		Marks	BL	CO	PO
	PART – B (5 × 10 = 50 Marks) Answer **ALL** Questions				
26. a.i.	Discuss the term LATCH as proposed by Richard Saul Wurman.	5	3	1	1
ii.	Write short notes on 3V's with respect to big data.	5	3	1	1
	(OR)				
b.i.	Write short notes on data scrubbing.	5	3	1	1
ii.	Discuss the relationship between the designer, reader and data.	5	3	1	1
27. a.i.	Write a R script to find the levels of factor of a given vector, V containing data 2, 5, 3, 3, 5, 2, NA, 3, 5, NA, 10.	5	4	2	2
ii.	Write short notes on missing values? In R how missing values are represented.	5	4	2	2
	(OR)				
b.	Create a histogram in R with data X = (22, 24, 57, 91, 21, 8, 95, 13, 58, 77, 70, 46, 35, 33, 50, 56, 58). Title of the chart should be "Distribution of Marks". "Marks" as X label and "Frequency" as Y label. Colour of the labels shared be in dark green.	10	4	2	2
28. a.	Discuss in detail the various types of Tableau filters.	10	3	3	3
	(OR)				
b.	Explain the various functions provided in trifacta tool.	10	3	3	3
29. a.	Discuss in detail the various features of D3.js.	10	3	4	4
	(OR)				
b.	Write a d3.js code snippet to create the following		4	4	4
(i)	Create a SVG image with width as 450 pixel and height as 450 pixel	2			
(ii)	Create a line between (100, 100) to (250,100)	2			
(iii)	Create a rectangle with width as 250 pixel and height as 150 pixel. It's one of the coordinates to be (50, 50)	2			
(iv)	Include select method to select all <p> tags	2			
(v)	Apply red colour using style ( ) for the above <p> selection	2			
30. a.	Discuss how will data visualization can be done to alternate pie charts. Provide suitable illustrations to support your answer.	10	3	5	5
	(OR)				
b.	There should be a logic in the order in which you display the information – Justify this with suitable case study.	10	3	6	6

* * * * *

Reg. No.														

B.Tech. DEGREE EXAMINATION, DECEMBER 2022
Sixth/ Seventh Semester

18CSE490T – BIG DATA VISUALIZATION
(For the candidates admitted from the academic year 2018-2019 to 2019-2020)

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** should be answered in answer booklet.

Time: 2½ Hours Max. Marks: 75

		Marks	BL	CO	PO
	PART – A (25 × 1 = 25 Marks) Answer **ALL** Questions				
1.	Which method shows hierarchical data in a nested format? (A) Tree map (B) Scatter plots (C) Population pyramids (D) Area chart	1	1	1	1
2.	_____ are column charts, where each column represents a range of values and the height of a column corresponds to how many values are in that range. (A) Bar graph (B) Histogram (C) Line chart (D) Pie chart	1	2	1	1
3.	An object that explains the symbols, colors or patterns used to differentiate the data. (A) Legend (B) Data label (C) Chart title (D) x-axis	1	1	1	1
4.	Which of the following is not used to represent data quality? (A) Accuracy (B) Relevance (C) Appropriateness (D) Visual art	1	1	1	1
5.	What are pros of data visualization? (A) It can misrepresent information (B) It can be distracting (C) It can be accessed quickly by a wider audience (D) No prose in data visualization	1	2	1	1
6.	Which of the following is TRUE for a vector is R? (A) It is a homogeneous 1-dimensional data structure (B) It is a heterogeneous 1-dimensional data structure (C) It is a homogeneous 2-dimensional data structure (D) It is a heterogeneous 2-dimensional data structure	1	2	2	2
7.	Which function cannot be used to import a CSV file in R? (A) read.table ( ) (B) read.csv ( ) (C) read_excel ( ) (D) read_csv ( )	1	2	2	2

8. _____ are generic data objects of R which are used to store the tabular data, which are made up of three principal components the data, rows and columns. 1 1 2 2
(A) Array (B) Factors
(C) Data frames (D) Lists

9. What will be the output of the following R code? 1 2 2 2
> x ← data.frame (foo = 1;4, bar = c(T,T,F,F)
> ncol (x)
(A) 2 (B) 4
(C) 8 (D) True

10. What will be the output of the following R commands listed below? 1 2 2 2
> data = list ("x", "y", "z", 1, 2, 3)
> dim (data) = c (2, 3)
> data
(A) "x" 1 (B) "x" "z" 2
"y" 2 "y" "1" 3
"z" 3
(C) "x" "y" "z" (D) Error
1 2 3

11. How do you identify a continuous field in tableau? 1 1 3 3
(A) It is identified by a blue pill in visualization (B) It is identified by a green pill in visualization
(C) It is preceded by a # symbol in the data window (D) When added to the visualization, it produced district values

12. The icon associated with the field that has been grouped in a _____. 1 1 3 3
(A) Paper clip (B) Set
(C) Hash (D) Equal to

13. To create a variable size bin, we use _____ in tableau. 1 2 3 3
(A) Sets (B) Groups
(C) Calculated fields (D) Table calculation

14. Which of these is not a tableau file format? 1 1 3 3
(A) tde (B) tdb
(C) twb (D) twbx

15. What is sets in tableau? 1 1 3 3
(A) Category purpose (B) Purpose based on the some condition
(C) Calculated field possible in sets (D) A type of dual axis

16. What does D3.js mean? 1 2 4 4
(A) It is a JavaScript framework to display D3 models (B) It is a JavaScript library for changing native objects to D3 objects
(C) It is a 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

17. D3 is an open source JavaScript library based upon its predecessor the _____ framework. 1 1 4 4
(A) SVG (B) Protovis
(C) Vi2 SQl (D) Drotovis

18. The _____ method is used to add or update the attribute of the selected elements. 1 1 4 4
(A) attrib ( ) (B) attribute ( )
(C) add_attrib ( ) (D) attr ( )

19. Which of the following task is not performed by D3? 1 2 4 4
(A) Manipulation of DOM (B) Working with charts
(C) Working with data (D) Compression of data

20. Which of the following are needed for D3.js development? 1 2 4 4
(A) D3.js library, SVG, web browser, web server (B) D3.js library, editor, web browser, web server
(C) D3.js library, editor, web server (D) D3.js library, MySQL, web browser, web server

21. A _____ is a line that provides an approximation of the relationship between the variables. 1 1 5 5
(A) Spark line (B) Grid line
(C) Trend line (D) Spark

22. Which of the following statement is false? 1 2 5 5
(A) Data visualization include the ability to absorb information quickly (B) Data visualization is another form of visual art
(C) Data visualization decreases the insights and take slower decisions (D) Data visualization makes complex data more accessible, understandable and usable

23. What are the benefits of data visualization? 1 2 6 6
(i) Better analysis
(ii) Identifying patterns
(iii) Exploring business insights
(A) (i) and (ii) (B) (ii) and (iii)
(C) (i) only (D) (i), (ii) and (iii)

24. Which method shows hierarchical data in a nested format? 1 1 6 6
(A) Tree maps (B) Scatter plots
(C) Area charts (D) Population pyramids

25. _____ is the presentation of quantitative information in a graphical form. 1 1 6 6
(A) Data exploration (B) Data visualization
(C) Data improvision (D) Data magnification