Milestones in the History of Data Visualization

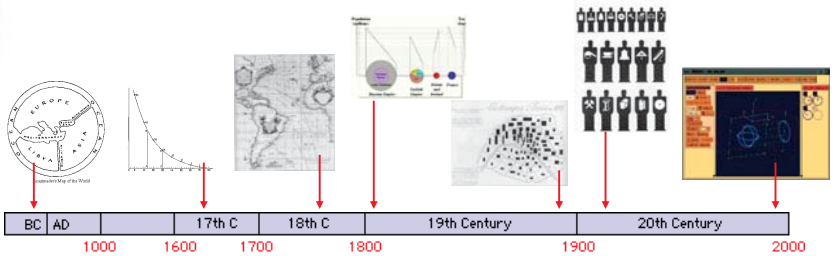
CARME 2003

A case study in statistical historiography

{flea bites man, bites flea, bites man}-wise

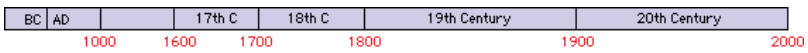
Michael Friendly, York University

CARME 2003



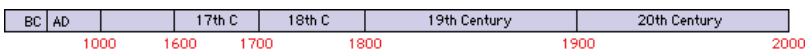
Outline

- Introduction
  - Milestones Project: overview
  - Background
  - Data and Stories
- Milestones tour
- Problems of statistical historiography
  - What counts as a milestone?
  - What is "data"
  - How to visualize?



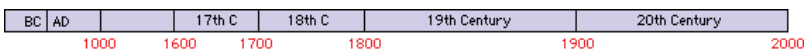
Milestones: Project Goals

- Comprehensive catalog of historical developments in *all fields* related to data visualization.
- → Collect representative bibliography, images, cross-references, web links, etc.
- → Enable researchers to find/study themes, antecedents, influences, patterns, trends, etc.
- Web: http://www.math.yorku.ca/SCS/Gallery/milestone/



Milestones: Conceptual Overview

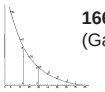
- Roots of Data Visualization
  - **Cartography**: map-making, geo-measurement thematic cartography, GIS, geo-visualization
  - **Statistics**: probability theory, distributions, estimation, models, stat-graphics, stat-vis
  - **Data**: population, economic, social, moral, medical, …
  - **Visual thinking**: geometry, functions, mechanical diagrams, EDA, …
  - **Technology**: printing, lithography, computing…



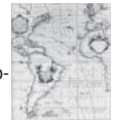
Milestones: Content Overview

Every picture has a story – Rod Stewart

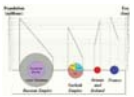
c. 550 BC: The first world map? (Anaximander of Miletus)



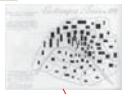
1669: First graph of a continuous distribution function (Gaunt's life table)– Christiaan Huygens.



1782: First topographical map- M. du Carla-Boniface



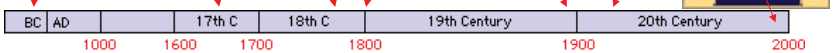
1801: Pie chart, circle graph - William Playfair



1896: Area rectangles to display two variables and their product- Jacques Bertillon

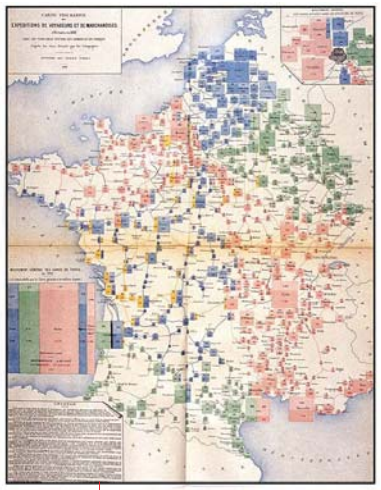
1924: Museum of Social Statistical Graphics- Otto Neurath

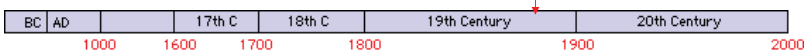
1991-1996: Interactive data visualization systems (Xgobi, ViSta)



Background: *Les Albums*

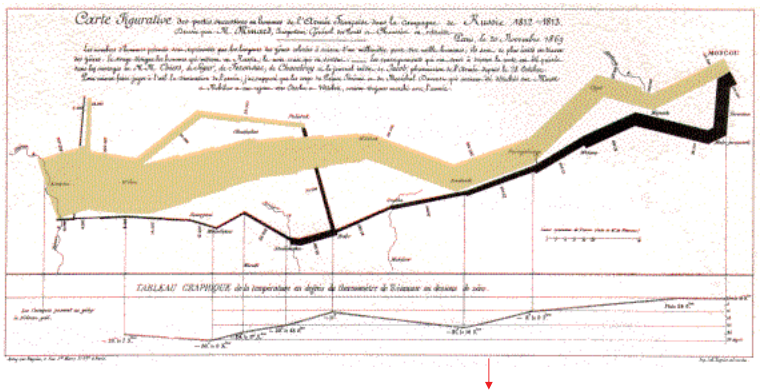
- *Album de Statistique Graphique, 1879-99*
- *Les Chevaliers des Albums*
- *Milestones* session, 5th Intl. Conf. Social Science Methodology (Cologne)

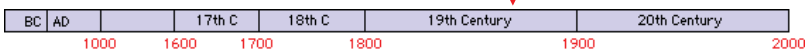




Background: C. J. Minard

- *"The best statistical graphic ever produced… defies the pen of the historian"*

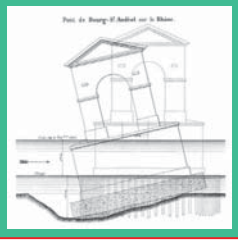


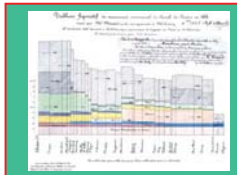


Why Minard?

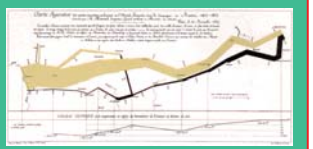
- Study breadth and depth of his work
  - How related to work in his time?
  - How related to modern statistical graphics?
  - How related to his personal history?

Civil Engineer for ENPC (1810-1842)



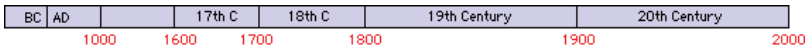


Visual Engineer for France (1843-1869)



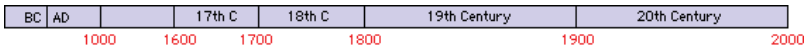
Background: C. J. Minard

- Bibliographic problem: Catalog & categorize his graphic work to make sense of patterns, trends, indications
  - The Graphic Works of Charles Joseph Minard (http://www.math.yorku.ca/SCS/Gallery/minbib.html)
  - All known graphic works, catalog entries (ENPC & BNF), keywords, cross-references
  - Online, searchable

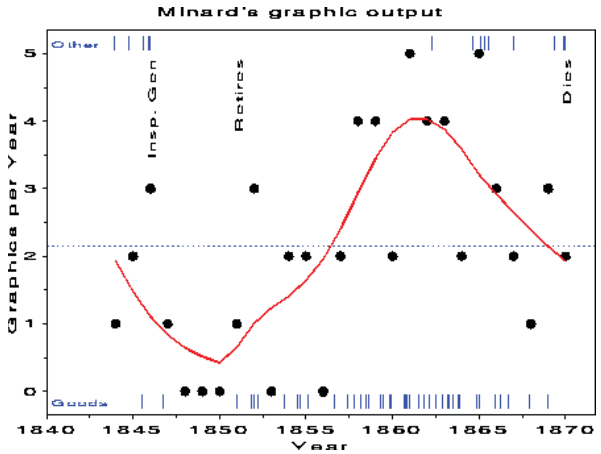


Meta Questions

- How to export advances in data visualization to an historical realm?
  - ❖How might a graphically-minded statistician look at history of data visualization?
  - ❖EDA → EBA (Exploratory bibliographic analysis)?
  - ❖What kinds of tools are needed?

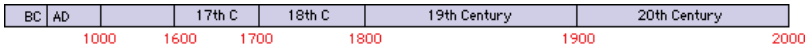


The Ebb and Flow of Minard's Graphic Output

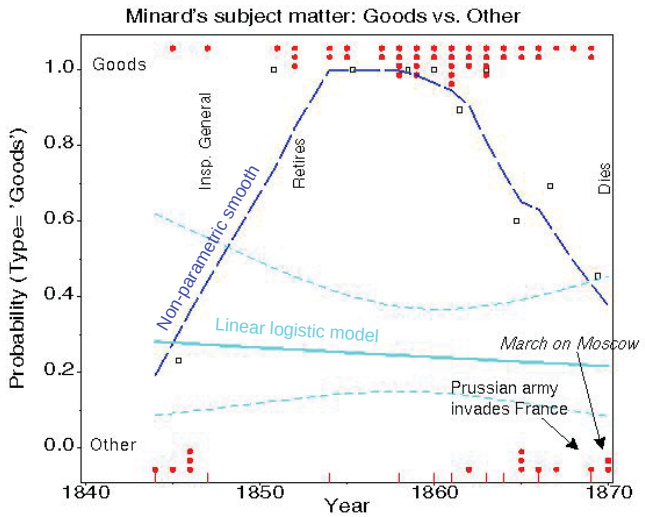


Graphical insight: smoothing helps the mind's eye

Life skills insight: retirement may not be a bad thing.

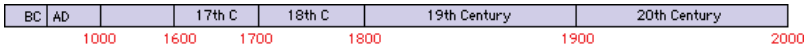


Minard's themes: Goods vs. Other

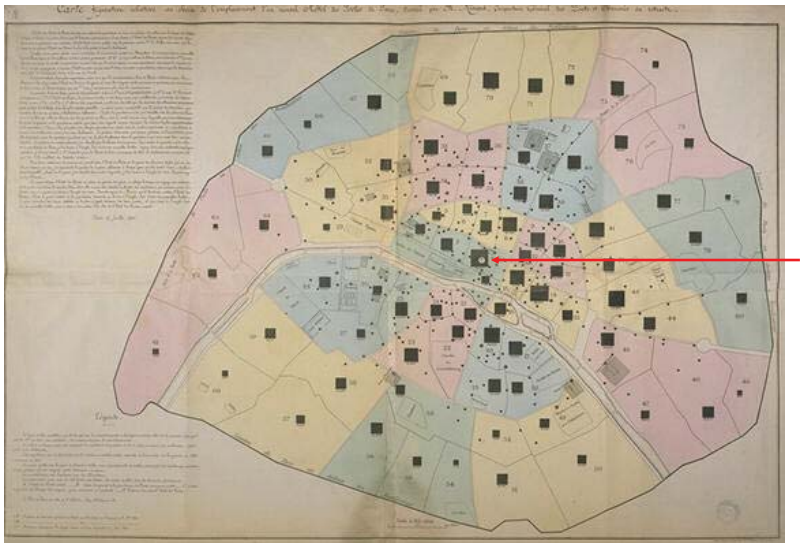


Graphical insight: discrete data are hard to show effectively

Statistical insight: models are often crude approximations



Where to build a new post office?(1867)

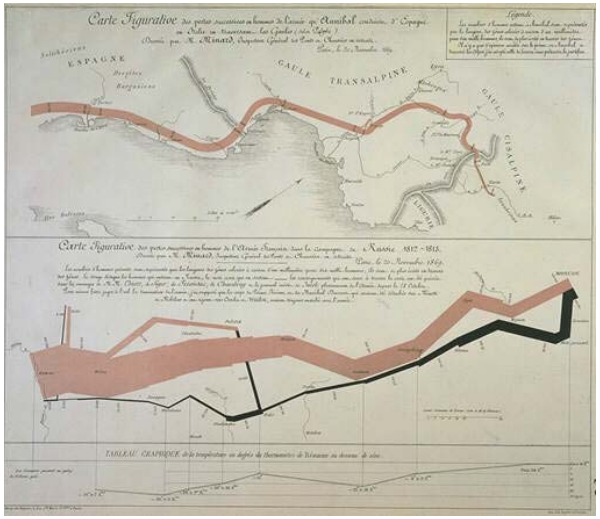


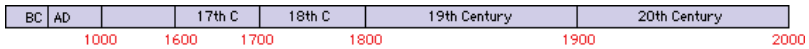
Center of gravity of pop. density

The March Re-Visited (1869)

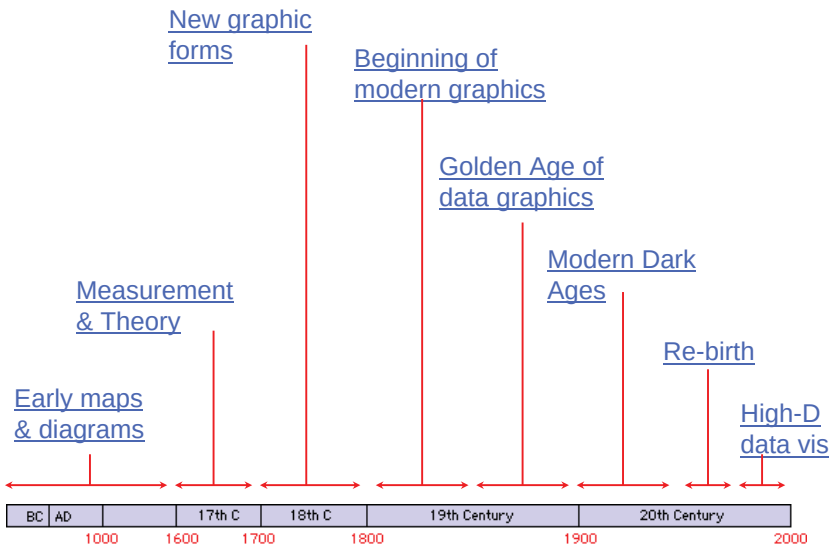
Hannibal's retreat

Napoleon's 1812 campaign



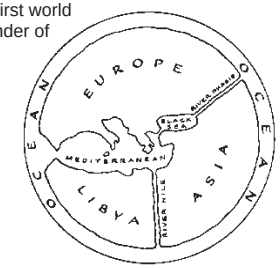


Milestones Tour



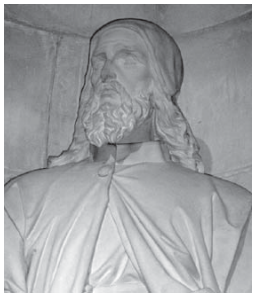
Pre 17th C.: Early maps & diagrams

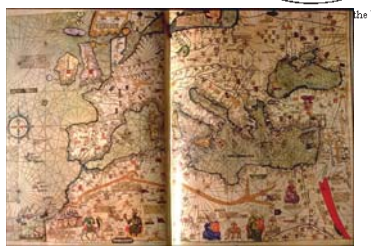
c. 550 BC: The first world map? (Anaximander of Miletus)



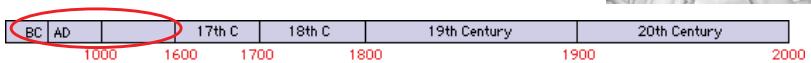
1350: Bar graph of theoretical function N. Oresme, France

c. 1280: Diagrams of paired comparisons for electoral systems- Ramon Llull, Spain

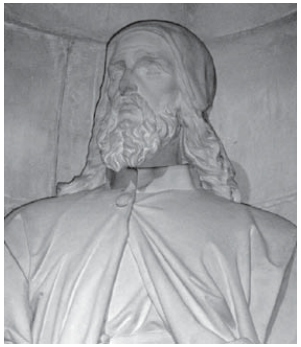




1375: Catalan Atlas, an exquisitely beautiful visual cosmography, perpetual calendar, and thematic representation of the known world- Abraham Cresques, Spain



c. 1280: Diagrams of paired comparisons for electoral systems- Ramon Llull, Spain

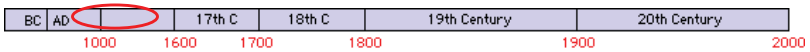


(Main hall, Univ. of Barcelona)

red = winner

bc	cd	de	ef	fg	gh	hi
bd	ce	df	eg	fh	gi	
be	cf	dg	eh	fi		
bf	cg	dh	ei			
bg	ch	di				
bh	ci					
bi						

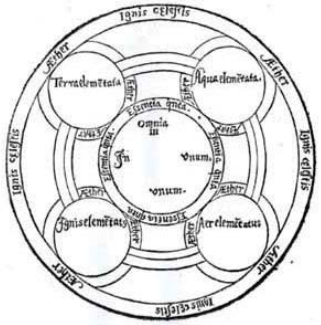
candidate	f	c	h	i	d	b	e
wins	7	6	5	4	3	2	0

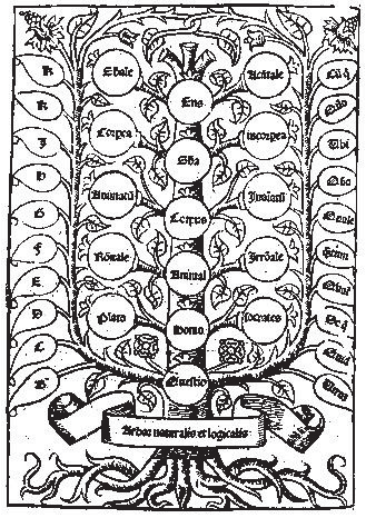


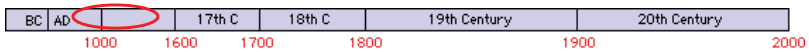
1305: Mechanical diagram of knowledge- Ramon Llull, Spain

Tree of porphyry: Aristotle's categories of knowledge (center)

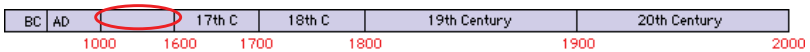
- Left: questions
- Right: rotating disks → answers





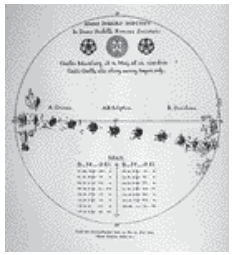


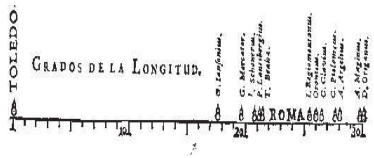
1375: Catalan Atlas, an exquisitely beautiful visual cosmography, perpetual calendar, and thematic representation of the known world- Abraham Cresques, Majorca, Spain [BNF: ESP 30]



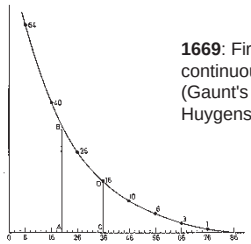
1600-1699: Measurement and Theory

1626: Visual representations used to chart the changes in sunspots over time- Christopher Scheiner

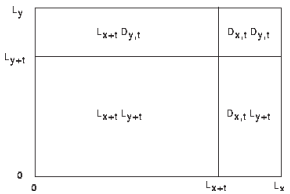




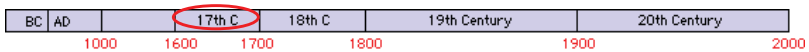
1644: First visual representation of statistical data- M.F. van Langren, Spain



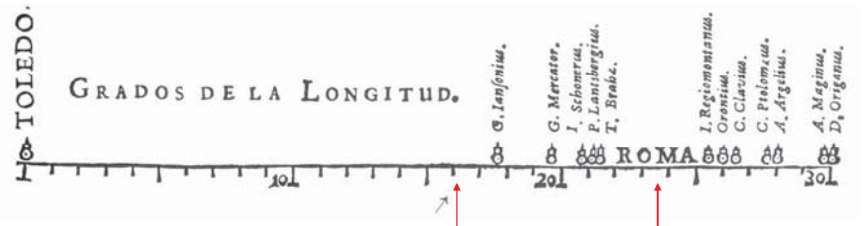
1669: First graph of a continuous distribution function (Gaunt's life table)– Christiaan Huygens.

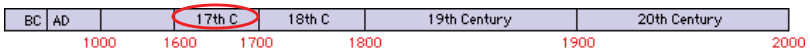


1693: First use of areas of rectangles to display probabilities of independent binary events- Edmund Halley, England



1644: First visual representation of statistical data: determination of longitude between Toledo and Rome- M. F. van Langren, Spain

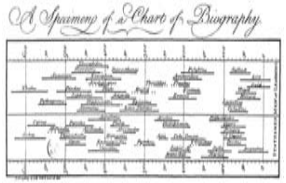




1700-1799: New graphic forms

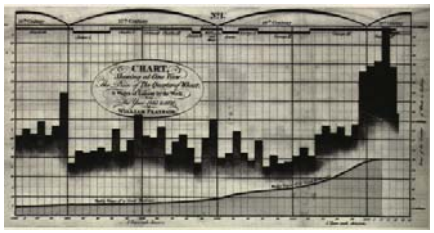
1701: Isobar map, lines of equal magnetic declination – Edmund Halley

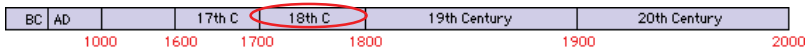
1765: Historical time line (life spans of famous people) Joseph Priestley



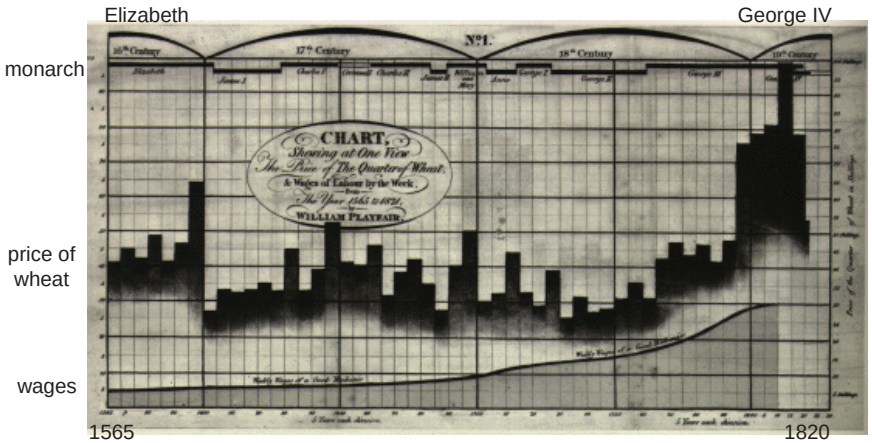
1782: First topographical map- Marcellin du Carla-Boniface

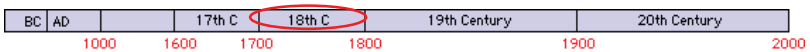
1786: Bar chart, line graphs of economic data- William Playfair





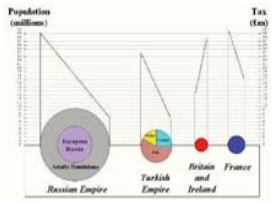
1786: Bar chart and line graph showing three time series: Price of wheat, weekly wages and reigning monarch over a 250+ year span- William Playfair



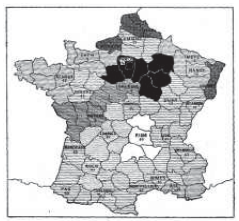


1800-1849: Beginning of modern data graphics

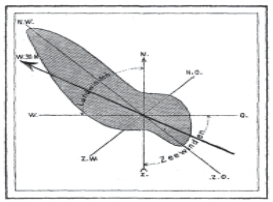
1801: Pie chart, circle graph invented- William Playfair

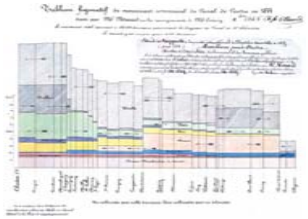


1819: First modern statistical map (illiteracy in France)- Charles Dupin

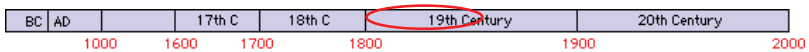


1843: Wind-rose (polar coordinates)- L. Lalanne

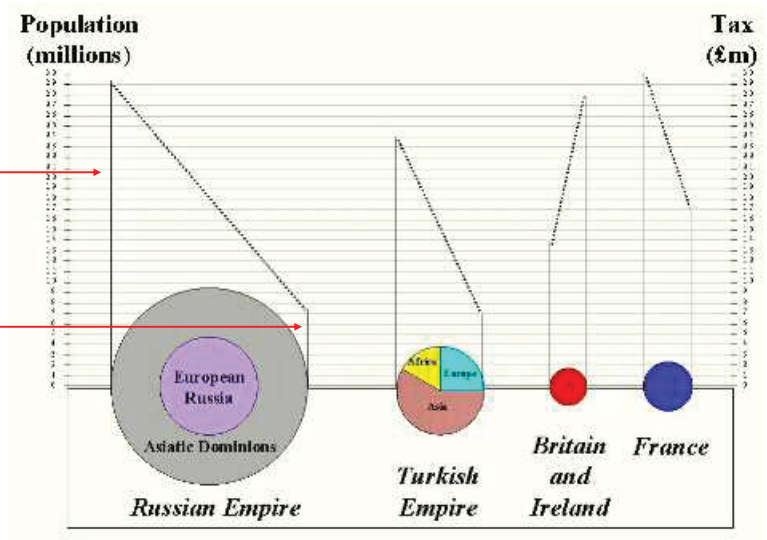




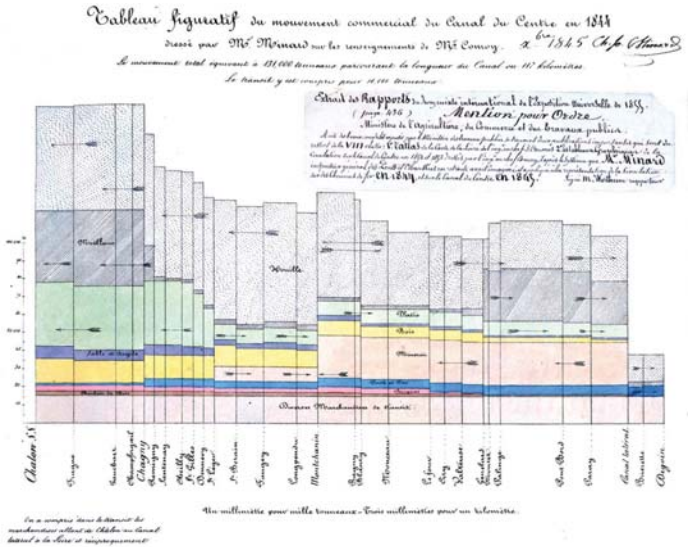
1844: variable-width, divided bars, area ~ cost of transport- C. J. Minard



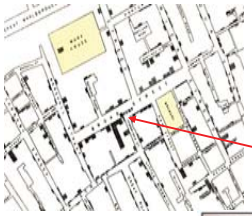
1801: Pie chart, circle graph invented- William Playfair



1844: *Tableau-graphique*: variable-width, divided bars, area ~ cost of transport- Charles Joseph Minard

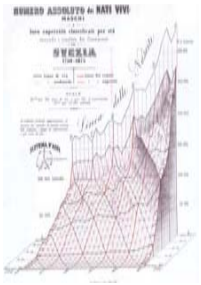


1850-1899: Golden Age



1855: Dot map of disease data (cholera)- John Snow

Broad St. pump

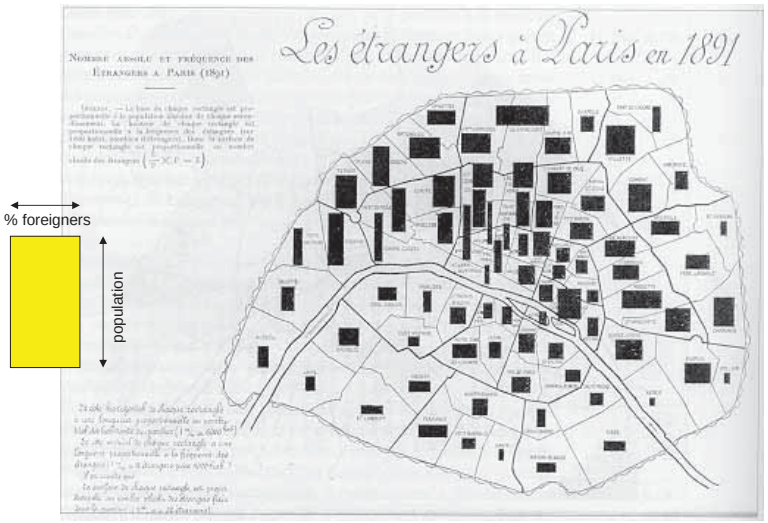


1879: Stereogram (3D population pyramid)- Luigi Perozzo

1884: Recursive multi-mosaic on a map- Emile Cheysson

1896: Area rectangles on a map to display two variables and their product- Jacques Bertillon

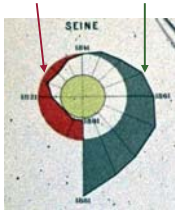
1896: Area rectangles on a map to display two variables and their product- Jacques Bertillon

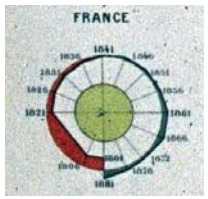


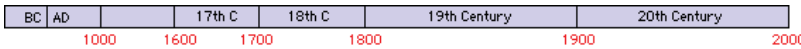
Album de Statistique Graphique, 1879-1899

1884: Changes in French population by department, 1801-1881

loss gain

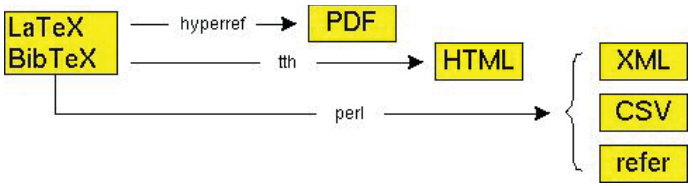




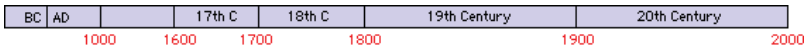


Under the Cover

- One set of sources to produce various hyper-linked versions

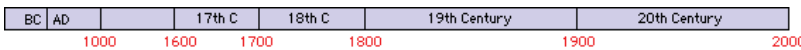


- Items: 216
- Images: 222
- Text links: 120
- References: 243



Problems of statistical historiography

- What counts as a milestone?
- Who gets credit? – *Stigler's Law of Eponomy*
- What is "data"
- How to display, visualize, search?



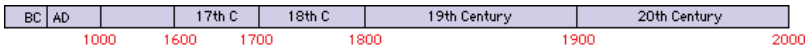
What counts as a "milestone"?

Innovations and/or developments in:

- **Graphic forms**:
  - *Statistical graphics*: bar chart, line plot, scatterplot, boxplot, mosaic plot
  - *Cartography*: isoline, choropleth
- **Graphic content**: data collection,recording
  - Bills of Mortality, vital statistics, census
  - Measurement, recording devices

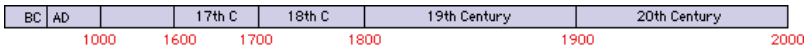
What counts as a “milestone”?

- **Technology and enablement**
  - *Reproduction*: printing press, lithography
  - *Imaging*: photography, motion picture
  - *Rendering*: computing, video display
- **Theory and practice**
  - Probability theory
  - Summarization: estimation & modeling
  - Exposure: EDA
  - Awareness & use



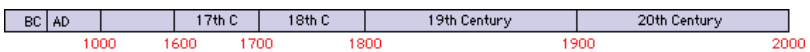
What counts as a “milestone”?

- **Theory of, and data on visual display**
  - Principles of graphics (Bertin, Tufte, …)
  - Empirical studies- what works?
- **Implementation/dissemination**
  - Techniques *available & accessible*
  - Printing, publication, web
  - Software



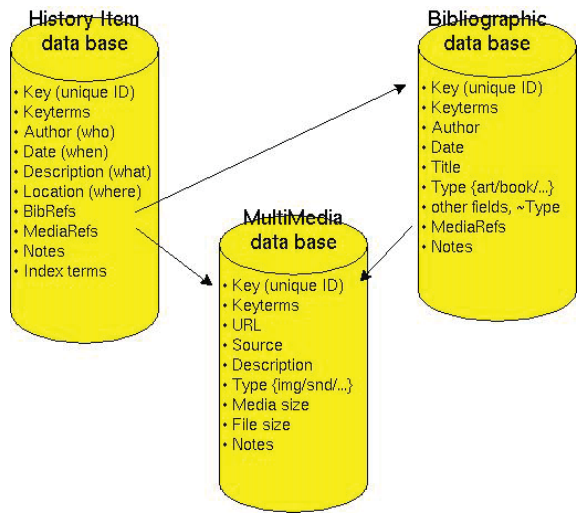
What is milestone “data”?

- **History-item data base** <hdbitem>
  - When, who, what, where, notes, …
- **Bibliographic data base** <bibitem>
  - Book, article, in-collection, map, misc, …
- **Multi-media database** <mediaitem>
  - Image, portrait, web link, audio, movie, …

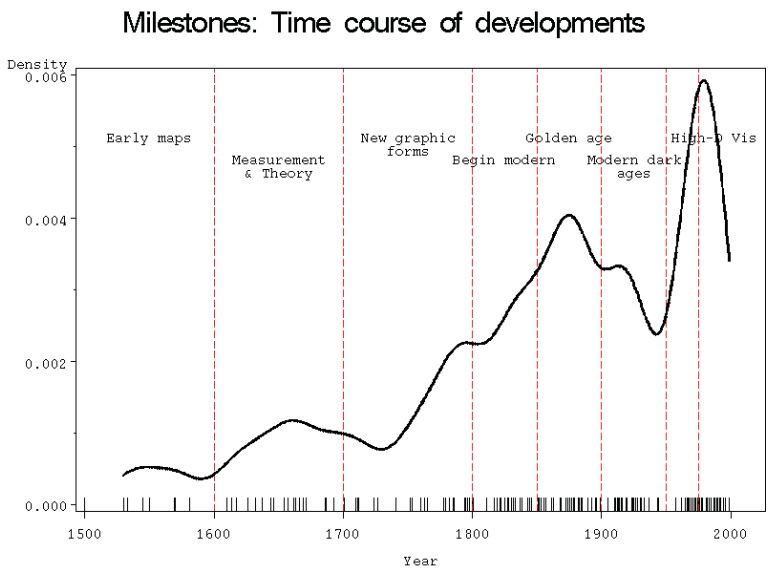


What is milestone “data”?

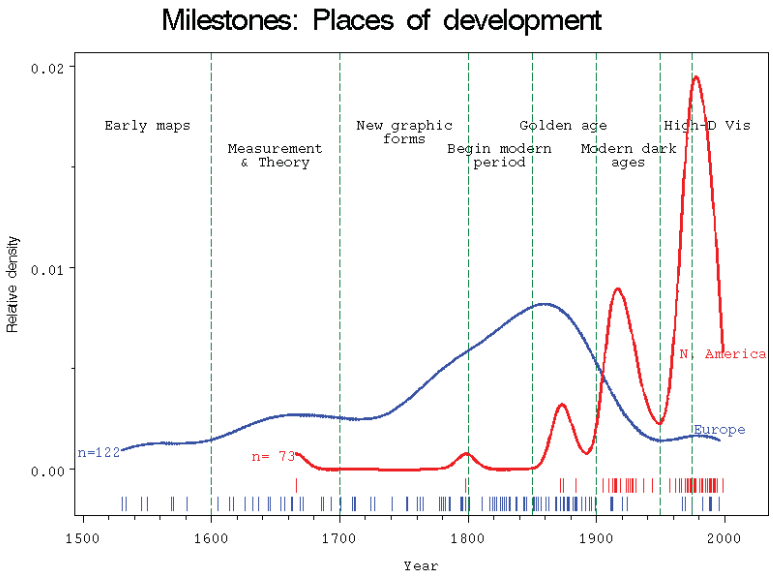
Milestones data as a relational data base



Analyzing milestones data?

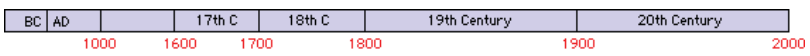


Analyzing milestones data?



How to visualize a history?

- Timeline: obvious, but:
  - 8000+ years, but most in last 300-400
  - Problems of display, resolution, access
  - Linear: no representation of content
- Lessons from the past?
  - Dubourg's Scroll of History
  - Priestly's Chart of Biography
  - Marey's life spans of British monarchs



Jacques Barbeu-Dubourg's Scroll of History

- 54' scroll, spanning 6,480 years (Creation→1753)
- Grouped vertically by theme or country
- Symbols for character & profession
- History=Geography + Chronology

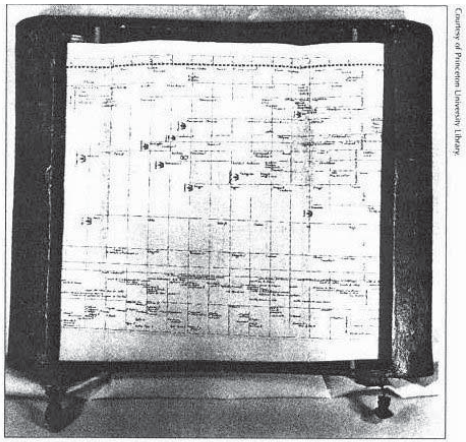
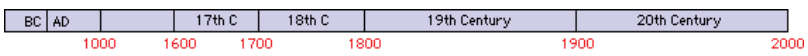
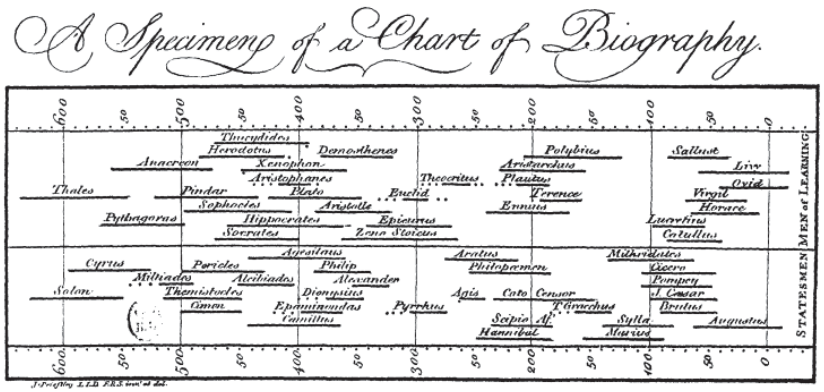


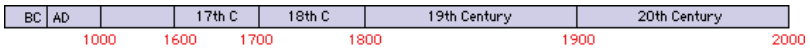
Figure 2. Photo of Dubourg's scroll opened to the years A.D. 360 to A.D. 510.



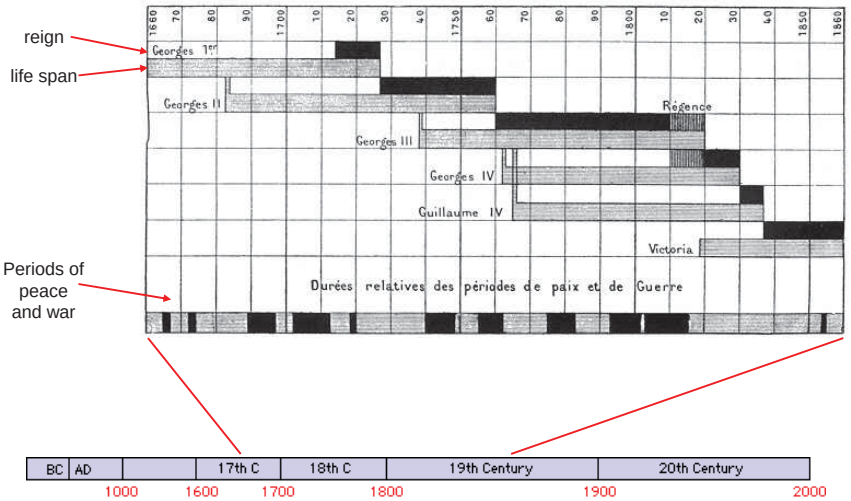
Priestly's Chart of Biography

- Life spans of famous people, 1200 BC to 1750



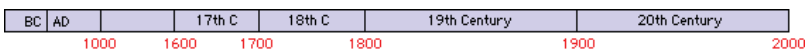


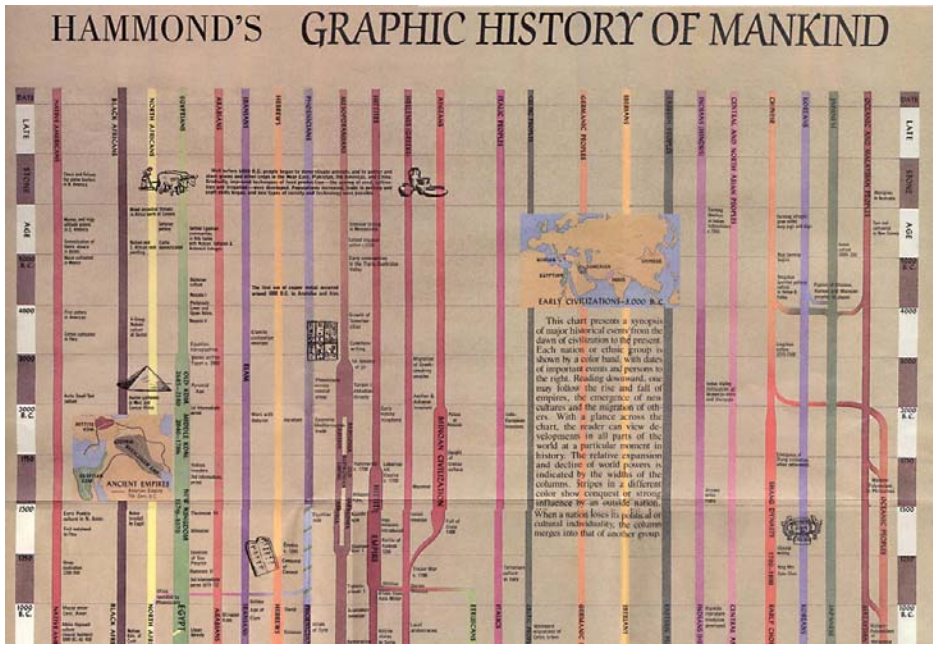
Marey's life spans of British monarchs (1878)

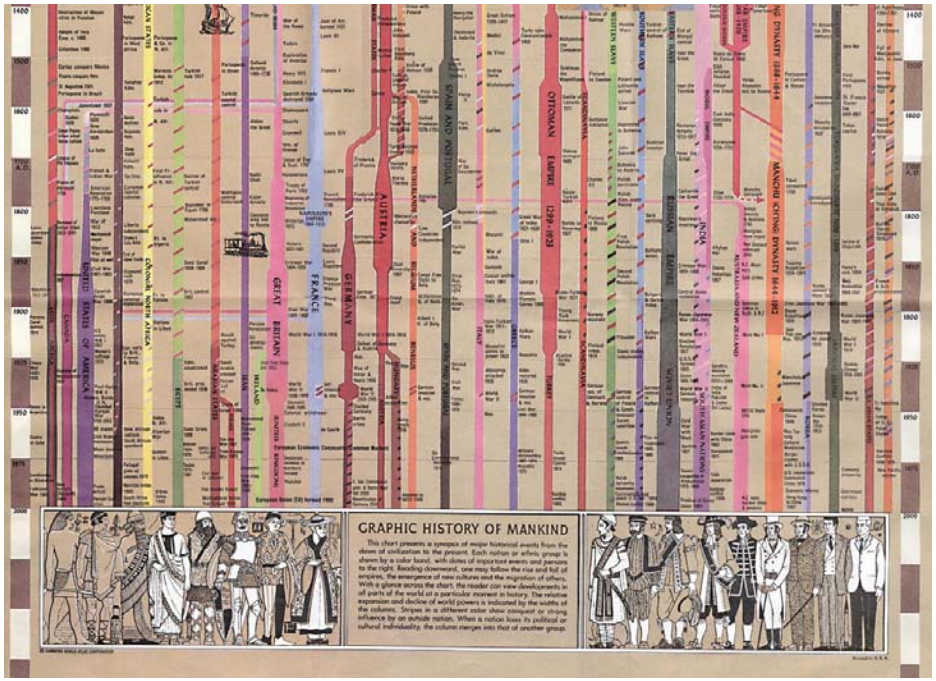


Lessons from the present

- Hammond's Graphic History of Mankind
  - Varying-resolution time scale
  - Separate time lines for nations/ethnic groups
    - Rise and fall of empires
    - Emergence of new cultures
  - Influence indicated by width of lines
  - Shading/stripes show conquest or outside influence

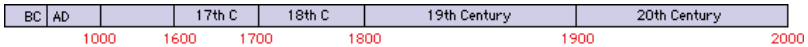






Lessons from the web

- Digital image libraries
  - AP Photo Archive: http://archivepix.ap.org/
  - David Rumsey Map Collection: http://www.davidrumsey.com/
  - American Memory: http://memory.loc.gov
- Provide:
  - Search
  - Zoom
  - Image data



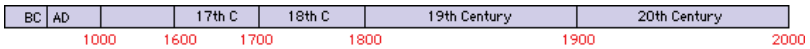
Searchable image catalogs

AP Photo Archive: http://archivepix.ap.org/

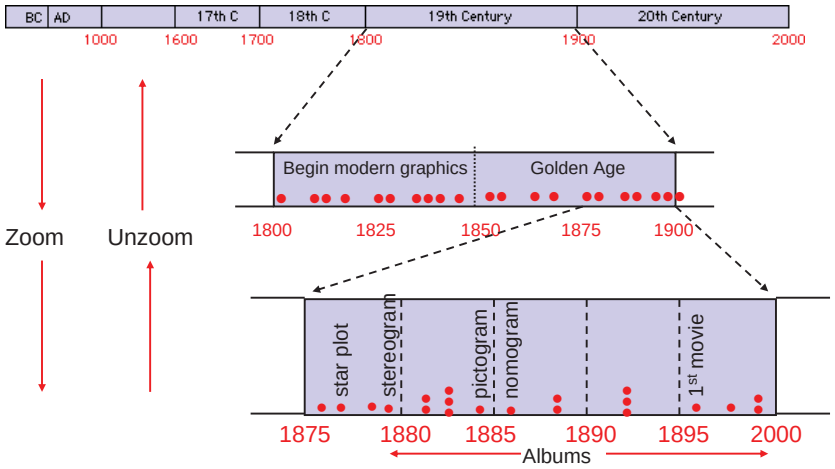
Search by: what, when, where, …

Lessons from data visualization

- Zoom, focus & resolution
  - Non-linear scales for space & time
  - Table lens
- Network representations
- Tree representations



Zoom, Focus & Resolution



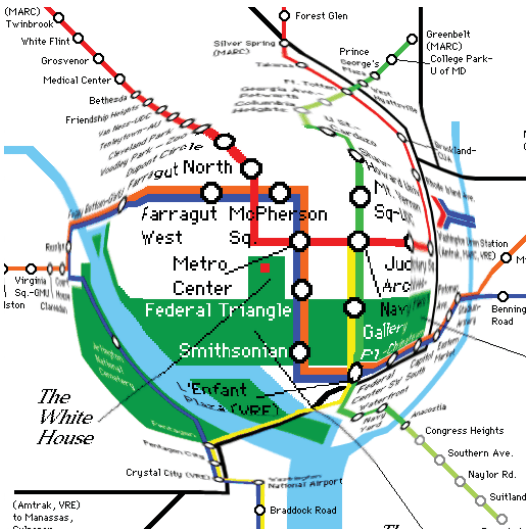
Non-linear Scales for Space

Hand with Sphere,
M. E. Esher

Non-linear Scales for Space

Fisheye view of central Washington D.C.

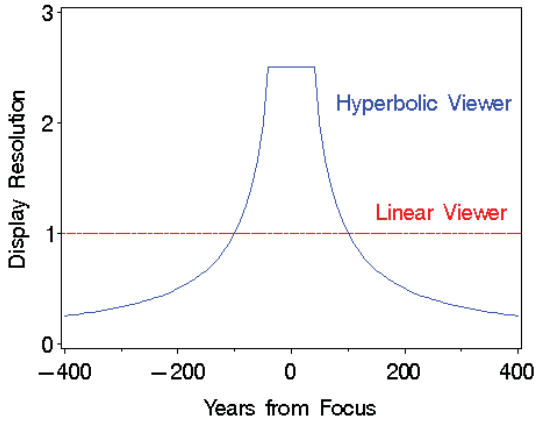
Dynamic: move the cursor to change the focal point



Non-linear Scales for Time

A hyperbolic viewer varies resolution smoothly, trading off against span of the view:

*Resolution * Span = constant*



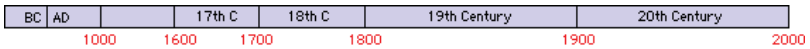
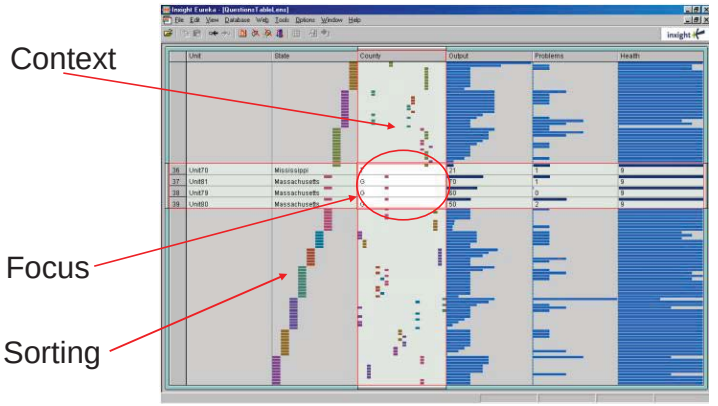
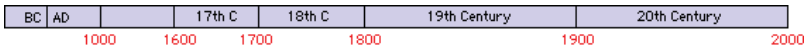


Table Lens

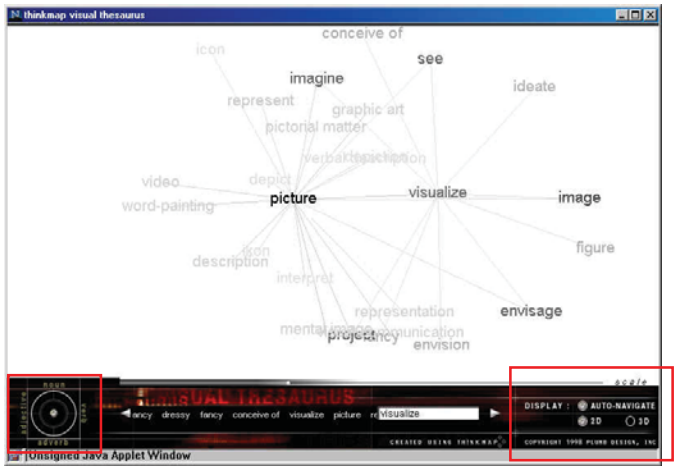
- Hyperbolic viewer: increased resolution at focus
- http://www.tablelens.com





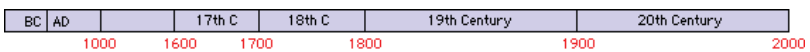
Semantic Network

Visual Thesaurus - http://www.visualthesaurus.com/



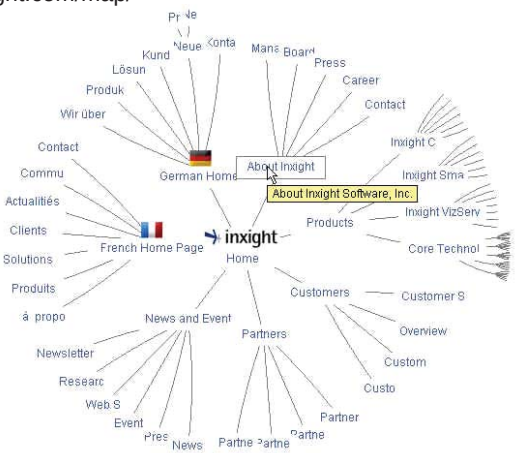
Part of speech controls

Zoom, animate controls

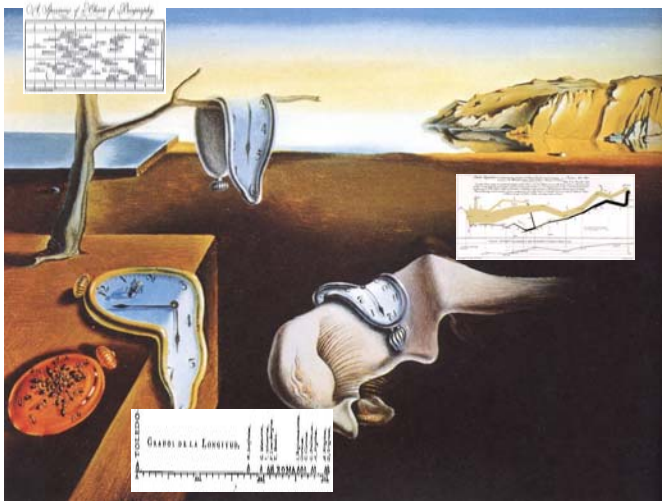


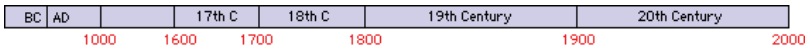
Star Tree

http://www.inxight.com/map/



Dali TimeScape

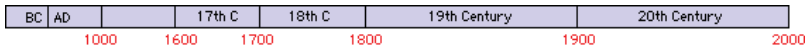




Conclusions

The only new thing… is the history you don't know – Harry Truman

- Modern data visualization has deep roots:
  - Cartography
  - Statistics
  - Data collection
  - Visual thinking
  - Technology
- *Milestones Project* attempts to document them all comprehensively.



Conclusions

- This leads to interesting problems in (statistical) historiography:
  - What counts as a Milestone?
  - How to organize and represent historical data?
  - What tools are needed?
  - How to visualize a history?