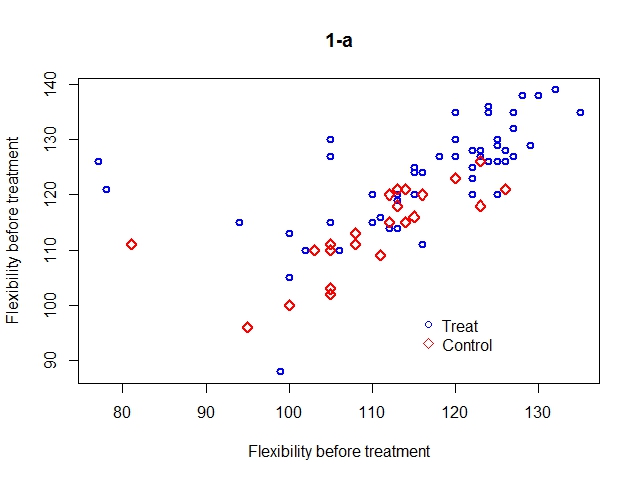
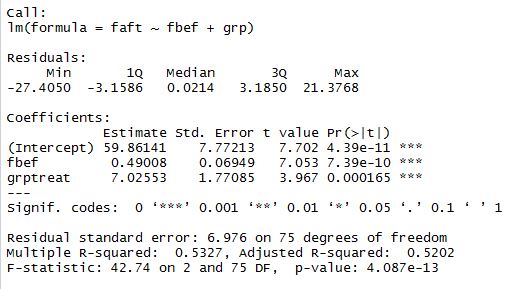
**HW 5**

1. According to the scatter plot below, we can see that for the control group, the relationship between before and after is roughly linear, while the relationship between before and after for the treatment group shows a curvilinear pattern.



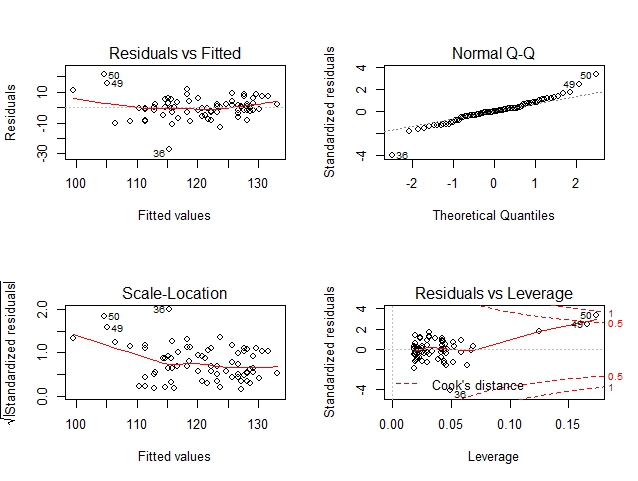
1. According to the summary below, there’s a significant treatment effect, in that the p-value for this regression coefficient is less than 0.05 (p-value = 0.000165). The flexibility before as well as the intercept are also significant at α= 0.05. Therefore, flexibility before and the treatment are both significant predictors for flexibility after.



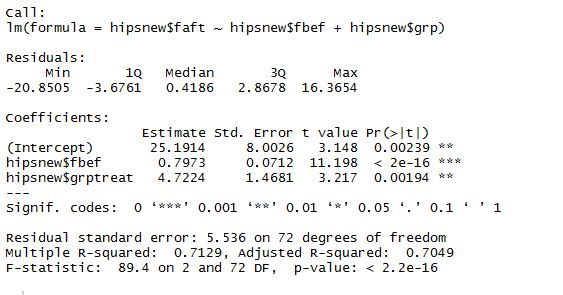
1. According to the output below, No. 36 and No. 50 are considered outliers.



According to the diagnostic plot below, we can tell that No. 36, 49 and 50 are pretty influential points. Refer to the data, we find that both No. 49 and 50 have before measurements less than 90. Therefore, it’s reasonable to include only before measurements under 90.



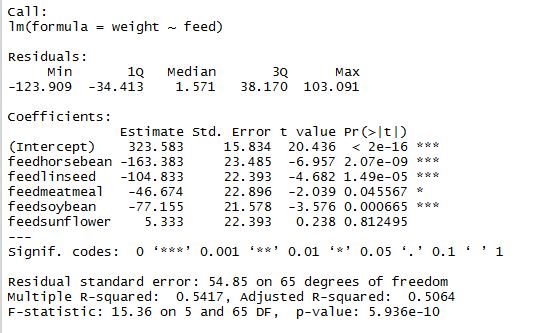
1. According to the model summary below, there’s a significant treatment effect, because the p-value for the corresponding regression coefficient is less than 0.05 (p-value = 0.00194).



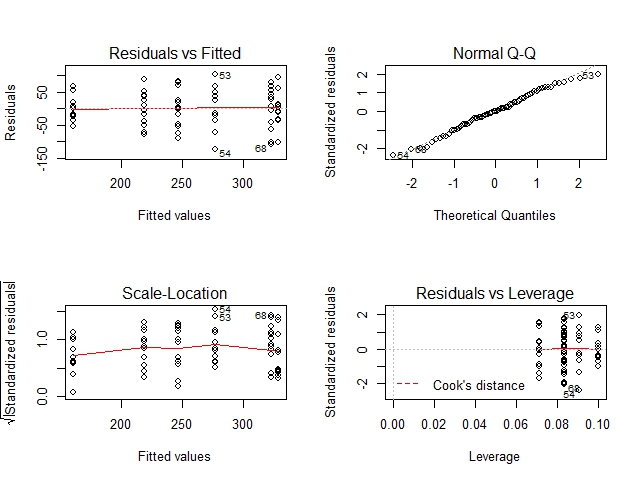
1. According to the model summary above, the estimated size of the treatment effect is 4.72. The 95% CI for the effect is (1.796, 7.649).



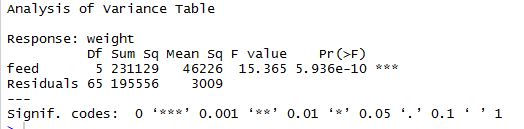
1. Summary



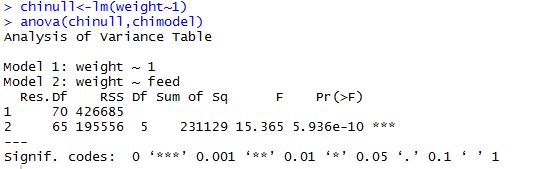
1. According to the diagnostic plots below, we can conclude that there’s no big problem with linearity and constant variance (Residual vs Fitted has a roughly flat line, and points spread equally vertically; Scale-Location has a roughly flat line). The normality assumption holds because the points are on a straight line in the Q-Q plot. Residual-Leverage plot indicates no problems with outliers or too influential observations.



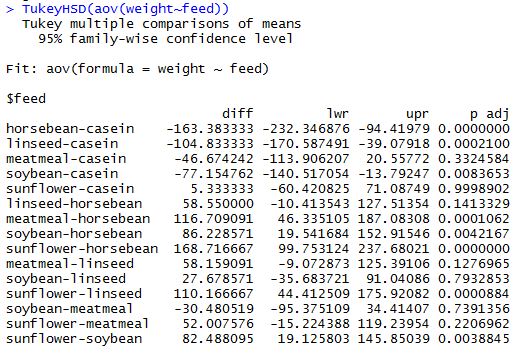
1. ANOVA table



1. Group comparison: according to the output below, there is a significant difference between the mean weights of the groups, in that the F-statistic has a p-value smaller than 0.05.



1. Tukey simultaneous 95% CI:



1. According to the output above, the pairs that have significantly different means are **:**

horsebean-casein

linseed-casein

soybean-casein

meatmeal-horseben

soybean-horsebean

sunflower-horsebean

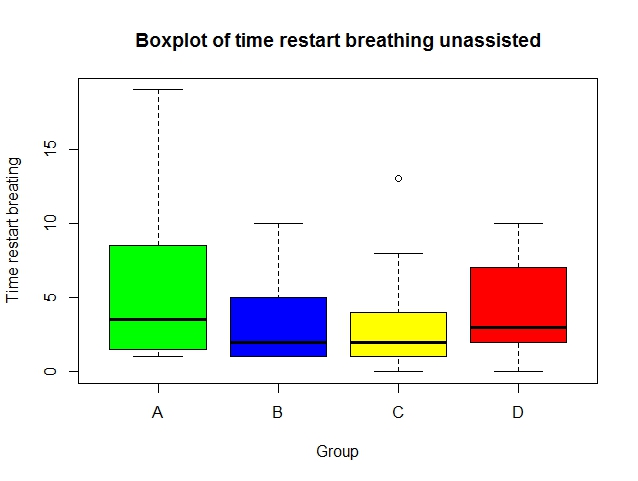
sunflower-linseed

sunflower-soybean

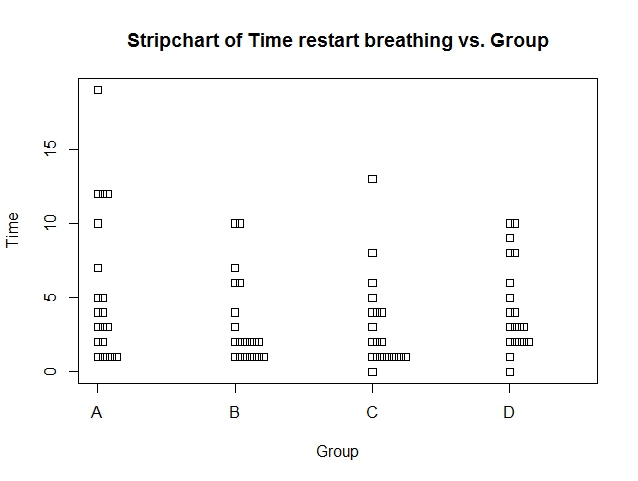
1. According to the boxplot below, there is more variability in Group A than the other 3, in that it has the longest box.

The four boxplots are not very different in term of their positions vertically (no boxplot is obviously higher or lower than the other ones), indicating that there might not be very large difference among groups.

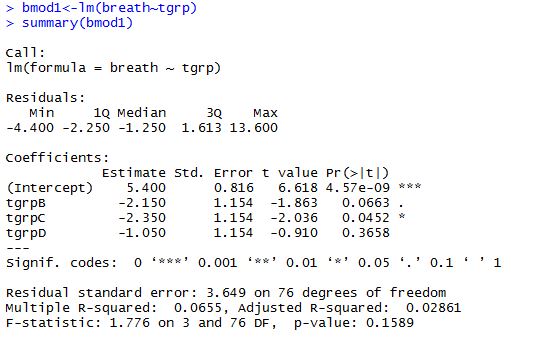
Medians of the 4 groups are not drastically different, but the distributions are pretty different. All 4 groups have more variability in the most positive quartile than in the least positive quartile, and the difference between the two is the biggest for Group A.



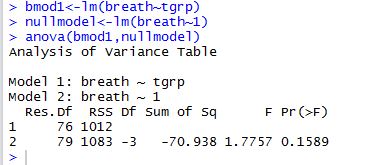
1. The stripchart:



1. One-factor model with default coding of effect:



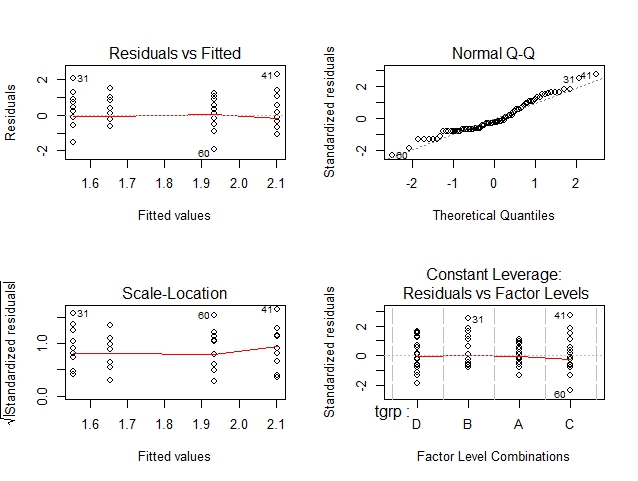
According to the ANOVA output below, there is no difference between the four groups, because p-value = .16>.05.



1. When the Box-Cox transformation was tried, an error message was returned, because this method can only be applied when responses are positive. However, two responses in the data are actually equal to 0.



1. According to the plots below, No. 31, 41 and 60 are problematic, in that they are marked in all 4 diagnostic plots. Other than theses 3 observation, generally the diagnostics look OK in term of variance constance, outliers, normality (except for the lower end), and influence.



According to the ANOVA output below, we can’t reject the null hypothesis that there’s no difference among groups, because the p-value is larger than 0.05 (equals 0.16).

