

## Project 2 - Austin Crime Report 2015

### Introduction - ([Google Slides](#) | [Project 2 Folder](#))

The domain covered in this project is criminal statistics and housing market analysis within the Austin, Texas area. Our investigation delves into the intersection of affordable housing and crime types across various Council Districts. Utilizing a comprehensive dataset that combines the 2015 Annual Crime Dataset and the 2014 Comprehensive Housing Market Analysis, we focused on understanding the housing affordability for teachers concerning crime incidence. Employing analytical techniques such as scatterplots and means, we visualized and assessed our data's distribution and central tendencies.

Our key analytical tool, the Pearson correlation, revealed statistically significant relationships ( $p\text{-value} < 0.05$ ) between housing affordability for teachers and the prevalence of specific crimes, including burglary, aggravated assault, and murder. These results underscore potential vulnerabilities in certain Council Districts where housing is more affordable for teachers. The implications of these findings could be important for urban planning and social policy development, suggesting a potential link between housing affordability and crime rates.

### Dataset -

The Austin Crime Housing 2015 dataset represents an amalgamation of two distinct data sources. It integrates the Annual Crime Dataset 2015 from [austintexas.gov](http://austintexas.gov), which details all Part 1 crimes recorded in Austin, Texas, spanning the duration of January 1 to December 31, 2015. It has been merged with the 2014 Housing Market Analysis Data by Zip Code, extracted from the City of Austin's 2014 Comprehensive Housing Market Analysis. This latter dataset provides a detailed overview of demographic and housing-related information for each zip code within Austin. Together, these datasets present a varied view of Austin's crime and housing landscape in 2015. Location attributes such as Council District, Location, Zip, and X/Y Coordinates exist and can be used to analyze crime data held in the Highest NIBRS/UCR Offense Description and GO Highest Offense Desc attributes.

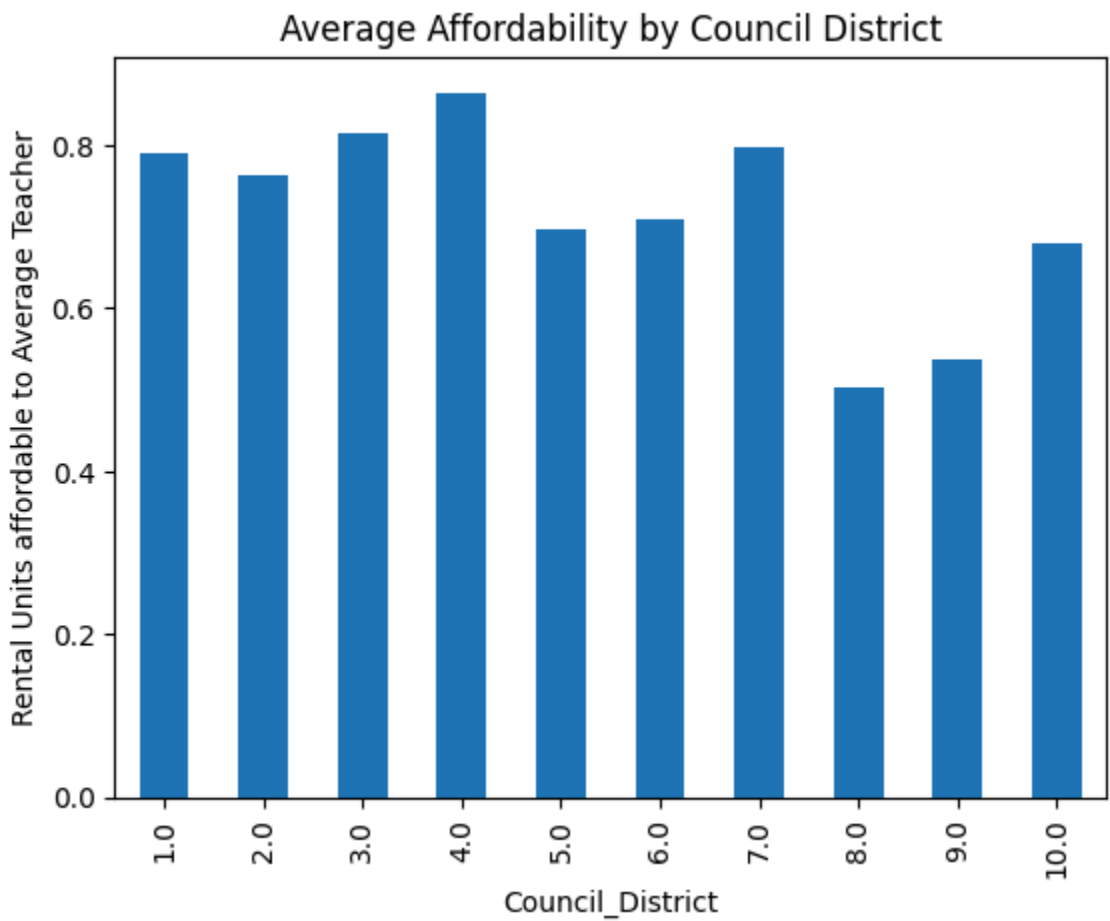
### Analysis technique -

In the Third Analysis of the Austin Crime Dataset, data transformation, aggregation, and visualization (using bar and scatter plots) alongside Pearson correlation were used to study the link between crime rates and housing affordability in Austin's Council Districts. Transforming rental affordability percentages into numerical values allowed for quantitative analysis. Aggregating this data by council district and crime type provided a

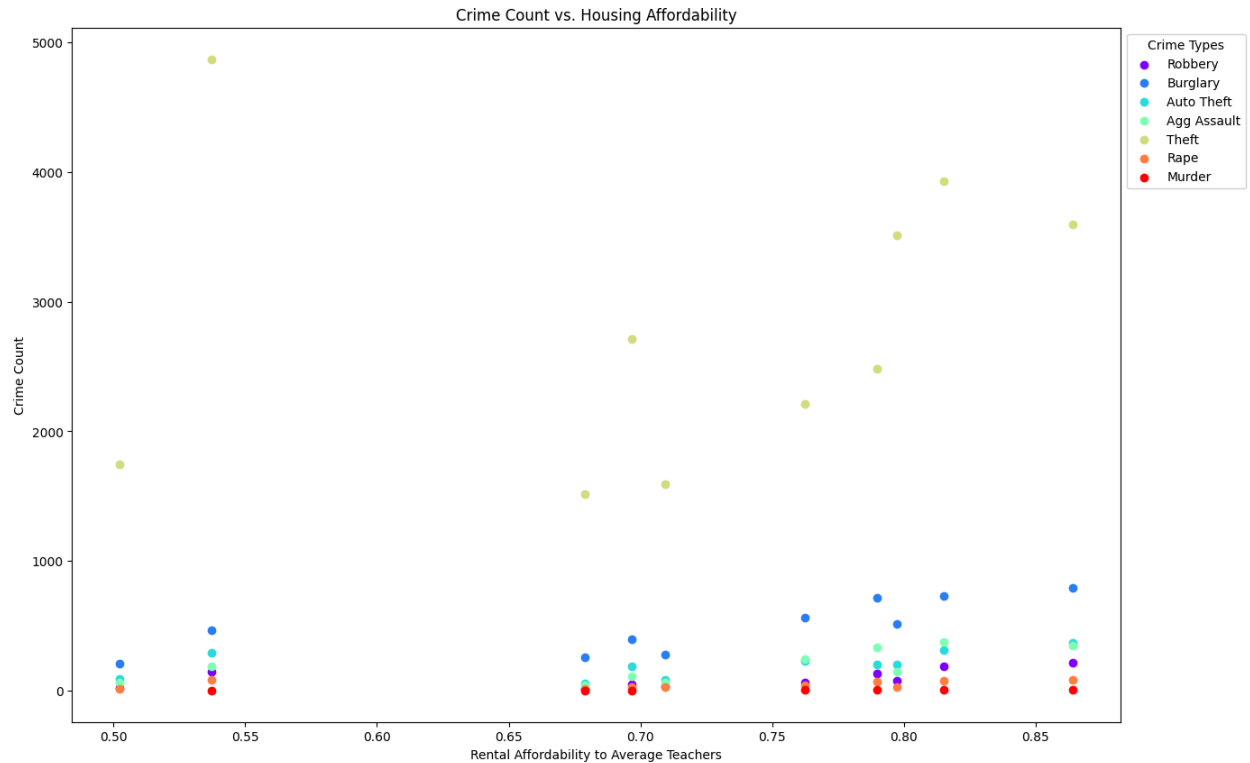
clear focus. Bar plots were used for a straightforward comparison of housing affordability across districts, while scatter plots showed the relationship between various crimes and housing affordability. A Pearson correlation was employed to measure the strength and direction of the relationship between these continuous variables. This approach was chosen for its directness and effectiveness in analyzing and interpreting the dataset, providing insights for urban planning in Austin.

**Results -**

Through our analysis, we gained insights into the relationship between crime rates and housing affordability in Austin's Council Districts. The bar plot below highlighted the variance in housing affordability across districts, revealing specific areas where affordability is higher.



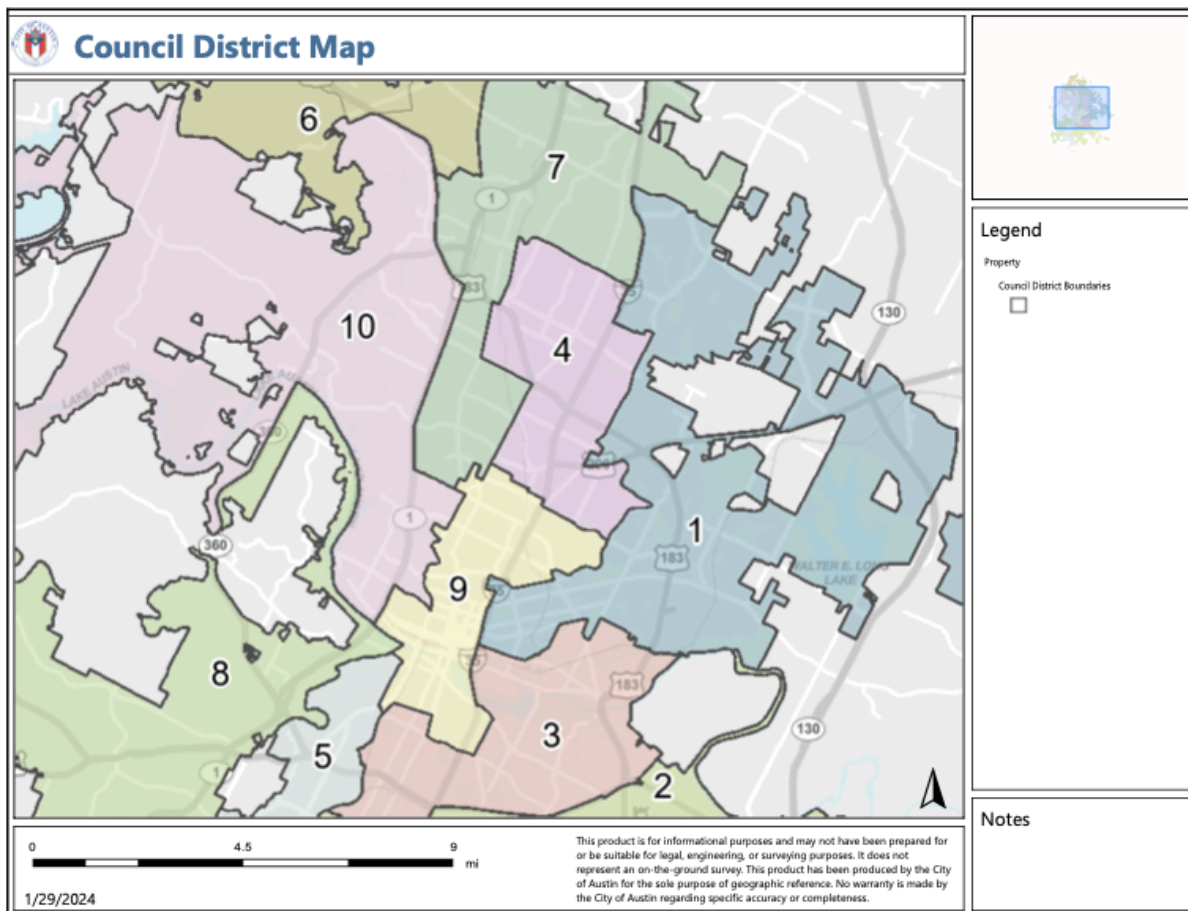
This scatter plot appears to reveal a correlation between the affordability of housing for teachers and the incidence of certain crimes. Specifically, districts with higher affordability tended to have increased crime rates.



The Pearson correlation coefficients reveal that burglary, aggravated assault, and murder all show a positive linear relationship with rental affordability for teachers in Council Districts, as evidenced by a p-value less than 0.05. This suggests a significant association between the affordability of housing for teachers and the rates of these specific crimes.

Crime Type	Statistic	P-Value
Robbery	0.4979731917297231	0.14299130987118389
Burglary	0.7524148439938884	0.012040988637993074
Auto Theft	0.4644297427752464	0.17628999052759065
Agg Assault	0.6396279144383077	0.04642298953435549
Theft	0.12502540179510488	0.7307420539041181
Rape	0.358049353615812	0.30968924312650786
Murder	0.7905787384700048	0.034312964590100706

The results indicate a clear bias in crime rates for these specific crimes in certain Council Districts. If we look at the four districts (1, 3, 4 & 7) with the most affordable rental housing for teachers we can see their geographic location likely has an impact as well.



## Technical -

The data preparation involved transforming the 'Rentalunitsaffordabletoaverageateacher' column from a percentage format into numerical values using the `percentage_to_float` function. The data was also aggregated by Council District, allowing for analysis specific to each geographical area.

The analysis techniques, including bar plots, scatter plots, and Pearson correlation, were chosen for their direct applicability to the dataset. Bar plots were used to visualize housing affordability, and scatter plots examined the relationship between crime rates and housing affordability. Pearson correlation provided a quantitative measure of the relationship between these variables. This approach allowed for a clear examination of the data, aligning with our goal of exploring crime and housing affordability relationships.

The analysis was iterative, with adjustments made to improve clarity at each step. One thing to consider in our Pearson correlation is the effect of outliers on the reliability of the test. In our initial scatterplot comparing affordability to crime rates there appeared to be significant outliers in the 54% range. This could potentially weaken the validity of our results or indicate alternate relationships between crime and affordability.