CALICUT UNIVERSITY – FOUR-YEAR UNDER GRADUATE PROGRAMME (CU-FYUGP)

BSc PHYSICS HONOURS

Programme	**B.Sc. Physics Honours**				
Course Title	**DATA ANALYSIS IN PHYSICS USING PYTHON**				
Type of Course	**Vocational Minor (SET II: DATA ANALYSIS IN PHYSICS)**				
Semester	**II**				
Academic Level	**100 – 199**				
Course Details	Credit	Lecture per week	Tutorial per week	Practical per week	Total Hours
	4	3	-	2	75
Pre-requisites	PHY1VN102- Python Basics				
Course Summary	This paper continues from the previous paper for data analysis. More data analysis tools are introduced to be used in machine learning, as well as in physical data analysis. In addition, essential statistics required for data analysis is also introduced.				

Course Outcomes (CO):

CO	CO Statement	Cognitive Level*	Knowledge Category#	Evaluation Tools used
CO1	Become familiar with data format & programs used in data analysis	U	F	Practical Assignment / Observation of Practical Skills
CO2	Understand & apply Pandas module for data analysis	U, Ap	P	Instructor-created exams, Practical Assignment / Observation of Practical Skills

CO3	Understand & apply Seaborn module for data visualization	U, Ap	P	Instructor-created exams, Practical Assignment / Observation of Practical Skills
CO4	Understand the significance of statistical analyses as well as error analysis in physical measurements.	U	F	Instructor-created exams
CO5	Understand the significance of few distributions commonly found in physical measurements.	U	F	Instructor-created exams/ Home Assignments
CO6	Apply statistical methods to physical measurements	U, E	P	Home Assignments

* - Remember (R), Understand (U), Apply (Ap), Analyse (An), Evaluate (E), Create (C)

\# - Factual Knowledge(F) Conceptual Knowledge (C) Procedural Knowledge (P) Metacognitive Knowledge (M)

Detailed Syllabus:

Module	Unit	Content	Hrs (45 +30)	Marks (70)
I		**Data file formats**	8	10
	1	Introducing different data file formats: csv, xls, tab, dat formats.	2	
	2	Jupyter Notebooks using Anaconda and Google Colab: introduction.	2	
	3	Familiarization with Google Colab	1	
	4	Familiarization with Anaconda	2	
	5	Reading data files in Jupyter Notebooks.	1	
Basic overview to be given about data formats and software used.				
II		**Using Pandas for Data Analysis**	12	20
	6	Data Analysis Using Pandas: Series and dataframe, creating data frame from an excel spreadsheet - creating dataframe from .csv files.	3	

	7	Creating data frame from a python dictionary - creating dataframe from python list of tuples - viewing data frame using loc() and iloc().	3	
	8	Operations on data frames series object - creating series from a dataframe - creating dataframe from series - creating series from numpy array.	2	
	9	Converting series into numpy array - creating series from a dictionary - accessing elements of a series.	2	
	10	Joining data frames - how to join when there is no common column - concatenation of tables - where() method - groupby() method - aggregate functions on data frames.	2	
Chapters 12,13 (SQL & Regular expressions not required) of Book 1				
	Data Visualization using Seaborn		**10**	
	10	Loading datasets in Seaborn, Distribution plot	1	
	11	Count plot, box plot, scatter plot, joint plot.	2	
III	12	Line Plot, displaying scatter plot with regression line	2	**20**
	13	Creating subplots	1	
	14	Heat map - cat plot	2	
	15	Violin plot - pair plot.	2	
Chapter 15 of Book 1				
	Basic Statistics & Error Analysis		**15**	
	16	Preliminaries of Error Analysis: errors as uncertainties, inevitability of uncertainty,	2	
	17	Importance of knowing the uncertainties.	2	
	18	Statistical analysis of random uncertainties: random and systematic errors, the mean and standard deviation.	2	
IV	19	Standard deviation as the uncertainty in a single measurement, the standard deviation of the mean, systematic errors.	2	**20**
	20	The Normal Distribution: Histograms and distributions, limiting distributions, the normal distribution.	3	
	21	The Standard deviation as 68% confidence limit, justification of the mean as best estimate.	2	
	22	The Poisson Distribution: Definition of the Poisson Distribution, Properties of the Poisson Distribution.	2	

Sections 1.1-1.3; 4.1-4.6; 5.1-5.5; and 11.1-11.3 of Book 2			
	PRACTICALS	**30**	
V	Conduct any 6 experiments from the given list and 1 additional experiment, decided by the teacher-in-charge, related to the content of the course. The 7th experiment may also be selected from the given list.		
	1. Familiarising Jupyter notebook using Colab/Anaconda and basic coding 2. Read data from different output format (csv, xls, tab, dat, txt) and save it in a specific format (csv, dat) 3. Heat map, Box plot, scatter plot 4. Violin plot, Pair plot 5. Basic statistics - plots including error bars 6. Grouping example using colab 7. Create series from a dataframe and dataframe from series using numpy array.		

Books and References:

1. Machine Learning in Data Science using Python, Dr. R. Nageswara Rao, Dreamtech press, 2022 (Book 1)
2. An Introduction to Error Analysis, John R. Taylor 2nd edition, University Science Books, 1996 (Book 2)

Mapping of COs with PSOs and POs :

	PSO1	PSO2	PSO3	PSO4	PSO5	PSO6	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO 1	0	0	0	0	1	0	1	0	0	2	1	0	0
CO 2	0	0	0	0	2	0	1	0	0	2	1	0	0
CO 3	0	0	0	0	2	0	1	0	0	2	1	0	0
CO 4	0	1	2	0	1	1	1	0	0	1	2	0	0
CO 5	0	1	1	0	1	1	1	0	0	1	2	0	0
CO 6	0	1	1	0	1	1	1	0	0	1	2	0	0

Correlation Levels:

Level	Correlation
0	Nil
1	Slightly / Low
2	Moderate / Medium
3	Substantial / High

Assessment Rubrics:
- Quiz / Assignment/ Quiz/ Discussion / Seminar
- Midterm Exam
- Programming Assignments (20%)
- Final Exam (70%)

Mapping of COs to Assessment Rubrics :

	Internal Exam	Assignment	Project Evaluation	End Semester Examinations
CO 1	✓			✓
CO 2	✓			✓
CO 3	✓			✓
CO 4		✓		✓
CO 5		✓		✓
CO 6			✓	