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

The chart shows the median income (in USD) for different locations:
- Honolulu: $70,000
- Maui: $60,000
- Kalawao county: $60,000
- Kauai: $60,000
- Hawaii: $50,000
CHARTING EXAMPLES

- Honolulu
- Maui
- Kalawao county
- Kauai
- Hawaii

<table>
<thead>
<tr>
<th>Area</th>
<th>Median Income (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Honolulu</td>
<td>$60,000</td>
</tr>
<tr>
<td>Maui</td>
<td>$50,000</td>
</tr>
<tr>
<td>Kalawao county</td>
<td>$45,000</td>
</tr>
<tr>
<td>Kauai</td>
<td>$40,000</td>
</tr>
<tr>
<td>Hawaii</td>
<td>$30,000</td>
</tr>
</tbody>
</table>
CHARTING EXAMPLES
CHARTING EXAMPLES

stock price (USD)

Nov 1, 2016  Dec 1, 2016  Jan 1, 2017
CHARTING EXAMPLES
May these charts be improved? Why? How?
CHARTING EXAMPLES

- **Pie Chart**
  - County: Providence, Kent, Washington, Newport, Bristol
  - Numbers: Providence 627,000, Kent 49,900, Washington 127,000, Newport 166,000, Bristol 82,900

- **Bar Chart**
  - Counties: Providence, Kent, Washington, Newport, Bristol
  - Number of inhabitants: Providence 600,000, Kent 200,000, Washington 150,000, Newport 100,000, Bristol 50,000
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.

Font size = (1) + (2) + (20) = (19) + (20)
Fonts: General Rules

- Leading should be 2 points larger than type size
- Avoid too small or condensed typefaces
- 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 tracking

Fonts are meant to describe, not to adorn
Typography in Charts

Don’t

- Use bold and italic.
- Use all caps or high contrast white type out of black.
- Use bold for axis.
- Use tilted text.

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

Title of y-axis

Title of x-axis

Do

Headline of the chart

Title of y-axis

Title of x-axis
# Typography in Charts

Many elements in bold. Which part is highlighted?

Give emphasis to relevant results.
DATA-INK RATIO

Visual Display of Quantitative Data
Edward Tufte, 1983
**DATA-INK RATIO**

Data-Ink Ratio = \[
\frac{\text{Data ink}}{\text{Total ink used in graphic}}
\]
**DATA-INK RATIO**

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

- **Categoria 1**
  - Serie 1
  - Serie 2
  - Serie 3
  - Serie 4

- **Categoria 2**
  - Serie 1
  - Serie 2
  - Serie 3
  - Serie 4
BAR CHARTS

Represent discrete quantities

Town A

Town B

Town C

Town D
BAR CHARTS

Avoid non-functional adornation
BAR CHARTS: BASELINE

Chart Title

Chart Title
BAR CHARTS: BASELINE

Chart Title

-10 -5 0 5 10 15
A B C D
BAR CHARTS: ORDERING

France 1,1
Germany 4,1
Italy 6,1
China 9,1
USA 7,1

France 1,1
Germany 4,1
Italy 6,1
US 7,1
China 9,1
PIE CHARTS

• Pie Charts compares relative sizes and contributions
PIE CHARTS: ORDERING SLICES

- Start here
- Largest segment
- 2nd largest
- 3rd
- 4th
May these charts be improved? Why? How?
TAKEAWAY MESSAGES

• Charts exploit position on scale VV
• Best practice to reduce biases and misinterpretation of charts
Information Graphics
A Comprehensive Illustrated Reference

Visual Tools for Analyzing, Managing, and Communicating
Robert L. Harris
Chart Suggestions—A Thought-Starter

Comparison
- What would you like to show?

Relationship
- Two Variables
- Three Variables

Distribution
- Few Data Points
- Many Data Points

Composition
- Few Periods
- Many Periods
- Static

Graph Types
- Variable Width Column Chart
- Table or Table with Embedded Charts
- Box Chart
- Column Chart
- Circular Area Chart
- Line Chart
- Column Chart
- Line Chart
- Scatter Chart
- Bubble Chart
- Column Histogram
- Line Histogram
- Stacked 100% Column Chart
- Stacked Column Chart
- Stacked Area Chart
- Stacked Area Chart
- Stacked 100% Column Chart with Subcomponents
- Pin Chart
- Waterfall Chart

Taxonomy
- Two Variables per Item
- Many Categories
- Many Items
- Few Items
- One Variable per Item
- Few Categories

Over Time
Bars vs. Lines

Line implies trends. Do not use for categorical data.
Trend over Time

William Playfair
1759-1823
PATTERNS OVER TIME

Marey’s diagram, 1885
TREND OVER TIME
TREND OVER TIME

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

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

Use the slider to isolate a single year’s forecast.

The 1995 forecast for 1999 did not predict a surplus ...

but the 2008 forecast for 2012 did.

Published: February 2, 2010

Make clear distinction between data and prediction
STREAMGRAPHS

The Ebb and Flow of Movies: Box Office Receipts 1986 — 2008

Summer blockbusters and holiday hits make up the bulk of box office revenue each year, while contenders for the Oscars tend to attract smaller audiences that build over time. Here’s a look at how movies have fared at the box office, after adjusting for inflation.

NYT
Vision Statement

Six Ways to Find Value in Twitter’s Noise

It’s easy to dislike Twitter in jibber, but smart marketers will recognize it as a source of free consumer data to be mined for rich insights. By visualizing this stream of data, we can help pinpoint what consumers are saying and doing, which gives companies a competitive edge.

1. Listen for unexpected currencies. 
2. Look for unexpected sources.
3. Listen for unexpected products.
4. Look for new revenue streams.
5. Look for new opportunities.
6. Listen for unexpected companies.

The iPad Launch by the Numbers

517,898

How to Read This Graph

This graph incorporates two different time bases over the quarter leading up to the product launch on April 3. The top line is sales data for the entire quarter; the bottom line represents sales data for the last four days.

June 2010, HBR by Jeff Clark
65% of the market is controlled by companies B and C.
PIES VS BAR CHARTS

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)

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.
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

- High value
- 75th percentile
- Median (50th percentile)
- 25th percentile
- Low value
- Spread (100% of the values)
- Midspread (50% of the values)
SCATTERPLOT
CLUTTERING, OVERPLOTTING
alpha=1/10

alpha=1/100
The Office for National Statistics (ONS) said gross domestic product (GDP) expanded by 0.7 per cent in the fourth quarter - an increase of 0.1 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
Demographics

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>13-17</td>
<td>0.31%</td>
<td>0.257%</td>
</tr>
<tr>
<td>18-24</td>
<td>52.9%</td>
<td>56.6%</td>
</tr>
<tr>
<td>25-34</td>
<td>24.1%</td>
<td>17%</td>
</tr>
<tr>
<td>35-44</td>
<td>7.12%</td>
<td>7.87%</td>
</tr>
<tr>
<td>45-54</td>
<td>8.05%</td>
<td>10.8%</td>
</tr>
<tr>
<td>55-64</td>
<td>7.43%</td>
<td>7.61%</td>
</tr>
</tbody>
</table>
Procent użytkowników rolnych w gospodarstwach > niż 50 ha:

1989

25%
Camera Industry Facts 2009-2016

35% drop in shipped cameras in 2016
12% drop in shipped lenses in 2016

DSLR vs. Mirrorless 2013-2016

Cameras Manufactured Between 2010-2016

Camera Market Overview 2013-2016

The entire camera market in 2016 saw 81% drop compared to 2010

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

Source: https://lensvid.com/gear/lensvid-exclusive-happened-photography-industry-2016/
VISUAL TAXONOMY

The Data Visualisation Catalogue

Search by Function

About  ·  Suggest  ·  Shop  ·  Resources

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

http://www.datavizcatalogue.com/
TAKEAWAY MESSAGES

• Appropriate chart type for specific data type and visualization task