**Multiple Linear Regression: (MLR)**

|  |
| --- |
| Linear Regression Model: Y = 0 + 1 X 1 + - - - +k X k +  Error Term:  ~ N(0,2)   is Normal i.i.d.  (independent & identically distributed)  (2 is constant)  Linear Regression Equation: Ŷ = b 0 + b 1 X 1 + - - - + b k X k  Residual: i = Y i – Ŷi  . . . |

|  |
| --- |
| **MLR: Model Building**  **Stage 1**. Data Preparation. Use exploratory plots to determine if linear relationships exist between the dependent variable and each independent variable. If not, use appropriate transformations to insure the linear assumption. Address possible outliers.  **Stage 2**. Significant Regression. Identify a statistically significant regression model for Predictive Regression. Identify statistically significant variables within the regression model for Descriptive Regression.  **Stage 3**. Valid Regression. Test entire model for violations of assumptions that include heteroscedasticity, non-normality, correlated errors, and multicollinearity.  **Stage 4**. Referent Regression. Select the final model to be used. The characteristics of the final model should include parsimony, tractability, and supportability.  . . . |

|  |
| --- |
| **Violations**  \* **Non-Linearity**.  The model must be a linear combination of independent variables.  The model must contain linearity in the coefficients.  \* **Heteroscedasticity**.  The variance of the error term must be constant.  \* **Non-Normality**.  The error terms must be normally distributed.  \* **Autocorrelation**.  The error terms must be independent.  \* **Multicollinearity**.  The independent variables must not contain large redundant correlations.  . . . |