PhenoQC Quality Control Report

Source file: kidney_disease.csv

Summary

Imputation Strategy	Mean
Schema Validation Score	100.00%
Missing Data Score	90.96%
Mapping Success Score	0.00%
Overall Quality Score	63.65%

Imputation Settings

Global Strategy	None
Global Params	None
Tuning Enabled	False

Data Quality Scores:

Schema Validation Score: 100.00%
Missing Data Score: 90.96%
Mapping Success Score: 0.00%
Overall Quality Score: 63.65%

Schema Validation Results

Format Validation: True
Duplicate Records: No issues found.
Conflicting Records: No issues found.
Integrity Issues: No issues found.
Referential Integrity Issues: No issues found.
Anomalies Detected: 42 issues found.
**Invalid Mask:** 400 issues found.

Missing Data Summary

Column	Missing Count
rbc	152
rbcc	131
wbcc	106
pot	88
sod	87
pcv	71
pc	65
hemo	52
su	49
sg	47
al	46
bgr	44
bu	19
sc	17
bp	12
age	9
ba	4
pcc	4
cad	2
dm	2
htn	2
ane	1
appet	1
pe	1

Records Flagged for Missing Data: 242

Ontology Mapping Summary

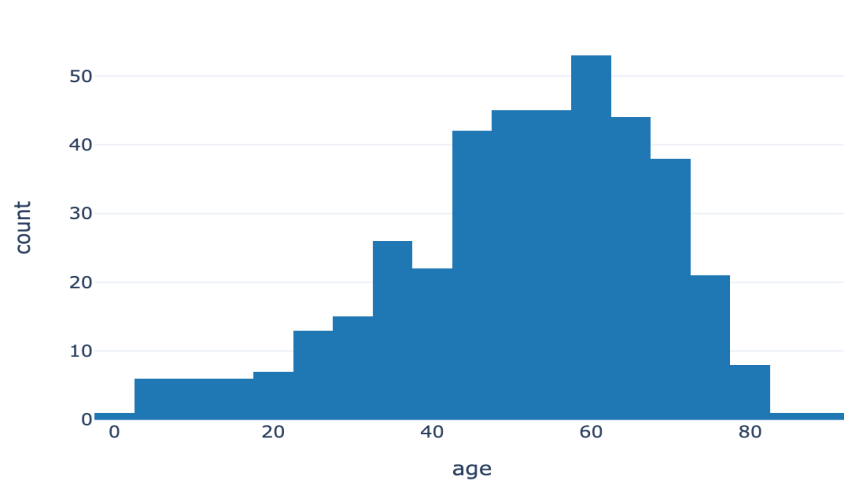| Ontology	Total Terms	Mapped	Success Rate
HPO	0	0	0.00%
DO	0	0	0.00%
MPO	0	0	0.00%

Visualizations

Missingness Heatmap
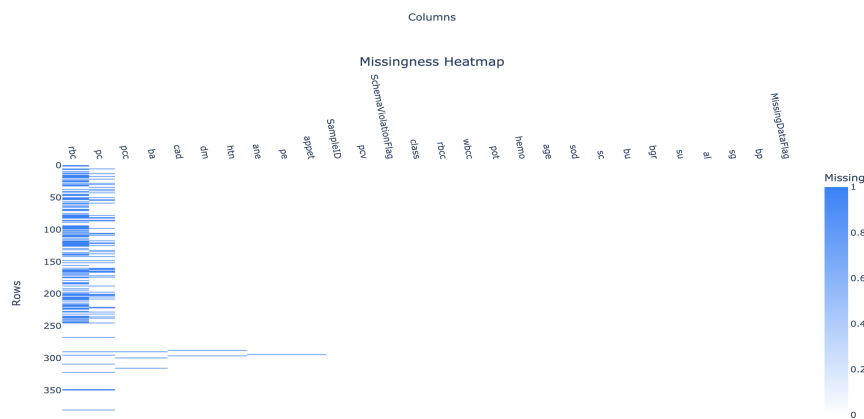
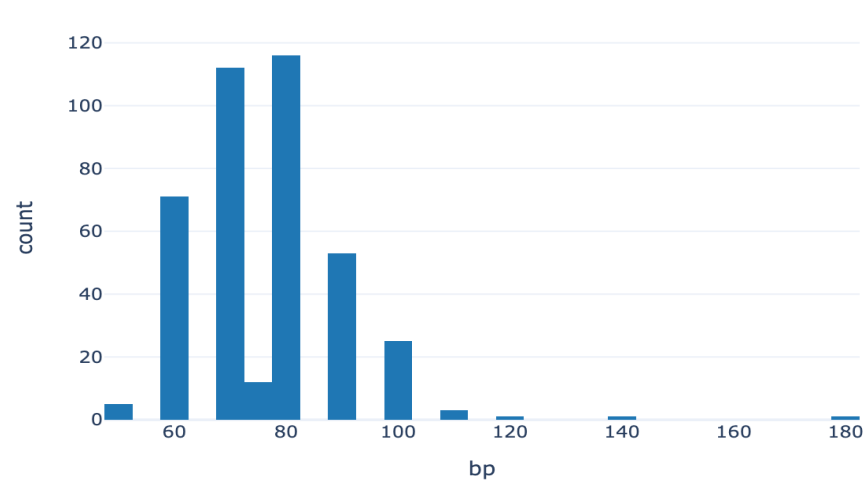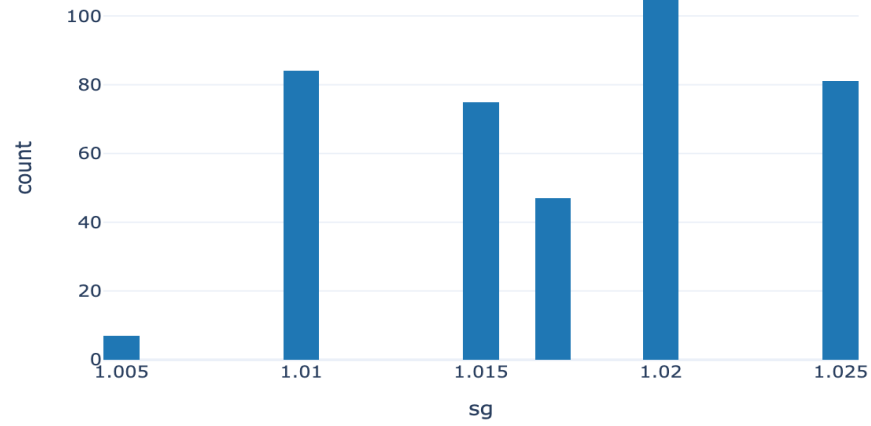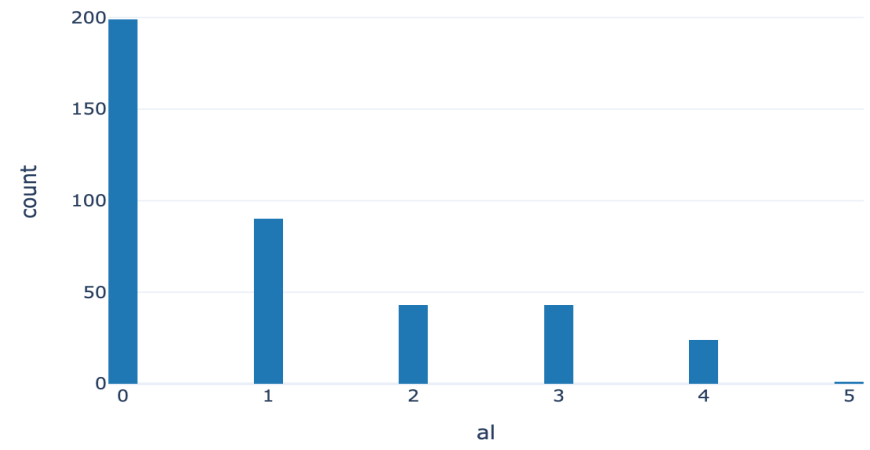Percentage of Missing Data by Column
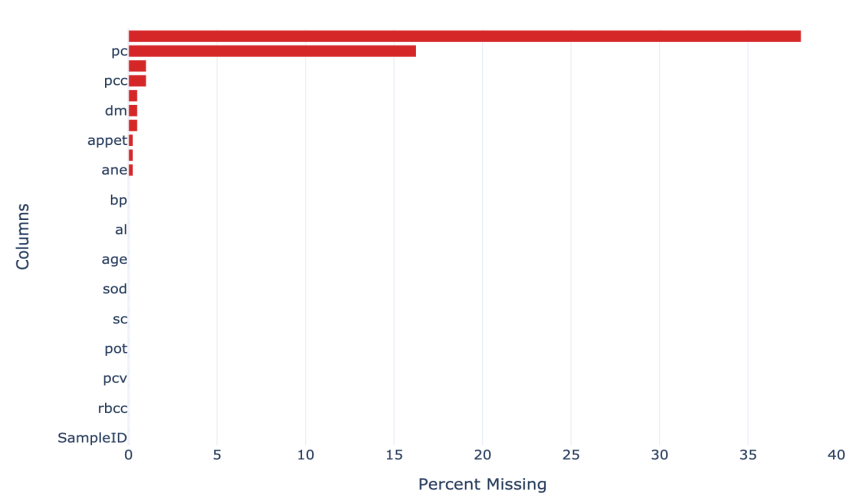
Distribution of age

Distribution of bp

Distribution of sg

Distribution of al

Distribution of su