ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

UNDER GRADUATE PROGRAMMES
SYLLABUS

SJC-UGP

(2025 Admission Onwards)

Faculty: Science

BoS: Statistics

Programme: Bachelor of Science (Honours) Statistics

St. Joseph's College Moolamattom (Autonomous)
Arakulam – 685591, Kerala, India

Contents

Sl.No	Title
1.	Syllabus Index
2.	Semester 1
3.	Semester 2
4.	Semester 3
5.	Semester 4

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

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
SJC1DSCSTA100	Fundamentals of Statistics and Data Visualisation	DSC A	4	5	3		2	
SJC1MDCSTA100	Statistical Data Collection	MDC	3	4	2		2	
SJC1MDCSTA101	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
SJC2DSCSTA100	Introduction to Statistical Modelling	DSC A	4	5	3		2	
SJC2MDCSTA100	Time Series Methods and their Applications	MDC	3	4	2		2	
SJC2MDCSTA101	Data Analysis using JAMOVI and Introduction to R							

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

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
SJC3DSCSTA202	Data Analysis in Inferential Statistics using R/Python	DSC B	4	5	3		2	
SJC3DSCSTA203	Statistical Research Techniques using Softwares							
SJC3DSCSTA204	Business Data Analytics							
SJC3MDCSTA200	Statistical Analysis of Related Data	MDC	3	3	3			
SJC3MDCSTA201	Data Analysis using R and Type Setting using LaTex							
SJC3VACSTA200	Applied Statistical Analysis: Ethical Data Collection, Interpretation and Decision making in Society	VAC	3	3	3			

ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

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
SJC4DSCSTA202	Statistical Inference using R/Python	DSC B	4	5	3		2	
SJC4DSCSTA203	Statistical Research Methods using Softwares							
SJC4DSCSTA204	Statistical Modelling in Data Science							
SJC4SECSTA200	Introduction to Spreadsheets and Latex Typing	SEC	3	3	3			
SJC4VACSTA200	Ethical Dimensions in Statistical Machine Learning through R/Python	VAC	3	3	3			
SJC4INTSTA200	Internship		2					

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

Semester-1

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme	**BSc (Hons) Statistics**					
Course Name	**Fundamentals of Statistics and Data Visualisation**					
Type of Course	**DSC A**					
Course Code	**SJC1DSCSTA100**					
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		**2**		**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

COURSE CONTENT

Content for Classroom Transaction (Units)

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.**		

ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

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** **Thory: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

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

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

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme						
Course Name	**Statistical Data Collection**					
Type of Course	**MDC**					
Course Code	**SJC1MDCSTA100**					
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		**2**		**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.

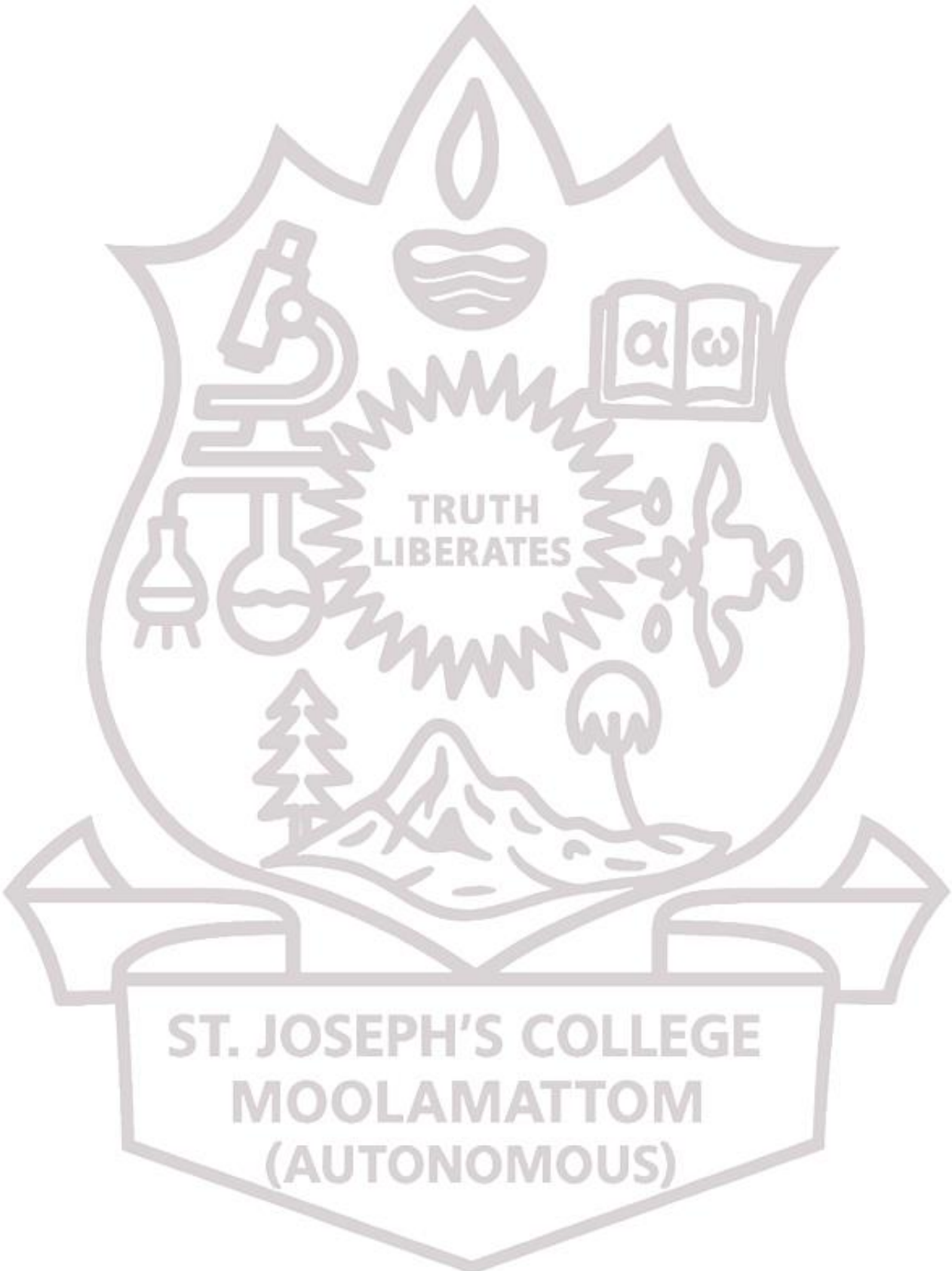
Suggested Readings:

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.

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme						
Course Name	**Data Analysis using Libre Calc**					
Type of Course	**MDC**					
Course Code	**SJC1MDCSTA101**					
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		**2**		**60**
Pre-requisites						


TRUTH
LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

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 descriptive 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**

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

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.

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

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

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme	**BSc (Hons) Statistics**					
Course Name	**Introduction to Statistical Modelling**					
Type of Course	**DSC A**					
Course Code	**SJC2DSCSTA100**					
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		**2**		**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	Develop skills required to effectively understand various distributions.	S	2
3	Analyse several applications and advantages of distributions.	An	2

4	Evaluate fitting procedure of distribution and its simulation using spreadsheet.	A,E & 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**	**15**	
1.1	Describe univariate random variables in discrete and continuous cases.	2	1
1.2	Demonstrate probability mass function, probability density function and their properties, distribution function of a random variable: Definition and properties.	3	1
1.3	Demonstrate functions of random variable, transformations of random variable (univariate).	2	1
Module 2	**Mathematical Expectation**	**15**	
2.1	Demonstrate mathematical expectation, its properties and simple problems.	4	1
2.2	Describe Arithmetic Mean (AM), Geometric Mean (GM), Harmonic Mean (HM), Mean Deviation and Variance in terms of expectation and evaluate simple problems.	5	1
2.3	Describe generating functions: Moment generating function, characteristic function, their properties and simple problems.	6	1

Module 3	**Discrete and Continuous Distributions**	**15**	
3.1	Discrete uniform distribution and Bernoulli distribution, explain binomial distribution and its properties, simple problems.	3	2
3.2	Explain Poisson distribution and its properties, simple problems. Explain geometric distribution, its properties	4	2
3.3	Explain continuous uniform distribution and its properties.	2	2
3.4	Explain exponential distribution, its properties.	3	2
3.5	Explain normal distribution and its properties. Discuss standard normal distribution and use of standard normal tables, problems.	3	3
Module 4	**Spreadsheet for Statistical Computing** (A practical record with minimum 10 problems has to be submitted).	**30**	
4.1	Use spreadsheet functions to solve numerical problems associated with topics covered in various modules.	30	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.

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

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

ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

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.

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme						
Course Name	**Time Series Methods and Their Applications**					
Type of Course	**MDC**					
Course Code	**SJC2MDCSTA100**					
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		**2**		**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

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

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)

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.**	**30**	3

ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

	(A practical record with minimum 5 problems has to be submitted)		
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).

ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

	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.

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme						
Course Name	**Data Analysis Using JAMOVI and Introduction to R**					
Type of Course	**MDC**					
Course Code	**SJC2MDCSTA101**					
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		**2**		**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 CONTENT

Content for Classroom transaction (Sub-units)

Module 1	**Course Description**	**Hours**	**CO No.**
	Data Visualization and *Inferential Statistics* using JAMOVI	**15**	
1.1	Types of Data-Ordinal Interval, ratio, measures of central tendency – mean, median, mode, measures of dispersion – Quartile Deviation, variance, standard deviation.	3	3
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
1.5	Non-parametric analogues of t-test, one sample ANOVA	4	4

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

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	4
3.2	Introduction to various charts- histogram, Bar chart, line chart, bar chart, pie chart	2	3
3.3	Generating various charts using real time data	2	4
3.4	Generating frequency table and cross tables and summary measures using JAMOVI	2	4
3.5	Scatter diagram and correlation – Pearson and Spearman's Correlation in JAMOVI	2	4
3.6	Regression Analysis in Jamovi and Spreadsheet.	2	4
3.8	t-test (one sample, paired sample t-test, independent sample t-test) – Interpreting results, one way and two way ANOVA	3	4

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

3.9	Assumptions of t-test and verifying the assumptions	2	3
3.10	Chi-square test for independence	2	4
3.11	Non-parametric analogues of t-test, one sample ANOVA	2	4
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

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

	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

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

Semester-3

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme	**BSc (Hons) Statistics**					
Course Name	**Data Analysis in Inferential Statistics Using R/Python**					
Type of Course	**DSC B**					
Course Code	**SJC3DSCSTA202**					
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		**2**		**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.**		

ST JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

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

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

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.

ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme	**BSc (Hons) Statistics**					
Course Name	**Statistical Research Techniques using Softwares**					
Type of Course	**DSC B**					
Course Code	**SJC3DSCSTA203**					
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		**2**		**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.

ST JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme	**BSc (Hons) Statistics**					
Course Name	**Business Data Analytics**					
Type of Course	**DSC B**					
Course Code	**SJC3DSCSTA204**					
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		**2**		**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

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

	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.

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme						
Course Name	**Statistical Analysis of Related Data**					
Type of Course	**MDC**					
Course Code	**SJC3MDCSTA200**					
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)			

COURSE CONTENT

Content for Classroom Transaction (Sub-units)

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.

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
(AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme						
Course Name	**Data Analysis Using R and Type Setting Using LaTex**					
Type of Course	**MDC**					
Course Code	**SJC3MDCSTA201**					
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)			

COURSE CONTENT

Content for Classroom transaction (Sub-units)

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	2	4

	effectively in LaTeX.		
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

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme	**STATISTICS**					
Course Name	**Applied Statistical Analysis: Ethical Data Collection, Interpretation and Decision making in Society.**					
Type of Course	**VAC**					
Course Code	**SJC3VACSTA200**					
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 CONTENT

Content for Classroom Transaction (Sub-units)

	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

St. Joseph's College Moolamattom (Autonomous)

Programme	**BSc (Hons) Statistics**					
Course Name	**Statistical Inference Using R/Python**					
Type of Course	**DSC B**					
Course Code	**SJC4DSCSTA202**					
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		**2**		**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

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

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.

ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme	**BSc (Hons) Statistics**					
Course Name	**Statistical Research Methods using Softwares.**					
Type of Course	**DSC B**					
Course Code	**SJC4DSCSTA203**					
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		**2**		**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.**		

ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

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

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

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.

ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme	**BSc (Hons) Statistics**					
Course Name	**Statistical Modelling in Data Science**					
Type of Course	**DSC B**					
Course Code	**SJC4DSCSTA204**					
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		**2**		**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.**		

ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

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

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE MOOLAMATTOM (AUTONOMOUS)

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.

ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme						
Course Name	**Introduction to Spreadsheets and LaTex typing**					
Type of Course	**SEC**					
Course Code	**SJC4SECSTA200**					
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.

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

St. Joseph's College Moolamattom (Autonomous)

Programme	**STATISTICS**					
Course Name	**Ethical Dimensions in Statistical Machine Learning through R/Python**					
Type of Course	**VAC**					
Course Code	**SJC4VACSTA200**					
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.

TRUTH LIBERATES
ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)

	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.

ST. JOSEPH'S COLLEGE
MOOLAMATTOM
(AUTONOMOUS)