

Fetch Rewards

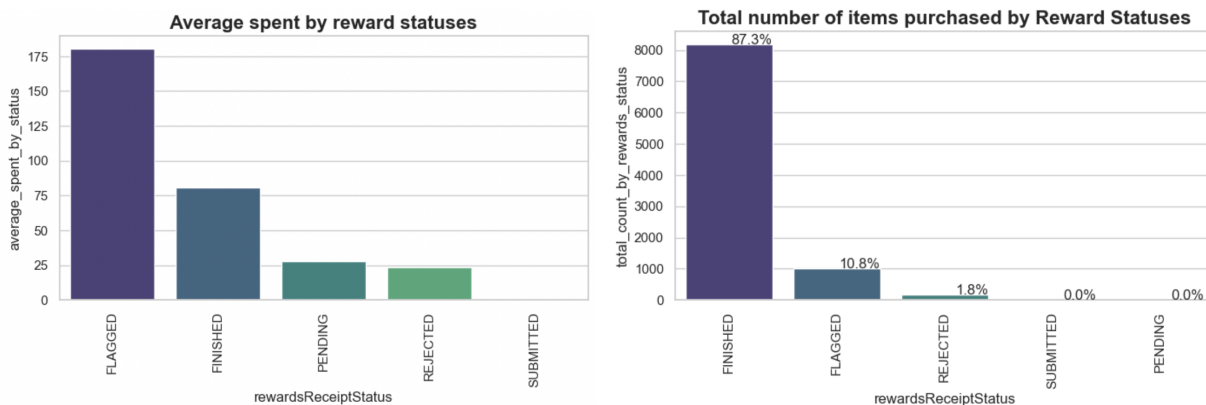
Email

Hello Member,

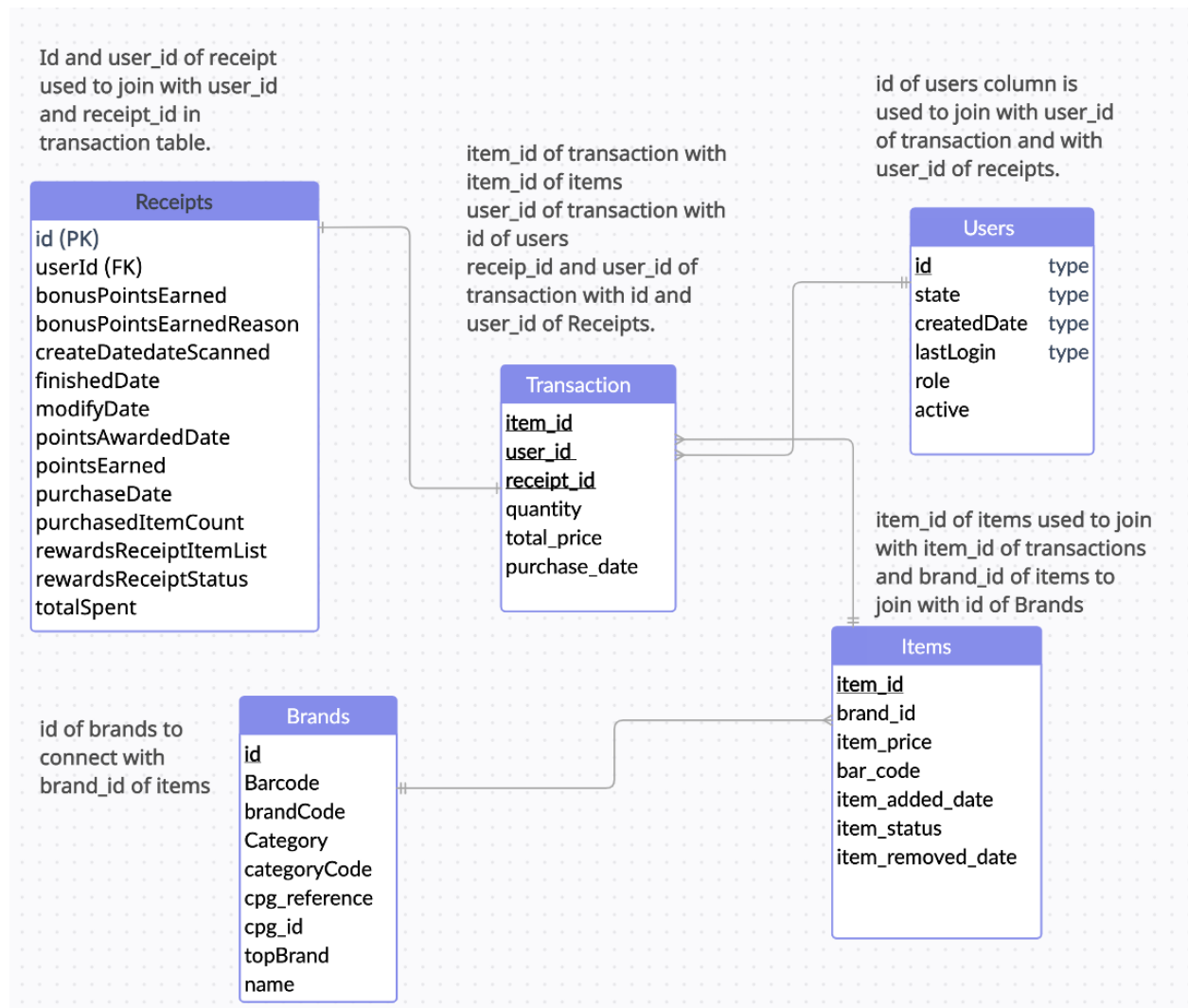
Hope you're doing really well! I have been working on the data analysis project and wanted to share the updates with you in brief in this email. I believe that it is highly important for us to try to rectify these data issues so that we maintain the standards at Fetch Works.

- There are almost 50% duplicate rows in the users' data set. This dataset lacks additional demographic information about customers (like city) which could potentially help us identify the buying habits of customers across different states and cities and thus lead us to build a personalized model.
- Furthermore, almost 60% of the columns have NaN values in the receipts dataset. Less than 10% of values in the brand ID column of the receipts data set consist of non-null values. This is making us difficult to identify the purchases of various brands across different customers. As a result, we are not able to work on different statistics across customers on the basis of their purchases made across different brands, etc. Not only this, but the “**purchasedItemCount**”, “**totalSpent**” contains a significant number of outliers. It could be interesting to cross-check it and see if there's a pattern arising due to a particular brand region, etc. Barcode on maximum data is 4011 which also is suggesting that even if barcode is missing we are imputing it with 4011. We need proper way to join Receipts table with brand table.
- Additionally, in the brands dataset category code is not different than category, although it should be, a code should be assigned. Therefore I dropped one column which is category code as it has more null values than category column.

I have also attached a graph with this email which seems to be of importance for our business.



Please take a look at the new schema proposed which tries to answer more business problems, I created two new tables transaction and items, entailing details about the transaction and item sold or purchased.



Please do let me know what you think about it and if there's anything else that you'd like me to work on, I would be more than happy to assist you with it. Please don't hesitate to reach out to me in case of any concerns with the analysis.

Hoping to hear from you soon!
Sincerely,
Nikunj Malpani