1. Use the dataset *Garlic* to perform an ANOVA. This dataset has information on the weight of garlic bulbs (**bulbwt**) for 32 different garlic plants. Each garlic plant was initially treated with 1 of 4 types of **fertilizer**. The purpose of the experiment is to determine whether or not there is any difference in the resulting bulb weight when a garlic plant is treated with a different fertilizer.
   a. Test the hypothesis that the means of bulb weight are equal regardless of the fertilizer used.
   - Verify that the assumption of normality in each sample you are testing. Are you comfortable with this assumption?

YES, The QQ plots look fine.

   - Verify that the groups satisfy the assumption of equal variance.

YES, using Levene's test.

   - State your conclusion to the hypothesis test (alpha = 0.05)

At least one fertilizer type is different.

   b. If you were to test if each type of fertilizer were different from each other fertilizer (pairwise), how many hypothesis tests would you be running?

4Choose2 = 6.

   c. If the probability of incorrectly rejecting a true null hypothesis is 0.05, and each hypothesis test is considered independent, what is the probability that you incorrectly reject at least one true hypothesis? How do we solve this problem when we're performing post-hoc analysis in ANOVA?

1-.95^6 = .26

We solve it by controlling for experiment wise error using some kind of p-value or alpha adjustment like Tukey's method or Dunnett's method.

   d. Which fertilizers are statistically different from each other and which fertilizers do not appear to produce different bulb weights?

Fertilizer 4 is different from Fertilizers 1 and 3, but all other pairs of fertilizers do not show a significant difference.

2. The *Bottle* dataset contains observations from a factory that is producing plastic water bottles along 3 different assembly lines. The number of

units produced by each assembly **line** are given for a number of days. The manufacturer wants to know if the assembly lines are producing the same number of bottles. He suspects that 2 of his lines are significantly better than the third, but he has had trouble demonstrating this due to the variability in production. Confirm or deny his hypothesis using the data provided. Be sure to verify your assumptions before proceeding with the analysis.

Our assumptions are verified looking at the QQplot/histogram of residuals and Levene's test for HOV. Lines 1 and 2 have larger average units produced and this difference proves to be statistically significant using Tukey's test.

3. The Trials dataset contains information from a clinical trial for blood pressure medicine. There are two different drugs being compared to a placebo, 1 of which is already FDA approved. The study aims to confirm that the new drug (**treatment** = New Drug) is effective at lowering blood pressure when compared to the placebo. Perform an ANOVA and post-hoc analysis which compares the reduction in blood-pressure of the new drug and the existing drug (**treatment** = Approved Drug) to the placebo. Do both drugs outperform the placebo in blood-pressure reduction? Be sure to verify your assumptions before proceeding with the analysis.

In verifying the assumptions for this data, we see that it fails Levene's test for homogeneity of variances. Therefore, the ANOVA output is invalid and we'd have to use a different procedure. In this case, we'd want to use Welch's ANOVA (Welch's p-value is less than 2.2×10^{-16}).

– for post-hoc test, you can use Games-Howell (which does NOT assume equal variances nor equal sample sizes…although it is a bit conservative for comparing to a control group!). You can find this algorithm in the package userfriendlyscience). They are both differenent than the placebo.

ALSO!! You could do 2 different t-tests (with unequal variance….one comparing Placebo to New Drug and one comparing Placebo to Approved Drug). Then use Bonferroni to adjust p-values (basically multiply them by 2). Both are significantly different than the placebo!