

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

Hi team,

Hope you all are doing well!

The team and I conducted an analysis of Fetch's performance for 2024 (June to September).

**Key Insight:** There has been negative YoY growth in the number of new users joining Fetch in 2023 and 2024 (-42% and -24% resp). Bringing the new user count to pre-COVID levels (2020). We need to understand why new users have reduced in recent years.

**Note:** The above insight is generated using the Created Date from User data and is therefore not impacted by the below-mentioned data quality issues.

**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. Only ~0.5% of users in the Transaction data were mapped to the User data
  - i. This limits the insights we can generate related to users (age/joining date/demographic)
  - ii. **Needed:** Improved mapping of the user data to generate useful user level insights
2. Transaction data with missing Barcodes
  - i. ~10% of transactions do not have a barcode
  - ii. **Needed:** Improved mapping of Barcodes to generate product level insights
3. Barcodes in the Transaction and Product data are stored as integers and are of varying length
  - i. This creates an issue, since any barcode with zero as the first digit(s) might get dropped while being transferred from/to a csv file
  - ii. **Solution:** Resolved by converting Barcodes to string and making it a string of length 14
4. Missing Brand and Manufacturer information in Product data
  - i. ~26% of products are missing the brand and manufacturer information
  - ii. **Needed:** Get Brand and Manufacturer information for the missing products
5. In Transaction data, we have some duplicate records where the combination of Receipt ID and Barcode appears more than once. Specifically, in these cases:
  - i. The Final Sale is missing
  - ii. The Final Volume is zero
  - iii. **Solution:** Resolved by dropping the duplicate rows

**Other Data Issues/Questions:**

1. Users with missing information:
  - i. Some users have missing information like State, Birth Date, Gender and Language
2. Products with missing information:
  - i. Some products have missing Category information

3. Final Quantity field has non-integer values, is this because of different units of each product?

Please let me know which team(s) I can reach out to in order to solve the above data quality issues.  
Happy to setup a call to discuss these issues in detail.

Best,

Adit Pradhan