Statistical foundation of Data Sciences

Practical- 05

Roll number: GF202344767

Name: Ishita Mehta

Workflow summary:

1. Import pandas for data handling.
2. Read the file using pd.read_csv("teacher_ratings_updated.csv").
3. Use df.head() and df.columns to see the data and column names.
4. Find total number of records using len(df).

Calculate probability for Q1:

1. Filter rows where eval > 4.5.
2. Divide the count by total records.

Calculate probability for Q2:

1. Filter rows where eval is between 3.5 and 4.2.
2. Divide the count by total records.
3. Print the calculated probabilities for both cases.

Questionl 3: (Two-Tailed Z-Test)

1. Import math and scipy.stats.norm.
2. Define given data:
    - Population mean (μ) = 12
    - Population SD (σ) = 5.5
    - Sample mean ($\bar{x}$) = 10.7
    - Sample size (n) = 36
3. Form hypotheses:
   H_0: μ = 12 (no difference)
   H_1: μ ≠ 12 (difference exists)
4. Compute Z-score:
5. Use formula

$$Z=(\bar{x}-\mu)/(\sigma/\sqrt{n})$$

6. Find critical Z value: For α = 0.05 (two-tailed), critical Z = ±1.96.
7. Decision rule:
   If |Z| > 1.96 → Reject H_0
   If |Z| ≤ 1.96 → Fail to reject H_0

Github Repository link:

https://github.com/pineapplesdontbelongonpizza/CSU1658_practical1_Testing_Pandas_and_Numpy.git