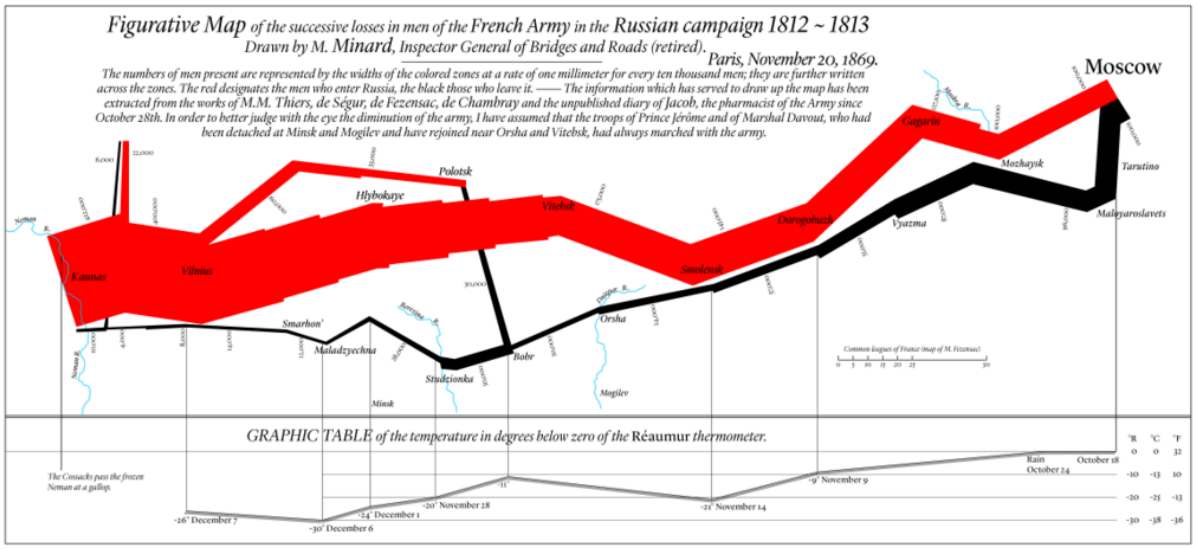
Homework 2

Carlos Rubio

January 20th, 2025

Part 1

The following is a reproduction translated into English of a diagram created by Charles Minard in 1869to visualize Napoleon's army during its 1812 invasion of Russia.



- What data values are encoded in this visual? There are at least 6.
- Describre how each value is presented visually. What visualization techniques

are used to distinguish data values?

- What is the story being told in this figure? Is the figure exploratory or

explanatory?

Answer

Some of the dimensions of the data presented include

1. Amount of men entering Russia
2. Amount of men leaving Russia,
3. Location
4. Date
5. Temperature at the time and place of the battle (in Celsius,

Farenheit or Reaumur)

1. distance traveled

The amount of man at given point is presented as two bars with varying width along time, one including all the men joining the French army during the campaign at a given point, the other one including all the men leaving.

The diagram represents the course of the march and the places transited. The diagram also includes a line chart underneath recording the temperaature at different points in time of the campaign.

It’s obvious from the diagram that, as the campaign advanced, the increase in numbers due to reinforcements droped. Today, it is accepted as history that the France-Russian invasion was aborted because of the extreme temperatures. But from this diagram, as we contemplate the history of the temperature the army was subject with the numbers of the army, there is a negative correlation between the two dimensions.

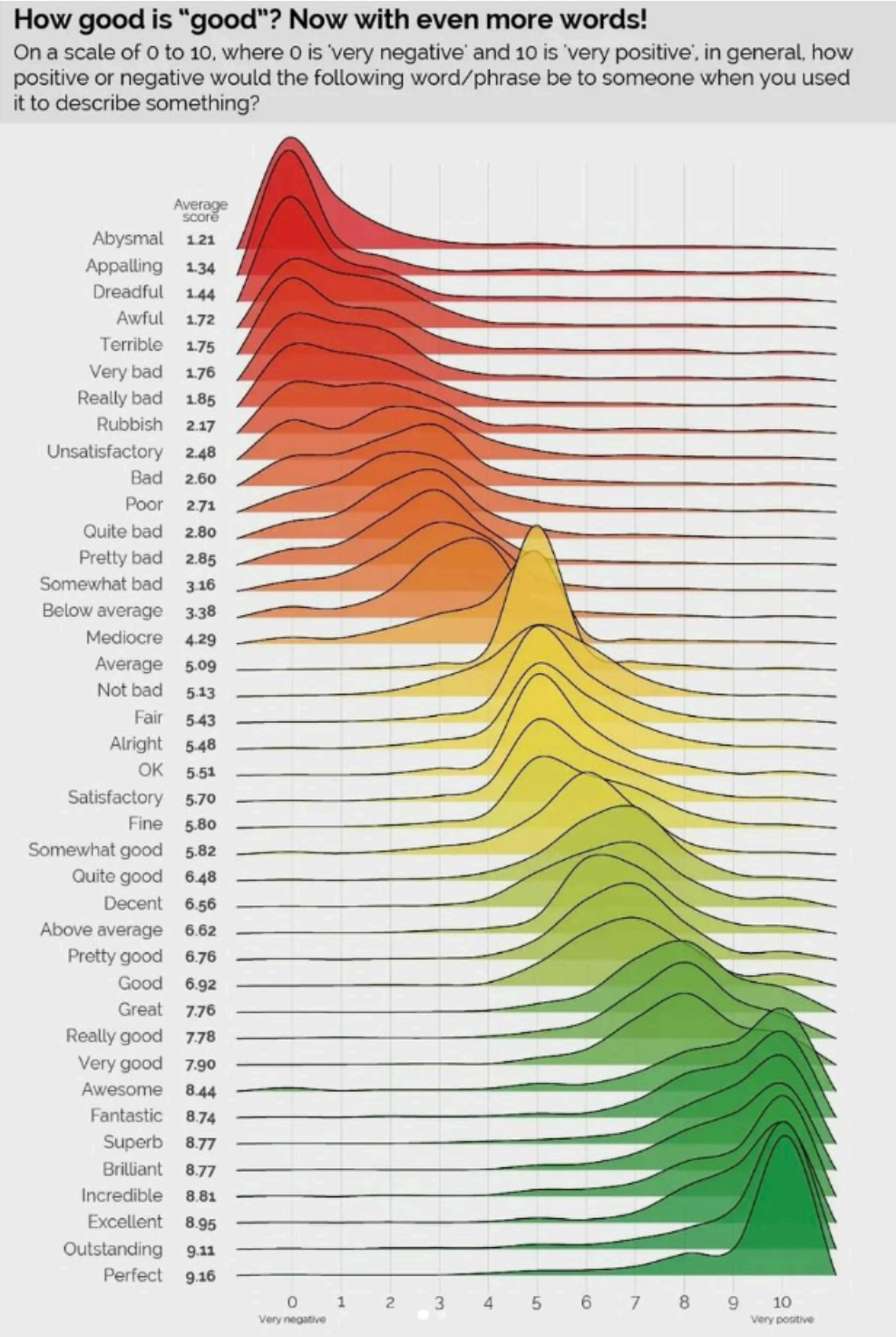
Part 2

Here are some additional figures. For each of the following, state

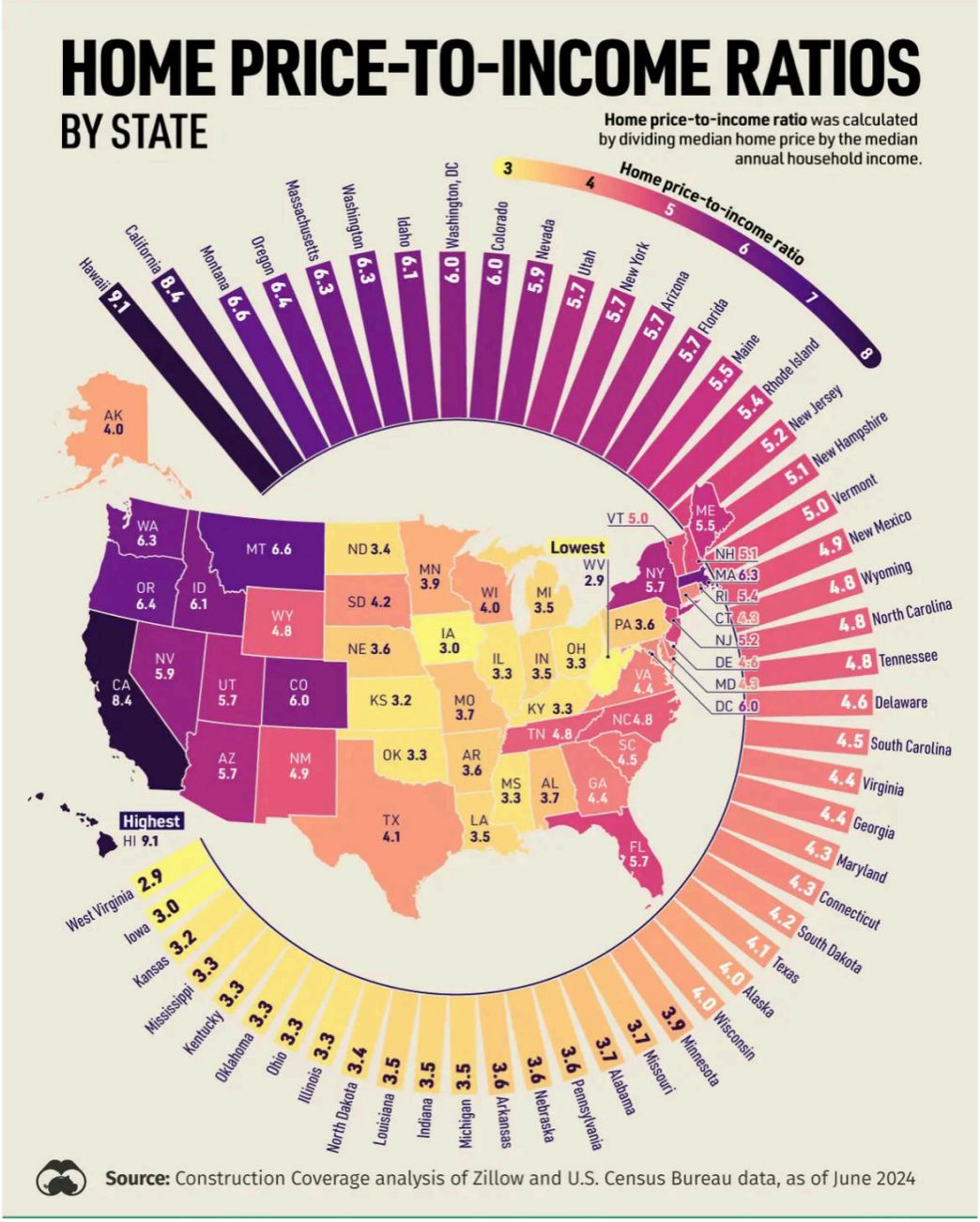
1. What data values are encodeed in the figure.
2. What visualization techniques are used to present the data.
3. Whether the figure is explonatory or explanatory.

Answers

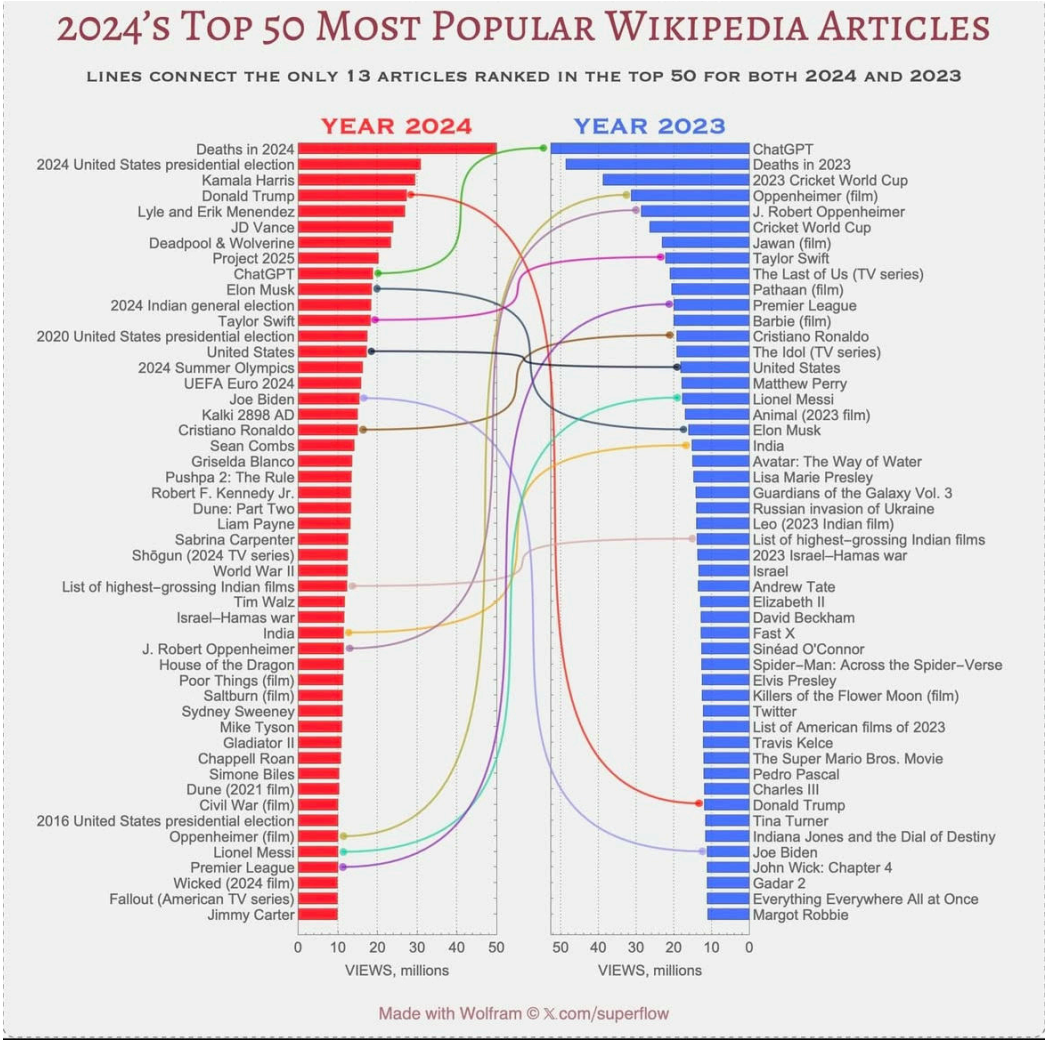
(Next page)



The data encoded is the distribution of scores of how "positive/negative" of the a given word is, as well a its average. The data is presented as a series of line charts with colored areas in a range from red to green, where red indicates negativity, and green positivity. The figure is mostly exploratory: there is no interpretation on the title or the accompaning text.



The data presented encodes the distribution of home price median-to-income median ratio among the fifty states. It includes two different visualizations: a heatmap of the United State, and a radial bar chart sorted decreasingly. Both visualizations use a spectrum of color going from yellow to purple to indicate the magnitude of the ratio, where yellow is the lowest and purple the highest. The figure is explanatory: there is not much information hidden, and both the highest and lowest ratios are highlighted by explainers.



The data encodes the frequency of visits of the 50 most popular articles in Wikipedia in 2023 and 2024. The data is presented over two bar charts (each year), organized decreasing descent. There are lines crossing from chart to chart, signifying the change in the relative order of popularity. Again, exploratory: The figure only presents the pages, the visits and the relative change, but it doesn't include explainers; only after exploring one might realize how, of the top 5 articles of 2024, 3 are related to U.S. Presidential Election, or how most of the articles are recent movie releases.