Subject: Regarding data quality issues on users, receipts, and Brands data

Hello stakeholders,

I hope this email finds you well.

Today, I'm emailing you to report my initial investigation of the Users, Receipts, and Brands datasets in the database.

Below are what I found regarding data quality issues:

1. I failed to find the brand identifier in the receipts dataset that I can link with two datasets. So, for now, I couldn't answer all the information you want to know. So, for creating a structured relational model, I need to have further information so that I can link the Receipts dataset with Brands dataset.

2. I found a lot of missing data and duplicate data on each dataset. Especially, for the analysis purpose, I separated the RewardsReceiptItemList column with multiple values in the Receipts dataset, and made the 4th dataset. I expected that I can find any identifier that can link with the Brands dataset, and other valuable information to give me some actionable insights, but there are over half of the columns inside the dataset is null values, and didn't have any identifier either.

With these data quality issues, I'd like to meet you and discuss them further.

I'll appreciate it if you suggest a few available time slots at your end so that I can set up a meeting.

Thanks and looking forward to your reply.

All the best,
Seungmin Yang