

Reg. No. :

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

Question Paper Code: U9273

B.E./B.Tech. DEGREE EXAMINATION, APRIL/MAY 2025

Open Elective

Computer Science Engineering

21UCS973– DATA ANALYSIS USING OPEN SOURCE TOOLS

(Common to All Engineering branches)

(Regulations 2021)

Duration: Three hours

Maximum: 100 Marks

Answer ALL Questions

PART A - (10 x 2 = 20 Marks)

- | | |
|---|---------|
| 1. What is data analysis? | CO1- U |
| 2. Differentiate Python and R in terms of data analysis? | CO1- U |
| 3. What is data cleaning? | CO1- U |
| 4. Write the Difference between deleting and imputing missing data? | CO1- U |
| 5. Name two open-source tools used for data visualization. | CO1- U |
| 6. Write a Matplotlib program to create a simple pie chart. | CO2-App |
| 7. What is the formula for calculating the mean of a dataset? | CO1- U |
| 8. What is the difference between AI and machine learning? | CO1- U |
| 9. Name two applications of data analysis in the healthcare sector. | CO1- U |
| 10. What are the key stages of an end-to-end data analysis project? | CO1- U |

PART – B (5 x 16= 80 Marks)

- | | | |
|--|--------|------|
| 11. (a) Explain the key steps in the data analysis process with examples. | CO1- U | (16) |
| Or | | |
| (b) Compare structured, unstructured, and semi-structured data. | CO1- U | (16) |
| | | |
| 12. (a) Discuss the role of data analysis in decision-making with examples from industries like healthcare, finance and marketing. | CO1- U | (16) |
| Or | | |
| (b) Explain the importance of Exploratory Data Analysis (EDA) using Pandas. | CO1- U | (16) |

13. (a) Explain different visualization techniques used in data analysis. CO1- U (16)
Or
(b) Discuss geospatial data visualization using open-source tools. CO1- U (16)
Explain its applications.
14. (a) Explain the concept of Supervised Learning with examples. How is it different from Unsupervised Learning? CO1- U (16)
Or
(b) What is Cross-Validation? Explain different types of cross-validation techniques in scikit-learn. CO1- U (16)
15. (a) Discuss how time series analysis can be used to forecast sales. CO1- U (16)
Or
(b) Explain the use of data analytics in healthcare for disease prediction. CO1- U (16)