

How can PROC REG be used in SAS, and can you provide an example?

Authored by
stats writer

June 23, 2024

RECOMMENDED CITATION

stats writer (2024). *How can PROC REG be used in SAS, and can you provide an example?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=148622>

PROC REG is a SAS procedure that is used to perform regression analysis. It allows for the exploration and modeling of relationships between variables in a dataset. This procedure can be used to predict the values of a dependent variable based on the values of one or more independent variables. It also allows for the identification of influential outliers and the assessment of the overall fit of the regression model. An example of using PROC REG in SAS would be to analyze the relationship between a person's income and their level of education. By using this procedure, one could determine if there is a significant correlation between these variables and use the results to make predictions or inform decision-making.

Use PROC REG in SAS (With Example)

You can use PROC REG in SAS to fit linear regression models.

You can use the following basic syntax to fit a :

```
proc regdata = my_data;  
model y = x;  
run;
```

This will fit the following linear regression model:

$$y = b_0 + b_1x$$

You can use the following basic syntax to fit a :

```
proc regdata = my_data;  
model y = x1 x2 x3;
```

run;

This will fit the following linear regression model:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3$$

The following example shows how to use PROC REG to fit a simple linear regression model in SAS along with how to interpret the output.

Example: How to Use PROC REG in SAS

Suppose we have the following dataset that contains information on hours studied and final exam score for 15 students in some class:

```
/*create dataset*/  
data exam_data;  
input hours score;  
datalines;  
1 64  
2 66  
4 76  
5 73  
5 74  
6 81
```

6 83

7 82

8 80

10 88

11 84

11 82

12 91

12 93

14 89

;

run;

/*view dataset*/

proc printdata=exam_data;

Obs	hours	score
1	1	64
2	2	66
3	4	76
4	5	73
5	5	74
6	6	81
7	6	83
8	7	82
9	8	80
10	10	88
11	11	84
12	11	82
13	12	91
14	12	93
15	14	89

We can use PROC REG to fit a simple linear regression model to this dataset, using hours as the predictor variable and score as the :

```
/*fit simple linear regression model*/  
proc regdata = exam_data;  
model score = hours;  
run;
```

The first table in the output shows a summary of the model fit:

The REG Procedure
Model: MODEL1
Dependent Variable: score

Number of Observations Read	15
Number of Observations Used	15

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	847.26698	847.26698	63.91	<.0001
Error	13	172.33302	13.25639		
Corrected Total	14	1019.60000			

Root MSE	3.64093	R-Square	0.8310
Dependent Mean	80.40000	Adj R-Sq	0.8180
Coeff Var	4.52852		

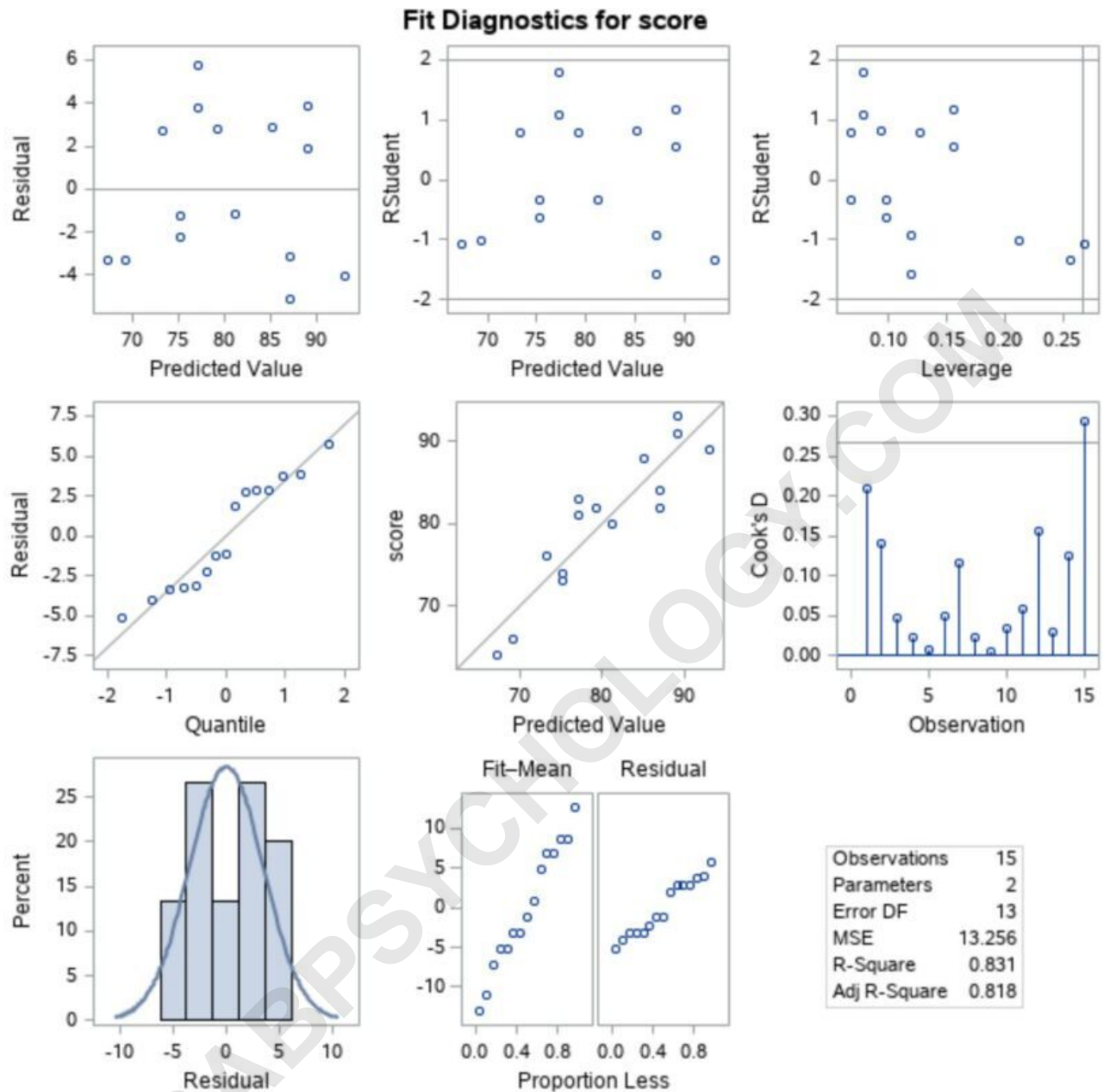
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	65.33395	2.10599	31.02	<.0001
hours	1	1.98237	0.24796	7.99	<.0001

The Parameter Estimates table contains the coefficient estimates for the model.

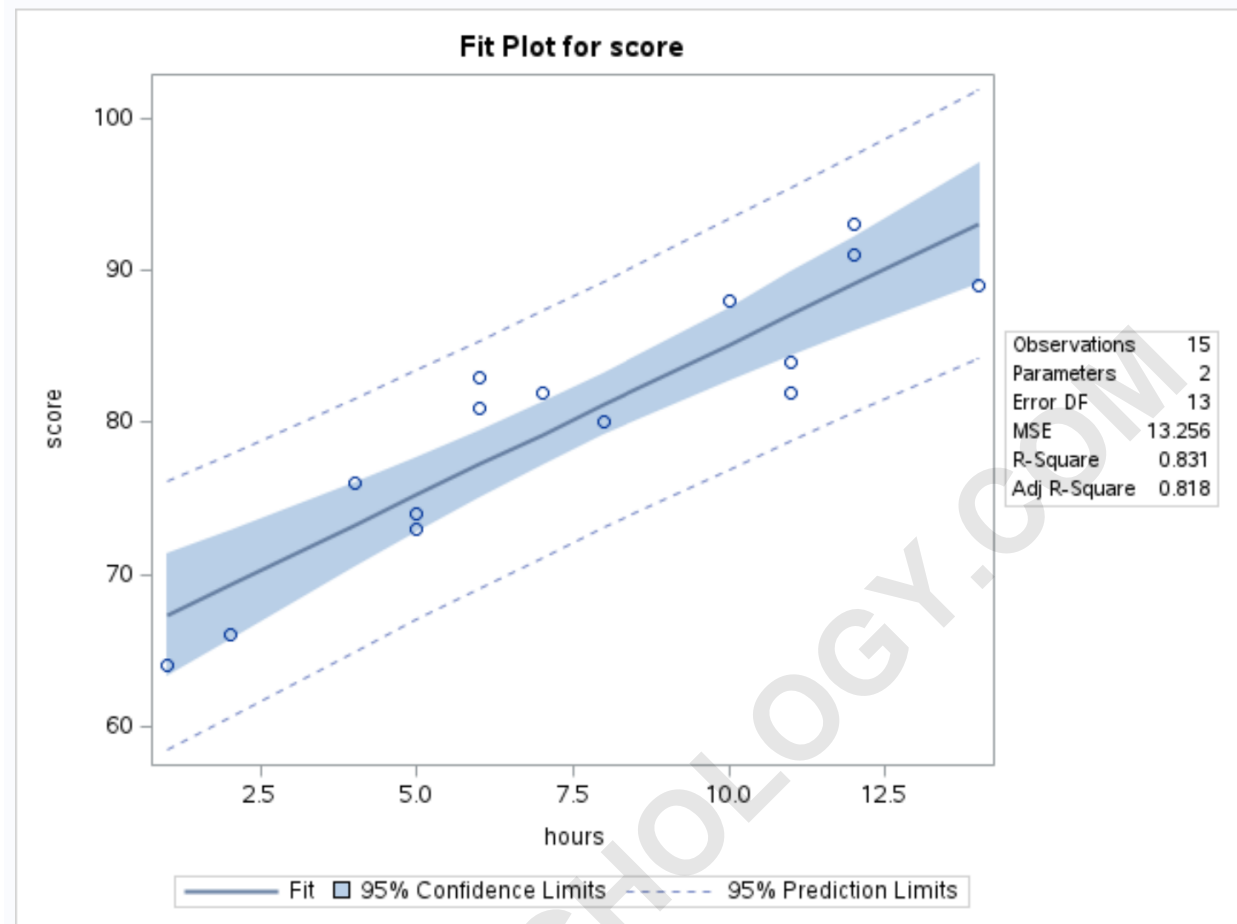
From this table we can see the fitted regression equation:

$$\text{Score} = 65.33 + 1.98 * (\text{hours})$$

The PROC REG procedure also produces residual plots that we can use to check if the are met:



Lastly, the PROC REG procedure produces a scatterplot of the raw data with the fitted regression line overlaid on top:



This plot allows us to visually see how well the regression line fits the data.

Note: You can find the complete documentation for PROC REG .

The following tutorials explain how to perform other common tasks in SAS: