**Introduction:**

The way that students spend their time out of the classroom is important as it effects how they perform in class, and ultimately if they pass or fail their classes as a result of time spent studying or time used in other ways. We decided to collect and analyze data that we have collected from students to see how they spend their time outside of class and what effect it has on their GPA. Also, as many students are working in collage in order to be able to afford it we are attempting to see if there is a correlation between students with full time jobs and the effects that those jobs have on grades.

**Data collection characteristics:**

The data that was collected was from the volunteer collection method that will be described later, the characteristics of the data collected are as follows:

GPA – the GPA that was collected had a range from 2.0 – 4.0 with class widths of 0.499. There were five different classes to chose from in the sampling method whose source was the individual whose data we were collecting.

Full time employee- the working status had either two choices, yes or no. The individual chooses what option best fitted their current employment status.

Hours spent in class, hours spent studying, hours spent with friends, hours spent online, hours spent with friends, hours spent with family, hours spent on a hobby, hours spent doing a other (non-specified) activity were all done on the individuals weekly schedule so as to have more a more normal timeline as people have reoccurring weekly schedule. There were five classes again for each of the different items with class sizes of 5, the ranges were 0-4, 5-9, 10-14, 15-19, 20-24.

**How Data was collected and what was collected:**

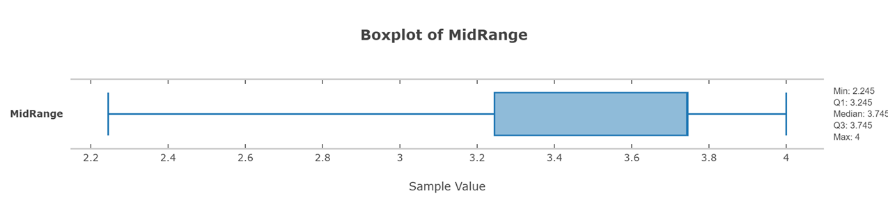
The data was collected using the Voluntary response sample method of data collection by posting the survey on an online forum for students in collage as that was the target of the study. This method does provide easy access to data collection as the survey is out there for people to post on and does provide easier access. However, there are some drawbacks as well, such as the very method of this data collection is biased in nature as there is no way to tell if the individual who is posting the information is being truthful in the information that they are offering.

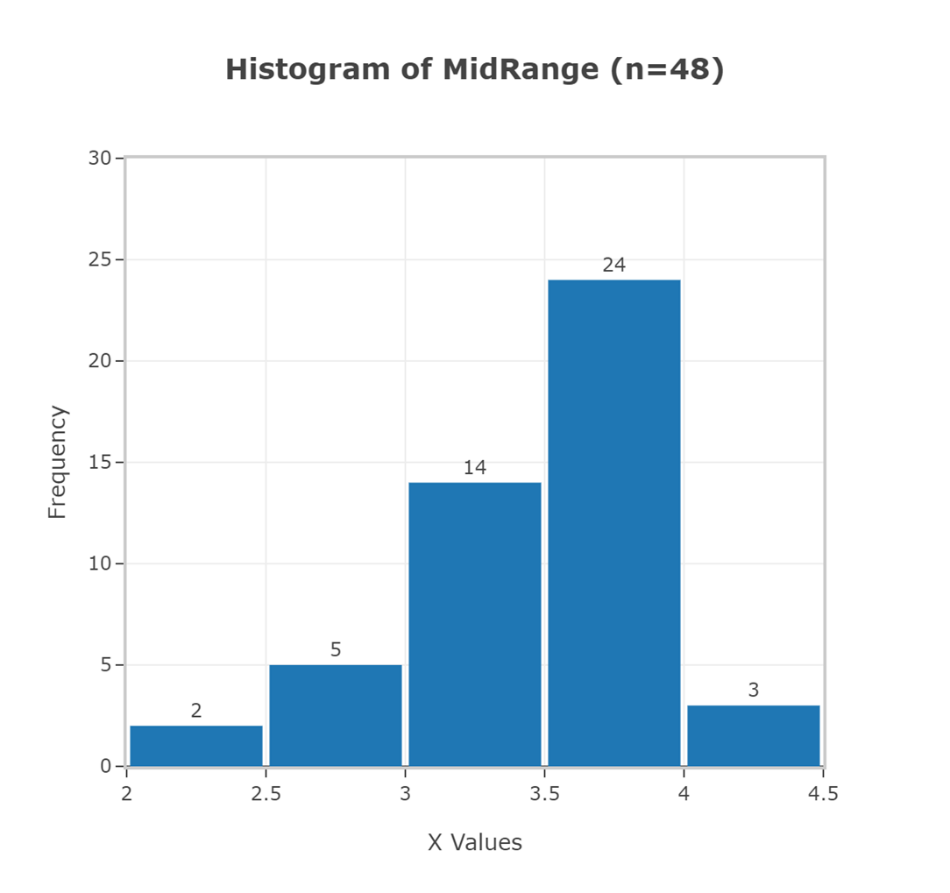
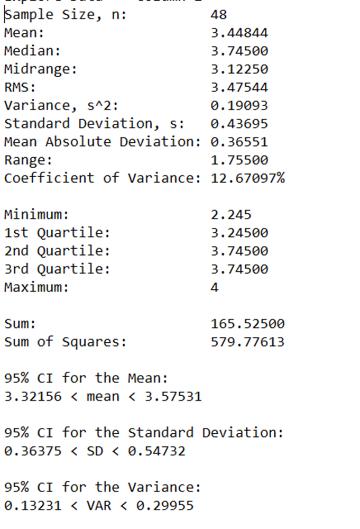
The data that was collected then assigned values for each of the data points and processed the information using technology (Statdisk) to analyze the data to see the different descriptive statistics of each of the different data sets that were generated through the study. To see if there is a correlation and regression between GPA and having a full-time job, GPA and hours in a week spent studying, and hours in a week spent with friends and GPA. Checking the hypothesis that students spend a certain amount of time per week studying.

**Exploratory Data Analysis:**

The raw data that was generated from the survey were lists of ranges for each of the different categories per individual who answered. To be able to process the data the Midpoint of the data ranges needed to be calculated for each of the data sets, for the full-time employment data set numeric values had to be assigned to be assigned in order to process the data. When the Midpoints had replaced each of the different data sets each column was processed separately in order to find the characteristics of the GPA data sets as well as to see if there were any outliers and to identify if the data was normally distributed and to see if there were any correlations between the GPA and other parameters.

When reviewing the characteristics of the different data sets it appears that the GPA data on a Normal Quantile Plot shows that the data is close to a straight line so it is approximately normally distributed. When reviewing the histogram from when the GPA goes from Low to High the chart can be loosely applied as normally distributed as the bell-shape is roughly there. GPA characteristic show that the range of data is 2.0 – 4.0 as was the range of the survey that was given. The mean is 3.45 GPA and the Median GPA is 3.745, this shows that the majority of surveyed students GPA are on the higher end of the scale as the interquartile range is from 3.245 to 3.745, showing that the top 75% of students have a B average or higher. See the charts below. There does not seem to be any outliers in the data as well.





When reviewing the Hours of classes that a student attends in class per week we see that the hours spent studying have a mean of 13.46 hours and a median of 17 hours per week, but when we review the mean and median of the hours spent studying, we see that the mean of hours spent studying is 9.08 hours and median of 7 hours per week. It seems that more time is spent in the classroom in an average week then studying for class. The mean of students hanging out in a week is 7.1 with the median having the value of 7. When also reviewing the amount of time spent in a week studying with the amount of time students spend with friends the mean is the only difference between the two boxplots.

When using technology (Statdisk) reviewing the correlation between the students GPA and the hours spent in a week studying we find that the P-Value is 0.1436 with a significance value of 0.05 the r is 0.21 and critical r is 0.2845. So, we see that the P value is greater than the significance value and that r is smaller than critical r. So, there is no correlation between the time in a week that a student spends studying and their GPA. See the visual data at the end. (Figure 1)

When using technology (Statdisk) reviewing the correlation between the students GPA and the hours spent in a week hanging out with friends we find that the P-Value is 0.1745 with a significance value of 0.05 the r is -0.199 and critical r is 0.2845. So, we see that the P value is greater than the significance value and that r is smaller than critical r. So, there is no correlation between the time in a week that a student spends studying and their GPA. See the visual data at the end. (Figure 2)

When using technology (Statdisk) reviewing the correlation between the students who have full time jobs and the hours spent in a week studying we find that the P-Value is 0.1401 with a significance value of 0.05 the r is -0.22 and critical r is 0.2845. So, we see that the P value is greater than the significance value and that r is smaller than critical r. So, there is no correlation between the time in a week that a student spends studying and their GPA. See the visual data at the end. (Figure 3)

Interpretations from the data generated show that despite students not spending that much time studying for class they are still producing good grades in class. That there is no correlation between a student’s grades and the amount of time spent studying for class. No correlation between a student’s GPA and how much time is spent hanging out with friends. It also shows that students who work tend to spend the same amount of time studying as those who don’t and are full time students.

**Inferential Statistics:**

When reviewing the amount of time students spend studying it appears that less than half of their free time outside of work and class are spent studying.

Hypothesis:

H0: p=0.5, where p is the probability that students spend half of their time outside class studying

H1: P<0.5, where p is the probability that students spend less than half of their time outside class studying

The α = 0.05

1. It does not meet the criteria that it is a simple random sample as the information was gathered using a voluntary response sample, however other conditions are meet, such as:
2. The conditions for Binomial Distributions are satisfied:

* There are a fixed number of trials. There is a fixed number of hours to choose from the survey (2421) hours.
* The trials are Independent. Each survey is separate from one another.
* Each trial has two categories of success or failure. The success is studying time and the failure is time not spent studying.
* The probability will not change.

1. The requirements that n\*p≥5 and n\*q≥5 is both satisfied as shown below:

N\*p = (2421)\*(0.5) = 1210.05≥5

N\*q = (2421)\*(0.5) = 1210.05≥5

p=0.5; q=0.5 x=436; n=2421; p1 = x/n

p1 = 436/2421 = 0.18009

Z = ((p1)-p)/Square root((p\*q)/n) = (0.18009-0.5)/Square root((0.5\*0.5)/2421) = -0.31991/Square root(0.000103) = -0.31991/0.10161846 = -31.48

Because the z value is far beyond the limits of the normal standard distribution the P-Value is +00000. Because the P-value is less than the α = 0.05 which is less than the significance level we reject the null hypothesis.

Because we reject the null hypothesis, we support the alternative hypothesis of p<0.5. We conclude that there I sufficient sample evidence to support the claim that more than half of student’s free time is spent on other items besides studying for class

Presentation of results:

These results show that studying outside of class has little impact on a student’s GPA as well as how much time they spend with friends or whether or not that they work a full-time job. It may imply that students are not spending as much time studying outside of class due to finding more efficient ways of studying, or that they are taking easier collage courses that to not require as much effort. It could also imply that there is dishonesty in collage with students violating academic standards as well.

The results of this study have a strong probability of not being correct or accurate. One reason is that the method of data collection was a voluntary response sample method, which carries an inherent bias as those who participate could be lying or adjusting numbers to look better despite the survey being anonymous. Another reason that the data might not be correct as it is unknown where the samples came from in terms of region or which school the students are in as some schools have different standards than others. There are also different disciplines that the students are studying in college that range in requirements and criteria that is given. So, there is no way to tell if a science major is being compared to a liberal art. Each have their own difficulties in the nature of their programs.

Going forward there are ways to remove or reduce the bias of the survey. Some ways include a larger sample size so that there is a larger pool that samples can be collected. The survey could also could have been left up longer as to allow more people to have access to it so that they can participate. There might have been other sites that had more people accessing them so there were more people who could have seen the survey. The survey should be limited to a particular area or even just one University so that students from the same area are being compared. The class sizes could be smaller in size and there be more in order to collect more accurate information. There are also other methods of data collection that may be better suited for gathering information as to alleviate the possibility of a bias, such as an experiment where people have time loggers that are used to track and see how much time they are doing each activity. There could also be time loggers that do not work so that people do not know if they have a working time logger or not, each bit of information is amanous and no one would know who has what time logger. It should also be included as to the general studies of the students that are taking part in the study as each major has their own criteria to graduate. The different studies should be compared to themselves and not to different majors.

**Appendix for Large figures:**

Chart, box and whisker chart

Description automatically generated

Boxplot of student’s GPA and Hour of classes they take.

Chart, box and whisker chart

Description automatically generated

Boxplot of student’s GPA and how many hours they hang out.

Chart, box and whisker chart

Description automatically generated

Boxplot of student’s GPA and how many hours they study.

Chart, box and whisker chart

Description automatically generated

Boxplot of working status and how many hours they studyChart

Description automatically generated

Boxplot of student’s GPA

Chart, line chart

Description automatically generated

Normal quantile plot GPA

Chart, histogram

Description automatically generated

Histogram GPA

Chart, histogram

Description automatically generated

Histogram Hour of Studying

Chart, treemap chart

Description automatically generated

Correlation Matrix Heatmap

Graphical user interface, text

Description automatically generated

Figure 1: Correlation GPA & hour of studying

Graphical user interface, text

Description automatically generated

Figure 2: Correlation GPA & hour of hanging out

Graphical user interface, text

Description automatically generated

Figure 3: Correlation working status & hour of studying

Graphical user interface, text

Description automatically generated

Correlation GPA & working status.

Pie Charts:

Chart, pie chart

Description automatically generated

Chart, pie chart

Description automatically generated