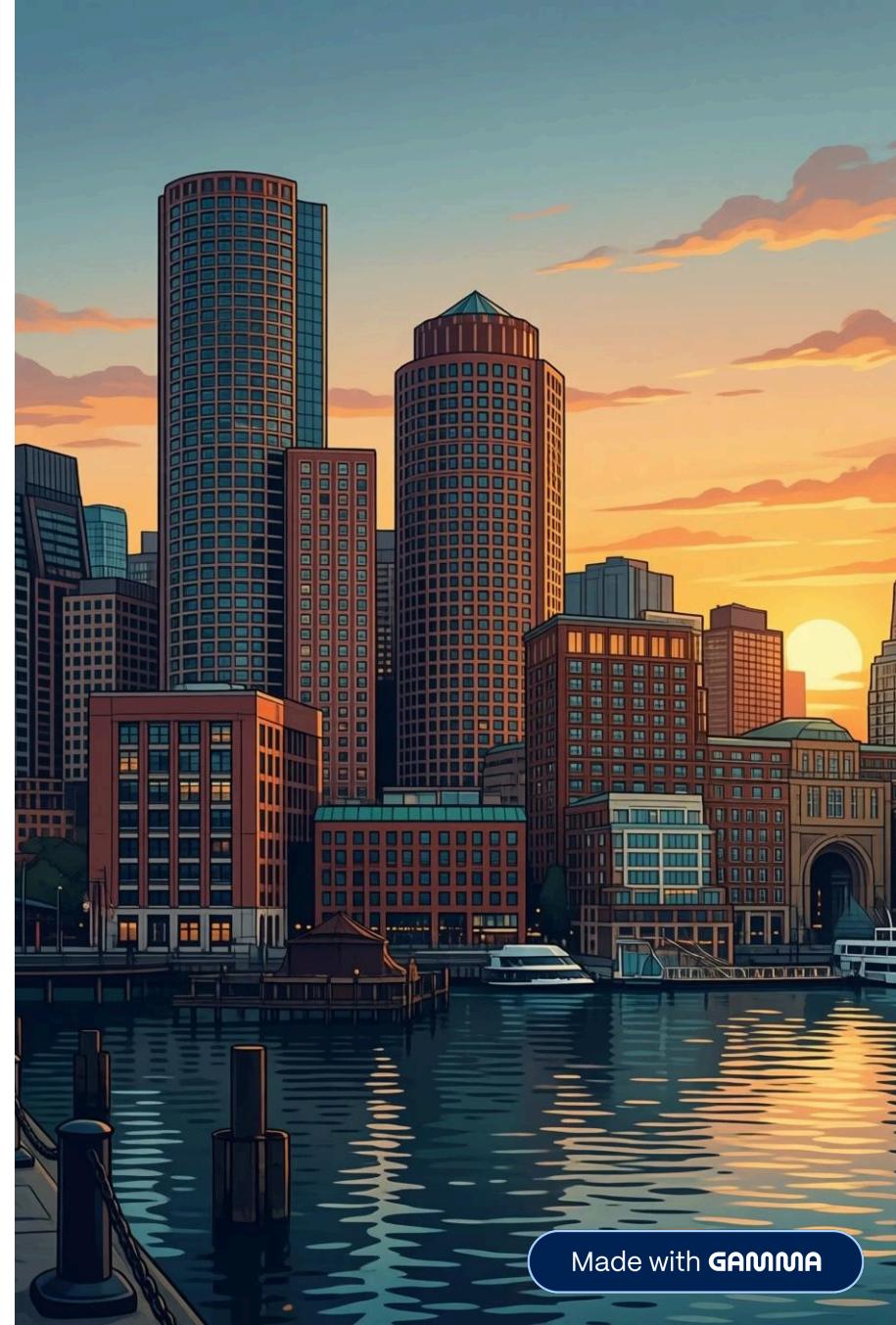


Boston Housing Price Prediction

Presented by Sharon Imaledo — Data Analyst



Executive Summary



Strong Predictive Performance

Model explains 73% of house price variation with low error.



Key Price Drivers Identified

Rooms increase price; pollution, crime, disadvantage reduce it.



Practical Use Case

Reliable price estimates for new properties based on characteristics.

Project Overview & Problem



Project Purpose

Analyze housing data to identify factors influencing property value and develop a predictive pricing model.



Problem Identified

Lack of quantitative insight into how structural, environmental, and socioeconomic variables affect house prices limits accurate property valuation and investment decisions.



Methodology

01

Data Source

Boston Housing dataset: housing, environmental, and demographic variables.

02

Data Preparation

Checked for missing values/duplicates. Log transformation for skewed variables (CRIM, ZN, LSTAT) as a result of detection of outliers.

03

Analysis Approach

Multiple linear regression model trained.

Evaluated using R^2 and RMSE.

Visualized actual vs. predicted prices.

Key Price Insights



Primary Value Driver — RM

Higher number of rooms strongly increases house price.



Socioeconomic Impact — LSTAT

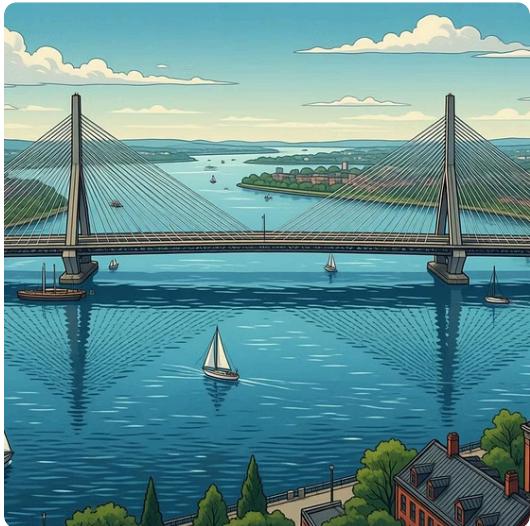
Higher socioeconomic disadvantage is associated with lower property value.



Environmental Effect — NOX

Higher pollution levels reduce house prices.

Additional Price Influencers



Location Advantage — **CHAS**

Properties near the Charles River show higher prices.



Education Proxy — **PTRATIO**

Higher pupil-teacher ratio is linked to lower house value.

Model Performance

0.73

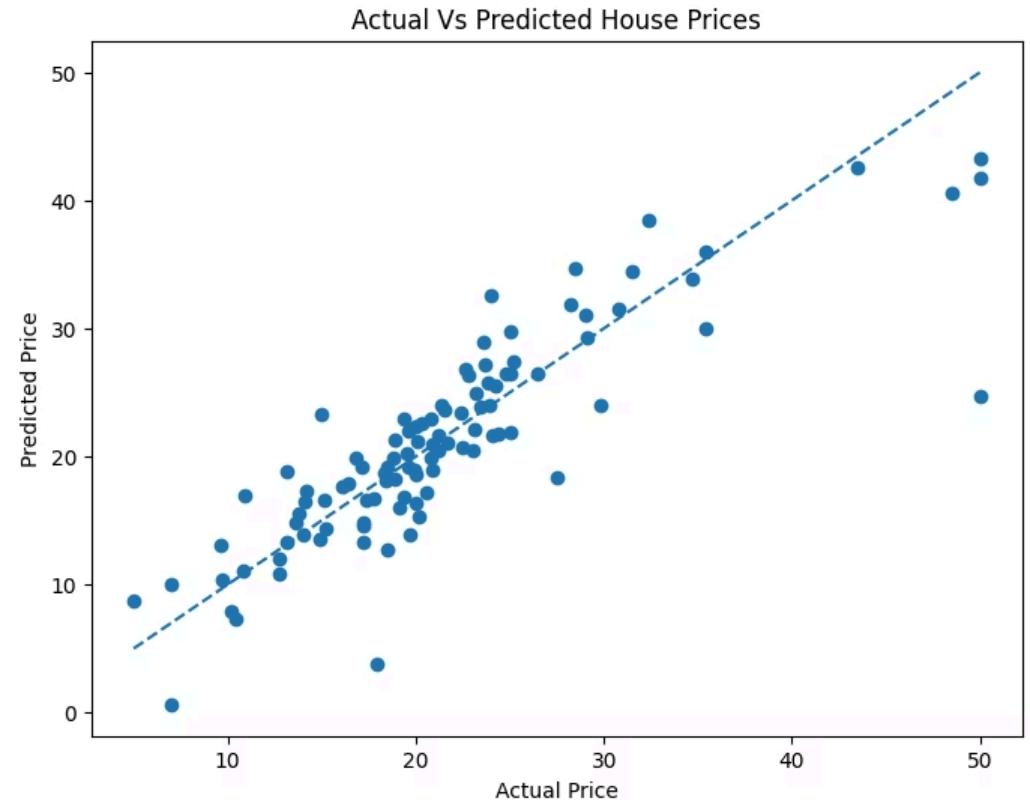
R² Value

Explains 73% of house price variation, indicating a strong relationship.

4.41

RMSE

Predictions differ from actual prices by about 4.41 units on average.



Scatter plot shows predicted values closely aligned with actual prices, demonstrating model accuracy.

Prediction Use Case



Valuing New Properties

The model estimated a house price of 11.21 for a new property based on its features, demonstrating practical application for valuation in real-time scenarios.

Strategic Recommendations

Property Valuation

Utilize the model to support data-driven pricing decisions for sales and acquisitions.

Investment Focus

Prioritize properties with more rooms and lower pollution levels for optimal returns.

Risk Assessment

Integrate crime rate and socioeconomic indicators into property value evaluations.

Location Strategy

Target properties near desirable environmental features for higher valuation potential.

Conclusion

The regression model provides accurate and interpretable house price predictions. Key structural, environmental, and socioeconomic variables significantly influence property value. This approach supports data-driven real estate decision-making and investment planning.

