

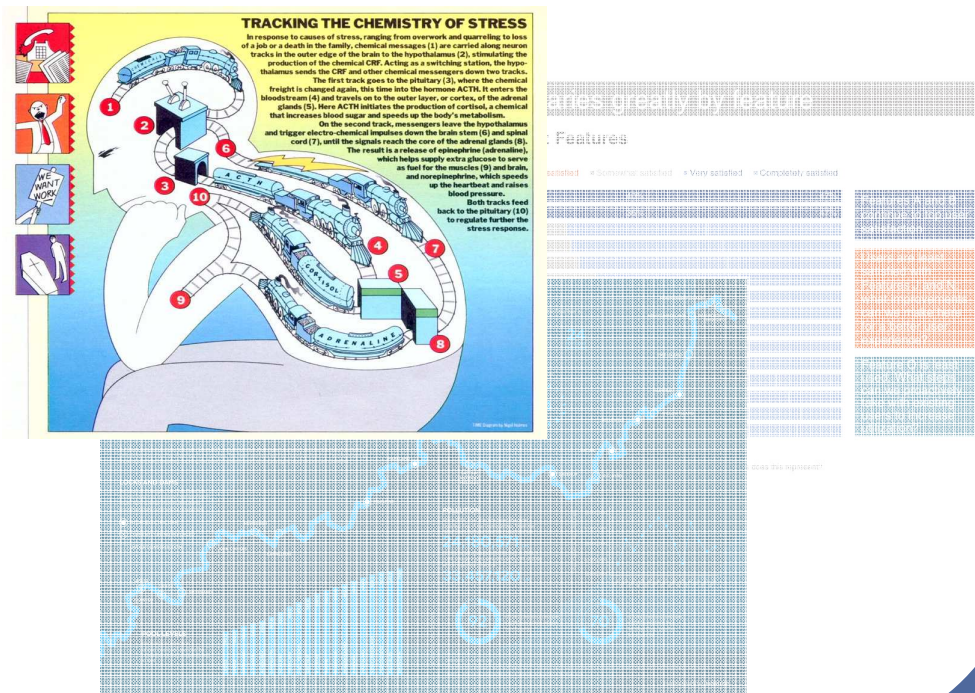
# Datenvisualisierung

- Ja. Powerpoint
- Power Corrupts. Power Point corrupts absolutely
- Nicht schnalzen oder mit finger schnippen
- Fehler nicht ansprechen
- Fragen wenn es interessiert finde ich es raus

## Gliederung

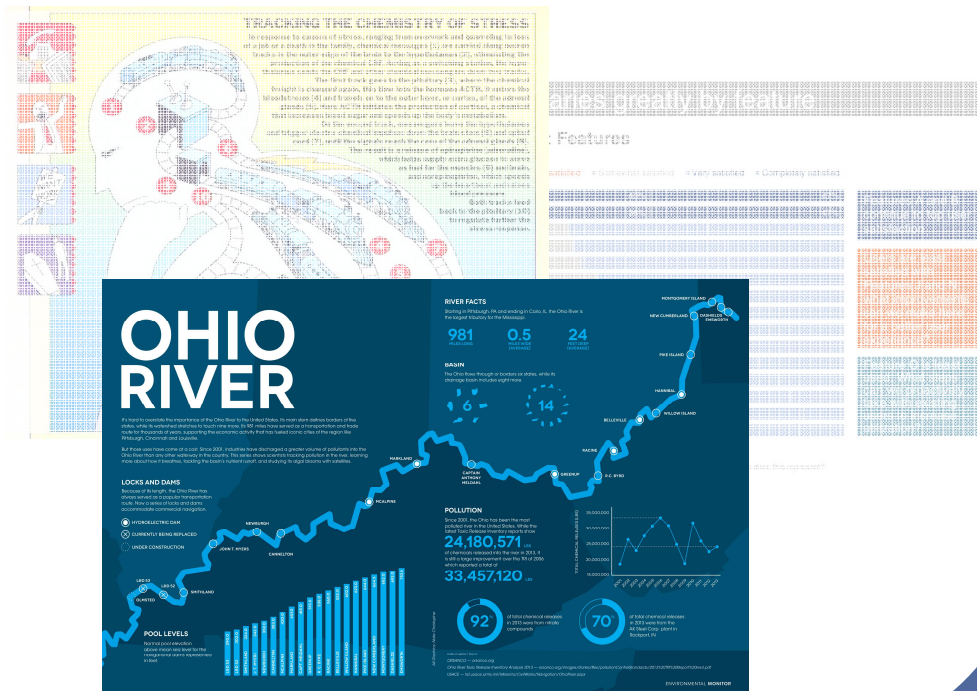
1. Was ist Datenvisualisierung?
2. Warum brauchen wir Visualisierungen?
3. Wir stehen auf den Schultern von Giganten
4. „Falsche“ Visualisierungen
5. Richtlinien für Visualisierungen
6. Software für Analyse und Visualisierungen
  - Software: Tableau vs. Excel
  - Software: Tableau vs. Power BI

# 1. Was ist Datenvisualisierung?



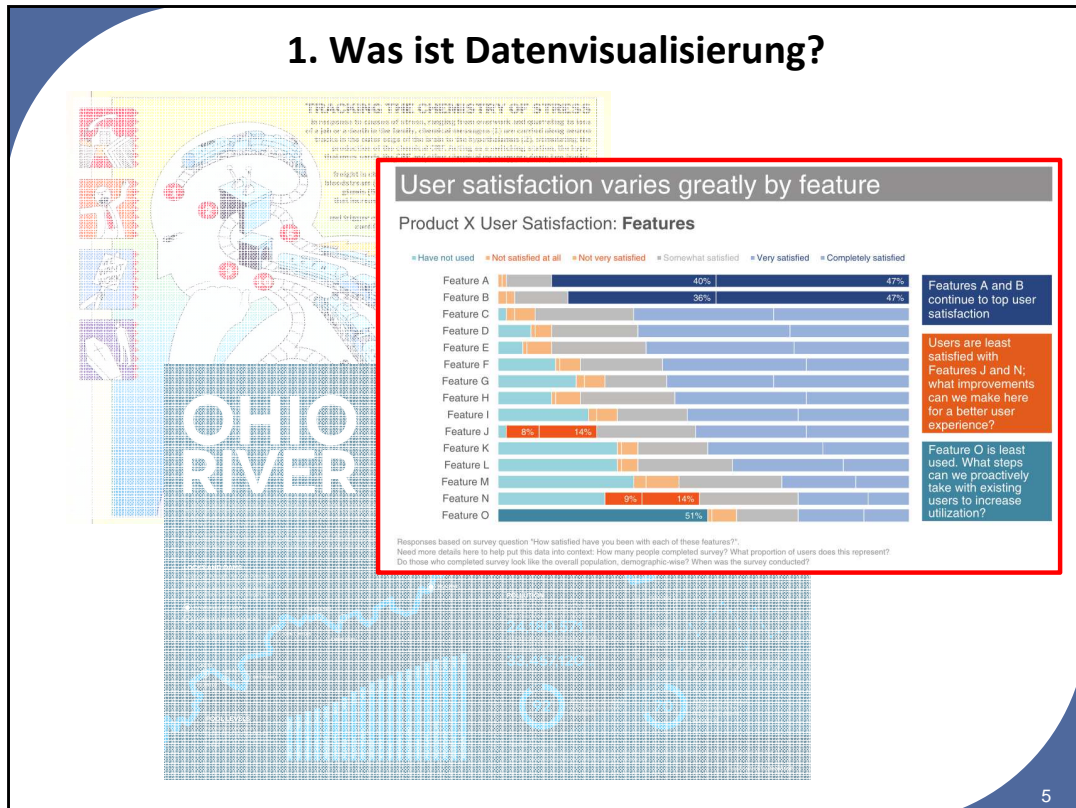
- 3 Bilder auf der Folie, damit die Angesprochenen alle im Hinterkopf haben, wenn ich umschalte und sage: Das alles sind Datenvisualisierungen
- Infografik mit viel Text: „Tracking the chemistry of stress“ - aber es gibt auch Infografiken mit weniger Text bis hin zu „Data Art“ – wie in der nächsten Folie

# 1. Was ist Datenvisualisierung?



- Infografik mit „Tendenz“ zu Data Art
- Aber keine komplette „Data Art“ – Beispiele dazu habe ich im Anhang
- Daten wunderschön präsentiert – aber nicht unbedingt so, dass man viele Informationen schnell daraus mitnehmen kann

# 1. Was ist Datenvisualisierung?



- Andere Ziele + Zielgruppe als Dashboards/ Datenauswertung für Unternehmen. Teilweise (!) auch andere Empfehlungen
- Dashboards sind interaktiv im Gegensatz zu „einfachen“ Infografiken. Nachher zeige ich ein Beispiel
- Ich konzentriere mich aber auf die Visualisierung mit den Balkendiagrammen
- Eines der wenigen Beispiele für „gute“ Gestaltung und Nutzung von gestapelten Balken
- Aus „Storytelling with Data“

## 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

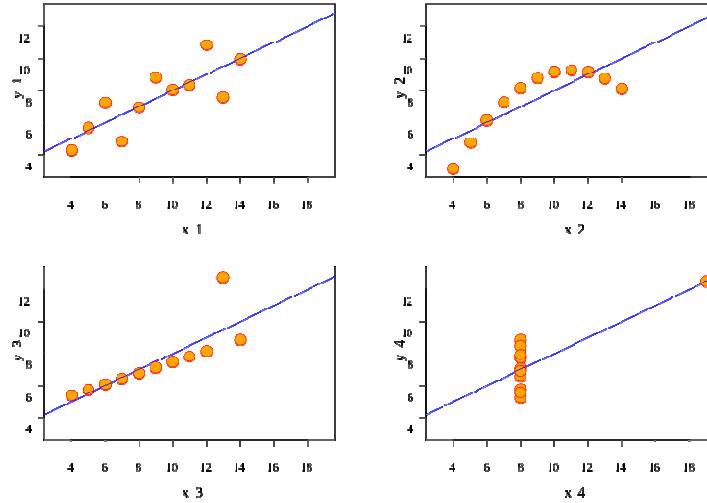
Regressionsgleichung:  $y = 3,00 + 0,500x$

6

- A picture is worth a thousand words. An interface is worth a thousand pictures. Ben Shneiderman
- Warum Datenvisualisierung? Ausspruch „Ein Bild...“
- Verarbeitungsgeschwindigkeit – und Aufwand – des Gehirns
- Einem angehenden Mathematiker – als Beispiel - 4 verschiedene Datenbündel geben, und jedes davon enthält 10 einzelne Datenpunkte
- So sind das nur „nichtssagende Daten“, aber im Kontext und im Vergleich werden daraus mglw. wertvolle Informationen
- Durchschnitt (Average), Standardabweichung, Korrelation/Gleichlauf, Regressionsgleichung/Geradengleichung

## 2. Warum brauchen wir Visualisierungen?

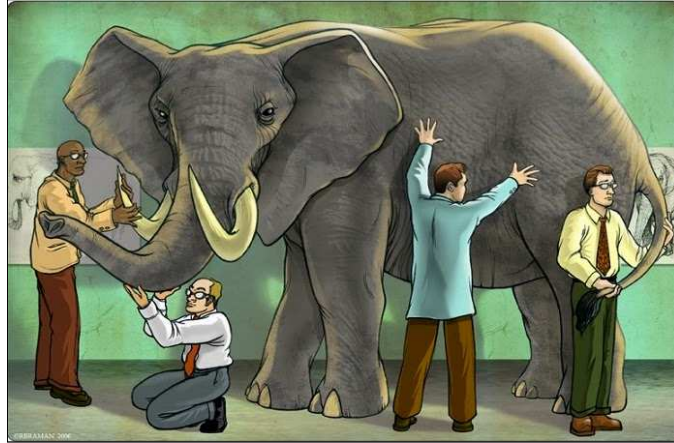
Francis Anscombes Quartett, 1973



- Seitdem weitere Beispiele veröffentlicht
- Summenstatistiken
- Martin Wattenberg (FlowingData): Wenn Du herausfinden willst, wie schlecht Deine Daten wirklich sind – versuche sie zu visualisieren
- Nächste Folie: komplexe Zusammenhänge nur anhand summierter Aussagewerte von oftmals (statistischen) Daten zu verstehen

## 2. Warum brauchen wir Visualisierungen?

„Ein Bild sagt mehr als 1000 Worte“



- Versuch 4 blinder Menschen, einen Elefanten zu beschreiben durch das, was sie direkt vor sich ertasten können
- Nicht nur einfacher, sondern auch notwendig, komplexe Zusammenhänge bildlich darzustellen
- Visualisierung ist auch oft der erste Schritt für Data-Science-Analysen
- Datenvisualisierung gibt es schon länger
- Nächste Folie: Isaac Newton Zitat verkürzt: Wir stehen...



### 3. Wir stehen auf den Schultern von Giganten

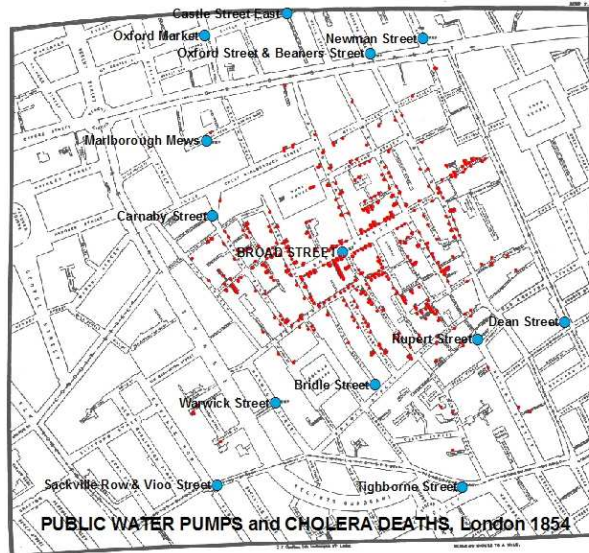


9

- „Game of Thrones“?
- John Snow

### 3. Wir stehen auf den Schultern von Giganten

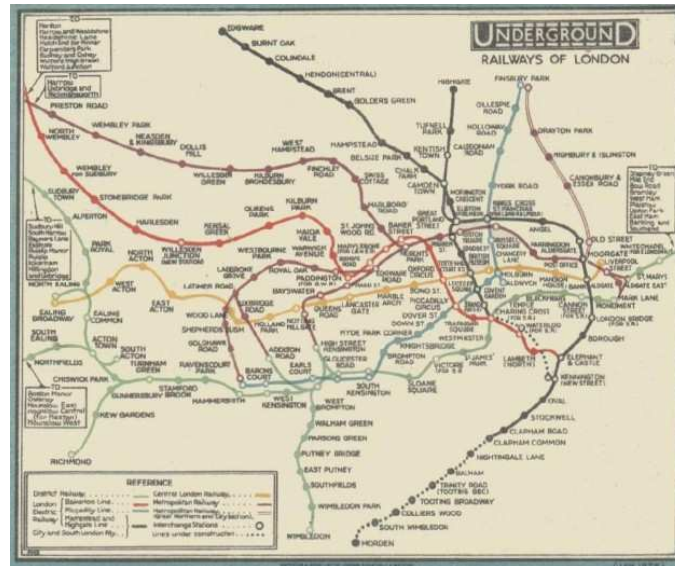
John Snow, 1854



- Arzt im vorletzten Jahrhundert
- Todesfälle auf Stadtkarte
- Wasserpumpe in Broad Street
- Geschichte geht so, dass er die Stadtverwaltung überzeugte, den dortigen Pumpenschlegel abzumontieren
- Neue Erkenntnisse durch Datenvisualisierung möglich

### 3. Wir stehen auf den Schultern von Giganten

1928

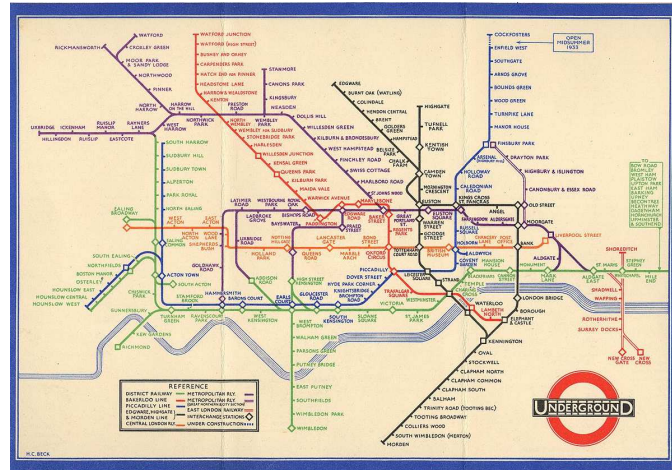


11

- Geographisch in etwa korrekt
- Ubahnnetz in London 1928

### 3. Wir stehen auf den Schultern von Giganten

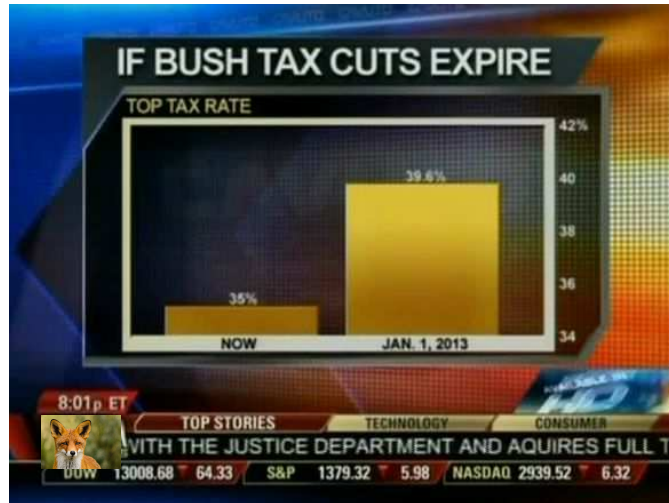
Harry Beck, 1933



- Ja, die Geographie stimmt nicht – Themse oder auch die Entfernungen zwischen Stationen bei Vorortzügen
- Verzerrung ok, da es ja nicht das Ziel ist, Entfernungsangaben korrekt darzustellen, sondern eine Orientierung geben soll, die für den Großteil der Nutzer hilfreich ist

## 4. „Falsche“ Visualisierungen

Aktive Manipulation?

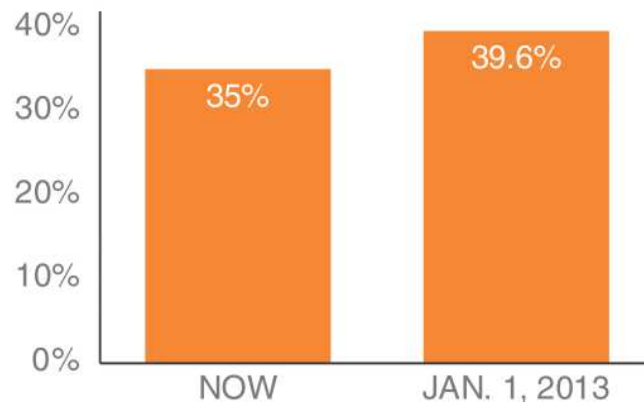


13

- Eng verwandt: Manipulation
- Hier sieht man, wie ein amerikanischer Fernsehsender – dessen Logo ich zur Wahrung der Anonymität überklebt habe – Daten präsentiert
- Aussage: Sobald die Steuererleichterungen von George Bush im Januar 2013 auslaufen, wird es allen viel schlechter gehen
- Aber: y-Achse beginnt nicht bei null, Balken „abgeschnitten“

#### 4. „Falsche“ Visualisierungen

Aktive Manipulation?  
IF BUSH TAX CUTS EXPIRE  
TOP TAX RATE

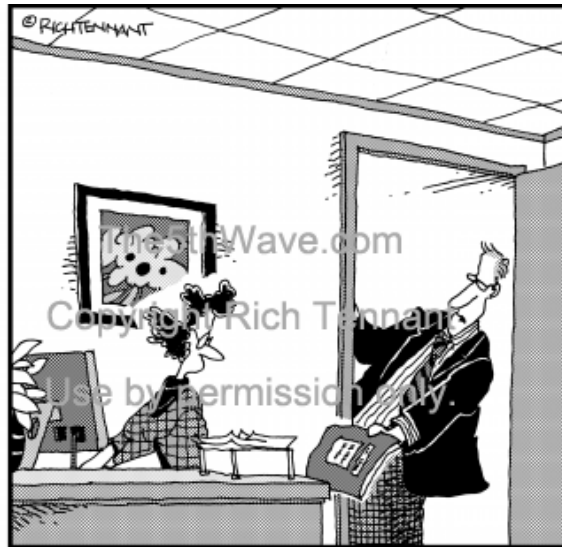


[https://en.wikipedia.org/wiki/Misleading\\_graph](https://en.wikipedia.org/wiki/Misleading_graph)

14

- Keine so gute Story – für Verkaufszahlen von Medien gilt ja oftmals „Bad news is good news“
- Link: bei Wikipedia weitere Beispiele für irreführende Graphen
- Auf manche Verzerrung oder Täuschung ist man wohl selbst schon reingefallen
- Weitere Beispiele: beliebige Zeitung oder Zeitschrift - Bild genauso wie Spiegel. Mglw. auch Cicero
- Oft ist es Anweisung oder vorauseilender Gehorsam des Mitarbeiters
- Manchmal kommt aber auch der Chef zur Tür rein und sagt: nächste Folie

#### 4. „Falsche“ Visualisierungen



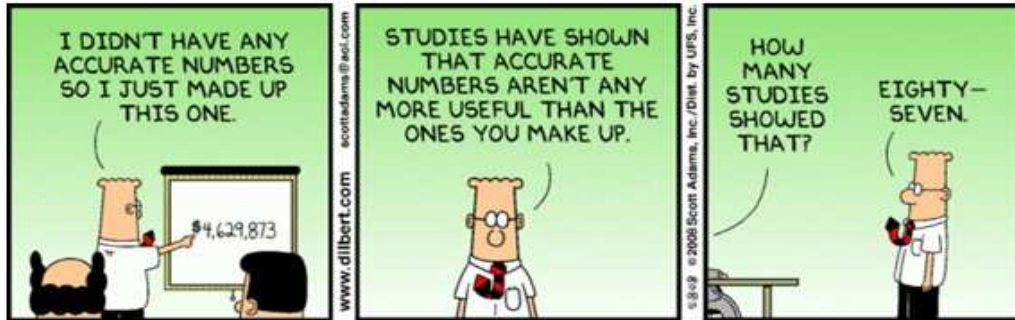
"I don't like the numbers on this report. Send it down to Marketing and see what they can do with it."

15

- Mir gefallen die Zahlen in diesem Bericht nicht. Schicken Sie ihn ans Marketing – mal schauen, was die draus machen können
- Menschen sind leicht zu täuschen
- Man kann einer Story einen beliebigen SPIN geben
- Skalen nicht nur abschneiden sondern künstlich verzerren, um z.B. das Quartalsergebnis visuell besser aussehen zu lassen
- Außer Irreführung und aktiver Täuschung gibt es aber auch offene Lügen

## 4. „Falsche“ Visualisierungen

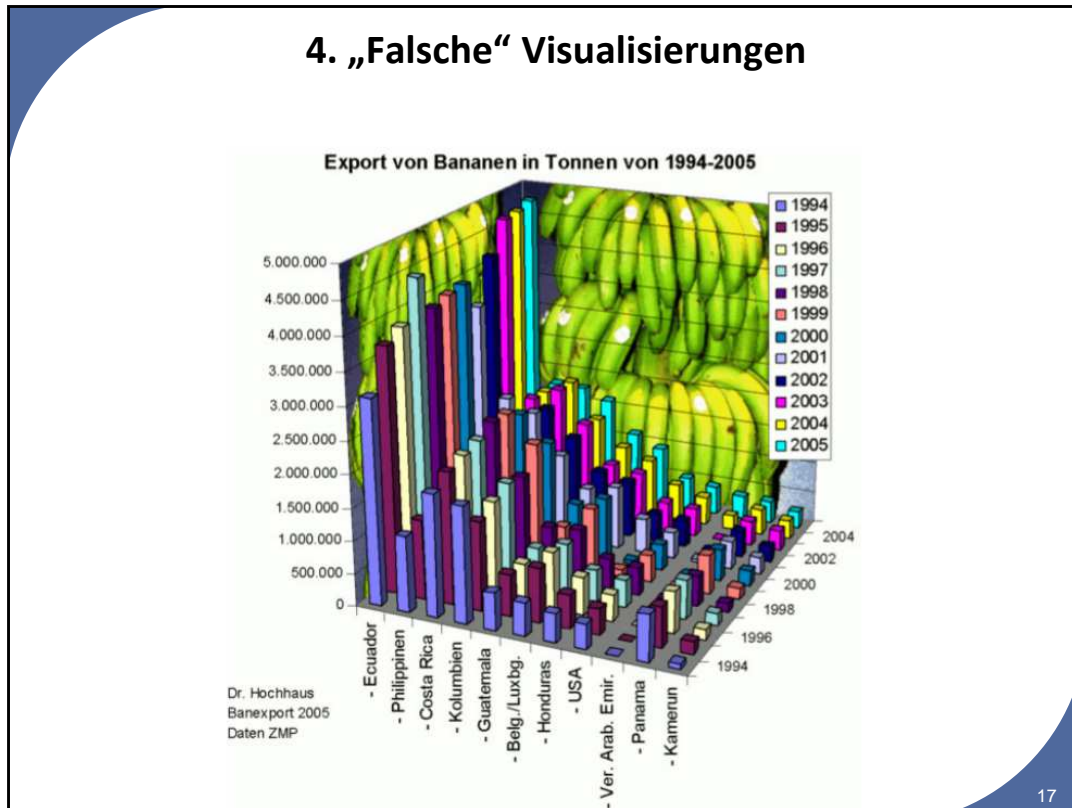
„Lügen mit Statistik“...



- Dilbert: Ich hatte keine genauen Zahlen, darum habe ich einfach diese erfunden
- Studien haben gezeigt, dass korrekte Zahlen in keiner Weise nützlicher sind als welche, die man erfunden hat
- Wie viele Studien zeigen das? 87
- Spielt mit „Lügen mit Statistik“
- Es geht natürlich – aber ja auch ohne
- Dilbert und xkcd sind Fundgruben!
- Manchmal ist es aber keine Lüge, bewusste Irreführung oder Selbst-Täuschung

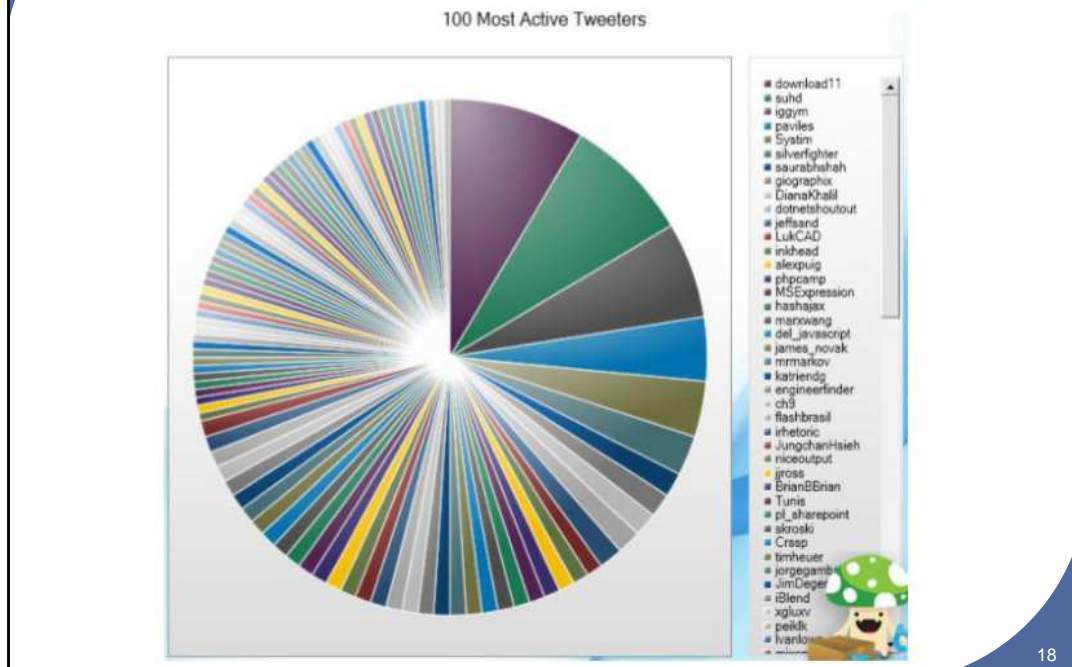


## 4. „Falsche“ Visualisierungen



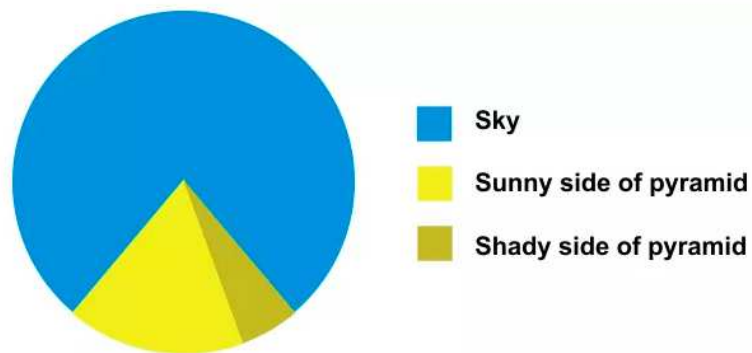
- Sondern einfach schlechtes Handwerk
- DatenVisualisierungen aus der Hölle
- Export von Bananen in Tonnen von 1994-2005
- Überlappende Daten in 3D, „schlechte“ Überschrift, Diagramm ist zu überfüllt, mit ablenkenden Bildern im Hintergrund, miserable Farbgestaltung, schlechte Skala
- Ich weiß nicht, ob es ein gezielt erstelltes Beispiel für schlechte Datenvisualisierung ist
- Man findet auch andere Graphen von Dr. Hochhaus ZMP, und bei Googlesuche sind schlechte Datenvisualisierungen häufiger als Gute
- 3D macht generell nur seeehr selten Sinn. Bei Kuchendiagrammwürde 3D Ansicht die Größe von Teilstücken verzerren – durch Perspektive
- Damit auch schon das Nächste Beispiel – ein Kuchendiagramm
- Liniendiagramm mit x Zeit und y Höhe der Exporte erstellen, je Land andersfarbige, manche Länder raus/ Gruppierungen

## 4. „Falsche“ Visualisierungen



- 100 Personen, die auf Twitter am Aktivsten sind
- Zu viele Kategorien, zu viele Farben, “größere” Kuchenstücke nicht beschriftet, schlechte Legende (scrollen)
- Fehlt nur noch, dass es verzerrt in 3D dargestellt ist

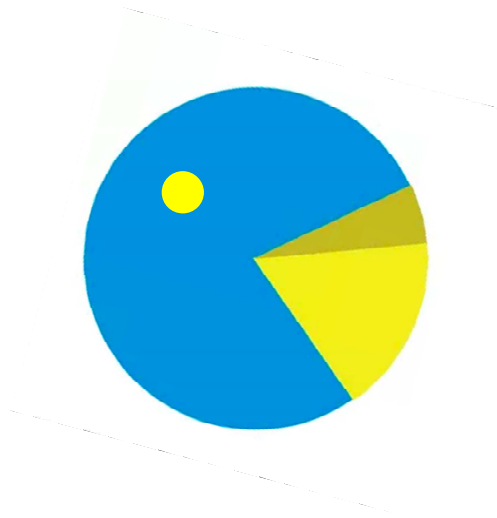
## 5. Richtlinien für Datenvisualisierungen



19

- Prima um den Himmel und eine Pyramide perspektivisch darzustellen
- Manche würden Kuchendiagramme generell nicht einsetzen
- ...andere halten sie für gute Tools, um UNTER BESTIMMTEN UMSTÄNDEN eine GERINGE ANZAHL von Kategorien gleichzeitig darzustellen wenn sich diese ZU 100% SUMMIEREN!
- MIT BEDACHT einsetzen!
- Im Gegensatz zu vielen anderen Diagrammarten sind bei Pie Charts/ Kuchendiagrammen Negative Anteile – wie z.B. manchmal Gewinn - nicht darstellbar
- Teilschitte bei 12 Uhr – oben – beginnen
- Dann nicht Pac-Man

## 5. Richtlinien für Datenvisualisierungen



20

- Dann nicht Pac-Man

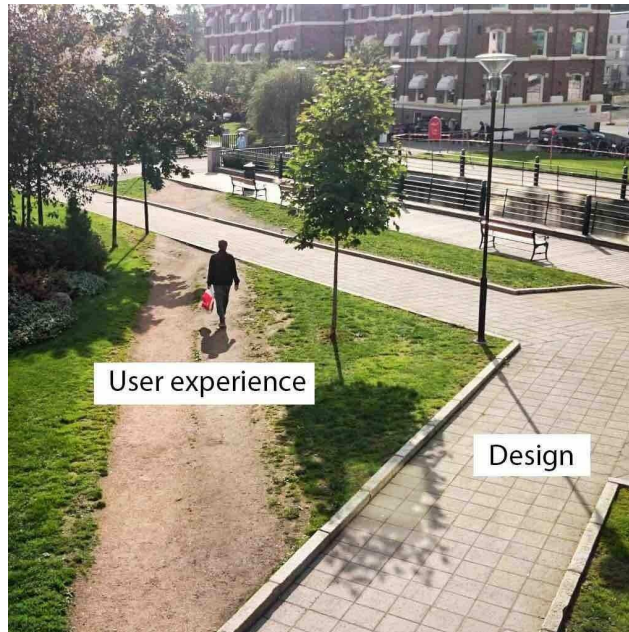
## 5. Richtlinien für Datenvisualisierungen

- 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?

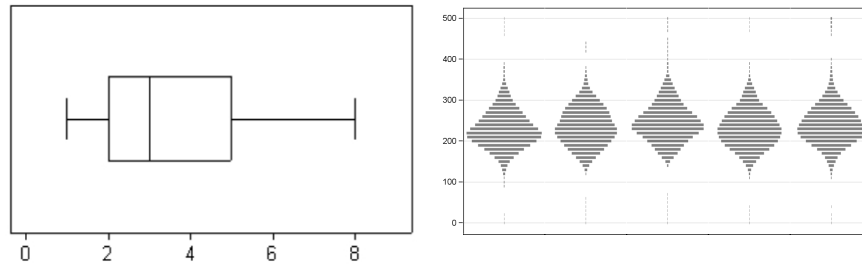
21

- Mitten drin in Richtlinien für gute Datenvisualisierungen
- 6 RICHTLINIEN, keine Regeln
- Zielen auf Dashboards ab, bei Infografiken – die ja meist statisch sind – oder bei „DataArt“ würde ich manche Dinge anders machen als bei Dashboards
- Zu den einzelnen Richtlinien
- „nur“ Daten präsentieren, Erkenntnisse daraus hervorheben und Handlungsempfehlungen geben
- oder überzeugen/bewilligen mit Daten als Hintergrundinfo z.B. Budget
- Erkundend oder Erst-Sichtbarmachen. Erklärend/ geht mehr in die Tiefe
- Zielgruppe: Ingenieure oder Manager; Thema bekannt – dann auch Tiefe möglich – oder neu, wieviel Zeit habe ich zum Erläutern, Aufnahmefähigkeit/ wann am Tag...

## 5. Richtlinien für Datenvisualisierungen - Ziel + Grundsätzlich



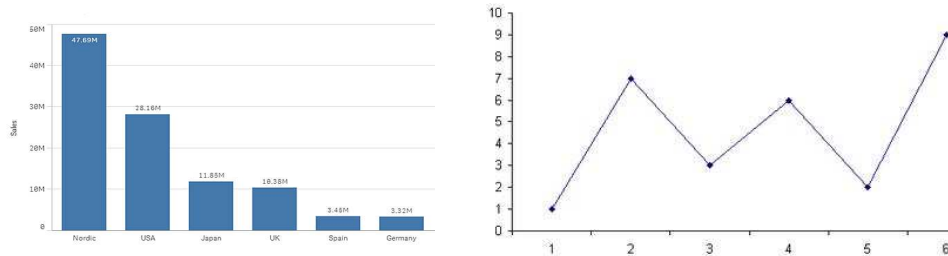
## 5. Richtlinien für Datenvisualisierungen - Zielgruppe



23

- Ingenieure oder Mathematiker
- Summarisch: Min, Max, Median, Q1+4 und damit Streubreite und damit auch Schiefe der Verteilung
- - oder sogar Violinplots, die Übersicht über Dichteverteilung geben

## 5. Richtlinien für Datenvisualisierungen - Zielgruppe

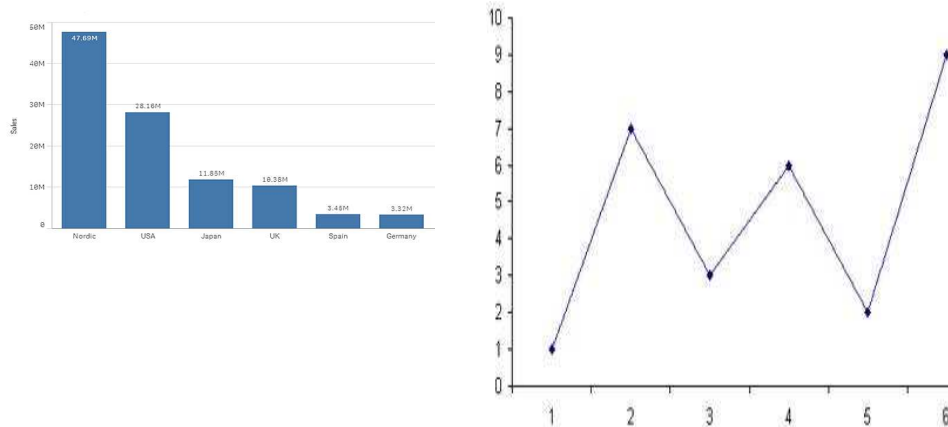


24

- Für Verkäufer oder nicht-technisch fokussierte Manager
- Linien oder Balken, da besser in kurzer Zeit verständlich und
- Unnützes Wissen, wie bei den Boxplots – rausgelassen wird
- Manchmal ist es besser Balken horizontal anzuordnen
- Manchmal auch Skala oben
- F bzw. Z-förmigen Lesebewegungen
- Skalen nicht verzerrend
- Sonst sieht es z.B....



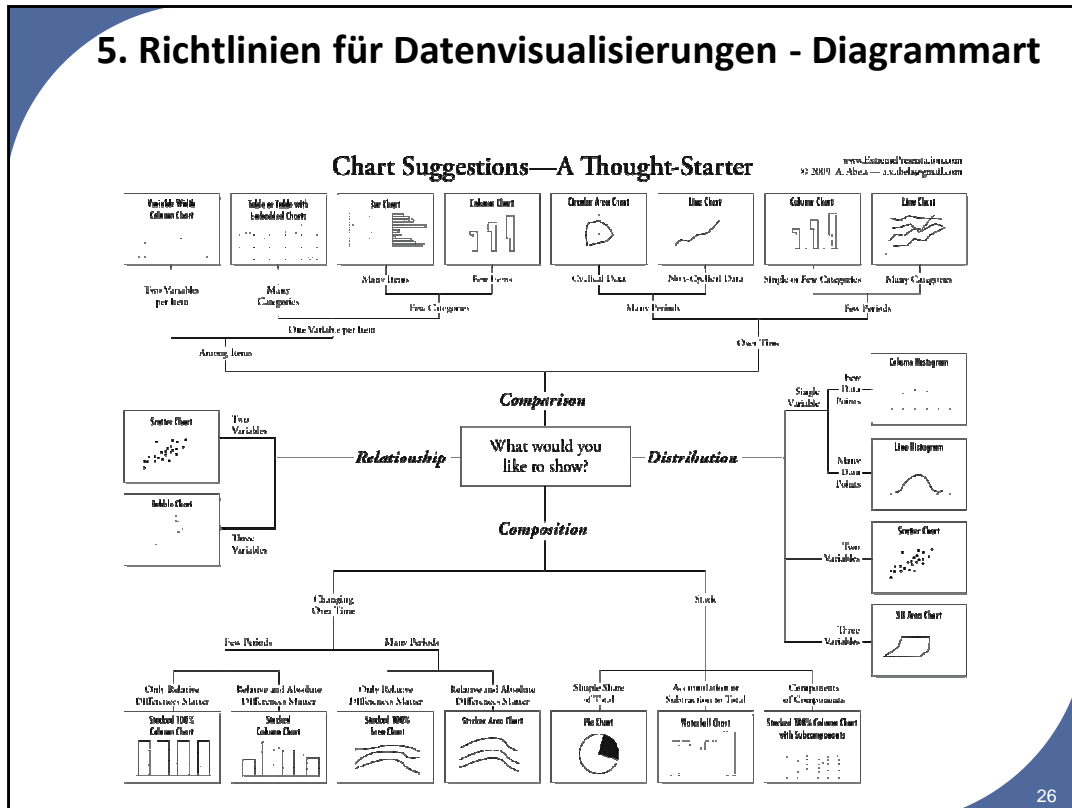
## 5. Richtlinien für Datenvisualisierungen - Zielgruppe



25

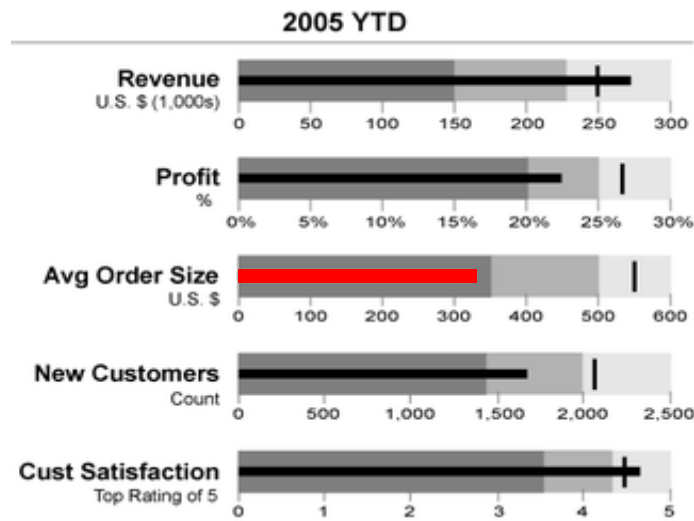
- Liniendiagramm: Zukunftsprojektionen/Trends nicht genau eintragen, sondern mit Unsicherheitsbereich
- Und z.B. Vorsicht bei Logarithmischen Skalen
- Für jede Darstellung gibt es Diagrammarten, die besser oder aber nicht so gut geeignet sind

## 5. Richtlinien für Datenvisualisierungen - Diagrammart



- Hier eine Übersicht - Nicht alles durchlesen, Folien schicke ich gerne zu
- Werte vergleichen – zwischen verschiedenen Items einmal vs. während eines Zeitraumes
- Verteilungen einzelnen Datenpunkte darstellen
- Zusammensetzung von Datenmenge darstellen – einmalig vs. Im Zeitverlauf
- Verhältnisse erfassen
- Gestapelte Balken: Am Anfang bei „Was ist Datenvisualisierung“. Eines der WENIGEN guten Beispiele. Selten sinnvoll
- Schattenwurf – einfach: NEIN
- Bei Infografik, Data Art mglw. manche Diagrammarten auswählen – aber BEWUSST wählen und Alternativen bedenken
- manches anders ... siehe nächste Folie

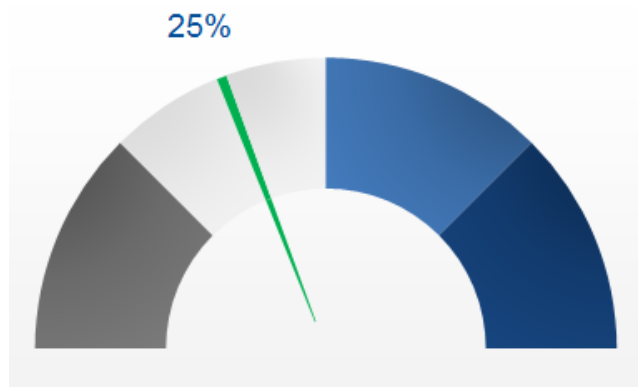
## 5. Richtlinien für Datenvisualisierungen - Diagrammart



27

- Bullet Graphs va. Für Dashboards
- Nächstes: Bullecharts Andreas Theos Stephen Few
- Informationsdichte
- Schwarze Balken zeigt Wert auf Skala
- Skala ist beschriftet
- Gedeckte Hintergrundfarben gut mittel schlecht
- Hervorhebungen möglich
- Vergleichswert durch Querstrich
- Data-to-Ink-Ratio/ Informationsdichte
- Chartjunk – weniger ist mehr
- „De-Cluttering“, nicht überladen, nicht „unnützen“ Data Junk
- De-Cluttering wird manchmal fälschlicherweise benutzt für „wir brauchen keine Skalenangaben“
- Das Nicht, aber Skalenlinien unaufdringlich gestalten, Leerflächen als Begrenzung nehmen statt Linien

## 5. Richtlinien für Datenvisualisierungen - Diagrammart



28

- Genauso wie Kuchendiagramme/ Pie charts
- Sind Tachometer oder Messuhren oder Druckmeter verpönt
- ...u.a. wegen der Platzverschwendung, geringen Informationsdichte/ Data-to-ink-ratio
- Aber manchmal → Ziel + Zielgruppe
- hier eines der noch besseren Beispiele – klare Aussage, gedämpfte Farben
- Wenn es um Informationsdichte geht

## 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

29

- Bei Farben: nicht übertreiben:
- 5 meistverkauften Medikamente A-E werden verglichen in verschiedenen Ländern
- Der Verkaufsrank wird in verschiedenen Ländern anhand der Farben und Nummern dargestellt
- Eingefasst wirkt es durch deutliche Zellumrandungen

## 5. Richtlinien für Datenvisualisierungen - Farbe

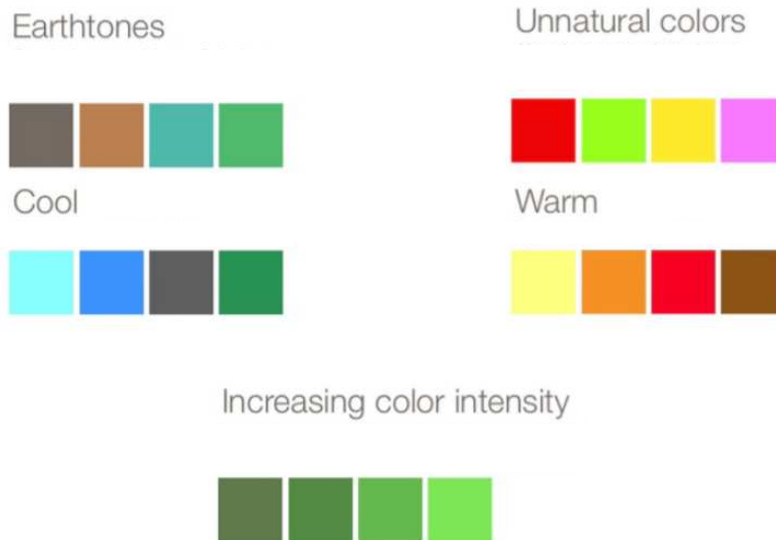
Top 5 drugs: country-level sales rank

COUNTRY   DRUG	RANK				
	1	2	3	4	5+
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

30

- Viel übersichtlicher, viel klarer und in kürzerer Zeit
- Ja, gewisser Informationsverlust, da nur die ersten 5 Kategorien – aber „welche sind wichtig“
- Welche Medikamente überall ankommen, welche in einzelnen Ländern gut ankommen
- Man könnte noch die Länder „besser“ organisieren – Umsatz statt A-Z
- Augenfreundlicher weniger gedrängt, da Umrandung fehlt und Zelumrandungen in weiß – „De-Cluttering“

## 5. Richtlinien für Datenvisualisierungen - Farbe

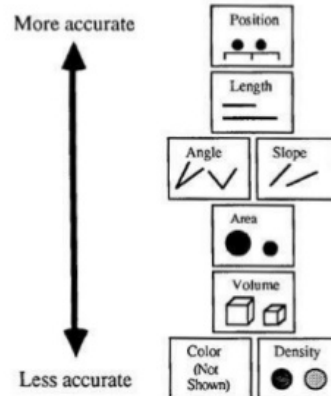


31

- Einerseits wirken für seriöse Darstellungen gedecktere Farben besser
- Aber es kommt auch auf das Ziel an: will ich hervorheben/alarmieren – oder ist es aus Gründen des „Corporate Design“ besser?
- Mein Rot ist nicht Dein Rot
- Aber auch: 10% der Bevölkerung sind auf eine oder andere Weise farbenblind bzw. haben Farbschwächen
- Darum ist rot-grün nicht gut, rot-blau oder einfach unterschiedliche Tönungen einer Farbe sind besser
- Andererseits auch kulturspezifisch: westlich Rot Warnung ggü. Chinas Glück
- [http://www.huffingtonpost.com/smartertravel/what-colors-mean-in-other\\_b\\_9078674.html](http://www.huffingtonpost.com/smartertravel/what-colors-mean-in-other_b_9078674.html)
- <http://enably.com/chromatic/> - AdobeAIR-Tool
- <http://colorfilter.wickline.org/> - Webpages
- <http://www.etre.com/tools/colourblindsimulator/> - Bilder hochladen
- Visual Miscellaneum: A colorful guide to the world's most consequential trivia 2012, Stephen Few

## 5. Richtlinien für Datenvisualisierungen – Häppchen

### Prä-attentive Wahrnehmung



Mackinlay, APT (A Presentation Tool), 1986

Slide: Cecilia Aragon, HCDE, UW

32

- Gehirngerechte Häppchen
- Gehirn sucht ständig nach Mustern + kann Ausreißer gut wahrnehmen
- durch Kurzzeiteinblendung der Bilder und Untersuchung im Magnetomographen
- Scheint nicht der Wahrheit zu entsprechen – aber „Prä-attentive Wahrnehmung hat sich durchgesetzt
- Bestimmte Musteränderungen kann das Gehirn schneller wahrnehmen – und auch dort gibt es eine Reihenfolge in der Wahrnehmungsgenauigkeit
- Feine Unterschiede: Position, Linienlänge
- Danach Winkel oder Steigung
- Schlechter: Größe eines Areals oder gar Volumen – 3D verzerrt
- Und am Schlechtesten sind feine Unterschiede in Farbstärke, Farbhelligkeit oder Farbton wahrnehmbar und „Datendichte“ ob es - 5 oder 10 Datenpunkte sind, können wir noch abschätzen, ob es 50 oder 100 sind, nicht mehr
- Trotzdem sind Daten eine gute Möglichkeit zur Hervorhebung – z.B. Warnfarben
- Colin Ware
- Farbverläufe auf Skalen: gut zu unterscheiden – aber wir können z.B. nicht sagen ob dieses rot für 0,5 Mio steht oder dieses, etwas auf der Folie entferntere



## 5. Richtlinien für Datenvisualisierungen – Häppchen

Gestalt - Prinzipien

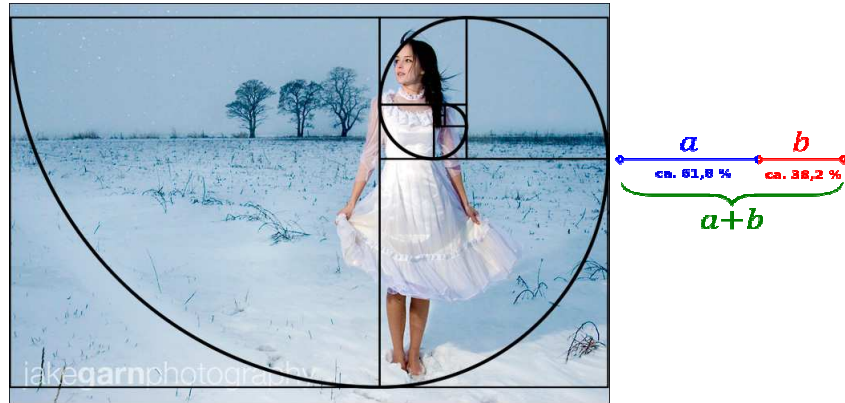


33

- Wahrnehmung
- Ok, Logo statt Datenmenge – aber brilliant!
- In dem Fall nicht wegen Tierliebe ausgewählt
- Umrandung/Enclosure
- Andere Gestalt-Prinzipien Symmetrie, Vorder-vs. Hintergrund
- Siehe z.B. Kippbilder

## 5. Richtlinien für Datenvisualisierungen - Design

### Ästhetik



34

- Schließlich noch: Was Ästhetik ausmacht, dafür kann ich keine Definition geben
- Fibonacci – Spirale und goldener Schnitt
- Vitruv, Parthenon
- <https://meiert.com/de/publications/translations/washington.edu/tufte/>
- <http://www.cs.umd.edu/~ben/goldenrules.html>
- Leonardo Da Vinci combined art and science and aesthetics and engineering, that kind of unity is needed once again. Ben Shneiderman
- Dieter Rams, Braun Elektrogeräte
- Anknüpfungspunkte und Überschneidungen zu Design, UI
- Firmitas (Festigkeit), Utilitas (Nützlichkeit, Usability) und Venustas (Schönheit). Vitruv
- Dino Citraro, one of the creators, told me that a simple chart "might be the best way of conveying the information, but it isn't effective at generating empathy or a deeper understanding of consequence." A data visualization that resonates, he argues, works on the level of metaphor: "We map the data to a metaphor and use interactivity to reveal information at a self-directed pace."

## 6. Software für Analyse und Visualisierungen

Jan. 2017: BI + Analytics Leader nach Gartner

- Tableau
- Microsoft (PowerBI)
- Qlik



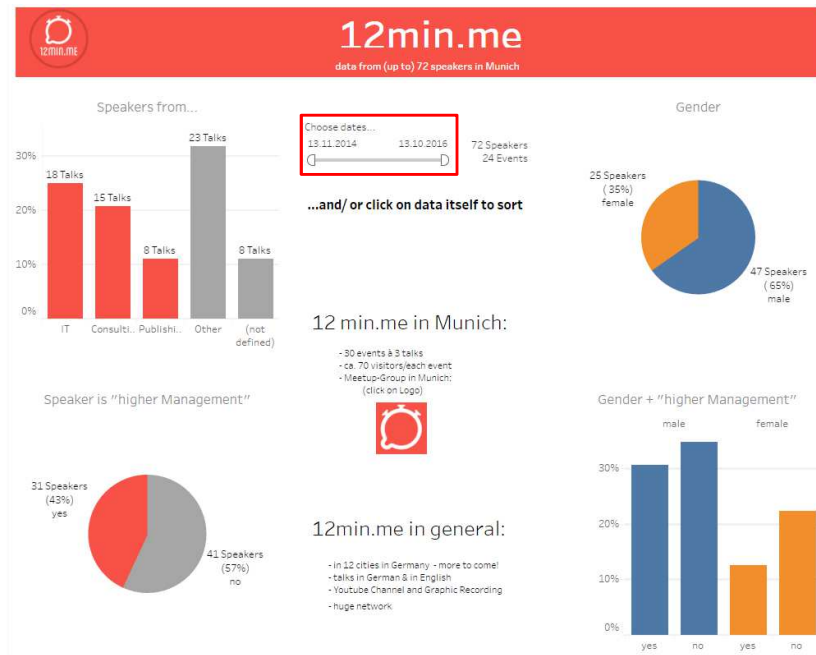
Wohlgemerkt:

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

35

- Bewusst so klein
- Qlik ist richtig geschrieben. View + QlikSense – ich werde nicht näher drauf eingehen
- Gartner Quadrant: oben rechts Leaders + ability to execute
- <https://www.gartner.com/doc/reprints?ct=160204&id=1-2XXET8P>
- Freakalytics nach Affordability Individuell auch diese drei recht gut
- RapidMiner, Knime, Alteryx oder Spotfire
- Easel.ly, Piktochart, Infogr.am
- Ggplot2 oder Shiny in R, Plot.ly in Python oder Julia, D3 für Java. GnuPlot oder Gephi
- Für Bedienung von Tableau gibt es gute Videotutorials
- Weitere: Cluvio, Domo, Periscope Data, Sisense...

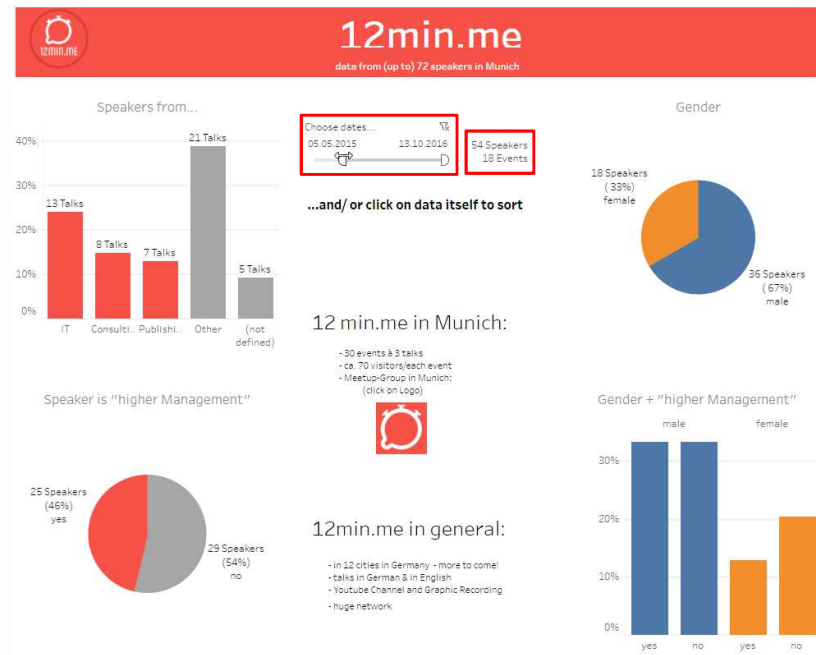
## 6. Software für Analyse und Visualisierungen



36

- Hier ein Ergebnis
- Zwischen Infografik und Dashboard – Fingerübung
- Corporate Design: Farbe, Logo links oben, Platzverschwendung durch Balken
- 12min.me ~ TED auf deutsch binnen 12 min
- Meetup
- Datenquellen: hier ein Excel-Sheet, aber auch DB – unternehmensintern wie auch extern (Teradata, Salesforce), Datenmarktplätze, Webseiten (z.B. Destatis, Gapminder, UN, WHO...)
- Verschiedene Diagramm-Seiten werden auf Dashboards zusammengefasst, angeordnet, formatiert – generell, aber auch für verschiedene Bildschirmgrößen
- Interaktiv: Datenpunkte und Zeitraum auswählen, Verlinkungen

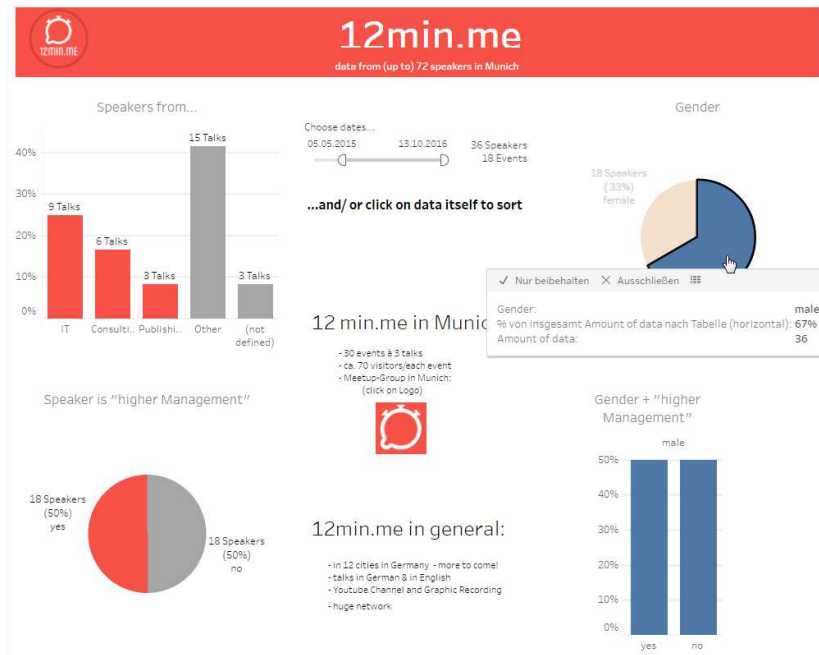
## 6. Software für Analyse und Visualisierungen



37

- Interaktivität: Zeitraum

## 6. Software für Analyse und Visualisierungen



38

- Auch möglich mit Tableau:
- ...Geschichten erzählen - es gibt auch „Story-Modus“
- ...auch Animationen sind möglich
- Da es ja ein Excelstammtisch ist...

## Software: Tableau vs. Excel

Tableau:		Excel:	
-	Kosten	0	
+	Gruppenzusammenarbeit	0	
+	gute Designvorlagen	-	
+	für Big Data geeignet	-	
-	Verbreitung	+	
0	Lernkurve	0	
~	„kann alles“ + Integration	+	<a href="http://www.cafepress.com/+excel+mugs">http://www.cafepress.com/+excel+mugs</a>
-	Einzelne Punkte bearbeiten	+	
-	„Bequeme“ Funktionen	+	
+	Community	+	
+	Dashboards, Storytelling	-	

39

- Zwar kostenlose Version, aber mit Veröffentlichungszwang
- Einzelne Punkte: anderer Ansatz
- Makros, Skriptinsprache
- Community: NICHT Stammtisch, sondern Events, Viz, Austausch, Erreichbarkeit...
- Aber: besser/ gerechter mit Power BI zu vergleichen

## 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	?

40

- Herkunft Power BI: Power Pivot + Power Query + Power Map + Excel mit manchen Funktionen
- Im Grunde genommen: eine Frage des Geschmacks + Coin Toss



# 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](mailto:ms@grauschattierung.de)

42

- Dilbert und xkcd sind Fundgruben!

## 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](#)

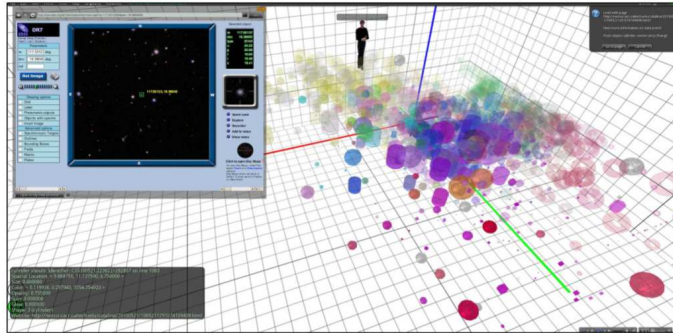
### **Bücher:**

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

43

- Tufte: The Visual Display of Quantitative Information
- Tukey: The future of data analysis

## Wo geht die Reise hin?



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

44

- Originlabs

## Nicht besprochen

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

45

- <http://www.creativebloq.com/design-tools/data-visualization-712402>

## Dashboards...



- Dilbert und xkcd sind Fundgruben!

# Infographics + Data Art

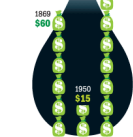
## GAS GUZZLING

Datagraphic by Nigel Holmes  
Research by Sarah Richardson

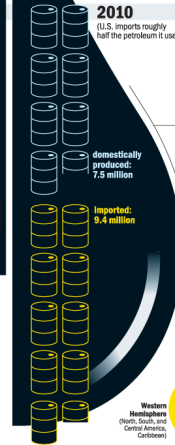
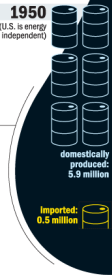
### BARRELS USED IN THE U.S. EVERY DAY

**52%**  
used for transportation

### CHANGE IN THE PRICE OF A BARREL OF CRUDE OIL (in 2010 dollars)



Sources: EIA.gov; Transportation Energy Data Book Edition 30-2011; www.doh.gov.uk



### WHAT REFINERS MAKE FROM A BARREL OF CRUDE OIL

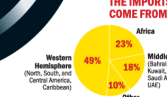


gasoline: 19 gallons  
diesel: 10 gallons  
jet fuel: 6 gallons  
other: 12 gallons (includes heating oil and lighter petroleum gases)

**70%**  
used for transportation



### WHERE THE IMPORTS COME FROM



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

- Chartjunk? Vs. Ziel+Zielgruppe
- Stephen Few, Edward Tufte
- Software Tools

# 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 chimps.

DAVID CONSTANTINE

### COMPARING CHROMOSOMES 1

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

DNA  
BASE  
PAIR

Species  
Rhesus Monkey  
Chimp  
Human

Regions of highest similarity tend to bundle together.

Gaps represent areas that either haven't been sequenced or can't be sequenced.

Lines join the 200 regions on each chromosome that are most similar to a human's (based on number of matching base pairs).

Thicker lines represent more similarity.

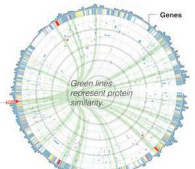
Chimp

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

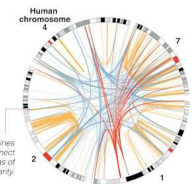
**Line charts** show what percentage of the human chromosome is similar to each of the other five genomes.

### 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).

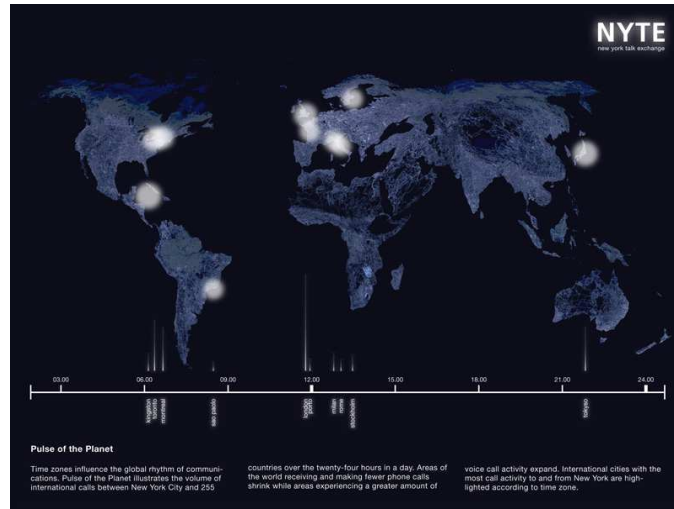
If a region of the human genome is very similar to a region in another's genome, there is reason to suspect that these two regions both generate basic functions that are vital to both species and do not permit variation.



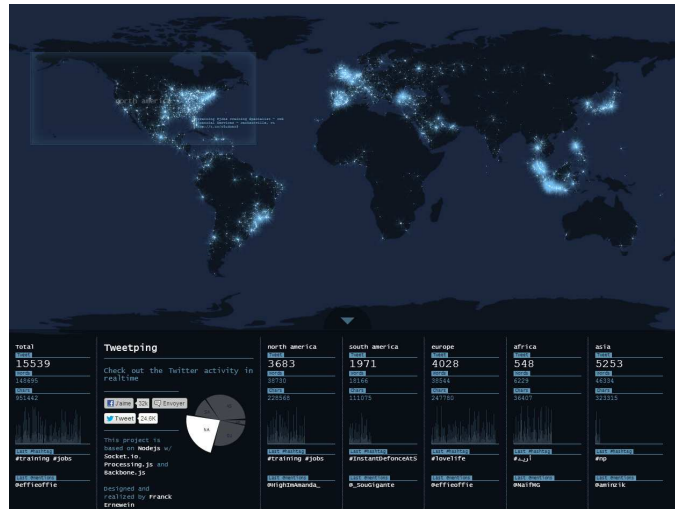
[illegible]

- 49

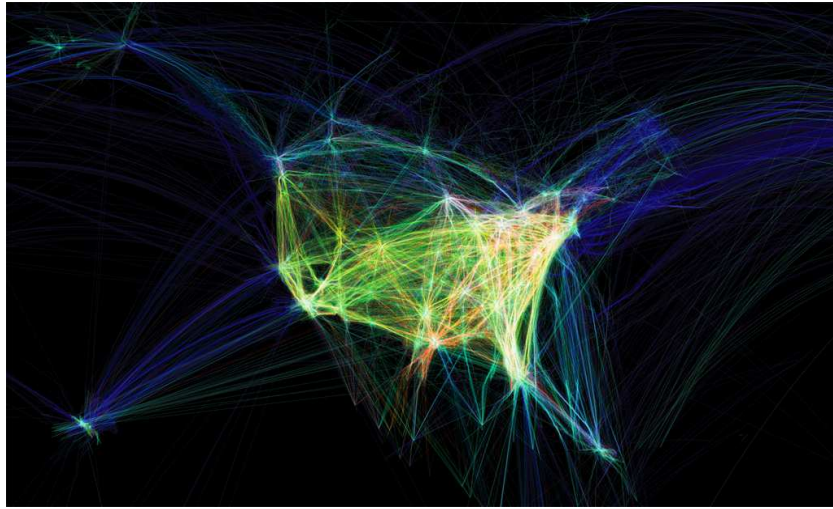
## Infographics + Data Art



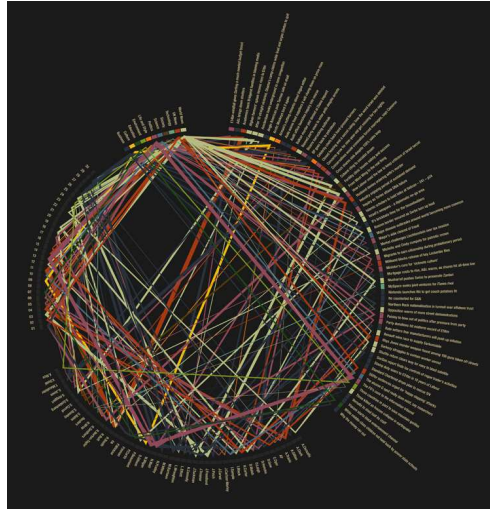
## Infographics + Data Art



## Infographics + Data Art



## Infographics + Data Art



53

- Guardian



## 4. „Falsche“ Datenvisualisierungen

„Lügen mit Statistik“...



- Dilbert und xkcd sind Fundgruben!

## 5. Richtlinien für Datenvisualisierungen

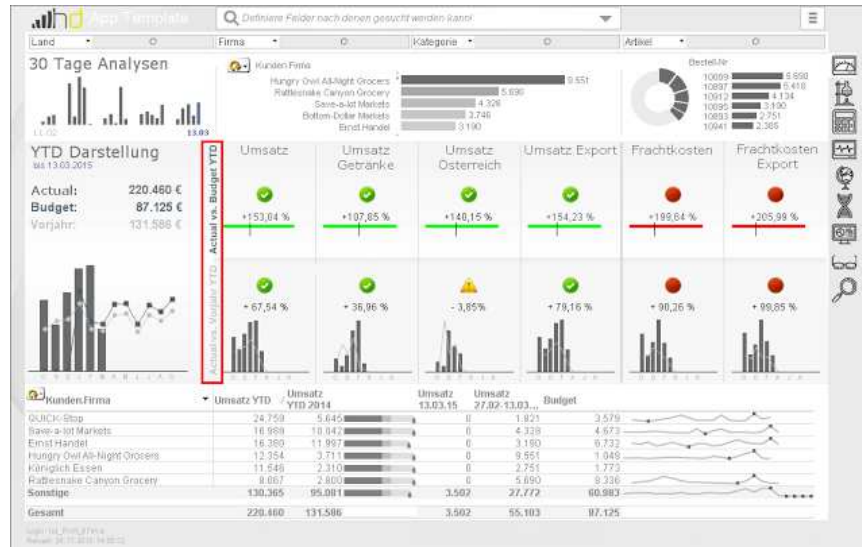


Seurat hebt die Kanten bewusst hervor



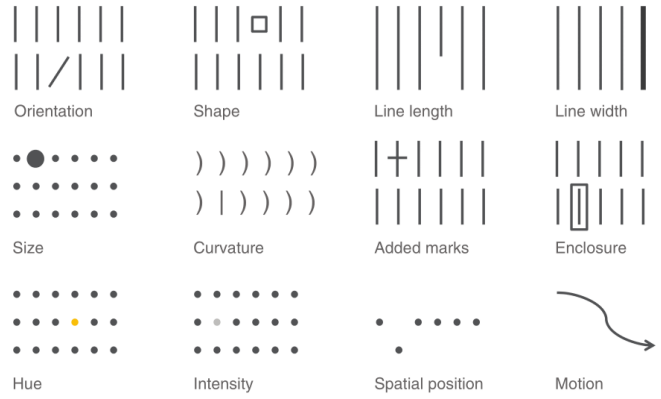
# Foliensammlung

# Qlikview



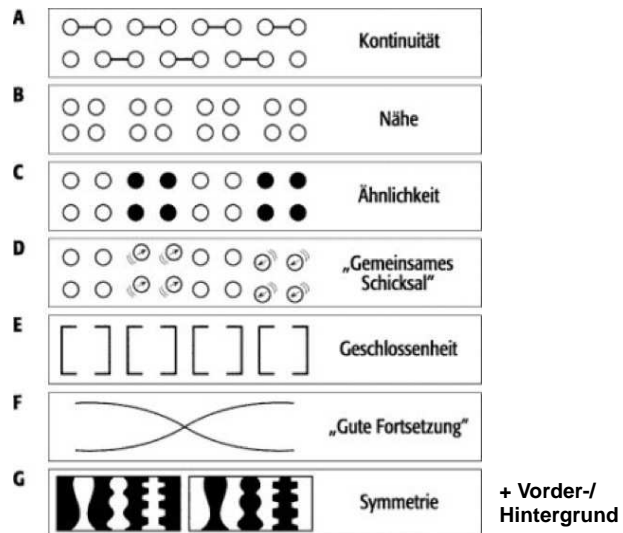
## 5. Richtlinien für Datenvisualisierungen

### Prä-attentive Wahrnehmung



## 5. Richtlinien für Datenvisualisierungen

### Weitere Gestalt - Prinzipien



## 5. Richtlinien für Datenvisualisierungen

### Sequential

Colors can be ordered from low to high



### Diverging

Two sequential schemes extending out from a critical midpoint value

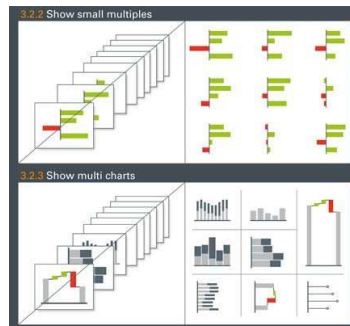


### Categorical

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

