Name _____

When Smart Machines Make Dumb Decisions Answer Key

1. The Job Application Filter

Where might the bias or error have entered?
The AI trained on past hiring data that reflected the company's own gender bias. Since the system learned that "successful hires" were mostly men, it assumed male candidates were naturally better - copying human prejudice rather than removing it.

How could this system be improved?

Use **diverse, bias-audited training data** that include successful employees of all genders.

Remove gender identifiers from résumés before training (blind review).

Test outputs regularly for discrimination and adjust algorithms when patterns appear.

2. The Predictive Policing Program

Where might the bias or error have entered?
The algorithm relied on **historical arrest data**, which already reflected uneven policing. Because police had focused more on certain neighborhoods, the data contained more crimes from those areas - creating a feedback loop.

How could this system be improved?

Use **broader datasets** beyond arrest records (e.g., community reports, social data, or crime victim surveys).

Include human oversight to question patterns before acting on predictions.

Apply algorithms transparently and pair them with community accountability.

3. The Loan Approval Assistant

Where might the bias or error have entered?
The model used **geographic data (ZIP codes)** that reflected a history of redlining and housing discrimination. It learned a correlation between certain neighborhoods and financial "risk," even though that pattern came from unfair systems, not real ability to repay loans.

How could this system be improved?

Remove or de-emphasize variables like ZIP code or neighborhood demographics.

Audit datasets for **historical bias** and retrain models with more equitable data.

Combine algorithmic scoring with human judgment for final decisions.