Summary of OLS Results - Model Variables

Variable	Coefficient, {'element': 'sup', 'data': 'a'}]	Std.Error	['Probability', {'element': 'sup', 'data': 'b'}] t-Statistic	Robust_SE, {'element': 'sup', 'data': 'b'}] Robust_Pr
Intercept	0.072085	0.005703	12.640531 0.000000*	0.005401 13.345601 0.000000*
MEAN	0.006065	0.001092	5.554264 0.000000*	0.001333 4.551307 0.000008*

OLS Diagnostics

Input Features	cancer_tracts.shp	Dependent Variable	CANRATE
Number of Observations	1401	[" Akaike's Information Criterion (AICc)", {'element': 'sup', 'data': 'd'}]	1445.439405
['Multiple R-Squared', {'element': 'sup', 'data': 'd'}]	0.021576	[' Adjusted R-Squared', {'element': 'sup', 'data': 'd'}]	0.020876
['Joint F-Statistic', {'element': 'sup', 'data': 'e'}]	30.849852	Prob(>F), (1,1399) degrees of freedom	0.000000*
['Joint Wald Statistic', {'element': 'sup', 'data': 'e'}]	20.714396	Prob(>chi-squared), (1) degrees of freedom	0.000005*
['Koenker (BP) Statistic', {'element': 'sup', 'data': 'f'}]	21.230999	Prob(>chi-squared), (1) degrees of freedom	0.000004*
['Jarque-Bera Statistic', {'element': 'sup', 'data': 'g'}]	2445.779373	Prob(>chi-squared), (2) degrees of freedom	0.000000*

Notes on Interpretation

* An asterisk next to a number indicates a statistically significant p-value (p < 0.01).

a Coefficient: Represents the strength and type of relationship between each explanatory variable and the dependent variable.

b Probability and Robust Probability (Robust_Pr): Asterisk (*) indicates a coefficient is statistically significant (p < 0.01); if the Koenker (BP) Statistic [f] is statistically significant, use the Robust Probability column (Robust_Pr) to determine coefficient significance.

c Variance Inflation Factor (VIF): Large Variance Inflation Factor (VIF) values (> 7.5) indicate redundancy among explanatory variables.

d R-Squared and Akaike's Information Criterion (AICc): Measures of model fit/performance.

e Joint F and Wald Statistics: Asterisk (*) indicates overall model significance (p < 0.01); if the Koenker (BP) Statistic [f] is statistically significant, use the Wald Statistic to determine overall model significance.

f Koenker (BP) Statistic: When this test is statistically significant (p < 0.01), the relationships modeled are not consistent (either due to non-stationarity or heteroskedasticity). You should rely on the Robust Probabilities (Robust_Pr) to determine coefficient significance and on the Wald Statistic to determine overall model significance.

g Jarque-Bera Statistic: When this test is statistically significant (p < 0.01) model predictions are biased (the residuals are not normally distributed).

Variable Distributions and Relationships
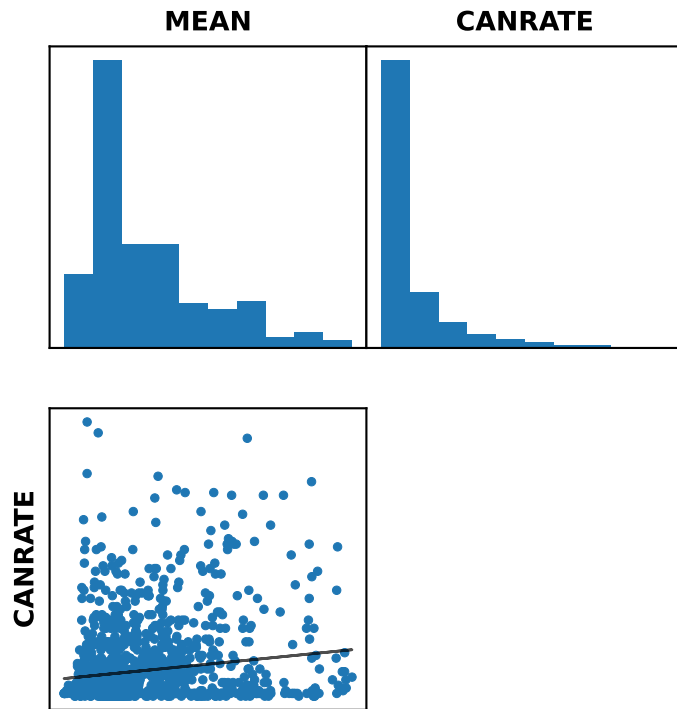
The above graphs are Histograms and Scatterplots for each explanatory variable and the dependent variable. The histograms show how each variable is distributed. OLS does not require variables to be normally distributed. However, if you are having trouble finding a properly-specified model, you can try transforming strongly skewed variables to see if you get a better result.

Each scatterplot depicts the relationship between an explanatory variable and the dependent variable. Strong relationships appear as diagonals and the direction of the slant indicates if the relationship is positive or negative. Try transforming your variables if you detect any non-linear relationships. For more information see the Regression Analysis Basics documentation.

Histogram of Standardized Residuals
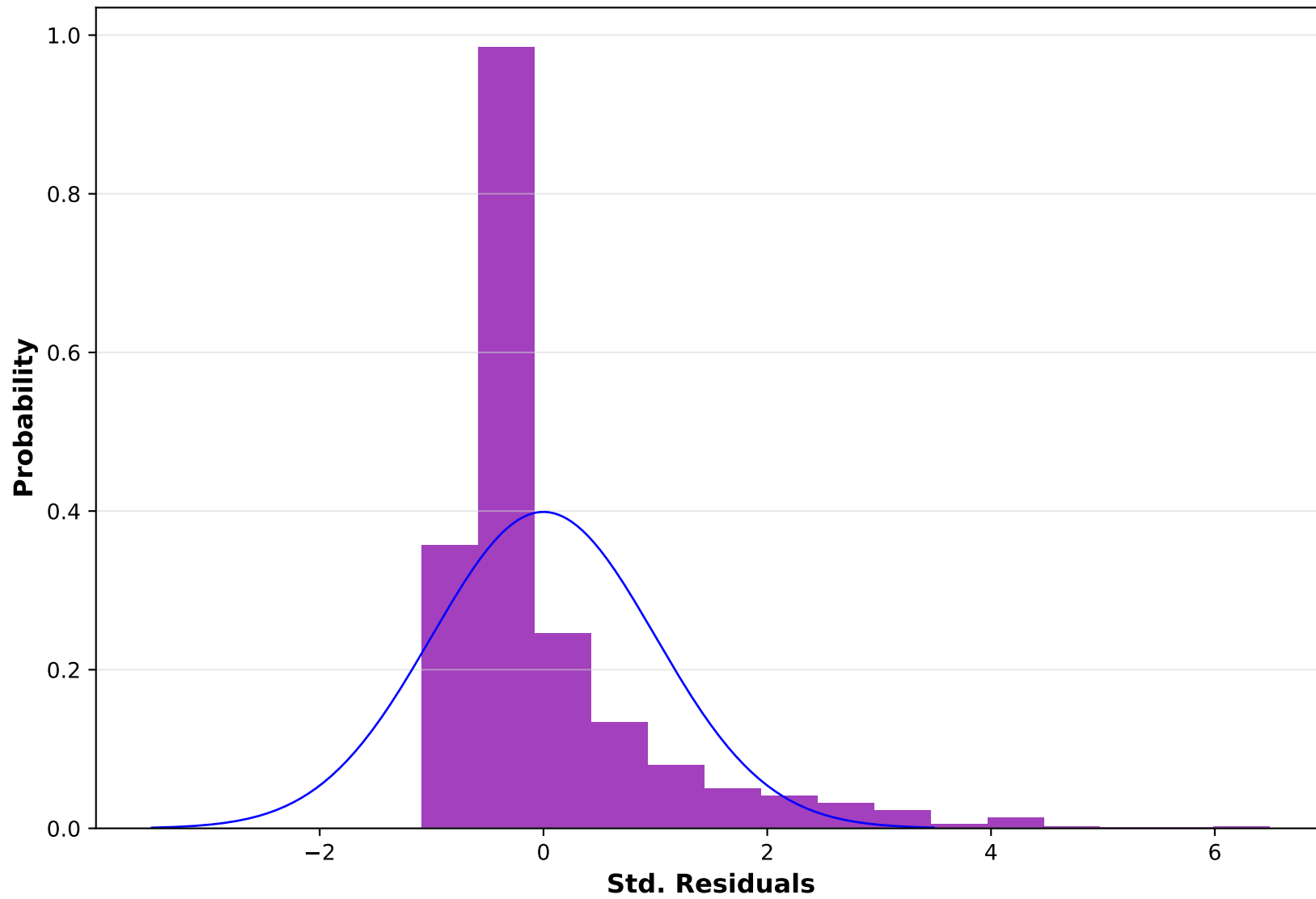
Ideally the histogram of your residuals would match the normal curve, indicated above in blue. If the histogram looks very different from the normal curve, you may have a biased model. If this bias is significant it will also be represented by a statistically significant Jarque-Bera p-value (*).

Residual vs. Predicted Plot
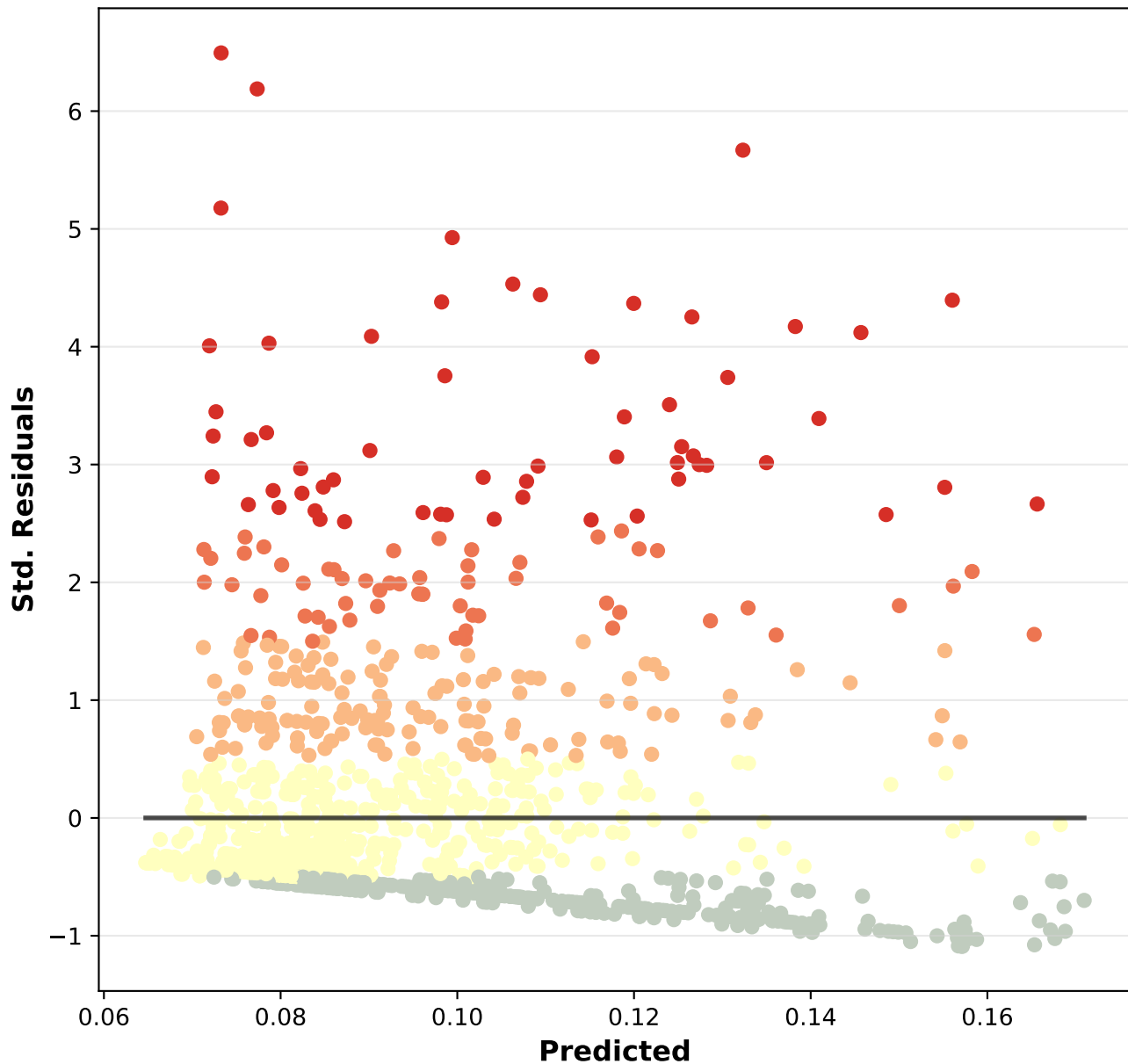
This is a graph of residuals (model over and under predictions) in relation to predicted dependent variable values. For a properly specified model, this scatterplot will have little structure, and look random (see graph on the right). If there is a structure to this plot, the type of structure may be a valuable clue to help you figure out what's going on.

Random Residuals
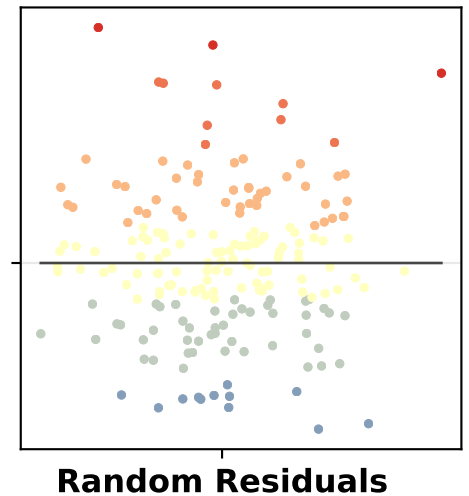
Ordinary Least Squares Parameters

Parameter Name	Input Value
Input Features	data/cancer_tracts.shp
Unique ID Field	UID
Output Feature Class	
Dependent Variable	CANRATE
Explanatory Variables	MEAN
Selection Set	False