

Samuel Withers: 1809713

Investigating differences in University Satisfaction
Between mature and young students

Abstract

A study was conducted on two sets of students, one aged 17-20, the other 45-62, with the purpose of determining whether age has an effect on student satisfaction. They were each given a university satisfaction test and their scores recorded. The mean of the young students was found to be 28.1 and the mean of the mature students found to be 44.5. The null hypothesis for the experiment was that age has no effect on student satisfaction. This was tested at significance levels of 0.10, 0.02, 0.01 and found to be significant at every level. Allowing us to conclude that age does have an effect on student satisfaction. The validity of the experimental set up is then evaluated.

Introduction

Data analysis is the practise of applying mathematical techniques to data with the purpose of obtaining useful information and evaluating the validity of the data [1]. It is an integral step to any experimental process as data must be valid and understandable to be of any use. Data analysis is deeply rooted in statistics and is believed to have started with the ancient Egyptians as they took a census for building pyramids [2]. An early example of the usefulness of data analysis is Henry Ford measuring the speed of his production lines. He collected this data and analysed it to optimise the build time and quality of the Ford Model T to the fullest extent. Thus allowing the car to become the most popular car in the world at the time whilst still meeting production demands [3].

Data analysis has evolved throughout history to become an integral part of our everyday lives today, with data being collected on almost everything we do. It is especially important at this point in time to have established techniques in place to analyse data and discover beneficial information from within.

In this study, 10 randomly selected young (aged 17 and 20) and mature (aged 45-62) students took a university satisfaction test and the results recorded. The purpose was to find out whether there is any difference between the satisfaction of mature and young students. The null hypothesis proposed for the study was that age has no effect on student satisfaction. The hypotheses tested to significance levels of $p = 0.10$, $p = 0.01$ and $p = 0.02$ with the goal of finding the statistical validity of the hypotheses.

Theory

Multiple tests were completed on the data collected from this study, this section will overview the techniques used.

Several measures of central tendency were used including the mean, median, mode, range, variance and standard deviation. The formula to calculate the mean was

$$\bar{X} = \frac{\sum x}{N} \quad [4](1.0)$$

Where $\bar{x}$ is equal to the mean, $\sum$ represents the sum of, x is a value in the data set (in this case, satisfaction score) and N is the number of data points (in this case, 10). The mean represents the average value of the data set by summing all numbers and dividing by the number of data points.

The median is calculated by taking the middle value of the data set, in this case that meant finding the mean of the 5th and 6th data point (when the data is sorted). The mode is simply calculated by finding the most commonly occurring value.

The range is calculated by subtracting the smallest value from the largest, this is useful to see how spread out the data is overall, but is heavily effected by outliers or anomalous results. Therefore, it can be more useful to use the variance to see the spread of data, the formula to calculate the variance is

$$\sigma^2 = \frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N - 1} \quad [4](2.0)$$

where σ is the standard deviation. Variance is a useful measure of how spread out a data set is, which can help assess the reliability of results drawn from the data. Standard deviation is then the square root of the variance, and is a measure of how far from the mean you would normally expect the data to deviate. This is useful because it can help you easily identify anomalous results and outliers. It is also an excellent measure of the volatility of the data.

The main test used to determine whether we could reject or accept the null hypothesis was the t test. This test is used to determine whether the means of two samples are different to a significant degree. As the goal of the study is to see simply whether there are differences between the satisfaction levels of mature and young students, and not to determine whether one group is more satisfied than the other, we will be using the two-tailed t test.

The formula to calculate the t value for the t test is

$$t_{obt} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\frac{(n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2}{n_1 + n_2 - 2} \right] \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad [5](3.0)$$

where t_{obt} is the value of t obtained from the data, and the subscripts 1 and 2 represent the two independent data sets. If the sample sizes are the same then this simplifies to

$$t_{obt} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{SD_1^2 + SD_2^2}{n}}} \quad [5](3.1)$$

The null hypothesis for this study is that age has no effect on student satisfaction. We have performed data analysis on the results of the satisfaction surveys for young and mature students and will test the significance of this hypothesis at levels 0.10, 0.02 and 0.01.

Results

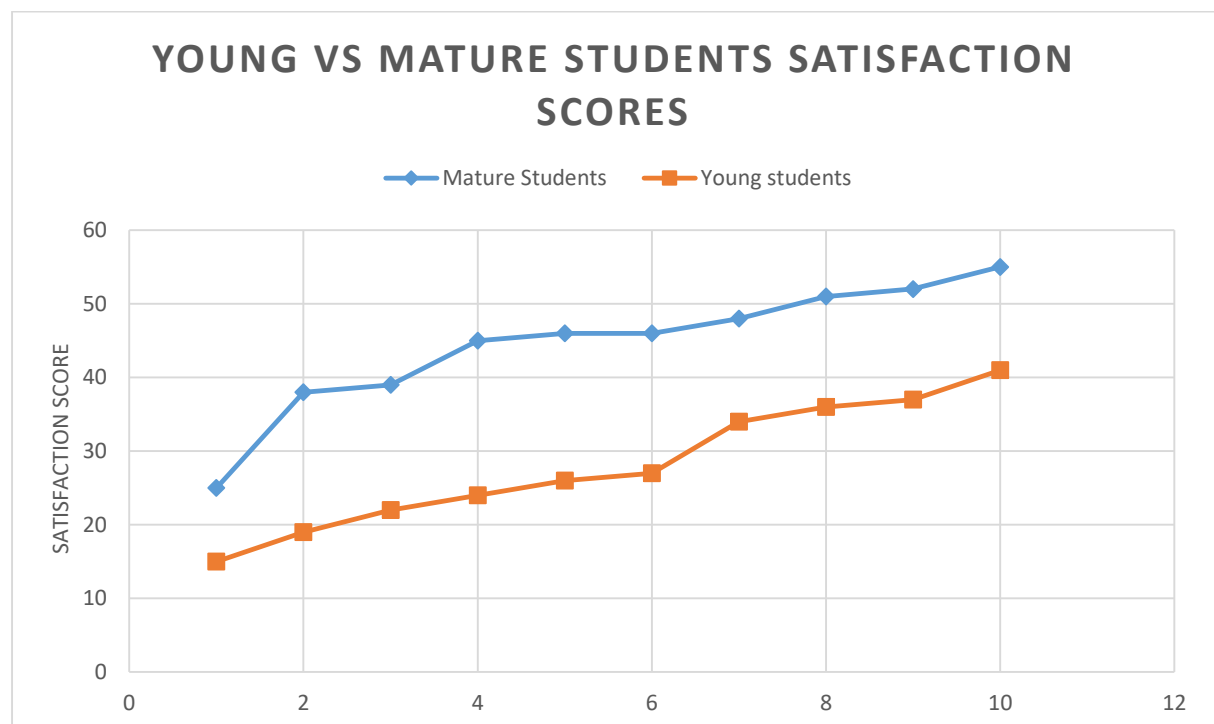


Figure 1: Graph showing the scores of the mature students (blue) compared with the scores of the young students (orange)
The data points are joined simply to help illustrate the separate age brackets.

From this graph we can see that on average the scores of the mature students appear to be higher than those of the young students, however we cannot draw any meaningful results for how spread out and varied that data in each set is.

The mean of the young students' results was 28.1 whilst the mean of the mature students was 44.5. Looking at these results it is tempting to draw the conclusion that mature students have a higher satisfaction score, however, we cannot conclusively say that this is true as we do not know whether the results are skewed by anomalous data or have any quantifiable result to tell us about the validity of the data.

The median of the mature students' results was 46 whilst the median of the young students' results was 26.5. This results tells us much the same as the mean does, with the medians of each set of data being very similar to the means. These two different types of averages being close together is a good indication of the consistency of the data meaning it is likely valid and usable.

The variance of the young students' results was 73.0 whilst the variance of the mature students' results was 75.4. These results are interesting as the mature students' data has a larger variance than that of the young students. Although the difference between the values is not large. This then leads on to the standard deviation. For the young students' the standard deviation was 8.5 whilst the standard deviation of the mature students' results was 8.7. This shows that there isn't really any significant difference between how far we would expect a result in each data set to differentiate from the mean with both standard deviations rounding up to 9 as the nearest integer.

The main test performed on the data to test the significance of the hypothesis was the two-tailed t test. This was performed at alpha levels of 0.10, 0.02 and 0.01. The t_{obt} value of the data was found to be 4.26. The number of degrees of freedom of the data is 18, so this value of 4.26 was then compared to the corresponding values found in the statistics tables. At alpha level 0.10, the t_{crit} value is 1.73 for a two-tailed distribution. Therefore, we can say that at the 0.10 level, $t_{obt} > t_{crit}$ which means it is statistically significant. We can then reject the null hypothesis and accept the alternative hypothesis and say that age does in fact have an effect on student satisfaction.

The story is very similar at alpha levels 0.02 and 0.01 as well, with t_{crit} values of 2.89 and 2.56 respectively. Our t_{obt} value of 4.26 is still significantly larger than either of these, therefore we can one again say that t_{obt} is statistically significant. This means that there is a less than 1% chance that the difference in the means of the two samples occurred by chance alone. Therefore, we can reject the null hypothesis at all significance levels tested and say for sure that according to this data set, age has an impact on student satisfaction.

However, the story changes when we consider outliers. Within the mature students' data set, there is a score of 25. Which is 15 points beneath the next nearest score and 2.25 standard deviations away from the mean. If we disregard this point as an outlier the analysis of the mature students' data changes significantly.

The mean becomes 46.7 instead of 44.5, which isn't a huge difference, however, it does move the mean of the mature students yet further away from that of the young students. The variance and standard deviation become 32 and 5.7 respectively. These values are vastly lower than those for that same data set with the outlier included, meaning the data is significantly closer together and more consistent without the outlier value of 25.

When we perform the t-test again using formula (3.0) this time as the sample sizes are unequal, we get a t_{obt} value of 5.63, which is 1.4 higher than the result when we included the outlier result of 25. However, as the degrees of freedom have changed, so do the t_{crit} values we have to compare to. The value of 5.63 is still significantly larger than the t_{crit} values at alpha levels of 0.10, 0.02 and 0.01 so the conclusion is the same as above. The results gained from analysing the data are statistically significant and we reject the null hypothesis and accept the alternative hypothesis.

Unfortunately, as this is a test based on the opinions of the students, we cannot conclusively say that the outlier results of 25 is truly an outlier. It would be morally wrong to reject the result as it may be a valid opinion of the student who gave the score. As such we must take the results from the full sets of both age groups.

Conclusions

We can conclude from this study that age does have an effect on student satisfaction and this statement is backed by statistics to prove that it is true. This result is the same whether we include potential outlier or not. However, it is always useful to take a common sense look at the experimental set up used to see if it could be improved in future.

The experimental set up used in the study could be changed in numerous ways if repeated in future. The most effective change to increase the reliability of results would be to take into account the different courses put on in a university. Within a university there are multiple courses and a students' satisfaction may vary between courses. Also increasing the sample size from 10 to 100 would better represent a university as a whole due to the number of students within.

Therefore, a proposed change would be to begin with a random sample of ten different courses, and then randomly select 10 students (from each age group) from within each course. This would both increase the sample size and decrease the probability of a result being counted as anomalous simply because it derives from a course with low overall student satisfaction.

If the sample size needed to be kept at 10 for time or budget constraints, then randomly selecting one student for each age group from 10 different courses could also accomplish the same goal.

Repeatability is always an important factor in any experiment, so the study could be repeated another two times throughout the year in order to check that time of year doesn't have any impact on the results. Or whether time of year causes a systematic error, causing the means of both groups to change, but the outcome of the experiment to be the same.

Bibliography

- [1] "Exploring Data Analysis," Spotless Data, [Online]. Available: <https://web.archive.org/web/20171018181046/https://spotlessdata.com/blog/exploring-data-analysis>. [Accessed 29 11 2018].
- [2] "A brief history of data analysis," FlyData, [Online]. Available: <https://www.flydata.com/blog/a-brief-history-of-data-analysis/>. [Accessed 29 11 2018].
- [3] "A brief history of analytics," [Online]. Available: <http://www.dataversity.net/brief-history-analytics/>. [Accessed 29 11 2018].
- [4] D. M. Abuhmida, *RMDA Lecture 2*, 2018.
- [5] D. M. Abuhmida, *RMDA Lecture 7*, 2018.

Appendix

Scores ordered low-high	
Mature Students	Young Students
25	15
38	19
39	22
45	24
46	26
46	27
48	34
51	36
52	37
55	41

Mean	
Mature	Young
44.5	28.1

Median	
Mature	Young
46	26.5

Variance	
Mature	Young
75.4	73.0

StDev	
Mature	Young
8.7	8.5

Variance w/o 25
Mature
32.0

Mean w/o 25	StDev w/o 25
46.66666667	Mature
	5.7

T-value
4.257546666

T-value w/o 25
1.337416852

Mode	
Mature	Young
46	#N/A

Standard Error	
Mature	Young
2.7	2.7

StError w/o 25
Mature
1.9

DoF	DoF w/o 25
18	17

These are all the basic tests performed on the data in excel, showing the results for mature and young students. Results labelled “w/o 25” are the results for the mature students without the potential outlier result of 25.

t-Test: Two-Sample Assuming Unequal Variances

Alpha = 0.10

	<i>Variable 1</i>	<i>Variable 2</i>	
Mean	44.5	28.1	
Variance	75.38888889	72.98889	
Observations	10	10	
Hypothesized Mean Difference	0		
df	18		
t Stat	4.257546666		Significant
P(T<=t) one-tail	0.000236832		
t Critical one-tail	1.330390944		
P(T<=t) two-tail	0.000473663		
t Critical two-tail	1.734063607		

t-Test: Two-Sample Assuming Unequal Variances

Alpha = 0.02

	<i>Variable 1</i>	<i>Variable 2</i>	
Mean	44.5	28.1	
Variance	75.38888889	72.98889	
Observations	10	10	
Hypothesized Mean Difference	0		
df	18		
t Stat	4.257546666		Significant
P(T<=t) one-tail	0.000236832		
t Critical one-tail	2.213703252		
P(T<=t) two-tail	0.000473663		
t Critical two-tail	2.55237963		

t-Test: Two-Sample Assuming Unequal Variances

Alpha = 0.01

	<i>Variable 1</i>	<i>Variable 2</i>	
Mean	44.5	28.1	
Variance	75.38888889	72.98889	
Observations	10	10	
Hypothesized Mean Difference	0		
df	18		
t Stat	4.257546666		Significant
P(T<=t) one-tail	0.000236832		
t Critical one-tail	2.55237963		
P(T<=t) two-tail	0.000473663		
t Critical two-tail	2.878440473		

These are the results of the t-tests performed in excel with their alpha levels labelled and the result of whether they are significant or not labelled on the right.