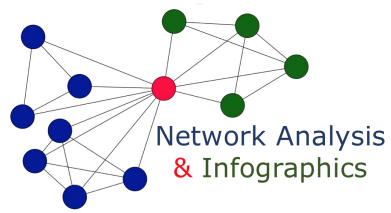
# Principles of Infographics

Dr Daniele Rotolo

SPRU (Science Policy Research Unit)
Business School
University of Sussex




*Week 7*

# Learning Outcomes

| | Learning outcome | Assessment mode |
|---|---|---|
| 1 | Explain the concept of network and list the main network indicators | ESS |
| 2 | Describe and apply the major techniques for the collection of network data and their statistical analysis | ESS, GPN + GWS |
| 3 | Identify the main characteristics of networks by means of network measures | ESS, GPN + GWS |
| 4 | Employ network analysis techniques to produce network data-based infographics | GPN + GWS |

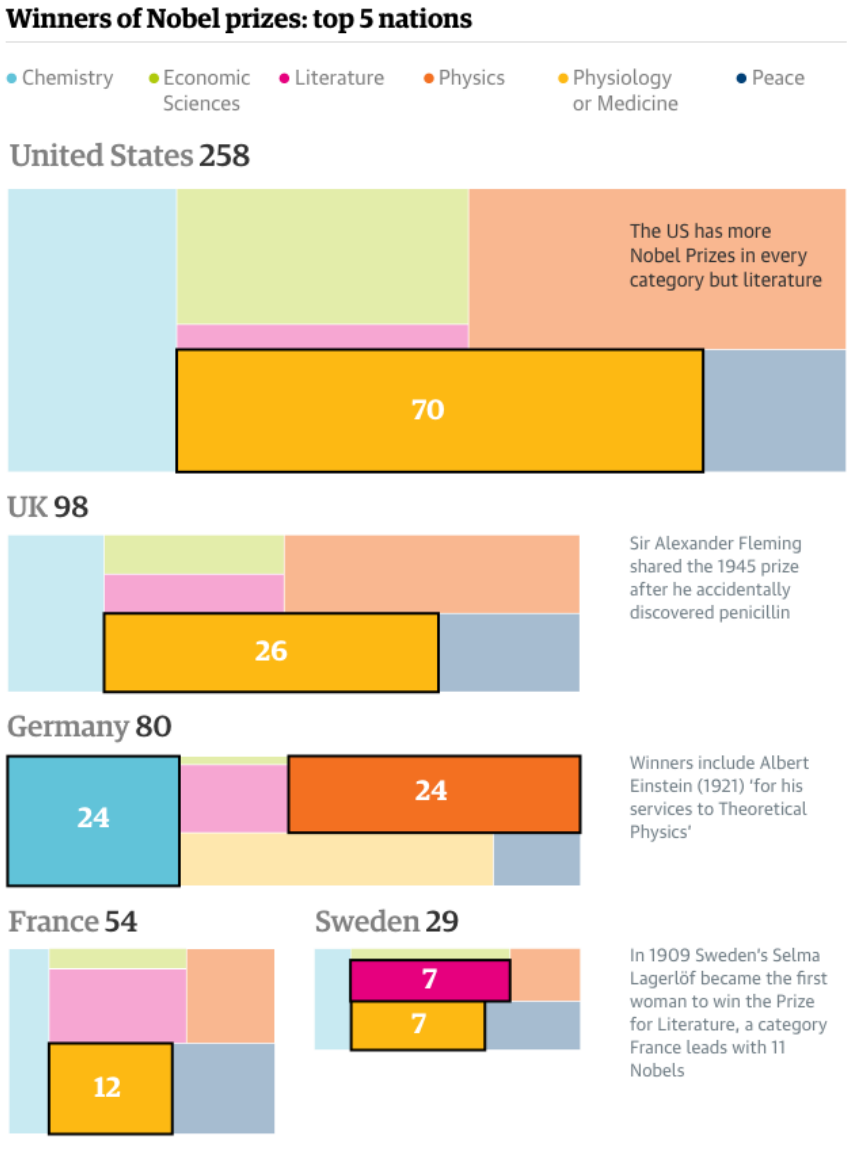
Note: ESS: Essay; GPN: Group Presentation; GWS: Group Written Submission

# Overview

# Why infographics?

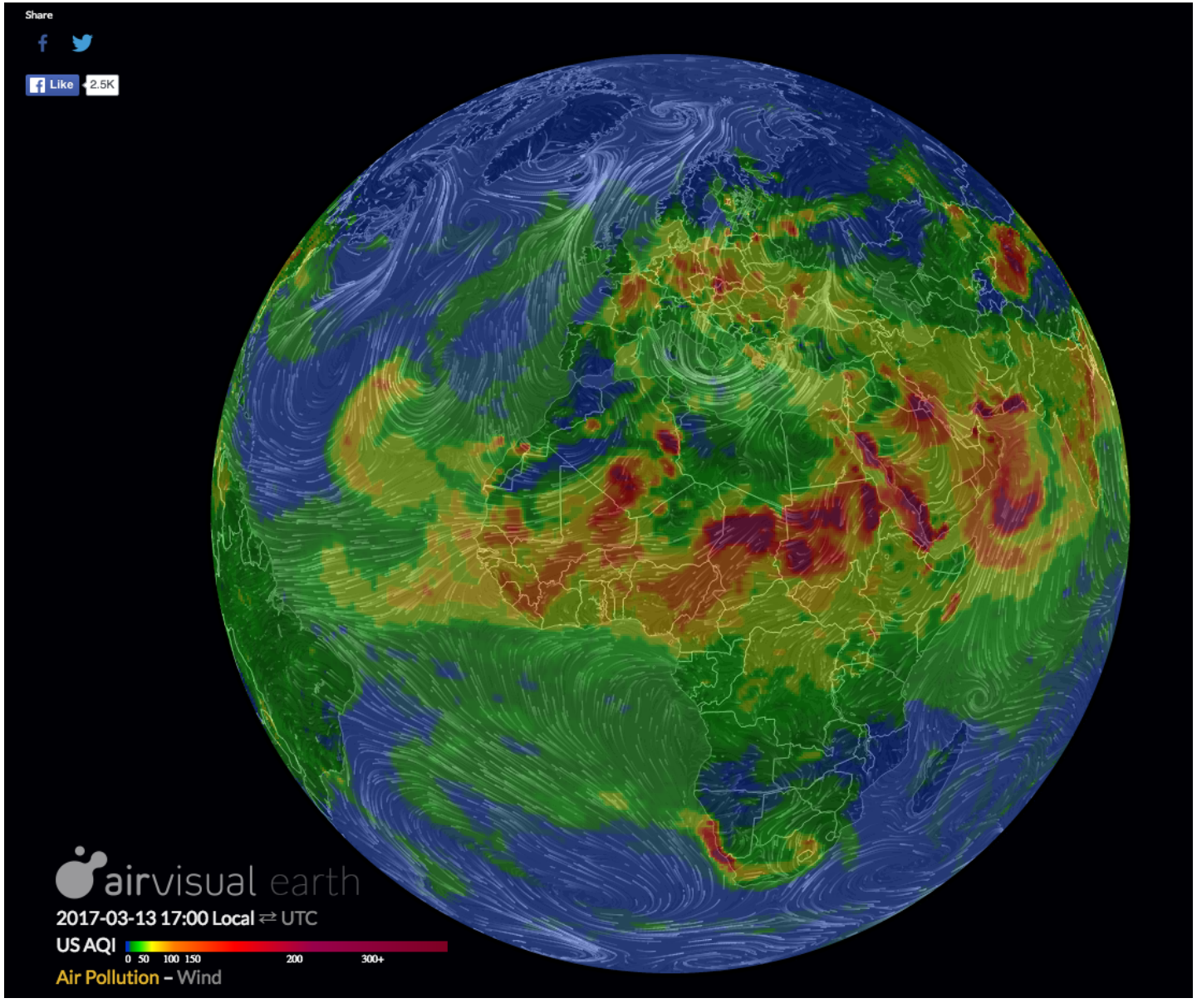
# Why infographics?

- Half of our brain is dedicated to processing visual signals (we can process images 60,000 times faster than text)
- Increasing capabilities to collect, store and analyse data (e.g. government data, APIs, Big Data, Data Science)
- Risk of information overload



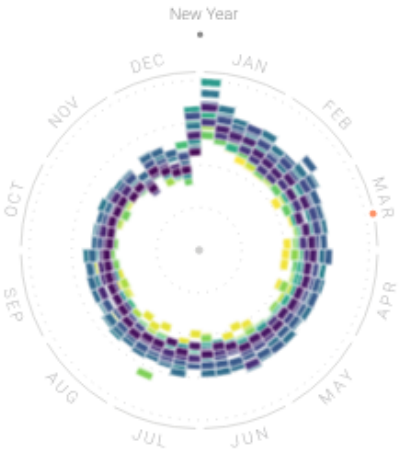

Source: The Guardian (October, 2016)

# Why infographics?

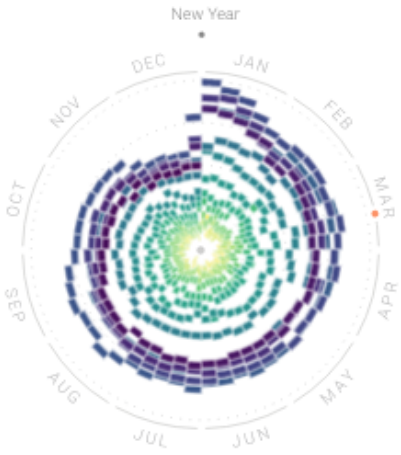



Source: Air pollution (March, 2017) [https://airvisual.com/earth]

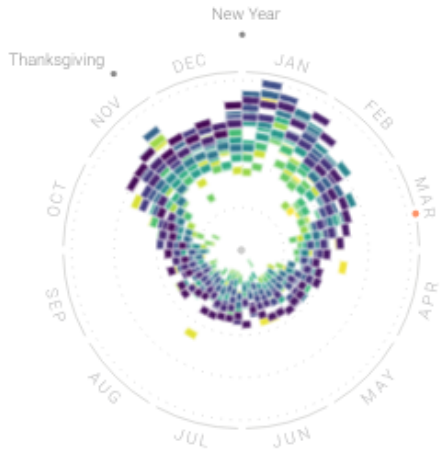
# Why infographics?



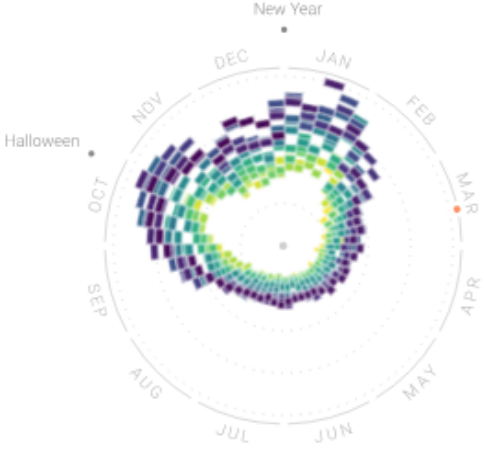

DIET



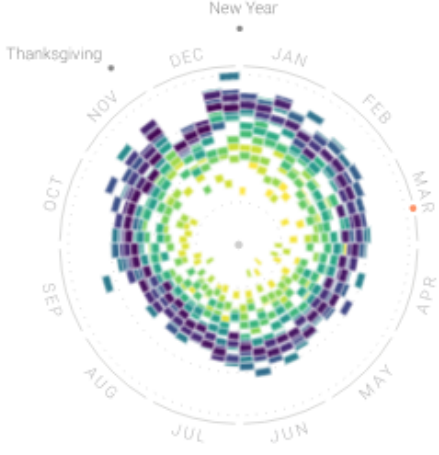

QUINOA



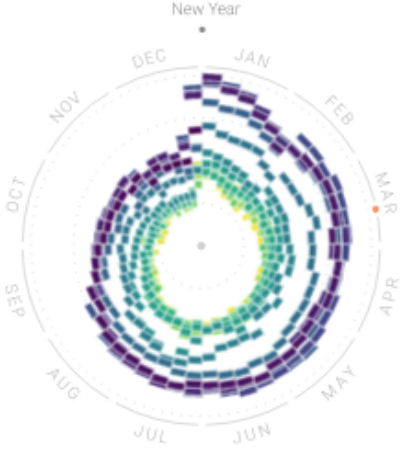

PUMMELO




CHILI CON CARNE



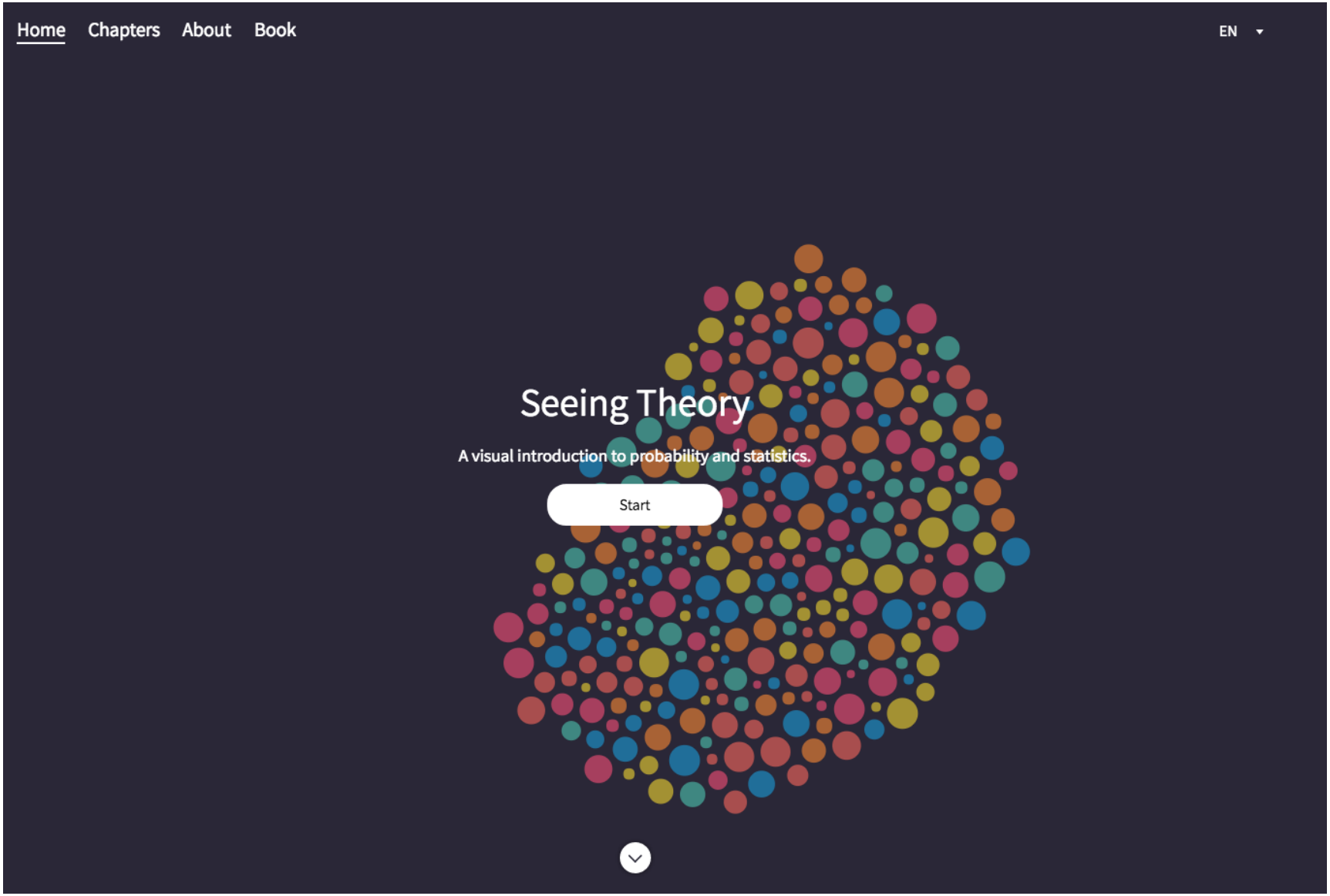

ENDIVE




SMOOTHIE

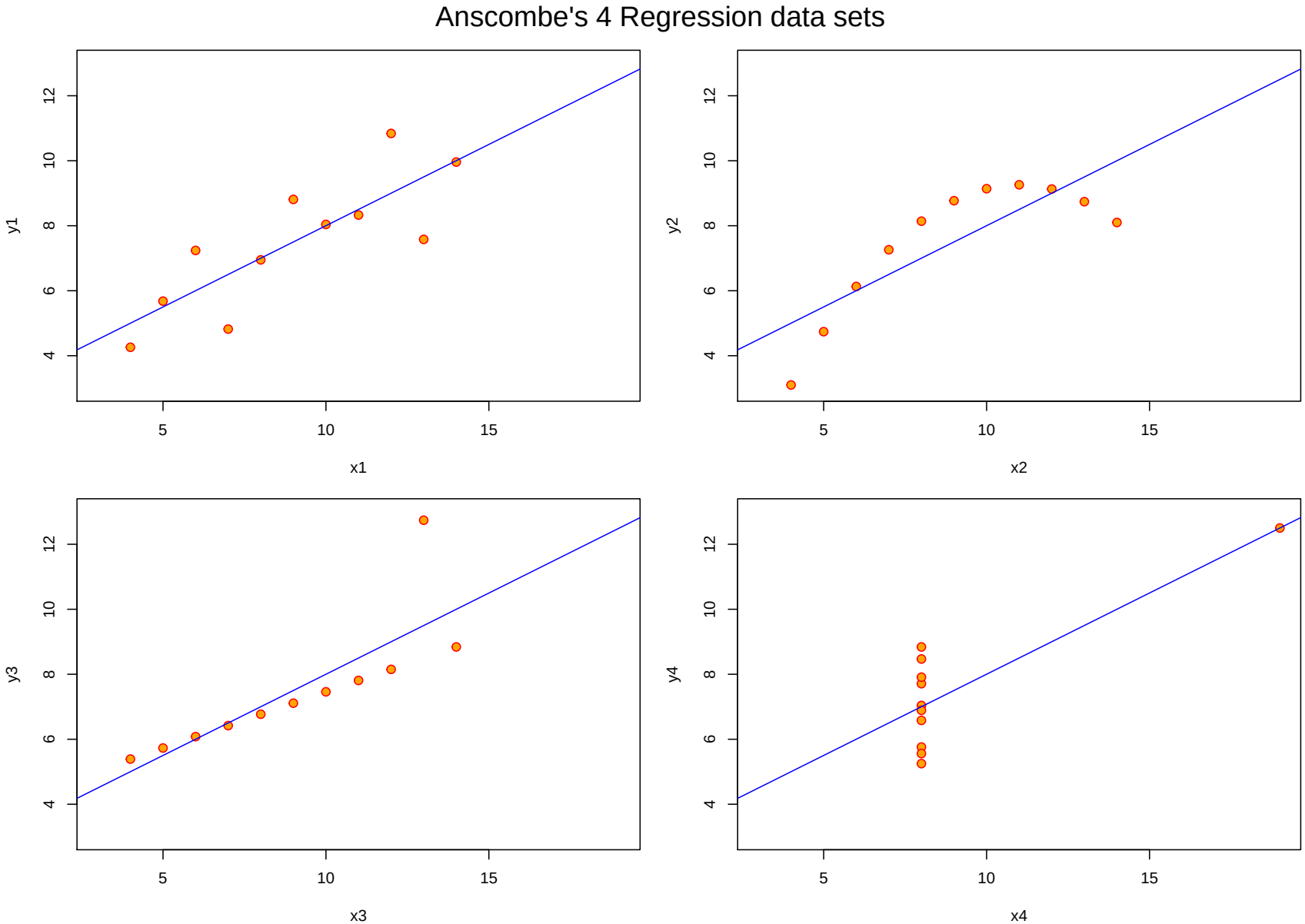
Source: Rhythm of food [`http://rhythm-of-food.net`]

# Why infographics?




Source: Seeing theory [http://students.brown.edu/seeing-theory]

# Why infographics?



# Why infographics?

- Raw data do not tell as much
- Statistics and visualisation help us to go beyond raw data (patterns, relationships)
- Infographics helps us to tell a story, to trigger emotions and actions


Source: Hans Rosling's 200 Countries, 200 Years (`www.youtube.com/watch?v=jbkSRLYSojo`)

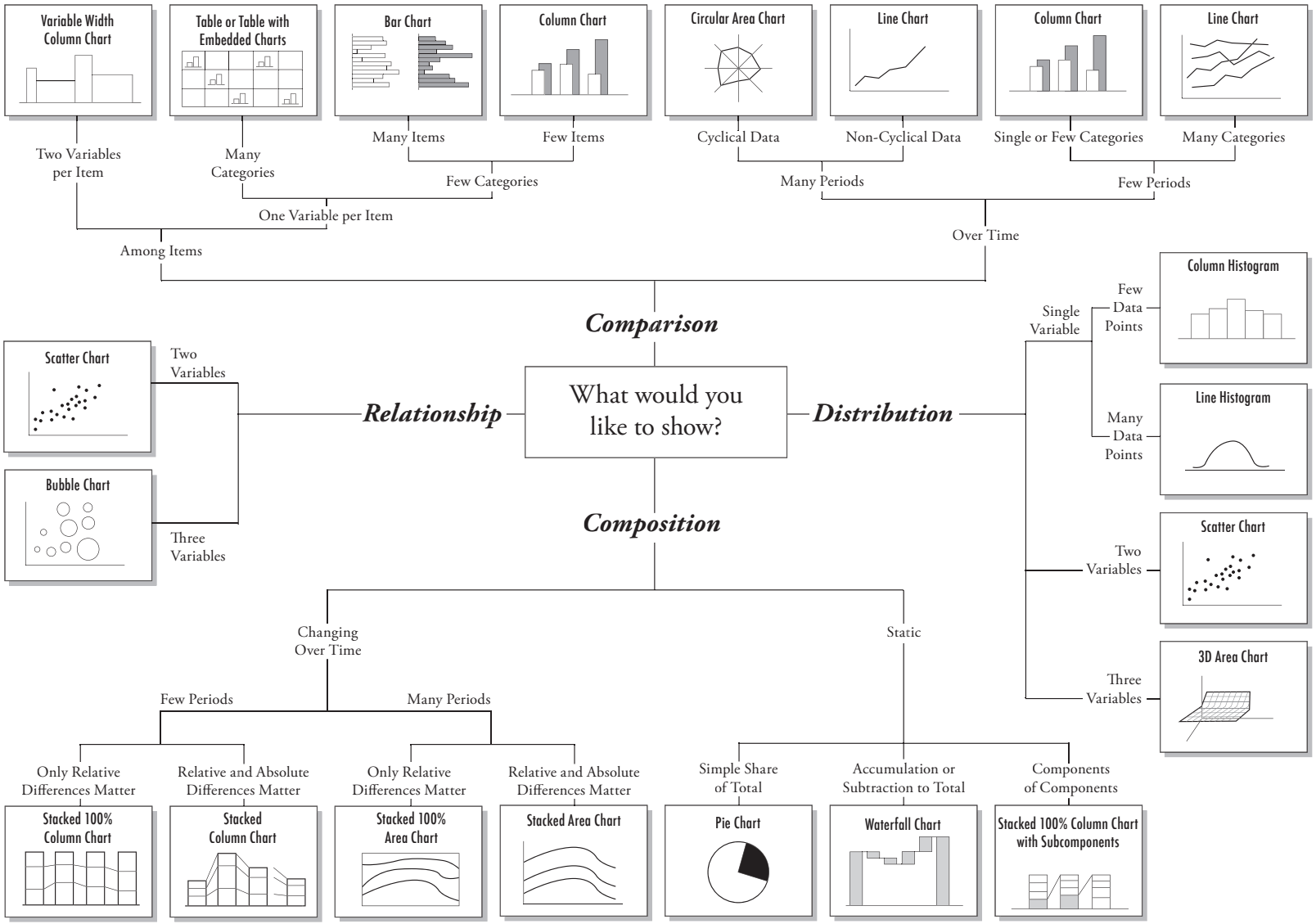
# Designing infographics

Check-list when producing an infographic

- Use appropriate charts to visualise your data
- Explain encoding (e.g. legend)
- Label axis
- Include sources (e.g. figures, data)
- Maximise the data-ink ratio
- Ensure graphical integrity

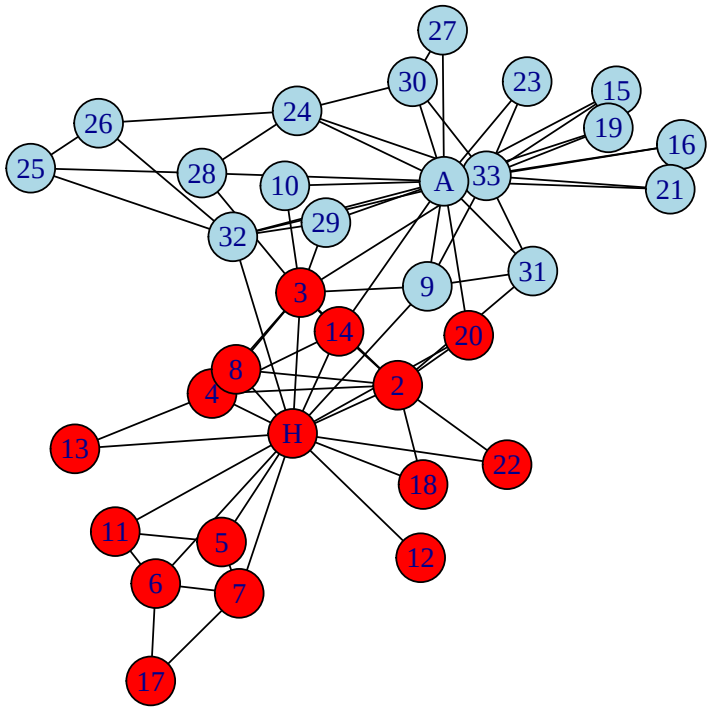
# Designing infographics

## Use appropriate charts



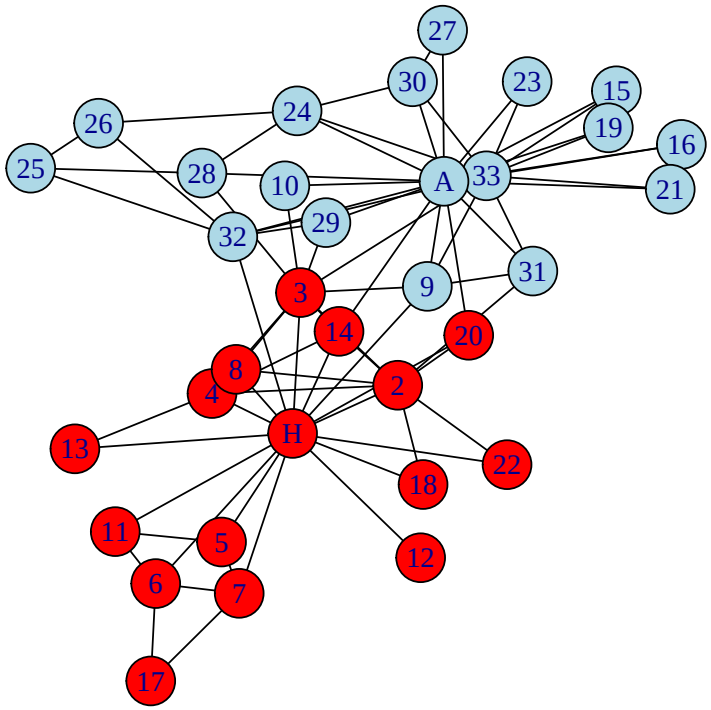

Source: Andrew Abela [https://extremepresentation.com/]

# Designing infographics
## Explain encoding

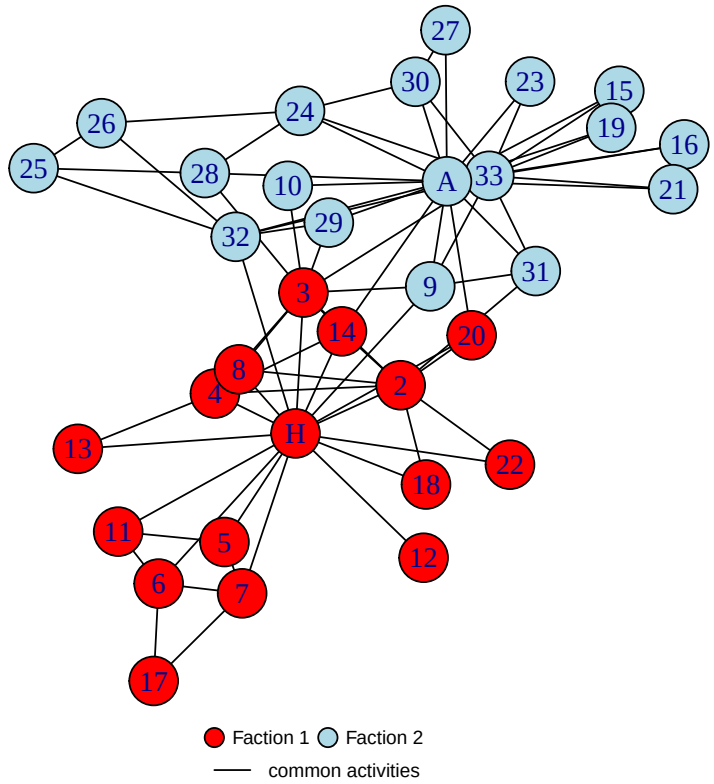



Source: Karate club network (igraphdata) [Zachary, 1977]

# Designing infographics
## Explain encoding

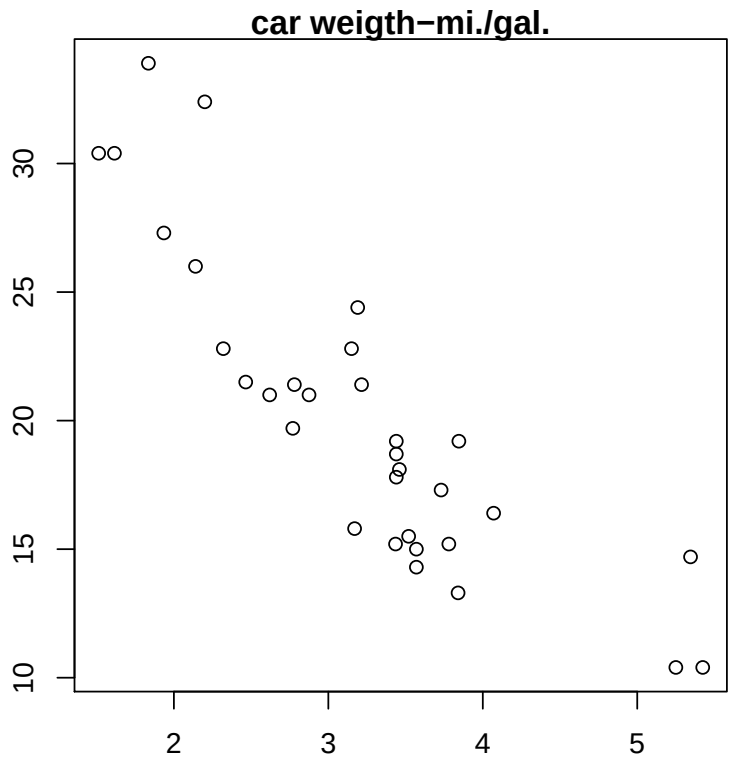



Source: Karate club network (igraphdata) [Zachary, 1977]



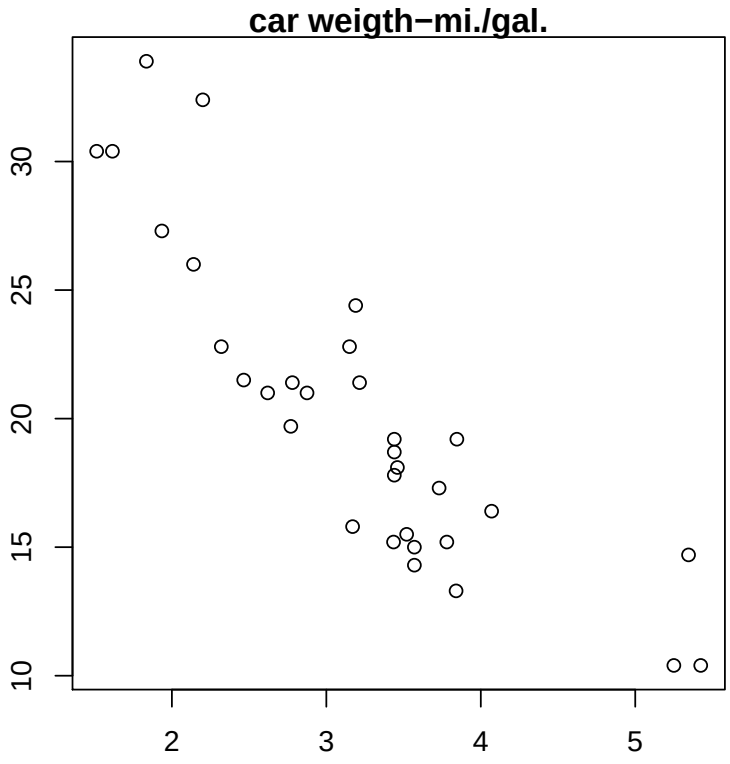

Source: Karate club network (igraphdata) [Zachary, 1977]

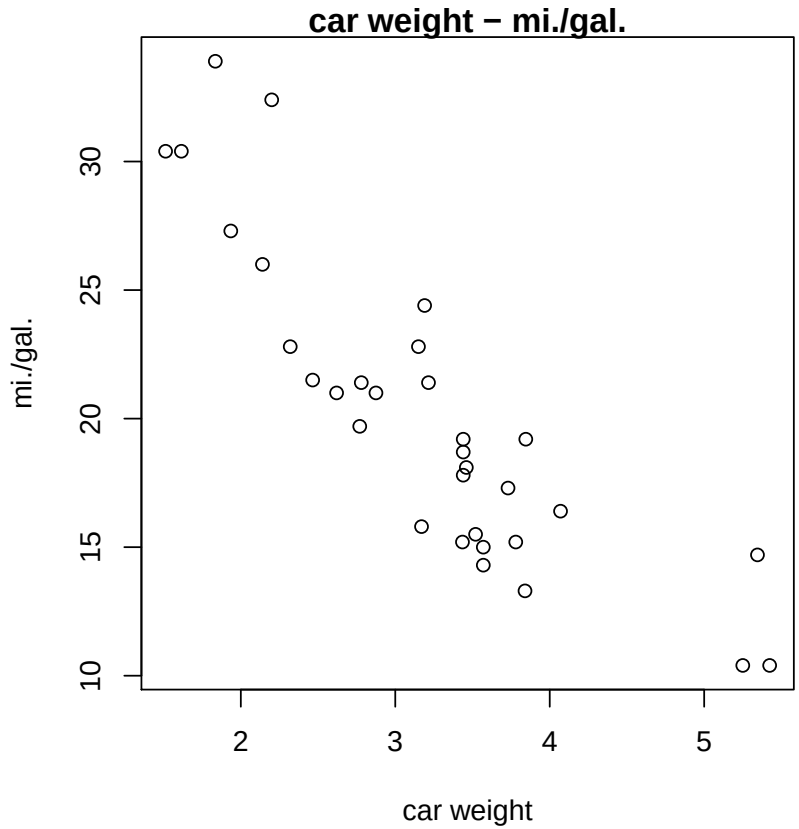
# Designing infographics
## Label axis



# Designing infographics

## Label axis



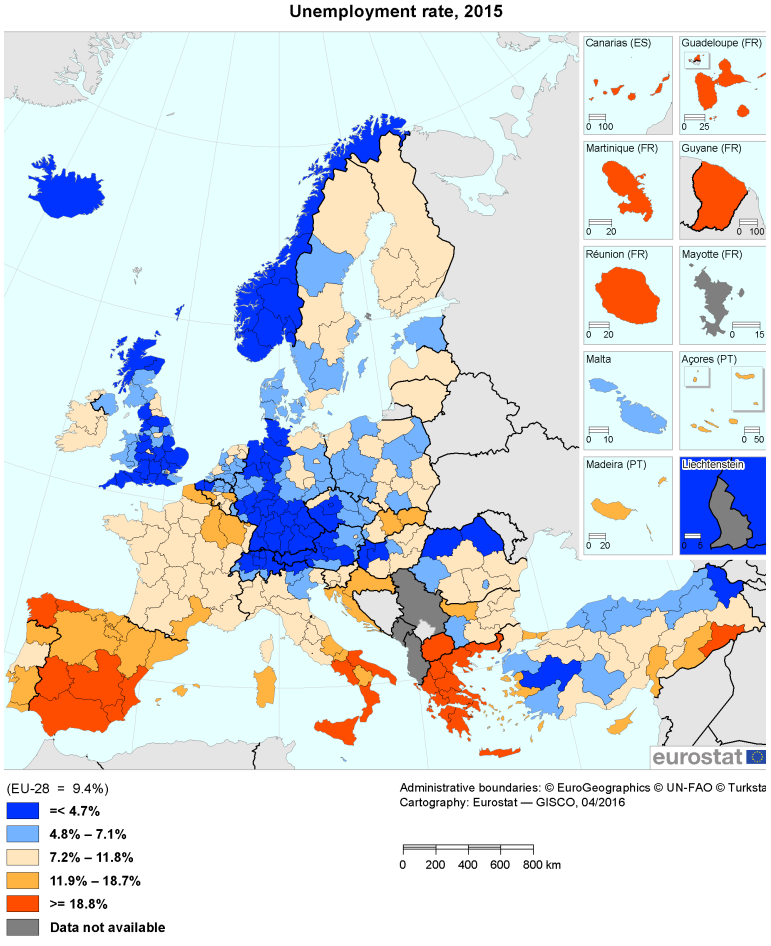



# Designing infographics

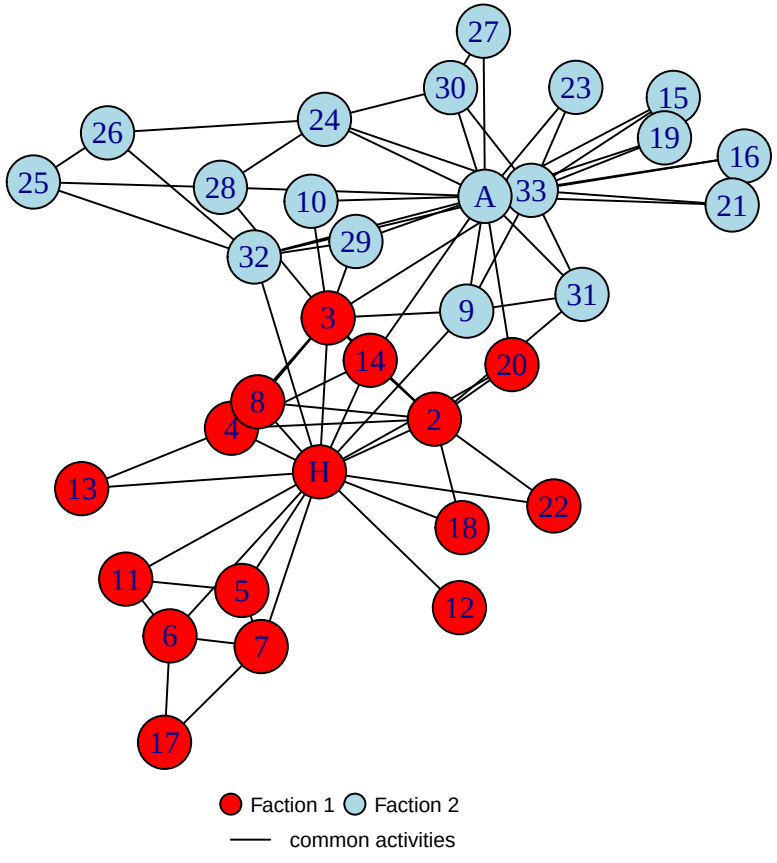
## Include sources

- Recognise others' work (IPR)
- Risk of plagiarism
- Text, images, and data

# Designing infographics
## Include sources



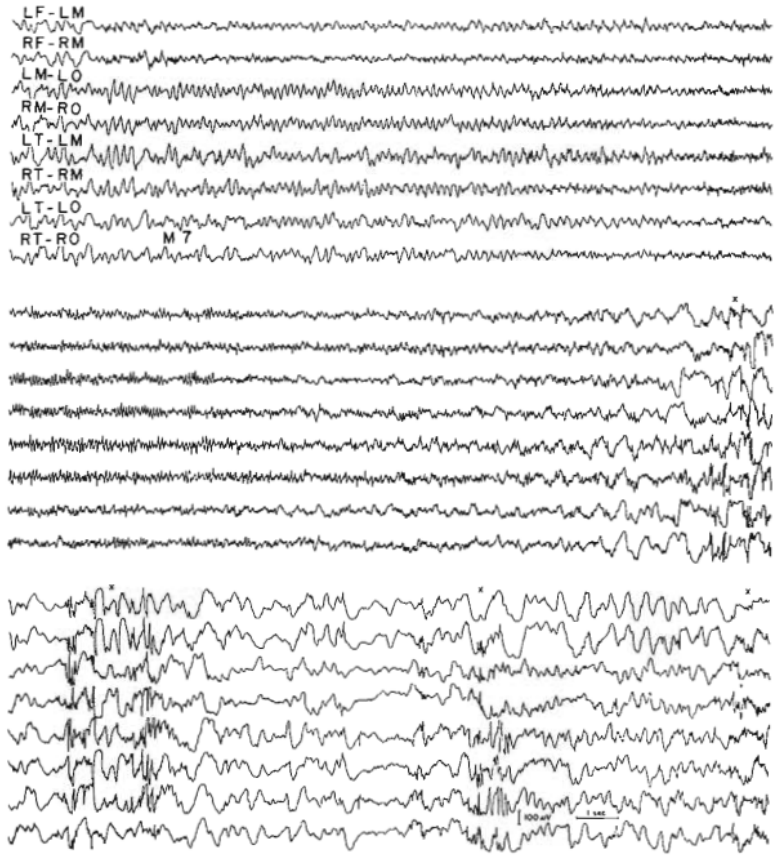

Source: Eurostat [http://ec.europa.eu]




Source: Karate club network (igraphdata) [Zachary, 1977]

# Designing infographics
## Maximise the data-ink ratio

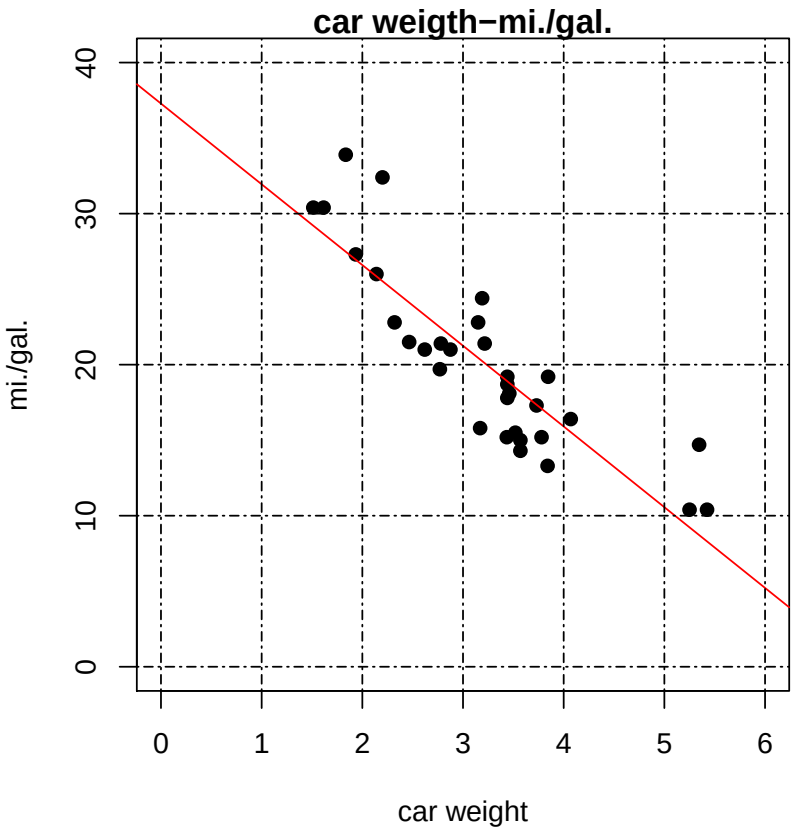
- Most of the 'ink' used to print a graphic should represent data-information
- Data-ink is defined as the part of the graphic that we cannot erase without losing information
- Data-ink ratio is defined as the proportion of a graphic's ink devoted to the display of non-redundant information (data-ink)




Source: Fundamentals of electroencephalography [Tufte, 2001]

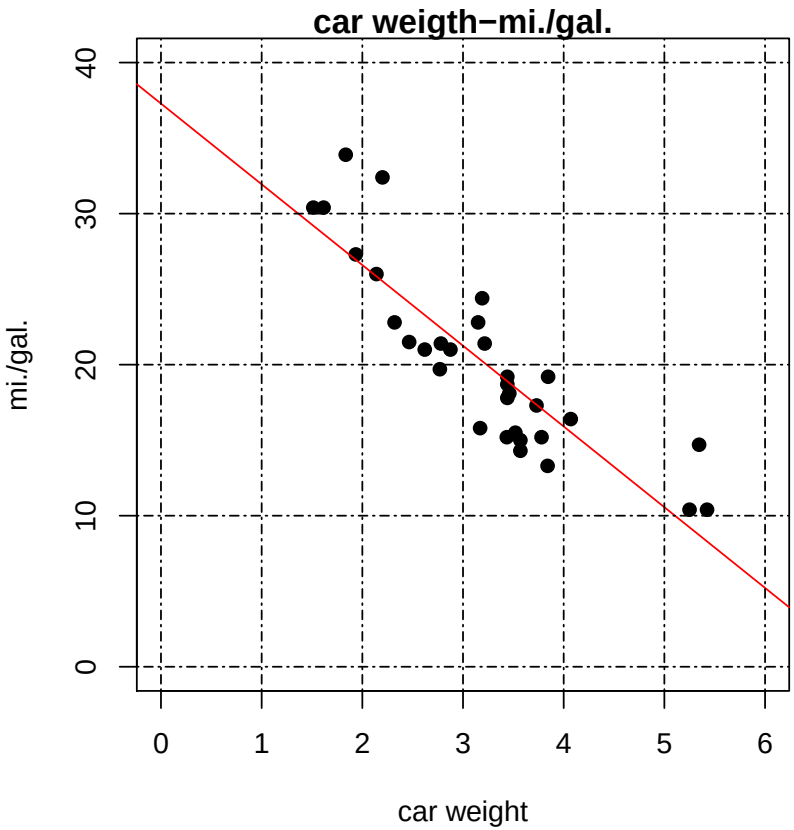
# Designing infographics

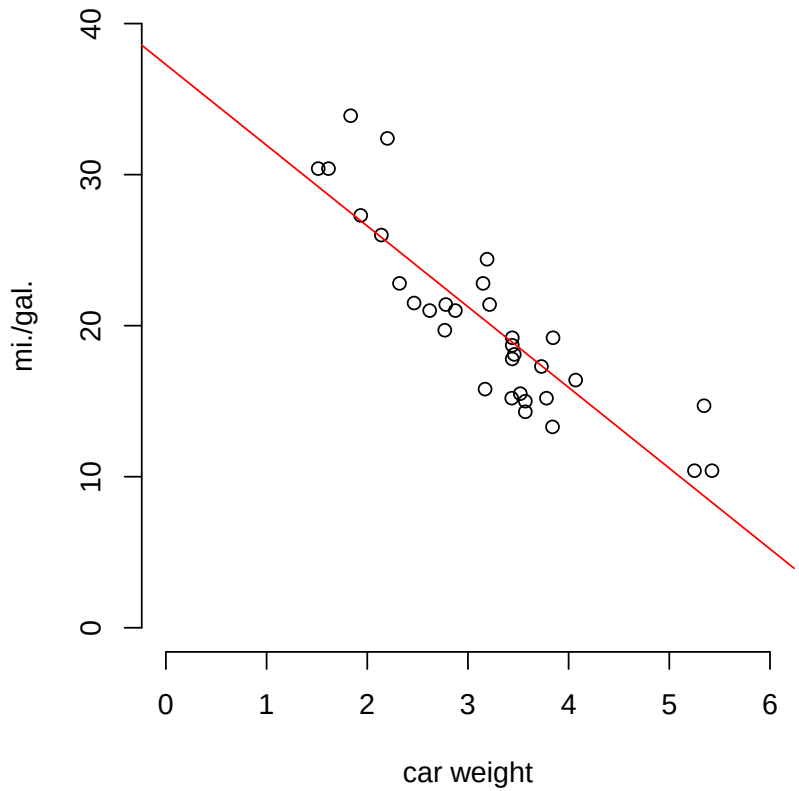
## Maximise the data-ink ratio



# Designing infographics

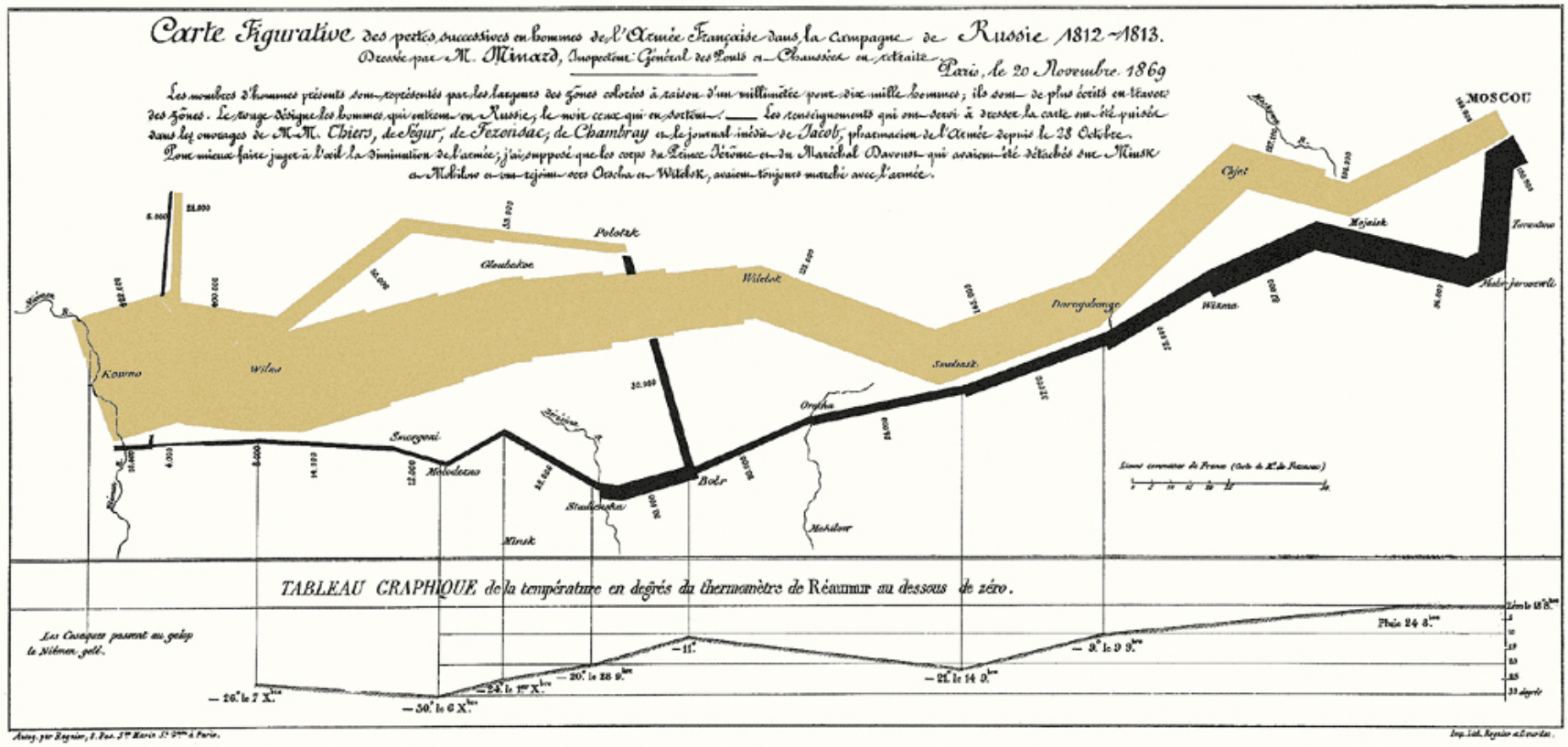
## Maximise the data-ink ratio



# Designing infographics

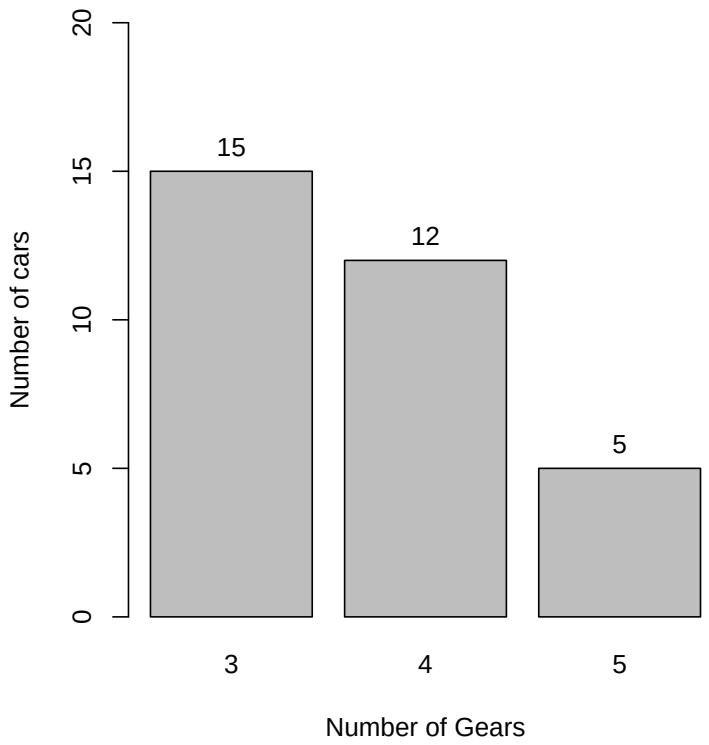
## Maximise the data-ink ratio




Source: Minard's visualisation of Napoleon's retreat from Moscow (`https://www.edwardtufte.com/tufte/minard`)

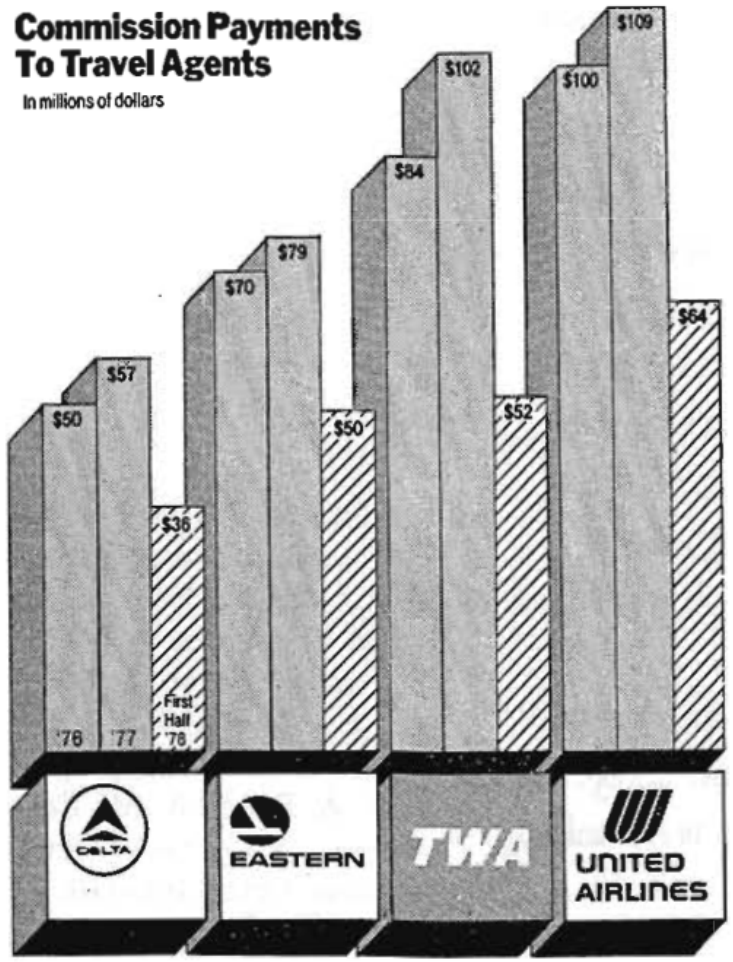
# Designing infographics
## Maximise the data-ink ratio

- Maximise data-ink ratio
- Erasing principles [Tufte, 2001]
  - Non-data-ink
  - Redundant data-information
  - Revise and edit



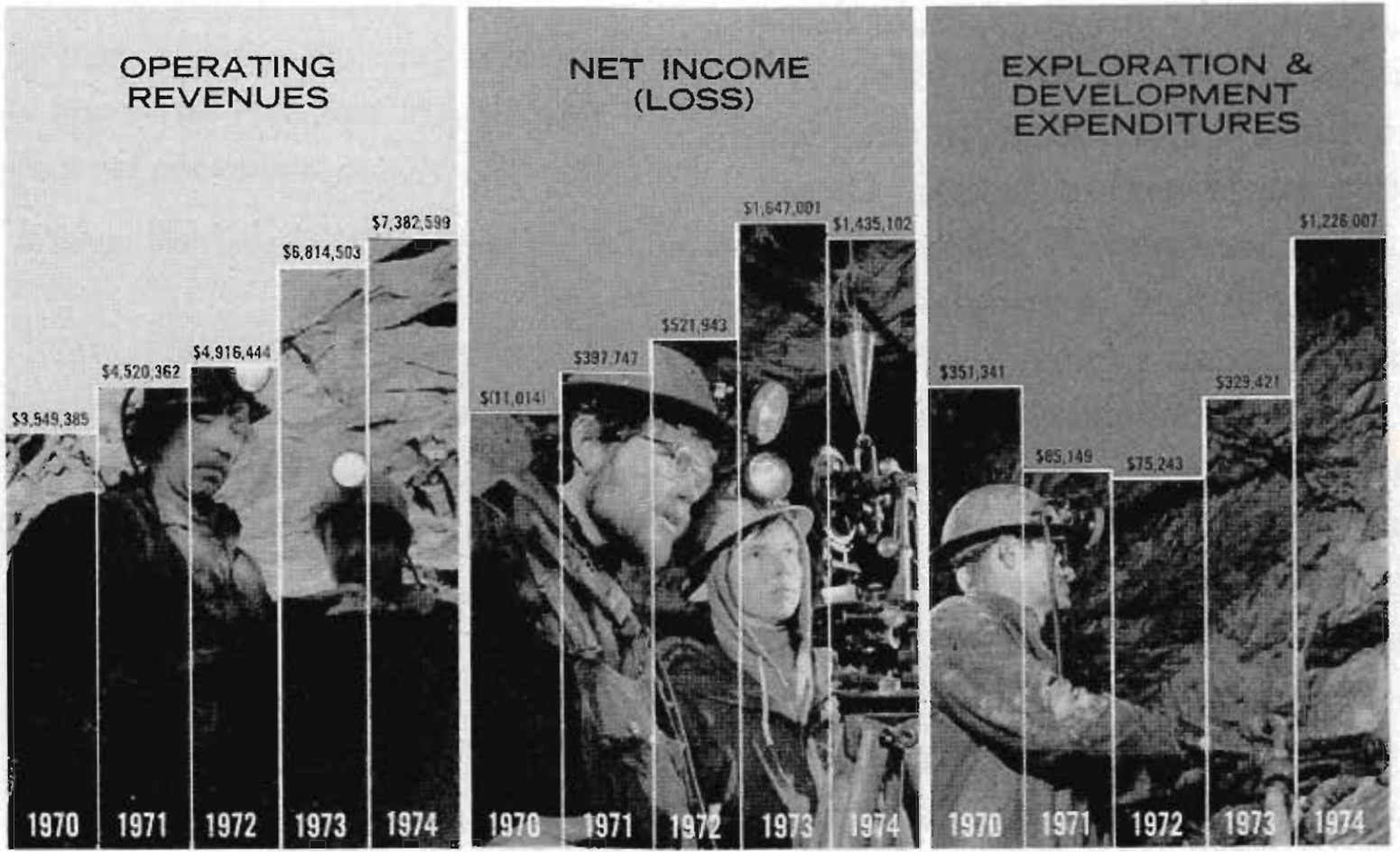
# Designing infographics
## Graphical integrity

- Graphics can distort underlying data
- Graphics can be used to deceive communication
- The visual representation may be consistent with data, while the perceived visual effect may deceive communication



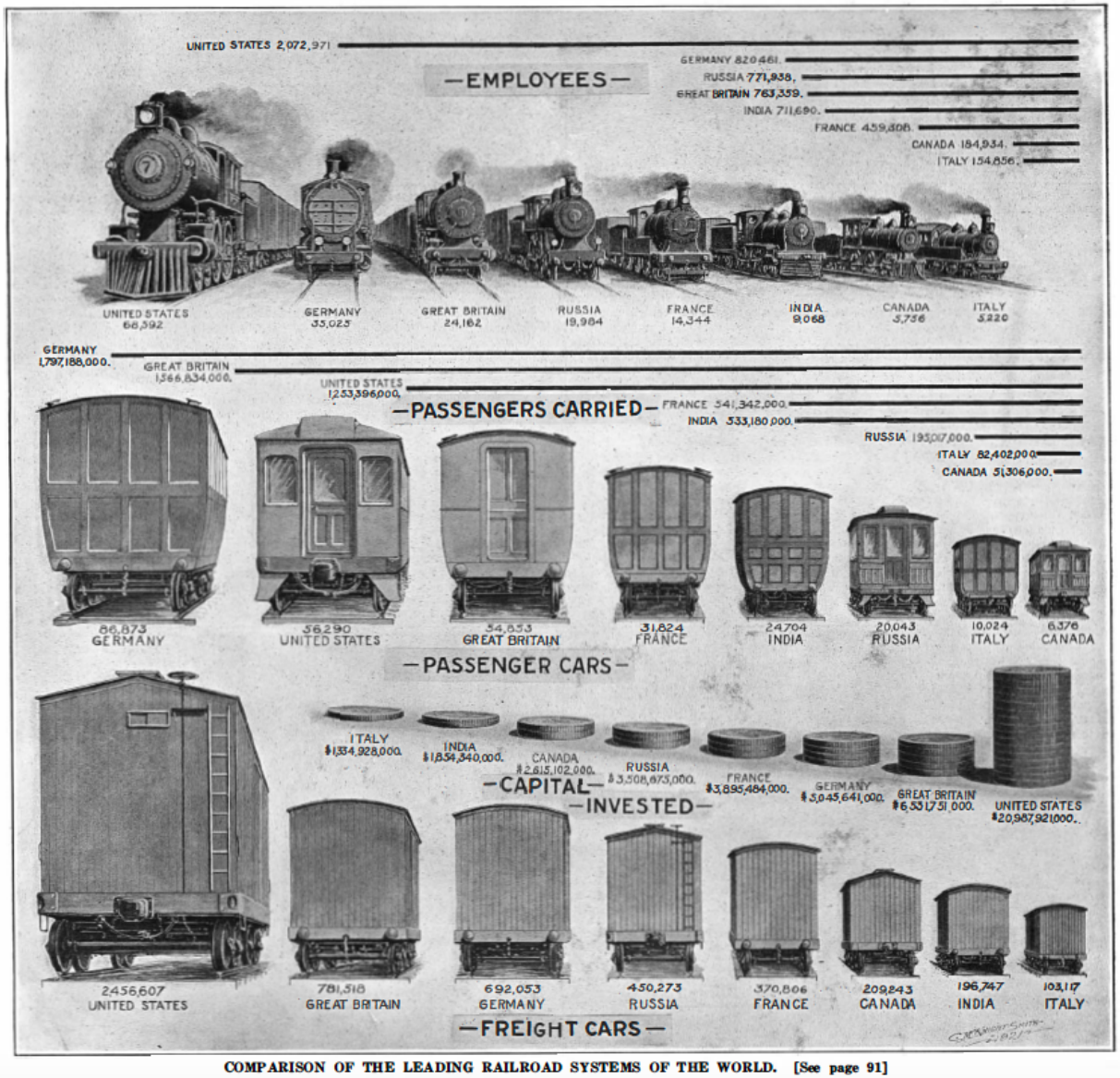

Source: Commission payments to travel agents [Tufte, 2001]

# Designing infographics
## Graphical integrity



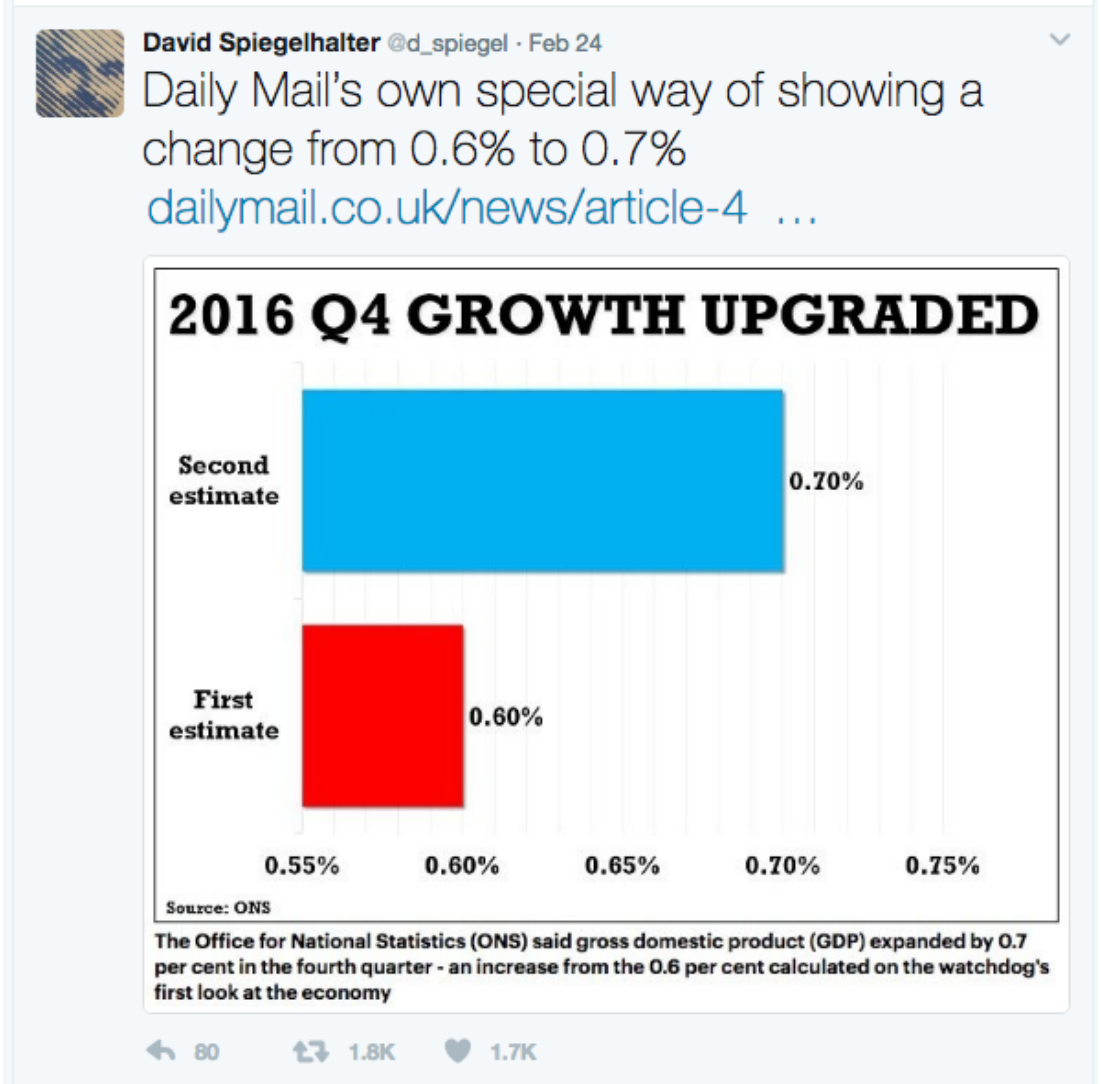

Source: Annual report of a company [Tufte, 2001]

# Designing infographics
## Graphical integrity




Source: Scientific American (1921)

# Designing infographics
## Graphical integrity




Source: Twitter (March, 2017)


Source: Campaign ad in Islington South (https://electionleaflets.org/brb.html

# Designing infographics
## Graphical integrity


Source: UK Lidl's leaflet (February 2022)

# Designing infographics
## Tools

Some of the tools that can support the creation of infographics

- Microsoft Office (e.g. PowerPoint)
- Google Documents (e.g. draw.io)
- Adobe Creative Cloud
- Tableau (`www.tableau.com`)
- R (ggplot2)
- Python, Java, D3js, etc.
- ArcGIS, GoogleMaps, etc.
- ...

# The good/bad infographic competition

# The good/bad infographic competition

1. Groups of 4 students (randomly selected)
2. Identify a good example and a bad example of infographics
3. Every group has a leader (name in bold) in charge of the submission
4. Please do send your examples on Canvas by 21 March 2022
5. Groups will showcase their examples during the lecture in Week
6. We will then vote for the best 'good infographic' and for the best 'bad infographic'

# The good/bad infographic competition

| |
|---|
| Group 1: **Ross**, Johanna, Samuel, Maria Belen |
| Group 2: **Maria Fernanda**, Poojani, Saradha, Charunan |
| Group 3: **Evi**, Ananya, Noemie, Harry |
| Group 4: **Perizat**, America, Ayesha, Oscar |
| Group 5: **Jongho**, Daniela, Tanya, Anas |

# Network visualisation

# Network visualisation

We can convey additional information through

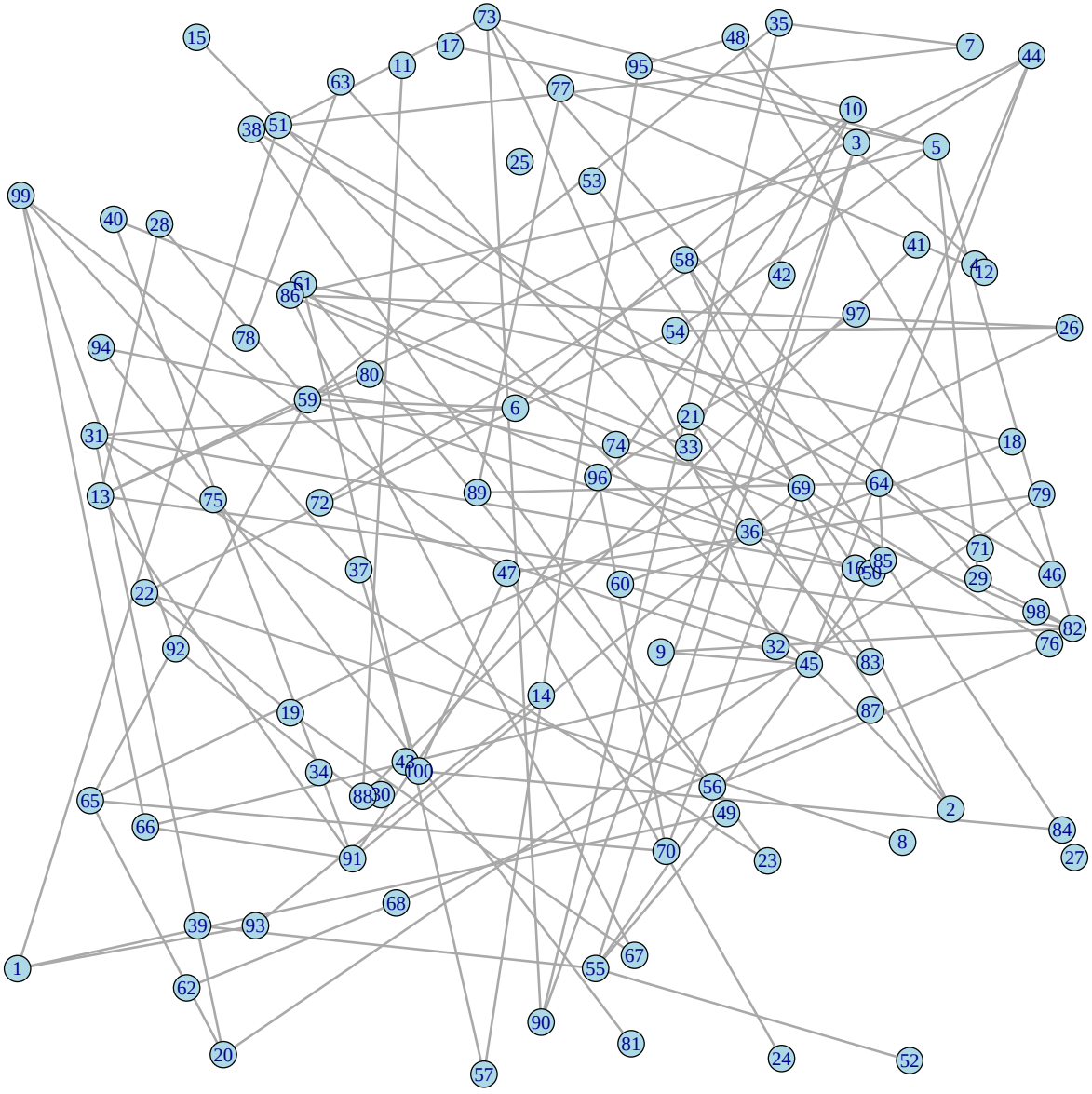
- Nodes: size, colour, and shape
- Edges: width, colour, and line type
- Layout: position of nodes

Let's generate a network $G(N, E)$ where:

- $N = 100$
- $x_{ij} = 1$ with $p = 0.02$

# Network visualisation

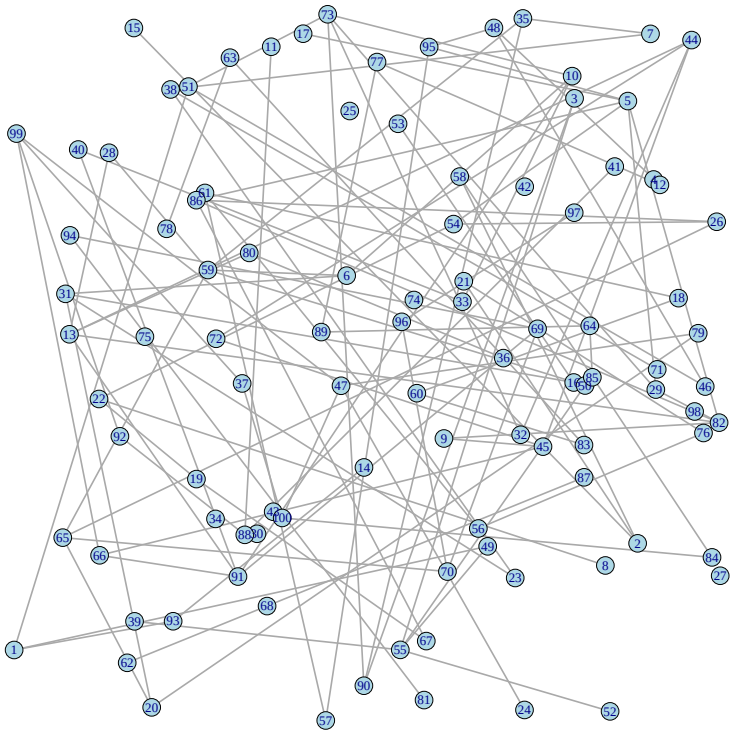
## Layout: Random



# Network visualisation

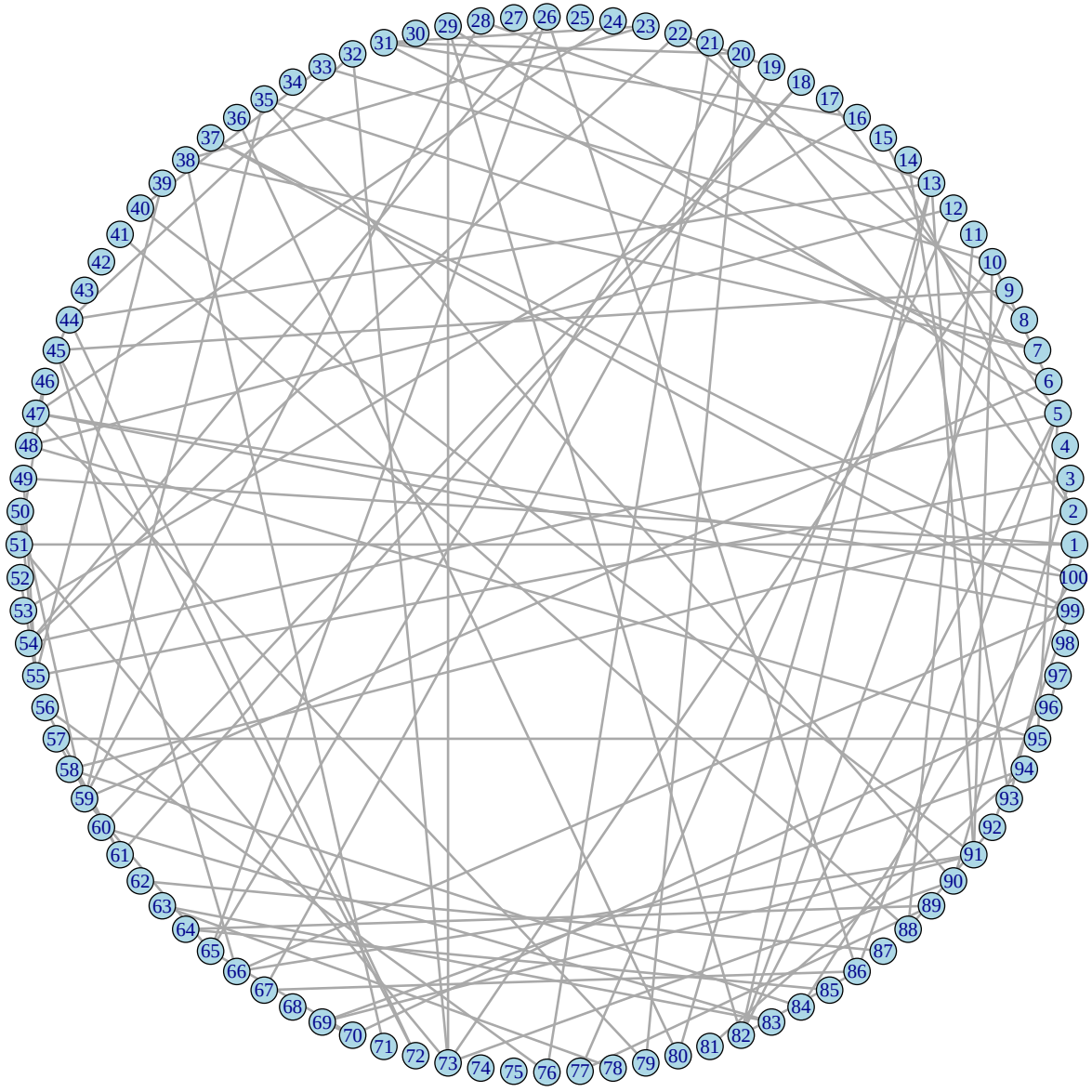
## Layout

- A network visualization should provide a relatively clear view of the structure of the network (e.g. central nodes)
- This requires positioning nodes in a 2D (or 3D) space according to certain criteria/rules



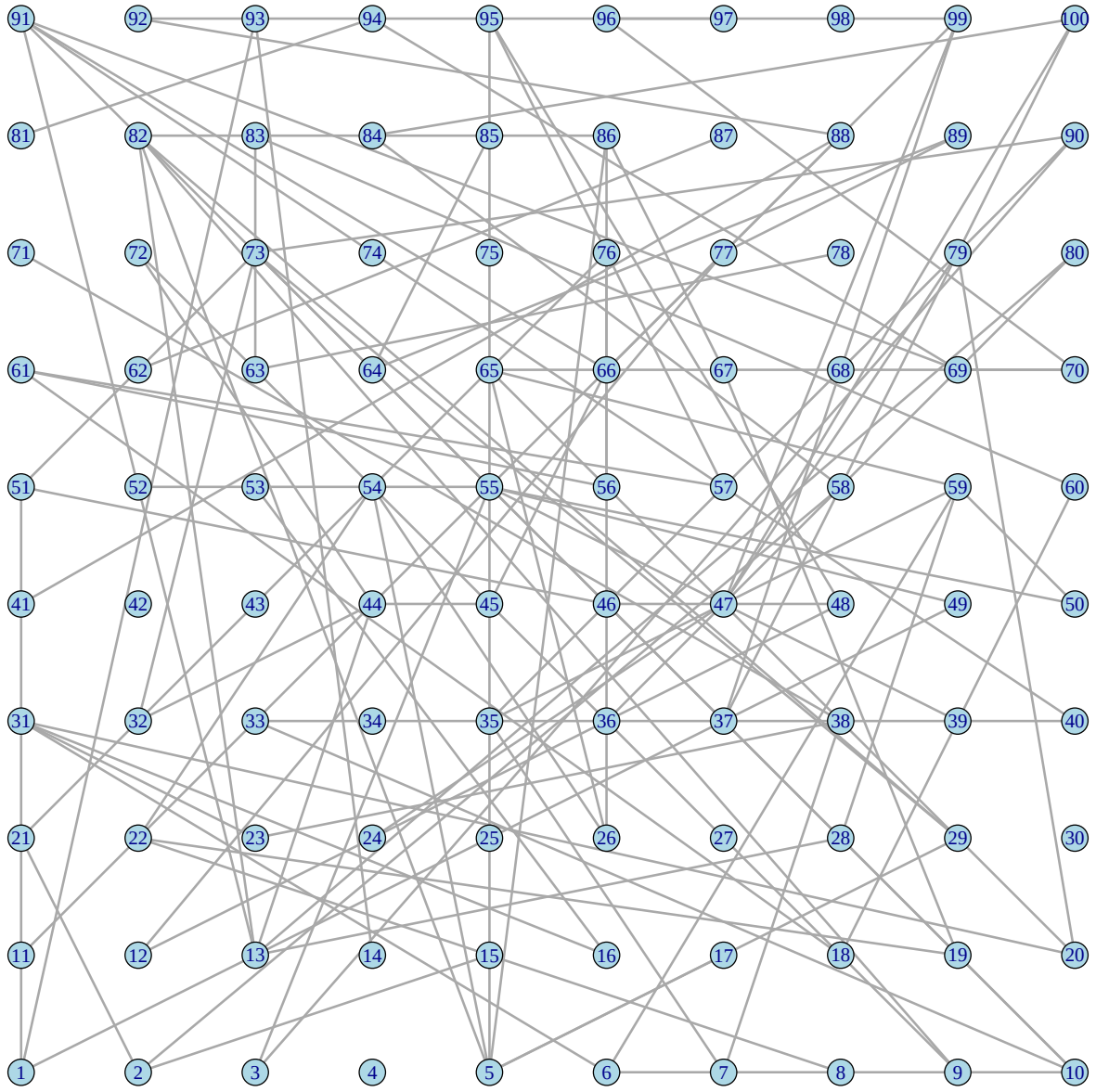
# Network visualisation

## Layout: Circle



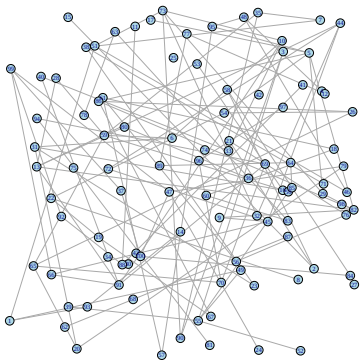
# Network visualisation

## Layout: Grid

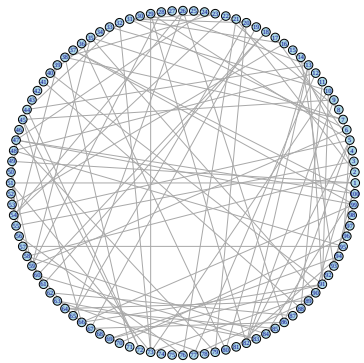


# Network visualisation

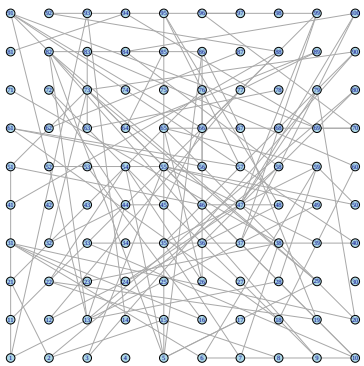
## Layout in igraph



```
grandom <- erdos.renyi.game(100, 0.02)
V(grandom)$color <- "lightblue"
V(grandom)$size  <- 5
E(grandom)$width <- 2
l_random <- layout.random(grandom)
pdf('random.pdf', width = 10, height = 10)
par(mfrow=c(1, 1), mar=c(0,0,0,0))
plot(grandom, layout = l_random)
dev.off()
```



```
l_circle <- layout.circle(grandom)
pdf('circle.pdf', width = 10, height = 10)
par(mfrow=c(1, 1), mar=c(0,0,0,0))
plot(grandom, layout = l_circle)
dev.off()
```



```
l_gr <- layout_on_grid(grandom, width=10)
pdf('grid.pdf', width = 10, height = 10)
par(mfrow=c(1, 1), mar=c(0,0,0,0))
plot(grandom, layout = l_gr)
dev.off()
```

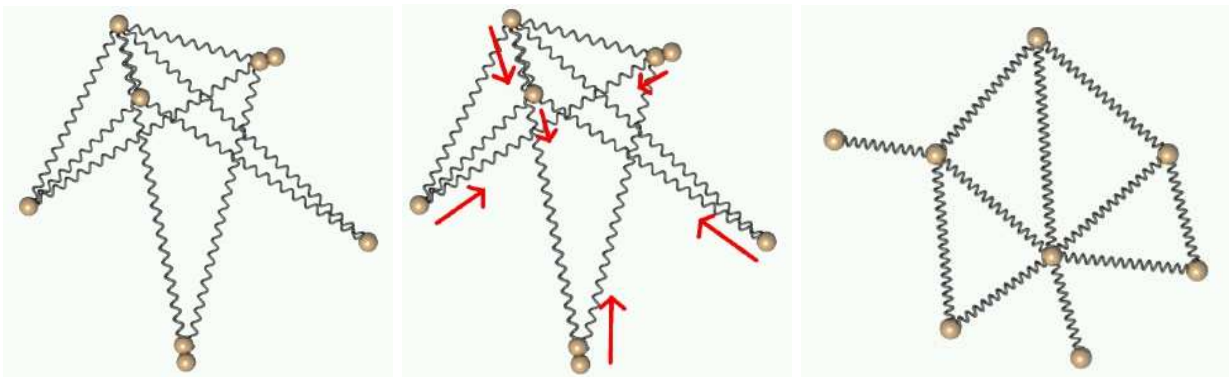
- A variety of layout algorithms have been developed to improve the layout of nodes and therefore network visualisations
- We focus on force-directed algorithms: The layout of a network is defined using the information contained within the structure of the network itself
  - Kamada-Kawai
  - Fruchterman and Reingold

# Network visualisation

## Layout: Kamada-Kawai

[Kamada and Kawai, 1989]

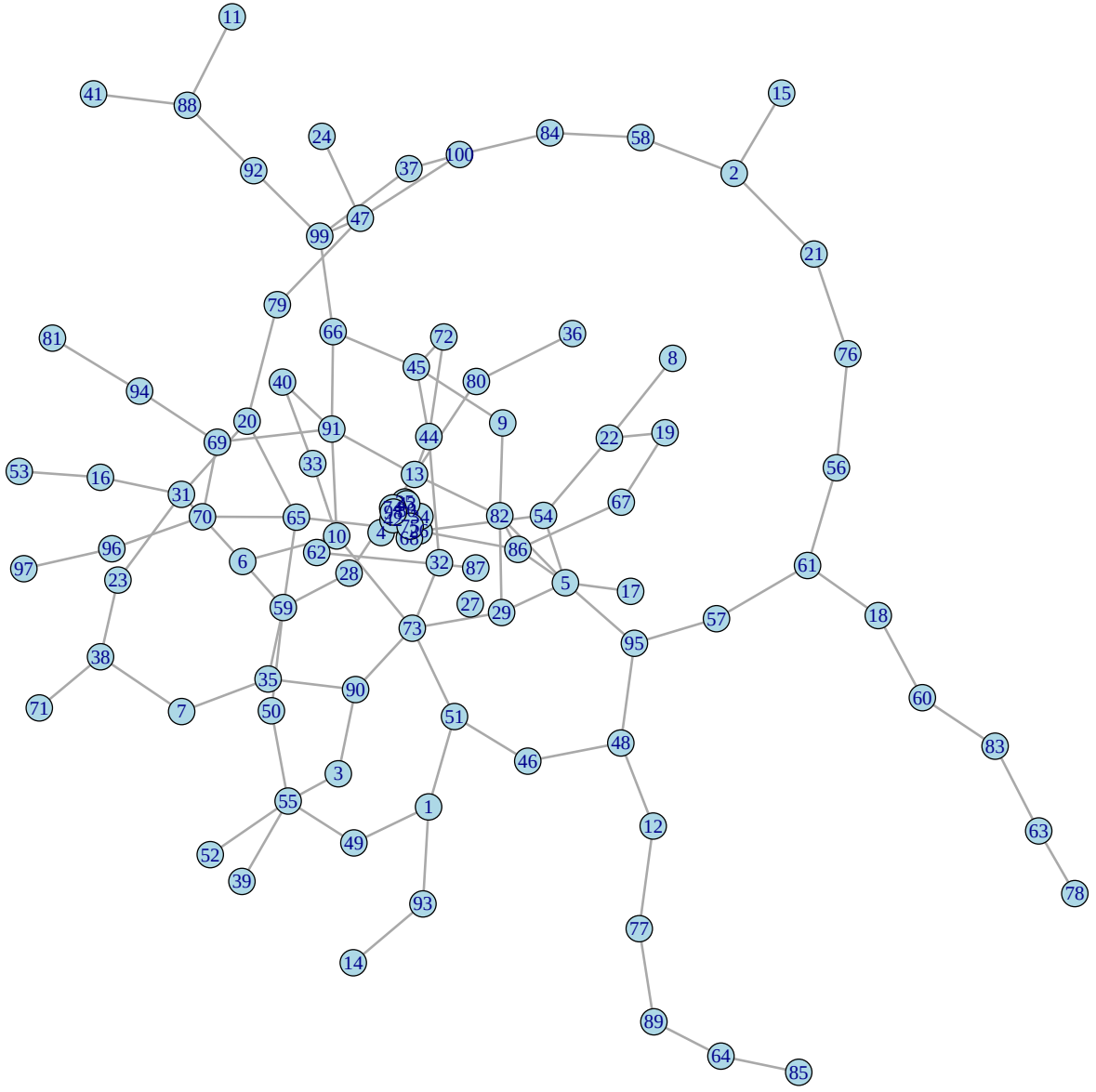
- Ties as springs exerting attraction and repulsion forces between nodes
- Length of spring: proportional to the shortest distance between two nodes
- Strength of the spring (i.e. repulsion and attraction forces): inversely proportional to the square of the shortest distance between two nodes



Source: [Kobourov, 2012]

# Network visualisation

## Layout: Kamada-Kawai

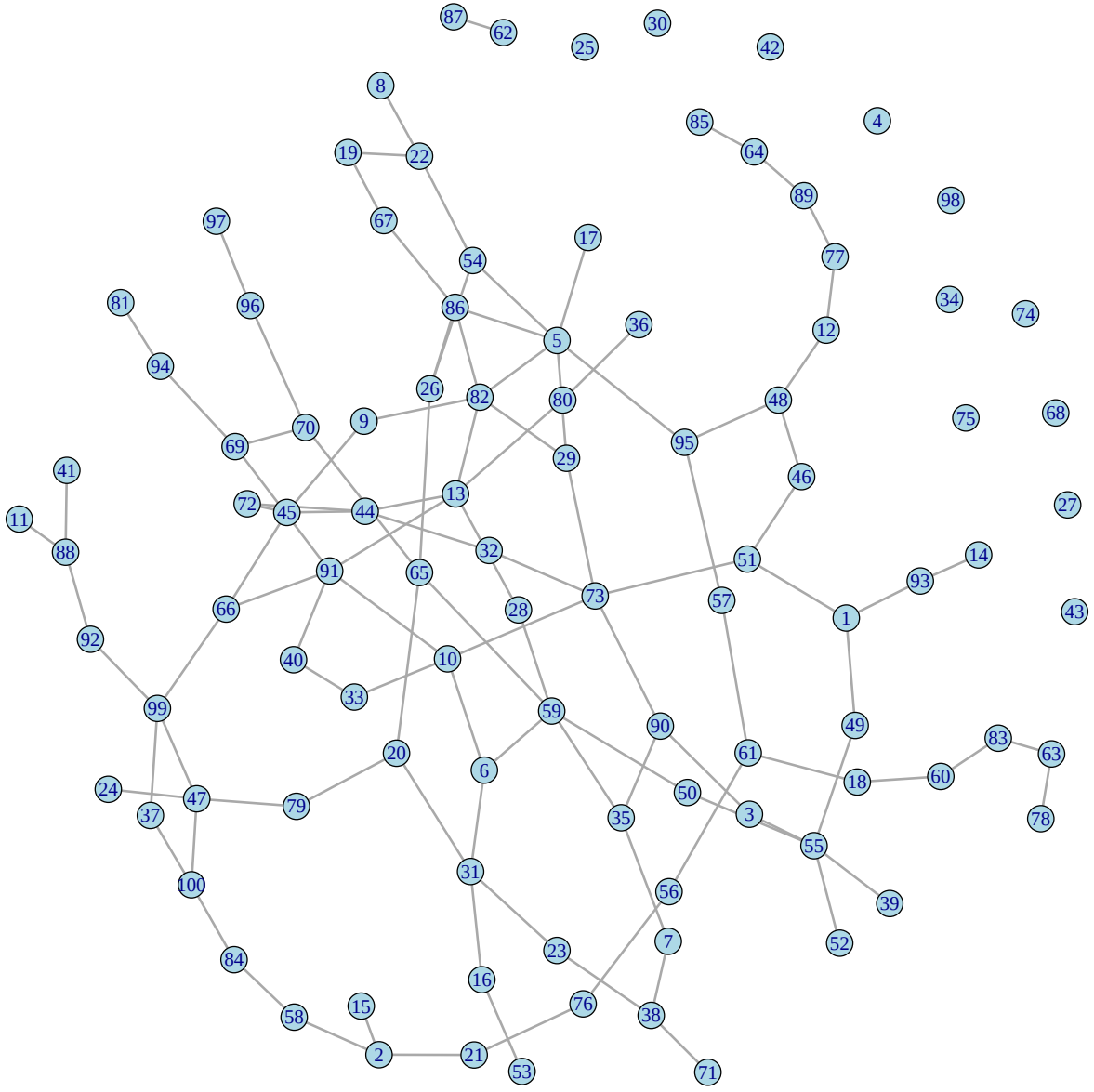


[Fruchterman and Reingold, 1991]

- Nodes as "atomic particles or celestial bodies exerting attractive and repulsive forces from one another"
- Attractive forces: proportional to the square of the shortest distance between two nodes
- Repulsive forces: inversely proportional to the shortest distance between two nodes

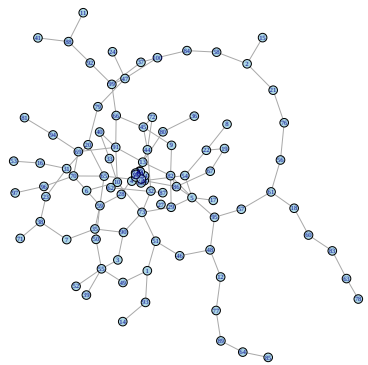
# Network visualisation

## Layout: Fruchterman and Reingold

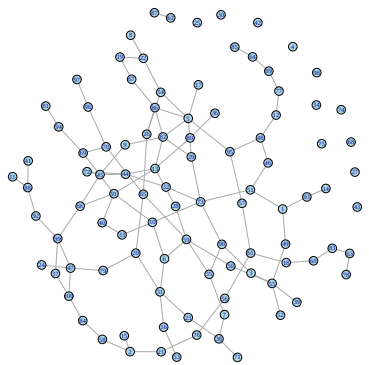


# Network visualisation

## Layout: Kamada-Kawau and Fruchterman and Reingold in igraph



```
l_kk <- layout_with_kk(grandom)
pdf('kamada.pdf', width = 10, height = 10)
par(mfrow=c(1, 1), mar=c(0,0,0,0))
plot(grandom, layout = l_kk)
dev.off()
```



```
l_fr <- layout_with_fr(grandom)
pdf('fr.pdf', width = 10, height = 10)
par(mfrow=c(1, 1), mar=c(0,0,0,0))
plot(grandom, layout = l_fr)
dev.off()
```

# Questions

# Next time ...

# Next time ...

- **Seminar: Principles of infographics**
  - Network layout algorithms in igraph
  - Import data in Gephi
  - Visualise and analyse networks in Gephi

- **Lecture: Network models**
  - Mathematical models of network analysis
  - Overview of statistical models of network analysis

# References I


Fruchterman, T. M. J. and Reingold, E. M. (1991).

Graph drawing by force-directed placement.
*Software: Practice and Experience*, 21(11):1129–1164.

Kamada, T. and Kawai, S. (1989).

An algorithm for drawing general undirected graphs.
*Information Processing Letters*, 31(1):7–15.

Kobourov, S. (2012).

Spring Embedders and Force Directed Graph Drawing Algorithms.
*arXiv preprint arXiv:1201.3011*, pages 1–23.

Tufte, E. R. (2001).

*The Visual Display of Quantitative Information*, volume 4.
Graphic Press, Cheshire, Connecticut.

Zachary, W. W. (1977).

An Information Flow Model for Conflict and Fission in Small Groups.
*Journal of Anthropological Research*, 33(4):452–473.