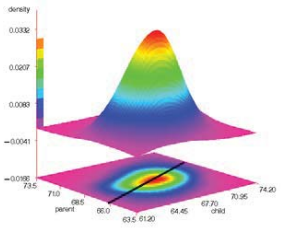


The rise of visual thinking and graphic communication

Stories & lessons from the history of data visualization

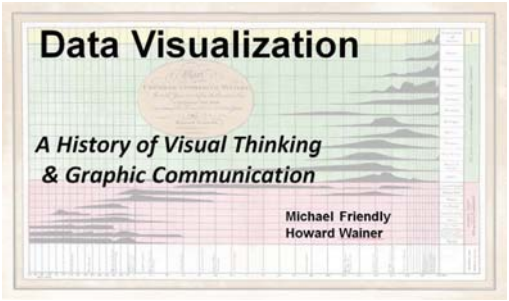
Michael Friendly

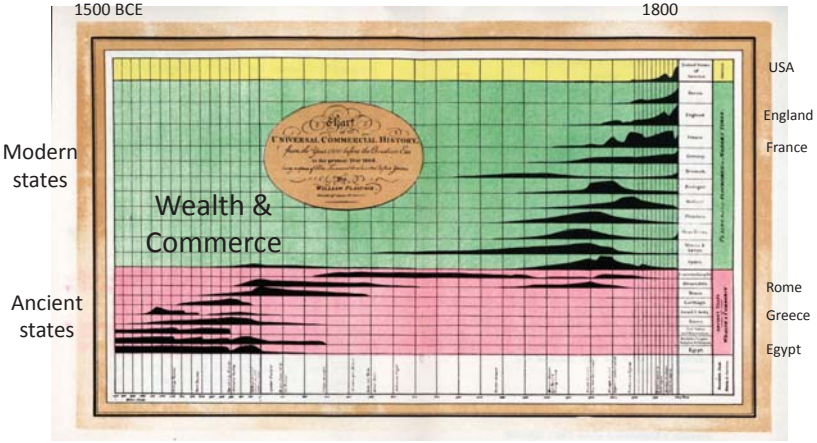
MSDS Lab, Utrecht, October 24, 2019

Slides: http://datavis.ca/papers/MSDS-talk-2x2.pdf

Obligatory shameless plug

This talk is based on our new book, Harvard University Press, 2020



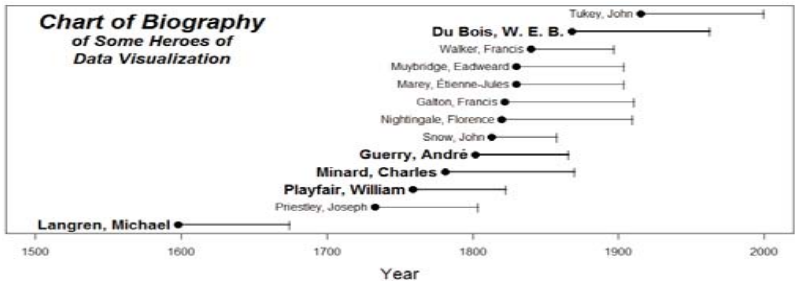


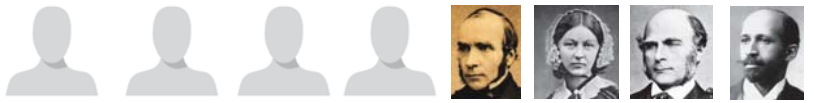
Cover image: William Playfair (1805), *Chart of Universal Commercial History*.
Q1: How and when did civilizations rise and fall from 1500 BCE to 1800 AD?
Q2: Why? How to explain?

Larger Q: How to visualize history?
Why is this a remarkable graph?

A long view of history

Dramatis personae





MF van Langren | William Playfair | C. J Minard | A.-M. Guerry | John Snow | Nightingale | Galton | WEB Du Bois

Plan for today

- Introduction: Context for history of data vis
  - Data visualization today: problems & challenges
  - Orienting Qs: When & why did visualization arise?
- Prehistory of visualization
- Some stories of the rise of visual thinking

Datavis today: Problems & challenges

- Today: Immersed in a sea of data
  - fitness trackers, election polls, economic forecasts, what's trending on Twitter
- Big data, complex, high-dimensional problems
  - Personal:
    - how to monitor my heart health? blood sugar?
    - how to manage my investment portfolio?
  - Societal:
    - Tracking disease outbreaks of measles, Ebola, etc.
    - Understanding crime, gun violence, poverty, etc.
    - Effects of climate change on extreme weather, forest fires, etc.
- How can data vis help?
  - Role of graphics in communication & persuasion?

Powerful graphs: Measels and vaccines

Visualizing the impact of health policy interventions

In 2015 Tynan DeBold & Dov Friedman in the *Wall Street Journal* show the effect of the introduction of vaccination programs in the US states on disease incidence, using color-coded heat maps for a variety of diseases

Measles was decimated!

The message hits you between the eyes!

Powerful graphs make comparison easy

In 2014, vaccination rates declined and measles re-emerged in those areas

Effective graphs can cure ignorance, but not stupidity.

Source: http://graphics.wsj.com/infectious-diseases-and-vaccines/

Orienting Questions

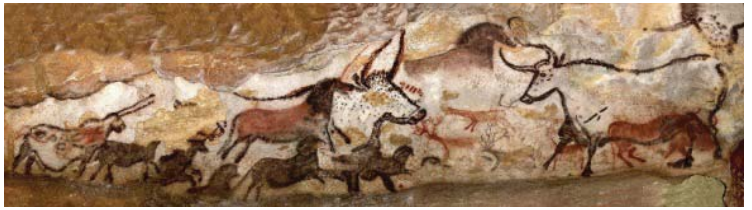
- Visualization in prehistory?
  - When did pictorial, symbolic representation arise?
  - Why? What purpose did it serve?
- How did the graphic depiction of numbers ("data") arise?
- Why?
  - What purpose did it serve?
  - What were the scientific questions promoting this?
  - How did graphic inventions make a difference?

Prehistory of visualization

Lascaux Cave, ~ 15000 BCE, the "Sistine Chapel of pre-historic art"

Lascaux II, Main chamber

Lascaux: What were they thinking?



Lascaux II, Chamber of the Bulls

- Visual features:
  - show perspective, a sense of motion, rich use of color & texture
- What was the purpose?
  - Hunting success? NO (they hunted reindeer)
  - mostly symbolic – visual language, story of communal myths
- How to understand them?
  - A **cognitive revolution:** evidence for the modern human mind in Cro Magnon man
  - inner vision, visual thinking, mental imagery– a gleam in the mind's eye
- Other cave art [20000BC – 10000BC]: Altamira (Spain); Chauvet (France), Cueva de las Manos (Argentina), …

Prehistory: Diagrams, graphic stories

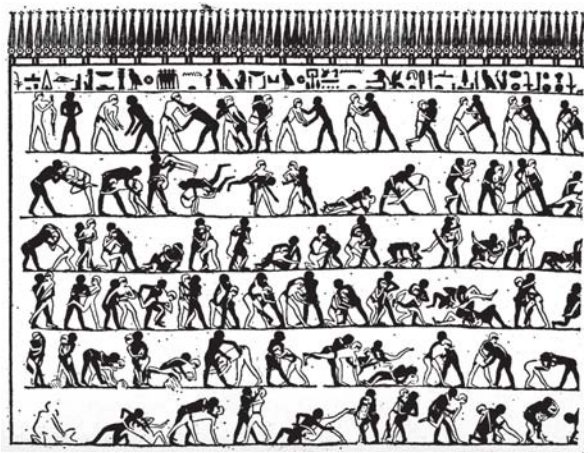
Early Egyptian animated graphic diagram

Wrestling scene on east wall, tomb of Baqt at Beni Hasan (ca. 2000 BCE).

A **visual explanation** of a wrestling match

Anticipates modern graphic novels

Why? Perhaps Baqt's last lesson as a wrestler in his youth and later as a coach



Visualizing the known world

A next step in visual thinking was to visualize **space** beyond what the eye could see.

How to show visually what we know about the known world?

This was the origin of maps.

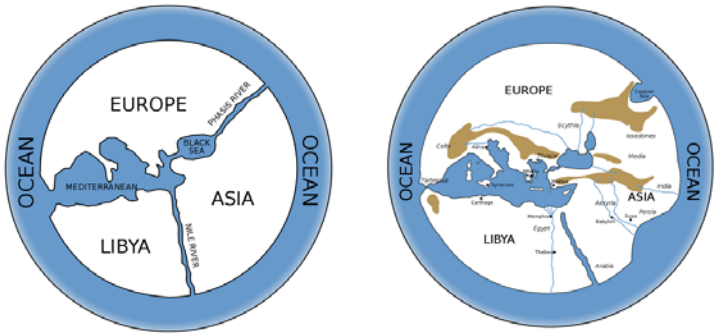
The epic poems of Homer, the *Iliad* and the *Odyssey* told stories of the Mycenaean Greeks, ~1600-1100 BCE: The siege of Troy, return of Odysseus to Ithaca & Penelope.

But only in words.

This early Babylonian world map, from ~ 600 BCE showed the known world in a circular form that would become the commonplace representation.

Visualizing the known world

By around 500 BCE Axiamander of Miletius produced the first world map.
This would later be filled in with more detail by the geographer Hecataeus
Goals: exploration, navigation, trade, …



Modern reconstructions of the ancient Greek maps of Axiamander & Hecataeus

Stories of the rise of visual thinking

Stories:

- M.F van Langren & the "secret" of Longitude
- A.-M. Guerry & the rise of social science
- Graphic vision of Minard
- Galton's graphical discoveries

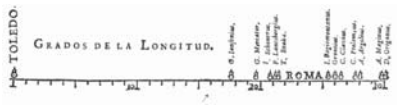
Themes:

- The idea of a GRAPH
- The birth of DATA
- Visual solutions to practical & scientific problems
- Visualization → Theory (graphic discovery)
- Data → Theory → Practice
- Escaping Flatland: 2D→ 3D

1. The IDEA of the Graph

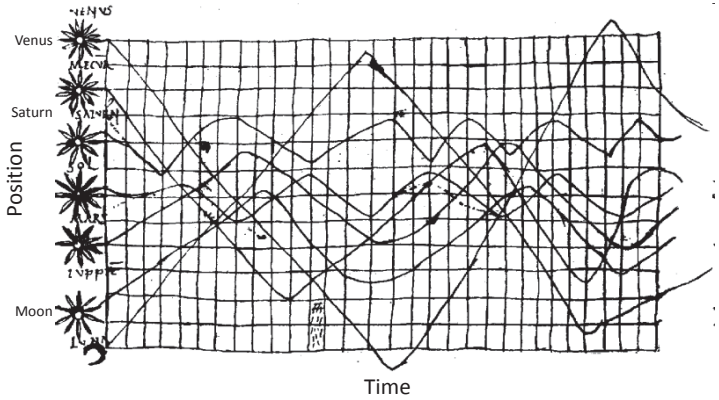
- When did the idea of an abstract visual representation of information or data arise?
  - What made this special, as distinct from earlier graphical forms?
- What was the first instance of something we can call a graph of data?
  - What does this tell us about the rise of visual thinking?

Hero of this story:
Michael Florent van Langren
- B: Apr. 1598 (Amsterdam)
- D: May 1675 (Brussels)



Early things called "graphs"

- Planetary movements ~ 10th C (Anonymous)
  - Conceptual depiction of cyclic movements of 7 heavenly bodies
  - On a grid (☺), but not really a graph of any data (☹)
  - Yet, first use of time as a horizontal axis, with a changing vertical variable



Source: Funkhouser, H. G. (1936). A note on a tenth century graph. *Osiris, 1, 260–262.*

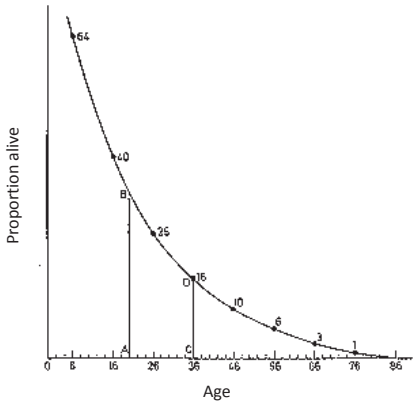
Early things called “graphs”

- Oresme (~1360): functional relations (e.g., time, velocity, distance)
  - Illustrates various functions
  - He even anticipates bad 3D graphs available in excel!
  - What’s not to like?
- Not based on data
  - “*If Oresme had data, we might have had statistical graphs 400 years before Playfair*” (Funkhouser, 1937)

Page from *Tractatus de Latitudinibus Formarum* (Oresme, ~ 1360). Often called “Oresme’s pipes”

Early things called “graphs”

A later entry, not quite not a data graph, but very close.



Christiaan Huygens (1669)

Distribution of **life expectancy by age,** derived by interpolation from John Graunt’s *Observations upon the Bills of Mortality*.

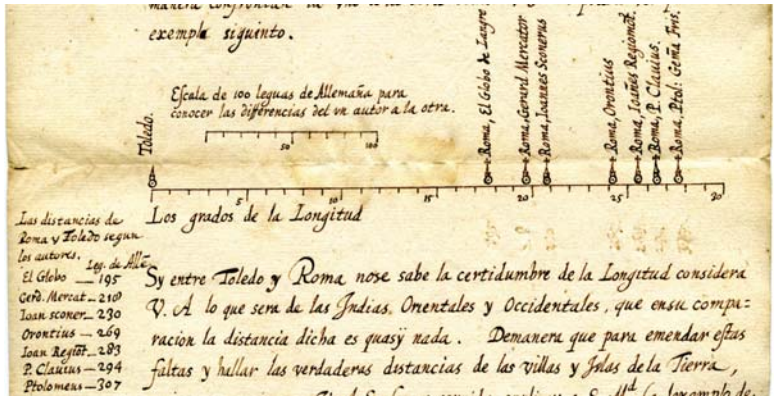
Essentially a graph of a calculated functional relation, not of statistical data

Huygens’ idea to plot these data, and then present them with a smoothed curve stands out as an important contribution to this history.

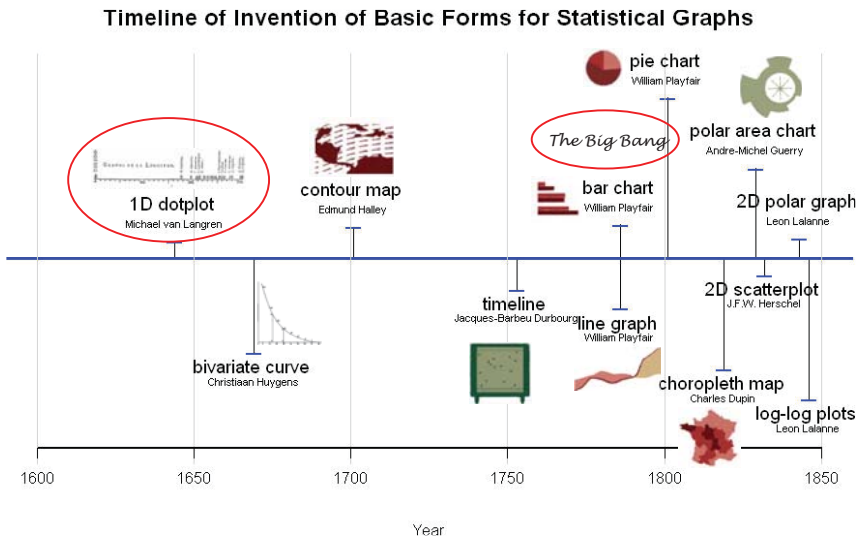
The first *real* graph of data

Michael Florent van Langren (1628) in a letter to the Infanta Isabella, regent for the Spanish court in the Netherlands

- Determinations of the longitude distance from Toledo to Rome

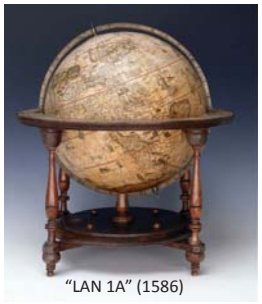


Why is MFvL’s graph important?

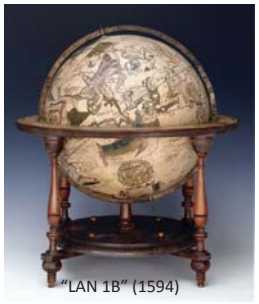


van Langren: Family

Terrestrial and celestial globes by Jacob & Arnold



"LAN 1A" (1586)



"LAN 1B" (1594)

Van Langren Family: Context

- Dutch revolt against Spanish rule, 1568 (80 Years War)
  - Van Langren globes prized by navigators
  - 1592: Jacob granted a 10-yr monopoly on globes
  - 1595: 1st Dutch merchant fleet → East India Company (1602)
  - Dutch fleets dominate trade in spices, coffee, etc. in Indonesia & elsewhere
  - Largely due to better maps & navigation

The problem of longitude: Context

- Latitude- easy to determine by inclination of sun, moon, stars above/below equator
- Longitude:
  - No fixed 0 reference; distance varies with latitude
  - Only known fact: 360º = 24hrs → 15º / hr
  - → Need to know Δtime precisely, for given event
- Solution classes:
  - 2 clocks: $\Delta time = time_{Here} - time_{There}$
  - Astrononomic: $\Delta time = time_{Here} - time_{There}$

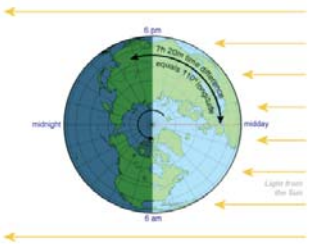
Longitude: Lunar distance method

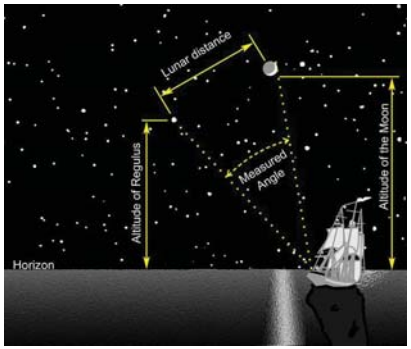
Measure the angle between moon and some other body

Lookup in almanac time at which that distance would be observed in Greenwhich: $time_{There}$

Determine $time_{Here}$ (sextant)

Longitude = 15º x Δtime

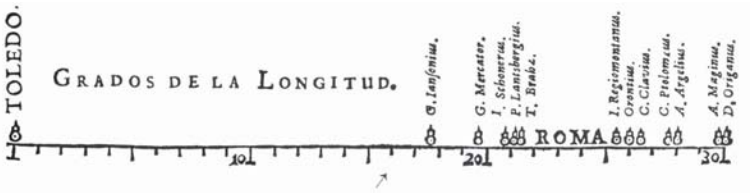




van Langren had a better idea.
He wanted to sell it to King Philip IV

Why the first graph got it right

- What was van Langren's communication goal?
- What else could he have done?
- Why did the idea of a graph occur to him?



M. F. van Langren (1644), *La Verdadera Longitud por Mar y Tierra* (The Truth about Longitude for Sea and Land)

What else could he have done?

- What would occur to men of his time to convey a message to the King?
- … he could used a *table* have sorted by *year* to establish *priority* (or show change).

Sorted by Priority

Year	Name	Longitude	Where
150	Ptolomeus, C.	27.7	Egypt
1463	Regiomontanus,	25.4	Germany
1530	Lantsbergius, P.	21.1	Belgium
1536	Schonerus, I.	20.8	Germany
1542	Ortonius	26.0	France
1567	Mercator, G.	19.6	Flanders
1567	Clavius, C.	26.5	Germany
1578	Brahe, T.	21.5	Denmark
1582	Maginus, A.	29.8	Italy
1601	Organus, D.	30.1	Germany
1605	Iansonius, G.	17.7	Flanders
1610	Argelius, A.	28.0	Italy

Answers: Who did it when?

- … he could have sorted by *longitude*, to show the *range*.

Answers: How much did they vary?

Sorted by Longitude

Longitude	Name	Year	Where
17.7	G. Iansonius	1605	Flanders
19.6	G. Mercator	1567	Flanders
20.8	I. Schonerus	1536	Germany
21.1	P. Lantsbergius	1530	Belgium
21.5	T. Brahe	1578	Denmark
25.4	I. Regiomontanus	1463	Germany
26.0	Orontius	1542	France
26.5	C. Clavius	1567	Germany
27.7	C. Ptolomeus	150	Egypt
28.0	A. Argelius	1610	Italy
29.8	A. Maginus	1582	Italy
30.1	D. Organus	1601	Germany

- … he could have sorted by *name*, to show *authority*.

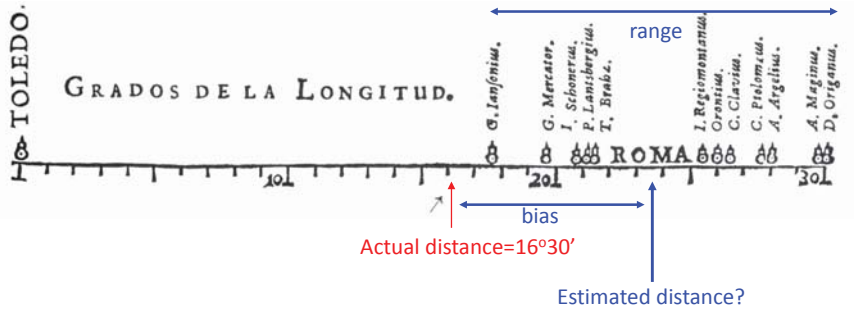
Answers: What did XXX say?

Sorted by Authority

Name	Longitude	Year	Where
Argelius, A.	28.0	1610	Italy
Brahe, T.	21.5	1578	Denmark
Clavius, C.	26.5	1567	Germany
Iansonius, G.	17.7	1605	Flanders
Lantsbergius, P.	21.1	1530	Belgium
Maginus, A.	29.8	1582	Italy
Mercator, G.	19.6	1567	Flanders
Organus, D.	30.1	1601	Germany
Orontius	26.0	1542	France
Ptolomeus, C.	27.7	150	Egypt
Regiomontanus, I.	25.4	1463	Germany
Schonerus, I.	20.8	1536	Germany

Only a graph shows…

- central location
- bias
- name labels– avoiding overplotting
- wide variability
- clustering, detached observations



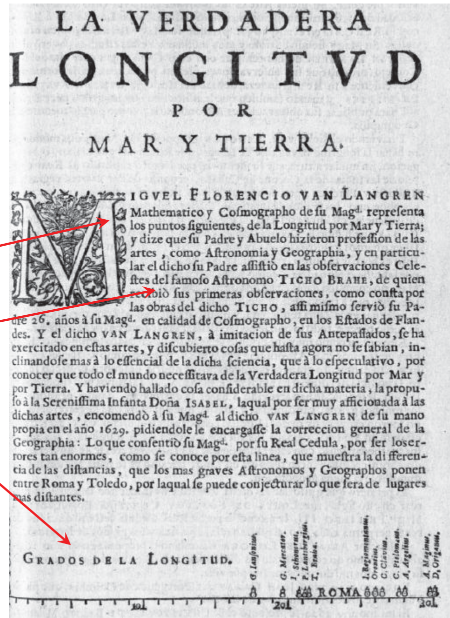
What was he thinking?

The first graph in context

From van Langren (1644), *The Truth about Longitude for Sea and Land.*

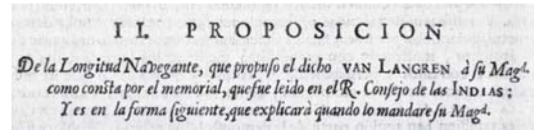
Patronage:

- **Credentials**: I am your chief mathematician & cosmographer
- **Problem**: Navigation at sea is most important problem for you to prosper. Many others have studied this, without success.
- **Demonstration**: I show the great errors from all previous scholars.
- **Supplication**: I have a solution, if you will grant me the magnificent awards you have given to others, less worthy than I am.

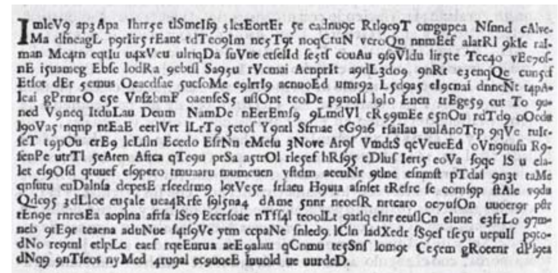

LA VERDADERA
LONGITVD
POR
MAR Y TIERRA.

MIGVEL FLORENCIO VAN LANGREN Mathematico y Cosmographo de su Magd. representa los puntos siguientes, de la Longitud por Mar y Tierra; y dize que su Padre y Abuelo hizieron professión de las artes, como Astronomia y Geographia, y en particular el dicho su Padre assistiò en las observaciones Celestes del famoso Astronomo TICHO BRAHE, de quien recibiò sus primeras observaciones, como consta por las obras del dicho TICHO, assi mismo serviò su Padre 26. años à su Magd. en calidad de Cosmographo, en los Estados de Flandes. Y el dicho VAN LANGREN, à imitacion de sus Antepassados, se ha exercitado en estas artes, y discubierto cosas que hasta agora no se sabian, inclinandose mas à lo essencial de la dicha sciencia, que à lo especulativo, por conocer que todo el mundo necessitava de la Verdadera Longitud por Mar y por Tierra. Y haviendo hallado cosa considerable en dicha materia, la propuso à la Serenissima Infanta Doña ISABEL, laqual por ser muy afficionada à las dichas artes, encomendò à su Magd. al dicho VAN LANGREN de su mano propia en el año 1629. pidiendole le encargasse la correccion general de la Geographia: Lo que consentiò su Magd. por su Real Cedula, por ser los errores tan enormes, como se conoce por esta linea, que muestra la differencia de las distancias, que los mas graves Astronomos y Geographos ponen entre Roma y Toledo, por laqual se puede conjecturar lo que sera de lugares mas distantes.

The "secret": van Langren's cipher


II. PROPOSICION
De la Longitud Navegante, que propuso el dicho VAN LANGREN à su Magd. como consta por el memorial, que fue leido en el R. Consejo de las INDIAS; Y es en la forma siguiente, que explicarà quando lo mandare su Magd.

On the Seafaring Longitude, that van Langren proposed to His Majesty, as is shown in the memorandum that was read in the Royal Council of the Indies. And it is in the following way, that van Langren will explain when his Majesty orders. (& PAYS ME!!)



What was he thinking: Graph or Map?

- Did van Langren think he invented something new called a "graph" or was he just making a 1D map?
  - With his family background in cartography, most likely a map
  - Toledo – Rome: Along a parallel of latitude
  - This corresponds to development of thematic cartography:

X → (X, Y) → (X, Y, Z)

then now

1D → 2D → 3D

2. The Birth of Data

- When & how did the idea of "data" arise?
- What do we mean by "data"?
  - Empirical observations ✓
  - Recorded ✓, quantified ✓, categorized ?
  - Suited to something that could be used to address some larger question or problem
  - More than just a collection of individual numbers

Hero of this story:
Andre-Michel Guerry

- B: Dec. 1802 (Tours)
- D: Apr. 1866 (Paris)

A.-M. Guerry

Early numbers, not quite "data"

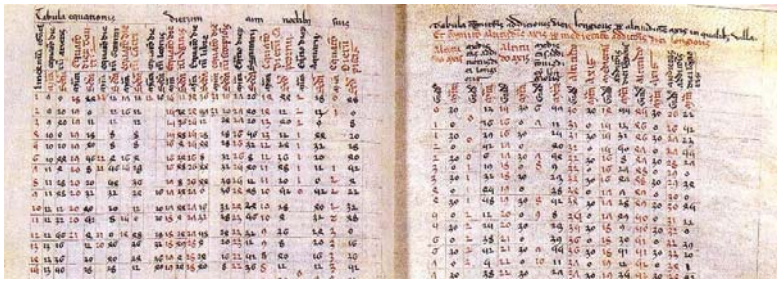
- Flooding of the Nile
  - Goes back 7000 years before construction of the Aswan Dam
  - Dates & heights of flooding recorded
  - Perhaps the longest time series data ever recorded
  - Why is this not "data"?

View of the pyramids in flood season

A nilometer was invented, ~ 700 AD

Ephemeris tables: not quite "data"

- Extensive tables of astronomical observations
  - Positions of planets, moon, etc. observed from given location
  - Tables of Toledo (~1150), Alfonsine Tables (~1260), Rudolphine tables (Kepler, 1627) using Tycho Brahe's catalog: 1' of arc
  - Included topgraphic tables – conversion to time in other cities
  - Why do I say not quite "data" as we understand this today?



Alfonsine tables, Toledo, ca.1260

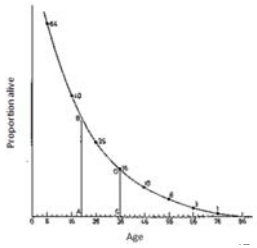
Political arithmetic

- The first inkling that "data" could be put to a larger use appeared in 1662, with John Graunt's *Natural & Political Observations Made upon the Bills of Mortality*
- It established a basis for numbers as **evidence for some proposition**.
- William Petty (1685-90) developed "political arithmetic" based on "the rule of three", allowing prediction & interpolation

a is to *b* as *c* is to *?*

$$\frac{a}{b}=\frac{c}{?} \Rightarrow ?=\frac{bc}{a}$$

- This is what Huygens (1669) used to calculate life expectancy from Graunt's data
- Annuities & life insurance could now be calculated
- This was the beginning of what we now call "statistics" (term only coined in 1749)

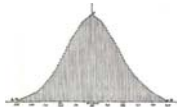


Big questions of the early 1800s

- 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
  - Birth, mortality, lifespan

- Beginnings of statistical theory and application
  - Normal dist^{n} (de Moivre, 1733)
  - *L'homme moyen* (Quetelet, 1835)



Big data of the early 1800s: France

"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-Duchâtelet (prostitution); Min. of War (desertions)
  - Suicide notes in Paris collected and analyzed for motives
- Social issues could now be addressed with **DATA**

3. 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


ESSAI
LA STATISTIQUE
MORALE
DE LA FRANCE,
A. M. GUERRY,

Social context of crime in 1820s France

- Social upheaval following Napoleon's defeat
- Crime a serious concern:
  - Explosive growth in Paris
  - Widespread unemployment,
  - Emergence of "dangerous classes"
- Liberal ("philanthrope") view
  - Increase education
  - Better prison conditions, diet (bread ***and*** soup)
  - Religious instruction
- Conservative view
  - Build more prisons
  - 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! ~ constant →	79	79	78	77	78	78
Female		21	21	22	23	22	22
Age				Accused of Theft (%)			
16–25	OMG! ~ constant →	37	35	38	37	37	37
25–25		31	32	30	31	32	31
Crime				Committed in summer (%)			
Indecent assault		.	36	36	35	38	36
Assault & battery		.	28	27	27	27	28

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?

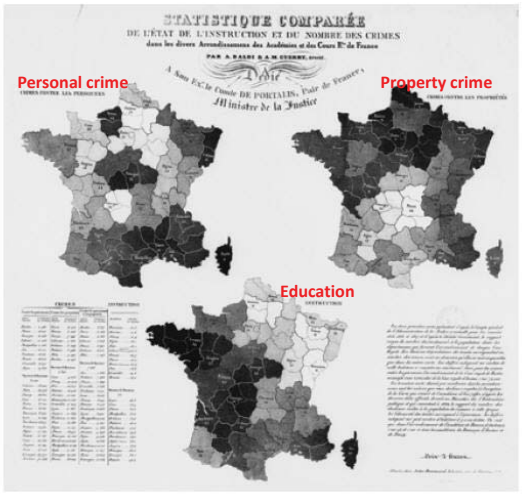
Guerry argued:

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)

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

- First shaded thematic maps of crime data
- First comparative maps of social data
- ↦ crime against persons seemed inversely related to crime against property!
- Instruction: ↦ *France obscure* and *France éclairée* (Dupin, 1826)
- North of France highest in education, but also in property crime!



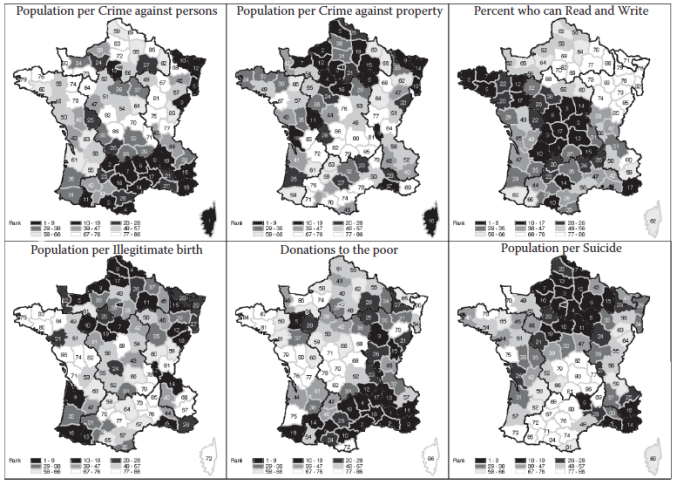
1833: *Essai sur la statistique morale de la France*

- 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

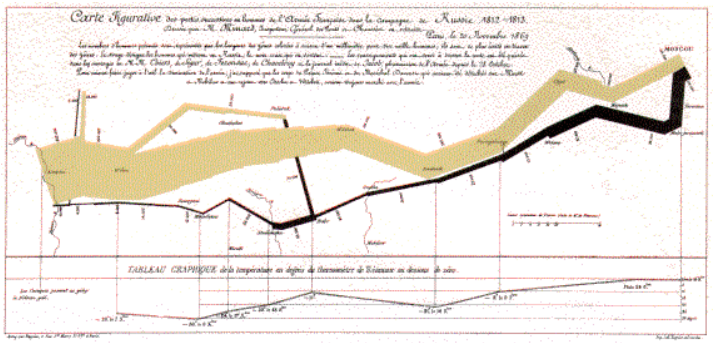
Guerry's 1833 map of literacy in France

Multivariate comparisons

Before the invention of correlation, maps of different phenomena allowed thinking about relations among disparate social variables [Darker = WORSE]



4. The graphic vision of C. J. Minard



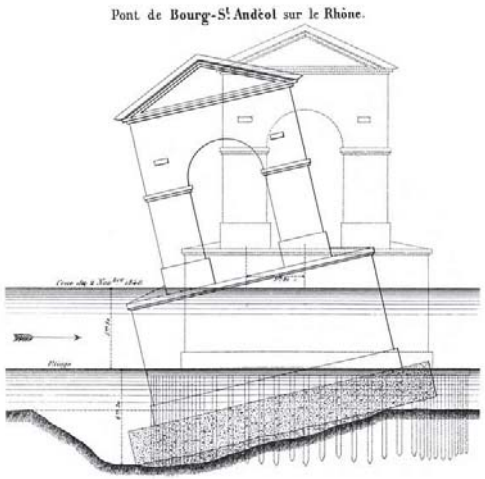
- Marey (1878): *"defies the pen of the historian in its brutal eloquence"*
- Tufte (1983): *"the best statistical graphic ever produced"*

Visual thinking, visual explanation

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

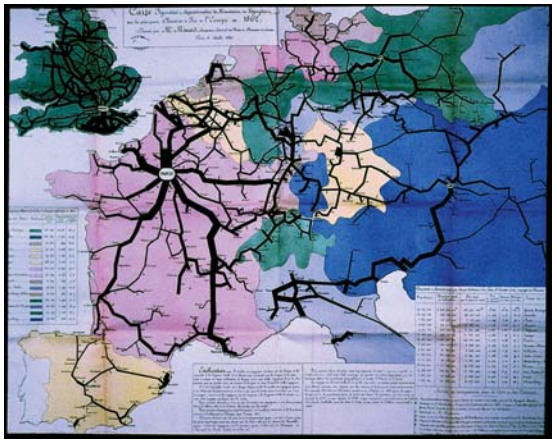
1840: Why did the bridge at Bourg-St. Andèol collapse?

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



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.

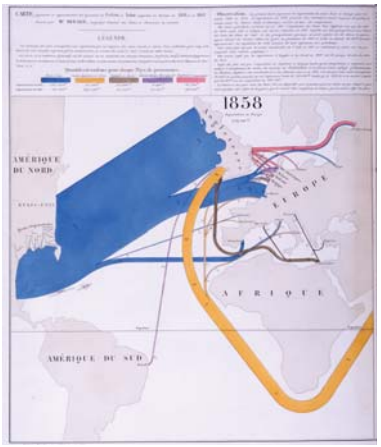
...Not only do my maps speak, but even more, they count, they calculate by the eye.

-- Minard (1862)

Effect of US Civil War on cotton trade

Visual explanation of the shift in cotton trade

Before



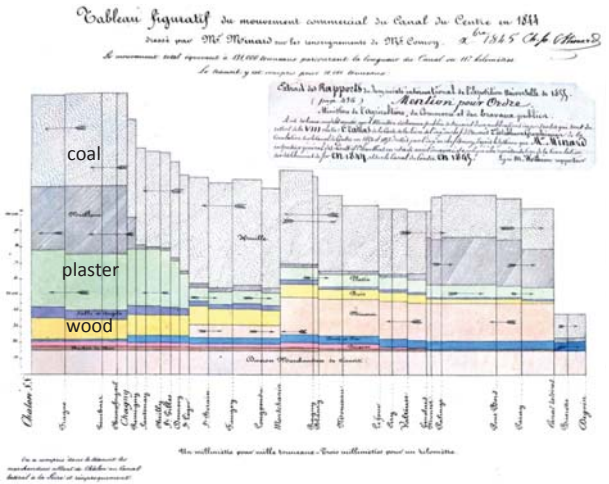
After

New graphic forms to answer questions

How to charge for the transportation of different goods on the Canal du Centre?

Visual answer:
Area ~ distance × amt
Show direction

This is an early ancestor of mosaic displays and related graphics

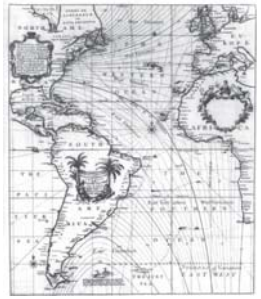


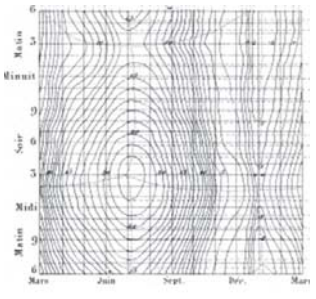
Escaping Flatland

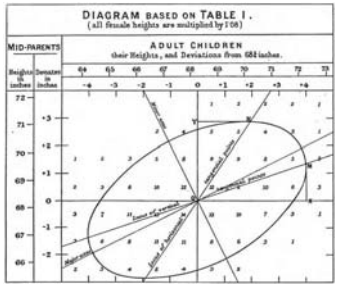
3D maps & graphs from Halley to Galton

A next step in visual thinking was the idea to show 3D+ phenomena on a 2D surface
This often involved (a) interpolation from scant data & (b) visual smoothing

1701: Halley | 1843: Lalanne | **1866**: Galton



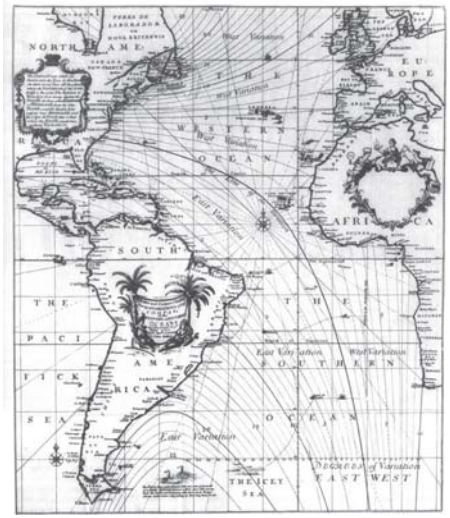




Early 3D maps & graphs

1701: Halley's contour maps showing curves of equal value (an isogonic map: lines of equal magnetic declination for the world) -- possibly the first thematic contour map of a data-based variable.

Visual ideas:
- curves showing equal value on a (lat, long) map.
- show a non-spatial variable
- Interpolation from observed data
- Regularity → smoothing



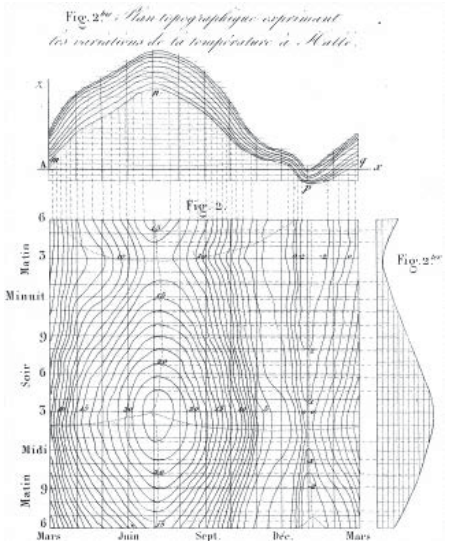
3D maps → Graphs

1843: Léon Lalanne, France
Contour diagram of a table: temperature ~ hour x month

Visual ideas:
- Ordered table like a map
- 3D level curves
- 2D marginal projections
- multiple views: plan, elevation, section
- Regularity → smoothing

month

hour



5. Galton's visual discoveries-
Bivariate normal correlation surface (1886)

Table 9.1 One of Galton's correlation tables

Height of the mid-parent in inches	Height of the adult child													
	<61.7	62.2	63.2	64.2	65.2	66.2	67.2	68.2	69.2	70.2	71.2	72.2	73.2	>73.7
>73.0	—	—	—	—	—	—	—	—	—	—	—	1	3	—
72.5	—	—	—	—	—	—	—	1	2	1	2	7	2	4
71.5	—	—	—	—	1	3	4	3	5	10	4	9	2	2
70.5	1	—	1	—	1	1	3	12	18	14	7	4	3	3
69.5	—	—	1	16	4	17	27	20	33	25	20	11	4	5
68.5	1	—	7	11	16	25	31	34	48	21	18	4	3	—
67.5	—	3	5	14	15	36	38	28	38	19	11	4	—	—
66.5	—	3	3	5	2	17	17	14	13	4	—	—	—	—
65.5	1	—	9	5	7	11	11	7	7	5	2	1	—	—
64.5	1	1	4	4	1	5	5	—	2	—	—	—	—	—
<64.0	1	—	2	4	1	2	2	1	1	—	—	—	—	—
Totals	5	7	32	59	48	117	138	120	167	99	64	41	17	14
Medians	—	—	66.3	67.8	67.9	67.7	67.9	68.3	68.5	69.0	69.0	70.0	—	—

Source: Galton (1886), p. 68.

Visual smoothing → Insight

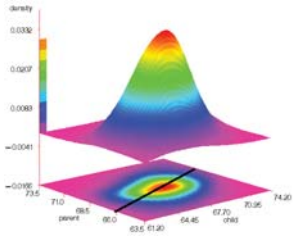
Table 9.1 One of Galton's correlation tables

Height of the mid-parent in inches	Height of the adult child													
	<61.7	62.2	63.2	64.2	65.2	66.2	67.2	68.2	69.2	70.2	71.2	72.2	73.2	>73.7
>73.0	—	—	—	—	—	—	—	—	—	—	—	1	3	—
72.5	—	—	—	—	—	—	—	1	2	1	2	7	2	4
71.5	—	—	—	—	1	3	4	3	5	10	4	9	2	2
70.5	1	—	1	—	1	1	3	12	18	14	7	4	3	3
69.5	—	—	1	16	4	17	27	20	33	25	20	11	4	5
68.5	1	—	7	11	16	25	31	34	48	21	18	4	3	—
67.5	—	3	5	14	15	36	38	28	38	19	11	4	—	—
66.5	—	3	3	5	2	17	17	14	13	4	—	—	—	—
65.5	1	—	9	5	7	11	11	7	7	5	2	1	—	—
64.5	1	1	4	4	1	5	5	—	2	—	—	—	—	—
<64.0	1	—	2	4	1	2	2	1	1	—	—	—	—	—
Totals	5	7	32	59	48	117	138	120	167	99	64	41	17	14
Medians	—	—	66.3	67.8	67.9	67.7	67.9	68.3	68.5	69.0	69.0	70.0	—	—

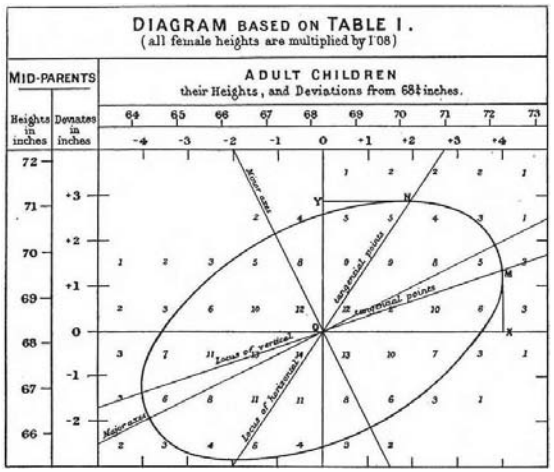
Source: Galton (1886), p. 68.

Visual insight → Theory

- Level curves are **ellipses**
- Regression lines are loci of conjugate **tangents**



… that Galton should have evolved all this … is to my mind one of the most note-worthy scientific discoveries arising from analysis of pure observation (Pearson 1920, p37)



Galton (1886, Pl X): Smoothed contours of heights of parents and children

Galton's essential ideas

- DATA: Gather data from available sources
  - Crowd sourcing, create standardized forms
- Organize
  - tables, maps, …
  - look for patterns
- Find regularities
  - visual smoothing of observed data
  - zoom out: abstract version of a map or graph
- Explain with some general theory

An early example of modern data science

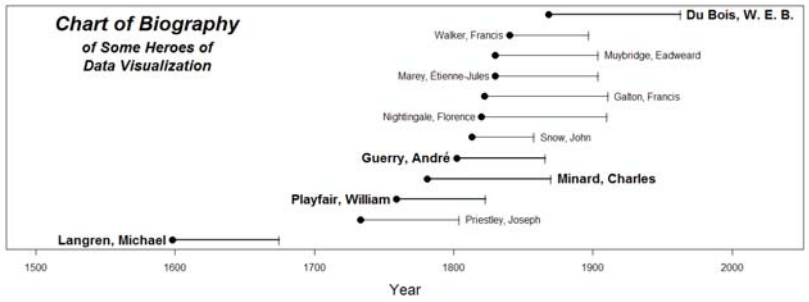
Afterword: Heroes of data visualization

- In our history of data visualization a number of individuals stand out as heroes
  - advanced the ideas of visualization through new graphic forms
  - visual thinking: idea → graphic → insight → theory
  - graphic communication: idea → insight → how to sell it?
- Some well-known, others less except for their graphic contributions
  - How to contribute to their biographies?
  - What were their personal lives? Married? Children?
  - Where were they buried?
  - How to place them in a wider context?

Afterword: Heroes of data visualization

Timeline chart of the lifespans of heroes of #datavis, in the style of Joseph Priestley

- How are they remembered today?
- How should they be celebrated?
- What were the connections in the history of visual thinking?



Afterword: Raiders of the lost tombs

- van Langren: known mainly for his naming of lunar features
  - Now know: married; 4 children 1626—1635 + 1 illegitimate daughter
  - Died in Brussels, May, 1672; buried in Notre-Dame de la Chapelle
- Guerry: death notice found by Michel Moser
  - Died April 9, 1866, residence 123 Boul. St. Michel, 5eme
  - Buried in Montparnasse, Section 9
- Minard: Lived at 32 rue du Bac, 5eme
  - Buried in Montparnasse, Section 7

M. F. van Langren: RIP

Notre-Dame la Chapelle, Brussels

Tomorrow, I go to Brussels with Gustavo & my grand-kids to pay tribute to MFvL and this Dutch connection to this history of Data Vis

Minard & Guerry in Montparnasse

Les Chevaliers: Minard's Tomb

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

Celebrated June 5, 2017

Celebrate #MinardDay on his birthday, Mar. 27

Jean-Pierre Airey-Jouglard Antoine de Falguerroles

MF Gilles Palsky

Conclusions

- Data Visualization has deep & wide roots:
  - **Cartography**: map-making, geo-measurement, thematic cartography, GIS, geo-visualization
  - **Statistics**: probability theory, distributions, estimation, models, stat-graphics, stat-visualization
  - **Data**: population, economic, social, moral, medical, …
  - **Visual thinking**: geometry, functions, mechanical diagrams, EDA, …
  - **Technology**: printing, lithography, computing…
- **Problem driven**: developments often driven by practical and theoretical problems of the day
- **Communication driven**: developments often arose from a desire to communicate better

Thank you!

Questions?

Further info:

http://datavis.ca

@datavisFriendly

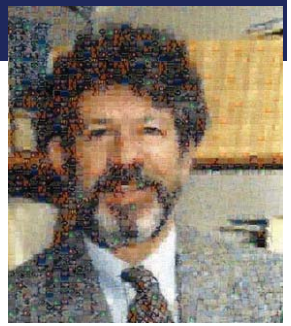


Photo mosaic of history of datavis