

Objective:

Use the housing dataset to practice fitting, predicting, and evaluating linear regression (predicting house prices) and logistic regression (classifying if house price is above median).

Dataset:

Use the EDA housing dataset or any dataset with features related to houses (size, bedrooms, location, etc.) and a continuous target variable (house price).

Tasks:

1. Linear Regression: Predicting House Prices
 - Load the housing dataset.
 - Visualize key features vs target price.
 - Split data into train and test sets.
 - Fit a Linear Regression model to predict house prices.
 - Predict on test data and compute Mean Squared Error (MSE).
 - Plot true vs predicted house prices.
 - Interpret what the MSE says about model accuracy.
2. Logistic Regression: Classifying Expensive Houses
 - Create a new binary target variable:
 - 1 if house price is above median price, else 0.
 - Split data into train and test sets.
 - Fit a Logistic Regression model to predict this binary label.
 - Predict on test data and compute classification accuracy.
 - Visualize confusion matrix or ROC curve.
 - Discuss when logistic regression is useful compared to linear.
3. Comparison and Reflection
 - Summarize differences in outputs, tasks, and metrics of the two models.
 - Discuss real-world use cases for predicting prices vs classifying expensive houses.
 - Optionally, experiment with features and see how adding/removing features affects both models.

Deliverables:

- Jupyter notebook or Python script with commented code.
- Plots and metrics shown inline.
- Short written reflection on differences between linear and logistic regression based on the housing data.