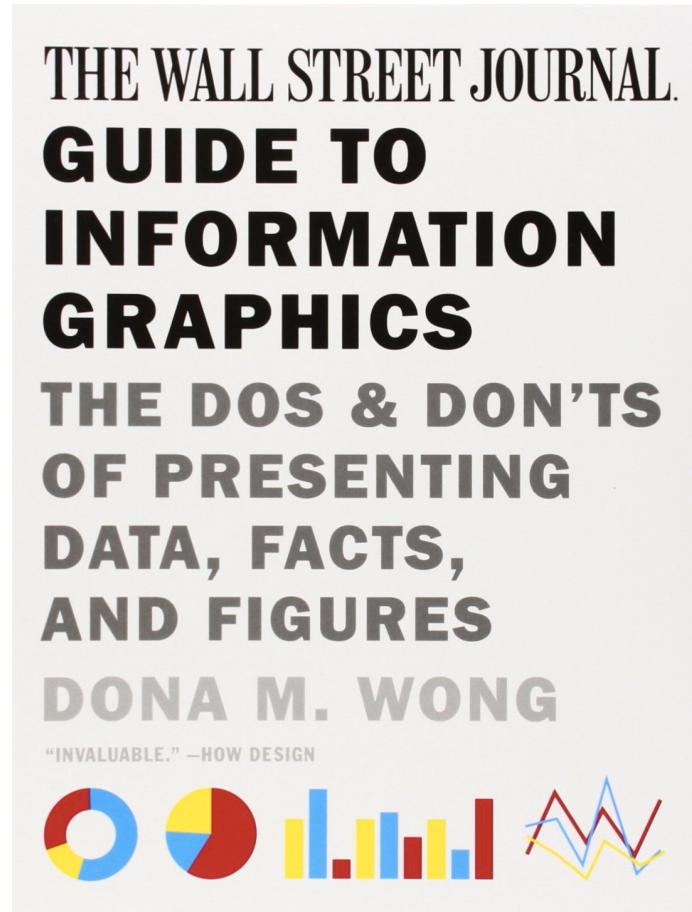


6 CHARTING TAXONOMY

S. Rinzivillo – rinzivillo@isti.cnr.it

CRASH COURSE ON EFFECTIVE CHARTING



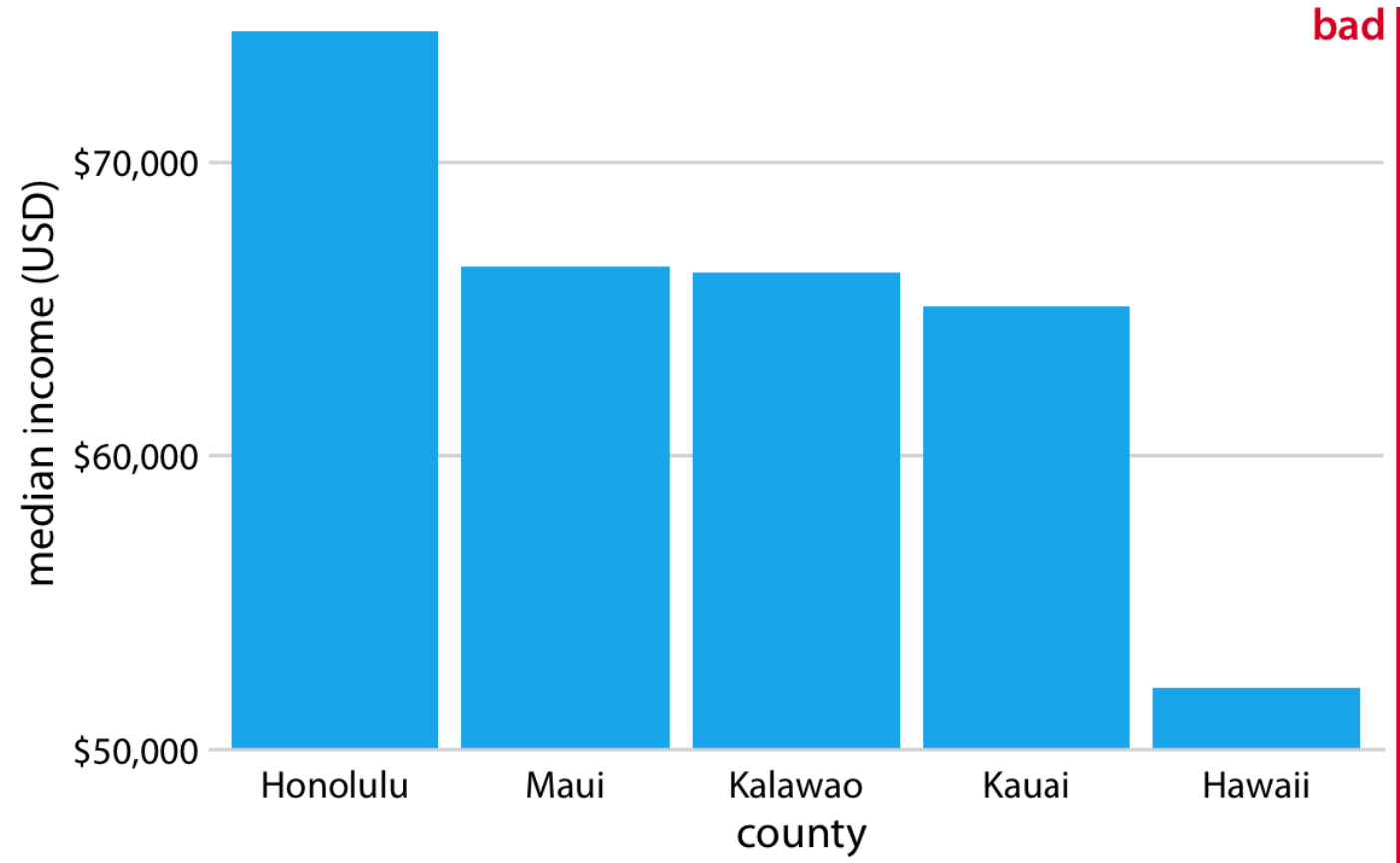
Dona M. Wong

Guide to Information Graphics

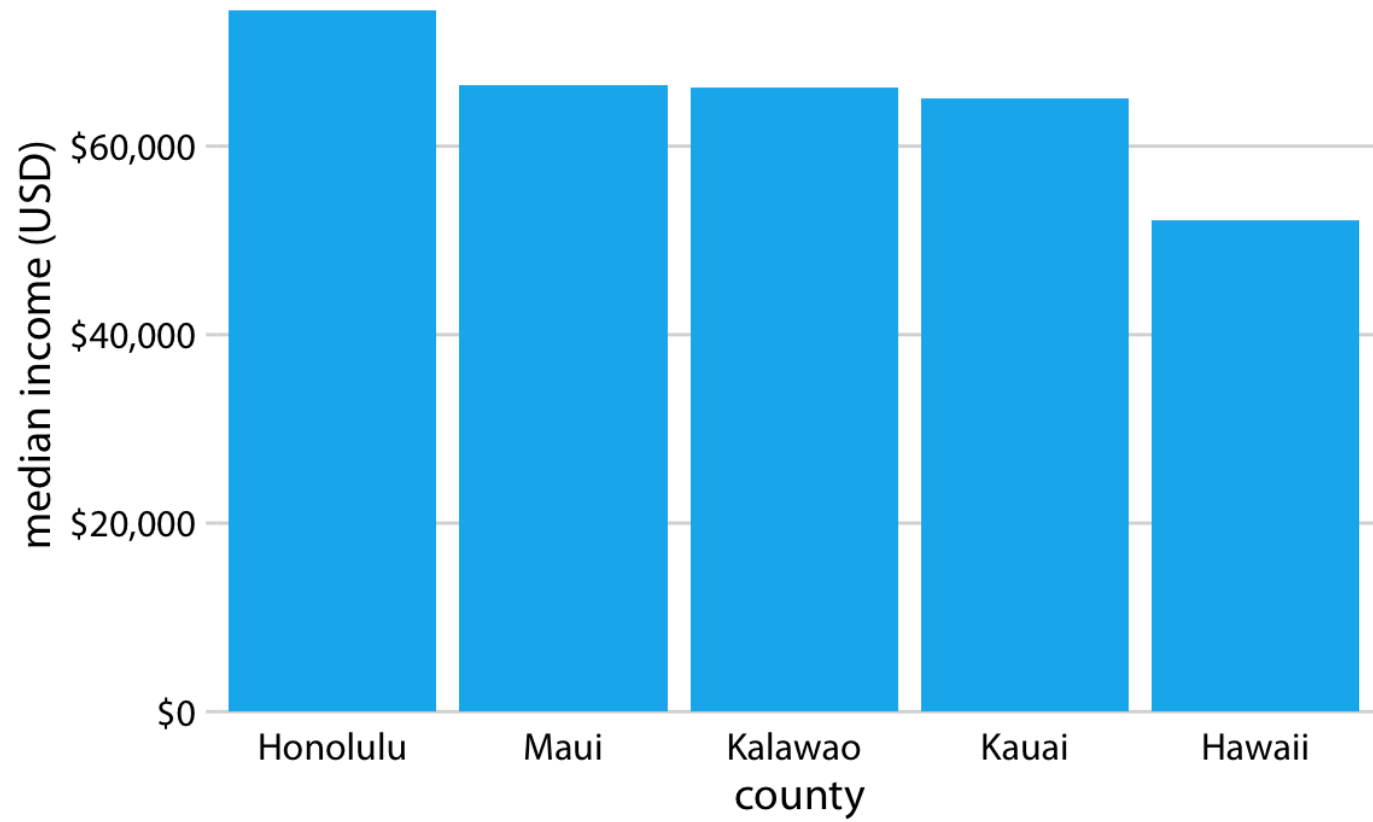
The Dos and Don'ts of Presenting Data, Facts, and Figures

W. W. Norton & Company

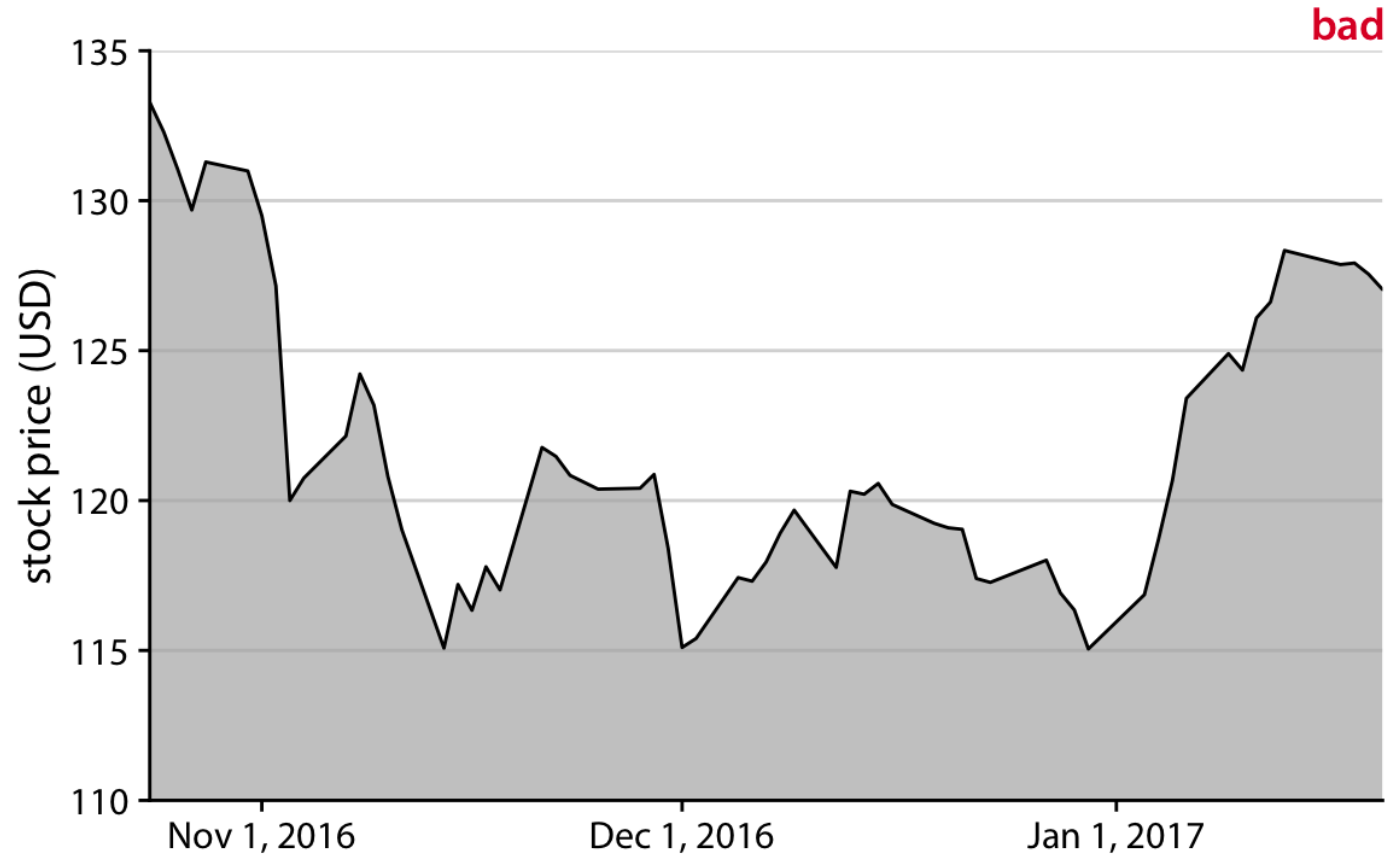
CHARTING EXAMPLES



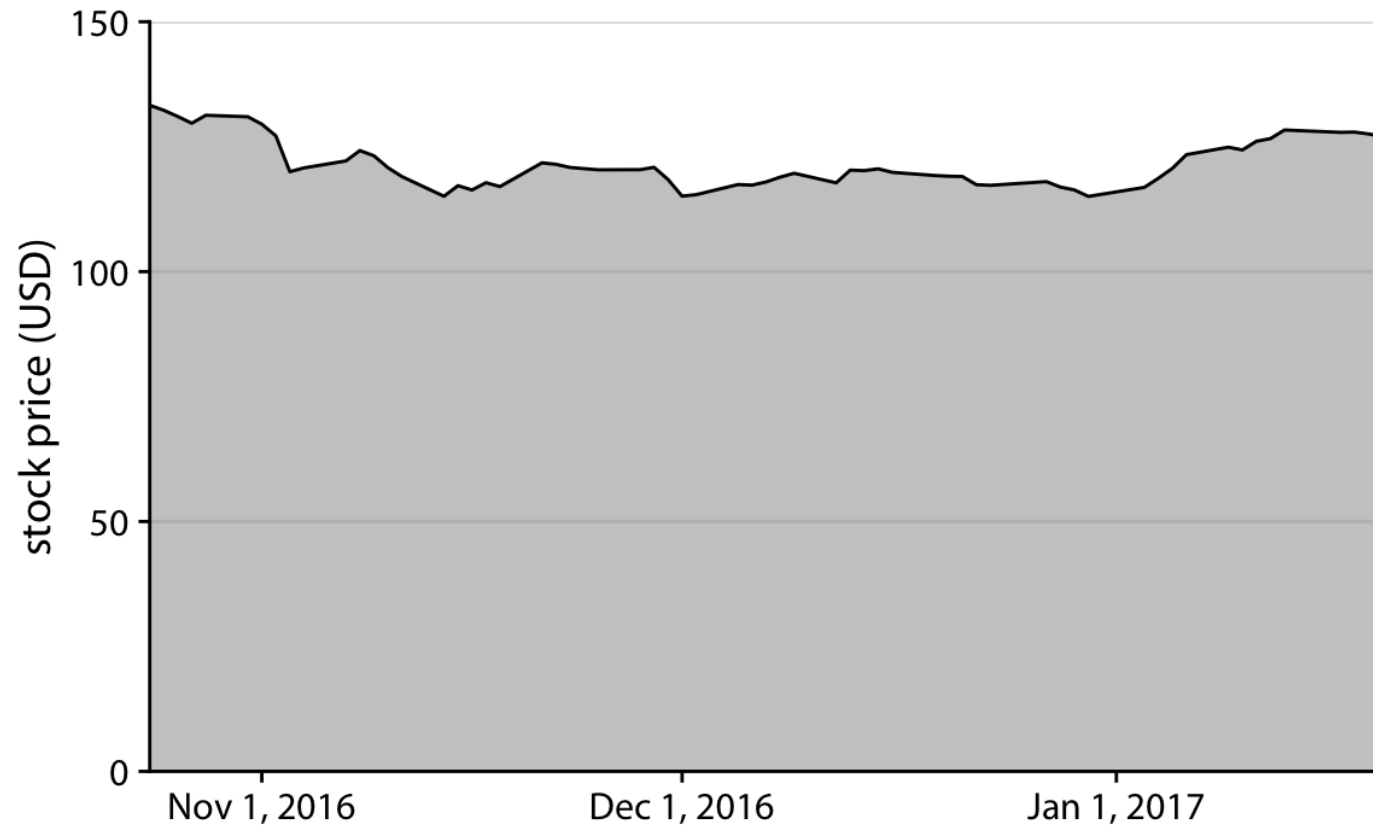
CHARTING EXAMPLES



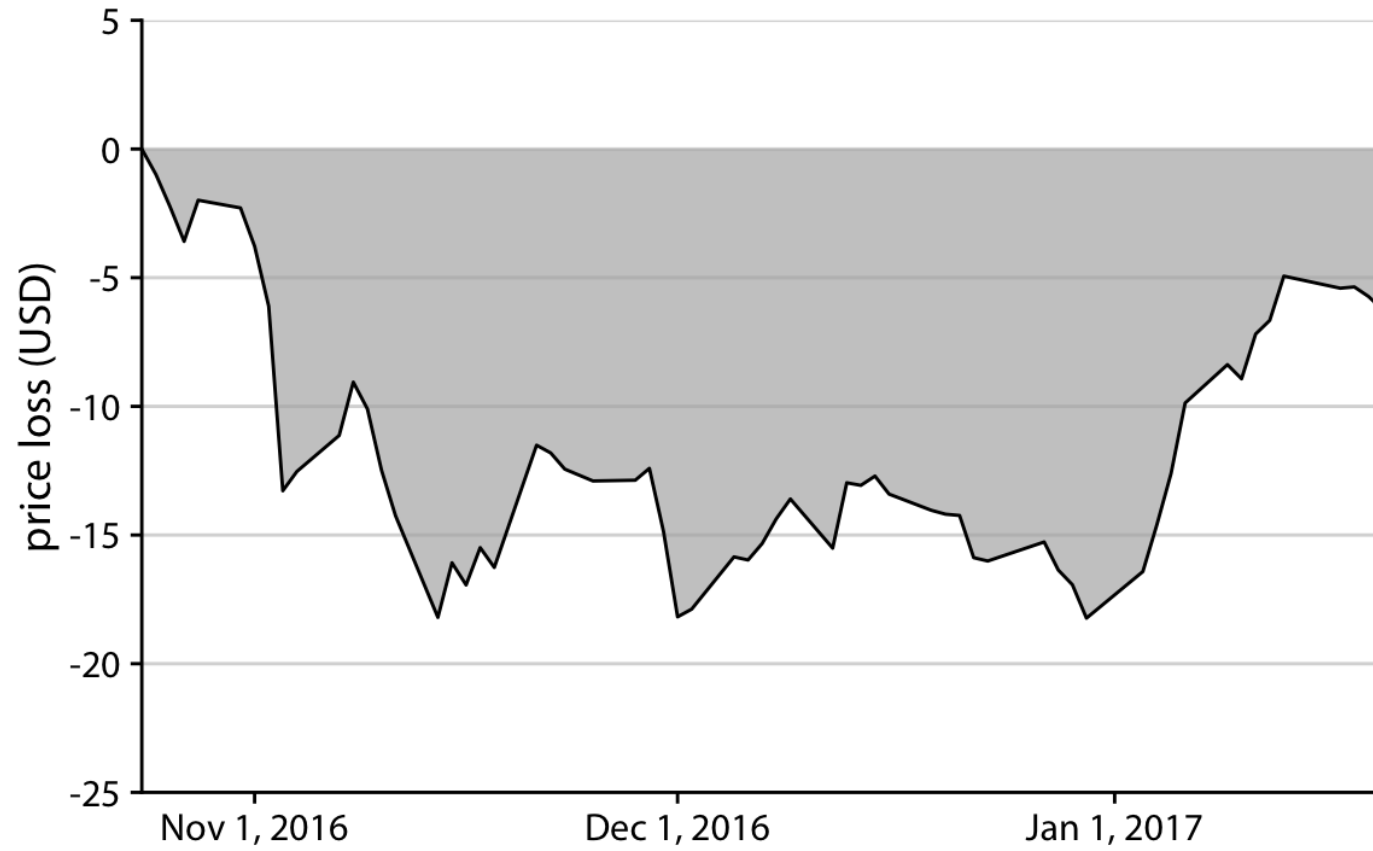
CHARTING EXAMPLES



CHARTING EXAMPLES

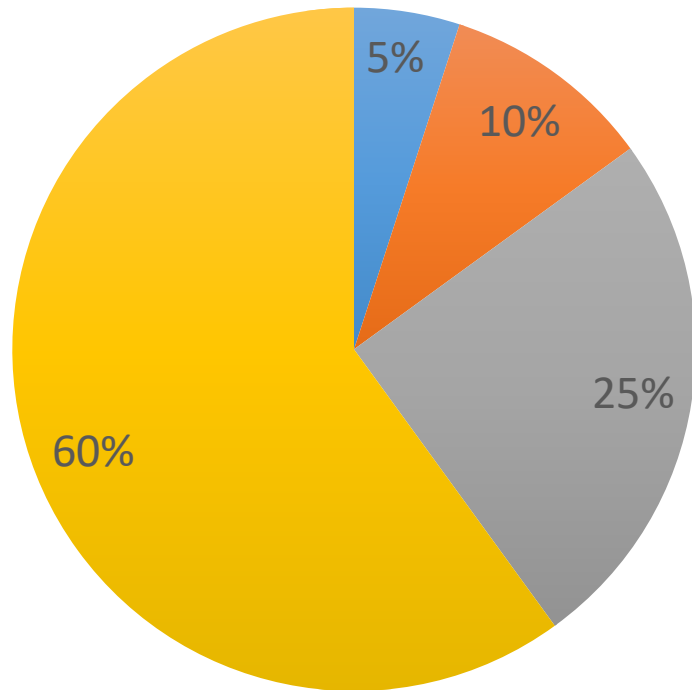


CHARTING EXAMPLES

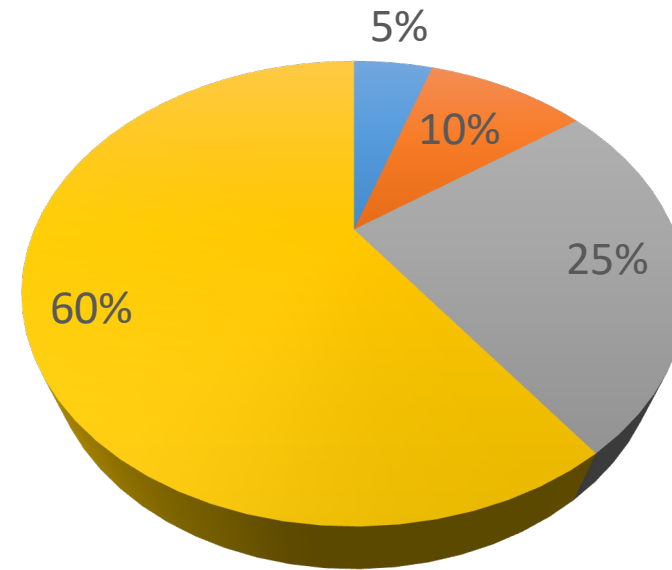


CHARTING EXAMPLES

Sales

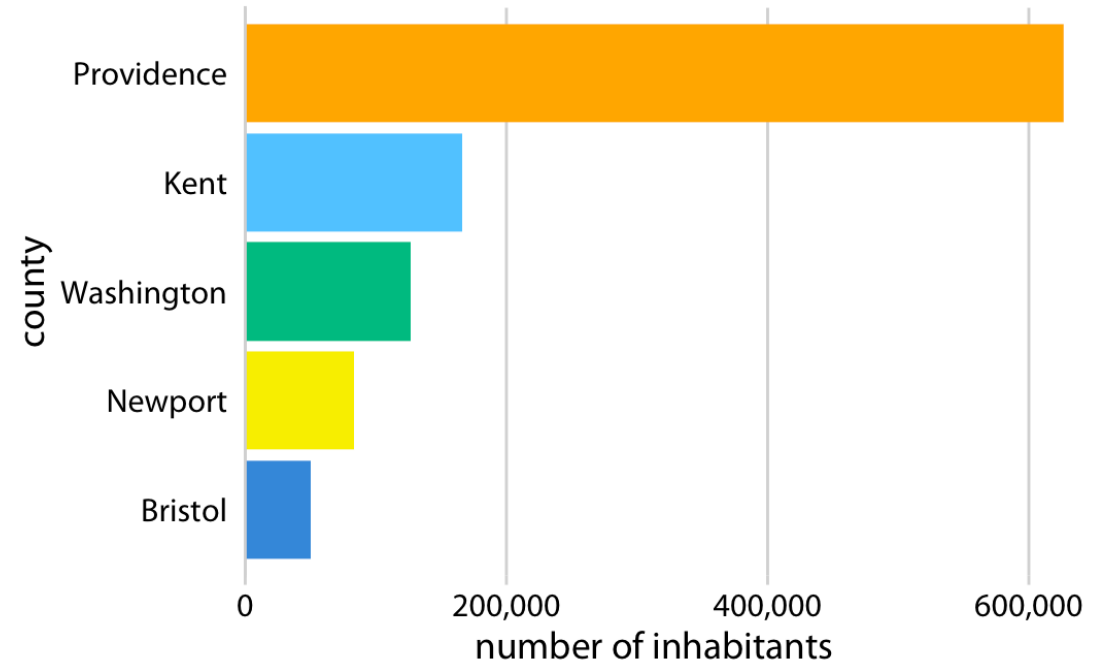
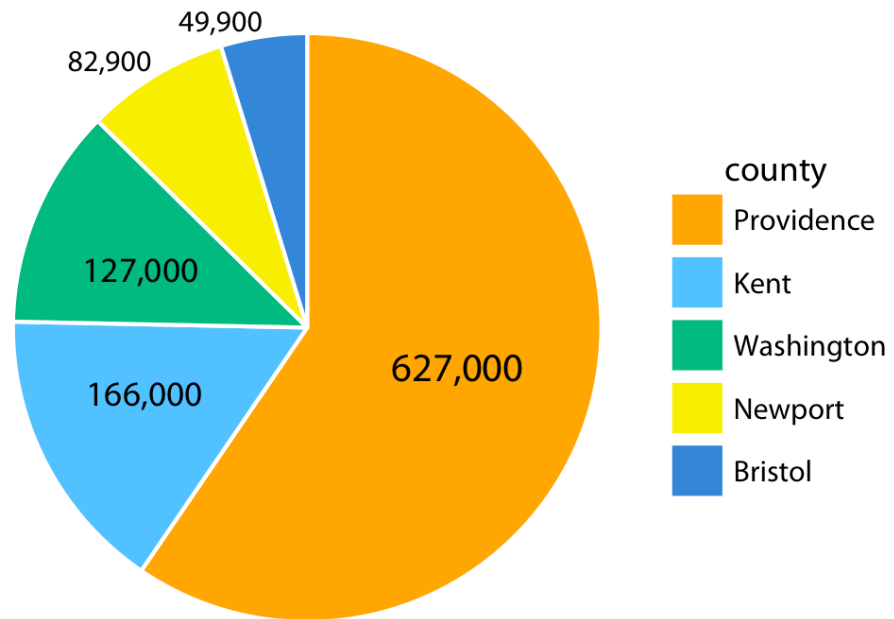


Sales



May these charts be improved? Why? How?

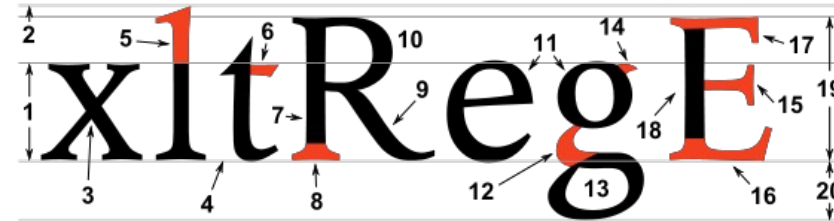
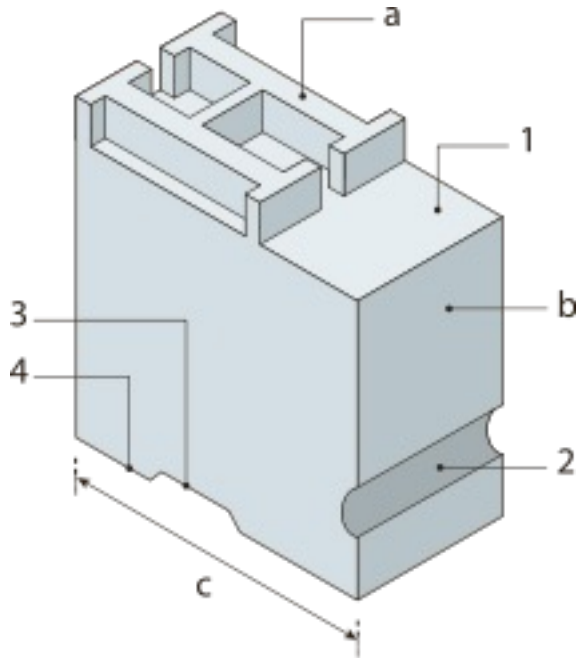
CHARTING EXAMPLES



FONTS



FONTS



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**.

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

FONTS: GENERAL RULES

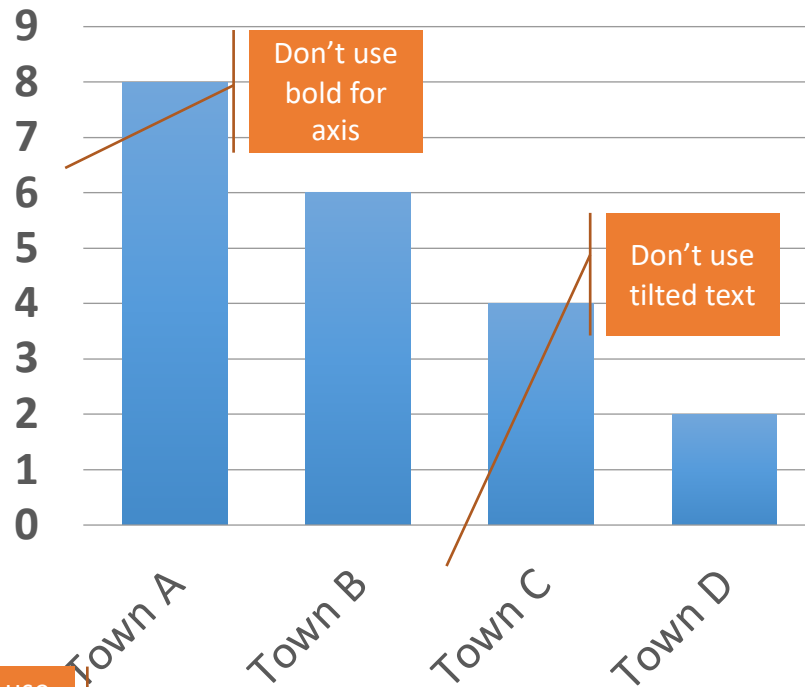
- Leading should be 2 points larger than type size
- Avoid too small or condensed type faces
- Keep style simple: use **bold** or *italic* to emphasize a word (better not ***both***)
- Avoid ALL CAPS
- Avoid *styled fonts*
- Avoid C***C Sans Serif
- Reduce type at an angle
- Avoid t r a c k i n g

TYPOGRAPHY IN CHARTS

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

Don't

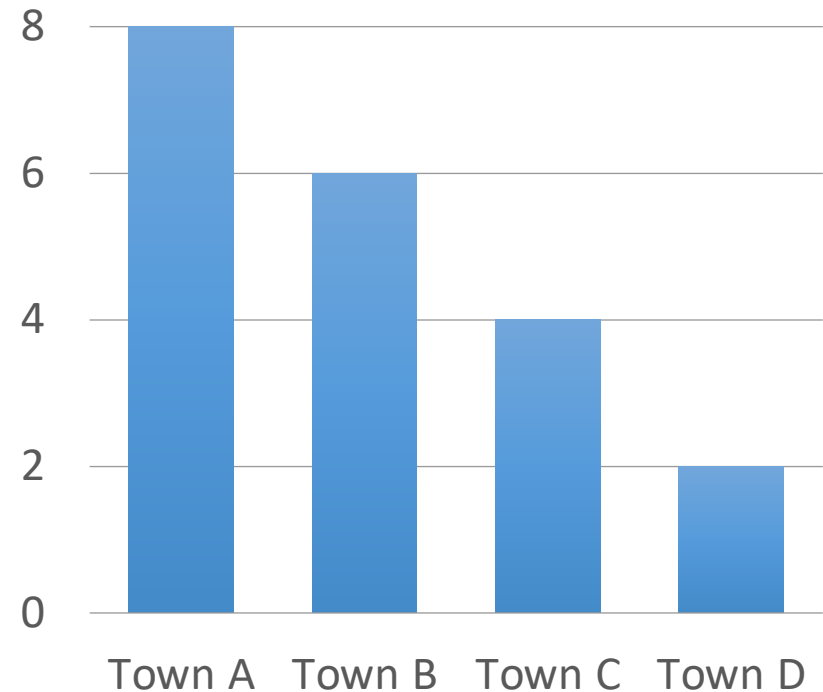
HEADLINE OF THE CHART



A brief description that outlines what the data shows

Do

Headline of the chart

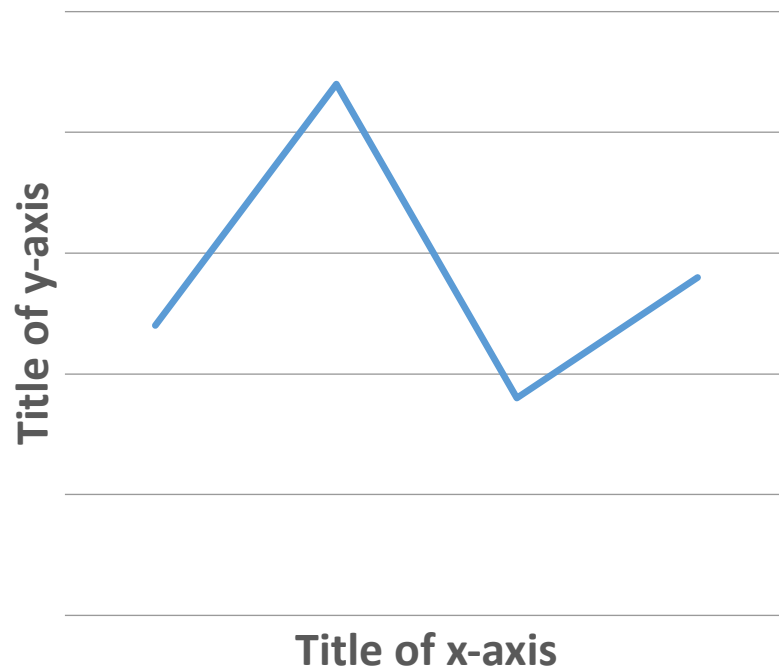


A brief description that outlines what the data shows

TYPOGRAPHY IN CHARTS

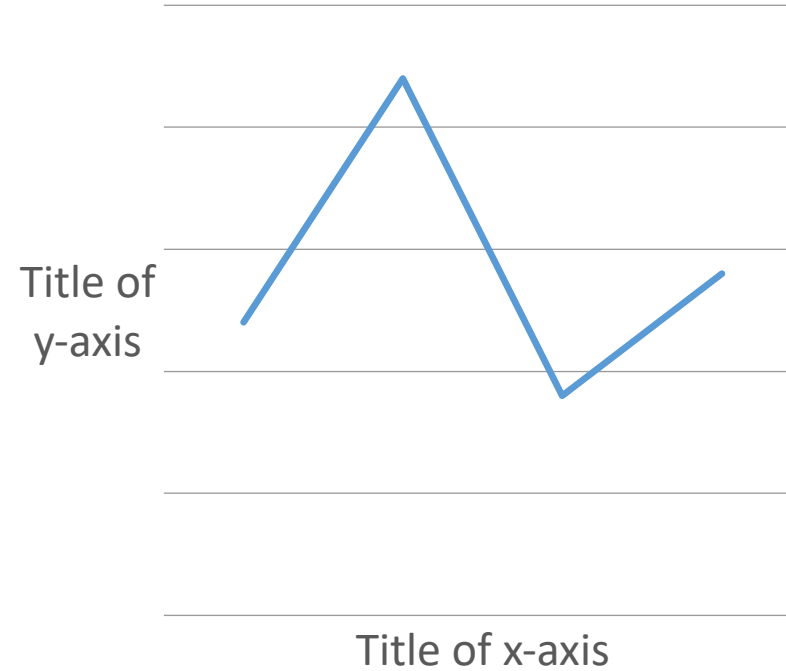
Don't

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

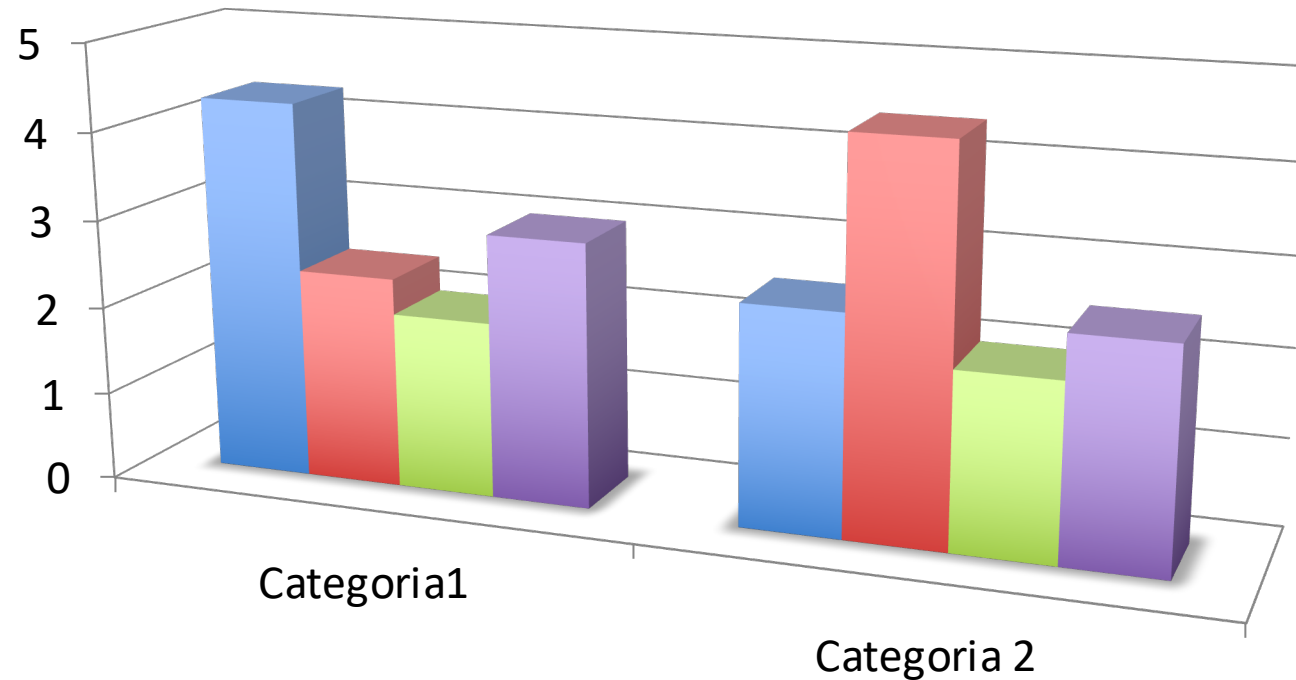
DATA-INK RATIO

Visual Display of Quantitative Data

Edward Tufte, 1983

DATA-INK RATIOⁿ

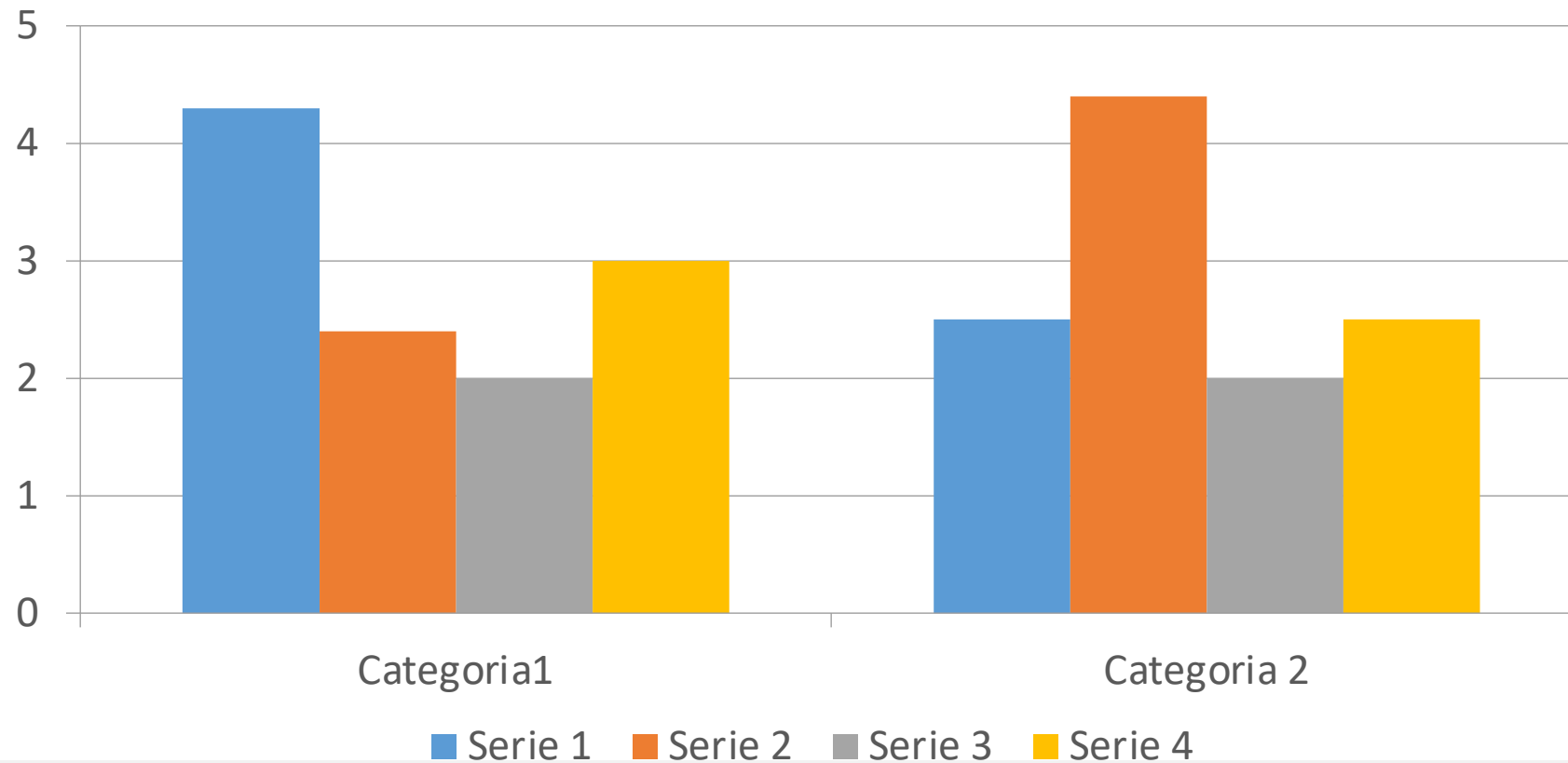
$$\text{Data-Ink Ratio} = \frac{\text{Data ink}}{\text{Total ink used in graphic}}$$



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

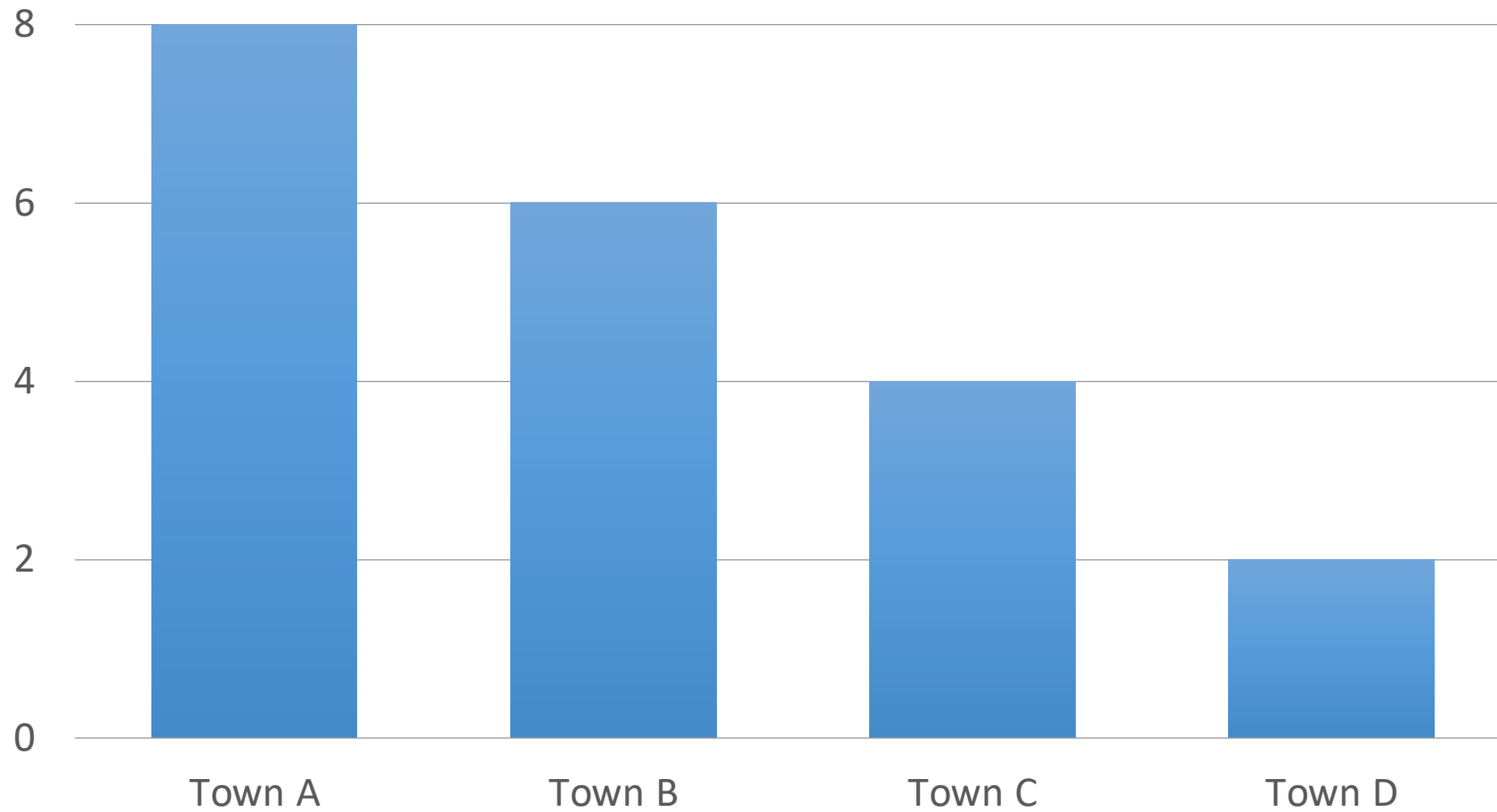
DATA-INK RATIOⁿ

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



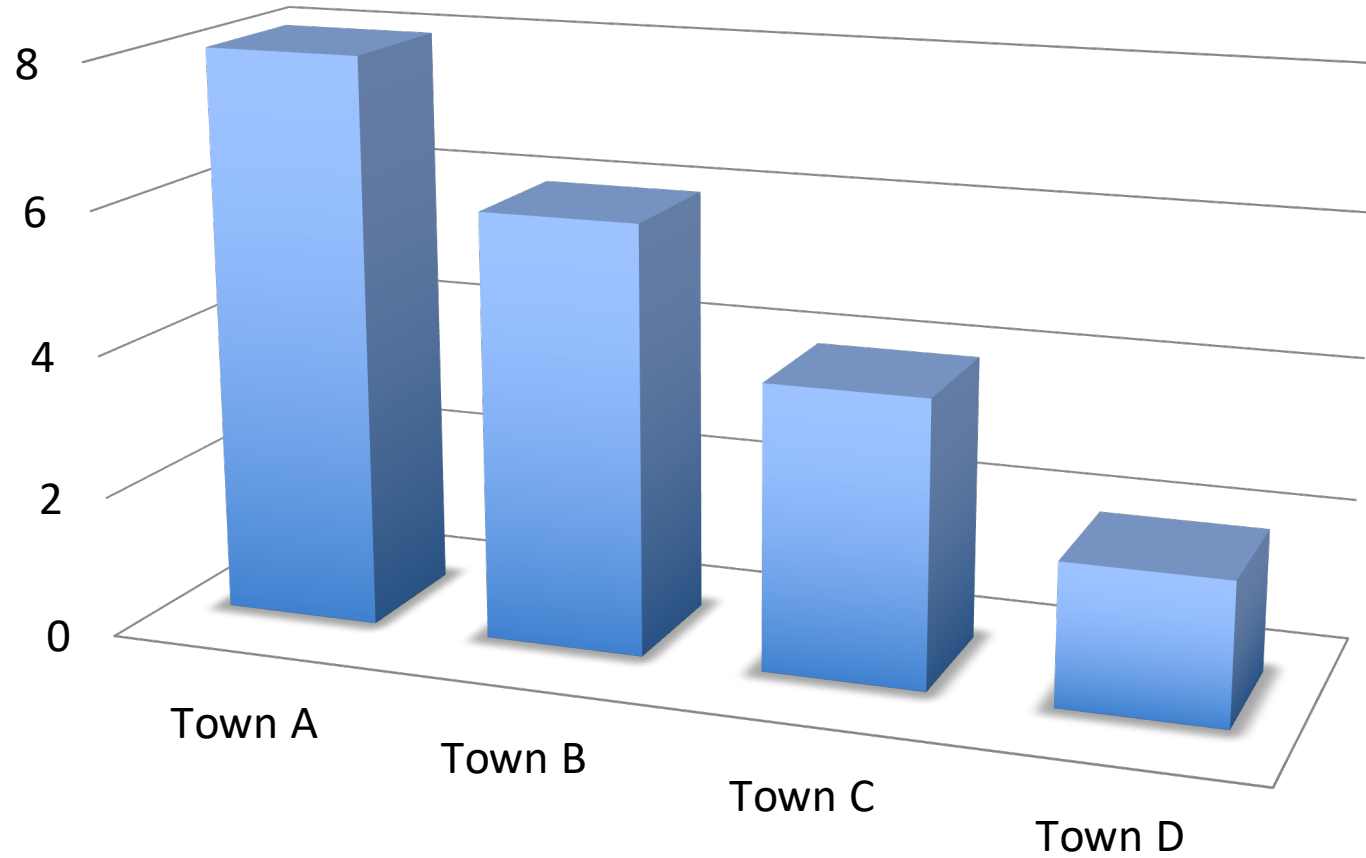
BAR CHARTS

Represent discrete quantities

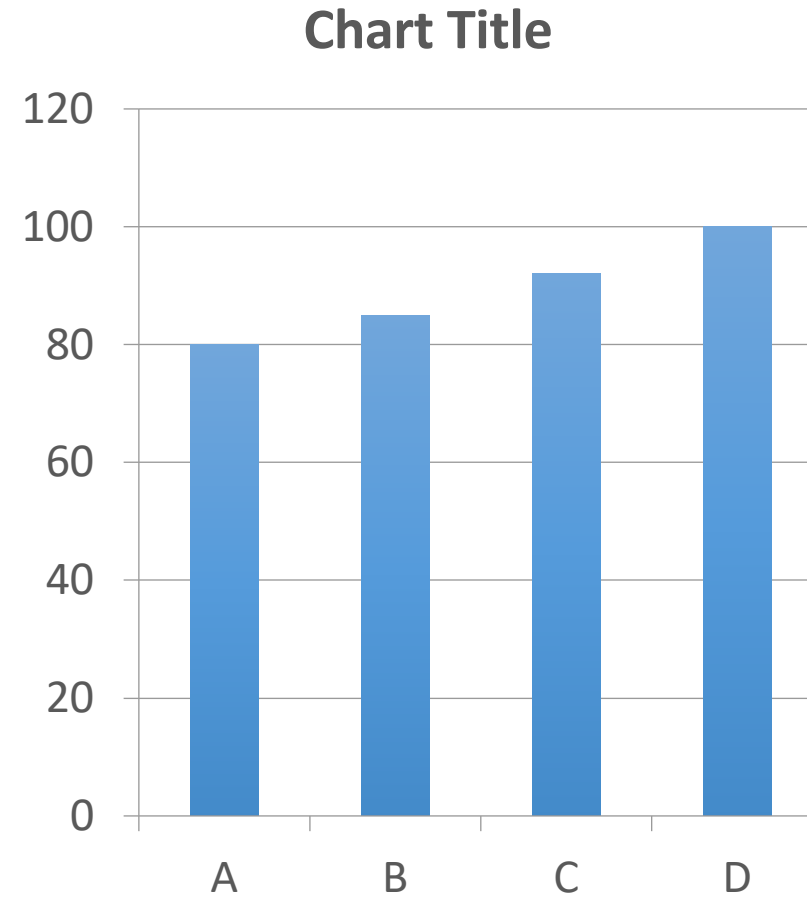
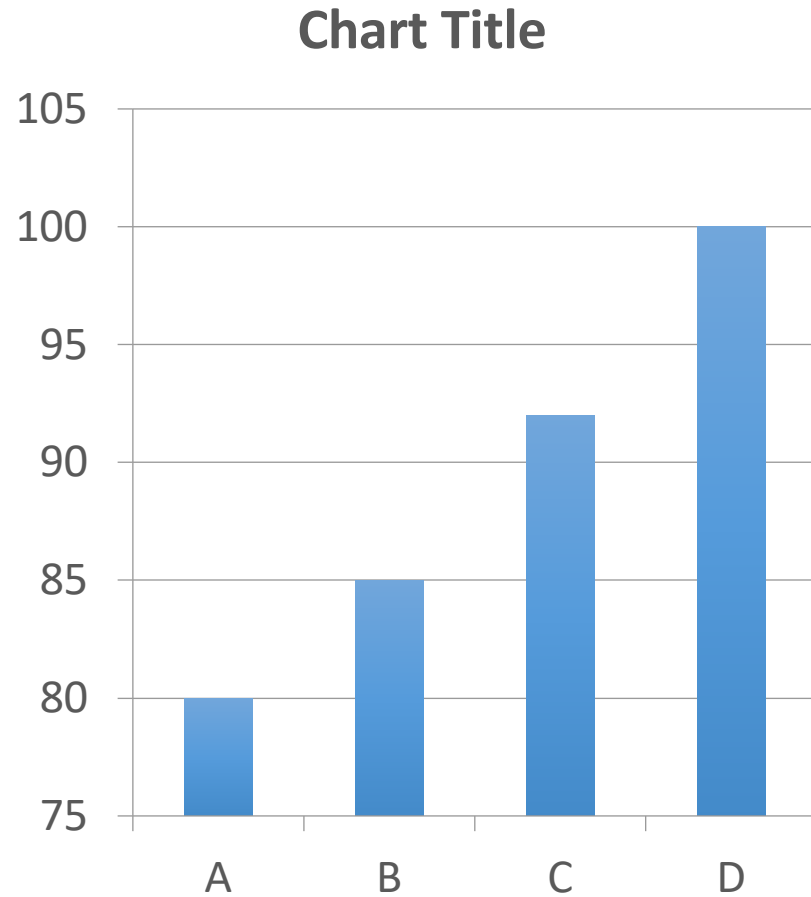


BAR CHARTS

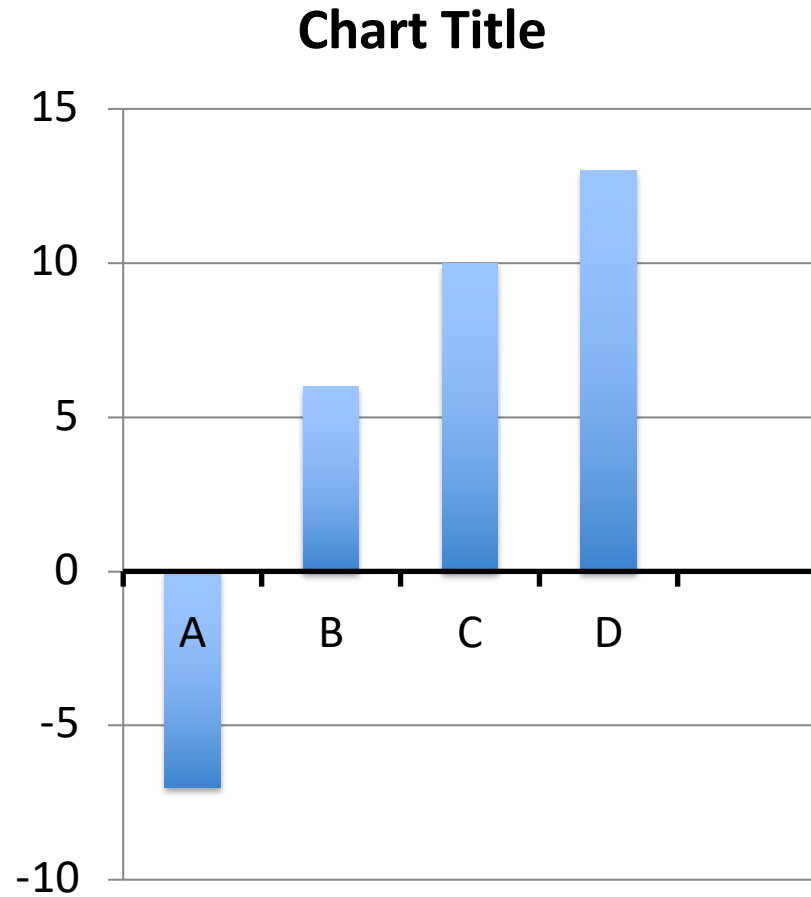
Avoid non-functional adornment



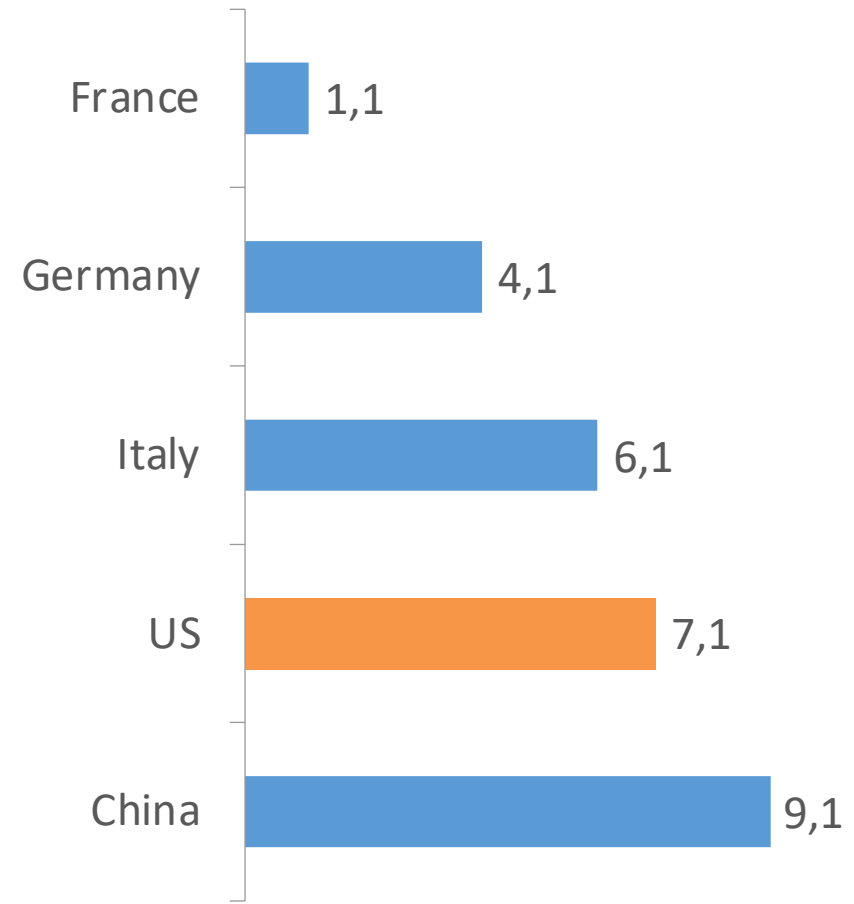
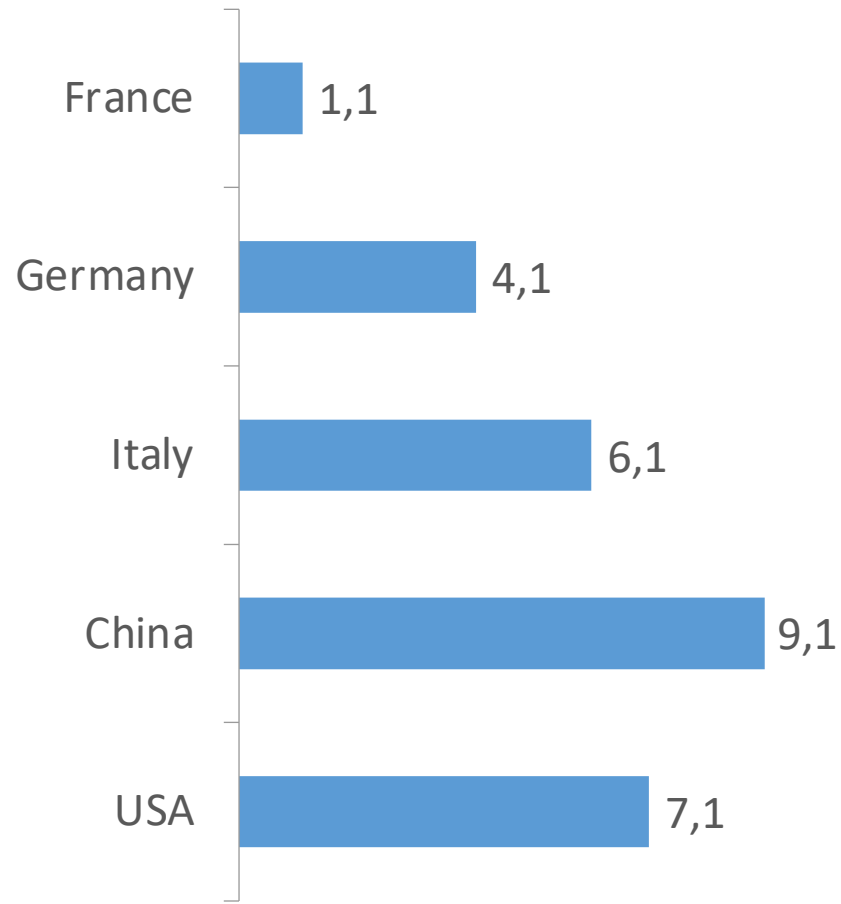
BAR CHARTS: BASELINE

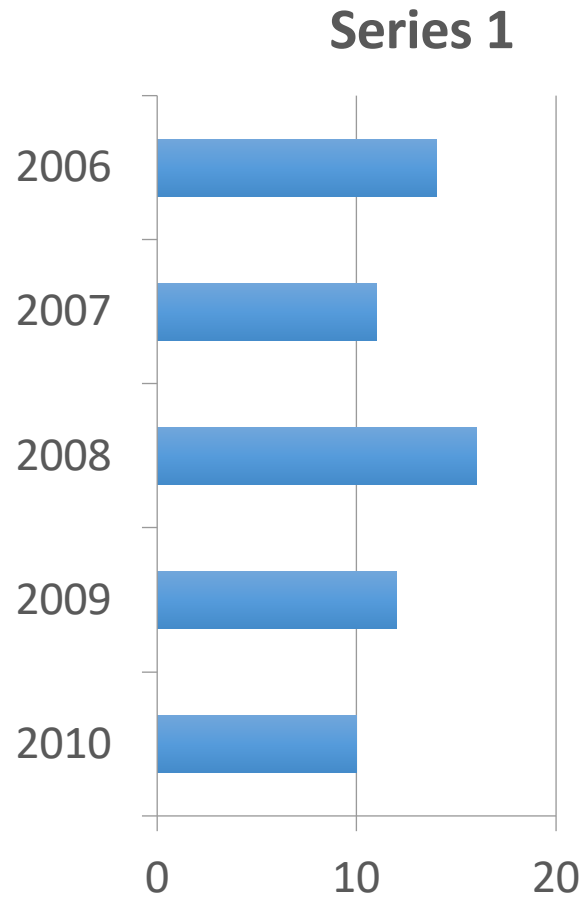


BAR CHARTS: BASELINE

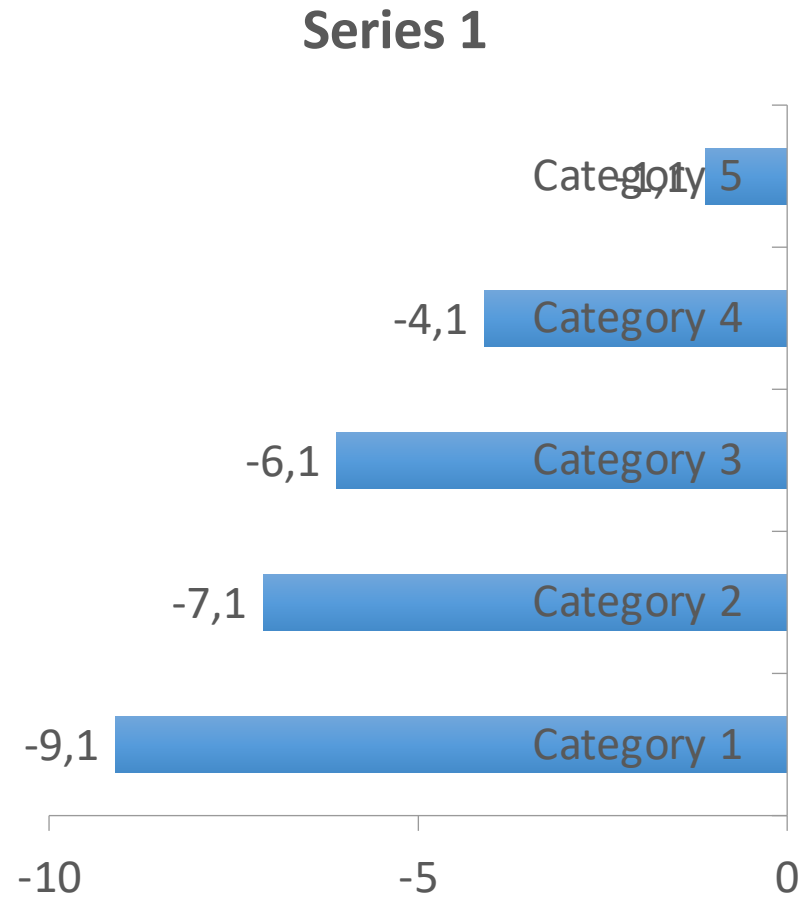


BAR CHARTS: ORDERING



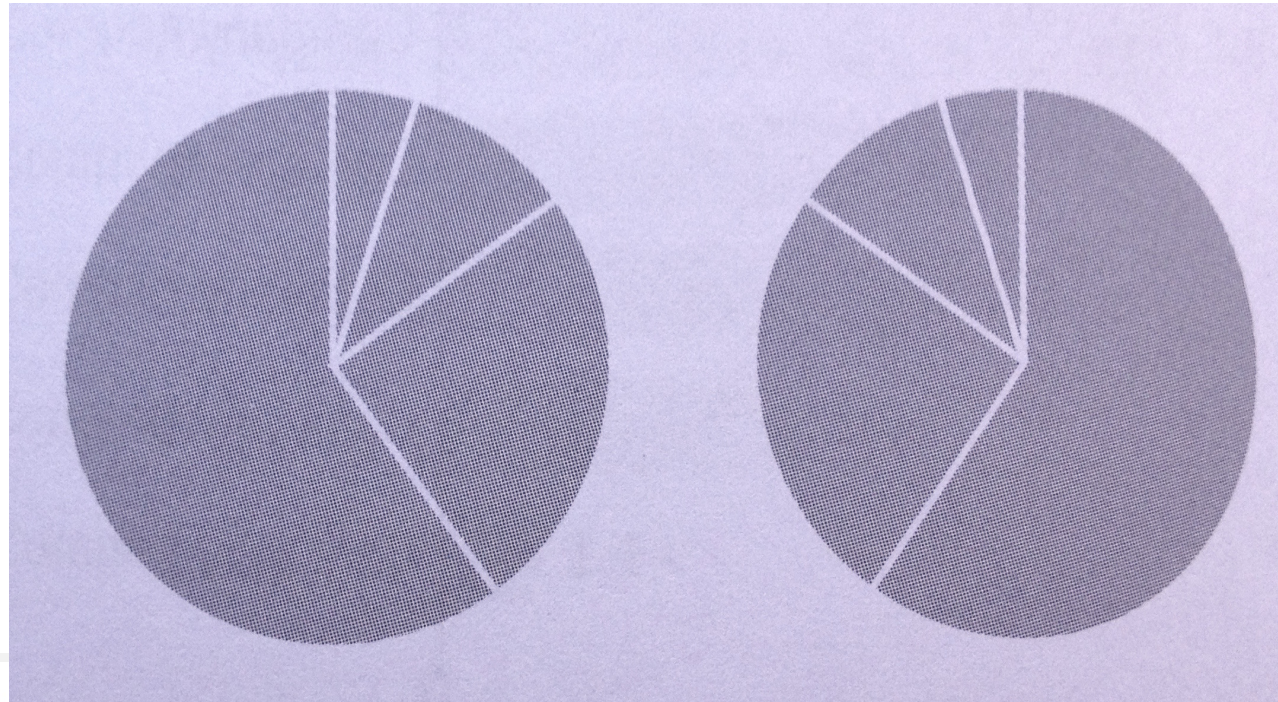


Series 1

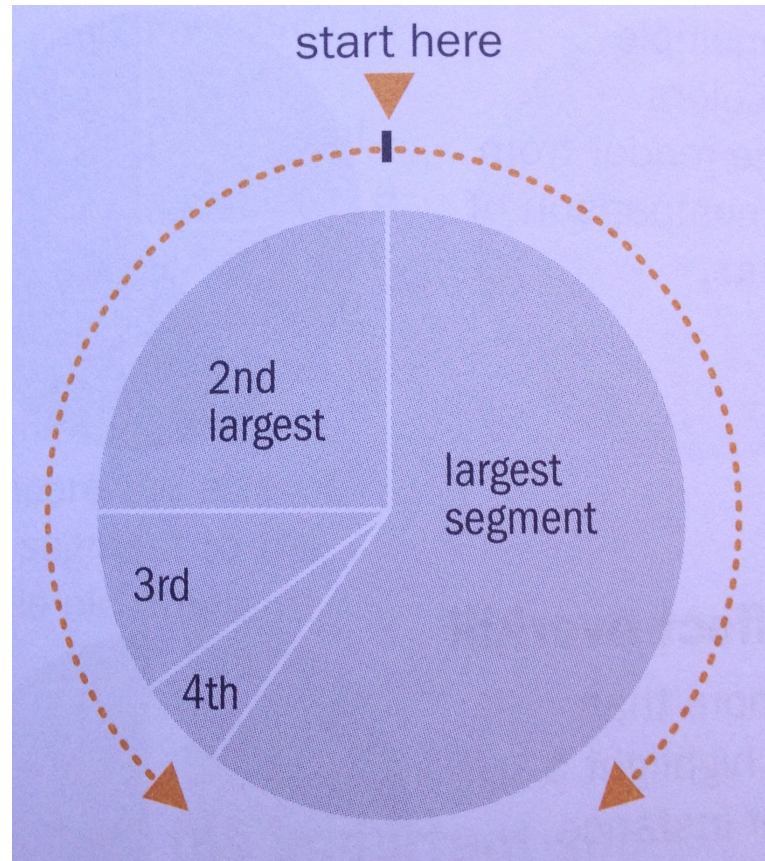


PIE CHARTS

- Pie Charts compares relative sizes and contributions

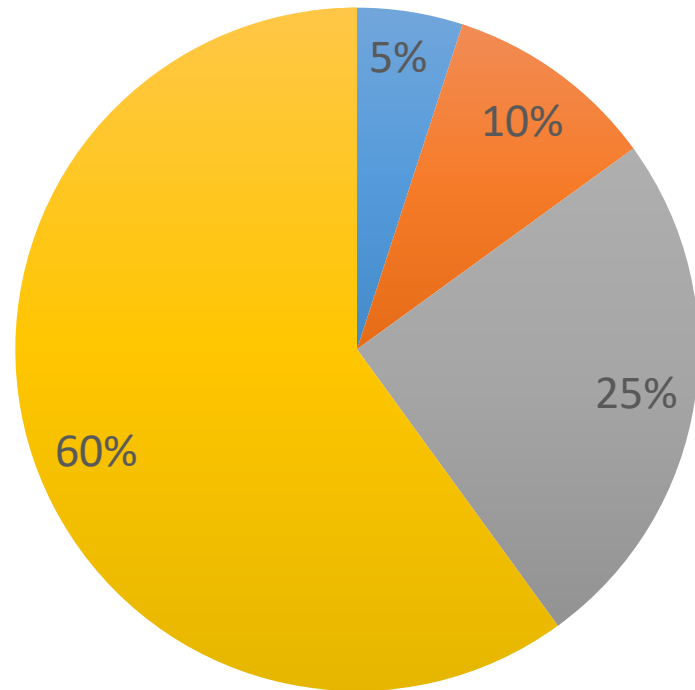


PIE CHARTS: ORDERING SLICES

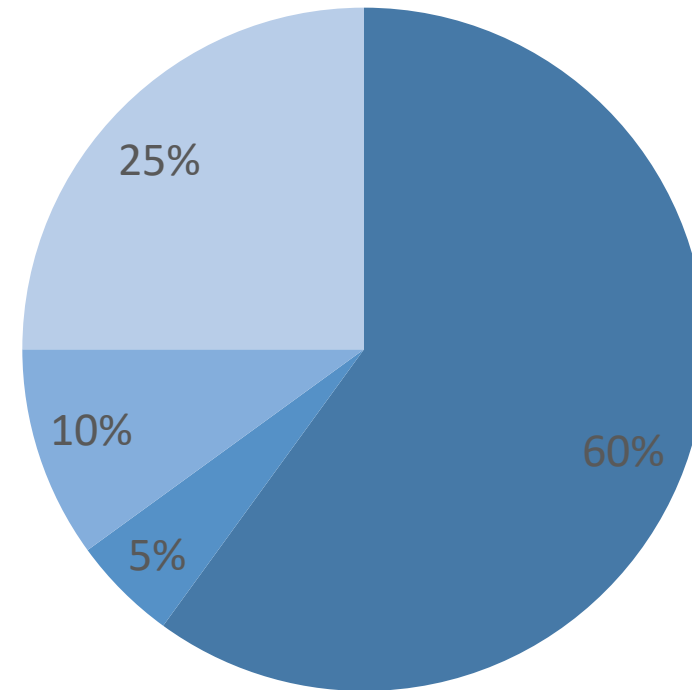


CHARTING EXAMPLES

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

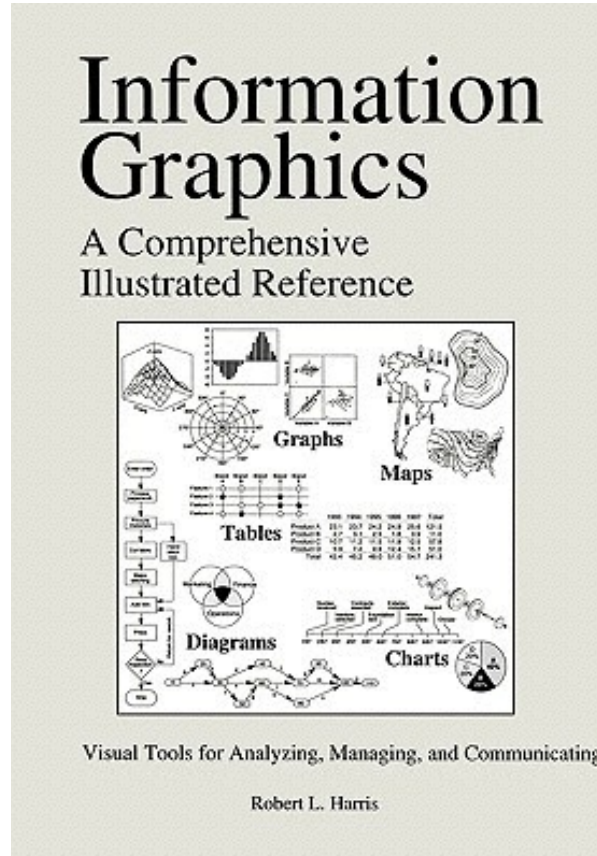
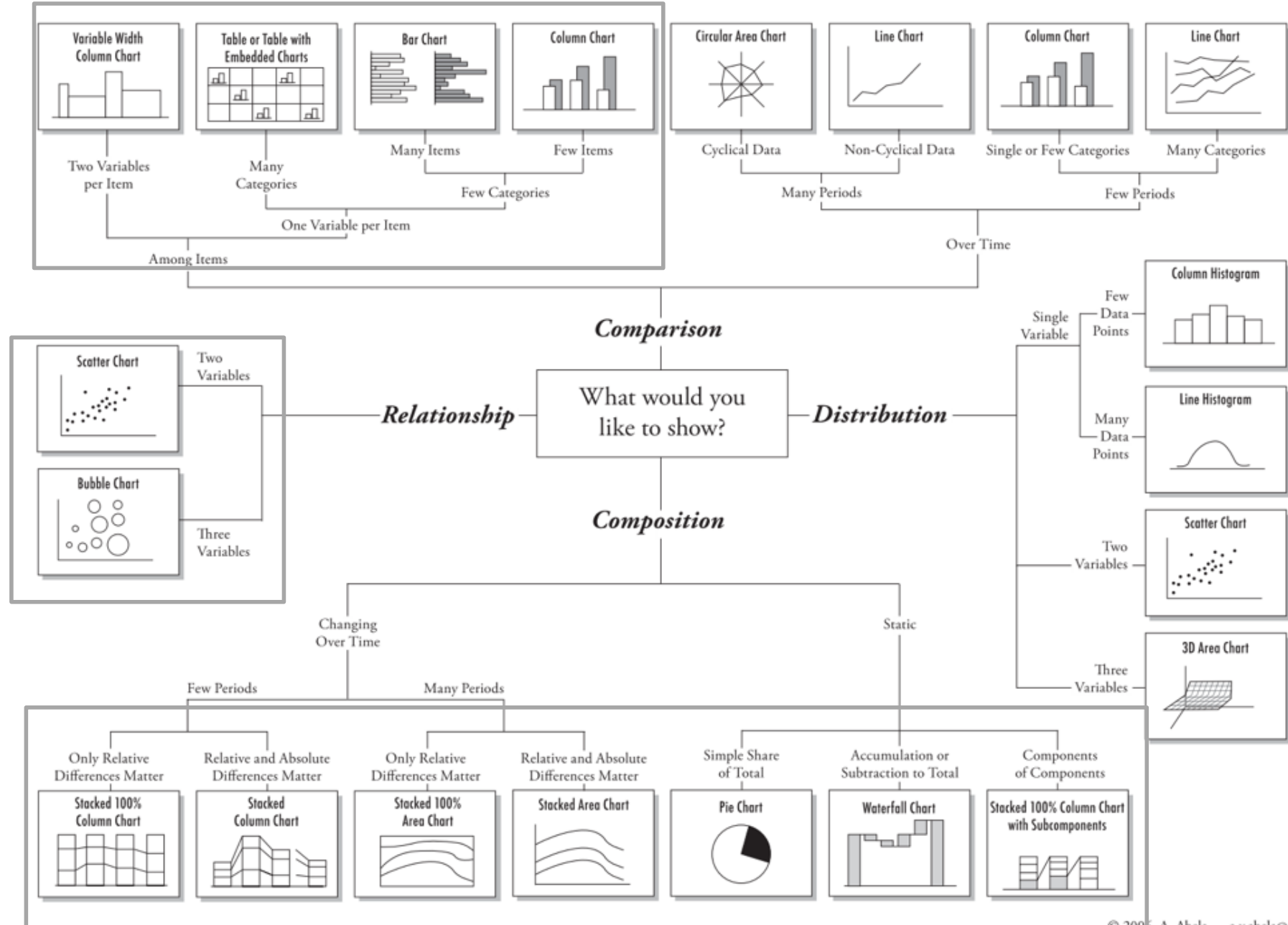
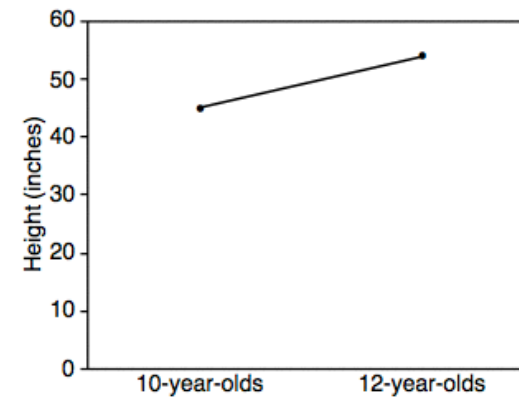
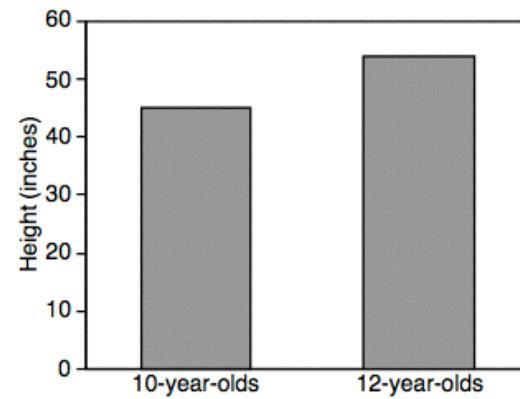
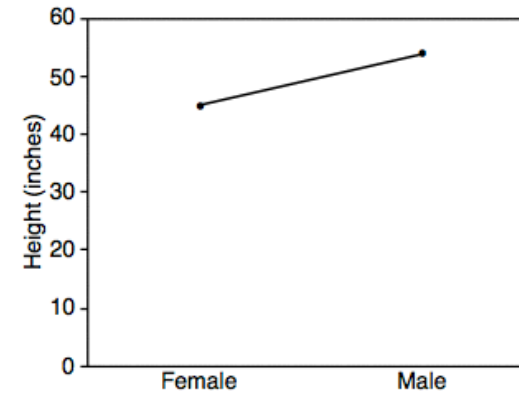
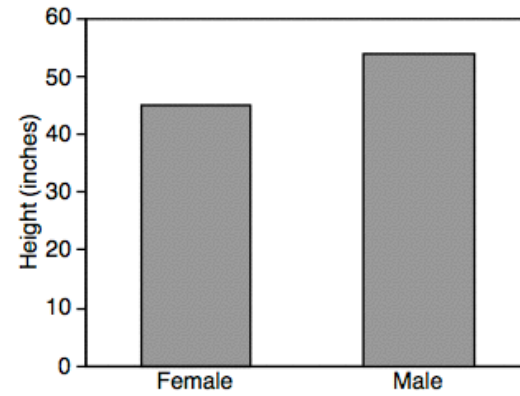


Chart Suggestions—A Thought-Starter



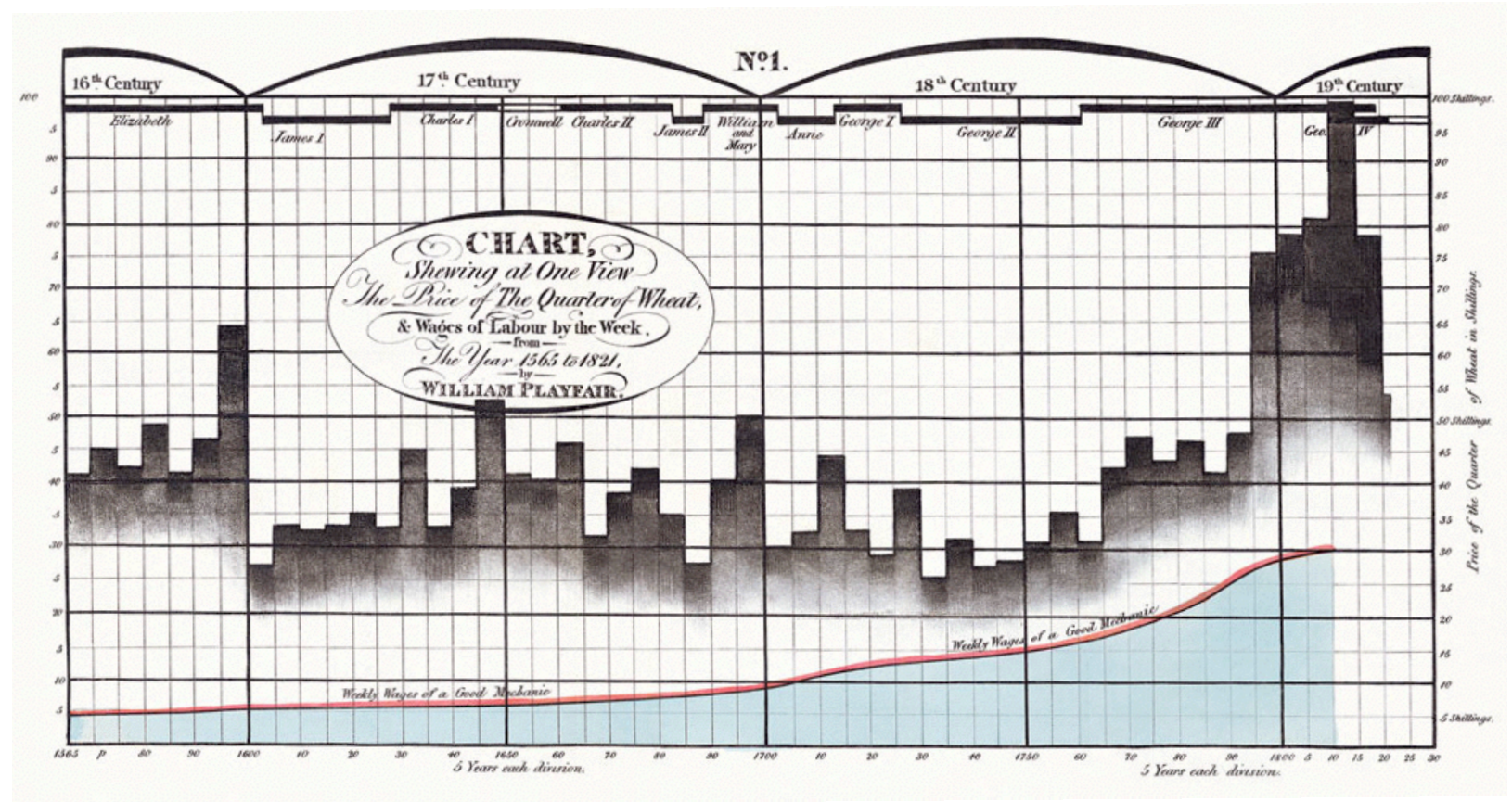
BARS VS. LINES



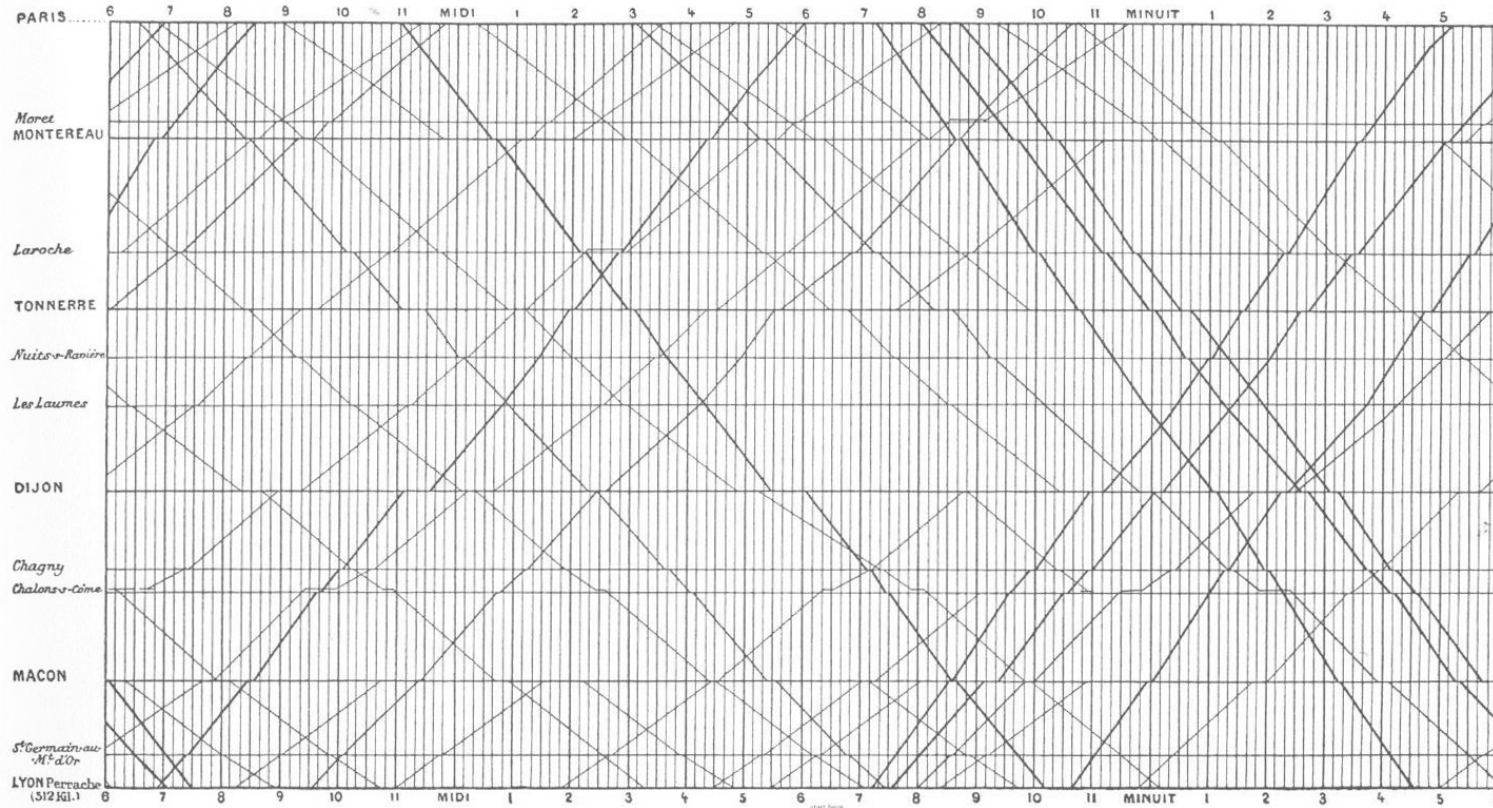
TREND OVER TIME

WILLIAM PLAYFAIR

1759-1823



PATTERNS OVER TIME



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

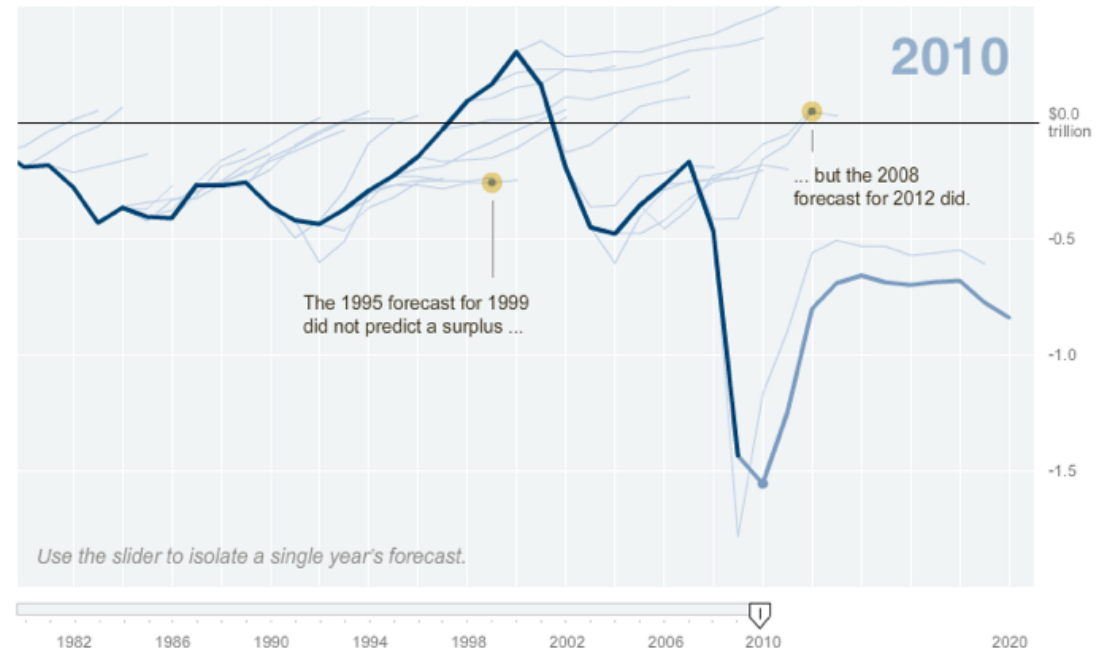
Budget Forecasts, Compared With Reality

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

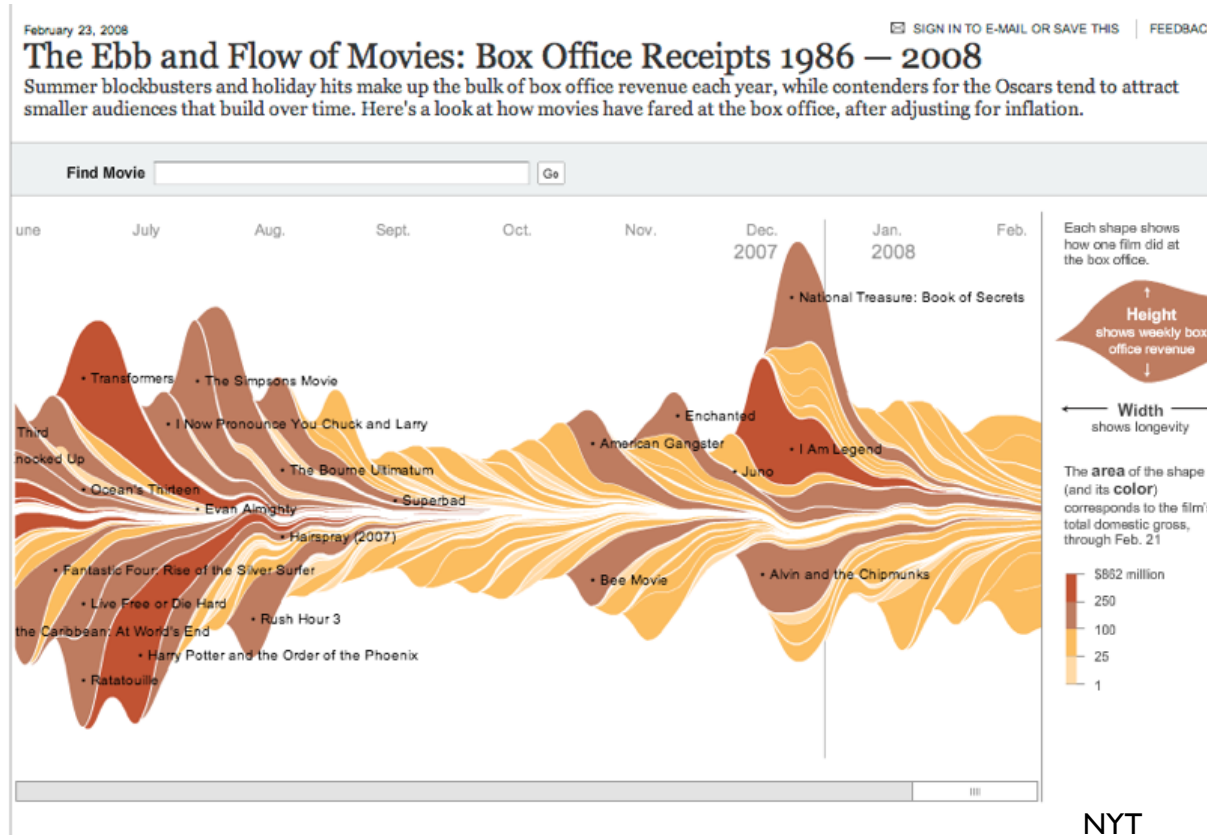
1 2 3 4 5 6 NEXT ▶

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**. **Jeff Clark** 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 Reported | 100000

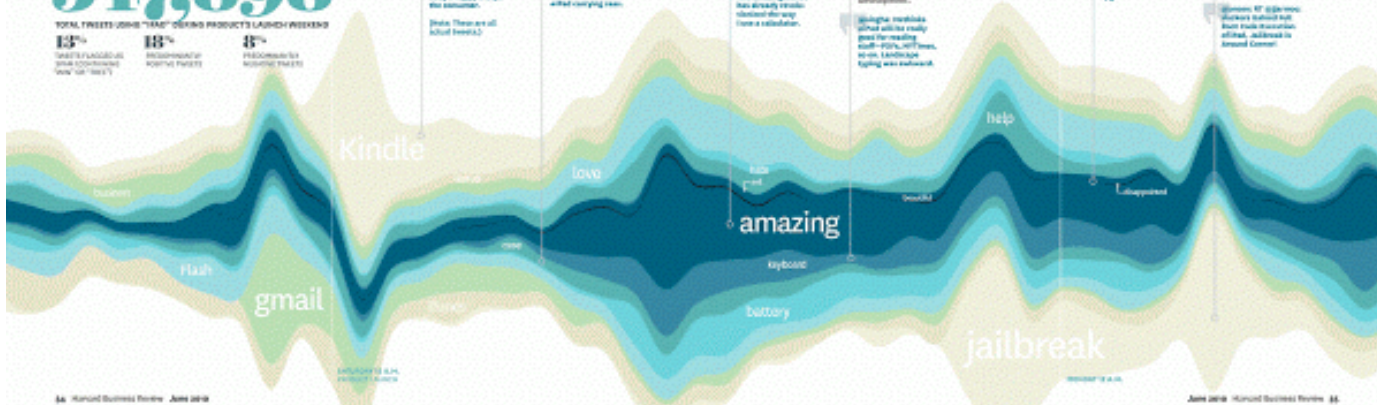
The iPad Launch by the Numbers

547,898

TOTAL TWEETS USING "IPAD" DURING PRODUCT'S LAUNCH WEEKEND
13% TWEETS FLAGGED AS SPAM (COMPARING NEW USE THIS YEAR)

18% RECOMMENDED BY OTHER TWEETS

4% RECOMMENDED BY FOLLOWERS



1 **Learn about the competitive landscape.** Tweets about your product that include the names of rival brands can reveal a lot about market positioning. Most of the **Kindle** tweets didn't focus on the iPad's being a Kindle killer (or, for that matter, a laptop killer). Instead, they hailed the arrival of the Kindle app for the iPad—warning Apple that **Kindle** books app will face serious competition.

2 **Look for unexpected themes.** Apple didn't want to promote ideas. The word **CRASH** became the most common right after people started trying iPads. The rumored iBooks and iApp Store trading apps about crashing their respective new gadgets.

3 **Go deeper into the stream.** This stream graph gives an overall impression of what people are saying about, it's important to know what other words are being used in relation to those in the stream. A lot of tweets were containing the word **HELP**. Some not from Apple fans but from gamers tweeting about the arrival of **Angry Birds** for the iPad. In our case, it might look as though a lot of tweets are more calling the iPad **AMAZING**. But in fact, most of this traffic came from users attempting a jail break connection (see below).

4 **Look for user experiences.** Product testing and reviews can't replace user reactions. Looking at iPad tweets with the words **TYPING** and **KEYBOARD**, and then digging out the most common words in those tweets (see above), gives you feedback from THOUSANDS of users. The negative events could help prioritize existing messages and product development.

5 **Learn why negative words are coming up.** Finding negative events is a good way to locate consumers' pain points. Though not frequently, **DISAPPOINTED** and **FRUSTRATED** appeared. Customers could be adjusting to address the most common complaints.

6 **Learn about conversation dominators.** Words that suddenly dominate the feed stream mean something has happened. But it's not always about the word. **LAUNCH** refers to hackers' use of iPad source code to run software not authorized by Apple. Twitter's most striking spike in activity, **crash**, and other failures on launch week, the device.

7 **Use a timeline of tweets.** The iPad app stores in memory, while Apple's software, which is likely to be updated as consumers.

8 **Use a timeline of tweets.** The iPad app stores in memory, while Apple's software, which is likely to be updated as consumers.

9 **Use a timeline of tweets.** The iPad app stores in memory, while Apple's software, which is likely to be updated as consumers.

10 **Use a timeline of tweets.** The iPad app stores in memory, while Apple's software, which is likely to be updated as consumers.

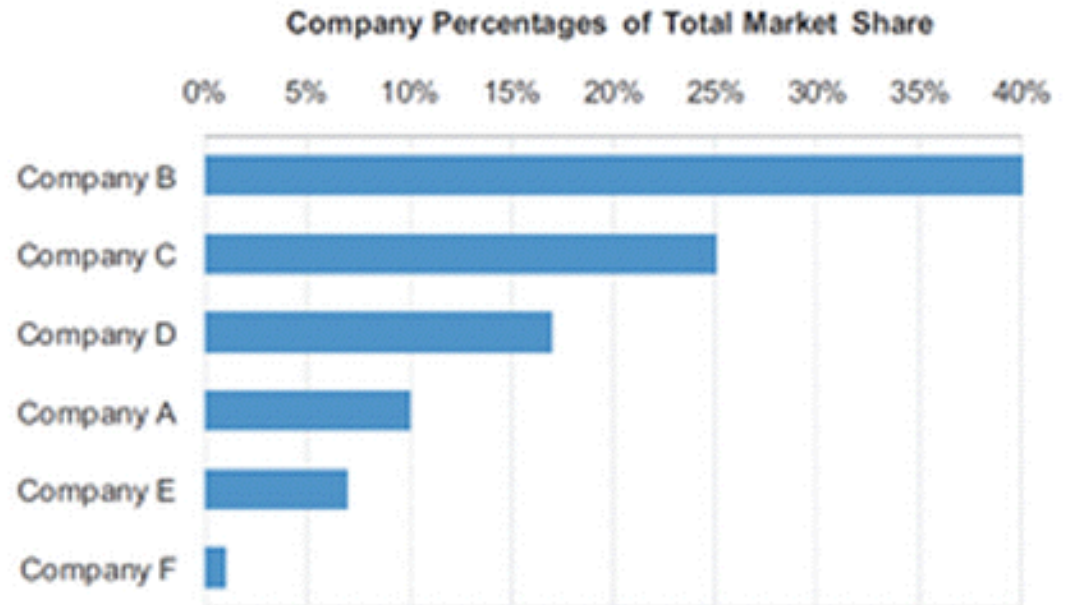
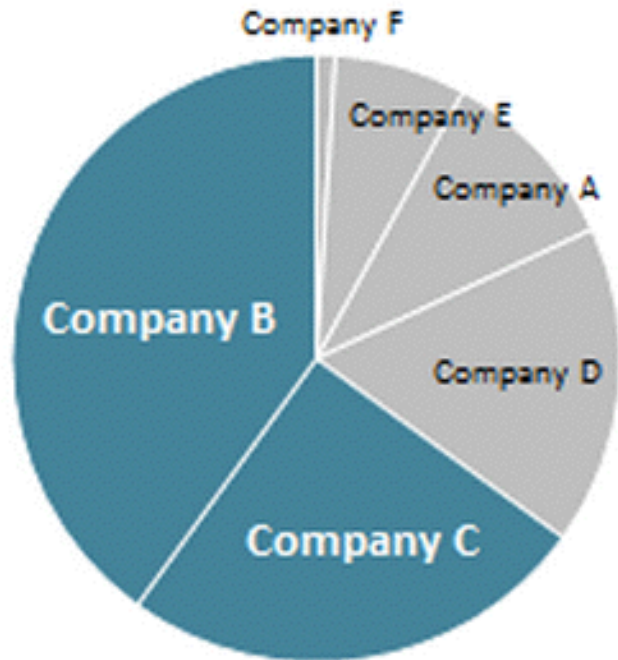
11 **Use a timeline of tweets.** The iPad app stores in memory, while Apple's software, which is likely to be updated as consumers.

12 **Use a timeline of tweets.** The iPad app stores in memory, while Apple's software, which is likely to be updated as consumers.

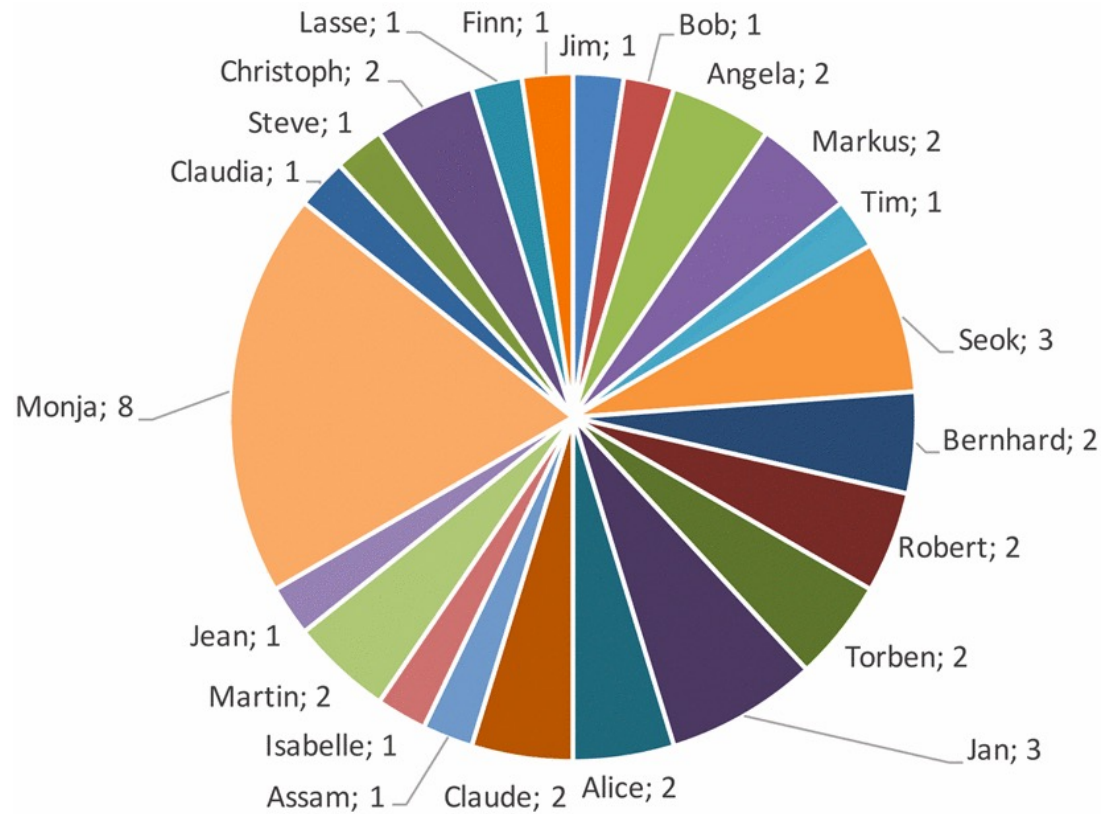


PIE VS BAR CHARTS

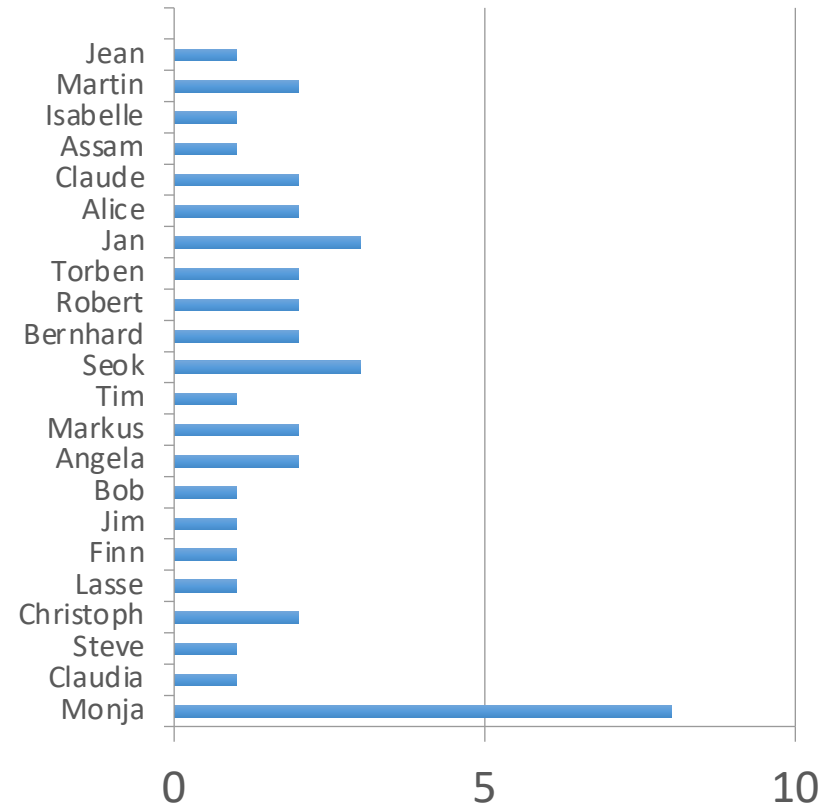
65% of the market is controlled by companies B and C



PIES VS BAR CHARTS

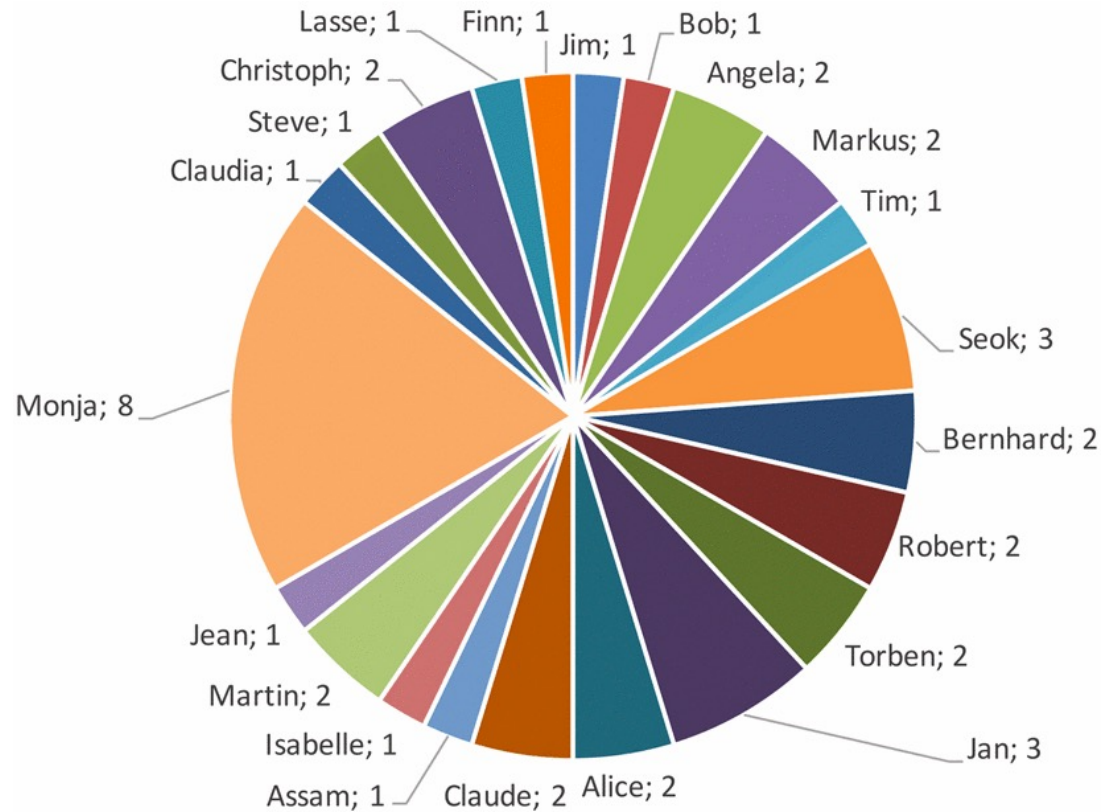


Episodes

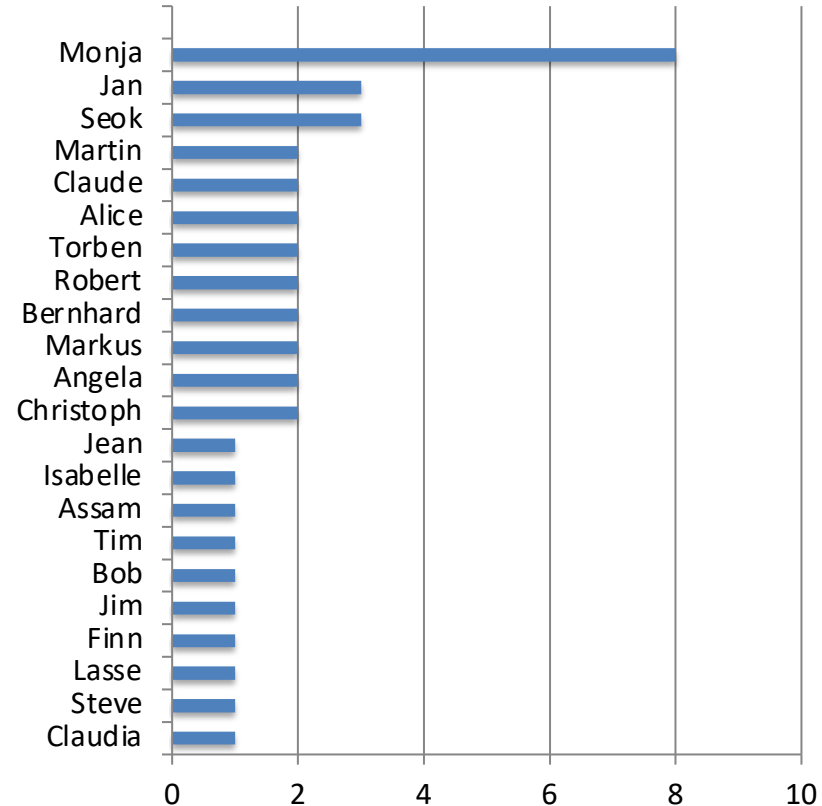


Furthermore, we present the distribution of attacks towards employees in detail in Fig. 10 right. The blue employees are secretaries, the green ones are administrators and the red ones are scientific employees. The number following the name is the number of times that person was attacked. All of the 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)



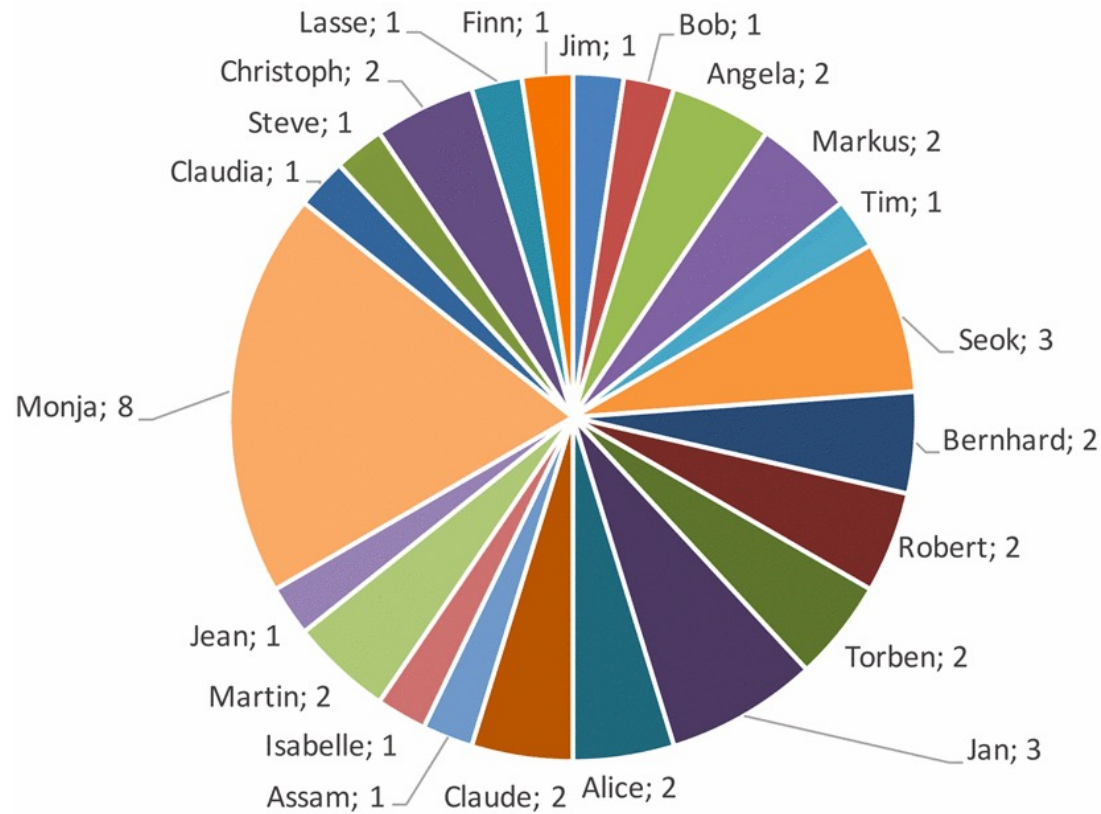
Episodes per person



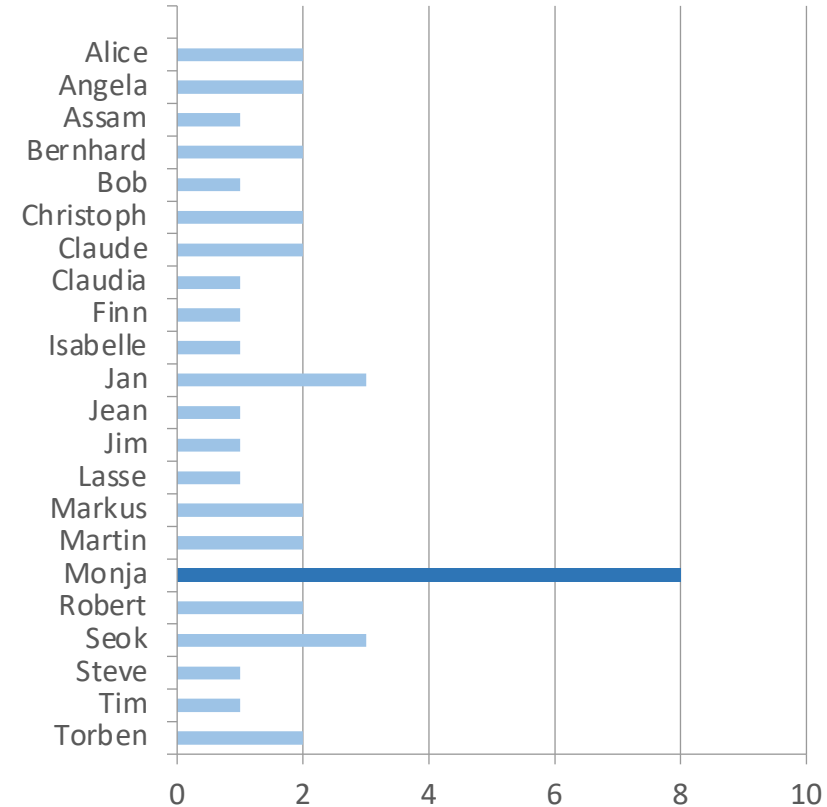
Furthermore, we present the distribution of attacks towards employees in detail in Fig. 10 right. The blue employees are secretaries, the green ones are administrators and the red ones are scientific employees. The number following the name is the number of times that person was attacked. All of the 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)

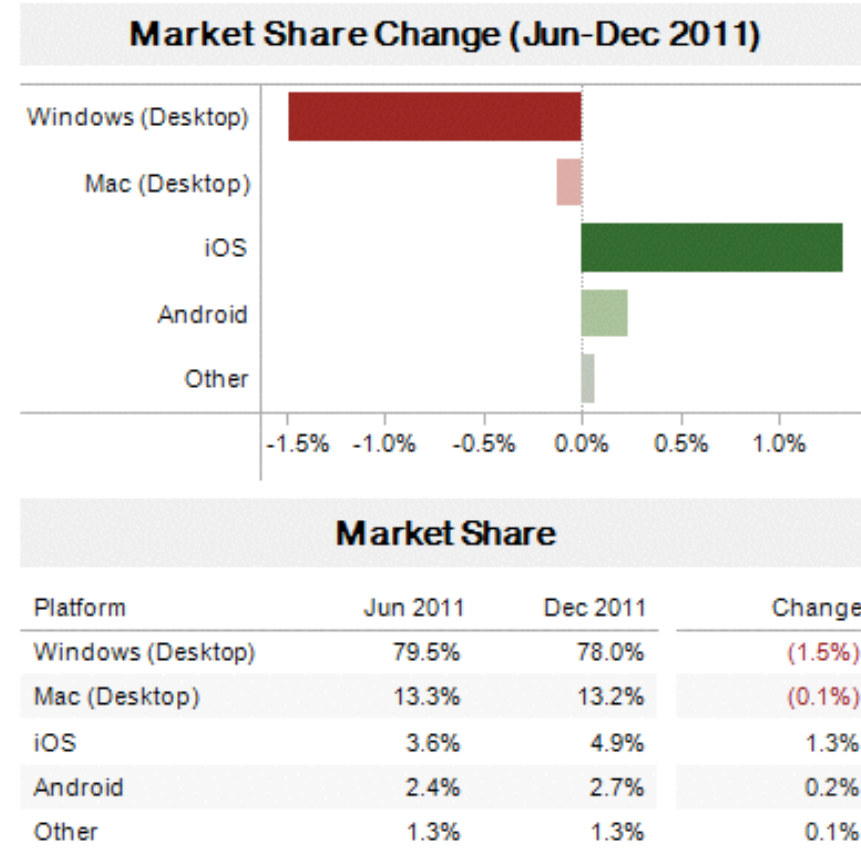
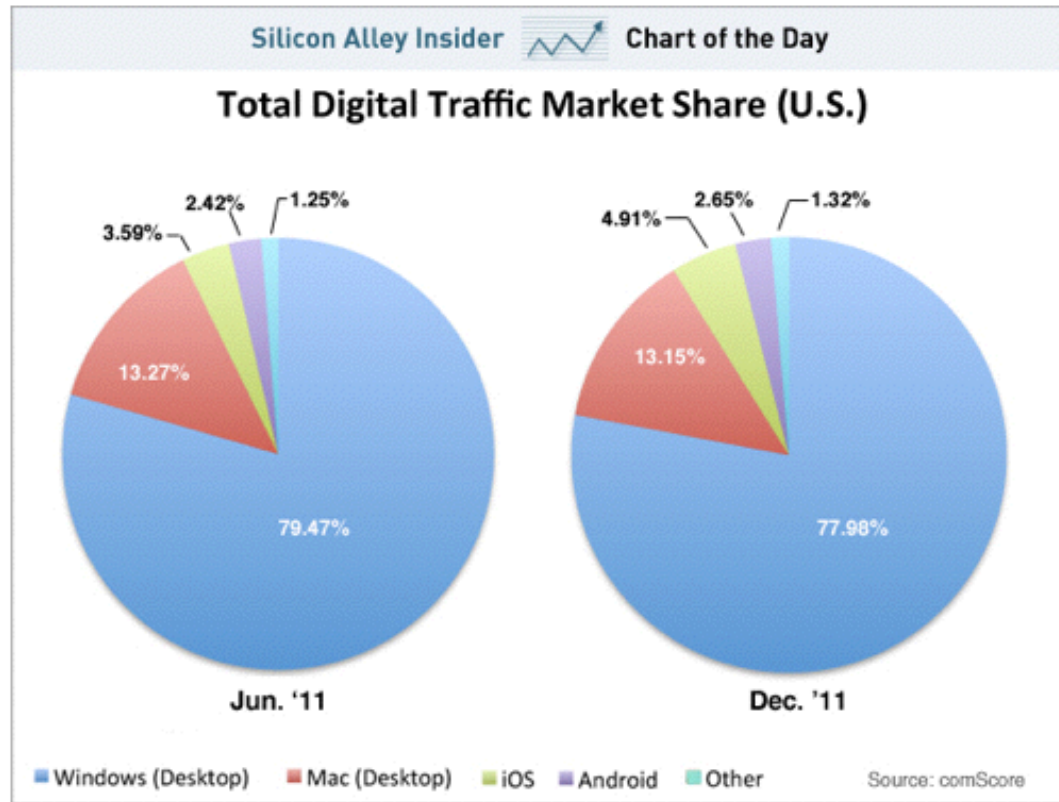


Episodes per person

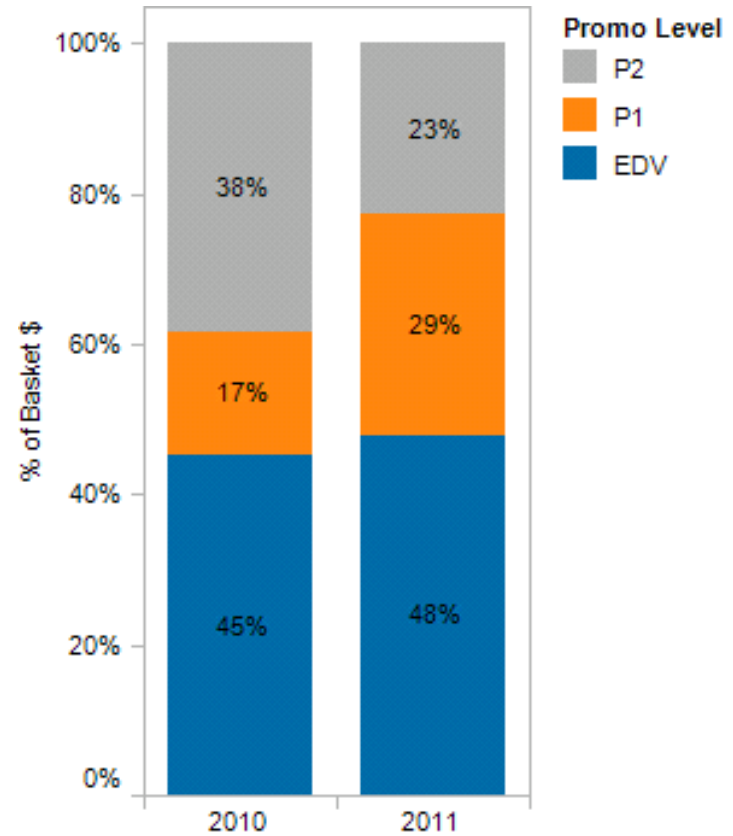
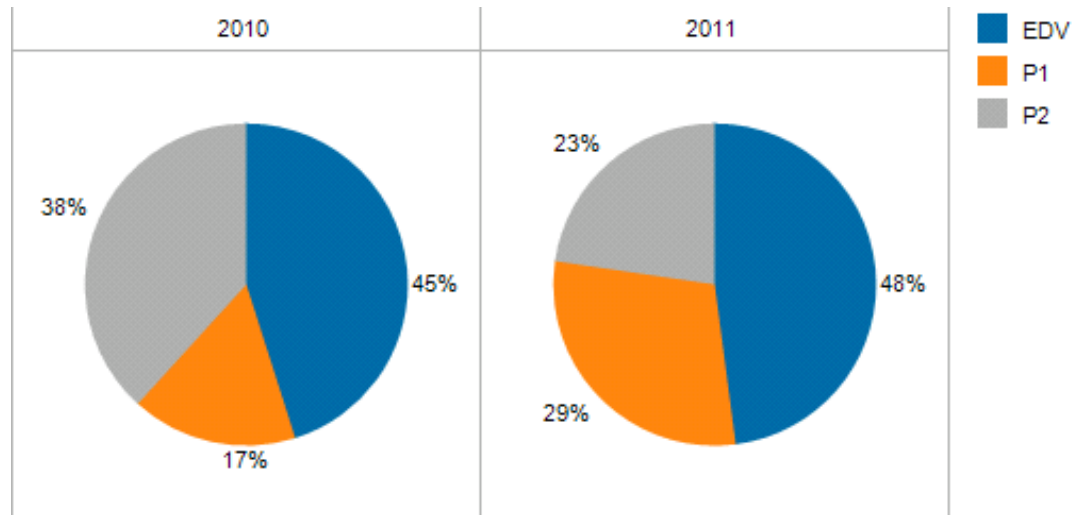


Furthermore, we present the distribution of attacks towards employees in detail in Fig. 10 right. The blue employees are secretaries, the green ones are administrators and the red ones are scientific employees. The number following the name is the number of times that person was attacked. All of the 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.

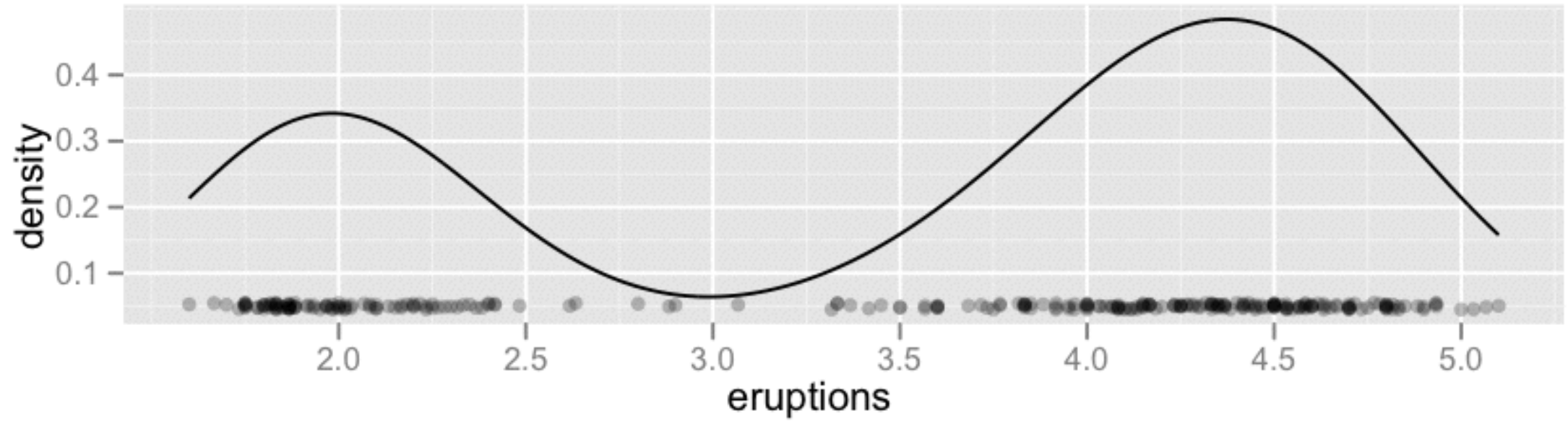
SHOWING CHANGES



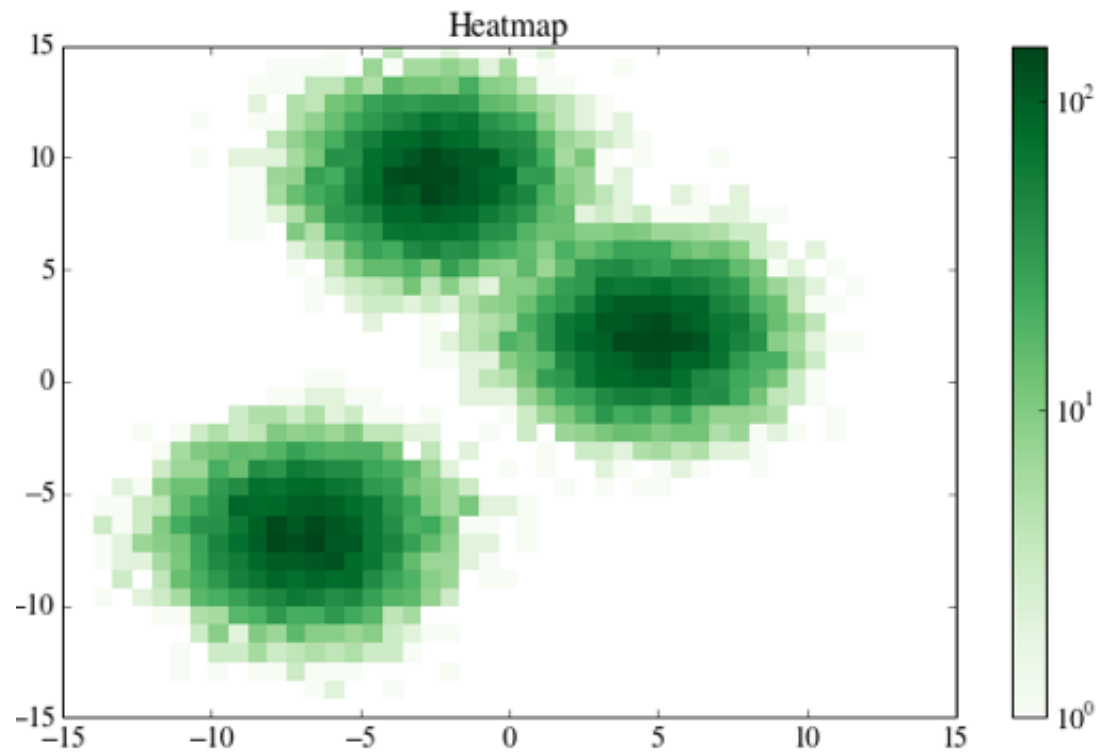
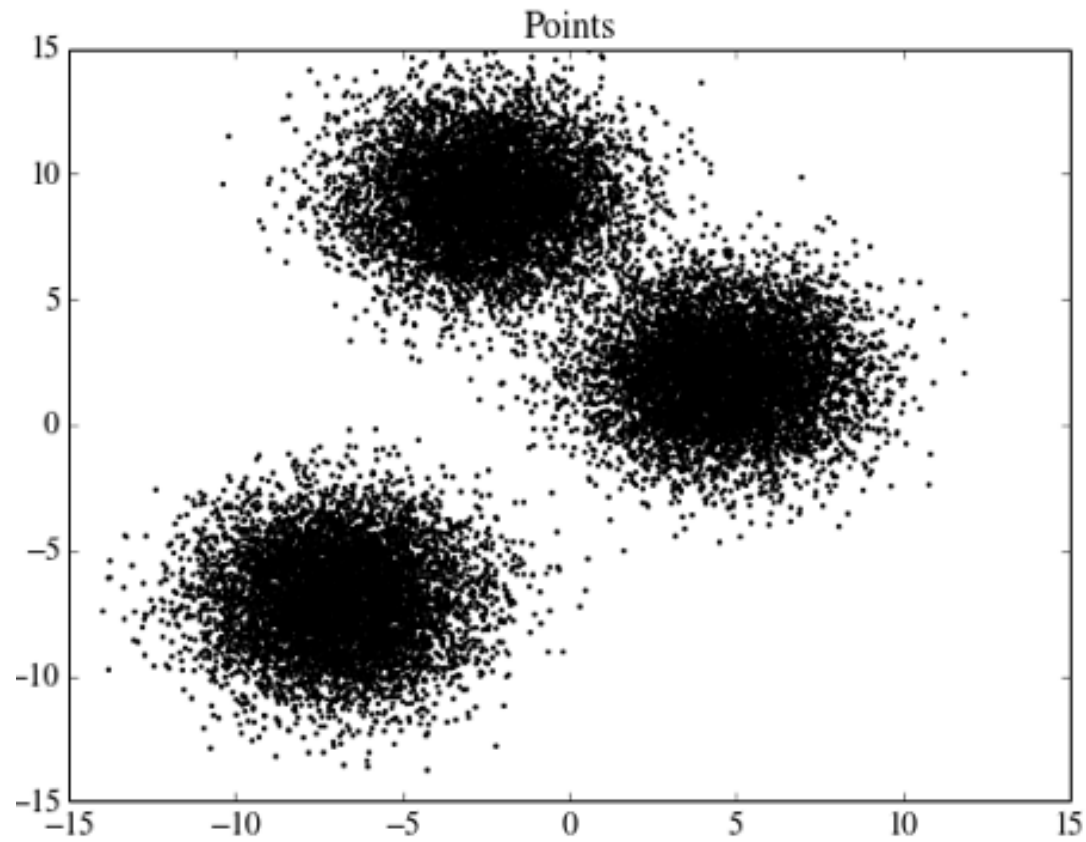
SHOWING CHANGES



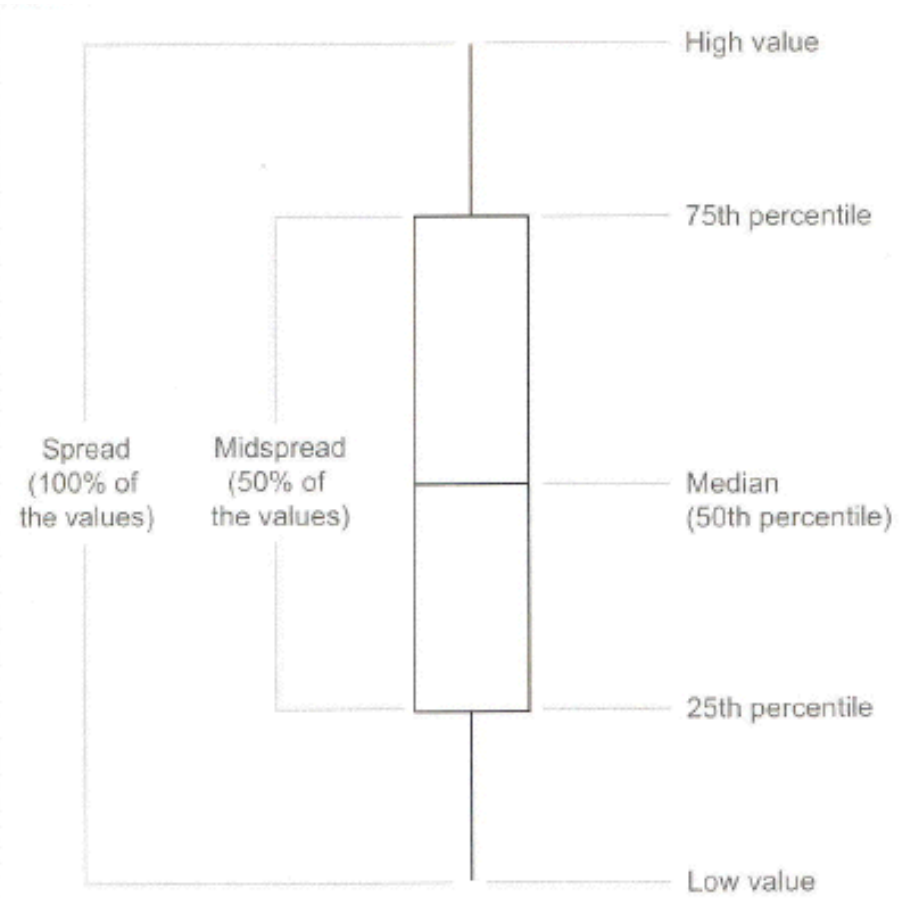
DENSITY PLOT



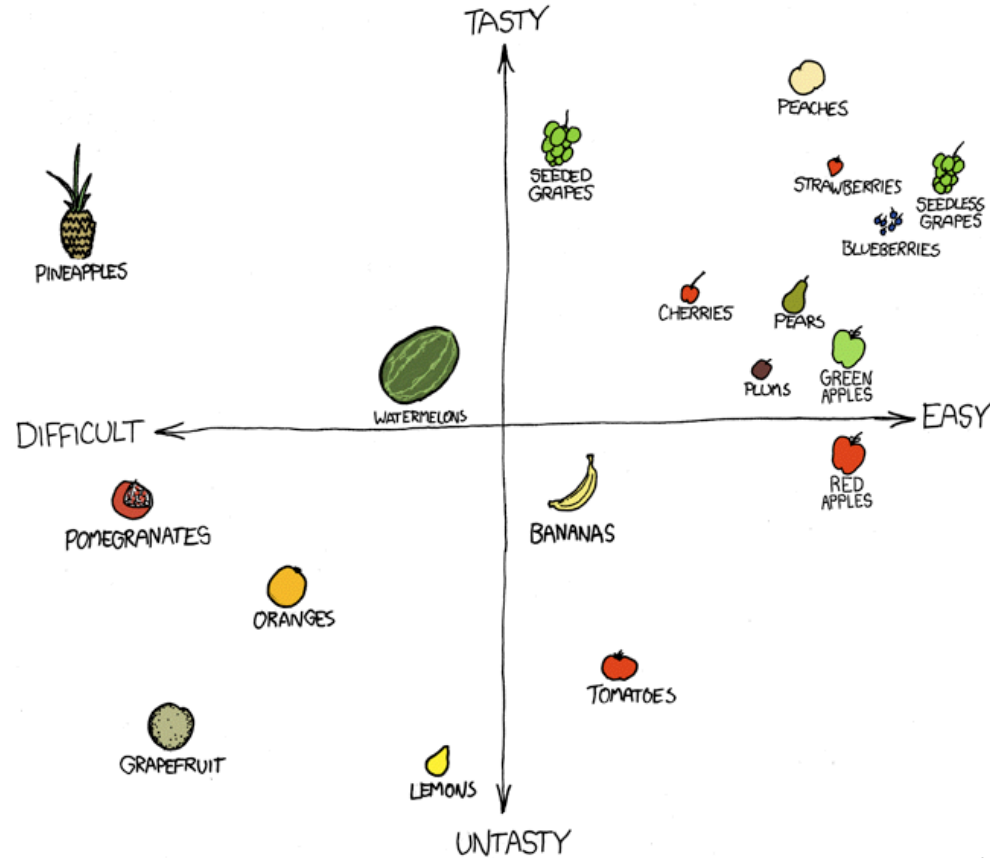
2D DENSITY PLOTS



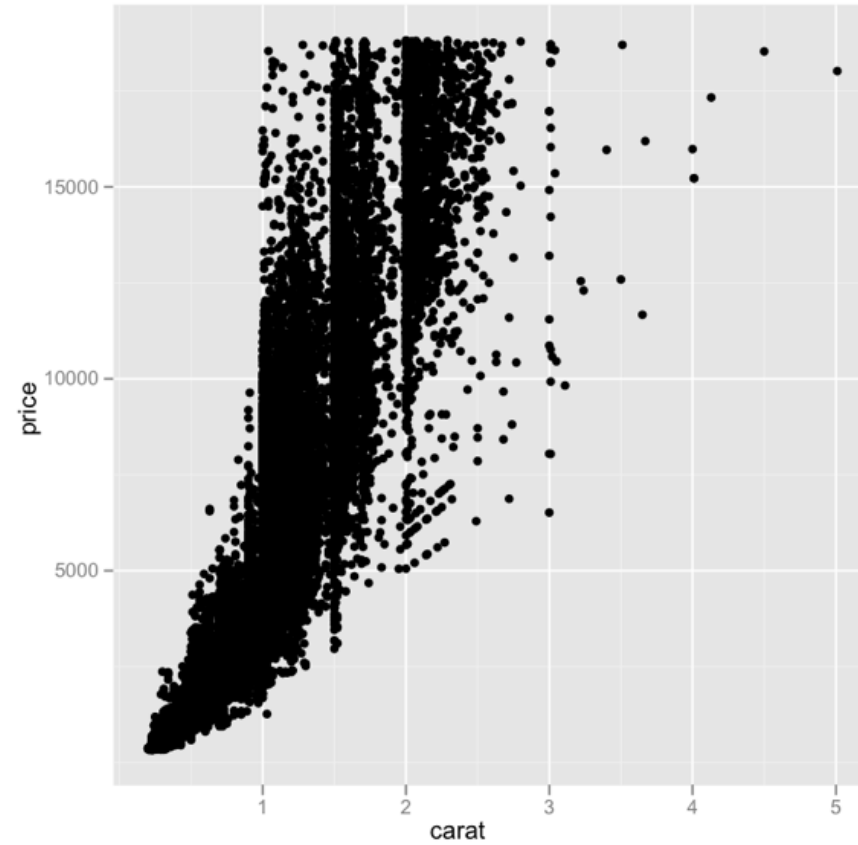
BOX PLOTS



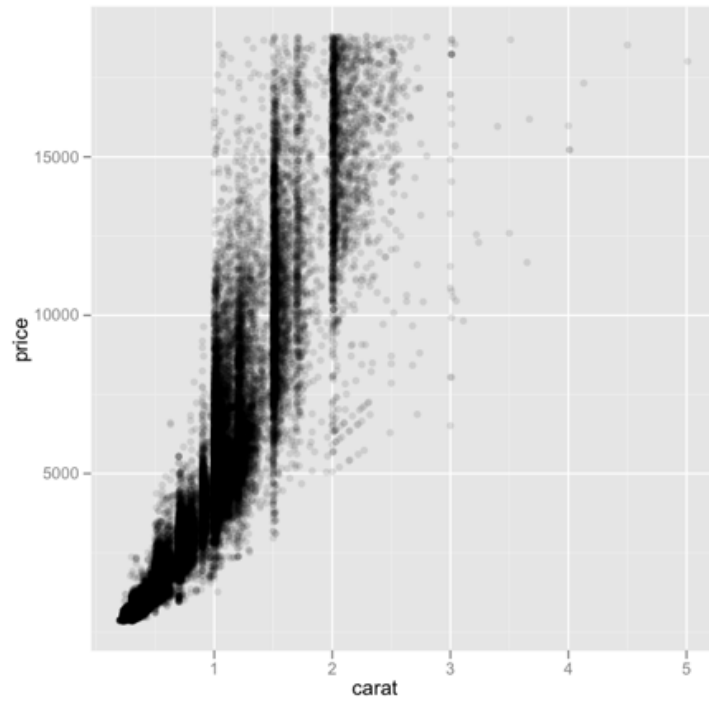
SCATTERPLOT



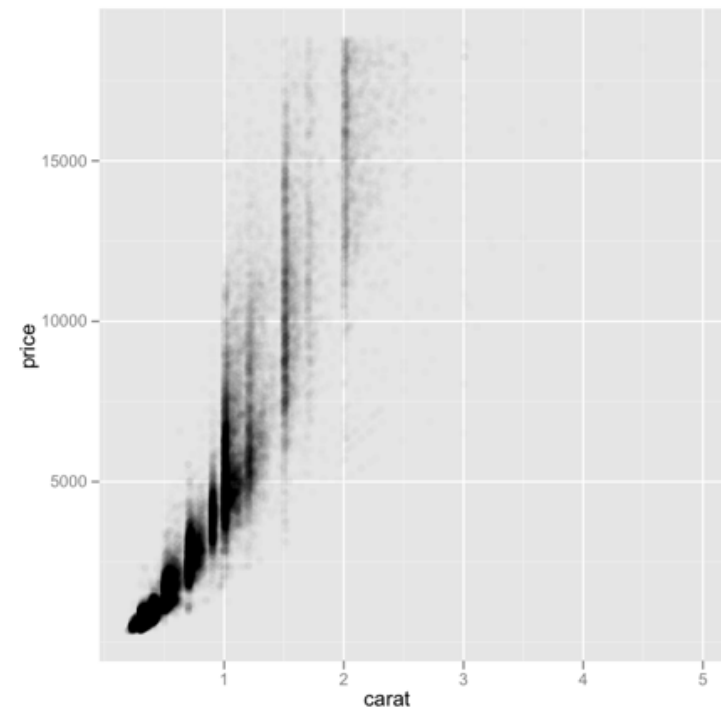
CLUTTERING, OVERPLOTTING



alpha=1/10

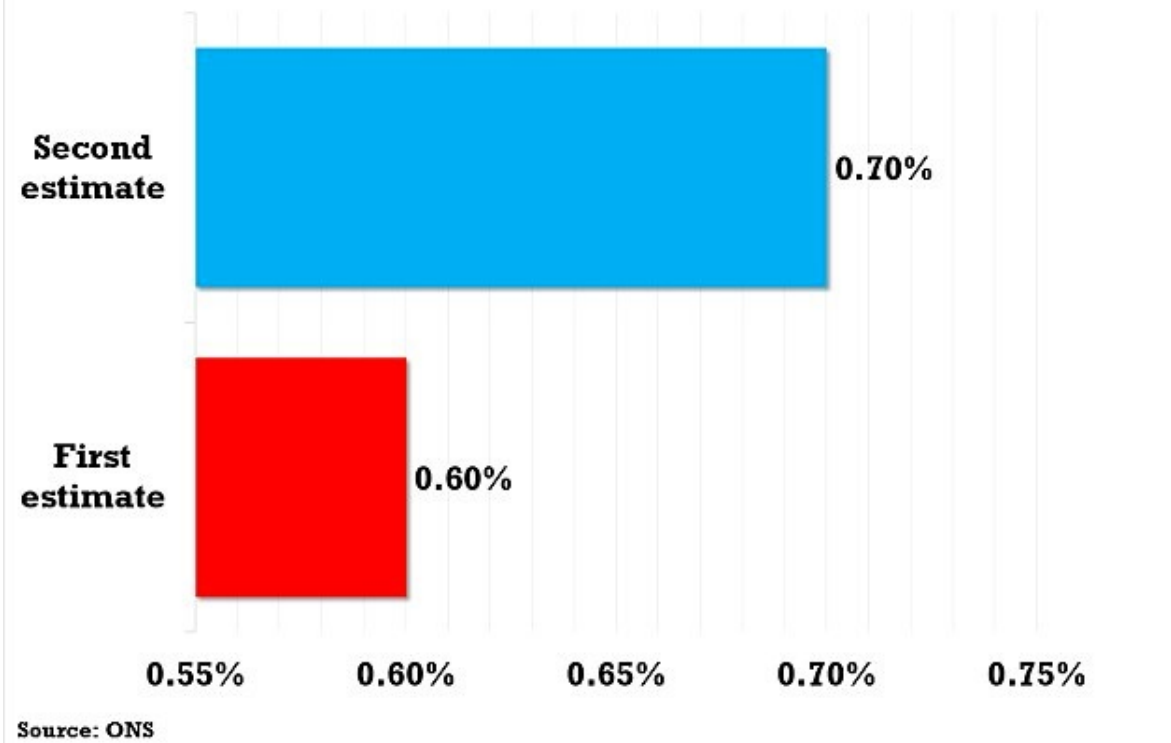


alpha=1/100



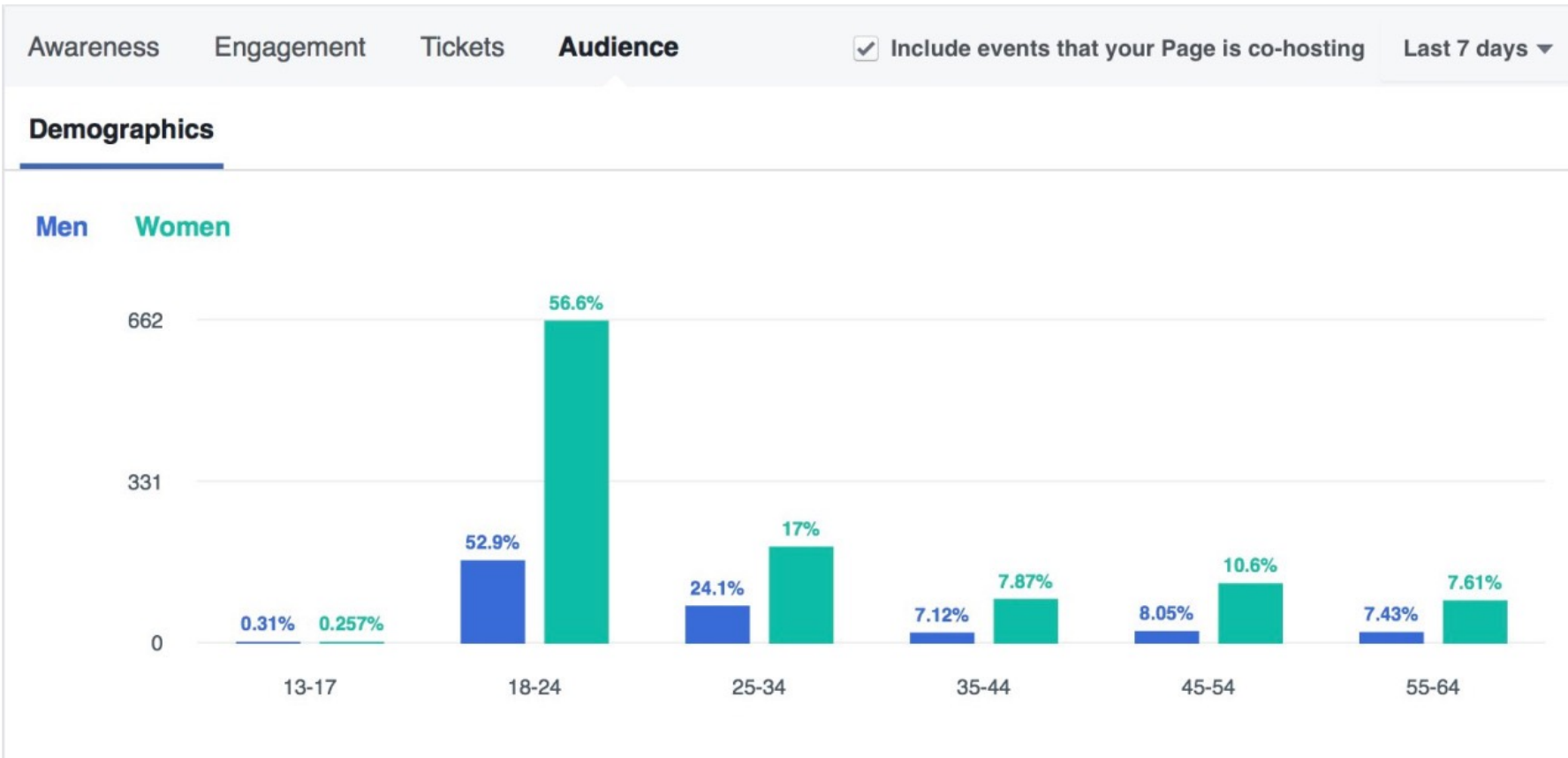
A FEW EXAMPLES AND CASE STUDIES

2016 Q4 GROWTH UPGRADED



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>



Procent użytków rolnych w gospodarstwach > niż 50 ha:

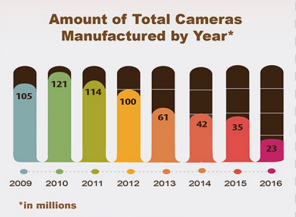
1989

25%

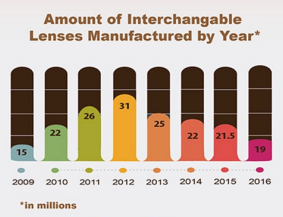


GAZETA.PL

CAMERA INDUSTRY FACTS 2009-2016



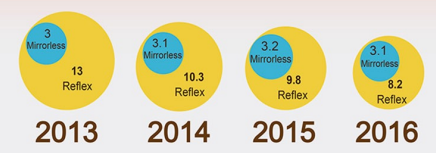
35% DROP IN SHIPPED CAMERAS IN 2016



12% DROP IN SHIPPED LENSES IN 2016

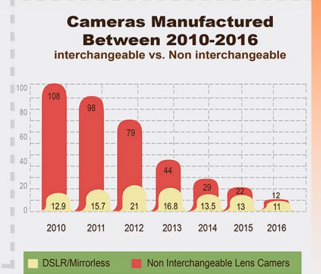
DSLR vs. Mirrorless 2013-2016

*in millions

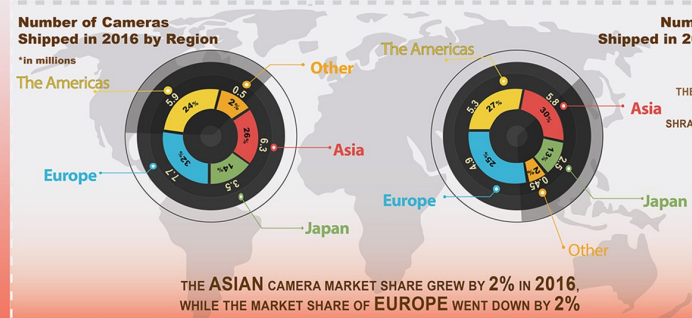


4% DECREASE IN MIRRORLESS PRODUCED & 17% DROP IN DSLR PRODUCED IN 2016

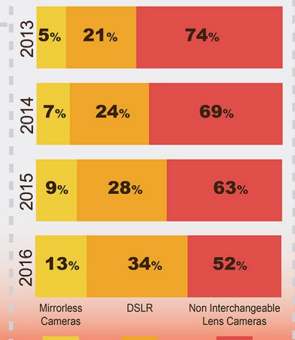
Based on CIPA (Camera & Imaging Products Association), Shipment of Digital Still Cameras & Lenses Data



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



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