

Course Code: 302
Course Title: Statistical Methods and Data Analysis

Course Code	302																																																															
Course Title	Statistical Methods and Data Analysis (Multi-Disciplinary Course – 03) [Title of the course will be the one selected by the student from courses offered by college/institute out of the course basket offered by the University under the Multi-Disciplinary courses or Inter-disciplinary courses.]																																																															
Credits	4																																																															
Course Category	Multidisciplinary Course (MDC-03)																																																															
Level of Course	200-299 (Intermediate Level Course)																																																															
Teaching per Week	4 Hrs.																																																															
Minimum weeks per Semester	15 (Including class work, examination, preparation etc.)																																																															
Review / Revision	-																																																															
Implementation Year:	A.Y. 2024-2025																																																															
Purpose of Course	To equip students with the fundamental principles and techniques necessary to analyze and interpret data across various disciplines. Through hands-on experience and theoretical understanding, students will gain proficiency in statistical methods essential for making informed decisions and drawing meaningful insights from complex datasets, fostering interdisciplinary problem-solving skills. [Student will opt any one course of multi-disciplinary nature from other than the computer Science and Application faculty. The course will be offered by the institute/college passed by the Board of Studies of University faculties other than the computer science and application faculty.]																																																															
Course Objective	1. Develop fundamental level knowledge of statistical data analysis, including data manipulation, visualization, and modelling using R programming language. 2. Understand and apply basic statistical concepts and techniques such as descriptive statistics, 3. Gain practical experience in cleaning, exploring, and preparing datasets for analysis, emphasizing reproducible research practices. 4. Enhance critical thinking and problem-solving skills by applying statistical methods to real-world datasets and interpreting results effectively using R.																																																															
Pre-requisite	Knowledge of Fundamentals of Statistics and Mathematics of 10 th Grade Level																																																															
Course Outcomes	CO1: Understand foundational statistical concepts including descriptive statistics, probability theory, and basic inferential statistics. CO2: Apply statistical techniques such as hypothesis testing, confidence intervals, and correlation analysis to analyze and interpret data accurately. CO3: Demonstrate proficiency in data visualization methods to effectively communicate statistical findings and insights. CO4: Utilize basic statistical software tools or programming languages like R or Python to perform data analysis and visualization tasks. CO5: Develop critical thinking skills to assess the validity and reliability of statistical analyses and draw appropriate conclusions from data. CO6: Apply statistical reasoning to real-world scenarios and make informed decisions based on data-driven insights.																																																															
Mapping between Course Outcomes(CO) with Program Specific Outcomes(PSO)	<table border="1"> <thead> <tr> <th></th> <th>PSO1</th> <th>PSO2</th> <th>PSO3</th> <th>PSO4</th> <th>PSO5</th> <th>PSO6</th> <th>PSO7</th> <th>PSO8</th> </tr> </thead> <tbody> <tr> <td>CO1</td> <td style="background-color: #cccccc;"></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>CO2</td> <td></td> <td style="background-color: #cccccc;"></td> <td></td> <td style="background-color: #cccccc;"></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>CO3</td> <td style="background-color: #cccccc;"></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td style="background-color: #cccccc;"></td> <td></td> </tr> <tr> <td>CO4</td> <td style="background-color: #cccccc;"></td> <td style="background-color: #cccccc;"></td> <td style="background-color: #cccccc;"></td> <td></td> <td style="background-color: #cccccc;"></td> <td></td> <td></td> <td></td> </tr> <tr> <td>CO5</td> <td></td> <td></td> <td></td> <td></td> <td style="background-color: #cccccc;"></td> <td style="background-color: #cccccc;"></td> <td style="background-color: #cccccc;"></td> <td></td> </tr> <tr> <td>CO6</td> <td style="background-color: #cccccc;"></td> <td></td> <td></td> <td></td> <td style="background-color: #cccccc;"></td> <td></td> <td></td> <td style="background-color: #cccccc;"></td> </tr> </tbody> </table>		PSO1	PSO2	PSO3	PSO4	PSO5	PSO6	PSO7	PSO8	CO1									CO2									CO3									CO4									CO5									CO6								
	PSO1	PSO2	PSO3	PSO4	PSO5	PSO6	PSO7	PSO8																																																								
CO1																																																																
CO2																																																																
CO3																																																																
CO4																																																																
CO5																																																																
CO6																																																																
Course Outcome	After studying the course, students will be able to Implement acquired skills in writing codes using programming languages.																																																															

Course Content	<p>Unit-1: Basic concepts of statistic 1.1 Population vs. sample, variables (categorical vs. numerical), datatypes. 1.2 Descriptive statistics: measures of central tendency (mean, median, mode), 1.3 Measures of dispersion (range, variance, standard deviation)</p> <p>Unit-2: Data Representation and Sampling technique 2.1 Graphical representation of data (histograms, box plots, scatter plots) 2.2 Probability theory: basic probability concepts 2.3 Probability distributions (binomial, normal distributions) 2.4 Sampling techniques: random sampling, stratified sampling, 2.5 sampling distributions. 2.6 Understanding Bell curve.</p> <p>Unit-3: Introduction to R and working with Data 3.1 Overview of R and its applications in data analysis and statistics. 3.2 Installing R and RStudio. 3.3 Basic R syntax, variables, and data types. 3.4 Importing data into R from different file formats (CSV, Excel, etc.). 3.5 read, write and view data using data frames.</p> <p>Unit-4: Data Filtering and cleaning 4.1 Subsetting and filtering data. 4.2 Adding, removing, and renaming variables/Attributes. 4.3 Data Cleaning and Transformation 4.4 Identifying and handling missing values. 4.5 Data type conversion and recoding variables.</p> <p>Unit-5: Working with Data in R 5.1 Reordering and reshaping data frames. 5.2 Merging and joining data frames. 5.3 Calculating summary statistics (mean, median, mode, standard deviation). 5.4 Generating frequency tables and cross-tabulations. 5.5 Commands to measures of central tendency and dispersion. 5.6 Concepts of normal distribution. 5.7 Commands to explore view data distributions graphically (Bell curve).</p>
Reference Books	<ol style="list-style-type: none"> 1. "An Introduction to Statistical Learning: with Applications in R" by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, Publisher: Springer, ISBN: 978-1461471370 2. "R for Data Science: Import, Tidy, Transform, Visualize, and Model Data" by Hadley Wickham and Garrett Golemund, Publisher: O'Reilly Media, ISBN: 978-1491910399 3. "Discovering Statistics Using R" by Andy Field, Jeremy Miles, and Zoe Field Publisher: SAGE Publications Ltd, ISBN: 978-1446200469 4. "Practical Data Science with R" by Nina Zumel and John Mount Publisher: Manning Publications, ISBN: 978-1617291562 5. "Statistics: Unlocking the Power of Data" by Robin H. Lock, Patti Frazer Lock, Kari Lock Morgan, and Eric F. Lock, Publisher: Wiley, ISBN: 978-1119325572 6. "The Art of R Programming: A Tour of Statistical Software Design" by Norman Matloff, Publisher: No Starch Press, ISBN: 978-1593273842 7. "Introduction to Probability and Statistics Using R" by G. Jay Kerns, Publisher: RStudio, PBC, ISBN: 978-1886529450 8. "Business Analytics – The science of Data-Driven Decision Making" by U.Dinesh Kumar, Publisher: Wiley, ISBN: 978-81-265-6872-2
Teaching Methodology	Class Work, Discussion, Self-Study, Seminars and/or Assignments
Evaluation Method	50% Internal assessment. 50% External assessment.