

## Data Quality Report

timestamp: 2026-01-22T21:17:01.350327

dataset\_info:

rows: 10050

columns: 5

memory\_usage\_mb: 1.5

schema:

has\_required\_columns: True

missing\_columns: []

extra\_columns: ['extra\_column']

total\_columns: 5

missing\_values:

total\_missing: 100

columns\_with\_missing: {'rating': 100}

missing\_percentages: {'rating': 1.0}

rows\_with\_any\_missing: 100

duplicates:

total\_duplicates: 278

duplicate\_percentage: 2.77

unique\_rows: 10000

rating\_range:

column: rating

expected\_range: [1, 5]

actual\_min: 0.0

actual\_max: 6.0

out\_of\_range\_count: 20

out\_of\_range\_percentage: 0.2

data\_types:

type\_mismatches: {'rating': {'expected': 'int', 'actual': 'float64'}}

all\_types\_correct: False

statistics:

unique\_users: 1000

unique\_items: 500

avg\_rating: 2.9785929648241205

rating\_std: 1.418557998392824

quality\_score: 81.03