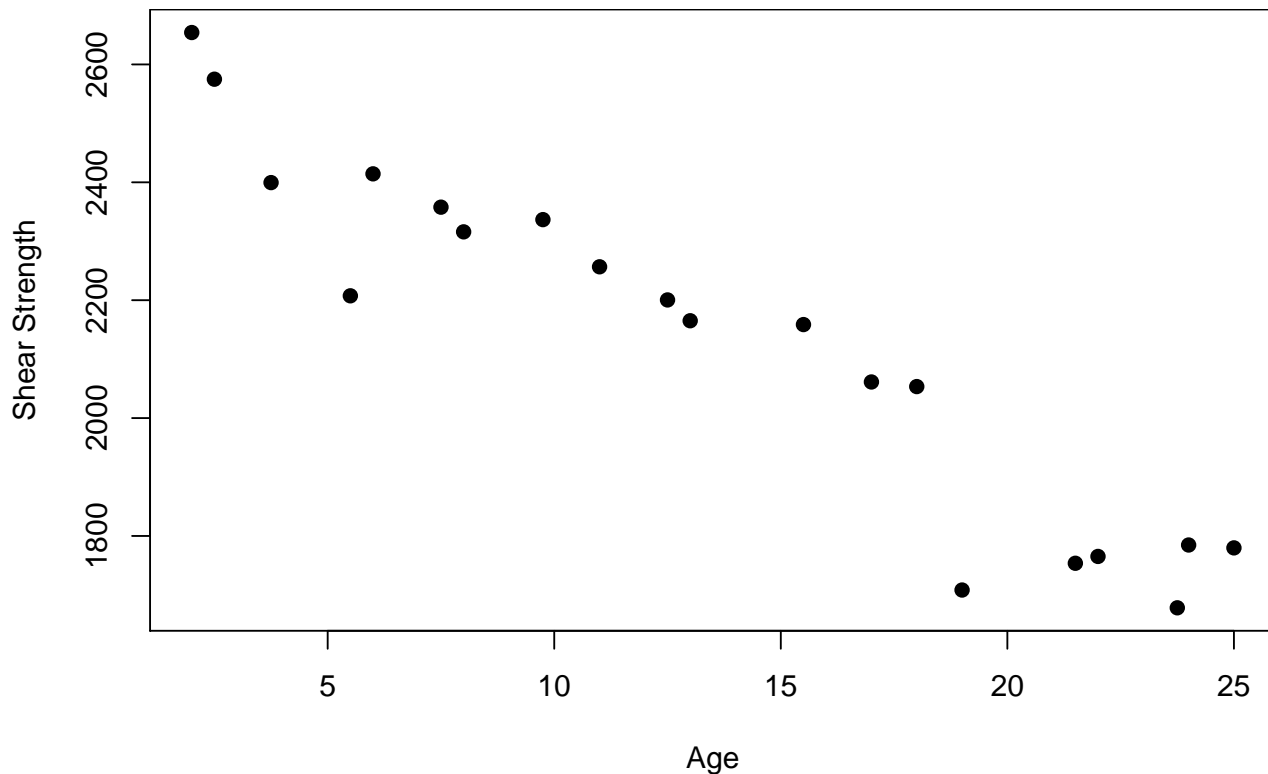# PREDICTION FOR THE ROCKET PROPULSION DATA

```
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
n<-length(x)
par(mar=c(4,4,0,1))
plot(x,y,pch=19,xlab='Age',ylab='Shear Strength')
```



```
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
```
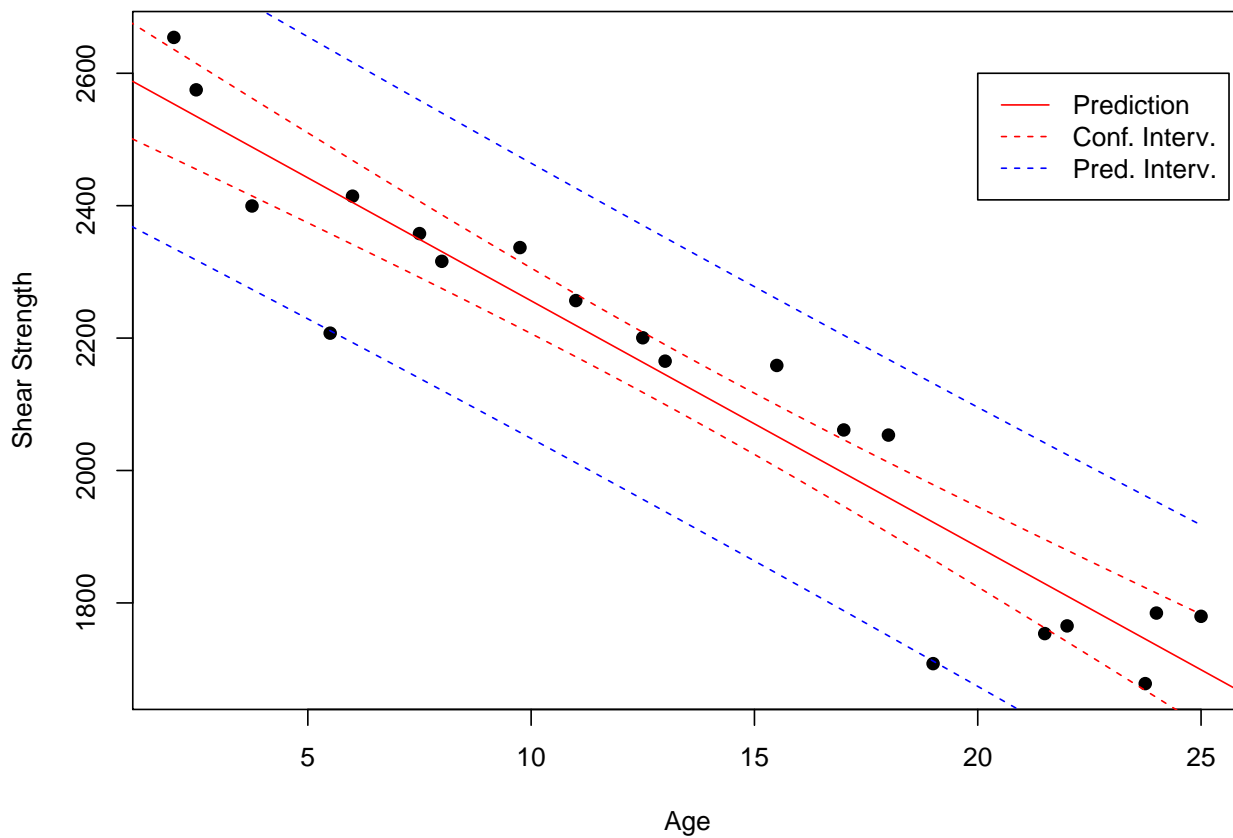
Predictions on a grid of new x values using the `predict` function:

```
xnew<-seq(0,25,by=0.1)

#Confidence interval
ynew.interval<-predict(fit.RP,newdata=data.frame(x=xnew),interval='confidence')
plot(x,y,pch=19,xlab='Age',ylab='Shear Strength')
abline(coef(fit.RP),col='red')
title('Prediction for Rocket Propulsion Data')
lines(xnew,ynew.interval[,2],lty=2,col='red')
lines(xnew,ynew.interval[,3],lty=2,col='red')

#Prediction interval
yonew.interval<-predict(fit.RP,newdata=data.frame(x=xnew),interval='prediction')
lines(xnew,yonew.interval[,2],lty=2,col='blue')
lines(xnew,yonew.interval[,3],lty=2,col='blue')
legend(20,2600,c('Prediction','Conf. Interv.','Pred. Interv.'),col=c('red','red','blue'),lty=c(1,2,2))
```

Prediction for Rocket Propulsion Data

Using the matrix formulae:

```
X<-cbind(rep(1,n),x)
(XtX<-t(X)%*%X)

+              x
+    20.00  267.250
+ x 267.25 4677.688

Xty<-t(X) %*% y
```

```r
beta.hat<-solve(XtX,Xty)
Hat.matrix<-X %*% solve(XtX) %*% t(X)

SS.Res<-(t(y)%*%(diag(1,n)-Hat.matrix)%*%y)[1,1]
MS.Res<-SS.Res/(n-2)
sigma.hat<-sqrt(MS.Res)

nnew<-length(xnew)
Xnew<-cbind(rep(1,nnew),xnew)

Ynew<-Xnew%*%beta.hat[,1]
Hnew<-diag(Xnew %*% solve(XtX) %*% t(Xnew))
Var.Ynew<-sigma.hat^2*Hnew
Var.Yonew<-sigma.hat^2*(1+Hnew)

#Look up the quantile
alpha<-0.05
q.val<-qt(1-alpha/2,n-2)

#Intervals
Conf.vals<-cbind(Ynew-q.val*sqrt(Var.Ynew),Ynew+q.val*sqrt(Var.Ynew))
Pred.vals<-cbind(Ynew-q.val*sqrt(Var.Yonew),Ynew+q.val*sqrt(Var.Yonew))

#Check computation - should be zero
sum(ynew.interval[,2:3]-Conf.vals)

+ [1] -7.260041e-10

sum(yonew.interval[,2:3]-Pred.vals)

+ [1] -7.264589e-10
```