

Symmetry and Skewness Distribution

Many parametric methods required that variable is approximately symmetrically distributed, such as

- Analysis of variance or ANOVA test
- Linear regression
- F-test and T-test

How to check whether data are symmetrically distributed using SPSS

The following **numerical** and **visual outputs** must be investigated:

- Skewness and Kurtosis z-values

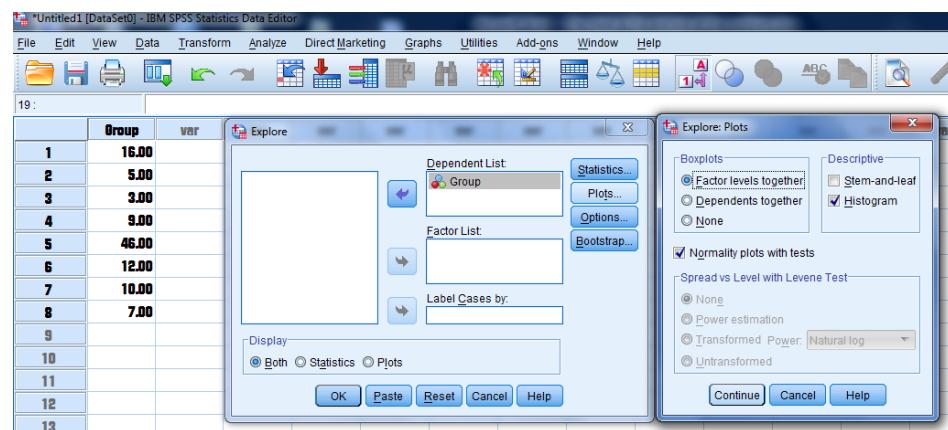
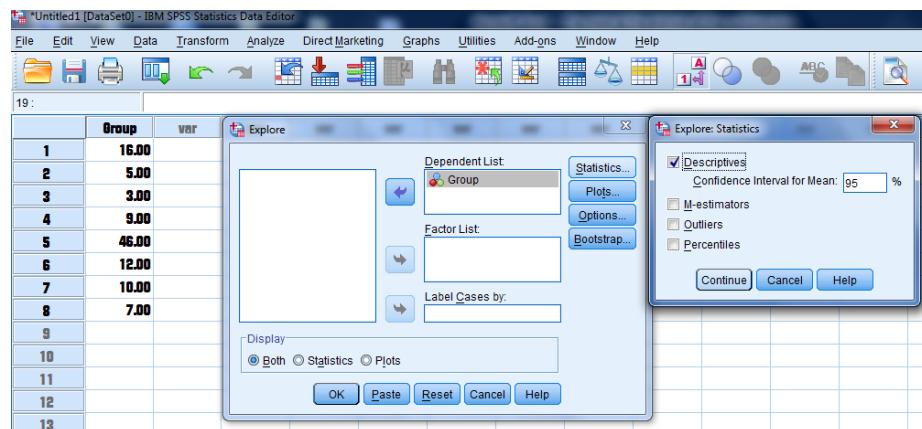
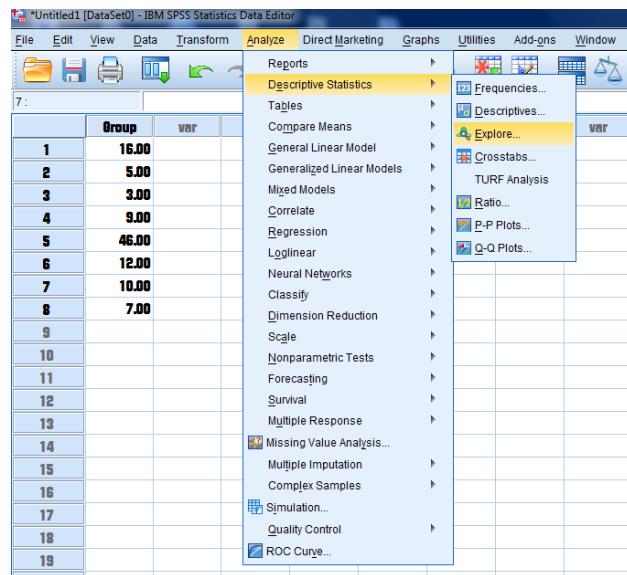
Should be somewhere in the span of **-1.96 to +1.96**

- The Shapiro-wilk test p-value

Should be **above 0.05**

- Histograms, normal Q-Q Plots, and box plots

Should visually that our data are **approximately symmetrically distributed**



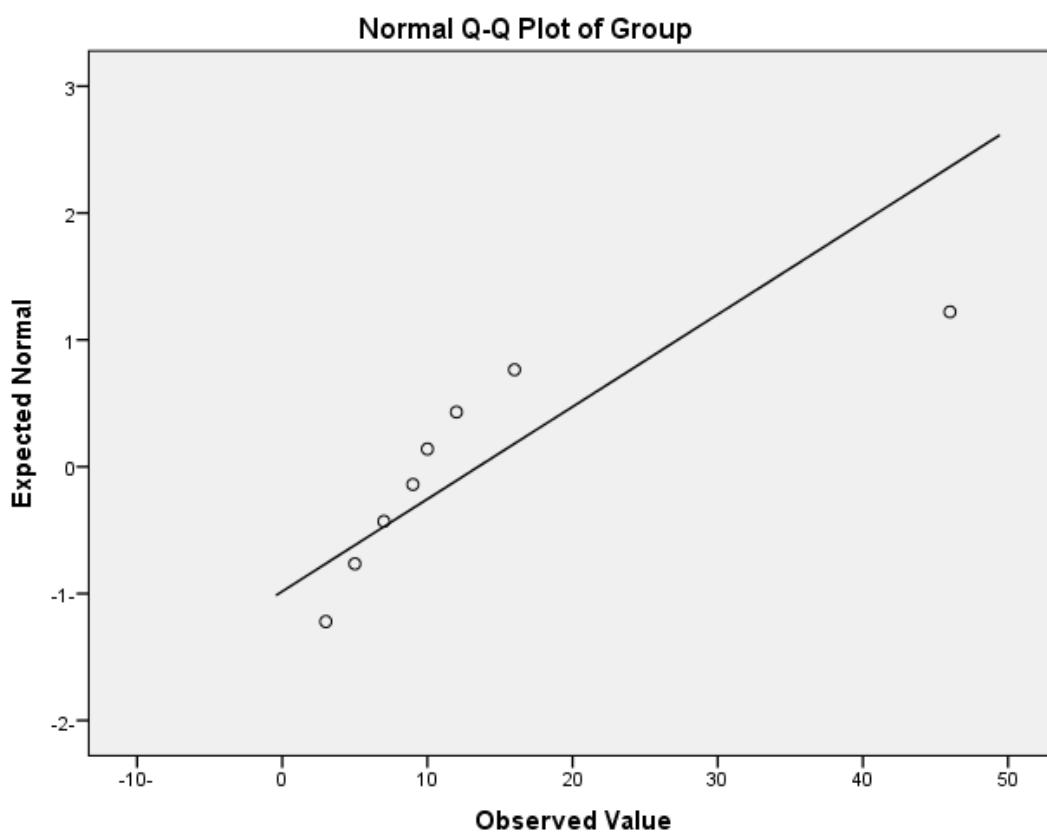
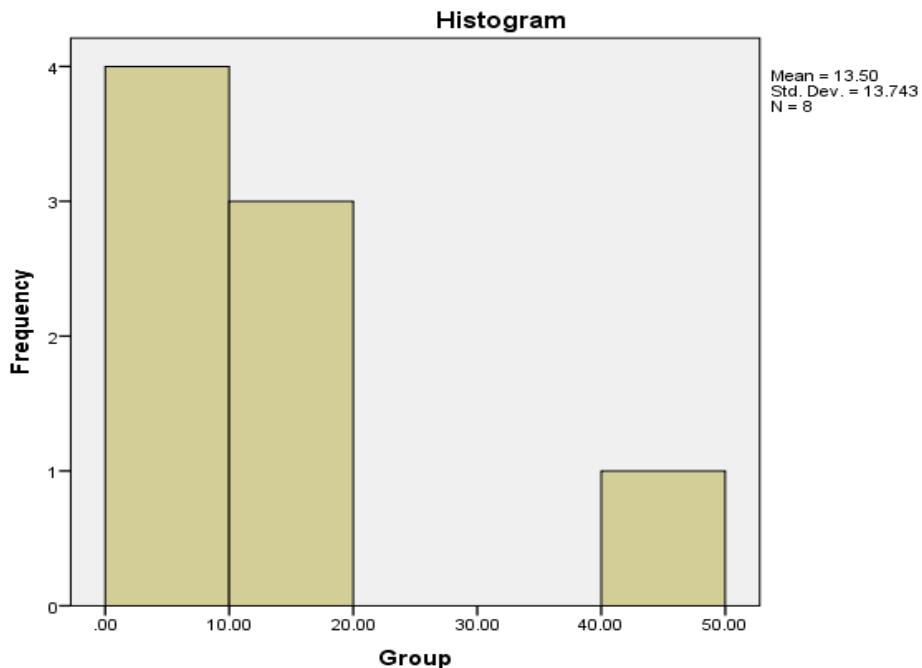
Output

Descriptives			
		Statistic	Std. Error
Group	Mean	13.5000	4.85872
	95% Confidence Interval for Mean	Lower Bound 2.0110	
		Upper Bound 24.9890	
	5% Trimmed Mean	12.2778	
	Median	9.5000	
	Variance	188.857	
	Std. Deviation	13.74253	
	Minimum	3.00	
	Maximum	46.00	
	Range	43.00	
	Interquartile Range	9.50	
	Skewness	2.360	.752
	Kurtosis	6.015	1.481
			2.360/0.752=3.138
			6.015/1.481=4.06

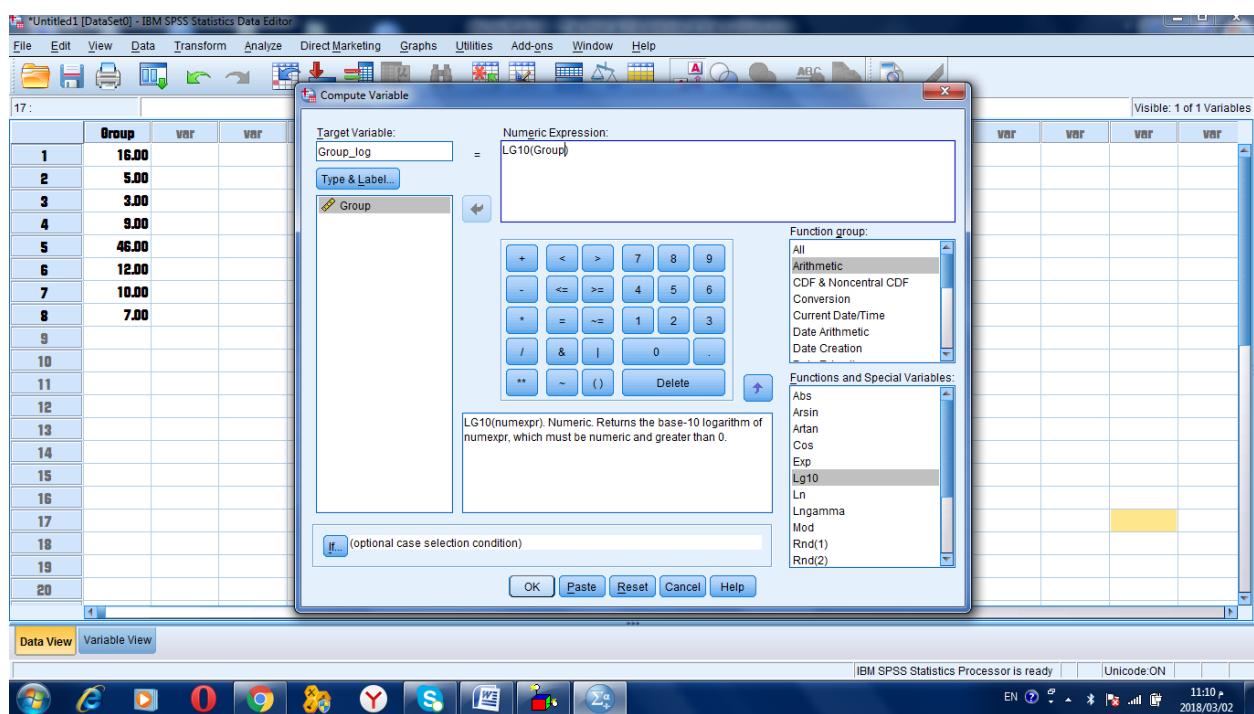
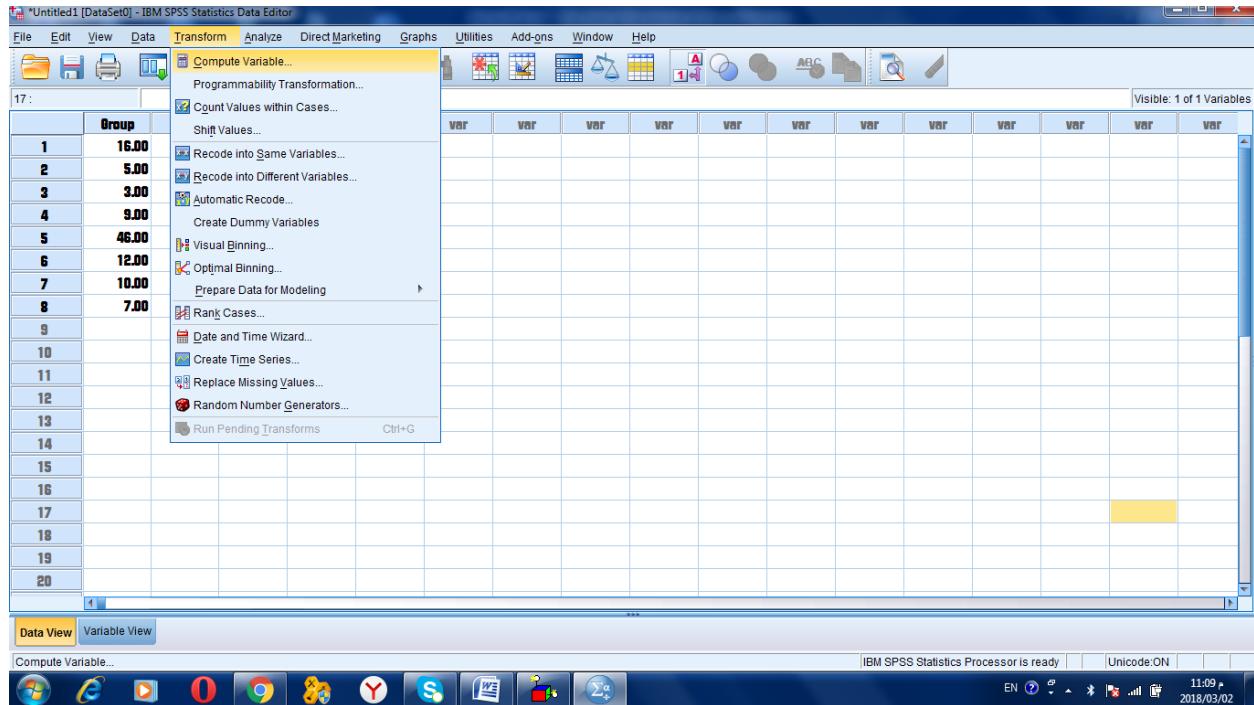
Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Group	.303	8	.030	.700	8	.002

a. Lilliefors Significance Correction

0.002 < 0.05



How to transform Non-Normalized data into normalized data in SPSS



*Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

	Group	Group_Log	var	var	
1	16.00	1.20			
2	5.00	.70			
3	3.00	.48			
4	9.00	.95			
5	46.00	1.66			
6	12.00	1.08			
7	10.00	1.00			
8	7.00	.85			
9					
10					

*Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

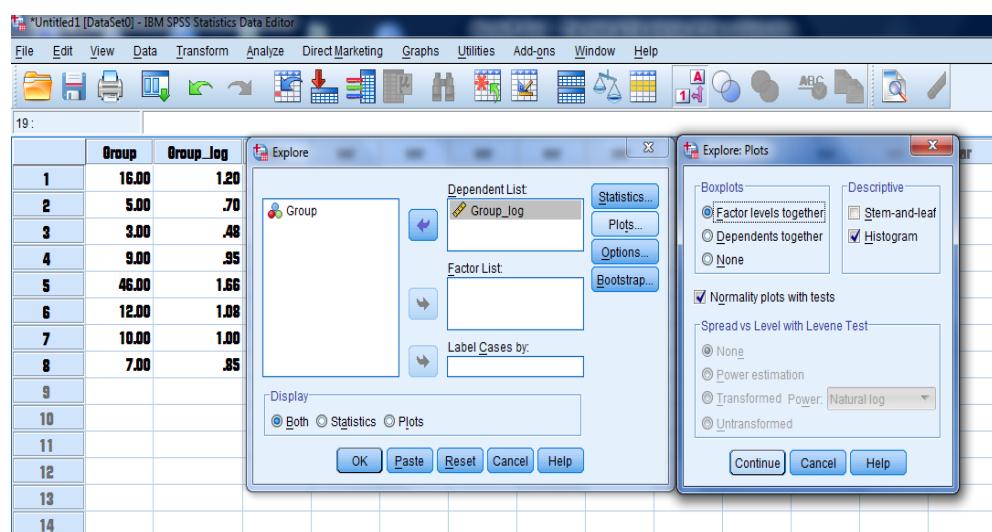
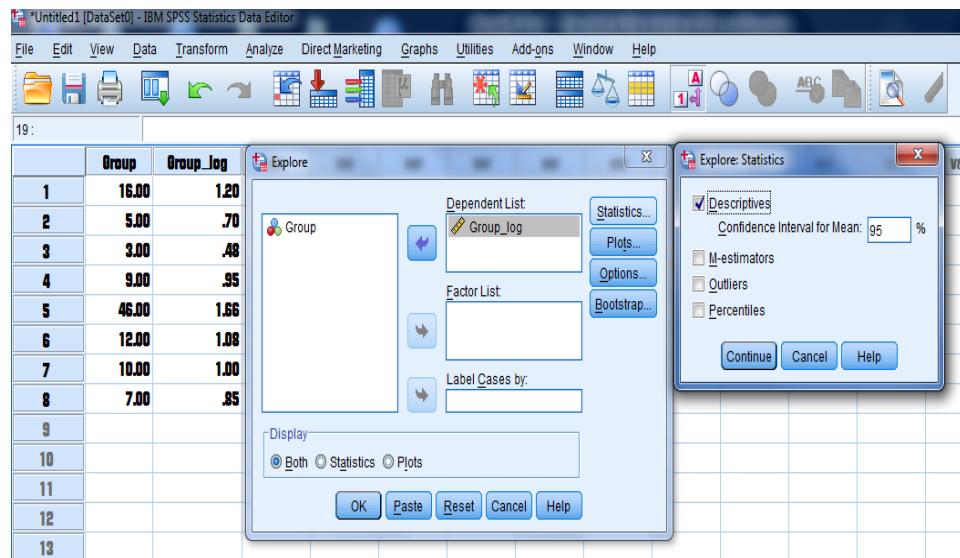
Visible: 2 of 2 Variables

Data View Variable View

Explore... IBM SPSS Statistics Processor is ready Unicode:ON 11:14 p 2018/03/02

Explore... 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

- Reports
- Descriptive Statistics
- Tables
- Compare Means
- General Linear Model
- Generalized Linear Models
- Mixed Models
- Correlate
- Regression
- Loglinear
- Neural Networks
- Classify
- Dimension Reduction
- Scale
- Nonparametric Tests
- Forecasting
- Survival
- Multiple Response
- Missing Value Analysis...
- Multiple Imputation
- Complex Samples
- Simulation...
- Quality Control
- ROC Curve...



Output

Descriptives

		Statistic	Std. Error
Group_log	Mean	.9902	.12511
	95% Confidence Interval for Mean	.6944	
	Lower Bound		
	Upper Bound	1.2860	
	5% Trimmed Mean	.9813	
	Median	.9771	
	Variance	.125	
	Std. Deviation	.35386	
	Minimum	.48	
	Maximum	1.66	
	Range	1.19	
	Interquartile Range	.44	
	Skewness	.653	.752
	Kurtosis	1.304	1.481
		0.653/0.752=0.87	
		1.304/1.481=0.88	

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Group_log	.151	8	.200 [*]	.967	8	.875

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

a. Lilliefors Significance Correction

0.875>0.05

