

## Multiple Linear Regression

This notebook implements Multiple Linear Regression using the sklearn library. It demonstrates:

- Data loading and preprocessing
- Exploratory Data Analysis (EDA)
- Model training
- Predictions
- Performance evaluation using common regression metrics

### Key Learnings:

- Pandas & NumPy are essential for handling structured data.
- Matplotlib & Seaborn are used for data visualization.
- scikit-learn provides out-of-the-box tools for model building and evaluation.
- Pairplot visualizes relationships between numeric features.
- Heatmap shows feature correlation. Helps in selecting impactful variables for regression.
- Coefficients tell how much the target variable changes with a unit change in the feature.
- Intercept is the predicted value when all features are zero.
- Mean Absolute Error (MAE): Average of absolute errors.
- Mean Squared Error (MSE): Punishes large errors more.
- Root Mean Squared Error (RMSE): Same unit as target; easy to interpret.

### Final Thought

- Multiple Linear Regression is a good baseline for regression problems.
- Always check multicollinearity and feature importance before deploying the model.
- Preprocessing, visualization, and interpretation are key in real-world applications.