**ADTA 5410 Section 400 - Applications and Deployment of Advanced Analytics**

**Housing Market Cycles and Shocks: An Empirical Study of Seasonal and Regional  
Variations in U.S. Home Values**

**TABLE OF CONTENTS:**

1. **Abstract……………………………………………………………….1**
2. **Introduction…………………………………………………………...1**

**2.1 Problem Statement……………………………………………….2**

**2.2 Literature Review………………………………………………...2**

**2.3 Scope of the Study ……………………………………………….3**

**2.4 Objectives …………………………………………………...........3**

**3. Exploratory Data …………………………………………………......4**

**3.1 Data Collection …………………………………………………...4**

**3.2 Data Preprocessing ……………………………………………......4**

**3.3 Data Overview …………………………………………………....4**

**3.4 Key Insights from EDA ………………………………………......5**

**3.5 Observations and Limitations …………………………………...9**

**4. Methodology ………………………………………………………….9**

**4.1 Feature Engineering ……………………………………………...9**

**4.2 Machine Learning Logic in Time Series Decomposition ……...10**

**4.3 Testing for Stationarity using the Augmented Dickey-Fuller Test……………………………………………………………………...12**

**4.4 Interpretation……………………………………………………13**

**4.5 Model Development …………………………………………….13**

**4.6 Model Training and Evaluation ………………………………...15**

**4.7 Forecasting the house prices…………………………………......16**

**4.8 Visualisation of Predictive Results …………………………….19**

**4.9 Challenges and limitations ……………………………………...21**

**5. Outcome ……………………………………………………………...22**

**5.1 Model Performance ……………………………………………...22**

**5.2 Business Insights ………………………………………………...22**

**5.3 Our Recommendations ………………………………………....23**

**6. Conclusion …………………………………………………………...23**

**7. References …………………………………………………………...24**

**Abstract**

This paper is a literature that reviews the impact of the 2008 financial crisis and the COVID-19 pandemic on housing price returns in American cities and uses the time-series analysis to provide short-term future estimates. The analysis based on Zillow housing price data at the city levels 2000-2024 will entail analysis of exploratory data analysis, seasonality, and stationarity tests (ADF), and prediction modelling (ARIMA and SARIMA). The concept of seasonal decomposition is used to decompose a trend, seasonal, and residual effect, which consequently determines any structural changes as a result of the two shocks. The best forecasting method is calculated by measuring the model comparison and forecast accuracy metrics (MAE, RMSE, MAPE). The results constantly show that SARIMA shows a better result compared to ARIMA, and the prices are growing rapidly between 2020-2024, and the price variations took a repetitive pattern of seasons. The use of the SARIMA model of the 20 most valuable regions gives six-month predictions of a rise in prices in 15 regions and in five ones the rise is minor. The findings prove that the 2008 crisis caused a long-term negative price shock, and COVID-19 served as a catalyst of price growth, and its impacts differed considerably across regions. Data and modelling policy implications and limitations are also described.

**Introduction**

In the U.S. economy housing market is the most critical part which reflects both the local and national economics.  when the housing values are disrupted, it influences people's wealth, financial stability and the overall economic growth. Understanding the housing price returns provides us the valuable insights on the dynamic change of housing prices in the market and reveals patterns and trends.

The two major shocks “2008 financial crisis” and “COVID-19” are the most notable disruptions to the housing market. The 2008 financial crisis was caused by the subprime mortgage market's collapse which led to significant drops in housing prices and had a long-term impact on household wealth. The COVID-19 pandemic impacted housing demand, supply restrictions, and buyer preferences are affected as there is a shift towards remote work, resulting in unstable house price returns.

These shocks did not have the equal impact on all the regions. Housing markets in various U.S cities had different reactions to the shocks depending on local economies, population growth, employees and install housing availability. Understanding house price changes at the city-level helps clarify long term differences in reaction to the shocks, as well as the ways different cities were impacted by the economic events.

**Problem Statement:**

**“How did the 2008 Financial Crisis and the COVID-19 pandemic reshape housing price returns across U.S. regions, and what short-term forecasts can be derived from these dynamic”**

The paper examines how COVID-19 and the 2008 financial crisis have influenced the housing price returns in the various regions in the United States by using time-series forecasting to determine short-term price dynamics. The study will use the Zillow house price data of the years 2000-2024 to compare the behaviour of the housing market prior, during, and after these economic shocks in order to find out the regional disparities, long term trends as well as the resilience patterns. Through a combination of exploratory data analysis, seasonal decomposition and predictive modelling (ARIMA and SARIMA), The research aims to find out how these crises have reshaped the dynamics of the housing market and what they hold in terms of predictive future prices.

**Literature Review:**

**Impact Of Economic Shocks on Housing Price Returns:**

The impact of economic shocks on housing price returns has been an important area of research, especially with regard to the 2008 financial crisis and the COVID-19 pandemic. Research has shown how both events disrupted housing markets in different ways across various regions of the U.S. Below are some of the findings from several studies that investigate the effects of these economic shocks on housing markets.

Ilmazkuday (2023) examined the impact of COVID-19 on housing prices in the U.S. at the county level. Their research found that COVID-19 cases had a negative and significant effect on housing prices, with the impact being more pronounced in counties with higher poverty rates

Mondragon and Wieland (2022) examined the link between the shift to remote work and housing demand. According to their findings, the dynamics of housing demand changed significantly during the pandemic, and at least half of the recent aggregate increase in house prices can be attributed to the shift to remote employment.

Doerner, W. M., & Lin, W. (2023) their research highlights the seasonality in the housing market is not same across geographies . The authors found that while seasonality in house prices has increased over the past decade, updating adjustments have made only small improvements, especially for over 400 areas that did not have them before.

Petersen (2024) researched how housing values changed during the COVID-19 and determined that homes for sale were increasing in value by around 1% more each month than projected. This shows that people speculated more and that the market behaved differently during the crisis.

Duca (2010) examined the role of housing markets during the 2007–2009 financial crisis. The study highlights that the unsustainable weakening of credit standards led to U.S. mortgage lending and a housing bubble. This bubble's consumption impact was amplified by innovations that altered the collateral role of housing.

Adelino, Schoar, and Severino (2018) explored how housing and mortgage markets contributed to the financial crisis. They pointed out that inflated house-price expectations caused households of all income levels, particularly the middle class, to raise their demand for housing and mortgage debt. Banks loaned a amid rising collateral values, ignoring the risk of default.

The Great Recession and Its Aftermath: According to the Federal Reserve History, home prices fell by over a fifth on average across the nation from the first quarter of 2007 to the second quarter of 2011. This reduction in house prices aided the development of the financial crisis of 200708 in that the players in the financial markets were quite uncertain on the prevalence of the losses in the mortgage-related assets.

The 2007-2009 Housing Crisis: Not every part of the United States was equally affected by the crisis. During the period of boom (2001-2006), more gains in the prices of homes were realized in the regions that had more investors whereas during the crisis period (2007-2009) and the years that followed, there were higher losses in the prices of homes. This local exception explains the complex nature of the housing market after economic shocks.

**Scope of the Study:**

This paper explores the housing price returns at the city level in the United States based on Zillow data, specifically around the times of the 2008 financial crisis as well as the COVID-19 pandemic. The analysis by focusing on the behaviour of the markets in the run up, during, and after these two huge shocks captures the regional variation in the dynamics of the housing market and predicts the short-term price patterns in the housing market. Both the economic shocks are understood by applying time-series models, that is, ARIMA and SARIMA, to understand the cyclical trends, evaluate the market strength, and predict the future growth and decline of housing prices. **Objectives:**

* This study examines how the 2008 financial crisis and the COVID-19 pandemic changed house price returns in different U.S. cities.
* To determine how the impact and recovery differed from one region to another.
* To analyse the trends in the housing price returns in the lead-up, the time, and the aftermath of these occurrences.
* To predict the short-term changes in housing prices in the best-value areas with the most effective time-series model.

**Exploratory Data**

**Data Collection**

The data was acquired on Zillow, a real estate market application on the internet. The data of Home Value Index (ZHVI) provided by Zillow monitors the average home values as of each month in 3.073 counties in the U. S in the period between Jan 2000 and July 2025. This data can be analyzed as a time-series measure of residential housing market trends, which can be consistent, and one can analyze the data before and after a specific event like the crisis in 2008 and the COVID-19 pandemic.

**Data Preprocessing**

The data was work sheeted in Python (Google Colab Application) and metadata columns were divided with time-series columns in which the values of Zillow Home Value Index (ZHVI) per month were extracted. The data was then transformed using the melt operation where the wide data was reformatted into long where each row corresponded to a single county-month record. In preprocessing, non-numeric and missing data were either changed or eliminated to provide accurate analysis. Price column was transformed into numeric format whereby invalid or text-based values could be changed to NaN. The date column was interpreted into correct proper datetime objects. The value of percentage change in ZHVI per month in values in each region was determined as returns, and in case of any missing value in observation in each county, it was eliminated. Our report has classified economic periods into 5 groups, namely: Pre-2008 Crisis, 2008 Crisis, Post-2008, COVID-19 Shock, and Post-COVID-19. Period and State were later coded to form dummy variables in order to carry out regression analysis to provide the opportunity to analyse the changes in housing prices returns differently across time and regional markets. All these preprocessing made sure that the dataset was clean, consistent and was ready to be statistically accurate.

**Data Overview**

This dataset contains 3,073 U.S counties with the monthly Zillow Home Value Index (ZHVI) data from January 2000 through July 2025. With 2,073 rows (counties) and 361 columns with 9 being meta data and 307 being monthly ZHVI data. The meta data characteristics are Region ID, Region Name, State Name, State, Metro, Size Rank, and FIPS codes. 307 columns being ZHVI that represent typical home values in dollars for each county/month.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Monthly Home Prices | Monthly Returns |  | Monthly Home Prices | Monthly Returns |
| Mean | 224,000 | .35 | Min | 27,248 | -27.4 |
| Median | 183,762 | .33 | Max | 3,015,796 | +16.81 |
| Standard Deviation | 125,084 | . 75 | 25th Division | 110,378 | -0.06 |
|  |  |  | 75th Division | 216,276 | .70 |

Table 1: Descriptive Statistics of Monthly Home Prices and Returns

**Key Insights From EDA**

The Exploratory Data Analysis (EDA) investigates the trends, variability, and spatial dynamics in United States housing prices from 2000 to 2025, using the Zillow’s Home Value Index i.e. (ZHVI) dataset. Each observation represents the median home value of a U.S. County in a month, allowing for insight into market behaviours.

**National Trends Over Time**

This figure introduces the national average ZHVI over time displaying notable points of growths and dips like the Global Financial Crisis in 2008, COVID-19 boom, and stabilization points for each disaster.

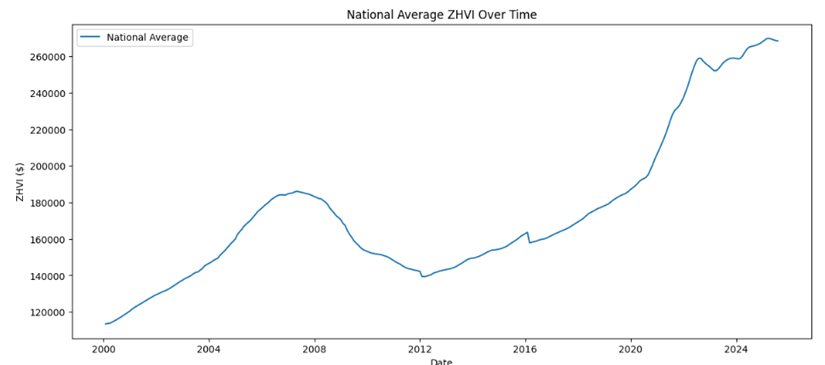


Figure 1: National Average ZHVI Over Time

**National Housing Value Distribution**

The ZHVI value distribution for July 2025 is seen in the figure below, with a pattern that is skewed to the right. According to these statistics, most counties have house values between $100,000 and $400,000, however a few high-value metropolitan areas have property prices exceeding $2 million. This disparity draws attention to how unbalanced the US housing market is.

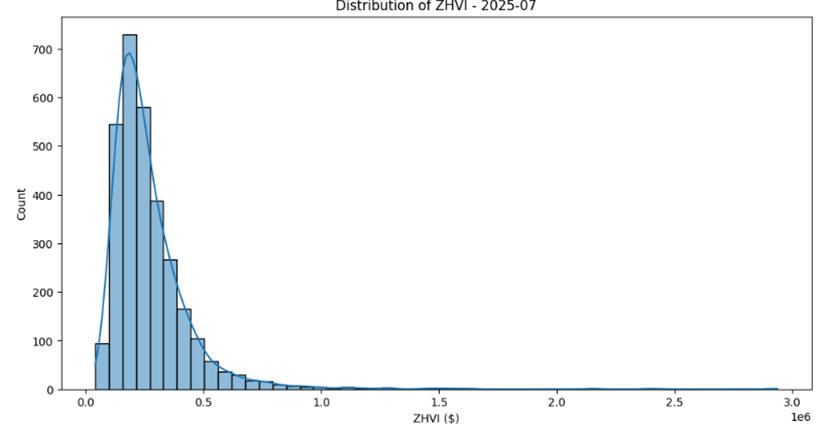


Figure 2: Distribution of ZHVI (July 2025)

**Regional and State-Level Variation**

Figure three focuses on the top 5 states by median, MA: Massachusetts, HI: Hawaii, NJ: New Jersey, DC: District of Columbia and RI: Rhode Island. We see that these states consistently are above national averages, most notably Hawaii and DC far exceeding the other states. The shaded bands which indicate the most within-state dispersion, showing mostly the increasing volatility after 2020.

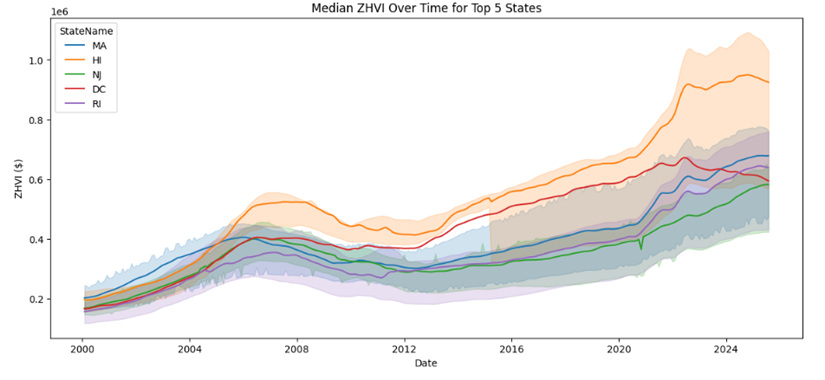


Figure 3: Median ZHVI Over Time for Top 5 States

**Year-over-Year Growth Performance**

A monetary indicator called year-over-year (YOY) development shows the percentage rise or decrease over a 12-month period by comparing a number from one period to the same period in the previous year. The YoY growth rates at the state level in July 2025 are contrasted in figure four below. After the epidemic, Florida and Washington, DC, show minor drops, while Arkansas, California, and Georgia show the highest rises. This demonstrates geography differences between chilling coastal hubs and developing Desert states

.

Figure 4: Median State-Level YoY Growth (July 2025)

**National Return Dynamics**

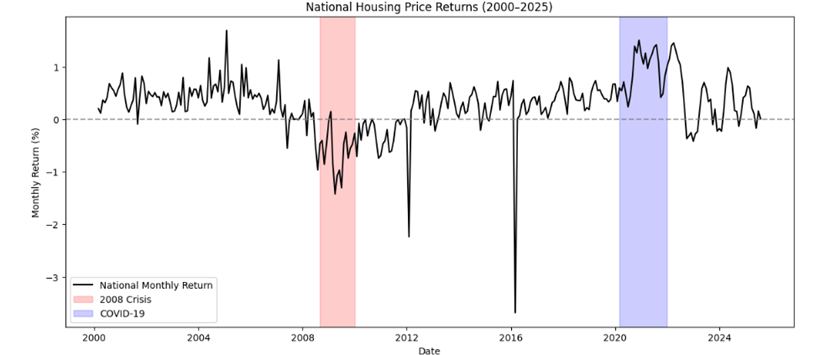
Major economic events are colored in Figure 5, which shows the monthly national housing returns from 2000 to 2025. Currency habits are impacted by crises, as seen by the stark departures to baseline tendencies displayed by the worldwide financial crisis of 2008 (pink) together with COVID-19 (purple).

Figure 5: National Housing Price Returns (2000–2025)

**Observations and Limitations**

Although the dataset is vast containing data starting from January 2000, it contains missing county-level observations due to limited Zillow coverage. Which may introduce regional inconsistencies into our report. The dataset provides a sound longitudinal perspective of housing market dynamics across economic regimes and geographies in spite of these drawbacks.

**Methodology:**

This paper focuses on the impact of the 2008 financial crisis and the COVID-19 pandemic on housing price returns in different regions of the United States using Zillow House Price data. Also, the study aims at predicting house prices in the next six months in the 20 leading cities that have the best time series model between ARIMA and SARIMA. The combination of statistical time series analysis, seasonal decomposition, and predictive modelling are used to achieve these goals.

**Feature Engineering**:

Features engineering was a key step to turn raw Zillow home cost information to meaningful predictions for time series analysis and regional impacts study. The following methods were used to get the dataset ready for decomposition, correlation analysis, and predictive forecasting.

1. Statistical and temporal trends: To ensure that the growth or decline trends of the region and alterations in the housing markets are taken into consideration, the mean every month house price indicators have been calculated.

To minimize the volatility and detect long-term trends, moving averages and moving means (3- and 6-month) were estimated.

1. Seasonal Decomposition:
2. Seasonal Decomposition:

We employed "statsmodels.tsa.seasonal import seasonal\_decompose"

1. Statsmodels: This Python library offers classes and functions for statistical data exploration, statistical testing, and estimation of numerous statistical models.
2. statsmodels.tsa - The statsmodels submodule dedicated to time series analysis. The acronym tsa represents "time series analysis."
3. statsmodels.tsa.seasonal: Another submodule under statsmodels.tsa that offers tools for interacting with the seasonality of time-series data.
4. The statsmodels.tsa.seasonal module contains a unique method, seasonal\_decompose.

A time series can be broken down into its component pieces using the seasonal\_decompose function:

1. Trend: The long-run direction of the data.
2. Seasonal: The periodic patterns concerning a period.
3. Residuals or errors - the unpredictable unthought-of variations
4. Normalization: The Zillow home prices and returns data were normalized using a z-score such that each feature was on the same scale.

Z = X - μ / σ

Stored the smoothed trend, seasonal, and residual factors in seasonal-trend decomposition (seasonal-trend-decompose) and further analysis relied on these factors as an independent analytical property.

**Machine Learning Logic in Time Series Decomposition**

Several fundamental statistical and machine learning ideas are involved in the seasonal decomposition process. Allow me to clarify the machine learning reasoning employed in your code:

Seasonal Decomposition Using Core Machine Learning Logic

1. Algorithm for Smoothing Moving Averages: Centred moving averages are used to extract the trend component.

Basis in mathematics: When calculating the moving average for a time series 𝑌ₜ, the formula is

MA(Y)ₜ = (Xₜ₋(p/2) +... + Xₜ +... + Xₜ₊(p/2)) / p

where p is the period ML concept: This eliminates high-frequency fluctuations by a type of low-pass filtering.

1. Identifying Patterns: By averaging values at a single point over several cycles, seasonal patterns are found. The mathematical method in the case of monthly data with yearly seasonality:

Average (all values for that month) = Seasonal[month] - overall average

1. Analysis of Correlation
   1. ML idea: Analysis of feature relationships. The goal is to determine which time periods exhibit comparable price patterns.
   2. Mathematics: Linear relationships are measured by the Pearson correlation coefficient which aids in identifying time series structural alterations

**Performed Seasonal Decomposition on raw data and below are the results:**

A graph of different types of graphs

AI-generated content may be incorrect.

Figure 6: Seasonal decomposition of the time series

From observed graph we can notice that

1. Pre-2008 Growth and Crisis: From 2000 to 2007, housing prices showed a robust growth pattern, peaking in 2006–2007. After that, they saw a steep and prolonged decrease, hitting a low point in 2012. Beginning in 2020, there is a clear acceleration in price growth, which lasts until 2024 at a rapid pace. In sharp contrast to the price fall observed during the 2008 crisis, this points to a major beneficial impact on house values both during and after the epidemic.

From Trend graph we can notice that

1. Impact that the 2008 Financial Crisis: The trend clearly demonstrates the observed data pattern, which consists of a strong exponential rise followed by an apparent multi-year reduction from approximately 2007 to 2012. This demonstrates that the housing market's basic long-term growth tendency underwent a significant reversal during the 2008 crisis.
2. The COVID-19 Pandemic's Effect: According to the trend, the growth curve starts to steepen significantly about 2020. It also suggests that the pandemic acted as an accelerator, initiating a new, much larger long-term growth phase, compared to the slower, slower recuperation seen between 2012 and 2020.In contrast to the COVID-19 pandemic, which caused a beneficial increase to the trend, the 2008 crisis caused a negative shock to the general trend (a notable fall).

From Seasonality graph we can notice that

The seasonality component shows a significant annual cycle and seems to be consistent across the whole period. At least at this aggregate level, the seasonal variations' range (approximately ±500 to ±750 units) seems consistent, indicating that neither the COVID-19 pandemic nor the 2008 crisis significantly changed the timing or amplitude of the annual cyclical pattern of price movements.

From Residual graph we can see that

The residuals which exhibit some enhanced volatility during the 2008 crisis era, especially a deep trough over the 2008–2012 period, but no very big spikes. There is a noticeable period of markedly elevated volatility and big positive and negative spikes around the COVID-19 pandemic (post-2020). This implies that all the short-term and irregular fluctuations in prices during this time were difficult for the model's predicted trend and seasonality to account for.

**Testing for Stationarity using the Augmented Dickey-Fuller Test**

The ADF test is a technique that helps determine whether a time is stationary, that is defined as having no trend or periodicity with a variance and means that stay constant throughout space.

H₀: The time series is non-stationary (a unit root).

H₁: There is no movement in the time series.

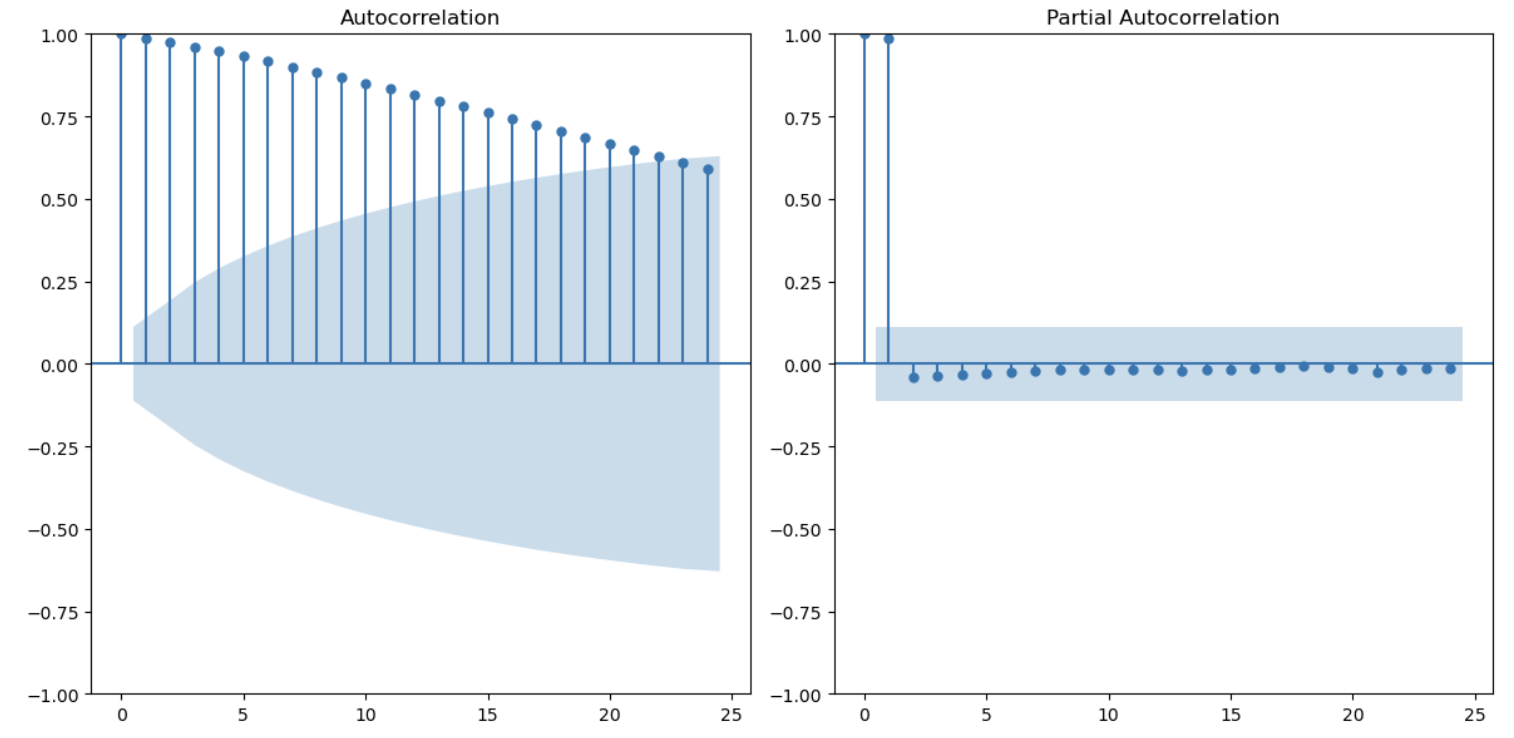


Figure 7: ACF and PACF plots of the time series.

ADF Statistic: -0.9734201783669494  
p-value: 0.7628348332420377  
Critical Values:  
 1%: -3.452789844280995  
 5%: -2.871421512222641  
 10%: -2.5720351510944512

**Interpretation**

The ADF statistic (-0.97) is larger when compared to all critical levels (e.g., -2.87 at 5%).

The null hypothesis cannot be rejected since the p-value (0.7628) is significantly higher than 0.05.

Conclusion: There is non-stationarity in the time series.

We go with Seasonal decomposition after differencing and decide on best Forecasting Model.

**Model Development**

The prediction model was developed iteratively, moving from more basic frameworks to more complicated designs that were specifically designed to account for the complexities of Zillow home markets.

I compared ARIMA and SARIMA, using average monthly Zillow home prices by all regions to determine what forecasting model performs best to predict future housing prices. This consolidation gives a country trend base level. Time-ordered data were used to train and test both models and assess their predictive accuracy. After identifying the improved model, it was implemented in each high value region to predict the price dynamics and evaluate the resilience of the region after the crisis.

1. ARIMA (Autoregressive Integrated Moving Average)

The ARIMA model was applied as the initial advanced statistical forecasting model. It was selected because it is strong in non-seasonal time-series modelling of data in which the future values would be determined as a linear function of the past and residual values.

A graph showing the growth of a house

AI-generated content may be incorrect.

Figure 8: ARIMA House Price Forecast-US

The ARIMA model was able to reproduce the long-term positive trend of house prices in the U.S. Nonetheless, it had a less jagged forecast curve that had done poorly during times of high change, like the steep increase seen after 2020.

1. SARIMA (Seasonal ARIMA) Model:

This model was created to enhance the forecasting power of the ARIMA model by directly adding seasonality. And is used to predict housing markets which tend to be cyclical (e.g., prices are high in some months because of demand spikes), and SARIMA is effective at capturing this seasonality as well as trend and noise.

The model parameters consisted of non-seasonal components (p, d, q) and seasonal components (P, D, Q, s), where s = 12, which are monthly data with annual cycles.

A graph showing a line going up

AI-generated content may be incorrect.

Figure 9: SARIMA House Price Forecast - US

The red SARIMA Forecast line predicts the gradual, steady, increasing home values, and it continues to increase in a relatively stable, controlled manner.

Throughout the duration of time (until the end of the chart, 2025) the orange "Test" line appears to reflect the actual historical factors. The observed price (Test/Orange) is much more than that which the model predicted (Forecast/Red) end of period.

**Model Training and Evaluation**

1. Data Split

* The data was separated into 70% training, 15% validation and 15% testing. subsets so that it is properly evaluated. Learning was performed on the training data, hyperparameter was guided with the validation set. Generalization testing and the test set measured optimization.

1. Evaluation Metrics:

* RMSE (Root Mean Square Error) which is a measure of the type size of prediction errors in which the large errors getting more weight. To evaluate the generality of ARIMA and SARIMA models it is helpful to evaluate their overall accuracy. The lower the RMSE, the higher is the stability and performance of the model.
* MAE is Mean absolute error which is average absolute difference between actual and predicted home prices. A lower with the MAE in which it indicates the model more reliably represents monthly price changes over time and across geographical boundaries.
* Mean Absolute Percentage Error (MAPE) – It gives the mean forecast error as a proportion of real price. In the Zillow house data, Here MAPE aids in the evaluating the model's performance in relation to regional price levels.

1. Baseline comparison:

SARIMA model outperformed ARIMA in all measures. It followed the 2020 price explosion in the test set closely and recorded periodic seasonal patterns. The red forecast line on the plotted graph tracks the actual (orange) trend more closely, which shows that SARIMA is sensitive to the complex temporal structure of housing price data.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Metric** | **ARIMA** | **SARIMA** | **Improvement (or Difference)** | **Better Model** |
| MAE | 63,912.58 | 53,087.83 | ↓∼17% | SARIMA |
| RMSE | 70,041.12 | 57,499.26 | ↓∼18% | SARIMA |
| MAPE | 20.00% | 16.70% | ↓∼3.3 percentage points | SARIMA |

Table 2: Model Performance Comparison Between ARIMA and SARIMA

Based on the results, we can say that SARIMA is the best model as errors made by SARIMA are less than those made by ARIMA with respect to MAE, RMSE, MAPE values. It implies that on average, SARIMA forecasts are more accurate.

The reduction which is in the RMSE implies that the SARIMA reduces large deviations, i.e., it is more adept at dealing with abrupt transitions or seasonality.

We continue with SARIMA model for further forecasting analysis

**Forecasting the house prices**

Since the SARIMA model had best performances in the prediction of the housing price at the national level (in terms of smaller MAE, RMSE, and MAPE), it was time to apply the predictive model to the top 20 regions (cities) in the U.S. (which represented the highest average price of houses).

The aim was to project the subsequent six months of price dynamics of these regions and deliver region-specific insights on the market trends to come.

First, we found the Top 20 regions based on highest price and then perform the model forecast for those regions.

A screenshot of a computer screen

AI-generated content may be incorrect.

A screenshot of a computer screen

AI-generated content may be incorrect.

A screenshot of a table

AI-generated content may be incorrect.

Figure 10: SARIMA Forecast Results for Top 20 U.S. Regions (Oct 2025 – Mar 2026)

SARIMA Forecast Results (Six months - Top 20 Regions)

A six-month future housing price predictions computed by the SARIMA (Seasonal Autoregressive Integrated Moving Average) model on the top 20 most valuable areas (Clyde Hill (WA), Sullivans Island (SC), and Monte Sereno (CA), etc.) are displayed in the table.

In both series, the expected average value of house prices, over the years 2025 to 2026, October to March, have been shown in accordance with past trends and seasonal variations. The values are given in U.S. dollars ($) and are the estimated average prices of the model of each month.

**Visualization Of Predictive Results**

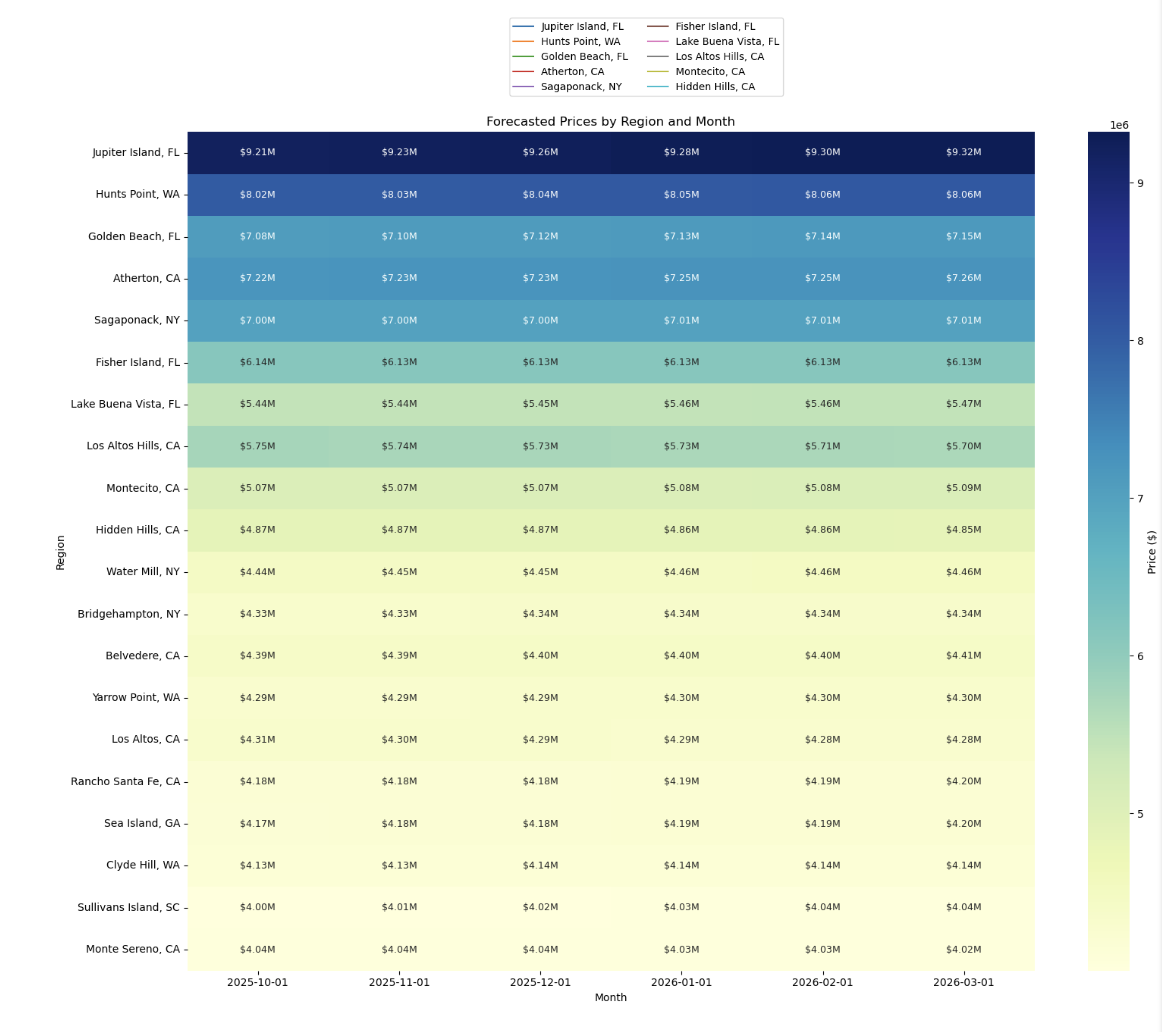


Figure 11: Forecasted Prices by Region and Month.

The above heatmap shows the next six months timeframe of median house prices starting from October 1, 2025, to March 2026.

The X-axis which shows every monthly interval (2025-10-01 to 2026-03-01)

On the Y-axis (Region) are listed geographical areas included in the prediction (e.g., Jupiter, FL; Harts Point, WA; Golden Beach, FL; Hidden Hills, CA), and in each cell is the exact projected median price in that location in that month, expressed in millions of dollars (e.g., $9.21M, $4.01M).

| **Region** | **First Month ($)** | **Last Month ($)** | **Change ($)** | **% Change** |
| --- | --- | --- | --- | --- |
| Jupiter Island, FL | 9,213,416 | 9,321,989 | 108,573 | 1.18% |
| Sullivans Island, SC | 4,001,119 | 4,041,930 | 40,811 | 1.02% |
| Golden Beach, FL | 7,084,941 | 7,154,088 | 69,147 | 0.98% |
| Sea Island, GA | 4,166,674 | 4,198,593 | 31,918 | 0.77% |
| Lake Buena Vista, FL | 5,437,636 | 5,469,091 | 31,454 | 0.58% |
| Water Mill, NY | 4,439,224 | 4,464,189 | 24,966 | 0.56% |
| Atherton, CA | 7,217,228 | 7,257,565 | 40,338 | 0.56% |
| Hunts Point, WA | 8,019,403 | 8,060,443 | 41,040 | 0.51% |
| Montecito, CA | 5,065,082 | 5,087,842 | 22,760 | 0.45% |
| Rancho Santa Fe, CA | 4,178,343 | 4,196,558 | 18,215 | 0.44% |
| Clyde Hill, WA | 4,129,717 | 4,142,152 | 12,435 | 0.30% |
| Belvedere, CA | 4,393,349 | 4,406,033 | 12,683 | 0.29% |
| Bridgehampton, NY | 4,331,669 | 4,341,924 | 10,255 | 0.24% |
| Yarrow Point, WA | 4,288,278 | 4,297,121 | 8,843 | 0.21% |
| Sagaponack, NY | 7,002,586 | 7,013,915 | 11,329 | 0.16% |
| Los Altos Hills, CA | 5,752,773 | 5,703,946 | -48,827 | -0.85% |
| Los Altos, CA | 4,310,842 | 4,276,594 | -34,248 | -0.79% |
| Monte Sereno, CA | 4,041,614 | 4,023,547 | -18,067 | -0.45% |
| Hidden Hills, CA | 4,872,734 | 4,851,581 | -21,153 | -0.43% |
| Fisher Island, FL | 6,135,467 | 6,125,274 | -10,193 | -0.17% |

Table3: Projected Six-Month Housing Price Changes for Top 20 U.S. Regions (October 2025 – March 2026)

From the predictive six months prices, we can see that:

Of the 20 regions, 15 have a price growth, and 5 have minor price decreases.

The luxury housing market remains resilient and recovering even following the recent economic shocks with most regions recording positive price growth.

**Top Performers:**

Jupiter Island, FL (+1.18%)

Sullivans Island, SC (+1.02%)

Golden Beach, FL (+0.98%)

These areas indicate robust recovery efforts following the pandemic and continued demand of high-end coastal housing, arguably because of the flexibility of remote work and migration to states with low taxes.

**Moderate Growth Regions:**

Regions such as Sea Island, GA (+0.77%), Lake Buena Vista, FL (+0.58%), and Atherton, CA (+0.56%) are experiencing consistent increases, which indicate consistent but less aggressive momentum than booms in the period around the pandemic.

Regions with Decline:

Los Altos, CA (-0.79%)

Hidden Hills, CA (-0.43%)

Fisher Island, FL (-0.17%)

Such decreases could reflect market rectifications in what were once overheated luxury markets or post-COVID-19 migration of people out of ultra-cost tech cities.

**Challenges & Limitations**

* Data Challenges
  + The dataset only concentrates on prices.It did not include critical macroeconomic variables, including interest rates, mortgage availability, unemployment rates, and housing inventory. Because of this, it is difficult to describe why/reason behind the prices differences.
  + Also, the amount of time available before and after every crisis is different. The COVID-19 event has a shorter post-recovery period and thus constrains the carrying-out of long-term validation, but the 2008 event provides additional post-crisis information in recovery research.
* Model Limitations
  + Housing markets often exhibit nonlinear dynamics due to such factors as investor attitude, policy and demographics.
  + The models are purely time-based and do not consider external shock such as mortgage restrictions, recovery plans or a reduction in interest rates. We cannot directly attribute recovery to policy actions, and this keeps us from being able to assess causality though we can see recovery.
  + The performance of SARIMA is extremely dependent on the rightly selected parameters (p, d, q, P, D, Q, s).However, optimizing these parameters usually takes a very long time while using manual tuning or a grid search option and may not be the most effective option.
* Evaluation Constraints
  + Though measures such as RMSE, MAE and R2 were effective in estimating the extent to which housing prices were accurately predictable, they were ineffective in explaining the economic or policy importance of prediction errors, including the impact on geographic housing affordability or the timing of market recovery. The analysis has largely focused on average pricing in general: this would have masked regional variations or post-crisis recovery patterns.

**Outcome**

The research we conducted offers important insights into how regional housing data and advanced time-series modelling might be utilized for predicting market trends and direct investment choices. Although the outcomes highlight both the strengths and weaknesses inherent within our methodology, they form a solid starting point intended for boosting correctness when predicting future events, as well as producing useful managerial and governmental understandings within the property domain.

**Model Performance**

Regression models were constructed using dummy variables that represented both economic eras or state regions to assess the impact of local features and economic developments on housing returns. Looking at the 5 economic periods (pre-2008, 2008 crisis, post-2008, COVID-19 shock, and post-COVID) these dummy variables allowed the model to quantify differences in average returns. In the early data exploration where it confirmed that both events caused statistically significant disruptions in housing performance. By using state dummy variables, the model was able to account for geographical differences, showing us that places such as the District of Columbia, Massachusetts, and Hawaii had larger returns than inland markets. Its explanatory power, however, was modest (R2 = 0.001) notwithstanding statistical significance, indicating that housing yields be extremely localized and influenced by supply-side, social, and economic influences that go beyond a given time or geographic area alone.

**Business Insights**

Viewing from a business standpoint, these findings highlight how drastic economic events and geographical factors come together to share the housing market’s pricing trends. This analysis confirms that the housing market are regular in nature but will react strongly to financial and health crises, with periods of elasticity once normalcy is restored. While rising markets in less expensive inland states like Wyoming, Louisville, Kentucky, and Alaska see faster rising prices due to financial migration, high-priced territory like Hawaii, Massachusetts, and the Washington, DC, District of Columbia continue to be appealing as they were having more volatility risk.

For real estate investors it means that crisis-related risk can be lowered by diversifying by region and property type. It emphasizes for legislators the value of interest-rate stabilization along with concentrated housing supply initiatives to stabilize markets with significant economic shocks.

**Our Recommendations**

* Expanding Data Sources:
  + To capture the bigger economic environment that influences the house price return, include macroeconomic information such as GDP growth, consumer mood, interest rates, unemployment rates, mortgage lending rates
* Model Refinements
  + Consider ways to better capture nonlinear and long-term correlations in housing price trends through hybrid models, which combine SARIMA with machine learning techniques (including Random Forest, XGBoost, and or LSTM networks).
  + Perform automated hyperparameter tuning (e.g. Grid Search or Bayesian Optimization) to get more precise regional predictions.
* Applications in Practice and Policy
  + To support financiers, real estate constructors, and policy researchers, develop a graphical dashboard that is interactive and provides historical, current, and projected housing value of various areas.
  + Request financial institutions and local governments to use model experience on the early warning system of market instability, risk management and affordable housing planning.

**Conclusion**

This research aimed to analyse how the 2008 Global Financial Crisis and COVID-19 have changed housing price returns across different regions at the U.S. and for this we utilized Zillow city-level data from 2000–2025. Through exploratory data analysis, seasonal decomposition and forecasting approaches (ARIMA and SARIMA), we found two crises had contrasting impacts on housing market. While the COVID-19 epidemic has caused sharp price spikes and elevated volatility, the 2008 recession set off a five-year decline and steadily declining prices. These two opposites aside, both occurrences underscored the cyclical nature and regional divergence of the U.S. housing market.

It was observed that the SARIMA models performed better than ARIMA in explaining seasonality and predicting near future trends as illustrated by lower errors including MAE, RMSE and MAPE. Projections for the 20 highest value areas were that most are projected to continue their price appreciation over the next six months, especially in coastal and high demand markets.

From a policy and business perspective, the findings suggest that recessions further exacerbate already an imbalance in the markets While some regions heal more slowly, other areas that have stronger economic underpinnings do as more quickly. thereby during immediate recovery periods, policymakers should concentrate on identifying policies to enhance housing affordability particularly stability in the markets.

**References**

* Duca, J. V., Muellbauer, J., & Murphy, A. (2010). Housing markets and the financial crisis of 2007–2009: Lessons for the future. Journal of Financial Stability, 6(4), 203–217. (https://doi.org/10.1016/j.jfs.2010.05.002)
* Adelino, M., Schoar, A., & Severino, F. (2018). The Role of Housing and Mortgage Markets in the Financial Crisis. Annual Review of Financial Economics, 10(1), 25–41. (https://doi.org/10.1146/annurev-financial-110217-023036)

# Rothstein, J. (2017). The Great Recession and Its Aftermath: What Role for Structural Changes? RSF: The Russell Sage Foundation Journal of the Social Sciences, 3(3), 22. (https://doi.org/10.7758/rsf.2017.3.3.02)

* The 2007-2009 Housing Crisis: Causes, Policy Responses and Long Term Implications (<https://www.nber.org/programs-projects/projects-and-centers/7500-2007-2009-housing-crisis-causes-policy-responses-and-long-term-implications?page=1&perPage=50>)
* ilmazkuday, Hakan (2023). “COVID-19 and housing prices: evidence from U.S. county-level data.” Review of Regional Research, vol. 43, no. 2, 2023, pp. 241-263. DOI: 10.1007/s10037-023-00187-4 (<https://link.springer.com/article/10.1007/s10037-023-00187-4>)
* Pandemic-Induced Remote Work and Rising House Prices. NBER Working Paper No. 30041, John A. Mondragon & Johannes Wieland, 2022 (<https://www.nber.org/digest/202207/pandemic-induced-remote-work-and-rising-house-prices?page=1&perPage=50>)
* Doerner, William M., and Wenzhen Lin. “Applying Seasonal Adjustments to Housing Markets.” Cityscape: A Journal of Policy Development and Research, Vol. 24, No. 3. HUD User. (<https://www.huduser.gov/portal/periodicals/cityscape/vol24num3/ch4.pdf>)
* Petersen, Alexander M. “Shift in house price estimates during COVID-19 reveals effect of crisis on collective speculation.” EPJ Data Science, Vol. 13, Article 47, 2024. (<https://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-024-00488-9>)