Third: communicate with stakeholders

Hi,

Following my analysis of the datasets, here are the key findings, trends, and recommended actions:

Key Data Quality Issues:

1. **GENDER Field:** Ambiguous values (e.g., "Non-Binary" vs. "non_binary") require standardization for accurate demographic segmentation.
2. **BARCODE Field:** Missing or duplicate values hinder product-level analysis. Collaboration is needed to resolve illogical entries (e.g., "-1").
3. **FINAL_SALE and FINAL_QUANTITY Fields:** Blank or invalid values require clarification from stakeholders. Decimal quantities likely represent weighted items but need confirmation.
4. **CATEGORY_4 Field:** High missing rate (92%) suggests insufficient product categorization or structural issues.
5. **USER_ID and RECEIPT_ID:** Validation is needed to confirm relationships and address potential batch transaction errors.
6. **Limited Transaction Data:** Current timeframe (June to August 2024) restricts long-term trend and seasonal sales analysis.

Key Trend in the Data:

Female users dominate spending across most age groups, especially in the Middle-Aged (36-60y) and Young Adult (18-35y) categories. Elderly Males (60+y) show higher spending than females in their age group, presenting a targeted opportunity.

Recommendations for Action:

1. **Standardize GENDER values** and confirm acceptable categories with stakeholders.
2. Investigate and resolve **BARCODE and CATEGORY_4 issues** to improve product-level insights.
3. **Expand the dataset timeframe** to enable robust seasonal and long-term trend analyses.
4. Leverage insights for targeted campaigns:
    - **Middle-Aged Females (36-60y):** Focus on family, health, and convenience.
    - **Young Adult Females (18-35y):** Promote trendy, eco-friendly products for loyalty building.
    - **Elderly Males (60+y):** Highlight health and wellness products, emphasizing trust and value-added features.

Please confirm if these approaches align with the business priorities or if additional data/context is required to refine the analysis.