

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

# Gandhi Institute of Engineering and Technology University, Odisha, Gunupur (GIET University)



B. Tech(Eighth Semester - Regular) Examinations, April – 2025

**21BCMPE48001 –Big Data & Visualization**

(CSE-AIML)

Time: 3 hrs

Maximum: 70 Marks

**Answer ALL questions**  
(The figures in the right hand margin indicate marks)

## PART – A

(2 x 5 = 10 Marks)

Q.1. Answer **ALL** questions

	CO #	Blooms Level
a. What is the key difference between exploration and explanation in data visualization?	CO2	K1
b. Why does proximity matter in visualizing relationships between data points?	CO1	K2
c. What is the significance of naming nodes in data visualization?	CO1	K2
d. What is data wrangling, and why is it important before visualization?	CO3	K1
e. How does the Nest function help in structuring hierarchical data?	CO4	K3

## PART – B

(15 x 4=60 Marks)

Answer **all** the questions

	Marks	CO #	Blooms Level
2. a. What are the major challenges in Big Data Visualization? Discuss issues related to scalability, data quality, real-time processing, and user interpretation. Provide examples of how these challenges can be addressed?	8	CO1	K3
b. Define the key factors that determine the quality of visualization. How can clarity, accuracy, and usability impact the effectiveness of data visualization?	7	CO2	K2
(OR)			
c. Why is understanding the reader's context important in data visualization? Discuss how compatibility with real-world scenarios can enhance or hinder the effectiveness of visualization?	7	CO1	K3
d. Different organizations use specific types of graphs and layouts. How do industry-specific requirements influence visualization choices? Provide examples from different fields like healthcare, finance, and social sciences?	8	CO2	K2
3.a. Why is avoiding loops in R beneficial for Big Data analysis? Explain how vectorized operations and functions like apply (), lapply(), and sapply() improve efficiency compared to traditional looping structures?	7	CO3	K3
b. Explain the concepts of central tendency and dispersion in data analysis. How do functions in R, such as mean (), median (), sd(), and IQR(), help in understanding the distribution of large datasets?	8	CO4	K3
(OR)			
c. What is the role of a Data Manager in handling Big Data? Explain how a Data Manager ensures data integrity, security, and accessibility while working with large-scale datasets?	8	CO2	K2
d. How do you assess the status and relevance of a dataset before visualizing it? Discuss pre-processing techniques such as filtering, outlier detection, and missing value treatment in R?	7	CO3	K1

- |      |   |   |     |    |
|------|---|---|-----|----|
| 4.a. | Why is experimentation important in data visualization? Discuss the impact of different chart types, color schemes, and interactivity features in improving insights from data?   | 8 | CO3 | K2 |
| b.   | What is a Data Pane in visualization tools, and how does it facilitate data exploration? Explain how users can manipulate data fields, apply filters, and create calculated fields?   | 7 | CO3 | K3 |
| (OR) |   |   |     |    |
| c.   | What is the role of filters and measure names in data visualization? Explain how they help users refine data views and enhance interactivity in dashboards?   | 8 | CO4 | K3 |
| d.   | How can visualization help analyze the effect of promotional spending on sales? Discuss different visualization techniques used to study correlations between marketing efforts and sales performance?                                  | 7 | CO4 | K2 |
| 5.a. | How does D3.js handle large-scale data visualization? Discuss the challenges of visualizing Big Data using D3 and the techniques used to optimize performance when dealing with complex datasets?                                       | 7 | CO3 | K2 |
| b.   | How does D3.js integrate with large datasets for visualization? Explain different methods of handling and displaying Big Data efficiently, including techniques like lazy loading, hierarchical data structures, and real-time updates? | 8 | CO4 | K3 |
| (OR) |   |   |     |    |
| c.   | What is a stacked view in data visualization, and why is it useful? Provide an example of how D3.js can be used to create a stacked bar or stacked area chart?  | 7 | CO4 | K3 |
| d.   | How does D3.js support dynamic formatting and data-driven visualization updates? Discuss techniques like responsive design, data filtering, and real-time updates to modify visualization formats based on user interactions?           | 8 | CO3 | K2 |

--- End of Paper ---