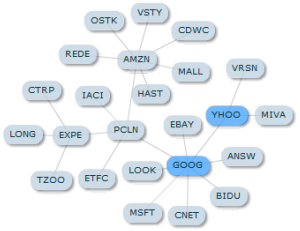
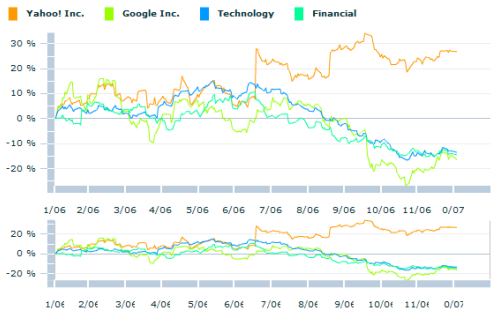
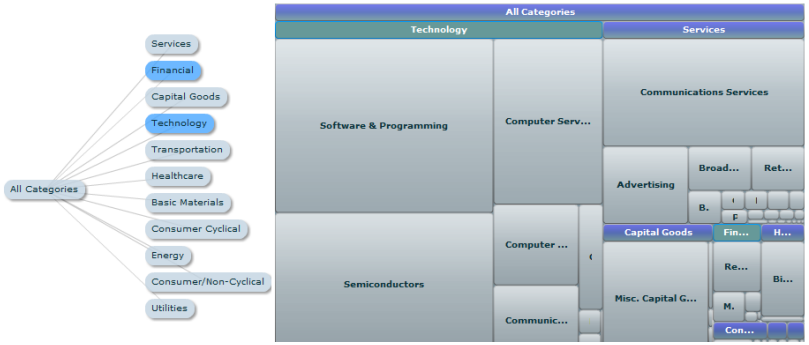
Informationsvisualisierung

Human Computer Systems

SS 2014







Dr. Tatiana von Landesberger

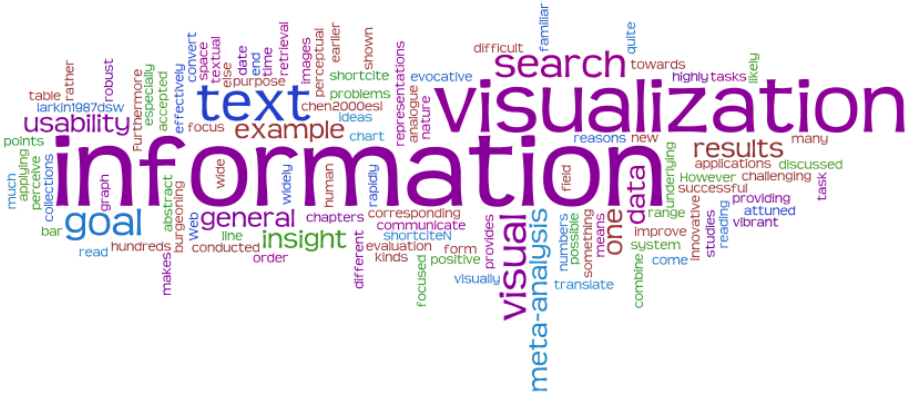
Technische Universität Darmstadt
Fachgebiet Graphisch-Interaktive Systeme
Fraunhoferstraße 5
64283 Darmstadt
Germany

Tel.: +49 (6151) 155 – 631
Fax: +49 (6151) 155 – 669

Email: ttekusov@gris.informatik.tu-darmstadt.de
http://www.gris.informatik.tu-darmstadt.de/~ttekusov

Was lernen wir heute…

- Was ist Informationsvisualisierung?
- Was ist (guter) Informationsdesign
- Welche Datentypen
  + Visualisierungstypen gibt es?
- Interaktion



INFORMATIONSVISUALISIERUNG

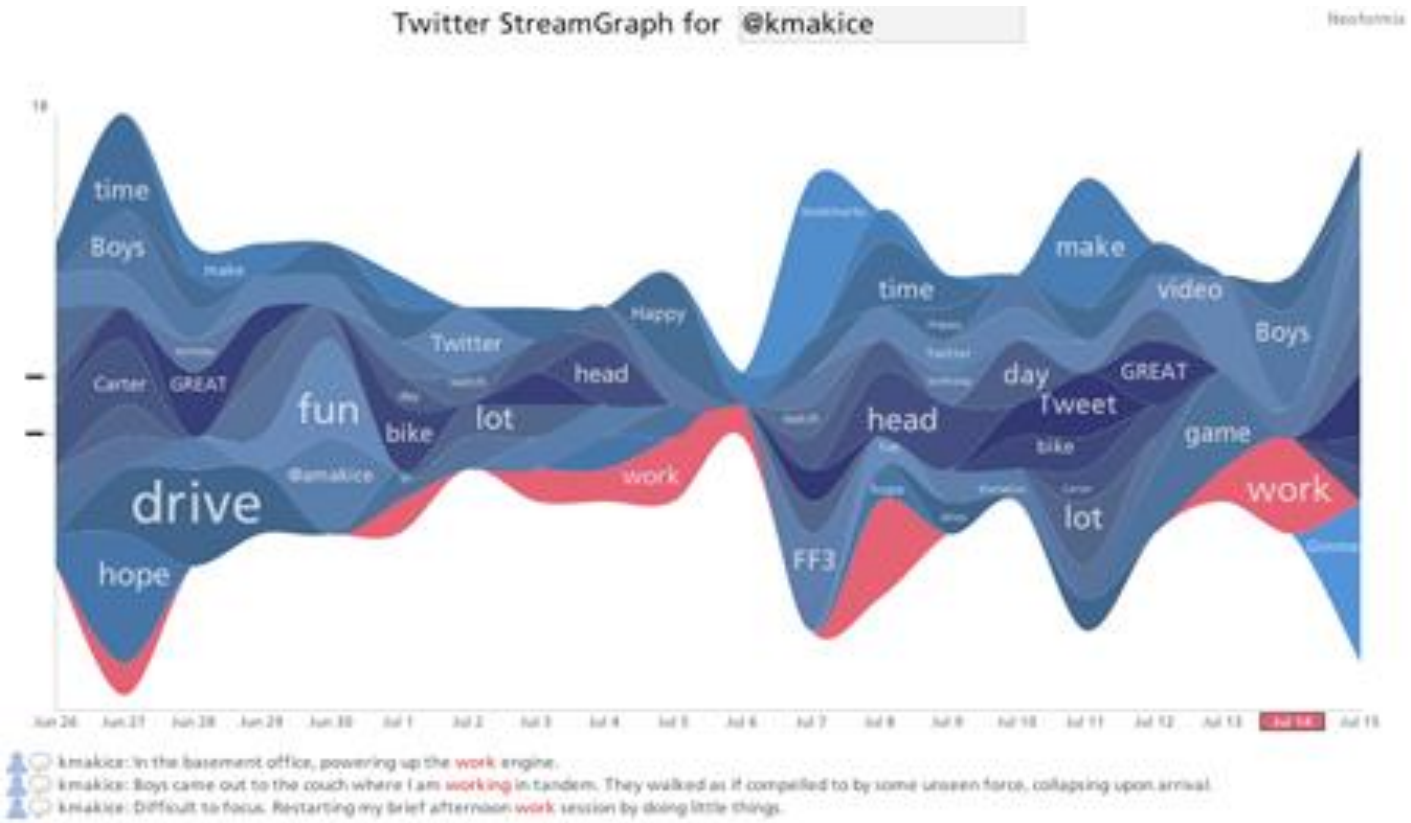
Finanzmarktvisualisierung

http://www.dashboardinsight.com/CMS/d317219f-99c0-4eaf-8193-2617c213b550/Dundas_data_visualization_stock_market_watch_digital_dashboard.png

Facebook Visualisierung

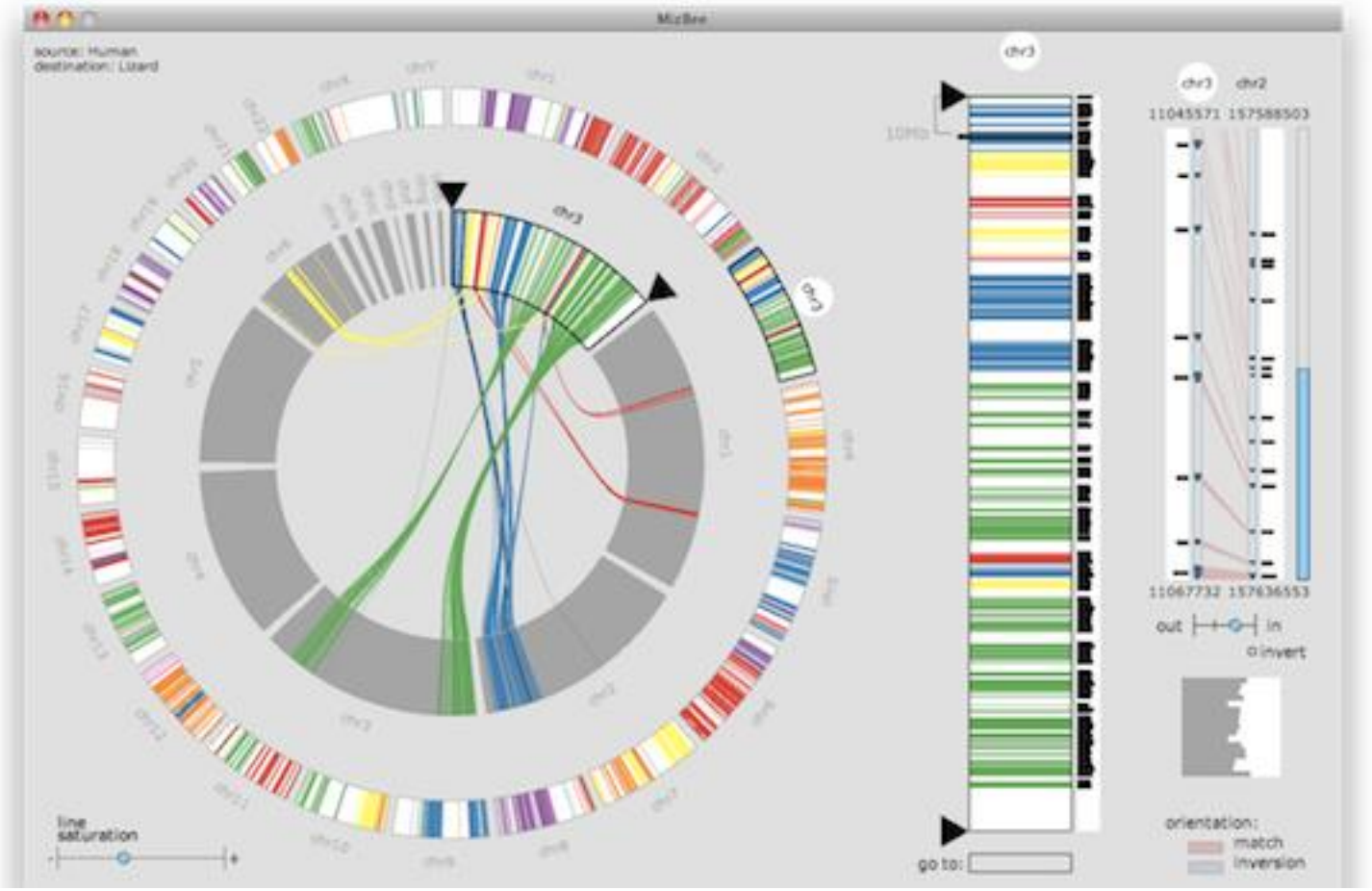
http://a6.sphotos.ak.fbcdn.net/hphotos-ak-snc4/163413_479288597199_9445547199_5658562_8388607_n.jpg

Twitter messages



http://www.blogschmog.net/wp-content/uploads/2008/07/streamgraph_kmakice.png

Genome Visualisierung



http://www.mizbee.org/Overview.html

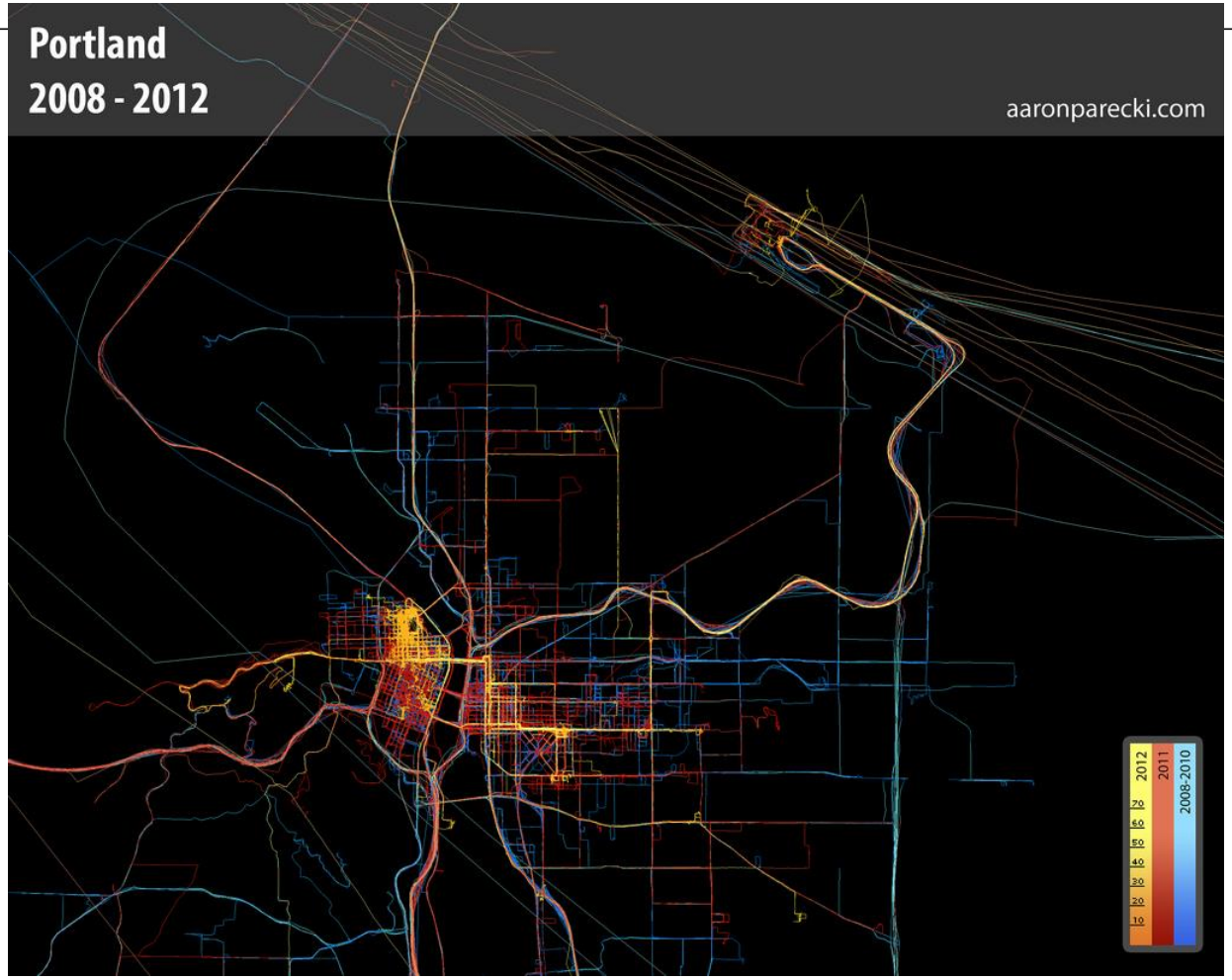
Musik

The Shape of Song

Madonna, Like a prayer

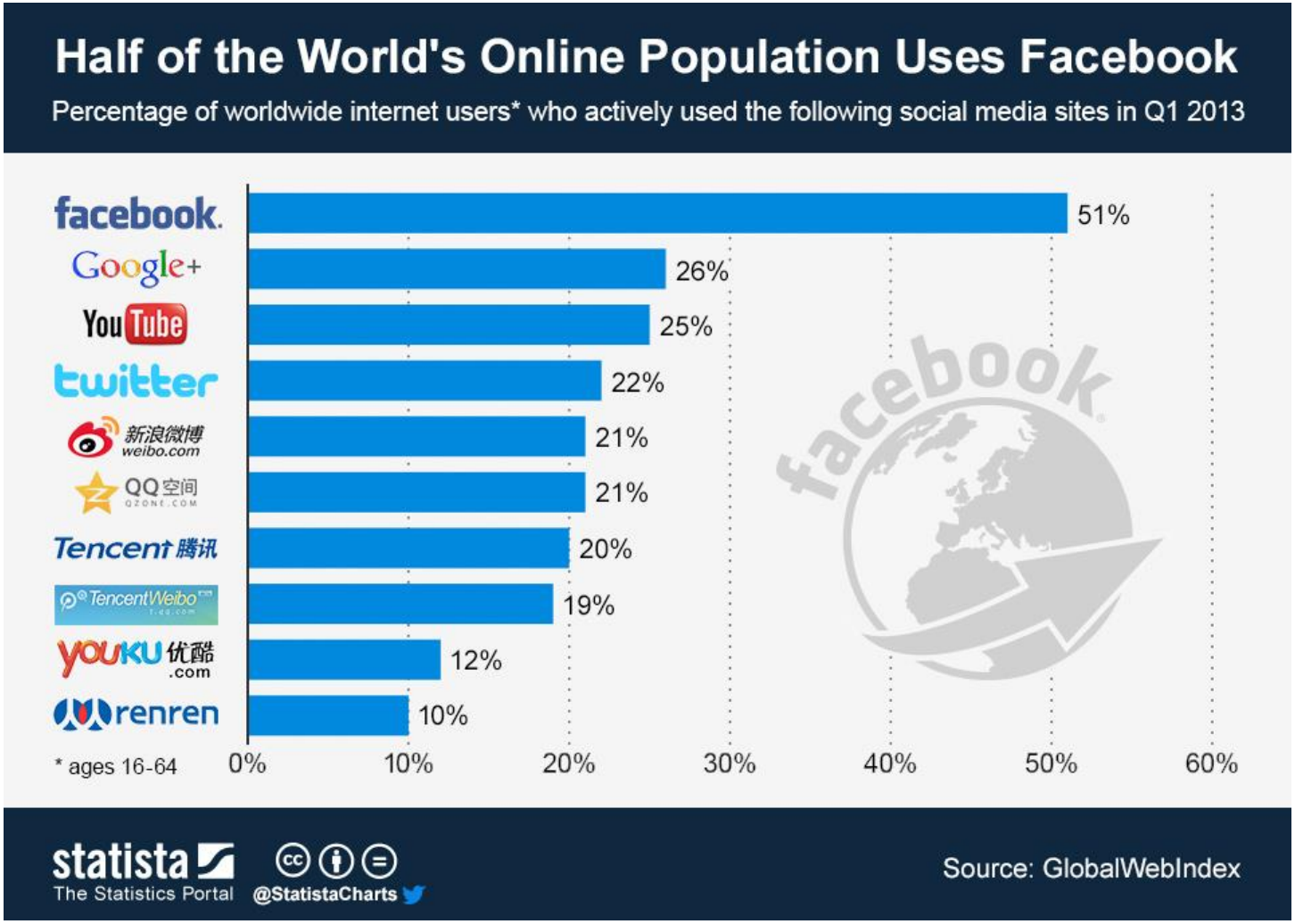
http://www.visualcomplexity.com/vc/project_details.cfm?id=371&index=7&domain=Music

GPS Bewegungsdaten



https://www.flickr.com/photos/aaronpk/6958126945/sizes/l/in/photostream/

Social service usage



WAS IST INFORMATIONSVISUALISIERUNG?

„Ein Bild sagt mehr als 1000 Worte“

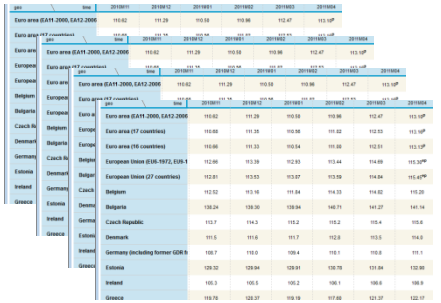
Was ist Informationsvisualisierung?

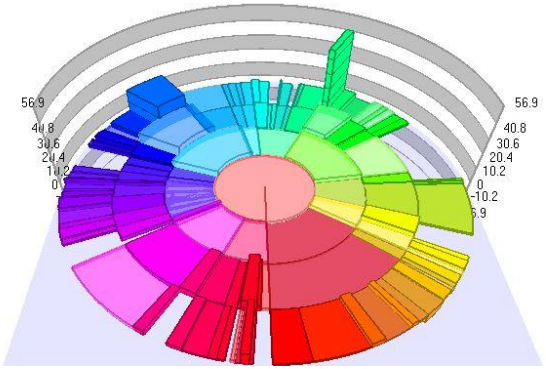
TECHNISCHE UNIVERSITÄT DARMSTADT

Visualisieren:

- Sich von etwas ein mentales Bild machen
- Sich etwas bildlich vorstellen
- „Ein Bild sagt mehr als 1000 Worte“

- **Visualisierung:**
  - Die bestmögliche Unterstützung der natürlichen Intelligenz… durch ein künstliches Gedächtnis, das unsere natürliche Wahrnehmung optimal unterstützt. (Bertin, 1983)
  - Die Transformation des Symbolischen in das Geometrische. (McCormick et al., 1987)





TUD Graphisch-Interaktive Systeme

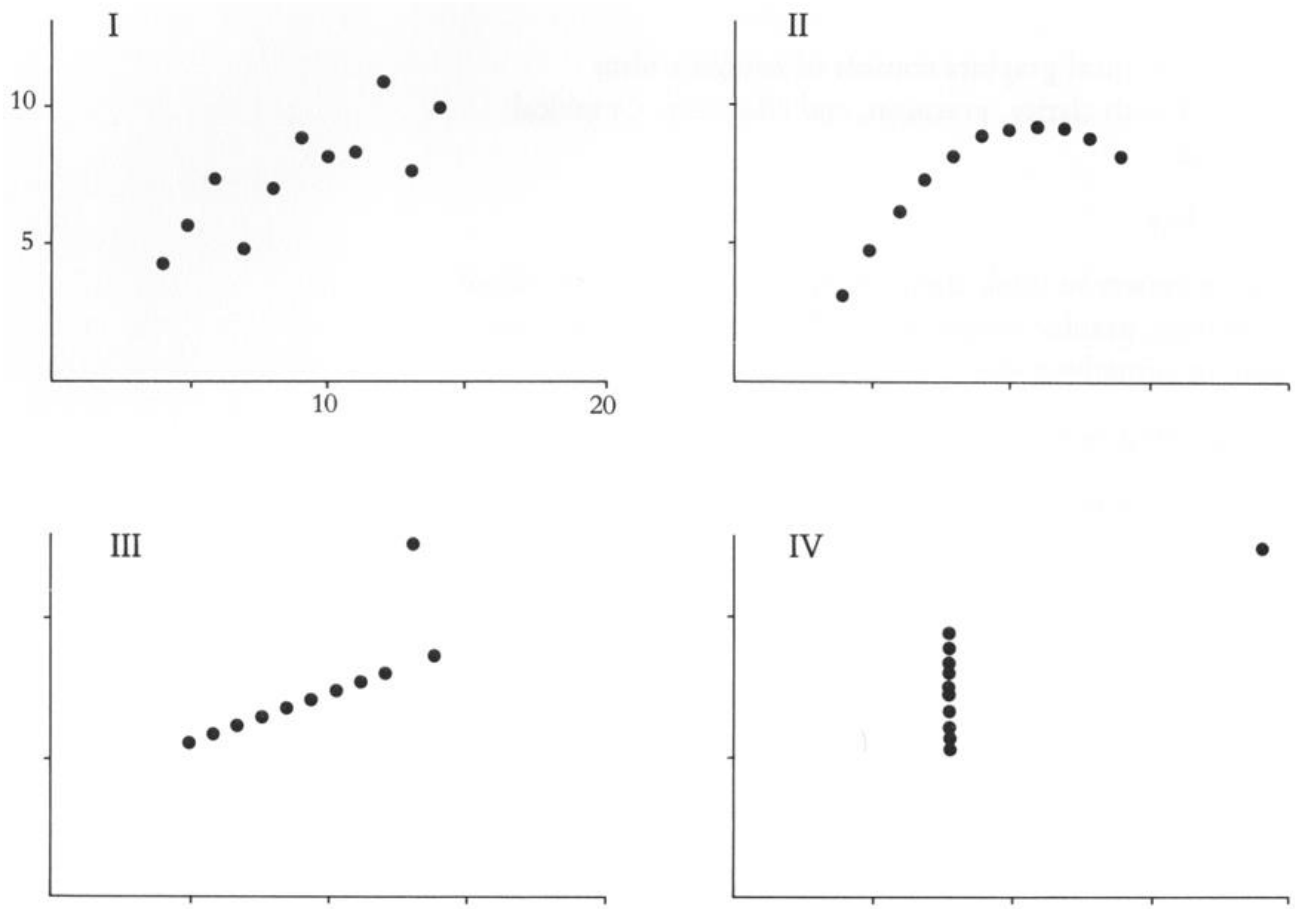
Graphische Darstellungen „enthüllen“ Daten

I		II		III		IV	
X	Y	X	Y	X	Y	X	Y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

$N = 11$
mean of X's $= 9.0$
mean of Y's $= 7.5$
equation of regression line: $Y = 3 + 0.5X$
standard error of estimate of slope $= 0.118$
$t = 4.24$
sum of squares $X - \overline{X} = 110.0$
regression sum of squares $= 27.50$
residual sum of squares of $Y = 13.75$
correlation coefficient $= .82$
$r^2 = .67$

Vier Datensätze mit identischem linearen Modell
(Tufte, 1983)

Graphische Darstellungen „enthüllen“ Daten

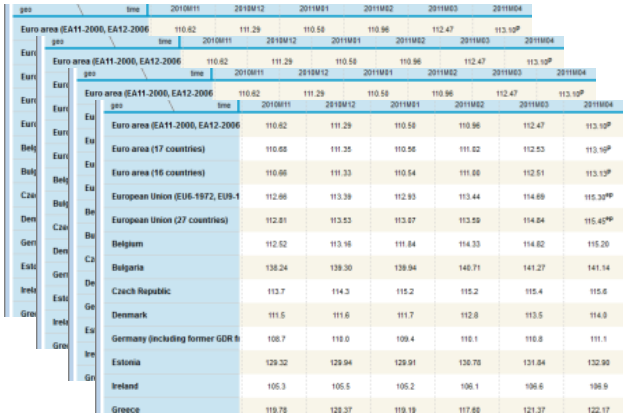


(Tufte, 1983)

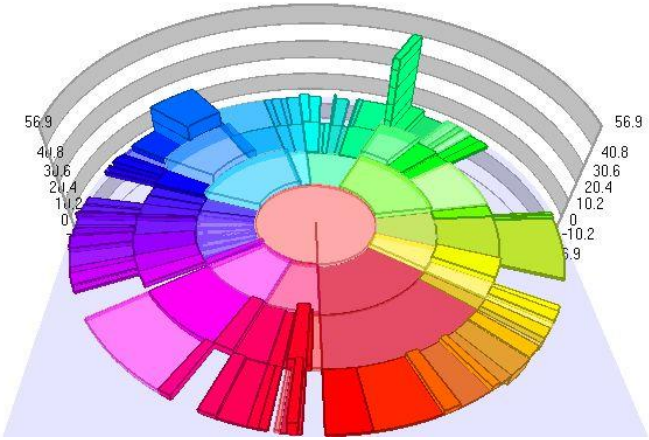
INFORMATIONSDESIGN

Informationsdesign

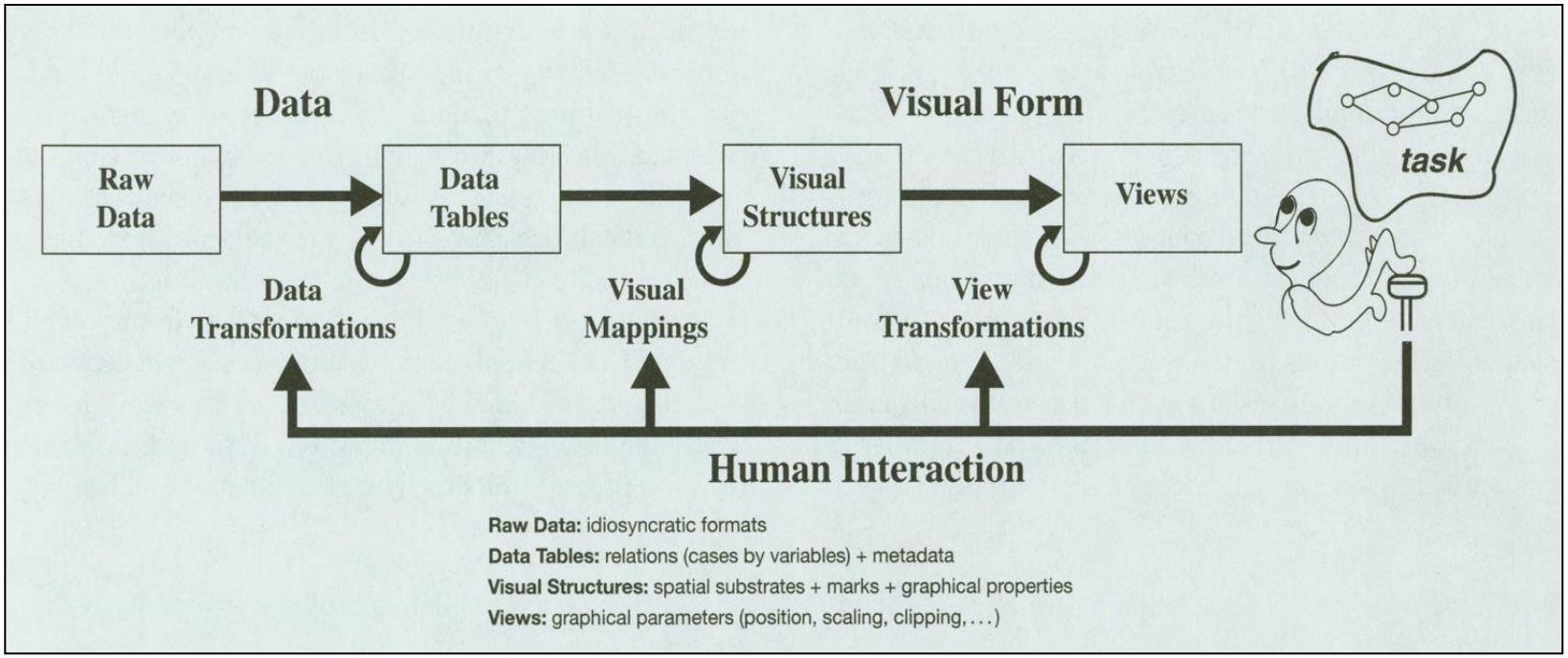
Wie können wir abstrakte Daten
in visuelle Representierung umwandeln?

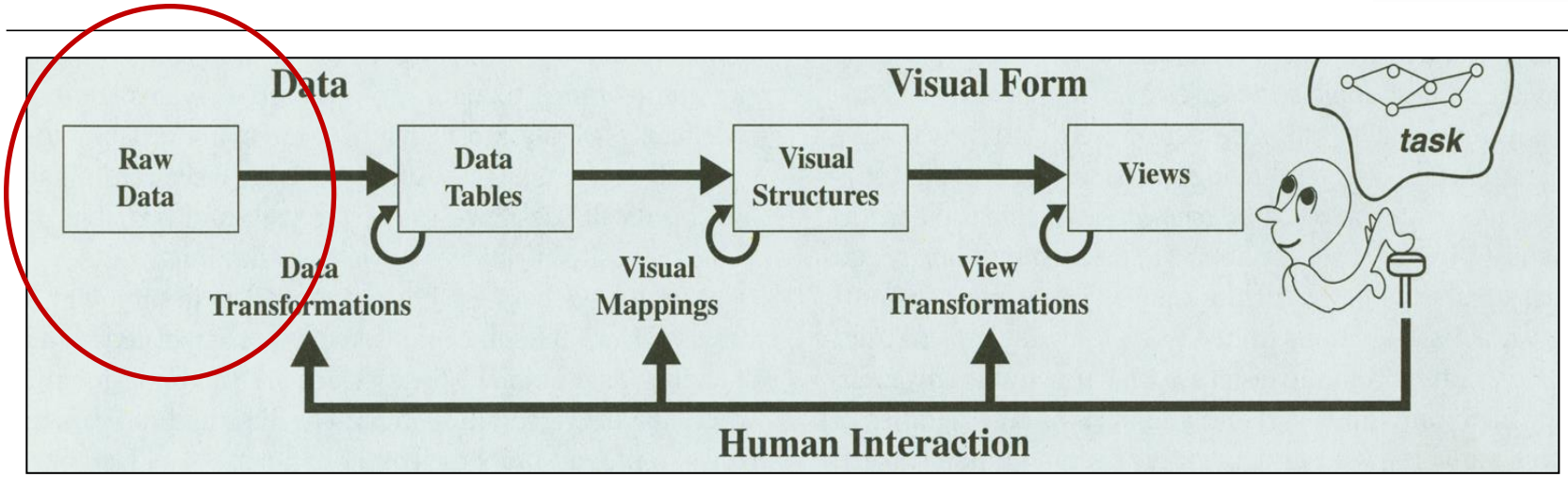


geo \ time	2010M11	2010M12	2011M01	2011M02	2011M03	2011M04
Euro area (EA11-2000, EA12-2006	110.62	111.29	110.58	110.96	112.47	113.10p
Euro area (EA11-2000, EA12-2006	110.62	111.29	110.58	110.96	112.47	113.10p
Euro area (EA11-2000, EA12-2006	110.62	111.29	110.58	110.96	112.47	113.10p
Euro area (EA11-2000, EA12-2006	110.62	111.29	110.58	110.96	112.47	113.10p
Euro area (17 countries)	110.68	111.35	110.56	111.02	112.53	113.10p
Euro area (16 countries)	110.66	111.33	110.54	111.00	112.51	113.13p
European Union (EU6-1972, EU9-1	112.66	113.39	112.93	113.44	114.69	115.39pp
European Union (27 countries)	112.81	113.53	113.07	113.59	114.84	115.45pp
Belgium	112.52	113.16	111.84	114.33	114.82	115.20
Bulgaria	138.24	139.30	139.94	140.71	141.27	141.14
Czech Republic	113.7	114.3	115.2	115.2	115.4	115.6
Denmark	111.5	111.6	111.7	112.0	113.5	114.0
Germany (including former GDR fr	108.7	110.0	109.4	110.1	110.8	111.1
Estonia	129.32	129.94	129.91	130.70	131.84	132.90
Ireland	105.3	105.5	105.2	106.1	106.6	106.9
Greece	119.78	120.37	119.19	117.60	121.37	122.17

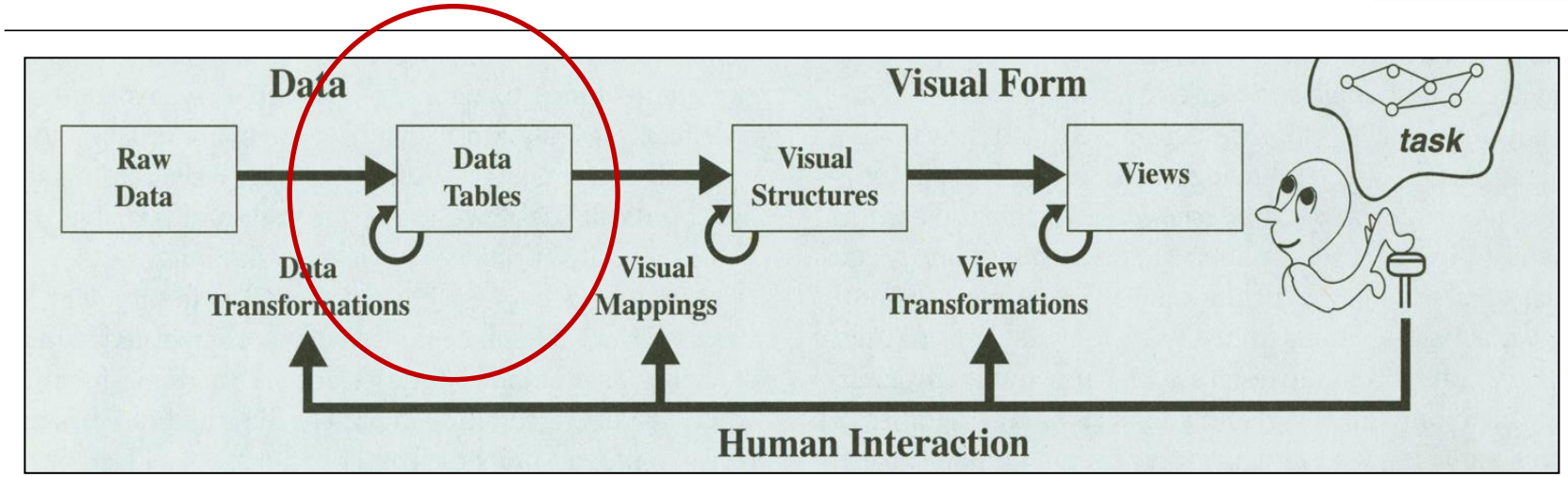


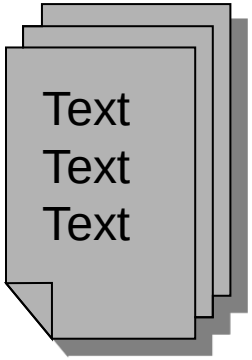
Referenzmodell von Card et al. (1999)



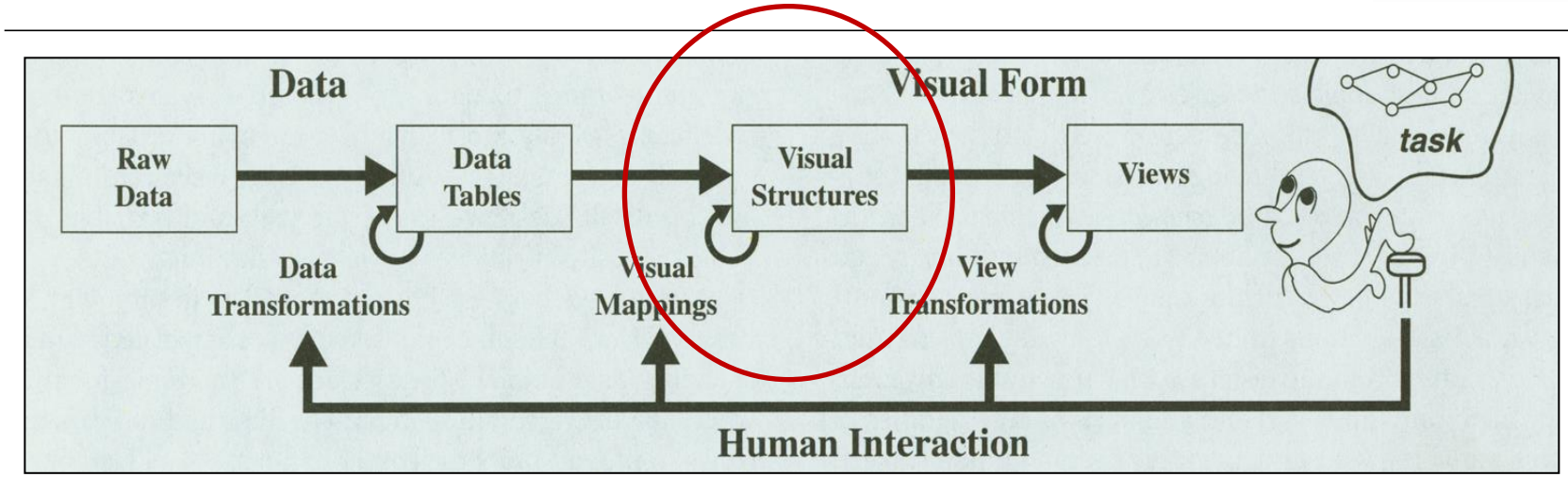


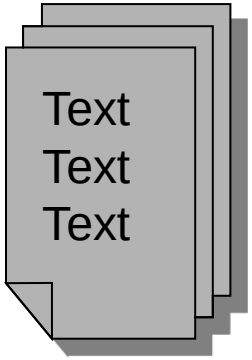
Text
Text
Text





#Worte	#Satze
50	9
100	20
80	7

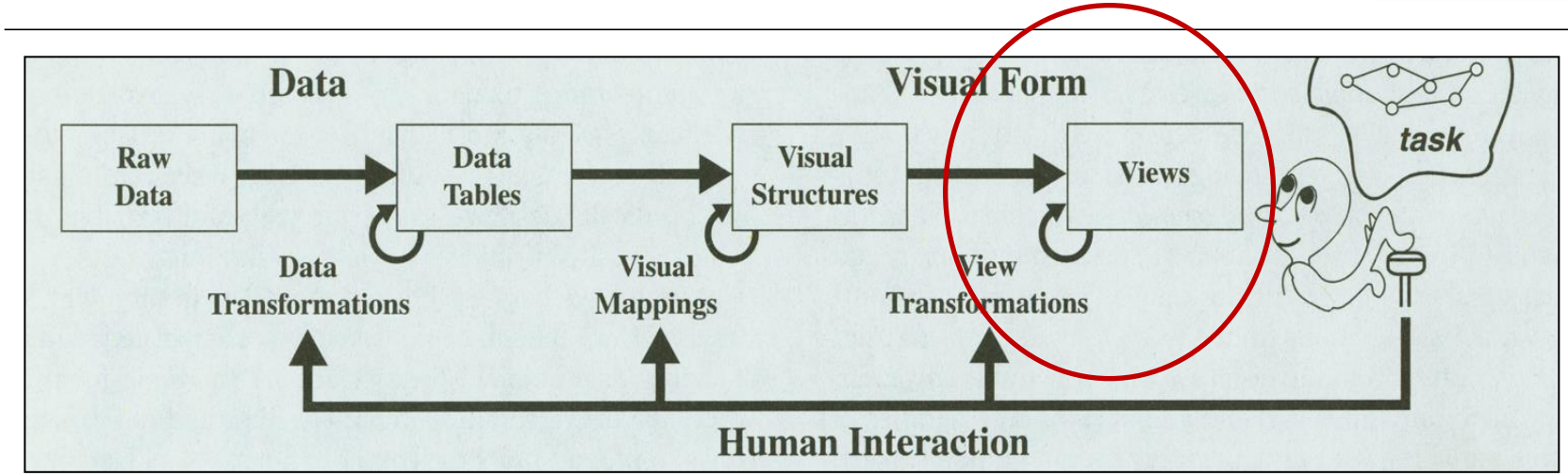


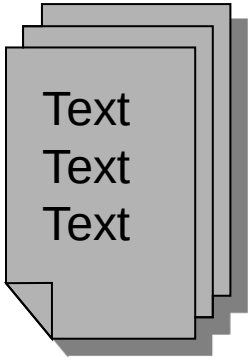


#Worte	#Satze
50	9
100	20
80	7

#Worte ⇨ Größe

#Satze ⇨ Farbe

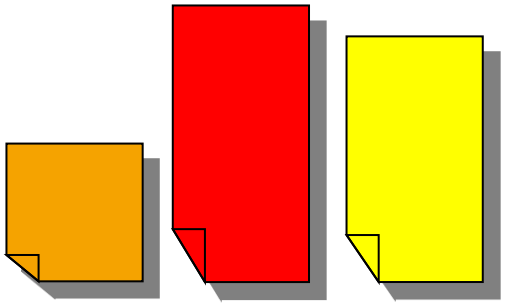


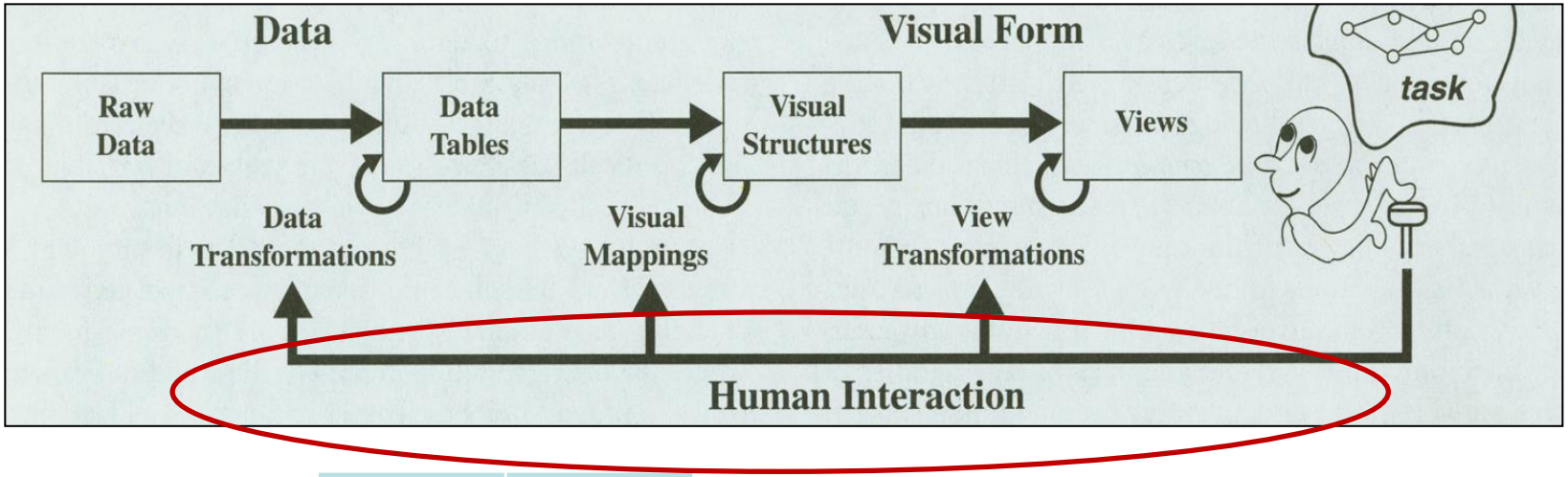


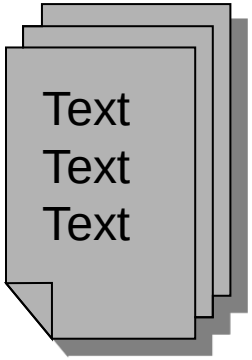
#Worte	#Satze
50	9
100	20
80	7

#Worte ⇨ Größe

#Satze ⇨ Farbe



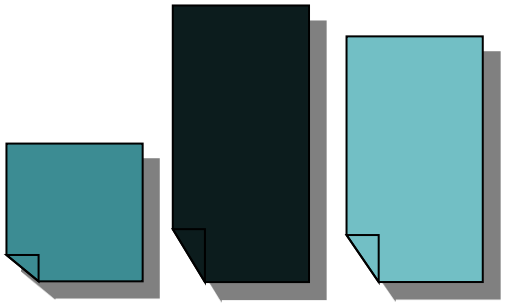




#Worte	#Satze
50	9
100	20
80	7

#Worte ⇨ Größe

#Satze ⇨ Farbe



Informationsdesign

Wie können wir abstrakte Daten
in visuelle Representierung umwandeln?

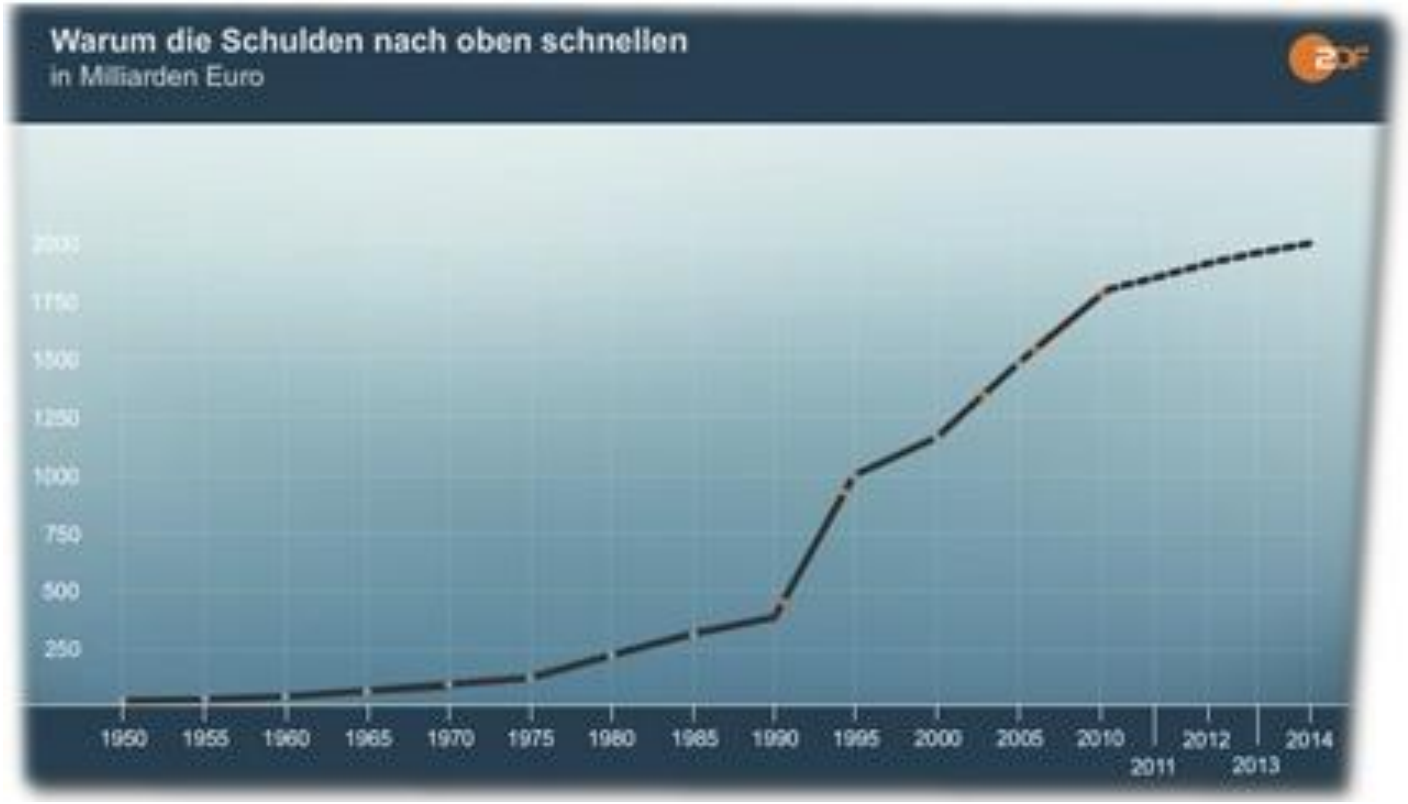
So dass…

wir immer noch die Bedeutung behalten…
und neue Informationen enthüllen?

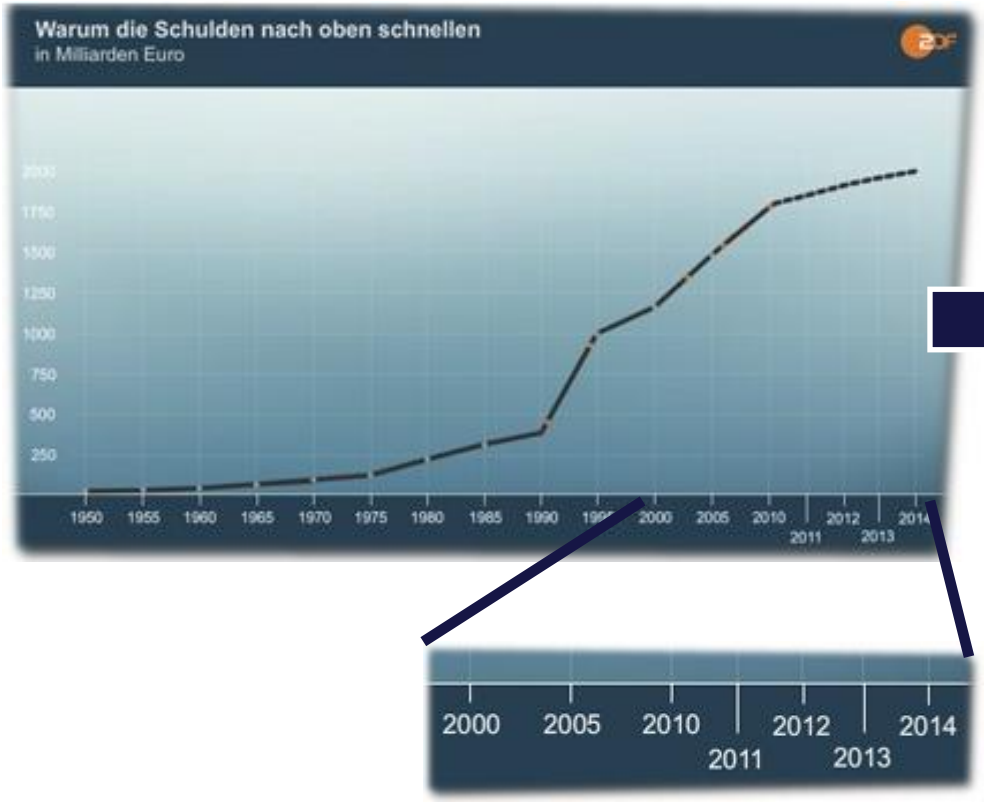
(Marti Hearst, 2004)

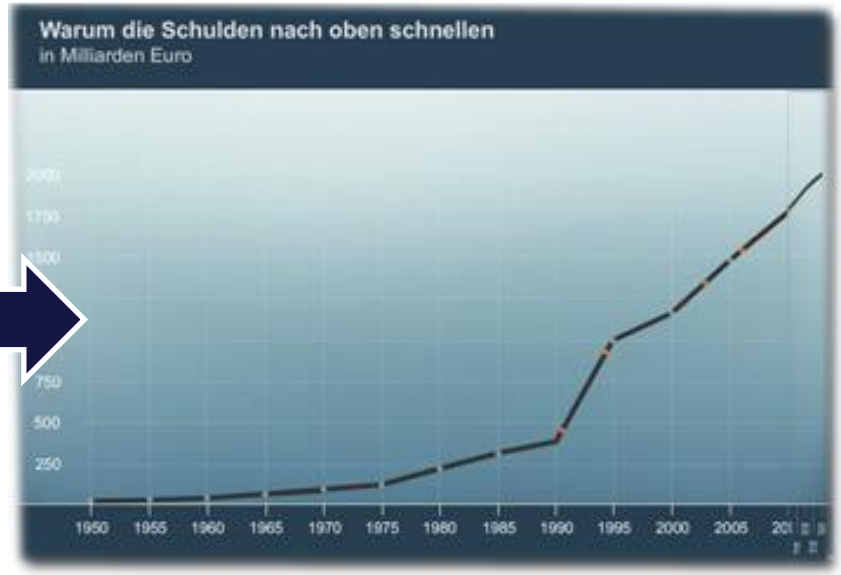
(GEGEN)BEISPIELE

Staatsverschuldung – was stimmt nicht?



Abstände (Zeiteinheiten) müssen stimmen

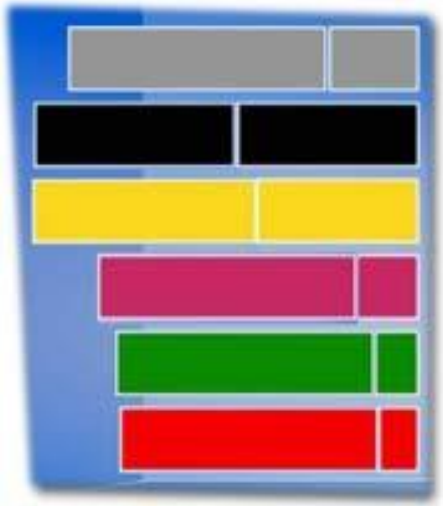




Quelle: Bildblog.de

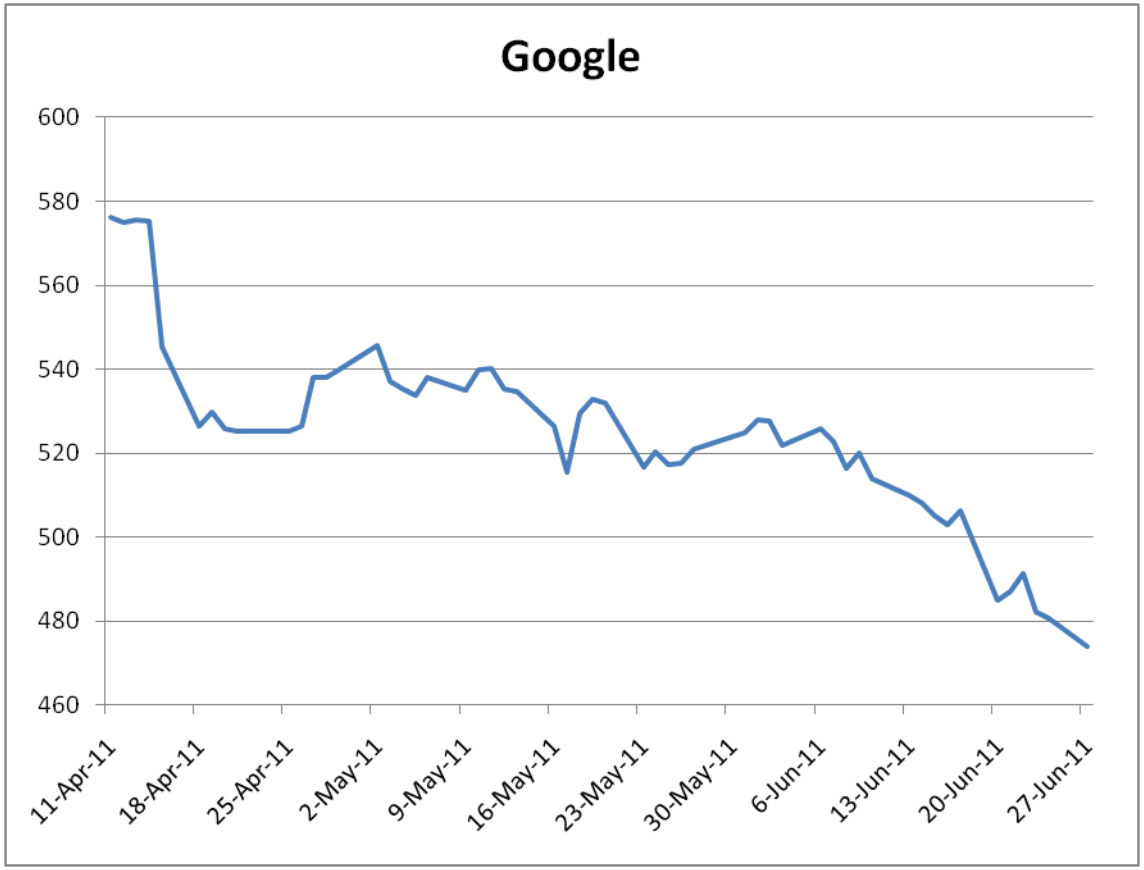
Regierungszufriedenheit: Was stimmt nicht?

Größenverhältnisse müssen stimmen…

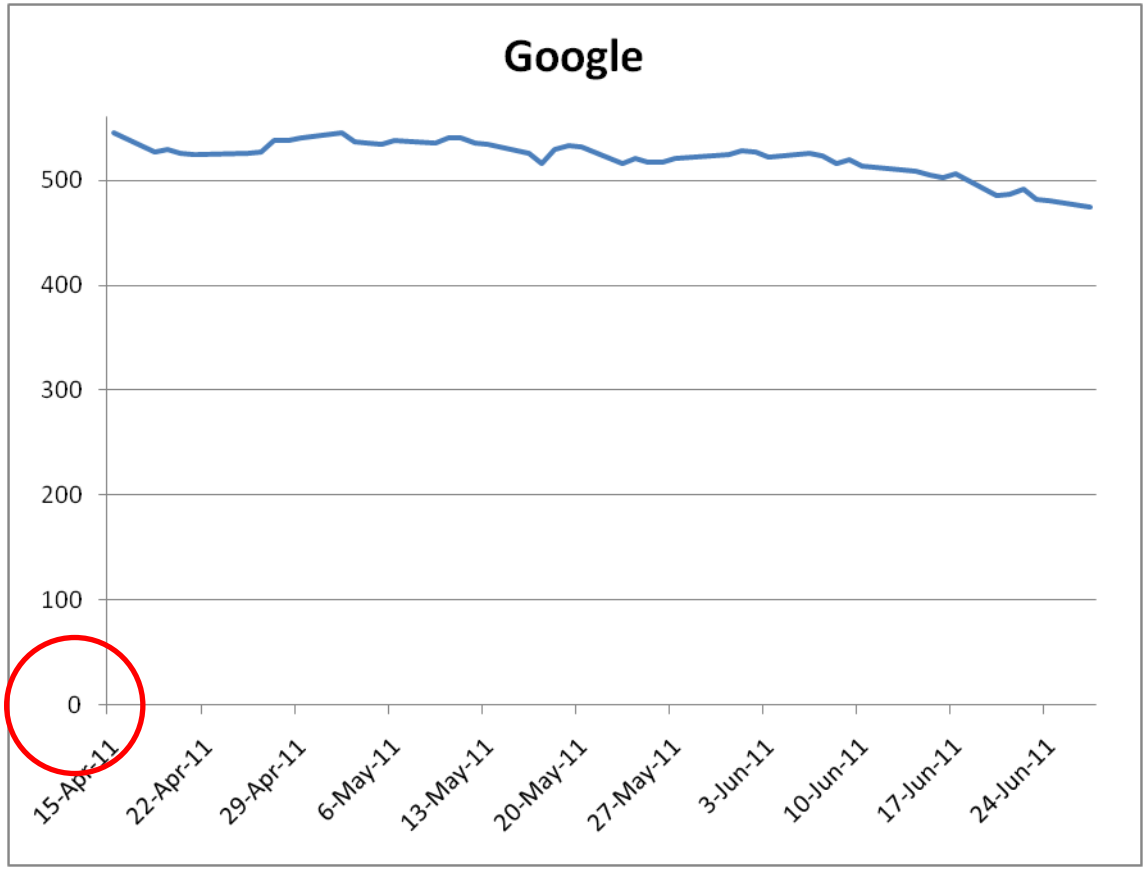


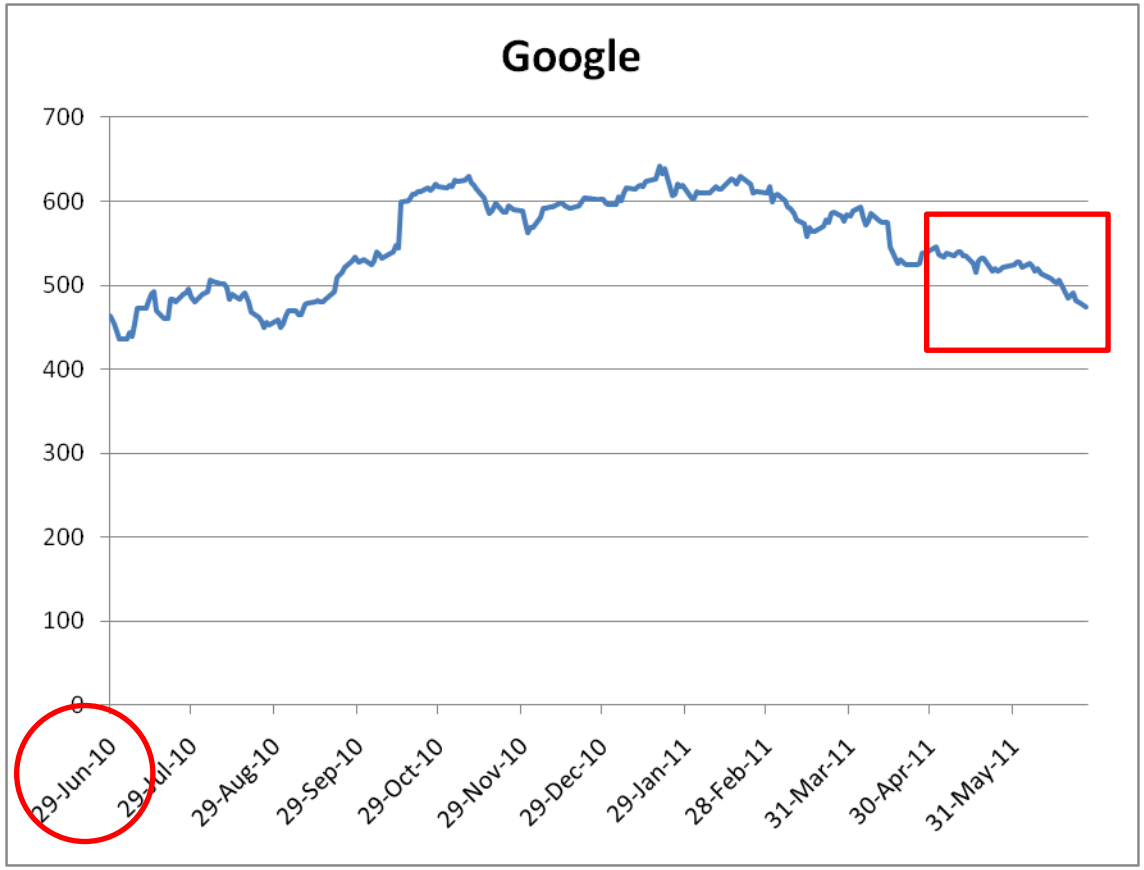
Quelle: Bildblog.de

Skalierung und Kontext ist wichtig

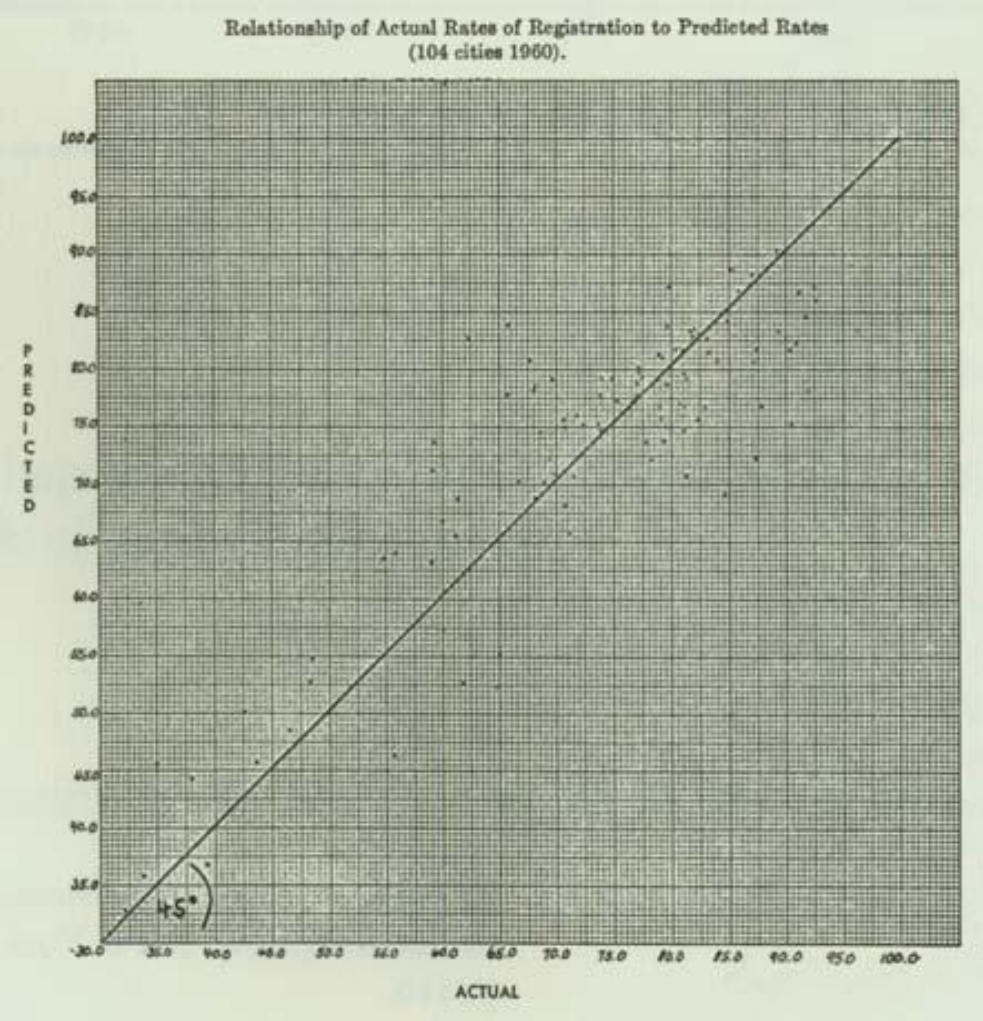


Skalierung und Kontext ist wichtig

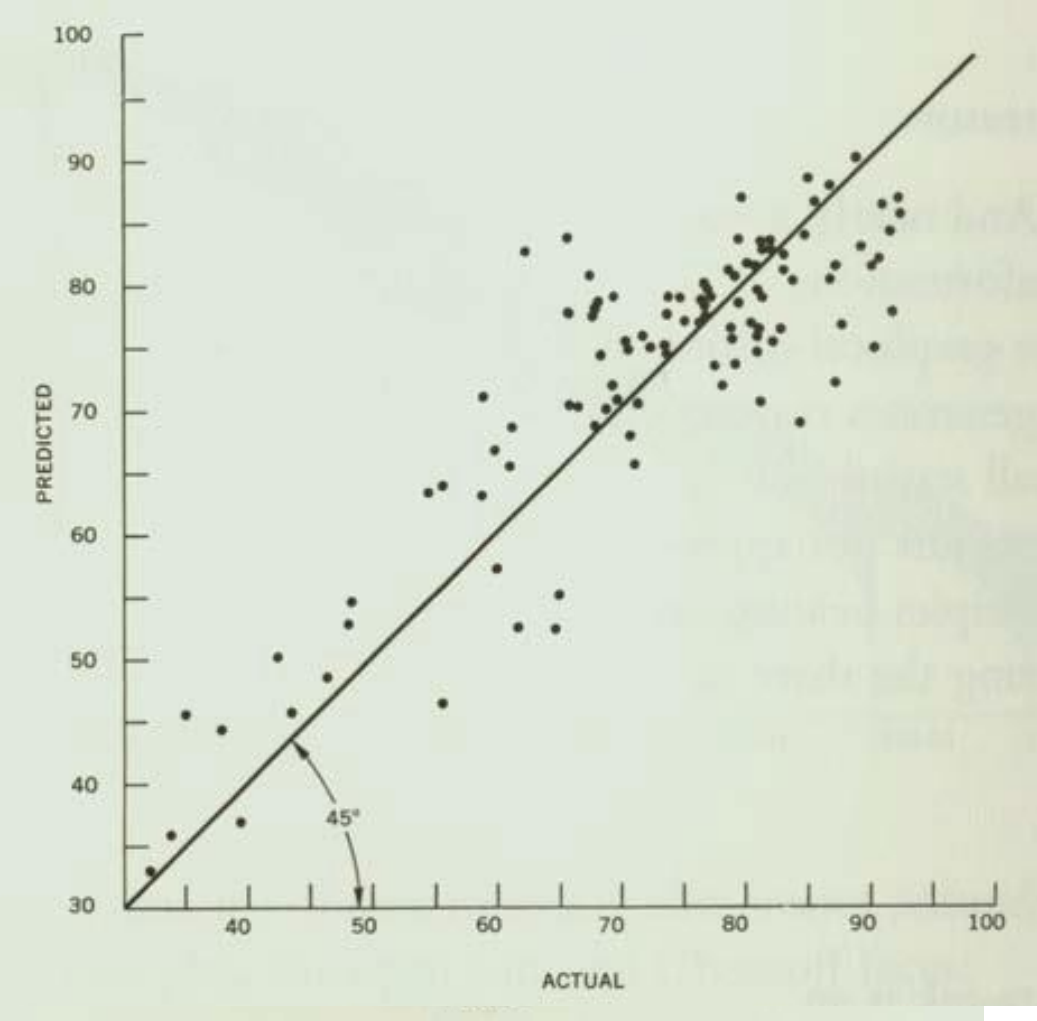



Google
700
600
500
400
300
200
100
0
29-Jun-10
29-Jul-10
29-Aug-10
29-Sep-10
29-Oct-10
29-Nov-10
29-Dec-10
29-Jan-11
28-Feb-11
31-Mar-11
30-Apr-11
31-May-11

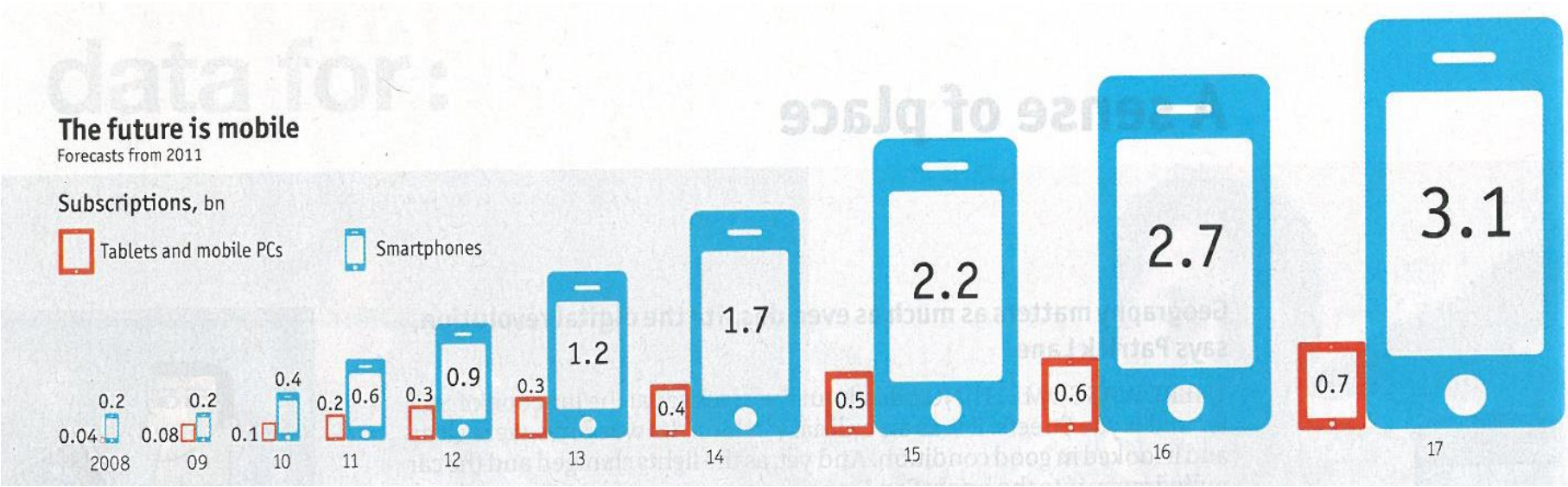
Was ist hier falsch?



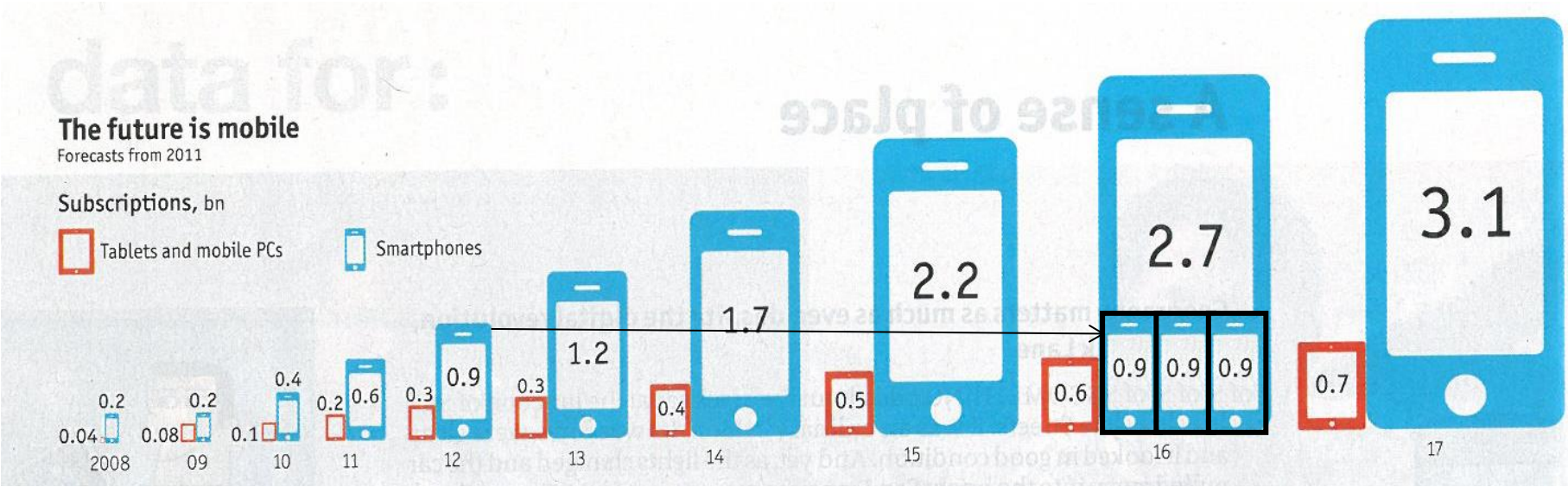
Nur wichtige Daten zeigen



Was ist hier irrefuehrend abgebildet?



Was ist hier irrefuehrend abgebildet?



Was ist hiern problematisch?

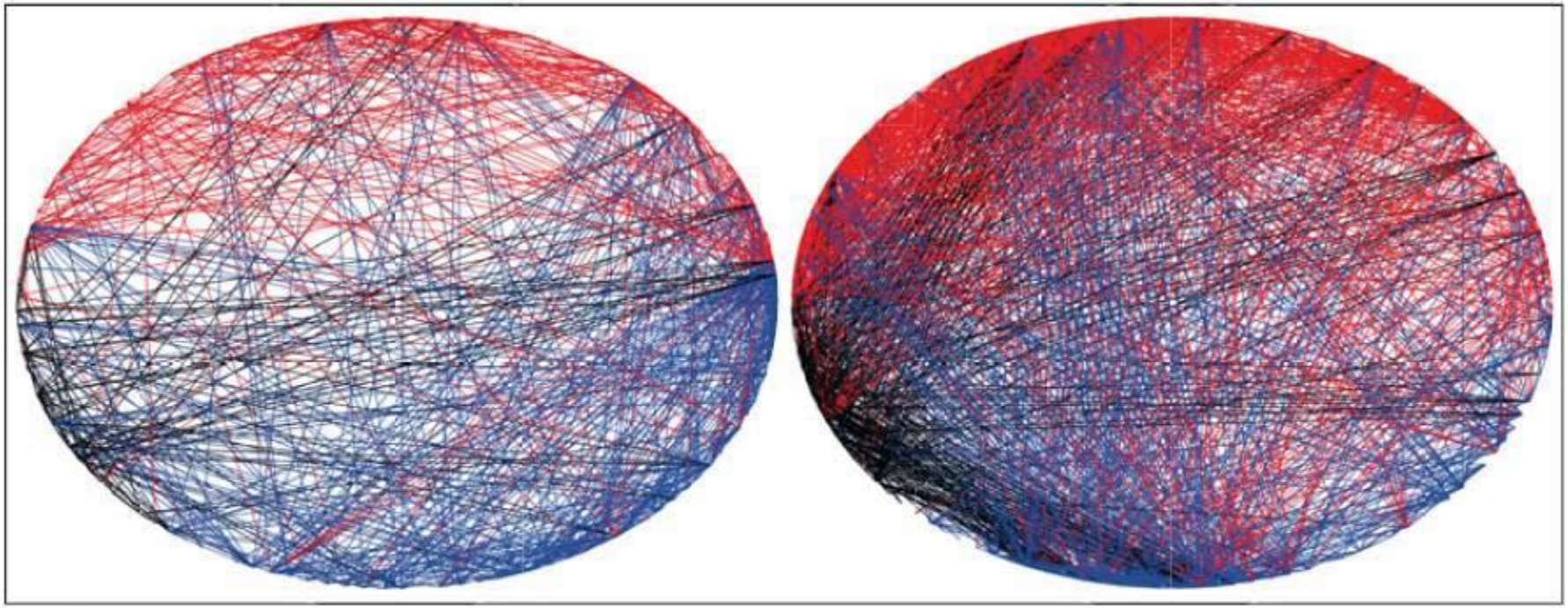
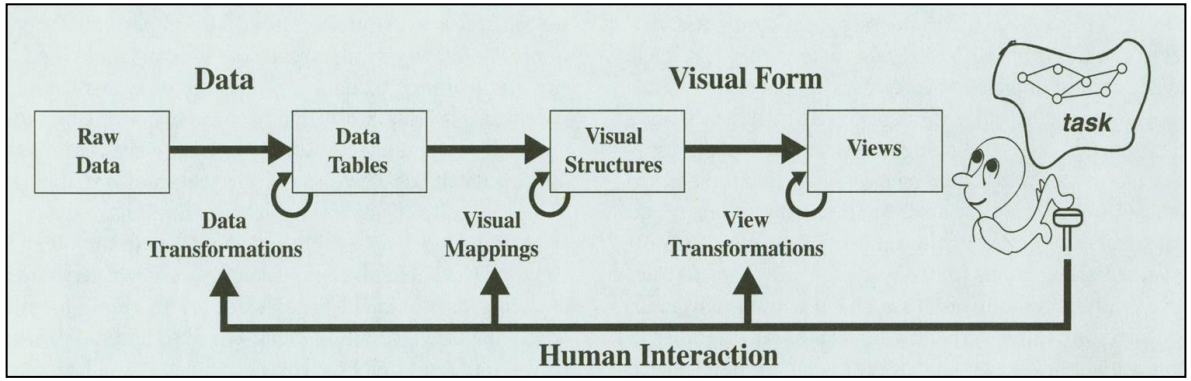


Figure 1: Network diagrams of linear Granger-causality relationships that are statistically significant at the 1% level between banks (red), insurers (black), and sovereigns (blue) – from July 2004 to June 2007 (left) and from April 2009 to March 2012 (right)

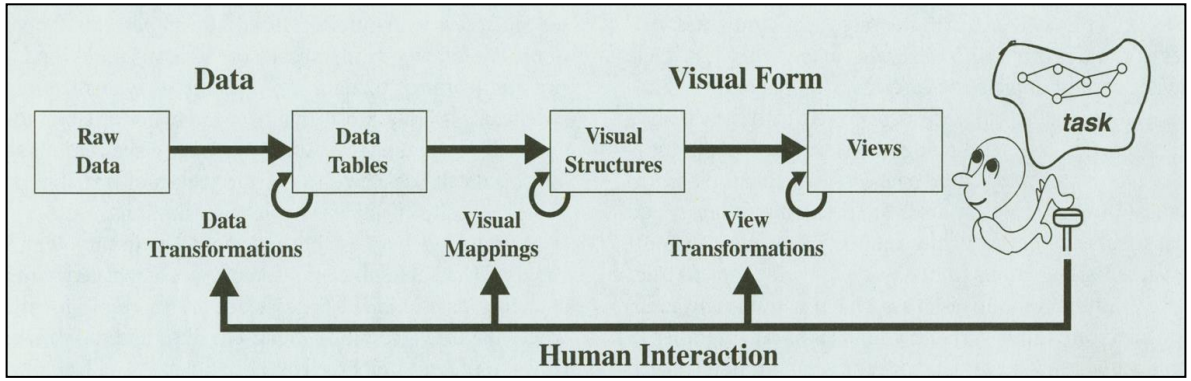
Informationsdesign: Zusammenfassung

- Prozess:
  Data → Visualisierung



Informationsdesign: Zusammenfassung

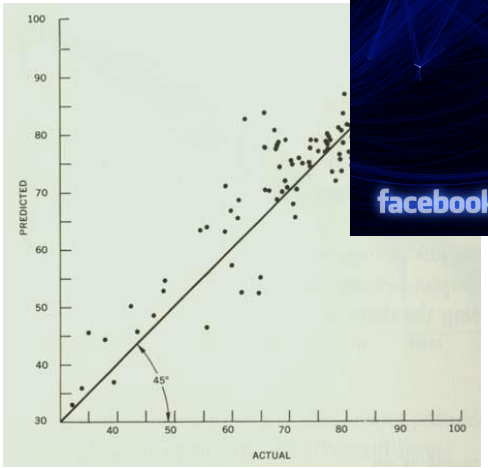
- Prozess:
  Data → Visualisierung

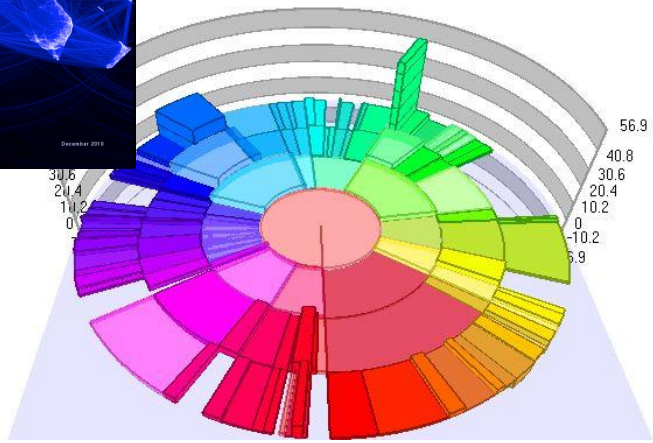


- Gute Visualisierungen sollten
  - Daten nicht Verzerren und Dateninformationen behalten
  - Probleme sollten vermieden werden:
    - Falsche/ irreführende Skalierung
    - Verzerrung, Groessenverhaeltnisse
    - Zu volle Darstellung
    - Keine Legende

INFORMATIONSVISUALISIERUNG TECHNIKEN

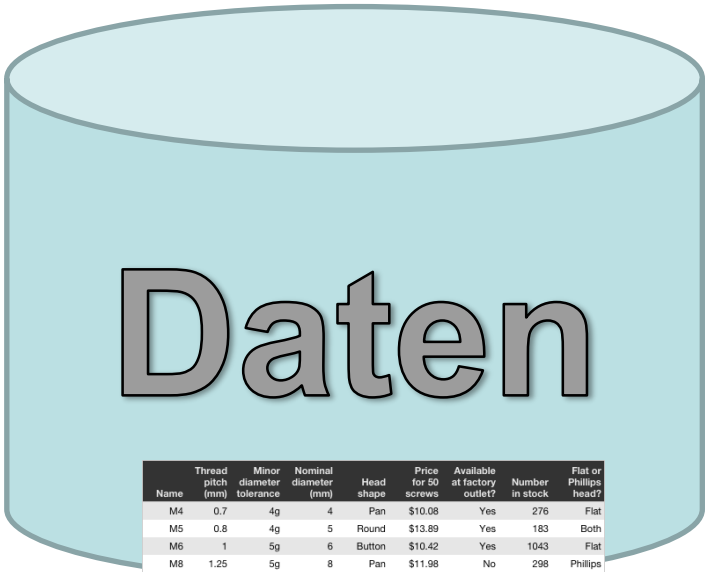
Informationsvisualisierung
Techniken und Datentypen



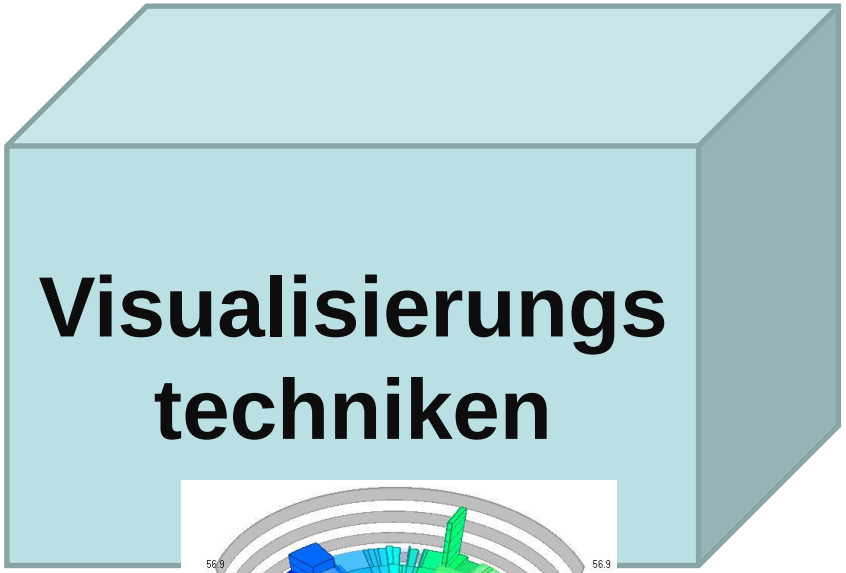


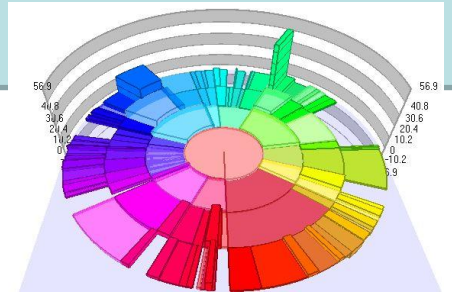
Informationsvisualisierung
Techniken und Datentypen

TECHNISCHE UNIVERSITÄT DARMSTADT

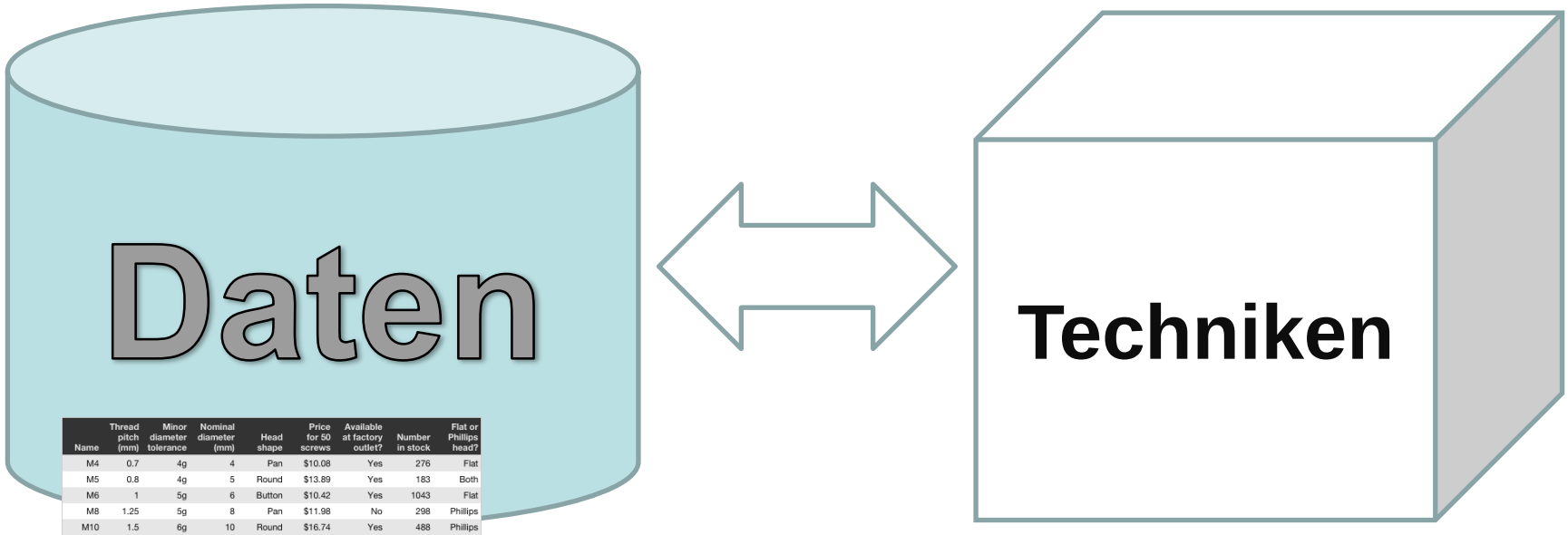


Name	Thread pitch (mm)	Minor diameter tolerance	Nominal diameter (mm)	Head shape	Price for 50 screws	Available at factory outlet?	Number in stock	Flat or Phillips head?
M4	0.7	4g	4	Pan	$10.08	Yes	276	Flat
M5	0.8	4g	5	Round	$13.89	Yes	183	Both
M6	1	5g	6	Button	$10.42	Yes	1043	Flat
M8	1.25	5g	8	Pan	$11.98	No	298	Phillips
M10	1.5	6g	10	Round	$16.74	Yes	488	Phillips
M12	1.75	7g	12	Pan	$18.26	No	998	Flat
M14	2	7g	14	Round	$21.19	No	235	Phillips
M16	2	8g	16	Button	$23.57	Yes	292	Both
M18	2.1	8g	18	Button	$25.87	No	664	Both
M20	2.4	8g	20	Pan	$29.09	Yes	486	Both
M24	2.55	9g	24	Round	$33.01	Yes	982	Phillips
M28	2.7	10g	28	Button	$35.66	No	1067	Phillips
M36	3.2	12g	36	Pan	$41.32	No	434	Both
M50	4.5	15g	50	Pan	$44.72	No	740	Flat





Informationsvisualisierung
Techniken und Datentypen



Name	Thread pitch (mm)	Minor diameter tolerance	Nominal diameter (mm)	Head shape	Price for 50 screws	Available at factory outlet?	Number in stock	Flat or Phillips head?
M4	0.7	4g	4	Pan	$10.08	Yes	276	Flat
M5	0.8	4g	5	Round	$13.89	Yes	183	Both
M6	1	5g	6	Button	$10.42	Yes	1043	Flat
M8	1.25	5g	8	Pan	$11.98	No	298	Phillips
M10	1.5	6g	10	Round	$16.74	Yes	488	Phillips
M12	1.75	7g	12	Pan	$18.26	No	998	Flat
M14	2	7g	14	Round	$21.19	No	235	Phillips
M16	2	8g	16	Button	$23.57	Yes	292	Both
M18	2.1	8g	18	Button	$25.87	No	664	Both
M20	2.4	8g	20	Pan	$29.09	Yes	486	Both
M24	2.55	9g	24	Round	$33.01	Yes	982	Phillips
M28	2.7	10g	28	Button	$35.66	No	1067	Phillips
M36	3.2	12g	36	Pan	$41.32	No	434	Both
M50	4.5	15g	50	Pan	$44.72	No	740	Flat

Datentypen: 1D

- 1 Spalte mit Datenwerten
  - Datenwerte über Objekte
  - Verteilungen

Bereich	Häufigkeit
0-10	1
10-20	3
20-30	6
30-40	4
40-50	2

Datentypen: 1D: Zeitreihen

- Zeitreihen
  - Meteorologie
  - Netzwerkmonitoring
  - Finanzen
  - …
- Sind spezielle 1D Daten, viele sagen es sind 2D Daten da 2 Spalten...
    - Irgendwie Beides richtig aber 1D passt besser da nur 1 Wertspalte

Zeit	Wert
1.1.2012	35
2.1.2012	40
3.1.2012	70
...	..
30	30

Datentypen: 2D

- 2 Spalten mit Datenwerten
  - Datenwerte über Objekte

Wife‘s age	Husband‘s age
50	45
80	70
60	65
...	..
30	30

Datentypen: 3D,… multi-dimensionale Daten

- 3,…, n Spalten mit Datenwerten

- **nD= multidimensionale Daten = multivariate Daten**
  - Datenwerte über Objekte
    - Befragungsdaten
    - Objektbeschreibungen,…

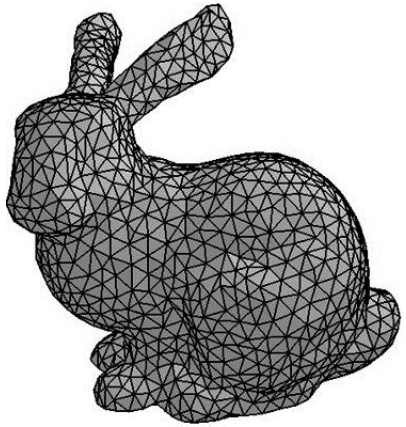
Platz	**Hersteller**
1	BMW
2	Audi
3	VW
4	BMW
5	Peugeut
6	BMW

Datentypen: 3D, ... , nD

- **3D**,…, n Spalten mit Datenwerten

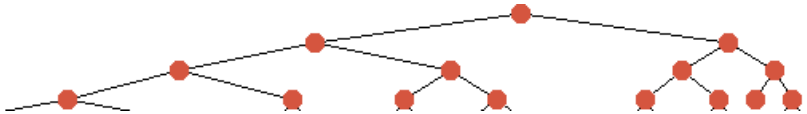
Achtung:

3D Daten sind in Komputergraphik meistens mit Volumendaten, Meshes und Punktwolken verbunden, in InfoVis anders



Datentypen: Hierarchien

- Hierarchien
  - Beziehungen zwischen Objekten Top-Down (Vater-Sohn)



  - Natürliche Art, Daten zu strukturieren bzw. zu organisieren
    - Unternehmensstruktur
    - Geographie
    - Produkt- und Teilekataloge
    - Dateisysteme
    - u.v.m.

Netzwerke/Graphen

- Graph G(V, E) - ein Objekt bestehend aus Knoten und Kanten
  Knoten: Objekte
  Kanten: Beziehungen zwischen Objekten

- Graphen sind sehr verbreitet
  - Computernetzwerke
  - Softwaremodulen/Klassen
  - WWW
  - Soziale Netzwerke
  - Biologische Reaktionen
  - Transportnetzwerke
  - Etc.

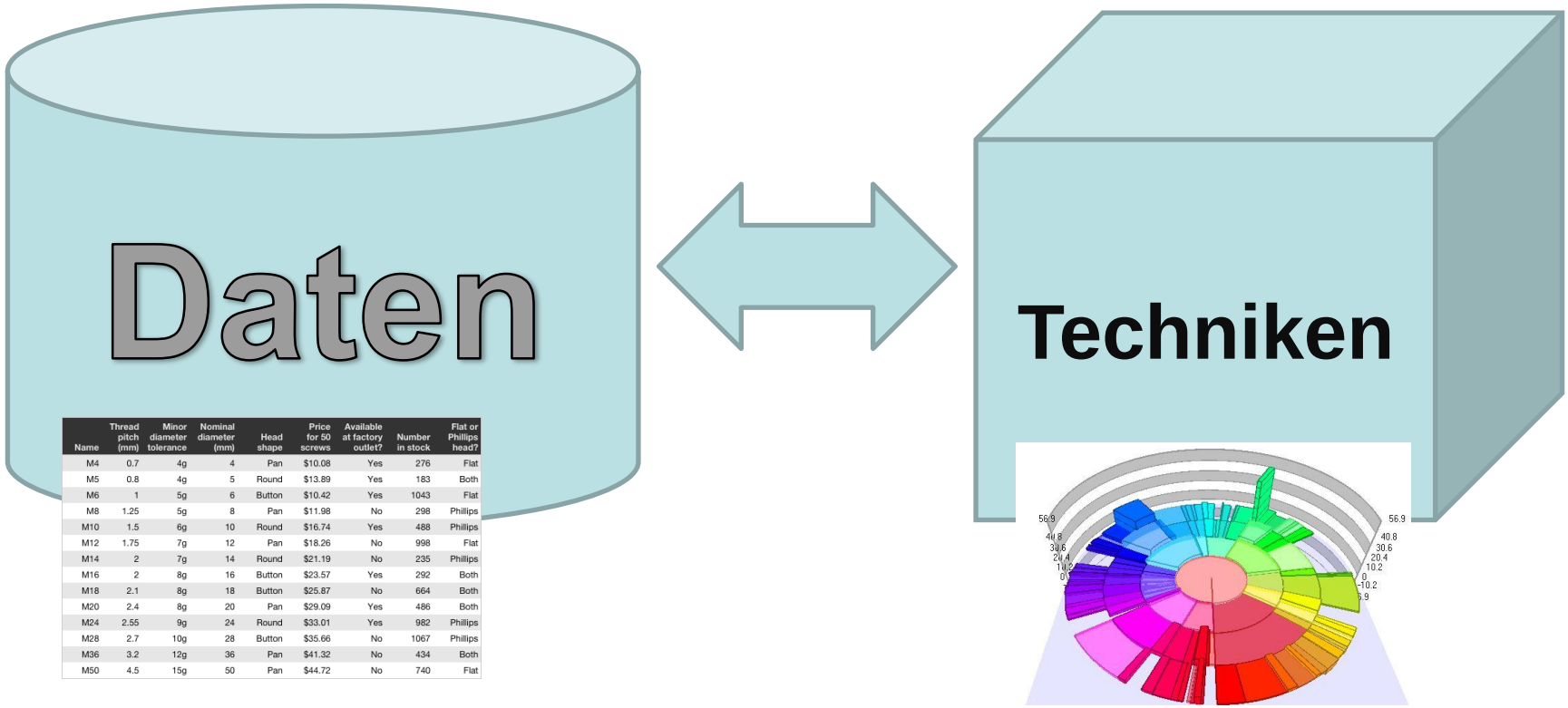
Datentypen: Überblick

- 1D
- Zeitreihen
- 2D
- 3D,… multi-dimensionale/multivariate Daten
- Hierarchien/Bäume
- Netzwerke/Graphen

- Geographische Daten, sonstige Daten,…

Informationsvisualisierung
Techniken und Datentypen

TECHNISCHE UNIVERSITÄT DARMSTADT



Name	Thread pitch (mm)	Minor diameter tolerance	Nominal diameter (mm)	Head shape	Price for 50 screws	Available at factory outlet?	Number in stock	Flat or Phillips head?
M4	0.7	4g	4	Pan	$10.08	Yes	276	Flat
M5	0.8	4g	5	Round	$13.89	Yes	183	Both
M6	1	5g	6	Button	$10.42	Yes	1043	Flat
M8	1.25	5g	8	Pan	$11.98	No	298	Phillips
M10	1.5	6g	10	Round	$16.74	Yes	488	Phillips
M12	1.75	7g	12	Pan	$18.26	No	998	Flat
M14	2	7g	14	Round	$21.19	No	235	Phillips
M16	2	8g	16	Button	$23.57	Yes	292	Both
M18	2.1	8g	18	Button	$25.87	No	664	Both
M20	2.4	8g	20	Pan	$29.09	Yes	486	Both
M24	2.55	9g	24	Round	$33.01	Yes	982	Phillips
M28	2.7	10g	28	Button	$35.66	No	1067	Phillips
M36	3.2	12g	36	Pan	$41.32	No	434	Both
M50	4.5	15g	50	Pan	$44.72	No	740	Flat

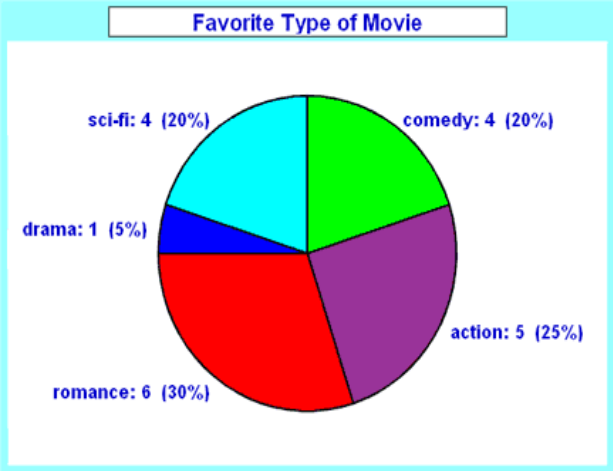
Datentyp: 1D Daten

- **1D Daten**

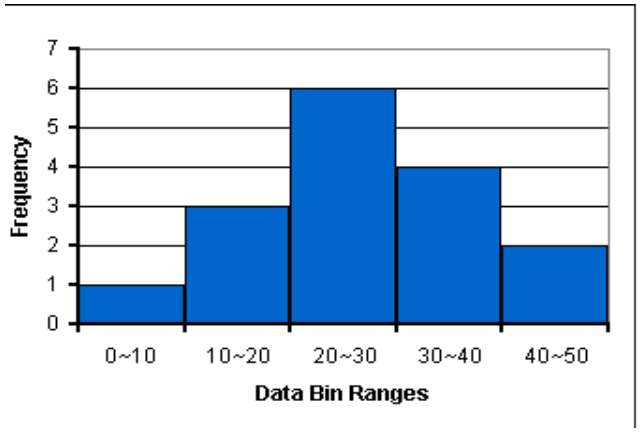
Bereich	Häufigkeit
0-10	1
10-20	3
20-30	6
30-40	4
40-50	2

Typische Visualisierungen:

Kuchendiagramm

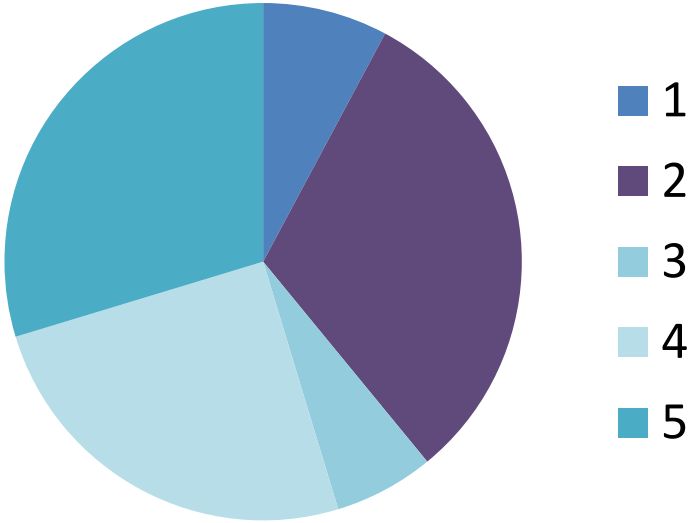


Balkendiagram



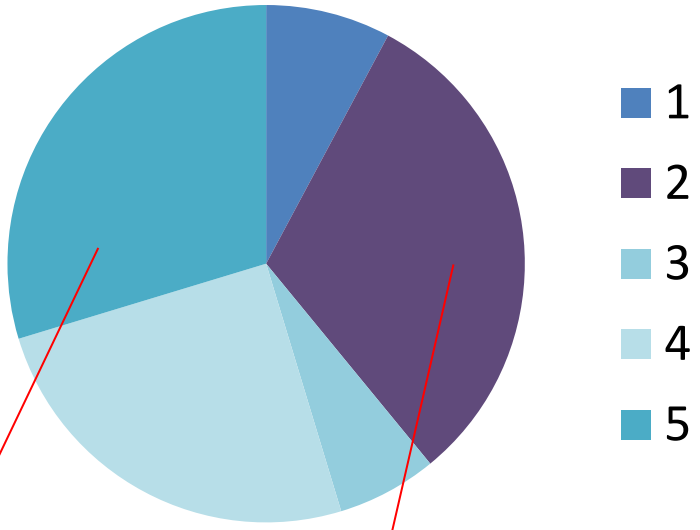
Datentypen: 1D Technik: Kuchendiagramm

- Kuchendiagramm
  - Beliebt für Anteiledaten
  - Werte werden als Größe abgebildet
    Größe der Teile – Winkel der Teile
  - Objekte als Farben abgebildet



Datentypen: 1D Technik: Kuchendiagramm

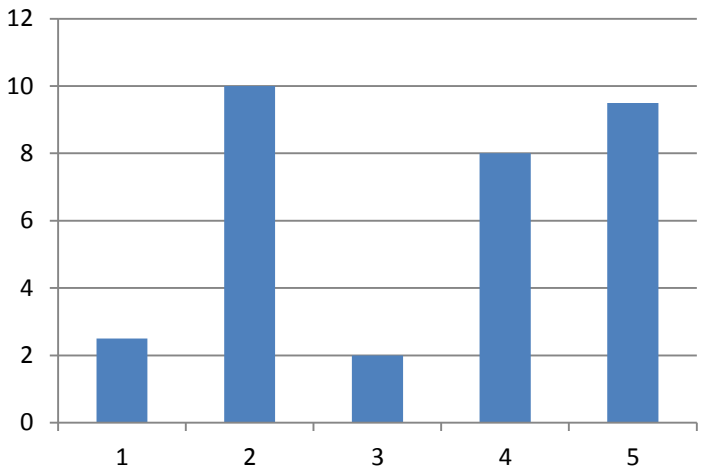
- Kuchendiagramm
  - Beliebt für Anteiledaten
  - Werte werden als Größe abgebildet
    Größe der Teile – Winkel der Teile
  - Man kann aber schlecht
    Wertunterschiede sehen



Was ist mehr (5) oder (2)?

Datentypen: 1D Technik: Balkendiagramm

- Balkendiagram
  - Gut zum Werte vergleichen und abbilden
  - Werte werden als Größe abgebildet



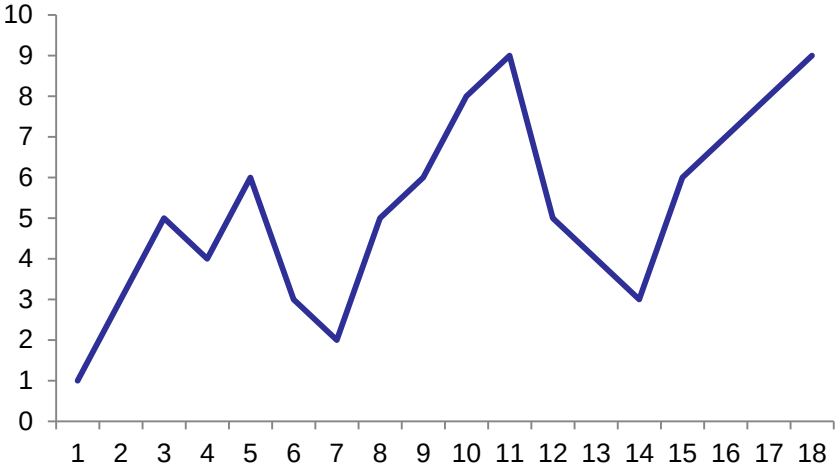
Datentyp: Zeitreihen

Technik: Liniengraphik

- Visualisierung von zeitbezogenen Daten
  - Meteorologie
  - Netzwerkmonitoring
  - Finanzen
  - …

Zeit	Wert
1.1.2012	35
2.1.2012	40
3.1.2012	70
...	..
30	30

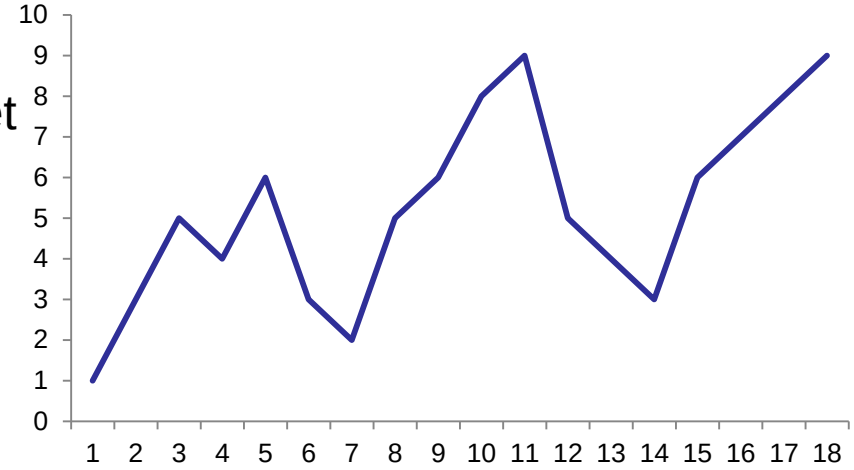
- Visualisierung: Liniengraphik



Datentyp: Zeitreihen Technik: Liniengraphik

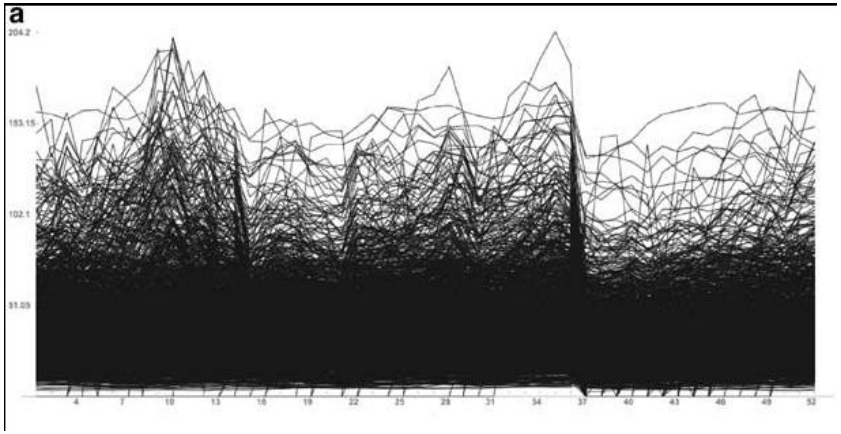
- Visualisierung: Liniengraphik

- Werte werden an die Position abgebildet (Y achse) und mit Linien verbunden



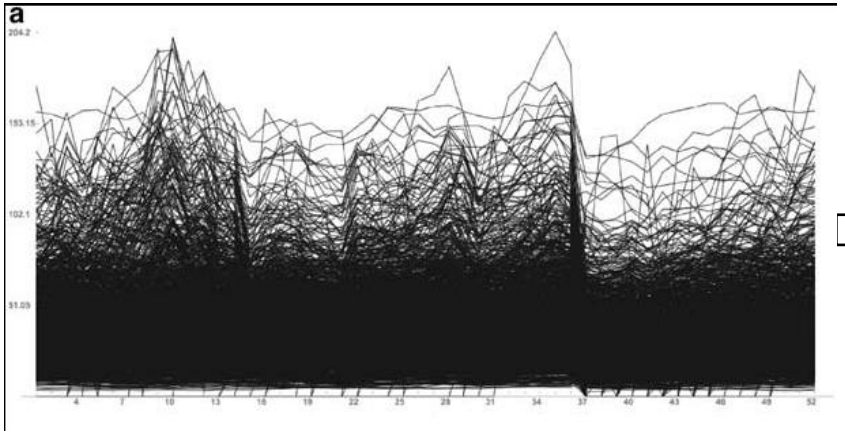
Datentyp: Zeitreihen Technik: Liniengraphik
Problem 1: Viele Zeitreihen

- Visualisierung von zeitbezogenen Daten
- Klassische Visualisierung: Liniengraphik
- Probleme:
  - Viele Zeitreihen

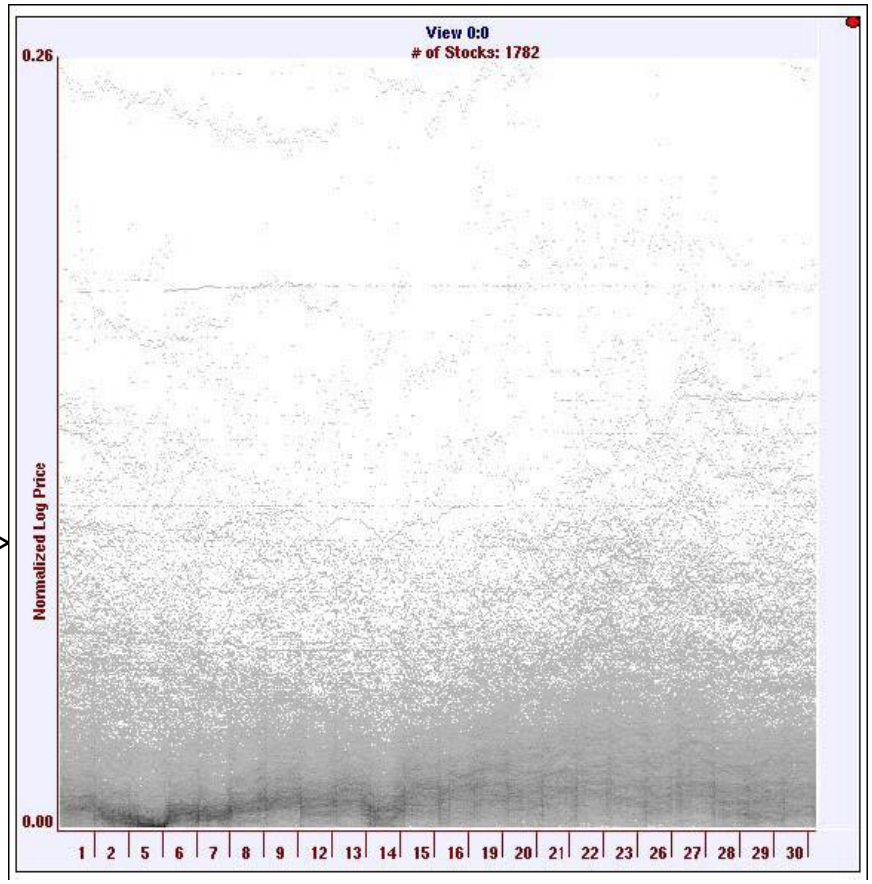


Datentyp: Zeitreihen Technik: Liniengraphik
Problem 1: Viele Zeitreihen

- Visualisierung von zeitbezogenen Daten
- Klassische Visualisierung: Liniengraphik
- Probleme:
  - Viele Zeitreihen



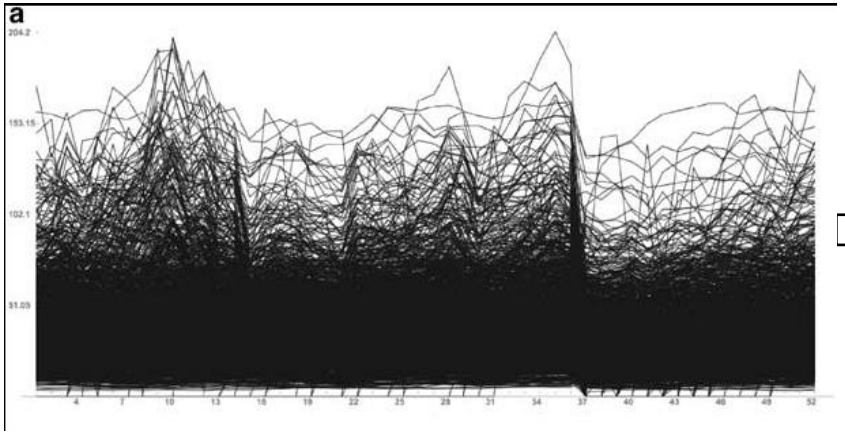
Lösung 1: Dichte

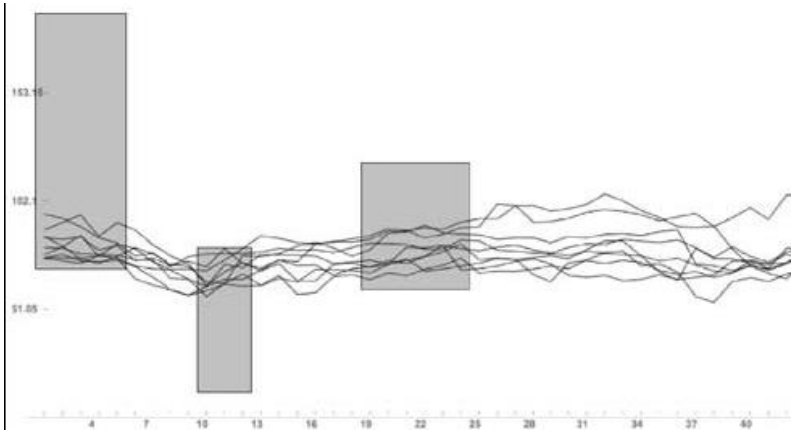


Datentyp: Zeitreihen Technik: Liniengraphik
Problem 1: Viele Zeitreihen

- Visualisierung von zeitbezogenen Daten
- Klassische Visualisierung: Liniengraphik
- Probleme:
  - Viele Zeitreihen

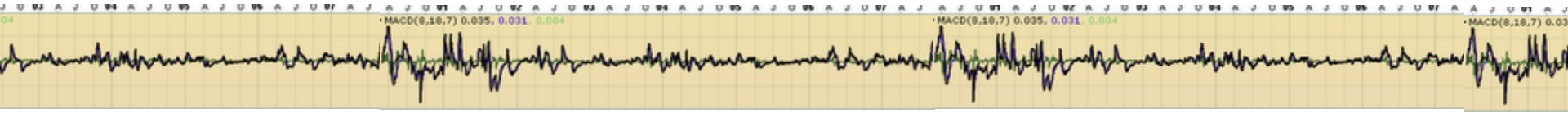
Lösung 2: Wertefilter





Datentyp: Zeitreihen Technik: Liniengraphik
Problem 2: Länge

- Visualisierung von zeitbezogenen Daten
- Klassische Visualisierung: Liniengraphik
- Probleme:
  - Lange Zeitreihen



Datentyp: Zeitreihen Technik: Liniengraphik
Problem 2: Länge

- Visualisierung von zeitbezogenen Daten
- Klassische Visualisierung: Liniengraphik
- Probleme:
  - Lange Zeitreihen

Lösung: Zeitfilter

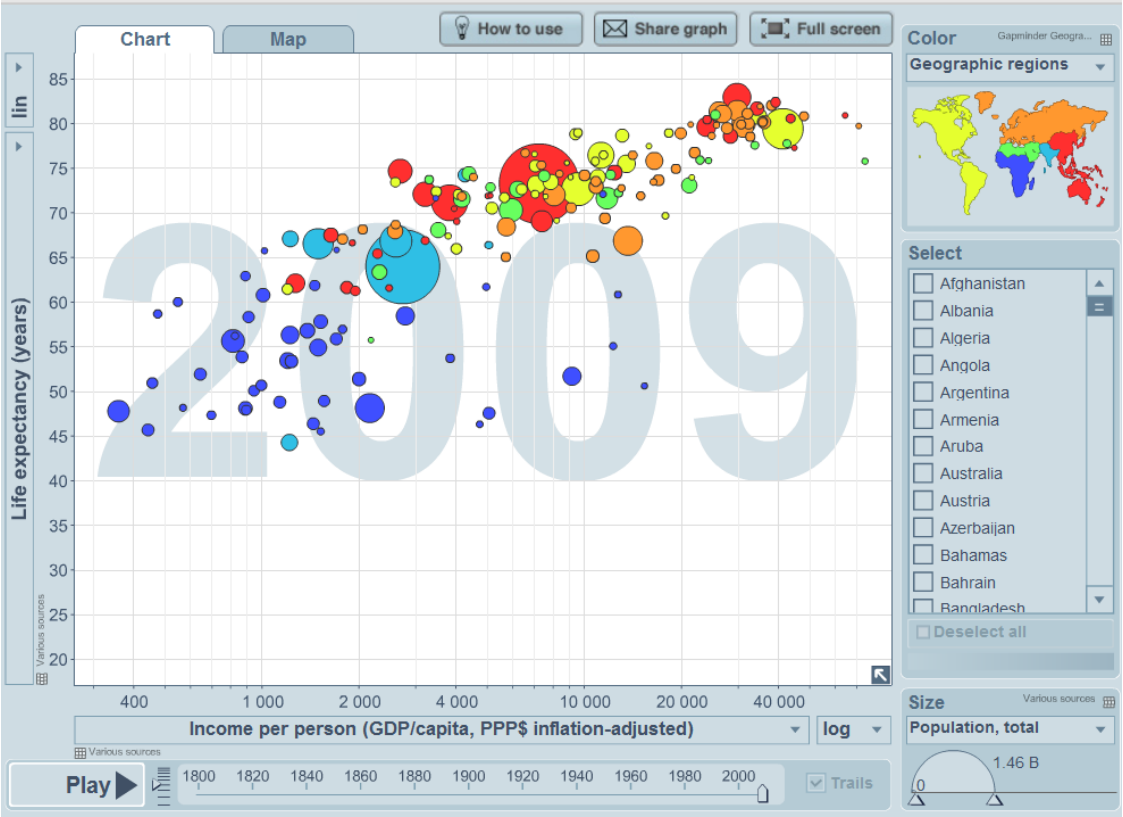
Datentyp: 2D, 3D

Technik: Scatterplot

- Datentyp: 2D, 3D

- Oft suche nach Korrelationen
  - Finanzen
  - Biologie
  - Etc.

Technik:Scatterplot

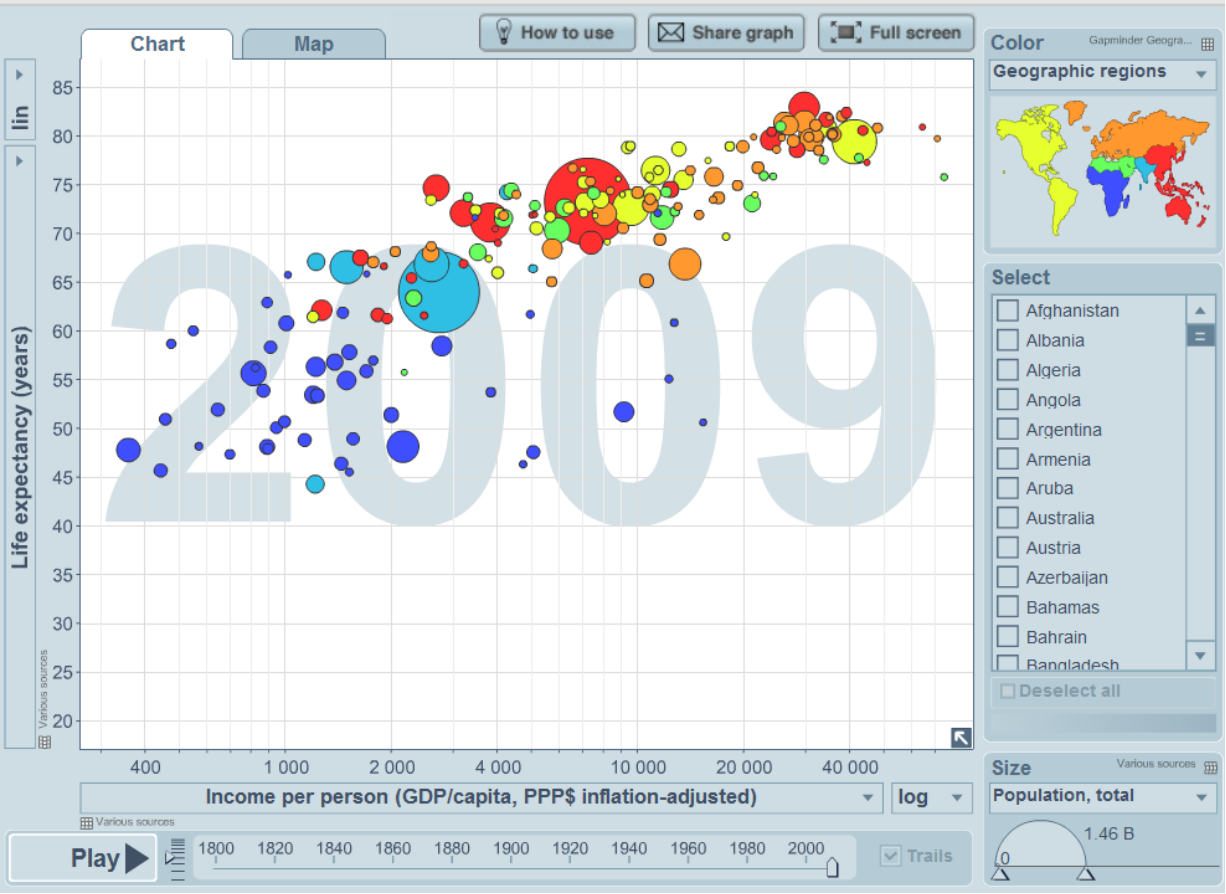


Datentyp: 2D, 3D

Technik: Scatterplot

- Daten werden abgebildet auf:
  - Position X Achse
  - Position Y Achse
  - (Größe)
  - (Farbe)

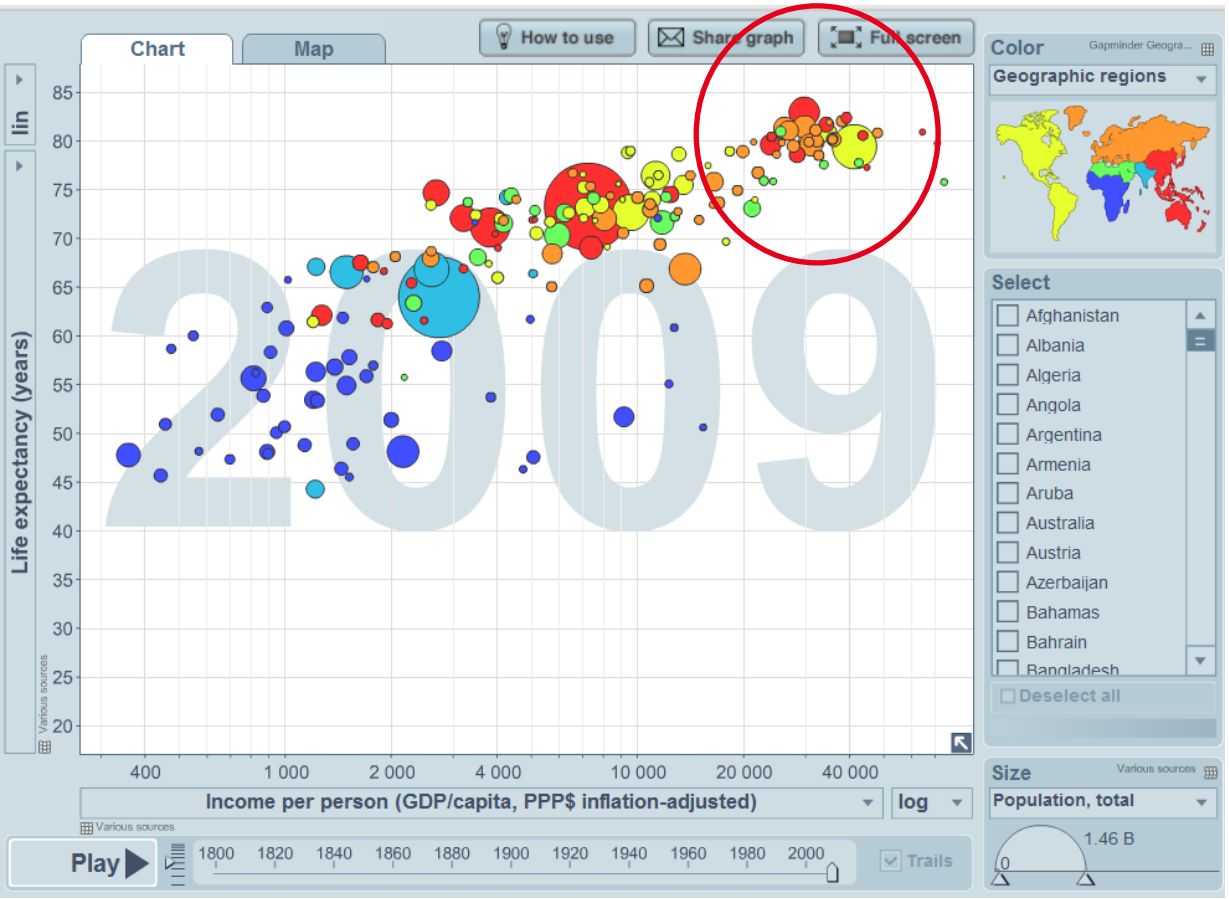
- Scatterplot
  - Intuitiv
  - Leicht lesbar



Datentyp: 2D, 3D
Problem: Overplotting

Technik: Scatterplot

- Probleme:
  - Overplotting



Datentyp: Multidimensionale Daten

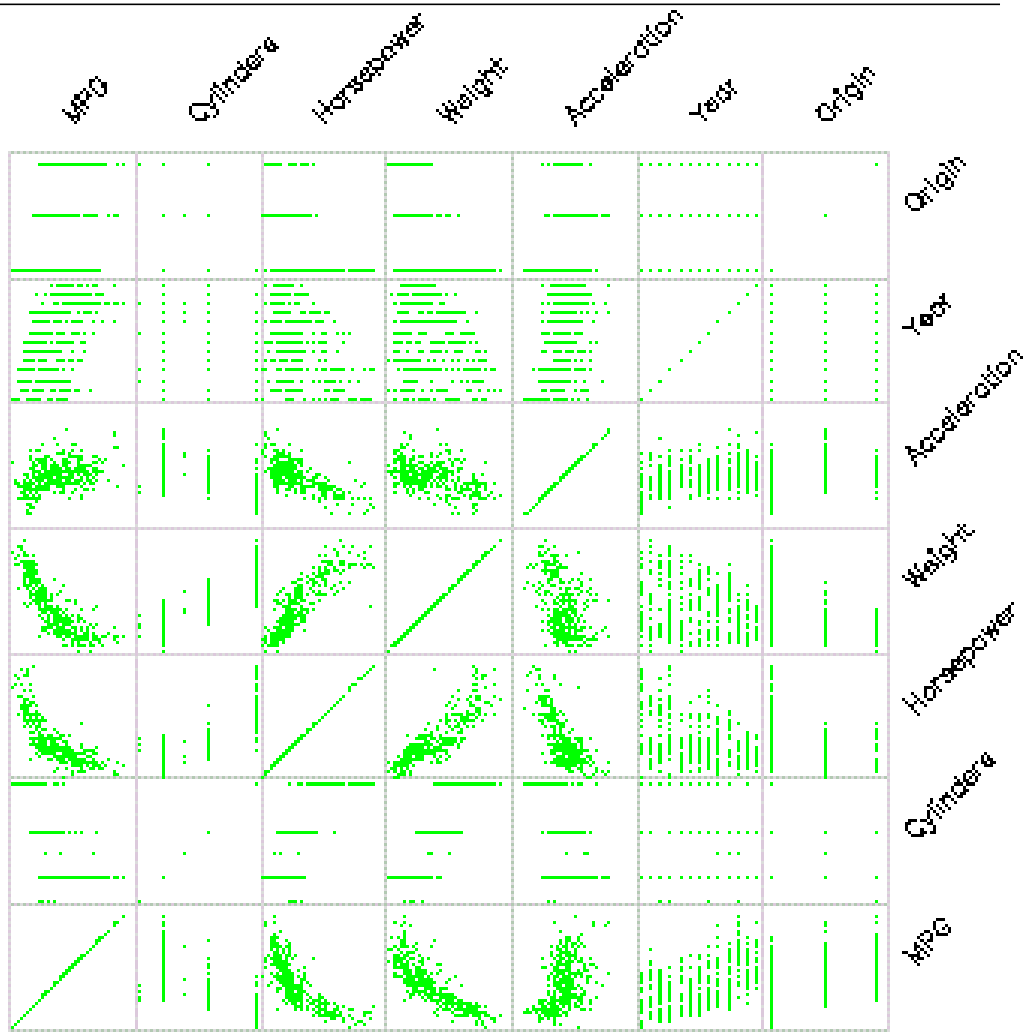
- Multidimensionale (multivariate) Daten
  - Erhebungen, Befragungen (Census)
  - Produktdaten, Marketingdaten

FZDNR	FNAME	NOTIERUI	PLZ	ORT	FTYP	NACE_CO	NACE_TXTE	BRA_COD	BESCHAE	UMSATZ	BILANZ	KAPITAL	GRUEND	RECHT
215078825	kassel tourist gmbh	ja	34117	Kassel	FIRMA	63301	Travel agencie	63301	24	1722	0	500	0	1
215163619	Augsburg Aktiengese	ja	86159	Augsburg	VERBAND	75130	Regulation of a	75130	15	0	0	125	2002	2
215273798	SGS-ICS Gesellscha	ja	20095	Hamburg	FIRMA	74301	Technical testi	74301	30	5770	0	26	1993	1
215273822	T?V S?D Product Se	ja	80339	M?nchen	FIRMA	74301	Technical testi	74301	300	37000	0	9300	1988	1
215273855	VdS Schadenverh?tu	ja	50735	K?ln	FIRMA	74301	Technical testi	74301	269	26230	0	2500	1995	1
215497397	T?V Nord MPA Gese	ja	6237	Leuna	FIRMA	74301	Technical testi	74301	90	5000	0	510	1993	3
215500307	T?V S?D Auto Servic	X	70794	Filderstadt	VERBAND	74301	Technical testi	74301	2933	280700	0	7000	1997	1
215541855	CFB Commerz Fond	nein	40470	D?sseldorf	FIRMA	65231	Investment cor	65231	0	36454	11000	26	0	1
215562398	CommerzLeasing Mo	nein	40470	D?sseldorf	FIRMA	71405	Renting of othe	71405	0	449260	880500	5139	1997	1
315000067	Philips Technologie (	X	20099	Hamburg	FIRMA	29710	Manufacture of	29710	10800	4.30E+06	0	587986	0	1
315000208	Kodak Verwaltung G	X	70327	Stuttgart	FIRMA	25241	Manufacture of	25241	0	0	0	111462	1927	1
315000273	Blohm + Voss Holdin	NEIN	20457	Hamburg	FIRMA	35111	Shipbuilding	35110	1019	171700	0	48573	0	1
315000331	ZF Sachs AG	ja	97424	Schweinfur	FIRMA	29140	Manufacture of	29140	16902	2.23E+06	0	170000	0	2
315000414	ITT Automotive Euro	nein	60489	Frankfurt a	FIRMA	34300	Manufacture of	34300	20	16400	44300	10200	1906	3
315000422	RWE Power Aktieng	nein	50935	K?ln	FIRMA	40108	Wind power ge	40115	13443	6.48E+06	1.72E+07	1.03E+06	1908	2
315000455	TKV Telekommunika	nein	10997	Berlin	FIRMA	32200	Manufacture of	32200	964	188890	144000	85000	0	3
315000489	mg verm?gensverwa	nein	60439	Frankfurt a	FIRMA	70320	Management o	70320	14	21700	130100	47321	1930	1
315001164	August Br?tje GmbH		26180	Rastede	FIRMA	28220	Manufacture of	28220	400	125000	0	5120	1919	1

- Problem: viele Dimensionen darstellen auf 2D Monitor
- Scatterplot nur limitierte Anzahl von Dimensionen

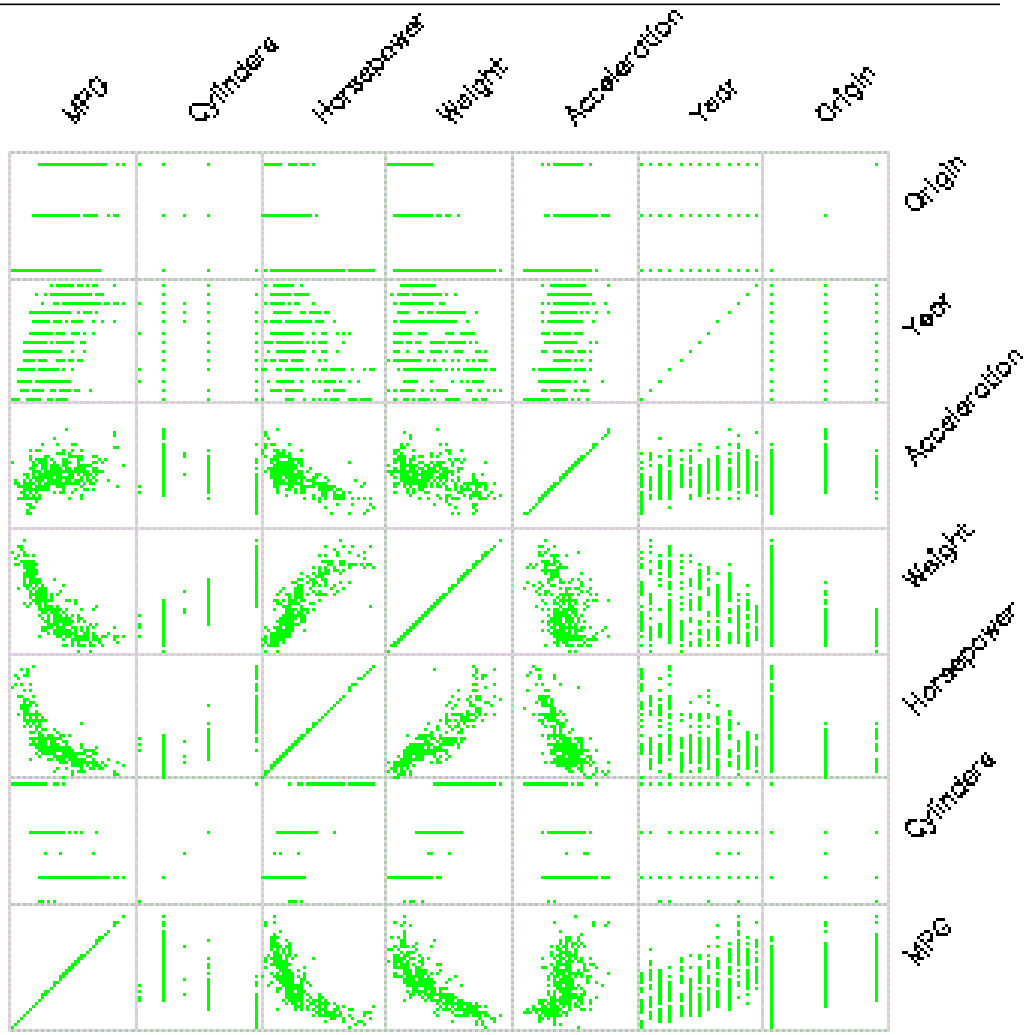
Datentyp: Multidimensionale Daten
Technik: Scatterplotmatrix

- Scatterplot Matrix
  - Alle Paare der Dimensionen als Scatterplots
  - Scatterplots in Matrix nach abgebildeten Dimensionen angeordnet



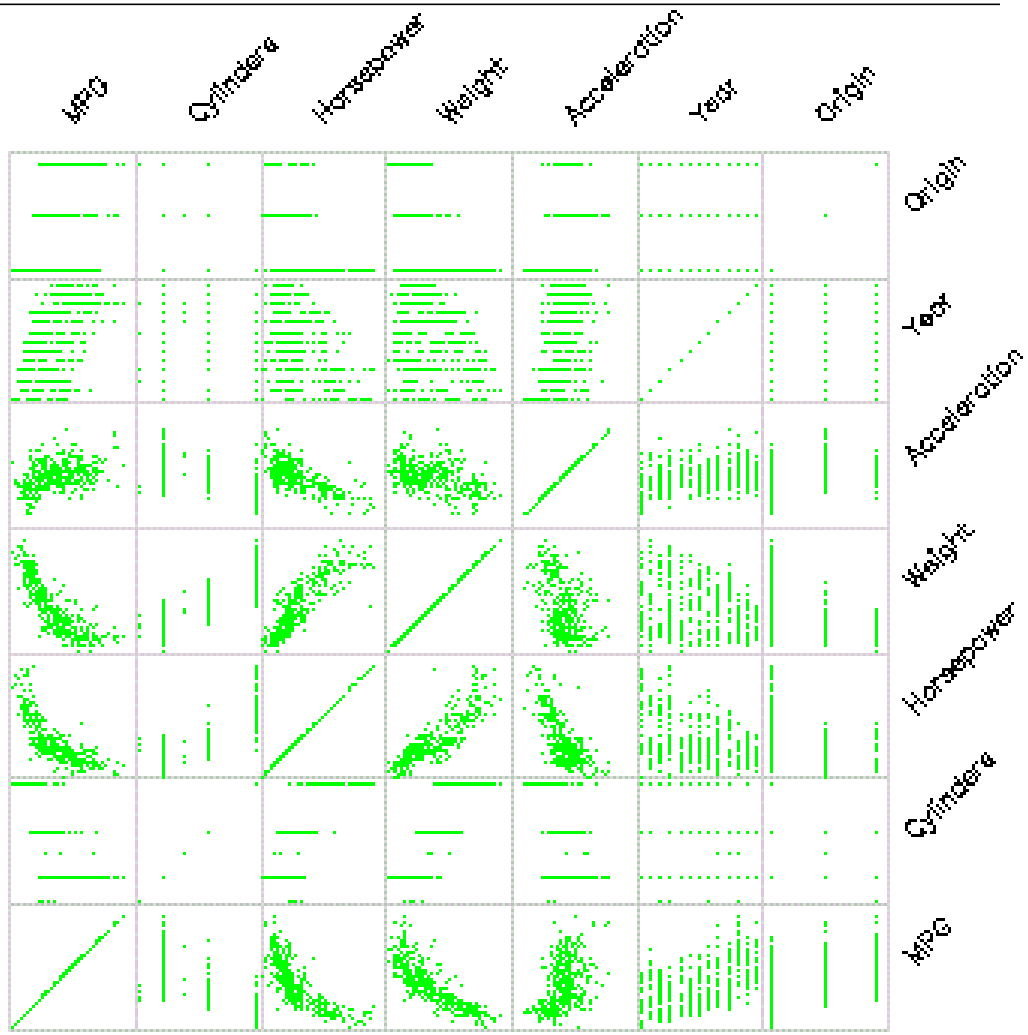
Datentyp: Multidimensionale Daten
Technik: Scatterplotmatrix

- Scatterplot Matrix
  - Gut für paarweise Korrelationen/
  - Abhängigkeiten
  - Beliebt in Statistik



Datentyp: Multidimensionale Daten
Technik: Scatterplotmatrix

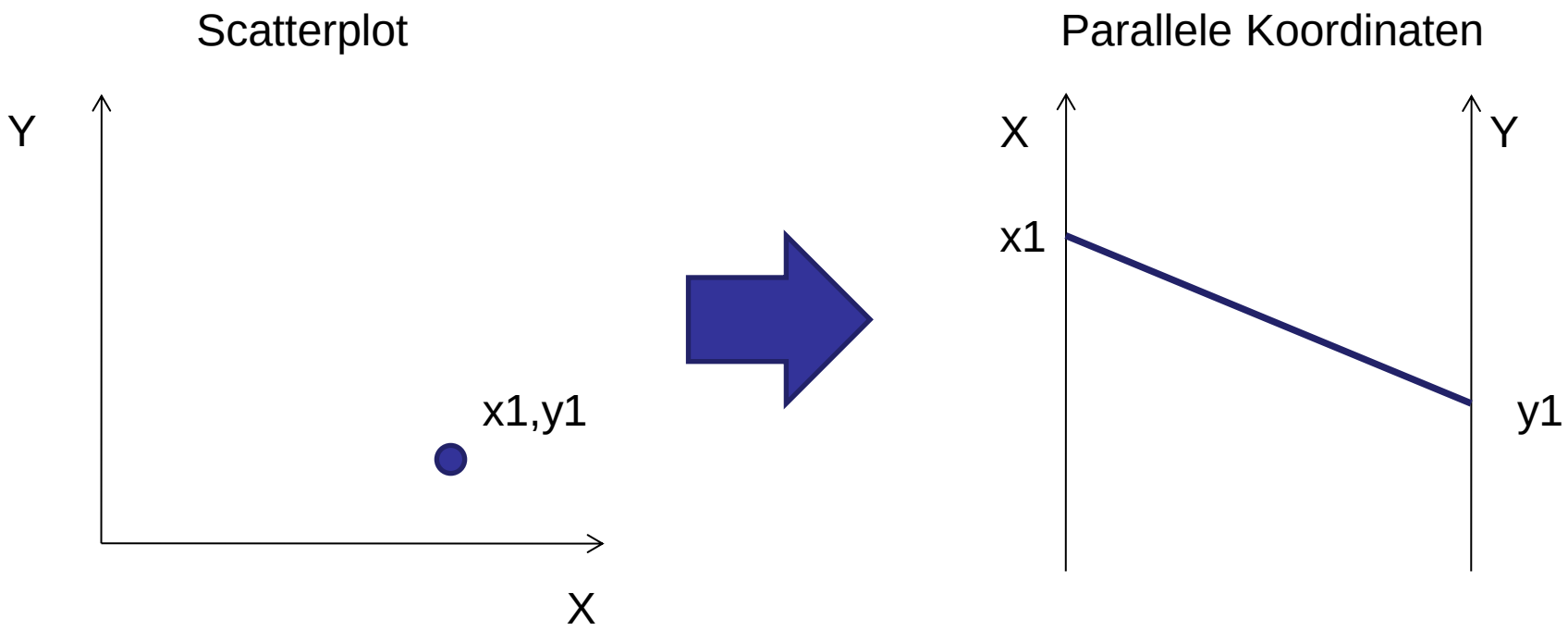
- Scatterplot Matrix
  - Problem:
    - Viele Dimensionen
    - Limitierter Platz für einzelne Scatterplots
    - Nur paarweise Abhängigkeiten sichtbar



Datentyp: Multidimensionale Daten
Technik: Parallele Koordinaten

Idee: Koordinaten parallel zueinander

- Aus Punkt im Scatterplot wird eine Linie in parallelen Koordinaten

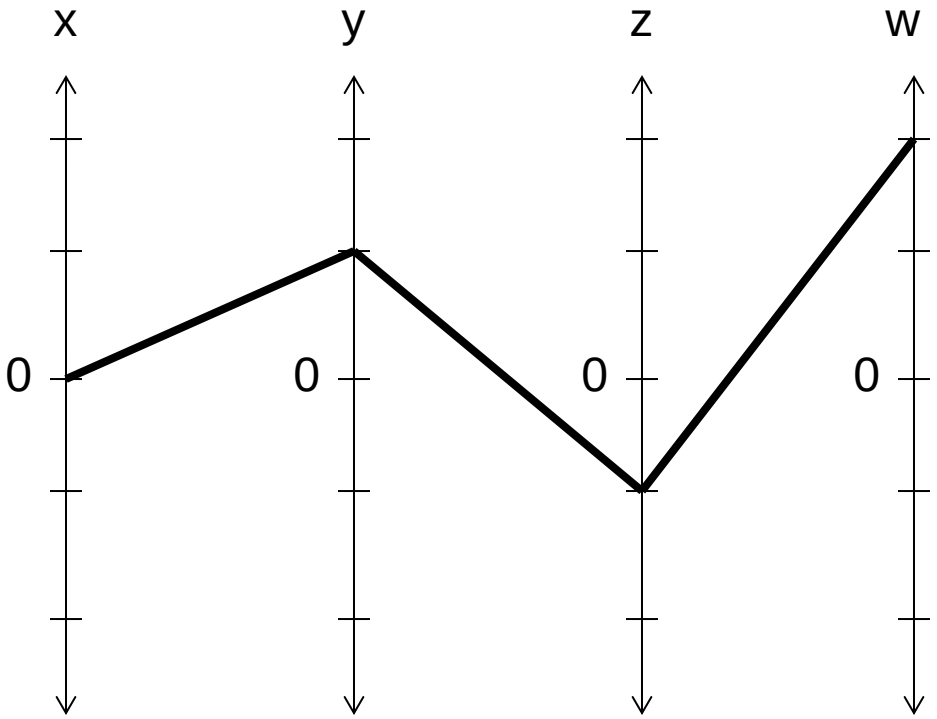


Datentyp: Multidimensionale Daten
Technik: Parallele Koordinaten

Idee: Koordinaten parallel zueinander

- Aus Punkt im Scatterplot wird eine Linie in parallelen Koordinaten
- Mehrere Dimensionen nebeneinander abgebildet

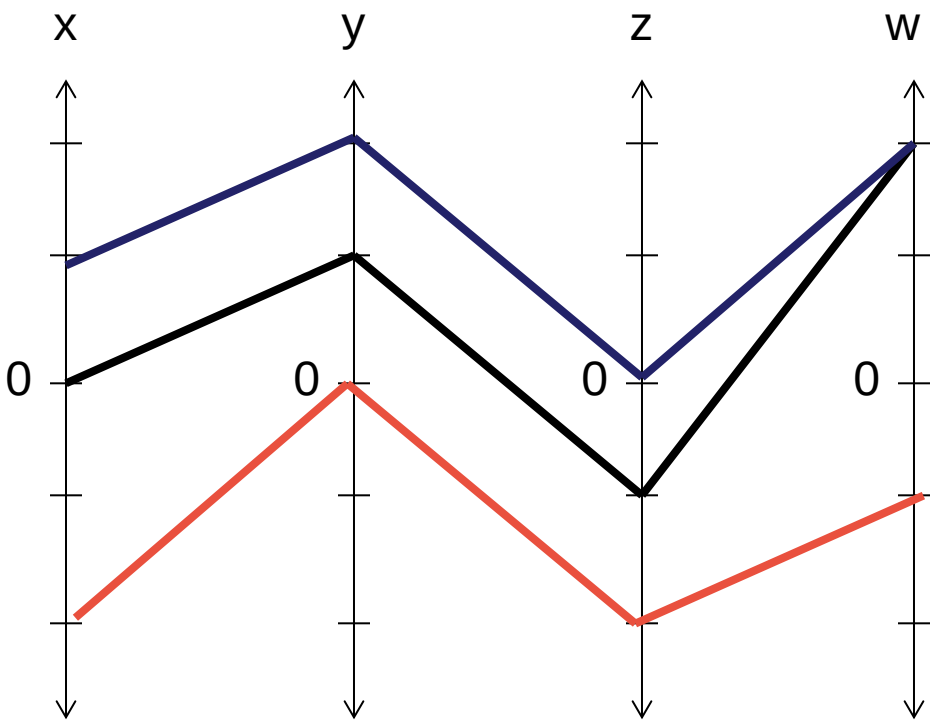
- (0,1,-1,2)=



Datentyp: Multidimensionale Daten
Technik: Parallele Koordinaten

- Parallele Koordinaten

  - (0,1,-1,2)
  - (1,2,0,2)
  - (-2,0,-2,-1)

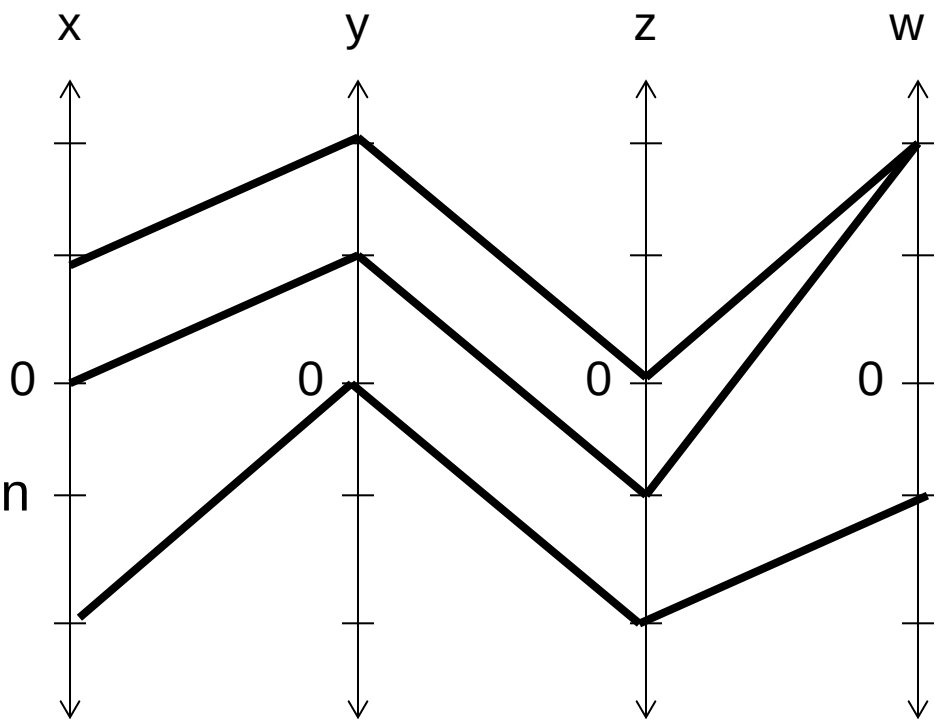


Datentyp: Multidimensionale Daten
Technik: Parallele Koordinaten

- Parallele Koordinaten

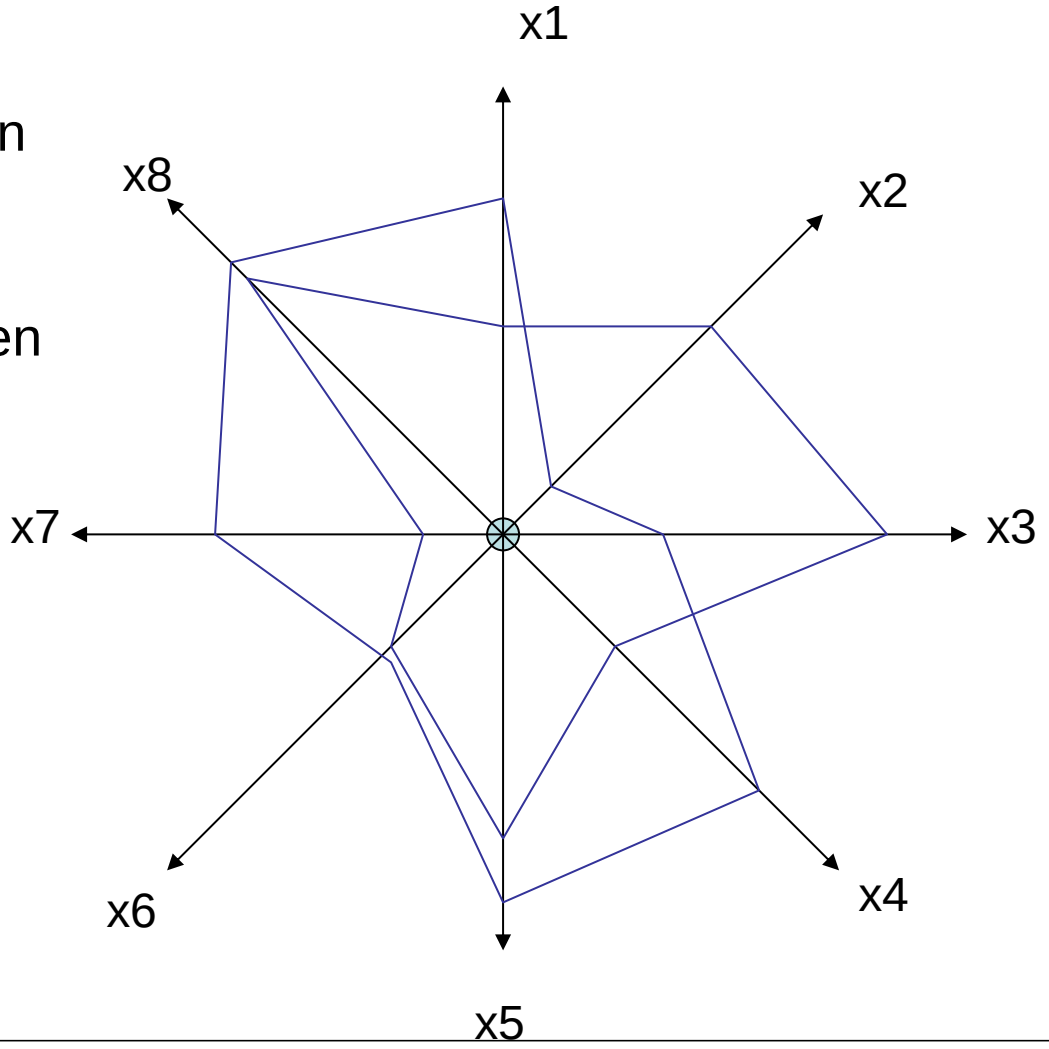
  - (0,1,-1,2)
  - (1,2,0,2)
  - (-2,0,-2,-1)

  - Abhängigkeiten zwischen mehreren Dimensionen leicht sichtbar



Datentyp: Multidimensionale Daten
Technik: Star Plot

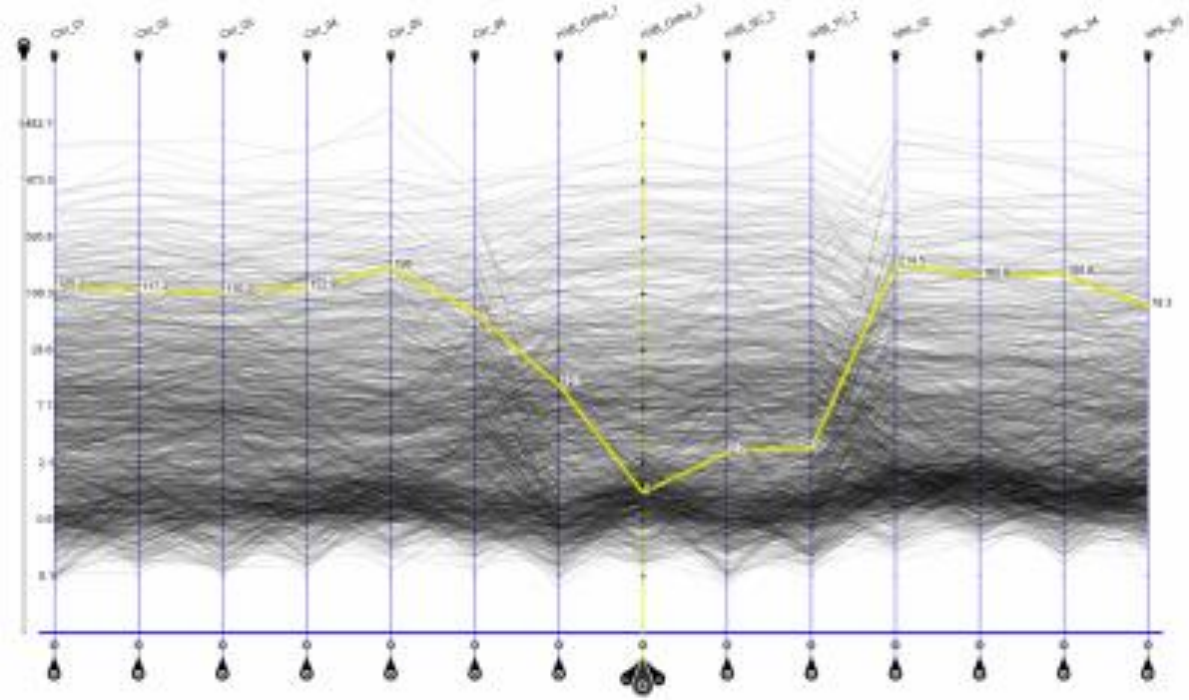
- Star Plot
  - Mehrere Dimensionen im Kreis abgebildet
  - Ähnlich zu Parallelen Koordinaten
  - Beliebt in
    - Psychologie
    - Wirtschafts-wissenschaften



Datentyp: Multidimensionale Daten
Technik: Parallele Koordinaten
Problem: Overplotting

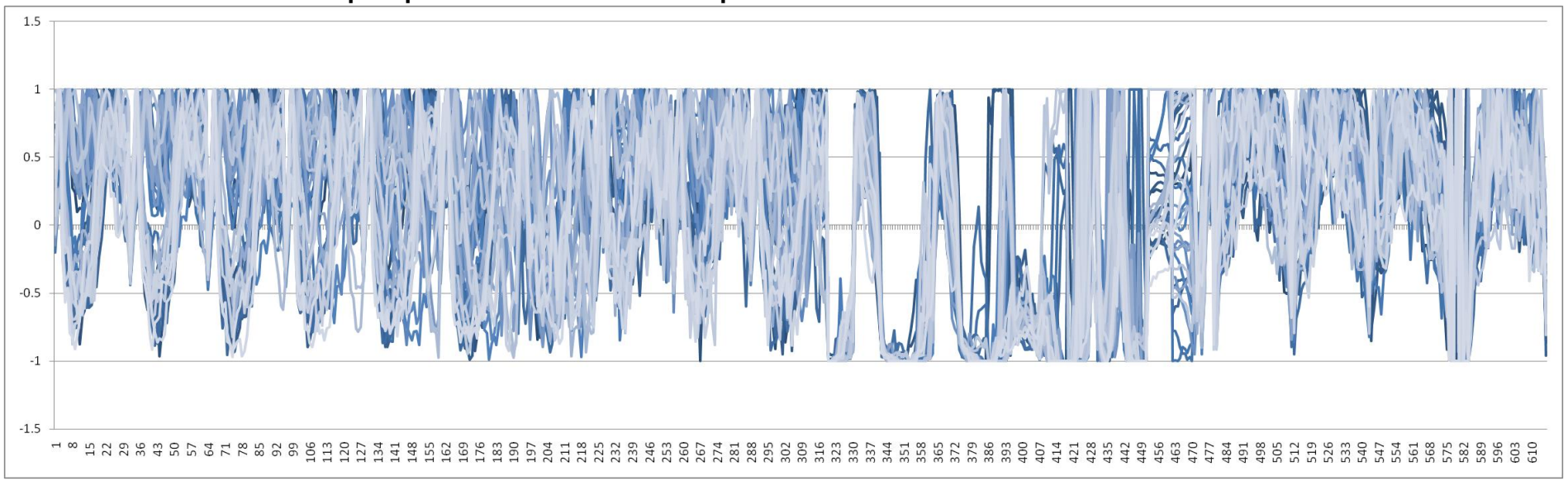
- Parallele Koordinaten

- Problem:
  - overplotting
  - Lösung:
    - Dichte
    - Transparenz



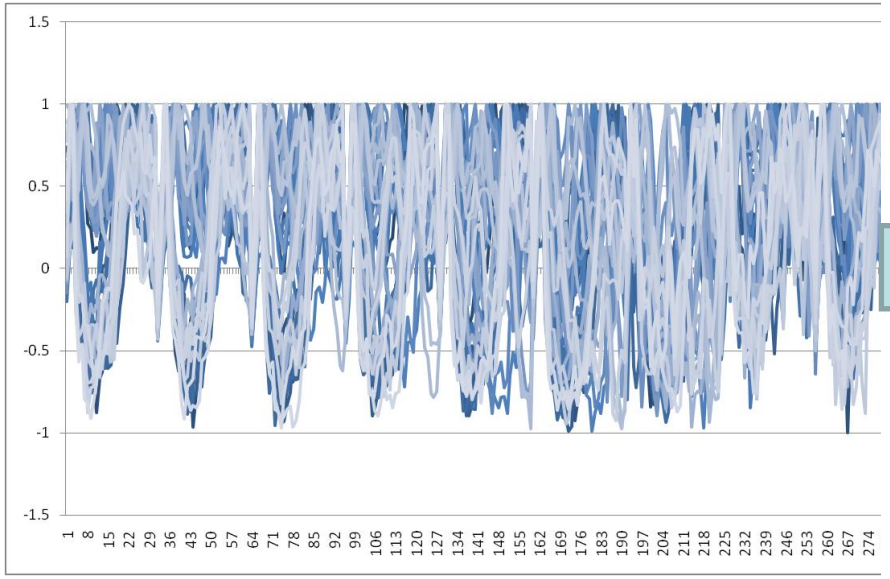
Datentyp: Multidimensionale Daten
Technik: Parallele Koordinaten
Problem: Viele Dimensionen

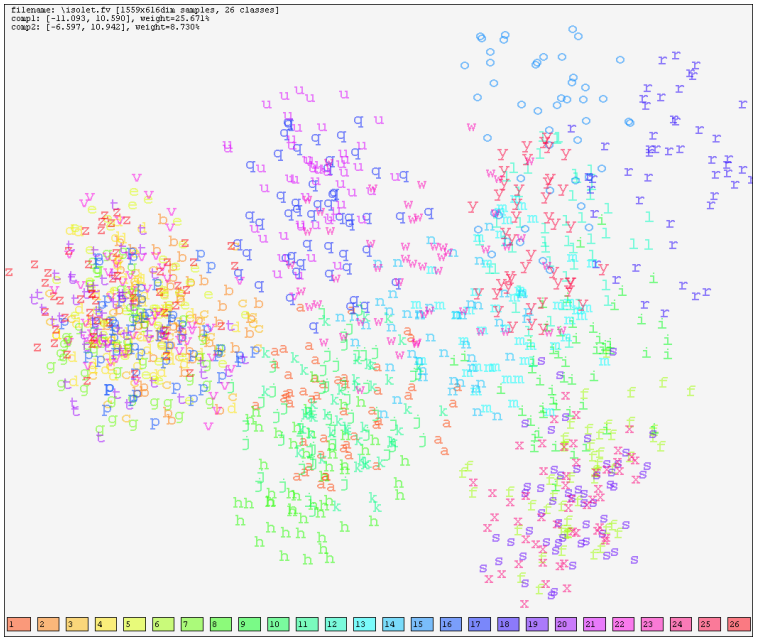
- **Problem von vielen Dimensionen**
  - Gesprochene Buchstaben – letter recognition
  - 1559 audio samples of the letters A to Z spoken by different persons (26 letters)
  - The samples are represented by 616-dimensional feature vectors encoding certain aural properties of the samples.



Datentyp: Multidimensionale Daten
Technik: Dimensionsreduktion + Scatterplot

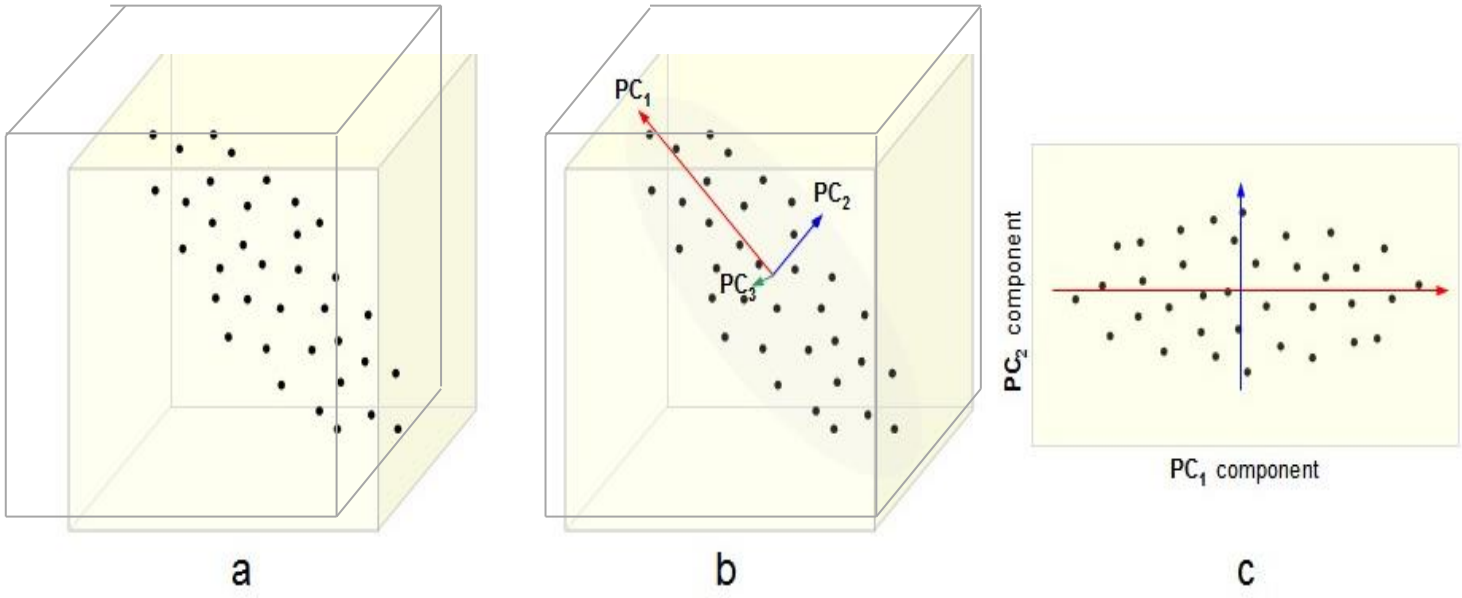
- **Alternative:**
  - **Dimensionsreduktion nD → 2D**
    - **z.B. durch principal component analysis (PCA)**
  - **Ergebnis als Scatterplot darstell**





FYI: Multidimensionale Daten – Principal Component Analysis

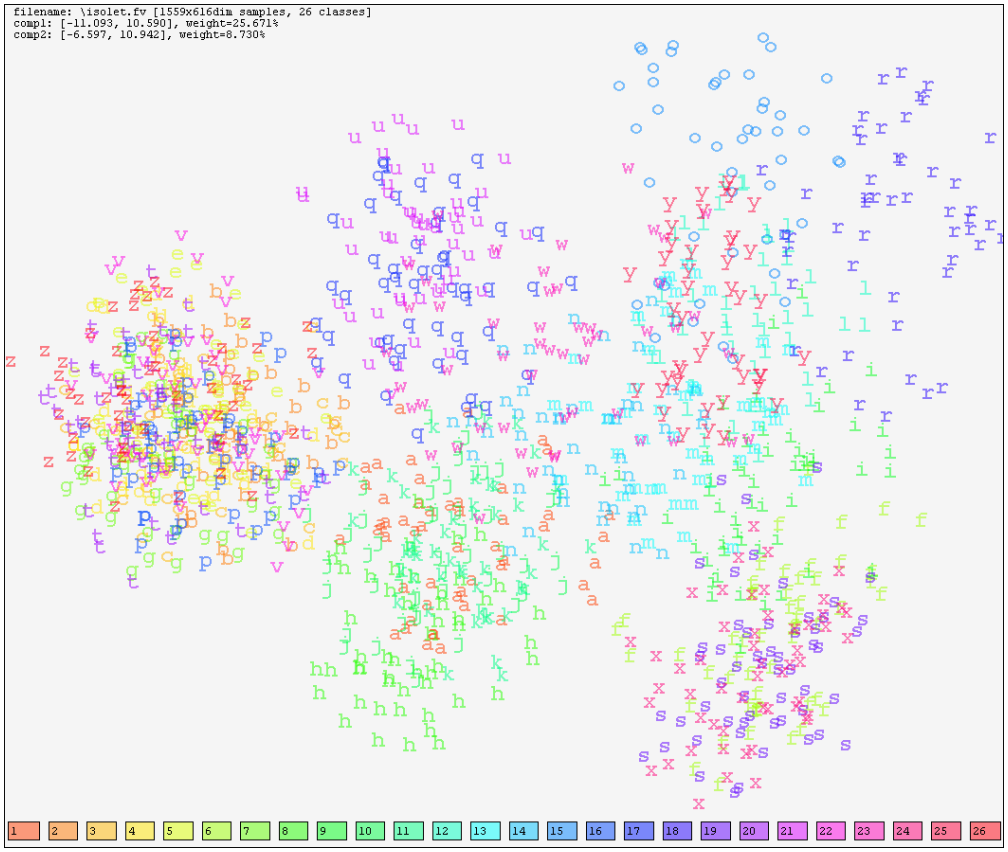
- Converts a set of observations of possibly correlated variables into a set of values of uncorrelated variables called **principal components**.
- The number of principal components is less than or equal to the number of original variables. → Projection to lower-dimensional space
- Orthogonal linear transformation



Datentyp: Multidimensionale Daten
Technik: Dimensionsreduktion + Scatterplot

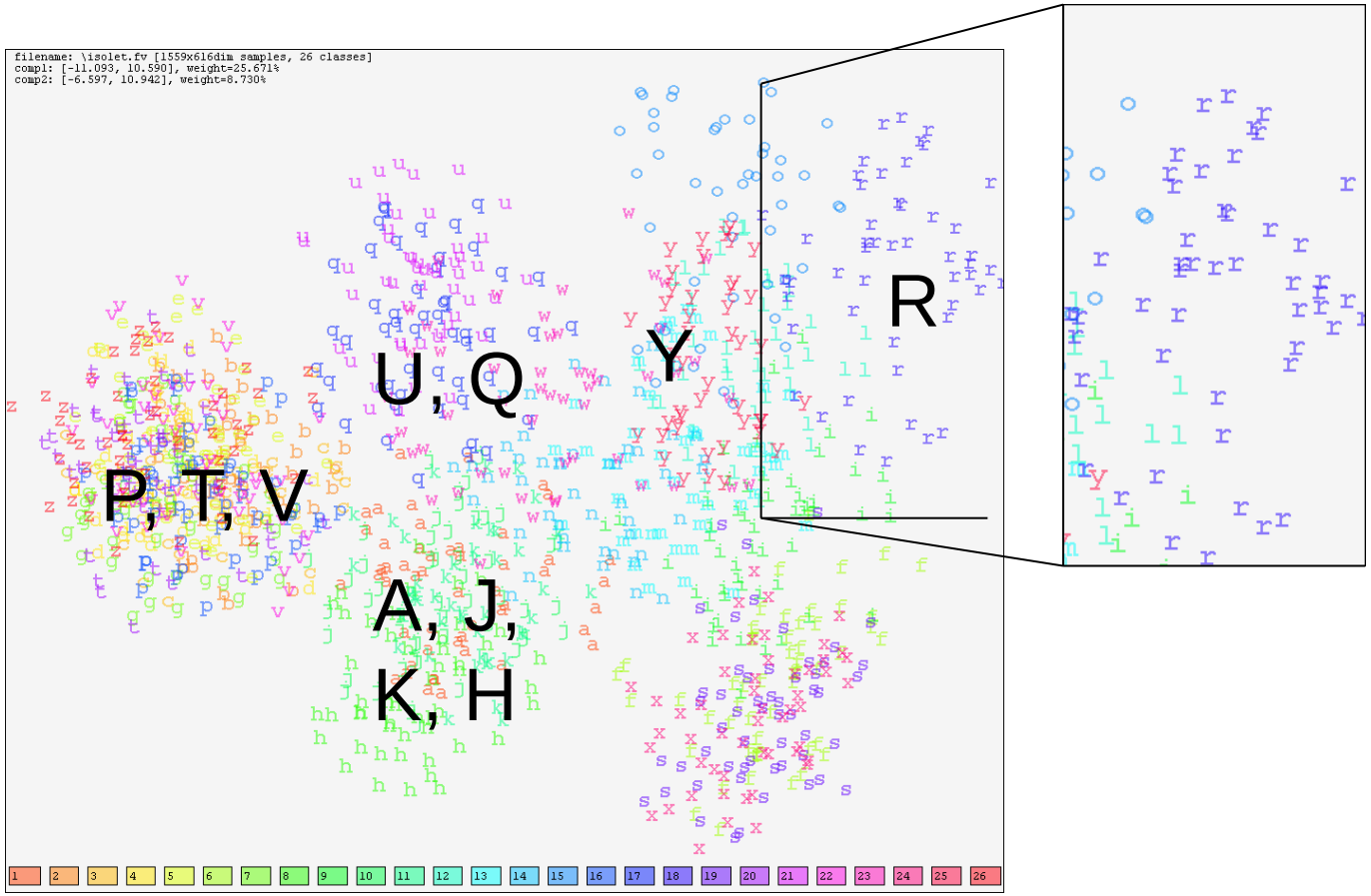
TECHNISCHE UNIVERSITÄT DARMSTADT

- Gesprochene Buchstaben (spoken letter recognition data set)
- PCA Projektion



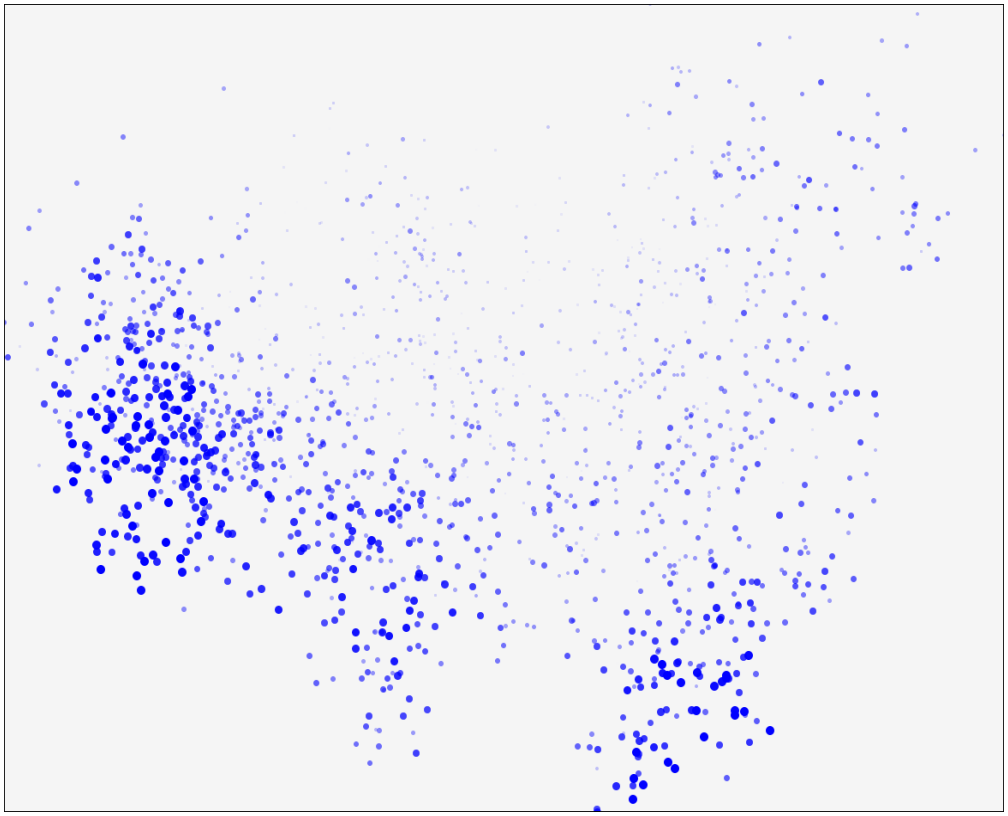
Datentyp: Multidimensionale Daten
Technik: Dimensionsreduktion + Scatterplot

- Gesprochene Buchstaben (spoken letter recognition data set)
- PCA Projektion



Datentyp: Multidimensionale Daten
Technik: Dimensionsreduktion + Scatterplot

- Gesprochene Buchstaben (spoken letter recognition data set)
- PCA Projektion → Informationsverlust → **Projektionsqualität**

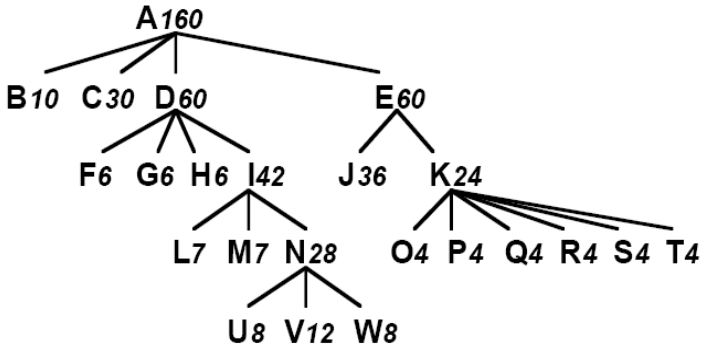


Datenyp: Hierarchien

- Hierarchien
  - Natürliche Art, Daten zu strukturieren bzw. zu organisieren
    - Unternehmensstruktur
    - Geographie
    - Produkt- und Teilekataloge
    - Dateisysteme
    - u.v.m.
  - **Wie visualisiert man hierarchische Daten?**

Datenyp: Hierarchien
Visualisierung: Node-Link Diagramm

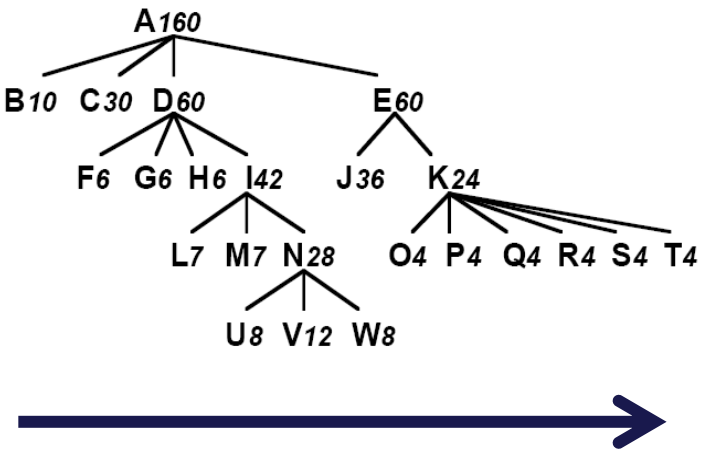
- **Klassische Verfahren**
  - Node-Link Diagramme
    - Jeder Knoten: 1 Punkt
    - Vater-Sohn Beziehung als Linie
    - Position der Knoten ist wichtig
      (z.B. Sohn unter Vater, Söhne nebeneinander)



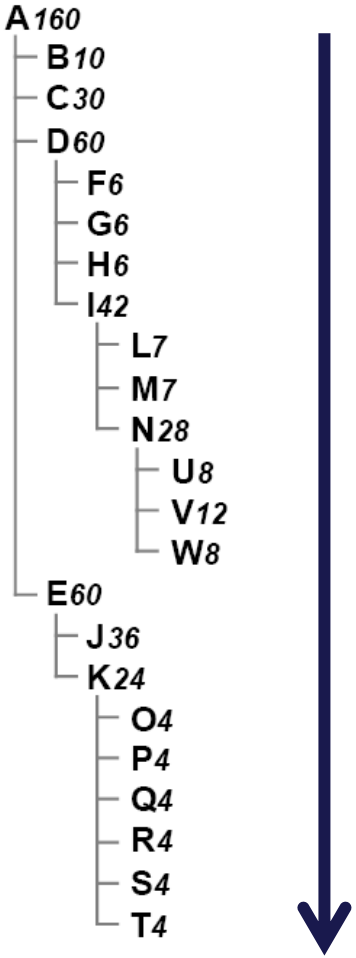
Datenyp: Hierarchien
Visualisierung: Node-Link Diagramm
Problem: Positionierung der Knoten (Layout)

- **Klassische Verfahren**
  - Node-Link Diagramme
- Probleme der klassischen Verfahren
  - Layoutprobleme

Datenyp: Hierarchien
Visualisierung: Node-Link Diagramm
Problem: Positionierung der Knoten (Layout)



Horisontal Layout



Vertikal Layout

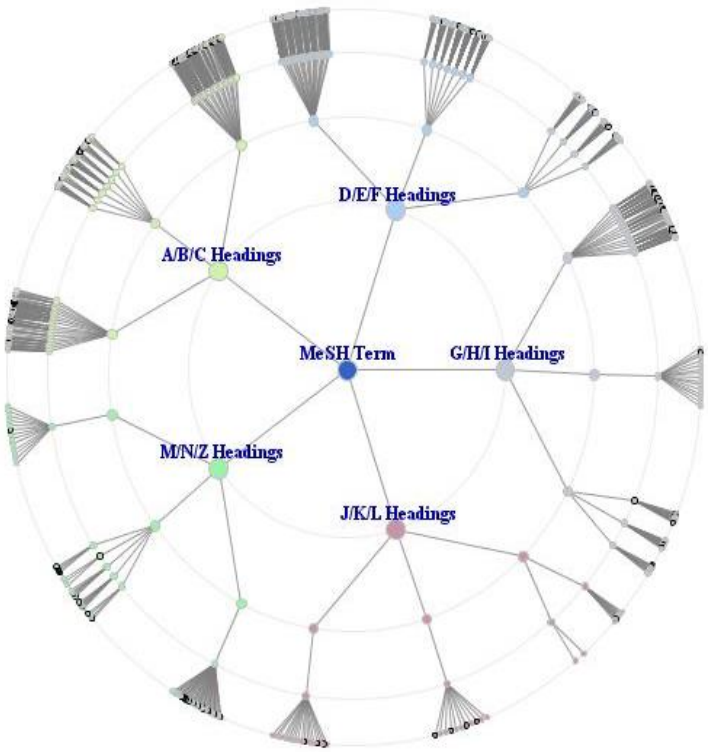
Lineare Anordnung:

Begrenzte Anzahl Knoten und Ebenen...

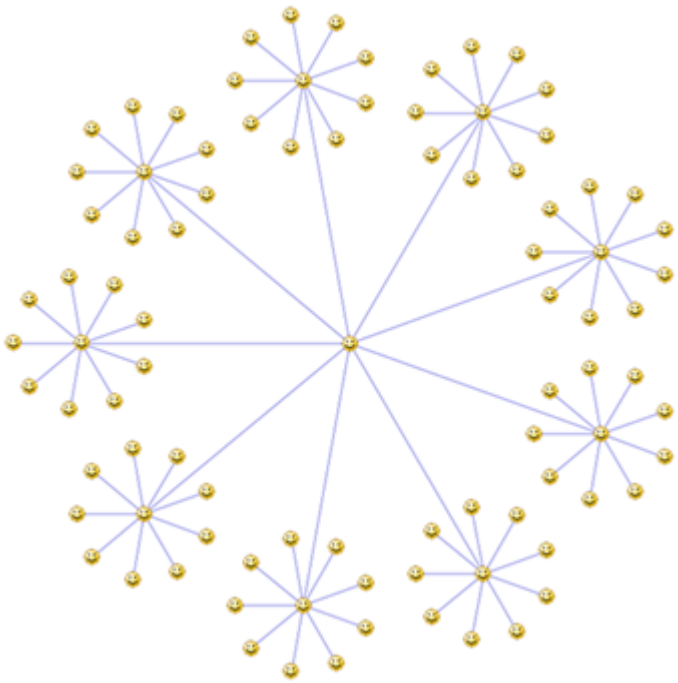
Datenyp: Hierarchien
Visualisierung: Node-Link Diagramm
Problem: Positionierung der Knoten (Layout)

Kreisangeordnete Knoten:

Bessere Platznutzung



Radial



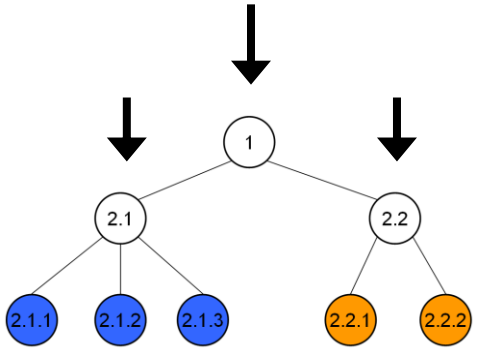
Baloon layout

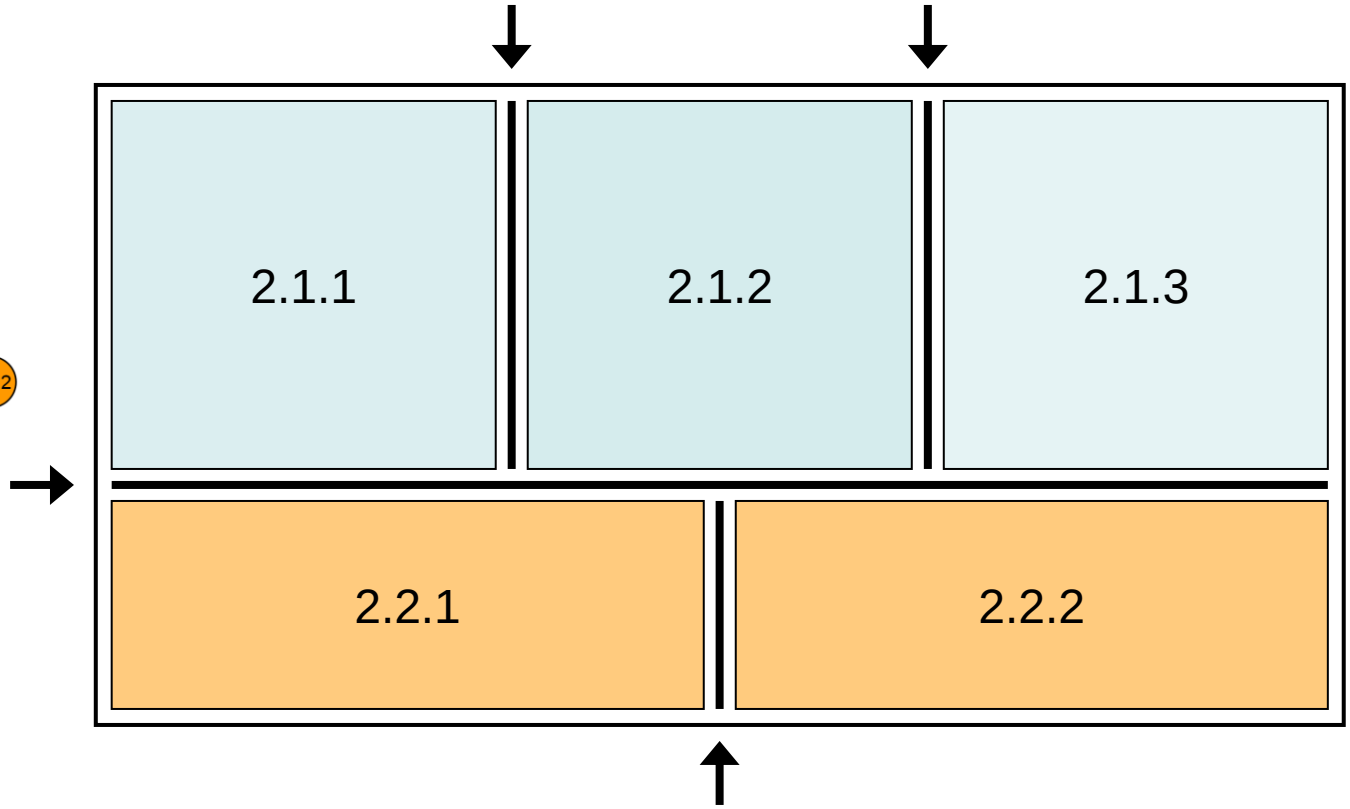
Datenyp: Hierarchien
Visualisierung: Node-Link Diagramm
Problem: Viele Knoten

- **Klassische Verfahren**
  - Node-Link Diagramme
- Problem der klassischen Verfahren
  - Skalierungsproblem

Datenyp: Hierarchien
Visualisierung: Treemap

Idee: Kindknoten **IM** Vaterknoten zeigen
Volle Platznutzung

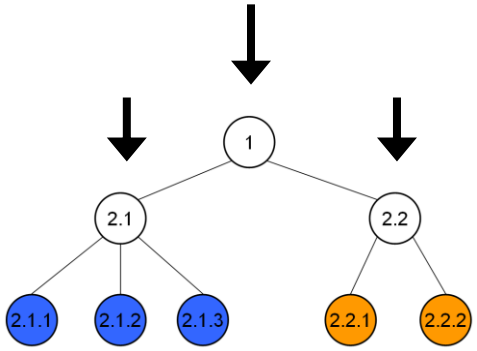


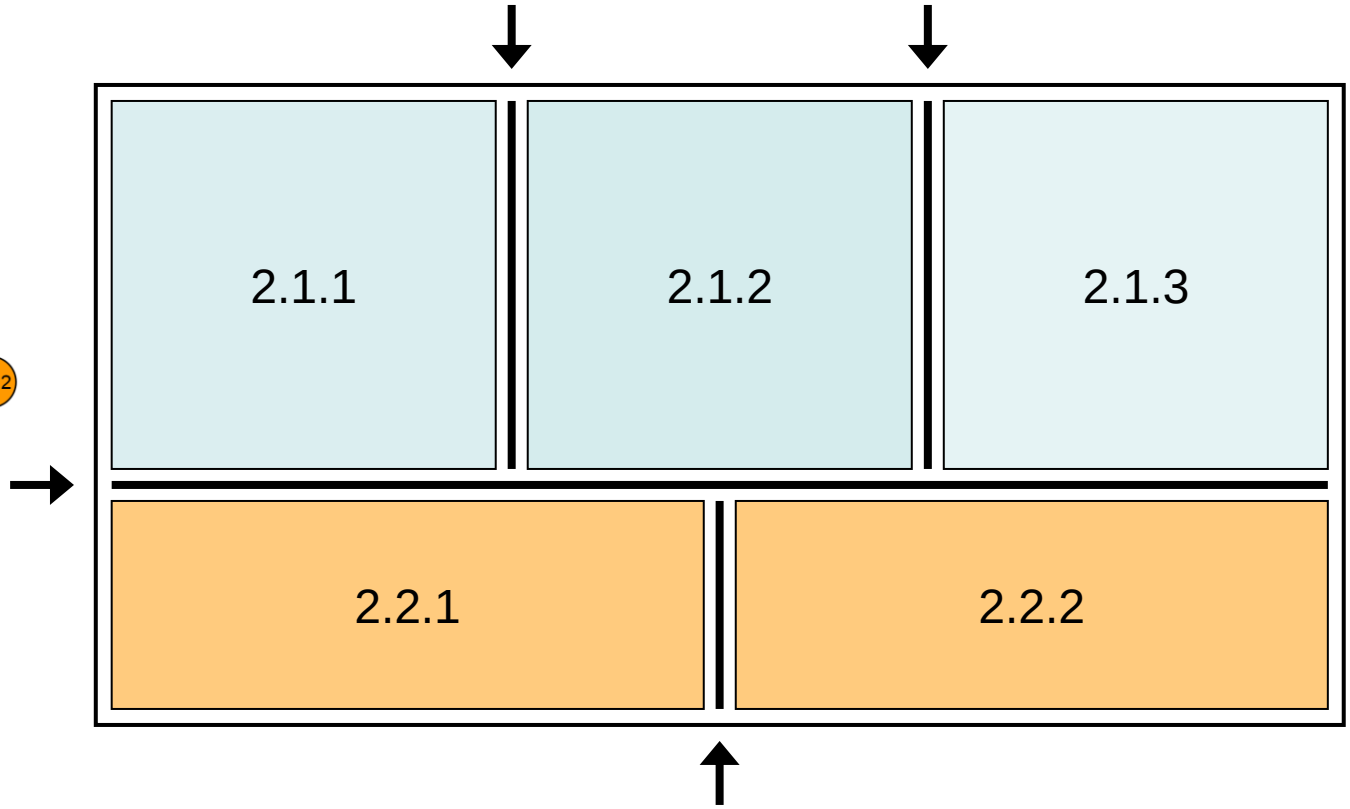


Rekursives „Slicing and Dicing“

Datenyp: Hierarchien
Visualisierung: Treemap

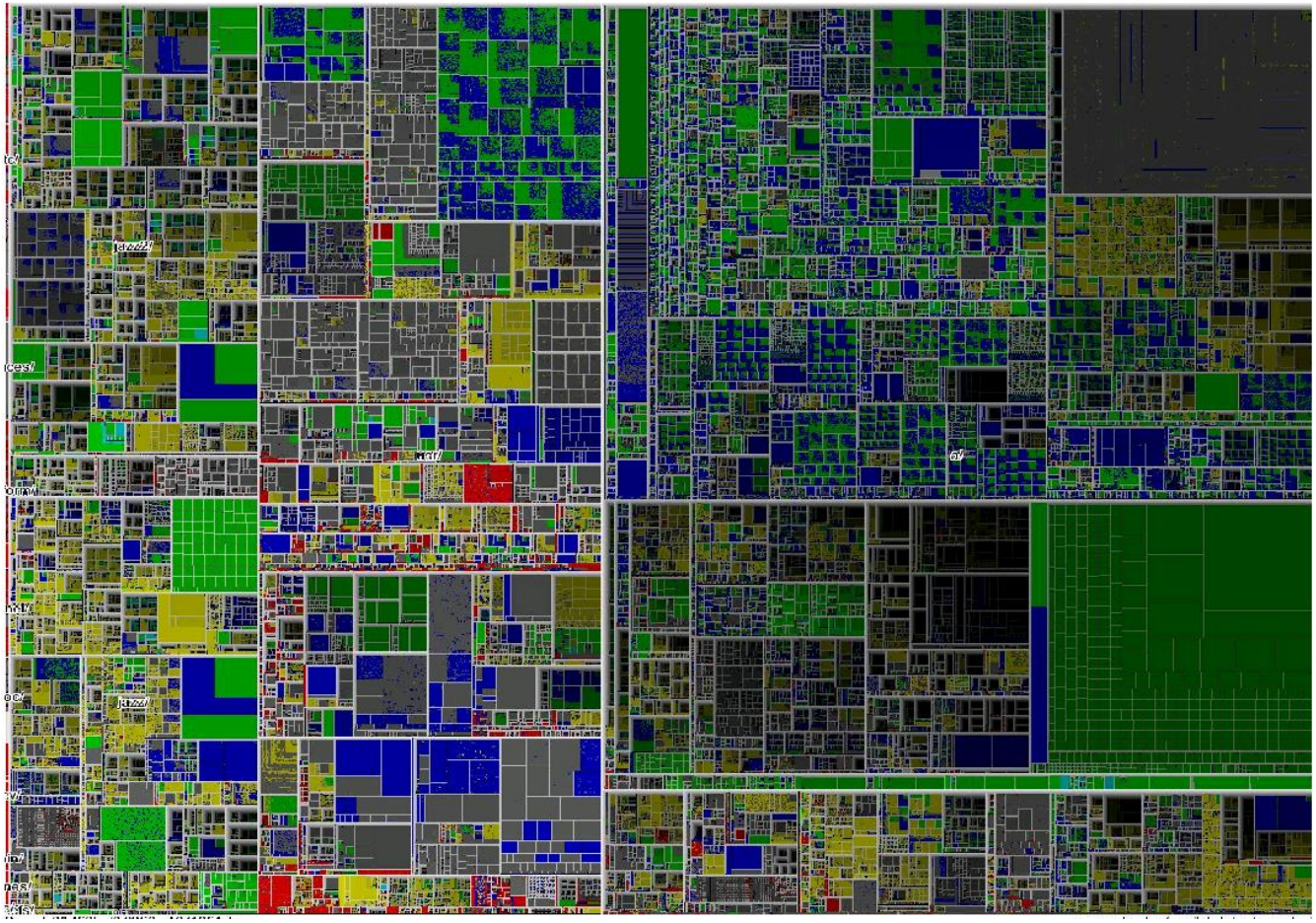
Idee: Kindknoten **IM** Vaterknoten zeigen
Volle Platznutzung





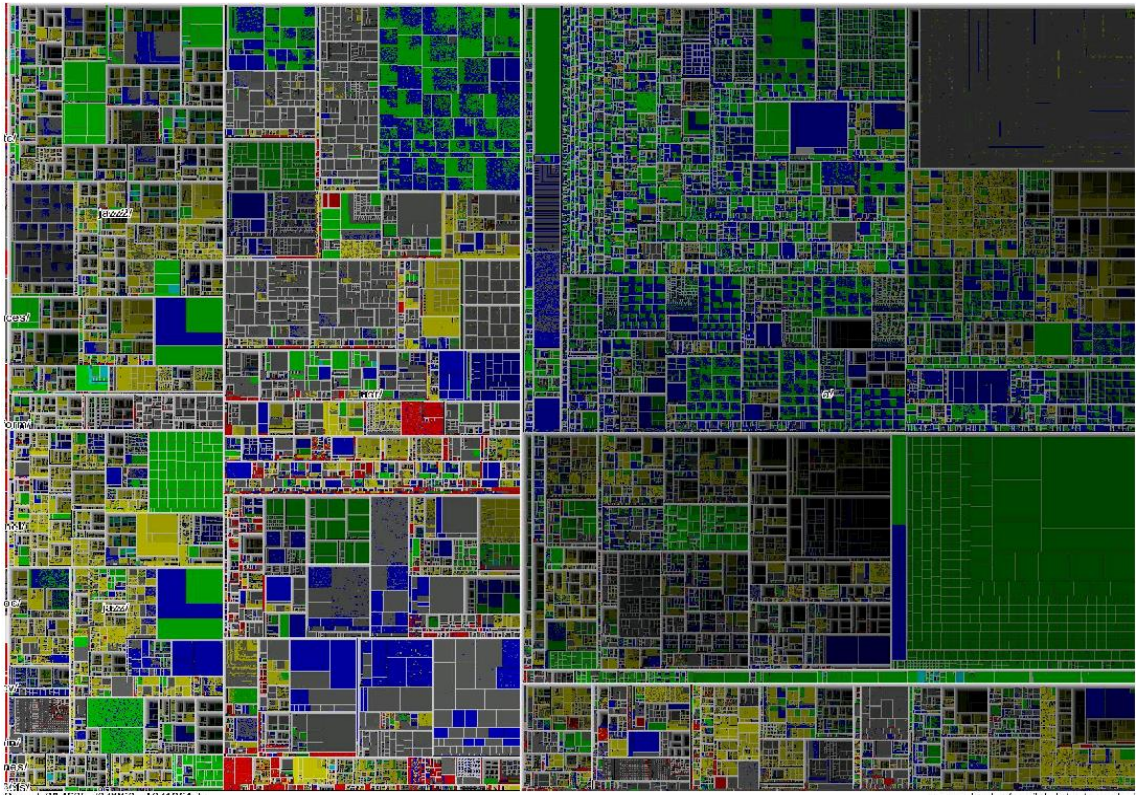
Rekursives „Slicing and Dicing“

Datenyp: Hierarchien
Visualisierung: Treemap
Beispiel: Visualisierung vom Filesystem



900.000 Elemente in einem Filesystem

Datenyp: Hierarchien
Visualisierung: Treemap
Problem: Lesbarkeit

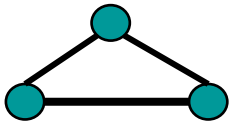


Probleme:

Überlappung der Elemente führt zur schlechter Lesbarkeit der Hierarchiestruktur

Datentyp: Graphen/Netzwerke

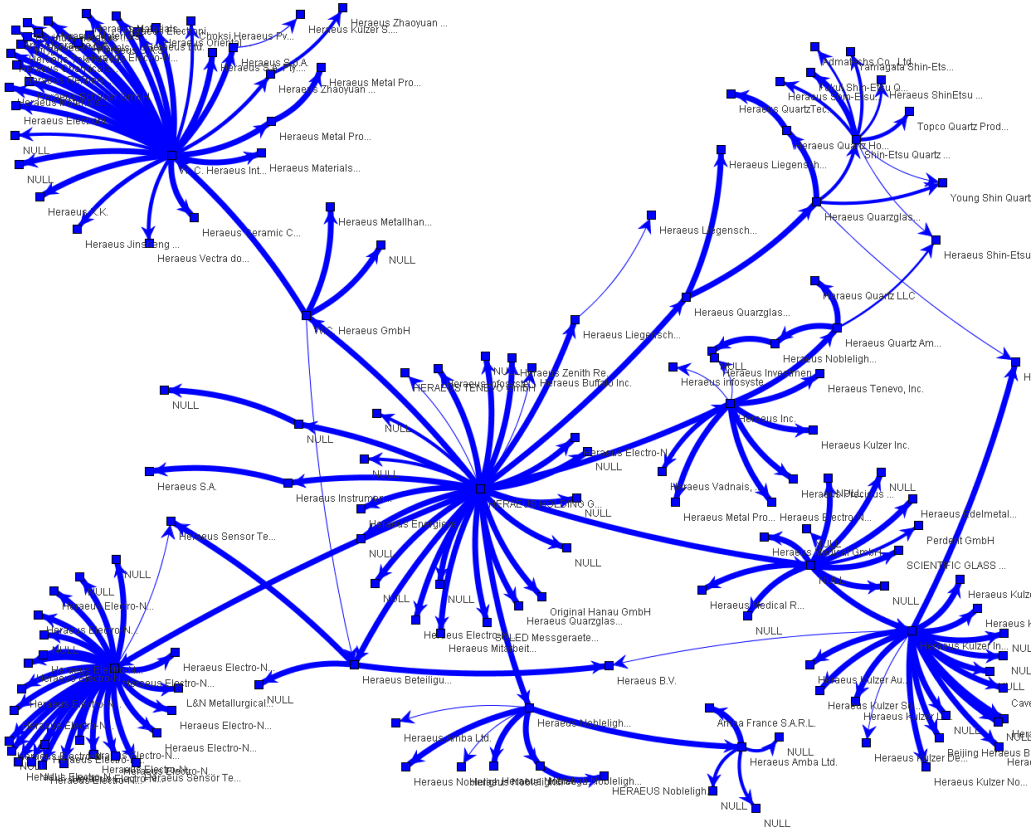
- Graph G(V, E) - ein Objekt bestehend aus Knoten und Kanten



Datenyp: Graphen/Netzwerke
Visualisierung: Knoten-Link Diagramm

- Klassische intuitive Graphdarstellung
  - Jeder Knoten: 1 Punkt
  - Beziehung als Linie
- Gut um Pfade zu verfolgen

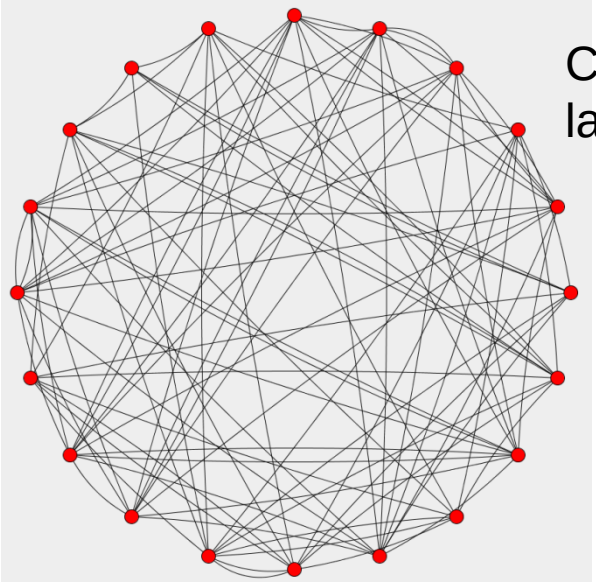
Knoten-Link Diagramm



Datenyp: Graphen/Netzwerke
Visualisierung: Knoten-Link Diagramm
Problem: Layout

- Klassische intuitive Graphdarstellung
- Gut um Pfade zu verfolgen

- Herausforderung: Graph Layout



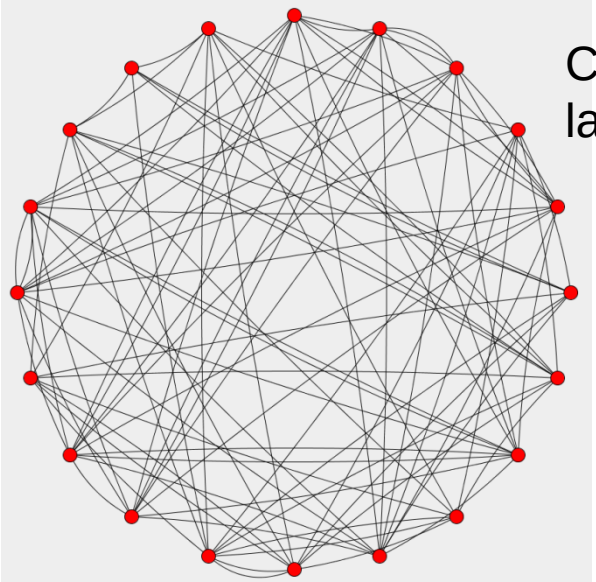
Circle layout

Datenyp: Graphen/Netzwerke
Visualisierung: Knoten-Link Diagramm
Problem: Layout

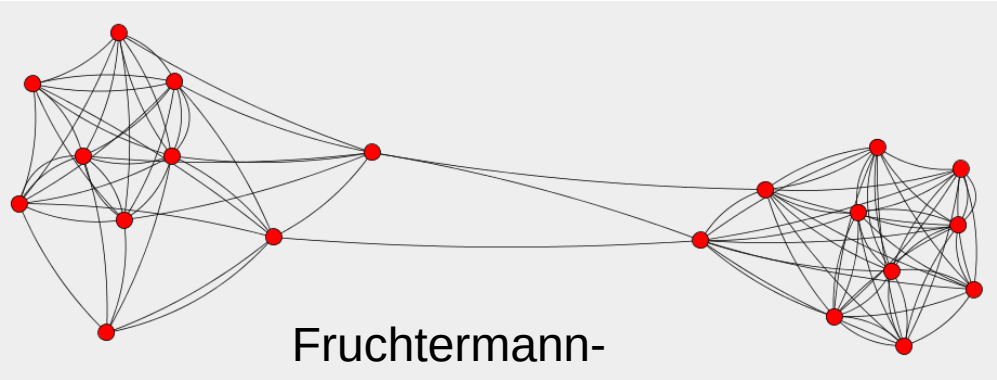
- Klassische intuitive Graphdarstellung
- Gut um Pfade zu verfolgen

- Herausforderung: Graph Layout

- Kriteria
  - Wenig Kantenüberschneidung
  - Möglichst gleich große Kanten
  - Wenig Knotenüberlappung
  - Klar deutliche Strukturen



Circle layout



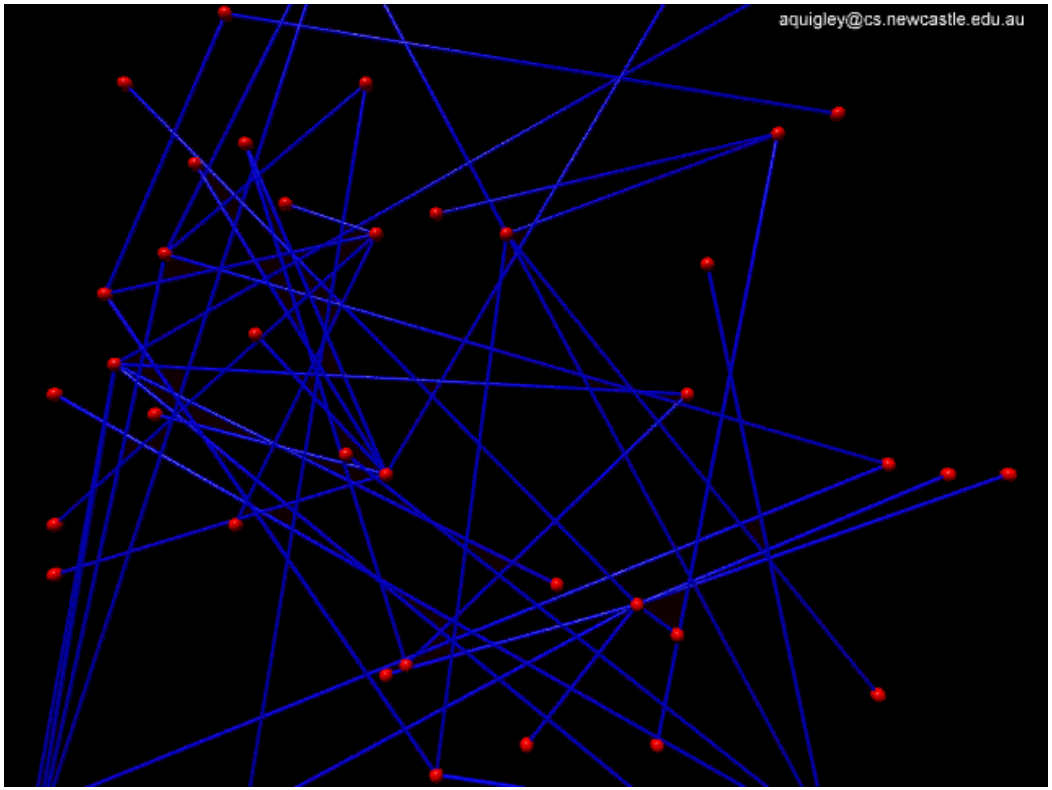
Fruchtermann-Reihngold layout

FYI: Graph Layout: Force-directed

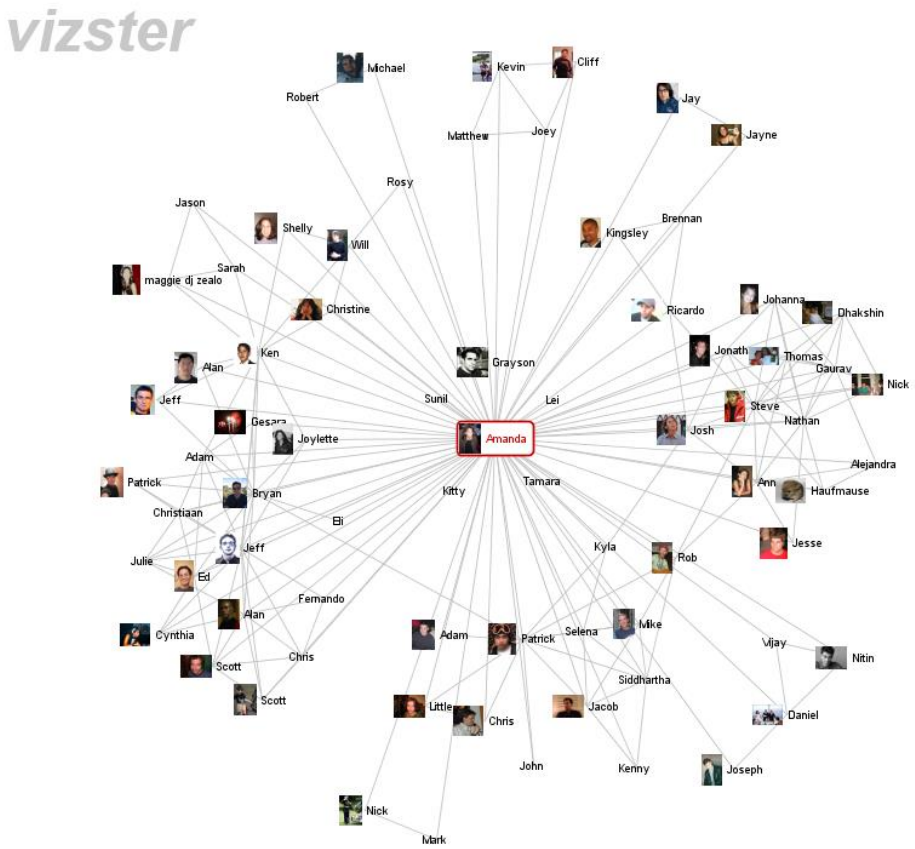
Häufigstes Layout

Idee:

- Knoten als abweisende elektrische Partikel
- Kanten als Federn
- Ziel: Kräfte zwischen Knoten und Kanten zu minimieren



Datenyp: Graphen/Netzwerke
Visualisierung: Adjazenzmatrix
Beispiel: Soziale Netzwerke



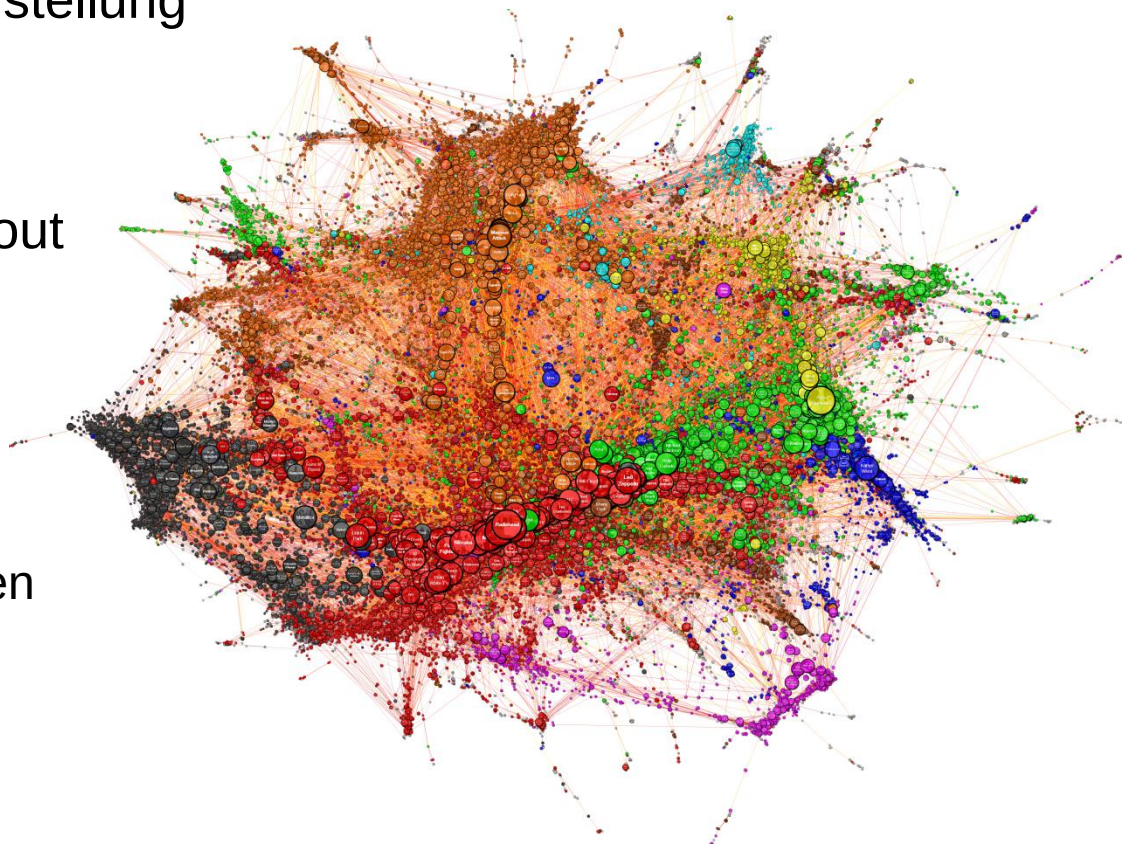
Visualisierung von Sozialen Netzwerken

Datenyp: Graphen/Netzwerke
Visualisierung: Knoten-Link Diagramm
Problem: Große Graphen

- Klassische intuitive Graphdarstellung
- Gut um Pfade zu verfolgen

- Herausforderung: Graph Layout

- Problem:
  - Große Graphen
  - Überlappungen von Kanten und Knoten bei großen Graphen

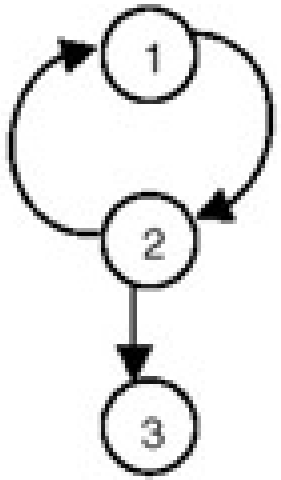


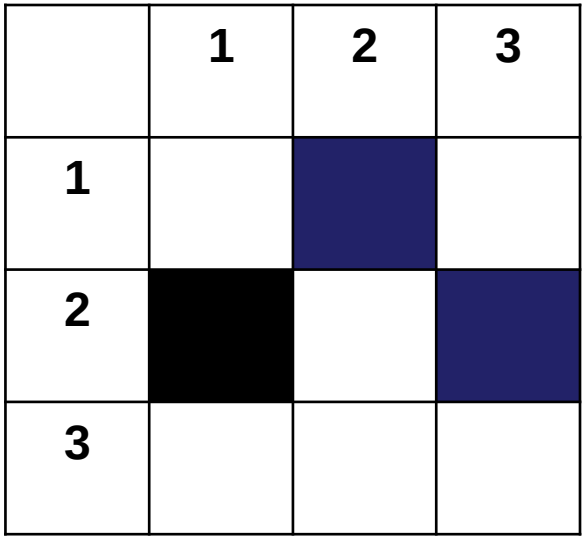
Music: This is Your Music on the Internet

https://relationary.wordpress.com/tag/visualization/

Datenyp: Graphen/Netzwerke
Visualisierung: Adjazenzmatrix

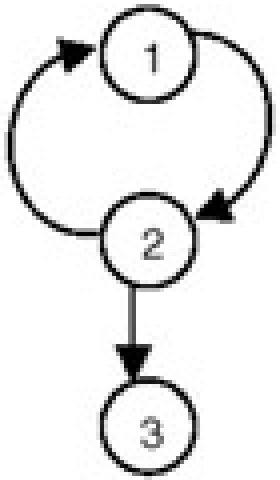
- Knoten: Spalten und Zeilen in Matrix
- Kanten: gefärbte Matrixzellen





Datenyp: Graphen/Netzwerke
Visualisierung: Adjazenzmatrix

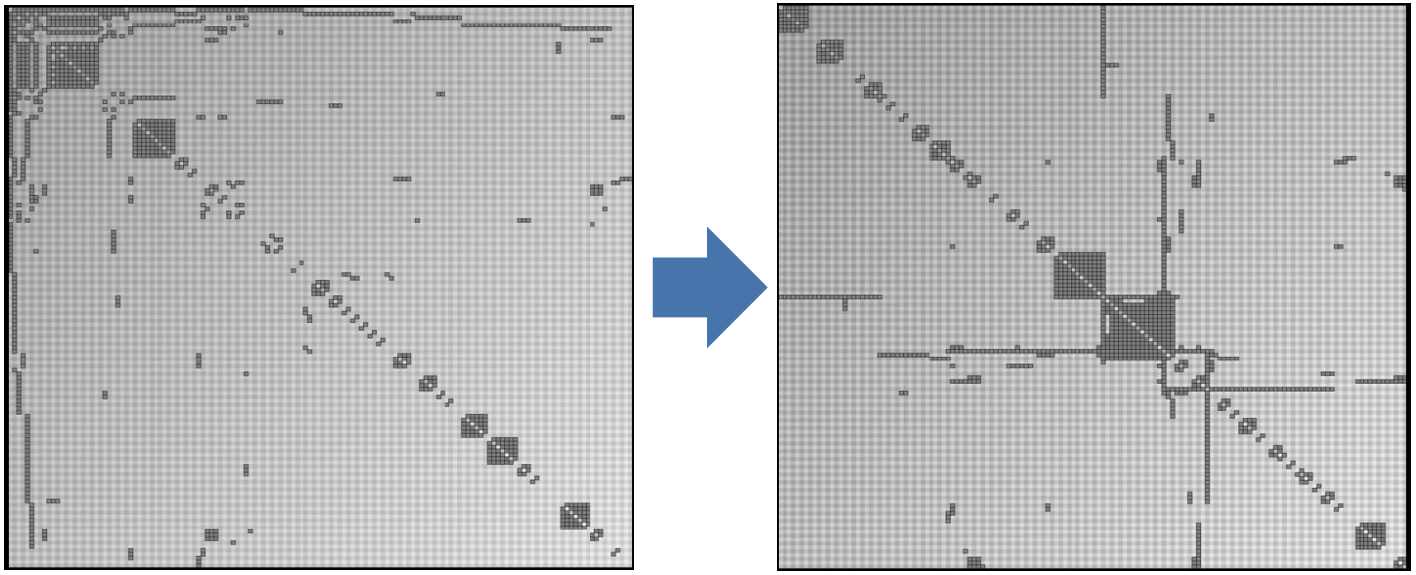
- Überlappungsfreie Darstellung
- Gut für Graphen mit vielen Kanten
- Keine Kantenüberschneidungen



	1	2	3
1			
2			
3			

Datenyp: Graphen/Netzwerke
Visualisierung: Adjazenzmatrix
Problem: Layout

- Überlappungsfreie Darstellung, Keine Kantenüberschneidungen
- Gut für Graphen mit vielen Kanten
- Herausforderung: Layout = Matrixsortierung um Datenstrukturen sichtbar



- Problem:
  - Schlechte Visualisierung von Pfaden

Zusammenfassung: Datentypen und Visualisierungstechniken

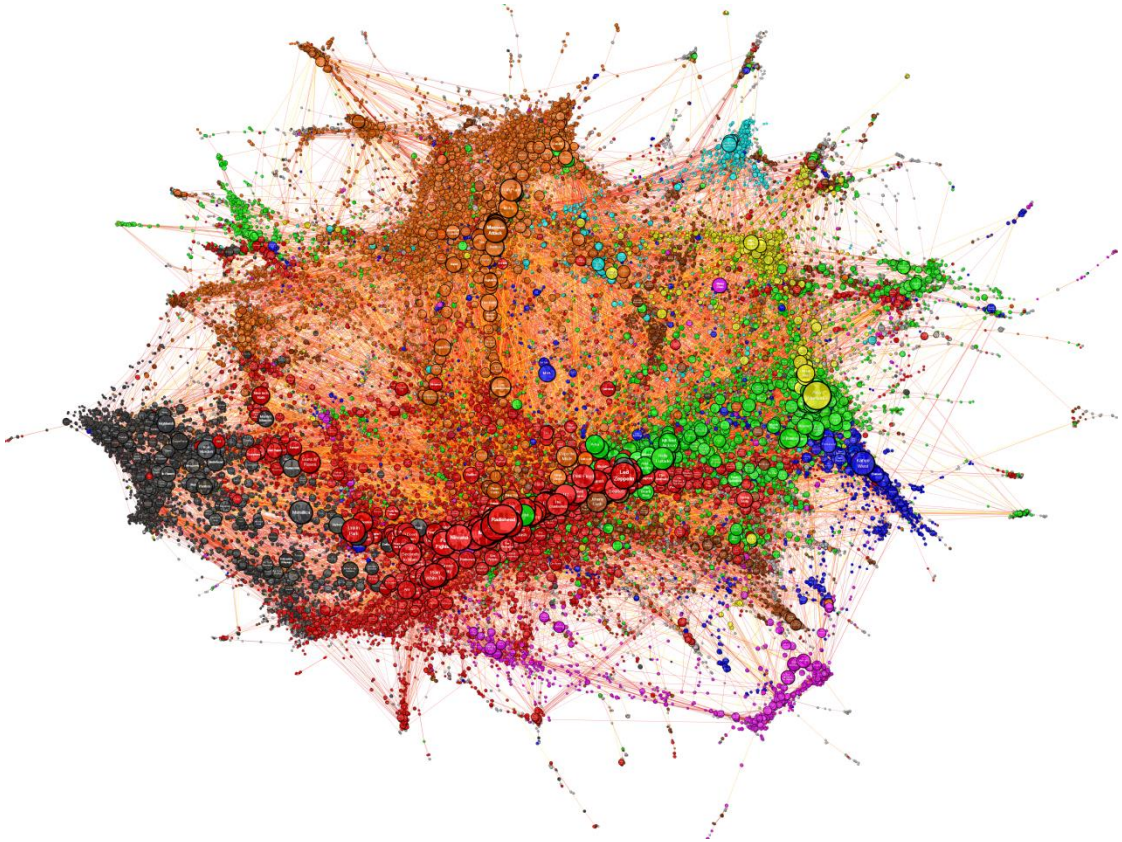
Datentyp	Visualisierungstechniken
1D	Balkendiagramm, Kuchendiagramm
Zeitreihe	Liniengraphik
2D	Scatterplott
3D	Scatterplot+, Parallele Koordinaten
nD, multivariat, multidimensional - wenige Dimensionen - viele Dimensionen	 ScatterplotMatrix, Parallele Koordinaten, Star Plot Dimensionsreduktion+Scatterplot
Hierarchien	Knoten-Link Diagram, Treemap
Graphen/Netzwerke	Knoten-Link Diagram, Adjazenzmatrix

Zusammenfassung: Datentypen und Visualisierungstechniken

TECHNISCHE UNIVERSITÄT DARMSTADT

Allgemeine Probleme:

- Große Datenmengen und wenig Platz am Bildschirm
  - Overplotting
    - → Filtrierung



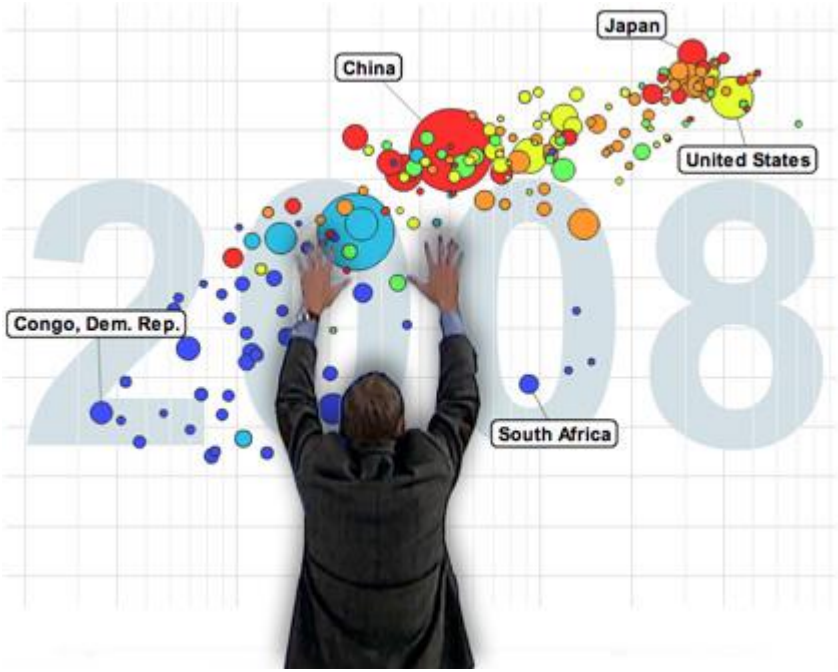
INTERAKTION

Informationsvisualisierung + Interaktion

Statische Visualisierung sind wie Bilder im Museum

Information visualization + Interaction

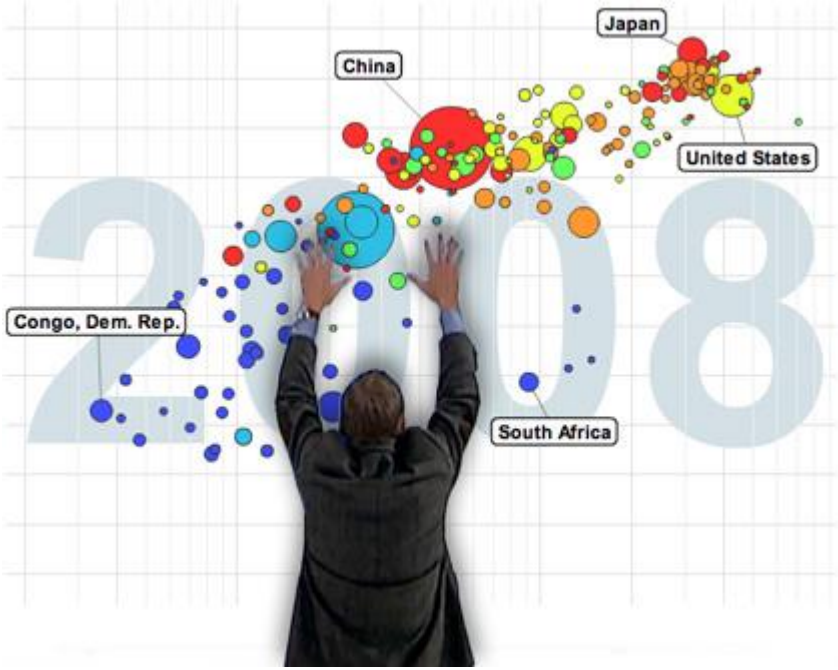
Interaktionstechniken ermöglichen
mit Daten zu Spielen, Daten zu explorieren…



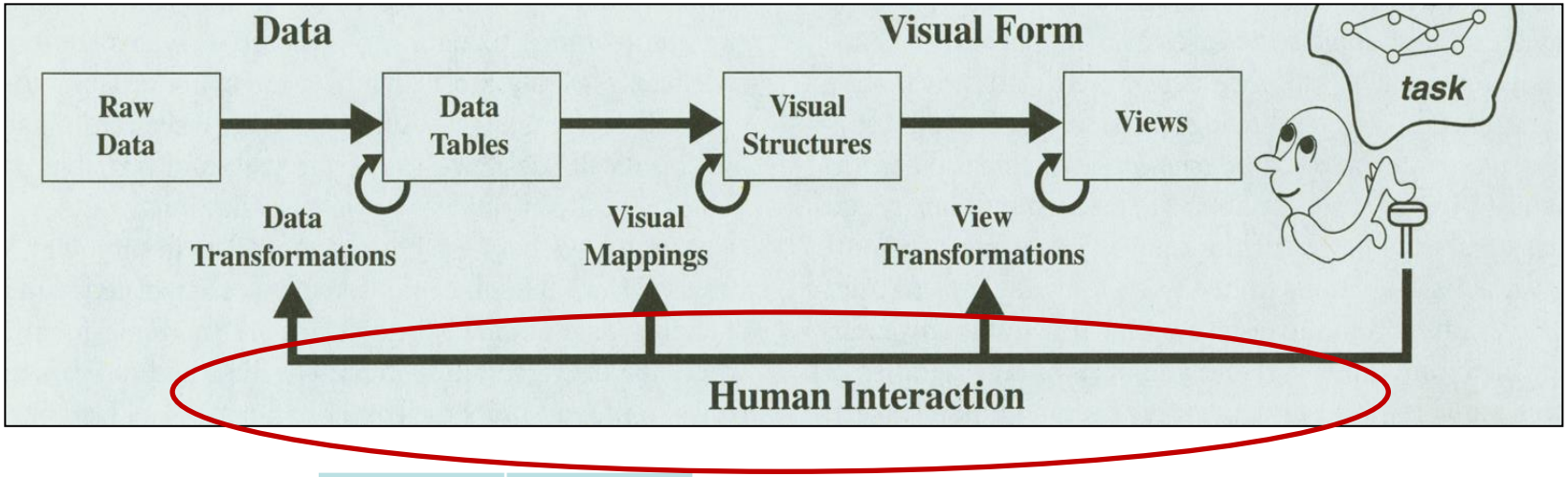
Information visualization + Interaction

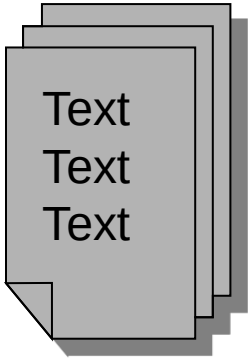
Interaktionstechniken ermöglichen
mit Daten zu Spielen, Daten zu explorieren…

Und sie reduzieren viele Probleme
der Visualisierung,
zB Überlappung



Information visualization + Interaction

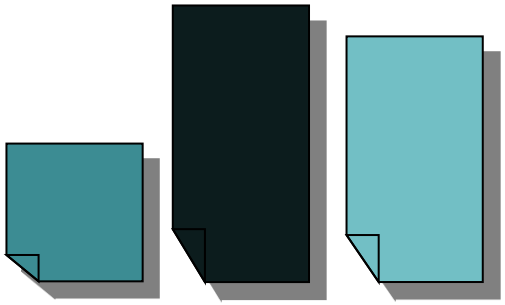




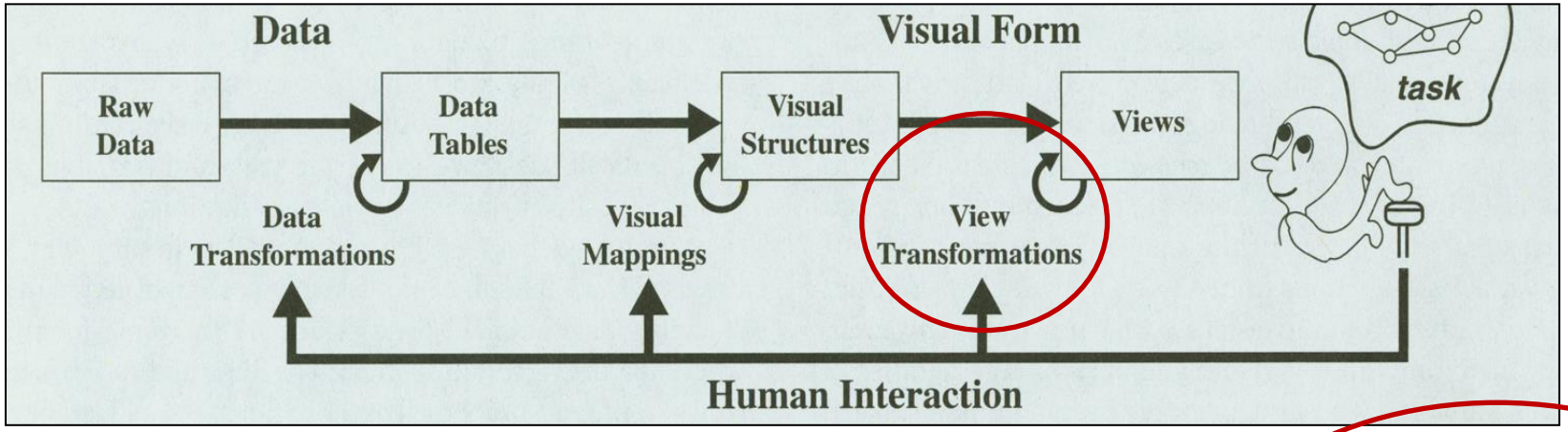
#Worte	#Satze
50	9
100	20
80	7

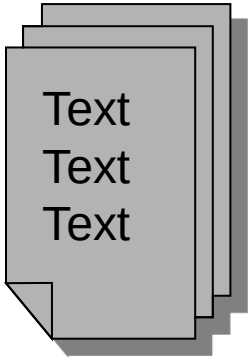
#Worte ⇨ Größe

#Satze ⇨ Farbe



Information visualization + Interaction

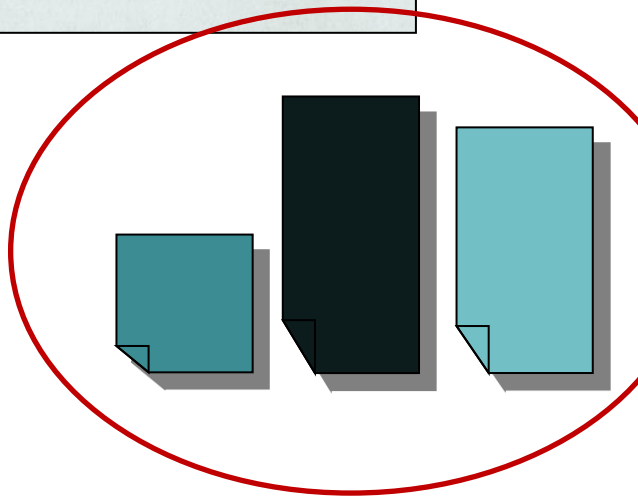




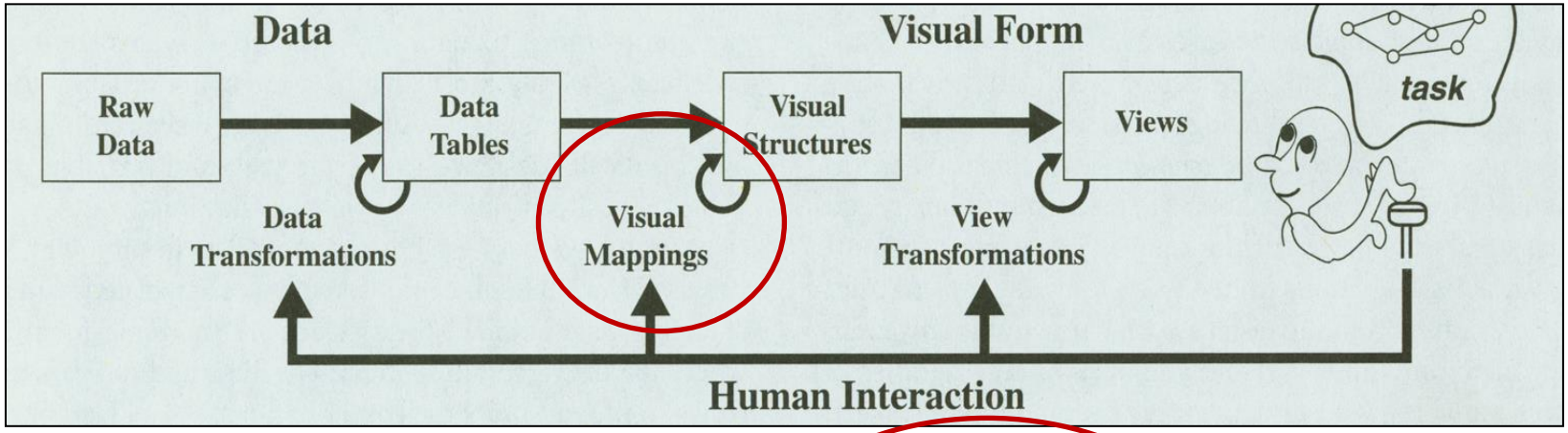
#Worte	#Satze
50	9
100	20
80	7

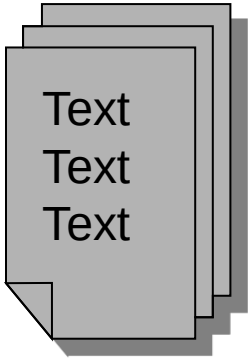
#Worte ⇨ Größe

#Satze ⇨ Farbe



Information visualization + Interaction

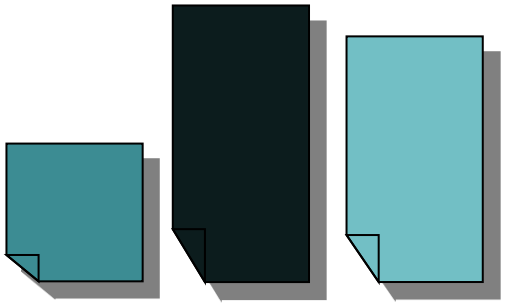




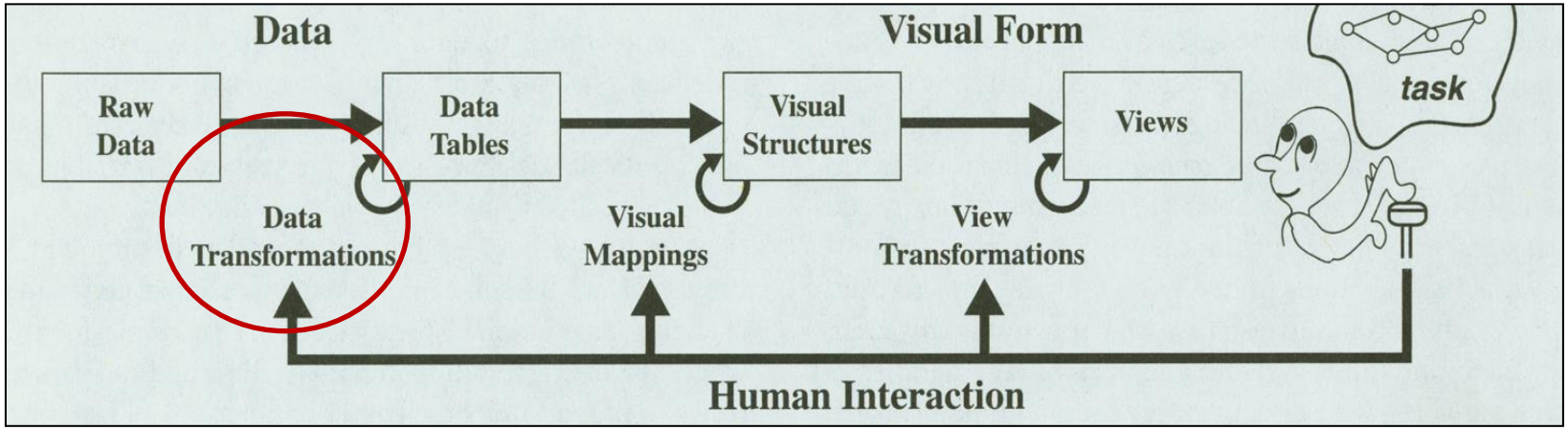
#Worte	#Satze
50	9
100	20
80	7

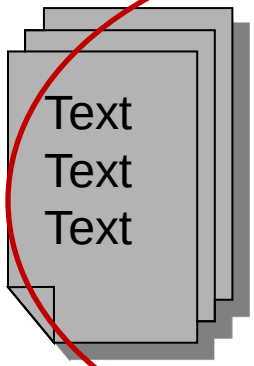
#Worte ⇨ Größe

#Satze ⇨ Farbe



Information visualization + Interaction

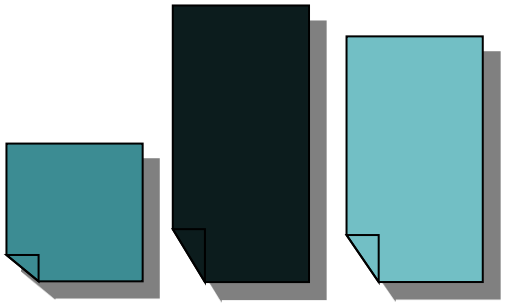




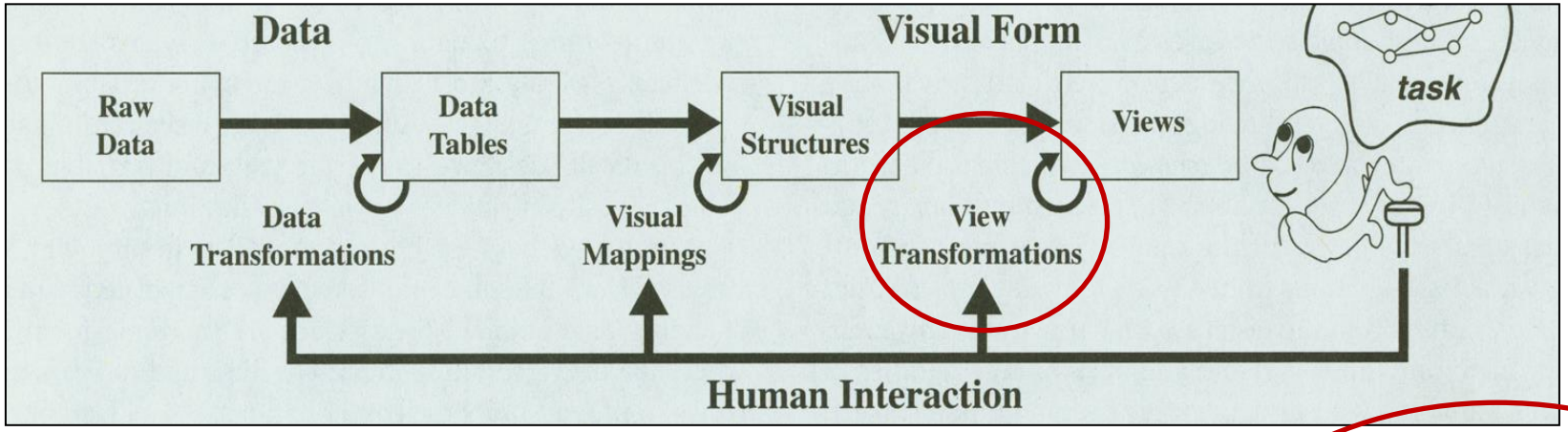
#Worte	#Satze
50	9
100	20
80	7

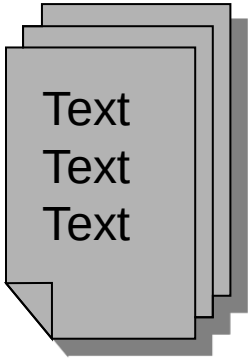
#Worte ⇨ Größe

#Satze ⇨ Farbe



Information visualization + Interaction

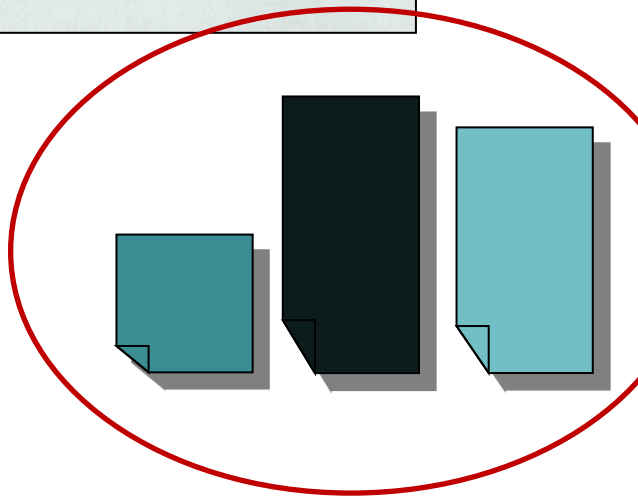




#Worte	#Satze
50	9
100	20
80	7

#Worte ⇨ Größe

#Satze ⇨ Farbe



Interaktionstechniken – View Transformation

- View transformation
  - Navigieren
    - Techniken
      - Zoom
      - pan

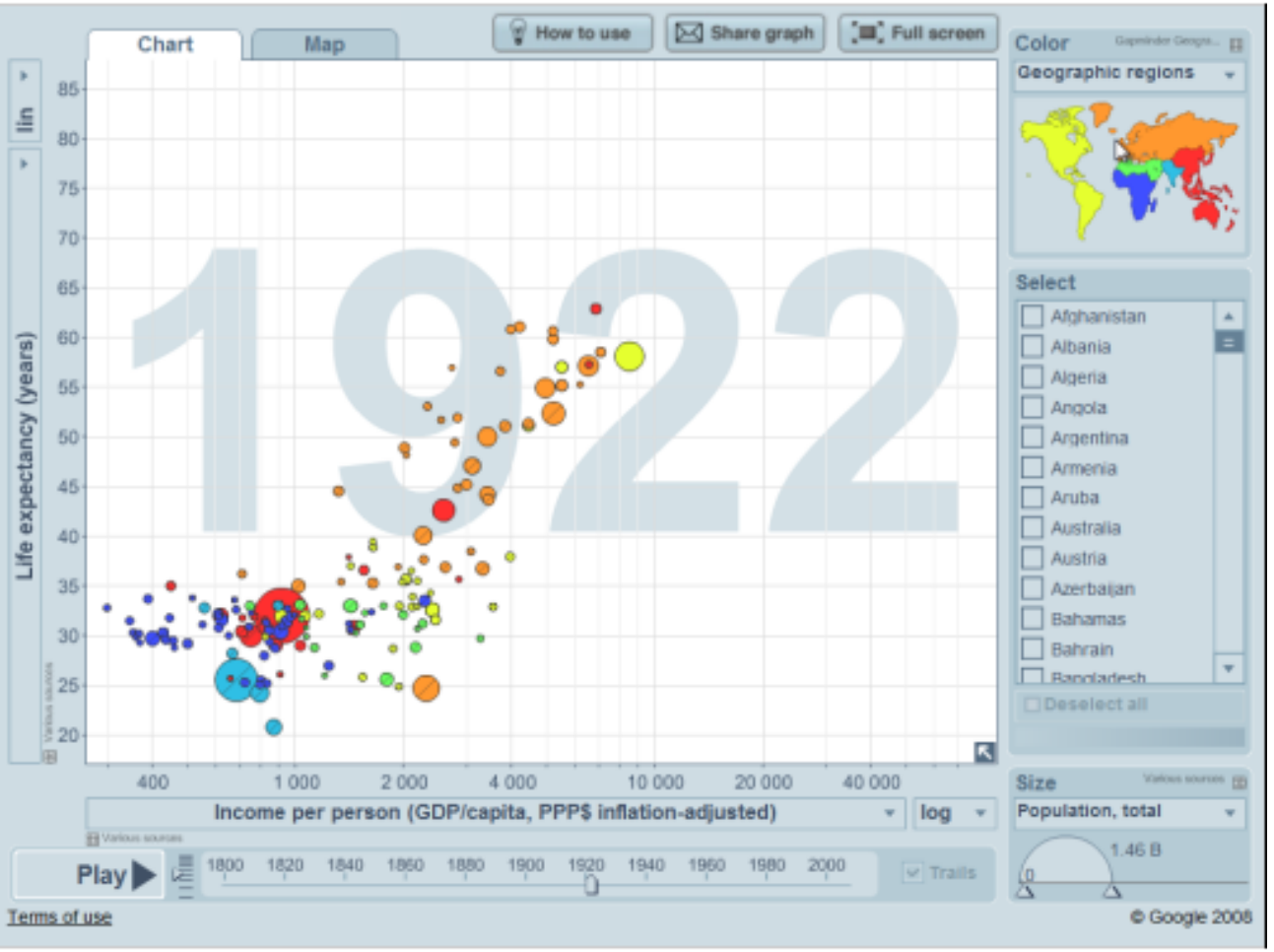
Beispiel: Google Maps

View transformation
Highlighting - Hervorheben

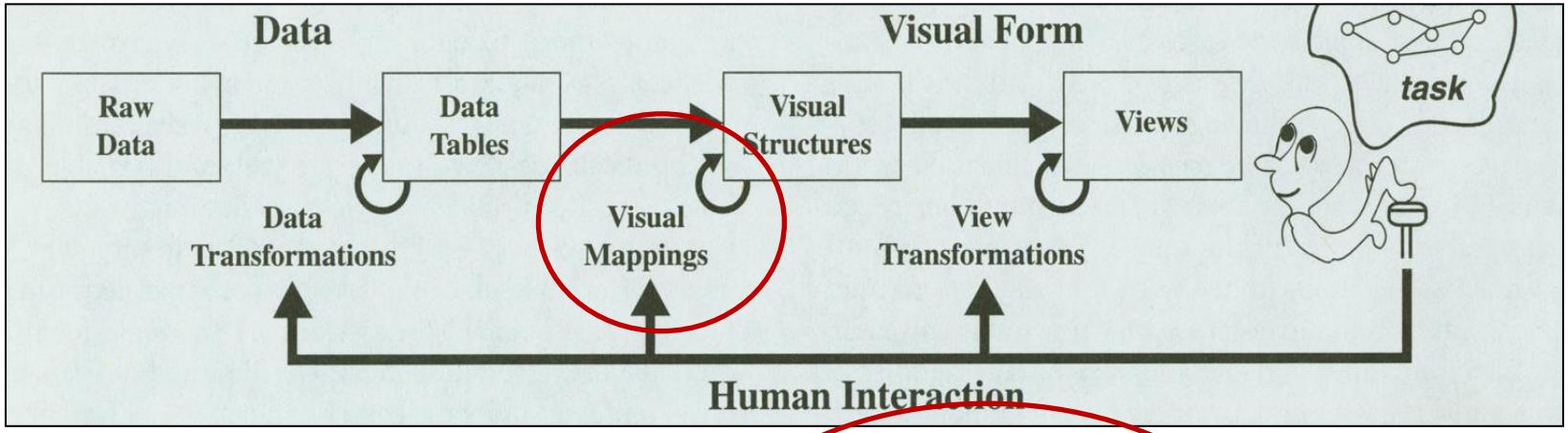
TECHNISCHE UNIVERSITÄT DARMSTADT

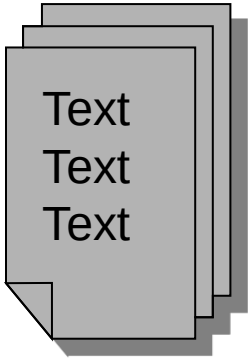
View transformation
Fokussieren
Technik
Hervorheben

Beispiel: Gapminder



Information visualization + Interaction

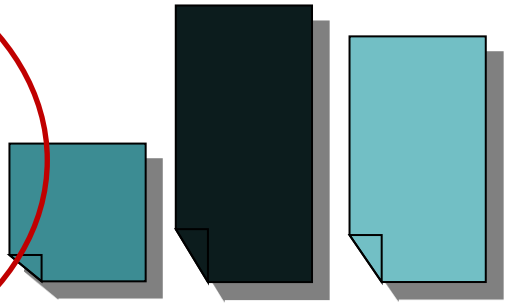




#Worte	#Satze
50	9
100	20
80	7

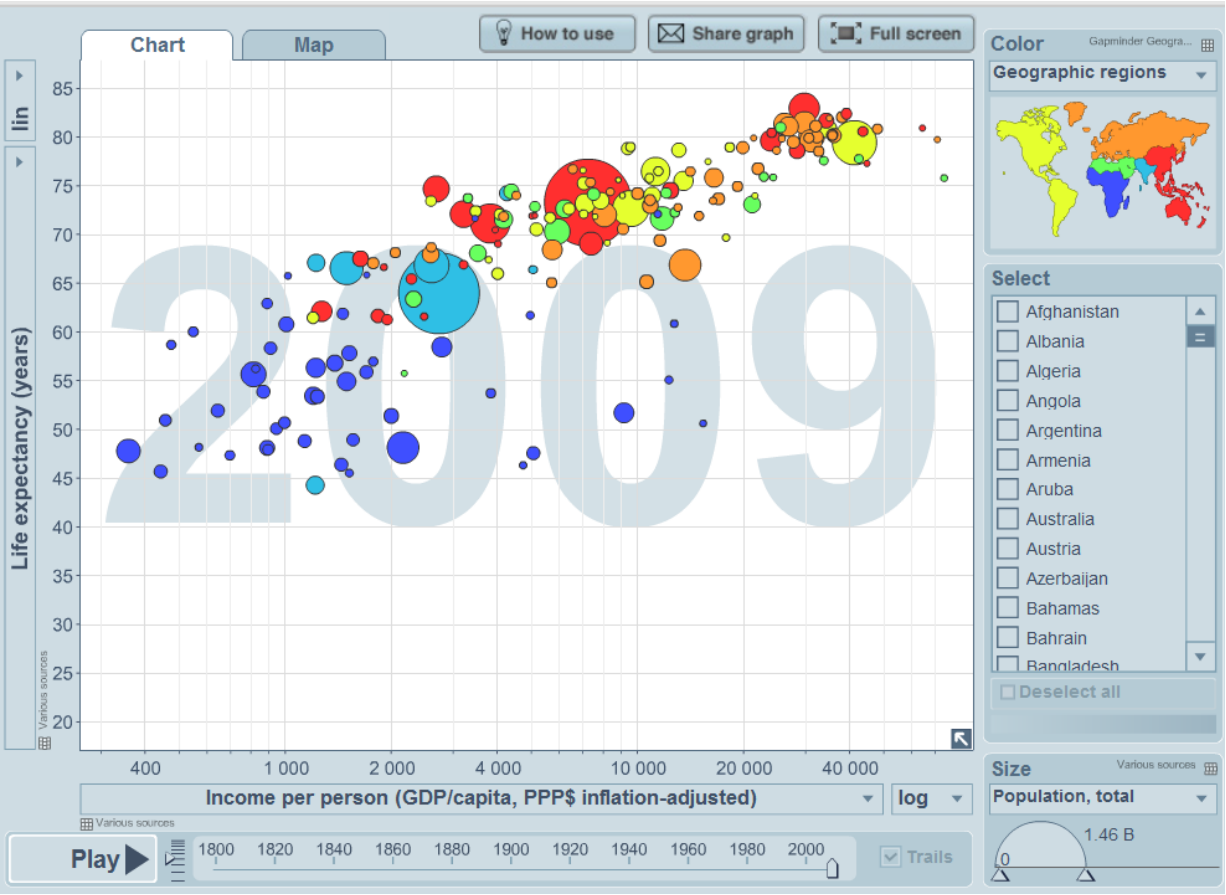
#Worte ⇨ Größe

#Satze ⇨ Farbe

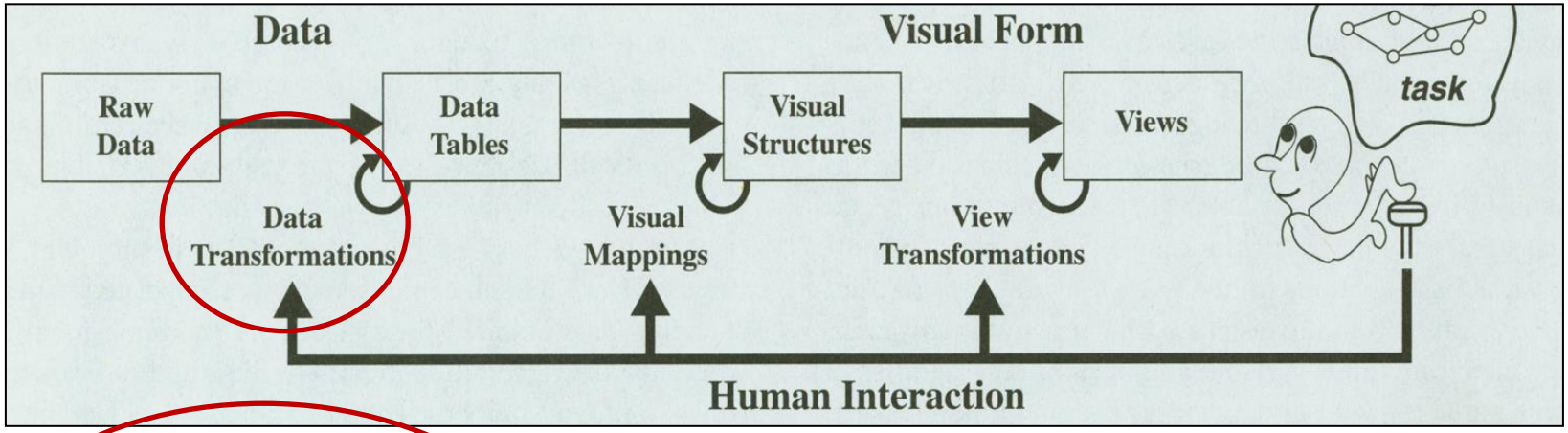


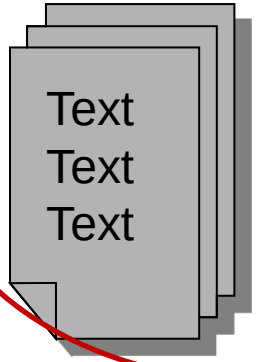
Interaction - Visual Mapping Aenderung

- Visual Mapping Aenderung Techniken:
  - Datenauswahl für Abbildung
  - Farbschema-änderung



Information visualization + Interaction

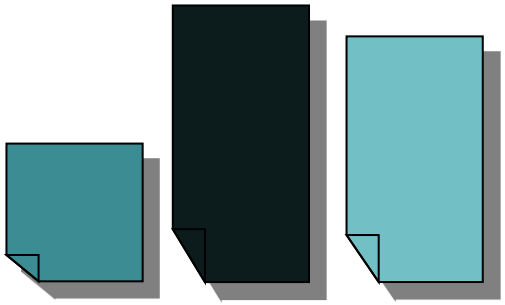




#Worte	#Satze
50	9
100	20
80	7

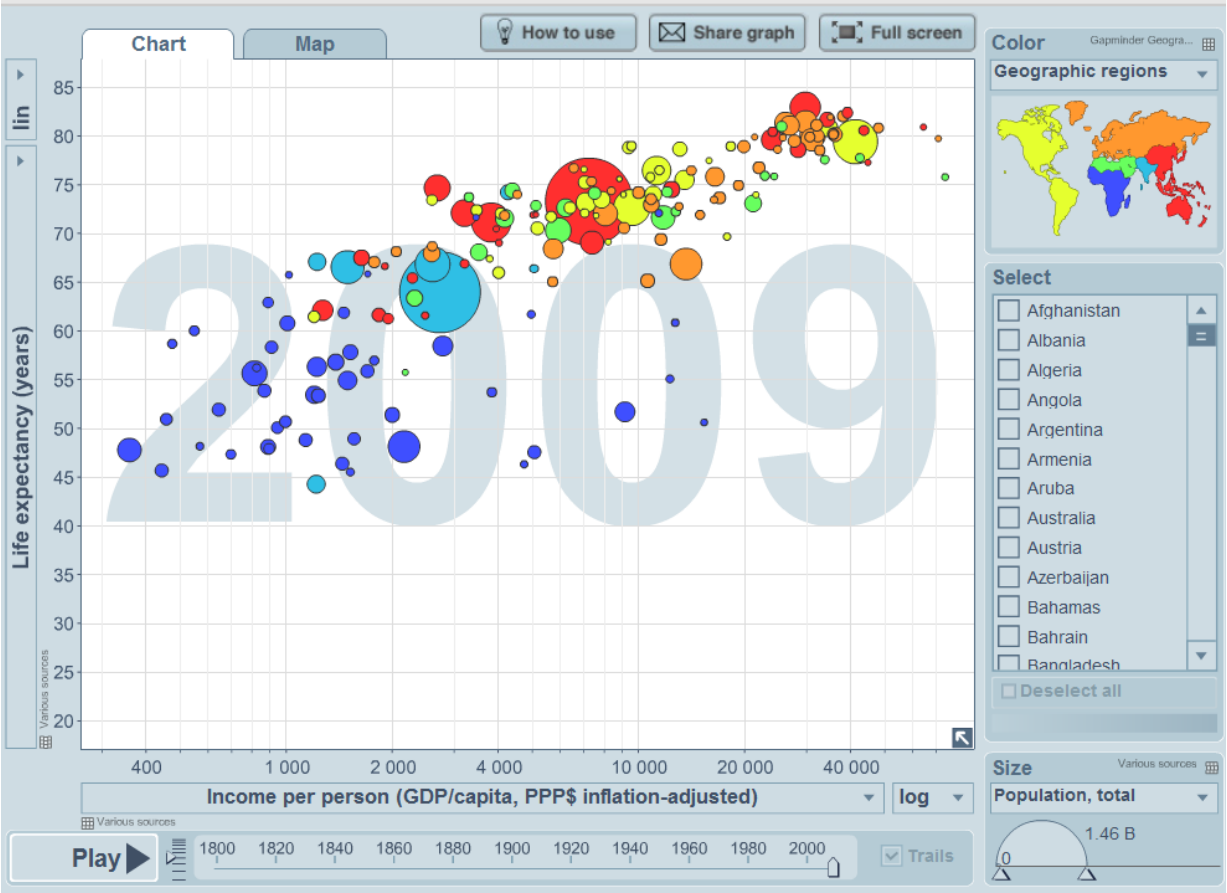
#Worte ⇨ Größe

#Satze ⇨ Farbe



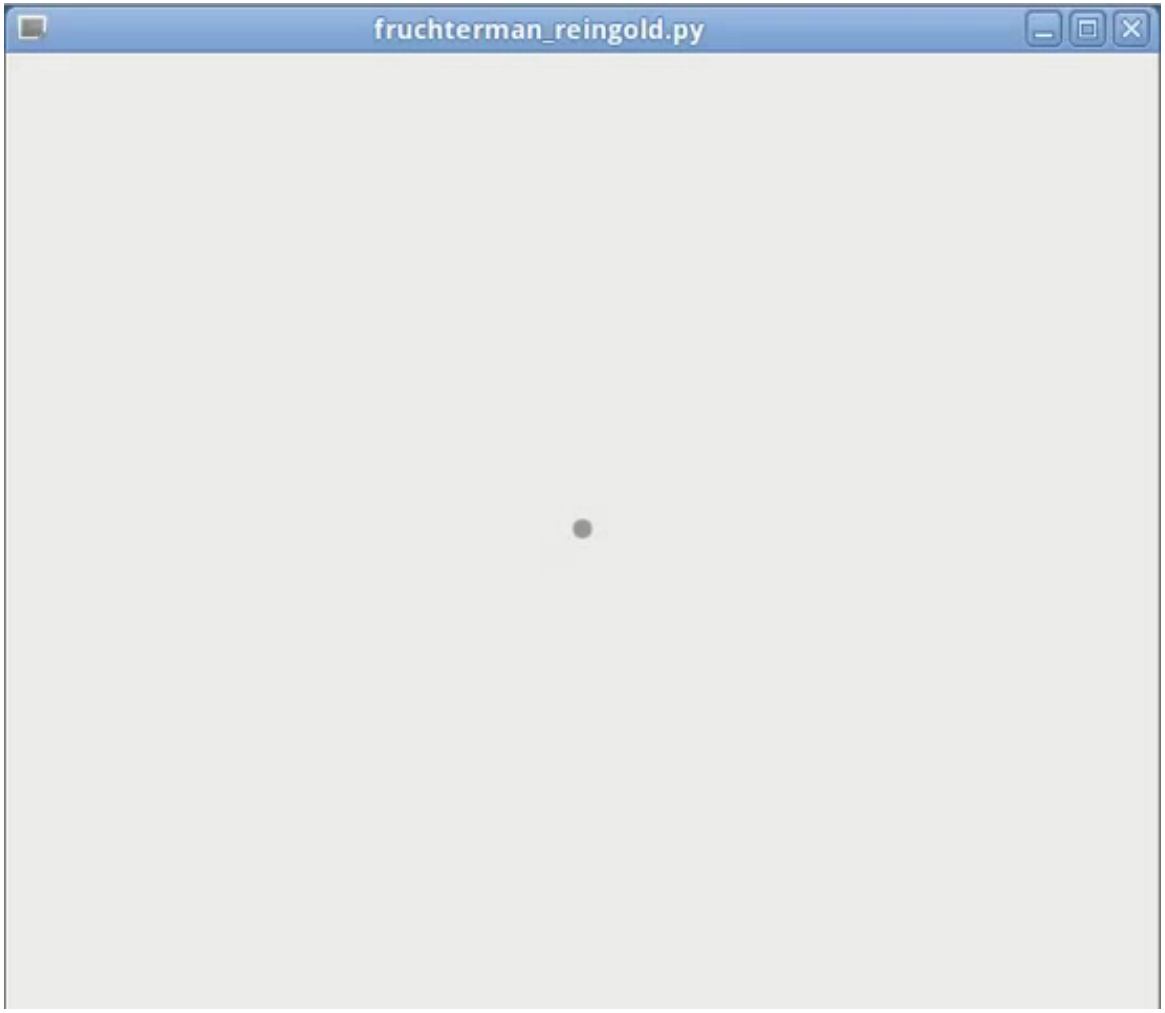
Interaction - Datentransformation

- Datentransformation Techniken:
  - Änderung der Datennormalisierung



Interaction - Datentransformation

- Datentransformation Techniken:
  - Änderung der Datennormalisierung
  - Dateneditieren



Zusammenfassung: Interaktionstechniken

- Interaktion erlaubt dem Benutzer
  - die Daten aus verschiedenen Perspektiven zu sehen (im Datensatz zu navigieren)
  - Interaktion greift in die Schritte der Informationsvisualisierungsipeline ein

Zusammenfassung: Interaktionstechniken

TECHNISCHE UNIVERSITÄT DARMSTADT

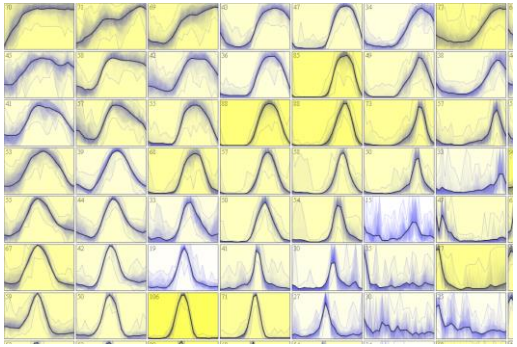
- Interaktion erlaubt dem Benutzer
  - die Daten aus verschiedenen Perspektiven zu sehen (im Datensatz zu navigieren)
  - Interaktion greift in die Schritte der Informationsvisualisierungsipeline ein

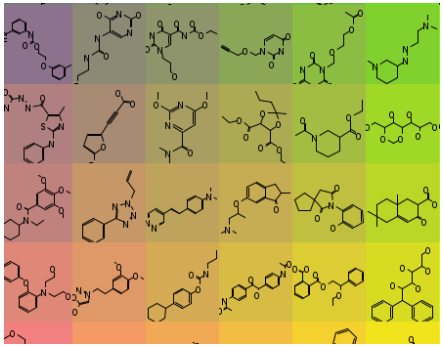
Pipelineschritt	Technik
View	Zoom Pan Highlight/Hervorheben
Visuelle Abbildung	Datenauswahl für Abbildung Farbschemaänderung
Datentransformation	Änderung der Datennormalisierung Dateneditieren

WEITERFÜHRENDE VERANSTALTUNGEN

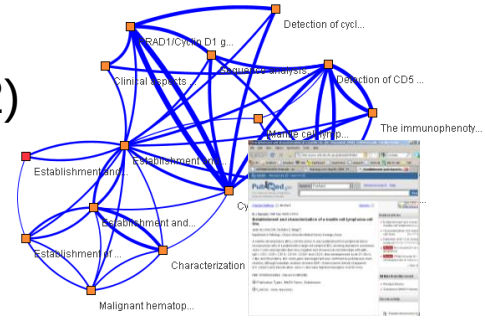
Lehrveranstaltungen zur Informationsvisualisierung

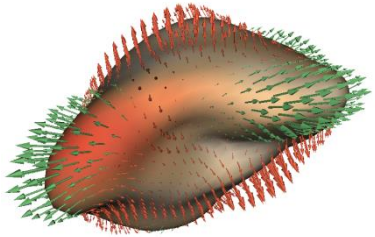
- **Vorlesung:**
  - Informationsvisualisierung und Visual Analytics (V2+Ü2)



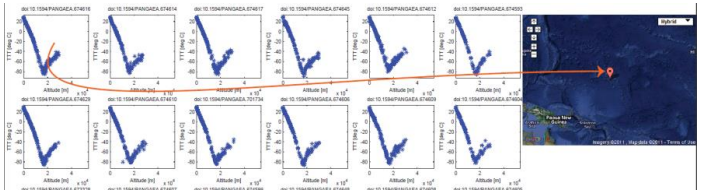


- **Seminare:**
  - Visualisierung großer Datenmengen (S2)



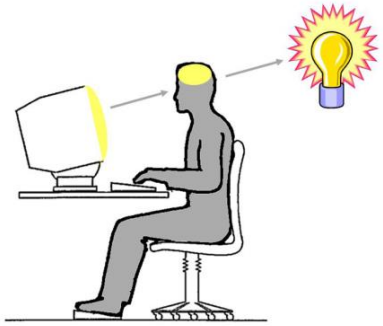


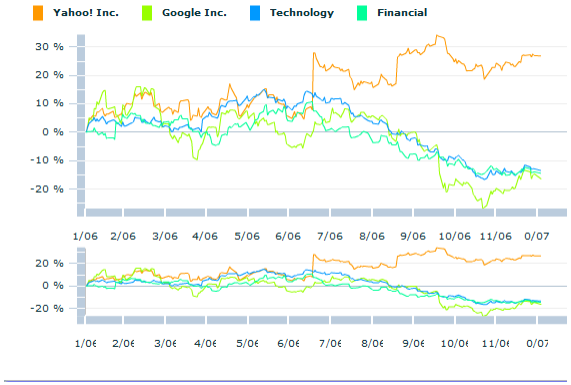
- **Praktikas**
- **Bachelorarbeiten, Masterarbeiten**



Aktuelle Projekte
Praktika, BA/MA/Hiwi

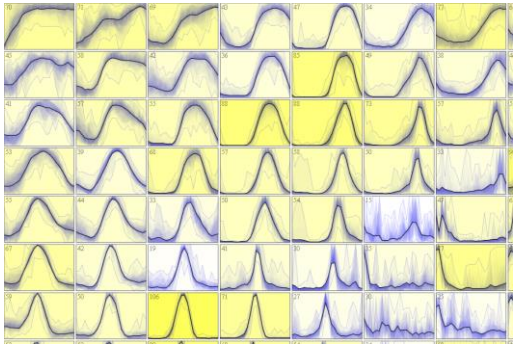
- Psychologie
- Medizin
- Wirtschaft/Finanzen
- Biologie & Chemie

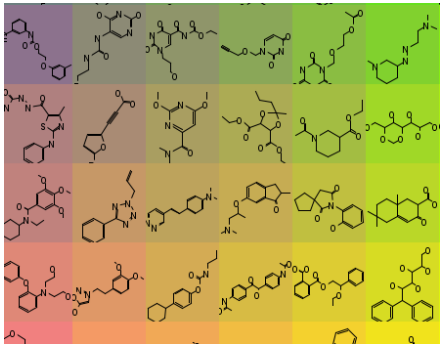


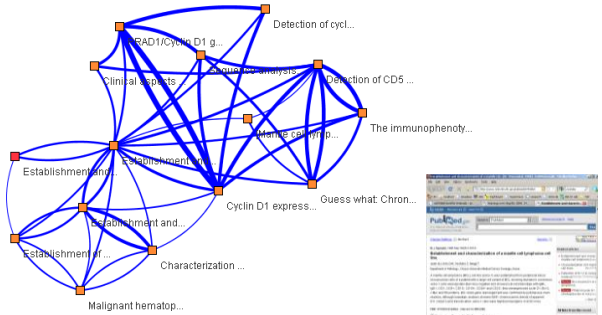


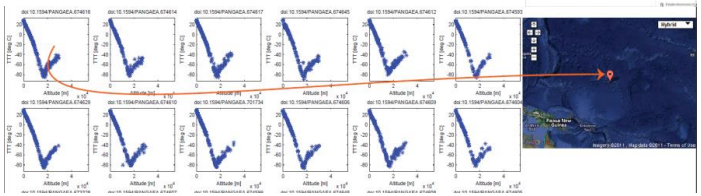
Kontakt

Tatiana von Landesberger
Visual Search and Analysis
GRIS, Fraunhoferstr. 5
Tel 155 – 631
Tatiana.von-landesberger@
gris.tu-darmstadt.de









http://www.gris.tu-darmstadt.de/vsa

ENDE...