Name _____

CSI: Code Scene Investigation

Artificial Intelligence (AI) learns from data - the information people give it. But what happens if that data isn't fair or balanced? Bias in data means some voices or groups are overrepresented while others are left out. When an AI trains on biased data, it can make unfair decisions without even realizing it!

In this activity, you'll take on the role of a **Data Detective**. You'll study a few "data cases," look for unfair patterns, and predict how those biases might affect an AI trained on them. Then you'll answer questions to show what you've discovered.

The Hiring Database - A company wants to use AI to help review job applications. They give it a dataset of past employees so it can "learn" what a good worker looks like. Here's what the dataset looks like:

Category	% in Data	Notes
Male employees	80%	Mostly from engineering roles
Female employees	20%	Mostly from HR and design roles
College graduates	95%	Few examples of people with trade or technical training
Employees over 40	10%	Most data shows younger workers
Employees from one region	85%	Mostly from the company's home city

1. What kind of bias can you already spot in this dataset?
A. Gender bias B. Age bias C. Geographic bias D. All of the above

2. If this AI learns from this data, who might it think is the "best" kind of worker?
A. Anyone with a trade skill B. Mostly younger men from the company's home city
C. Mostly older women from different cities D. Equal chances for everyone

3. How might this bias affect the AI's hiring choices?
A. It might favor workers who look like the old data - mostly young men
B. It would focus on fairness and diversity
C. It would randomly choose people to be fair
D. It would only look at college degrees and nothing else

4. What could the company do to make its data more balanced?
A. Erase some data
B. Add examples of more diverse workers and skills
C. Train the AI faster
D. Ignore the problem