

Dear Mr. X,

I hope you are doing well. Thank you for having your team send over the Fetch receipt, user, and brands data. We have successfully imported the unstructured data in to our database and have begun our preliminary diagnostics. In our preliminary examination of the data, we have noticed some inconsistencies in the data and have clarifying questions. We are hoping you or someone on your team can help answer questions relating to this data so that we may better provide key insights and metrics.

#### General

1. For what purposes and for whom were these data originally made? Are they documents that you keep in the normal course of business?
2. The date values in all of these files are 13 digit long numeric values. We would like to know in what timestamp format these are kept so that we may decode them to get date ranges.

#### Receipts Data

1. Why are some receipt items indicated as "ITEM NOT FOUND". Why would this be the case?

#### User Data

1. We only seem to have user data for users in the following states: WI, NH, AL, OH, IL, CO, SC and KY. Would it be possible to determine if a significant portion of users are being excluded because of these state limitations?
2. We have also noticed that the userIds from the receipts data do not seem to match up with the userIds from the user data.

#### Brands Data

1. We have noticed a handful of observations where a barcode is associated with more than one brand. Can we get a more thorough understanding of how brands are assigned to barcodes?
2. Can we please get a definition and understanding for what CPGs are?

Pulling the data, especially pulling the receipt items data, had long runtimes. In production, I anticipate we will implement a parallel programming algorithm to reduce these runtimes to ensure that dashboards are always up to date.

Looking forward to speaking with you.

Sincerely,

Claire