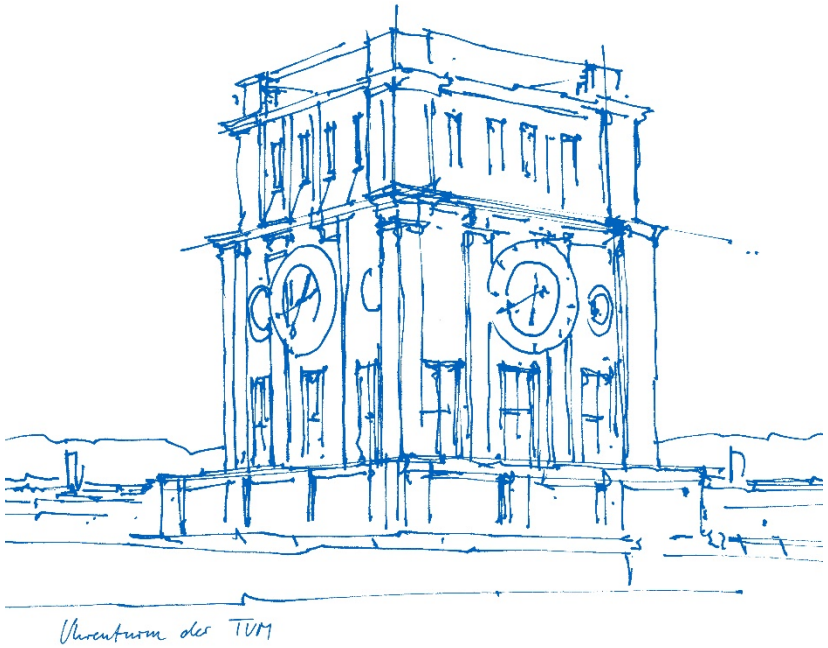
Visualisierung

Jürgen Pfeffer

Technical University of Munich

Bavarian School of Public Policy

juergen.pfeffer@tum.de | @JurgenPfeffer



Why Pictures?

1. Parallel processing of pictures

Language and writing required sequential coding

Graphical communication has a high „bandwidth“

2. Relational cognition of the human brain

We think in pictures

Mental models

Why Pictures?

“Efficient communication of information” [Tufte 2001]

Graphical representation of multivariate (high dimensional) data [Tufte 2001]:

- visual evidence
- visual reasoning
- visual understanding

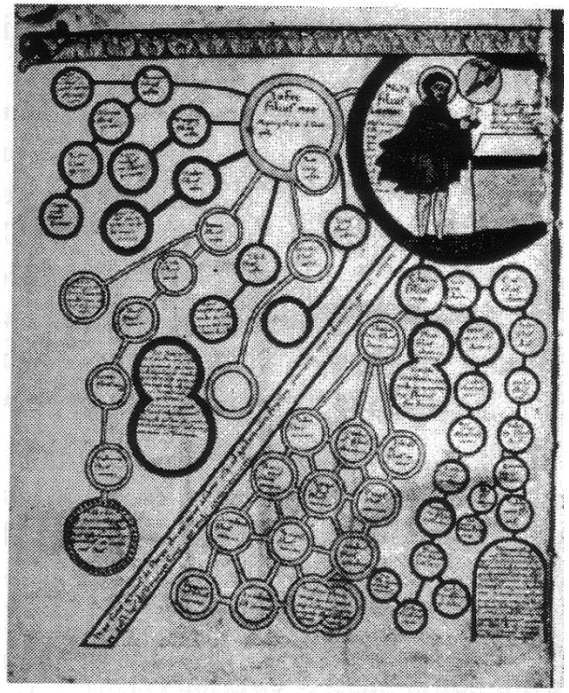
„Effective translation of information to a system of visual elements“ [Bertin 1984]

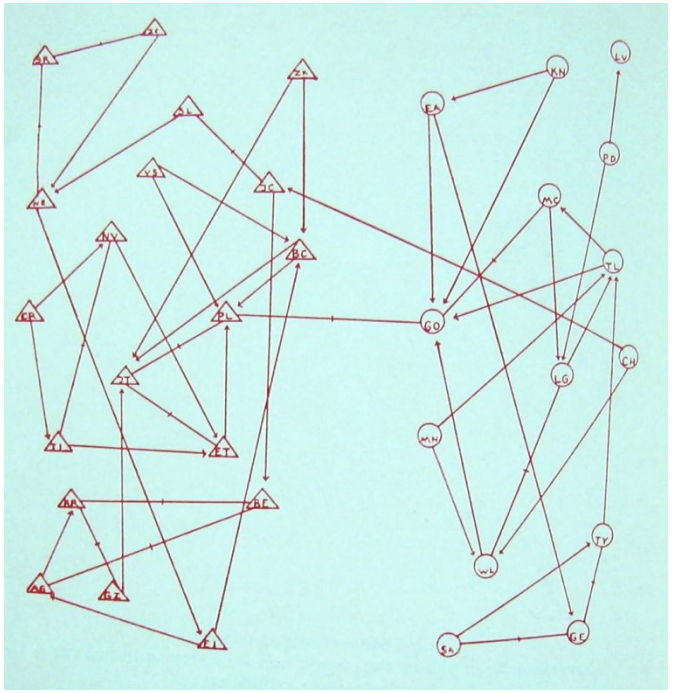
„The faster the information is understood, the more effective is the visualization“ [Krempel 2005]

Network Visualizations

Family Trees (medieval)

Sociometry, Moreno (1934)





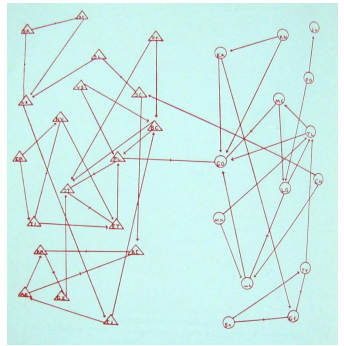
Visualizing Networks

Explorative visualizations – find something

- First impressions of the data
- Validate your network data
- Find new research questions

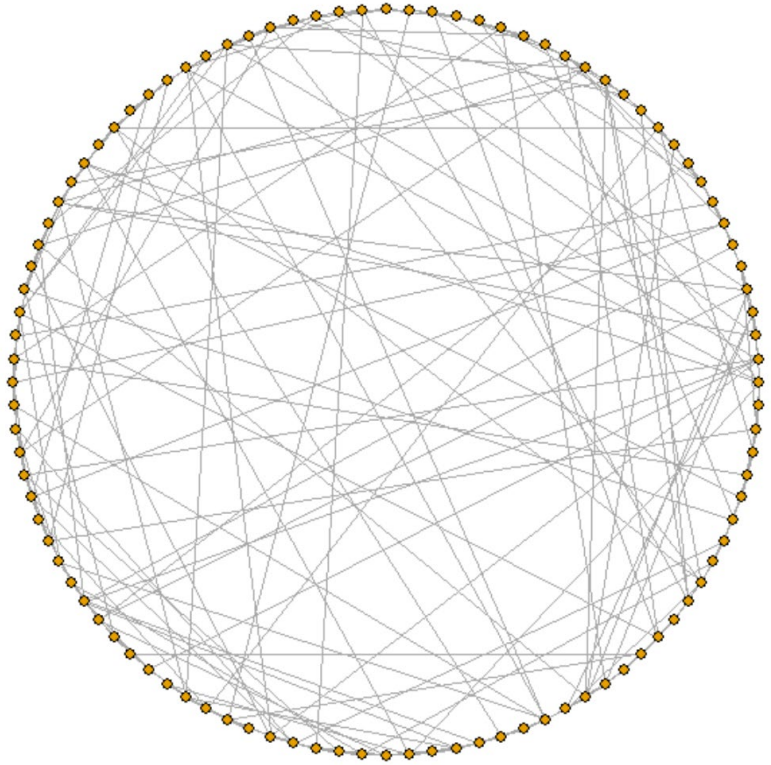
Information visualization – show something

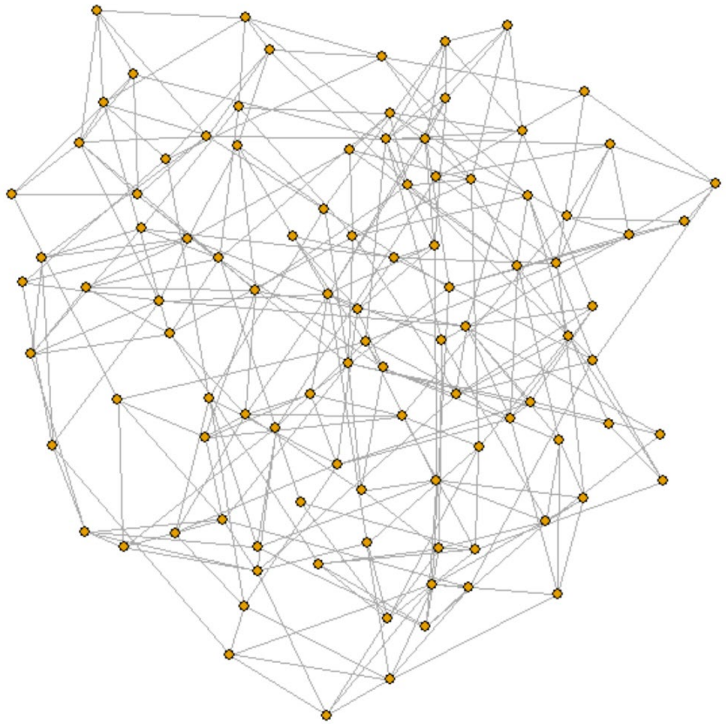
- What is the information that you want to visualize (substance)?
- How is it possible to represent this information with your network in a useful way (design)?
- How to realize this with satisfying approaches (algorithm)?



“It is first of all a method of exploration”, (J. L. Moreno 1953, 96)

Was ist eigentlich “layout_with_fr(g)”





Was ist eigentlich “layout_with_fr(g)”

Layout Algorithmus

„Graph Drawing by Force-Directed Placement”

Thomas M.J. Fruchterman und Edward M. Reingold, 1991

Weit verbreitet in Netzwerk Tools

Populär weil intuitive verständlich

Kriterien für das Erzeugen von Netzwerklayouts

Gleichmäßige Verteilung der Knoten innerhalb der Zeichenfläche

Minimierung der Kantenkreuzungen

Gleichmäßige Kantenlängen

Wiedergabe der inhärenten Graphsymmetrien

Algorithmus bearbeitet diese Kriterien nicht explizit, werden aber implizit optimiert (mit Ausnahme der Kantenkreuzungen)

Fruchterman/Reingold

Metapher aus der Physik: force-directed placement

Der Ablauf des Algorithmus basiert auf zwei Prinzipien:

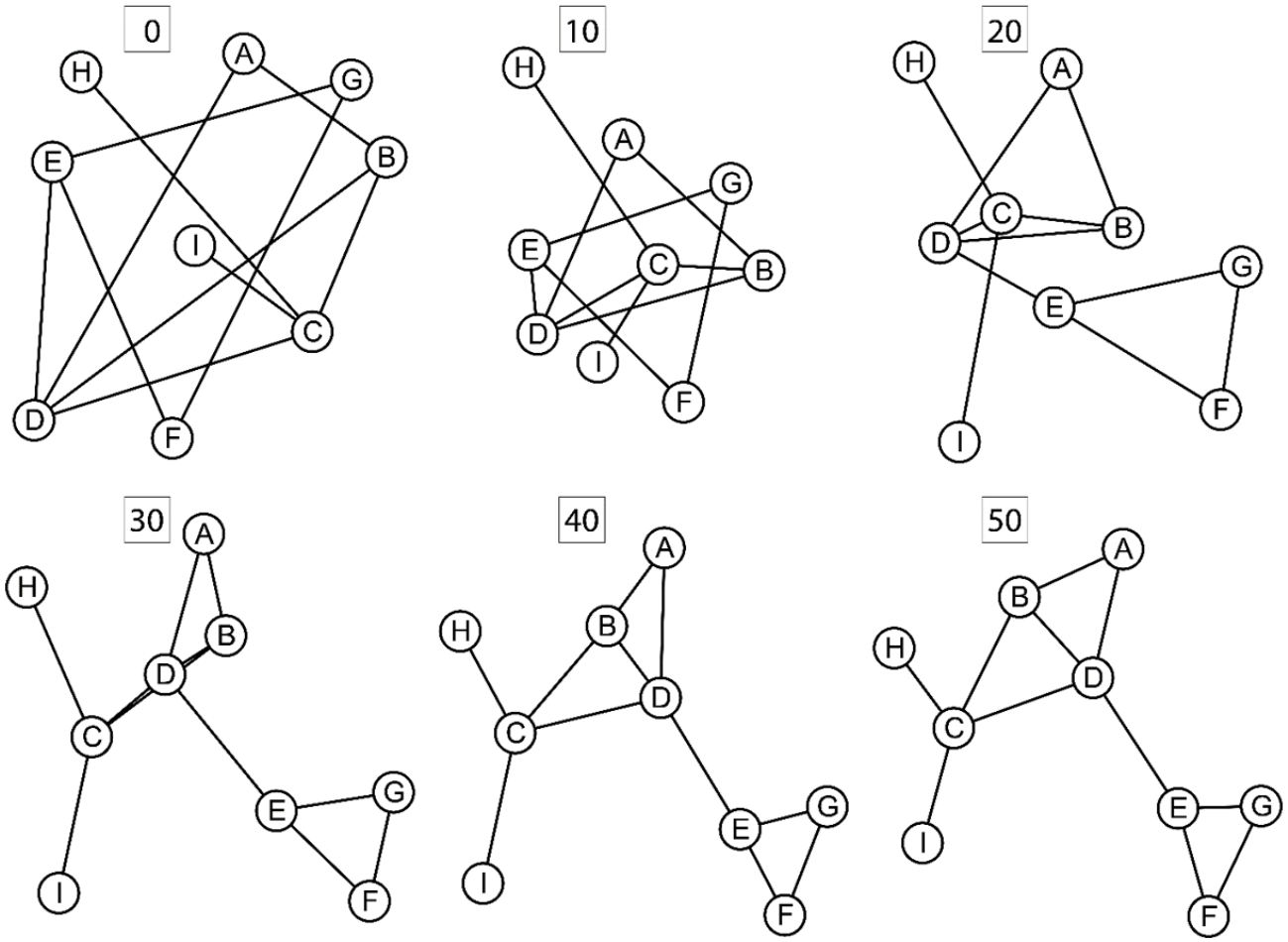
- Knoten, die durch eine Kante miteinander verbunden sind, sollen nah beieinander platziert werden.
- Generell aber sollen Knoten nicht *zu* nah an anderen Knoten platziert werden

Algorithmische Umsetzung: Berechnung von Anziehungs- und Abstoßungskräften:

1. Berechnen des Effekts der gegenseitigen Abstoßung von Knoten
2. Berechnen des Effekts der gegenseitigen Anziehung unter den Knoten, die durch eine Kante verbunden sind
3. Neupositionierung der Knoten unter Einbeziehung beider Effekte, wobei die maximale Distanz, die ein Knoten bewegt werden darf, begrenzt ist

Iteratives Vorgehen: keine Abbruchbedingung. Realität: 50-100 Durchläufe

Fruchterman/Reingold: Beispiel



Probleme

Algorithmische Komplexität: Sehr lange Berechnungszeiten für Netzwerke mit >1.000 Knoten

Lokale Optimierungen: Wiederholung auf gleichem Netzwerk erzeugt unterschiedliche Layouts

Kamada und Kawai (1989)

Kamada und Kawai (1989)

Der ideale Abstand zwischen zwei Knoten ist proportional zur Länge der kürzesten Pfaddistanzen zwischen diesen Knoten.

Vorteil: Bezieht Struktur des Graphen mit ein

Nachteil: Nochlangsamer als F/R

Multidimensionale Skalierung

Distanzmatrix des Netzwerks wird mit einer zweidimensionalen Matrix von euklidischen Abständen bestmöglich angenähert

Basiert auf Eigenvektorberechnungen

Vorteile: Eindeutige globale Lösung. Ansätze für große Netzwerke

Nachteile: Mathematisch aufwendig.

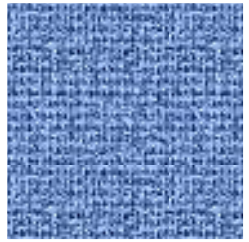
Problem: Knotenüberlappungen:

- Knoten mit gleicher Struktur (gleichen Nachbarn) liegen exakt übereinaner
- Lösungsansatz: MDS + F/R

Perception

Preattentive perception…

Request for attention!

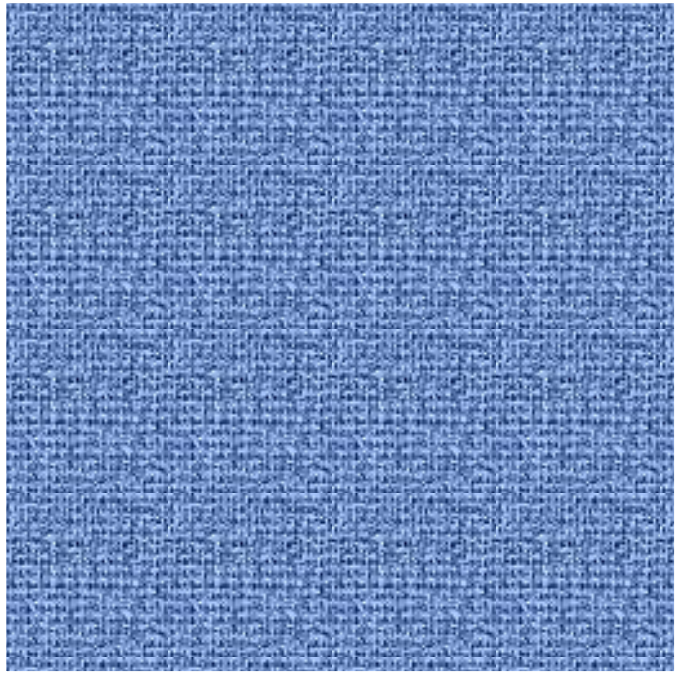


What Did You See?

Preattentive Perception

Preattentive elements:

- **Position**
- Size
- Shape
- Color
- Saturation
- Texture



Preattentive Perception

Preattentive perception:

- Unconscious collection of information
- Nervous system can react, no brain activity
- All information we see, hear,…
- 200-250 msec.

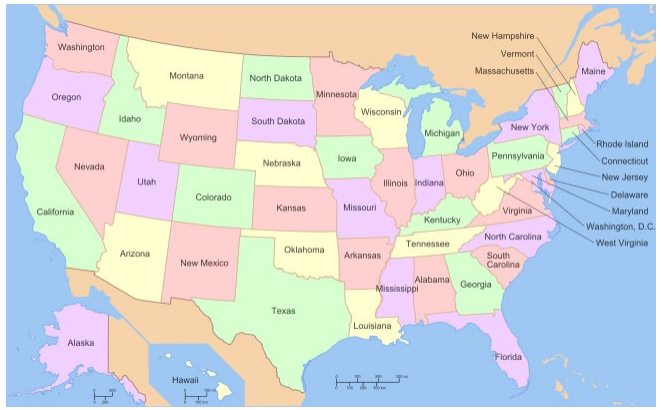
Attentive perception

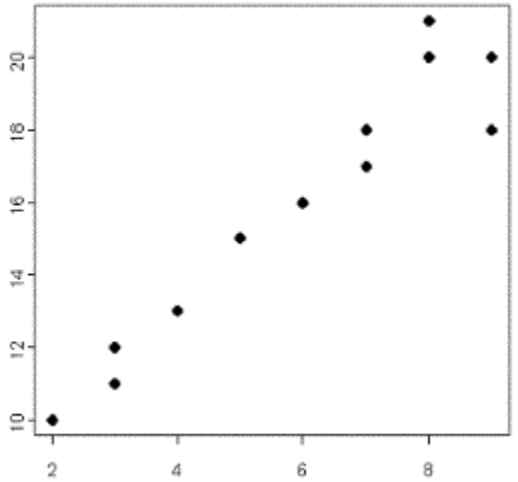
- Conscious processing of information
- Analyzing and interpretation

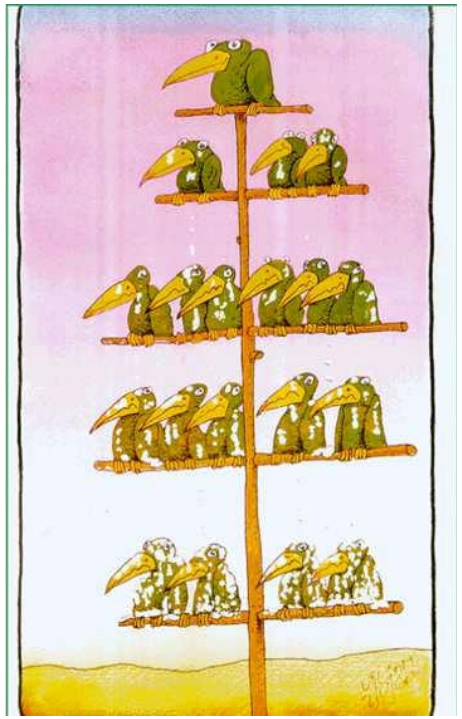
Position

x-axis, y-axis of elements

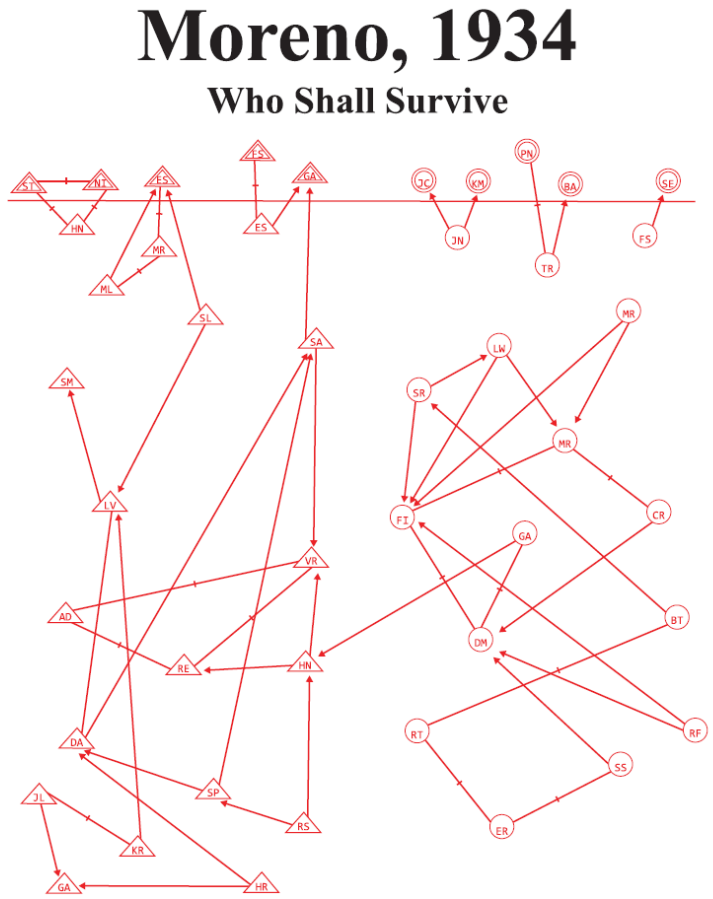
- Left, right, top, bottom
- Central, peripheral







Position to Show Structure

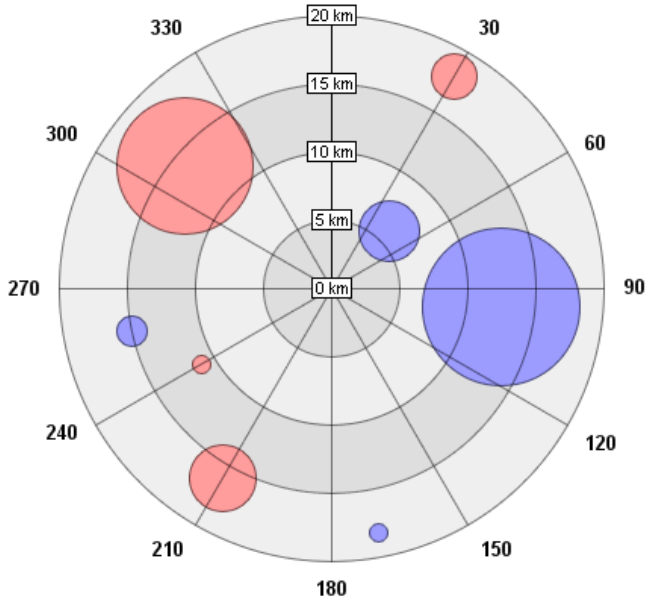


Class structure, 5th grade. Girls (circles) and boys (triangles).
Links show two best friends. Top line defines group border.

Area/Volume

Size of elements

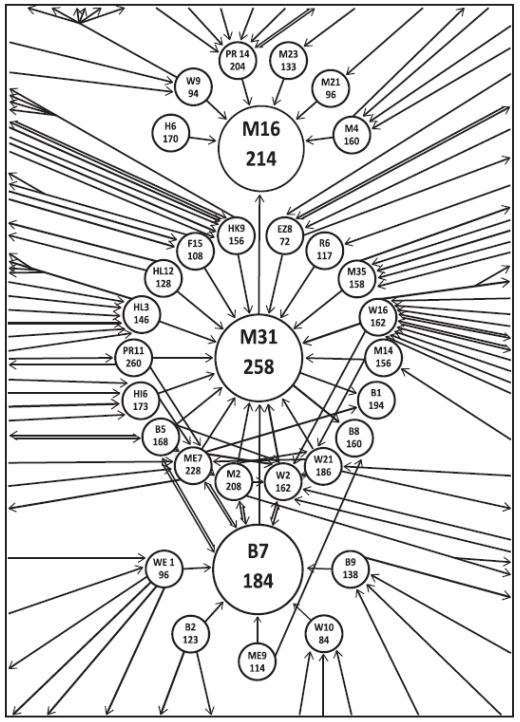
- Larger, smaller



Node Size to Show Importance

Lundberg & Steele, 1938

Social Attraction-Patterns in a Village

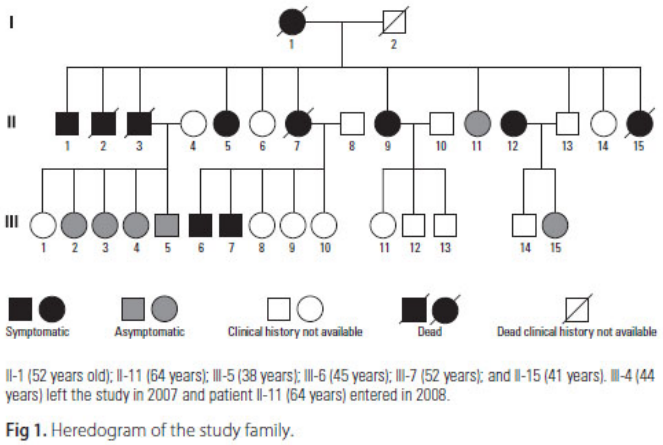


. Most imortant friendships in the village. Number is socio-economic status. M31 = "Lady Bountiful".

Shape

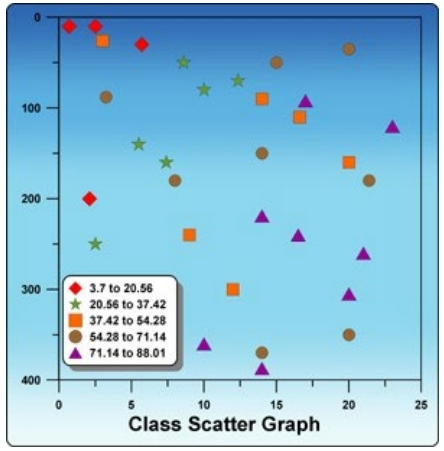
The form of the elements

- Squares, circles, triangles,…



II-1 (52 years old); II-11 (64 years); III-5 (38 years); III-6 (45 years); III-7 (52 years); and II-15 (41 years). III-4 (44 years) left the study in 2007 and patient II-11 (64 years) entered in 2008.

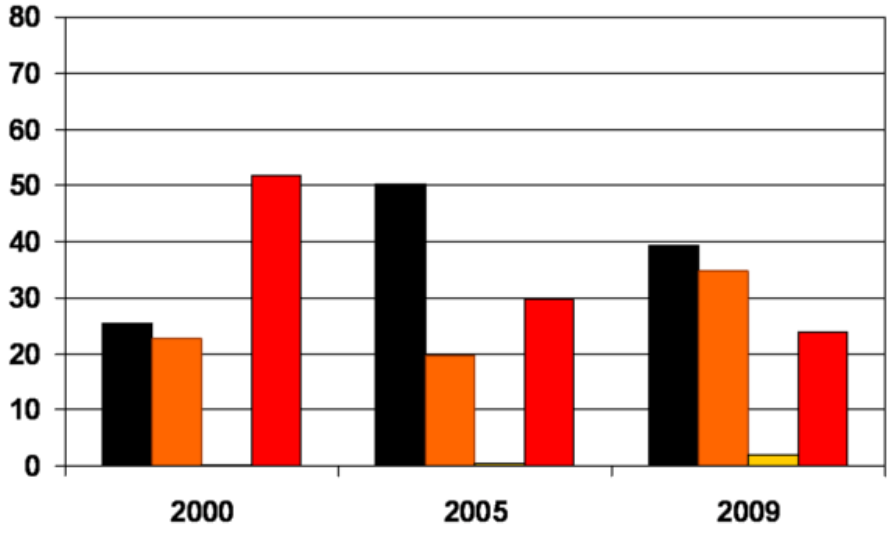
Fig 1. Heredogram of the study family.

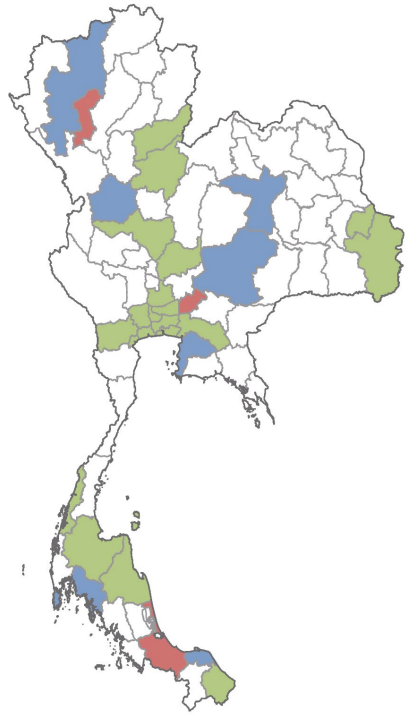


Color Hue

Color of elements

- Red, black, blue,…

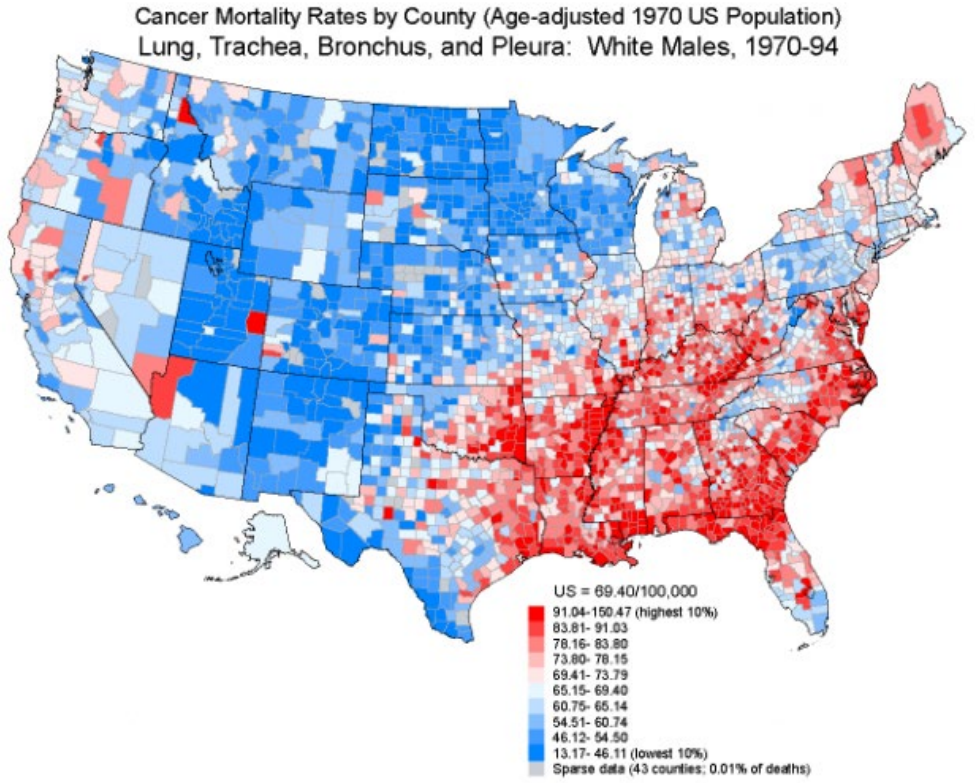




Color Saturation

Saturation of colors of elements

- Light, dark, color gradient

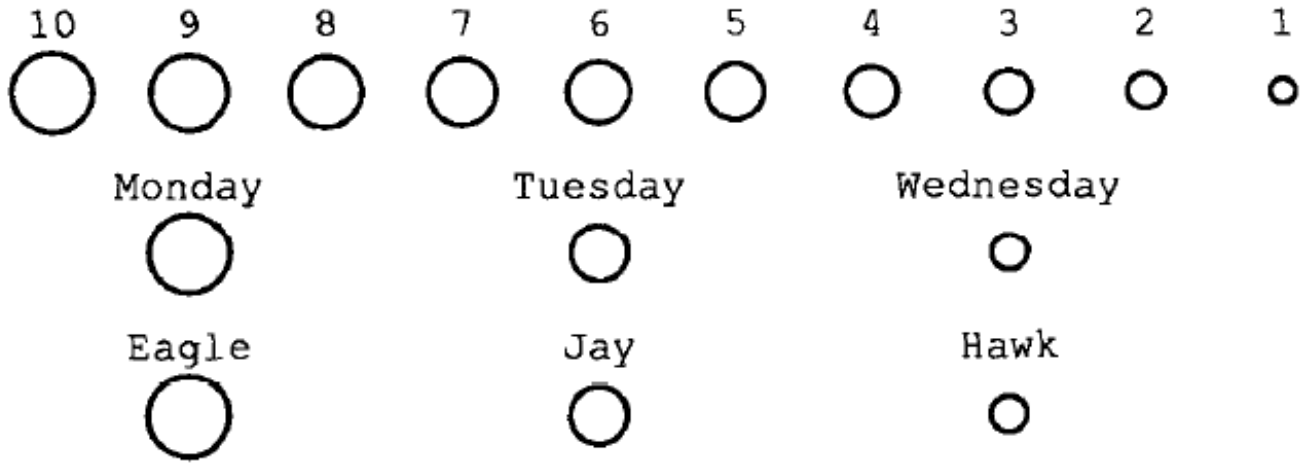


Elements and Data Type

Quantitative data

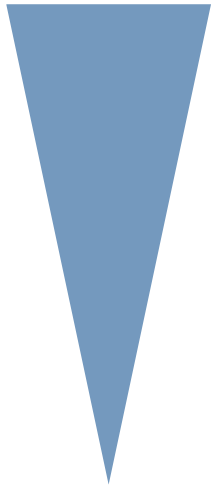
Ordinal data

Nominal data



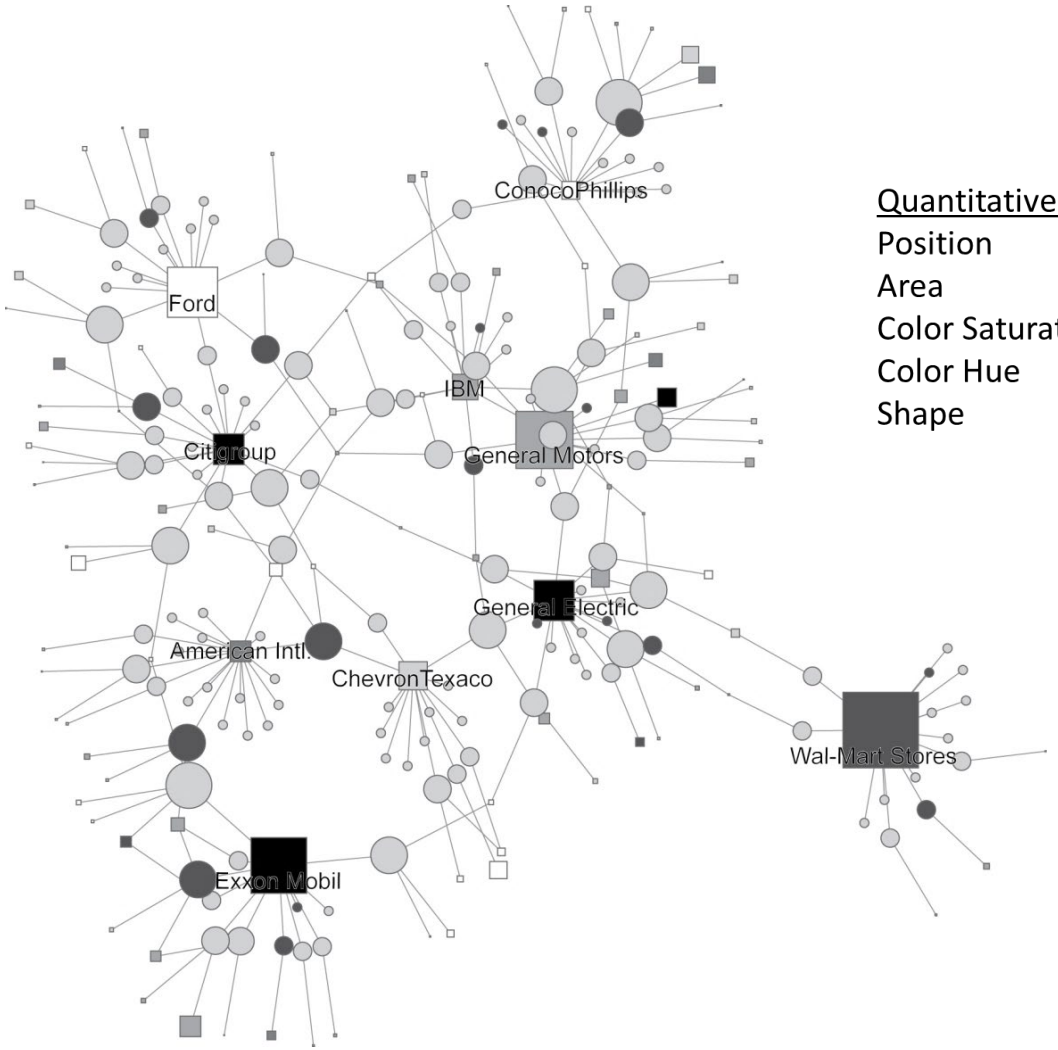
[Mackinlay 1986]

Relevance of Elements



Quantitative	Ordinal	Nominal
Position	Position	Position
Area	Color Saturation	Color Hue
Color Saturation	Color Hue	Color Saturation
Color Hue	Area	Shape
Shape	Shape	Area

Multivariate Network Visualizations

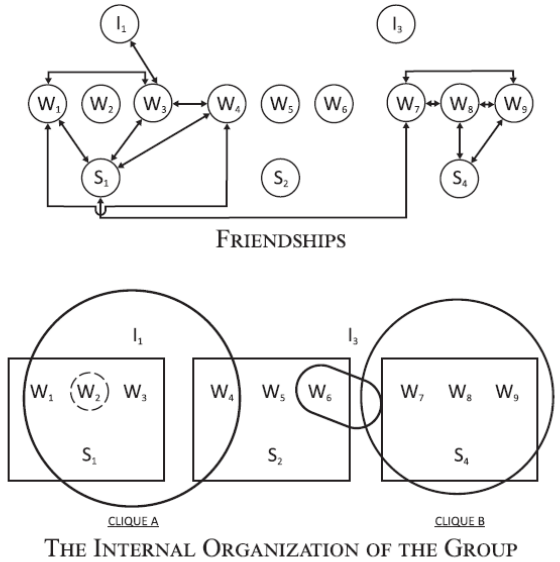


Quantitative	Ordinal	Nominal
Position	Position	Position
Area	Color Saturation	Color Hue
Color Saturation	Color Hue	Color Saturation
Color Hue	Area	Shape
Shape	Shape	Area

Substance-Based Layout

Roethlisberger et al., 1939

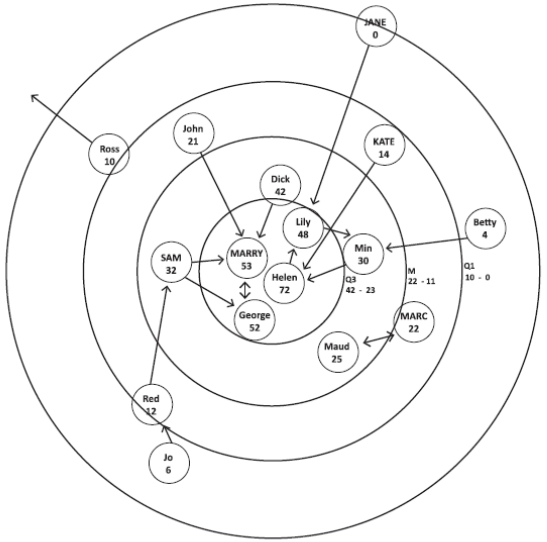
Management and the Worker



Observed friendship ties and cliques in a factory. Position reflects the location of their workspace.

Northway, 1940

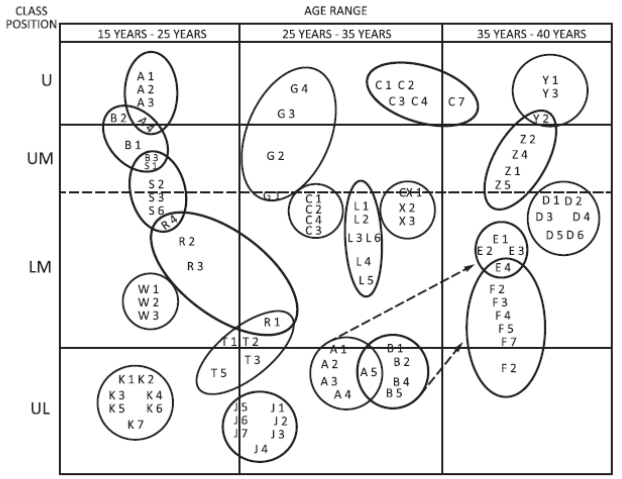
A Method for Depicting Social Relationships Obtained by Sociometric Testing



Targeted diagram showing scores of acceptability and predominating choices in a social group.

Davis et al., 1941

Deep South: A Social Anthropological Study of Caste and Class



Subgroups and their overlaps. Arrangement in terms of both social class and age reveals groups.

Graphical Excellence

Graphical excellence [Tufte 2001]…

is a matter of substance, statistics, and design

consists of complex ideas communicated with clarity, precision, and efficiency

give the viewer the greatest number of ideas in shortest time with the least ink in the smallest space

is nearly always multivariate

requires telling the truth about the data

induce the viewer to think about the substance rather than the methodology

→ Above all **show the data**

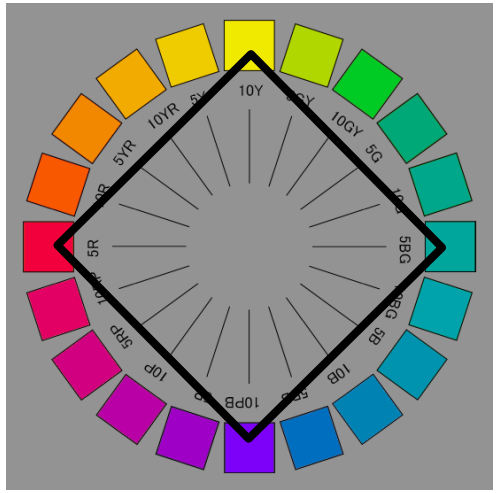
Smart Use of Colors

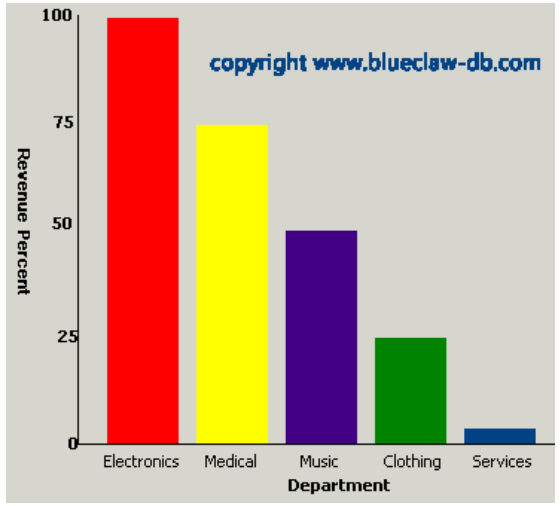
Just use colors when they carry additional information

Color hue and saturation are used for different data

Colors often have meaning

But, learn to visualize without colors





And Finally…

Explain what you did

- Describe mapping of data to visual elements
- Use a legend or caption

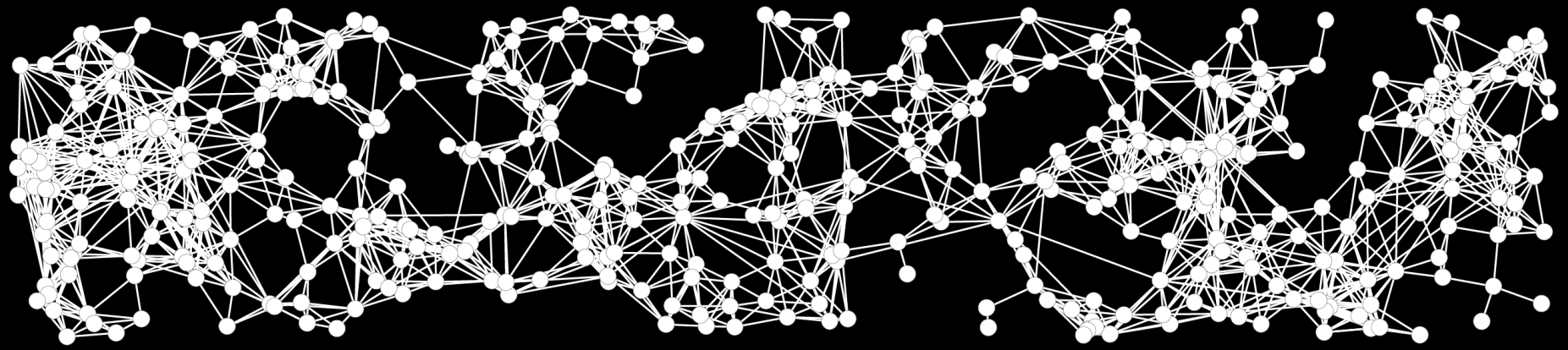
Be consistent across visualizations

- Easier to memorize and recognize repeated designs
- Find your style – a good one

It is all about the story

- Narrative quality of the visualization

“Our mission is to go forward, and it has only just begun.
There's still much to do, still so much to learn. Engage!”
Jean-Luc Picard, Star Trek TNG, Season 1 Episode 26



Juergen.Pfeffer@tum.de @JurgenPfeffer
Mirco.Schoenfeld@tum.de