

Regression Diagnostics

Test name: Regression Diagnostics

Dataset: D:\DATA ANALYSIS\H Regression Tests and Models\Regression Diagnostics\dataset.csv

Rows in original dataset: 649

Rows used in model: 649

Outcome: G3

Predictors: G1, G2, studytime, failures, absences, age

Formula: $G3 \sim G1 + G2 + studytime + failures + absences + age$

R-squared: 0.850632

Adjusted R-squared: 0.849236

RMSE: 1.254409

Durbin-Watson: 1.860956

Condition number on scaled predictors: 4.250230

Residual outliers |studentized| > 3: 10

High leverage cases: 45

High Cook's distance cases: 26

High DFFITS cases: 26

COVRATIO flagged cases: 49

Cases with any diagnostic flag: 62

Diagnostic rule: review flagged cases; do not automatically delete observations.

Interpretation should combine graphical patterns, formal tests, and subject-matter context.

Charts generated: 10

Ramsey RESET: statistic=10.7741, p=0.00108523, decision=Potential issue

Shapiro-Wilk residual normality: statistic=0.754697, p=6.24448e-30, decision=Potential issue

Jarque-Bera residual normality: statistic=10152.7, p=0, decision=Potential issue

Breusch-Pagan: statistic=33.0542, p=1.0237e-05, decision=Potential issue

White test: statistic=75.5474, p=1.75286e-06, decision=Potential issue

Durbin-Watson: statistic=1.86096, p=NA, decision=Acceptable range

Model summary

Regression Diagnostics	test_name	outcome	predictors	is_complete	number_of_predictions	r_squared	adjusted_r_squared	rmse	f_statistic	f_p_value	aic	bic	odds_ratio	number_scaled
	G1, G2, G3	Study time, failures	649.00000	6.00000	0.85063	0.84924	1.25441	609.35274	0.00000	2142.95478	2174.28281	1.86096	4.25023	

Formal diagnostic tests

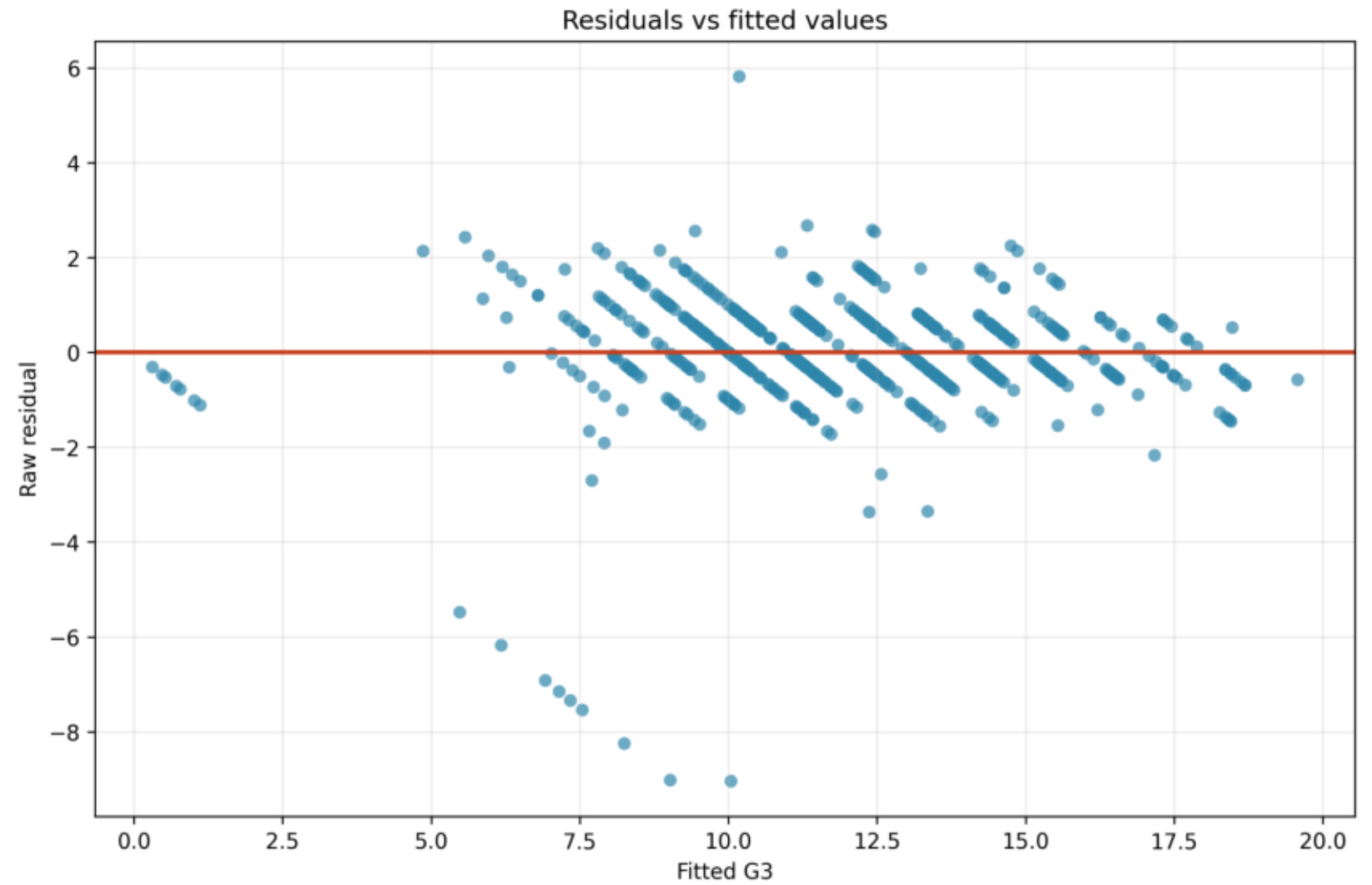
diagnostic	statistic	p_value	purpose	decision
Ramsey RESET	10.77406	0.00109	Functional form / omitted nonlinearity	Potential issue
Shapiro-Wilk residual normality	0.75470	0.00000	Residual normality	Potential issue
Jarque-Bera residual normality	10152.72945	0.00000	Residual normality	Potential issue
Breusch-Pagan	33.05419	0.00001	Homoscedasticity	Potential issue
White test	75.54743	0.00000	Homoscedasticity	Potential issue
Durbin-Watson	1.86096		Residual independence; values near 2 are good	Acceptable range

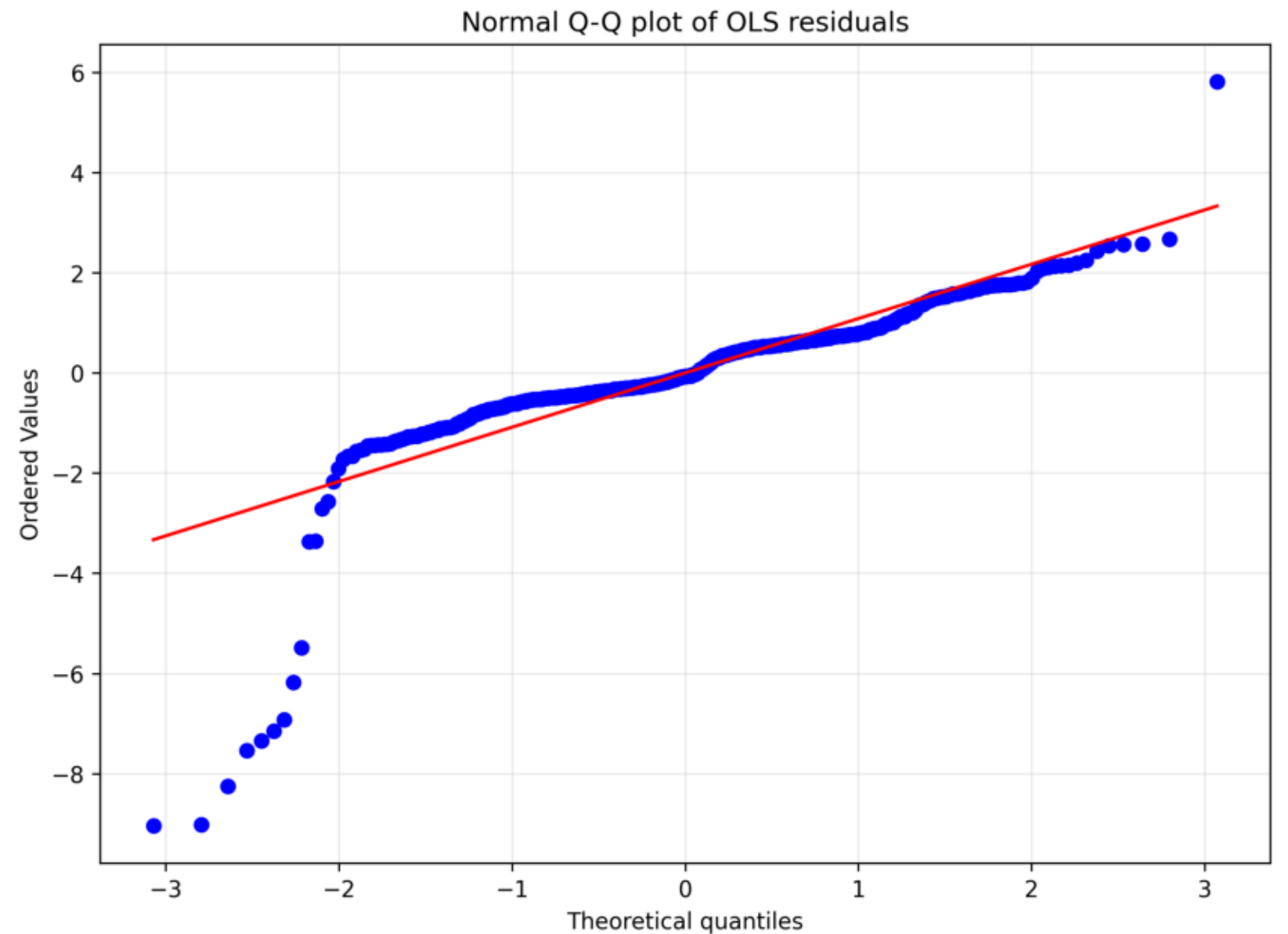
Influence flag summary

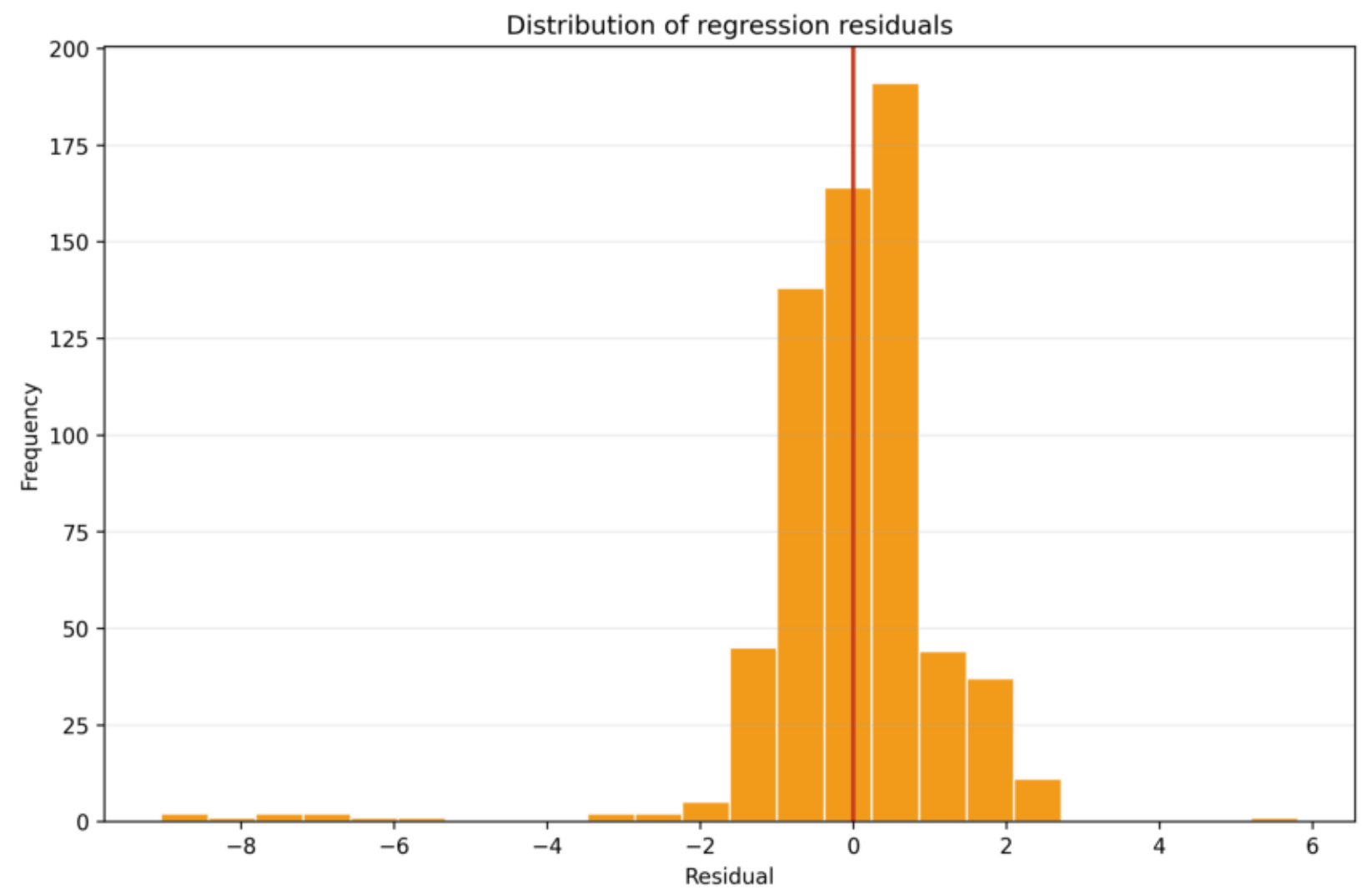
n_cases	residual_outliers	high_leverage_cases	high_cooks_cases	high_dffits_cases	ratio_flagged_cases	maxmany_flag	absolute_student	maximum_leverage	num_cooks_max	num_absolute_dffits
649.00000	10.00000	45.00000	26.00000	26.00000	49.00000	62.00000	7.59071	0.10331	0.17695	1.16099

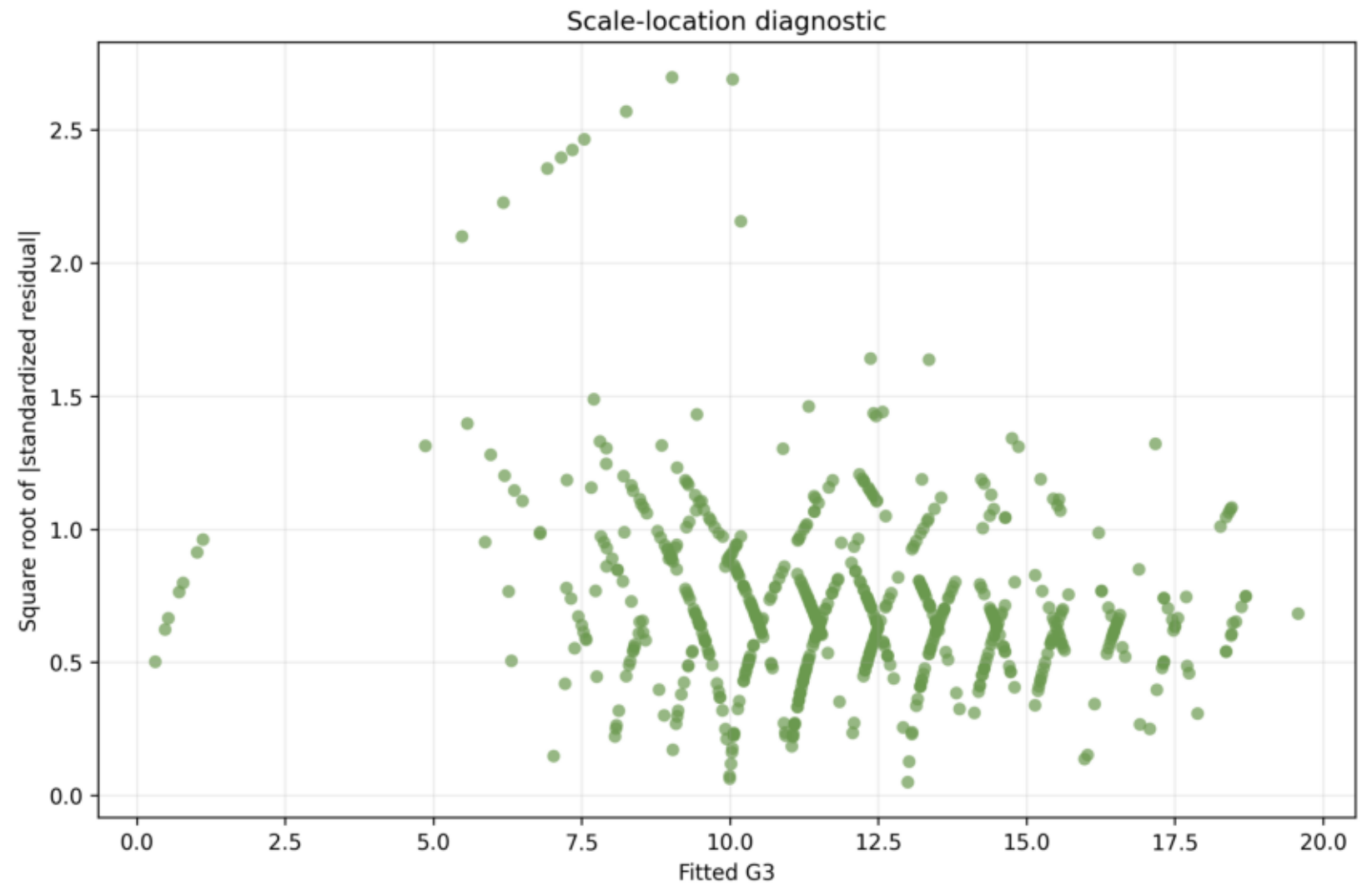
Top influential cases

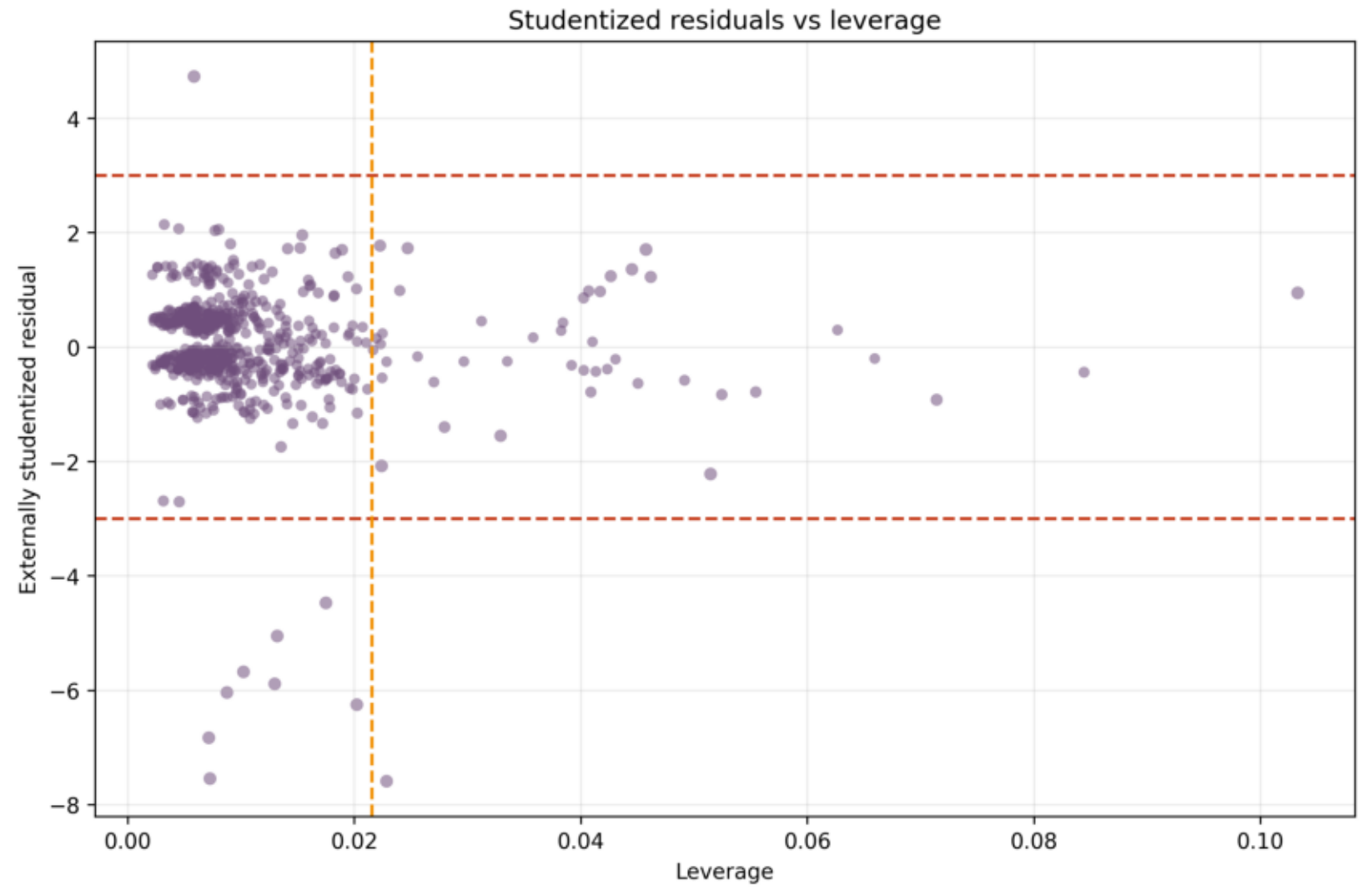
[illegible]

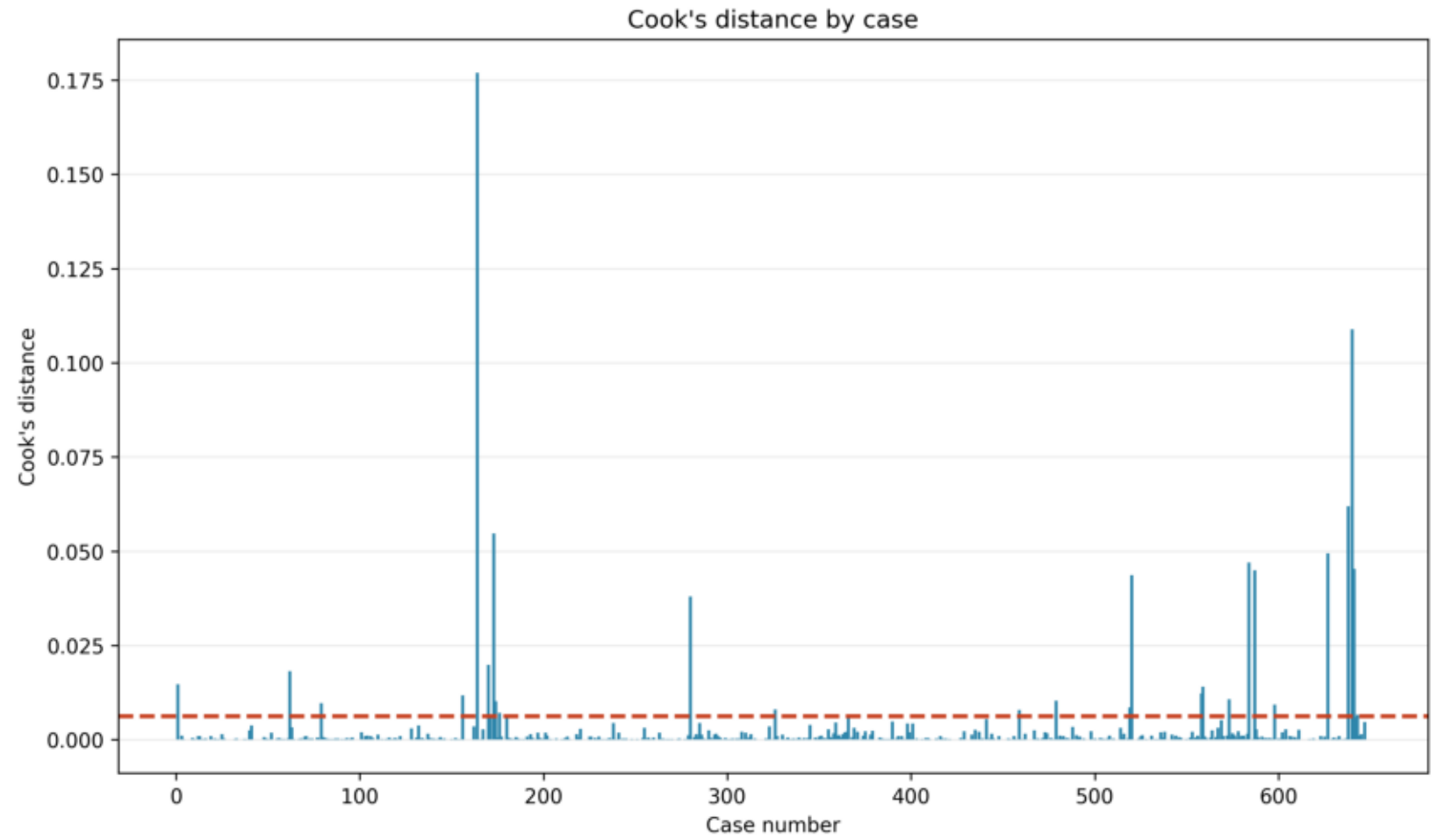


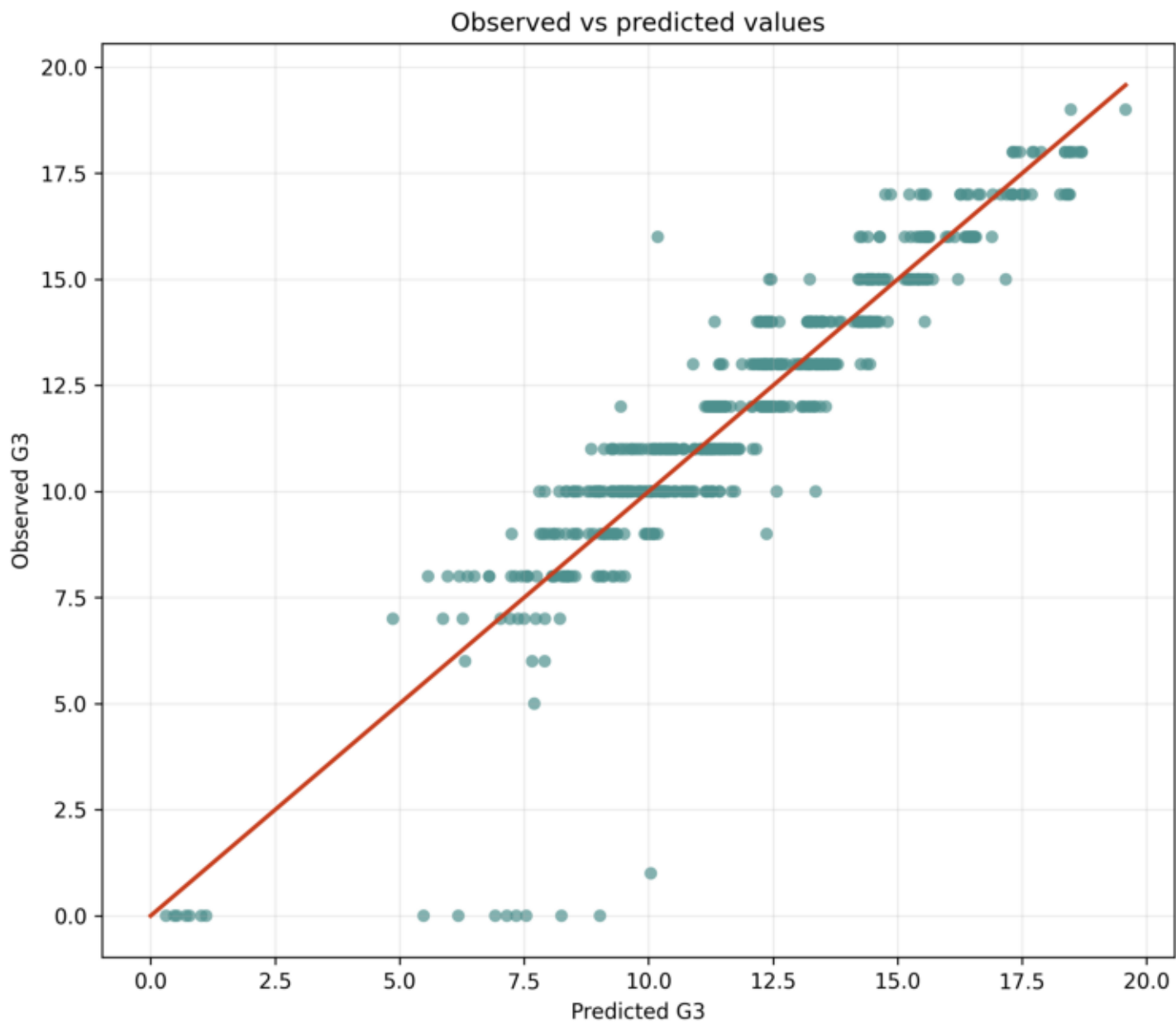












Observed vs predicted values

