Correlation vs Regression Analysis

Predictor / X variable: G2
Outcome / Y variable: G3

Correlation summary:

metric	value	interpretation
x_variable	G2	Predictor used for both correlation and simple regression.
y_variable	G3	Outcome used for both correlation and simple regression.
valid_pairs_n	649	Number of complete x−y pairs.
mean_x	11.5701078582435	Average predictor value.
mean_y	11.9060092449923	Average outcome value.
sample_sd_x	2.91363866430387	Sample standard deviation of predictor.
sample_sd_y	3.2306562428048	Sample standard deviation of outcome.
sample_covariance	8.6462601533223	Shared linear variation in original units.
pearson_r	0.918548003560351	Direction and strength of the linear relationship.
pearson_r_squared	0.843730434844706	Same value as simple regression R−squared.
correlation_t	59.1039894287048	t statistic for H0: r = 0.
correlation_p_value	5.64240148958541e−263	Two−tailed significance test for Pearson correlation.
spearman_rho	0.944451186692252	Rank−order monotonic association context.
spearman_p_value	7.06637721976557e−315	Two−tailed significance test for Spearman correlation.

Model fit:

metric	value	interpretation
r	0.9185480	Correlation shows strength and direction.
r_squared	0.8437304	Simple regression explains this proportion of y variance.
adjusted_r_squared	0.8434889	R−squared adjusted for one predictor.
slope	1.0184903	Expected y change for a one−unit increase in x.
intercept	0.1219661	Predicted y when x equals zero.
rmse	1.2761247	Typical prediction error in y units.
mae	0.8080228	Average absolute prediction error in y units.
correlation_vs_regression_link	0.8437304	In simple linear regression, R−squared equals Pearson r squared.

ANOVA summary:

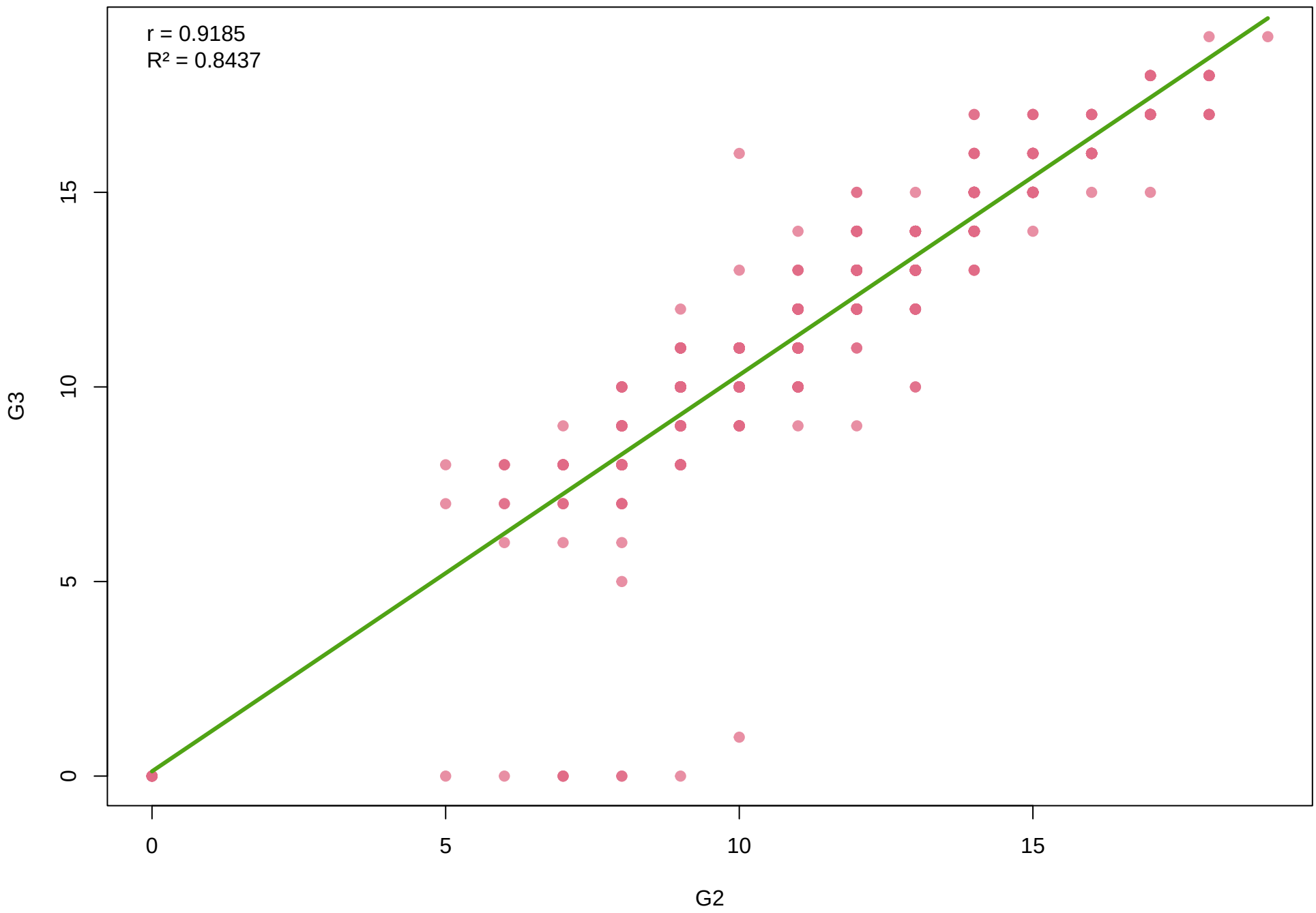
source	sum_of_squares	df	mean_square	f_value	p_value
Regression	5706.374	1	5706.373839	3493.282	5.642401e−263
Residual	1056.893	647	1.633528	NA	NA
Total	6763.267	648	NA	NA	NA

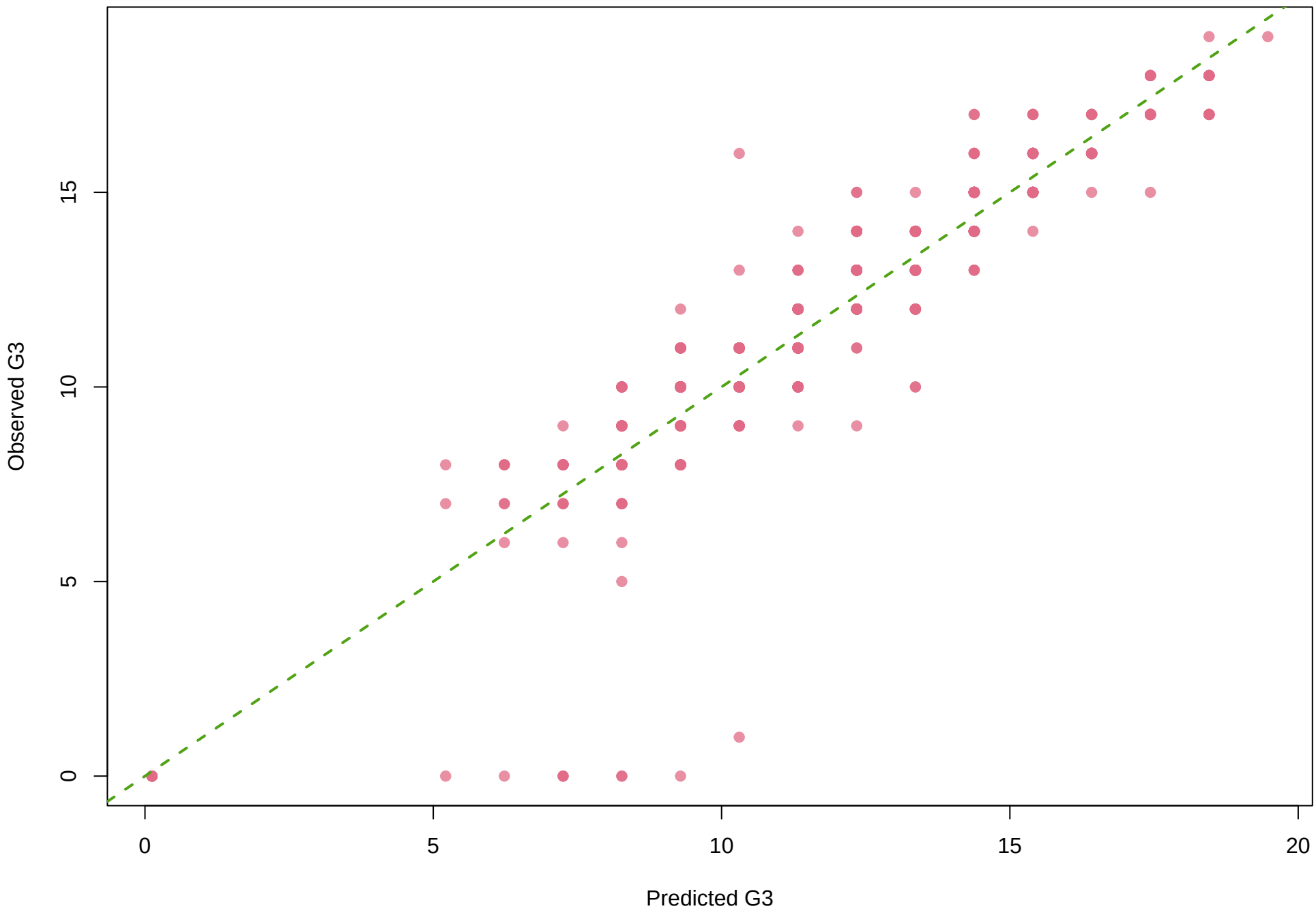
Regression coefficients:

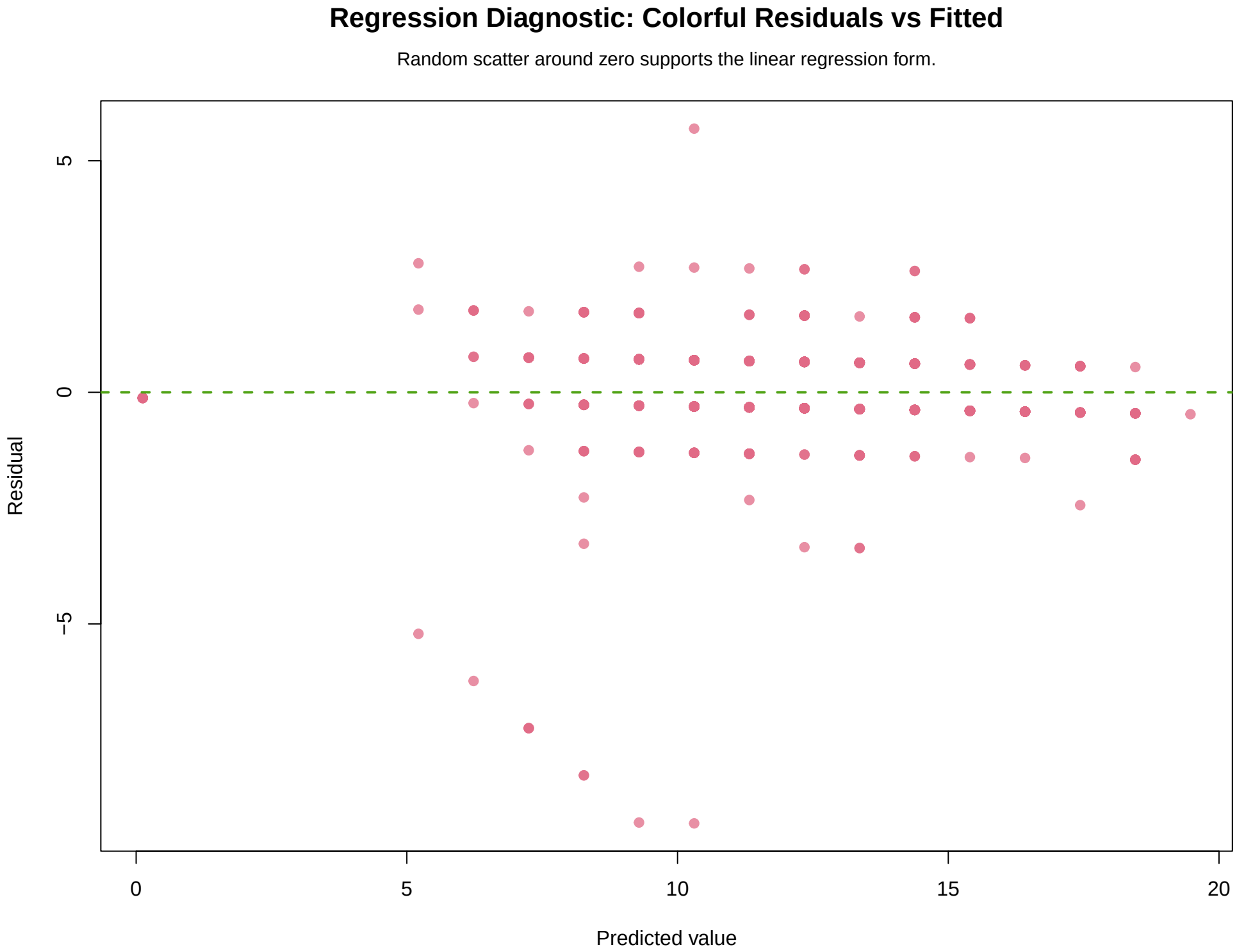
term	estimate	standard_error	t_value	p_value	ci95_lower	ci95_upper
Intercept	0.1219661	0.20559337	0.5932396	5.532281e−01	−0.2817447	0.5256769
G2	1.0184903	0.01723218	59.1039894	5.642401e−263	0.9846526	1.0523281

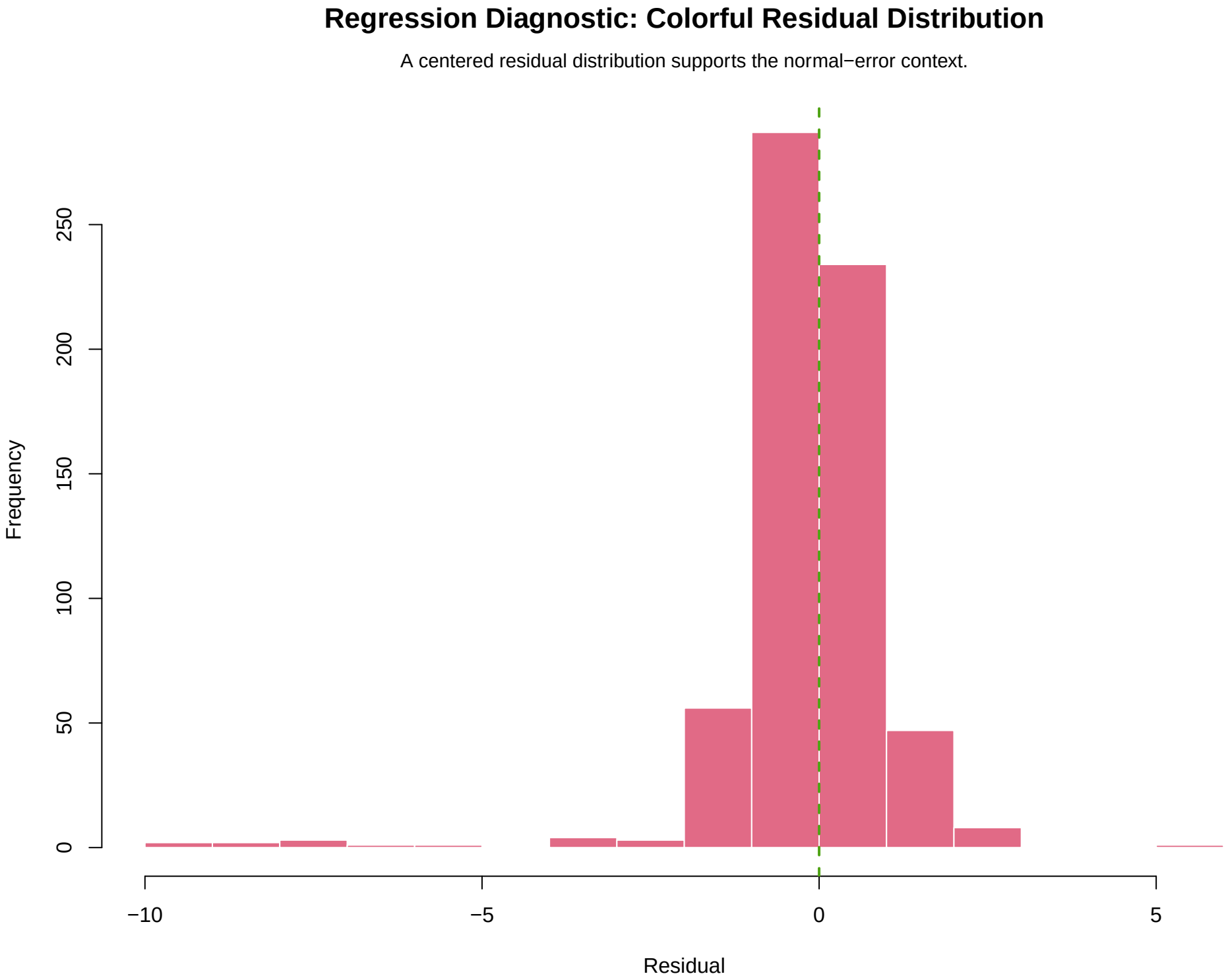
Diagnostic context:

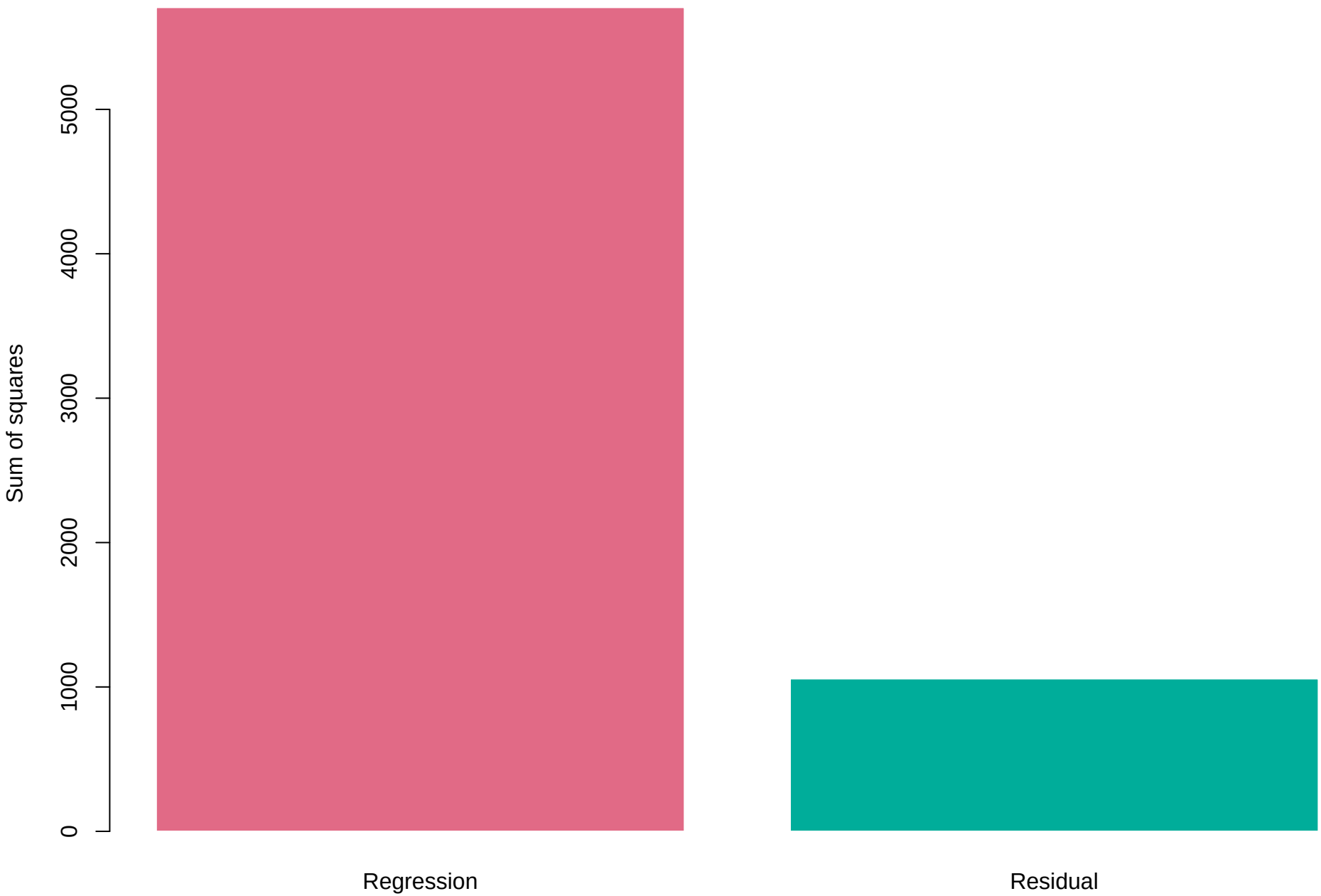
check	statistic	p_value	note
Residual mean	8.252591e−17	NA	Residual mean should be close to zero.
Residual standard deviation	1.277109e+00	NA	Residual spread in outcome units.
Residual vs fitted correlation	2.502581e−17	NA	Should be close to zero; inspect residual plot.
Predictor vs residual correlation	3.625127e−17	NA	Should be close to zero in simple OLS regression.

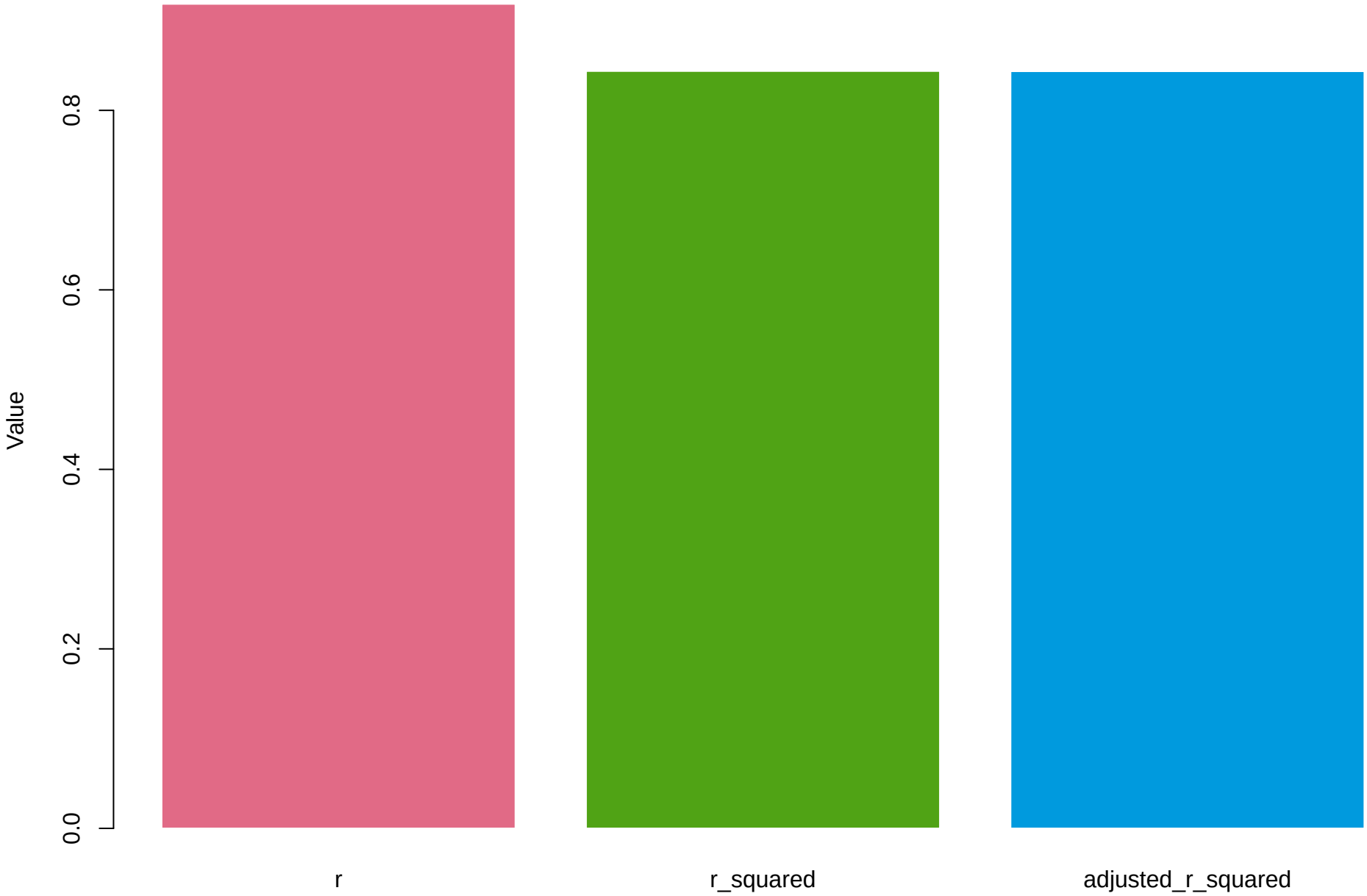

Correlation vs Regression: Colorful Scatter with Regression Line
Correlation measures strength; regression provides a prediction equation.
r = 0.9185
R² = 0.8437
G3
G2
0
5
10
15


Regression Accuracy: Colorful Observed vs Predicted
Points near the diagonal indicate predictions close to observed outcomes.
Observed G3
Predicted G3
0
5
10
15
20


Regression Diagnostic: Colorful Residuals vs Fitted
Random scatter around zero supports the linear regression form.
Residual
5
0
−5
0
5
10
15
20
Predicted value


Regression Diagnostic: Colorful Residual Distribution
A centered residual distribution supports the normal−error context.
Frequency
0
50
100
150
200
250
−10
−5
0
5
Residual


Regression Variance Decomposition
Regression SS is explained variance; residual SS is unexplained variance.
Sum of squares
0
1000
2000
3000
4000
5000
Regression
Residual


Correlation and Regression Fit Metrics
In simple regression, R−squared equals Pearson r squared.
Value
0.0
0.2
0.4
0.6
0.8
r
r_squared
adjusted_r_squared