Elective in Software and Services
(Complementi di software e servizi per la società dell'informazione)

Section Information Visualization

Numbers of credit : 3

Giuseppe Santucci

Course Introduction

Thanks to John Stasko, Robert Spence, Ross Ihaka, Marti Hearst

Outline

- Facts about the course
- Historical examples
- Definitions
  - The Power of Information Visualization
- The problem and the involved issues

Outline

- Facts about the course
- Historical examples
- Definitions
  - The Power of Information Visualization
- The problem and the involved issues

Course resources

- Giuseppe Santucci home page (news, exams, etc.)
  - http://www.dis.uniroma1.it/~santucci/didattica.html
  - Follow the link for this course
  - Or Google giuseppe santucci

- Reference books:
  1. Robert Spence : Information Visualization – Design for interaction 2nd Ed. - Pearson Prentice Hall
  2. Colin Ware : Information Visualization, Second Edition: Perception for Design 2nd Ed. – Elsevier
  3. Stephen few : Show me the numbers – Analytics Press

- Office Hours: Tuesday 14.30-16.30 Via Ariosto 25 room B218
  - Always have a look at news before coming !

Lectures

- The lectures of this section will be given in the second semester in Via Ariosto 25, as follows:

- https://docs.google.com/a/dis.uniroma1.it/spreadsheet/ccc?key=0AhWTPlxe-sdhdDE1MzFvSXdjMVhoUi04S3FCVTE0c2c&usp=sharing#gid=1

- Have a look at

https://piazza.com/uniroma1.it/spring2015/1038134/resources

- For the detailed schedule

Program

- Introduction (today)
- Number visualization
  - Common errors & lies
  - Kinds of number
  - Table and graphs
- Representation
  - Encoding
  - Type of data – Univariate / Multivariate data
  - Data and relationship
  - Perceptual issues
- Presentation
  - Space limitation
  - Time limitation
- Interaction
  - Continuous interaction
  - Stepped interaction
- Case studies

What we are *not* covering

- Scientific visualization
- Statistics
- Cartography (maps)
- Education
- Games
- Computer graphics in general

Outline

- Facts about the course
- Historical examples
- Definitions
  - The Power of Information Visualization
- The problem and the involved issues

Visualization ?

1. Problem solving / Analyzing
2. Explaining
3. Making decision

Problem Solving/Analyzing

Mystery: what is causing a cholera epidemic in London in 1854?

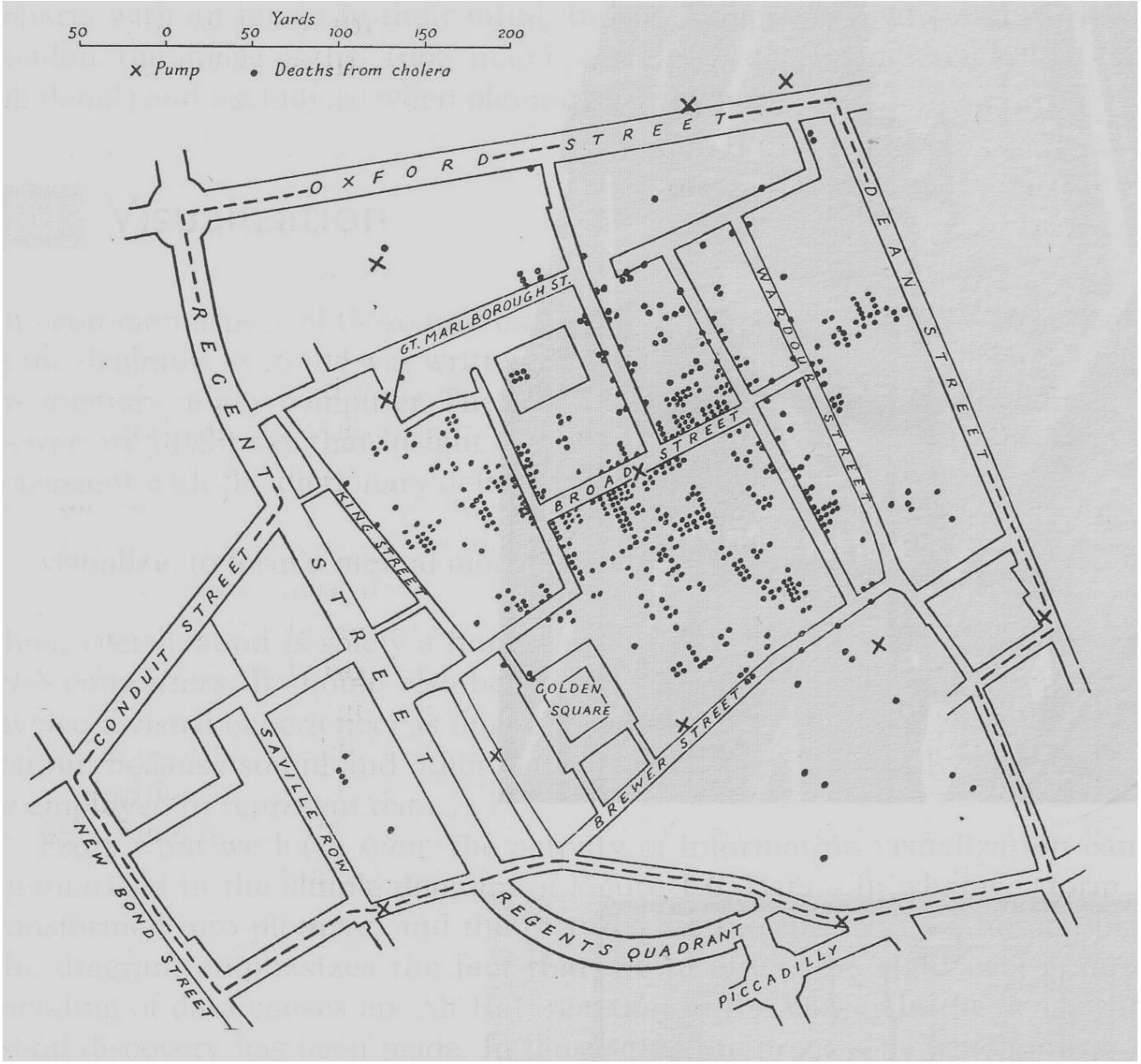
Visualization for Problem Solving

Illustration of Dr. John Snow (1854).

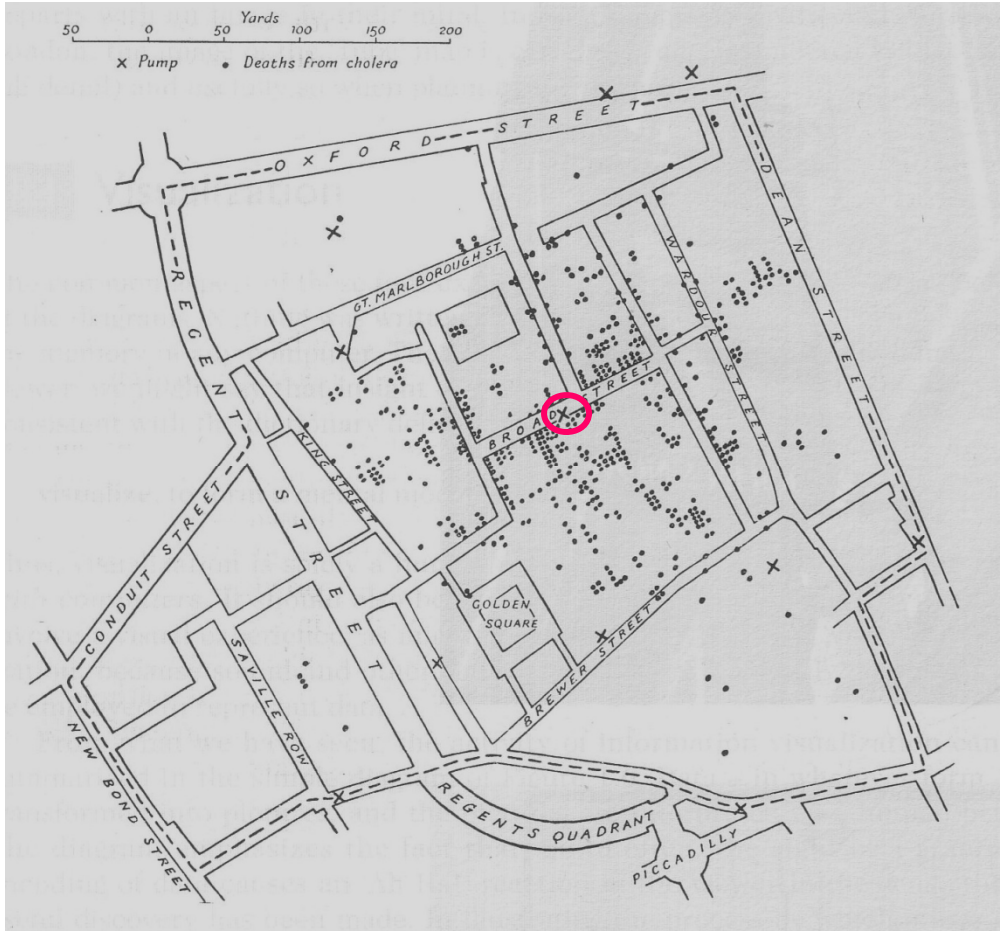
Dots indicate location of deaths.

X indicate the location of water pumps

From Visual Explanations by Edward Tufte, Graphics Press, 1997



Visualization for Problem Solving



The actual John Snow pub in London close to the water pump !!!

John Snow deducted that the cholera epidemic was caused by a bad water pump !!!
Closing that pump quickly solved the problem

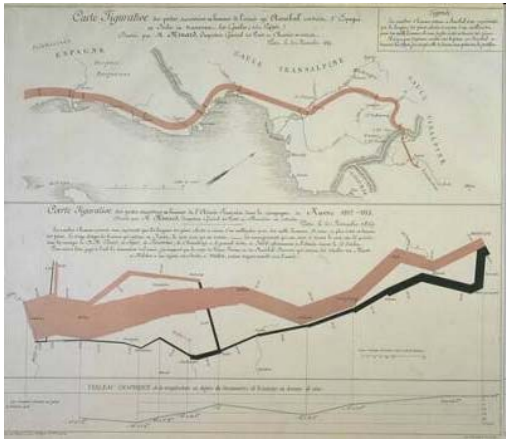
B.T.W., workers at the nearby brewery were noted to be relatively free of cholera…

Explaining

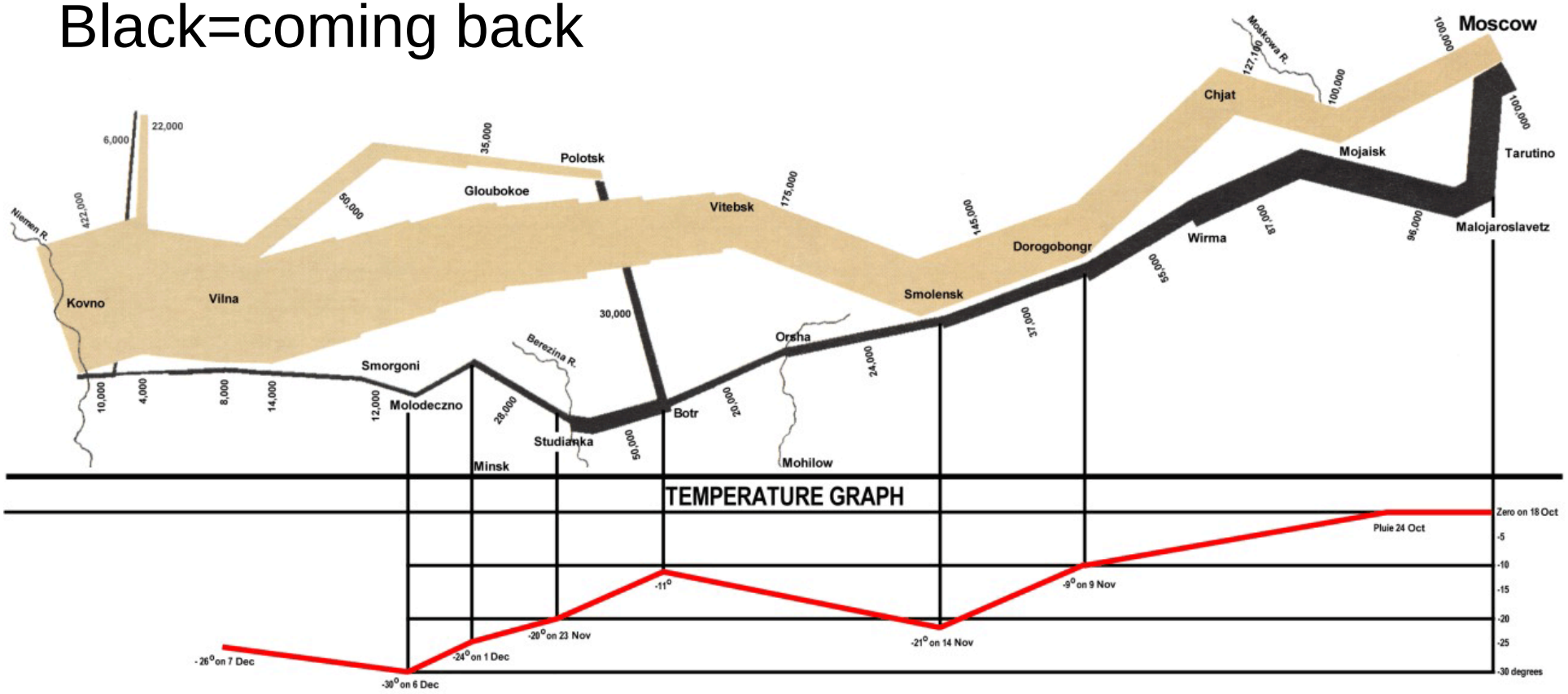
What happened during the Napoleon’s Russian Campaign?

Russian campaign of 1812

Charles Joseph Minard (1781 – 1870)



Size= number of soldiers
Brown=going to Moscow
Black=coming back



Visualization for Making decision

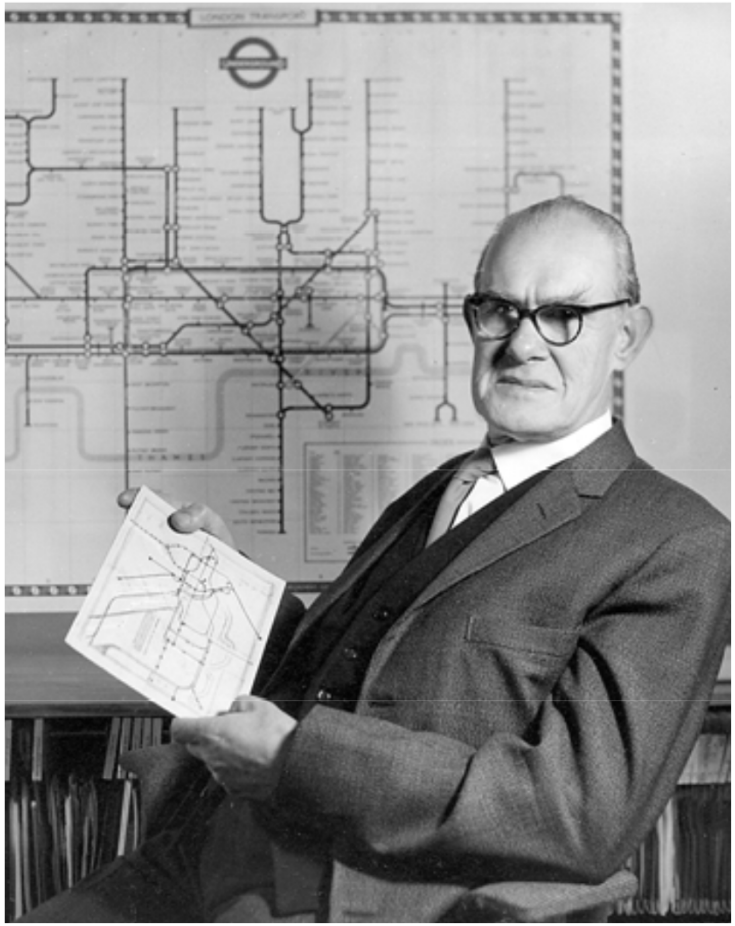
Traveling in London by underground

How can I get Queens Park from Victoria?

London Underground Map 1927

The Harry Beck's idea

- Real position (when traveling in underground) does not matter
- Only station sequences matter together with their connections
- Beck proposed a "distorted" map
- Actually all the underground maps in the world follow the Beck's approach
- He got a little payment (London underground was not sure about the idea)
- Still true right now: infovis people do not become rich…



London Underground Map 1990s

Outline

- Facts about the course
- Historical examples
- Definitions
  - The Power of Information Visualization
- The problem and the involved issues

Moving to the present time

- What is Information Visualization ?
- First of that, what is Visualization ?
- Visualize: to form a mental model or mental image of something.
- It is a **cognitive activity** and it has nothing to do with computers.

What is Information Visualization?

“Transformation of the symbolic into the geometric”

(McCormick et al., 1987)

“... finding the artificial memory that best supports our natural means of perception.‘‘

(Bertin, 1983)

“Information visualization is the use of computer-supported, *interactive*, visual representations of *abstract data* to *amplify cognition*.”

(Card et al., 1999)

What is Information Visualization?

Information visualization is the use of *computer-supported*, *interactive*, visual representations of *abstract data* to *amplify cognition*.

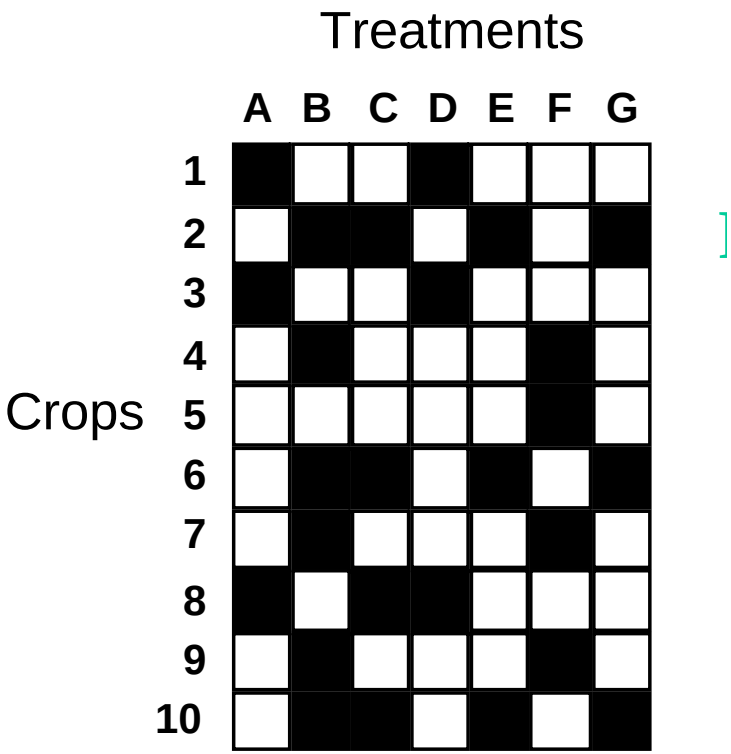
[Card et al. ‘99]

…computer supported and interactive

- **Computer-supported**
  - Even beautiful examples of paper based visualizations exist the actual understanding of information visualization (infovis) is about computer based visualization, **but we have to always remember that a cognitive activity is involved in the process**
- **Interactive**
  - To exploit the full power of infovis techniques interaction is mandatory. The user must be allowed for manipulating the visualization to better reach his goals

Interaction example

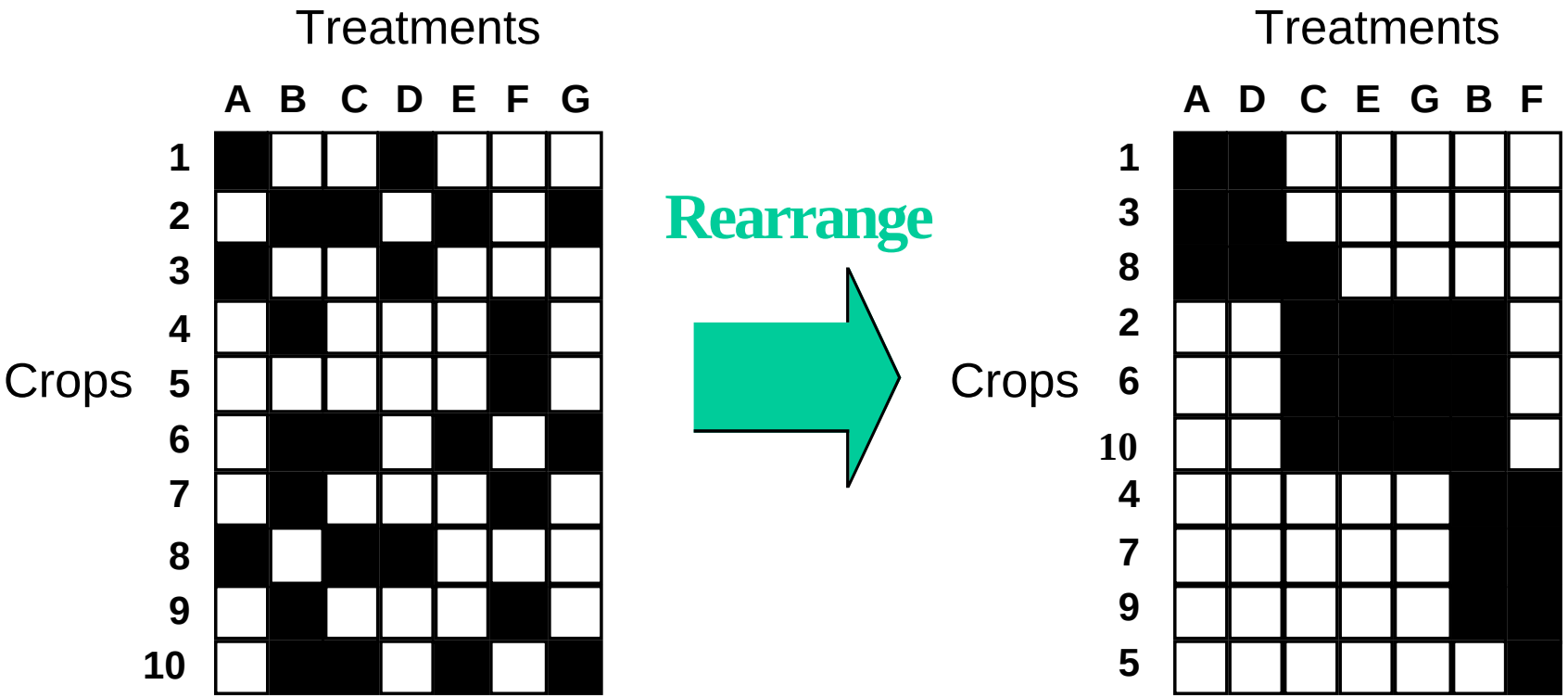
- Agronomists are experimenting 7 treatments (anti-parasite, fertilizer, etc.) on 10 different crops
- A black square indicates success
- Does this visualization help?



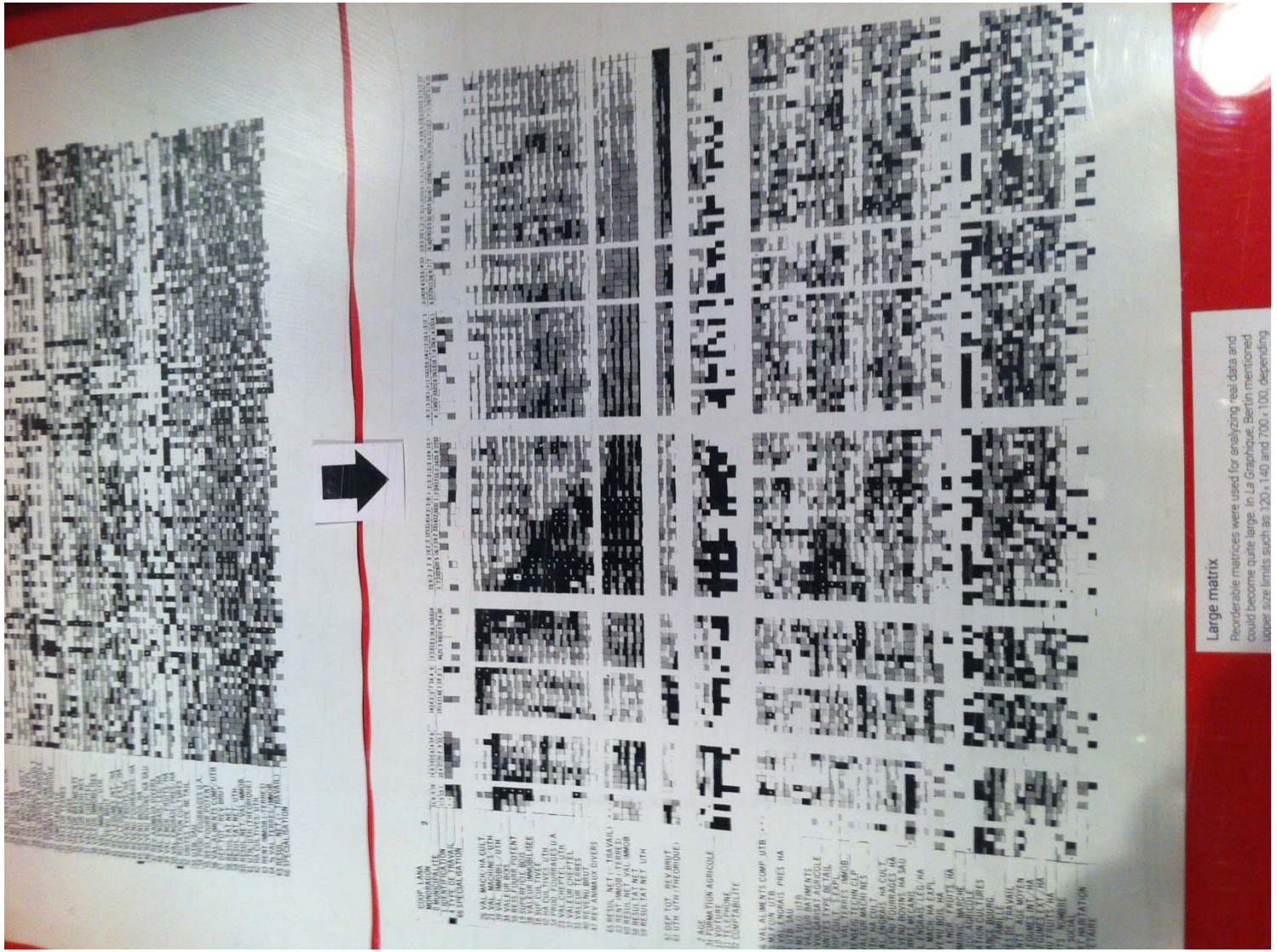
Interaction example

Let’s rearrange the columns

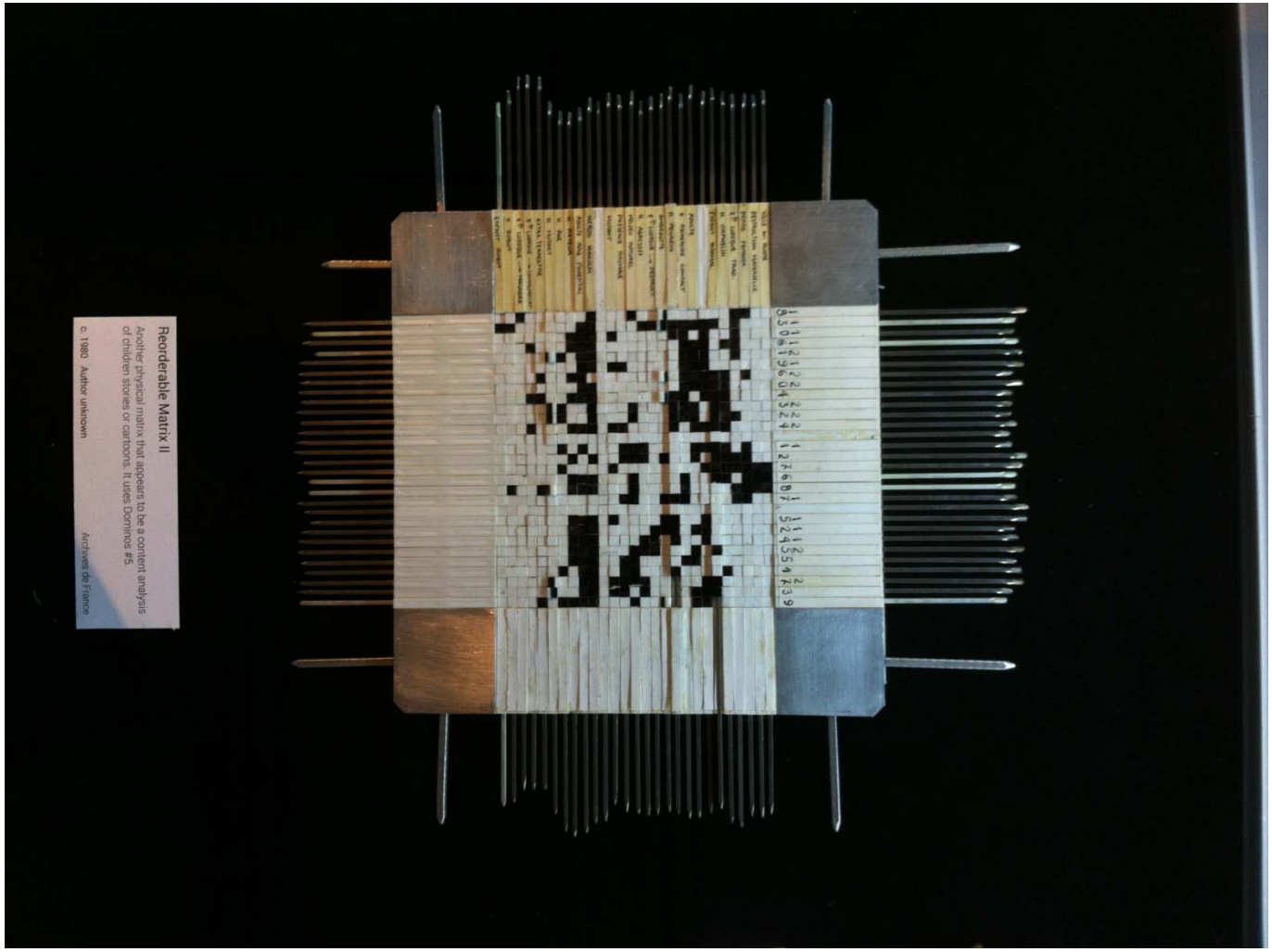
- thanks to Jacques *Bertin* (27 July 1918 – 3 May 2010)



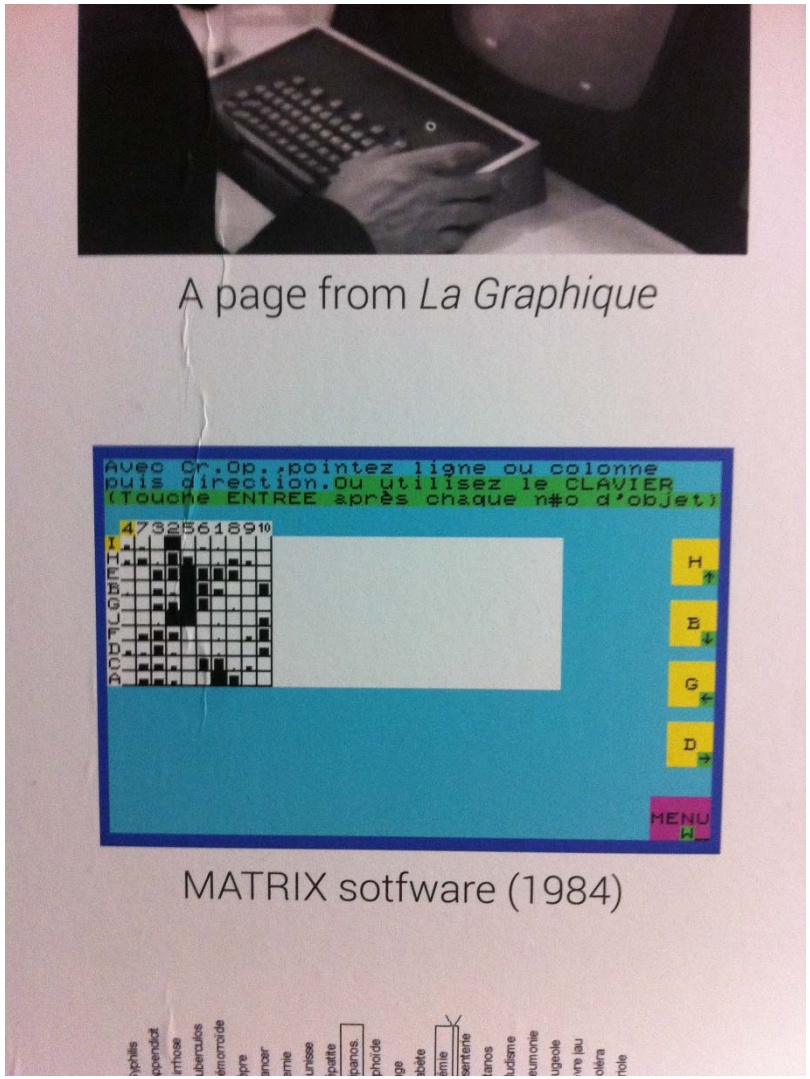
Bertin’s work



Manually!


Reorderable Matrix II
Another physical matrix that appears to be a content analysis of children stories or cartoons. It uses Dominos #3.
c. 1980 Author unknown
Archives de France

Computer !



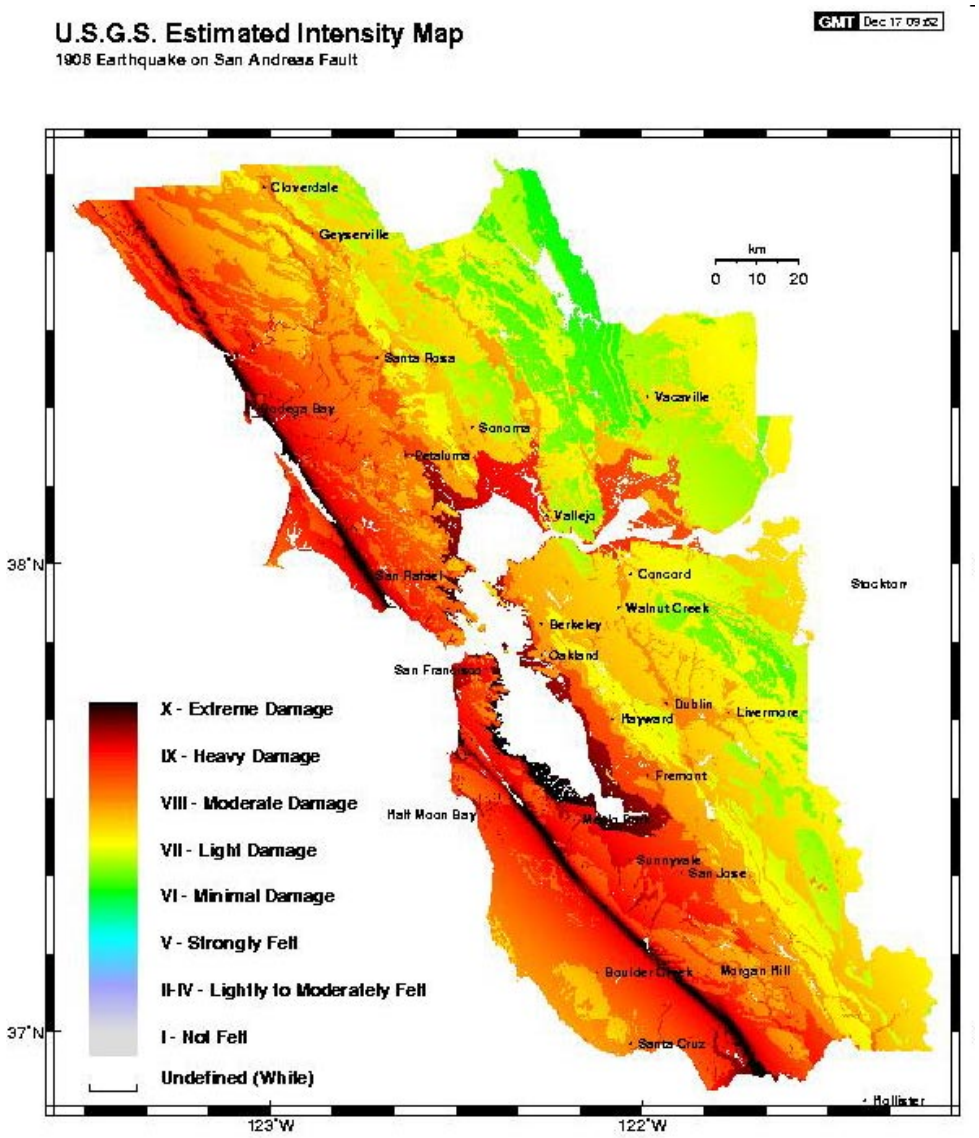
A page from *La Graphique*

MATRIX sotfware (1984)

…it is about abstract data

- Abstract data
  - Information visualization is about visualizing abstract data, i.e., presenting images that does not refer to physical situation. In other words it is NOT scientific visualization/geographic visualization
- Scientific visualization primarily relates to and represents something physical or geometric
- Examples
  - Air flow over a wing
  - Weather over Italy
  - Torrents inside a tornado
  - Organs in the human body
  - Molecular bonding…

Scientific/geographic visualization

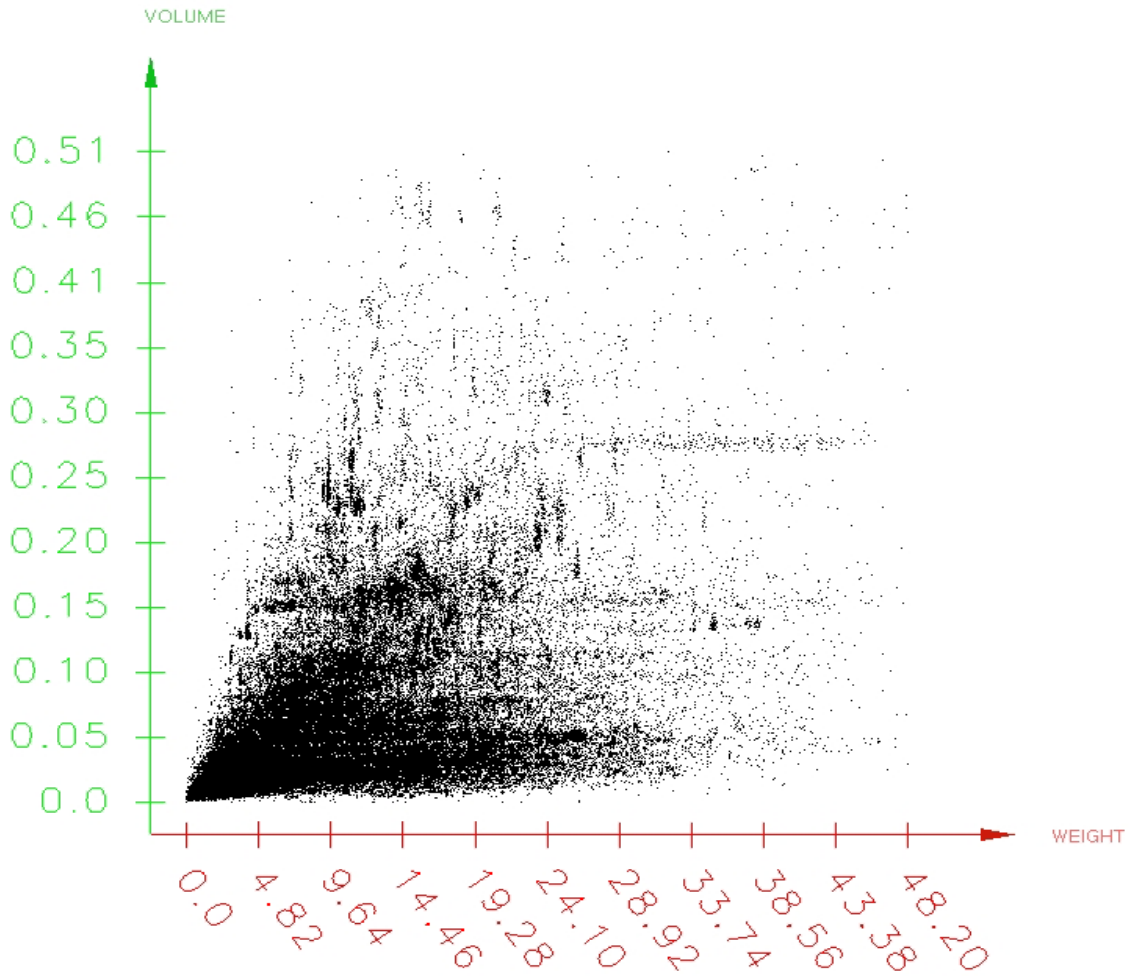


Earthquake intensity

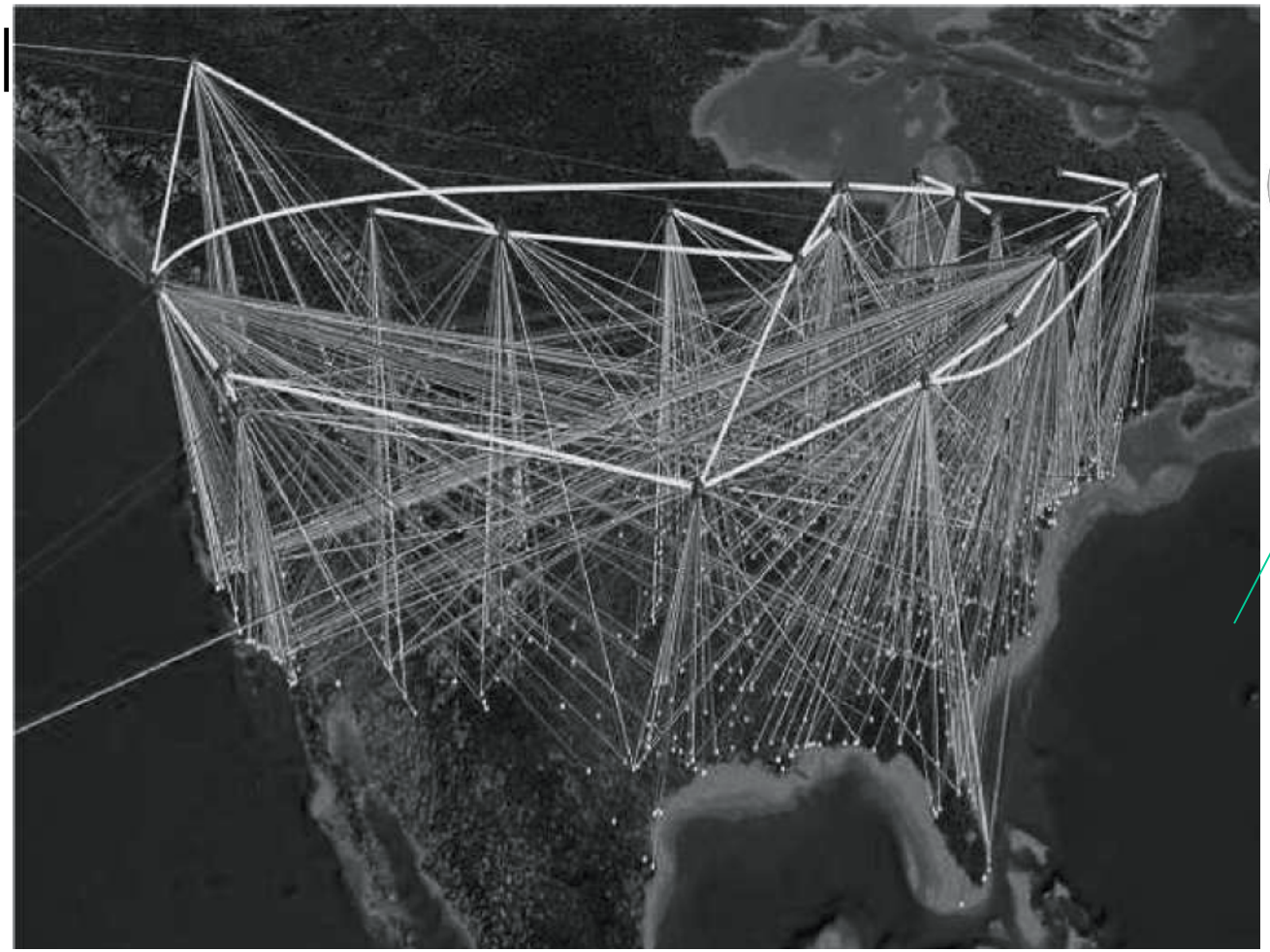
…abstract data

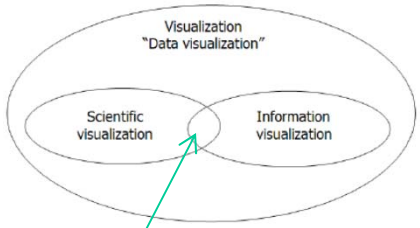
- What is “information”
  - Items that do not have a direct **physical/visual** correspondence (or such a correspondence is not relevant for the application)
  - Examples: sport statistics, stock trends, query results, software data, etc…
- Items are represented on a 2D / 3D physical space using their numerical characteristics (attributes)
- The visualization is useful for analysis and decision-making (not just fun or colors)
- E.g. Postal parcels
  - Shipping date
  - Volume
  - Weight
  - Sender country
  - Receiver country
  - …

Abstract data



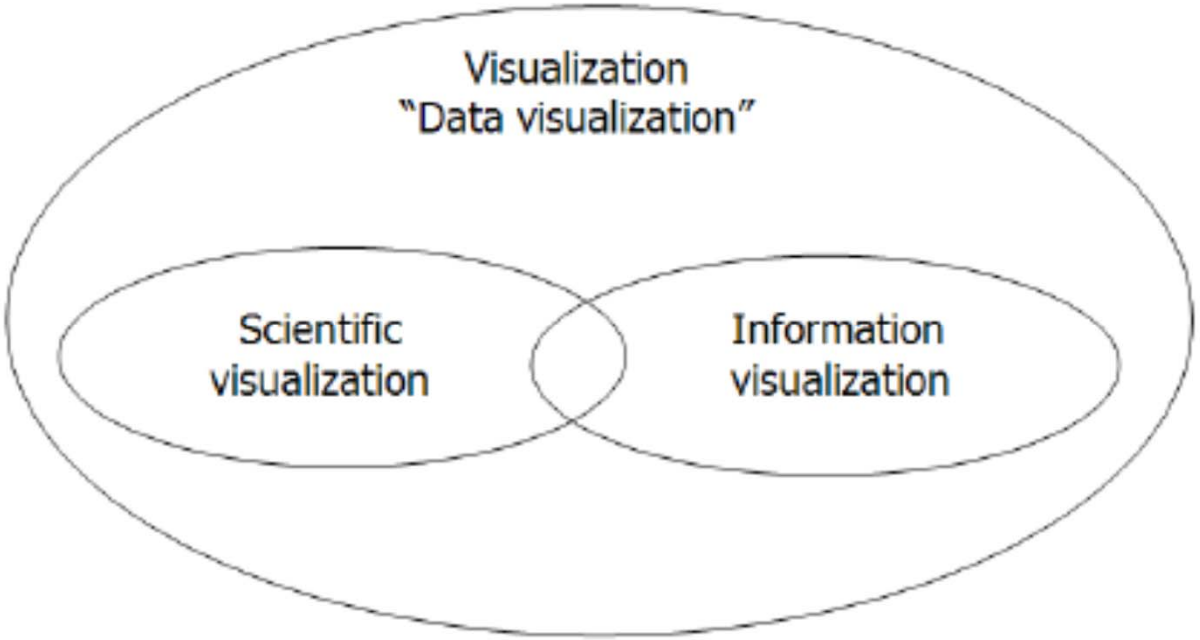
A 2D Scatterplot showing about 200.000 postal parcels





Byte traffic into the ANS/NSFNET T3 backbone in 1993

Overview


Visualization
"Data visualization"
Scientific
visualization
Information
visualization

… it amplifies cognition

- Facts about the course
- Historical examples
- Definitions
  - The Power of Information Visualization
- The problem and the involved issues

Amplify cognition using the human vision

- Highest bandwidth sense
- Fast, parallel
- Pattern recognition
- Pre-attentive
- Extends memory and cognitive capacity
  - Multiplication test
- People think visually (I see… means also I understand in most languages)
- Amplify cognition
  - Presenting data in the right way, taking into account the way in which human vision system works can greatly improve the comprehension of complex phenomena
- Three very simple examples (put away pencil and paper…)

Amplify cognition

- Example: multiplication (Card, Moran, & Shneiderman.)

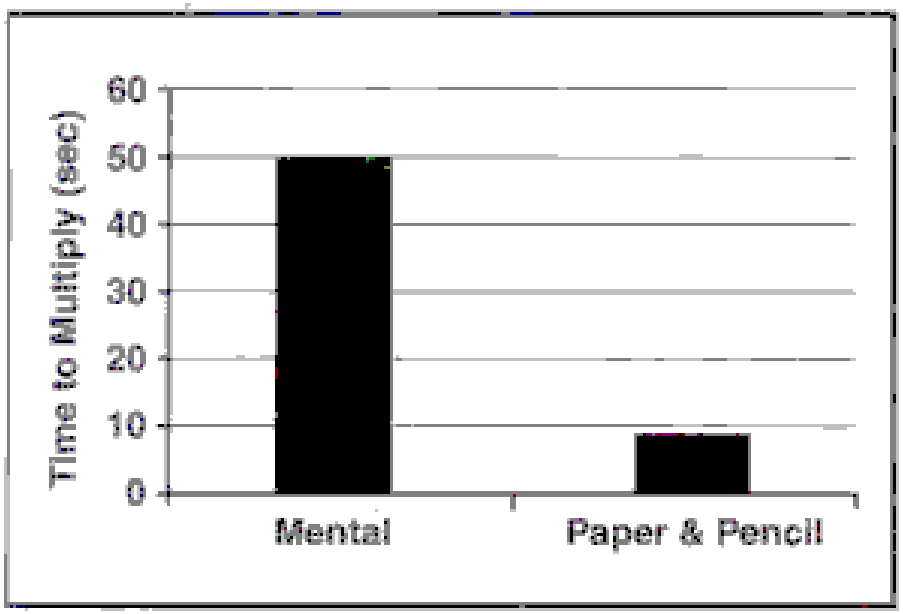
- In your head, multiply 35 x 95

Amplify cognition

- Now do it on paper

Visual Aids for Thinking

– People are 5 times faster with the visual aid



Three simple questions

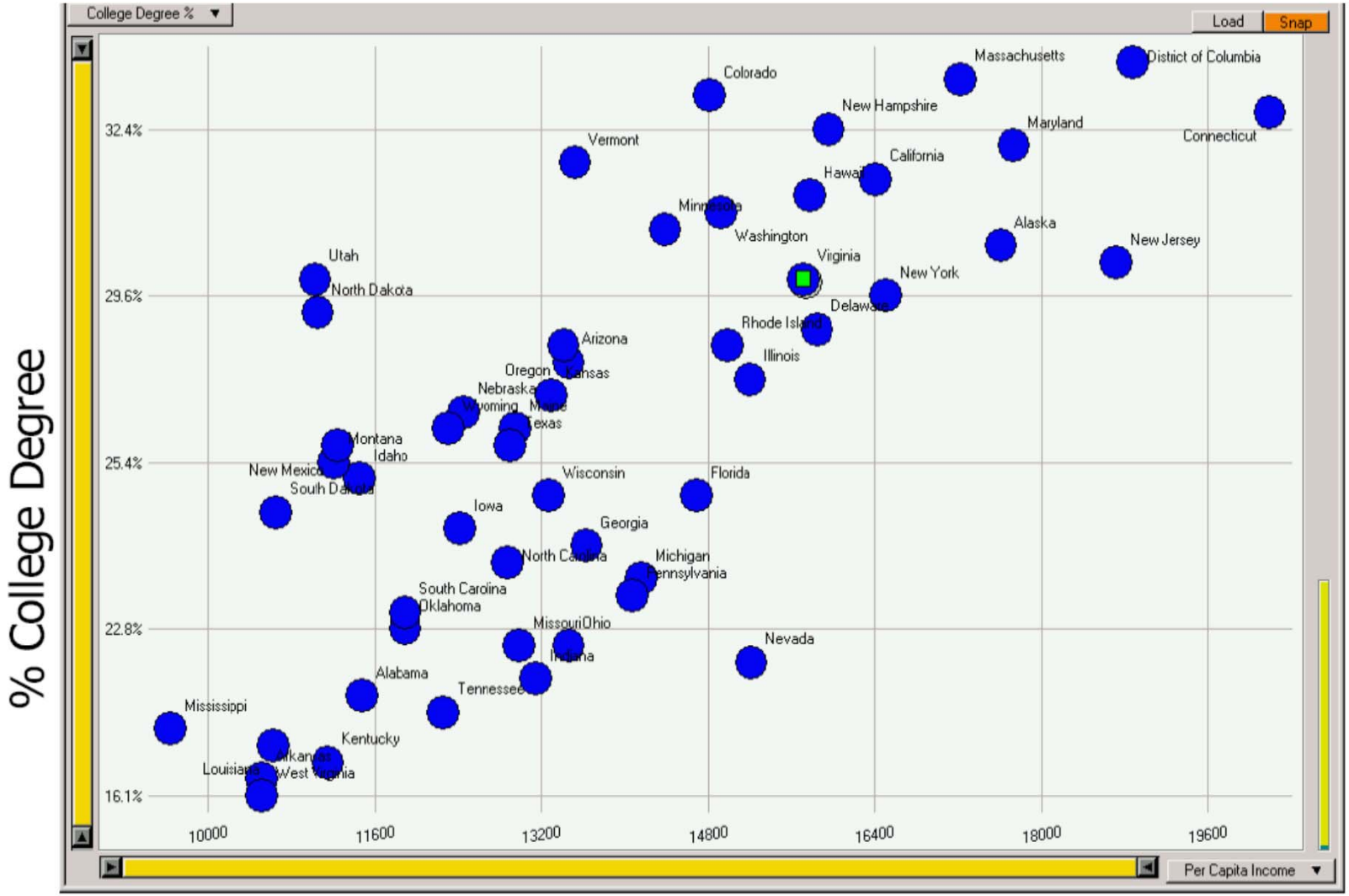
Table - StateData ()

Load | Snap

State	College Degree %	Per Capita Income
Alabama	20.6%	11486
Alaska	30.3%	17610
Arizona	27.1%	13461
Arkansas	17.0%	10520
California	31.3%	16409
Colorado	33.9%	14821
Connecticut	33.8%	20189
Delaware	27.9%	15854
District of Columbia	36.4%	18881
Florida	24.9%	14698
Georgia	24.3%	13631
Hawaii	31.2%	15770
Idaho	25.2%	11457
Illinois	26.8%	15201
Indiana	20.9%	13149
Iowa	24.5%	12422
Kansas	26.5%	13300
Kentucky	17.7%	11153
Louisiana	19.4%	10635
Maine	25.7%	12957
Maryland	31.7%	17730
Massachusetts	34.5%	17224
Michigan	24.1%	14154
Minnesota	30.4%	14389
Michigan	24.1%	14154
Minnesota	30.4%	14389
Mississippi	19.9%	9648
Missouri	22.3%	12989
Montana	25.4%	11213
Nebraska	26.0%	12452
Nevada	21.5%	15214
New Hampshire	32.4%	15959
New Jersey	30.1%	18714
New Mexico	25.5%	11246
New York	29.6%	16501
North Carolina	24.2%	12885
North Dakota	28.1%	11051
Ohio	22.3%	13461
Oklahoma	22.8%	11893
Oregon	27.5%	13418
Pennsylvania	23.2%	14068
Rhode Island	27.5%	14981
South Carolina	23.0%	11897
South Dakota	24.6%	10661
Tennessee	20.1%	12255
Texas	25.5%	12904
Utah	30.0%	11029
Vermont	31.5%	13527
Virginia	30.0%	15713
Washington	30.9%	14923
West Virginia	16.1%	10520
Wisconsin	24.9%	13276
Wyoming	25.7%	12311

Which state has the highest % college degree?
Highest Income? Relationship between college and income?

The quick answers



One (very) simple question

- How many 3s here ?
- You have 4 seconds…

Game over!

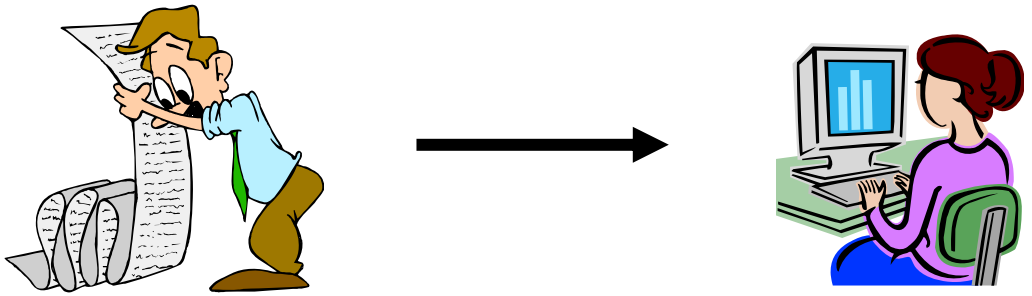
So ?

- Time is not enough?
- You can do that in less than one second !

458757626808609928083982698028
747976296262867897187743671947
746588786758967329667287682085

- Color is pre-attentive (pops up)
- No cognitive effort is required
- A lot of issues are already clear
- Most of people ignore them...

Information visualization

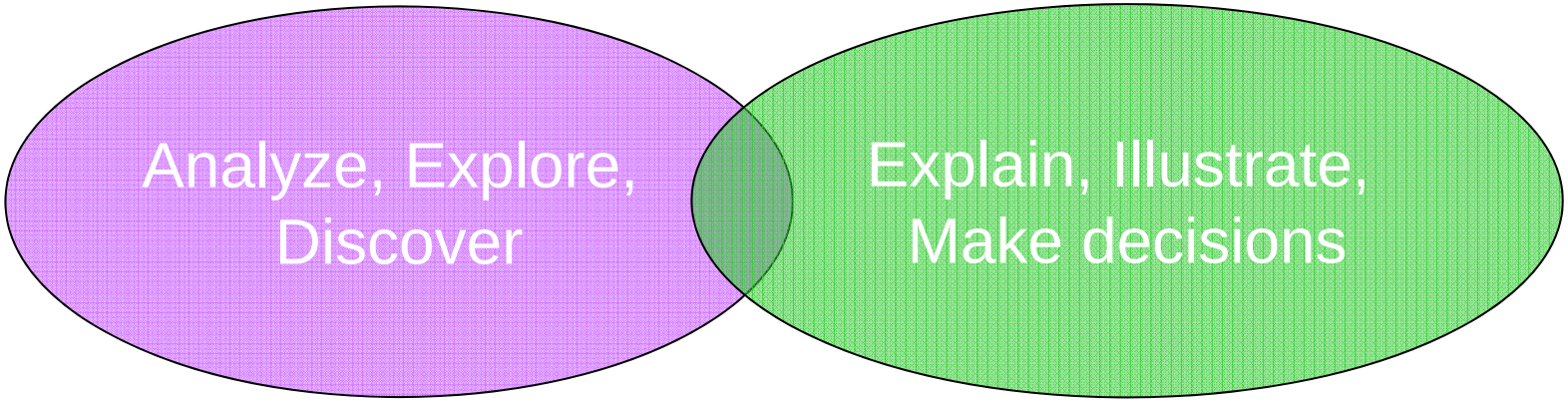


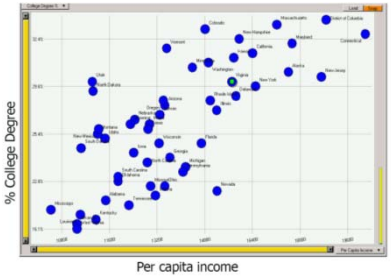
1. Infovis is perfect for exploration, when we don't know exactly what to look at. It supports vague goals

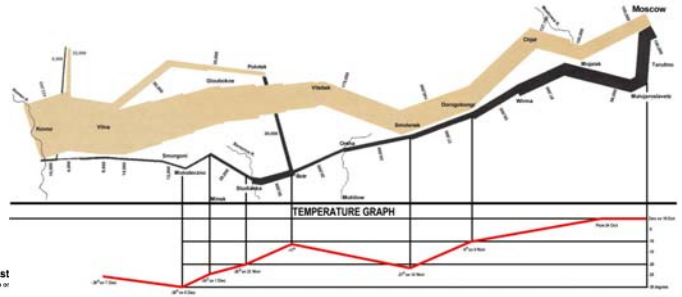
2. Infovis is perfect to explain complex data and to support decisions

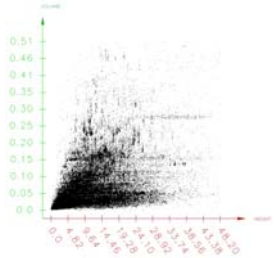
- Other approaches to data analysis
  - Statistics: strong verification but does not support exploration and vague goals
  - Data mining: actionable and reliable but black box, not interactive, question-response style
  - **Visual analytics** (formerly Visual Data Mining) is trying to join the two worlds

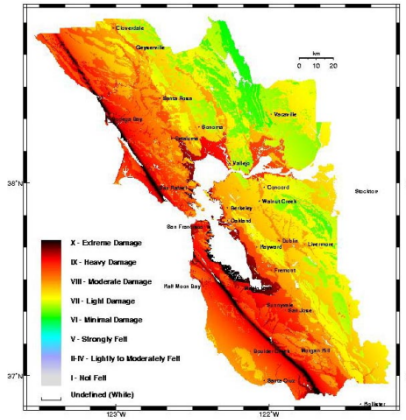
Visualization: Two Primary Goals





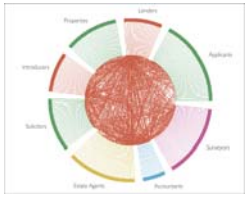


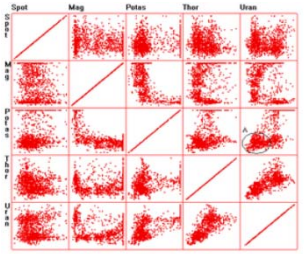


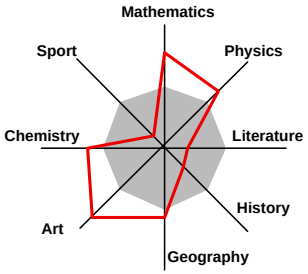


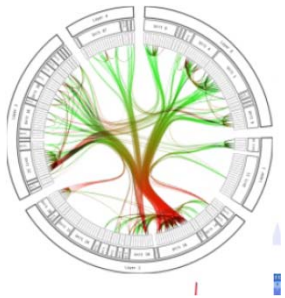
Canonical steps in infovis – STEP 1

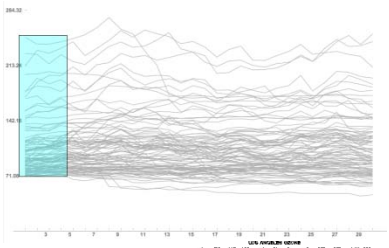
DATA → Internal **Representation**

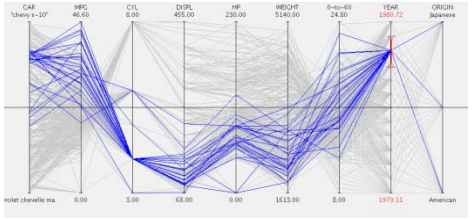


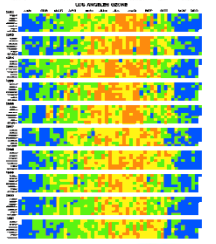












Encoding of values
- Univariate data
- Bivariate data
- Trivariate data
- Multidimensional data

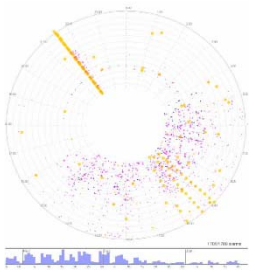
Encoding of relations

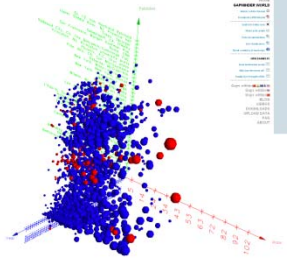
Temporal data

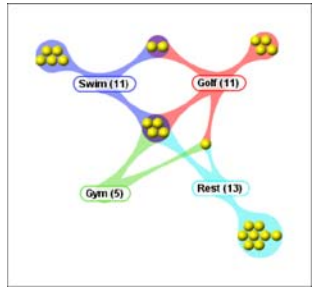
Map & Diagrams

Graphs/Trees

Data streams

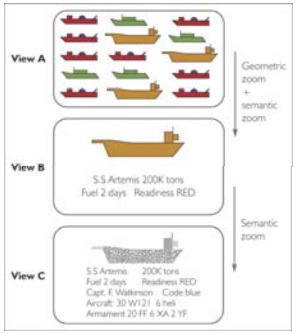




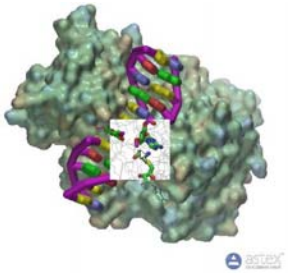


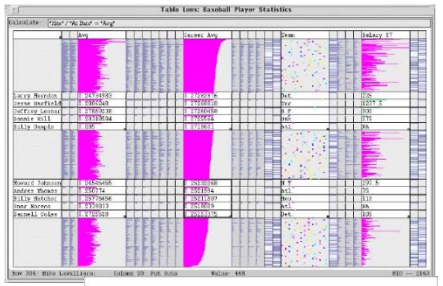
Canonical steps in infovis – STEP 2

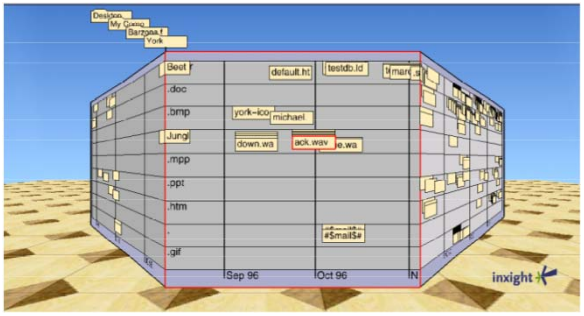
Internal **Representation**



Presentation





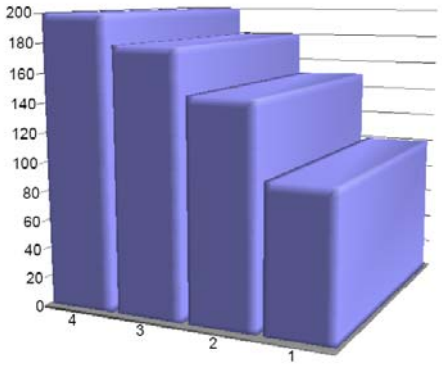


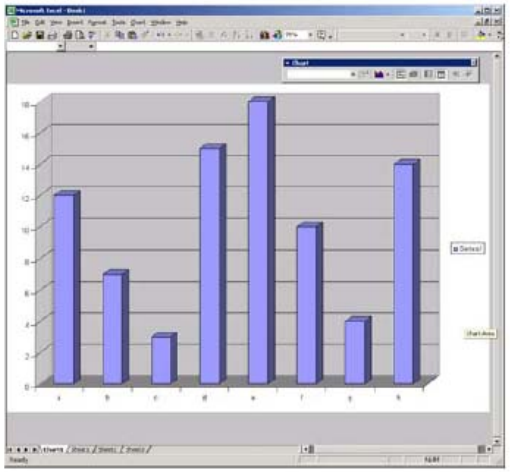
- Space limitations
  - Scrolling
  - Overview + details
  - Distortion
  - Suppression
  - Zoom & pan
  - Semantic zoom
- Time limitation
- Perceptual issues
- Cognitive issues

Outline

- Facts about the course
- Historical examples
- Definitions
  - The Power of Information Visualization
- The problem and the involved issues

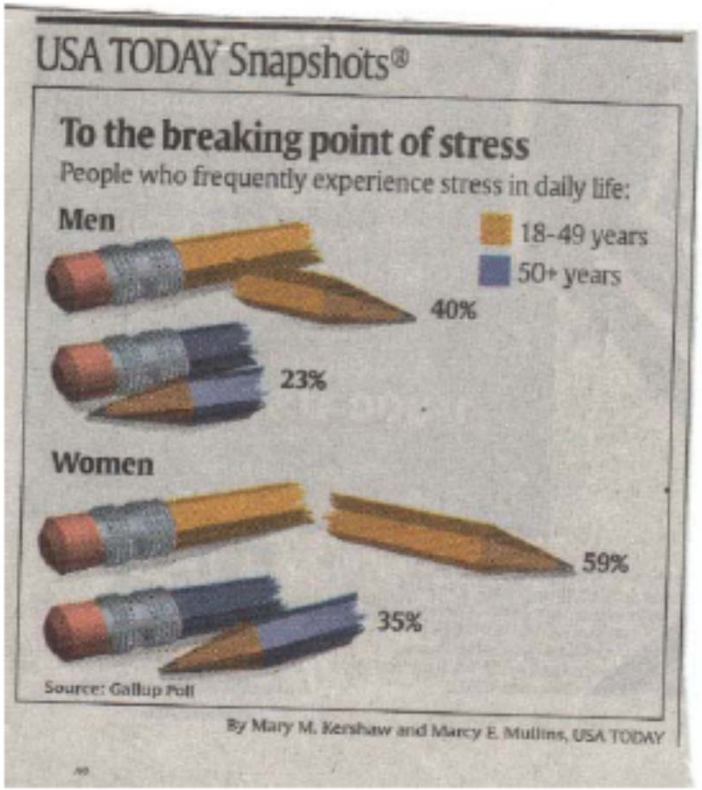
So, let’s visualize





- Please, GET RID of those darn 3D bars !!!!!

So, let's visualize



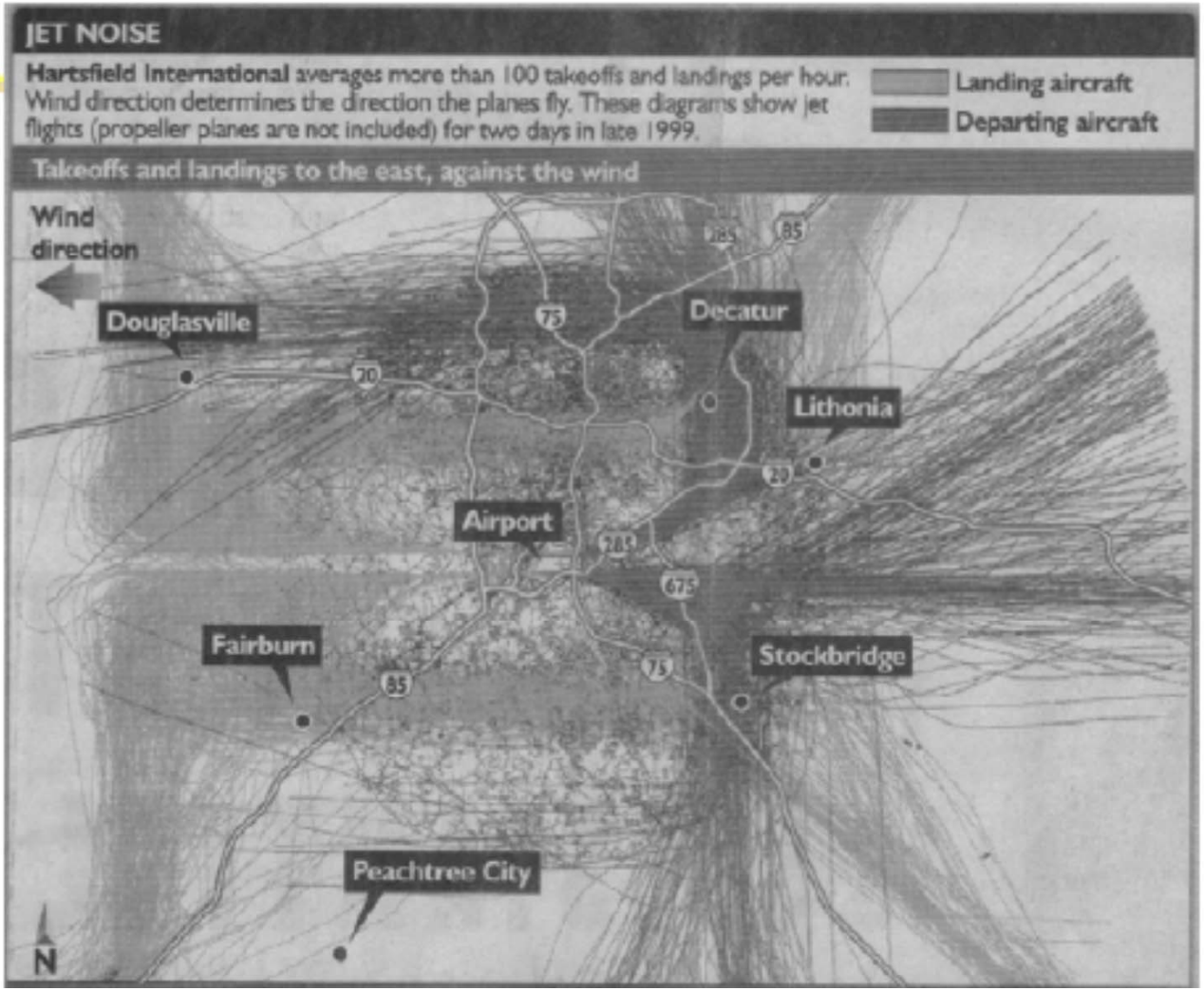
- Or worse yet…

So, let's visualize

- Make a simple decision here …

So, let's visualize

Atlanta Flight Traffic



Atlanta Journal
April 30, 2000

…

So, let's visualize

Power Costs

Average cost per month to use

Wall Street Journal
August 16, 2001

- …

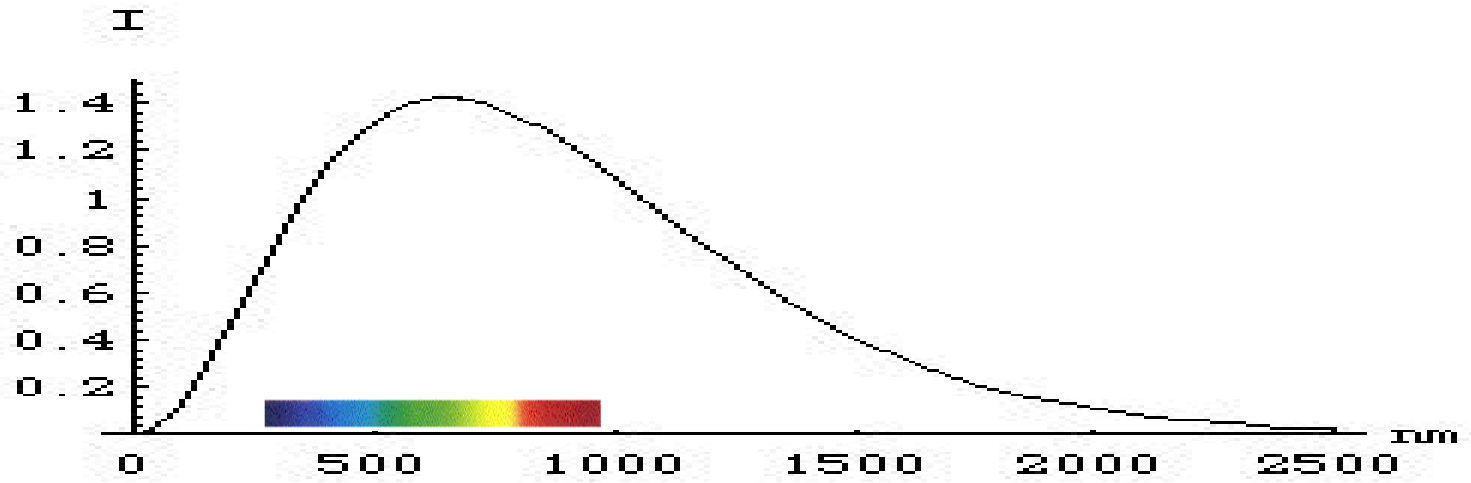
Visualization is not enough

- Perceptual, cognitive issues

- Why this green laser is so strong ?
  - Big battery?
  - More power ?

Visualization is not enough

- ?
- Human eyes are more sensible to green (555 nm) than low-red and hi-blue !
- More means 100 times!!!!



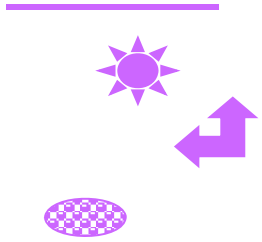
Human Perceptual Facilities

Use the eye for pattern recognition; people are good at

- scanning
- recognizing
- remembering images

Graphical elements facilitate comparisons via

- length
- shape
- orientation
- texture



Animation shows changes across time

Color and other pre-attentive features helps make distinctions

But … How many colors can human eyes distinguish in pre-attentive way? (about 6 / 12 ☹ …)

Focusing problem / effect

Most people see the red

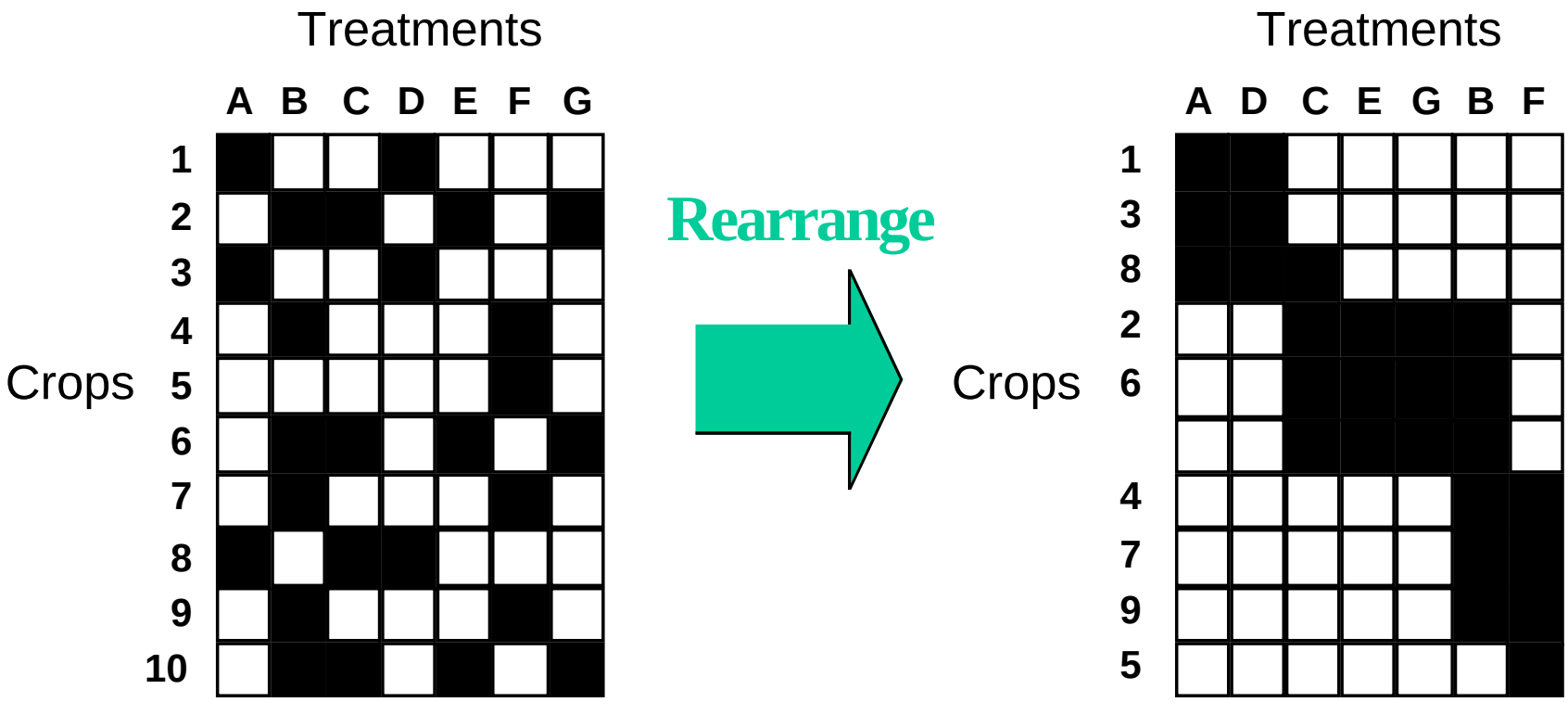
Closer than the blue

But some see the

Opposite effect

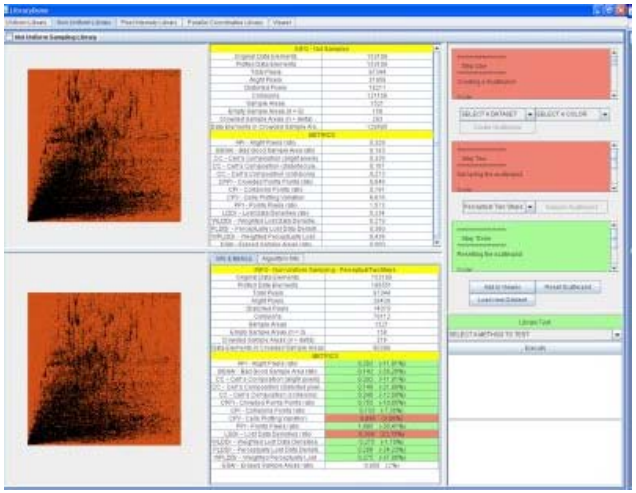
Visualization is not enough

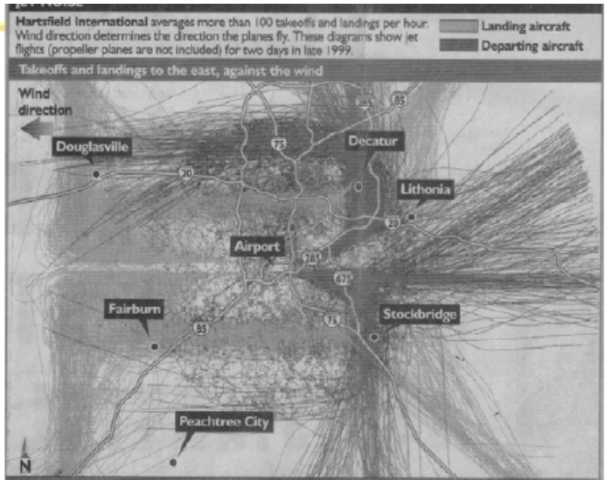
- Interaction is not a plus !



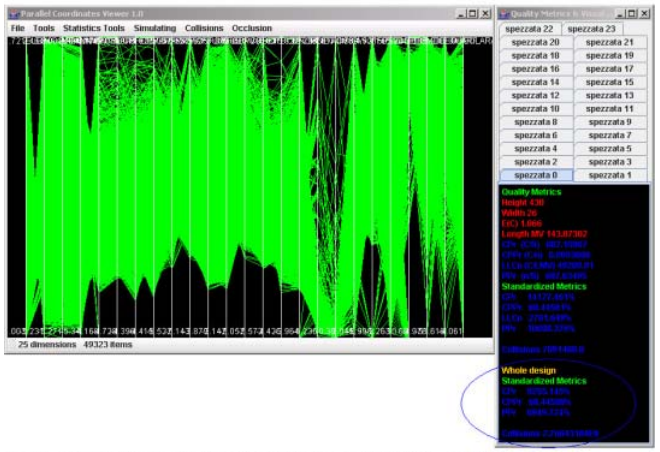
Visualization is not enough

- Data are complex (a lot of attributes) and dataset can be very large !
- How to manage complexity?

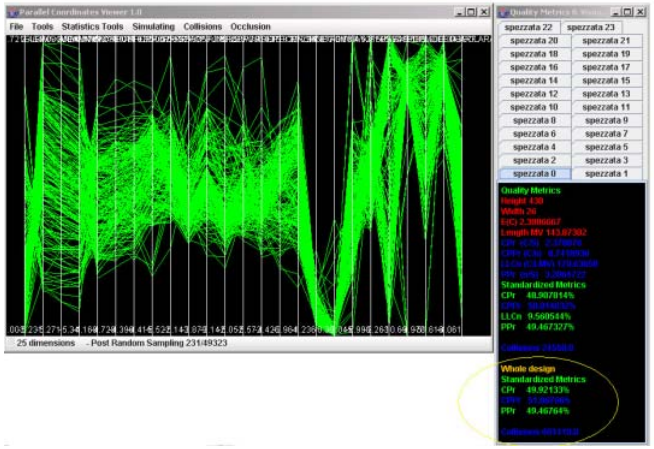




Rappresentazione iniziale



Rappresentazione dopo il campionamento guidato mediante la metrica



Visualization is not enough

- User task (and skill) must be considered !
- Learning time: visualizations are not simple!!!!!
- Evaluation of visual system is not trivial

- Others:
  - Analysis
  - Monitoring
  - Planning
  - Communication
- Tufte:
  - Description
  - Exploration
  - Tabulation
  - Decoration

- Others:
  - Aid to thinking
  - Problem solving/Decision making
  - Insight
  - Clarifying
  - Entertainment / Art

Another way to think about user tasks

- Answer this question: Do you know the answer?
  - If yes,
    - Presentation, communication, education
  - If no,
    - Exploration, analysis
    - Problem solving, planning,
    - Aid to thinking, reasoning
- Answer this question: Are you the creator or the viewer of the information?
  - Often there is a loop between analysis and presentation

Visualization is not enough

- How develop the software ?
- From scratch or using an infovis toolkit ?

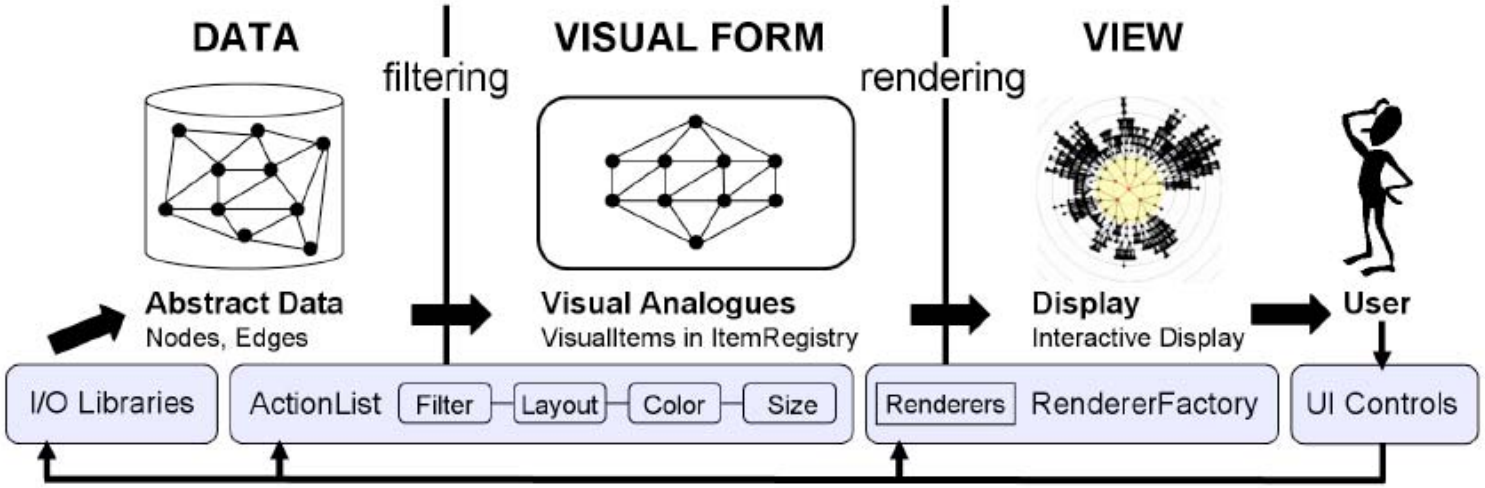


Figure 2. The prefuse visualization framework. Lists of composable actions filter abstract data into visualizable content and assign visual properties (position, color, size, font, *etc*). Renderer modules, provided on a per-item basis by a RendererFactory, draw the VisualItems to construct interactive Displays. User interaction can then trigger changes at any point in the framework.