Task 1: Data Cleaning & Preprocessing - Titanic Dataset

✅ Objective:
To clean and preprocess the Titanic dataset for use in machine learning models.

📌 Steps Performed:
1. Loaded the dataset using Pandas.
2. Handled missing values:
   - Replaced missing 'Age' with median.
   - Replaced missing 'Embarked' with mode.
   - Dropped 'Cabin' due to many missing values.
3. Converted categorical variables:
   - Encoded 'Sex' using label encoding.
   - One-hot encoded 'Embarked'.
4. Scaled numerical features ('Age', 'Fare') using StandardScaler.
5. Removed outliers in 'Fare' using the IQR method.
6. Saved the cleaned dataset as `cleaned_titanic.csv`.

🛠 Tools Used:
- Python
- Pandas
- NumPy
- scikit-learn

📁 Files in this Repo:
- `titanic_cleaning.py`: Python script with all data cleaning steps.
- `cleaned_titanic.csv`: Output dataset after preprocessing.

📥 Dataset Source:
[Kaggle Titanic Dataset](https://www.kaggle.com/datasets/yasserh/titanic-dataset)

📝 Submission:
Submit this GitHub repository link using the form provided in the internship task.