

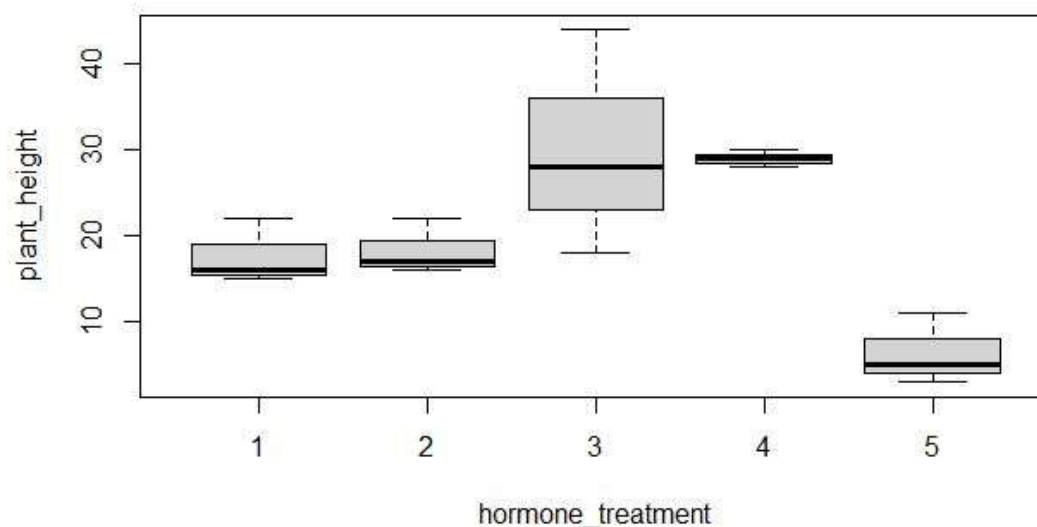
Question 1:

A Randomized Complete Block Design (RCBD) is a designed experiment which treatment combinations are located randomly in blocks to reduce effect of nuisance factors.

In this experiment, it is designed to compare outcomes of plants' growth and yields when two different hormones, auxin and gibberellin are applied. These hormones are treatment factors which may directly relate and affect to response. Plants are arranged in sand trays which are block factors which should not affect to outcomes. In this experiment, the height(cm) and the yield were recorded for each plant and these are two responses.

Question 2(a):

```
> dat1 = read.csv("ass1_data.csv", header = T)
> str(dat1)
'data.frame': 15 obs. of 4 variables:
 $ tray      : int  1 1 1 1 1 2 2 2 2 2 ...
 $ hormone_treatment: int  4 5 3 2 1 1 2 4 3 5 ...
 $ plant_height  : int  30 3 44 16 22 15 17 28 18 5 ...
 $ yield        : int 11181 31 20226 40084 19305 23074 8147 3045 3967 22153 ...
> dat1$hormone_treatment = factor(dat1$hormone_treatment)
> dat1$tray = factor(dat1$tray)
> str(dat1)
'data.frame': 15 obs. of 4 variables:
 $ tray      : Factor w/ 3 levels "1","2","3": 1 1 1 1 1 2 2 2 2 2 ...
 $ hormone_treatment: Factor w/ 5 levels "1","2","3","4",...: 4 5 3 2 1 1 2 4 3 5 ...
 $ plant_height  : int  30 3 44 16 22 15 17 28 18 5 ...
 $ yield        : int 11181 31 20226 40084 19305 23074 8147 3045 3967 22153 ...
> view(dat1)
> boxplot(plant_height~hormone_treatment, data=dat1)
```



Question 2(b):

According to the boxplot in 2(a), first, there may be differences between the groups, especially control group (5) and treatment group (1~4) seems to have significant distinction. Second, other groups look like they have less variability, but group 3 relatively has large range of responses. Third, each treatment group (e.g., 1&2, 3&4) has similar mean but seems 3&4 groups are little bit higher. In conclusion, it is possible to differ in responses by treatments, but it needs more analysis to figure out more precisely.

Question 2(c):

Since we already designated factors, directly carry out the analysis of variance

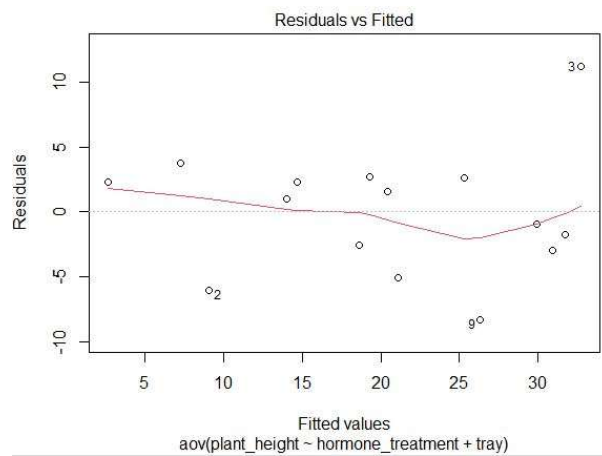
```
> dat1.aov1 = aov(plant_height~hormone_treatment + tray, data = dat1)
> summary(dat1.aov1)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
hormone_treatment	4	1126.9	281.73	7.020	0.00994	**
tray	2	108.9	54.47	1.357	0.31082	
Residuals	8	321.1	40.13			

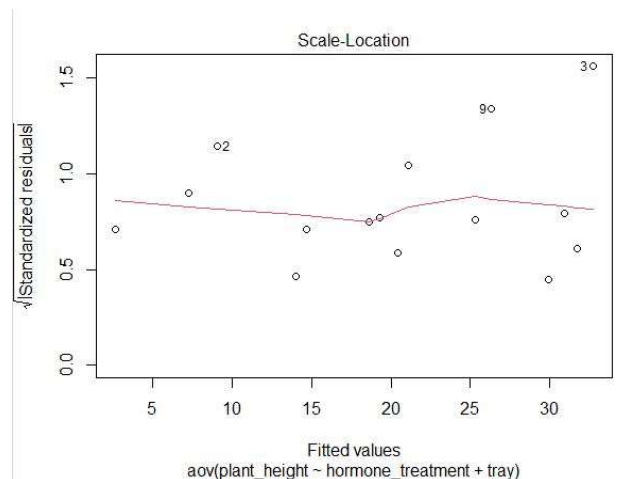
```
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> |
```

Question 2(d):

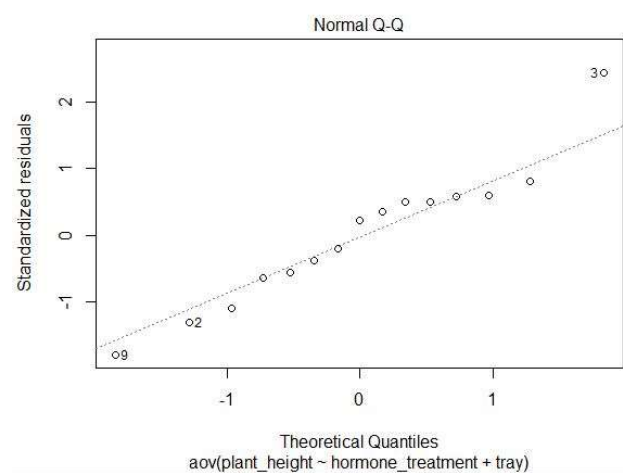
i. Residuals vs Fitted Values



ii. Standardized Residuals vs Fitted Values



iii. Q-Q plot for the Standardized Residuals



Question 2(e):

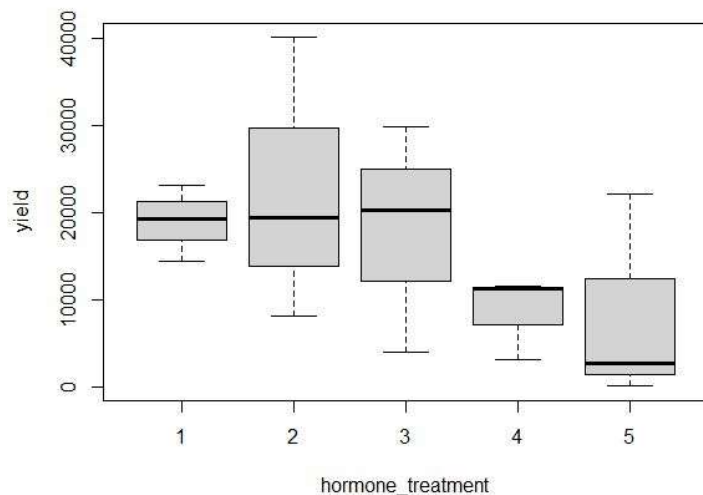
According to the three diagnostic plots in 2(d), there appears to be neither non-normality symptom from Q-Q plot for the Standardized Residuals nor heterogeneity from Residuals vs Fitted Values. Therefore, this experiment did not violate any of assumptions in ANOVA.

Question 2(f):

Based on the ANOVA table in 2(c), F-statistic of hormone treatment is 7.020 which can be calculated (mean sum of square factor / mean sum of square residuals) $281.73/40.13$. In F probability density function with 4 degree of freedom for factor and 8 degree of freedom for residual (F pdf with df (4,8)), F value (7.020) will take probability of 0.99 there for p-value is $1 - 0.992 = 0.01$ as defined in ANOVA table above which is very significant. This is enough evidence that there is sufficient difference in group. In conclusion, there are meaningful changes in plant's height growth by the hormone treatments. However more tests need to be conducted to see which treatment is effective.

Question 3(a):

```
> str(dat1)
'data.frame': 15 obs. of 4 variables:
 $ tray      : Factor w/ 3 levels "1","2","3": 1 1 1 1 1 2 2 2 2 2 ...
 $ hormone_treatment: Factor w/ 5 levels "1","2","3","4",...: 4 5 3 2 1 1 2 4 3 5 ...
 $ plant_height  : int  30 3 44 16 22 15 17 28 18 5 ...
 $ yield        : int 11181 31 20226 40084 19305 23074 8147 3045 3967 22153 ...
> boxplot(yield~hormone_treatment, data=dat1)
> |
```



Question 3(b):

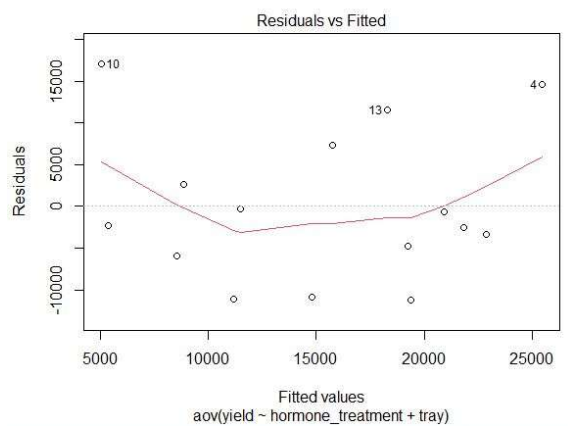
According to the boxplot in 3(a), it is hard to say if there are important data in the plot. There may be difference between control group and treatment groups. Other than that, each group has dynamic variability and means are not constant, so it is hard to see the pattern to analyze data.

Question 3(c):

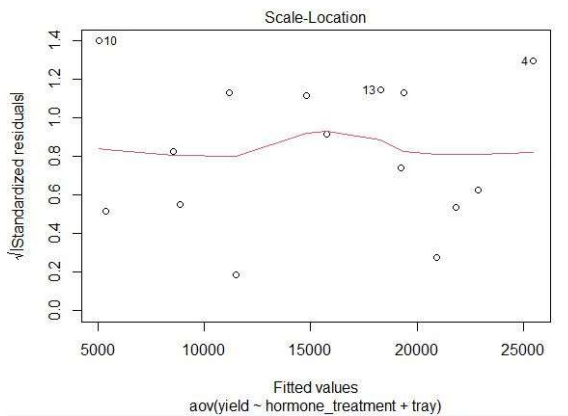
```
> dat1.aov2 = aov(yield~hormone_treatment + tray, data=dat1)
> summary(dat1.aov2)
              Df    Sum Sq   Mean Sq F value    Pr(>F)
hormone_treatment  4  5.038e+08 125940469   0.877    0.518
tray                2  9.332e+07  46659068   0.325    0.732
Residuals          8  1.149e+09 143631474
> |
```

Question 3(d):

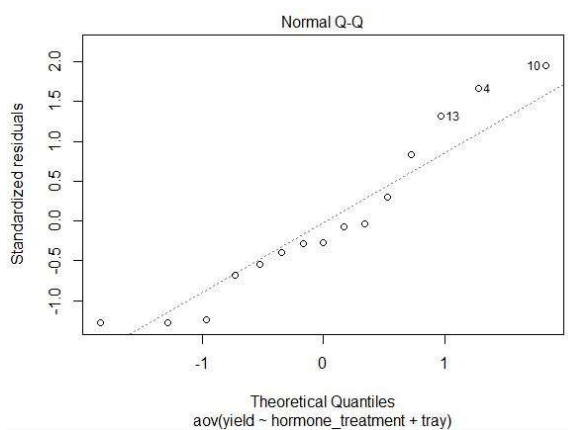
i. Residuals vs Fitted Values



ii. Standardized Residuals vs Fitted Values



iii. Q-Q plot for the Standardized Residuals



Question 3(e):

According to three diagnostic plots in 3(d), there appears to be no non-normality symptom from Q-Q plot for the Standardized Residuals, however there is an evidence that it may not have constant variance since its red line is steep and far from zero line in Residuals vs Fitted Values plot. It may violate assumptions of ANOVA.

Question 3(f):

Based on the ANOVA table in 3(c), F-statistic of hormone treatment is 0.877 which can be calculated (mean sum of square factor / mean sum of square residuals) $125940469/143631474$. In F probability density function with 4 degree of freedom for factor and 8 degree of freedom for residual (F pdf with df (4,8)), F value (0.877) will take probability of 0.482 there for p-value is $1 - 0.482 = 0.518$ which is not significant at all. This result tells that hormone treatments did not affect on yields of plants.

Question 4(a):

The hormone treatment affects significantly to growth of plant height.

Question 4(b):

```
> str(dat1$hormone_treatment)
Factor w/ 5 levels "1","2","3","4",...: 4 5 3 2 1 1 2 4 3 5 ...
> contrastmatrix = cbind( c(-1/4,-1/4,-1/4,-1/4,1), c(-1/2,-1/2,1/2,1/2,0), c(-1,1,0,
0,0), c(0,0,-1,1,0))
> contrasts(dat1$hormone_treatment) = contrastmatrix
> dat1$forcont = aov(plant_height~hormone_treatment + tray, data=dat1)
> summary(dat1$forcont)

Df Sum Sq Mean Sq F value Pr(>F)
hormone_treatment 4 1126.9 281.73 7.020 0.00994 **
tray 2 108.9 54.47 1.357 0.31082
Residuals 8 321.1 40.13

---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> summary(dat1$forcont, split=list(hormone_treatment = list("Treatment vs Control"=1,
"Auxin vs Gibberellin"=2, "Auxin1 vs Auxin2"=3, "Gibberellin1 vs Gibberellin2"=4)))

Df Sum Sq Mean Sq F value
hormone_treatment 4 1126.9 281.7 7.020
hormone_treatment: Treatment vs Control 1 728.0 728.0 18.140
hormone_treatment: Auxin vs Gibberellin 1 396.7 396.7 9.886
hormone_treatment: Auxin1 vs Auxin2 1 0.7 0.7 0.017
hormone_treatment: Gibberellin1 vs Gibberellin2 1 1.5 1.5 0.037
tray 2 108.9 54.5 1.357
Residuals 8 321.1 40.1

Pr(>F)
hormone_treatment 0.00994 **
hormone_treatment: Treatment vs Control 0.00276 **
hormone_treatment: Auxin vs Gibberellin 0.01372 *
hormone_treatment: Auxin1 vs Auxin2 0.90063
hormone_treatment: Gibberellin1 vs Gibberellin2 0.85152
tray 0.31082
Residuals

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

Question 4(c):

According to the analysis, there is a significant difference between control group and treatment groups with the P-value of 0.0099. Then if we look at the comparison between Auxin and Gibberellin (0.0137), these two hormone treatments have very different impacts on growth of plant as well which describes only one of two is effective. Other two contrast comparisons at the last tells that there is no significant difference between same kind of hormone treatments.

Question 5:

In this experiment, aim was not distinguishing which hormone had impact on plants but determining whether there is a significant difference between treatment groups and control group and within treatments. We could evaluate that there was a significant difference in impact of treatments and only one of two hormone preparations is effective. According to the first box plot, Gibberellin has higher mean in growth, so we could know Gibberellin has effectiveness. In conclusion if Gibberellin preparation influence on plants there would be significant change in growth. However, it is hard to determine how much it differs to other outcomes. Therefore, more analyses must be conducted such as LSD or pair-wise comparisons to figure out the scale.