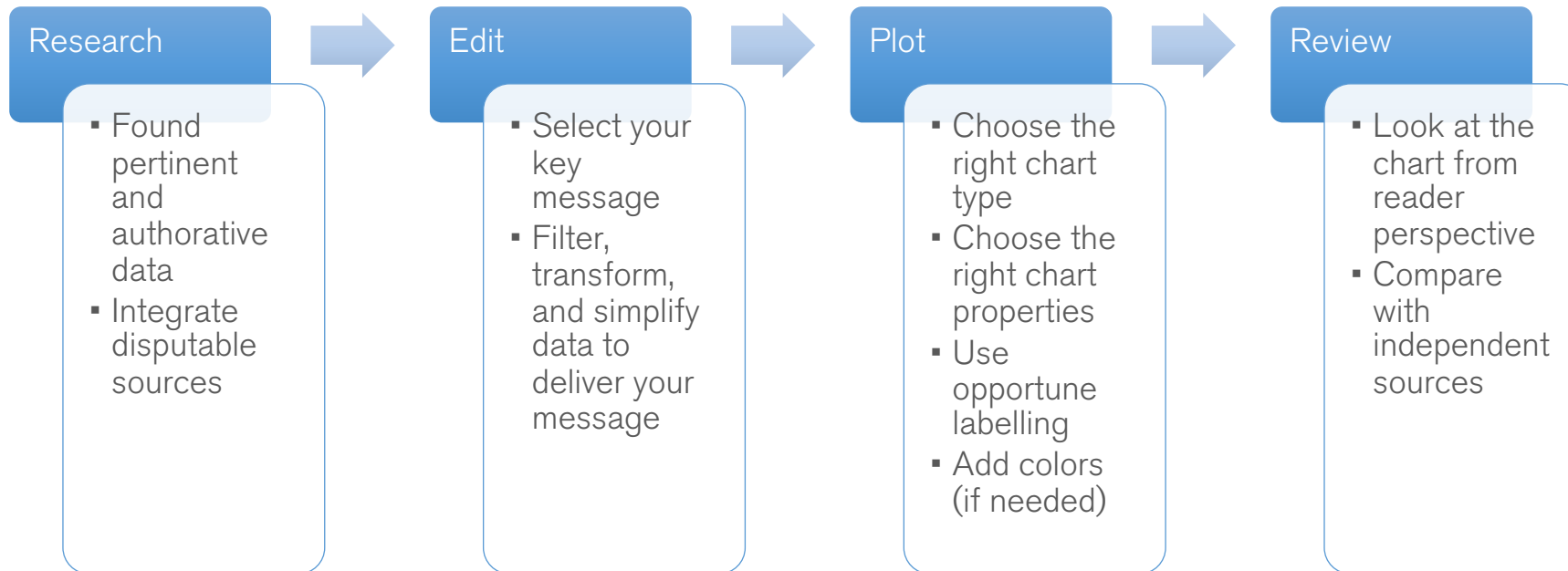
Dona M. Wong

Guide to Information Graphics

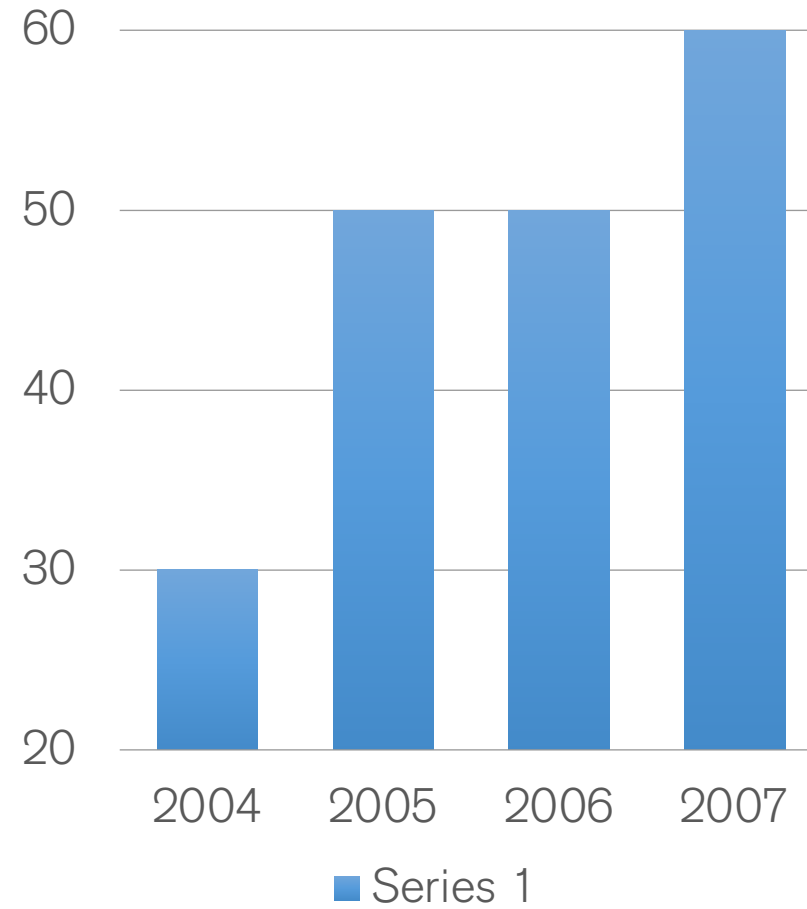
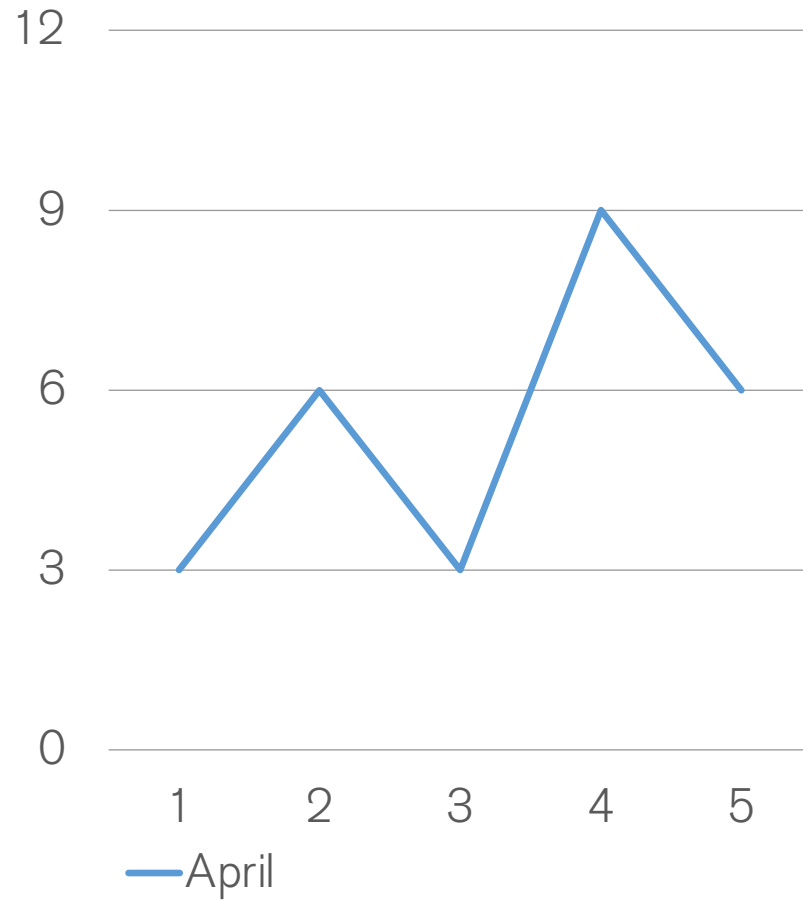
The Dos and Don'ts of Presenting Data, Facts, and Figures

W. W. Norton & Company

CHARTING PIPELINE



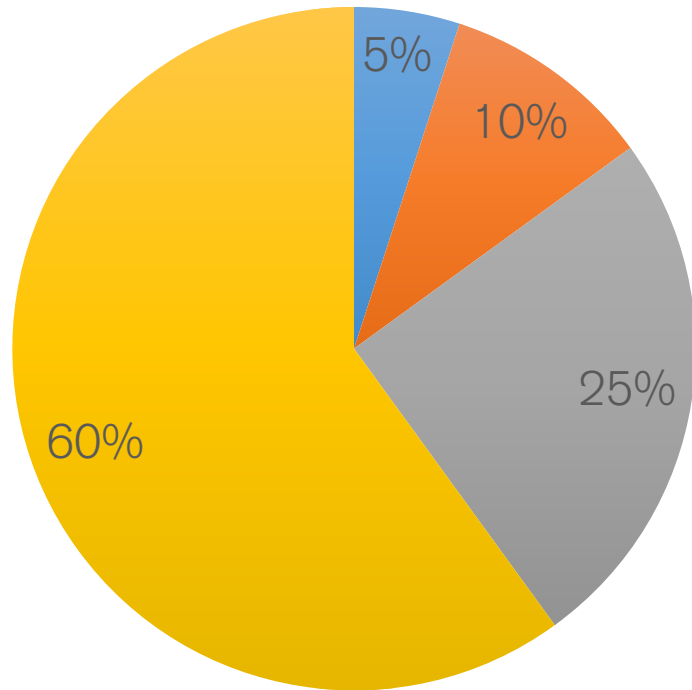
CHARTING EXAMPLES



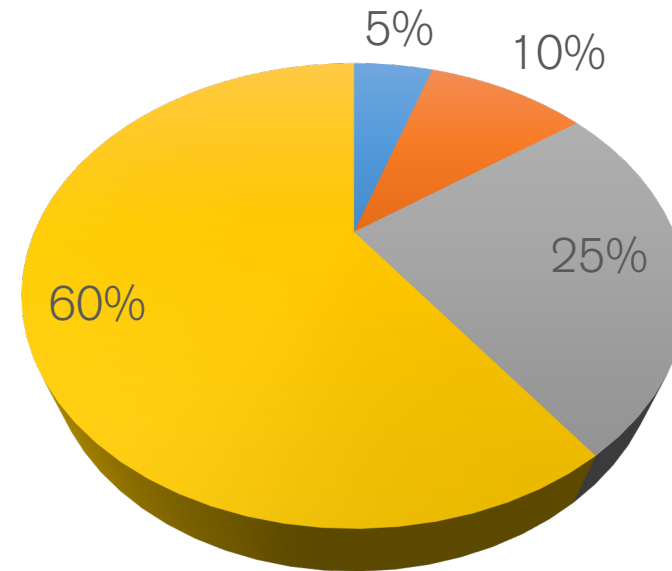
May these charts be improved? Why? How?

CHARTING EXAMPLES

Sales

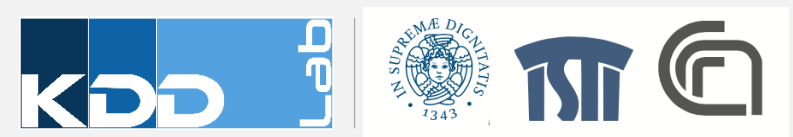


Sales

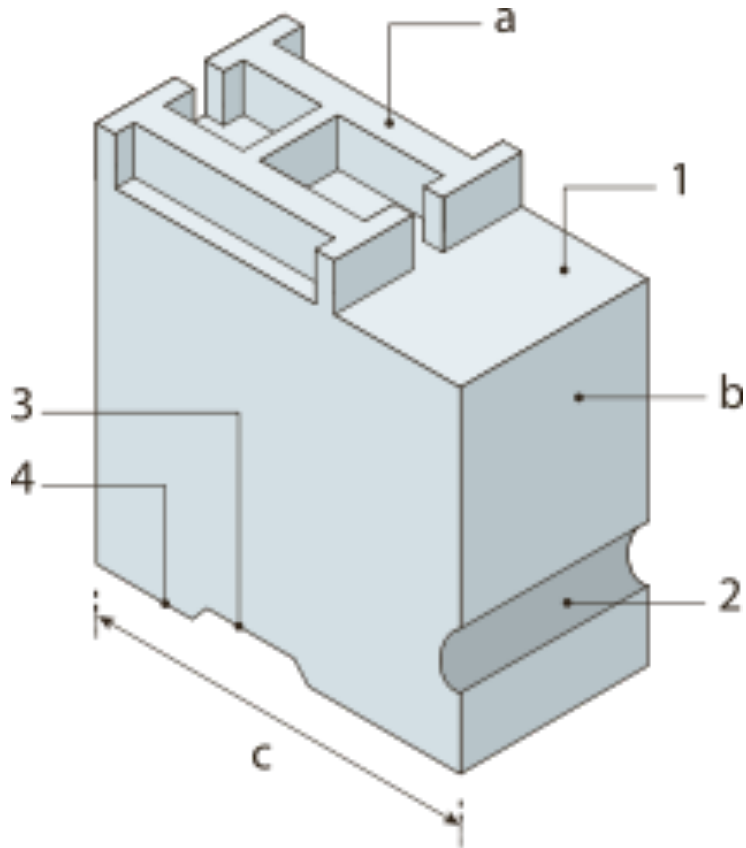


May these charts be improved? Why? How?

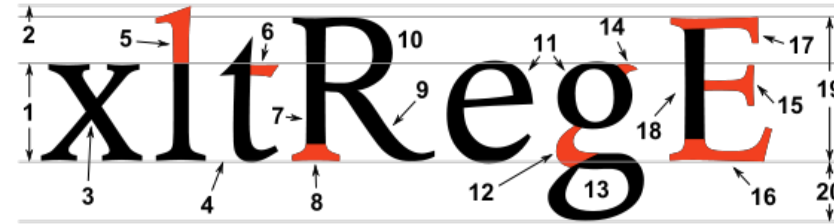
FONTS



FONTS



"Metal type". Licensed under Public Domain via Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Metal_type.svg#mediaviewer/File:Metal_type.svg



Typographic parts of a glyph:

1) x-height; 2) ascender line; 3) apex; 4) baseline; 5) ascender; 6) crossbar; 7) stem; 8) serif; 9) leg; 10) bowl; 11) counter; 12) collar; 13) loop; 14) ear; 15) tie; 16) horizontal bar; 17) arm; 18) vertical bar; 19) cap height; 20) descender line.

$$\begin{aligned} \text{Font size} &= (1) + (2) + (20) \\ &= (19) + (20) \end{aligned}$$

"Typoghaphia" by F l a n k e r (typographic font designed by myself, named Emperor). Licensed under Public Domain via Wikimedia Commons - <http://commons.wikimedia.org/wiki/File:Typoghaphia.svg#mediaviewer/File:Typoghaphia.svg>

FONTS: GENERAL RULES

- Leading should be 2 points larger than type size
- Avoid too small or condensed type faces
- Keep style simple: use **bold** or *italic* to emphasize a word (better not **both**)
- Avoid ALL CAPS
- Avoid *styled fonts*
- Avoid C***C Sans Serif
- Reduce type at an angle
- Avoid t r a c k i n g

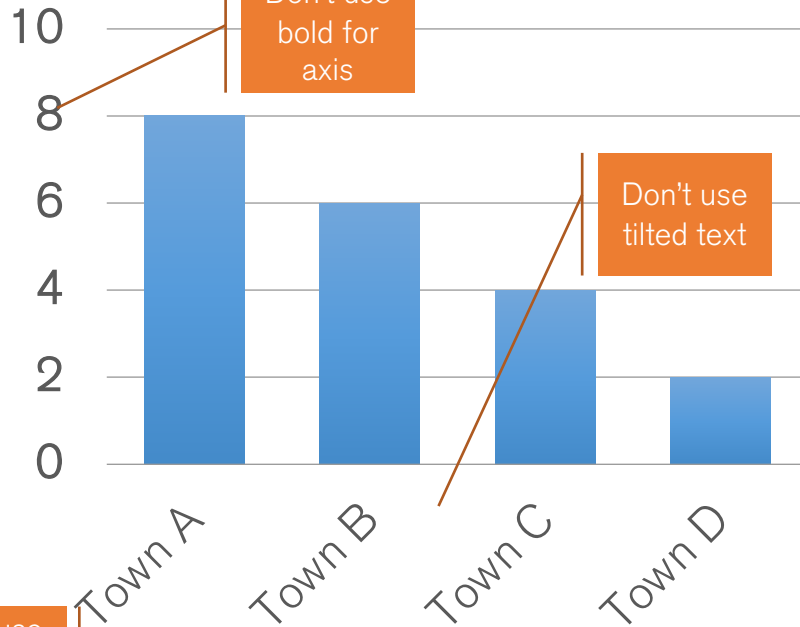
Fonts are meant to describe, not to adorn

TYPOGRAPHY IN CHARTS

Don't use all caps or high contrast white type out of black

Don't

HEADLINE OF THE CHART



Don't use bold for axis

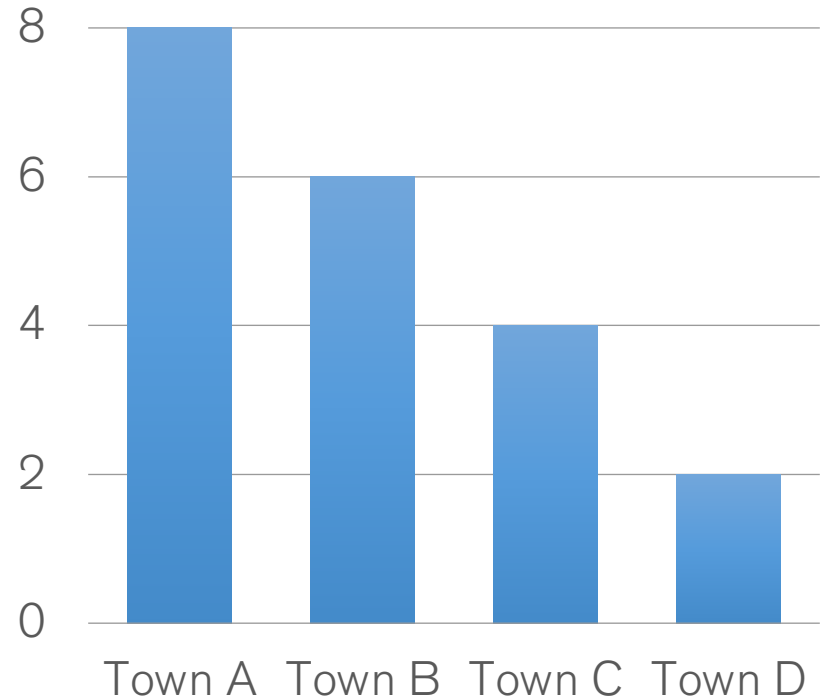
Don't use tilted text

Don't use bold and italic

A brief description that outlines what the data shows

Do

Headline of the chart

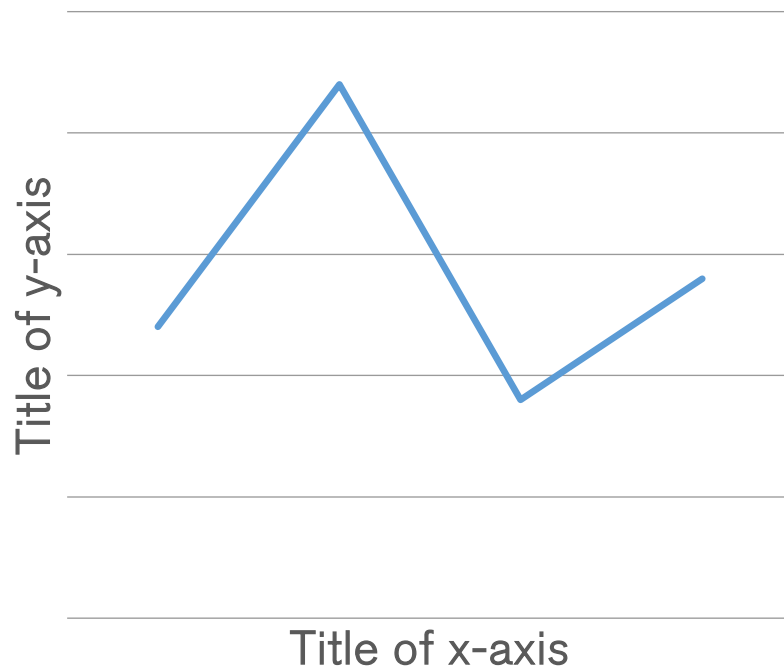


A brief description that outlines what the data shows

TYPOGRAPHY IN CHARTS

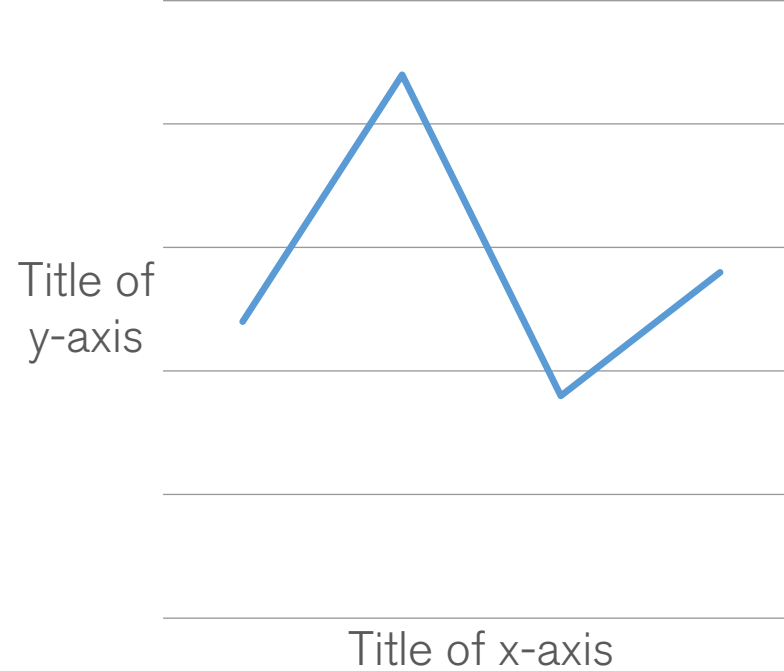
Don't

Headline of the chart



Do

Headline of the chart



TYPOGRAPHY IN CHARTS

Name	Data	Data	Data
Company A	0.0	0.0	0.0
Company B	0.0	0.0	0.0
Company C	0.0	0.0	0.0
Company D	0.0	0.0	0.0

Many elements in bold. Which part is highlighted?

Name	Data	Data	Data
Company A	0.0	0.0	0.0
Company B	0.0	0.0	0.0
Company C	0.0	0.0	0.0
Company D	0.0	0.0	0.0

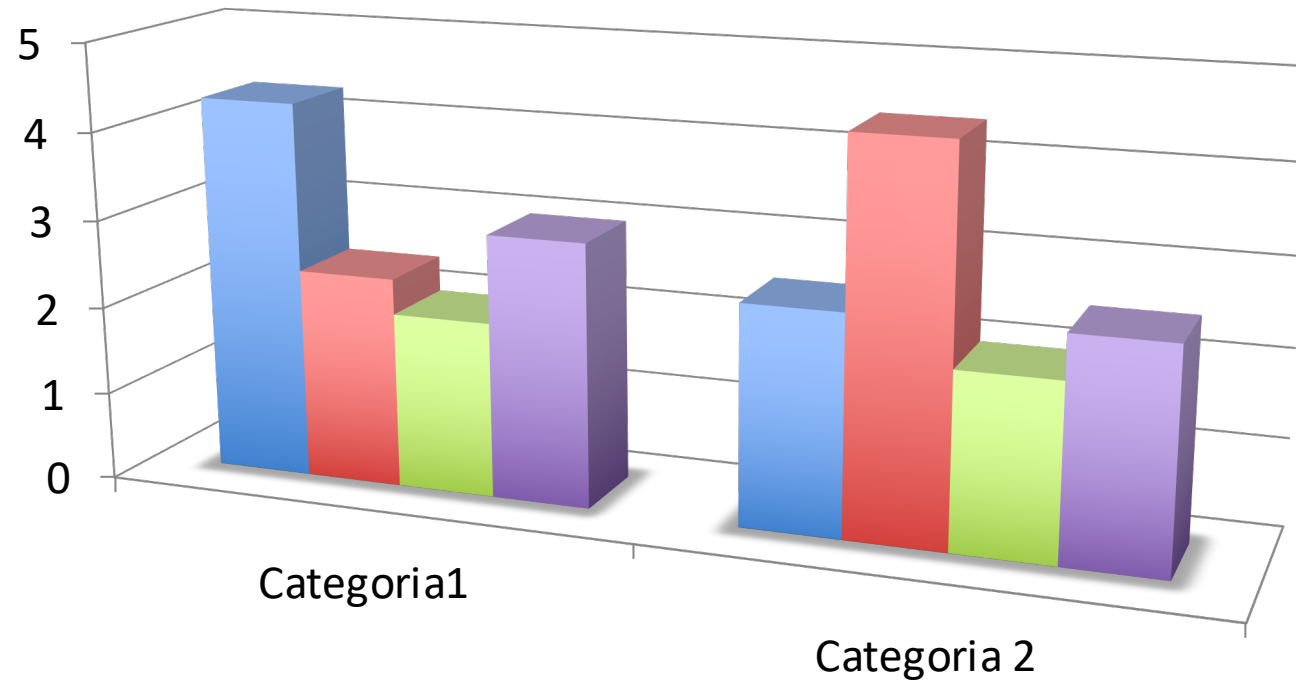
Give emphasis to relevant results

DATA-INK RATIO

Visual Display of Quantitative Data
Edward Tufte, 1983

DATA-INK RATIOⁿ

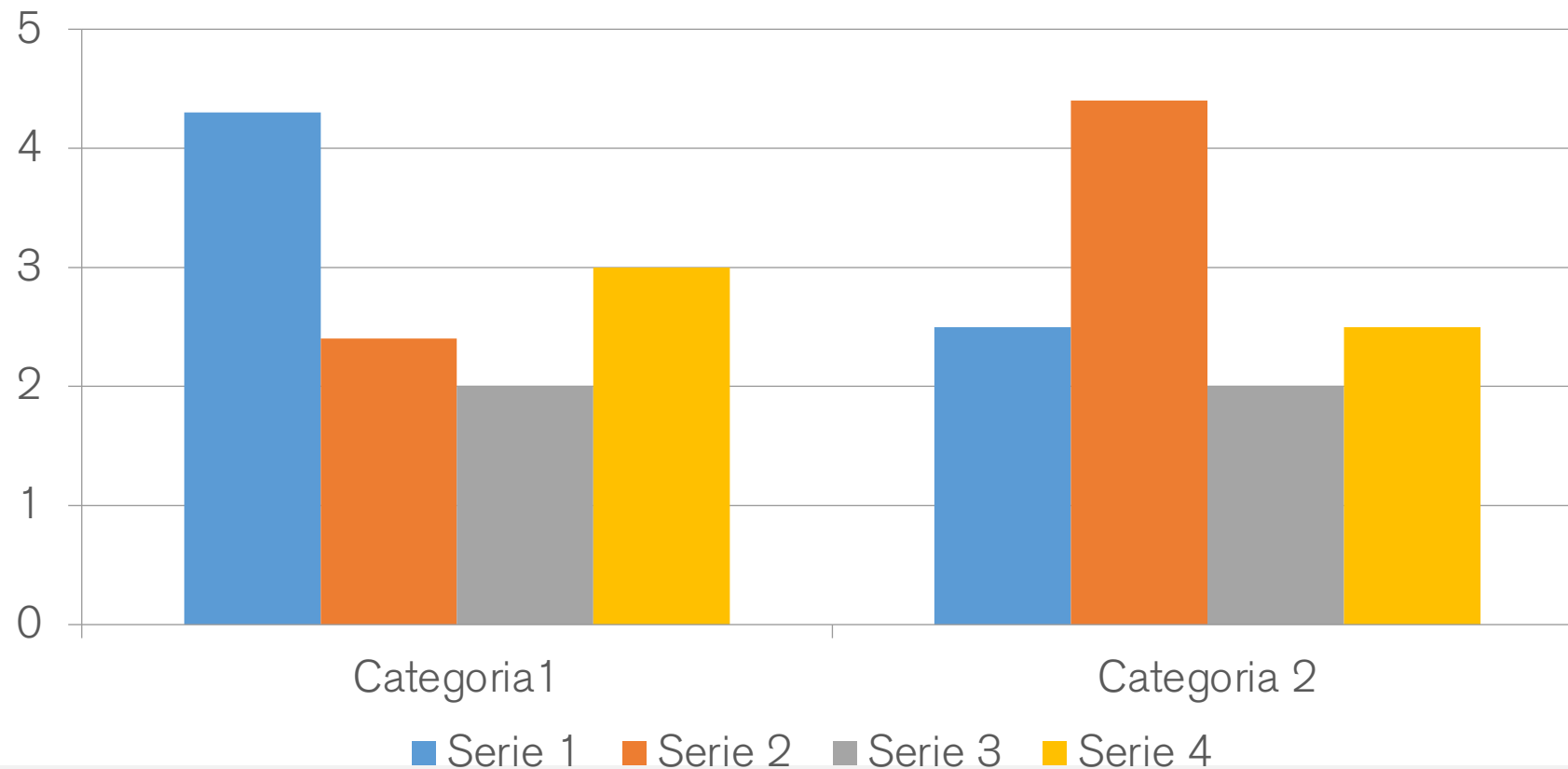
$$\text{Data-Ink Ratio} = \frac{\text{Data ink}}{\text{Total ink used in graphic}}$$



■ Serie 1 ■ Serie 2 ■ Serie 3 ■ Serie 4

DATA-INK RATIOⁿ

$$\text{Data-Ink Ratio} = \frac{\text{Data ink}}{\text{Total ink used in graphic}}$$



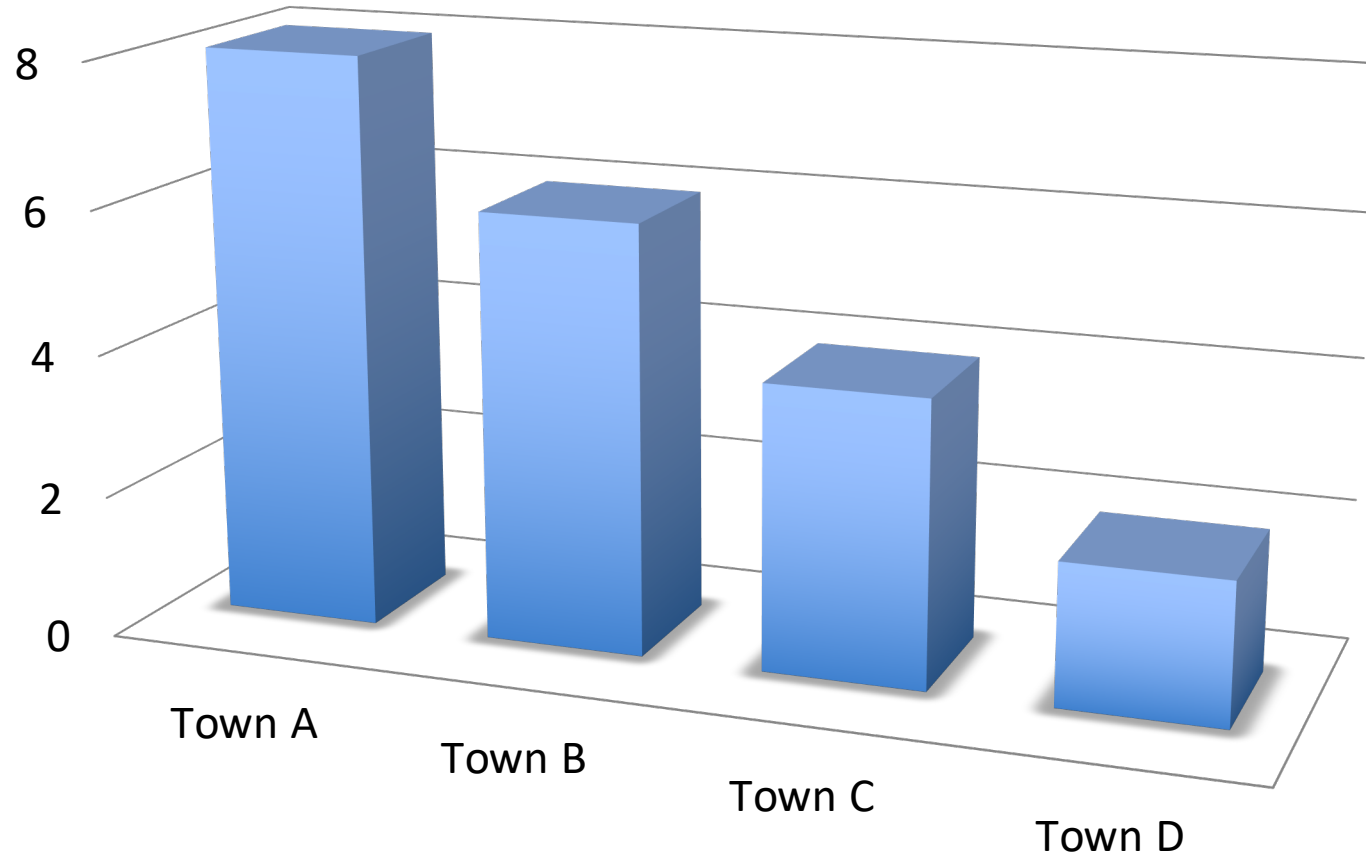
BAR CHARTS

Represent discrete quantities

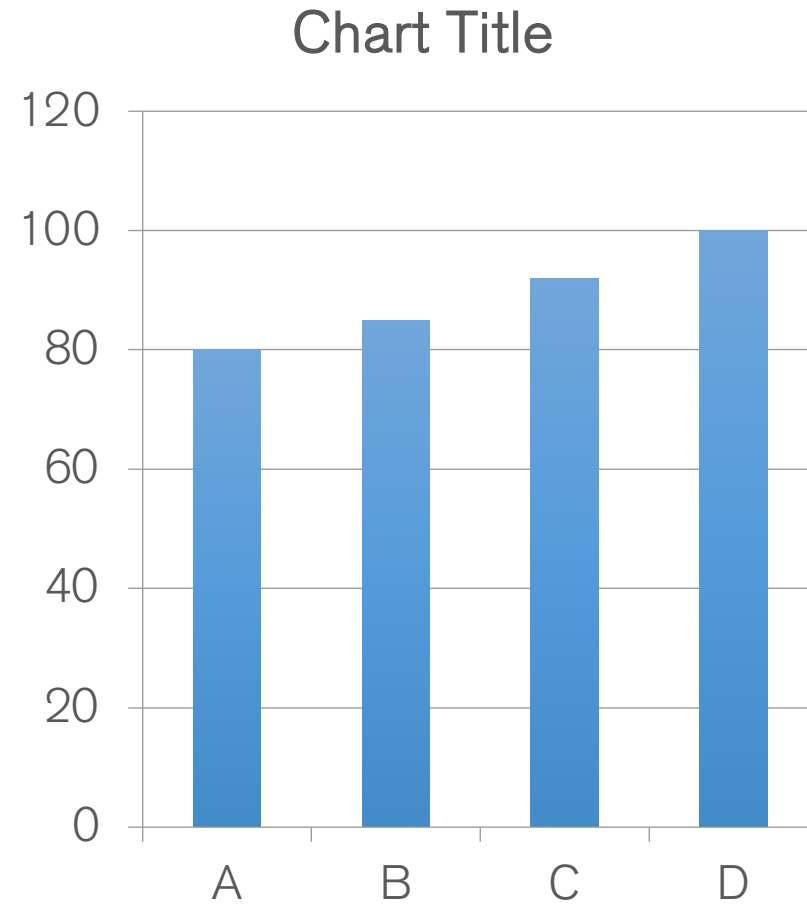
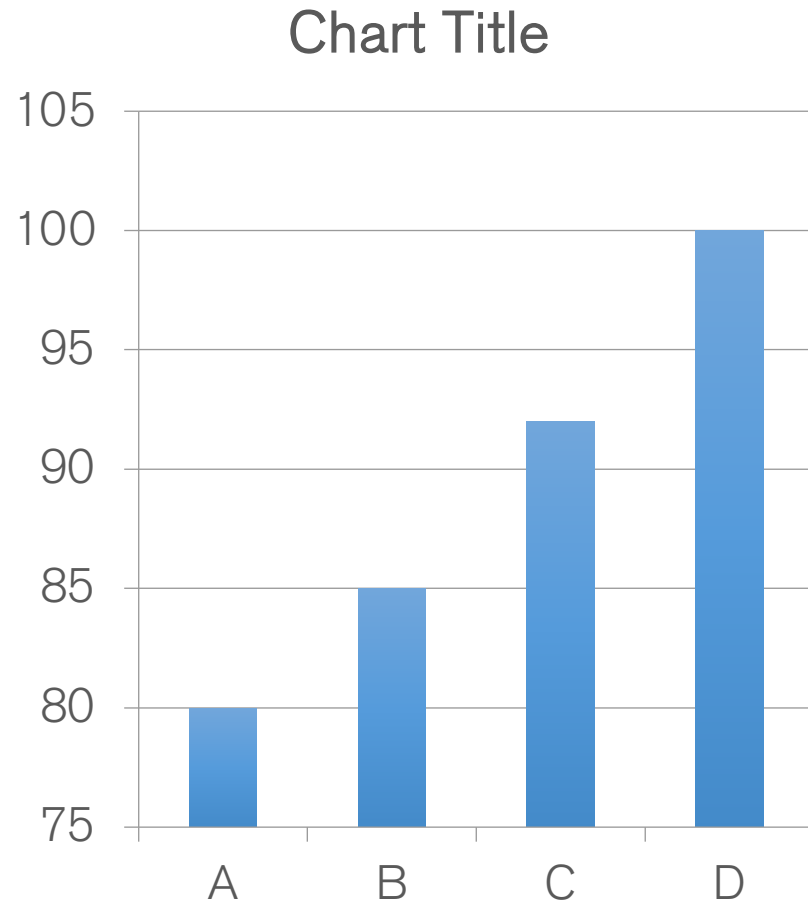


BAR CHARTS

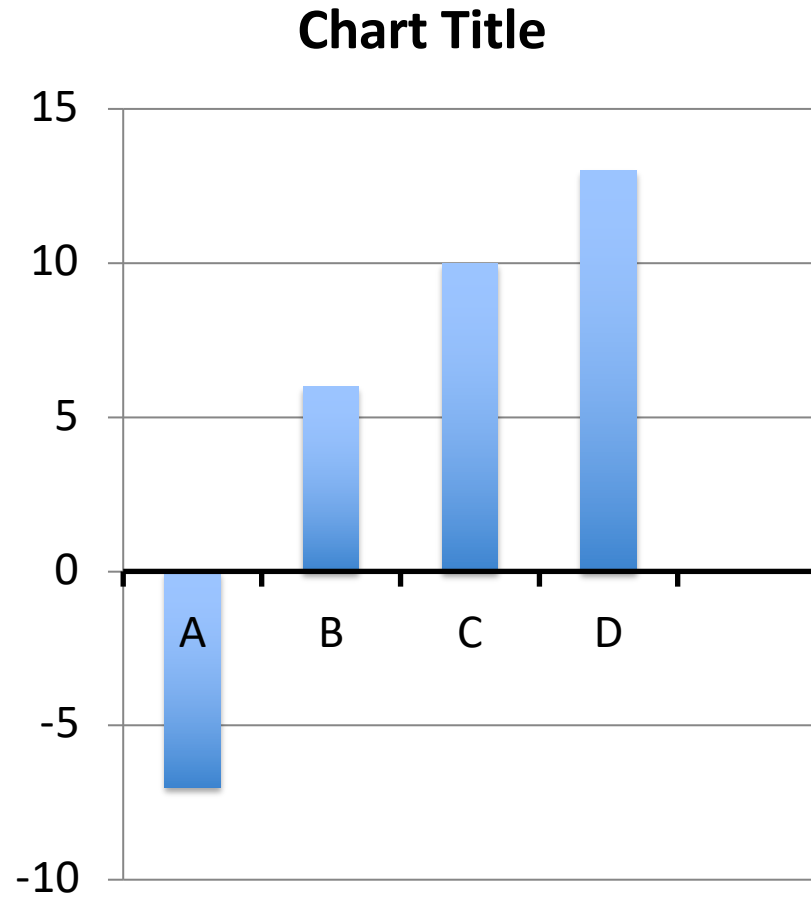
Avoid non-functional adornation



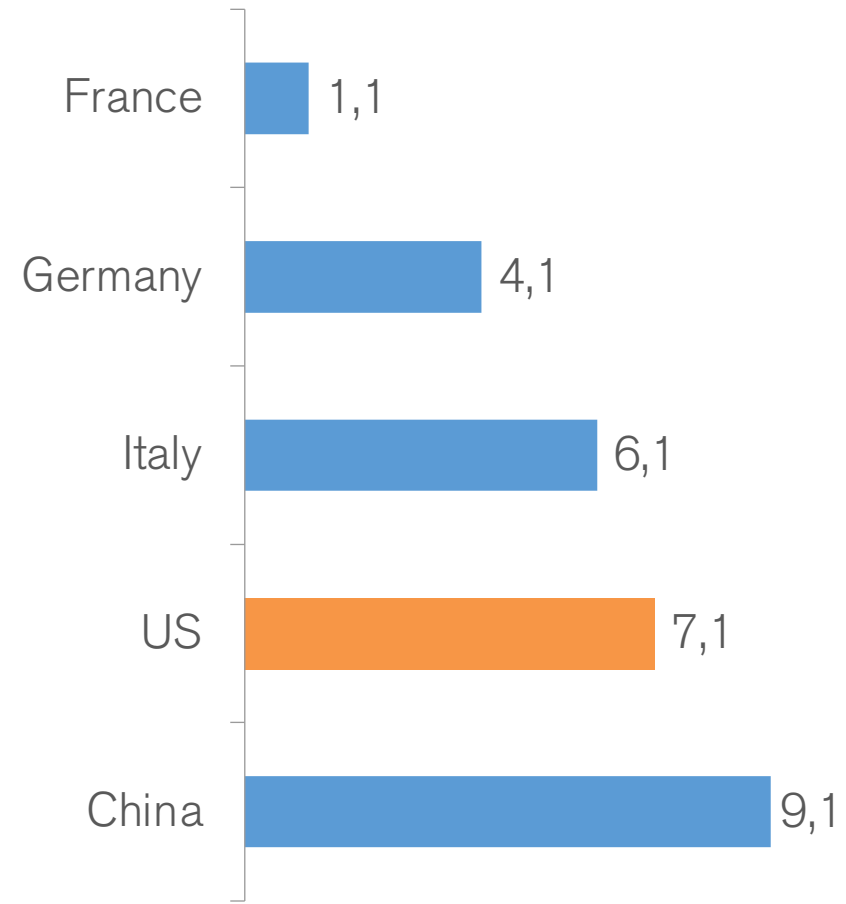
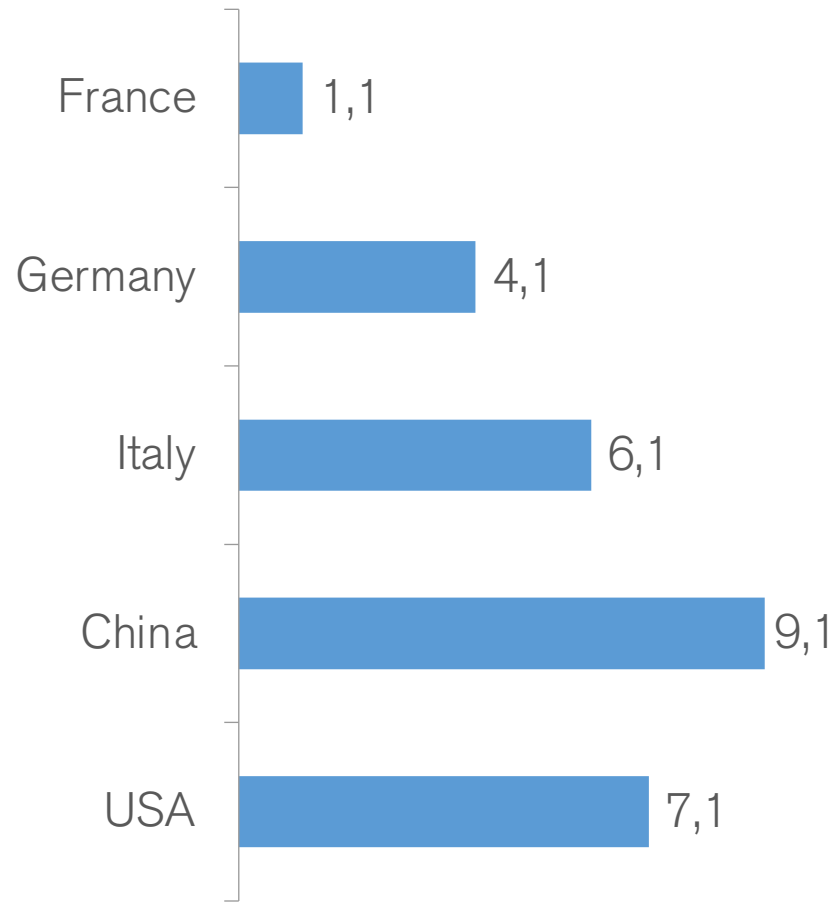
BAR CHARTS: BASELINE

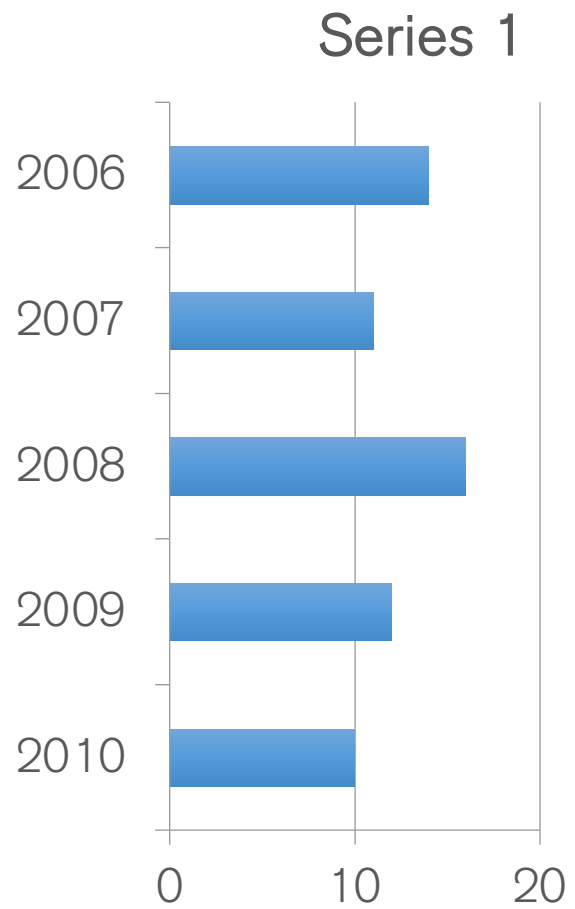


BAR CHARTS: BASELINE

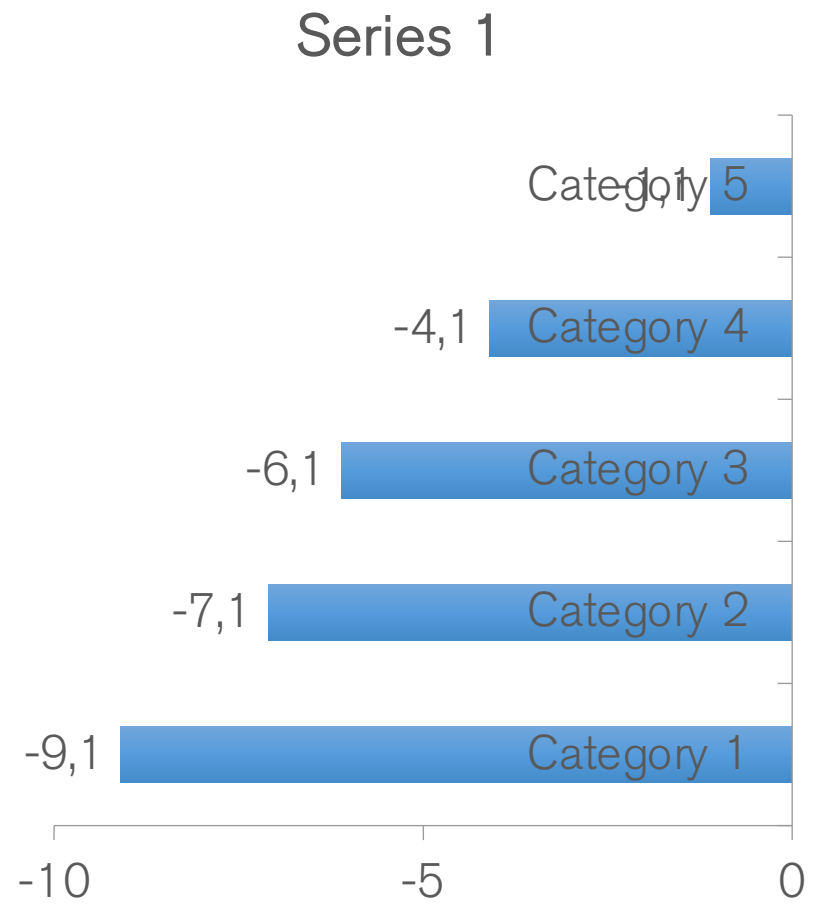


BAR CHARTS: ORDERING



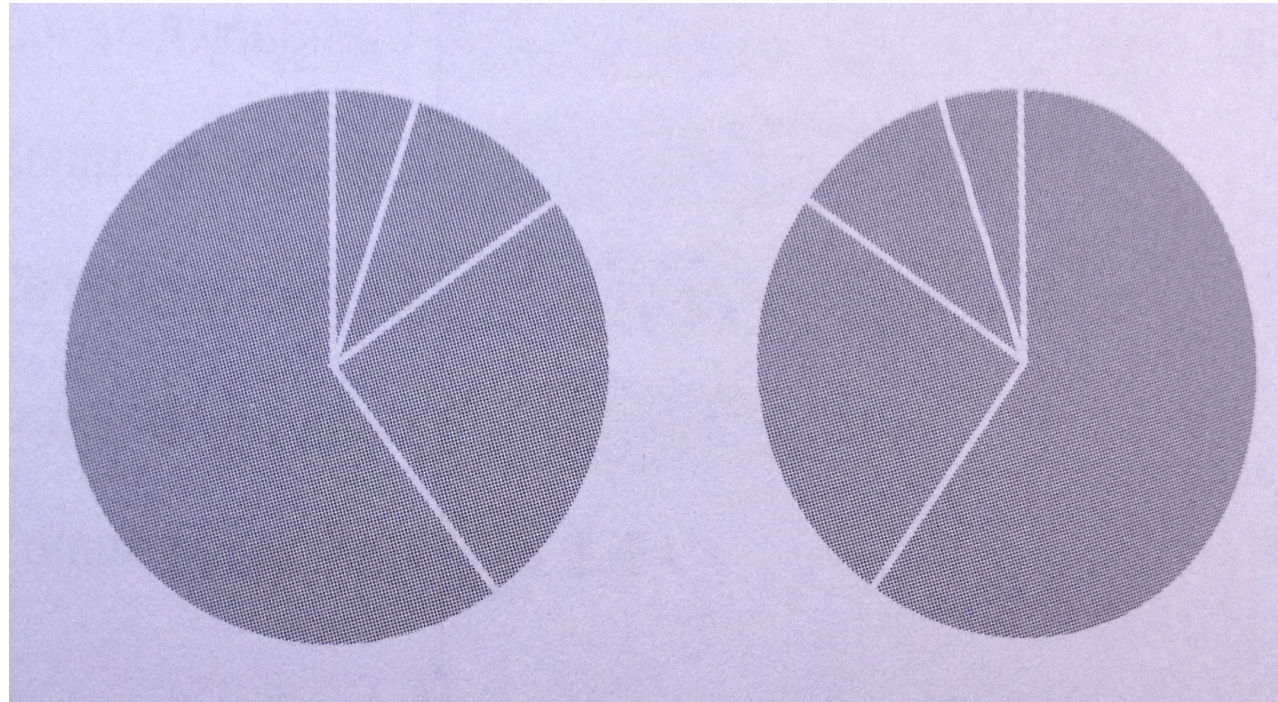


■ Series 1

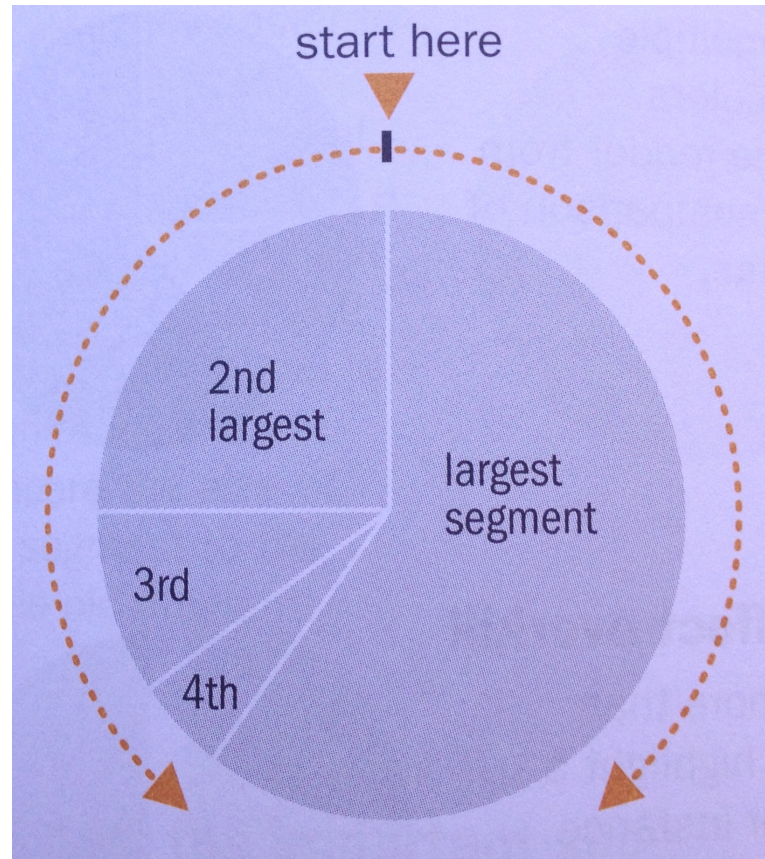


PIE CHARTS

- Pie Charts compares relative sizes and contributions

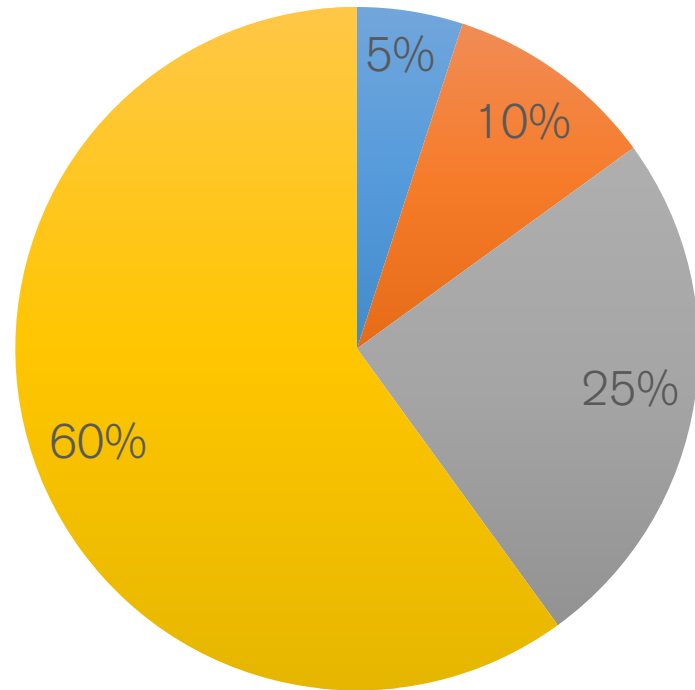


PIE CHARTS: ORDERING SLICES

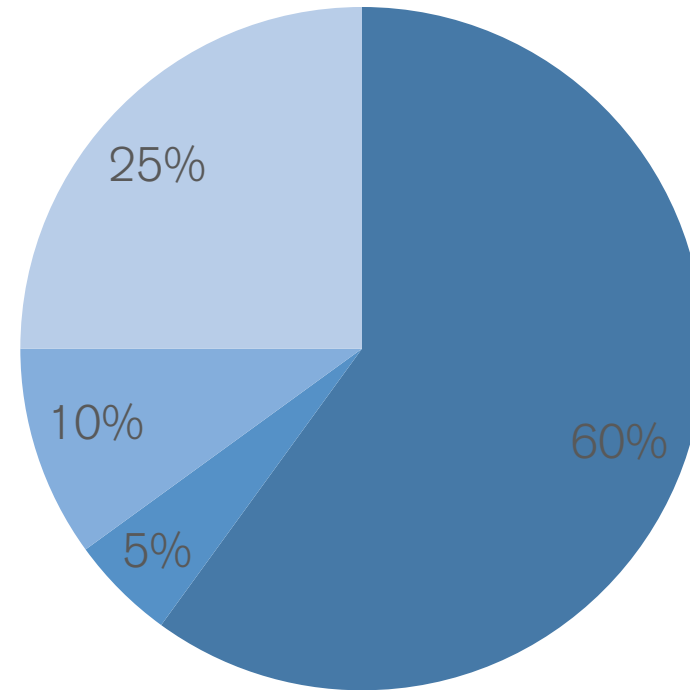


CHARTING EXAMPLES

Sales



Sales



May these charts be improved? Why? How?

TAKEAWAY MESSAGES

- Charts exploit position on scale VV
- Best practice to reduce biases and misinterpretation of charts

VISUALIZATION TAXONOMY

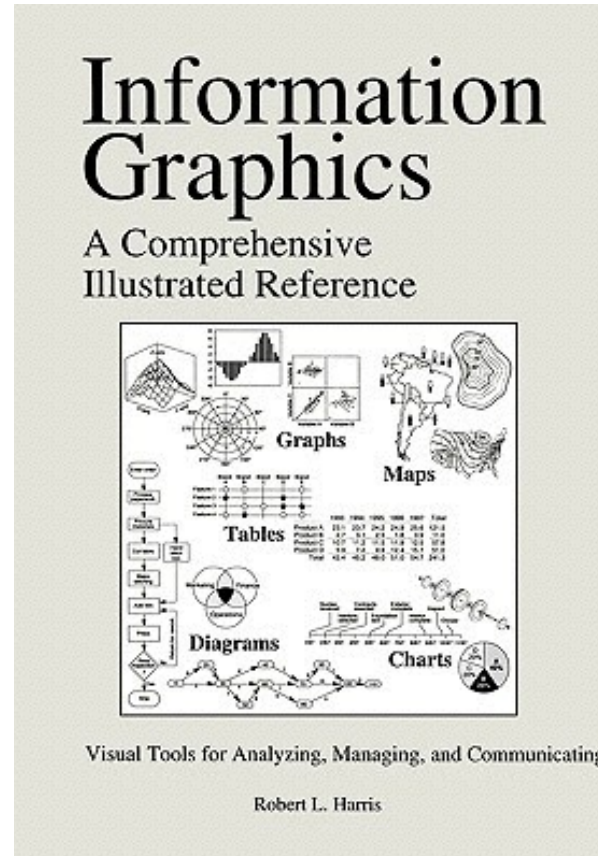
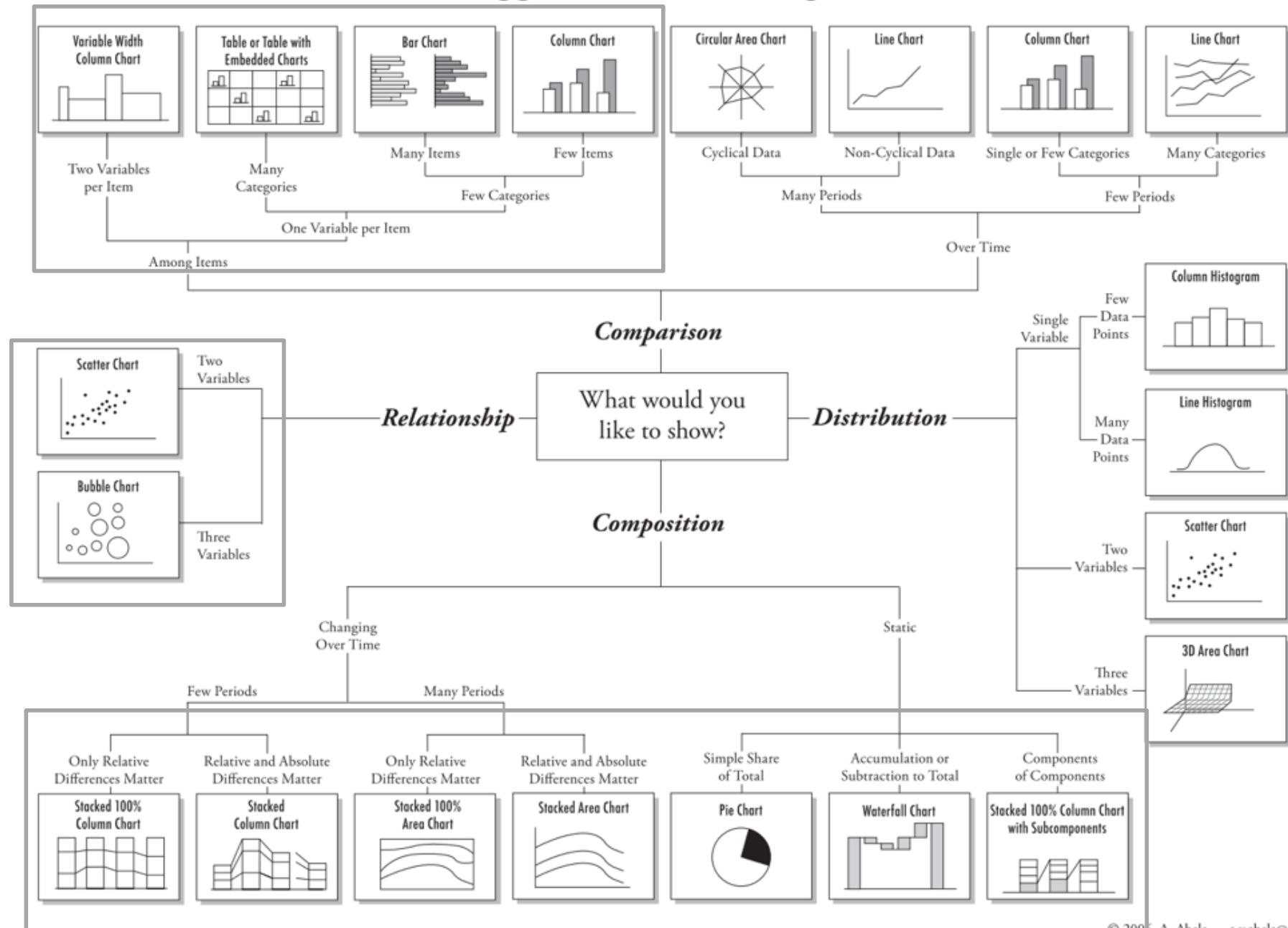
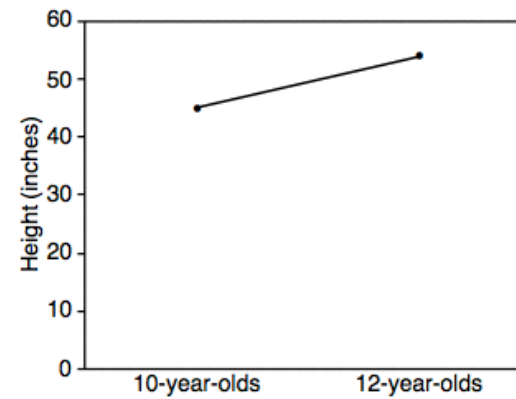
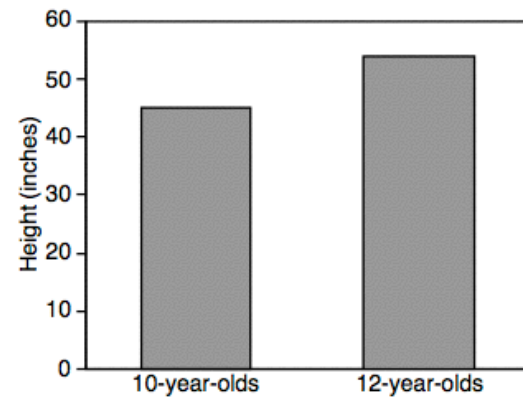
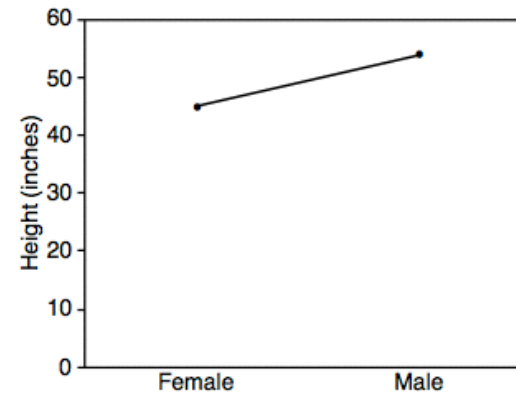
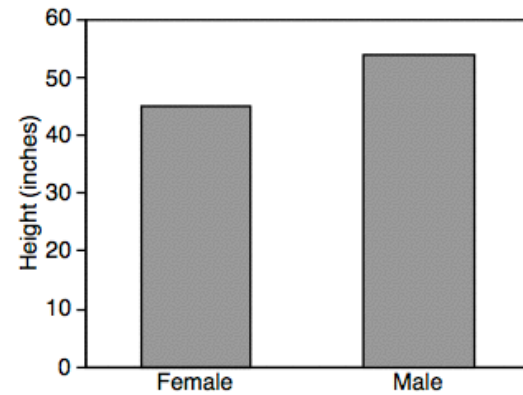


Chart Suggestions—A Thought-Starter

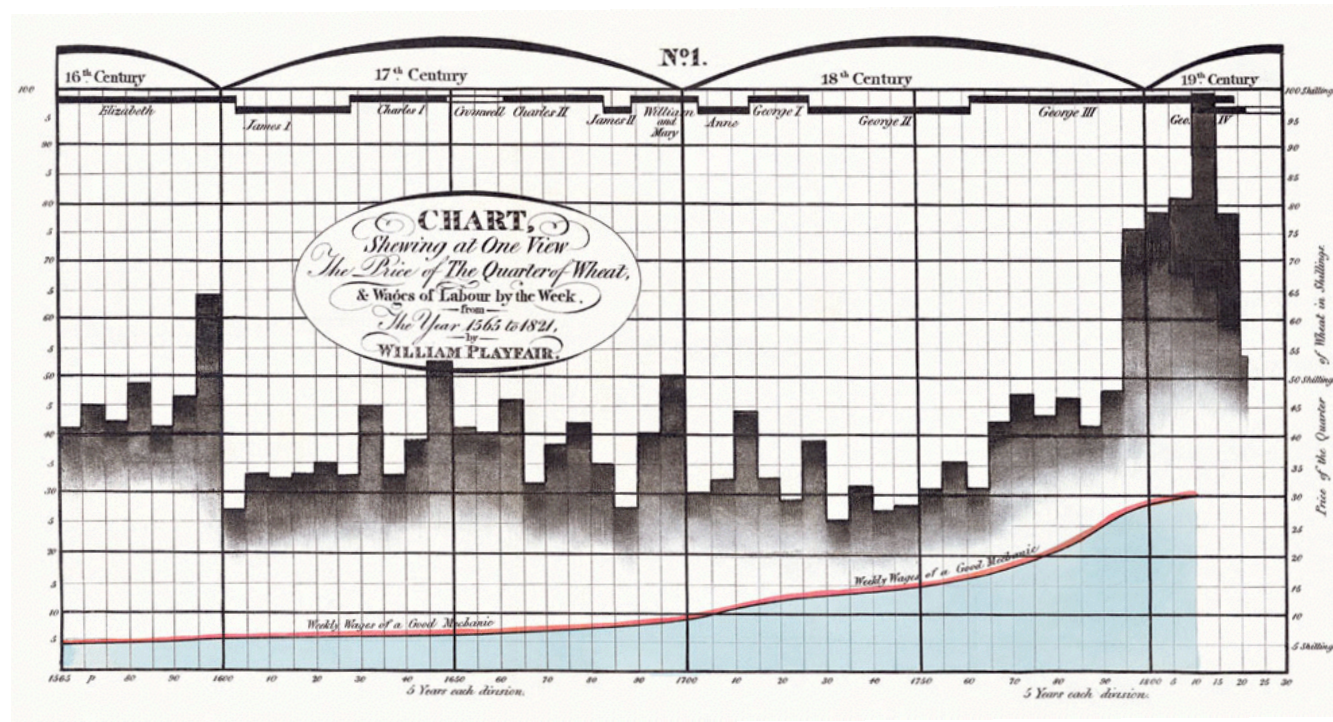


BARS VS. LINES



TREND OVER TIME

WILLIAM PLAYFAIR 1759-1823



TREND OVER TIME

Apple Inc. (AAPL) - NasdaqGS

[+ Add to Portfolio](#) [Like](#) 6k

601.10 ↑ 15.53 (2.65%) 4:00PM EDT | After Hours: **604.60** ↑ 3.50 (0.58%) 7:15PM EDT - Nasdaq Real Time Price



TREND OVER TIME

Published: February 2, 2010

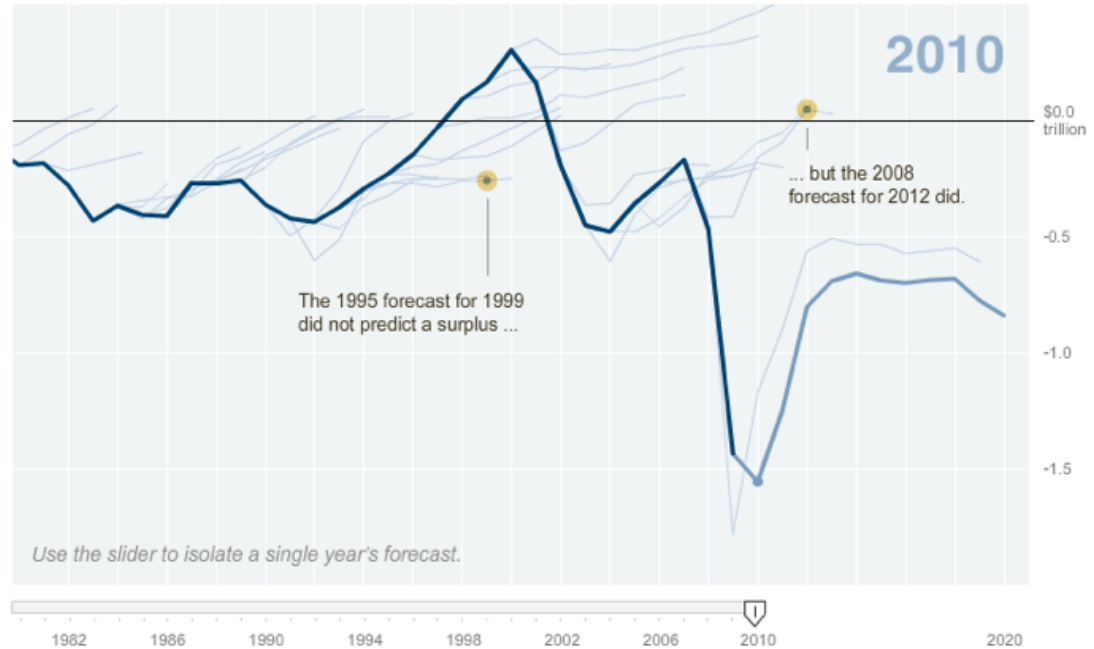
Budget Forecasts, Compared With Reality

Just two years ago, surpluses were predicted by 2012. How accurate have past White House budget forecasts been?

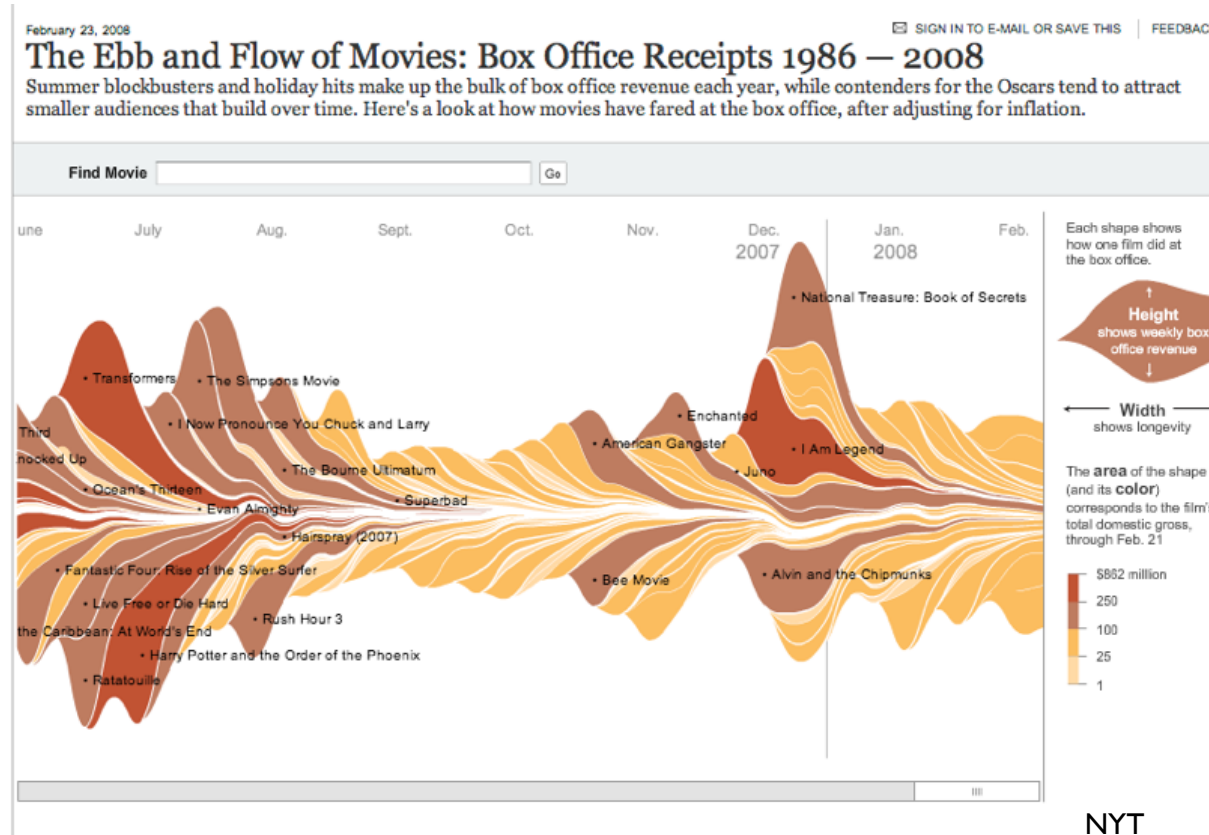
1 2 3 4 5 6 NEXT ▶

Latest forecast

Today, with a better understanding of the severity of the economic downturn, the deficit situation is much more dire.



STREAMGRAPHS



Vision Statement



David Berkeley is an editor at HBR. **APFlicker** is a developer and information scientist based in Toronto.

How to Read This Graph

This stream graph shows tweet volume over time. Each color stream represents the proportion of tweets containing a given word, such as **HELP**. Color is used only to distinguish words.

Six Ways to Find Value in Twitter's Noise

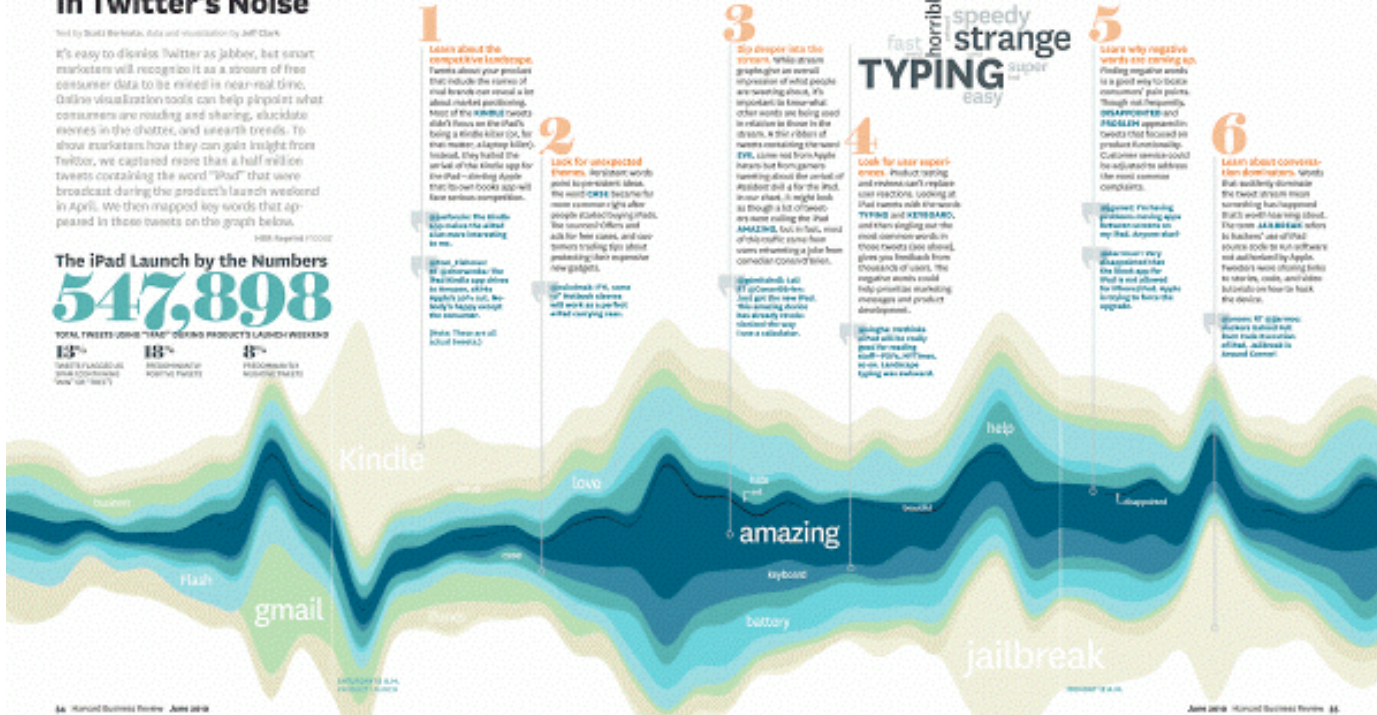
It's easy to dismiss Twitter as jabber, but smart marketers will recognize it as a stream of free consumer data to be mined in near-real time. Online visualization tools can help pinpoint what consumers are reading and sharing, elucidate voices in the chatter, and unearth trends. To show marketers how they can gain insight from Twitter, we captured more than a half million tweets containing the word "iPad" that were broadcast during the product's launch weekend in April. We then mapped key words that appeared in those tweets on the graph below.

HBR Special Report

The iPad Launch by the Numbers

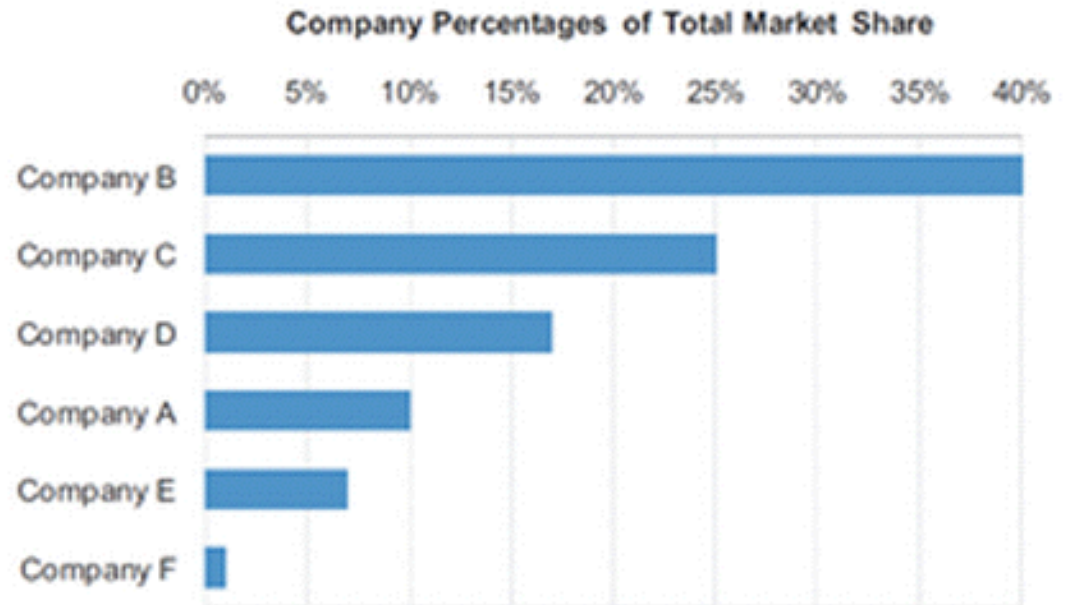
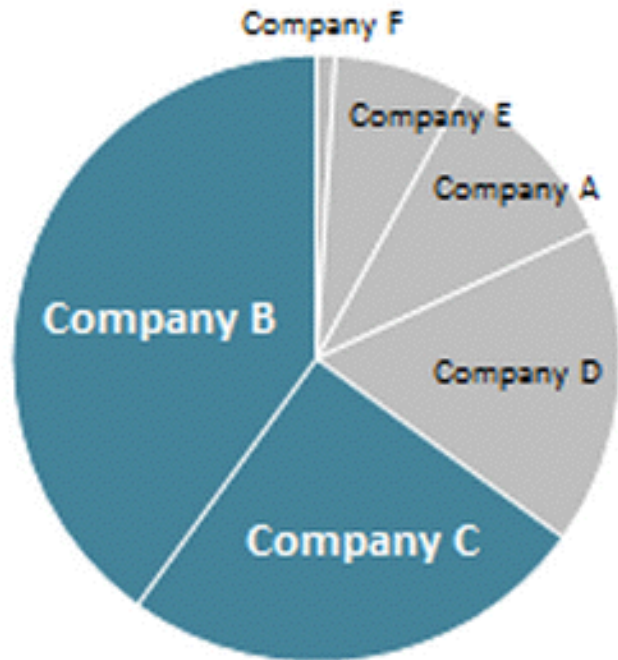
547,898

TOTAL TWEETS USING "IPAD" DURING PRODUCT'S LAUNCH WEEKEND
13% TWEETS FLAGGED AS SPAM (COMPARING NEW YORK TWEETS)
18% RECOMMENDED BY OTHER TWEETS
4% RECOMMENDED BY FOLLOWERS

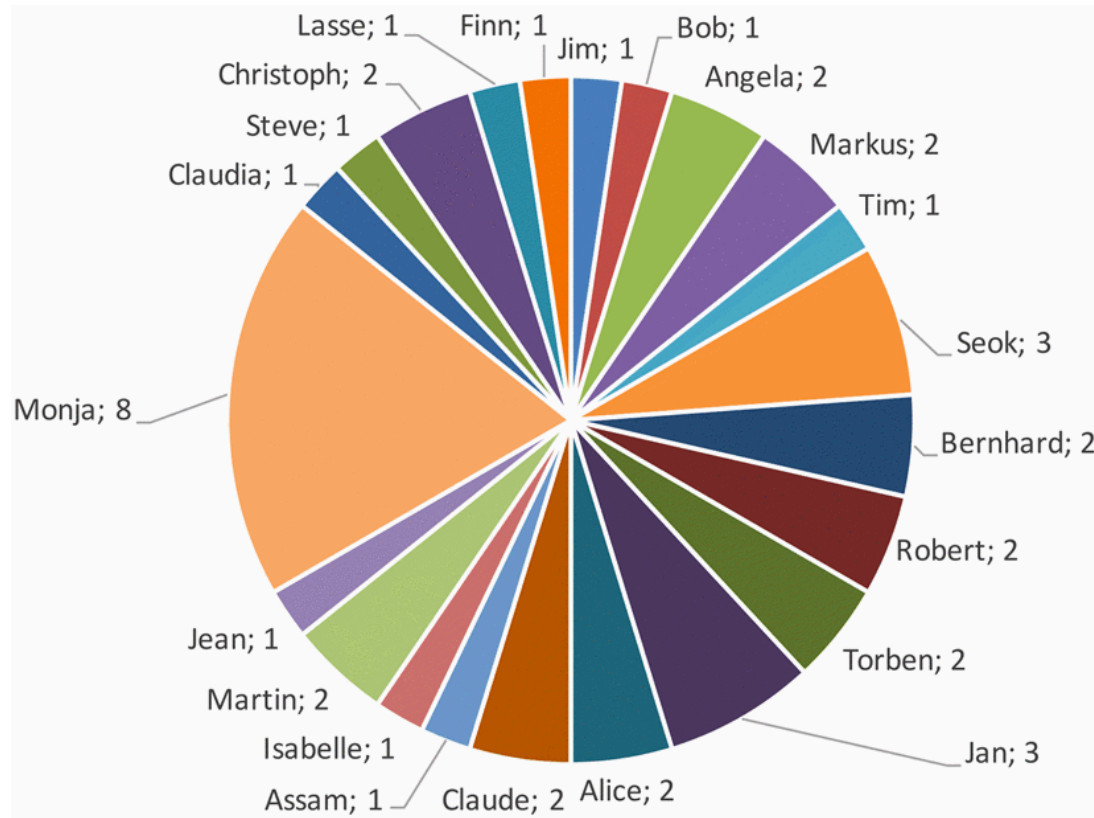


PIE VS BAR CHARTS

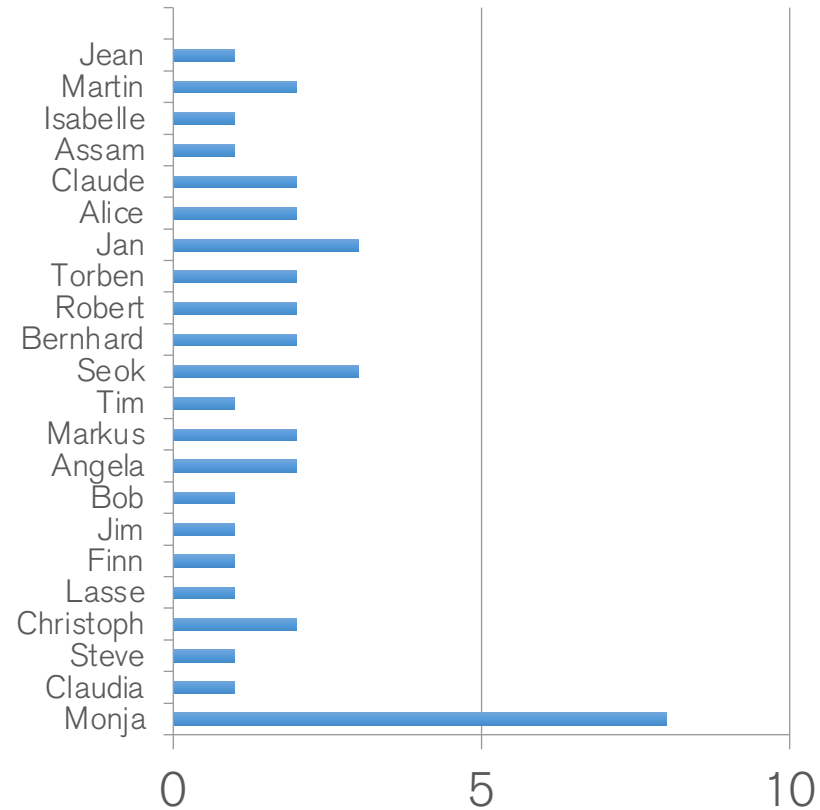
65% of the market is controlled by companies B and C



PIES VS BAR CHARTS

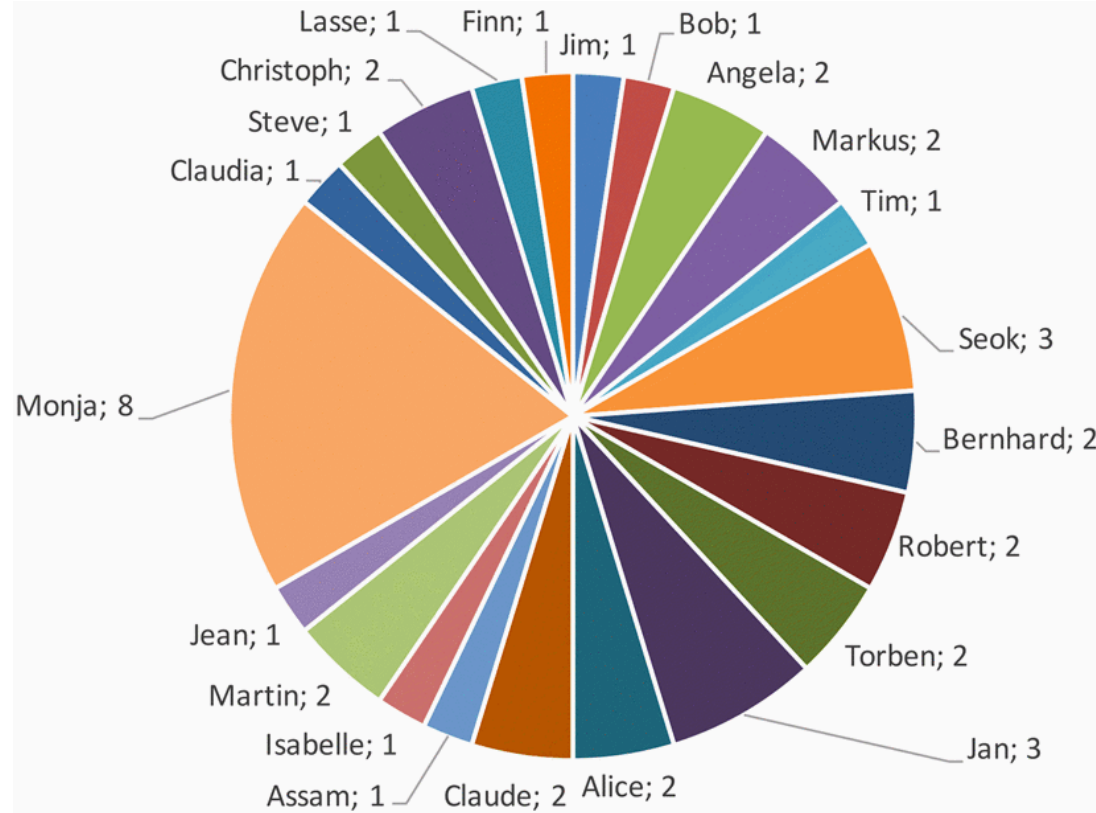


Episodes

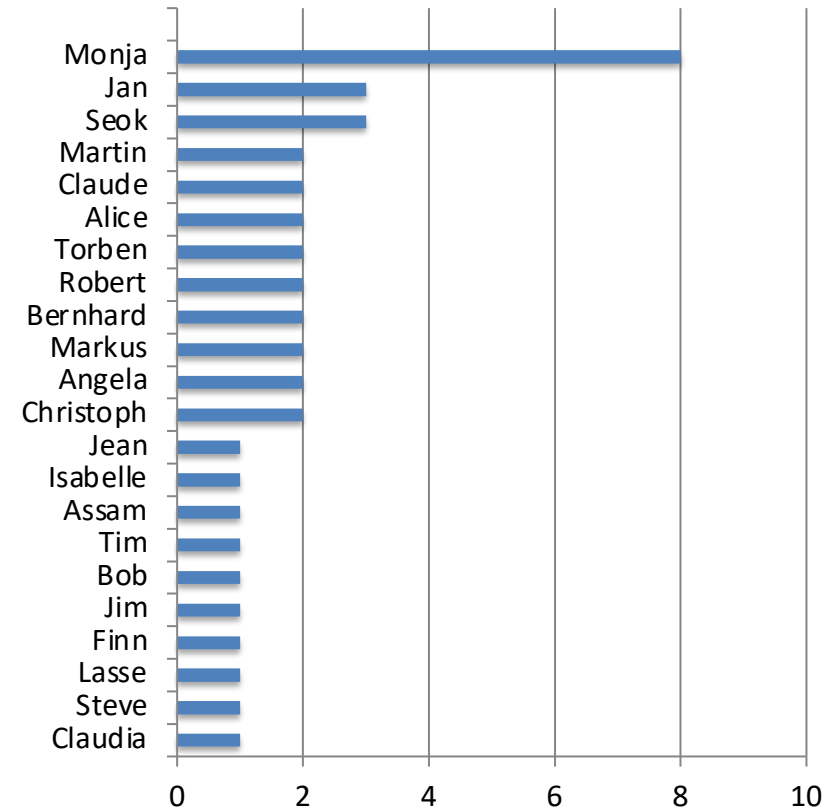


Furthermore, we present the distribution of attacks towards employees in detail in Fig. 10 right. The blue employees are secretaries, the green ones are administrators and the red ones are scientific employees. The number following the name is the number of times that person was attacked. All of the names are pseudonyms for real people. The person that suffered the most attacks is Monja a secretary with overall 8 attacks. In contrast, all other victims suffered between 1 and 3 attacks.

PIES VS BAR CHARTS (IMPROVED)

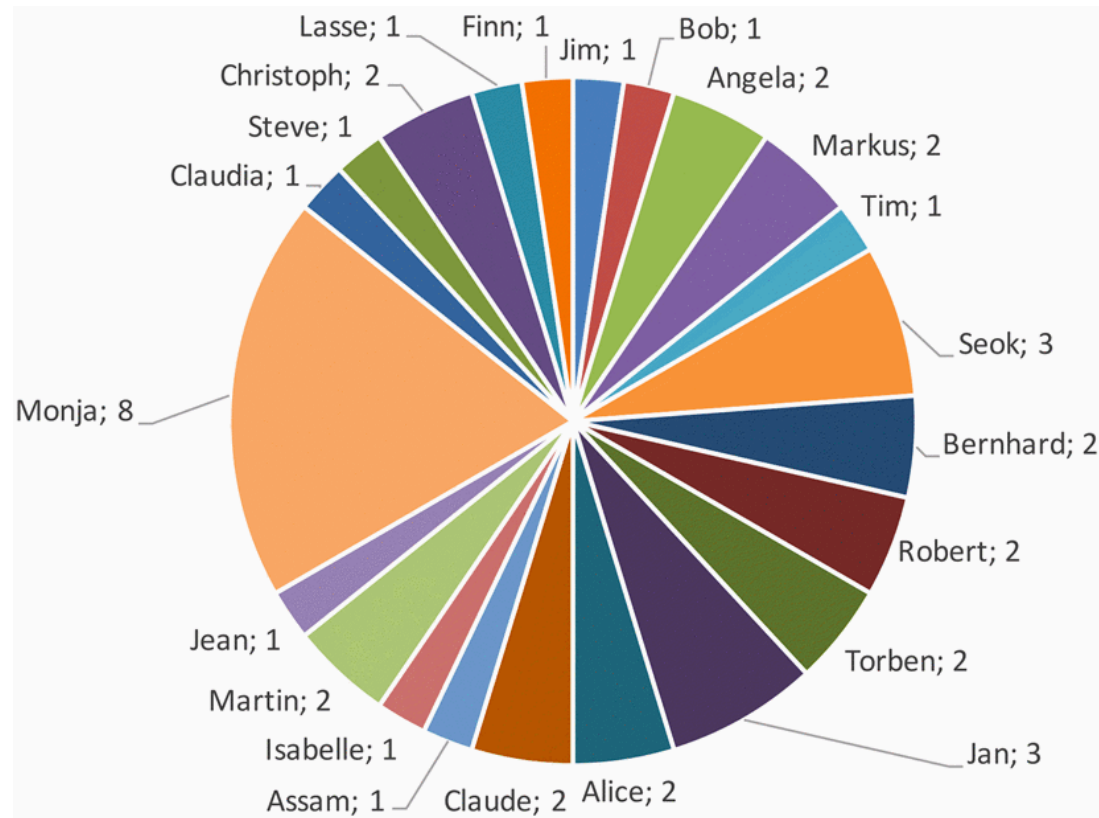


Episodes per person

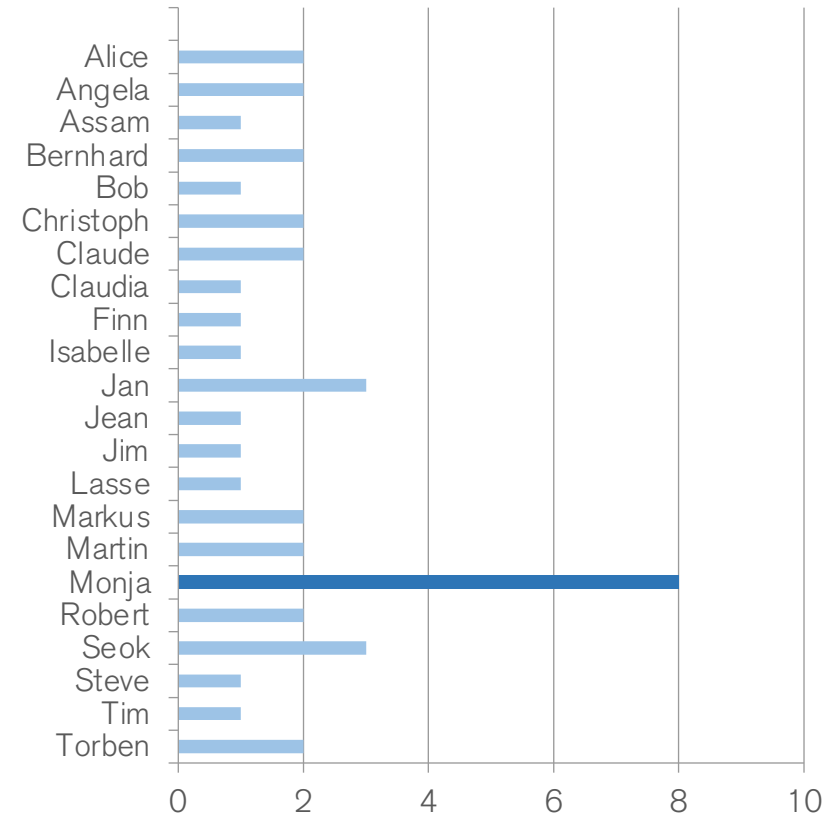


Furthermore, we present the distribution of attacks towards employees in detail in Fig. 10 right. The blue employees are secretaries, the green ones are administrators and the red ones are scientific employees. The number following the name is the number of times that person was attacked. All of the names are pseudonyms for real people. The person that suffered the most attacks is Monja a secretary with overall 8 attacks. In contrast, all other victims suffered between 1 and 3 attacks.

PIES VS BAR CHARTS (IMPROVED)

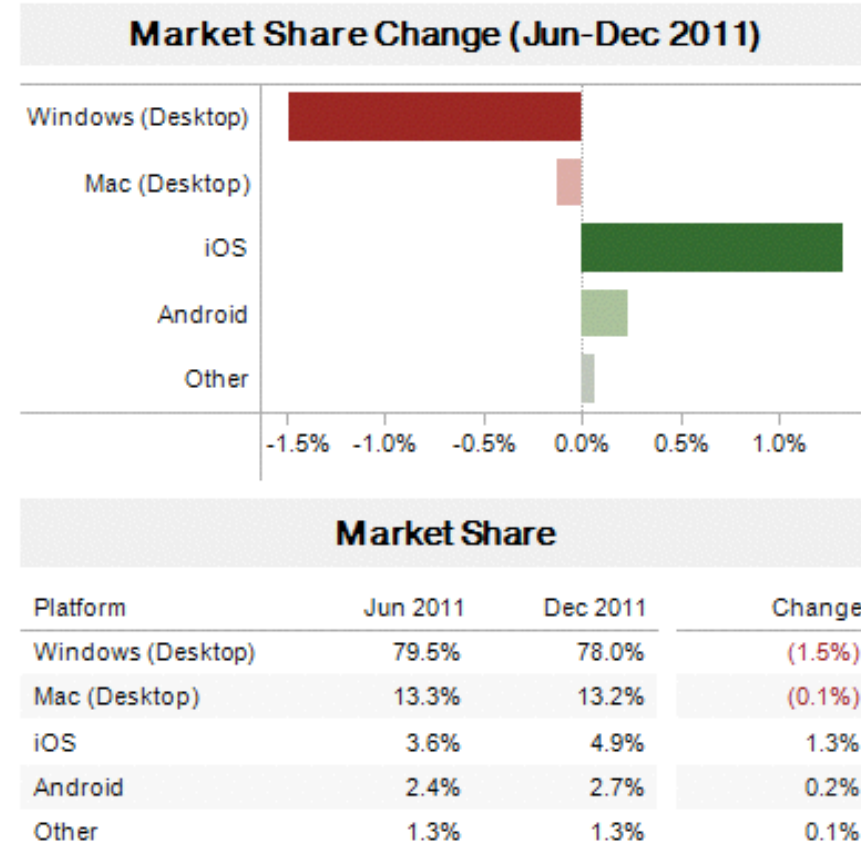
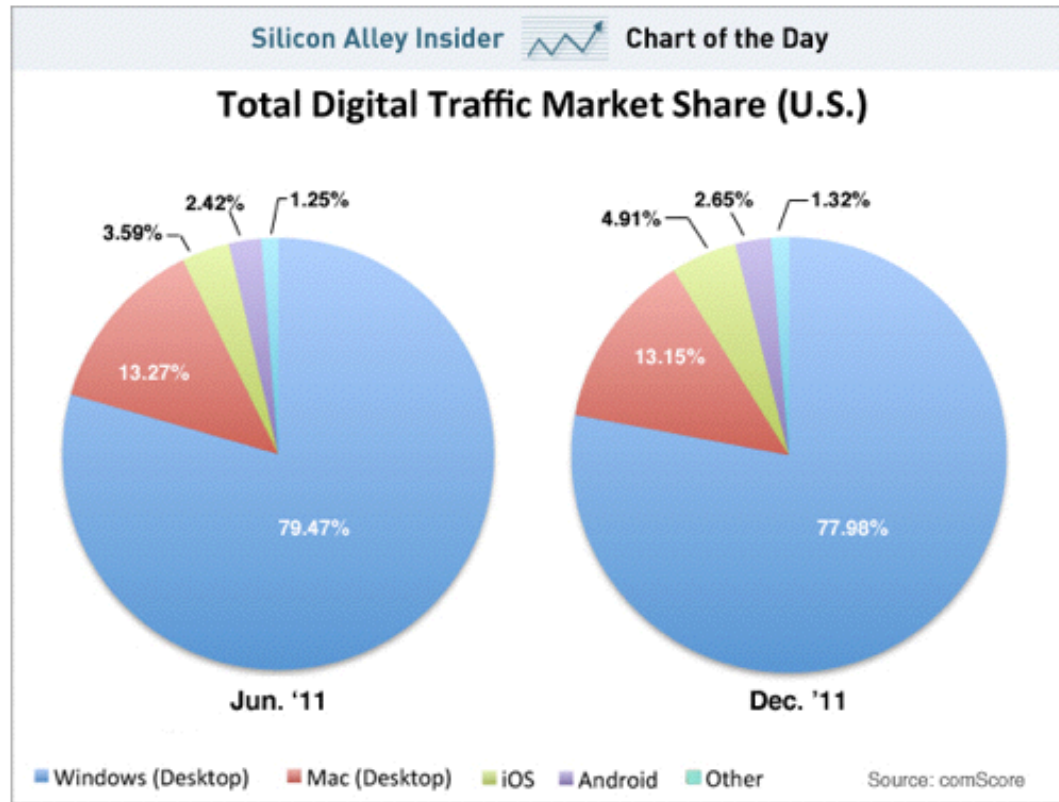


Episodes per person

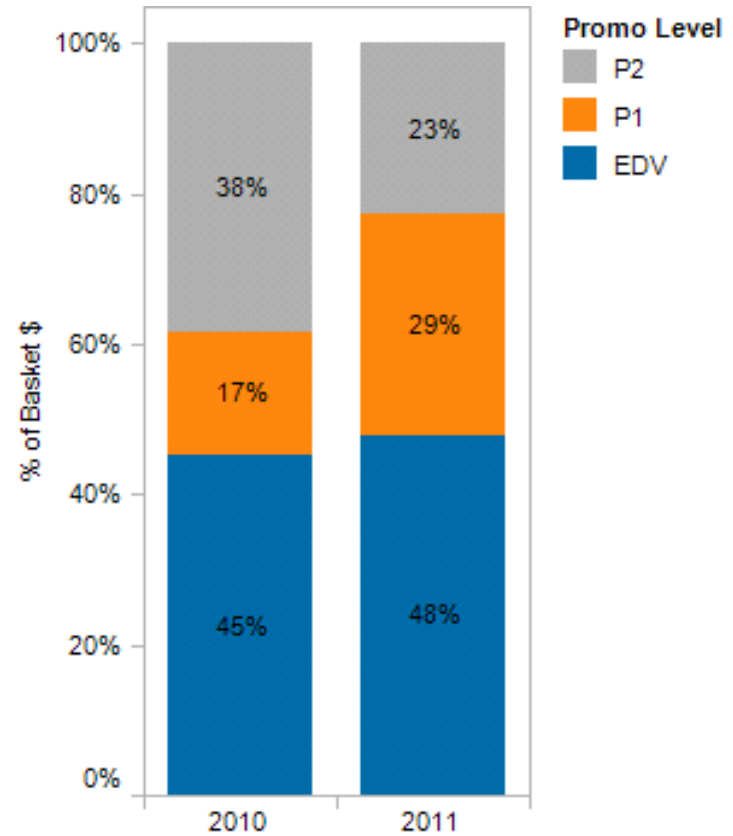
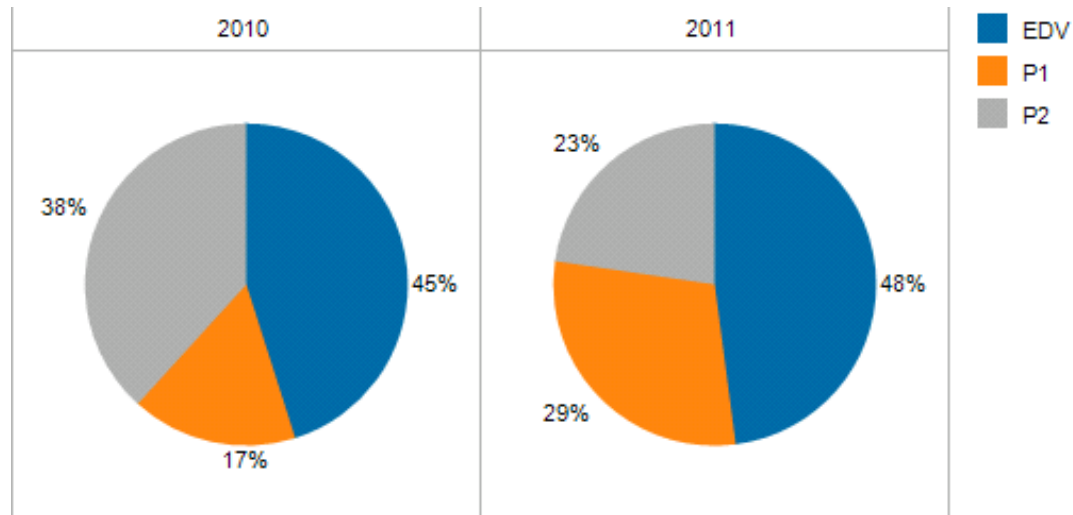


Furthermore, we present the distribution of attacks towards employees in detail in Fig. 10 right. The blue employees are secretaries, the green ones are administrators and the red ones are scientific employees. The number following the name is the number of times that person was attacked. All of the names are pseudonyms for real people. The person that suffered the most attacks is Monja a secretary with overall 8 attacks. In contrast, all other victims suffered between 1 and 3 attacks.

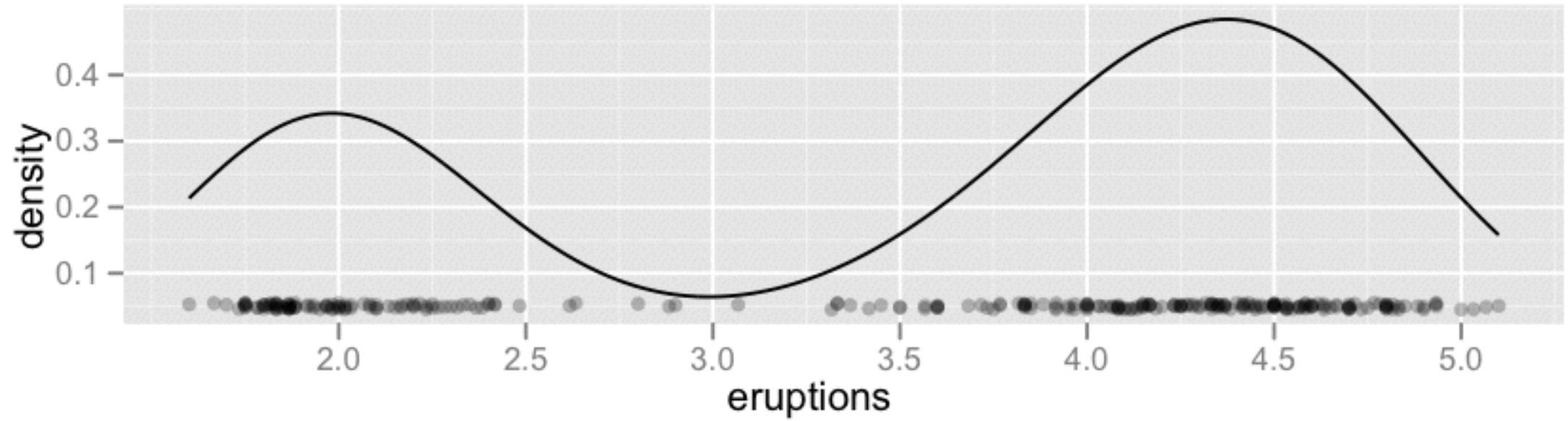
SHOWING CHANGES



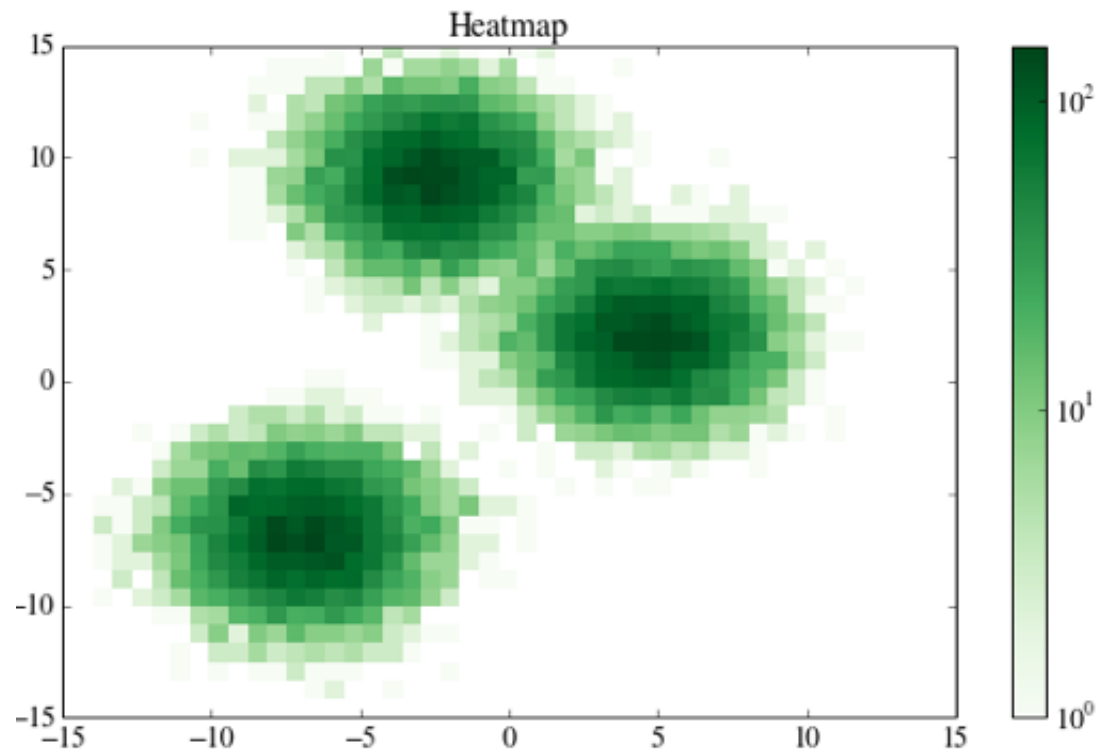
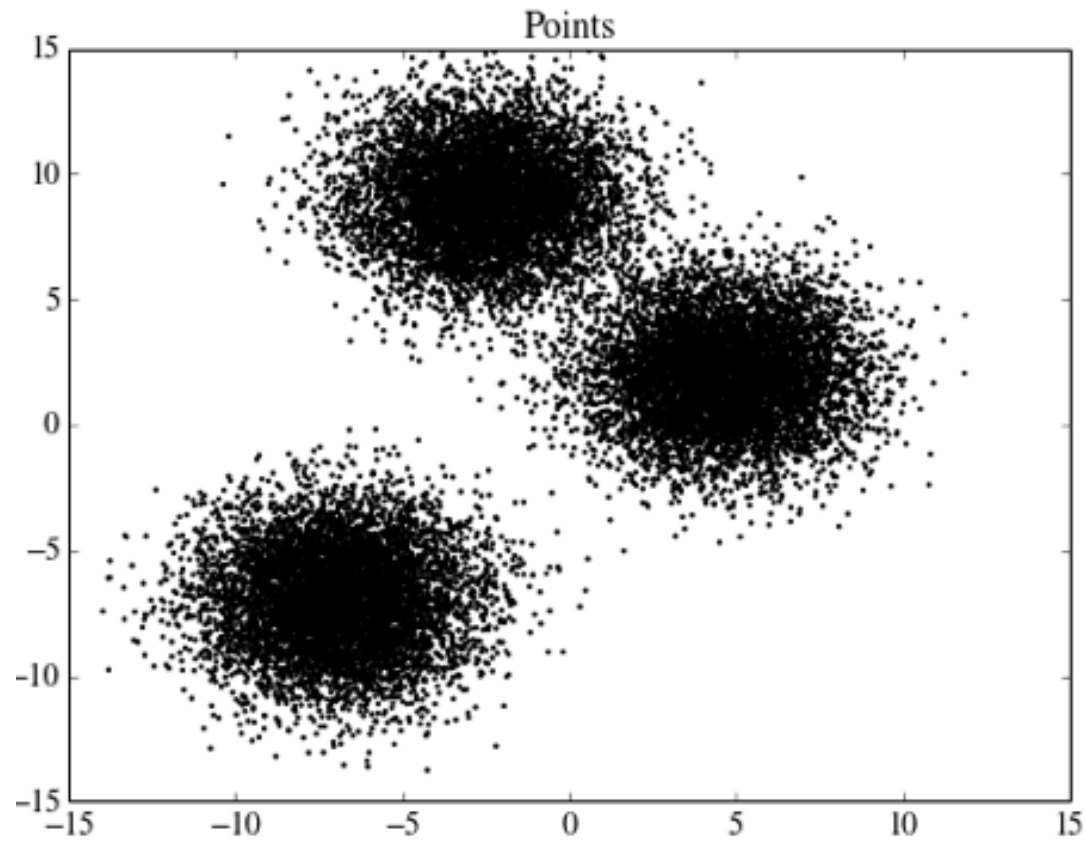
SHOWING CHANGES



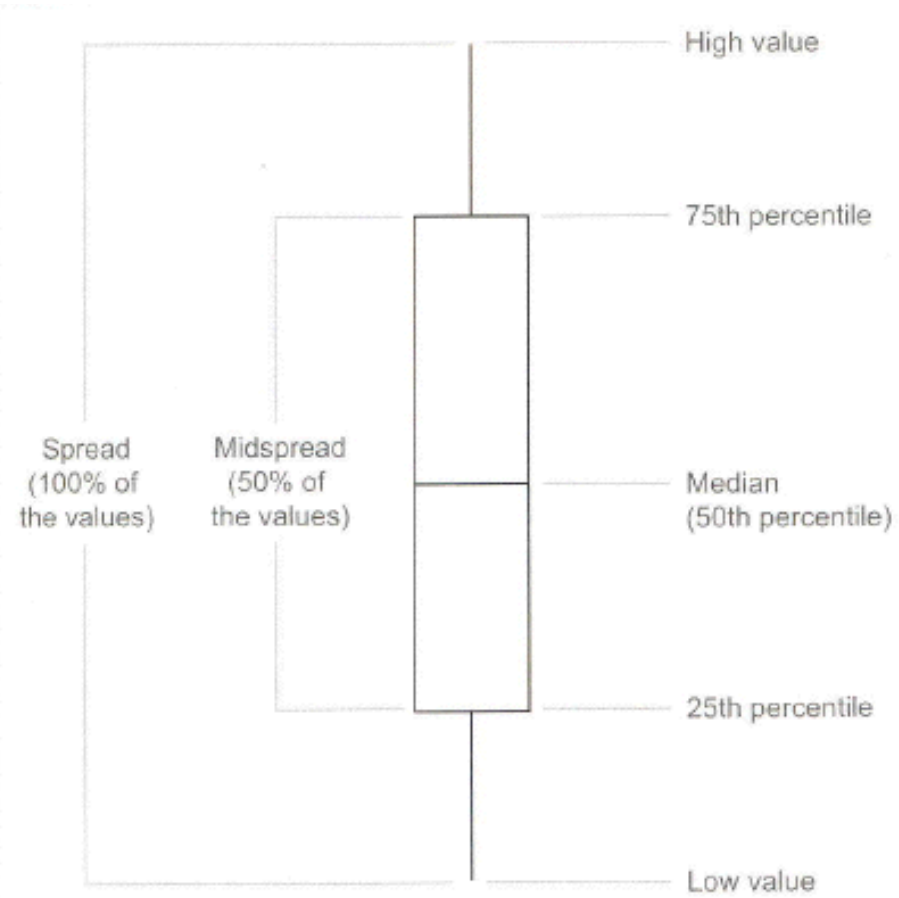
DENSITY PLOT



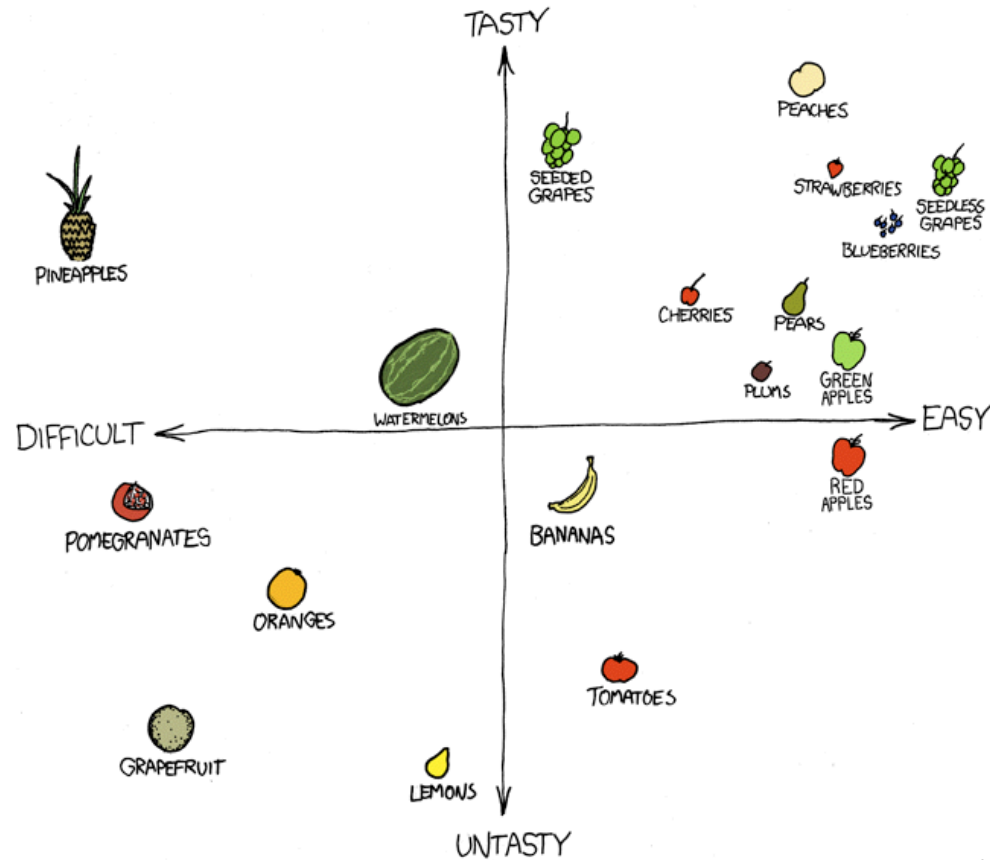
2D DENSITY PLOTS



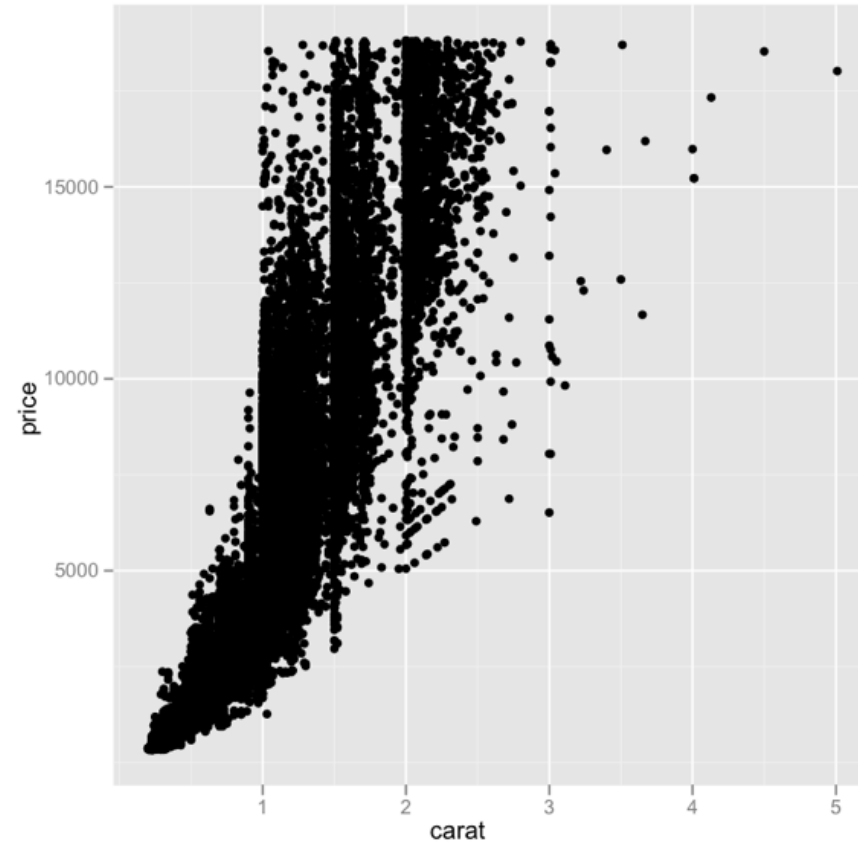
BOX PLOTS



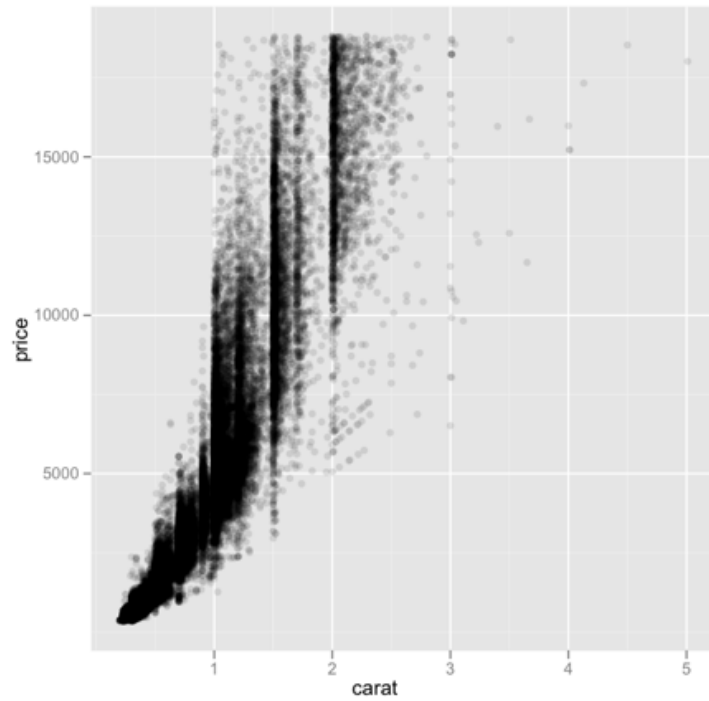
SCATTERPLOT



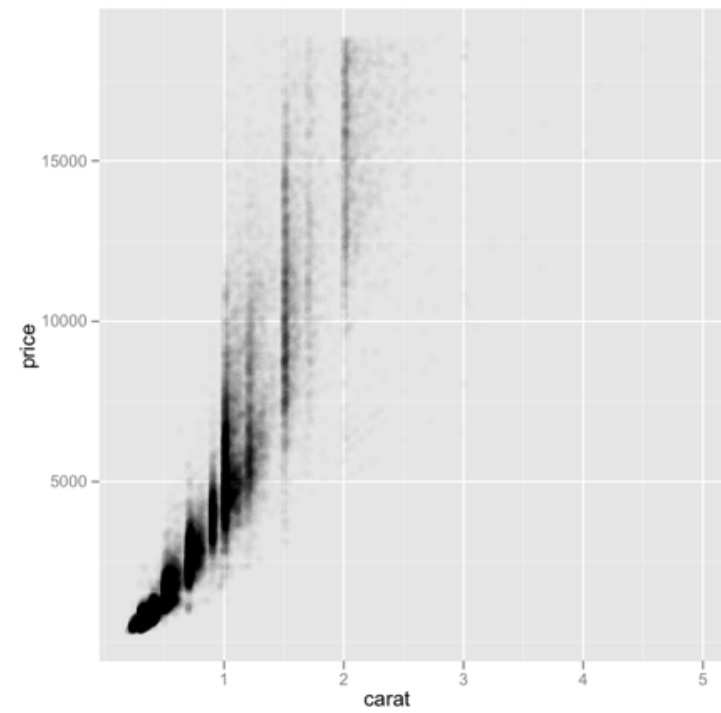
CLUTTERING, OVERPLOTTING



alpha=1/10

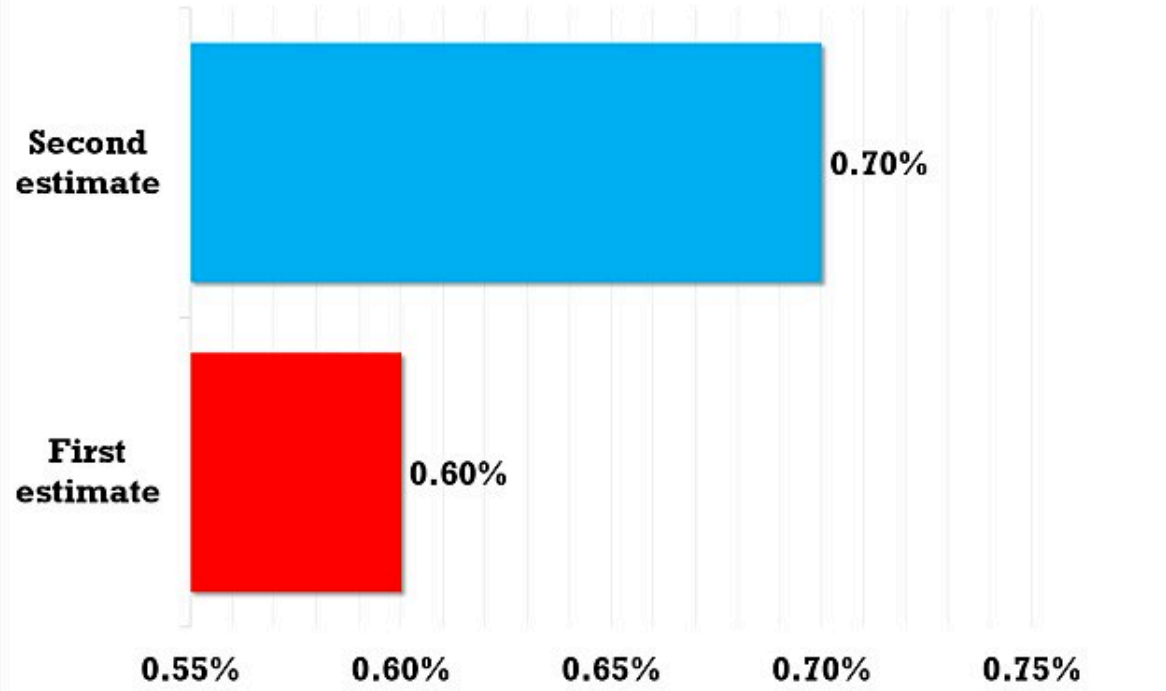


alpha=1/100



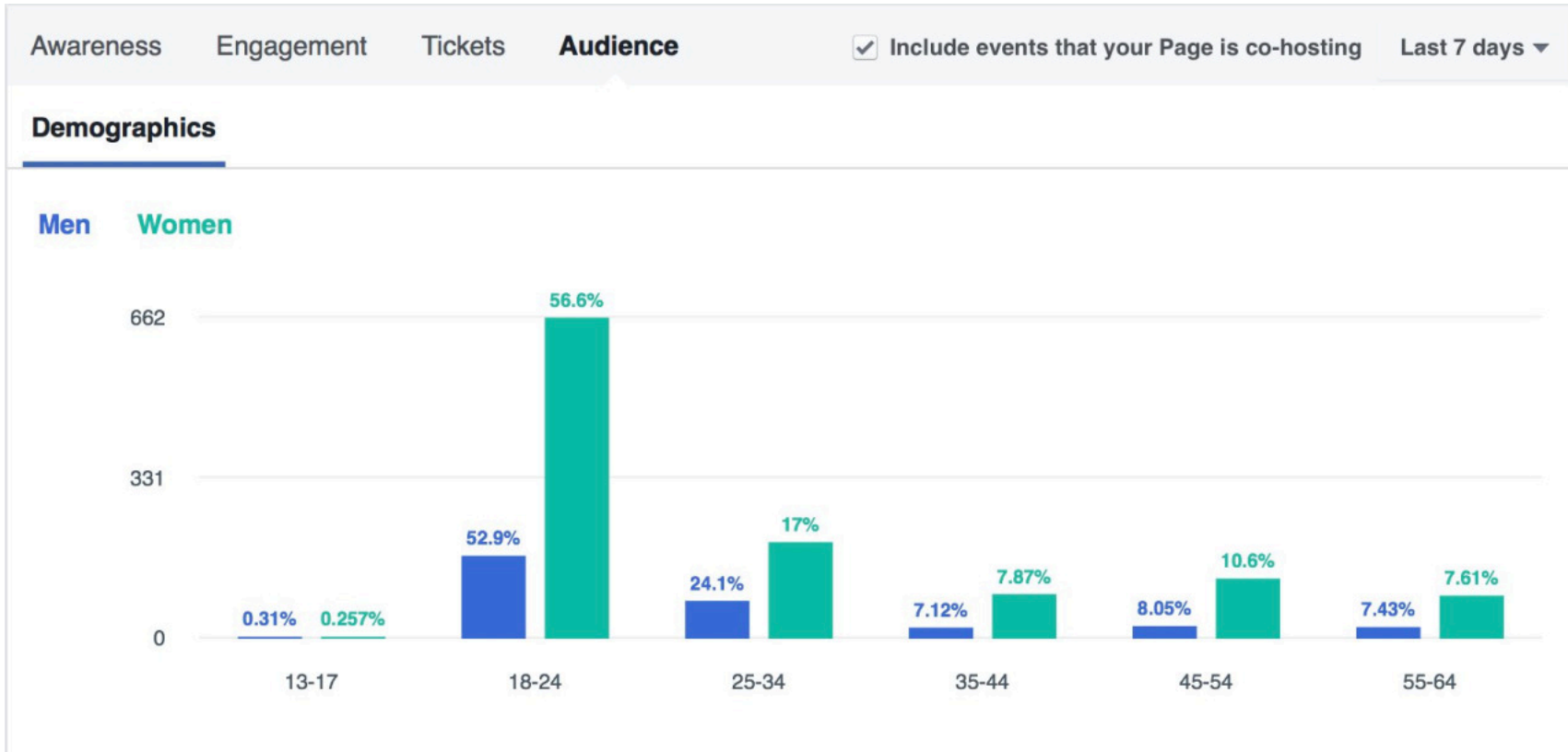
A FEW EXAMPLES AND CASE STUDIES

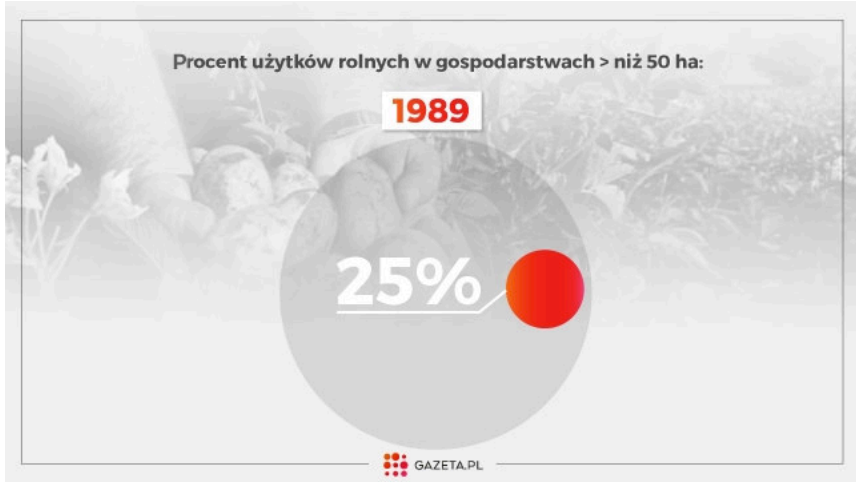
2016 Q4 GROWTH UPGRADED



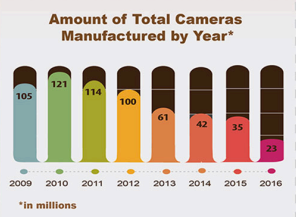
Source: ONS

The Office for National Statistics (ONS) said gross domestic product (GDP) expanded by 0.7 per cent in the fourth quarter - an increase from the 0.6 per cent calculated on the watchdog's first look at the economy
Source: <http://www.dailymail.co.uk/news/article-4248690/Economy-grew-0-7-final-three-months-2016.html>

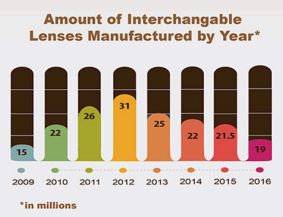




CAMERA INDUSTRY FACTS 2009-2016



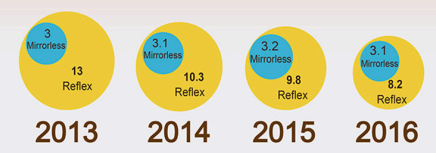
35% DROP IN SHIPPED CAMERAS IN 2016



12% DROP IN SHIPPED LENSES IN 2016

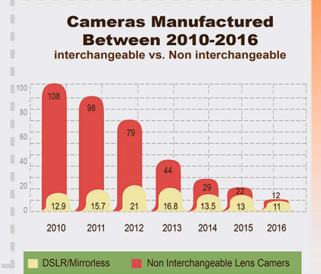
DSLR vs. Mirrorless 2013-2016

*in millions

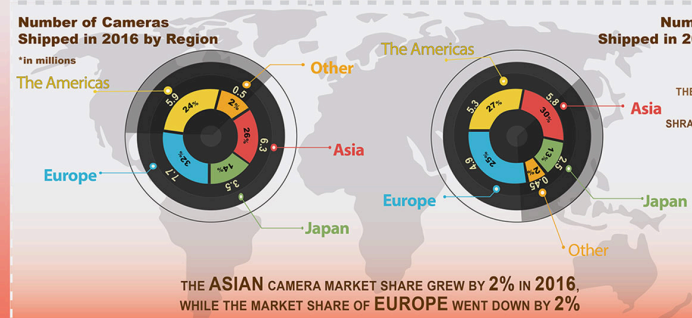


4% DECREASE IN MIRRORLESS PRODUCED & 17% DROP IN DSLR PRODUCED IN 2016

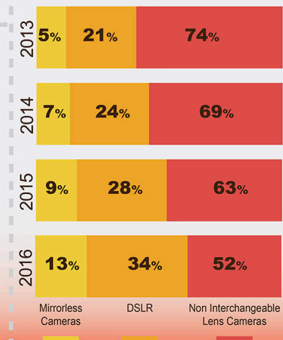
Based on CIPA (Camera & Imaging Products Association), Shipment of Digital Still Cameras & Lenses Data



THE ENTIRE CAMERA MARKET IN 2016 SAW 81% DROP COMPARED TO 2010



Camera Market Overview 2013-2016



Lensvid.com is THE place to find the most interesting, informative, professional and inspiring photography videos on the web.

make sure you follow us on:
 Facebook.com/Lensvid
 Twitter.com/LensVidcom



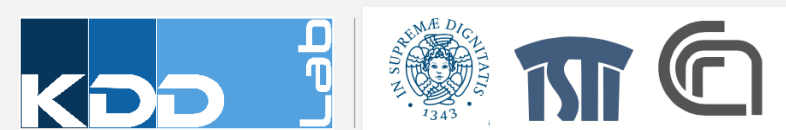
VISUALIZATION TAXONOMY

Borkin MA, VoAA, Bylinskii Z, Isola P, Sunkavalli S, Oliva A, Pfister H.

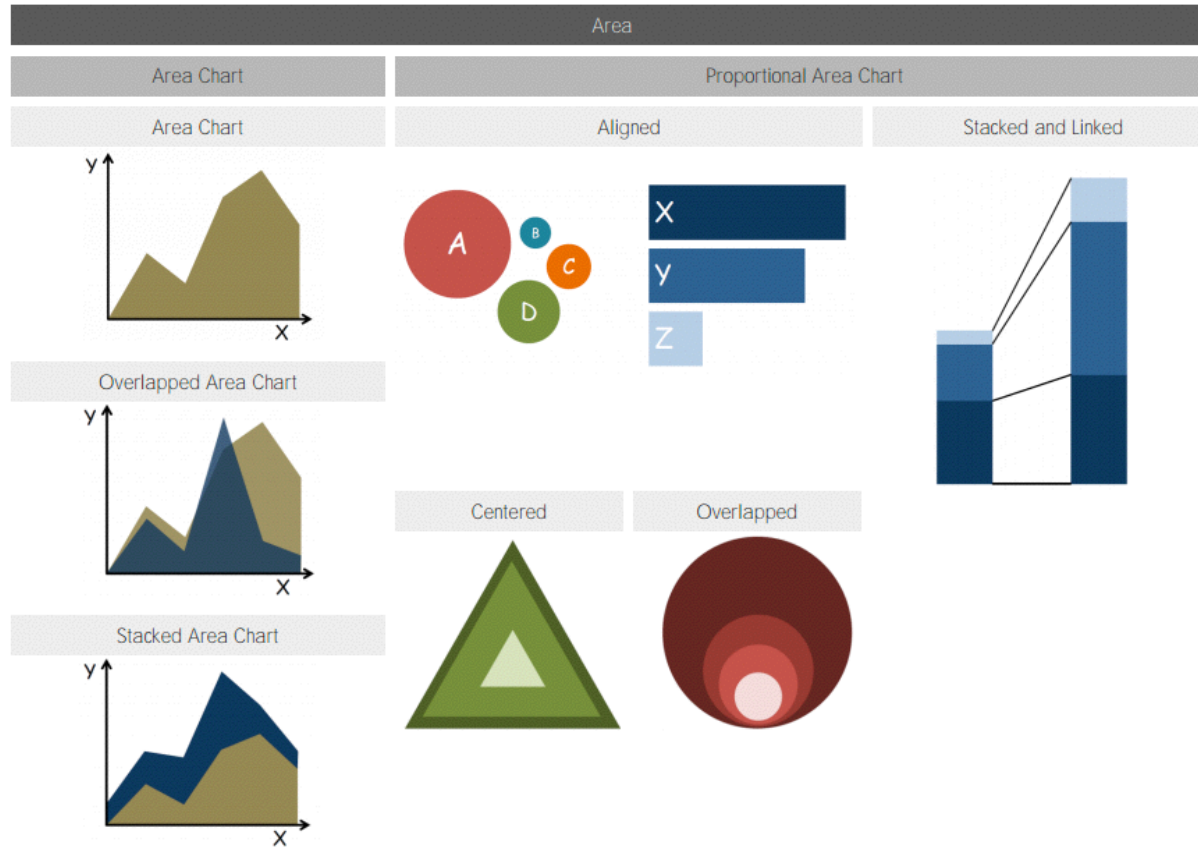
What Makes a Visualization Memorable?

IEEE Transactions on Visualization and Computer Graphics (InfoVis 2013).

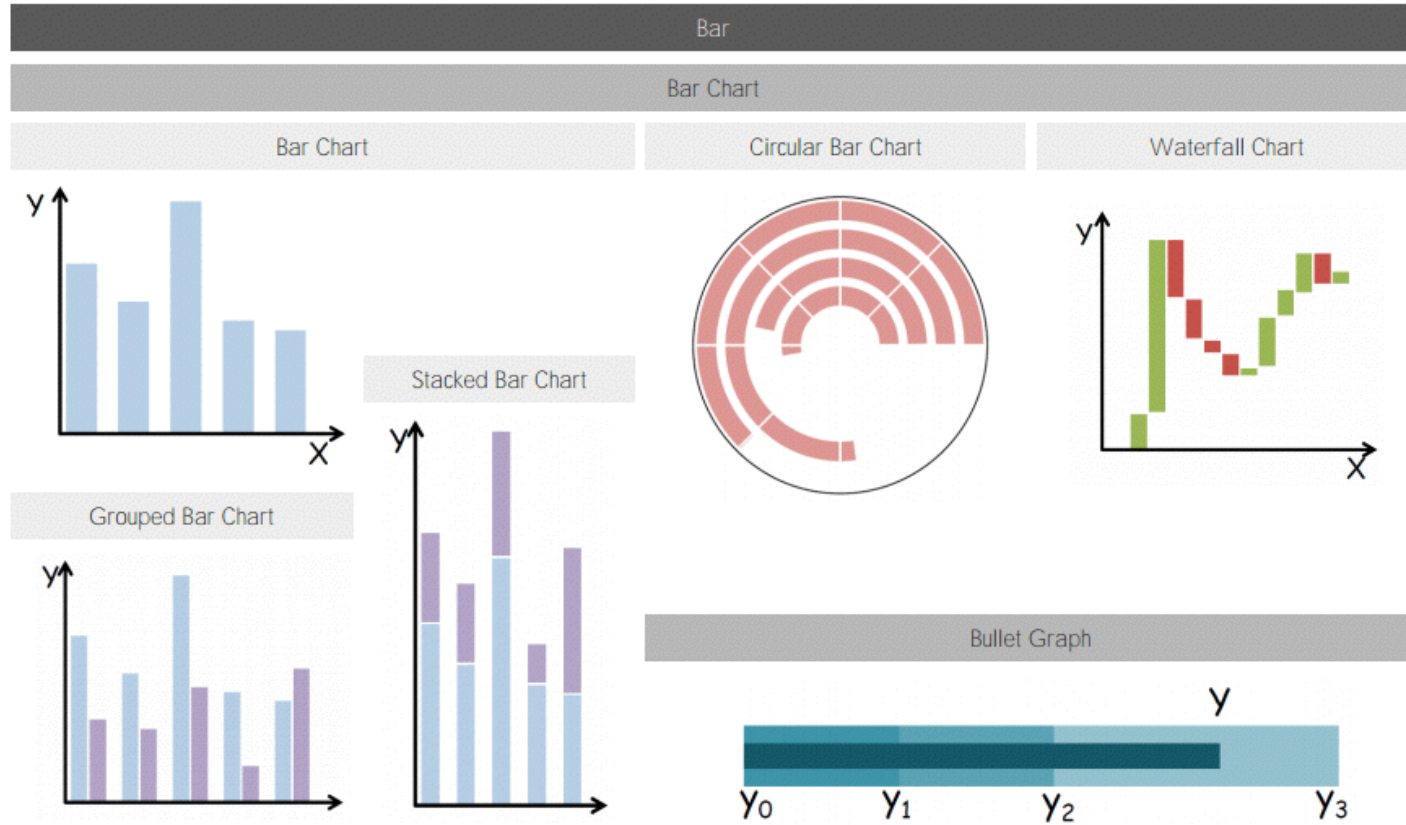
<http://vcg.seas.harvard.edu/publications/what-makes-visualization-memorable>

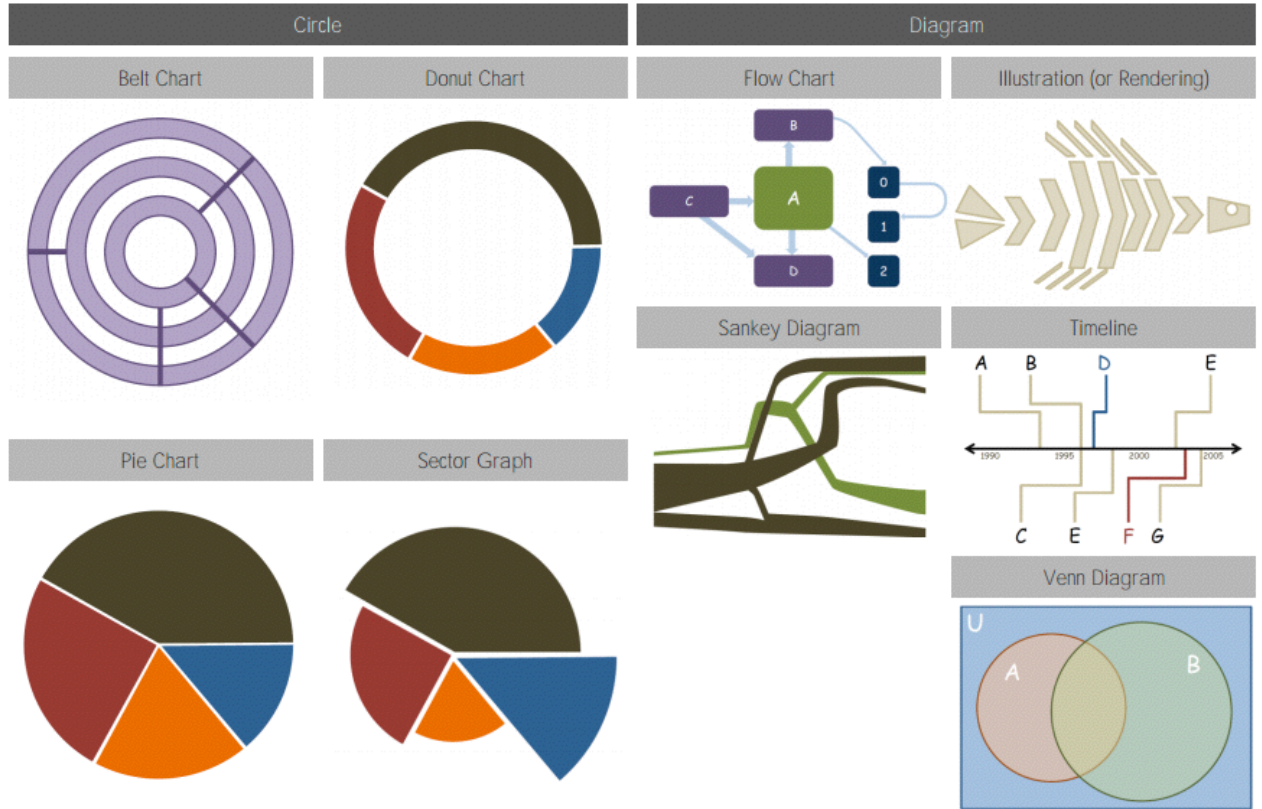


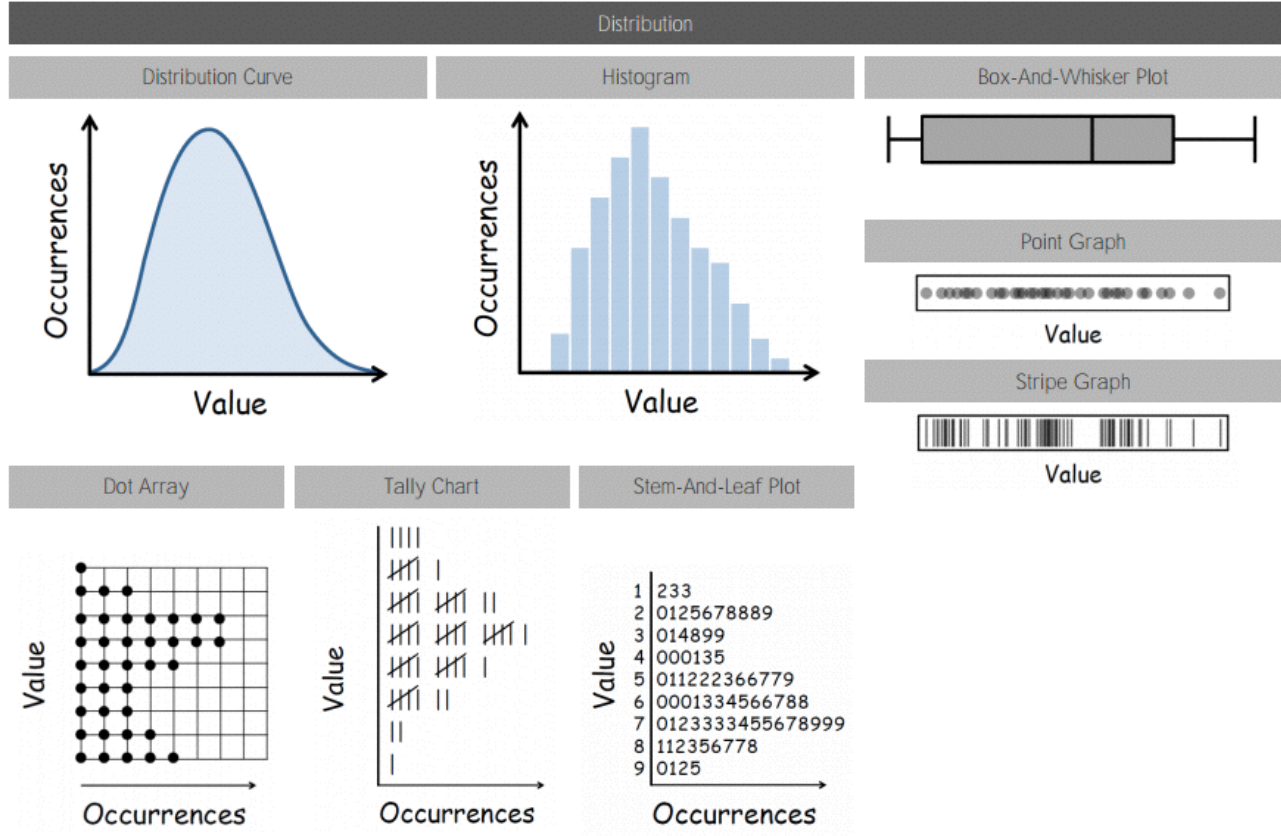
AREA

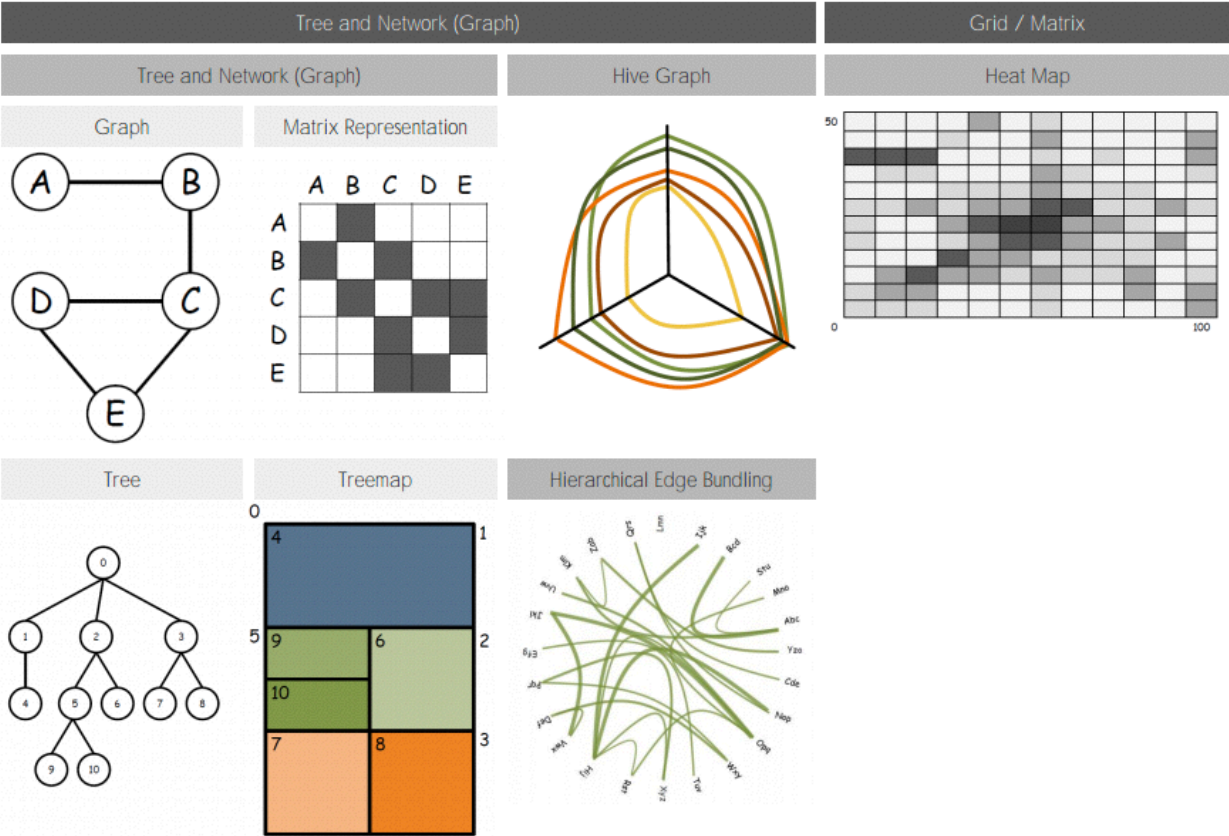


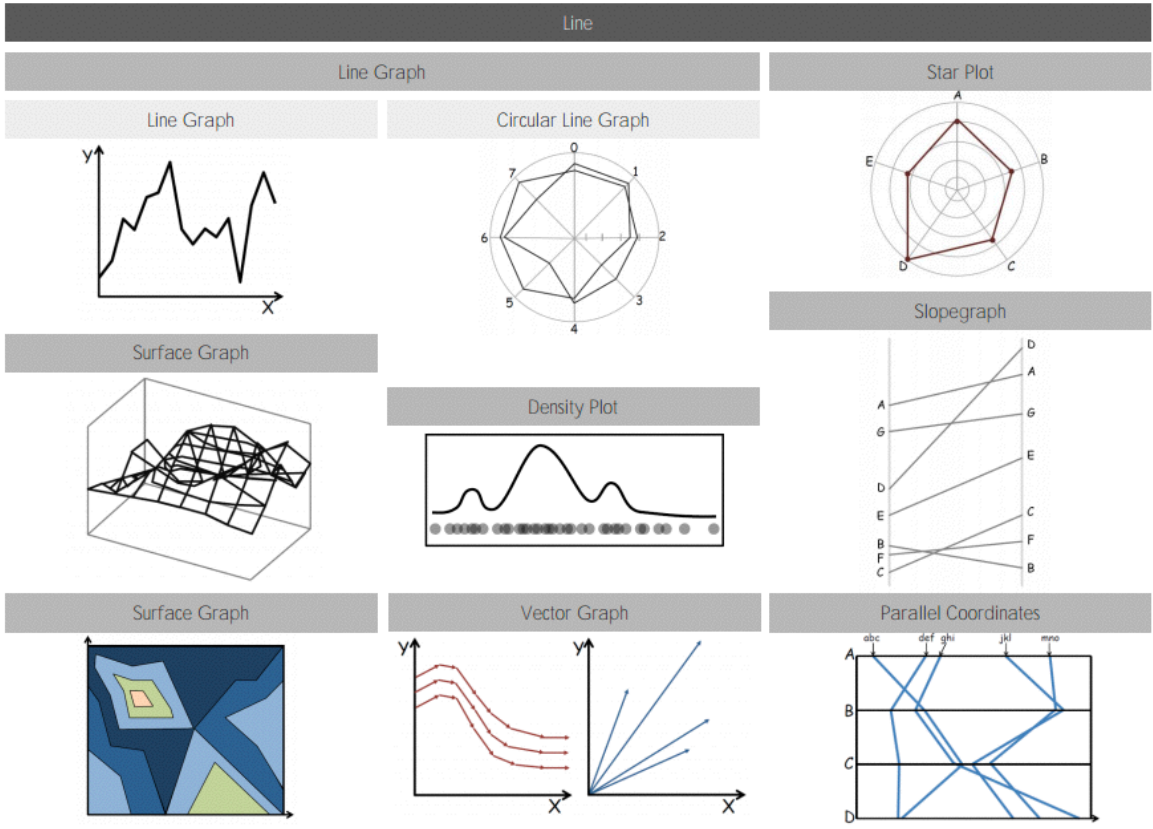
BAR











Map

Geographic Map

Flow Map

Geographic Map

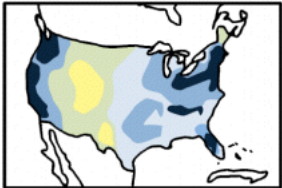


Statistical Map

Street Map

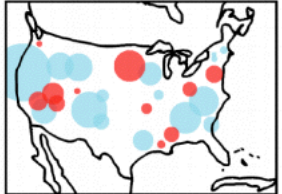
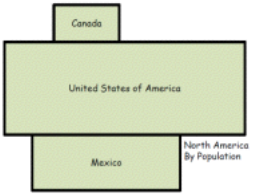
Choropleth Map

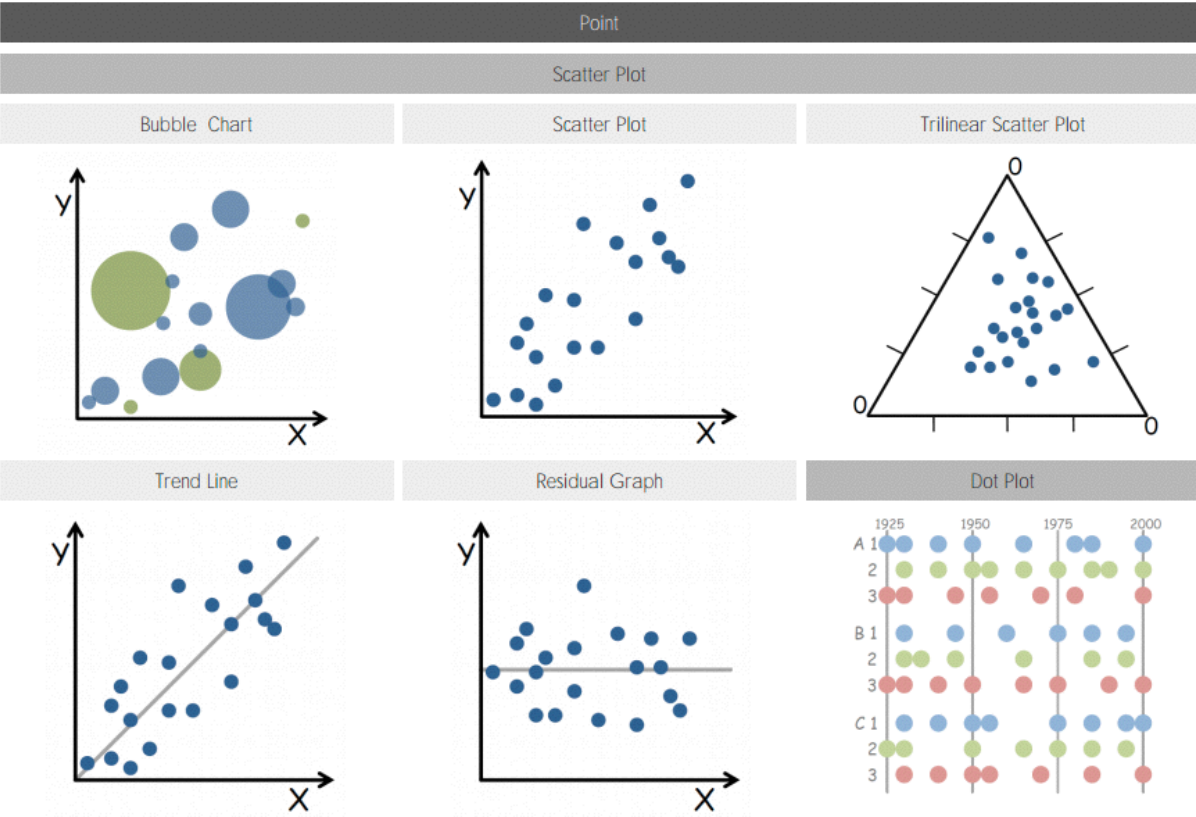
Contour Map (Isopleth)



Distorted Map (Cartogram)

Statistical Plot Map





Table

Table

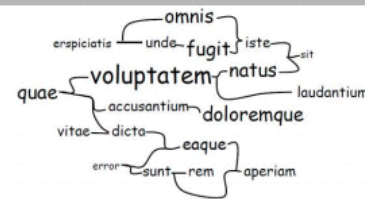
ABC	1234	X45
Category	543.2109	7%
Group	45.67	45%
Unit	9876	98%
Class	123.78	12%

Text Chart

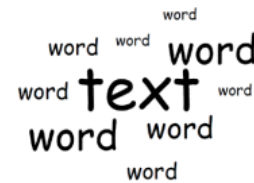
<p><i>Title</i></p> <ul style="list-style-type: none"> •Sed dignissim vehicula •Nisi quis congue •Sed vitae rhoncus odio •Integer at odio 	<p><i>Heading 1</i></p> <p>"Nunc aliquam turpis at tellus varius hendrerit. Ut nec magna terten. Praes adipiscing dolor eget odio semper ut commodo locus imperdiet."</p> <p>- Lorem</p>
<p><i>Heading 2</i></p> <p>Aenean tincidunt sem vel massa cursus non tempus quam auctor. In nisi mi, commodo sit.</p> <p>Amet rutrum vitas, fringilla non urna. Quisque sagittis ultrices sapien, quis posuere massa interdum quis.</p>	<p><i>Heading 3</i></p> <ul style="list-style-type: none"> ✓Chart 1 ✓Chart 2 ✓Chart 3 ✓Chart 4

Text Based

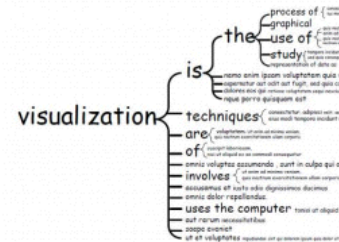
Phrase Net



Word Cloud



Word Tree



VISUAL TAXONOMY

The Data Visualisation Catalogue

About • Suggest • Shop • Resources

Search by Function

View by List



Arc Diagram



Area Graph



Bar Chart



Box & Whisker Plot



Brainstorm



Bubble Chart



Bubble Map



Calendar



Chord Diagram



Choropleth Map



Circle Packing



Connection Map



TAKEAWAY MESSAGES

- Appropriate chart type for specific data type and visualization task