SHRI GNANAMBICA DEGREE COLLEGE: MADANAPALLE

(AUTONOMOUS)

COURSE-2: Statistical Foundations of Data Science

SEMESTER I

(W.E.F.2025 - 26)

Program: B.Sc. Data Science Honors

Hours per week: 4 **Credits: 3**

Course Objectives:

1. To introduce the fundamental concepts and importance of statistics in data science.
2. To understand data types, collection methods, and graphical presentation.
3. To develop the ability to summarize and interpret data using statistical measures.
4. To apply probability theory and random variable concepts to data-driven problems.
5. To understand the role of expectation, moments, skewness, and kurtosis in describing distributions.

Learning Outcomes:

After completion of this course, the student will be able to:

1. Understand the basic concepts, importance, and scope of statistics.
2. Distinguish between types of data and methods of data collection.
3. Summarize, represent, and interpret data through tables, diagrams, and graphs.
4. Apply measures of central tendency, dispersion, skewness, and kurtosis in data analysis.
5. Use probability and random variable theory in real-world data science problems.

UNIT -I

Introduction to statistics and Data Collection: Definition, Meaning, nature and scope of Statistics.

Importance of statistics: Role in data science, business, economics, and research. Functions of Statistics, Limitations of Statistics. Types of data: Qualitative and quantitative.

Primary Data and Secondary data: Methods of collection, merits and demerits. Questionnaire and schedule.

UNIT-II

presentation of Data: Classification and Tabulation of data: Objectives and rules of tabulation.

Graphic presentation: Line graph, pie Chart, histogram, frequency polygon, and Ogive.

Diagrammatic Representation: Simple, Sub-divided, multiple bar diagrams.

Data visualization in Data Science: Introduction to modern tools (R/Python/Excel)

UNIT-III

Measures of Central Tendency and Dispersion: Measures of Central Tendency: Arithmetic Mean, Median, Mode, Geometric Mean, Harmonic Mean. Merits and demerits of each measure.

Measures of Dispersion: Range, Quartile Deviation, Mean Deviation, Standard Deviation, Co-efficient of Variation. Numerical problems using raw and grouped-data.

UNIT-IV

Moments, Skewness and Kurtosis: Moments: Definition, Computation of raw (Non-Central) and Central moments, and their interpretation.

Academic Council ★ Shri Gnanambica Degree College(A) ★ Chairperson

C Mahesh Babu
CHAIRMAN
BOARD OF STUDIES
Shri Gnanambica Degree College (A)
MADANAPALLE - 517 325

Skewness: Meaning, types (Positive, Negative, No skewness) and measures (Karl Pearson's, Bowley's. Skewness based on moments).
Kurtosis: Concept, types of Curves (Leptokurtic, mesokurtic, platykurtic).
Applications in Data Science: Understanding data Symmetry and shape of distributions.

UNIT-V

Probability, Random Variables and Mathematical Expectation:

Probability: Definition (Classical, empirical and axiomatic). Addition and Multiplication theorems, Conditional probability. Boole's inequality, ad Baye's theorem and applications.
Random Variables: Definition, Discrete and Continuous random variables, probability mass function and probability density function.
Mathematical Expectation: Definition, properties, Variance of random Variable, and linearity expectation.

Recommended Books:

1. S.C.Gupta & V.K.Kapoor-Fundamentals of Mathematical Statistics
2. Goon, Gupta & Dasgupta – Fundamentals of Statistics
3. S.P. Gupta - Statistical Methods
4. R.P. Hooda - Statistics for Business and Economics
5. Montgomery & Runger - Applied Statistics and Probability for Engineers
6. James et al. – An Introduction to Statistical Learning

Academic Council
Chairperson
Shri Gnanambica Degree College(A)

C Mahesh Babu
CHAIRMAN
BOARD OF STUDIES
Shri Gnanambica Degree College (A)
MADANAPALLE - 517 325

SHRI GNANAMBICA DEGREE COLLEGE: MADANAPALLE

(AUTONOMOUS)

COURSE-2: STATISTICAL FOUNDATIONS OF DATA SCIENCE

SEMESTER I

(W.E.F.2025 - 26)

Program: B.Sc. Data Science

Question Paper Blue Print

Time : 3 Hrs **Max Marks: 70**

Part-A

Answer any **FOUR** of the following questions. Each question carries FIVE marks **4×5=20**

1. **Question**
2. **Question**
3. **Question**
4. **Question**
5. **Question**
6. **Question**
7. **Question**
8. **Question**

Answer any Five of the following choosing one from each unit **5x10=50**

Unit-1

9. **Question**

or

10. **Question**

Unit-II

11. **Question**

or

12. **Question**

Unit-III

13. **Question**

or

14. **Question**

Unit-IV

15. **Question**

or

16. **Question**

Unit-V

17. **Question**

or

18. **Question**

Academic Council Chairperson Shri Gnanambica Degree College (A)

C. Mahesh Babu

CHAIRMAN
BOARD OF STUDIES
Shri Gnanambica Degree College (A)
MADANAPALLE