

## Communicate with stakeholders

Construct an email or Slack message that is understandable to a product or business leader who is not familiar with your day-to-day work. Summarize the results of your investigation. Include:

- Key data quality issues and outstanding questions about the data
- One interesting trend in the data
  - Use a finding from part 2 or come up with a new insight
- Request for action: explain what additional help, info, etc. you need to make sense of the data and resolve any outstanding issues

### Email for Stakeholders

Hi Team,

I've completed an initial investigation of the user, product, and transaction data. Here's a summary of what I found:

#### Key Data Quality Issues:

- **Missing Values:** Several columns have missing data, particularly in the users and products tables. I've filled in some of these gaps where appropriate, but we need to investigate why they're missing and whether we can recover it.
- **Inconsistent Date Formats:** The date columns were inconsistent, which I've standardized to a SQL-friendly format. We should ensure that dates are consistently collected in the future.
- **Non-ASCII Characters:** Some product and store names contained non-ASCII characters, causing encoding issues. I've cleaned these up, but we should address this at the source to avoid future problems.
- **Data Integrity:** Some user IDs in the transactions table don't exist in the user table. This suggests potential data loss or inconsistencies that need further investigation.

#### Interesting Trend:

The majority of our users (70%) are female. This presents an opportunity to tailor our marketing and product development efforts to better serve this demographic.

#### Request for Action:

- To make better sense of the data and address the outstanding issues, I need the following:
- **Information on missing data:** Can we recover any of the missing values? What are the reasons for data being missing?
- **Clarification on category hierarchy:** Is there a defined structure for the product categories?
- **Input on data collection:** Can we improve the consistency of date formats and prevent non-ASCII characters from being entered?
- **Collaboration on data integrity:** Can we work together to identify and resolve the user ID inconsistencies?

Once I have this information, I can perform a more in-depth analysis and provide further insights.