

Regression Residual Analysis

Test name: Regression Residual Analysis

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

Rows in original dataset: 649

Rows used in model: 649

Outcome: G3

Predictors: G1, G2, studytime, failures, absences

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

R-squared: 0.850563

Adjusted R-squared: 0.849401

RMSE: 1.253722

Durbin-Watson: 1.857347

Condition number on scaled predictors: 4.130506

Residual outliers |studentized| > 3: 10

High leverage cases: 47

High Cook's distance cases: 29

High DFFITS cases: 29

COVRATIO flagged cases: 49

Cases with any diagnostic flag: 66

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=9.86846, p=0.00175867, decision=Potential issue

Shapiro-Wilk residual normality: statistic=0.756572, p=7.72186e-30, decision=Potential issue

Jarque-Bera residual normality: statistic=10019.3, p=0, decision=Potential issue

Breusch-Pagan: statistic=30.129, p=1.39107e-05, decision=Potential issue

White test: statistic=66.4542, p=6.8318e-07, decision=Potential issue

Durbin-Watson: statistic=1.85735, p=NA, decision=Acceptable range

Model summary

	test_name	outcome	predictors	is_complete	number	number_of_predictors	r_squared	adjusted_r_squared	rmse	f_statistic	f_p_value	aic	bic	odds_ratio	number_scaled
Progression Residual	G1, G2 study	time, failure	649.00000	5.00000	0.85056	0.84940	1.25372	731.96586	0.00000	2141.25366	2168.10626	1.85735	4.13051		

Formal diagnostic tests

diagnostic	statistic	p_value	purpose	decision
Ramsey RESET	9.86846	0.00176	Functional form / omitted nonlinearity	Potential issue
Shapiro-Wilk residual normality	0.75657	0.00000	Residual normality	Potential issue
Jarque-Bera residual normality	10019.31661	0.00000	Residual normality	Potential issue
Breusch-Pagan	30.12902	0.00001	Homoscedasticity	Potential issue
White test	66.45417	0.00000	Homoscedasticity	Potential issue
Durbin-Watson	1.85735		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	47.00000	29.00000	29.00000	49.00000	66.00000	7.58173	0.10329	0.20288	1.15074

Top influential cases

[illegible]



















