

# Cherry Trees

```
# Cherry tree data is already a dataframe in R
```

```
help(trees)
```

```
## starting httpd help server ... done
```

```
str(trees)
```

```
## 'data.frame':   31 obs. of  3 variables:
## $ Girth : num  8.3 8.6 8.8 10.5 10.7 10.8 11 11 11.1 11.2 ...
## $ Height: num  70 65 63 72 81 83 66 75 80 75 ...
## $ Volume: num  10.3 10.3 10.2 16.4 18.8 19.7 15.6 18.2 22.6 19.9 ...
```

```
head(trees)
```

```
##   Girth Height Volume
## 1   8.3     70   10.3
## 2   8.6     65   10.3
## 3   8.8     63   10.2
## 4  10.5     72   16.4
## 5  10.7     81   18.8
## 6  10.8     83   19.7
```

```
#tail(trees)
```

```
#x <- trees$Girth (this is mislabeled and is really the tree diameter in inches)
#y <- trees$Volume (timber amount in cubic feet)
```

```
Diameter <- trees$Girth
trees1 <- cbind(trees, Diameter)
```

```
model <- lm(Volume ~ Diameter, data = trees1)
summary(model)
```

```
##
## Call:
## lm(formula = Volume ~ Diameter, data = trees1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.065 -3.107  0.152  3.495  9.587
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -36.9435     3.3651  -10.98 7.62e-12 ***
## Diameter       5.0659     0.2474   20.48 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 4.252 on 29 degrees of freedom
## Multiple R-squared:  0.9353, Adjusted R-squared:  0.9331
## F-statistic: 419.4 on 1 and 29 DF,  p-value: < 2.2e-16
```

```
#str(model)
```

```
anova(model)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: Volume
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Diameter   1 7581.8   7581.8   419.36 < 2.2e-16 ***
```

```
## Residuals 29  524.3     18.1
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
confint(model)
```

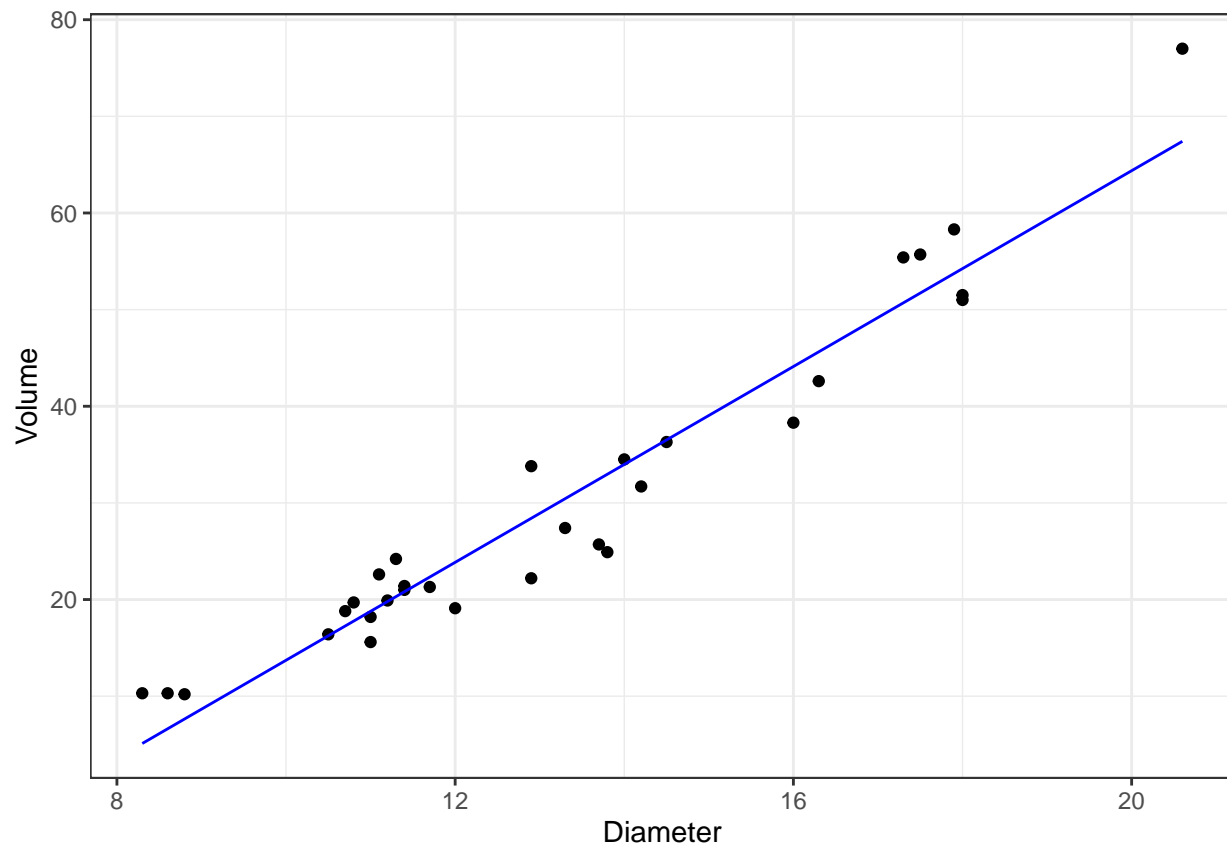
```
##           2.5 %    97.5 %
```

```
## (Intercept) -43.825953 -30.060965
```

```
## Diameter      4.559914   5.571799
```

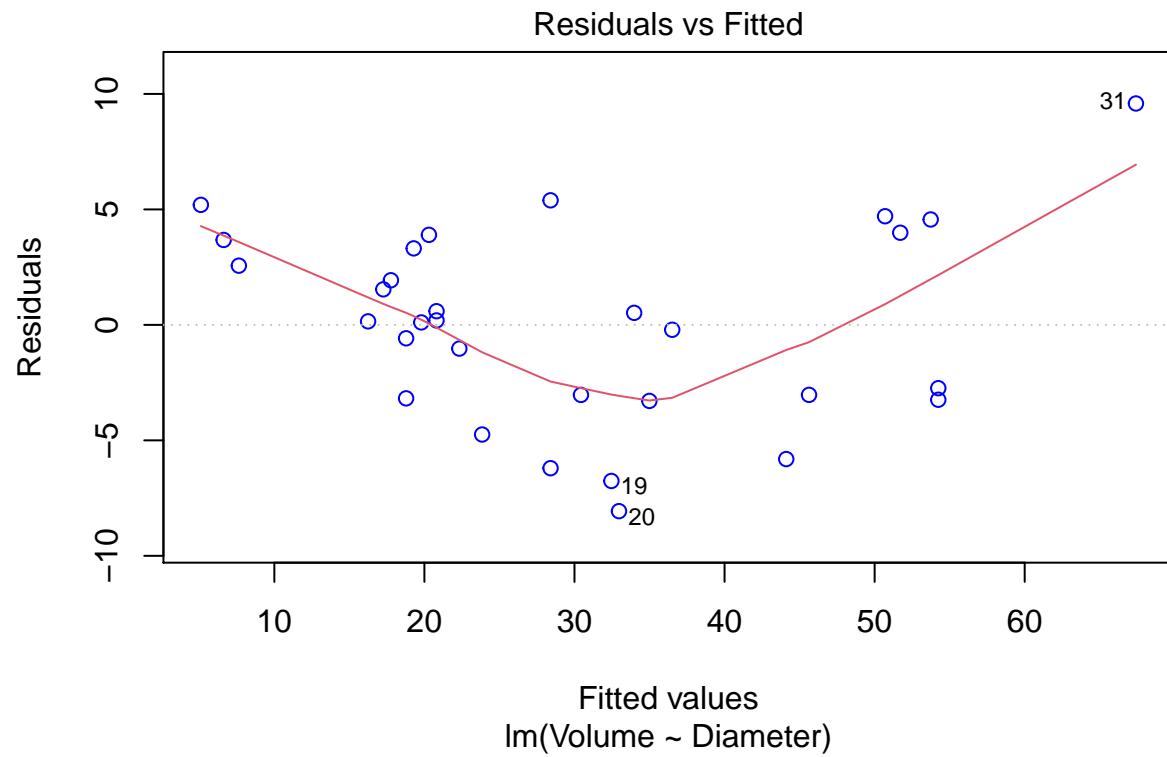
```
library(ggplot2)
```

```
figure <- ggplot(trees1) + geom_point( aes(x=Diameter, y=Volume) ) + geom_line(aes(x=Diameter, y=model))
figure
```

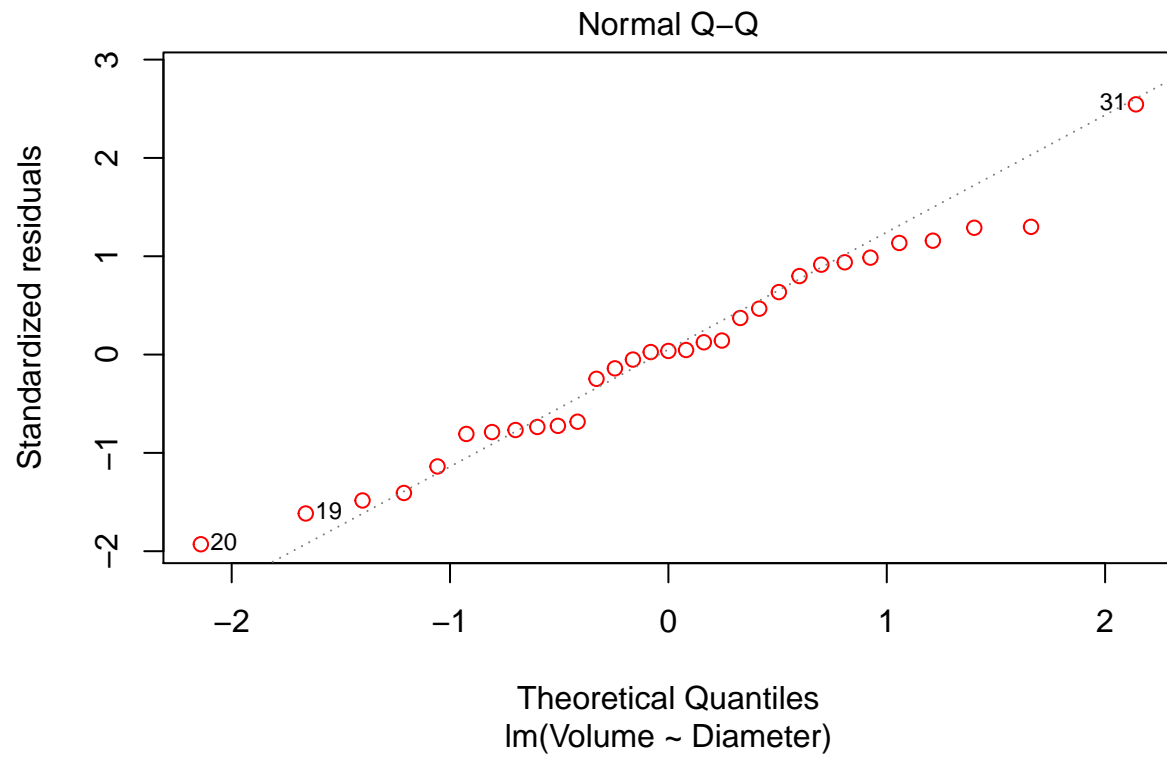


```
# residual diagnostics
```

```
plot(model, which=1, col=c("blue")) # Residuals vs Fitted Plot
```



```
plot(model, which=2, col=c("red")) # Q-Q Plot
```



```
shapiro.test(residuals(model))
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  residuals(model)  
## W = 0.97889, p-value = 0.7811
```