**Email to Product Leader – Interview Case Assignment**

Subject: The data quality of Receipts, Users, and Brands data has problems.

Hi Jeff,

I conducted an exploratory analysis of the Receipts, Users, and Brands data and found some data quality issues that I believe are important to bring to your attention. Here's what I discovered:

Certain fields have a considerable amount of missing data, including finishedDate, pointsEarned, purchasedItemCount, totalSpent, rewardsReceiptItemList, topBrand, and categoryCode. These missing values may cause problems in deciding if users qualify for special offers or bonus points or in determining the points earned for transactions.

Some values in the pointsEarned, purchasedItemCount, and totalSpent columns appear to be out of place, and I recommend investigating the app's processes that produce these values to determine if they are legitimate or the result of errors.

More than half of the Users data has duplicate records, which should be eliminated to ensure there are no more anomalies in the future. The date formats are inconsistent with the standard MM/DD/YYYY or similar formats, and I suggest reviewing the database to ensure that the date fields are consistently captured and stored.

To tackle the data quality issues, we require a data cleansing process that guarantees completeness, accuracy, and consistency of the data. We also intend to enforce unique constraints for each dataset and establish foreign keys to establish the relationship between datasets, ensuring consistent and accurate data across all tables.

Furthermore, to maximize the value of our data, we require more information on customer demographics and purchase history. It is particularly essential to fill the gaps and correct incomplete data. The additional information will enhance the precision and effectiveness of our analysis, enabling us to devise more targeted marketing campaigns and improve our product offerings.

I have a plan to address these data quality issues, and I'd like to discuss it with you further. Please let me know when a good time would be to schedule a meeting.

Regards,

Group 1 (Aakash Mundra, Aayush Jhingan, Yashika Chauhan