Subject: Urgent: Data Quality Issues and Business Impact

Hello Team,

I hope you're well. I've completed a detailed analysis of our Receipts, Users, and Brands data and identified several critical issues that are affecting our ability to answer key business questions.

Key Questions and Issues:

- **Data Uncertainty**: We need clarification on the meaning and relevance of several columns in the "rewardsReceiptItemList" (e.g., preventTargetGapPoints, pointsPayerId, originalMetaBriteBarcode, etc.).

- **Data Quality**: We've found a high percentage of null values across multiple columns and significant relationship inconsistencies, such as receipt IDs missing from the main receipts table and user IDs in the receipts table without matching entries in the users table.

- **Data Integrity**: Over half of the Users data contains duplicates, and date formats are inconsistent.

Proposed Solutions:

- **Immediate Action**: Implement data deduplication for the Users table, standardize date formats, and review database indexing.

- **Long-term Fixes**: Establish data validation rules, automated quality monitoring, and governance policies to ensure ongoing data integrity and alignment with business objectives.

Production and Scaling:

- To support future growth, we recommend data partitioning, optimizing indexing, and introducing caching mechanisms to handle increased data volume and ensure system performance.

I'd be happy to discuss these findings and next steps in more detail. Please let me know when we can schedule a meeting.

Best regards,
Upendra