Yikuan Ye
Dec/6/2023
GGIS 480
Final Project Report

Geographic Clustering Analysis of the Impact of COVID-19 Pandemic on Housing Prices in the United States

Introduction

The COVID-19 pandemic has significantly impacted various sectors globally, including real estate. Understanding its effect on housing prices across different regions in the U.S. is crucial for multiple stakeholders including policymakers, urban planners, investors, and the public. In this research, I expect finding out how has the COVID-19 pandemic affected housing prices in different regions of the United States? What patterns of geographic clustering can be observed in these changes?

Background

The impact of the COVID-19 pandemic on real estate markets has been a focal point of recent research, particularly concerning shifts in housing demand and pricing dynamics. Preceding studies have revealed notable shifts in housing preferences and prices, influenced largely by the pandemic-induced changes in lifestyle and work modalities.

Key studies have documented a migration from urban centers to lower-density suburban areas, attributed to an increased feasibility of working from home (WFH) and a preference for more spacious living environments conducive to home-bound lifestyles (Tsai et al., 2022; Gamber et al., 2022). This shift aligns with broader trends observed during health crises, where housing markets react significantly to such events. Historical parallels, such as the impact of the cholera epidemic on 19th-century London housing prices or the SARS epidemic in Hong Kong, demonstrate long-lasting effects on real estate values following major health crises (Qian et al., 2021). This historical context provides a foundational understanding of the potential impacts of COVID-19 on housing markets.

Despite the extensive research, there remains a gap in comprehensively understanding the specific impact of COVID-19 on housing prices using geographic clustering analysis. While previous studies have focused on the shifts within cities and the general pricing trends, the granularity provided by a geographic clustering approach, particularly at the county level across the U.S., is less explored. This approach is crucial to understanding regional disparities that might be obscured in broader analyses.

The methodological foundation for analyzing the impact of the COVID-19 pandemic on U.S. housing prices draws inspiration from the geographic clustering techniques used in studies like "The Obama Effect". Employing Local Indicators of Spatial Association (LISA) and Moran's I statistic, which were effective in revealing spatial dependencies and patterns in electoral data (Nicholas, 2009), is well-suited for housing market analysis during the pandemic. Moran's I will be instrumental in detecting global spatial autocorrelation, indicating whether housing price changes in one area are linked to changes in neighboring areas, thereby highlighting overarching market trends across the U.S. Simultaneously, LISA allows for a localized examination of specific regions, identifying clusters where housing prices have been notably affected. This dual approach of global and local analysis is essential for a comprehensive understanding of the pandemic's impact on housing prices.

Methods

Data: Housing price data by month of each county in the year of 2019 and 2022 were collected from Zillow (Housing Data - Zillow Research) and were pre-processed in Excel to get the average housing price of the first half of 2019 and the same period of 2022 , representing the situations before and after the Covid-19 Pandemic respectively. The housing price of a total of 3005 counties and county-equivalent areas in the coterminous 48 U.S. states were included. In addition, a shapefile of U.S. county boundaries and a shapefile of U.S. state boundaries were collected from U.S. Census Bureau (Cartographic Boundary Files - Shapefile (census.gov)). The housing price data were then joined with these shapefiles by state-code and municipal-code.

Methods: Spatial join, data visualization, global Moran's I and LISA are used in the analysis.

Limitations: According to the county boundary shapefile from U.S. Census Bureau, there are 3108 counties and county-equivalent areas, so there are 103 counties whose housing data are not available. They are excluded when conducting Moran analysis.

Delimitations: This study focuses on the mainland of the U.S. (Alaska, Hawaii and remote islands are not included for analysis) as the spatial weight matrix is calculated based on contiguity. Multiple data sources are used to get the data as complete as possible.

Results

To begin understanding and describing the geographical clustering patterns in the housing price by county, it's essential to first visually analyze the geographic distribution of average housing price by county in 2019 and 2022 (Fig.1, Fig.2). Counties are color-coded based on the quantile range in which their house prices fall. All housing price data were classified into quintiles. Red counties indicate their housing prices are in the top 40% while blue counties indicate their housing prices fall into the bottom 40%, with darker colors indicating the prices are closer to the top or bottom. Purple areas then indicate those whose housing prices fall into 40^{th}-60^{th} percentile. These maps visually display areas with high, moderate, and low housing prices, which are showing clusters rather than random distribution.

Thus, the global Moran's I statistics were calculated to further reveal the magnitude of the spatial autocorrection in the housing prices. Moran's I statistics were computed using a first-order queen spatial weight matrix, which show significant and strong positive spatial autocorrection in the data for both years (Fig.3, Fig.4). The values of Moran's I are 0.697 and 0.679 for 2019 and 2022, respectively, suggesting housing prices in one area are strongly related to the prices in neighboring areas.

Since a high degree of clustering patterns are observed on the maps and Moran's I values indicate that there is significant spatial autocorrection in the data, the local Moran statistics (LISA) were then calculated to identify local clustering areas. By calculating the local Moran for each county, better Moran scatterplots were created (Fig.5, Fig.6) which can show how the values are similar or dissimilar to the values in surrounding areas. Maps that display significant local clustering patterns were also created (Fig.7, Fig.8). Counties are color-coded according to whether the area contains high values surrounded by high values (High-High), high values surrounded by low values (High-Low), or similarly, Low-Low or Low-High. Grey areas indicate the local Moran of these areas is not statistically significant. Due to missing data from some counties and the spatial weight matrix being calculated based on contiguity, many blanks appear in the map, making it look somewhat fragmented.

The next section focuses on the housing price change from 2019 to 2022, in order to isolate the effects of the Covid-19 and gain greater insight into how specifically the pandemic have influenced the housing market in the U.S. To examine the spatial patterns of the change in housing prices from 2019 to 2022, the growth rate of housing prices between 2019 and 2022 was first calculated and visualized to see its geographic distribution (Fig.9). Instead of random distribution, clear clustering patterns were observed.

Correspondingly, the global Moran's I was calculated for the change of housing prices during the 3-year period (Fig.10). Just like the raw data, the housing price change also demonstrates a high degree of spatial dependence. The value of Moran's I is 0.694, indicating that price change in one county was strongly associated with the changes in neighboring counties. In the same manner, the local Moran statistics was computed to identify local clustering patterns (Fig.11) and a hot-spot and cold-spot map was also created (Fig.12). In the map, the clustering patterns of the price change can be easily observed which is quite different from the patterns observed in the raw data, indicating the potential influence of the pandemic.

Discussion and Conclusion

First of all, from the geographic distribution of housing price in the two years, we can see that the overall pattern of housing prices in the United States has not changed much, with the highest prices in states along the East and West coasts, as well as parts of the Mountain West and Southwest regions, and lower prices in the Central region. There has been a noticeable increase in prices in Western states such as Montana, Idaho, Oregon, and Nevada, and in the Eastern state of Florida. Conversely, in some Southern states like Mississippi, Louisiana, Arkansas, and in the Midwest region, housing prices have remained relatively low or even shown a declining trend.

From the local clustering maps of the two years, we can see that high values are concentrated in areas mentioned above, such as the states on the West Coast, the Mountain West, and the New England and Mid-Atlantic regions of the East Coast. Low values are concentrated in the Rust Belt, the Midwest, and the Deep South. South Florida and the area around Austin, Texas, have seen an increase in housing prices, while southern Louisiana has experienced a significant decline.

We can more easily see the changes when using the growth rate data. From spatial distribution map, it is observed that housing prices have fallen along the Mississippi River areas such as Louisiana, western Mississippi, and Illinois. Additionally, there has been a significant decline in New Mexico, North Dakota, and Wyoming. On the contrary, in the Mountain West as well as Arizona, central Taxes and in the East, including Maine, Tennessee, North Carolina, North Georgia and Florida, there has been a substantial increase in housing prices.

The local clustering map of the change in housing prices from 2019 to 2022 confirms the above conclusions through clearer patterns. However, it can be added that significant low-value clusters have emerged in some major metropolitan areas such as Chicago in the Central region, New York State, Massachusetts, Pennsylvania, Washington D.C. in the East Coast. In California's San Francisco Bay Area, the change in housing prices is also not significant.

From the interpretations above, two major conclusions can be summarized as follows:

(1) Suburban and rural areas are favored: During the pandemic, many jobs shifted to remote working, reducing the demand for living in city centers. This may have led to a slowdown in urban housing price growth, or even a decline, as seen in areas such as Chicago, New York, and Boston. On the other hand, suburban and rural areas typically offer more space for home offices and lower population density, with more room for activities. These conditions became more desirable during the pandemic, and as people moved from cities to these areas, housing prices in suburban and rural regions rose, with significant increases in areas such as the Mountain West, Southwest, and Florida.

(2) Regional economic disparities have intensified: The pandemic aggravated economic disparities between regions, which is reflected in housing prices. Areas whose economies were heavily impacted by the pandemic, such as those reliant on tourism or maritime industries like southern Louisiana, may have experienced significant declines in housing prices. Conversely, areas that were able to adapt well to the pandemic and maintain economic stability saw housing prices rise. Regions surrounding Austin, Texas, with a strong presence in the tech industry, experienced less economic disruption and therefore maintained or increased housing prices.

Works Cited

Barış K. Yörük, Early effects of COVID-19 pandemic-related state policies on housing market activity in the United States, Journal of Housing Economics, Volume 57, 2022, 101857, ISSN 1051-1377, https://doi.org/10.1016/j.jhe.2022.101857.

I-Chun Tsai, Ying-Hui Chiang, Shih-Yuan Lin, Effect of COVID-19 lockdowns on city-center and suburban housing markets: Evidence from Hangzhou, China, Journal of Asian Economics, Volume 83, 2022, 101544, ISSN 1049-0078, https://doi.org/10.1016/j.asieco.2022.101544.

Seabrook, Nicholas R. "The Obama Effect: Patterns of Geographic Clustering in the 2004 and 2008 Presidential Elections" The Forum 7, no. 2 (2009): 0000102202154088841308. https://doi.org/10.2202/1540-8884.1308

S. Sisman, A.C. Aydinoglu, A modelling approach with geographically weighted regression methods for determining geographic variation and influencing factors in housing price: A case in Istanbul, Land Use Policy, Volume 119, 2022, 106183, ISSN 0264-8377, https://doi.org/10.1016/j.landusepol.2022.106183.

William Gamber, James Graham, Anirudh Yadav, Stuck at home: Housing demand during the COVID-19 pandemic, Journal of Housing Economics, Volume 59, Part B, 2023, 101908, ISSN 1051-1377, https://doi.org/10.1016/j.jhe.2022.101908.

Xianhang Qian, Shanyun Qiu, Guangli Zhang, The impact of COVID-19 on housing price: Evidence from China, Finance Research Letters, Volume 43, 2021, 101944, ISSN 1544-6123, https://doi.org/10.1016/j.frl.2021.101944.

Figures and Tables

Fig. 1 Average Housing Price by County, 2019

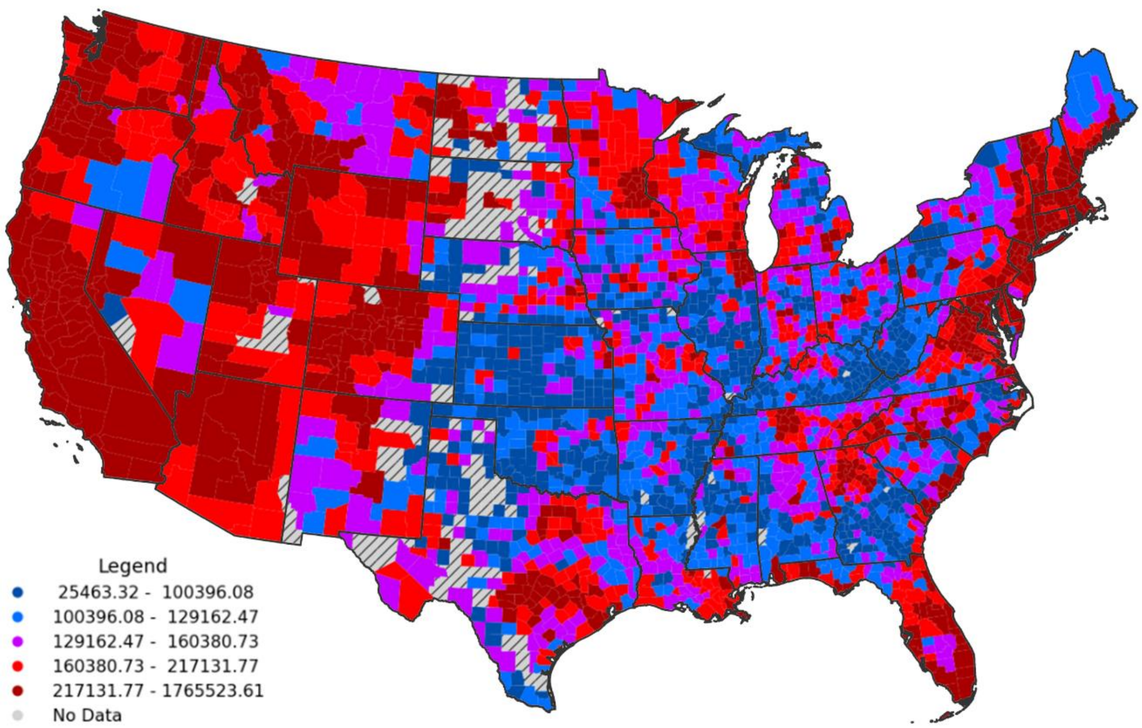


Fig. 2 Average Housing Price by County, 2022

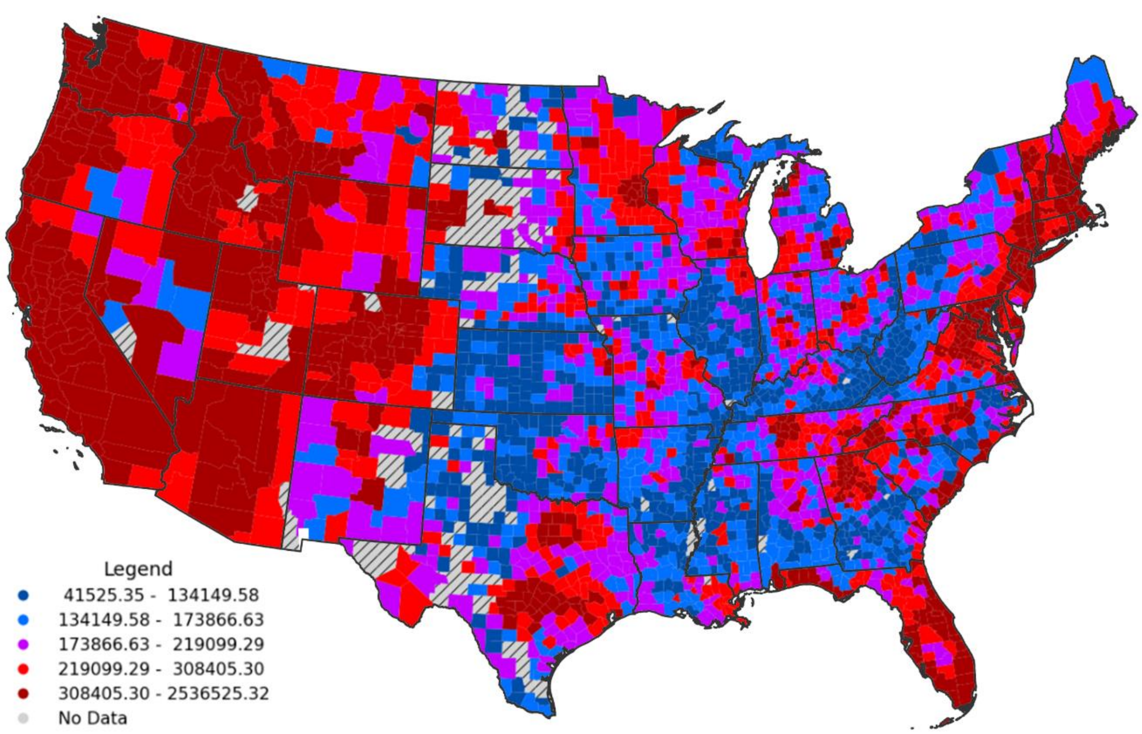


Fig. 3 Moran Scatterplot for Housing Price by County in 2019

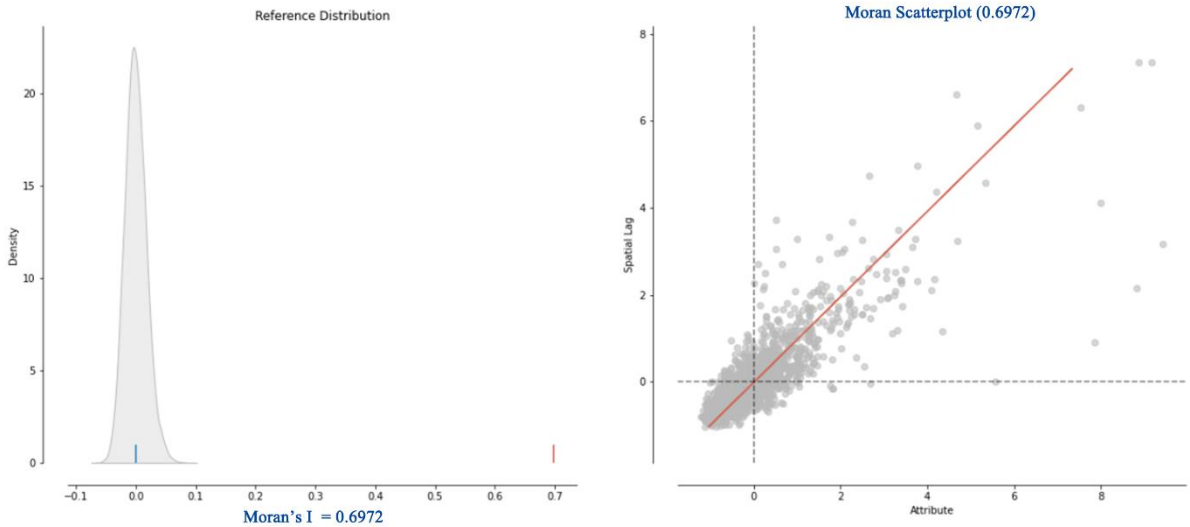


Fig. 4 Moran Scatterplot for Housing Price by County in 2022

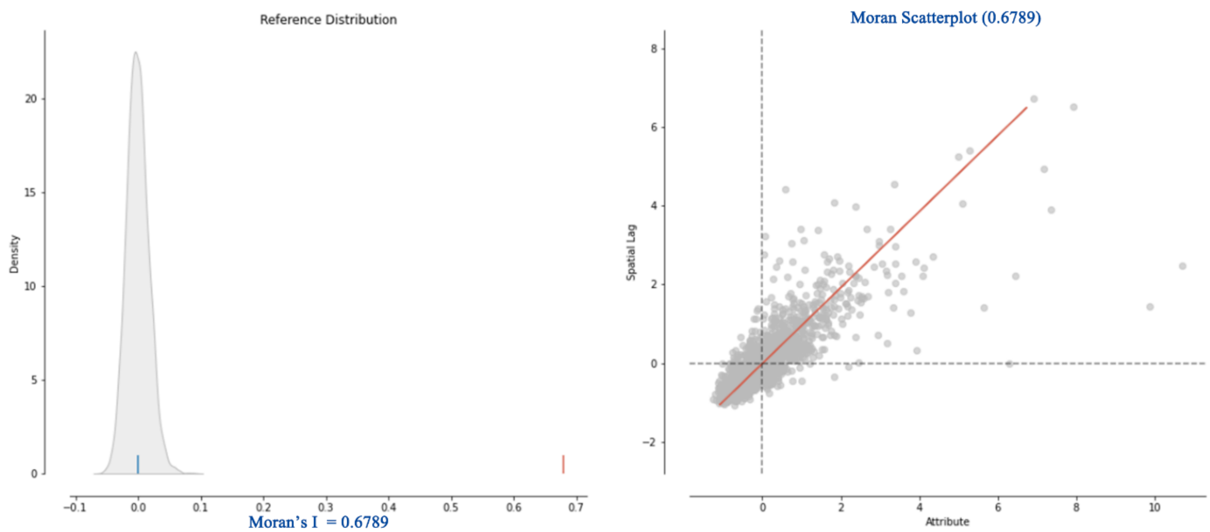


Fig. 5 Local Moran Scatterplot for Housing Price by County in 2019

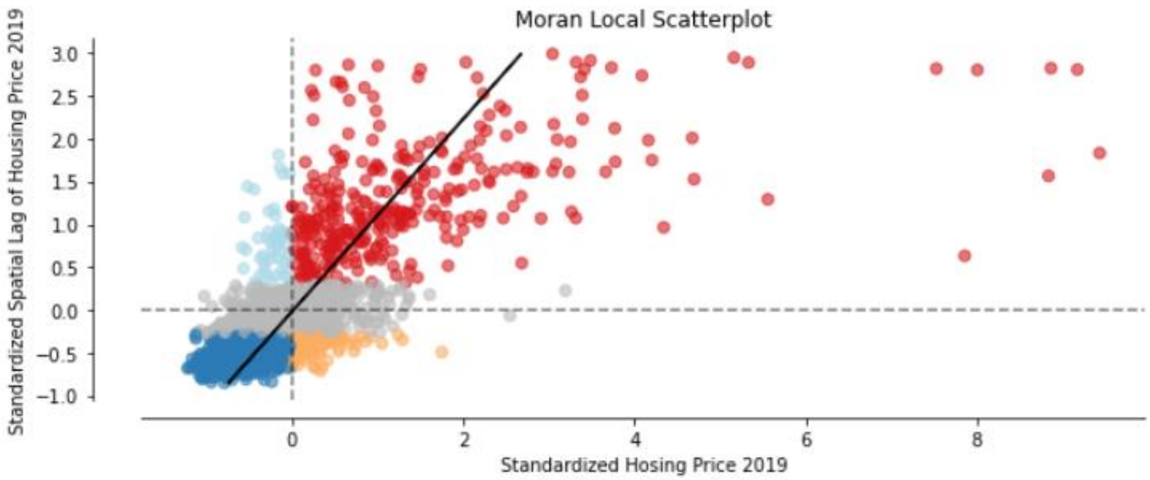


Fig. 6 Local Moran Scatterplot for Housing Price by County in 2022

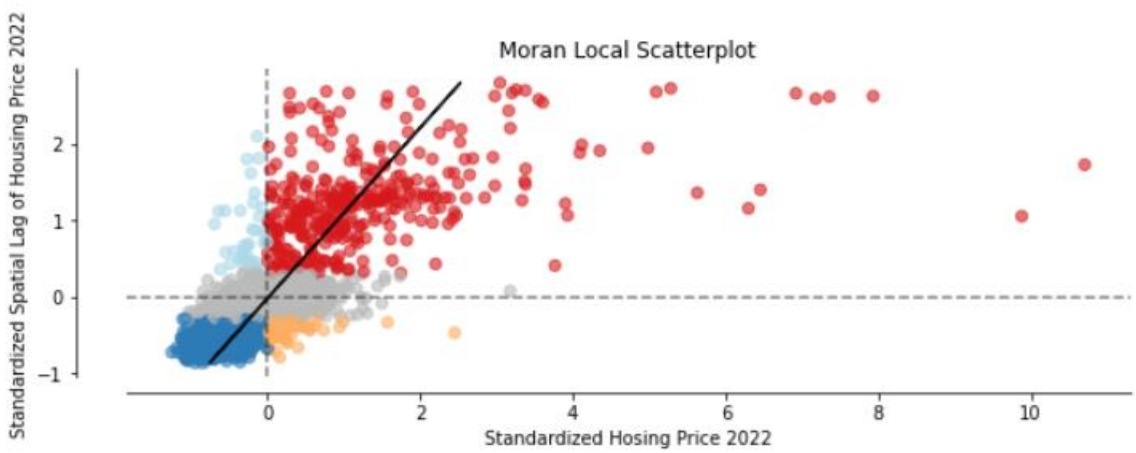


Fig. 7 Geographic Clustering of Housing Price by County, 2019

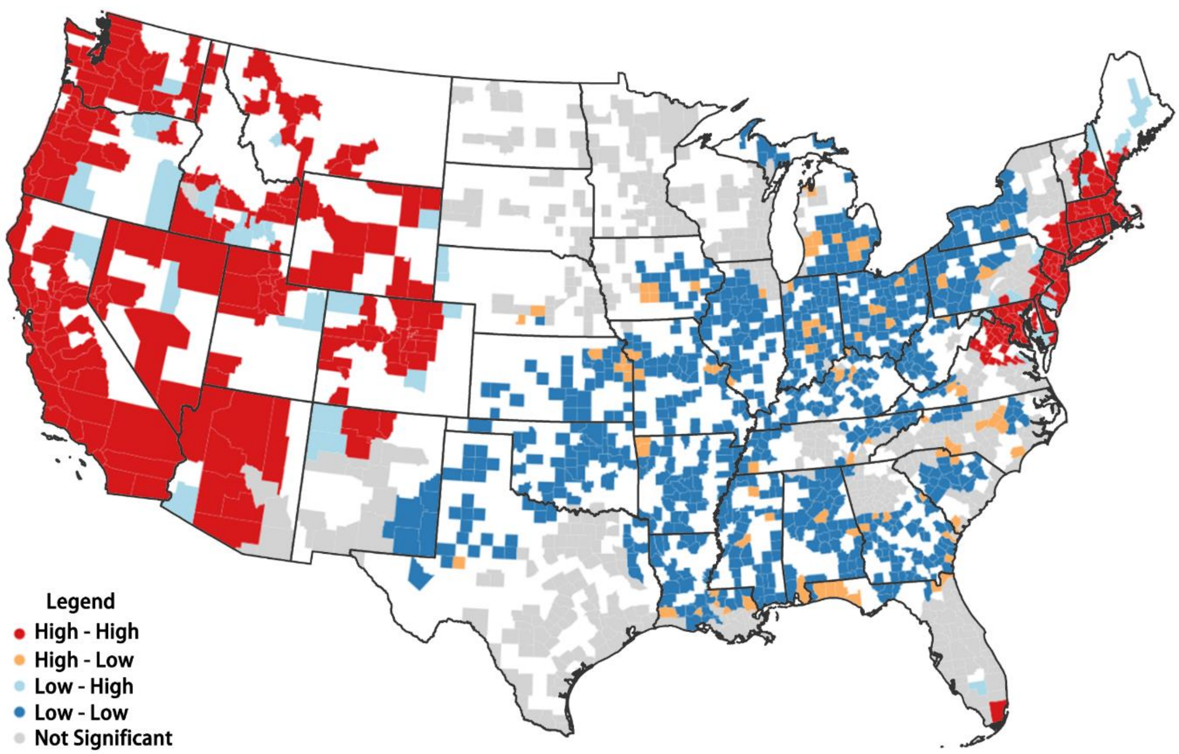


Fig. 8 Geographic Clustering of Housing Price by County, 2022

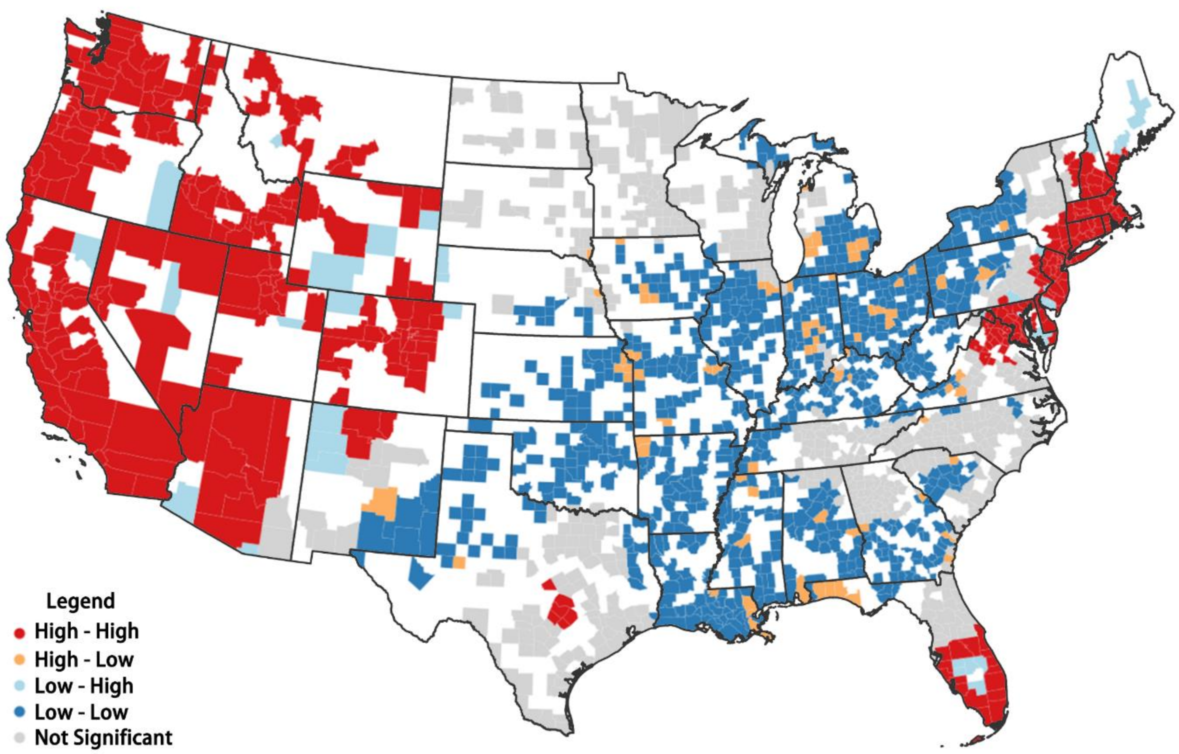


Fig. 9 Change in Housing Price from 2019-2022 by County

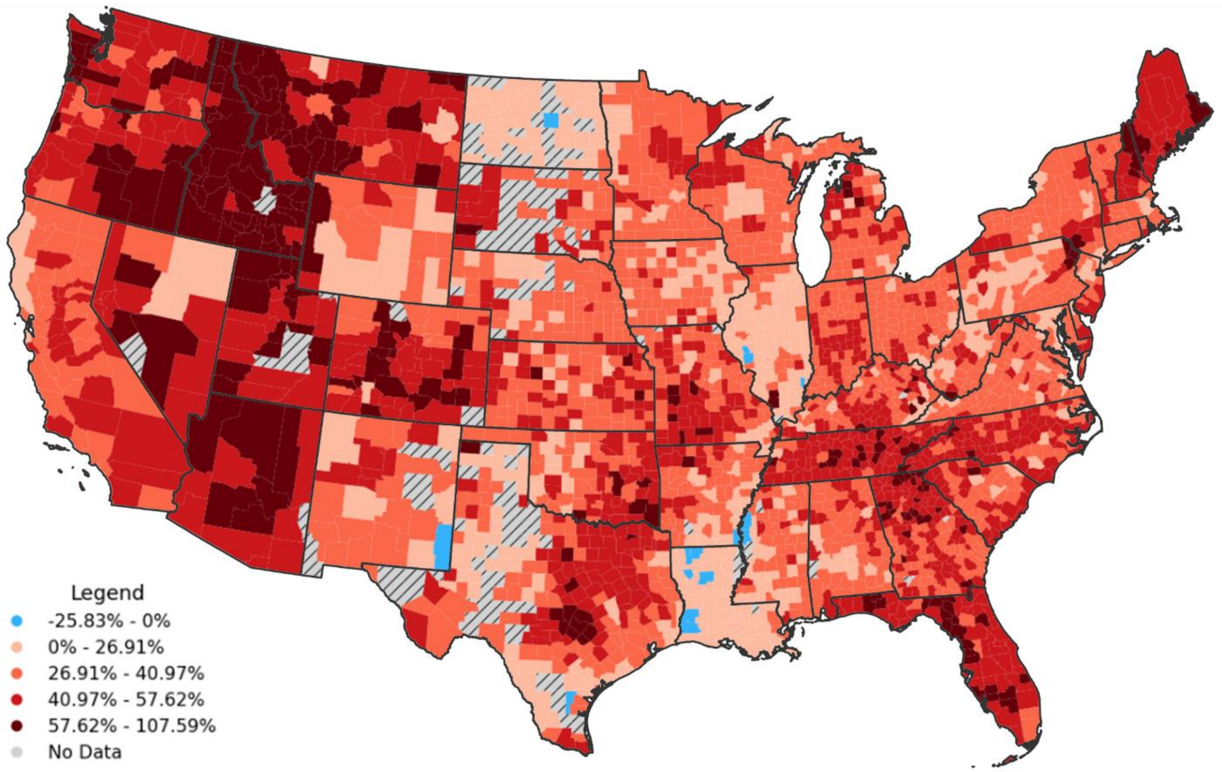
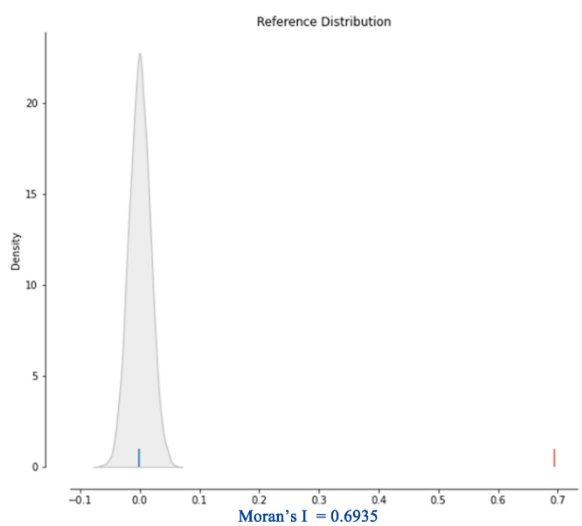


Fig. 10 Moran Scatterplot for the Change in Housing Price by County



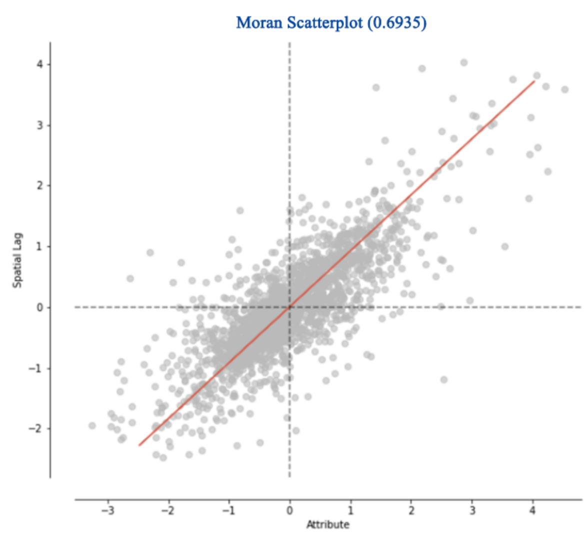


Fig. 11 Local Moran Scatterplot for the Change in Housing Price by County

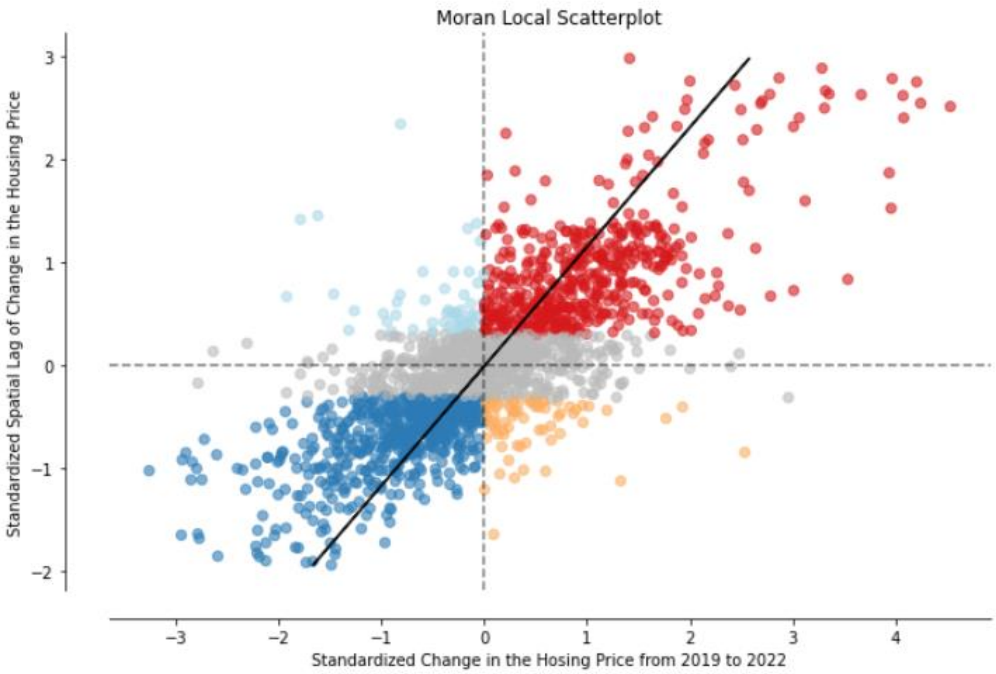


Fig. 12 Geographic Clustering of the Change in Housing Price from 2019-2022 by County