Assignment for Students

Objective:
Perform data preprocessing on a dataset of employee salaries and job satisfaction scores.

Instructions:

Dataset:
Use the dataset named employee_data.csv with the following columns:

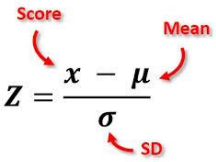
- **Employee_ID** (Integer)
- **Age** (Integer, has missing values)
- **Salary** (Float, has missing values)
- **Job_Satisfaction** (Scale of 1-10, has missing values)
- **Work_Hours_Per_Week** (Integer, with varying scales)

Tasks:

1. Load the dataset using Pandas.
2. Handle missing values using mean, median, or mode imputation.
3. Apply feature scaling using:
    - **Min-Max Scaling** (for Age and Salary)
        - One using built-in MinMaxScaler from sklearn
        - One manually using the formula:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

    - **Standardization (Z-score normalization)** (for Job_Satisfaction, Work_Hours_Per_Week)
        - One using built-in StandardScaler from sklearn
        - One manually using the formula:



4. Display the dataset before and after preprocessing.

Submission:

- Submit the Python script (.ipynb) with and pdf.