Student Survey

Brett Sutow

October 1st 2020

summary(studentsurvey)

## TimeReading TimeTV Happiness Gender   
## Min. :1.000 Min. :50.00 Min. :45.67 Min. :0.0000   
## 1st Qu.:2.000 1st Qu.:67.50 1st Qu.:65.34 1st Qu.:0.0000   
## Median :4.000 Median :75.00 Median :75.92 Median :1.0000   
## Mean :3.636 Mean :74.09 Mean :73.31 Mean :0.5455   
## 3rd Qu.:5.000 3rd Qu.:82.50 3rd Qu.:83.83 3rd Qu.:1.0000   
## Max. :6.000 Max. :95.00 Max. :89.52 Max. :1.0000

#Covariance of Time Reading and Time TV 1.The covariance between Time Reading and Time Watching TV is -20.35364. Covariance is a tool that is used to show a relationship between two variables. In this case it is showing a negative or inverse relationship between the two variables being looked at.

## [1] -20.36364

1. The variables are using various different measurements to describe the information. For example the they are describing the Time Reading in single integer variables like: 1,2,3,4,5,6 to describe hours. Whereas Time TV is being described in minutes varying from 50 to 95. This simple miscalculation for variables is throwing off the description. Happiness is also being described as a percentage with no true description of what it is representing and gender is described as 0 and 1 with no showing of interpretation of what that represents. Changing the variables to would create a very different representation whether that is a positive or negative correlation. When changing the Time Reading to represent hourly times it creates a covariance of -1221.818, which is very different than our -20.35364. Through the process of converting the Time Watching TV to single digit integers makes the covariance process more understandable because it creates a -.33939 covariance. Which is much better than the other two presented from the variable changes.

#Correlation Test

## TimeReading TimeTV Happiness Gender  
## TimeReading 1.00000000 -0.883067681 -0.4348663 -0.089642146  
## TimeTV -0.88306768 1.000000000 0.6365560 0.006596673  
## Happiness -0.43486633 0.636555986 1.0000000 0.157011838  
## Gender -0.08964215 0.006596673 0.1570118 1.000000000

1. To observe correlation I decided to run the cor.test for Pearson. I believe Pearson works best for this because the variables are normally distributed. It utilizes information that I believe shows the proper correlation for the test shown. This was chosen because I believe the Pearson works best for this simple data frame. It would create the best distribution for the outcomes given.
2. Between the variables time spent reading and time watching tv a negative almost linear correlation between the two variables. This shows as time reading increases the time watching tv would decrease or vice versa. This tells me that the two variables have an adverse effect on one another. Looking more in depth at the actual matrix we can see there is a positive correlation between happiness and time watching tv, and a negative correlation between happiness and time reading. The p-value shows a highly confidence significance level. That the data will fall in the specific range that is given. #Full Correlation Test

##   
## Pearson's product-moment correlation  
##   
## data: timereading and timetv  
## t = -5.6457, df = 9, p-value = 0.0003153  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.9694145 -0.6021920  
## sample estimates:  
## cor   
## -0.8830677

##   
## Pearson's product-moment correlation  
##   
## data: timereading and timetv  
## t = -5.6457, df = 9, p-value = 0.0003153  
## alternative hypothesis: true correlation is not equal to 0  
## 99 percent confidence interval:  
## -0.9801052 -0.4453124  
## sample estimates:  
## cor   
## -0.8830677

1. Part1- The Coefficient of Determination presents a .7798085 or almost 78% Part2- This information presents us with showing that there is a negative relationship between time spent reading and time spent watching tv. It also tells us that almost 78% of the data points fall on or close to the regression model line.

#Coefficient of Determination

## [1] 0.7798085

## [1] 0.7798085

1. Based off the information provided I can determine that there was a negative impact of watching tv on reading. Students who watched tv more seemed to read less. This information can be see in the corr value of -.88, the coefficient of determination of 78%, and the low pvalue of .0003. Thus, causing us to reject the null hypothesis

#Partial Correlation

## Loading required package: MASS

## estimate p.value statistic n gp Method  
## 1 -0.872945 0.0009753126 -5.061434 11 1 pearson

1. I utilized the two variables as time spent reading and tv spent watching tv, with the control being happiness. The estimate still shows a similar strong correlation. The p-value for this is slightly larger than the original correlation that was created, however it still validates the hypothesis of there being a negative almost linear relationship between time reading and time watching in regards to happiness.