U.S. Housing Affordability: A State-Level Study of Economic, Demographic, and Risk Indicators

Introduction

Housing affordability has become a growing concern across the United States. In many areas, home prices and rents are rising faster than household incomes, making it increasingly difficult for individuals and families to find affordable places to live. According to external sources such as Zillow, the U.S. Census Bureau, and FEMA, factors like income levels, employment opportunities, mortgage rates, population changes, state-level economic sectors, and natural disasters all play a role in shaping housing costs and accessibility. These pressures not only affect homeowners and renters but also influence broader economic resilience and community well-being.

This project analyzes housing affordability trends across all 50 U.S. states from 2015 to 2023. By combining data from multiple reliable sources, the project offers a clearer understanding of where affordability is declining, why these trends are occurring, and how different states are affected.

The study addresses the following clearly formulated research questions:

1. How have home prices and rental rates changed over time in U.S?
2. Which areas are the most and least affordable for homeownership and renting, and how does affordability stress vary across the U.S.?
3. How do income levels, mortgage rates, and key socioeconomic stress indicators (poverty, unemployment, insurance) impact housing affordability?
4. How does population growth or decline affect housing prices and affordability?
5. How do broader economic sectors (GDP, agriculture) and demographic characteristics (age, race, gender, risk exposure) shape housing affordability and access?

These questions provide the foundation for an integrated, evidence-based analysis of housing patterns across the country, aiming to generate insights that can inform future housing policy, economic planning, and decision-making at multiple levels.

Methodology

Data Sources and Links:

This project uses multiple datasets from well-established public and research platforms to analyze U.S. housing affordability. Below is a summary of the data sources, what they provide, and where they were sourced from:

1. **Zillow (ZHVI & ZORI)** — Provides yearly estimates of home values and rental rates by state.
   Source: https://www.zillow.com/research/data/
2. **U.S. Census Bureau (ACS DP03 & DP05)** — Supplies state-level economic and demographic data, including income, poverty rates, health insurance coverage, unemployment, age, race, travel time, and gender.
   Sources:
   - DP03 (Economic): https://data.census.gov/table/ACSDP5Y2023.DP03
   - DP05 (Demographic): https://data.census.gov/table/ACSDP5Y2023.DP05
3. **Census Population Estimates Program (PEP)** — Tracks population changes, births, deaths, and migration across states.
   Sources:
   - https://www.census.gov/data/datasets/time-series/demo/popest/2020s-state-total.html
   - https://www.census.gov/data/datasets/time-series/demo/popest/2010s-state-total.html
4. **Bureau of Labor Statistics (BLS CES)** — Provides monthly employment statistics at the state level, aggregated to annual averages.
   Source: https://data.bls.gov/cgi-bin/dsrv?sm
5. **Bureau of Economic Analysis (BEA GDP)** — Offers annual gross domestic product (GDP) data by state.
   Source: https://apps.bea.gov/iTable/iTable.cfm?ReqID=70&step=1
6. **FEMA National Risk Index (NRI)** — Provides disaster risk scores, vulnerability indices, and hazard ratings for each state.
   Source: https://hazards.fema.gov/nri/data-resources
7. **Federal Reserve Economic Data (FRED)** — Supplies national-level 30-year fixed mortgage rate data.
   Source: https://fred.stlouisfed.org/series/MORTGAGE30US
8. **Kaggle (USA States to Region Mapping)** — Provides a mapping of U.S. states into broader regions (Northeast, Midwest, South, West) for regional trend analysis.
   Source: https://www.kaggle.com/datasets/omer2040/usa-states-to-region

Data Integration Approach:

After collecting these datasets, we applied a structured integration and preparation process to combine, align, and enrich the data for analysis.

1. **Pre-processing:**
   Gathered and cleaned all datasets, fixed missing or mismatched values, and transformed them to the state and year level (except mortgage data, which is only at the year level, and FEMA

NRI data, which is only at the state level). Combined ZORI and ZHVI as a single housing dataset, and merged Census DP03 and DP05 as one combined demographic-economic dataset. All datasets were then joined using State and Year as the main keys.

2. **Regional Grouping:**
   Used the Kaggle state-to-region mapping to group states into Northeast, Midwest, South, and West for regional-level analysis.

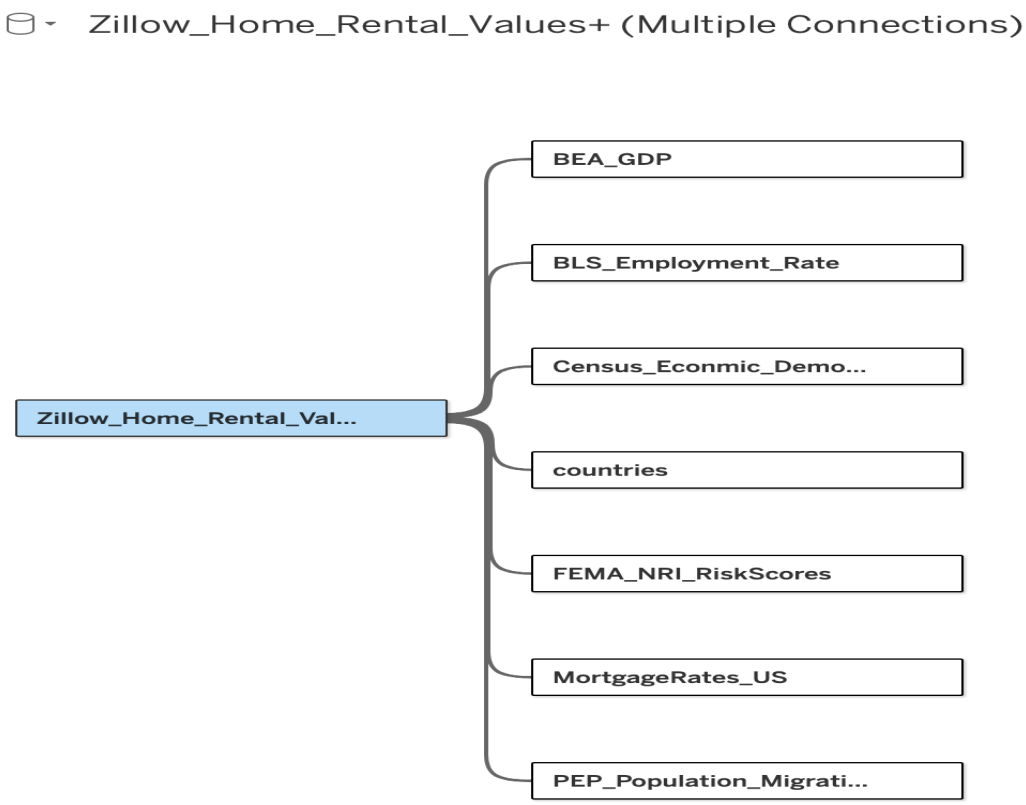
3. **Calculated Metrics:**

Metric / Parameter	Description
Cost_to_buy	Annual cost of buying a home, including mortgage payments and maintenance.
Cost_to_rent	Annual cost of renting a home.
Affordability_index	Overall index comparing combined housing costs to median household income.
AgriValue_per_Population	Agricultural sector value normalized by state population.
Avg Home Price-to-GDP Ratio (%)	Share of the housing market relative to the state's total economic output (GDP).
Employment-to-Population Ratio	Employment rate adjusted relative to total state population.
Home Affordability Index	Ratio of home-buying costs specifically to median household income.
Home Value vs Comparison Metric Ratio	Comparison of home value against socioeconomic stress indicators (poverty, unemployment, insurance gaps).
Home Value vs Income Ratio	Direct ratio of average home price to average income.
Top5_States_Set	Identifies the top 5 states with the highest home value-to-income burden.
Bottom5_States_Set	Identifies the bottom 5 states with the lowest home value-to-income burden.
HomeValue_percentage	Year-over-year percentage change in home value.
MortgageRate_percentage	Year-over-year percentage change in mortgage rates.
Rent Affordability Index	Ratio of annual rent costs to median household income.
Selected Demographic Value	Filtered metric focusing on a selected demographic factor (such as age, race,Population or gender).

4. **Parameters Used:**

Parameter	Description
Affordability_index (Both Home, Rent exceed Threshold)	Defines what affordability stress means by setting thresholds for both homeownership and rental costs.
Comparison Metric	User-defined metric to compare home values against selected socioeconomic indicators (e.g., poverty, unemployment).
Demographic Factor Selector	Parameter allowing selection of which demographic variable (such as age, race, or gender) to analyze.

Table Mapping:

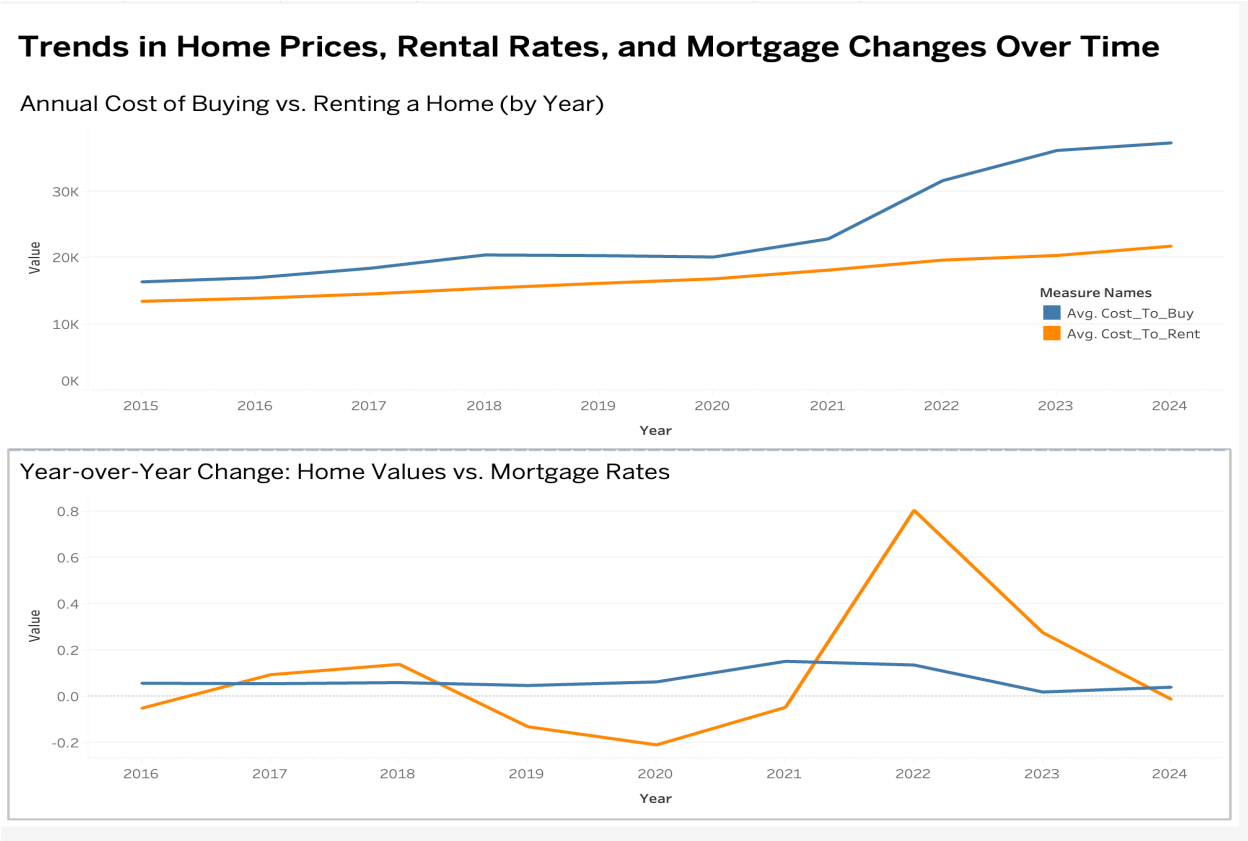


The tables were linked using the following keys:

- Between **State** and **Year** in Zillow_Home_Rental_Values and BEA_GDP.
- Between **State** and **Year** in Zillow_Home_Rental_Values and BLS_Employment_Rate.
- Between **State** and **Year** in Zillow_Home_Rental_Values and Census_Economic_Demographic.
- Between **State** in Zillow_Home_Rental_Values and FEMA_NRI_RiskScores
- Between **State** and **Year** in Zillow_Home_Rental_Values and PEP_Population_Migration
- Between **State** in Zillow_Home_Rental_Values and the countries table from Kaggle (state-to-region mapping).
- Between **Year** in Zillow_Home_Rental_Values and MortgageRates_US

Analysis

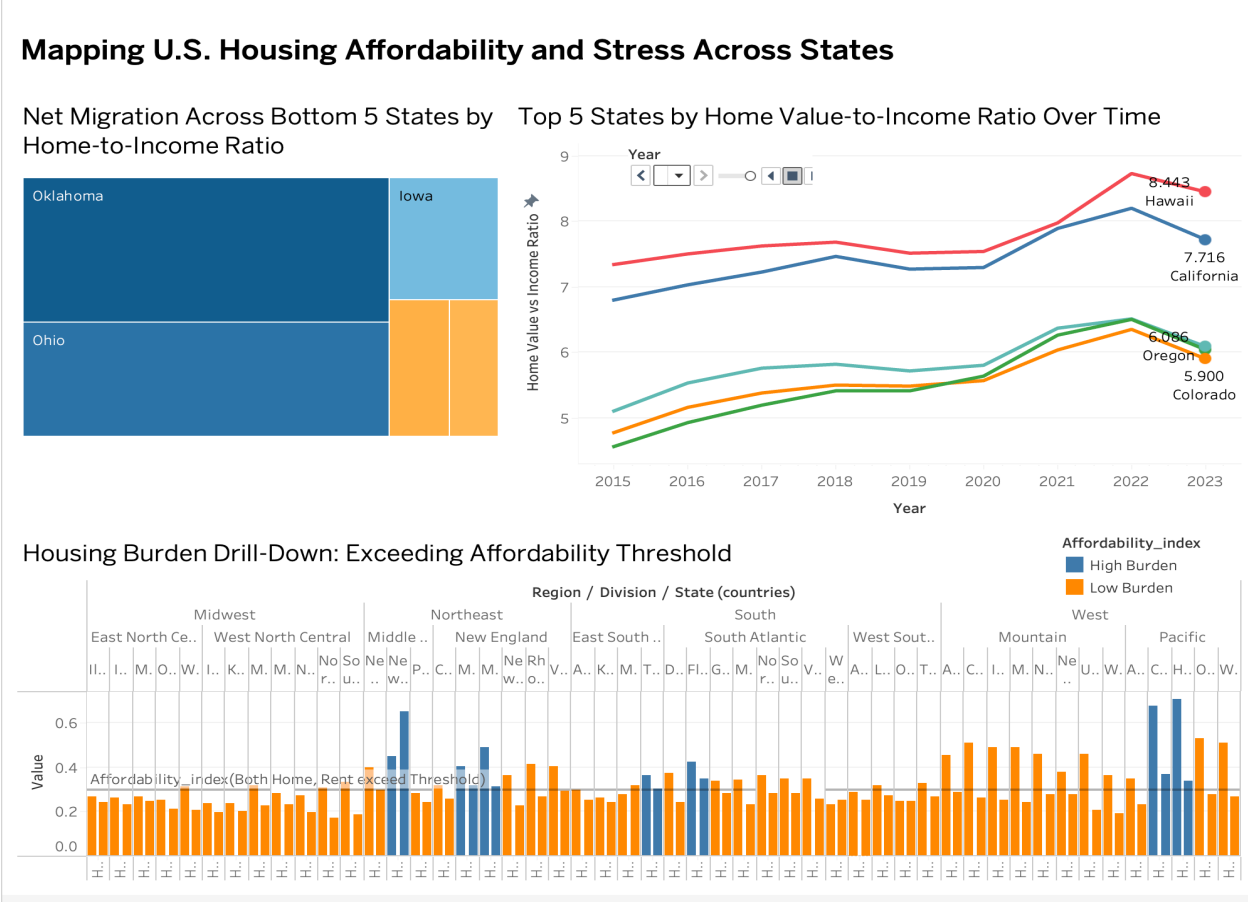
1. **How have home prices and rental rates changed over time in U.S.?**



Interpretation:

- The "**Annual Cost of Buying vs. Renting a Home (by Year)**" graph shows that over time, the cost to buy a home has increased much more sharply than the cost to rent, especially from 2021 onward. This widening gap suggests that while renting costs have gone up moderately, the costs for potential homebuyers have surged, making homeownership increasingly difficult.
- The "**Year-over-Year Change: Home Values vs. Mortgage Rates"** graph highlights that mortgage rates spiked dramatically in 2022, even though home price changes remained more gradual. This means that even if home prices weren't rising steeply, the increased mortgage rates still raised the overall cost of purchasing a home.
- Overall, this highlights that rising mortgage rates and home prices together are making homeownership less affordable over time, pushing more people toward renting as the more accessible option. how much you rate for this interpretation.

2. **Which areas are the most and least affordable for homeownership and renting, and how does affordability stress vary across the U.S.?**



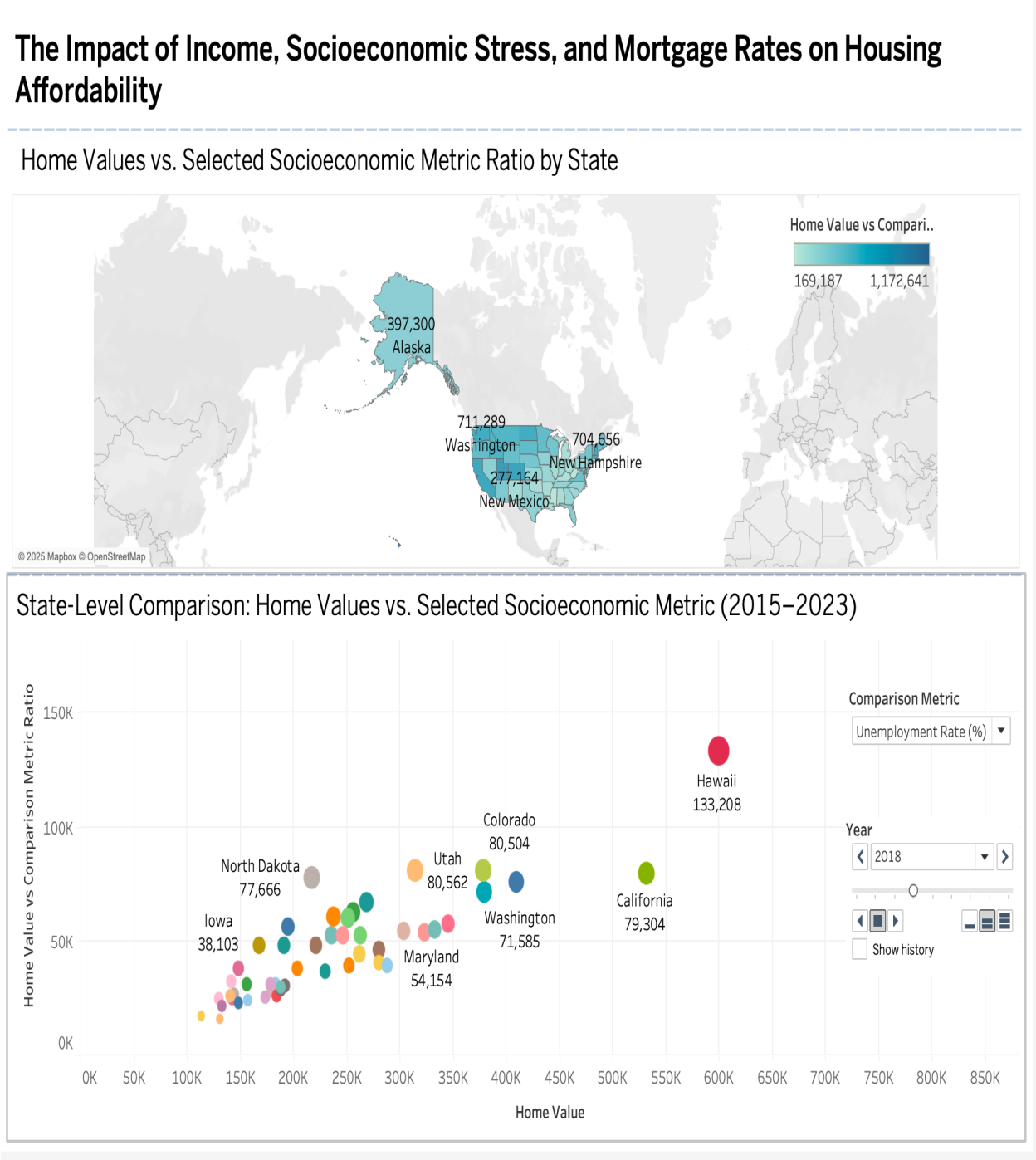
Interpretation:

- The "**Net Migration Across Bottom 5 States by Home-to-Income Ratio"** treemap highlights states like Oklahoma, Iowa, and Ohio, which have the most affordable ratios. Interestingly, even though these states are more affordable, they still show net out-migration, which may reflect broader factors beyond housing costs, such as local job markets or lifestyle preferences.

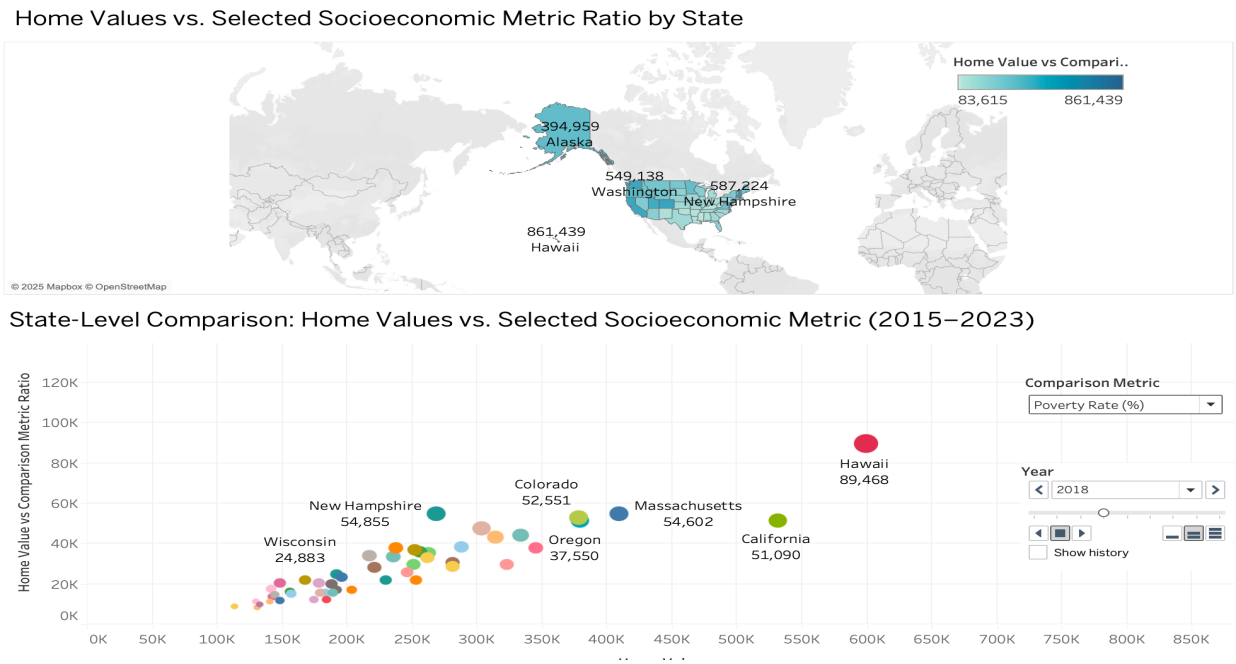
- The "**Top 5 States by Home Value-to-Income Ratio Over Time**" chart shows that Hawaii and California have consistently been the least affordable, where home values are much higher than incomes. On the other hand, Colorado, Oregon, and Washington, although expensive, show more moderate ratios, meaning they are slightly more affordable than Hawaii and California.
- The "**Housing Burden Drill-Down"** bar chart shows how affordability varies by region, with the West and Northeast having more high-burden counties, while the Midwest and South show more low-burden areas. Some states like New York, Hawaii, and Florida cross the affordability threshold, showing that in these places, both rent and home values are putting pressure on residents.
- Overall, this shows that the most and least affordable areas are not evenly spread across the U.S., and affordability challenges are more severe in certain high-cost states and regions.

3. **How do income levels, mortgage rates, and key socioeconomic stress indicators (poverty, unemployment, insurance) impact housing affordability?**

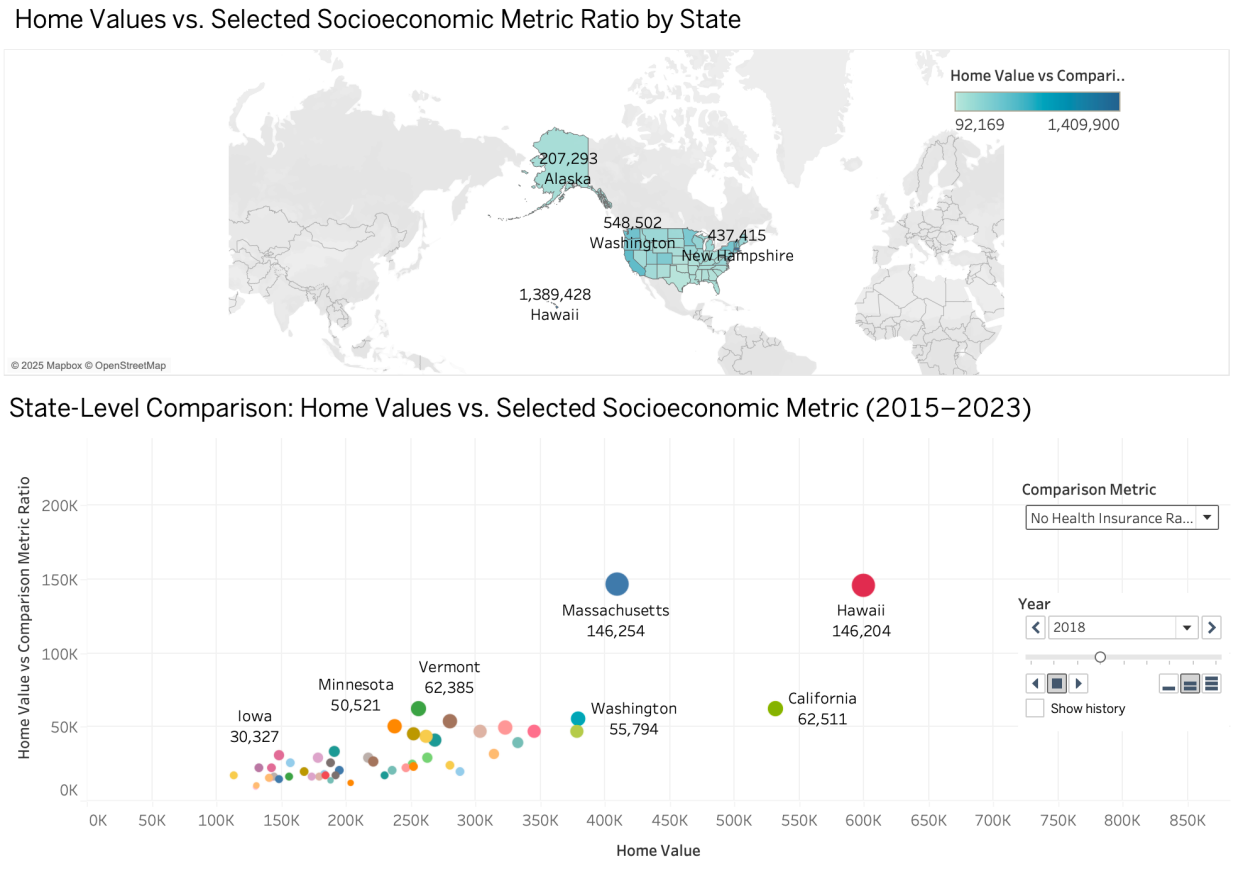
Unemployment Rate:



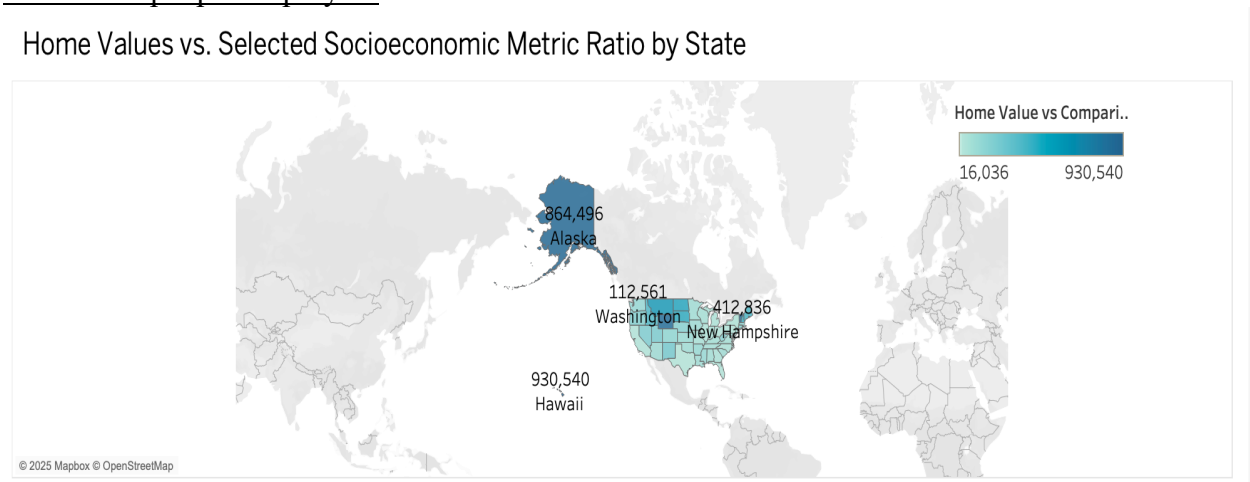
Poverty Rate:

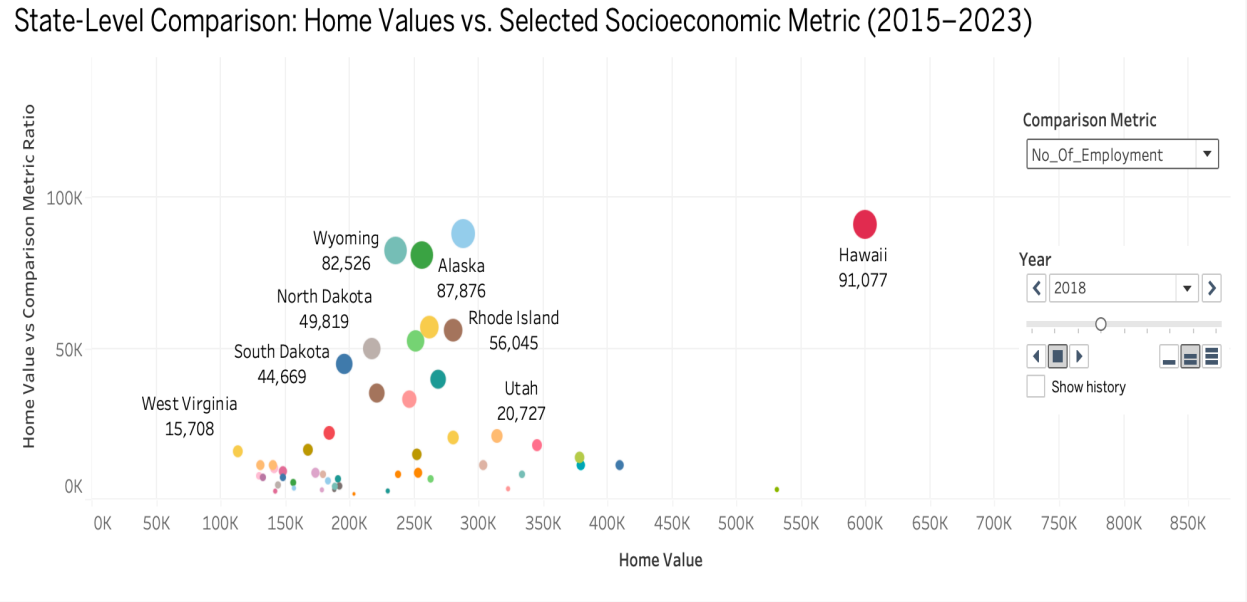


No health insurance rate:



Number of people employed:

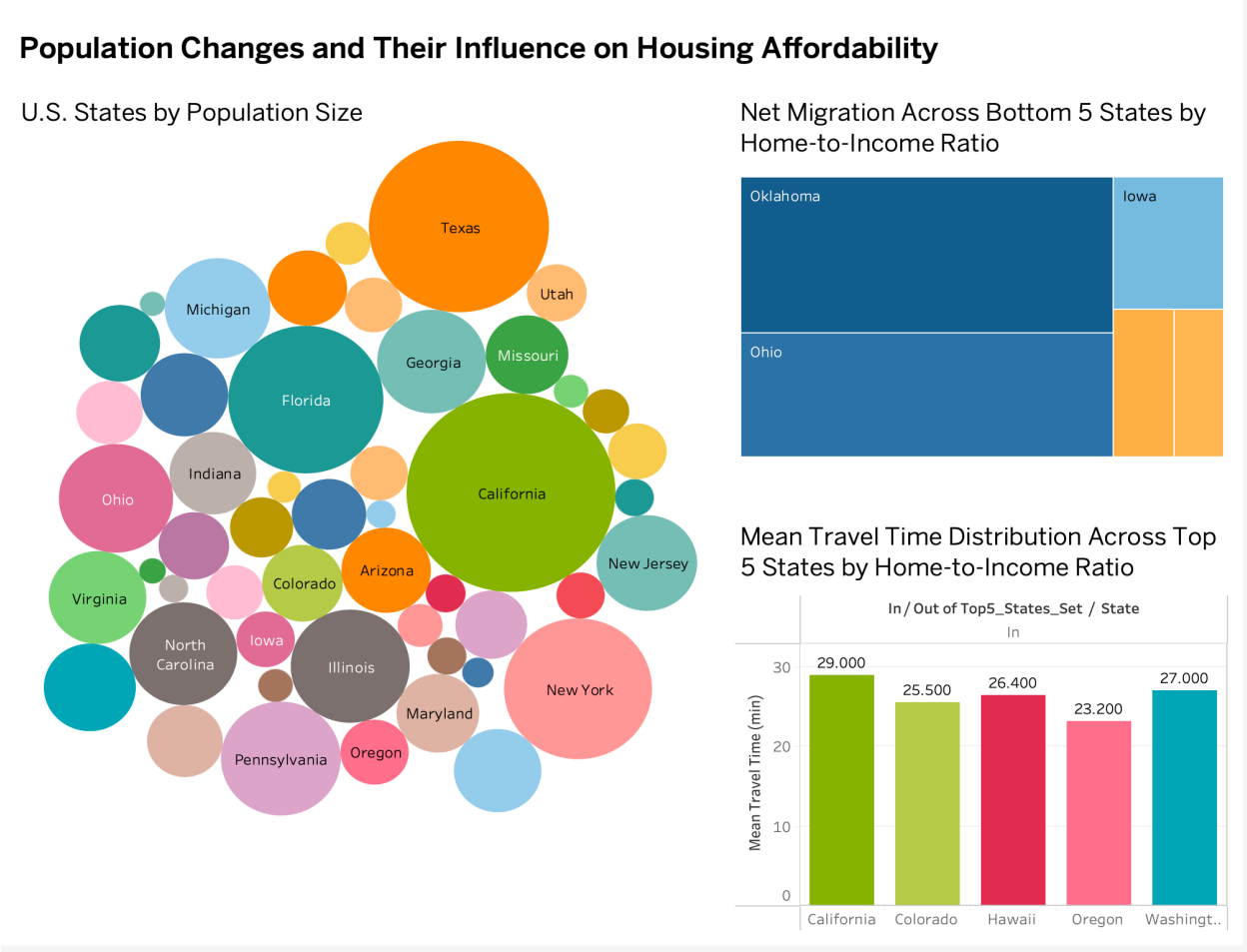




Interpretation:

- The **Home Values vs. Selected Socioeconomic Metric Ratio by State** map shows clear state-level differences, where places like California and Washington and show the most pressure which means not only are home prices high, but factors like unemployment, poverty, and lack of insurance make affordability even worse.
- The **State-Level Comparison: Home Values vs. Selected Socioeconomic Metric (2015–2023)** scatter plot adds to this by showing that states with higher unemployment or poverty tend to have steeper home value burdens compared to their local conditions. Hawaii stands out with extremely high home values despite its socioeconomic stress, while states like North Dakota and Utah show more balanced conditions, meaning their housing markets are less crushing on residents.
- Overall, this shows that income levels and socioeconomic stress directly affect housing costs, making homes either more or less affordable depending on the state.

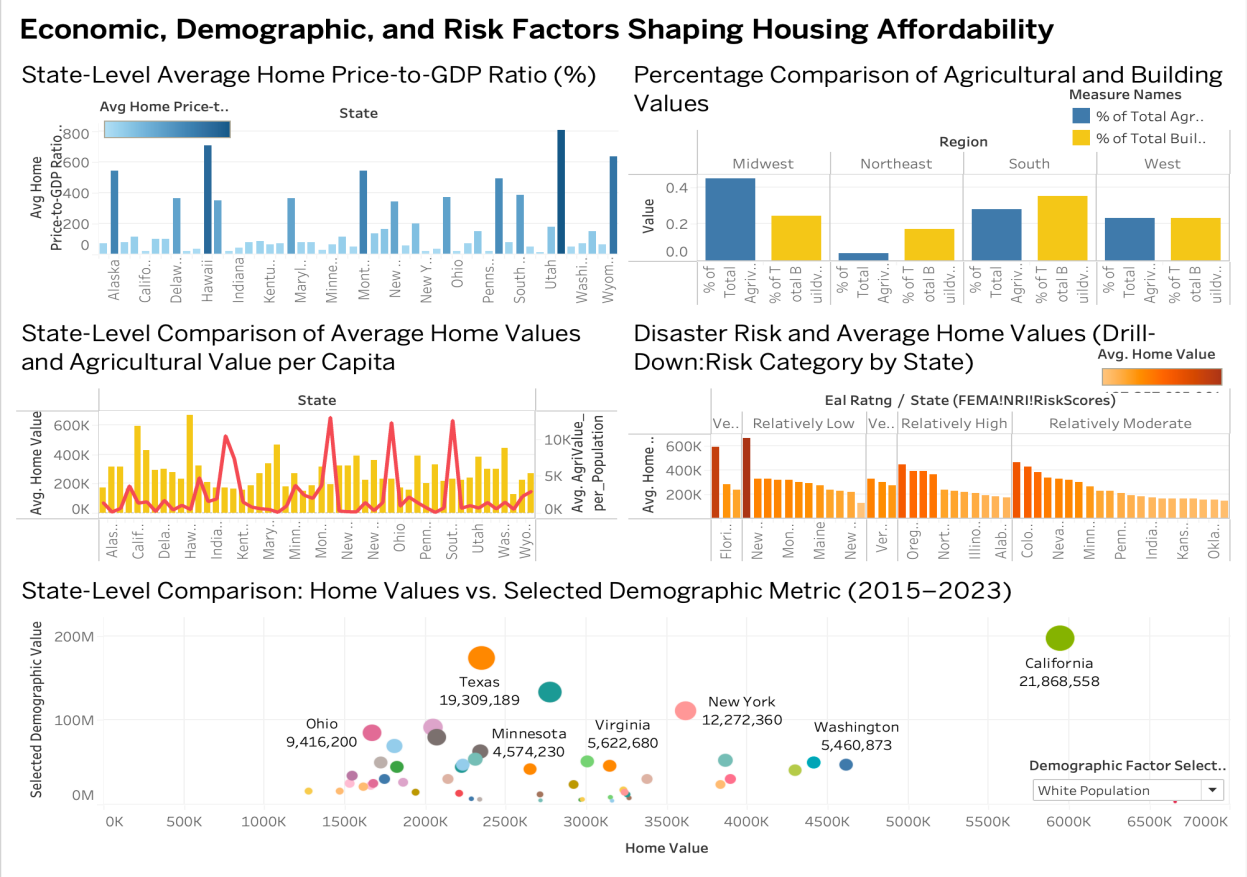
4. **How does population growth or decline affect housing prices and affordability?**



Interpretation:

- The **U.S. States by Population Size** bubble chart shows that big states like California, Texas, and Florida face major housing market pressure just from their size, where high demand pushes up home prices.
- The **Net Migration Across Bottom 5 States by Home-to-Income Ratio** treemap highlights that more affordable states like Oklahoma and Ohio are seeing positive net migration, which may suggest that lower housing costs are a factor attracting people relocating for better affordability.
- The **Mean Travel Time Distribution Across Top 5 States by Home-to-Income Ratio** bar chart points out that in expensive states like California and Hawaii, residents deal with longer commutes, suggesting that high housing prices may contribute to longer commutes, as residents adjust to living farther from where they work.
- Overall, this shows that population trends, whether from large state populations or migration shifts, directly influence housing affordability by either driving up costs in high-demand areas or easing pressures in more affordable regions.

5. **How do broader economic sectors (GDP, agriculture) and demographic characteristics (age, race, gender, risk exposure) shape housing affordability and access?**



Interpretation:

- The **State-Level Average Home Price-to-GDP Ratio (%)** shows that in some states, housing costs take up a large part of the economy, indicating that affordability in these states may be more sensitive to price spikes and economic shifts. The **Percentage Comparison of Agricultural and Building Values** highlights that the Midwest's strong agricultural base may help explain more stable housing pressures compared to urban-heavy regions. The Average Home Values vs. Agricultural Value per Capita graph suggests that states with more agriculture often see more stable or moderate housing costs.
- The **Disaster Risk vs. Home Values chart** shows that areas with high risk (like wildfires or hurricanes) often have surprisingly high housing prices, adding hidden cost burdens. The Scatter Plot on Demographic Factors shows how age, race, and gender patterns connect to home value differences, with places like California standing out for extreme disparities.
- Overall, these factors together show that housing affordability is not just about prices or income, and it's also deeply shaped by local economic structures and who lives there.

Conclusion

This analysis shows that U.S. housing affordability is shaped by a combination of rising home prices, surging mortgage rates, local income levels, socioeconomic stress, demographic patterns, and broader economic factors like GDP and agriculture. While rising home prices and mortgage rates have steadily made homeownership more difficult, regional differences in income, employment, and population trends create varying affordability levels across states.

We observe that larger states with booming populations face stronger price pressures, while states with strong agricultural bases or lower migration inflows experience more stability. Additionally, states with high disaster risk or demographic imbalances often carry hidden burdens that increase the burden on affordability.

Finally, housing markets are not just economic systems but also social systems, shaped by who lives where, local industry makeup, and environmental pressures. Rising rates or prices alone don't explain everything. It's their interaction with local conditions that ultimately determines housing affordability.

Future Research Questions:

1. How does crime rate in a region affect housing prices and people's decisions to rent or buy?
2. How do real estate development patterns, like new housing projects or urban renewal, impact long-term housing affordability?