

Hi <Stakeholder name / team name>,

I performed an initial review of the data and came across a couple of issues. In order to move to a structured data model, we need to properly define each field and we will need to ensure the data is adhering to our set standards.

For example, in case an item is not found, we are currently storing incorrect barcode values and the records are incomplete. I need to know how we would like to process those records so that we can properly account for the inventory. Additionally, I would like to know if we have a process in place to find the “not found” items and load missing information in our systems so that the same issue doesn't happen again in future.

Also, the initial review showed that the receipts file has a column `rewardsreceiptsItemList` that is supposed to be a nested json file which means it has multiple columns in one column so if we want to create a relational database that this might be a hindrance for us, and it would be better to create a separate table. Also, the file receipts, brands have multiple null values and I'm not sure how we can consider those as a mean of the other or do we just delete that particular row value.

Lastly, I would like to point out some concerns I have with the relational data model:

Scaling is going to be a problem if our usage increases. In order to ensure we don't lose data in case of a server failure; I recommend maintaining a replica of the server in production environment.

Also, if we anticipate increased usage over a period of time, we should think about switching to a NoSQL model. It will allow us to keep up with the demand and reduce the cost associated with maintaining massive relational data servers.