



VIT
Vellore Institute of Technology
(Chartered by the Government of India under Section 3 of the U.R. Act, 1956)

School of Computer Science and Engineering

Winter Semester 2023-24

Continuous Assessment Test – I

SLOT: G2+TG2

Programme Name & Branch: B.Tech & BCE, BDS

Course Name & Code: BCSE206L & Foundations of Data Science

Class Number (s): ALL

Faculty Name (s): ALL

Exam Duration: 90 Min.

Maximum Marks: 50

Q. No.	Question	Max Marks																		
1.	Consider a logistics industry management system. Identify the need of data science in logistics industry management system to enhance business and management. Also describe in detail about uses of data science in logistics industry automation system.	10																		
2.	(a) Consider a fraud detection system in banking sector that requires to implement a set of proactive measures to detect and avoid fraudulent activities and financial losses. Illustrate the different steps involved in data science process with neat sketch with respect to the above scenario.	7																		
	(b) What is the role of a data scientist? What are the prerequisites to become a data scientist?	3																		
3.	(a) Consider the following table. Write SQL commands to find the mean, median and mode of gasoline cost and also write the mean, median and mode of gasoline cost. Table Name: Statecost <table><tr><th>State</th><th>Gasoline Cost (Rs.)</th></tr><tr><td>Tamilnadu</td><td>10</td></tr><tr><td>Kerala</td><td>5</td></tr><tr><td>Karnataka</td><td>7</td></tr><tr><td>Assam</td><td>8</td></tr><tr><td>UP</td><td>4</td></tr><tr><td>Andrapradesh</td><td>7</td></tr><tr><td>Bihar</td><td>20</td></tr><tr><td>Goa</td><td>25</td></tr></table>	State	Gasoline Cost (Rs.)	Tamilnadu	10	Kerala	5	Karnataka	7	Assam	8	UP	4	Andrapradesh	7	Bihar	20	Goa	25	6
State	Gasoline Cost (Rs.)																			
Tamilnadu	10																			
Kerala	5																			
Karnataka	7																			
Assam	8																			
UP	4																			
Andrapradesh	7																			
Bihar	20																			
Goa	25																			

(b) Consider the following table and write SQL queries to perform the following.

Table Name: Employee

Name	Department	City	Salary
anto	CSE	london	10000
cyan	IT	delhi	20000
sikkin	ECE	vellore	15000
sin	CSE	kurnool	25000
mathiew	IT	madhuboni	21000
suren	CSE	rajnandgaon	19000

- (i) Display the employee details whose name starts with either a, c, d, s
(ii) Display the employee details who is getting highest salary in the department "CSE"
(iii) Display all the employees whose city contains the pattern "on"
(iv) Display all the employees whose name starts with "s" and are at least 5 characters in length.

4. (a) Consider the following tables and write SQL commands to perform inner join, left join and full outer join. Also give the output.

Employee Table:

EmpID	Empname	Age
1	Kumar	22
2	Sharma	32
3	Shreshth	24
4	Sharma	25
5	Kapoor	26

Project Table:

ProjectID	EmpID	ClientID	ProjectName
111	1	3	Project1
222	2	1	Project2
333	3	5	Project3
444	3	2	Project4
555	5	4	Project5
666	9	1	Project6

(b) Consider the following table and identify the output of following queries.

Runners Table:

Name	Time
andy	101
bob	103
cedric	104
dave	104
eric	108

rank	dense rank	row	cume
1	1	0	0.2
2	2	0.25	0.4
3	3	0.5	0.8
3	3	0.5	0.8
5	4	1	1

- (i) select name, time, row_number() over (order by time), rank() over (order by time), dense_rank() over (order by time) FROM runners order by time;
(ii) select name, time, percent_rank() over (order by time), cume_dist() over (order by time) FROM runners order by time;

5. Identify the different phases of data analytics lifecycle and illustrate the key activities need to be performed in each phase with neat sketch.