NOTE: This is the R session used to find the regression equation, and some plots for the Highway sign data. There is also information for later use on finding confidence intervals and prediction intervals. Download the file from the course webpage before you begin, and store it in the appropriate directory. Lines that begin with # are comments and not executed in R.
For Lecture 6, scroll down to highlighted lines.

```
> #Read in the data.
> #sep="\t" shows that the columns are separated with a tab.
> #header=F says there is no beginning line with variable names.
> #col.names provides names for the two columns.
> Highway<-read.table("HighwaySign.txt", header=F, sep="\t",
col.names=c("Age","Distance"))

> #Make sure it worked by printing out the first 6 lines:
> head(Highway)
  Age Distance
1  18      510
2  20      590
3  22      560
4  23      510
5  23      460
6  25      490

> #Create the regression model. Call it "HWModel"
> HWModel<-lm(Distance~Age,data=Highway)

> #See a summary of the model, including coefficients, etc.
> summary(HWModel)

Call:
lm(formula = Distance ~ Age, data = Highway)

Residuals:
    Min      1Q  Median      3Q     Max
-78.231 -41.710   7.646  33.552 108.831

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 576.6819    23.4709  24.570  < 2e-16 ***
Age          -3.0068     0.4243  -7.086 1.04e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 49.76 on 28 degrees of freedom
Multiple R-squared: 0.642,      Adjusted R-squared: 0.6292
F-statistic: 50.21 on 1 and 28 DF,  p-value: 1.041e-07
```

```
> #Create standardized and studentized residuals
> Highway$StandardResids <- rstandard(HWModel)
> Highway$StudentResids <- rstudent(HWModel)

> #Plot Age versus distance and add least squares line
> plot(Highway$Age,Highway$Distance)
> abline(HWModel)
> #Get four residual plots all together
> plot(HWModel)
> #Get a normal probability plot of the residuals and add a line
> qqnorm(HWModel$resid)
> qqline(HWModel$resid)

> #The remainder of this information is for later.
> #Get 95% confidence intervals for the intercept and slope.
> confint(HWModel)
                2.5 %     97.5 %
(Intercept) 528.604017 624.759857
Age          -3.876051  -2.137620

> #Get 95% confidence interval for the mean of Y's at X = 30.
> #se.fit=T includes value of the standard error in the output
> predict(HWModel, list(Age=30),se.fit=T,interval="confidence")
$fit
       fit      lwr      upr
1 486.4769 460.4091 512.5447

$se.fit
[1] 12.72588

$df
[1] 28

$residual.scale
[1] 49.76158

> #Get 95% prediction interval for an individual Y when X = 30.
> predict(HWModel, list(Age=30),se.fit=T,interval="p")
$fit
       fit      lwr      upr
1 486.4769 381.2644 591.6893

$se.fit
[1] 12.72588

$df
[1] 28
```

```
$residual.scale
[1] 49.76158

> #Get confidence interval and prediction interval for Age = 50
> predict(HWModel, list(Age=50),se.fit=T,interval="confidence")
$fit
       fit      lwr      upr
1 426.3402 407.7097 444.9706

$se.fit
[1] 9.095085

$df
[1] 28

$residual.scale
[1] 49.76158

> predict(HWModel, list(Age=50),se.fit=T,interval="predict")
$fit
       fit      lwr      upr
1 426.3402 322.7196 529.9607

$se.fit
[1] 9.095085

$df
[1] 28

$residual.scale
[1] 49.76158

> #Quit R
> q()
```