Real Estate Housing Price Prediction in Boston (1970) Report

Table of Contents

Background and Business Questions

This project aims to study the different factors that affect housing prices in Boston. The data, collected in 1970, is used to understand what influences house prices in the area. Several things can impact housing prices, such as the number of rooms, location, and access to utilities.

The main goal of this research is to figure out how each of these factors affects housing prices and to build a model that can predict these prices. The key questions guiding this project are:

1. What are the top five factors that drive housing prices?

2. What are the top three factors that cause housing prices to go up?

3. What are the top three factors that cause housing prices to go down?

The aim is to gain helpful insights into how the housing market worked in Boston as at the time this data was collected (1970).

Data Structure Overview

The data used in this study comes from the 1970 census, collected by a group of researchers. The dataset has 14 variables: one response (dependent) variable and 13 predictor

(independent) variables. The response variable represents the median housing price in each town. There are 511 observations (or data points) in total.

The 13 predictor variables include:

- The crime rate per person in each town

- The percentage of land used for residential purposes

- The percentage of land used for non-retail businesses

- Proximity to the Charles River

- Air quality, measured by nitrogen oxide levels

- The average number of rooms per house

- The distance to employment centers

- An accessibility index for highways

- The pupil-teacher ratio in schools

- The percentage of Black residents in each town

- The percentage of people with lower income status

- The proportion of older buildings

- The percentage of non-retail business land

These predictor variables help explain what factors influence housing prices. The dataset originally had shortened headers, but these were standardized for clarity and consistency throughout the analysis.

Below are images showing an overview of the original dataset, the variable attributes, and the cleaned version:
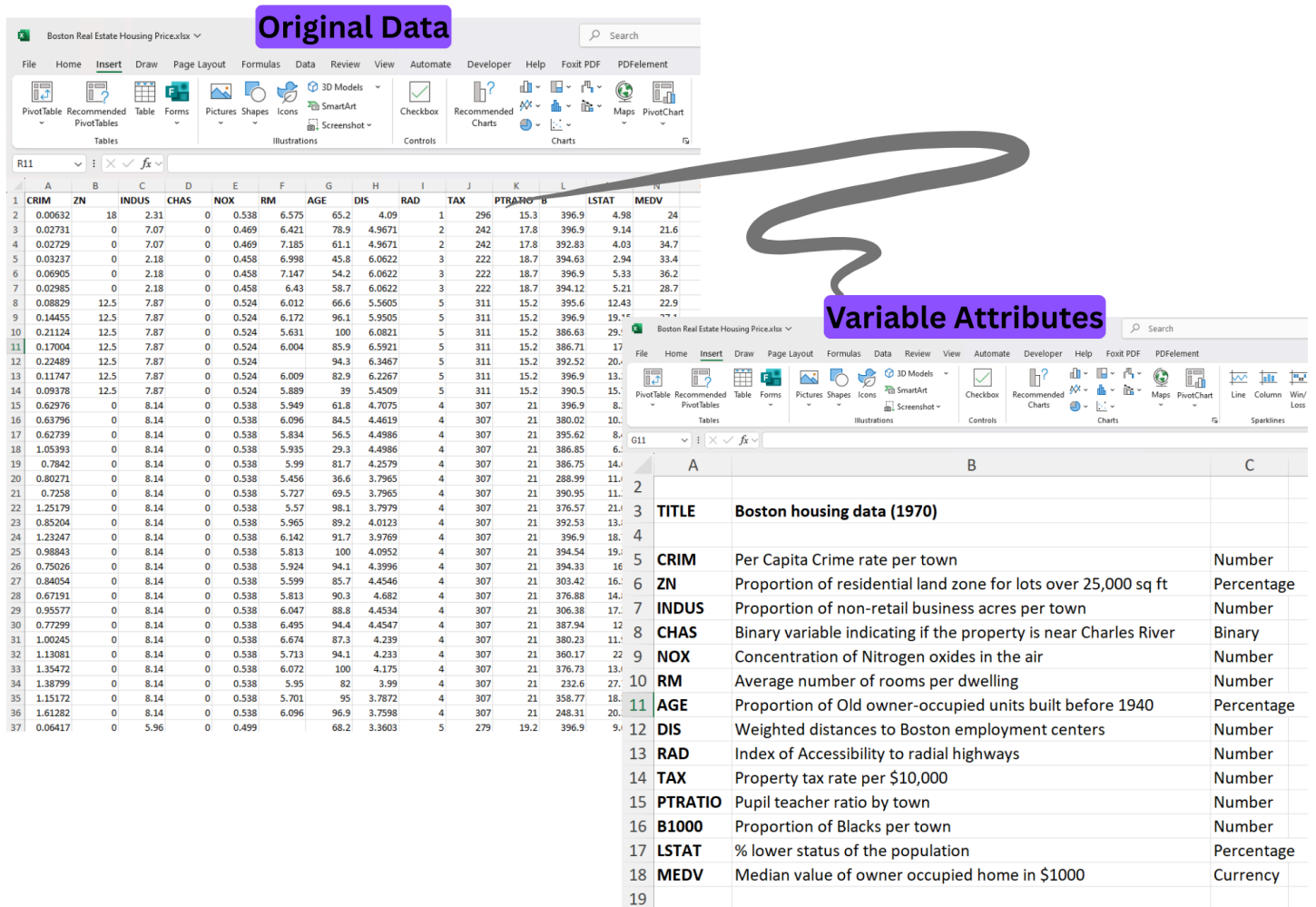

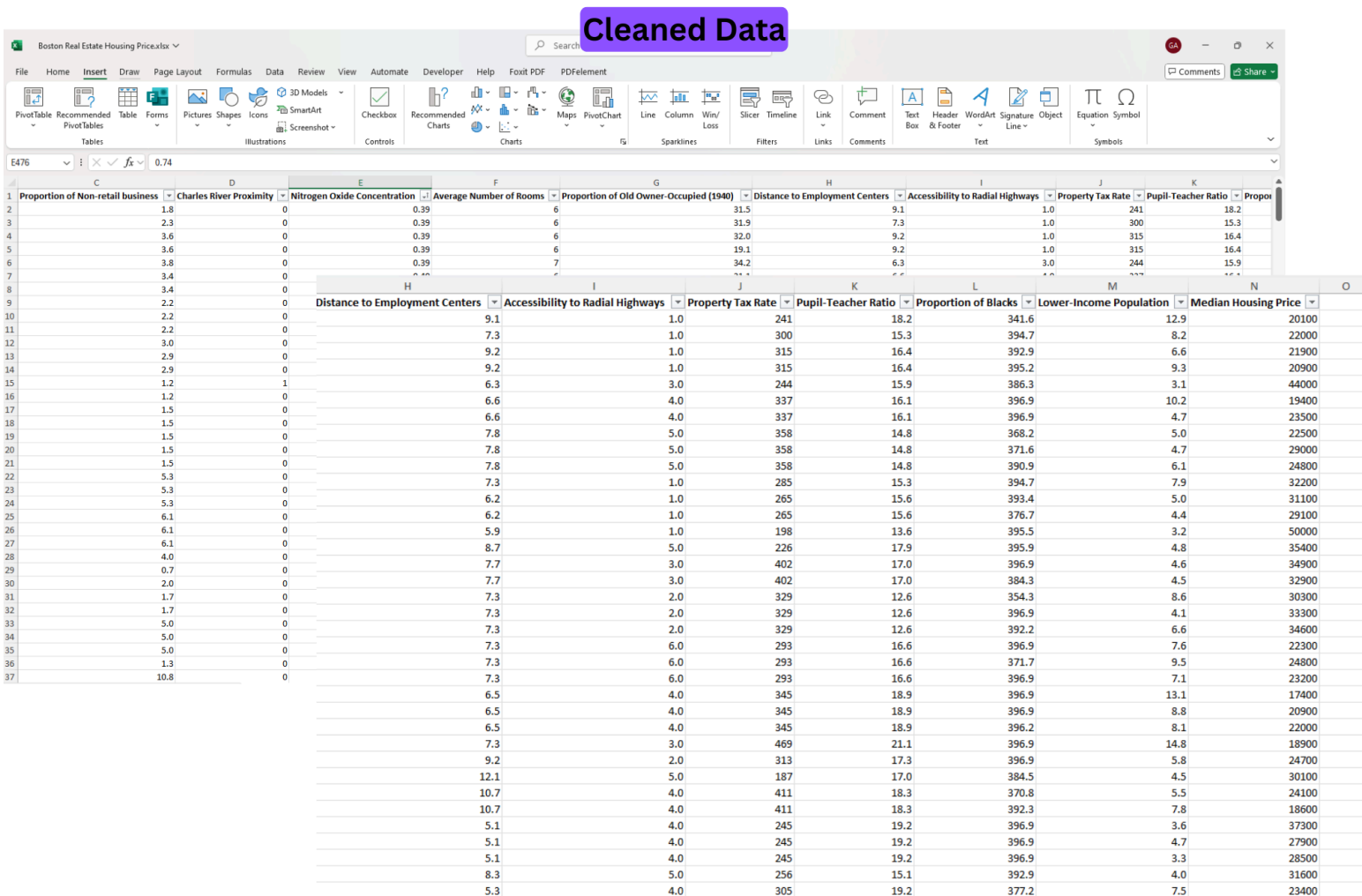Figure 1: The Original Dataset and the variable attributes

Figure 2: Cleaned Data

Executive Summary

Using data from 511 housing units and 13 independent variables, the model explains approximately 62.7% of the variation in housing prices. This is indicating a strong fit for such a complex market.

Key Drivers of Housing Prices:

Size Matters: The average number of rooms is the most influential variable. Homes with more rooms command significantly higher prices.

Location & Accessibility: Proximity to highways and employment centers plays a vital role in boosting housing value. Shorter commutes and better road access increase desirability.

Environmental Quality: Areas with lower air pollution and properties near the Charles River tend to have higher prices, showing the importance of clean and scenic environments.

Public Services Impact Value: High property taxes, poor school quality (measured by pupil-teacher ratio), and high crime rates negatively affect home values.

Age and Condition: A higher percentage of older homes tends to reduce property prices, possibly due to outdated infrastructure or maintenance costs.

Implications for Stakeholders:
- Real Estate Developers should focus on building spacious homes in well-connected and environmentally clean areas.

- Urban Planners should prioritize infrastructure, environmental health, and public safety to increase neighborhood appeal.

- Policy Makers can consider tax reforms and investments in education to indirectly raise property values.

- Investors should target locations near rivers, with low crime and short commutes, for better returns.

Overall, the analysis reveals that both structural features and location-based factors significantly shape housing prices in Boston. These insights are valuable for anyone involved in urban development, investment planning, or housing policy.
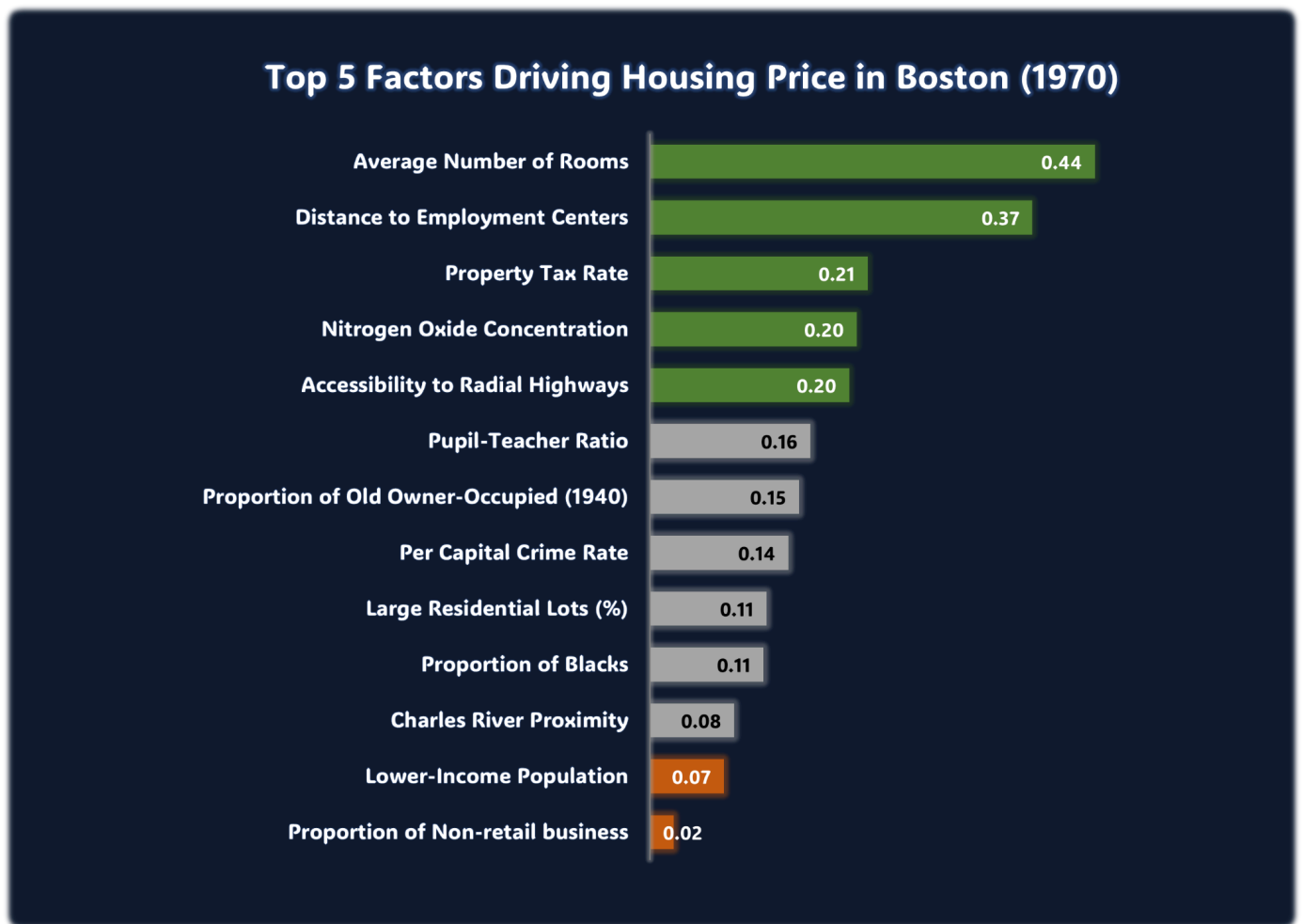
Figure 3: The Top 5 Factors Driving Housing Price in Boston (1970)

Further analysis was done to find out what factors lead to an increase in housing prices. The results showed that the number of rooms in a house has the biggest positive impact on price. This is followed by how easy it is to access highways and the size of the residential lot. These are important points for property developers to keep in mind when planning high-value projects. On the other hand, housing prices tend to drop when homes are located far from employment centers, when property taxes are high, and when the air quality is poor (see Figure 4).
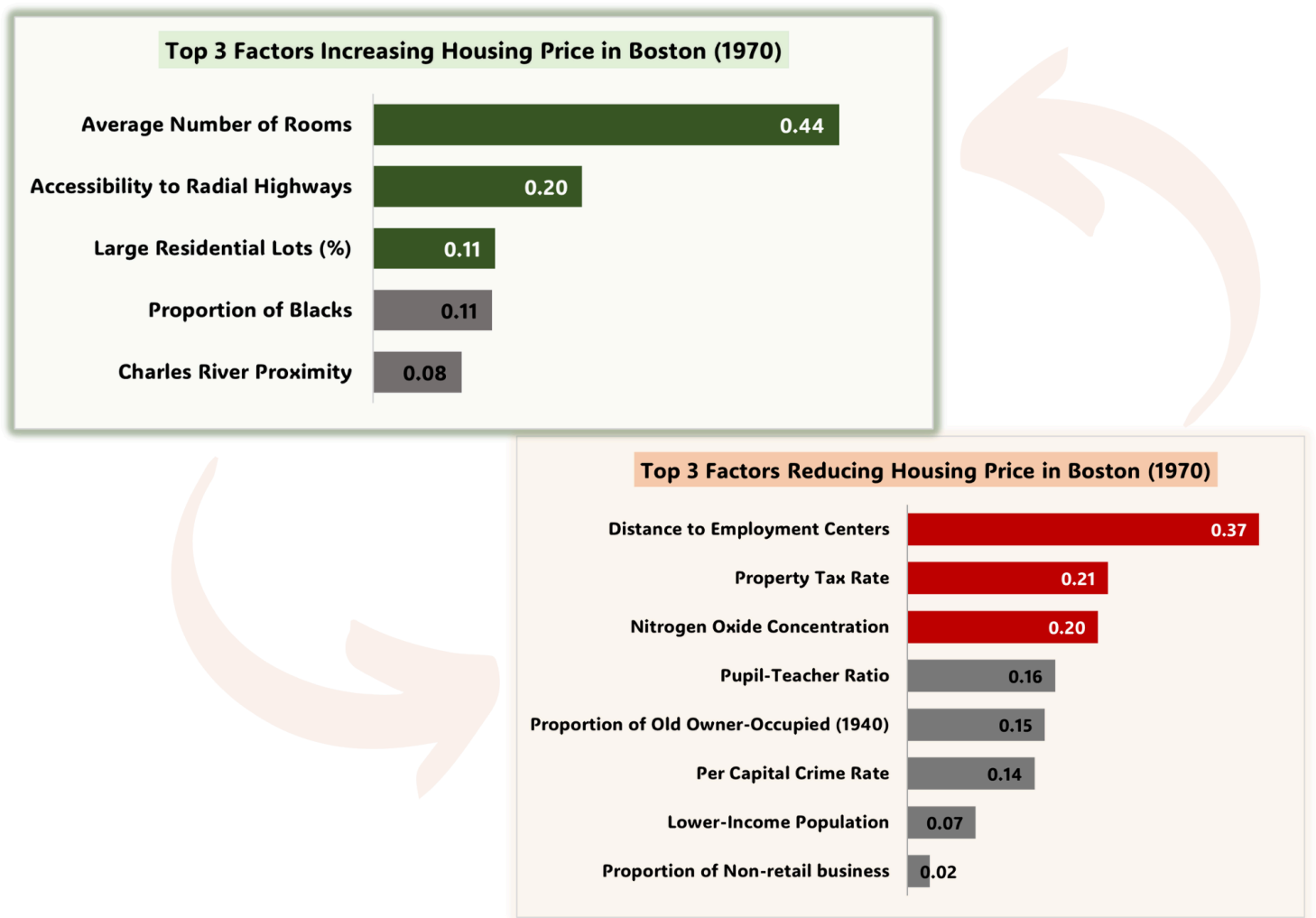
Top 3 Factors Increasing Housing Price in Boston (1970)

Factor	Value
Average Number of Rooms	0.44
Accessibility to Radial Highways	0.20
Large Residential Lots (%)	0.11
Proportion of Blacks	0.11
Charles River Proximity	0.08

Top 3 Factors Reducing Housing Price in Boston (1970)

Factor	Value
Distance to Employment Centers	0.37
Property Tax Rate	0.21
Nitrogen Oxide Concentration	0.20
Pupil-Teacher Ratio	0.16
Proportion of Old Owner-Occupied (1940)	0.15
Per Capital Crime Rate	0.14
Lower-Income Population	0.07
Proportion of Non-retail business	0.02

Figure 4: Top 3 Factors Increasing and Reducing Housing Price in Boston

Analytics Tools and Techniques

The analytics process was completed using Microsoft Excel. Multiple Linear Regression analysis was used to build the predictive model. A simple linear regression was previously used to detect the impact of each predictor variable before the MLR was conducted.

Insights Deep Dive

The MLR model (see Figure 5) revealed several key insights that have strong implications for urban planning, property development, and investment decision-making.

A7 fx Standard Error

	A	B	C	D	E	F
1	SUMMARY OUTPUT					
2						
3	*Regression Statistics*					
4	Multiple R	0.79172291				
5	R Square	0.62682517				
6	Adjusted R Square	0.61696486	**62%**			
7	Standard Error	5892.2337				
8	Observations	506				
9						
10	ANOVA					
11		*df*	*SS*	*MS*	*F*	*Significance F*
12	Regression	13	28691887186	2207068245	63.5705303	2.00794E-96
13	Residual	492	17081461667	34718418		
14	Total	505	45773348854			
15						
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	
17	Intercept	17358.4212	6056.031631	2.86630292	0.00433092	
18	Per Capital Crime Rate	-150.1776	40.65660616	-3.69380561	0.00024569	
19	Large Residential Lots (%)	46.6567157	17.06476695	2.73409627	0.0064807	
20	Proportion of Non-retail business	-32.803901	76.68829071	-0.42775632	0.66901571	
21	Charles River Proximity	3094.17074	1069.470082	2.89318121	0.00398273	
22	Nitrogen Oxide Concentration	-16710.0924	4746.437781	-3.5205544	0.00047069	
23	Average Number of Rooms	5898.33273	479.625324	12.2977925	1.7047E-30	
24	Proportion of Old Owner-Occupied (1940)	-49.4669386	15.68959643	-3.15284965	0.00171558	
25	Distance to Employment Centers	-1702.66993	250.1462699	-6.80669726	2.9153E-11	
26	Accessibility to Radial Highways	213.582169	82.01911868	2.60405345	0.00949133	
27	Property Tax Rate	-12.0686384	4.68000018	-2.5787688	0.01020472	
28	Pupil-Teacher Ratio	-679.900704	158.3424454	-4.2938626	2.1161E-05	
29	Proportion of Blacks	11.6191611	3.327650306	3.49170137	0.00052319	
30	Lower-Income Population	-88.308464	49.7615865	-1.7746312	0.07657708	
31						

Figure 5: MLR Summary Output

1. Size and Accessibility Are Prime Drivers of Value

The number of rooms in a house was the strongest single predictor of housing prices, contributing nearly half of the overall explained variance. This underscores a fundamental buyer preference: larger living spaces command higher prices. Additionally, proximity to highways and employment centers significantly affect pricing. Homes closer to major roads and work hubs enjoy a price premium, likely due to the convenience they offer. This means location and accessibility remain critical levers in property valuation.

2. Environmental Quality Matters More Than Expected

Air pollution measured via nitric oxides concentration (NOx) had a highly negative effect on housing prices. This shows that buyers are willing to pay less for homes in polluted areas,

likely due to concerns over health and long-term livability. Conversely, homes near the Charles River attracted a price premium, indicating that natural and scenic features enhance perceived value.

3. Taxes and Commute Time Are Deal Breakers
High property tax rates and longer distances from employment centers both had strong negative relationships with housing price. These factors make homeownership less affordable and less attractive, especially for working-class buyers. Therefore, tax incentives or infrastructure improvements that reduce commute times could stimulate property demand in lower-priced neighborhoods.

4. Indicators of Neighborhood Quality Influence Price
The pupil-teacher ratio (an indicator of school quality) was found to negatively affect housing prices. This reaffirms that parents and family-oriented buyers value education infrastructure, and poor school quality can drive prices down. Similarly, higher crime rates sharply reduce property value, emphasizing the need for public safety investments in urban areas.

5. Some Assumptions May Not Hold
Interestingly, variables like the proportion of non-retail business and lower-income population were not statistically significant. This could mean that economic class and commercial zoning alone may not determine housing prices, or that their effects are more nuanced and possibly masked by multicollinearity.

Key Takeaways

Bigger homes near highways and job centers are highly valued.

Clean air and scenic locations significantly boost home prices.

Pollution, high taxes, and long commutes decrease housing desirability.

School quality and neighborhood safety are critical for maintaining property value.

Developers and planners should prioritize infrastructure, education, and environment to raise housing market performance.

Recommendations

For Real Estate Developers
- Prioritize Spacious Designs

Focus on constructing residential units with more rooms, as this is the strongest driver of housing prices.

- Choose Locations with Strong Accessibility

Select sites with good proximity to major highways and employment centers to attract buyers who value short commutes and convenience.

- Incorporate Environmental Amenities

Projects near clean natural features like rivers or green spaces should be prioritized or marketed as premium offerings.

For Urban Planners and City Officials
- Invest in Infrastructure Connectivity

Enhance road networks and reduce commuting times through efficient transport planning to boost property values across neighborhoods.
- Improve Environmental Conditions

Enforce stricter pollution controls and promote green living initiatives, especially in residential zones.
- Enhance Education Quality

Reduce pupil-teacher ratios through better funding and policy support for schools—school quality has a measurable impact on local property values.

- Promote Safety

Continue efforts in crime prevention and community policing. Safer neighborhoods see significantly higher home prices.

For Policymakers
- Review and Adjust Property Tax Policies

High property tax rates negatively impact home prices. Consider balanced tax regimes to maintain housing affordability without discouraging investment.

- Support Urban Renewal of Older Housing Stocks

Incentivize the renovation or redevelopment of older properties, especially in areas with a high percentage of pre-1940 homes.

For Real Estate Investors

- Target High-Value Zones

Focus on properties with multiple rooms, close to highways or rivers, and in low-crime areas with good schools.

- Avoid Overburdened Areas

Steer clear of locations with poor air quality, long commute distances, or high tax burdens unless there's clear long-term upside or redevelopment potential.