Reg. No. : ☐☐☐☐☐☐☐☐☐☐☐☐☐

Question Paper Code : 20011

B.E./B.Tech. DEGREE EXAMINATIONS, NOVEMBER/DECEMBER 2023.

Third Semester

Artificial Intelligence and Data Science

AD 3301 – DATA EXPLORATION AND VISUALIZATION

(Regulations 2021)

Time : Three hours   Maximum : 100 marks

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1.   What does EDA mean in data?

2.   State the purpose of data aggregation.

3.   What is Matplot lib used for?

4.   Differentiate plot vs subplot.

5.   List the three main types of univariate analyses.

6.   What is the purpose of smoothing a time series data?

7.   Name the two main types of statistical testing in bivariate analysis.

8.   Is bivariate qualitative or quantitative?

9.   What is multivariate analysis?

10.   What are the two common techniques used to perform dimension reduction?

PART B — (5 × 13 = 65 marks)

11.   (a)   Provide an explanation of the various EDA tools that are used for data analysis.

Or

      (b)   What is cross-tabulation and PivotTable? How to Build Pivot Table and Cross Tab Reports?

12. (a) Why matplotlib is used for data visualization? Which module of matpotlib is used for data visualization?

Or

(b) How do you Visualize a Three-Dimensional Function in python? Illustrate with a code.

13. (a) What is scaling and standardization? When and why to standardize a variable? Illustrate with suitable example.

Or

(b) Explain the Smoothing Techniques for time series data with suitable example.

14. (a) How do you analyze contingency tables? Give examples.

Or

(b) Discuss the best Practices for Designing Scatter Plots.

15. (a) What are the characteristics of multivariate analysis? How do you explain multivariate analysis?

Or

(b) What is TSA analysis? Explain ARIMA, smooth-based and moving average.

PART C — (1 × 15 = 15 marks)

16. (a) Give a case study on univariate and multivariate analysis with example.

Or

(b) Describe the various distributions module of Seaborn for visualization. Consider a sample application to illustrate.