Aria Askaryar

Professor Pineda

Math 338-85

8 December 2020

Does the amount of Sleep have correlation with GPA

Whether you are a billionaire or penniless we all have to sleep. However the question arises does the amount of sleep you receive determine your success? If the penniless man slept more during highschool would he be more successful today? Although there has been numerous studies on covering connections between sleep and cognitive function, there still remains a lack of data using quantitative statistical measures to directly represent a link between sleep and academic performance. So in my project I have decided to analyze and test, in various statistical methods, whether the amount of Sleep has a strong correlation to academic performance. My prediction is that the more sleep a student obtains, the greater their GPA will be. I think that the relationship between the amount of sleep a teenager receives compared to their GPA is a direct relationship. The belief is increased amounts of sleep duration will lead to higher academic performance. With various complex steps we will test this hypothesis.

Gathering Data: Since there seems to be a lack of credible data on the subject, I have decided to make my own data set using Simple Random Sampling type test. With software tools/platforms such as Instagram pools and SurveyMonkey I was able to obtain a large population of data. My survey consisted of 3 questions. "What is your academic position? A) Senior B) Junior C) Sophomore D) Freshman" "How many hours do you sleep on a school night?" and "What is your overall grade/GPA?" Since I am conducting the survey myself I had

to create a strata which was the academic position. After receiving 120 data results I continued with only 107 since I disposed of the outliers and corrupt data, meaning applicants who failed to complete the entire pool or provided false information. The data was then transferred onto an excel sheet (Figure 1). Some of the sample sizes answered "What is your overall grade/GPA?" in a grade percentage so I converted all data for this question into GPA. However on the excel for more accuracy I calculated and displayed both percentage and GPA average in the data set.

Issues with the Data: Unfortunately due to the circumstances of the year 2020 there is no available data on the subject because of the effects of Covid-19, leaving the project with no control group. There are some biases which must be addressed for this particular study, this data is uniquely bounded by schools within southern California, 2020 Fall semester, and within the Covid-19 situation period. The condition of the Corona-virus plays various other obstacles such as resulting in a 6+month quarantine where individuals can not leave their homes and traditional schooling converted to virtual schooling. There is also a strict time restraints, which means the collected data is only performed over the span of a week. With all these factors we are reduced to a moderate sample size with no control group and controversial effects of the corona-virus to our sample. However despite all the setbacks this project will still be able to prove the hypothesis true or false. Even more interesting if someone post-covid plans to further test this hypothesis can use my data and results to see how the virus affected these factors.

Testing Data: Now that the biases have been discussed and the parameter has been set, we can move on and start testing the hypothesis. The null hypothesis states that the amount of sleep is not significant enough to have correlation to academic performance. While the hypothesis claims that the amount of sleep will have a strong correlation to academic

performance (Figure 2). I will first test my hypothesis through various types of graphs to see if there is a direct correlation. In Rstudio, I imported my dataset from the excel sheet in order to create some graphs. The first graph is a histogram of both the GPA and hours slept of all students (Figure 3 & 4). Then I continued by creating histogram for each specific class denoted in the figure below (figure 6). I began the test with a T-test in R to see if there are differences or correlation between the data, or in other words if the differences are by chance or not. The T-test gave a result of a p-value 0.2025 (Figure 5). To further test my results I made a Boxplot of both GPA and Hours spent sleeping for each class (Figure 7). I then tried by computing a chi-squared test in Rstudio (Figure 8). After the chi-square test I decided to create a scatterplot graph for each class (Figure 9). I also ran a linear regression model to further analyze the results (Figure 10).

Interpreting Data: The first test allowed us to see the mean, median, Q1, Q3, and standard deviation for each class's GPA and time slept. The GPA mean for Freshman was 89.6, Sophomore 90.2, Junior 90.8, and Senior 89.7. As for the Hours spent sleeping mean was Freshman 7.4, Sophomore 6.8, Junior 7.1, and Senior 6.9. With the results from T-test, I conclude with 95% confidence that the difference of the sample means resides somewhere between 3.385725 and 3.846051. By the law of the T-test since 0 is not in the interval there is not a difference and we accept the null to be true. In the graphs we see there is no class that has a unique GPA which means it is a true representation of the dataset, also meaning there is no outliers. The mean of all GPA's and Hours are very close showing there is limited to no evidence that there is a corelation between sleep and academic grades. When viewing the scatter plot graphs, we can see the results are scattered around and does not following a linear pattern which means the relationships are direct and have a weak correlation. This makes it difficult to fit a line because taking a line which has constant change and fitting it to data that is always changing

with no pattern. In the Linear relationship we see the corelation between a student's GPA and the amount of hours of sleep they receive on school nights that is positive because the slope shows positive growth. However the slope is miniture, which means that the relationship is not as strong as other tests. The boxplot test also shows the outliers of the expirment as seen in the Juniors GPA and the hours spent sleeping in the Freshman.

Conclusion: In final conclusion after checking multiple tests we come to see the null hypothesis ended up being true which was rejected by hypothesis and prediction made in the beginning of the experiment. All the results show that the correlation between the amount of sleep a student receives and their GPA is weak and not sufficient enough alone to make a difference. Although there seems to be a slight correlation as seen in the linear example however as stated before its too small of a difference to effect the entire statistical project. Even with all the obsticals made from the year 2020 and covid-19 this expirment taught me a lesson. Now thinking about the expirement it makes sense since other factors such as how one spends his time is more important then the amount used sleeping. A better expirment would be how much time spent studying has an effect of academic performance and GPA. As the null hypothesis stated the "Amount of sleep has no significant correlation to academic performance" has been proven true with a 95% confidence. So if the penniless man had slept more during highschool I would be 95% sure he would not be more successful today than he is.

References

Figure 1: Dataset collected from students

Class_name	Students	Hours_Spent_Sleeping	Grade	GPA
Freshman	1	5	84%	3
Freshman	2	6	83%	2.8
Freshman	3	6	86%	3.1
Freshman	4	7	89%	3.4
Freshman	5	7	89%	3.4
Freshman	6	7	89%	3.4
Freshman	7	7	90%	3.5
Freshman	8	7	90%	3.5
Freshman	9	7	90%	3.5
Freshman	10	7	90%	3.5
Freshman	11	7	90%	3.5
Freshman	12	7	91%	3.6
Freshman	13	7	91%	3.6
Freshman	14	7	96%	4
Freshman	15	7.5	85%	3
Freshman	16	7.5	90%	3.5
Freshman	17	7.5	94%	3.4
Freshman	18	8	81%	2.6
Freshman	19	8	85%	3
Freshman	20	8	92%	3.7
Freshman	21	8	93%	3.8
Freshman	22	8	96%	4
Freshman	23	8.5	91%	3.6
Freshman	24	8.5	92%	3.7
Freshman	25	9	79%	2.4
Freshman	26	9	90%	3.5
Freshman	27	9	94%	3.9
Sophmore	28	3	92%	3.7
Sophmore	29	5	86%	3.1
Sophmore	30	5.5	91%	3.6
Sophmore	31	5.5	93%	3.8
Sophmore	32	6	79%	2.4
Sophmore	33	6	88%	3.3
Sophmore	34	6	96%	4
Sophmore	35	6.5	89%	3.4
Sophmore	36	7	87%	3.2
Sophmore	37	7	89%	3.4
Sophmore	38	7	90%	3.5
Sophmore	39	7	91%	3.6
Sophmore	40	7	92%	3.7
Sophmore	41	7	96%	4
Sophmore	42	7	96%	4
Sophmore	43	7.5	84%	2.9
Sophmore	44	8	85%	3
Sophmore	45	8	89%	3.4
Sophmore	46	8	90%	3.5
Sophmore	47	8	91%	3.6
Sophmore	48	8	94%	3.9
Sophmore	49	8	94%	3.9
Sophmore	50	8	96%	4
Sophmore	51	8.5	85%	3
Junior	52	5	83%	2.8
Junior	53	6	87%	3.2
Junior	54	6	90%	3.5
Junior	55	6	90%	3.5
Junior	56	6	95%	4
Junior	57	6	96%	4
Junior	58	6.5	95%	4
Junior	59	7	75%	2
Junior	60	7	78%	2.2
Junior	61	7	85%	3
Junior	62	7	90%	3.5
Junior	63	7	92%	3.7
Junior	64	7	92%	3.7
Junior	65	7	92%	3.7
Junior	66	7	94%	3.9
Junior	67	7	93%	3.8
Junior	68	7	96%	4
Junior	69	7.5	91%	3.6
Junior	70	7.5	94%	3.9
Junior	71	8	89%	3.4
Junior	72	8	93%	3.8
Junior	73	8	93%	3.8
Junior	74	8	95%	4
Junior	75	9	93%	3.8
Junior	76	9	93%	3.8
Junior	77	9	95%	4
Senior	78	4.5	90%	3.5
Senior	79	5	82%	2.7
Senior	80	5	83%	2.8
Senior	81	5	88%	3.3
Senior	82	5	93%	3.8
Senior	83	6	83%	2.8
Senior	84	6	87%	3.2
Senior	85	6	90%	3.5
Senior	86	6	96%	3.6
Senior	87	6	96%	3.6
Senior	88	6.5	95%	3.5
Senior	89	7	80%	2.5
Senior	90	7	84%	2.9
Senior	91	7	89%	3.4
Senior	92	7	90%	3.5
Senior	93	7	92%	3.7
Senior	94	7	93%	3.8
Senior	95	7	94%	3.9
Senior	96	7.5	90%	3.5
Senior	97	8	85%	3
Senior	98	8	86%	3.1
Senior	99	8	89%	3.4
Senior	100	8	91%	3.6
Senior	101	8	91%	3.6
Senior	102	8	91%	3.6
Senior	103	8	93%	3.8
Senior	104	8	95%	4
Senior	105	8	96%	4
Senior	106	8.5	85%	3
Senior	107	9	91%	3.6

Figure 2: Hypothesis

H0 : Amount of sleep has no significant correlation to academic performance.
H1 : Amount of sleep does have significant correlation to academic performance.
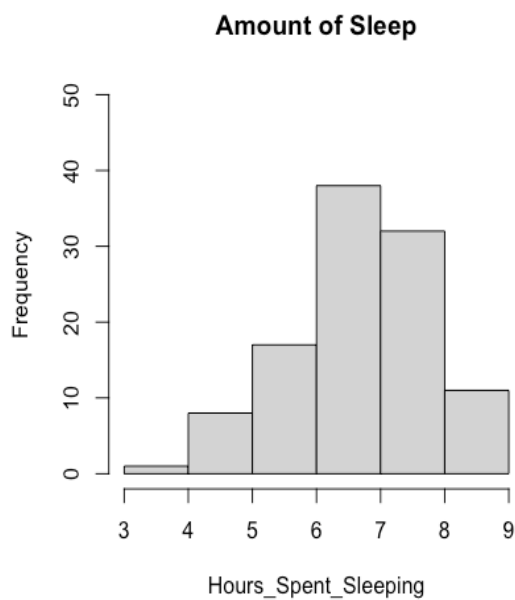
Figure 3: Histogram of Hours spent sleeping   Figure 4: GPA level of all Students
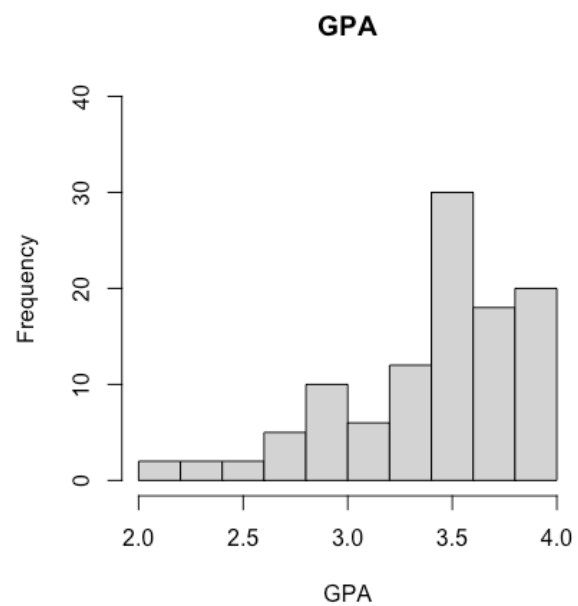

Figure 5: T-test

Welch Two Sample t-test

data: x1 and x2
t = 31.065, df = 137.18, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 3.385725 3.846051
sample estimates:
mean of x mean of y
 7.079439 3.463551

Figure 6: Histogram for each class
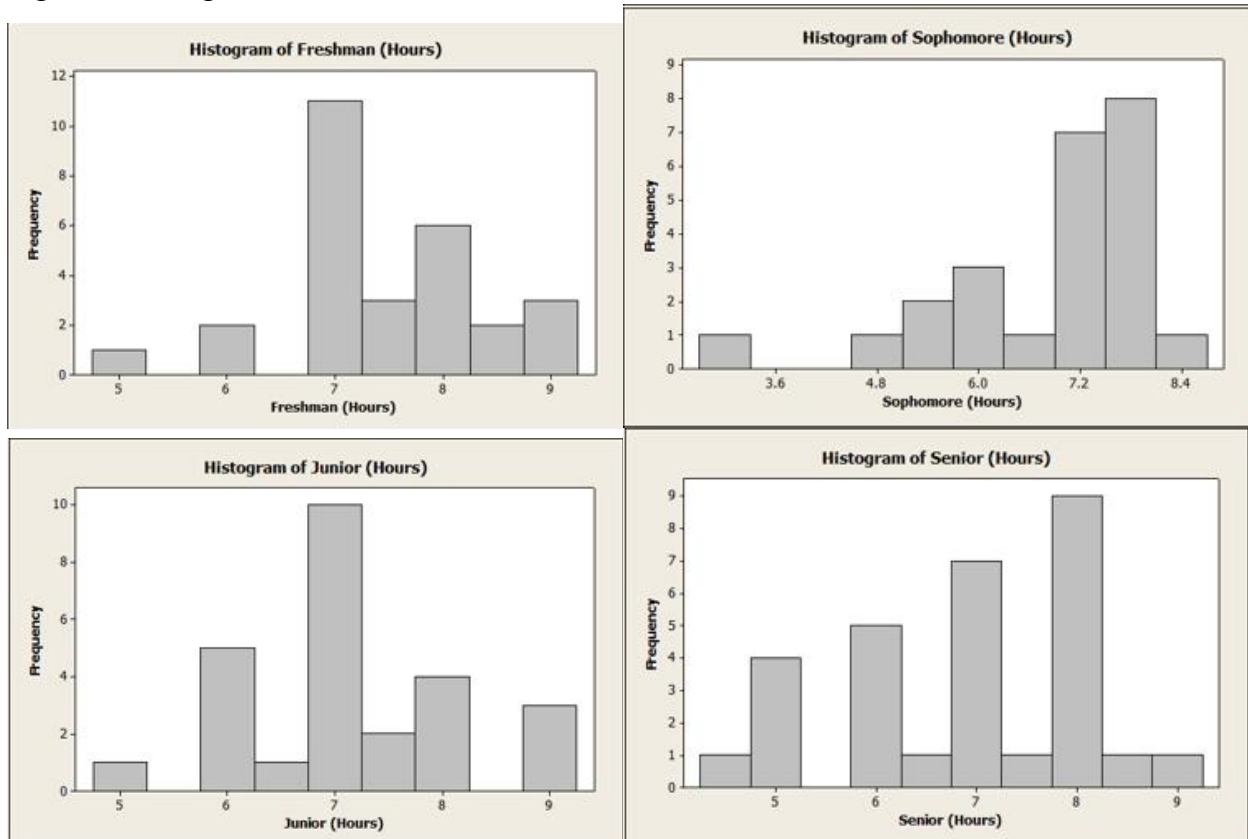
Figure 7: Boxplot graphs
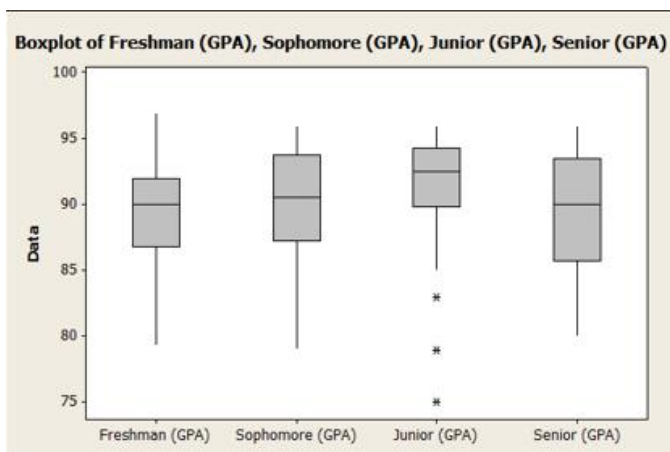
GPA   Hours Spent Sleeping
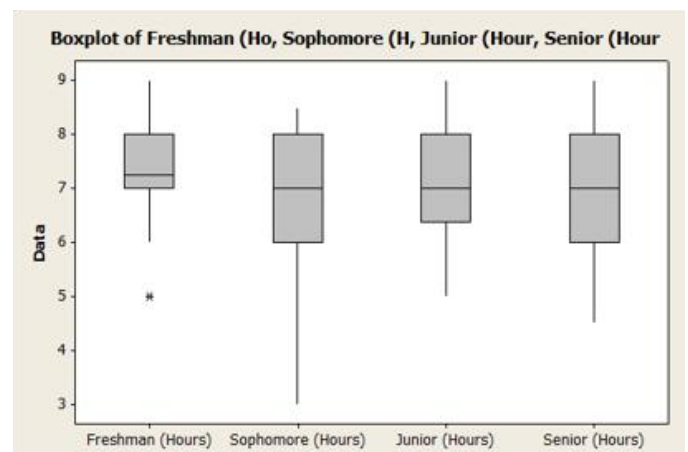
Figure 8: Chi-square Test

> chisq.test(Table)

Pearson's Chi-squared test with Yates' continuity correction

data: Table
X-squared = 1.6243, df = 1, p-value = 0.2025
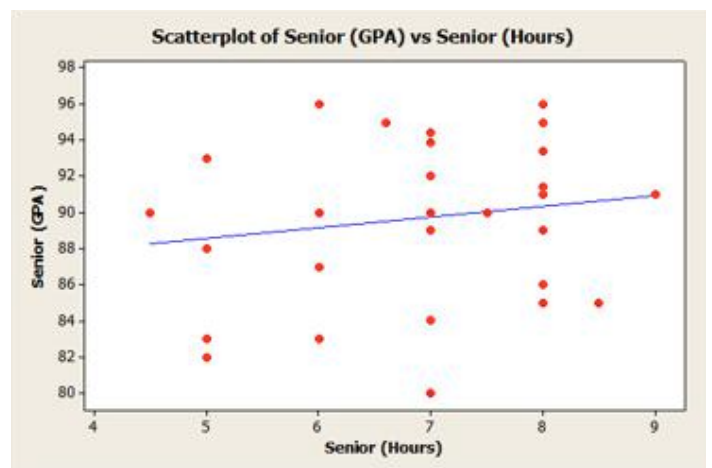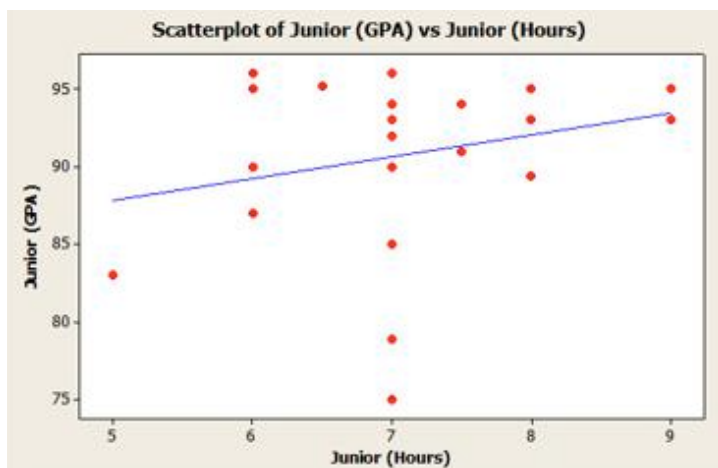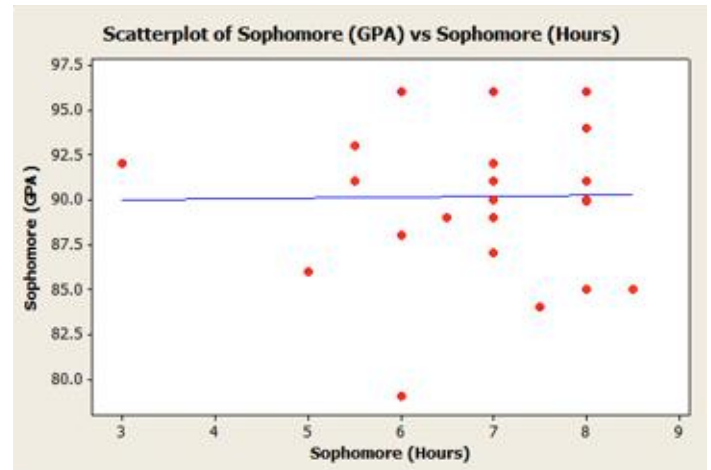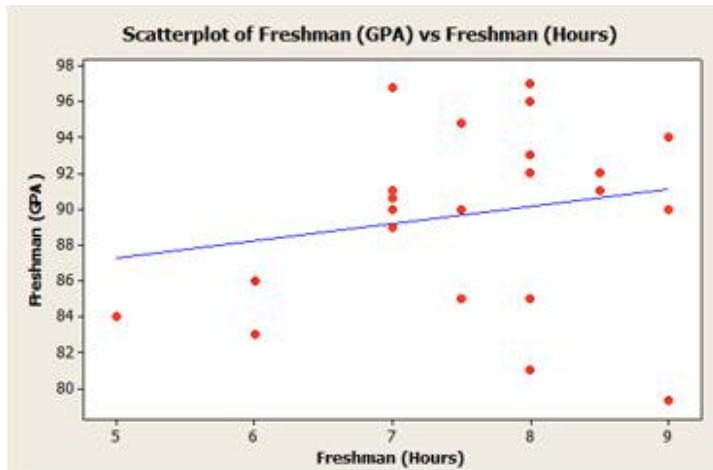
Figure 9: Scatterplot for each Class

Figure 10: Linear Regression

Freshman:
 $y=0.4405907634 x^3 + 8.151204402x^2+-46.43029169x+166.6243928.$
$r^2=0.1815780088.$

Sophomore:
$y=-0.3483974929x^3+6.163352972x^2+-34.3813295+149.1820719.$
$r^2=0.0579922685.$

Junior:
$y=0.6797516965x^3+-14.5135393x^2++102.8688261x+-151.3902294.$
$r^2=0.1078822327.$

Senior:
$y=0.00060410406x^3+-0.3858679201x^2+5.660452883x+69.31512618.$
$r^2=0.0388401488.$

Work Cited

Okano, K., Kaczmarzyk, J., Dave, N., Gabrieli, J., & Grossman, J. (2019, October 01). Sleep quality, duration, and consistency are associated with better academic performance in college students. Retrieved December 12, 2020, from https://www.nature.com/articles/s41539-019-0055-z