

Subject: Questions regarding business & data quality issues.

Hi Product/Business Leader,

I am Chengwu Weng, a current Data Analyst at Fetch Rewards. I am reaching out to discuss **four topics** that will help keep in line with what business expectation and data analysis orientation and improve efficiency.

- Business data problems
- Data quality issues
- Data scaling concerns
- Other info needed

I would appreciate it if you would like to share your perspectives to these topics.

Here are a few **business data related questions** that will be helpful during exploratory data analysis and modeling process.:

- What is our typical user persona?
- What are pain points of target users?
- What are use cases of our products?
- What are the KPIs you will use to keep track on our product's performance, such as users' engagement, retention or monetization or other dimensions?

Moreover, there are three types of **data quality issues** and potential resolutions I want to share with you and solve together.

- Missing/Null Values
 - Res: we can fill NA value with 0 or drop them if those attributes are not significant in business.
- Duplicate records
 - Res: we can build a unique primary key constraint or normalize tables following 3 Normal Format in data modeling.
- Data inconsistency
 - Desc: some users can be found in receipts but are not included in users table.
 - Res: we can build foreign key index constraint or build data ETL pipeline to convert those data into integrated database or reconstruct data modeling.
 - Desc: there is not a connection between table of brand & receipt.
 - Res: we can convert reward receipt item list into a new table as a connection between brand & receipts.

Data scaling concerns

As our company continues growing, there are some anticipated data scaling concerns. Thus, it's critical that scaling is done in a way that ensures the security and integrity of our data. In this context, we will need to adopt data sharding (vertical & horizontal partitioning). Additionally, we can also use some techniques like indexing, caching to reduce data processing time and improve efficiency. During this process, we will need help from business perspective to ensure our data consistency and accuracy.

Other info needed

Since our revenue mainly comes from Affiliate Earnings paid by our brand partners, a clear requirement document will make a difference on increasing and developing our business. Thus, if there are some researches or documents that clarify what our partners want from us, how users' receipts data help partners grow sales, it will greatly helping in optimizing our data asset and guiding us to build data models/database.

Please do let me know if you have any questions regarding any of the details described above. It will be my pleasure to schedule a meeting to discuss further.

Thanks & Regards,

Chengwu Weng
Data Analyst