B. Sc. Semester – I – Statistics (Multi Disciplinary)
Basics of Data Science

Prerequisite:

BSc Statistics is an undergraduate course that deals primarily with statistics, probability, and permutations. Students who are thinking of pursuing a BSc Statistics must have completed 10+2 in the fields of Mathematics, Physics and Chemistry. A fundamentally sound knowledge of limit and continuity, derivatives - Total and partial, and integration is desirable. Students must have basic know how of numerical or qualitative information, methods of collecting numerical or qualitative information through attributes, graphical presentation and some primary measures, like arithmetic mean, median, mode.

Co-requisite

Fundamental knowledge about use of scientific calculator and functionality of computers is necessary. Knowledge and basic understanding of MS – Office is recommended.

Vision and Outcome
The aim of introducing statistics as a subject and scientific tool as well, at an undergraduate level is to provide students a strong theoretical foundation, which is on par with other institutions and colleges with reputation of national level. At the same time, enough care is taken to emphasize on the course contents that enhance the ability of students to gain knowledge of open-source statistical software. This enables students' understanding in dealing with real life problems from statistical viewpoint. The weightage is given to fieldwork and projects that make students develop statistical thinking and work independently.

Outcomes	
Programme Outcome	Students will demonstrate an understanding of major concepts in statistics. Students tend to think critically and apply their understanding to develop ability to design, collection, presentation, analyse and interpret of data based problems of real life situations.
Programme Specific Outcome	The ability to identify type of observable phenomena and probability distributions that are associated with observable phenomena. This helps them to collect the relevant data and to verify different properties of associated probability distribution. The design and execution of the proper statistical analysis reveals their understanding of good analytical skills and proper handling of statistical data.

Course	Outcomes
Statistics **MDC-STA-114T** **(A+B)** **(Basics of Data Science – I)**	The main objective of this course is to acquaint students with some basic concepts in Statistics. They will be introduced to some elementary statistical methods of analysis of data. At the end of this course students are expected to be able, 1. to tabulate statistical information given in descriptive form. 2. to use graphical techniques and interpret. 3. to compute various measures of central tendency, dispersion, skewness and kurtosis. 4. to analyze data pertaining to attributes and to interpret the results. 5. to summarize and analyze the data using computer. 6. to apply statistics in the various fields.
Statistics **MDC-STA-114P (Practical)**	At the end of the semester, students can identify nature of the problem and type of data to be collected. Also, He/She can ably obtain certain summary statistics in order to present the data in meaningful way. This will help them to understand the other associated methods and procedures used in analysis in a better way. The topics for this paper are based on Theory paper M**DC-STA – 114T**

Course Structure with Credits, Lecture Hours and Marks

Course Code	Course Title	Credit	Lecture Hours Per Week	Exam Hours	Marks		
					Internal	**External**	**Total**
MDC- STA - 114T-A	**Basics of Data Science-I (Theory)**	2	2	2	25	25	50
MDC-STA - 114P -A	**Basics of Data Science-I (Practical)**	2	4	2	25	25	50

Course Code	Course Title	Credit	Lecture Hours Per Week	Exam Hours	Marks		
					Internal	**External**	**Total**
MDC- STA - 114T-B	**Basics of Data Science-II (Theory)**	2	2	2	25	25	50
MDC-STA - 114P-B	**Basics of Data Science-II (Practical)**	2	4	2	25	25	50

Syllabus for Statistics (UG)
B. Sc. Semester I

MDC-STA-114T - A- Basics of Data Science-I

Hours per week: 2   **Credit: 2**

Unit: 1 Data Types and Classification
Types of data: Primary, Secondary, Internal and External data and their sources.
Designing a questionnaire schedule.
Classification of data: Qualitative, Quantitative: Discrete, Continuous; Chronological
(Time series) data.
Nominal, Ordinal, Interval and Ratio data.
Frequency: grouped and ungrouped data; Construction of frequency and cumulative
frequency distribution.
Presentation of qualitative data: Tabulation (up to four attributes).

Unit: 2 Data visualisation
Graphical representation of grouped data: Histogram, frequency curve, frequency polygon, ogives
(cumulative frequency curves), Diagrammatic representation of data: Bar diagrams- simple Bar,
multiple bar, sub-divided bar and percentage bar diagrams.
Two dimensional diagrams: Rectangles and Pie diagrams. Stem - Leaf plot.
Bivariate : Frequency distribution, Marginal and Conditional frequency distributions.

Reference Books for MDC-STA-114T - A: Basics of Data Science-I
1. Statistical analysis: Graphs and diagrams, S. M. Nair and M. Garg, Spectrum books (P) Ltd,
New Delhi.
2. Introduction to the Practice of Statistics, Moore, S. David; McCabe, P. George
W. H. Freeman and Company, New York.
3. Basic Statistics, Agarwal, B. L., New Age International (P) Ltd.
4. Introduction to the theory of Statistics, Mood, A. M., Greybill, F.A., Boes, D.C.,
McGraw Hill.
5. Fundamentals of Mathematical Statistics, S. C. Gupta and V. K. Kapoor, Sultan
Chand and Sons, New Delhi.

Syllabus for Statistics (UG)
B. Sc. Semester I

MDC-STA-114P – A Basics of Data Science-I (Practical)

Hours per week: 4   **Credit: 2**

---

1. Classification of the variable/data into various category and tabulation.
2. Construction of frequency table (one way and two way).
3. Data visualization: Bar Chart, Divided bar chart, multiple bar chart, pie char, Stem – Leaf chart and box – plot chart
4. Data visualisation: Histogram, frequency curve, frequency polygon, ogives (cumulative frequency curves),

Websites for secondary data sources

www.actuaries.org.uk

www.actuariesindia.org

www.soa.org

FBLD (Flip Blended Learning Design Template)
- Any One Unit from the above syllabus can be discussed by the faculty through online mode.
- Online mode can be SWAYAM MOOC Course or any other suggested by the UGC or Gujarat University.

Syllabus for Statistics (UG)
B. Sc. Semester I

MDC-STA-114T – B – Basics of Data Science-II

Hours per week: 2   **Credit: 2**

Unit: 1 Measures of Central Tendency
Concept of central tendency, various measures of central tendency and their interrelationship. Their merits and demerits. Empirical relation between mean, median and mode. Properties and applications of measures of central tendency. Partition values (quartiles, deciles and percentiles.)

Unit: 2 Measures of Dispersion and Moments
Concept of variation/dispersion, quartile deviation, Absolute and relative measures of dispersion with their merits, demerits and applications.
Moments: raw moments, central moments, factorial moments and their interrelationship.
Skewness, Kurtosis and their measures. Box plot.

Reference Books for Reference Books for MDC-STA-114T – B Basics of Data Science-II

1. Statistical analysis: Graphs and diagrams, S. M. Nair and M. Garg, Spectrum books (P) Ltd, New Delhi.
2. Introduction to the Practice of Statistics, Moore, S. David; McCabe, P. George W. H. Freeman and Company, New York.
3. Basic Statistics, Agarwal, B. L., New Age International (P) Ltd.
4. Programmed Statistics, Agarwal, B. L., New Age International (P) Ltd.
4. Introduction to the theory of Statistics, Mood, A. M., Greybill, F.A., Boes, D.C., McGraw Hill.
5. Fundamentals of Mathematical Statistics, S. C. Gupta and V. K. Kapoor, Sultan Chand and Sons, New Delhi.

Syllabus for Statistics (UG)
B. Sc. Semester I

MDC-STA-114P – B Basics of Data Science-II (Practical)

Hours per week: 4   **Credit: 2**

1. Problems based on measures of central tendency.
2. Problems based on single mean, combined mean for two variables.
3. Problems based on median and mode.
4. Problems based on Measures of variations : Standard deviation, quartile deviation
5. Problems based on various measures of Skewness and Kurtosis.

Websites for secondary data sources

www.actuaries.org.uk

www.actuariesindia.org

www.soa.org

FBLD (Flip Blended Learning Design Template)
   • Any One Unit from the above syllabus can be discussed by the faculty through online mode.
• Online mode can be SWAYAM MOOC Course or any other suggested by the UGC or Gujarat University.