Impact of Population Shifts on the Housing Market

Maria Berg, ShaNiece Twitty

Department of Mathematics and Statistics, Northern Arizona University, Flagstaff, AZ 86001

NAU NORTHERN ARIZONA UNIVERSITY — Office of Undergraduate Research and Creative Activity

NAU NORTHERN ARIZONA UNIVERSITY — Department of Mathematics and Statistics

Abstract

This project analyzed the growth and decline of the housing market in five major U.S. cities—Boston, Chicago, Los Angeles, Philadelphia, and New York—between the years 2012 and 2023, in conjunction with population data. Housing data was collected from Redfin, and corresponding population statistics were obtained from the United States Census Bureau. The analysis identified trends and relationships between housing market activity and population changes by examining factors such as home close-to-list sales, market inventory, and demographic shifts. The findings provided insights into the broader dynamics that influenced the housing markets in these cities over time and revealed how certain demographic changes affected housing affordability and availability.

Introduction

For our project, we collected datasets focused on housing market trends, specifically analyzing the impact of population growth or decline on key real estate metrics. The datasets was used to identify housing indicators and population demographics from all 5 regions. Our main metric for housing was the close-to-list percentage and the discrepancies found within each regions housing market performance. This data allows one to understand housing market dynamics in relation to population shifts and can aid in predicting market trends and guiding real estate investments. Analyzing how housing prices and sales activity responded to population changes, we identified patterns that indicated whether a region was experiencing growth, stability, or decline.

Key Questions

- How did racial demographic shifts effect housing sales?
- What is close-to-list percentage
  - what does it tell us about the housing market?
- Which racial demographic indicators showed the strongest correlation with close-to-list price ratios?

Data Breakdown

Monthly housing metrics from Redfin were joined with annual demographic estimates from the U.S. Census Bureau's 1-Year ACS, spanning 2012–2023. Collected variables are described in **Table 1.**
- CSV files were directly downloaded from these sources.
- The 2020 Census ACS is excluded due to unreliable COVID-19 disruptions.
- Census data was collected by region. This was used to determine racial demographics in the area.
- Housing data metrics are recorded month to month.

Table 1. Key variable used in analysis and visualizations.

Key Variable Name	Description
Region	Name of metro area.
Average Sale To List	Regional average ratio of sale to listing price.
Month of Period End	Month and Year stored as 'MM-YY' date format.
Est Total Population	Estimated total population.
Pct Hispanic or Latino	Percentage of a Region identifying as Hispanic or Latino.
Pct White	Percentage of a Region identifying as White alone.
Pct African American	Percentage of a Region identifying as Black or African American.
Pct Asian	Percentage of a Region identifying as Asian alone.

Methods

- Data handling and analysis was done using the R statistical language and RStudio.
- We used packages like those in the tidyverse—such as dplyr for data manipulation and tidyr for cleaning—to help clean and organize our data efficiently.
- Separate CSV files of housing and census data with frames were joined by region and year.
- R package was created to host our cleaned data and linked to GitHub for easy collaboration.
- Exploratory data analysis (EDA) focused on overall population, ethnicity, housing inventory, and close-to-list price ratio.
- Time-dependent trends were explored to analyze changes across years and detect patterns in housing and demographic variables.
- Linear modeling of ethnicity and regional close-to-list price.
  - Assessed univariate and multivariate relationships using combinations of additive and interactive effects.
  - Model assessment using residual analysis.
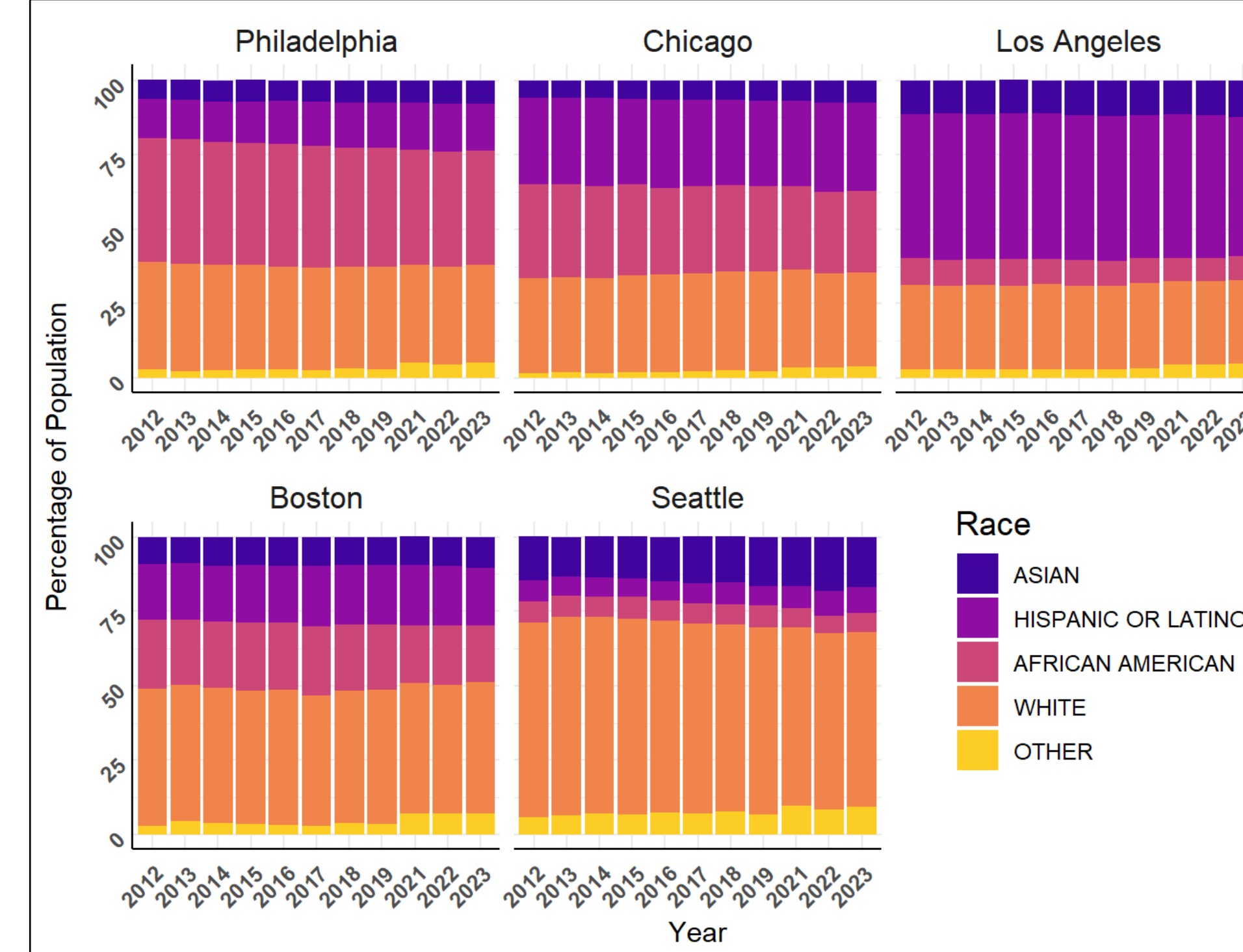
Racial Demographic Shifts

Figure 1. Mosaic plot illustrating the percent of regional population stratified into 5 racial groupings.
- White populations makes up the largest racial group in each city.
- Hispanic or Latino populations are growing in all cities.
- Seattle has the highest proportion of White population, while also showing the greatest increase in other racial demographics.

Project's GitHub

Close-to-List (%) by Demographic

Figure 3: African American Population vs. Close-to-List (%) by Region
- Regions with higher African American populations tend to close at or below listing price
- Trend holds both across and within regions.

Figure 4: Asian Population vs. Close-to-List (%) by Region
- Regions with higher Asian populations tend to close at or above listing price
- Trend holds both across and within regions.

Figure 5: Hispanic or Latino Population vs. Close-to-List (%) by Region
- Regions with lower Hispanic or Latino populations tend to close below list price, though the overall trend within most regions is upward.
- Los Angeles stands out, showing the opposite pattern compared to other regions.

Figure 6: White Population vs. Close-to-List (%) by Region
- Regions with higher White populations tend to close at higher close-to-list ratios.
- Trend holds across regions, despite some having negative within-region slopes.
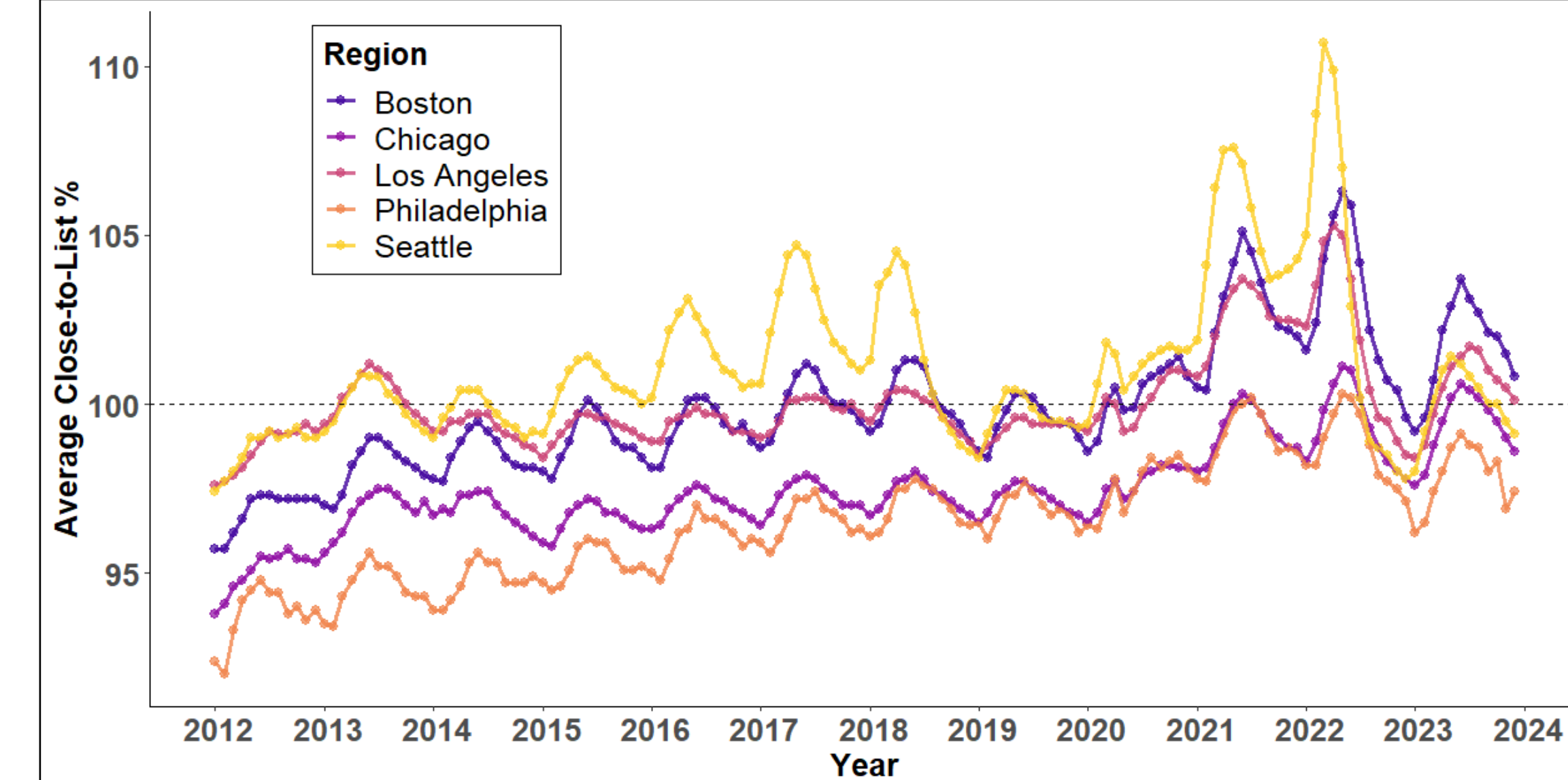
Close-to-List (%)

Figure 2. Scatter of housing sale discrepancy over time grouped by region.
- Seattle consistently demonstrated periods above asking price.
- Philadelphia and Chicago typically trended below asking price.
- All markets showing a notable payment surge during the pandemic.

Conclusion

Our analysis across five major U.S. cities reveals that racial demographics significantly influence close-to-list price percentages. Initial visualizations of demographic trends (**Figure 1**) and close performance by region (**Figure 2**) laid the groundwork. Additive and interaction models were tested, and the interaction model was found to be significant for all races.
- In Boston (our model's reference city), higher Asian population percentage had significantly higher close-to-list percentages (**Figure 4**). This effect weakens in Seattle and Los Angeles, indicating regional variation.
- Inversely, higher African American population percentages correlated with lower close-to-list percentages (**Figure 3**), with stronger negative effects in Seattle and L.A.
- For the Hispanic or Latino population, Boston showed a strong positive slope (**Figure 5**), while Los Angeles displayed a negative interaction, further highlighting how regional context shapes these dynamics.
- White populations show a negative slope within regions but inversely displaying an overall positive trend (**Figure 6**).

Finally, we observed strong negative correlations between racial group proportions, particularly between Asian and African American populations, suggesting persistent racial proportion shifts across regions. Demographic distribution patterns may be reinforcing disparities in housing access and outcomes.

Acknowledgment

References

- Redfin. (n.d.). *Downloadable Housing Market Data.* Redfin News. https://www.redfin.com/news/data-center/
- U.S. Census Bureau. (n.d.). *DP05: ACS Demographic and Housing Estimates.* data.census.gov. https://data.census.gov/table?q=DP05:+ACS+Demographic+and+Housing+Estimates&g=1 60XX00US0644000,1714000,2507000,4260000,5363000