Hi Team,

I've analyzed our user, transaction, and product data and uncovered some important findings that could impact decision-making. Below is a quick summary of the key data quality issues, an interesting business trend, and what I need from you to move forward.

Key Data Quality Issues

Missing & Inconsistent Data:

User Data: Birth date, state, language, and gender fields have significant gaps, making it hard to segment and analyze user demographics.

Transaction Data: Around 11.5% of transactions are missing barcodes, and 25% are missing final sale values, making it difficult to track product performance.

Product Data: Over 226K records lack brand or manufacturer info, and some barcodes are linked to multiple brands. Could this be due to data mismatches, inconsistencies in product sourcing, or legitimate multi-brand assignments?

Potential Data Errors & Anomalies:

1. We found transactions where the quantity is 0, but there's still a sale amount. Could this be due to refunds, pricing issues, or mistakes in data entry?
2. Also, some gender entries are written differently ("non_binary" vs. "non-binary"), and some state names may need to be checked for accuracy.
3. We also noticed that FINAL_QUANTITY sometimes has decimal values instead of whole numbers. Is this because the quantity is measured by weight or volume, or could it be a data error?

Open Questions:

1. Should we try to fill in the missing barcode information ourselves, or should we leave out these transactions from our analysis since the data is incomplete?
2. If the quantity is **0**, but there's still a sale amount, does this mean the transaction is valid, a refund, or a possible error?
3. Should each product have a unique barcode, or is there a valid business reason for different brands using the same barcode?

<u>Interesting Trend: Fetch's User Growth & Recent Decline</u>

Fetch grew rapidly from 2017 to 2020, reaching 16,886 new users in 2020 (a 138% increase from the previous year).

After that, growth slowed in 2021 (+13%), picked up again in 2022 (+40%), but then dropped sharply in 2023 (-42%) and 2024 (-25%).

Request for Action – Need Your Input!

To move forward, I need your thoughts on:
1. Do we attempt to fill in missing barcodes manually or exclude these records?
2. If a sale amount is recorded, but the quantity is **0**, is this a valid transaction, a refund, or a mistake?
3. Do we have any insights on why Fetch's growth has slowed down?

Your feedback will help us improve data reliability and optimize future analysis. Let me know your thoughts, and let's align on the best next steps.

Thanks,
Prachi Yadav
Senior Data Analyst