

CBCS Syllabus as per NEP 2020 for F.Y.B.Sc. (Computer Science) (2023 Pattern)

Name of the Programme	: B.Sc. Computer Science.
Programme Code	: USCOS
Class	: F.Y.B.Sc.(Computer Science)
Semester	: II
Course Type	: Minor (Theory)
Course Code	: COS-161-MN(A)
Course Title	: Exploratory Data Analysis
No. of Credits	: 02
No. of Teaching Hours	: 30

Course Objectives:

1. To understand the characteristics and examples of each data type.
2. To understand the importance of sampling in statistics.
3. To understand the importance of data visualization.
4. To understand concept of central tendency.
5. To understand the various measures of central tendency in data analysis.
6. To understand concept of dispersion and its types.
7. To understand the various measures of dispersion in data analysis.

Course Outcomes:

By the end of the course, students will be able to:

- CO1.** Identify and categorize different types of data accurately.
- CO2.** Recognize real-world examples of structured, unstructured, and semi-structured data.
- CO3.** Apply various measures of central tendency on real data with usual interpretation.
- CO4.** Apply various measures of dispersion on the data with usual interpretation.
- CO5.** Understand Concept of inventory which is used in software development.
- CO6.** Grasp the significance of dispersion measures in data analysis.
- CO7.** Understand the concept of inventory and its significance in various industries.

Topics and Learning Points

UNIT 1: Introduction to Data

(8L)

- 1.1 Types of Data: Structured, Unstructured and semi structured.
- 1.2 Qualitative and Quantitative Data.
- 1.3 Data Levels of measurement: Nominal, ordinal, interval and ratio scale.

1.4 Data Condensation: Types of data (Primary and secondary), Attributes and Variables, Discrete and Continuous variables, classification and construction of frequency distribution.

1.5 Graphical Representation: Histogram, Frequency polygon, Frequency curve, Ogive Curves, Steam and leaf chart.

UNIT 2: Sampling (4L)

2.1 Population and Sample.

2.2 Concept of population and sample with illustration.

2.3 Methods of sampling –SRSWR, SRSWOR, Stratified, Systematic (Description of sampling procedures only).

UNIT 3: Measures of central tendency (6L)

3.1 Concept of central tendency, requisites of good measures of central tendency.

3.2 Arithmetic mean: Definition, computation for ungrouped and grouped data, combined mean, weighted mean, merits and demerits.

3.3 Median and Mode: Definition, formula for computation for ungrouped and grouped data, graphical method, merits and demerits.

3.4 Quartiles: Definition, formula for computation for ungrouped and grouped data.

3.5 Numerical problems related to real life situations.

UNIT 4: Measures of Dispersion (6L)

4.1 Concept of dispersion and measures of dispersion, requisites of good measures of dispersion, absolute and relative measures of dispersion.

4.2 Range and Quartile Deviation: definition for ungrouped and grouped data and their coefficients, merits and demerits.

4.3 Variance and Standard deviation: definition for ungrouped and grouped data, coefficient of variation, combined variance & standard deviation, merits and demerits.

4.4 Numerical problems related to real life situations.

UNIT 5: Inventory Control (6L)

5.1 Introduction.

5.2 Principal items of inventories.

5.3 Need and importance of inventory.

5.4 ABC analysis.

5.5 Basic parameters to operate the inventory system: Maximum level, Re-order level, Minimum level.

5.6 The pay off matrix approach to select an optimal strategy (EMV criterion).

5.7 Numerical problems related to real life situations.

References:

1. An Introductory Statistics, Kennedy and Gentle.
2. Fundamentals of Applied Statistics (3rd Edition), Gupta and Kapoor, S.Chand and Sons, New Delhi, 1987.
3. Kantiswroop, Gupta : Operations Research,Sultan Chand and Sons Publisher.
4. Mukhopadhy Parimal (1999): Applied Statistics, New Central Book Agency, Pvt. Ltd. Calcutta.11.
5. Statistical Methods, G.W. Snedecor, W.G. Cochran, John Wiley & sons, 1989.

Course Outcomes	Programme Outcomes (POs)						
	PO1	PO2	PO3	PO4	PO5	PO6	PO7
CO1	3	3	3	1	1	2	2
CO2	3	2	3	1	1	2	2
CO3	3	2	3	1	1	2	2
CO4	3	3	3	1	1	2	2
CO5	3	3	3	1	1	2	2
CO6	2	3	3	1	1	2	2
CO7	2	2	3	1	1	2	2

Weight: 1 - Partially related 2 - Moderately Related 3 - Strongly related

PO1 Computer knowledge

CO1: 3 (Strongly Related)

CO2: 3 (Strongly Related)

CO3: 3 (Strongly Related)

CO4: 3 (Strongly Related)

CO5: 3 (Strongly Related)

CO6: 3 (Strongly Related)

CO7: 3 (Strongly Related)

CO8: 2 (Moderately Related)

CO9: 2 (Moderately Related)

CO10: Not provided

Justification: All objectives (CO1 to CO9) require the use of computer tools and software to compute, analyze, and interpret statistical data, thus strongly related to computer knowledge.

PO2 Design / Development of solution

CO1: 3 (Strongly Related)

CO2: 2 (Moderately Related)

CO3: 2 (Moderately Related)

CO4: 3 (Strongly Related)

CO5: 3 (Strongly Related)

CO6: 3 (Strongly Related)

CO7: 3 (Strongly Related)

CO8: 2 (Moderately Related)

CO9: 2 (Moderately Related)

CO10: Not provided

Justification: Designing and developing statistical solutions (CO1 to CO9) involves planning, organizing, and implementing strategies to address statistical problems effectively, thus strongly related to the design and development of solutions.

PO3 Modern tool usage

CO1: 3 (Strongly Related)

CO2: 3 (Strongly Related)

CO3: 3 (Strongly Related)

CO4: 3 (Strongly Related)

CO5: 3 (Strongly Related)

CO6: 3 (Strongly Related)

CO7: 3 (Strongly Related)

CO8: 3 (Strongly Related)

Justification: The use of modern statistical tools and software is essential for achieving objectives (CO1 to CO9), thus strongly related to modern tool usage.

PO4 Environment and sustainability

CO1: 1 (Partially Related)

CO2: 1 (Partially Related)

CO3: 1 (Partially Related)

CO4: 1 (Partially Related)

CO5: 1 (Partially Related)

CO6: 1 (Partially Related)

CO7: 1 (Partially Related)

CO8: 1 (Partially Related)

CO9: 1 (Partially Related)

Justification: The objectives (CO1 to CO9) primarily focus on statistical analysis and computation and do not directly relate to environmental or sustainability concerns.

PO5 Ethics

CO1: 1 (Partially Related)

CO2: 1 (Partially Related)

CO3: 1 (Partially Related)

CO4: 1 (Partially Related)

CO5: 1 (Partially Related)

CO6: 1 (Partially Related)

CO7: 1 (Partially Related)

CO8: 1 (Partially Related)

CO9: 1 (Partially Related)

Justification: The objectives (CO1 to CO9) do not explicitly address ethical considerations in statistical analysis.

PO6 Individual and Team work

CO1: 2 (Moderately Related)

CO2: 2 (Moderately Related)

CO3: 2 (Moderately Related)

CO4: 2 (Moderately Related)

CO5: 2 (Moderately Related)

CO6: 2 (Moderately Related)

CO7: 2 (Moderately Related)

CO8: 2 (Moderately Related)

CO9: 2 (Moderately Related)

Justification: While statistical analysis can involve both individual and team efforts, the objectives mainly focus on individual skills in analyzing and interpreting statistical data, thus moderately related to individual and team work.

PO7 Innovation, employability and Entrepreneurial skills

CO1: 2 (Moderately Related)

CO2: 2 (Moderately Related)

CO3: 2 (Moderately Related)

CO4: 2 (Moderately Related)

CO5: 2 (Moderately Related)

CO6: 2 (Moderately Related)

CO7: 2 (Moderately Related)

CO8: 2 (Moderately Related)

CO9: 2 (Moderately Related)

Justification: While the objectives (CO1 to CO9) provide essential statistical skills, they do not explicitly address innovation, employability, or entrepreneurial skills. However, these skills may be indirectly developed through the application of statistical methods in various contexts.