1. **Carefully read the background and collection plan again. What types of potential bias exist in your team lead's collection plan? Why was it biased? Please explain your answer. You may also think of biases that go beyond this reading (e.g., cultural bias).**

1. Selection Bias

- **Description**: This occurs when certain groups or individuals are systematically excluded from the data collection process.
- **Reason**: The focus on ATM transactions near the U.S.-Mexico border may lead to a selection bias, as it may disproportionately capture transactions related to money laundering while neglecting normal transactions from Mexican citizens working in the U.S. This selection could skew the model towards flagging more transactions from this demographic.

2. Confirmation Bias

- **Description**: This bias arises when individuals favor information that confirms their pre-existing beliefs or hypotheses.
- **Reason**: If the analysts believe that transactions involving Mexican citizens are more likely to be suspicious, they may unconsciously flag more of these transactions as positive, thereby reinforcing this bias in the model. The existing perception of Mexican citizens in relation to money laundering could influence their scoring decisions.

3. Cultural Bias

- **Description**: This occurs when the data collection and analysis do not account for cultural differences or socioeconomic factors that might affect behavior.
- **Reason**: The model may not adequately consider legitimate behaviors of Mexican citizens who live and work in the U.S. For instance, many may regularly send money home or convert currencies for legitimate reasons. Failing to distinguish these from suspicious activities can lead to disproportionate targeting of this group.

4. Measurement Bias

- **Description**: This bias arises from using flawed measurement instruments or criteria that do not accurately capture the intended phenomenon.
- **Reason**: If the model's rules and thresholds are overly simplistic or do not account for the complexities of cross-border transactions, it may lead to incorrect classifications. For instance, frequent small withdrawals by Mexican citizens may be legitimate, but the model may flag them as suspicious based on a rigid set of rules.

5. Presentation Bias

- **Description**: This bias occurs when the way data is presented influences how it is interpreted or perceived, potentially leading to skewed conclusions.
- **Reason**: In this case, if the results of the analysts' evaluations are displayed in a manner that emphasizes the number of suspicious transactions flagged by Mexican citizens, it

could reinforce stereotypes or assumptions about this demographic. For example, if the findings highlight the high percentage of flagged transactions involving Mexican citizens without context, it may lead stakeholders to draw misleading conclusions about their behavior.

2. **How might these biases distort the results? What could you do to avoid these biases?**

Biases in data collection and analysis can significantly distort results, leading to inaccurate conclusions and potentially harmful decisions. Here's how these biases might affect the results and some strategies to mitigate them:

How Biases Might Distort Results

1. **Selection Bias**
   - Distortion: This can lead to overrepresentation of certain groups, such as transactions from Mexican citizens, while underrepresenting others. This skew can result in false assumptions about who is more likely to engage in suspicious behavior.
   - Mitigation: Ensure a diverse sample that represents the entire population of customers. Include various demographics and transaction types to create a more balanced dataset.
2. **Confirmation Bias**
   - Distortion: Analysts may focus on confirming their beliefs, leading to an inflated number of suspicious flags for certain demographics, reinforcing stereotypes.
   - Mitigation: Implement blind scoring where analysts do not know the demographic details of the transactions. Additionally, encourage critical discussions about assumptions and regularly review scoring criteria for fairness.
3. **Cultural Bias**
   - Distortion: Cultural misunderstandings can lead to misinterpretation of legitimate behaviors as suspicious. For instance, frequent currency conversions by migrant workers could be flagged incorrectly.
   - Mitigation: Train analysts on cultural contexts and legitimate transaction behaviors. Incorporate expert insights on cultural norms related to financial transactions.
4. **Measurement Bias**
   - Distortion: Overly simplistic or rigid models might miss the nuances of transaction behavior, leading to false positives or negatives.
   - Mitigation: Use advanced statistical methods and machine learning algorithms that can adapt to complex patterns in the data. Continuously test and validate the model against real-world scenarios.

5. **Presentation Bias**
    - Distortion: Data presented in a way that emphasizes certain findings can lead to misinterpretation, causing stakeholders to overreact or misallocate resources.
    - Mitigation: Present data neutrally, providing context and explanations for findings. Use balanced visualizations that do not disproportionately highlight certain groups without appropriate context.

3. **If you know that there is bias in the collection method, what could you do to communicate your concerns to your team lead? Please be as specific as possible.**

1. Prepare Your Observations

- **Gather Evidence**: Compile specific examples of potential biases you've identified, such as selection bias or confirmation bias. Document instances where the data collection might misrepresent certain demographics or lead to skewed results.
- **Impact Analysis**: Consider how these biases could impact the outcomes of the project, such as false positives in fraud detection or misallocation of resources.

2. Schedule a Meeting

- **Timing**: Request a dedicated meeting to discuss your concerns rather than bringing them up casually. This shows you take the matter seriously and gives your lead space to engage thoughtfully.
- **Agenda**: Provide a brief agenda in advance, highlighting that you'd like to discuss data collection methods and potential biases.

3. Present Your Concerns Clearly

- **Structured Approach**: Use a structured format to communicate your concerns. For example:
    - **Introduction**: State the purpose of the meeting and the importance of unbiased data collection.
    - **Observation**: Clearly outline the specific biases you've identified in the collection method (e.g., focus on transactions near the border leading to selection bias).
    - **Implications**: Explain the potential consequences of these biases on the model's effectiveness and the bank's reputation.

4. Suggest Solutions

- **Propose Changes**: Offer concrete suggestions for addressing the identified biases. For example:
    - **Broader Sampling**: Recommend expanding the sample to include a wider range of transaction types and demographics.

    - **Analyst Training**: Suggest training sessions for analysts on cultural contexts and objective scoring methods.
    - **Validation Procedures**: Advocate for implementing regular audits of the data collection process to identify and correct biases early.

5. Encourage Open Dialogue

- **Invite Feedback**: Ask your team lead for their perspective on your concerns and suggestions. Be open to discussing any challenges they foresee in implementing changes.
- **Collaborative Approach**: Emphasize that you're raising these issues to improve the project collaboratively and ensure its success.

6. Follow Up

- **Document the Discussion**: After the meeting, summarize the key points discussed and any agreed-upon actions in an email. This provides a record of your concerns and the steps to address them.
- **Continued Engagement**: Offer to assist with implementing any changes or monitoring for biases in future data collection efforts. This shows your commitment to the project and willingness to contribute positively

4. **Read through the details of testing. How might the lack of transparency around the experience and training of the investigators allow for bias?**

The discrepancy between the proportion of Mexican citizens in the overall population (11%) and their overrepresentation in flagged transactions (75%) raises concerns about potential **bias**. A lack of transparency regarding the experience and training of analysts involved in scoring transactions could contribute to this issue in several ways:

1. Implicit Bias in Scoring

Even with good intentions, analysts may hold unconscious stereotypes about certain nationalities or transaction patterns. Without proper anti-bias training, they could perceive activities involving Mexican citizens as more suspicious, leading to a higher number of false positives.

Impact:

- Analysts may conflate legitimate cross-border transactions with illegal activities, especially since many Mexican nationals living legally in the U.S. convert dollars to pesos for practical reasons.

2. Inconsistent or Inadequate Training

If analysts are not uniformly trained on the key variables to look for (such as account age, transaction frequency, amounts, and patterns), they might place too much weight on easily identifiable characteristics, like nationality or location, rather than more objective red flags.

Impact:

- A reliance on subjective factors like nationality can result in skewed outcomes, falsely identifying routine transactions as suspicious.
- Analysts with inconsistent interpretations may introduce noise, making it harder to build a reliable model.

3. Limited Experience and Groupthink

If analysts have varying levels of experience in anti-money laundering (AML) investigations, those with less experience might follow patterns they assume to be suspicious based on limited knowledge. In a group setting, inexperienced analysts may also rely on trends observed among their peers, creating a feedback loop of incorrect assumptions.

Impact:

- **Overflagging** transactions involving Mexican citizens due to assumptions reinforced by other analysts.
- Inconsistent scores across analysts with different experience levels lead to unreliable data for model development.

4. Absence of Guidance on Cross-Border Context

The analysts need to distinguish between **legitimate cross-border activity** and suspicious patterns, but if they aren't trained in the nuances of cross-border transactions (e.g., seasonal travel, remittances, or cash conversions), they may incorrectly label legal activity as suspicious.

Impact:

- Mexican nationals engaging in legitimate transactions could be disproportionately flagged.
- Cross-border transactions are inherently complex, and the lack of contextual training introduces bias into the decision-making process.

5. Pressure to Perform and Score Large Volumes of Items

The task of scoring 1,000 transactions per analyst can lead to **fatigue**, errors, or reliance on cognitive shortcuts. If analysts are overwhelmed or under time pressure, they might focus on certain "shortcuts," such as nationality, location, or frequency, rather than deeper investigative patterns.

Impact:

- Analysts may rely on heuristics (e.g., nationality) as an easy way to mark transactions, leading to biased outcomes.
- The data that the model relies on could perpetuate this bias in future iterations.

5. **Analyze the bar chart showing the scores of individual analysts and see where their scores fall on the distribution curve. If the mean of the scores was 307 and the standard deviation is 166, which score, or scores might you eliminate to control for bias? Why?**

Let's analyze the **individual analyst scores** with the provided summary statistics:

- **Mean (μ):** 307
- **Standard Deviation (σ):** 166

Step 1: List of Analyst Scores

Based on the bar chart, the scores are:

1. Analyst 1: 179
2. Analyst 2: 225
3. Analyst 3: 230
4. Analyst 4: 232
5. Analyst 5: 250
6. Analyst 6: 275
7. Analyst 7: 278
8. Analyst 8: 280
9. Analyst 9: 358
10. Analyst 10: 759

Step 2: Calculate Z-scores

We calculate the z-score for each analyst's score to see how far it deviates from the mean.

Z=X−μσZ = \frac{X - \mu}{\sigma}Z=σX−μ

Where XXX is the individual score, μ=307\mu = 307μ=307, and σ=166\sigma = 166σ=166.

Analyst	**Score**	**Z-Score**
1	179	-0.77
2	225	-0.49

Analyst	Score	Z-Score
3	230	-0.46
4	232	-0.45
5	250	-0.34
6	275	-0.19
7	278	-0.17
8	280	-0.16
9	358	+0.31
10	759	+2.72

Step 3: Identify Outliers

A **z-score greater than +3 or less than -3** is usually considered an outlier. While none of the scores hit the strict +3 threshold, **Analyst 10** has a **z-score of +2.72**, which is **very close** to being an outlier.

Score of 759 stands out significantly compared to the other scores, most of which cluster between 200 and 300. It's possible that **Analyst 10's scoring introduces bias**, as the analyst flagged more than **double the average number of suspicious transactions** (759 vs. 307).

Step 4: Recommendation

Given the large deviation of **759** from the rest of the scores:

- **Analyst 10's score** should be reviewed for potential bias or inconsistency.
- It might be wise to **exclude this score** from the dataset to ensure that one outlier doesn't skew the model's accuracy.

Reason:

- This extreme score suggests that the analyst may have applied different or incorrect criteria when identifying suspicious transactions, which could distort the overall results. Excluding it would prevent the model from learning biased patterns.