

SOLUTIONS

► For all questions in this quiz, use the following information:

You would like to estimate whether there is linear relationship between the vehicle fatality rate and tax on case of beer. You have data for 48 states where variable mortrate denotes the number of vehicle fatalities per thousand of population in a given state and beertax denotes tax on case of beer in %. Below is a Stata output for the regression of the vehicle fatality rate on beer tax. Some information is deliberately erased. For all computations, use standard rounding rules and round the intermediate values to the third decimal point.

Source	SS	df	MS			
Model	.017107723	1	.017107723	Number of obs =	48	
Residual	.110559141	46	.00240346	F(1, 46) =	7.12	
				Prob > F =	0.0105	
				R-squared =	0.1340	
				Adj R-squared =	0.1152	
Total	.127666863	47	.002716316	Root MSE =	.04903	

mortrate	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
beertax	.0438755	.0164454	2.67	0.011	.0107726	.0769783
_cons	.1859073	.0105989	17.54	0.000	.1645729	.2072417

(1) The standard error of estimate for this regression is: 0.049. To find it, first divide SSE by the number of degrees of freedom, and then take a square root of that: $s_e = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{0.110559}{46}} = 0.049$.

(2) The coefficient of correlation between the vehicle fatality rate and tax on beer is square root of R-squared: 0.366. $r = \sqrt{\frac{SSR}{SST}} = \sqrt{\frac{0.0171}{0.1277}} = 0.366$. Please note that the sign of the slope estimate is also the sign of the correlation coefficient.

(3) The test-statistic for β_1 is distributed with 46 degrees of freedom (sample size is 48, so $n-2=48-2=46$).

(4) The p-value for the test of statistical significance for the slope coefficient is between 0.01 and 0.02 (to find p-value, compute t -statistic $= 0.0439/0.0164 = 2.6768$ and find it inside the Student t -table)

(5) We have enough evidence to infer that the beer tax and the vehicle fatality rate are linearly related at 5% significance level (Since p-value in previous question is between 1% and 2%, at 5% significance level we reject H_0 and conclude that the slope coefficient is statistically significant).

(6) The 95% confidence interval for the slope estimate is (0.012, 0.076). Confidence interval estimate is $b_1 \pm t_{\alpha/2, n-2} se(b_1) = 0.044 \pm 2.013 * 0.016 = (0.012, 0.076)$

(7) The variation in tax on case of beer explains about 13% of variation in the vehicle mortality rate. This is just the value of R-squared for this regression expressed in percent. To find R-squared, divide SSR by SST, or $0.0171/0.1277 = 0.1339$