NAME: LEKAN AHMED NURENI
MODULE: DATA ANALYSIS AND VISUALISATION
MODULE CODE: ACCB4002
LECTURER: ASIM SATTI
ASSIGNMENT TITLE: EXPERIMENTAL DATA ANALYSIS

DATA DESCRIPTION

The hypothesis will be tested using a class test result from 20 of 40 local primary school pupils acquiring mathematical topics.

Descriptive statistics were used to narrate the properties of the data gathered, by giving narration of the sample and data measures, which helped in narrating and understanding the properties of a certain data set. There two sorts of descriptive statistics that can be used to explain and understand the properties of data set.[6]

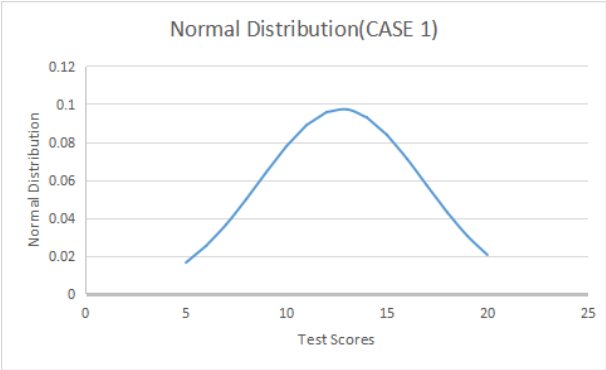
Measures of central tendency are used to describe where the data comes from. This involves the usage of the mean, median and mode, which measurement of variability or spread are used to describe the dispersion of data within a collection. It describes how data are distributed within the set [6]

Table 1 below shows the values for each description of the data for both Case1 and Case2. Table 1:

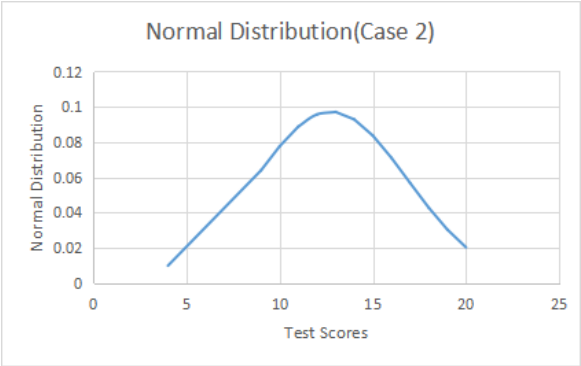
Descriptive Statistics		
	CASE 1(Before)	CASE 2(After)
Observations	40	40
Mean	12.75	15.3
Median	12.5	15.5
Mode	17	20
Maximum	20	20
Minimum	5	4
Range	15	16
Standard Deviation	4.10	3.91
Variance	16.81	15.29
Skew	-0.21	-0.73

1st Quartile	9.75	13
3rd Quartile	17	19
IQR	7.25	6
Lower Bound	-1.125	4
Upper Bound	27.875	28

The data has a normal distribution. When plotted on a graph, the data follows a bell shape as shown in graph1 and graph 2.

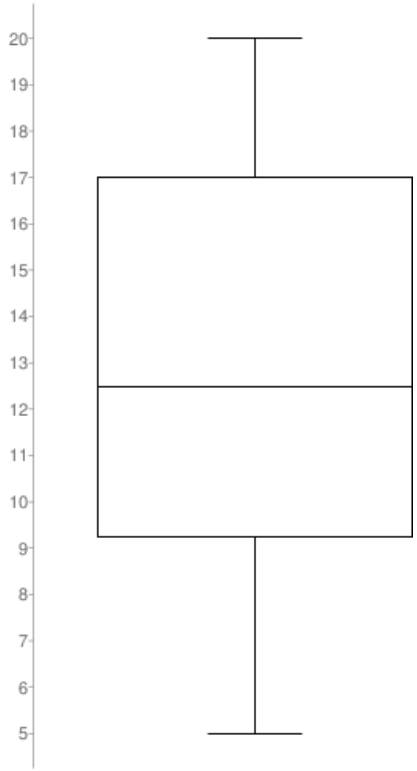


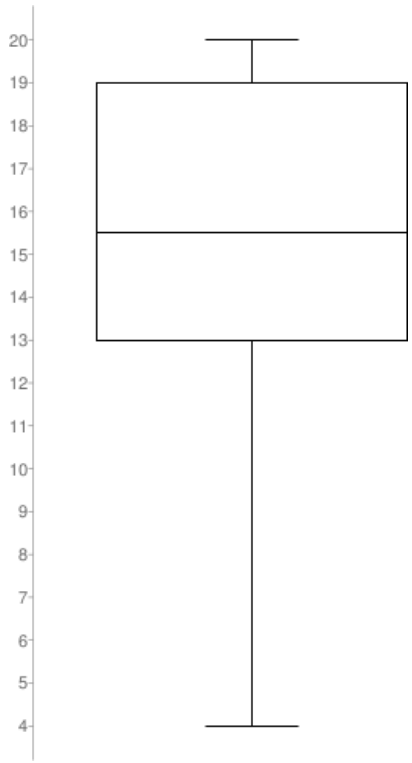
Graph1



Graph2

There are no outliers in both cases of the test scores. The after-test scores case is more skewed than the before. Both can be viewed in the boxplots, graph 4 and graph 5 below.





Graph4 Graph5

DEVELOPMENT OF HYPOTHESES

Hypothesis is a result to a scientific inquiry that is entirely theoretical. Testing, data collecting and experience can all be applied in verifying or disprove a testable hypothesis. In the scientific process, only testable hypothesis is acceptable to be used while designing and conducting an experiment [1]

According to Dr Helmenstine, for a hypothesis to be considered testable, it must meet the following criteria: [1]

- The theory must be able to stand up to critical observation
- It should be possible to show that the hypothesis is false
- The results of the hypothesis must be repeatable

A set of testable hypotheses comprises two opposing claims, namely the Null hypothesis and the Alternative hypothesis. The Null hypothesis, abbreviated as H_0, is the initial hypothesis. The null hypothesis asserts that the parameter has the same value as the stated value or the groups have the same average (their difference is equal to zero). An equal symbol appears in all null hypotheses. It is the hypothesis to be rejected.[3]

- $H_0: \mu_1 = \mu_2$

The alternative hypothesis is a prediction of an effect or relationship based on your study. It is denoted as H_a. Unlike the null hypothesis, it states that the parameter is different from the value. The Alternative hypothesis can be grouped into three (3) which are:

- Two-tailed non-directional, when the claimed value of the population parameter is not equal to the actual value

$$H_a: \mu_1 \neq \mu_2$$

- The right-tailed directional. The population parameter is higher or better than the stated value

$$H_a: \mu_1 > \mu_2$$

- The population parameter is lower or worse than the stated value. This is the left tailed directional.

$$H_a: \mu_1 < \mu_2$$

where
H_0 = null hypothesis,
H_a = alternative hypothesis, μ_1 = mean of sample 1, and μ_2 = mean of sample 2. [4]

According to the task description, to analyze whether the development of the online learning application should continue after seeing the test scores of children who used it, a set of testable hypotheses for the experiment will be.

The Null Hypothesis:
The null hypothesis to be proved states thus

- **The new online learning application does not affect the student test performance.** A more specific null hypothesis while comparing the mean will be

The mean test scores of the children are the same while they use traditional and also for the new online learning application.

$$H_0: \mu_1 = \mu_2$$

The Alternative Hypothesis:

- **Test scores of children when using the new online app are higher than when using the traditional materials.**

$$H_a: \mu_1 < \mu_2$$

The set of hypotheses chosen above is testable and provable and the outcome below can be replicated by any researcher.

HYPOTHESIS TESTING

Test statistic is a numerical summary of data to a single number that may be applied in hypothesis test which provides information about the data can help decide Whether or not to reject the null hypothesis [3]

The major types of test statistics are *Z- Test, F-Test, Chi-Square Test and T-Test statistics.*

Under the null hypothesis, a Z-test is statistical test in which the test statistics distribution can be approximated by a normal distribution. The Z-test can be used to determine the mean of a distribution. Both the Z -tests have the same purpose: to identify the significance of a collection of data. However, because determining the population deviation is challenging, the Z-test is rarely utilized.[3]

A Z-test is a statistic that can be used to examine statistical distributions that can be approximated by a normal distribution when testing the null hypothesis

When determining if data groupings by category are meaningful, F-tests (analysis of variance, ANOVA) are usually utilized.

Chi squared test is used to evaluate whether two variables are connected or not.

Any statistical hypothesis test in which the statistics follows the student’s t-distribution is referred to as a t-test under null hypothesis.

t-test is employed when the value of a scaling term in a test statistic is known. When the scaling term is unknown, the test statistics will follow a student’s distribution and replaced by a data-driven approximation. t-test, for example can be used to determine whether two sets of data have significantly different means.

The commonly used t-tests are one-sample and two-sample. The population mean is compared to the null hypothesis value in a one-sample location test. If the means of two populations are equal, the null hypothesis is tested using a two-sample location test.

In a t-test comparing the means of the two independent samples, the following assumptions must be met:

- The means of the two populations to be compared should be distributed normally.
- The two populations being compared should have the same variance if following the student's original definition of the t-test.
- The test data should either be collected independently from the two populations being compared or properly matched.

Table 2 below shows the Test Statistics and the possible hypothesis test that can be performed with them

Table 2:

Test Statistic	Hypothesis Test
Z- Statistics	One sample Z test
	Two sample Z test
	Two Proportions Z test
T - Statistics	One sample T-test
	Two sample T-test
	Paired T-tests
Chi-Square Statistics	Goodness of fit test
	Chi-Square test of independence
F - Statistics	One way ANOVA
	Poc Host Test
	Analysis of covariance
	Repeated Measure of Variance

T - TEST

T-test statistics is the best fit for the testing, although the z-test statistics will give similar results. The paired sample t-test, also called dependent samples t-test is used with all conditions satisfied for the testing:

- Two sample means are to be compared
- The population standard deviations of both samples are known
- The sample compared is small (40 children)
- Equal pair sample size, i.e the number of students is equal for both cases
- The variances are slightly different
- The test scores are independent of one another.
- The test scores are not affected by any potential confounding factors.
- Each case of the scores is distributed normally (This has been proved in graph1 and graph 2 above).
- There should not be any outliers.
- The degree of freedom(df) used is $n - 1 = 39$, where n represents the number of children

The formula for the paired T-test

$$t = \frac{\sum d}{\sqrt{\frac{n\left(\sum d^2\right) - \left(\sum d\right)^2}{n-1}}}$$

Where d is the sum of the differences

(The total of the differences of each pair divided by the square root of n, times the sum of the differences squared, minus the sum of the squared differences, overall n-1.)

To calculate the paired t-test, following steps was followed:

- Assume that the actual mean difference is zero, which is the nullhypothesis.
- Calculate the difference between the set of observations $d_i = y_i - x_i$.
- Calculate the mean difference.
- Calculate the standard error of the mean difference, which is $S_d / \sqrt{n}$, where n is the total number and S_d is the difference's standard deviation.
- Calculate the value of the t-statistic.
- Check the T-distribution table with the df distribution to get the p-value

The p-value is compared with the level of statistical significance. The level of statistical significance is the set chance of incorrectly eliminating the null hypothesis when it is true. It is expected to determine if the finding is statistically significant enough to reject the null hypothesis. It is denoted by α. The levels of significance for this hypothesis are 0.05, 0.1, and 0.01.

Using Excel software implementation for the steps above, the p-value at each level of significance is shown in table 3

Table 3:

T-Test: Paired Two Sample for Means for alpha (0.5, 0.1, 0.01)		
	Test Score Before	Test Score After
Mean	12.75	15.3
Variance	16.81	15.29
Observations	40	40
Pearson Correlation	0.53	
Hypothesized Mean Difference	0	
df	39	
t Stat	-4.15	
Alpha 0.05, P(T<=t)	0.00	
Alpha 0.1, P(T<=t)	0.00	
Alpha 0.01, P(T<=t)	0.00	

For each level of significance (alpha) the p-value is approximately 0, which means that the test scores before and the test scores of the children after using the online learning app are significantly different. Therefore, the null hypothesis (H_0) is false and should be rejected. The alternative hypothesis (H_a) which states the online learning app improved the test scores should be accepted in each case. The app is a success and the development of it can continue to help the children perform better in future tests.

As stated above the data contains no outliers.
The result of this analysis could have been affected by the selected sample of children, the number of children, the hours the student used in preparing for each test, and the level of difficulty of the tests.

These factors can be controlled by selecting children at random with different backgrounds, a higher number of children should be considered for sampling, the children should be allowed to study for equal hours for both cases and the level of difficulty of the test should be the same.

REFERENCE

[1]A. Helmenstine, Ph.D., "How to Write a Testable Hypothesis", ThoughtCo, 2022. [Online]. Available: https://www.thoughtco.com/testable-hypothesis-explanation-and-examples-609100. [Accessed: 12- Apr- 2022]

[2]K. Gonzalez and R. Ferdinand, https://study.com/academy/lesson/what-is-a-null-hypothesis-definition-examples.html, 2022. [Online]. Available: https://study.com/academy/lesson/what-is-a-null-hypothesis-definition-examples.html. [Accessed: 23- Apr- 2022]

[3]"Importance Of Hypothesis Testing In Data Science", https://analyticsindiamag.com/importance-of-hypothesis-testing-in-data-science/, 2020. [Online]. Available: https://analyticsindiamag.com/importance-of-hypothesis-testing-in-data-science. [Accessed: 18- Apr- 2022]

[4]"Level of Significance (Statistical Significance) | Definition & Steps", BYJUS. [Online]. Available: https://byjus.com/maths/level-of-significance/. [Accessed: 19- Apr- 2022]

[5]"Paired T Test", https://www.chegg.com/homework-help/definitions/paired-t-test-31. [Online]. Available: https://www.chegg.com/homework-help/definitions/paired-t-test-31. [Accessed: 19- Apr- 2022]

[6]"Understanding Descriptive and Inferential Statistics | Laerd Statistics", Statistics.laerd.com. [Online]. Available: https://statistics.laerd.com/statistical-guides/descriptive-inferential-statistics.php. [Accessed: 16- Apr- 2022]

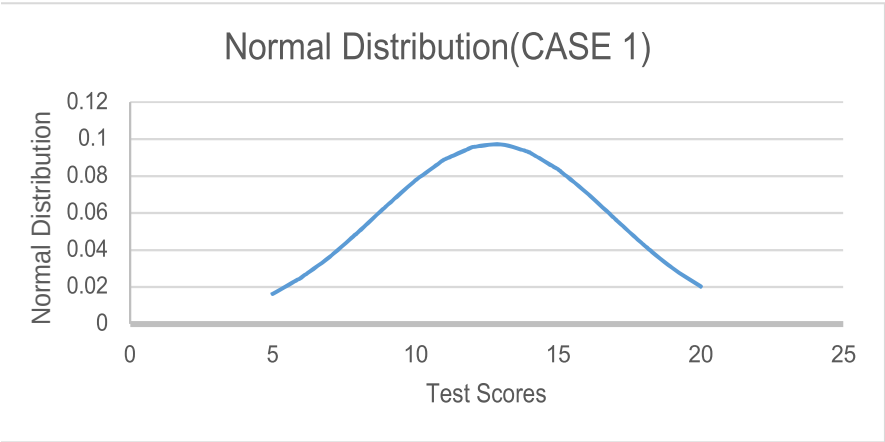
[6]"Understanding Descriptive and Inferential Statistics | Laerd Statistics", Statistics.laerd.com. [Online]. Available: https://statistics.laerd.com/statistical-guides/descriptive-inferential-statistics.php. [Accessed: 16- Apr- 2022]

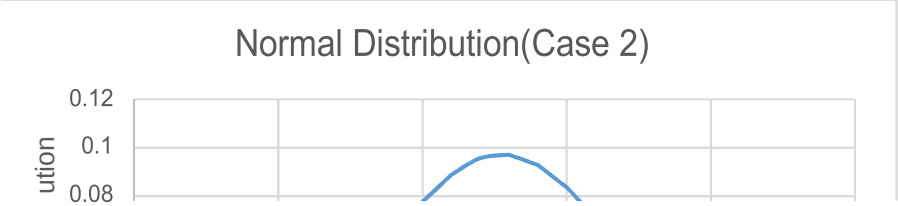
Student	Test Score Before Using App	Test Score After Using App	Normal Distribution(Before)	Normal Distribution (After)
21	5	4	0.016299625	0.00997713
19	6	9	0.025090574	0.064043541
20	6	9	0.025090574	0.064043541
32	6	10	0.025090574	0.077705576
18	7	10	0.036391884	0.077705576
3	8	10	0.049734694	0.077705576
17	8	11	0.049734694	0.088836207
33	8	11	0.049734694	0.088836207
2	9	12	0.064043541	0.0956949
4	9	13	0.064043541	0.097128918
15	10	13	0.077705576	0.097128918
34	10	14	0.077705576	0.092890067
16	11	14	0.088836207	0.092890067
30	11	14	0.088836207	0.092890067
1	12	14	0.0956949	0.092890067
5	12	15	0.0956949	0.083704918
14	12	15	0.0956949	0.083704918
28	12	15	0.0956949	0.083704918
31	12	15	0.0956949	0.083704918
39	12	15	0.0956949	0.083704918
29	13	16	0.097128918	0.071071198
6	14	16	0.092890067	0.071071198
11	14	17	0.092890067	0.056858745
12	14	17	0.092890067	0.056858745
13	14	17	0.092890067	0.056858745
7	15	17	0.083704918	0.056858745
10	15	17	0.083704918	0.056858745
35	15	18	0.083704918	0.042860958
26	16	18	0.071071198	0.042860958
8	17	19	0.056858745	0.030443
22	17	19	0.056858745	0.030443
25	17	19	0.056858745	0.030443
27	17	19	0.056858745	0.030443
36	17	20	0.056858745	0.020373891
37	17	20	0.056858745	0.020373891
40	17	20	0.056858745	0.020373891
9	18	20	0.042860958	0.020373891
38	18	20	0.042860958	0.020373891
23	19	20	0.030443	0.020373891

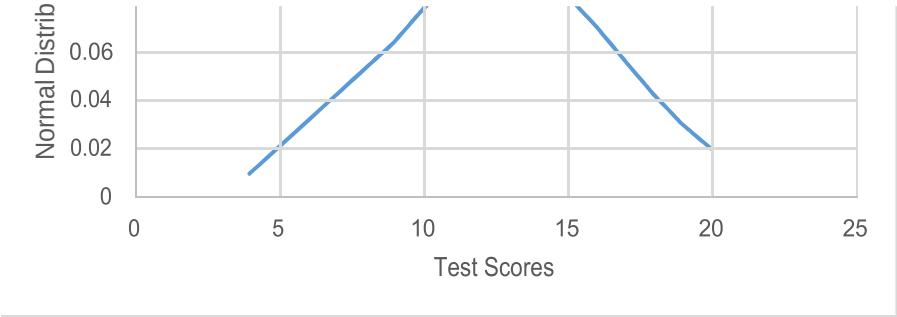
24	20	20	0.020373891	0.020373891

Descriptive Statistics		
	CASE 1(Before)	CASE 2(After)
Observations	40	40
Mean	12.75	15.3
Median	12.5	15.5
Mode	17	20
Maximum	20	20
Minimum	5	4
Range	15	16
Standard Deviation	4.10	3.91
Variance	16.81	15.29
Skew	-0.21	-0.73
1st Quartile	9.75	13
3rd Quartile	17	19
IQR	7.25	6
Lower Bound	-1.13	4
Upper Bound	27.88	28

MIN	5
Q1-MIN	4.75
Q2-Q1	2.75
Q3-Q2	7.25
MAX-Q3	3






Normal Distrib
0.06
0.04
0.02
0
0
5
10
15
20
25
Test Scores

4
9
2.5
6
1

Student	Test Score Before Using App	Test Score After Using App
1	12	14
2	9	9
3	8	10
4	9	11
5	12	15
6	14	15
7	15	15
8	17	10
9	18	20
10	15	20
11	14	18
12	14	16
13	14	17
14	12	14
15	10	20
16	11	15
17	8	4
18	7	10
19	6	11
20	6	19
21	5	16
22	17	17
23	19	19
24	20	18
25	17	20
26	16	19
27	17	20
28	12	14
29	13	14
30	11	13
31	12	12
32	6	9
33	8	20
34	10	17
35	15	17
36	17	17
37	17	19
38	18	20
39	12	13
40	17	15

t-Test: Paired

Mean	
Variance	
Observations	
Pearson Correlation	
Hypothesized Mean Difference	
df	
t Stat	
P(T<=t) one-tail	
t Critical one-tail	
P(T<=t) two-tail	
t Critical two-tail	

t-Test: Paired

Mean	
Variance	
Observations	
Pearson Correlation	
Hypothesized Mean Difference	
df	
t Stat	
P(T<=t) one-tail	
t Critical one-tail	
P(T<=t) two-tail	
t Critical two-tail	

t-Test: Paired

Mean	
Variance	
Observations	
Pearson Correlation	
Hypothesized Mean Difference	
df	
t Stat	
P(T<=t) one-tail	

t Critical one-tail	
P(T<=t) two-tail	
t Critical two-tail	

Two Sample for Means(alpha = 0.05)

Test Score Before	Test Score After
12.75	15.3
16.81	15.29
40	40
0.53	
0	
39	
-4.15	
0.00	
1.68	
0.00	
2.02	

Two Sample for Means(alpha = 0.1)

Test Score Before Using App	*Test Score After Using App*
12.75	15.3
16.81	15.29
40	40
0.53	
0.00	
39	
-4.15	
0.00	
1.30	
0.00	
1.68	

Two Sample for Means(alpha = 0.01)

Test Score Before Using App	*Test Score After Using App*
12.75	15.3
16.81	15.29
40	40
0.53	
0.00	
39	
-4.15	
0.00	

2.43
0.00
2.71

z-Test: Two Sample for Means

	Test Score Before Using App	*Test Score After Using App*
Mean	12.75	15.3
Known Variance	16.81	15.29
Observations	40.00	40
Hypothesized Mean Difference	0.00	
z	-2.85	
P(Z<=z) one-tail	0.00	
z Critical one-tail	1.64	
P(Z<=z) two-tail	0.00	
z Critical two-tail	1.96	