QP Code: D142143	**Total Pages:** 2	**Name:**
		Register No.

SECOND SEMESTER (CUFYUGP) DEGREE EXAMINATION, APRIL 2026

APPLIED PHYSICS / PHYSICS

PHY2VN102/APH2VN102-Data Analysis in Physics Using Python

2024 Admission onwards

Maximum Time :2 Hours	**Maximum Marks :70**

Section A

All Questions can be answered. Each Question carries 3 marks (Ceiling : 24 Marks)

1	Explain what a Python virtual environment is and outline why it's particularly beneficial for data analysis projects.
2	How do the structure and usage of a pandas Series differ from those of a pandas DataFrame?
3	Mention two methods for dealing with missing values in a pandas DataFrame and provide short code snippets demonstrating these methods.
4	In Matplotlib, what is the difference between an axis object and a figure object?
5	What are quartiles in a dataset? Discuss their importance in describing data distribution.
6	Compare Jupyter Notebook and Google Colab in terms of features and typical usage scenarios.
7	Name some common file formats for storing data in pandas, and briefly explain how you can load them into a DataFrame.
8	Show how to filter rows in a DataFrame using a specific condition. Provide a short code example.
9	Distinguish between accuracy and precision in measurements.
10	Define the median and describe a situation in which it is a more suitable measure of central tendency than the mean.

Section B

All Questions can be answered. Each Question carries 6 marks (Ceiling : 36 Marks)

11	Demonstrate how to concatenate two DataFrames both vertically and horizontally using pandas. Provide short code examples.
12	Explain methods for detecting and removing duplicate rows in a DataFrame. Include a brief Python example.

13	Write a Python program to read data from a CSV file, sort its rows by a given column, and save the sorted output to a new file.
14	Describe the concept of the coefficient of variation (CV) and illustrate its usage with a simple example.
15	How can you save a pandas DataFrame to both CSV and Excel files? Provide the essential Python code.
16	What is the difference between a bar chart and a count plot? When is each type of plot more appropriate?
17	In what ways does the standard deviation help us understand the spread of a dataset? Provide a brief example.
18	State the central limit theorem (CLT) and discuss why it is so important in statistical analysis.
Section C	
Answer any ONE. Each Question carries 10 marks (1x10=10 Marks)	
19	Discuss the importance of understanding both random and systematic errors in experimental data. How do these uncertainties influence data analysis, and what approaches can minimize their impact?
20	Explain the Poisson distribution and show how it arises as a limiting case of the Binomial distribution. Include an example to illustrate the relationship.