Hello,

In response to your questions about receipts, below are my findings regarding Accepted Vs Rejected receipts and some follow-up questions regarding receipts by brands in recent months.

Based on the data provided and making an assumption that receipts with a reward receipt status of 'Finished' were Accepted, both average spend and total items purchased are higher for 'Finished' receipts than for 'Rejected' receipts:

Finished receipts contained 8184 purchased items with an average spend of $80.85.

Rejected receipts contained 173 purchased items with an average spend of $23.33.

Before moving forward with analysis of receipts by brand, I have some questions about the data. When exploring the data on receipt items as it relates to brands, I found a large number of items missing barcodes, and very few items with barcodes matched the barcodes on record for our brands. Additionally, there were some brand records that appeared to be duplicates or test cases.

In order to match brands to receipt items, is the barcode the appropriate identifier? Alternatively, there is some information available on the receipts relating to brand that can be explored further. Identifying the correct identifier or data point to capture brand name for items will help optimize our data, and if barcode is a key data point, an analysis of the duplicates on the brand side is needed.

Considering the large number of items to parse on each receipt, there may be performance concerns with generating an item-level flattened data source in this manner. I would recommend only parsing items for accepted receipts, updating data incrementally, and monitoring the performance. If performance issues do arise, it is also an option to keep the item-level data in an unstructured format within the receipt-level data source. While analysts could access the item-level data as needed, this option is a trade-off between ease of access and performance.

Thank you for your questions and review,

Lili