Est. in 1921

UNION CHRISTIAN COLLEGE (AUTONOMOUS) ALUVA

Affiliated to Mahatma Gandhi University, Kottayam, India
NAAC Accredited with A++ Grade in Vth cycle
0484 2609194, +91-7012626868
email: ucc@uccollege.edu.in

Department of Statistics

Statistics Syllabus

UNDERGRADUATE (HONOURS) PROGRAMMES
{UCC UGP (HONOURS)}

Adopted from THE MAHATMA GANDHI UNIVERSITY
UNDER GRADUATE PROGRAMMES
(HONOURS) SYLLABUS
MGU-UGP (Honours)
(2024 Admission Onwards)

Est. in 1921
THE
TRUTH SHALL MAKE YOU
FREE

Sl. No.	Content	Page No
	Preamble and Preface	1
	Motto, Vision, Mission, Coat of arms	2
	Foreword	3
1	Preliminary Definitions and Nomenclature	4
2	Scope and Application	4
3	Definitions	5
4	Features and Objectives	7
5	Eligibility for Admission and Reservation of Seats	11
6	Academic Monitoring and Student Support	12
7	Course Registration	13
8	Re—Admission and Scheme Migration	14
9	Duration of Programmes, Credit Requirements, Options	15
10	Credit Structure	15
11	Course Structure of the UCC UGP Programmes	17
12	Academic Levels of Pathway Courses	19
13	Signature Courses	19
14	Programme Pathways and Curriculum Structure	20
15	Credit from Other Institutions/Online/Distance Mode	25
16	Attendance	26
17	Workload	26
18	Credit Transfer and Credit Accumulation	27
19	Outcome Based Approach	27
20	Assessment and Evaluation	27
21	Practical Examination	30
22	Evaluation of Project/ Dissertation	31
23	Evaluation of Internship	31
24	Letter Grades and Grade Points	31
25	Computation of SGPA and CGPA	33
26	Implementation and Monitoring Approaches	34
27	Modifications to the Regulations	36
28	Conclusion	36

THE BOARD OF STUDIES IN STATISTICS

1. Chairperson: Head, Department of Mathematics
   Union Christian College, Aluva

2. Faculty of the Department
   a. Ms. Riswana T N
   b. Ms. Anu Mary John

3. Subject Experts from outside parent University:

   Dr. Asha Gopalakrishnan
   Professor
   Department of Statistics
   Cochin University of Science and Technology
   Kochi - 682022

   Dr. Jerin Paul
   Assistant Professor
   Department of Statistics
   Vimala College, Thrissur, Kerala

4. Representative from Industry/ Corporate Sector/ Allied area for Placement:

   Ms. Soumya Sivasankar, CEO, Support Sages Consultancy Services Pvt Ltd

5. Notable Alumni:

   Dr. Joseph Justin Rebello
   Principal, Department of Statistics
   St. Alberts College, Ernakulam

6. Subject Expert nominated by VC of parent University:

   Dr. Angel Mathew
   Professor
   Maharajas College, Ernakulam

Programme Outcomes (POs)

PO1	Critical Thinking and Analytical Reasoning
PO2	Scientific Reasoning and Problem Solving
PO3	Multidisciplinary /Interdisciplinary/ Transdisciplinary Approach
PO4	Communication Skills
PO5	Leadership Skills
PO6	Social Consciousness and Responsibility
PO7	Equity, Inclusiveness and Sustainability
PO8	Moral and Ethical Reasoning
PO9	Networking and Collaborating
PO10	Lifelong Learning

Components	**Percentage**
Continuous Comprehensive Assessment (CCA)	30
End Semester Evaluation (ESE)	70
Total	100

Syllabus Index

Name of the Major: **STATISTICS**

Semester: 1

Course Code	Title of the Course	Type of the Course DSC, MDC, SEC etc.	Credit	Hours/ week	Hour Distribution /week			
					L	T	P	O
UC1DSCSTA100	Fundamentals of Statistics and Data Visualisation	DSC A	4	5	3		2	
UC1MDCSTA100	Statistical Data Collection	MDC	3	4	2		2	
UC1MDCSTA101	Data Analysis using Libre Calc							

L — Lecture, T — Tutorial, P — Practical/Practicum , O — Others

Semester: 2

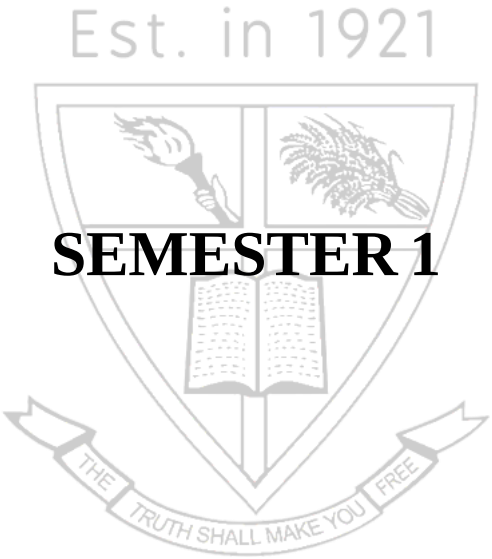
Course Code	Title of the Course	Type of the Course DSC, MDC, SEC etc.	Credit	Hours/ week	Hour Distribution /week			
					L	T	P	O
UC2DSCSTA100	Introduction to Statistical Modelling	DSC A	4	5	3		2	
UC2MDCSTA100	Time Series Methods and their Applications	MDC	3	4	2		2	
UC2MDCSTA101	Data Analysis using JAMOVI and Introduction to R							

Semester: 3

Course Code	Title of the Course	Type of the Course DSC, MDC, SEC etc.	Credit	Hours/ week	Hour Distribution /week			
					L	T	P	O
UC3DSCSTA200	Statistical Distributions	DSC A	4	4	4			
UC3DSCSTA201	Analytical Tools for Multivariate Analysis	DSC A	4	5	3		2	
UC3DSESTA200	Statistical Techniques for Data Science and Machine Learning **(Data Analytics Specialization)**	DSE	4	5	3		2	
UC3DSESTA201	Statistical Computing using R							
UC3DSESTA202	Vital Statistics and Index Numbers							
UC3DSCSTA202	Data Analysis in Inferential Statistics using R/Python	DSC B	4	5	3		2	
UC3DSCSTA203	Statistical Research Techniques using Softwares							
UC3DSCSTA204	Business Data Analytics							
UC3MDCSTA200	Statistical Analysis of Related Data	MDC	3	3	3			
UC3MDCSTA201	Data Analysis using R and Type Setting using LaTex							
UC3VACSTA200	Applied Statistical Analysis: Ethical Data Collection, Interpretation and Decision making in Society	VAC	3	3	3			

Semester: 4

Course Code	Title of the Course	Type of the Course DSC, MDC, SEC etc.	Credit	Hours/ week	Hour Distribution /week			
					L	T	P	O
UC4DSCSTA200	Basics of Multivariate Distributions	DSC A	4	5	3		2	
UC4DSCSTA201	Statistical Inference	DSC A	4	5	3		2	
UC4DSESTA200	Data Analysis Using JAMOVI **(Data Analytics Specialization)**	DSE	4	4	4			
UC4DSESTA201	Statistical Quality Control							
UC4DSESTA202	Biostatistics							
UC4DSESTA203	Econometrics							
UC4DSCSTA202	Statistical Inference using R/Python	DSC B	4	5	3		2	
UC4DSCSTA203	Statistical Research Methods using Softwares							
UC4DSCSTA204	Statistical Modelling in Data Science							
UC4SECSTA200	Introduction to Spreadsheets and Latex Typing	SEC	3	3	3			
UC4VACSTA200	Ethical Dimensions in Statistical Machine Learning through R/Python	VAC	3	3	3			
UC4INTSTA200	Internship		2					



SEMESTER 1

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Fundamentals of Statistics and Data Visualisation**					
Type of Course	**DSC A**					
Course Code	**UC1DSCSTA100**					
Course Level	**100**					
Course Summary	This course helps to acquire basic knowledge of various types of data, probability theory, correlation, regression and their real world applications. Additionally, spreadsheet functions are used to address numerical challenges associated with the topics discussed.					
Semester	**1**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3		**1**		**75**
Pre-requisites						

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains	Program Outcome
1	Explain and understand the concepts of different types of data, sampling and sampling techniques.	U	1
2	Summarise data using various measures of central tendency, dispersion, skewness and kurtosis.	U	1
3	Analyse relationships between variables using scatter diagrams, correlation coefficients and regression analysis.	A, An	1
4	Develop skills in solving real- world problems through the application of regression techniques, particularly in predicting outcomes and understanding the limitations of predictions.	An, A	2, 3
5	Understand basic probability concepts including random experiments, sample space and elementary ideas of probability.	U	2
6	Apply Bayes' theorem to update probabilities based on new information and evidence.	E	1

7	Understand how statistical concepts are relevant across disciplines, fostering interdisciplinary thinking.	U	2
8	Apply using spreadsheets to illustrate and analyse statistical concepts, enhancing practical skills.	A, An	2

Module1	**Course Description**	**Hours**	**CO NO.**
	Data and Variables, Measures of Central Tendency, Dispersion and Moments.	**15**	
1.1	Types of data and variables: Concepts of primary data and secondary data, examples of univariate and bivariate data type, Diagrams and Graphs: Bar diagrams, pie diagram and frequency graphs.	2	1
1.2	Scales of measurements: Ordinal, nominal, ratio and interval.	2	1,7
1.3	Population and sample, Types of sampling: Non-probability and Probability sampling: Simple random sampling, systematic sampling, stratified random sampling and cluster sampling with real life examples (derivations not required).	3	2
1.4	Measures of central tendency: Arithmetic Mean (AM), Geometric Mean (GM), Harmonic Mean (HM), median and mode (examples using raw data).	3	2
1.5	Measures of dispersion: Range, Quartile Deviation (QD), Mean Deviation (MD) and Standard Deviation (SD), Coefficient of Variation (CV). (examples using raw data). Box Plot.	3	2
1.6	Moments, skewness and kurtosis with examples using raw data. (derivations not required).	2	1,2
Module 2	**Correlation and Regression**	**15**	
2.1	Correlation, scatter diagram, Karl Pearson's correlation coefficient, Spearman's rank correlation coefficient. (Only the concepts, problems and properties-without proof of the above topics).	8	3
2.2	Regression: Two types of regression lines, formula and numerical problems.	7	4,7
Module 3	**Elementary Probability Theory**	**15**	
3.1	Random experiment, sample space and event with examples.	4	5
3.2	Elementary ideas of probability: Frequency, classical and axiomatic definitions with examples.	5	5
3.3	Conditional probability, independence of events, total probability law, Bayes' theorem (without proof) with examples.	6	5,6,7

Module 4	**Problem Solving using Spreadsheets** (A practical record with minimum 5 problems has to be submitted).	**30**	
4.1	Introduction to spreadsheet	**5**	1
4.2	Using spreadsheet, solve numerical problems associated with topics covered in various modules	25	7,8
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 15 marks** Quiz, Assignments **Practical: 15 marks** Lab involvement, Practical Record,Viva voce ***Summative assessment*** **Theory: 10 marks** Written tests
	B. End Semester Evaluation (ESE) **Theory : 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12). **Practical: 35 marks** Problem solving skills: 30 marks Record: 5 marks

References:

1. Gupta, S. C. and Kapoor, V. K. (2020). Fundamentals of Mathematical Statistics, 12th Edition, Sultan Chand and Sons.
2. Gupta, S.P. (2021). Statistical Methods, 46th Edition, Sultan Chand and Sons: New Delhi.
3. Beverly J. Dretzke. (2008). Statistics with Microsoft Excel, 4th Edition, Pearson.

Suggested Readings:

1. Medhi, J. (2006). Statistical Methods, 2nd Edition, New Age International Publishers.
2. Mukhopadhyay, P. (1999). Applied Statistics, New Central Book Agency Private Limited, Kolkata.

Mahatma Gandhi University

Kottayam

Programme						
Course Name	**Statistical Data Collection**					
Type of Course	**MDC**					
Course Code	**UC1MDCSTA100**					
Course Level	**100**					
Course Summary	To acquire the basic knowledge of statistical data collection and basic principles of experimental design. Also students will be able to design experiments incorporating the principles of experimentation and perform basic exploratory data analysis.					
Semester	**1**	**Credits**			**3**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		2		**1**		**60**
Pre-requisites						

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Understand the characteristics of scientific research.	U	1
2	Understand different sampling schemes.	U	1
3	Describe concepts of data, methods of data collection and levels of measurements.	U	1
4	Apply a proper sampling scheme for the concerned problem.	A	2
5	Develop a research problem and formulate the research hypothesis.	C	2
6	Prepare a questionnaire for a problem.	C	2
7	Design experiments and perform basic exploratory data analysis.	A, An	2

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

COURSE CONTENT

Content for Classroom Transaction (Units)

	Course Description	Hours	CO. No
Module 1	**Scientific Research**	**15**	
1.1	Characteristics of scientific research: Qualitative studies, quantitative studies, longitudinal studies, experimental studies and survey studies.	2	1
1.2	Stating hypothesis or research question, concepts and constructs, units of analysis and characteristics of interest, independent and dependent variables, extraneous or confounding variables.	4	1
1.3	Concepts of statistical population and sample, complete enumeration and sampling, probability and non-probability sampling, simple random sampling and stratified random sampling (Outline only).	4	2
1.4	Primary and secondary data, different types of data: quantitative and qualitative data, continuous and discrete data, time series and cross-sectional data, methods of collection of primary data, sources of secondary data.	5	3
Module 2	**Design of Experiments**	**15**	
2.1	Levels of measurement: Nominal, ordinal, interval and ratio.	2	3
2.2	Designing a questionnaire.	2	4
2.3	Planning of experiments: Basic principles of experimental design, uniformity trials.	5	7
2.4	Completely Randomised Design (CRD), Randomised Block Design (RBD), Latin Square Design (LSD), Factorial	6	6
	experiments, Split plot experiments.(**Only the concepts and outline of the designs are needed)**		
Module 3	**Practical problems from the above topics.**	**30**	
	Develop a research problem from the relevant disciplines of the students. Formulate research hypotheses. Identify the target population, determine the variables of interest and decide the proper sampling scheme.	10	4,5,6,7
	Prepare a questionnaire for the problem in (1), collect data using it and basic Exploratory Data Analysis (EDA) using any statistical software.	10	4,5,6,7
	If experimentation is needed, design experiments incorporating the principles of experimentation and perform basic EDA using the data.	10	4,5,6,7
Module 4	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** **Formative assessment** **Theory: 10 marks** Quiz,Assignment **Practical: 15 marks** Lab involvement, Practical book, Viva voce **Summative assessment** **Theory: 5 Marks** written test

	B. End Semester Evaluation (ESE) **Theory : 35 marks** i) MCQ : 10 questions (10*1=10). ii) Short essay type questions: Answer any 3 questions out of 5 (3*5=15). iii) Essay type questions: Answer any 1 question out of 2 (1*10=10). **Practical: 35 marks** Problem solving skills: 35 marks

References:

1. Gupta, S.C. and Kapoor, V.K. (2007). Fundamentals of Applied Statistics, Sultan Chand and Sons.
2. Gupta, S.P. (2021). Statistical Methods, 46th Edition, Sultan Chand and Sons: New Delhi.
3. Kothari, C.R. (2014). Research methodology, Second revised edition, New Age International publishers.

Suggested Readings:

1. Mukhopadhyay, P. (2009). Theory and Methods of Survey Sampling, Second Edition, PHI Learning (P) Ltd.
2. Das, M.N. and Giri, N.C. (1994). Design and analysis of experiments, Wiley Eastern Ltd.
3. Rangaswamy, R. (2010). A textbook on Agricultural Statistics, New Age International publishers

Mahatma Gandhi University

Kottayam

Programme						
Course Name	**Data Analysis using Libre Calc**					
Type of Course	**MDC**					
Course Code	**UC1MDCSTA101**					
Course Level	**100**					
Course Summary	This comprehensive course covers fundamental spreadsheet operations, including basic calculations, data entry, and manipulation using mathematical operators and built-in functions. Students will learn data visualisation techniques using Google Looker Studio, as well as how to categorise data types and perform basic statistical analysis, including mean, median, mode, and hypothesis testing. Through hands-on exercises, participants will gain proficiency in generating frequency and cross tables, conducting t-tests and chi-square tests, and analysing correlations using both parametric and non-parametric methods. By the end of the course, students will have the skills to effectively manage and analyse data, making informed decisions based on statistical insights. **Upon completion of this course student acquires NOS1,2,3,5 of Data Analysis Associate available in NQR**						
Semester	**1**	**Credits**			**3**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		2		**1**		**60**
Pre-requisites						

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Apply basic Mathematical formula in Spreadsheet	A	1

2	Analyse the information in the data using visual tools	An	2
3	Analyse the data using desc riptive statistics tools in spreadsheet	An	2
4	Perform basic inference tools in the data and arrive at conclusions about populations using spreadsheet	An	1
Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)			

COURSE CONTENT

Content for Classroom transaction (Sub-units

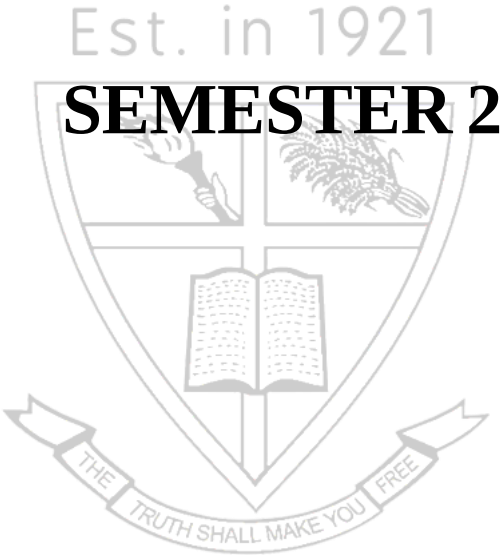
Module 1	**Course Description**	**Hours**	**CO No.**
	Introduction to Spreadsheets, Data Visualization and random number generation	**15**	
1.1	Entering data into cells, importing data from other formats and exporting data into other formats, Introduction to Google spreadsheets	2	**1**
1.2	Using mathematical operators (+, -, *, /), Using built-in functions (SUM, AVERAGE, MIN, MAX)	2	**1**
1.3	Understanding cell references (relative vs. Absolute), Sorting data alphabetically or numerically or in a custom order, Filtering data based on specific criteria,	2	**1**
1.4	Removing duplicates from datasets, Formatting Spreadsheets, Data validations, conditional formatting, conditional statements and vlookup and hlookup operators	2	**1**
1.5	Types of Data based on information – Ordinal, nominal, interval, ratio scale, Introduction to various charts- histogram, Bar chart, line chart, bar chart, pie chart	2	**2**
1.6	Random number generation – uniform random numbers, generation of binomial, bernoulli, other custom discrete random numbers, exponential and Erlangian random numbers	3	**3**

1.7	Generating normal and beta random numbers using Acceptance rejection sampling	2	**3**
Module 2	**Descriptive and Inferential Statistics**	**15**	
2.1	Various Measures of central tendency and measures of dispersion and contexts of their usage	3	**3**
2.2	Pivot tables and interpretations	2	**4**
2.3	T-test (one sample, paired sample t-test, independent sample t-test) – Interpreting results, one way and two way ANOVA	3	**4**
2.4	Assumptions of t-test and verifying the assumptions	2	**4**
2.5	Chi-square test for independence, Spearman and Peasron correlation in Spreadsheet directly and without using function and interpreting results	3	**4**
2.6	Non-parametric analogues of t-test, one sample ANOVA	2	**4**
Module 3	**Practicals**	**30**	
3.1	Formatting data using spreadsheets incorporating all methods in module 1	5	**1**
3.2	Generating Random numbers from exponential, binomial, normal, beta distributions using theory discussed in module 1	6	**2**
3.3	Creating a dashboard using google vlooker and apply it in 5 real data sets	6	**2**
3.4	Applying various Data visualisation in 20 real time data and 5 generated datasets	5	**2**
3.5	Analysing 10 real data sets of size minimum 30 based on the module 2 (All descriptive statistics and test procedures should be used	8	**3, 4**
Module 4	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** **Formative assessment** **Theory: 10 marks** Quiz, Assignment **Practical: 15 marks** Lab involvement, Practical book, Viva voce **Summative assessment** **Theory: 5 Marks** Written test
	B. End Semester Evaluation (ESE) **Theory : 35 marks** i) MCQ : 10 questions (10*1=10). ii) Short essay type questions: Answer any 3 questions out of 5 (3*5=15). iii) Essay type questions: Answer any 1 question out of 2 (1*10=10). **Practical: 35 marks** Problem solving skills: 35 marks

References

1. Sam O A(2023), Excel Mastering Data Analysis, Visualization, and Automation for Success with Microsoft 365, SA Press,
2. D Narayana, Sharad Ranjan, and Nupur Tyagi (2023), Basic Computational Techniques For Data Analysis, Routledge
3. David Ray Anderson, Dennis J. Sweeney, Thomas Arthur Williams (2011), Essentials of Statistics for Business and Economics, West Publishing Company
4. Sheldon M. Ross(2006), Simulation, Elsevier
5. Nussbaumer Knaflic, Cole(2015), Storytelling With Data: A Data Visualization Guide For Business Professionals, Wiley.



SEMESTER 2

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Introduction to Statistical Modelling**					
Type of Course	**DSC A**					
Course Code	**UC2DSCSTA100**					
Course Level	**100**					
Course Summary	To acquire the basic knowledge of theory of random variables, various probability functions and their applications. Also spreadsheet functions are used to solve numerical problems associated with the topics discussed.					
Semester	**2**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3		**1**		**75**
Pre-requisites						

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Examine major components of random variable theory and distribution theory.	U	1
2	Evaluate Mathematical Expectation.	S	2
3	Analyse various probability distributions.	An	2
4	Evaluate fitting procedures of probability distributions and numerical problems.	S	2

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

COURSE CONTENT
Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module1	**Random Variable Theory**	**10**	
1.1	Describe univariate random variables in discrete and continuous cases.	3	1
1.2	Demonstrate probability mass function and its properties, probability density function (concept only), distribution function of a random variable: Definition and properties.	4	1
1.3	Demonstrate functions of random variable, transformations of random variable (univariate – one to one functions only).	3	1
Module 2	**Mathematical Expectation**	**15**	
2.1	Demonstrate mathematical expectation, its properties and simple problems (univariate only).	7	2
2.2	Describe Arithmetic Mean (AM), Geometric Mean (GM), Harmonic Mean (HM), Mean Deviation and Variance in terms of expectation and evaluate simple problems on AM and Variance.	5	2
2.3	Describe generating functions: Moment generating function, characteristic function and their properties.	3	2
Module 3	**Discrete and Continuous Distributions**	**20**	
3.1	Describe uniform distribution and Bernoulli distribution, explain binomial distribution and its properties (without proof), simple problems.	6	3
3.2	Explain Poisson distribution and its properties (without proof), simple problems. Explain geometric distribution, mean and variance (without proof).	6	3
3.3	Explain continuous uniform distribution, its mean and variance. Explain exponential distribution, mean and variance (without proof).	2	3
3.4	Explain normal distribution and its properties. Discuss standard normal distribution and use of standard normal tables, problems.	6	3
Module 4	**Spreadsheet for Statistical Computing** (A practical record with minimum 10 problems has to be submitted).	**30**	
4.1	Computation of Binomial, Poisson and Normal probabilities.	10	4

4.2	Fitting of Binomial, Poisson and Normal Distributions.	10	4
4.3	Random number generation from Uniform, Binomial, Poisson and Normal Distributions.	10	4
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 15 marks** Quiz, Assignments **Practical: 15 marks** Lab involvement, Practical Record,Viva voce ***Summative assessment*** **Theory: 10 marks** Written tests
	B. End Semester Evaluation (ESE) **Theory : 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12). **Practical: 35 marks** Problem solving skills: 30 marks Record: 5 marks

References:

1. Mukhopadhaya, P. (1996). Mathematical Statistics. New Central Book Agency (P) Ltd., Calcutta.
2. Beverly J. Dretzke. (2008). Statistics with Microsoft Excel, Fourth Edition, Pearson.
3. Gupta, S.C. and Kapoor, V.K. (2002). Fundamentals of Mathematical Statistics. Sulthan Chand, New Delhi.

Suggested Readings:

1. Bhat, B.R., Venkata Ramana, T. and Rao Madhava, K.S. (1977). Statistics*:* A Beginners Text Vol-2, New Age International (P) Ltd., New Delhi.
2. Goon, A. M., Gupta, N.K., and Das Gupta, B. (1999). Fundamentals of Statistics-Vol.2. World Press, Kolkatha.
3. Rohatgi, V.K. and Saleh, A.M.E. (2001). An Introduction to Probability and Statistics. 2nd Edition. John Wiley & Sons, Inc, New York.
4. Wilks, S.S. (1964). Mathematical Statistics, John Wiley, New York.

**Remember (K), Understand (U), Apply (A), Analyze (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)*

Mahatma Gandhi University

Kottayam

Programme						
Course Name	**Time Series Methods and Their Applications**					
Type of Course	**MDC**					
Course Code	**UC2MDCSTA100**					
Course Level	**100**					
Course Summary	Introductory R programming, time series analysis and forecasting methods using statistical packages.					
Semester	**2**	**Credits**			**3**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		2		**1**		**60**
Pre-requisites						

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	To critically analyse and summarise time series data.	An	1
2	To familiarise the basic concepts of time series model building and its applications.	S	2
3	Illustrate the time series models with different live data.	I	2
4	Apply R built in functions to solve numerical problems.	A	2

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

Module 1	Course Description	Hours	CO No.
	Exploratory Time Series Data Analysis and Forecasting	**15**	
1.1	Introduction to time series, real world examples and applications of time series for social science in GDP, inflation etc.	3	1
1.2	Time series plots, interpretations using different tools, sampling frequency, basic assumption of time series, components of time series.	4	1

1.3	Trend spotting: Linear, rapid growth, periodic, examples of increasing variance trends over time, sample transformations.	3	1,2
1.4	White noise model, simulations of white noise models in R and examples.	3	2
1.5	Random walk model (simple examples of non-stationary model), stationary processes.	2	2
Module 2	**Correlation Analysis**	**15**	
2.1	Scatter plots, covariance and correlations.	3	3
2.2	Covariance and correlation: Log returns, autocorrelation.	3	1,3
2.3	Auto regressive model estimation and forecasting.	5	1,2,3
2.4	Introduce simulation and live data explanations with AR model.	4	2,3
Module 3	**Illustrate the concepts in Module 1 and 2 Using R.** (A practical record with minimum 5 problems has to be submitted)	**30**	4
Module 4	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** **Formative assessment** **Theory: 10 marks** Quiz, Assignment **Practical: 15 marks** Lab involvement, Practical record, Viva voce **Summative assessment** **Theory: 5 Marks** Written test

	B. End Semester Evaluation (ESE) **Theory : 35 marks** i) MCQ : 10 questions (10*1=10). ii) Short essay type questions: Answer any 3 questions out of 5 (3*5=15). iii) Essay type questions: Answer any 1 question out of 2 (1*10=10).
	Practical: 35 marks Problem solving skills: 35 marks

References:

1. Cowpertwait, Paul, S.P., and Andrew V. Metcalfe. (2009). Introductory time series with R. Springer Science & Business Media.
2. Box, George EP, et al. (2015). Time series analysis: Forecasting and Control. John Wiley & Sons.

Suggested Readings:

1. Chatfield, Christopher. (2013). The analysis of time series: Theory and Practice. Springer.
2. Chan, Kung-Sik, and Jonathan D. Cryer. (2008). Time series analysis with applications in R. springer publication.
3. Chatfield, Chris, and Haipeng Xing. (2019).The analysis of time series: An introduction with R. CRC press.

Mahatma Gandhi University

Kottayam

Programme						
Course Name	**Data Analysis Using JAMOVI and Introduction to R**					
Type of Course	**MDC**					
Course Code	**UC2MDCSTA101**					
Course Level	**100**					
Course Summary	This course provides comprehensive training in statistical analysis using JAMOVI and Introduces R programming. Students will learn to analyse real data sets, conduct various statistical tests, and apply regression analysis using JAMOVI, enhancing their proficiency in statistical analysis for research and data-driven decision-making. **Upon completion of this course student acquires NOS1,2,3,5 of Data Analysis Associate available in NQR**					
Semester	**2**	**Credits**			**3**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		2		**1**		**60**
Pre-requisites						

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Analyse the information in the data using visual tools from JAMOVI	An	1
2	Analyse the data using descriptive statistics tools in JAMOVI	An	1

3	Perform basic inference tools in the data and arrive at conclusions about populations using JAMOVI	An	1
4	Apply loops and conditional statements in R	A	2
***Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)**			

	Course Description	**Hours**	**CO No.**
Module 1	**Correlation and Regression Analysis**	**10**	
1.1	Types of Data-Ordinal Interval, ratio, measures of central tendency – mean, median, mode, measures of dispersion – Quartile Deviation, variance, standard deviation.	2	2
1.2	Introduction to correlation and regression- simple and multiple.	3	3
1.3	Verifying the assumptions of Linear Regressions.	2	3
1.4	Logistic Regression and interpreting results.	3	3
Module 2	**Introduction to R Programing**	**20**	
2.1	Introduction to R and arithmetic operations in R	4	4
2.2	IF THEN statements and FOR, WHILE loops in R and basic Programs in R	6	4
2.3	Data Frames, subsetting, filtering and other data manipulations	6	4
2.4	R Markdown	4	4
Module 3	**Practicals using JAMOVI and Basic Operations in R**	**30**	
3.1	Entering data into JAMOVI, importing data from other formats to JAMOVI	2	1
3.2	Introduction to various charts- histogram, Bar chart, line chart, bar chart, pie chart	2	2
3.3	Generating various charts using real time data	2	1
3.4	Generating frequency table and cross tables and summary measures using JAMOVI	2	2

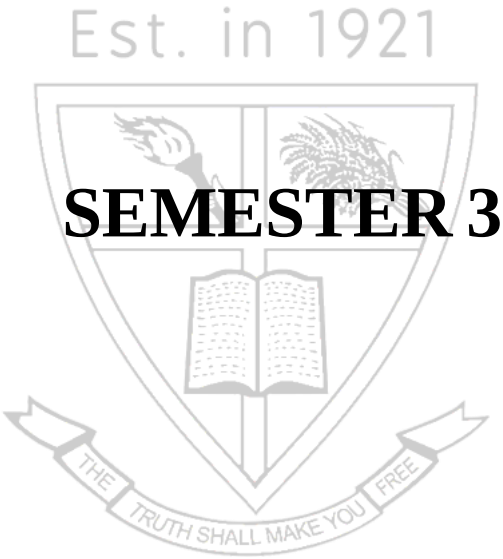
3.5	Scatter diagram and correlation – Pearson and Spearman's Correlation in JAMOVI	2	3
3.6	Regression Analysis in Jamovi and Spreadsheet.	2	3
3.7	t-test (one sample, paired sample t-test, independent sample t-test) – Interpreting results.	2	3
3.8	One way and two way ANOVA	3	3
3.9	Assumptions of t-test and verifying the assumptions	2	3
3.10	Chi-square test for independence	2	3
3.11	Non-parametric analogues of t-test, one sample ANOVA	2	3
3.12	Logistic Regression in JAMOVI	2	3
3.13	Analyse atleast 10 data sets using all the methods in 3.1- 3.12	2	1,2,3
3.14	Practicals of R	3	4
Module 4	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.

Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** **Formative assessment** **Theory: 10 marks** Quiz, Assignment **Practical: 15 marks** Lab involvement, Practical Record, Viva voce. **Summative assessment** **Theory: 5 Marks** Written test
	B. End Semester Evaluation (ESE) **Theory : 35 marks** i) MCQ : 10 questions (10*1=10). ii) Short essay type questions: Answer any 3 questions out of 5 (3*5=15). iii) Essay type questions: Answer any 1 question out of 2 (1*10=10). **Practical: 35 marks** Problem solving skills: 35 marks

References

1. D Narayana, Sharad Ranjan, and Nupur Tyagi (2023), Basic Computational Techniques For Data Analysis, Routledge
2. Navarro DJ and Foxcroft DR (2022). learning statistics with jamovi: a tutorial for psychology students and other beginners. (Version 0.75). DOI: 10.24384/hgc3-7p15r
3. Nussbaumer Knaflic, Cole(2015), Storytelling With Data: A Data Visualization Guide For Business Professionals, Wiley
4. Andy Field, Jeremy Miles, Zoe Field (2012) DISCOVERING STATISTICS USING R, Sage Publications



SEMESTER 3

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Statistical Distributions**					
Type of Course	**DSC A**					
Course Code	**UC3DSCSTA200**					
Course Level	**200**					
Course Summary	Gain foundational knowledge in random variables, explore discrete distributions like Binomial, Poisson, Uniform and Geometric, understand continuous distributions such as Uniform, Exponential, Gamma, Beta (two types), Normal, Lognormal, Cauchy and Laplace distributions and their basic properties. Students will get an idea about sampling distributions and their inter relationships. Spreadsheet is applied for practical applications.					
Semester	**3**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		4				**60**
Pre-requisites	Level 100 knowledge of Statistics					

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Understand various concepts such as probability density functions and cumulative distribution functions etc. of random variables.	U	1
2	Derive various generating functions of random variables such as moment generating functions, characteristic functions etc.	C	2
3	Find out characteristics of random variables like moments from either probability density (mass) functions or the generating functions.	E	1
4	Fitting of Binomial, Poisson and Normal distributions.	A, E & S	2
5	Derivation of the sampling distribution of sample mean and variance for a normal population.	C & S	2
6	Establish relationships between t, F and $\chi 2$ distributions.	A	1

Remember (K), Understand (U), Apply (A), Analyze (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

	Course Description	Hours	CO No.
Module 1	**Discrete Distributions**	**15**	
1.1	Random variables: Discrete random variables, probability mass function, distribution function, change of variables.	3	1
1.2	Definition of mathematical expectation, properties, mean and variance using expectation.	2	1
1.3	Moment generating function (mgf), characteristic function, important properties.	4	2
1.4	Binomial, Poisson, uniform, geometric distributions: Mean, variance, mgf and characteristic functions, lack of memory property of geometric distribution.	3	2
1.5	Fitting of Binomial and Poisson distributions.	3	4
Module 2	**Continuous Distributions**	**15**	
2.1	Continuous random variables, probability density function, distribution function and change of variable.	2	1
2.2	Definition of mathematical expectation, properties, mean and variance using expectation. mgf, characteristic function, properties.	4	1,2
2.3	Uniform, exponential, gamma, beta (two types), Laplace distributions: Mean, variance, mgf and characteristic functions, Cauchy distribution, lack of memory property of exponential distribution.	4	2
2.4	Normal distribution, standard normal distribution, use of standard normal tables for various probability computation, properties of normal distribution, normal distribution as a limiting case of binomial and Poisson under suitable assumptions. Fitting of normal distribution. Lindeberg-Levy central limit theorem (without proof).	4	3, 4
2.5	Lognormal distribution: Definition and properties only (Derivation not required).	1	1
Module 3	**Sampling Distributions**	**15**	
3.1	Derivation of the sampling distribution of sample mean and variance for a normal population, standard errors of sample mean and sample variance.	3	5
3.2	Definition and derivation of pdf of χ^2 with n degrees of freedom, nature of pdf curve for different degrees of freedom, mean, variance, mgf, additive property of χ^2 distribution.	4	5
3.3	Student's t-distribution, derivation of its pdf, nature of probability curve with different degrees of freedom, mean, variance.	3	5

3.4	Snedecor's F-distribution: Derivation of pdf, nature of pdf curve with different degrees of freedom, mean, variance. Distribution of 1/F .	3	5
3.5	Relationship between t, F and χ^2 distributions.	2	6
Module 4	**Statistical Analysis Using Spreadsheet** (A record with minimum 10 problems has to be submitted).	**15**	
	1. Fitting of binomial distribution for given n and p. 2. Fitting of binomial distribution after computing mean and variance. 3. Fitting of Poisson distribution for given value of λ. 4. Fitting of Poisson distribution after computing mean. 5. Problems based on binomial distribution.	15	4
	6. Problems based on Poisson distribution 7. Fitting of normal distribution when parameters are given and not given. 8. Problems based on Normal distribution. 9. Random number generation from Binomial distribution, Poisson distribution and their histograms. 10. Random number generation from Normal distribution and its histogram.		
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 20 Marks** Quiz, Assignments, Seminar ***Summative assessment*** **Theory: 10 marks** Two written tests: 10 marks (5 marks each)

	B. Semester End Examination: (Theory based examination) **Total:70 marks** i) Short answer type questions: Answer any 10 questions out of 12 (10*3=30). ii) Short essay type questions: Answer any 4 questions out of 6 (4*7=28). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12).

References:

1. Gupta, S.C. and Kapoor, V.K. (2020). Fundamentals of Mathematical Statistics, 12th Edition. Sultan Chand & Sons, New Delhi.
2. Mood, A.M., Graybill, F.A. and Boes, D.C. (2007). Introduction to the Theory of Statistics, 3rd Edition. (Reprint). Tata McGraw-Hill Pub. Co. Ltd.
3. Beverly J. Dretzke. (2008). Statistics with Microsoft Excel, 4th Edition, Pearson.

Suggested Readings:

1. Hogg, R.V., McKean, J.W. and Craig, A.T. (2014). Introduction to Mathematical Statistics, 7th Edition, Pearson Education Publication.
2. Rohatgi, V.K. and Saleh, A.K.MD.E. (2015).An Introduction to Probability and Statistics, 3rd Edition, John Wiley & Sons Inc.
3. Johnson, N.L., Kotz, S. and Balakrishnan, N (1994).Continuous Univariate Distributions, Vol.I, 2nd Edition. John Wiley, New York.
4. Johnson, N.L., Kemp, A.W. and Kotz, S. (2005).Univariate Discrete Distributions, 3rd Edition , John Wiley, New York.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Analytical Tools for Multivariate Analysis.**					
Type of Course	**DSC A**					
Course Code	**UC3DSCSTA201**					
Course Level	**200**					
Course Summary	Students will comprehend real vectors, orthogonality and Gram-Schmidt orthogonalization process. They will also grasp the concepts of matrices, determinants, G-inverse, quadratic forms and characteristic roots. Additionally, students will gain the ability to apply this knowledge practically using R/Python software.					
Semester	**3**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3		**1**		**75**
Pre-requisites	Level 100 knowledge of Statistics.					

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Interpret vector space, linear dependence and independence of vectors, spanning vector space, projection of vectors.	U	1
2	Evaluate matrices, trace, determinant, adjoint and inverse of a matrix, product of determinants, related results.	An	1
3	Solve theory of equations, generalised inverse of matrix, quadratic forms, linear transformations.	A	2
4	Obtain the characteristic roots, characteristic vectors, and different related methods.	E	2
5	Find inner product and norm.	An	2
6	Applications of linear algebra in Statistics as the foundation to the courses like Multivariate Analysis and Linear Models.	C	3

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module 1	**Real vectors**	**15**	
1.1	Real vectors (generalisation of coordinates), angle and norm of vectors, orthogonality and Gram-Schmidt orthogonalization process, Axiomatic approach and examples.	6	1
1.2	Subspaces, intersection and sum of subspaces, span of a set, linear dependence and independence, dimension and basis, dimension theorem.	5	1
1.3	Direct sum and complement subspace, orthogonal projection of a vector.	4	1
Module 2	**Matrices and Determinants**	**15**	
2.1	Algebra of matrices, theorems related to triangular, symmetric and skew symmetric matrices, idempotent matrices, orthogonal matrices and their properties.	3	2
2.2	Trace of a matrix, determinant, singular and non-singular matrices, adjoint and inverse of a matrix and related properties.	3	2
2.3	Product of determinants, rank of a matrix, row-rank, column-rank, standard theorems on ranks, rank of the sum and the product of two matrices. Rank factorization and Sylvester's Inequality.	5	2
2.4	Partitioning of matrices, determinant and inverse of partitioned matrices, elementary transformations, Echelon form and Normal form.	4	2
Module 3	**G-inverse , Quadratic forms and Characteristic roots**	**15**	
3.1	System of homogeneous and non-homogeneous linear equations, Cramer's rule, projection matrix and its applications to least square method.	3	3
3.2	Generalised inverse, Moore-Penrose inverse, quadratic forms: Classification and canonical reduction, linear transformations.	3	3
3.3	Characteristic roots and characteristic vectors, properties of characteristic roots (symmetric and general matrices).	3	4
3.4	Diagonalization of matrices, spectral decomposition, and singular value decomposition, power method, Cayley- Hamilton theorem, extrema of quadratic forms.	3	4
3.5	General concepts of inner product and norm. Applications of Linear Algebra in Statistics.	**3**	5,6

Module 4	**Practicals Using R/Python** (A practical record with minimum 10 problems has to be submitted).	**30**	
	1. Linear independence and dependence. 2. Orthogonality and Gram-Schmidt orthogonalization process. 3. Basis and Dimension 4. Basis of sum intersection and complement of subspaces. 5. Projection of vectors on a subspace 6. Determinant of a matrix. 7. Inverse of a matrix. 8. Rank and rank factorization of matrices. 9. Elementary transformations. 10. Solutions of system of linear equations. 11. Finding G-inverse of a matrix 12. Problems on quadratic forms 13. Problems related to characteristic roots and vectors. 14. Power method of finding characteristic roots.	30	3
	15. Problems related to linear transformations.		
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.

Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 15 marks** Quiz, Assignments **Practical: 15 marks** Lab involvement, Practical Record,Viva voce ***Summative assessment*** **Theory: 10 marks** Written tests
	B. End Semester Evaluation (ESE) **Theory : 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12). **Practical: 35 marks** Problem solving skills: 30 marks
	Record: 5 marks

References:

1. Shanti Narayan and Mittal, P.K. (2007). A Textbook of Matrices, S Chand & Co Ltd.
2. Mathai, A.M. (1997). Jacobians of Matrix Transformation and Functions of Matrix Arguments, World Scientific Publishing Company.
3. Lipschutz, S. and Lipson, M. (2017). Schaum's Outline of Linear Algebra , 3rd Edition, McGraw Hill Education.
4. Nick Fieller. (2021). Basics of Matrix Algebra for Statistics with R, 1st Edition, Chapman & Hall.
5. Archana Jadhav and Nandani Sakhare. (2018). Linear Algebra Using Python, Himalaya Publishing House.

Suggested Readings:

1. Hadley G.(2020). Linear Algebra, Narosa Publishing House.
2. Rao A.R. and Bhimasankaram P. (2000): Linear Algebra, 2nd Edition,

Hindustan Book Agency.
3. Searle S.R. and Khuri, A.I. (2017). Matrix Algebra Useful for Statistics 2nd Edition, Wiley.
4. Rao, C.R.(2009). Linear Statistical Inference & its Applications 2nd Edition , Wiley.
5. Strang G.(2023). Introduction to Linear Algebra 6th Edition, Wellesley-Cambridge Press, U.S.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Statistical Techniques for Data Science and Machine Learning (Data Analytics Specialization)**					
Type of Course	**DSE**					
Course Code	**UC3DSESTA200**					
Course Level	**200**					
Course Summary	Students will be able to navigate the realms of inferential statistics, non-parametric tests, ANOVA, machine learning and data science.					
Semester	**3**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3		**1**		**75**
Pre-requisites	Level 100 knowledge of Statistics					

EXPECTED COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains	Program Outcome
1	Operate parametric tests.	A	2
2	Relate non parametric tests.	An	2
3	Apply Machine learning tools in Statistics.	A	2
4	Understand the basics of Data science.	U	1
5	Conduct data analysis using R/Python.	E	2
Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)			

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module1	**Inferential Statistics, Non-Parametric Tests and ANOVA**	**15**	1
1.1	Introduction, sampling distribution: Normal and t.	3	1

1.2	Hypothesis testing: z test(One sample), t test(One sample).	3	1
1.3	Introduction, Chi-square test for goodness of fit.	3	2
1.4	Chi -square test for independence.	3	2
1.5	F test, ANOVA (one way classification)	3	2
Module 2	**Introduction to Machine Learning and its Applications**	**15**	
2.1	Introduction: Techniques of Machine Learning: Supervised learning, unsupervised learning, reinforcement learning.	4	3
2.2	Applications of machine learning, statistical tools for machine learning.	3	3
2.3	Simple linear regression (concepts and simple applications).	2	3
2.4	Multiple linear regression (concepts and simple applications).	3	3
2.5	Logistic regression (concepts and simple applications).	3	3
Module 3	**Introduction to Data Science**	**15**	
3.1	Introduction, definition.	1	4
3.2	Data Science in various fields, Examples.	2	4
3.3	Impact of data science.	2	4
3.4	Understating data: Introduction, types of data, numeric, categorical, graphical, high dimensional data.	3	4
3.5	Classification of digital data: Structured, Semi-structured, Unstructured, Example, Applications.	3	4
3.6	Sources of data: Time series data, transactional data, biological data, spatial data, social network data.	2	4
3.7	Data evolution, introduction to big data.	2	4
Module 4	**Data analysis Using R /Python** (A practical record with minimum 5 problems has to be submitted).	**30**	
4.1	Categorical data analysis.	8	5
4.2	Correlation and Regression.	12	5
4.3	Testing, ANOVA (one-way classification).	10	5
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.

Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 15 marks** Quiz, Assignments **Practical: 15 marks** Lab involvement, Practical Record,Viva voce ***Summative assessment*** **Theory: 10 marks** Written tests
	B. End Semester Evaluation (ESE) **Theory : 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12). **Practical: 35 marks** Problem solving skills: 30 marks Record: 5 marks

References:

1. Gupta, S.P. (2021). Statistical Methods. Sultan Chandand Sons: NewDelhi.
2. Gupta, S.C. and Kapoor, V.K. (2020). Fundamentals of Mathematical Statistics, Sultan Chand and Sons.
3. Sudha G. Purohit, Sharad D. Gore and Shailaja R. Deshmukh. (2019). Statistics Using R, 2nd Edition, Narosa Publishing House.

Suggested Readings:

1. Tilman M. Davies. (2016).The Book of R, A First Course in R Programming and Statistics, NoStarch Press.
2. Dirk P. Kroese, Zdravko Botev, Thomas Taimre and Radislav Vaisman ·(2019). Data Science and Machine Learning, Mathematical and Statistical Methods, CRC Press.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Statistical Computing Using R**					
Type of Course	**DSE**					
Course Code	**UC3DSESTA201**					
Course Level	**200**					
Course Summary	Through this course, students will comprehend R software, adept at conducting descriptive statistics, handling probability distributions and gaining insights into correlation and regression.					
Semester	**3**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3		**1**		**75**
Pre-requisites	Level 100 knowledge of Statistics					

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Understand various methods of data input and commands in R software.	U	1
2	Manipulate data using various commands and functions in R.	A	1
3	Analyse data using R software.	An	2
4	Apply various R graphics.	A	2
5	Evaluate various measures of central tendency, dispersion, skewness and kurtosis.	E	1

6	Fitting probability distributions using R software.	C	2
7	Generate random numbers from important probability distributions.	C	2
8	Develop correlation and regression analysis using R software.	A	3
Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)			

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module 1	**Introduction to Statistical Software R and Descriptive Statistics.**	**15**	
1.1	Introduction to statistical software R, data objects in R.	2	1
1.2	Manipulating vectors, matrices, lists, importing of files, data frame, Controlling Loops : For, repeat, while, if, if else etc., functions in R.	3	2
1.3	Diagrammatic and graphical representation of data: Bar diagram, histogram, pie diagram, box plot, Q-Q plot, the plot function and curve function, stem and leaf plot, scatter plot, Plot options: The plot function and curve function, multiple plots in a single graphic window.	4	4
1.4	Frequency table, measures of central tendency and dispersion.	2	5
1.5	Measures of skewness and kurtosis.	2	5
1.6	Selection of representative samples.	2	5
Module 2	**Probability Distributions Using R**	**15**	
2.1	Probability distributions, some discrete distributions: Bernoulli, binomial, Poisson, geometric and uniform, plotting of these distributions, fitting of discrete distributions.	4	6
2.2	Continuous probability distributions, some continuous distributions (Normal, exponential, rectangular), plotting of these distributions, fitting of normal distribution.	4	6
2.3	Methods for generating random numbers: Introduction, random number generation-discrete and continuous distributions in R.	4	7
2.4	Quantiles, inverse transform method, and transformation methods.	3	7
Module 3	**Correlation and Regression Analysis**	**15**	
3.1	Correlation, inference procedures for correlation coefficient, linear regression, coefficient of determination.	7	8

3.2	Simple regression, logistic regression, inference procedures for simple linear model.	8	8
Module 4	**Practical Using R** (A practical record with minimum 10 problems has to be submitted)	**30**	8
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** **Formative assessment** **Theory: 15 Marks** Quiz, Assignments **Practical: 15 Marks** Lab involvement, Practical Record, Viva voce **Summative assessment** **Theory: 10 Marks** Two written tests.
	B. End Semester Evaluation (ESE) **Theory : 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12). **Practical: 35 marks** Problem solving skills: 30 marks Record: 5 marks

References:

1. Purohit, S.G, Gore, S.D and Deshmukh, S.R. (2015).Statistics Using R, 2nd Edition, Narosa Publishing House.
2. W. N. Venables, D. M. Smith and the R Development Core Team (2009). An Introduction to R, 2nd Edition, Network Theory Limited.

Suggested Readings:

1. Zuur, A.F, Leno, E.N. and Meesters, E.H.W.G. (2009): Use R, Springer.
2. Rizzo, M.L. (2007). Statistical Computing with R, Chapman and Hall/CRC.
3. Dalgaard, P. (2008). Introductory Statistics with R, Springer.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Vital Statistics and Index Numbers**					
Type of Course	**DSE**					
Course Code	**UC3DSESTA202**					
Course Level	**200**					
Course Summary	By combining theoretical knowledge of vital statistics, mortality and fertility measurement, population growth and index numbers with practical applications using spreadsheets. This course equips students with the skills and understanding necessary for proficient demographic analysis and decision-making in various fields.					
Semester	**3**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3		**1**		**75**
Pre-requisites	Level 100 knowledge of Statistics.					

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Understand the sources of vital statistics including census, registration, adhoc survey and hospital records.	U	1
2	Determine the measurement of mortality including Crude Death Rate, Specific Death Rate, Infant Mortality Rate and Standardised Death Rate.	U,E	2
3	Understand complete life tables and its characteristics.	U,A	1
4	Understand abridged life tables and its characteristics.	U,A	1
5	Determine the measurement of fertility including Crude Birth Rate, General Fertility Rate, Age Specific Fertility Rate, Total Fertility Rate, Gross Reproduction Rate and Net Reproduction Rate.	U,E	2
6	Obtain the measurement of population growth including Crude rates of natural increase, Pearl's vital index, Gross Reproduction Rate and	An, E	2 & 3

	Net Reproduction Rate.		
7	Understand the concepts of index numbers including price, quantity and value indices.	U	1
8	Explain the tests for index numbers, various formulae and their comparisons.	U,A & E	2
**** Understand (U), Apply (A), Analyse (An), Evaluate (E)***			

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	**Hours**	**CO No.**
Module 1	**Sources of Vital Statistics**	**15**	
1.1	Introduction and sources of collecting data on vital statistics.	6	1
1.2	Census, registration, adhoc surveys, hospital records.	9	1
Module 2	**Measurement of Mortality, Fertility and Population Growth.**	**15**	
2.1	Crude Death Rate (CDR) and Specific Death Rate (SDR), Standardised Death Rates and Infant Mortality Rate (IMR).	3	2
2.2	Complete life tables and its characteristics, Abridged life tables and its characteristics.	4	3
2.3	Crude Birth Rate, General Fertility Rate, Age-Specific Fertility Rate, Total Fertility Rate.	4	5
2.4	Crude rates of natural increase and Pearl's Vital Index,Gross Reproduction Rate (GRR) and Net Reproduction Rate (NRR).	4	6
Module 3	**Index Numbers**	**15**	
3.1	Price, Quantity and Value indices.	1	7
3.2	Construction, uses and limitations of index number.	3	7
3.3	Tests for index numbers, various formulae, and their comparisons.	4	8
3.4	Chain-index numbers.	3	7
3.5	Formulae and uses of some important indices: Consumer Price Index, wholesale price index and index of industrial production.	4	7
Module 4	**Practical Using Spreadsheet** (A practical record with minimum 10 problems has to be submitted.)	**30**	
	1. Calculate CDR and ASDR. 2. Calculate STDR by direct method. 3. Calculate STDR by indirect method.	8	2
	4. Find the missing values in the Life Table.	4	3

	5. Calculate CBR and GFR. 6. Calculate age-specific fertility rate and Total Fertility rate. 7. Calculate Gross Reproduction Rate and Net Reproduction Rate.	12	5
	8. Calculate various types of weighted index numbers. 9. Check whether the index numbers satisfy the factor reversal test and time reversal test. 10. Calculate the consumer price index.	6	7
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 15 marks** Quiz, Assignments **Practical: 15 marks** Lab involvement, Practical Record, Viva voce ***Summative assessment*** **Theory: 10 marks** Written tests
	B. End Semester Evaluation (ESE) **Theory : 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12). **Practical: 35 marks**

	Problem solving skills: 30 marks Record: 5 marks

References:

1. Gupta, S.C. and. Kapoor, V.K. (2018). Fundamentals of Applied Statistics, Sultan Chand & Co. New Delhi.
2. Srivastava, O.S. (1983). A Text Book of Demography, Vikas Publishing House, New Delhi.
3. Parimal Mukhopadhyay. (2005). Applied Statistics. Books & Allied (p) Ltd.

Suggested Readings:

1. Goon, A.M. Gupta, M.K. and Das Gupta, B. (2016): Fundamentals of Statistics, Vol. II, World press, Calcutta.
2. Newsholme, A. (2021). The Elements of Vital Statistics, Routledge, Taylor & Francis Group.
3. Keyfitz,N, and Beekman, J.A.(2010), Demography through Problems, 1st Edition, Springer- Verlag.
4. Jhingan, M.L., Bhatt, B.K. and Desai, J.N.(2016). Demography, 3rdEdition, Vrinda Publications (P) Ltd, Delhi.
5. Benjamin B (1960).Elements of Vital Statistics, Quadrangle Books.
6. Whipple, G.C.(2022).Vital Statistics: An Introduction to the Science of Demography, Legare Street Press.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Data Analysis in Inferential Statistics Using R/Python**					
Type of Course	**DSC B**					
Course Code	**UC3DSCSTA202**					
Course Level	**200**					
Course Summary	This course covers key concepts in Statistics including sampling distribution, estimation of parameters, testing of hypothesis and non-parametric tests. Emphasis is placed on practical applications using R or Python.					
Semester	**3**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3		**1**		**75**
Pre-requisites	Level 100 knowledge of Statistics.					

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains	Program Outcome
1	Understand different Sampling Distributions.	U	1
2	Describe estimation and methods.	U	1
3	Relate different parametric tests in testing the hypothesis.	An	1
4	Organise different non-parametric tests in testing the hypothesis.	An	1
5	Conduct data analysis using R/Python.	E	2
Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)			

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module 1	**Sampling Distributions**	**15**	
1.1	Statistic, parameter.	2	1
1.2	Distribution of sample mean and variance.	2	1
1.3	Normal distribution, Student's t-distribution.	5	1
1.4	Chi- square distribution, F distribution.	4	1
1.5	Inter-relationship between Normal, t, Chi-square and F distributions.	2	1
Module 2	**Statistical Inference**	**15**	
2.1	Estimation, point estimation and interval estimation.	2	2
2.2	Desirable properties of a good point estimator.	2	2
2.3	Methods of estimation – MLE, Method of moments.	4	2
2.4	Testing of hypothesis: Statistical test, null and alternative hypothesis, types of errors, significance level, power, critical region, p value.	3	3
2.5	Parametric test: Testing of population mean (One sample and two sample) (z test, t-test), testing of population proportion (One sample and two sample), paired t test. ANOVA(one way only).	4	3
Module 3	**Non- Parametric Tests**	**15**	
3.1	Goodness of fit, Chi-Square test(independence of attributes).	4	4
3.2	Sign test, median test.	5	4
3.3	Kruskal Wallis H test, Wilcoxon test.	6	4
Module 4	**Data Analysis using R /Python**	**30**	
4.1	Introduction to R/Python.	6	5
4.2	Categorical data analysis.	6	5
4.3	Correlation and Regression.	8	5
4.4	Testing, ANOVA (one-way classification). (A practical record with minimum 5 problems has to be submitted).	10	5
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.

Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** **Formative assessment** **Theory: 15 Marks** Quiz,Two Assignments (5 marks each) **Practical: 15 Marks** Lab involvement, Practical Record, Viva voce(5 marks each) **Summative assessment** **Theory: 10 Marks** Two written tests: (5 marks each)
	B. End Semester Evaluation (ESE) **Theory : 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12). **Practical: 35 marks** Problem solving skills: 30 marks Record: 5 marks

References:

1. Rohatgi V.K. and Saleh, A.K. Md.E. (2009): An Introduction to Probability andStatistics.2nd Edition. (Reprint)John Wiley andSons.
2. Gupta,S.P. (2021) Statistical Methods. Sultan Chandand Sons: NewDelhi.
3. Gupta,S.C.and Kapoor, V.K.(2020) Fundamentals of Mathematical Statistics, Sultan Chand and Sons.
4. Sudha G Purohit, Sharad D. Gore, Shailaja Deshmukh (2019) Statistics using R, 2nd Edition, Narosa Publishing House.
5. Python for Everybody: Exploring Data Using Python3, ADS 2016.

Suggested Readings:

1. Mood, A.M. Graybill, F.A. and Boes, D.C. (2007) Introduction to the Theory of Statistics,3rd Edition., (Reprint), Tata Mc Graw-Hill Pub. Co.Ltd.
2. John E Freund, Mathematical Statistics, Pearson Edn, NewDelhi
3. Tilman M. Davies. (2016) The Book of R, A First Course in Programming and Statistics, No Starch Press.
4. Python for Data Analysis (2012)WesMc Kinney, O'REILLY.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Statistical Research Techniques using Softwares**					
Type of Course	**DSC B**					
Course Code	**UC3DSCSTA203**					
Course Level	**200**					
Course Summary	This course aims to equip students with a solid foundation in Research Methodology, Statistical Testing and Data Analysis.					
Semester	**3**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3		**1**		**75**
Pre-requisites	Level 100 knowledge of Statistics					

EXPECTED COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains	Program Outcome
1	Understand different research methods in social science.	U	1
2	Understand the statistical testing procedures.	A	2
3	Illustrate the parametric tests.	An	2
4	Describe the non-parametric tests.	An	2
5	Conduct a Social survey and data analysis using R/Python/Spreadsheet.	E	2

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module 1	**Introduction to Research Methodology**	**15**	
1.1	Research design, qualitative and quantitative research.	3	1
1.2	Data collection methods and sampling techniques.	3	1
1.3	Research reporting and communication: Writing Research proposal.	4	1
1.4	Apply research methods to real-world social issues.	5	1
Module 2	**Testing of hypothesis**	**10**	
2.1	Parameter, Statistic.	2	1
2.2	Statistical hypothesis: Simple and composite hypothesis, null and alternative hypothesis.	4	1
2.3	Types of Errors, significance level.	3	1
2.4	p-value, power, testing procedure.	4	1
2.5	Critical region.	2	1
Module 3	**Parametric and Non-parametric Tests**	**20**	
3.1	Large sample test: z test for single mean and equality of two means.	3	2
3.2	Small sample test: t test for single mean and equality of two means, paired t test.	5	3
3.3	ANOVA (one way only).	2	3
3.4	Non- parametric tests: Testing association of attributes using Chi square test.	2	4
3.5	Sign test, Median test, Wilcoxon ranked test-simple problems only.	6	4
3.6	Applications of statistical tests in various fields.	2	4
Module 4	**Data Analysis using R/Spreadsheet/Python** (A practical record with minimum 5 problems has to be submitted).	**30**	
4.1	Conduct a social survey and prepare a project report (Questionnaire, geographical and diagrammatic representation, analysis - Descriptive Statistics).	15	5
4.2	Statistical analysis and interpretation of a social problem by using Spreadsheet/ Python/ R programming.	15	5
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 15 marks** Quiz, Assignments **Practical: 15 marks** Lab involvement, Practical Record,Viva voce ***Summative assessment*** **Theory: 10 marks** Written tests
	B. End Semester Evaluation (ESE) **Theory : 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12).
	Practical: 35 marks Problem solving skills: 30 marks Record: 5 marks

References:

1. Rohatgi V.K. and Saleh, A.K. Md.E. (2009): An Introduction to Probability andStatistics.2nd Edn. (Reprint)John Wiley andSons.
2. Gupta,S.P. (2021) Statistical Methods. Sultan Chandand Sons: NewDelhi.
3. Gupta,S.C.and Kapoor, V.K.(2020) Fundamentals of Mathematical Statistics, Sultan Chand and Sons.
4. Sudha G Purohit, Sharad D. Gore, Shailaja Deshmukh (2019) Statistics using R, 2nd Edition, Narosa Publishing House.
5. Python for Everybody: Exploring Data Using Python3, ADS 2016.
6. Kothari, C. R. (2014)-Research-methodology-2nd-revised Edition New age International publications.

Suggested Readings:

1. Mood, A.M. Graybill, F.A. and Boes, D.C. (2007) Introduction to the Theory of Statistics,3rd Edition, (Reprint), Tata Mc Graw-Hill Pub. Co.Ltd.
2. John E Freund, Mathematical Statistics, Pearson Edition, NewDelhi
3. Tilman M. Davies. (2016) The Book of R, A First Course in Programming and Statistics, No Starch Press.
4. Python for Data Analysis (2012)WesMc Kinney, O'REILLY.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Business Data Analytics**					
Type of Course	**DSC B**					
Course Code	**UC3DSCSTA204**					
Course Level	**200**					
Course Summary	Students will be equipped with a comprehensive set of skills ranging from handling different types of data to apply time series analysis, statistical quality control, optimization techniques and statistical software for effective data analysis.					
Semester	**3**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3		**1**		**75**
Pre-requisites	Level 100 knowledge of Statistics					

EXPECTED COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains*	Program Outcome
1	Understand different types of data and data sources.	U	1
2	Analyze trends in time series.	A	2
3	Implement Statistical quality assurance in business.	An	2
4	Apply optimization techniques in decision-making problems.	An	2
5	Conduct a market survey and data analysis using R/Python/Spreadsheet.	E	2

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module 1	**Introduction to Different types of Data and Time series Analysis**	**15**	
1.1	Data in various fields, example.	2	1
1.2	Understating of data, types of data: numeric, categorical, graphical, high dimensional data. Classification of digital data: Structured, semi-structured, unstructured, example, applications.	3	1
1.3	Sources of data: Time series data, financial data, actuarial data, transactional data, biological data, spatial data, social and network data. Big data. Data Evolution.	3	1
1.4	Components of Time Series. Different Models.	2	2
1.5	Methods of finding components (Only Trend and Seasonal Variation- Simple average method).	3	2
1.6	Forecasting Sales and Profits (Trend Analysis).	2	2
Module 2	**Statistical Quality Assurance**	**15**	
2.1	Quality and Quality Assurance.	1	3
2.2	Methods of Quality Assurance.	1	3
2.3	Introduction to TQM and ISO 9000 standards.	1	3
2.4	Statistical Quality Control.	1	3
2.5	Acceptance Sampling for Attributes.	3	3
2.6	Single Sampling.	1	3
2.7	Double Sampling.	1	3
2.8	Multiple and Sequential Sampling Plans.	2	3
2.9	Control charts : Mean and Range charts.	4	3
Module 3	**Optimization Techniques**	**15**	
3.1	Decision Theory.	3	4
3.2	Decision making under uncertainty.	4	4
3.3	Decision making under risks.	4	4
3.4	Decision trees.	4	4
Module 4	**Data Analysis Using R/Python/Spreadsheet** (A practical record with minimum 5 problems has to be submitted).	**30**	
4.1	Conduct a market survey and prepare a project report (Questionnaire, geographical and diagrammatic representation, analysis - Descriptive Statistics) by using Spreadsheet/ Python/ R programming.	15	5

4.2	Statistical analysis and interpretation of a social problem by using Spreadsheet/ Python/ R programming.	15	5
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 15 marks** Quiz, Assignments **Practical: 15 marks** Lab involvement, Practical Record,Viva voce ***Summative assessment*** **Theory: 10 marks** Written tests
	B. End Semester Evaluation (ESE) **Theory : 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12). **Practical: 35 marks** Problem solving skills: 30 marks Record: 5 marks

References:

1. Gupta, S.P. (2021). Statistical Methods. Sultan Chandand Sons: NewDelhi.
2. Gupta, S.C. and Kapoor, V.K. (2020). Fundamentals of Mathematical Statistics, Sultan Chand and Sons.
3. Sudha G. Purohit, Sharad D. Gore and Shailaja R. Deshmukh. (2019). Statistics Using R, 2nd Edition, Narosa Publishing House.

Suggested Readings:

1. Tilman M. Davies. (2016). The Book of R, A First Course in R Programming and Statistics, No Starch Press.
2. Python for Data Analysis. (2012). Wes McKinney, O'REILLY.
3. Jason R Brigs: Python for kids- A playful introduction to programming, No Starch Press.
4. Amit Saha. (2015). Doing Math with Python, No Starch Press.

Mahatma Gandhi University

Kottayam

Programme						
Course Name	**Statistical Analysis of Related Data**					
Type of Course	**MDC**					
Course Code	**UC3MDCSTA200**					
Course Level	**200**					
Course Summary	This course focuses on a fundamental aspect of data analysis and machine learning-identifying and understanding the relationships or associations between variables. The curriculum covers the exploration of relationships among variables, considering various types of data scales such as nominal, ordinal, interval and ratio. Practical applications involve leveraging the Google Looker Studio and gretl for the computation and analysis of these relationships, providing students with a comprehensive skill set to navigate and interpret data across different scales.					
Semester	**3**	**Credits**		**3**		**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3				**45**
Pre-requisites						

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Understand the basic concepts of Google Looker Studio and gretl.	U	1
2	Apply Google Looker Studio for visualising the relationship between related variables.	A	2
3	Analyze and interpret measures of associations and dependencies.	An	2
4	Utilise gretl for practical demonstration and problem-solving in association between related variables.	A	2

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

Module 1	Course Description	Hours	CO No.
	Exploratory Time Series Data Analysis	**15**	
1.1	Google Looker Studio - Understanding the user interface, navigating through dashboards and reports, connecting to various data sources, creating charts, graphs, and tables, customising visualisations for effective communication.	8	1
1.2	Implementing filters and drill-downs in Google looker Studio and analysing real-world datasets using Google Looker Studio.	5	1
1.3	Gretl: Introduction, data entry and import, descriptive statistics and data exploration.	2	1,3
Module 2	**Correlation and Regression Analysis**	**15**	
2.1	Correlation: Definition, properties and range of correlation coefficient, invariance under linear transformation - Demonstration using gretl.	2	2,3
2.2	Importance of scatter diagram and construction of scatter diagram using Google Looker Studio.	2	1, 2
2.3	Rank correlation: Definition and examples, solving problems using gretl, illustrating the situations where Pearson	3	1,2,3
	correlation coefficient and rank correlation is used using Google Looker Studio.		
2.4	Principle of least squares: Introduction and basic problems, demonstration using Google Looker Studio.	2	1,2
2.5	Fitting of straight line and parabola using gretl with visual representation using google looker studio.	2	1,2, 3
2.6	Regression coefficients and regression lines: Basics and illustrations using gretl.	2	1,2,3
2.7	Relationship between correlation coefficient and regression coefficients and validating the relationships using data, analysis of real data for regression.	2	1,2,3
Module 3	**Statistical Analysis Using gretl**	**15**	
3.1	Categorical data: Definition, examples, frequency distributions, contingency table.	3	2,3
3.2	Visual representation of categorical data using different charts.	2	1
3.3	Chi-square test for association between variables.	2	2,3
3.4	Ordinal and logistic regression, Mantel- Haenszel test.	3	2,3
3.5	Measures of associations and dependencies - Odds Ratio, Kendall's Tau.	5	2,3
Module 4	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 15 marks** Quiz, Two Assignments(5 marks each) ***Summative assessment*** **Theory: 10 marks** Two written tests
	B. End Semester Evaluation (ESE): (Theory based examination.) **Total: 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12).

References:

1. Hurst, L. (2020). Hands On With Google Data Studio: A Data Citizen's Survival Guide. John Wiley & Sons.
2. Arnold, J. (2023). Learning Microsoft Power Bi: Transforming Data Into Insights. O'Reilly Media.

Suggested Readings:

1. Pulipati,S. and Kelly,N. (2022). Data Storytelling with Google Looker Studio: A hands-on guide to using Looker Studio for building compelling and effective dashboards
2. Lucchetti, R. and Cottrell, A. .Gretl - Gnu Regression, Econometrics and Time-series Library by Gnu Regression, Econometrics and Time-series Library, Allin Cottrell.
3. Agresti, A. (2013). Categorical Data Analysis. 3rd Edition, John Wiley & Sons Inc.

Mahatma Gandhi University

Kottayam

Programme						
Course Name	**Data Analysis Using R and Type Setting Using LaTex**					
Type of Course	**MDC**					
Course Code	**UC3MDCSTA201**					
Course Level	**200**					
Course Summary	This comprehensive course covers fundamental statistical analysis techniques, including generating frequency tables, conducting t-tests, chi-square tests, ANOVA tests, and correlation analysis. Students will also learn advanced data visualisation skills using ggplot2, delve into principles of curve fitting and linear regression models, and gain proficiency in LaTeX typesetting for creating professional documents with tables, equations, images, and bibliographies. By the end of the course, students will be equipped with essential statistical analysis tools and LaTeX formatting skills to conduct data analysis and produce high-quality research documents.					
Semester	**3**	**Credits**			**3**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3				**45**
Pre-requisites	Basic Knowledge in R programming					

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Analyse the information in the data using visual tools from R	An	1

2	Analyse the data using descriptive statistics tools in R	An	1
3	Perform basic inference tools in the data and arrive at conclusions about populations using R	A n	2
4	Understand the Basic Typesetting using Latex	U	2

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

Module 1	**Course Description**	**Hours**	**CO No.**
	Data Visualization using R	**8**	
1.1	Introduction to R and importing data into R from Other formats	3	1
1.2	Introduction to various charts and Data Visualization using ggplot2 - histogram, Bar chart, line chart, bar chart, pie chart	2	1
1.3	Generating various charts using real time data	2	1
1.4	Generating frequency table and cross tables and summary measures using R	1	1
Module 2	**Inferential Statistics and Regression Analysis using R**	**16**	
2.1	T-test (one sample, paired sample t-test, independent sample t-test) – Interpreting results, one way and two way ANOVA	4	2
2.2	Assumptions of t-test and verifying the assumptions	1	2
2.3	Non-parametric analogues of t-test, one sample ANOVA, Chi-square test for independence	4	2
2.4	Scatter diagram and correlation – Pearson and Spearman's Correlation in R	2	3
2.5	Regression Analysis in R – Linear and Multiple, Verifying the assumptions of Linear Regressions and Box Cox Transformations	3	3
2.6	Logistic Regression in R and interpreting results	2	3
Module 3	**Type Setting using Latex**	**21**	

3.1	Introduction to LaTeX and typesetting: Understand the basics of LaTeX and its role in document preparation and Learn how to customise fonts and adjust the size of text in LaTeX documents.	4	4
3.2	Explore different document classes and page styles available in LaTeX for various types of documents	3	4
3.3	Learn how to create a table of contents, index, and glossary in LaTeX for better document navigation.and Bibliography	6	4
3.4	Create lists with bullets and numbering, and format them effectively in LaTeX.	2	4
3.5	Gain proficiency in creating tables, writing equations, and inserting images into LaTeX documents for comprehensive document preparation.	6	4
Module 4	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 15 marks** Quiz, Assignments ***Summative assessment*** **Theory: 10 marks** Written tests

	B. End Semester Evaluation (ESE): (Theory based examination.) **Total: 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12).

References

1. D Narayana, Sharad Ranjan, and Nupur Tyagi (2023), Basic Computational Techniques For Data Analysis, Routledge
2. Nussbaumer Knaflic, Cole(2015), Storytelling With Data: A Data Visualization Guide For Business Professionals, Wiley
3. Andy Field, Jeremy Miles, Zoe Field (2012) DISCOVERING STATISTICS USING R, Sage Publications
4. LATEX Tutorials : A PREMIER by Indian TEX Users Group, Edited by E. Krishnan, 2003. A free PDF document from the URL https://www.tug.org/twg/mactex/tutorials/ltxprimer-1.0.pdf
5. LATEX , a Document Preparation System by Leslie Lamport (second edition, Addison Wesley, 1994)
6. Hadley Wickham and Garrett Grolemund , R for Data Science

Mahatma Gandhi University

Kottayam

Programme	**STATISTICS**					
Course Name	**Applied Statistical Analysis: Ethical Data Collection, Interpretation and Decision making in Society.**					
Type of Course	**VAC**					
Course Code	**UC3VACSTA200**					
Course Level	**200**					
Course Summary	Students will critically assess ethical implications in statistical analysis, communicate findings responsibly and synthesise information to make ethical decisions based on statistical outcomes. They will assess the reliability of statistical inferences in societal scenarios considering both the statistical significance and ethical implications of their findings.					
Semester	**3**	**Credits**			**3**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3				**45**
Pre-requisites	Level 100 knowledge of Statistics.					

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Demonstrate various data collection methods, sampling strategies, and statistical tools used for organising, summarising, and visualising data in societal contexts.	A	1
2	Apply statistical techniques such as hypothesis testing, correlation and regression analysis to real-world data.	A	2

3	Evaluate ethical considerations in data collection, statistical analysis and interpretation of results in societal contexts using statistical software packages.	E	8

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

	Course Description	**Hours**	**CO No.**
Module 1	**Sampling, Data Collection, Organizing and Summarizing Data: Case study based on a relevant topic taken from society**	**15**	
1.1	Nature of data, sampling strategies, questionnaire designing, data collection (primary/secondary) interview- designing, conduct and ethics.	3	1,3
1.2	Classification of data, tabulation of data and scaling of data.	2	1
1.3	Measures of central tendency (mean, median, mode), Measure of dispersion (Standard deviation).	3	1
1.4	Visualisation of data: Histogram, frequency polygon and ogives.	2	1
1.5	Concepts of correlation and regression.	2	1
1.6	Theory of attributes: Introduction, independence of attributes, criterion of independence, association of attributes, Yule's coefficient of association, coefficient of colligation.	3	1
Module 2	**Tests of Significance**	**15**	
2.1	Parameter, statistic, statistical inference, null and alternative hypotheses, level of significance, p-value, large sample tests for single mean, difference of means and test for proportion (one sample and two samples).	6	2
2.2	Small sample tests-t test of significance for single mean, difference in means, paired t - test for related samples.	5	2
2.3	Chi square test for independence of attributes.	4	2
Module 3	**Analysis using Statistical Software.**	**15**	
3.1	Working with real life data using statistical software packages, Introduction to R and R commander and its application. : Defining variables: Numeric and String Variables Assigning names and labels to variables and values - Entering Data.	5	1,2,3

3.2	Summary Statistics: Frequencies, Descriptive Statistics: Means, Crosstab, Graphs, Histograms and Bar charts, Scatter diagram, Pie diagram, Bivariate correlation - Linear regression.	3	1
3.3	Inferential Statistics: Statistical Tests: Testing a mean, t-test for a mean, two sample Z test for Means- Two sample t-test for means, Paired t- test, Chi-square test for independence of attributes.	4	2,3
3.4	Ethical theories and principles in data science, Group discussions on ethical frameworks and their application in data analysis.	3	3
Module 4	**Teacher Specific Content.**		

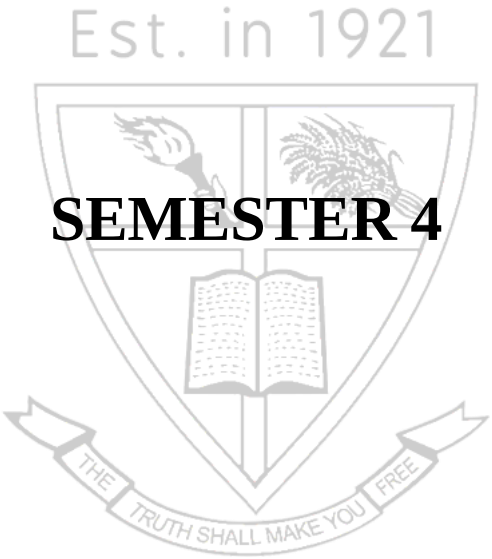
Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 15 marks** Quiz, Assignments ***Summative assessment*** **Theory: 10 marks** Two written tests.
	B. End Semester Examination(ESE) **Total: 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12).

References:

1. Powers, Daniel, and Yu Xie. (2008) Statistical methods for categorical data analysis. Emerald Group Publishing.
2. Kapoor, V.K. and Gupta, S.C. (2020): Fundamentals of Mathematical Statistics,Sultan Chand & Sons, New Delhi.
3. Fox, J. (2005). The R Commander: A basic-statistics graphical user interface to R. Journal of Statistical Software,19(9):1–42.

Suggested Readings:

1. Davis, K.(2012) Ethics of Big Data: Balancing risk and innovation. " O'Reilly Media, Inc."
2. Chiang, Chin Long.(2003) Statistical methods of analysis. World Scientific.
3. Fox, J. (2007).Extending the R Commander by “plug-in” packages. R News,7(3):46–52.



SEMESTER 4

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Basics of Multivariate Distributions**					
Type of Course	**DSC A**					
Course Code	**UC4DSCSTA200**					
Course Level	**200**					
Course Summary	Students will be proficient in conducting correlation and regression analysis, understanding bivariate and multivariate distributions, interpreting results from the distribution of quadratic forms and applying these skills in practical scenarios using R/Python software.					
Semester	**4**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3		**1**		**75**
Pre-requisites						

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Illustrate bivariate and multivariate data and analyze them.	U & A	1
2	Analyze the bivariate data using a scatter diagram.	A	2
3	Elucidate various types of correlation measures.	Ap	2
4	Construct regression models and estimate values of dependent variables.	An & C	3
5	Describe bivariate distributions and obtain marginal and conditional distributions and examine the independence of random variables.	U ,An & E	1

6	Obtain mathematical expectations and correlation.	A	2
7	Apply multivariate normal distribution in real-life situations.	U & A	2
8	Build characterizations of multivariate distribution.	C	3
Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)			

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module 1	**Correlation and Regression**	**15**	
1.1	Bivariate data, multivariate data, scatter diagram, types of correlation: Karl Pearson, Spearman's rho and Kendall's tau.	6	1,2,3
1.2	Curve fitting and regression analysis.	4	4
1.3	Multiple linear regression, multiple correlation and partial correlation : Their properties and related results.	5	4
Module 2	**Bivariate and Multivariate Distributions**	**15**	
2.1	Bivariate random vector, joint pmf, joint pdf, and bivariate cdf, marginal and conditional distributions and independence of random variables.	4	5
2.2	Mathematical expectation, conditional expectation, covariance and correlation.	2	6
2.3	Random vectors, mean vector and dispersion matrix.	2	8
2.4	Bivariate normal distribution: pdf, marginal distributions, conditional distributions and independence.	2	5
2.5	Multivariate normal distribution: mgf, characteristic function, marginal distributions and conditional distributions, properties, characterizations and orthogonal transformation.	3	7,8
2.6	Multinomial distribution and its basic properties.	2	7
Module 3	**Distribution of Quadratic Forms**	**15**	
3.1	Quadratic forms: Types, independence, Scalar quadratic forms: properties.	8	9

3.2	Distribution of quadratic forms, Cochran's theorem.	7	9
Module 4	**Practical Using R/Python** (A practical record with minimum 10 problems has to be submitted).	**30**	
4.1	1. Multiple correlation and regression. 2. Partial correlation. 3. Curve fitting. 4. Karl Pearson's correlation coefficient. 5. Spearman's rho. 6. Kendall's tau. 7. Multivariate normal distribution (variance-covariance matrix). 8. Quadratic forms (positive definite). 9. Multinomial distribution.		1, 2, 3, 4, 7, 8
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** **Formative assessment** **Theory: 15 Marks** Quiz,Two Assignments **Practical: 15 Marks** Lab involvement, Practical Record, Viva voce **Summative assessment** **Theory: 10 Marks** Two written tests.

	B. End Semester Evaluation (ESE) **Theory : 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12). **Practical: 35 marks** Problem solving skills: 30 marks Record: 5 marks

References:

1. Gupta, S.C. and Kapoor, V.K. (2020).Fundamentals of Mathematical Statistics,12th Edition, Sultan Chand & Sons, New Delhi.
2. Anderson, T.W. (2009).An Introduction to Multivariate Statistical Analysis,3rd Edition, John Wiley.
3. Rencher, A.C. (1998).Multivariate Statistical Inference and Applications,1st Edition, Wiley-Interscience.
4. Sudha G. Purohit, Sharad D. Gore and Shailaja R. Deshmukh. (2019). Statistics Using R, 2nd Edition, Narosa Publishing House.
5. F. Mary Harin Fernandez. (2022) R Programming Language, Booknetz.
6. Mathai, A.M. ,Serge B. Provost , Hans J. Haubold (2022). Multivariate Statistical Analysis in the Real and Complex Domains, Springer.
7. Mathai, A.M. (1997). Jacobians of Matrix Transformation and Functions of Matrix Arguments, World Scientific Publishing Company.

Suggested Readings:

1. Rohatgi, V.K. and Saleh, A.K.MD.E.(2015).An Introduction to Probability and Statistics, 3rd Edition, John Wiley & Sons Inc.
2. Johnson , R.A. and Wichern, D.W.(2013).Applied Multivariate Statistical Analysis, 6th Edition, Pearson Education.
3. Hogg, R.V., McKean, J.W. and Craig, A.T. (2014).Introduction to Mathematical Statistics,7th Edition, Pearson Education Publication.
4. Mood, A.M., Graybill, F.A. and Boes, D.C. (2007): Introduction to the Theory of Statistics, 3rd Edition (Reprint), Tata McGraw-Hill Pub. Co. Ltd.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Statistical Inference**					
Type of Course	**DSC A**					
Course Code	**UC4DSCSTA201**					
Course Level	**200**					
Course Summary	Students will be well-equipped to apply statistical hypothesis testing, parametric and non - parametric tests, and conduct data analysis using R / Python programming.					
Semester	**4**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3		**1**		**75**
Pre-requisites						

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Understand Chebychev's inequality, Analyse basic concepts of stochastic convergence.	U, An	1
2	Apply Law of large numbers and CLT to sequences of random variables.	A	2
3	Examine properties of a good estimator, apply Cramer-Rao inequality.	A	1,2
4	Obtain minimum variance bound estimator, estimate parameters using various methods.	E	2, 3
5	Construct confidence intervals for parameters.	C	2
6	Understand basic concepts of statistical hypotheses and their applications.	U & A	1
7	Explain various parametric test procedures and perform various parametric tests.	U,A & An	1
8	Understand the importance of normality assumption in data analysis and construct tests for normality.	U, A & C	1

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module 1	**Point and Interval Estimation**	**15**	
1.1	Chebychev's inequality, sequence of random variables, convergence of sequence of random variables, Law of large numbers (statement only).	6	1
1.2	Properties of a good estimator, Cramer-Rao inequality (without proof) and its applications.	4	2
1.3	Confidence interval, confidence coefficient, construction of confidence intervals for the mean, difference of means, variance, ratio of variances, proportion, difference of proportions and Odds ratio.	5	3
Module 2	**Methods of Estimation**	**15**	
2.1	Method of moments.	4	7
2.2	Method of maximum likelihood, properties of maximum likelihood estimation (statement only).	6	7
2.3	Method of minimum variance.	5	7
Module 3	**Statistical Hypothesis**	**15**	
3.1	Introduction to statistical hypothesis testing, Neyman-Pearson test procedure, Neyman-Pearson lemma (without proof),	**3**	4
3.2	**Parametric Tests:** Tests concerning mean, equality of means, proportion and equality of proportions, paired-t test, tests for variance and equality of variance: Chi- square test, F test, Bartlett's test and Levene's test, One way ANOVA, tests for sphericity. (Problem oriented approach)	5	4
3.3	**Non - parametric tests:** Chi-square tests: Goodness of fit, independence and homogeneity, Tests for normality- Anderson- Darling test, Shapiro-Wilk test, one sample and paired sample: Sign test, Wilcoxon signed rank test, Mann-Whitney U test and Kruskal-Wallis test. (Problem oriented approach).	7	4
Module 4	**Practical Using R/Python**	**30**	
4.1	A practical record with minimum 10 problems has to be submitted.		1
Module 5	**Teacher Specific content.** This can be classroom teaching, practical session, field visit etc. as specified by the teacher concerned. This content will be evaluated internally.		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** **Formative assessment** **Theory: 15 Marks** Quiz,Two Assignments(5 marks each) **Practical: 15 Marks** Lab involvement, Practical record, Viva voce(5 marks each) **Summative assessment** **Theory: 10 Marks** Two written tests.
	B. End Semester Evaluation (ESE) **Theory : 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12). **Practical: 35 marks** Problem solving skills: 30 marks Record: 5 marks

References:

1. Gupta, S.C. and Kapoor, V.K. (2014). Fundamentals of Mathematical Statistics, Sultan Chand & Sons, New Delhi.
2. Mood, A.M., Graybill, F.A. and Boes, D.C. (2001). Introduction to the Theory of Statistics, 3rd Edition, McGraw Hill Education (India) Private Limited.
3. Sudha G. Purohit, Sharad D. Gore and Shailaja R. Deshmukh. (2019).Statistics Using R, 2nd Edition, Narosa Publishing House.
4. Srivastava, M., Hamid Khan, A., Srivastava, N. (2014). Statistical Inference : Theory of Estimation. PHI Learning.
5. Srivastava, M., Srivastava, N. (2019) Statistical Inference : Testing of Hypotheses. PHI

Learning.

Suggested Readings:

1. Hogg, R.V., McKean, J.W. and Craig, A.T. (2014).Introduction to Mathematical Statistics, 7th Edition , Pearson Education Publication.
2. Spiegel, M.R. and Stephens L.J. (2014). Statistics,5th Edition, Schaum's outlines, McGraw-Hill Education.
3. Lehmann, E.L. and Casella, G.(2003).Theory of Point Estimation, 2nd Edition, Springer.
4. Rohatgi, V.K. and Saleh, A.K.MD.E.(2015).An Introduction to Probability and Statistics, 3rd Edition, John Wiley & Sons Inc.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Data Analysis Using JAMOVI (Data Analytics Specialization)**					
Type of Course	**DSE**					
Course Code	**UC4DSESTA200**					
Course Level	**200**					
Course Summary	**To make the students proficient in the open source statistical data analysis software JAMOVI**					
Semester	**4**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		4				**60**
Pre-requisites						

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Understand basics of JAMOVI	U	1, 2
2	Apply EDA procedures to real life datasets in JAMOVI	A, Ap, S	1, 2
3	Apply Regression modelling techniques in JAMOVI	A, Ap, S	1, 2
4	Apply Factor analysis for identification of latent variables in JAMOVI	A, Ap, S	2
5	Test statistical hypothesis in JAMOVI	A, Ap, S	2
6	Apply PCA for dimension reduction in JAMOVI	A, Ap, S	2

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

COURSE CONTENT

Content for Classroom transaction (Sub-units)

Module 1	Course Description	Hours	CO No.
	Title- Introduction to JAMOVI	**15**	
1.1	Introduction to JAMOVI, Downloading and installing JAMOVI, Exploring-Variable Types in JAMOVI	4	1, 2
1.2	Sample datasets in JAMOVI, Menus in JAMOVI, Syntax mode, Adding modules to JAMOVI, Rj Editor	11	1,2
Module 2	**Intermediate JAMOVI**	**15**	
2.1	Computing columns, Data& Label Editing, Filtering columns, Descriptive statistics and basic plots, Distributions module- random number generation, estimation of parameters	8	3,4, 6
2.2	Scatter plots, Correlation coefficients, Linear regression, Log-linear regression	7	5
Module 3	**More with JAMOVI**	**15**	
3.1	Testing of Hypothesis- Binomial test, One Sample t-test, two sample t-test, paired t-test, ANOVA, tests for association	15	5
Module 4	**Advanced JAMOVI**	**15**	
4.1	Factor Analysis-EFA, Logistic regression, Principal Component Analysis, Reliability analysis.	15	6
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.

Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 20 marks** Quiz, Assignments ,Seminar(5 marks each). ***Summative assessment*** **Theory: 10 marks** Written tests.
	B. End Semester Evaluation (ESE) **Total:70 marks** i) Short answer type questions: Answer any 10 questions out of 12 (10*3=30). ii) Short essay type questions: Answer any 4 questions out of 6 (4*7=28). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12).

References:

1. Navarro, Danielle, and David Foxcroft. "Learning statistics with jamovi: A tutorial for psychology students and other beginners (Version 0.70)." Tillgänglig online: http://learnstatswithjamovi. com [Hämtad 14 december] (2019).

2. Heo, I., Veen, D., & Van de Schoot, R. (2020, July). Tutorial: JASP for beginners. Zenodo. https://doi.org/10.5281/zenodo.4008280

3. Anderson T. W. (2010) An Introduction to Multivariate Statistical Analysis (3rd ed.) John Wiley.

4. Johnson R.A. and Wichern DAV. (2008) Applied Multivariate Statistical Analysis, (fi^edn) Pearson education.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Statistical Quality Control**					
Type of Course	**DSE**					
Course Code	**UC4DSESTA201**					
Course Level	**200**					
Course Summary	To acquire the basic knowledge of process and product control techniques. Also, built in functions in R programming are used to solve numerical problems associated with the topics discussed.					
Semester	**4**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		4				**60**
Pre-requisites						

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Understand quality and dimensions.	U	1
2	Describe statistical process control and causes of variations.	U, A	2
3	Learn statistical control charts and its construction.	K. A	2
4	Understand Control charts for variables and attributes.	A	2
5	Analyse the patterns on the control chart.	An	2
6	Learn process capability analysis and process capability indices	K,A	2
7	Understand the concept of Acceptance sampling plans.	A, An	2
8	Use R built in functions to solve numerical problems associated with topics covered in various modules	A, S	2

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module 1	**Control Charts**	**15**	
1.1	Quality: Definition, dimensions of quality, Quality system and standards: Introduction to ISO quality standards, Quality registration.	2	1
1.2	Statistical Process Control: Seven tools of SPC, chance and assignable causes of quality variation.	2	2
1.3	Statistical Control Charts: Construction and Statistical basis of 3-σ Control charts, Rational Sub-grouping.	3	3
1.4	Control charts for variables: X-bar and R-chart, X-bar and s-chart. Control charts for attributes: np-chart, p-chart, c-chart and u-chart.	4	4
1.5	Comparison between control charts for variables and control charts for attributes.	2	4
1.6	Analysis of patterns on control charts.	2	5
Module 2	**Process Capability Analysis**	**15**	
2.1	Process capability analysis, process capability indices – Cp Cpk, Cpm., estimation of process capability.	8	6
2.2	Introduction to Six-Sigma: Overview of Six Sigma, Lean Manufacturing and Total Quality Management (TQM).	7	6
Module 3	**Acceptance Sampling Plans**	**15**	
3.1	Principle of acceptance sampling plans. Single and Double sampling plan.	4	7
3.2	OC, AQL, LTPD, AOQ, AOQL, ASN, ATI functions with graphical interpretation of SSP and DSP.	6	7
3.3	Use and interpretation of Dodge and Romig sampling inspection plan tables.	5	7
Module 4	**Statistical Analysis Using R programming** (Record with minimum 5 problems has to be submitted.)	**15**	
4.1	Introduction to R	4	8
4.2	Use R built in functions to solve numerical problems associated with topics covered in various modules.	11	8
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 20 marks** Quiz, Assignments, Seminar (5 marks each). ***Summative assessment*** **Theory: 10 marks** Two written tests
	B. End Semester Evaluation (ESE) **Total: 70 marks** i) Short answer type questions: Answer any 10 questions out of 12 (10*3=30). ii) Short essay type questions: Answer any 4 questions out of 6 (4*7=28). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12).

References:

1. Montgomery, D. C. (2009): Introduction to Statistical Quality Control, 6th Edition, Wiley India Pvt. Ltd.
2. Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I & II, 8th Edition. The World Press, Kolkata.
3. Mukhopadhyay, P (2011). Applied Statistics, 2nd edition revised reprint, Books and Allied(P) Ltd.

Suggested Readings:

1. Montgomery, D. C. and Runger, G.C. (2008): Applied Statistics and Probability for Engineers, 3rd Edition reprint, Wiley India Pvt. Ltd.
2. Ehrlich, B. Harris. (2002): Transactional Six Sigma and Lean Servicing, 2nd Edition, St. Lucie Press.
3. Hoyle, David. (1995): ISO Quality Systems Handbook, 2nd Edition, Butterworth Heinemann Publication.
4. Purohit,S.G.,Deshmukh,S.R.,& Gore,S.D.(2008).Statistics using R.Alpha Science International, United Kingdom
5. Wilks S.S. (1964). Mathematical Statistics, John Wiley, New York.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Biostatistics**					
Type of Course	**DSE**					
Course Code	**UC4DSESTA202**					
Course Level	**200**					
Course Summary	This course equips students to understand the problems in Biomedical Research with the Principles of Biostatistical designs and application of different distributions.					
Semester	**4**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		4				**60**
Pre-requisites						

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	PO No
1	Understand the problems in Biomedical Research.	U	1
2	Understand the Principles of Biostatistical designs and application of different distributions.	U, A	2
3	Describe Type 1, Type 2 ,progressive censoring and random censoring.	K	2
4	Evaluate mean survival time.	E	2
5	Explain categorical data analysis.	K	2
6	Evaluate probabilities of death under competing risks models.	E	2
7	Planning and design of clinical trials.	K,An	2
8	Describe different types of clinical trials and apply in different situations.	K,A, S	2

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module 1	**Introduction to Biostatistics**	**15**	
1.1	Biostatistics: Example on statistical problems in Biomedical Research-Types of Biological data.	3	1
1.2	Principles of Biostatistical design of medical studies, functions of survival time, survival distributions and their applications viz. exponential, gamma, Weibull, Rayleigh, lognormal, distribution having bath-tub shape hazard function.	7	2
1.3	Parametric methods for comparing two survival distributions (L.R test and Cox's F- test).	5	2
Module 2	**Types of Censoring**	**15**	
2.1	Type I, Type II and progressive or random censoring with biological examples.	4	3
2.2	Estimation of mean survival time and variance of the estimator for type I and type II censored data with numerical examples.	4	4
2.3	Non-parametric methods for estimating survival function and variance of the estimator viz. Actuarial and Kaplan –Meier methods.	7	4
Module 3	**Categorical Data Analysis**	**15**	
3.1	Categorical data analysis (logistic regression) : competing risk theory, indices for measurement of probability of death under competing risks and their inter-relations.	6	5
3.2	Estimation of probabilities of death under competing risks by ML method.	4	6
3.3	Stochastic epidemic models: Simple and general epidemic models.	5	6
Module 4	**Basic Biological concepts in Genetics**	**15**	
4.1	Basic biological concepts in genetics, Mendel's law, Hardy-Weinberg equilibrium, random mating, natural selection, mutation, genetic drift, detection and estimation of linkage in heredity.	4	7
4.2	Planning and design of clinical trials, Phase I, II, and III trials. Sample size determination in fixed sample designs.	5	7
4.3	Planning of sequential, randomised clinical trials, designs for comparative trials; randomization techniques and associated distribution theory and permutation tests (basic ideas only); ethics behind randomised studies involving human subjects; randomised dose-response studies(concept only).	6	8

Module 5	Teacher Specific Content.

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 20 marks** Quiz, Assignments ,Seminar(5 marks each). ***Summative assessment*** **Theory: 10 marks** Written tests.
	B. End Semester Evaluation (ESE) **Total:70 marks** i) Short answer type questions: Answer any 10 questions out of 12 (10*3=30). ii) Short essay type questions: Answer any 4 questions out of 6 (4*7=28). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12).

References:

1. Biswas, S. (1995). Applied Stochastic Processes. A Biostatistical and Population Oriented Approach. Wiley EasternLtd., New Delhi.
2. Cox, D.R. and Oakes, D. (1984). Analysis of Survival Data. Chapman & Hall, New York

Suggested Readings

1. Elandt, R.C. and Johnson (1975). Probability Models and Statistical Methods in Genetics. John Wiley & Sons, New York.
2. Lawless, J.F.(2003). Statistical Methods for Lifetime - Second Edition. John Wiley & Sons, New York.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Econometrics**					
Type of Course	**DSE**					
Course Code	**UC4DSESTA203**					
Course Level	**400**					
Course Summary	To acquire the basic knowledge of econometric models and its applications. Also learn tests and solutions of multicollinearity and heteroscedasticity concepts.					
Semester	**4**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		4				**60**
Pre-requisites						

CO No.	**Expected Course Outcome**	**Learning Domains***	**Program Outcome**
1	Apply the challenges of empirical modelling in economics and business.	A	2
2	Analyze economic data by using regression analysis.	An	2
3	Explain theoretical background for the standard methods used in empirical analyses, like properties of least squares estimators and the statistical testing of hypotheses.	A	2
4	Describe the concept of structural econometric models and their applications in econometric modelling.	U	1

	Course Description	**Hours**	**CO No.**
Module 1	**Homogeneous functions**	**15**	
1.1	Demand and supply functions, elasticity of demand, equilibrium of market.	6	1

1.2	Production functions: Homogeneous functions, elasticity of production.	5	1
1.3	Input- output analysis, Leontief's open and closed models.	4	1
Module 2	**Linear Regression Models**	**15**	
2.1	Simple linear regression models, multiple linear regression models.	3	2
2.2	Estimation of the model parameters, tests concerning the parameters, confidence intervals,	4	2
2.3	Prediction, heteroscedasticity, tests, consequences,	4	2
2.4	Multicollinearity- consequences, Farrar-Glauber test, remedial measures. Residual Analysis.	4	
Module 3	**Generalised Least Square Methods**	**15**	
3.1	Aitken's generalised least square method, tests for autocorrelation, consequences and estimation procedures.	5	3
3.2	stochastic regressors, errors in variables, use of Dummy variables in regression.	4	3
3.3	polynomial regression models in one variable, basic ideas of logistic regression and stepwise regression.	6	3
Module 4	**Simultaneous Equation Models and its Identification**	**15**	
4.1	Simultaneous equation models, Identification problems, rank and order condition.	5	4
4.2	Methods of estimation- indirect least squares, least variance ratio(LVR) or LIML.	6	4
4.3	Two-stage least squares, FIML- methods.	4	4
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.

Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 20 marks** Quiz, Assignments, Seminar ***Summative assessment*** **Theory:10 marks** Written tests
	B. End Semester Evaluation(ESE) **Total:70 marks** i) Short answer type questions: Answer any 10 questions out of 12 (10*3=30). ii) Short essay type questions: Answer any 4 questions out of 6 (4*7=28). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12).

References:

1. Johnston J. (1984). Econometric Methods (Third edition), McGraw Hill, New York.
2. Montgomery D.C., Peck E.A. and Vining G.G. (2007). Introduction to Linear Regression Analysis, John Wiley, India.
3. Gujarati D. (2009). Basic Econometrics, Fifth edn McGraw Hill.
4. Apte P.G. (1990). Text book of Econometrics, Tata McGraw Hill.
5. Theil H. (1982). Introduction to the Theory and Practice of Econometrics, John Wiley.

Suggested Readings:

1. Gujarati, D. and Sangeetha, S. (2007). Basic Econometrics, 4th Edition, McGraw Hill Companies.
2. Johnston, J. (1972). Econometric Methods, 2nd Edition, McGraw Hill International.
3. Maddala, G.S. and Lahiri, K. (2009). Introduction to Econometrics, 4th Edition, John Wiley & Sons.
4. Koutsoyiannis, A. (2004). Theory of Econometrics, 2nd Edition, Palgrave Macmillan Limited.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Statistical Inference Using R/Python**					
Type of Course	**DSC B**					
Course Code	**UC4DSCSTA202**					
Course Level	**200**					
Course Summary	This course equips students with a comprehensive understanding of different sampling distributions, estimation methods, parameter testing, and non - parametric testing for hypothesis evaluation. The practical aspect of the course involves hands-on experience in conducting data analysis using R or Python.					
Semester	**4**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3		**1**		**75**
Pre-requisites	Level 100 knowledge of Statistics					

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains	Program Outcome
1	Understand different sampling distributions.	U	1
2	Describe estimation and methods.	U	1
3	Relate different parametric tests in testing the hypothesis.	An	1
4	Organise different non-parametric tests in testing the hypothesis.	An	1
5	Conduct data analysis using R/Python.	E	2

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module 1	**Sampling Distributions**	**15**	
1.1	Statistic, parameter.	1	1
1.2	Distribution of sample mean and variance.	2	1
1.3	Normal distribution.	3	1
1.4	Student's t-distribution.	2	1
1.5	Chi- square distribution.	2	1
1.6	F distribution.	2	1
1.7	Inter-relationship between normal, t, Chi-square and F distributions.	3	1
Module 2	**Estimation of Parameters and methods of Estimation**	**15**	
2.1	Estimation, point estimation and interval estimation.	2	2
2.2	Desirable properties of a good point estimator.	6	2
2.3	Methods of estimation – MLE, method of moments.	7	2
Module 3	**Testing of Hypothesis**	**15**	
3.1	Testing of hypothesis, Statistical test, null and alternative hypothesis, types of errors, significance level, power, critical region and p- value.	2	3
3.2	Parametric test: Testing of population mean (One sample and two samples) (z test, t-test), paired t test.	6	3
3.3	Testing of population proportion (One sample and two samples).	3	3
3.4	ANOVA(one way only).	1	3
3.5	Non-parametric tests: Chi-square test, sign test, median test. Kruskal Wallis H test and Wilcoxon test.	3	3
Module 4	**Data analysis using R /Python.**	**30**	
4.1	Introduction to Python/R.	4	5
4.2	Categorical data analysis.	4	5
4.3	Random number Generation.	2	5
4.4	Descriptive and inferential statistical analysis using R/Python, Data visualisation, Descriptive measures, Correlation and Regression, Statistical Tests, ANOVA.	20	5
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.

Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** ***Theory:* 15 marks** Quiz, Assignments ***Practical:* 15 marks** Lab involvement, Practical record,Viva voce. ***Summative assessment*** **Theory: 10 marks** Written tests.
	B. End Semester Evaluation (ESE) **Theory : 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12). **Practical: 35 marks** Problem solving skills: 30 marks Record: 5 marks

References:

1. Rohatgi V.K. and Saleh, A.K. Md.E. (2009): An Introduction to Probability andStatistics.2nd Edn. (Reprint)John Wiley andSons.
2. Gupta, S.P. (2021) Statistical Methods. Sultan Chandand Sons: NewDelhi.
3. Gupta, S.C. and Kapoor, V.K.(2020) Fundamentals of Mathematical Statistics, Sultan Chand and Sons.
4. Sudha G Purohit, Sharad D. Gore, Shailaja Deshmukh (2019) Statistics using R, 2nd edition, Narosa Publishing House.
5. Python for Everybody: Exploring Data Using Python3, ADS 2016.

Suggested Readings:

1. Mood, A.M. Graybill, F.A. and Boes, D.C. (2007) Introduction to the Theory of Statistics, 3rd Edition., (Reprint), Tata Mc Graw-Hill Pub. Co.Ltd.
2. John E Freund, Mathematical Statistics, Pearson Edn, NewDelhi
3. Tilman M. Davies. (2016) The Book of R, A First Course in Programming and Statistics, No Starch Press.
4. Python for Data Analysis (2012).WesMc Kinney, O'REILLY.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Statistical Research Methods using Softwares.**					
Type of Course	**DSC B**					
Course Code	**UC4DSCSTA203**					
Course Level	**200**					
Course Summary & Justification	This course aims to equip students with a solid foundation in Research Methodology, Statistical Testing and Data Analysis.					
Semester	**4**	**Credits**			**4**	**Total Hours**
Total Student Learning Time (SLT)	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3		**1**		**75**
Pre-requisites	Level 100 knowledge of Statistics					

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains	Program Outcome
1	Understand different research methods in social science	U	1
2	Understand the statistical testing procedure in sociology	U	1
3	Illustrate the large sample tests	A	2
4	Describe the small sample tests	A	2
5	Conduct a social survey and data analysis using R/Python/Spreadsheet.	E	2

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module 1	**Introduction to Research Methodology**	**15**	
1.1	Research design, Qualitative and quantitative research.	3	1
1.2	Data collection methods & sampling techniques.	5	1
1.3	Research reporting and Communication-Writing Research proposal.	4	1
1.4	Apply research methods to real-world social issues.	3	
Module 2	**Testing of Hypothesis**	**10**	
2.1	Parameter, statistic.	1	2
2.2	Statistical hypothesis, Simple and composite hypothesis.	1	2
2.3	Null and alternative hypotheses, type I and type II Errors.	2	2
2.4	Critical region, size of the test, p value, power.	2	2
2.5	Sociological research problems in Statistical perspective.	4	2
Module 3	**Parametric and Non-parametric Tests**	**20**	
3.1	Large sample test: z test for single mean and equality of two means.	5	2
3.2	Small sample test: t test for single mean and equality of two means, paired t test.	5	3
3.3	ANOVA (one way only).	1	3
3.4	Non- parametric tests: Testing association of attributes using Chi square test.	2	4
3.5	Sign test, median test, Wilcoxon Ranked test-simple problems only.	6	4
3.6	Applications of statistical tests in various fields.	1	4
Module 4	**Data analysis using R/spreadsheet/Python**	**30**	
4.1	Conduct a social survey and prepare a project report (Questionnaire, geographical and diagrammatic representation, analysis - Descriptive Statistics).	12	5
4.2	Statistical analysis and interpretation of a social problem by using Spreadsheet/ Python/ R programming.	18	5
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** ***Theory:* 15 marks** Quiz, Assignments ***Practical:* 15 marks** Lab involvement, Practical Record, Viva voce. ***Summative assessment*** **Theory: 10 marks** Written tests
	B. End Semester Evaluation (ESE) **Theory : 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12). **Practical: 35 marks** Problem solving skills: 30 marks Record: 5 marks

References:

1. Rohatgi V.K. and Saleh, A.K. Md.E. (2009): An Introduction to Probability andStatistics.2nd Edition (Reprint)John Wiley andSons.
2. Gupta, S.P. (2021) Statistical Methods. Sultan Chandand Sons: NewDelhi.
3. Gupta, S.C. and Kapoor, V.K.(2020) Fundamentals of Mathematical Statistics, Sultan Chand and Sons.
4. Sudha G Purohit, Sharad D. Gore, Shailaja Deshmukh (2019) Statistics using R, 2nd Edition, Narosa Publishing House.
5. Python for Everybody: Exploring Data Using Python3, ADS 2016.

Suggested Readings:

1. Mood, A.M. Graybill, F.A. and Boes, D.C. (2007) Introduction to the Theory of Statistics, 3rd Edition., (Reprint), Tata Mc Graw-Hill Pub. Co.Ltd.
2. John E Freund, Mathematical Statistics, Pearson Edn, New Delhi
3. Tilman M. Davies. (2016). The Book of R, A First Course in Programming and Statistics, No Starch Press.
4. Python for Data Analysis (2012).WesMc Kinney, O'REILLY.

Mahatma Gandhi University

Kottayam

Programme	**BSc (Hons) Statistics**					
Course Name	**Statistical Modelling in Data Science**					
Type of Course	**DSC B**					
Course Code	**UC4DSCSTA204**					
Course Level	**200**					
Course Summary	This course provides a comprehensive introduction to Data Sciences, covering Inferential Statistics, Non-parametric Tests, ANOVA and Analysis of AI models in Statistics. The focus is on developing practical skills for data analysis and interpretation in real-world scenarios.					
Semester	**4**	**Credits**			**4**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3		**1**		**75**
Pre-requisites	Level 100 knowledge of Statistics					

EXPECTED COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains	Program Outcome
1	Understand the basics of Data science	U	1
2	Operate Parametric tests	A	2
3	Relate Non parametric tests	An	1
4	Compare AI models in statistics	An	1
5	Conduct statistical data analysis using R/Python	E	2
Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)			

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module 1	**Introduction to Data Science**	**15**	
1.1	Introduction, definition.	1	1
1.2	Data Science in various fields, examples.	1	1
1.3	Impact of data science.	1	1
1.4	Understating data: Introduction, types of data, numeric, categorical, graphical, high dimensional data.	3	1
1.5	Classification of digital data: structured, semi-structured, unstructured, example, applications.	3	1
1.6	Sources of data: Time series data, transactional data, biological data, spatial data, social network data.	3	1
1.7	Data evolution.	1	1
1.8	Introduction of big data.	2	1
Module 2	**Inferential Statistics, Non parametric test and ANOVA**	**18**	
2.1	Introduction, sampling distribution: z distribution, t distribution.	5	2
2.2	Hypothesis testing: z test, t test (one sample), problems.	5	2
2.3	Introduction, chi square test for goodness of fit and independence.	4	3
2.4	F test. ANOVA (one way classification).	4	3
Module 3	**AI models in Statistics**	**12**	
3.1	Linear and Multiple Regression.	4	4
3.2	Logistic Regression.	4	4
3.3	Decision Trees.	4	4
Module 4	**Exploratory Data Analysis using R/Python**	**30**	
4.1	Random number generation.	6	5
4.2	Descriptive and inferential statistical analysis using R/Python Data visualisation, Descriptive measures, Correlation and Regression, Statistical Tests, ANOVA.	24	5
Module 5	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** ***Theory:* 15 marks** Quiz, Assignments ***Practical:* 15 marks** Lab involvement, Practical Record,Viva voce ***Summative assessment*** **Theory: 10 marks** Written tests
	B. End Semester Evaluation (ESE) **Theory : 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12). **Practical: 35 marks** Problem solving skills: 30 marks Record: 5 marks

References:

1. Rohatgi V.K. and Saleh, A.K. Md.E. (2009). An Introduction to Probability andStatistics.2nd Edition. (Reprint)John Wiley andSons.
2. Gupta, S.P. (2021). Statistical Methods. Sultan Chandand Sons: NewDelhi.
3. Gupta, S.C. and Kapoor, V.K.(2020). Fundamentals of Mathematical Statistics, Sultan Chand and Sons.
4. Sudha G Purohit, Sharad D. Gore, Shailaja Deshmukh (2019) Statistics using R, 2nd Edition, Narosa Publishing House.
5. Python for Everybody: Exploring Data Using Python3, ADS 2016.

Suggested Readings:

1. Mood, A.M. Graybill, F.A. and Boes, D.C. (2007) Introduction to the Theory of Statistics, 3rd Edition., (Reprint), Tata Mc Graw-Hill Pub. Co.Ltd.
2. John E Freund, Mathematical Statistics, Pearson Edition, New Delhi
3. Tilman M. Davies. (2016). The Book of R, A First Course in Programming and Statistics, No Starch Press.
4. Python for Data Analysis (2012).WesMc Kinney, O'REILLY.

Mahatma Gandhi University

Kottayam

Programme						
Course Name	**Introduction to Spreadsheets and LaTex typing**					
Type of Course	**SEC**					
Course Code	**UC4SECSTA200**					
Course Level	**200**					
Course Summary	To get basic knowledge and skills of data analysis using spreadsheets and be able to create printed materials with professional quality using LaTex.					
Semester	**4**	**Credits**			**3**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3				**45**
Pre-requisites	Level 100 knowledge in Statistics /Computer					

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	Program Outcome
1	Illustrate how to present data in a presentable format using pictures, tables and create well-presented documents.	U	1
2	Analyze the data and compare the distributions with statistical believes.	A	2
3	Elucidate new conclusions, if any, shown by the data based on the thorough analysis.	Ap	2
4	Critically examine and compare the results of the data analysis.	A	2
5	Describe the data based on the analysis using the spreadsheet.	U	1
6	Explain how to create documents and powerpoints.	U	1
7	Build documents using LaTex.	C	1
8	Appraise the need for presenting data and documents suitable for different situations.	E	2

Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

COURSE CONTENT
Content for Classroom Transaction (Sub-units

	Course Description	Hours	CO No.
Module 1	**Spreadsheet and Data**	**15**	
1.1	Basics of spreadsheet and data types, creation of worksheets, editing, formatting and saving.	3	1
1.2	Introduction to functions in a spreadsheet, if function, freeze panes, vlookup, hlookup, sorting, filtering.	5	1,2
1.3	Pivot tables, Statistics in spreadsheets, conditional formatting,data validation.	4	2,3,4
1.4	Data visualisation, Statistical analysis using spreadsheets.	3	4,5
Module 2	**Basics of LaTex**	**15**	
2.1	Introduction to LaTex interfaces, understanding Latex compilation, basic syntax.	3	7
2.2	Writing equations, matrices, tables. Page Layout: Titles,abstract,chapters,sections,references,equatio references, citation.List Making Environments.	4	7
2.3	Table of contents, generating commands, figure handling numbering, list of figures,list of tables, generating index.	3	7
2.4	Classes: Article, book, report, beamer, slides. Applicationsto: Writing articles / Projects.	3	7,8
2.5	Presentation using beamer.	2	6,8
Module 3	**Statistical Computing using spreadsheet and LaTex.** (Exercises based on the above concepts. Both spreadsheet & LaTex).	**15**	2,6,7,8
Module 4	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.

Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 15 marks** Quiz, Assignments. ***Summative assessment*** **Theory: 10 marks** Written tests
	B. End Semester Evaluation(ESE) **Total: 50 marks** i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14). ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24). iii) Essay type questions: Answer any 1 question out of 2 (1*12=12).

References:

1. Excel 2022: From Basic to Advanced. (2022). George Wahlberg.
2. Stefan Kottwitz: LATEX Cookbook. (2015). Packt Publishing.
3. David F. Griffths and Desmond J. Higham. (2016). Learning LATEX (2nd edition) Siam.

Suggested Readings:

1. Excel Formulas and Functions. (2020). Basics: Step-by-Step Guide with Examples for Beginners (Excel Academy Book 2) Adam Ramirez .
2. Excel 2022 : Three books-in-one: a to z mastery guide on excel basic operations, excel formulas, functions, pivot tables & dashboards (2022). Joe Webinar.
3. M.R.C. van Dongen:LATEX and Friends (2012). Springer-Verlag Berlin Heidelberg.

Mahatma Gandhi University

Kottayam

Programme	**STATISTICS**					
Course Name	**Ethical Dimensions in Statistical Machine Learning through R/Python**					
Type of Course	**VAC**					
Course Code	**UC4VACSTA200**					
Course Level	**200**					
Course Summary	The course delves into the crucial intersection of ethics and data analysis tools. Students examine real-world ethical dilemmas and learn strategies to mitigate biases and ensure responsible data handling within software-driven analyses. The course also gives an introduction to statistical machine learning and enables the student to up-skill his technical presentation skills.					
Semester	**4**	**Credits**			**3**	**Total Hours**
Course Details	**Learning Approach**	**Lecture**	**Tutorial**	**Practical**	**Others**	
		3				**45**
Pre-requisites						

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains *	Program Outcome
1	To critically analyze summarising data and testing a hypothesis.	An	1
2	To familiarise the basic concepts of model building and Statistical Machine Learning.	S	2
3	To articulate and present, both orally and in written form, the ethical implications of real life data using R/Python.	Ap	8

Remember (K), Understand (U), Apply (A), Analyze (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

	Course Description	Hours	CO No.
Module 1	**Foundation of Data Analysis and Ethical Framework**	**15**	
1.1	Basic on data collection, questionnaire preparation, interview methods for collecting data, organising and cleaning data.	2	1
1.2	Descriptive statistics, correlation and scatter plot. Visualisation of data: Histogram, frequency polygon and ogives.	3	1
1.3	Theory of attributes: Introduction, independence of attributes, criterion of independence, association of attributes, Yule's coefficient of association and coefficient of colligation.	4	1
1.4	Small sample tests: t test and F test-t test of significance for single mean, difference in means, paired t - test for related samples, F test of significance for equality of population variances, chi- square test.	6	1
Module 2	**Introduction to Model Building and Statistical Machine Learning**	**15**	
2.1	Regression, simple linear regression, multiple linear regression and logistic regression.	4	1, 2
2.2	Bayesian inference: Prior, posterior, map, regularisation in Bayesian setup, introduction to mcmc (markov chain monte carlo).	5	2
2.3	Classification, introduction, example of supervised learning, classification model, classification learning steps, common classification algorithms- KNN, decision tree, random forest models, support vector machine.	6	2
Module 3	**Ethical Decision Making and Communication in Data Analysis**	**15**	
3.1	Ethical theories and principles in data science, group discussions on ethical frameworks and their applications in data analysis.	6	3
3.2	Introduction to R/ Python.	4	1,3
3.3	Presentation on the implemented data analysis using real life data using R/Python.	5	1,2,3
Module 4	**Teacher Specific Content.**		

Teaching and Learning Approach	**Classroom Procedure (Mode of transaction)** Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
Assessment Types	**MODE OF ASSESSMENT** **A. Continuous Comprehensive Assessment (CCA)** ***Formative assessment*** **Theory: 15 marks** Quiz, Assignments ***Summative assessment*** **Theory: 10 marks** Written tests.
	B. End Semester Evaluation: (Theory based examination.) **Total: 50 marks** i) MCQ: Answer 10 questions (10*1=10). ii) Short essay type questions: Answer any 5 questions out of 7 (5*6=30). iii) Essay type questions: Answer any 1 question out of 2 (1*10=10).

References:

1. Wickham, Hadley, Mine Çetinkaya-Rundel, and Garrett Grolemund.(2023). R for data science. " O'Reilly Media, Inc.".
2. V.K.Kapoor and S.C.Gupta (2010). Fundamentals of Mathematical Statistics, Sultan Chand & Sons, New Delhi.
3. Chiang, Chin Long.(2003). Statistical methods of analysis. World Scientific.

Suggested Readings:

1. Davis, Kord. (2012). Ethics of Big Data: Balancing risk and innovation." O'Reilly Media, Inc.".
2. Powers, Daniel, and Yu Xie.(2008). Statistical methods for categorical data analysis. Emerald Group Publishing.
3. Sugiyama, Masashi.(2015). Introduction to statistical machine learning. Morgan Kaufmann.