

LINEAR REGRESSION ASSUMPTIONS

1. **Linearity:** The relationship between the dependent variable and independent variables should be linear. This means that the change in the dependent variable should be proportional to the change in the independent variables.
2. **Independence of Errors:** The errors (residuals), which are the differences between the observed values and the predicted values, should be independent of each other. There should be no pattern or correlation in the residuals.
3. **Homoscedasticity:** The variance of the errors should be constant across all levels of the independent variables. In other words, the spread of residuals should remain roughly the same as you move along the predicted values.
4. **Normality of Residuals:** The residuals should follow a normal distribution. This assumption is not critical for large sample sizes, but for smaller samples, it can affect the validity of hypothesis tests and confidence intervals.
5. **No or Little Multicollinearity:** Independent variables should not be highly correlated with each other. High multicollinearity can make it difficult to distinguish the individual effects of each independent variable on the dependent variable.
6. **No Endogeneity:** The independent variables should not be influenced by the errors. Endogeneity occurs when there is a feedback loop between the independent and dependent variables.
7. **No Autocorrelation:** Autocorrelation, or serial correlation, occurs when there is a correlation between the residuals at different time points (in time series data) or observations (in cross-sectional data).