



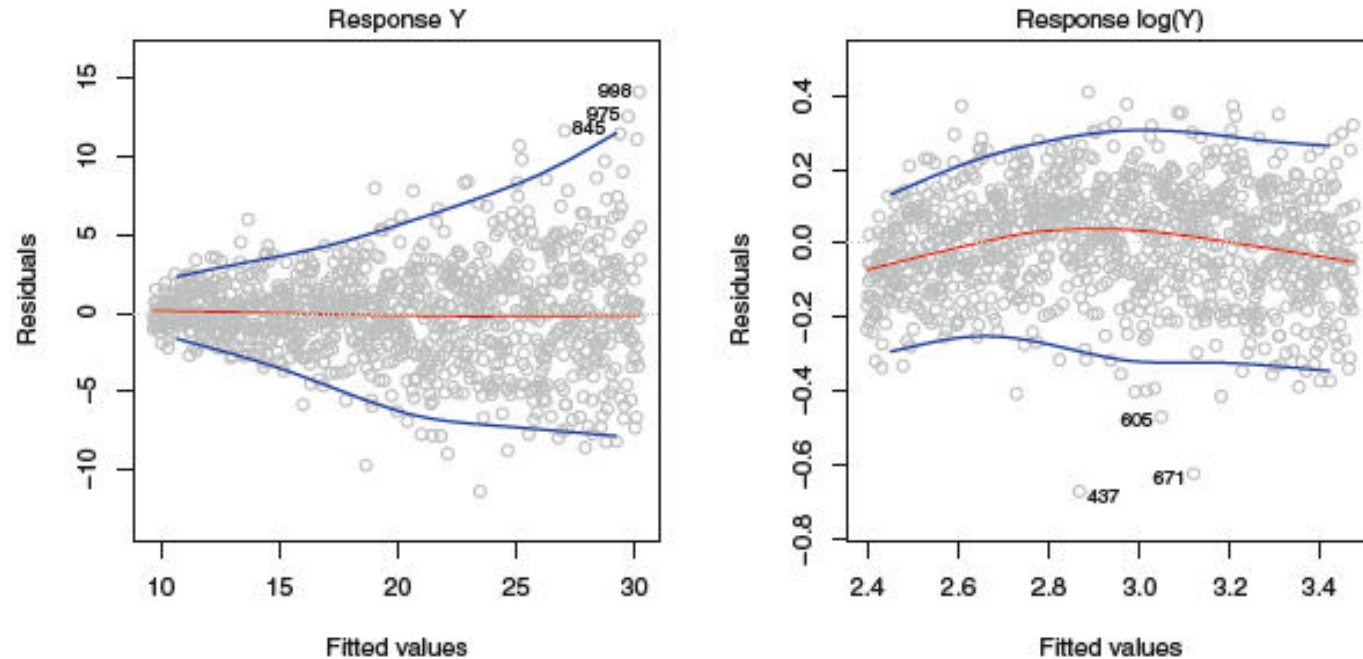
Start-Tech Academy

Multiple Linear Regression

heteroscedasticity.

$$\text{Var}(\epsilon_i) = \sigma^2.$$

Assumption : Variance of error term is independent of values of Y



Solution : Scaled down Y variable by using $\log(y)$ or \sqrt{y} instead



Multiple Linear Regression

Categorical Variables

airportYES	waterbodyLake	waterbodyLake and River	waterbodyRiver
1	0	0	1
0	1	0	0
0	0	0	0
1	1	0	0
0	1	0	0
1	0	0	0
1	0	0	1
0	1	0	0
1	0	0	0

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
waterbodyLake	0.264086	0.641963	0.411	0.6810
`waterbodyLake and River`	-0.687556	0.714023	-0.963	0.3361
waterbodyRiver	-0.291319	0.546656	-0.533	0.5943

