**Problem 1**

> 127\*(2014-32/1.43)

[1] 252936

> sqrt(5213.4/14.3)

[1] 19.09381

> log(19\*\*5.5)

[1] 16.19441

> exp(2.47/1.05)

[1] 10.51057

**Problem 2**

> X

[1] 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120

[25] 125 130 135 140 145 150 155 160 165 170 175 180 185 190

> Y <- seq(86,49,-1)

> Y

[1] 86 85 84 83 82 81 80 79 78 77 76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59 58 57 56 55

[33] 54 53 52 51 50 49

> Z <- seq(53.0,108.5,1.5)

> Z

[1] 53.0 54.5 56.0 57.5 59.0 60.5 62.0 63.5 65.0 66.5 68.0 69.5 71.0 72.5 74.0 75.5

[17] 77.0 78.5 80.0 81.5 83.0 84.5 86.0 87.5 89.0 90.5 92.0 93.5 95.0 96.5 98.0 99.5

[33] 101.0 102.5 104.0 105.5 107.0 108.5

> M <- X\*Y-Z

> M

[1] 377.0 795.5 1204.0 1602.5 1991.0 2369.5 2738.0 3096.5 3445.0 3783.5 4112.0 4430.5 4739.0 5037.5

[15] 5326.0 5604.5 5873.0 6131.5 6380.0 6618.5 6847.0 7065.5 7274.0 7472.5 7661.0 7839.5 8008.0 8166.5

[29] 8315.0 8453.5 8582.0 8700.5 8809.0 8907.5 8996.0 9074.5 9143.0 9201.5

> M[15]

[1] 5326

> M[20]

[1] 6618.5

> M[32]

[1] 8700.5

> intersect(M[M<4900], M[M>4000])

[1] 4112.0 4430.5 4739.0

> max(M)

[1] 9201.5

> min(M)

[1] 377

> max(M)-min(M)

[1] 8824.5

**Problem 3**

> sum(M)

[1] 224171.5

> mean(M)

[1] 5899.25

> median(M)

[1] 6499.25

> sd(M)

[1] 2705.801

> quantile(M,c(.4,.8))

40% 80%

5548.8 8530.6

**Problem 4**

> A <- matrix(c(1,6,7,10,2,5,8,11,3,4,9,12),3,4,TRUE)

> A

[,1] [,2] [,3] [,4]

[1,] 1 6 7 10

[2,] 2 5 8 11

[3,] 3 4 9 12

> B <-matrix(c(5,8,4,9,7,1,6,2),4,2)

> B

[,1] [,2]

[1,] 5 7

[2,] 8 1

[3,] 4 6

[4,] 9 2

> Z <-matrix(c(6,-3,7,4),2,2)

> A\*B\*Z

Error in A \* B : non-conformable arrays

Must be a syntax error – perfectly fine to multiply a 3x4 by a 4x2

**Problem 5**

> summary(`Assignment+2\_R\_Data\_CherryTrees`$Girth)

Min. 1st Qu. Median Mean 3rd Qu. Max.

8.30 11.05 12.90 13.25 15.25 20.60

> summary(`Assignment+2\_R\_Data\_CherryTrees`$Height)

Min. 1st Qu. Median Mean 3rd Qu. Max.

63 72 76 76 80 87

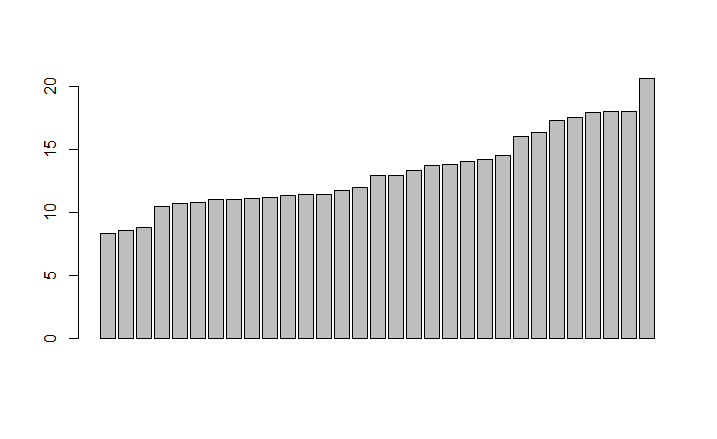
> summary(`Assignment+2\_R\_Data\_CherryTrees`$Volume)

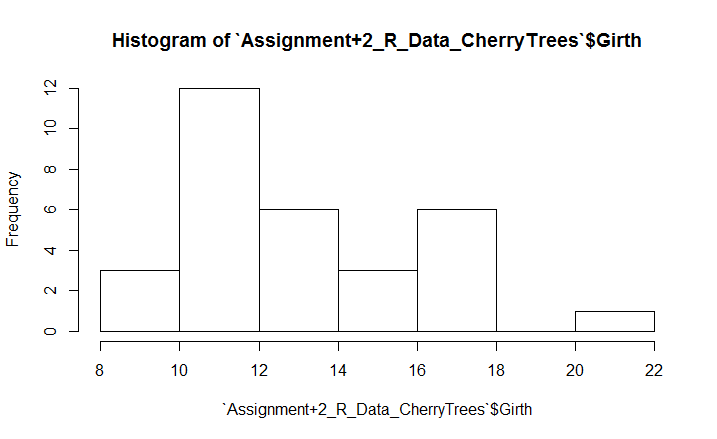
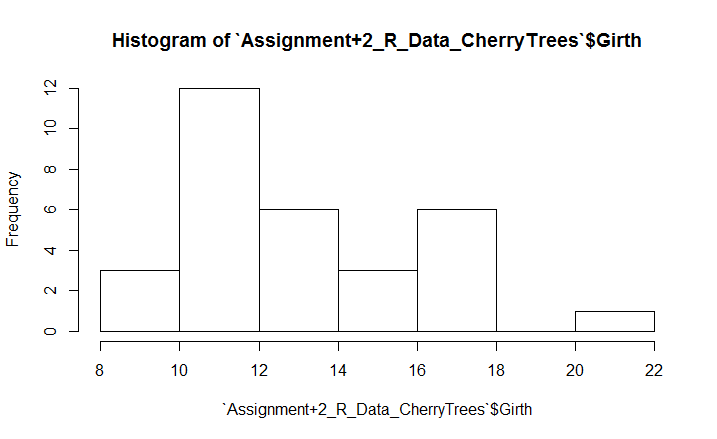
Min. 1st Qu. Median Mean 3rd Qu. Max.

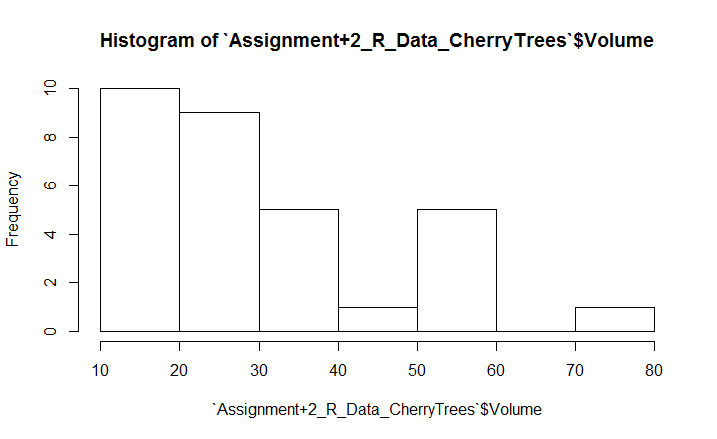
10.20 19.40 24.20 30.17 37.30 77.00

**Problem 6**

Tree girth bar plot

****

****

****

**Problem 7**

> CherryTrees$Level <- L3

> CherryTrees$Level[CherryTrees$Height < 80] <- "L2"

> CherryTrees$Level[CherryTrees$Height < 70] <- "L1"

> as.data.frame(table(CherryTrees$Level))

Var1 Freq

1 L1 5

2 L2 14

3 L3 12

> ggplot(CherryTrees, aes(x=CherryTrees$Girth, y=CherryTrees$Height)) +

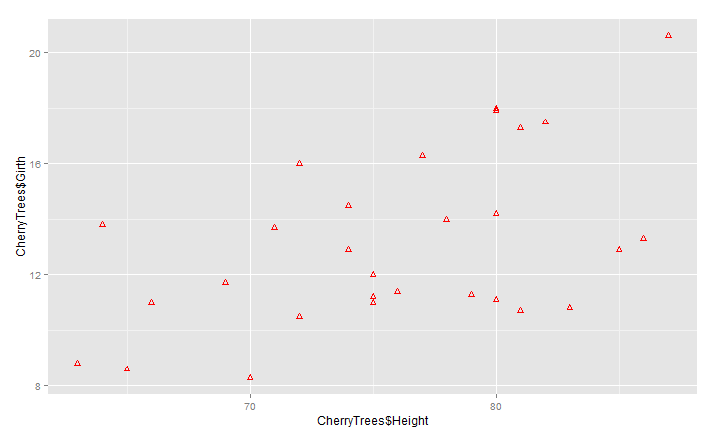
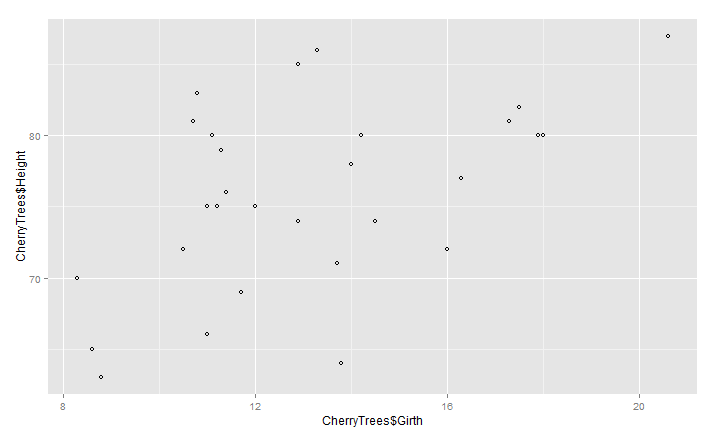
+ geom\_point(shape=1, color=1)

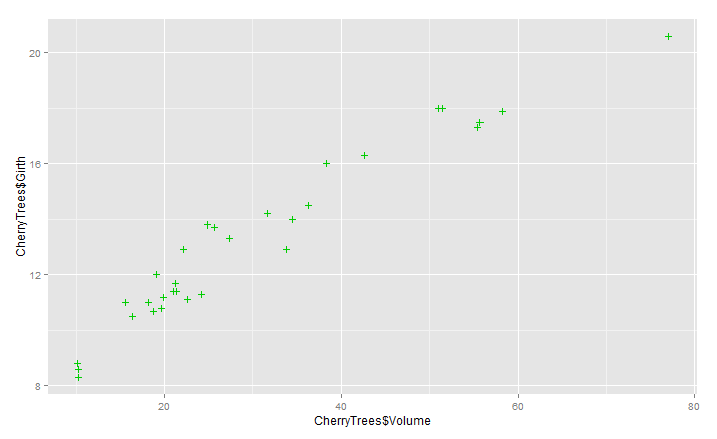
> ggplot(CherryTrees, aes(x=CherryTrees$Height, y=CherryTrees$Girth)) +

+ geom\_point(shape=2, color=2)

> ggplot(CherryTrees, aes(x=CherryTrees$Volume, y=CherryTrees$Girth)) +

+ geom\_point(shape=3, color=3)

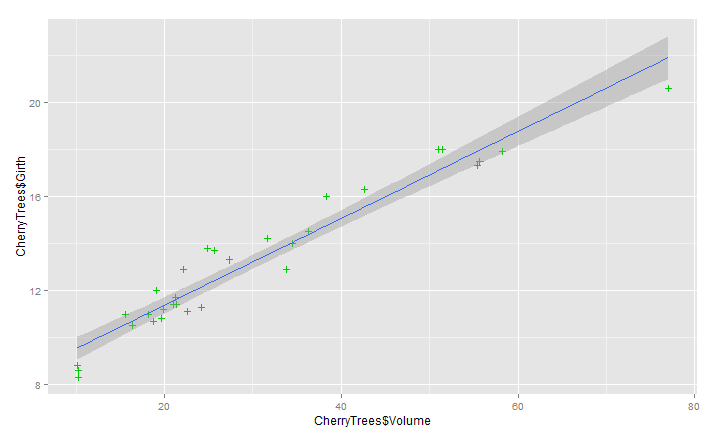




> ggplot(CherryTrees, aes(x=CherryTrees$Volume, y=CherryTrees$Girth)) +

+ geom\_point(shape=3, color=3) +

+ geom\_smooth(method=lm, se =TRUE)



Looking quickly at the graph, this looks like a pretty good fit but human intuition is notoriously falliable here. Before making any inference, I would like to dive deeper into why we’re making this linear estimation, any domain specific logic for a linear relationship, and evaluate goodness of fit through tool likes R\*\*2, residual plot (making sure to see if there’s any pattern), etc.