University of Central Punjab

Faculty of Information Technology and Computer Science

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

Assignment 1

Peer Assignment

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

|  |  |  |  |
| --- | --- | --- | --- |
| **CLO #** | **Course Learning Outcome (CLO)** | **Taxonomy Level** | **Mapping to PLO** |
| CLO 2 | Apply statistical data using set theory, Venn diagrams, descriptive measures, probability distributions, and hypothesis testing to support  decision-making | **C3** | **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 8, 2025. Till 12:00 pm**
7. **No Late submission accepted**

# Problem 1: Data Types in Statistics

# Identify and select one data type for the following scenarios

1. A teacher records student grades as “A”, “B”, “C”, “D”, and “F” for an exam. She wants to analyze the distribution of grades in the class. \_\_\_\_\_\_\_\_\_\_\_\_\_\_
2. A weather department records the temperature in degrees Celsius across five cities. \_\_\_\_\_\_\_\_\_\_\_\_\_\_
3. A group of 500 people is asked to choose their favorite movie genre (Action, Comedy, Horror, Drama, Sci-Fi). \_\_\_\_\_\_\_\_\_\_\_\_
4. A company asks customers to rate their satisfaction with a product on a scale of 1 to 5 (1 = Very Dissatisfied, 5 = Very Satisfied). \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
5. A researcher records the height of 50 students in centimeters. \_\_\_\_\_\_\_\_\_\_\_\_
6. A mall manager counts the number of cars parked in the parking lot every hour. \_\_\_\_\_\_\_\_\_\_\_
7. A homeowner records their electricity bill amount (in dollars) each month. \_\_\_\_\_\_\_\_\_\_\_
8. The ATP ranks professional tennis players from #1 to #100 based on their performance. \_\_\_\_\_\_\_\_\_\_\_
9. A company collects the birth years of its employees to analyze age distribution. \_\_\_\_\_\_\_\_\_\_
10. A hospital records patients’ blood types as A, B, AB, or O. \_\_\_\_\_\_\_\_\_\_\_

# Problem 2: Graphical Visualization of data

# A Identify the graphical visualization type

# B justify your reason

# C write code with appropriate logic for decision

1. COVID-19 Cases in Different Countries

A health organization wants to compare the total number of COVID-19 cases in different countries. Which type of chart would best illustrate this comparison? Justify your reason and also wr

1. Stock Market Trends

An investor wants to track the stock prices of five major companies over the past year to identify trends and make investment decisions. Which type of graph should be used?

1. Income Distribution Among Employees

A company wants to analyze the distribution of salaries among its employees. They have categorized salaries into different ranges (e.g., $20K-$40K, $41K-$60K, etc.). Which type of chart would best represent this data?

# Problem 3 Measure of Centre tendency and dispersion

# Given the dataset {2, 5, 8, 5, 9, 12, 5}, find the mean, median, and mode.

1. Calculate the interquartile range (IQR) for the dataset {3, 7, 8, 12, 15, 18, 21}.

# Identify variance for the following scenario

# The HR manager wants to measure how much the salaries vary within the department of information technology in university of central Punjab.

|  |  |
| --- | --- |
| Employee | Salary |
| A | 30000 |
| B | 32000 |
| C | 29000 |
| D | 31000 |
| E | 33000 |

1. Calculate the **variance** of the salaries.
2. What does the variance tell you about salary distribution?
3. If variance is high, what does that indicate about salary differences among employees?

# Problem 4 Correlation analysis

# Using Pearson correlation analysis formula identify the value of r

# Identify the type of correlation between both variables for decision

# Draw graph and plot the data points

# A university professor wants to analyze whether there is a relationship between the number of hours students spend studying and their exam scores. She collects data from 10 students, recording the total hours they studied for an exam and their corresponding scores.

# ..

|  |  |  |
| --- | --- | --- |
| Sr no. | Student hour (x) | Score (y) |
| 1 | 2 | 50 |
| 2 | 3 | 55 |
| 3 | 5 | 65 |
| 4 | 7 | 70 |
| 5 | 8 | 75 |
| 6 | 10 | 85 |
| 7 | 12 | 90 |
| 8 | 14 | 92 |
| 9 | 15 | 94 |
| 10 | 16 | 96 |