Math644 Regression Project

Chhavi Tyagi

11/4/2019

setwd("/Users/chhavityagi/Desktop/math644project/")  
getwd()

## [1] "/Users/chhavityagi/Desktop/math644project"

library(readxl)  
xy<- read\_excel('/Users/chhavityagi/Desktop/math644project/ConcreteData.xls')  
attach(xy)  
xy

## # A tibble: 1,030 x 9  
## Cement BlastFurnaceSlag FLyAsh Water Superplasticizer CoarseAggregate  
## <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 540 0 0 162 2.5 1040   
## 2 540 0 0 162 2.5 1055   
## 3 332. 142. 0 228 0 932   
## 4 332. 142. 0 228 0 932   
## 5 199. 132. 0 192 0 978.  
## 6 266 114 0 228 0 932   
## 7 380 95 0 228 0 932   
## 8 380 95 0 228 0 932   
## 9 266 114 0 228 0 932   
## 10 475 0 0 228 0 932   
## # … with 1,020 more rows, and 3 more variables: FineAggregate <dbl>,  
## # Age <dbl>, ConcreteCompressiveStrength <dbl>

# Rename column to rmeove non unicode characters  
names(xy)[names(xy) == "ï..Cement"] <- "Cement"  
names(xy)[names(xy) == "BlastFurnaceSlag"] <- "BlastF"  
names(xy)[names(xy) == "Superplasticizer"] <- "SuperP"  
names(xy)[names(xy) == "CoarseAggregate"] <- "Coarse"  
names(xy)[names(xy) == "FineAggregate"] <- "Fine"  
names(xy)[names(xy) == "FLyAsh"] <- "FlyAsh"  
names(xy)[names(xy) == "ConcreteCompressiveStrength"] <- "Strength"  
# get column names  
colnames(xy)

## [1] "Cement" "BlastF" "FlyAsh" "Water" "SuperP" "Coarse"   
## [7] "Fine" "Age" "Strength"

xy

## # A tibble: 1,030 x 9  
## Cement BlastF FlyAsh Water SuperP Coarse Fine Age Strength  
## <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 540 0 0 162 2.5 1040 676 28 80.0  
## 2 540 0 0 162 2.5 1055 676 28 61.9  
## 3 332. 142. 0 228 0 932 594 270 40.3  
## 4 332. 142. 0 228 0 932 594 365 41.1  
## 5 199. 132. 0 192 0 978. 826. 360 44.3  
## 6 266 114 0 228 0 932 670 90 47.0  
## 7 380 95 0 228 0 932 594 365 43.7  
## 8 380 95 0 228 0 932 594 28 36.4  
## 9 266 114 0 228 0 932 670 28 45.9  
## 10 475 0 0 228 0 932 594 28 39.3  
## # … with 1,020 more rows

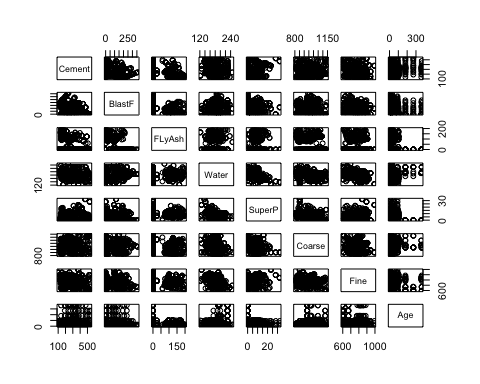
Cement = xy$Cement  
BlastF= xy$BlastF  
FlyAsh = xy$FlyAsh  
Water = xy$Water  
SuperP = xy$SuperP  
Coarse = xy$Coarse  
Fine = xy$Fine  
Age = xy$Age  
Strength = xy$Strength  
xy

## # A tibble: 1,030 x 9  
## Cement BlastF FlyAsh Water SuperP Coarse Fine Age Strength  
## <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 540 0 0 162 2.5 1040 676 28 80.0  
## 2 540 0 0 162 2.5 1055 676 28 61.9  
## 3 332. 142. 0 228 0 932 594 270 40.3  
## 4 332. 142. 0 228 0 932 594 365 41.1  
## 5 199. 132. 0 192 0 978. 826. 360 44.3  
## 6 266 114 0 228 0 932 670 90 47.0  
## 7 380 95 0 228 0 932 594 365 43.7  
## 8 380 95 0 228 0 932 594 28 36.4  
## 9 266 114 0 228 0 932 670 28 45.9  
## 10 475 0 0 228 0 932 594 28 39.3  
## # … with 1,020 more rows

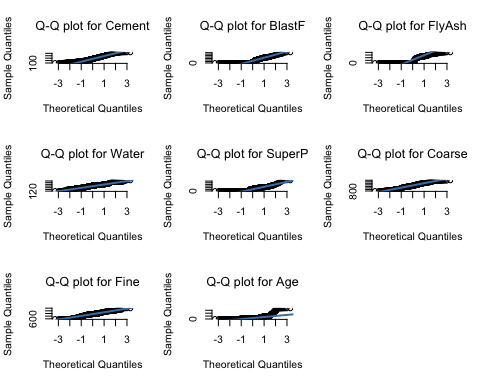
summary(xy)

## Cement BlastF FlyAsh Water   
## Min. :102.0 Min. : 0.0 Min. : 0.00 Min. :121.8   
## 1st Qu.:192.4 1st Qu.: 0.0 1st Qu.: 0.00 1st Qu.:164.9   
## Median :272.9 Median : 22.0 Median : 0.00 Median :185.0   
## Mean :281.2 Mean : 73.9 Mean : 54.19 Mean :181.6   
## 3rd Qu.:350.0 3rd Qu.:142.9 3rd Qu.:118.27 3rd Qu.:192.0   
## Max. :540.0 Max. :359.4 Max. :200.10 Max. :247.0   
## SuperP Coarse Fine Age   
## Min. : 0.000 Min. : 801.0 Min. :594.0 Min. : 1.00   
## 1st Qu.: 0.000 1st Qu.: 932.0 1st Qu.:731.0 1st Qu.: 7.00   
## Median : 6.350 Median : 968.0 Median :779.5 Median : 28.00   
## Mean : 6.203 Mean : 972.9 Mean :773.6 Mean : 45.66   
## 3rd Qu.:10.160 3rd Qu.:1029.4 3rd Qu.:824.0 3rd Qu.: 56.00   
## Max. :32.200 Max. :1145.0 Max. :992.6 Max. :365.00   
## Strength   
## Min. : 2.332   
## 1st Qu.:23.707   
## Median :34.443   
## Mean :35.818   
## 3rd Qu.:46.136   
## Max. :82.599

pairs(~Cement+BlastF+FLyAsh+Water+SuperP+Coarse+Fine+Age, data =xy)



par(mfrow = c(3,3)) # Create a 2 x 3 plotting matrix  
# The next 4 plots created will be plotted next to each other  
  
# Plot 1  
qqnorm(xy$Cement, pch = 1, frame = FALSE,main = expression("Q-Q plot for Cement"))  
qqline(xy$Cement, col = "steelblue", lwd = 2)  
  
# Plot 2  
qqnorm(xy$BlastF, pch = 1, frame = FALSE,main = expression("Q-Q plot for BlastF"))  
qqline(xy$BlastF, col = "steelblue", lwd = 2)  
  
# Plot 3  
qqnorm(xy$FlyAsh, pch = 1, frame = FALSE,main = expression("Q-Q plot for FlyAsh"))  
qqline(xy$FlyAsh, col = "steelblue", lwd = 2)  
  
# Plot 4  
qqnorm(xy$Water, pch = 1, frame = FALSE,main = expression("Q-Q plot for Water"))  
qqline(xy$Water, col = "steelblue", lwd = 2)  
  
# Plot 5  
qqnorm(xy$SuperP, pch = 1, frame = FALSE,main = expression("Q-Q plot for SuperP"))  
qqline(xy$SuperP, col = "steelblue", lwd = 2)  
  
#Plot 6  
qqnorm(xy$Coarse, pch = 1, frame = FALSE,main = expression("Q-Q plot for Coarse"))  
qqline(xy$Coarse, col = "steelblue", lwd = 2)  
  
# Plot 7  
qqnorm(xy$Fine, pch = 1, frame = FALSE,main = expression("Q-Q plot for Fine"))  
qqline(xy$Fine, col = "steelblue", lwd = 2)  
  
# Plot 8  
qqnorm(xy$Age, pch = 1, frame = FALSE,main = expression("Q-Q plot for Age"))  
qqline(xy$Age, col = "steelblue", lwd = 2)



par(mfrow = c(2,4)) # Create a 2 x 3 plotting matrix  
# The next 4 plots created will be plotted next to each other  
  
# Plot 1  
boxplot\_Cement=boxplot(Cement,col='red',xlab="Cement")  
boxplot\_Cement$output

## NULL

# Plot 2  
boxplot\_BlastF=boxplot(BlastF,col='green',xlab="BlastF")  
boxplot\_BlastF$output

## NULL

# Plot 3  
boxplot\_FLyAsh=boxplot(FLyAsh,col='blue',xlab="FLyAsh")  
boxplot\_FLyAsh$output

## NULL

# Plot 4  
boxplot\_Water=boxplot(Water,col='orange',xlab="Water")  
boxplot\_Water$output

## NULL

#Plot 5  
boxplot\_SuperP=boxplot(SuperP,col='red',xlab="SuperP")  
boxplot\_SuperP$output

## NULL

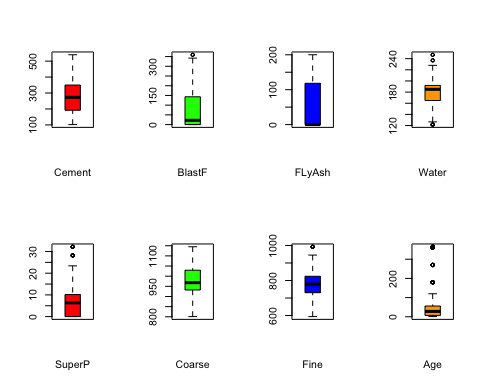
#Plot 6  
boxplot\_Coarse=boxplot(Coarse,col='green',xlab="Coarse")  
boxplot\_Coarse$output

## NULL

#Plot 7  
boxplot\_Fine=boxplot(Fine,col='blue',xlab="Fine")  
boxplot\_Fine$output

## NULL

#Plot 8  
boxplot\_Age=boxplot(Age,col='orange',xlab="Age")



boxplot\_Age$output

## NULL

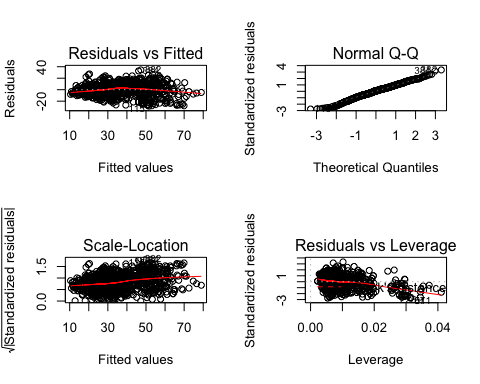
reg = lm(Strength~Cement+BlastF+FlyAsh+Water+SuperP+Coarse+Fine+Age)  
summary(reg)

##   
## Call:  
## lm(formula = Strength ~ Cement + BlastF + FlyAsh + Water + SuperP +   
## Coarse + Fine + Age)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -28.653 -6.303 0.704 6.562 34.446   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -23.163756 26.588421 -0.871 0.383851   
## Cement 0.119785 0.008489 14.110 < 2e-16 \*\*\*  
## BlastF 0.103847 0.010136 10.245 < 2e-16 \*\*\*  
## FlyAsh 0.087943 0.012585 6.988 5.03e-12 \*\*\*  
## Water -0.150298 0.040179 -3.741 0.000194 \*\*\*  
## SuperP 0.290687 0.093460 3.110 0.001921 \*\*   
## Coarse 0.018030 0.009394 1.919 0.055227 .   
## Fine 0.020154 0.010703 1.883 0.059968 .   
## Age 0.114226 0.005427 21.046 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 10.4 on 1021 degrees of freedom  
## Multiple R-squared: 0.6155, Adjusted R-squared: 0.6125   
## F-statistic: 204.3 on 8 and 1021 DF, p-value: < 2.2e-16

anova(reg)

## Analysis of Variance Table  
##   
## Response: Strength  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Cement 1 71172 71172 658.0463 < 2.2e-16 \*\*\*  
## BlastF 1 22957 22957 212.2606 < 2.2e-16 \*\*\*  
## FlyAsh 1 21636 21636 200.0464 < 2.2e-16 \*\*\*  
## Water 1 11459 11459 105.9488 < 2.2e-16 \*\*\*  
## SuperP 1 1360 1360 12.5785 0.0004079 \*\*\*  
## Coarse 1 253 253 2.3435 0.1261178   
## Fine 1 1 1 0.0058 0.9393393   
## Age 1 47905 47905 442.9232 < 2.2e-16 \*\*\*  
## Residuals 1021 110428 108   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

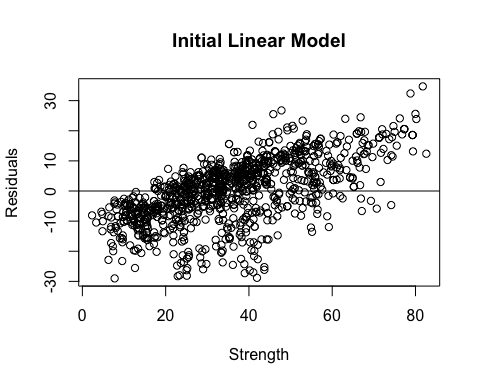
#residuals versus leverage  
# https://medium.com/data-distilled/residual-plots-part-4-residuals-vs-leverage-plot-14aeed009ef7  
par(mfrow=c(2,2))  
plot(reg)



reg1 = lm(Strength~Cement+BlastF+FlyAsh+Water+SuperP+Age)  
summary(reg1)

##   
## Call:  
## lm(formula = Strength ~ Cement + BlastF + FlyAsh + Water + SuperP +   
## Age)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -29.014 -6.474 0.650 6.546 34.726   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 29.030224 4.212476 6.891 9.64e-12 \*\*\*  
## Cement 0.105427 0.004248 24.821 < 2e-16 \*\*\*  
## BlastF 0.086494 0.004975 17.386 < 2e-16 \*\*\*  
## FlyAsh 0.068708 0.007736 8.881 < 2e-16 \*\*\*  
## Water -0.218292 0.021128 -10.332 < 2e-16 \*\*\*  
## SuperP 0.239003 0.084586 2.826 0.00481 \*\*   
## Age 0.113495 0.005408 20.987 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 10.41 on 1023 degrees of freedom  
## Multiple R-squared: 0.614, Adjusted R-squared: 0.6117   
## F-statistic: 271.2 on 6 and 1023 DF, p-value: < 2.2e-16

residuals1<-residuals(reg1)  
plot(xy$Strength, residuals1,ylab = "Residuals", xlab = "Strength",main = "Initial Linear Model")   
abline(0, 0) # the horizon



#Scale the addtives  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(tidyverse)

## ── Attaching packages ───────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────── tidyverse 1.2.1 ──

## ✔ ggplot2 3.2.1 ✔ readr 1.3.1  
## ✔ tibble 2.1.3 ✔ purrr 0.3.2  
## ✔ tidyr 1.0.0 ✔ stringr 1.4.0  
## ✔ ggplot2 3.2.1 ✔ forcats 0.4.0

## ── Conflicts ──────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

## Add new column that is the output of the ANN, using mutate() in dplyr  
# Calculate water to binder ratio as WaterBind = W / (cemenet+flyash+Blastf)  
  
xy<- mutate(xy, WaterBind = (Water/(Cement + FlyAsh + BlastF + SuperP)))  
xy

## # A tibble: 1,030 x 10  
## Cement BlastF FlyAsh Water SuperP Coarse Fine Age Strength WaterBind  
## <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 540 0 0 162 2.5 1040 676 28 80.0 0.299  
## 2 540 0 0 162 2.5 1055 676 28 61.9 0.299  
## 3 332. 142. 0 228 0 932 594 270 40.3 0.48   
## 4 332. 142. 0 228 0 932 594 365 41.1 0.48   
## 5 199. 132. 0 192 0 978. 826. 360 44.3 0.580  
## 6 266 114 0 228 0 932 670 90 47.0 0.6   
## 7 380 95 0 228 0 932 594 365 43.7 0.48   
## 8 380 95 0 228 0 932 594 28 36.4 0.48   
## 9 266 114 0 228 0 932 670 28 45.9 0.6   
## 10 475 0 0 228 0 932 594 28 39.3 0.48   
## # … with 1,020 more rows

WaterBind<-xy$WaterBind

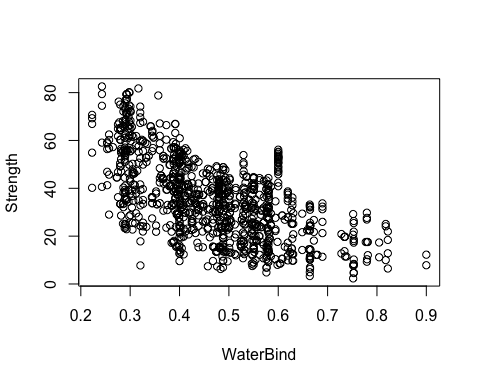
summary(xy)

## Cement BlastF FlyAsh Water   
## Min. :102.0 Min. : 0.0 Min. : 0.00 Min. :121.8   
## 1st Qu.:192.4 1st Qu.: 0.0 1st Qu.: 0.00 1st Qu.:164.9   
## Median :272.9 Median : 22.0 Median : 0.00 Median :185.0   
## Mean :281.2 Mean : 73.9 Mean : 54.19 Mean :181.6   
## 3rd Qu.:350.0 3rd Qu.:142.9 3rd Qu.:118.27 3rd Qu.:192.0   
## Max. :540.0 Max. :359.4 Max. :200.10 Max. :247.0   
## SuperP Coarse Fine Age   
## Min. : 0.000 Min. : 801.0 Min. :594.0 Min. : 1.00   
## 1st Qu.: 0.000 1st Qu.: 932.0 1st Qu.:731.0 1st Qu.: 7.00   
## Median : 6.350 Median : 968.0 Median :779.5 Median : 28.00   
## Mean : 6.203 Mean : 972.9 Mean :773.6 Mean : 45.66   
## 3rd Qu.:10.160 3rd Qu.:1029.4 3rd Qu.:824.0 3rd Qu.: 56.00   
## Max. :32.200 Max. :1145.0 Max. :992.6 Max. :365.00   
## Strength WaterBind   
## Min. : 2.332 Min. :0.2228   
## 1st Qu.:23.707 1st Qu.:0.3733   
## Median :34.443 Median :0.4671   
## Mean :35.818 Mean :0.4634   
## 3rd Qu.:46.136 3rd Qu.:0.5504   
## Max. :82.599 Max. :0.9000

# Regress with Water to Binder Ratio  
  
regwb = lm(Strength ~ Coarse+Fine+Age+WaterBind)  
summary(regwb)

##   
## Call:  
## lm(formula = Strength ~ Coarse + Fine + Age + WaterBind)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -36.312 -7.283 0.819 7.487 35.627   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 60.463138 6.126821 9.869 < 2e-16 \*\*\*  
## Coarse 0.001605 0.004664 0.344 0.73080   
## Fine 0.014084 0.004620 3.048 0.00236 \*\*   
## Age 0.118999 0.005607 21.223 < 2e-16 \*\*\*  
## WaterBind -91.800716 2.911627 -31.529 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 10.95 on 1025 degrees of freedom  
## Multiple R-squared: 0.5722, Adjusted R-squared: 0.5706   
## F-statistic: 342.8 on 4 and 1025 DF, p-value: < 2.2e-16

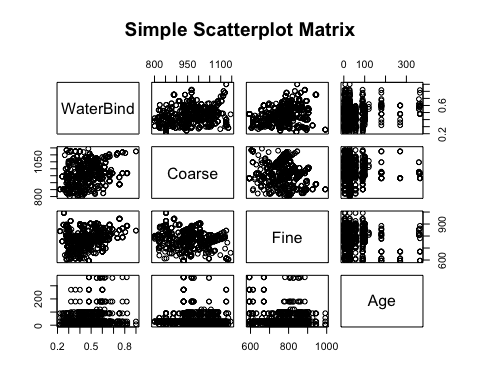
plot(WaterBind, Strength)



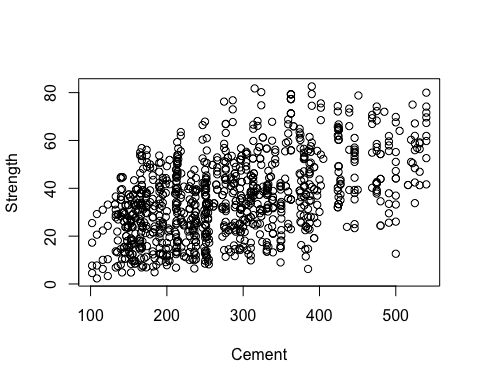
res <- cor(xy)  
round(res, 3)

## Cement BlastF FlyAsh Water SuperP Coarse Fine Age Strength  
## Cement 1.000 -0.275 -0.397 -0.082 0.093 -0.109 -0.223 0.082 0.498  
## BlastF -0.275 1.000 -0.324 0.107 0.043 -0.284 -0.282 -0.044 0.135  
## FlyAsh -0.397 -0.324 1.000 -0.257 0.377 -0.010 0.079 -0.154 -0.106  
## Water -0.082 0.107 -0.257 1.000 -0.657 -0.182 -0.451 0.278 -0.290  
## SuperP 0.093 0.043 0.377 -0.657 1.000 -0.266 0.223 -0.193 0.366  
## Coarse -0.109 -0.284 -0.010 -0.182 -0.266 1.000 -0.179 -0.003 -0.165  
## Fine -0.223 -0.282 0.079 -0.451 0.223 -0.179 1.000 -0.156 -0.167  
## Age 0.082 -0.044 -0.154 0.278 -0.193 -0.003 -0.156 1.000 0.329  
## Strength 0.498 0.135 -0.106 -0.290 0.366 -0.165 -0.167 0.329 1.000  
## WaterBind -0.458 -0.269 -0.161 0.547 -0.628 0.225 0.231 0.157 -0.618  
## WaterBind  
## Cement -0.458  
## BlastF -0.269  
## FlyAsh -0.161  
## Water 0.547  
## SuperP -0.628  
## Coarse 0.225  
## Fine 0.231  
## Age 0.157  
## Strength -0.618  
## WaterBind 1.000

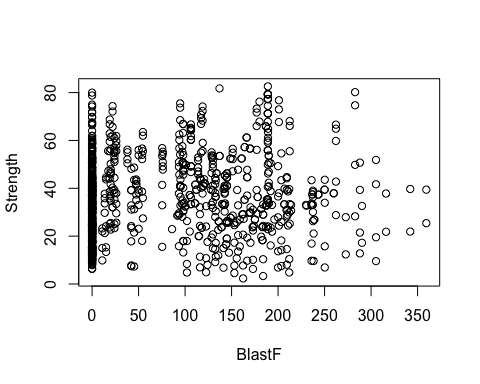
# Basic Scatterplot Matrix  
pairs(~WaterBind+Coarse+Fine+Age,data=xy,  
 main="Simple Scatterplot Matrix")



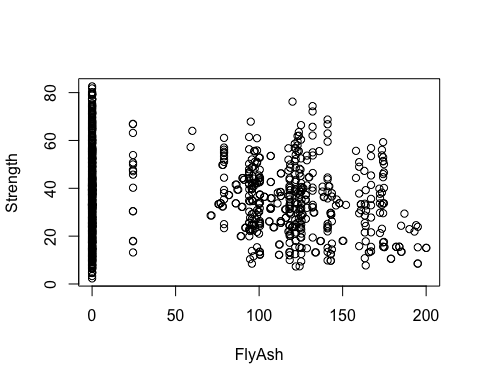
plot(Cement, Strength)



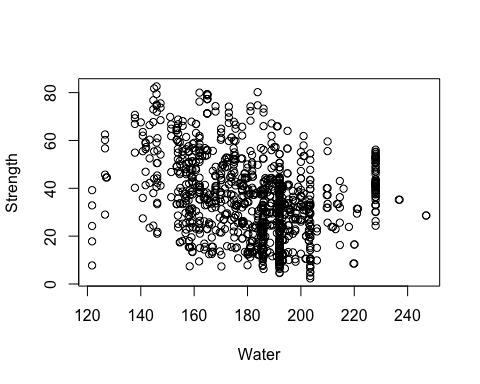
plot(BlastF, Strength)



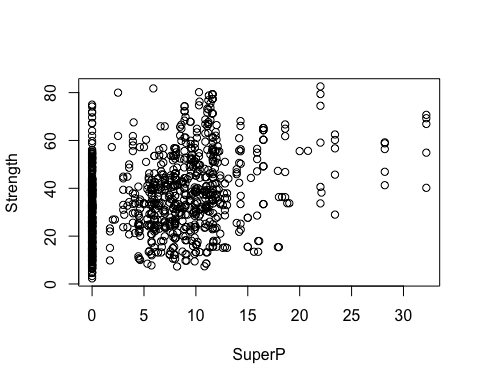
plot(FlyAsh, Strength)



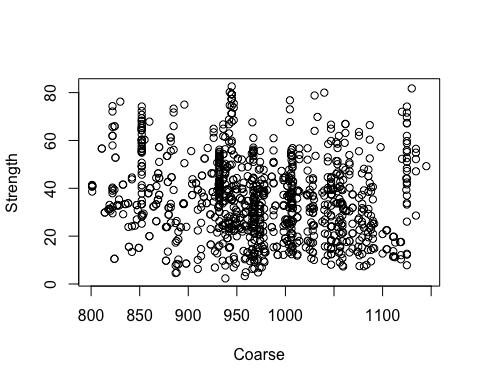
plot(Water, Strength)



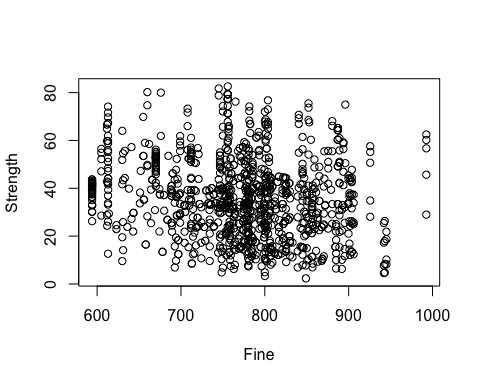
plot(SuperP, Strength)



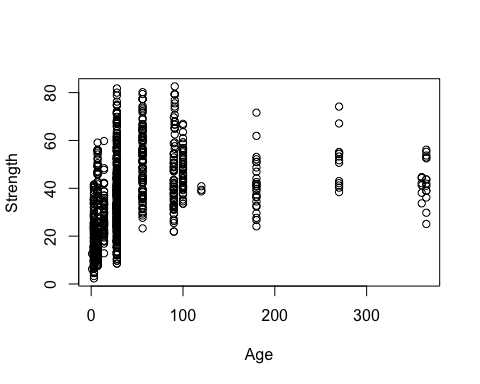
plot(Coarse, Strength)



plot(Fine, Strength)



plot(Age, Strength)



WBSquared = WaterBind\*WaterBind  
RegBIC<-lm(Strength~Fine+Age+WBSquared)  
summary(RegBIC)

##   
## Call:  
## lm(formula = Strength ~ Fine + Age + WBSquared)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -34.452 -7.803 0.241 7.302 36.991   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 37.782930 3.610414 10.46 < 2e-16 \*\*\*  
## Fine 0.016783 0.004755 3.53 0.000434 \*\*\*  
## Age 0.116783 0.005883 19.85 < 2e-16 \*\*\*  
## WBSquared -87.722479 2.976041 -29.48 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 11.54 on 1026 degrees of freedom  
## Multiple R-squared: 0.5246, Adjusted R-squared: 0.5232   
## F-statistic: 377.3 on 3 and 1026 DF, p-value: < 2.2e-16

AgeSquared = Age\*Age  
RegA2<-lm(Strength~Fine+WaterBind+Age+AgeSquared)  
summary(RegA2)

##   
## Call:  
## lm(formula = Strength ~ Fine + WaterBind + Age + AgeSquared)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -30.572 -6.311 -0.566 5.750 35.916   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 5.847e+01 2.921e+00 20.018 <2e-16 \*\*\*  
## Fine 9.024e-03 3.818e-03 2.364 0.0183 \*   
## WaterBind -8.797e+01 2.384e+00 -36.896 <2e-16 \*\*\*  
## Age 3.465e-01 1.257e-02 27.555 <2e-16 \*\*\*  
## AgeSquared -7.723e-04 3.947e-05 -19.564 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 9.342 on 1025 degrees of freedom  
## Multiple R-squared: 0.6885, Adjusted R-squared: 0.6873   
## F-statistic: 566.4 on 4 and 1025 DF, p-value: < 2.2e-16

AgeCubed = Age\*Age\*Age  
RegAgeCubed<-lm(Strength~Fine+WaterBind+Age+AgeCubed)  
summary(RegAgeCubed)

##   
## Call:  
## lm(formula = Strength ~ Fine + WaterBind + Age + AgeCubed)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -32.195 -6.276 -0.156 6.190 36.214   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 5.906e+01 3.033e+00 19.473 < 2e-16 \*\*\*  
## Fine 1.122e-02 3.961e-03 2.832 0.00472 \*\*   
## WaterBind -8.923e+01 2.473e+00 -36.082 < 2e-16 \*\*\*  
## Age 2.519e-01 9.366e-03 26.890 < 2e-16 \*\*\*  
## AgeCubed -1.418e-06 8.474e-08 -16.729 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 9.703 on 1025 degrees of freedom  
## Multiple R-squared: 0.664, Adjusted R-squared: 0.6626   
## F-statistic: 506.3 on 4 and 1025 DF, p-value: < 2.2e-16

LogAge = log(Age+1)  
RegLogAge<-lm(Strength~Fine+WaterBind+LogAge)  
summary(RegLogAge)

##   
## Call:  
## lm(formula = Strength ~ Fine + WaterBind + LogAge)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -25.735 -5.480 -0.067 4.681 32.850   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 37.597671 2.807246 13.393 < 2e-16 \*\*\*  
## Fine 0.012161 0.003439 3.536 0.000425 \*\*\*  
## WaterBind -87.801703 2.136904 -41.088 < 2e-16 \*\*\*  
## LogAge 9.097508 0.242220 37.559 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 8.526 on 1026 degrees of freedom  
## Multiple R-squared: 0.7403, Adjusted R-squared: 0.7395   
## F-statistic: 974.7 on 3 and 1026 DF, p-value: < 2.2e-16

LogWB = log(WaterBind+1)  
RegLogWB<-lm(Strength~Fine+LogWB+LogAge)  
summary(RegLogWB)

##   
## Call:  
## lm(formula = Strength ~ Fine + LogWB + LogAge)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -26.157 -5.250 -0.013 4.756 32.728   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.745e+01 2.771e+00 17.128 < 2e-16 \*\*\*  
## Fine 1.069e-02 3.346e-03 3.197 0.00143 \*\*   
## LogWB -1.314e+02 3.076e+00 -42.706 < 2e-16 \*\*\*  
## LogAge 9.131e+00 2.364e-01 38.621 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 8.321 on 1026 degrees of freedom  
## Multiple R-squared: 0.7526, Adjusted R-squared: 0.7519   
## F-statistic: 1040 on 3 and 1026 DF, p-value: < 2.2e-16

FineSquared = Fine\*Fine  
RegBIC<-lm(Strength~FineSquared+LogWB+LogAge)  
summary(RegBIC)

##   
## Call:  
## lm(formula = Strength ~ FineSquared + LogWB + LogAge)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -26.108 -5.223 -0.017 4.763 32.767   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 5.162e+01 1.768e+00 29.201 < 2e-16 \*\*\*  
## FineSquared 6.774e-06 2.183e-06 3.103 0.00197 \*\*   
## LogWB -1.313e+02 3.075e+00 -42.693 < 2e-16 \*\*\*  
## LogAge 9.122e+00 2.362e-01 38.614 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 8.323 on 1026 degrees of freedom  
## Multiple R-squared: 0.7525, Adjusted R-squared: 0.7518   
## F-statistic: 1040 on 3 and 1026 DF, p-value: < 2.2e-16

OneOverAge = 1/Age  
RegBIC<-lm(Strength~FineSquared+LogWB+OneOverAge)  
summary(RegBIC)

##   
## Call:  
## lm(formula = Strength ~ FineSquared + LogWB + OneOverAge)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -31.275 -6.508 -0.024 6.225 42.710   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 8.752e+01 1.860e+00 47.053 <2e-16 \*\*\*  
## FineSquared 2.488e-06 2.569e-06 0.968 0.333   
## LogWB -1.239e+02 3.618e+00 -34.242 <2e-16 \*\*\*  
## OneOverAge -7.658e+01 2.748e+00 -27.869 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 9.835 on 1026 degrees of freedom  
## Multiple R-squared: 0.6544, Adjusted R-squared: 0.6534   
## F-statistic: 647.6 on 3 and 1026 DF, p-value: < 2.2e-16

BW <- 1/WaterBind  
RegBIC<-lm(Strength~Fine+BW+LogAge)  
summary(RegBIC)

##   
## Call:  
## lm(formula = Strength ~ Fine + BW + LogAge)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -26.469 -4.729 0.350 4.547 33.841   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -3.314e+01 2.893e+00 -11.457 <2e-16 \*\*\*  
## Fine -9.611e-04 3.124e-03 -0.308 0.758   
## BW 1.711e+01 3.700e-01 46.245 <2e-16 \*\*\*  
## LogAge 9.169e+00 2.244e-01 40.863 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 7.896 on 1026 degrees of freedom  
## Multiple R-squared: 0.7772, Adjusted R-squared: 0.7766   
## F-statistic: 1193 on 3 and 1026 DF, p-value: < 2.2e-16

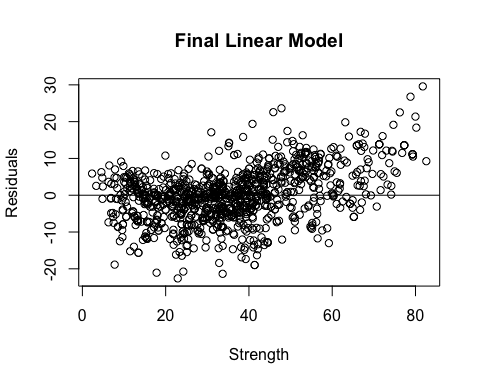
LogBW <- log(1/WaterBind)  
RegBW<-lm(Strength~SuperP+Fine+LogBW+LogAge)  
summary(RegBW)

##   
## Call:  
## lm(formula = Strength ~ SuperP + Fine + LogBW + LogAge)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -28.928 -4.836 0.165 4.307 33.900   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -43.705285 3.347364 -13.057 < 2e-16 \*\*\*  
## SuperP -0.413537 0.061430 -6.732 2.78e-11 \*\*\*  
## Fine 0.017682 0.003546 4.987 7.20e-07 \*\*\*  
## LogBW 47.379106 1.297408 36.518 < 2e-16 \*\*\*  
## LogAge 9.288272 0.223580 41.543 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 7.843 on 1025 degrees of freedom  
## Multiple R-squared: 0.7804, Adjusted R-squared: 0.7796   
## F-statistic: 910.8 on 4 and 1025 DF, p-value: < 2.2e-16

FineSuperP <- Fine\*SuperP  
FineBW<-Fine\*BW  
FineCement<-Fine\*Cement  
FineFly<-Fine\*FlyAsh  
FineW<-Fine\*Water  
FineB<-Fine\*BlastF  
FSW<-Fine\*SuperP\*Water  
FFW<-Fine\*FlyAsh\*Water  
RegInteract<-lm(Strength~FineSuperP+FineBW+FSW+FFW+FineCement+FineB+FineFly+FineW+SuperP+Fine+LogBW+LogAge)  
summary(RegInteract)

##   
## Call:  
## lm(formula = Strength ~ FineSuperP + FineBW + FSW + FFW + FineCement +   
## FineB + FineFly + FineW + SuperP + Fine + LogBW + LogAge)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -22.608 -4.252 -0.067 4.452 29.579   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -3.674e+01 9.204e+00 -3.991 7.04e-05 \*\*\*  
## FineSuperP -2.979e-03 6.117e-04 -4.870 1.29e-06 \*\*\*  
## FineBW -3.925e-03 5.583e-03 -0.703 0.482224   
## FSW 8.428e-06 4.006e-06 2.104 0.035658 \*   
## FFW -4.497e-07 2.867e-07 -1.569 0.117036   
## FineCement 7.746e-05 4.169e-05 1.858 0.063454 .   
## FineB 4.191e-05 4.201e-05 0.998 0.318720   
## FineFly 7.674e-05 6.411e-05 1.197 0.231591   
## FineW -1.368e-04 9.648e-05 -1.418 0.156538   
## SuperP 1.244e+00 4.174e-01 2.980 0.002955 \*\*   
## Fine 3.116e-02 1.383e-02 2.253 0.024467 \*   
## LogBW 3.165e+01 9.153e+00 3.458 0.000568 \*\*\*  
## LogAge 9.280e+00 2.085e-01 44.512 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 7.163 on 1017 degrees of freedom  
## Multiple R-squared: 0.8183, Adjusted R-squared: 0.8162   
## F-statistic: 381.7 on 12 and 1017 DF, p-value: < 2.2e-16

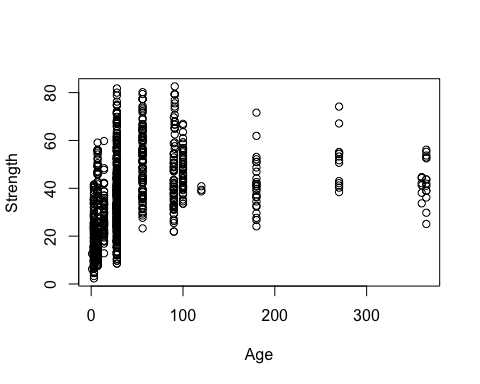
residuals2<-residuals(RegInteract)  
plot(xy$Strength, residuals2,ylab = "Residuals", xlab = "Strength",main = "Final Linear Model")   
abline(0, 0) # the horizon



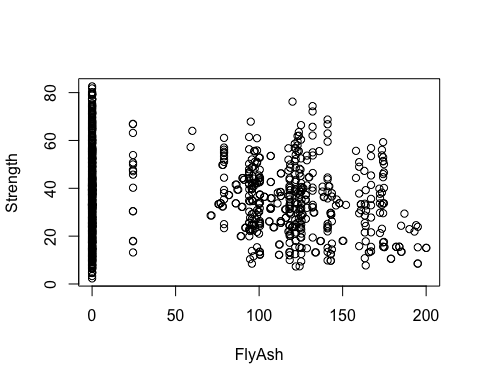
anova(RegInteract)

## Analysis of Variance Table  
##   
## Response: Strength  
## Df Sum Sq Mean Sq F value Pr(>F)   
## FineSuperP 1 36132 36132 704.2859 < 2.2e-16 \*\*\*  
## FineBW 1 61589 61589 1200.5016 < 2.2e-16 \*\*\*  
## FSW 1 2085 2085 40.6430 2.767e-10 \*\*\*  
## FFW 1 6373 6373 124.2218 < 2.2e-16 \*\*\*  
## FineCement 1 3360 3360 65.4940 1.653e-15 \*\*\*  
## FineB 1 691 691 13.4769 0.0002541 \*\*\*  
## FineFly 1 5601 5601 109.1847 < 2.2e-16 \*\*\*  
## FineW 1 14139 14139 275.6092 < 2.2e-16 \*\*\*  
## SuperP 1 148 148 2.8759 0.0902187 .   
## Fine 1 1874 1874 36.5265 2.111e-09 \*\*\*  
## LogBW 1 1361 1361 26.5261 3.121e-07 \*\*\*  
## LogAge 1 101645 101645 1981.2941 < 2.2e-16 \*\*\*  
## Residuals 1017 52175 51   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

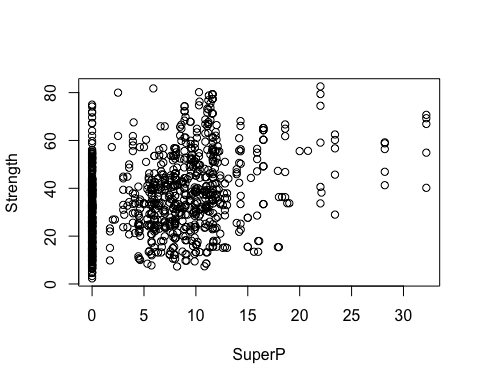
plot(Age, Strength)



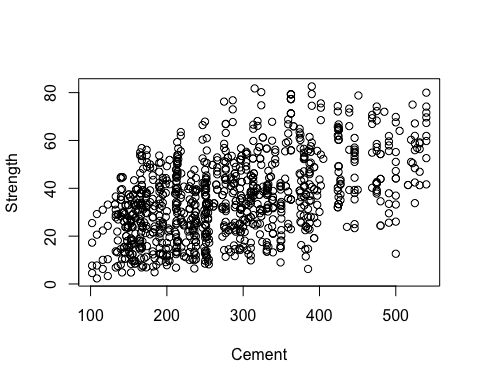
plot(FlyAsh, Strength)



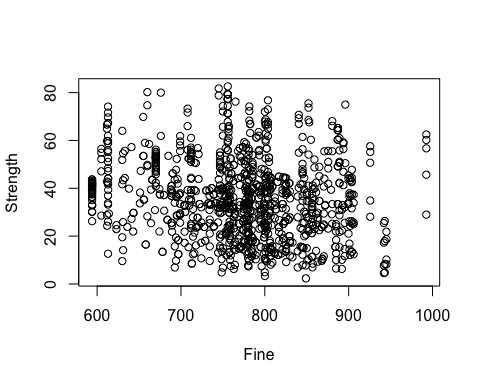
plot(SuperP, Strength)



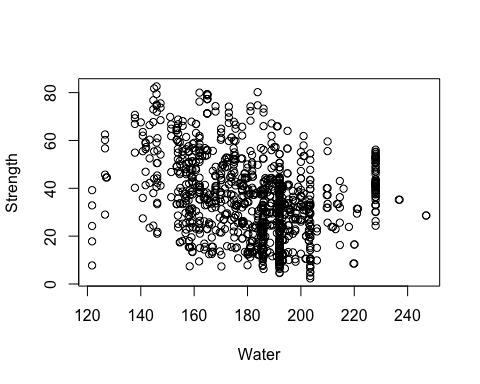
plot(Cement, Strength)



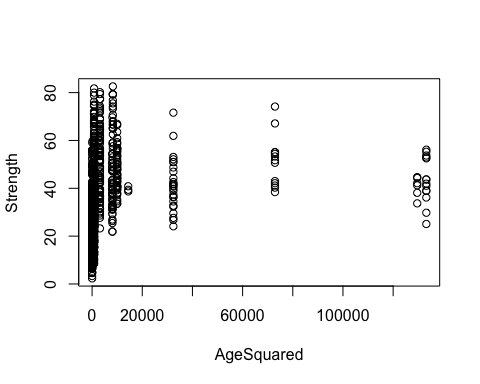
plot(Fine, Strength)



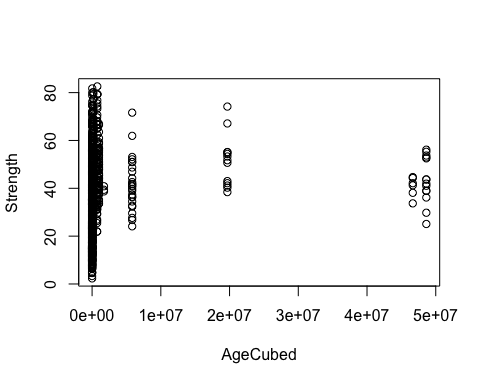
plot(Water, Strength)



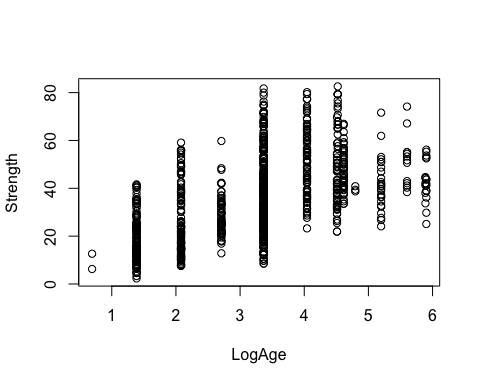
plot(AgeSquared, Strength)



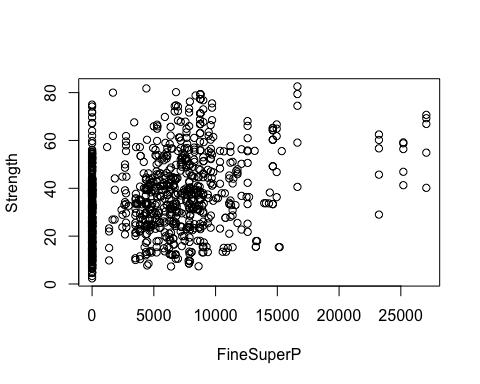
plot(AgeCubed,Strength)



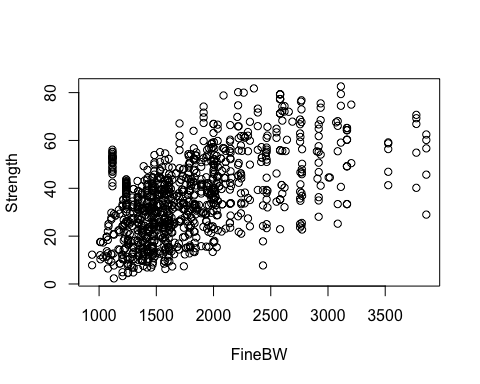
plot(LogAge, Strength)



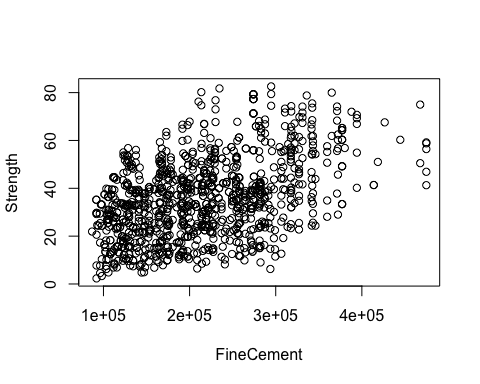
LogOneOverAge=log(1/Age)  
  
#FineSuperP, FineBW, FineCement, FineFly, FSW, FFW, LogOneOverAge, SuperP, Strength)  
  
#par(mfrow = c(3,4)) # Create a 2 x 3 plotting matrix  
# The next 4 plots created will be plotted next to each other  
  
# Plot 1  
plot(FineSuperP,Strength)



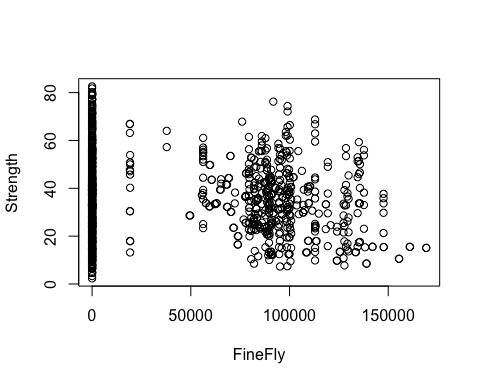
plot(FineBW,Strength)



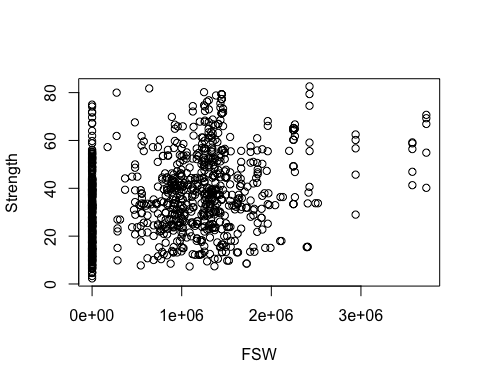
plot(FineCement,Strength)



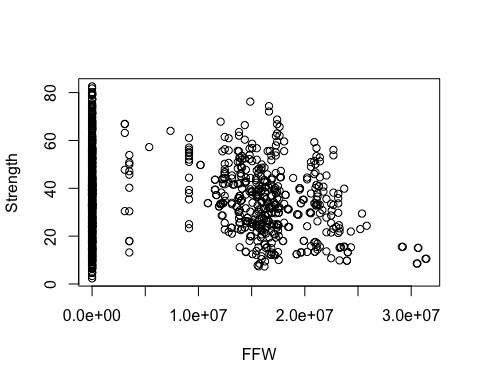
plot(FineFly,Strength)



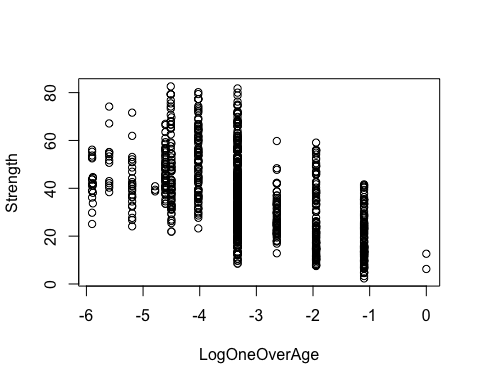
plot(FSW,Strength)



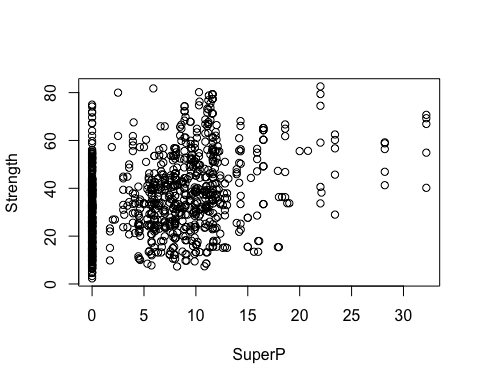
plot(FFW,Strength)



plot(LogOneOverAge, Strength)



plot(SuperP, Strength)



#Scale the addtives  
library(dplyr)  
library(tidyverse)  
  
## Add new column using mutate() in dplyr  
# Calculate water to binder ratio as WaterBind = W / (cement+flyash+Blastf)  
OneOverLogAge=1/LogAge  
LogOneOverAge=log(1/Age)  
  
xy<- mutate(xy, WaterBind = (Water/(Cement + FlyAsh + BlastF + SuperP)))  
xy

## # A tibble: 1,030 x 10  
## Cement BlastF FlyAsh Water SuperP Coarse Fine Age Strength WaterBind  
## <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 540 0 0 162 2.5 1040 676 28 80.0 0.299  
## 2 540 0 0 162 2.5 1055 676 28 61.9 0.299  
## 3 332. 142. 0 228 0 932 594 270 40.3 0.48   
## 4 332. 142. 0 228 0 932 594 365 41.1 0.48   
## 5 199. 132. 0 192 0 978. 826. 360 44.3 0.580  
## 6 266 114 0 228 0 932 670 90 47.0 0.6   
## 7 380 95 0 228 0 932 594 365 43.7 0.48   
## 8 380 95 0 228 0 932 594 28 36.4 0.48   
## 9 266 114 0 228 0 932 670 28 45.9 0.6   
## 10 475 0 0 228 0 932 594 28 39.3 0.48   
## # … with 1,020 more rows

WaterBind<-xy$WaterBind

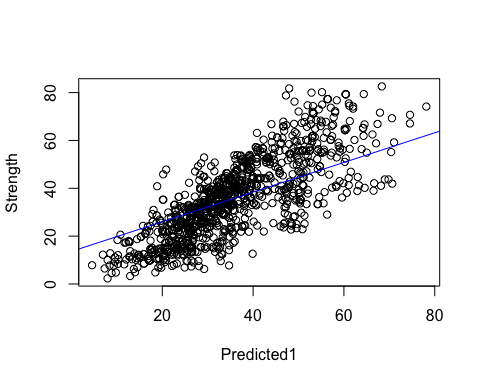
# Regress simple linear model  
WB = WaterBind  
BW = 1/WB  
Reg1<-lm(Strength ~ WB + Cement + BlastF + FlyAsh + Water + SuperP + Coarse + Fine + Age)  
summary(Reg1)

##   
## Call:  
## lm(formula = Strength ~ WB + Cement + BlastF + FlyAsh + Water +   
## SuperP + Coarse + Fine + Age)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -28.714 -6.465 0.578 6.555 33.804   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -8.548970 26.786228 -0.319 0.749674   
## WB -37.825735 10.981248 -3.445 0.000595 \*\*\*  
## Cement 0.079825 0.014349 5.563 3.38e-08 \*\*\*  
## BlastF 0.062463 0.015685 3.982 7.31e-05 \*\*\*  
## FlyAsh 0.039942 0.018733 2.132 0.033228 \*   
## Water -0.057234 0.048242 -1.186 0.235747   
## SuperP 0.268000 0.093200 2.876 0.004117 \*\*   
## Coarse 0.017781 0.009345 1.903 0.057354 .   
## Fine 0.024460 0.010719 2.282 0.022702 \*   
## Age 0.113384 0.005404 20.980 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 10.34 on 1020 degrees of freedom  
## Multiple R-squared: 0.6199, Adjusted R-squared: 0.6165   
## F-statistic: 184.8 on 9 and 1020 DF, p-value: < 2.2e-16

# Reg1 predicted  
Predicted1<-predict.lm(Reg1)  
  
# model actual vs predicted  
line<-lm(Predicted1~Strength)  
coef(line)

## (Intercept) Strength   
## 13.6148470 0.6198864

# predicted versus actuals  
plot(Predicted1, Strength)  
abline(13.613, 0.6199, col=4)



## CReate full model Reg2  
FineSuperP <- Fine\*SuperP  
FineBW<-Fine\*BW  
FineCement<-Fine\*Cement  
FineFly<-Fine\*FlyAsh  
FineW<-Fine\*Water  
FineB<-Fine\*BlastF  
FSW<-Fine\*SuperP\*Water  
FFW<-Fine\*FlyAsh\*Water  
  
### adding more inverse  
# OneOverCement,OneOverFlyAsh,OneOverWater,OneOverBlast, OneOverSuperP, OneOverCoarse, OneOverFine, OneOverAge  
  
# no need for 1/WB = BW or 1/WB^2 = BW^2  
OneOverCement=1/(1+Cement)  
OneOverFlyAsh=1/(1+FlyAsh)  
OneOverWater=1/(1+Water)  
OneOverBlastF=1/(1+BlastF)  
OneOverSuperP=1/(1+SuperP)  
OneOverCoarse=1/(1+Coarse)  
OneOverFine=1/(1+Fine)  
OneOverAge=1/(1+Age)  
  
### adding inverse square  
# OneOverCement,OneOverFlyAsh,OneOverWater,OneOverBlast, OneOverSuperP, OneOverCoarse, OneOverFine, OneOverAge  
  
OneOverCement2=1/(1+Cement)^2  
OneOverFlyAsh2=1/(1+FlyAsh)^2  
OneOverWater2=1/(1+Water)^2  
OneOverBlastF2=1/(1+BlastF)^2  
OneOverSuperP2=1/(1+SuperP)^2  
OneOverCoarse2=1/(1+Coarse)^2  
OneOverFine2=1/(1+Fine)^2  
OneOverAge2=1/(1+Age)^2  
  
### adding more squares  
# SquaredCement, SquaredFlyAsh, SquaredWater, SquaredBlastF, SquaredSuperP, SquaredCoarse, SquaredFIne, SquaredAge, Squared BW  
SquaredCement=Cement^2  
SquaredFlyAsh=FlyAsh^2  
SquaredWater=Water^2  
SquaredBlastF=BlastF^2  
SquaredSuperP=SuperP^2  
SquaredCoarse=Coarse^2  
SquaredFine=Fine^2  
SquaredAge=Age^2  
SquaredBW=WaterBind^2  
  
### adding more log  
# LogCement, LogFlyAsh, LogWater, LogBlastF, LogSuperP, LogCoarse, LogFine, LogAge, LogBW  
LogCement=log(Cement+1)  
LogFlyAsh=log(FlyAsh+1)  
LogWater=log(Water+1)  
LogBlastF=log(BlastF+1)  
LogSuperP=log(SuperP+1)  
LogCoarse=log(Coarse+1)  
LogFine=log(Fine+1)  
LogAge=log(Age+1)  
LogBW=log(BW+1)  
LogWB=log(WB+1)

# Reduced model 4 originally made by hand has higher error  
  
Reg4<-lm(formula = Strength ~ WB + FSW + FFW + FineCement + FineB +   
 FineW + Fine + LogWB + LogAge + OneOverFlyAsh + OneOverWater +   
 OneOverBlastF + OneOverFine + OneOverAge + SquaredFlyAsh +   
 SquaredWater + SquaredBlastF + SquaredSuperP + SquaredCoarse +   
 SquaredFine + SquaredAge + SquaredBW + LogCement + LogFlyAsh +   
 LogWater + LogSuperP + LogCoarse + LogFine)  
summary(Reg4)

##   
## Call:  
## lm(formula = Strength ~ WB + FSW + FFW + FineCement + FineB +   
## FineW + Fine + LogWB + LogAge + OneOverFlyAsh + OneOverWater +   
## OneOverBlastF + OneOverFine + OneOverAge + SquaredFlyAsh +   
## SquaredWater + SquaredBlastF + SquaredSuperP + SquaredCoarse +   
## SquaredFine + SquaredAge + SquaredBW + LogCement + LogFlyAsh +   
## LogWater + LogSuperP + LogCoarse + LogFine)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -26.9493 -3.7268 0.0589 3.8745 23.9077   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.852e+05 5.611e+04 3.302 0.000995 \*\*\*  
## WB 6.242e+03 9.916e+02 6.295 4.58e-10 \*\*\*  
## FSW -1.757e-05 2.098e-06 -8.373 < 2e-16 \*\*\*  
## FFW -6.739e-07 2.627e-07 -2.566 0.010444 \*   
## FineCement -3.546e-04 4.557e-05 -7.782 1.77e-14 \*\*\*  
## FineB -2.225e-04 4.827e-05 -4.609 4.57e-06 \*\*\*  
## FineW -1.307e-03 2.454e-04 -5.324 1.25e-07 \*\*\*  
## Fine 3.595e+01 1.207e+01 2.979 0.002966 \*\*   
## LogWB -7.964e+03 1.190e+03 -6.694 3.62e-11 \*\*\*  
## LogAge 8.889e+00 5.089e-01 17.468 < 2e-16 \*\*\*  
## OneOverFlyAsh -4.464e+01 8.429e+00 -5.296 1.45e-07 \*\*\*  
## OneOverWater -1.600e+05 2.035e+04 -7.862 9.77e-15 \*\*\*  
## OneOverBlastF 5.953e+00 1.037e+00 5.742 1.24e-08 \*\*\*  
## OneOverFine -7.738e+06 2.320e+06 -3.336 0.000882 \*\*\*  
## OneOverAge -1.425e+01 6.112e+00 -2.331 0.019962 \*   
## SquaredFlyAsh -3.593e-04 1.097e-04 -3.276 0.001090 \*\*   
## SquaredWater 7.812e-03 9.139e-04 8.548 < 2e-16 \*\*\*  
## SquaredBlastF -2.534e-04 3.418e-05 -7.415 2.59e-13 \*\*\*  
## SquaredSuperP 2.345e-02 6.040e-03 3.883 0.000110 \*\*\*  
## SquaredCoarse 7.220e-05 1.651e-05 4.373 1.36e-05 \*\*\*  
## SquaredFine -7.358e-03 2.622e-03 -2.806 0.005113 \*\*   
## SquaredAge -9.275e-05 1.431e-05 -6.481 1.42e-10 \*\*\*  
## SquaredBW -1.217e+03 2.190e+02 -5.557 3.52e-08 \*\*\*  
## LogCement 1.151e+01 3.226e+00 3.568 0.000376 \*\*\*  
## LogFlyAsh -1.311e+01 2.334e+00 -5.616 2.54e-08 \*\*\*  
## LogWater -1.083e+03 1.969e+02 -5.500 4.81e-08 \*\*\*  
## LogSuperP 8.809e+00 1.037e+00 8.492 < 2e-16 \*\*\*  
## LogCoarse -1.145e+02 3.146e+01 -3.639 0.000288 \*\*\*  
## LogFine -2.871e+04 9.189e+03 -3.125 0.001831 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.893 on 1001 degrees of freedom  
## Multiple R-squared: 0.8789, Adjusted R-squared: 0.8756   
## F-statistic: 259.6 on 28 and 1001 DF, p-value: < 2.2e-16

# Note: All intercations using lm(d~(a+b+c)^2) which gives a, b, c, ab, bc, and ac  
  
RegFull<-lm(Strength~(WB+BW+Cement+FlyAsh+Water+BlastF+SuperP+Fine+Coarse+Age)^2+LogBW+LogWB+LogAge+OneOverCement+OneOverFlyAsh+OneOverWater+OneOverBlastF+OneOverSuperP+OneOverCoarse+OneOverFine+OneOverAge+OneOverCement2+OneOverFlyAsh2+OneOverWater2+OneOverBlastF2+OneOverSuperP2+OneOverCoarse2+OneOverFine2+OneOverAge2+SquaredCement+SquaredFlyAsh+SquaredWater+SquaredBlastF+SquaredSuperP+SquaredCoarse+SquaredFine+SquaredAge+SquaredBW+LogCement+LogFlyAsh+LogWater+LogBlastF+LogSuperP+LogCoarse+LogFine)  
  
summary(RegFull)

##   
## Call:  
## lm(formula = Strength ~ (WB + BW + Cement + FlyAsh + Water +   
## BlastF + SuperP + Fine + Coarse + Age)^2 + LogBW + LogWB +   
## LogAge + OneOverCement + OneOverFlyAsh + OneOverWater + OneOverBlastF +   
## OneOverSuperP + OneOverCoarse + OneOverFine + OneOverAge +   
## OneOverCement2 + OneOverFlyAsh2 + OneOverWater2 + OneOverBlastF2 +   
## OneOverSuperP2 + OneOverCoarse2 + OneOverFine2 + OneOverAge2 +   
## SquaredCement + SquaredFlyAsh + SquaredWater + SquaredBlastF +   
## SquaredSuperP + SquaredCoarse + SquaredFine + SquaredAge +   
## SquaredBW + LogCement + LogFlyAsh + LogWater + LogBlastF +   
## LogSuperP + LogCoarse + LogFine)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -26.0617 -3.0723 -0.2553 3.1652 23.8616   
##   
## Coefficients: (4 not defined because of singularities)  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.036e+07 1.842e+06 -5.623 2.47e-08 \*\*\*  
## WB 2.311e+04 2.691e+04 0.859 0.390795   
## BW -7.341e+02 1.260e+03 -0.583 0.560357   
## Cement -5.561e+00 5.298e+00 -1.050 0.294201   
## FlyAsh -2.497e+00 5.445e+00 -0.459 0.646603   
## Water -1.618e+02 2.534e+02 -0.638 0.523406   
## BlastF -4.422e+00 5.062e+00 -0.874 0.382547   
## SuperP 5.363e+00 1.029e+01 0.521 0.602546   
## Fine -1.322e+03 2.280e+02 -5.801 9.02e-09 \*\*\*  
## Coarse -3.857e+01 1.500e+01 -2.571 0.010297 \*   
## Age -7.747e-01 4.078e-01 -1.900 0.057775 .   
## LogBW 3.556e+03 8.305e+03 0.428 0.668578   
## LogWB -2.680e+04 4.224e+04 -0.635 0.525882   
## LogAge 1.482e+01 2.568e+00 5.769 1.08e-08 \*\*\*  
## OneOverCement 3.016e+04 9.352e+04 0.322 0.747157   
## OneOverFlyAsh -1.142e+03 3.975e+03 -0.287 0.773916   
## OneOverWater 5.450e+06 7.826e+06 0.696 0.486345   
## OneOverBlastF -2.090e+02 2.198e+02 -0.951 0.341943   
## OneOverSuperP 4.833e+01 1.703e+02 0.284 0.776633   
## OneOverCoarse -3.784e+07 1.378e+07 -2.745 0.006161 \*\*   
## OneOverFine 7.818e+08 1.318e+08 5.930 4.24e-09 \*\*\*  
## OneOverAge 4.536e+01 3.298e+01 1.375 0.169380   
## OneOverCement2 -3.743e+05 2.544e+06 -0.147 0.883072   
## OneOverFlyAsh2 8.121e+02 3.358e+03 0.242 0.808923   
## OneOverWater2 -1.202e+08 1.682e+08 -0.714 0.475124   
## OneOverBlastF2 1.929e+02 1.826e+02 1.056 0.291143   
## OneOverSuperP2 -4.760e+00 8.660e+01 -0.055 0.956174   
## OneOverCoarse2 9.002e+09 3.275e+09 2.749 0.006095 \*\*   
## OneOverFine2 -7.445e+10 1.243e+10 -5.990 2.98e-09 \*\*\*  
## OneOverAge2 -7.508e+01 5.125e+01 -1.465 0.143292   
## SquaredCement 4.993e-04 2.769e-03 0.180 0.856924   
## SquaredFlyAsh -3.907e-03 3.550e-03 -1.101 0.271299   
## SquaredWater 1.158e-01 1.736e-01 0.667 0.504902   
## SquaredBlastF -1.645e-04 2.669e-03 -0.062 0.950876   
## SquaredSuperP 1.262e-01 5.729e-02 2.203 0.027848 \*   
## SquaredCoarse 1.027e-02 3.924e-03 2.617 0.009017 \*\*   
## SquaredFine 2.136e-01 3.722e-02 5.739 1.29e-08 \*\*\*  
## SquaredAge 2.581e-04 7.264e-05 3.553 0.000399 \*\*\*  
## SquaredBW -3.795e+03 4.050e+03 -0.937 0.348999   
## LogCement 3.261e+02 6.062e+02 0.538 0.590755   
## LogFlyAsh -9.845e+01 1.570e+02 -0.627 0.530766   
## LogWater 4.651e+04 6.754e+04 0.689 0.491217   
## LogBlastF -3.657e+00 9.852e+00 -0.371 0.710546   
## LogSuperP 3.766e+01 4.121e+01 0.914 0.361147   
## LogCoarse NA NA NA NA   
## LogFine 1.531e+06 2.608e+05 5.870 6.04e-09 \*\*\*  
## WB:BW NA NA NA NA   
## WB:Cement 9.901e+00 4.595e+00 2.155 0.031430 \*   
## WB:FlyAsh 9.929e+00 4.779e+00 2.078 0.038015 \*   
## WB:Water -8.700e+00 3.978e+00 -2.187 0.029007 \*   
## WB:BlastF 9.881e+00 4.619e+00 2.139 0.032674 \*   
## WB:SuperP NA NA NA NA   
## WB:Fine -9.558e-02 4.339e-01 -0.220 0.825683   
## WB:Coarse -3.925e-01 3.691e-01 -1.063 0.287842   
## WB:Age -4.190e-02 1.079e-01 -0.388 0.697891   
## BW:Cement 7.576e-01 7.960e-01 0.952 0.341444   
## BW:FlyAsh 7.321e-01 7.622e-01 0.960 0.337076   
## BW:Water NA NA NA NA   
## BW:BlastF 7.291e-01 7.970e-01 0.915 0.360504   
## BW:SuperP -2.868e+00 1.345e+00 -2.132 0.033292 \*   
## BW:Fine -6.471e-02 1.438e-01 -0.450 0.652842   
## BW:Coarse -6.648e-02 1.439e-01 -0.462 0.644235   
## BW:Age 1.145e-01 7.246e-02 1.581 0.114265   
## Cement:FlyAsh -3.509e-04 5.062e-03 -0.069 0.944741   
## Cement:Water -1.570e-03 1.409e-02 -0.111 0.911305   
## Cement:BlastF -3.253e-04 5.269e-03 -0.062 0.950784   
## Cement:SuperP 1.204e-02 1.001e-02 1.203 0.229364   
## Cement:Fine -2.824e-04 9.075e-04 -0.311 0.755686   
## Cement:Coarse -5.615e-04 8.808e-04 -0.638 0.523954   
## Cement:Age -3.991e-04 4.114e-04 -0.970 0.332259   
## FlyAsh:Water -2.508e-03 1.363e-02 -0.184 0.854103   
## FlyAsh:BlastF -2.509e-04 5.141e-03 -0.049 0.961092   
## FlyAsh:SuperP 9.564e-03 9.591e-03 0.997 0.318953   
## FlyAsh:Fine 2.294e-06 8.881e-04 0.003 0.997939   
## FlyAsh:Coarse -7.686e-04 8.579e-04 -0.896 0.370511   
## FlyAsh:Age 6.795e-05 4.182e-04 0.162 0.870950   
## Water:BlastF -1.725e-03 1.402e-02 -0.123 0.902075   
## Water:SuperP -4.256e-02 2.629e-02 -1.619 0.105832   
## Water:Fine -4.692e-03 2.685e-03 -1.748 0.080851 .   
## Water:Coarse -3.922e-03 2.437e-03 -1.609 0.107861   
## Water:Age 1.258e-03 1.116e-03 1.127 0.259860   
## BlastF:SuperP 1.380e-02 9.948e-03 1.387 0.165646   
## BlastF:Fine -3.452e-05 9.048e-04 -0.038 0.969572   
## BlastF:Coarse -6.202e-04 8.751e-04 -0.709 0.478679   
## BlastF:Age -1.584e-04 4.172e-04 -0.380 0.704219   
## SuperP:Fine -1.419e-03 3.019e-03 -0.470 0.638469   
## SuperP:Coarse -4.303e-04 2.978e-03 -0.144 0.885154   
## SuperP:Age -2.265e-03 2.086e-03 -1.086 0.277878   
## Fine:Coarse -6.531e-04 3.007e-04 -2.172 0.030122 \*   
## Fine:Age 2.201e-04 1.373e-04 1.603 0.109218   
## Coarse:Age 1.297e-04 9.971e-05 1.300 0.193788   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.17 on 943 degrees of freedom  
## Multiple R-squared: 0.9122, Adjusted R-squared: 0.9042   
## F-statistic: 114 on 86 and 943 DF, p-value: < 2.2e-16

# Reg1 AIC  
step(RegFull, direction="both", trace=0, k=2)

##   
## Call:  
## lm(formula = Strength ~ WB + BW + Cement + FlyAsh + Water + BlastF +   
## SuperP + Fine + Coarse + Age + LogBW + LogWB + LogAge + OneOverCement +   
## OneOverFlyAsh + OneOverBlastF + OneOverSuperP + OneOverCoarse +   
## OneOverFine + OneOverAge + OneOverBlastF2 + OneOverCoarse2 +   
## OneOverFine2 + OneOverAge2 + SquaredCement + SquaredFlyAsh +   
## SquaredSuperP + SquaredCoarse + SquaredFine + SquaredAge +   
## SquaredBW + LogCement + LogFlyAsh + LogWater + LogSuperP +   
## LogFine + WB:Cement + WB:FlyAsh + WB:Water + WB:BlastF +   
## WB:Coarse + BW:Cement + BW:FlyAsh + BW:BlastF + BW:SuperP +   
## BW:Fine + BW:Age + Cement:SuperP + Cement:Fine + Cement:Coarse +   
## Cement:Age + FlyAsh:Water + FlyAsh:SuperP + FlyAsh:Coarse +   
## Water:SuperP + Water:Fine + Water:Coarse + Water:Age + BlastF:SuperP +   
## BlastF:Coarse + BlastF:Age + SuperP:Age + Fine:Coarse + Fine:Age +   
## Coarse:Age)  
##   
## Coefficients:  
## (Intercept) WB BW Cement   
## -1.027e+07 2.382e+04 -8.582e+02 -6.456e+00   
## FlyAsh Water BlastF SuperP   
## -4.050e+00 5.535e+00 -5.572e+00 1.350e+00   
## Fine Coarse Age LogBW   
## -1.340e+03 -3.756e+01 -8.480e-01 4.134e+03   
## LogWB LogAge OneOverCement OneOverFlyAsh   
## -2.775e+04 1.492e+01 1.686e+04 -1.764e+02   
## OneOverBlastF OneOverSuperP OneOverCoarse OneOverFine   
## -1.170e+02 4.113e+01 -3.670e+07 7.935e+08   
## OneOverAge OneOverBlastF2 OneOverCoarse2 OneOverFine2   
## 4.592e+01 1.154e+02 8.743e+09 -7.563e+10   
## OneOverAge2 SquaredCement SquaredFlyAsh SquaredSuperP   
## -7.535e+01 6.062e-04 -3.068e-03 1.311e-01   
## SquaredCoarse SquaredFine SquaredAge SquaredBW   
## 9.960e-03 2.161e-01 2.612e-04 -3.961e+03   
## LogCement LogFlyAsh LogWater LogSuperP   
## 2.411e+02 -5.926e+01 1.138e+03 3.712e+01   
## LogFine WB:Cement WB:FlyAsh WB:Water   
## 1.552e+06 9.945e+00 1.001e+01 -8.791e+00   
## WB:BlastF WB:Coarse BW:Cement BW:FlyAsh   
## 9.902e+00 -3.052e-01 8.225e-01 8.096e-01   
## BW:BlastF BW:SuperP BW:Fine BW:Age   
## 8.058e-01 -2.728e+00 -6.820e-02 1.364e-01   
## Cement:SuperP Cement:Fine Cement:Coarse Cement:Age   
## 1.273e-02 -2.304e-04 -8.305e-04 -4.722e-04   
## FlyAsh:Water FlyAsh:SuperP FlyAsh:Coarse Water:SuperP   
## -9.508e-04 1.098e-02 -1.021e-03 -3.669e-02   
## Water:Fine Water:Coarse Water:Age BlastF:SuperP   
## -5.269e-03 -3.245e-03 1.400e-03 1.445e-02   
## BlastF:Coarse BlastF:Age SuperP:Age Fine:Coarse   
## -8.895e-04 -2.306e-04 -2.554e-03 -6.340e-04   
## Fine:Age Coarse:Age   
## 2.109e-04 1.421e-04

# Reg2 BIC  
step(RegFull, direction="both", trace=0, k=log(1000))

##   
## Call:  
## lm(formula = Strength ~ WB + BW + Cement + FlyAsh + Water + BlastF +   
## SuperP + Fine + Coarse + Age + LogBW + LogAge + OneOverCement +   
## OneOverFlyAsh + OneOverBlastF + OneOverCoarse + OneOverFine +   
## OneOverBlastF2 + OneOverFine2 + SquaredCement + SquaredFlyAsh +   
## SquaredFine + SquaredAge + LogCement + LogFlyAsh + LogWater +   
## LogSuperP + LogFine + WB:Cement + WB:FlyAsh + WB:BlastF +   
## BW:Cement + BW:FlyAsh + BW:BlastF + BW:Age + Cement:Fine +   
## Cement:Age + FlyAsh:Water + Water:Fine + Water:Coarse + BlastF:SuperP)  
##   
## Coefficients:  
## (Intercept) WB BW Cement   
## -6.715e+06 2.703e+02 -3.866e+02 -2.912e+00   
## FlyAsh Water BlastF SuperP   
## -5.841e-02 -1.845e+00 -1.809e+00 -4.837e-01   
## Fine Coarse Age LogBW   
## -8.631e+02 3.913e-01 -1.608e-01 1.783e+03   
## LogAge OneOverCement OneOverFlyAsh OneOverBlastF   
## 1.179e+01 2.122e+04 -2.046e+02 -1.278e+02   
## OneOverCoarse OneOverFine OneOverBlastF2 OneOverFine2   
## -6.923e+04 5.198e+08 1.264e+02 -4.997e+10   
## SquaredCement SquaredFlyAsh SquaredFine SquaredAge   
## 6.898e-04 -3.717e-03 1.380e-01 1.728e-04   
## LogCement LogFlyAsh LogWater LogSuperP   
## 2.944e+02 -6.924e+01 7.822e+02 6.259e+00   
## LogFine WB:Cement WB:FlyAsh WB:BlastF   
## 1.008e+06 5.078e+00 4.852e+00 4.937e+00   
## BW:Cement BW:FlyAsh BW:BlastF BW:Age   
## 3.199e-01 2.633e-01 2.770e-01 6.145e-02   
## Cement:Fine Cement:Age FlyAsh:Water Water:Fine   
## -2.443e-04 -2.491e-04 -9.008e-04 -2.935e-03   
## Water:Coarse BlastF:SuperP   
## -2.497e-03 1.912e-03

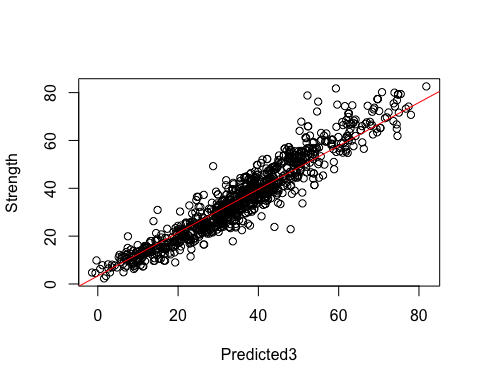
# Use BIC reduce model Reg3 which has fewer terms than AIC  
  
Reg3<-lm(formula = Strength ~ WB + BW + Cement + FlyAsh + Water + BlastF +   
 SuperP + Fine + Coarse + Age + LogBW + LogAge + OneOverCement +   
 OneOverFlyAsh + OneOverBlastF + OneOverCoarse + OneOverFine +   
 OneOverBlastF2 + OneOverFine2 + SquaredCement + SquaredFlyAsh +   
 SquaredFine + SquaredAge + LogCement + LogFlyAsh + LogWater +   
 LogSuperP + LogFine + WB:Cement + WB:FlyAsh + WB:BlastF +   
 BW:Cement + BW:FlyAsh + BW:BlastF + BW:Age + Cement:Fine +   
 Cement:Age + FlyAsh:Water + Water:Fine + Water:Coarse + BlastF:SuperP)  
  
summary(Reg3)

##   
## Call:  
## lm(formula = Strength ~ WB + BW + Cement + FlyAsh + Water + BlastF +   
## SuperP + Fine + Coarse + Age + LogBW + LogAge + OneOverCement +   
## OneOverFlyAsh + OneOverBlastF + OneOverCoarse + OneOverFine +   
## OneOverBlastF2 + OneOverFine2 + SquaredCement + SquaredFlyAsh +   
## SquaredFine + SquaredAge + LogCement + LogFlyAsh + LogWater +   
## LogSuperP + LogFine + WB:Cement + WB:FlyAsh + WB:BlastF +   
## BW:Cement + BW:FlyAsh + BW:BlastF + BW:Age + Cement:Fine +   
## Cement:Age + FlyAsh:Water + Water:Fine + Water:Coarse + BlastF:SuperP)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -25.1368 -3.1221 -0.0441 3.2470 26.5815   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -6.715e+06 1.062e+06 -6.322 3.91e-10 \*\*\*  
## WB 2.703e+02 4.669e+01 5.789 9.50e-09 \*\*\*  
## BW -3.866e+02 6.487e+01 -5.960 3.51e-09 \*\*\*  
## Cement -2.912e+00 4.163e-01 -6.994 4.92e-12 \*\*\*  
## FlyAsh -5.841e-02 3.425e-01 -0.171 0.86462   
## Water -1.845e+00 6.898e-01 -2.675 0.00761 \*\*   
## BlastF -1.809e+00 2.059e-01 -8.787 < 2e-16 \*\*\*  
## SuperP -4.837e-01 2.559e-01 -1.890 0.05901 .   
## Fine -8.631e+02 1.393e+02 -6.198 8.40e-10 \*\*\*  
## Coarse 3.913e-01 5.216e-02 7.500 1.41e-13 \*\*\*  
## Age -1.608e-01 2.191e-02 -7.341 4.42e-13 \*\*\*  
## LogBW 1.783e+03 2.443e+02 7.300 5.91e-13 \*\*\*  
## LogAge 1.179e+01 4.155e-01 28.387 < 2e-16 \*\*\*  
## OneOverCement 2.122e+04 7.377e+03 2.876 0.00412 \*\*   
## OneOverFlyAsh -2.046e+02 2.538e+01 -8.062 2.15e-15 \*\*\*  
## OneOverBlastF -1.278e+02 2.921e+01 -4.373 1.36e-05 \*\*\*  
## OneOverCoarse -6.923e+04 2.580e+04 -2.683 0.00741 \*\*   
## OneOverFine 5.198e+08 8.118e+07 6.403 2.35e-10 \*\*\*  
## OneOverBlastF2 1.264e+02 2.835e+01 4.458 9.22e-06 \*\*\*  
## OneOverFine2 -4.997e+10 7.676e+09 -6.509 1.20e-10 \*\*\*  
## SquaredCement 6.898e-04 2.193e-04 3.145 0.00171 \*\*   
## SquaredFlyAsh -3.717e-03 5.102e-04 -7.285 6.57e-13 \*\*\*  
## SquaredFine 1.380e-01 2.261e-02 6.106 1.47e-09 \*\*\*  
## SquaredAge 1.728e-04 3.829e-05 4.513 7.18e-06 \*\*\*  
## LogCement 2.944e+02 9.133e+01 3.223 0.00131 \*\*   
## LogFlyAsh -6.924e+01 8.710e+00 -7.949 5.10e-15 \*\*\*  
## LogWater 7.822e+02 7.398e+01 10.573 < 2e-16 \*\*\*  
## LogSuperP 6.259e+00 7.899e-01 7.924 6.19e-15 \*\*\*  
## LogFine 1.008e+06 1.600e+05 6.301 4.46e-10 \*\*\*  
## WB:Cement 5.078e+00 7.521e-01 6.752 2.49e-11 \*\*\*  
## WB:FlyAsh 4.852e+00 8.326e-01 5.827 7.63e-09 \*\*\*  
## WB:BlastF 4.937e+00 7.480e-01 6.600 6.68e-11 \*\*\*  
## BW:Cement 3.199e-01 4.476e-02 7.148 1.71e-12 \*\*\*  
## BW:FlyAsh 2.633e-01 5.081e-02 5.182 2.66e-07 \*\*\*  
## BW:BlastF 2.770e-01 4.137e-02 6.695 3.61e-11 \*\*\*  
## BW:Age 6.145e-02 7.455e-03 8.243 5.30e-16 \*\*\*  
## Cement:Fine -2.443e-04 3.296e-05 -7.412 2.66e-13 \*\*\*  
## Cement:Age -2.491e-04 3.429e-05 -7.264 7.63e-13 \*\*\*  
## FlyAsh:Water -9.008e-04 2.444e-04 -3.686 0.00024 \*\*\*  
## Water:Fine -2.935e-03 2.334e-04 -12.579 < 2e-16 \*\*\*  
## Water:Coarse -2.497e-03 2.468e-04 -10.117 < 2e-16 \*\*\*  
## BlastF:SuperP 1.912e-03 5.841e-04 3.273 0.00110 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.214 on 988 degrees of freedom  
## Multiple R-squared: 0.9065, Adjusted R-squared: 0.9026   
## F-statistic: 233.5 on 41 and 988 DF, p-value: < 2.2e-16

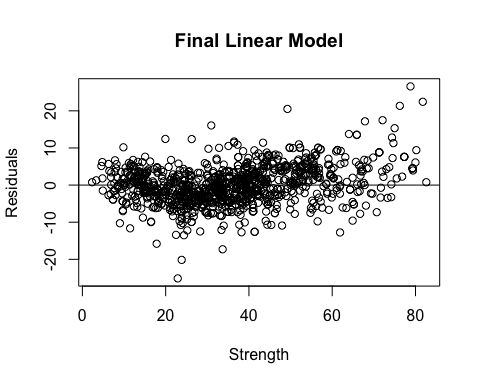
# Reg3  
Predicted3<-predict.lm(Reg3)  
  
# model actual vs predicted  
line<-lm(Predicted3~Strength)  
coef(line)

## (Intercept) Strength   
## 3.3502507 0.9064642

# graph predicted versus actuals  
plot(Predicted3, Strength)  
abline(3.355, 0.9063, col=2)



Residuals3<-residuals(Reg3)  
  
plot(xy$Strength, Residuals3,ylab = "Residuals", xlab = "Strength",main = "Final Linear Model")   
abline(0, 0) # the horizon



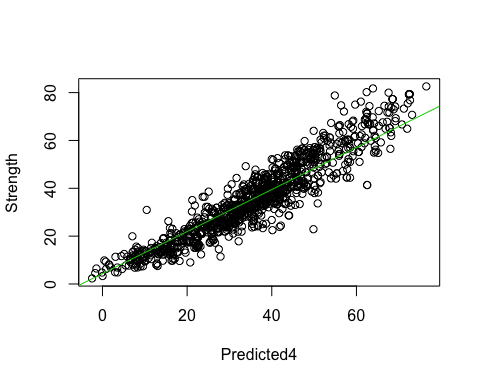
anova(Reg3)

## Analysis of Variance Table  
##   
## Response: Strength  
## Df Sum Sq Mean Sq F value Pr(>F)   
## WB 1 109854 109854 4040.6352 < 2.2e-16 \*\*\*  
## BW 1 7673 7673 282.2258 < 2.2e-16 \*\*\*  
## Cement 1 11216 11216 412.5376 < 2.2e-16 \*\*\*  
## FlyAsh 1 1700 1700 62.5262 7.004e-15 \*\*\*  
## Water 1 247 247 9.0747 0.002658 \*\*   
## BlastF 1 796 796 29.2624 7.931e-08 \*\*\*  
## SuperP 1 117 117 4.3205 0.037912 \*   
## Fine 1 13 13 0.4861 0.485822   
## Coarse 1 252 252 9.2548 0.002411 \*\*   
## Age 1 46255 46255 1701.3599 < 2.2e-16 \*\*\*  
## LogBW 1 1137 1137 41.8288 1.564e-10 \*\*\*  
## LogAge 1 59657 59657 2194.2935 < 2.2e-16 \*\*\*  
## OneOverCement 1 857 857 31.5086 2.579e-08 \*\*\*  
## OneOverFlyAsh 1 837 837 30.8016 3.671e-08 \*\*\*  
## OneOverBlastF 1 174 174 6.3824 0.011681 \*   
## OneOverCoarse 1 22 22 0.8196 0.365523   
## OneOverFine 1 939 939 34.5256 5.746e-09 \*\*\*  
## OneOverBlastF2 1 689 689 25.3259 5.750e-07 \*\*\*  
## OneOverFine2 1 22 22 0.8162 0.366525   
## SquaredCement 1 0 0 0.0115 0.914677   
## SquaredFlyAsh 1 0 0 0.0005 0.982423   
## SquaredFine 1 271 271 9.9661 0.001643 \*\*   
## SquaredAge 1 0 0 0.0002 0.990101   
## LogCement 1 892 892 32.8212 1.341e-08 \*\*\*  
## LogFlyAsh 1 1585 1585 58.2922 5.317e-14 \*\*\*  
## LogWater 1 18 18 0.6756 0.411318   
## LogSuperP 1 1820 1820 66.9426 8.548e-16 \*\*\*  
## LogFine 1 24 24 0.8652 0.352516   
## WB:Cement 1 199 199 7.3238 0.006922 \*\*   
## WB:FlyAsh 1 467 467 17.1646 3.721e-05 \*\*\*  
## WB:BlastF 1 762 762 28.0328 1.470e-07 \*\*\*  
## BW:Cement 1 1 1 0.0388 0.843973   
## BW:FlyAsh 1 648 648 23.8437 1.218e-06 \*\*\*  
## BW:BlastF 1 3465 3465 127.4489 < 2.2e-16 \*\*\*  
## BW:Age 1 660 660 24.2720 9.803e-07 \*\*\*  
## Cement:Fine 1 1027 1027 37.7574 1.160e-09 \*\*\*  
## Cement:Age 1 1038 1038 38.1795 9.420e-10 \*\*\*  
## FlyAsh:Water 1 17 17 0.6431 0.422777   
## Water:Fine 1 1758 1758 64.6751 2.514e-15 \*\*\*  
## Water:Coarse 1 2913 2913 107.1472 < 2.2e-16 \*\*\*  
## BlastF:SuperP 1 291 291 10.7109 0.001102 \*\*   
## Residuals 988 26861 27   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

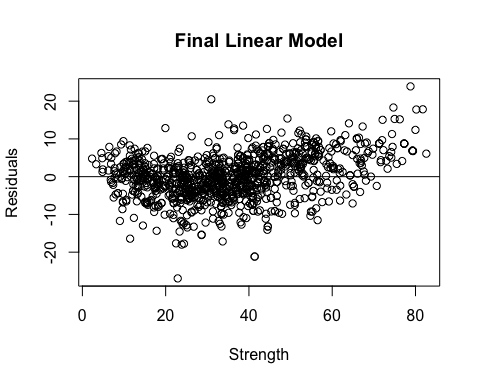
# Reg4  
Predicted4<-predict.lm(Reg4)  
  
# model actual vs predicted  
line<-lm(Predicted4~Strength)  
coef(line)

## (Intercept) Strength   
## 4.3358466 0.8789473

# predicted versus actuals  
plot(Predicted4, Strength)  
abline(4.337, 0.8789, col=3)



residuals4<-residuals(Reg4)  
  
plot(xy$Strength, residuals4,ylab = "Residuals", xlab = "Strength",main = "Final Linear Model")   
abline(0, 0) # the horizon



anova(Reg4)

## Analysis of Variance Table  
##   
## Response: Strength  
## Df Sum Sq Mean Sq F value Pr(>F)   
## WB 1 109854 109854 3163.2256 < 2.2e-16 \*\*\*  
## FSW 1 450 450 12.9550 0.0003347 \*\*\*  
## FFW 1 12313 12313 354.5642 < 2.2e-16 \*\*\*  
## FineCement 1 5552 5552 159.8673 < 2.2e-16 \*\*\*  
## FineB 1 0 0 0.0027 0.9583288   
## FineW 1 237 237 6.8218 0.0091399 \*\*   
## Fine 1 1301 1301 37.4592 1.338e-09 \*\*\*  
## LogWB 1 2626 2626 75.6033 < 2.2e-16 \*\*\*  
## LogAge 1 100812 100812 2902.8622 < 2.2e-16 \*\*\*  
## OneOverFlyAsh 1 1642 1642 47.2816 1.081e-11 \*\*\*  
## OneOverWater 1 100 100 2.8735 0.0903581 .   
## OneOverBlastF 1 41 41 1.1939 0.2748007   
## OneOverFine 1 1669 1669 48.0630 7.389e-12 \*\*\*  
## OneOverAge 1 1096 1096 31.5720 2.490e-08 \*\*\*  
## SquaredFlyAsh 1 653 653 18.8099 1.590e-05 \*\*\*  
## SquaredWater 1 255 255 7.3456 0.0068378 \*\*   
## SquaredBlastF 1 1760 1760 50.6740 2.078e-12 \*\*\*  
## SquaredSuperP 1 809 809 23.3065 1.597e-06 \*\*\*  
## SquaredCoarse 1 1513 1513 43.5780 6.597e-11 \*\*\*  
## SquaredFine 1 101 101 2.8960 0.0891108 .   
## SquaredAge 1 1418 1418 40.8369 2.532e-10 \*\*\*  
## SquaredBW 1 1809 1809 52.0763 1.054e-12 \*\*\*  
## LogCement 1 584 584 16.8042 4.480e-05 \*\*\*  
## LogFlyAsh 1 1124 1124 32.3669 1.675e-08 \*\*\*  
## LogWater 1 1931 1931 55.6164 1.906e-13 \*\*\*  
## LogSuperP 1 2051 2051 59.0588 3.641e-14 \*\*\*  
## LogCoarse 1 370 370 10.6454 0.0011408 \*\*   
## LogFine 1 339 339 9.7638 0.0018311 \*\*   
## Residuals 1001 34763 35   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Cross valid of Reg4 hand model  
  
Reg4xy<-data.frame(Strength, WB, FSW, FFW, FineCement, FineB,   
 FineW, Fine, LogWB, LogAge, OneOverFlyAsh, OneOverWater,   
 OneOverBlastF, OneOverFine, OneOverAge, SquaredFlyAsh,  
 SquaredWater, SquaredBlastF, SquaredSuperP, SquaredCoarse,   
 SquaredFine, SquaredAge, SquaredBW, LogCement, LogFlyAsh,   
 LogWater, LogSuperP, LogCoarse, LogFine)  
  
  
library(tidyverse)  
library(caret)

## Loading required package: lattice

##   
## Attaching package: 'caret'

## The following object is masked from 'package:purrr':  
##   
## lift

# Define training control  
set.seed(123)   
train.control <- trainControl(method = "cv", number = 10)  
# Train the model  
model <- train(Strength ~., data = Reg4xy, method = "lm",  
 trControl = train.control)  
# Summarize the results  
print(model)

## Linear Regression   
##   
## 1030 samples  
## 28 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (10 fold)   
## Summary of sample sizes: 926, 926, 927, 928, 926, 927, ...   
## Resampling results:  
##   
## RMSE Rsquared MAE   
## 6.035286 0.8694434 4.689512  
##   
## Tuning parameter 'intercept' was held constant at a value of TRUE

# Reg3 cross validaiton  
  
WBCement<-WB\*Cement  
WBFlyAsh<-WB\*FlyAsh  
WBBlastF<-WB\*BlastF  
BWCement<-BW\*Cement  
BWFlyAsh<-BW\*FlyAsh  
BWBlastF<-BW\*BlastF  
BWAge<-BW\*Age  
CementFine<-Cement\*Fine  
CementAge<-Cement\*Age  
FlyAshWater<-FlyAsh\*Water  
WaterFine<-Water\*Fine  
WaterCoarse<-Water\*Coarse  
BlastFSuperP<-BlastF\*SuperP  
  
  
  
Reg3xy<-data.frame(Strength, WB, BW, Cement, FlyAsh, Water, BlastF,   
 SuperP, Fine, Coarse, Age, LogBW, LogAge,OneOverCement,   
 OneOverFlyAsh, OneOverBlastF, OneOverCoarse, OneOverFine,   
 OneOverBlastF2,OneOverFine2, SquaredCement, SquaredFlyAsh,  
 SquaredFine, SquaredAge, LogCement, LogFlyAsh, LogWater,   
 LogSuperP, LogFine, WBCement, WBFlyAsh, WBBlastF,   
 BWCement, BWFlyAsh, BWBlastF, BWAge, CementFine,   
 CementAge, FlyAshWater, WaterFine, WaterCoarse, BlastFSuperP)

###  
  
  
library(tidyverse)  
library(caret)  
  
# Define training control  
set.seed(123)   
train.control <- trainControl(method = "cv", number = 10)  
# Train the model  
model <- train(Strength ~., data = Reg3xy, method = "lm",  
 trControl = train.control)  
# Summarize the results  
print(model)

## Linear Regression   
##   
## 1030 samples  
## 41 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (10 fold)   
## Summary of sample sizes: 926, 926, 927, 928, 926, 927, ...   
## Resampling results:  
##   
## RMSE Rsquared MAE   
## 5.375276 0.8969393 4.10565  
##   
## Tuning parameter 'intercept' was held constant at a value of TRUE

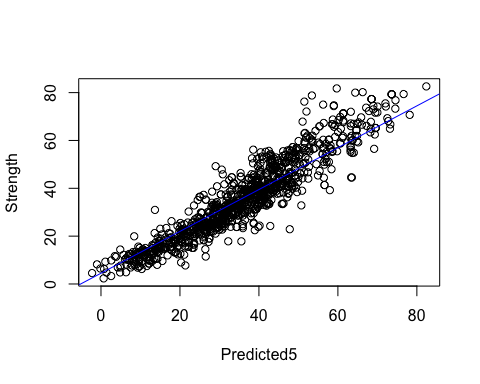
# Reg5 eliminates insignificant features from ANOVA table for Reg3  
  
Reg5<-lm(formula = Strength ~ WB + BW + Cement + FlyAsh + Water + BlastF +   
 SuperP + Coarse + Age + LogBW + LogAge + OneOverCement +   
 OneOverFlyAsh + OneOverBlastF + OneOverFine +   
 OneOverBlastF2 + SquaredFine + LogCement + LogFlyAsh +   
 LogSuperP + WB:Cement + WB:FlyAsh + WB:BlastF + BW:FlyAsh + BW:BlastF + BW:Age + Cement:Fine +   
 Cement:Age + Water:Fine + Water:Coarse + BlastF:SuperP)  
  
summary(Reg5)

##   
## Call:  
## lm(formula = Strength ~ WB + BW + Cement + FlyAsh + Water + BlastF +   
## SuperP + Coarse + Age + LogBW + LogAge + OneOverCement +   
## OneOverFlyAsh + OneOverBlastF + OneOverFine + OneOverBlastF2 +   
## SquaredFine + LogCement + LogFlyAsh + LogSuperP + WB:Cement +   
## WB:FlyAsh + WB:BlastF + BW:FlyAsh + BW:BlastF + BW:Age +   
## Cement:Fine + Cement:Age + Water:Fine + Water:Coarse + BlastF:SuperP)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -24.9153 -3.6587 -0.4049 3.5458 25.3866   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.882e+02 1.425e+02 -1.321 0.186744   
## WB 8.906e+01 3.684e+01 2.417 0.015820 \*   
## BW -2.647e+01 1.761e+01 -1.503 0.133145   
## Cement 1.498e-01 4.895e-02 3.060 0.002275 \*\*   
## FlyAsh 3.982e-01 1.521e-01 2.617 0.009007 \*\*   
## Water -1.190e+00 7.838e-01 -1.518 0.129409   
## BlastF -1.793e-01 9.828e-02 -1.825 0.068342 .   
## SuperP -1.826e-01 2.847e-01 -0.641 0.521401   
## Coarse 1.290e-01 3.115e-02 4.141 3.75e-05 \*\*\*  
## Age -4.591e-02 1.353e-02 -3.394 0.000716 \*\*\*  
## LogBW 1.836e+02 8.554e+01 2.146 0.032125 \*   
## LogAge 1.050e+01 3.212e-01 32.693 < 2e-16 \*\*\*  
## OneOverCement -2.978e+03 2.981e+03 -0.999 0.317989   
## OneOverFlyAsh -3.102e+01 1.001e+01 -3.099 0.001995 \*\*   
## OneOverBlastF -2.216e+01 3.261e+01 -0.680 0.496884   
## OneOverFine -9.773e+04 1.457e+04 -6.706 3.34e-11 \*\*\*  
## OneOverBlastF2 2.443e+01 3.163e+01 0.772 0.440037   
## SquaredFine 1.416e-05 1.071e-05 1.322 0.186436   
## LogCement 2.743e+00 2.550e+01 0.108 0.914338   
## LogFlyAsh -7.875e+00 2.736e+00 -2.878 0.004092 \*\*   
## LogSuperP 5.287e+00 8.856e-01 5.970 3.29e-09 \*\*\*  
## WB:Cement 2.283e+00 7.858e-01 2.905 0.003750 \*\*   
## WB:FlyAsh 1.963e+00 8.865e-01 2.214 0.027043 \*   
## WB:BlastF 2.661e+00 7.982e-01 3.334 0.000888 \*\*\*  
## BW:FlyAsh -7.459e-02 3.110e-02 -2.399 0.016635 \*   
## BW:BlastF 2.680e-02 2.042e-02 1.313 0.189643   
## BW:Age 3.356e-02 8.115e-03 4.135 3.85e-05 \*\*\*  
## Cement:Fine -1.536e-04 3.430e-05 -4.478 8.40e-06 \*\*\*  
## Cement:Age -1.842e-04 3.886e-05 -4.740 2.45e-06 \*\*\*  
## Water:Fine -6.613e-04 1.474e-04 -4.486 8.10e-06 \*\*\*  
## Water:Coarse -6.105e-04 1.636e-04 -3.731 0.000201 \*\*\*  
## BlastF:SuperP 2.452e-03 6.231e-04 3.936 8.88e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.01 on 998 degrees of freedom  
## Multiple R-squared: 0.8745, Adjusted R-squared: 0.8706   
## F-statistic: 224.3 on 31 and 998 DF, p-value: < 2.2e-16

# Reg5  
Predicted5<-predict.lm(Reg5)  
  
# model actual vs predicted  
line<-lm(Predicted5~Strength)  
coef(line)

## (Intercept) Strength   
## 4.4954939 0.8744901

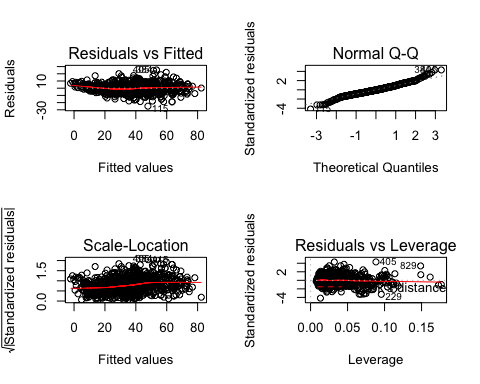
# predicted versus actuals  
plot(Predicted5, Strength)  
abline(4.4954939, 0.8744901 , col=4)



anova(Reg5)

## Analysis of Variance Table  
##   
## Response: Strength  
## Df Sum Sq Mean Sq F value Pr(>F)   
## WB 1 109854 109854 3041.7472 < 2.2e-16 \*\*\*  
## BW 1 7673 7673 212.4566 < 2.2e-16 \*\*\*  
## Cement 1 11216 11216 310.5539 < 2.2e-16 \*\*\*  
## FlyAsh 1 1700 1700 47.0691 1.201e-11 \*\*\*  
## Water 1 247 247 6.8314 0.0090919 \*\*   
## BlastF 1 796 796 22.0284 3.061e-06 \*\*\*  
## SuperP 1 117 117 3.2525 0.0716178 .   
## Coarse 1 142 142 3.9362 0.0475313 \*   
## Age 1 45900 45900 1270.9387 < 2.2e-16 \*\*\*  
## LogBW 1 813 813 22.5079 2.398e-06 \*\*\*  
## LogAge 1 58693 58693 1625.1691 < 2.2e-16 \*\*\*  
## OneOverCement 1 1052 1052 29.1169 8.512e-08 \*\*\*  
## OneOverFlyAsh 1 1358 1358 37.6030 1.248e-09 \*\*\*  
## OneOverBlastF 1 135 135 3.7485 0.0531385 .   
## OneOverFine 1 1743 1743 48.2711 6.689e-12 \*\*\*  
## OneOverBlastF2 1 695 695 19.2504 1.268e-05 \*\*\*  
## SquaredFine 1 236 236 6.5462 0.0106575 \*   
## LogCement 1 64 64 1.7801 0.1824470   
## LogFlyAsh 1 91 91 2.5127 0.1132497   
## LogSuperP 1 1634 1634 45.2515 2.914e-11 \*\*\*  
## WB:Cement 1 442 442 12.2392 0.0004886 \*\*\*  
## WB:FlyAsh 1 638 638 17.6728 2.859e-05 \*\*\*  
## WB:BlastF 1 798 798 22.0964 2.957e-06 \*\*\*  
## BW:FlyAsh 1 674 674 18.6492 1.728e-05 \*\*\*  
## BW:BlastF 1 1241 1241 34.3501 6.250e-09 \*\*\*  
## BW:Age 1 192 192 5.3207 0.0212780 \*   
## Cement:Fine 1 603 603 16.7087 4.708e-05 \*\*\*  
## Cement:Age 1 889 889 24.6038 8.273e-07 \*\*\*  
## Water:Fine 1 468 468 12.9510 0.0003355 \*\*\*  
## Water:Coarse 1 466 466 12.9109 0.0003426 \*\*\*  
## BlastF:SuperP 1 559 559 15.4883 8.877e-05 \*\*\*  
## Residuals 998 36043 36   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#residuals versus leverage  
# https://medium.com/data-distilled/residual-plots-part-4-residuals-vs-leverage-plot-14aeed009ef7  
par(mfrow=c(2,2))  
plot(Reg5)



Reg5xy<-data.frame(Strength, WB, BW, Cement, FlyAsh, Water, BlastF,   
 SuperP, Coarse, Age, LogBW, LogAge,OneOverCement,   
 OneOverFlyAsh, OneOverBlastF, OneOverFine,   
 OneOverBlastF2, SquaredFine, LogCement, LogFlyAsh,   
 LogSuperP, WBCement, WBFlyAsh, WBBlastF, BWFlyAsh, BWBlastF, BWAge, CementFine,   
 CementAge, WaterFine, WaterCoarse, BlastFSuperP)  
  
#10-fold Xval of Reg 5  
  
library(tidyverse)  
library(caret)  
  
# Define training control  
set.seed(123)   
train.control <- trainControl(method = "cv", number = 10)  
# Train the model  
model <- train(Strength ~., data = Reg5xy, method = "lm",  
 trControl = train.control)  
  
  
# Summarize the results  
print(model)

## Linear Regression   
##   
## 1030 samples  
## 31 predictor  
##   
## No pre-processing  
## Resampling: Cross-Validated (10 fold)   
## Summary of sample sizes: 926, 926, 927, 928, 926, 927, ...   
## Resampling results:  
##   
## RMSE Rsquared MAE   
## 6.190597 0.8635658 4.706972  
##   
## Tuning parameter 'intercept' was held constant at a value of TRUE