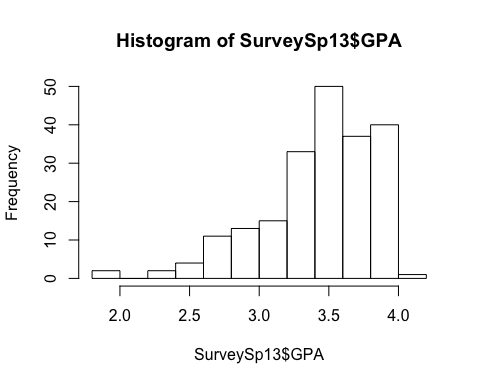
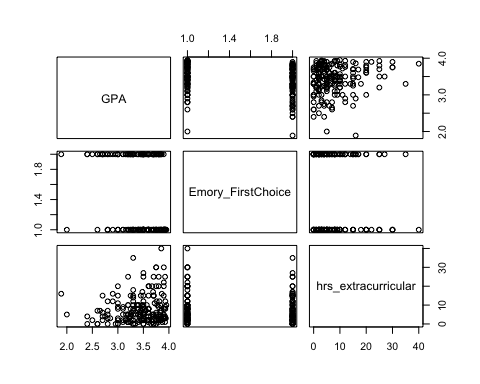
extra\_credit.R

Sat Apr 15 19:29:31 2017

#extra credit  
  
SurveySp13 <- read.csv("SurveySp13.csv", header = TRUE)  
  
hist(SurveySp13$GPA)



cleanedSurvey <-subset(SurveySp13,SurveySp13$GPA<4.0)   
  
#Examining relationship between Emory FirstChoice, hrs extracurricular and GPA   
pairs(GPA~Emory\_FirstChoice+hrs\_extracurricular,data=cleanedSurvey)



#Calculating a linear model   
m4<-lm(GPA~Emory\_FirstChoice+hrs\_extracurricular,data=cleanedSurvey)  
  
summary(m4)

##   
## Call:  
## lm(formula = GPA ~ Emory\_FirstChoice + hrs\_extracurricular, data = cleanedSurvey)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.50473 -0.18918 0.05934 0.29085 0.54800   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.470174 0.046926 73.949 < 2e-16 \*\*\*  
## Emory\_FirstChoiceYes -0.188908 0.053170 -3.553 0.000477 \*\*\*  
## hrs\_extracurricular 0.006911 0.003615 1.912 0.057363 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3748 on 197 degrees of freedom  
## Multiple R-squared: 0.08024, Adjusted R-squared: 0.0709   
## F-statistic: 8.593 on 2 and 197 DF, p-value: 0.0002642

#viewing results   
#Finding confidence interval   
confint(m4)

## 2.5 % 97.5 %  
## (Intercept) 3.3776318612 3.56271670  
## Emory\_FirstChoiceYes -0.2937626424 -0.08405350  
## hrs\_extracurricular -0.0002181709 0.01403948

#1. What is the estimated linear regression equation? ŷ = 3.47 - 0.188908(Emory\_FirstChoiceYes) + 0.006911(hrs\_extracurricular)   
#2. The intercept is 3.47 - at 0 hours of extracurriculars and 0 Emory\_FirstChoiceYes the GPA is 3.47. The confidence interval is (3.3776318612, 3.56271670) which means it is very different from 0.   
#3 The slope is -0.188908 an the confidence interval is (-0.0002181709, 0.01403948)   
#4 The slope is 0.006911 and the confidence interval is (-0.2937626424, -0.08405350)   
#5 Yes.