Assignment 4

Sainivedhitha Arunajatesan

Exercise 1

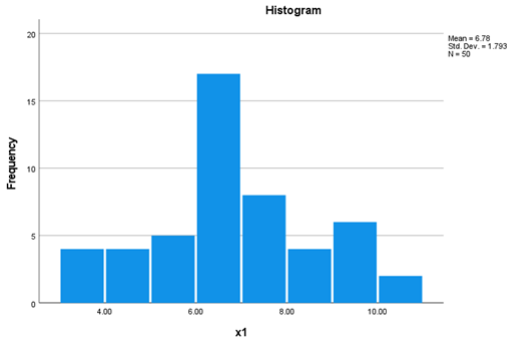
1. **Shapiro-Wilk:**
   The value of α is assumed to be 0.05 and the value of this test is greater than α. Therefore, x1 is normally distributed.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
x1	.110	50	.177	.970	50	.241

a. Lilliefors Significance Correction

Histogram:
The histogram looks more or less like a bell shaped curve which proves it is normally distributed.



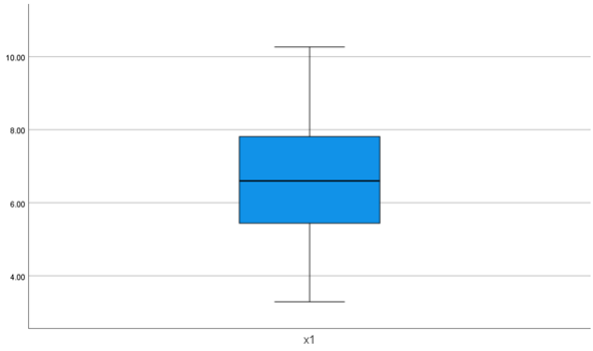
Skewness and Kurtosis:
The skewness of 0.052 is closer to zero which indicates that it is normally distributed.
Generally, a negative kurtosis defines flat peaks. But, here the kurtosis is a very small negative value of -0.454 which indicates a normal distribution with slightly thin tails.

Statistics

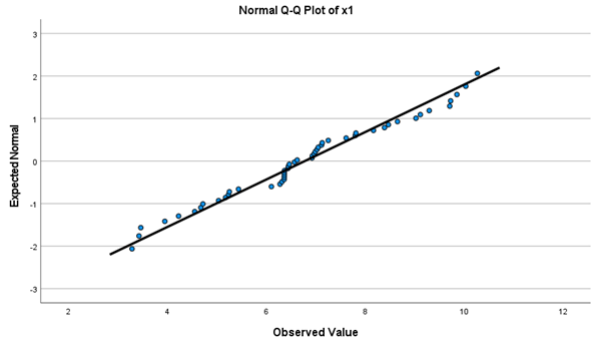
x1

N	Valid	50
	Missing	0
Skewness		.052
Std. Error of Skewness		.337
Kurtosis		-.464
Std. Error of Kurtosis		.662

Box Plot:
The median is in the middle of the box and hence it is normally distributed.



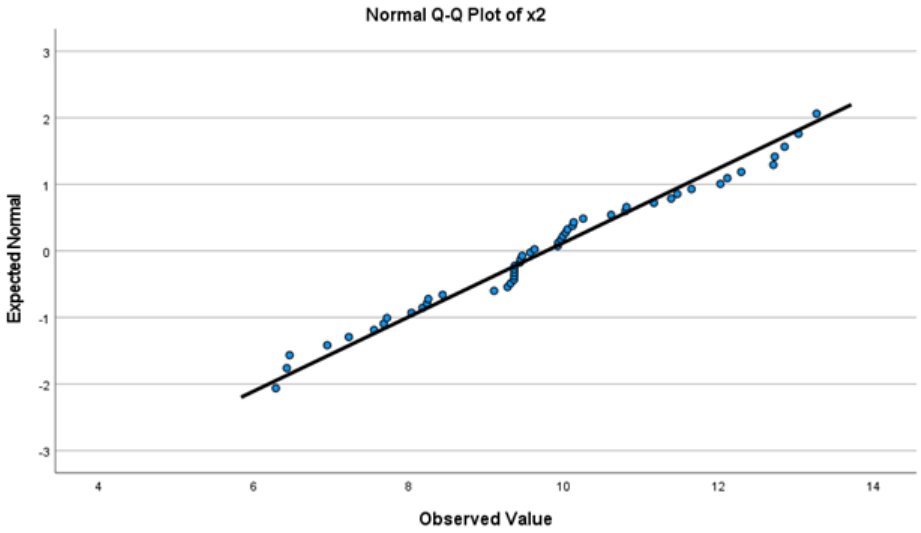
QQ Plot:
The points in the QQ-normal plot lie almost near the straight diagonal line. Thus, the data is normally distributed.



2. The data set x1 was used from Biostatistics 4 simulation.sav file. The data set x2 was calculated using the RV.NORMAL by randomly choosing the mean to be 10 and variance of 4.
For the data set containing 50, 40, 25 and 10 values, the plots were as follows:

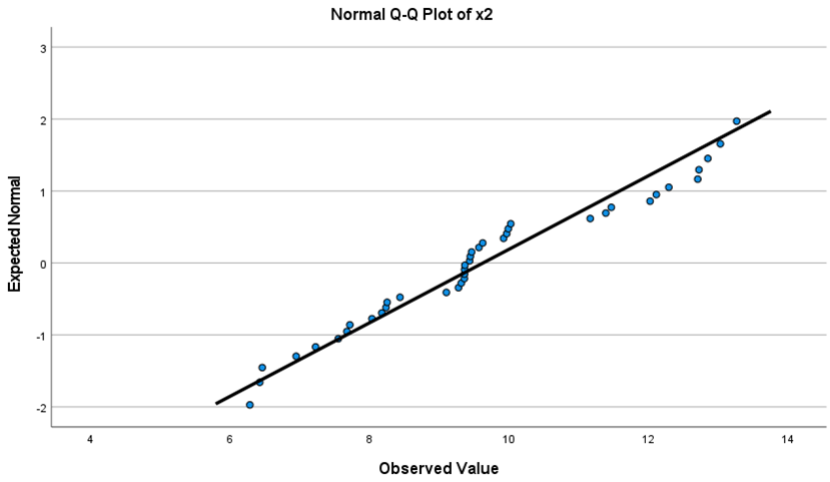
Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
x2	50	100.0%	0	0.0%	50	100.0%



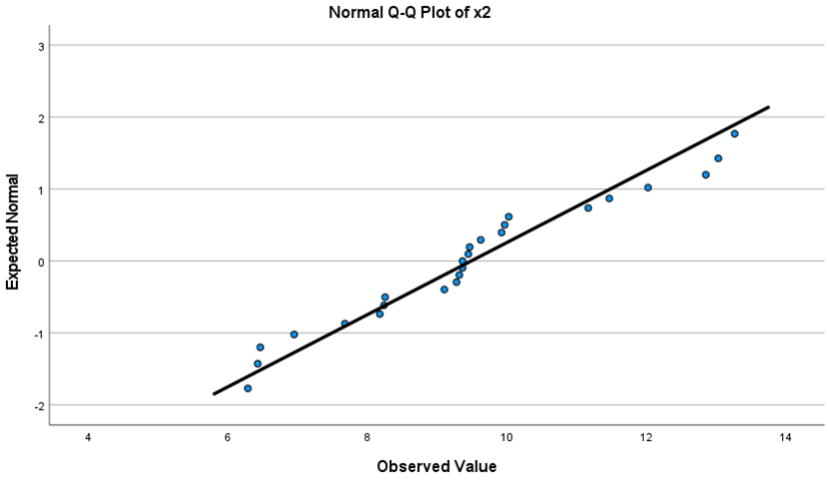
Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
x2	40	100.0%	0	0.0%	40	100.0%



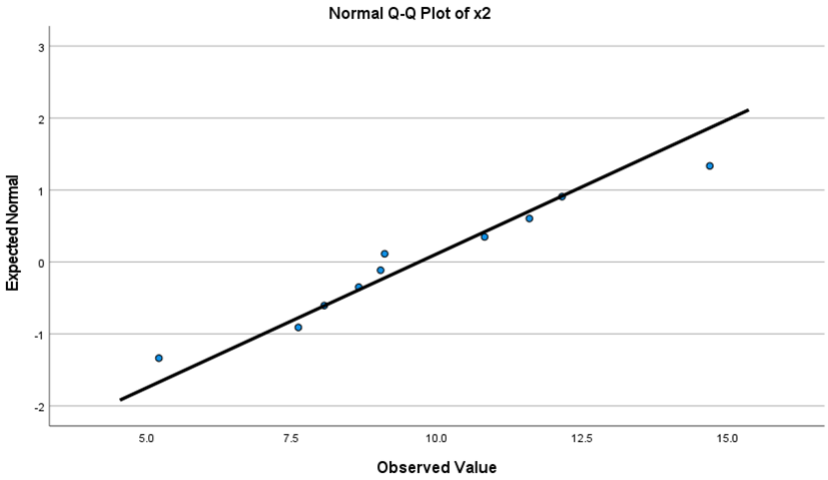
Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
x2	25	100.0%	0	0.0%	25	100.0%



Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
x2	10	100.0%	0	0.0%	10	100.0%



CONCLUSION:
From these graphs obtained through SPSS software, the data sets with higher value of n have dots closer towards the line indicating normal distribution while reduced value of n show deviations of the values from the line more clearly. The outliers are visible accurately with lesser n values.

Exercise 2

1. The 2-samples-t-test is known to robust against violations of the assumption of normal distributions in case of large samples. This means that the 2-samples-t-test holds true for all the assumptions of normal distributions for large samples.
My last name contains **12** letters and hence the variable x2 of both the data sets is taken into consideration. For easy identification, x2 of type A skulls is denoted as **V2** and x2 of type B skulls is denoted as **V8**.

Shapiro-Wilk:

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
V2	.121	17	.200*	.961	17	.647

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

a. Lilliefors Significance Correction

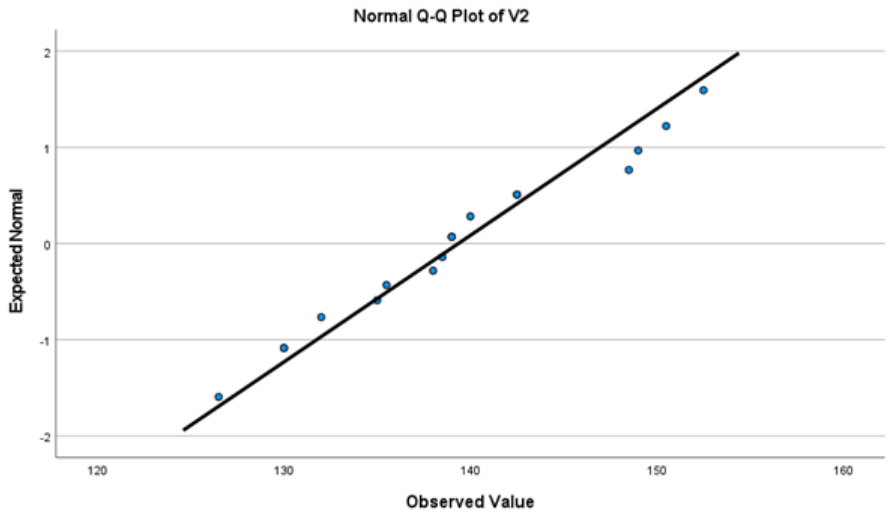
Tests of Normality

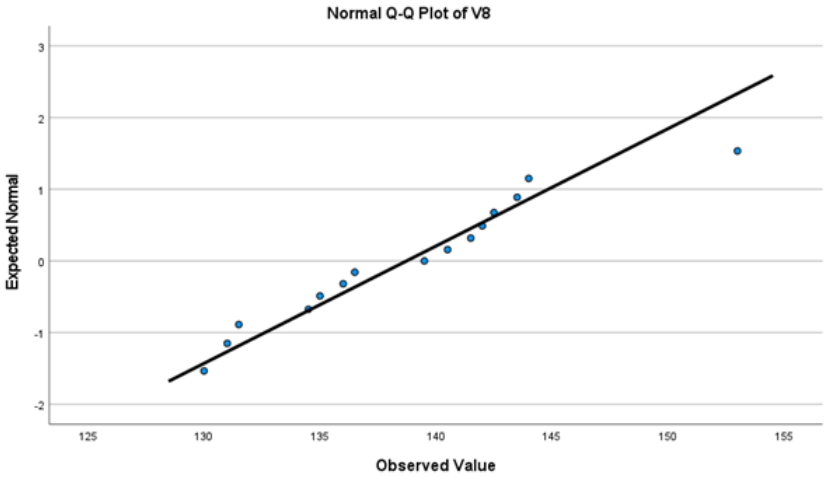
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
V8	.128	15	.200*	.946	15	.467

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

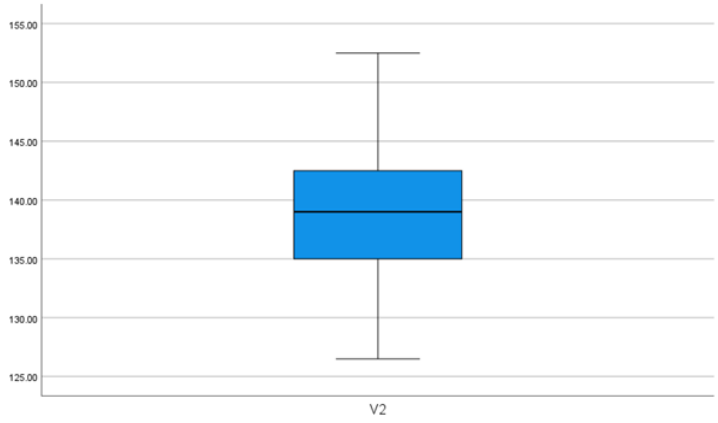
a. Lilliefors Significance Correction

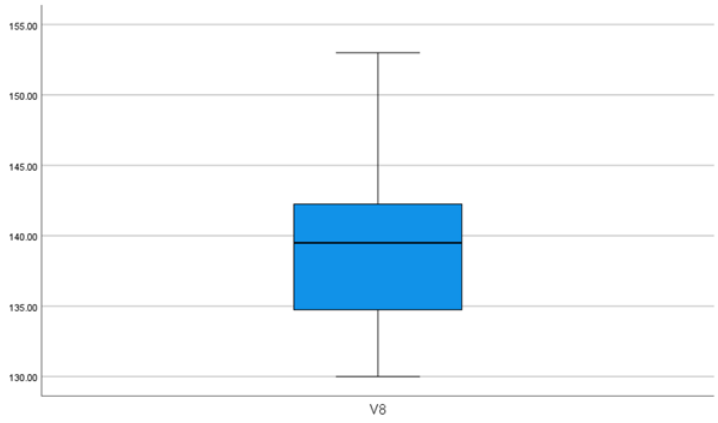
QQ-Plot:



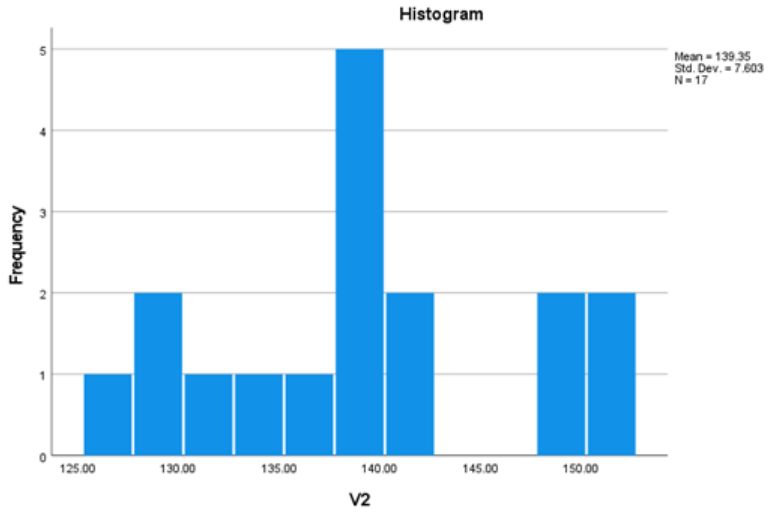

Normal Q-Q Plot of V8
Expected Normal
Observed Value
3
2
1
0
-1
-2
125
130
135
140
145
150
155

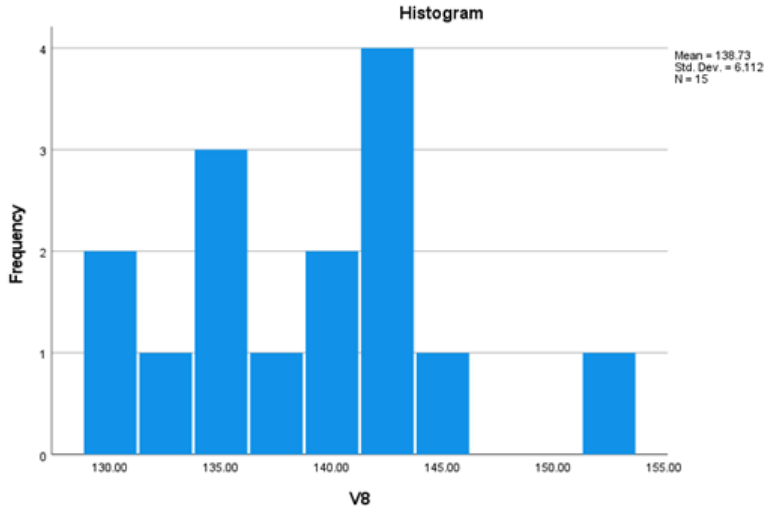
Box-Plot:


155.00
150.00
145.00
140.00
135.00
130.00
125.00
V2


155.00
150.00
145.00
140.00
135.00
130.00
V8

Histogram:





Skewness and Kurtosis:

Statistics

V2

N	Valid	17
	Missing	0
Skewness		.175
Std. Error of Skewness		.550
Kurtosis		-.733
Std. Error of Kurtosis		1.063

Statistics

V8

N	Valid	15
	Missing	2
Skewness		.580
Std. Error of Skewness		.580
Kurtosis		.628
Std. Error of Kurtosis		1.121

CONCLUSION: As the data sets contain very less number of samples, it is considered to satisfy the assumptions of two sample-t-test, if most of the conditions of normal distribution is satisfied. In this case, the Shapiro-Wilk test, QQ-Plot and box plot are satisfied which clearly denotes that x2 of both the types of skulls satisfy the assumptions of two sample-t-test.

2. **Two Sample-t-test:**

Assuming the value of α to be 5%, the following conclusions has been made. Considering the Levene's test for the equality of variances, the p-value (sig.) is higher than 0.05 meaning that the null hypothesis should not be rejected. This means that for the t-statistic, the value of equal variances assumed is taken for calculation. The value of c, with degree of freedom as 30, is found to be 1.697. From the figure below, the value of t=0.293 is less than c. This means that the null hypothesis is not rejected. Therefore, it is concluded that there is **no difference** between the groups i.e., **Type A skulls do not differ from Type B skulls**.

Group Statistics

	V8	N	Mean	Std. Deviation	Std. Error Mean
V2	1.00	17	139.3529	7.60297	1.84399
	2.00	15	138.6333	6.06650	1.56636

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
V2	Equal variances assumed	.565	.458	.293	30	.771	.71961	2.45439	-4.29292	5.73213
	Equal variances not assumed			.297	29.730	.768	.71961	2.41946	-4.22347	5.66269

[Note: In these figures, V2 denotes the greatest horizontal breadth of both the types of skull. The values '1' and '2' of V8 denotes 'type A' and 'type B' respectively.**]**