Fourth: Communicate with Stakeholders

Hi Team,

This is Dhruv. I've been working on reviewing and cleaning up the dataset for our current project, and I wanted to provide an update on my progress and clarify a few questions I have.

There are some issues within the data that I would like your feedback on to ensure that, what I deliver is as helpful as possible:

1. **Item-Specific Information**: Is there a complete list of items available that contains item-specific details, such as original price, original description, etc.? Having this data would help potentially reduce the amount of space needed for setting up the database.

2. **Missing Values**: I noticed several missing values in key columns, such as barcodes for many items in the item list. This could lead to challenges in establishing proper database relationships. Could you please clarify the relationship between itemNumber, brandCode, and barCode so that I can ensure the database is designed correctly?

3. **Extreme Values**: While examining the numerical columns, I found some extreme values—for instance, transactions earning 800+ points, purchasing 500+ items, or spending over 4000 dollars. Could you confirm whether these values are expected or if they might be incorrect?

4. **Data Storage Considerations**: As the dataset grows, the current relational database approach may become increasingly resource intensive. Would it be possible to explore other storage options, such as NoSQL, which might be more suitable for handling JSON-formatted data?

Please let me know your thoughts. I'd be happy to schedule a quick meeting to discuss these questions further if needed.

Looking forward to your feedback.

Best regards,
Dhruv Shetty