

**Subject Name: Statistical Foundation of Data Sciences**

**Subject Code: CSU1658 Practical**

**Instruction:**

- Please use either Google Colab or Jupyternotebook.
- You are required to make the synthetic data set which has some Nan Values as well. Keep the random seed to 42 in the beginning and then as per requirement can be changed.

**Problem 1:** Compute (a) mean, (b) median, and (c) age-weighted mean of income. Ignore NaNs where appropriate. Explain when a weighted mean is preferable.

**Problem 2:** Standardize income (z-score). Report how many incomes are outliers using rule  $|z| > 3$ . Handle NaNs correctly (do not drop entire rows unnecessarily).

**Problem 3:** Create age bins: [18-25), [25-35), [35-45), [45-60) and compute for each bin:

- count of observations,
- mean income,
- median score.

Show result as a tidy DataFrame sorted by age bin.

**Problem 4:** Create an array it cannot be of 1 Dimension. And then showcase the operation for the following:

- Shape and Resize → shape, size, Transpose, Flatten
- Showcasing negative indexing and display error while doing slicing
- Arithmetic Operations → Broadcasting, Dot Product
- Linear Algebra → Determinant, Inverse