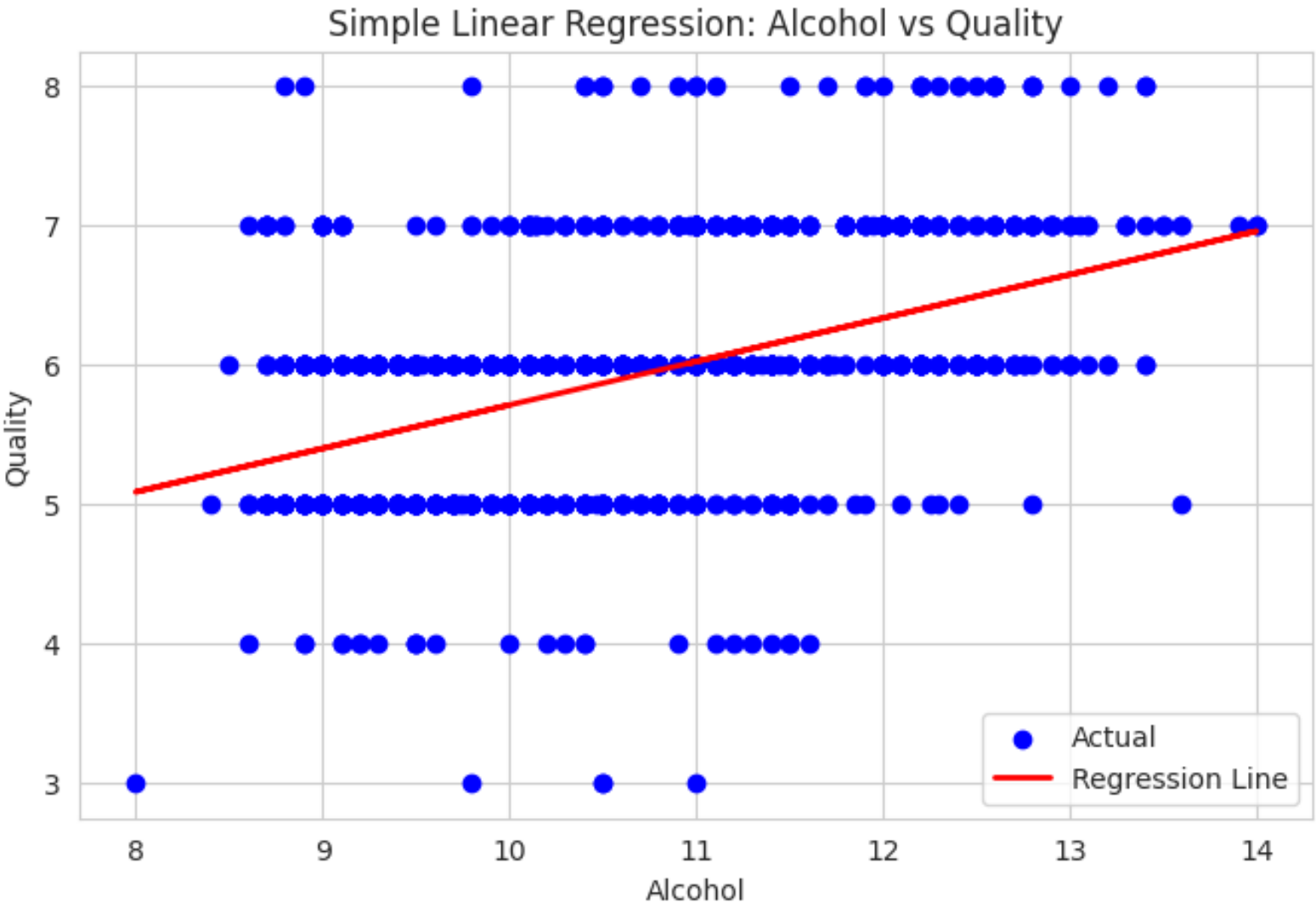
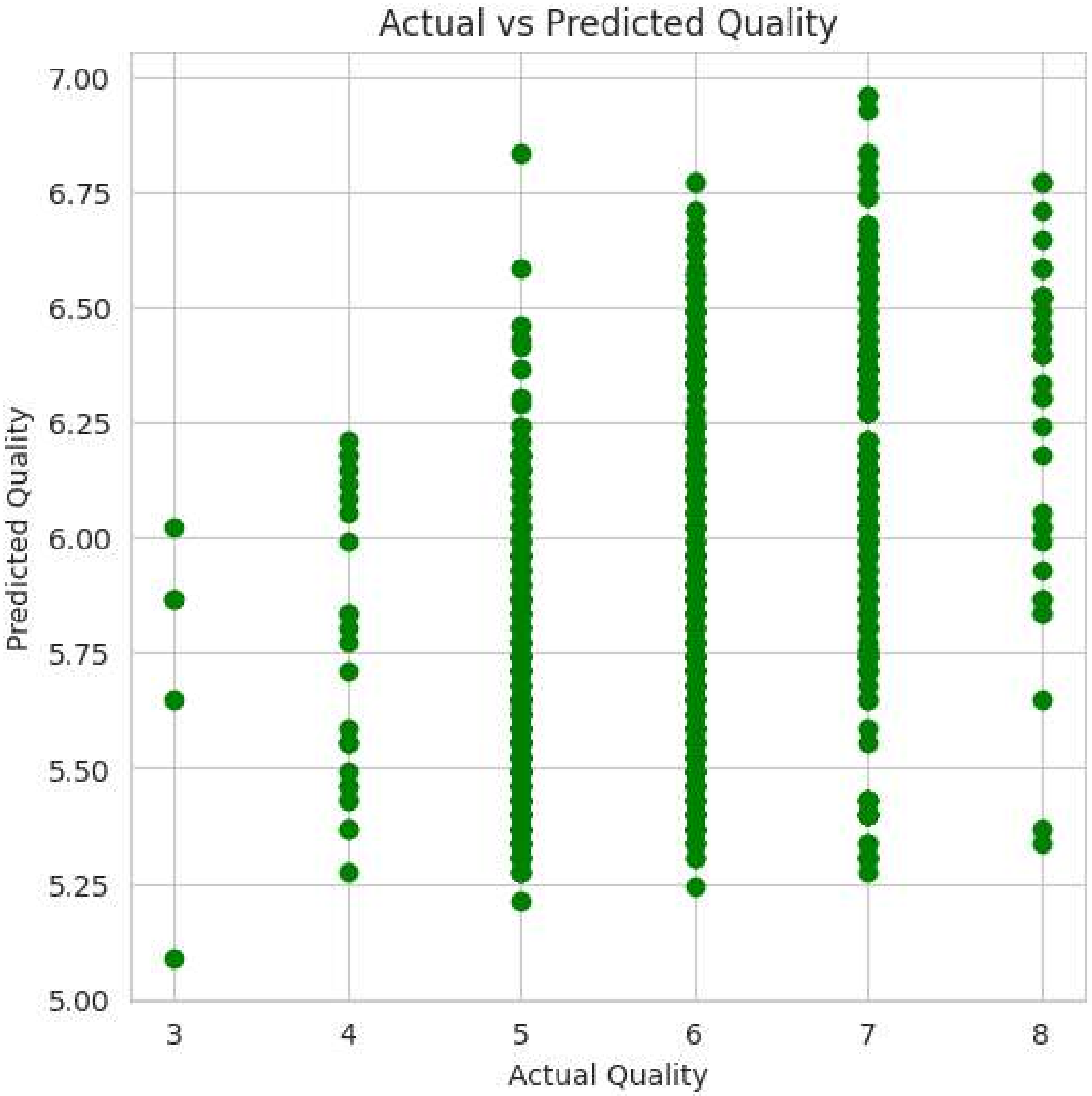
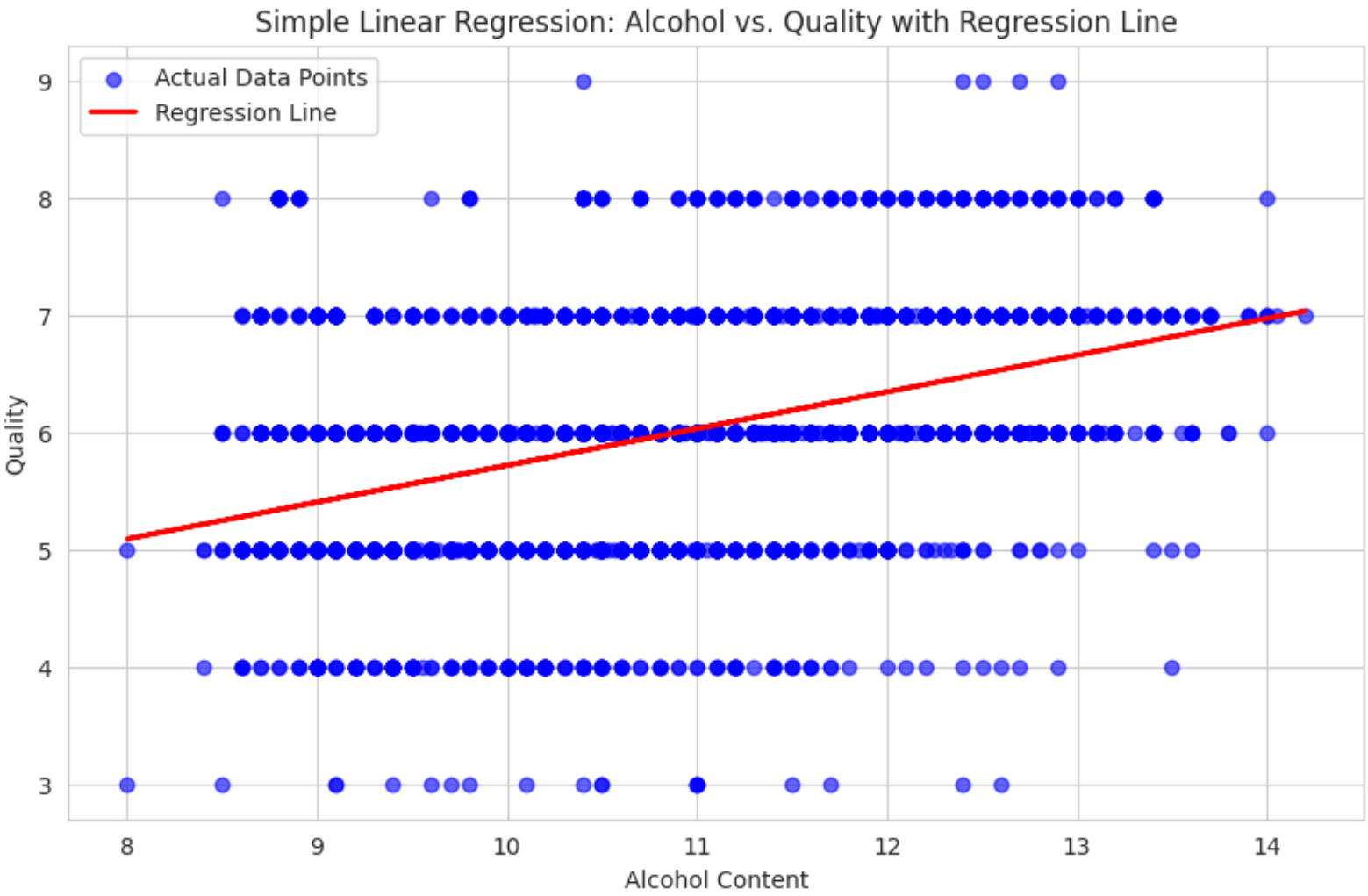
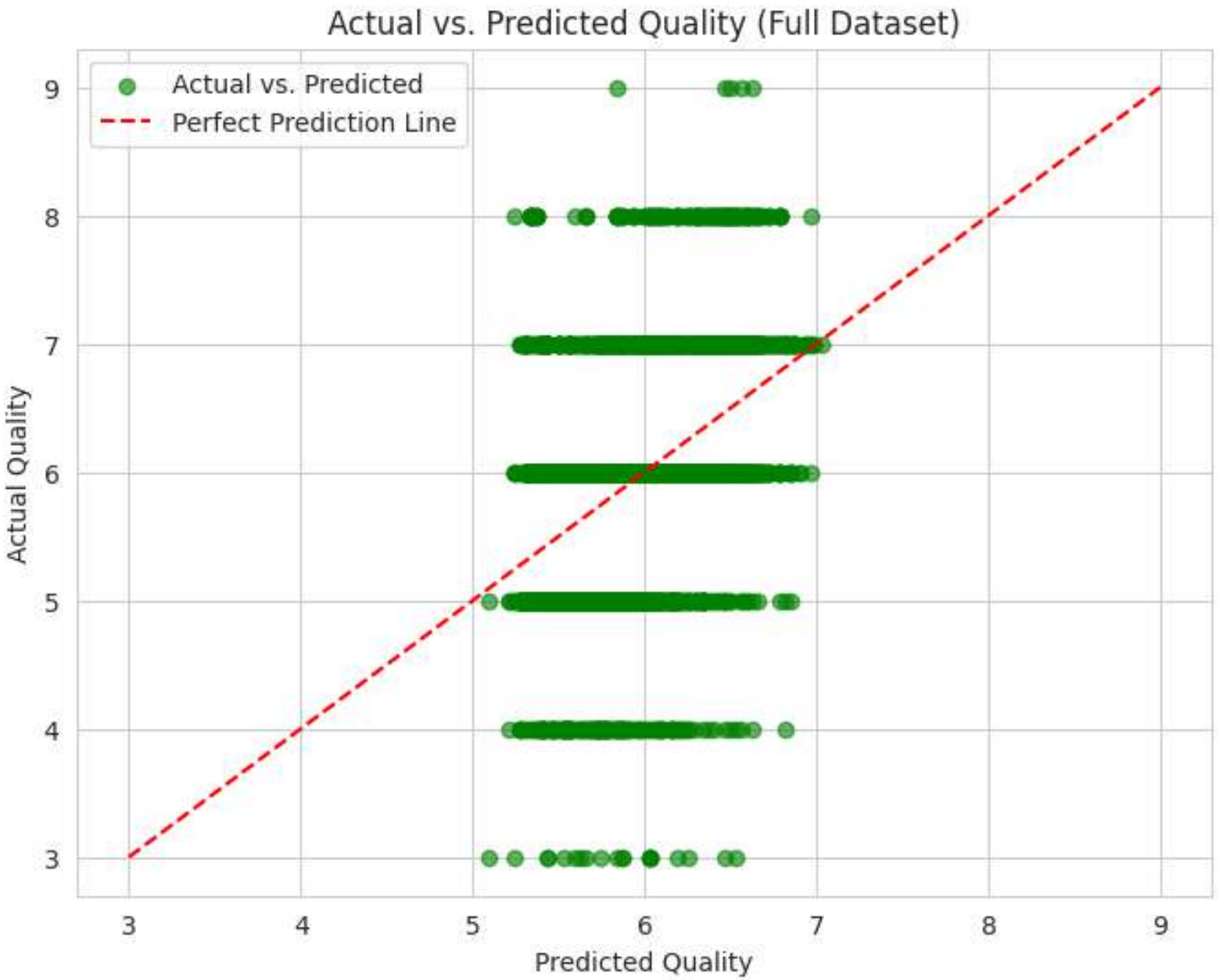
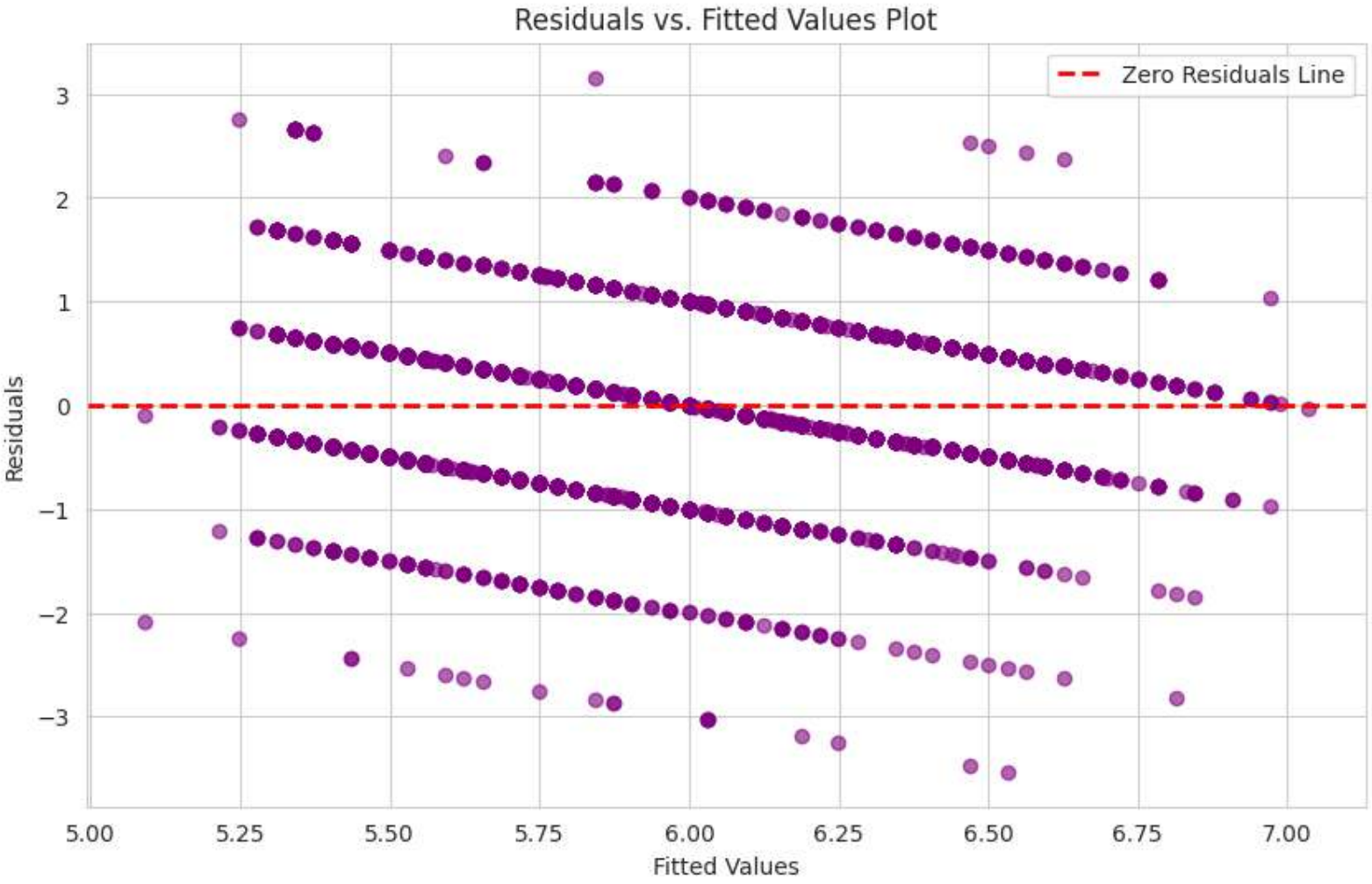
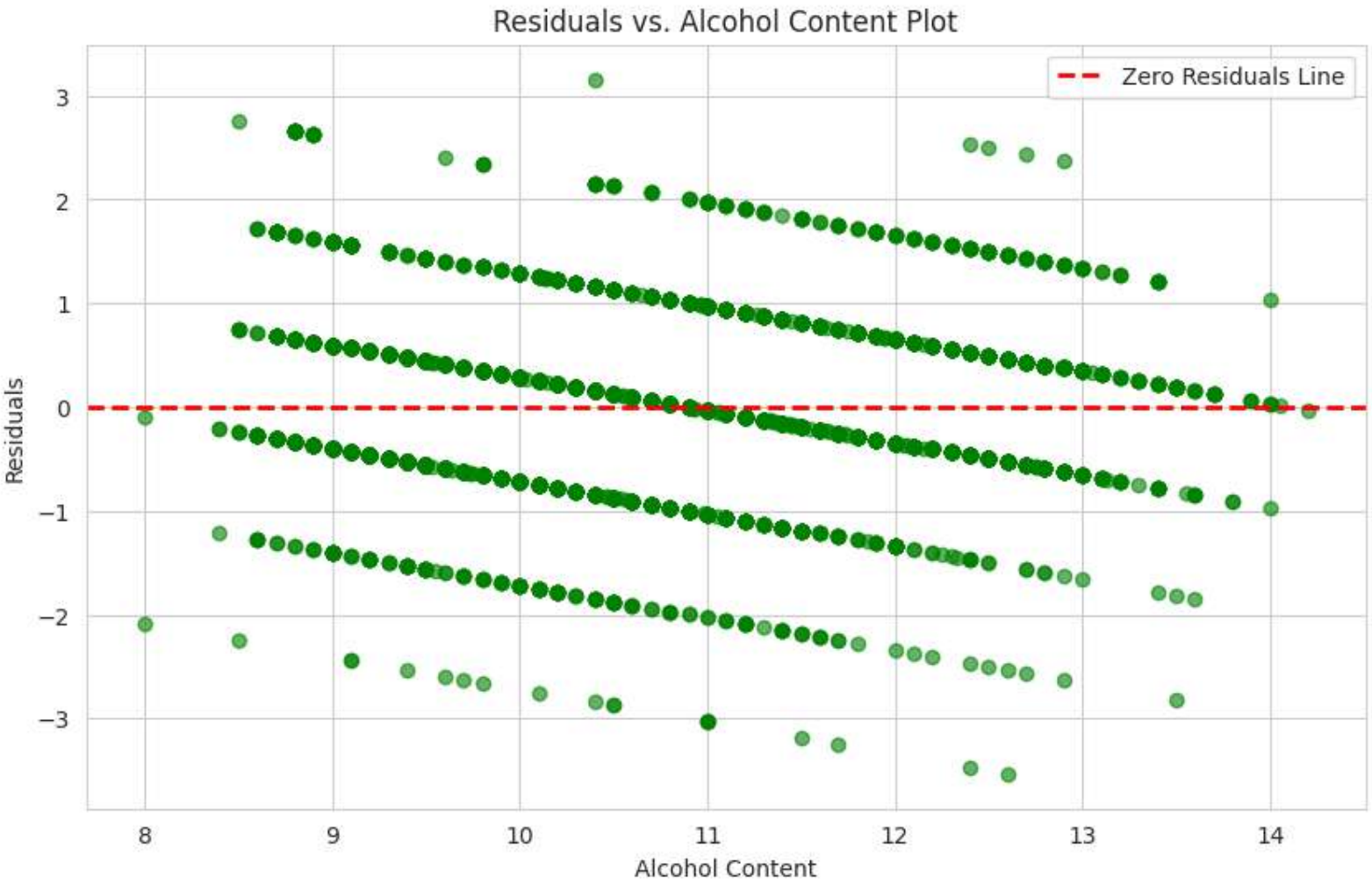

Simple Linear Regression: Alcohol vs Quality
Quality
Alcohol
8
7
6
5
4
3
8
9
10
11
12
13
14
Actual
Regression Line

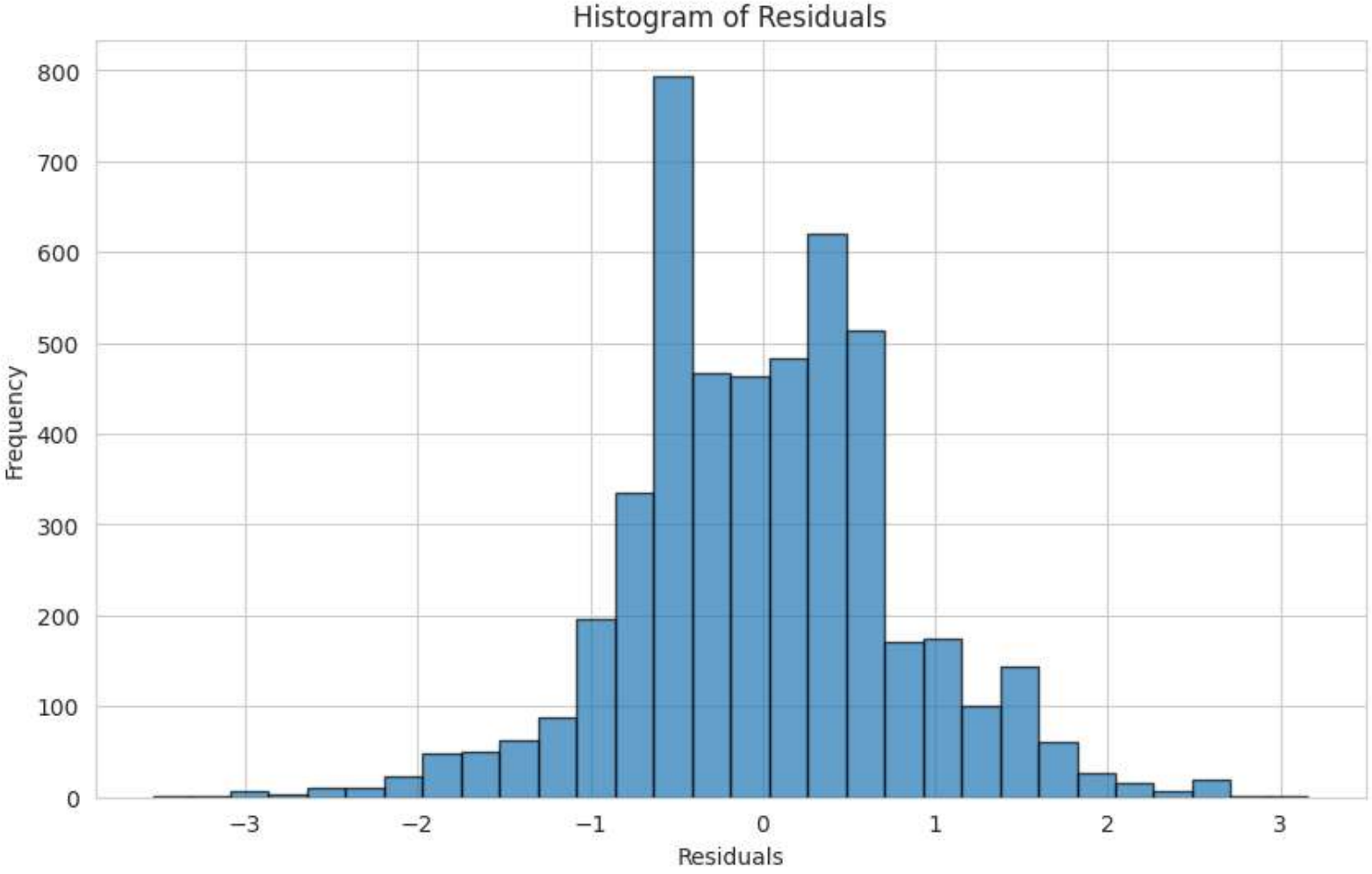

Actual vs Predicted Quality
Predicted Quality
7.00
6.75
6.50
6.25
6.00
5.75
5.50
5.25
5.00
3
4
5
6
7
8
Actual Quality

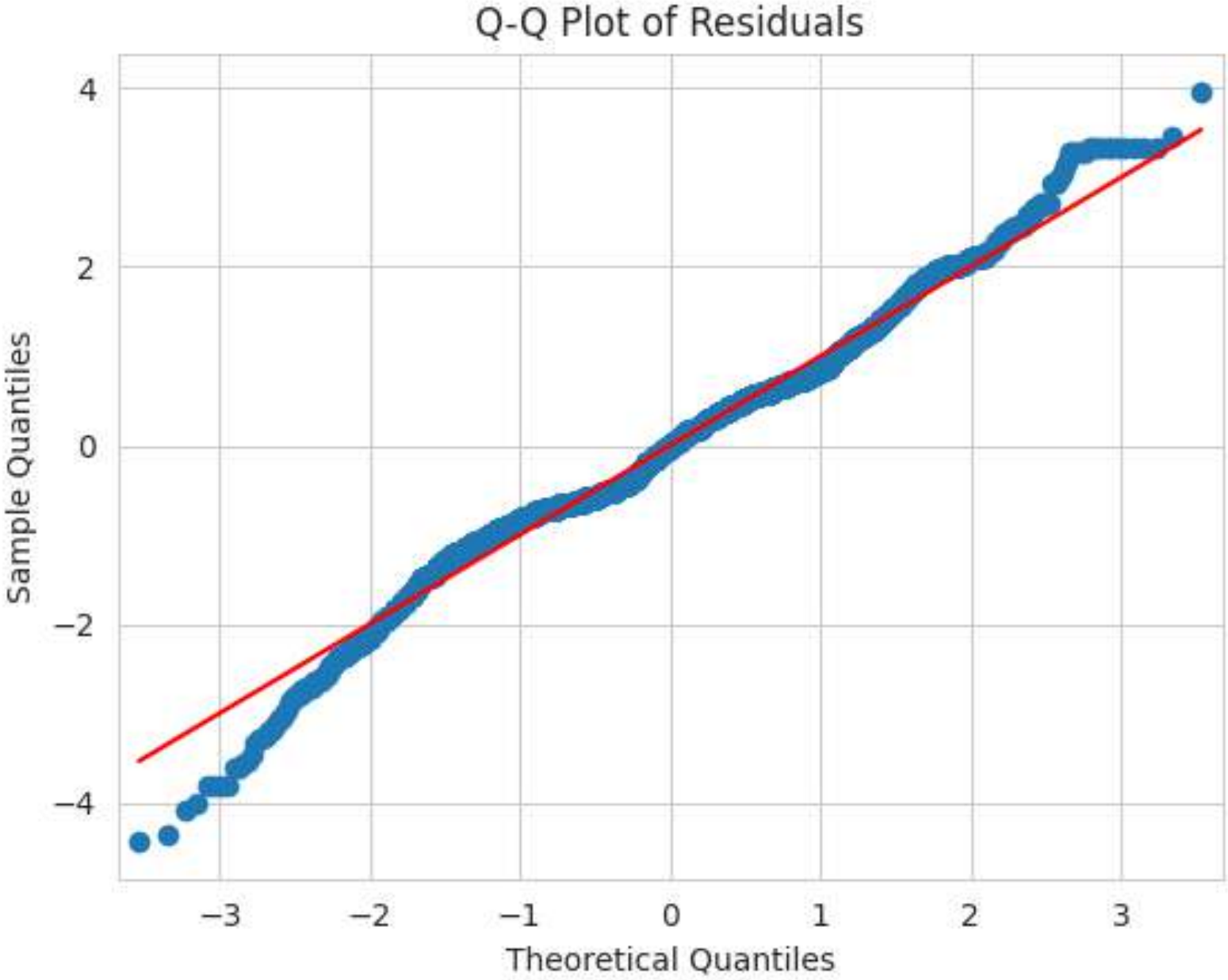

Simple Linear Regression: Alcohol vs. Quality with Regression Line
Actual Data Points
Regression Line
Quality
Alcohol Content
9
8
7
6
5
4
3
8
9
10
11
12
13
14

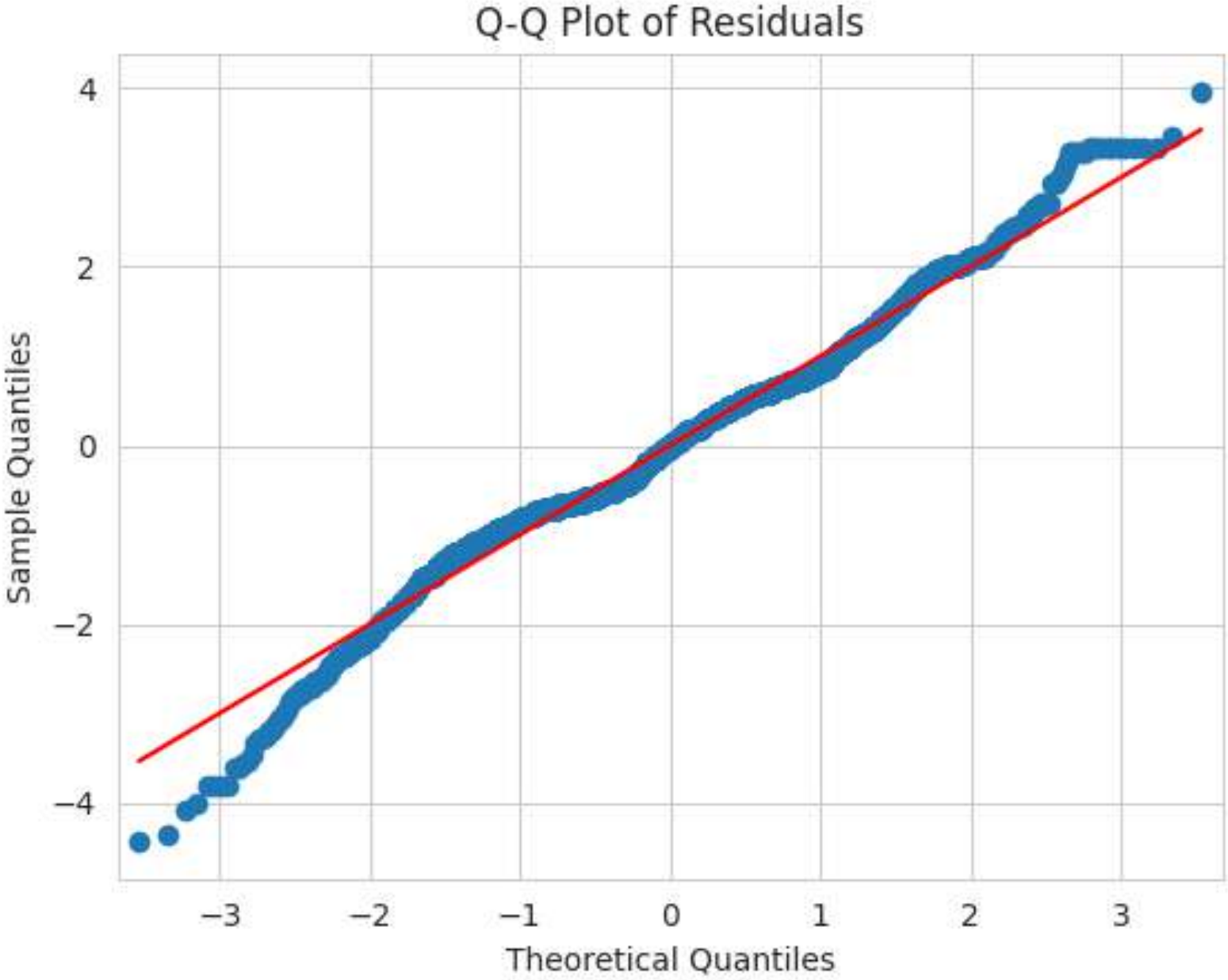

Actual vs. Predicted Quality (Full Dataset)
Actual vs. Predicted
Perfect Prediction Line
Actual Quality
Predicted Quality
9
8
7
6
5
4
3
3
4
5
6
7
8
9

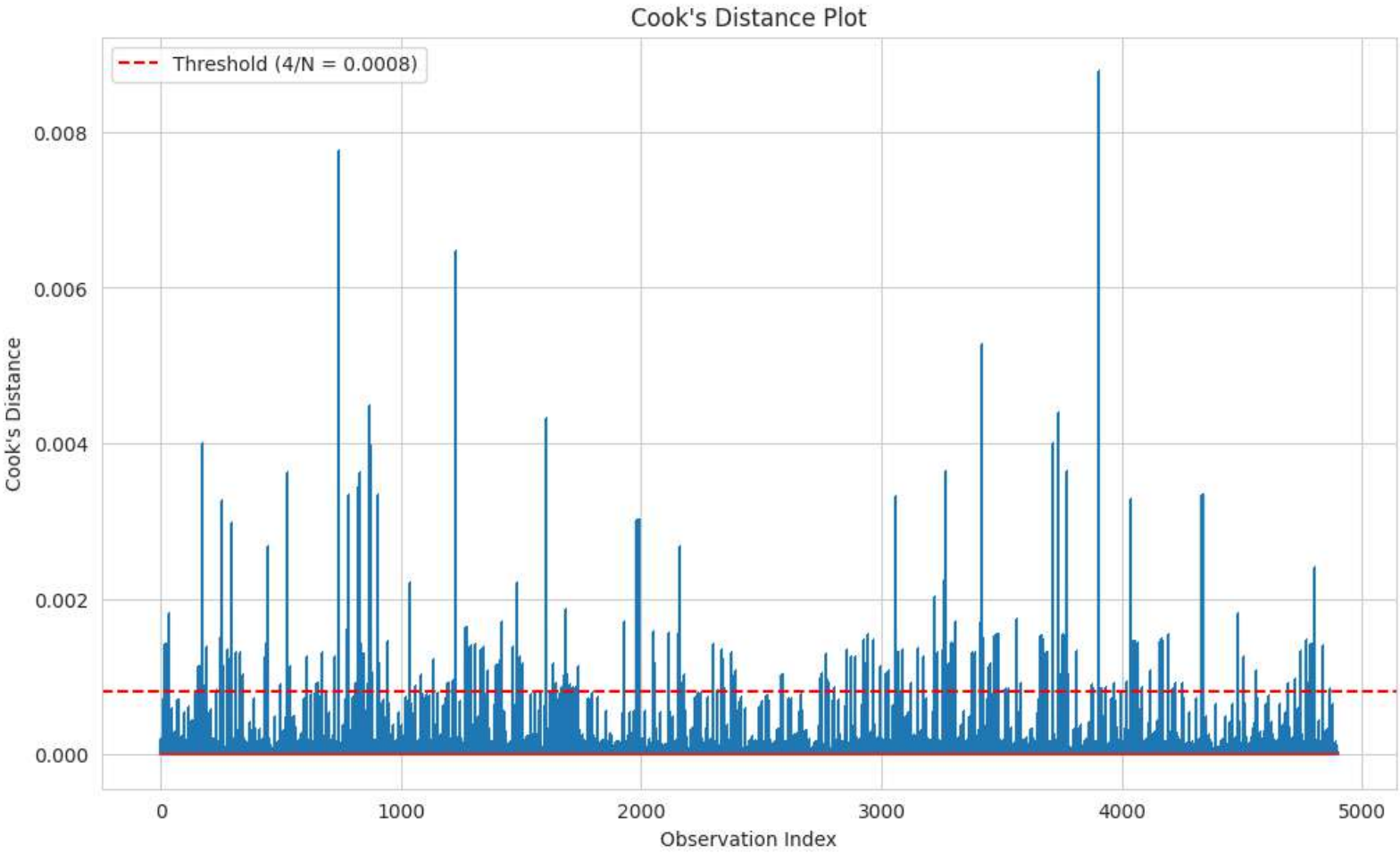

Residuals vs. Fitted Values Plot
Zero Residuals Line
Residuals
Fitted Values
3
2
1
0
−1
−2
−3
5.00
5.25
5.50
5.75
6.00
6.25
6.50
6.75
7.00

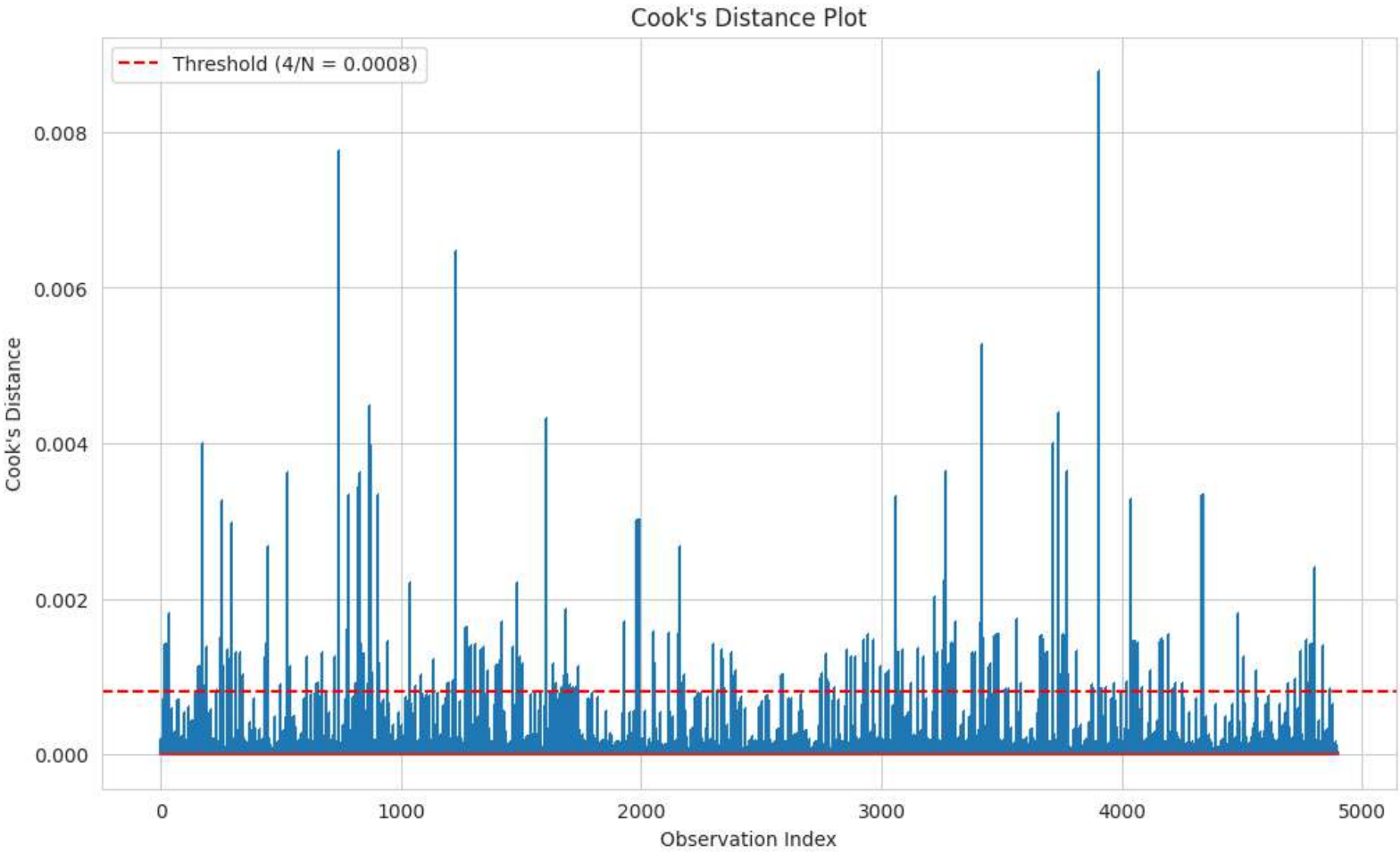

Residuals vs. Alcohol Content Plot
Zero Residuals Line
Residuals
Alcohol Content
3
2
1
0
−1
−2
−3
8
9
10
11
12
13
14

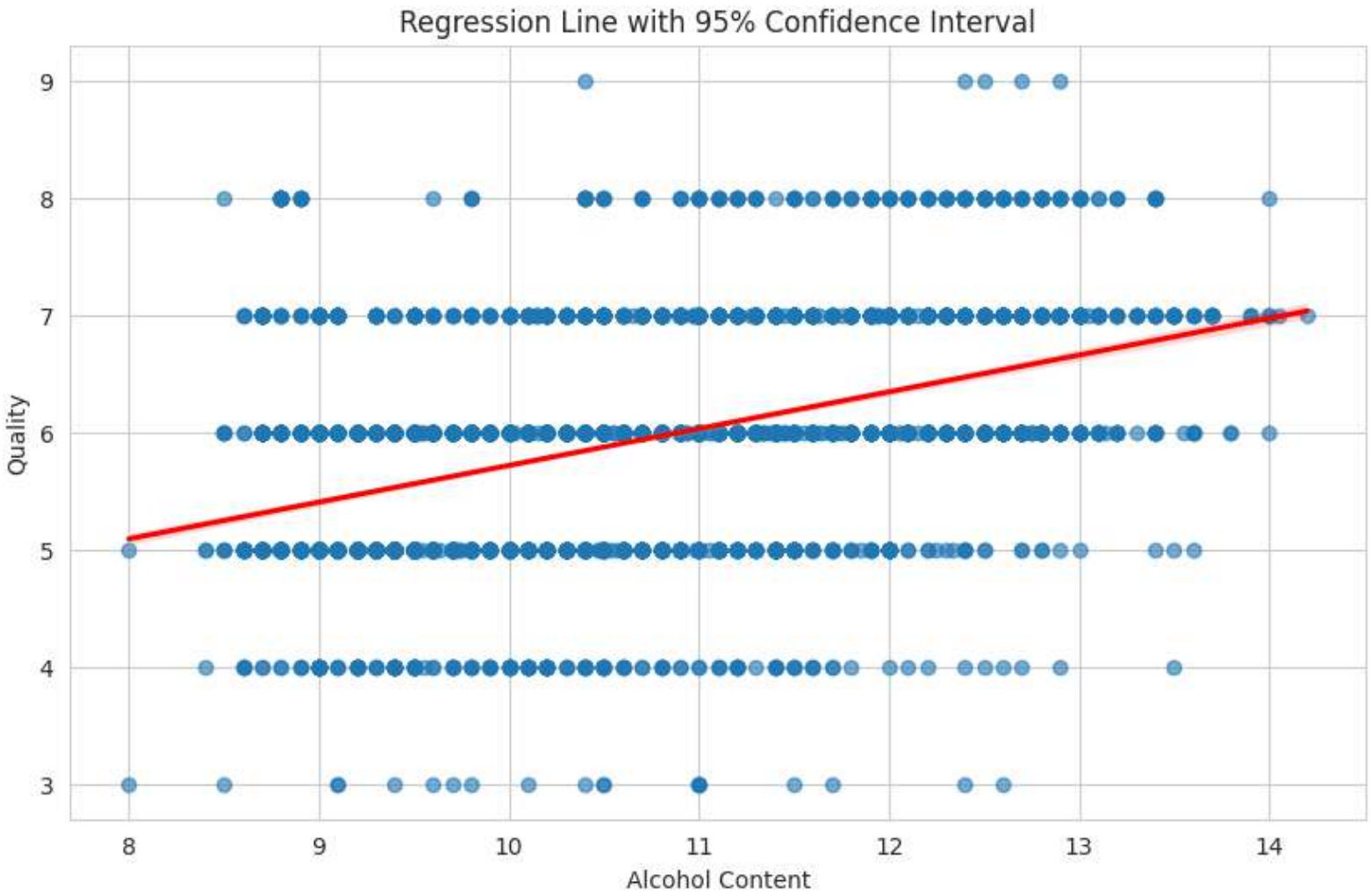

Histogram of Residuals
Frequency
Residuals
0
100
200
300
400
500
600
700
800
−3
−2
−1
0
1
2
3

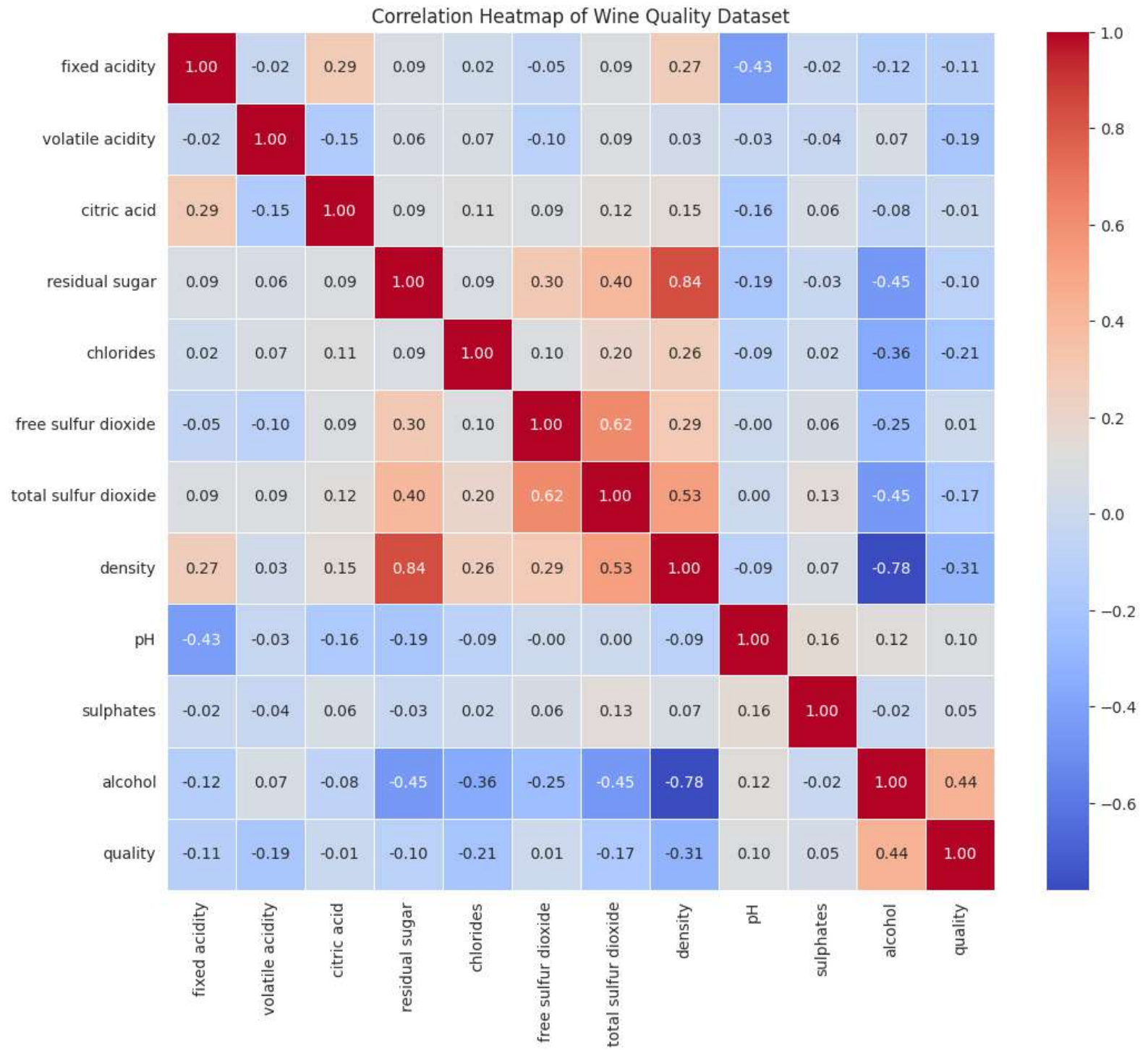

Q-Q Plot of Residuals
Sample Quantiles
Theoretical Quantiles
4
2
0
−2
−4
−3
−2
−1
0
1
2
3

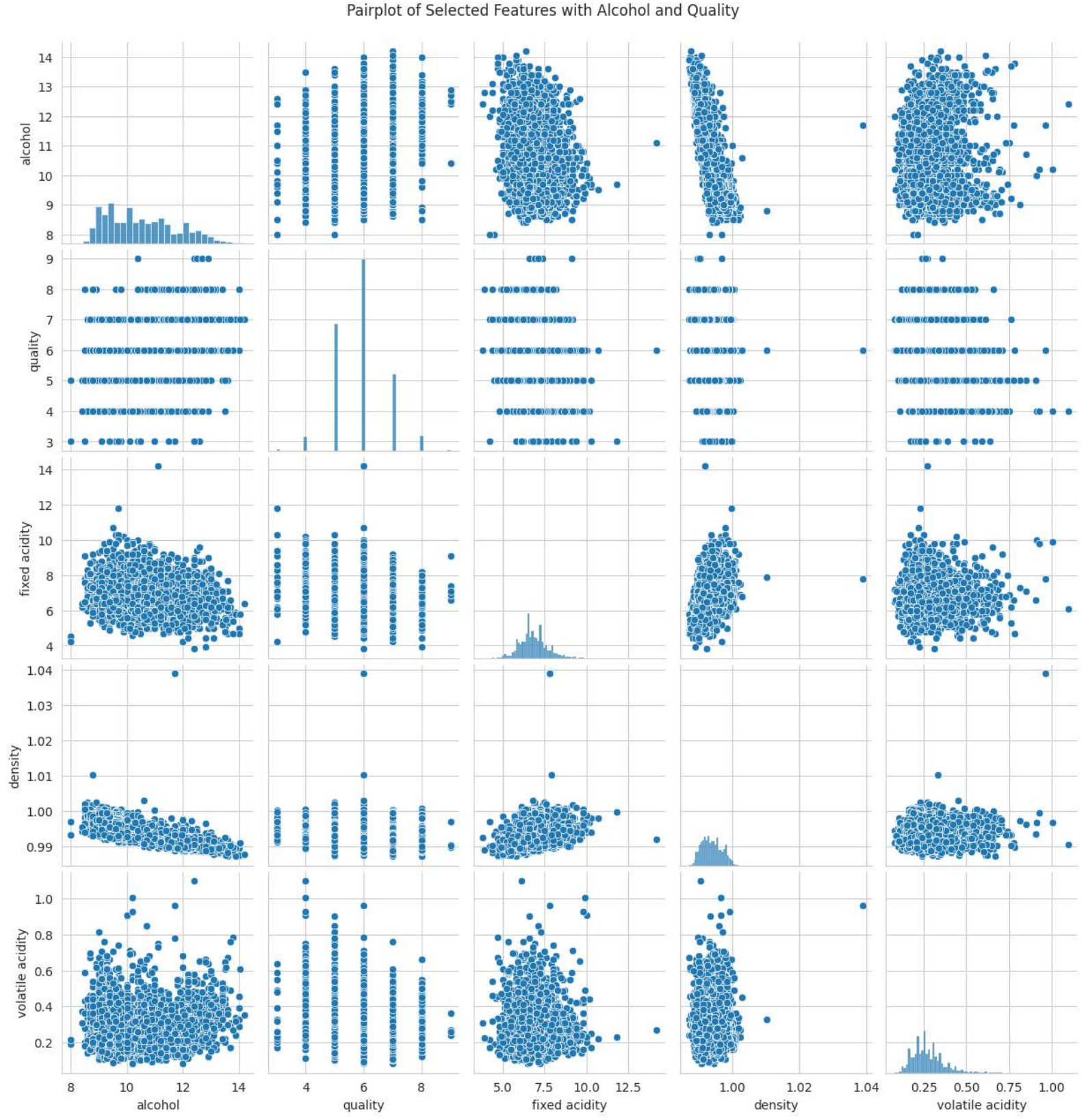

Q-Q Plot of Residuals
Sample Quantiles
Theoretical Quantiles
4
2
0
−2
−4
−3
−2
−1
0
1
2
3

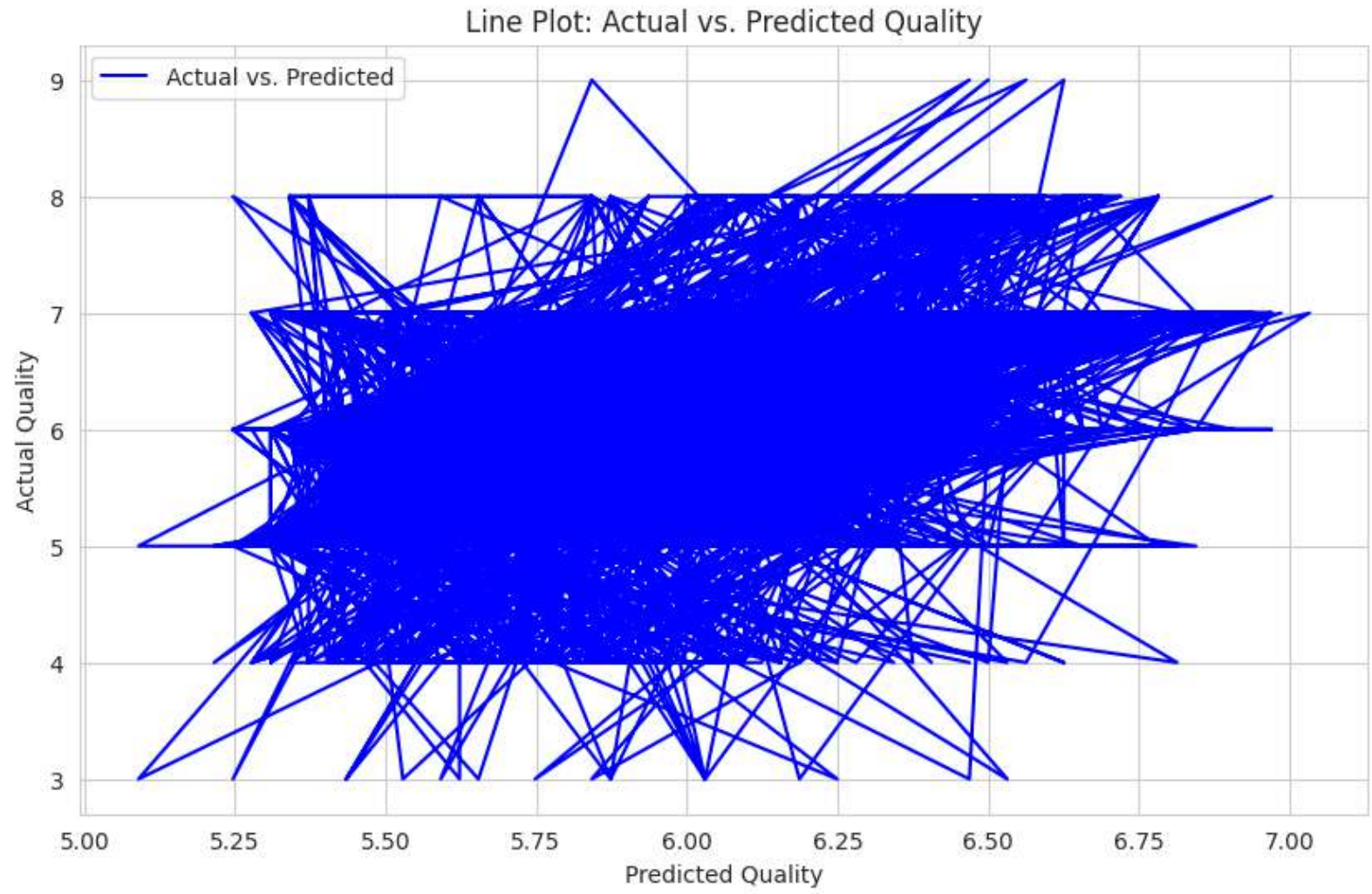

Cook's Distance Plot
Threshold (4/N = 0.0008)
0.008
0.006
0.004
0.002
0.000
Cook's Distance
0
1000
2000
3000
4000
5000
Observation Index

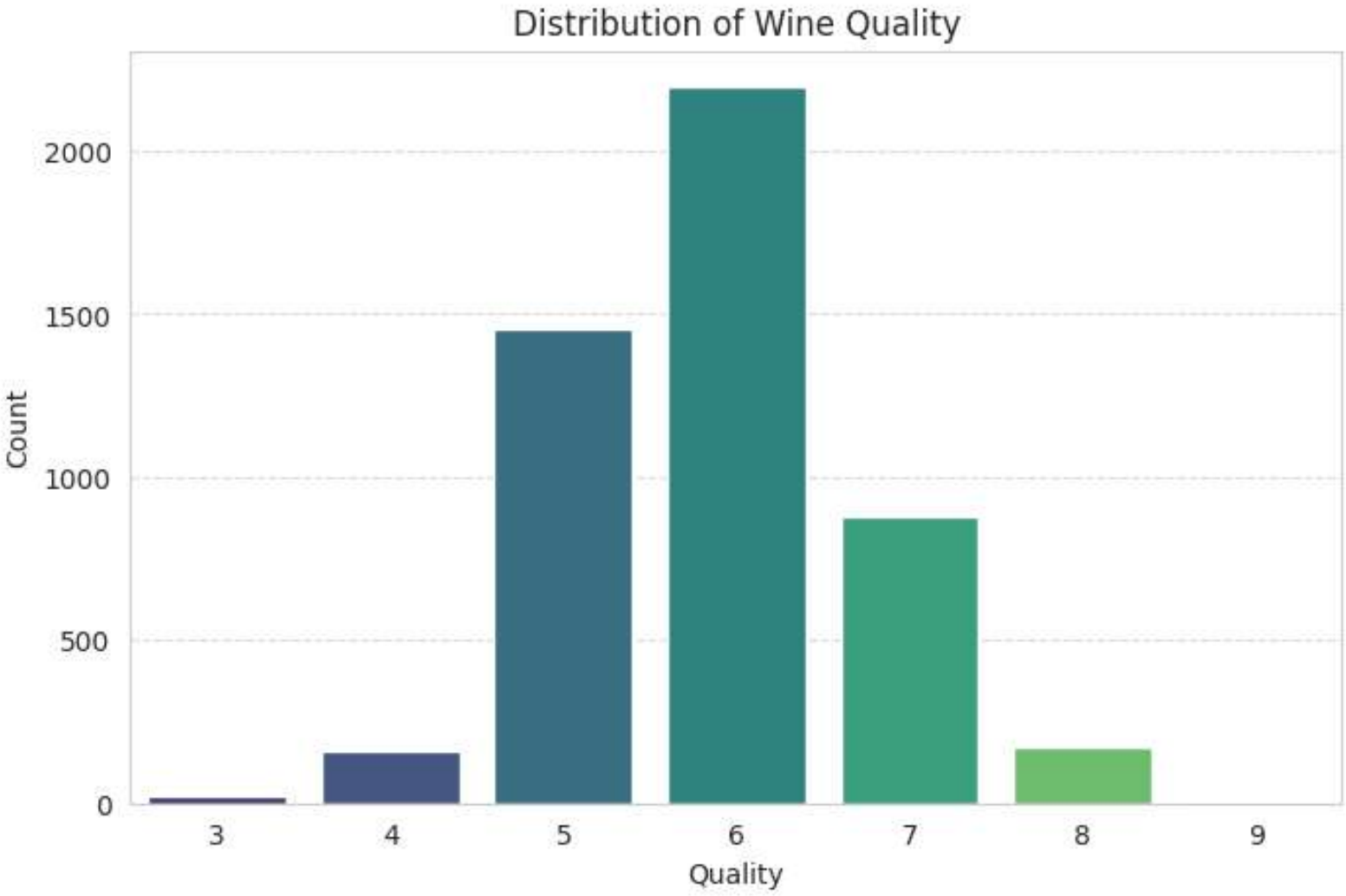

Cook's Distance Plot
Threshold (4/N = 0.0008)
0.008
0.006
0.004
0.002
0.000
Cook's Distance
0
1000
2000
3000
4000
5000
Observation Index

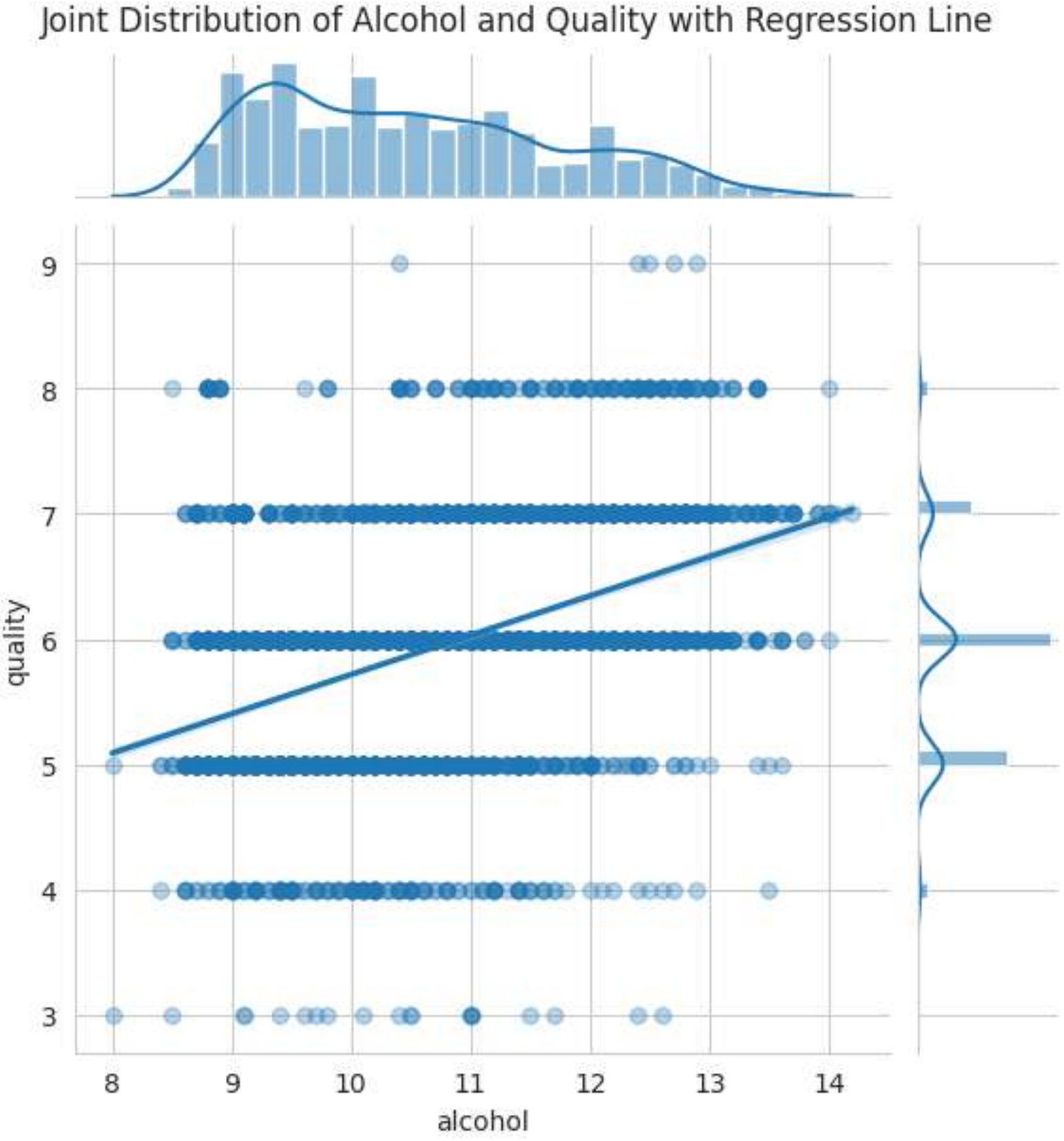

Regression Line with 95% Confidence Interval
Quality
Alcohol Content
9
8
7
6
5
4
3
8
9
10
11
12
13
14

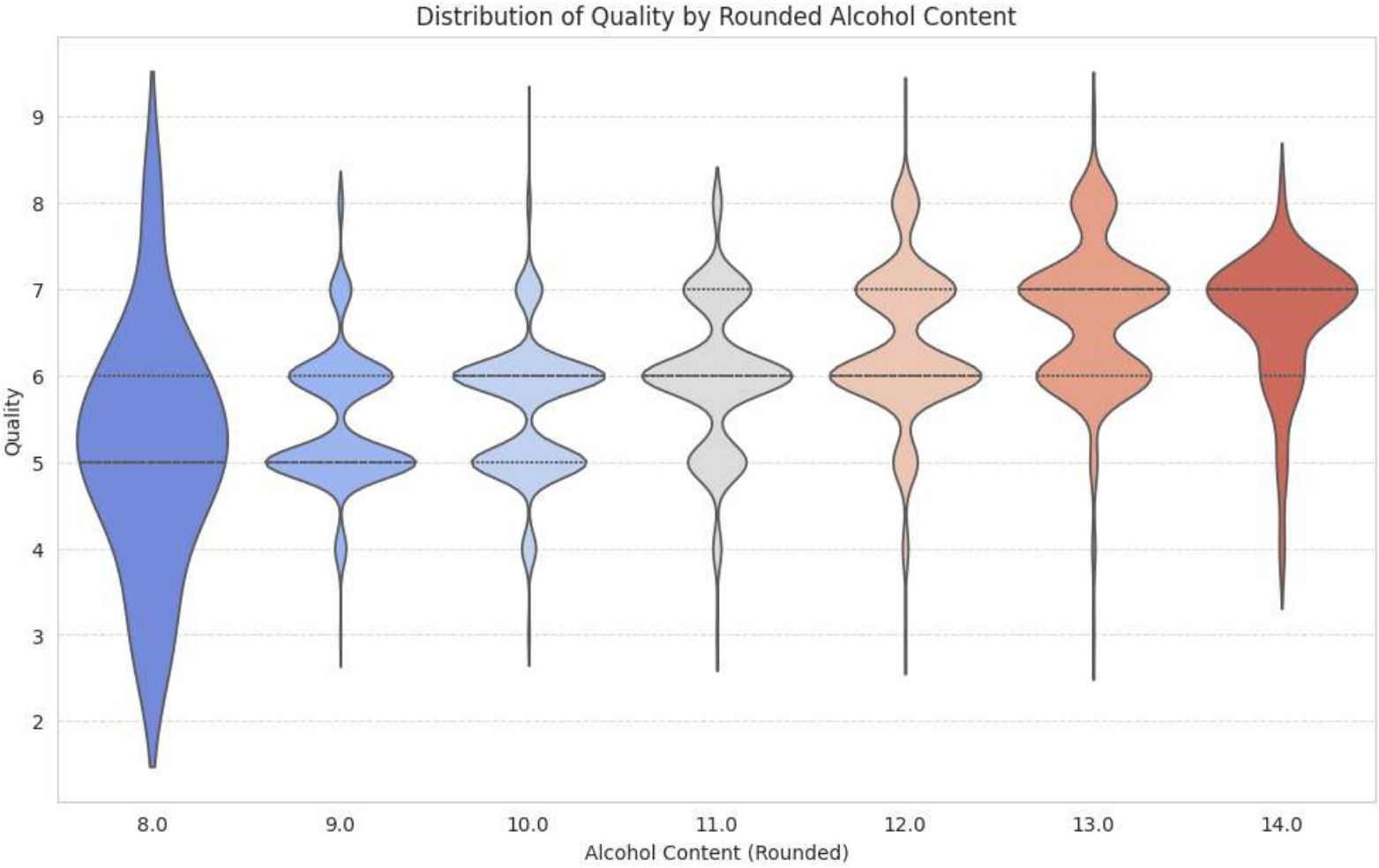

Correlation Heatmap of Wine Quality Dataset
fixed acidity 1.00 -0.02 0.29 0.09 0.02 -0.05 0.09 0.27 -0.43 -0.02 -0.12 -0.11
volatile acidity -0.02 1.00 -0.15 0.06 0.07 -0.10 0.09 0.03 -0.03 -0.04 0.07 -0.19
citric acid 0.29 -0.15 1.00 0.09 0.11 0.09 0.12 0.15 -0.16 0.06 -0.08 -0.01
residual sugar 0.09 0.06 0.09 1.00 0.09 0.30 0.40 0.84 -0.19 -0.03 -0.45 -0.10
chlorides 0.02 0.07 0.11 0.09 1.00 0.10 0.20 0.26 -0.09 0.02 -0.36 -0.21
free sulfur dioxide -0.05 -0.10 0.09 0.30 0.10 1.00 0.62 0.29 -0.00 0.06 -0.25 0.01
total sulfur dioxide 0.09 0.09 0.12 0.40 0.20 0.62 1.00 0.53 0.00 0.13 -0.45 -0.17
density 0.27 0.03 0.15 0.84 0.26 0.29 0.53 1.00 -0.09 0.07 -0.78 -0.31
pH -0.43 -0.03 -0.16 -0.19 -0.09 -0.00 0.00 -0.09 1.00 0.16 0.12 0.10
sulphates -0.02 -0.04 0.06 -0.03 0.02 0.06 0.13 0.07 0.16 1.00 -0.02 0.05
alcohol -0.12 0.07 -0.08 -0.45 -0.36 -0.25 -0.45 -0.78 0.12 -0.02 1.00 0.44
quality -0.11 -0.19 -0.01 -0.10 -0.21 0.01 -0.17 -0.31 0.10 0.05 0.44 1.00
fixed acidity
volatile acidity
citric acid
residual sugar
chlorides
free sulfur dioxide
total sulfur dioxide
density
pH
sulphates
alcohol
quality
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6


Pairplot of Selected Features with Alcohol and Quality
alcohol
quality
fixed acidity
density
volatile acidity


Line Plot: Actual vs. Predicted Quality
Actual vs. Predicted
Actual Quality
Predicted Quality
9
8
7
6
5
4
3
5.00
5.25
5.50
5.75
6.00
6.25
6.50
6.75
7.00


Distribution of Wine Quality
Count
2000
1500
1000
500
0
3
4
5
6
7
8
9
Quality


Joint Distribution of Alcohol and Quality with Regression Line
quality
alcohol
9
8
7
6
5
4
3
8
9
10
11
12
13
14


Distribution of Quality by Rounded Alcohol Content
Quality
9
8
7
6
5
4
3
2
8.0
9.0
10.0
11.0
12.0
13.0
14.0
Alcohol Content (Rounded)