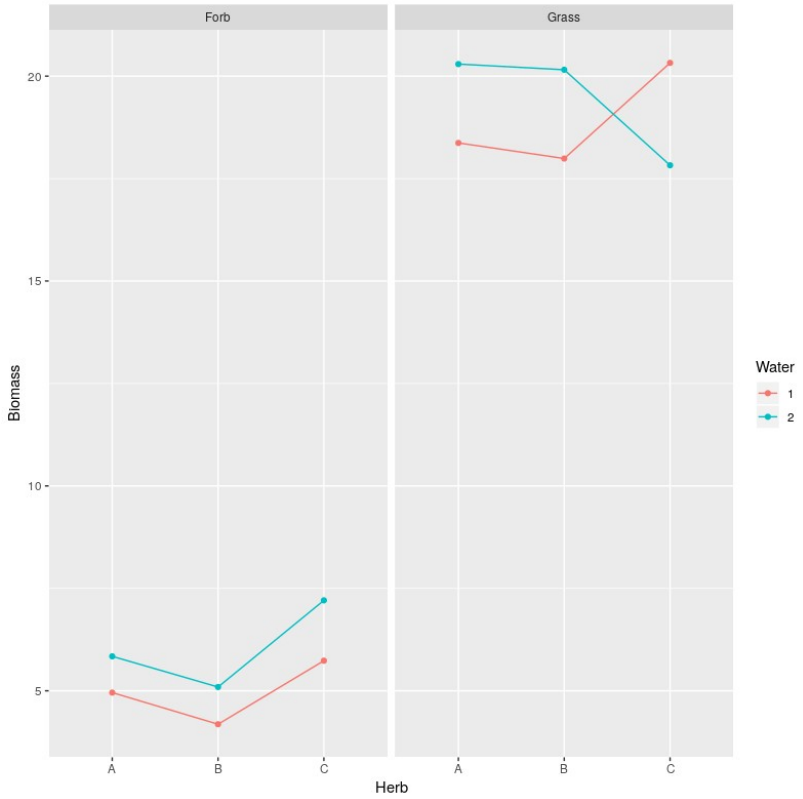
STAT 512 – Assignment 8

Vignesh M. Pagadala
Vignesh.Pagadala@ColoState.Edu

1. Create two-way interaction plots separately for each plant type (Grass/Forb). (4 pts) This can be done by first calculating treatment means (using aggregate or dplyr), then using code something like this:
library(ggplot2) qplot(x = Herb, y = Biomass, colour = Water, group = Water, data = SumStats) + geom_line() + facet_grid(. ~ Type)
ANSWER:

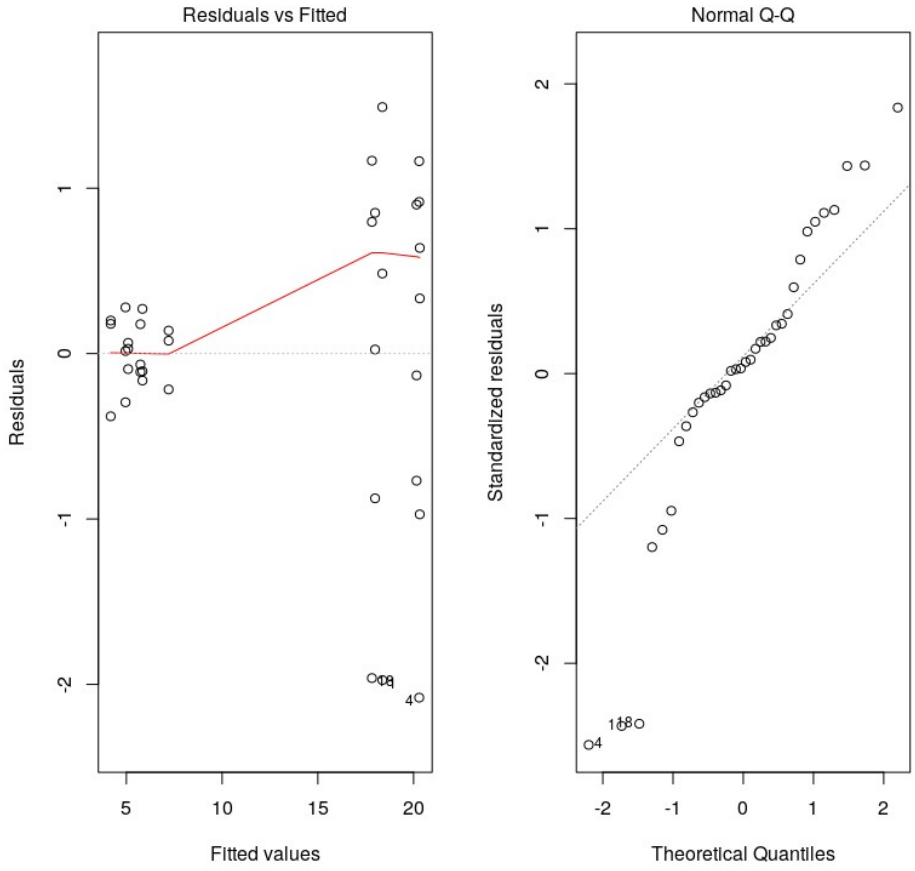


2. Fit the three-way model with all interactions. Include the Type 3 ANOVA table in your assignment. You should find evidence of a 3 way interaction.
ANSWER:

Anova Table (Type III tests)

```
Response: Biomass
                 Sum Sq Df   F value    Pr(>F)
(Intercept)     5473.8  1 5539.1323 < 2.2e-16 ***
Type            1678.7  1 1698.7721 < 2.2e-16 ***
Herb               5.1  2    2.5644  0.097872 .
Water              5.9  1    5.9711  0.022271 *
Type:Herb          5.4  2    2.7185  0.086259 .
Type:Water         0.7  1    0.7092  0.408021
Herb:Water         7.9  2    3.9973  0.031742 *
Type:Herb:Water   13.1  2    6.6440  0.005055 **
Residuals         23.7 24
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

3. Consider the diagnostics plots and discuss whether model assumptions are satisfied.
ANSWER:



The Resid vs Fitted plot does not look good, as there is unequal scatter. So the model assumptions may not be satisfied.

4. Use emmeans to calculate pairwise comparisons of Water (1 vs 2) for each level of Herb and Type. You can use code something like this:
emmeans(Model1, pairwise ~ Water|Herb*Type)
ANSWER:

```
Herb = A, Type = Forb:
 contrast estimate    SE df t.ratio p.value
 1 - 2      -0.885 0.812 24 -1.090  0.2865

Herb = B, Type = Forb:
 contrast estimate    SE df t.ratio p.value
 1 - 2      -0.908 0.812 24 -1.119  0.2741

Herb = C, Type = Forb:
 contrast estimate    SE df t.ratio p.value
 1 - 2      -1.473 0.812 24 -1.815  0.0821

Herb = A, Type = Grass:
 contrast estimate    SE df t.ratio p.value
 1 - 2      -1.924 0.812 24 -2.371  0.0261
```

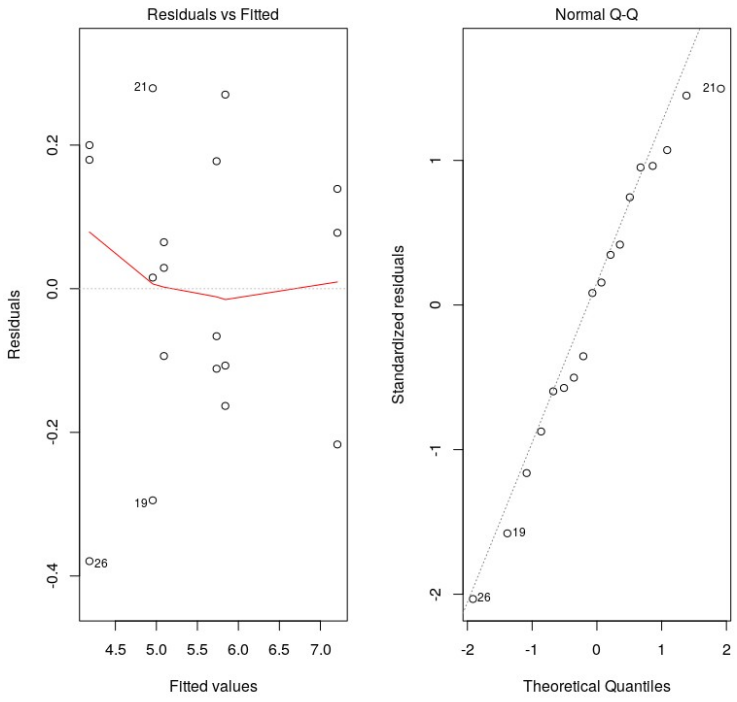
```
Herb = B, Type = Grass:
 contrast estimate    SE df t.ratio p.value
 1 - 2      -2.167 0.812 24 -2.669  0.0134

Herb = C, Type = Grass:
 contrast estimate    SE df t.ratio p.value
 1 - 2       2.499 0.812 24  3.079  0.0051
```

5. Include the Type 3 ANOVA table in your assignment.
ANSWER:

Anova Table (Type III tests)

```
Response: Biomass
             Sum Sq Df    F value    Pr(>F)
(Intercept) 544.92  1 10432.1562 < 2.2e-16 ***
Herb         10.17  2    97.3895 3.820e-08 ***
Water         5.33  1   102.1229 3.197e-07 ***
Herb:Water    0.33  2     3.1846   0.07772 .
Residuals     0.63 12
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

6. Consider the diagnostics plots and discuss whether model assumptions are satisfied.
ANSWER:



The scatter seems okay in the Resid vs Fitted plot. The Q-Q plot looks good too. These plot are better than the earlier ones.

7. Use emmeans to calculate pairwise comparisons of Water (1 vs 2) for each level of Herb.
ANSWER:

```
$emmeans
Herb = A:
 Water emmean    SE df lower.CL upper.CL
 1       4.96 0.132 12     4.67     5.24
 2       5.84 0.132 12     5.55     6.13

Herb = B:
 Water emmean    SE df lower.CL upper.CL
 1       4.18 0.132 12     3.90     4.47
 2       5.09 0.132 12     4.80     5.38

Herb = C:
 Water emmean    SE df lower.CL upper.CL
 1       5.73 0.132 12     5.45     6.02
 2       7.21 0.132 12     6.92     7.49

Confidence level used: 0.95

$contrasts
Herb = A:
 contrast estimate    SE df t.ratio p.value
 1 - 2      -0.885 0.187 12 -4.742  0.0005

Herb = B:
 contrast estimate    SE df t.ratio p.value
 1 - 2      -0.908 0.187 12 -4.868  0.0004

Herb = C:
 contrast estimate    SE df t.ratio p.value
 1 - 2      -1.473 0.187 12 -7.894  <.0001
```

8. Calculate the LSD value (ME for difference between means) corresponding to the comparisons from the previous question. Show your work for full credit. Hint: See ExpDesign2 slide 81. Recall that you can check your calculation using the SE value given in the output from the previous question.
ANSWER:

```
> qt(0.975, df = 12)*sqrt(2*(0.63/12)/3)
[1] 0.4076186
```

9. Use emmeans to calculate the comparison of Water (1 vs 2) averaging over the levels of Herb.
ANSWER:

```
> emout3$contrasts
 contrast estimate    SE df t.ratio p.value
 1 - 2       -1.09 0.108 12 -10.106 <.0001

Results are averaged over the levels of: Herb
```

10. Calculate the LSD value (ME for difference between means) corresponding to the comparisons from the previous question. Show your work for full credit. See hints from #8.
ANSWER:

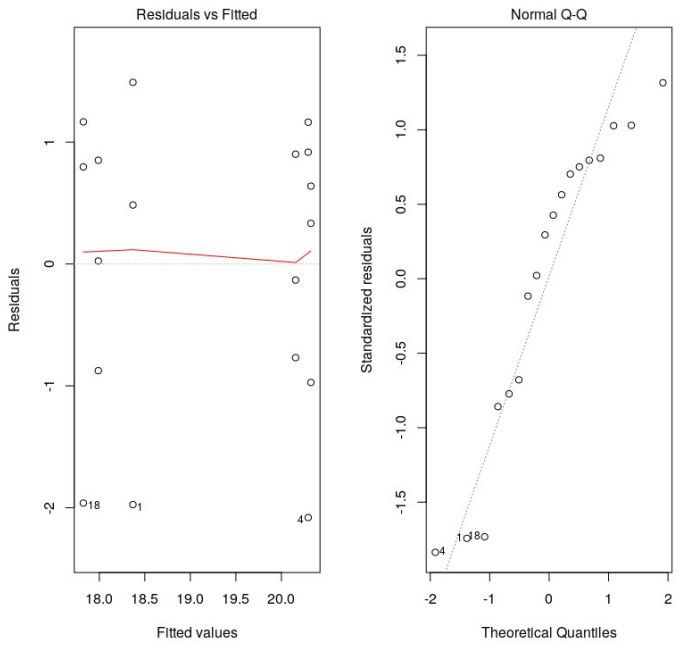
```
 qt(0.975, df = 12)*sqrt(2*(0.63/12)/9)
[1] 0.2353387
```

11. Comparing your LSD values from questions #8 (interaction comparison) and #10 (main effect comparison), which comparison has higher power? Briefly discuss.
ANSWER:
The LSD value for the Main Effect comparison is lesser, hence, the power is higher for the main effect comparison than the interaction comparison.

12. Include the Type 3 ANOVA table in your assignment.
ANSWER:

```
Anova Table (Type III tests)

Response: Biomass
             Sum Sq Df   F value    Pr(>F)
(Intercept) 6607.6  1 3433.9972 4.029e-16 ***
Herb           0.3  2    0.0694    0.9333
Water          1.3  1    0.6586    0.4329
Herb:Water    20.7  2    5.3786    0.0215 *
Residuals     23.1 12
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

13. Consider the diagnostics plots and discuss whether model assumptions are satisfied.
ANSWER:



From the plots, the scatter seems more or less equal in the Resid vs Fitted plot, and the Q-Q plot looks good as well indicating a good chance that the model assumptions are satisfied.

14. Use emmeans to calculate pairwise comparisons of Water (1 vs 2) for each level of Herb.
ANSWER:

```
Herb = A:
 contrast estimate   SE df t.ratio p.value
 1 - 2       -1.92 1.13 12 -1.699  0.1150

Herb = B:
 contrast estimate   SE df t.ratio p.value
 1 - 2       -2.17 1.13 12 -1.913  0.0799

Herb = C:
 contrast estimate   SE df t.ratio p.value
 1 - 2        2.50 1.13 12  2.207  0.0476
```

15. Would it be appropriate to calculate the comparison of Water (1 vs 2) averaging over the levels of Herb? Briefly discuss.
ANSWER:

No. Because, there is significant interaction between Water and Herb.

16. The comparisons from #7 and #14 were also considered in #4. Considering the corresponding results from the combined analysis (#4) versus the split analysis (#7, #14), briefly discuss what is the same and what is different. Specifically, consider the estimated difference, corresponding SE and df. (4 pts)
ANSWER:
From the results, we can see that the estimated difference is the same, but the SE and DF are quite different.

17. Give (at least) one reason why we might prefer to split the analysis by Type (running separate 2way ANOVAs for Grass and Forb). Your answer should be based on the output.
ANSWER:
From looking at the Resid vs Fitted and Q-Q plot for the analysis where we do not split by Type, the model assumptions are clearly not satisfied. But when we do the splitting, the assumptions are satisfied.

18. Give (at least) one weakness of splitting the analysis "by type" as compared to the full 3way ANOVA model.
ANSWER:
By splitting the analysis by Type, we are unable to observe and analyze the behavior of Forb against Grass, which could be quite useful.