

# How do I perform quantile regression in SAS?

Authored by  
**stats writer**

June 26, 2024

## RECOMMENDED CITATION

stats writer (2024). *How do I perform quantile regression in SAS?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=153230>

Quantile regression is a statistical method used to estimate the relationship between a set of independent variables and a continuous dependent variable at different quantiles of the data distribution. This allows for a more comprehensive analysis of the data, as it takes into account the variability and distribution of the dependent variable. In SAS, quantile regression can be performed by using the QUANTREG procedure, which allows for the selection of different quantiles and the specification of various model options. This procedure also provides diagnostic tools for assessing the model's fit and allows for the comparison of multiple quantile regression models. Overall, performing quantile regression in SAS provides a robust and effective approach for analyzing and understanding the relationship between variables in a dataset.

## Perform Quantile Regression in SAS

**Linear regression is a method we can use to understand the relationship between one or more predictor variables and a .**

**Typically when we perform linear regression, we're interested in estimating the mean value of the response variable.**

**However, we could instead use a method known as quantile regression to estimate *any* percentile value of the response value such as the 30th percentile, 90th percentile, 98th percentile, etc.**

**To perform quantile regression in SAS we can use the `proc quantreg` statement.**

**The following example shows how to perform quantile**

## regression in SAS in practice.

### Example: Performing Quantile Regression in SAS

Suppose we have the following dataset in SAS that shows the number of hours studied and corresponding exam score for students in some class:

```
/*create dataset*/  
data original_data;  
input hours score;  
datalines;  
1 75  
1 79  
2 78  
2 83  
2 85  
3 84  
3 84  
3 89  
4 93  
4 88  
4 79  
4 94  
5 96
```

**5 98**

**;**

**run;**

**/\*view dataset\*/**

**proc printdata=original\_data;**

Obs	hours	score
1	1	75
2	1	79
3	2	78
4	2	83
5	2	85
6	3	84
7	3	84
8	3	89
9	4	93
10	4	88
11	4	79
12	4	94
13	5	96
14	5	98

**Next, we'll fit a quantile regression model using hours studied as the predictor variable and exam score as the response variable.**

**We'll use the model to predict the expected 90th**

**percentile of exam scores based on the number of hours studied:**

```
/*perform quantile regression*/  
proc quantregdata=original_data;  
model score = hours / quantile = 0.9;  
run;
```

ARABPSYCHOLOGY.COM

## The QUANTREG Procedure

Model Information	
Data Set	WORK.ORIGINAL_DATA
Dependent Variable	score
Number of Independent Variables	1
Number of Observations	14
Optimization Algorithm	Simplex
Method for Confidence Limits	Inv_Rank

Number of Observations Read	14
Number of Observations Used	14

Summary Statistics						
Variable	Q1	Median	Q3	Mean	Standard Deviation	MAD
hours	2.0000	3.0000	4.0000	3.0714	1.3281	1.4826
score	79.0000	84.5000	93.0000	86.0714	7.1840	8.1543

Quantile Level and Objective Function	
Quantile Level	0.9
Objective Function	5.2500
Predicted Value at Mean	89.8214

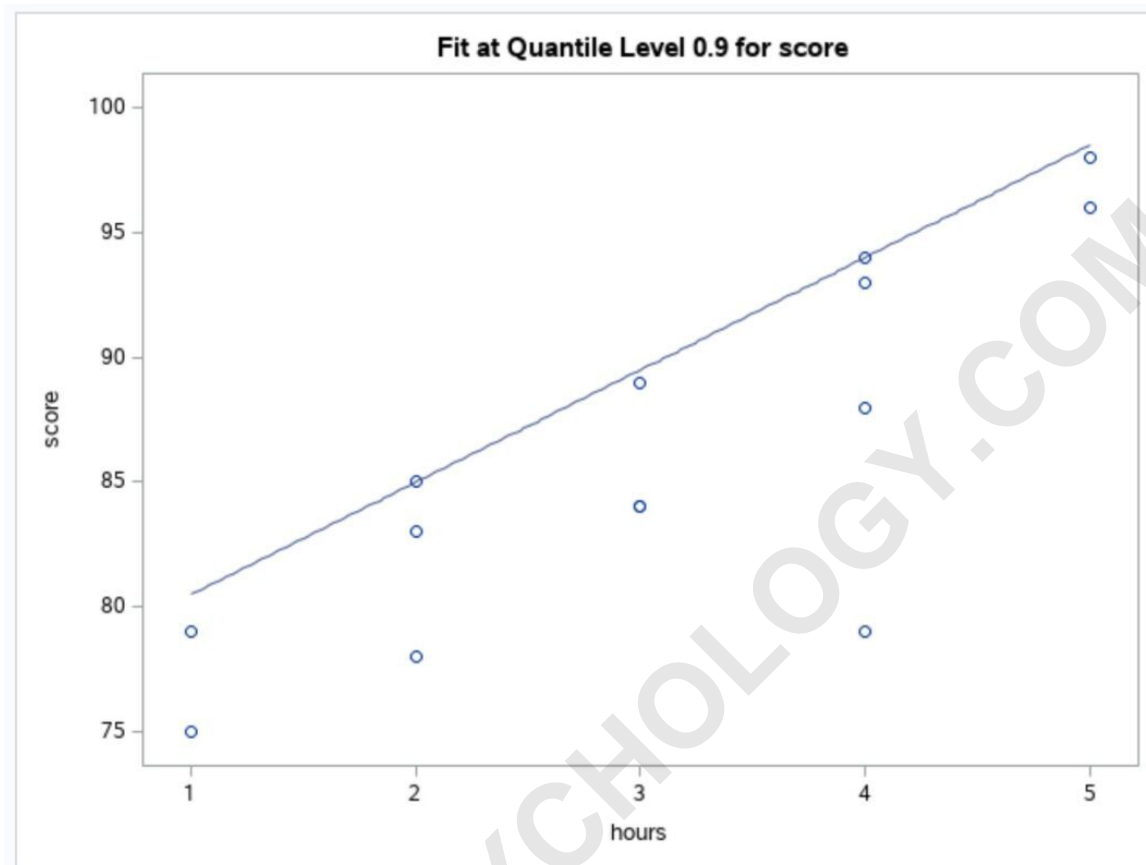
Parameter Estimates				
Parameter	DF	Estimate	95% Confidence Limits	
Intercept	1	76.0000	73.9037	Infy
hours	1	4.5000	-Infy	Infy

From the output, we can see the estimated regression equation:

**90th percentile of exam score = 76 + 4.5(hours)**

For example, the 90th percentile of scores for all students who study 2 hours is expected to be 85:

**90th percentile of exam score =  $76 + 4.5 \cdot (2) = 85$ .**



**Unlike a typical regression model, the fitted line for this regression model goes through the 90th percentile of each value of the predictor variable instead of the mean value.**