





The Golden Age of Statistical **Graphics**



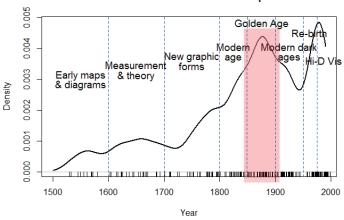
Michael Friendly https://friendly.github.io/6135



The Golden Age: ~ 1850 -- 1900

Why do I call this the "Golden Age of Statistical Graphics"? The most obvious is as a peak in developments over the course of history.

Milestones: Time course of development



What makes an "Age"? What makes one "Golden"?

- Age:
 - Qualitatively distinct from before & after
- Golden age:
 - Recognizable period in a field where great tasks were accomplished
 - Years following some innovations
 - Artists apply skills to new areas
 - New ideas expressed, art forms flourish
 - Often ends with some turning point event(s)

Some Golden Ages

- Athens (Pericles): 448 BC— 404 BC: growth & culture
- Islam: 750—1258 (sack of Baghdad): science, math ...
- England: Elizabeth I (1558-1603): literature, poetry, ...

Piracy: 1690--1730
Radio: 1920—1940
Animation: 1928 (sound) –
1960s (TV)

• Senior citizens: 60+



Pietro Da Cortona, The Golden Age (Fresco, Sala della Stufa, Palazzo Pitti, Florence)

Preludes to the Golden Age

Infrastructure required:

- Data: collection & dissemination
- Statistical theory: combining & summarizing quantitative information
- Technology: printing & reproduction of maps & diagrams
- Visual language: new graphic forms for maps and diagrams
- → a perfect storm for data graphics

What does this imply for today?

Preludes: data

- Population: ~ 1660--
 - Bills of mortality: Graunt (1662)
 - Political arithmetic: Petty (1665)
 - Demography: Süssmilch (1741)
- Economic data: ~ 1770--
 - Revenue, expenditures, taxes
 - Imports, exports
 - Transport
- Social data: ~ 1820--
 - Literacy, education
 - Crime, suicides, illegitimate births, prostitution
 - Poverty, debtors, disease
- → An avalanche of data, waiting to be visualized!

"Data! Data! I can't make bricks without clay." -Sherlock Holmes. Copper Beeches

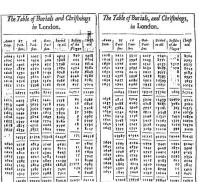




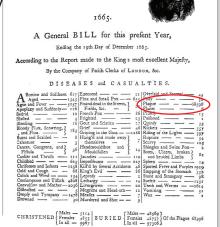


deaths, causes of mortality

Population data: keeping track of births,



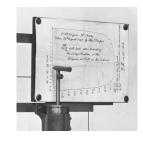
J. Graunt (1662) Natural and Political Observations on the Bills of Mortality



Preludes: technology

- Copperplate \rightarrow Lithography (1800+) \rightarrow color printing (1850+)
- Automatic recording: James Watt (1822)
- Calculation: Babbage (1822/33), Guerry ~1850
- Photography: Niépce (1827), Deguerre (1839), trichromatic process (1861)
- Motion: Muybridge (1872), Marey (1882)



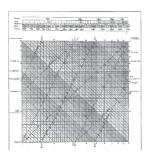


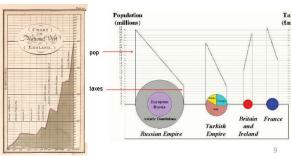




Preludes: visual language

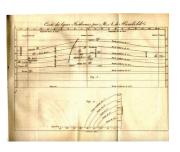
- Graphs & diagrams
 - Line, bar, pie charts-Playfair (1786, 1801)
 - Scatterplot- Herschel (1832)
 - Polar plots- Guerry (1829), Nightingale (1857)
 - Nomograms & graphical calculation
 – Lalanne (1846)

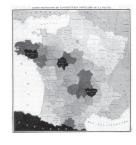




Preludes: visual language

- Thematic maps
 - Isopleth– Humboldt (1817)
 - Choropleth Dupin (1826)
 - Dot-Frère de Montizon (1830)
 - Flow– Harness (1837)







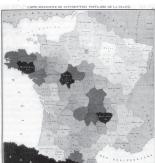


10

Data visualization: Diffusion of ideas

Those who developed thematic maps often not cartographers

Dupin (1826): literacy in France

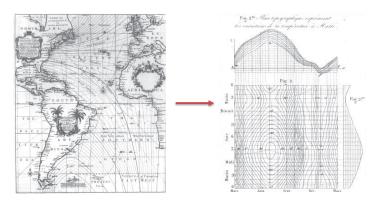


Galton (1881): travel time from London



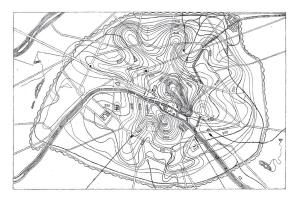
Data visualization: Diffusion of ideas

- Those who developed data graphics often borrowed from cartography
 - Halley (1701): contour map -> Lalanne (1843): contour diagrams of soil temperature



Data visualization: Diffusion of ideas

- ... and vice-versa
 - Lalanne → L.L. Vauthier (1874) contour map of population density of Paris, seen as mountains
 - Map-based data visualization was extended widely



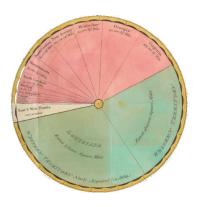
14

Data visualization: Diffusion of ideas

- Graphical inventions often applied to maps
 - Playfair (1805): pie chart -> Minard (1858): pie map

What are the sizes of US territories?

Where does meat sold in Paris come from?





15

Stories from the Golden Age (1850-1900)

Stories:

- A.-M. Guerry & the rise of social science
- Graphic vision of C. J. Minard
- Galton's graphical discoveries
- Statistical albums

Themes:

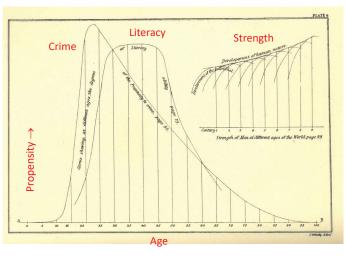
- Statistics: numbers of the state
- Rise of visual thinking
- Escaping flatland: 2D → 3D
- Visualization → Theory (graphic discovery)
- Data \rightarrow Theory \rightarrow Practice
- Graphical excellence

Big questions of the early 1800s

- Issues for European states
 - Demography: taxes, raising an army (Süssmilch, 1741)
 - "Statistik": Numbers of the state (Achenwall, 1748)
 - Social problems: crime, suicide, literacy, etc.
 - Disease epidemics, e.g., cholera
- Anthropometry: the measure of Man
 - Distributions of human characteristics (Quetelet)
 - Mortality, suicide, propensity to crime
- Beginnings of statistical theory and application
 - Normal distⁿ (de Moivre, 1733)
 - L'homme moyen (Quetelet, 1835)

Quetelet: Anthropometry

Quetelet (1842), A Treatise on Man and the Development of His Faculties, uses graphs to illustrate various themes: measurement, graphical comparison, ...





Human characteristics could be measured and studied graphically

18

Big data of the early 1800s:

"An avalanche of social numbers"

- J.-B.J. Fourier: *Recherches statistique sur la ville de Paris* (1821-1829)
 - Massive tabulations: births, deaths (by cause), admission to insane asylums (age, sex, affliction)
- Ministry of Justice: Compte generale (1825--)
 - First national compilation of criminal justice data
 - All charges & dispositions, quarterly, 86 departments
- Other sources:
 - Bureau de Longitudes (illegitimate births)
 - Parent-Duchatelet (prostitution); Min. of War (desertions)
 - Suicide notes in Paris collected and analyzed for motives
- Social issues could now be addressed with DATA

9

22

1. A. M. Guerry and the rise of social science

Essai sur la statistique moral de la France

The launching pad of modern social science

- Presented to Academie des Sciences Français July 2, 1832
- First systematic analysis of comprehensive data on crime, suicide, and other social variables.
- ► Along with Quetelet (1831, 1835), established the study of "moral statistics" → modern social science, criminology, sociology



Social context of crime in 1820s France

- Crime a serious concern:
 - Explosive growth in Paris after Napoleon's defeat (Waterloo, 1815)
 - Widespread unemployment,
 - Emergence of perception of "dangerous classes": what to do???
 - Victor Hugo (Les Misérables); Honoré de Balzac; Emile Zola
- Liberal ("philanthrope") view:
 - Increase education
 - Better prison conditions, diet (bread and soup)
 - Religious instruction
- Conservative view:
 - Build more prisons; longer prison sentences
 - Harsher treatment of recidivists
- Now, there was finally some DATA!

The discovery of "social facts"

Stability and Variation

Guerry's results were both compelling and startling:

- Rates of crime and suicide remained remarkably invariant over time, yet varied sytematically by region, sex of accused, type of crime, etc.
- In any given French city or department, almost the same number committed suicide, stole, gave birth out of wedlock, etc.

Year	1826	1827	1828	1829	1830	Avg
Sex	All accused (%)					
Male	OMG! \sim constant \longrightarrow 79	79	78	77	78	78
Female	21	21	22	23	22	22
Age	Accused of Theft (%)					
16-25	37	35	38	37	37	37
25–25	OMG! ~ constant → 31	32	30	31	32	31
Crime	Committed in summer (%)					
Indecent assault .		36	36	35	38	36
Assault & battery .		28	27	27	27	28

"We are forced to conclude that the facts of the moral order are subject, like those of the physical order to invariable laws." (Guerry, 1833, p14)

The discovery of "social facts"

Social laws á la physical laws

Do crime and other moral variables represent:

- structural, lawful characteristics of society, or are they
- simply indicants of individual behaviour?

Guerry argued:

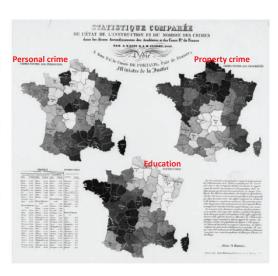
Each year sees the same number of crimes of the same degree reproduced in the same regions. (Guerry, 1833, p.10)

... We are forced to recognize that the facts of the moral order are subject, like those of the physical order, to invariable laws (Guerry, 1833, p14)

24

1829: Statistique comparée de l'état de l'instruction...

- First shaded thematic maps of crime data
- First comparative maps of social data
- ► Instruction: → France obscure and France éclairée (Dupin, 1826)
- North of France highest in education, but also in property crime!



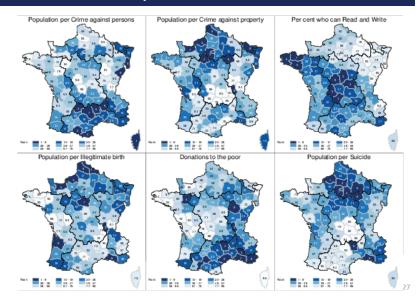
1833: Essai sur la statistique morale de la France

- Divided the 86 departments into 5 regions
- Supplemented data from the Compte général with:
 - Suicides in Paris, 1794–1832
 - Prostitutes in Paris (Parent-Duchâtelet)
 - Wealth (taxes per inhabitant)
 - Distribution of clergy
 - ..
- First study to use crime data to 'test' hypotheses
- Attracted widespread interest in Europe



Guerry's 1833 map of literacy in France

Guerry's moral variables



1864: Statistique morale de l'Angleterre comparée...

- Proposes to replace simple "moral statistics" (tables) with "analytical statistics"
 - ► calculation, graphic display
 - ▶ → general, abstract results
- ▶ 17 large color plates (56 × 39 cm):
 - data for France (1825–1855), England (1834–1855)
 - crimes against persons and property decomposed in various ways
 - first attempt to delineate multivariate relations among moral variables
- ▶ Voluminous data:
 - ▶ 85,564 suicide records (1836–1860), classified by motive
 - 226,224 accused of personal crime
 - ▶ numbers, in a line → 1170 meters!

STATISTIQUE MORALE
L'ANGLETERRE

I A STATISTICE P MODELE

LA FRANCE

media del campio pe glaricomissione en la agranta distintació de amenicami en es secreta de conflicto de la mella de producto di glaricomi, el deplicació, en por media-fillence el ex o el confessió difensi de mella officio del conflicto de confessione de confessione de confessione.

PAR A.-N. GUERRY

to an other bases of the first

ATLAS

CARD I COMPOSITO SECURITY SECURI

PROSPECTUS

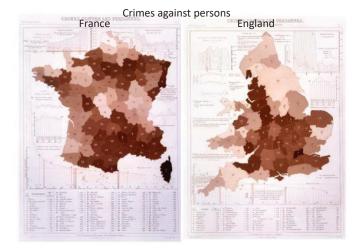
PARIS L.- BAILLIERE ET FILS.

HEW-YORK, St. Micro-State. States or regions. Id. Michigan. MADRO, C. 12525-041004, state on the control of the

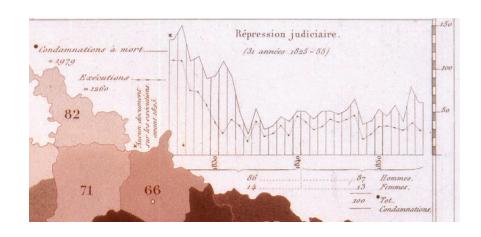
Further details: Friendly, M. (2007). A.-M. Guerry's *Moral Statistics of France*: Challenges for Multivariable Spatial Analysis, *Statistical Science*, 22, 368-399

29

1864: Statistique morale de l'Angleterre comparée... Comparing France and England



Graphs and tables around the outside give details: data, trends over time, or season, ...



Detail: Trends in death sentences and executions over 31 years

Statistique analytique: General causes of crime

Plate XVII: M. Guerry's magnum opus

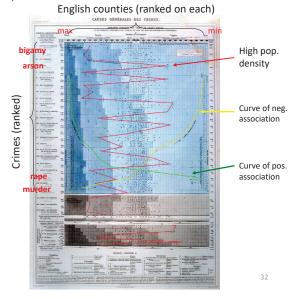
Goal:

- Show multivariate factors associated with distribution of crime
- Before invention of correlation

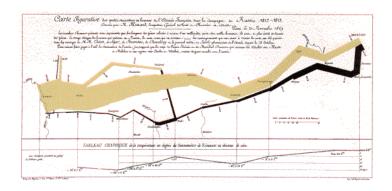
Entries: Codes for factors

- Pop: (% Irish, domestics, ...)
- Criminality: (male, young, ...)
- Religion (Anglicans, dissenters, ...)





2. The graphic vision of C. J. Minard



- Marey (1878): "defies the pen of the historian in its brutal eloquence"
- Tufte (1983): "the best statistical graphic ever produced"

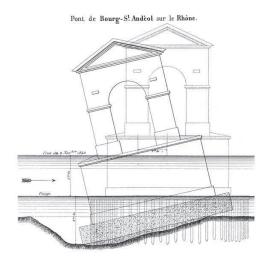
33

Visual thinking, visual explanation

Minard's main career was a as a civil engineer for the ENPC (bridges & roads)

1840: Why did the bridge at Bourg-St. Andèol collapse?

Minard's report consisted essentially of this self-explaining diagram.



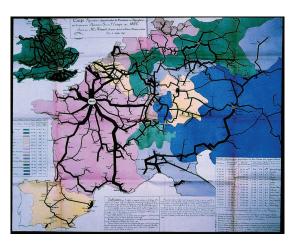
Big questions of the mid 1800s

- 1830—1860: emergence of modern French state, dawn of globalization
- Trade, commerce, transportation:
 - Where to build railroads, canals?
 - How to compete with imports/exports?
 - Visualizing changes over time, differences over space
 - → Flow maps and other graphical innovations
- These questions motivated the "Golden Age" of statistical graphics.
 - data, statistics, technology & visual thinking

See: Friendly, M. (2008). The Golden Age of Statistical Graphics, Statistical Science, 23, 502-535, https://www.datavis.ca/papers/golden-age.pdf

Flow maps as visual tools

Transport of passengers on the principal railroads in Europe in 1862



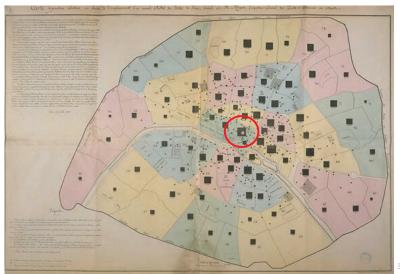
The dominant principle which characterizes my graphic tables and my figurative maps is to make immediately appreciable to the eye, as much as possible, the proportions of numeric results.

...Not only do my maps speak, but even more, they count, they calculate by the eye.

-- Minard (1862)

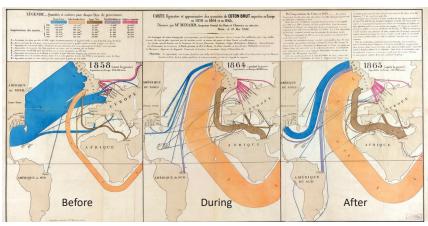
Q: Where to build a new post office in Paris?

• Visual solution: at the center of gravity of population



Visual explanation

What was the effect of the US Civil War on trade in cotton?

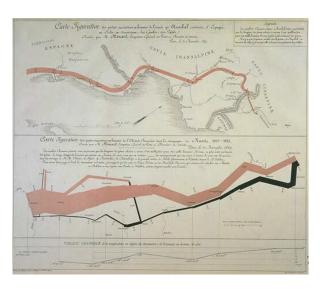


"Carte figurative et approximative des quantités de coton brut importées en Europe en 1858, en 1864 et en 1865" by Charles Joseph Minard (1866)

The March Re-Visited (1869)

Hannibal's retreat

Napoleon's 1812 campaign



Les Chevaliers: Minard's Tomb

Jean-Pierre Airey-Jouglard

Antoine de Falguerroles

Recent discovery of Minard's tomb in Montparnasse Cemetery, Paris.

Celebrated June 5, 2017

CHARLES JOSEPH MINARD
Ingicieur geferril des Prots et Chametes,
Promier de la méthode graphique

Celebre par ses amis,

Les Chevaliers des Alliums de Sazinistique Graphique,
le 5 juin 2017

MF Gilles Palsky

40

3. Galton's discovery of weather patterns-Perhaps the most notable *purely graphic* discovery ever!

METEOROGRAPHICA,

METHODS OF MAPPING THE WEATHER;

ILLUSTRATED BY UPWARDS OF 600 PRINTED AND LITHOGRAPHED DIAGRAMS

THE WEATHER OF A LARGE PART OF EUROPE,

During the Month of December 1861.

By FRANCIS GALTON, F.R.S.

(Galton, 1863)

42

Method: All weather stations across Europe asked to record data 3x/day for all of Dec., 1861

Galton's data collection form:

Contributors, according to the Conditions of my Circular Letter, are requested to enter their Observations in one of the blank forms, to enclose it in a stamped envelope, and to post it to my address on January 1st, 1862. FRANCIS GALTON, 42, Rutland Gate, London. Name of Station : Its Latitude : Full Address to which the Charts are to be forwarded when ready: Its Longitude from Greenwich: Its Height above Sea Level, in English Feet REMARKS December 1861. 9 л.м. 3 P.M. 9 P.M. 9 л.м. 3 P.M. 9 P.M. 9 A.M. 3 P.M.

Method: All weather stations across Europe asked to record data 3x/day for all of Dec., 1861

Data: recordings of barometric pressure, wind dir/speed, rain, temp., cloud: 3x/day, 50 weather stations in Europe.

Graphic analysis: 3x31=93 maps, each with multivariate glyphs showing all variables

Visual ideas:

- Iconic symbols
- Multivariate glyphs (stamps!)



EXPLANATION OF THE SYMBOLS USED IN THE WEATHER CHARTS.

(LOUD)

Rain. Snow. Entirely and Entirely clouded, clouded, clouded, clouded, clouded, clouded, clouded, clouded.

DIRECTION OF WIND.

S SSW SW WSW W. Code. Street. Moderner. Gentle. Almost calm. Calm.

Visual abstraction \rightarrow Patterns

How to see patterns of geographical variation over time?

- Iconic symbols on a geographical grid
- "Small multiples:" separate graphs laid out for direct comparison



Symbols in Barometrical Charts.

Black - 29.95 to 29.71

29.70 to 29.46 ⊙ 30.21 to 30.45 Inches. Inches. 29.45 to 29.21 29.20 and below. 30.46 to 30.70 30.71 and above.

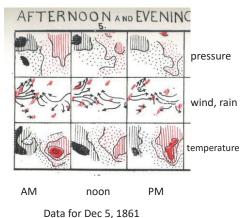
45

Visual abstraction → Patterns

What varies with what, over time and space?

- mini, abstract maps: vars x TOD
- iso-contours, shading to show equivalence
- arrows to show wind direction





46

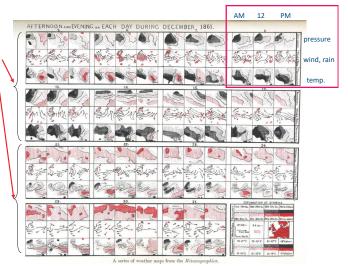
The large picture → Insight

Pattern:

Low pressure (black) in early Dec. → CCW wind High pressure (red) in late Dec. → CW wind

Graphic: 3x3x31 grid, mapping {pressure, wind/ rain, temperature} x {AM, 12, PM} x day {1:31}

(try this with your software!)



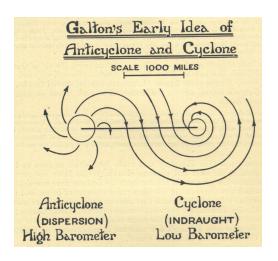
Visual insight → Theory

Visual insight from 93 (3x31) high-D graphs:

- Changes in wind dir w/ pressure over time
- → Winds revolve inwardly (CCW) in low pressure areas— as in a cyclone;
- $\bullet \to {\sf revolve}$ outwardly (CW) in high pressure areas—"anticyclone"

Theory:

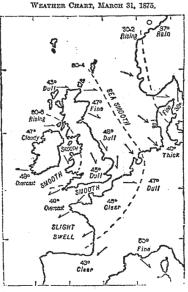
- Explained by Dove's 'Law of Gyration'
- Prediction: reversed pattern (CW/CCW) in southern hemisphere confirmed!



Theory \rightarrow Practice

The first modern weather map, London Times, Apr. 1, 1875

Galton did for weathermen what Kepler did for Tycho Brahe. This is no small accomplishment. (Wainer 2005)



The dotted lines indicate the gradations of barometric pressure. The variations of the temperature are marked by figures, the state of the sea and sky by descriptive words, and the direction of the wind by arrows—barbed and feathered according to its force. O denotes

4. Statistical atlases: Data → practice, national identity & graphical excellence

- Collection of gov't statistics on pop., trade, moral & political issues widespread in Europe & US, starting ~ 1820
- Statistical albums ~ 1870—1910
 - France: Album de Statistique Graphique: 1879-1899
 - USA: Census atlases: 1870/80/90
 - Germany: local albums (Berlin, Frankfurt, etc.)
 - Switzerland: Atlas graphique de la Suisse:1897, 1914
 - Others: Latvia, Romania, Bulgaria, etc.

50

Album de statistique graphique

- Published by the Statistical Graphics Bureau, Ministry of Public Works, Émile Cheysson, director
- 18 volumes: 1879-1899, 12—34 plates each, ~ 11"x15" pages
- Graphic forms:
 - Flow maps (simple, double, multi)
 - Pie maps, star, radial, polar time-series, proportional circles
 - Mosaic maps, anamorphic maps, planetary diagrams
 - Choropleth, bi-polar scales
 - Charts: line, bar, time-series
- Pinnacle of the Golden Age: exquisite sampler of all known graphic forms!









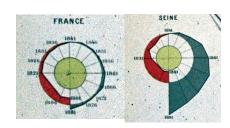
Album de statistique graphique

Spiral time-series on a map

Changes in the population of France from 1801—1881, by department [Album, 1881, plate 25]

Where is population growing most? least? declining?

Why use this graphic form?



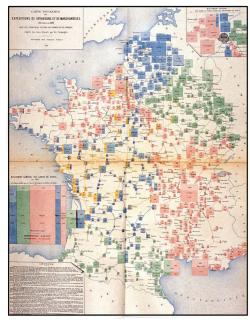


Recursive multi-mosaic map

Distribution of passengers and goods from the Paris railways to the rest of France [Album, 1884, pl. 11]

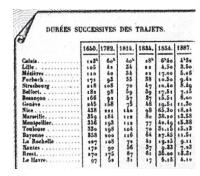
(The image that launched my interest in the history of data vis.)





Anamorphic map

Shrinking France to show change in travel time over 200 years [Album, 1888, plate 8]



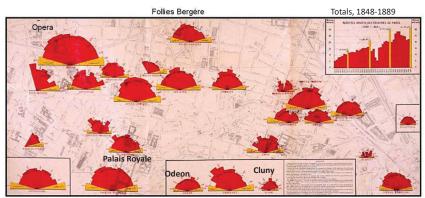


Album de statistique graphique

Q: How did Paris benefit from various int'l expos? How to show this visually?

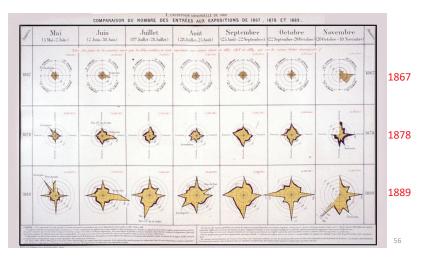
Polar area diagrams on Paris map

Gross receipts in theaters in Paris, 1878—1889, related to universal expositions [Album, 1889, plate 26]



Two-way table of star/radar diagrams

Attendance at the universal expositions in 1867, 1878, 1889 (rows), by month (cols) and days (rays). [Album, 1889, plate 21]

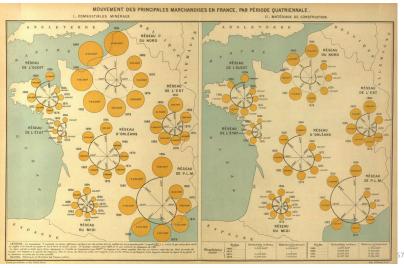


Planetary diagrams

Movement of principal merchandise by region. Spiral ~ distance; circles ~ tonnage [Album, 1895, plate 9]

Combustible minerals

Construction materials



Classed choropleth maps,

- bipolar color scale
- visualizing change

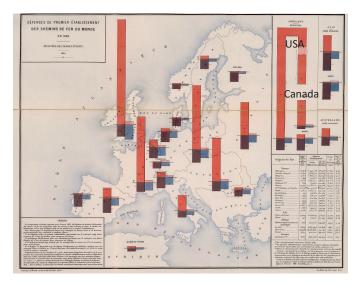
Circulation on the national roads in 'colliers réduits', a standard measure Left: 1894; Right: % change, 88-94, [Album, 1895, plate 21]



2x2 graphic bar charts

- 4 variables shown
- creative folding of long bars

Expenses of the first establishment of railroads of the world as of 1883 [Album 1886, p. 11]





ASG now online: David Rumsey

All 18 volumes, https://www.davidrumsey.com/luna/servlet/s/nl72bu



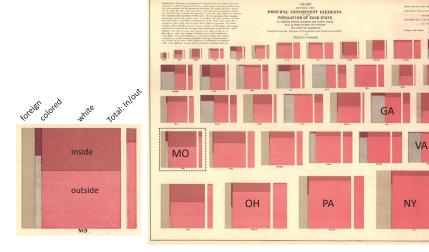
U.S. Census Atlases

- Statistical Atlas of the Ninth Census (1872) Francis Walker
 - 60 plates: First graphic portrait of the nation
 - Topics: geology, minerals, weather, pop. by ethnicity, wealth, literacy, death rates by age, sex, cause, rates of blindness, insanity, etc.
- Tenth Census (1880) Henry Gannett
 - 151 plates
- Eleventh Census (1890) Henry Gannett
 - 126 plates

51

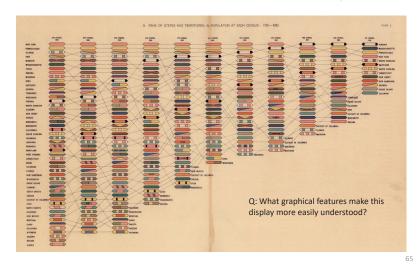
Mosaics/treemaps: Area ~ state population

State populations: Foreign born / Native colored / White + Born inside/outside [Atlas, 1870, plate 20]



Linked parallel-coordinates time-series diagram

Rank of states & territories in each census, 1790—1890. [Atlas, 1898, plate 2]

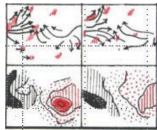


Golden Lessons

- What are the lessons for the future?
- Phenomena, not numbers or simply pretty pictures
 - Playfair, Guerry, Minard, Galton, etc. all developed new graphic forms to show phenomena of deep interest:
 - balance of trade, rates of crime, patterns in weather data, ...
- 1st lesson: data visualization today should have a similar focus







Golden Lessons: Graphical Impact

- Impact: Early ideas
 - Playfair, Guerry: data should "speak to the eyes"
 - Minard, Lalanne: allow "calculation by the eyes"
 - Nightingale: graphs should speak to the heart and mind, influence public policy & practice
- Graphical impact (Tukey, 1990)
 - Interocularity: the message hits you between the eyes
 - Immediacy: it hits you fast
 - Inescapability: it is hard to avoid the message
- **2**nd **lesson**: strive for visual impact in graphs and tables

Golden Lessons: Expressive power

- Hand-made graphics were often beautiful but entailed much sweat and hard work.
- Today: software— ease of use vs. expressive power
- Theories of graphics → graphic "languages"
 - Bertin: Semiology of graphics
 - Wilkinson: Grammar of Graphics
 - Wickham: ggplot2 R package
 - In all: the devil is in the details!
- 3rd lesson: continue to reduce the distance between a graphic idea and appearance on screen or paper.



Conclusions

The only new thing... is the history you don't know - Harry Truman

- Data visualization has deep roots:
 - Cartography
 - Statistical theory
 - Data collection
 - Visual thinking
 - Technology

All combine to give insightful views of data

Each area fed from, and nourished the others

- The Golden Age:
 - Qualitatively distinct, deserves recognition
 - Works of unparalleled beauty & scope
 - Statistical graphics had a purpose: tell a story, inform decision
 - Provides lessons for today and tomorrow