

The following is the email that I have constructed to send to Chris (A product/business leader)

Subject: Data Analysis discussion and Next Steps.

Hi Chris,

I hope this email finds you well. I am a Data Analyst at Fetch Rewards working on analyzing the Fetch Rewards database to find impactful insights and problematic errors present in the data. I am reaching out to you to discuss the same matter.

While analyzing the data, I encountered following questions that I feel you could answer and help me to get a better understanding of the data:

1. How do you ensure that a single receipt is scanned only once?
2. What is the end-goal of this reward program?

In addition to the above questions, I would also like to know several technicalities about the data such as:

1. How is the totalSpent calculated for each receipt? I need to know this because the totalSpent in the receipt does not match with the individual items listed in the rewardsReceiptItemList.
2. How is the pointsEarned value decided in the database? I ask this question as I observed that 45% of the receipts did not have information on the points earned.
3. How is the scanned data parsed to be stored in the database? I ask this as I noticed that there were 13 rows where the date of receipt scanning occurs before the date of item purchase.
4. What current measures are taken to ensure that the data is normalized. I need to know this as I found that there are redundant entries in the user table for the same user_id.

The above information will help me to resolve the listed pressing data quality issues that I found while analyzing the data in python. It is crucial to address these issues as they can cause severe load on the computation time and power. Not only this but the presence of these issues can create difficulties in scaling for the data engineers. Additionally, such issues can tend the data scientists to predict misleading results and can eventually lead to bad customer experience.

I hope you can help me on this matter so that by using your inputs, I can create optimal data assets that are free of data quality issues for the company. Please feel free to set a meeting for us this week to discuss more about this project and plan for the next steps.

Thank you for your time and consideration.

Sincerely,
Binny.