Data Quality Visualization Report

Summary

Completeness

Uniqeness

Accuracy

Validity

Consistency

Recommendations

SpikeUp

Deloitte.

Data Quality Summery

Completeness :Data is not complete

97.85

Accuracy :Data is not accurate

89.00

Consistency :Data is not consistent

79.84

Quality vs Threshold

Uniqueness :Data is not unique

96.28

Validity :Data is not valid

92.74

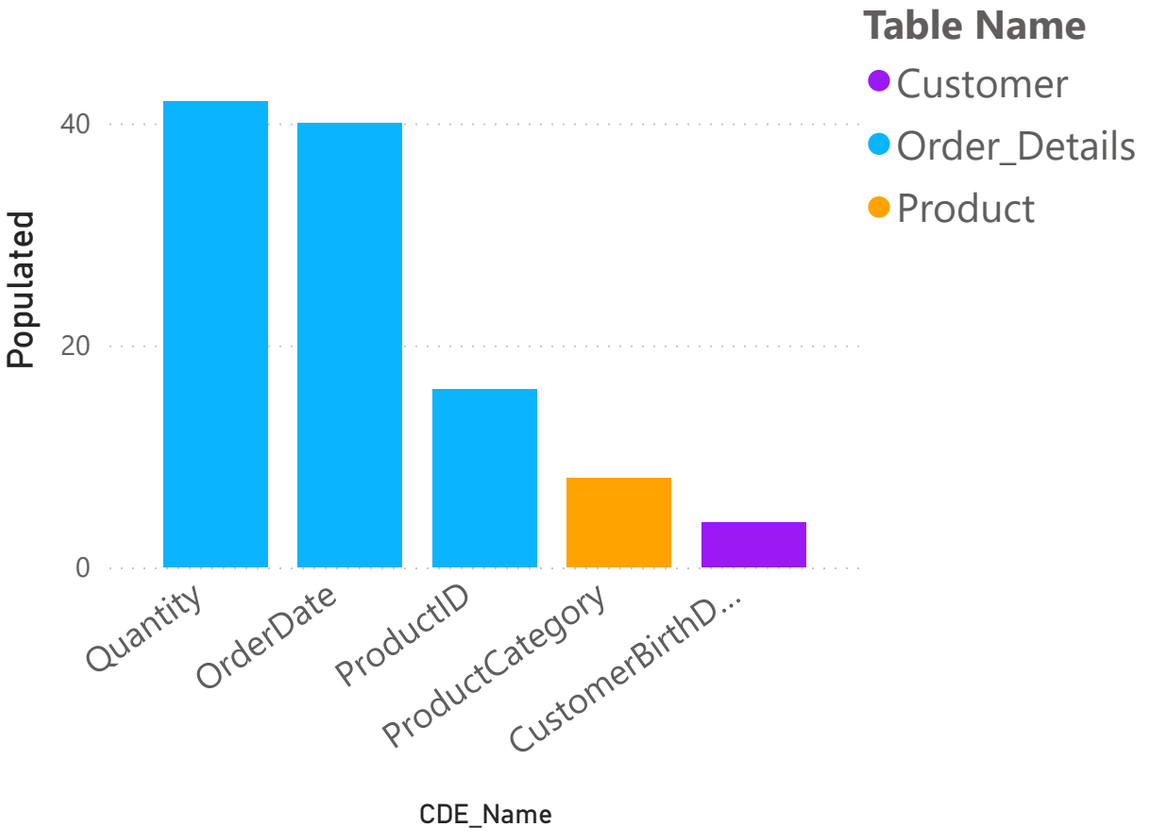
Data Quality (Completeness)

26K
Total Rows

97.85	0%	2.15%
Completeness	Threshold	Error

R_01

The CDE must be populated



Data Quality (Uniqueness)

96.28	0%	3.72%
Uniqueness	Threshold	Error

R_02

CustomerID must be unique in Customer table

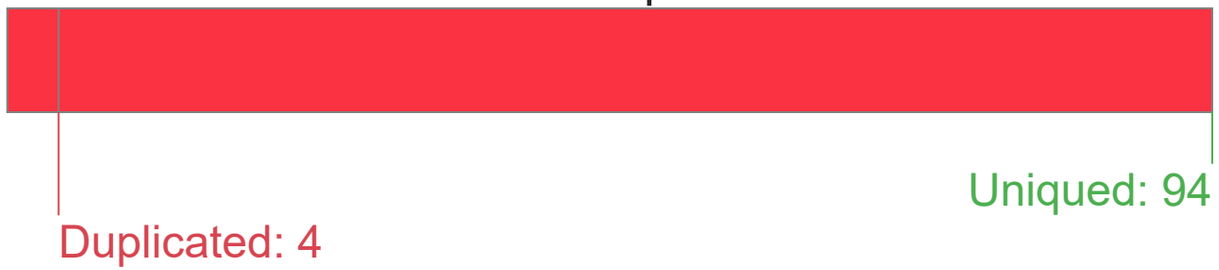
D_03
CDE_ID

0% Threshold

7.00% Error

R_03

ProductID must be unique in Product table



D_04
CDE_ID

0% Threshold

4.00% Error

R_04

OrderID must be unique in Order_Details table

D_01
CDE_ID

0% Threshold

0.16% Error

Data Quality (Accuracy)

89.00	5%	11.00%
Accuracy	Threshold	Error

R_05

D_08

CDE_ID

ListedPrice of the same product category should fall within ± 2 standard deviations

5% Threshold

5.00% Error

R_06

D_09

CDE_ID

DiscountedPrice of the same product category should fall within ± 2 standard deviations

5% Threshold

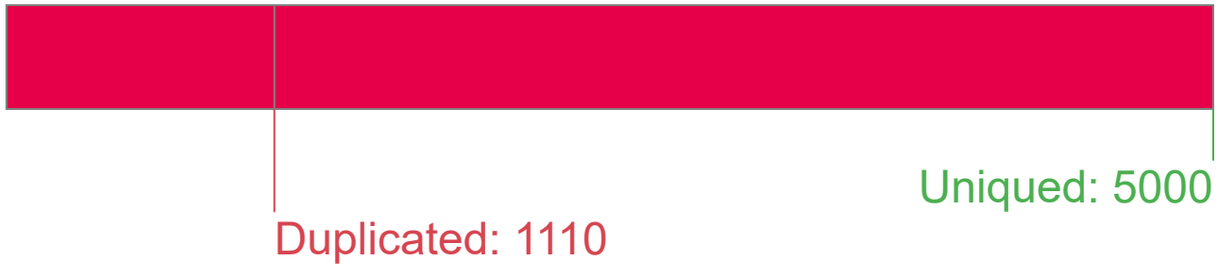
6.00% Error

R_07

D_05

CDE_ID

Total Quantity of sales per month should fall within ± 3 standard deviations

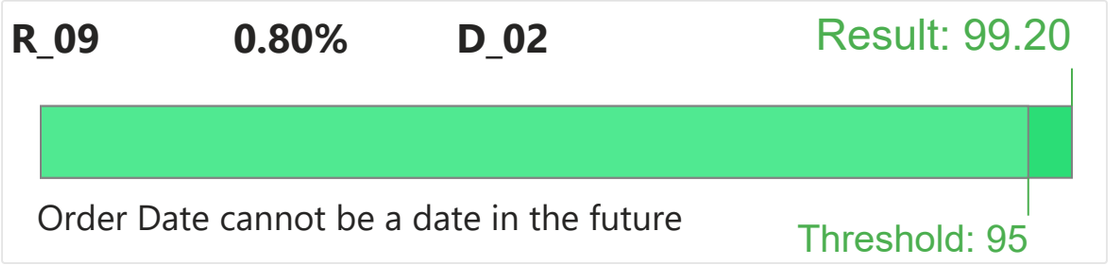


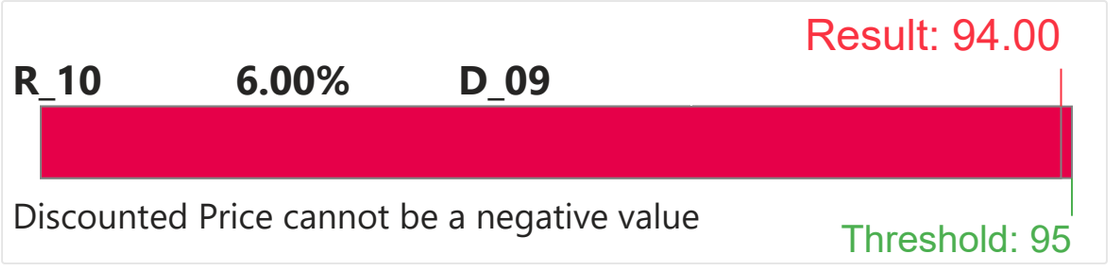
5% Threshold

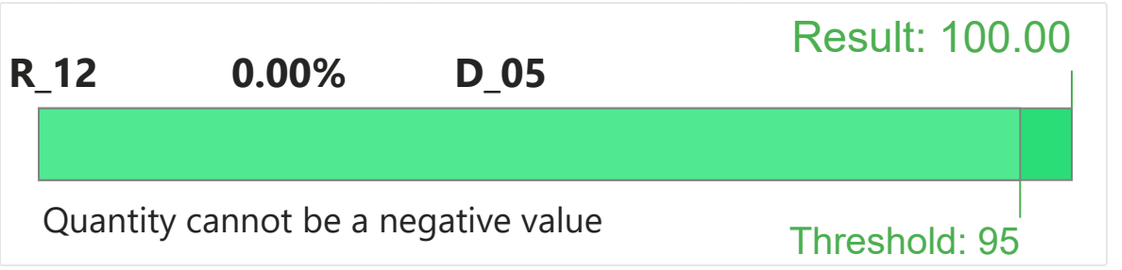
22.00% Error

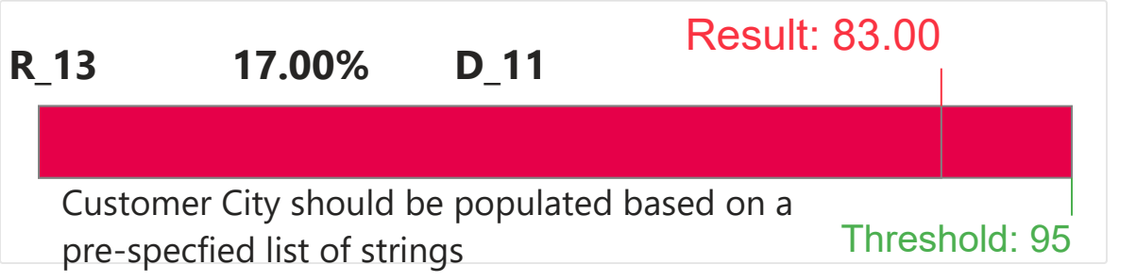
Data Quality (Validity)

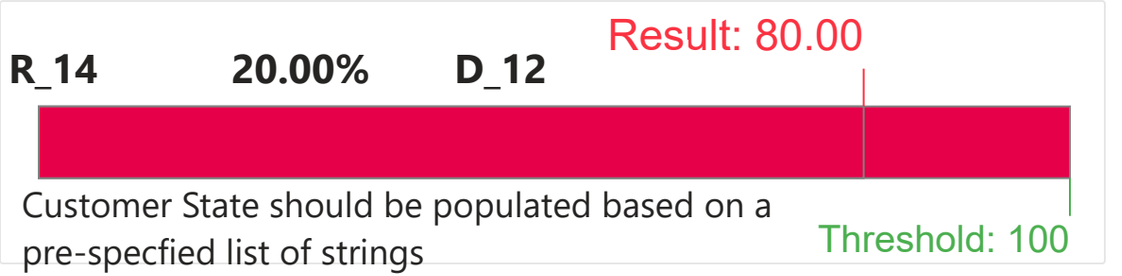
92.78	4%	7.23%
Validity	Threshold	Error

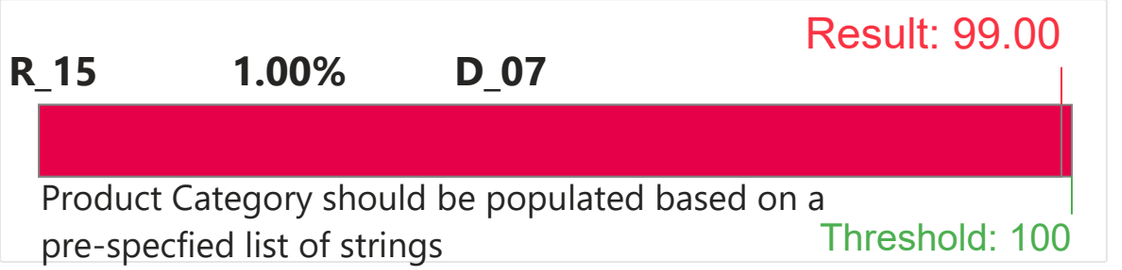












Data Quality (Consistency)

79.84	5%	20.17%
Consistency	Threshold	Error

R_16

DiscountedPrice cannot be higher than ListedPrice for one specific product

D_08; D_09

CDE_ID

5%
Threshold

17.00%
Error

R_17

CustomerCity should be consistent with CustomerState for one specfic customer

D_11; D_12

CDE_ID

5%
Threshold

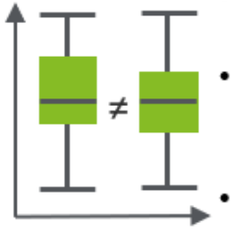
23.33%
Error

Recommendations

Completeness

- Conduct a comprehensive data audit to identify missing information.
- Establish clear data entry protocols and responsibilities.
- Implement regular data validation checks to ensure completeness.

Accuracy



- Establish data entry standards and guidelines for the team.
- Implement validation rules and data validation checks during data entry.
- Conduct periodic data accuracy audits to identify and rectify errors.

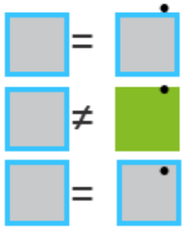
Uniqueness

- Implement unique constraints on key fields to prevent duplicate records.
- Regularly audit and cleanse the database to remove duplicate entries.

Validity

Define data validation rules to ensure that data conforms to predefined standards.
Conduct regular data validation checks and correct any data that does not meet the defined criteria.

Consistency



Develop and enforce consistent naming conventions and data formats.
Establish data update and modification protocols to maintain data consistency.
Implement version control and change tracking mechanisms for data updates.

Thank you!

Q&A