Reg. No															

B.Tech DEGREE EXAMINATION, NOVEMBER 2023

Fifth & Seventh Semester

18CSE490T - BIG DATA VISUALIZATION

(For the candidates admitted during the academic year 2020 - 2021 & 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** and **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
1.	Which of the following requires a routine and regular review and evaluation of the data? (A) Data quality assurance (B) Big data (C) Software quality assurance (D) Software project management	1	1	1
2.	To represent Annual seasons, 24-hour days, and biological processes such as the Krebs Cycle, which of the following charts can be used? (A) Circular layout (B) Rectangle layout (C) Square layout (D) Hierarchy layout	1	1	1
3.	Not Manually drawn, aesthetically poor and relatively data-rich are the characteristics of ---------- (A) Machine Learning (B) Infographics (C) Data Visualization (D) Big Data	1	1	1
4.	Which of the following is NOT an example of structured data? (A) Captured data (B) Created data (C) Compiled data (D) Transactional data	1	1	1
5.	Data Manager supports reading from various----------data sources for greater flexibility. (A) Java Database Connectivity (ODBC) (B) Open Database Connectivity (ODBC) (C) Components-based (D) ActiveX control	1	1	2
6.	Which script is used to execute external commands in the data manager? (A) Java Script (B) Python (C) R Script (D) Visual Basic Script	1	1	2
7.	---------- is used to describe the range, variance, and standard deviation of the dataset. (A) Dispersion (B) Contrasts (C) Compare (D) Tendencies	1	1	2
8.	Which SQL function is similar to the table function in R? (A) sort (B) count (C) group by (D) where	1	1	2
9.	Which Tableau product offers more options for data profiling, preprocessing, and manipulation? (A) Desktop (B) Trifacta (C) VizQL (D) SQL	1	1	3

10. Which is the shortcut for swap or flip the sheets visualization from a vertical to a horizontal display in Tableau? 1 1 3

(A) Ctrl - W (B) Ctrl - S
(C) Ctrl - H (D) Ctrl – V

11. Which option is used to add calculations directly to a visualization? 1 1 3

(A) Dimensions (B) Measures
(C) Create Calculated Field (D) Data pane

12. Which Tableau extension is suitable to save all the sheets and their connection information in a workbook file? 1 1 3

(A) .twbx (B) .twb
(C) .twbs (D) .twbv

13. Which is actually an open-source JavaScript library, designed with the intention of visualizing data using today's web standards? 1 1 4

(A) D3 (B) D3.ts
(C) D3.java (D) Both a and b

14. In Visual transitions, which button is used to change the format of the visualization? 1 1 4

(A) Reset (B) Radio
(C) Submit (D) Select

15. When dealing with manufacturing data, which is collects production data directly from the machine, stores it in memory, and periodically sends it off to the data repository? 1 1 5

(A) Data Manager (B) Data Logger
(C) Data Aggregator (D) Data Handler

16. Visual transitions handle missing data by filling 1 1 5

(A) Null values (B) Zero
(C) Either Null or zero (D) Variables

17. _____ background support is used to focus more on background? 1 1 6

(A) White background (B) Dark background
(C) Light Element on Dark Background (D) Grey background

18. Which Programming language does not support greater flexibility and control of the specific elements of the graphs? 1 1 6

(A) R (B) D3
(C) Python (D) Java

19. Some people use---------- either alone or together with excel or via a programming language, for easier manipulation of graph elements and a professional look and feel. 1 1 6

(A) Adobe Illustrator (B) Adobe photoshop
(C) Adobe editor (D) Adobe pdf

20. _____site allows you to submit the visualization designs and to receive feedback on data visualization work in progress from readers. 1 1 6

(A) VizWiz (B) Junk Charts
(C) HelpMeViz (D) Eager Eyes

PART - B (5 × 4 = 20 Marks) **Marks BL CO**

Answer **any 5** Questions

21. Summarize the data scrubbing process in visualization. 4 2 1

22. Explain the ways to import or read a CSV file into a data frame in R with a sample code. 4 2 2

23. Differentiate join and union in Tableau. 4 2 3

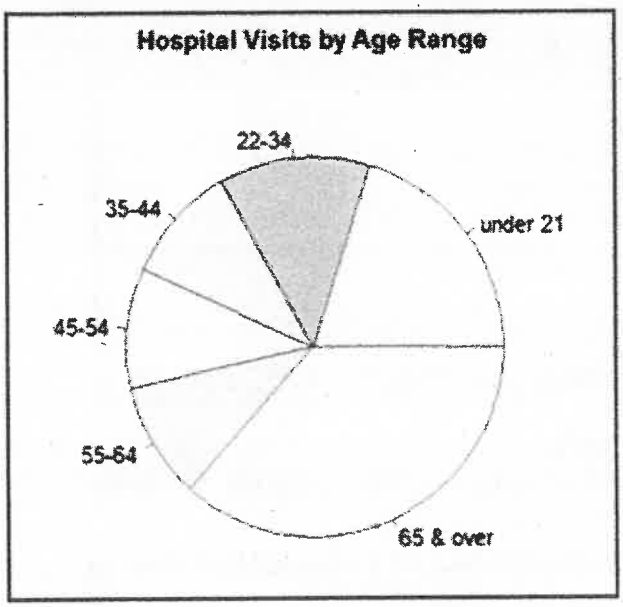
24. Compare dimensions and measures in Tableau with an example. 4 2 3

		Marks	BL	CO
25.	Create the web page to display the 'Hello world' using d3.js.	4	3	4
26.	Differentiate domain, range and scale in D3.	4	2	5
27.	Describe any five data blogs used in data visualization.	4	2	6

PART - C (5 × 12 = 60 Marks)

Answer **all** Questions

		Marks	BL	CO
28.	(a) Explain Infographics and data visualization with example. Differentiate them	12	2	1
	(OR)			
	(b) Discuss the approaches to data visualization.			
29.	(a) An Online Survey collected the medical data set. The file includes many fields such as Patient_ID, Sex, Date_of_Birth, age, Height, weight, Bloodtype, Blood_pressure, Heart_rate, Smoker, Non_smoker, No_of_hospital_vist, State, Country_code, Phone_no. Write an R code to implement the age ranges and create the following pie chart visualization.	12	3	2



(OR)

(b) Consider the medical data set to calculate BMI from the existing patient information such as the patient's weight and height.Write the procedure to calculate BMI and visualise the same

30. (a) Consider the following productSalesPromotionBurn table and write the procedure to draw the data pane, visualize the Promotion type as columns, Measure Values (the monthly promotion burn total) as rows, and Save the workbook. 12 3 3

	Promotion type	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2	Social media	1332	295	70	402	7921	26588	32437	36330	29064	2906	2180	2812
3	Television	670	557	1179	363	2670	25837	28937	32410	3241	324	243	314
4	Radio	1159	942	330	535	6290	10376	11621	13016	1302	130	98	126
5	Print	702	1675	241	508	14820	15919	17829	19969	1997	200	150	193
6	Internet	2279	415	321	396	12532	2229	2496	2796	280	28	21	27

(OR)

(b) Consider the following totals table, create a new column as an indicator using Create Calculated Field and display the output. The formula indicates that if the value of the current cell in the Change column is positive (greater than zero) then it's Up from the prior year (otherwise it is Dn from the prior year).

	A	B	C	D	E	F	G
1	Measure Names	Totals	Change				
2	CY Sales	$1,365,869	27%				
3	PY Sales	$1,079,037	21%				
4	CY Spend	$887,683	27%				
5	PY Spend	$701,270	14%				
6							

31. (a) Explain the procedure to convert Stacked-to-Multiples using D3 with an example. 12 2 5

(OR)

(b) Discuss the features used in Stacked Area via Nest example for better visualization.

32. (a) Explain how color can impact the overall tone of a visualization with an example. 12 2 6

(OR)

(b) Describe about spaghetti graph. Discuss the strategies for avoiding the spaghetti graph.

* * * * *