Subject: Important Findings and Questions Regarding Our Data Quality

Dear xxx,

I hope you're doing well. I've been working on analyzing and refining the data we have in our 'receipts', 'users', and 'brands' datasets and wanted to share some findings and thoughts with you.

Here's what I've found and what it could mean for us:

Duplicate Information: In the 'receipts' dataset, it seems like we have some information that is repeated. For example, we have 'barcode' information appearing in both the 'barcode' field and the 'userFlaggedBarcode' field. I've noticed similar patterns with item descriptions and prices. It's like having two drawers in your desk with the same stuff - takes up extra space without giving us extra benefit.

Potential Errors: The 'users' dataset has quite a few identical entries, just like having the same contact saved multiple times in your phonebook. This could lead to confusion. Also, the 'brands' dataset has some issues, particularly with 'barcode' values which seem inconsistent or missing, like some items in our inventory missing labels.

To fix these and to ensure we make the most out of our data, it would be helpful to know:

- Is there a specific reason why our data is structured this way? Understanding this could help us to organize it better.
- Has there been any recent changes in the way we collect or organize our data?
- As we continue to grow and handle more data, we will also need to think about how we store and process it efficiently. Just like a highway can only handle so many cars before there's traffic, our current systems can only handle so much data.

I'm looking at different ways of managing this and am excited to talk more about it with you. If possible, could we schedule some time for a quick chat so that I can demonstrate the details? I believe this will be beneficial for our mutual understanding and future collaboration.

Thank you for your attention, and I am looking forward to your reply.

Best,

Runqing Jia
Data Analyst, Fetch Rewards