**Subject**: Data Quality Issues and Key Insights from Fetch Data Analysis

Dear [Stakeholder's Name],

I hope you're doing well. I wanted to share some key findings from my analysis of the Fetch data, including data quality issues, an interesting trend, and next steps.

**Key Data Quality Issues**

During our initial review, we identified several data quality concerns that could impact analysis:

1. **Products Data**
   - **Missing Values:** Several product entries have missing barcodes, which disrupts the connection between transactions and products.
   - **Unclear Product Fields:** Columns such as CATEGORY_1, CATEGORY_2, etc., lack clear descriptions, making it difficult to analyze product performance.
2. **Users Data**
   - **Incomplete Demographics:** Missing birthdate values hinder generational analysis.
   - **Inconsistent Gender Labels:** Variations like "prefer_not_to_say" vs. "Prefer not to say" create redundancy and inconsistency in DEI-related analysis.
3. **Transactions Data**
   - **Invalid Transaction Dates:** Some purchases have scan dates that precede the purchase date, suggesting potential data entry errors.
   - **Quantity and Sales Mismatches:** Some transactions have a final quantity > 0 but missing final sale values, making revenue calculations unreliable.

Our analysis reveals significant patterns in user behavior and product interactions, shedding light on Fetch's competitive positioning and growth opportunities.

1. **User Engagement Trends**
   - Fetch saw explosive user growth between 2017 and 2020, with the highest spike in 2017 (820% YoY growth) and continued strong expansion through 2020 (138% growth). However, we observed a significant slowdown in new user registrations starting in 2021, with sign-ups dropping 42% in 2023 and another 25% in 2024.
   - This decline could indicate market saturation, reduced marketing effectiveness, or changing user preferences. Further investigation into acquisition channels and retention strategies may help clarify the underlying cause.
2. **Power Users Insights**
   - I have assumed power users as the top 1% of users by number of receipts scanned. These users contribute disproportionately to engagement and transaction volume.

Understanding their behaviors can help optimize rewards, incentives, and user retention efforts.

- We should explore what drives their engagement and whether similar behaviors can be encouraged among the broader user base.

**Next Steps & Request for Support**

To improve the reliability of our analysis, we need to:

- Clean up missing values, particularly barcodes and demographic data.
- Address transaction duplicates to ensure accurate reporting.
- Clarify product category fields—additional documentation or input from the product team would be helpful to understand their definitions.

Once these issues are addressed, we can refine our insights and provide more accurate recommendations. Please let us know if you can provide additional context on product categories or if we should schedule a meeting to discuss next steps.

Looking forward to your thoughts.


Best,

Nikhila Chowdary Vaitla