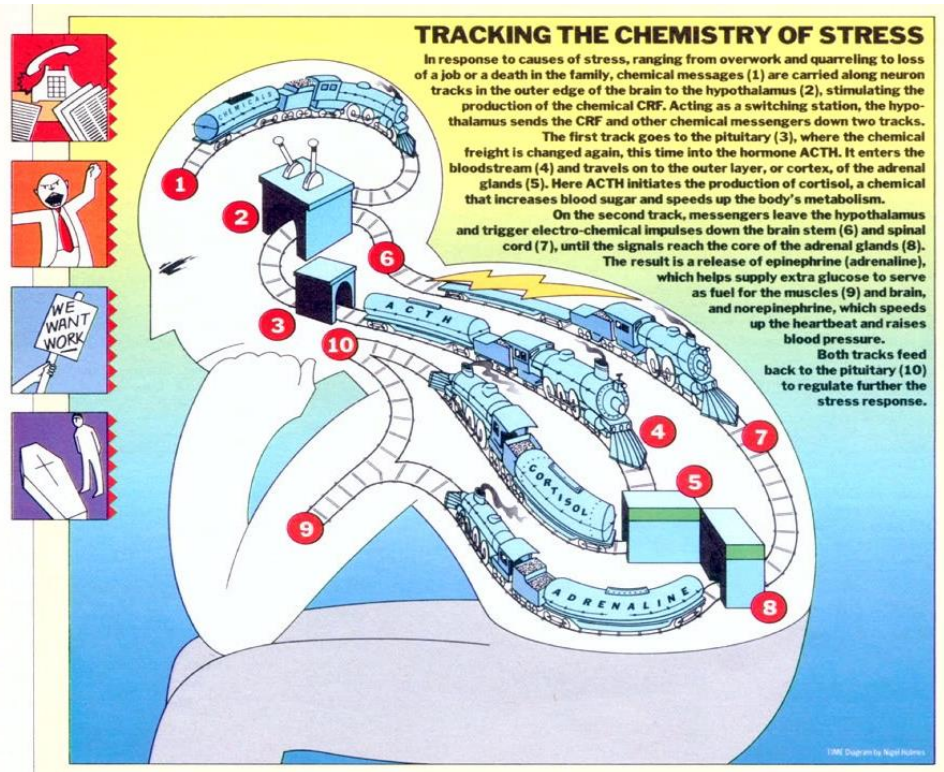


Datenvisualisierung

Gliederung

1. Was ist Datenvisualisierung?
2. Warum brauchen wir Visualisierungen?
3. Wir stehen auf den Schultern von Giganten
4. „Falsche“ Visualisierungen
5. Richtlinien für Visualisierungen
6. Software für Analyse und Visualisierungen
 - Software: Tableau vs. Excel
 - Software: Tableau vs. Power BI

1. Was ist Datenvisualisierung?



aries greatly by feature

: Features

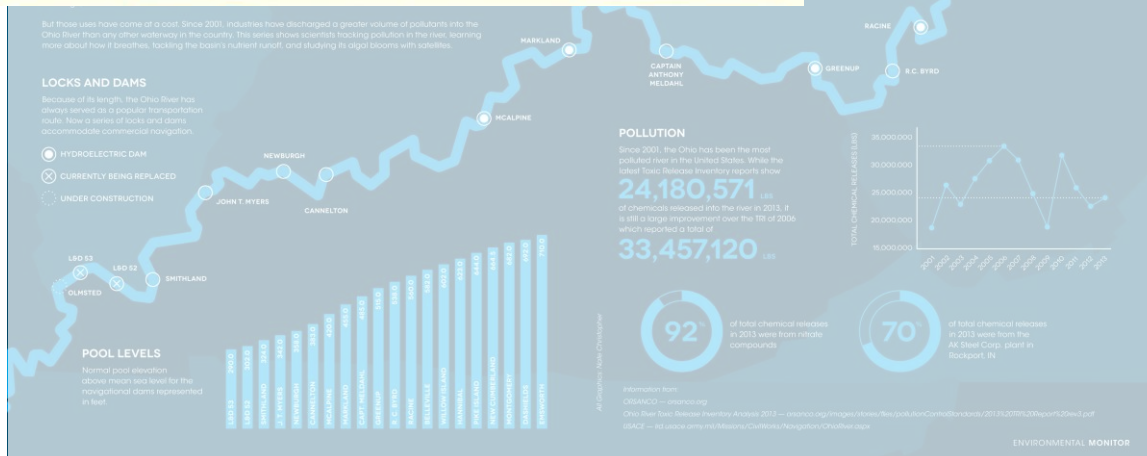
satisfied ■ Somewhat satisfied ■ Very satisfied ■ Completely satisfied

40%	47%
36%	47%

Features A and B continue to top user satisfaction

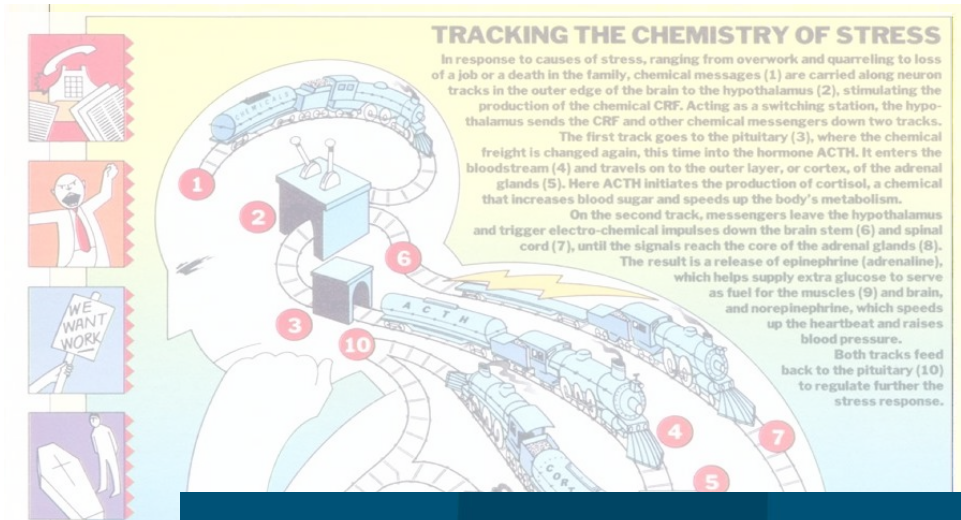
Users are least satisfied with Features J and N; what improvements can we make here for a better user experience?

Feature O is least used. What steps can we proactively take with existing users to increase utilization?



does this represent?

1. Was ist Datenvisualisierung?



aries greatly by feature

: Features

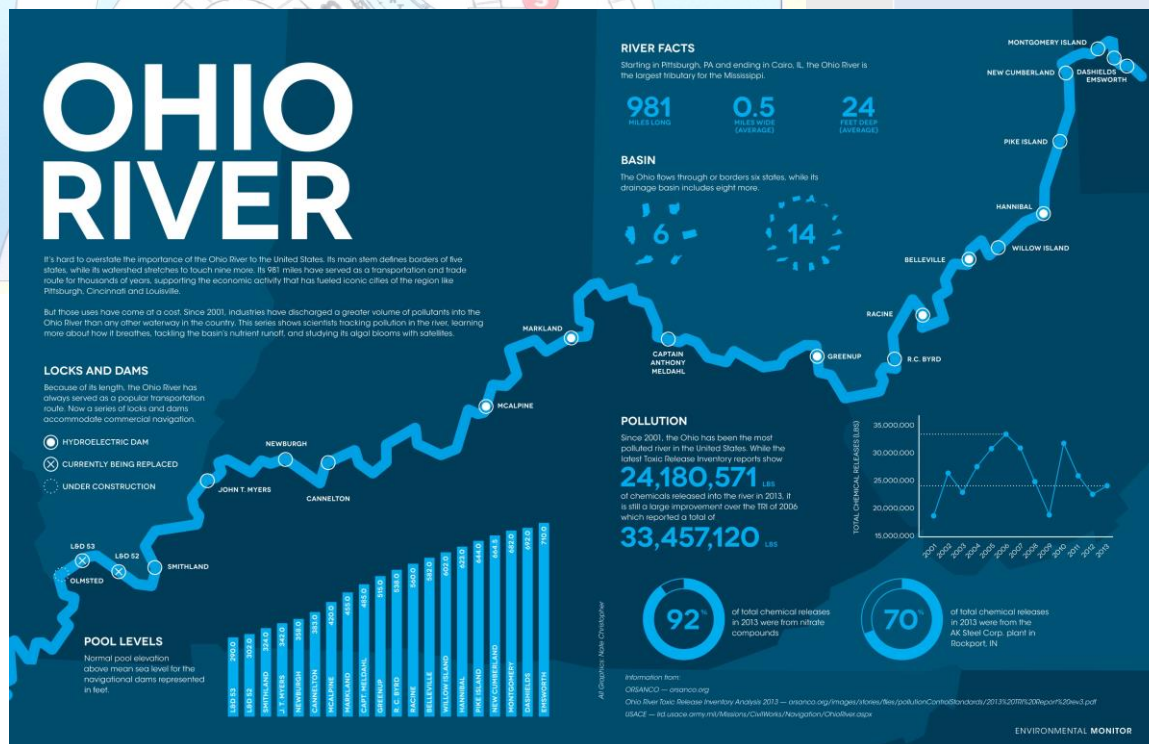
satisfied ■ Somewhat satisfied ■ Very satisfied ■ Completely satisfied

40%	47%
36%	47%

Features A and B continue to top user satisfaction

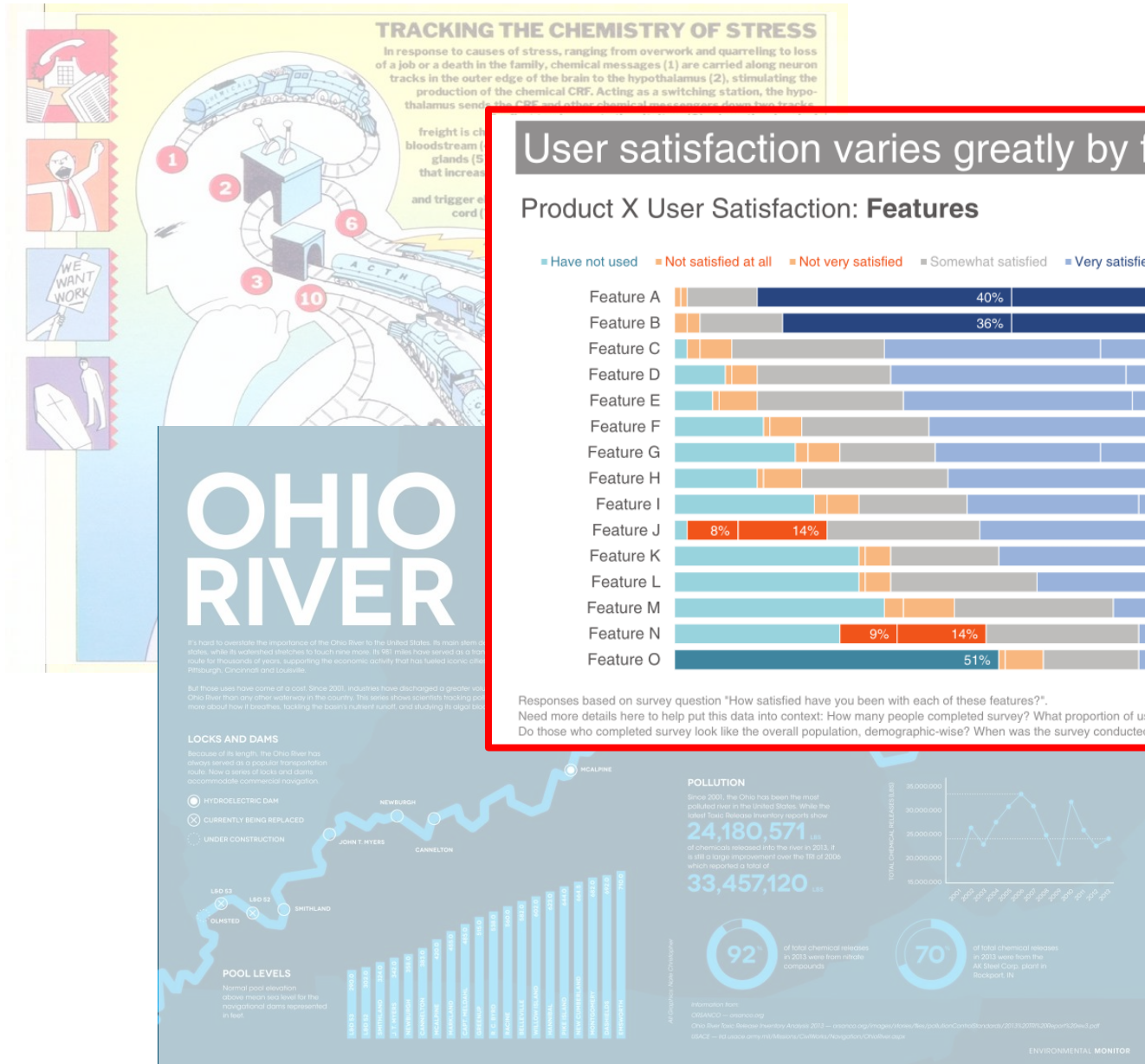
Users are least satisfied with Features J and N; what improvements can we make here for a better user experience?

Feature O is least used. What steps can we proactively take with existing users to increase utilization?



does this represent?

1. Was ist Datenvisualisierung?



2. Warum brauchen wir Visualisierungen?

A) Verarbeitung des Gehirns

B) Komplexität des Themas

Durchschnitt (Avg) von x: 9

Standardabweichung von x: 11

Durchschnitt (Avg) von y: 7,50

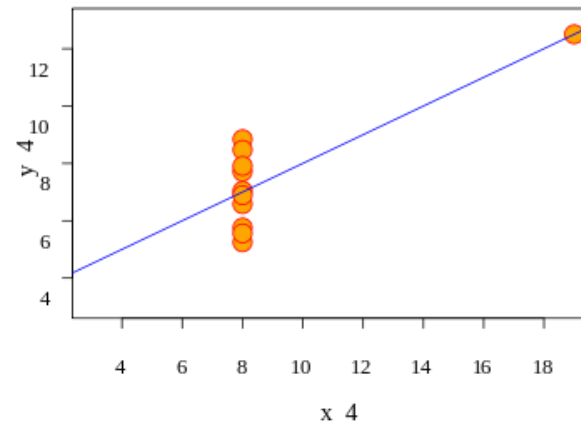
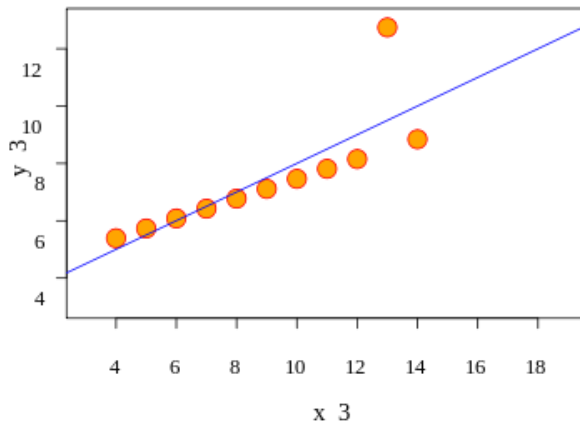
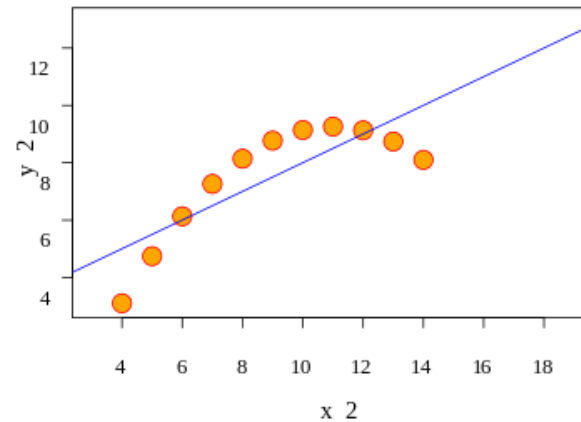
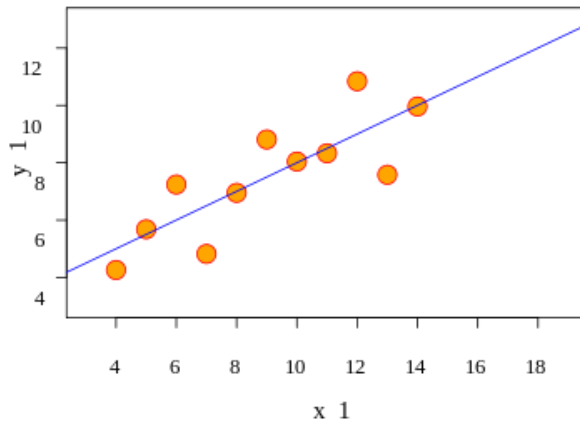
Standardabweichung von y: 4,1

Korrelationswert: 0,816

Regressionsgleichung: $y = 3,00 + 0,500x$

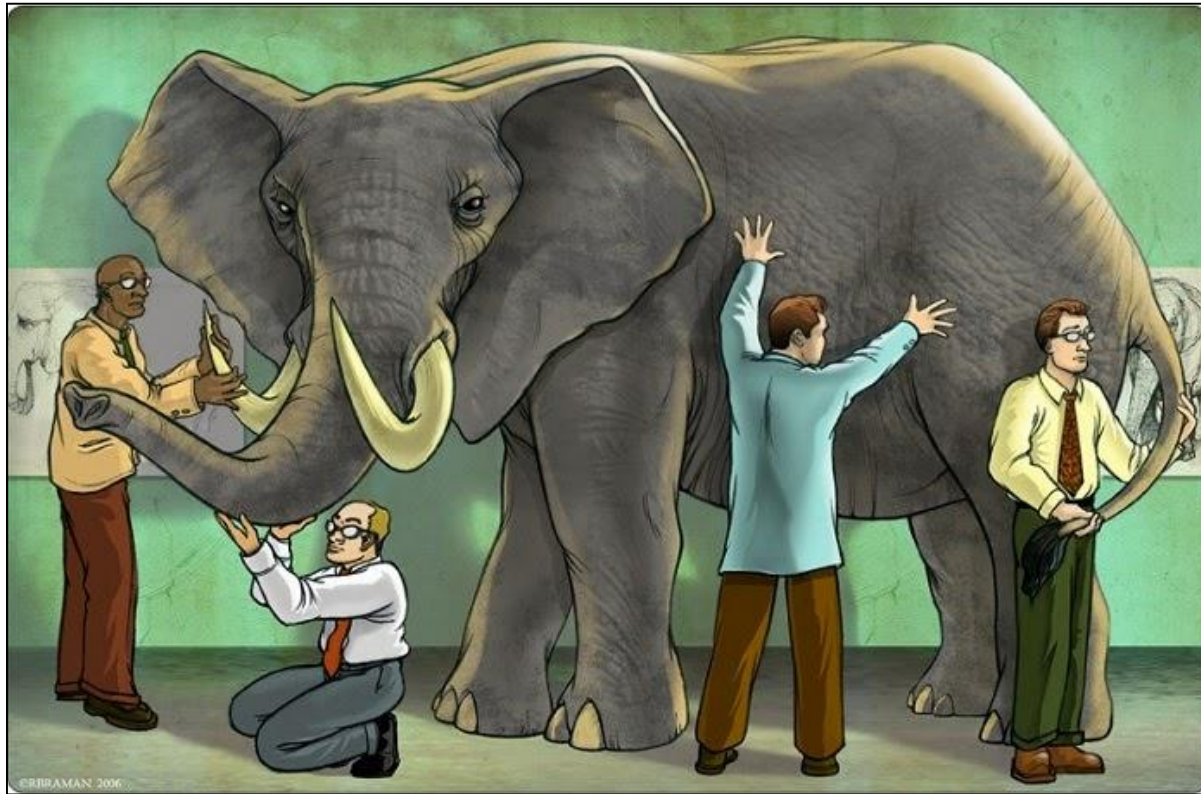
2. Warum brauchen wir Visualisierungen?

Francis Anscombes Quartett, 1973



2. Warum brauchen wir Visualisierungen?

„Ein Bild sagt mehr als 1000 Worte“

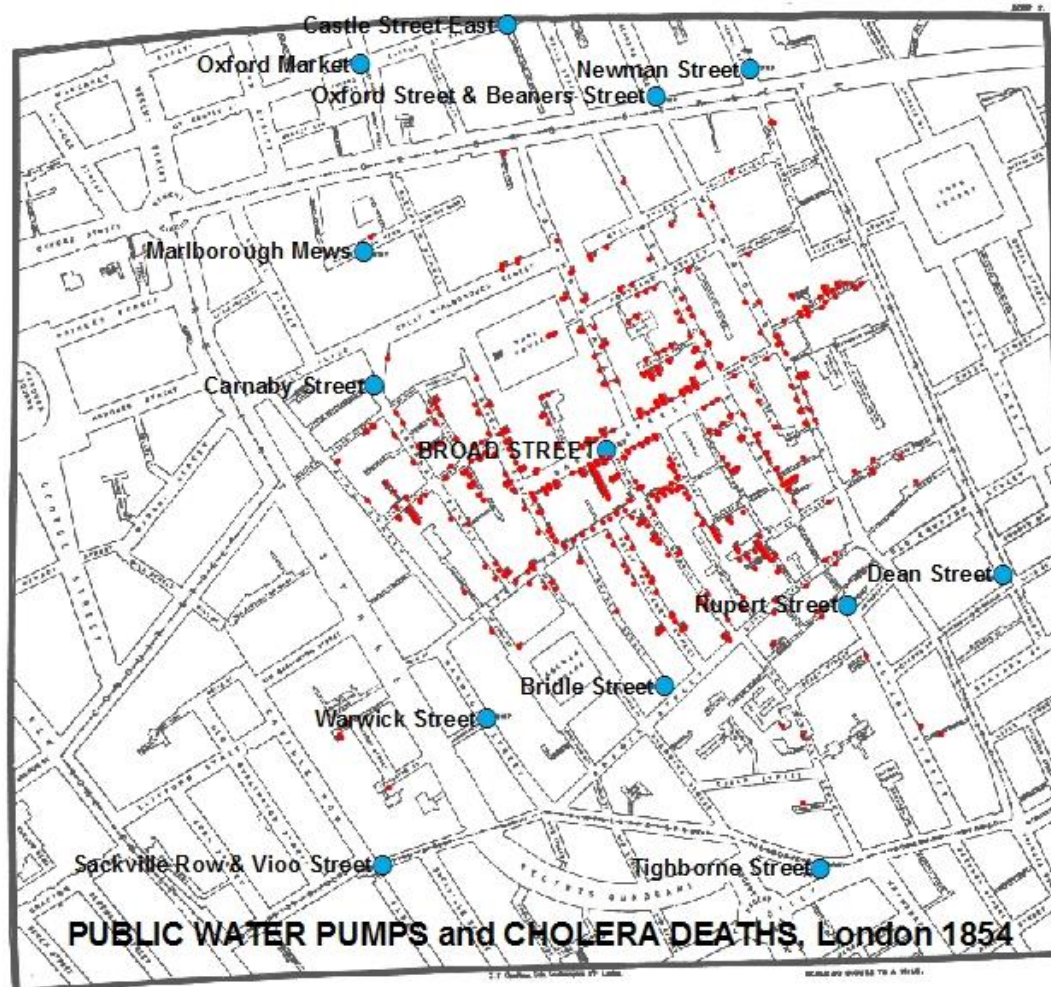


3. Wir stehen auf den Schultern von Giganten



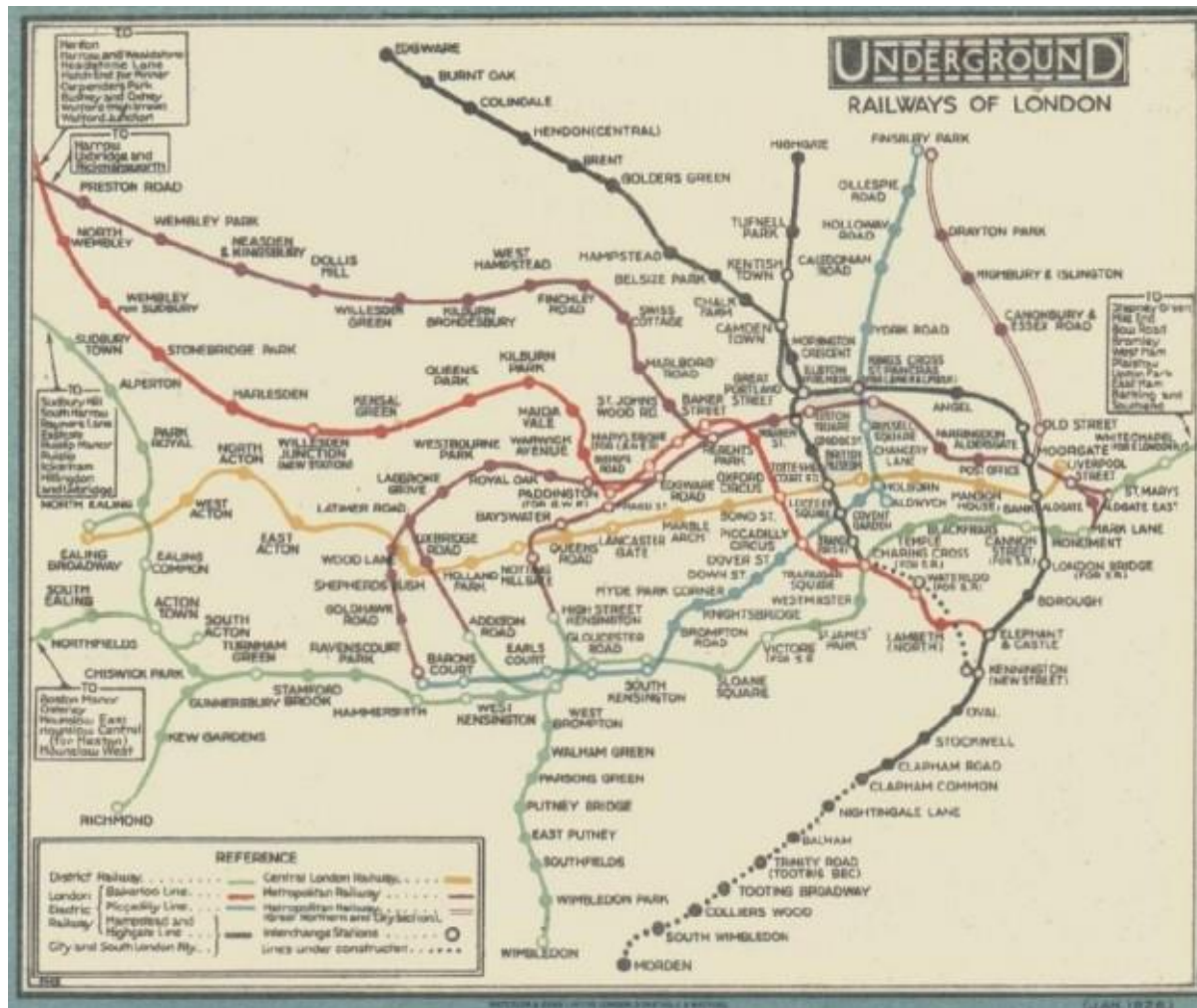
3. Wir stehen auf den Schultern von Giganten

John Snow, 1854



3. Wir stehen auf den Schultern von Giganten

1928

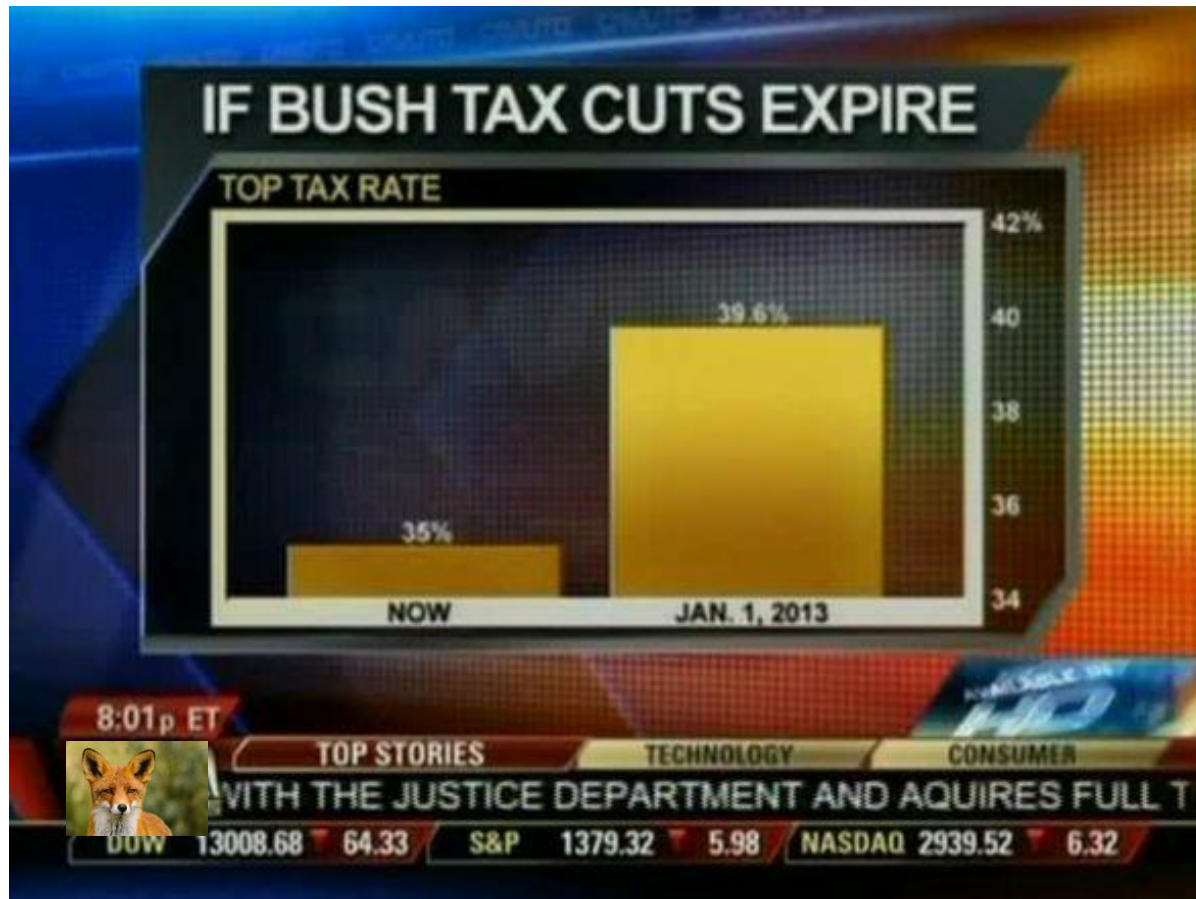


Harry Beck, 1933



4. „Falsche“ Visualisierungen

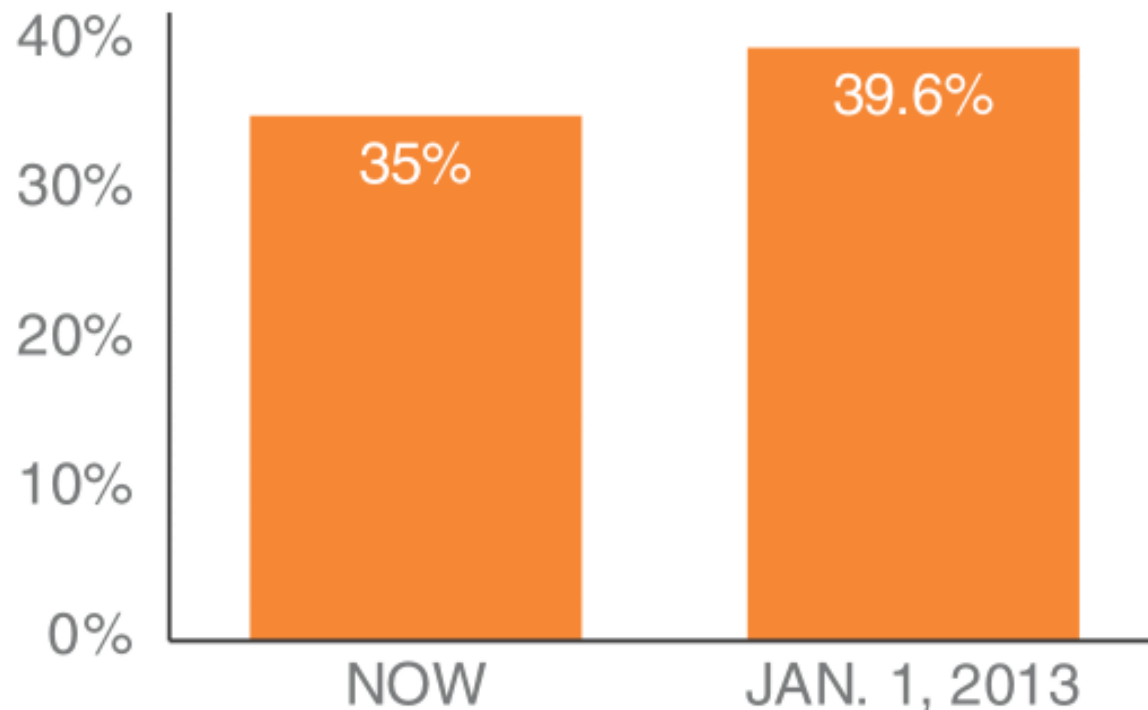
Aktive Manipulation?



4. „Falsche“ Visualisierungen

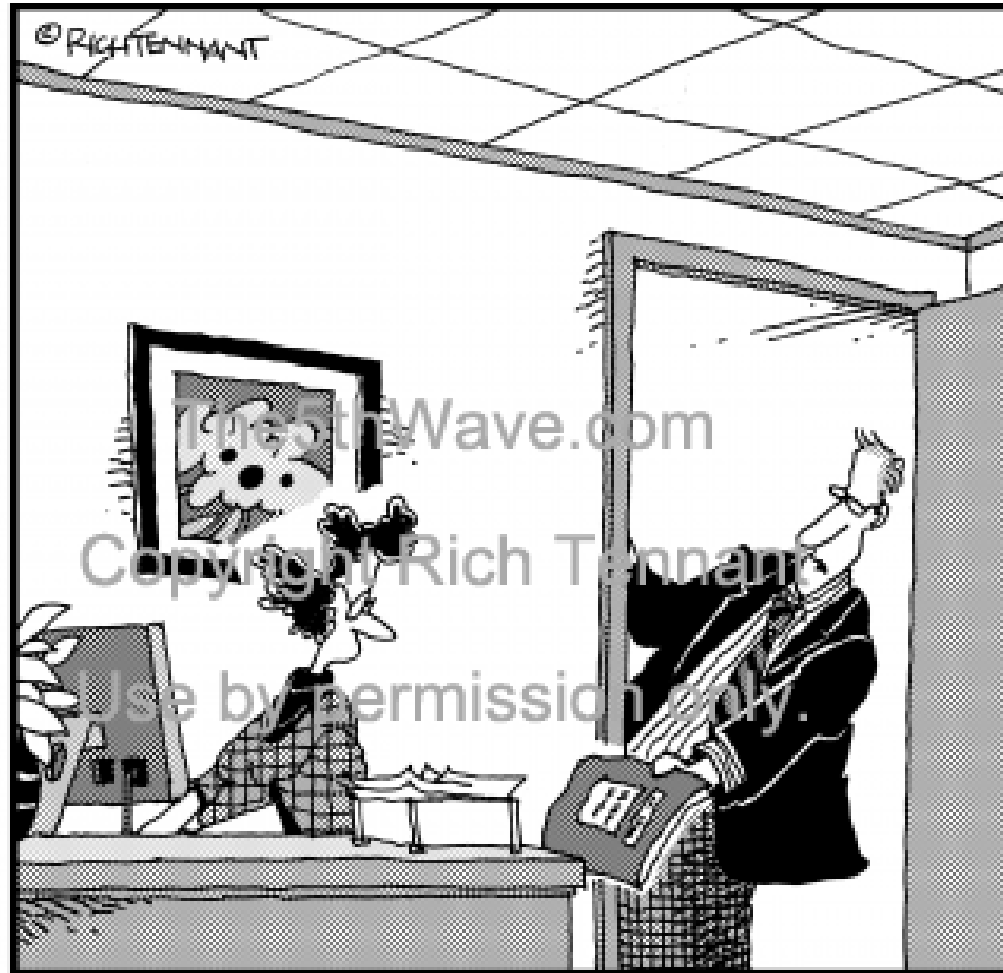
Aktive Manipulation?

IF BUSH TAX CUTS EXPIRE
TOP TAX RATE



https://en.wikipedia.org/wiki/Misleading_graph

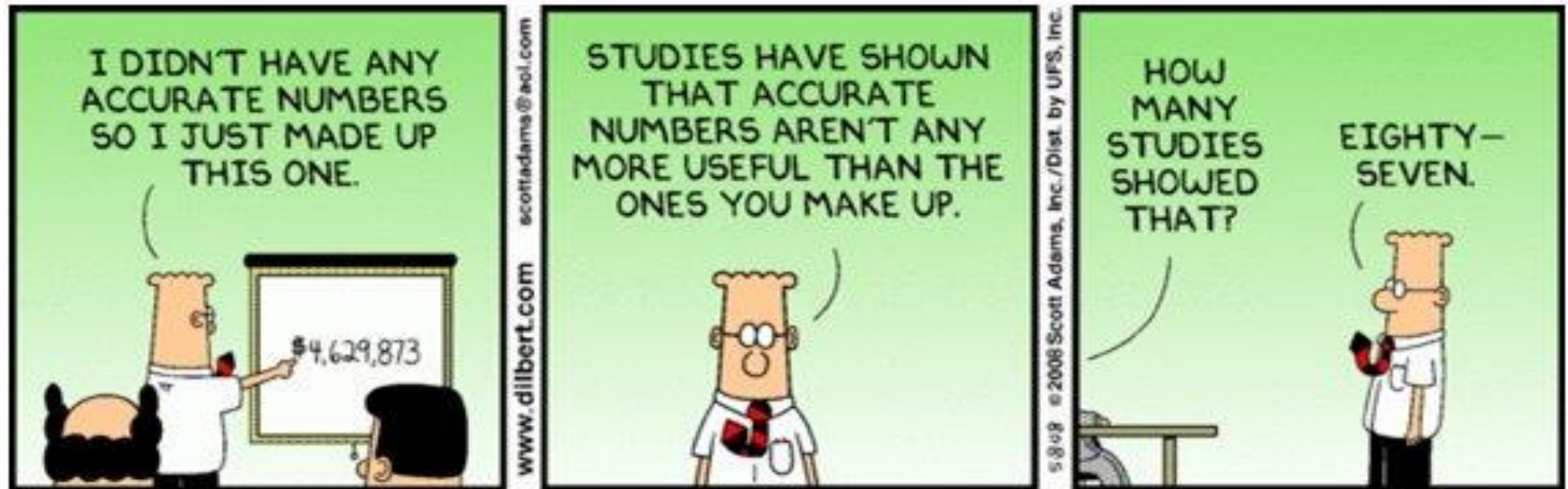
4. „Falsche“ Visualisierungen



"I don't like the numbers on this report. Send it down to Marketing and see what they can do with it."

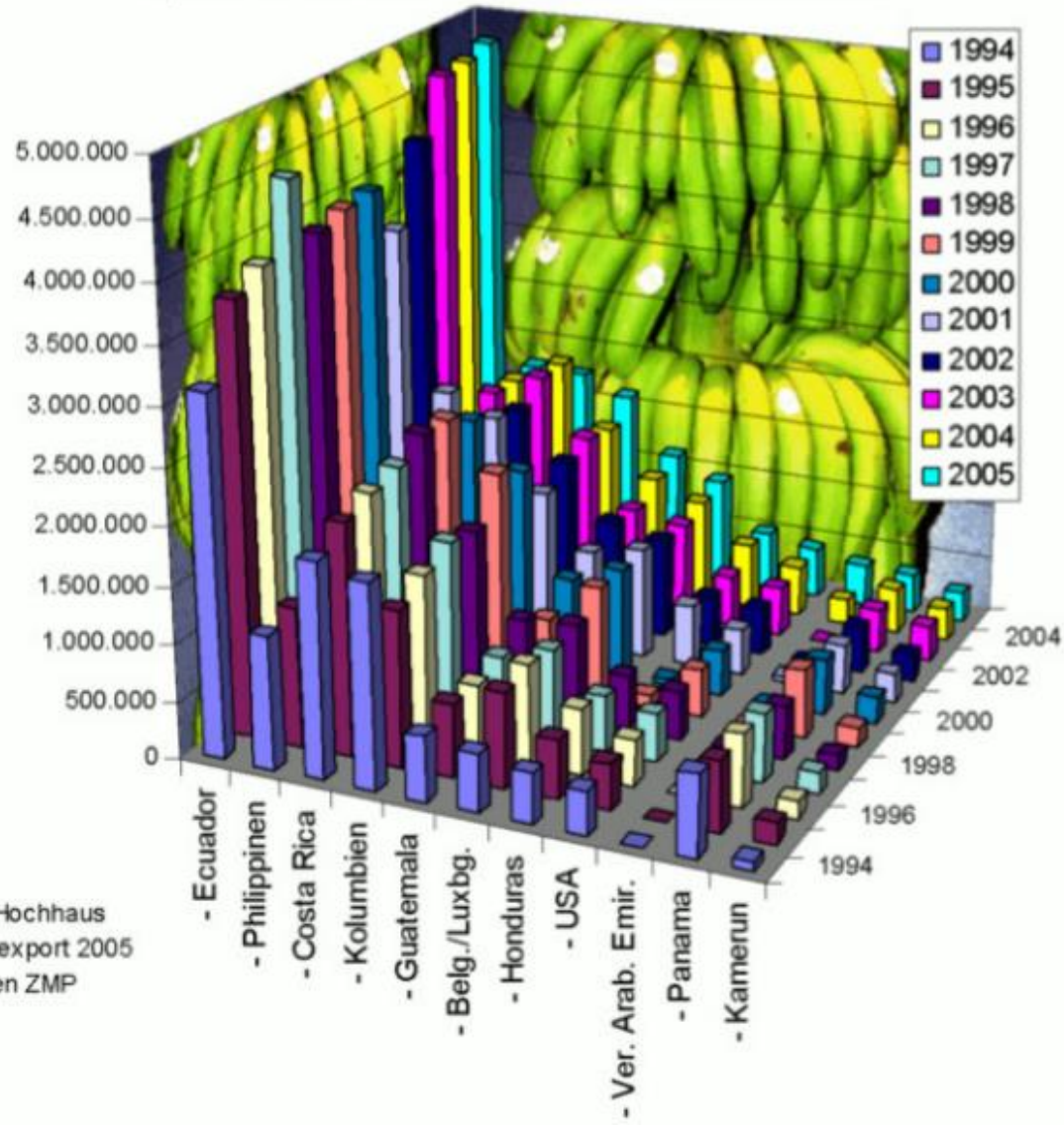
4. „Falsche“ Visualisierungen

„Lügen mit Statistik“...



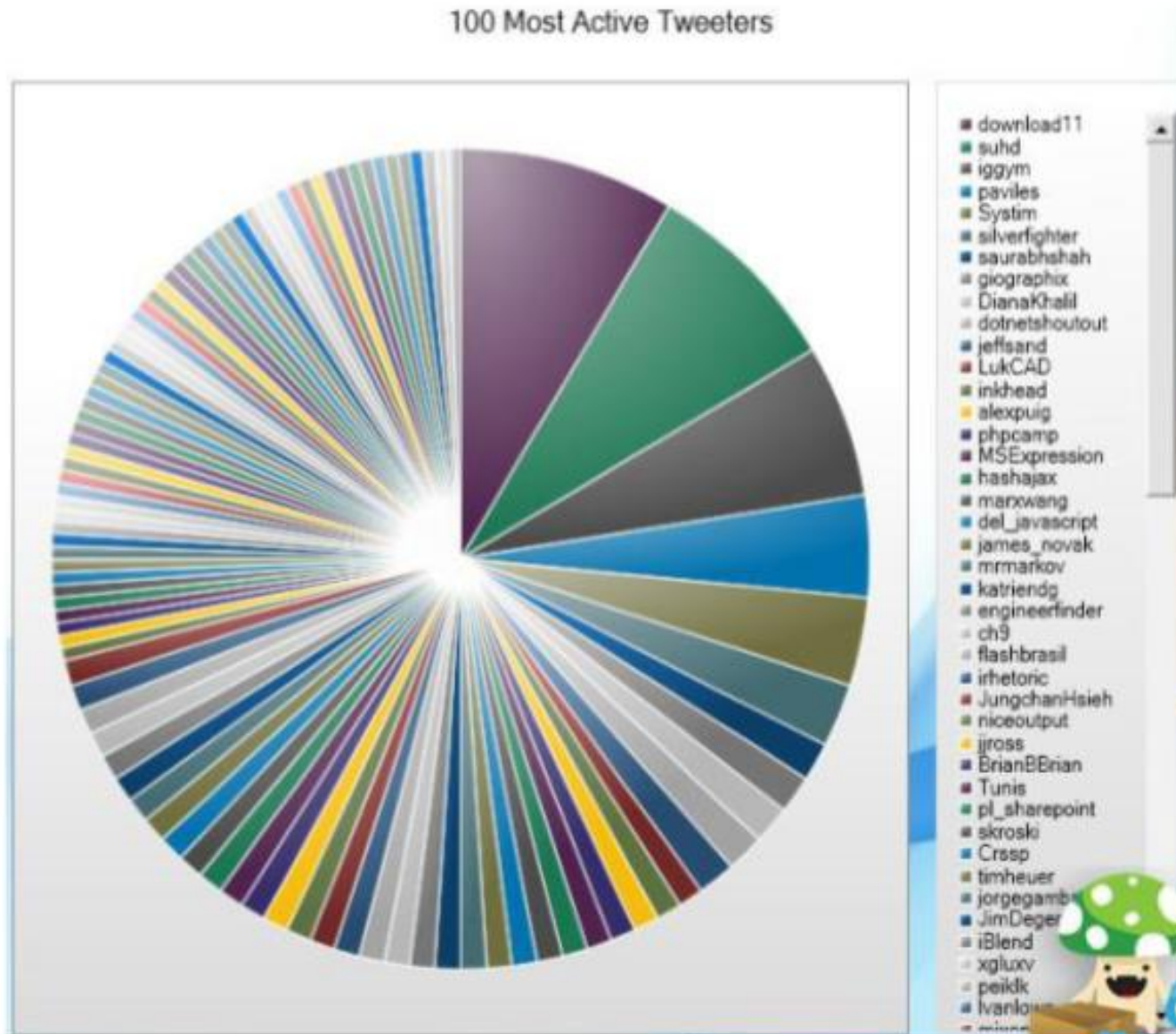
4. „Falsche“ Visualisierungen

Export von Bananen in Tonnen von 1994-2005

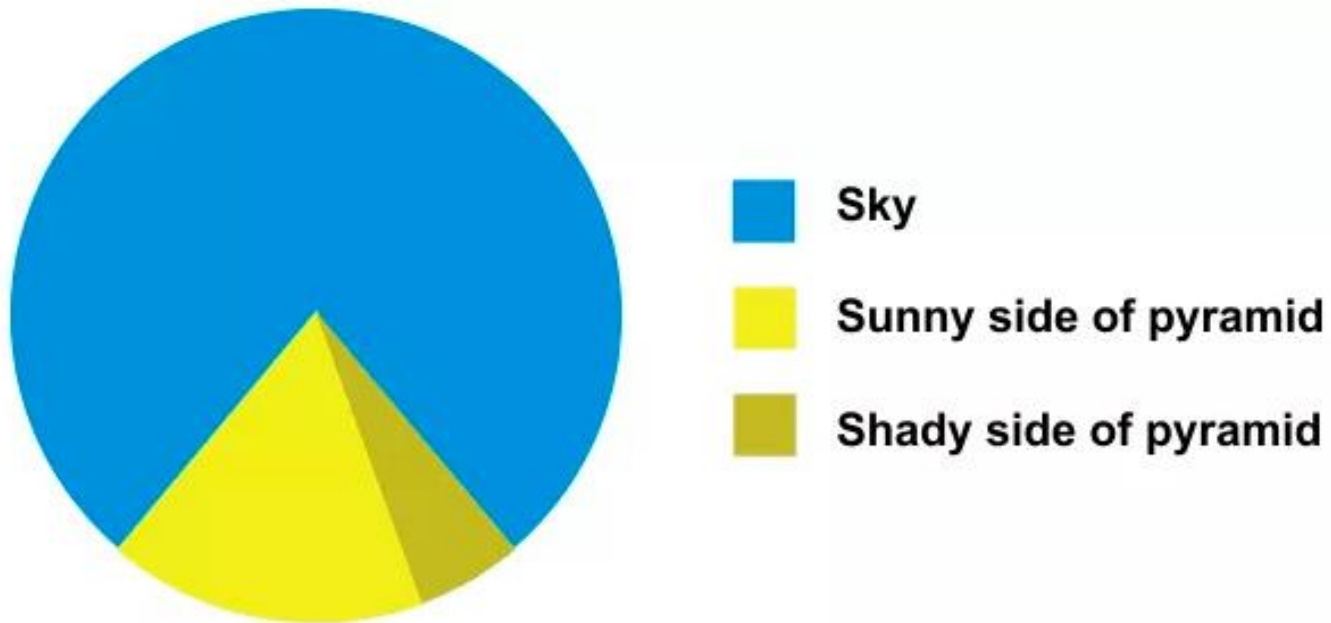


Dr. Hochhaus
Banexport 2005
Daten ZMP

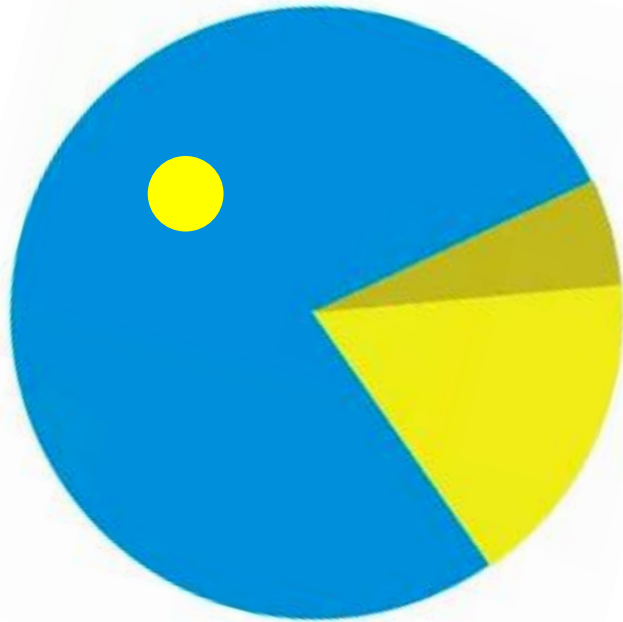
4. „Falsche“ Visualisierungen



5. Richtlinien für Datenvisualisierungen



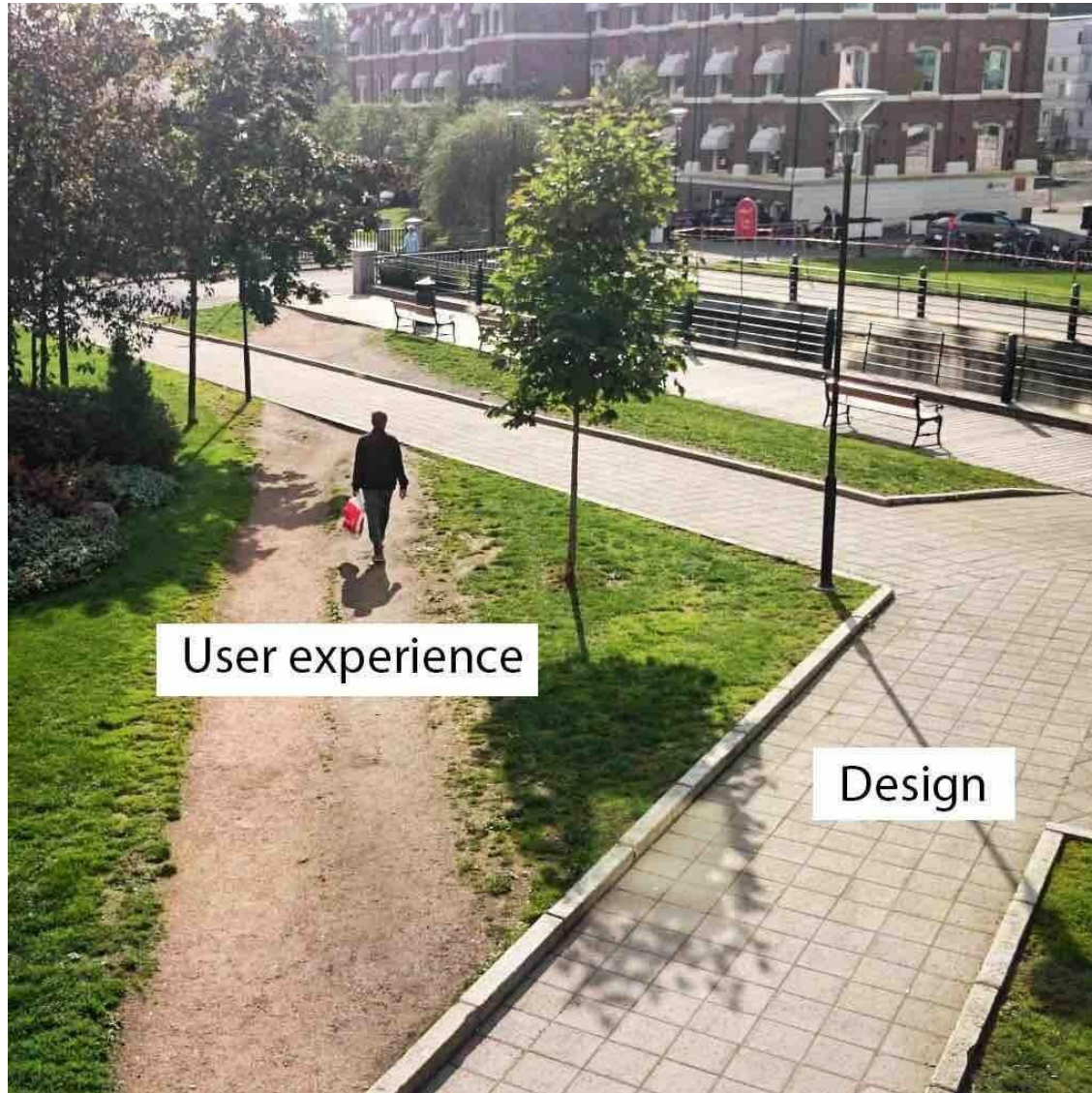
5. Richtlinien für Datenvisualisierungen



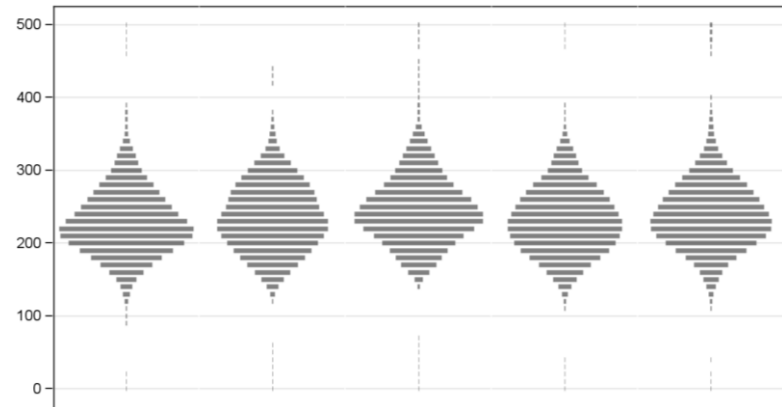
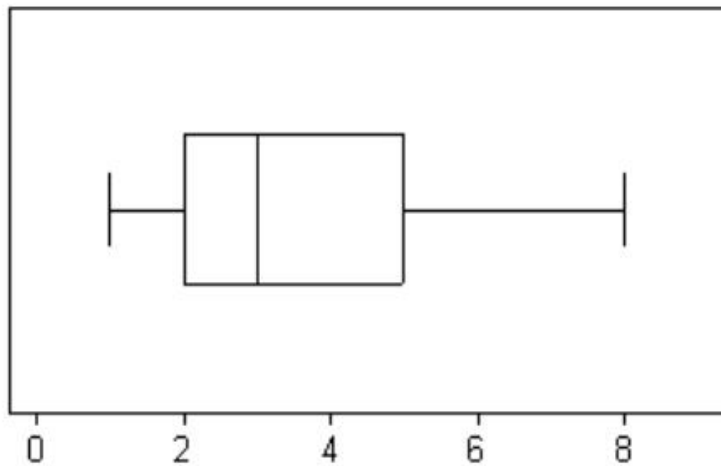
5. Richtlinien für Datenvisualisierungen

- Ziel festlegen – was will ich ausdrücken?
- Auf Zielgruppe ausrichten – wie ist der Kenntnisstand?
- Das passende Diagramm – welches eignet sich für diese Aussage/ Ziel/Zielgruppe am Besten?
- Farbwahl – wie wird sie wahrgenommen?
- Psychologie – wie präsentiere ich „gehirngerechte Häppchen“?
- Design – was ist ästhetisch?

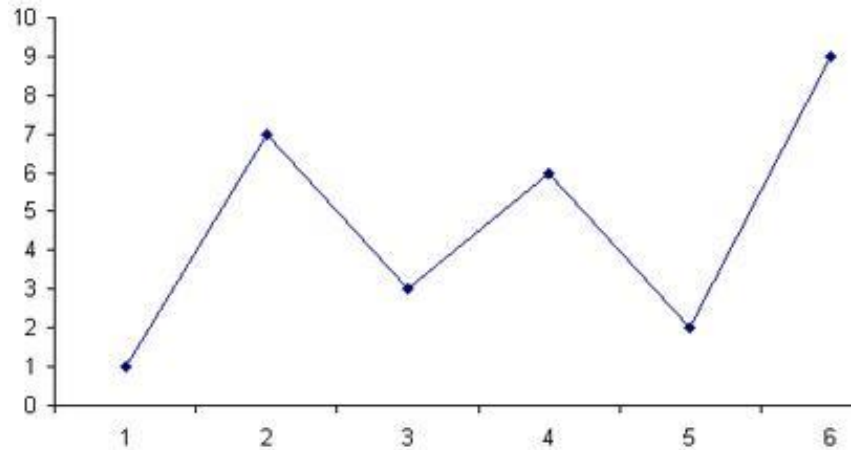
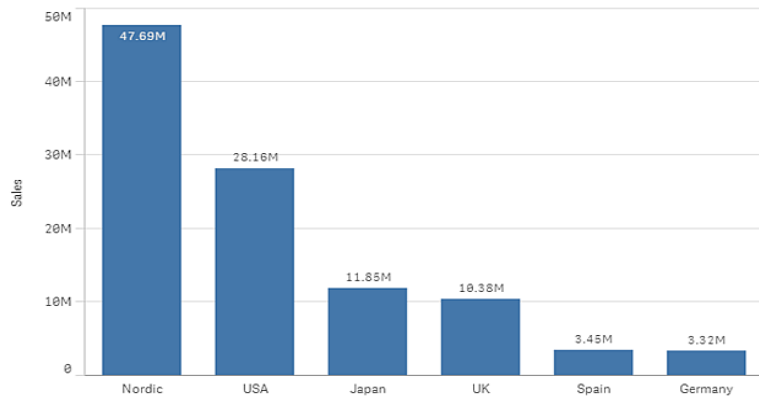
5. Richtlinien für Datenvisualisierungen - Ziel + Grundsätzlich



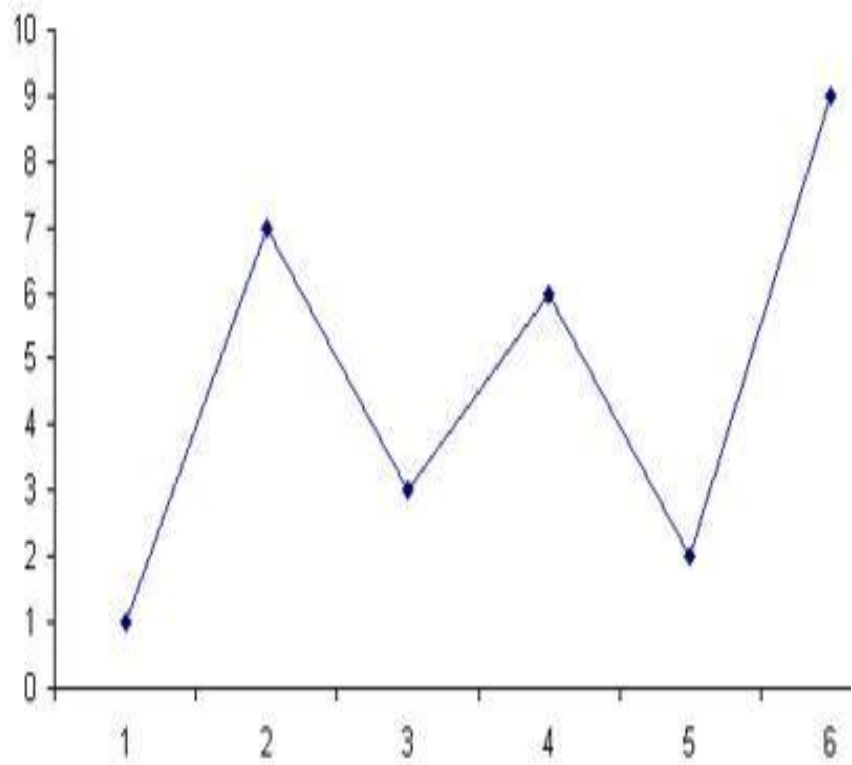
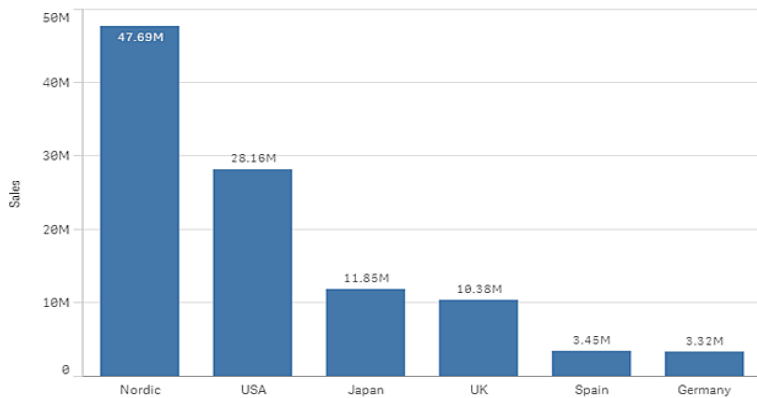
5. Richtlinien für Datenvisualisierungen - Zielgruppe



5. Richtlinien für Datenvisualisierungen - Zielgruppe



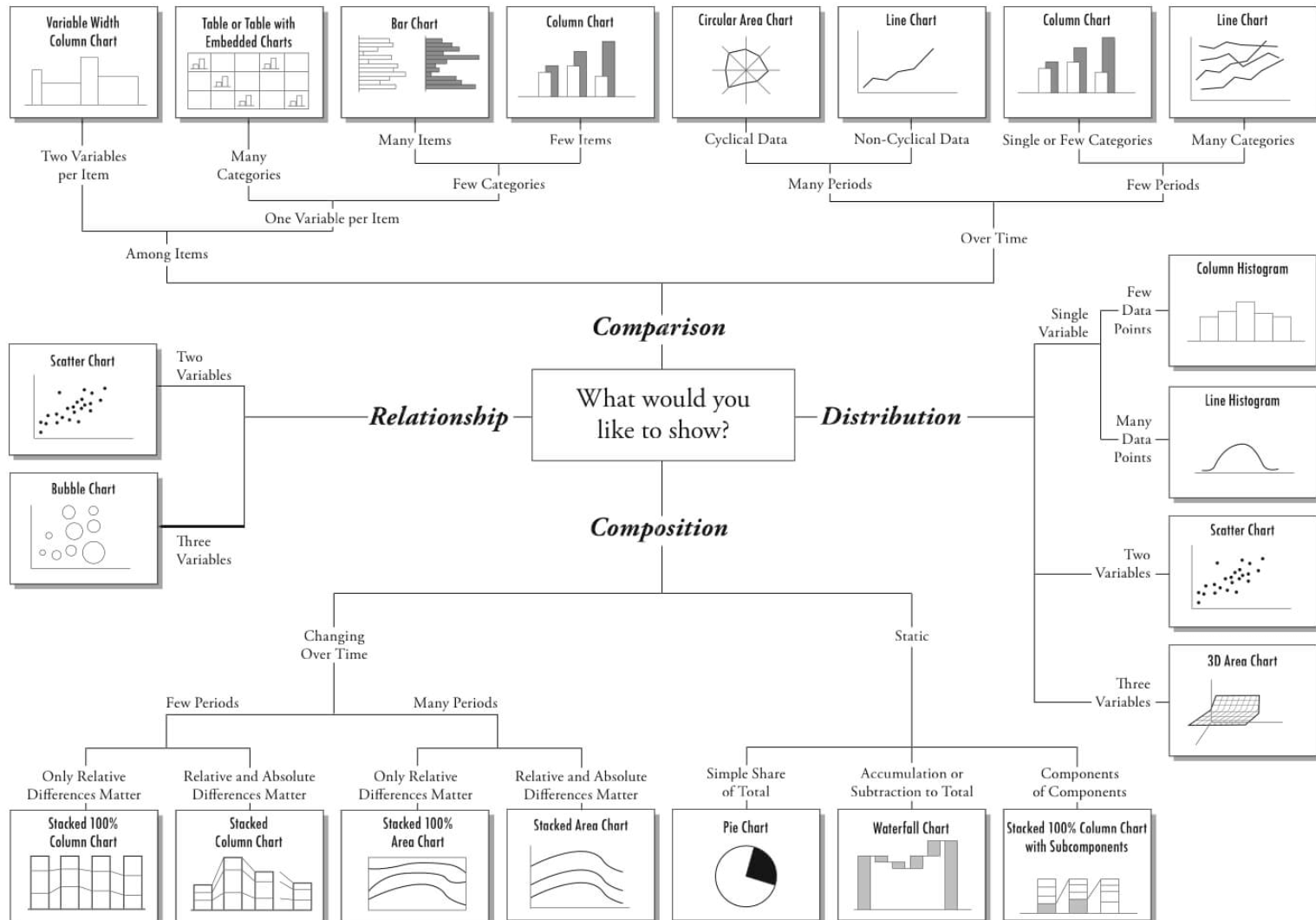
5. Richtlinien für Datenvisualisierungen - Zielgruppe



5. Richtlinien für Datenvisualisierungen - Diagrammart

Chart Suggestions—A Thought-Starter

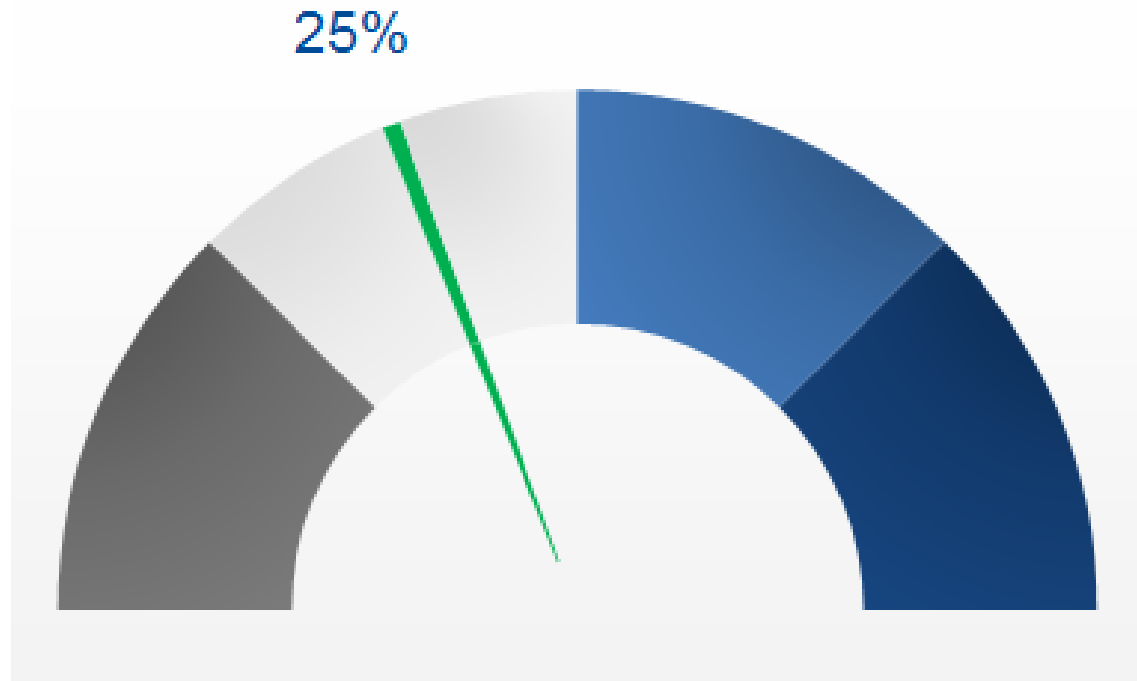
www.ExtremePresentation.com
© 2009 A. Abela — a.v.abela@gmail.com



5. Richtlinien für Datenvisualisierungen - Diagrammart



5. Richtlinien für Datenvisualisierungen - Diagrammart



5. Richtlinien für Datenvisualisierungen - Farbe

Country Level Sales Rank Top 5 Drugs

Rainbow distribution in color indicates sales rank in given country from #1 (red) to #10 or higher (dark purple)

Country	A	B	C	D	E
AUS	1	2	3	6	7
BRA	1	3	4	5	6
CAN	2	3	6	12	8
CHI	1	2	8	4	7
FRA	3	2	4	8	10
GER	3	1	6	5	4
IND	4	1	8	10	5
ITA	2	4	10	9	8
MEX	1	5	4	6	3
RUS	4	3	7	9	12
SPA	2	3	4	5	11
TUR	7	2	3	4	8
UK	1	2	3	6	7
US	1	2	4	3	5

5. Richtlinien für Datenvisualisierungen - Farbe

Top 5 drugs: country-level sales rank

RANK		1	2	3	4	5+
COUNTRY DRUG		A	B	C	D	E
Australia		1	2	3	6	7
Brazil		1	3	4	5	6
Canada		2	3	6	12	8
China		1	2	8	4	7
France		3	2	4	8	10
Germany		3	1	6	5	4
India		4	1	8	10	5
Italy		2	4	10	9	8
Mexico		1	5	4	6	3
Russia		4	3	7	9	12
Spain		2	3	4	5	11
Turkey		7	2	3	4	8
United Kingdom		1	2	3	6	7
United States		1	2	4	3	5

5. Richtlinien für Datenvisualisierungen - Farbe

Earthtones



Cool



Unnatural colors



Warm

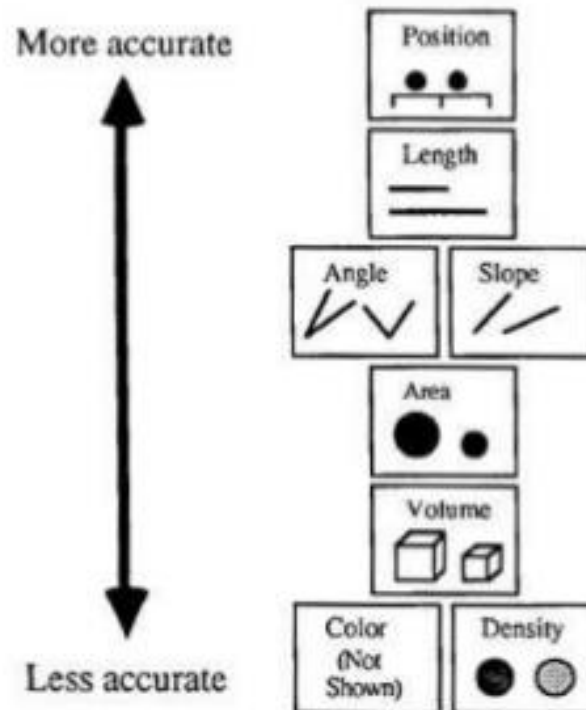


Increasing color intensity



5. Richtlinien für Datenvisualisierungen – Häppchen

Prä-attentive Wahrnehmung



Mackinlay, APT (A Presentation Tool), 1986

Slide: Cecilia Aragon, HCDE, UW

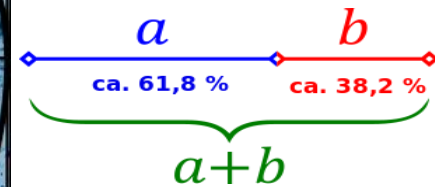
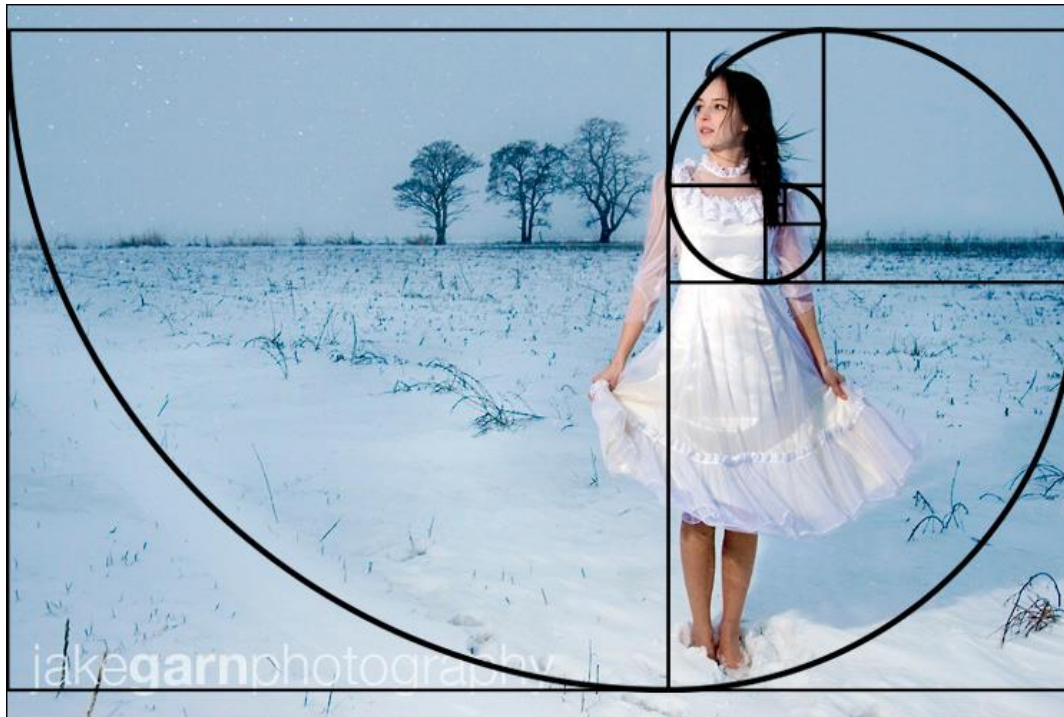
5. Richtlinien für Datenvisualisierungen – Häppchen

Gestalt - Prinzipien



5. Richtlinien für Datenvisualisierungen - Design

Ästhetik



6. Software für Analyse und Visualisierungen

Jan. 2017: BI + Analytics Leader nach Gartner

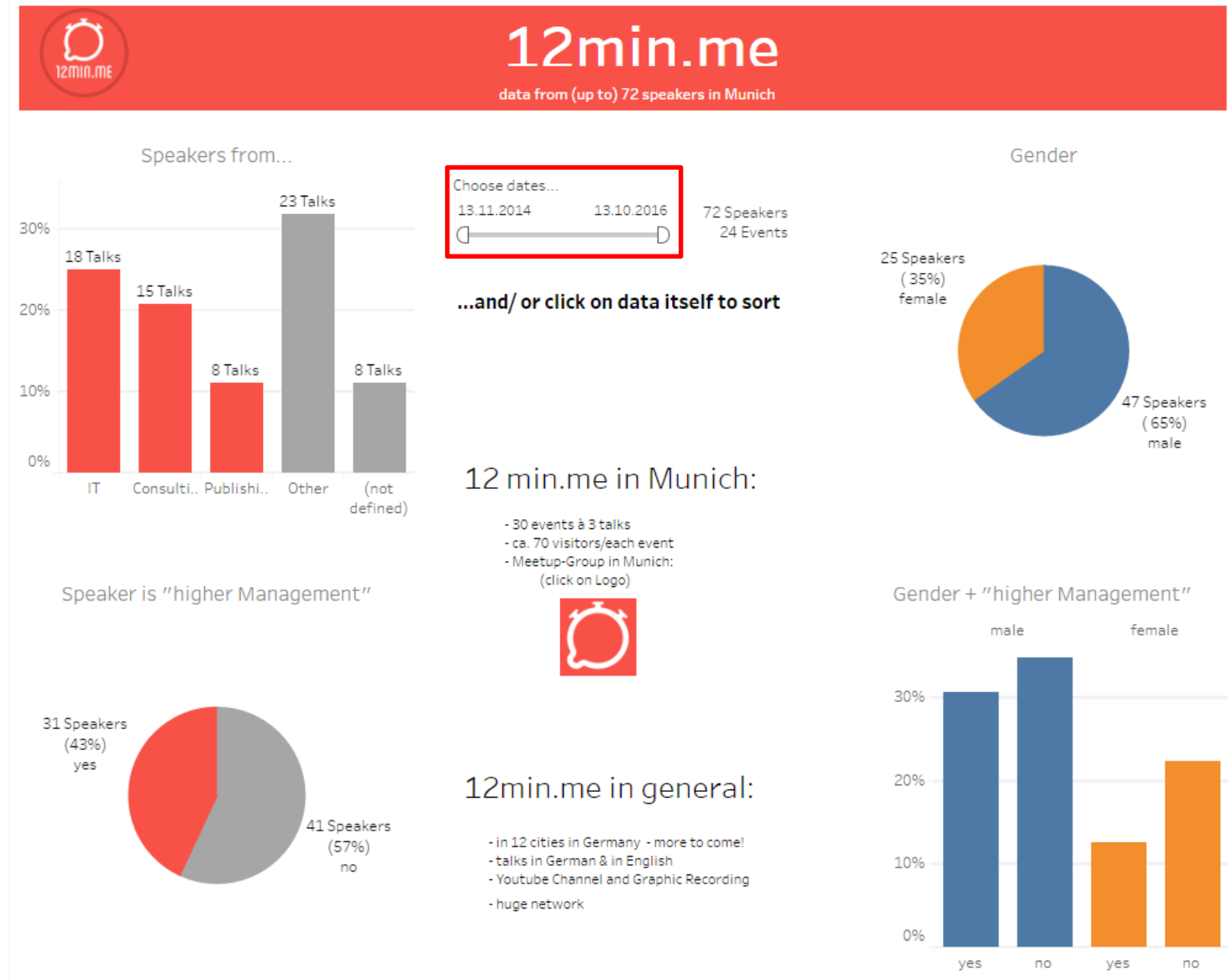
- Tableau
- Microsoft (PowerBI)
- Qlik



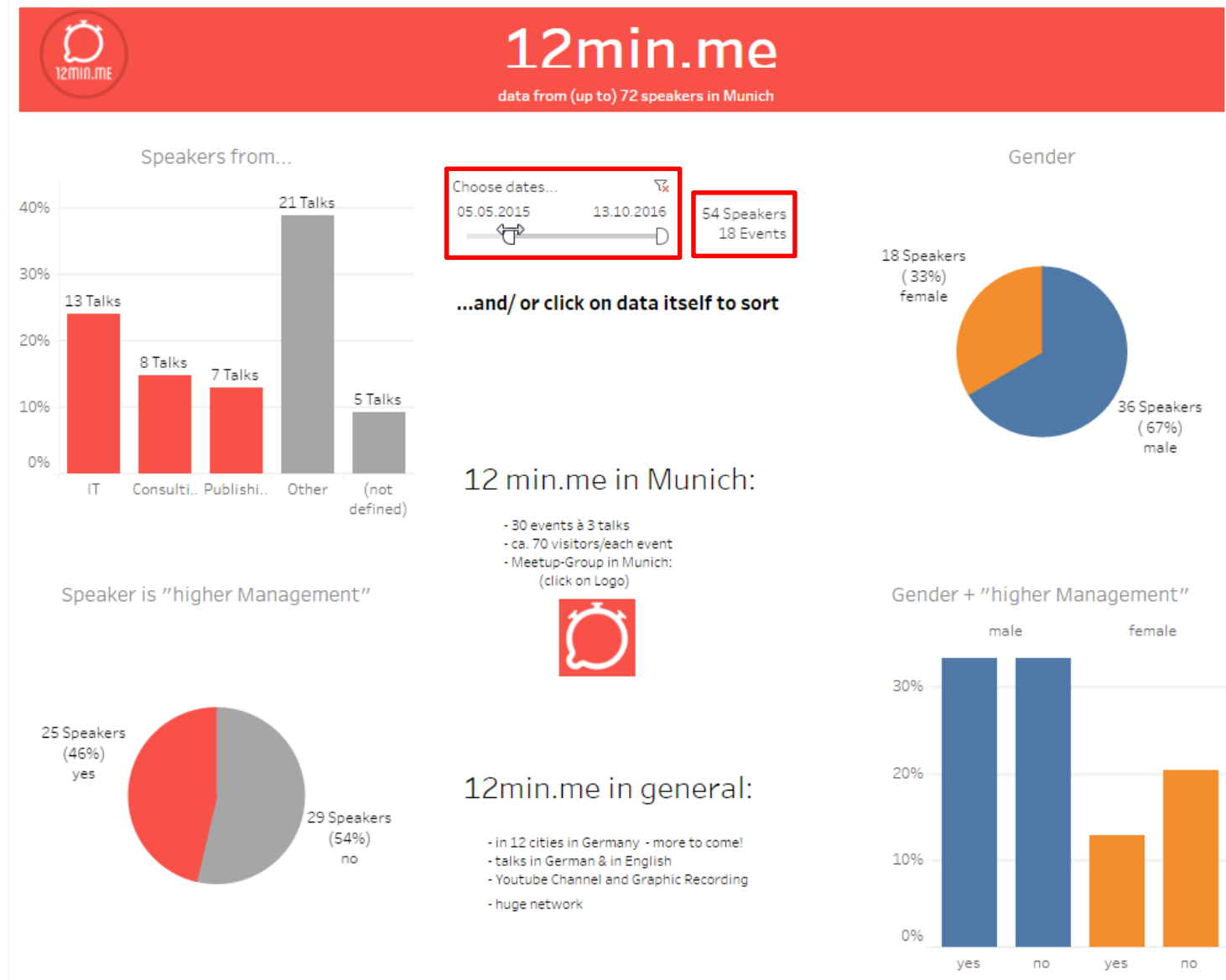
Wohlgemerkt:

- Nicht „Advanced Analytics“
- Nicht „Infografik-Programme“
- Nicht „Viz-Packages“ für Sprachen

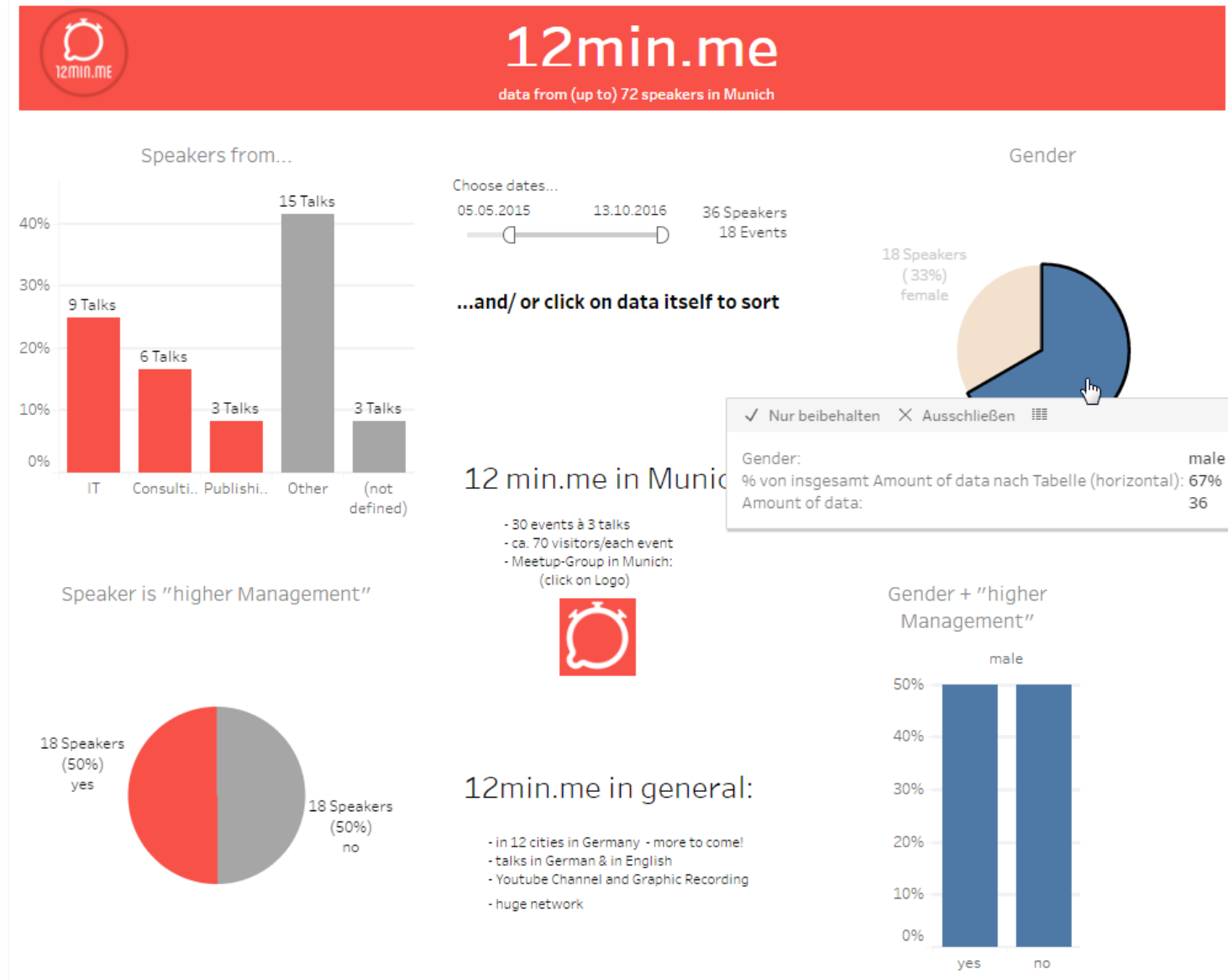
6. Software für Analyse und Visualisierungen



6. Software für Analyse und Visualisierungen



6. Software für Analyse und Visualisierungen



Software: Tableau vs. Excel

Tableau:

Excel:

-	Kosten	0
+	Gruppenzusammenarbeit	0
+	gute Designvorlagen	-
+	für Big Data geeignet	-
-	Verbreitung	+
0	Lernkurve	0
~	„kann alles“ + Integration	+
-	Einzelne Punkte bearbeiten	+
-	„Bequeme“ Funktionen	+
+	Community	+
+	Dashboards, Storytelling	-

<http://www.cafepress.com/+excel+mugs>

Software: Tableau vs. Power BI

Tableau:

Power BI:

-	Kosten	+
+	Gruppenzusammenarbeit	+
+	gute Designvorlagen	~
+	für Big Data geeignet	+
-	Verbreitung	~
0	Lernkurve	?
-	„kann alles“ + Integration	?
~	Einzelne Punkte bearbeiten	?
-	„Bequeme“ Funktionen	?
+	Community	?
+	Dashboards, Storytelling	?

Mitnahmefolie

- Datenvisualisierung ist notwendig
- ...es gibt den Bereich schon länger
- Wir täuschen uns oft selbst - oder lassen uns täuschen

Richtlinien für Visualisierungen:

- Ziel + Zielgruppe im Auge behalten
 - Ein passendes Diagramm auswählen
 - Passende Farben
 - Wahrnehmungshäppchen + Ästhetik im Kopf behalten
-
- Tableau ist eine sehr gute Dashboard-Software
 - ...aber es gibt auch Power BI oder Qlikview...

**Für Fragen stehe ich
gerne zur Verfügung.**

Michael Schmidt, ms@grauschattierung.de

Quellen und weitergehende Informationen

allgemein + Blogs:

William Cleveland
Andy Cotgreave
Andy Kriebel
Ben Shneiderman
Edward Tufte
John Tukey
Colin Ware
Martin Wattenberg
Steve Wexler
Nathan Yau

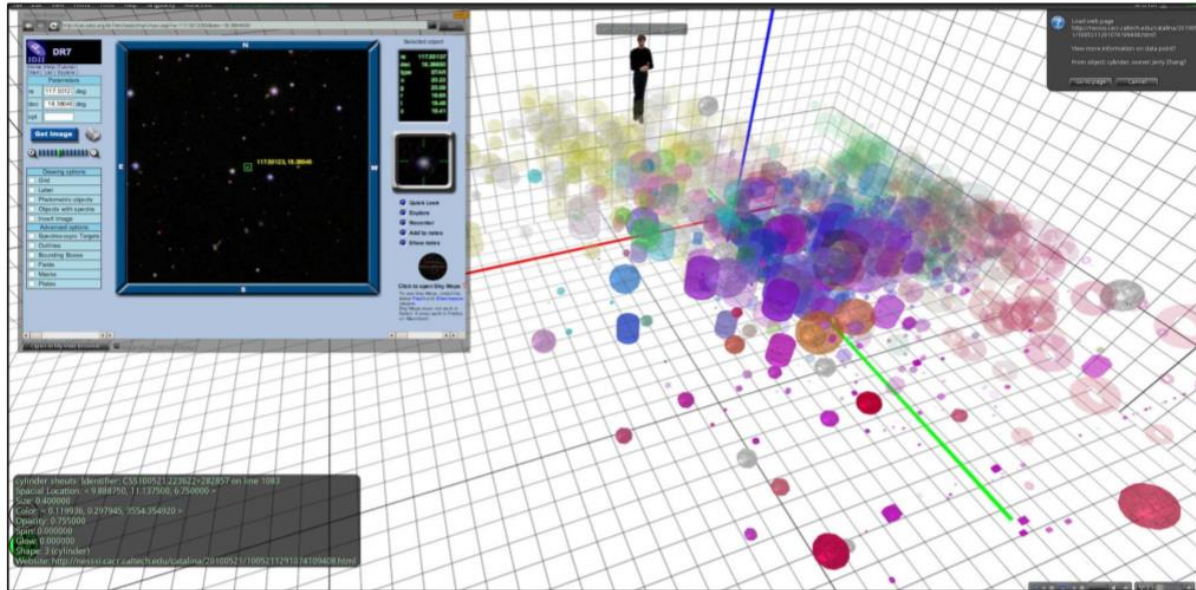
Videos:

Hans Rosling
[The best stats you've ever seen](#)
David McCandless
[The beauty of data visualization](#)
John Rauser
[How Humans See Data](#)
Steve Wexler
[Visualizing Survey Data with Tableau](#)

Bücher:

Albert Cairo: The Truthful Art: Data, Charts...
Cole – Nussbaumer – Knaflic: Storytelling with Data
Stephen Few: Information Dashboard Design

Wo geht die Reise hin?



- Mehr Datenpunkte
- zusätzliche Dimensionen
- Augmented Reality, Virtual Reality
- Automatisierung der Auswertung

Nicht besprochen

- Vertrauenswürdigkeit der Daten
- ETL – Extract, Transform, Load
- Big Data + Skalierung + Parallelisierung
- Advanced Analytics + Data Science
- Automatisierung
- Probleme beim Verständnis von statistischen Daten
- ...

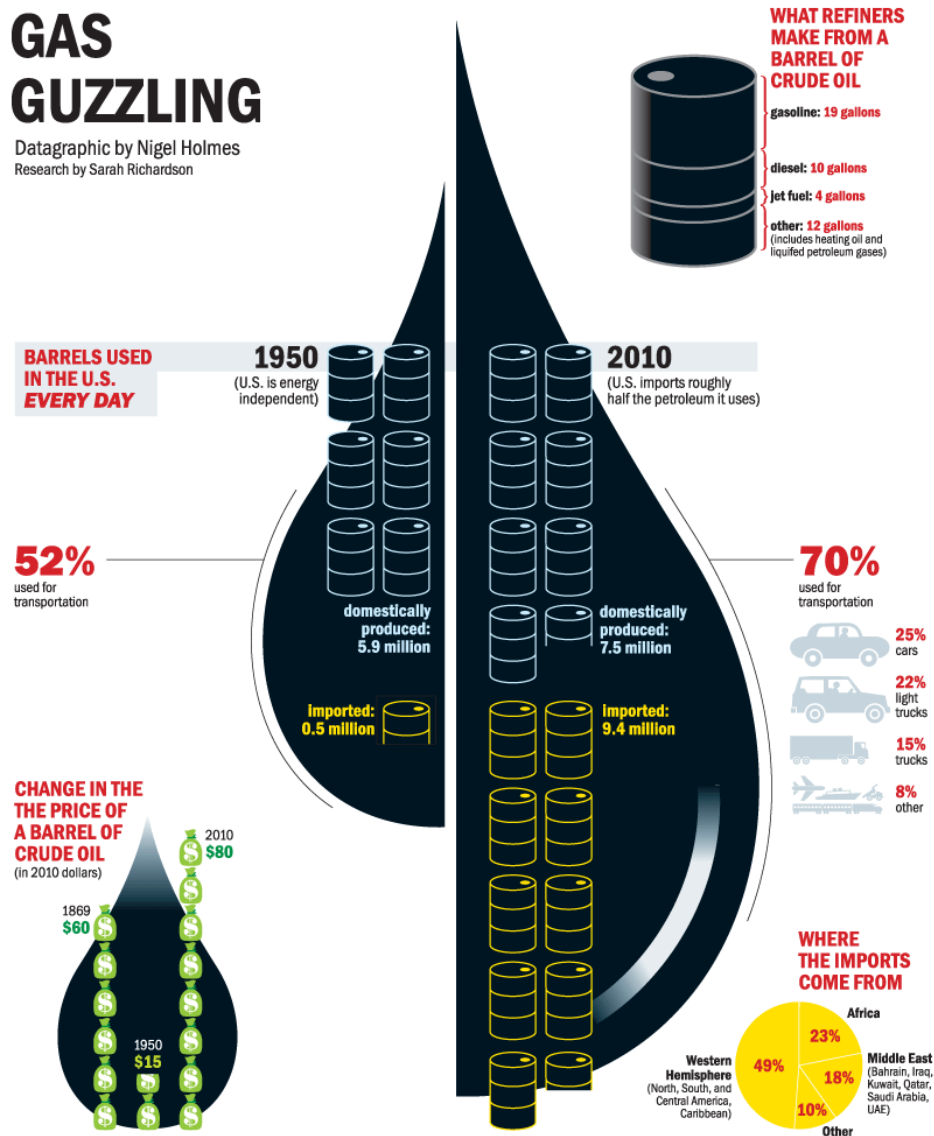
Dashboards...



Infographics + Data Art

GAS GUZZLING

Datagraphic by Nigel Holmes
Research by Sarah Richardson



Sources: EIA.gov; Transportation Energy Data Book Edition 30–2011; www.defra.gov.uk

Petroleum refers to crude oil plus products made in the refining of oil and natural gas

Infographics + Data Art

Close-Ups of the Genome, Species by Species by Species

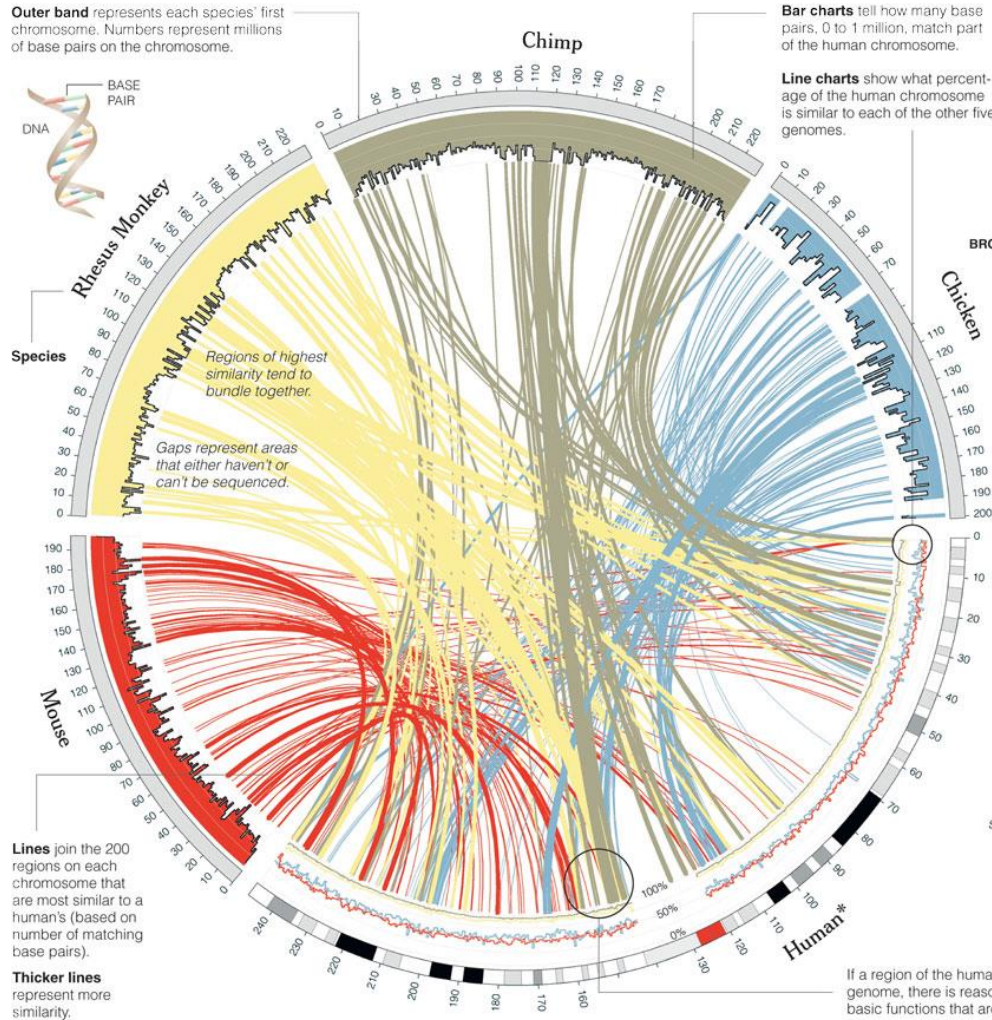
Scientists are sequencing the genomes of more than 70 organisms. The availability of these sequences has given rise to the field of comparative genomics, which seeks to answer questions about one animal's genome using information derived from another. A Canadian genomics scientist, Martin Krzywinski, has created a computer program called

Circos that aids in visualizing and comparing the data. The large diagram below illustrates the large degree of similarity between the first chromosomes of four animals to that of a human. Not surprisingly, the humans' is closest to the chimp's.

DAVID CONSTANTINE

COMPARING CHROMOSOMES 1

Outer band represents each species' first chromosome. Numbers represent millions of base pairs on the chromosome.

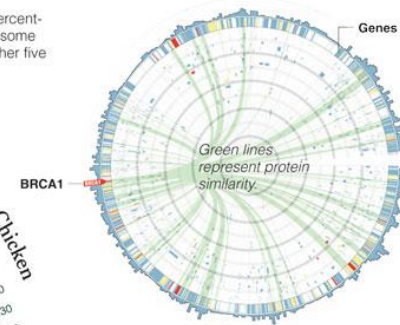


Bar charts tell how many base pairs, 0 to 1 million, match part of the human chromosome.

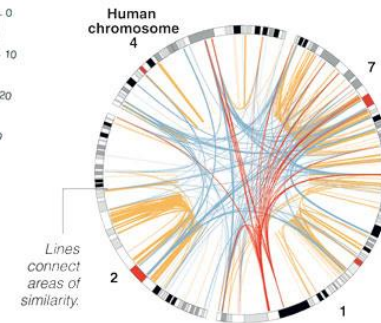
Line charts show what percentage of the human chromosome is similar to each of the other five genomes.

OTHER TYPES OF COMPARISONS

To download the free program or view other examples: <http://mkweb.bcgsc.ca/circos/>



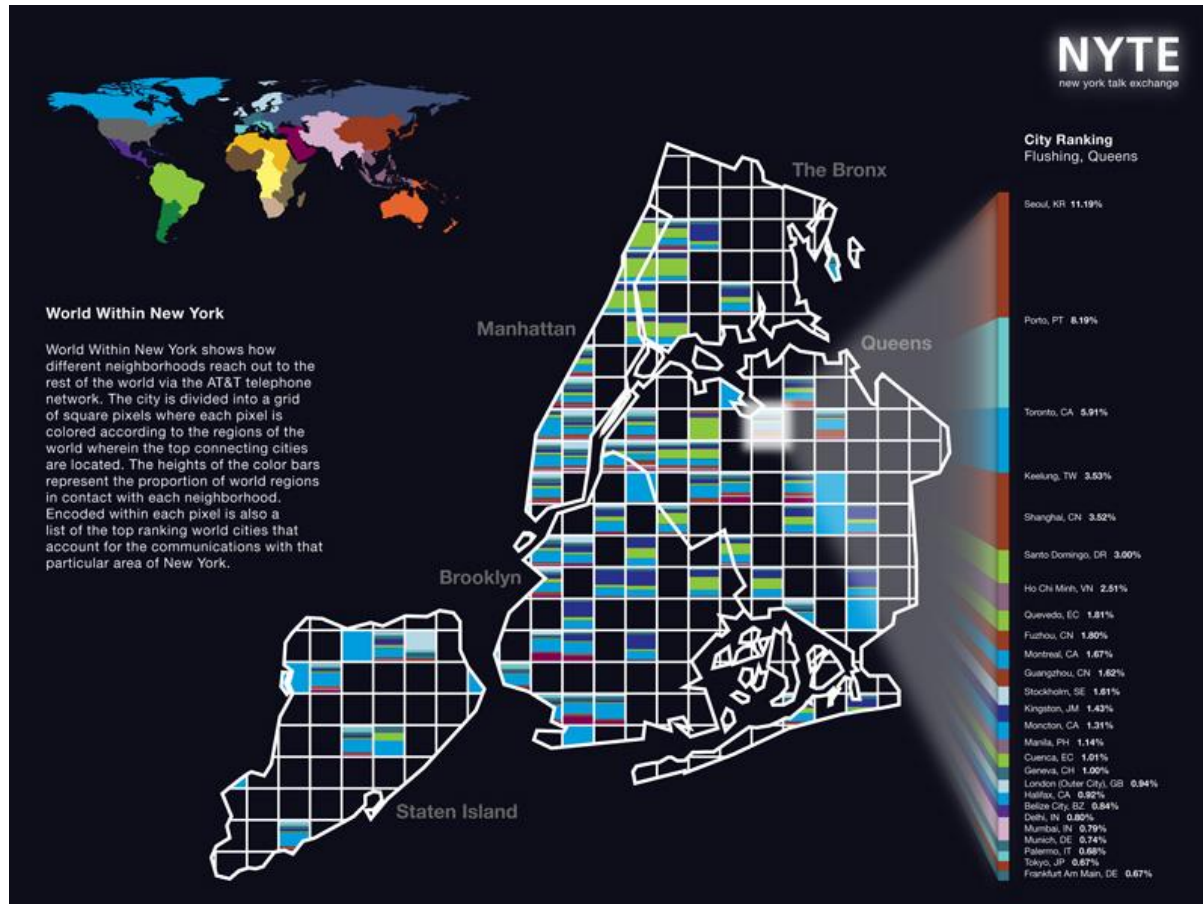
The chart above shows the similarity of the BRCA1 protein, implicated in early breast cancer, to other genes on human chromosome 17.



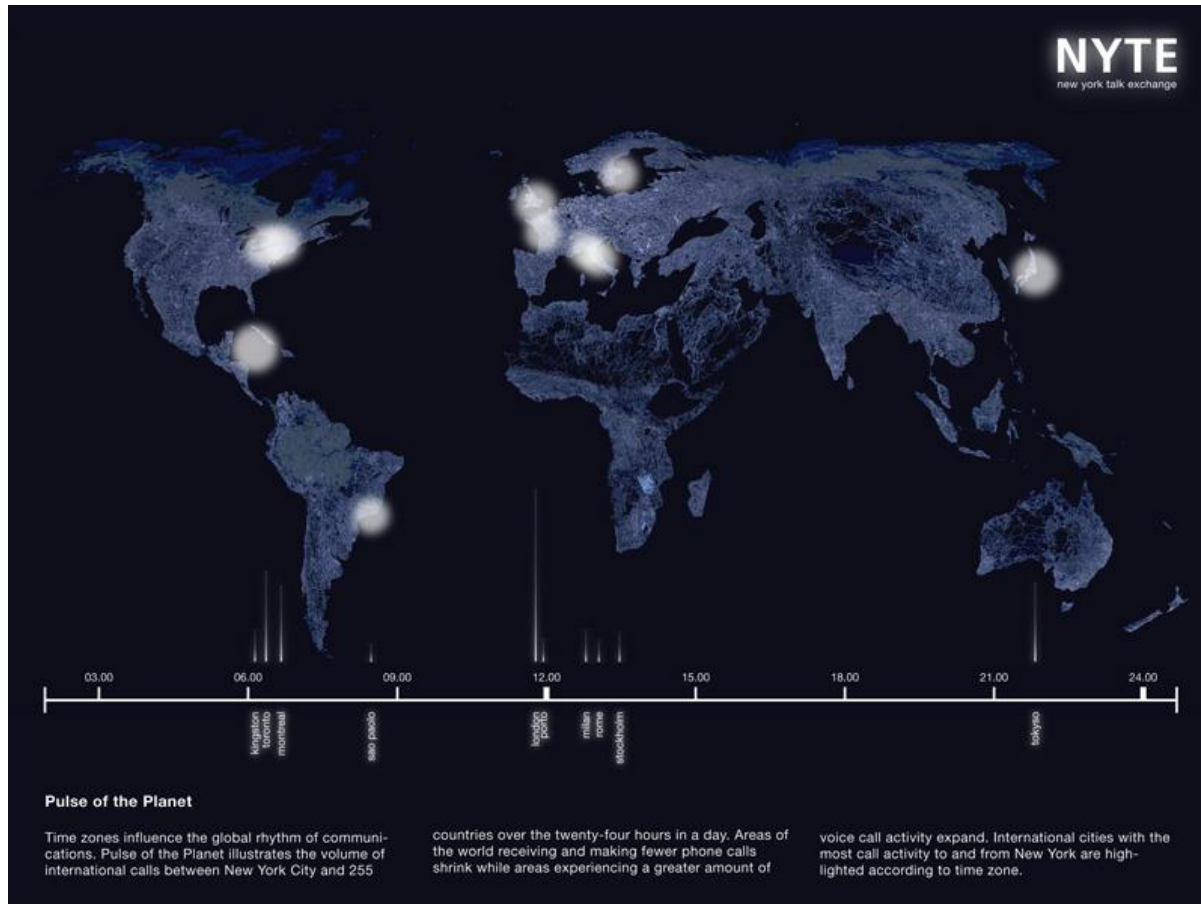
The image above illustrates the duplication within the human genome. Here, chromosomes 1, 2, 4 and 7 are shown (arbitrarily chosen).

If a region of the human genome is very similar to a region in another's genome, there is reason to suspect that these two regions both generate basic functions that are vital to both species and do not permit variation.

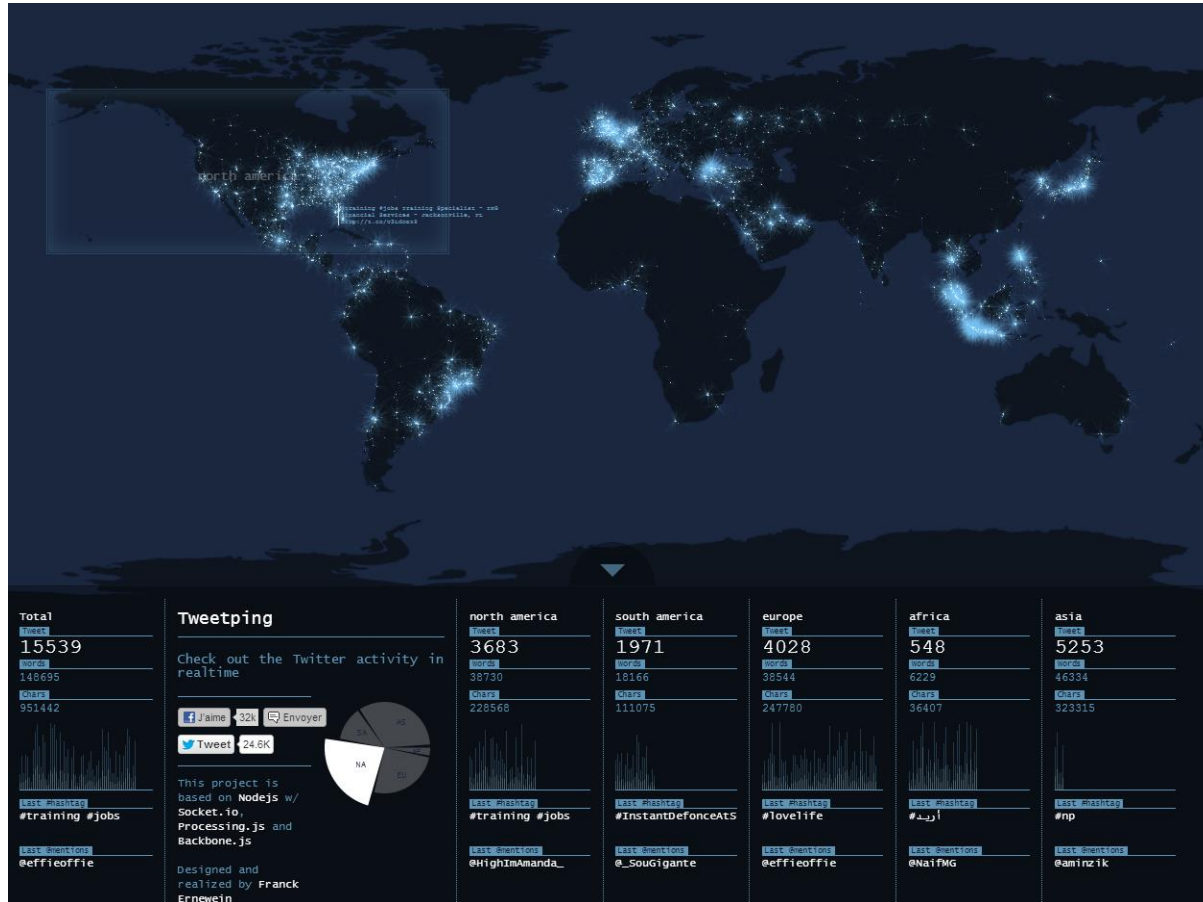
Infographics + Data Art



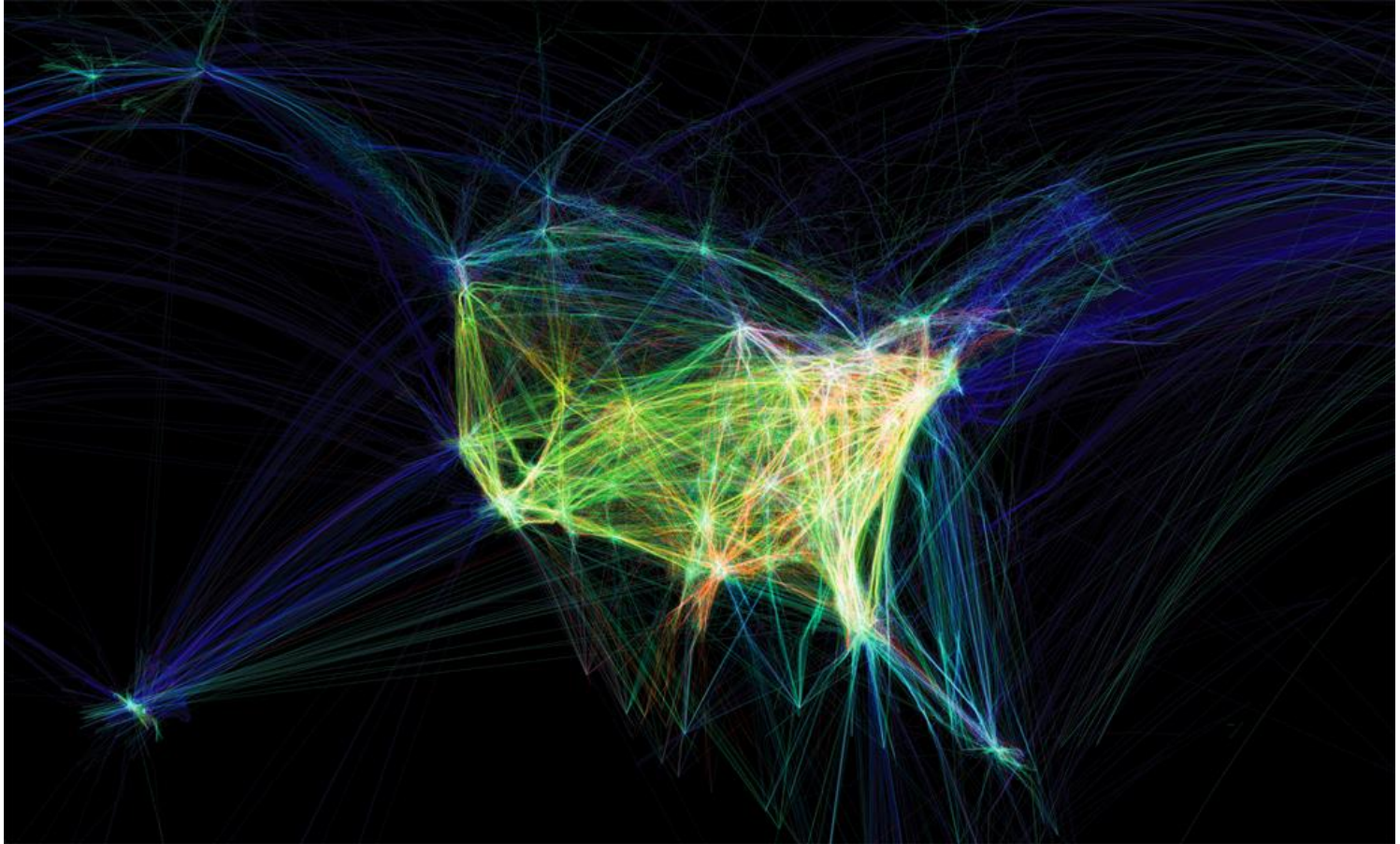
Infographics + Data Art



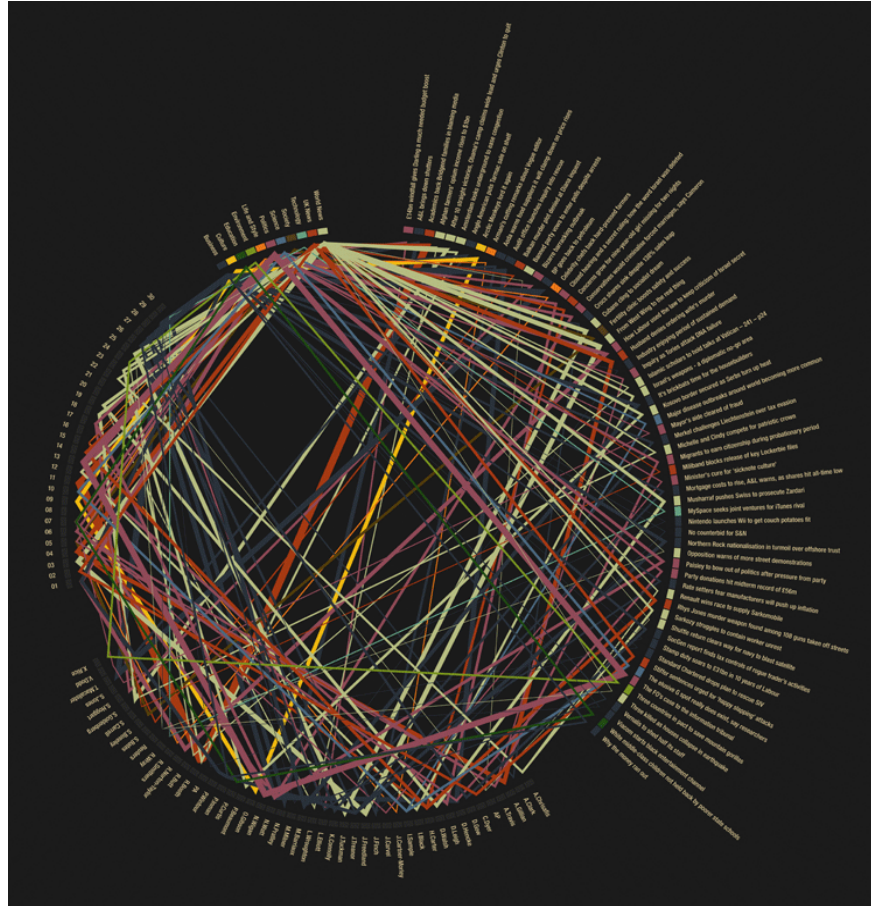
Infographics + Data Art



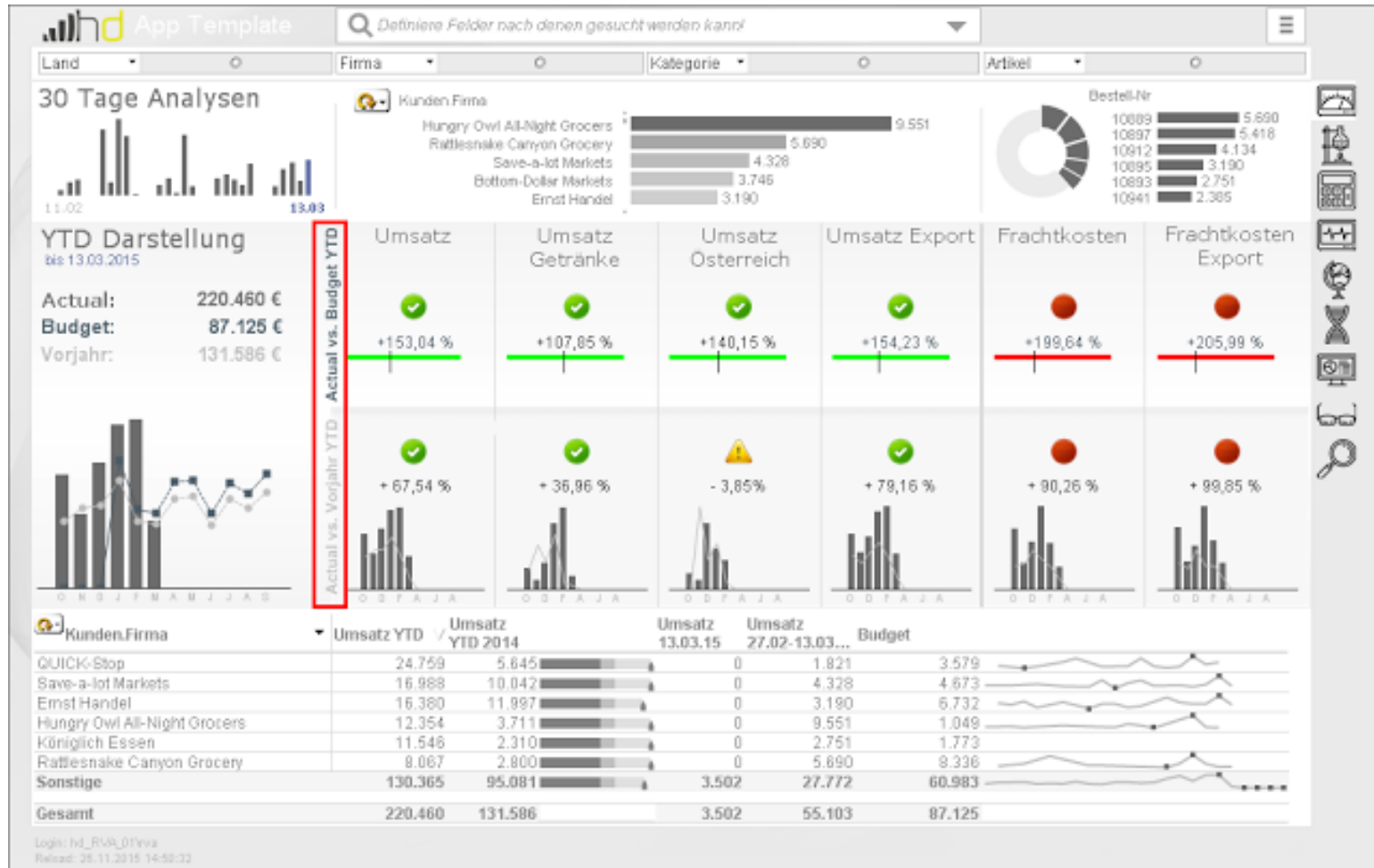
Infographics + Data Art



Infographics + Data Art

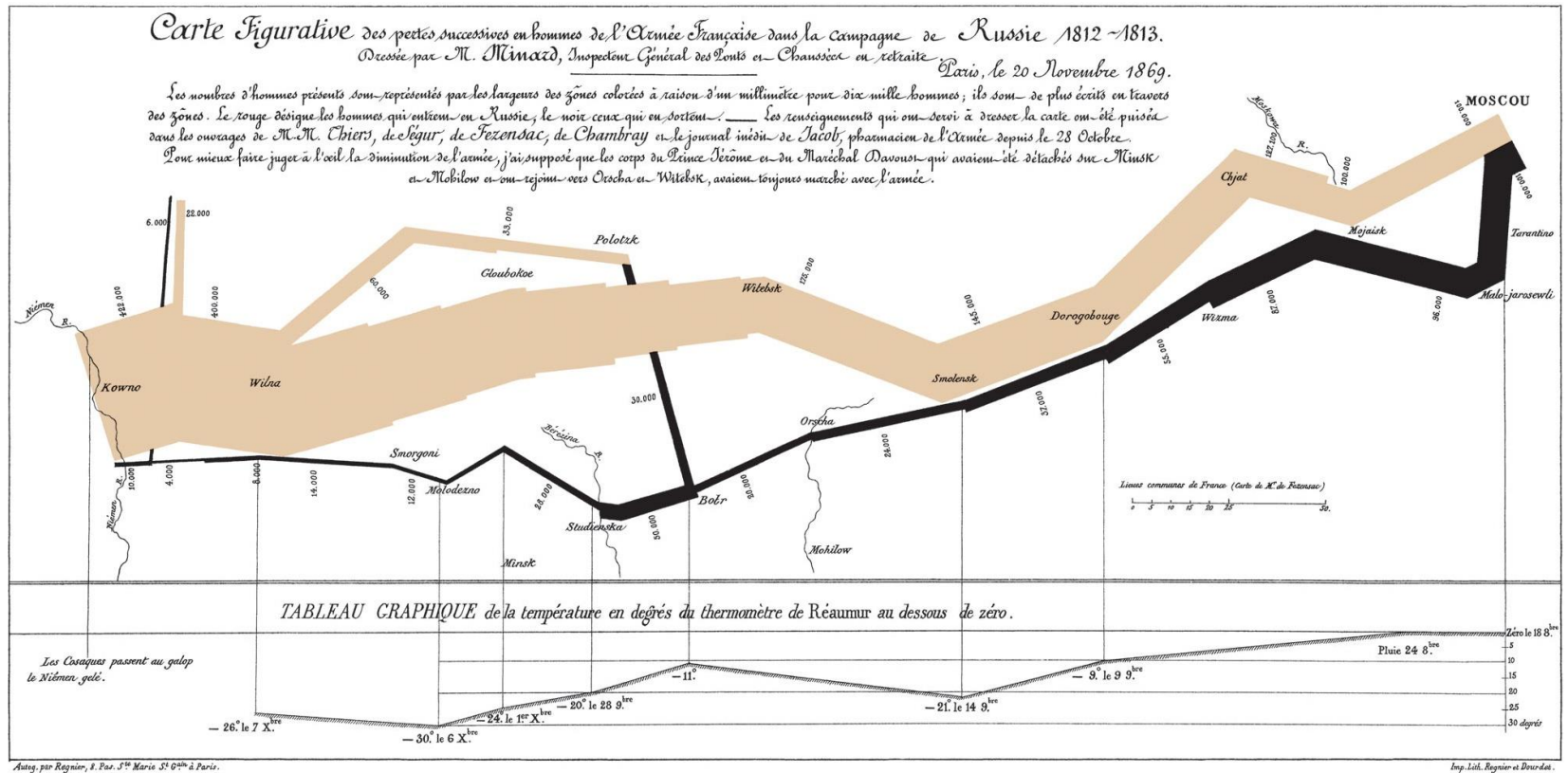


Qlikview



3. Wir stehen auf den Schultern von Giganten

Charles Joseph Minard, 1869



4. „Falsche“ Datenvisualisierungen

„Lügen mit Statistik“...



5. Richtlinien für Datenvisualisierungen



Seurat hebt die Kanten bewusst hervor

5. Richtlinien für Datenvisualisierungen

Prä-attentive Wahrnehmung



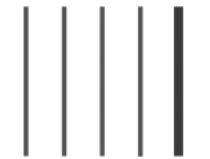
Orientation



Shape



Line length



Line width



Size



Curvature



Added marks



Enclosure



Hue



Intensity



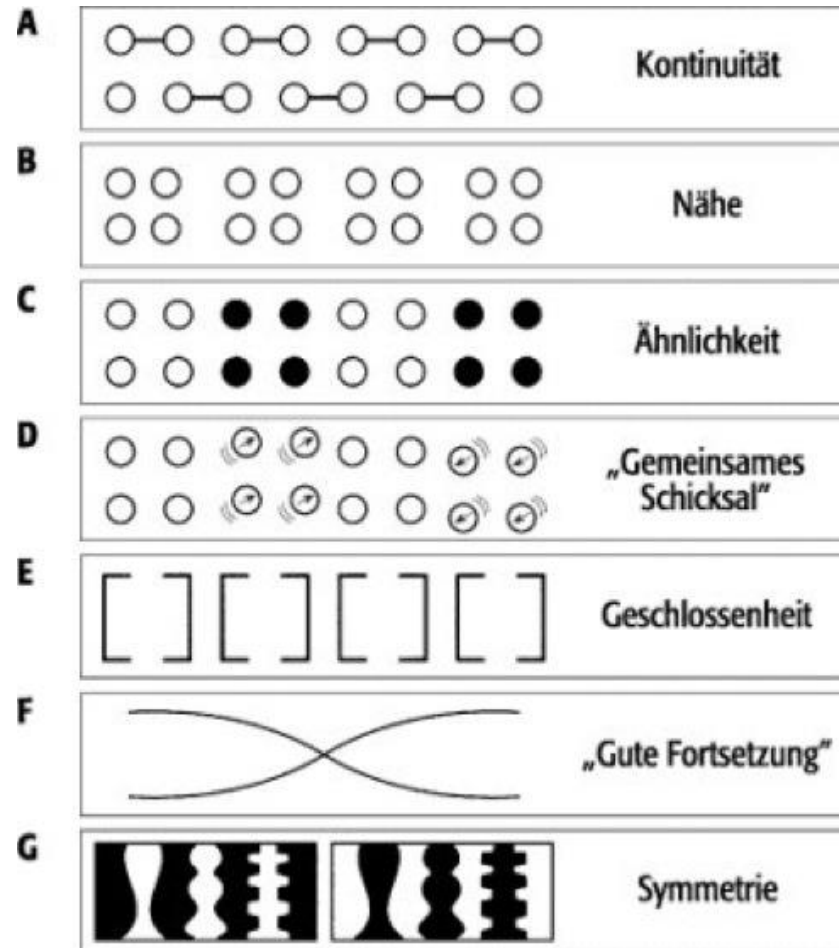
Spatial position



Motion

5. Richtlinien für Datenvisualisierungen

Weitere Gestalt - Prinzipien



+ Vorder-/
Hintergrund

5. Richtlinien für Datenvisualisierungen

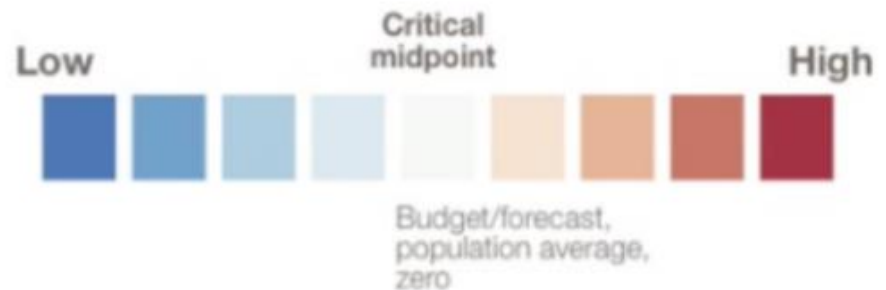
Sequential

Colors can be ordered from low to high



Diverging

Two sequential schemes extending out from a critical midpoint value

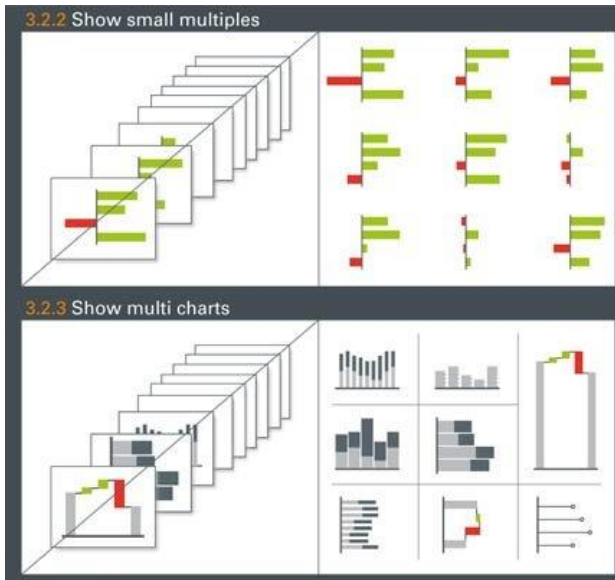


Categorical

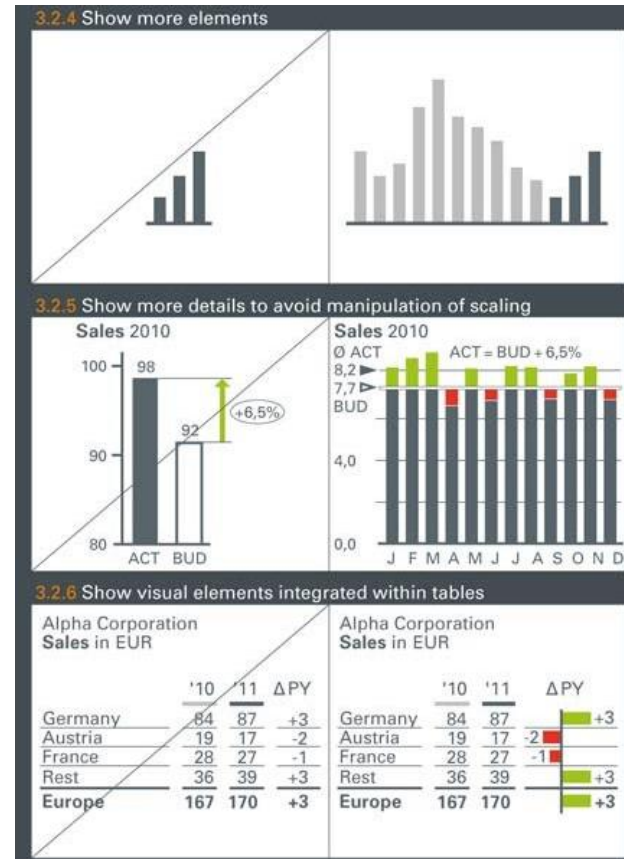
Lots of contrast between each adjacent color



5. Richtlinien für Datenvisualisierungen



<http://update.hanser-fachbuch.de/2015/04/von-atlantis-bis-excel-die-kunst-der-visualisierung/>



5. Richtlinien für Datenvisualisierungen

