TASK:

This assignment aims to enhance proficiency in statistical analysis and hypothesis testing in Python utilizing real-world datasets to conduct hypothesis tests for paired data, multiple population means, and variance comparison, cultivating a deeper understanding of the statistical method's limitation and reliability through exploring assumptions, potential consequences of violating them, and their remedies.

Task 1: Paired Data Analysis

1. Find a suitable dataset with paired data from a reliable source.
2. Write a Jupyter notebook to import the dataset and perform a hypothesis test for paired data.
3. Explain the problem, data, research question, and the appropriate statistical test that you would use to address the research question.
4. Discuss the assumptions underlying the hypothesis test, potential consequences of violating these assumptions, and possible remedies.
5. Propose a Python tool to assist with the interpretation of the results of the statistical test and visualizations.
6. Record a 3- to 4-minute video explaining the analysis, results, and interpretation of the hypothesis test.

Task 2: Multiple Population Means Analysis

1. Find a suitable dataset with more than two populations means from a reliable source.
2. Continuing in the same Jupyter notebook, write a code to import the dataset and perform a hypothesis test for more than two population means.
3. Explain the problem, data, research question, and the appropriate statistical test that you would use to address the research question.
4. Discuss the assumptions underlying the hypothesis test, potential consequences of violating these assumptions, and possible remedies.
5. Propose a Python tool to assist with the interpretation of the results of the statistical test and visualizations.
6. Record a 3- to 4-minute video explaining the analysis, results, and interpretation of the hypothesis test.

Task 3: Variance Comparison Analysis

1. Find a suitable dataset with two or more populations with variances to compare from a reliable source.
2. Continuing in the same Jupyter notebook, write a code to import the dataset and perform a hypothesis test for variance comparison.
3. Explain the problem, data, research question, and the appropriate statistical test that you would use to address the research question.
4. Discuss the assumptions underlying the hypothesis test, potential consequences of violating these assumptions, and possible remedies.
5. Propose a Python tool to assist with the interpretation of the results of the statistical test and visualizations.

6. Record a 3- to 4-minute video explaining the analysis, results, and interpretation of the hypothesis test.