Department of Computer Applications
National Institute of Technology, Tiruchirappalli

1. Course Outline

Course Title	STATISTICAL COMPUTING		
Course Code	CA601		
Department	CA	**No. of Credits**	3
Programme	M.Tech.	**Learning Hours**	3
Course Type	Programme Core	**Course Teacher**	Dr. P.J.A. Alphonse
Pre-requisites	NA		
E-mail	alphonse@nitt.edu	**Telephone No.**	0431-2503742
Course Type	Core Course	**Office**	Lyceum 107

2. Course Content

Covers practical issues in statistical computing and basic principles of the theory of probability and its applications. Topics include probability theory, Sampling distributions & descriptive statistics, analysis of variance ANOVA, tabular data, density estimation, introduction of statistical programming language R, organizing and commenting code, debugging, functions, creating R packages with documentation, and introduction to Python Packages

3. Course Objectives

1. To learn the probability distributions and density estimations to perform analysis of various kinds of data.
2. To explore the statistical analysis techniques using Python and R programming languages.
3. To expand the knowledge in R and Python to use it for further research.

4. Course Learning Outcomes (CO)

Students will be able to:

1. Understand axioms of Probability theory and have working knowledge with probability calculations.
2. Implement statistical analysis techniques for solving practical problems.
3. Perform statistical analysis on variety of data.
4. Perform appropriate statistical tests using R and visualize the outcome.
5. Obtain the knowledge on python packages

6. Course Teaching and Learning Activities

Week	Mode of Delivery	Topics
1.	Classroom activity	Introduction of probability theory, Sample Spaces and Events
		Axioms, Counting, Conditional Probability and Bayes' Theorem
		Mean and Variance of a Random variable-Binomial distributions
2.	Classroom activity	Poison distribution
		Exponential and Normal distributions
3.	Classroom activity	Curve Fitting and Principles of Least Squares
		Regression and correlation.
		The Central Limit Theorem
4.	Classroom activity	distributions of the sample mean and the sample variance for a normal population
		Sampling distributions (Chi-Square, t, F, z).
		Test of Hypothesis and Testing for Attributes
5.	Classroom activity	Mean of Normal Population and One-tailed and two-tailed tests,
		F-test and Chi-Square test
		Analysis of variance ANOVA and One way and two way classifications.
6.	Classroom activity	Tabular data and Power and the computation of sample size
		Advanced data handling and Multiple regression
		Linear models and Logistic regression
7.	Classroom activity	Rates and Poisson regression Nonlinear curve fitting.
		Density Estimation and Recursive Partitioning
		Smoothers and Generalized Additive Models
8.	Classroom activity	Survivals Analysis, Analyzing Longitudinal, Data Simultaneous Inference and Multiple Comparisons
		Meta-Analysis and Principal Component Analysis
		Multidimensional Scaling and Cluster Analysis
9.	Classroom activity	Introduction to R- Packages, Scientific Calculator, Inspecting Variables, Vectors and Matrices and Arrays
		Lists and Data Frames, Functions and Strings and Factors
		Flow Control and Loops, Advanced Looping and Date and Times
10.	Classroom activity	Introduction to Python Packages and Fundamentals of Python
		Inserting and Exporting Data and Data Cleansing
		Checking and Filling Missing Data- Merging Data- Operations-Joins

- All relevant material will be made available to the students in advance. Classroom activity includes lectures, tutorials, quiz, simulation exercise, laboratory exercise, mini-project, group task and seminar.

The assessment details for this course are given below. The assessment will be done for a total of 100 marks.

7. Course Assessment Methods – Theory

Sl. No.	Mode of Assessment	Portions	Schedule	Duration in Minutes	Weightage (%)
1.	Test – 1	Unit I	4th week	60	10
2.	Test – 2	Unit V	6th week	60	10
3.	Test - 3	Unit II	8th week	60	10
4.	Test - 4	Unit III	10th week	60	10
5.	Class activity	Periodic	Course duration	30	10
6.	End Semester Exam	All Units	November	180	50
				Total	100

* Evaluated based on participation in lecture class and class activity.

8. Essential Learning material (Textbooks, Reference books, Websites, Journals, etc.)

1. Richard Cotton, "Learning R", O'Reilly, 2013.
2. Dalgaard, Peter, "Introductory statistics with R", Springer Science & Business Media, 2008.
3. Brain S. Everitt, "A Handbook of Statistical Analysis Using R", Second Edition, 4 LLC, 2014.
4. Samir Madhavan, "Mastering Python for Data Science", Packt, 2015.
5. Sheldon M. Ross,"Introduction to Probability and Statistics for Engineers and Scientists", 4 th edition, Academic Press; 2009.
6. Paul Teetor, "R Cookbook, O'Reilly, 2011.
7. Mark Lutz ,"Learning Python", O'Reilly,5th Edition,2013

9. Course Exit Survey

- The students may give their feedback at any time to the course Teacher or through an email message, which will be duly addressed.
- The students may also give their feedback during Class Committee meeting and fill up the feedback form at the end of the semester.

10. Course Policy (including plagiarism, academic honesty, attendance, etc.)

Classroom Behavior

- Ensure that the course atmosphere, both in the class and discussions outside the class room with Teacher, is conducive for learning. Participate in discussions but do not dominate or be abusive. Be considerate of your fellow students and avoid disruptive behavior.

Exam policy

- Each student is required to take all exams at the scheduled times. All exceptions must be cleared with the professor prior to the exam time. Exams missed for insufficient reason and without being cleared with the professor prior to the exam time will be assigned a score of zero.

Assignments

- All assignments are due on or before the mentioned date and time.

Late assignments
- Late submissions are not accepted.

Plagiarism
- The students are expected to come out with their original work on term activity, assignments and tests/examinations. If found to be plagiarized, it will be assigned a score of zero.

Attendance
- Attendance is expected. If a student misses a class, the student is still responsible for the material that is studied and for completing any assignments by the due date that may have been handed out by the instructor during class.

Academic Honesty

i) No type of academic dishonesty will be tolerated. If the student is caught cheating (on the assignments, exams, or project) the punishment will be the most severe penalty allowed by the Institute policy.
ii) Possession of any electronic device, if any, found during the test or exam, the student will be debarred for 3 years from appearing for the exam and this will be printed in the Grade statement/Transcript.
iii) Tampering of MIS records, if any, found, then the results of the student will be withheld and the student will not be allowed to appear for the Placement interviews conducted by the Office of Training & Placement, besides (i).

11. Additional Course Information
- The students can get their doubts clarified during class.
- Prior request for appointment through mail, stating the subject matter to be discussed, is required to fix a time for discussion of subject matter outside class. Appointment time will be communicated through reply mail.

For Senate's Consideration

(Dr. P.J.A. Alphonse)
Course Faculty

(Dr. R.Eswari)
PAC Chairperson

(Dr. S. R. Balasundaram)
HoD