Subject: Data Quality Challenges, Key Insights & Clarifications Needed

Hi Team,

I've been analyzing Fetch's **user, transaction, and product data** to uncover insights that can help us better understand user behavior and drive value for **CPG brands and retail partners**. While this data provides valuable opportunities, I've encountered some complexities that require further clarification. Below is a summary of key data quality issues, interesting trends, and some questions where your input would be helpful.

Key Data Quality Issues & Challenges

1. **Limited User Match Rates:**
   ○ Only **0.5% of users** appear in transactions.
   ○ This suggests we may not be capturing the full scope of purchases or that our user directory is incomplete.
   ○ **Question:** Do we not have user data for these transactions, or is it missing from our dataset?
2. **Duplicate Transactions:**
   ○ **50% of transaction records** were redundant and required aggregation at the **quantity and sale level**.
   ○ **Question:** Is the duplication occurring because users are attempting to scan the same receipt multiple times, or is it a data processing issue where the system isn't consolidating transaction details correctly?
3. **Unclassified Manufacturers & Brands:**
   ○ **14.13% of products** lack manufacturer labels, reducing visibility at the **brand level**.
   ○ Some brands are labeled as **"BRAND NOT KNOWN"** or **"PLACEHOLDER MANUFACTURER"**, which limits our ability to track performance.
   ○ **Question:** Could this be due to private label items, supplier inconsistencies, or a need for better manufacturer data integration?

Primary Insight: User Growth Slowdown

Fetch's growth has evolved in three phases:

1. **2014-2018:** Early adoption.
2. **2018-2022:** Exponential growth.
3. **2022-Present:** Growth plateauing, indicating **market saturation or shifting user behavior**.

A **Q3 2024 spike in new user signups** was followed by a **sharp drop-off**, suggesting an opportunity to refocus efforts on **user acquisition, engagement, and loyalty-driven promotions**.

Additional Notable Insights

- **Basket Size Trends:** 99% of transactions contain only one item, indicating a strong single-item redemption pattern.
- **Retailer Dependence:** 42% of transactions occur at Walmart, suggesting strong engagement but also over-reliance on one retailer.
- **Sales Spike on July 1st:** Notable transaction increase, probably this was linked to a marketing push.

Next Steps & Questions

Would love to get your thoughts on the following:

1. **Can we establish a structured data processing framework** to ensure data quality and provide access to a more complete 2024 dataset for deeper analysis and better trend identification?
2. **Was the July 1st transaction spike driven by a marketing campaign, retailer promotion, or an external factor?** Understanding this will help us optimize future engagement strategies.
3. **Are there existing strategies in place to encourage multi-item purchases**, or should we explore new incentives to increase basket sizes and transaction value?

Looking forward to your insights. Let me know if you'd like to discuss further.

Best,

Sarnav