



HSNC UNIVERSITY, MUMBAI

(2023-2024)

Ordinances and Regulations With

Respect to

Choice Based Credit System (CBCS)

For the Programmes Under

The Faculty of Science and Technology

Framed According to the National Education Policy (NEP 2020)

To be implemented from Academic Year: 2023-2024

For the Course

Statistics

Curriculum – First Year Undergraduate

Semester-I and Semester -II

2023-2024



HSNC UNIVERSITY, MUMBAI

Board of Faculty of Science & Technology

Board of Studies in the Subjects of Statistics and Data Science & Business Analytics

1) Name of Chairperson/Co-Chairperson/Coordinator:-

- a) **Dr Asha Jindal**, Professor and (UG: Head & PG: Coordinator), Department of Statistics, K. C. college, HSNC University Churchgate, Mumbai –400 020. Email ID- asha.jindal@kccollege.edu.in
Mobile no- 9821235627

2) Two to five teachers each having minimum five years teaching experience amongst the full time teachers of the Departments, in the relevant subject.

- a) **Dr. S. B. Muley**, Assistant Professor, Department of Statistics, K. C. college, HSNC University Churchgate, Mumbai – 400 020. Email ID sakharam.muley@kccollege.edu.in,
Mobile No- 9323817918
- b) **Mrs. Pratiksha Kadam**, Assistant Professor, Department of Statistics, K. C. college, HSNC University Churchgate, Mumbai – 400 020. Email ID pratiksha.kadam@kccollege.edu.in,
Mobile No- 7507162816
- c) **Ms. Shailaja Rane**, Assistant Professor, Department of Statistics, K. C. college, HSNC University Churchgate, Mumbai – 400 020. Email ID shailaja.rane@kccollege.edu.in, Mobile No- 7506986359

3) One Professor / Associate Professor from other Universities or professor / Associate Professor from colleges managed by Parent Body;

- a) **Dr Anjum Ara Ahmed**; Professor and I/C Principal, Rizvi College, Mumbai. Email ID anjumahmed8@gmail.com, Mobile No- 8451046220

4) Four external experts from Industry / Research / eminent scholar in the field relevant to the subject nominated by the Parent Body;

- a. **Prof. Suresh Kumar Sharma**, Senior Professor, Department of Statistics, Panjab University, Chandigarh.
Email ID ssharma643@yahoo.co.in, **Mobile No-**9815911381
- b. **Mr Mukesh Jain**, Vice President and Chief Technological Officer, Capgemini. Email ID mdjain@hotmail.com, **Mobile No-**7972637347.
- c. **Dr Santosh Gite**, Professor, Dept. of Statistics, University of Mumbai, Mumbai. Email ID santgite@yahoo.com, **Mobile No-** 9167157717.
- d. **Mr Prashant Kumar Nair**, Director, Geo Spatial Analytics Global Lead, Intelligent Analytics, Nielsen Connect, Email ID prashantkumar.nair@nielsen.com , **Mobile No-**9833747057.

5. Top rankers of the Final Year Graduate and Final Year Post Graduate examination of previous year of the concerned subject as invitee members for discussions on framing or revision of syllabus of that subject or group of subjects for one year.

- a) **Mr. Chinmay Mokal** (Postgraduate student 22-23) Email ID- chinmay30mokal@gmail.com
; Mobile no- 9372323901
- b) **Ms. Aanchal Goyal** (undergraduate student 22-23) Email ID- aanchalgoyal2703@gmail.com
; Mobile no- 7738886488

Statistics

Part 1-

Preamble

B. Sc. Statistics program is of minimum 120 credits cover six semesters. Statistics is the language of the uncertainties riddled modern information age. Statistics facilitates the choice making process by quantifying the element of chance or uncertainties. It's descriptive and inferential roles not only formulate the basis of the expansion of almost all the disciplines of the modern world, but also provide an array of non-traditional employment avenues starting from that of sport analysts to business analysts to actuaries. The thrust of the course is to prepare students to enter into a promising career even after graduation, as also provide to them a platform for pursuing higher studies resulting in post-graduate or doctorate degrees. The program has some unique features like number of elective courses and extensive computer training of statistical computations including standard software packages like IBM SPSS, Excel, MINITAB, R and PYTHON. The integration of continuous assessment, project work, and applied learning ensures that students develop not only academic knowledge but also critical thinking, ethical decision-making, and effective communication skills, essential for careers in statistics, data science, finance, research, and beyond. The curriculum supports the objectives of the National Education Policy (NEP 2020) by focusing on interdisciplinary collaboration, sustainable practices and lifelong learning.

1. Programme Outcomes

1. **Understanding Core Scientific Concepts:** Students will gain clear insight and understanding to recall key scientific principles across various fields. A well-established foundational knowledge of the subject will play a crucial role for deeper learning and future studies.
2. **Commit to Lifelong Scientific Learning:** Students will cultivate a habit of continuous learning and shall learn to stay updated with the latest scientific and technological advancements. This mindset will ensure that they remain relevant, engaged and informed throughout their future academic journey.
3. **Abilities to Analyse and Evaluate:** Students will learn to classify and scrutinize complex problems into manageable parts, critically analyse data, and evaluate potential solutions to scientific problems.
4. **Assessing Ethical Implications:** Students will be trained to evaluate the ethical dimensions of research and technological innovations, ensuring that their decisions consider societal impacts and they adhere to ethical standards. This is vital for responsible and sustainable practices.
5. **Design Experiments and Innovate:** Students will learn to design and conduct experiments, developing innovative solutions to challenges through Research Projects. They will also learn to evaluate their results and refine their experimental approaches over time.
6. **Application of Scientific and Technical Knowledge to Real-World Problems:** Students will use their scientific and Technical knowledge and expertise to identify and solve real-world problems. This would involve applying theoretical concepts to practical situations, bridging the gap between classroom learning, Industry-academia and real-life applications.
7. **Communicating Scientific Findings Effectively:** Students will develop the ability to communicate scientific information clearly and effectively, both in writing and verbally. Whether presenting research findings or writing technical reports, clear communication is key to knowledge sharing and collaboration.
8. **Foster an Interdisciplinary Approach:** Students will cultivate leadership and teamwork skills, enabling them to collaborate effectively in diverse, interdisciplinary teams. Leadership qualities such as decision-making and delegation will help them achieve successful outcomes in various projects.
9. **Promote Environmental Sustainability:** Students will understand the environmental impact of scientific activities and advocate for sustainable practices. By considering environmental factors in their work, they will contribute to the long-term health of the planet.

10. **Enhanced Skills for Entrepreneurship and Employability:** Students will be equipped with essential skills for entrepreneurship and employability, focusing on job readiness, soft skills, and practical business knowledge. Hands-on experience through internships and mentorship will further enhance their readiness for the job market and entrepreneurial ventures.

2. Course Objective: The main objectives of the course are-

- 1) Graduates will build a strong foundation in core statistical principles and methodologies, allowing them to apply statistical techniques to real-world problems.
- 2) Graduates will commit to lifelong learning and remain updated on emerging statistical methods, computational tools, and technological advancements.
- 3) Graduates will develop problem-solving and analytical skills, enabling them to apply statistical models and methods in various industries, including finance, healthcare, and research.
- 4) Graduates will demonstrate professionalism, ethical responsibility, and effective communication in multidisciplinary and collaborative environments.
- 5) Graduates will possess the skills necessary for employability and entrepreneurship, and they will be well-prepared to pursue higher education or careers in data analysis, statistical consulting, or related fields.

3. Process adopted for curriculum designing.

The members of Department of Statistics initially drafted the syllabus. The draft syllabus was shown to Industry Partners, Academic Partners and Research Institute Partners through mail and in person invited to college. They suggested some changes. These changes were incorporated.

4. Salient features, how it has been made more relevant.

Statistics deals with collection, organization, analysis and interpretation of data. Statistical knowledge is very important as it helps to use appropriate methodologies for collecting data, tools for employing analysis and interpretation of results. It also provides us with techniques which are important in designing and planning of experiments.

A lot of data is generated at each and every moment. Data literacy has become crucial and indispensable to the society. Statistics has the quality of quantifying and measuring uncertainty

which helps in assessing risk. It helps in extracting the meaningful information from the data, making predictions and taking decisions. Study of data has become an integral part of education, business, and overall human progress. This has put Statistics on the center stage of teaching, research, policy making and development all over the globe.

The F.Y.B.Sc Statistics syllabus is a Choice based credit system comprising of one paper having three units each in both the semesters. The current course is designed to enhance the knowledge of the subject. While designing of the syllabus care has been taken to balance the fundamental techniques of Statistics with soft skills like analysis using Statistical Software.

Several radical changes have been made in the syllabi. Firstly, the concept of having separate papers in practicals has been abandoned. All the numerical / practical work has been integrated with the teaching of theory courses. Secondly, majorly the numerical /practical work be carried out on computers.

The course would give the students option to develop skills in areas which have direct relevance to employability in insurance and finance industries, banks, econometrics, quality control, pharmaceutical, medical statistics, agricultural statistics, weather forecasting, civil services, stock market, machine learning and artificial intelligence related job opportunities in Statistics.

5. Program-Specific Outcomes (PSOs)

1. **Comprehension of Statistical Concepts:** Develop a strong foundational understanding of probability, statistical distributions, and inference.
2. **Analytical and Critical Thinking:** Apply statistical techniques to analyze data, solve problems, and make informed decisions.
3. **Technological Proficiency:** Gain hands-on expertise in statistical software such as R, Python, Excel, and Power BI for data analysis and visualization.
4. **Data Interpretation and Experimentation:** Design experiments, interpret data meaningfully, and apply statistical reasoning in real-world applications.
5. **Interdisciplinary Integration:** Collaborate with other disciplines and integrate statistical methods for diverse applications like business, health, and environment.
6. **Communication and Collaboration:** Communicate statistical concepts effectively in both written and verbal formats, fostering teamwork and interdisciplinary collaboration.
7. **Ethics and Sustainability:** Ensure ethical data handling, emphasizing sustainable practices and societal responsibility.

6. Learning Outcomes: Learning Outcomes of various courses are

Semester I

1. Descriptive Statistics (STA101B)

- **Unit I:** Identify and classify data types; use graphical methods to summarize data.
- **Unit II:** Calculate measures of central tendency, dispersion, skewness, and kurtosis.
- **Unit III:** Analyze relationships between two variables using correlation, Cross Tab and regression and Learn Interpretation.

2. Introduction to Spreadsheet (STA101C)

- **Unit I:** Learn Excel functionalities for data entry, manipulation, and basic visualizations.
- **Unit II:** Learn Multiple Spreadsheets & Functions

3. Introduction to Power BI (STA103D)

- Understand Power BI basics; create dashboards and interactive visualizations.
-

Semester II

1. Probability and Probability Distributions (STA102B)

- **Unit I:** Understand basic probability, conditional probability, and Bayes' theorem.
- **Unit II:** Learn Concepts and Analyze discrete random variables and associated standard distributions viz Binomial, Poisson and Uniform.
- **Unit III:** Study Concepts of continuous random variables and associated standard distributions viz Normal, Exponential and Exponential.

2. Data Analysis and Visualization using Spreadsheets (STA102C)

- **Unit I:** Use advanced spreadsheet techniques for statistical computations.
- **Unit II:** Use advanced spreadsheet techniques to create advanced plots and dashboards using real-life datasets.

3. Power BI: Advanced Visualization Techniques (STA104D)

- Create advanced plots and dashboards using real-life datasets.

Part 2- The Scheme of Teaching and Examination is as under: Semester – I Summary

Sr. No.	Choice Based Credit System		Subject Code	Remarks
1	Core Course (Statistics)		STA101B STA101D	Nil
2	Elective Course	Discipline Specific Elective (DSE) Course		
		2.1	Interdisciplinary Specific Elective (IDSE) Course	
		2.2	Dissertation/Project	
		2.3	Generic Elective (GE) Course	STA101C
3	Ability Enhancement Courses (AEC)			
	Skill Enhancement Courses (SEC)		STA103D	

First Year Semester I Internal and External Detailed Evaluation Scheme

Sr. No.	Subject Code	Subject Title	Periods Per Week						Seasonal Evaluation Scheme		Total Marks
			Units	S. L.	L	T	P	Credit	S. L. E	SA	
1	STA101B	DESCRIPTIVE STATISTICS	3	20%*	3	0	0	3	15	60	100
	STA101D	Practical Based on STA101B	1		0	0	2	1		25	
2	STA101C	Introduction to Spreadsheet	2		0	0	2	2		50	50
3	STA103D	Introduction to Power BI and Basic Visualizations	1		0	0	1	1		25	25
								04			175

*One to two lectures to be taken for CONTINUOUS self-learning Evaluation.

First Year Semester I - Units – Topics – Teaching Hours

S. N	Subject Code	Subject Unit Title		Hours/Lectures	Total No. of hours/lectures	Credit	Total Marks
1	STA101B	I	Types of Data and Data Condensation	15	45 H	3	100
		II	Measures of Central Tendency, Dispersion, Skewness & Kurtosis	15			
		III	Bivariate Correlation and Regression Analysis	15			
	STA101D	I	Practical based on STA101B	30	30H	1	

2	STA101C	I	Basics of Spreadsheets	15	30H	2	50
		II	Multiple Spreadsheets & Functions	15			
3	STA103D	I	Introduction to Power BI and Basic Visualizations	15	15H	1	25
TOTAL						4	175

- **Lecture Duration – One hour**
- **One Credit =15 class room teaching hours.**

L: Lecture: Tutorials P: Practical Ct-Core Theory, Cp-Core Practical, SLE- Self learning evaluation CT-Commutative Test, SEE- Semester End Examination , PA-Project Assessment, AT- Attendance

Part -3 Detailed Scheme Theory

Curriculum Topics along with Self-Learning topics - to be covered, through self-learning mode along with the respective Unit. Evaluation of self-learning topics to be undertaken before the concluding lecture instructions of the respective UNIT

Course Code: STA101B

Course Title: Descriptive Statistics

Statistics Major Paper I		
	Pre-requisites: Knowledge of basic concepts of mathematics from XII standard.	
Unit	Content	No. of Hours
I	<p>Types of Data and Data Condensation</p> <p>Types of Data:</p> <p>1.1 Introduction, Definition and scope of Statistics</p> <p>1.2 Types of Characteristics, Different types of scales: nominal, ordinal, interval and ratio.</p> <p>1.3 Collection and Scrutiny of Data</p> <p>1.3.1 Primary data: concept of a questionnaire and a schedule and distinction between them, designing a questionnaire and checking their consistency.,</p> <p>1.3.1 Secondary data- their major sources including some government publications</p> <p>1.4 Types of data: Qualitative and quantitative data; Time series data and cross section data, discrete and continuous data.</p> <p>Classification of Data:</p> <p>1.5 Presentation of Data: Construction of tables with one or more factors of classification and Tabulation.</p> <p>1.6 Univariate frequency distribution of discrete and continuous variables. Cumulative frequency distribution.</p> <p>1.7 Diagrammatic and graphical representation of Histogram, Boxplot, Stem and leaf Diagram, Ogives, Frequency Polygon and Curve</p>	15H

II	<p>Measures of Central Tendency, Dispersion, Skewness & Kurtosis:</p> <p>Measures of Central Tendency</p> <ol style="list-style-type: none"> 1. Measures of Central Tendency <ol style="list-style-type: none"> 1.1 Concept of central tendency of data. Requirements of good measure. 1.2 For Raw Data, Ungrouped Frequency Data, Grouped Frequency Data: 1.3 Mathematical averages Arithmetic mean (Simple, weighted mean, combined mean), Geometric mean, Harmonic mean, 1.4 Locational averages: Median, Mode and 1.5 Partition Values: Quartiles, Deciles and Percentiles. Empirical relation between mean, median and mode Merits and demerits of different measures & their applicability 2. Measures of Dispersion, Skewness & Kurtosis <ol style="list-style-type: none"> 2.1 Concept of dispersion. Requirements of good measure. 2.2 Absolute and Relative measures of dispersion: Range, Quartile Deviation, Mean absolute deviation, Standard deviation. 2.3 Variance and Combined variance 2.4 Concept of Skewness and Kurtosis: Measures of Skewness: Karl Pearson's, Bowley's and Coefficient of skewness based on moments, Measure of Kurtosis. 	15H
III	<p>Bivariate Correlation and Regression Analysis</p> <ol style="list-style-type: none"> 3.1 Scatter Diagram, Karl Pearson's Product moment correlation coefficient, it's properties. Spearman's Rank correlation (With and without ties). 3.2 Analysis of Categorical Data: Theory of attributes and consistency of data, independence and association of attributes, measures of association and contingency for 2 x 2 and r x s contingency tables. 3.3 Concept of linear regression. Principle of least squares. Fitting a straight line by method of least squares. Properties of regression. 3.4 Relation between regression coefficients and correlation coefficient. 3.5 Fitting of curves reducible to linear form by transformation. Concept and use of coefficient of determination. 3.6 Fitting a quadratic curve by method of least squares. 	15H

Self-Learning topics (Unit wise)	
Unit	Topics
1	1.1 Introduction to definition of Statistics
1	1.3 Types of data: Qualitative and quantitative data
1	1.4 Different types of scales: nominal, ordinal, interval and ratio.
2	2.2.3 Mathematical averages Arithmetic mean (Simple, weighted mean, combined mean), Geometric mean, Harmonic mean
2	2.2.4 Locational averages: Median, Mode.
2	2.2.5 Partition Values: Quartiles, Deciles and Percentiles.

Online Resources

‘Business Statistics’ by Dr Mukesh Kumar Barua from IIT Roorkee available on the Swayam portal, <https://nptel.ac.in/courses/110/107/110107114/> for US-FST-101 for unit I and unit II

Course Code: STA101C

Course Title: Introduction to Spreadsheet

Unit	Content	No. of Hours
1	<p>Basics of Spreadsheets</p> <p>1.1 Creating and Navigating worksheets</p> <p>1.2 Adding information to worksheets</p> <p>1.3 Types of data, entering different types of data such as texts, numbers, dates, functions.</p> <p>1.4 Importing data from different sources</p> <p>1.5 Quick way to add data Auto complete, Autocorrect, Auto fill, Auto fit.</p> <p>1.6 Undo and Redo.</p> <p>1.7 Moving data</p> <p>1.8 Contiguous and Non-contiguous selections</p> <p>1.9 Selecting with keyboard</p> <p>1.10 Cut-Copy, Paste</p> <p>1.11 Adding and moving columns or rows</p> <p>1.12 Inserting columns and rows.</p> <p>1.13 Find and replace values</p> <p>1.14 Spell check.</p> <p>1.15 Formatting cells, Numbers, Date, Times, Font, Colours, Borders, Fills.</p> <p>1.16 Adding, removing, hiding, and renaming worksheets.</p> <p>1.17 Add headers/Headers to a Workbook</p> <p>1.18 Page breaks and Preview.</p> <p>1.19 Creating formulas, Inserting functions</p> <p>1.20 Cell references: Absolute, Relative</p>	15

2	<p>Multiple Spreadsheets & Functions</p> <p>2.1 Creating and using templates, using predefined templates, Adding protection option.</p> <p>2.2 Creating and Linking Multiple Spreadsheets.</p> <p>2.3 Using formulas and logical operators.</p> <p>2.4 Creating and using named ranges.</p> <p>2.5 Creating Formulas that use reference to cells in different worksheets.</p> <p>2.6 Financial functions: FV, PV, PMT, PPMT, IPMT, NPER, RATE</p> <p>2.7 Mathematical functions: ROUND, ROUNDDOWN, ROUNDUP, CEILING, FLOOR, INT, MAX, MIN, MOD, SQRT, ABS, SUM, COUNT</p> <p>2.8 Database Functions LOOKUP, VLOOKUP, HLOOKUP</p> <p>2.9 Boolean Functions: AND, OR, NOT</p> <p>2.10 Conditional Logic functions IF, Nested IF, COUNTIF, SUMIF, AVERAGEIF</p> <p>2.11 String functions LEFT, RIGHT, MID, LEN, UPPER, LOWER, PROPER, TRIM, FIXED</p> <p>2.12 Date functions TODAY, NOW, DATE, TIME, DAY, MONTH, YEAR, WEEKDAY, DAYS360</p> <p>2.13 Statistical Functions COUNTA, COUNTBLANK, LARGE, SMALL, AVERAGE, CORREL, STDEV, MEDIAN, MODE</p>	15
---	--	----

Course Code: STA103D

Course Title: Introduction to Power BI and Basic Visualizations

Unit	Content	No. of Hours
1	<p>Introduction to Power BI and Basic Visualizations:</p> <p>1.1 Overview of Power BI and its features</p> <p>1.2 Understanding the Power BI ecosystem</p> <p>1.3 Installing Power BI Desktop</p> <p>1.4 Introduction to Power BI Service and Power BI Mobile</p> <p>1.5 Importing data into Power BI</p> <p>1.6 Connecting to various data sources (Excel, CSV, databases, etc.)</p> <p>1.7 Data preparation and transformation</p> <p>1.8 Overview of Power Query for data cleaning</p> <p>1.9 Understanding the importance of data presentation</p> <p>1.10 Using colors, fonts, and layouts effectively</p> <p>1.11 Creating Visualizations in Power BI: Overview of Power BI visuals and their types, Building basic visuals (bar charts, line charts, pie charts etc)</p> <p>1.12 Formatting visuals (colors, labels, titles, etc.)</p> <p>1.13 Working with tables and matrices</p> <p>1.14 Using interactive filters and slicers</p> <p>1.16 Creating drill-through and drill-down experience</p>	15

Part -4 Detailed Scheme Practicals

Course Code: STA101D

Total Credit:

01

Course Title: Computer Applications and Practical_

Content	No. of Lectures
<p>Course Outcome: Introduce students</p> <ol style="list-style-type: none">1. To understand and practice the analytical skills while doing statistics practical.2. To understand the concepts of data visualization, probability theory, etc.3. To understand the concepts of univariate and bivariate random variables and theory probability distributions.4. To equip the student with nature and properties of various discrete and continuous probability distributions.5. To equip the student with statistical problem solving and statistical interpretation skills through practical problems.6. To analyze practical implication of statistical techniques used to measure association between two variables.7. To apply correlation, regression between two variables <ol style="list-style-type: none">1) Univariate Visualization :Constructing and interpreting Histogram, Piechart,-Stem and leaf diagram, Box Plot-2) Bivariate Visualization: Constructing and interpreting Ogives, Frequency Curve and Frequency Polygon3) Classification and Frequency Distribution4) Tabulation5) Mathematical Averages6) Positional Averages7) Measures of dispersion8) Skewness and Kurtosis9) Correlation10) Analysis of Categorical Data11) Regression12) Fitting of Curves	<p>02 Hours per Practical per Batch</p>

References:

Course Code: **STA101B**

Course Title: Descriptive Statistics

- 1) **Medhi J.:** *Statistical Methods, An Introductory Text*, Second Edition, New Age International Ltd.
- 2) **Spiegel M. R.:** *Theory and Problems of Statistics*, Schaum's Publications series, Tata McGraw-Hill.
- 3) **Kothari C. R.:** *Research Methodology*, Wiley Eastern Limited.
- 4) **Goon A. M., Gupta M. K., Dasgupta B.:** *Fundamentals of Statistics*, Volume II, The World Press Private Limited, Calcutta.

Course Code: **STA101C**

Course Title: Introduction to Spreadsheet

1. Excel 2019 Advanced Topics: Leverage More Powerful Tools to Enhance Your Productivity (Excel 2019 Mastery), George, Nathan

Software Requirements:

1. Office 2016 or Higher Version

Course Code: **STA103D**

Course Title: Introduction to Power BI and Basic Visualizations

- Power Pivot and Power BI by Rob Collie and Avichal Singh
- Data Analytics With MS Excel & Power BI by Punit Prabhu

Software Requirements:

- Power BI Desktop

Part 5- The Scheme of Teaching and Examination is as under:First Year Semester -II Summary

Sr. No.	Choice Based Credit System		Subject Code	Remarks
1	Core Course (Statistics)		STA102B STA102D	
2	Elective Course	Discipline Specific Elective (DSE) Course		
		2.1	Interdisciplinary Specific Elective (IDSE) Course	
		2.2	Dissertation/Project	
	2.3	Generic Elective (GE) Course	STA102C	
3	Ability Enhancement Courses (AEC)			
	Skill Enhancement Courses (SEC)		STA104D	

First Year Semester -II Internal and External Detailed Evaluation Scheme

Sr. No.	Subject Code	Subject Title	Periods Per Week						Seasonal Evaluation Scheme		Total Marks
			Units	S. L.	L	T	P	Credit	S. L. E	SA	
1	STA102B	Probability and Probability Distributions	3	20%*	3	0	0	3	15	60	100
	STA102D	Practical Based on STA102B	1		0	0	2	1		25	
2	STA102C	Data Analysis and Data Visualization using Spreadsheet	2		0	0	2	2		50	50
3	STA104D	Power BI: Advanced Visualization Techniques	1		0	0	1	1		25	25
								04			175

***One to two lectures to be taken for CONTINUOUS self -learning Evaluation.**

First Year Semester – II Units – Topics – Teaching Hours

S. N	Subject Code	Subject Unit Title		Hours/Lectures	Total No. of hours / lectures	Credit	Total Marks
1	STA102B	I	Probability	15	45 H	3	100
		II	Discrete random variable and Standard Discrete Distributions	15			
		III	Continuous Random Variable and Standard Continuous Distributions	15			
	STA102D	I	Practical based on STA102B	30	30H	1	
2	STA102C	I	Data Analysis	15	30H	2	50
		II	Data Visualization	15			
3	STA104D	I	Power BI: Advanced Visualization Techniques	15	15H	1	25
TOTAL						4	175

● **One Credit =15 Classroom hours**

L: Lecture: Tutorials P: Practical Ct-Core Theory, Cp-Core Practical, SLE- Self learning evaluation CT-Commutative Test, SEE- Semester End Examination , PA- Project Assessment, AT- Attendance

Part -6 - Detailed Scheme Theory

Curriculum Topics along with Self-Learning topics - to be covered, through self-learning mode along with the respective Unit. Evaluation of self-learning topics to be undertaken before the concluding lecture instructions of the respective UNIT

Course Code: STA102B

Course Title: Probability and Probability Distributions

STATISTICS MAJOR PAPER II		
	Pre-requisites: Knowledge of basic concepts of mathematics from XII standard.	
Unit	Content	No. of Hours
I	Probability: Introduction, random experiments, sample space, events and algebra of events. Definitions of Probability – classical, statistical, and axiomatic. Conditional Probability, laws of addition and multiplication, independent events, theorem of total probability, Bayes’ theorem and its applications.	15
II	Concept of Discrete random variable and properties of its probability distribution	

	<p>2.1 Random variables and Types of random variables. Definition and properties of probability mass function and cumulative distribution function of discrete random variable.</p> <p>2.2 Central and Non- Central moments (definition only) and their relationship (up to order four). Calculation of moments, Skewness and Kurtosis using probability concept.</p> <p>2.3 Expectation of a random variable. Theorems on Expectation & Variance.</p> <p>Some Standard Discrete Distributions</p> <p>2.4 Discrete Uniform, Binomial, Poisson distributions and Hypergeometric distribution: derivation of their mean and variance.</p> <p>2.5 Recurrence relation for probabilities of Binomial and Poisson distributions. Poisson approximation to Binomial distribution, Binomial approximation to hypergeometric distribution.</p>	
III	<p>Continuous Random Variable</p> <p>Concept of Continuous random variable and properties of its probability distribution.</p> <p>Probability density function and cumulative distribution function. Their graphical representation.</p> <p>Expectation of a random variable and its properties. Measures of location, dispersion and kurtosis. Raw and central moments (simple illustrations).</p> <p>Some Standard Continuous Distributions</p> <p>Rectangular and Normal distribution.</p> <p>Derivations of Cumulative distribution function, mean, median and variance for Uniform distribution</p> <p>Properties of Normal distribution (without proof). Additive Property of Normal Distribution (Statement only). Properties of normal curve. Use of standard normal tables.</p> <p>Normal approximation to Binomial and Poisson distribution (Statement and Problems only).</p> <p>Exponential, memory less property of exponential distribution.</p> <p>Derivations of Cumulative distribution function, mean, median and variance for Exponential distribution</p>	

Self-Learning topics (Unit wise)

Unit	Topics
3	<p>Standard probability distributions: MGF and CGF of Binomial, Poisson, geometric, negative binomial, hypergeometric, uniform, normal, exponential, Cauchy, beta and gamma distributions.</p>

Online Resources

“Probability and Statistics” by Prof. Somesh Kumar, Department of Mathematics, IIT Kharagpur
 Source : <https://nptel.ac.in/courses/111/105/111105090/> on unit 3.

Course Code: STA102C

Course Title: Data Analysis and Data Visualization Using Spreadsheet

Unit	Content	No. of Hours
1	<p>Data Analysis</p> <p>1.1 Sorting, Subtotal. 1.2 Pivot Tables- Building Pivot Tables, Pivot Table regions, Rearranging Pivot Table 1.3 Filter with customized condition. 1.4 Scenario Manager: Scenarios, creating and managing a scenario. 1.5 Goal Seek 1.6 Solver: Using solver to solve the optimization problems. 1.7 Measures of Central Tendency: Arithmetic mean (Simple & Weighted), Median, Mode Geometric mean –Harmonic mean, Applications, Merits and Limitations, Interpretation of these measures. 1.8 Measures of Dispersion (Absolute and Relative): Importance, Range, Quartile Deviation, Mean Deviation, Standard Deviation, Coefficient of Skewness and Kurtosis 1.9 Correlation and regression: Scatter diagram, Coefficient of Correlation, Scatter Plot, Fitting Regression Line, Coefficient of Determination 1.10 Time Series Analysis: Moving averages, Trend Fitting, Simple Exponential Smoothing</p>	15
2	<p>Data Visualization</p> <p>2.1 The Graphical representation of data: Bar, Column, Line, Pie and Bar charts 2.2 Area Chart 2.3 Scatter Plot 2.4 Hierarchy Charts: Treemap, Sunburst 2.5 Histogram, Box and Whisker Plot 2.6 Waterfall, Funnel, Stock, Surface and Radar Charts 2.7 Map Charts 2.8 Combo Charts 2.9 Stem and Leaf Plot 2.10 Frequency Polygon and Frequency Curve 2.11 Cumulative Frequency Curves 2.12 Curve Fitting: Linear, Quadratic, Logarithmic and Exponential 2.13 Pivot Charts</p>	15

Course Code: STA104D

Course Title: Power BI: Advanced Visualization Techniques

	Learning Objectives: This course is design to equip students from basic to advance visualization techniques.	
2	<p>Advanced Visualization Techniques:</p> <p>2.1 Advanced visualizations (tree-maps, scatter plots, maps, etc.)</p> <p>2.2 Adding custom visuals from Power BI market place</p> <p>2.3 Implementing custom formatting using themes</p> <p>2.4 Applying conditional formatting for data highlighting</p> <p>2.5 Power Query and data transformation</p> <p>2.6 Advanced DAX functions and calculations</p> <p>2.7 Creating calculated tables and advanced relationships</p> <p>2.8 Power BI data modeling best practices</p> <p>2.9 Designing Interactive Dashboards</p> <p>2.10 Utilizing bookmarks for navigation and storytelling</p> <p>2.11 Adding and configuring slicers for interactivity</p> <p>2.12 Introduction to bookmarks, buttons, and tooltips</p> <p>2.13 Using Q&A feature for natural language queries</p> <p>2.14 Introduction to Power BI mobile reports</p> <p>2.15 Data exploration and filtering</p> <p>2.16 Aggregating and summarizing data</p> <p>2.17 Applying statistical functions</p> <p>2.18 Time intelligence calculations</p> <p>2.19 Analyzing trends and patterns in data</p>	15

Part – 7- Detailed Scheme Practical

Course Code: STA102D

Total Credit: 01

Title of Paper: Probability and Probability Distribution using Statistical Software

Unit	Content	No. of Lectures
I	<p>Course Outcome: Introduce students</p> <p>1. To understand and practice the analytical skills while doing statistics practical.</p> <p>2. To understand the concepts of data visualization and probability theory, etc.</p> <p>3. To understand the concepts of univariate and bivariate random variables and theory probability distributions.</p> <p>4. To equip the student with nature and properties of various discrete and continuous probability distributions.</p> <p>5. To equip the student with statistical problem solving and statistical interpretation skills through practical problems.</p> <p>6. To understand probability distribution of discrete and continuous random variables also to study its properties like</p>	02Hours per Practical per Batch

	<p>mean, variance, etc.</p> <ol style="list-style-type: none"> 1) Elementary Probability 2) Advance Probability 3) Random Variables 4) Plotting of pmf, pdf and cdf of standard probability distributions 5) Application problems based on binomial and Poisson distribution. 6) Application problems based on Hypergeometric and uniform distribution. 7) Fitting of Discrete Distributions <ol style="list-style-type: none"> a. Fitting of binomial distributions for n and $p = q = \frac{1}{2}$ and Fitting of binomial distributions for given n and p. b. Fitting of binomial distributions after computing mean and variance. c. Fitting of Poisson distributions for given value of λ and d. Fitting of Poisson distributions after computing mean. e. Fitting of negative binomial. 8) Problems based on area property of normal distribution and to find the ordinate for a given area for normal distribution. 9) Application based problems using normal distribution. Fitting of normal distribution when parameters are given and Fitting of normal distribution when parameters are not given 10) Application problems based on exponential, Uniform distribution 	
--	--	--

Reference:

Course Code: : STA102B

Course Title: Probability and Probability Distributions

- 1) **David S.:** *Elementary Probability*, Cambridge University Press.
- 2) **Hogg R. V. and Tannis E. P.:** *Probability and Statistical Inference*, McMillan Publishing Co. Inc.
- 3) **Miller I. & Miller M.:** *John E. Freund's Mathematical Statistics with Applications*, 7th Edition, Pearson Education Asia.
- 4) **Gupta, S. C. and Kapoor, V. K.:** *Fundamentals of Mathematical Statistics*, Eighth Edition, Sultan Chand and Sons Publishers, New Delhi.

Course Code: STA102C

Course Title: Data Analysis and Data Visualization Using Spreadsheet

1. Excel 2019 Advanced Topics: Leverage More Powerful Tools to Enhance Your Productivity (Excel 2019 Mastery), George, Nathan

Software Requirements:

2. Office 2016 or Higher Version

Course Code: STA104D

Course Title: Power BI: Advanced Visualization Techniques

- Power Pivot and Power BI by Rob Collie and Avichal Singh
- Data Analytics With MS Excel & Power BI by Punit Prabhu

Software Requirements:

- Power BI Desktop

The Scheme of Teaching and Examination:

The performance of the learners shall be evaluated in two components: Internal Assessment with 40% marks by way of continuous evaluation and by Semester End Examination with 60% marks by conducting the theory examination.

Examination Pattern for First Year Degree as per NEP 2020 Academic Year 2023-2024**1) Evaluation of Major and Minor Subjects**

Subject	Formative Assessment (Marks)	Summative Assessment (Marks)
Major Subject	40	60
Minor Subject	40	60
Major (Practical based Subject)	-	25
Minor (Practical based Subject)	-	25
General Elective (GE/OE)	20	30

FORMATIVE ASSESSMENT:- It is defined as the assessment of the learners on the basis of continuous evaluation as envisaged in the credit based system by way of participation of learners in various academic and correlated activities in the given semester of the programme.

A). Formative Assessment – 40

B). Practical's (internal Components of the Practical Course)

40 marks

1. For Theory Courses

Sr. No.	Particulars	Marks
1	Self-Learning Evaluation with Active participation in routine class instructional deliveries	10+5 Marks

2. For Courses with Practicals

Each practical course can be conducted out of 50 marks with 10 marks for internal **component of the Practical** and 40 marks for formative assessment which will be converted to 25 marks.

Practical's (Internal component of the Practical Course)

Sr. No	Evaluation type	Marks
1	Journal	5
2	Viva	5

C). SUMMATIVE ASSESSMENT =SEMESTER END EXAMINATION :-
It is defined as the examination of the learners on the basis of performance in the semester end theory / written examinations.

The semester end examination (external component) of 60 % for each course will be as follows:

a. For Theory Courses

i) Duration – 2 Hours

Theory Question Paper pattern for Main Papers of 60 Marks:

Sr. No.	All questions are Compulsory.	Marks
Q. 1	Attempt either a & b or p & q based on unit 1.	15
Q. 2	Attempt either a & b or p & q based on unit 2.	15
Q. 3	Attempt either a & b or p & q based on unit 3.	15
Q. 4	Attempt either a & b or p & q based on all three units.	15

ii) Duration – 1 Hours

Theory Question Paper pattern for GE/ OE Papers of 30 Marks:

Sr. No.	All questions are Compulsory.	Marks
Q. 1	a) Attempt any 5 out of 7 based on MCQ based on all units. b) Attempt any 5 out of 7 based on True/False based on all units.	10
Q. 2	Attempt any two out of four based on unit 1.	10
Q. 3	Attempt any two out of four based on unit 2.	10

b. For Practical Courses

Duration – 2 Hours

Practical Question Paper Pattern:-

Sr. No.	All questions are Compulsory.	Marks
Q. 1	Attempt any two out of three based on unit 1.	10
Q. 2	Attempt any two out of three based on unit 2.	10
Q. 3	Attempt any two out of three based on unit 3.	10
Q. 4	Attempt any two out of three based on all three units using Statistical Software.	10