

Regression in SPSS (Quiz)

Regression practical questions

We have seen that self-efficacy is a significant predictor of science test scores - students who are confident in their ability to perform science tasks well tend to accumulate more knowledge and understanding, at least as captured by the test. Another factor emphasised in theories of student learning is instrumental motivation, or the belief that learning science (in this example) is useful because it will be rewarded later in life. PISA contains a measure of instrumental motivation, INSMOVSCI, derived from four items in which students rated their agreement with such statements as "Many things I learn in my school science subjects will help me to get a job" (see PISA datafile description for further details).

Here you will explore the evidence for whether instrumental motivation for science learning is associated with better performance on the PISA science test.

Begin by looking at some basic summary information about the response variable **SCISCORE** and the predictor variable **INSMOVSCI**.

- Question: What is the range of the predictor?
- Question: What is the mean of the predictor?

Next use a scatter plot to visualise the association between **INSMOVSCI** and **SCISCORE**, and add a regression line to the plot to show the linear relationship .

- Question: What does the chart tell you about the relationship?

Next run the linear regression to formally estimate the effect of **INSMOVSCI** on **SCISCORE**. Make sure that you select the option to save the standardized residuals from the regression so that you can explore these with some plots.

- Question: Looking at the model summary table what proportion of the variance in the response is explained by the predictor?
- Question: Looking at the ANOVA table is the regression significant?
- Question: Looking at the Coefficients table what is the value of the intercept?
- Question: Looking at the Coefficients table what is the value of the slope?
- Question: What is the standardised slope coefficient and what does it mean?
- Question: Is the slope significant?
- Question: What is the 95% CI for the intercept?
- Question: What is the 95% CI for the slope?

Next plot a histogram of the residuals from the regression model.

- Question: Do the residuals look approximately Normal?

Now use a scatterplot to check whether there appears to be any relationship between the residuals and the predictor variable **INSMOVSCI**.

- Question: Is there any specific pattern in the plot?

Solutions to Regression practical questions

The instructions for creating the required descriptive statistics in SPSS are as follows:

- Select **Descriptives** from the **Descriptive Statistics** submenu available from the **Analyze** menu.
- Copy the **Science test score[SCISCORE]** and **Science instrumental motivation score[INSMOVSCI]** variables into the **Variable(s)** box.
- Click on the **Options** button.
- Ensure that the **Mean, Std. deviation, Minimum** and **Maximum** options are selected only.
- Click on the **Continue** button to return to the main window.
- Click on the **OK** button to run the command.

- Question: What is the range of the predictor?
- Question: What is the mean of the predictor?

Solutions: The output from SPSS is as follows:

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Science test score	5194	175.22	853.13	523.5944	102.57804
Science instrumental motivation score	4791	-1.93	1.74	.3819	.96666
Valid N (listwise)	4791				

The answers are as follows:

- The predictor takes values between -1.93 and 1.74.
- The mean of the predictor is .3819.

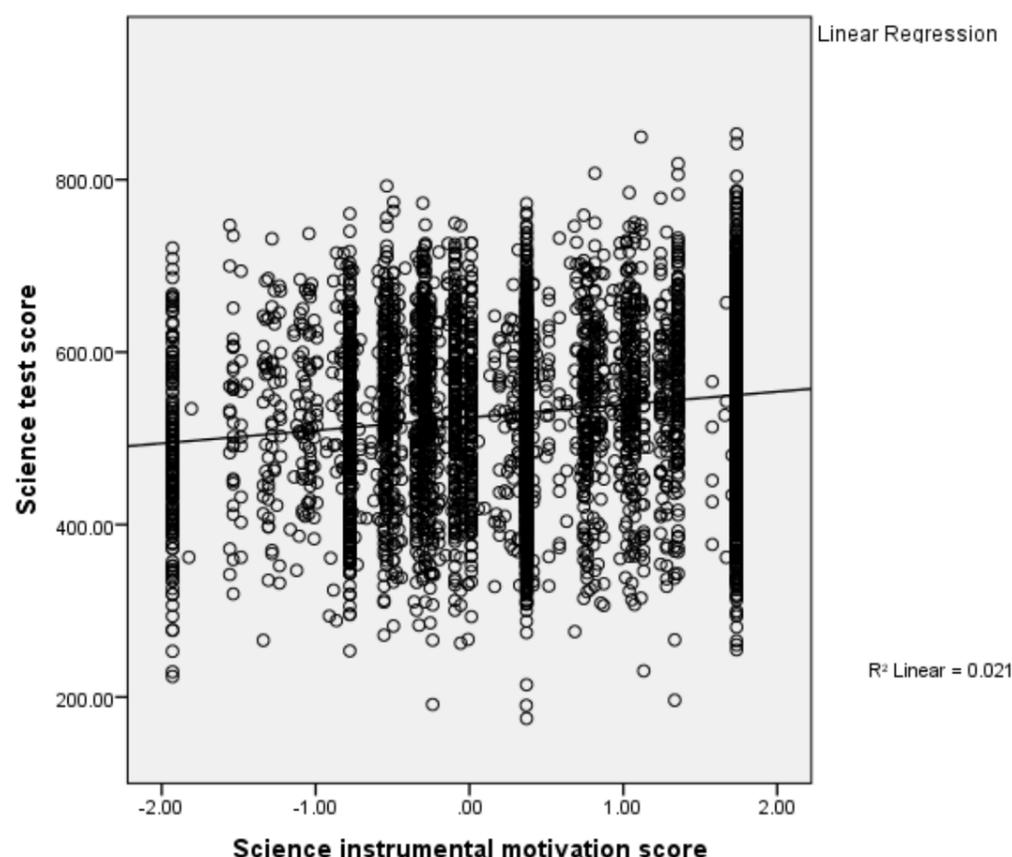
The instructions for creating the required scatterplot in SPSS are as follows:

- Select **Scatter/Dot** from the **Legacy diagnostics** available from the **Graphs** menu.
- Select **Simple Scatter** and click on Define to bring up the Simple Scatterplot window.
- Copy the **Science test score[SCISCORE]** variable into the **Y Axis** box.
- Copy the **Science instrumental motivation score[INSMOVSCI]** variable into the **X Axis** box.
- Click on the **OK** button.

- Double click with the left mouse button on the plot and it will pop out into a Chart Editor window.
- On the window click on the 5th button from the left on the bottom row of icons (It will say **Add Fit Line at Total** if you hover the mouse over it)
- On the **Properties** window that appears remove the tick next to **Attach label to line** as otherwise the equation is superimposed on the plot which looks untidy.
- Click on the **Close** button and the line will be added in the Chart Editor window.
- Finally close the **Chart Editor** window and the graph in the output window will now have the fixed line as shown below:

- Question: What does the chart tell you about the relationship?

Solution: The output from SPSS is as follows:



The scatterplot appears to indicate a significant positive relationship.

The instructions for running the required regression in SPSS are as follows:

- Select **Linear** from the **Regression** submenu available from the **Analyze** menu.
- Copy the **Science test score[SCISCORE]** variable into the **Dependent** box.
- Copy the **INSMOVSCI** variable into the **Independent(s)** box.
- Click on the **Statistics** button.
- On the screen appears add the tick for **Confidence Interval** to those for **Estimates** and **Model fit**.
- Click on the **Continue** button to return to the main window.
- Click on the **Save** button.
- On the screen appears select the tick for **Standardized** found under **Residuals**.
- Click on the **Continue** button to return to the main window.
- Click on the **OK** button to run the command.

- Question: Looking at the model summary table what proportion of the variance in the response is explained by the predictor?
- Question: Looking at the ANOVA table is the regression significant?
- Question: Looking at the Coefficients table what is the value of the intercept?
- Question: Looking at the Coefficients table what is the value of the slope?
- Question: What is the standardised slope coefficient and what does it mean?
- Question: Is the slope significant?
- Question: What is the 95% CI for the intercept?
- Question: What is the 95% CI for the slope?

Solutions: The output from SPSS is as follows:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.144 ^a	.021	.020	99.59691

a. Predictors: (Constant), Science instrumental motivation score

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1004032.097	1	1004032.097	101.218	.000 ^b
	Residual	47504701.272	4789	9919.545		
	Total	48508733.369	4790			

b. Predictors: (Constant), Science instrumental motivation score

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	524.263	1.547		338.857	.000	521.230	527.296
	Science instrumental motivation score	14.977	1.489	.144	10.061	.000	12.059	17.896

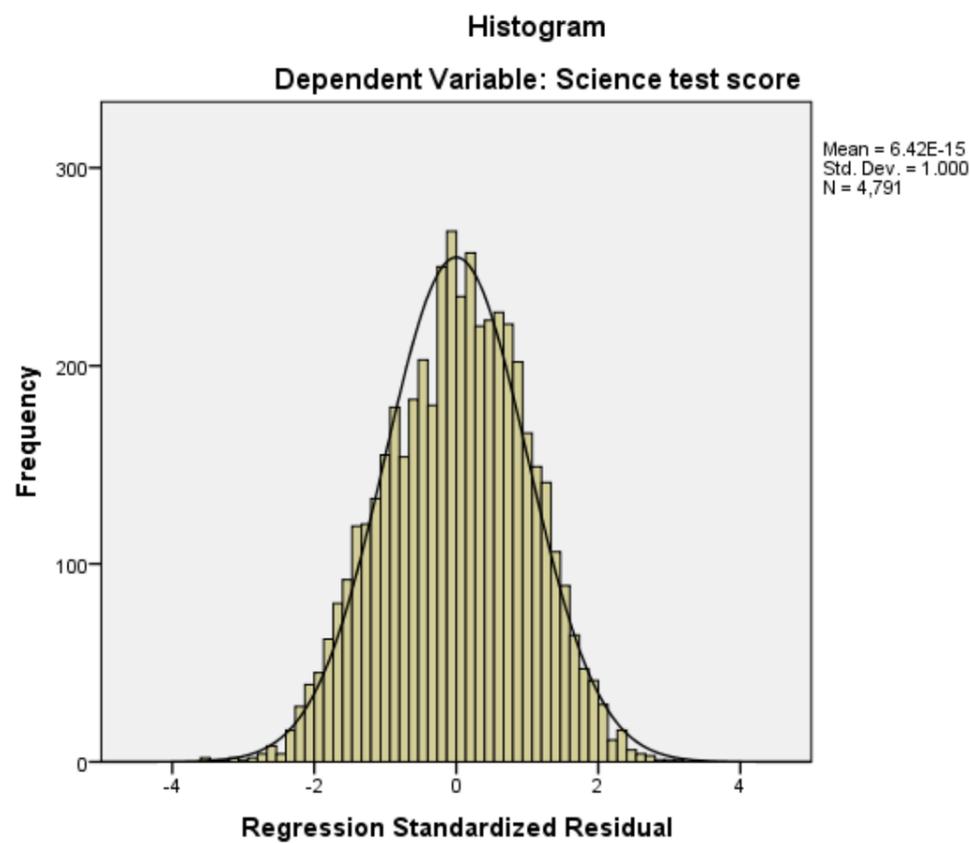
- The R squared provides the proportion of the variance explained by the regression and has the value .021.
- Here p is quoted as .000 (reported as $p < .001$) which is less than the conventional .05 level used to judge significance. We therefore have sufficient evidence to reject the null hypothesis that the predictors have no effect.
- The regression intercept takes value 524.263.
- The regression slope takes value 14.977.
- The standardised slope coefficient takes value .144 which represents the predicted change in **SCISCORE** for an increase of 1 standard deviation in **INSMOVSCI**.
- Here p is quoted as .000 (reported as $p < .001$) so we have sufficient evidence to reject the null hypothesis that the slope coefficient on **INSMOVSCI** is zero. We therefore have significant evidence to reject the null hypothesis that the slope is zero.
- The confidence interval for the intercept is 521.230 to 527.296.
- The confidence interval for the slope is 12.059 to 17.896.

The instructions for creating the residual histogram in SPSS are as follows:

- Select **Histogram** from the **Legacy dialogs** available from the **Graphs** menu.
- Copy the **Standardized Residual [ZRE_1]** variable into the **Variable** box.
- Click on the **Display normal curve** tick box.
- Click on the **OK** button to produce the graph.

- Question: Do the residuals look approximately Normal?

Solution: The output from SPSS is as follows:

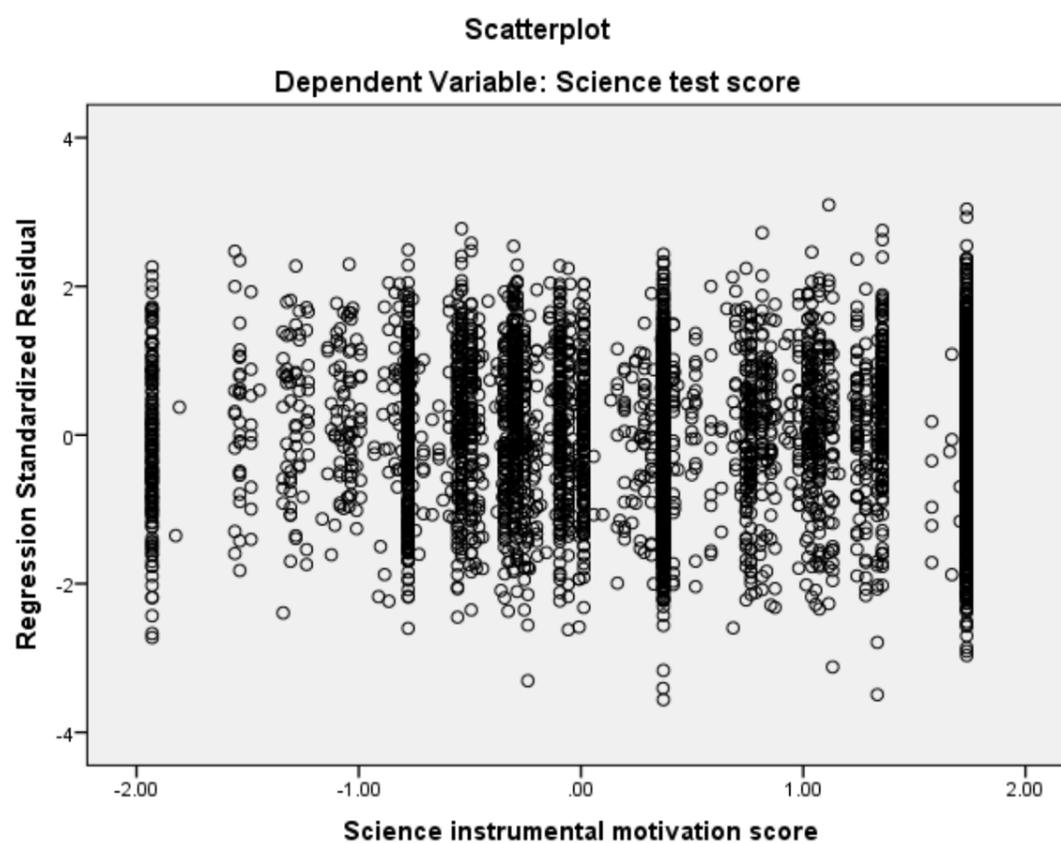


We hope to see the histogram of residuals roughly following the shape of the normal curve that is superimposed over them. The instructions for creating the required scatterplot in SPSS are as follows:

- Select **Scatter/Dot** from the **Legacy dialogs** available from the **Graphs** menu.
- Select Simple Scatter and click on Define to bring up the Simple Scatterplot window.
- Copy the **Standardized Residual [ZRE_1]** variable into the **Y Axis** box.
- Copy the **Science instrumental motivation score[INSMOVSCI]** variable into the **X Axis** box.
- Click on the **OK** button and the plot will appear.

- Question: Is there any specific pattern in the plot?

Solution: The output from SPSS is as follows:



We hope that the residuals show a random scatter when plotted against the predictor variable and also that their variability is constant across different values of the predictor variable.

As expected, science instrumental motivation is significantly positively related to achievement in science, but it is a less strong predictor than self-efficacy, accounting for only around 2% of the variance in science test scores, compared to the 12% accounted for by individual differences in students' self-efficacy.