Subject: Data Quality Check and Recommendations

Dear Client,

I hope this message finds you well. We have conducted a comprehensive data quality check for the datasets you provided, and I would like to share our findings and recommendations with you.

Transactions Data

- Identified a small proportion of missing values in columns such as `online_order`, `brand`, and `product_line`.
- Investigate high missing values for the `medium` product size and product class for the `online_order` column.
- Address missing values in the `brand` column, corresponding to missing values in related columns with the same `product_id`.

New Customer List Data

- High number of missing values in the `job_industry_category`, `job_title`, `last_name`, and `DOB` columns.
- Suggested actions:
    - Contact customers with missing last names and fill in `last_name`.
    - Address missing values in `job_title` and `job_industry` columns.
    - Contact customers with gender `U` (Unidentified) to collect `DOB`.
- Identify and label columns indicating quantities (`Unnamed: 16`, `Unnamed: 17`, etc.).
- Remove or clarify columns with ambiguous information (`default`, `Unnamed: 20`, `Rank`, `Value`).
- Specify units of measurement in the `property_valuation` column.

Customer Demographics Data

- Identified missing values in several columns, including `job_industry_category`, `job_title`, `default`, `last_name`, `DOB`, and `tenure`.
- Recommended actions:
    - Fill missing values for `last_name`, `job_title`, and `job_industry_category` by contacting relevant customers.
    - Remove the `default` column, as it contains non-informative values.
    - Contact customers with gender `U` (Unidentified) to collect `DOB`.
    - Reconfirm the date of birth for the customer with an unusual age value.
    - Specify the unit of measurement for the `tenure` column (years or months).

Customer Address Data

- Main recommendation is to mention the unit of currency (e.g., million, billion) for the `property_valuation` column.

We believe that addressing these data quality issues will significantly improve the accuracy and reliability of your datasets. Please let us know how you would like to proceed, and if you have any questions or need further assistance.

Thank you for entrusting us with your data, and we look forward to working together to enhance its quality.

Best regards,

Rajat Bahuguna
KPMG