RMarkdown Assignment Template

Author Name

March 23rd 2018

## Note - Remove example code and comments before submitting assignment. Producing a professional R Markdown document is the goal.

## R Markdown

# Assignment: ASSIGNMENT 5 student survey

# Name: Burden, Joshua

# Date: 2022-05-01

ss\_df <- read.csv(url(‘<http://content.bellevue.edu/cst/dsc/520/id/resources/student-survey.csv>’)) ss\_df

head(ss\_df)

#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.

cov(ss\_dfTimeTV) # The result is negative # TimeReading and TimeTV are negatively related. cov(ss\_dfHappiness) # The result is negative # TimeReading and Happiness are negatively related. cov(ss\_dfHappiness) # The result is in positive # TimeTV and Happiness are positively related. cov(ss\_dfGender) # The result is negative # TimeReading and Gender are negatively related. cov(ss\_dfGender) # The result is positive # TimeTV and Gender are positively related. cov(ss\_dfGender) # The result is positive # Happiness and Gender are positively related.

#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.

#TimeReading - In hours #TimeTV - In minutes #Happiness - In Scale 0-100 #Gender - Binary 0 and 1

#computing Covariance between Gender and the others since its not really relevant to the research question. #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?

cov(ss\_dfTimeTV)

#Answer - Is there relationship between time spent reading and watching TV?

# both variables are related to time

# they are negatively related as a person who watches more tv they reading time will go down and vise versa

#Perform a correlation analysis of: #All variables cor(ss\_df, use = “complete.obs”, method = “pearson”)

#A single correlation between two a pair of the variables cor(ss\_dfHappiness)

#Repeat your correlation test in step 2 but set the confidence interval at 99% cor.test(ss\_dfHappiness, conf.level = .99)

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

# TimeReading and timeTV have a negative correlation

# TimeReading and Happiness have a negative correlation

# TimeTV and Happiness have a negative correlation

# All above three comparison shows that if one variable goes up

# the opposite variable will go down.

#Calculate the correlation coefficient and the coefficient of determination, describe what you conclude about the results. ss\_df1 <- ss\_df[, c(“TimeReading”, “TimeTV”, “Happiness”)] cor(ss\_df1) # TimeTV and Happiness have a moderate correlation # TimeReading and TimeTV have a strong correlation

#Based on your analysis can you say that watching more TV caused students to read less? Explain. #Answer - Based on the above correlation, there is a strong evidence that more reading leads to less watching TV

#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. install.packages(“ggm”) library(ggm) partial\_correlation <- pcor(c(“TimeReading”, “TimeTV”, “Happiness”), var(ss\_df1)) partial\_correlation^2

# Happiness is the variable that is controllable as when TV time and time reading are changed, Happiness is effected.