



COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE


Programme: M.Sc. (Five Year Integrated) in Computer Science (Artificial Intelligence & Data Science)

Course Code & Title: 21-805-0504 R for Data Science

Name of Examination: Series 1	Max. Marks: 20	Semester: V
Batch: 2023-28	Duration: 2 Hours	Date: 18.09.2025
		Time: 9:00 AM – 11 AM

Answer all questions

1. Create a numeric vector named `my_numbers` containing the values 5, 10, 15, and 20. Then, calculate and print the sum and mean of the vector.
(2 marks) [CO1, DL-2, BTL-3]
2. Use a for loop to iterate through the vector `c(1, 2, 3, 4, 5)` and print the square of each number.
(2 marks) [CO1, DL-1, BTL-2]
3. Write a while loop that starts with a variable `i = 1` and prints `i` as long as `i` is less than or equal to 5. Increment `i` by 1 in each iteration
(2 marks) [CO1, DL-1, BTL-2]
4. Write an ifelse statement in R that assigns the value "positive" to a variable `x` if a given number `y` is greater than zero, "negative" if `y` is less than zero, and "zero" if `y` is equal to zero.
(3 marks) [CO1, DL-1, BTL-2]
5. Write a code in R that creates a matrix `m` with 3 rows and 4 columns, filled with random numbers between 0 and 1. Then, write a code that extracts the second column of `m` as a vector `v`.
(3 marks) [CO1, DL-1, BTL-2]
6. Write a code in R that converts a character vector `c("red", "green", "blue", "yellow")` into a factor variable `f` with levels ordered alphabetically.
(3 marks) [CO1, DL-1, BTL-3]
7. Define a dataframe. Create a dataframe of student having variables Student Name, Register No, Programme and Date of joining
(2 marks) [CO1, DL-1, BTL-2]
8. Give any 3 data visualisation tools including an example script.
(3 marks) [CO2, DL-1, BTL-2]

 COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY DEPARTMENT OF COMPUTER SCIENCE			
Programme: Fifth Semester Five Year Integrated MSc in Computer Science (Artificial Intelligence and Data Science) Course Code & Title 23-813-0501: Regression Analysis			
Name of Examination: Series I		Max. Marks: 30	Semester: V
Batch: 2023-28	Duration: 2 Hours	Date: 15..09.2024	Time: 9:00 AM – 11:00 AM

Answer all questions

1. Explain simple linear regression and the methods to find β_0 and β_1
(5 marks) [CO1, DL-1, BTL-4]
2. The effect of temperature of the deodorizing process on the color of the finished product was determined experimentally. The data collected were as follows. Fit a simple linear regression model $Y = \beta_0 + \beta_1 X + \epsilon$. Find the least square estimate of β_1

Temperature (X)	Color (Y)
100	0.3
150	0.4
140	0.4
130	0.4
120	0.5
110	0.5
150	0.5
140	0.6
130	0.6
120	0.6
110	0.7
100	0.6
120	0.6
110	0.6
100	0.6

3. Explain the properties of fitted regression lines?
(5 marks) [CO1, DL-1, BTL-1]
4. Explain the analysis of variance
(5 marks) [CO2, DL-2, BTL-2]
5. What are the tests for finding the significance of a regression model?
(5 marks) [CO2, DL-2, BTL-2]
6. Fit the model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$ to the data below.

X_1	X_2	Y
-5	5	11
-4	4	11
-1	1	8
2	-3	2
2	-2	5
3	-2	5
3	-3	4


What is the estimate of β_0 .

(5 marks) [CO4, DL-2, BTL-2]

SFI-DCS

IMSC

IMSC

 COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY DEPARTMENT OF COMPUTER SCIENCE			
Programme: V Semester Five Year Integrated M.Sc in Computer Science (Artificial Intelligence and Data Science)			
Course Code & Title: 23-813-0505 Number Theory and Cryptography			
Name of Examination: Series I			
Batch: 2023-28	Duration: 2 Hours	Max. Marks: 20	Semester: V
		Date: 19.09.2025	Time: 09:00 AM – 11:00 AM

Answer all questions

1.
 - a) Compute the greatest common divisor (gcd) of 54 and 67 and find the multiplicative inverse of 54 mod 67 using the Extended Euclidean Algorithm. (2 marks) [CO1, DL-2, BTL-3]
 - b) Find the residue classes modulo 5 and identify any two complete residue classes. (1 mark) [CO1, DL-1, BTL-3]
2.
 - a) Find $4^{532} \bmod 11$ using Fermat's theorem. (2 marks) [CO1, DL-2, BTL-3]
 - b) Prove that 17 is prime using Wilson's theorem. (1 mark) [CO1, DL-1, BTL-3]
3. Solve the following congruence equations to find x using Chinese Remainder Theorem.

$$x \equiv 2 \pmod{3}$$

$$x \equiv 3 \pmod{4}$$

$$x \equiv 1 \pmod{5}$$
 (3 marks) [CO1, DL-2, BTL-3]
4. Evaluate Jacobi's Symbol $\left(\frac{215}{253}\right)$. (2 marks) [CO1, DL-2, BTL-3]
5. Find primitive roots of Group $\langle \mathbb{Z}_5^*, X \rangle$ and check whether it is cyclic or not, where 'X' denotes multiplication. (3 marks) [CO1, DL-1, BTL-3]
6.
 - a) Express the irrational number $\sqrt{2}$ into infinite simple continuous fraction. (2 marks) [CO1, DL-2, BTL-3]
 - b) Use Pollard rho method to factor 1111 using $x_0 = 2$ and $f(x) = x^2 + 1$. (4 marks) [CO2, DL-2, BTL-3]



COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF COMPUTER SCIENCE


Programme: V Semester Five Year Integrated M.Sc in Computer Science (Artificial Intelligence and Data Science)

Course Code & Title: 23-813-0505 Number Theory and Cryptography

Name of Examination: Series I	Max. Marks: 20	Semester: V
Batch: 2023-28	Duration: 2 Hours	Date: 19.09.2025
		Time: 09:00 AM – 11:00 AM

Answer all questions

1.
 - a) Compute the greatest common divisor (gcd) of 54 and 67 and find the multiplicative inverse of 54 mod 67 using the Extended Euclidean Algorithm. (2 marks) [CO1, DL-2, BTL-3]
 - b) Find the residue classes modulo 5 and identify any two complete residue classes. (1 mark) [CO1, DL-1, BTL-3]
2.
 - a) Find $4^{532} \bmod 11$ using Fermat's theorem. (2 marks) [CO1, DL-2, BTL-3]
 - b) Prove that 17 is prime using Wilson's theorem. (1 mark) [CO1, DL-1, BTL-3]
3. Solve the following congruence equations to find x using Chinese Remainder Theorem.
$$\begin{aligned}x &\equiv 2 \pmod{3} \\x &\equiv 3 \pmod{4} \\x &\equiv 1 \pmod{5}\end{aligned}$$
(3 marks) [CO1, DL-2, BTL-3]
4. Evaluate Jacobi's Symbol $\left(\frac{215}{253}\right)$. (2 marks) [CO1, DL-2, BTL-3]
5. Find primitive roots of Group $\langle \mathbb{Z}_5^*, X \rangle$ and check whether it is cyclic or not, where 'X' denotes multiplication. (3 marks) [CO1, DL-1, BTL-3]
6.
 - a) Express the irrational number $\sqrt{2}$ into infinite simple continuous fraction. (2 marks) [CO1, DL-2, BTL-3]
 - b) Use Pollard rho method to factor 1111 using $x_0 = 2$ and $f(x) = x^2 + 1$. (4 marks) [CO2, DL-2, BTL-3]

 COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY DEPARTMENT OF COMPUTER SCIENCE			
Programme: M.Sc. (Five Year Integrated) in Computer Science Course Code & Title: 23-813-0502: Big Data Analytics			
Name of Examination: Series I		Max. Marks: 20	Semester: V
Batch: 2023-28	Duration: 2 Hours	Date: 16.09.2025	Time: 9:00 AM – 11:00 AM

Answer all questions

1. Two regression models are built to predict house prices. Their predictions are as follows

House	Actual Price (In Lakhs)	Model A Prediction	Model B prediction
1	50	52	49
2	60	63	61
3	70	65	72
4	80	78	85
5	90	88	95

- Compute MAE, MSE, RMSE for both models.
 - Find the R^2 for each model.
 - Decide which model is better, and justify your answer
- (5 marks) [CO1, DL-2, BTL-2]
2. Given the covariance matrix below, identify the first principal component, $C =$
- (5 marks) [CO1, DL-2, BTL-2]
3. Explain two techniques used for detecting outliers and two methods for handling outliers in machine learning datasets
- (4 marks) [CO1, DL-2, BTL-2]
4. Explain why KNN regression is considered a non-parametric and non-linear regression model. Explain its working principle.
- (3 marks) [CO2, DL-2, BTL-2]
5. Explain how regularization helps in reducing overfitting. Differentiate between L1, L2 and Elastic net.
- (3 marks) [CO2, DL-2, BTL-2]

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY DEPARTMENT OF COMPUTER SCIENCE	
---	--



COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE

Programme: Fifth Semester Five Year Integrated MSc in Computer Science
(Artificial Intelligence and Data Science)

Course Code & Title 23-813-0501: Regression Analysis

Name of Examination: Series I	Max. Marks: 30	Semester: V
Batch: 2023-28	Duration: 2 Hours	Date: 15.09.2024
		Time: 9:00 AM – 11:00 AM

Answer all questions

1. Explain simple linear regression and the methods to find β_0 and β_1
(5 marks) [CO4, DL-1, BTL-4]
2. The effect of temperature of the deodorizing process on the color of the finished product was determined experimentally. The data collected were as follows. Fit a simple linear regression model $Y = \beta_0 + \beta_1 X + \epsilon$. Find the least square estimate of β_1

Temperature (X)	Color (Y)
300	0.3
350	0.3
400	0.4
450	0.4
500	0.5
550	0.5
600	0.6
650	0.6
700	0.6
750	0.6
800	0.7
850	0.6
900	0.6
950	0.6
1000	0.6

- (5marks) [CO1, DL-1, BTL-1]
3. Explain the properties of fitted regression lines? (5 marks) [CO1, DL-1, BTL-1]
4. Explain the analysis of variance (5 marks) [CO2, DL-2, BTL-2]
5. What are the tests for finding the significance of a regression model?
(5 marks) [CO2, DL-2, BTL-2]

6. Fit the model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$ to the data below.

X_1	X_2	Y
-5	5	11
-4	4	11
-1	1	8
2	-3	2
2	-2	5
3	-2	5
3	-3	4

What is the estimate of β_0 .

(5 marks) [CO2, DL-2,

SFI-DCS

COCHI

Programme: Five Year Inter

Name