Final_2.R

VinaRahimian

Mon Dec 05 18:55:53 2016

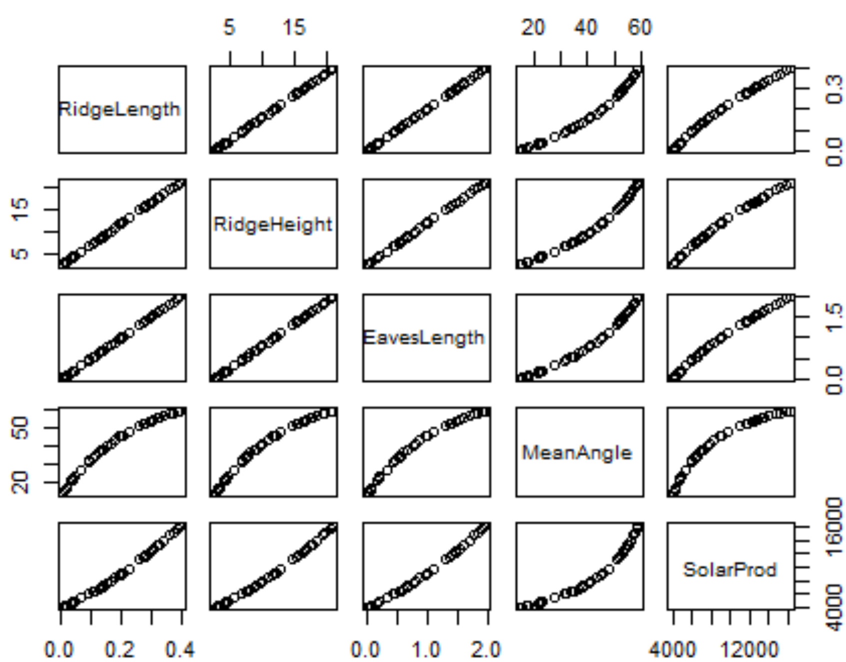
```
#Importing all excel files first
energy0 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation data/0.csv")
energy1 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation data/1.csv")
energy2 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation data/2.csv")
energy3 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation data/3.csv")
energy4 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation data/4.csv")
energy5 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation data/5.csv")
energy6 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation data/6.csv")
energy7 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation data/7.csv")
energy8 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation data/8.csv")
energy9 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation data/9.csv")
energy10 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/10.csv")
energy11 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/11.csv")
energy12 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/12.csv")
energy13 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/13.csv")
energy14 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/14.csv")
energy15 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/15.csv")
energy16 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/16.csv")
```

```
energy17 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/17.csv")
energy18 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/18.csv")
energy19 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/19.csv")
energy20 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/20.csv")
energy21 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/21.csv")
energy22 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/22.csv")
energy23 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/23.csv")
energy24 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/24.csv")
energy25 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/25.csv")
energy26 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/26.csv")
energy27 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/27.csv")
energy28 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/28.csv")
energy29 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/29.csv")
energy30 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/30.csv")
energy31 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/31.csv")
energy32 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/32.csv")
energy33 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
```

```
data/33.csv")
energy34 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/34.csv")
energy35 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/35.csv")
energy36 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/36.csv")
energy37 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/37.csv")
energy38 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/38.csv")
energy39 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/39.csv")
energy40 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/40.csv")
energy41 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/41.csv")
energy42 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/42.csv")
energy43 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/43.csv")
energy44 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/44.csv")
energy45 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/45.csv")
energy46 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/46.csv")
energy47 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/47.csv")
energy48 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/48.csv")
energy49 <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Solar generation
data/49.csv")
angle <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
```

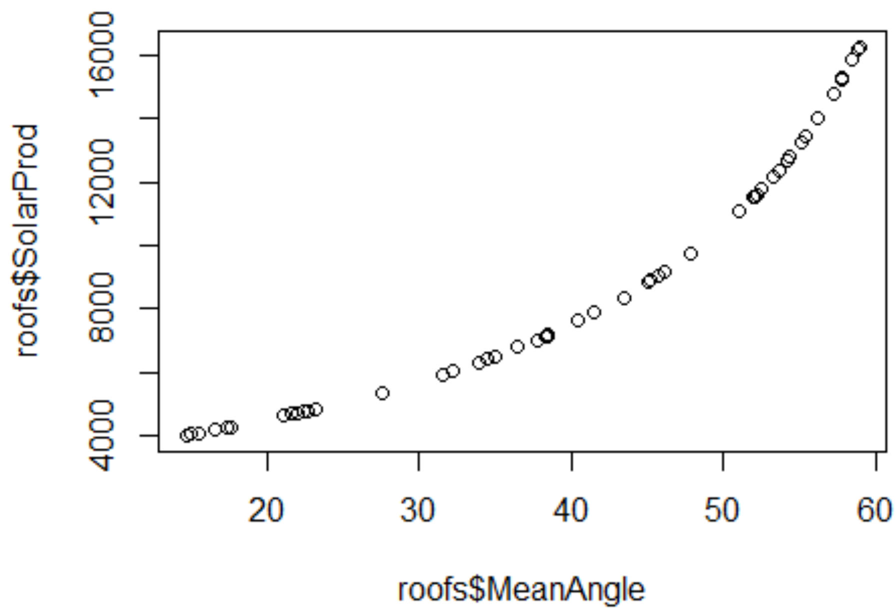
```r
State/PhD/Microgrid Case Studies/2500 R Midtown/Average Roof Angle.csv")
roofs <- read.csv("C:/Users/VinaRahimian/Google Drive/Penn
State/PhD/Microgrid Case Studies/2500 R Midtown/Roof parameters.csv")

#Now adding columns from energy and angle data to roof data.
#The reason is to accumulate all necessary data into one file so I'd be able
#to conduct analysis on it
roofs[,4] <- angle[,1] #adding the roof angle to the main dataframe
#About the produced energy data, I'm calculating the average annual solar
energy production, converting it from joules to kwH, and adding it to the 5th
column of my main dataframe:
roofs[1,5] <- mean(energy0[,2])*0.0000002778
roofs[2,5] <- mean(energy1[,2])*0.0000002778
roofs[3,5] <- mean(energy2[,2])*0.0000002778
roofs[4,5] <- mean(energy3[,2])*0.0000002778
roofs[5,5] <- mean(energy4[,2])*0.0000002778
roofs[6,5] <- mean(energy5[,2])*0.0000002778
roofs[7,5] <- mean(energy6[,2])*0.0000002778
roofs[8,5] <- mean(energy7[,2])*0.0000002778
roofs[9,5] <- mean(energy8[,2])*0.0000002778
roofs[10,5] <- mean(energy9[,2])*0.0000002778
roofs[11,5] <- mean(energy10[,2])*0.0000002778
roofs[12,5] <- mean(energy11[,2])*0.0000002778
roofs[13,5] <- mean(energy12[,2])*0.0000002778
roofs[14,5] <- mean(energy13[,2])*0.0000002778
roofs[15,5] <- mean(energy14[,2])*0.0000002778
roofs[16,5] <- mean(energy15[,2])*0.0000002778
roofs[17,5] <- mean(energy16[,2])*0.0000002778
roofs[18,5] <- mean(energy17[,2])*0.0000002778
roofs[19,5] <- mean(energy18[,2])*0.0000002778
roofs[20,5] <- mean(energy19[,2])*0.0000002778
roofs[21,5] <- mean(energy20[,2])*0.0000002778
roofs[22,5] <- mean(energy21[,2])*0.0000002778
roofs[23,5] <- mean(energy22[,2])*0.0000002778
roofs[24,5] <- mean(energy23[,2])*0.0000002778
roofs[25,5] <- mean(energy24[,2])*0.0000002778
roofs[26,5] <- mean(energy25[,2])*0.0000002778
roofs[27,5] <- mean(energy26[,2])*0.0000002778
roofs[28,5] <- mean(energy27[,2])*0.0000002778
roofs[29,5] <- mean(energy28[,2])*0.0000002778
roofs[30,5] <- mean(energy29[,2])*0.0000002778
roofs[31,5] <- mean(energy30[,2])*0.0000002778
roofs[32,5] <- mean(energy31[,2])*0.0000002778
roofs[33,5] <- mean(energy32[,2])*0.0000002778
roofs[34,5] <- mean(energy33[,2])*0.0000002778
roofs[35,5] <- mean(energy34[,2])*0.0000002778
roofs[36,5] <- mean(energy35[,2])*0.0000002778
roofs[37,5] <- mean(energy36[,2])*0.0000002778
roofs[38,5] <- mean(energy37[,2])*0.0000002778
roofs[39,5] <- mean(energy38[,2])*0.0000002778
```

```
roofs[40,5] <- mean(energy39[,2])*0.0000002778
roofs[41,5] <- mean(energy40[,2])*0.0000002778
roofs[42,5] <- mean(energy41[,2])*0.0000002778
roofs[43,5] <- mean(energy42[,2])*0.0000002778
roofs[44,5] <- mean(energy43[,2])*0.0000002778
roofs[45,5] <- mean(energy44[,2])*0.0000002778
roofs[46,5] <- mean(energy45[,2])*0.0000002778
roofs[47,5] <- mean(energy46[,2])*0.0000002778
roofs[48,5] <- mean(energy47[,2])*0.0000002778
roofs[49,5] <- mean(energy48[,2])*0.0000002778
roofs[50,5] <- mean(energy49[,2])*0.0000002778

#renaming the headers
names(roofs)[1] <- "RidgeLength"
names(roofs)[2] <- "RidgeHeight"
names(roofs)[3] <- "EavesLength"
names(roofs)[4] <- "MeanAngle"
names(roofs)[5] <- "SolarProd"

#Now let's plot the data that we have to have a visual understanding that how
#the relationship between different roof variables may look like...
pairs(roofs)
```



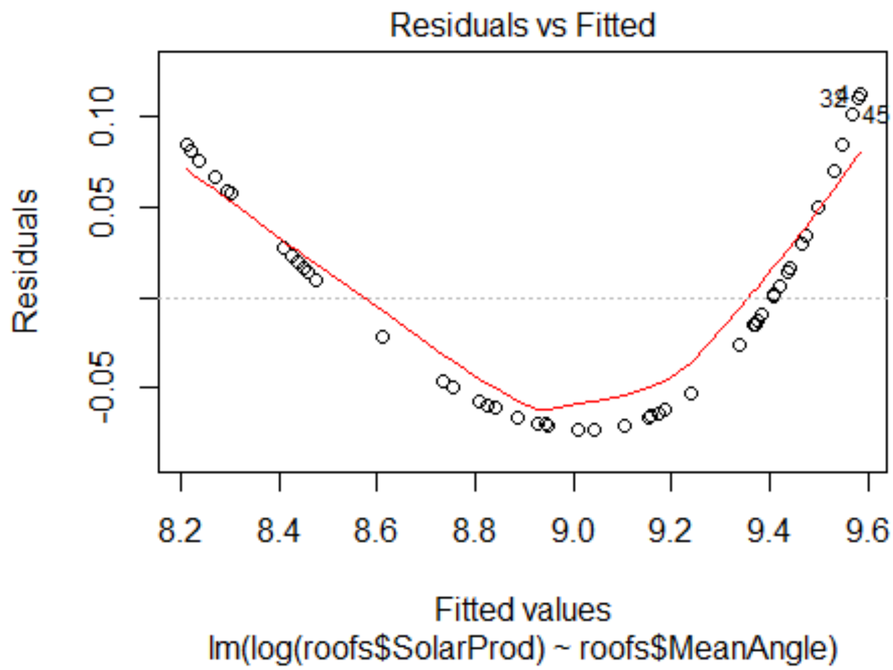

```
#We're interested in finding how different roof variables may effect amount
of solar energy production.
#So we would look at the last row.
#We can see from the visualization that solar production has
```

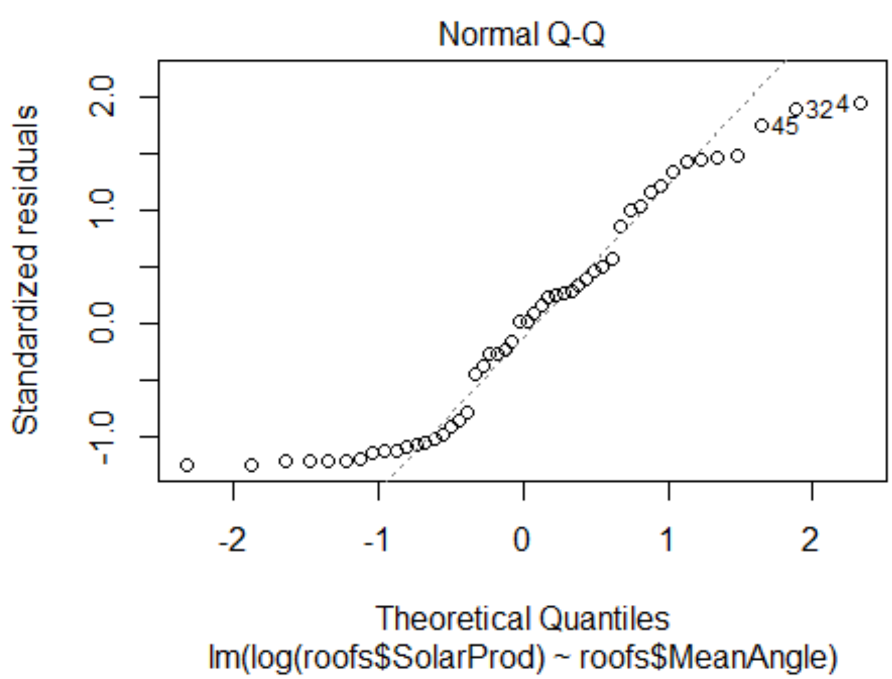
```
#positvie relationships with all roof's features but these relations are not linear...
#let's start from analyzing the relation between solar energy production and mean angle of roofs
plot(roofs$SolarProd~roofs$MeanAngle)
```

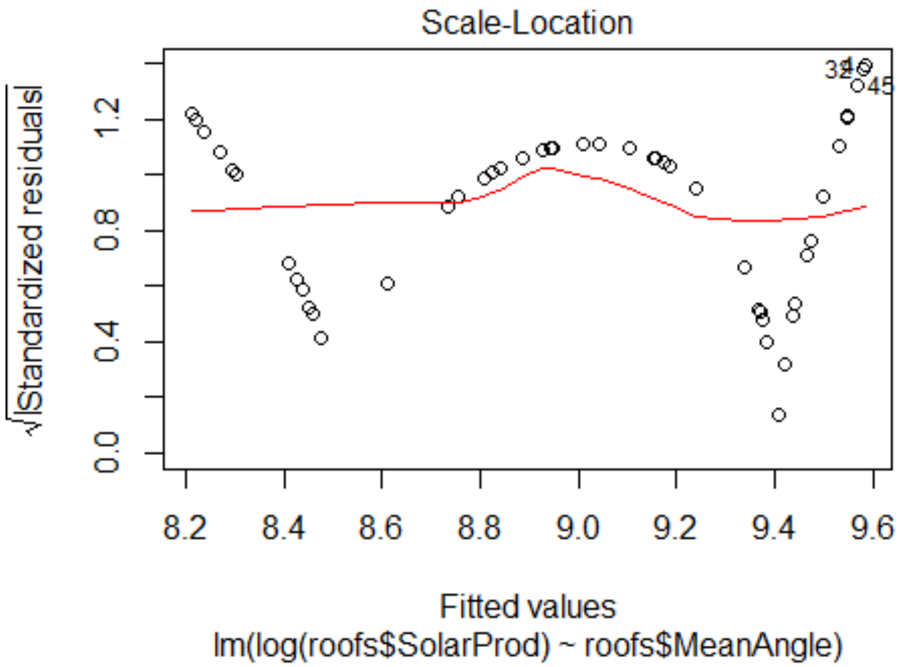


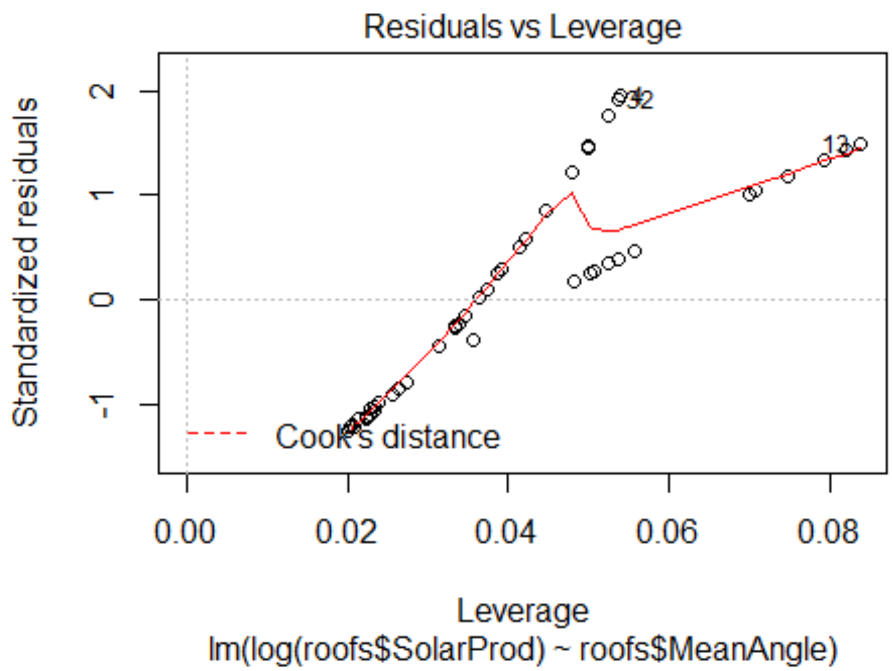

```
#this plot looks like an exponential regression model...let's fit an exponential regression model and see how the R-square looks like
angle.exponential.model <- lm(log(roofs$SolarProd)~roofs$MeanAngle)
summary(angle.exponential.model)

##
## Call:
## lm(formula = log(roofs$SolarProd) ~ roofs$MeanAngle)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.073093 -0.060909  0.001123  0.045736  0.112586
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     7.7550432  0.0250309  309.82   <2e-16 ***
## roofs$MeanAngle 0.0310542  0.0005858   53.01   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05937 on 48 degrees of freedom
```

```
## Multiple R-squared:  0.9832, Adjusted R-squared:  0.9829
## F-statistic:  2810 on 1 and 48 DF,  p-value: < 2.2e-16

plot(angle.exponential.model)
```







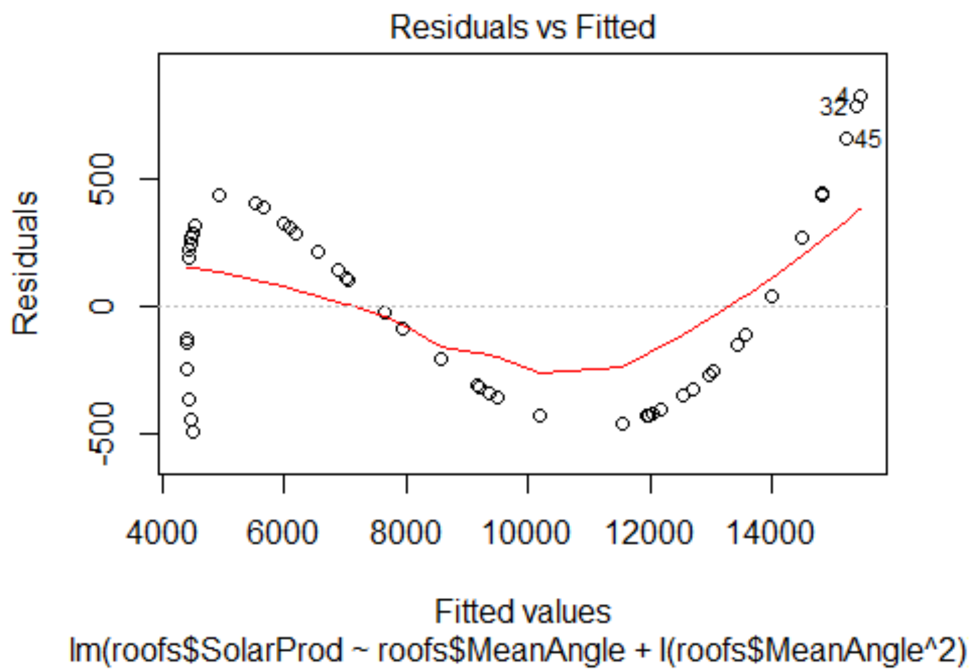


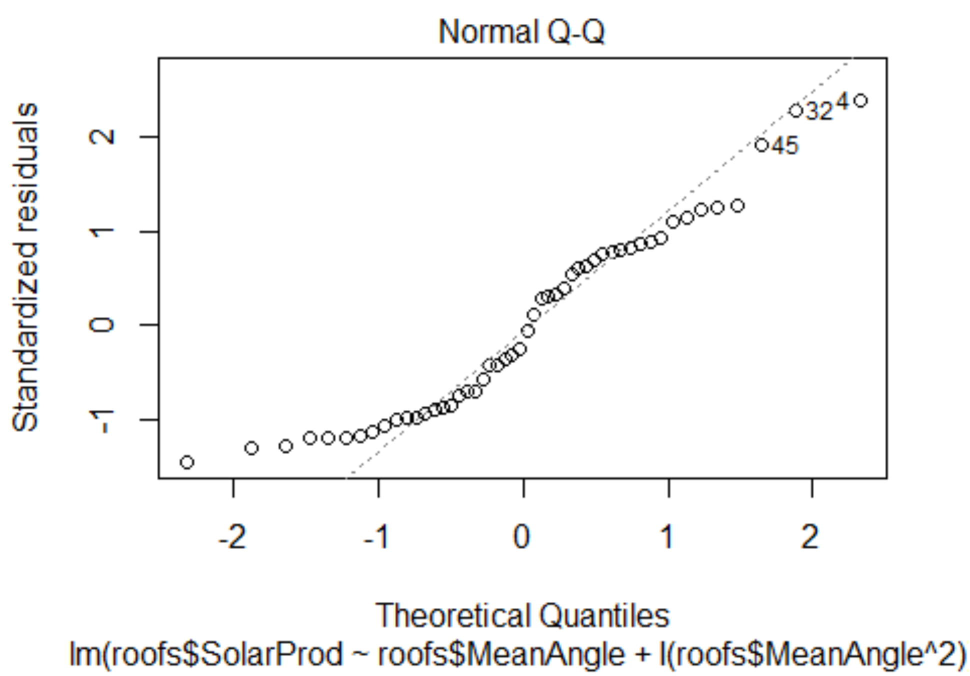

```
#from the summary we can see the coefficient of roofs' mean angle is almost 0.031 and from the R-square we can see that this model explains almost 98% of our data.
#this model looks very good. The p-value is also very low
```

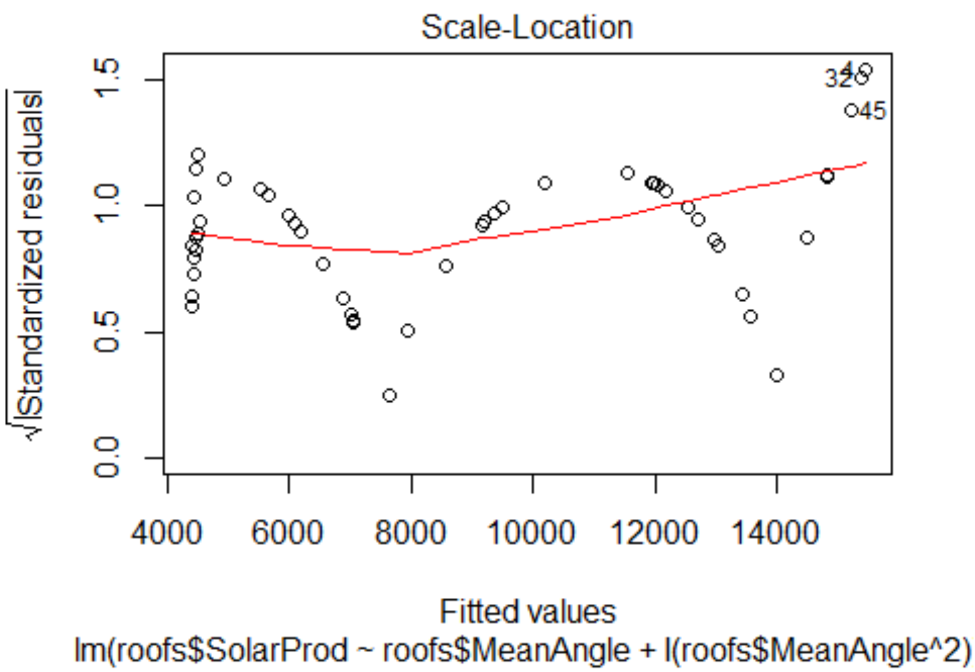
```r
#Checking the residuals they're somehow normal
#But just in case, let's now test a quadratic model for this relationship:
angle.quadratic.model <-
lm(roofs$SolarProd~roofs$MeanAngle+I(roofs$MeanAngle^2))
summary(angle.quadratic.model)
```

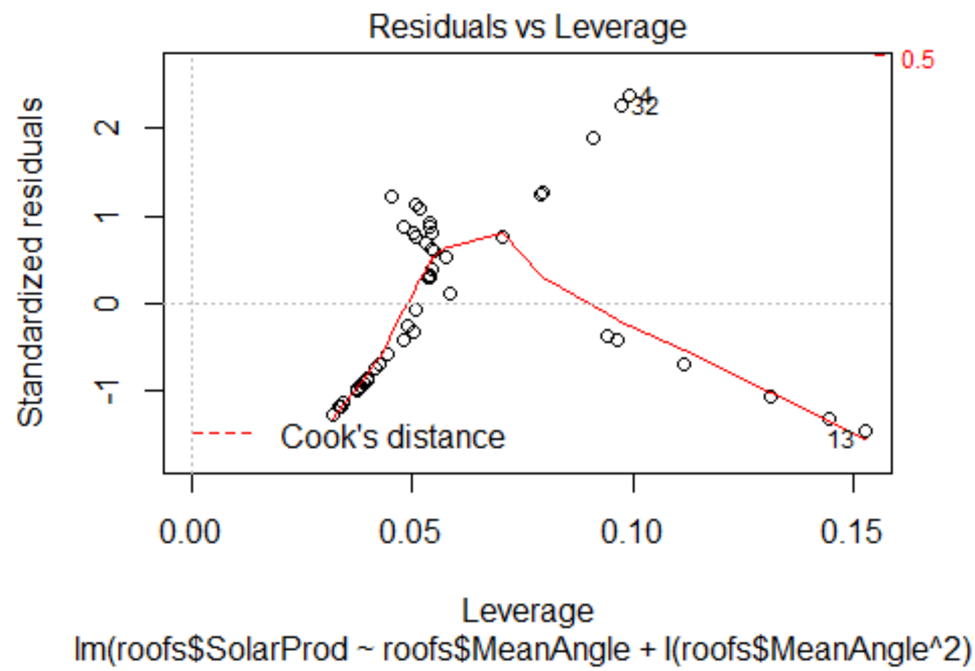
```
##
## Call:
## lm(formula = roofs$SolarProd ~ roofs$MeanAngle + I(roofs$MeanAngle^2))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -487.29 -333.02  -55.88  280.40  822.37
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          6744.2794   403.4475   16.72  < 2e-16 ***
## roofs$MeanAngle      -253.1519    23.7701  -10.65 4.03e-14 ***
## I(roofs$MeanAngle^2)    6.8050     0.3137   21.69  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 364.3 on 47 degrees of freedom
## Multiple R-squared:  0.9916, Adjusted R-squared:  0.9913
## F-statistic:  2782 on 2 and 47 DF,  p-value: < 2.2e-16
```

```r
plot(angle.quadratic.model)
```


Residuals vs Fitted
Residuals
Fitted values
lm(roofs$SolarProd ~ roofs$MeanAngle + I(roofs$MeanAngle^2)
32
4
45
500
0
-500
4000
6000
8000
10000
12000
14000

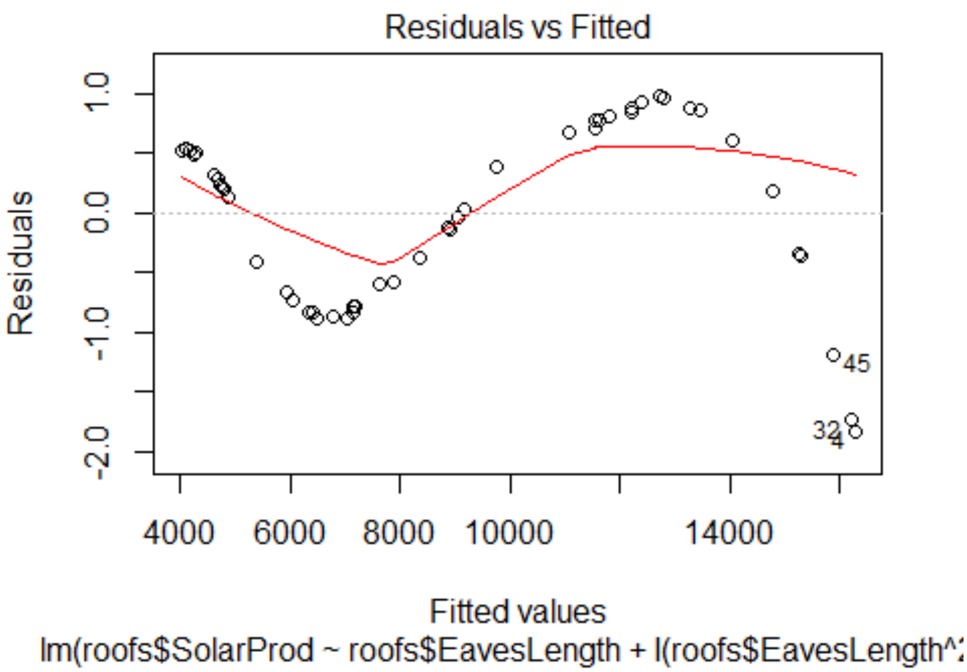

Normal Q-Q
Standardized residuals
Theoretical Quantiles
lm(roofs$SolarProd ~ roofs$MeanAngle + I(roofs$MeanAngle^2)
32
4
45
2
1
0
-1
-2
-1
0
1
2

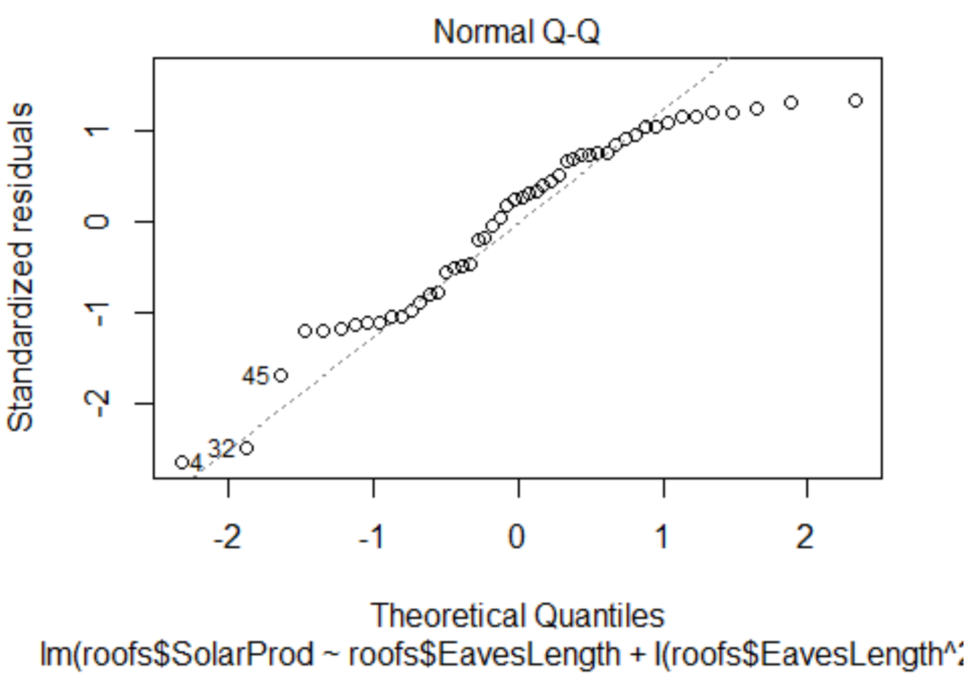


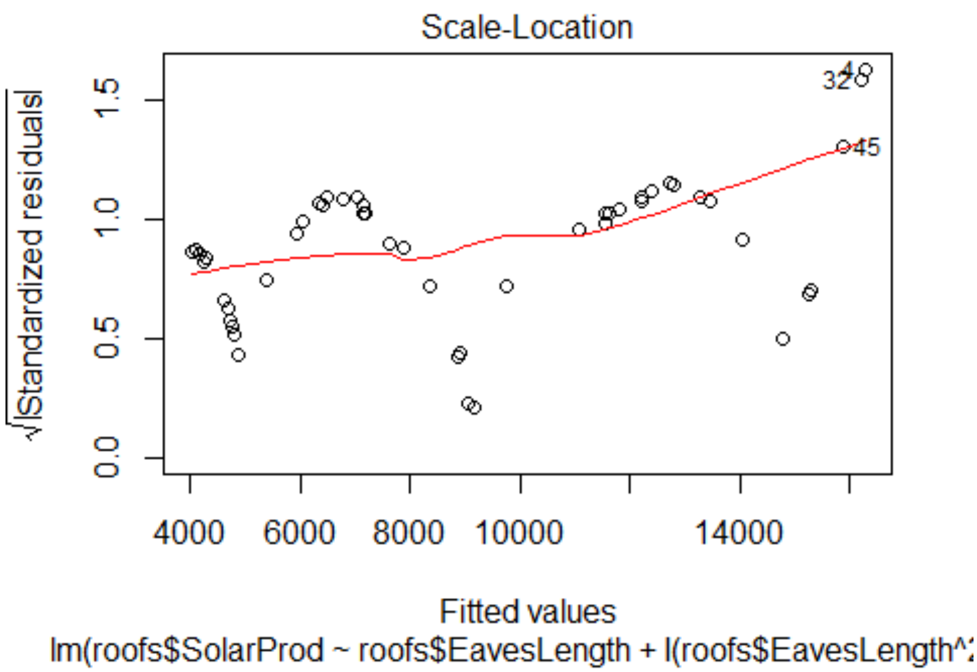


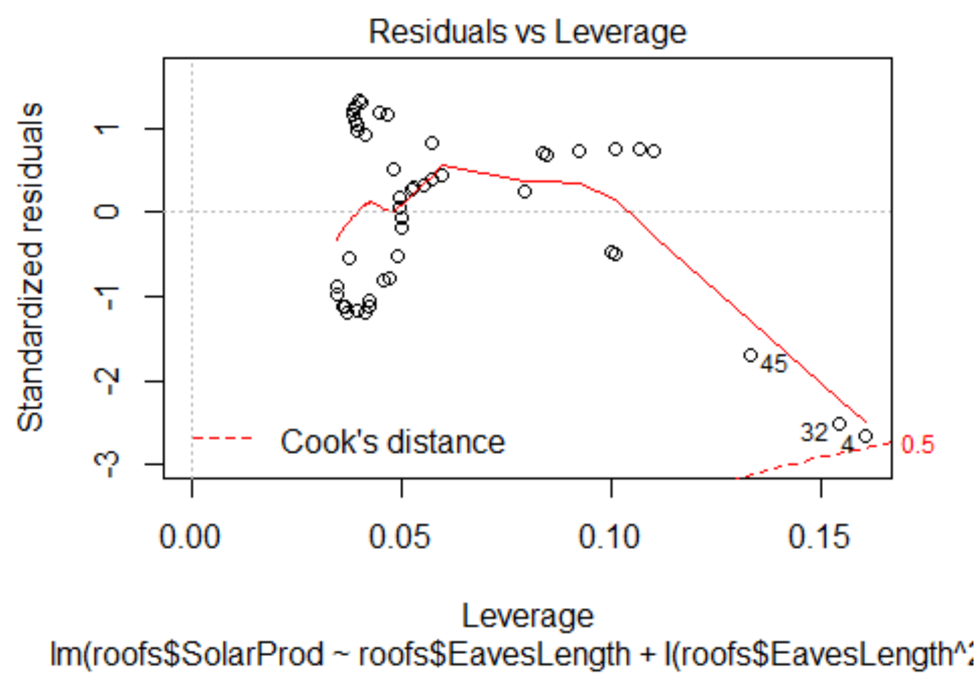

```
#The R-square quadratic model explains almost 99% of our data. P-value is
also very low
#But the residuals for this model are less normal than the exponential model
#So we'll proceed with the exponential model.
```

```r
#the mathematical equation for this relationship looks like below:
#log(SolarProd) = 6744.279 + 0.031*MeanAngle

#Now let's fit models for the rest of roofs' features and the solar
production
eaves.quadratic.model <-
lm(roofs$SolarProd~roofs$EavesLength+I(roofs$EavesLength^2))
summary(eaves.quadratic.model)

##
## Call:
## lm(formula = roofs$SolarProd ~ roofs$EavesLength + I(roofs$EavesLength^2))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8284 -0.6420  0.1874  0.5927  0.9782
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)            3943.3759     0.2585   15256   <2e-16 ***
## roofs$EavesLength      3884.2966     0.6642    5848   <2e-16 ***
## I(roofs$EavesLength^2) 1206.4658     0.3366    3584   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.753 on 47 degrees of freedom
## Multiple R-squared:      1,  Adjusted R-squared:      1
## F-statistic: 6.567e+08 on 2 and 47 DF,  p-value: < 2.2e-16

plot(eaves.quadratic.model)
```


Residuals vs Fitted
Residuals
1.0
0.0
-1.0
-2.0
45
32
4
4000
6000
8000
10000
14000
Fitted values
lm(roofs$SolarProd ~ roofs$EavesLength + I(roofs$EavesLength^2


Normal Q-Q
Standardized residuals
1
0
-1
-2
45
32
4
-2
-1
0
1
2
Theoretical Quantiles
lm(roofs$SolarProd ~ roofs$EavesLength + I(roofs$EavesLength^2

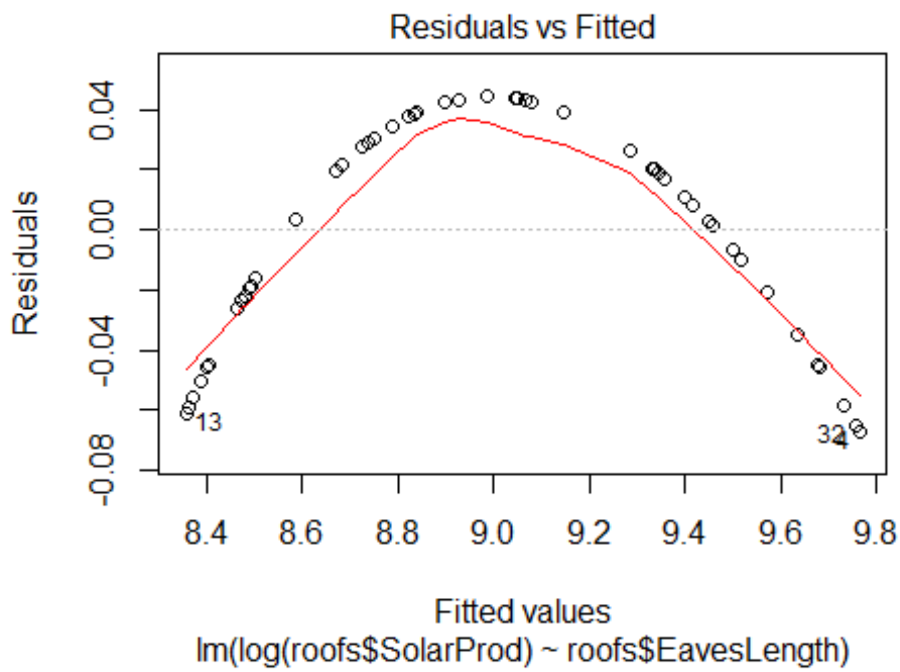


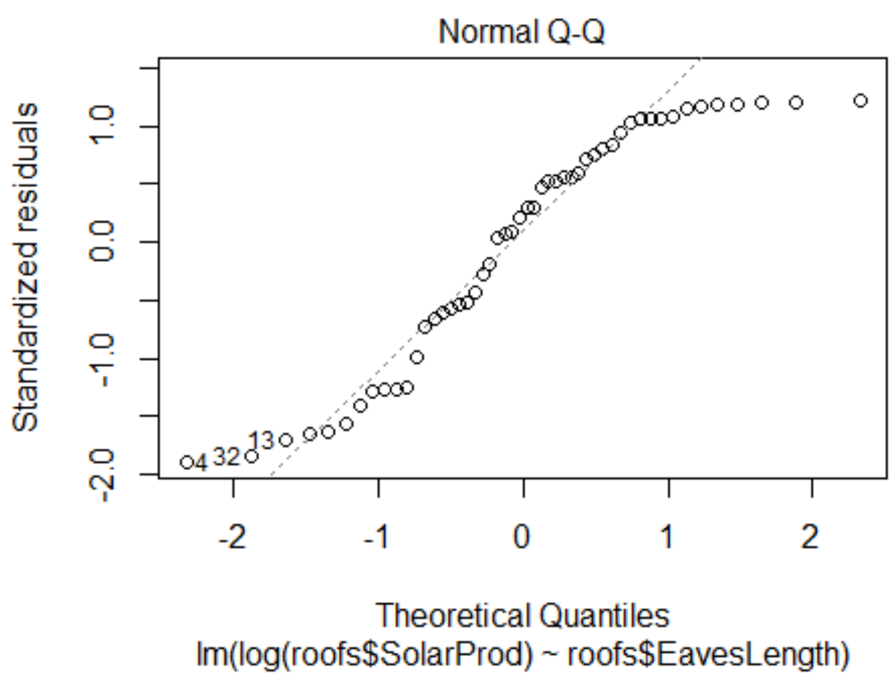


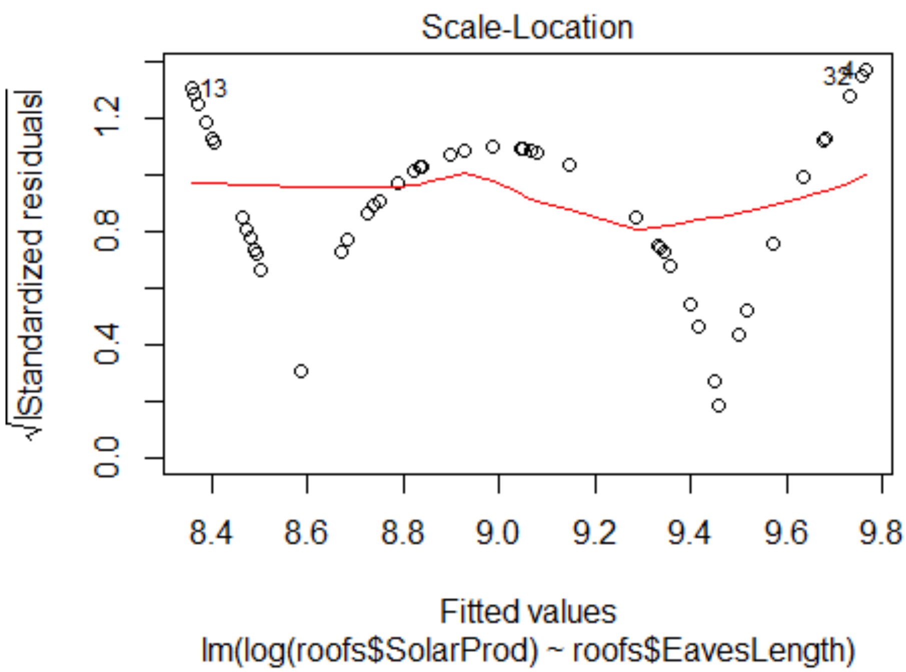

```
eaves.exponential.model <- lm(log(roofs$SolarProd)~roofs$EavesLength)
summary(eaves.exponential.model)
```

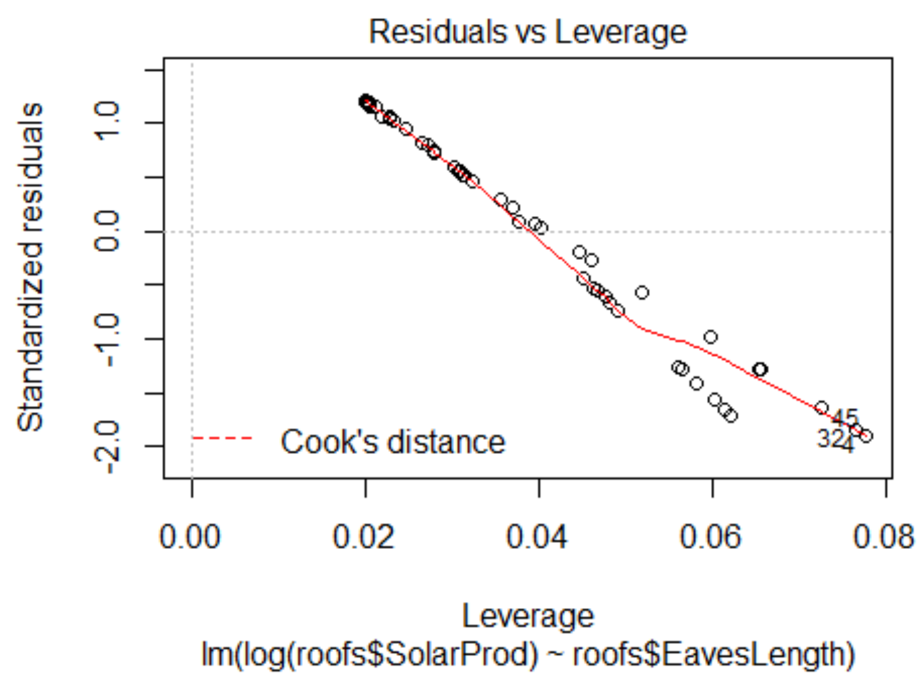
```
##
## Call:
## lm(formula = log(roofs$SolarProd) ~ roofs$EavesLength)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.06695 -0.02559  0.00932  0.03342  0.04428
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       8.345181   0.009292   898.1   <2e-16 ***
## roofs$EavesLength 0.720494   0.008397    85.8   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03687 on 48 degrees of freedom
## Multiple R-squared:  0.9935, Adjusted R-squared:  0.9934
## F-statistic:  7362 on 1 and 48 DF,  p-value: < 2.2e-16
```

```r
plot(eaves.exponential.model)
```


Residuals vs Fitted
Residuals
0.04
0.00
-0.04
-0.08
13
32
4
8.4
8.6
8.8
9.0
9.2
9.4
9.6
9.8
Fitted values
lm(log(roofs$SolarProd) ~ roofs$EavesLength)


Normal Q-Q
Standardized residuals
1.0
0.0
-1.0
-2.0
4
32
13
-2
-1
0
1
2
Theoretical Quantiles
lm(log(roofs$SolarProd) ~ roofs$EavesLength)






```
#Here we're comparing exponential and quadratic model for the relationship
between eaves length and solar production.
#The R-squared for the quadratic model is more than the exponential model and
the p-value is very low.
```

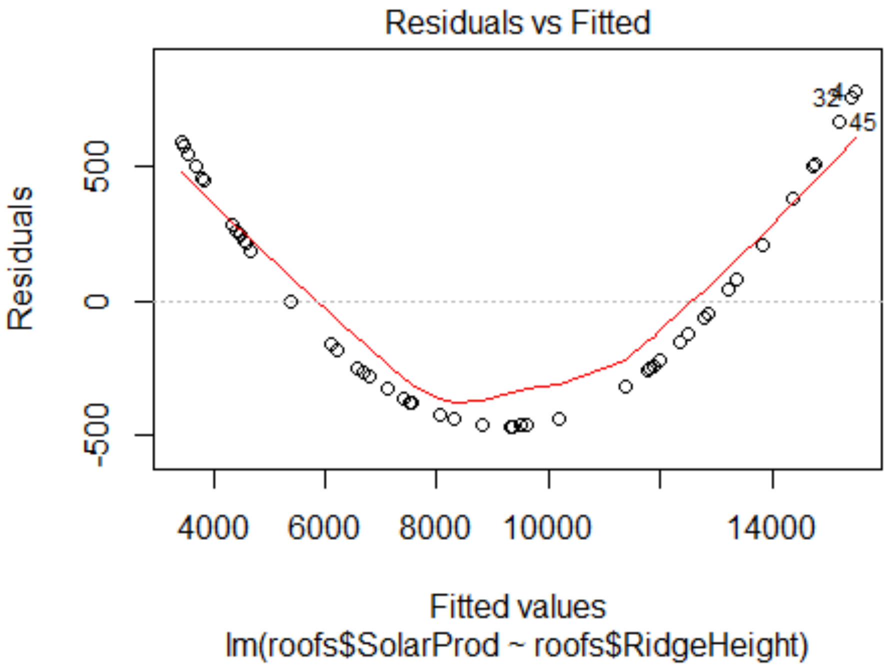
```r
#But the residauls distribution in the exponential model is more normal than the quadratic model
#So we'll go with the second mathematical equation for this relationship:
# log(SolarProd) = 8.345+0.72*EavesLength

#We're going to repeat the same procedure for the rest and select the most reliable fit model
ridgeHeight.quadratic.model <-
lm(roofs$SolarProd~roofs$RidgeHeight+I(roofs$RidgeHeight^2))
summary(ridgeHeight.quadratic.model)

##
## Call:
## lm(formula = roofs$SolarProd ~ roofs$RidgeHeight + I(roofs$RidgeHeight^2))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8286 -0.6431  0.1882  0.5928  0.9785
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)            2.939e+03  4.490e-01    6546   <2e-16 ***
## roofs$RidgeHeight      3.429e+02  9.112e-02    3764   <2e-16 ***
## I(roofs$RidgeHeight^2) 1.385e+01  3.867e-03    3583   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7532 on 47 degrees of freedom
## Multiple R-squared:      1,  Adjusted R-squared:      1
## F-statistic: 6.563e+08 on 2 and 47 DF,  p-value: < 2.2e-16

ridgeHeight.exponential.model <- lm(log(roofs$SolarProd)~roofs$RidgeHeight)
summary(ridgeHeight.exponential.model)

##
## Call:
## lm(formula = log(roofs$SolarProd) ~ roofs$RidgeHeight)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.06695 -0.02558  0.00932  0.03342  0.04428
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       8.1409142  0.0113414   717.8   <2e-16 ***
## roofs$RidgeHeight 0.0772082  0.0008999    85.8   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03687 on 48 degrees of freedom
```

```
## Multiple R-squared:  0.9935, Adjusted R-squared:  0.9934
## F-statistic:  7362 on 1 and 48 DF,  p-value: < 2.2e-16
```

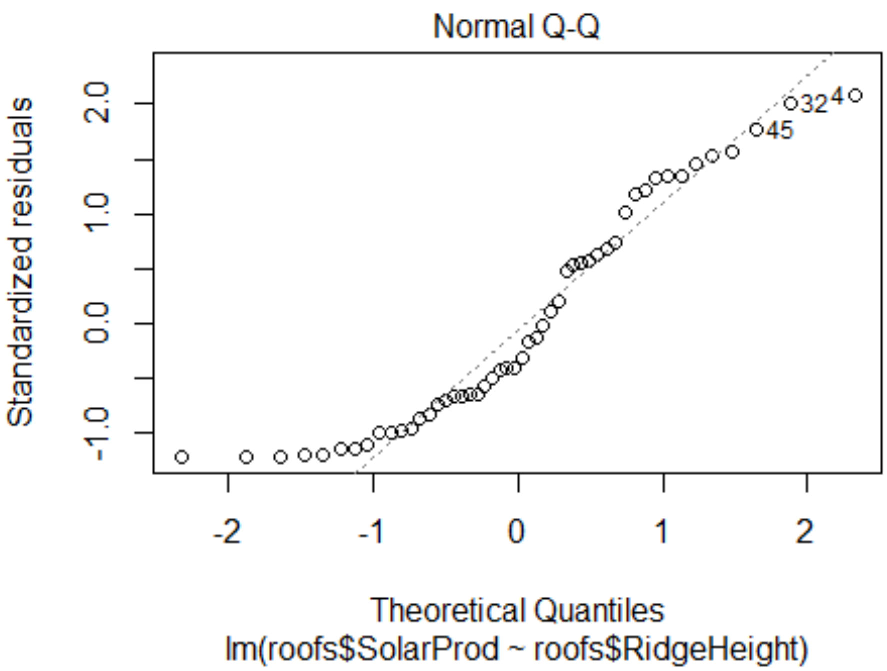
```
ridgeHeight.linear <- lm(roofs$SolarProd~roofs$RidgeHeight)
summary(ridgeHeight.linear)
```

```
##
## Call:
## lm(formula = roofs$SolarProd ~ roofs$RidgeHeight)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -466.7 -325.5 -137.5  274.6  776.5
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       1560.992    119.815   13.03   <2e-16 ***
## roofs$RidgeHeight  662.697      9.507   69.71   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 389.6 on 48 degrees of freedom
## Multiple R-squared:  0.9902, Adjusted R-squared:   0.99
## F-statistic:  4859 on 1 and 48 DF,  p-value: < 2.2e-16
```

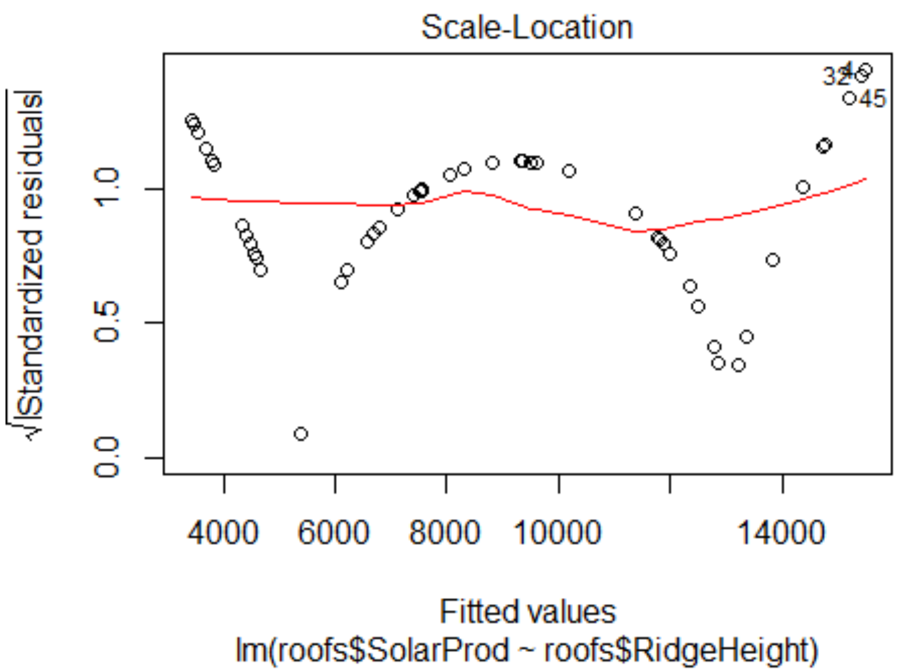
```
plot(ridgeHeight.linear) # normal residuals
```

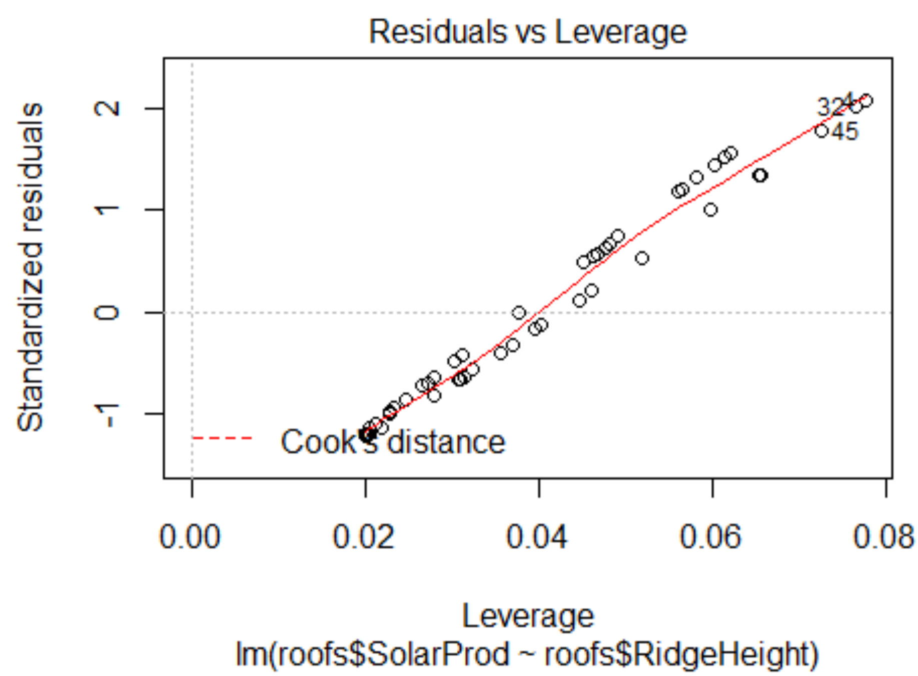


Residuals vs Fitted
Residuals
500
0
-500
4000
6000
8000
10000
14000
32
4
45
Fitted values
lm(roofs$SolarProd ~ roofs$RidgeHeight)

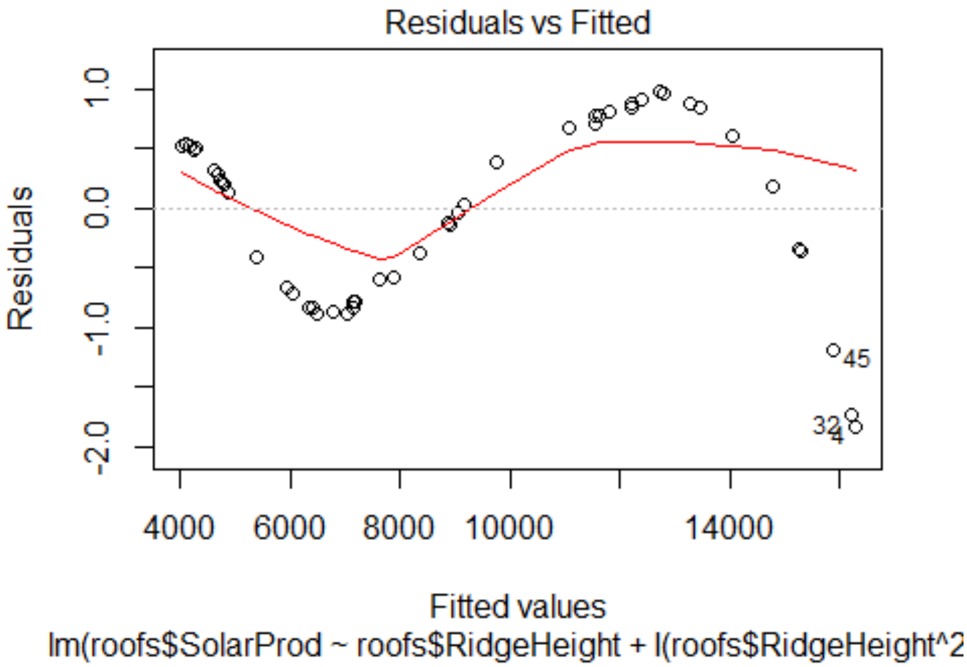


Normal Q-Q
Standardized residuals
2.0
1.0
0.0
-1.0
-2
-1
0
1
2
32
4
45
Theoretical Quantiles
lm(roofs$SolarProd ~ roofs$RidgeHeight)

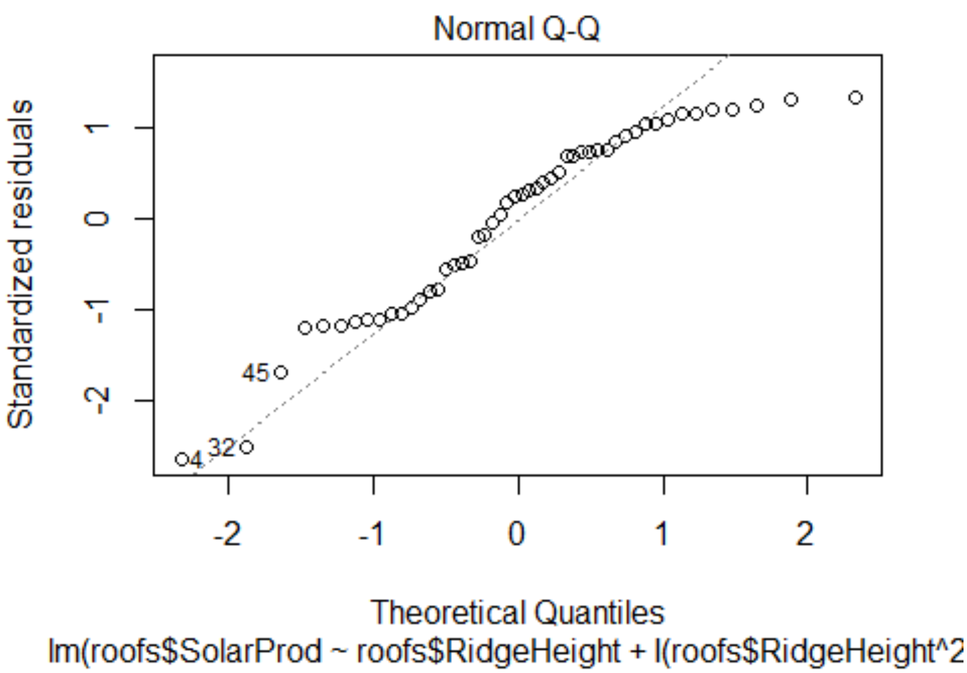




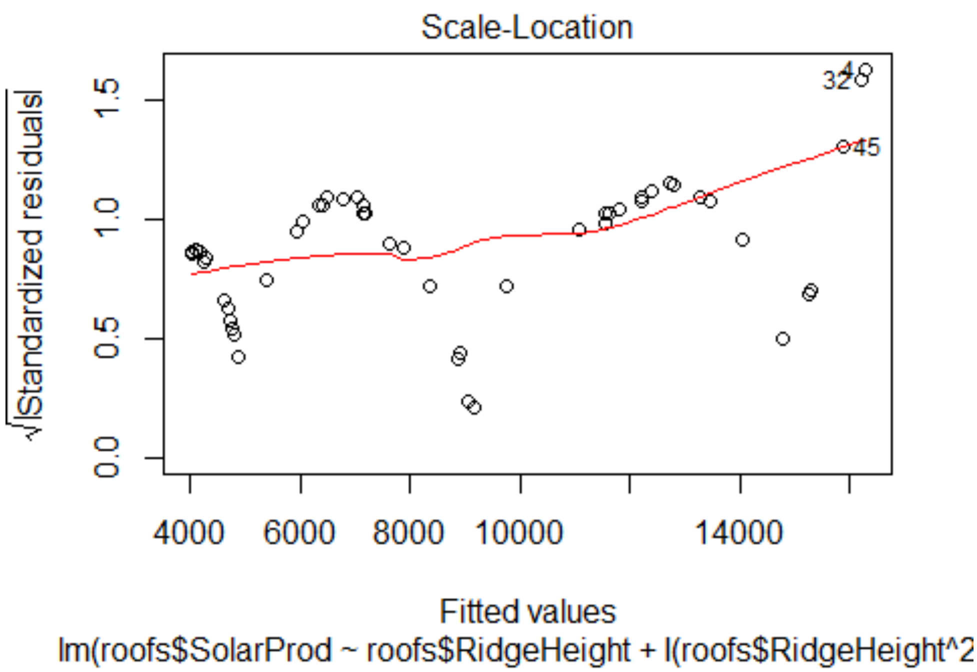

```r
plot(ridgeHeight.quadratic.model) #residuals not normal
```

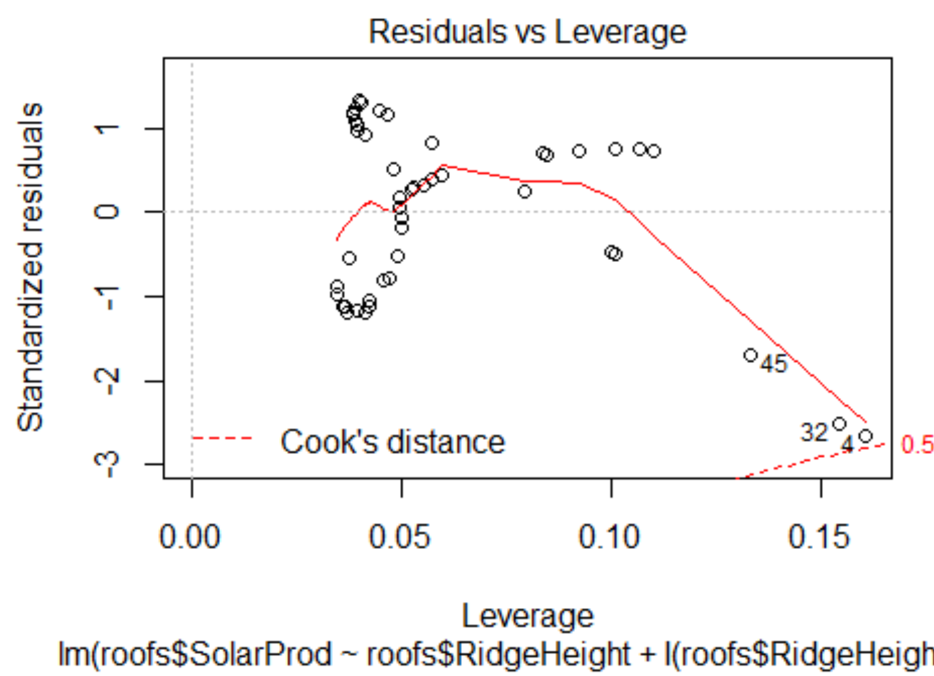


Residuals vs Fitted
Residuals
1.0
0.0
-1.0
-2.0
4000
6000
8000
10000
14000
45
32
4
Fitted values
lm(roofs$SolarProd ~ roofs$RidgeHeight + I(roofs$RidgeHeight^2

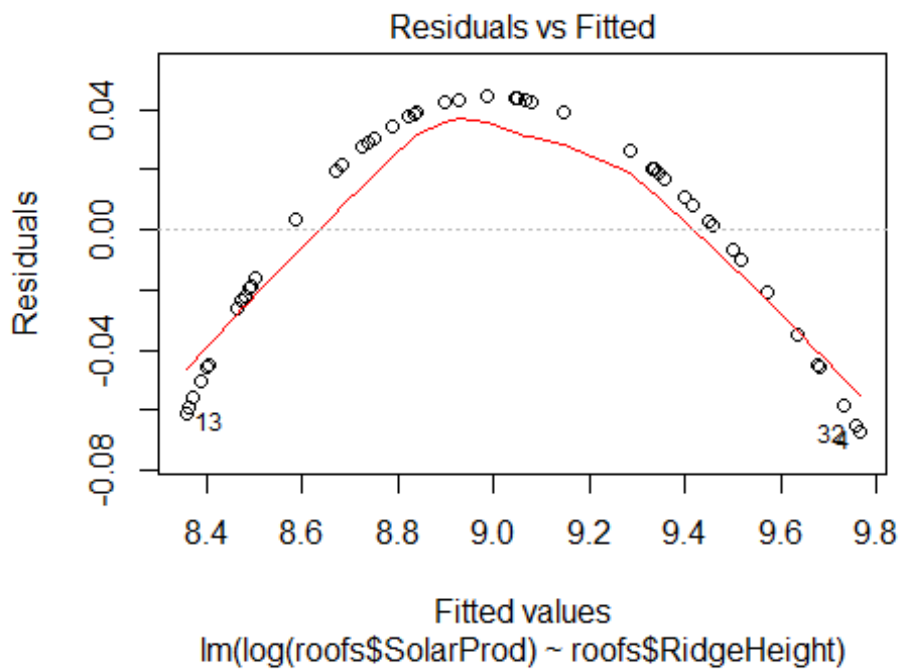


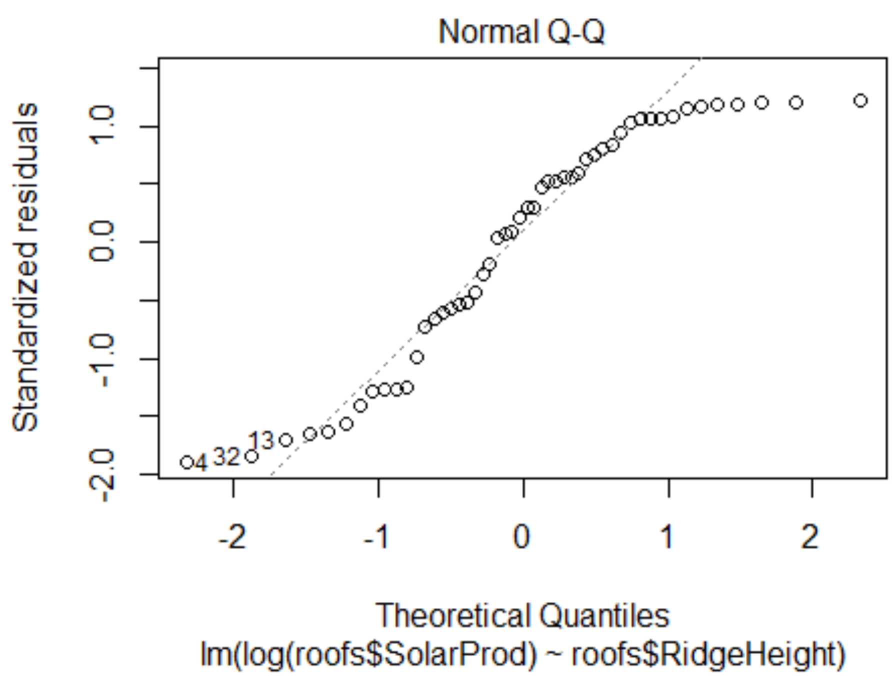
Normal Q-Q
Standardized residuals
1
0
-1
-2
45
4
32
-2
-1
0
1
2
Theoretical Quantiles
lm(roofs$SolarProd ~ roofs$RidgeHeight + I(roofs$RidgeHeight^2

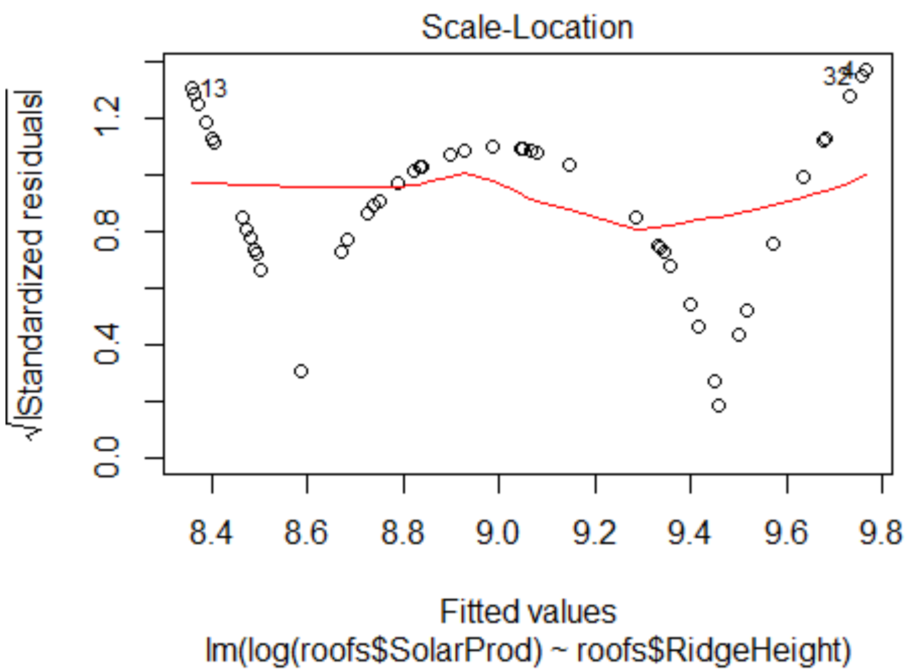


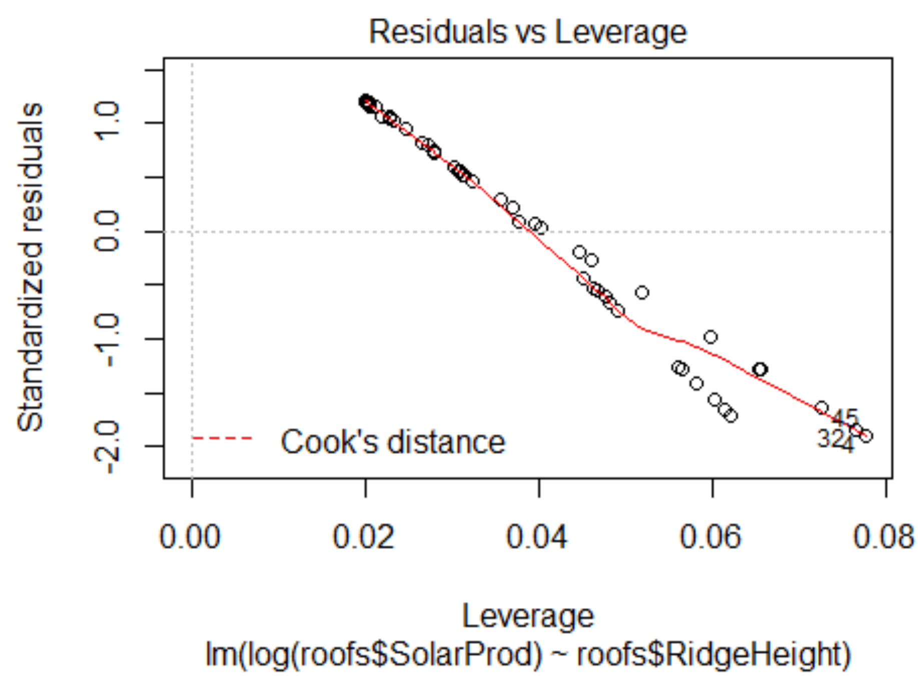



```
plot(ridgeHeight.exponential.model) #normal residuals
```


Residuals vs Fitted
Residuals
0.04
0.00
-0.04
-0.08
8.4
8.6
8.8
9.0
9.2
9.4
9.6
9.8
13
32
4
Fitted values
lm(log(roofs$SolarProd) ~ roofs$RidgeHeight)

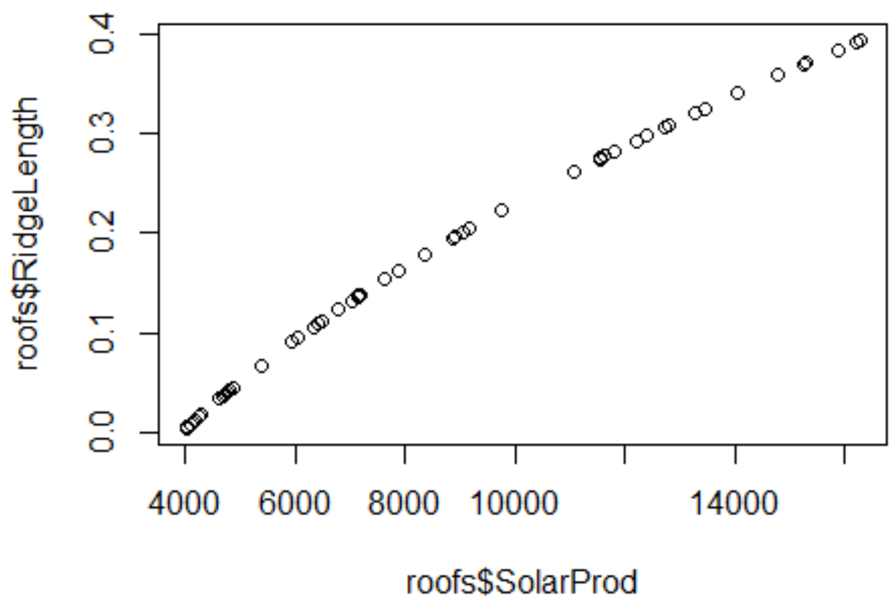

Normal Q-Q
Standardized residuals
1.0
0.0
-1.0
-2.0
-2
-1
0
1
2
4
32
13
Theoretical Quantiles
lm(log(roofs$SolarProd) ~ roofs$RidgeHeight)



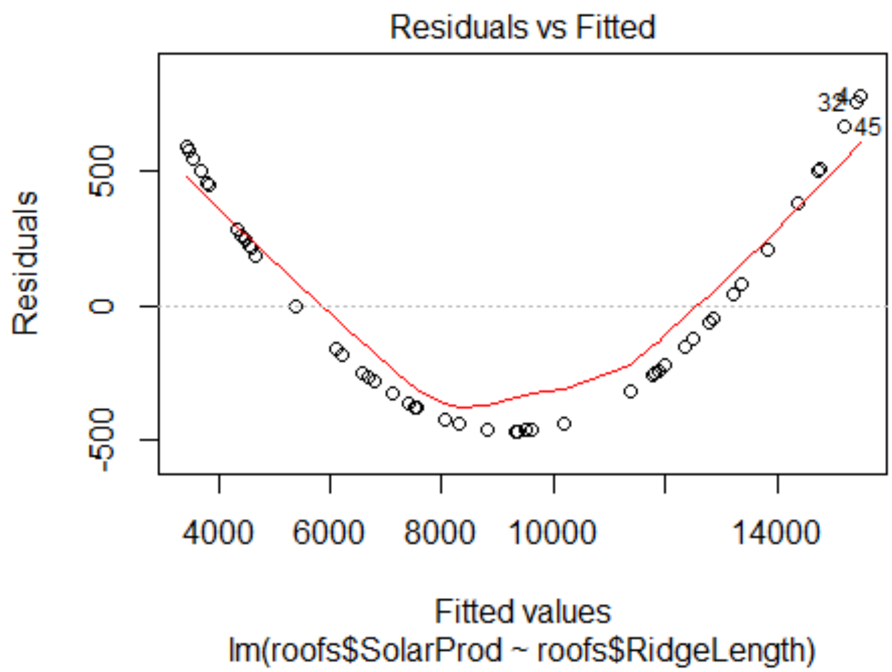


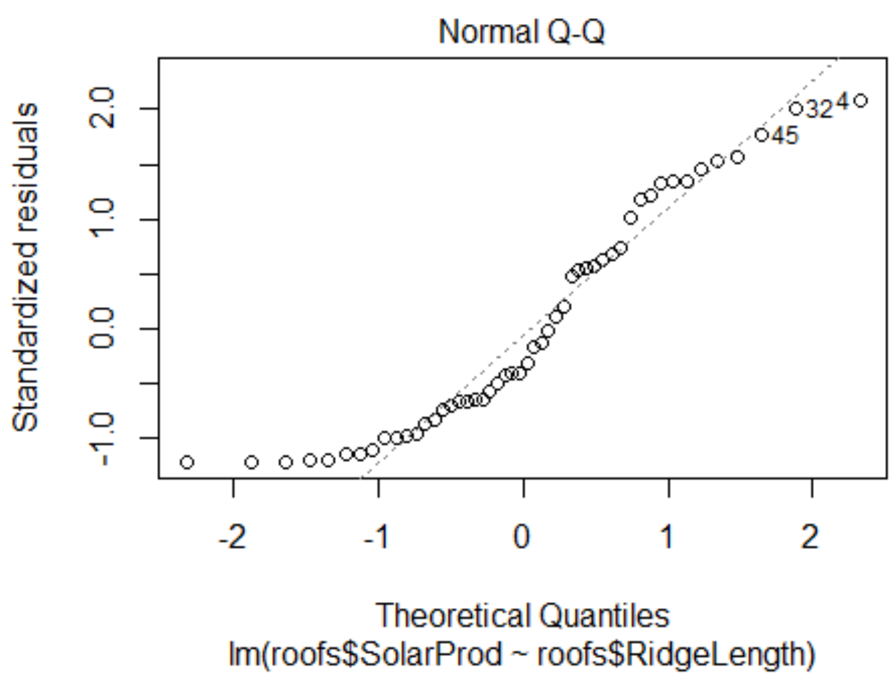

```
#this relationship is very close to being a linear relationship. As we can
see both in the linear and exponential models the R value is very high, the
residuals are normal, and the p-value is very low
```

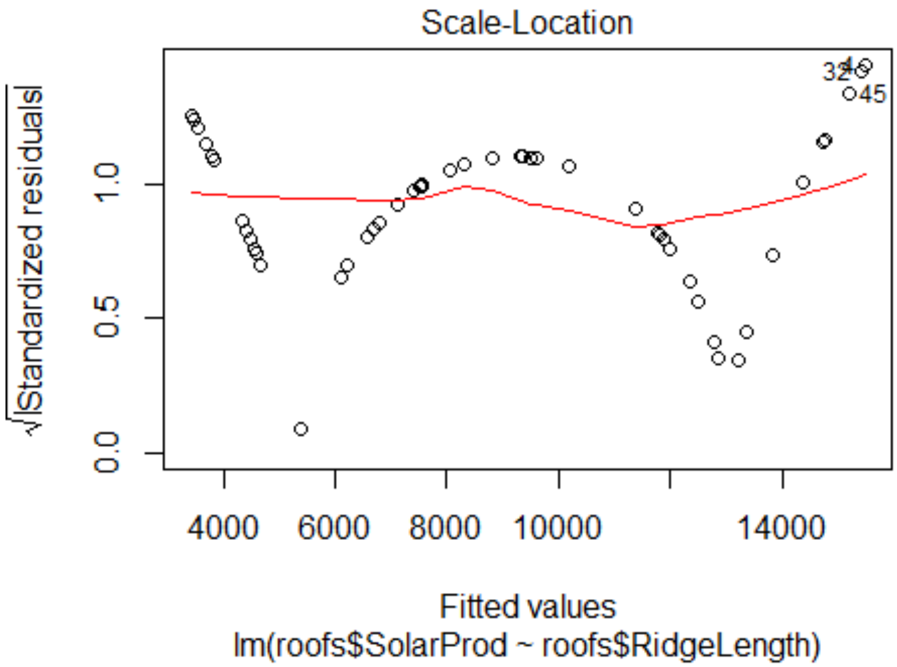
```r
plot(roofs$SolarProd,roofs$RidgeLength)#plot almost linear, let's proceed
with simple linear regression fit
```

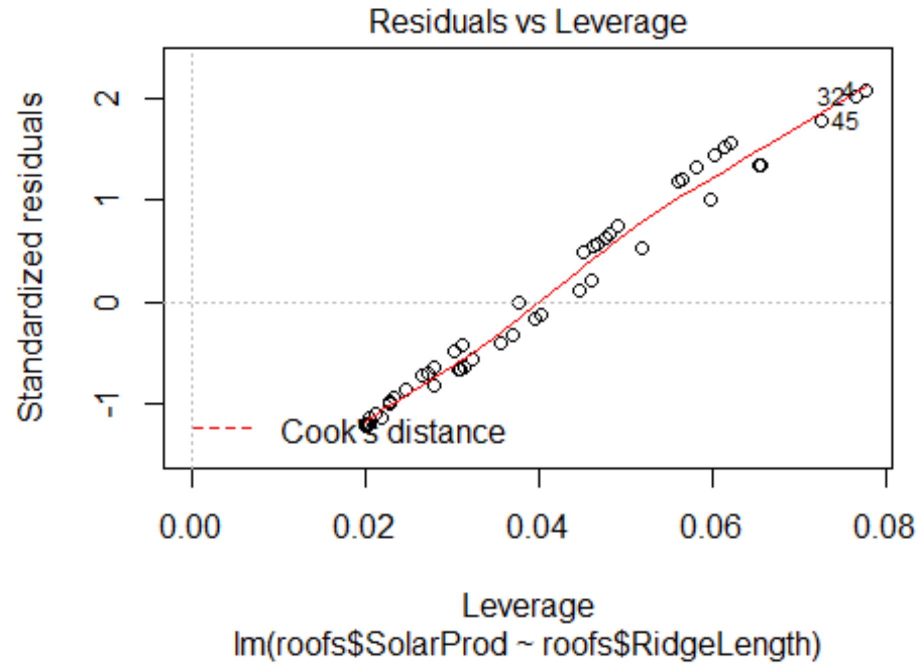



```r
ridgeLength.linear <- lm(roofs$SolarProd~roofs$RidgeLength)
summary(ridgeLength.linear)

##
## Call:
## lm(formula = roofs$SolarProd ~ roofs$RidgeLength)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -466.7 -325.5 -137.5  274.6  776.5
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        3289.49      98.46   33.41   <2e-16 ***
## roofs$RidgeLength 30967.40     444.24   69.71   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 389.6 on 48 degrees of freedom
## Multiple R-squared:  0.9902, Adjusted R-squared:   0.99
## F-statistic:  4859 on 1 and 48 DF,  p-value: < 2.2e-16

plot(ridgeLength.linear)
```


Residuals vs Fitted
Residuals
500
0
-500
4000
6000
8000
10000
14000
Fitted values
lm(roofs$SolarProd ~ roofs$RidgeLength)
32
4
45


Normal Q-Q
Standardized residuals
2.0
1.0
0.0
-1.0
-2
-1
0
1
2
Theoretical Quantiles
lm(roofs$SolarProd ~ roofs$RidgeLength)
32
4
45






```
#the mathematical equation is: SolarProd = 3289.49 + 30697.40*RidgeLength
```