■ DataScribe Analysis Report

Generated on: 2025-08-31 23:04:18

■ *Available Data Keys:*

- overview
- data_quality
- statistics
- univariate
- bivariate
- multivariate
- target_analysis
- insights

■ Table of Contents

■ Dataset Overview

Metric	Value
Rows	1599
Columns	12
Memory Usage	0.15 MB
Numerical Columns	12
Categorical Columns	0
Datetime Columns	0

■ Data Quality Assessment

Metric	Value
Quality Score	25.0%
Missing Values	0
Duplicate Rows	240
Constant Columns	0

■ Statistical Summary

numerical

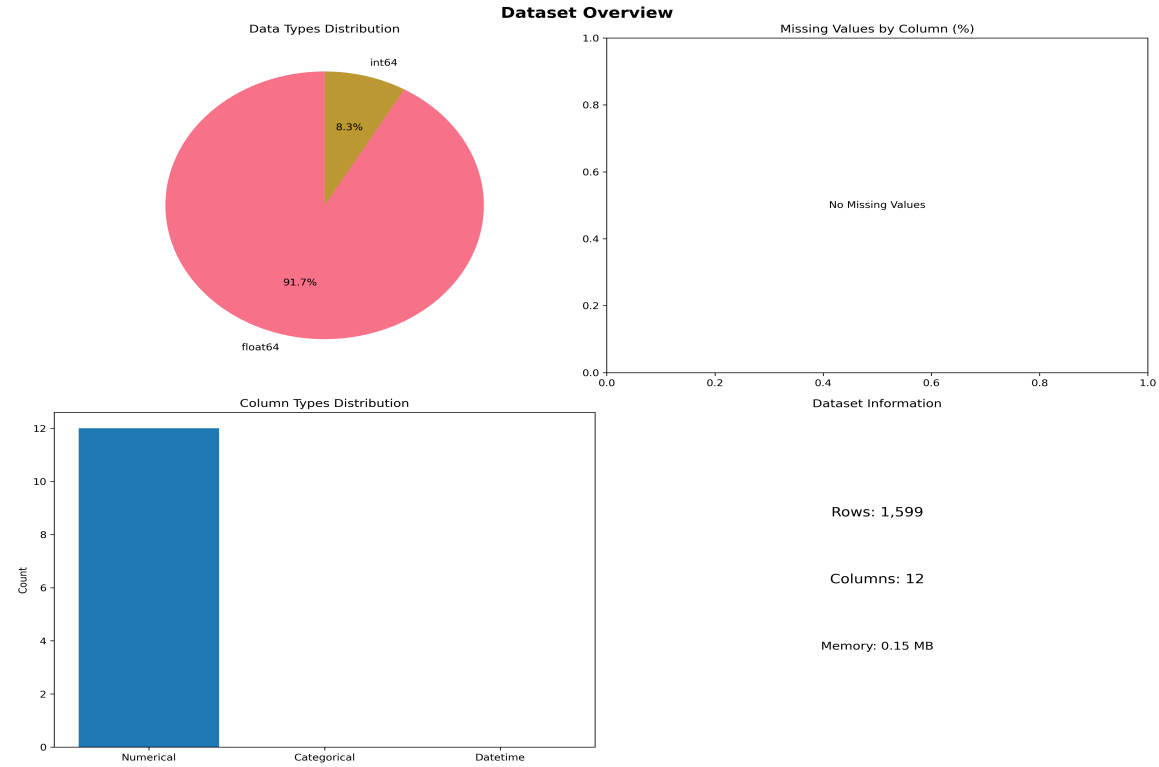| Value |
| --- |
| {'count': 1599.0, 'mean': 8.31963727329581, 'std': 1.7410963181276953, 'min': 4.6, '25%': 7.1, '50%': 7.9, '75%': 9.2, 'max': 15.9} |
| {'count': 1599.0, 'mean': 0.5278205128205128, 'std': 0.17905970415353537, 'min': 0.12, '25%': 0.39, '50%': 0.52, '75%': 0.64, 'max |
| {'count': 1599.0, 'mean': 0.2709756097560976, 'std': 0.19480113740531857, 'min': 0.0, '25%': 0.09, '50%': 0.26, '75%': 0.42, 'max': |
| {'count': 1599.0, 'mean': 2.53880550343965, 'std': 1.4099280595072798, 'min': 0.9, '25%': 1.9, '50%': 2.2, '75%': 2.6, 'max': 15.5} |
| {'count': 1599.0, 'mean': 0.08746654158849279, 'std': 0.0470653020100901, 'min': 0.012, '25%': 0.07, '50%': 0.079, '75%': 0.09, 'm |
| {'count': 1599.0, 'mean': 15.8749218261411339, 'std': 10.460156969809725, 'min': 1.0, '25%': 7.0, '50%': 14.0, '75%': 21.0, 'max': 72 |
| {'count': 1599.0, 'mean': 46.46779237023139, 'std': 32.895324478299074, 'min': 6.0, '25%': 22.0, '50%': 38.0, '75%': 62.0, 'max': 28 |
| {'count': 1599.0, 'mean': 0.9967466791744841, 'std': 0.0018873339538425554, 'min': 0.99007, '25%': 0.9956, '50%': 0.99675, '75% |
| {'count': 1599.0, 'mean': 3.3111131957473416, 'std': 0.15438646490354277, 'min': 2.74, '25%': 3.21, '50%': 3.31, '75%': 3.4, 'max': |
| {'count': 1599.0, 'mean': 0.6581488430268917, 'std': 0.16950697959010996, 'min': 0.33, '25%': 0.55, '50%': 0.62, '75%': 0.73, 'max |
| {'count': 1599.0, 'mean': 10.422983114446529, 'std': 1.0656675818473946, 'min': 8.4, '25%': 9.5, '50%': 10.2, '75%': 11.1, 'max': 14 |
| {'count': 1599.0, 'mean': 5.6360225140712945, 'std': 0.807569439734705, 'min': 3.0, '25%': 5.0, '50%': 6.0, '75%': 6.0, 'max': 8.0} |

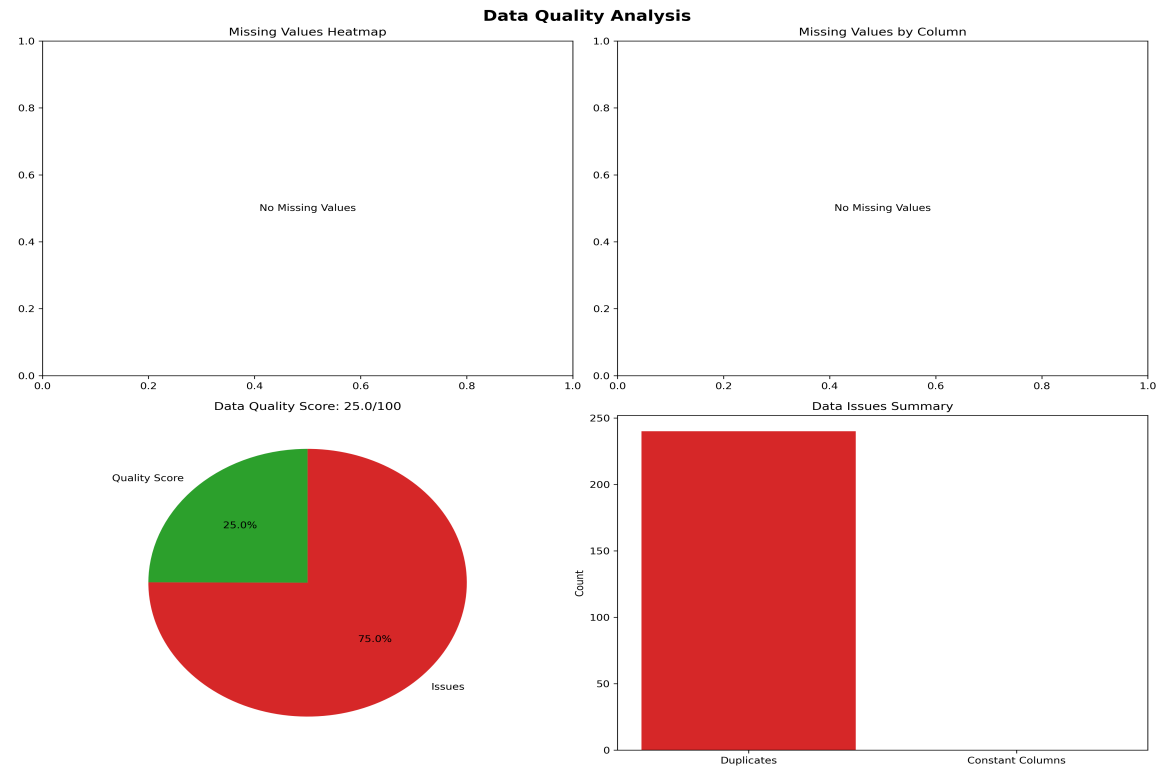|	
	0.9827514413284587
	1.1321433977276252
	0.6715925723840199
	1.2255422501791422
	0.3183372952546368
	-0.7889975153633966
	4.54065542590319
	28.617595424475443
	5.680346571971722
	41.7157872475766
ss	1.250567293314441
	2.023562045840575
ss	1.515531257594554
	3.8098244878645744
	0.07128766294927483
	0.9340790654648123
	0.19368349811284427
	0.806942508246574
	2.4286723536602945
	11.720250727147674
	0.8608288068888538
	0.2000293113417695
	0.21780157547366327
	0.2967081197538759

■ Visualizations

■ *Available Visualizations:*

- overview: static\050bedcb-834b-402f-93a0-63e42d4254e3_overview.png
- data_quality: static\050bedcb-834b-402f-93a0-63e42d4254e3_data_quality.png
- univariate: static\050bedcb-834b-402f-93a0-63e42d4254e3_univariate.png
- bivariate: static\050bedcb-834b-402f-93a0-63e42d4254e3_bivariate.png
- multivariate: static\050bedcb-834b-402f-93a0-63e42d4254e3_multivariate.png
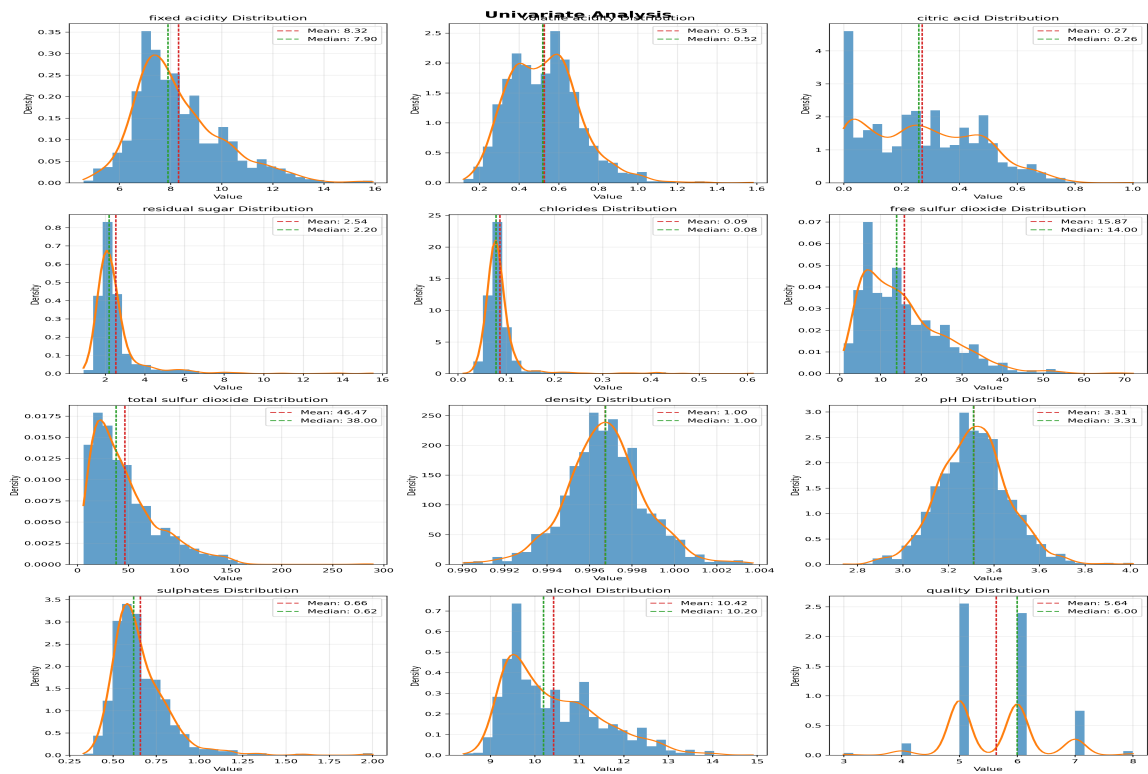- target_analysis: static\050bedcb-834b-402f-93a0-63e42d4254e3_target_analysis.png
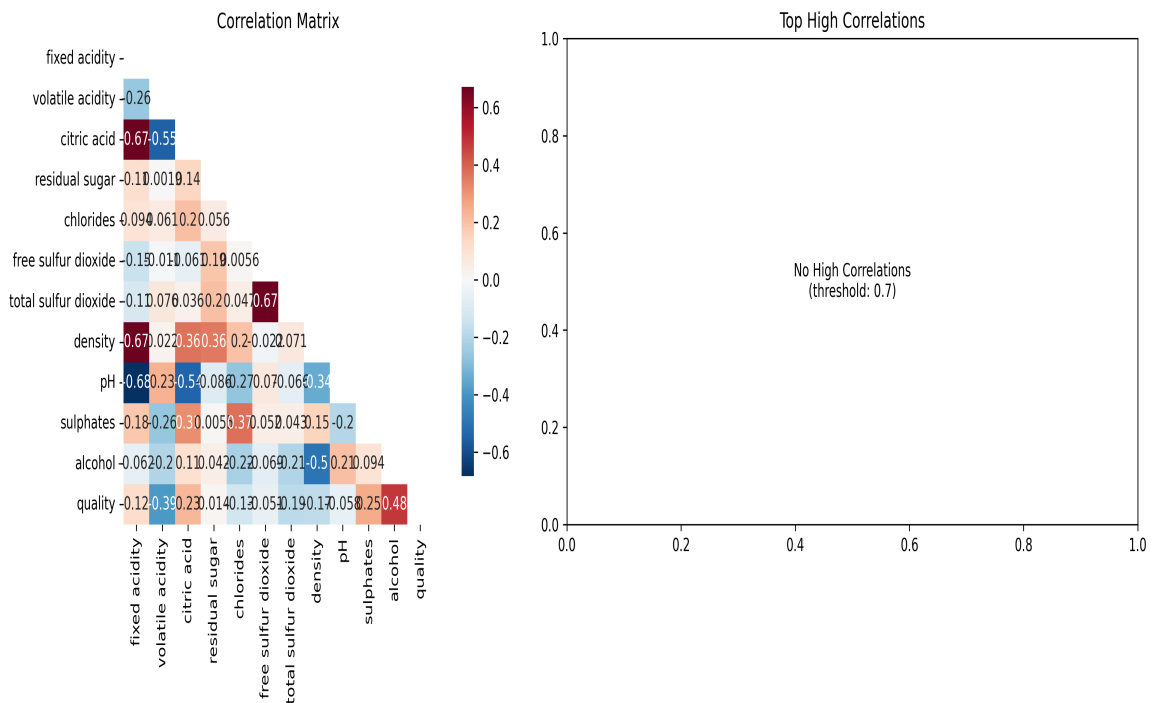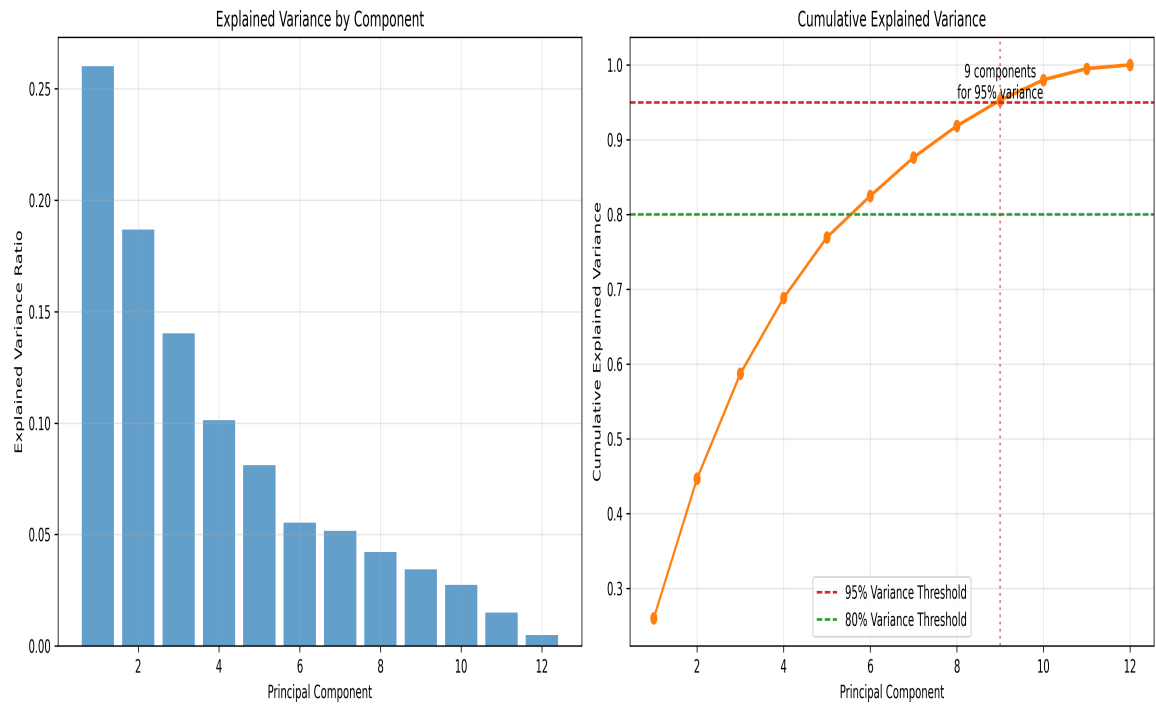
Overview

Dataset Overview

Data Types Distribution

int64 8.3%

float64 91.7%

Missing Values by Column (%)

No Missing Values

Column Types Distribution

Dataset Information

Rows: 1,599

Columns: 12

Memory: 0.15 MB

Data Quality

Data Quality Analysis

Missing Values Heatmap

No Missing Values

Missing Values by Column

No Missing Values

Data Quality Score: 25.0/100

Quality Score 25.0%

Issues 75.0%

Data Issues Summary

Univariate

Univariate Analysis

Bivariate

Correlation Analysis

Correlation Matrix

Top High Correlations

No High Correlations
(threshold: 0.7)

Multivariate

Principal Component Analysis

Explained Variance by Component

Cumulative Explained Variance

Target Analysis

Target Variable Analysis: quality

quality Distribution

Target Statistics

Mean: 5.64

Median: 6.00

Std: 0.81

Range: 3.00 - 8.00
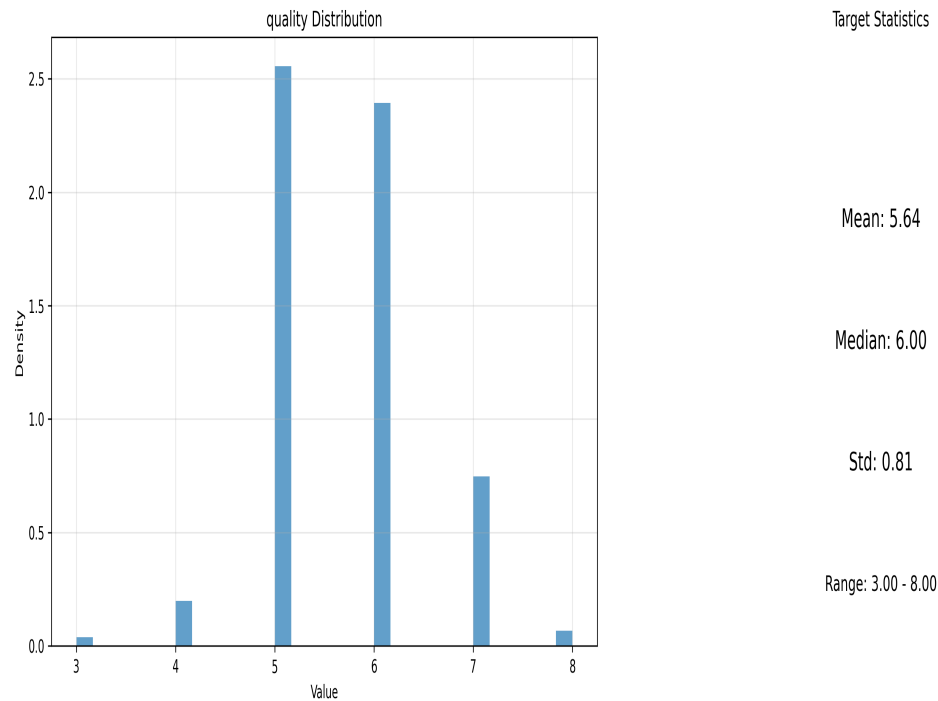
■ AI Insights & Recommendations

Data Quality Insights:

• Found 240 duplicate rows that should be removed

Feature Insights:

• Dataset has 12 numerical features for analysis

Recommendations:

• Consider data cleaning before analysis
• Consider dimensionality reduction for high-dimensional data