

Name: _____

Date: _____

Answer Key: Data Detectives: Spotting Bias and Bad Graphs in Grade 6

Analyze how misleading visualizations and skewed sample groups can distort reality, building the critical eyes needed for modern citizen-journalism.

1. A researcher wants to find the average height of 6th graders globally but only collects data from a professional basketball youth camp. What is the primary issue with this data's reliability?

Answer: B) Selection bias in the sample

Selection bias occurs when the individuals chosen for a study are not representative of the general population being studied.

2. True or False: If two variables show a strong correlation on a scatter plot, it proves that one variable must be causing the other to change.

Answer: B) False

Correlation does not equal causation; two things can increase at the same time due to a third factor or coincidence.

3. When a news report uses a bar chart where the y-axis starts at 50 instead of 0 to make a small difference look huge, they are using a _____ axis.

Answer: B) Truncated

A truncated axis omits the bottom portion of the scale, which can visually exaggerate differences between data points.

4. You find a dataset about sugar's health benefits, but notice it was funded by a large candy corporation. Which evaluation criteria is most concerning here?

Answer: C) Conflict of interest

A conflict of interest arises when the source of funding has a financial stake in the outcome of the data analysis.

5. True or False: Using 'Anonymization' means removing identifying details like names from a dataset to protect person-level privacy.

Answer: A) True

Anonymization is a critical data management practice used to ensure individuals cannot be identified from the data.

Name: _____

Date: _____

6. Before analyzing a spreadsheet of global temperatures, you notice some entries are in Celsius and some in Fahrenheit. The process of fixing this is called data ____.

Answer: C) Cleaning

Data cleaning (or cleansing) involves identifying and correcting errors or inconsistencies in a dataset before analysis.

7. Which of these is an example of 'Metadata' for a digital photograph used in a citizen-science project?

Answer: C) The GPS coordinates of where it was taken

Metadata is 'data about data.' GPS coordinates provide context about the file rather than being the content itself.

8. A scientist publishes their full dataset online so other scientists can check their work. This practice is known as ____ Data.

Answer: C) Open

Open Data is information that is free to use, reuse, and redistribute by anyone, which promotes transparency in science.

9. If you are looking for specific, non-biased statistics on national unemployment, which source is generally considered the most authoritative?

Answer: B) A government labor bureau

Official government agencies are generally bound by strict data collection standards and provide the most reliable national statistics.

10. True or False: An 'Outlier' is a data point that is significantly different from the rest of the dataset and should always be deleted immediately.

Answer: B) False

Outliers should be investigated, not just deleted; they can represent errors, but they can also represent important rare events.