Correlation Assumptions Report

X variable: G2
Y variable: G3

Correlation results:

method	x_variable	y_variable	n	correlation	p_value	ci95_lower_fisher_z	ci95_upper_fisher_z	assumption_note
Pearson correlation	G2	G3	649	0.918548	5.642401e-263	0.905606	0.929781	Use when relationship is linear, variables are approximately continuous and normal, and influential outliers are not driving the result.
Spearman rank correlation	G2	G3	649	0.944451	0.000000e+00	NaN	NaN	Use as a robust alternative when monotonic relationship is present but Pearson assumptions are doubtful.

Assumption decision summary:

assumption	status	evidence
Complete paired cases	Pass	n = 649
Approximate normality for Pearson	Flag	See normality table and Q-Q plots
Linearity / monotonic agreement	Pass	|Pearson - Spearman| = 0.0259
Constant residual spread	Flag	Breusch-Pagan p = 7.195e-05
No influential bivariate outliers	Flag	Mahalanobis flags = 19; Cook flags = 20
Recommended reporting choice	Decision	Pearson can be reported with caution, but Spearman should also be reported because at least one assumption needs attention.

Normality diagnostics:

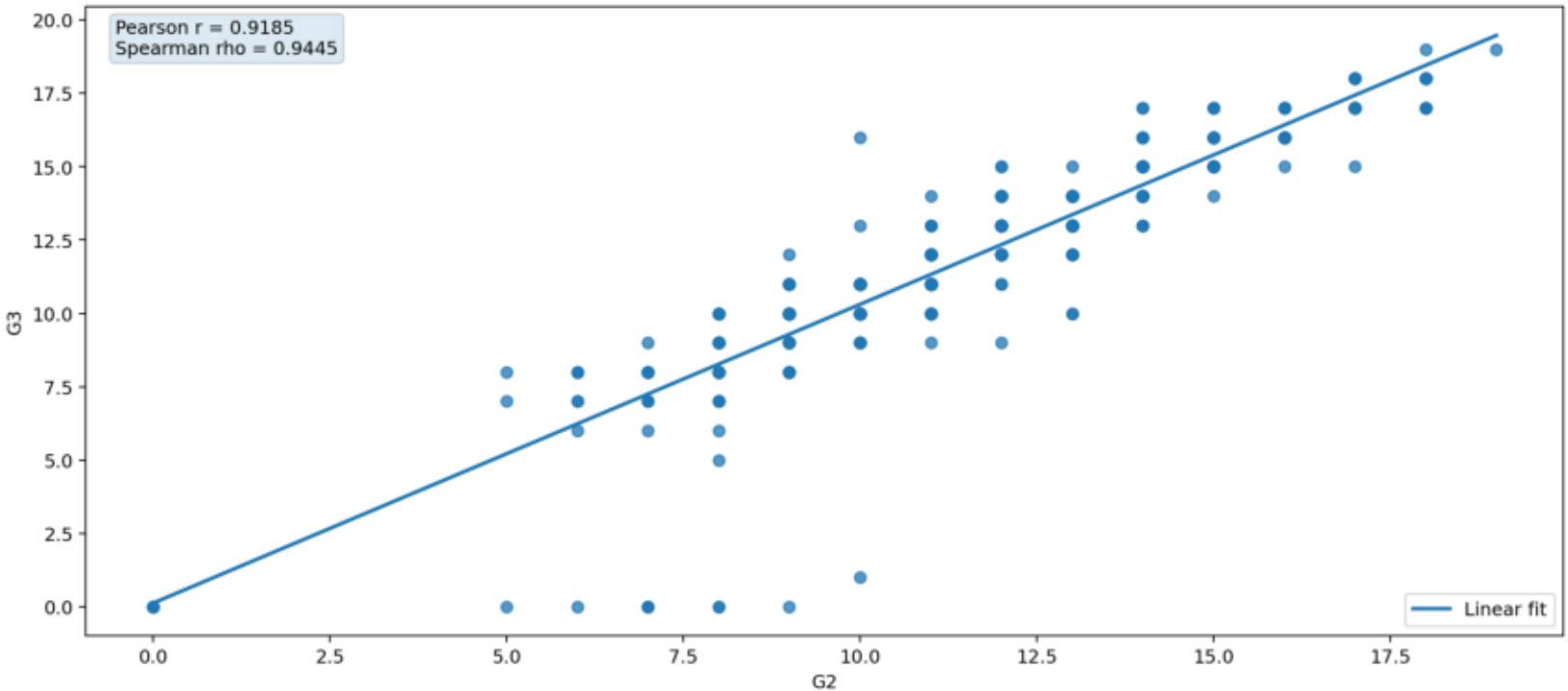
variable	n	mean	standard_deviation	minimum	maximum	skewness	excess_kurtosis	normality_method	normality_statistic	normality_p_value	normality_decision_alpha_0_05
G2	649	11.570108	2.913639	0.0	19.0	-0.360283	1.662465	Shapiro-Wilk normality test	0.961667	5.583292e-12	Possible non-normality
G3	649	11.906009	3.230656	0.0	19.0	-0.912909	2.712204	Shapiro-Wilk normality test	0.925981	2.415986e-17	Possible non-normality

Linearity, residual, and outlier diagnostics:

n_complete_pairs	pearson_r	pearson_p_value	spearman_rho	spearman_p_value	absolute_difference_pearson_spearman	regression_intercept	regression_slope	rmse	breusch_pagan_statistic	breusch_pagan_p_value	homoscedasticity_decision_alpha_0_05	mahalanobis_critical_chi_square_975_df2	mahalanobis_outlier_count_975	cooks_distance_threshold_4_over_n	cooks_distance_flag_count
649	0.918548	5.642401e-263	0.944451	0.0	0.025903	0.121966	1.01849	1.278096	15.758766	0.000072	Possible heteroscedasticity	7.377759	19	0.006163	20

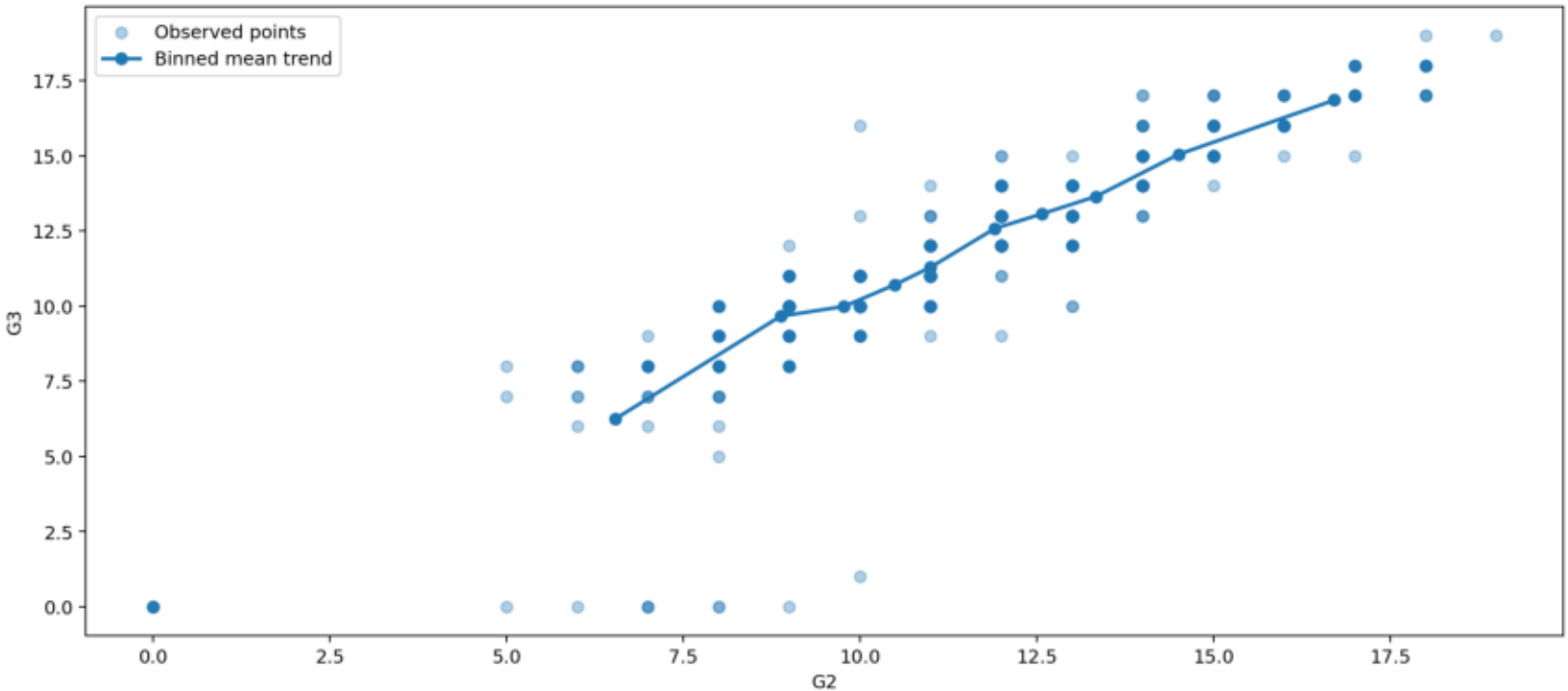
Correlation Assumptions: Scatterplot with Linear Fit

The main Pearson correlation assumption is a reasonably linear relationship without extreme influential points.



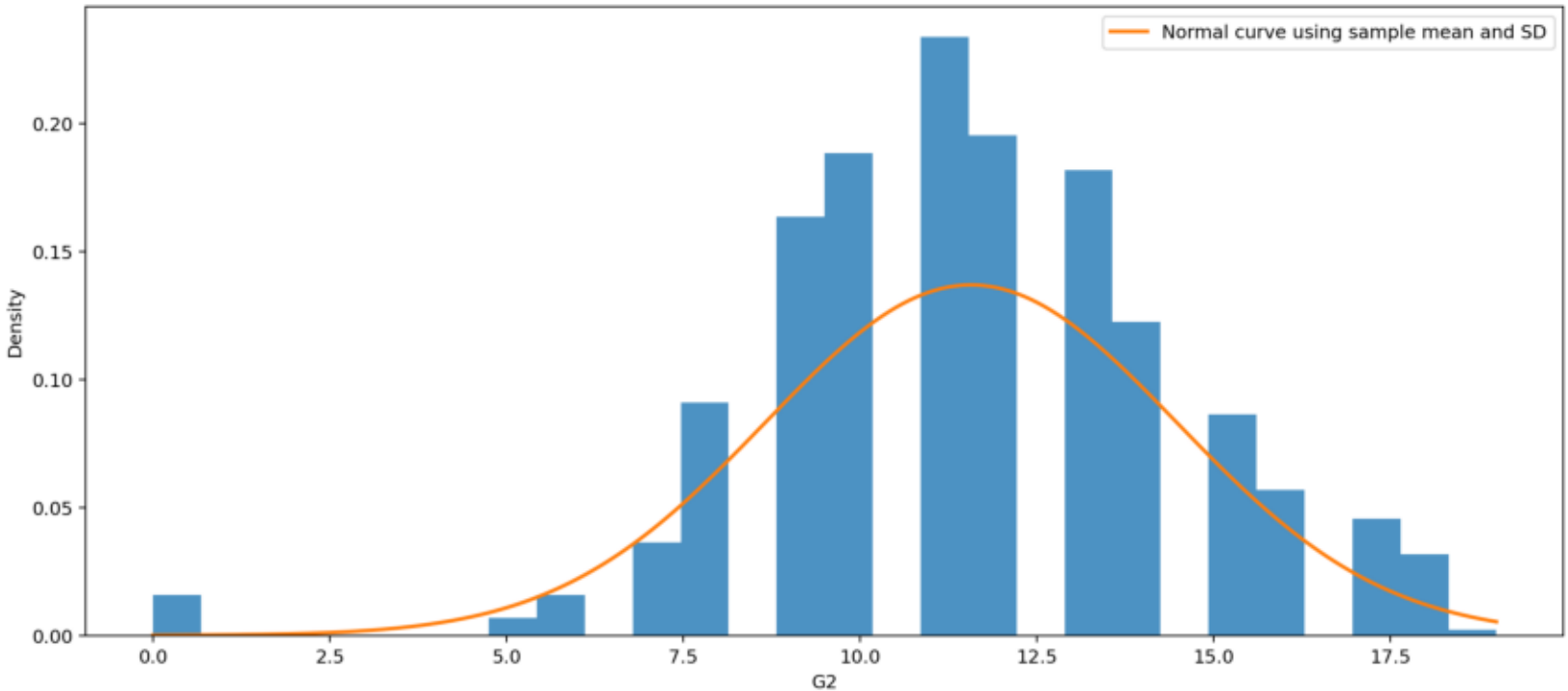
Correlation Assumptions: Linearity Trend Check

A smooth binned trend close to a straight line supports Pearson correlation interpretation.



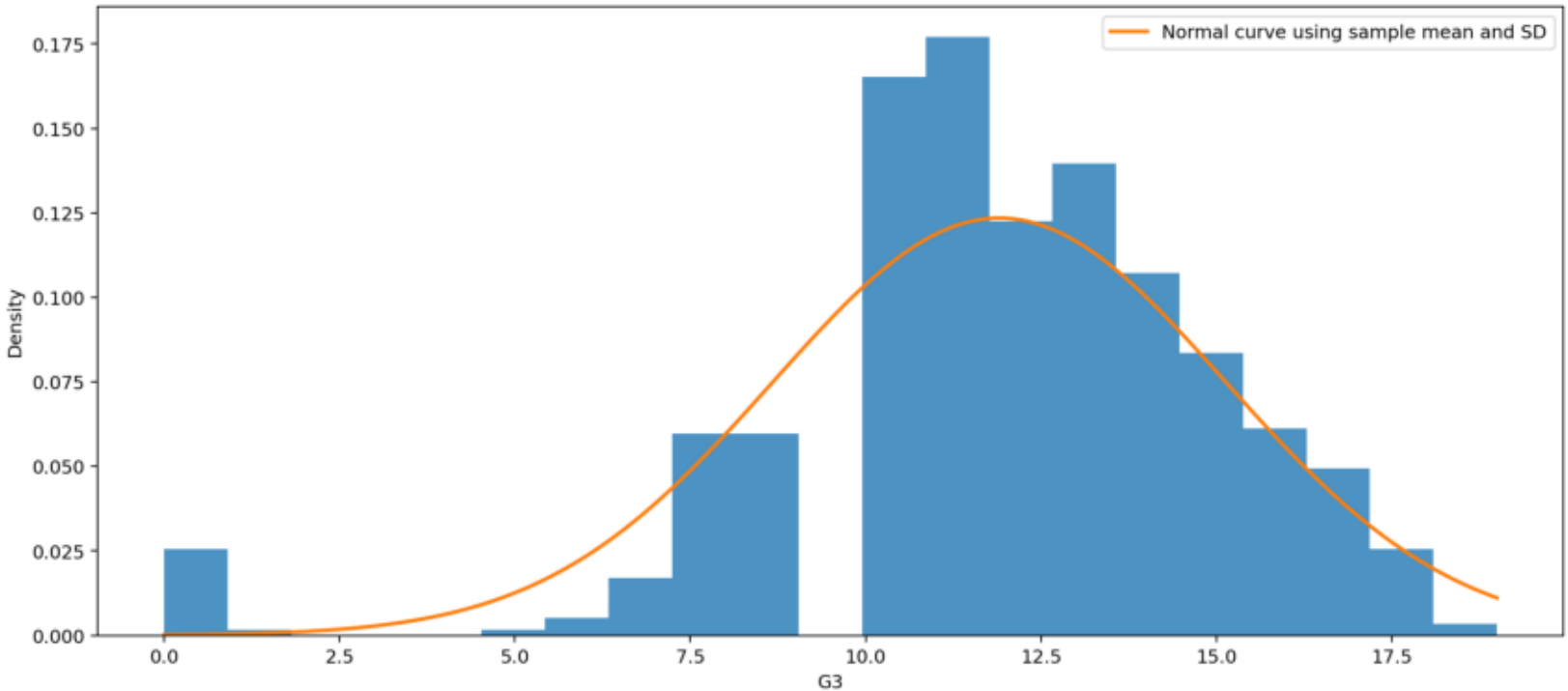
Correlation Assumptions: Histogram of G2

Histogram shape is used with the Q-Q plot and normality test for Pearson correlation assumptions.



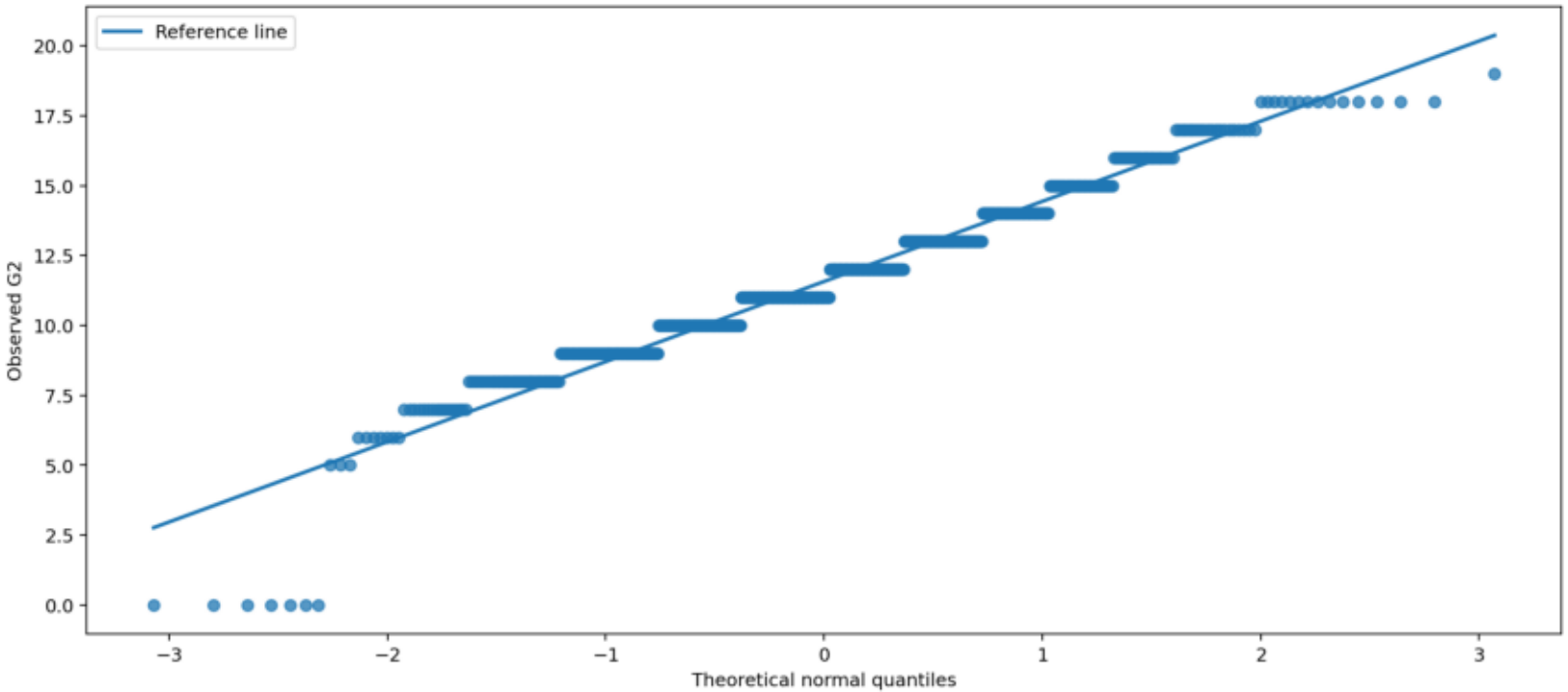
Correlation Assumptions: Histogram of G3

Histogram shape is used with the Q-Q plot and normality test for Pearson correlation assumptions.



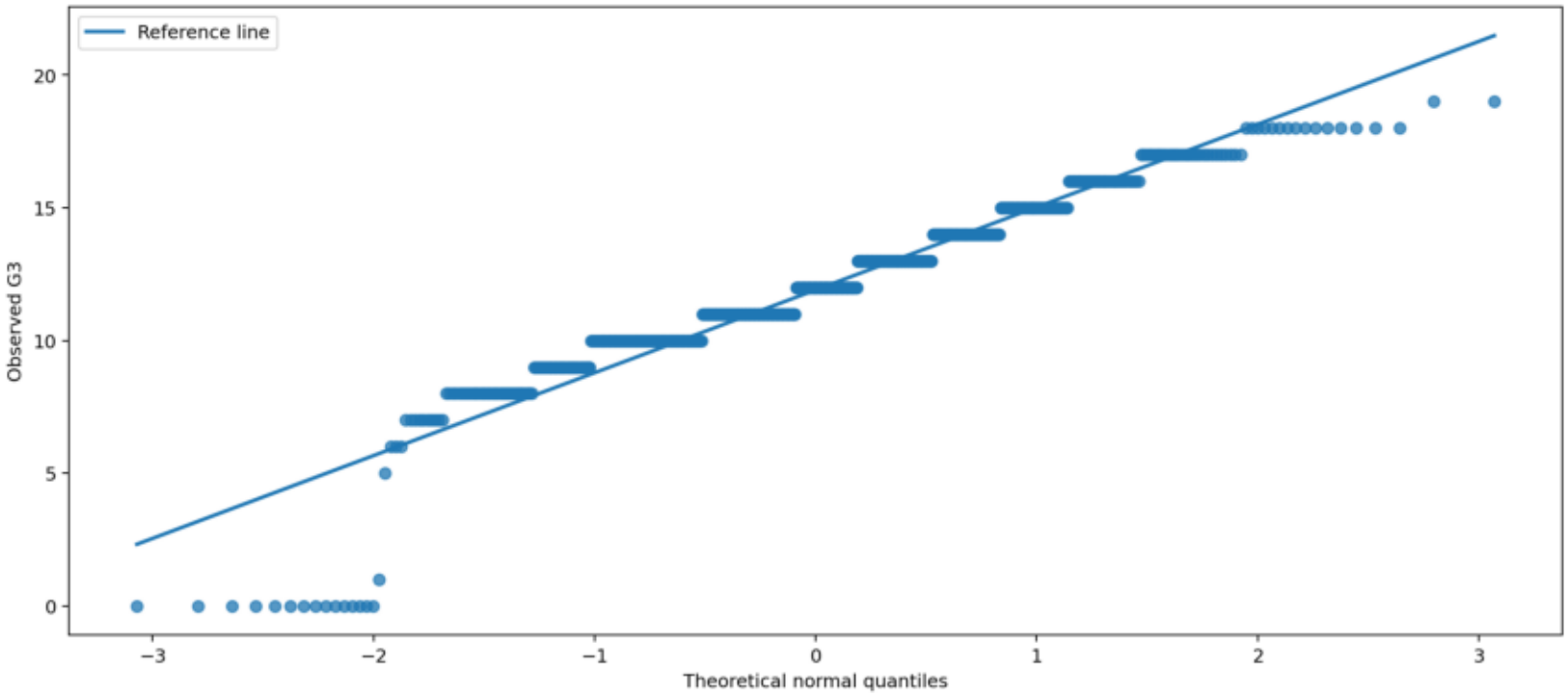
Correlation Assumptions: Q-Q Plot of G2

Points close to the reference line suggest the variable is approximately normal.



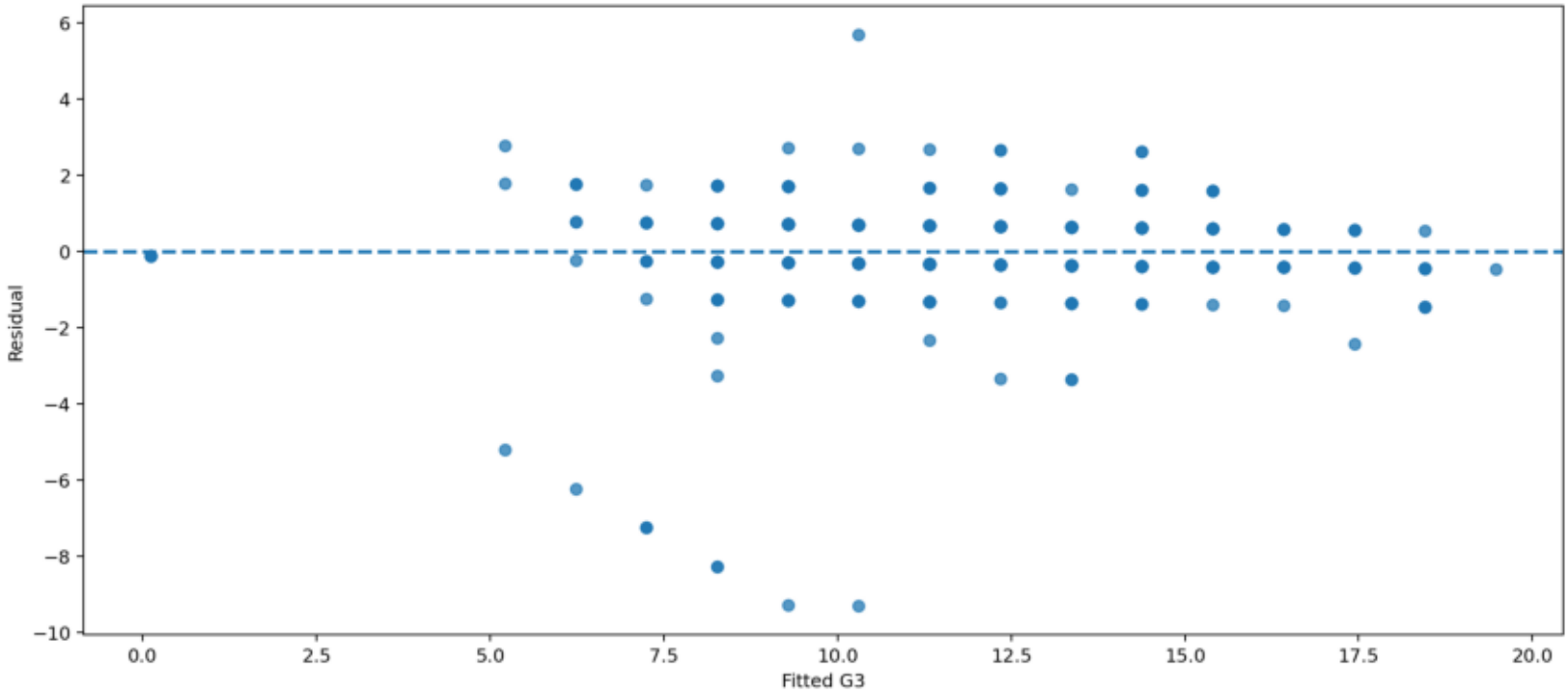
Correlation Assumptions: Q-Q Plot of G3

Points close to the reference line suggest the variable is approximately normal.



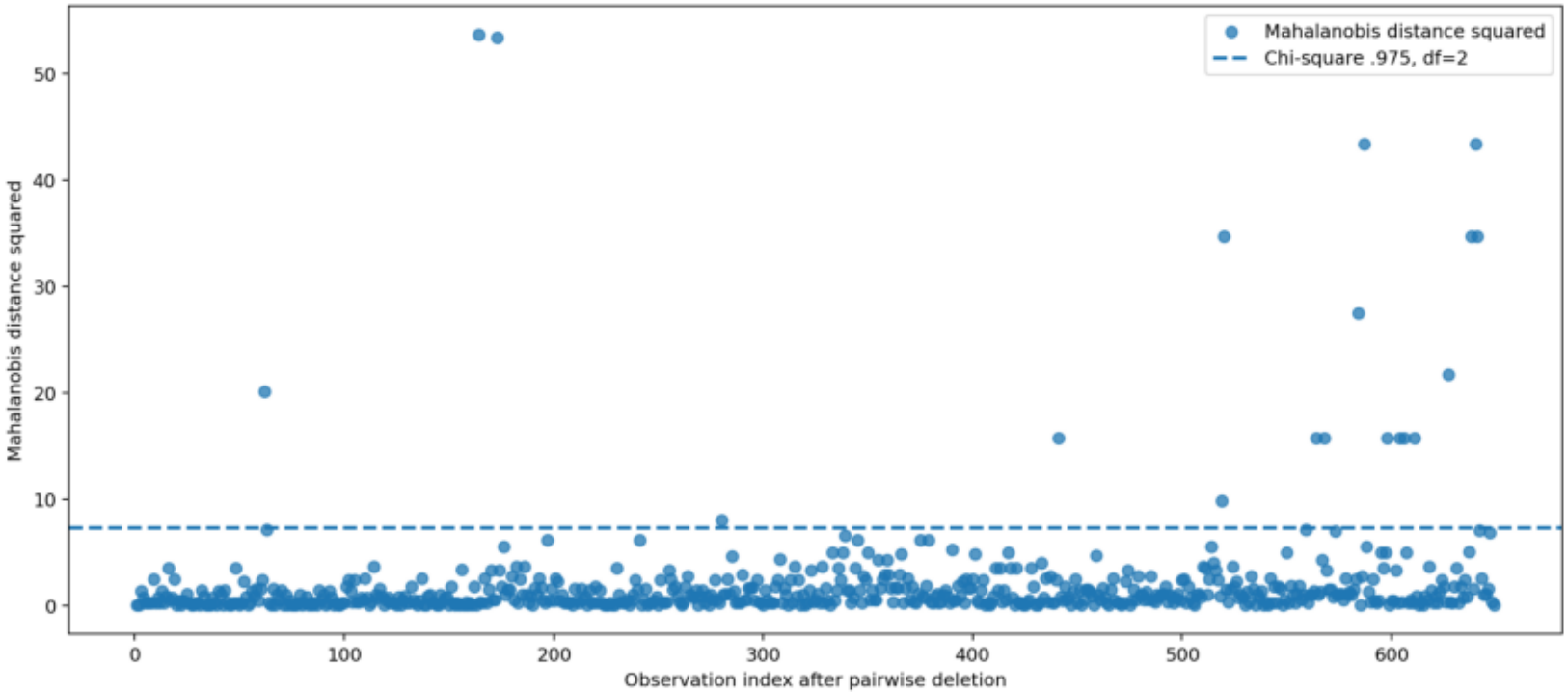
Correlation Assumptions: Residuals vs Fitted Values

A roughly even vertical spread around zero supports the constant variance assumption.



Correlation Assumptions: Bivariate Outlier Check

Large Mahalanobis distances or high Cook's distance can show points that drive the correlation.



Correlation Assumptions: Numeric Correlation Matrix

The selected pair is G2 with G3; nearby variables help check context and possible redundancy.