You must attach your Python script output as an HTML file and respond to the questions below.

In this discussion, you will apply the statistical concepts and techniques covered in this week's reading about one-way analysis of variance (ANOVA). An investment analyst is evaluating the 10-year mean return on investment for industry-specific exchange-traded funds (ETFs) for three sectors: financial, energy, and technology. The analyst obtains a random sample of 30 ETFs for each sector and calculates the 10-year return of each ETF. The analyst has provided you with this data set. Run Step 1 in the Python script to upload the data file.

Using the sample data, perform one-way analysis of variance (ANOVA). Evaluate whether the average return of *at least one* of the industry-specific ETFs is significantly different. Use a 5% level of significance.

In your initial post, address the following items:

1. Define the null and alternative hypothesis in mathematical terms and in words.
2. Report the level of significance.
3. Include the test statistic and the P-value. See Step 2 in the Python script.
4. Provide your conclusion and interpretation of the test. Should the null hypothesis be rejected? Why or why not?
5. Does a side-by-side boxplot of the 10-year returns of ETFs from the three sectors confirm your conclusion of the hypothesis test? Why or why not? See Step 3 in the Python script.

In your follow-up posts to other students, review your peers' results and provide some analysis and interpretation:

1. What does a post-hoc test (like Tukey's HSD test) contribute after one-way ANOVA is performed?
2. Comment on your peers' results and compare them with your own.

Remember to attach your Python output and respond to all questions in your initial and follow-up posts. Be sure to clearly communicate your ideas using appropriate terminology. Finally, be sure to review the [Discussion Rubric](https://learn.snhu.edu/d2l/common/dialogs/quickLink/quickLink.d2l?ou=1230325&type=content&rcode=snhu-702316) to understand how you will be graded on this assignment.

The null hypothesis, signified by H0, states that there is no significant difference in the mean return of the three sectors. The alternative hypothesis, denoted by Ha, states that at least one of the three sectors has a significantly different mean return.

H0: μ1 = μ2 = μ3

Ha: At least one μi is different from the others, where i=1,2,3, representing financial, energy, and technology sectors, respectively.

The level of significance is 5% or 0.05.

The test statistic is 55.07, and the P-value is 0.0.

Based on the obtained P-value of 0.0, which is less than the level of significance of 0.05, we can reject the null hypothesis. Therefore, we can conclude that there is sufficient evidence to suggest that at least one of the three sectors has a significantly different mean return.

The side-by-side boxplot of the 10-year returns of ETFs from the three sectors confirms the hypothesis test's conclusion. The boxplots show that the technology sector has a higher median return compared to the financial and energy sectors, which aligns with the ANOVA result. The boxplots also show that the financial and energy sectors have more significant variability in returns than the technology sector, which is consistent with the ANOVA's hypothesis of equal variances.

Response 1:

A post-hoc test, such as Tukey's HSD test, is used after performing one-way ANOVA to determine which specific groups have significant differences in means. ANOVA only determines whether there is a significant difference between at least two groups, but it doesn't indicate which specific groups have differences. Post-hoc tests provide a way to compare all possible pairs of groups and determine which pairs have statistically significant differences in means.

It's great to see that you have correctly stated the null and alternative hypotheses and the level of significance used in the hypothesis test. Their conclusion that the null hypothesis should be rejected because the technology sector has significantly greater ETF returns than both the financial and energy sectors is also correct.

Response 2:   
 A posthoc test like Tukey's HSD (Honestly Significant Difference) test determines which groups have significantly different means after a significant ANOVA result. Tukey's HSD test can be used to compare all possible pairs of groups and provide a confidence interval for the difference in means. This can help identify which specific groups have significantly different standards and provide more detailed information about the differences among the groups.

Your analysis and conclusion are consistent with mine. They correctly identified the null and alternative hypotheses, reported the significance level, and correctly interpreted the test results. Their inclusion of a side-by-side boxplot to visualize the data is also a helpful addition. Our analysis provides a thorough and accurate interpretation of the hypothesis test results.