Lab 16 R Script

Alexander Hernandez

10/24/2022

```
library(PASWR)
```

```
## Loading required package: lattice
```

1) Study Hours and Test Score

```
students = read.table("C:\\repos\\STAT 50001\\Lab 15\\hours_scores.txt",
           header=TRUE)
```

a) Pearson Correlation Coefficient

```
cor.test(students$Score, students$hours)
```

```
##
##  Pearson's product-moment correlation
##
## data:  students$Score and students$hours
## t = 3.7578, df = 8, p-value = 0.005562
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.3407671 0.9504633
## sample estimates:
##       cor
## 0.7989697
```

b) Spearman Correlation Coefficient

```
cor.test(students$Score, students$hours,
         method="spearman", exact=FALSE)
```

```
##
##  Spearman's rank correlation rho
##
## data:  students$Score and students$hours
## S = 20.562, p-value = 0.0009053
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.875384
```

c) Kendall's Tau

```
cor.test(students$Score, students$hours,
         method="kendall", exact=FALSE)
```

```
##
##  Kendall's rank correlation tau
##
## data:  students$Score and students$hours
## z = 3.2329, p-value = 0.001225
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.8090398
```

d) For each method, test the hypothesis that the correlation is nonzero

```
# Ho: p == 0
# Ha: p != 0
cor.test(students$Score, students$hours)$p.value
```

```
## [1] 0.005561874
```

```
cor.test(students$Score, students$hours,
         method="spearman", exact=FALSE)$p.value
```

```
## [1] 0.0009053258
```

```
cor.test(students$Score, students$hours,
         method="kendall", exact=FALSE)$p.value
```

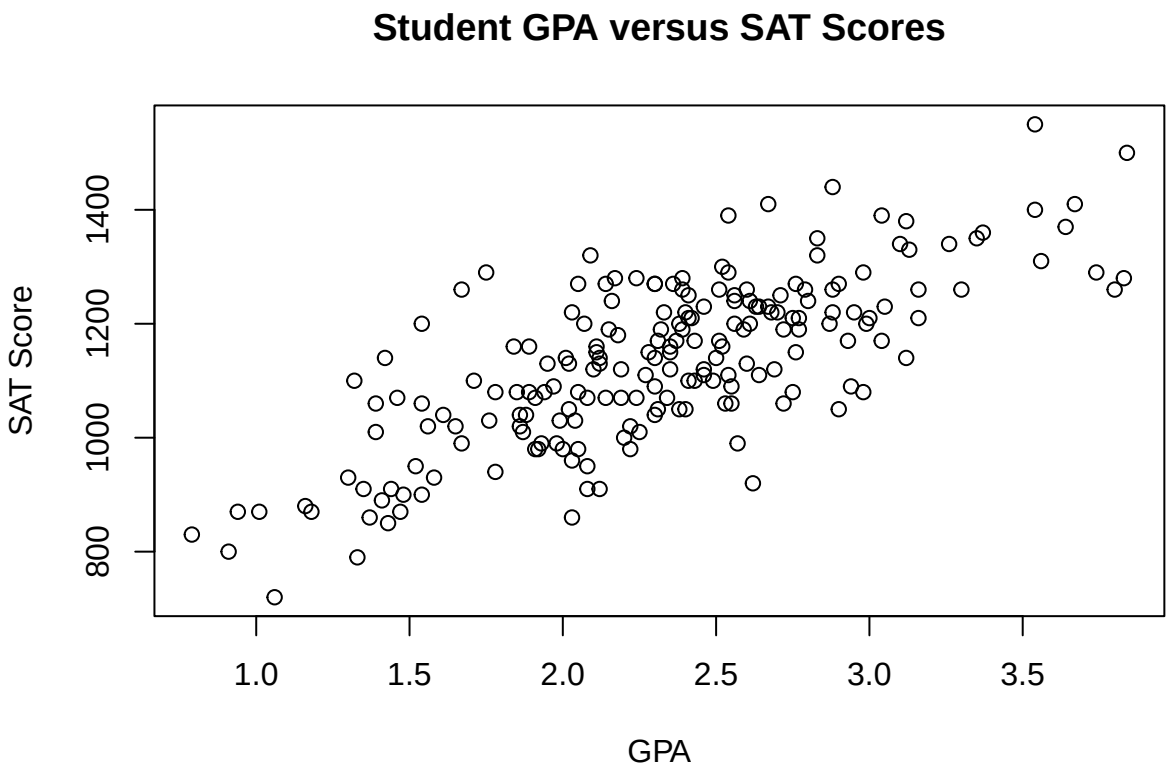
```
## [1] 0.001225424
```

```
# With p-values of:
#  0.0056 for Pearson's method,
#  0.0009 for Spearman,s method,
#  0.0012 for Kendall's method,
# we have enough evidence for all three cases to reject the null hypothesis
# and claim that the Correlation Coefficient is non-zero
# and that there is a linear correlation between student study hours and scores.
```

2) GPA and SAT scores in Grades

a) Create scatterplot of the data for GPA and SAT scores

```
plot(Grades$gpa, Grades$sat,
     main="Student GPA versus SAT Scores",
     xlab="GPA", ylab="SAT Score")
```



b) Obtain the least squares estimates for B0 and B1

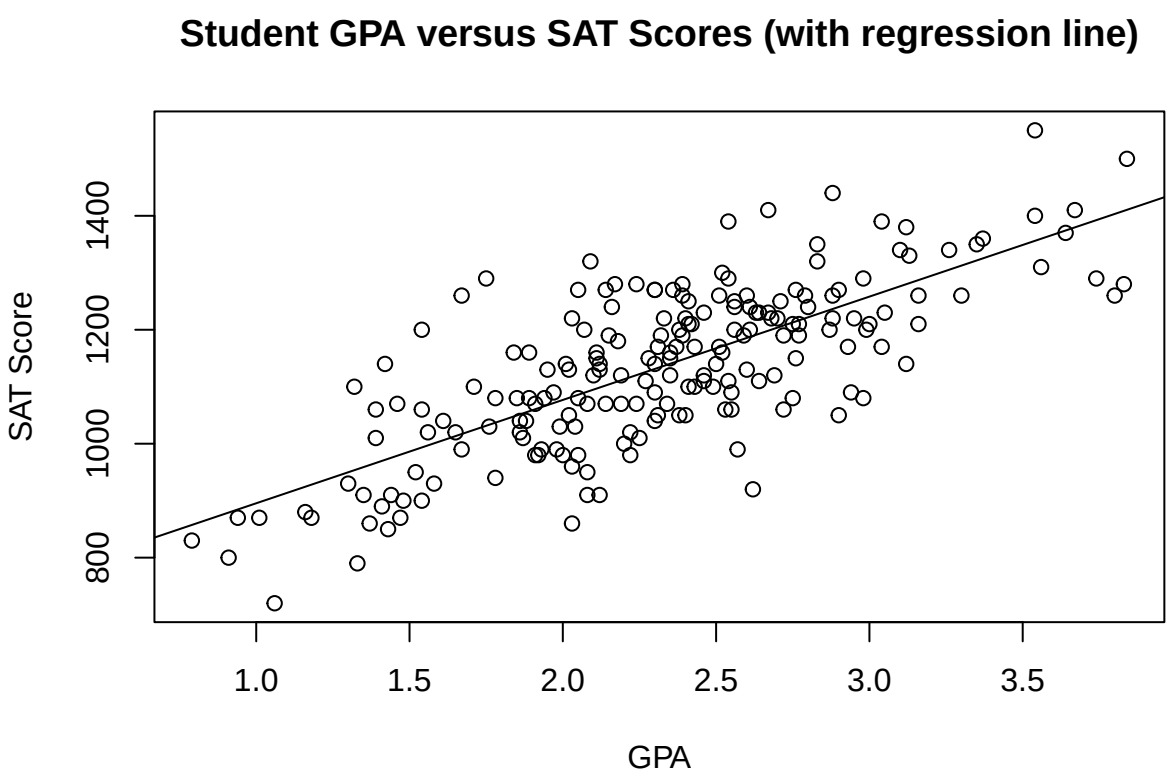
```
model = lm(Grades$sat ~ Grades$gpa)
model
```

```
##
## Call:
## lm(formula = Grades$sat ~ Grades$gpa)
##
## Coefficients:
## (Intercept)   Grades$gpa
##       714.1        181.4
```

```
# B0 = 714.1
# B1 = 181.4
```

c) Display the regression model along with the scatterplot

```
plot(Grades$gpa, Grades$sat,
     main="Student GPA versus SAT Scores (with regression line)",
     xlab="GPA", ylab="SAT Score")
abline(model)
```


Student GPA versus SAT Scores (with regression line)
SAT Score
800
1000
1200
1400
1.0
1.5
2.0
2.5
3.0
3.5
GPA