<h1 align="center">🚢 Titanic Data Preprocessing | ML Internship Task 1</h1>

<p align="center">

<img src="https://img.shields.io/badge/Python-3.10-blue?logo=python&logoColor=white" />

<img src="https://img.shields.io/badge/Pandas-Used-lightgrey?logo=pandas" />

<img src="https://img.shields.io/badge/Scikit--Learn-Used-orange?logo=scikitlearn" />

<img src="https://img.shields.io/badge/Internship-Task\_1-important" />

</p>

---

## 📌 Objective

The goal of this task is to \*\*clean and preprocess the Titanic dataset\*\* to prepare it for machine learning models.

This step is \*\*critical before model building\*\* to ensure accurate predictions and model performance.

---

## 🔧 Tools & Libraries Used

- 📘 Python

- 📊 Pandas, NumPy

- 📈 Matplotlib, Seaborn

- 🧠 Scikit-learn

---

## 🧼 Steps Performed

- ✅ Loaded the Titanic dataset

- ✅ Handled missing values (`Age`, `Embarked`)

- ✅ Dropped irrelevant column (`Cabin`)

- ✅ One-Hot Encoded categorical columns (`Sex`, `Embarked`)

- ✅ Standardized numeric features (`Age`, `Fare`)

- ✅ Visualized and removed outliers using Boxplot + IQR

---

## 🧠 Concepts Practiced

- Data Cleaning

- Feature Engineering

- One-Hot Encoding

- Feature Scaling (StandardScaler)

- Outlier Detection & Removal

---

## 📂 Folder Structure