COMPUTATIONAL STATISTICS LAB		Semester	V
Course Code	**BCBL504**	CIE Marks	50
Teaching Hours/Week (L:T:P: S)	0:0:2:0	SEE Marks	50
Credits	01	Exam Hours	100
Examination type (SEE)	Practical		

Course objectives:

● To understand the mean, variance, regression models and error term for use in Multivariate data analysis.

● To understand the correlation between the data for decision making.

● To understand the various tests used for the data analysis.

● To explore various techniques for data analysis and visualize the results.

Sl.NO	Experiments (Implementation using Python/R Programming)
1	Program on data wrangling: Combining and merging datasets, Reshaping and Pivoting
2	Program on Data Transformation: String Manipulation, Regular Expressions
3	Program on Time series: GroupBy Mechanics to display in data vector, multivariate time series and forecasting formats
4	Program to measure central tendency and measures of dispersion: Mean, Median, Mode, Standard Deviation, Variance, Mean deviation and Quartile deviation for a frequency distribution/data.
5	Program to perform cross validation for a given dataset to measure Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and R^2 Error using Validation Set, Leave One Out Cross-Validation(LOOCV) and K-fold Cross-Validation approaches
6	Program to display Normal, Binomial Poisson, Bernoulli disrtibutions for a given frequency distribution and analyze the results.
7	Program to implement one sample, two sample and paired sample t-tests for a sample data and analyse the results.
8	Program to implement One-way and Two-way ANOVA tests and analyze the results
9	Program to implement correlation, rank correlation and regression and plot x-y plot and heat maps of correlation matrices.
10	Program to implement PCA for Wisconsin dataset, visualize and analyze the results.
11	Program to implement the working of linear discriminant analysis using iris dataset and visualize the results.
12	Program to Implement multiple linear regression using iris dataset, visualize and analyze the results.

Course outcomes (Course Skill Set):
At the end of the course the student will be able to:

● Design the experiment for the given problem using statistical methods.

● Develop the solution for the given real world problem using statistical techniques.

● Analyze the results and produce substantial written documentation.

Assessment Details (both CIE and SEE)

The weightage of Continuous Internal Evaluation (CIE) is 50% and for Semester End Exam (SEE) is 50%. The minimum passing mark for the CIE is 40% of the maximum marks (20 marks out of 50) and for the SEE minimum passing mark is 35% of the maximum marks (18 out of 50 marks). A student shall be deemed to have satisfied the academic requirements and earned the credits allotted to each subject/ course if the student secures a minimum of 40% (40 marks out of 100) in the sum total of the CIE (Continuous Internal Evaluation) and SEE (Semester End Examination) taken together

Continuous Internal Evaluation (CIE):
CIE marks for the practical course are **50 Marks**.
The split-up of CIE marks for record/ journal and test are in the ratio **60:40**.

- Each experiment is to be evaluated for conduction with an observation sheet and record write-up. Rubrics for the evaluation of the journal/write-up for hardware/software experiments are designed by the faculty who is handling the laboratory session and are made known to students at the beginning of the practical session.

- Record should contain all the specified experiments in the syllabus and each experiment write-up will be evaluated for 10 marks.

- Total marks scored by the students are scaled down to **30 marks** (60% of maximum marks).

- Weightage to be given for neatness and submission of record/write-up on time.

- Department shall conduct a test of 100 marks after the completion of all the experiments listed in the syllabus.

- In a test, test write-up, conduction of experiment, acceptable result, and procedural knowledge will carry a weightage of 60% and the rest 40% for viva-voce.

- The suitable rubrics can be designed to evaluate each student's performance and learning ability.

- The marks scored shall be scaled down to **20 marks** (40% of the maximum marks).

The Sum of scaled-down marks scored in the report write-up/journal and marks of a test is the total CIE marks scored by the student.

Semester End Evaluation (SEE):

- SEE marks for the practical course are 50 Marks.

- SEE shall be conducted jointly by the two examiners of the same institute, examiners are appointed by the Head of the Institute.

- The examination schedule and names of examiners are informed to the university before the conduction of the examination. These practical examinations are to be conducted between the schedule mentioned in the academic calendar of the University.

- All laboratory experiments are to be included for practical examination.

- (Rubrics) Breakup of marks and the instructions printed on the cover page of the answer script to be strictly adhered to by the examiners. **OR** based on the course requirement evaluation rubrics shall be decided jointly by examiners.

- Students can pick one question (experiment) from the questions lot prepared by the examiners jointly.

- Evaluation of test write-up/ conduction procedure and result/viva will be conducted jointly by examiners.

• General rubrics suggested for SEE are mentioned here, writeup-20%, Conduction procedure and result in -60%, Viva-voce 20% of maximum marks. SEE for practical shall be evaluated for 100 marks and scored marks shall be scaled down to 50 marks (however, based on course type, rubrics shall be decided by the examiners)
Change of experiment is allowed only once and 15% of Marks allotted to the procedure part are to be made zero.
The minimum duration of SEE is 02 hours

Suggested Learning Resources:
- Chris Chatfield, The Analysis of Time Series: An Introduction, 6th Edition Chapman and Hall/CRC, 2003
- Garett Grolemund, Hands-on Programming with R, 1st Edition, O'Reilly, 2014
- Jobson J Dave, Applied Multivariate data analysis Vol I and II, 2012 Springer-Verlag New York Inc...
- Anderson T W, An Introduction to Multivariate Statistical Analysis, 3rd Edition, Wiley publications, 2009
- Mark Lutz, Programming Python, 4th Edition, O'Rielly Medeia, 2012
- https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data
- https://www.kaggle.com/datasets/arshid/iris-flower-dataset
- https://www.youtube.com/watch?v=VSRUm3HRoiU
- https://www.youtube.com/watch?v=DkwvAn9AAU0
- https://www.youtube.com/playlist?list=PLoROMvodv4rPP6braWoRt5UCXYZ71GZIQ