



VIT

Vellore Institute of Technology
CHENNAI

Reg. Number:

29MC51017

காலை நடவடிக்கை

Continuous Assessment Test (CAT) – I - AUGUST 2024

Programme	:	M.Tech CSE & its Specializations	Semester	:	Fall 2024-25
Course Code & Course Title	:	MCSE501L & Data Structures and Algorithms	Class Number	:	CH2024250103140 CH2024250103180
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	A software application that processes a list of items where each item requires a certain amount of additional processing as the list grows. For each additional item added to the list, the processing time increases linearly with the number of items processed so far. Base cost for the first item is $T(1)=O(1)$. Calculate the total cost for the recurrence relation $T(n)=T(n-1)+n$ using the back substitution method in order to model the total processing cost.	10
2	Transform the given expression $((5+3)*(10-2)/4)$ into postfix notation showing the steps involved, so that calculators can evaluate them using an appropriate data structure without worrying about operator precedence or parentheses. Following this, evaluate the resulting postfix expression by detailing each step of the process.	10
3	Develop a task management application where users can create and manage a list of tasks. Each task has attributes such as a description, due date, and priority. The tasks need to be stored and managed in a dynamic data structure such as a singly linked list that allows for efficient traversal and display operations. Write the Pseudocode for creation and display of the tasks and print the tasks in reverse order using an auxiliary data structure.	10
4	A multi-stage manufacturing process uses sensors to monitor various stages of production. These sensors record data such as temperature, pressure, or other metrics at different points in the production line. The data collected from these sensors can be complex, often exhibiting a pattern where values initially increase and then decrease due to the nature of the process (e.g., temperature rising and then falling as the machine cycles through stages). Efficiently sort sensor data that exhibits a bitonic pattern to analyze trends and anomalies. Sensor data collected from a production line might look like this: [3, 7, 12, 8, 6, 2, 1, 4] Use bitonic sorting to efficiently sort the data, leveraging the known bitonic pattern or sequence to reduce computational complexity. Provide a step-by-step trace using the given values.	10

5	A financial application provides real-time stock price information to investors. The application maintains a large, sorted dataset of historical stock prices for quick lookup using interpolation search. Users can input a specific price, and the app quickly retrieves information related to that price. Dataset: The sorted dataset of stock prices, e.g., prices = [100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150]. Input: The investor gives target stock price they are interested in, say target_price = 115. Demonstrate the step-by-step procedure for determining the given target price.	10
---	--	----

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



Continuous Assessment Test (CAT) – I - August 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.		To complete a given job two persons have written two different programs namely Program A and B. Program A is run on machine A whose cycle time is 1 nano seconds (ns). Program A uses 10 instructions. 50% of the instructions in the program are completed in one clock cycle. Remaining instructions consume 2 clock cycles to complete. Program B is run on machine B whose clock cycle time is 1.1 ns. Program B uses 8 instructions and each of them takes 2 clock cycles to complete. Evaluate the performance of these two programs A and B and identify the optimal program.	10
2.	a)	Consider two alternatives for speeding up computation. In the first alternative, we make 20% of a program 90 times faster than original processor. In the second alternative, we make 95% of the program 15 times faster than original processor. Calculate the ratio of the speedups for the above two cases. [5 Marks]	10
		b) Compare the features of RISC and CISC architectures in terms of the instruction set architecture, instruction formats and lengths, and addressing modes. [5 Marks]	
3.		A block – set associative cache memory consists of 128 blocks divided into four block sets. The main memory consists of 16384 blocks and each blocks contains 256 eight-bit words. i) Calculate the number of bits required for addressing the main memory. [5 Marks] ii) Calculate the number of bits needed to represent the TAG, SET and WORD fields. [5 Marks]	10
4.		ON chip cache hit rate is 90%, hit time is 5 ns, OFF chip cache hit rate is 96%, hit time is 10 ns, main memory hit rate is 99.8%, hit time is 60 ns, memory miss penalty is 10 milliseconds. Calculate the Average Memory Access time (AMAT).	10 6.32×1.56
5.		Consider an Associative cache with a total of 4 cache blocks. The main memory block requests are as follows: 10, 55, 11, 4, 13, 8, 132, 129, 212, 129, 64, 8, 48, 32, 73, 92 Calculate the number of misses and the miss ratio for the block replacement strategy as follows: i) First In First Out (FIFO) [5 Marks] ii) Least Recently Used (LRU) [5 Marks]	10

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



VIT

Vellore Institute of Technology
CHENNAI

Reg. Number: 24MCS1017

HESITATION

Continuous Assessment Test (CAT) – 1 - AUG 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>The CPU is executing a text editing application, actively managing user inputs, display updates, and autosave features. During this, the network interface card (NIC) receives an incoming network packet, possibly related to a file download or a message notification. The NIC triggers a hardware interrupt, signalling the CPU to process the incoming packet.</p> <p>i) Explain in details how the CPU handles this scenario. Your explanation should include the steps involved in saving the current process state, handling the interrupt, and resuming the original process. (7 Marks)</p> <p>ii) Use a timing diagram for this scenario to explain the whole process. (3 Marks)</p>	10
2		<p>Consider a scenario where an operating system is being designed for a new autonomous vehicle system. The system must handle real-time processing for critical tasks like sensor data analysis, navigation, and collision avoidance, while also supporting non-critical tasks such as user interface management and infotainment services.</p> <p>i) Given the need for high reliability, modularity, and real-time performance, which operating system structure (monolithic, microkernel, modular, or layered) would be most suitable for this application? Justify your choice with the help of a diagram, by discussing how the chosen structure would manage the critical and non-critical tasks, ensure system stability, and allow for easy updates and maintenance. (7 Marks)</p> <p>ii) State one reason each to reject the others in the list of operating system structures for the given scenario. (3 Marks)</p>	10

3		<p>In a scenario where a CPU is managing multiple processes, describe how non-pre-emptive and pre-emptive scheduling methods affect process state transitions. Using individual process state transition diagrams, illustrate how each scheduling approach impacts the movement of processes between states. Also, mention the different schedulers and their respective roles in managing the process scheduling.</p>	10																								
4		<p>You are managing a real-time operating system that uses the Priority Scheduling algorithm with pre-emption. The processes have the following details, where a lower number indicates a higher priority:</p> <table border="1"> <thead> <tr> <th>Process ID</th> <th>Arrival Time</th> <th>Burst Time</th> <th>Priority</th> </tr> </thead> <tbody> <tr> <td>P1</td> <td>1</td> <td>10</td> <td>2</td> </tr> <tr> <td>P2</td> <td>2</td> <td>5</td> <td>1</td> </tr> <tr> <td>P3</td> <td>2</td> <td>3</td> <td>2</td> </tr> <tr> <td>P4</td> <td>4</td> <td>8</td> <td>1</td> </tr> <tr> <td>P5</td> <td>3</td> <td>2</td> <td>3</td> </tr> </tbody> </table> <p>Draw a Gantt Chart for the processes and find the sequence of completion of the processes for this scenario. Calculate the average turnaround time and average waiting time. Compare these metrics with the non-pre-emptive version of the same scheduling algorithm.</p>	Process ID	Arrival Time	Burst Time	Priority	P1	1	10	2	P2	2	5	1	P3	2	3	2	P4	4	8	1	P5	3	2	3	10
Process ID	Arrival Time	Burst Time	Priority																								
P1	1	10	2																								
P2	2	5	1																								
P3	2	3	2																								
P4	4	8	1																								
P5	3	2	3																								
5		<p>You are working on a multi-threaded application where two threads, A and B, need to access few shared variables. Device a mechanism to ensure the consistency of the variables. Elaborate the implementation with the help of a pseudo code and thereby explaining it in a step-wise manner. Discuss any two potential limitations of the suggested solution.</p>	10																								

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



VIT

Vellore Institute of Technology
VELLORE - 600 024
CHENNAI - 600 097

Reg. Number:

24MCS1017

மத்தீர் அமைதியு

Continuous Assessment Test (CAT) – I - AUG 2024

Programme	:	M.Tech. CSE & 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>Assume that you are planning to develop an online shopping cart system. Discuss in detail how three schema architecture guarantees proper insulation between application, program code and data for the online shopping cart system. Clearly state all your assumptions in detail pertaining to the architecture considered for this system. (7 marks)</p> <p>b) Consider that a College schema contains two relations namely Student and Department. The Student relation captures the essential details of students such as RegNo, Name, BirthDate, MobileNumber, City and DeptID. The Department relation contains the necessary details of a department such as DepartmentID, DepartmentName, HoD. Identify the various possible keys and constraints that need to be imposed over the relations Student and Department such that the Data Integrity is not compromised. (3 marks)</p>	[10]
2		<p>A furniture shop wants to sell different types of furniture in online mode. The customers can view the furniture and place orders online.</p> <ul style="list-style-type: none"> • The portal sells different types of furniture such as chairs, tables, sofa etc. • For each of the furniture, the portal maintains its unique code, furniture name, the furniture description (such as wooden, plastic), the year it was manufactured and price. • Customers can place an order for the furniture of their interest. Information about customers such as their Name, PAN number, address, phone number and email are collected during ordering. • A customer can place more than one order. Each order may comprise of more than one furniture. The customers who place orders with order amount exceeding Rs. 50000 are given a discount of 5% on the order price. • In each order, the furniture code, quantity, furniture name, order amount and discount are recorded. <p>Draw the ER-diagram and make sure to indicate key attributes, cardinality and participation constraints.</p>	[10]

3	<p>Consider the relational schema R of a TV Showroom as given in the following schema. It maintains the orders for TV items placed by different customers in the relation:</p> <p>$R(\text{OrderId}, \text{OrderDate}, \text{CustomerID}, \text{CustomerName}, \text{CustomerAddress}, \text{ItemId}, \text{ItemName}, \text{ItemBrand}, \text{Price}, \text{Quantity})$.</p> <p>$\text{OrderId} \rightarrow \text{OrderDate}, \text{CustomerID}, \text{CustomerName}, \text{CustomerAddress}, \text{Quantity}$</p> <p>$\text{CustomerID} \rightarrow \text{CustomerName}, \text{CustomerAddress}$</p> <p>$\text{ItemId} \rightarrow \text{ItemName}, \text{ItemBrand}, \text{Price}$</p> <p>i. Decompose the schema into highest possible normal form and do not forget to indicate keys of the resulting relations. (7 marks)</p> <p>ii. Also state whether your decompositions are lossless and dependency preserving. (3 marks)</p>	[10]																																																												
4	<p>Consider the following schema of a Garment Shop which sells a variety of garments retailed from popular brands. Keys of the table are underlined.</p> <p>Assume that there are 10,000 tuples in Order table, 10,000 tuples in Customer table, 100 tuples in Brand table, and 1500 tuples in Item table.</p> <p>$\text{Item}(\underline{\text{Itemid}}, \text{Iname}, \text{Itype}, \text{Price})$</p> <p>$\text{Customer}(\underline{\text{Cid}}, \text{Cname}, \text{Mobile}, \text{City})$</p> <p>$\text{Brand}(\underline{\text{Bid}}, \text{Bname}, \text{Bcity})$</p> <p>$\text{Order}(\underline{\text{Order\#}}, \text{Itemid}, \text{Cid}, \text{Bid}, \text{NetAmount})$</p> <p>i. Write relational algebra expression to find names of customers who have ordered all types of garments from Levis brand. (2 marks)</p> <p>ii. Construct at least two possible query trees to list names of all customers who have ordered garments from Brands located in East zone of Mumbai. Also show which of the trees is optimal using the query optimization heuristic rules. (8 marks)</p> <p>iii.</p>	[10]																																																												
5	<p>Assume that you are working as a database administrator in an organization and your responsibilities include designing and managing database objects. Following is a sample of collection of tuples from one of the database table "Employee".</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left; padding: 2px;">EmpId</th> <th style="text-align: left; padding: 2px;">Ename</th> <th style="text-align: left; padding: 2px;">Address</th> <th style="text-align: left; padding: 2px;">Designation</th> <th style="text-align: left; padding: 2px;">Salary</th> </tr> </thead> <tbody> <tr> <td style="text-align: left; padding: 2px;">10</td> <td style="text-align: left; padding: 2px;">Rohan</td> <td style="text-align: left; padding: 2px;">Chennai</td> <td style="text-align: left; padding: 2px;">Project Lead</td> <td style="text-align: left; padding: 2px;">86000</td> </tr> <tr> <td style="text-align: left; padding: 2px;">198</td> <td style="text-align: left; padding: 2px;">Dheeksha</td> <td style="text-align: left; padding: 2px;">Mumbai</td> <td style="text-align: left; padding: 2px;">Team Lead</td> <td style="text-align: left; padding: 2px;">59000</td> </tr> <tr> <td style="text-align: left; padding: 2px;">22</td> <td style="text-align: left; padding: 2px;">Sudeep</td> <td style="text-align: left; padding: 2px;">Chennai</td> <td style="text-align: left; padding: 2px;">Developer</td> <td style="text-align: left; padding: 2px;">48000</td> </tr> <tr> <td style="text-align: left; padding: 2px;">200</td> <td style="text-align: left; padding: 2px;">Ajay</td> <td style="text-align: left; padding: 2px;">Mumbai</td> <td style="text-align: left; padding: 2px;">Project Lead</td> <td style="text-align: left; padding: 2px;">56000</td> </tr> <tr> <td style="text-align: left; padding: 2px;">110</td> <td style="text-align: left; padding: 2px;">Wilson</td> <td style="text-align: left; padding: 2px;">Mumbai</td> <td style="text-align: left; padding: 2px;">Team Lead</td> <td style="text-align: left; padding: 2px;">60000</td> </tr> <tr> <td style="text-align: left; padding: 2px;">100</td> <td style="text-align: left; padding: 2px;">Ahmed</td> <td style="text-align: left; padding: 2px;">Chennai</td> <td style="text-align: left; padding: 2px;">Developer</td> <td style="text-align: left; padding: 2px;">45000</td> </tr> <tr> <td style="text-align: left; padding: 2px;">125</td> <td style="text-align: left; padding: 2px;">Supriya</td> <td style="text-align: left; padding: 2px;">Delhi</td> <td style="text-align: left; padding: 2px;">Developer</td> <td style="text-align: left; padding: 2px;">48000</td> </tr> <tr> <td style="text-align: left; padding: 2px;">68</td> <td style="text-align: left; padding: 2px;">Ashwin</td> <td style="text-align: left; padding: 2px;">Ahmedabad</td> <td style="text-align: left; padding: 2px;">Developer</td> <td style="text-align: left; padding: 2px;">48000</td> </tr> <tr> <td style="text-align: left; padding: 2px;">99</td> <td style="text-align: left; padding: 2px;">Rakesh</td> <td style="text-align: left; padding: 2px;">Hyderabad</td> <td style="text-align: left; padding: 2px;">Developer</td> <td style="text-align: left; padding: 2px;">48000</td> </tr> <tr> <td style="text-align: left; padding: 2px;">45</td> <td style="text-align: left; padding: 2px;">Laxmi</td> <td style="text-align: left; padding: 2px;">Chennai</td> <td style="text-align: left; padding: 2px;">Developer</td> <td style="text-align: left; padding: 2px;">45000</td> </tr> <tr> <td style="text-align: left; padding: 2px;">156</td> <td style="text-align: left; padding: 2px;">Preeti</td> <td style="text-align: left; padding: 2px;">Noida</td> <td style="text-align: left; padding: 2px;">Team Lead</td> <td style="text-align: left; padding: 2px;">59000</td> </tr> </tbody> </table>	EmpId	Ename	Address	Designation	Salary	10	Rohan	Chennai	Project Lead	86000	198	Dheeksha	Mumbai	Team Lead	59000	22	Sudeep	Chennai	Developer	48000	200	Ajay	Mumbai	Project Lead	56000	110	Wilson	Mumbai	Team Lead	60000	100	Ahmed	Chennai	Developer	45000	125	Supriya	Delhi	Developer	48000	68	Ashwin	Ahmedabad	Developer	48000	99	Rakesh	Hyderabad	Developer	48000	45	Laxmi	Chennai	Developer	45000	156	Preeti	Noida	Team Lead	59000	[10]
EmpId	Ename	Address	Designation	Salary																																																										
10	Rohan	Chennai	Project Lead	86000																																																										
198	Dheeksha	Mumbai	Team Lead	59000																																																										
22	Sudeep	Chennai	Developer	48000																																																										
200	Ajay	Mumbai	Project Lead	56000																																																										
110	Wilson	Mumbai	Team Lead	60000																																																										
100	Ahmed	Chennai	Developer	45000																																																										
125	Supriya	Delhi	Developer	48000																																																										
68	Ashwin	Ahmedabad	Developer	48000																																																										
99	Rakesh	Hyderabad	Developer	48000																																																										
45	Laxmi	Chennai	Developer	45000																																																										
156	Preeti	Noida	Team Lead	59000																																																										

i. Assume you are using a dynamic multi-level indexing structure for the Employee table. Then which attribute of the Employee relation is chosen for indexing? Justify. (2 marks)

ii. Assume the tuples are inserted into the Employee table in the same order as shown above. Create dynamic multi-level indexing for the Employee table. You need to illustrate every step in building the indexing. (8 marks)

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

Reg. No.: 24MC51017

Name : மதீரா நகதாப்

**VIT[®]**

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

CONTINUOUS ASSESSMENT TEST 1 – AUGUST 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 o	Question Description	Mark																																																																								
1.	<p>You have been hired as a data visualization expert in a logistics company that manages shipments, inventory, and customer feedback. The company requires to deploy the following tasks:</p> <ul style="list-style-type: none"> a) Develop a comprehensive visualization strategy for i) Shipment Tracking, ii) Inventory Management and iii) Customer feedback analysis. Describe how you would abstract the data for each task, the types of visualizations you would use, and why these visualizations are appropriate. (6 Marks) b) For each of the visualizations proposed, explain how you would validate them using the four levels of validation: Internal Consistency, External Consistency, Face Validity, and Predictive Validity. Provide specific examples for each validation level. (4 Marks) 	10																																																																								
2	<p>You have been hired by a healthcare organization to visualize the various aspects of patient data.</p> <ul style="list-style-type: none"> • Patient Demographics: Understand the distribution of patients by age group, gender, and geographic location. • Treatment Outcomes: Evaluate the effectiveness of different treatments by comparing success rates and recovery times for dengue cases. • Hospital Performance Metrics: Track key performance indicators (KPIs) such as average wait times, patient satisfaction scores, and number of admissions in Out Patients Department. <p>Design 5 different statistical charts to monitor and analyze the above requirements.</p>	10																																																																								
3.	<p>You are tasked with developing a vector-based visualization system for a weather forecasting company that needs to display different requirements with the given meteorological data</p> <table border="1"> <thead> <tr> <th>Location ID</th> <th>Latitude</th> <th>Longitude</th> <th>Date</th> <th>Wind Speed (km/h)</th> <th>Wind Direction (degree)</th> <th>Temp (°C)</th> <th>Rainfall intensity (mm/hr)</th> </tr> </thead> <tbody> <tr><td>1</td><td>34.0522</td><td>-118.2437</td><td>2024-08-01</td><td>15</td><td>270</td><td>22</td><td>5.0</td></tr> <tr><td>1</td><td>34.0522</td><td>-118.2437</td><td>2024-08-02</td><td>12</td><td>260</td><td>23</td><td>0.0</td></tr> <tr><td>2</td><td>34.0522</td><td>-118.2537</td><td>2024-08-01</td><td>20</td><td>90</td><td>24</td><td>7.0</td></tr> <tr><td>2</td><td>34.0522</td><td>-118.2537</td><td>2024-08-02</td><td>18</td><td>85</td><td>25</td><td>2.0</td></tr> <tr><td>3</td><td>34.0622</td><td>-118.2437</td><td>2024-08-01</td><td>10</td><td>180</td><td>20</td><td>3.0</td></tr> <tr><td>3</td><td>34.0622</td><td>-118.2437</td><td>2024-08-02</td><td>8</td><td>175</td><td>21</td><td>1.0</td></tr> <tr><td>4</td><td>34.0622</td><td>-118.2537</td><td>2024-08-01</td><td>25</td><td>45</td><td>26</td><td>8.0</td></tr> <tr><td>4</td><td>34.0622</td><td>-118.2537</td><td>2024-08-02</td><td>22</td><td>50</td><td>27</td><td>6.0</td></tr> </tbody> </table>	Location ID	Latitude	Longitude	Date	Wind Speed (km/h)	Wind Direction (degree)	Temp (°C)	Rainfall intensity (mm/hr)	1	34.0522	-118.2437	2024-08-01	15	270	22	5.0	1	34.0522	-118.2437	2024-08-02	12	260	23	0.0	2	34.0522	-118.2537	2024-08-01	20	90	24	7.0	2	34.0522	-118.2537	2024-08-02	18	85	25	2.0	3	34.0622	-118.2437	2024-08-01	10	180	20	3.0	3	34.0622	-118.2437	2024-08-02	8	175	21	1.0	4	34.0622	-118.2537	2024-08-01	25	45	26	8.0	4	34.0622	-118.2537	2024-08-02	22	50	27	6.0	10
Location ID	Latitude	Longitude	Date	Wind Speed (km/h)	Wind Direction (degree)	Temp (°C)	Rainfall intensity (mm/hr)																																																																			
1	34.0522	-118.2437	2024-08-01	15	270	22	5.0																																																																			
1	34.0522	-118.2437	2024-08-02	12	260	23	0.0																																																																			
2	34.0522	-118.2537	2024-08-01	20	90	24	7.0																																																																			
2	34.0522	-118.2537	2024-08-02	18	85	25	2.0																																																																			
3	34.0622	-118.2437	2024-08-01	10	180	20	3.0																																																																			
3	34.0622	-118.2437	2024-08-02	8	175	21	1.0																																																																			
4	34.0622	-118.2537	2024-08-01	25	45	26	8.0																																																																			
4	34.0622	-118.2537	2024-08-02	22	50	27	6.0																																																																			

5	34.0722	-118.2437	2024-08-01	18	135	21	2.0
5	34.0722	-118.2437	2024-08-02	20	130	22	3.0
6	34.0722	-118.2537	2024-08-01	22	225	23	6.0
6	34.0722	-118.2537	2024-08-02	24	230	24	4.0

The company wants to use this system to monitor i) Wind Speed and Direction, ii) Temperature Data and iii) Rainfall Intensity and justify your answers with detailed insights and appropriate visualizations.

- 4 You are a data analyst working for a retail company that wants to enhance its customer segmentation strategy. Explain how you would segment customers based on given attributes, steps involved in the clustering process and interpretation aspects with K=3, initial centroids (Customer id – 1,5,9) 10

Customer ID	Purchases Per Month	Average Purchase Amt	Visits Per Year	Age Group
(1)	5	\$50	30	18-24
2	2	\$200	10	25-34
3	10	\$30	50	35-44
4	1	\$100	5	45-54
(5)	8	\$40	25	18-24
6	3	\$150	15	35-44
7	6	\$70	35	25-34
8	4	\$60	20	45-54
(9)	7	\$80	22	18-24

- 5 You are a data analyst at a financial firm that needs to create a comprehensive dashboard to monitor and analyse stock market given data as shown in the Table. The firm wants to make informed investment decisions and track market performance.

Month-Year	Avg Closing Price	Total Trading Volume
2024-01	145.00	40,000
2024-02	148.50	35,000
2024-03	150.75	42,000
2024-04	155.25	50,000
2024-05	157.00	45,000
2024-06	159.50	48,000
2024-07	155.00	13,550
2024-08	158.25	52,000
2024-09	160.00	47,000
2024-10	162.75	54,000
2024-11	164.50	49,000
2024-12	167.00	51,000

- a) Design a time series data visualization strategy for the following three aspects of stock market data: (6 Marks)
- Stock Price Trends: Visualize the historical prices of a specific stock over time.
 - Trading Volume Analysis: Display the trading volumes of the stock to identify periods of high activity.
 - Moving Averages and Market Trends: Evaluate moving averages of 3-mon, 6-mon and 9-mon fluctuations and represent it with corresponding visualizations.
- b) Discuss the potential challenges in visualizing time series data and suggest strategies to address these challenges. (4 Marks).

Total

50