Paper DSC 203: DATA ANALYTICS ESSENTIALS

Hours Per Week: (3T+4P) **Credits**: 5
Exam Hours: 1 ½ **Marks:** 50T+35P+15I

Objective: *To make students to learn Essentials of data Analytics.*

UNIT I: VARIABLES FOR DATA ANALYTICS:
Types of Variables: Determine the nature of variables in data analysis - Differentiate between numerical and categorical. Variables - Distinguish between nominal and ordinal variables - Differentiate between interval and ratio - Distinguish between continuous and discrete

UNIT II: ESSENTIAL STATISTICS DATA ANALYTICS:
Central Tendency of Data: Identify the components of central tendency - Calculate mean/median/mode - Identify the steps in calculating weighted/geometric/harmonic means - Measurement and Variability: Determine core aspects of measurement and variability - Calculate range - Calculate quartiles - Calculate interquartile range - Calculate variance - Calculate standard deviation - Analyze permutation with repetition - Analyze combinations without repetition

UNIT III: PROBABILITY FOR DATA ANALYTICS :
Basic Probability: Uses of probability - Differentiate between sample space, event, independent and dependent - Calculate probability - Probability and Ven Diagramming: Analyze "this" OR "that" diagram - Analyze "this" AND "that" diagram - Analyze exclusive diagram - Joint probability - Conditional probability - Calculating Probability: Calculate P using a contingency table - Calculate P from trees - Calculate Bayes' theorem - Calculate the mean in terms of probabilities - Calculate the variance and standard deviation in terms of probabilities - Calculate conditional probability

UNIT IV: DISTRIBUTIONS:
Distributions: Analyze distributions - Discrete distributions - Binomial distributions - Poisson distributions - Continuous Distributions: Identify continuous distributions - Calculate continuous distributions - Identify cumulative distributions - Identify normal distributions - Calculate normal distributions - Compare quartiles and normal distributions - Identify skew

UNIT V: CASE STUDIES USING R:
Statistics in R Case Study: Apply Vectors in R - Use Data Frames in R - Use data from an external file in R - Apply mean/median/standard deviation in R - Distributions in R Case Study: Use Normal distribution function in R - Use Poisson distribution function in R - Apply Scatter plot in R - Apply Histogram in R - Apply Box Plot in R - Fraud Detection Case Study: Apply scripts in R - Create reusable, user defined function in R - Use Bayes' Theorem in R - Choose a function flexible to allow for different input parameters

SUGGESTED READING:

1. Application of Data Analysis Essentials Certificate; AICPA
2. Fundamentals of Business Analytics, 2nd Edition; R N Prasad, SeemaAcharya; Wiley
3. Business Analysis with Microsoft Excel and Power BI, 5th edition; Conrad G. Carlberg; Pearson
4. Data Analytics with R; BhartiMotwani; Wiley.