MCA 5141 – Machine Learning Lab
Week – 5

Exercise 1

Using the given IPL 2013 dataset, apply Linear Regression techniques to predict player auction prices and performance statistics.

1. Price Prediction from Batting Stats
   - Build a linear regression model to predict SOLD PRICE based on batting stats such as RUNS-S, HS (High Score), SR-B (Strike Rate), SIXERS.
   - Evaluate model accuracy using R^2 score and RMSE.
   - Visualize the predicted vs actual prices with a scatter plot.

2. Price Prediction from Bowling Stats
   - Train a linear regression model to predict SOLD PRICE using bowling features: WKTS, AVE-BL, ECON, SR-BL.
   - Compare predicted vs actual prices and check which bowling feature contributes most (via coefficients).
   - Plot residuals to check model errors.

3. Base Price vs Sold Price Relationship
   - Fit a regression model with BASE PRICE as input and SOLD PRICE as output.
   - Check accuracy using Mean Absolute Error (MAE).
   - Visualize using a regression line plot.

4. Multi-feature Model for Auction Price
   - Build a multiple linear regression model using both batting and bowling features (e.g., RUNS-S, WKTS, SR-B, ECON, AVE-BL, etc.) to predict SOLD PRICE.
   - Compare performance against single-feature models.
   - Show feature importance (coefficients bar chart).

5. Age Impact on Price
   - Train a regression model to see if AGE and performance stats (RUNS, WKTS) explain variations in SOLD PRICE.
   - Evaluate correlation between AGE and price using regression plots.
   - Visualize with heatmap of correlations.

6. Country-wise Price Prediction
   - Build regression models for different COUNTRY groups (e.g., IND vs AUS vs SA).
   - Compare model accuracies for each country.
   - Use boxplots to visualize SOLD PRICE distribution by country.

7. Prediction of Strike Rate (Batting)
   - Use player stats (RUNS-S, HS, SIXERS, AGE) to predict SR-B (Strike Rate) with linear regression.
   - Evaluate prediction accuracy.
   - Plot actual vs predicted strike rates.