

**Subject:** Detailed Data Quality Assessment and Potential Issues

**Dear Client,**

I trust this message finds you in good health. I would like to provide you with an extensive data quality assessment of the datasets you've shared with our team. In our effort to ensure data accuracy, completeness, consistency, relevance, and validity, we've performed a comprehensive analysis for each dataset, detailing the issues found in each category.

Datasets	Total No of Records	Unique ID records
Transactions	20000	3427
Customer Demographic	4000	4000
Customer Addresses	4000	4000

**Transactions Sheets**

Issue	Accuracy	Completeness	Consistency	Relevance	Validity
Online_order blanks	Ensure orders have valid online order entries	Ensure no orders are missing		Ensure relevant orders are included	Ensure online order format is valid
Canceled Orders	Exclude canceled orders	Ensure canceled orders are not included		Exclude irrelevant canceled orders	Ensure order statuses are valid
Brand blanks	Eliminate entries with blank brands	Ensure all products have a brand			Ensure brands are valid
Data type issues	Correct data type inconsistencies				Ensure data types are valid
Currency in list price	Ensure uniform currency format	Ensure all prices have valid currency format			Ensure currency is valid

Date formatting	Standardize date format in 'product first sold'	Ensure all dates are in valid format			Ensure dates are valid
Gross profit calculation	Calculate gross profit accurately	Ensure gross profit values are complete	Ensure consistent profit calculations	Ensure relevant profit values are used	Ensure profit calculations are valid
Profit margin and markup %	Compute profit margin and markup percentage	Ensure profit margin and markup values are complete	Ensure consistent calculations	Ensure relevant calculations are used	Ensure calculations are valid

### Customers Demographic Sheet

Issue	Accuracy	Completeness	Consistency	Relevance	Validity
Gender replacements	Replace 'M' with 'Male', 'Femal' with 'Female'			Ensure gender entries are relevant	Ensure gender entries are valid
Date of birth blanks	Remove blank entries and entries prior to 1883	Ensure no blank dates of birth are present		Exclude irrelevant or invalid dates	Ensure date of birth entries are valid
Age calculation	Calculate age based on date of birth	Ensure age values are complete	Ensure consistent age calculations	Ensure relevant age values are used	Ensure age calculations are valid
Job title and industry	Eliminate blank entries	Ensure no blank job titles or industries exist		Exclude irrelevant job titles/industries	Ensure job titles/industries are valid
Deceased indicator	Filter out deceased customers	Ensure deceased indicators are complete	Ensure consistent deceased indicators	Ensure relevance of active customers	Ensure deceased indicators are valid
Default column removal	Remove any unnecessary	Ensure default columns are removed	Ensure consistency in		Ensure column removal is valid

	default columns		column removal		
--	--------------------	--	-------------------	--	--

### Customers Address Sheet

Issue	Accuracy	Completeness	Consistency	Relevance	Validity
Customer ID blanks	Eliminate entries with blank customer IDs	Ensure all customer IDs are complete	Ensure consistent removal of blank IDs		Ensure customer IDs are valid
State abbreviation standard	Replace 'New South Wales' with 'NSW', 'Victoria' with 'VIC'	Ensure all state abbreviations are standardized	Ensure consistent state abbreviation changes	Ensure relevant state abbreviations	Ensure state abbreviations are valid

### New Customer Sheet

Issue	Accuracy	Completeness	Consistency	Relevance	Validity
Blank first and last names	Remove entries with blank first or last names	Ensure complete first and last names	Ensure consistent removal of blank names		Ensure names are valid
Sales format conversion	Convert 'past 3 years sales' to a number format	Ensure all sales values are in number format	Ensure consistent conversion of sales values	Ensure relevance of sales data	Ensure sales data is valid
Blank job industry and title	Eliminate entries with blank job industry/title	Ensure no blank job titles or industries exist	Ensure consistent removal of blank entries	Exclude irrelevant job titles/industries	Ensure job titles/industries are valid
Blank wealth segment, rank	Remove entries with blank wealth segment/rank	Ensure complete wealth segment/rank values	Ensure consistent removal of blank values	Exclude irrelevant wealth data	Ensure wealth data is valid

Blank property value	Eliminate entries with blank property values	Ensure complete property values	Ensure consistent removal of blank values	Exclude irrelevant property data	Ensure property data is valid
----------------------	--	---------------------------------	---	----------------------------------	-------------------------------

We believe addressing these issues meticulously will significantly enhance the quality and reliability of your data. Should you require further clarification or assistance with any of these points, please do not hesitate to reach out to us.

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

Best Regards,

Ali Haider (alishaider.ah1510@gmail.com)

---