

(OR)		6	2	1	1
b.i.	Illustrate the need of data categorization.				
	ii. Describe, how to choose appropriate visual encoding.	6	2	1	1
29. a.	Illustrate using data manager how to reformat any two-character length record date years to the desired four-character length years.	12	2	2	1
(OR)					
b.	Explain dispersion in R and its syntax using an example.	12	2	2	1
30. a.	Summarize various file types and extensions in tableau.	12	2	3	4
(OR)					
b.i.	Describe on the tool trifecta.	6	2	3	4
	ii. Examine how a .CSV file is loaded in tableau desktop with clear steps.	6	2	3	4
31. a.	Consider the manufacturing data set, having a simple comma-delimited text file with the total number of products produced for each machine ID, broken out by shift.	12	3	4	1
	(i) Explain the steps for creating a grouped bar chart				
	(ii) Explain about stacked and multiple view for downtime				
(OR)					
b.	Demonstrate the way to create a custom animation using D3.JS.	12	2	4	1
32. a.	Explain the alternatives to pie chart by referring the example. Assume your own dataset.	12	2	5	1
(OR)					
b.	Imagine that you manage an Information Technology (IT) team. Your team receives tickets, or technical issues, from employees. In the past year, you've had a couple of people leave and decided at the time not to replace them. You have heard a rumbling of complaints from the remaining employees about having to "Pick Up the Slack". You've just been asked about your hiring needs for the coming year and are wondering if you should hire a couple more people. Discuss the changes required to reduce the clutter.	12	3	6	1

* * * * *

Reg. No.																			
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

B.Tech. DEGREE EXAMINATION, JUNE 2023
Fifth & Sixth Semester

18CSE490T – BIG DATA VISUALIZATION
(For the candidates admitted during the academic year 2018-2019 to 2021-2022)

Note:

- (i) **Part - A** should be answered in OMR sheet within first 40 minutes and OMR sheet should be handed over to hall invigilator at the end of 40th minute.
- (ii) **Part - B & Part - C** should be answered in answer booklet.

Time: 3 hours

Max. Marks: 100

PART – A (20 × 1 = 20 Marks)

Answer **ALL** Questions

	Marks	BL	CO	PO
1. Point out the correct statement	1	1	1	1
(A) Hadoop do need specialized hardware to process the data				
(B) Hadoop 2.0 allows live stream processing of real-time data				
(C) In the Hadoop programming framework output files are divided into lines or records				
(D) Hadoop 2.0 does not allow live stream processing of real-time data				
2. _____ helps in designing effective tables and charts for data visualization	1	1	1	1
(A) Pivot table				
(B) Data-ink ratio				
(C) Scatter charts				
(D) Cross tabulation				
3. Hadoop is a framework that works with a variety of related tools common cohorts include _____.	1	1	1	1
(A) Mapreduce, Hive and hbase				
(B) Mapreduce, MySQL and Google Apps				
(C) Mapreduce, Hummer and Iguana				
(D) Mapreduce, Heron and Trumpet				
4. Data visualization is also an element of the broader _____.	1	1	1	1
(A) Data process architecture				
(B) Data presentation architecture				
(C) Deliver presentation architecture				
(D) Deliver process architecture				
5. Point out the wrong statement	1	1	2	1
(A) R is a language for data analysis and graphics				
(B) K is a language for statistical modelling and graphics				
(C) One key limitation of the S language was that it was only available in a commercial package, S-plus				
(D) C is a language for data and graphics				
6. Choose the language which are most used in data visualization.	1	1	2	1
(A) C				
(B) C++				
(C) Ruby				
(D) R programming				

7. Numbers in R are generally treated as _____ precision real numbers. 1 1 2 1
 (A) Single (B) Double
 (C) Real (D) Imaginary
8. Which of the following method make vector of repeated values? 1 1 2 1
 (A) Rep () (B) Data ()
 (C) View () (D) Read ()
9. What does D3.js mean? 1 1 3 2
 (A) It is a JavaScript framework to display D3 methods (B) It is a JavaScript library for chaining native objects to D3 objects
 (C) It is node.js to parse a servers data to objects with D3 features (D) It is a JavaScript library for creating and manipulating documents based on data
10. What are the selector methods in D3.js? 1 1 3 2
 (A) Par () (B) Show ()
 (C) Read () (D) HTML ()
11. Which of these is the most fundamental contribution of D3? 1 1 4 2
 (A) D3 allows a web page to include vector graphics (B) D3 allows a web page to be interactive
 (C) D2 allows a web page to be accessed as a database (D) D3 allows a web page to display a data visualization
12. The data method of D3 is designed to associated each data item with _____. 1 1 4 2
 (A) A unique CSS style in the web page (B) A unique SVG canvas in the web page
 (C) A unique section of javascript code in the web page (D) A unique element in the web page
13. How do you identify a continuous field in tableau? 1 1 5 4
 (A) It is identified by a blue pill in the visualization (B) It is identified by a green pill in a visualization
 (C) It is preceded by a # symbol in the data window (D) When added to the visualization, it produces distinct values
14. The icon associated with the field that has been grouped is a _____. 1 1 5 4
 (A) Paper clip (B) Set
 (C) Hash (D) Equal to
15. A good reason to use a bullet graph is _____. 1 1 5 4
 (A) Analyzing the trend for a time period (B) Comparing the actual against the target sales
 (C) Adding data to bins and calculating count measure (D) Displaying the sales growth for a particular year
16. Tableau was introduced in the year of 1 1 5 4
 (A) 2000 (B) 2002

- (C) 2003 (D) 2005
17. What type of title should you use for a proper data story? 1 1 6 1
 (A) An action title (B) A predictive title
 (C) Don't use titles (D) An arbitrary title
18. What helps you to show differences across various categories between the two data points? 1 2 6 1
 (A) Slope graph (B) Waterfall chart
 (C) Line graphs (D) Scatter plot
19. When should we use a table instead of a chart? 1 2 6 1
 (A) When the data is very complex (B) When there are no real patterns in the data
 (C) When we are intersected in looking up specific values in the data set (D) When the data is very simple
20. What is an appropriate gap width between the bars in a bar chart? 1 1 6 1
 (A) 20% (B) 30%
 (C) 50% (D) 100%

PART – B (5 × 4 = 20 Marks)

Answer ANY FIVE Questions

Marks BL CO PO

21. Explain the general categories of data quality. 4 2 1 1
22. Differentiate info graphics and data visualization. 4 2 1 1
23. List the various data types/ objects in R and describe any two. 4 1 2 2
24. Outline any three components of dash board. 4 2 3 4
25. Show the command used to create simple axis in D3.js. 4 1 4 2
26. Illustrate the way to create a responsive design using D3.js. 4 2 5 2
27. Demonstrate the strategies that you can leverage for gaining acceptance in the design of your data visualization. 4 2 6 1

PART – C (5 × 12 = 60 Marks)

Answer ALL Questions

Marks BL CO PO

28. a. Scenario: You are a data scientist at a healthcare organization, and you have been asked to analyse patient data to identity trends in patient health. The data includes patient demographics, medical history and health outcomes. 12 3 1 1

Question: How would you use data visualization to effectively communicate the trends in patient health to the healthcare providers and stakeholders?