

Ashley Tignor

Math 1040-043

Term Project Report

Purpose of the study:

Our group was interested in finding out if student's exercise habits influence their GPAs. The research question our group chose was, For SLCC students attending the Redwood Campus, does hours of exercise per week relate to GPA?

Data Collection Plan:

The sampling method we choose for this project was the systematic sampling method. We choose to interview students who walked through the door of the Student Center building at the SLCC Redwood Campus. We randomly selected $k=5$ and each started with the 4th person that walked in the door. So our selections started with the 4th person that walked in the door, then continued with the 9th, 14th, 19th, etc. We continued until we each sampled 14-15 students so that as a group we collectively interviewed 100 people.

Group 3 data:

Hours of exercise per week	GPA
4	3.1
3	3.92
4	2.79
6	3.8
2	2.5
3	3.06
1	2.5
6	3.7
3	3
3.5	2.5
2.5	2.96
7	2
2	3.78
8	2.5
3	3.4
0.5	3.9
3	3.8
0	3.4
0	2.7
0.5	2.7
1	3.9
1	3.7
4	2.9

2	3.3
2	3.3
3	3.8
4	2.5
5	3
1	3
0	2.6
8	2.9
3	2.1
6	3
0	3.5
0	2.5
10	3
8	2.5
8	4
0	3.7
8	3
6	3.5
8	2.5
10	2
5	3.8
4	2.9
10	3.5
2	3
8	3
2	3.8
1	3.4
15	3.2
3	4
5	2.2
8	3.1
4	1.5
2	2.5
3	3.5
3	3
1	2.5
3	3.35
3.5	2.9
6	3.8
5	3
2	2.3
5.5	3.1
3	2.5
10	3
2.5	3.5
2	2

Mean: 3.92 $392/100=3.92$

Standard deviation: 2.88388066

5 point summary: Minimum- 0

1st Quartile- 2

Median- 3

3rd Quartile – 5

Maximum- 15

Range: 15 $15-0=15$

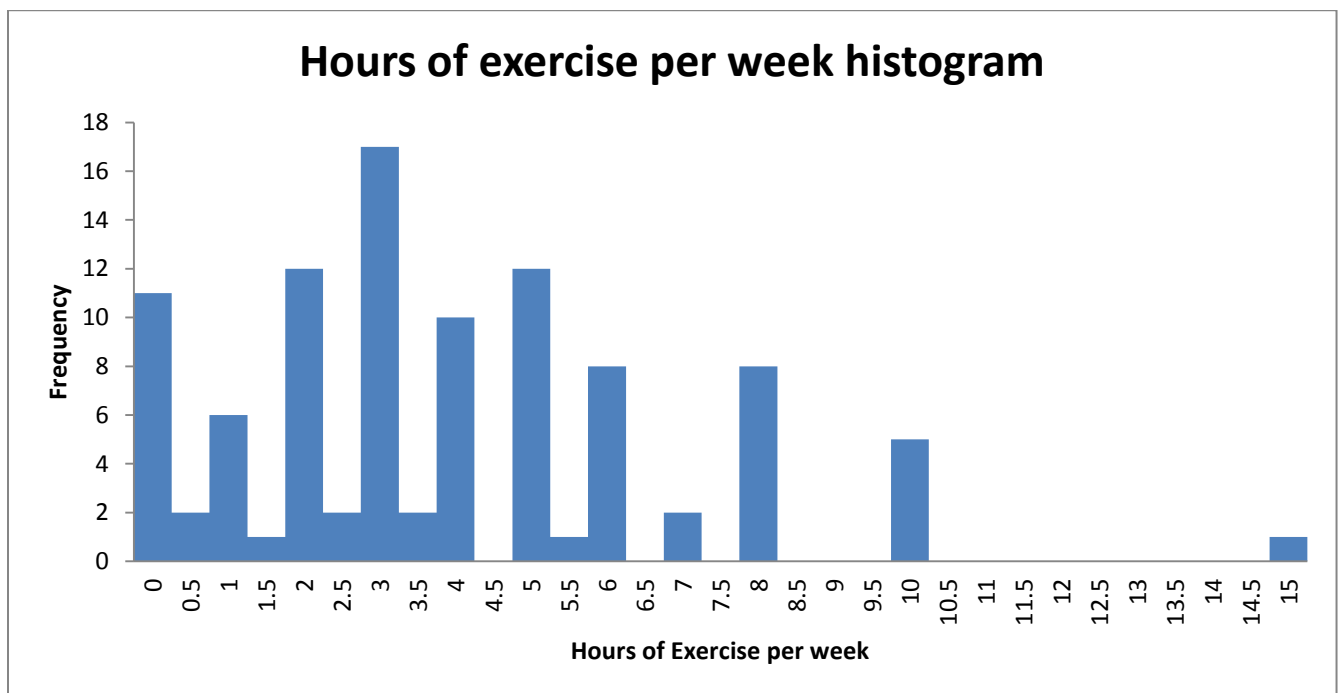
Mode: 3

Outliers: $Q3-Q1=IQR$ $5-2=3$

$Q1-1.5(IQR)=$ Outlier $2-1.5(3)= -2.5$

$Q3+1.5(IQR)=$ Outlier $5+1.5(3)= 9.5$

The outliers for this data are 10,10,10,10,10, and 15.



3rd Quartile- 3.5

Maximum- 4

Range: $4 - 1.5 = 2.5$

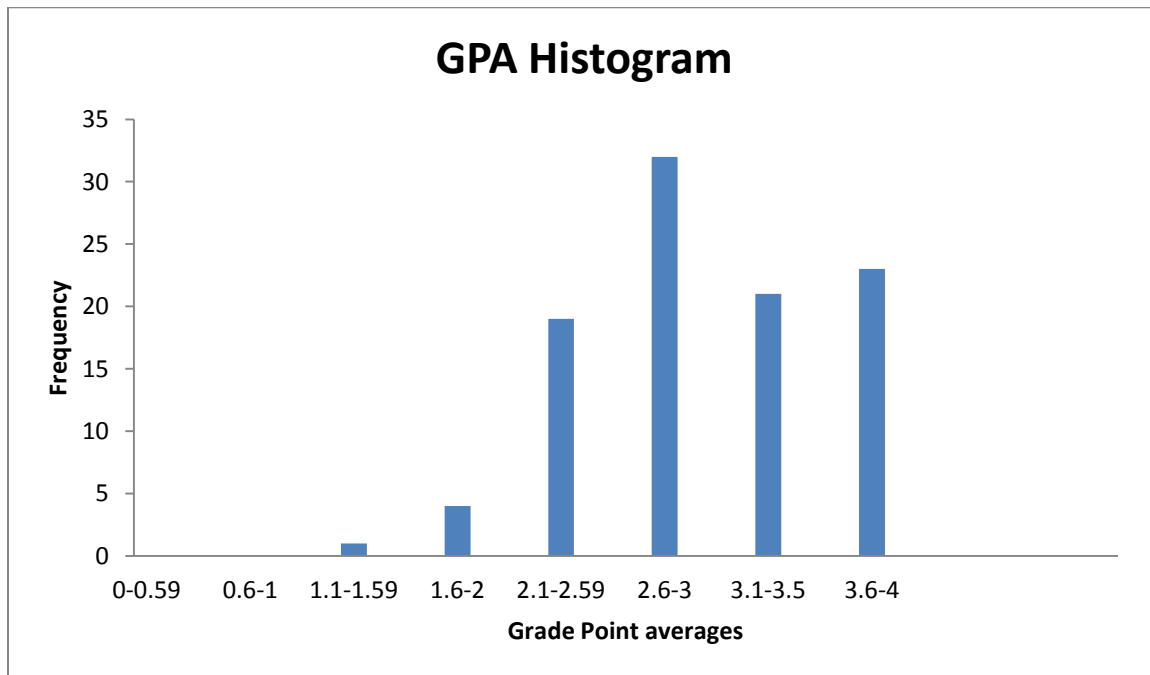
Mode: 3

Outliers: $Q3 - Q1 = IQR$ $3.5 - 2.6 = 0.9$

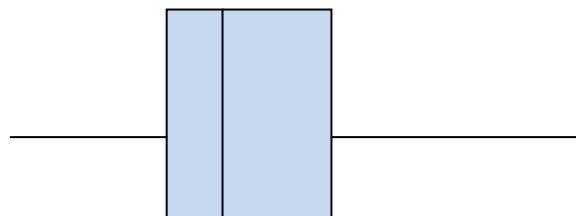
$Q1 - 1.5(IQR) = \text{Outlier}$ $2.6 - 1.5(0.9) = 1.25$

$Q3 + 1.5(IQR) = \text{Outlier}$ $3.5 + 1.5(0.9) = 4.85$

There are no outliers for this data.



GPA Boxplot



0 0.5 1 1.5 2 2.5 3 3.5 4 4.5 5 5.5 6

Testing Correlation:

The correlation coefficient can be used to tell us the strength and direction of the hours of exercise per week and the student's GPA.

$$r = \frac{1}{n-1} \sum \left(\frac{x - \bar{x}}{s_x} \right) \left(\frac{y - \bar{y}}{s_y} \right)$$

$$\bar{X} = 3.92$$

$$\bar{Y} = 3.069$$

$$S_x = 2.88388066$$

$$S_y = 0.58034$$

Using this equation our correlation coefficient is **-0.05493452**. This coefficient tells us that there is a moderate negative relation between the two qualitative variables.

Line of Regression:

$Y = a + bX$ X is the explanatory variable, Y is the dependent variable, b is the slope, and a is the intercept.

Using our data, hours of exercise per week is our explanatory variable and GPA is the dependent variable.

The sum of our X variable is 392 and the sum of our Y variable is 306.9. The sum of $X*Y$ is 1188.88 and the sum of X^2 is 2360.

$$306.9 * 2360 - 392 * 1188.88$$

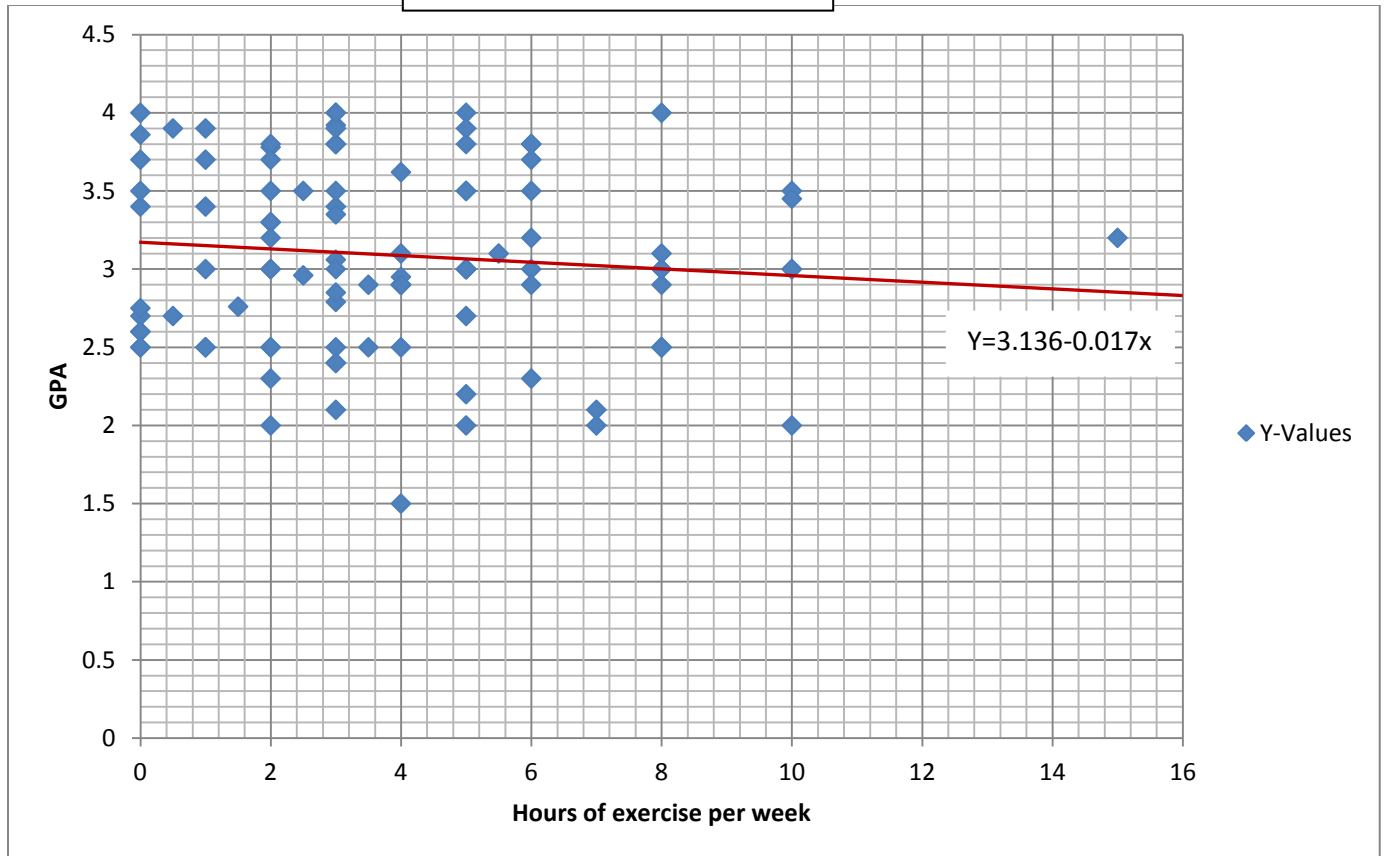
$$\frac{100 * 2360 - 392^2}{100 * 2360 - 392^2} = 3.136 = \text{intercept (b)}$$

$$100 * 1188.88 - 392 * 306.9$$

$$\frac{100 * 1188.88 - 392 * 306.9}{100 * 2360 - 392^2} = -0.017 = \text{slope (a)}$$

Line of regression equation: $y = 3.136 - 0.017x$

Line of Regression Scatterplot



Difficulties/Surprises encountered:

I found it difficult dealing with the calculations of so much data. It took quite a bit of time to do the calculations for 100 data points for each of the variables and double check each of my results. Often I found that I missed a number and would have to go back through all the information and start over again.

I honestly wasn't sure if our data would show a strong correlation between the hours of exercise per week and GPA. I suspected that increased exercise might lead to doing better in school but I know that my exercise levels are low and my GPA is high so I wouldn't follow my own prediction. I was curious to see what other student's response would be and how their data would fit into my prediction. Our data ended up surprising me by showing a moderate negative correlation between the hours of exercise and GPA for the students we interviewed.

Analysis:

The distribution of the variable, hours of exercise per week, is skewed to the right because the data is mostly distributed on the left side of the graph. The distribution of the variable, GPA, is

skewed to the left because the data is mostly distributed on the right side of the graph. Our data ended up showing a moderately negative correlation between the two variables with our value R at -0.05493452. The critical value for our sample size is 0.205 using the degree of freedom at 98 and the 0.05 as the level of significance. This shows that there is not a significant relationship between the hours of exercise per week and GPA.

Interpretation and conclusions:

To answer our original question, For SLCC students attending the Redwood Campus, does hours of exercise per week relate to GPA? , I would say that no, there is not a significant relationship between the two variables. Some individuals may do better in school when they get significant exercise throughout the week but as a population there isn't a strong enough correlation between them to prove that there is a definite relation between the two. I feel that our sample accurately reflected the SLCC Redwood Campus population. Our samples were taken at different times and on different days using systematic sampling method to choose our 100 interviewees which allowed us to interview a variety of SLCC students. If I was to do something differently in repeating this project I would probably broaden my research question to all college students. I would adjust the question to say, For college students, does hours of exercise per week relate to GPA? I would be interested in seeing how the results of college student's exercise and GPA would be in compared to SLCC Redwood Campus student's results.