



UNIVERSITÄT
LEIPZIG

Informationsvisualisierung

Sommersemester 2025

Dirk Zeckzer

Institut für Informatik



Teil I

Einführung

Übersicht der Vorlesung

1. Einführung
2. Daten und Aufgaben
3. Datentransformationen
4. Wahrnehmung
5. Darstellung von mehrdimensionalen Daten
6. Zeitabhängige Daten
7. Darstellung von Graphen
8. Darstellung von Bäumen
9. Visuelle Abbildung
10. Ansichtstransformationen
11. Design und Validierung

Übersicht

1. Einführung

1.1 Warum Informationsvisualisierung?

1.2 Definitionen

1.3 Ziele

1.4 Überblick

1.5 Weitere Beispiele

Warum Informationsvisualisierung?

- ▶ Allgemeine Ziele
 - ▶ Suche in Daten
 - ▶ Analyse von Daten
 - ▶ Extraktion von Information, welche implizit in den Daten enthalten ist
 - ▶ Identifikation von
 - ▶ Mustern
 - ▶ Strukturen
 - ▶ Trends
 - ▶ Beziehungen
 - ▶ Anomalien
 - ▶ Entdeckung neuer Informationen
- ▶ Automatisierter Ansatz
 - ▶ Statistik
 - ▶ Machine-Learning
 - ▶ Knowledge Discovery in Databases (KDD, Extraktion von “Wissen” aus Datenbanken)
 - ▶ Data Mining
 - ▶ ...

Warum Informationsvisualisierung? – Statistik

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 $X - \bar{X} = 110.0$
regression sum of squares = 27.50
residual sum of squares of Y = 13.75
correlation coefficient = .82
 $r^2 = .67$

Abbildung: Vier Datensätze mit gleichen statistischen Eigenschaften. [F.J. Anscombe. Graphics in Statistical Analysis. American Statistician 27, 17–21, 1973], [Tufte1983]

Warum Informationsvisualisierung?

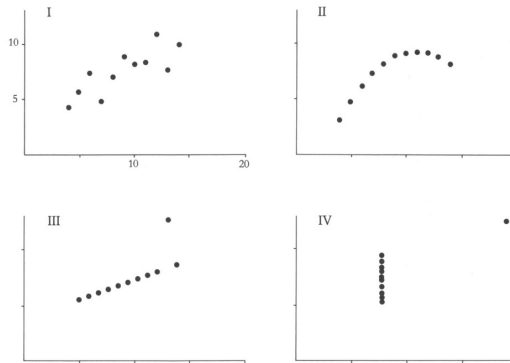


Abbildung: Einfache kartesische Darstellung der vier Datensätze. [F.J. Anscombe. Graphics in Statistical Analysis. American Statistician 27, 17–21, 1973], [Tufte1983]

Warum Informationsvisualisierung?



Warum Informationsvisualisierung?

- ▶ Visualisierung
 - ▶ zur Analyse
 - ▶ von hochdimensionalen Daten,
 - ▶ die miteinander in Beziehung stehen können
- ▶ Vorteile
 - ▶ Interaktiv
 - ▶ Flexibel

Definitionen

► **Information:**

Datenmenge, der eine bestimmte Bedeutung zugeordnet werden kann

► **Visualisierung:**

Vorgang, etwas in sichtbare Beziehungen zu setzen oder in sichtbare Form zu überführen

► **Informationsvisualisierung:**

Konzepte, Methoden und Tools zur visuellen Darstellung von Informationen aus Datenbanken, digitalen Bibliotheken oder anderen großen Dokumentensammlungen (Däbler / Plam, 1998)

Definitionen

► **Visual Thinking:**

“What information visualization is really about is external cognition, that is, how resources outside the mind

can be used to boost cognitive capabilities of the mind.”

[Stuard Card im Vorwort zu Colin Ware's Buch

“Information Visualization”]

► Computergestützte Informationsvisualisierung:

- Erzeugen von Darstellungen in Echtzeit
- Dynamische Darstellungen
- Integration in den Prozess des Verstehens und des Findens
- Veränderung der Darstellung während des Analyseprozesses

Ziele

Grundlegende Aufgaben der Visualisierung

- ▶ Unterstützung der angeborenen menschlichen Fähigkeit, die Umwelt zu erfassen und zu verstehen
 - ▶ Der Mensch nimmt seine Umgebung zu 75% visuell wahr und erhält daraus die notwendigen Informationen
- ▶ Erleichterung
 - ▶ der Analyse
 - ▶ der Interpretation von Daten
- ▶ Sichtbarmachung verborgener Trends
- ▶ Unterstützung von Mustererkennung

Ziele

Grundlegende Aufgaben der Visualisierung

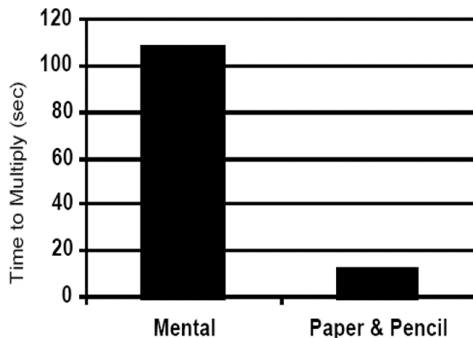
- ▶ Vereinfachter Zugang zu großen Datenmengen durch
 - ▶ Klassifikation
 - ▶ Datenstrukturierung
 - ▶ ...
- ▶ Unterstützung bei der Analyse komplexer Prozessabläufe und Objektbeziehungen in der realen Welt durch
 - ▶ Symbole
 - ▶ Diagramme
 - ▶ Animationen

Ziele

Visualisierung unterstützt Denkprozesse

- ▶ Strategie: verwende externe Informationen [CMS1999, Seite 2]
- ▶ Interne und externe Repräsentationen und Prozesse sind miteinander verbunden
- ▶ Beispiel: Multiplikation $34 \cdot 72$

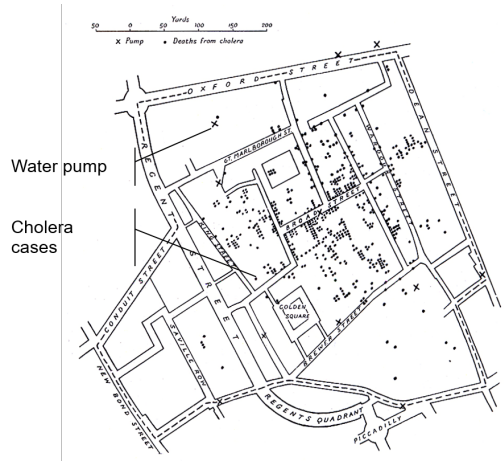
$$\begin{array}{rcl} 34 & \cdot & 2 & = & 68 \\ 34 & \cdot & 70 & = & 2380 \\ 34 & \cdot & 72 & = & \underline{2448} \end{array}$$



Ziele

Visualisierung unterstützt Analyse und Interpretation

- Epidemic cholera, 1845
[Tufte1983, page 24]



Ziele

- ▶ Hauptziele
 - ▶ Zusammenfassen
 - ▶ Verknüpfen
 - ▶ Klassifizieren
 - ▶ Gruppieren
 - ▶ Analysieren des zeitlichen Verhaltens
 - ▶ Vorhersagen

Ziele

[Tufte2001, S. 13]: Für numerische Daten sollten graphische Darstellungen

- ▶ die Daten zeigen
- ▶ viele Zahlen auf kleinem Raum präsentieren
- ▶ große Datenmengen zusammenhängend darlegen
- ▶ die Daten auf mehreren Detailebenen zeigen
 - ▶ vom breiten Überblick
 - ▶ zur Feinstruktur
- ▶ vermeiden die Aussagen der Daten zu verzerren
- ▶ das Auge zum Vergleich verschiedener Teile der Daten anregen
- ▶ ein klares Ziel verfolgen:
 - ▶ Beschreibung
 - ▶ Untersuchung
 - ▶ Tabellarische Aufstellung
 - ▶ Ausschmückung
- ▶ eng mit statistischen und verbalen Ausführungen verknüpft sein

Ziele

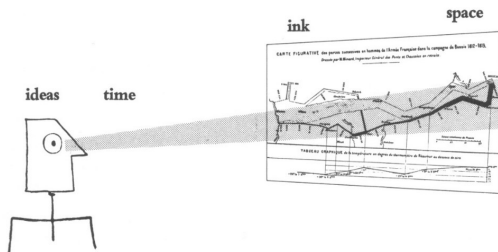
[Tufte2001, S. 13]

- ▶ Der Betrachter
 - ▶ soll über den Inhalt nachdenken
 - ▶ sollte **nie** nachdenken über
 - ▶ die Methode
 - ▶ das Design
 - ▶ die Graphiktechnologie

Ziele

[Tufte1983] (Graphische Exzellenz)

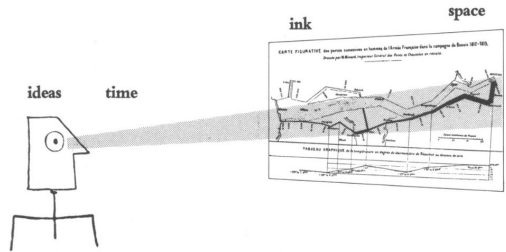
- ▶ Wohlgestaltete Präsentation interessanter Daten ist eine Frage von
 - ▶ Inhalt
 - ▶ Statistik
 - ▶ Gestaltung (Design)
- ▶ und besteht aus
 - ▶ komplexen Ideen
 - ▶ die mit Klarheit und Präzision effizient mitgeteilt werden



Ziele

[Tufte1983] (Graphische Exzellenz)

- ▶ Die Präsentation liefert dem Betrachter die
 - ▶ größte Anzahl an Ideen
 - ▶ in der kürzesten Zeit
 - ▶ mit der wenigsten Tinte
 - ▶ auf kleinstem Raum
- ▶ Es ist fast immer ein multivariates Problem
- ▶ Es erfordert, die Wahrheit über die Daten zu erzählen



Ziele

[Tufte1983] (Graphische Exzellenz)

- ▶ Charles Joseph Minard (1781–1870, französischer Ingenieur)
- ▶ Graphische Darstellung des Russlandfeldzuges von Napoleon von 1812–1813
- ▶ Multivariate Daten

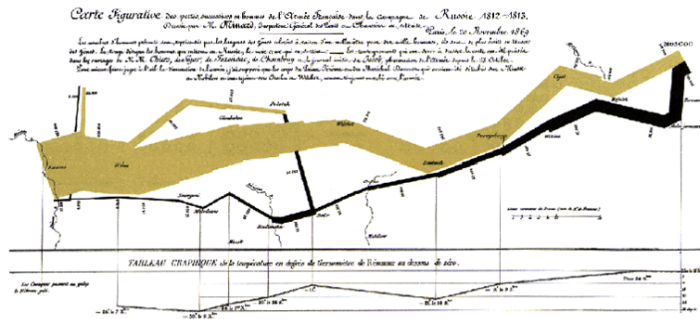


Bild 1. Joseph Minard: Karte des Russlandfeldzuges von Napoleon von 1812-1813.

Überblick

Verwendete Methoden und Techniken

▶ Aus der Statistik

- ▶ Beschreibende Statistik
- ▶ Vergleichende Statistik
- ▶ Normalisierung
- ▶ Fehler- und Abweichungsanalyse
- ▶ Statistische Sicherheit

▶ Aus anderen Bereichen

- ▶ Dimensionsreduktion
- ▶ Clustering
- ▶ Regression (linear und nicht-linear)
- ▶ Naïve Bayes
- ▶ Künstliche Neuronale Netze (KNN)
- ▶ Heuristische Methoden
- ▶ Verknüpfungsregeln
- ▶ Entscheidungsbäume
- ▶ Hauptkomponentenanalyse (PCA)

Weitere Beispiele

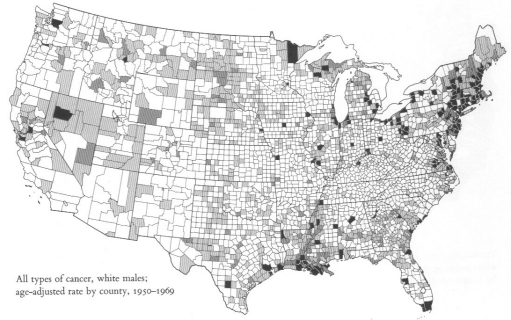
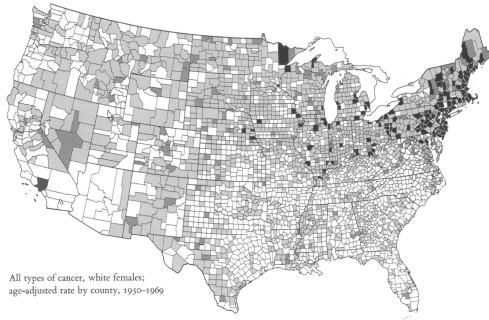


Abbildung: Datenkarten: 3056 Werte (Counties), 2D Position, mindestens vier Werte bezeichnen die Lage des Counties (Rand)
→ mindestens $7 \cdot 3056 = 21392$ Werte in einem Bild [Tuft1983]

Weitere Beispiele

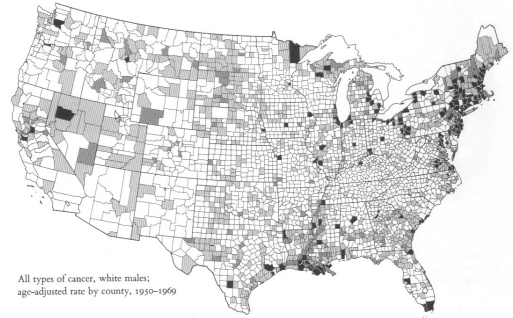
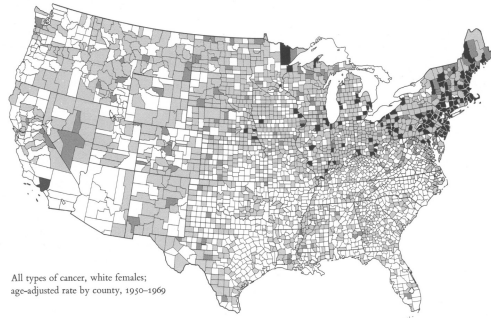


Abbildung: Vergleich zweier Datensätze: links: Krebs bei weißen Frauen, rechts: Krebs bei weißen Männern

Weitere Beispiele

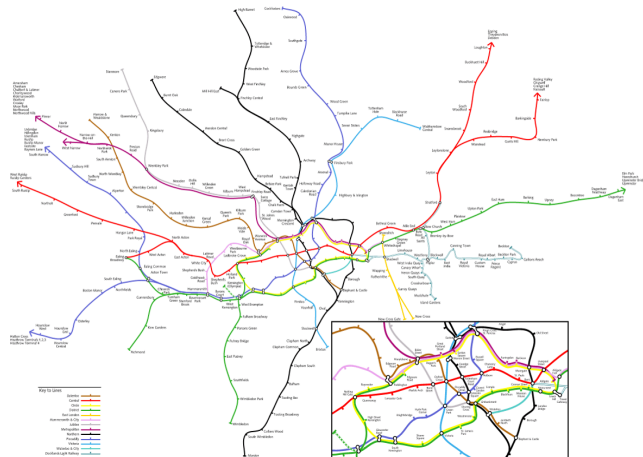


Abbildung: The Tube: Geography [<http://www.kottke.org/plus/misc/images/tubegeo.gif>]

Weitere Beispiele



Abbildung: The Tube: Beck's Map, 1931 [Spence2001], p. 2, 8

Weitere Beispiele



Abbildung: The Tube: Schematic

Weitere Beispiele

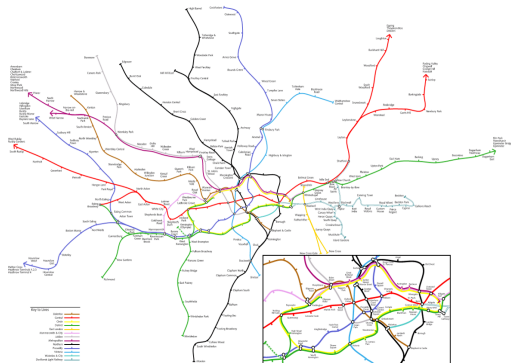


Abbildung: The Tube: Geography & Schematic [<http://www.kottke.org/plus/misc/images/tubegeo.gif>]

Literatur

Klassische Lehrbücher

- [Tufte1983] E.R. Tufte.
The Visual Display of Quantitative Information.
2. Auflage, Graphics Press, Cheshire, CT, USA, 1983, 2001
- [Tufte1990] E.R. Tufte.
Envisioning Information.
Graphics Press, Cheshire, CT, USA, 1990
- [Tufte1997] E.R. Tufte.
Visual Explanations.
Graphics Press, Cheshire, CT, USA, 1997
- [Tufte2006] E.R. Tufte.
Beautiful Evidence.
Graphics Press, Cheshire, CT, USA, 2006

Literatur

Klassische Lehrbücher

[Bertin1982] J. Bertin.
Graphische Darstellungen.
De Gruyter, Berlin, 1982.

Literatur

Moderne Lehrbücher

- [Spence2001] R. Spence.
Information Visualization.
Addison-Wesley, Reading, MA, USA, 2001.
- [WGK2010] M. Ward, G. Grinstein, D. Keim.
Interactive Data Visualization.
A K Peters, Ltd., ISBN 978-1-56881-73-5, 2010.

Literatur

Moderne Lehrbücher

- [Ware2012] C. Ware.
Information Visualization – Perception for Design
3. Auflage, Elsevier, Amsterdam, NL, 2012.
- [Ware2008] C. Ware
Visual Thinking: For Design.
Morgan Kaufman, San Francisco, 2008.

Literatur

Moderne Lehrbücher

- [Few2004] S. Few.
Show Me the Numbers, Designing Tables and Graphics to Enlighten.
Analytics Press, 2004.
- [Few2006] S. Few.
Information Dashboard Design: The Effective Visual Communication
of Data.
O'Reilly Media, 2006.

Literatur

Wichtige Konferenzen

- ▶ IEEE Symposium on Information Visualization (IEEE InfoVis)
 - ▶ jährlich seit 1995
 - ▶ ab 2008 IEEE Information Visualization Conference
 - ▶ ab 2009 IEEE Information Visualization
- ▶ IEEE Symposium on Visual Analytics Science and Technology (IEEE VAST)
 - ▶ jährlich seit 2006
 - ▶ ab 2009 IEEE Visual Analytics Science and Technology
- ▶ seit 2021
 - ▶ IEEE VIS
 - ▶ umfasst alle ehemaligen Visualisierungskonferenzen:
 - ▶ Scientific Visualization
 - ▶ Information Visualization
 - ▶ Visual Analytics

Literatur

Wichtige Konferenzen

▶ EuroVis

- ▶ jährlich seit 2005
- ▶ 1990–1998: EG ViSC Workshop on Visualization Scientific Computing (Vorgänger)
- ▶ 1999–2004: VisSym Symposium on Visualization

▶ PacificVis

- ▶ jährlich seit 2008

Literatur

Wichtige Fachzeitschriften

- ▶ Information Visualization Journal
 - ▶ Sage Publications, seit 2002
- ▶ IEEE Transactions on Visualization and Computer Graphics
 - ▶ seit 1995
- ▶ IEEE Computer Graphics and Applications
 - ▶ seit 1981