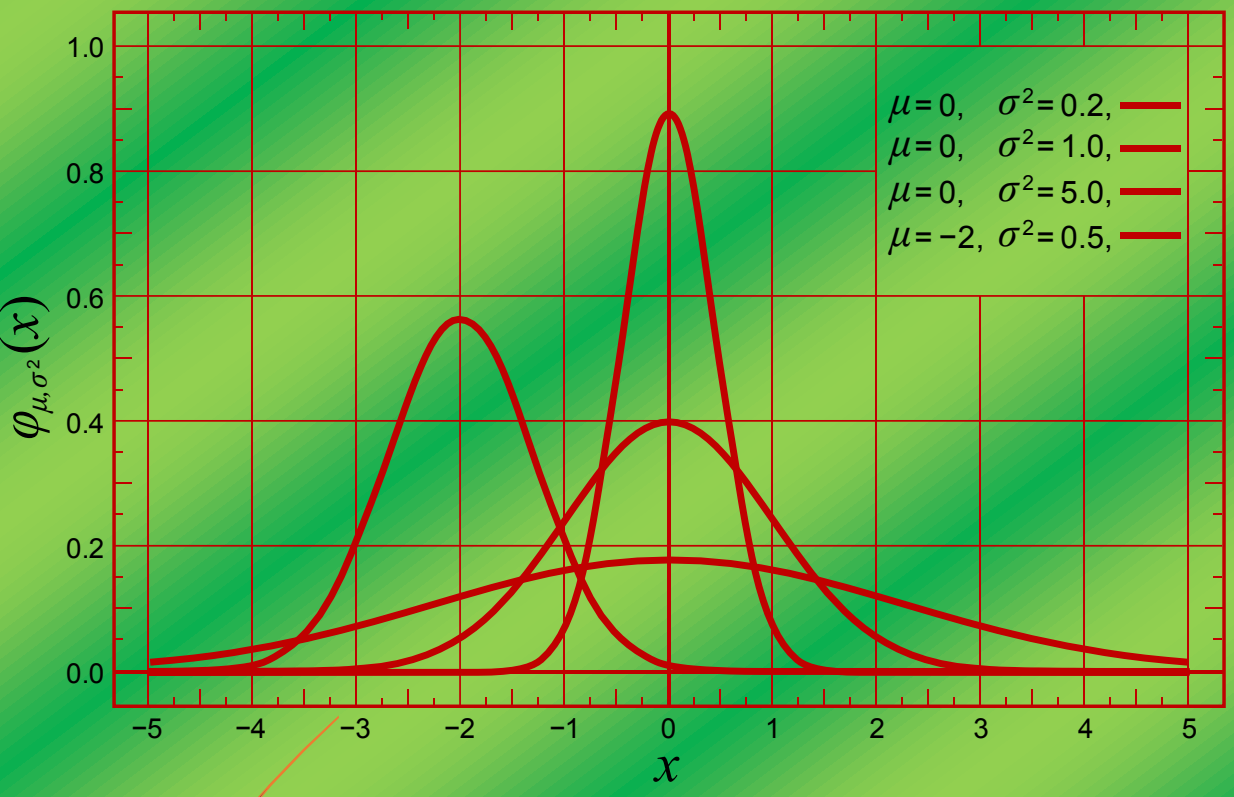
SPSS Practical Manual on Normality Checking of Data



D. S. Dhakre, D. Bhattacharya and Bhola Nath

Institute of Agriculture, Visva-Bharati, Sriniketan
West Bengal -731 236, India

SPSS Practical Manual on Normality Checking of Data

D. S. Dhakre, D. Bhattacharya and Bhola Nath
Institute of Agriculture, Visva-Bharati, Sriniketan
West Bengal -731 236, India

Example:

The following table gives the yields in pound per plot, of seven varieties of a crop after being applied to each of 4 plots, tested in a Completely Randomized Design. Test data is normal or not

treatment	R1	R2	R3
t1	17.8	24.4	18.1
t2	20.2	22.7	23
t3	15.7	16.4	18.8
t4	19.7	15.3	19.8
t5	19.4	20.3	23.7
t6	17.7	18.6	21.1
t7	21.7	21.6	17.4

There are three methods to check Normality of Data

1) the Z value of Skewness and Kurtosis should be -1.96 to +1.96
2) Shapiro Wilk's test significant value should be above 0.05
3) with the help of Histogram, Q-Q plot, Box plot

Arrangements of data for SPSS

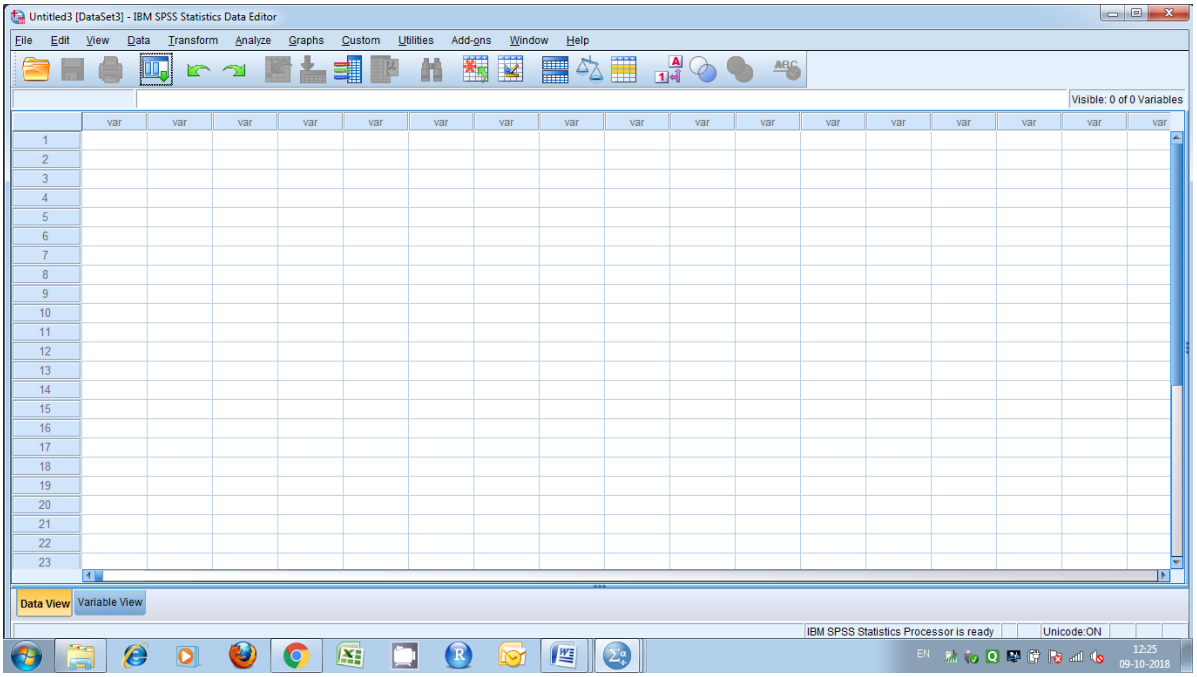
Rep	Trt	yld
1	1	17.8
1	2	20.2
1	3	15.7
1	4	19.7
1	5	19.4
1	6	17.7
1	7	21.7
2	1	24.4

2	2	22.7
2	3	16.4
2	4	15.3
2	5	20.3
2	6	18.6
2	7	21.6
3	1	18.1
3	2	23
3	3	18.8
3	4	19.8
3	5	23.7
3	6	21.1
3	7	17.4

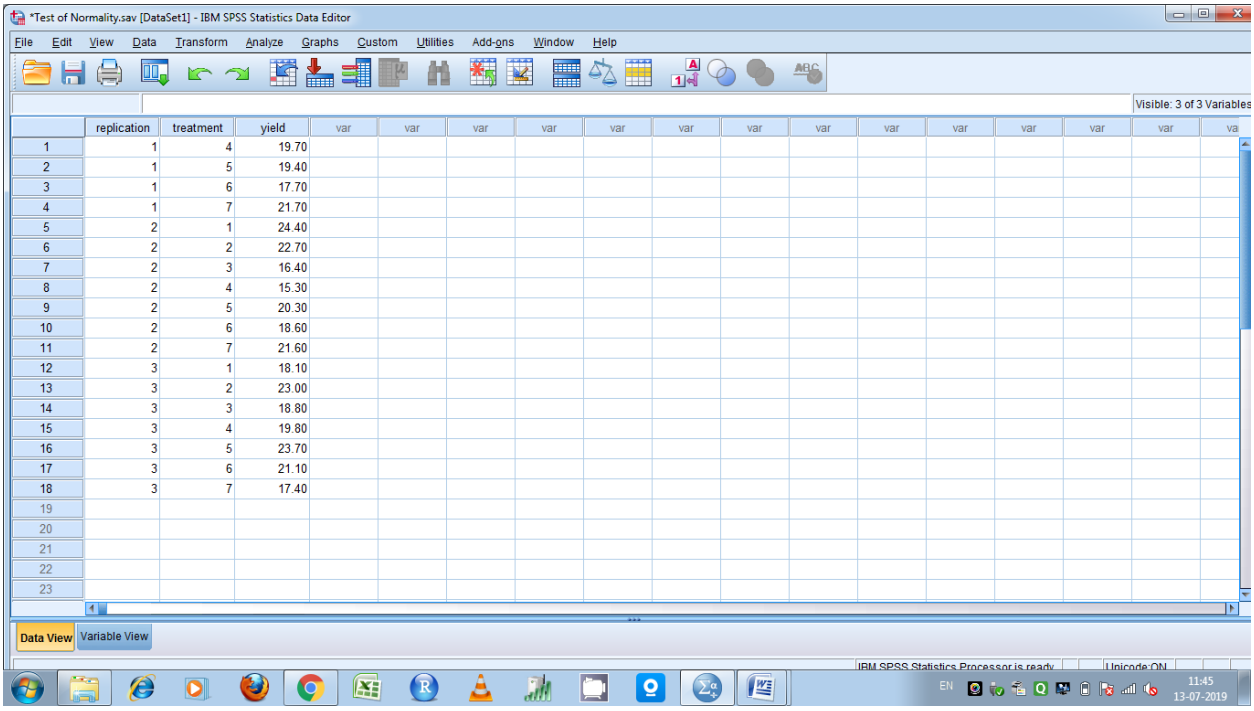
SPSS commands for Analysis

The input data file can be created as shown below:

Step 1: File → New → Data →

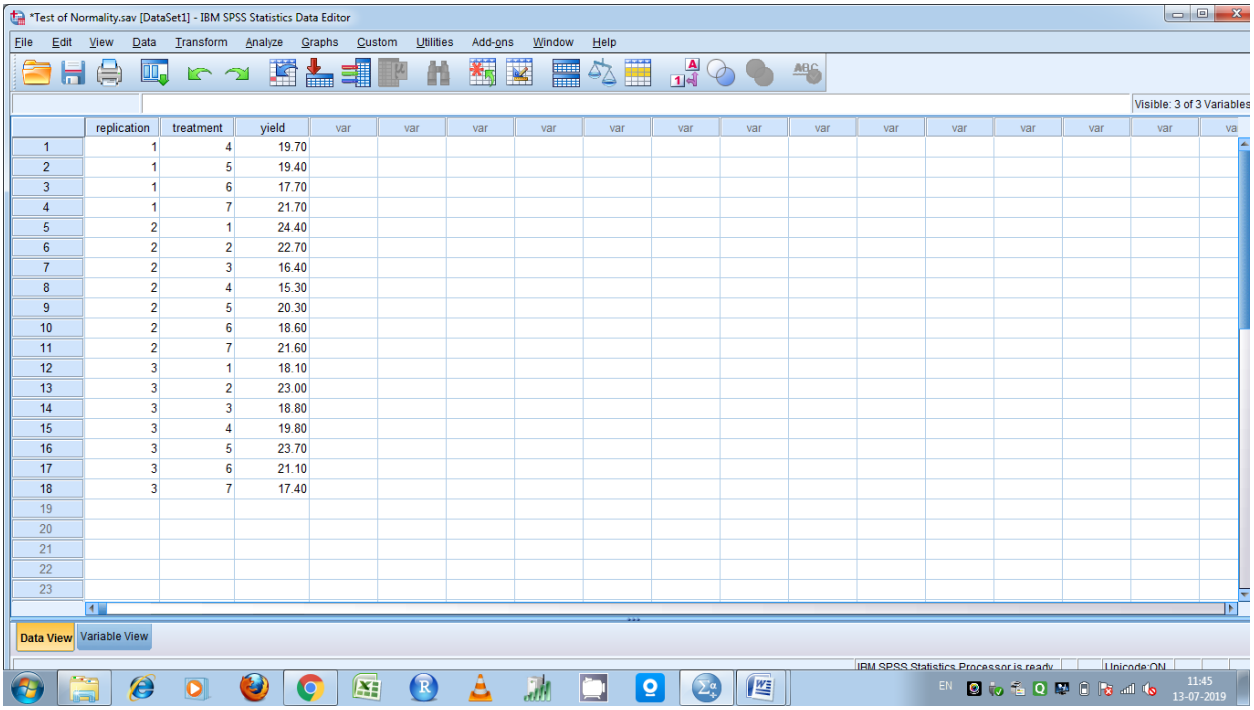


Step 2: Variable view → Name (replication, treatment, yield) →



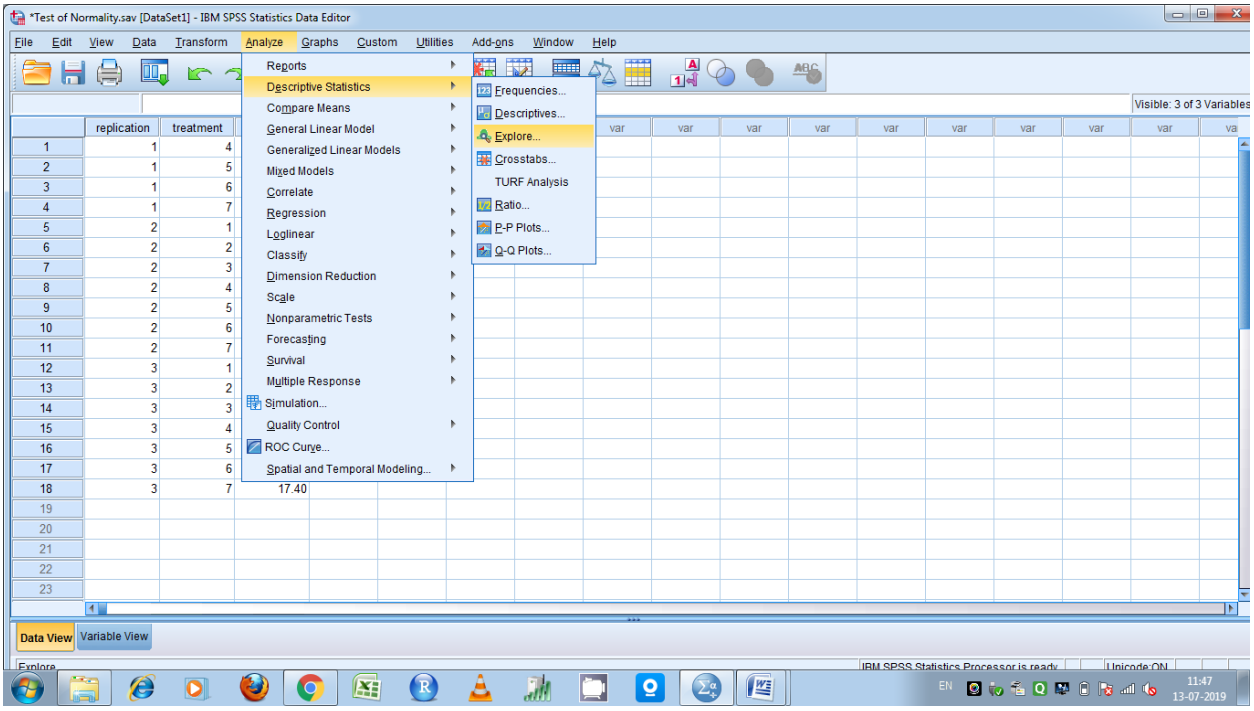
	replication	treatment	yield
1	1	4	19.70
2	1	5	19.40
3	1	6	17.70
4	1	7	21.70
5	2	1	24.40
6	2	2	22.70
7	2	3	16.40
8	2	4	15.30
9	2	5	20.30
10	2	6	18.60
11	2	7	21.60
12	3	1	18.10
13	3	2	23.00
14	3	3	18.80
15	3	4	19.80
16	3	5	23.70
17	3	6	21.10
18	3	7	17.40

Step 3: Data view → Enter data → File → Save (with any file name)

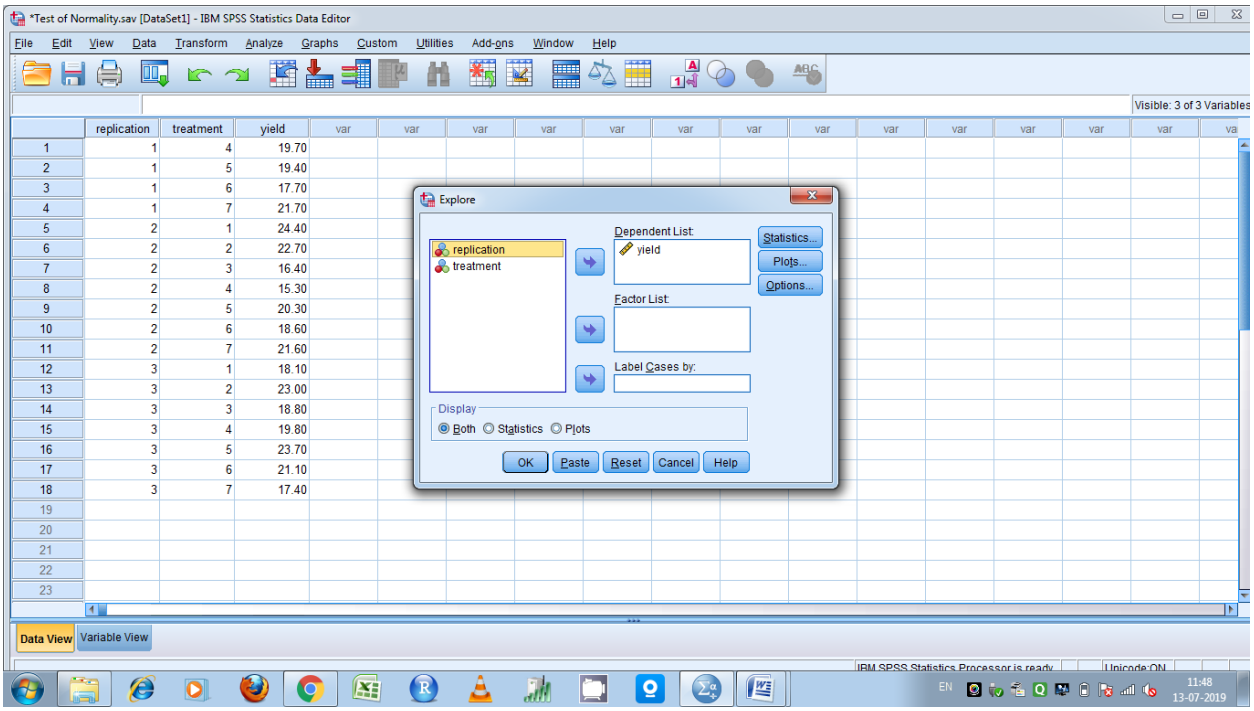


	replication	treatment	yield
1	1	4	19.70
2	1	5	19.40
3	1	6	17.70
4	1	7	21.70
5	2	1	24.40
6	2	2	22.70
7	2	3	16.40
8	2	4	15.30
9	2	5	20.30
10	2	6	18.60
11	2	7	21.60
12	3	1	18.10
13	3	2	23.00
14	3	3	18.80
15	3	4	19.80
16	3	5	23.70
17	3	6	21.10
18	3	7	17.40

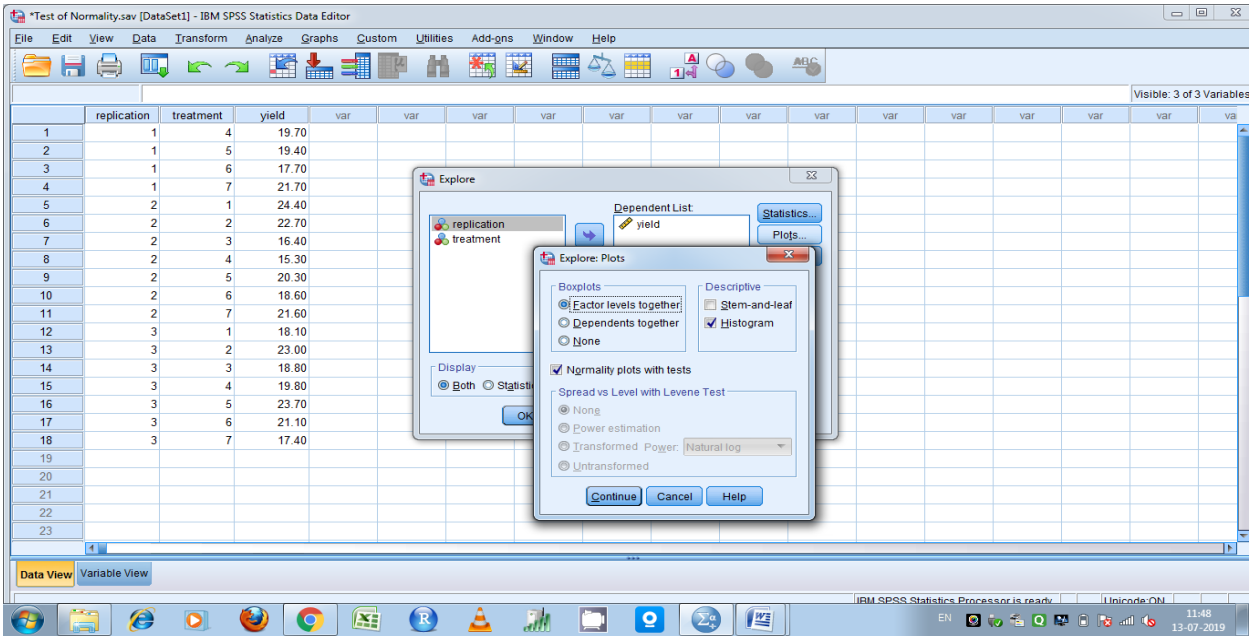
Step 4: Analyze → Descriptive Statistics → Explore →



Step 4: Analyze → Descriptive Statistics → Explore → Select dependent variable



Step 5: Plots..→Normality plots with tests → OK



Output:

Descriptives

			Statistic	Std. Error
yld	Mean		19.9833	.60143
	95% Confidence Interval for Mean	Lower Bound	18.7144	
		Upper Bound	21.2522	
	5% Trimmed Mean		19.9981	
	Median		19.7500	
	Variance		6.511	
	Std. Deviation		2.55164	
	Minimum		15.30	
	Maximum		24.40	
	Range		9.10	
	Interquartile Range		3.95	
	Skewness		.026	.536
	Kurtosis		-.722	1.038

Z value of skewness = 0.026/0.536= 0.049
Z value of Kurtosis = -0.722/1.038= -0.696
Both are under -1.96 to +1.96 so our data is approximately normally distributed.

Tests of Normality

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
yld	.084	18	.200*	.983	18	.978

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

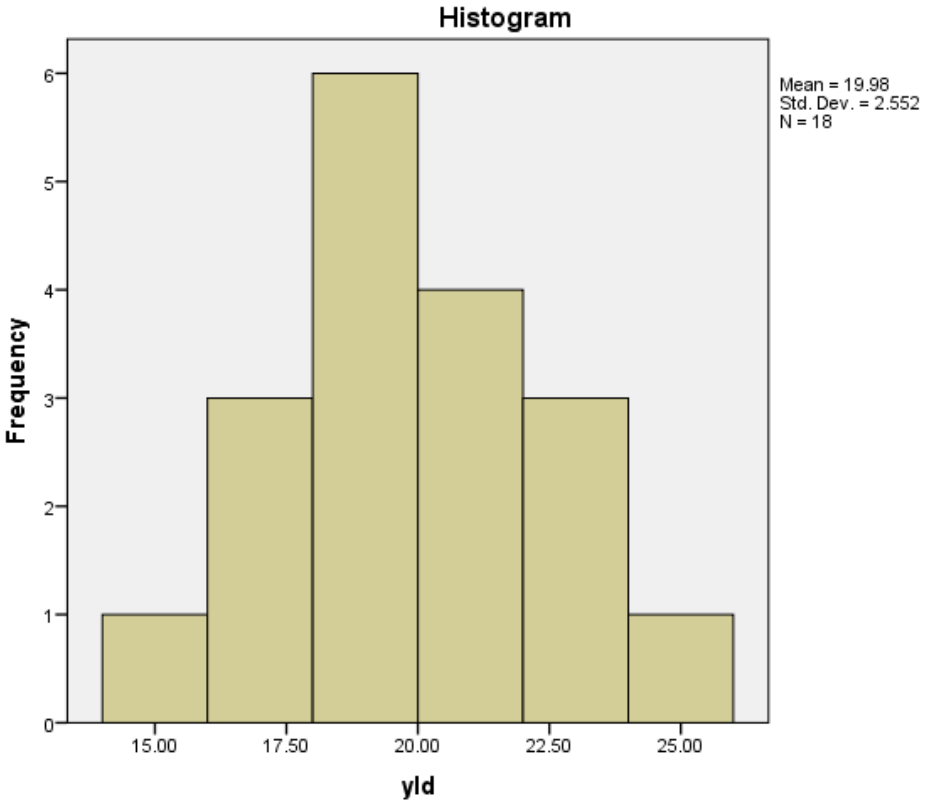
a. Lilliefors Significance Correction

Hypothesis under Shapiro-Wilk's test statistic is defined as:

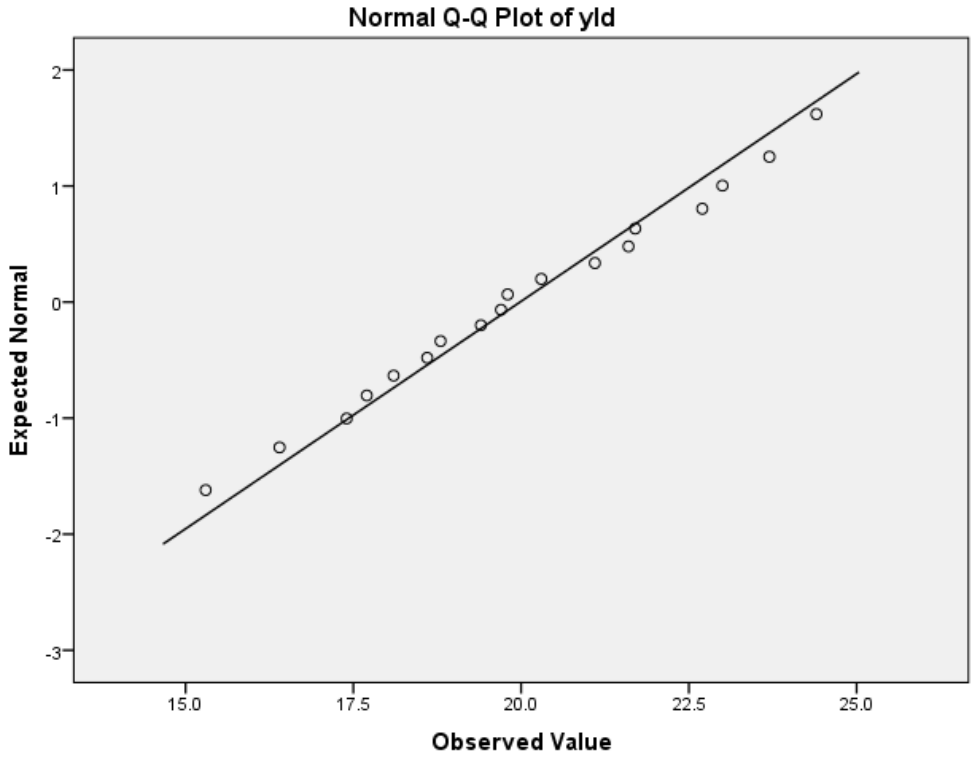
H_o: The data are normally distributed

H_1: The data are not normally distributed

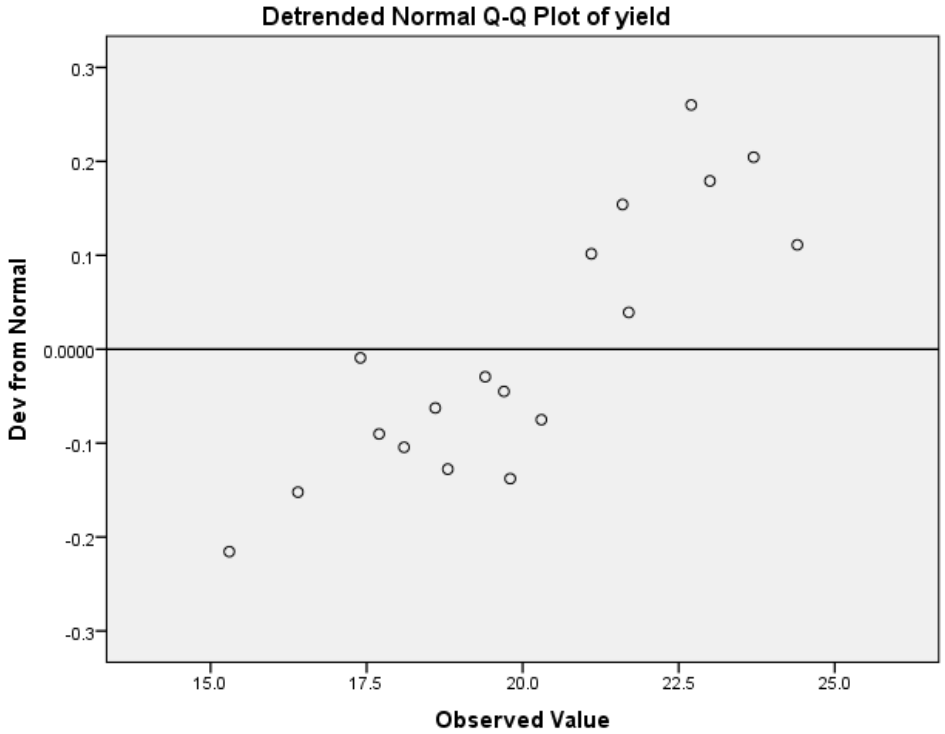
According to Shapiro-Wilk's test result is non-significant it means, we cannot reject H_o it means data are normally distributed

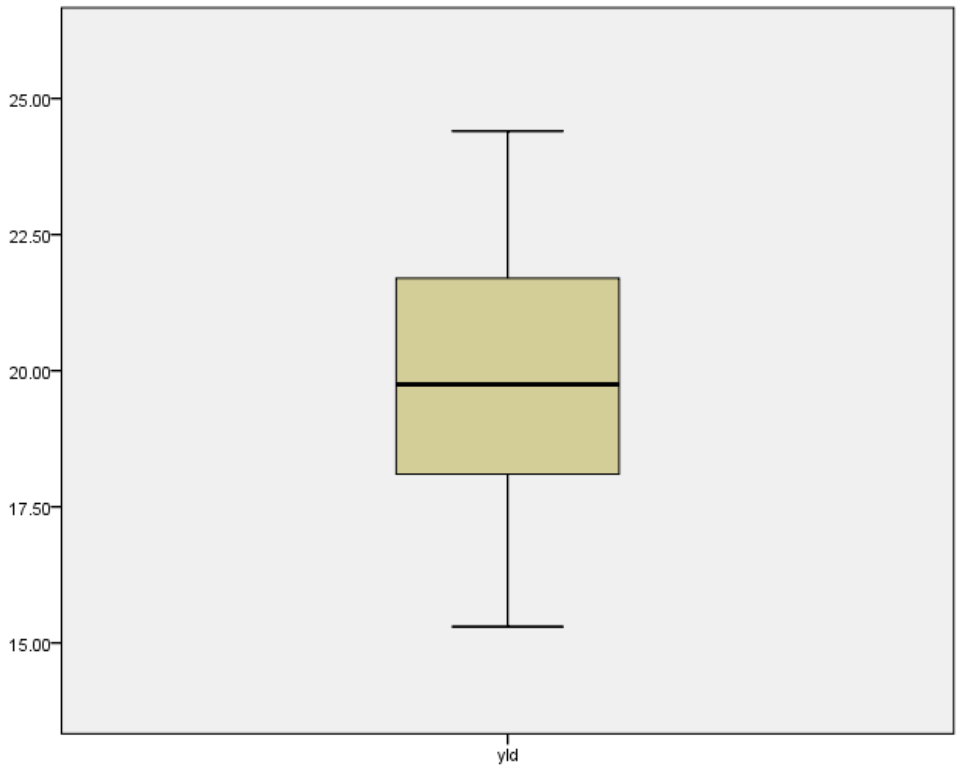


Histogram should have the approximate shape of normal curve.



All points are near about straight line so we can say data approximately normal.





Box plot should be symmetrical as possible

A Shapiro Wilk's test (p>0.05) and a visual inspection of histogram, normal Q-Q plots and box plots showed that the yield was approximately normally distributed.