

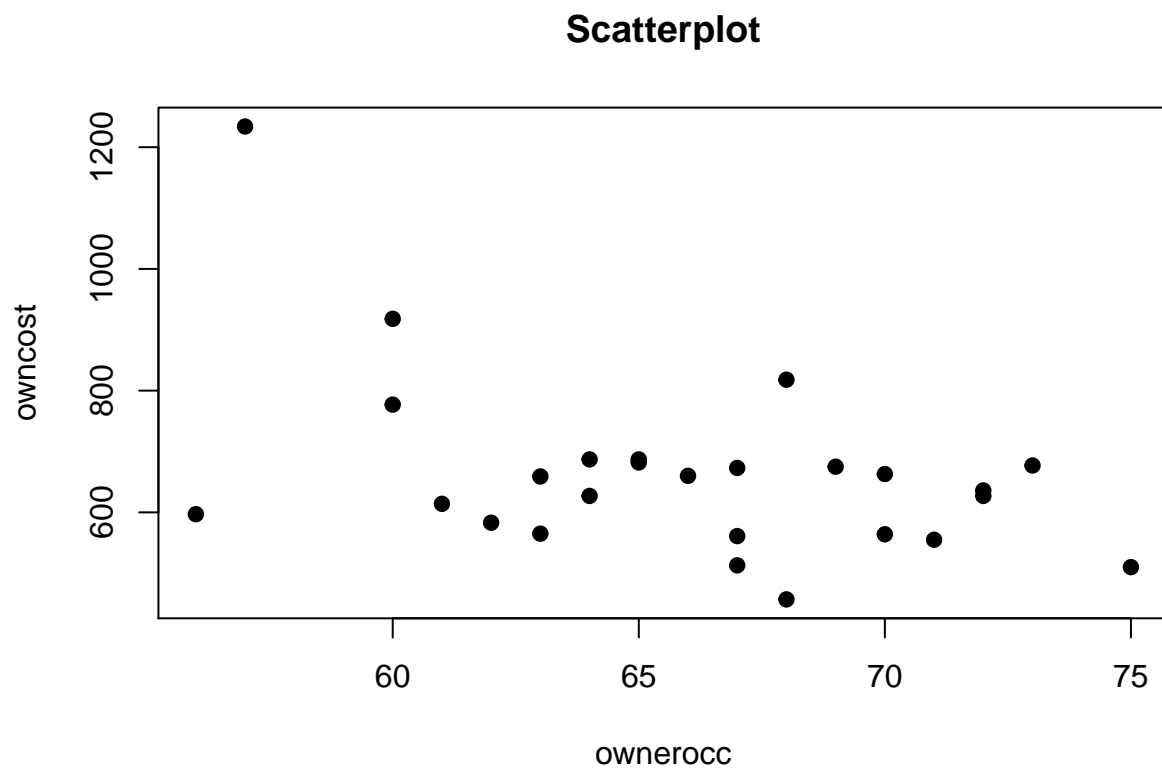
# Linear Regression Project

Luc Barengnien

August 20, 2020

###Part 1: Making a scatterplot

```
setwd("~/Downloads")
load("ch05q19.rdata")
X<-ch05q19$owncost
Y<-ch05q19$ownerocc
plot(Y, X, main="Scatterplot", xlab="ownerocc", ylab="owncost", pch=19)
```



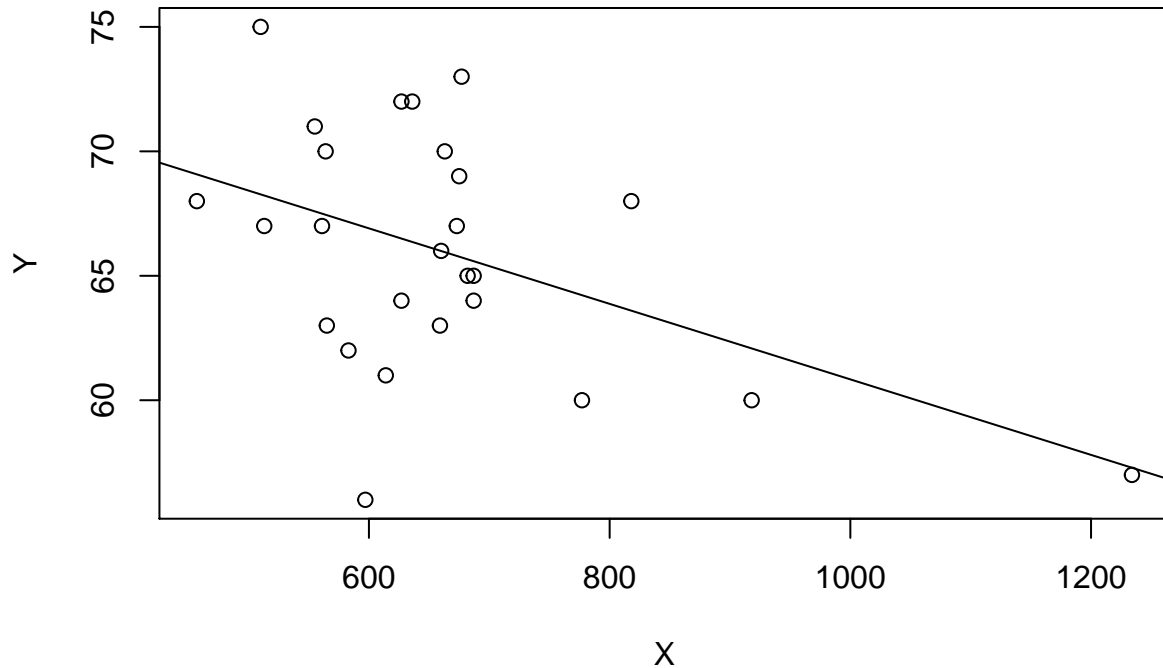
###Part 2: Finding the regression equation.

```
Equation<-lm(Y~X)
```

So the equation is:  $\hat{owncost} = -0.01517 \times ownerocc + 76.00764$ .

###Part 3: Remaking the scatterplot and adding the regression line.

```
plot(Y~X)
abline(Equation)
```



####Part 4: Test to see if the slope is significantly different than 0.

```
summary(Equation)
```

```
##
## Call:
## lm(formula = Y ~ X)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.9516  -2.3930  -0.5421   3.3663   7.2619
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  76.007640   3.949795  19.243 4.29e-16 ***
## X            -0.015169   0.005819  -2.607  0.0155 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.417 on 24 degrees of freedom
## Multiple R-squared:  0.2207, Adjusted R-squared:  0.1882
## F-statistic: 6.796 on 1 and 24 DF, p-value: 0.01546
```

The relevant p-value given in this output is 0.0155, and is significant at the 5% level. In other words, this data leads us to reject the null hypothesis that the slope of the equation predicting weight from age is 0. More succinctly, the slope of this regression line is significantly different from 0.

###Part 5: Finding the 95% confidence interval for the slope.

```
confint(Equation,level = 0.95)
```

```
##                2.5 %        97.5 %  
## (Intercept) 67.85566408 84.159615708  
## X          -0.02717863 -0.003159796
```