Example Chapter 6

Problem 45

get the data

```
require(magrittr)
```

```
## Loading required package: magrittr
```

```
require(ggplot2)
```

```
## Loading required package: ggplot2
```

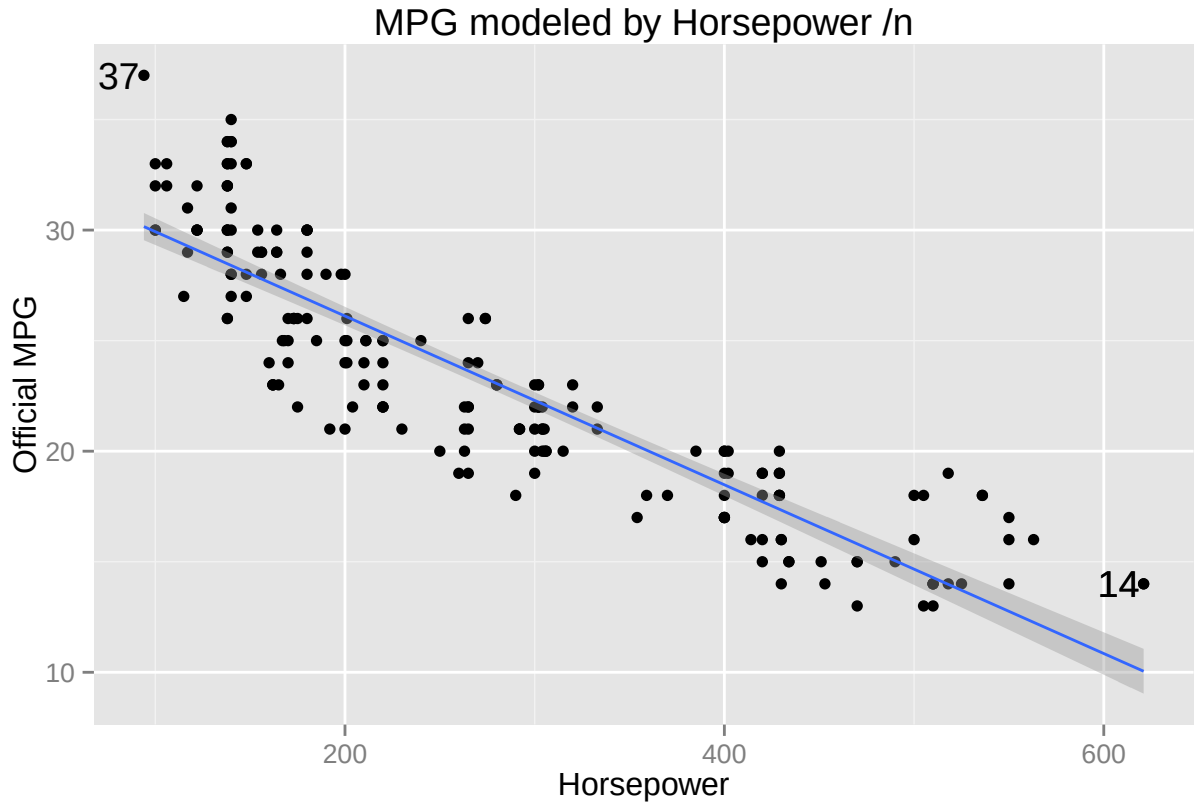
```
cars <- read.table("C:\\Users\\Jonathan\\Google Drive\\Stats Camp\\Stine&Foster\\Data by Chapter\\Chapt
```

Part a)

We need to make a scatterplot using the combined or "Official Mpg" and the rated horsepower. We need to figure out which of the columns should be the response variable and which should be the explanitory variable.

Looking at our options. The best combination would be to put horsepower as the explanitory variable and Mpg as the responce varible. Horsepower does a better job explaining mpg than mpg does explaining horsepower.

```
cars %>%
  ggplot(aes(x = Horsepower, y = CombinedMPG)) +
    geom_point() +
    geom_smooth(method = "lm") +
    geom_text(aes(label=ifelse((CombinedMPG < median(CombinedMPG) - 1.5 * IQR(CombinedMPG) | CombinedMP
  geom_text(aes(label=ifelse((Horsepower < median(Horsepower) - 1.5 * IQR(Horsepower) | Horsepower > me
    labs(x = "Horsepower", y = "Official MPG", title = "MPG modeled by Horsepower /n")
```



Examing the graph you can see that they have a strong negitive correlation. This means that the points on the graph closely resemble a straight line it has a strong correlation. It's negitive because the line it follows have a negitive slope. The cars to the far right are manufactured by Mercedez Benz. The car in the top left is Scion iQ.

Part C)

To find the correlation we'll use the cor function. Remember that correlations go from - 1 to 1.

```
cor(cars$CombinedMPG, cars$Horsepower)
```

```
## [1] -0.8918865
```

Our value is -.892, or a strong negitive.

Part D)

This correlation provides us with a good summary of the data. As horsepower increases the miles per gallon goes down.

Part E

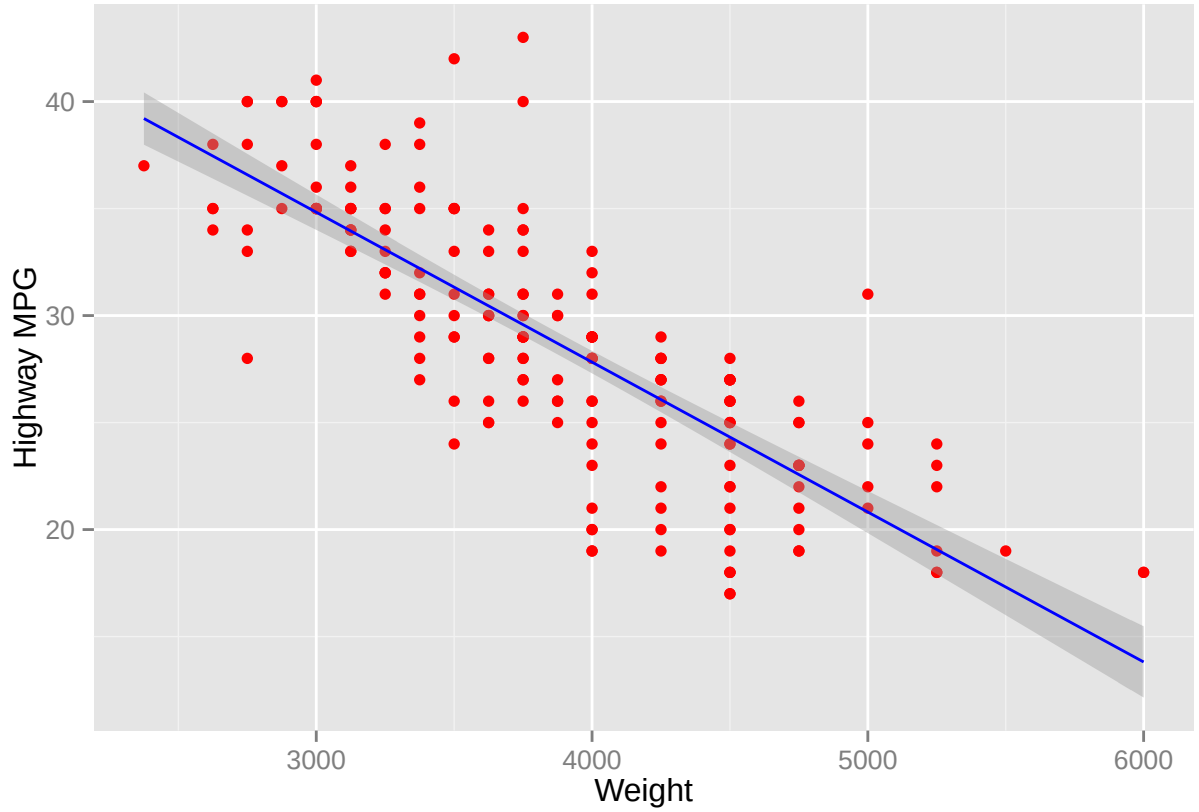
Using the line we've superimposed in the graph follow 200 up until we run into the correlation line. Now move over to the MPG axis it should put us at about 26 miles per gallon.

Problem 46

Part A

For problem 46 we'll use weight as our explanitory variable and highway mileage as our responce variable. Highway mileage is a variable that we'd want to predict. Weight is also something we have under control.

```
ggplot(data = cars, aes(x = Weight, y = HighwayMPG)) +
  geom_point(color = "red") +
  geom_smooth(method = "lm", color = "blue") +
  labs(x = "Weight", y = "Highway MPG")
```



Part B

From the plot we can see that The correlation is fairly strong and negitive. Some outliers to pay attention to may include the point at Weight = 2500, miles per gallon = 26. (Mazda MX-5) The cluster of three less greater than 3500 and greater than 39 are also of interest. (Volkswagen GOLF and Passat–two types). Another one would be the point at 5000 lbs thatis at about 31 miles per gallon. (Mercedes S).

Part C

```
cor(cars$HighwayMPG, cars$Weight)
```

```
## [1] -0.8010125
```

Part D

refering back to the scatter plot, if we find the point where the correlation line and vertical line at 4000 lbs intersect it would predict the highway mpg to be 28 mpg.