University of Central Punjab

Faculty of Information Technology and Computer Science

Course Title: Probability and Statistics Course Code: SESS-2733

Assignment 2

Peer Assignment

Total marks: 50 Obtained marks:\_\_\_

|  |  |
| --- | --- |
| Name | Roll number |
|  |  |
|  |  |
|  |  |
|  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **CLO #** | **Course Learning Outcome (CLO)** | **Taxonomy Level** | **Mapping to PLO** |
| CLO 3 | **Evaluate** regression models and use inferential statistical techniques, including hypothesis testing, to interpret and assess data in software engineering applications. | **C5** | **PLO 2** |

**Instructions:**

1. **Attempt all questions.**
2. **Write your answer showing all steps required to perform the task.**
3. **Assignment should be Hand Written. Computerized assignment is not accepted**
4. **Assignment should be submitted on A4 sheets or Assignment sheets only. Violation will result to deduction of 5 mark from the scored marks.**
5. **Each student will have attached this front page with his/her assignment. Violation will result to deduction of 3 mark from the scored marks.**
6. **Due Date for Assignment on portal is April 29, 2025. Till 12:00 pm**
7. **No Late submission accepted**

# Problem 1: Evaluate a linear regression model for productivity (15 marks)

**A company tracks how training hours affect employee productivity scores.**

| **Employee** | **Training Hours** | **Productivity Score** |
| --- | --- | --- |
| **1** | **2** | **50** |
| **2** | **4** | **55** |
| **3** | **6** | **60** |
| **4** | **8** | **68** |
| **5** | **10** | **75** |
| **6** | **12** | **78** |

**Tasks:**

1. Calculate the slope (m) and intercept (c) manually.
2. Predict productivity for 9 hours of training.
3. Explain what the slop and intercept mean in this scenario.
4. Would this model apply to someone with 0 hours of training?
5. Evaluate this regression model by using R-squared Method.

# Problem 2: Evaluate a linear regression model for sales (15 marks)

**A startup company wants to know if the number of followers affects sales on launch day.**

| **Influencer** | **Followers (1000s)** | **Launch Day Sales ($1000s)** |
| --- | --- | --- |
| **A** | **5** | **15** |
| **B** | **10** | **22** |
| **C** | **15** | **30** |
| **D** | **20** | **38** |
| **E** | **25** | **45** |
| **F** | **30** | **50** |

**Tasks:**

1. Calculate the slope (m) and intercept (c) manually.
2. Predict sales for an influencer with 18,000 followers.
3. Explain what the slop and intercept mean in this scenario.
4. Would this model apply to someone with 0 Followers?
5. Evaluate this regression model by using R-squared Method.

# Problem 3: Evaluate a multiple regression model for given scenario (20 marks)

# A training manager at a logistics company wants to determine how job tenure and training hours completed affect the employee performance score. The company suspects that both factors contribute significantly to how employees perform in operational tasks.

# The manager collects data from a sample of employees:

| **Employee** | **Job Tenure (Years)** | **Training Hours** | **Performance Score** |
| --- | --- | --- | --- |
| **L1** | **1** | **5** | **60** |
| **L2** | **3** | **10** | **75** |
| **L3** | **4** | **15** | **85** |
| **L4** | **6** | **20** | **95** |

**Tasks:**

1. Estimate the coefficients using Cramer Rule.
2. Predict the performance score for an employee with 5 years of tenure and 12 training hours completed.
3. Interpret the meaning of each coefficient in this scenario.
4. Evaluate whether both predictors are significant contributors to performance score.