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)!**
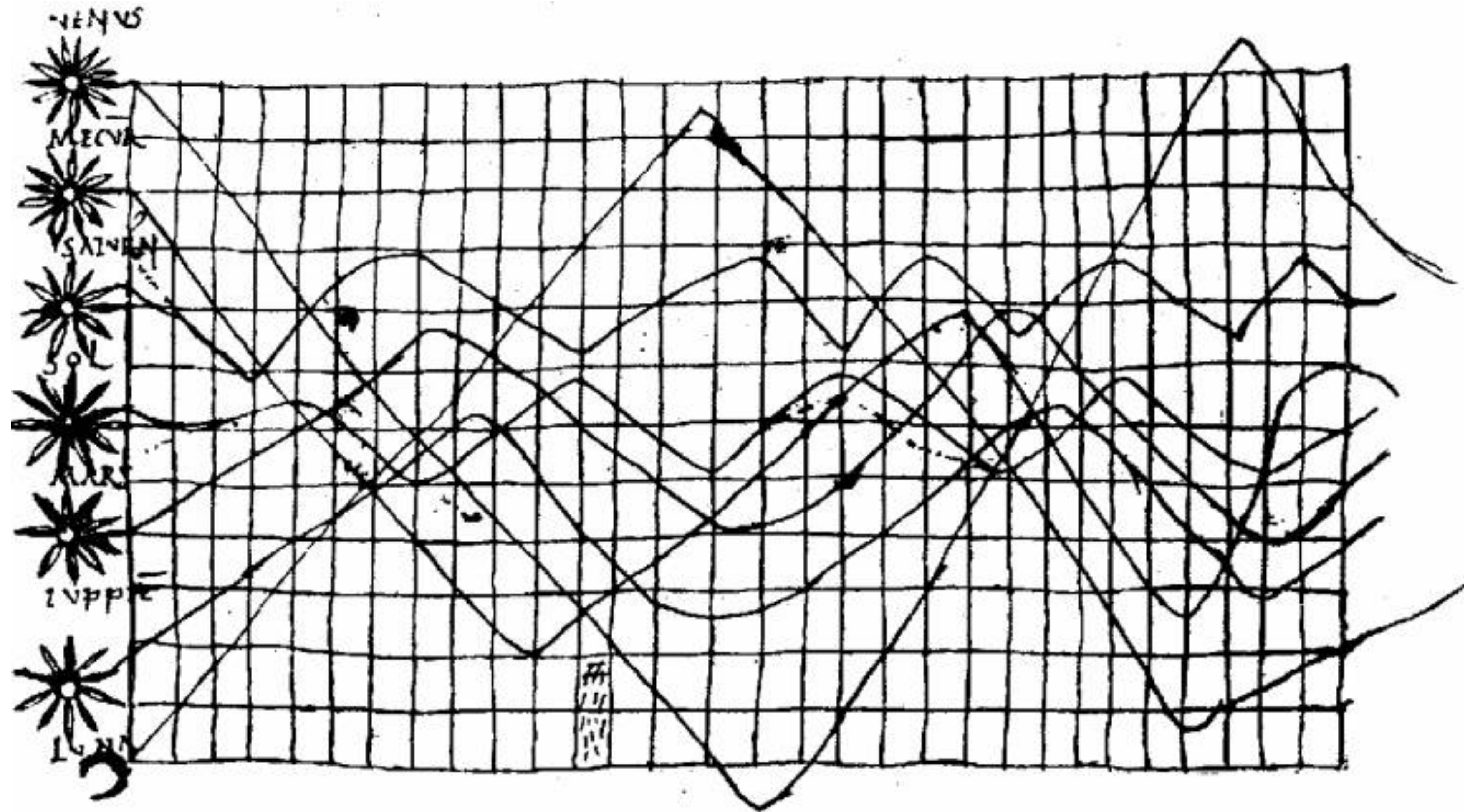
Agenda

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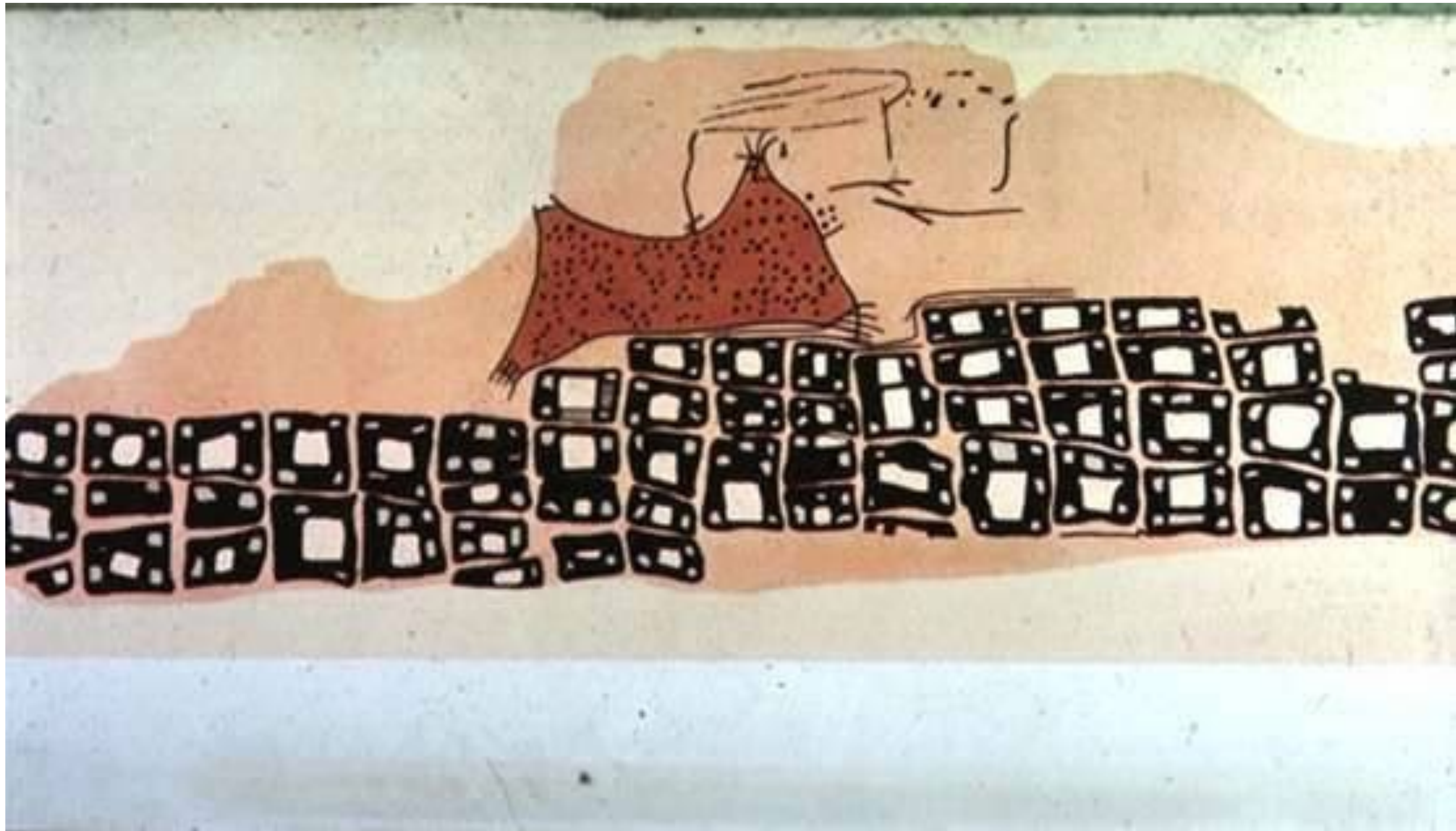
6200 BC: A map depicting the town of Konya, Turkey.

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

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

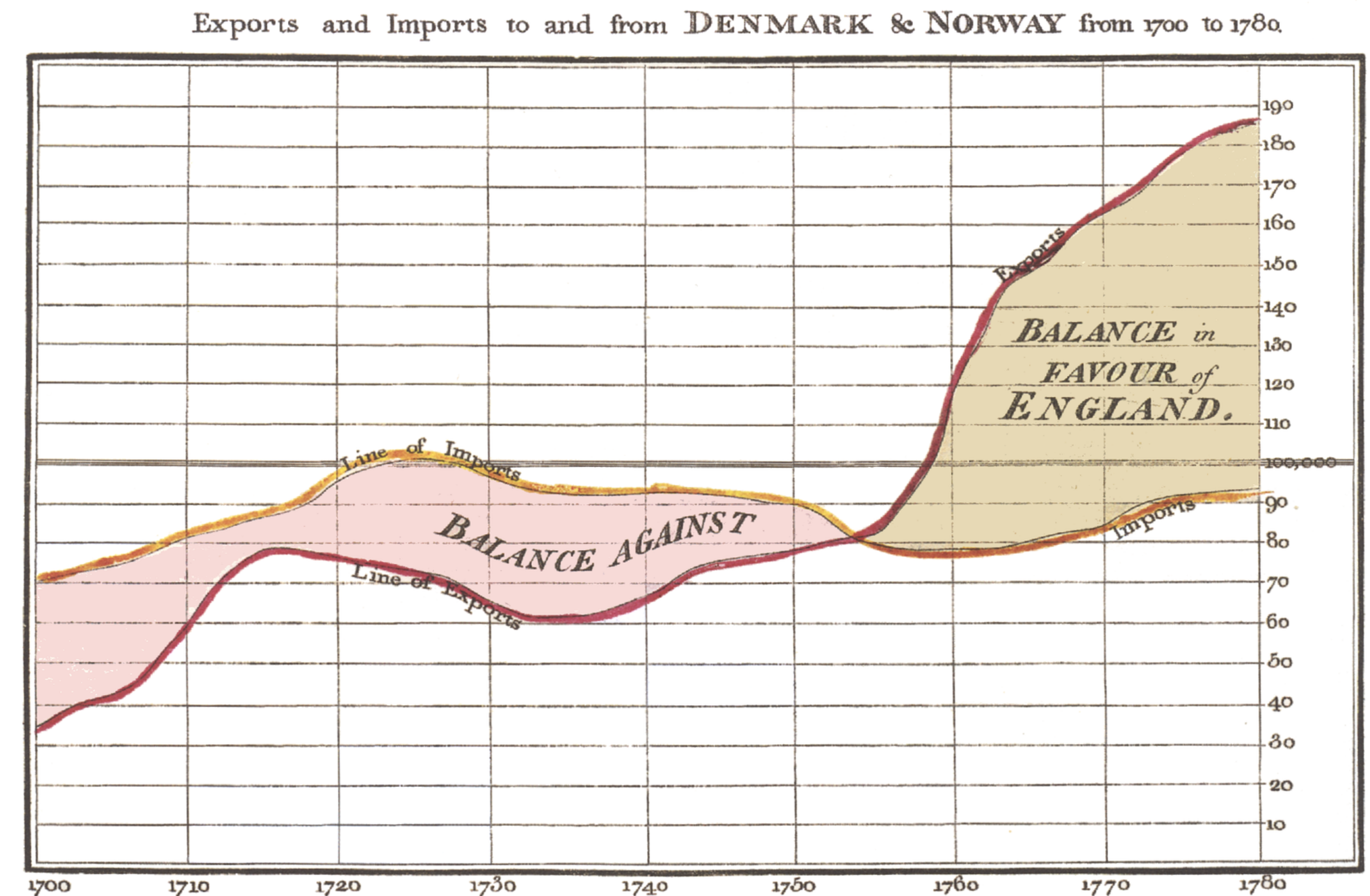Today, we would use line charts or scatter plots to show the same information,.

"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.

Milestones: Time course of developments
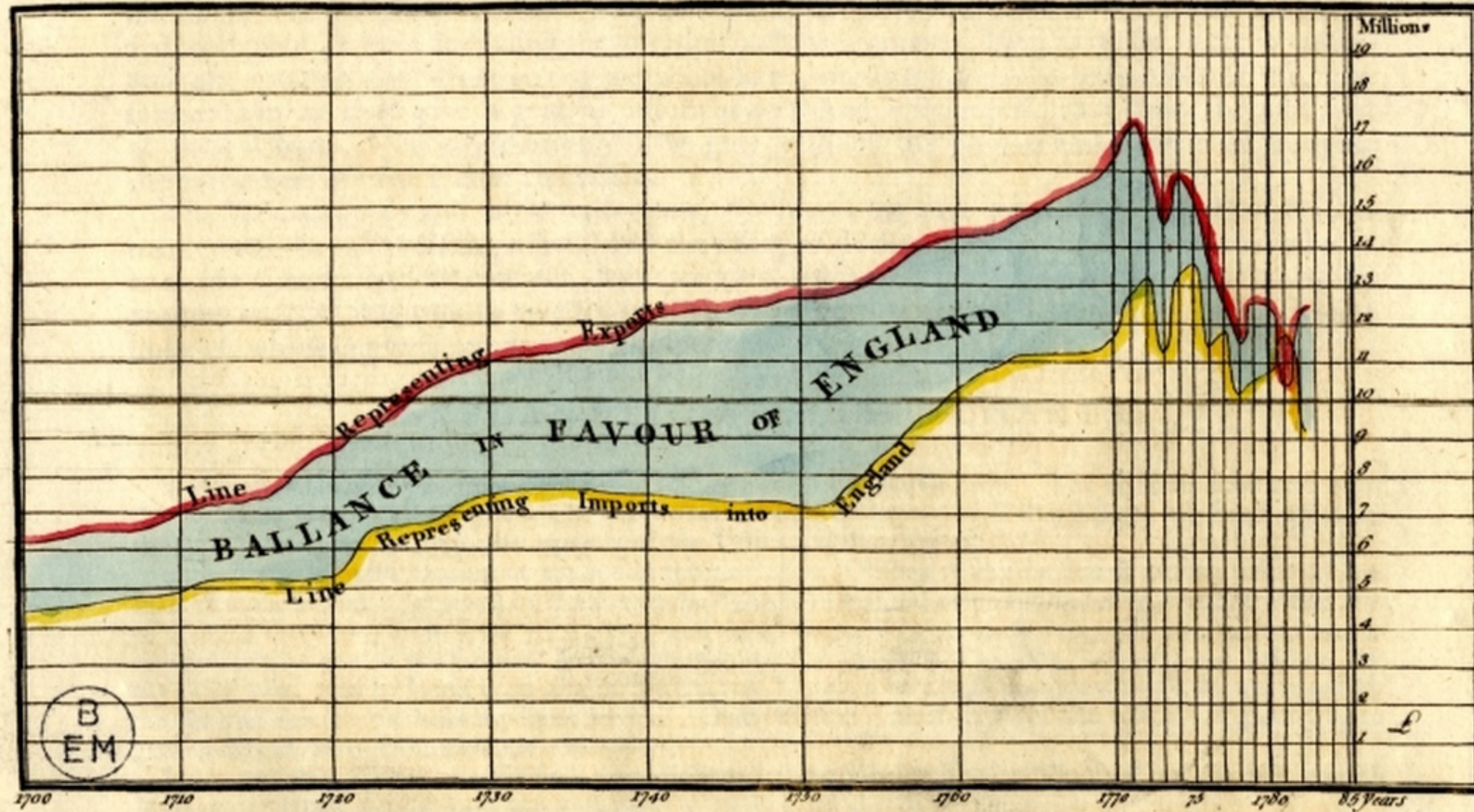
Source
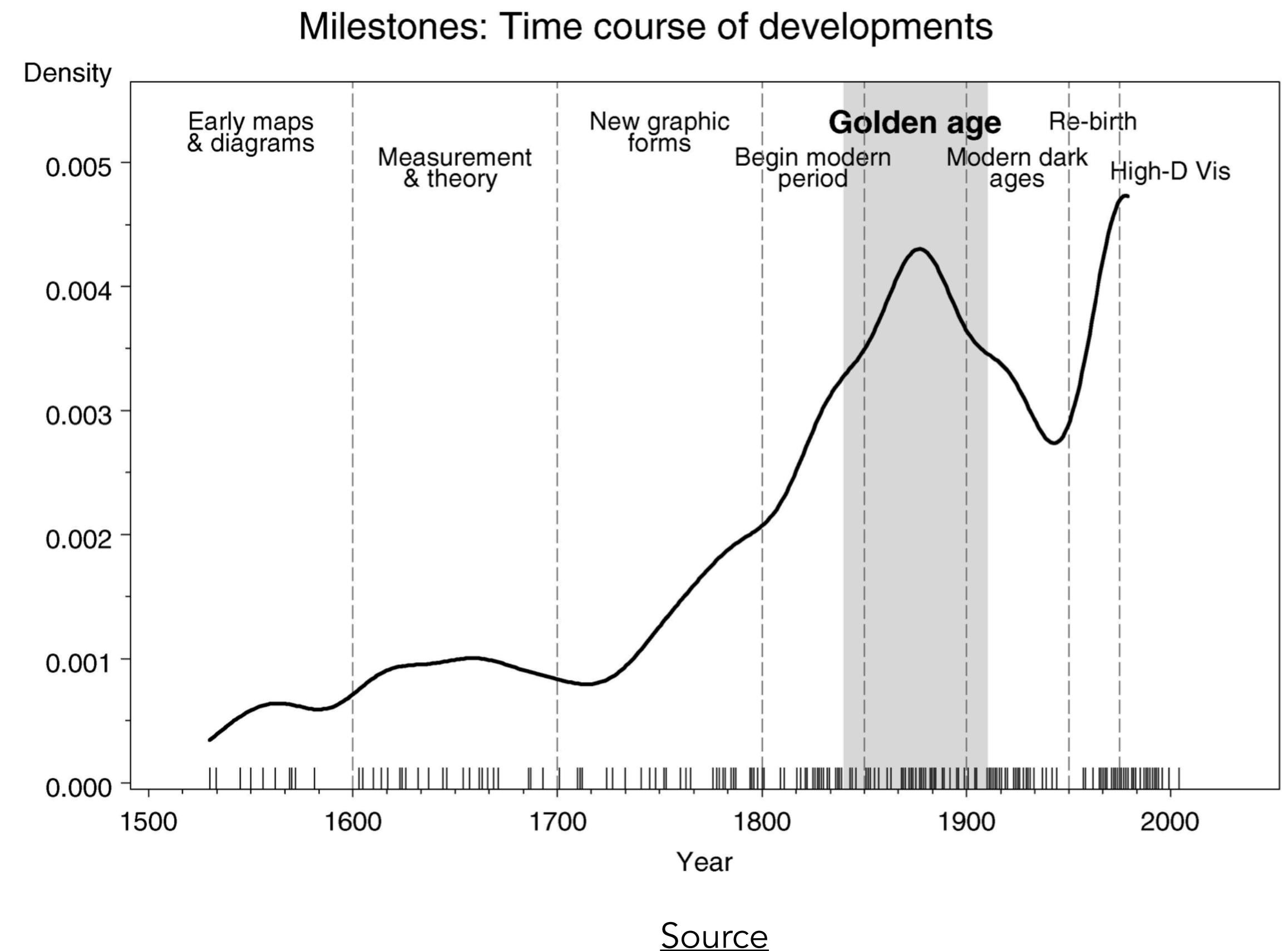
William Playfair

- 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.

Exports and Imports to and from DENMARK & NORWAY from 1700 to 1780.

BALANCE in FAVOUR of ENGLAND.

Line of Imports

BALANCE AGAINST

Line of Exports

The Bottom line is divided into Years, the Right hand line into £10,000 each.

Published as the Act directs, 1st May 1786, by Wm Playfair

Neele sculpt 352, Strand, London.

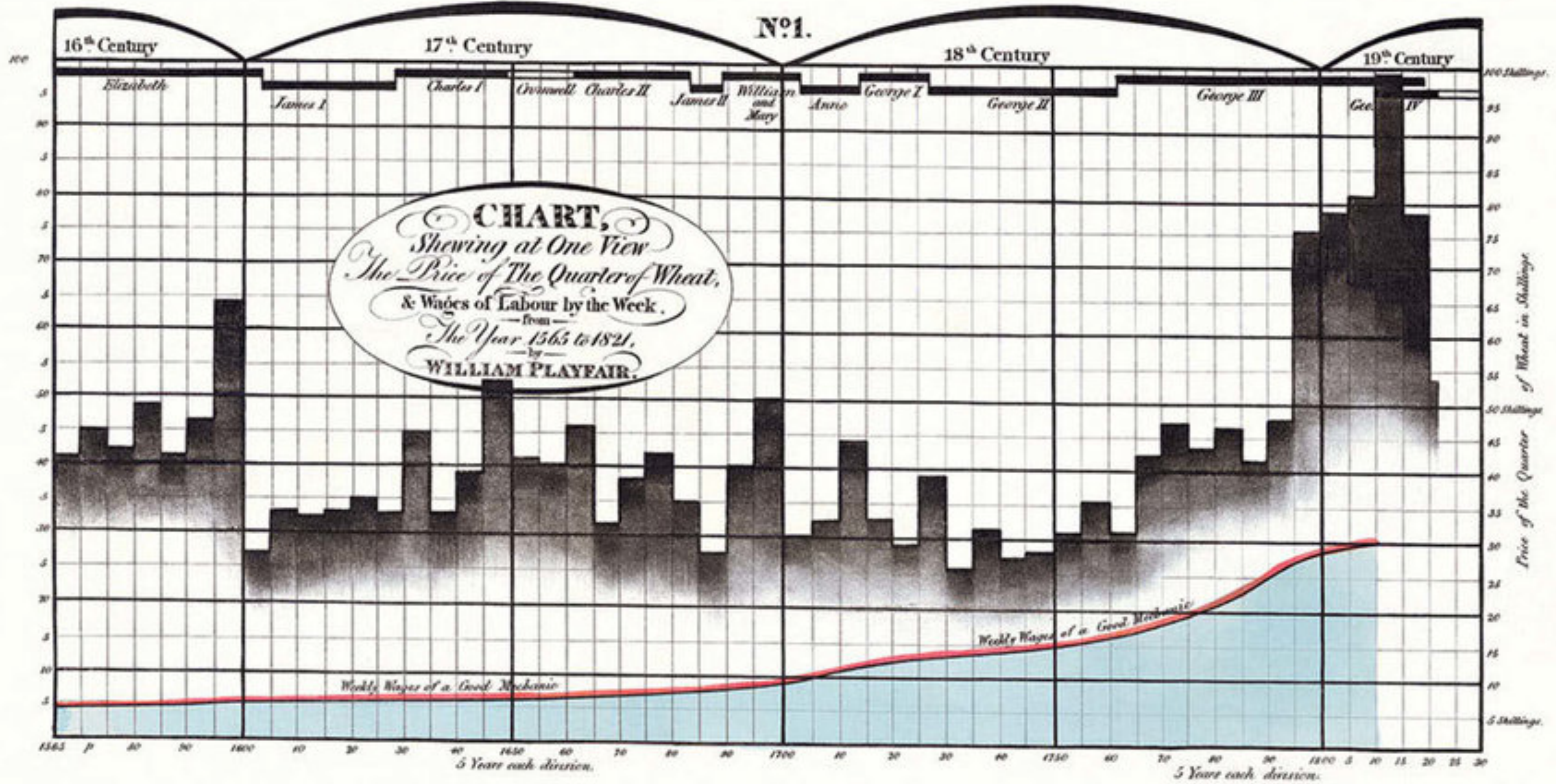**1785:** Another line chart by Playfair, depicting the total imports and exports to England over a period of 85 years.
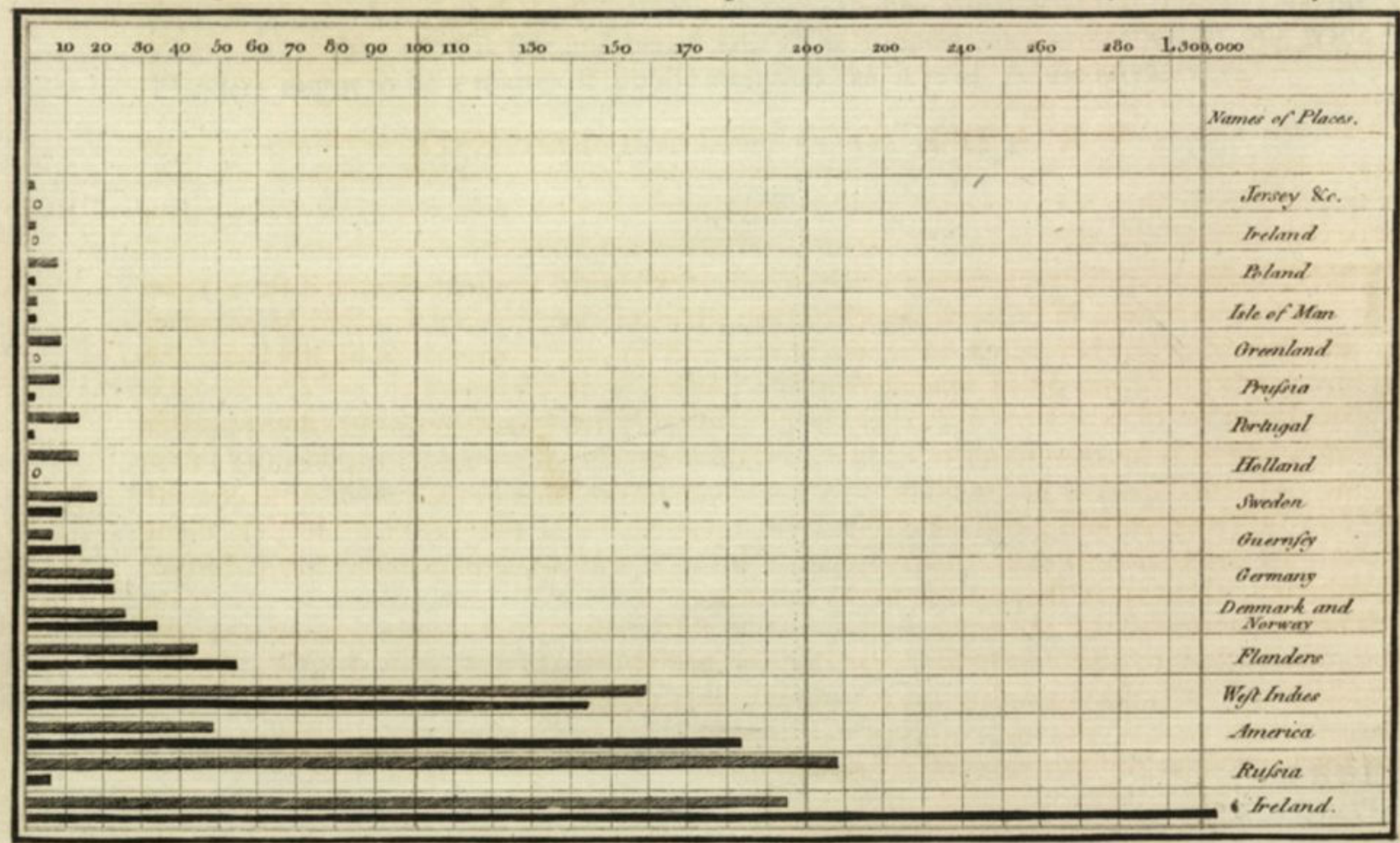
1796: The first known example of a bar chart, also by Playfair, depicting the imports and exports of Scotland to various countries in 1780.

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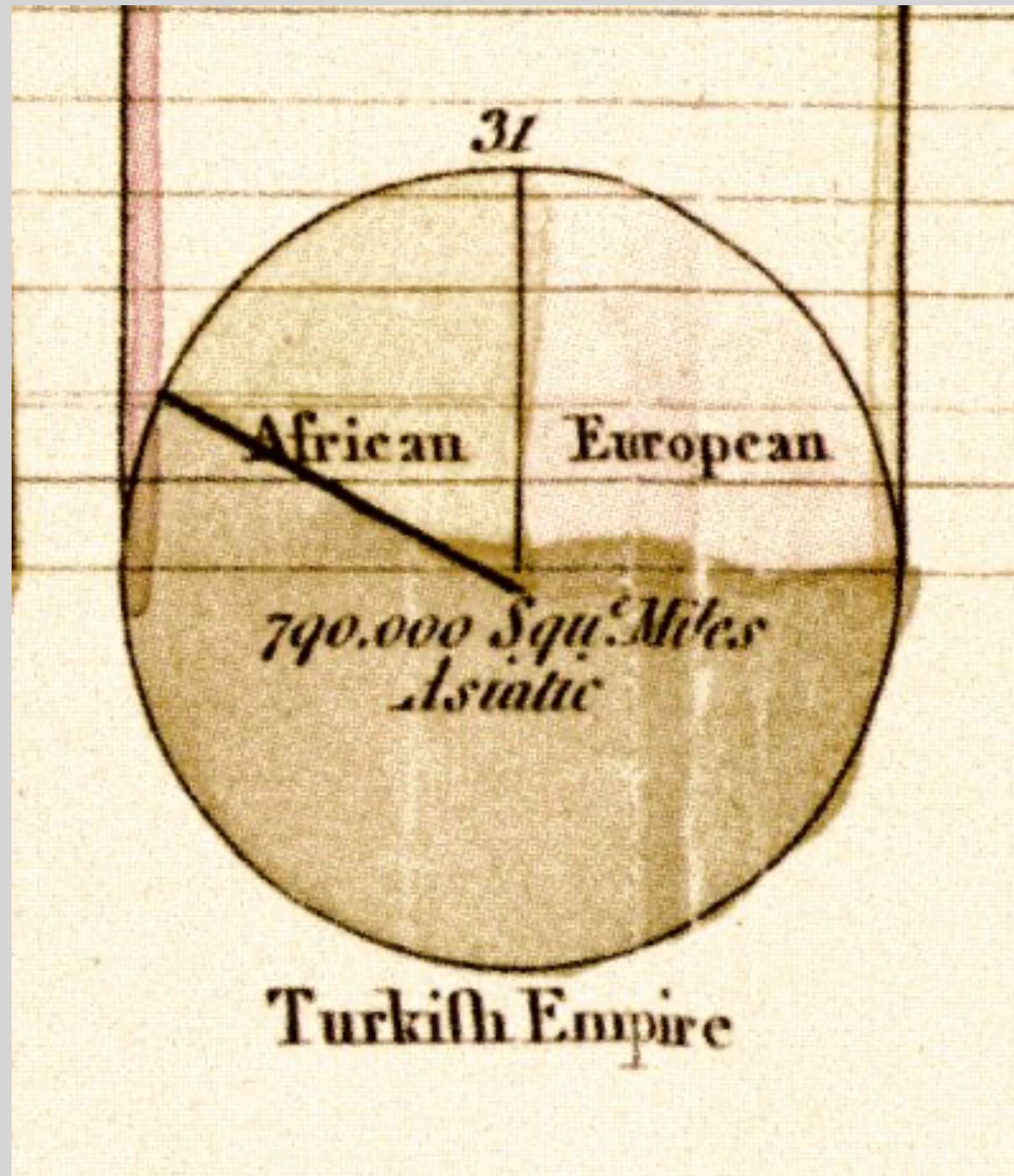
1801: Playfair's pie chart depicting the distribution of the Turkish Empire.

CARTE FIGURATIVE DE L'INSTRUCTION POPULAIRE DE LA FRANCE.
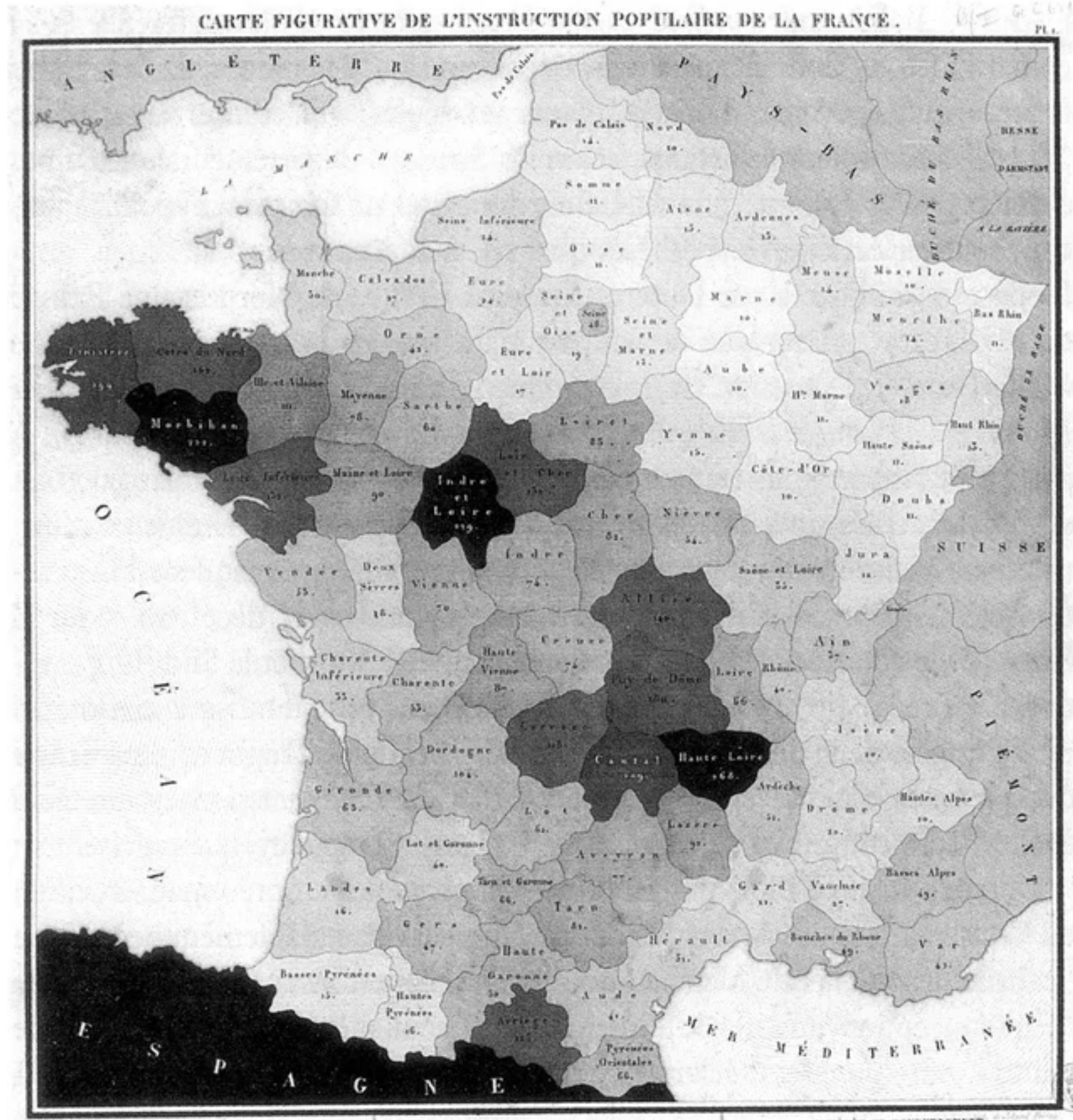
1826: Charles Dupin creates a choropleth, which describes the distribution of some quantity for each of several physical regions. His choropleth depicted rates of literacy in different parts of France.
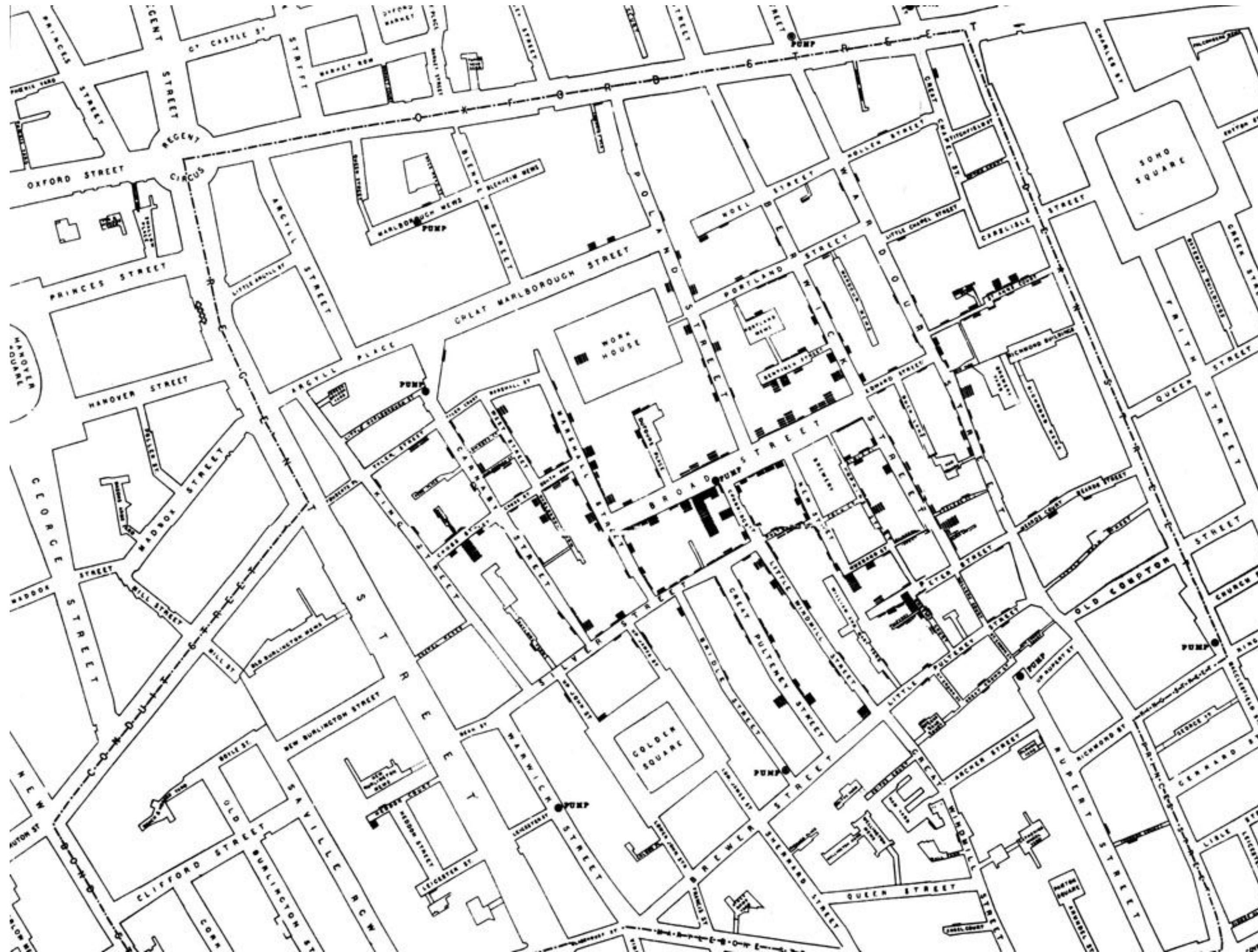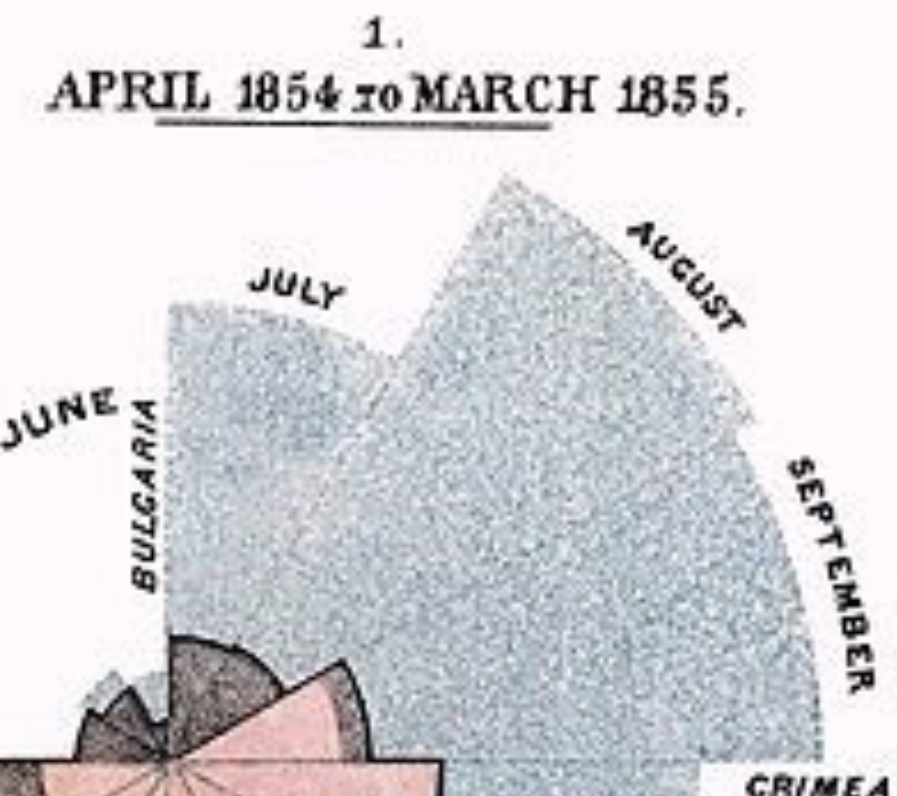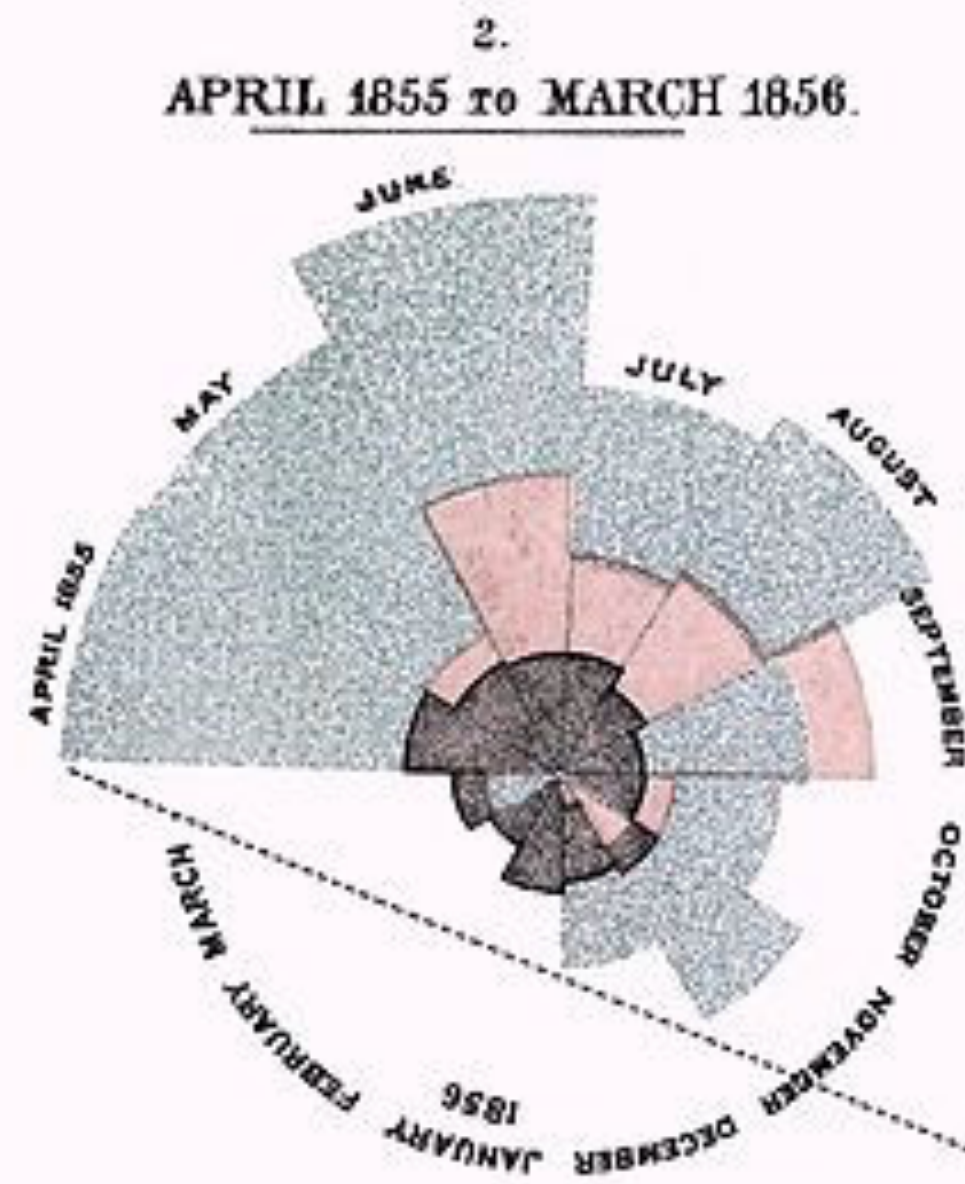
1854: John Snow mapped cholera deaths in SoHo, London. He noticed that many deaths were clustered around the Broad Street pump.

2020: The site of the Broad Street pump.

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.

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

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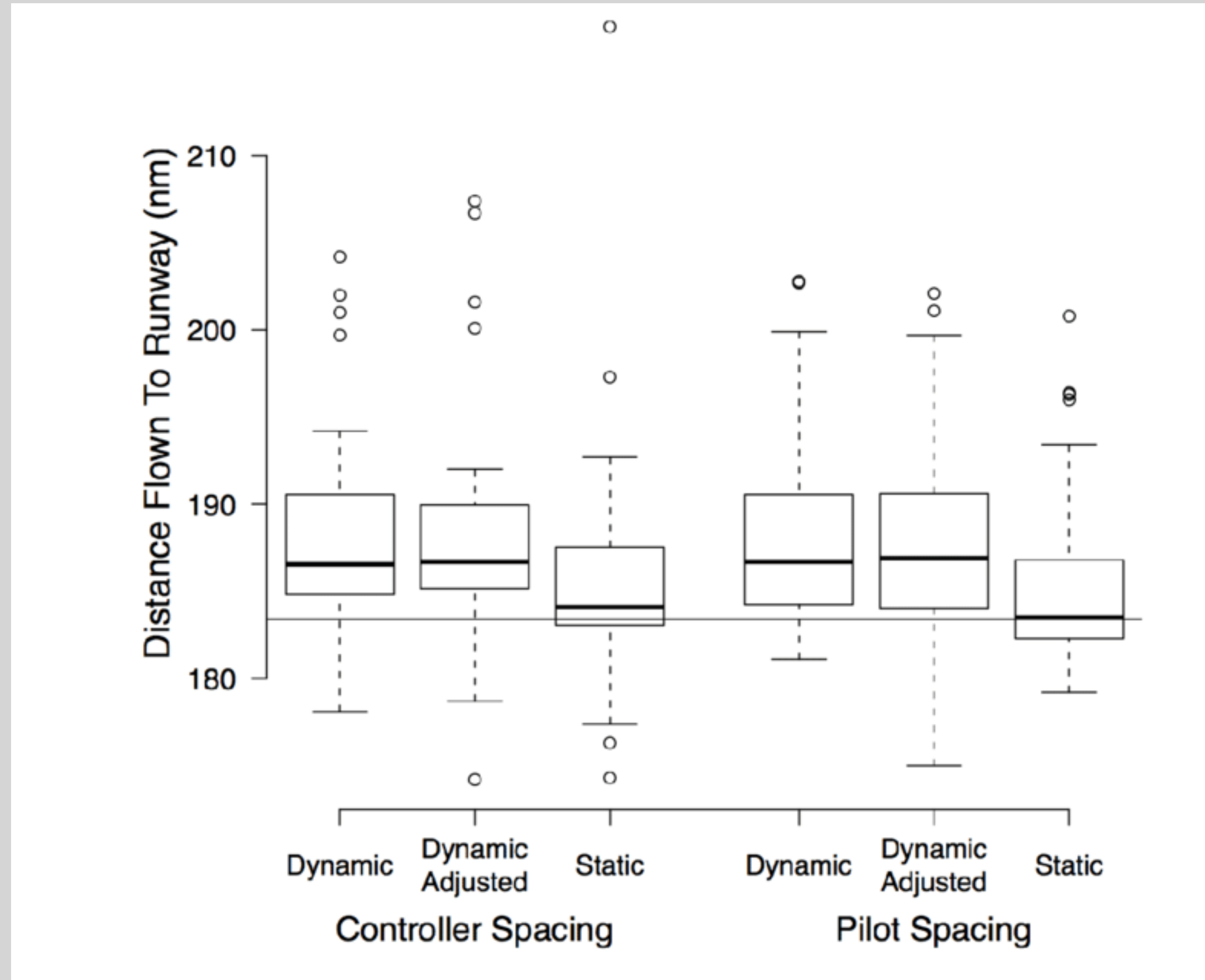
1973: John Tukey, who defined the term "Exploratory Data Analysis", created the box plot, which describes a numerical distribution using a 5 number summary.

That's all!