Subject: Data quality issues and questions with Receipts, Users and Brand data.

Hi,

After looking at the data and analyzing it, I have the following concerns and questions about the data, which I would like to convey to you.

1. A majority of the data in all the tables are missing, it can hinder with the process of analyzing the data to get some insights. I did a basic analysis and I can let you know the attributes in each table which have significant missing values.
2. The ‘rewardsReceiptItemList’ in Receipts table, should either be reformatted or made into a separate table. To get any sort of insights into that column, we need a better way to look at it.
3. I would also like to know about some of the attributes and what they signify in the data, the information is lost a little just based on the attribute names. We can talk about it and reframe the data better.
4. We can talk about the restructuring of data, I feel like it’s the best approach to better understand the data and frame it in a way that easier to analyze and derive insights from.
   1. For restructuring, we can start by breaking the receipt table into a smaller table, while looking at all of the attributes.
   2. Breaking the data would also help optimize queries when run on the data.
   3. We can also restructure the data formats, especially all the date formats which are not in human readable format but instead EPOCH formats, which is hard to understand unless they are converted.

I hope we can connect some time and talk about the data, I can also share my plans on how to deal with the data. Let me know what time works for you.

Thank you,

Saksham