Project Report: Housing Affordability in Johnson County

Teague Stotlar

Github Link

Johnson County, Iowa is home to Iowa City and the University of Iowa. The county now faces a serious housing affordability problem. As of December 2024, the median home price in Johnson County was $334,812, which represents a 3 % increase over the previous year . By comparison, the statewide median in Iowa was only $226,900 .

A report from the National Low Income Housing Coalition found that in 2023 there were just 18 affordable units for every 100 low-income renters in Johnson County . When households spend more than 30 % of their income on housing, they are considered cost-burdened. Cost-burdened families often must cut back on food, healthcare, or other essentials . In Iowa as a whole about 23.6 % of households are cost-burdened, but Johnson County's higher prices and large student population make the problem worse (Pew Research Center).

In this project I examine home-sale prices and household incomes in Johnson County for the years 2013, 2018, and 2023. I ask three questions:

1. How have sale prices changed over this decade, and what patterns appear by property characteristics?

2. How do trends in median household income compare with changes in home prices?

3. What share of households spends more than 30 % of their income on housing, and how has that proportion shifted?

To answer these questions I combine parcel-level sales data from the Iowa City Assessor's website with income and cost-burden figures from the U.S. Census Bureau's American Community Survey.

Data:

I began by scraping residential property sales from the Iowa City Assessor's website for the years 2013, 2018, and 2023. I limited my scrape to single-family, owner-occupied homes sold through standard arm's-length transactions. During parsing I filtered out any records labeled as foreclosures or family transfers. Each annual pull resulted in a raw CSV: historical_transactions_2013.csv (595 records), historical_transactions_2018.csv (655 records), and historical_transactions_2023.csv (535 records).

Next, I performed an initial cleaning of each year's CSV in a Jupyter notebook (DataScrape and Integration.ipynb). I standardized column names (for example converting "Sale Price" to "Sale Amount" and stripped non-numeric characters from currency and square-footage fields. I also added a new column, Year, to each record so I could keep track of its origin when all three files were combined. Once each file passed basic quality checks (no missing parcel numbers, sale dates in the correct range, and sale amounts above zero) I concatenated them into iowa_city_housing_combined.csv.

To add socioeconomic context I downloaded American Community Survey 5-year estimates from data.census.gov for Johnson County. I retrieved Table B19013 for median household income and Table B25106 for the number and proportion of households spending more than 30 % of income on housing, for the same three years. I saved these county-level estimates in ACS_JohnsonCo_combined.csv. During a second cleaning pass I converted all percentage columns to floating-point fractions (e.g., 0.3818 for 38.18 %) and ensured the Year field matched exactly the sales data.

With both files ready, I performed a horizontal merge in my integration notebook. I joined iowa_city_housing_combined.csv to ACS_JohnsonCo_combined.csv on the Year column so that each sale record carried its county's median income and cost-burden proportion for that year. The merge was one-to-many from ACS (one row per year) onto each of the thousands of sale records. The resulting master file (master_housing_plus_ACS.csv), contains detailed property-level attributes alongside the matching county-year income and affordability metrics.

Throughout this process I logged my steps and checks in DataScrape and Integration.ipynb. I also used it to spot anomalies such as outlier sale prices above $1.5 million and to validate that each year's record count matched the original assessor lists.

Data Dictionary:

Field	Type	Description
Parcel Number	Integer	County parcel identifier
Sale Date	Date	Date of property sale
Sale Amount	Float	Sale price in U.S. dollars
Address	String	Full street address
Style	String	Architectural style (e.g., Ranch, 2-Story)
Year Built	Integer	Year the structure was built
Total SF	Integer	Total residential square footage
Lot Area	Float	Lot size in square feet
Appraised Value	Float	County assessed value in U.S. dollars
Recording	String	County recording reference
Year	Integer	Transaction year (2013, 2018, or 2023)
Geography	String	FIPS code for Johnson County ("0500000US19")
Geographic Area Name	String	"Johnson County, Iowa"
Total_Units	Integer	Total occupied housing units (from ACS)
Cost_Burden_Units	Integer	Households spending > 30 % on housing (from ACS)
Cost_Burden_Proportion	Float	Fraction of households spending > 30 % on housing (from ACS)
Median_Income	Integer	Median household income in U.S. dollars (from ACS)

Data Sources:

US. Census: Table B25106: "Tenure by Housing Costs as a Percentage of Household Income in the Past 12 Months" , Table B19013: "Median Household Income in the Past 12 Months"

Johnson County Real Estate: https://iowacity.iowaassessors.com/showResSaleSearch.php?

Analysis: (Visualizations (1).ipynb)

Changes in Sale-Price Distributions (Figure 1 & Figure 2)

To understand how Johnson County home-sale prices evolved, I first plotted the full price distributions for 2013, 2018, and 2023 on a single histogram (Figure 1). In 2013, the bulk of sales clustered between $150 000 and $250 000, with only a handful above $500 000. By 2018, that bulk shifted upward to roughly $200 000–$300 000, and by 2023 the densest band moved again to $250 000–$400 000. The long tails in 2018 and 2023 show more frequent high-end transactions, indicating that the luxury segment of the market gained steam.

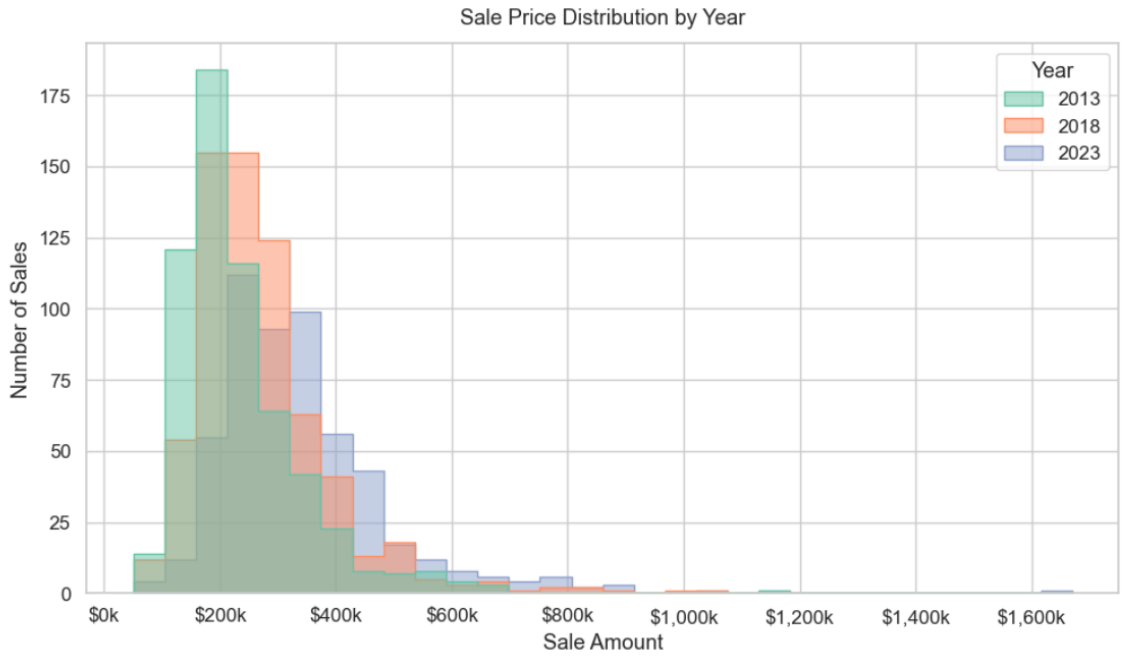

Figure 1: Changes in Sale Distributions

To highlight central tendencies and dispersion, I created side-by-side boxplots for 2013 and 2023 (Figure 2). The median sale price increased from $203 000 in 2013 to $310 000 in 2023—an

absolute jump of $107 000, or +52.7 %. The interquartile range also widened, from about $175 000–$248 000 in 2013 to $245 000–$380 000 in 2023, signaling growing variability. Finally, a Welch's t-test ($t = 13.35$, $p < 0.001$) confirms that the upward shift is statistically significant and not due to chance. Taken together, these plots demonstrate a clear increase in both the typical and high-end segments of the market over the past decade.

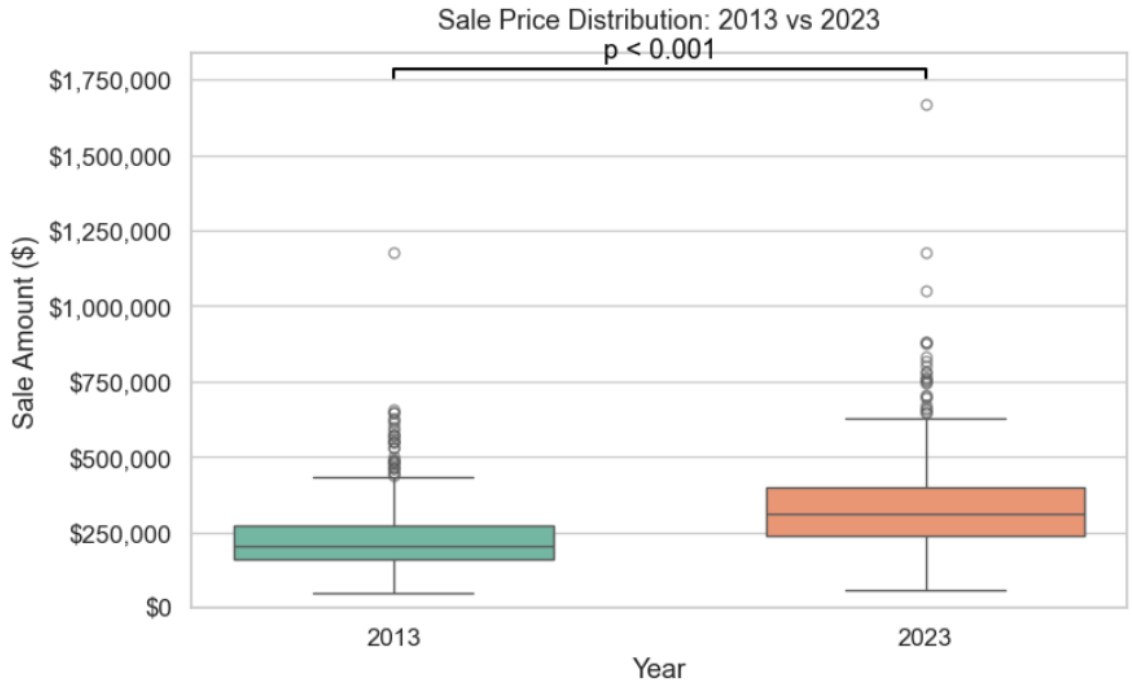

Figure 2: Changes in Sale Distributions

Price vs. Income Trends (Figure 3)

Next, I compared the county's median sale price with its median household income over the same period (Figure 3). From 2013 to 2023, median sale price rose from $203 000 to $310 000 (+52.7 %), while median income grew from $53 424 to $74 721 (+39.8 %). Although both series climbed steadily, home-price growth outpaced income growth by roughly 13 percentage points. This divergence implies that housing became less affordable: families saw their earnings rise, but not fast enough to keep pace with the market, pushing more households toward cost-burdened status. The parallel lines and percentage annotation on the chart make this widening gap immediately apparent.

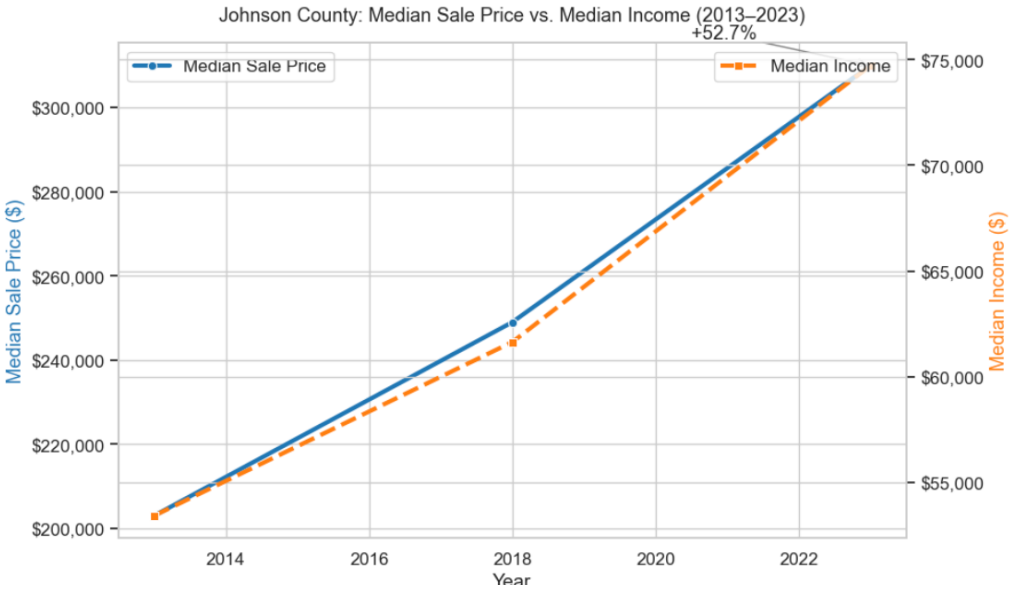

Figure 3: Price vs. Income Trends

Trends in Cost-Burdened Households (Figure 4)

I then examined how these price and income shifts translated into financial strain, measured as the share of households spending more than 30 % of income on housing. In 2013, 38.2 % of Johnson County households were cost-burdened. That share fell to 35.8 % in 2018, likely reflecting a temporary convergence between income gains and price growth. However, by 2023 the cost-burden share rebounded to 38.2 % (Figure 4). In effect, any relief experienced midway through the decade was erased as home prices again surged faster than incomes. This U-shaped pattern underscores the fragility of affordability improvements and suggests that long-term relief requires incomes to grow at least as fast as housing costs.

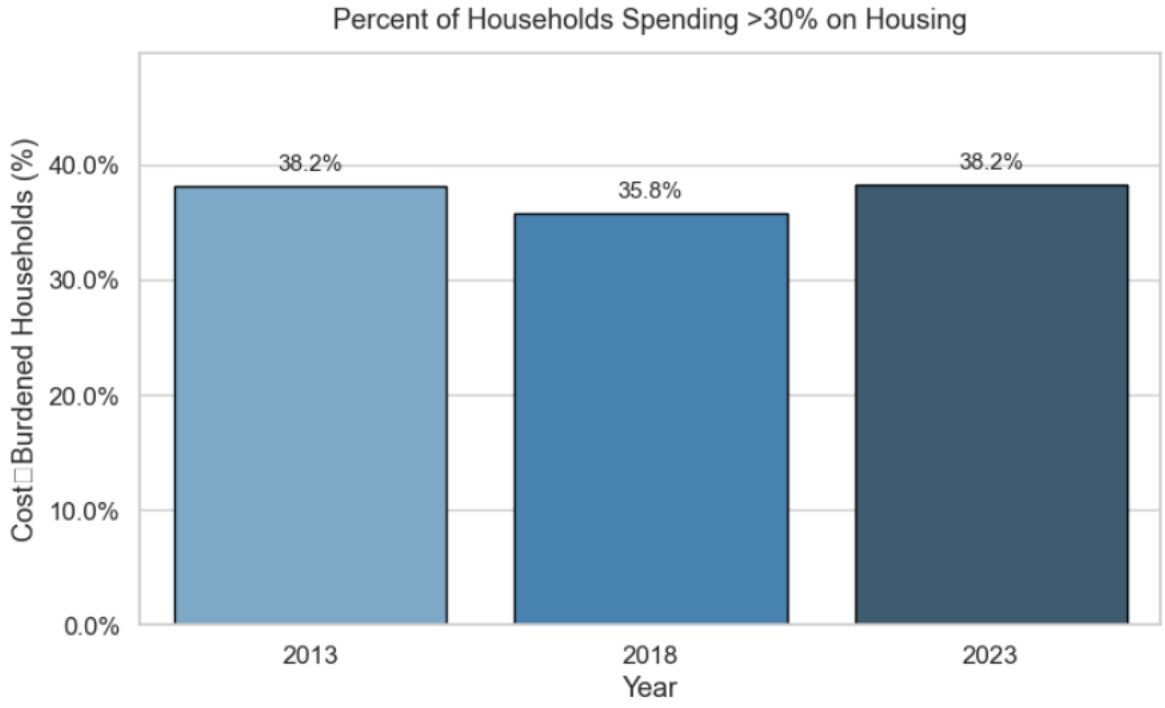

Figure 4: Trends in Cost-Burdened Households

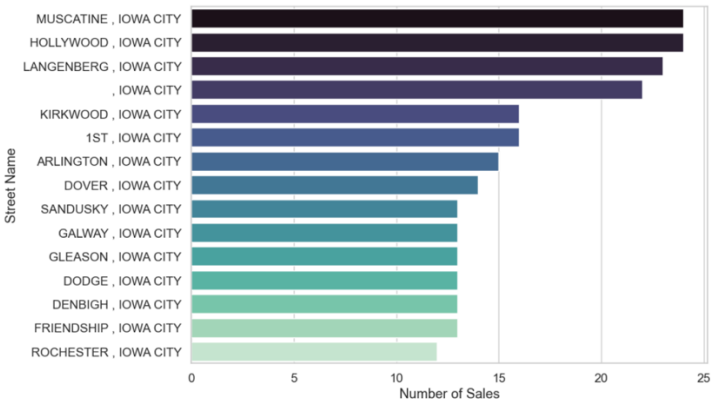

Figure 5: Top 15 Streets by sales, Bar Chart

Figure 6: Word Cloud

I pulled street names from each address and counted the top 15 by number of sales. In Figure 5, Muscatine Avenue and Hollywood Drive each had 24 sales, Langenberg Road had 22, and 21 records had no street name after parsing. Kirkwood Avenue and 1st Avenue each had 16 sales, and Arlington Drive and Dover Street each had 15. Figure 6 shows a word cloud of these top streets. Larger names like Muscatine, Hollywood, and Langenberg indicate higher sales volume. Smaller names like Sandusky, Galway, Gleason, Dodge, Denbigh, and Friendship show moderate activity. These results suggest that home sales are spread across several main roads rather than focused in one area. This simple street-level view can help planners target areas with the most market activity.

Predictive Modeling of Sale Prices and High-End Sales

Finally, I tested whether basic property attributes like square footage, lot area, assessor's appraised value, and year built— ould predict sale prices or flag above-median transactions. On a held-out 20 % test set, three regression models achieved nearly identical performance ($R^2 \approx 0.742$ for Linear, Ridge, and Lasso). This indicates that 74 % of the variance in sale prices can be explained by those four features alone. Coefficient inspection shows that newer homes (Year Built) carry the highest positive weight (≈ +53 $/year built), followed by appraised value (+0.76 $/appraised dollar), lot area (+0.95 $/sq. ft.), and tightly controlled square-footage effects (≈ –11.98 $/sq. ft.).

For classification(predicting whether a sale exceeded the overall median price) the models performed as follows:

Logistic Regression: 84.6 % accuracy, F1 = 0.848

Decision Tree: 82.6 % accuracy, F1 = 0.830
KNN (k=5): 82.1 % accuracy, F1 = 0.827
SVM (RBF): 84.6 % accuracy, F1 = 0.847

These results demonstrate that even a small set of structural and assessment features can distinguish high-end from the lower-end sales. The close performance between Logistic and SVM suggests a roughly linear decision boundary in feature space, and the tree and KNN reveal that non-linear interactions add only modest gains.

Conclusion:

This project set out to answer three questions about housing affordability in Johnson County from 2013 to 2023.

1. How have sale prices changed, and what patterns appear by property characteristics?
Sale prices have risen sharply, from a median of $203 000 in 2013 to $310 000 in 2023 (+52.7 %). Both the central range and the upper tail of the distribution shifted upward which reflects growth across the market and includes high-end segments. In our predictive models, appraised value, square footage, lot area, and year built explained about 74 % of the variance in sale price, showing these characteristics remain strong drivers of home values.

2. How do trends in median household income compare with changes in home prices?
Median household income rose from $53 424 to $74 721 (+39.8 %) over the same period. Although incomes grew steadily, home-price growth outpaced income by roughly 13 percentage points. This widening gap signals that housing is becoming less affordable, as earnings have not kept pace with market prices.

3. What share of households spends more than 30 % of their income on housing, and how has that shifted?
The share of cost-burdened households was 38.2 % in 2013, dipped to 35.8 % in 2018, and returned to 38.2 % in 2023. Any temporary relief in the middle of the decade was erased as

prices again outstripped incomes. This U-shaped pattern shows the fragility of affordability improvements and suggests that long term relief will require incomes to accelerate along with housing costs.

Overall, Johnson County has seen strong home-price growth that exceeds income gains, resulting in a stable but high rate of cost burden. Future work could explore neighborhood-level differences, incorporate more socioeconomic variables, or model short-term price forecasts to guide policy and planning.

Works Cited

"Johnson County, Iowa Housing Market Report December 2024." Rocket Homes, 2024, https://www.rockethomes.com/real-estate-trends/ia/johnson-county

"Iowa Housing Market: Home Prices & Trends." Houzeo, 2024, https://www.houzeo.com/blog/iowa-real-estate-market/

"Johnson County seeks proposals for affordable housing project." The Gazette, February 20, 2025, https://www.thegazette.com/news/johnson-county-seeks-proposals-for-affordable-housing-project/

"Nearly Half of Renter Households Are Cost-Burdened, Proportions Differ by Race." U.S. Census Bureau, September 12, 2024, https://www.census.gov/newsroom/press-releases/2024/renter-households-cost-burdened-race.html

"The state of affordable housing in the US." Pew Research Center, October 25, 2024, https://www.pewresearch.org/short-reads/2024/10/25/a-look-at-the-state-of-affordable-housing-in-the-us/