

Communicate with Stakeholders

Email –

Subject: Data Quality Insights and Key Trends from Fetch Data Analysis

Hi [Business Leader's Name],

I hope this email finds you well. I've conducted a thorough analysis of our user, transaction, and product datasets, and here are some key findings and outstanding issues that need to be addressed:

Key Data Quality Issues:

User Data -

Missing Language Values: Around 30% of users don't have their preferred language recorded. This can significantly affect personalized marketing efforts.

Inconsistent Gender Data: There are inconsistencies and missing values in the gender column. We should standardize gender categories to get a clearer understanding of our user demographics.

State and Birth Date: Minor gaps in the data that need attention for accurate demographic profiling.

Transaction Data -

Missing Final Sale Values: 25% of the transactions are missing final sale amounts, which is critical since it could lead to incorrect reward points for users, potentially affecting user satisfaction and loyalty. It may also distort sales data used for brand partnerships and marketing efforts.

Final Quantity Data Issues: Some transactions show decimal values for the number of items purchased, which is likely incorrect and could misrepresent the volume of purchases.

Outliers: Both the Final Sale and Final Quantity columns have high-value outliers. Such values can result in misinformed business decisions, such as revenue forecasts. This may require further investigation to ensure data accuracy.

Product Data -

Missing Brand and Manufacturer Information: 26% of product records are missing brand and manufacturer details. This could limit our ability to enhance the user experience through product recommendations and might impact loyalty partnerships with brands. Additionally, some of the values are incorrect (placeholder values) which can obstruct accurate data analysis of brands and manufacturers.

Categorization Gaps: A significant number of records are missing detailed subcategory information, making it harder to precisely categorize products.

Interesting Trend

From the analysis, I found that females are Fetch's major user base. By focusing on the top categories and subcategories that resonate with female users, such as Health & Wellness and Snacks, we can tailor marketing efforts to increase engagement and revenue. For example, popular subcategories like Skin Care and Medicines & Treatments within Health & Wellness and Candy within Snacks have the highest purchase frequency among female users.

Request for Action

To fully address these data quality issues and make better-informed business decisions, I recommend -

Collaboration with the data engineering team to resolve missing and inconsistent values across all datasets.

Clear guidelines from business stakeholders on how to handle missing data for all datasets – whether through imputation, targeted surveys, or another strategy? This is a very important step as it could impact the integrity of our reward system.

Support from the product team to review the current product categorization process to ensure products are classified appropriately at all levels.

These steps will ensure a cleaner dataset for more reliable future analyses and drive targeted marketing efforts based on the trends observed.

Let me know if we can set up a meeting to discuss these findings further.

Best,

Aditi Patel