ISM 6419 Data Visualization

Prof. Johannes Reichgelt

FINAL PROJECT REPORT

On

Real Estate Market Analysis

By

Sai Chowdary Penikalapati

Contents

1. Introduction

The housing market plays a crucial role in the economic landscape of any nation, influencing individual wealth, regional development, and broader economic trends. Over time, the dynamics of housing markets have shifted dramatically, driven by factors such as economic policies, population growth, technological advancements, and even global crises. Understanding these changes requires an interdisciplinary approach, integrating data from various sectors to paint a holistic picture. The dataset under analysis provides a comprehensive view of home values, sales counts, housing inventory, and forecast trends across the United States from 2000 to 2025. This report seeks to uncover insights by analyzing visualizations drawn from this data, identifying trends, key findings, and proposing actionable recommendations to address the challenges facing the housing market today.

2. Research Questions

To provide a structured analysis, the following research questions are formulated based on the visualizations:

1. **What are the historical trends in the U.S. housing market over the past 25 years?**
    - This question aims to explore the evolution of home values, identifying periods of growth, decline, and stability.
2. **How do regional disparities in house value indices manifest across different states?**

- This question focuses on geographical variations, using a heat map of house values by state to identify high and low-value regions.

3. **What is the projected trend for home values up to 2025, and which states are expected to see the largest changes?**
    - By examining forecast data, this question aims to predict which states will experience the most significant increases or decreases in home values.
4. **Is there a correlation between housing inventory and sales count, and how does this relationship vary by state?**
    - This question investigates the relationship between the availability of housing and the volume of home sales.
5. **What is the link between sales price and housing inventory, and how does it impact overall market dynamics?**
    - This question examines the pricing trends relative to inventory levels and how they shape the housing market.

3. Methodology

The datasets for this analysis are sourced from three key platforms: Zillow, the U.S. Census Bureau, and the U.S. Department of Justice's Criminal Justice Information Services (CJIS) division. Each of these datasets offers valuable insights into housing and population trends, with a specific focus on housing market dynamics and demographics.

- Zillow: Zillow is one of the largest online real estate databases in the U.S. The data covers key housing market indicators such as house values, sales counts, housing inventories, and forecasts for future housing market trends. Zillow's comprehensive housing data allows for detailed analysis of trends in home prices, market dynamics, and regional comparisons.

- U.S. Census Bureau: The Census Bureau provides data on housing units, including the number and characteristics of occupied and vacant housing units across the U.S. This dataset is crucial for understanding the distribution of housing, population growth patterns, and demographic shifts in various regions.

- U.S. Department of Justice, Federal Bureau of Investigation (CJIS): The CJIS database offers crime data collected across the United States, which can be used to assess the correlation between housing trends and crime rates. It helps to analyze social factors that may influence housing market conditions, such as public safety and neighborhood demographics.

Each dataset has been cleaned and processed to remove missing values, correct any data inconsistencies, and ensure accuracy.

Data Cleaning

The initial step in the data cleaning process involved handling missing data points, either by interpolating values based on adjacent data or by removing incomplete records. Outliers were also identified and addressed, particularly in the case of extreme home values or abnormal sales counts. Duplicate entries were removed to ensure data integrity.

Data Processing

Geographical coordinates (latitude and longitude) were generated for states without them, enabling the creation of heat maps for visual representation. Advanced statistical models were employed to predict future home values, using a combination of historical trends and forecast indicators.

4. Visual Analysis of Insights

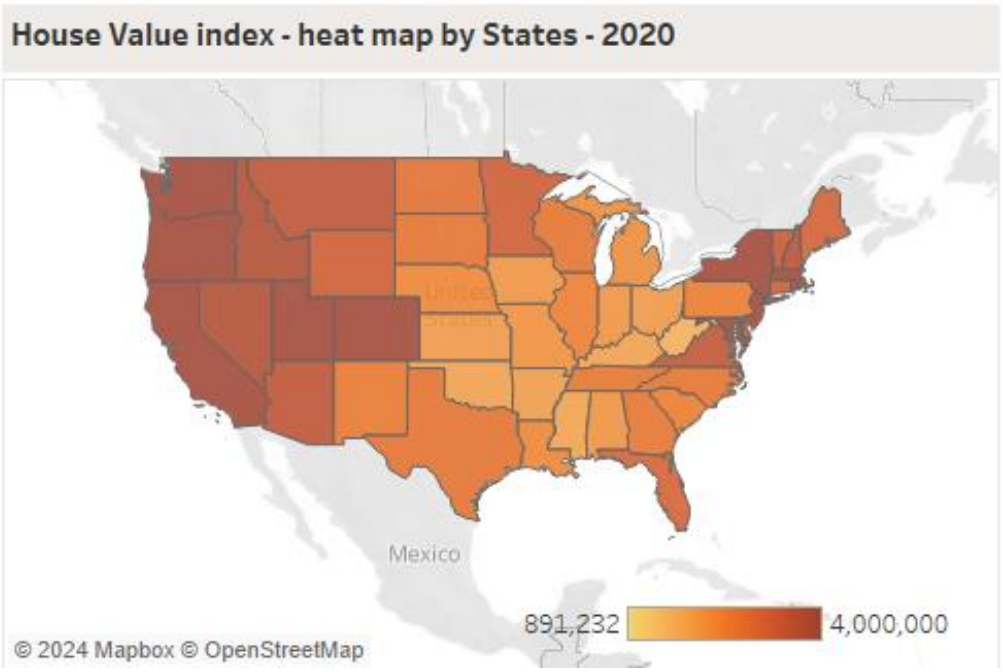
House Value Index - Heat Map by States

The heat The U.S. housing market is a fascinating and ever-changing landscape, heavily influenced by economic forces, policy changes, and societal trends. The visuals in our dataset offer a deep dive into various aspects of this market—trends in home values, sales volumes, and housing inventories—highlighting both regional disparities and nationwide patterns. In this analysis, we'll walk through nine key visuals, integrating insights drawn from them to form a compelling narrative that reflects the dynamics shaping the housing market from 2000 to 2025.

Insight 1: House Value Index - Heat Map by States

Our story begins in the year 2020 with a heat map of house values across the United States, where dramatic disparities are immediately evident. States like **California, Washington, and Massachusetts** show significantly higher house value indices, painting a picture of vibrant economies and strong demand for real estate. California stands out with its staggering home values averaging **$6.75 million**, driven by the tech boom, high-income residents, and severe housing shortages in cities like San Francisco and Los Angeles.

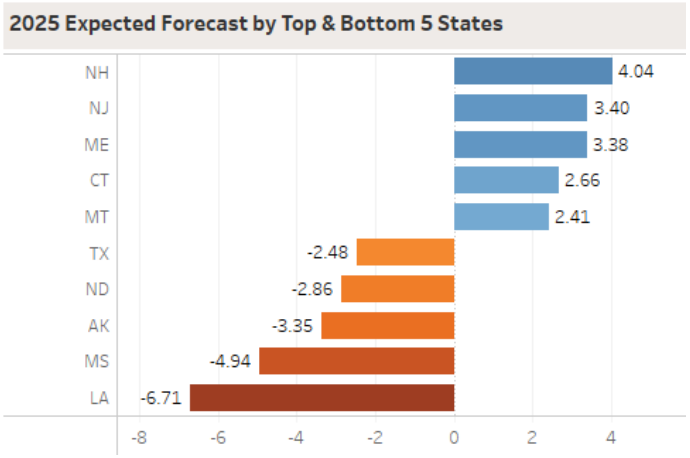
In contrast, states such as **Arkansas, Mississippi, and West Virginia** display far lower house value indices. Here, the visual captures a different story—a struggle to attract investment and population, leading to stagnating or slowly growing property values. This map gives us our first crucial insight: **the U.S. housing market is not homogenous; it is a patchwork quilt of prosperity and stagnation, heavily influenced by location and local economies.**



Insight 2: 2025 Expected Forecast by Top & Bottom 5 States

Looking into the future, the second visual focuses on the projected changes in house values up to **2025**, emphasizing both growth and decline. States like **New Hampshire and New Jersey** are expected to see house value increases of over 4%, while **Louisiana and Mississippi** may experience slight declines. This forecast invites us to consider the economic and demographic factors driving these trends. For example, New Hampshire's projected rise may be attributed to its proximity to booming urban centers and a steady influx of residents seeking suburban or rural lifestyles post-pandemic. Meanwhile, the expected decline in Louisiana's home values might reflect economic instability or vulnerability to natural disasters, which make long-term investments in property riskier.

This forecast draws attention to the uneven recovery expected in the coming years, **while some states will continue to prosper, others may lag, impacted by factors beyond immediate control, such as climate risks or economic shifts.**

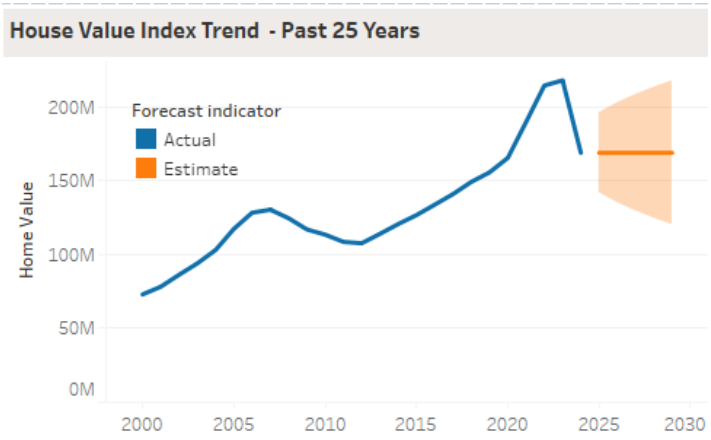


Insight 3: House Value Index Trend - Past 25 Years

To understand how we got here, we look back at the broader **25-year trend in house values**. This visual chart shows a journey through the early 2000s real estate boom, the **2008 financial crisis**, and the recovery that followed. From **2000 to 2007**, home values surged as easy credit, speculative investment, and high demand fueled a housing bubble. However, the **2008 crisis** marked a sharp downturn, with home values plummeting across the country, leading to foreclosures and a loss of wealth for millions of Americans.

Interestingly, the recovery phase, beginning around **2012**, shows a steady upward trend in home values, which accelerated dramatically post-**2020** due to the pandemic. The pandemic created an unexpected housing market boom such as low interest rates, remote work, and lifestyle shifts led

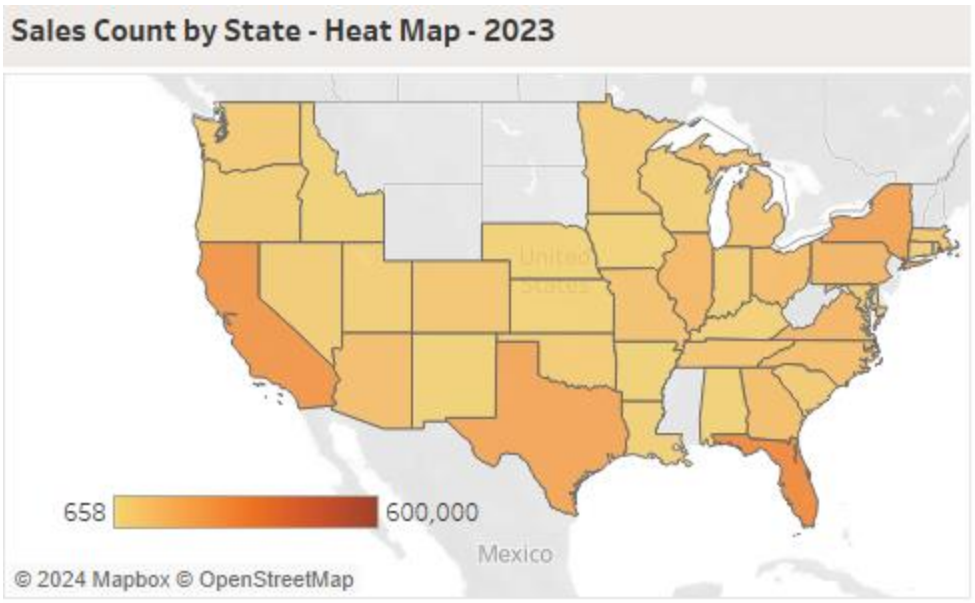
to increased demand for housing in suburban and rural areas. The visual shows the **peak in 2021**, followed by a slight moderation as the market begins to stabilize. The story here is one of resilience—the U.S. housing market, despite its ups and downs, has continually recovered and even thrived after periods of economic downturn.



Insight 4: Sales Count by State - Heat Map

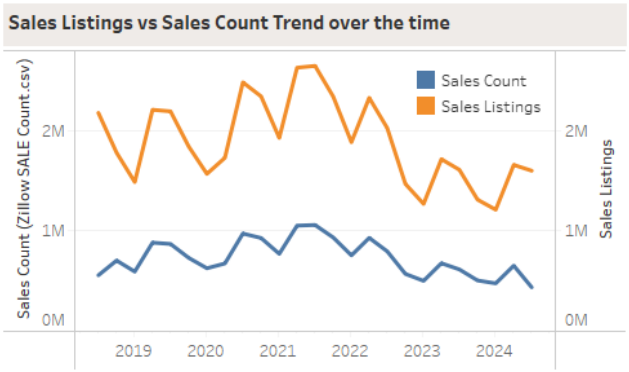
A heat map of home sales per 100,000 residents reveals significant disparities among states. For instance, states like Florida (1,608 per 100k), Virginia (1,699 per 100k), and Arizona (1,472 per 100k) show exceptionally high home sales rates, while states such as Kansas (325 per 100k) and Iowa (373 per 100k) are on the lower end.

This prompts me to question what makes these states different. The numbers hint at a complex interplay between factors like job opportunities, climate, and affordability. However, simply examining sales figures feels like trying to complete a puzzle with only some of the pieces.



Insight 5: Sales Listings vs Sales Count Trend Over Time

Next, the fifth visual tracks the relationship between **sales listings and sales counts from 2018 to 2024**. There is a noticeable cyclicality in the data, with peaks typically occurring in the second and third quarters of each year, reflecting the traditional housing market cycles where spring and summer see the highest activity. However, the visual reveals a **declining trend** in sales listings despite a steady rise in sales counts leading up to 2021, followed by a slight downturn by 2024. This indicates that while fewer homes are being listed, the homes that do go on the market are selling quickly, hinting at **a seller's market where demand outstrips supply**.

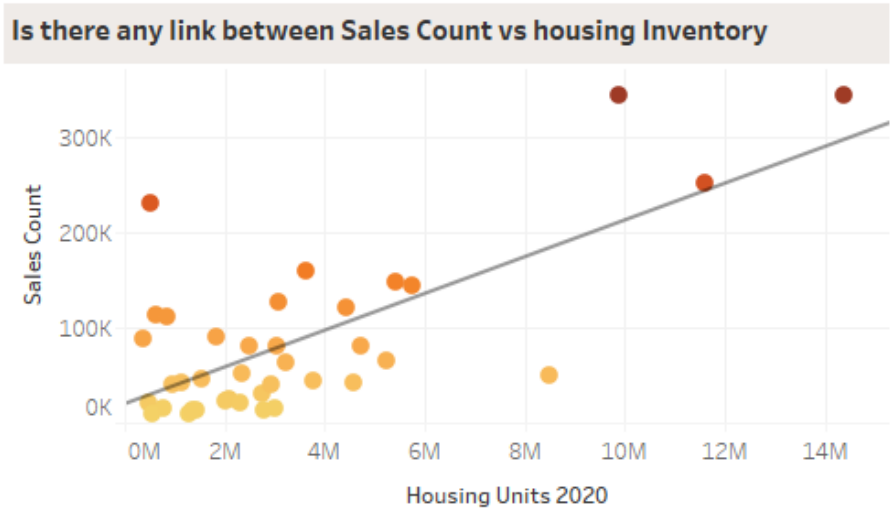


Insight 6: Link Between Sales Count vs Housing Inventory

Curious if housing inventory plays a role in driving these sales figures, I turn to another dataset. Patterns begin to emerge:

- **Arizona** has a moderate inventory of 41,880 homes per 100k but also boasts one of the highest sales counts.
- **Florida**, with an inventory of 44,349 homes per 100k, tops the sales chart.
- In contrast, **Nebraska** has a high inventory of 152,371 homes per 100k, yet its sales rate is moderate (657 per 100k).

These findings suggest that high inventory doesn't always translate to higher sales. States with moderate inventory but strong demand, like Arizona, are rapidly selling homes. Meanwhile, oversupply in areas like Nebraska may indicate a mismatch between what's available and what buyers desire.

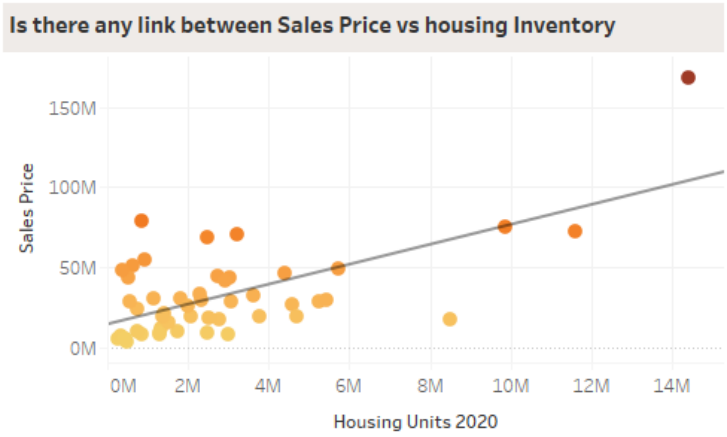


Insight 7: Link Between Sales Price vs Housing Inventory

Shifting focus to the relationship between housing prices and inventory levels, I notice another intriguing pattern:

- **California**, with a tight inventory of 36,875 homes per 100k, commands a sky-high average sales price of $224.6 million.
- **Texas**, despite a larger inventory of 38,593 homes per 100k, sees average sales prices around $93.3 million.
- Meanwhile, states like **West Virginia**, with relatively high inventory (48,201 per 100k), see much lower sales prices of about $9.9 million.

It becomes evident that while limited supply can drive up prices in competitive markets like California, other regions with higher inventories may struggle to increase prices due to weaker demand.

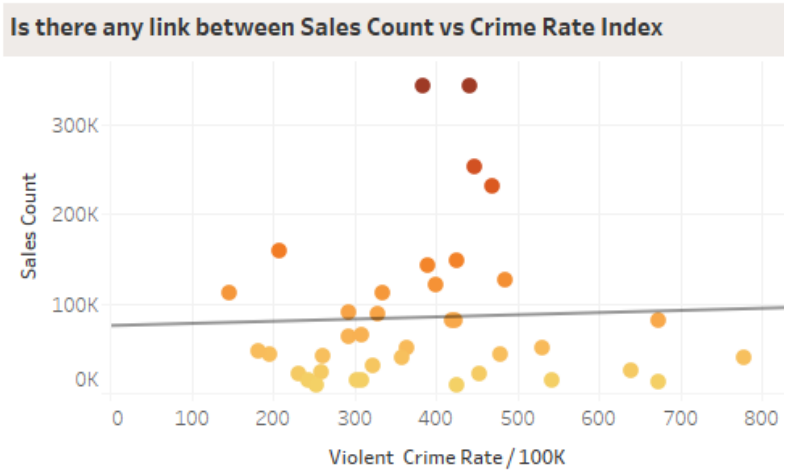


Insight 9: Link between Sales vs Crime Rate

I explore whether safety might influence home sales. The question arises: "Do buyers prefer states with lower crime rates?"

- **Virginia**, leading in sales per capita, also has a low violent crime rate of 208.7 per 100k.
- **Arizona** shows a high sales count alongside a moderate crime rate of 484.8 per 100k.
- Interestingly, states like **Illinois** (425.9 per 100k) and **Florida** (383.6 per 100k) maintain robust home sales despite moderate crime rates.

The data suggests that while crime rates might affect buyer sentiment, they are not the sole driver of housing activity. Factors such as affordability, job market strength, and tax incentives may play a more significant role in some states.



5. Key Findings

1. **Geographical Disparities in Home Values**: There are significant regional differences in home values, with states like California and Washington having much higher property values than states like Arkansas and Mississippi.
2. **Forecast Trends**: By 2025, home values are expected to rise in states like New Hampshire and New Jersey, while states like Louisiana and Mississippi are likely to see declines.
3. **Sales Trends**: States with higher housing inventories generally have higher sales counts, although this relationship is influenced by regional economic conditions.
4. **Market Saturation**: The declining trend in sales listings relative to sales counts suggests a potential saturation in the housing market, with fewer homes being sold despite an increase in listings.

6. Recommendations and Implementation

1. **Regional Policy Adjustments**: Policymakers should focus on addressing the economic disparities that contribute to the differences in home values and sales across states. Investing in infrastructure and job creation in lower-value regions could help to balance the market.
2. **Affordable Housing Initiatives**: In states where home values are expected to decline, affordable housing programs could be expanded to prevent market destabilization and provide housing for low-income populations.
3. **Sustainable Development**: Encouraging sustainable housing developments in high-demand areas can help alleviate pressure on the housing market and ensure long-term stability.
4. **Technology-Driven Forecasting**: Real estate agencies and policymakers should adopt more advanced forecasting tools, incorporating machine learning models to predict market trends with greater accuracy.

7. Conclusion

The U.S. housing market is complex and influenced by numerous factors, including geographical location, economic conditions, and market demand. While some states continue to see rising home values and strong sales counts, others are facing challenges related to affordability and market saturation. By analyzing the trends and patterns in house values, sales counts, and

housing inventories, we can gain a deeper understanding of the market's dynamics and make informed decisions to promote a more balanced and sustainable housing market.

8. Additional Research Questions

1. How does population migration affect regional home values and sales trends?
2. What is the impact of rising interest rates on housing affordability in different states?
3. How does housing demand in urban vs. rural areas shift in response to economic changes?

9. Reference

Zillow. (n.d.). Zillow research data. Zillow. Retrieved November 11, 2024, from https://www.zillow.com/research/data/

U.S. Census Bureau. (n.d.). Housing units: Data. U.S. Census Bureau. Retrieved November 11, 2024, from https://data.census.gov/table?q=Housing%20Units

U.S. Department of Justice, Federal Bureau of Investigation. (n.d.). Criminal justice information services (CJIS) division. U.S. Department of Justice. Retrieved November 11, 2024, from https://cde.ucr.cjis.gov/LATEST/webapp/#/pages/explorer/crime/crime-trend