MATH 423/533 – SIMPLE LINEAR REGRESSION EXAMPLE IN R

The following code in R fits the simple linear regression model

$$\mathbb{E}[Y_i|x_{i1}] = \beta_0 + \beta_1 x_{i1} \qquad i = 1, 2, \ldots, n$$

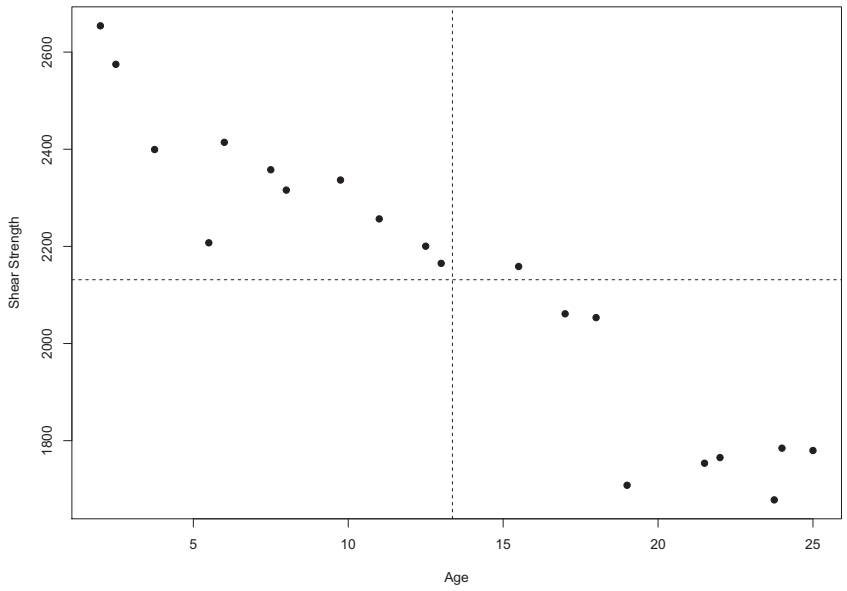
where $n = 20$ and

- x_{i1} is Age, the single continuous predictor;
- Y_i is the Shear Strength outcome random variable;
- y_i is the observed version of Y_i.

Simple Linear Regression: Plot Data

```
> data.source<-
+"http://www.math.mcgill.ca/dstephens/Regression/Data/2-1-RocketProp.csv"
>
> RocketProp<-read.csv(file=data.source)
>
> names(RocketProp)<-c('i','Strength','Age')
>
> x<-RocketProp$Age
> y<-RocketProp$Strength
> plot(x,y,pch=19,xlab='Age',ylab='Shear Strength')
> (xmean<-mean(x))
[1] 13.3625
> (ymean<-mean(y))
[1] 2131.358
>
> abline(v=xmean,h=ymean,lty=2)
> dev.copy2pdf(file='RocketPropPlot.pdf',paper='USr',width=11,height=9)
```

The resulting plot of the data is given below:



```
> ##############################################
> #Fit the simple linear regression using lm
> 
> fit.RP<-lm(y ~ x)
> summary(fit.RP)

Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max 
-215.98  -50.68   28.74   66.61  106.76 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 2627.822     44.184   59.48  < 2e-16 ***
x            -37.154      2.889  -12.86 1.64e-10 ***
---
Signif. codes:  0  ***  0.001  **  0.01  *  0.05  .  0.1     1

Residual standard error: 96.11 on 18 degrees of freedom
Multiple R-squared:  0.9018,    Adjusted R-squared:  0.8964 
F-statistic: 165.4 on 1 and 18 DF,  p-value: 1.643e-10

> 
> coef(fit.RP)
(Intercept)           x 
 2627.82236   -37.15359 
> abline(coef(fit.RP),col='red')
> title('Line of best fit for Rocket Propulsion Data')
> 
> ##############################################
> #Using the matrix formulae
> n<-length(x)
> X<-cbind(rep(1,n),x)
> 
> (XtX<-t(X)%*%X)
                  x
   20.00  267.250
x 267.25 4677.688
> 
> Xty<-t(X) %*% y
> 
> (beta.hat<-solve(XtX,Xty))
        [,1]
  2627.82236
x  -37.15359
> 
```