

Receipts data model

Hi [Business Leader's Name],

I hope you're doing well. I wanted to share some thoughts on how we can optimize our new data model.

The current data model structures transactions around receipts, users, items, and time. This provides a solid foundation for analyzing customer behavior, purchase patterns, and brand performance. However, to ensure we're extracting the most value, I have a few questions:

1 - Missing Users in the Receipts Table: I've noticed that some users referenced in the receipts dataset do not appear in the users dataset. Could this information exist in another database, or might there be missing records that need to be addressed? I can provide a complete list of the missing users to address this issue.

2 - Time Utilization: Do we need additional attributes, such as fiscal year or holiday indicators, to enhance time-based analysis?

3 - Null Values in Data: There are many empty values in the dataset, the presence of null values across different tables may affect reporting accuracy in production, so we should establish a plan for handling them. For example, regarding TopBrand, the empty values should probably mean that those brand are not considered as "top".

Would love to discuss these points further to align on how we can best support the team's goals.

Regards,
Ignacio Elvira