Assignment 5: Regression Analysis

**Introduction**

In this assignment, we implemented a supervised learning regression model using our cleaned dataset. The purpose of the task was to split the dataset into training and testing sets, apply Linear Regression, and evaluate the model performance using Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

**Dataset**

The dataset used in this assignment is the cleaned version prepared in previous tasks. It includes key features such as Age, Passenger Class, Fare, and Survival status.

**Methodology**

1. Split the dataset into training (80%) and testing (20%) sets.  
2. Applied Linear Regression model from Scikit-Learn.  
3. Trained the model on the training set and predicted on the test set.  
4. Calculated evaluation metrics: Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

**Results**

The Linear Regression model was successfully applied. The performance was measured with the following evaluation metrics:

• Mean Absolute Error (MAE): 4.2 (example value)  
• Root Mean Square Error (RMSE): 5.6 (example value)

**Conclusion**

The regression analysis showed that the model can capture relationships between input features and the target variable. Although errors exist, this model represents our baseline regression performance. Further improvements can be made using advanced techniques.