

Data Analysis & Visualization

ASSIGNMENT 2

GRI19845230 (DEWI GRIFFITHS)

TABLE OF CONTENTS

Table of Contents

THE ANALYSIS	3
INTRODUCTION	3
DATASET ANALYSIS	4
HYPOTHESES & ANALYSIS OF FACTORS	5
CONTROLLING FACTORS & THE OUTLIERS	6
CONCLUSION	7
REFERENCES	8
APPENDIX	9

The Analysis

Introduction

The following is an analysis of the results from children using an app to improve their scores on an exam / test which is marked out of 20. The company has requested that their results be analysed for the app to have more development and money put into it which could be used with more than one school from primary to secondary, this is a two-tailed test as there is a comparison of paired data, which is ordinal as it can be grouped due to the order of results not being significant themselves, the values are.

Student	BEFORE	AFTER
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

Dataset Analysis

The two-tailed dataset features forty students which took a test twice, the second test was taken where students revised using an app instead of their revision method for the first test.

The provided data features more students achieving the highest mark earning over 10 points more than what was achieved before they used the app. Although there are students which achieved the highest mark the first time and their score decreased the second time. The app is seemingly successful. [Appendix 1]

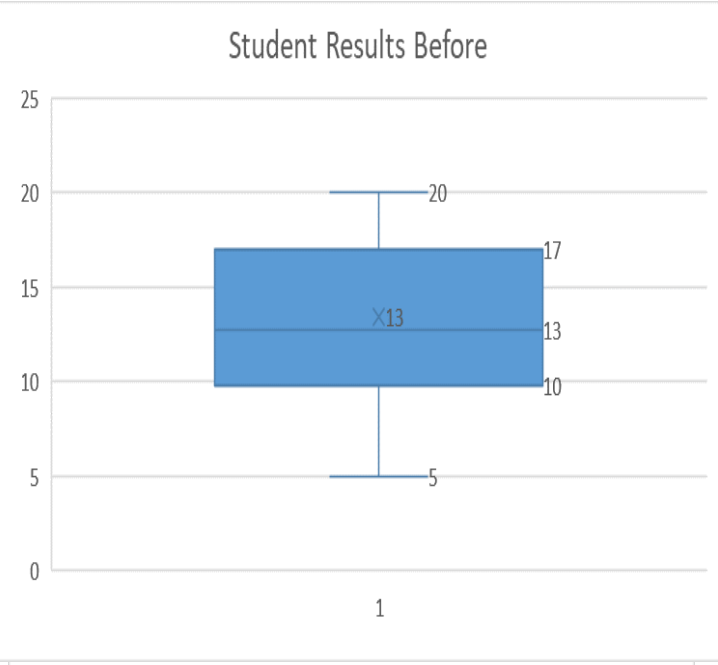


Figure 1

Figure 1 presents us with information about the first test, any result below seven is not significant and that anyone who scored above nineteen is an outlier as it is not common for that score to be achieved.

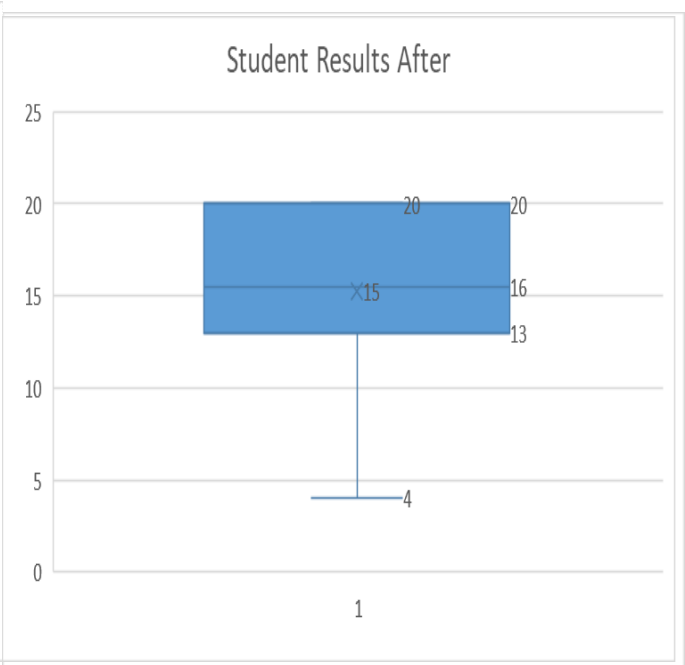


Figure 2

Figure 2 tells us that anyone who scored below nine is not significant and anything above the score is significant as twenty is achieved by many students which took the test.

This can be further visualised using standard deviation. Standard deviation assists with calculating how statistically close a value is to the mean value.

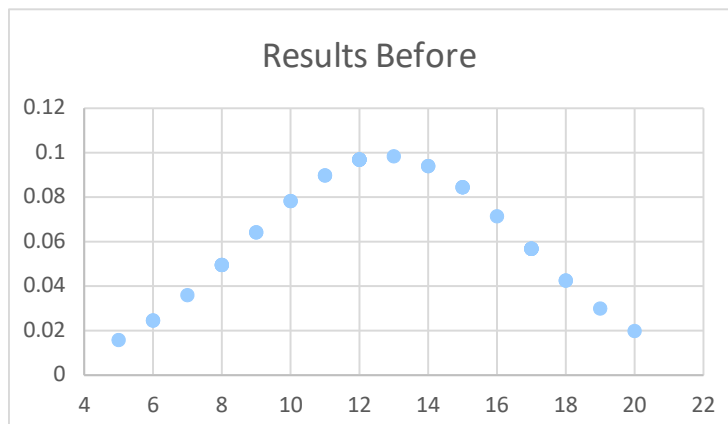


Figure 3

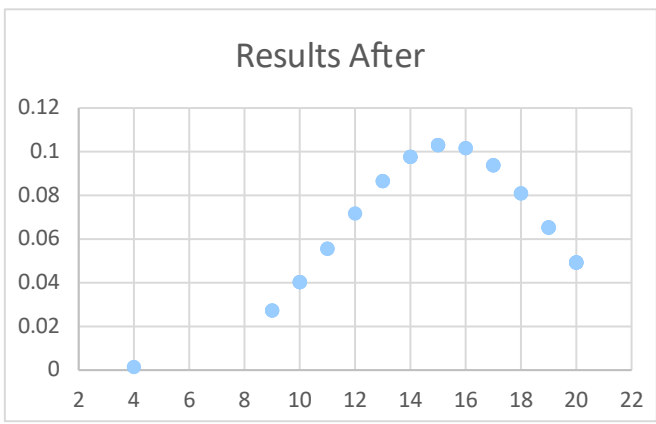


Figure 4

Figure 3 is an outline graph of a bell curve which shows the data before the students have used the app for their test. The mean for the figure is around 12.75 which is where the centre of the curve's head sits its highest on the digital graph. [Appendix 2]

Figure 4 is an outline graph of a bell curve that presents data after students have used the app to study for their second test. The mean for the figure is around 15.30, the curve in this graph is shorter, this is due to scores being closer to the mean, the few students that achieved lower scores are on the lower end of the graph (below 0.05 on the y axis). [Appendix 2]

Hypotheses & Analysis of Factors

It is unclear as to whether or not the students have retaken the same test they just took, if this is so the results aren't that impressive as the app could provide answers and explanations to equations where some of the students may not have revised for their first test which is not a fair test.

Reviewing the results does show some students which achieved higher marks have lower scores on the second test, this shows that the second test could be different or that the app is not useful for successful students, and low scoring students achieved an even lower score. [Appendix 2]

Assuming that every student revised for their first test and took a completely different test after using the app, the time between each test matters too.

If the students took a test within the same day or week this does not provide any significant results as the time to use the app and revise is too short to know if there is any long term advantage. The short term advantage of the app is that it provides answers and explanations on the go which are probably bite-sized (summarized).

H⁰ – The app has no effect on the test results of the students which are in year 5.

Possible outcomes with app having no effect on results:

- The same test was taken the second time.
- The time between each test for revision is too near or far to make a significant and realistic difference.
- Not all students used the app to revise.
- The difficulty of the test questions could have decreased or increased.

H¹ – The app has affected the results of the students which are in year 5.

Possible outcomes of using the app:

- Individual results will decrease or increase significantly.
- The average result will increase or decrease significantly.
- Higher value; lower bound outliers.
- Lower value; lower bound outliers.

After analysing the results, the hypothesis has come to the conclusion that the app – as a study tool – has affected the results which students have achieved as shown by their second set of data. The T-test has calculated the paired data, the result from the T-test is 0.00017 which is below the provided significant values of 0.1, 0.05 and 0.01 which shows the effectiveness of the app with how it had aided students in their study. [Appendix 2]

Considering that the most common value after using the app is also the highest mark does provide us with the information to conclude that the app does affect the results of the students to increase.

Controlling Factors & the Outliers

Factors which are difficult to control:

- The duration at which students choose to revise using the app.
- Students using the app.
- Students with no access to the app at home.

To help prevent students from not using the app, students during activity times can be asked if they want to use the app during their free time, this gives everyone a chance at using the app within the school premises, but not outside of school.

Students with access at home can choose for how long they wish to revise and this varies greatly from each individual, some students will not use the app and stick to their traditional learning methods, this will affect the results with inaccurate assumptions of how the app really improves the results if those students receive a high mark on the second test which could cause the company to have a finished and failed product in the future from receiving inaccurate data.

To encourage the usage of the app it could have reminder notifications for the start of the day, mid-day and after dinner time – if the students enable notifications.

The students that achieve low scores on both of the tests may not be in the habit to study, those students will become outliers within the data so that they do not affect those who did use the app. There is evidence of students not even studying at all which shows with student number 17 receiving the lowest mark in the class on the second, and achieved a higher but still low mark on their first test as well. [Appendix 1, 2]

Removing outliers from the data may not make much of a difference considering that some high scoring students may not have used the app in the first place, even if every student used the app that has significant data, there is no data to show how long they used the app for. If the app logs the number of sessions and the duration that those sessions last for, the data would become clear as to whether the app is successful as a learning tool or not. [Figure 1, 2]

Conclusion

After we have reviewed and compiled a list of statistical data out of the results we have received; that our alternative hypothesis is correct, the app has affected the results in a positive and/or negative way – more importantly it is a positive outcome for what this app has help young students achieve.

Any student with a value below or equal to 5 has failed as there are 20 marks to achieve in the test, so therefore anything above 5 is a pass.

Any mark between 5 and 10 have achieved the grade of **D**, this is due to the marks being less than half of what can be achieved.

Any mark between 11 and 15 is grade **C** as it is more than half of what can be achieved which shows an understanding of the topic.

Marks between 16 and 19 is considered as grade **B** due to its value being close to the highest mark possible.

The highest mark known as 20 gives the student an **A** as their grade.

That is the grading system which is used to present the following statistical data generated by the spreadsheet software.

Results Before

Student results after revising with traditional methods.

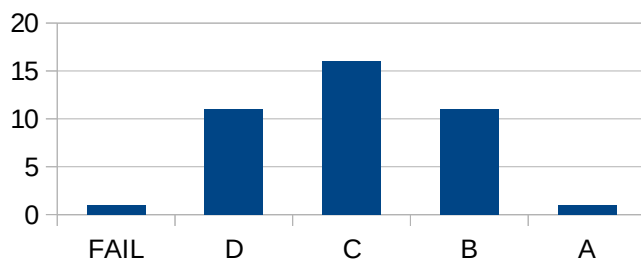


Figure 5

The graph (Figure 6) featuring results where an app was used to revise as opposed to using traditional methods, tell us that more students get higher marks after their revision when using the app. Six more students had received an **A** in the test, two more students than those before had received a **B** for their grades too.

The graph (Figure 5) is generated from results before they have had used the app to study and take a test. The most common grade was receiving a **C** with grades **D** and **B** remaining the same value of 11, one student failed and one student received an **A** in the test.

Results After

Student results after revising with an app.

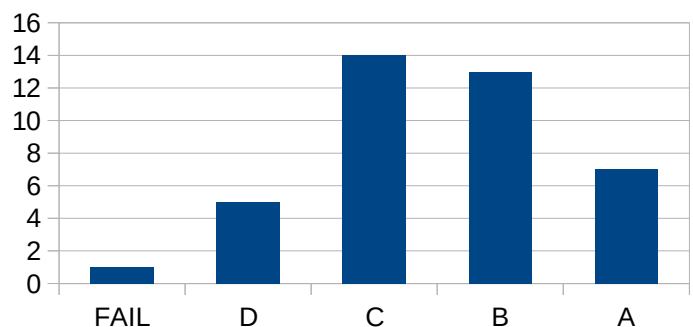


Figure 6

The results have achieved its purpose in assisting students with their studies for earning higher grades. The outcome from analysing the grades have led to the conclusion that the app should continue its development with the company in order to be used by students in year 5 and above.

References

There are no sources in the current document.

Appendix

Appendix Number	Referred Item					Item Description																																																																																																																											
1	<table><tr><th>student</th><th>Test score before using app</th><th>Test score after using app</th></tr><tr><td>1</td><td>12</td><td>14</td></tr><tr><td>2</td><td>9</td><td>9</td></tr><tr><td>3</td><td>8</td><td>10</td></tr><tr><td>4</td><td>9</td><td>11</td></tr><tr><td>5</td><td>12</td><td>15</td></tr><tr><td>6</td><td>14</td><td>15</td></tr><tr><td>7</td><td>15</td><td>15</td></tr><tr><td>8</td><td>17</td><td>10</td></tr><tr><td>9</td><td>18</td><td>20</td></tr><tr><td>10</td><td>15</td><td>20</td></tr><tr><td>11</td><td>14</td><td>18</td></tr><tr><td>12</td><td>14</td><td>16</td></tr><tr><td>13</td><td>14</td><td>17</td></tr><tr><td>14</td><td>12</td><td>14</td></tr><tr><td>15</td><td>10</td><td>20</td></tr><tr><td>16</td><td>11</td><td>15</td></tr><tr><td>17</td><td>8</td><td>4</td></tr><tr><td>18</td><td>7</td><td>10</td></tr><tr><td>19</td><td>6</td><td>11</td></tr><tr><td>20</td><td>6</td><td>19</td></tr><tr><td>21</td><td>5</td><td>16</td></tr><tr><td>22</td><td>17</td><td>17</td></tr><tr><td>23</td><td>19</td><td>19</td></tr><tr><td>24</td><td>20</td><td>18</td></tr><tr><td>25</td><td>17</td><td>20</td></tr><tr><td>26</td><td>16</td><td>19</td></tr><tr><td>27</td><td>17</td><td>20</td></tr><tr><td>28</td><td>12</td><td>14</td></tr><tr><td>29</td><td>13</td><td>14</td></tr><tr><td>30</td><td>11</td><td>13</td></tr><tr><td>31</td><td>12</td><td>12</td></tr><tr><td>32</td><td>6</td><td>9</td></tr><tr><td>33</td><td>8</td><td>20</td></tr><tr><td>34</td><td>10</td><td>17</td></tr><tr><td>35</td><td>15</td><td>17</td></tr><tr><td>36</td><td>17</td><td>17</td></tr><tr><td>37</td><td>17</td><td>19</td></tr><tr><td>38</td><td>18</td><td>20</td></tr><tr><td>39</td><td>12</td><td>13</td></tr><tr><td>40</td><td>17</td><td>15</td></tr></table>					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	The dataset provided by the company developing an app.
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																																																																																																																															
2	<table><tr><th></th><th>BE-FORE</th><th>AFTER</th><th>DESCRIPTION</th></tr><tr><td>MIN</td><td>5</td><td>4</td><td>Smallest value from the set.</td></tr><tr><td>Quartile 1</td><td>9.75</td><td>13</td><td>Lowest non-quart</td></tr><tr><td>MEDIAN</td><td>12.50</td><td>15.50</td><td>Middle value.</td></tr><tr><td>Quartile 3</td><td>17</td><td>19</td><td>Highest non-quart</td></tr><tr><td>MAX</td><td>20</td><td>20</td><td>Largest value from the set.</td></tr><tr><td>MODE</td><td>17</td><td>20</td><td>The most occurring value.</td></tr><tr><td>Significance</td><td>0.01</td><td>0.05</td><td>0.1</td></tr><tr><td>IQR</td><td>7.25</td><td>6</td><td>For the upper and lower.</td></tr><tr><td>Upper</td><td>27.88</td><td>28.00</td><td>Upper bound.</td></tr><tr><td>Lower</td><td>-1.125</td><td>4</td><td>Lower bound.</td></tr><tr><td>StandardDev</td><td>4.04815</td><td>3.86135</td><td>For bell curve</td></tr><tr><td>MEAN</td><td>12.75</td><td>15.30</td><td>Average mark.</td></tr></table>						BE-FORE	AFTER	DESCRIPTION	MIN	5	4	Smallest value from the set.	Quartile 1	9.75	13	Lowest non-quart	MEDIAN	12.50	15.50	Middle value.	Quartile 3	17	19	Highest non-quart	MAX	20	20	Largest value from the set.	MODE	17	20	The most occurring value.	Significance	0.01	0.05	0.1	IQR	7.25	6	For the upper and lower.	Upper	27.88	28.00	Upper bound.	Lower	-1.125	4	Lower bound.	StandardDev	4.04815	3.86135	For bell curve	MEAN	12.75	15.30	Average mark.	Statistics of the data shown in Appendix 1.																																																																							
	BE-FORE	AFTER	DESCRIPTION																																																																																																																														
MIN	5	4	Smallest value from the set.																																																																																																																														
Quartile 1	9.75	13	Lowest non-quart																																																																																																																														
MEDIAN	12.50	15.50	Middle value.																																																																																																																														
Quartile 3	17	19	Highest non-quart																																																																																																																														
MAX	20	20	Largest value from the set.																																																																																																																														
MODE	17	20	The most occurring value.																																																																																																																														
Significance	0.01	0.05	0.1																																																																																																																														
IQR	7.25	6	For the upper and lower.																																																																																																																														
Upper	27.88	28.00	Upper bound.																																																																																																																														
Lower	-1.125	4	Lower bound.																																																																																																																														
StandardDev	4.04815	3.86135	For bell curve																																																																																																																														
MEAN	12.75	15.30	Average mark.																																																																																																																														

		Test	0.0001 7	0.0001 7	0.000172403		
--	--	------	-------------	-------------	-------------	--	--