Printed Pages – 4 Roll No. : 20D211732011

C117514(022)

B. Tech. (Fifth Semester) Examination, Nov.-Dec. 2024

(AICTE Scheme)

(CSE Branch)

STATISTICAL THINKING for DATA SCIENCE

Time Allowed : Three hours

Maximum Marks : 100

Minimum Pass Marks : 35

Note : *Attempt all questions. Part (a) of each question is compulsory and carries 4 marks. Attempt any two parts from (b), (c) and (d) which carry 8 marks each.*

Unit-I

1. (a) Define Data Science and explain its importance in modern industries. 4

PTO

(b) Explain the difference between supervised and unsupervised learning. Provide examples of algorithms used in both. 8

(c) Discuss the role of statistics and probability in Machine Learning with suitable examples. 8

(d) Explain the concept of Data Warehousing and its importance in supporting analytical processes. 8

Unit-II

2. (a) Define Data Structures and explain their importance in designing efficient algorithms. 4

(b) Discuss the role of Data Visualization in Business Intelligence. Provide examples of tools commonly used for visualization. 8

(c) Explain the significance of Optimization Techniques in Scientific Computing. Discuss how they contribute to solving real-world problems. 8

(d) Describe the steps involved in Exploratory Data Analysis (EDA) and its importance in data-driven decision-making. 8

Unit-III

3. (a) Define Matrix Computations and explain their role in data science applications. 4

(b) Discuss the use of Scholastic Models in predictive analytics and their importance in experimentation and evaluation. 8

(c) Explain the significance of clustering in Predictive Analytics and Segmentation. Provide examples of clustering algorithms and their applications. 8

(d) Describe the role of Project Deployment Tools in data science workflows. Discuss their importance in ensuring successful project implementation. 8

Unit-IV

4. (a) Define Statistical Analysis and explain its meaning and importance in decision-making. 4

(b) Discuss the different types of Statistical Analysis. Provide examples of where each type can be applied. 8

(c) Outline the process of conducting Statistical Analysis. Explain how this process ensures accuracy and reliability in data interpretation. 8

(d) Describe the benefits of Statistical Analysis in solving real-world problems. Discuss commonly used methods and tools for performing statistical analysis. 8

Unit-V

5. (a) Explain the Statistical Foundations necessary for Data Science. Highlight their role in analyzing data effectively. 4

(b) Discuss key applications of Statistical Analysis in Data Science. Provide examples of real-world scenarios where statistical techniques are utilized. 8

(c) Present a case study showcasing the use of Statistical Analysis to solve a specific problem. Explain the steps and statistical methods used in the study. 8

(d) Discuss the importance of Statistical Analysis in Data Science workflows including its limitations and areas for improvement. 8