US Housing Market Analysis: Feature Impact Summary

Introduction

This document provides a detailed analysis of how various economic and socio-economic factors influence the housing prices in the United States. The analysis is based on a comprehensive dataset spanning the years 2002 to 2023. Three machine learning models, namely Linear Regression, Random Forest, and XGBoost, were employed to understand and predict the impact of these factors on the proxy Home Price Index.

Selected Factors

1. **Unemployment Rate (2002-2023):**
    - **Analysis:** The unemployment rate is a crucial economic indicator. During periods of high unemployment, the demand for homes may decrease as individuals are less likely to invest in real estate.
    - **Model Impact:** The models highlight a negative correlation between the unemployment rate and home prices, emphasizing the importance of economic stability.
2. **Per Capita GDP (2003-2023):**
    - **Analysis:** Higher per capita GDP generally indicates a prosperous economy, contributing positively to the real estate market.
    - **Model Impact:** All models underscored the positive impact of higher per capita GDP on home prices.
3. **Household Income (2002-2023):**
    - **Analysis:** A rise in household income is expected to drive housing demand as individuals are better positioned to invest in homeownership.
    - **Model Impact:** Models consistently indicated a positive relationship between household income and home prices.
4. **New Constructed Units (2002-2023):**
    - **Analysis:** The number of new constructed units directly affects the supply-demand balance, influencing home prices.
    - **Model Impact:** Models indicated a positive relationship, with increased construction positively impacting home prices.
5. **Working Age Population (2022-2023):**
    - **Analysis:** A growing working-age population often correlates with increased housing demand.

- **Model Impact:** The models highlighted a positive correlation, emphasizing the role of demographic factors in shaping the housing market.

6. **Consumer Price Index (2002-2023):**
    - **Analysis:** The Consumer Price Index reflects inflation and cost of living changes, impacting affordability and housing demand.
    - **Model Impact:** Models highlighted a nuanced impact, suggesting that while some inflation may boost home prices, excessively high inflation could be detrimental.
7. **Monthly Supply of New Houses (2002-2023):**
    - **Analysis:** The supply-demand balance is crucial in determining home prices, and the monthly supply of new houses provides insights into market dynamics.
    - **Model Impact:** Models consistently indicated a negative relationship, emphasizing the importance of supply constraints in driving prices.
8. **Building Permits (2002-2023):**
    - **Analysis:** Building permits signify future construction activity, impacting future housing supply.
    - **Model Impact:** Positive correlations were observed, indicating that a higher number of building permits tend to positively influence home prices.
9. **Personal Consumption Expenditures Price Index (2002-2023):**
    - **Analysis:** The PCE Price Index reflects overall inflation and consumption patterns, influencing housing affordability.
    - **Model Impact:** Models indicated a nuanced impact, emphasizing the need for a balanced inflation rate for sustainable home price growth.
10. **Delinquency Rate on Credit Card Loans (2002-2023):**

- **Analysis:** Credit market conditions influence borrowing capacity and, consequently, the ability to invest in real estate.
- **Model Impact:** Models highlighted a negative correlation, indicating that higher delinquency rates may negatively impact home prices.

11. **Rental Income of Persons with Capital Consumption Adjustment (2002-2023):**

- **Analysis:** Rental income trends provide insights into the investment attractiveness of real estate.
- **Model Impact:** The models consistently indicated a positive relationship, emphasizing the role of rental income in shaping home prices.

12. **Employment by Economic Activity - Construction (2002-2023):**

- **Analysis:** Employment in the construction sector is a key driver of economic activity and housing supply.

- **Model Impact:** Positive correlations were observed, underlining the importance of a thriving construction sector in supporting home prices.

13. **Proxy Home Price Index (2002-2023):**

- **Analysis:** This is the dependent variable representing the overall trend in home prices.
- **Model Impact:** All models highlighted the proxy Home Price Index as the most significant factor, reflecting the inherent relationship between historical prices and future trends.

Model Analysis

Model Selection: Random Forest

After evaluating three models, the Random Forest model emerged as the most robust choice, displaying high accuracy scores, particularly in terms of cross-validation and adjusted R2.

Feature Importance Analysis using SHAP

- The SHAP (SHapley Additive exPlanations) values were utilized to explain the importance of each feature in predicting home prices.
- Key Takeaways:
    - **Unemployment Rate:** Negative SHAP values suggest that a higher unemployment rate lowers home prices.
    - **Per Capita GDP:** Positive SHAP values affirm the positive impact of higher GDP on home prices.
    - **Household Income:** Positive SHAP values underscore the importance of rising household income in driving prices.
    - **New Constructed Units:** Positive SHAP values indicate that an increase in new construction positively influences home prices.
    - **Working Age Population:** Positive SHAP values highlight the positive correlation between a growing working-age population and home prices.
    - **Consumer Price Index:** Both positive and negative SHAP values suggest a nuanced impact of inflation on prices.
    - **Monthly Supply of New Houses:** Negative SHAP values emphasize the negative impact of an increased supply on home prices.
    - **Building Permits:** Positive SHAP values indicate that a higher number of building permits positively influences prices.
    - **Personal Consumption Expenditures Price Index:** Nuanced impact with both positive and negative SHAP values.
    - **Delinquency Rate on Credit Card Loans:** Negative SHAP values indicate that higher delinquency rates negatively impact home prices.
    - **Rental Income:** Positive SHAP values highlight the positive correlation between rental income and home prices.

- **Construction Employment:** Positive SHAP values emphasize the importance of a thriving construction sector in supporting prices.

Conclusion

The US Housing Market Analysis underscores the multifaceted nature of factors influencing home prices. The Random Forest model, supported by SHAP values, provides a comprehensive understanding of these dynamics. Key factors such as economic indicators, demographic trends, and market conditions collectively shape the housing market. Continued monitoring of these factors is essential for making informed decisions in real estate investment and policymaking.

Feel free to delve into the detailed analysis presented in the project notebooks for a more granular understanding of the relationships between each factor and home prices.

References:

All the data set are download form Federal Reserve Economic Data (FRED).

|- Unemployment_Rate_US_2002_2023

|- Per_Capita_GDP_US_2003_2023

|- New_Constructed_Units

|- Working_Age_Population_2022_2023

|- Consumer_Price_Index_2002_2023

|- Monthly_Supply_Of_New_Houses_2002_2023

|- Building_Permits_US_2002_2023

|- Personal_Consumption_Expenditures_Price_Index_2002_2023

|- Delinquency_Rate_On_Credit_Card_Loans_All_Banks_2002_2023

|- Rental_Income_Of_Persons_With_Capital_Consumption_Adjustment_In_Billions_2002_2023

|- Employment_by Economic_Activity_Construction_2002_2023

|- Proxy_Home_Price_Index_2002_2023 --> Dependent Variable