3-4

a) 10 pt
The stem and leaf plots will vary based on the width used for the stem. As long as valid plots are made that is fine with me. I used R to create these (and the students should do them 'by hand') so my plots have a lot more than necessary.

```
Ohm: 20
_____________________________________________
  1 | 2: represents 1.2, leaf unit: 0.1
        Quarter       Half
_____________________________________________
   3         110| 19* |
  (8)   33332222| 19  |
   4         544| 19  |
   1           6| 19  |77          2
                | 19  |9           3
                | 20  |011         6
                | 20  |223        (3)
                | 20  |445         6
                | 20  |6           3
                | 20. |8           2
                | 21* |
_____________________________________________
                      HI: 24.4
n:             15       15
_____________________________________________
```

```
Ohm: 75
________________________________________
  1 | 2: represents 1.2, leaf unit: 0.1
      Quarter       Half
                    LO: 68.6
________________________________________
              | 70* |
   1         9| 70. |
              | 71* |
   2         8| 71. |7          2
   7     43310| 72* |1          3
  (6)   997655| 72. |8          4
   2        42| 73* |2          5
              | 73. |89         7
              | 74* |02        (2)
              | 74. |68         6
              | 75* |0          4
              | 75. |
              | 76* |2          3
```

```
             | 76. |57       2
             | 77* |
_________________________________
n:          15       15
_________________________________

Ohm: 100
_______________________________
  1 | 2: represents 1.2, leaf unit: 0.1
     Quarter       Half
_______________________________
   5   98861|  94 |
  (3)    841|  95 |5        1
   7      00|  96 |68       3
   5     743|  97 |229      6
   2      53|  98 |57      (2)
            |  99 |2        7
            | 100 |00       6
            | 101 |
            | 102 |000      4
            | 103 |0        1
_______________________________
n:         15       15
_______________________________

Ohm: 150
____________________________________________
  1 | 2: represents 1.2, leaf unit: 0.1
            Quarter        Half
                           LO: 145 146
____________________________________________
    3           000| 147* |0            3
                   | 147. |
  (11)  00000000000| 148* |0            4
                   | 148. |
    1             0| 149* |0000         8
                   | 149. |
                   | 150* |00           7
____________________________________________
                           HI: 151 152 153
                           154 155
n:                15       15
____________________________________________
Ohm: 200_________________________________
  1 | 2: represents 12, leaf unit: 1
      Quarter       Half
_________________________________
             | 19  |
   1        3| 19  |2        1
   5     5554| 19  |5        2
  (6)  666666| 19  |677      5
   4     9998| 19  |9        6
```

```
                   | 20  |1         7
                   | 20  |2        (1)
                   | 20  |5         7
                   | 20  |77        6
                   | 20  |
                   | 21  |01        4
                   | 21  |
                   | 21  |4         2
                   | 21  |
                   | 21  |
                   | 22  |
___________________________________
                         HI: 257
n:          15           15
___________________________________
```

These plots show us that the quarter Watt resistors have a distribution that seems to have a lower average than the ½ Watt resistors. The ¼ watt resistors also seem to have less spread than the ½ watt resistors.

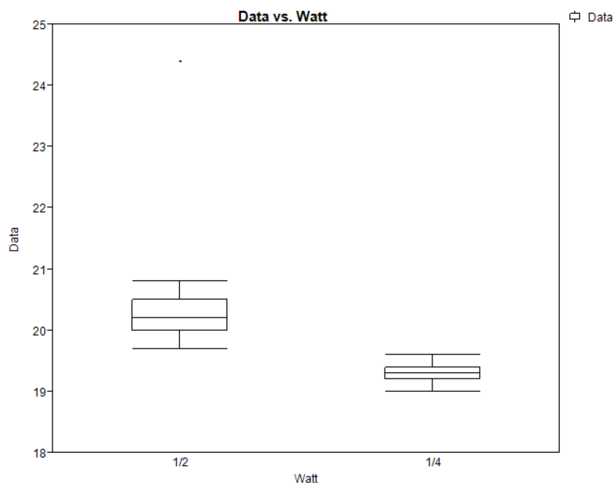
b) 10 pt

Following are the numeric summaries necessary to build the boxplots

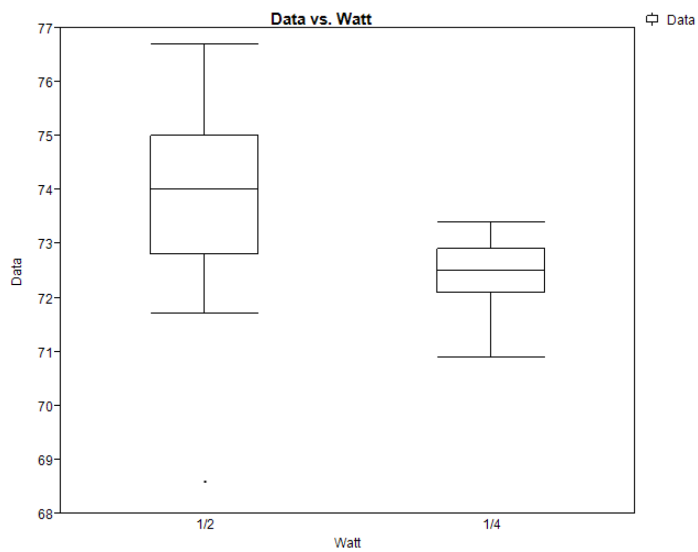
	1/2.20	1/4.20	1/2.75	1/4.75	1/2.100	1/4.100	1/2.150	1/4.150	1/2.200	1/4.200
lower.whisker	19.7	19	71.7	71.8	95.5	94.1	145	148	192	193
q1	20.025	19.2	72.9	72.15	97.2	94.825	148.25	148	197	195
median	20.2	19.3	74	72.5	98.7	95.8	149	148	202	196
q3	20.475	19.375	74.95	72.85	101.5	97.375	151.75	148	209.25	197.5
upper.whisker	20.8	19.6	76.7	73.4	103	98.5	155	148	214	199
outliers	24.4	Numeric,0	68.6	70.9	Numeric,0	Numeric,0	Numeric,0	147x3,149	257	Numeric,0

“Numeric,0” in the outliers row just means there aren’t any outliers

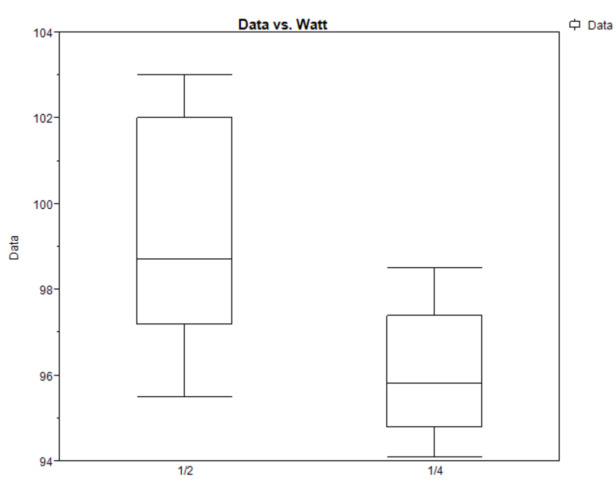
Ohms: 20



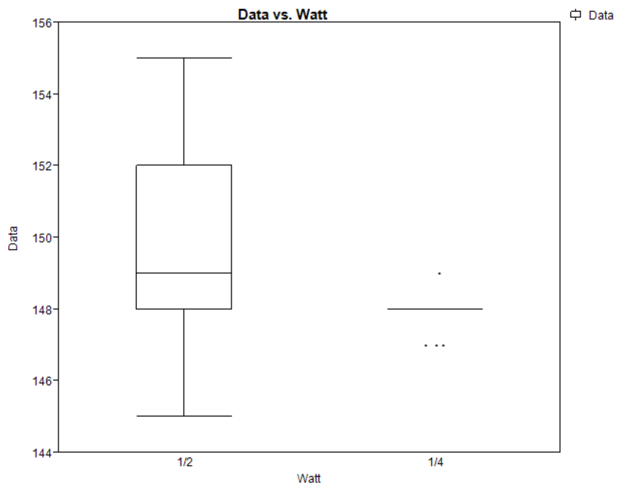
Ohms: 75



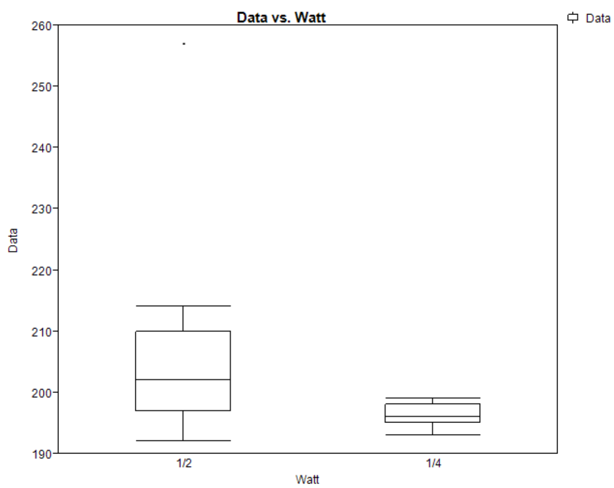
Ohms: 100



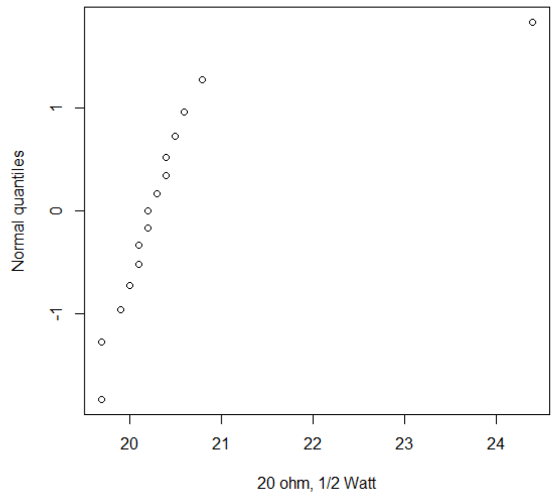
Ohms: 150

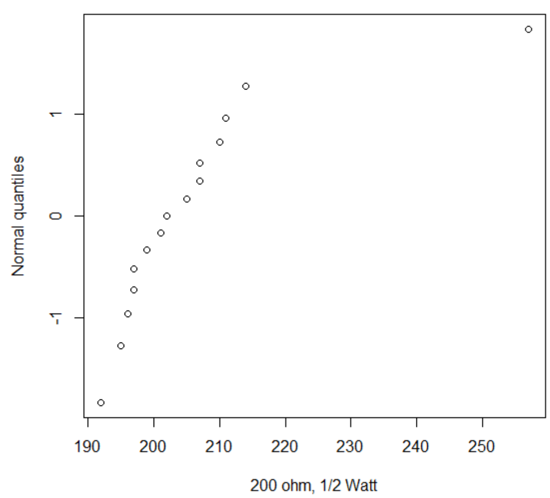


Ohms: 200



c) 10 pts





For the most part the plots look like a straight line could go through them fairly well except for the upper ‘outlier’ point. So it appears like we would underestimate the number of ‘Large’ values we would see in the data.

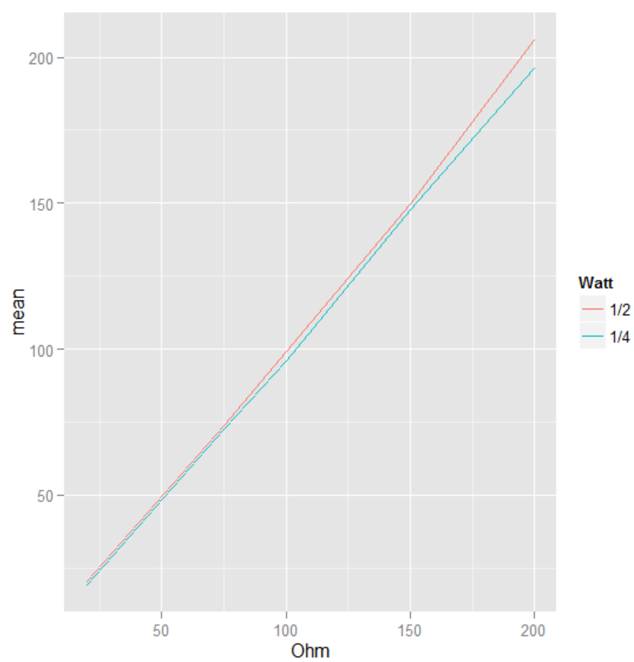
d) 10 pts

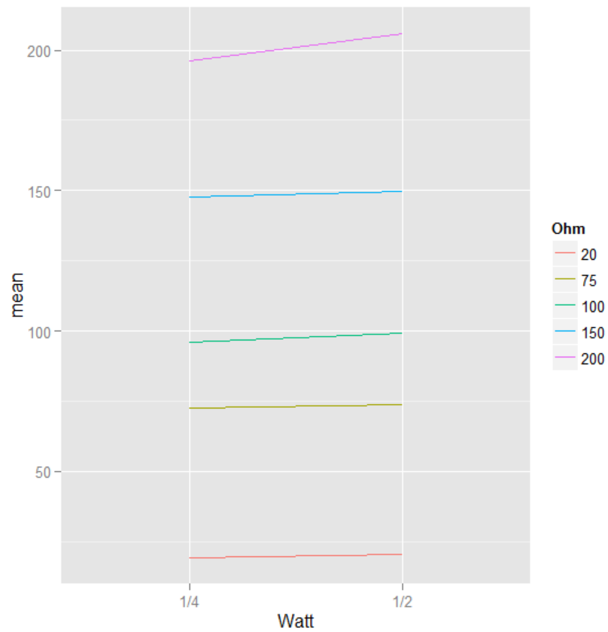
```
    Label       mean     StdDev
1  1/2W100  99.10667  2.3260840
2  1/2W150 149.80000  2.8334734
3   1/2W20  20.48667  1.1262242
4  1/2W200 206.00000 15.5333742
5   1/2W75  73.87333  2.0765585
6  1/4W100  96.04667  1.4371930
7  1/4W150 147.86667  0.5163978
8   1/4W20  19.27333  0.1579632
9  1/4W200 196.20000  1.8205180
10  1/4W75  72.43333  0.6067085
```

These do appear to agree with the statement that the 1/4Watt resistors tend to have smaller averages and small spreads

e) 5 pts

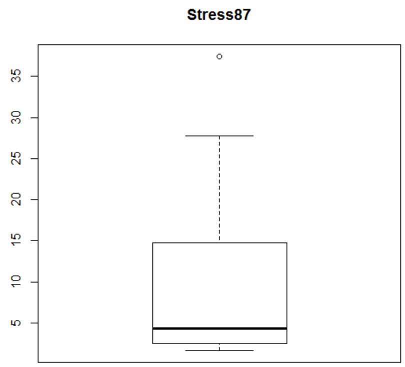
It doesn't specify whether to break make 5 lines (one for each ohm) or 2 lines (one for each Watt) so I'll present both…

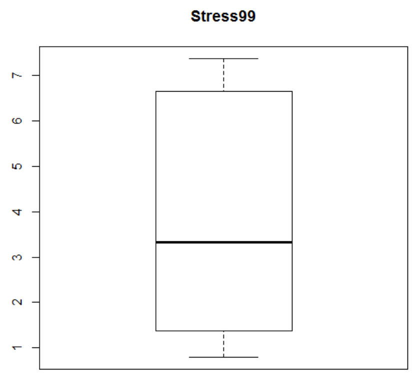


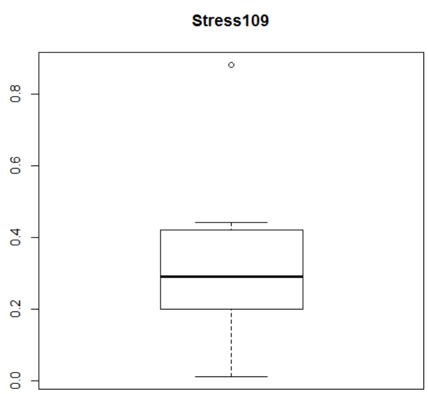


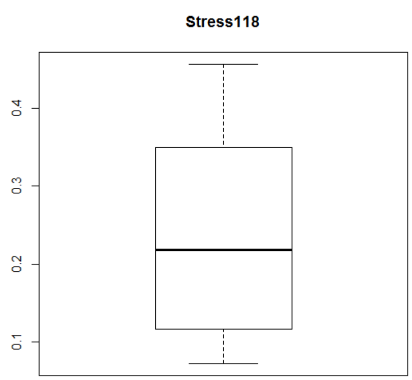
Both plots show the ¼ Watt resistors having a lower mean than the ½ plot resistors but we also see that the difference between the means gets larger as Ohms increases.

3-11)



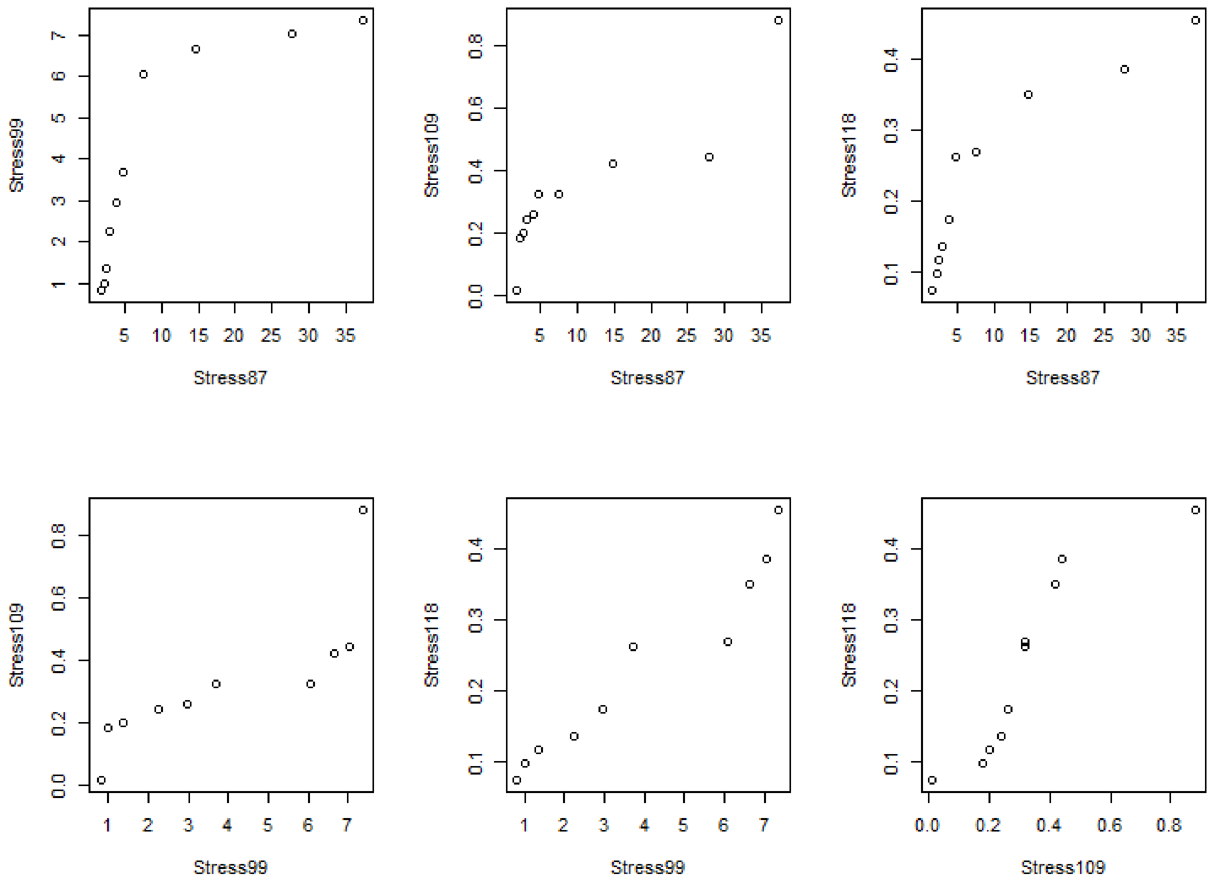






a) 10 pts
The problem said “side-by-side boxplots” which typically means we put them all on the same scale. But since we want to analyze shape it makes more sense to give them each their own axis to compare the actual shape of the boxplots. We only have 10 data values for each plot so we would need to see major differences to try to claim there are big differences. The .87x10^6psi distribution looks more skewed than the others.

b) 10 pts



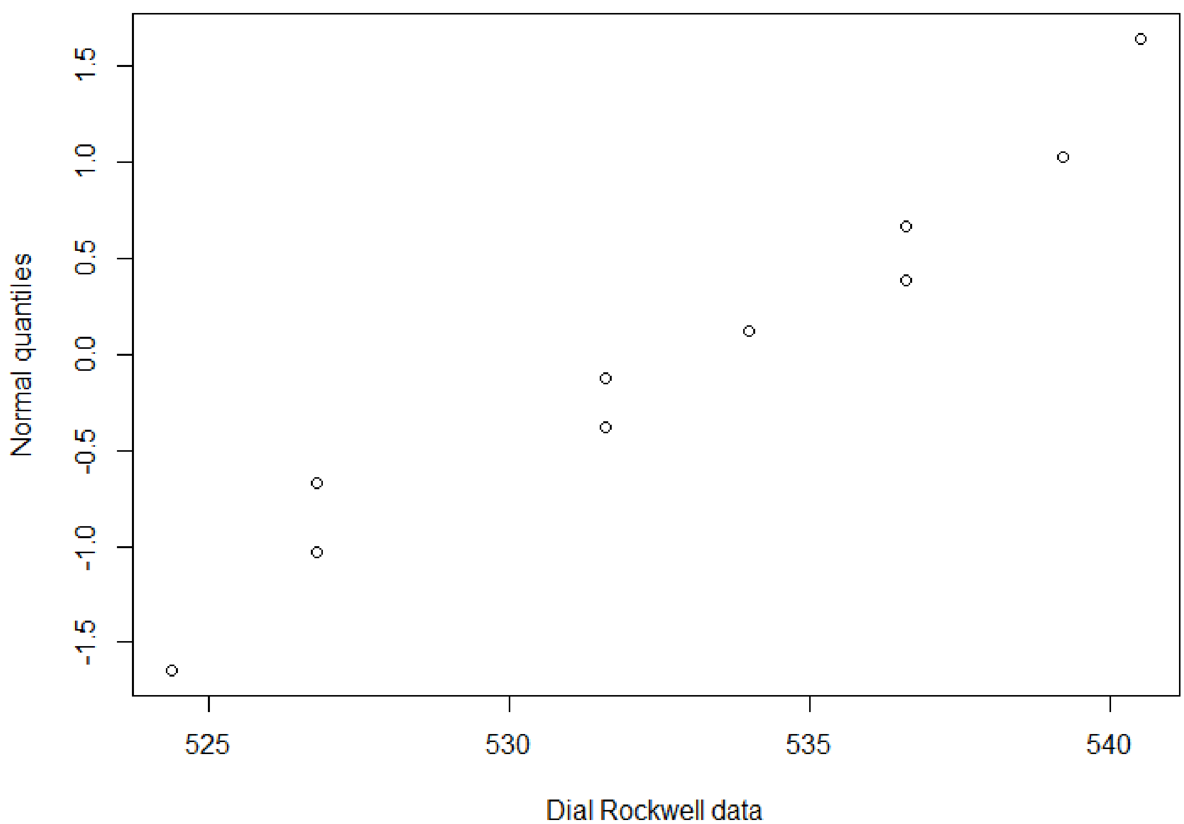
From the plots on the top row we can see that stress87 seems to be more right skewed than stress99 and stress118. There does appear to be a little variation in the shape of the distributions but with so few data points its hard to put too much faith into an assessment like that. Although it wouldn't be perfect there probably wouldn't be too much of an issue assuming that these had similar shapes (but larger difference centers and scales). Answers will vary – as long as there is an honest attempt to analyze the shapes that should suffice.

3-21

a) 5 pts
The median will be the mean of the 5th and 6th smallest sorted values in this case:
Median = 532.8
The first and third quartiles are actually just the 3rd and 8th observations:
Q1 = 526.8
Q3 = 536.6

For Q(.27) we have I' = 10*.27 + .5 = 3.2 so Q(.27) = .8*x[3] + .2*x[4] where x[3] is the 3rd smallest observation:
Q(.27) = .8*526.8 + .2*531.6 = 527.76

b) 5 pts



This looks like the data does fall pretty much on a straight line so there I wouldn't have an issue with somebody describing it as bell-shaped.

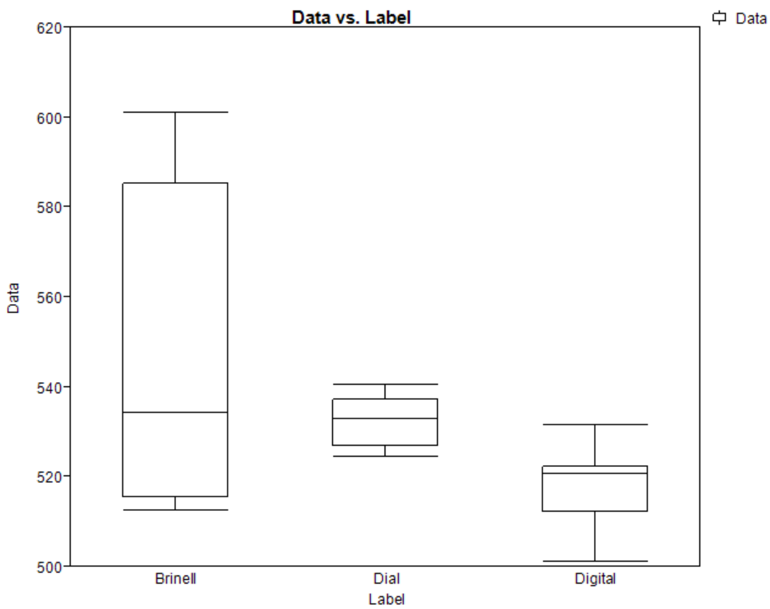
c) 10 pts

Label	N Rows	Mean(Data)	Std Dev(Data)	Range(Data)
Brinell	10	545.97	35.31773	88.4
Dial	10	532.81	5.53784	16.1
Digital	10	518.01	8.776287	30.4

d) 5 pts

e) 5 pts

Using JMP



f) 5 pts

The dial Rockwell produced the most precise results. It has the most compact boxplot. Looking at the summaries from part c suggest that it has the least amount of variation as well.

g) 5 pts

It is not possible to tell which machine produced the most accurate results. To know if the machines are accurate we would need to know what the “true” value was for each measurement. Since we don’t have that information we can’t assess accuracy.

4-1)

x	Y	x - xbar	y - ybar	(x-xbar)*(y-ybar)	(x-xbar)^2	(y - ybar)^2
0.45	2954	-0.05	188.1111	-9.405555556	0.0025	35385.79012
0.45	2913	-0.05	147.1111	-7.355555556	0.0025	21641.67901
0.45	2923	-0.05	157.1111	-7.855555556	0.0025	24683.90123
0.5	2743	0	-22.8889	0	0	523.9012346
0.5	2779	0	13.11111	0	0	171.9012346
0.5	2739	0	-26.8889	0	0	723.0123457
0.55	2652	0.05	-113.889	-5.694444444	0.0025	12970.67901
0.55	2607	0.05	-158.889	-7.944444444	0.0025	25245.67901
0.55	2583	0.05	-182.889	-9.144444444	0.0025	33448.34568
sum 4.5	24893	1.67E-16	1.82E-12	-47.4	0.015	154794.8889

a) 12 pts
xbar = 4.5/9 = 0.5
ybar = 24893/9 = 2765.89
b1 = -47.4/.015 = -3160
b0 = ybar – xbar*b1 = 4345.89

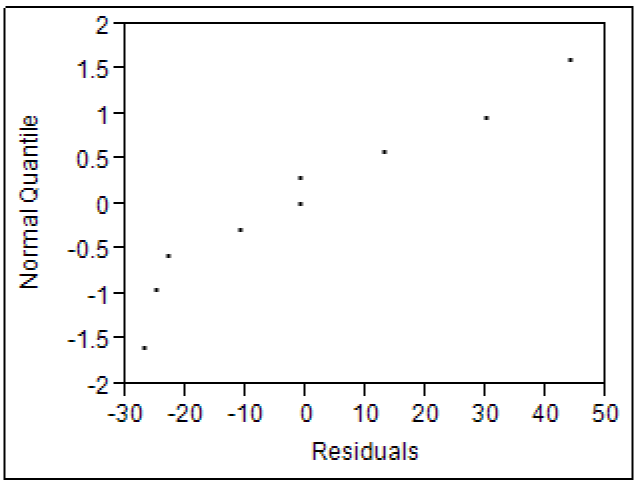
Our fitted regression line is then:
y-hat = 4345.89 – 3160*x

b) 3 pts
r = -47.4/sqrt(.015 * 154794.8889) = -.9837
The correlation of -.98 tells us that we have a strong negative linear relationship between Water/Cement Ratio and 14-Day Compressive Strength.

c) 3 pts
R^2 = r^2 = (-.9837)^2 = 0.967628848
The fitted line accounts for 96.7% of the raw variability in 14-Day Compressive Strength

d) 5 pts

i	p	Residuals	Normal Q
1	0.055555556	-26.8889	-1.59162
2	0.166666667	-24.8889	-0.96361
3	0.277777778	-22.8889	-0.58621
4	0.388888889	-10.8889	-0.28044
5	0.5	-0.88889	0
6	0.611111111	-0.88889	0.280438
7	0.722222222	13.11111	0.586205
8	0.833333333	30.11111	0.963615
9	0.944444444	44.11111	1.591622



The points don't fall perfectly on a straight line but the deviation isn't too bad. It would be reasonable to use a normal distribution to model the residuals.

e) 3 pts
We would just plug in .48 in for x in our regression equation
Predicted value = 4345.89 – 3160*0.48 = 2829.088889

So for a Water/Cement Ratio of .48 we would predict a 14-Day Compressive Strength of 2829.1 psi.

f) 8 pts

Using JMP

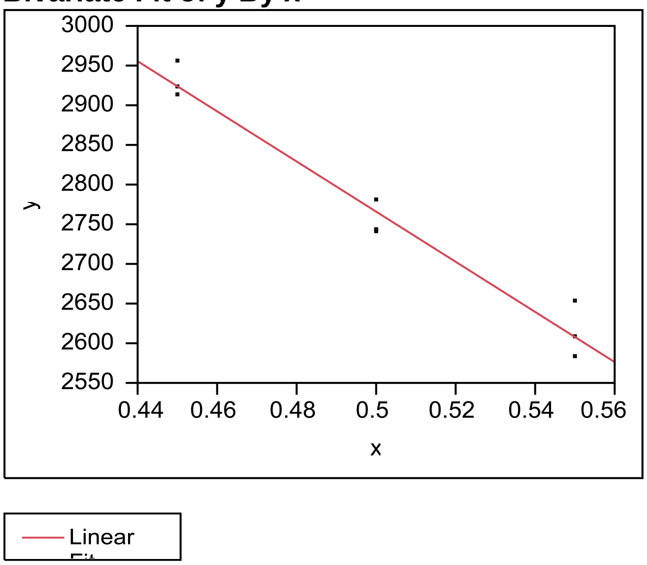
y = 4345.8889 - 3160*x

r = -0.9837 (obtained from Multivariate Correlations)

R^2 = .9676

Residual plots can be seen below

Bivariate Fit of y By x



Linear Fit

y = 4345.8889 - 3160*x

Summary of Fit

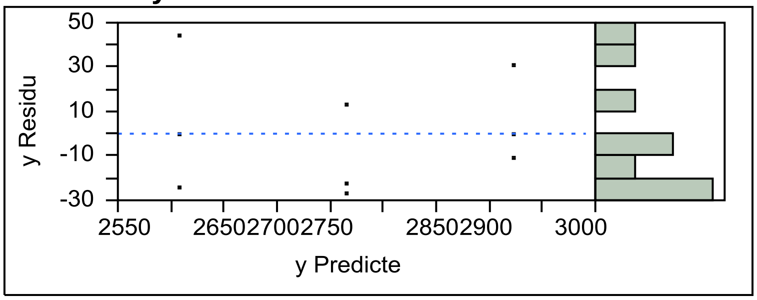
RSquare	0.967629
RSquare Adj	0.963004
Root Mean Square Error	26.75521
Mean of Response	2765.889
Observations (or Sum Wgts)	9

Parameter Estimates

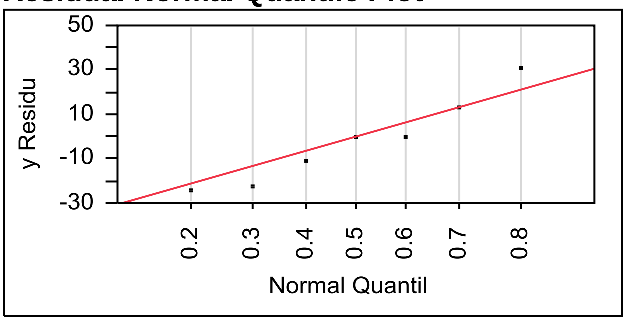
Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	4345.8889	109.5912	39.66	<.0001*
x	-3160	218.4554	-14.47	<.0001*

Diagnostics Plots

Residual by Predicted Plot



Residual Normal Quantile Plot



Multivariate

Correlations

	x	y
x	1.0000	-0.9837
y	-0.9837	1.0000

4-3)

a) 3 pts

They used a complete factorial design. An obvious weakness is that they only have a single replicate for each temperature*time combination.

b) 10 pts

Equation	R^2
Y = -515.15 + .356*x1 + .0107*x2	.886
Y = -528.46 + .356*x1 + 3.71*ln(x2)	.888
Y = -42.35 + .031*x1 -93.71*ln(x2) + .065*x1*ln(x2)	.962

Response y
Whole Model

Summary of Fit

RSquare	0.88616
RSquare Adj	0.848214
Root Mean Square Error	6.966912
Mean of Response	22.55556
Observations (or Sum Wgts)	9

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-515.1503	84.98199	-6.06	0.0009*
x1	0.3566667	0.056885	6.27	0.0008*
x2	0.01069	0.003932	2.72	0.0347*

Response y
Whole Model

Summary of Fit

RSquare	0.888598
RSquare Adj	0.851464
Root Mean Square Error	6.891927
Mean of Response	22.55556
Observations (or Sum Wgts)	9

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-528.4622	84.31093	-6.27	0.0008*
x1	0.3566667	0.056272	6.34	0.0007*
ln(x2)	3.7106558	1.338465	2.77	0.0323*

Response y
Whole Model

Summary of Fit

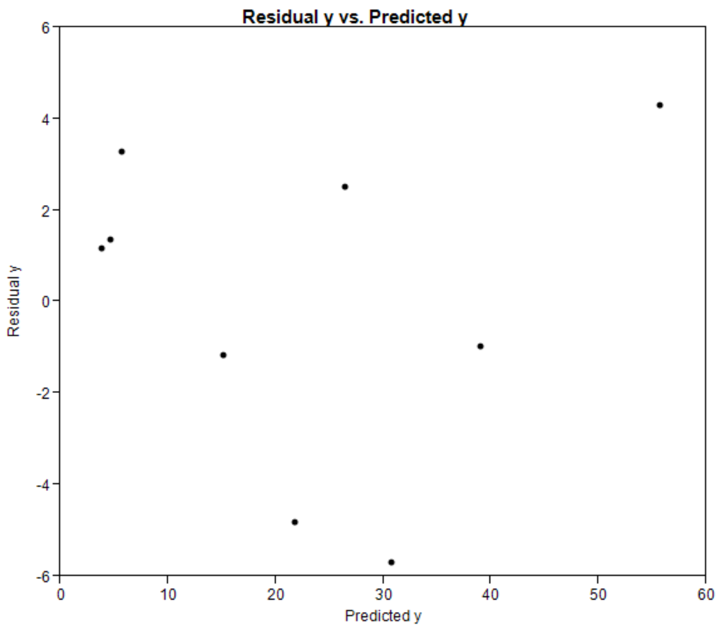
RSquare	0.962152
RSquare Adj	0.939444
Root Mean Square Error	4.400523
Mean of Response	22.55556
Observations (or Sum Wgts)	9

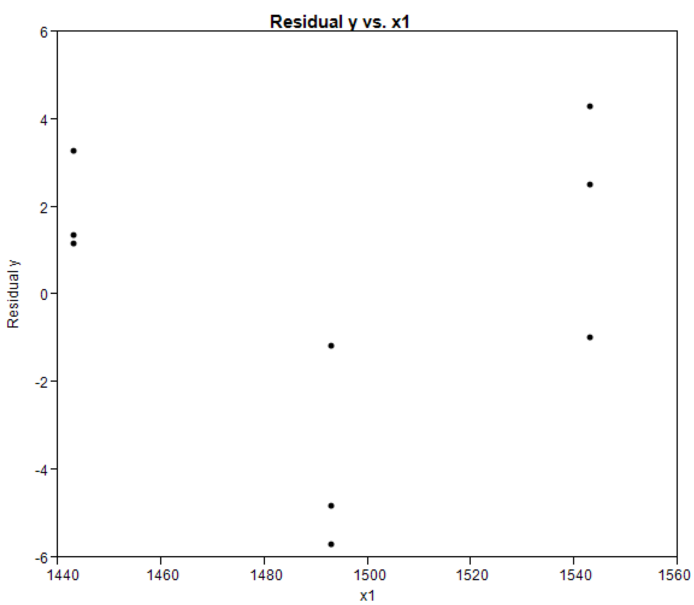
Parameter Estimates

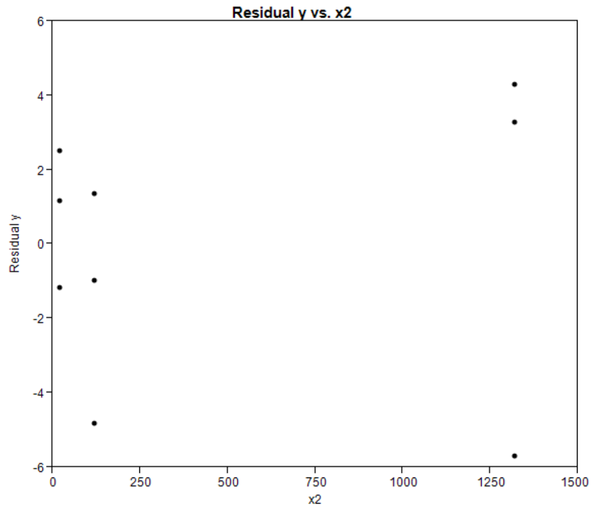
Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-42.35063	164.9735	-0.26	0.8076
x1	0.0310728	0.110457	0.28	0.7897
ln(x2)	-93.71554	31.26572	-3.00	0.0302*
x1*ln(x2)	0.0652553	0.020934	3.12	0.0263*

c) 10 pts

Residuals plots







Looking at the residuals by predicted plots it looks like there is still some trend left that could be modeled. The plot of residual by x1 makes it appear like we could probably do a better job of modeling if we used a transformation of x1 instead of just using x1 linearly.

d) Extra credit: + 10 pts if they did a decent job

e) 5 pts

Y = -42.35 + .031*x1 -93.71*ln(x2) + .065*x1*ln(x2) which gives…

Y-hat = -42.35063 + .0310728*1500 -93.71554*log(500) + .0652553*1500*log(500)

= 30.15739