

Visualizing Data

Sina Rüeger

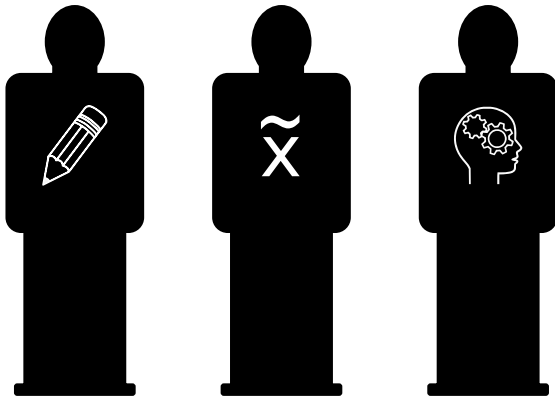
Institute of Data Analysis and Process Design (IDP)

29. November 2011

Example: temperature in train

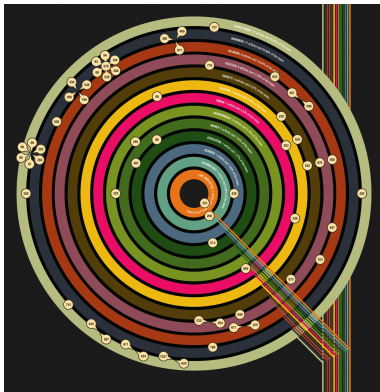


Three aspects



After isotype, Otto Neurath, 1882 – 1945

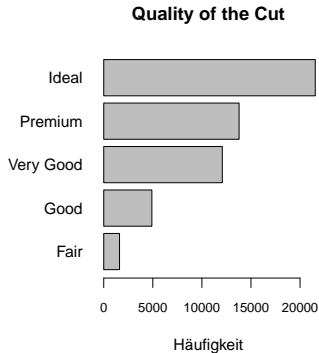
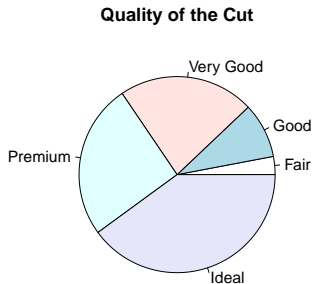
Example graphic designer



One Week of The Guardian - Designing the News.

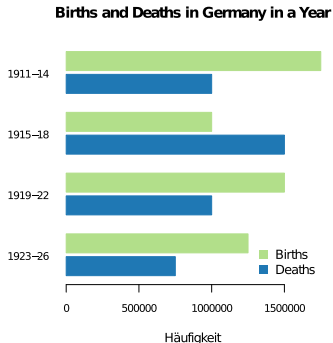
(http://www.designingthenews.com/wp-content/uploads/2008/04/03_wednesday_a1_72.jpg, letzter Zugriff:

Example statistician



Piechart and barplot of diamonds data.

Example psychologist



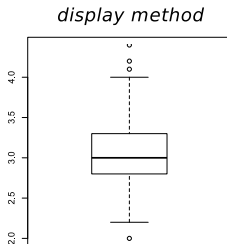
Birth and deaths in germany: barplot and pictogramms.

(http://www.latebytes.nl/archives/2008/04/17/T1.N112_A6_Isotype-Neurath.jpg, last access: 2011-11-29)

Processing information

```
3.5 3.0 3.2 3.1 3.6 3.9 3.4 3.4 2.9  
3.1 3.7 3.4 3.0 3.0 4.0 4.4 3.9 3.5  
3.8 3.8 3.4 3.7 3.6 3.3 3.4 3.0 3.4  
3.5 3.4 3.2 3.1 3.4 4.1 4.2 3.1 3.2  
3.5 3.6 3.0 3.4 3.5 2.3 3.2 3.5 3.8  
3.0 3.8 3.2 3.7 3.3 3.2 3.2 3.1 2.3  
2.8 2.8 3.3 2.4 2.9 2.7 2.0 3.0 2.2  
2.9 2.9 3.1 3.0 2.7 2.2 2.5 3.2 2.8  
2.5 2.8 2.9 3.0 2.8 3.0 2.9 2.6 2.4  
2.4 2.7 2.7 3.0 3.4 3.1 2.3 3.0 2.5  
2.6 3.0 2.6 2.3 2.7 3.0 2.9 2.9 2.5  
2.8 3.3 2.7 3.0 2.9 3.0 3.0 2.5 2.9  
2.5 3.6 3.2 2.7 3.0 2.5 2.8 3.2 3.0  
3.8 2.6 2.2 3.2 2.8 2.8 2.7 3.3 3.2  
2.8 3.0 2.8 3.0 2.8 3.8 2.8 2.8 2.6  
3.0 3.4 3.1 3.0 3.1 3.1 3.1 2.7 3.2  
3.3 3.0 2.5 3.0 3.4 3.0
```

encoding



decoding



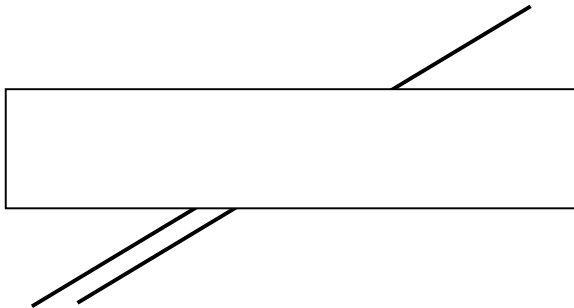
Encoding

Graph data = encode information

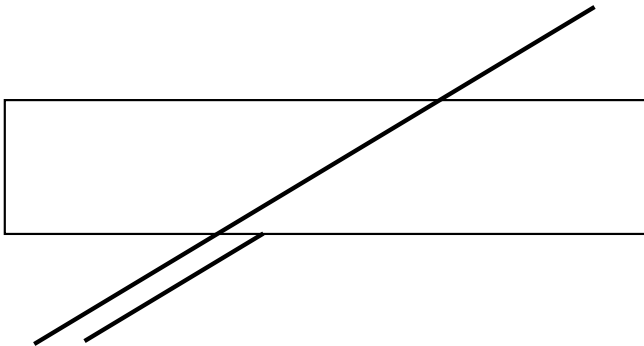
Decoding

Visual perception = decode information

The Elements of Graphing from William S. Cleveland (1994)



Poggendorff illusion



Poggendorff illusion

<http://www.youtube.com/watch?v=JhjUJTw2i1M>

02:28 to 04:15

Visual illusion

- Visual illusions give a wrong interpretation of data.
- Do not create visual illusions, e.g. by making pie charts.

Tipps (if possible)

- Do not (accidentally) cause *visual illusions* → leads to wrong conclusions.
- Do not use *colors* or *shades* if not necessary → difficult to decode (colorbrewer2.org)
- Do not use *areas* or *volumes* (e.g. 3D plots) → difficult to estimate.
- When using areas, they should be *proportional* to the actual values (e.g. circles).
- Do not only present graphs to the audience - also explain your data and your mapping technique.
- Do not use more *dimensions* in a graph than you have data.
- Reduce the graph - every ink on a graph needs a *reason* (data - ink - ratio)

R: software for statistics and computing

`http://www.r-project.org/`

2 purposes of visualizing data

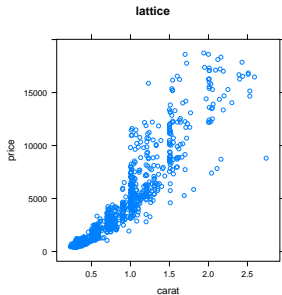
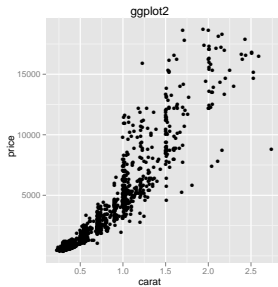
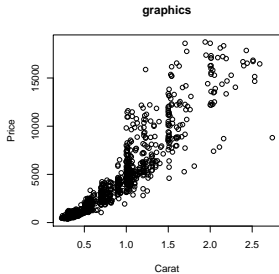
Purpose 1: analyse data

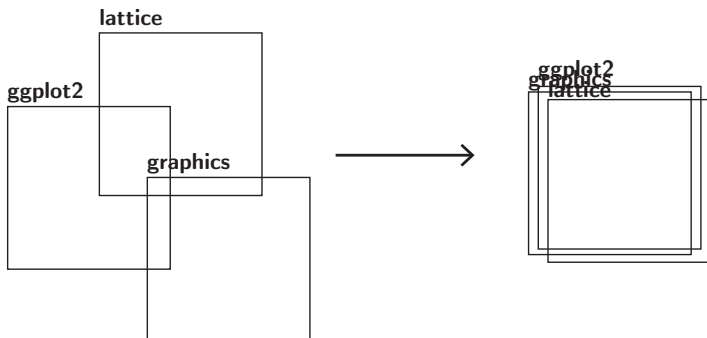
I only graph data for myself or some members of my research group.

Purpose 2: graphs as communication

I produce graphs and show them to others (e.g. in an article).

└ Conclusion 1: Implementation in software

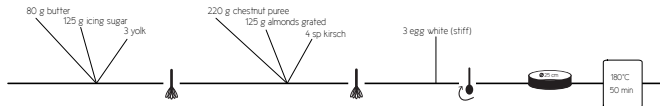




chestnut cake

by sina
2011-11-06

80 g butter
125 g icing sugar
3 yolk
220 g chestnut puree
125 g almonds grated
4 sp kirsch
3 egg white (stiff)



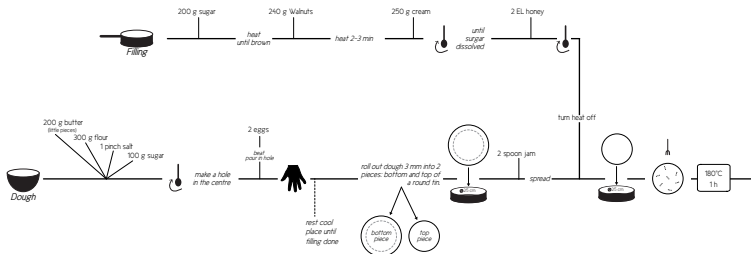
copyright 2011, sina ruegger

engadiner nusstorte

2011-10-24

copyright 2011, sina rueger

300 g flour
200 g butter (little pieces)
1 pinch salt
100 g sugar
2 eggs
2 spoon jam
200 g sugar
240 g Walnuts
250 g cream
2 EL honey



Edward Tufte in *The Visual Display of Quantitative Information*
(1995)

"Above all else show the data!"

"Data graphics should draw the viewers attention to the sense and substance of the data, not something else."