

○○○

CSCI 4341 BIG DATA ANALYTICS

MUHAMMAD AFIQ HUSYAIRI
20187691

UNVEILING HOUSING DYNAMICS: A COMPREHENSIVE ANALYSIS OF FACTORS INFLUENCING PROPERTY VALUES, OWNERSHIP, AND REGIONAL PATTERNS IN THE AMERICAN HOUSING MARKET

○○○

TABLE OF CONTENTS

- Overall Summary
- Objectives
- Data Science Questions
- Methodology
- Analysis
- Conclusion



OVERALL SUMMARY

The research aimed to analyze various aspects of housing dynamics using the American Housing Survey dataset spanning from 1985 to 2009. By exploring this dataset, we sought to uncover insights into factors influencing housing choices, property values, and the relationship between income and housing decisions. The study focused on key variables such as income, region, structure type, adequacy index, age, household size, and more. Through data visualization and analysis, we gained a deeper understanding of how these factors interact and shape housing trends.

OBJECTIVES

1. Exploration of Income and Housing:

Examined the correlation between income quintiles and uncover their housing choices .

2. Regional Variation in Housing:

Regional Variation in Housing:
Explored how region impacts housing patterns and if it plays a role in determining the price and rent of a unit.

3. Income and Property Values

Find out if income plays a role in deciding the house market value.

4. Adequacy Index and Housing Value:

The analysis aimed to unveil if housing value and fair market rent related to the Housing Adequacy Index (ZADEQ)

DATA SCIENCE QUESTIONS

1. How many percentage are renter and owners in the whole dataset?
2. According to the Household Income Distribution 2009, what percentage are renters and owners
3. Do people with higher income have higher house value?
4. Do number of people in a household increase as the age of the head of a household increase?
5. Do older people prefer single unit or building with multiple units
6. Did the Adequacy Index (ZADEQ) affects the value of the house and the Fair Market Rent?
7. Do median income in an area affect what type of home structure that they built?
8. What is the most popular type of structure in each region?
9. Do region correspond to the market value of the unit, based on the number of rooms in the unit?
10. Are people segregated by the status of their income?



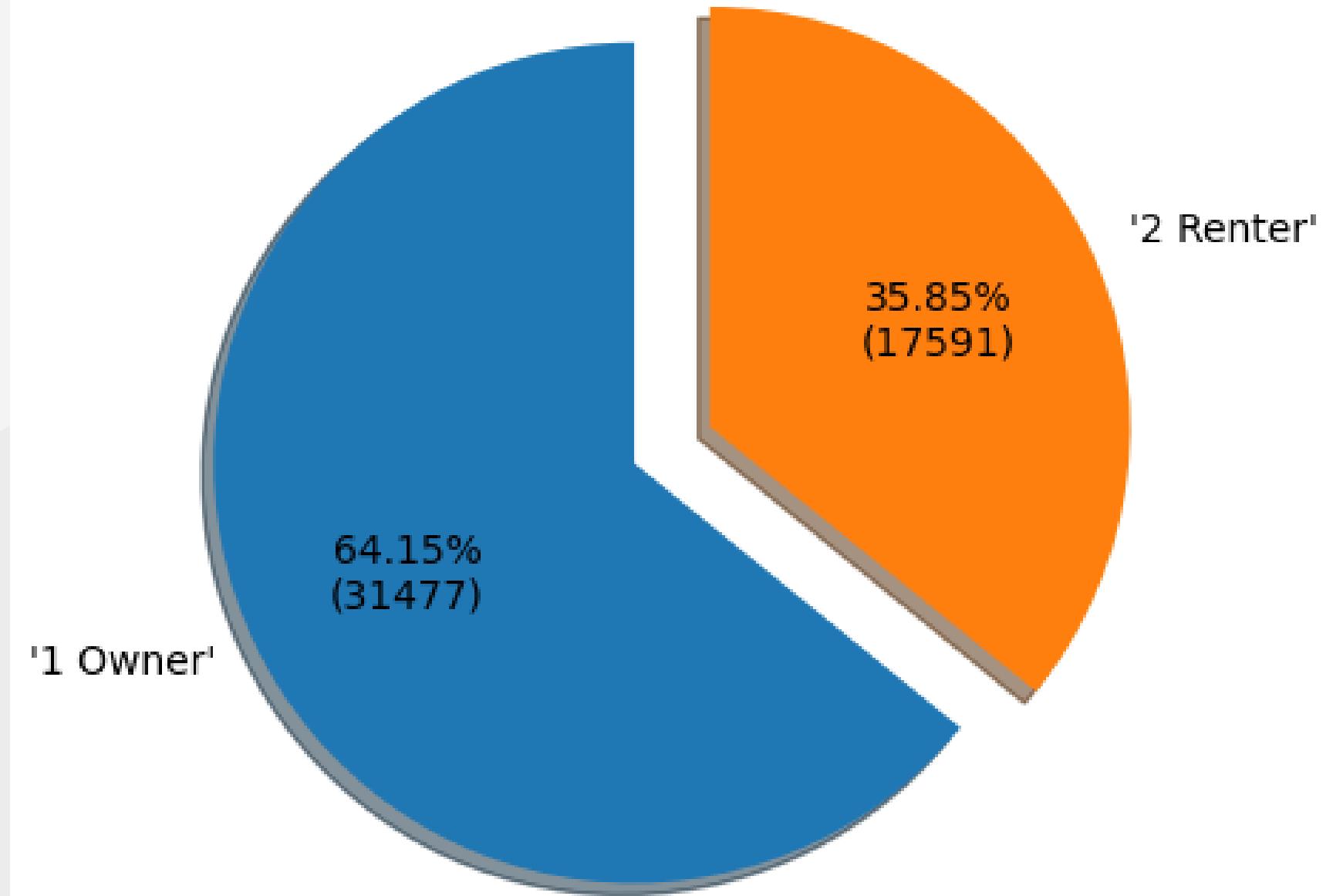
ANALYSIS

1. HOW MANY PERCENTAGE ARE RENTER AND OWNERS IN THE WHOLE DATASET?

- The majority of people in this dataset owned a home.

○ ○ ○ ○

Distribution of Renters and Owners



2. ACCORDING TO THE HOUSEHOLD INCOME DISTRIBUTION 2009, WHAT PERCENTAGE ARE RENTERS AND OWNERS

An income quintile is a measure of neighbourhood socioeconomic status that divides the population into 5 income groups (from lowest income to highest income) so that approximately 20% of the population is in each group([Link](#)). So, we are taking the Household Income Distribution (HIQ) in the year 2009 as our dataset are from 1985-2009.

So, the division is as below:

- Lowest Quintile <20,034
- 20,035 < Second Quintile < 38,549
- 38,550 < Middle Quintile < 61,800
- 61,801 < Fourth Quintile < 100,000
- 100,001 < Top Quintile < 180,001
- Top 5% > 180,001

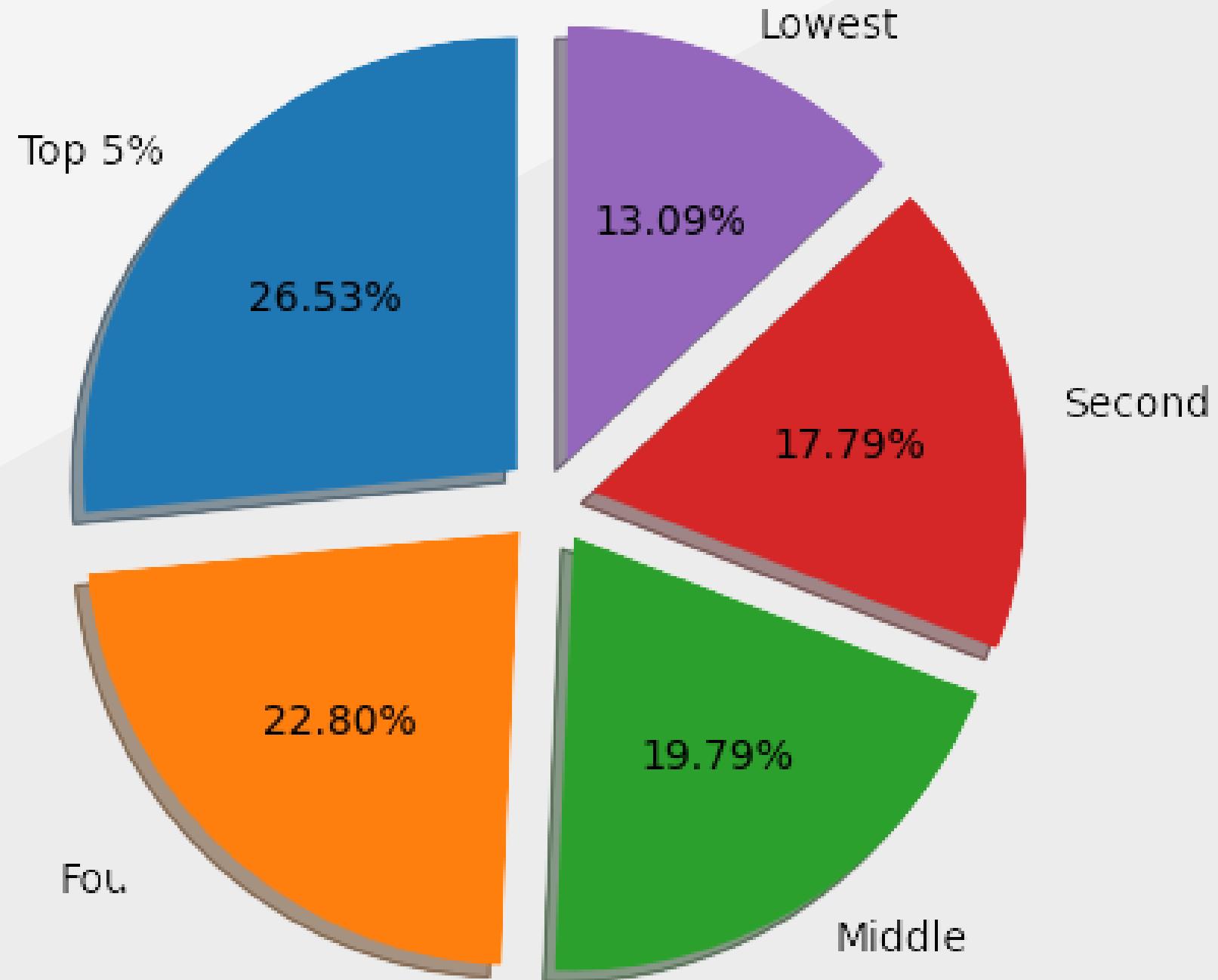
2009	
Upper Limit	Mean
\$20,453	\$11,552
\$38,550	\$29,257
\$61,801	\$49,534
\$100,000	\$78,694
—	\$170,844
\$180,001	\$295,388

○ ○ ○ ○

DISTRIBUTION OF OWNER OF THE HOUSE BY QUANTILE

- Approximately half (61.44%) of the owners fall into the fourth quintile and top 5% income groups. This suggests that higher-income individuals are more likely to own homes. This trend aligns with the perception that property ownership can be considered an investment for wealthier individuals.
- The idea that property can be seen as an investment rather than just a place to live is a valid one. High-income individuals may have the financial capacity to invest in real estate and take advantage of potential appreciation in property value.

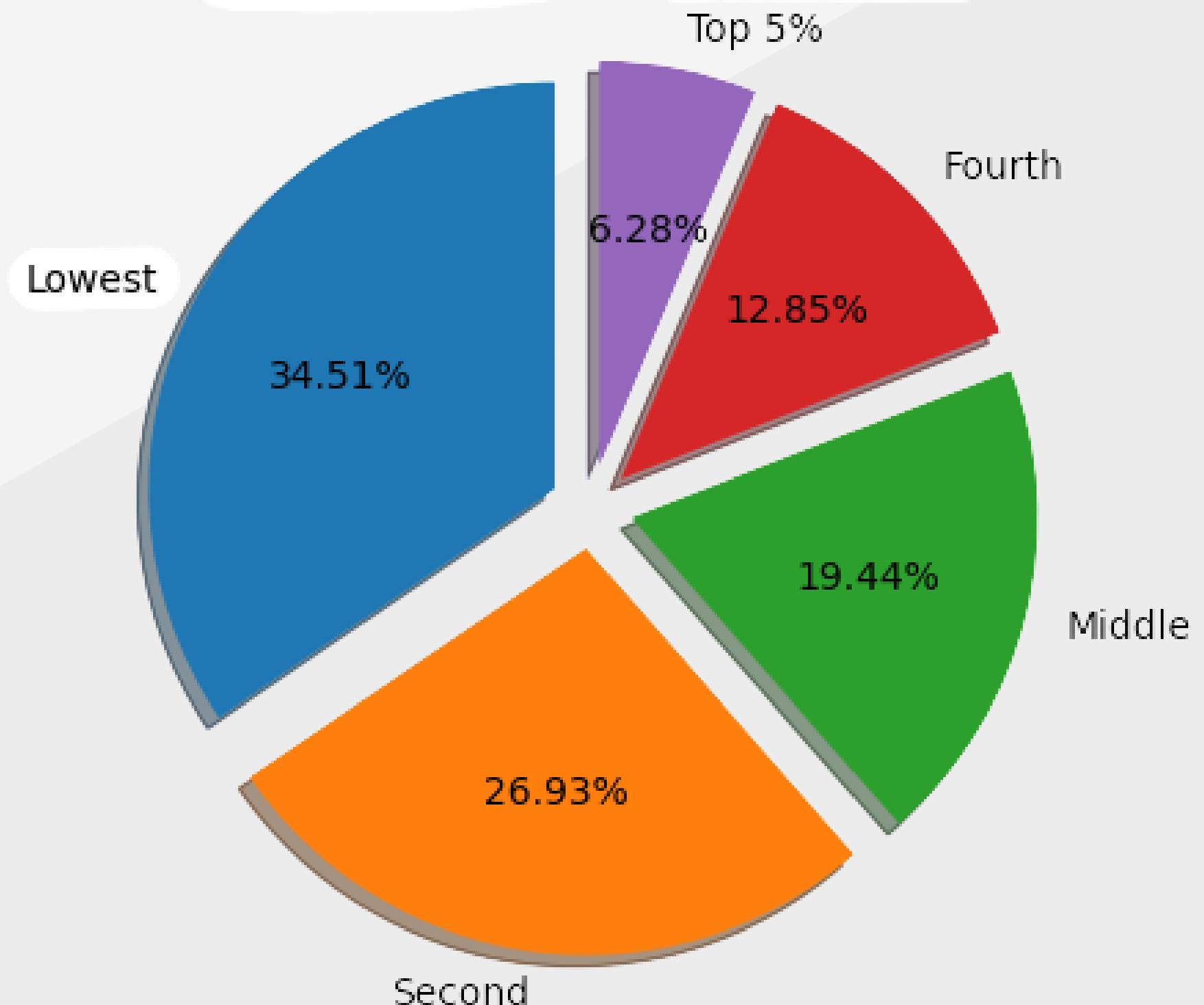
Distribution of Owner by Quintile



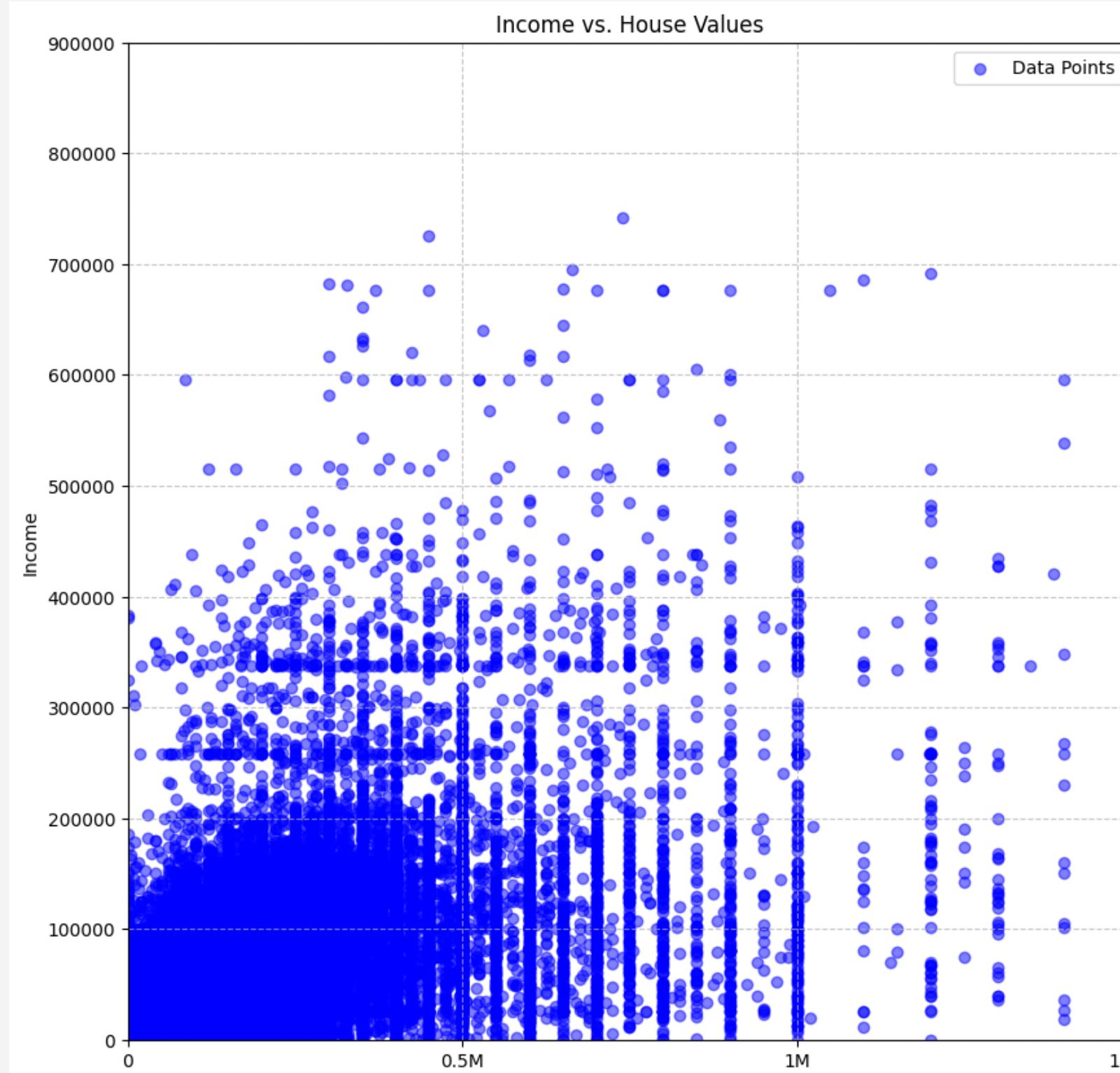
DISTRIBUTION OF NUMBER OF RENTERS BY QUANTILE

- The observation that the lowest and second quintiles dominate the rental market (constituting two-thirds of the data) is consistent with the economic realities of housing affordability. Individuals with lower incomes often find it more challenging to secure loans for home purchases, and they may be more reliant on renting as a housing option.
- Economic constraints, including low income and challenges in obtaining loans, can indeed drive a significant portion of the population toward renting rather than ownership.

Distribution of Renter by Quintile



3. DO PEOPLE WITH HIGHER INCOME HAVE HIGHER HOUSE VALUE?

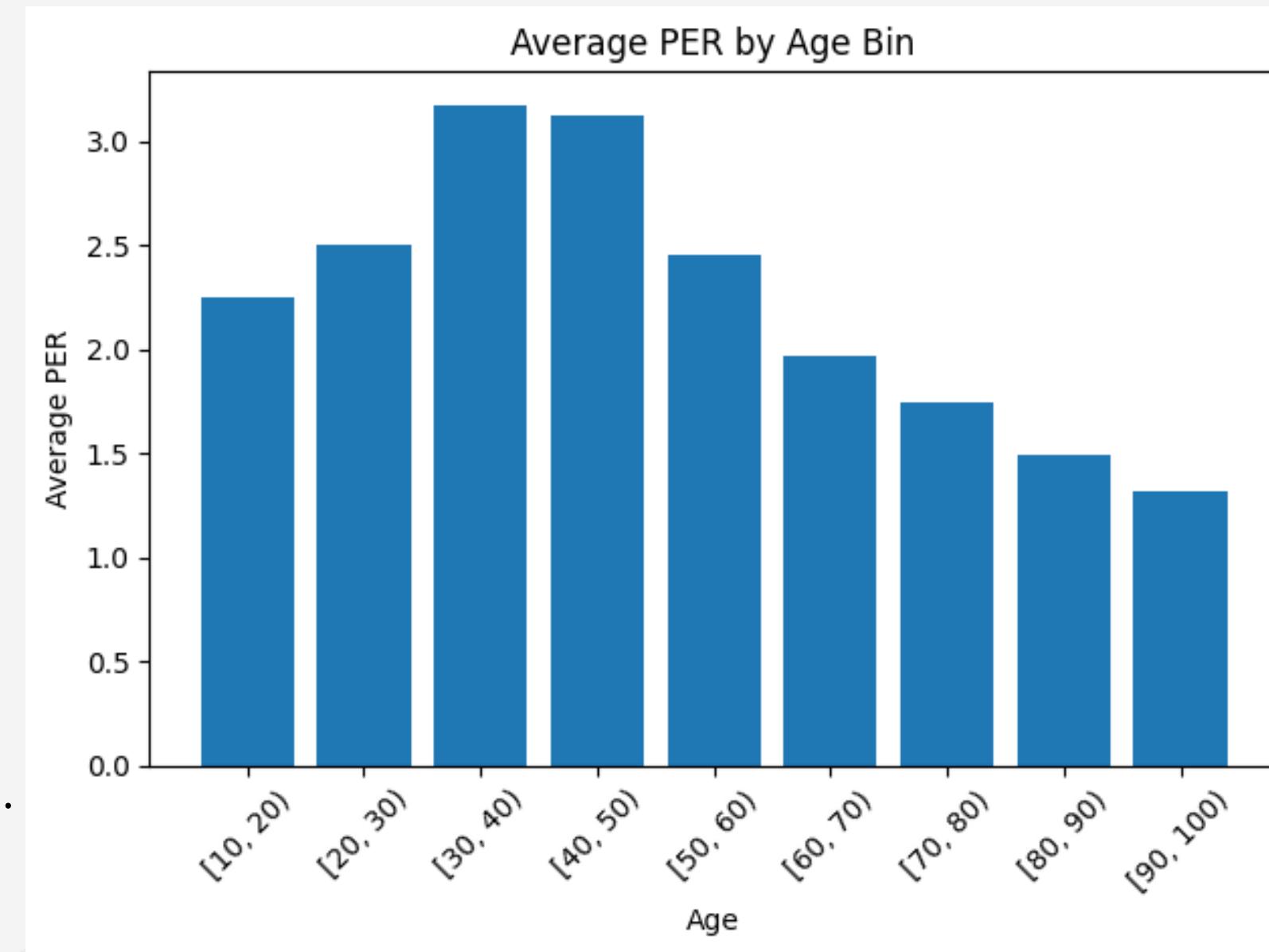


- Dominance of Lower House Values: Individuals with lower incomes tend to dominate the housing market for houses with values of \$500,000 and below. This reflects the affordability factor. Lower-income individuals are more likely to purchase homes within their budget constraints, which are typically in the lower price range.
- As income increases, individuals have more purchasing power and can afford higher-priced homes. This is why there is a more uniform scatter plot at higher house values. People with higher incomes have the financial capacity to consider a wider range of housing options



4. DO THE NUMBER OF PEOPLE IN A HOUSEHOLD INCREASE AS THE AGE OF THE HEAD OF THE HOUSEHOLD INCREASE?

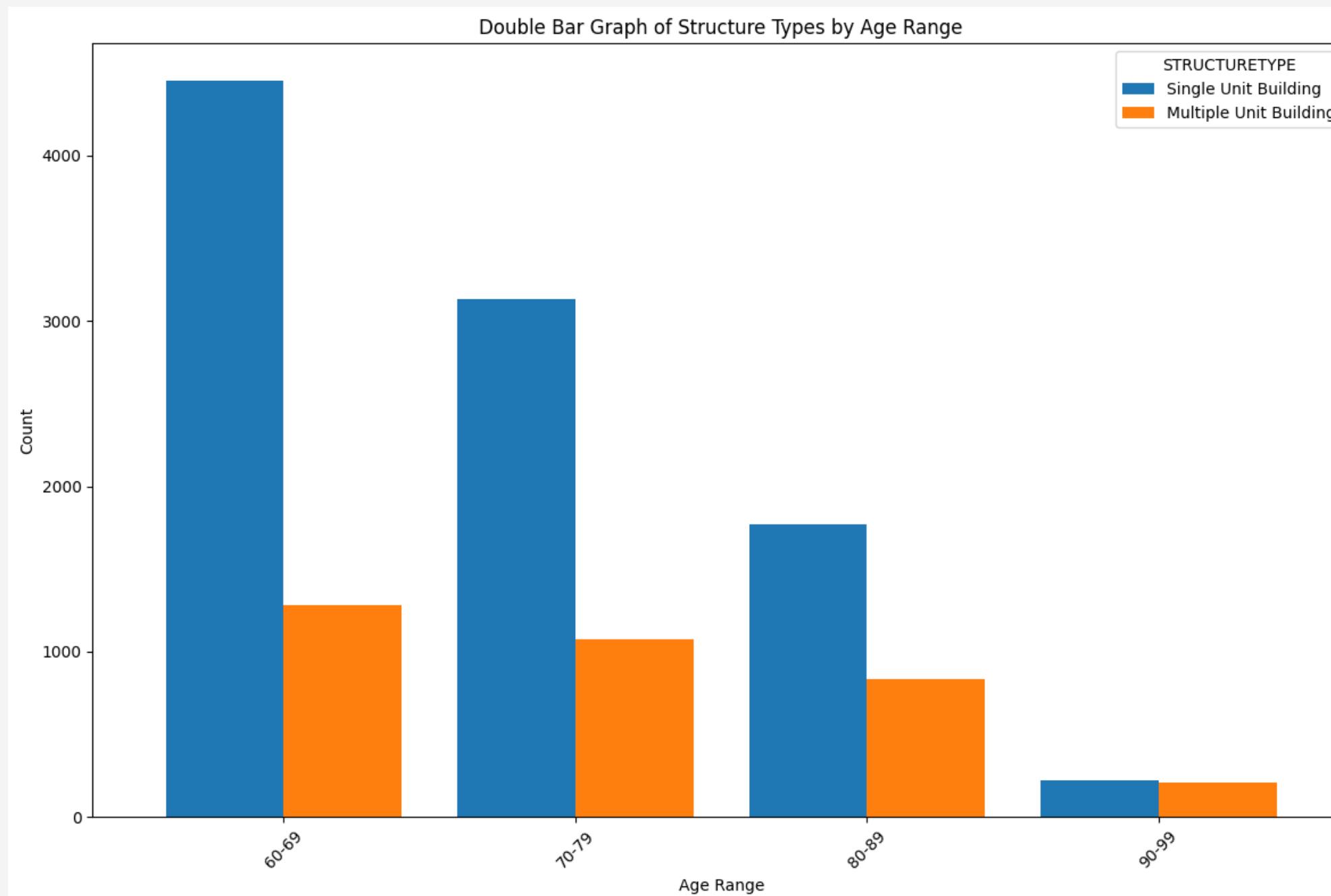
- As the age of the head of the household increases, the average number of people in the household tends to increase as well. This aligns with the typical life stages and family dynamics observed in many societies.
- In the earlier stages of adulthood, individuals often start families and have children. This is when the household typically includes not only the parents but also their children. The presence of children can significantly contribute to a larger household size.



Later Life Stages: After the age of 50 or so, children tend to move out of their parents' home as they become independent. This can lead to a decrease in the average number of people in the household, as the children are no longer counted as part of the household.



5. DO OLDER PEOPLE PREFER SINGLE UNIT OR BUILDING WITH MULTIPLE UNITS



- Older individuals, up to the age of 80, tend to prefer living alone. This could be due to a variety of factors, including maintaining their independence, privacy, and possibly their own established routines. At this stage, they might still feel capable of managing their daily activities and have a desire for self-sufficiency.
- Ages 90 and above: However, a shift occurs for individuals aged 90 and above, where they tend to opt for living in a community or with others. This change could be influenced by several reasons, including health considerations, safety concerns, and the desire for social interaction.



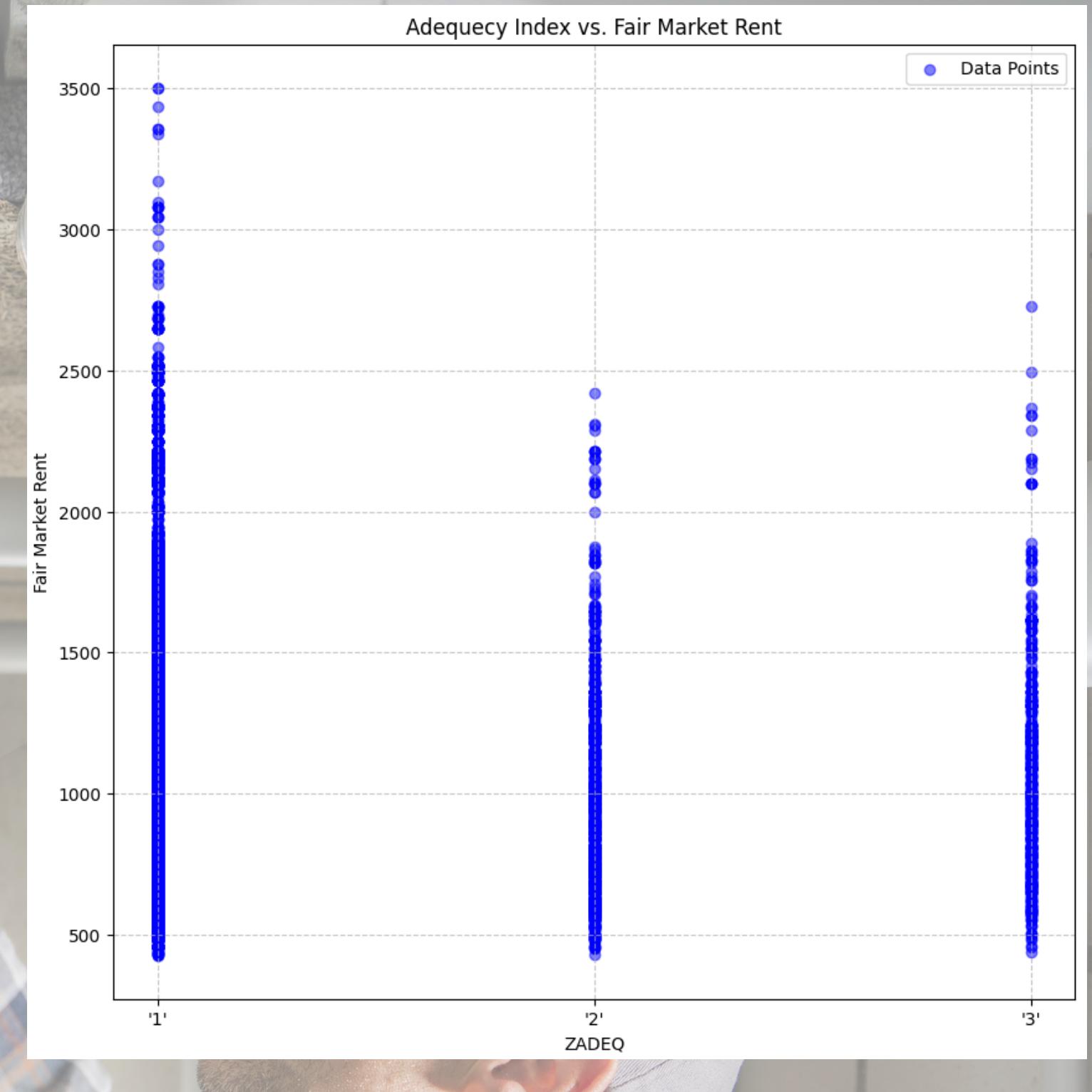
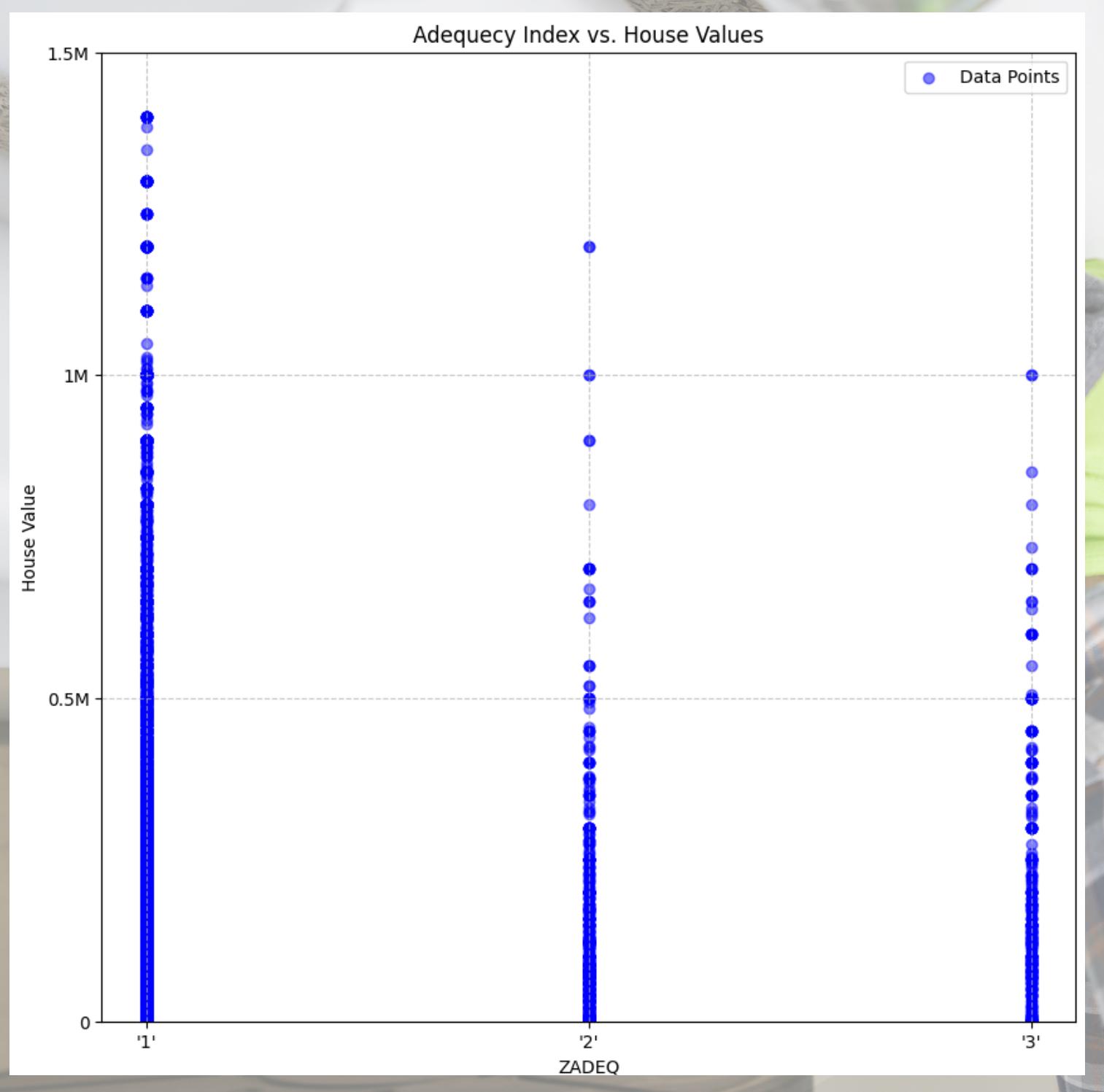
6. DID THE ADEQUACY INDEX (ZADEQ) AFFECTS THE VALUE OF THE HOUSE AND THE FAIR MARKET RENT?

According to American Housing Survey, The Adequacy Index (named ZADEQ prior to the 2015 AHS) is comprised of eight criterion capturing severe physical deficiencies like no running water, plumbing, etc. US Department of Housing and Urban Development (HUD) created this measure to assess the extent to which the housing stock met the standard of “a decent home and a suitable living environment,”

The HUD measure is contained in the variable ZADEQ, which takes three values:

- 1 - if a unit is considered adequate,
- 2 - if a unit is considered moderately inadequate, and
- 3 - if a unit is considered severely inadequate

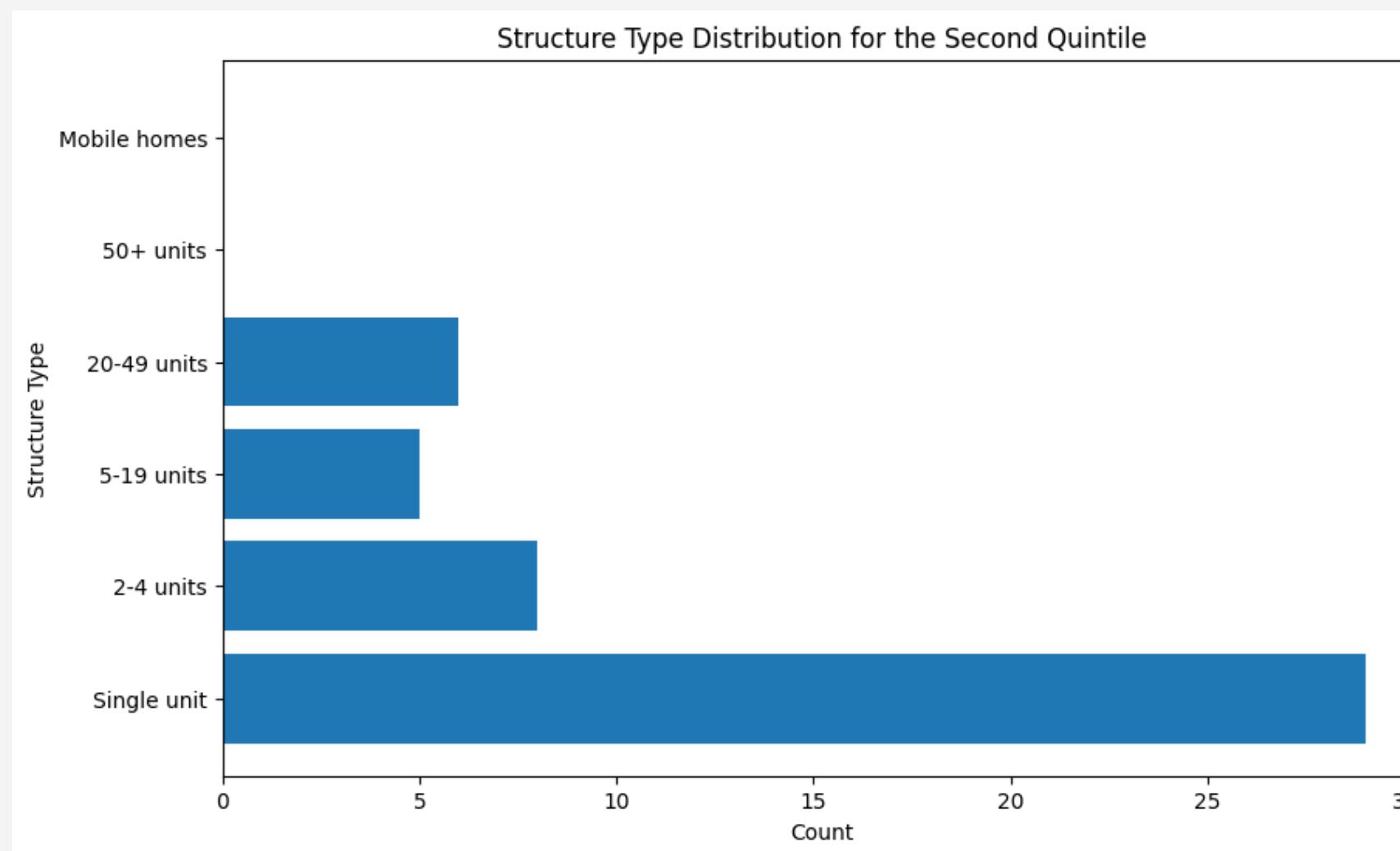
0 0 0 0



Both the value of the house and Fair Market Rent tend to be higher when the Adequacy Index score is higher. This observation suggests that people generally prefer to live in houses that meet basic standards of adequacy. As a result, houses with higher Adequacy Index scores are likely to have better demand, leading to increased values and rents.

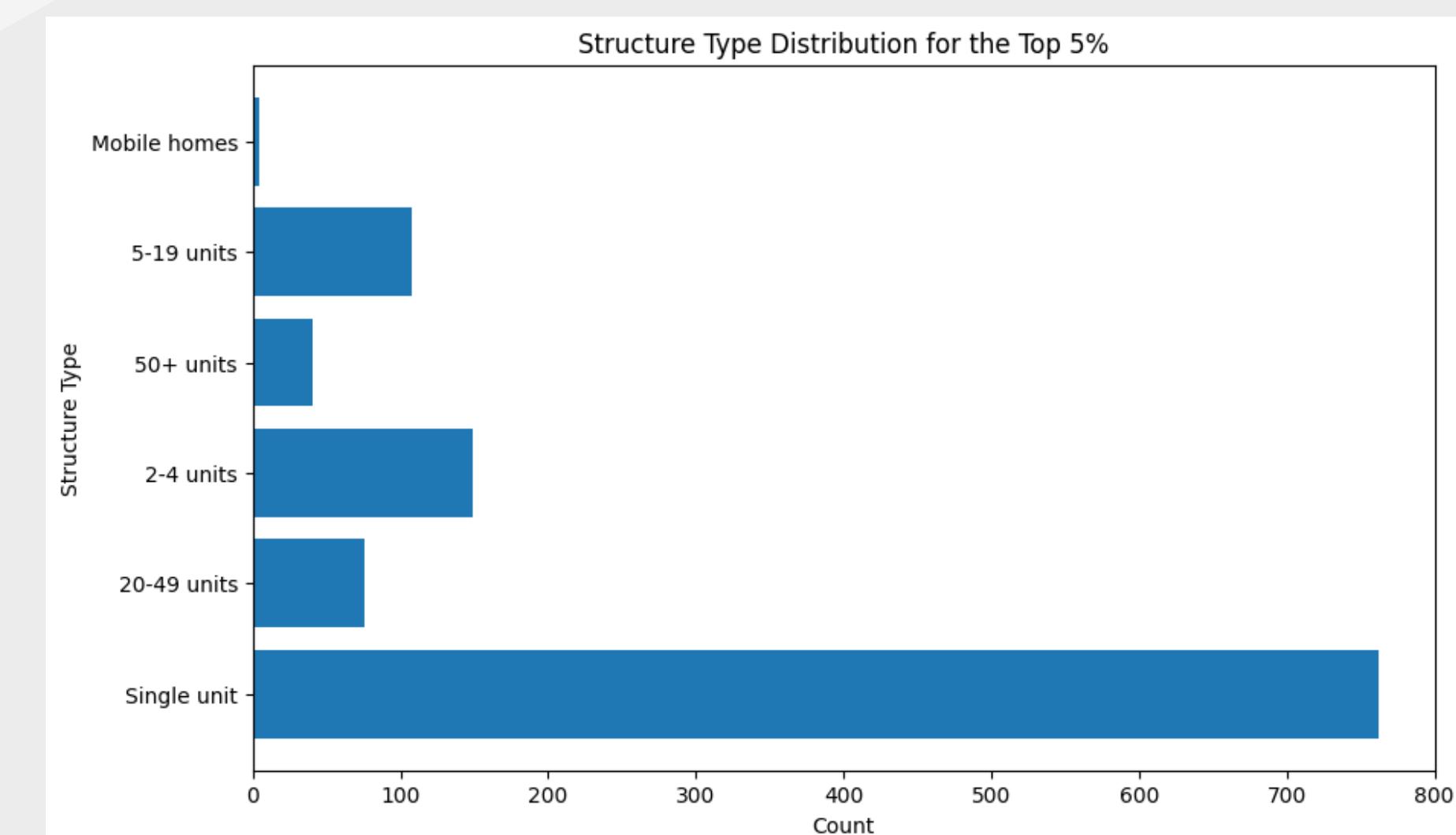
This analysis is in line with the concept of supply and demand in the housing market

7. DOES THE MEDIAN INCOME IN AN AREA AFFECT THE TYPE OF STRUCTURE THAT THEY BUILT?



The number of building may be explain by the Top 5% having multiple houses to generate passive income.

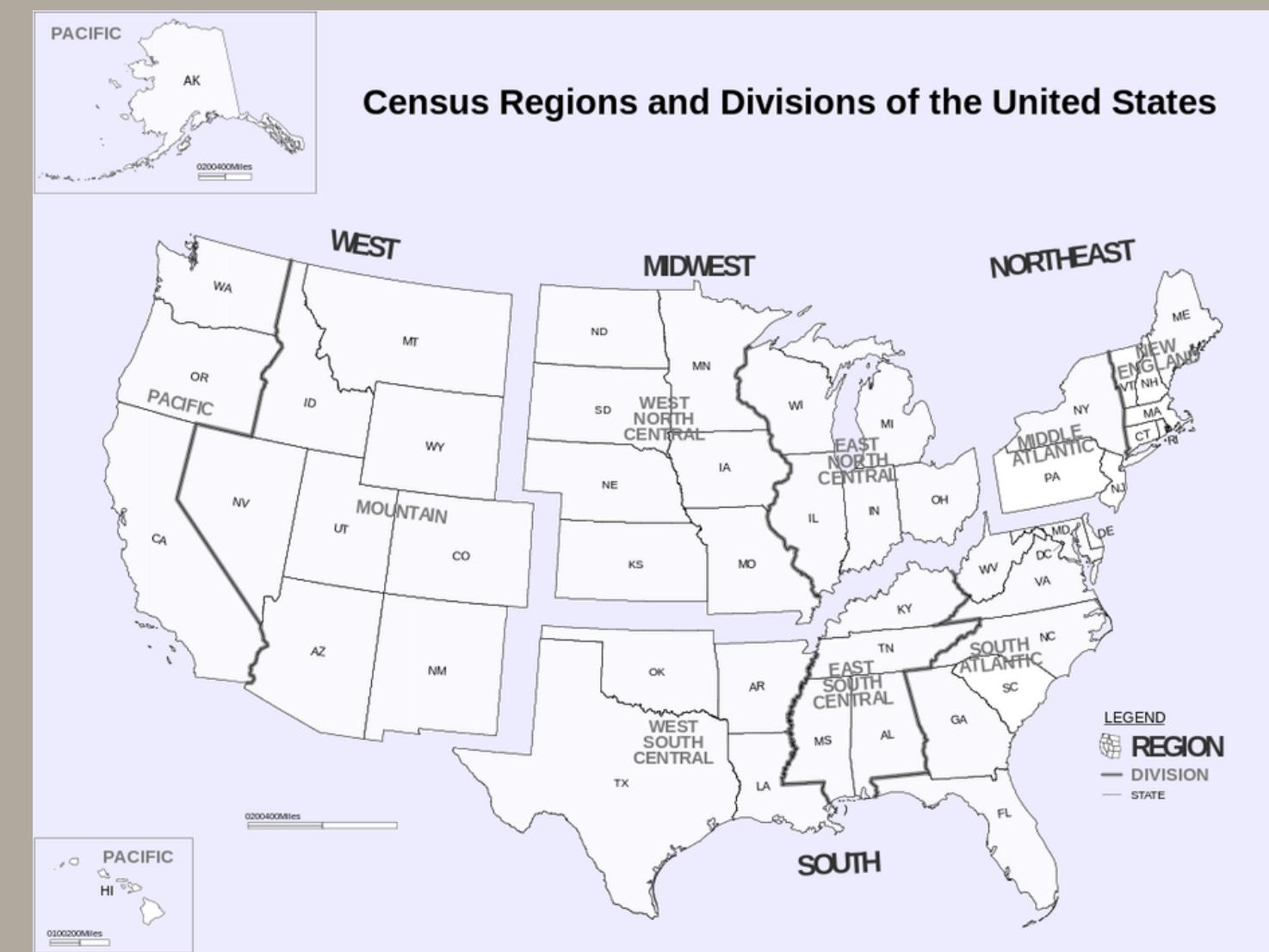
The second lowest quintile have many single unit but they also have the largest number of building that have several units. This is because they dont have the income to afford a single unit house which usually pricier than an apartment.

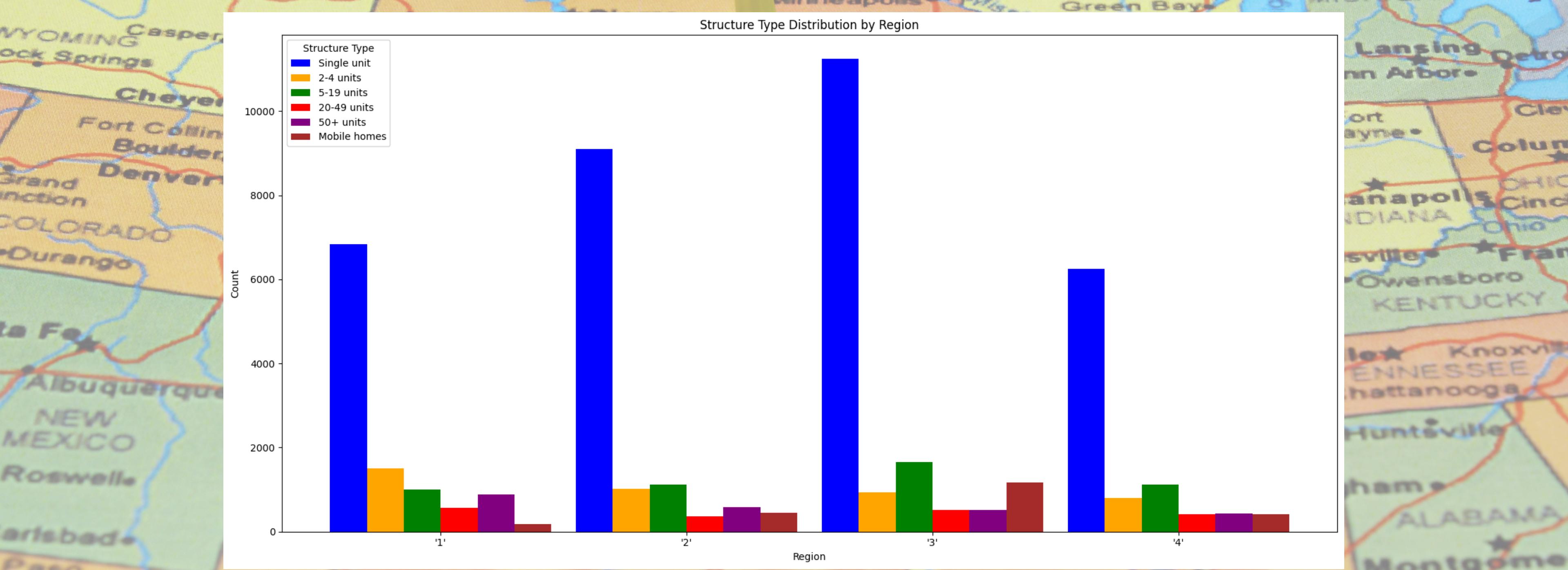


8. WHAT IS THE MOST POPULAR TYPE OF STRUCTURE IN EACH REGION?

Since 1950, the United States Census Bureau defines four statistical regions, with nine divisions. The Census Bureau region definition is "widely used for data collection and analysis", and is the most commonly used classification system. [[Wikipedia](#)]

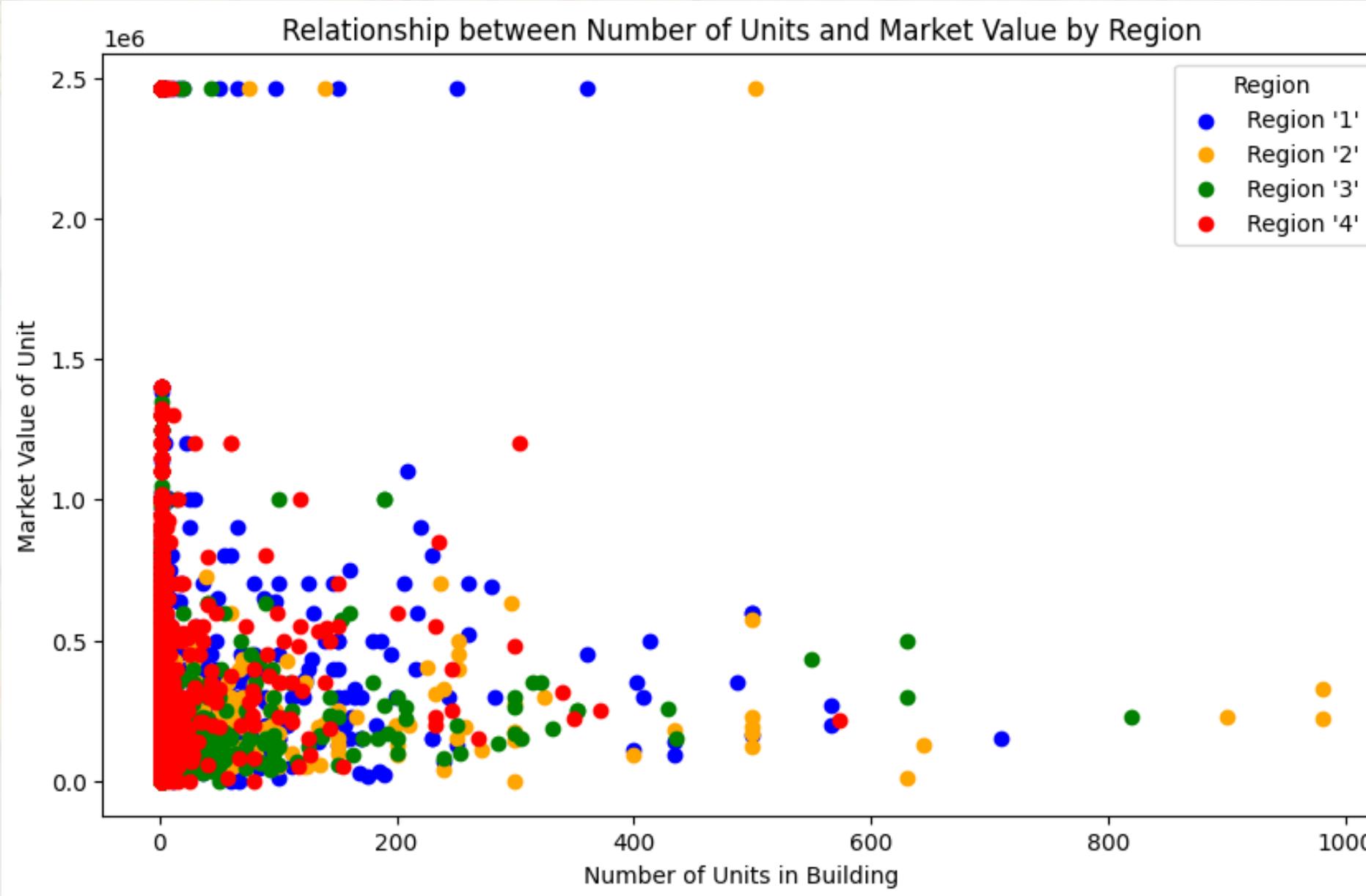
- Region 1 - Northeast (massachuset, rhode island)
- Region 2 - Midwest (michigan, ohio, nebraska)
- Region 3 - South (Washington DC, texas)
- Region 4 - West (california, arizona)





- The majority of homes in the United States are single unit which is not surprising considering the US is wide and big so they would much prefer single unit homes.
- In this context, the high number of multi-unit buildings in Region 1 aligns with the concentration of densely populated areas. Metropolises like New York City within this region are famously known for their towering apartment buildings that utilize available space to house a significant number of residents.

9. DO REGION CORRESPOND TO THE MARKET VALUE OF THE UNIT, BASED ON THE NUMBER OF ROOMS IN THE UNIT?



Region 1 - Northeast

- Diverse housing types and unit values.
- Potentially represents urban and densely populated areas with various housing options.

Region 2 - Midwest

- Contains the most expensive units.
- Suggests a mix of both high-value properties and more affordable housing.

Region 3 - South

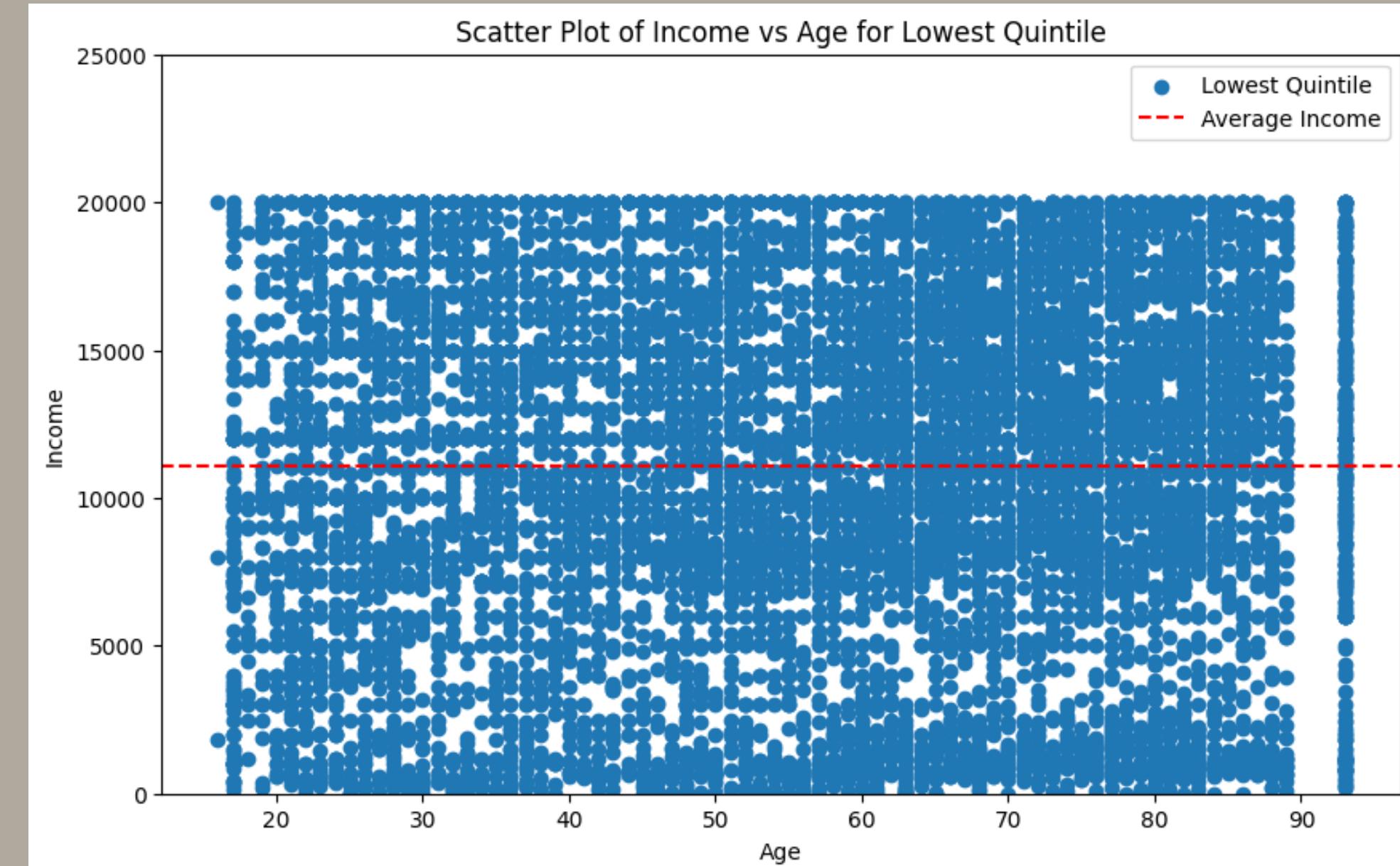
- Units are not very expensive.
- Implies a relatively stable and affordable housing market.

Region 4 - West

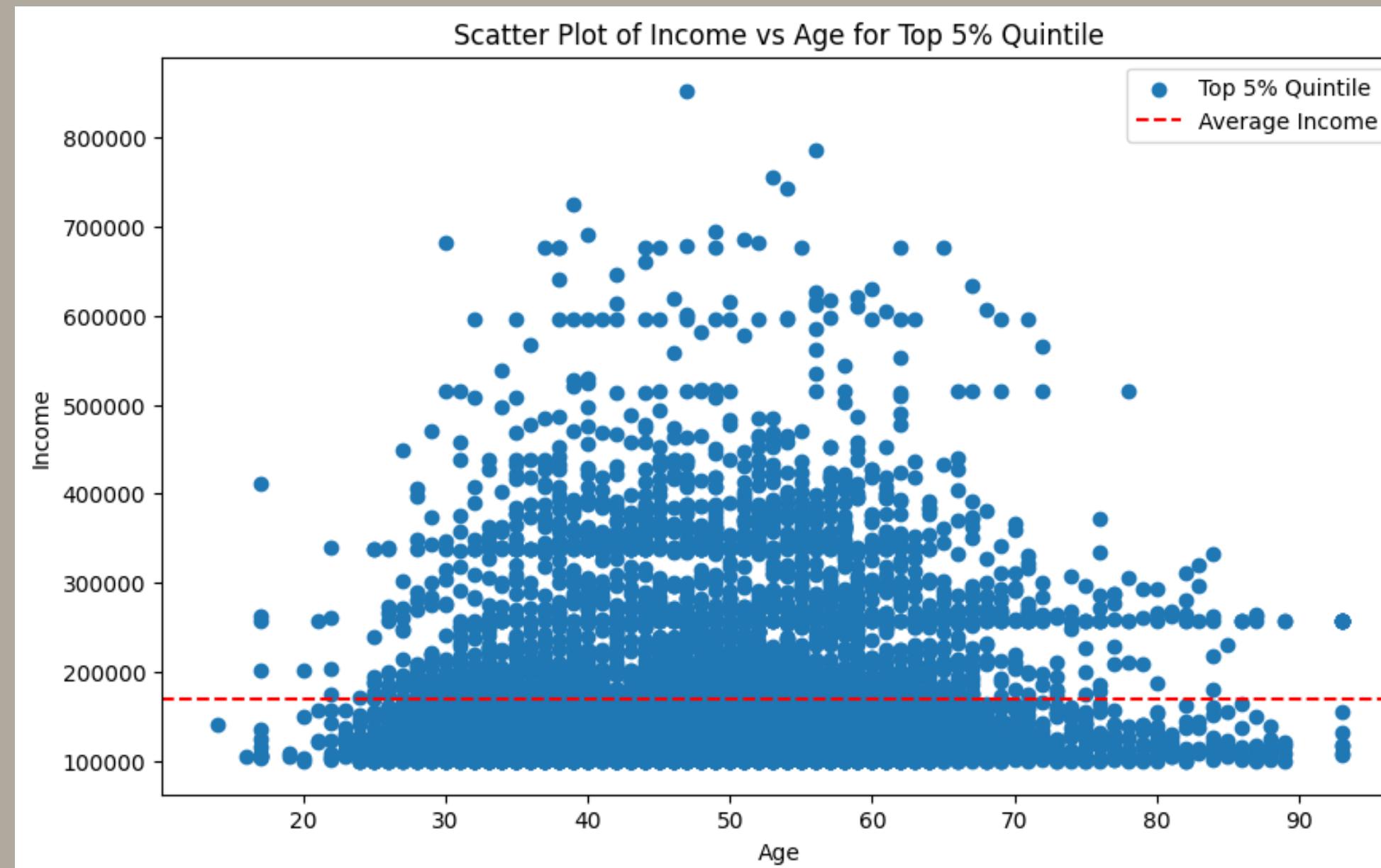
- Highest number of single-unit houses
- Units are not very expensive.
- Land availability in this region likely contributes to the prevalence of single-unit houses and lower prices.

10. ARE PEOPLE SEGREGATED BY THE STATUS OF THEIR INCOME?

Scatter Plot of Income vs. Age for the Lowest Quintile appears relatively uniform. This uniformity could be due to the fact that individuals within the lowest quintile tend to have more consistent income and age demographics, resulting in less variation in the data points.



10. ARE PEOPLE SEGREGATED BY THE STATUS OF THEIR INCOME?



This shows that for the top 5%, the income ranges are so wide that 800k income-person can live with 100k income-person.

A wide income range within the top 5% quintile, which can indeed result in individuals with significantly different income levels falling within the same category. This variation could be due to various factors, such as differences in household composition, location, or other socio-economic factors.

CONCLUSION

The research illuminated the intricate relationship between income, region, age, household size, and housing choices. Income emerged as a pivotal factor, impacting ownership, renting, and property values.

Regional disparities were evident in housing preferences, with factors like open space availability influencing choices. Age also played a role, shaping preferences for communal living in later years. The findings emphasize that housing dynamics are multifaceted, influenced by economic, social, and regional factors.