

Lecture 6

Visualization

History of Data Science, Winter 2022 @ UC San Diego

Suraj Rampure

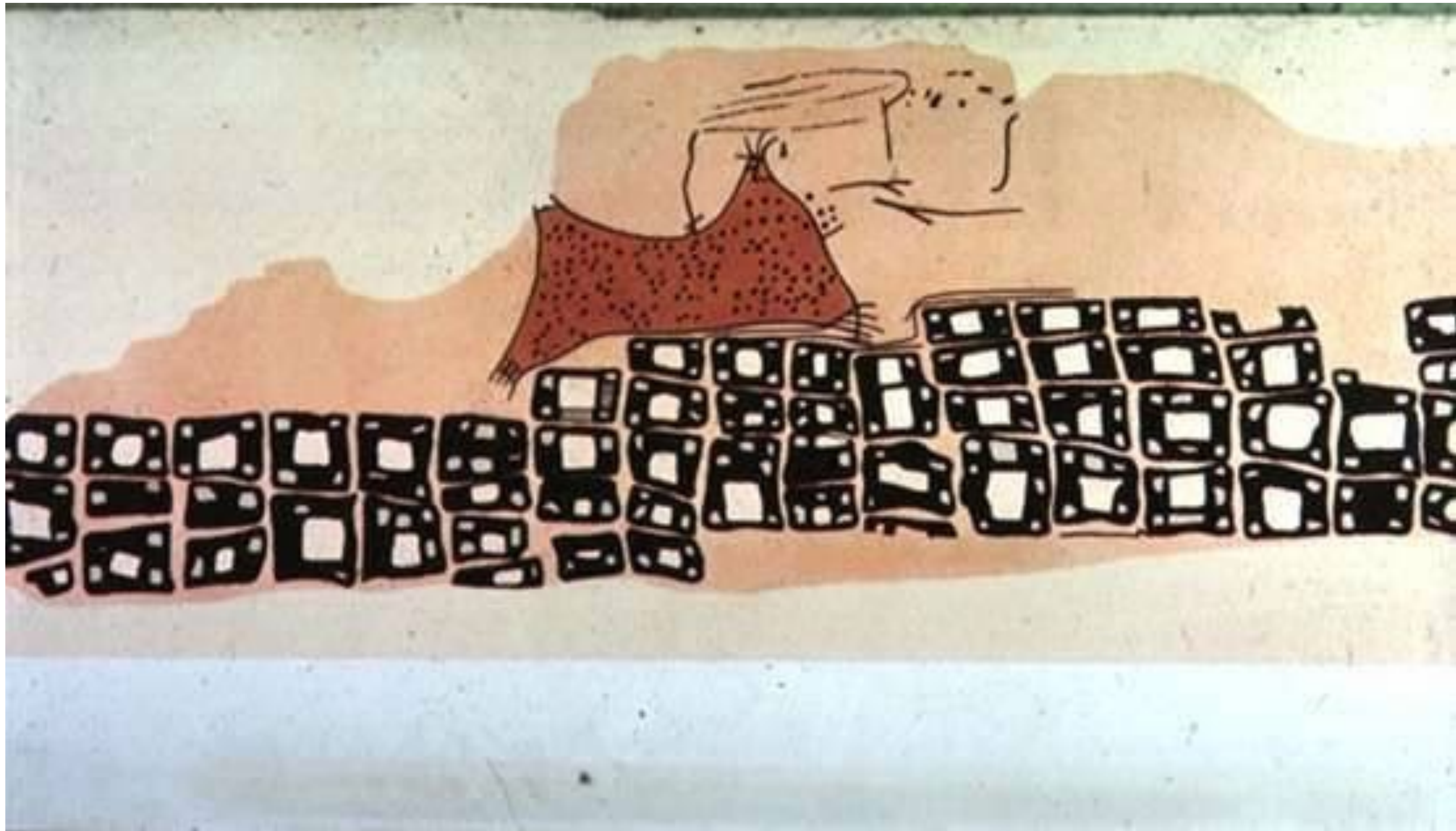
Announcements

- Homework 6 will be released by Wednesday, and will be due **Sunday, February 27th at 11:59PM.**
- Remember: no class next week due to the Presidents' Day holiday.
- **Make sure to read homework solutions (posted on Campuswire)!**

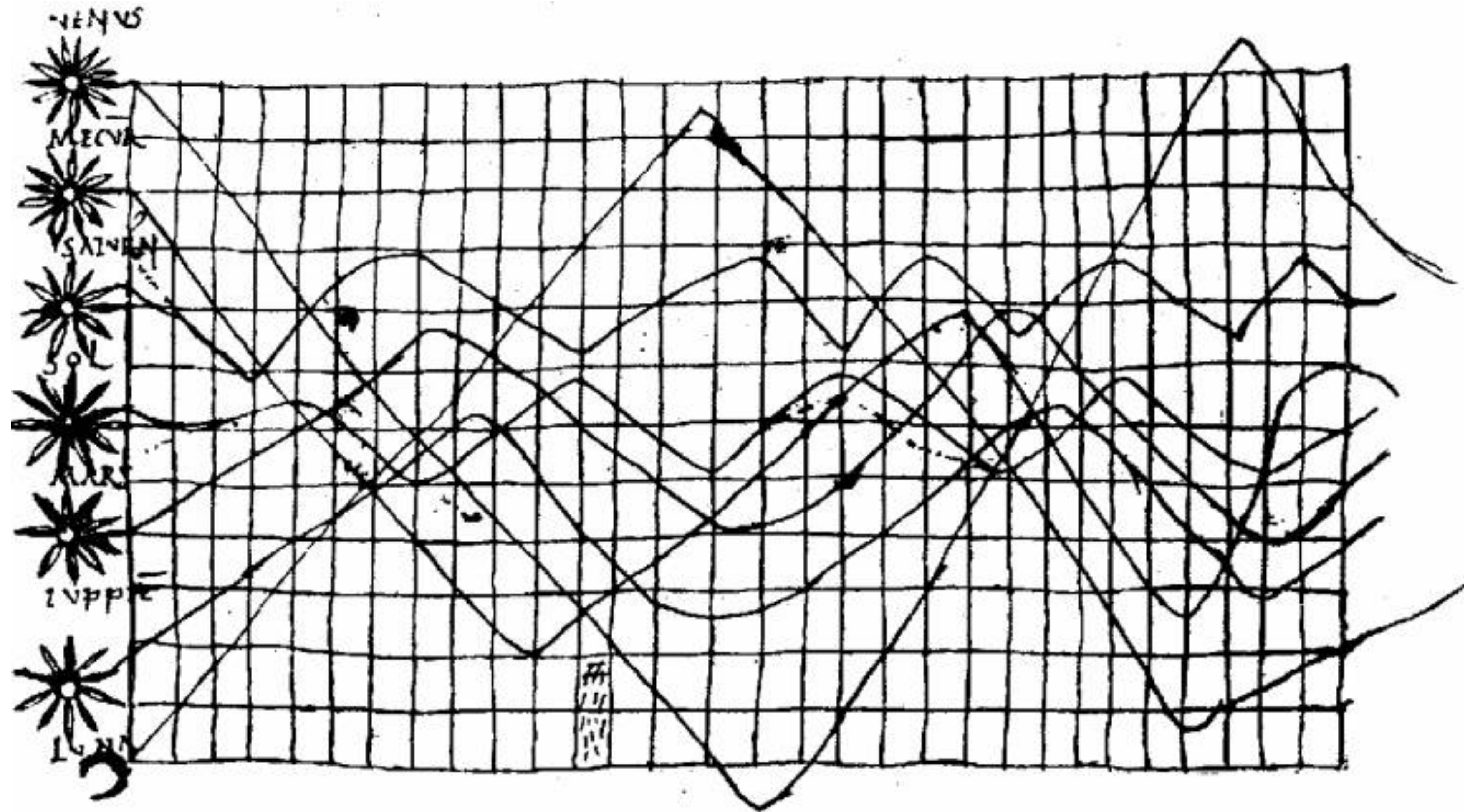
Agenda

- 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!

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.

difformis uniformiter variatio reddit unifor-
 miter difformiter difforme). ¶ Latitudo: uni-
 formis est dicitur est illa que inter excessus graduum
 equidistantium fuerit eadem proportio etiam in e-
 proportio equalitatis. Nam si inter excessus graduum
 inter se equidistantium fuerent proportio equalita-
 tis. ut est. autem: uniformiter difformis ut per ex-
 diffinitionibus membrorum secundum divisionis
 Rursus si nulla proportio seruetur tunc nulla
 posset attendi uniformitas in latitudine tali et
 sic non esset uniformiter difformis est difformis
 ¶ Latitudo: difformiter difformiter difformis
 Est illa que inter excessus graduum equidistantium
 non seruetur eandem proportionem sicut in se-
 cunda parte patebit. Notandum tamen est
 quod sicut in supradictis diffinitionibus ubi loquitur
 de excessu graduum inter se equidistantium
 debet accipi distantia secundum partes latitudinis
 extensae et non inter se ut loquuntur dicitur de dif-
 ferentia dicitur distantia sed dum simul in eum graduali

scilicet et fiat ad unum



difformiter difformis



difformiter difformis



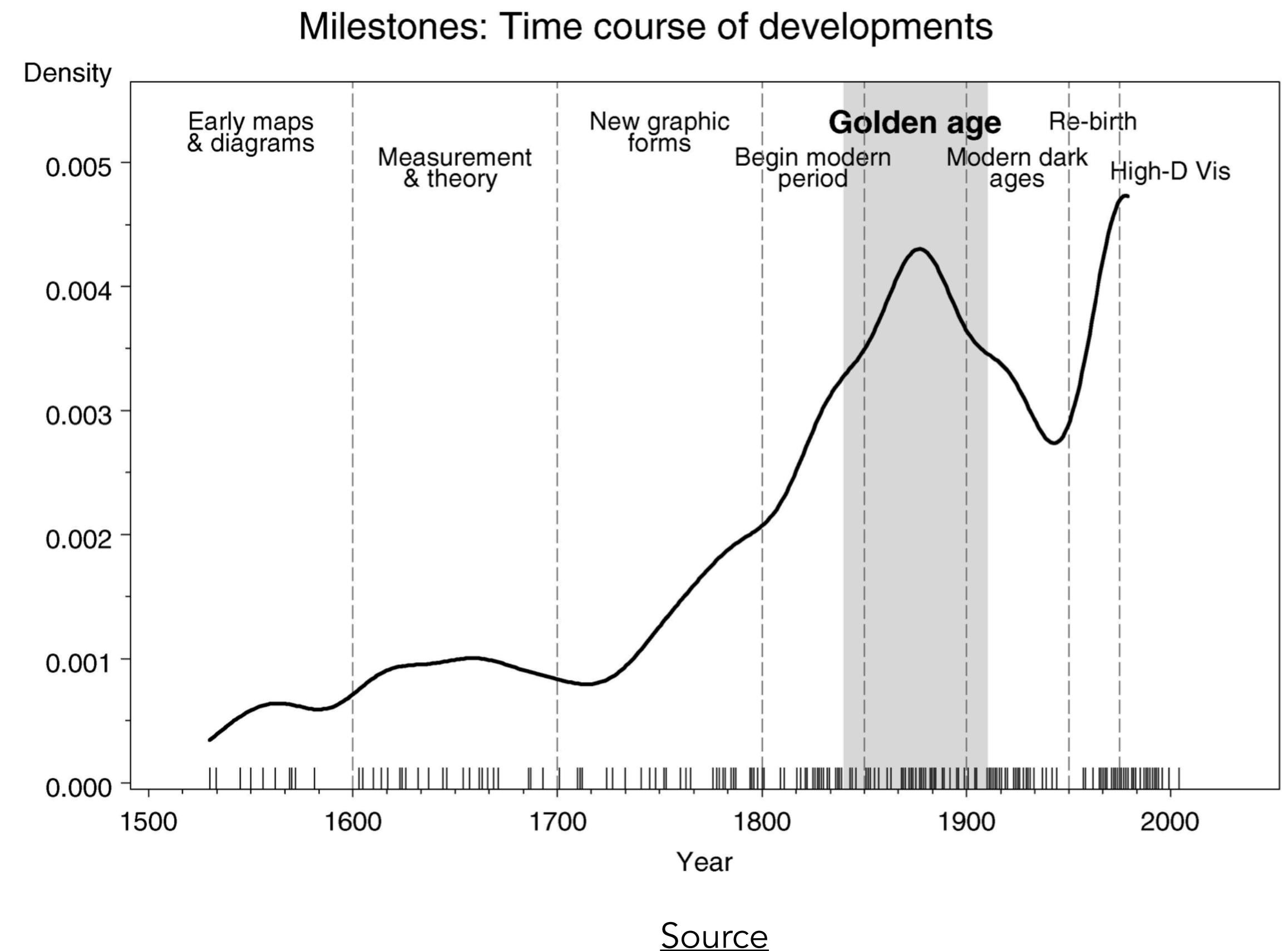
Equit. scilicet pars in qua ut
 supradicta intelligatur ad-
 sentium per figuras geome-
 tricas ostenduntur. Et ut
 omnem speciem latitudinis
 in presenti materia via oc-
 currat apparentior et studietur ad figuras geo-
 metricas applicantur. Hinc pars dividit per tria ca-
 pitula quorum primum continet dictiones. secundum 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.

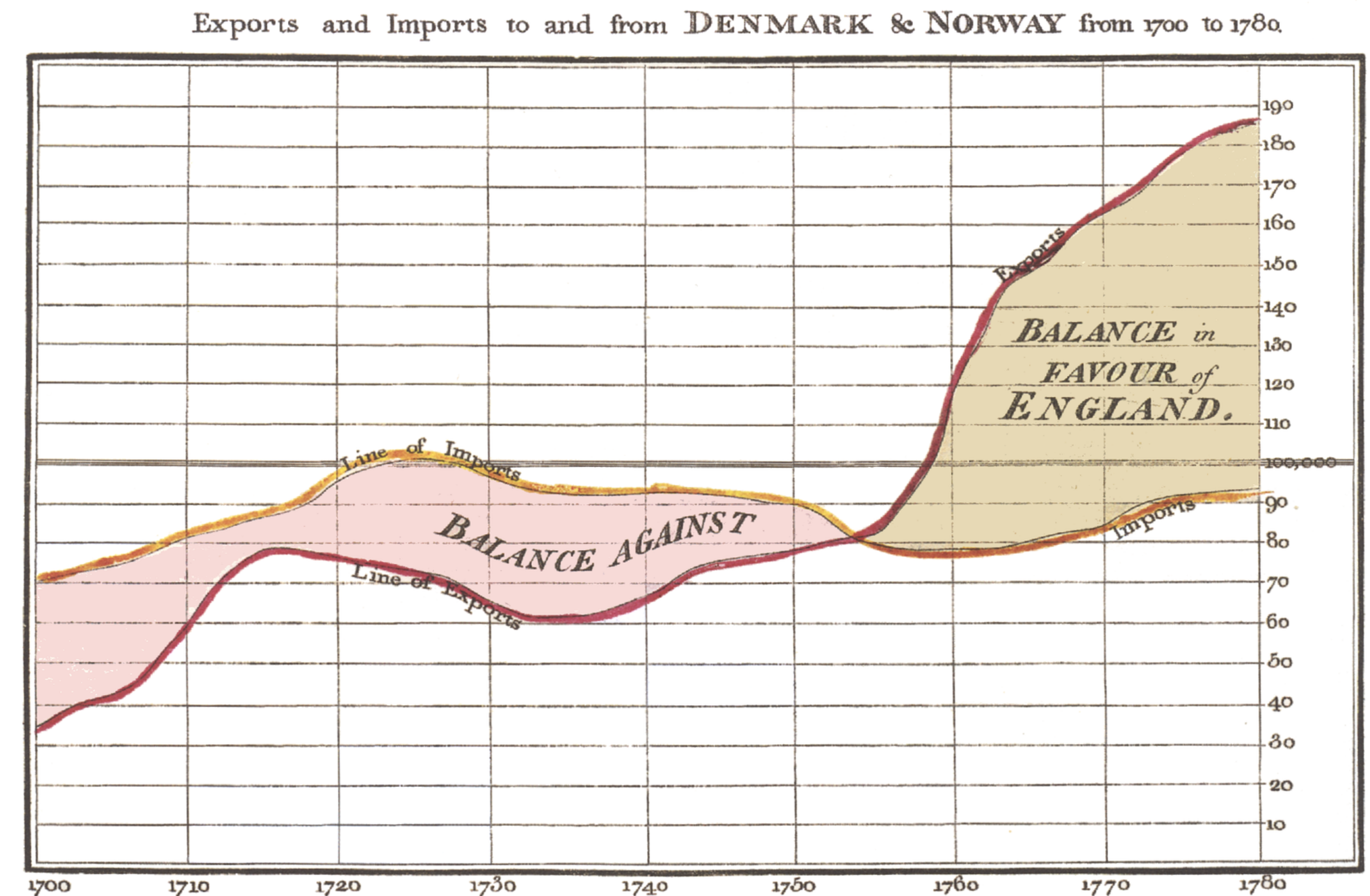
“Modern” data visualization

- 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.



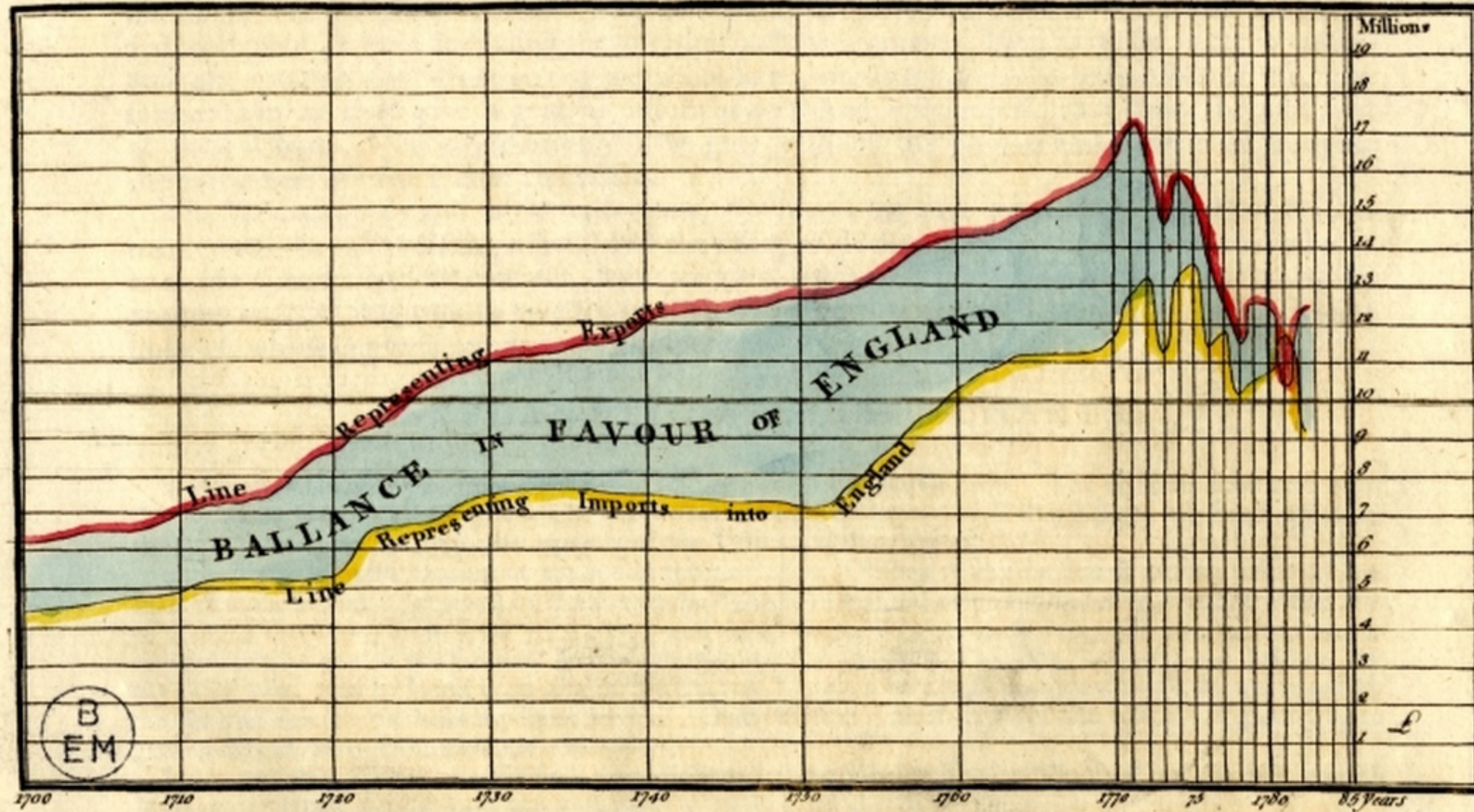
William Playfair

- 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.](#)



The Bottom line is divided into Years, the Right hand line into £10,000 each.
Published as the Act directs, 1st May 1786, by W^m Playfair
Neale sculpt 352, Strand, London.

*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

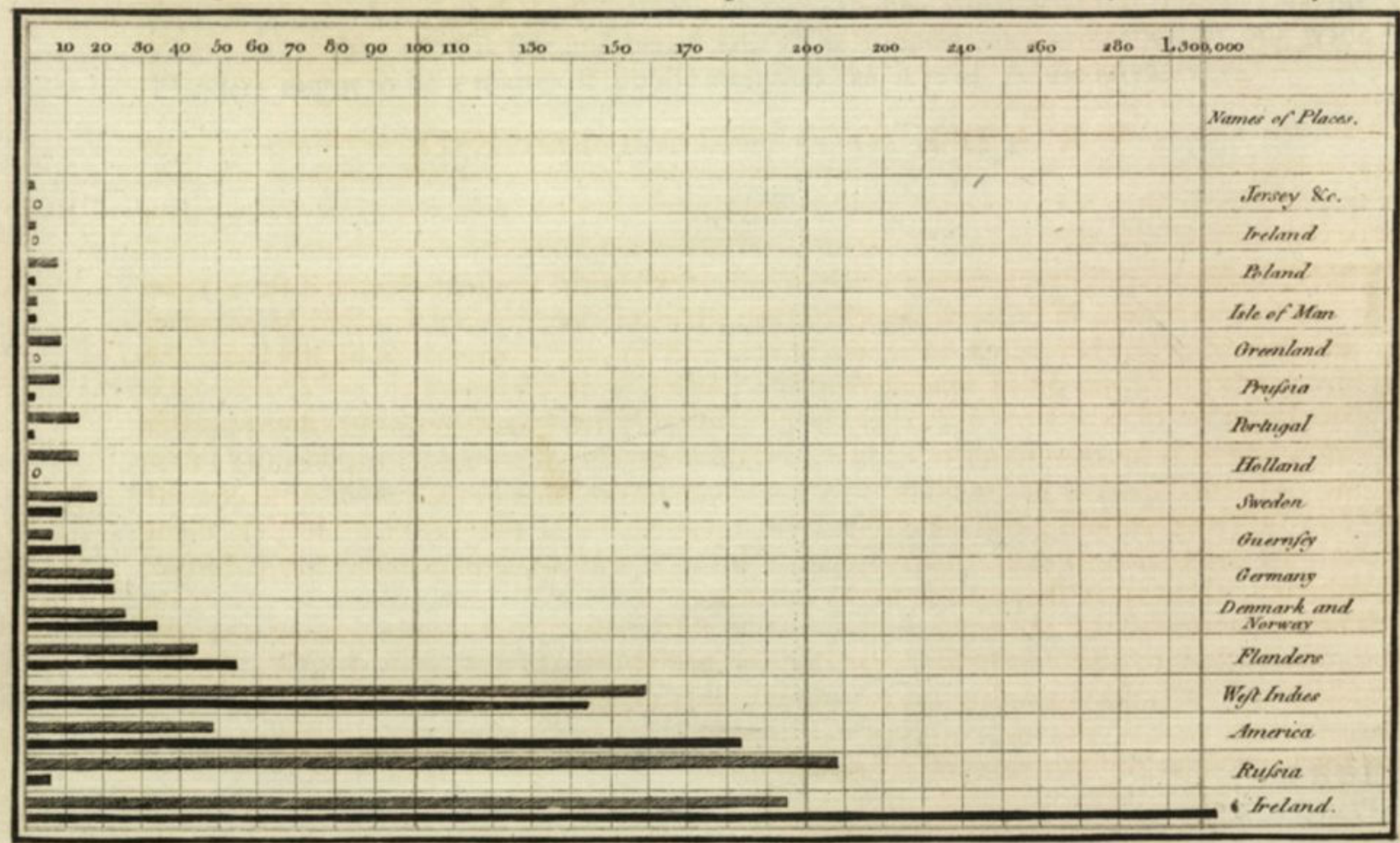
J. Smith Sculp.

to Face Page 37

Published as the Act direct. 20th Aug^r 1785

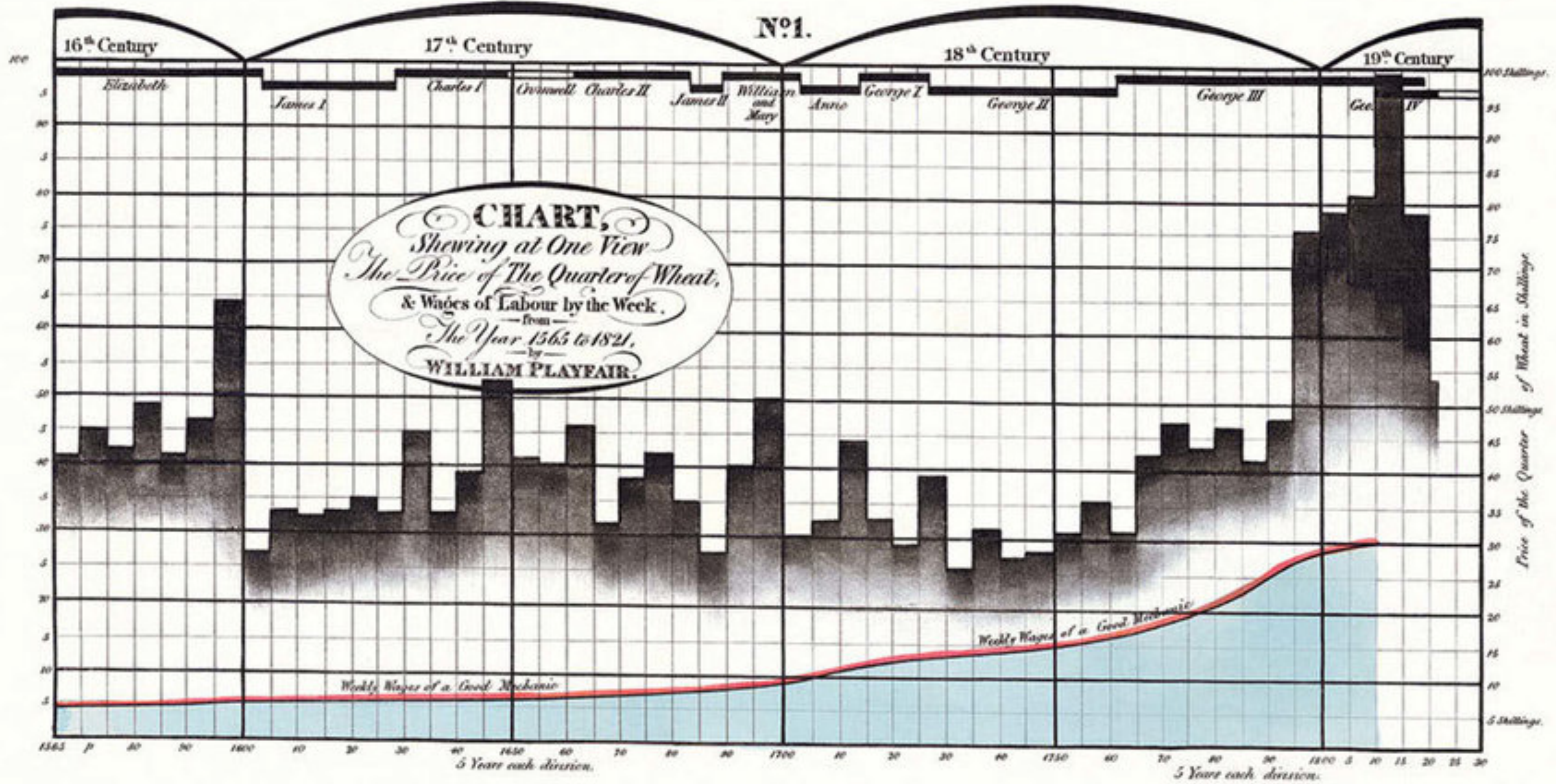
1785: Another line chart by Playfair, depicting the total imports and exports to England over a period of 85 years.

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

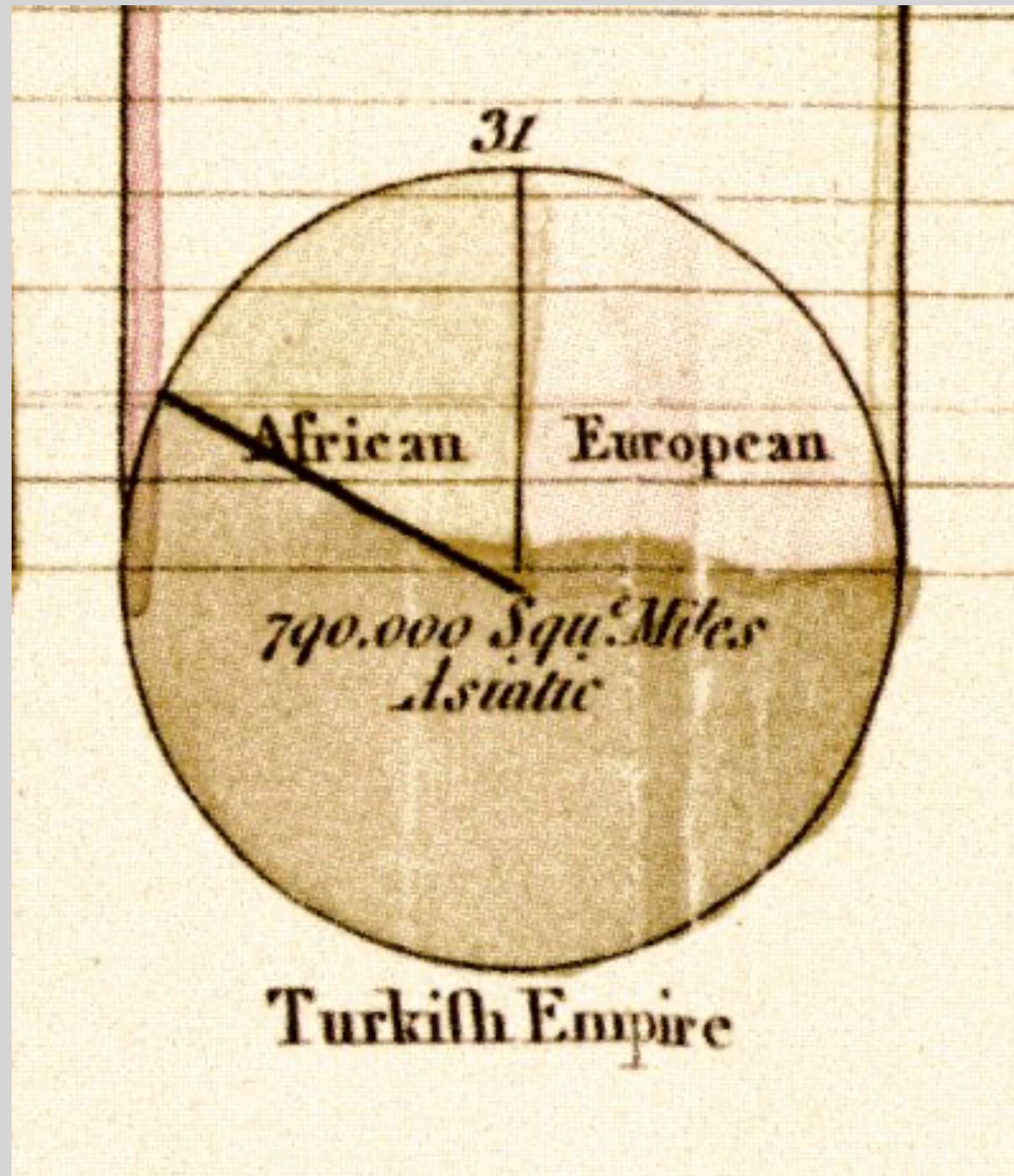


The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed lines Imports.

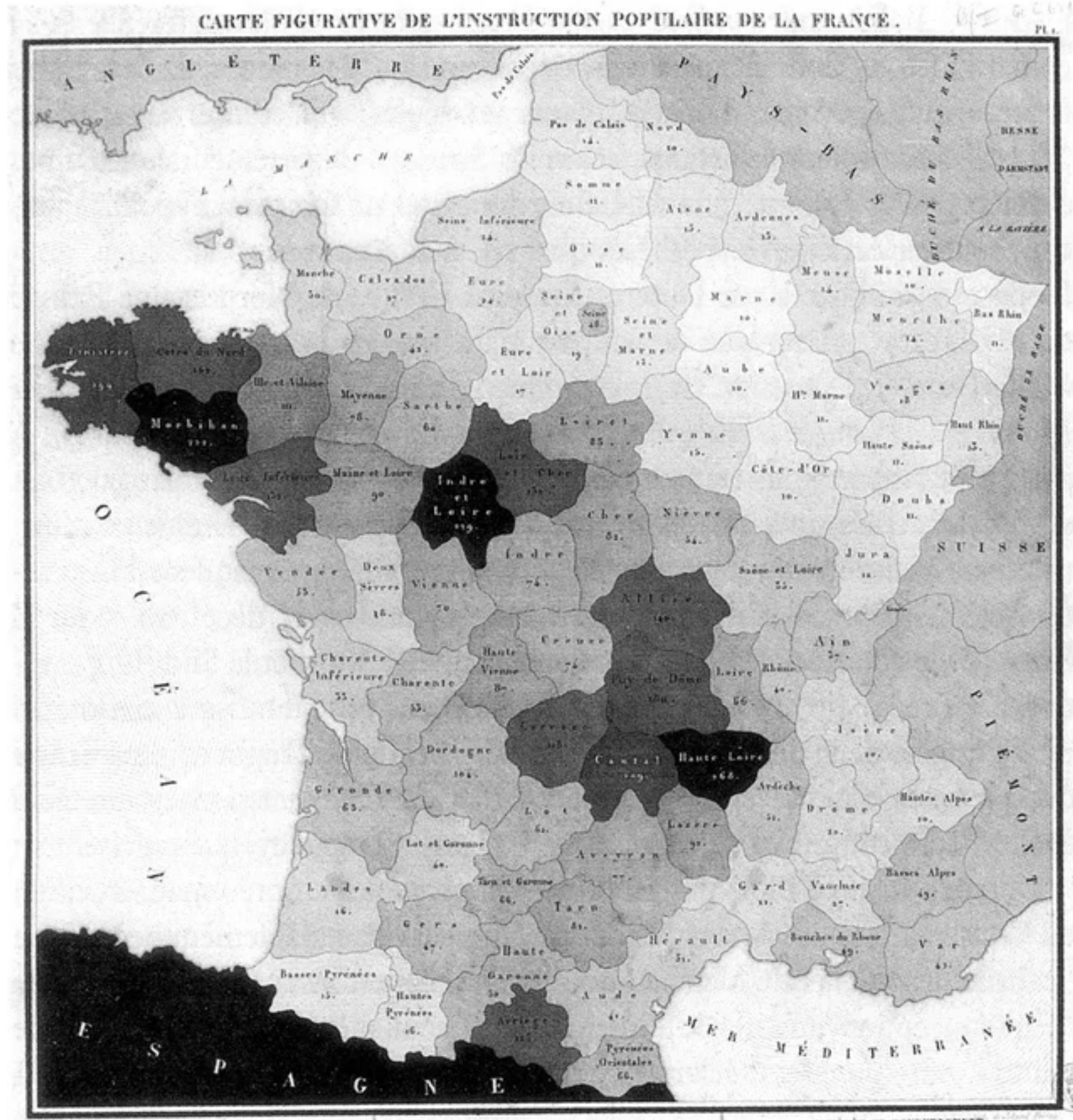
1796: The first known example of a bar chart, also by Playfair, depicting the imports and exports of Scotland to various countries in 1780.



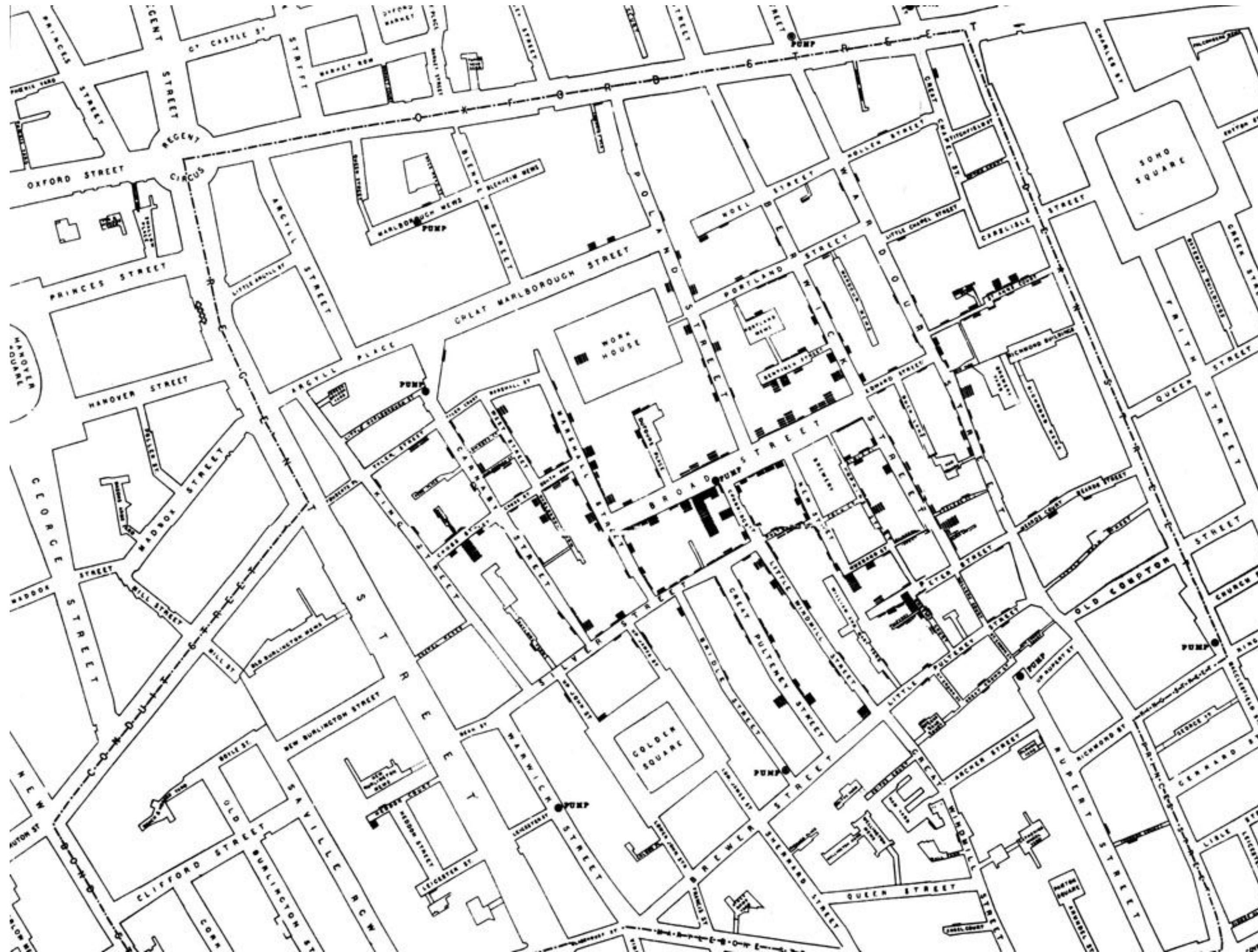
1821: Another Playfair visualization, showing the relationship between weekly labor wages and the cost of a "quarter" of wheat, along with a timeline of English monarchs, from 1565 to 1821.



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



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.

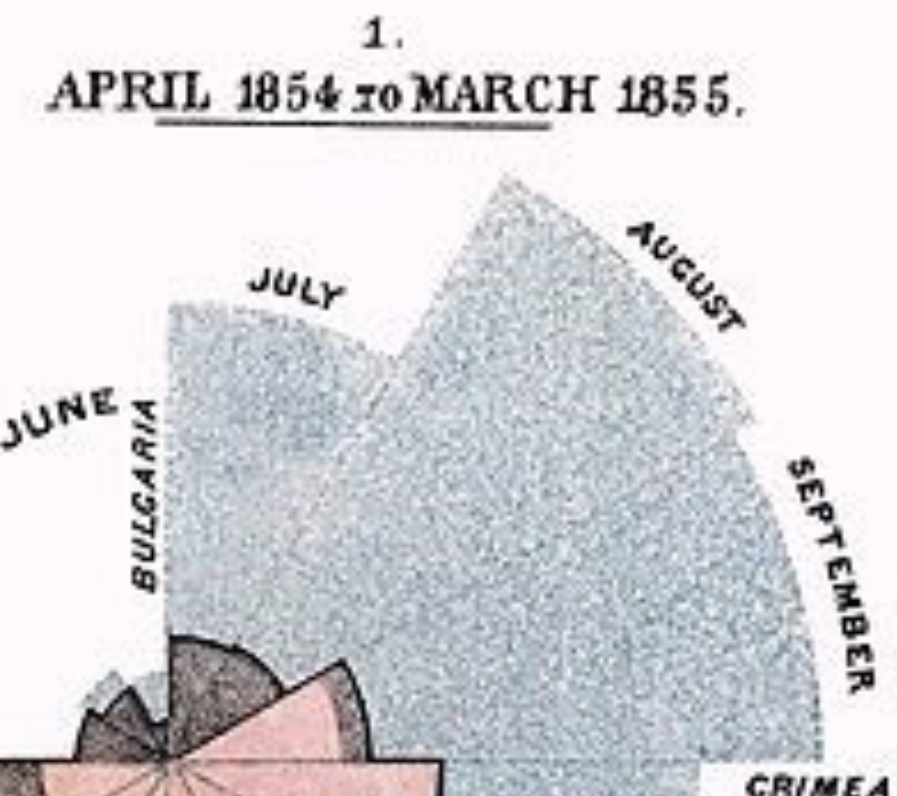
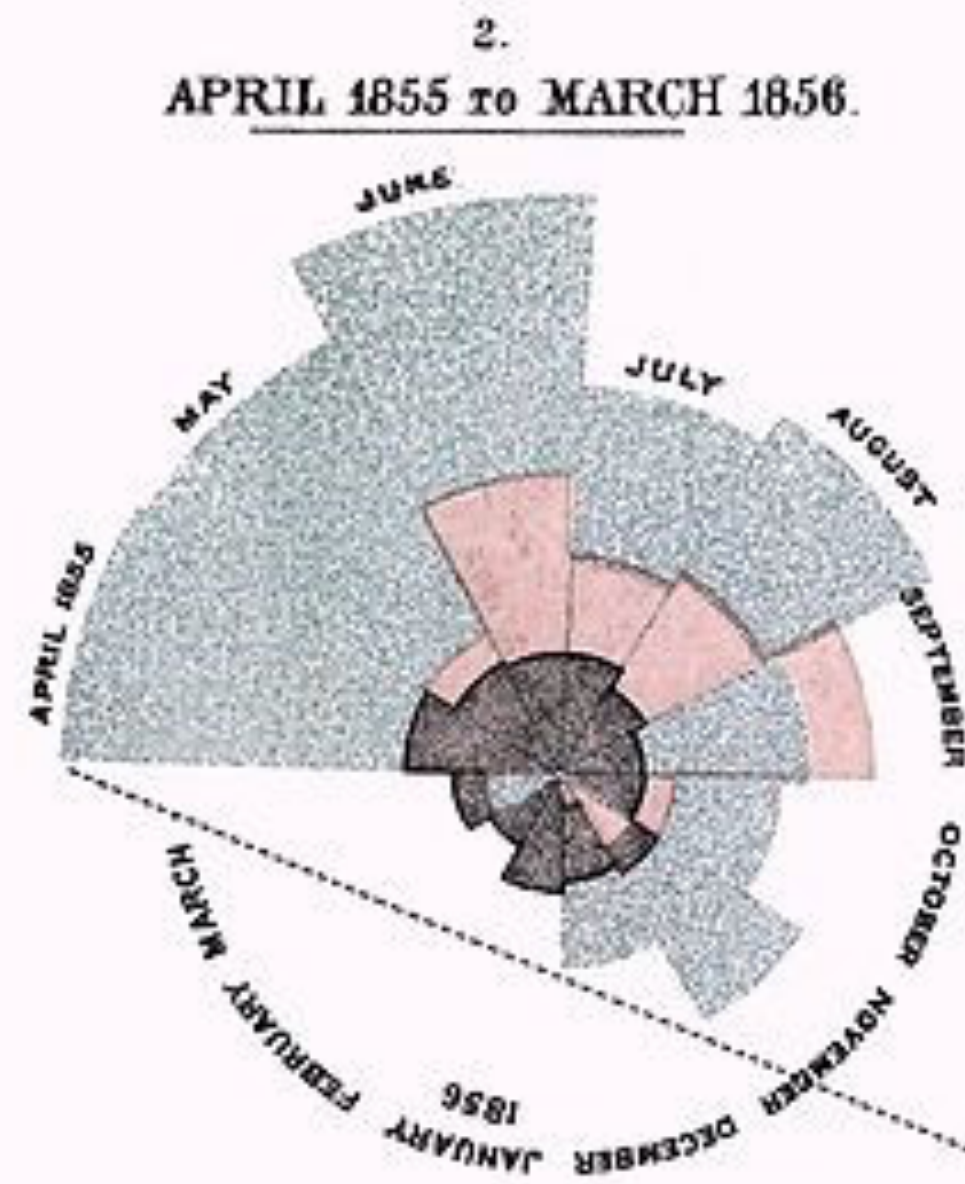


1854: John Snow mapped cholera deaths in SoHo, London. He noticed that many deaths were clustered around the Broad Street pump.



2020: The site of the Broad Street pump.

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



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^r 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.

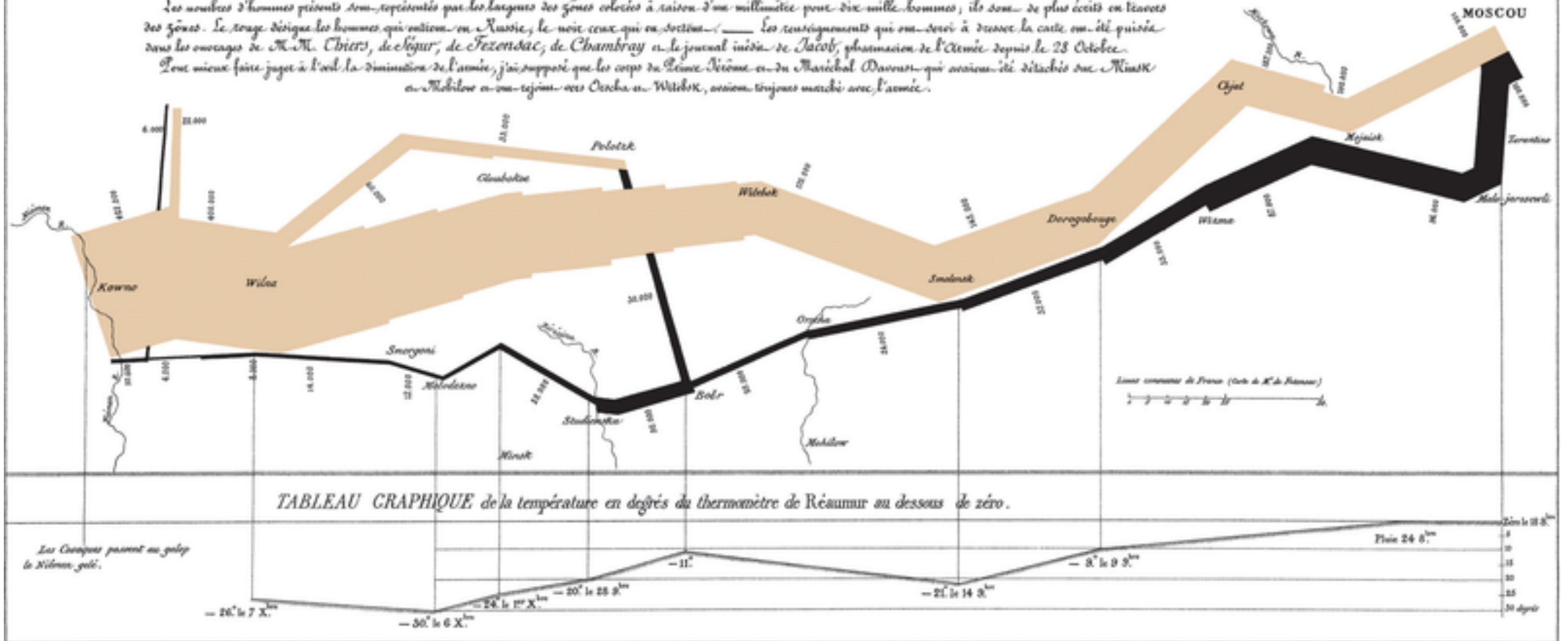
The entire areas may be compared by following the blue, the red & the black lines enclosing them.

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.

Devisé par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite. Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs 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 entrent en Russie, le noir ceux qui en sortent. Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Légar, de Fozensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 23 Octobre. Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supprimé que les corps du Prince Nicôme et du Maréchal Davoust qui avaient été détachés sur Minsk et Mohilew et qui rejoignent l'armée à Wilna, ainsi toujours marchés avec l'armée.

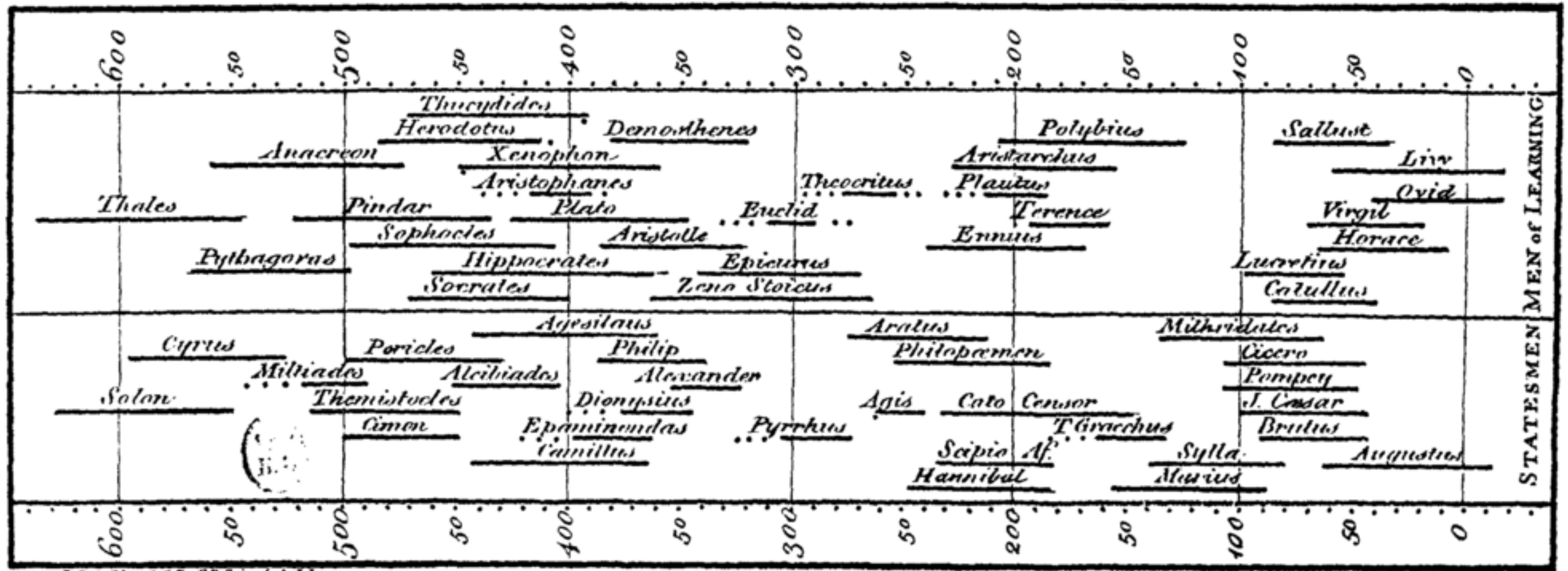


Atty. par Regnier, à Paris 27 Mars 27 077 à Paris.

Imp. Lit. Regnier et Borellet.

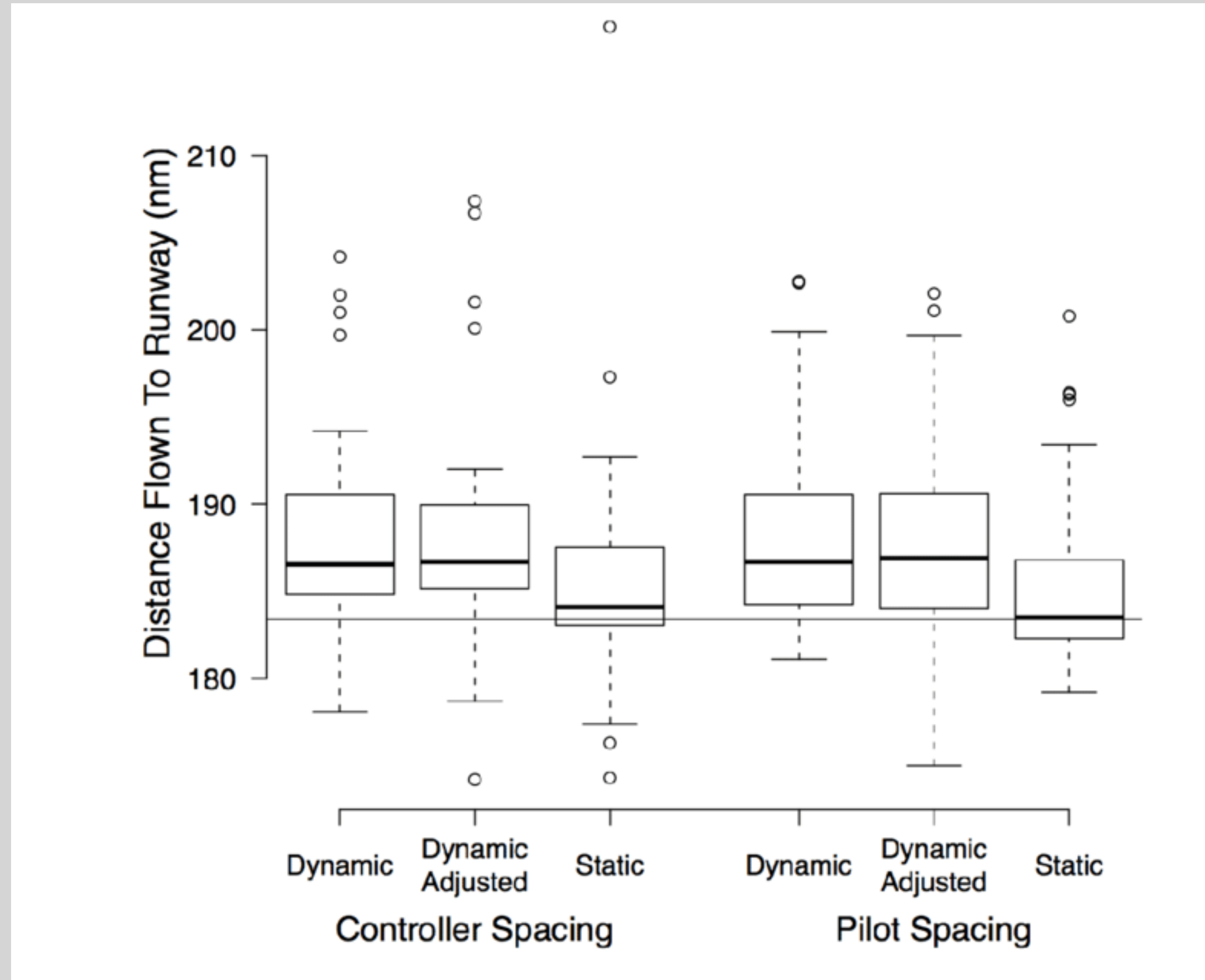
1869: Charles Joseph Minard's visualization of the French invasion of Russia (led by Napoleon).

A Specimen of a Chart of Biography.



J. Priestley L.L.D. F.R.S. scilicet del.

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!