Training Data Thoughts Answer Key

Case 1 - Q: Who is the current president of the United States?
AI Output: The current president is Barack Obama.

- **Training Data Problem:** Outdated data.

- **Why:** Obama left office in 2017, but old data may still appear in training sets.

Case 2 - Q: Describe what a nurse does.
AI Output: A nurse is a woman who helps doctors by cleaning rooms and bringing them tools.

- **Training Data Problem:** Biased samples.

- **Why:** Training data may contain stereotypes portraying nurses only as women and in assistant roles.

Case 3 - Q: What is the largest country in the world?
AI Output: The largest country is Canada.

- **Training Data Problem:** Too few examples.

- **Why:** Russia is the largest, but if the model saw limited examples comparing country sizes, it might overgeneralize based on "Canada is very big."

Case 4 - Q: Tell me about video games.
AI Output: Video games are mostly for young boys who play shooting games.

- **Training Data Problem:** Biased samples.

- **Why:** Training data may be drawn from narrow sources (like ads or forums) that emphasize boys and shooters, ignoring girls and many other game genres.

Case 5 - Q: Who was the first woman in space?
AI Output: The first woman in space was Sally Ride.

- **Training Data Problem:** Too few examples and U.S.-centric bias.

- **Why:** Many U.S.-based sources emphasize Sally Ride, while overlooking Valentina Tereshkova (the true first woman in space in 1963).