Housing Affordability Vs Income

Patrick Chen

George Mason University, Fairfax, VA, USA
pchen12@gmu.edu

Abstract

The topic of this research project is housing prices vs affordability. The reason why this is an important topic is because we are becoming increasingly more aware about people not being able to afford buying a home. The research showcases that median housing prices have increased while people have not been able to produce capital gain.

Introduction

The topic focuses on housing affordability and comparison of home prices to capital gain made by individuals.

The reason why this topic is important to know is because our standard of living in the United States has increased significantly. According to research by Susan Moritz, housing prices have increased by 30 percent (Moritz n.d.).

Below are the two research questions.

Q1: Do housing prices vary by location.

Q2: What is the average capital gain?

The following software R studio and Microsoft Excel was used to generate data visualizations.

Related Work

According to researcher Susan Moritz, one out of three American households spend 30 percent of income on housing, and one out of seven spend up to 50 percent of income on housing. Workers are more and more not able to afford housing in the United States (Moritz n.d). Michelle Delgado reports that housing prices have increased significantly and exponentially more than income. Income has not changed at all, and this has marked a major problem in our housing market (Delgado 2021).

Data

Three datasets were used to analyze the research questions.

Adult Income Dataset

The first dataset was retrieved from Kaggle, with the source being from toronto.edu website Delve. This dataset is often used for predictions and modeling. The size is around 5.33 mb. Various columns show the working class, education obtained, marriage status, occupation status, relationship status, native country, income, and finally capital gain. Capital gain does not include income, but other sources of monetary gain not generated from work.

Housing Price Dataset Across 50 States

The data was obtained from Kaggle also, with the data source being from Zillow. The columns show the names of the states in the United States, the median price of houses, and dates from which the housing prices were listed. Only the earliest housing prices which is year 2000 and the latest date marked 2022 January are kept for comparison.

Survey of More Than 900 Homes

The data was obtained from Kaggle also from Zillow showing more than 900 homes priced in various locations of the United States. There are also dates included from the year 2000 up to the year 2022.

Preprocessing

Here are the steps to preprocessing.

1. Data was downloaded and obtained from Kaggle.
2. Home prices were filtered to show the earliest and latest dates.

Figure 2: Workflow for processing data

Methodology

The datasets were processed using R studio for the choropleth maps, and Excel was used to calculate the summary statistics.

Histogram Analysis

A histogram analysis is conducted on the capital gains of adults as well as the median housing prices of the more than 900 homes surveyed.
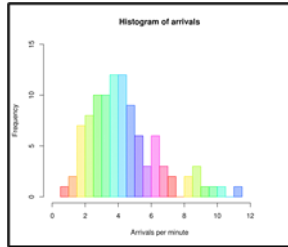

Figure 1: Histogram

Summary Statistics In Excel

Summary statistics is calculated for the prices of the 900 homes in the dataset.
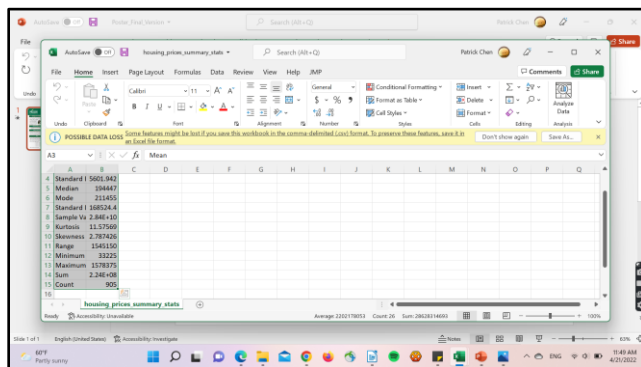

Figure 2: Microsoft Excel

Choropleth Map

Choropleth maps are generated using median housing prices across the 50 states.

3-D Scatterplot

3D scatterplots are used to show age, capital gain, and hours worked per week.

Results

The results here show a few visualizations processed from R studio.

Results subsection 1

A Choropleth map of housing prices across the 50 states in the year 2000 is shown. California is the most expensive state to live in. In the legend, the maximum price is at 220,000 dollars.
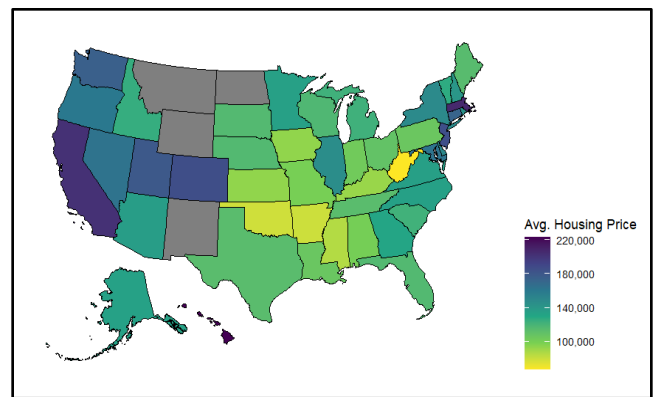

Figure 3: Housing Prices In Year 2000

The housing prices for year 2022 are shown in the second choropleth map. The maximum housing price is now at 800,000 dollars, with the minimum price being at 220,000 dollars.
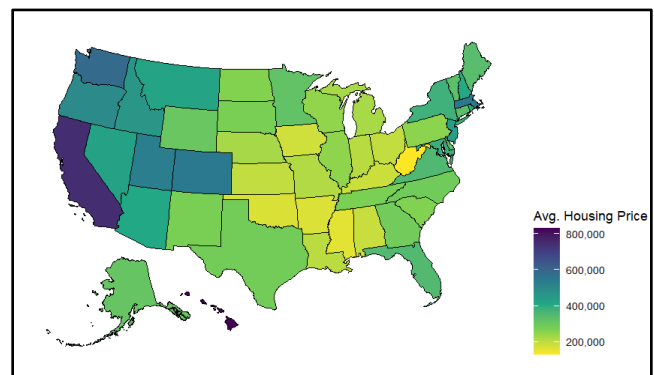

Figure 4: Housing Prices 2022

Results subsection 2

This histogram shows the amount of capital gain made by the individuals in the dataset. Most of the people have 0 capital gain.
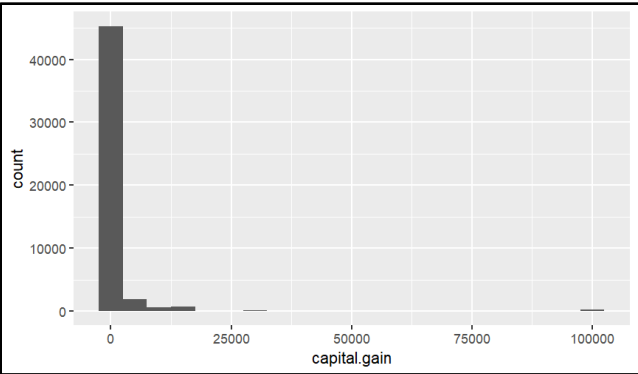

Figure 5: Capital Gain

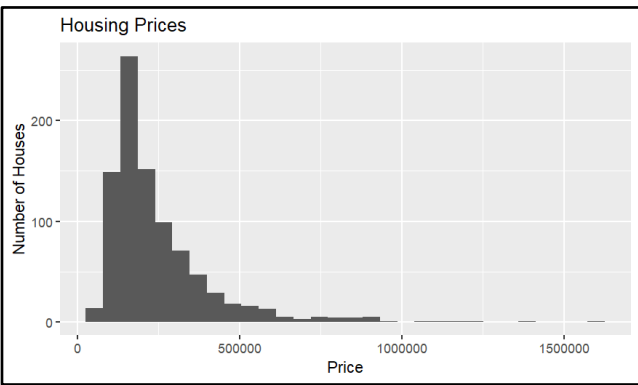Most of the following home prices are around the 200,000 dollar price range.

Figure 6: Distribution of Median Home Prices in Year 2022

Results subsection 3

A 3D scatterplot is generated to show the age, capital gain, and number of hours worked. Regardless of work hours being produced at any age, most people still have a capital gain of zero.
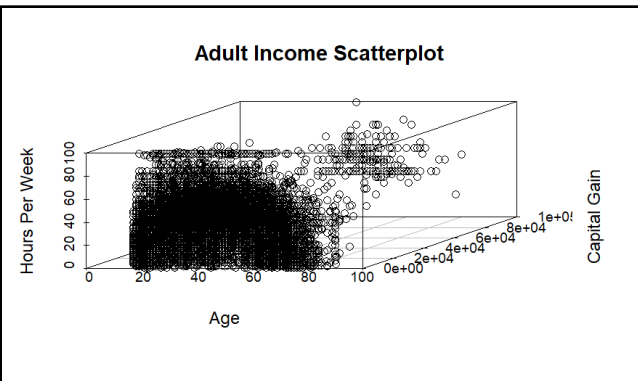

Figure 8: 3D scatterplot

Results subsection 4

Here are the summary statistics produced from Microsoft Excel over the home prices dataset.

Mean	247338.3
Standard Error	5601.942
Median	194447
Mode	211455
Standard Deviation	168524.4
Sample Variance	2.84E+10
Kurtosis	11.57569
Skewness	2.787426
Range	1545150
Minimum	33225
Maximum	1578375
Sum	2.24E+08
Count	905

Discussion

The results, based on the choropleth maps, indicate that housing prices have increased significantly. According to the 3d scatterplot on adult income and the histogram on capital gain, most people make zero amount of money. It can be inferred that people are living paycheck to paycheck.

Conclusion

It can be concluded that housing prices are becoming more and more unaffordable for most Americans. Analysis of demographics can also be useful to provide more insight on what the data holds. Because income has not increased to adjust for inflation as well as rising prices, it is becom-

ing a reminder that we must make change in terms of how to make housing affordable for everyone.

References

Moritz, S. (n.d.). *Affordable Housing Research Report | Pace Law School*. Pace University. Retrieved March 18, 2022, from https://law.pace.edu/affordable-housing-research-report

Delgado, M. (2021, October 22). *U.S. House Prices Are Rising Exponentially Faster Than Income (2021 Data)*. Real Estate Witch. Retrieved March 18, 2022, from https://www.realestatewitch.com/house-price-to-income-ratio-2021/#:%7E:text=From%202019%20to%202021%2C%20the,person%20earns%20in%20one%20year.

Adult income dataset. (2016, October 6). Kaggle. Retrieved March 3, 2022, from
https://www.kaggle.com/wenruliu/adult-income-dataset

Delve adult dataset. (1996, October 8). Adult Dataset. Retrieved March 3, 2022, from
http://www.cs.toronto.edu/%7Edelve/data/adult/desc.html

Housing data. (2021, March 25). Zillow Research. Retrieved March 3, 2022, from
https://www.zillow.com/research/data/

Mooney, P. (2020, December 8). *Zillow House Price Data*. Kaggle. Retrieved March 3, 2022, from
https://www.kaggle.com/paultimothymooney/zillow-house-price-data