PRESENTATION
BIG DATA ANALYTICS

ANALYZING THE RELATIONSHIP OF HOUSING-RELATED FACTORS IN VARIOUS REGIONS AND LOCATIONS AND THEIR IMPACT ON HOUSING AFFORDABILITY

Prepared by Ibrahim Abu Saiid

BACKGROUND

The housing landscape in the United States is a dynamic ecosystem, influenced by diverse factors ranging from regional variations to assisted housing programs. As housing affordability remains a critical concern, understanding the intricate relationships among housing-related factors becomes paramount. This analysis seeks to reveal the complex interplay between housing aspects and affordability across different regions, locations, and household characteristics. By examining these relationships, we aim to provide valuable insights that can inform policymakers, housing developers, and realtors in making informed decisions to enhance housing accessibility and affordability.

OBJECTIVE

The primary objective of this analysis is to comprehend the intricate relationship between various housing-related factors such as household size, age, location, region, cost burden, household income and assistance. And show how they impact housing affordability. Our exploration involves four key aspects:

1 REGIONAL & LOCATIONAL AFFORDABILITY

- Investigate cost burden distributions across different regions for both renters and owners.
- Examine how the location (city center vs. suburbs) impacts overall housing affordability.

2 HOUSING ADEQUACY & COST RELATIONSHIPS

- Explore the relationship between perceived housing adequacy and cost burdens.
- Identify specific cost-related patterns associated with different adequacy categories.

3 ASSISTED VS. NON-ASSISTED HOUSING AFFORDABILITY

- Compare the affordability of assisted housing units to non-assisted units

4 HOUSEHOLD DYNAMICS & AFFORDABILITY

- Analyze the relationship between utility costs and housing burden based on household size and available bedrooms.
- Investigate how household size influences the affordability of housing, especially in top burden categories.
- Examine how the burden of housing costs changes with the age of the head of the household.

DATA SCIENCE QUESTIONS

- What is the distribution of cost burden of renters and owners across different regions?
- How does the location, like being in a city center or suburbs, impact how affordable housing is for people?
- What is the relationship between perceived housing adequacy and housing cost burdens based on the location, and are there specific cost-related patterns associated with different adequacy categories?
- How does the affordability of assisted housing units compare to non-assisted units?
- Despite similar average incomes between assisted and non-assisted housing units in both Poverty and AMI income categories, what factors contribute to the provision of assistance to these households?
- How does the relationship between utility costs and housing burden change based on the number of people in a household and the available bedrooms?
- How does the household size influence the affordability of housing, particularly in relation to the top burden categories associated with different household sizes?
- How does the burden of housing costs change with the age of the head of the household?

METHODOLOGY

1 **DATA LOADING AND IMPORTING LIBRARIES:**

Load the HADS dataset, import required libraries, and set up the Spark environment for analysis

2 **DATA PREPROCESSING - DATA CLEANING AND PREPARATION:**

Check for null values and handle duplicates to ensure data quality.

3 **EXPLORATORY DATA ANALYSIS (EDA):**

Conduct preliminary data analysis to identify patterns and outliers, using visualizations for insights.

3 **ANALYSIS:**

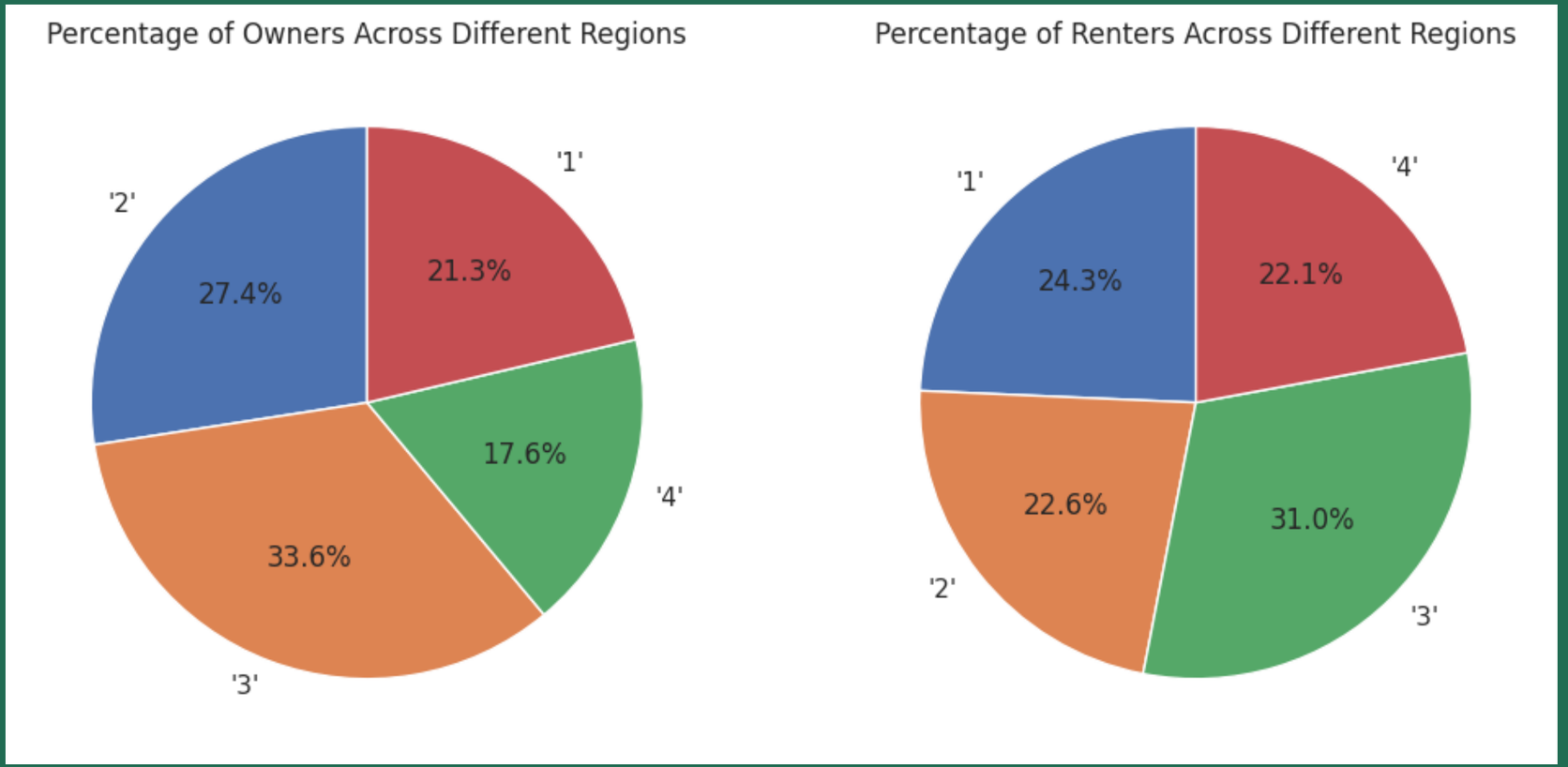
Utilize statistical tools and visualizations to address each of the eight data science questions, focusing on housing affordability, regional disparities, and other relevant factors.

4 **INTERPRETATION AND CONCLUSION:**

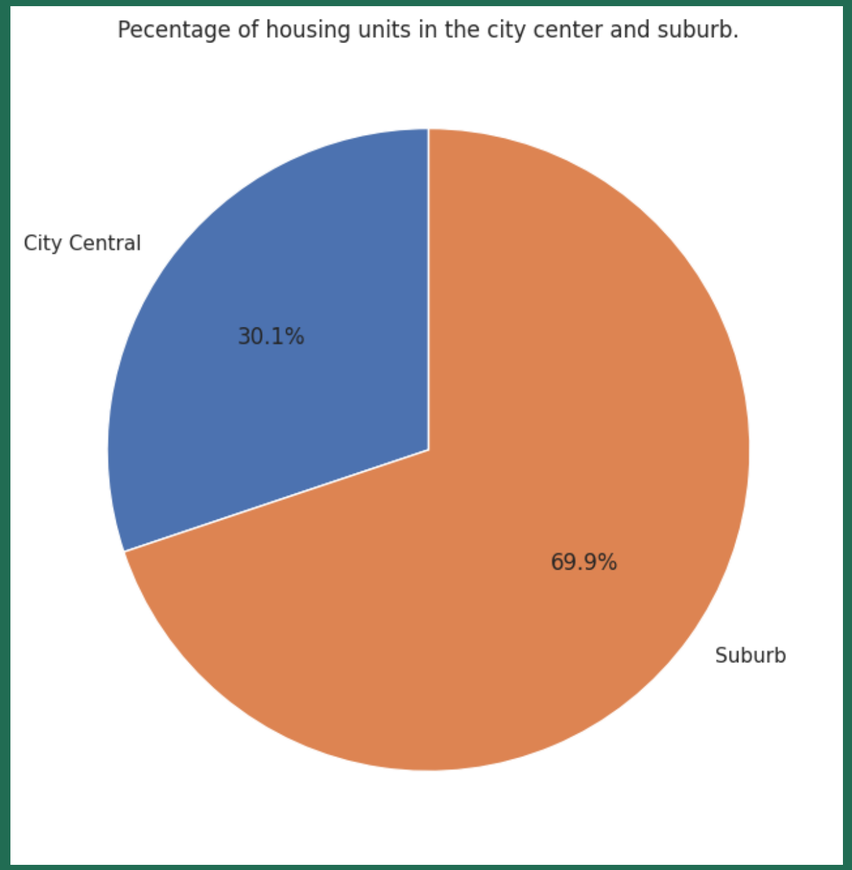
Derive meaningful insights from the analyses, summarizing key findings and concluding with implications for stakeholders and potential strategies for improving housing accessibility and affordability.

EXPLORATORY DATA ANALYSIS

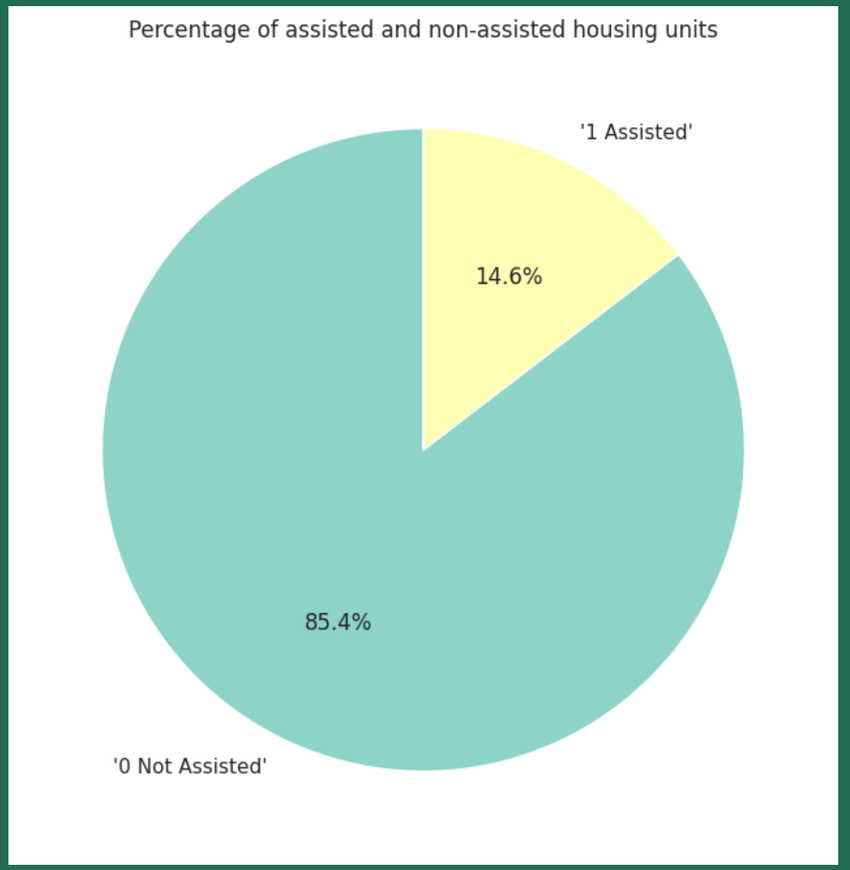
1 PERCENTAGE OF RENTERS AND OWNERS ACROSS DIFFERENT REGIONS



2 PECENTAGE OF HOUSING UNITS IN THE CITY CENTER AND SUBURB



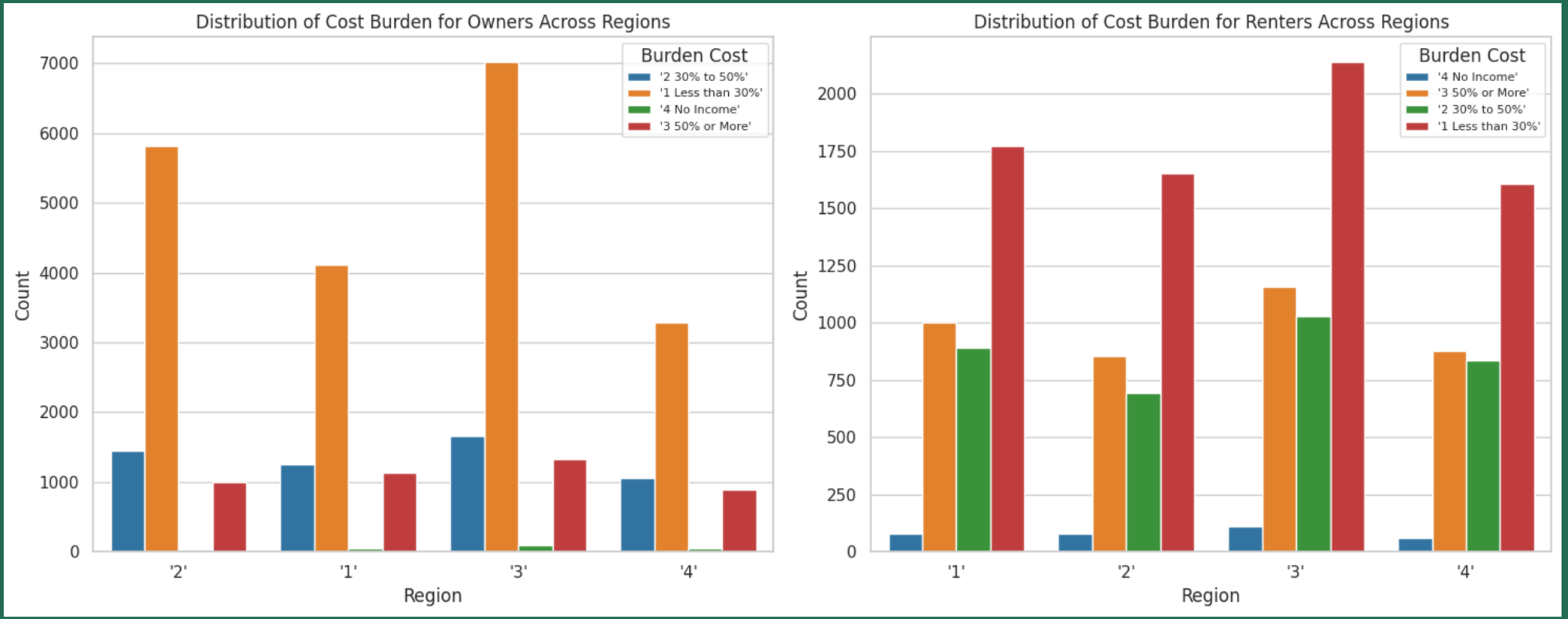
3 PERCENTAGE OF ASSISTED AND NON-ASSISTED HOUSING UNITS



RESULTS & VISUALIZATION

INVESTIGATEING COST BURDEN DISTRIBUTIONS ACROSS DIFFERENT REGIONS FOR BOTH RENTERS AND OWNERS

1 What is the distribution of cost burden of renters and owners across different regions?



1 Insights and Interpretations

- **General Comparison:** Across all regions, renters consistently showed a higher count in each burden category compared to owners. This suggests that a **larger proportion of renters, regardless of the region, experience various levels of cost burden compared to homeowners**.

- **Predominance of "Less than 30%":** The burden category "Less than 30%" emerged as the most common among both owners and renters in all regions. This indicates that a significant portion of both groups, regardless of the region, experiences a relatively lower cost burden, **suggesting a degree of affordability in housing costs.**

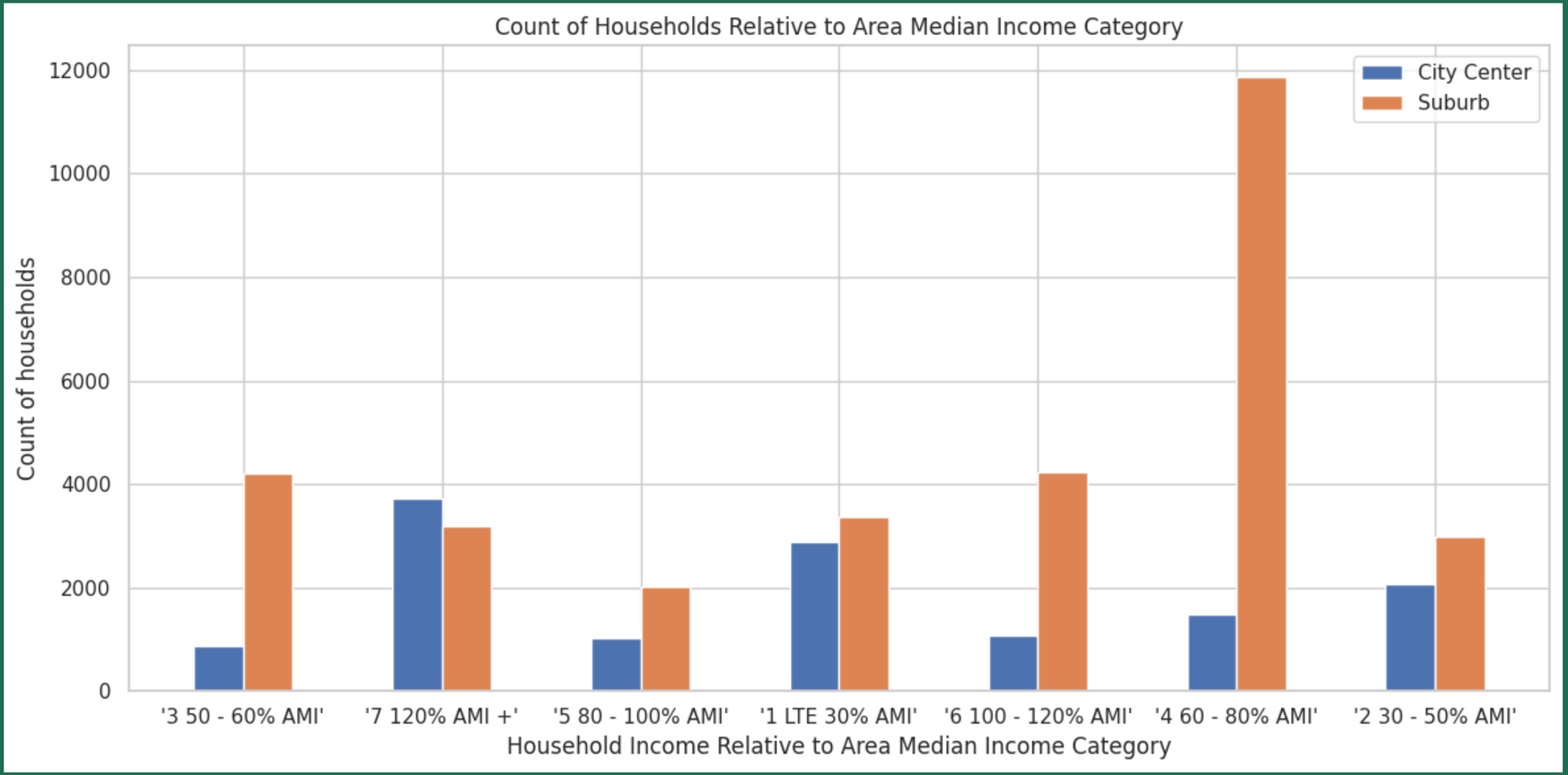
- **Regional Disparities:** Region 3 stands out as having the highest count in the "Less than 30%" burden category for both renters and owners. This might **suggest that Region 3 is relatively more affordable compared to other regions**, as it has a larger proportion of individuals with a lower housing cost burden.

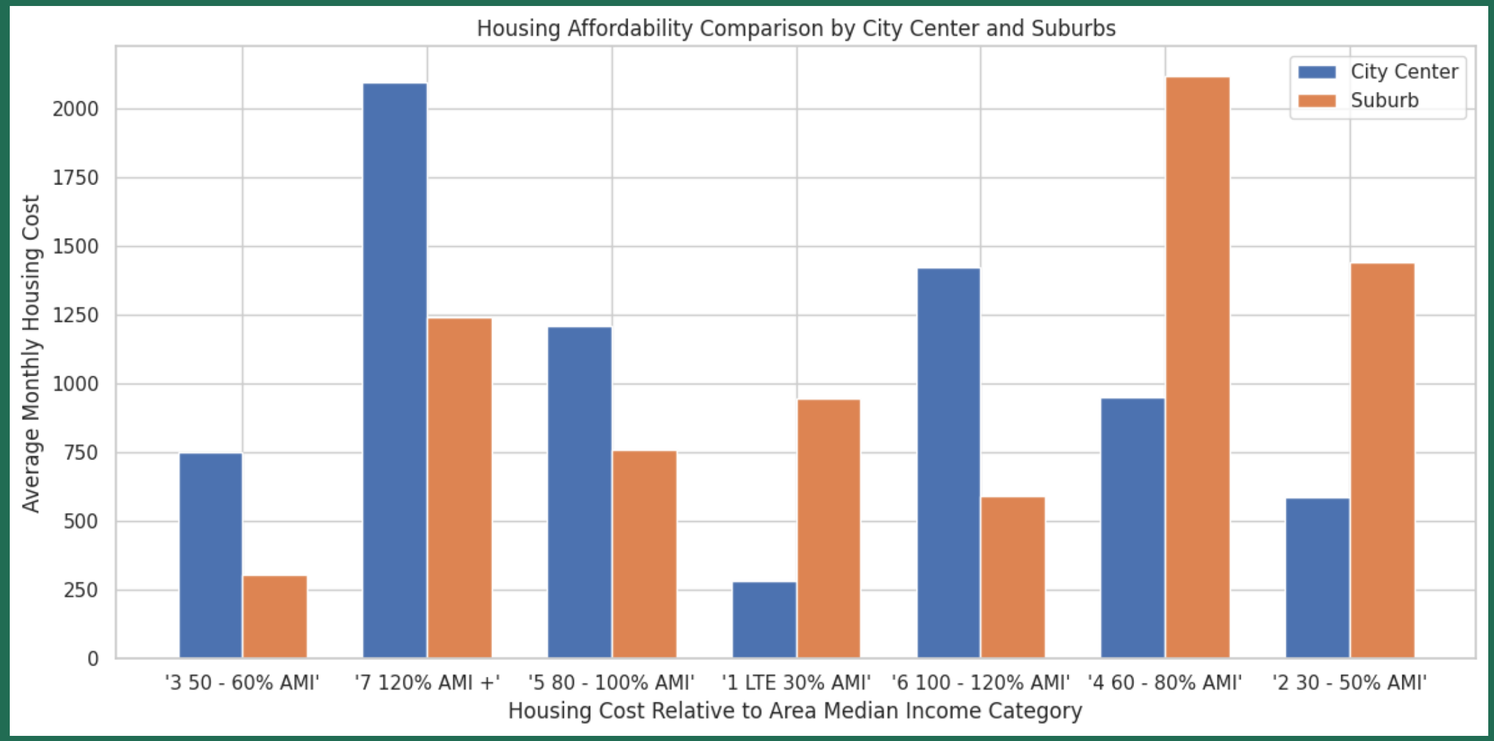
- **Policy Implications:** The findings may have implications for policymakers aiming to address housing affordability. **Regions with consistently higher counts in the "Less than 30%" category could serve as models** for implementing successful affordability strategies.

EXAMINE HOW THE LOCATION (CITY CENTER VS. SUBURBS) IMPACTS OVERALL HOUSING AFFORDABILITY.

2

How does the location, like being in a city center or suburbs, impact how affordable housing is for people?





2 Insights and Interpretations

- **Income Distribution:** In both the city center and suburbs, households in the lower AMI categories (LTE 30% and 30 - 50% AMI) had a higher count. This suggests that **both locations have a significant proportion of households with relatively lower incomes.**

- **Suburban Dominance in AMI Categories:** Suburbs consistently showed a higher count across most AMI categories (1, 2, 3, 4, 5, 6), **indicating a larger population in these income brackets** compared to the city center.

- **Suburban Mass in Category 4:** Category 4 (60 - 80% AMI) stood out with a significantly higher count in the suburbs. This might suggest that **a substantial number of households** in the suburbs **fall within the moderate-income range.**

- **Affordability Standard - Average Monthly Costs:** City households generally had higher average monthly costs in AMI categories 3, 5, 6, and 7. These categories represent higher income brackets, **indicating that housing affordability** tends to be **more challenging in the city for households with relatively higher incomes.**

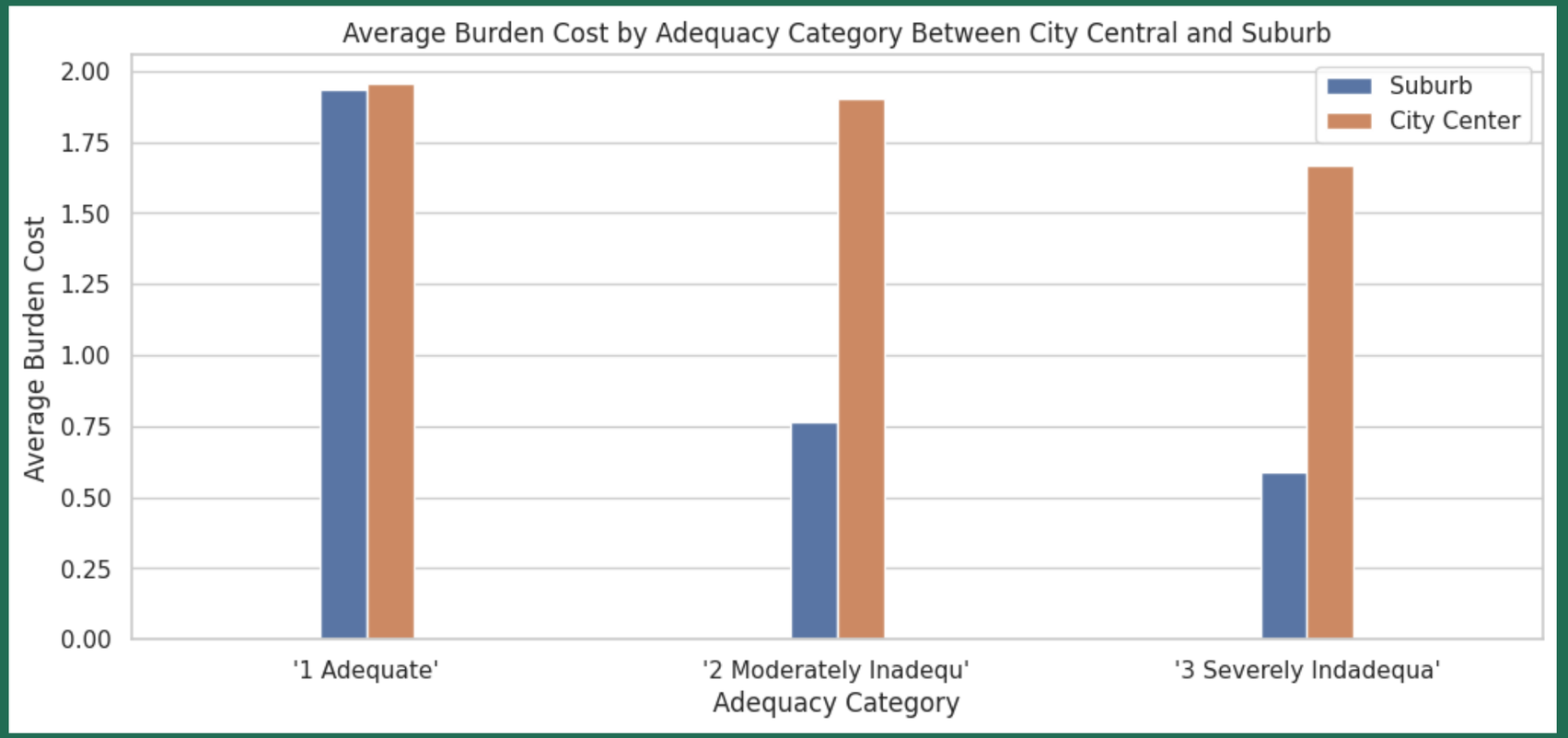
- **Suburban Affordability in Categories 1, 3, 4:** Suburbs had higher average monthly costs in AMI categories 1, 3, and 4, highlighting the affordability challenges in these income brackets in suburban areas

- **Impact of Higher AMI on City Costs:** The observation that city households face higher costs in the higher AMI categories (3, 5, 6) may indicate the influence of higher living standards or housing prices in the city center.

- **Policy Considerations:** Policymakers could explore strategies to enhance affordability for households in higher income brackets in both city and suburban areas. This may involve targeted housing policies or incentives for moderate-income households.

EXPLORING THE RELATIONSHIP BETWEEN PERCEIVED HOUSING ADEQUACY AND COST BURDENS

3 What is the relationship between perceived housing adequacy and housing cost burdens based on the location, and are there specific cost-related patterns associated with different adequacy categories?



3 Insights and Interpretations

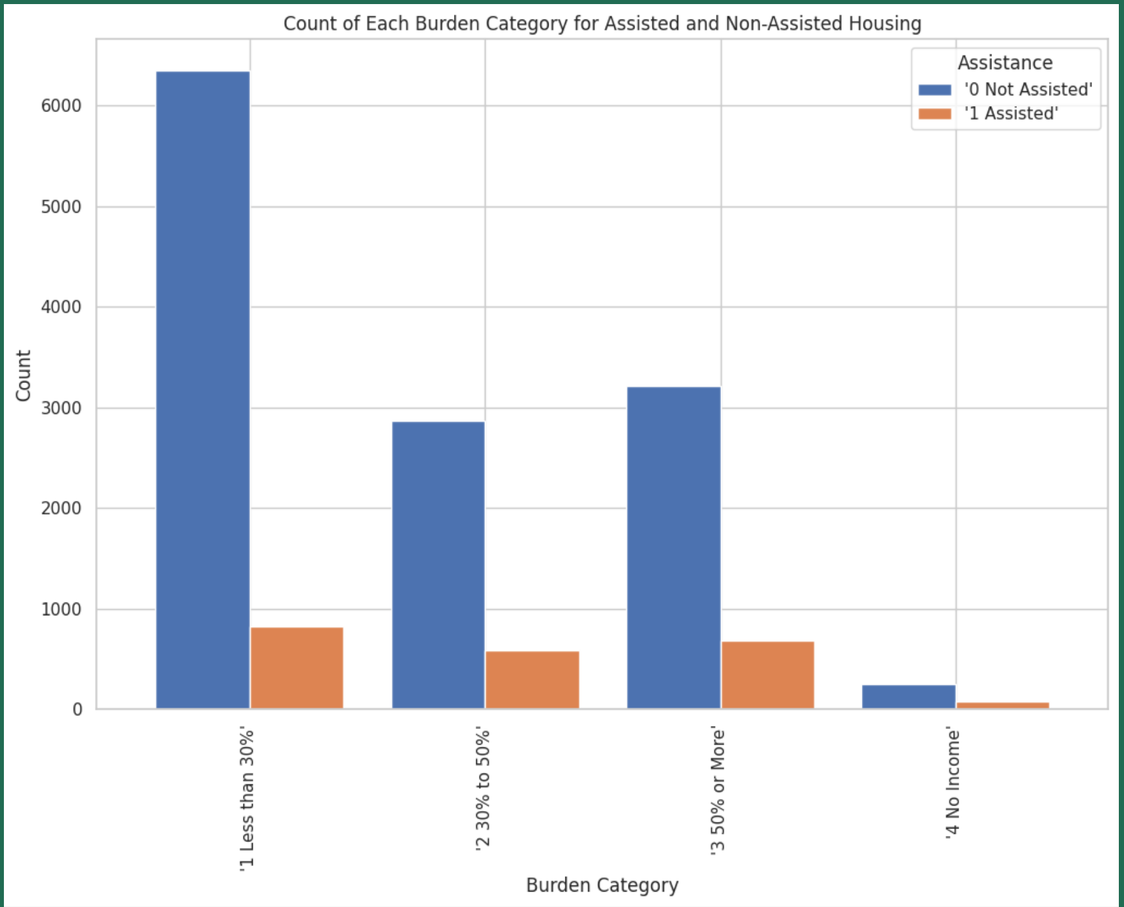
- **Consistent Burden in City Central:** City Central exhibited relatively consistent average burden costs across all adequacy categories (Adequate, Moderately Inadequate, Severely Inadequate). This suggests that, regardless of perceived housing adequacy, **the average burden cost remained relatively stable in the city.**

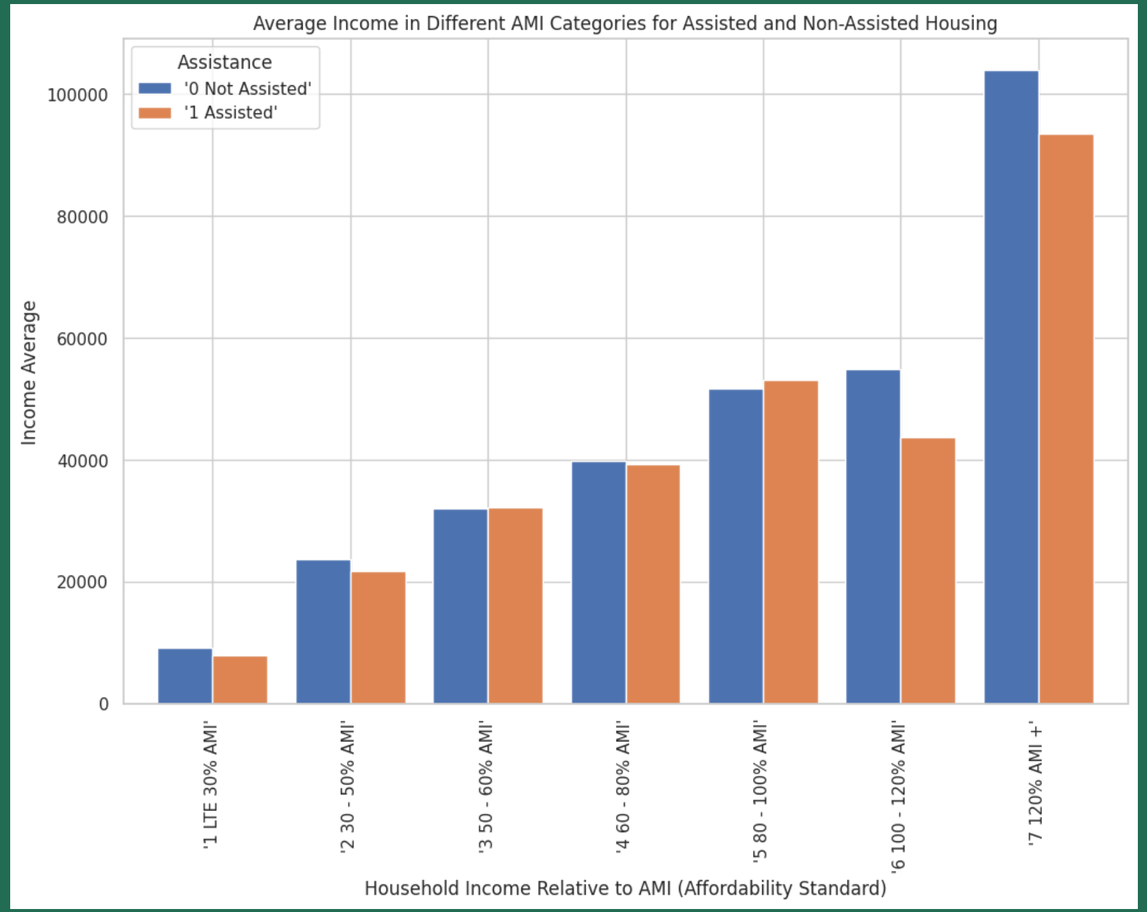
- **Similarities in Adequacy Category 1 - Adequate:** In the Adequate category, both City Central and Suburban areas showed similar average burden costs. This might indicate that households with perceived adequate housing face comparable cost burdens, irrespective of location.

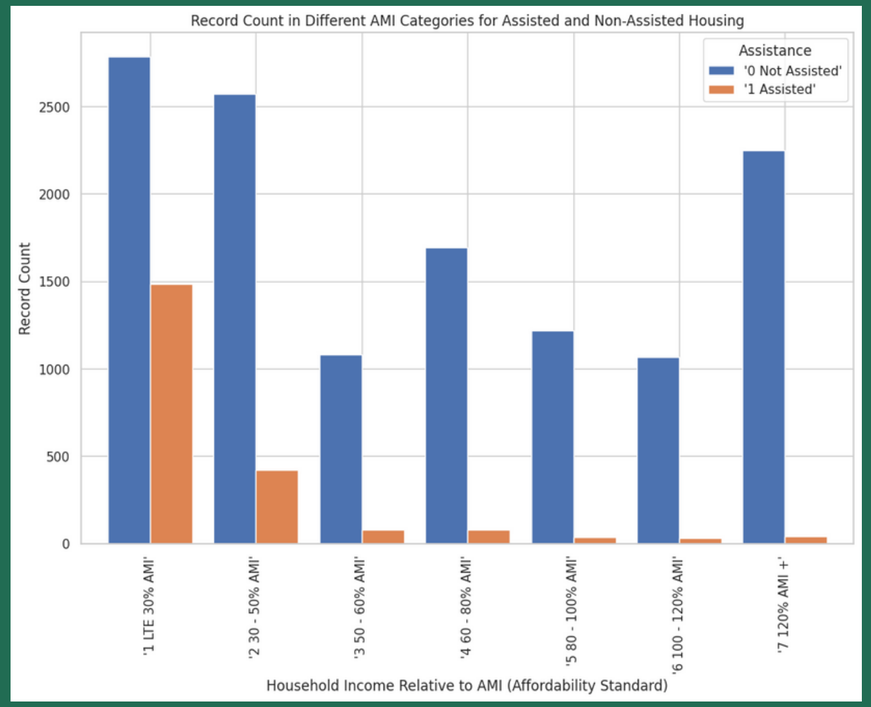
- **Possible Affordability Patterns:** The lower burden costs in Suburbs for moderately and severely inadequate housing categories may point to potential affordability advantages in these inadequacy situations. **Suburban housing might be more affordable for households facing housing inadequacies** compared to City Central.

COMPARING THE AFFORDABILITY OF ASSISTED HOUSING UNITS TO NON-ASSISTED UNITS

4 How does the affordability of assisted housing units compare to non-assisted units?







4 Insights and Interpretations

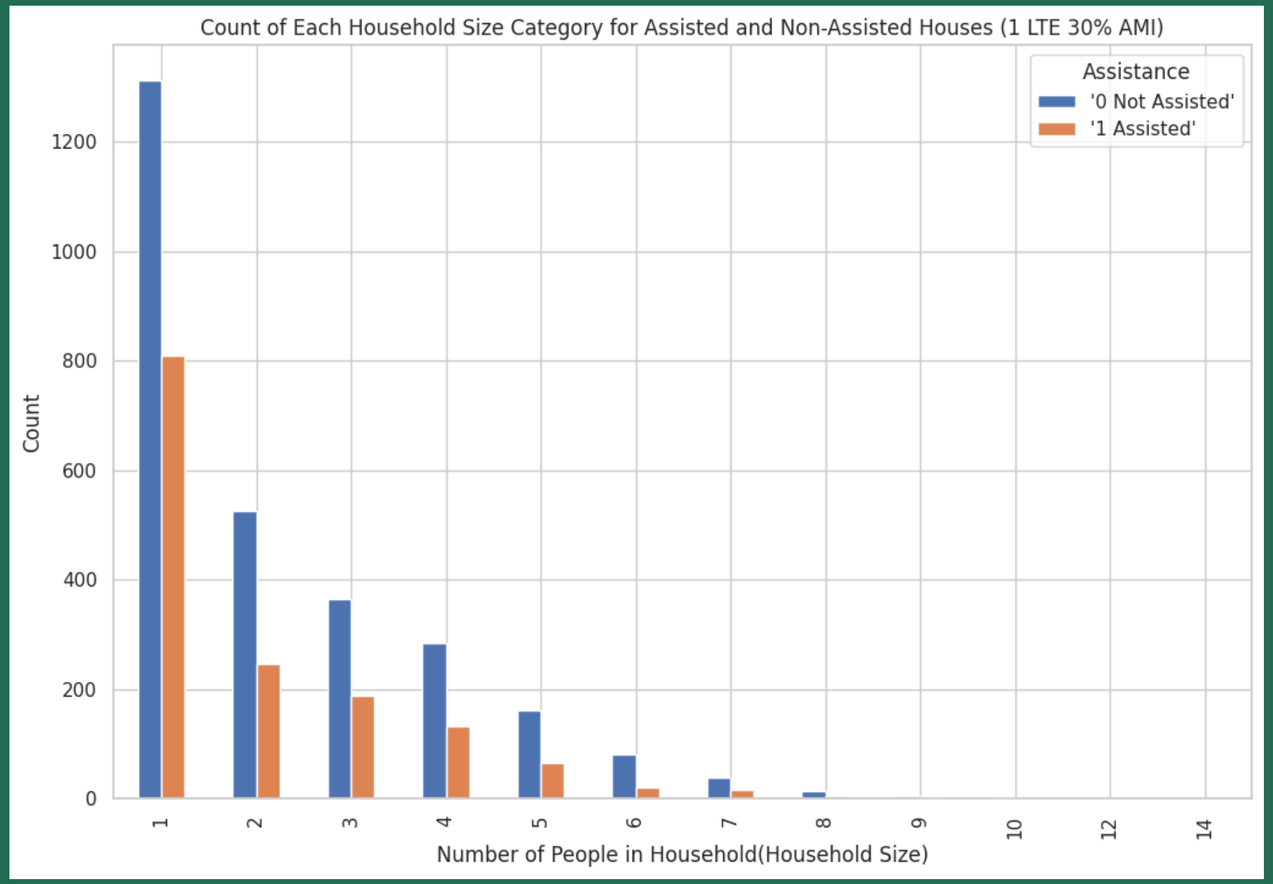
- **Count Disparity:** The first visualization reveals a significant disparity in counts between assisted and non-assisted housing units for each burden category. Non-assisted units have a considerably higher count across all burden categories, indicating a larger presence in the dataset.

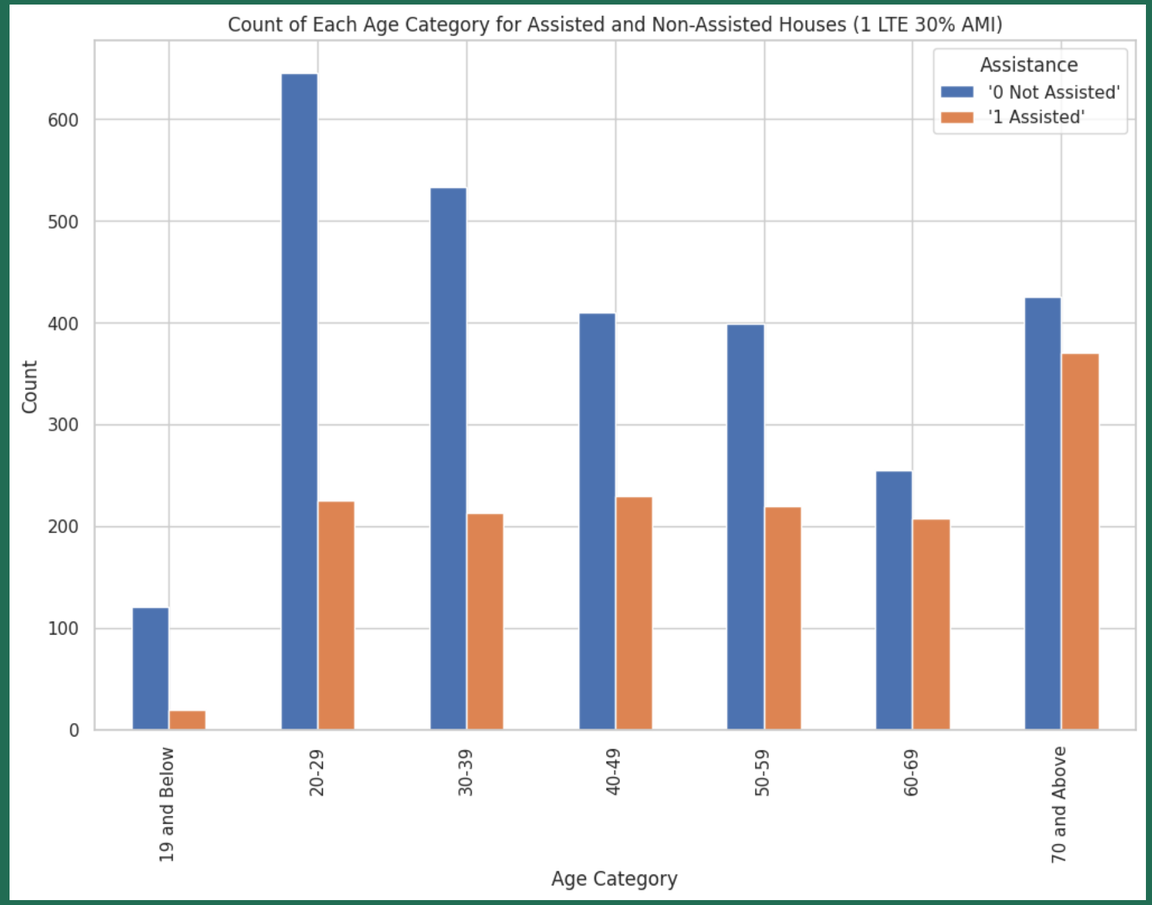
- **Similar Average Incomes:** The second visualization indicates that, on average, assisted and non-assisted households have very similar incomes. Despite a slight advantage for non-assisted households in some categories, the overall pattern shows a close resemblance in average incomes across different AMI categories.

- **Concentration in Categories 1 and 2 for Assisted Housing:** The third visualization shows that most of the assisted households fall within categories 1 (LTE 30% AMI) and 2 (30 - 50% AMI). This concentration in the lower AMI categories suggests that assisted housing is serving households with lower incomes, potentially addressing the needs of those with the greatest affordability challenges.

5

Despite similar average incomes between assisted and non-assisted housing units in both Poverty and AMI income categories, what factors contribute to the provision of assistance to these households?





5 Insights and Interpretations

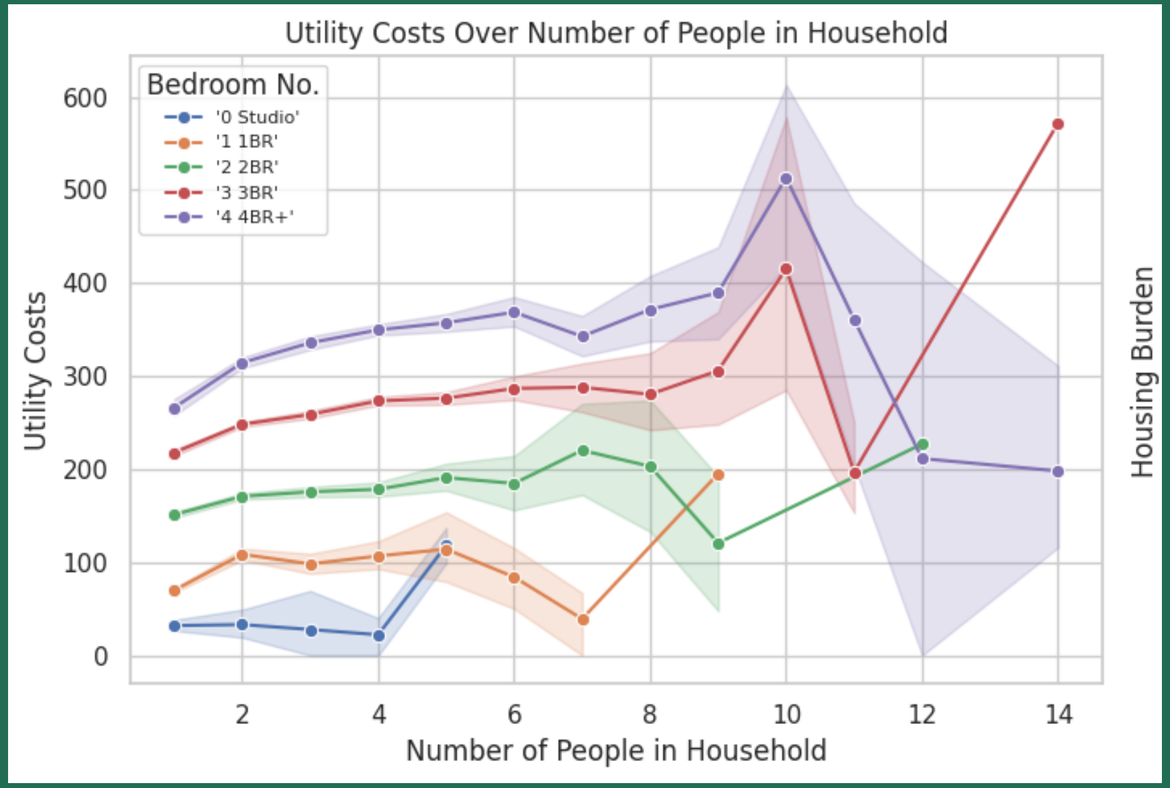
- Household Size and Assistance: The first visualization suggests that the count of both assisted and non-assisted houses decreases as the number of people in the household increases. This trend is observed in the "1 LTE 30% AMI" category. The similar pattern for both assisted and non-assisted households indicates that the number of people may not be a decisive factor in determining assistance, as the count decreases uniformly with increasing household size.

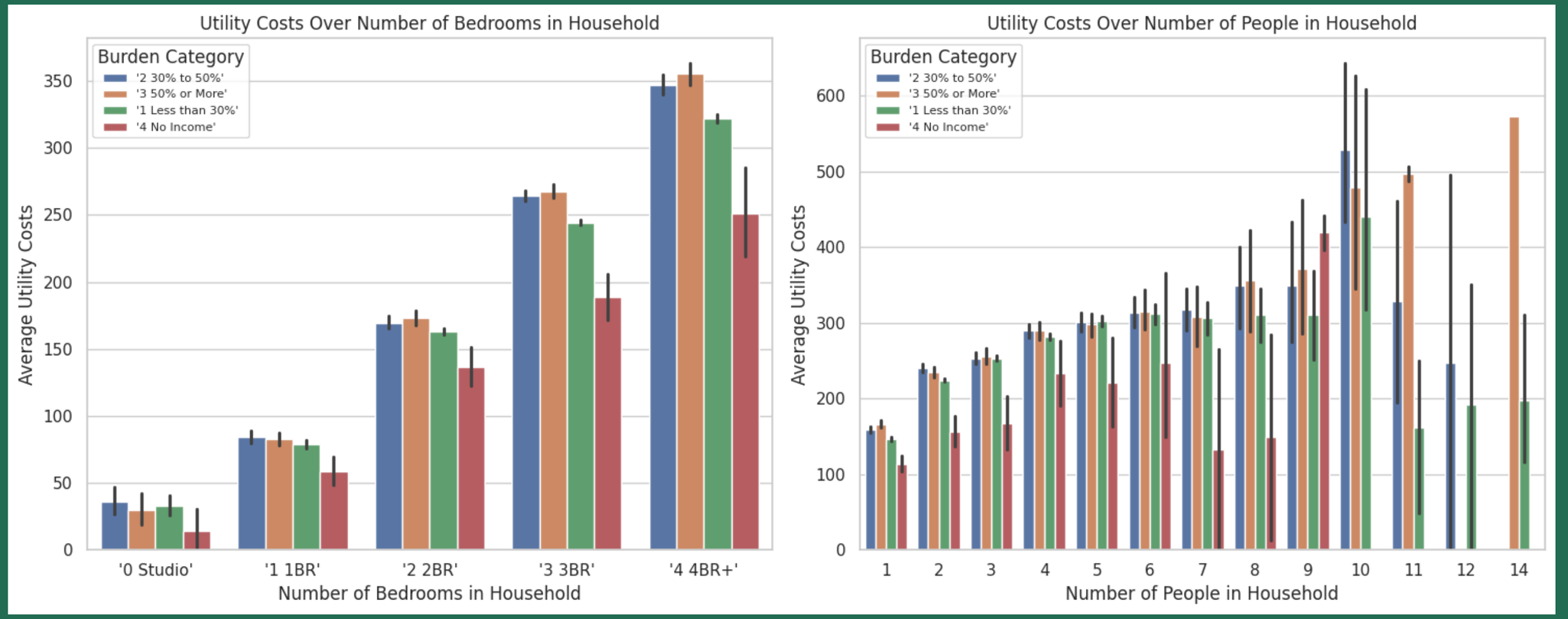
- Age and Assistance: The second visualization highlights a noteworthy distinction in the count of assisted and non-assisted houses for the age category "70 and Above" within the "1 LTE 30% AMI" income category. While the counts for ages 20 to 69 are relatively similar for assisted houses, there is a substantial increase in the count for assisted houses among individuals aged 70 and above.

- Potential Age-Related Assistance: The higher count of assisted houses for individuals aged 70 and above suggests that age may be a factor influencing the provision of assistance in the "1 LTE 30% AMI" category. This could indicate that certain assistance programs or housing initiatives may prioritize or be more accessible to elderly individuals in this income category.

ANALYZING THE RELATIONSHIP BETWEEN UTILITY COSTS AND HOUSING BURDEN BASED ON HOUSEHOLD SIZE AND AVAILABLE BEDROOMS

6 How does the relationship between utility costs and housing burden change based on the number of people in a household and the available bedrooms?





6 Insights and Interpretations

- **Utility Costs and Household Size:** The line chart indicating utility costs over the number of people with different numbers of bedrooms in the household shows a clear positive trend. As the number of people increases, along with the number of bedrooms, utility costs also increase. This suggests that larger households with more bedrooms tend to have higher utility costs, potentially due to increased consumption.

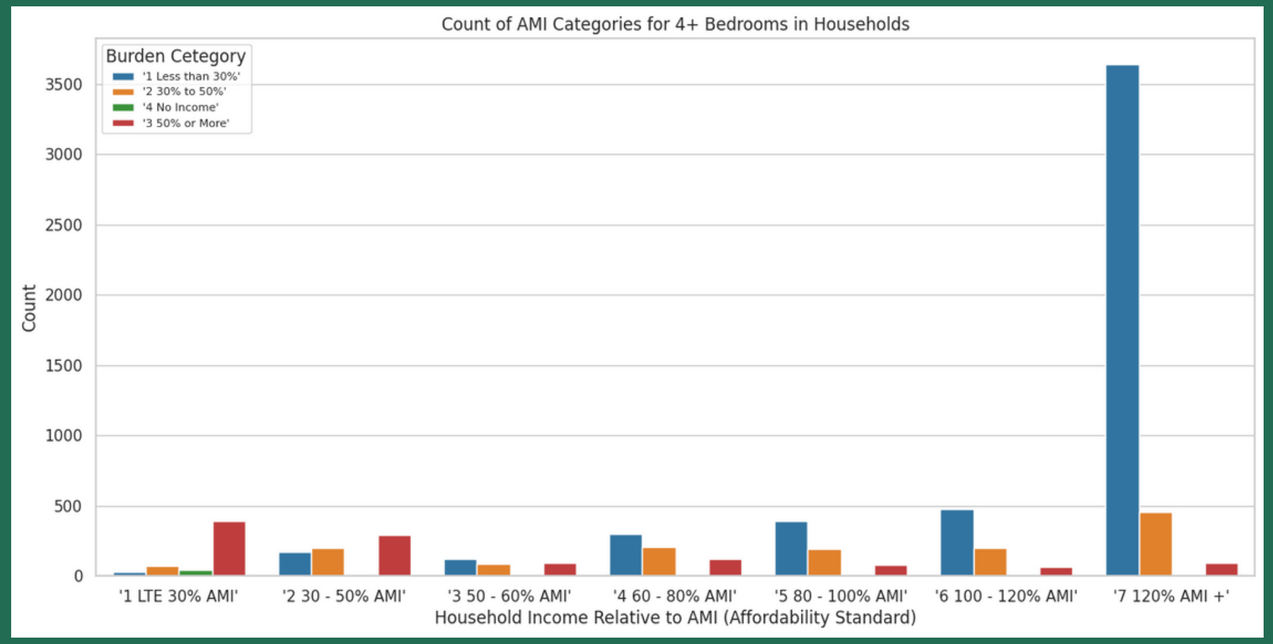
- **Average Utility Costs and Burden Categories:** The pair of bar charts illustrating average utility costs over the number of bedrooms and people in different burden categories provide valuable insights. The consistent increase in utility costs with an increase in the number of people and bedrooms aligns with common expectations. Larger households with more bedrooms typically consume more utilities, leading to higher costs.

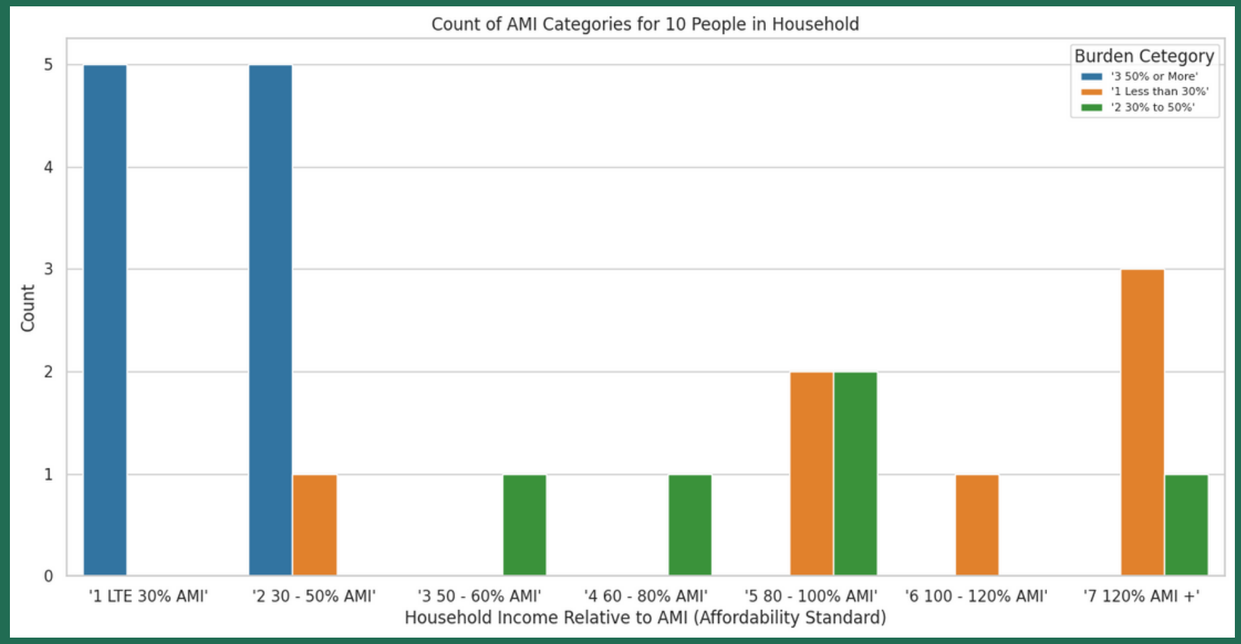
- **Direct Proportionality:** The observed direct proportionality between utility costs, household size, and the number of bedrooms underscores the interconnected nature of these factors. As households grow in size and complexity (more bedrooms), the associated utility costs rise accordingly.

- **Impact on Cost Burden:** The direct proportionality observed between utility costs and both the number of people and bedrooms also extends to the cost burden. Larger households with higher utility costs are likely to experience a higher cost burden. This emphasizes the importance of considering not only housing-related factors but also the associated utility costs when assessing affordability.

INVESTIGATING HOW HOUSEHOLD SIZE INFLUENCES THE AFFORDABILITY OF HOUSING, ESPECIALLY IN TOP BURDEN CATEGORIES.

7 How does the household size influence the affordability of housing, particularly in relation to the top burden categories associated with different household sizes?





7 Insights and Interpretations

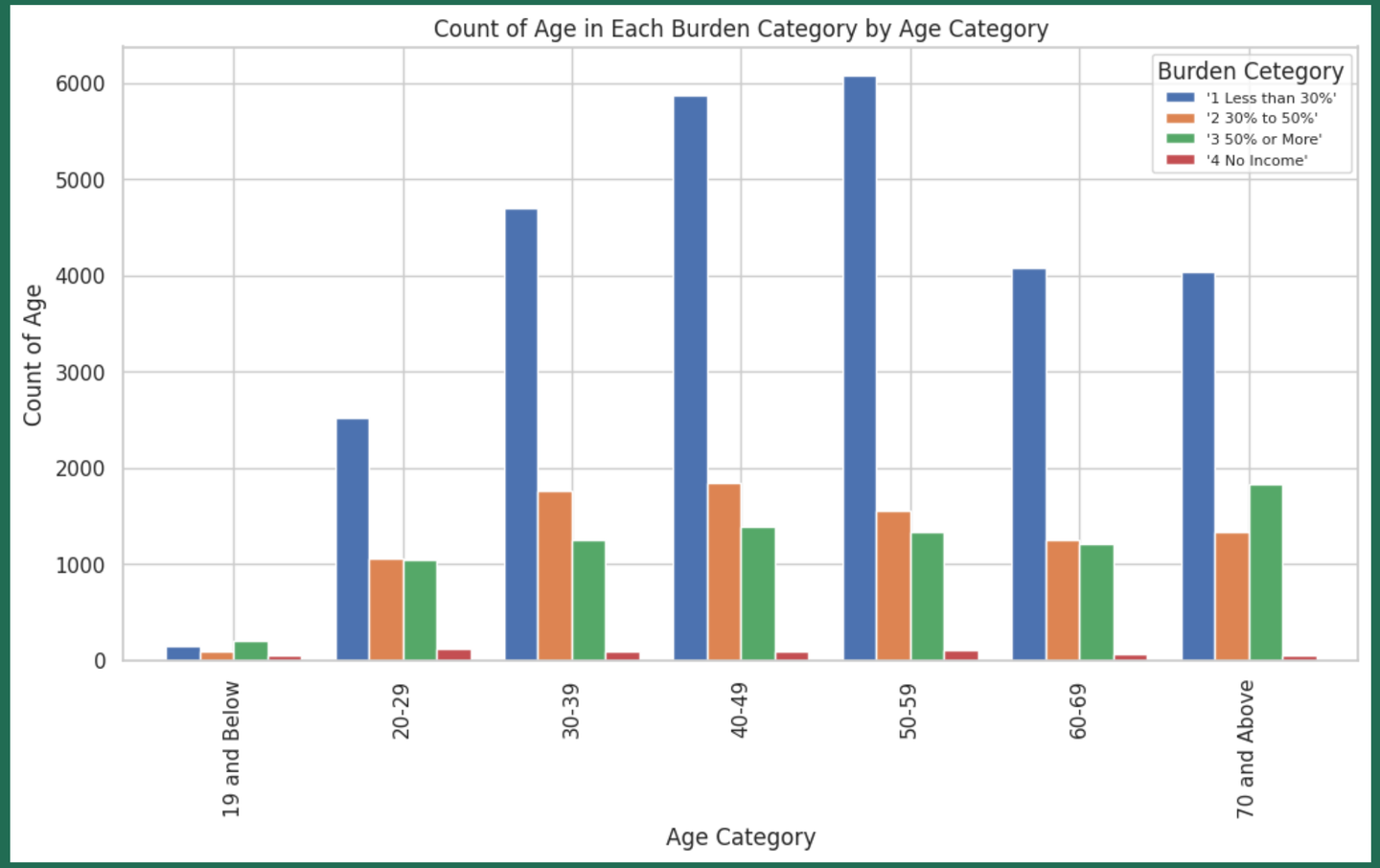
- **Affordability and Household Size in 4+ Bedrooms:** The bar chart depicting the count of AMI categories for households with 4+ bedrooms across different burden categories reveals a notable insight. Burden category 1 (less than 30%) stands out with the highest count, particularly in the AMI category 7 (120%). This suggests that larger households with 4+ bedrooms, especially those falling into burden category 1, are associated with higher income levels (AMI category 7), potentially indicating a more affluent demographic.

- **AMI Categories for 10 People in Household:** Examining the count of AMI categories for households with 10 people in different burden categories provides additional insights. Burden categories 1 and 2 exhibit the highest count in the AMI category 3 (50% or more). This finding suggests that larger households with 10 people are more likely to fall into burden categories indicating higher housing cost burdens, particularly in the mid-range AMI category.

- **Implications for Affordability Strategies:** The concentration of higher AMI categories in burden category 1 for households with 4+ bedrooms may imply that larger households in this burden category are more likely to afford housing with 4+ bedrooms. This could be attributed to the presence of more affluent households in this category.

EXAMINING HOW THE BURDEN OF HOUSING COSTS CHANGES WITH THE AGE OF THE HEAD OF THE HOUSEHOLD.

8 How does the burden of housing costs change with the age of the head of the household?



8 Insights and Interpretations

- **Dominance of Burden Category 1 (Less than 30%):** The bar chart illustrates a consistent trend across different age categories, with burden category 1 (less than 30%) having the highest count. This suggests that a significant portion of households, regardless of the age of the head, falls into the burden category with the least housing cost burden.

- **Age-Related Patterns in Burden Categories 1, 2, and 3:** The observed increase in burden categories 1, 2, and 3 from age categories 20 to 59 indicates a potential correlation between the age of the head of the household and the likelihood of experiencing varying levels of housing cost burdens. This pattern may reflect the evolving financial situations and housing needs of individuals during their working years.

- **Decrease in Burden Categories Beyond Age 59:** The subsequent decrease in burden categories beyond age 59 suggests a shift in housing cost burdens as individuals approach retirement age. This decline may be influenced by factors such as changes in income, housing choices, or assistance programs available to older individuals.

CONCLUSION

In conclusion, our comprehensive analysis of housing-related factors has shed light on nuanced patterns influencing affordability across different dimensions. From the prevalence of burden category 1, suggesting a common trend of lower housing cost burdens, to age-related variations and distinctions between assisted and non-assisted housing, our findings offer a multifaceted perspective. The interplay of utility costs with household dynamics further underscores the complexity of affordability considerations. These insights provide a solid foundation for informed decision-making by policymakers and stakeholders, enabling the development of tailored strategies that address the diverse needs of households and contribute to enhancing overall housing accessibility and affordability.