Assignment: Data Cleaning with Pandas

Objective: Your task is to clean and preprocess the given dataset using pandas. The dataset contains various inconsistencies, missing values, and formatting issues. Your goal is to identify and resolve these problems to prepare the data for further analysis.

Dataset: Download the datase: pandas_data_cleaning_assignment.csv

Tasks:

1. **Handling Missing Values:**
   ○ Identify columns with missing values.
   ○ Use appropriate strategies to fill or remove missing data.
2. **Fixing Inconsistent Data:**
   ○ Standardize gender values (e.g., 'M', 'Male', 'male' should be uniform).
   ○ Convert Join_Date into a standard date format.
3. **Handling Duplicates:**
   ○ Identify and remove duplicate records if necessary.
4. **Detecting and Treating Outliers:**
   ○ Check for outliers in the Salary column and handle them appropriately.
5. **Data Type Conversion:**
   ○ Ensure all columns have appropriate data types.
6. **Performance Rating:**
   ○ Fill in missing values in the Performance_Rating column using an appropriate method.
7. **Final Data Export:**
   ○ After cleaning, save the processed data into a new CSV file.

Submission:

● Submit a Jupyter Notebook (.ipynb) file with your pandas code, explanations, and the final cleaned dataset.
● Include markdown cells to explain the steps taken.

Bonus Challenge:

● Generate a summary report showing the before-and-after state of the dataset.

Good luck!