

# Display numeric data

- ⌚ is mandatory in science
- ⌚ as data is the root of meaningful critical thinking
- ⌚ can be difficult and time consuming
- ⌚ need to have an clearly stated objective

# Display numeric data

- ⌚ is mandatory in science
- ⌚ as data is the root of meaningful critical thinking
- ⌚ can be difficult and time consuming
- ⌚ need to have an clearly stated objective

## Display numeric data

- ⌚ is mandatory in science
- ⌚ as data is the root of meaningful critical thinking
- ⌚ can be difficult and time consuming
- ⌚ need to have an clearly stated objective

## Display numeric data

- ⌚ is mandatory in science
- ⌚ as data is the root of meaningful critical thinking
- ⌚ can be difficult and time consuming
- ⌚ need to have an clearly stated objective

## Display numeric data

- ⌚ is mandatory in science
- ⌚ as data is the root of meaningful critical thinking
- ⌚ can be difficult and time consuming
- ⌚ need to have an clearly stated objective

# The Art of Graphical Display of Quantitative Data

MEREC (Doctoral Studies)

Mathieu Lagrange 



# Credits

## Edward R. Tufte

- ✉ statistician and artist
- ✉ Professor Emeritus of Political Science, Statistics, and Computer Science at Yale University
- ✉ "Leonardo da Vinci of data" (New York Times)
- ✉ "Galileo of graphics" (Business Week)
- ✉ <http://www.edwardtufte.com>



# Credits

## Edward R. Tufte

- ε statistician and artist
- ε Professor Emeritus of Political Science, Statistics, and Computer Science at Yale University
- ε "Leonardo da Vinci of data" (New York Times)
- ε "Galileo of graphics" (Business Week)
- ε <http://www.edwardtufte.com>



# Credits

## Edward R. Tufte

- ε statistician and artist
- ε Professor Emeritus of Political Science, Statistics, and Computer Science at Yale University
- ε "Leonardo da Vinci of data" (New York Times)
- ε "Galileo of graphics" (Business Week)
- ε <http://www.edwardtufte.com>



# Credits

## Edward R. Tufte

- ε statistician and artist
- ε Professor Emeritus of Political Science, Statistics, and Computer Science at Yale University
- ε "Leonardo da Vinci of data" (New York Times)
- ε "Galileo of graphics" (Business Week)
- ε <http://www.edwardtufte.com>

# Credits

## Edward R. Tufte

- ε statistician and artist
- ε Professor Emeritus of Political Science, Statistics, and Computer Science at Yale University
- ε "Leonardo da Vinci of data" (New York Times)
- ε "Galileo of graphics" (Business Week)
- ε <http://www.edwardtufte.com>

# Credits

## Edward R. Tufte

- ε statistician and artist
- ε Professor Emeritus of Political Science, Statistics, and Computer Science at Yale University
- ε "Leonardo da Vinci of data" (New York Times)
- ε "Galileo of graphics" (Business Week)
- ε <http://www.edwardtufte.com>



# Outline

① Graphical excellence

② Space

③ Time

④ Space and Time

⑤ Relation

⑥ Principles



# Outline

① Graphical excellence

② Space

③ Time

④ Space and Time

⑤ Relation

⑥ Principles



# Outline

① Graphical excellence

② Space

③ Time

④ Space and Time

⑤ Relation

⑥ Principles



# Outline

① Graphical excellence

② Space

③ Time

④ Space and Time

⑤ Relation

⑥ Principles



# Outline

① Graphical excellence

② Space

③ Time

④ Space and Time

⑤ Relation

⑥ Principles



# Outline

① Graphical excellence

② Space

③ Time

④ Space and Time

⑤ Relation

⑥ Principles



# Graphical excellence

# Graphical Excellence

## Graphical displays should

- ↪ show the data
- ↪ present many numbers in a small space
- ↪ avoid distorting what the data have to say
- ↪ serve a reasonably clear purpose: description, exploration, tabulation, or decoration
- ↪ induce the viewer to think about the substance rather than about methodology, graphic design, the technology of graphic production, or something else

# Graphical Excellence

## Graphical displays should

- ⌚ show the data
- ⌚ present many numbers in a small space
- ⌚ avoid distorting what the data have to say
- ⌚ serve a reasonably clear purpose: description, exploration, tabulation, or decoration
- ⌚ induce the viewer to think about the substance rather than about methodology, graphic design, the technology of graphic production, or something else

# Graphical Excellence

## Graphical displays should

- ⌚ show the data
- ⌚ present many numbers in a small space
- ⌚ avoid distorting what the data have to say
- ⌚ serve a reasonably clear purpose: description, exploration, tabulation, or decoration
- ⌚ induce the viewer to think about the substance rather than about methodology, graphic design, the technology of graphic production, or something else

# Graphical Excellence

## Graphical displays should

- ⌚ show the data
- ⌚ present many numbers in a small space
- ⌚ avoid distorting what the data have to say
- ⌚ serve a reasonably clear purpose: description, exploration, tabulation, or decoration
- ⌚ induce the viewer to think about the substance rather than about methodology, graphic design, the technology of graphic production, or something else

# Graphical Excellence

## Graphical displays should

- ε show the data
- ε present many numbers in a small space
- ε avoid distorting what the data have to say
- ε serve a reasonably clear purpose: description, exploration, tabulation, or decoration
- ε induce the viewer to think about the substance rather than about methodology, graphic design, the technology of graphic production, or something else

# Graphical Excellence

## Graphical displays should

- ε show the data
- ε present many numbers in a small space
- ε avoid distorting what the data have to say
- ε serve a reasonably clear purpose: description, exploration, tabulation, or decoration
- ε induce the viewer to think about the substance rather than about methodology, graphic design, the technology of graphic production, or something else

# Graphics shall reveal Data

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

A table



# Graphics shall reveal Data

$N = 11$

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

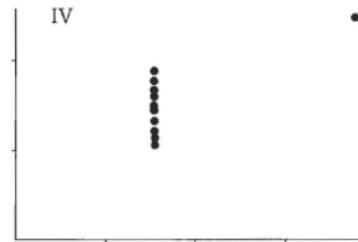
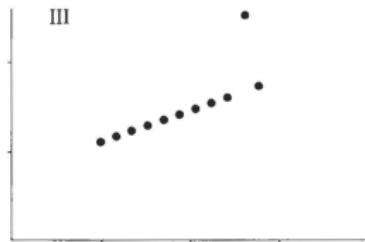
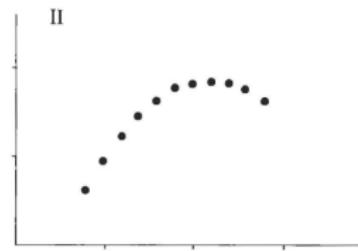
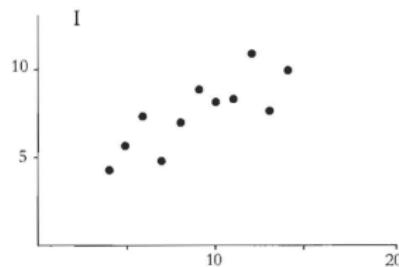
residual sum of squares of Y = 13.75

correlation coefficient = .82

$r^2 = .67$

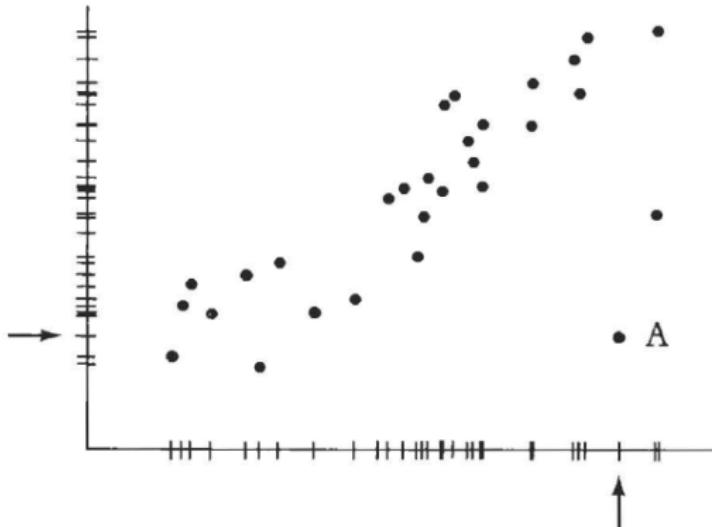
## Linear regression

# Graphics shall reveal Data

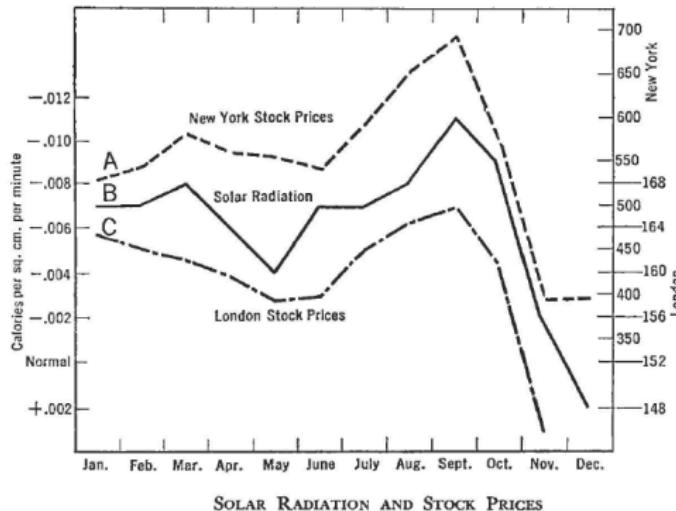


Plots

# Graphics shall reveal Data



## Graphics can be badly used



A. New York stock prices (Barron's average). B. Solar Radiation, inverted, and C. London stock prices, all by months, 1929 (after Garcia-Mata and Shaffner).

"Correlation is not causation"

<http://tylervigen.com/spurious-correlations>

# Space

# Maps

The notion of space is the thing to be visualized

- ⌚ Clay tablets : -3000
- ⌚ China : 1000
- ⌚ West : 1500
- ⌚ Data Maps : 1700

# Maps

The notion of space is the thing to be visualized

- ⌚ Clay tablets : -3000
- ⌚ China : 1000
- ⌚ West : 1500
- ⌚ Data Maps : 1700

# Maps

The notion of space is the thing to be visualized

- ⌚ Clay tablets : -3000
- ⌚ China : 1000
- ⌚ West : 1500
- ⌚ Data Maps : 1700

# Maps

The notion of space is the thing to be visualized

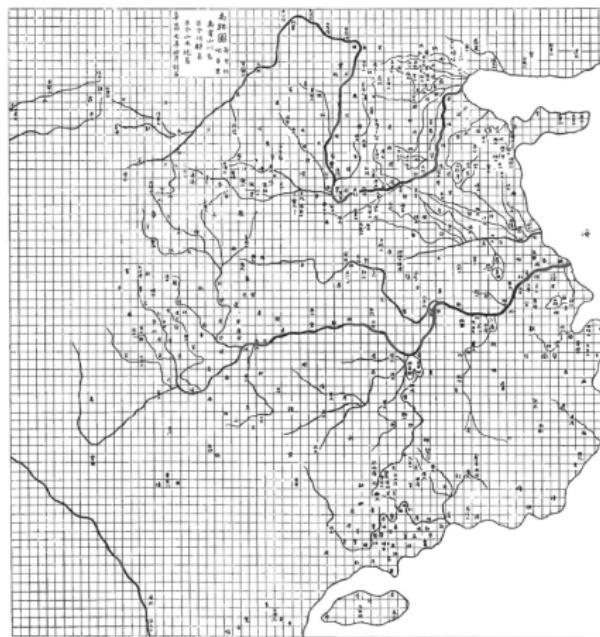
- ⌚ Clay tablets : -3000
- ⌚ China : 1000
- ⌚ West : 1500
- ⌚ Data Maps : 1700

# Maps

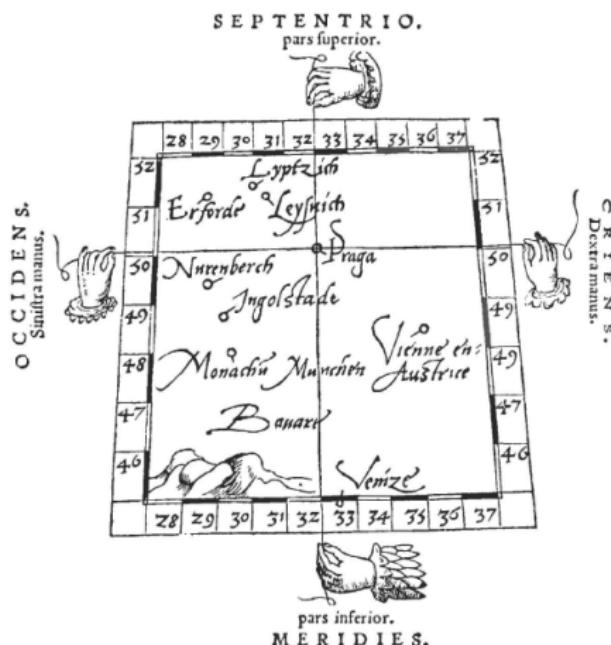
The notion of space is the thing to be visualized

- ε Clay tablets : -3000
- ε China : 1000
- ε West : 1500
- ε Data Maps : 1700

# China : 1000



# Europe : 1546



# Data Map: Trade Winds



# Data Map: Cholera



# Time

# Time Series

Time series are the second thing to be discovered

- ⌚ Much later
- ⌚ Improve the ability to predict things
- ⌚ First record: 1000
- ⌚ Scientific usage: 1800

# Time Series

Time series are the second thing to be discovered

- ⌚ Much later
- ⌚ Improve the ability to predict things
- ⌚ First record: 1000
- ⌚ Scientific usage: 1800

# Time Series

Time series are the second thing to be discovered

- ⌚ Much later
- ⌚ Improve the ability to predict things
- ⌚ First record: 1000
- ⌚ Scientific usage: 1800

# Time Series

Time series are the second thing to be discovered

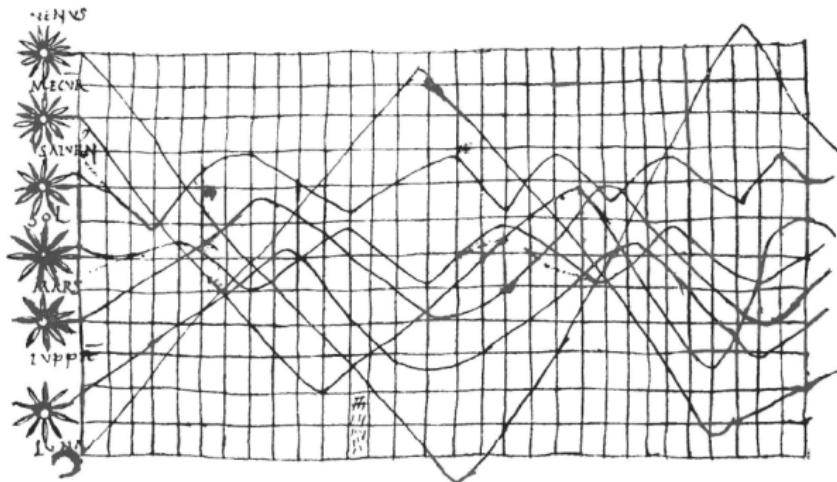
- ⌚ Much later
- ⌚ Improve the ability to predict things
- ⌚ First record: 1000
- ⌚ Scientific usage: 1800

# Time Series

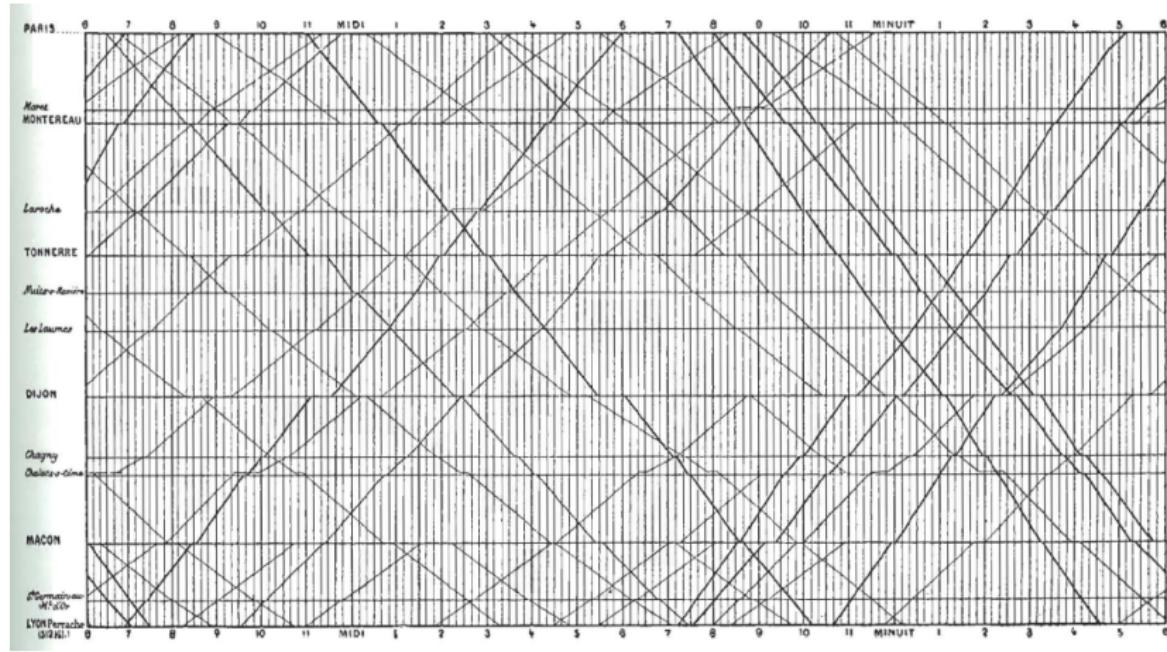
Time series are the second thing to be discovered

- ε Much later
- ε Improve the ability to predict things
- ε First record: 1000
- ε Scientific usage: 1800

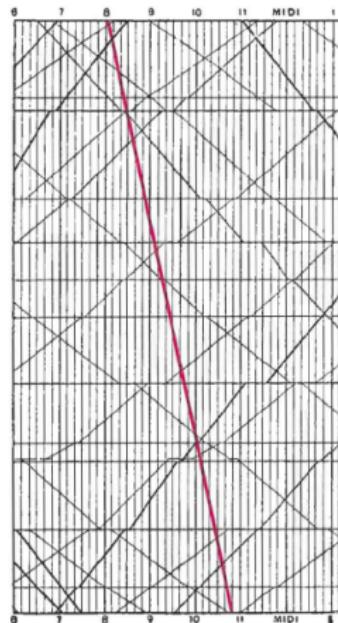
# First attempt



# Railway Scheduling : Ibry

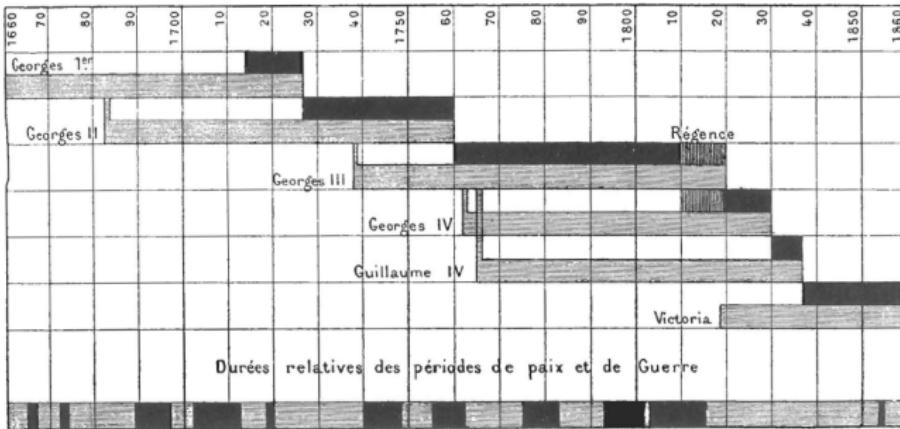


# Railway Scheduling

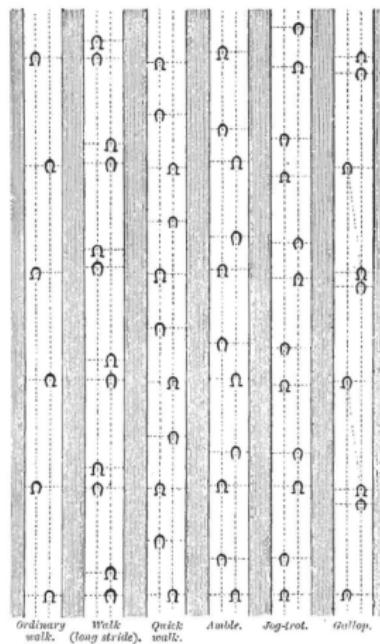


An update

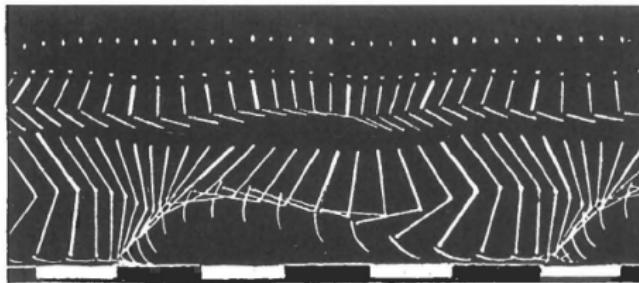
# Implicit Advanced Design : Marey (1900)



# Movement : Marey



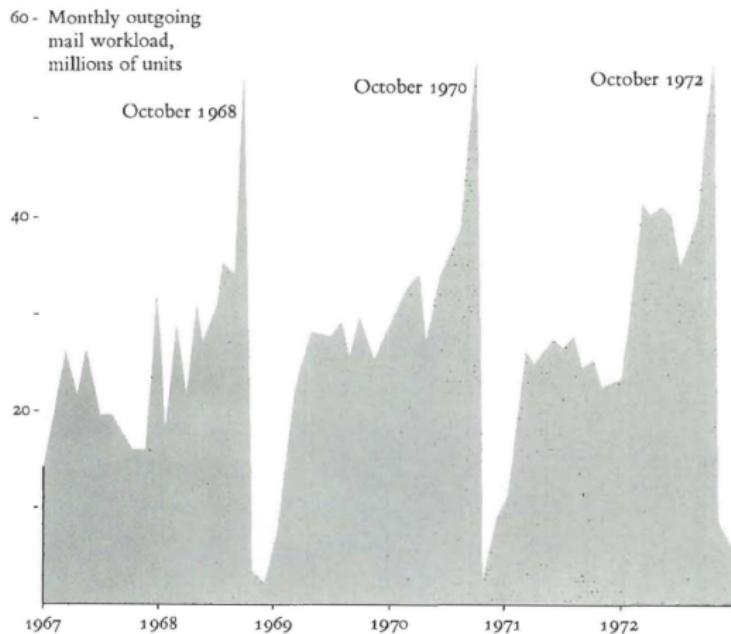
## Movement : Marey



# Movement : Marey

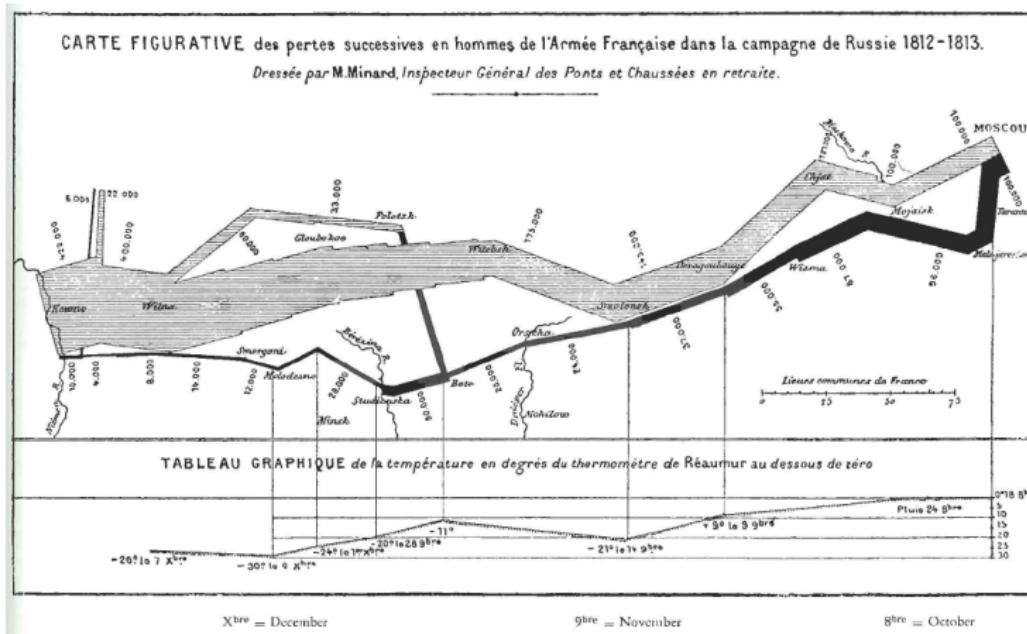


# Let the data speak



# Space and Time

# Minard : a masterpiece



# Relation

# Relation between quantities

## Relation is the most important thing

- Space and time are the more obvious quantities
- Moving away from the started in 1800 with Lambert and Playfair
- It took 15 years to Playfair to move way from space and time analogy to abstract design rules

# Relation between quantities

## Relation is the most important thing

- ⌚ Space and time are the more obvious quantities
- ⌚ Moving away from the started in 1800 with Lambert and Playfair
- ⌚ It took 15 years to Playfair to move way from space and time analogy to abstract design rules

## Relation between quantities

### Relation is the most important thing

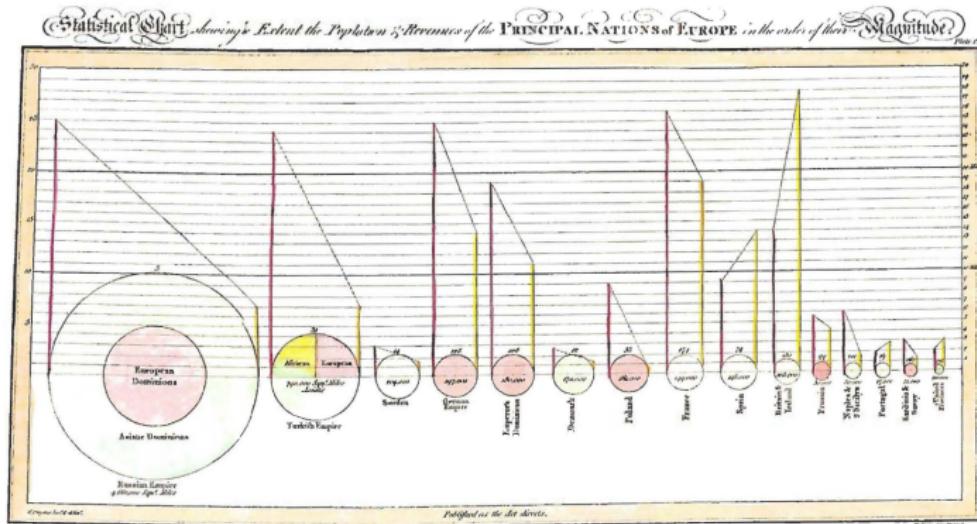
- ⌚ Space and time are the more obvious quantities
- ⌚ Moving away from the started in 1800 with Lambert and Playfair
- ⌚ It took 15 years to Playfair to move way from space and time analogy to abstract design rules

## Relation between quantities

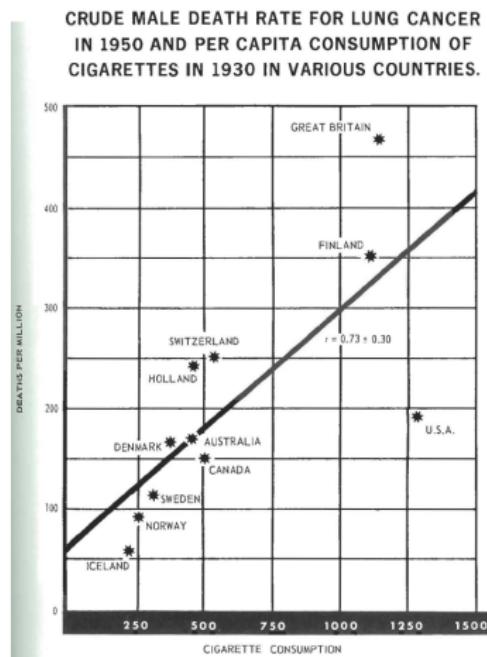
### Relation is the most important thing

- ⌚ Space and time are the more obvious quantities
- ⌚ Moving away from the started in 1800 with Lambert and Playfair
- ⌚ It took 15 years to Playfair to move way from space and time analogy to abstract design rules

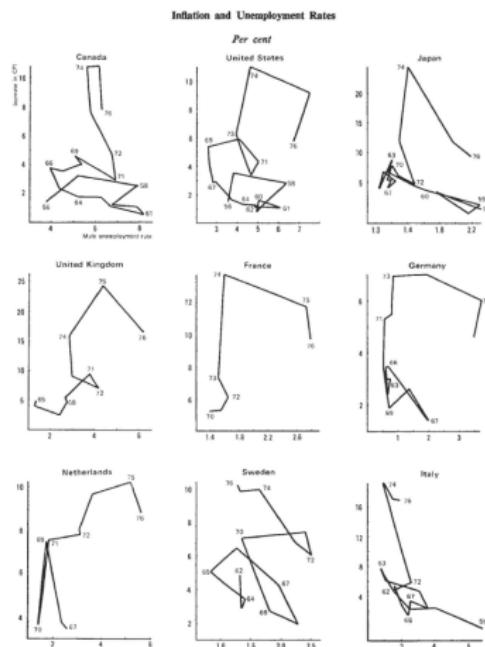
## Playfair: about tax



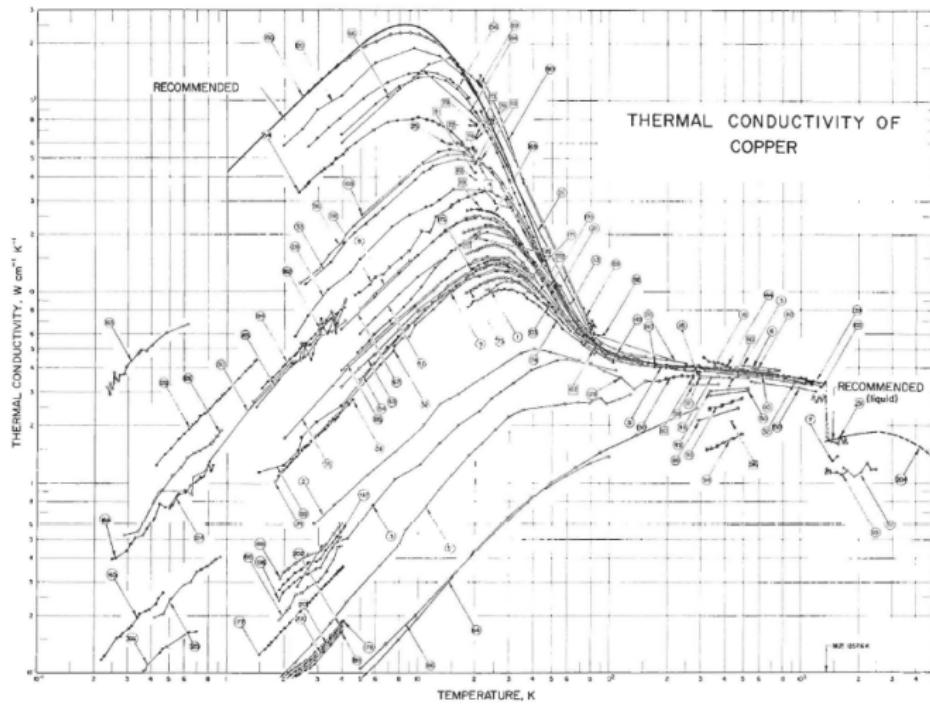
# The scatter plot: 1955



# Evolution of relation in time



# The quest for the truth is NOT a lonely work



# Principles

# Principles of Graphical Excellence

Graphical excellence is the well-designed presentation of interesting data

- ⌚ substance
- ⌚ statistics
- ⌚ design

# Principles of Graphical Excellence

Graphical excellence is the well-designed presentation of interesting data

- ✗ substance
- ✗ statistics
- ✗ design

# Principles of Graphical Excellence

Graphical excellence is the well-designed presentation of interesting data

- ≡ substance
- ≡ statistics
- ≡ design

# Principles of Graphical Excellence

Graphical excellence is the well-designed presentation of interesting data

- ε substance
- ε statistics
- ε design

# Principles of Graphical Excellence

Graphical excellence consists of complex ideas communicated with

- ε clarity
- ε precision
- ε efficiency

# Principles of Graphical Excellence

Graphical excellence consists of complex ideas communicated with

- ε clarity
- ε precision
- ε efficiency

# Principles of Graphical Excellence

Graphical excellence consists of complex ideas communicated with

- ε clarity
- ε precision
- ε efficiency

# Principles of Graphical Excellence

Graphical excellence consists of complex ideas communicated with

- ε clarity
- ε precision
- ε efficiency

# Principles of Graphical Excellence

Graphical excellence is that which gives to the viewer

- ↪ the greatest number of ideas
- ↪ in the shortest time
- ↪ with the least ink in the smallest space.

# Principles of Graphical Excellence

Graphical excellence is that which gives to the viewer

- ↪ the greatest number of ideas
- ↪ in the shortest time
- ↪ with the least ink in the smallest space.

# Principles of Graphical Excellence

Graphical excellence is that which gives to the viewer

- ε the greatest number of ideas
- ε in the shortest time
- ε with the least ink in the smallest space.

# Principles of Graphical Excellence

Graphical excellence is that which gives to the viewer

- ε the greatest number of ideas
- ε in the shortest time
- ε with the least ink in the smallest space.

# Principles of Graphical Excellence

## Graphical excellence

- is nearly always multivariate
- requires telling the truth about the data.

# Principles of Graphical Excellence

## Graphical excellence

- is nearly always multivariate
- requires telling the truth about the data.

# Principles of Graphical Excellence

## Graphical excellence

- is nearly always multivariate
- requires telling the truth about the data.

# To the digital age

## Mike Bostock

- ⌚ data visualizer
- ⌚ Phd Stanford
- ⌚ "Digital Superstar" (New York Times)
- ⌚ Software contribution: <https://d3js.org/>
- ⌚ Example (Paths to victory):  
<http://www.nytimes.com/interactive/2012/11/02/us/politics/paths-to-the-white-house.html>
- ⌚ <https://bost.ocks.org/mike>

# To the digital age

## Mike Bostock

- ⌚ data visualizer
- ⌚ Phd Stanford
- ⌚ "Digital Superstar" (New York Times)
- ⌚ Software contribution: <https://d3js.org/>
- ⌚ Example (Paths to victory):  
<http://www.nytimes.com/interactive/2012/11/02/us/politics/paths-to-the-white-house.html>
- ⌚ <https://bost.ocks.org/mike>

# To the digital age

## Mike Bostock

- ⌚ data visualizer
- ⌚ Phd Stanford
- ⌚ "Digital Superstar" (New York Times)
- ⌚ Software contribution: <https://d3js.org/>
- ⌚ Example (Paths to victory):  
<http://www.nytimes.com/interactive/2012/11/02/us/politics/paths-to-the-white-house.html>
- ⌚ <https://bost.ocks.org/mike>

# To the digital age

## Mike Bostock

- ε data visualizer
- ε Phd Stanford
- ε "Digital Superstar" (New York Times)
- ε Software contribution: <https://d3js.org/>
- ε Example (Paths to victory):  
<http://www.nytimes.com/interactive/2012/11/02/us/politics/paths-to-the-white-house.html>
- ε <https://bost.ocks.org/mike>

# To the digital age

## Mike Bostock

- ε data visualizer
- ε Phd Stanford
- ε "Digital Superstar" (New York Times)
- ε Software contribution: <https://d3js.org/>
- ε Example (Paths to victory):  
<http://www.nytimes.com/interactive/2012/11/02/us/politics/paths-to-the-white-house.html>
- ε <https://bost.ocks.org/mike>

# To the digital age

## Mike Bostock

- ⌚ data visualizer
- ⌚ Phd Stanford
- ⌚ "Digital Superstar" (New York Times)
- ⌚ Software contribution: <https://d3js.org/>
- ⌚ Example (Paths to victory):  
<http://www.nytimes.com/interactive/2012/11/02/us/politics/paths-to-the-white-house.html>
- ⌚ <https://bostocks.org/mike>

# To the digital age

## Mike Bostock

- ⌚ data visualizer
- ⌚ Phd Stanford
- ⌚ "Digital Superstar" (New York Times)
- ⌚ Software contribution: <https://d3js.org/>
- ⌚ Example (Paths to victory):  
<http://www.nytimes.com/interactive/2012/11/02/us/politics/paths-to-the-white-house.html>
- ⌚ <https://bostocks.org/mike>