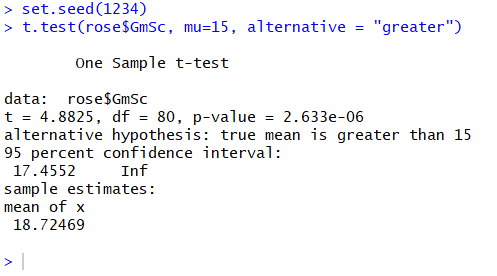
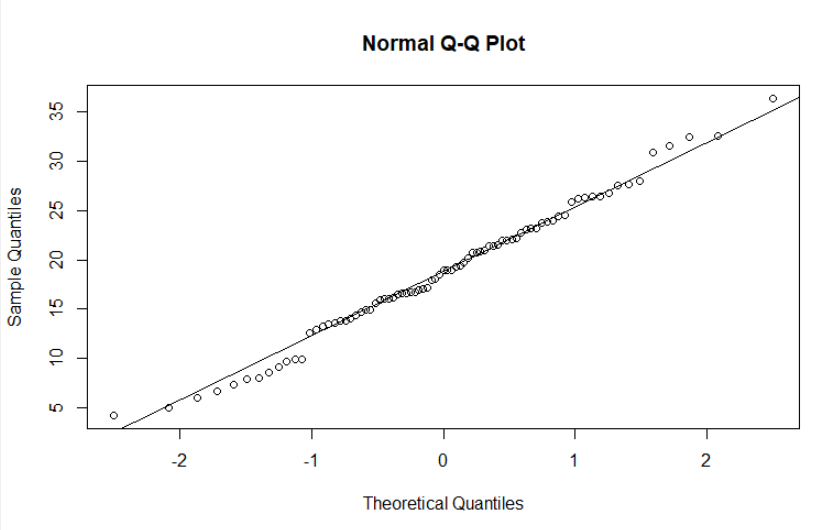
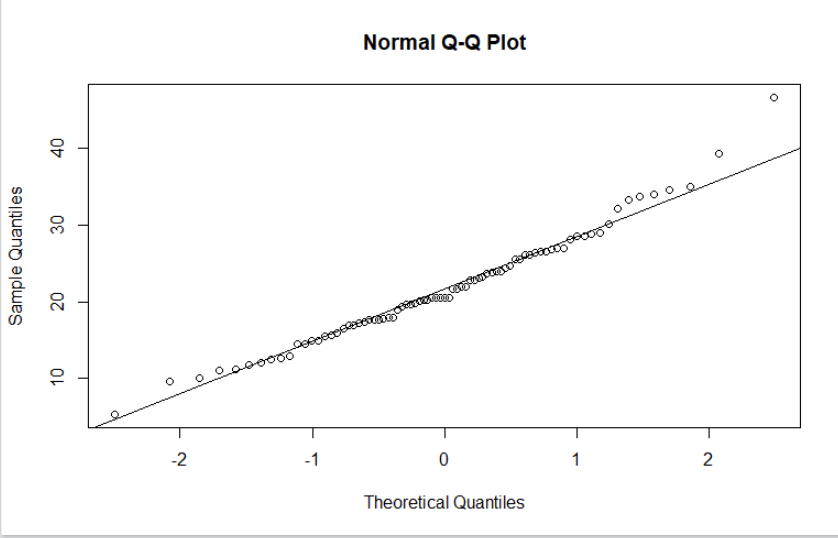
1)

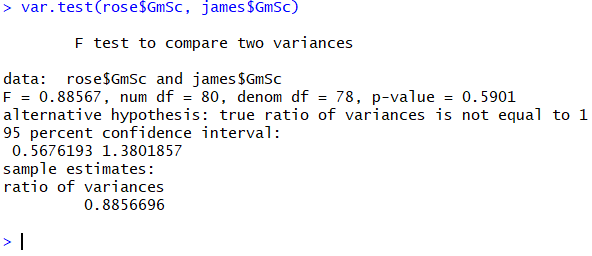


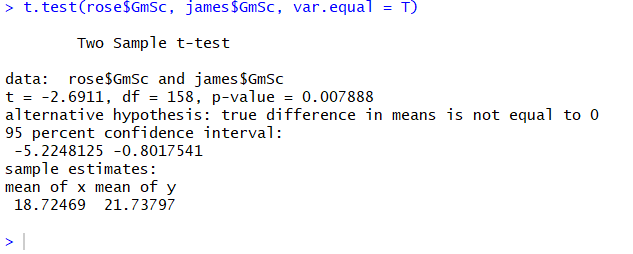
In this one sample one-sided t-test, we are comparing the mean of Derrick Rose’s game scores to the population mean of 15. The null hypothesis is that they are equal, while the alternative hypothesis is that the mean of Rose’s game scores is significantly greater than 15. The results of the t-test show that the mean of Rose’s game scores is 18.725, the p-value is significantly less than the alpha of .05, and the confidence intervals show that 95 times out of 100, the mean will be greater than 17.455, which is obviously greater than the null hypothesis value of 15. We can then reject the null hypothesis, and conclude that the mean of Rose’s game scores is significantly greater than the population mean of 15.

2)





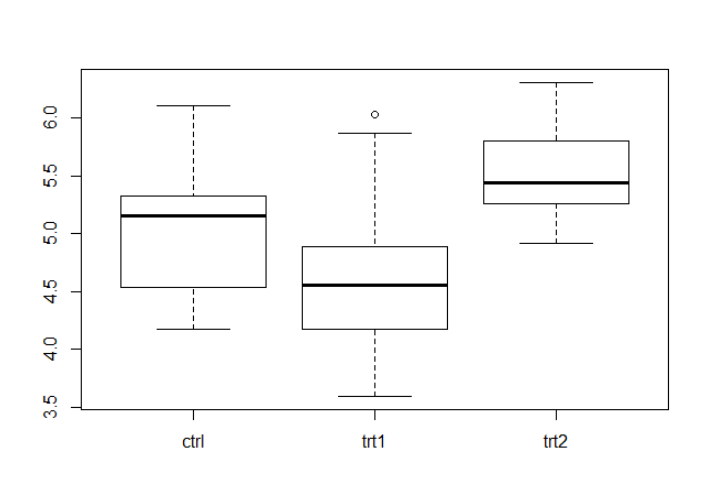




Before performing the two-sample t-test, we must confirm the assumptions of normality and homogeneity of variances. The first four lines of code are creating QQplots for both Rose and James’s game scores. Being that both fit closely to a straight diagonal line, we can assume that both samples are normal. For testing equal variances, we can use an f-test which is done through the built-in method var.test(). The results of this show that the p-value is .5901 which is significantly higher than the alpha of .05. Since we want the variances to be equal, this higher number confirms there is not a significant difference in variance between the two samples and we can accept the null hypothesis that the true ratio of variances is equal to 1. Now we can run the t-test. This two-sample t-test is comparing whether there is a significant statistical difference between Derrick Rose’s game scores and LeBron James’s game scores. The null hypothesis is that they are the same, while the alternative hypothesis is that they are not. The results show that the mean of Rose’s game scores is 18.725 while the mean of James’s game scores is 21.738. Additionally, the p-value of .00789 is less than the alpha of .05, and the 95% confidence interval shows that the difference is likely between -5.22 and -0.80 (the null hypothesis of zero is not in this range). With the true difference in means of -3.02 and the p-value below .05, we can reject the null and conclude that there is a significant statistical difference between Rose and James’s game scores.

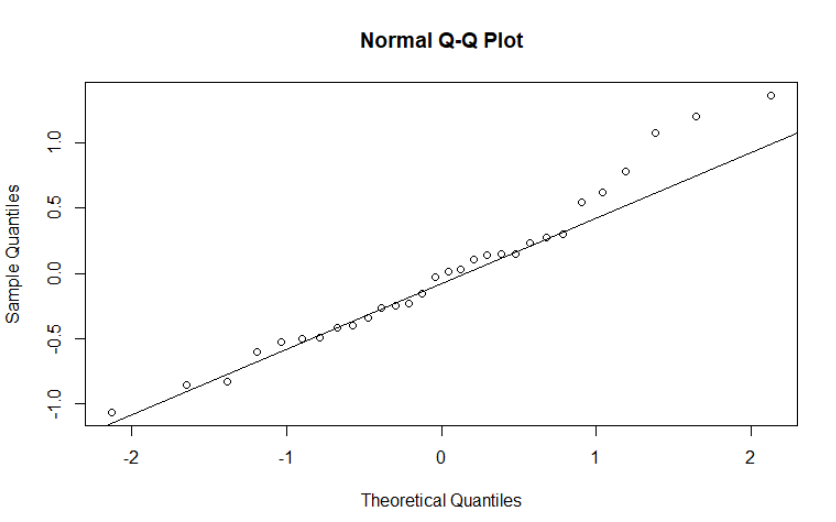
3)

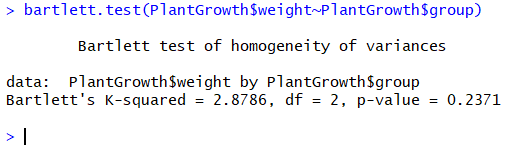


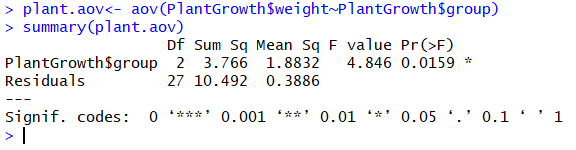


This line of code runs a boxplot with weight on the y-axis, and the groups on the x-axis, and shows the summary statistics for all three groups. It looks like the control group is generally heavier than the treatment1 group, and lighter than the treatment2 group. It also looks like there is slightly more variance on the control group than the treatment groups, with the treatment2 group having the least variance. There is also an outlier in treatment1 as signified by the blank circle above the upper maximum.

4)

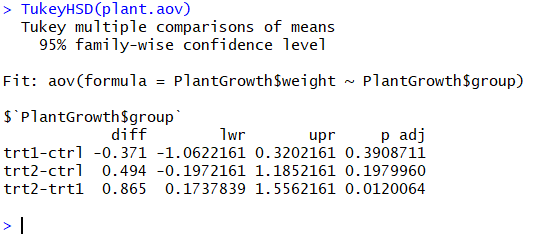






The first line of code creates the ANOVA model under the object plant.aov. Before we run the model, we must check that the normality and equal variance assumptions are met. The next two lines plot the residuals of the model on a QQplot and fit it with a diagonal. Since the points fit nicely on the diagonal, we can assume that the model is normal. To test homogeneity of variances, we cannot use the var.test method because there are more than two groups in the model. Instead, we use the Bartlett test, which is called in the next line of code to test the variances in our model. The p-value is .2371, which is significantly higher than the alpha of .05. This indicates that we can accept the null, and assume that there is not a statistically significant difference between the ratio of variances in our model. Finally, we can run the summary on the model, which is what the next line of code does. The Weight column of the data frame PlantGrowth is the dependent variable, and the Group column of the data frame is the independent variable. Based on this model, it looks like there is a significant statistical difference between the mean weights of the control and treatment groups (p-value of .0159 < .05 alpha), however it is not clear where the difference is coming from.

5)



This line of code runs Tukey’s HSD test on the plant.aov ANOVA model, allowing us to see the difference between each of the groups to determine what is driving the statistical difference between them indicated by the summary in question 4. The p-value for the difference between treatment1 and control is .391, while the p-value for the difference between treatment2 and control is .197. This indicates that there is not a significant statistical difference between either of the two treatment groups and the control group, as these are both above the alpha of .05. The statistically significant difference between the groups shown in question 4 is derived from the difference between the treatment1 and treatment2 as shown by the p-value for this difference of .012, which is less than the alpha and therefore statistically significant. This makes sense when comparing to the summary of the ANOVA model, as even though there was a significant difference based on an alpha of .05, the p-value of .016 was not a hugely significant indicator. If there was a significant difference between all three (with Tukey’s indicating a p-value less than .05 for all three comparisons), then the p-value from the summary of the model would have been even smaller.