

## Step 4: Communicate with Stakeholders

### Email/Slack Message

**Subject:** Data Quality Issues and Next Steps

Hello Sir/Madam

I've reviewed the data and identified a few data quality issues that need attention:

1. **Missing Data:** Some receipts lack critical fields like user\_Id or dateScanned.
2. **Inconsistent Data:** The rewardsReceiptStatus field contains unexpected values.
3. **Duplicate Records:** There are duplicate entries in the Receipts table.

To resolve these issues, I need clarification on:

- What should we do with records that have missing or invalid data?
- Are there specific rules for handling duplicates?

Additionally, to optimize the data assets, it would be helpful to know:

- Are there any business rules for validating receipts or users?
- Do we have a process for updating or archiving outdated records?

Finally, I anticipate performance concerns as the dataset grows. To address this, I recommend:

- Indexing critical fields like user\_Id and dateScanned.
- Partitioning large tables (e.g., Receipts) by date.

Let me know your thoughts, and I'll proceed accordingly!

Best,  
William Armah