hw 7

2022-11-29

```
setwd("~/Documents/GitHub/stats100")
```

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.4.0      v purrr   0.3.5
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(stats)
library(ggplot2)
library(RColorBrewer)
```

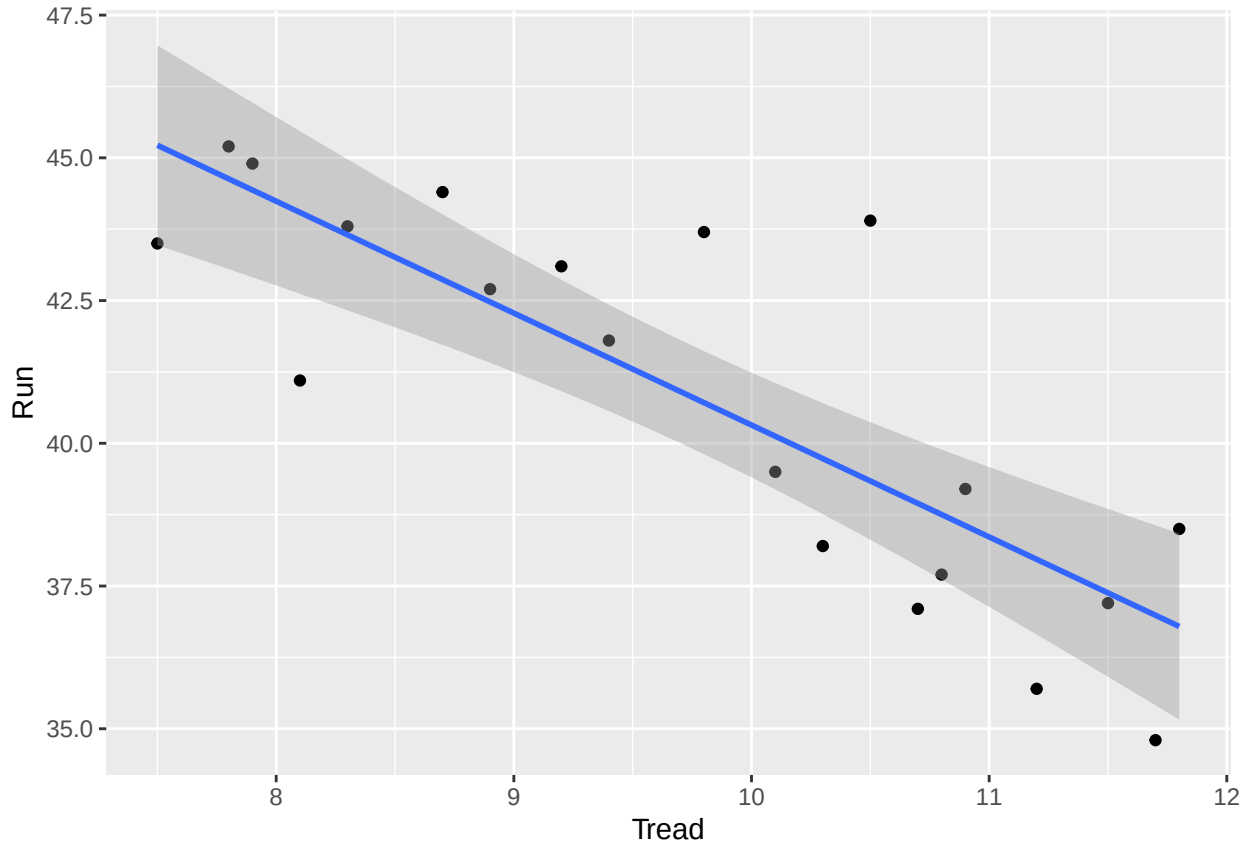
```
fitness<- read.csv("Fitness.csv")
fitness
```

```
##    Tread  Run
## 1    7.5 43.5
## 2    7.8 45.2
## 3    7.9 44.9
## 4    8.1 41.1
## 5    8.3 43.8
## 6    8.7 44.4
## 7    8.9 42.7
## 8    9.2 43.1
## 9    9.4 41.8
## 10   9.8 43.7
## 11  10.1 39.5
## 12  10.3 38.2
## 13  10.5 43.9
## 14  10.7 37.1
## 15  10.8 37.7
## 16  10.9 39.2
## 17  11.2 35.7
## 18  11.5 37.2
## 19  11.7 34.8
## 20  11.8 38.5
```

a. Find estimated regression line No outliers

```
fitness %>%
  ggplot(aes(x=Tread, y=Run))+
  geom_point()+
  geom_smooth(method = lm)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
fitlm<- lm(Run ~ Tread, data = fitness)
fitlm
```

```
##
## Call:
## lm(formula = Run ~ Tread, data = fitness)
##
## Coefficients:
## (Intercept)        Tread
##       59.92        -1.96
```

b. 95% confidence interval for slope

```
bothci<- confint(fitlm, level = 0.95)
slopeci<- bothci[2,]
slopeci
```

```
##     2.5 %    97.5 %
## -2.624957 -1.295313
```

c. Find value of Se Se = 1.921

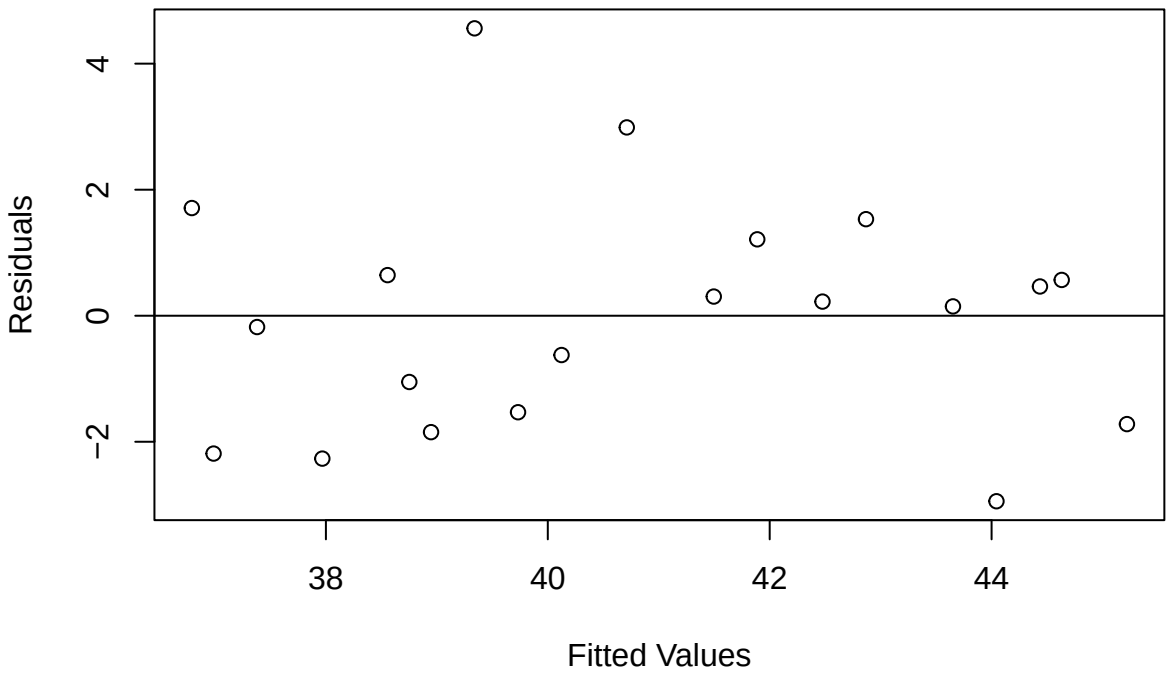
```
summary.lm(fitlm)
```

```
##
## Call:
## lm(formula = Run ~ Tread, data = fitness)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.9440 -1.5788  0.1860  0.7863  4.5603
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  59.9211     3.1166  19.226 1.90e-13 ***
## Tread        -1.9601     0.3164  -6.194 7.59e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.921 on 18 degrees of freedom
## Multiple R-squared:  0.6807, Adjusted R-squared:  0.6629
## F-statistic: 38.37 on 1 and 18 DF,  p-value: 7.589e-06
```

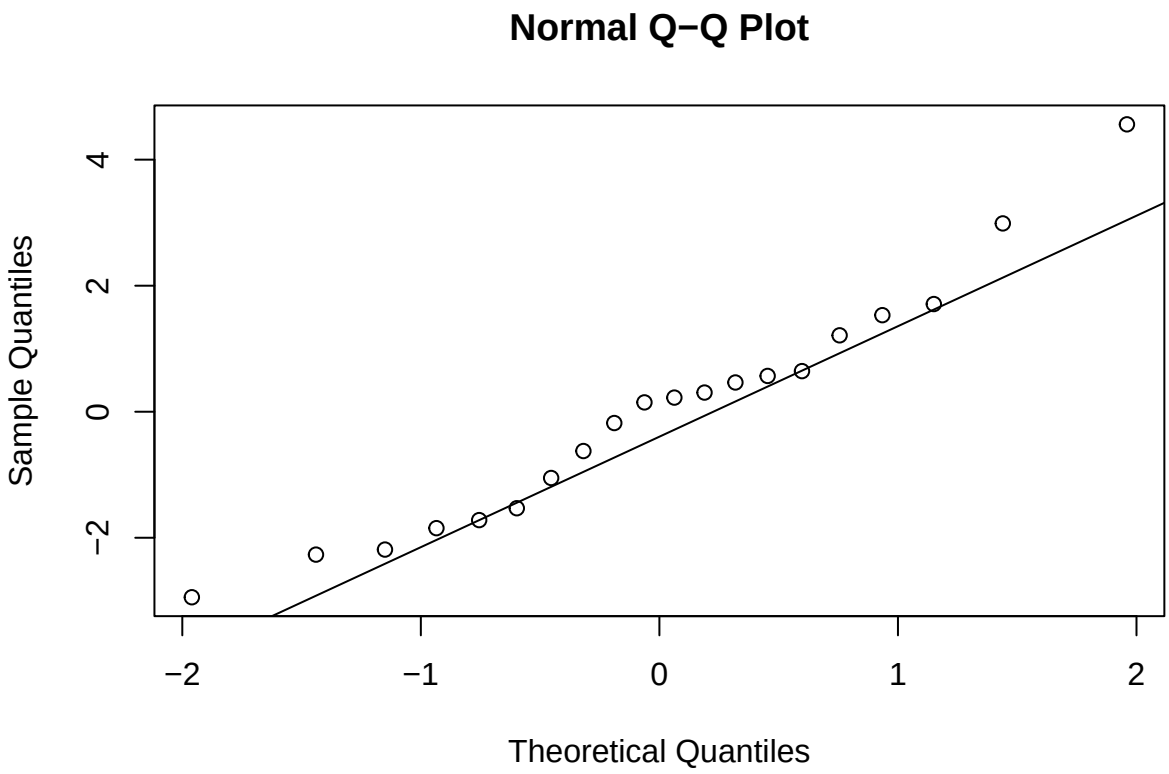
d. Find value of r squared = 0.6807

e. Does assumption of normality of error hold?

```
plot(fitlm$fitted.values, fitlm$residuals, ylab = "Residuals", xlab = "Fitted Values")
abline(h=0)
```




```
qqnorm(fitlm$residuals)
qqline(fitlm$residuals)
```




```
shapiro.test(fitlm$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  fitlm$residuals
## W = 0.96165, p-value = 0.5773
```