

Hi [Leader Name],

I just finished working on the Receipt-User-Brand data schema, and I have some progress to share with you. After reviewing the metadata, I would like to propose this new structured relational data model [readme figure 1]. In the jupyter notebook [Fetch\_Reward.ipynb], I parse the json files into the following five tables: *receipts*, *rewardsReceiptItemList*, *users*, *brands*, and *cpg*.

I wrote a SQL query to compare *average spend* and *the total number of items purchased* from receipts with 'rewardsReceiptStatus' of 'Accepted' or 'Rejected.' However, the 'Accepted' label is not present in the data set. Is it applicable to assume 'Finished' and 'Flagged' as accepted labels instead? I checked these are the only two labels containing both *spend* and *the number of items purchased*.

Multiple values were missing in the lastLogin column from the user table. I will need to know the March 2021 data to determine whether these null values are the monthly user churn of February 2021.

It is helpful to connect with someone from the user experience design team to understand the user journey map and further investigate the causes of user churn with data.

I am concerned with the user retention rate and would love to help improve it.

Best,  
Jason