VIT
Vellore Institute of Technology

Continuous Assessment Test (CAT - I) – January 2023

Programme	:	B.Tech (CSE)	Semester	:	WINTER 2022 - 23
Course Title	:	Data Visualization	Code	:	CSE3020
			Slot	:	D2
Faculty	:	Dr. Parvathi. R, Dr. Joshan Athanesious J	Class Nbr	:	CH2022235000658/ CH2022235000660
Time	:	1 Hour 30 Minutes	Max. Marks	:	50

Answer ALL questions

Q. No.	Sub. Sec.	Question Description	Marks
1.		The Students detail contains attributes such as Name, Gender, Email ID, Language they Speak, Nationality, and Attendance Percentage, Internals Marks. For the above details, Assume the appropriate values of each attributes and analyse its types with Justification. Illustrate the appropriate Visualization Techniques to represent the Attendance Percentage and Internals Marks for calculating the eligible criteria for Final Assessment Test.	10
2.		Jack was listening to latest music from his smartphone using his wireless headphone. Explain and Analyse the importance of Data and Task abstraction to implement the operations such as signal transmits from the Phone Bluetooth to the headphones and Changing the Songs using Headphones.	10
3.		Construct the Color Look-Up table with the following constraints. i) It should contains Minimum and Maximum index value. ii) Color table is represented in 8 bit coding with color scheme RGB. Calculate the color index for any scalar value and find the appropriate color mapping from the color look-up table. Assume your own data.	10
4.		For the given attribute of MTCARS dataset such as Miles/(US) gallon (mpg) and Transmission (0 = automatic, 1 = manual), Number of cylinders (cyl), Gross horsepower(hp) and Weight (1000 lbs). Identify the suitable Visual representation for Composition and Distribution to analyse which type of transmission is better for obtaining higher mpg using Visualization. Assume your own data.	10
5.		The attributes of Iron rod with length (cm) and Weight(mm) was given in following table:	10

Length (cm)	Width (mm)
40	78
21	70
25	60
31	55
48	80
47	66

Write the R snippet to implement the statistical analysis and Illustrate with suitable Visualization.

		Total	[50]

VIT
Vellore Institute of Technology

Continuous Assessment Test (CAT - II) – March 2023

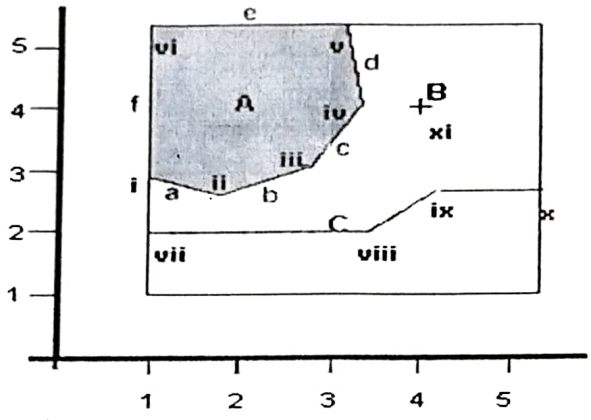
Programme	:	B.Tech (CSE)	Semester	:	WINTER 2022 - 23
Course Title	:	Data Visualization	Code	:	CSE3020
			Slot	:	D2
Faculty	:	Dr. Parvathi. R, Dr. Joshan Athanesious J	Class Nbr	:	CH2022235000658/ CH2022235000660
Time	:	1 Hour 30 Minutes	Max. Marks	:	50

Answer ALL questions

Q. No.	Sub. Sec.	Question Description	Marks
1.		In School of Computer Science and Engineering, various programmes are offered namely B.Tech CSE, B.Tech (AI&ML), B.Tech (CPS), B.Tech (AIR). In each Programmes different courses were offered. Construct an appropriate Network for the above scenario and apply suitable visualization based on the link, influential and adjacency characteristics. Justify your answer.	10
2.		The 10 month sales information of a particular bookstall is given in Table 1. Analyse the sales of November month with $\alpha = 0.3$ based on Trends, Seasonality, Cyclical components and apply any two suitable plots to visualize the results and justify your answer.	10

Table 1

Month	No. of Books Sold
January	30
February	25
March	35
April	25
May	20
June	30
July	35
August	40
September	30
October	35
November	?

3.		Figure – 1 is a digital form representation of a real-time instance of a Highway scenario. For the given Digital World Model, Analyse the following events: • Find the Coordinates of nodes and vertexes. • Find the Topology of each segments. • Find the Polygons topology.  Figure – 1	10
4	a.	Apply the different Visual Variables on **Full moon** to **New moon day cyclic process** with suitable visualizations.	10
	b.	Illustrate the characteristics of each visual variable with respect to above scenario.	
5.	a.	Compute $\nabla * \boldsymbol{A}$ at a point (1, -2, -1), if A = $xz^3\, i - 2x^2yz\, j + 2yz^4\, k$.	10
	b.	Describe the technique used to compute the above equation with two suitable visualizations.	
		Total	[50]

VIT Vellore Institute of Technology

Reg. No. : 20BCE1847

Final Assessment Test (FAT) - APRIL/MAY 2023

Programme	**B.Tech**	Semester	**Winter Semester 2022-23**
Course Title	**DATA VISUALIZATION**	Course Code	**CSE3020**
Faculty Name	**Prof. Parvathi R**	Slot	**D2**
		Class Nbr	**CH2022235000660**
Time	**3 Hours**	Max. Marks	**100**

Assume your own attributes and own data wherever needed
Apply suitable visual representation wherever needed

PART A (10 X 10 Marks)
Answer All questions

01. The Indian census database for poverty analysis contains data fields such as, Name, Age, Language known, Skin color, Marital status, Gender, Education qualifications, Experience and Income. **[10]**

1. Assume the appropriate values for each attribute and analyze its types with justification (2 Marks)
2. Illustrate any five visualization techniques to calculate the Economic status (Lower, Middle, Upper) of the people based on Education qualifications versus Income. (4 Marks)
3. Apply the data abstraction and task abstraction for the poverty analysis.(4 Marks)

02. Consider the HealthCare Dataset with the details like Department, Disease, Doctor Specialization, Inpatient and Outpatient . Apply Composition and Distribution techniques to analyse the dataset and provide your insights with suitable visualizations **[10]**

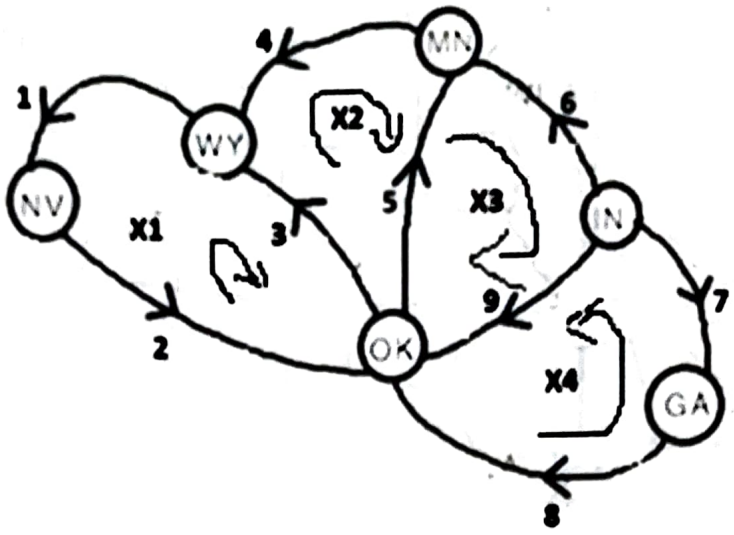
03. Consider the Velocity of Wind in the Ocean is measured using the vector function $F(x, y, z) = xyz^2i + 5xy^2zj - 8x^2yzk$. **[10]**

Identify and calculate the different metrics used to find the nature of wind. Then calculate the divergence at the point (4,5,4). Use suitable visualization to represent the result.

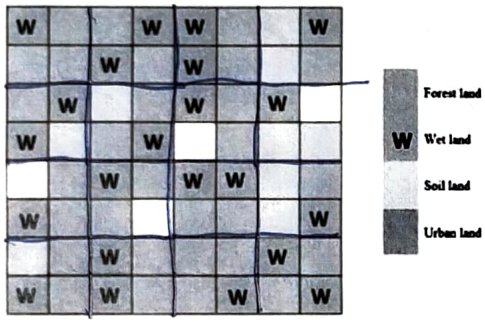
04. An Agritech company released various products such as Plough, Harrow, Tiller and Planter in different parts of India. Now they want to analyze the relation between the marketing and profit of their products in all locations. **[10]**

a. Illustrate how color theory and color mapping helps to the analysis (5 Marks)
b. Suggest the different types of colormap design for the better analysis (5 Marks)

05. For the given geospatial map in the following figure, find the topology of polygons, nodes and lines based on the locations (NV, WY, OK, MN, IN, GA) marked in the geomap and explain your observations on multilevel topologies. **[10]**



06. Consider the matrix representation of geospatial data in the following figure. Apply different types of encoding techniques only on wet land (marked as "W"). **[10]**



07. Consider the following dataset with attributes description: **[10]**

Attributes	**Description**
Age	Age limit: 20 ~ 64
Gender	F (Female) & M (Male)
Body fat_%	Body fat in percentage
Height_cm	If you want to convert to feet, divide by 30.48
Weight_kg	Weight in Kilogram
Class	A,B,C,D
BMI_Status	Below 18.5 - Underweight(A). 18.5 to 24.9 - Healthy(B). 25.0 to 29.9 - Overweight(C). 30.0 Above - Obese (D).

Write the R snippet to classify the BMI_Status based on your own data and illustrate with suitable visualization techniques.

08. The following table relates to the profits of one of the commercial industry for past 8 years **[10]**

Year	1986	1987	1988	1989	1990	1991	1992	1993
Profit (K)	15	16	18	21	26	31	35	34

A. Describe the different Time Series Components based on Profit and Illustrate with suitable three Time Series Visualization (3 Marks)

B. Visualize the profit's trend using three years moving average and Exponential Smoothing with alpha value as 0.20 (7 Marks)

09. Construct the database with 10 records based on following constrains: **[10]**

i. L-CORE (patient's internal temperature in C): high (> 37), mid (>= 36 and <= 37), low (< 36)

ii. L-SURF (patient's surface temperature in C): high (> 36.5), mid (>= 36.5 and <= 35), low (< 35)

iii. L-O2 (oxygen saturation in %):excellent (>= 98), good (>= 90 and < 98), fair (>= 80 and < 90), poor (< 80)

iv. L-BP (last measurement of blood pressure): high (> 130/90), mid (<= 130/90 and >= 90/70), low (< 90/70)

v. SURF-STBL (stability of patient's surface temperature): stable, mod-stable, unstable

vi. CORE-STBL (stability of patient's core temperature) :stable, mod-stable, unstable

vii. BP-STBL (stability of patient's blood pressure) :stable, mod-stable, unstable

viii. COMFORT (patient's perceived comfort at discharge, measured as an integer between 0 and 20)

ix. Decision ADM-DECS (discharge decision): I (patient sent to Intensive Care Unit), S (patient prepared to go home), A (patient sent to general hospital floor).

a. Write a procedure to design the various tableau worksheets (minimum 5 worksheets) with different visualization.(5 Marks)

b. Write the procedure to create a dashboard with R integration to classify the test data "L-CORE= 34, L-SURF =37, L-O2 = 100, L-BP = 120/80, SURF_STBL = unstable, CORE_STBL = stable, BP_STBL = mod-stable, COMFORT = 15". (5 Marks)

10. A carbig dataset contains various measured variables about 400 automobiles from 1970's to 1980's. Illustrate the multivariate visualization for the following attributes in given dataset. fuel efficiency (in miles per gallon, MPG), acceleration (0-60MPH in sec), engine displacement (in cubic inches), weight, number of cylinders (4, 6, 8) and horsepower. Write a procedure to create a dashboard with three ranking charts , three statistical graph and group the observations based on number of cylinders. **[10]**