SIR C.R.REDDY COLLEGE OF ENGINEERING, ELURU DEPARTMENT OF INFORMATION TECHNOLOGY <u>COURSE HANDOUT</u>



SUBJECT: STATISTICS WITH R CLASS: II/IV B.Tech. II SEMESTER, A.Y.2022-23

INSTRUCTOR: SRI P RAMAIAH CHOWDARY

Course Handout Index

S. No	Description
1	College Vision & Mission
2	Department Vision & Mission
3	Program Educational Objectives (PEOs)
4	Program Outcomes (POs)
5	Program Specific Outcomes (PSOs)
6	JNTUK Academic Calendar
7	Department Academic Calendar
8	Course Description
9	Course Objectives
10	Course Outcomes
11	Lesson Plan
12	Evaluation Pattern
13	Timetable
14	Unit wise Questions

College Vision & Mission

Vision: To emerge as a premier institution in the field of technical education and research in the state and as a home for holistic development of the students and contribute to the advancement of society and the region.

Mission: To provide high quality technical education through a creative balance of academic and industry oriented learning; to create an inspiring environment of scholarship and research; to instill high levels of academic and professional discipline; and to establish standards that inculcate ethical and moral values that contribute to growth in career and development of society in general.

Department Vision & Mission

Vision: To be a premier department in the region in the field of Information Technology through academic excellence and research that enable graduates to meet the challenges of industry and society.

Mission: To Provide dynamic teaching-learning environment to make the students industry ready and advancement in career; to inculcate professional and leadership quality for better employability and entrepreneurship; to make high quality professional with moral and ethical values suitable for industry and society.

Program Educational Objectives (PEOs)

PEO1: Solve real world problems through effective professional skills in Information Technology industry and academic research.

PEO2: Analyze and develop applications in Information Technology domain and adapt to changing technology trends with continuous learning.

PEO3: Practice the profession in society with ethical and moral values.

Program Outcomes (POs)

PO1: Engineering Knowledge: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

PO2: Problem Analysis: Identify, formulate, research literature, and analyze complex engineering problems reaching substantiated conclusions using the first principles of mathematics, natural sciences, and engineering sciences.

PO3: Design/Development of Solutions: Design solutions for complex engineering problems and system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, society, and environmental considerations.

PO4: Conduct Investigations of Complex Problems: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

PO5: Modern Tool Usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

PO6: The Engineer and Society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

PO7: Environment and Sustainability: Understand the impact of the professional engineering solutions in society and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

PO8: Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

PO9: Individual and Team Work: Function effectively as an individual, and as a member or leader in diverse teams, and in multi-disciplinary settings.

PO10: Communication: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

PO11: Project Management and Finance: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multi-disciplinary environments.

PO12: Life-long Learning: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

Program Specific Outcomes (PSOs)

PSO1: Design Skill: Design and develop software's in the area of relevance under realistic constraints.

PSO2: New Technology: Adapt new and fast emerging technologies in the field of Information Technology.

JNTUK Academic Calendar

Website: www.jntuk.edu.in Email: dap@jntuk.edu.in



Phone: 0884-2300991

Directorate of Academic Planning

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY KAKINADA KAKINADA-533003, Andhra Pradesh, INDIA

(Established by AP Government Act No. 30 of 2008)

Date 02.11.2022

Lr. No. DAP/RAC/II Year /B. Tech/2022 Dr. KVSG Murali Krishna,

M.E. Ph.D., Director, Academics & Planning JNTUK, Kakinada

То

All the Principals of Affiliated Colleges, JNTUK, Kakinada.

Revised Academic Calendar for II Year - B. Tech for the AY 2022-23 (2021-22 Admitted Batch)

I SEMESTER			
Description	From	То	Weeks
Community Service Project	22.08.2022	03.09.2022	2 W
I Unit of Instruction	05.09.2022	29.10.2022	8W
I Mid Examinations	24.10.2022	29.10.2022	
II Unit of Instructions	31.10.2022	24.12.2022	8W
II Mid Examinations	19.12.2022	24.12.2022	
Community Service Project for Lateral Entry Students, Preparation & Practicals	26.12.2022	14.01.2022	3 W
End Examinations	18.01.2023	28.01.2023	2W
Commencement of II Semester Class Work	28.01.2023		
II SEMESTER			
I Unit of Instructions	30.01.2023	25.03.2023	8W
I Mid Examinations	20.03.2023	25.03.2023	
II Unit of Instructions	27.03.2023	20.05.2023	8 W
II Mid Examinations	15.05.2023	20.05.2023	
Preparation & Practicals	22.05.2023	27.05.2023	1 W
End Examinations	29.05.2023	10.06.2023	2W

* As per the APSCHE Guidelines Out of the Total 180 hours of Community Service Project leading to 4 Credits, two weeks will be offline and remaining project work can be done during the II-I semester weekends and holidays.

Director,

Academics & Planning Copy to the Secretary to the Hon'ble Vice Chancellor, JNTUK Copy to Rector, JNTUK

Copy to Registrar, JNTUK

Copy to Director Academic Audit, JNTUK Copy to Director of Evaluation, JNTUK

Course Description

In this course, you will learn to analyse and visualize data in R and create reproducible data analysis reports, demonstrate a conceptual understanding of the unified nature of statistical inference, perform frequents and Bayesian statistical inference and modelling to understand natural phenomena and make data-based decisions, communicate statistical results correctly, effectively, and in context without relying on statistical jargon, critique data-based claims and evaluated data-based decisions, and wrangle and visualize data with R packages for data analysis.

You will produce a portfolio of data analysis projects from the Specialization that demonstrates mastery of statistical data analysis from exploratory analysis to inference to modelling, suitable for applying for statistical analysis or data scientist positions.

Course Objectives

After taking the course, students will be able to

- Use R for statistical programming, computation, graphics, and modeling,
- Write functions and use R in an efficient way,
- Fit some basic types of statistical models
- Use R in their own research,
- expand their knowledge of R on their own

Course Outcomes

At the end of this course, students will be able to:

CO Number	COs	Lev el
C01	Understand the R programming basic concepts required for statistics.	L2
CO2	Apply basic Math and statistics in R programming	L3
CO3	Sketch visualizations/graphics in R Programming.	L3
CO4	Analyze the R programming constructs/models required to perform certain statistics.	L4

Remembering	Understanding	Applying	Analyzing	Evaluating	Creating
L1	L2	L3	L4	L5	L6

II Vara II Consistent	L	Т	Р	С
II Year – II Semester	2	0	2	3
STATISTICS	VITH R			s

<u>Syllabus</u>

UNIT-I: Introduction, How to run R, R Sessions and Functions, Basic Math, Variables, Data Types, Vectors, Conclusion, Advanced Data Structures, Data Frames, Lists, Matrices, Arrays, Classes.

UNIT-II: R Programming Structures, Control Statements, Loops, - Looping Over Nonvector Sets,-IfElse, Arithmetic and Boolean Operators and values, Default Values for Argument, Return Values, Deciding Whether to explicitly call return- Returning Complex Objects, Functions are Objective, No Pointers in R, Recursion, A Quicksort Implementation-Extended Extended Example: A Binary Search Tree.

UNIT-III: Doing Math and Simulation in R, Math Function, Extended Example Calculating ProbabilityCumulative Sums and Products-Minima and Maxima- Calculus, Functions Fir Statistical Distribution, Sorting, Linear Algebra Operation on Vectors and Matrices, Extended Example: Vector cross Product- Extended Example: Finding Stationary Distribution of Markov Chains, Set Operation, Input /output, Accessing the Keyboard and Monitor, Reading and writer Files,

UNIT-IV: Graphics, Creating Graphs, The Workhorse of R Base Graphics, the plot() Function – Customizing Graphs, Saving Graphs to Files. Probability Distributions, Normal Distribution-Binomial Distribution- Poisson Distributions Other Distribution, Basic Statistics, Correlation and Covariance, T-Tests,-ANOVA

UNIT-V: Linear Models, Simple Linear Regression, -Multiple Regression Generalized Linear Models, Logistic Regression, - Poisson Regression- other Generalized Linear Models-Survival Analysis, Nonlinear Models, Splines- Decision- Random Forests,

TEXT BOOKS: 1) The Art of R Programming, Norman Matloff, Cengage Learning 2) R for Everyone, Lander, Pearson

REFERENCE BOOKS: 1) R Cookbook, PaulTeetor, Oreilly. 2) R in Action, Rob Kabacoff, Manning

Lesson Plan

S. No	Unit	Торіс	Teaching Aids	CO
1.		Introduction	BB	CO1
2.		How to run R	BB	CO1
3.		R Sessions and Functions	BB	CO1
4.		Basic Math.	BB	CO1
5.		Variables	BB	CO1
6.		Data Types	PPT	CO1
7.	Ι	Vectors	PPT	CO1
8.		Advanced Data Structures	PPT	CO1
9.		Data Frames	BB/PPT	CO1
10.		Lists	BB	CO1
11.		Matrices	BB/PPT	CO1
12.		Arrays	BB	CO1
13.		Classes	BB/PPT	CO1
14.		R Programming Structures,	BB	CO1
15.		Control Statements	BB/PPT	CO1
16.		Loops	BB/PPT	CO1
17.		Looping Over Non-vector Sets,- IfElse	BB/PPT	CO1
18.		Arithmetic and Boolean Operators and values	BB	CO1
19.		Default Values for Argument	BB	CO1
20.	II	Return Values	BB	CO1
21.		Deciding Whether to explicitly call return	BB/PPT	CO1
22.		Returning Complex Objects,	BB	CO1
23.		Functions are Objective	BB/PPT	CO1
24.		No Pointers in R, Recursion	BB	CO1
25.		A Quicksort Implementation-Extended	BB	CO1
26.		Extended Example: A Binary Search Tree	BB/PPT	CO1
27.		Doing Math and Simulation in R	BB/PPT	CO2
28.	III	Math Function, Extended Example	BB	CO2
29.		Calculating Probability Cumulative Sums and Products	BB/PPT	CO2

30.		Minima and Maxima	BB	CO2
31.		Calculus,	BB / PPT	CO2
32.		Functions For Statistical Distribution,	BB / PPT	CO2
33.		Sorting	BB	CO2
34.		Linear Algebra Operation on Vectors and Matrices	BB	CO2
35.		Extended Example: Vector cross Product-	BB	CO2
36.		Extended Example: Finding Stationary Distribution of Markov Chains	BB/PPT	CO2
37.		Set Operation,	BB	CO2
38.		Input /output	BB	CO2
39.		Accessing the Keyboard and Monitor	BB /PPT	CO2
40.		Reading and writer Files	BB	CO2
41.		Graphics, Creating Graphs,	BB	CO3
42.		The Workhorse of R Base Graphics	BB / PPT	CO3
43.		the plot() Function	BB	CO3
44.		Customizing Graphs	BB / PPT	CO3
45.		Saving Graphs to Files	BB	CO3
46.		Probability Distributions	BB	CO3
47.		Normal Distribution	BB	CO3
48.		Binomial Distribution	BB	CO3
49.	IV	Poisson Distributions	BB /PPT	CO3
50.		Other Distribution	BB	CO3
51.		Basic Statistics	BB /PPT	CO3
52.		Correlation and Covariance	BB	CO3
53.		T-Tests,-ANOVA	BB	CO3
54.		Linear Models	BB/PPT	CO4
55.		Simple Linear Regression	BB	CO4
56.		Multiple Regression Generalized	BB/PPT	CO4
57.	V	Logistic Regression	BB	CO4
58.		Poisson Regression	BB/PPT	CO4
59.		other Generalized Linear Models	BB	CO4
60.		Survival Analysis	BB/PPT	CO4

61.	Nonlinear Models	BB/PPT	CO4
62.	Splines	BB	CO4
63.	Decision	BB/PPT	CO4
64.	Random Forests	BB	CO4

Evaluation Pattern

The assessment of the student's performance in each course will be as per the details given:

S. No	Components	Internal	External	Total
1	Theory	30	70	100
2	Engineering Graphics/Design/Drawing	30	70	100
3	Practical	15	35	50
4	Mini Project/Internship/Industrial Training/ Skill Development programmes/Research Project	2 2)	50	50
5	Project Work	60	140	200

Continuous Internal Theory Evaluation:

a) For theory subjects, during a semester, there shall be two mid-term examinations. Each mid-term examination consists of (i) one online objective examination (20 multiple choice questions) for 10 marks for a duration of 20 minutes (ii) one descriptive examination (3 full questions for 5 marks each) for 15 marks for a duration of 90 minutes and (iii) one assignment for marks. All the internal exams shall be conducted as per university norms from first 50% of the syllabi.

b) In the similar lines, the second online, descriptive examinations assignment shall be conducted on the rest of the 50% syllabus.

c)Internal marks can be calculated with 80% weightage for better of the two mids and 20% Weightage for other mid exam.

Example:

Mid-1 marks = Marks secured in (online examination-1+descriptive examination-1 + one assignment-1)

Mid-2 marks = Marks secured in (online examination-2+descriptive examination-2 +one assignment-2)

Final internal Marks = (Best of (Mid-1/Mid-2) marks x 0.8 + Least of (Mid-1/Mid-2) marks x 0.2)

Timetable

SIR CRREDDY COLLEGE OF ENGINEERING, ELURU DEPARTMENT OF INFORMATION TECHNOLOGY

II/IV B.Tech IT - II SEMESTER A.Y 2022-2023

TIME TABLE A-SECTION

With effect from : 30-01-2023 Class Teacher:Sri.Y.SivaKoteswaraRao/P.Ramaiahchowdary LH-45

	1	2	3	4		5	6	7	8
DAY	09:00 To 09:50	09:50 To 10:40	11:00 To 11:50	11:50 To 12:40		1:40 To 2:30	02:30 To 03:20	03:20 To 04:10	04:10 To 5:00
MON	SRP	LAB	AT&CD	PSE(T)	L		UML LAI	3	AT&CD(M)
TUE	JAVA	PSE	MEFA	SRP(T)	U	FOSS LAB			MEFA(M)
WED	AT&CD		SOC LAB		N	SRP	MEFA	JAVA(T)	SRP(M)
THU	MEFA	SRP	AT&CD (T)	JAVA	C		JAVA LAB		
FRI	PSE	AT&CD	MEFA(T)	JAVA	H	PSE	SRP	Counseling	JAVA(M)
SAT	MEFA	AT&CD	JAVA	PSE		AT&CD (R)	JAVA (R)	PSE/SRP (R)	MEFA (R)

*T - Tutorial

*R- Remedial Class *M- Make Up Class STAFF: Statistics with R R2022121

Principles of Software Engineering R2022122

Automata Theory and Compiler Design R2022123

Java Programming R2022054

UML Lab R2022124

FOSS Lab R2022125

:Sri. P.Ramaiah chowdary
: Dr. G.Chamundeswari
: Sri. Ch. Yallamanda

: Dr. K.Satyanarayana

Managerial Economics and financial AccountancyR2022055:Sri.Y.Siva Koteswara Rao

:Dr.G.Chamundeswari/Sri.P.RamaiahChowdary

: Sri. E.B.K.Manash / Sri. G.Vihari

:Dr. K.Satyanarayana /Sri. P.Rajendra Kumar

:Sri.S.Uma Sheshagiri Rao/Sri.P.RamaiahChowdary

G. Krishaler

Dept. Time Table Incharge

Java Programming Lab R2022126

Skill oriented Course 41 R2022128

HOD, HEAD OF DEFIDERARTMENT Information Technology Sir C.R.R. College of Engg. ELURU-524 007,

PRINCIPAL

Principal Sir C.R.R.College of Engineering ELURU - 534 007

SIR CRREDDY COLLEGE OF ENGINEERING, ELURU **DEPARTMENT OF INFORMATION TECHNOLOGY**

II/IV B.Tech IT - II SEMESTER A.Y 2022-2023

TIME TABLE B-SECTION

With effect from :30-01-2023 Class Teacher: Dr. G.Chamundeswari/Ch.Yallamanda LH-46

	1	2	3	4		5	6	7	8
DAY	09:00 To 09:50	09:50 To 10:40	11:00 To 11:50	11:50 To 12:40		1:40 To 2:30	02:30 To 03:20	03:20 To 04:10	04:10 To 5:00
MON	PSE		JAVA LA	В	L	AT&CD	MEFA(T)	Counseling	AT&CD(M)
TUE	SRP	AT&CD	JAVA	PSE	U	MEFA	SRP	JAVA(T)	MEFA(M)
WED	JAVA	MEFA	PSE	AT&CD	N		FOSS LAB		SRP(M)
THU	AT&CD		UML LA	В	C	PSE	PSE MEFA SRP(T)		
FRI	MEFA	JAVA	SRP	AT&CD(T)	H	Real Mr.	SOC LAB		JAVA(M)
SAT	JAVA	PSE(T)	SF	RP LAB		AT&CD (R)	JAVA (R)	PSE/SRP (R)	MEFA (R)

*T - Tutorial

*R- Remedial Class *M- Make Up Class STAFF: Statistics with R R2022121

Principles of Software Engineering R2022122

Automata Theory and Compiler Design R2022123

R2022054 Java Programming

Managerial Economics and financial Accountancy R2022055 : Sri. Y.Siva Koteswara Rao

UML Lab R2022124

FOSS Lab R2022125

Java Programming Lab R2022126

Skill oriented Course 41 R2022128

g. Knisharlen

Dept. Time Table Incharge

: Sri. P.Ramaiah Chowdary : Dr. G. Chamundeswari : Sri. Ch. Yallamanda : Dr. K.Satyanarayana

:Dr.G.Chamundeswari/Sri.S.Uma Sheshagiri Rao

:Sri. G.Vihari /Sri. E.B.K.Manash

: Dr. K.Satyanarayana /Sri. P.Rajendra Kumar

: Sri. S.Uma Sheshagiri Rao / Sri.V.Gopinath

HEAD OPT TOTE HEPARTMENT HEAD METHE REPARTMENT SIr C.R.R.College of Engineering Sir C.R.P. College of Engineering ELURU - 534 007 R.P. Colication of Emgg.

Principal

Google Class room link:

https://classroom.google.com/c/NTk1ODk4NDExODYy?cjc=2lengeo class code: 2lengeo

Unit-wise questions

<u>UNIT – 1</u> Introduction to R

- 1. Describe the features of R.
- 2. Can you list the applications of R?
- 3. What are the merits and demerits of R programming language?
- 4. How do you install and run R?
- 5. Discuss R sessions and functions.
- 6. Explain data types in R.
- 7. Discuss the data structures in R with examples.
- 8. Discuss object-oriented systems (classes) in R with suitable examples.
- 9. Explain the concept of user-defined function with an example.
- 10. What are vectors? Elaborate with the help of R program.
- 11. What do you understand from vector arithmetic? Explain with an example.
- 12. Explain matrix function in R with an example.
- 13. Explain array function in R with an example.
- 14. Discuss the differences between built-in function and user-defined function.
- 15. Write R code to create two different data frames with some common information, and then merge those two data frames.
- 16. Discuss about the operations on vectors?
- 17. Describe the concept of lists with suitable examples.
- 18. Explain about apply() family in R.
- 19. Explain the following with suitable examples.
 - a) apply()
 - b) sapply()
 - c) lapply()
 - d) tapply()

<u>UNIT – 2</u> <u>R Programming Structures</u>

- 1. Explain R programming structures.
- 2. Discuss control structures in R with suitable examples.
- 3. Discuss the concepts of operators in R.
- 4. How do use default value for an argument? Elaborate with an example.
- 5. Discuss the importance of return with an example.
- 6. Discuss complex objects in R.
- 7. Describe the concept of "No pointers in R".
- 8. Write an R program to find a factorial of a given number using recursion.
- 9. Explain quick sort algorithm with a suitable example.
- 10. Write R script to implement quick sort algorithm.
- 11. Discuss binary search tree (BST) with suitable example.
- 12. Write R code to implement binary search tree (BST).
- 13. Write an R program to check whether the given number is a prime or not.
- 14. Write an R program to print a multiplication table for the given number.
- 15. Write code in R to print the sum of n natural numbers.
- 16. Discuss reserved words in R.
- 17. Write a program to print the factorial of a given number (use loops).
- 18. What is the difference between break and next? Explain with an example.
- 19. Explain switch statement with an example.
- 20. Write an R program to print the Fibonacci series.

<u>UNIT – 3</u> <u>Math and Simulation in R</u>

- 1. Discuss Math and Simulation in R.
- 2. Discuss various math functions in R.
- 3. How do you calculate probability? Explain with an example.
- 4. Write a short note on the following with suitable example.
 - a. Cumulative sum
 - b. Cumulative product
 - c. Maxima
 - d. Minima
- 5. How do you perform calculus functions in R?

- 6. Explain about the functions for statistical distribution.
- 7. Discuss sorting with suitable examples.
- 8. Write short note on the following with suitable examples.
 - a. sort()
 - b. order()
 - c. rank()
- 9. Describe linear algebra operations on vectors and matrices.
- 10. Discuss vector cross product with an example.
- 11. Elaborate the concept of finding stationary distribution of Markov chains with an example.
- 12. Discuss set operations in R.
- 13. How do you access the keyboard and monitor in R?
- 14. Briefly discuss input and output operations in R.
- 15. Write a brief note on readLine() function.
- 16. How do you read and write files in R?

<u>UNIT – 4</u> <u>Graphics in R</u> (You should write R script)

- 1. Write short notes on graphics in R.
- 2. What are the primary styles of graphics? Discuss at least three primary styles.
- 3. Discuss the workhorse of R base graphics.
- 4. Explain graphical parameters in R.
- 5. Explain about plot() function with example.
- 6. Draw a pie chart for the given data.
- 7. Draw a bar plot for the given data.
- 8. Draw a histogram for the given data.
- 9. Draw a scatter plot for the given data.
- 10. Write the difference between high-level and low-level graphics functions.
- 11. Discuss strip chart with some example.
- 12. Discuss saving graphs to files.
- 13. How do you create a polygon in R?
- 14. How do you draw multiple curves on the same plot?
- 15. Discuss box-plot with an example.
- 16. Draw a bar plot and use legend function.

<u>Basic Statistics and Probability Distributions</u> (You should write R code after solving problem)

- 1. Descriptive statistics vs. inferential statistics.
- 2. Briefly discuss basic statistics.
- 3. What is probability? Discuss probability distributions.
- 4. Discuss measures of central tendency.
- 5. Discuss measures of dispersion.
- 6. Calculate mean, median and mode for the given data.
- 7. Calculate Range, Interquartile range, variance, standard deviation for the given data.
- 8. Calculate coefficient of variation for the given data.
- 9. Discuss covariance with an example.
- 10. Explain correlation with an example.
- 11. Calculate covariance for the given data.
- 12. Calculate correlation for the given data.
- 13. Calculate correlation coefficient for the given data (use Karl Pearson's method).
- 14. Discuss ANOVA with an example.
- 15. One-way ANOVA problem.
- 16. Two-way ANOVA problem.
- 17. Discuss t-test.
- 18. t-test problem.
- 19. z-test problem
- 20. Binomial distribution problems.
- 21. Poisson distribution problems.
- 22. Normal distribution problems.

<u>Unit-5</u> Linear Models and Non-linear Models

(You should write R code)

- 1. Write short notes on linear models.
- 2. Discuss non-linear models.
- 3. Discuss linear regression.
- 4. Discuss multiple regressions.
- 5. Explain generalized linear models.

- 6. Write R script and discuss logistic regression.
- 7. Write R script and discuss poisson regression.
- 8. What is survival analysis?
- 9. Explain decision trees.
- 10. Write short notes on random forest.
- 11. Explain splines.
- 12. Fit straight line for the given data.
- 13. Fit parabola for the given data.
- 14. Fit exponential function for the given data.