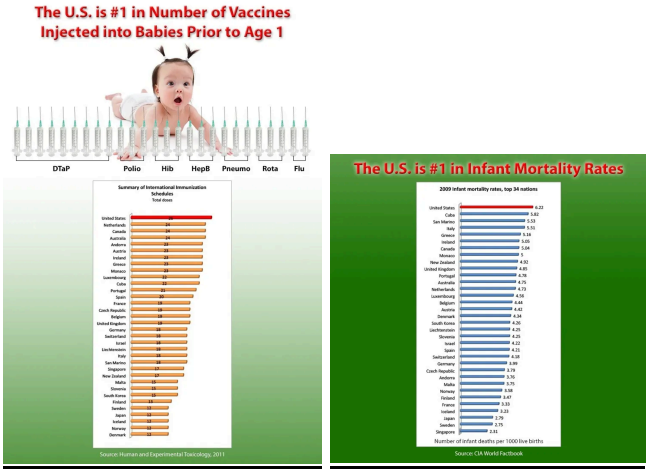
Identifying Techniques in Misleading Visualizations

We just spent some time understanding the commonly used techniques to create misleading data visualizations. Let's try to identify them in real world examples! Do **not use** SearchEngines/LLMs for this exercise, solve it using common sense and the techniques presented in the lecture.

Exercise:

1) Infant Mortality and Vaccinations

Identify the techniques used in the following misleading data visualization:

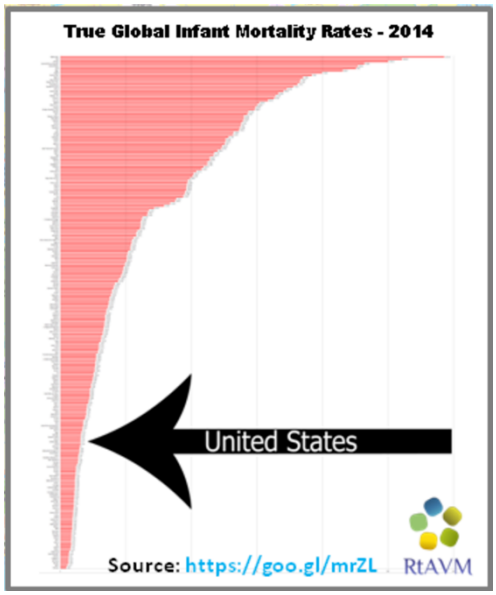


(source: https://healthimpactnews.com/2013/is-there-any-logic-behind-vaccine-claims/)

Misleading Techniques used:

- **Correlation and Causation**: the (postulated, but actually untrue – see below) correlation between Infant mortality immunizations (charts show the U.S. as #1 in both areas) is used to imply causality between the two.

- **Cherry Picking**: the right-hand side graph shows ONLY the countries with lower infant mortality ("top 34 nations"), but the world contains a lot more countries. On the left is the "full" graph (source: https://www.vislies.org/2016/gallery/InfantMortalityResponse.png) showing that the U.S. is nowhere near the worst country with regard to infant mortality. In fact, it is quite near the top. The graphic cheats by arbitrarily cutting off the list.



2) Lake Mead Water Levels

The Lake Mead reservoir (a lake close to Las Vegas, photo on the right by Nikola Majksner on unsplash) is a major water source for multiple states in the USA. The levels of water in Lake Mead therefore can have critical consequences.

All of the following visualizations are based on the same dataset of historical levels of water in Lake Mead and all of them are misleading, even though the contradict each other (Chart A and B argue for constant water levels, Chart C and D for decreasing water levels and Chart E for increasing water levels). (source: https://www.vislies.org/2017/gallery/)

a) Identify the techniques used in each of the following misleading data visualization.

Solutions:

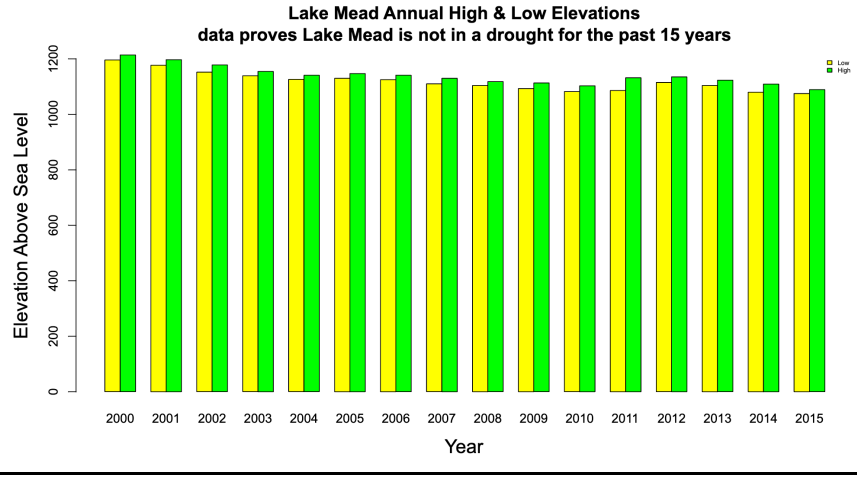
Important facts, you have to know/research about the Lake Mead water reservoir:

- The "real" minimum level of the reservoir is 895 feet: Water capacity at 895 ft elevation is considered "dead pool," which is when downstream releases from Hoover Dam are no longer possible.
- The maximum level of the reservoir is 1229 feet (Maximum designed water-surface elevation – top of Hoover Dam). The reservoir cannot hold more water than that.

  (source: https://www.nps.gov/lake/learn/nature/storage-capacity-of-lake-mead.htm)

➔ the valid ranges to consider for the reservoir water level are between 895 ft and 1229 ft. Anything significantly narrower or wider than that range is either exaggerating or diminishing perceived changes!

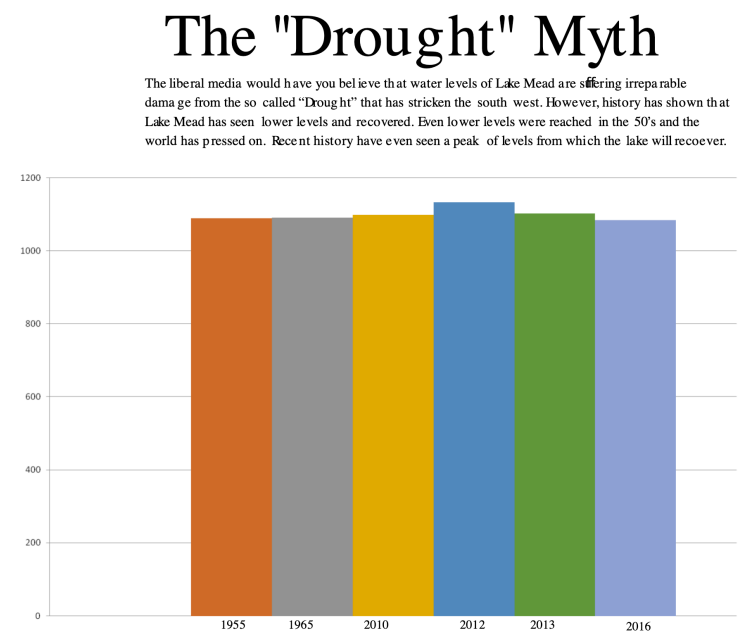
Chart A:



Misleading Techniques used:

- **Y-Axis Manipulated**: using the full scale from 0ft to 1200ft is misleading for this specific case, as the valid range is between 895ft and 1229ft. The range used shows mostly "dead storage".

Chart B:

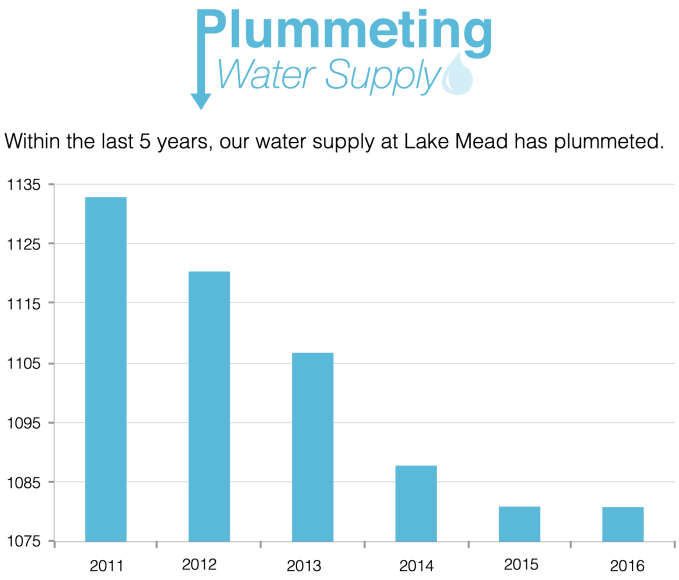


Misleading Techniques used:

- **Y-Axis Manipulated**: using the full scale from 0ft to 1200ft is misleading for this specific case, as the valid range is between 895ft and 1229ft. The range used shows mostly "dead storage", hiding the effect of the droughts.

- **Cherry Picking**: data is chosen so that the drought effects are hidden.

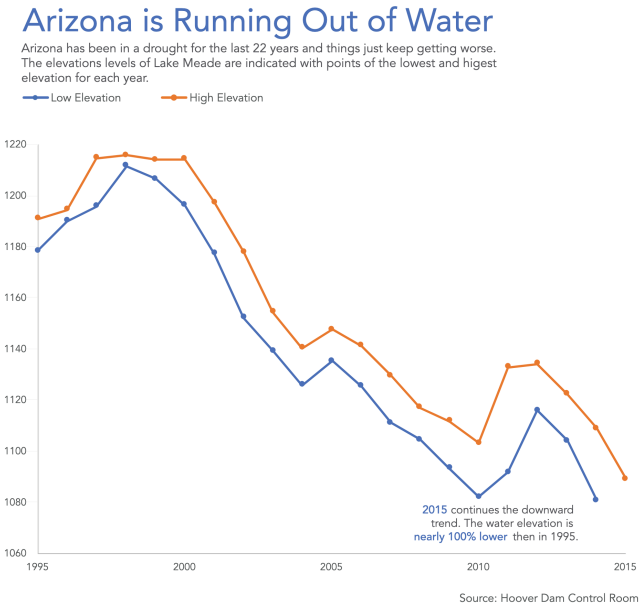
Chart C:



Misleading Techniques used:

- **Y-Axis Manipulated**: this Chart does the opposite of Charts A and B – the y-axis is much smaller than it should be, thereby exaggerating the effect of the drought.
- **Cherry Picking**: data is chosen (very narrow range of years) so it seems as if the water level is constantly and consistently falling.

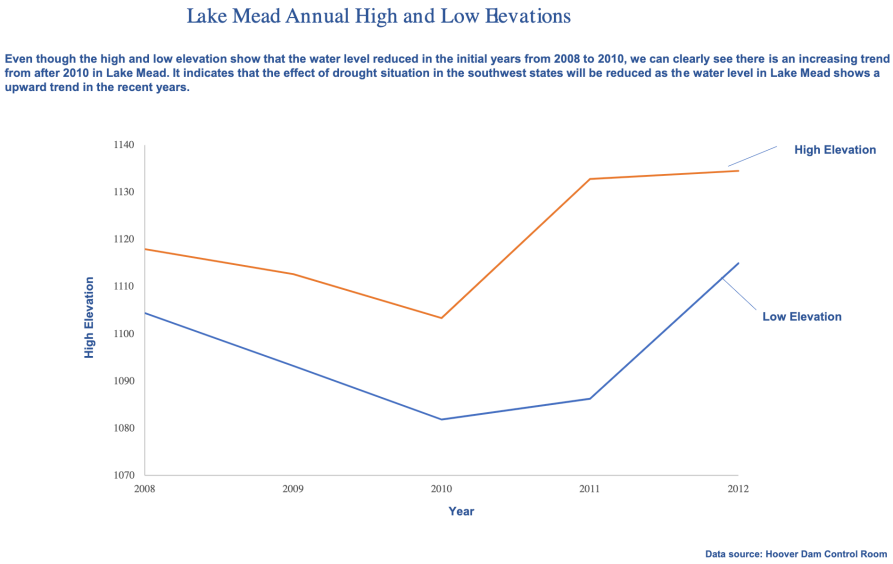
Chart D:



Misleading Techniques used:

- **Y-Axis Manipulated**: again, this Chart does the opposite of Charts A and B – the y-axis is much smaller than it should be, thereby exaggerating the effect of the drought. The Heading and texts on the chart further reinforce this argument.

Chart E:



Misleading Techniques used:

- **Cherry Picking**: once again, the data (years) on the x-axis is cherry-picked, but this time to make it look like the water levels are increasing.
- **Y-Axis Manipulated**: this cherry-picking effect is in addition exaggerated further by choosing a very small range of values for the y-axis.

b) So, what is the truth???
Solution: the water levels are decreasing, but not as much as it is implied in Charts C and D. A trustworthy chart would be similar to Chart D, but with a y-axis going from around 895ft to around 1229ft and showing all the years from 2000 to 2016 (or, even better, going back to 1936 when Hoover Damn was completed and Lake Mead became a water reservoir). Below is a non-misleading chart (from 2024):
(source: https://arachnoid.com/NaturalResources/)