✅ Scaling, Encoding & Feature Creation

🎯 Aim

Apply **scaling** (Min-Max, Standardization) and **encoding** (One-Hot, Label) to numerical and categorical features.
Create a **new feature** by combining or transforming existing columns.

📌 Overview

This project demonstrates the **preprocessing steps** required before building a machine learning model. It covers:

- Handling **numerical** and **categorical** data
- Applying **scaling techniques**
- Performing **encoding methods**
- Creating **new features** through transformation or combination

---

🧩 Problem Statement

Raw datasets contain numerical and categorical values in different formats.
Without preprocessing:

- Machine learning models may misinterpret the data
- Feature scales may affect performance
- Categorical columns may be unreadable for algorithms

This project solves these issues through structured preprocessing.

🛠️ Tools and Technologies

- **Python**
- **Pandas** – Data manipulation
- **NumPy** – Numerical operations
- **Scikit-learn** – Scaling & Encoding
- **Jupyter Notebook / VS Code** (optional)

🔧 Methods Used

✅ **1. Scaling (Numerical Features)**

- **Min-Max Scaling**
- **Standardization (Z-score)**

✅ **2. Encoding (Categorical Features)**

- **Label Encoding**
- **One-Hot Encoding**

✅ **3. Feature Engineering**

- Create a new column by combining or transforming existing ones
  - Example: Total_Score = Marks1 + Marks2
  - Example: Age_Group = Age // 10 * 10

💡 **Key Insights**

✓ Scaling brings all numerical values to a comparable range
✓ Encoding helps convert categories into machine-readable format
✓ New feature creation improves data richness and prediction power

---

✅ **Results & Conclusion**

- All numerical features are **scaled properly**
- Categorical values are **encoded correctly**
- A **new feature** is created using transformation
- Data is now ready for **model training or EDA**