Student Survey

DeBoris Leonard

10/4/2020

Use R to calculate the covariance of the Survey variables and provide an explanation of why you would use this calculation and what the results indicate

```
## [1] -1221.818
```

The reasons for using a covariance calculation is to determine the strength relationships or lack thereof between variables. The results of this calculation show a very week relationship.

Examine the Survey data variables. What measurement is being used for the variables? Explain what effect changing the measurement being used for the variables would have on the covariance calculation. Would this be a problem? Explain and provide a better alternative if needed.

The measurement types used in the Survey data are hours and minutes. By converting the hours for reding to minutes it really didn't have a major effect. The negative or weeak relationship was still there.

Choose the type of correlation test to perform, explain why you chose this test, and make a prediction if the test yields a positive or negative correlation?

I will use the Spearman correlation test. I chose this test for a couple of reasons; I have never used it and the sample size is small. I do not think there is a positive correlation between between TV and Reading but I do believe there is one between the two and happiness.

Perform a correlation analysis of:

1. All variables

2. A single correlation between two a pair of the variables

3. Repeat your correlation test in step 2 but set the confidence interval at 99%

4. Describe what the calculations in the correlation matrix suggest about the relationship between the variables. Be specific with your explanation.

ALL VARIABLES

```
##             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
```

SINGLE CORRELATION

```
## [1] 0.5662159
```

CONFIDENCE INTERVAL 99%

```
## [1] 0.5662159
```

CIr(student_df$TimeTV$, student_df$Happiness,level = 1.00)

The relationships between Happiness and TV time are positive as are the relationships between Gender and TV as well as Gender and Happiness. All other relationships are negative. The positive relationships indicates that as TV time increases so does Happiness.

Calculate the correlation coefficient and the coefficient of determination, describe what you conclude about the results.

Based on your analysis can you say that watching more TV caused students to read less? Explain.

I can find no relationship between TV time and Reading. There is not sufficient data to make the correlation.

Pick three variables and perform a partial correlation, documenting which variable you are "controlling". Explain how this changes your interpretation and explanation of the results.

```
## Warning: package 'ggm' was built under R version 3.6.3
```

```
##
## Attaching package: 'ggm'

## The following object is masked from 'package:Hmisc':
##
##     rcorr

## [1] 0.6435158
```

As I summerized before, TV Time, Happiness and Gender all have a positive correlation.