Dear Manager,

I hope this email finds you well. As I progress in our analysis of the dataset provided, I have come across a few areas where additional information would be greatly beneficial to ensure the accuracy and effectiveness of our work.

During my preliminary assessment with Python and SQL, I noticed several challenges, including missing values in key fields, instances of duplicate records, and some concerns regarding the integrity of the data relationships[receipts with unknown user]. These issues could potentially impact our analytical outcomes and the insights I hope to derive from the data.

To address these challenges effectively and enhance our dataset, I would appreciate your support in providing further details on the following:

* Insight into how the data was initially collected and processed would help me understand the origin of the current issues.
* Any information on the procedures currently in place for data cleaning and preprocessing would be meaningful.
* Access to related datasets might be necessary to fully resolve the relational integrity concerns. If additional data could be made available, it would greatly assist in our efforts.

Looking ahead, I am also considering the performance implications as I scale our data assets for production. Information on the frequency of data updates and the volume of new data expected would be helpful.

Thank you very much for your attention and support. I am looking forward to your guidance and any additional documentation.

Best,

Rui Xi