

Lecture 6

Visualization

History of Data Science, Spring 2022 @ UC San Diego

Suraj Rampure

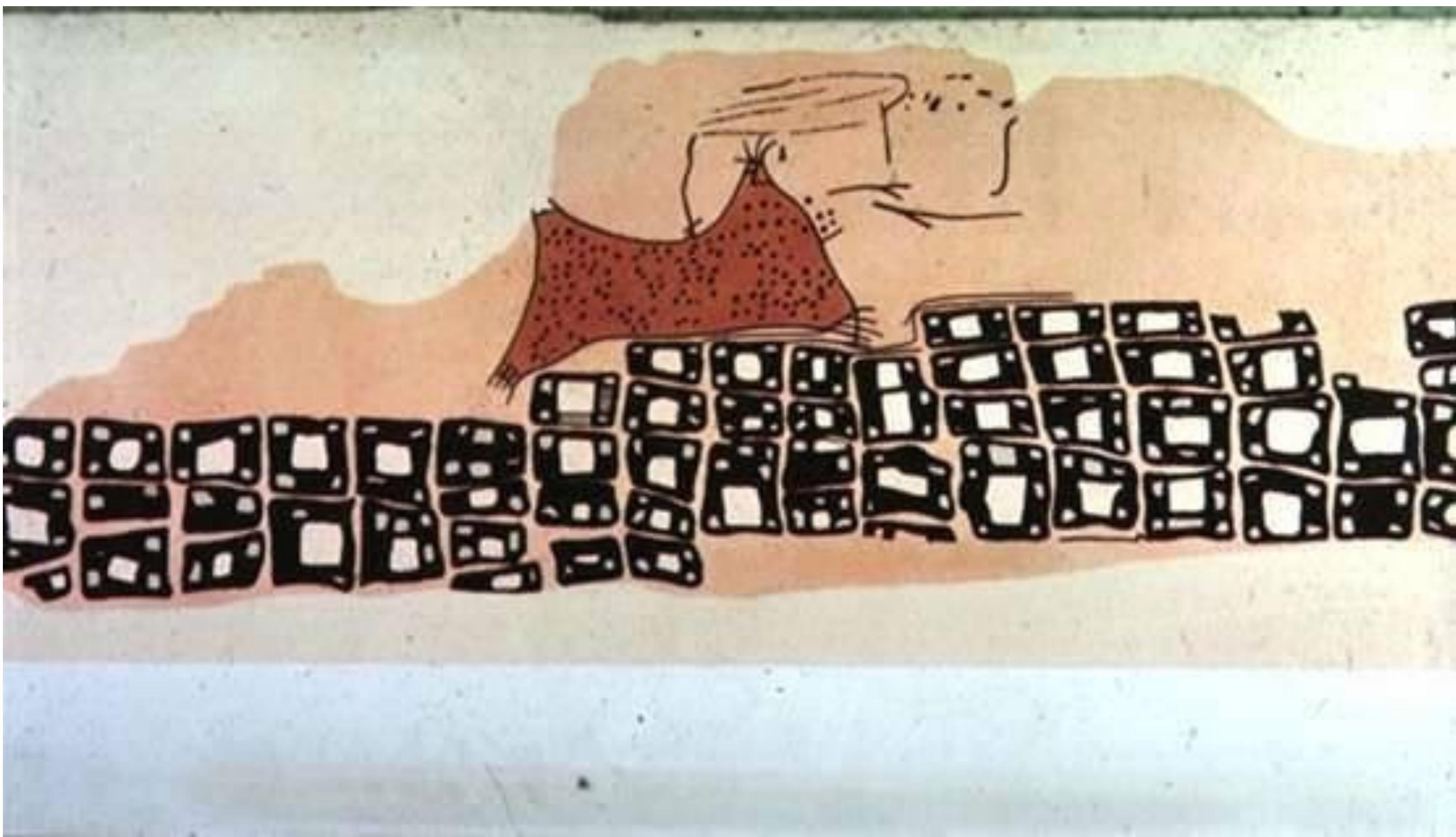
Announcements

- Homework 6 is released and is due on **Sunday, May 8th at 11:59PM.**
 - You'll get to make a website!
- Homework 4 is graded! Make sure to look at the solutions, posted on Slack and on the course website.
 - **Make sure to make an honest attempt on each question; you may not get credit for the homework if you do not.**
- Please try and attend in-person if you are able to!

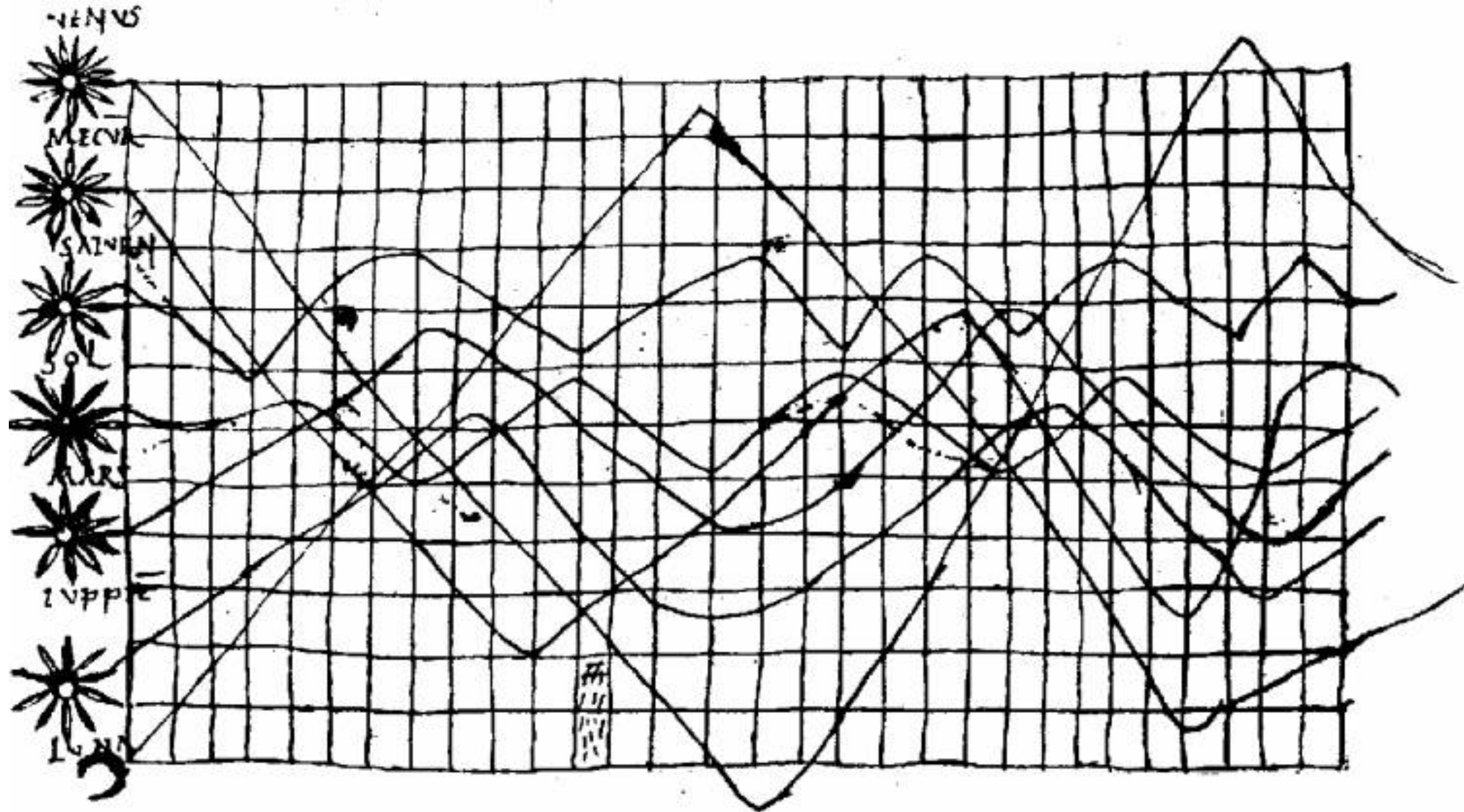
Agenda

- Today, we'll look at several examples of old visualizations.
- In addition, we will try to re-create some of these visualizations on our own in a Jupyter Notebook.
 - If the background of a slide is grey, it means that we'll re-create the visualization on that slide in the lecture notebook.
- Follow along!
- Last time, we didn't cover the derivation of the Gaussian distribution – we will revisit it later.

Early examples of visualizations



6200 BC: A map depicting the town of Konya, Turkey.



950 (AD): A line plot depicting the positions of the sun, moon, and planets throughout the year.

z

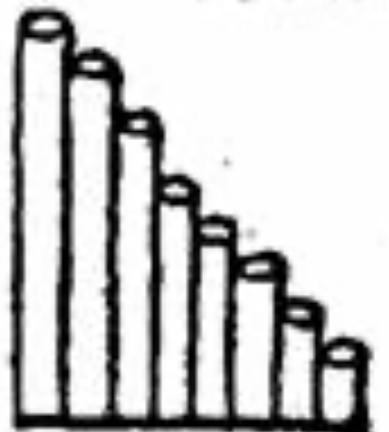
Differentia uniformiter variatio reddit uniforme scilicet et si sit ad hunc
mater differentia est differentia. **C** Lautu: uniformiter variatio
est differentia est illa quae ex celsus graduum
et ex distantiarum suarum cunctarum proportionis a ratio non est pro-
portionis explicitatio. Nam a ratio ex celsus graduum
sunt, sed ex distantiarum suarum proportionis et ratio
est. ut est. autem uniformiter differentia est per
differentiationibus membrorum secundum divisiones.
Rursum si nulla proportionis seruatur tunc nulla
possit attendi uniformitas in latitudine talis et
sic non est uniformiter differentia est differentia
C Lautu: differentia est differentia differentia
Est illa quae inter celsus graduum eque distantiarum
non ieruit eandem proportionem sicut in se-
cunda parte patet. Nonandum tamen est
quod sicut in supradictis differentiationibus ubi logitur
de excessu graduum inter se eque distantiarum
debet accipi distance scilicet partes latitudinis
est latitudine et non intensio recta ut loquuntur de differentiis
differentiis de divisione secundum simili in eis graduali



différētia



dīf̄ dīf̄, dīf̄



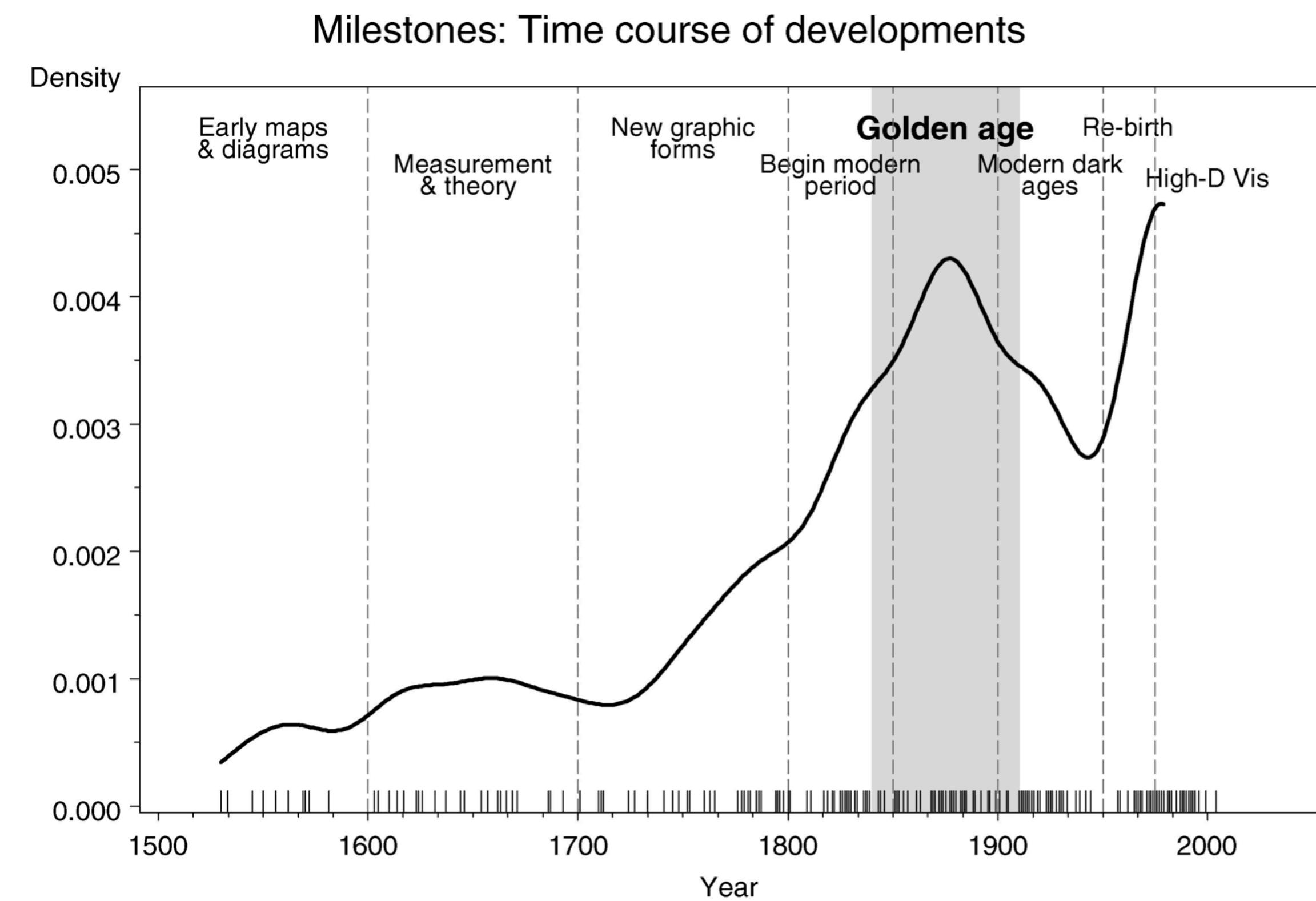
Sequitur scilicet per in quo ut
supradicta intelligatur ad
senius per figuratas geometricas
tricas ostenduntur. Et ut
omnem speciem latitudinis
in presenti materia via oc-
currit apparentior et studieret figurae geo-
metricas applicans. Hoc per dividitur per tria ca-
pitula quae p'nt' dicitur divisiones et suppositiones

1350: Nicole Oresme plotted functions of time
(e.g. velocity) as bar charts.

Today, we would use line charts or scatter plots to
show the same information.,

“Modern” data visualization

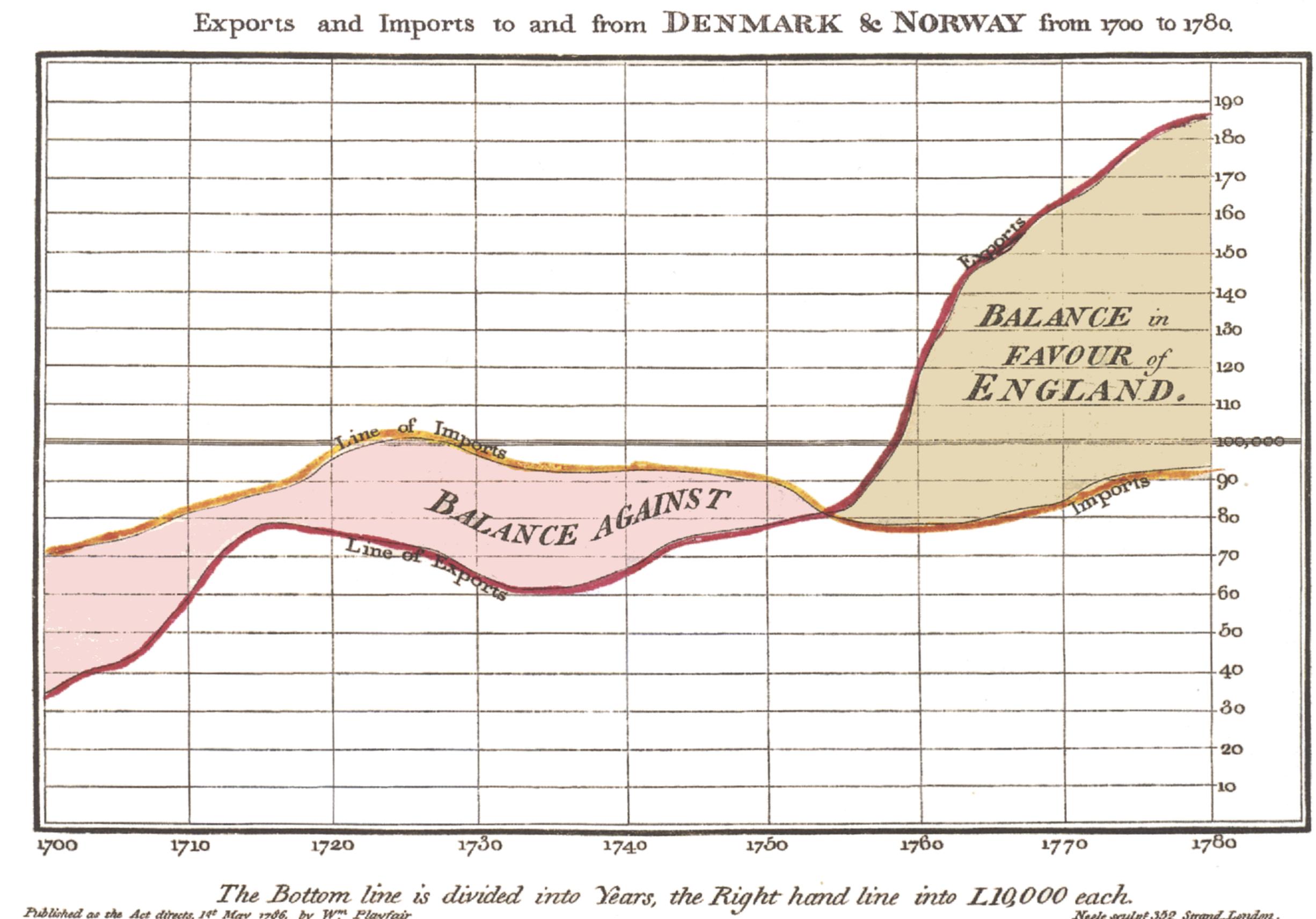
- In the late 1600s, civilizations started to gather large amounts of information about their citizens (e.g. births and deaths) and trade (e.g. imports and exports).
- The term *statistics* comes from the latin term *statisticum*, which means “of the state,” and was introduced around 1750.



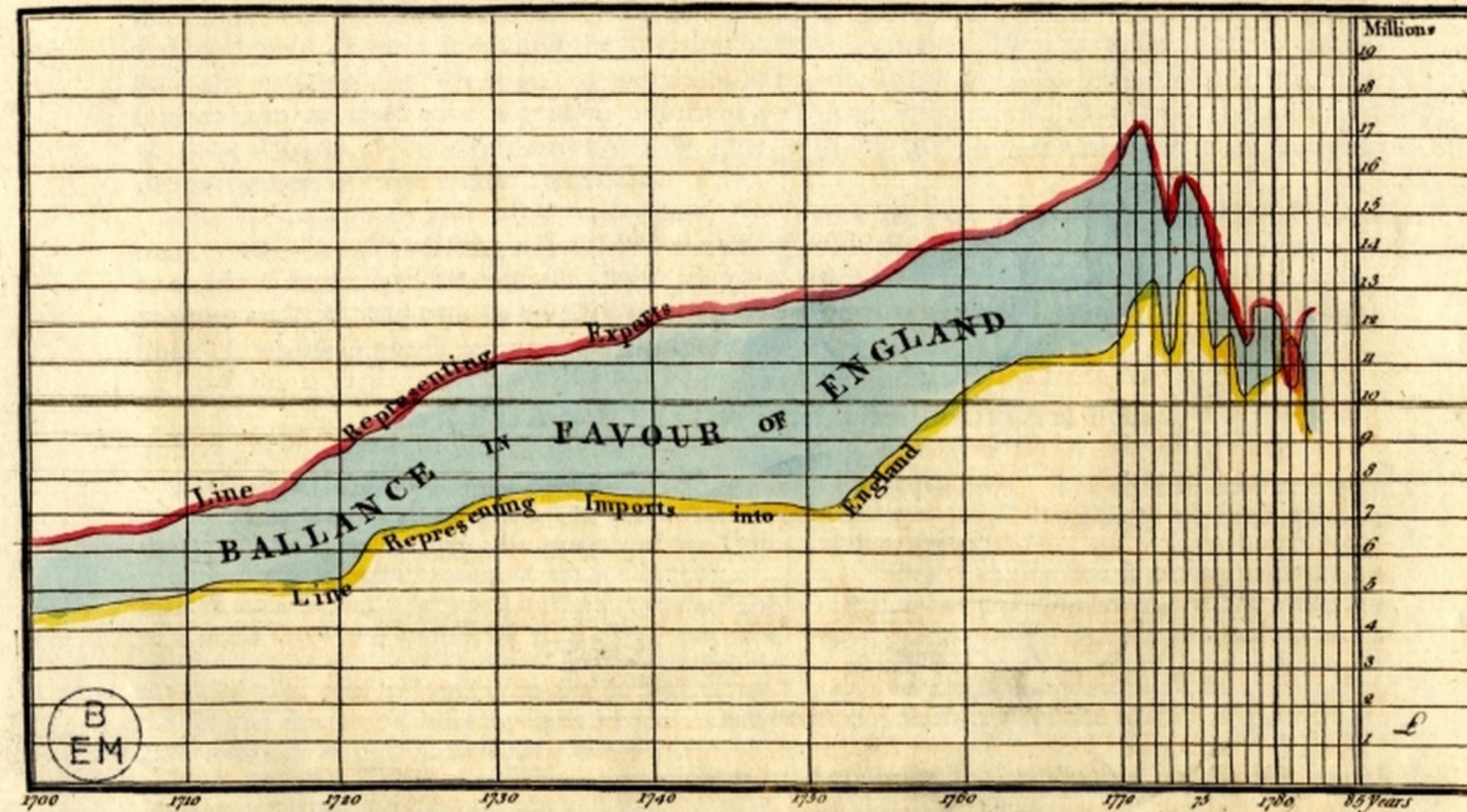
[Source](#)

William Playfair

- William Playfair (1759-1823), of Scotland, is known as the “father of data visualization”. He is credited for developing:
 - line charts
 - bar charts
 - pie charts
- One of his most famous visualizations, shown to the right, depicts England's imports and exports to Denmark and Norway (1786).
- [Interactive version here.](#)



*CHART of all the IMPORTS and EXPORTS to and from ENGLAND
From the Year 1700 to 1782 by W. Playfair*



The Divisions at the Bottom, express YEARS, & those on the Right hand, MILLIONS of POUNDS

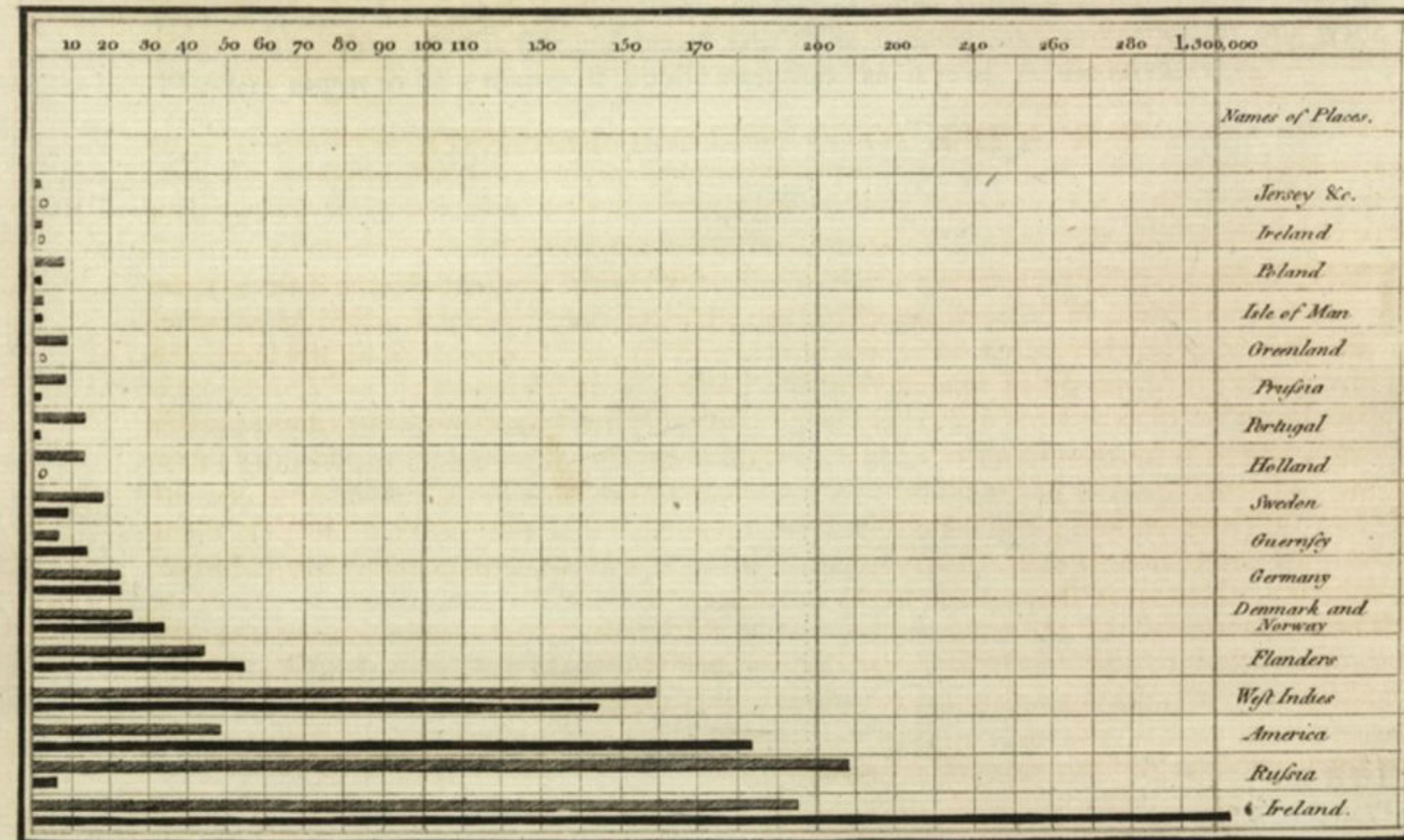
J. Finnie Sculp.

to Face Page 3rd

Published as the Act directs, 20th Aug^r. 1785

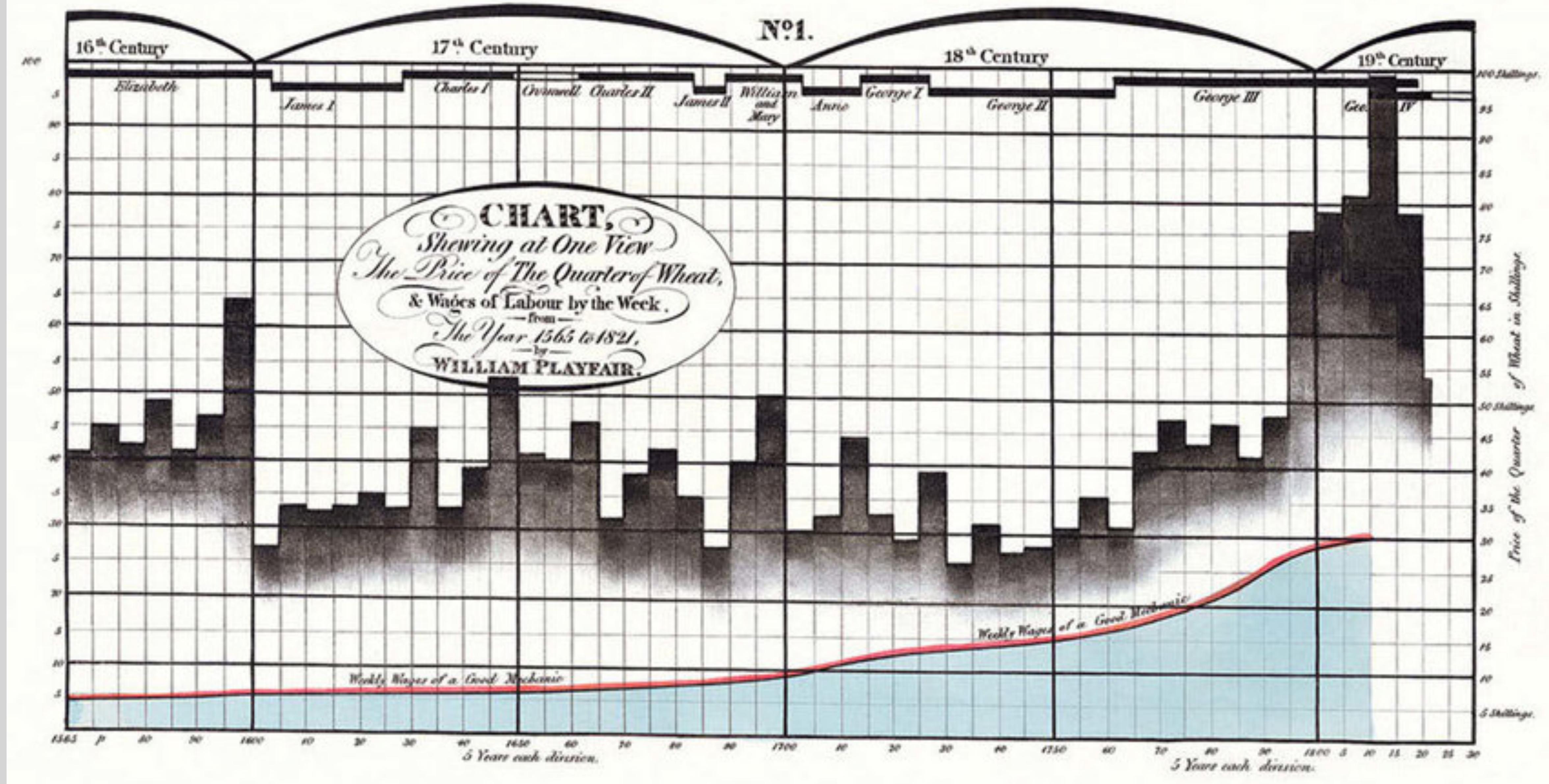
1785: Another line chart by Playfair, depicting the total imports and exports to England over a period of 85 years.

Exports and Imports of SCOTLAND to and from different parts for one Year from Christmas 1780 to Christmas 1781.

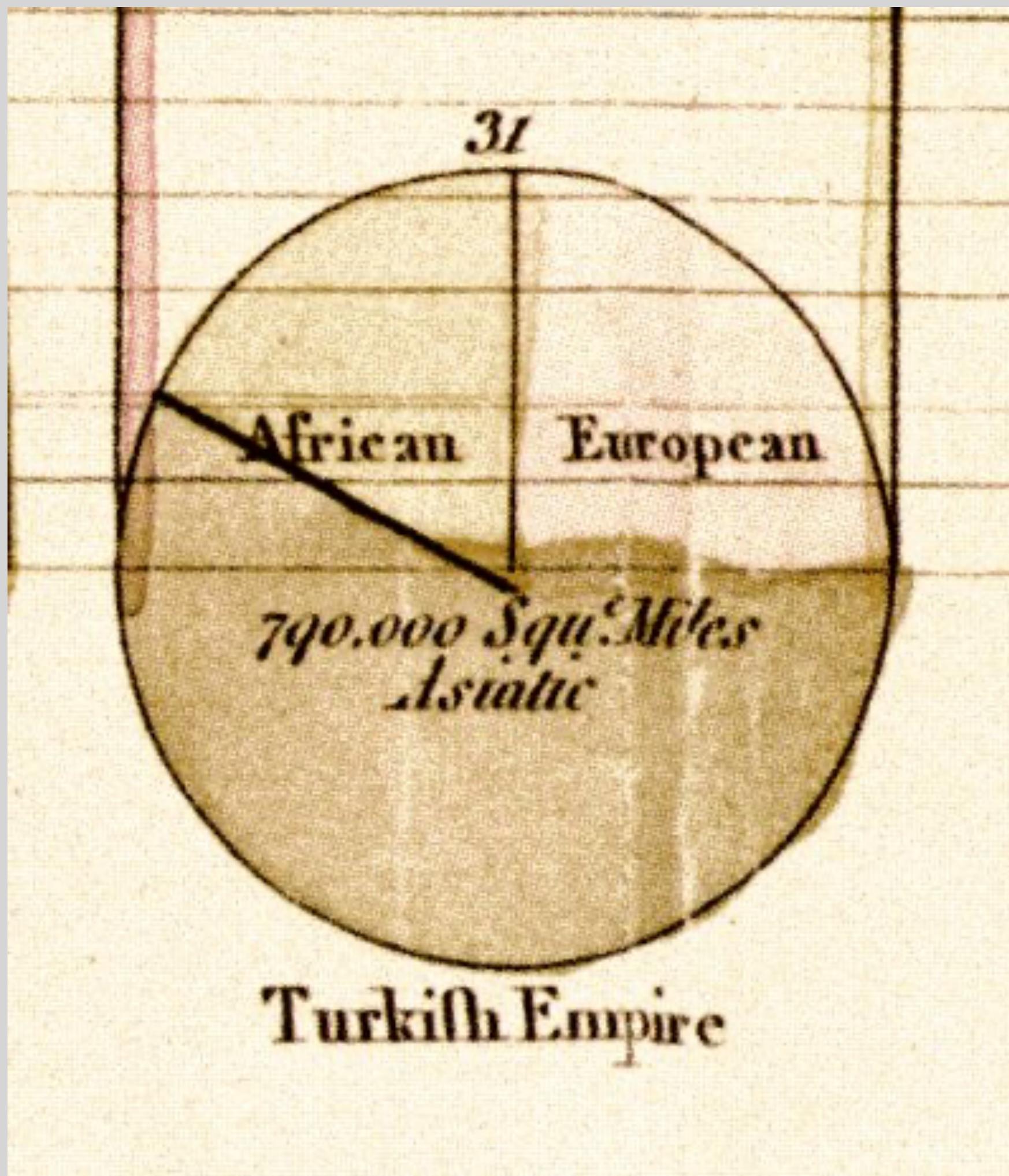


The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed lines Imports.

1796: The first known example of a bar chart, also by Playfair, depicting the imports and exports of Scotland to various countries in 1781.

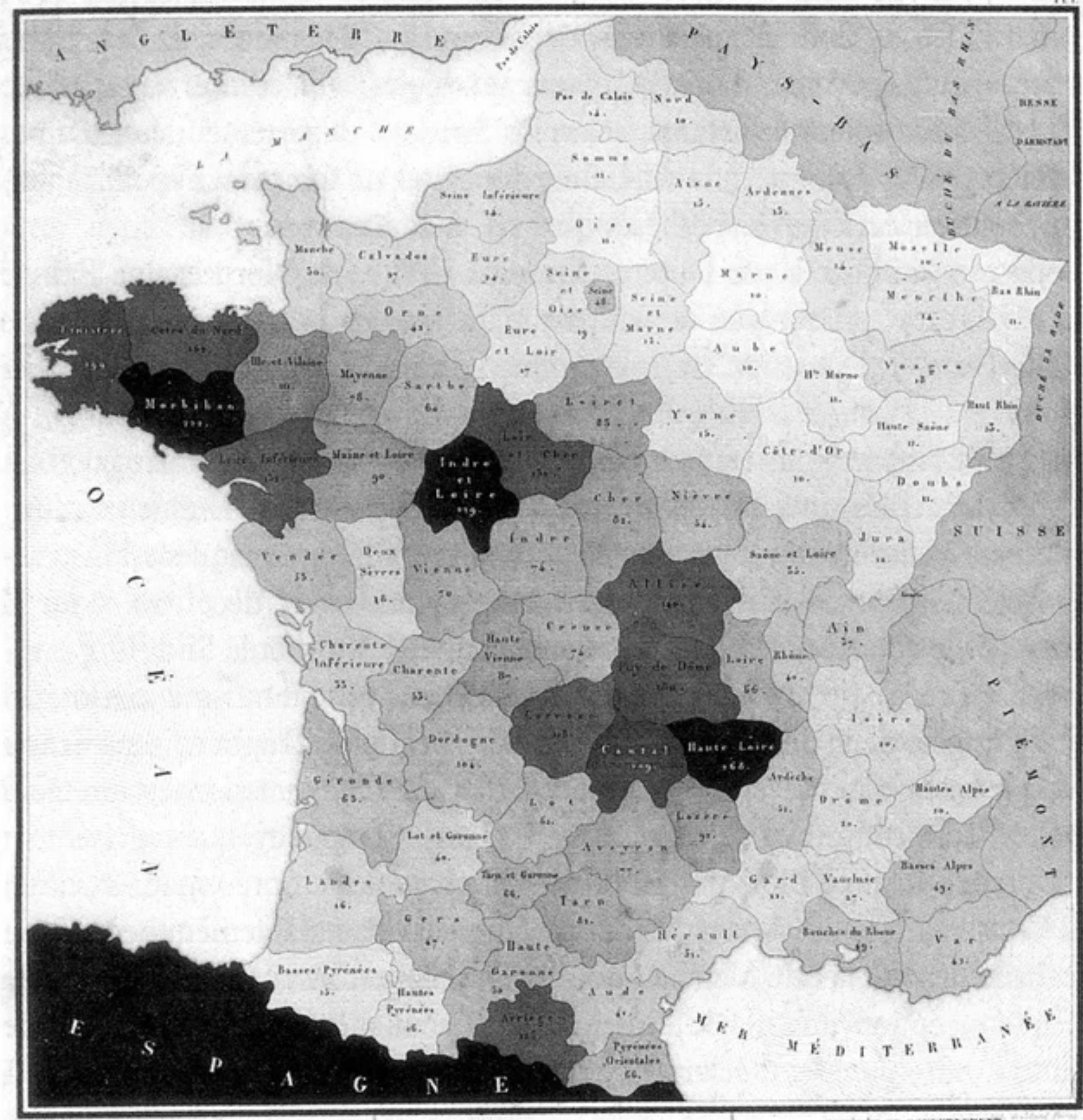


1821: Another Playfair visualization, showing the relationship between weekly labor wages and the cost of a “quarter” of wheat, along with a timeline of English monarchs, from 1565 to 1821.

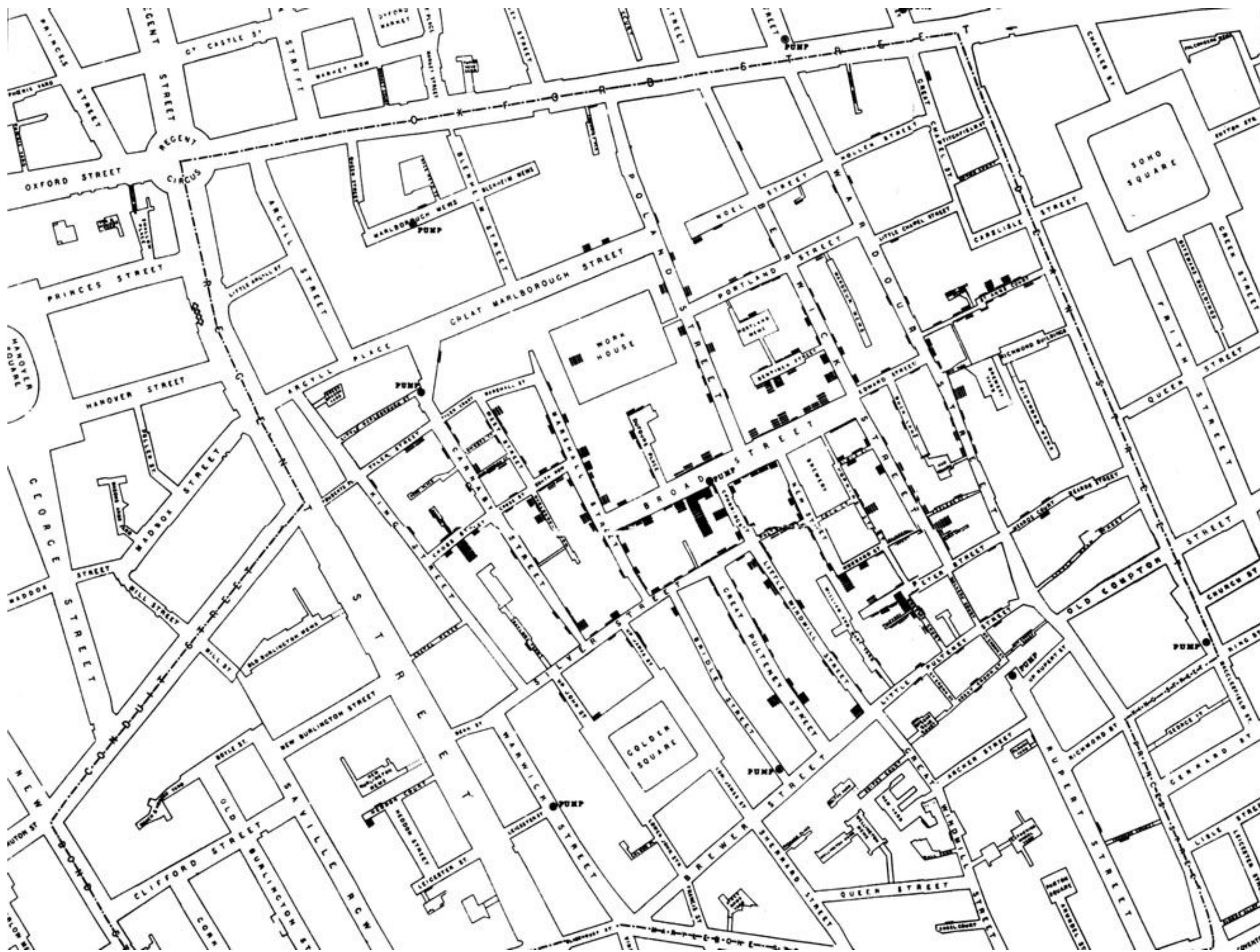


1801: Playfair's pie chart depicting the distribution of the Turkish Empire.

CARTE FIGURATIVE DE L'INSTRUCTION POPULAIRE DE LA FRANCE.



1826: Charles Dupin creates a choropleth, which describes the distribution of some quantity for each of several physical regions. His choropleth depicted rates of literacy in different parts of France.



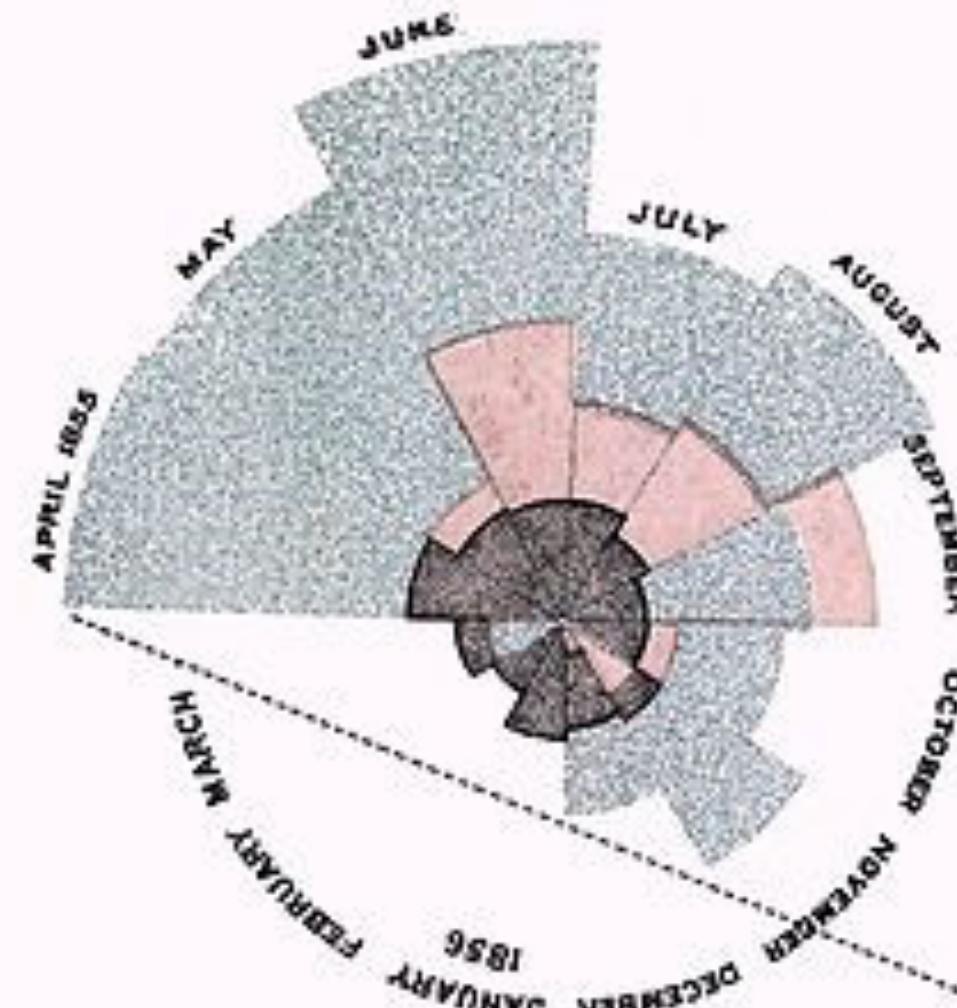
1854: John Snow mapped cholera deaths in SoHo, London. He noticed that many deaths were clustered around the Broad Street pump.



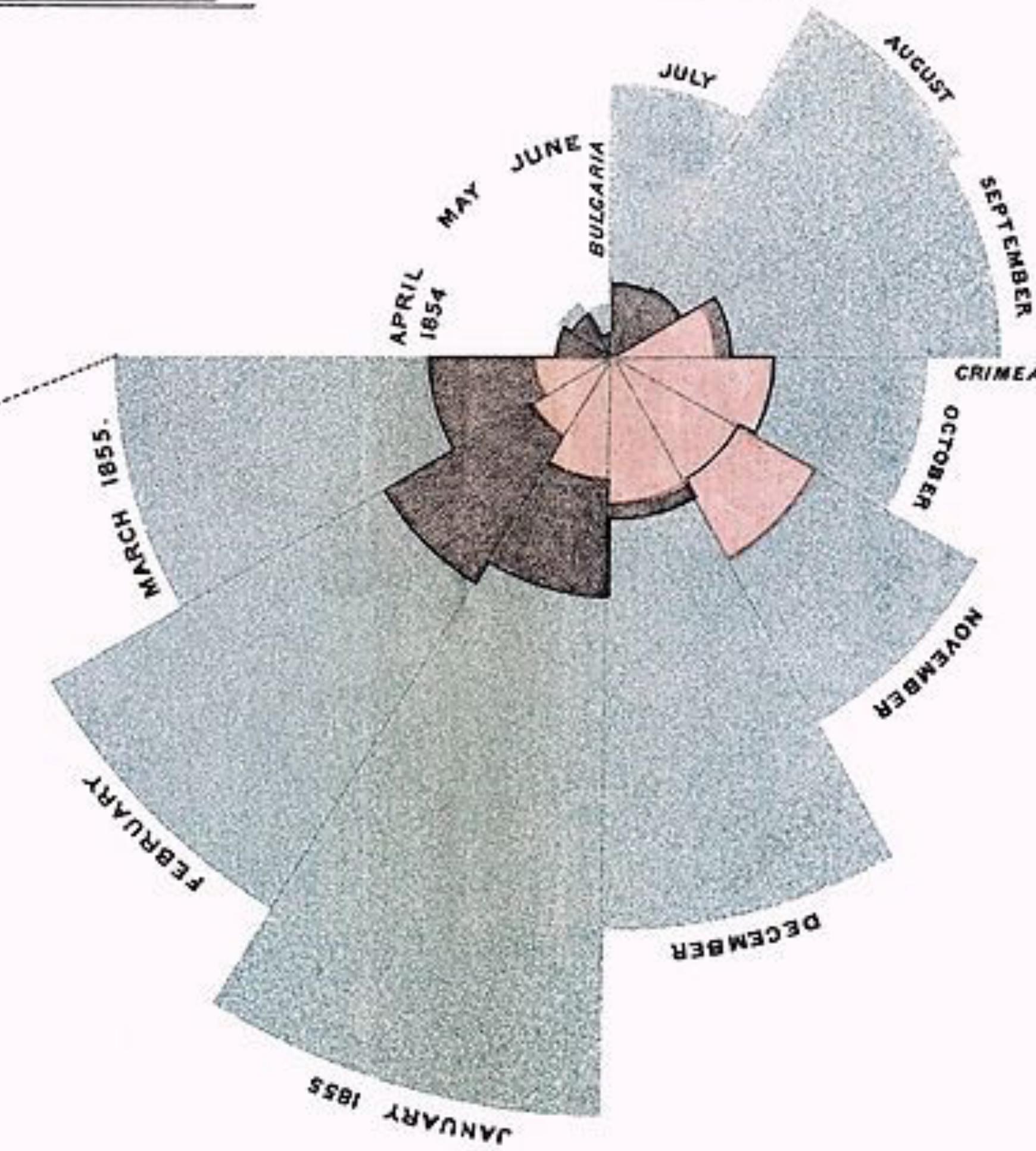
2020: The site of the Broad Street pump.

DIAGRAM OF THE CAUSES OF MORTALITY
IN THE ARMY IN THE EAST.

2.
APRIL 1855 TO MARCH 1856.



1.
APRIL 1854 TO MARCH 1855.



The Areas of the blue, red, & black wedges are each measured from the centre as the common vertex.

The blue wedges measured from the centre of the circle represent area for area the deaths from Preventible or Mitigable Zymotic diseases, the red wedges measured from the centre the deaths from wounds, & the black wedges measured from the centre the deaths from all other causes.

The black line across the red triangle in Nov. 1854 marks the boundary of the deaths from all other causes during the month.

In October 1854, & April 1855, the black area coincides with the red; in January & February 1856, the blue coincides with the black.

The entire areas may be compared by following the blue, the red & the black lines enclosing them.

1855: Florence Nightingale's depiction of the deaths of British soldiers in the Crimean war. Florence Nightingale is known as the founder of modern nursing.

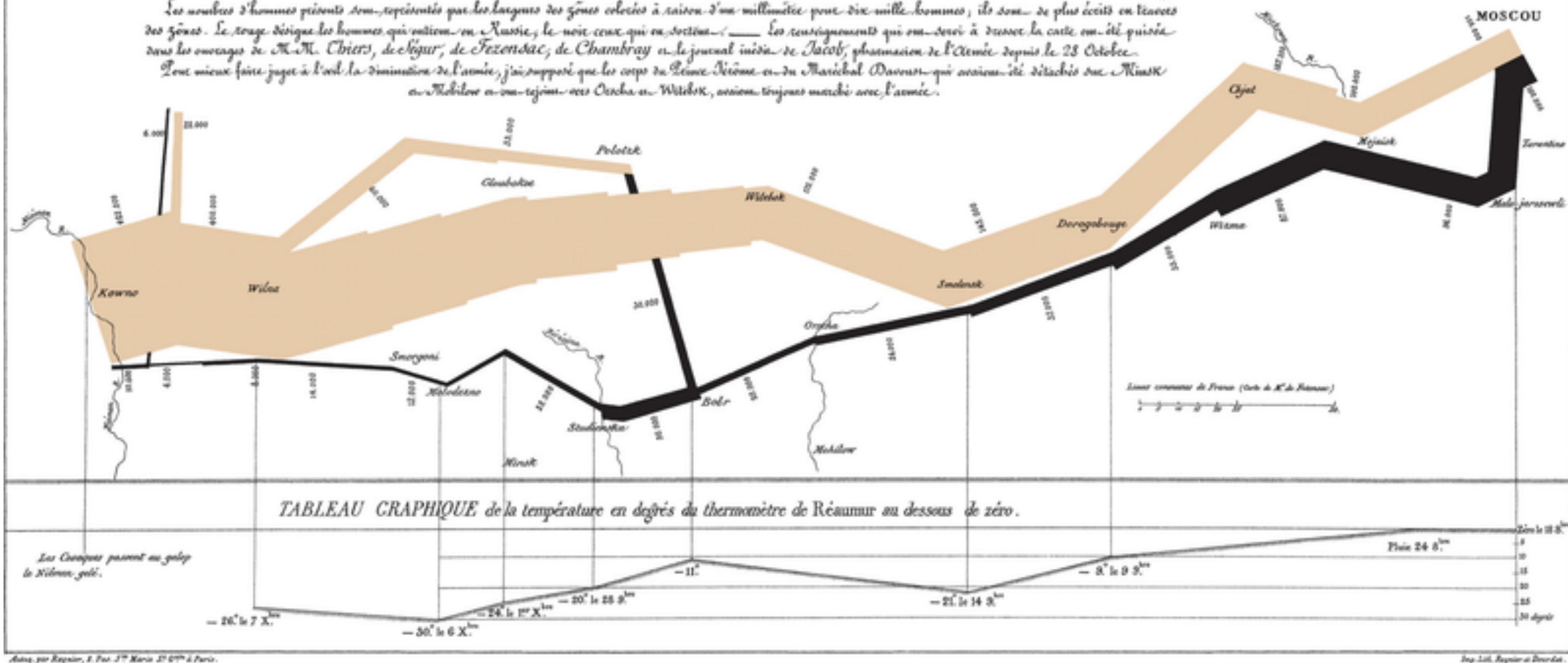
Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.

Dessiné par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite.

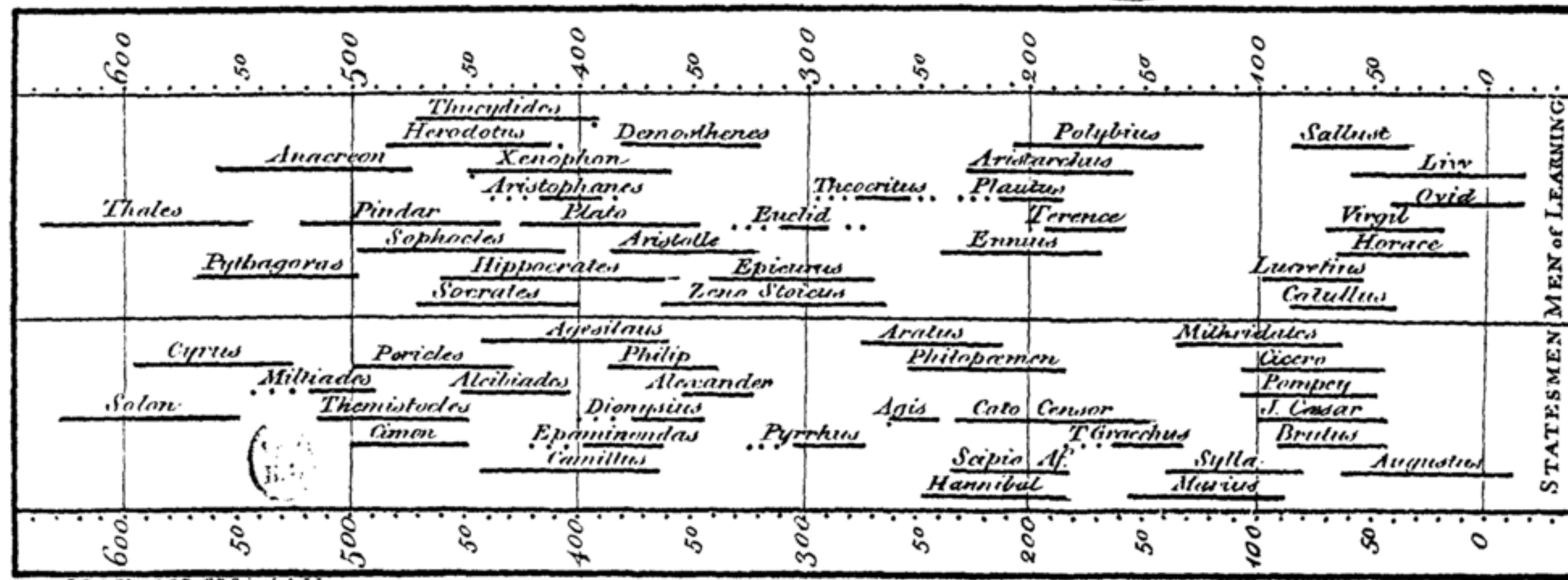
Paris, le 20 Novembre 1869.

Les nombres d'hommes présentés sont, représentés par les larges des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en lettres des zones. Le rouge désigne les hommes qui restent en Russie, le noir ceux qui se retirent. — Les renseignements qui ont servi à dresser la carte ont été pris dans les ouvrages de M.-A. Chiers, de Clémier, de Fezensac, de Chambray et le journal intime de Jacob, publication de l'Armée depuis le 28 Octobre.

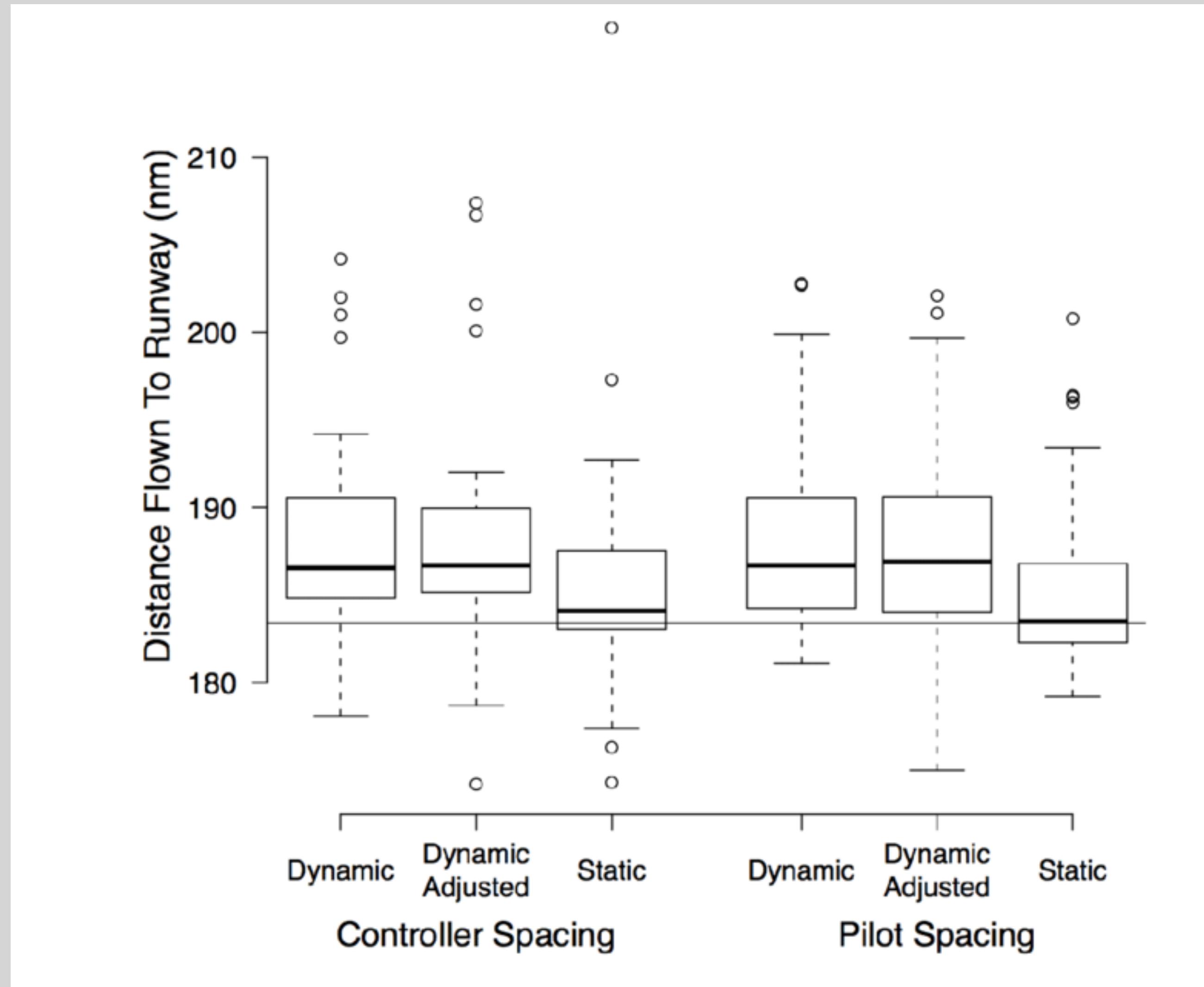
Pour mieux faire juger à l'œil la diminution de l'armée, j'ai rapporté que les corps de l'Armée Néerlandaise du Maréchal Davout qui avaient été détruits sur Niensk et Malibolz au mois d'août, avec Orelia et Witsch, avaient toujours marché avec l'armée.



A Specimen of a Chart of Biography.



1765: Joseph Priestley creates the “Chart of Biography”, a timeline of the lifespans of several prominent figures in BC. This type of visualization is now occasionally called a “Gantt chart.”



1973: John Tukey, who defined the term “Exploratory Data Analysis”, created the box plot, which describes a numerical distribution using a 5 number summary.

That's all!