

ECN 3972 – Lab 3

Start by doing three things:

1. Import HTWT1.xls into R. You should see 20 observations of three variables: observation number, height (inches above 5') which is labeled as "X," and weight (in pounds) which is labeled as "Y."
2. Type `library(mosaic)`
3. Type `library(psych)`

Simple Linear Regression

Let's use the data set on height and weight to see if we can use R to generate a linear regression to help us estimate weight, given a person's height (as measured as inches beyond 5').

Now, we can create a linear regression model. The command to use is "lm" which stands for "linear model."

The dependent variable will be "Y" (which is weight) and the independent variable will be "X" (which is height in inches above 5').

We're also going to tell R where the data are located in our command line.

```
model1 = lm(Y~X, data=HTWT1)
```

The model is then created, but needs to be displayed.

```
model1
##
## Call:
## lm(formula = Y ~ X, data = HTWT1)
##
## Coefficients:
## (Intercept)          X
##    103.397         6.377
```

Another way to display the coefficients is to use:

```
coefficients(model1)
```

```
## (Intercept)      X
## 103.397083      6.377093
```

R can also calculate the predicted value of our dependent variable, y-hat:

```
print(fitted(model1))

##      1      2      3      4      5      6      7      8
## 135.2825 160.7909 186.2993 179.9222 167.1680 173.5451 154.4138 160.7909
##      9     10     11     12     13     14     15     16
## 167.1680 179.9222 173.5451 160.7909 167.1680 179.9222 154.4138 160.7909
##     17     18     19     20
## 167.1680 199.0535 186.2993 173.5451
```

Now, we can calculate the residual by hand:

```
res1 <- HTWT1$Y - fitted(model1)
res1

##      1      2      3      4      5
## 4.71745003 -3.79092382 18.70070232 18.07779579 -5.16801729
##      6      7      8      9     10
## 0.45488925 -4.41383036 4.20907618 2.83198271 0.07779579
##     11     12     13     14     15
## -3.54511075 1.20907618 -2.16801729 0.07779579 5.58616964
##     16     17     18     19     20
## -5.79092382 -2.16801729 -9.05348460 -1.29929768 -18.54511075
```

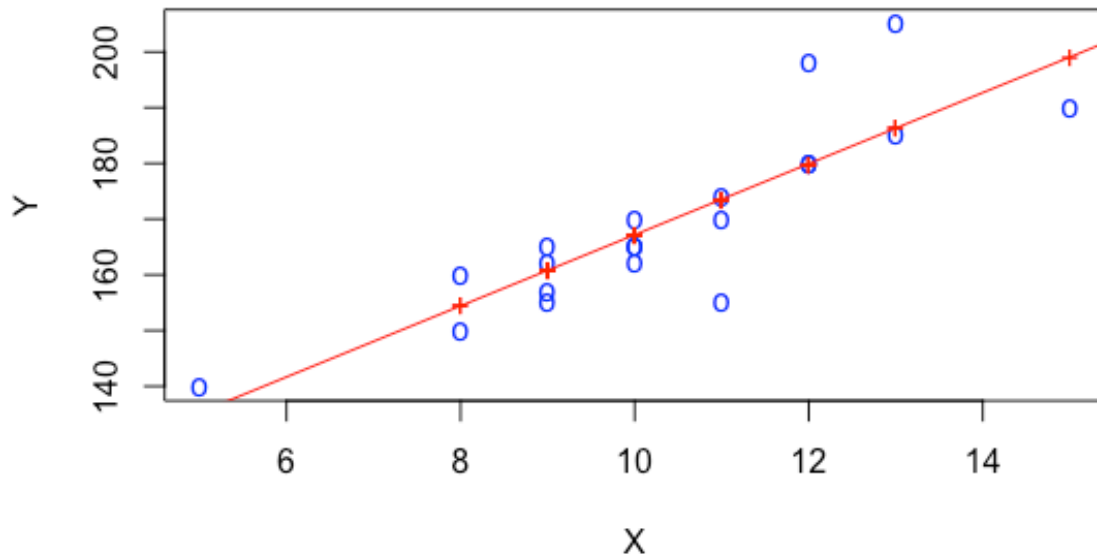
Or, we can use the residual function:

```
res2 <- resid(model1)
res2

##      1      2      3      4      5
## 4.71745003 -3.79092382 18.70070232 18.07779579 -5.16801729
##      6      7      8      9     10
## 0.45488925 -4.41383036 4.20907618 2.83198271 0.07779579
##     11     12     13     14     15
## -3.54511075 1.20907618 -2.16801729 0.07779579 5.58616964
##     16     17     18     19     20
## -5.79092382 -2.16801729 -9.05348460 -1.29929768 -18.54511075
```

Let's plot the actual values and then the predicted values:

```
plot(Y~X, data=HTWT1, col = "blue", pch = "o")
points(fitted(model1)~X, data = HTWT1, col = "red", pch = "+")
abline(model1, col="red")
```



Here's the picture of the regression line (**the red + marks**) and the actual points taken from your textbook. (The only difference is the starting point of the vertical axis.)

