Performing Normality in PASW (SPSS)

When do we do normality test?

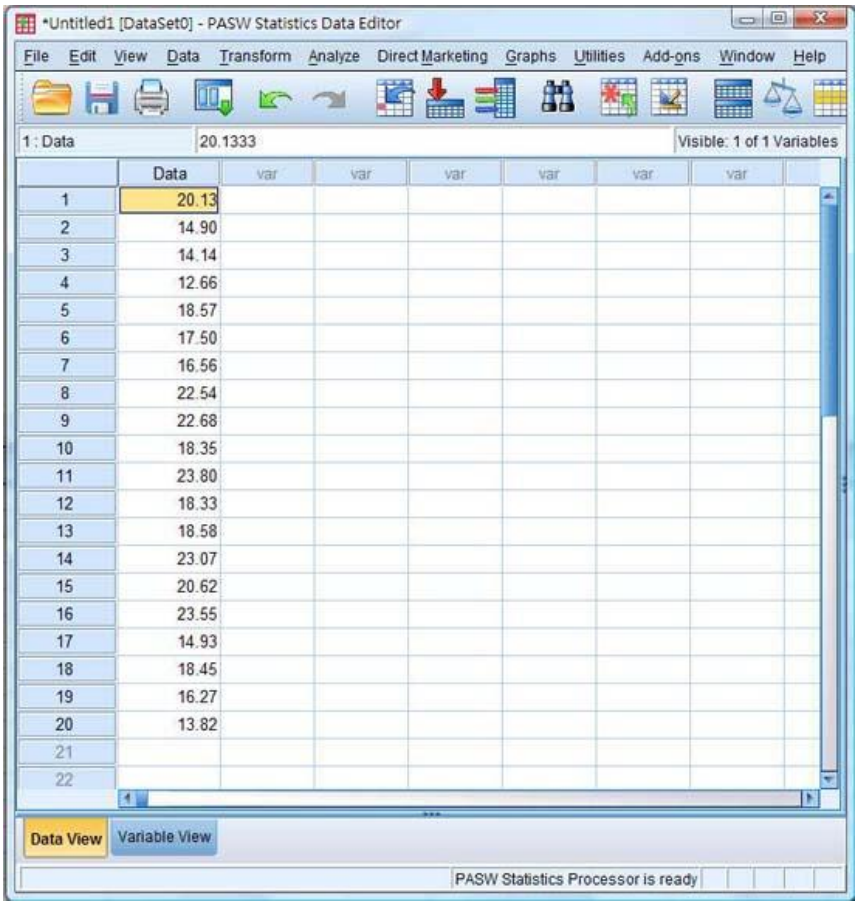
A lot of statistical tests (e.g. t-test) require that our data are normally distributed and therefore we should always check if this assumption is violated.

Example Scenario

Given a set of data, we would like to check if its distribution is normal.

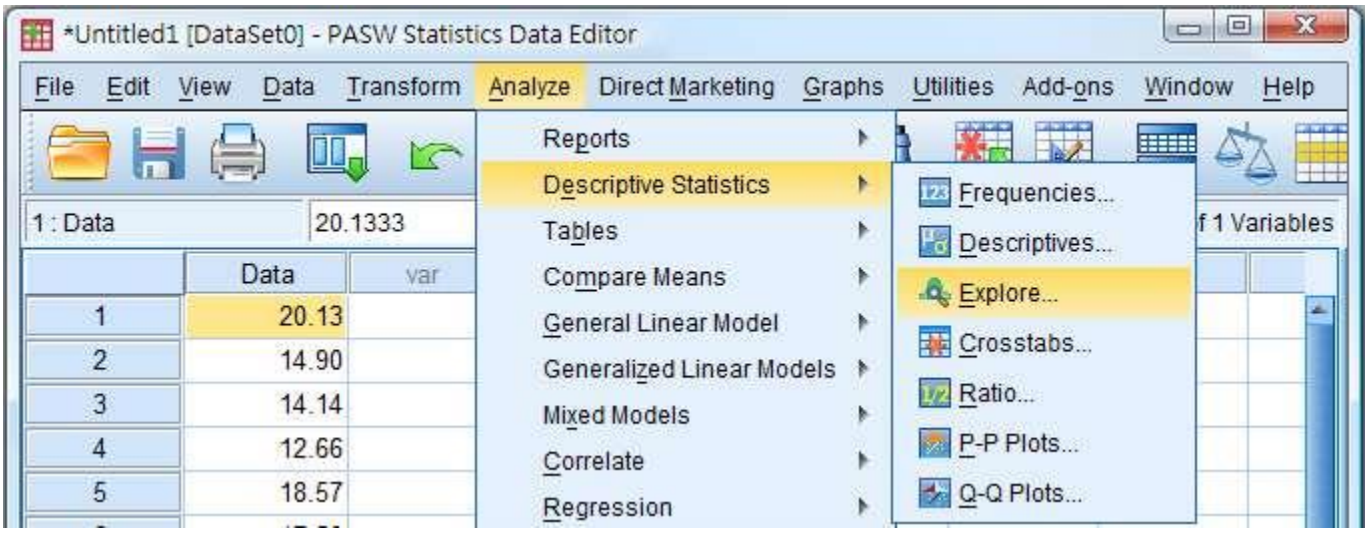
In this example, the null hypothesis is that the data is normally distributed and the alternative hypothesis is that the data is not normally distributed. The dataset can be obtained here.

The data to be tested in stored in the first column.

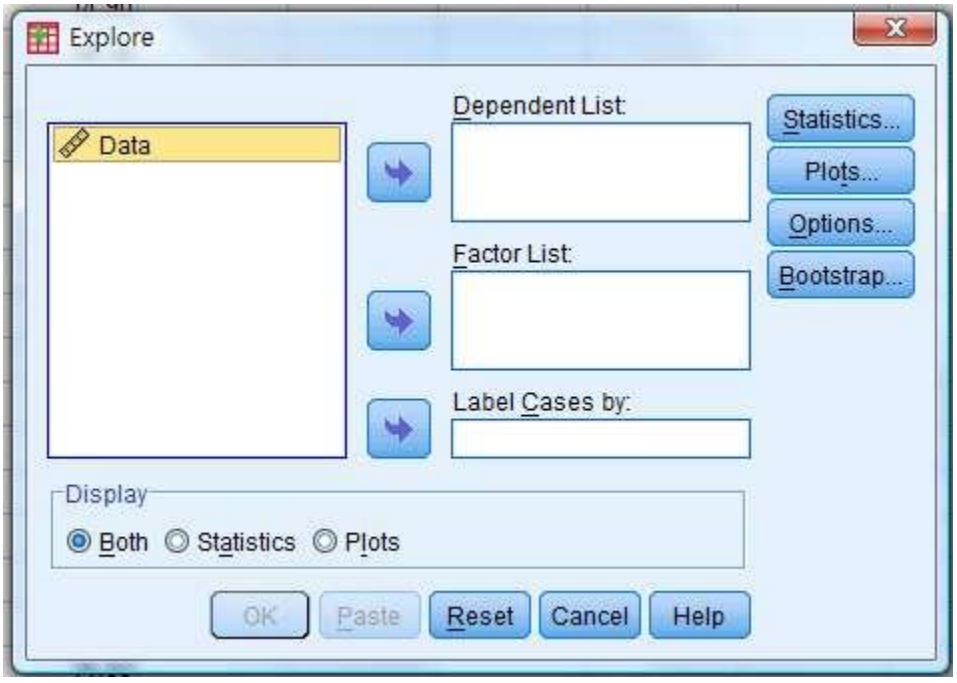


Step 1

Select "Analyze -> Descriptive Statistics -> Explore".

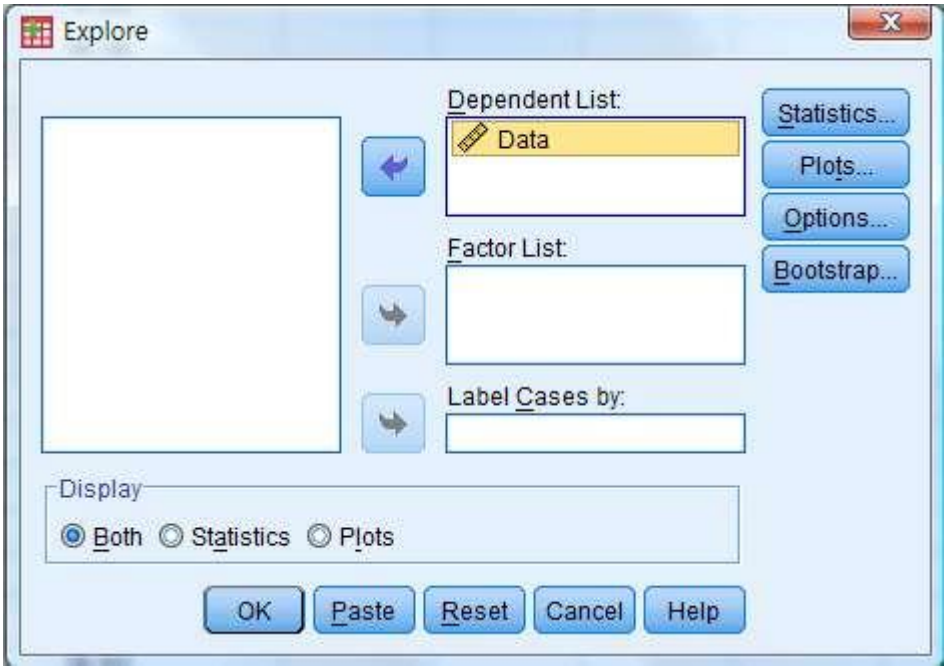


A new window pops out.

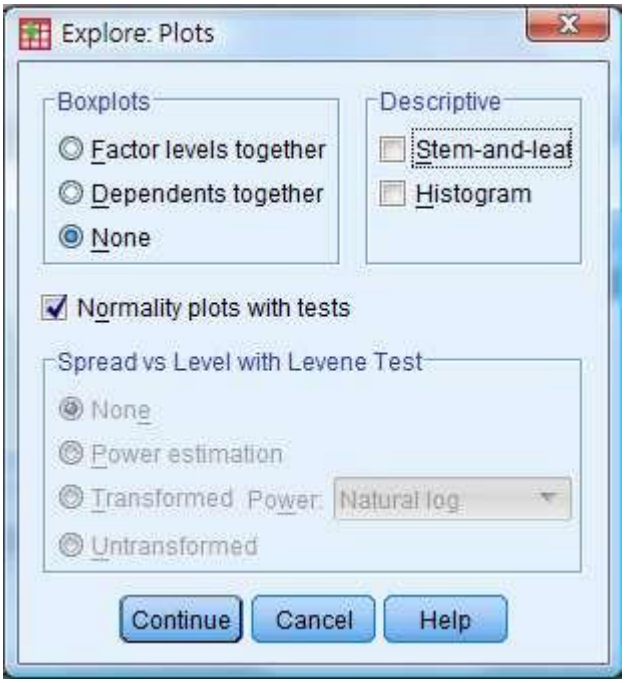


Step 2

From the list on the left, select the variable "Data" to the "Dependent List".

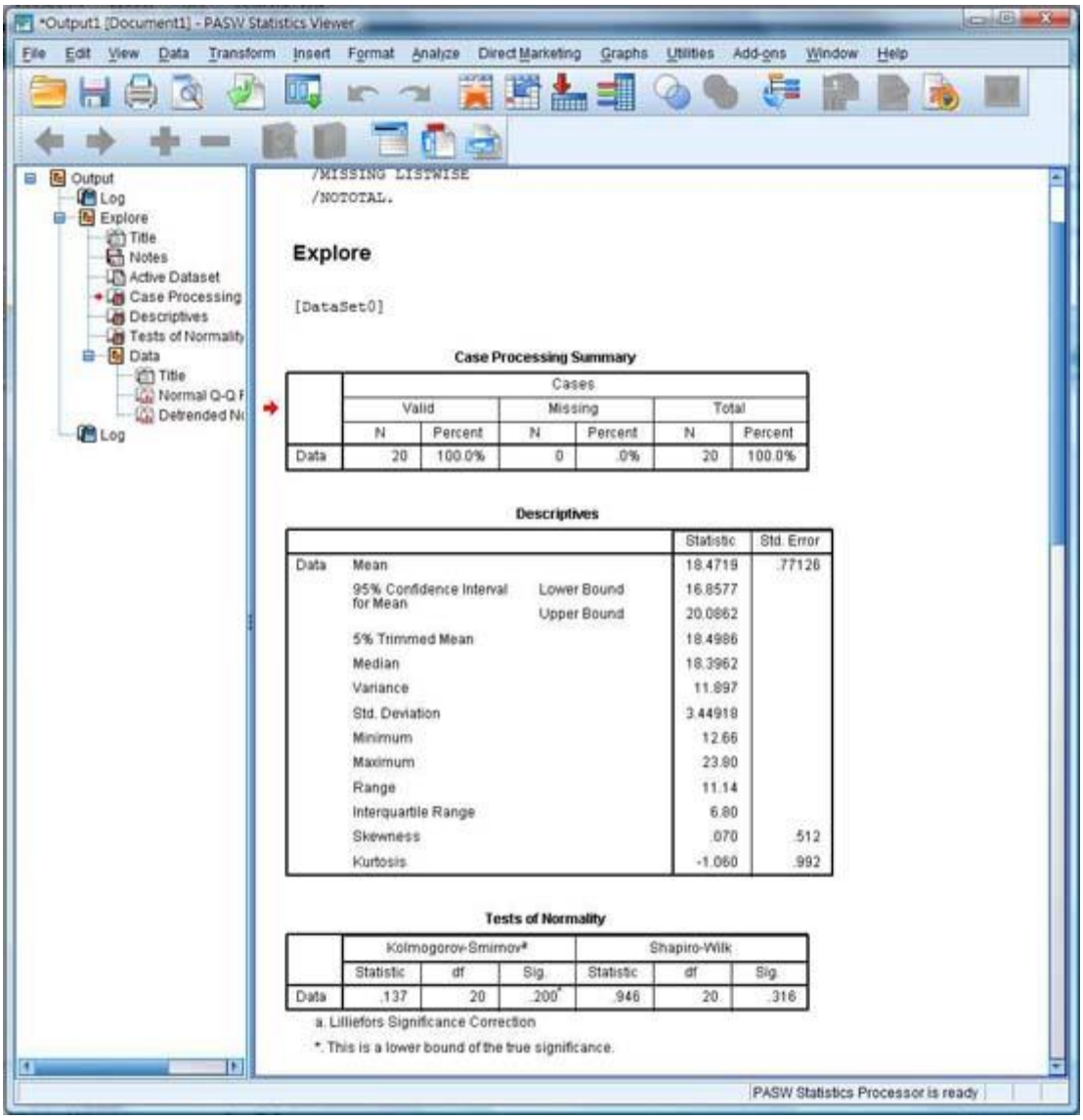


Click "Plots" on the right. A new window pops out. Check "None" for boxplot, uncheck everything for descriptive and make sure the box "Normality plots with tests" is checked.



Step 3

The results now pop out in the "Output" window.



Step 4

We can now interpret the result.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Data	20	100.0%	0	.0%	20	100.0%

Descriptives

			Statistic	Std. Error
Data	Mean		18.4719	.77126
	95% Confidence Interval for Mean	Lower Bound	16.8577	
		Upper Bound	20.0862	
	5% Trimmed Mean		18.4986	
	Median		18.3962	
	Variance		11.897	
	Std. Deviation		3.44918	
	Minimum		12.66	
	Maximum		23.80	
	Range		11.14	
	Interquartile Range		6.80	
	Skewness		.070	.512
	Kurtosis		-1.060	.992

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Data	.137	20	.200*	.946	20	.316

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

The test statistics are shown in the third table. Here two tests for normality are run. For dataset small than 2000 elements, we use the Shapiro-Wilk test, otherwise, the Kolmogorov-Smirnov test is used. In our case, since we have only 20 elements, the Shapiro-Wilk test is used. From A, the p-value is 0.316. We can reject the alternative hypothesis and conclude that the data comes from a normal distribution.