Lab 4a

Jas Sur

2/27/2022

Question 1

Setup data for pressure, temperature and yield

```
temp.data <- rep(c(150,160,170), each=6)
pressure.data <- rep(c(200,215,230), each=3, times=2)
yield <- c(90.4, 90.2, 90.7,90.6,90.2,90.4,90.1,90.3,90.5,90.6,89.9,90.1,90.5,90.7,90.8,90.9,90.4,90.1)
chem.process <- data.frame(temp.data, pressure.data, yield)
attach(chem.process)
```

```
## The following objects are masked _by_ .GlobalEnv:
##
##     pressure.data, temp.data, yield
```

```
chem.process$temp <-factor(temp.data, levels =c(150,160,170), labels=c("150C", "160C", "170C"))
chem.process$pressure <-factor(pressure.data, levels =c(200,215,230), labels=c("200Psig","215Psig", "230

head(chem.process)
```

```
##   temp.data pressure.data yield temp pressure
## 1       150           200  90.4 150C  200Psig
## 2       150           200  90.2 150C  200Psig
## 3       150           200  90.7 150C  200Psig
## 4       150           215  90.6 150C  215Psig
## 5       150           215  90.2 150C  215Psig
## 6       150           215  90.4 150C  215Psig
```

```
str(chem.process)
```

```
## 'data.frame':    18 obs. of  5 variables:
##  $ temp.data    : num  150 150 150 150 150 150 160 160 160 160 ...
##  $ pressure.data: num  200 200 200 215 215 215 230 230 230 200 ...
##  $ yield        : num  90.4 90.2 90.7 90.6 90.2 90.4 90.1 90.3 90.5 90.6 ...
##  $ temp         : Factor w/ 3 levels "150C","160C",..: 1 1 1 1 1 1 2 2 2 2 ...
##  $ pressure     : Factor w/ 3 levels "200Psig","215Psig",..: 1 1 1 2 2 2 3 3 3 1 ...
```

```
attach(chem.process)
```

```
## The following objects are masked _by_ .GlobalEnv:
##
##     pressure.data, temp.data, yield

## The following objects are masked from chem.process (pos = 3):
##
##     pressure.data, temp.data, yield

## The following object is masked from package:datasets:
##
##     pressure
```

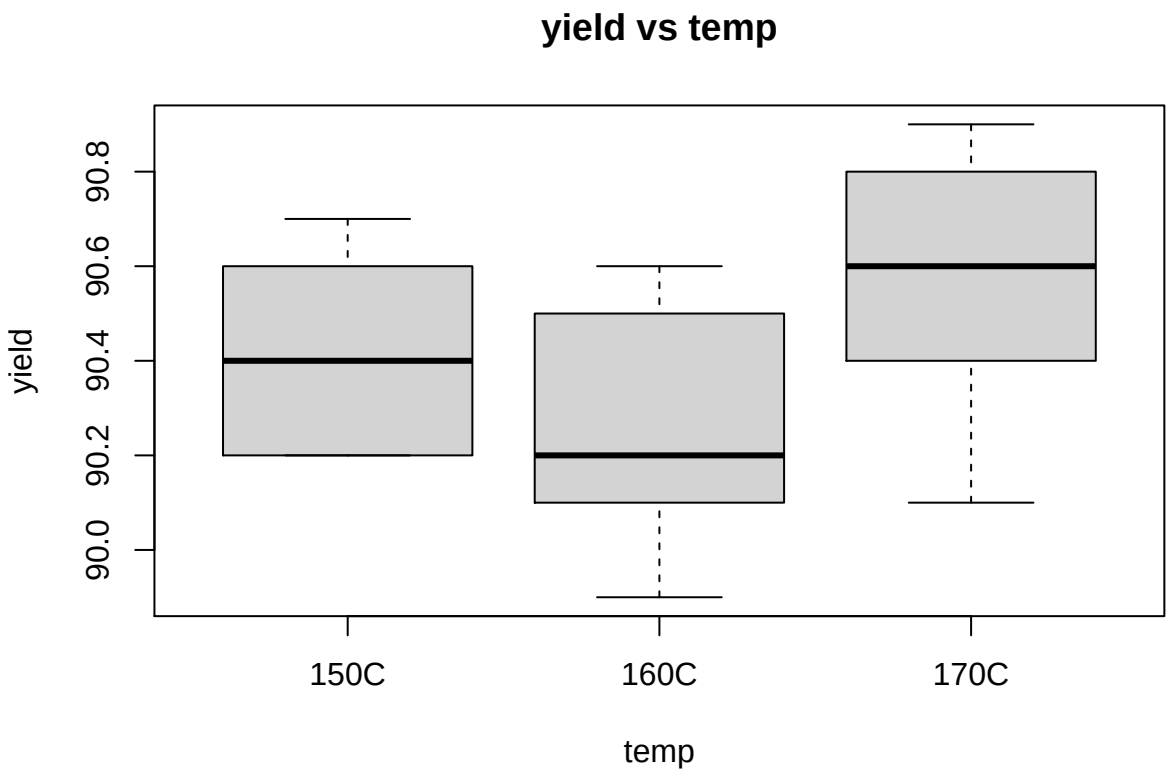
a. Analyze the data and draw conclusions, Use alpha = 0.05

a1. Plotting: boxplot of yield vs temp, yield vs pressure, yeld vs temp and pressure, interaction plot

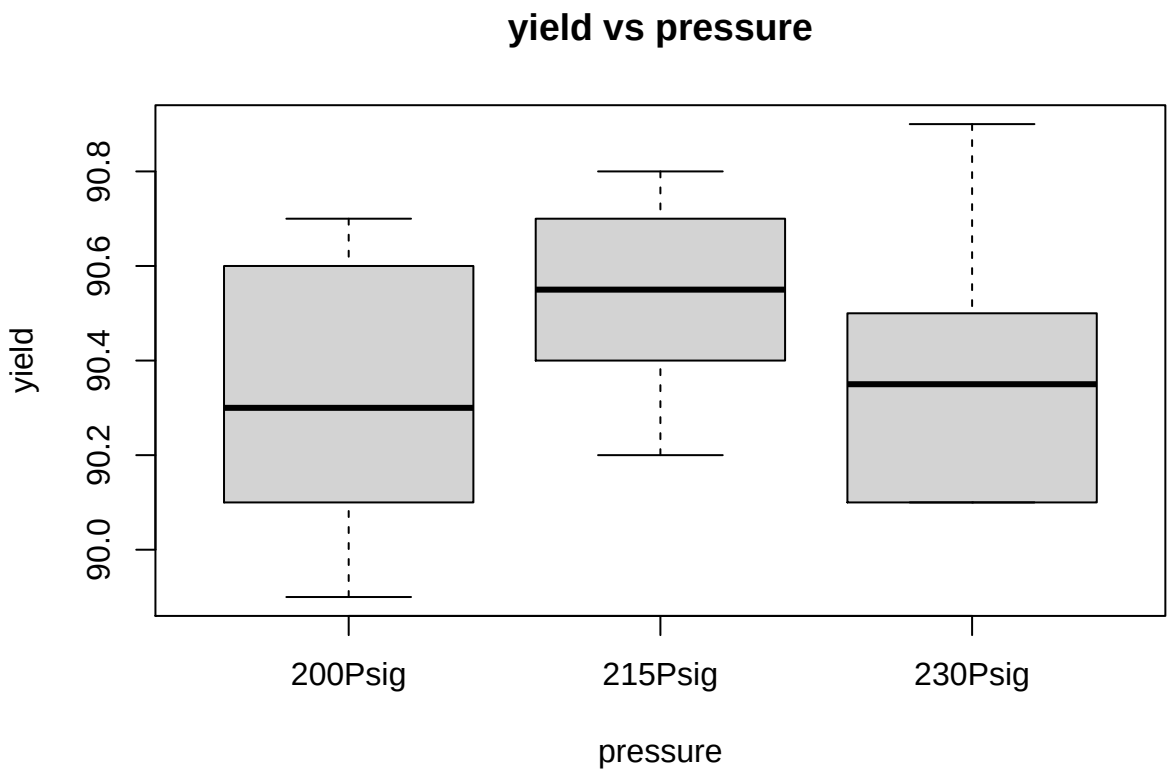
```
tapply(yield, list(temp,pressure), mean)
```

```
##       200Psig  215Psig  230Psig
## 150C 90.43333 90.40000       NA
## 160C 90.20000       NA 90.30000
## 170C       NA 90.66667 90.46667
```

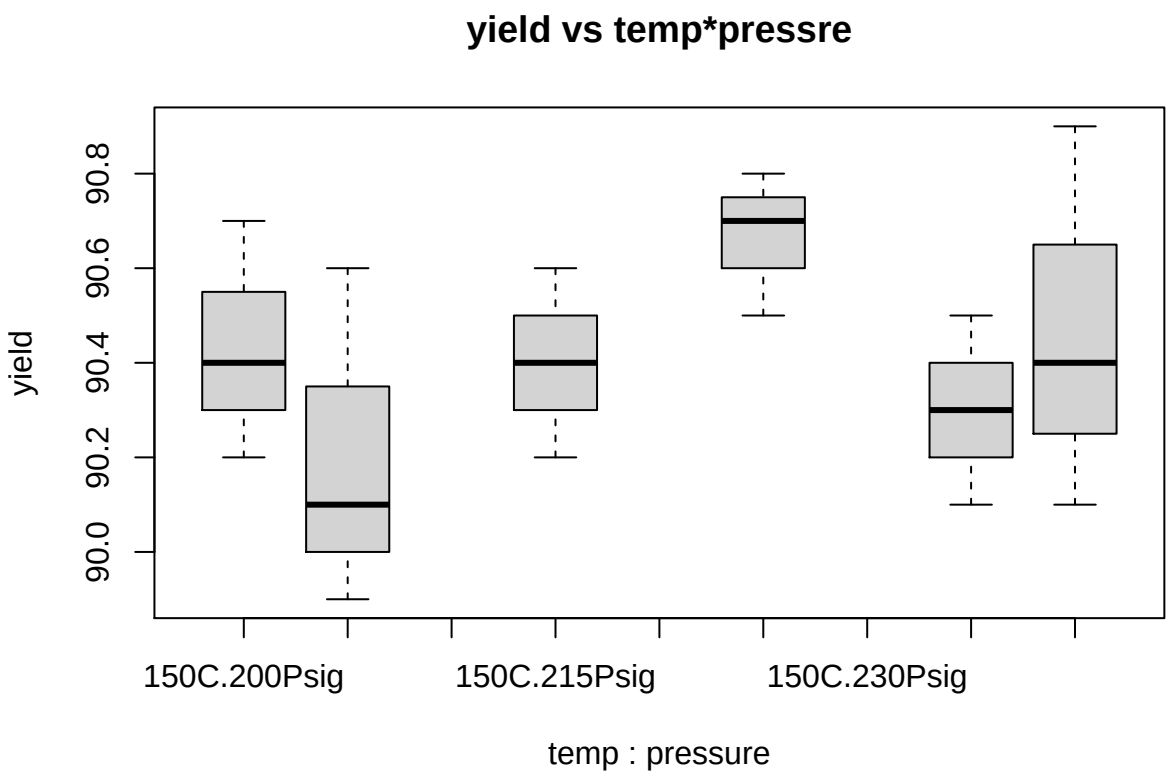
```
boxplot(yield~temp, main="yield vs temp")
```



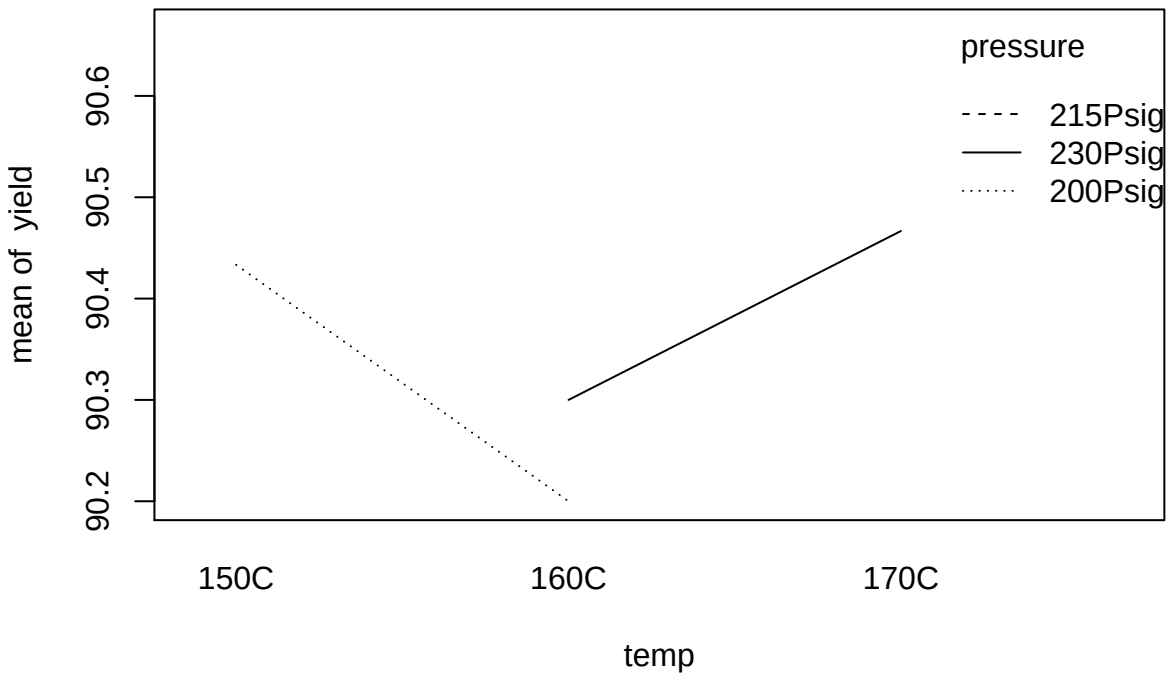
```
boxplot(yield~pressure, main="yield vs pressure")
```



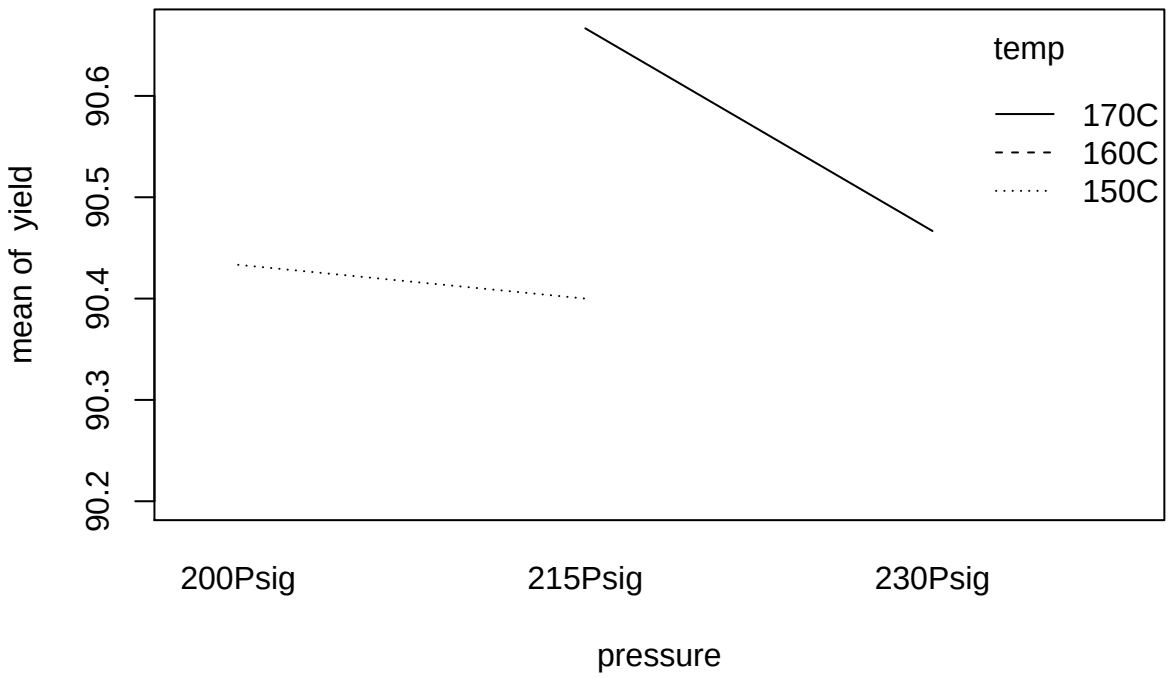

```
boxplot(yield~temp*pressure, main="yield vs temp*pressre")
```




```
interaction.plot(temp, pressure, yield)
```




```
interaction.plot(pressure, temp, yield)
```



Slopes are changing, so there is some kind of interaction

a2. Build a linear model using aov(). Print ANOVA output. Are main factors and interaction factor significant?

```
chem.process.mod <- aov(yield~temp*pressure)
summary.aov(chem.process.mod)
```

```
##               Df Sum Sq Mean Sq F value Pr(>F)
## temp           2 0.3011 0.15056   1.964  0.183
## pressure       2 0.0211 0.01056   0.138  0.873
## temp:pressure  1 0.0556 0.05556   0.725  0.411
## Residuals     12 0.9200 0.07667
```

```
summary.lm(chem.process.mod)
```

```
##
## Call:
## aov(formula = yield ~ temp * pressure)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.36667 -0.19167 -0.01667  0.18333  0.43333
##
## Coefficients: (3 not defined because of singularities)
##                           Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)              90.43333    0.15986 565.700   <2e-16 ***
## temp160C                 -0.23333    0.22608  -1.032    0.322
## temp170C                 -0.06667    0.31972  -0.209    0.838
## pressure215Psig          -0.03333    0.22608  -0.147    0.885
## pressure230Psig           0.10000    0.22608   0.442    0.666
## temp160C:pressure215Psig       NA         NA      NA       NA
## temp170C:pressure215Psig  0.33333    0.39158   0.851    0.411
## temp160C:pressure230Psig       NA         NA      NA       NA
## temp170C:pressure230Psig       NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2769 on 12 degrees of freedom
## Multiple R-squared:  0.2911, Adjusted R-squared:  -0.004281
## F-statistic: 0.9855 on 5 and 12 DF,  p-value: 0.4657
```

```
chem.process.nointeraction.mod <- aov(yield~temp+pressure)
anova(chem.process.mod, chem.process.nointeraction.mod)
```

```
## Analysis of Variance Table
##
## Model 1: yield ~ temp * pressure
## Model 2: yield ~ temp + pressure
##   Res.Df     RSS Df Sum of Sq      F Pr(>F)
## 1     12 0.92000
## 2     13 0.97556 -1 -0.055556 0.7246 0.4113
```

Since P-value >.05, we accept null and reject Alternate. The interaction term is in-significant, so we can discard the interaction term. We take the partial model instead of full model

a3. Print out regression model with coefficients

```
summary.lm(chem.process.nointeraction.mod)
```

```
##
## Call:
## aov(formula = yield ~ temp + pressure)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.42222 -0.15278 -0.01667  0.17778  0.37778
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     90.37778    0.14438 625.977   <2e-16 ***
## temp160C        -0.12222    0.18263  -0.669    0.515
## temp170C         0.15556    0.18263   0.852    0.410
## pressure215Psig  0.07778    0.18263   0.426    0.677
## pressure230Psig -0.01111    0.18263  -0.061    0.952
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2739 on 13 degrees of freedom
```

```
## Multiple R-squared:  0.2483, Adjusted R-squared:  0.01699
## F-statistic: 1.073 on 4 and 13 DF,  p-value: 0.4089
```

Only the intercept is significant, so it means that the linear model can remove other terms while predicting the values.

b. Prepare appropriate residuals plot and comment on the model's adequecy

```
summary.lm(chem.process.nointeraction.mod)
```

```
##
## Call:
## aov(formula = yield ~ temp + pressure)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.42222 -0.15278 -0.01667  0.17778  0.37778
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     90.37778    0.14438 625.977   <2e-16 ***
## temp160C        -0.12222    0.18263  -0.669    0.515
## temp170C         0.15556    0.18263   0.852    0.410
## pressure215Psig  0.07778    0.18263   0.426    0.677
## pressure230Psig -0.01111    0.18263  -0.061    0.952
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2739 on 13 degrees of freedom
## Multiple R-squared:  0.2483, Adjusted R-squared:  0.01699
## F-statistic: 1.073 on 4 and 13 DF,  p-value: 0.4089
```

```
summary.aov(chem.process.nointeraction.mod)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## temp         2 0.3011 0.15056   2.006  0.174
## pressure     2 0.0211 0.01056   0.141  0.870
## Residuals   13 0.9756 0.07504
```

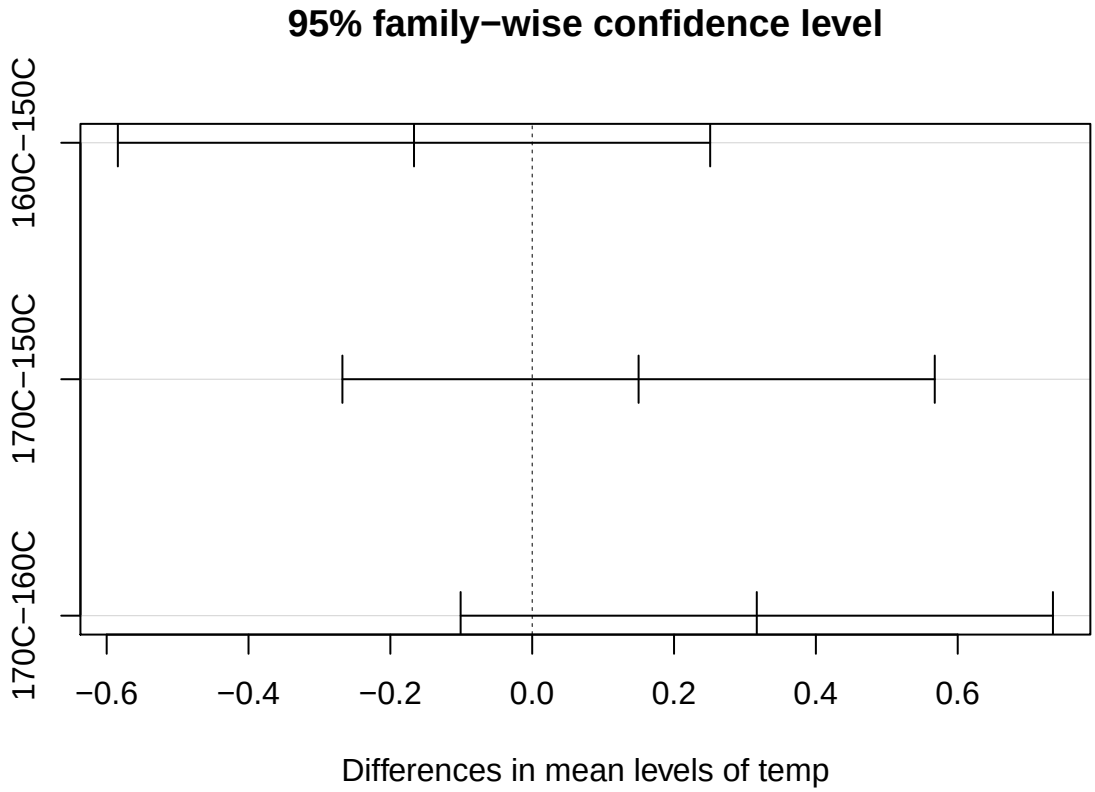
It looks like the difference of mean for both temperature and pressure is the same, so there is not much difference becaue of temperatuer or pressure.

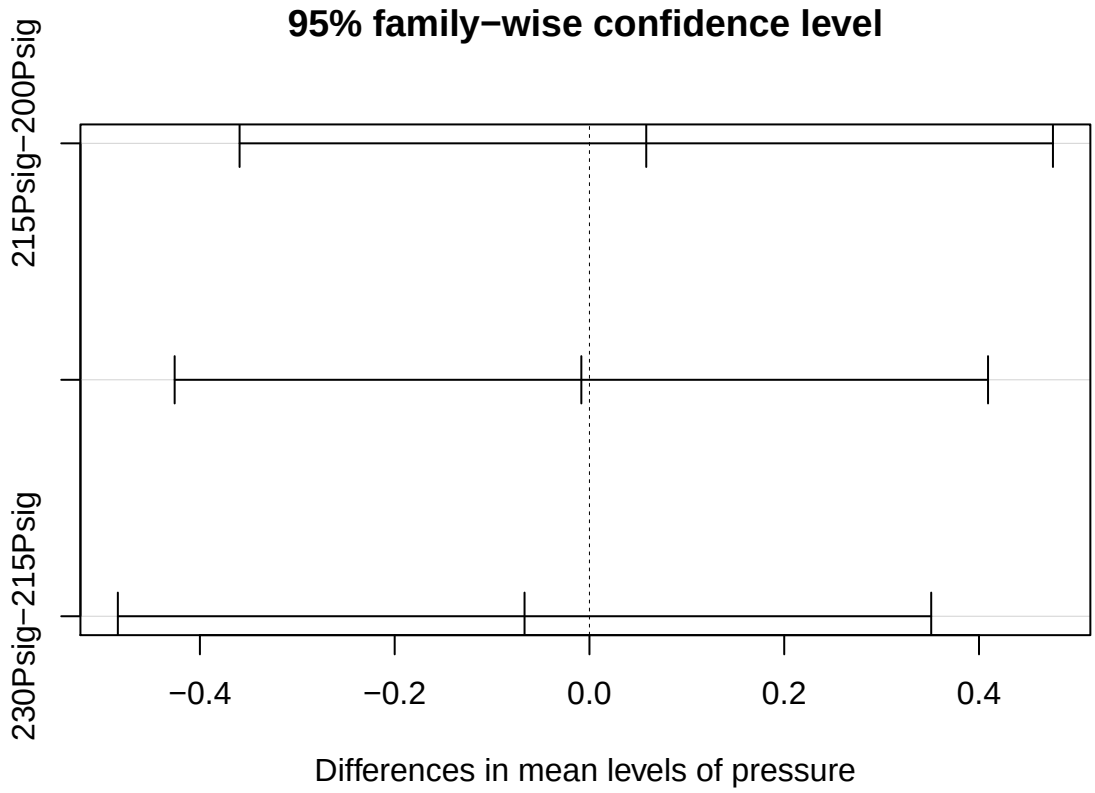
c. Under what condition would you operate this process

```
tukeys <- TukeyHSD(chem.process.nointeraction.mod)
tukeys
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = yield ~ temp + pressure)
##
## $temp
##                 diff        lwr       upr     p adj
## 160C-150C -0.1666667 -0.5842754 0.2509420 0.5578113
## 170C-150C  0.1500000 -0.2676087 0.5676087 0.6206552
## 170C-160C  0.3166667 -0.1009420 0.7342754 0.1509184
##
## $pressure
##                           diff        lwr       upr     p adj
## 215Psig-200Psig  0.058333333 -0.3592754 0.4759420 0.9281651
## 230Psig-200Psig -0.008333333 -0.4259420 0.4092754 0.9984708
## 230Psig-215Psig -0.066666667 -0.4842754 0.3509420 0.9073869
```

```
plot(tukeys)
```





Pvalue is large for both pressure, and temperature at all conditions, so we accept null hypothesis that there is not much difference due to the changing conditions

```
library(agricolae)
```

```
## Warning: package 'agricolae' was built under R version 4.0.5
```

```
lsd <- LSD.test(chem.process.nointeraction.mod, c("temp", "pressure"), console=T)
```

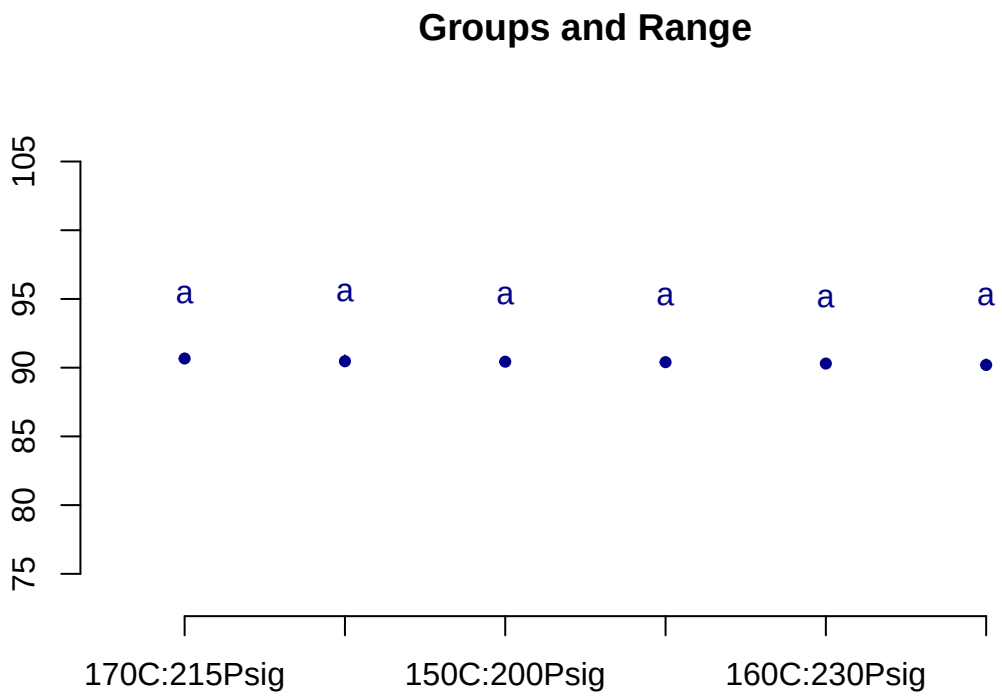
```
##
## Study: chem.process.nointeraction.mod ~ c("temp", "pressure")
##
## LSD t Test for yield
##
## Mean Square Error:  0.07504274
##
## temp:pressure,  means and individual ( 95 %) CI
##
##                 yield       std r      LCL      UCL  Min  Max
## 150C:200Psig 90.43333 0.2516611 3 90.09165 90.77501 90.2 90.7
## 150C:215Psig 90.40000 0.2000000 3 90.05832 90.74168 90.2 90.6
## 160C:200Psig 90.20000 0.3605551 3 89.85832 90.54168 89.9 90.6
## 160C:230Psig 90.30000 0.2000000 3 89.95832 90.64168 90.1 90.5
## 170C:215Psig 90.66667 0.1527525 3 90.32499 91.00835 90.5 90.8
## 170C:230Psig 90.46667 0.4041452 3 90.12499 90.80835 90.1 90.9
```

```
##
## Alpha: 0.05 ; DF Error: 13
## Critical Value of t: 2.160369
##
## least Significant Difference: 0.4832107
##
## Treatments with the same letter are not significantly different.
##
##                 yield groups
## 170C:215Psig 90.66667      a
## 170C:230Psig 90.46667      a
## 150C:200Psig 90.43333      a
## 150C:215Psig 90.40000      a
## 160C:230Psig 90.30000      a
## 160C:200Psig 90.20000      a
```

```
lsd
```

```
## $statistics
##      MSerror Df     Mean        CV  t.value       LSD
##   0.07504274 13 90.41111 0.3029929 2.160369 0.4832107
##
## $parameters
##         test p.ajusted        name.t ntr alpha
##   Fisher-LSD      none temp:pressure   6  0.05
##
## $means
##                 yield       std r      LCL      UCL  Min  Max   Q25  Q50   Q75
## 150C:200Psig 90.43333 0.2516611 3 90.09165 90.77501 90.2 90.7 90.30 90.4 90.55
## 150C:215Psig 90.40000 0.2000000 3 90.05832 90.74168 90.2 90.6 90.30 90.4 90.50
## 160C:200Psig 90.20000 0.3605551 3 89.85832 90.54168 89.9 90.6 90.00 90.1 90.35
## 160C:230Psig 90.30000 0.2000000 3 89.95832 90.64168 90.1 90.5 90.20 90.3 90.40
## 170C:215Psig 90.66667 0.1527525 3 90.32499 91.00835 90.5 90.8 90.60 90.7 90.75
## 170C:230Psig 90.46667 0.4041452 3 90.12499 90.80835 90.1 90.9 90.25 90.4 90.65
##
## $comparison
## NULL
##
## $groups
##                 yield groups
## 170C:215Psig 90.66667      a
## 170C:230Psig 90.46667      a
## 150C:200Psig 90.43333      a
## 150C:215Psig 90.40000      a
## 160C:230Psig 90.30000      a
## 160C:200Psig 90.20000      a
##
## attr(,"class")
## [1] "group"
```

```
plot(lsd)
```




```
detach()
```

All terms generate almost the same yield, so we should choose the one which can be operatored at least cost. I would choose the one at 150C and 200Psig

Question 2

setup data

```
operator<- rep(c(1,2,3), each=8)
machine<- rep(c(1,2,3,4), each=2, times=3)

strength <- c(109, 110,110,115,108,109,110,108,110,112,110,111,111,109,114,112,116,114,112,115,114,119,

fiber <- data.frame(operator, machine, strength)

fiber$operator <-factor(operator, levels =c(1,2,3), labels=c("Op1", "Op2", "Op3"))
fiber$machine <-factor(machine, levels =c(1,2,3,4), labels=c("Mach1","Mach2","Mach3","Mach4"))
operator <-fiber$operator
machine <-fiber$machine

head(fiber)
```

```
##   operator machine strength
```

```
## 1      Op1   Mach1      109
## 2      Op1   Mach1      110
## 3      Op1   Mach2      110
## 4      Op1   Mach2      115
## 5      Op1   Mach3      108
## 6      Op1   Mach3      109
```

```
str(fiber)
```

```
## 'data.frame':    24 obs. of  3 variables:
##  $ operator: Factor w/ 3 levels "Op1","Op2","Op3": 1 1 1 1 1 1 1 1 2 2 ...
##  $ machine : Factor w/ 4 levels "Mach1","Mach2",..: 1 1 2 2 3 3 4 4 1 1 ...
##  $ strength: num  109 110 110 115 108 109 110 108 110 112 ...
```

```
attach(fiber)
```

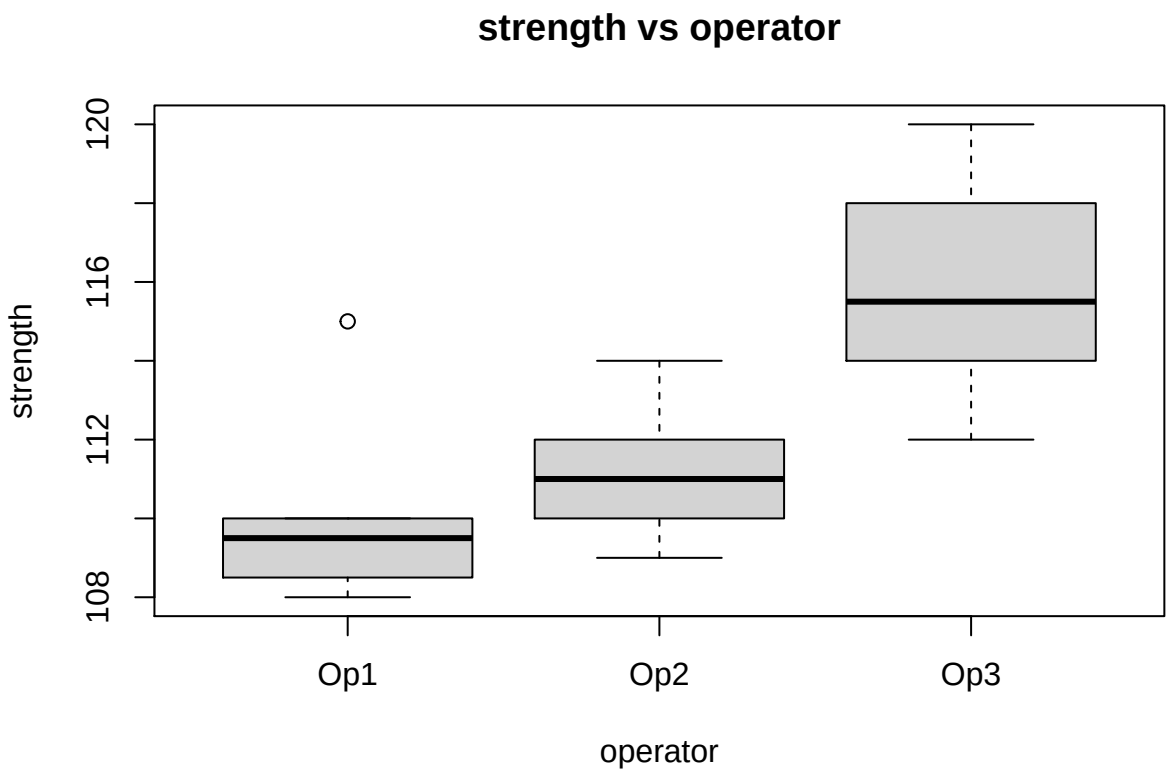
```
## The following objects are masked _by_ .GlobalEnv:
##
##     machine, operator, strength
```

a1. Plotting: boxplot of strength vs operator, strength vs machine, strength vs operator and machine, interaction plot

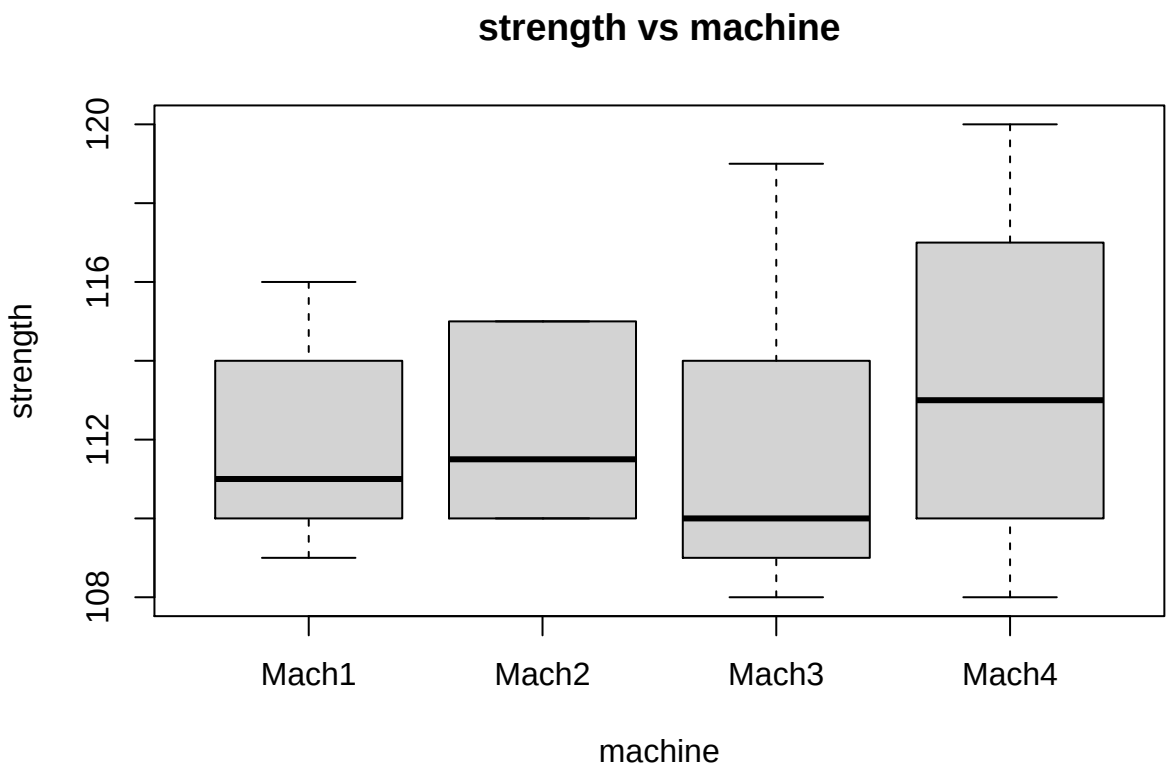
```
tapply(strength, list(operator,machine), mean)
```

```
##     Mach1 Mach2 Mach3 Mach4
## Op1 109.5 112.5 108.5 109.0
## Op2 111.0 110.5 110.0 113.0
## Op3 115.0 113.5 116.5 118.5
```

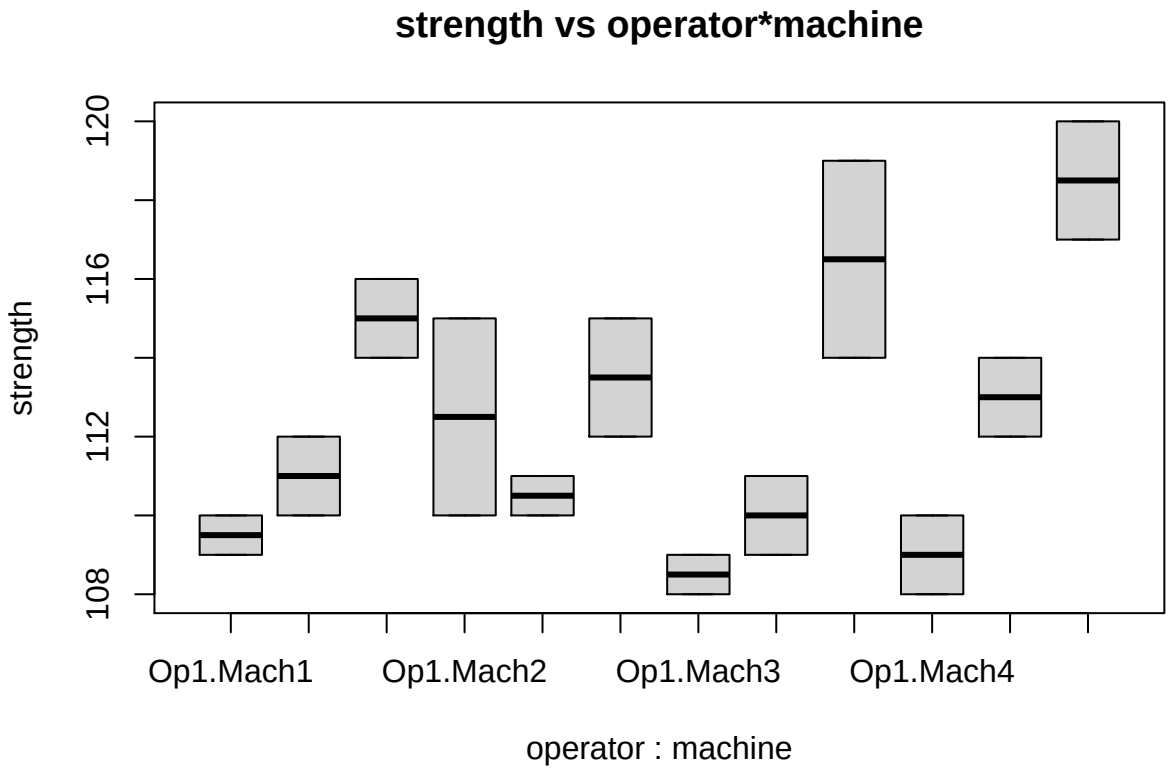
```
boxplot(strength~operator, main="strength vs operator")
```



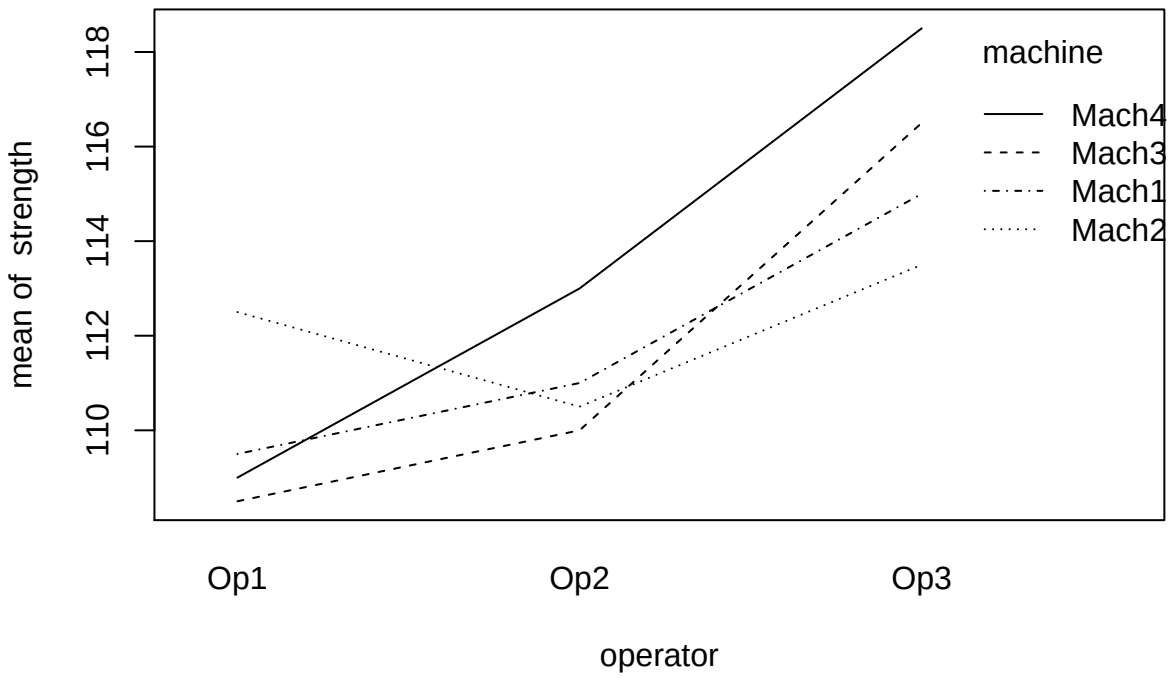

```
boxplot(strength~machine, main="strength vs machine")
```



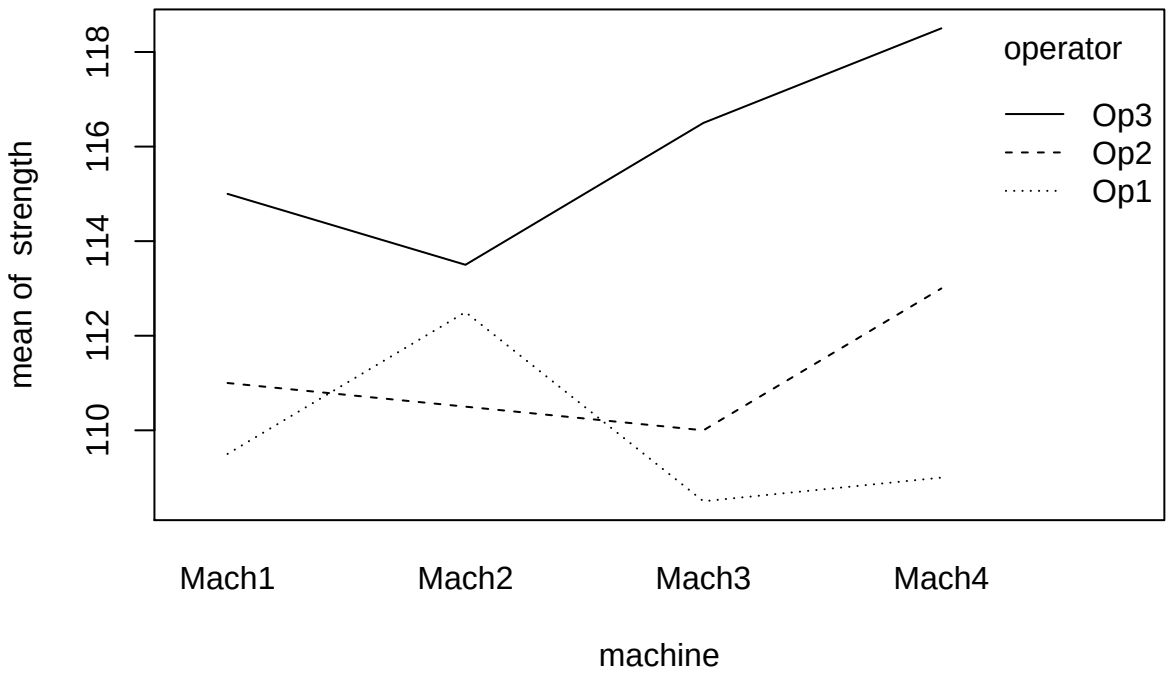

```
boxplot(strength~operator*machine, main="strength vs operator*machine")
```




```
interaction.plot(operator, machine, strength)
```




```
interaction.plot(machine, operator, strength)
```



Interaction plots are intersecting, so there is a indication of strong interaction

a2. Build a linear model using aov(). Print ANOVA output. Are main factors and interaction factor significant?

```
fiber.mod <- aov(strength~operator*machine)
summary.aov(fiber.mod)
```

```
##                  Df Sum Sq Mean Sq F value   Pr(>F)
## operator          2 160.33   80.17  21.143 0.000117 ***
## machine           3  12.46    4.15   1.095 0.388753
## operator:machine  6  44.67    7.44   1.963 0.150681
## Residuals        12  45.50    3.79
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary.lm(fiber.mod)
```

```
##
## Call:
## aov(formula = strength ~ operator * machine)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##   -2.5   -1.0    0.0    1.0    2.5
##
```

```
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                1.095e+02  1.377e+00  79.527   <2e-16 ***
## operatorOp2                1.500e+00  1.947e+00   0.770   0.4560
## operatorOp3                5.500e+00  1.947e+00   2.825   0.0153 *
## machineMach2               3.000e+00  1.947e+00   1.541   0.1493
## machineMach3              -1.000e+00  1.947e+00  -0.514   0.6169
## machineMach4              -5.000e-01  1.947e+00  -0.257   0.8017
## operatorOp2:machineMach2 -3.500e+00  2.754e+00  -1.271   0.2278
## operatorOp3:machineMach2 -4.500e+00  2.754e+00  -1.634   0.1282
## operatorOp2:machineMach3  8.573e-14  2.754e+00   0.000   1.0000
## operatorOp3:machineMach3  2.500e+00  2.754e+00   0.908   0.3818
## operatorOp2:machineMach4  2.500e+00  2.754e+00   0.908   0.3818
## operatorOp3:machineMach4  4.000e+00  2.754e+00   1.453   0.1720
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.947 on 12 degrees of freedom
## Multiple R-squared:  0.827,  Adjusted R-squared:  0.6684
## F-statistic: 5.214 on 11 and 12 DF,  p-value: 0.004136
```

```
operator
```

```
##  [1] Op1 Op1 Op1 Op1 Op1 Op1 Op1 Op1 Op2 Op2 Op2 Op2 Op2 Op2 Op2 Op2 Op3 Op3 Op3
## [20] Op3 Op3 Op3 Op3 Op3
## Levels: Op1 Op2 Op3
```

```
fiber.nointeraction.mod <- aov(strength~operator+machine)
anova(fiber.mod, fiber.nointeraction.mod)
```

```
## Analysis of Variance Table
##
## Model 1: strength ~ operator * machine
## Model 2: strength ~ operator + machine
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1     12 45.500
## 2     18 90.167 -6   -44.667 1.9634 0.1507
```

Since P-value >.05, we accept null and reject Alternate. The interaction term is in-significant, so we can discard the interaction term. We take the partial model instead of full model

a3. Print out regression model with coefficients

```
summary.lm(fiber.nointeraction.mod)
```

```
##
## Call:
## aov(formula = strength ~ operator + machine)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.7500 -1.1250 -0.2917  0.5833  5.2500
```

```
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  109.4167     1.1191  97.775  < 2e-16 ***
## operatorOp2    1.2500     1.1191   1.117    0.279
## operatorOp3    6.0000     1.1191   5.362 4.27e-05 ***
## machineMach2   0.3333     1.2922   0.258    0.799
## machineMach3  -0.1667     1.2922  -0.129    0.899
## machineMach4   1.6667     1.2922   1.290    0.213
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.238 on 18 degrees of freedom
## Multiple R-squared:  0.6571, Adjusted R-squared:  0.5619
## F-statistic: 6.899 on 5 and 18 DF,  p-value: 0.00093
```

Intercept and Operator are both siginficant

b. Prepare appropriate residuals plot and comment on the model's adequecy

```
summary.lm(fiber.nointeraction.mod)
```

```
##
## Call:
## aov(formula = strength ~ operator + machine)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.7500 -1.1250 -0.2917  0.5833  5.2500
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  109.4167     1.1191  97.775  < 2e-16 ***
## operatorOp2    1.2500     1.1191   1.117    0.279
## operatorOp3    6.0000     1.1191   5.362 4.27e-05 ***
## machineMach2   0.3333     1.2922   0.258    0.799
## machineMach3  -0.1667     1.2922  -0.129    0.899
## machineMach4   1.6667     1.2922   1.290    0.213
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.238 on 18 degrees of freedom
## Multiple R-squared:  0.6571, Adjusted R-squared:  0.5619
## F-statistic: 6.899 on 5 and 18 DF,  p-value: 0.00093
```

```
summary.aov(fiber.nointeraction.mod)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## operator     2 160.33   80.17  16.004 0.000101 ***
## machine      3  12.46    4.15   0.829 0.495098
## Residuals   18  90.17    5.01
```

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

It looks like the difference of mean is same for machine while it's different for operator. We need to run Tukey's test to see which one are different

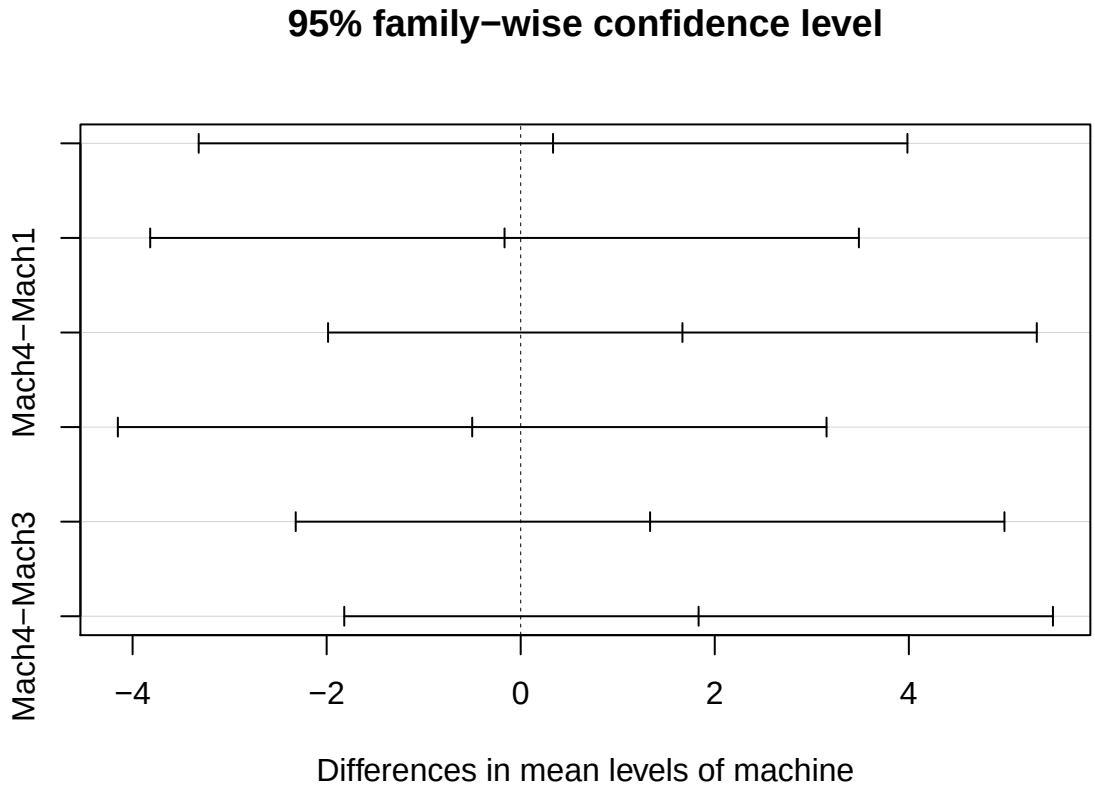
c. Under what condition would you operate this process

```
tukeys <- TukeyHSD(fiber.nointeraction.mod)
tukeys
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = strength ~ operator + machine)
##
## $operator
##         diff       lwr      upr     p adj
## Op2-Op1 1.25 -1.606046 4.106046 0.5162537
## Op3-Op1 6.00  3.143954 8.856046 0.0001210
## Op3-Op2 4.75  1.893954 7.606046 0.0013473
##
## $machine
##                    diff       lwr      upr     p adj
## Mach2-Mach1  0.3333333 -3.318768 3.985435 0.9937651
## Mach3-Mach1 -0.1666667 -3.818768 3.485435 0.9992035
## Mach4-Mach1  1.6666667 -1.985435 5.318768 0.5807291
## Mach3-Mach2 -0.5000000 -4.152102 3.152102 0.9796977
## Mach4-Mach2  1.3333333 -2.318768 4.985435 0.7334070
## Mach4-Mach3  1.8333333 -1.818768 5.485435 0.5043547
```

```
plot(tukeys)
```


95% family−wise confidence level
Op2−Op1
Op3−Op1
Op3−Op2
−2
0
2
4
6
8
Differences in mean levels of operator



Pvalue is large for both pressure, and temperature at all conditions, so we accept null hypothesis that there is not much difference due to the changing conditions

```
library(agricolae)
```

```
## Warning: package 'agricolae' was built under R version 4.0.5
```

```
lsd <- LSD.test(fiber.nointeraction.mod, c("operator", "machine"), console=T)
```

```
##
## Study: fiber.nointeraction.mod ~ c("operator", "machine")
##
## LSD t Test for strength
##
## Mean Square Error:  5.009259
##
## operator:machine,  means and individual ( 95 %) CI
##
##           strength       std r      LCL      UCL Min Max
## Op1:Mach1    109.5 0.7071068 2 106.1751 112.8249 109 110
## Op1:Mach2    112.5 3.5355339 2 109.1751 115.8249 110 115
## Op1:Mach3    108.5 0.7071068 2 105.1751 111.8249 108 109
## Op1:Mach4    109.0 1.4142136 2 105.6751 112.3249 108 110
## Op2:Mach1    111.0 1.4142136 2 107.6751 114.3249 110 112
## Op2:Mach2    110.5 0.7071068 2 107.1751 113.8249 110 111
```

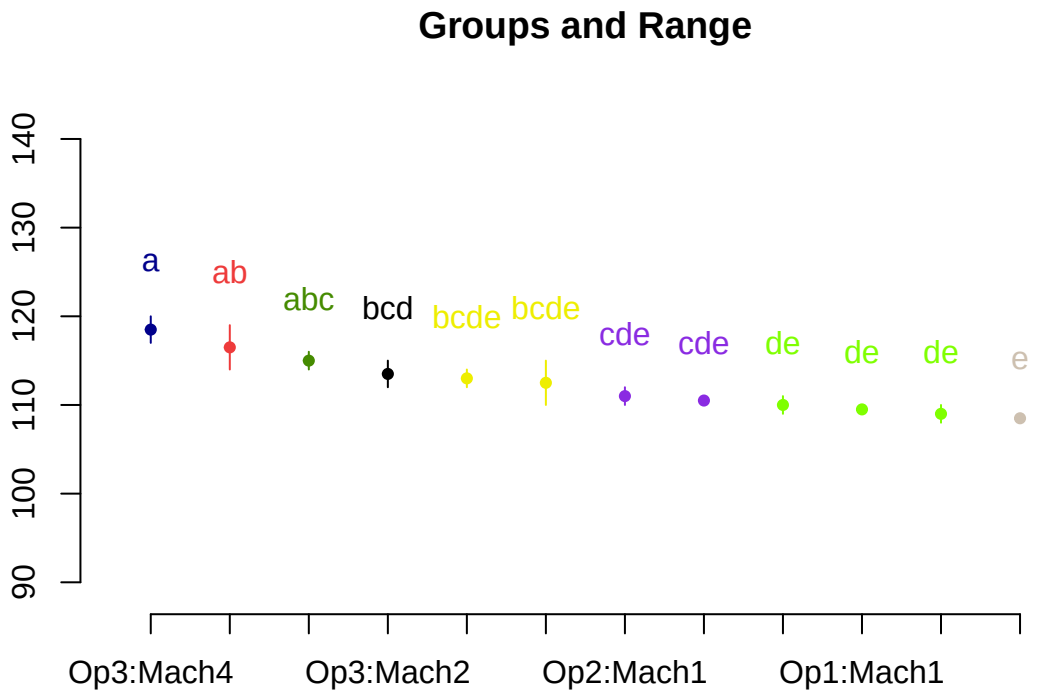
```
## Op2:Mach3    110.0 1.4142136 2 106.6751 113.3249 109 111
## Op2:Mach4    113.0 1.4142136 2 109.6751 116.3249 112 114
## Op3:Mach1    115.0 1.4142136 2 111.6751 118.3249 114 116
## Op3:Mach2    113.5 2.1213203 2 110.1751 116.8249 112 115
## Op3:Mach3    116.5 3.5355339 2 113.1751 119.8249 114 119
## Op3:Mach4    118.5 2.1213203 2 115.1751 121.8249 117 120
##
## Alpha: 0.05 ; DF Error: 18
## Critical Value of t: 2.100922
##
## least Significant Difference: 4.702152
##
## Treatments with the same letter are not significantly different.
##
##           strength groups
## Op3:Mach4    118.5      a
## Op3:Mach3    116.5     ab
## Op3:Mach1    115.0    abc
## Op3:Mach2    113.5    bcd
## Op2:Mach4    113.0   bcde
## Op1:Mach2    112.5   bcde
## Op2:Mach1    111.0    cde
## Op2:Mach2    110.5    cde
## Op2:Mach3    110.0     de
## Op1:Mach1    109.5     de
## Op1:Mach4    109.0     de
## Op1:Mach3    108.5      e
```

```
lsd
```

```
## $statistics
##    MSerror Df     Mean       CV  t.value      LSD
##   5.009259 18 112.2917 1.993147 2.100922 4.702152
##
## $parameters
##         test p.ajusted           name.t ntr alpha
##   Fisher-LSD      none operator:machine  12  0.05
##
## $means
##           strength       std r      LCL      UCL Min Max    Q25   Q50    Q75
## Op1:Mach1    109.5 0.7071068 2 106.1751 112.8249 109 110 109.25 109.5 109.75
## Op1:Mach2    112.5 3.5355339 2 109.1751 115.8249 110 115 111.25 112.5 113.75
## Op1:Mach3    108.5 0.7071068 2 105.1751 111.8249 108 109 108.25 108.5 108.75
## Op1:Mach4    109.0 1.4142136 2 105.6751 112.3249 108 110 108.50 109.0 109.50
## Op2:Mach1    111.0 1.4142136 2 107.6751 114.3249 110 112 110.50 111.0 111.50
## Op2:Mach2    110.5 0.7071068 2 107.1751 113.8249 110 111 110.25 110.5 110.75
## Op2:Mach3    110.0 1.4142136 2 106.6751 113.3249 109 111 109.50 110.0 110.50
## Op2:Mach4    113.0 1.4142136 2 109.6751 116.3249 112 114 112.50 113.0 113.50
## Op3:Mach1    115.0 1.4142136 2 111.6751 118.3249 114 116 114.50 115.0 115.50
## Op3:Mach2    113.5 2.1213203 2 110.1751 116.8249 112 115 112.75 113.5 114.25
## Op3:Mach3    116.5 3.5355339 2 113.1751 119.8249 114 119 115.25 116.5 117.75
## Op3:Mach4    118.5 2.1213203 2 115.1751 121.8249 117 120 117.75 118.5 119.25
##
## $comparison
```

```
## NULL
##
## $groups
##           strength groups
## Op3:Mach4    118.5      a
## Op3:Mach3    116.5     ab
## Op3:Mach1    115.0    abc
## Op3:Mach2    113.5    bcd
## Op2:Mach4    113.0   bcde
## Op1:Mach2    112.5   bcde
## Op2:Mach1    111.0    cde
## Op2:Mach2    110.5    cde
## Op2:Mach3    110.0     de
## Op1:Mach1    109.5     de
## Op1:Mach4    109.0     de
## Op1:Mach3    108.5      e
##
## attr(,"class")
## [1] "group"
```

```
plot(lsd)
```



Operator3 and machine 4 appears to give the best strength

Question 3.

a. create data frame with days and solution and growth as response variable

```
solution <-rep(c(1,2,3), each=4)
days <-  rep(c(1,2,3,4), each=3)
solution <- factor(solution, levels=c(1,2,3), labels = c("Sol1", "Sol2", "Sol3"))
days <- factor(days, levels =c(1,2,3,4), labels = c("Day1", "Day2", "Day3", "Day4"))

growth<-c(13,22,18,39,16,24,17,44,5,4,1,22)

powder <-data.frame(solution, days, growth)

head(powder)
```

```
##   solution days growth
## 1     Sol1 Day1     13
## 2     Sol1 Day1     22
## 3     Sol1 Day1     18
## 4     Sol1 Day2     39
## 5     Sol2 Day2     16
## 6     Sol2 Day2     24
```

```
str(powder)
```

```
## 'data.frame':    12 obs. of  3 variables:
##  $ solution: Factor w/ 3 levels "Sol1","Sol2",..: 1 1 1 1 2 2 2 2 3 3 ...
##  $ days    : Factor w/ 4 levels "Day1","Day2",..: 1 1 1 2 2 2 3 3 3 4 ...
##  $ growth  : num  13 22 18 39 16 24 17 44 5 4 ...
```

```
attach(powder)
```

```
## The following objects are masked _by_ .GlobalEnv:
##
##     days, growth, solution
```

Any evidence that the solution affect bacteria growth

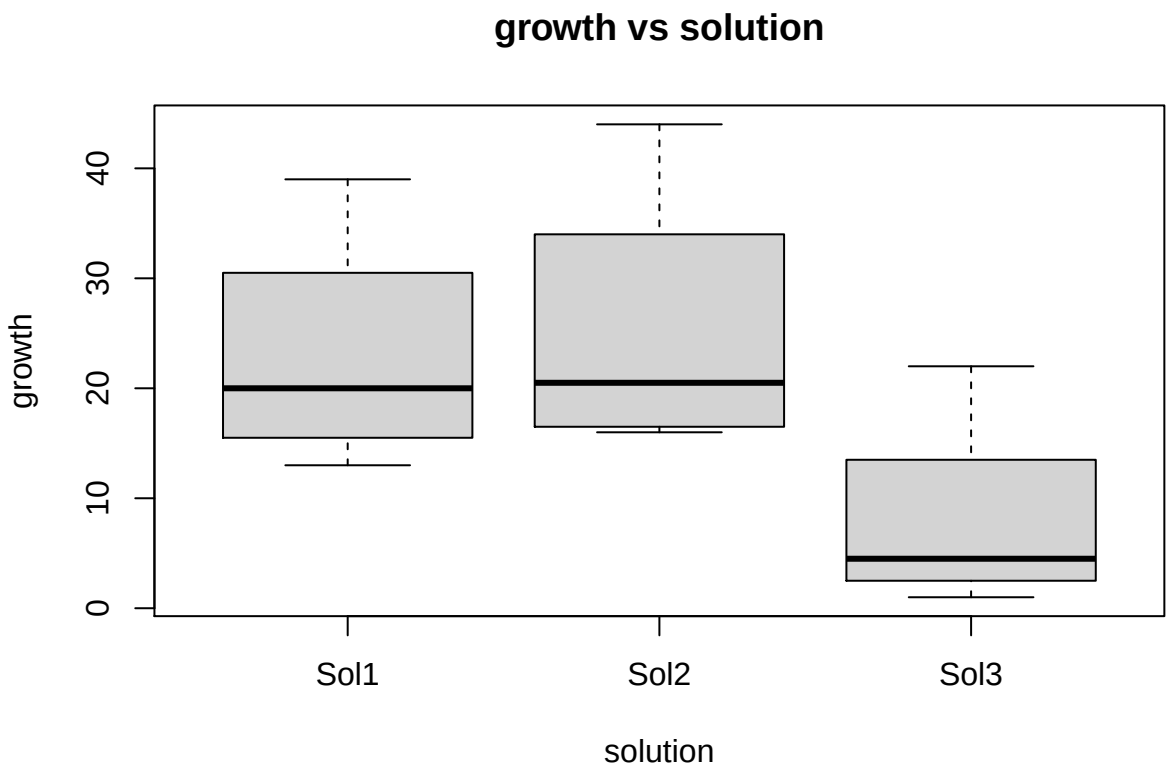
```
length(growth)
```

```
## [1] 12
```

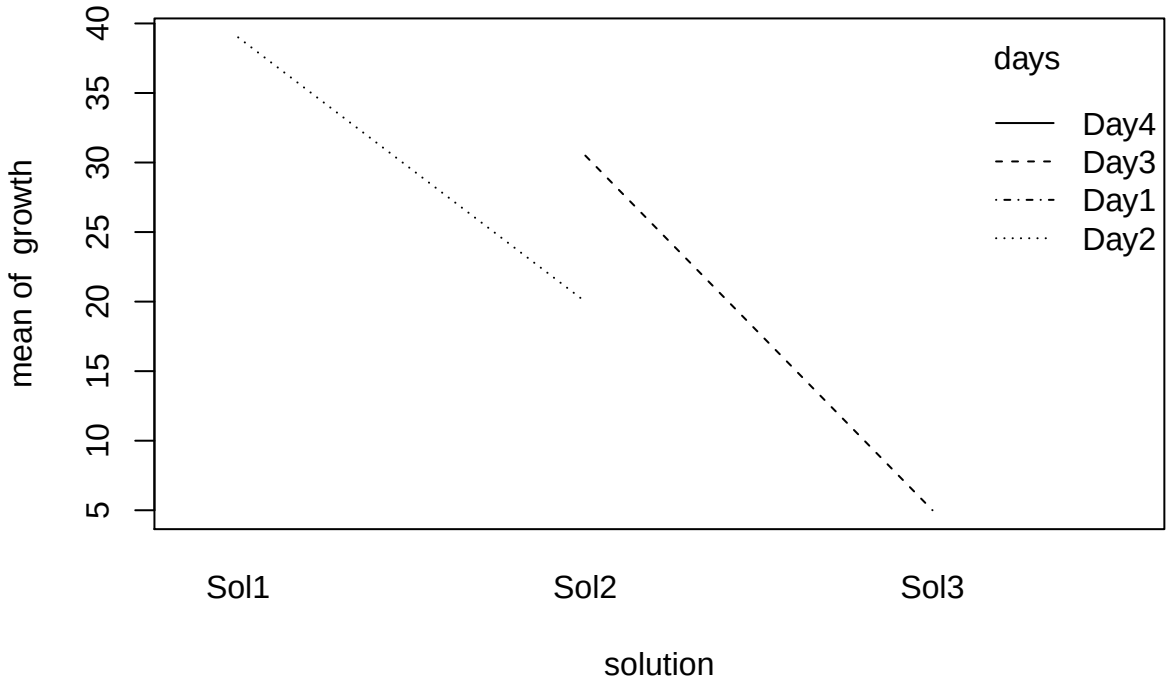
```
length(solution)
```

```
## [1] 12
```

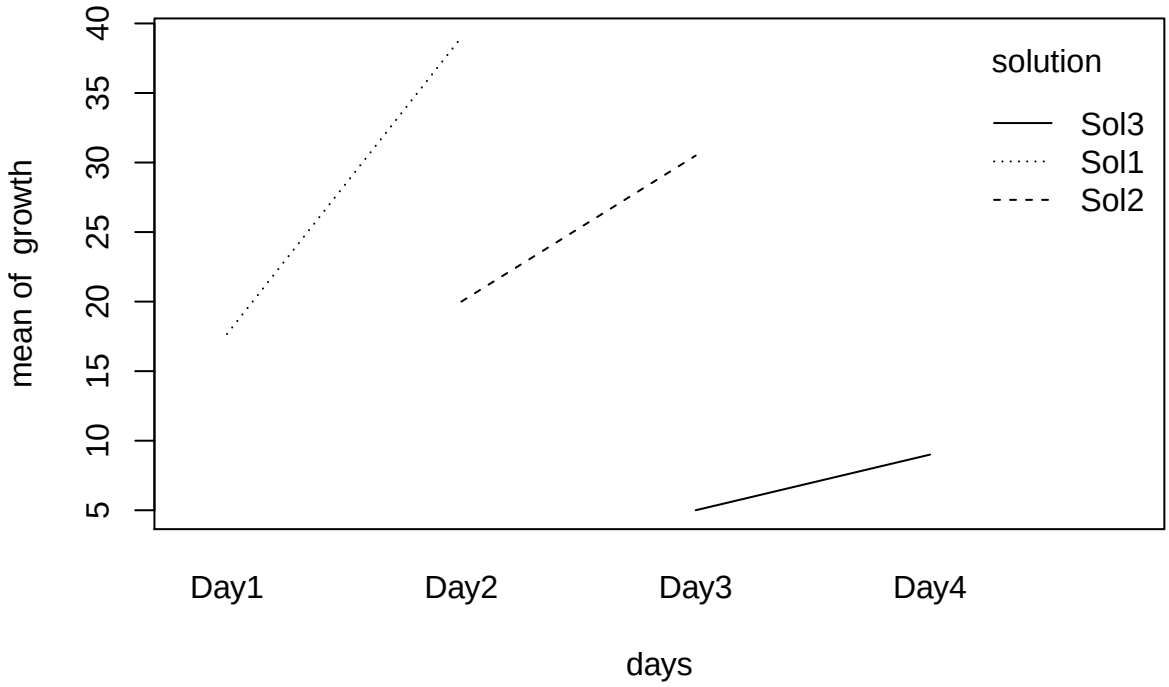
```
boxplot(growth~solution, main="growth vs solution")
```




```
interaction.plot(solution, days, growth)
```




```
interaction.plot(days, solution, growth)
```



The interaction appears to be weak cause these interaction plots are intersecting.

```
powder.mod <- aov(growth~solution*days)
summary.aov(powder.mod)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## solution     2  703.5   351.8   3.036  0.123
## days         3  463.6   154.5   1.334  0.348
## Residuals    6  695.2   115.9
```

```
summary.lm(powder.mod)
```

```
##
## Call:
## aov(formula = growth ~ solution * days)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -13.500  -4.750   0.000   4.083  13.500
##
## Coefficients: (6 not defined because of singularities)
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)             17.667      6.215   2.843   0.0295 *
## solutionSol2           -19.000     13.183  -1.441   0.1996
## solutionSol3           -44.500     18.644  -2.387   0.0543 .
```

```
## daysDay2                21.333     12.429   1.716   0.1369
## daysDay3                31.833     16.442   1.936   0.1010
## daysDay4                35.833     20.611   1.739   0.1328
## solutionSol2:daysDay2       NA         NA      NA       NA
## solutionSol3:daysDay2       NA         NA      NA       NA
## solutionSol2:daysDay3       NA         NA      NA       NA
## solutionSol3:daysDay3       NA         NA      NA       NA
## solutionSol2:daysDay4       NA         NA      NA       NA
## solutionSol3:daysDay4       NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.76 on 6 degrees of freedom
## Multiple R-squared:  0.6267, Adjusted R-squared:  0.3156
## F-statistic: 2.015 on 5 and 6 DF,  p-value: 0.2093
```

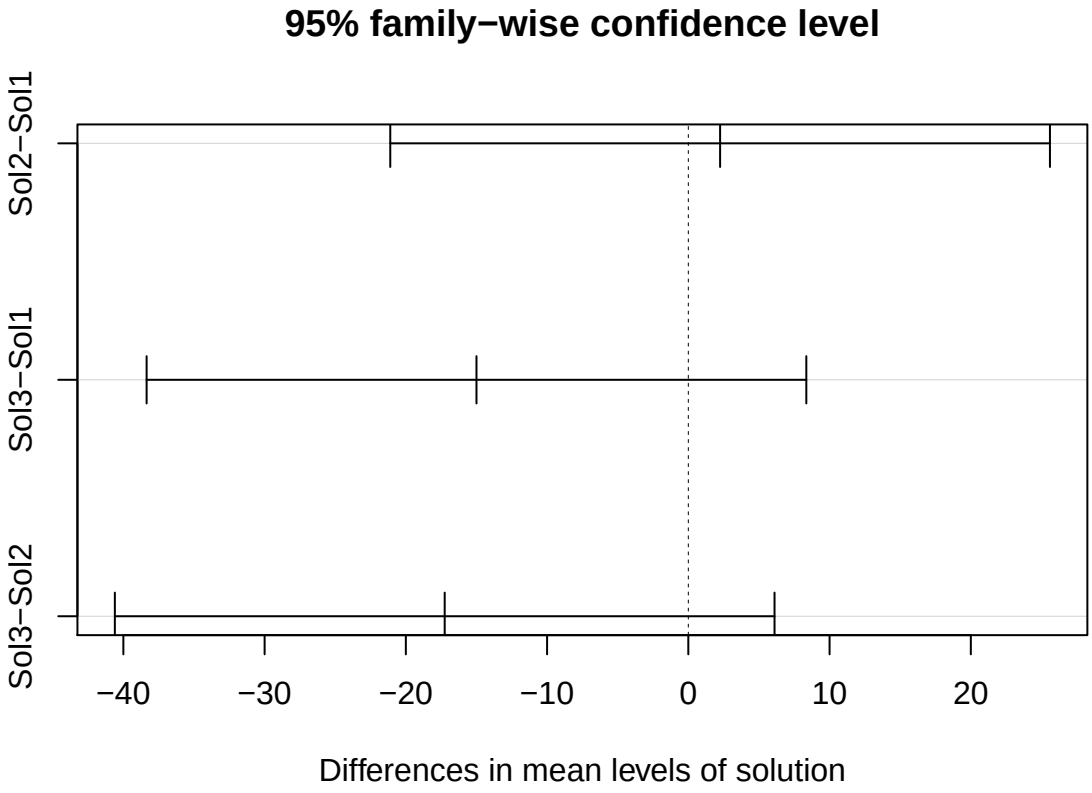
```
powder.nointeraction.mod <- aov(growth~solution+days)
anova(powder.mod, powder.nointeraction.mod)
```

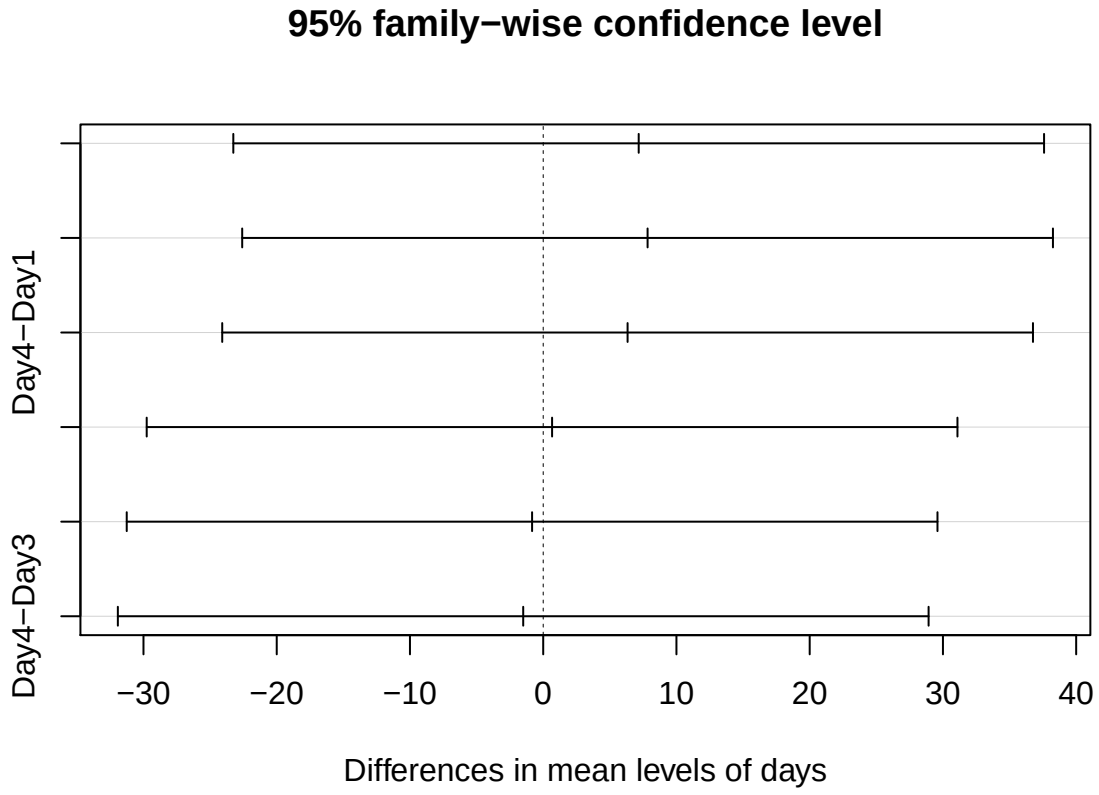
```
## Analysis of Variance Table
##
## Model 1: growth ~ solution * days
## Model 2: growth ~ solution + days
##   Res.Df    RSS Df Sum of Sq F Pr(>F)
## 1      6 695.17
## 2      6 695.17  0         0
```

Pvalue is 0, so there appears to be no interaction, so we chose non-interactive model.

b. Perform a TukeyHSD to compare solution treatment means

```
tukeys <- TukeyHSD(powder.nointeraction.mod)
plot(tukeys)
```


95% family−wise confidence level
Sol2−Sol1
Sol3−Sol1
Sol3−Sol2
−40
−30
−20
−10
0
10
20
Differences in mean levels of solution



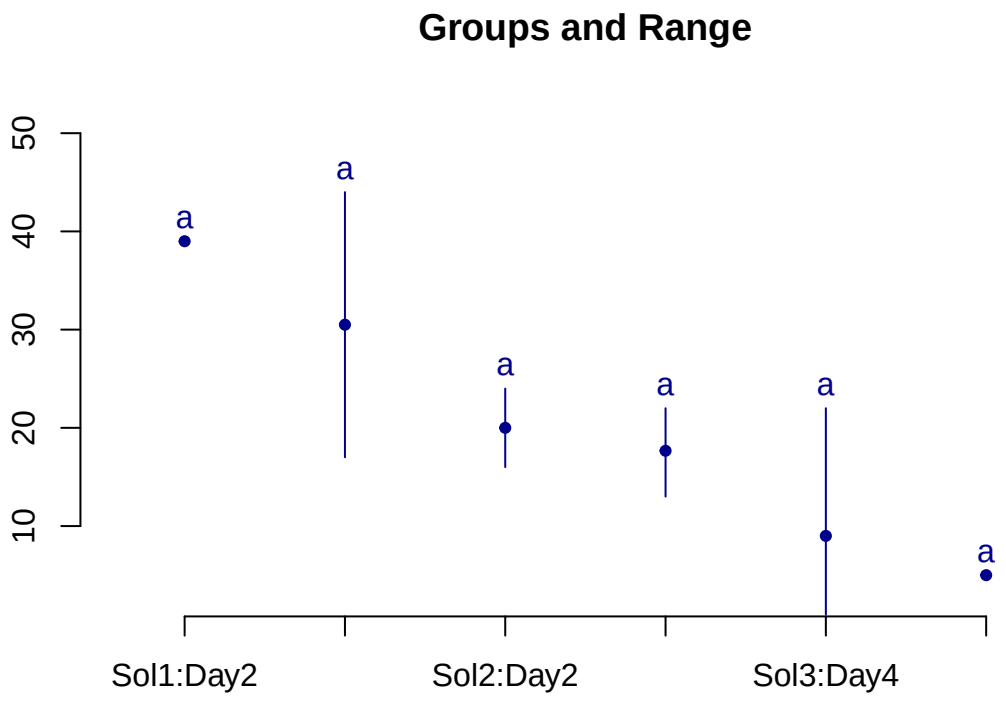
Difference of mean is the same so the solution or days doesn't affect much of the bacterial growth

```
lsd <- LSD.test(powder.nointeraction.mod, c("solution", "days"), console = T)
```

```
##
## Study: powder.nointeraction.mod ~ c("solution", "days")
##
## LSD t Test for growth
##
## Mean Square Error:  115.8611
##
## solution:days,  means and individual ( 95 %) CI
##
##             growth       std r        LCL      UCL Min Max
## Sol1:Day1 17.66667  4.509250 3   2.460262 32.87307  13  22
## Sol1:Day2 39.00000        NA 1  12.661735 65.33827  39  39
## Sol2:Day2 20.00000  5.656854 2   1.376034 38.62397  16  24
## Sol2:Day3 30.50000 19.091883 2  11.876034 49.12397  17  44
## Sol3:Day3  5.00000        NA 1 -21.338265 31.33827   5   5
## Sol3:Day4  9.00000 11.357817 3  -6.206405 24.20640   1  22
##
## Alpha: 0.05 ; DF Error: 6
## Critical Value of t: 2.446912
##
## Groups according to probability of means differences and alpha level( 0.05 )
##
```

```
## Treatments with the same letter are not significantly different.
##
##             growth groups
## Sol1:Day2 39.00000      a
## Sol2:Day3 30.50000      a
## Sol2:Day2 20.00000      a
## Sol1:Day1 17.66667      a
## Sol3:Day4  9.00000      a
## Sol3:Day3  5.00000      a
```

```
plot(lsd)
```

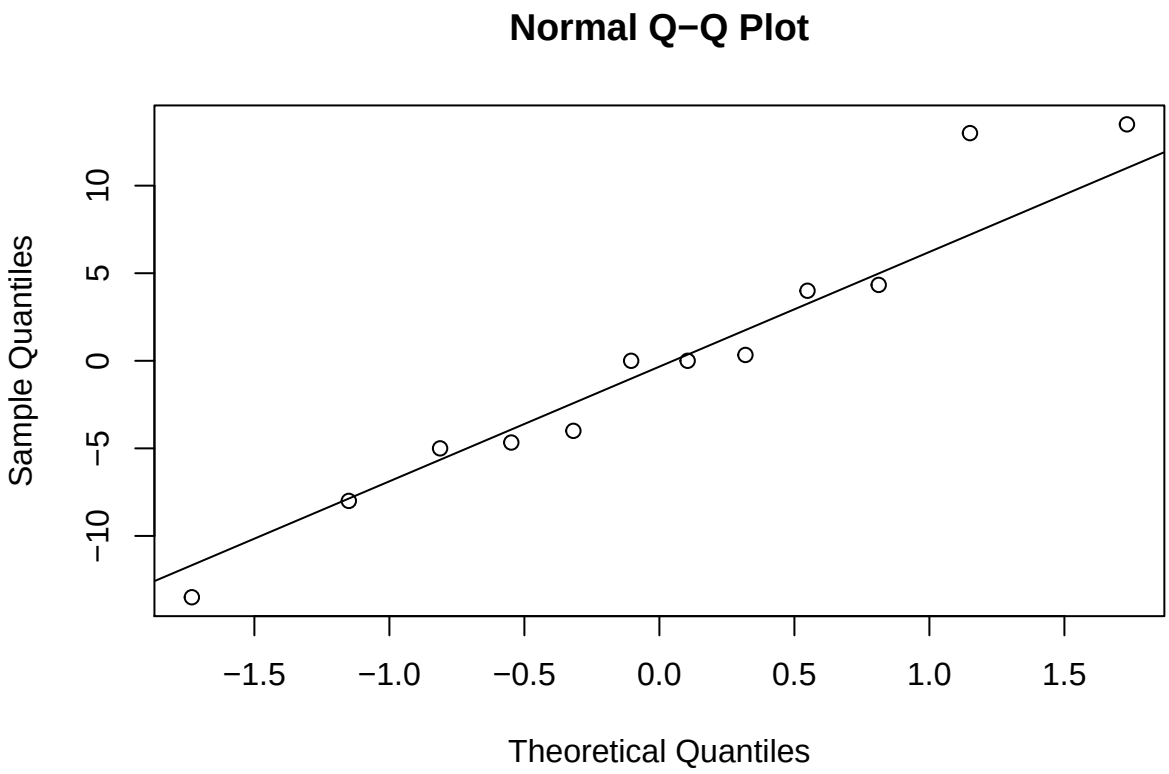


The solution and days doesn't appear to have much affect on growth.

c. Check assumptions of the residuals

```
res <- residuals(powder.nointeraction.mod)

qqnorm(res)
qqline(res)
```




```
shapiro.test(res)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  res
## W = 0.95242, p-value = 0.6726
```

```
#check variance
bartlett.test(res~solution)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  res by solution
## Bartlett's K-squared = 2.8626, df = 2, p-value = 0.239
```

```
bartlett.test(res~days)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  res by days
## Bartlett's K-squared = 3.3249, df = 3, p-value = 0.3442
```

The data is normal and variance is equal.