Assignment 5

Tucker Harris

2023-11-14

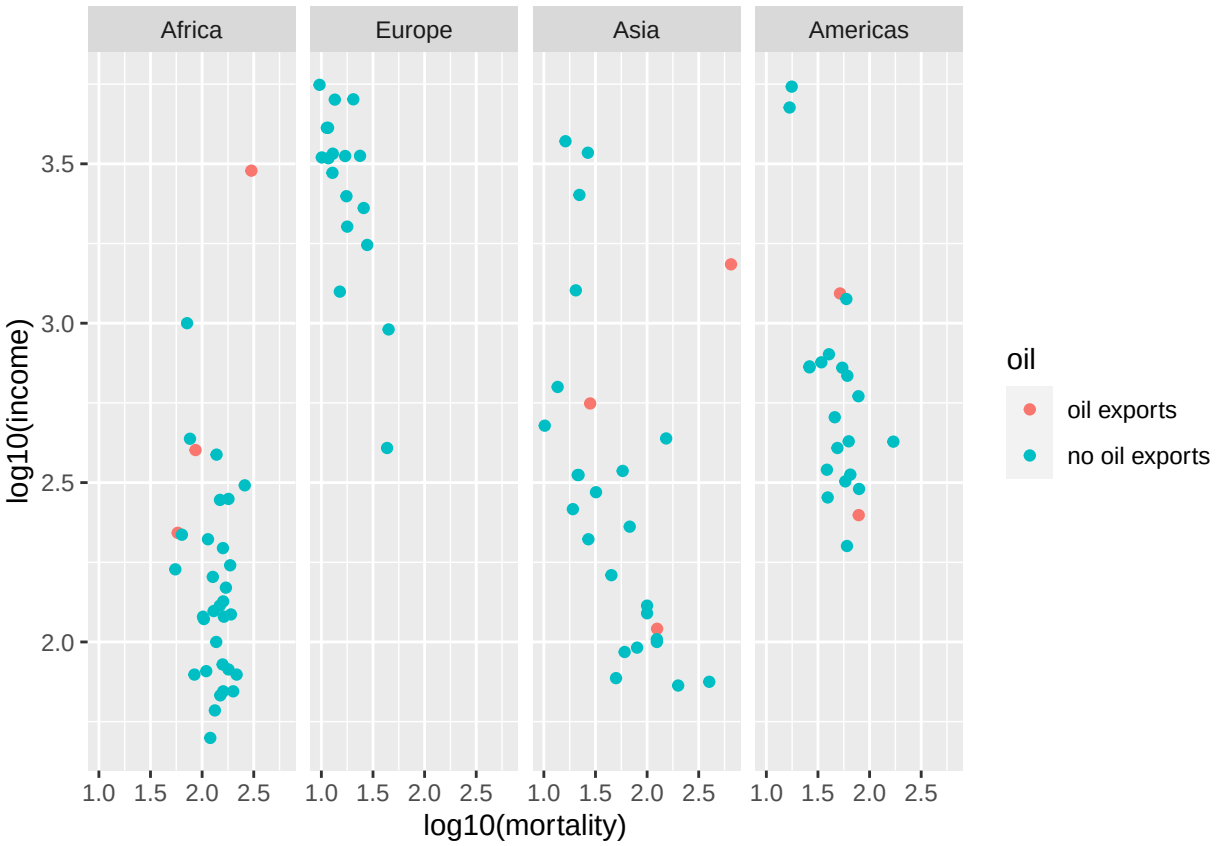
Question 1 ab)

I know we never learned the force command, but for some reason this was the only way it would let me load the dataset.

I also made this graph before looking at parts a-b. I thought those were adjustments for after making the first graph.

```
MortRate <- force(infmort)

Plot1 <- ggplot( MortRate, aes( x = log10( mortality ), y = log10(income), color = oil ) )+
  geom_point() +
  facet_grid(.~region)


Plot1
```

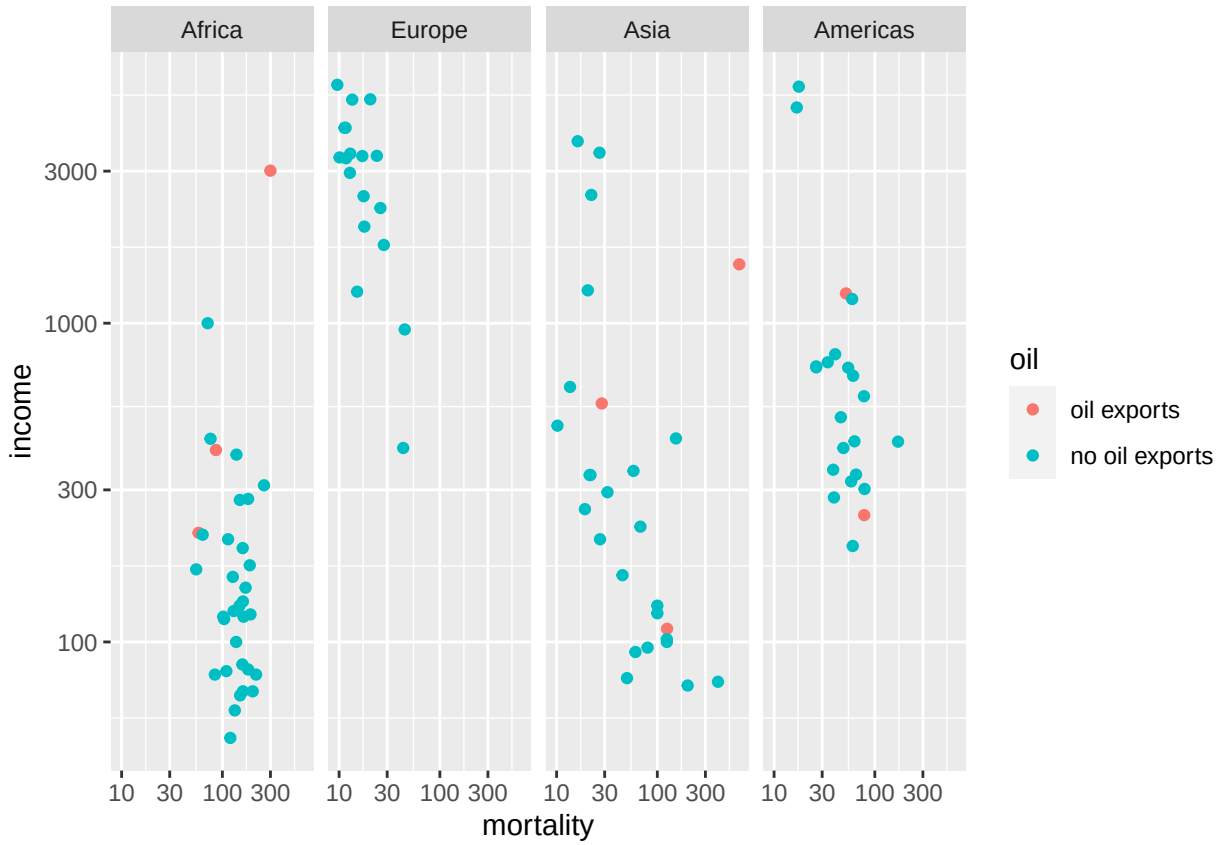
```
## Warning: Removed 4 rows containing missing values (`geom_point()`).
```



Question 1c)

```
Plot2 <- ggplot( MortRate, aes( x = mortality, y = income, color = oil ) )+
  geom_point() +
  facet_grid(.~region) +
  scale_x_log10() +
  scale_y_log10()



Plot2
```

```
## Warning: Removed 4 rows containing missing values (`geom_point()`).
```



I prefer to just use the log10 function inside the aes function.

##1d

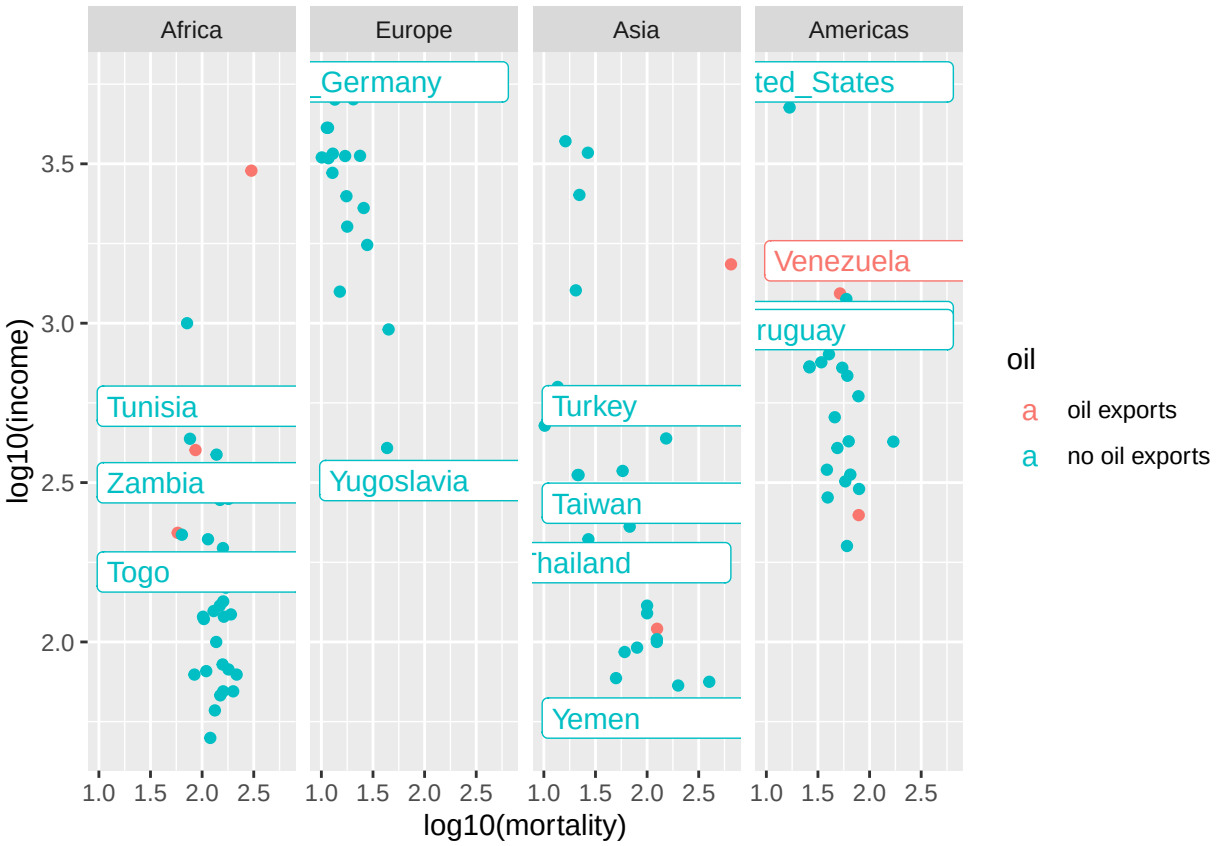
I had this graph label all countries that are below S on an alphabetical list.

```
MortRate2 <- MortRate %>% mutate(country = rownames(MortRate))

MortRate2 <- MortRate2 %>% mutate(partialCountry = str_remove_all(country, "^[A-S].*"))



Plot3 <- ggplot( MortRate2, aes( x = log10( mortality ), label = partialCountry, y = log10(income), col
  geom_point() +
  facet_grid(.~region) +
  geom_label_repel()


Plot3
```

```
## Warning: Removed 4 rows containing missing values (`geom_point()`).
```

```
## Warning: Removed 4 rows containing missing values (`geom_label_repel()`).
```

```
## Warning: ggrepel: 4 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```



Question 2a

```
data(trees)



treeModel <- lm(trees$Volume~trees$Height)

treeModel
```

```
##
## Call:
## lm(formula = trees$Volume ~ trees$Height)
##
## Coefficients:
##  (Intercept)  trees$Height
##      -87.124         1.543
```

Question 2b)

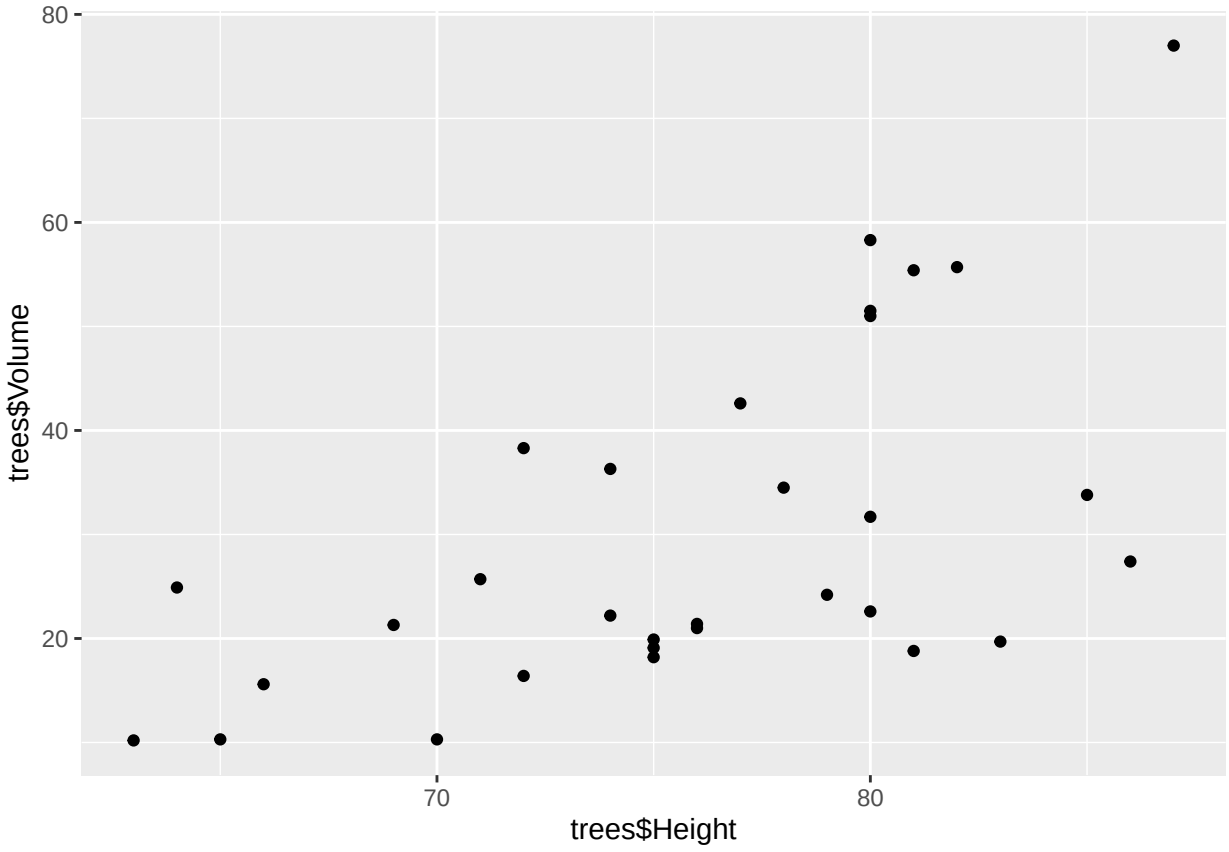
```
summary(treeModel)
```

```
##
## Call:
## lm(formula = trees$Volume ~ trees$Height)
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -21.274  -9.894  -2.894  12.068  29.852
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -87.1236    29.2731  -2.976 0.005835 **
## trees$Height   1.5433     0.3839   4.021 0.000378 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 13.4 on 29 degrees of freedom
## Multiple R-squared:  0.3579, Adjusted R-squared:  0.3358
## F-statistic: 16.16 on 1 and 29 DF,  p-value: 0.0003784
```

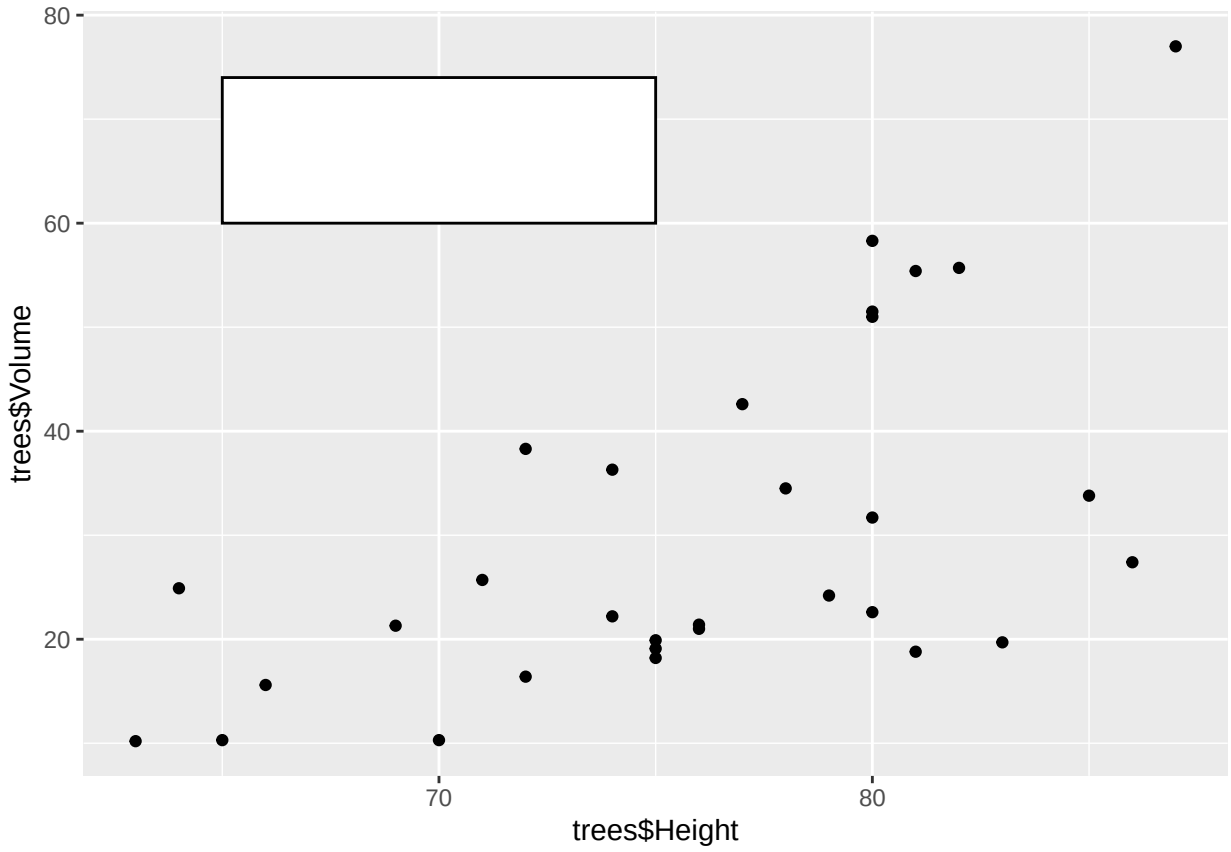
Question 2c

```
ggplot(treeModel, aes( x = trees$Height, y = trees$Volume ) ) +
  geom_point()
```



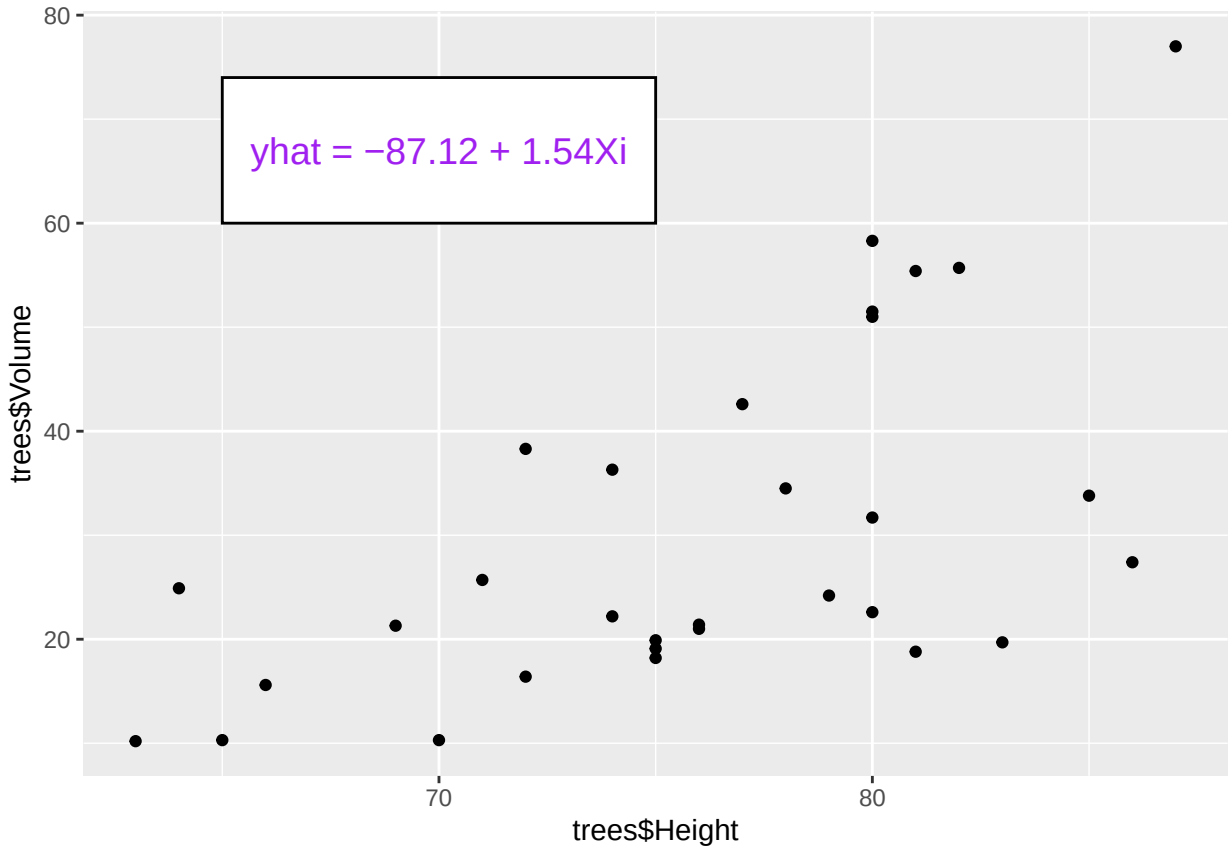
Question 2d

```
ggplot(treeModel, aes( x = trees$Height, y = trees$Volume ) ) +
  geom_point() +
  annotate('rect', xmin=65, xmax=75, ymin=60, ymax=74,
         fill='white', color='black')
```



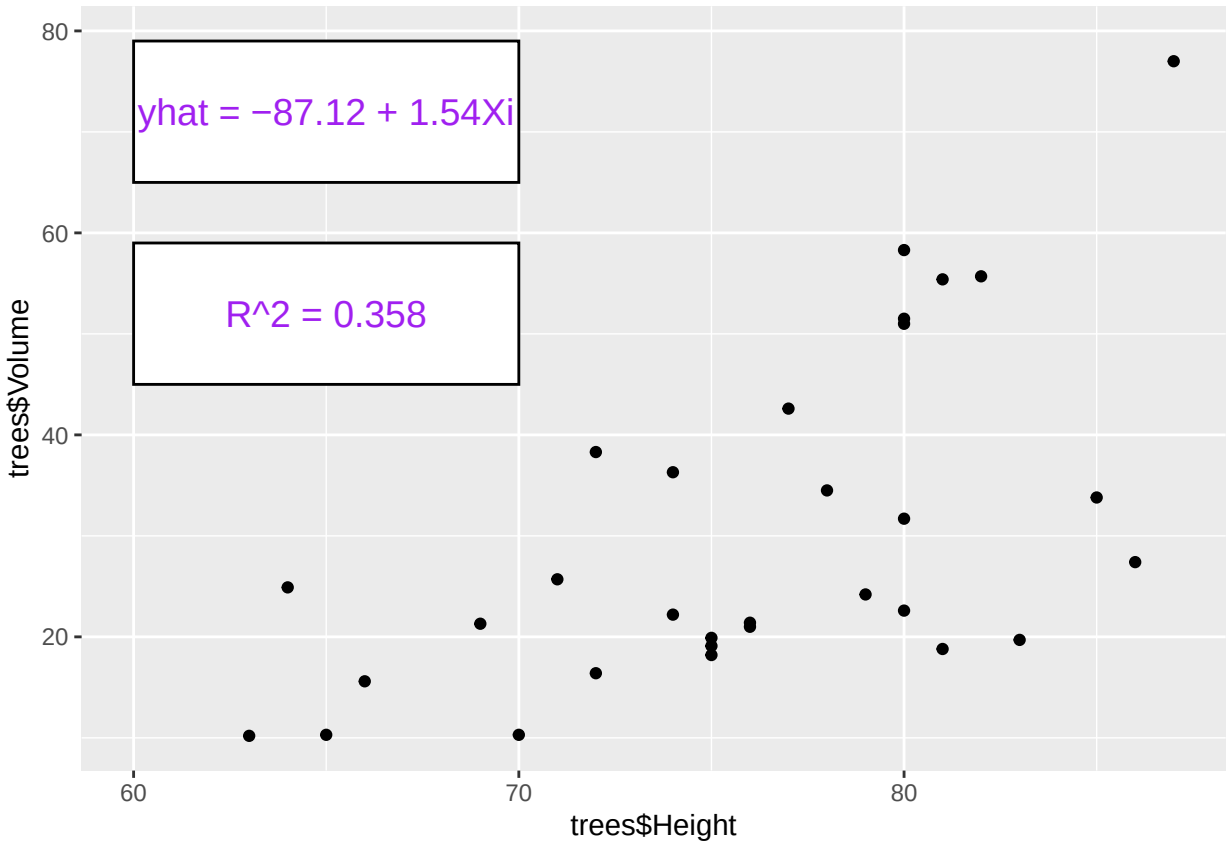
Queston 2e

```
ggplot( treeModel, aes( x = trees$Height, y = trees$Volume ) ) +
  geom_point() +
  annotate( 'rect', xmin=65, xmax=75, ymin=60, ymax=74,
         fill='white', color='black' ) +
  annotate( 'text', x = 70, y = 67, size=5, color='purple', label='yhat = -87.12 + 1.54Xi' )
```



Question 2f

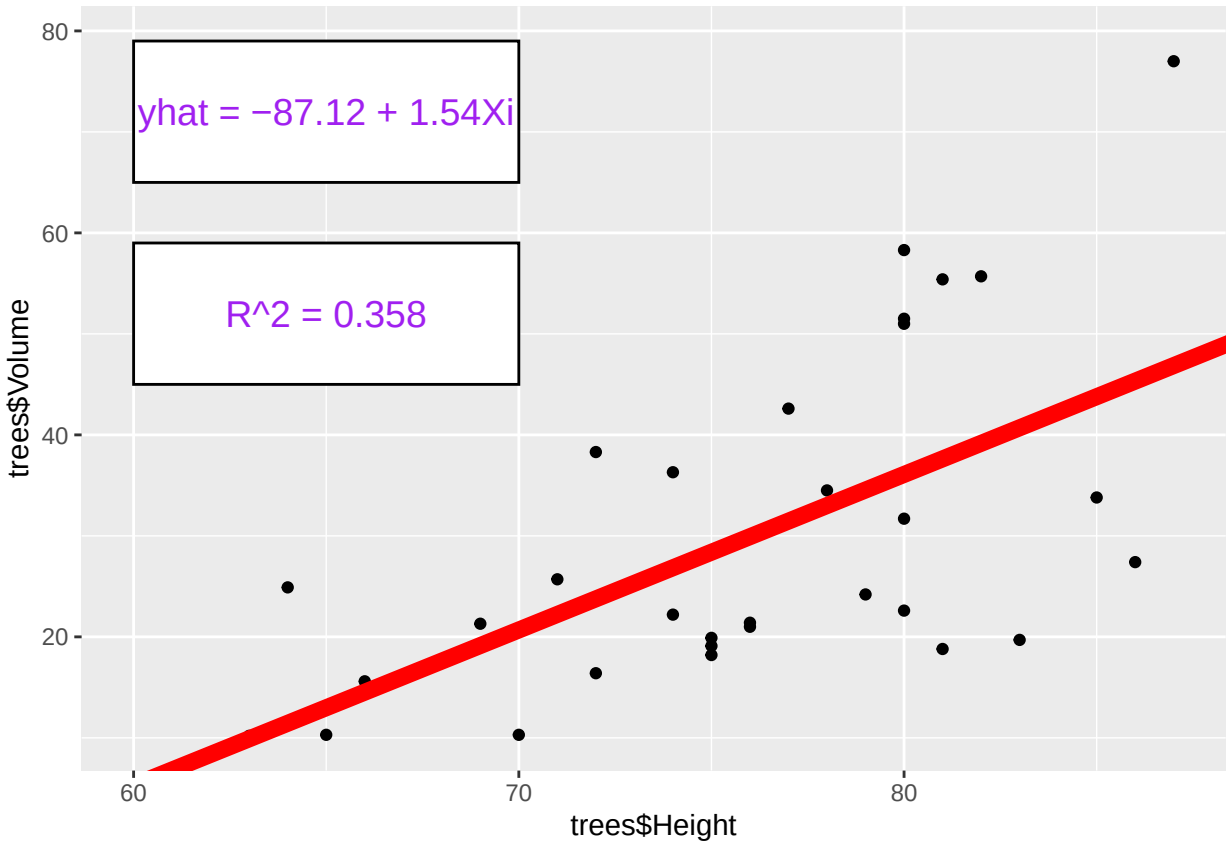
```
TwolabelTreePlot <- ggplot( treeModel, aes( x = trees$Height, y = trees$Volume ) ) +
  geom_point() +
  annotate( 'rect', xmin=60, xmax=70, ymin=65, ymax=79,
         fill='white', color='black' ) +
  annotate( 'text', x = 65, y = 72, size=5, color='purple', label='yhat = -87.12 + 1.54Xi' ) +
  annotate( 'rect', xmin=60, xmax=70, ymin=45, ymax=59,
            fill='white', color='black' ) +
  annotate('text', x = 65, y = 52, size = 5, color ='purple', label='R^2 = 0.358')

TwolabelTreePlot
```



Question 2g

```
TwolabelTreePlot + geom_abline(intercept = -87.12, slope = 1.54, color = 'red', size = 3)
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```


yhat = −87.12 + 1.54Xi
R^2 = 0.358
trees$Volume
trees$Height
80
60
40
20
60
70
80