**Problem**

Vital capacity is a measure of the amount of air that someone can exhale after taking a deep breath, Data was collected on brass players and a control group.

Brass Player Control Group

4.7 4.2

4.6 4.7

4.3 5.1

4.5 4.7

5.5 5.0

4.9 -

5.3 -

1. Put the data into a "long format" data frame. That is one column labeled "Vital" for vital measure and second character or factor column labeled "Group" with the group labels of "Brass" or "Control”.
2. Conduct a test using "t.test" using the S3 method for class 'formula', i.e. t.test(Vital~Group, data, ...) to determine whether the population mean for brass is larger than that for control. Do NOT use t.test(Brass,Control,data) which is the default S3 method.
3. Provide the equivalent 95% confidence interval for the difference of two population means. Meaning of equivalent: If you conducted a one-sided hypothesis test in step 2, then you would produce the same corresponding one-sided confidence interval here. If you conducted a two-sided hypothesis test in step 2, then you would produce a two-sided confidence interval here.
4. A researcher claims that in theory the "spread/variance" in the two populations is the same. Repeat step 2 (and 3) utilizing this assumption with the argument "var.equal" within the "t.test" function. Use the same information in step 3 regarding the confidence interval and hypothesis test direction

**Solution**

**Part (1) – Long Format**

Step one for this part was to copy the provided data into Excel and manually put it into a long format. Then, this data was used to create an R dataframe with the following code:



The results of which are shown here:



**Part (2) – t.test**

A t-test was run on the above dataframe at the 0.05 level of significance. Additionally, the argument *alternative="greater"* was included due to the question specifically asking whether the mean for the brass group was larger than that of the control group. Specifically, the null hypothesis for this test will be that the mean of the brass group will be less than or equal to the mean of the control group, and the alternative hypothesis will be that the mean of the brass group is higher than the mean of the control group. Both the code and the results are shown below:



In the results for this test we see that both the resulting p-value is large, and that the confidence interval contains 0. Therefore, we do not reject the null hypothesis and conclude that it is likely that the mean for brass group is not meaningfully larger than for the control group.

**Part (3) – 95% CI**

The 95% confidence interval, as calculated above, is -0.324 to positive infinity.

**Part (4) – Equal Variances**

Adding the specified assumption of equal variances can be done by adding the argument *var.equal=TRUE* to the previous test. The null and alternative hypotheses remain the same from part (2). The code and the results are shown below:



This assumption does not change the conclusion from the previous two sections. We will still fail to reject the null hypothesis and conclude that it is likely that the mean for brass group is not meaningfully larger than for the control group. Additionally, the 95% confidence interval, as calculated above, is now -0.339 to positive infinity.

**Summary**

In this assignment we worked with data in a long format. We then performed a one-sided t-test to determine whether the mean of the brass group was meaningfully larger than that of the control group. We then calculated a 95% confidence interval for the data provided. Finally, we assumed that the populations had equal variances and performed the same tests as before with this new assumption.