



VIT

Vellore Institute of Technology
Affiliated to University under section 3 of UGC Act 1956
CHENNAI

Reg. Number:

24MS1017

Mahesh Jagtap

Continuous Assessment Test (CAT) – II - OCTOBER 2024

Programme	:	M.Tech CSE & its specializations	Semester	:	Fall 2024-2025
Course Code & Course Title	:	MCSE501L & Data Structures and Algorithms	Class Number	:	CH202425010 3140 CH202425010 3180
Faculty	:	Dr.R.Kanniga Devi Dr.R.Rajakumar	Slot	:	A2+TA2
Duration	:	90 minutes	Max. Mark	:	50

General Instructions:

- Write only your registration number on the question paper in the box provided and do not write other information.
- Only non-programmable calculator without storage is permitted

Answer all questions

Q. No	Description	Marks
1	<p>A company's product management system hierarchy is represented using a complete binary tree where each node represents a product feature, arranged in lexicographical order at each level and the tree is traversed to make key decisions about product enhancements. The only information given is the post-order traversal of the binary tree:</p> <ul style="list-style-type: none"> Post-order Traversal: H, I, D, J, K, E, B, L, F, G, C, A <p>Given this post-order traversal, using logical reasoning reconstruct the following binary tree:</p> <ol style="list-style-type: none"> Pre-order traversal. (5 Marks) In-order traversal. (5 Marks) 	10
2	<p>You are managing an online bookstore's inventory system using a B-tree of order 3. The current B-tree contains the following character keys in the order they were inserted: A, C, E, G, I, K, M, O, Q, S.</p> <ol style="list-style-type: none"> Show the B-tree structure after all the insertions. (5 Marks) Delete the keys C, I, and M from the B-tree one at a time from the resulting tree of every deletion. Show the intermediate B-tree structures after each deletion. (5 Marks) 	10
3	<p>A database management system for a university's course registration is designed using a B+ tree of order 3. The keys represent unique course codes, and the leaf nodes contain pointers to course details, including the course title, instructor, and available seats.</p> <p>You are inserting the following course codes into the B+ tree in the given order:</p> <p>CS101, MATH201, PHYS102, CHEM101, BIO202, CS202, MATH30,</p>	10

	<p>STAT101, ENG201, HIST101, MATH101</p> <p>a. Construct the B+ tree step by step, showing the intermediate structure after each insertion. Indicate when splits occur and how the keys are redistributed at every step. (8 Marks)</p> <p>b. After inserting all the course codes, demonstrate a search operation for the course code PHYS102 and describe the steps taken to find it within the B+ tree. (2 Marks)</p>	
4	<p>A retail company has developed a product catalogue management system using a threaded binary tree to facilitate quick access to product details. Each node in the threaded binary tree represents a product, containing the Product ID (an integer) attribute:</p> <p>The company wants to insert the following products into the threaded binary tree in the given perfect order:</p> <ul style="list-style-type: none"> • Product ID: 1001 • Product ID: 1002 • Product ID: 1003 • Product ID: 1004 • Product ID: 1005 • Product ID: 1006 • Product ID: 1007 <p>Perform the following operations:</p> <ol style="list-style-type: none"> a. Show the binary tree structure after completing insertions. (2 Marks) b. Mention the rules for how the threading is established and indicate the final threaded binary tree structure. (4 Marks) c. List the product IDs in the order they are accessed during the in-order traversal and explain how threading improves traversal efficiency. (4 Marks) 	10
5	<p>A library management system uses a splay tree to manage its collection of books efficiently. Each node in the splay tree represents a book ID. Insert the following book IDs in this structure so as to get the most recently inserted book in the constant time.</p> <p>The book IDs are: 15, 23, 8, 42, 30, 50, 18, 7</p> <p>Perform a splay operation after each insertion and indicate the type of splay operation done in the intermediate steps.</p>	10

*****All the best*****



Continuous Assessment Test (CAT) – 2 - October 2024

Programme	:	M.Tech. CSE & its specialization	Semester	:	Fall Sem 2024-25
Course Code & Course Title	:	MCSE503L & Computer Architecture and Organisation	Class Number	:	CH2024250103182 CH2024250103188
Faculty	:	Dr.Thanikachalam V Dr. Gayathri Devi S	Slot	:	C2 + TC2
Duration	:	90 Minutes	Max. Mark	:	50 Marks

General Instructions:

- Write only your registration number on the question paper in the box provided and do not write other information.
- Only non-programmable calculator without storage is permitted

Answer all questions

Q. No	Sub Sec.	Description	Marks
1.		<p>Consider the following MIPS code fragments, each containing two instructions. For each code fragment identify the type of hazard that exists between the two instructions and the registers involved. [5 x 2 = 10 Marks]</p> <p>a. LD R1, 0(R2) DADD R3, R1, R2</p> <p>b. MULT R1, R2, R3 DADD R1, R2, R3</p> <p>c. MULT R1, R2, R3 MULT R4, R5, R6</p> <p>d. DADD R1, R2, R3 SD 2000(R0), R1</p> <p>e. DADD R1, R2, R3 SD 2000(R1), R4</p>	10
2		<p>Assume that there are 5 number of stages in the pipeline. Draw the diagram of the optimal pipeline schedule using forwarding and stalls for the given code. [2 X 5 = 10 Marks]</p> <p>a. LW R1, 45(R2) ADD R5, R1, R7 SUB R8, R6, R7 OR R9, R6, R7</p> <p>b. LW R1, 45(R2) ADD R5, R6, R7 SUB R8, R1, R7 OR R9, R6, R7</p>	10

3.	<p>Consider the given program sequence</p> <pre>//Code int a=0; while(a<5) { //branch instruction, condition either true or false if(a%2==0) { a++; } }</pre> <p>Discuss how the compiler predicts the branch using</p> <ol style="list-style-type: none"> 2-bit Dynamic branch prediction technique. [5 Marks] Correlating branch prediction technique. [5 Marks] 	10
4.	<p>OpenMP runs a user program in parallel. Create a program that finds the largest number in an array A.</p> <ol style="list-style-type: none"> Write a C program and parallelize it using OpenMP directives. Use Sections in OpenMP. [7 Marks] Find the Thread number (Thread ID) and master thread in the above program using OMP function. [3 Marks] 	10
5.	<p>You are given two array a & b with integer/real values and the expression given below</p> $c[i] = a[i] + b[i]$ <ol style="list-style-type: none"> Write a C Program to add the elements of arrays in the expression given above. [3 Marks] Write the OpenMP version of C code for adding elements of two array and store into another array. Use Scheduling options in OpenMP. [7 Marks] 	10

***** All the best *****



Continuous Assessment Test (CAT) - 2 - October 2024

Programme	:	M.Tech. CSE & its specialization	Semester	:	Fall 2024-25
Course Code & Course Title	:	MCSE504L & Operating Systems	Class Number	:	CH2024250103143 CH2024250103184
Faculty	:	Dr. Manas Ranjan Prusty Dr. Manimegalai T	Slot	:	D2+TD2
Duration	:	90 Minutes	Max. Mark	:	50

General Instructions:

- Write only your registration number on the question paper in the box provided and do not write other information.
- Only non-programmable calculator without storage is permitted

Answer all questions

Q. No	Sub Sec.	Description	Marks
1		<p>In a two-way traffic system, traffic lights manage the flow of vehicles from Direction A and Direction B. Only one direction can have a green light at any time. The traffic lights switch every 10 seconds to ensure smooth traffic flow. Two semaphores, S1 and S2, are initialized as S1 = 1 (allows Direction A to proceed) and S2 = 0 (blocks Direction B).</p> <p>i. Write the pseudocode for Process P1 that manages the green light for Direction A and the pseudocode for Process P2 that manages the green light for Direction B. Ensure that your pseudo must allow the next direction to proceed after appropriately simulating a delay for the green light of its own. (6 Marks)</p> <p>ii. Explain how your solution is eligible to overcome the critical section problem in this case. (4 Marks)</p>	10
2		<p>In a library management system, there are ten computers, six printers, and eight books available for use. Five groups of students are participating in a project, each with specific resource allocations and needs.</p> <ul style="list-style-type: none"> Group A has allocated 3 computers, 1 printer, and no books, while waiting for 4 computers and 1 book to complete their project. Group B has 2 computers, 0 printers, and 2 books but is waiting for 3 printers and 1 computer to finalize their work. Group C has 0 computers, 1 printer, and 3 books, waiting for 2 computers and 1 printer to finish its tasks. Group D is allocated 1 computer, 1 printer, and 1 book but is waiting for 2 computers and 2 printers. Finally, Group E has 2 computers, 0 printers, and 2 books, waiting for 1 computer, 1 printer, and no books. 	10

		The task is to perform a deadlock detection algorithm step by step to determine if a deadlock exists. If a deadlock exists, suggest a technique to resolve it immediately; if no deadlock exists, find the safe sequence of execution.	
3		<p>A computer system with 200K of available main memory employs variable-sized partitioning for memory allocation. The sequence of job arrivals and completions is as follows: Job A requiring 50K arrives first, followed by Job B needing 70K and Job C requesting 40K. After Job A finishes, it frees 50K of memory, at which point Job D arrives, requiring 25K. Subsequently, Job B finishes and frees 70K of memory, leading to the arrival of Job E, which needs 60K. Finally, Job F arrives requiring 30K, and after Job C finishes, it frees 40K of memory. The sequence concludes with Job E finishing and freeing 60K.</p> <p>Using the optimal space allocation fit and most wasteful allocation fit algorithms, draw the memory allocation table showing how memory is assigned and freed for each job. Compare the performance of both algorithms in terms of memory utilization and fragmentation.</p>	10
4		<p>A laptop which is byte addressable uses 33 bits to represent the physical address and contains a secondary memory of 512GB. Each page contains 16K words. Find the below:</p> <ol style="list-style-type: none"> Size of the RAM in GB (2 Marks) $2^{31}B$ Total number of frames (3 Marks) 2^{19} Total number of pages (3 Marks) 2^{25} Total size of a page table (2 Marks) $608MB$ 	10
5		<p>A disk with 500 cylinders (0 to 499) has its head at cylinder 340, moving towards lower values. The queue of pending requests (P) with is:</p> <p>120(P1), 450(P2), 200(P3), 460(P2), 50(P1), 370(P3), 80(P2), 300(P1), 15(P1), 490(P3), 220(P1)</p> <p>Using the Elevator algorithm with priority scheduling, calculate the total seek distance and access time (seek time: 2ns per cylinder), processing higher-priority requests first towards the larger value of the cylinder. (Priority order = P1>P2>P3) $1560 \quad 310$</p>	10

***** All the best *****

महाराष्ट्राप्राप्त



VIT

Vellore Institute of Technology
Affiliated to the University under section 2(f) of UGC Act, 1956
CHENNAI

Reg. Number: 24mcs/017

Continuous Assessment Test (CAT) – II - OCT 2024

Programme	:	M.Tech. CSE and its specialization	Semester	:	Fall 2024-25
Course Code & Course Title	:	MCSE506L / Database Systems	Slot	:	E2+TE2
Faculty	:	Dr. A. Balasundaram Dr. R.Jothi	Class Number	:	CH2024250103186 CH2024250103146
Duration	:	1 Hr. 30 Mins.	Max. Mark	:	50

General Instructions:

- Write only your registration number on the question paper in the box provided and do not write other information.
- Use statistical tables supplied from the exam cell as necessary
- Use graph sheets supplied from the exam cell as necessary
- Only non-programmable calculator without storage is permitted

Answer all questions

Q. No	Sub Sec.	Description			Marks																																	
1	a	<p>Consider the below schedule S1 having three transactions T1, T2 and T3: $S1 : r_1(X) r_1(Z) w_1(X) r_2(Y) w_1(Z) r_2(X) r_3(Z) w_2(Y) r_3(Y) w_3(Z) w_2(X) w_3(Y)$</p> <p>Check whether the schedule S1 is serializable or not (4 marks)</p>			10																																	
	b	<p>Consider the following three transactions as given below:</p> <table border="1"> <thead> <tr> <th>T1</th> <th>T2</th> <th>T3</th> </tr> </thead> <tbody> <tr> <td>read(X);</td> <td></td> <td></td> </tr> <tr> <td>read(Y);</td> <td></td> <td></td> </tr> <tr> <td>if X=15 then Y := Y-10;</td> <td></td> <td></td> </tr> <tr> <td></td> <td>read(Y)</td> <td></td> </tr> <tr> <td>write(Y)</td> <td></td> <td></td> </tr> <tr> <td></td> <td>if Y=5 then Z := Z-10;</td> <td></td> </tr> <tr> <td></td> <td></td> <td>read(Z);</td> </tr> <tr> <td></td> <td>write(Z)</td> <td></td> </tr> <tr> <td></td> <td></td> <td>read(X);</td> </tr> <tr> <td></td> <td></td> <td>if Z=5 then X := X + 1; write(X)</td> </tr> </tbody> </table> <p>i) Check whether the above transactions will adhere to Two Phase Locking protocol. (4 marks) ii) Will execution of these transactions result in a deadlock scenario even with enforcement of Two Phase Locking protocol? Justify. (2 marks)</p>	T1	T2	T3	read(X);			read(Y);			if X=15 then Y := Y-10;				read(Y)		write(Y)				if Y=5 then Z := Z-10;				read(Z);		write(Z)				read(X);			if Z=5 then X := X + 1; write(X)			
T1	T2	T3																																				
read(X);																																						
read(Y);																																						
if X=15 then Y := Y-10;																																						
	read(Y)																																					
write(Y)																																						
	if Y=5 then Z := Z-10;																																					
		read(Z);																																				
	write(Z)																																					
		read(X);																																				
		if Z=5 then X := X + 1; write(X)																																				

2	a	<p>Assume that a histogram is formed where the values are between 1 and 20000 and are partitioned into 10 ranges namely 1-2000, 2001-4000, 4001-6000,.....,18001-20000. The respective frequency values are 2000, 3000, 1000, 1500, 2500, 3000, 1000, 3000, 1000, 2000. Provide the load balanced range partitioning function to divide the values into 5 partitions. Also provide the range partitioning vector values. (5 marks)</p>	10
	b	<p>A system with 3 processors was able to complete executing a query over 1000 records in 3 mins achieving a speedup of 3 times. What would have been the original execution time of the query before speedup was obtained. (2 marks)</p>	
	c	<p>Differentiate between speedup and scaleup. Supplement your answer by suggesting when the system has to speedup and when the system should be scaled up. (3 marks)</p>	
3		<p>You are working as a database administrator for a large e-commerce platform. The platform processes millions of transactions daily, and the sales data is stored in a database system across three nodes (say P_1, P_2 and P_3). Initially the data is partitioned into D_1, D_2 and D_3 using round robin technique, where D_i belongs to P_i.</p> <p>Given below is a set of tuples in the sales relation (SalesId, SalesDate, SalespersonId, SalesAmount)</p> <p>(101, 23/07/2023, "A", 3000) (102, 23/07/2023, "B", 1500) (103, 24/07/2023, "A", 2990) (104, 25/07/2023, "A", 3300) (105, 25/07/2023, "B", 2000) (106, 25/07/2023, "C", 4800) (107, 27/07/2023, "A", 1000) (108, 27/07/2023, "B", 5000) (109, 28/07/2023, "B", 4000) (110, 02/08/2023, "A", 5300)</p> <p>i. Your team needs to generate daily reports based on this data, and a key requirement is to sort the sales transactions by the transaction amount. To optimize performance, you decide to implement parallel sorting across the distributed database system. Explain how you sort the above data using parallel sorting algorithm. Show all the steps. (5 marks)</p> <p>ii. Suppose you need to execute <i>Sales JOIN Salesperson</i> to generate a report of transaction amount group by salesperson names. The relation SalesPerson (SalespersonId, Name, Region) has the following instances. ("A", "Amir", 7) ("B", "John", 8) ("C", "Rishab", 6)</p> <p>Explain how you will optimize this join operation using parallelization. (5 marks).</p>	10

4	<p>You are working as a database architect for a logistics company that operates across multiple regions. The company's database is distributed across different locations, with separate tables stored in different regions:</p> <ul style="list-style-type: none"> • Drivers: Contains information about drivers (DriverID, Name, LicenseNo, Region), and is stored in a regional database for Region A. • Deliveries: Contains information about deliveries made by drivers (DeliveryID, DriverID, Date, DeliveryLocation), and is stored in a different database for Region B. <p>The company wants to generate a report listing all deliveries made by drivers who are based in Region A. However, the size of the Deliveries table is extremely large, and transferring the entire table from Region B to Region A for the join operation is inefficient and time-consuming.</p> <ol style="list-style-type: none"> i. Explain a suitable algorithm to carry out this operation with the focus on reducing the communication cost. (4 marks) ii. Illustrate working of the algorithm for the following database state. (6 marks) <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">DriverID</th><th style="text-align: left;">Name</th><th style="text-align: left;">LicenseNo</th><th style="text-align: left;">Region</th></tr> </thead> <tbody> <tr><td>1</td><td>Amit</td><td>ABC123</td><td>A</td></tr> <tr><td>2</td><td>Bavan</td><td>XYZ456</td><td>A</td></tr> <tr><td>3</td><td>Chandru</td><td>PWE328</td><td>A</td></tr> <tr><td>4</td><td>Dave</td><td>RST234</td><td>B</td></tr> </tbody> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">DeliveryID</th><th style="text-align: left;">DriverID</th><th style="text-align: left;">Date</th><th style="text-align: left;">DeliveryLocation</th></tr> </thead> <tbody> <tr><td>101</td><td>1</td><td>2024-09-21</td><td>Chennai</td></tr> <tr><td>102</td><td>2</td><td>2024-09-22</td><td>Mumbai</td></tr> <tr><td>103</td><td>3</td><td>2024-09-24</td><td>Madurai</td></tr> <tr><td>104</td><td>4</td><td>2024-09-24</td><td>Ahmedabad</td></tr> <tr><td>105</td><td>5</td><td>2024-09-25</td><td>Delhi</td></tr> <tr><td>106</td><td>1</td><td>2024-09-26</td><td>Chennai</td></tr> </tbody> </table>	DriverID	Name	LicenseNo	Region	1	Amit	ABC123	A	2	Bavan	XYZ456	A	3	Chandru	PWE328	A	4	Dave	RST234	B	DeliveryID	DriverID	Date	DeliveryLocation	101	1	2024-09-21	Chennai	102	2	2024-09-22	Mumbai	103	3	2024-09-24	Madurai	104	4	2024-09-24	Ahmedabad	105	5	2024-09-25	Delhi	106	1	2024-09-26	Chennai	10
DriverID	Name	LicenseNo	Region																																															
1	Amit	ABC123	A																																															
2	Bavan	XYZ456	A																																															
3	Chandru	PWE328	A																																															
4	Dave	RST234	B																																															
DeliveryID	DriverID	Date	DeliveryLocation																																															
101	1	2024-09-21	Chennai																																															
102	2	2024-09-22	Mumbai																																															
103	3	2024-09-24	Madurai																																															
104	4	2024-09-24	Ahmedabad																																															
105	5	2024-09-25	Delhi																																															
106	1	2024-09-26	Chennai																																															
5	<p>You work as a database engineer for a logistics company that specializes in last-mile delivery services in a large metropolitan area. The company has a spatial database that stores geographic information about the city, including:</p> <ul style="list-style-type: none"> • Locations: The coordinates (latitude, longitude) of customers' addresses. • Warehouses: The coordinates of various warehouse locations. • Road Network: The spatial representation of the city's road network, stored as polylines (sequences of coordinates that represent roads). • Traffic Zones: Areas in the city where there are specific traffic regulations (e.g., no-entry zones, restricted access, or heavy traffic areas) represented as polygons. <p>The logistics team is working to optimize delivery routes for drivers, and they have asked you to use the spatial database.</p> <ol style="list-style-type: none"> i. Write a spatial query to find the Nearest Warehouse for each 	10																																																

- | | | | |
|-------------------------------|--|---|--|
| | | <p>ii. customer. (3 marks)
Write a spatial query to find all customers within a 5 km radius of a warehouse (2 marks)
iii. Suggest a suitable indexing mechanism to optimize the above queries (5 marks)</p> | |
| <p>*****All the best*****</p> | | | |

Reg. No.: 24MCS1017

Name : சிவாஸ்காலாப்

**VIT**

Vellore Institute of Technology

CONTINUOUS ASSESSMENT TEST 2 – OCTOBER 2024

Programme	: M.Tech. CSE (Big-Data Analytics)	Semester	: FS 2024-25
Course Title	: Data Visualization	Code	: MCSE616L
Faculty	: Sellam V	Class Nbr(s)	: CH2024250103237
Time	: 90 Minutes	Slot	: F2
		Max. Marks	: 50

Answer ALL Questions

*Assume Necessary data

**Q.N
0****Question Description****Mar
k
10**

1. Create an interactive choropleth map for a city council to visualize the relationship between socioeconomic status and infrastructure quality across various districts. The dataset includes:

District	Median Household Income (\$)	Infrastructure Quality Score (1-10)	Population Density (people/sq mi)
1	75,000	8	5,500
2	40,000	5	15,000
3	90,000	9	3,200
4	55,000	6	8,000
5	30,000	4	12,000

- a) Create a colour mapping strategy for **Median Household Income** and **Infrastructure Quality Score**. Explain your choices of colour gradients and how to visually incorporate **Population Density**. (4 Marks)
- b) Calculate the correlation coefficients between the socioeconomic factors and propose a method to visualize these correlations (e.g., scatter plots). Provide recommendations for urban planning based on these insights. (6 Marks)

2

You are analysing a dataset from a Neolithic archaeological site with the following variables:

10

Artifact Type	Material	Estimated Age (years)	Location (X, Y)	Function	Size (cm)	Condition	Quantity Found
Pottery	Clay	3500	(10, 20)	Domestic	15	Intact	5
Tools	Stone	4000	(12, 18)	Toolmaking	25	Fragmented	3
Animal Remains	Bone	4200	(15, 30)	Ritual	10	Weathered	2
Pottery	Clay	3000	(5, 5)	Domestic	20	Intact	10
Tools	Metal	4500	(25, 40)	Toolmaking	30	Intact	1
Animal Remains	Bone	3700	(20, 25)	Ritual	5	Fragmented	4
Pottery	Stone	3300	(30, 35)	Domestic	18	Weathered	7
Tools	Clay	3800	(40, 10)	Toolmaking	28	Intact	2
Animal Remains	Bone	3600	(22, 28)	Ritual	12	Intact	6

- a) Use appropriate visualisation techniques to represent the relationships among the following four variables: **Estimated Age**, **Size**, **Condition**, and **Quantity Found**. Justify the effectiveness of this technique in illustrating the correlations and distributions among these variables. (5 Marks)
- b) Use a hyperbox method and scatter plot method to analyse relationships among **Artifact Type**, **Material**, **Function**, and compare the results of both findings. (5 Marks)

3. Spatial data for a city planning project involves various urban features:

10

Feature Type	Location (X, Y)	Height (m)	Area (sq m)	Traffic Density (vehicles/hr)	Public Transport Availability
Building	(10, 20)	30	N/A	150	Medium
Park	(15, 25)	N/A	500	N/A	High
Road	(5, 15)	N/A	N/A	300	Low
Transport Route	(20, 30)	N/A	N/A	100	High
Building	(25, 10)	45	N/A	200	Medium

- a) Propose an icon-based geometric projection technique. Specify icons for each feature type, a colour scheme for public transport availability, and how icon size reflects height and area. Justify how your design communicates spatial relationships. (5 Marks)
- b) Discuss how pixel-based visualization can be used for urban planning decisions like resource allocation and infrastructure improvement, by considering insights obtained from traffic density and public transport availability. (5 Marks)

4. The company tracks sales data across 5 regions (North, South, East, West, Central) and for each month of the year. Your task is to create a heat map to visualize this data.

10

Region

Region	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec
North	1000	1200	1100	1050	1150	1250	1300	1400	1350	1200	1100	1400
South	950	1000	1050	980	1150	1100	1200	1300	1250	1150	1050	1300
East	850	900	950	910	1000	950	1100	1200	1150	1050	1000	1150
West	800	850	900	870	950	900	950	1000	980	900	850	1000
Central	1050	1100	1200	1150	1250	1300	1350	1400	1300	1250	1200	1450

- a) Based on the provided sales data, design a heat map to represent the sales performance across different regions and months. Describe the suitable elements of your heat map design. (6 Marks)
- b) Explain how this heat map can be used to gain insights into sales performance. Discuss how you would use it to identify data trends, patterns, and anomalies. Also, describe how interactive features like filtering or tooltips could enhance the analysis. (4 Marks)

Visualize student performance data from a university with the following variables:

10

Student ID	Course	Exam Score	Attendance (%)	Term	Gender
001	Mathematics	85	95	Fall 2023	Female
002	Science	78	85	Fall 2023	Male
003	Literature	92	90	Fall 2023	Female
004	Mathematics	65	70	Spring 2024	Male
005	Science	88	80	Spring 2024	Female
006	Literature	75	60	Spring 2024	Male

Tasks:

- a) Create a Tableau visualization using Exam Score, Attendance, and Gender. Explain your design choices. (5 Marks)
- b) Interpretation and Insights:
 - a. Analyze the visualization to answer:
 - i. How do exam scores correlate with attendance? (2 Marks)
 - ii. Are there performance differences between genders across courses? (2 Marks)
 - iii. What recommendations can you provide to improve performance for lower-scoring groups? (1 Marks)

*****All the best *****

Thank you
Maan