

Lectures 1 & 2

Regression Output Supplement

Correlation

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.806307604
R Square	0.650131952
Adjusted R Square	0.64656187
Standard Error	151.5687515
Observations	100

The correlation coefficient is the square root of the R Square, with the sign of the slope coefficient, -0.8063

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	4183527.721	4183527.721	182.1056015	4.44346E-24
Residual	98	2251362.469	22973.08642		
Total	99	6434890.19			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	6533.383035	84.51232199	77.30686935	1.22253E-89
Odometer	-0.031157739	0.002308896	-13.49465085	4.44346E-24

R-Square

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.806307604
R Square	0.650131952
Adjusted R Square	0.64656187
Standard Error	151.5687515
Observations	100

The R-Square, the Coefficient of Determination, tells us the proportion of the variability of the dependent variable that is explained by the variation in the independent variable.

$$R^2 = SSR/SST$$

ANOVA

	df	SS	MS	F	Significance F
Regression	1	4183527.721	4183527.721	182.1056015	4.44346E-24
Residual	98	2251362.469	22973.08642		
Total	99	6434890.19			

	Coefficients	Standard Error	t Stat	P-value
Intercept	6533.383035	84.51232199	77.30686935	1.22253E-89
Odometer	-0.031157739	0.002308896	-13.49465085	4.44346E-24

“Standard Error”

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.806307604
R Square	0.650131952
Adjusted R Square	0.64656187
Standard Error	151.5687515
Observations	100

There are three "standard errors":

SEE, Standard Error of Estimate

Standard error of the slope coefficient

Standard error of the intercept coefficient

ANOVA

	df	SS	MS	F	Significance F
Regression	1	4183527.721	4183527.721	182.1056015	4.44346E-24
Residual	98	2251362.469	22973.08642		
Total	99	6434890.19			

	Coefficients	Standard Error	t Stat	P-value
Intercept	6533.383035	84.51232199	77.30686935	1.22253E-89
Odometer	-0.031157739	0.002308896	-13.49465085	4.44346E-24

SEE

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.806307604
R Square	0.650131952
Adjusted R Square	0.64656187
Standard Error	151.5687515
Observations	100

The Standard Error of Estimate, SEE or s_e , is the square root of the MSE. This is the Std. Deviation of the residuals about the regression line.

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	4183527.721	4183527.721	182.1056015	4.44346E-24
Residual	98	2251362.469	22973.08642		
Total	99	6434890.19			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	6533.383035	84.51232199	77.30686935	1.22253E-89
Odometer	-0.031157739	0.002308896	-13.49465085	4.44346E-24

S_y^2, S_y

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.806307604
R Square	0.650131952
Adjusted R Square	0.64656187
Standard Error	151.5687515
Observations	100

The SS(Total) is the numerator of the dependent variable's variance. Therefore, the MS(Total) is the variance and it's square root is the standard deviation.

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	4183527.721	4183527.721	182.1056015	4.44346E-24
Residual	98	2251362.469	22973.08642		
Total	99	6434890.19	64998.89081	254.9488004	

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	6533.383035	84.51232199	77.30686935	1.22253E-89
Odometer	-0.031157739	0.002308896	-13.49465085	4.44346E-24

F-Ratio

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.806307604
R Square	0.650131952
Adjusted R Square	0.64656187
Standard Error	151.5687515
Observations	100

The MSR is the amount of the SST explained per df. The MSE is the amount of the SST unexplained per remaining df. The F-ratio is the ratio of these two measures.

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	4183527.721	4183527.721	182.1056015	4.44346E-24
Residual	98	2251362.469	22973.08642		
Total	99	6434890.19			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	6533.383035	84.51232199	77.30686935	1.22253E-89
Odometer	-0.031157739	0.002308896	-13.49465085	4.44346E-24

t-ratio

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.806307604
R Square	0.650131952
Adjusted R Square	0.64656187
Standard Error	151.5687515
Observations	100

The t-ratio measures the number of standard deviations the slope estimate is from zero.

$$-0.03115.../0.0023... = -13.49...$$

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	4183527.721	4183527.721	182.1056015	4.44346E-24
Residual	98	2251362.469	22973.08642		
Total	99	6434890.19			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	6533.383035	84.51232199	77.30686935	1.22253E-89
Odometer	-0.031157739	0.002308896	-13.49465085	4.44346E-24

t-ratio and F-Ratio

SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.806307604
R Square	0.650131952
Adjusted R Square	0.64656187
Standard Error	151.5687515
Observations	100

The square of a t-ratio is an F-ratio with 1 df in the numerator and the same number of df in the denominator, $182.1... = (-13.49...)^2$. This is why the significance is the same.

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	4183527.721	4183527.721	182.1056015	4.44346E-24
Residual	98	2251362.469	22973.08642		
Total	99	6434890.19			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	6533.383035	84.51232199	77.30686935	1.22253E-89
Odometer	-0.031157739	0.002308896	-13.49465085	4.44346E-24