

Reg. No. :

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Question Paper Code: 2436111

B.E. / B.Tech. DEGREE EXAMINATIONS, NOV / DEC 2024

Sixth Semester

Artificial Intelligence and Data Science

U20AI603 – FOUNDATION OF DATA SCIENCE

(Regulation 2020)

Time: Three Hours

Maximum: 100 Marks

Answer ALL questions

PART – A

(10 x 2 = 20 Marks)

1. What is Data Science?
2. Define Data discretization.
3. Summarize Data visualization.
4. Construct a stem and leaf plot from the following list of numbers [1,14,16,23,23,29,35,44,50,67,67,70,78].
5. What is Bayes theorem?
6. Compare Correlation versus Covariance.
7. Illustrate Mean Squared Error?
8. What is empirical distribution?
9. How to Calculate Percentile?
10. List out the difference between Standard Deviation and Standard Error.

PART – B

(5 x 16 = 80 Marks)

11. (a) Elaborate on Data Cleaning in Data Science. How to Clean Data, Components of Quality Data, Advantages and Benefits of Data Cleaning. (16)

(OR)

- (b) Build Data preprocessing stages. (16)

12. (a) Summarize Exploratory Data Analysis (EDA). How does EDA differ from Classical Data Analysis Techniques? (16)

(OR)

- (b) Determine Overlaid Graphs and its methods. (16)

13. (a) What is sampling? Why is data sampling important? Give the types, Advantages, Challenges and Data sampling process. (16)

(OR)

- (b) Measure the Four Common Moments in Statistics. (16)

14. (a) Examine Hypothesis Testing? Explain different Types, Methods of Hypothesis Testing, and Basic Steps in Hypothesis Testing and Error Terms in Hypothesis Testing. (16)

(OR)

- (b) Distinguish between mean error and standard error. Explain how the maximum likelihood with suitable example. (16)

15. (a) Explain Central Limit Theorem with formula. Proof, Steps to Solve Problems on Central Limit Theorem, application with an example. (16)

(OR)

- (b) Discuss Point, Interval Estimation, Properties, and types of Interval Estimation, Limitations and Assumptions of Interval Estimation. (16)