Teacher Answer Key

Case Study

1. The questionnaire asks about family members' arrests and incarceration, and about friends who have been arrested — factors rooted in Jerome's family history and neighborhood rather than his own actions. Critics argue this is problematic because these factors reflect social conditions and historical inequities that Jerome has no control over, not individual choices he has made. It effectively penalizes people for growing up in communities with higher policing and incarceration rates.

2. Jerome's question points to the due process concern about confronting evidence. In the American legal system, defendants have the right to examine and challenge the evidence used against them. But when a risk score is based on a proprietary algorithm whose formula cannot be disclosed, Jerome and his defender cannot determine how the score was calculated, whether it was applied correctly, or what factors drove it — making meaningful challenge impossible.

Multiple Choice

1. D
2. A
3. B

Matching

1-E, 2-A, 3-D, 4-C, 5-B

Fill-in-the-Blank

1) 2016, 2) 2020, 3) Sanctions

Short Answer (Sample Responses)

Q1: Historical arrest data does not measure crime — it measures where police have previously been deployed and made arrests. When an algorithm directs more patrols to a neighborhood based on high past arrest rates, those patrols make more arrests, including for minor offenses. That new arrest data confirms the neighborhood's "high crime" status, directing still more police there. The algorithm continuously validates its own predictions rather than measuring actual crime patterns.

Q2: This matters because it means there is no technical solution that is simply "correct." Different communities and stakeholders define fairness differently — some prioritize equal false positive rates across groups, others prioritize equal overall accuracy, others prioritize equal false negative rates. Since satisfying all

simultaneously is mathematically impossible, choosing which version of fairness to optimize for is a policy and ethical decision that cannot be delegated to an algorithm.

Q3: The strongest argument is that data-driven decision-making, done responsibly, would be more consistent and less susceptible to individual human bias than subjective human judgment — a judge's or officer's biases might be less systematic and harder to identify than an algorithm's. Critics respond that if the data itself was generated under conditions of historical discrimination, then "more consistent" just means more consistently reproducing that discrimination, which is not an improvement over imperfect but challengeable human judgment.