

**Subject:** Fetch 2024 Analysis Update – Key Insight and Data Quality Observations

Hi team,

Hope you all are doing well!

The team and I recently completed an analysis of Fetch's performance for June to September 2024.

**Key Finding:** We observed a decline in the number of new users joining Fetch, with a negative year-over-year growth of -42% in 2023 and -24% in 2024. This has brought the new user acquisition numbers back to pre-COVID (2020) levels. It's important that we explore the reasons behind this continued decline in new users over recent years.

**Note:** This insight was derived using the *Created Date* field from the User data and is **not** impacted by the data quality concerns outlined below.

**Caveats:** We faced some data quality issues while performing the analysis (listed below). Due to this, the analysis results need to be re-evaluated once the data quality issues are fixed to give a better picture of the Fetch landscape.

Here are the issues along with additional help needed and solutions implemented:

**High Priority Data Issues:**

**1. Low User Mapping in Transaction Data**

- i. Only about 0.5% of users in the Transaction dataset are successfully linked to the User dataset.
- ii. **Impact:** This limits the ability to generate meaningful user-level insights (such as age, sign-up date, and demographics).
- iii. **Action Needed:** Improve the mapping between Transaction and User data.

**2. Missing Barcodes in Transaction Data**

- i. Approximately 10% of transactions are missing barcode information.
- ii. **Action Needed:** Better barcode mapping to support product-level analysis.

**3. Barcode Format Issues Across Datasets**

- i. Barcodes are stored as integers of varying lengths in both Transaction and Product datasets.
- ii. **Impact:** Leading zeros may be dropped during CSV file transfers, causing inconsistencies.
- iii. **Solution Implemented:** Converted barcodes to string type and standardized to a length of 14 characters.

**4. Missing Brand and Manufacturer Details in Product Data**

- i. Around 26% of products are missing brand and manufacturer information.
- ii. **Action Needed:** Retrieve missing brand and manufacturer details to enrich the product dataset.

**5. Duplicate Transactions in Data**

- i. Some Receipt ID and Barcode combinations are duplicated, often associated with:
  - > Missing Final Sale values
  - > Final Volume recorded as zero
- ii. **Solution Implemented:** Removed duplicate entries where these conditions were found.

**Other Observations/Questions:**

- **Users with Incomplete Profiles:**  
Some users have missing fields such as State, Birth Date, Gender, and Language.
- **Products with Missing Categories:**  
Certain products lack Category information.
- **Non-Integer Values in Final Quantity Field:**  
Observed non-integer values — could this be due to varying unit types across different products?

Please let me know which team(s) would be the right point of contact to help address these data quality issues. I'd be happy to set up a call to walk through these points in detail if needed.

Best,  
Dhanesh Akolu