U.S. Residential Real Estate Investment Opportunities: A Multi-Index Framework Using Zillow Market Data

Ali Akram *

Posted Date: 23 March 2026

doi: 10.20944/preprints202603.1783.v1

Keywords: real estate investment; data analytics; Zillow Home Value Index; return on investment; housing market analysis; rental market trends; urban planning

 doi:10.20944/preprints202603.1783.v1

Article

U.S. Residential Real Estate Investment Opportunities: A Multi-Index Framework Using Zillow Market Data

Ali Akram

CliftonLarsonAllen LLP (CLA), Minneapolis, MN 55402, USA; aakramm989@gmail.com

Abstract

The U.S. residential real estate market represents a significant component of national wealth, yet investment decision-making in this sector remains largely dependent on heuristic judgment rather than systematic, data-driven analysis. This study presents a multi-dimensional analytical framework that integrates multiple Zillow housing market indices to evaluate residential real estate investment opportunities across the United States. Using publicly available data from the Zillow Home Value Index (ZHVI), Zillow Observed Rent Index (ZORI), days-on-market metrics, and regional listing data spanning Q1 2018 through Q1 2023, the research conducts six interconnected analyses: (1) identification of high-growth regions in the single-family home segment, (2) rental market trend analysis across major metropolitan areas, (3) short-term property value forecasting, (4) market liquidity assessment through days-on-market analysis, (5) return on investment (ROI) distribution for rental strategies, and (6) ROI distribution for sale strategies. The framework employs Python-based data manipulation and visualization techniques to synthesize these indices into actionable investment intelligence. Key findings indicate that the top 3% of U.S. regions exhibited disproportionately high single-family home price appreciation between December 2022 and March 2023, rental indices demonstrated a consistent upward trajectory across major metropolitan areas, and property values were projected to increase by approximately 1.6% through March 2024. Furthermore, state-level variation in days-on-market and significant regional disparities in ROI distributions for both rental and sale strategies were identified, with notable outlier regions offering substantially higher returns. The proposed framework demonstrates the practical utility of integrating publicly available housing market indices for systematic investment evaluation, offering a reproducible methodology that can inform both individual and institutional decision-making in the residential real estate sector.

Keywords: real estate investment; data analytics; Zillow Home Value Index; return on investment; housing market analysis; rental market trends; urban planning

1. Introduction

The U.S. residential real estate market constitutes one of the largest asset classes in the national economy, with total housing value estimated at over $40 trillion (Baum & Hartzell, 2021). Residential properties, particularly single-family homes and rental apartments, have historically served as vehicles for both wealth creation and income generation, attracting investors across individual and institutional categories. However, the complexity and heterogeneity of the U.S. housing market — characterized by substantial regional variation in price dynamics, rental yields, and market liquidity — present significant challenges for systematic investment evaluation.

Traditional approaches to real estate investment analysis have relied heavily on localized expertise, broker recommendations, and simplified financial metrics that fail to capture the multidimensional nature of housing market dynamics (Kołodziejczyk et al., 2019). While institutional investors have increasingly adopted quantitative methods, individual and mid-sized investors

 doi:10.20944/preprints202603.1783.v1

frequently lack access to systematic analytical frameworks that integrate multiple market indicators for comprehensive investment assessment.

The proliferation of publicly available housing data, particularly through platforms such as Zillow, has created unprecedented opportunities for data-driven real estate analysis. Zillow's research division publishes a suite of standardized indices, including the Zillow Home Value Index (ZHVI), Zillow Observed Rent Index (ZORI), and various market activity metrics, covering metropolitan and regional geographies across the United States (Holt & Borsuk, 2020). Despite the availability of these data, there remains a gap in the academic literature regarding the systematic integration of multiple Zillow indices into a coherent analytical framework for investment evaluation.

Recent research has demonstrated the utility of machine learning and statistical techniques in housing price prediction (Choy & Ho, 2023; Truong et al., 2020). However, much of this work has focused on point-price estimation rather than on the broader investment decision framework that includes rental yield analysis, market timing assessment, and liquidity evaluation. The present study addresses this gap by developing a multi-index analytical framework that synthesizes ZHVI, ZORI, days-on-market data, and ROI calculations into a comprehensive investment evaluation methodology.

1.1. Research Objectives

This study pursues five specific research objectives: (1) to identify and characterize high-growth regions in the U.S. single-family home segment based on recent price appreciation patterns; (2) to analyze rental market trends across major metropolitan areas using ZORI data to assess income-generating potential; (3) to provide short-term property value forecasts to inform investment timing decisions; (4) to evaluate market liquidity across states through days-on-market analysis for both rental and sale properties; and (5) to compute and compare ROI distributions for rental and sale investment strategies, including the identification of outlier regions with disproportionately high or low returns.

1.2. Contribution

The primary contribution of this research is the development and demonstration of a reproducible, multi-dimensional framework that integrates publicly available housing market indices for systematic investment evaluation. Unlike prior studies that have examined individual market indicators in isolation, this work synthesizes growth trends, rental dynamics, market timing signals, liquidity metrics, and dual-strategy ROI analysis into a unified analytical approach. The framework is designed to be accessible to practitioners and can be replicated using freely available data and standard analytical tools.

2. Literature Review

2.1. Real Estate as an Investment Class

Real estate has long been recognized as a distinct asset class characterized by relatively low correlation with equity markets, potential for steady income generation through rental yields, and long-term capital appreciation (Baum & Hartzell, 2021). The residential segment, in particular, benefits from fundamental demand drivers including population growth, household formation rates, and the finite supply of developable land. These characteristics have made residential real estate a cornerstone of both individual and institutional investment portfolios.

However, the performance of real estate investments is highly sensitive to geographic, economic, and temporal factors. Research has consistently demonstrated that macroeconomic variables — including interest rates, employment levels, and inflation — exert significant influence on housing market dynamics (Kołodziejczyk et al., 2019). Furthermore, localized factors such as zoning

 doi:10.20944/preprints202603.1783.v1

regulations, infrastructure development, and neighborhood amenities can create substantial variation in property values and rental yields within and across metropolitan areas (Kuhlmann, 2021).

2.2. Data-Driven Approaches in Real Estate Analysis

The emergence of large-scale housing data platforms has facilitated a shift toward quantitative approaches in real estate analysis. Zillow's suite of indices, in particular, has been widely adopted in both academic and practitioner contexts. The Zillow Home Value Index (ZHVI) provides a smoothed, seasonally adjusted measure of median home values at various geographic levels, while the Zillow Observed Rent Index (ZORI) offers analogous metrics for the rental market (Holt & Borsuk, 2020). These indices have been employed in diverse research applications, from environmental valuation studies to urban planning analyses.

Recent advances in machine learning have expanded the toolkit available for real estate analysis. Choy and Ho (2023) provide a comprehensive review of machine learning applications in real estate research, documenting the growing adoption of ensemble methods, neural networks, and gradient boosting techniques for property valuation. Truong et al. (2020) demonstrated that machine learning models can achieve superior predictive performance compared to traditional hedonic models, particularly when incorporating high-dimensional feature sets. Despite these advances, there remains limited research on the integration of multiple market indices into comprehensive investment evaluation frameworks, a gap that the present study seeks to address.

2.3. Return on Investment Analysis in Real Estate

The assessment of ROI in real estate investment involves distinct considerations for rental and sale strategies. Rental ROI depends on the relationship between acquisition costs, ongoing expenses (maintenance, property management, taxes), and rental income, while sale ROI is determined by capital appreciation minus transaction costs (Baum & Hartzell, 2021). The choice between these strategies is influenced by market conditions, investment horizon, tax considerations, and investor preferences.

Market liquidity, as reflected in days-on-market metrics, represents an additional dimension of investment evaluation that has received comparatively less attention in the quantitative literature. Properties in markets with shorter days-on-market periods offer greater flexibility and reduced holding costs, which can significantly impact realized investment returns (Bełej et al., 2020). The present study incorporates this dimension alongside traditional price and return metrics to provide a more complete assessment of investment opportunities.

3. Methodology

3.1. Data Source

The primary data source for this study is the Zillow Research Data repository (https://www.zillow.com/research/data/), which provides publicly accessible housing market indices at multiple geographic levels. Zillow's data products are derived from a proprietary database of over 100 million property records and are widely used in academic and industry research. The specific datasets utilized in this analysis include the following: the Zillow Home Value Index (ZHVI), which represents a smoothed, seasonally adjusted measure of the median estimated home value for single-family residences at the regional level; the Zillow Observed Rent Index (ZORI), which captures median observed rental prices for listed properties; days-on-market data, encompassing both "Days to Close" and "Days to Pending" metrics that measure the duration properties remain listed before completing a transaction; and regional listing data including median sale prices and inventory counts.

 doi:10.20944/preprints202603.1783.v1

3.2. Temporal Scope

The analysis spans a five-year period from Q1 2018 through Q1 2023, with growth calculations focused on the most recent three-month window (December 2022 to March 2023) for the identification of high-growth regions. This temporal scope was selected to capture both longer-term market trends and recent price dynamics, encompassing the pre-pandemic period, the pandemic-induced market disruption, and the subsequent recovery and adjustment phases.

3.3. Analytical Framework

The analytical framework comprises six interconnected analyses, each addressing a specific research objective.

Analysis 1: Regional Growth Identification. Relative price increases were calculated for each region over the three-month period from December 2022 to March 2023 using ZHVI data. Regions were ranked by growth rate, and the top 3% were identified as high-growth investment targets. This threshold was selected to focus on statistically exceptional performers while maintaining a practically useful shortlist.

Analysis 2: Rental Market Trend Analysis. ZORI data were filtered for the top 10 regions by population size to assess rental income trends in major metropolitan markets. Time-series visualization was employed to identify directional trends and assess the consistency of rental price appreciation.

Analysis 3: Property Value Forecasting. Historical ZHVI data were used to project near-term growth rates through March 2024. The forecasting approach utilized trend extrapolation from the most recent historical period to estimate percentage growth in property values.

Analysis 4: Market Liquidity Assessment. Days-on-market data were aggregated at the state level, with median values calculated for both sale and rental properties. Box plot analysis was employed to characterize the distribution of listing durations across states, providing a measure of market liquidity and turnover velocity.

Analysis 5: Rental ROI Distribution. ROI for rental strategies was calculated by relating projected rental income (derived from ZORI) to property acquisition costs (derived from ZHVI) for a future target date of March 2024. Statistical analysis including mean, standard deviation, and outlier identification (defined as values exceeding three standard deviations from the mean) was performed to characterize the ROI distribution across regions.

Analysis 6: Sale ROI Distribution. ROI for sale strategies was calculated by relating projected property value appreciation to current acquisition costs. The same statistical methodology as Analysis 5 was applied, enabling direct comparison between rental and sale strategy returns.

3.4. Tools and Technologies

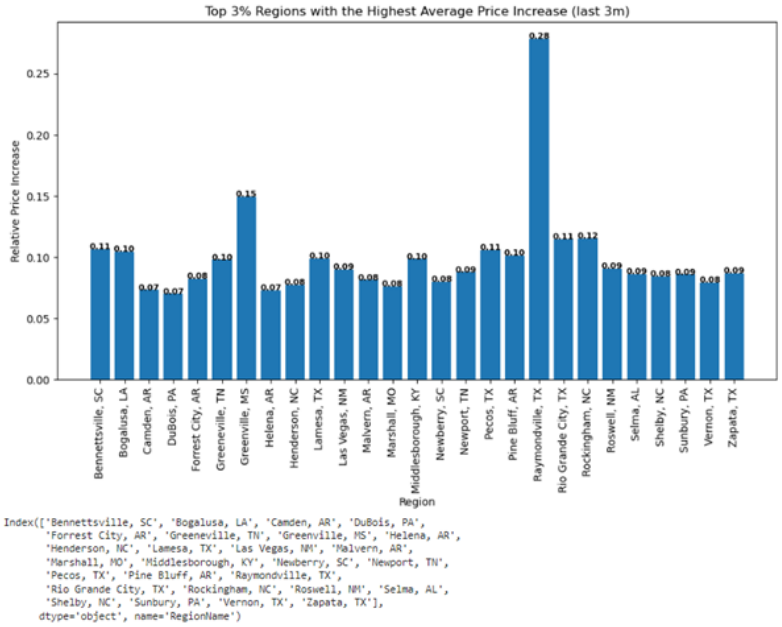
All analyses were conducted using Python 3.x with the Pandas library for data manipulation and aggregation, Matplotlib and Seaborn for data visualization, and NumPy for statistical computations. The analytical pipeline is fully reproducible using the publicly available Zillow datasets and standard Python data science libraries.

4. Results

4.1. Single-Family Home Growth Analysis

The regional growth analysis identified a concentrated set of high-performing regions within the top 3% of all U.S. regions by ZHVI price appreciation between December 2022 and March 2023. The results, visualized in Figure 1 (bar graph of highest-growth regions), revealed that the top-performing regions exhibited substantially higher average price increases compared to the national median.

 doi:10.20944/preprints202603.1783.v1



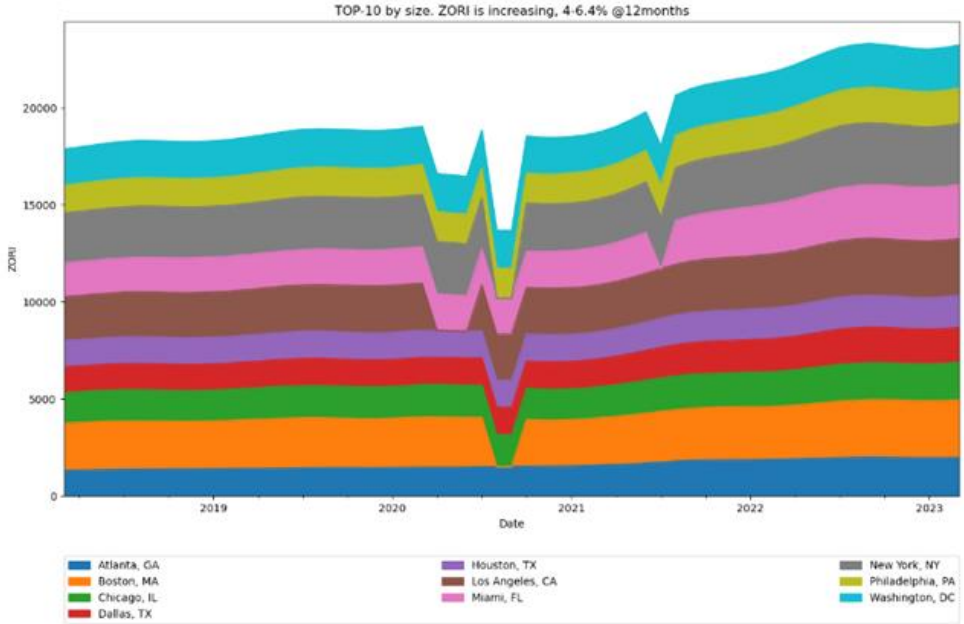

```
Index(['Bennettsville, SC', 'Bogalusa, LA', 'Camden, AR', 'DuBois, PA',
       'Forrest City, AR', 'Greeneville, TN', 'Greenville, MS', 'Helena, AR',
       'Henderson, NC', 'Lamesa, TX', 'Las Vegas, NM', 'Malvern, AR',
       'Marshall, MO', 'Middlesborough, KY', 'Newberry, SC', 'Newport, TN',
       'Pecos, TX', 'Pine Bluff, AR', 'Raymondville, TX',
       'Rio Grande City, TX', 'Rockingham, NC', 'Roswell, NM', 'Selma, AL',
       'Shelby, NC', 'Sunbury, PA', 'Vernon, TX', 'Zapata, TX'],
      dtype='object', name='RegionName')
```

These high-growth regions were geographically distributed across multiple states, suggesting that localized market conditions rather than broad regional trends were driving exceptional price appreciation. The identification of these regions provides a focused shortlist for targeted investment evaluation, reducing the search space from thousands of potential markets to a curated set of high-performance candidates.

4.2. Rental Market Trends

Analysis of ZORI data for the top 10 regions by population size revealed a consistent upward trend in rental prices over the study period. The area graph (Figure 2) demonstrated that rental values exhibited steady appreciation, with a notably accelerated increase observed during the most recent 12-month period.

```
count    10.000000
mean      5.647432
std       1.577551
min       4.011448
25%       4.408640
50%       4.902784
75%       6.915324
max       8.657860
Name: mar23_mar22, dtype: float64
```

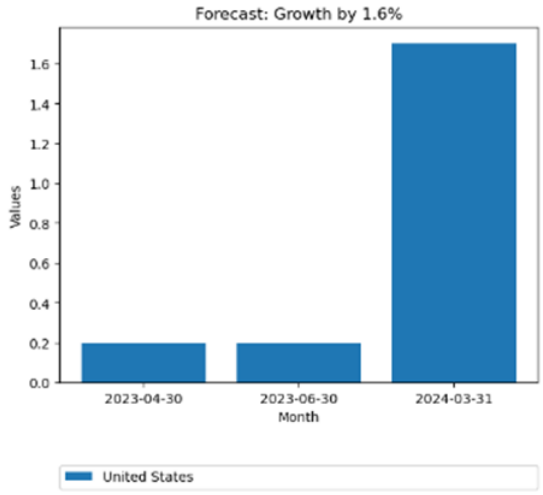


 doi:10.20944/preprints202603.1783.v1

The consistency of this upward trend across multiple major metropolitan areas suggests structural rather than transient drivers of rental price appreciation, including sustained housing demand, constrained rental supply, and demographic shifts toward renting in high-cost urban markets. For income-focused investment strategies, these markets demonstrated favorable conditions for generating and growing rental revenue streams.

4.3. Investment Timing Analysis

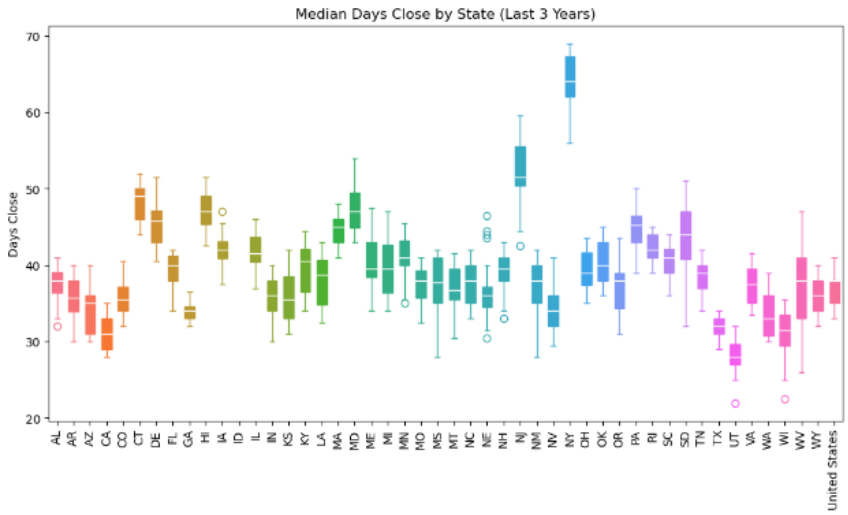
The property value forecast based on historical ZHVI trends projected an approximate 1.6% increase in home values through March 2024 (Figure 3). While modest in percentage terms, this projected appreciation represents a continuation of the positive growth trajectory and suggests that the market environment remains conducive to new investment.



The projected growth rate also serves as a baseline expectation against which regional performance can be evaluated. Regions exhibiting growth rates substantially above this national projection may represent above-average investment opportunities, while regions falling below this threshold warrant additional scrutiny regarding potential headwinds.

4.4. Days on Market Analysis

The state-level analysis of days-on-market (Figure 4) revealed considerable geographic variation in market liquidity. Box plot analysis demonstrated that the median number of days before property transactions close varies substantially across states, with some markets exhibiting notably faster turnover than others.

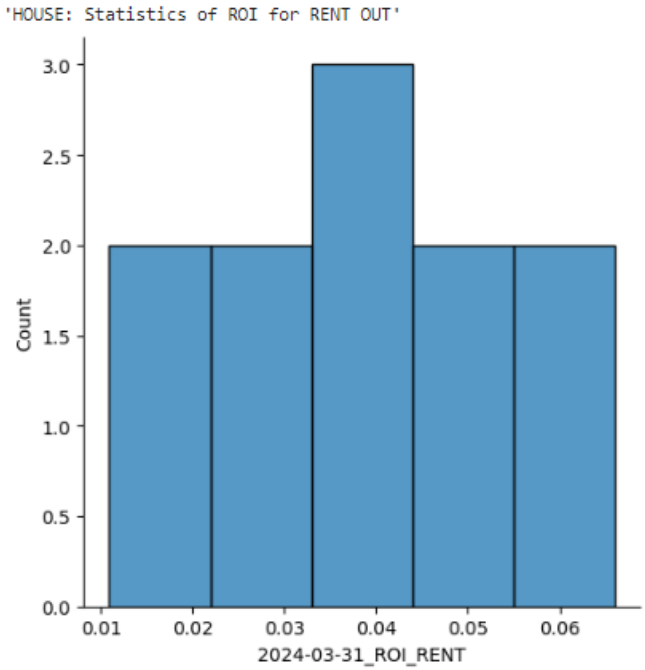


 doi:10.20944/preprints202603.1783.v1

Markets with shorter days-on-market periods indicate higher demand relative to supply, faster transaction completion, and reduced holding costs for investors. Conversely, markets with extended listing periods may signal softer demand conditions, potential pricing misalignment, or structural barriers to transaction completion. For investors prioritizing liquidity and rapid portfolio turnover, this analysis identifies favorable market environments.

4.5. ROI Distribution for Rental Strategy

Statistical analysis of the rental ROI distribution across regions (Figure 5) revealed a moderately right-skewed distribution, indicating that while most regions cluster around a central tendency, a meaningful subset of regions offer substantially higher-than-average rental returns. Outlier analysis identified specific regions where rental ROI exceeded three standard deviations from the mean, representing exceptional income-generating opportunities.




```
count    11.000000
mean      0.039112
std       0.017150
min       0.010885
25%       0.030354
50%       0.036568
75%       0.048323
max       0.066002
Name: 2024-03-31 ROI RENT, dtype: float64
```

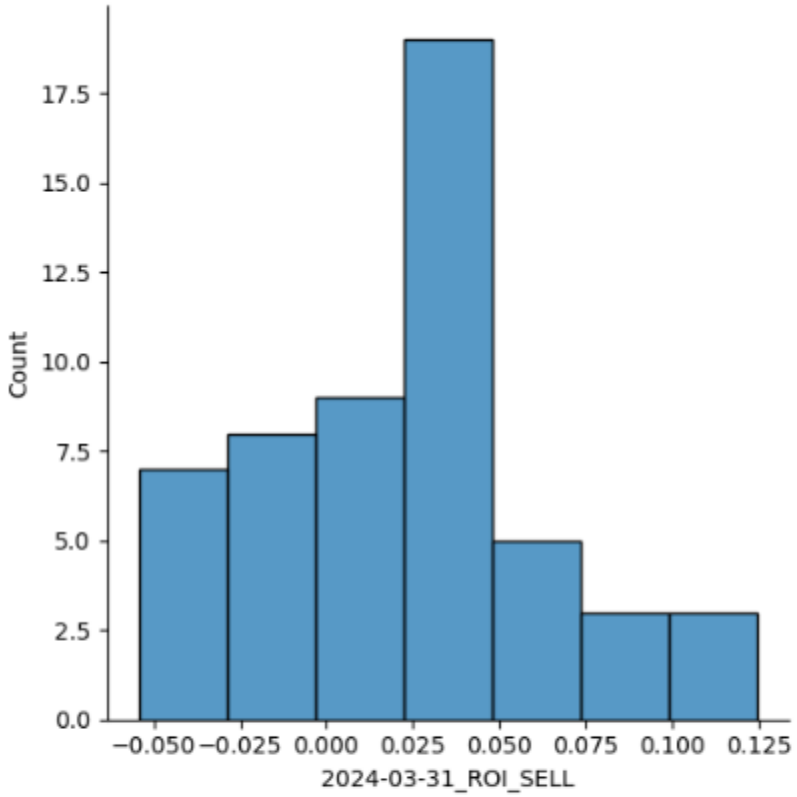
The identification of positive outlier regions is particularly valuable for rental-focused investment strategies, as these markets may offer superior income streams relative to acquisition costs. However, high-ROI outlier regions should be evaluated in conjunction with liquidity and growth metrics to assess the sustainability and risk profile of elevated returns.

4.6. ROI Distribution for Sale Strategy

The ROI distribution for sale strategies (Figure 6) exhibited distinct characteristics compared to the rental distribution, with different regions emerging as top performers. This divergence between rental and sale ROI rankings underscores the importance of aligning investment strategy with regional market conditions rather than applying a uniform approach across geographies.

 doi:10.20944/preprints202603.1783.v1

'HOUSE:Statistics of ROI for SELL'




```
count    54.000000
mean      0.023051
std       0.042757
min      -0.054286
25%      -0.015476
50%       0.031429
75%       0.042857
max       0.124762
Name: 2024-03-31_ROI_SELL, dtype: float64
```

Outlier analysis identified regions with exceptionally high sale ROI, driven by rapid property value appreciation in concentrated markets. The comparative analysis of rental versus sale ROI distributions provides a foundation for strategy selection, enabling investors to match their investment approach (income-focused vs. appreciation-focused) to the specific characteristics of their target markets.

5. Discussion

5.1. Synthesis of Findings

The multi-index analytical framework developed in this study demonstrates the value of integrating diverse housing market indicators for investment decision-making. The six analyses collectively provide a comprehensive view of the U.S. residential real estate landscape, addressing growth potential, income generation, market timing, liquidity, and comparative strategy assessment.

A key finding is the substantial regional heterogeneity observed across all dimensions of analysis. The top 3% of growth regions, the highest-ZORI metropolitan areas, the most liquid state-level markets, and the highest-ROI outlier regions do not perfectly overlap, indicating that no single market dominates across all investment criteria. This finding reinforces the importance of a multi-dimensional assessment framework that enables investors to prioritize criteria aligned with their specific objectives and risk preferences.

 doi:10.20944/preprints202603.1783.v1

5.2. Alignment with Existing Literature

The observed upward trend in rental indices across major metropolitan areas is consistent with broader demographic and economic trends documented in the housing literature. Kuhlmann (2021) noted that zoning reforms and densification efforts in major cities have been accompanied by continued pressure on rental prices, a pattern evident in the ZORI trajectories identified in this study. Similarly, the regional variation in days-on-market aligns with Bełej et al. (2020), who documented significant geographic heterogeneity in housing market dynamics across different market conditions.

The finding that rental and sale ROI distributions exhibit distinct characteristics supports the strategic differentiation between income-focused and appreciation-focused investment approaches, as theoretically motivated by Baum and Hartzell (2021). The practical implication is that investors should not treat rental and sale strategies as interchangeable but should instead select their approach based on regional market conditions and personal investment horizons.

5.3. Practical Implications

The framework presented in this study has several practical applications. For individual investors, it provides a systematic methodology for screening the U.S. housing market across multiple dimensions, reducing reliance on anecdotal information or geographically limited expertise. For institutional investors, the framework can be integrated into existing portfolio management processes to identify geographic diversification opportunities and monitor market conditions across target regions.

Furthermore, the methodology has implications for urban planning and public policy. The identification of high-growth regions and the analysis of rental price trajectories can inform housing policy discussions, particularly regarding affordability and supply-demand dynamics in rapidly appreciating markets. The days-on-market analysis provides a data-driven indicator of market tightness that can supplement traditional housing market assessments conducted by municipal planning departments.

5.4. Limitations

Several limitations of this study warrant acknowledgment. First, the analysis relies exclusively on Zillow's proprietary indices, which, while widely used, represent smoothed estimates derived from the platform's algorithmic models rather than direct transaction records. Indices based on public records, such as the Federal Housing Finance Agency (FHFA) House Price Index or the S&P CoreLogic Case-Shiller Index, may yield somewhat different results.

Second, the analytical framework does not incorporate macroeconomic variables such as interest rates, employment conditions, or inflation, which are known to significantly influence housing market dynamics. The integration of such variables would enhance the predictive and explanatory power of the framework.

Third, the ROI calculations represent simplified estimates that do not account for transaction costs, property management expenses, maintenance costs, property taxes, or financing structures. Actual realized returns may differ substantially from the gross ROI figures presented in this analysis.

Fourth, the property value forecast utilizes trend extrapolation rather than a formal predictive model, and should be interpreted as a directional indicator rather than a precise prediction.

Future research should address these limitations by incorporating macroeconomic covariates, developing more sophisticated forecasting models (potentially leveraging machine learning techniques), and extending the analysis to include net ROI calculations that account for realistic cost structures.

6. Conclusion

This study developed and demonstrated a multi-dimensional analytical framework for evaluating U.S. residential real estate investment opportunities using publicly available Zillow

 doi:10.20944/preprints202603.1783.v1

housing market data. Through six interconnected analyses spanning growth identification, rental trend assessment, investment timing, liquidity evaluation, and dual-strategy ROI comparison, the framework provides a comprehensive methodology for systematic investment evaluation.

The key findings can be summarized as follows. First, the top 3% of U.S. regions exhibited substantially above-average single-family home price appreciation between December 2022 and March 2023, providing a focused set of high-growth investment targets. Second, rental indices demonstrated a consistent and accelerating upward trend across major metropolitan areas, indicating favorable conditions for income-focused investment strategies. Third, property values were projected to appreciate by approximately 1.6% through March 2024, suggesting continued market growth. Fourth, significant state-level variation in days-on-market metrics was observed, providing a measure of relative market liquidity. Fifth, both rental and sale ROI distributions exhibited meaningful regional variation and distinct outlier patterns, underscoring the importance of strategy-market alignment in investment decision-making.

The proposed framework demonstrates that the systematic integration of publicly available housing market indices can yield actionable investment intelligence. By providing a reproducible, multi-dimensional approach to market evaluation, this research contributes both to the academic literature on data-driven real estate analysis and to the practical toolkit available to investors and urban planners operating in the U.S. residential market.

Data Availability Statement: The datasets analyzed in this study are publicly available from the Zillow Research Data repository at https://www.zillow.com/research/data/. The Zillow Home Value Index (ZHVI), Zillow Observed Rent Index (ZORI), and related market activity data can be accessed and downloaded without registration.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Baum, A. E., & Hartzell, D. (2021). Real estate investment: Strategies, structures, decisions. Wiley.
2. Bełej, M., Cellmer, R., & Głuszak, M. (2020). The impact of airport proximity on single-family house prices — Evidence from Poland. Sustainability, 12(19), 7928.
3. Holt, J. R., & Borsuk, M. E. (2020). Using Zillow data to value green space amenities at the neighborhood scale. Urban Forestry & Urban Greening, 56, 126794.
4. Kołodziejczyk, B., Mielcar, P., & Osiichuk, D. (2019). The concept of the real estate portfolio matrix and its application for structural analysis of the Polish commercial real estate market. Economic Research-Ekonomska Istraživanja, 32(1), 301-320.
5. Kuhlmann, D. (2021). Upzoning and single-family housing prices: A (very) early analysis of the Minneapolis 2040 plan. Journal of the American Planning Association, 87(3), 383-395.
6. Choy, L. H. T., & Ho, W. K. O. (2023). The use of machine learning in real estate research. Land, 12(4), 740. https://doi.org/10.3390/land12040740
7. Truong, Q., Nguyen, M., Dang, H., & Mei, B. (2020). Housing price prediction via improved machine learning techniques. Procedia Computer Science, 174, 433-442. https://doi.org/10.1016/j.procs.2020.06.111
8. Zillow Research. (2023). Zillow Home Value Index (ZHVI): Methodology and data documentation. Zillow Group.