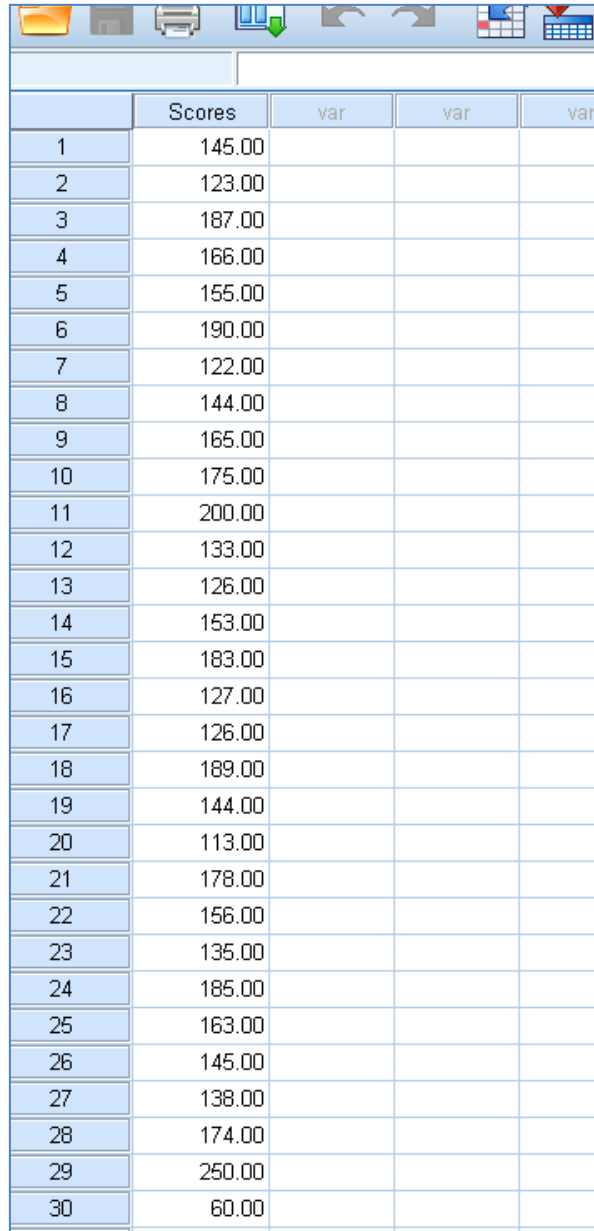


Exploring your data

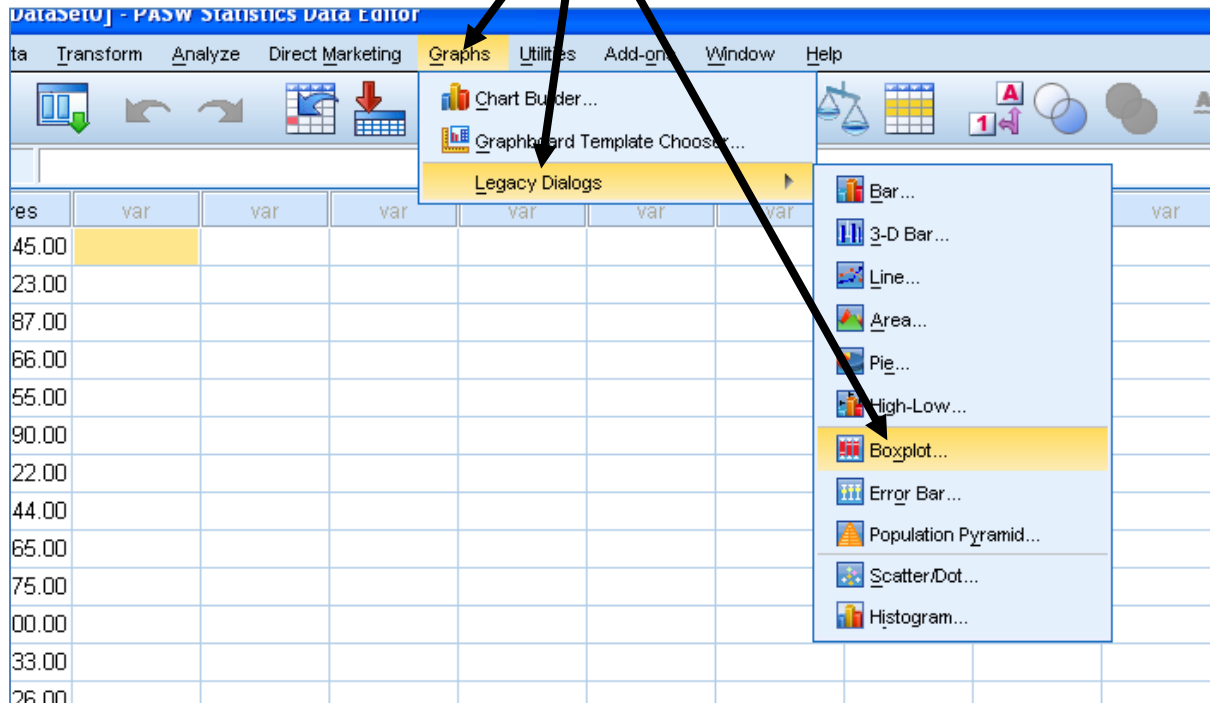
A. Checking for Outliers



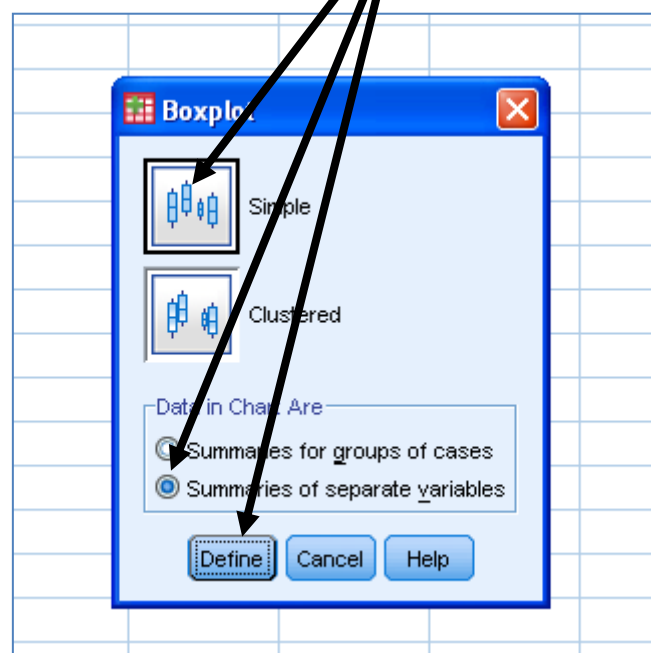
	Scores	var	var	var
1	145.00			
2	123.00			
3	187.00			
4	166.00			
5	155.00			
6	190.00			
7	122.00			
8	144.00			
9	165.00			
10	175.00			
11	200.00			
12	133.00			
13	126.00			
14	153.00			
15	183.00			
16	127.00			
17	126.00			
18	189.00			
19	144.00			
20	113.00			
21	178.00			
22	156.00			
23	135.00			
24	185.00			
25	163.00			
26	145.00			
27	138.00			
28	174.00			
29	250.00			
30	60.00			

You may have a data set which looks similar to the one above

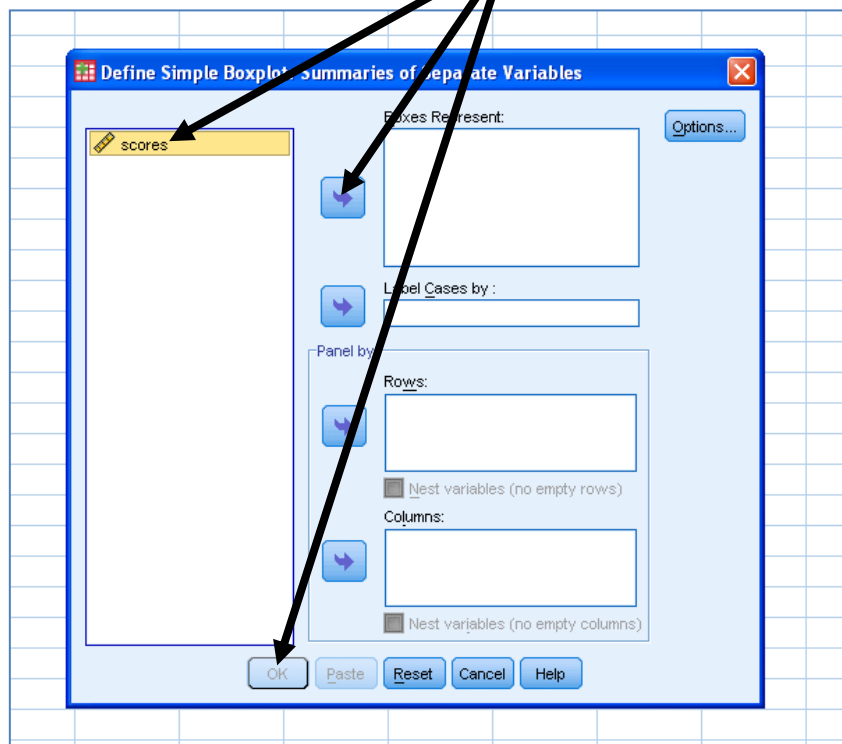
Click on **Graphs, Legacy Dialogs and Boxplot**



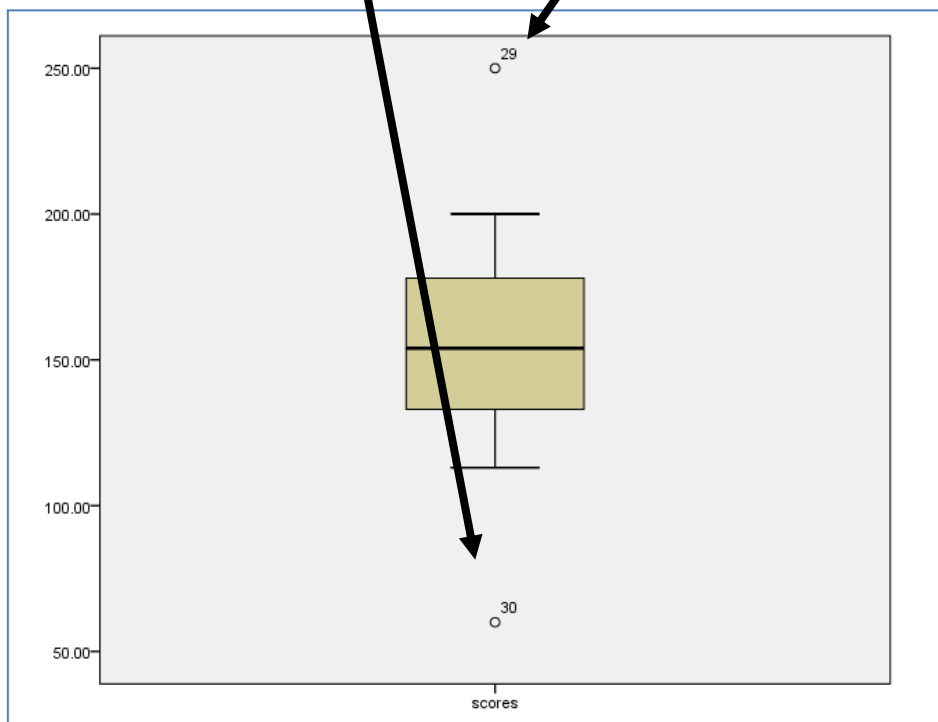
Click on **Simple, Summaries of separate variables and Define**



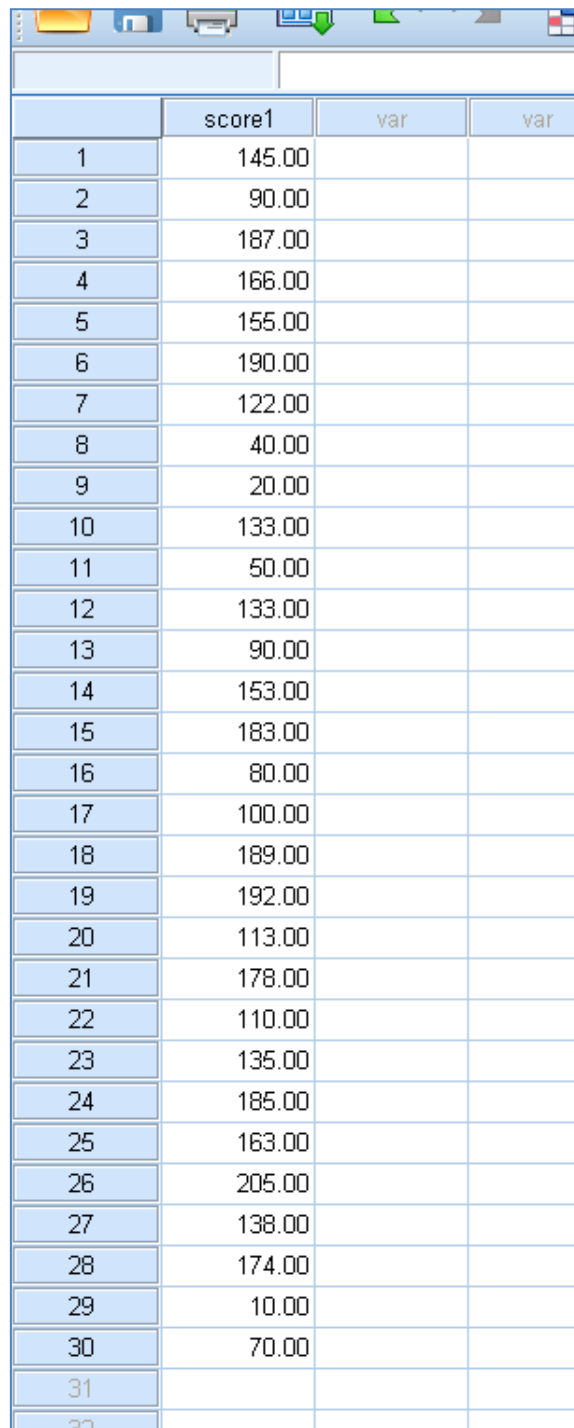
Click on the 'scores' and move it to the 'Boxes Represent' section, and click 'OK'



The following will appear in the output window. Cases 29 and 30 in our fictitious data set are outliers.



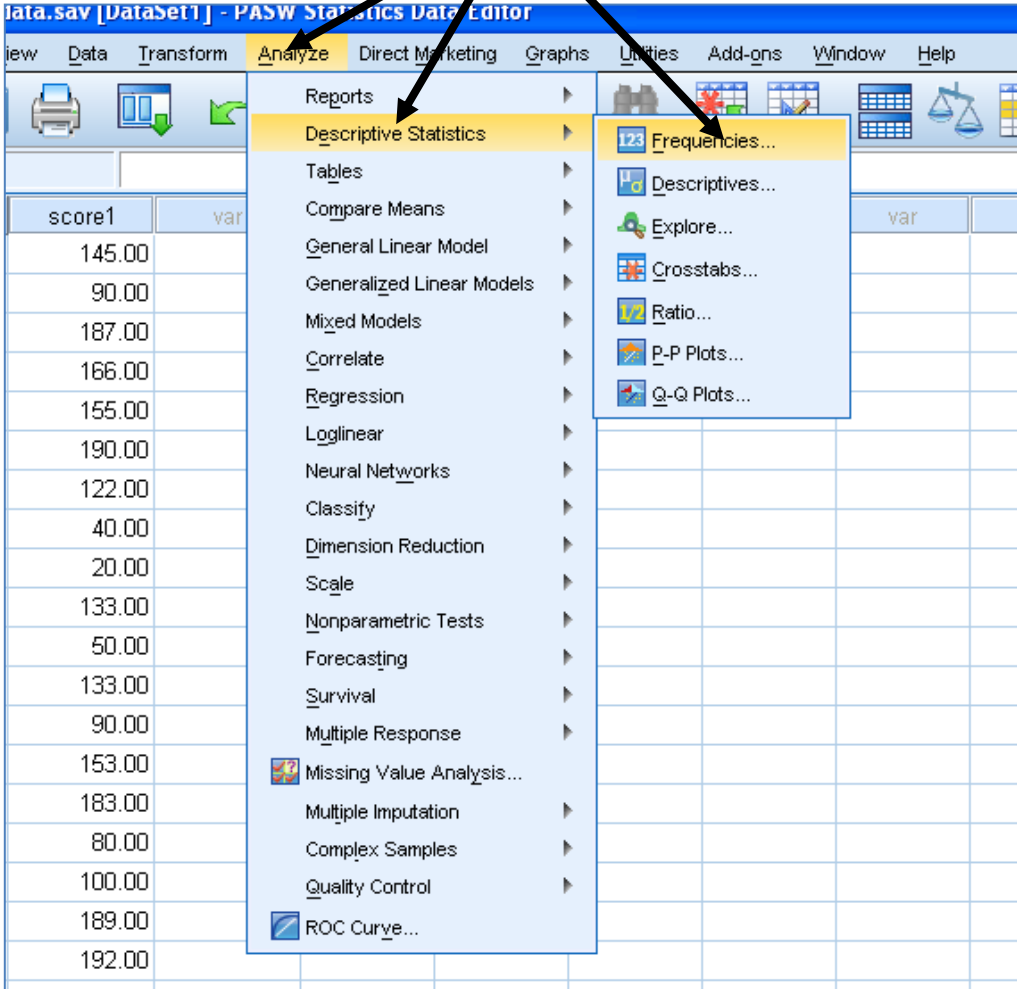
B. Checking for Normality



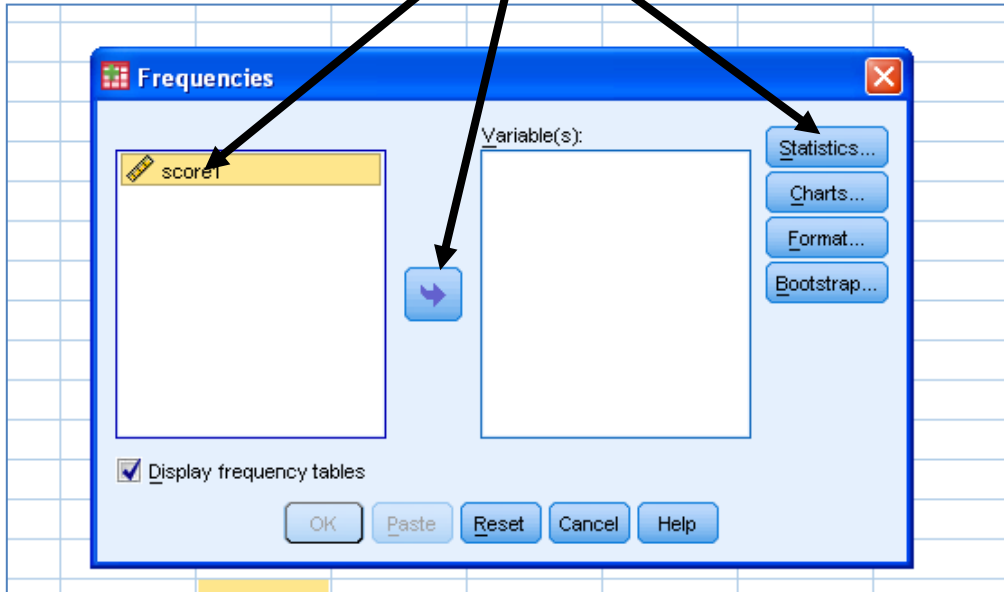
	score1	var	var
1	145.00		
2	90.00		
3	187.00		
4	166.00		
5	155.00		
6	190.00		
7	122.00		
8	40.00		
9	20.00		
10	133.00		
11	50.00		
12	133.00		
13	90.00		
14	153.00		
15	183.00		
16	80.00		
17	100.00		
18	189.00		
19	192.00		
20	113.00		
21	178.00		
22	110.00		
23	135.00		
24	185.00		
25	163.00		
26	205.00		
27	138.00		
28	174.00		
29	10.00		
30	70.00		
31			
32			

You may have a data set which looks similar to the one above

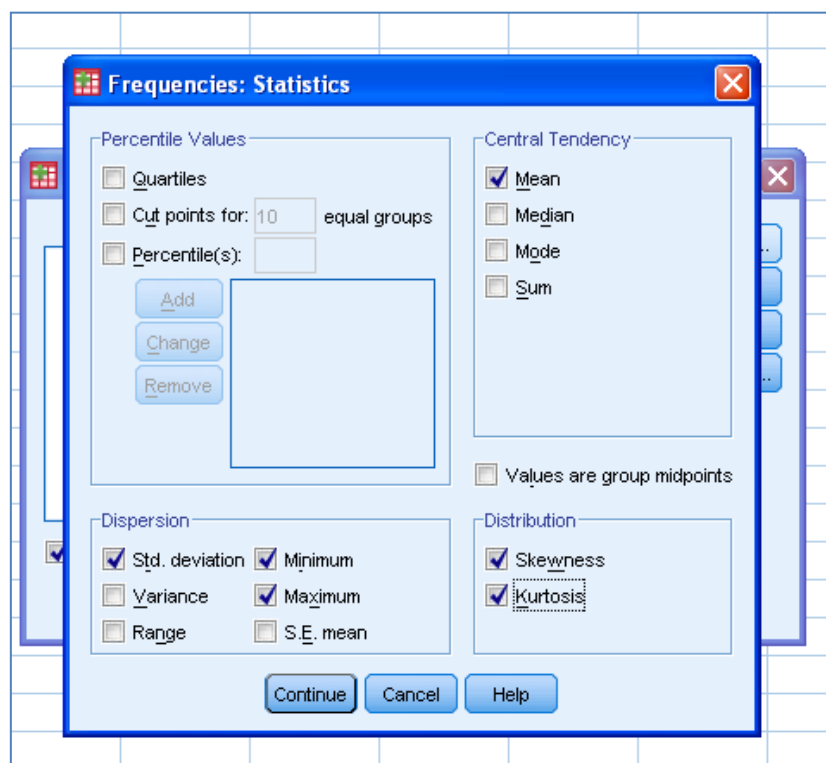
Click on **Graphs**, **Legacy Dialogs** and **Box-plot**



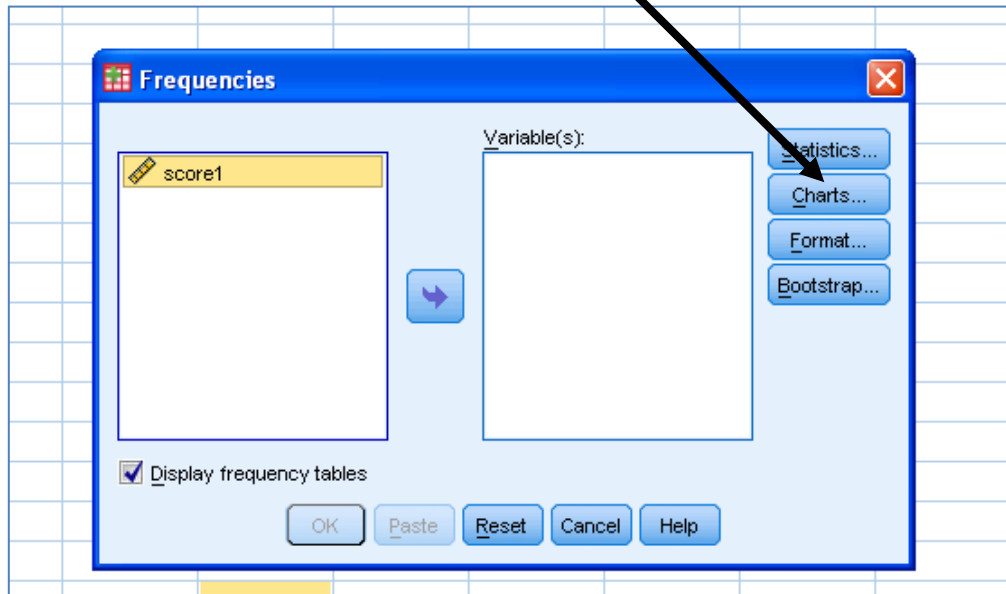
Click on the 'score1' and move it to the 'Variables (s)' section, and click on the box marked 'Statistics'



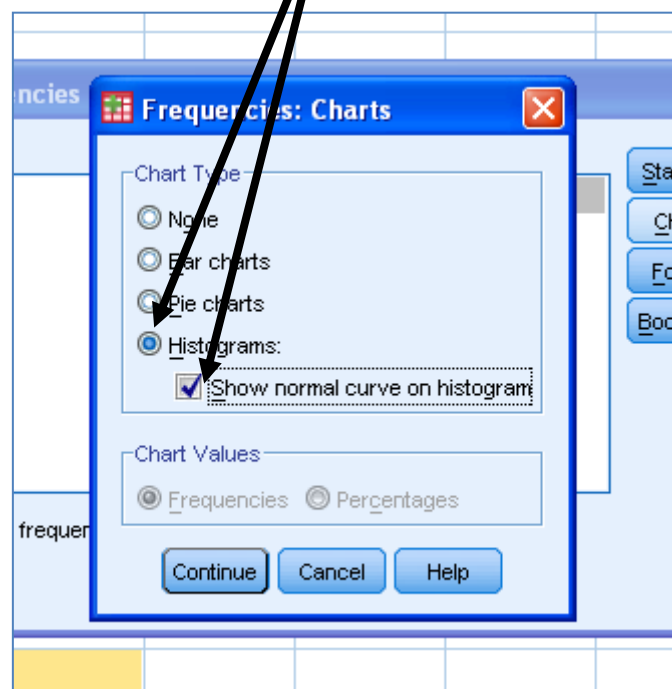
Under 'Central Tendency' check the box marked 'Mean' Under 'Dispersion' check 'Std deviation', 'minimum', 'maximum', and under 'Distribution' check the boxes marked 'Skewness' and 'Kurtosis' Then 'Continue'



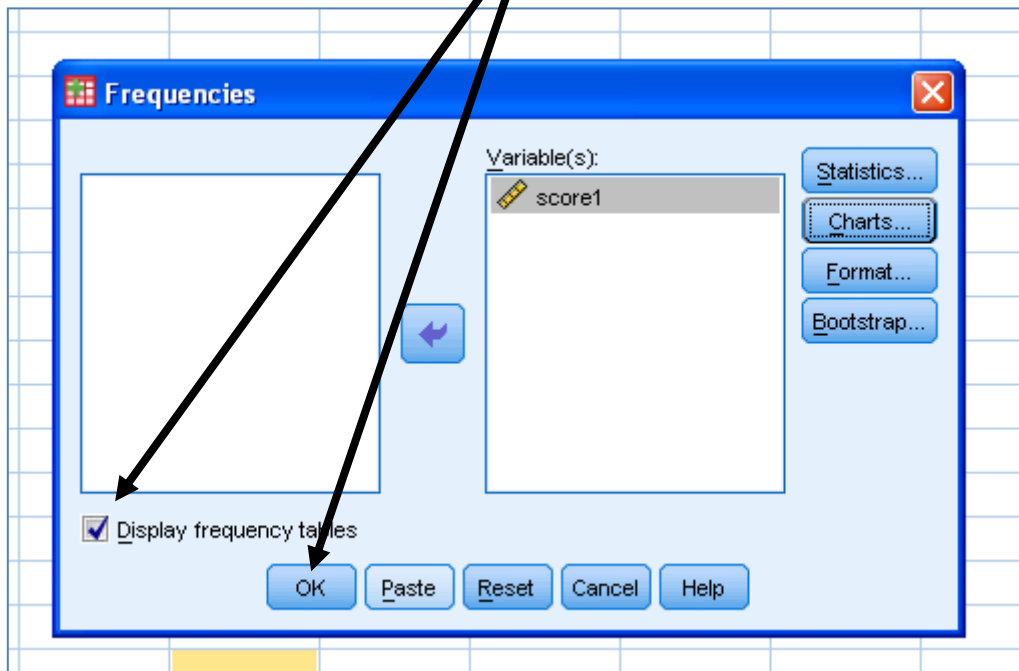
Click on the 'Charts'



Check the radio button for 'Histograms' and also 'Show normal curve on histogram'. Then 'Continue'



Uncheck the box marked 'Display frequency tables' and Click OK



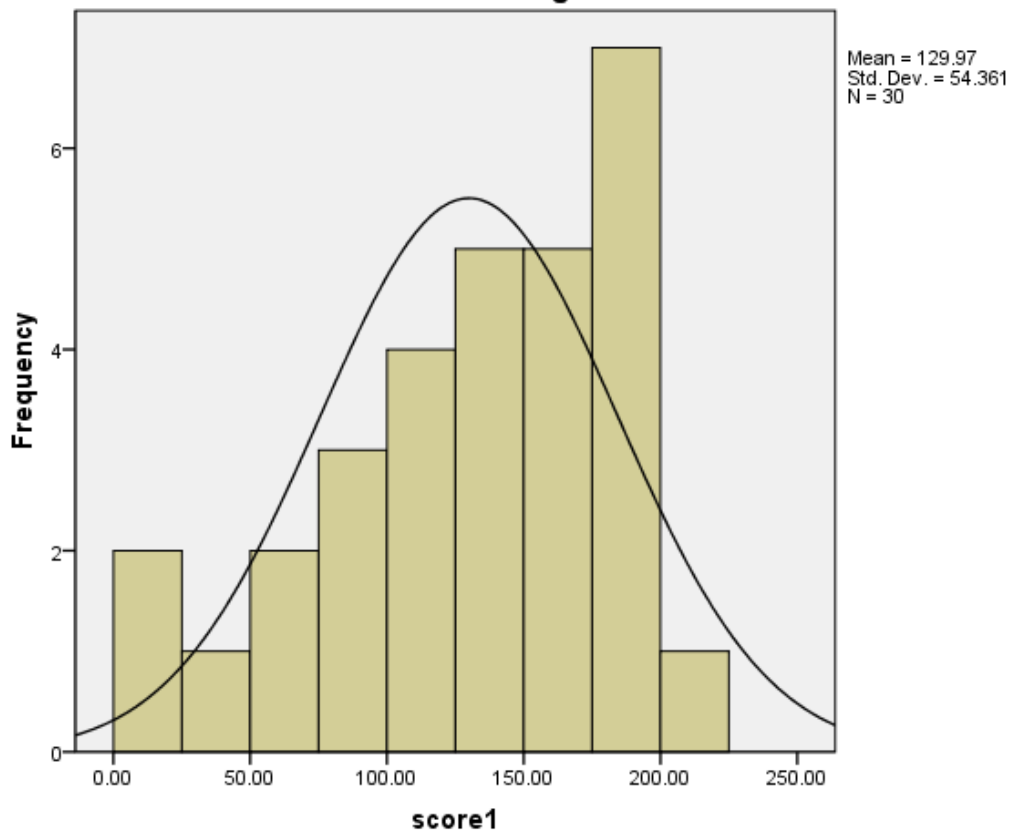
The output file appears as below

Statistics

score1

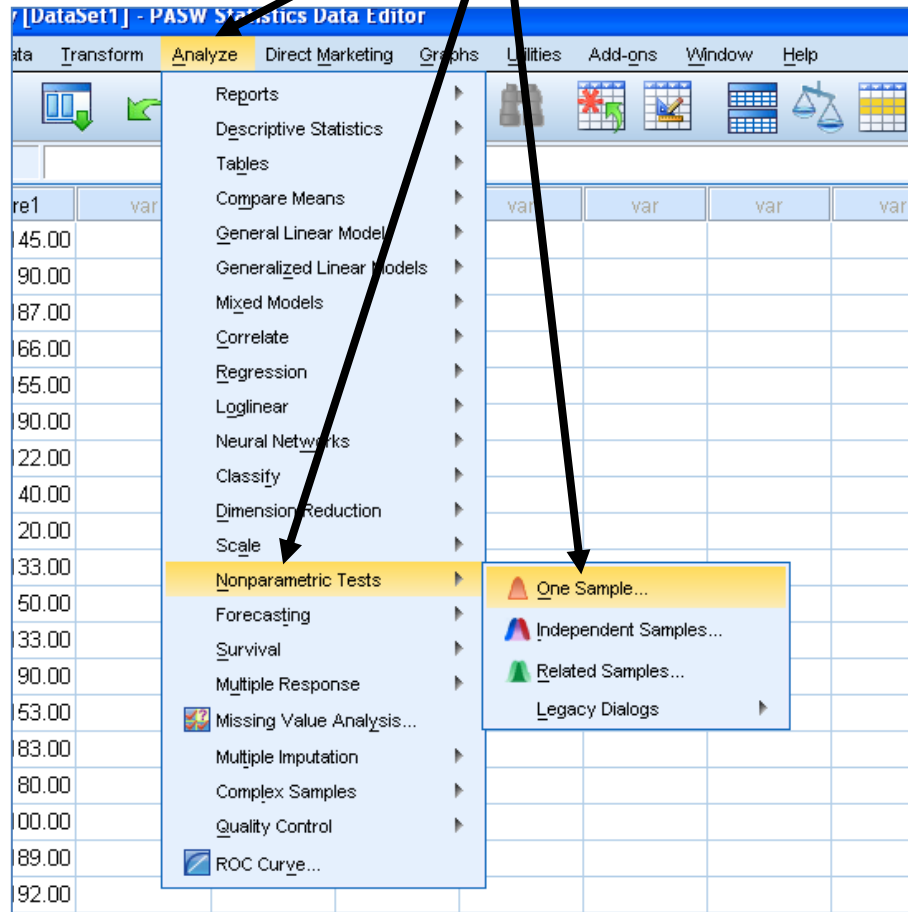
N	Valid	30
	Missing	0
Mean		129.9667
Std. Deviation		54.36054
Skewness		-.669
Std. Error of Skewness		.427
Kurtosis		-.437
Std. Error of Kurtosis		.833
Minimum		10.00
Maximum		205.00

Histogram

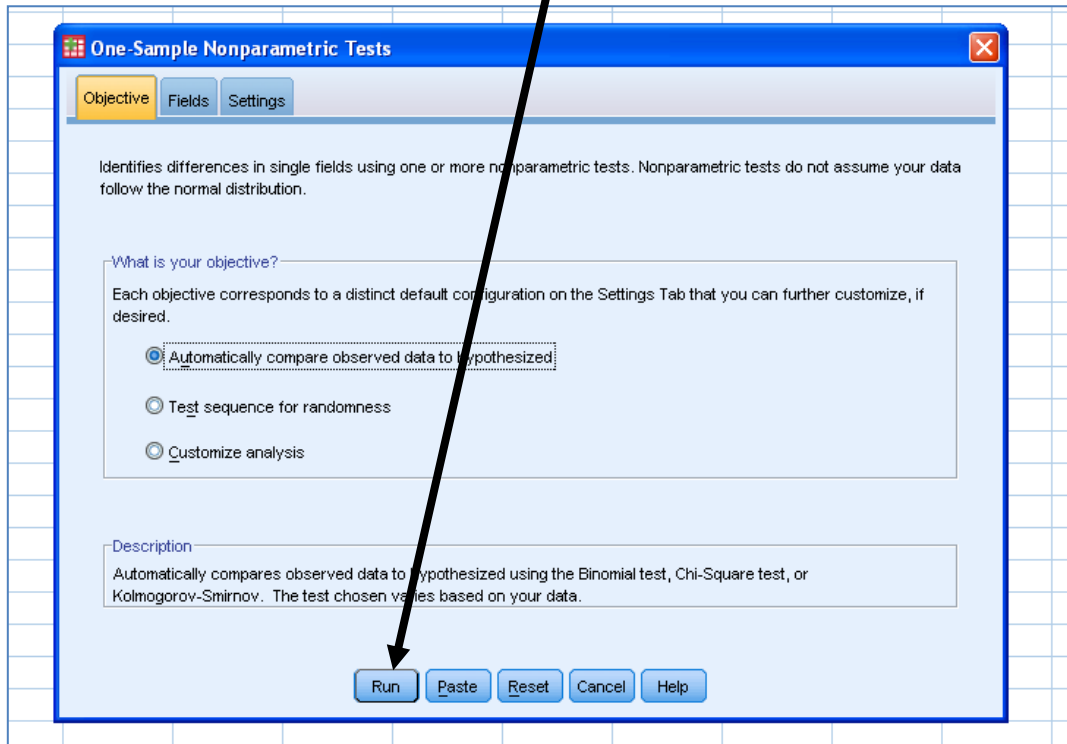


Running the K-S test

Click on 'Analyze', Nonparametric Tests' and 'One Sample'



Click on 'Run'



The following output will appear.

▶ Nonparametric Tests

```
[DataSet1] N:\d. Other work\b. Writing\3. Methods Bk\Normal data.sav
```

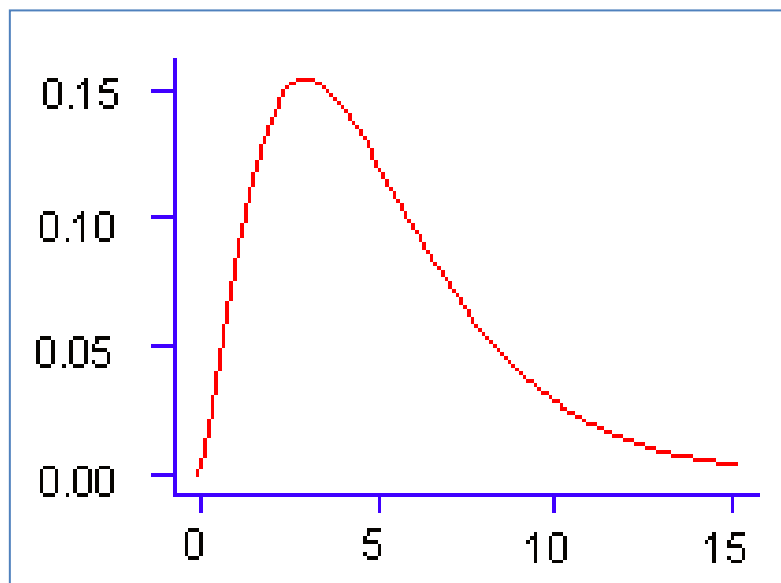
Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of score1 is normal with mean 129.967 and standard deviation 54.361.	One-Sample Kolmogorov-Smirnov Test	.761	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

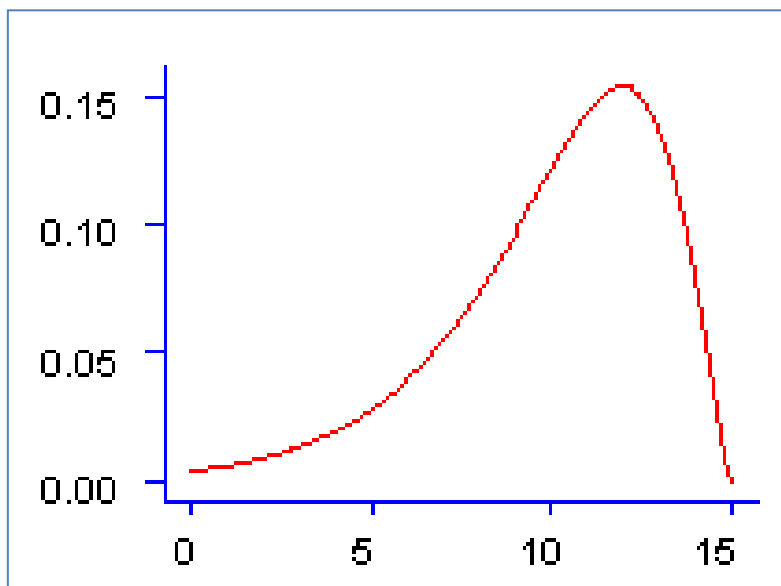
Transforming your data

Positive Skew



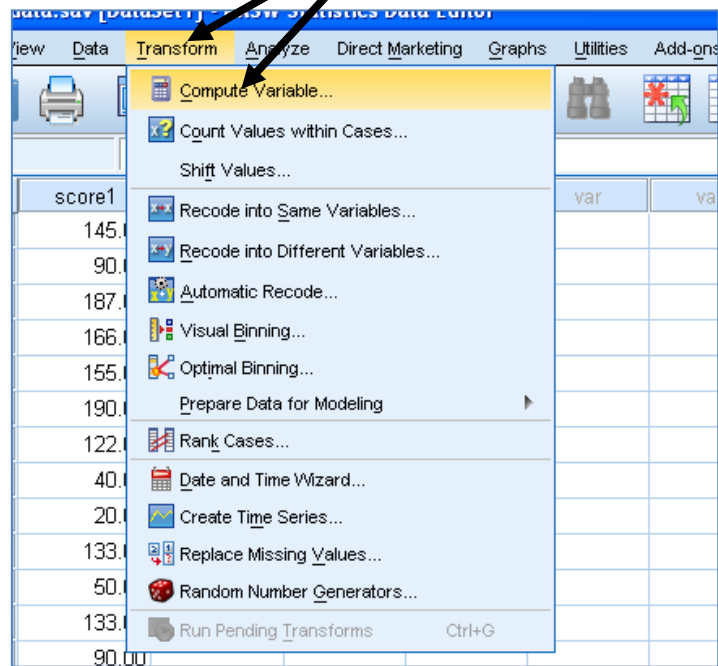
New Variable = SQRT (old variable)

Negative Skew



New variable = SQRT (K – old variable) where K = largest possible value + 1

Click on 'Transform' and 'Compute Variable'



Put the name of your new variable here, and the formula you will use here. Click on OK. Your new transformed data should then appear as a new variable.

