

Subject: Important Findings and Concerns: Receipts, Users, and Brands Data

Hello everyone,

I hope this message finds you all well. I recently had the chance to work with the data related to Fetch Rewards, specifically focusing on the receipts, users, and brands data. I would like to share my findings, raise some concerns, and gain a better understanding of the data. Allow me to provide an overview of my approach and highlight a few key points that require attention:

To begin, I utilized Python to read, clean, and restructure the data from the zipped JSON files into a more user-friendly format. After completing the data cleaning process, I performed several checks to ensure data quality and came across some *noteworthy observations*:

Duplicate Records in Users Data: More than half of the Users data contains duplicate records. It is imperative that we thoroughly review our database to eliminate redundant entries and prevent such anomalies from recurring in the future. It would be beneficial to understand the underlying reasons for this duplication.

Inconsistent Brand Codes: I observed that a significant number of receipts in the data contained brand codes that were not present in the brands data. This discrepancy could potentially impact brand-specific analyses and aggregations. Addressing this issue would be crucial to maintain accuracy in our analyses.

Numeric Representation of Brand Codes: It came to my attention that the brand codes were not represented numerically. Drawing from my experience with databases, I would recommend assigning numerical foreign keys to avoid potential issues arising from slight variations in string representations when joining the data.

Inconsistent Date Formats: I have observed inconsistencies in the date formats used, deviating from the usual MM/DD/YYYY or similar standard formats. I suggest conducting a comprehensive review of our database to ensure consistent and accurate capturing and storage of date fields.

Considering the importance of addressing these concerns, I would greatly appreciate the opportunity to discuss them further and schedule a meeting. Your insights and expertise in this matter would be invaluable.

Thank you for your attention to these findings and concerns. I look forward to the discussion.

Best regards,

Sai Soham Ramagiri