Dat 121 Term Paper

Group #80

Jess Piard [*501677*], Safiya Alexander [*510508*], Salman Yonis [*510477*]

Factors Influencing Housing Ownership In Missouri Counties

Introduction

In the US, the housing market and wealth distribution are relevant topics that currently influence economic policies and social attitudes toward homeownership. This is especially relevant in MO, where its history of slavery and redlining has had a lasting impact on the generational wealth of its black residents. By analyzing variables like, the percentage of Black homeowners, the percentage of homeowners with a bachelor's degree or higher, and age (35-54) we seek to uncover how socioeconomic factors and specific demographics shape access to housing in the state. Additionally, by looking at each county in MO individually, we will produce further research on the state of the current housing market in specific areas. We intend to contribute data supporting already researched factors (race and education level) while contributing a new perspective on the correlation between certain areas in Missouri and access to housing.

The perceived relationship between homeownership and demographic factors continues to be a significant area of inquiry. Identifying which variables influence homeownership and how they do so is crucial in order to better understand the inequalities that exist and make more informed decisions when formulating policies to address them. A study conducted by Patrick Bayer, Fernando Ferreira, and Stephen Ross titled “What Drives Racial and Ethnic Differences in High-Cost Mortgages? The Role of High-Risk Lenders," explores the role of race in mortgage

lending in seven diverse metropolitan areas from 2004-2007. Their research concluded that racial and ethnic minorities, particularly Black and Hispanic households, face unique barriers to homeownership due to discriminatory lending practices and "elevated exposure to subprime mortgage products." This study points to the structural inequities within the housing finance system, which disproportionately affects minority groups, making race an important factor to consider when evaluating homeownership disparities.

A more contemporary study conducted by Goodman, Kaul, and Zhu titled "The Impact of Tight Credit Standards on 2009–2015 Lending," examines how educational level can influence credit access. Their findings support the assumption that higher educational attainment is associated with higher credit scores, which in turn increases the probability of homeownership. This reinforces the idea that education plays an important role in enhancing financial literacy and income potential and thus remains a key determinant of homeownership.

Our hypothesis has many parts: First, we believe our data will support that there is a positive statistically significant relationship between having higher education and older age and increased homeownership. On the contrary, we hypothesize that our dependent variable black homeowners, will show that this demographic will have decreased rates of homeownership in many counties in MO. Identifying which variables influence homeownership and how they do so is crucial in order to better understand the inequalities that exist and make more informed decisions when formulating policies to address them.

Key Results

Through the creation of our model, our team found key implications about variables and their influence on homeownership. We were ultimately able to explore these multi-variable

relationships and test our hypotheses with scatter plots. The scatter plots in section b of the appendix highlight two key findings from our analysis: the relationships between homeownership rates and both median age and the rate of Black homeowners. The first scatter plot, examining homeownership vs. median age, shows a clear positive relationship with a trendline equation of $y=0.1505x+4.104$ and an R^2 value of 0.2939. This finding highlights the role of age in housing access, as older populations are often more financially stable and better positioned to own homes. The second scatter plot, showing homeownership vs. the rate of Black homeowners, also depicts a positive trend ($y=0.1786x+8.4461$) but with a weaker R^2 value of 0.1951. While this correlation is modest, it highlights the importance of representation in understanding demographic disparities in homeownership.

Overall, the results provide critical insights into how demographic factors shape homeownership trends. Education emerges as the most significant predictor, reiterating its central role in improving housing access. Age also demonstrates a meaningful influence, reflecting the stability and financial capacity associated with middle age. Additionally, the percentage of Black homeowners is significantly and positively associated with overall homeownership rates, highlighting the importance of representation within demographic contexts. This analysis emphasizes the value of considering these variables in future research and housing policies, offering a foundation for addressing disparities in the housing market.

Data

Dependent Variable: Home Owners per County

Independent Variables: % of Black Homeownership, % of Bachelor's Degrees or Higher, Homeownership % by age range 35-54

Our study draws primarily on data from one reputable source: the U.S. Census Bureau. This is a government-run institution and is widely used by policymakers, making the data highly credible and useful for analyzing homeownership against demographic factors. Despite the credibility of our source, there are some criticisms that could be raised. The Census Bureau collects data in set intervals, resulting in slight delays between the data collection and publication. As a result, some might argue that the data doesn't *fully* reflect more recent economic shifts. Another potential critique is regarding the level of granularity available in the data. Although it offers state-level data, it can be argued that more localized data is often more variable and possibly less accurate.

US Census Bureau Data sets Used:

- *PEPANNRES Annual Estimates of the Resident Population: April 1, 2010 to July 1, 2019*
- *S2502 Demographic Characteristics for Occupied Housing Units (2022: ACS 5-Year Estimates Subject Tables)*

Unemployment rates per county, poverty rates per county, and loan denial rates by race per county are a few supplemental data sets we would like to have in the future to further advance the scope and implications of our findings. To continue, as shown in section (a) of the Appendix, our descriptive statistics further suggest the correlation between our independent and dependent variables. Some key stats include the mean of *** and standard deviation of *** for our dependent variable % of black homeownership. The mean for age is 30.5 and the standard deviation is 3.96 which means that the standard deviation is 12% of the mean - which suggests fairly low variability, the mean and standard deviation of % Black homeowners is relatively large which suggest high variability at (1.39, 2.72) and mean and standard deviation of % with a Bachelor's degree or higher are , (24.02, 9.08) respectively. Lastly the mean and standard deviation of the 35-54 age group are (27.54 and 3.10). It also should be noted that the scatter

plots show positive linear relationships between our independent variable of homeownership and our chosen dependent variables. All variables excluding %Black homeownership have low variability and demonstrate statistically significant relationships with the dependent variable homeownership, with this, we can verify that the data is internally consistent and the descriptive statistics make sense.

Modeling

Model 1

Running the first model, we used the Multiple Linear Regression modeling technique, and the model we developed is as follows:

Homeownership Rate in Missouri = 222.3027552 + 2417.8908 (Age (35-54)) --60385.66151 (%Black Homeowners) -3062.620792 (%Bachelors or Higher) + 0.285152415 (Total Population)

Metrics: The overall model was significant at Significance F 4.704E-119. The R^2 was extremely high at .99344073 which raised immediate suspicions. The standard error of the mean is .195 which indicates an intermediate variability to the mean. The coefficients for %Black Homeownership (-60385.66151) and %Bachelors degree or higher (-3062.620792) were negative and abnormally large indicating a need for a potential rescaling of the data. Two out of the four Independent variables – %Black Homeowners (p-value = 9.6885E-10) and Total Population (p-value= 2.702E-107) were significant. The other two independent variables Age 35-54 (p-value = 0.76006357) and %Bachelors or Higher (p-value= 0.44742123) were not significant. The correlation table shows that while the r_{xy} for % Bachelor's Degree or higher and

Age 35-54 are above .2 *the threshold for when for when n=100 and therefore are sufficient to reject the Ho*, they are not significant in the Multiple Linear Regression -which suggests multicollinearity. Total population while significant in both the MLR and the Correlation tables is above .7 on the correlation table at .99527316 which is also indicative of multicollinearity. Please see table 7 and table 6in the appendix to view the MLR and the correlation table for model 1 respectively. In addition to multicollinearity, this model's Descriptive Statistics suggested a need to rescale the data as the standard deviation for both the Total Population and the dependent variable, Homeownership were significantly large. Please see table 3 to view the descriptive statistic for model 1.

Our hypothesis argues that there is a statistically significant linear relationship between Homeownership and age (35-54), race(black), education (bachelor's degree or higher), and total population per county, therefore the MLR is the appropriate model. We validate the assumption through homoscedasticity as all p-values are less than $\alpha=0.05$. While our model shows the linear relationships between variables, -a disadvantage is that if there is a nonlinear relationship with one of the variables we won't be able to analyze it with an MLR. Additionally, MLR can be highly impacted by outliers and leverage points making it harder to discern the relationship between the variables especially if the data is nonlinear. The advantage of having a nonlinear model, however, is contrasted by the disadvantage of selecting the right nonlinear model as it can be a challenge. Ultimately our hypothesis suggested that there was a linear relationship and MLR was therefore the selected model, but the outcome of skewed data and multicollinearity proved a need to rescale our variables.

Model 2

For Model 2 we rescaled the data by taking the rates of each of the dependent variables (age, race, and bachelors degree or higher) and multiplying by 100 to convert the decimals into percentages. Following this, to minimize skew, we replaced population with rate of homeownership by dividing homeownership per county by the total population per county and took the natural log of the dependent variable to normalize the distribution. After all of these changes, all variables in addition to the model were statistically significant, this model shows that the relationship between the dependent variable and the independent variable is linear when the dependent variable is transformed using LN. The model we developed is as follows:

Homeownership Rate in Missouri = 3.292290362 + 0.0744(Age (35–54)) +0.0584 (%Black Homeowners + 0.0586 (%Bachelors or Higher) + 0.0594(Rate of Homeowners)

Metrics: After developing our regression model in excel, we then used SPSS to elevate our model and further expand on our findings. The model's overall statistical significance is supported by an F-statistic of 2.04749E-19. The R squared for this model is .57 which is ok, meaning that 57.1% of the variation in homeownership rates is explained by the four independent variables we chose to explore: percentage of Black homeowners, percentage of individuals with at least a bachelor's degree, median age (35-54), and rate of homeownership. An adjusted R-squared value of 0.555 further confirms the model's reliability while accounting for the number of predictors included. All coefficients are positive, and all p-values are less than $\alpha=0.05$, this model is statistically significant. Please see table 8 for the MLR for model 2. The standard error of the mean is 0.085551215 which is good because it shows a relatively low variability to the mean. In the correlation table for model 2, depicted as table 5 in the appendix, the relationship between the dependent variable Homeownership and the rate of Homeowners

has a correlation of (- 0.032289149) and the scatter plot in figure 4, representing the relationship between those two variables, has an R^2 of less than .001. From these results we have concluded that this variable rate of homeownership (homeownership in the county /total population) as an effector of the dependent variable homeownership is somewhat redundant, however, we chose to keep it in the model as it is statistically significant in the MLR with a p-value of 0.0141086250.

In our analysis, we identified an outlier with a leverage value exceeding the threshold of 4 and a Cook's Distance value that was relatively high, indicating its influence on the model. It became evident that this observation in St.Louis City in particular was disproportionately affecting the regression results, potentially misrepresenting the relationships between the independent and dependent variables. This outlier can also be seen in the scatterplot for Homeownership and %Black homeowners, indicating that St.Louis City has a much larger percentage of black homeowners. To ensure the integrity and soundness of our findings, we removed this data point, which led to improved significance and stability across the model's predictors.

Model 3

After removing the data point for St.Louis City, Mo, our final model shows a strong fit, with an R-squared value of 0.561, meaning that 56.1% of the variation in homeownership rates is explained by the three independent variables we chose to explore: percentage of Black homeowners, percentage of individuals with at least a bachelor's degree, and median age. An adjusted R-squared value of 0.545 further confirms the model's reliability while accounting for the number of predictors included. The model's overall statistical significance is supported by the

Significance F 1.01618E-18, further exemplifying its strength in identifying the key patterns in the data.

The coefficients provide further valuable insight into how these specific variables influence homeownership levels. The percentage of Black homeowners has a coefficient of 0.085, indicating a statistically significant positive relationship with homeownership rates, also represented by a p-value of 0.004. This suggests that higher representation of Black homeowners in a county is associated with slight increases in the overall homeownership rate. Educational attainment, measured as the rate of individuals with a bachelor's degree or higher, remains the strongest predictor in the model. With a coefficient of 0.073 and a p-value of less than 0.001, this variable demonstrates that higher levels of education are strongly associated with increased homeownership rates. The rate of homeowners also exhibits a positive relationship with homeownership rates, with a coefficient of 0.051 and a p-value of 0.041, indicating statistical significance. While its effect is weaker than the other variables, this finding suggests that the general rate of homeownership in a county contributes to higher overall homeownership. Finally, the rate of individuals aged 35–54 shows a statistically significant positive relationship with homeownership rates. With a coefficient of 0.073 and a p-value of less than 0.001, this variable indicates that a one unit increase in the percentage of this age group is associated with a 0.073 increase in the homeownership rate, holding all other variables constant. This finding supports the hypothesis that a higher representation of middle-aged individuals in a county positively influences homeownership, likely reflecting the life stage factors associated with housing stability and financial capacity.

To ensure the accuracy of our results, we validated the regression model by evaluating its assumptions. Multicollinearity was not an issue, as shown by all the Variance Inflation Factor

values being under 4. Residual analysis showed no notable patterns or deviations from normality, strengthening the model's reliability. Cook's Distance values were also below 1 across all observations, confirming that no single data point had an extreme influence on the results. These diagnostic checks indicate that the model meets the necessary assumptions for interpreting the findings confidently.

Economic Significance

The coefficient for Age (35–54) is 0.072555039. This indicates that a one-unit increase in the percentage of individuals aged 35–54 is associated with a .0725 increase in Homeownership, on average, holding all else constant. This suggests a positive relationship between Age (35–54) and Homeownership, meaning that counties with a higher proportion of individuals in this age range tend to have more homeowners.

The coefficient for % Black Homeowners is.0.085773136. This indicates that a one-unit increase in the percentage of Black homeowners is associated with a .0858 increase in Homeownership, holding all else constant. This suggests a positive relationship between % Black Homeowners and Homeownership, meaning that as the percentage of Black homeowners increases, homeownership also increases.

The coefficient for % Bachelors or Higher is 0.05765802. This indicates that a one percentage point increase in the proportion of individuals with a bachelor's degree or higher is associated with a 0.057 increase in homeownership , on average, holding all else constant.This suggests a positive relationship between %Bachelors or higher and Homeownership.

The coefficient for Rate of Homeowners is.0.061493403. This indicates that a one percentage point increase in the rate of homeowners is associated with a 0.0615 unit increase in the homeownership, on average, holding all else constant. This suggests a positive relationship between Rate of Homeowners and Homeownership. These coefficients give us the tools to analyze how much a unit increase in each independent variable could impact the dependent variable. Taking these stats into account, our data implies that when targeted, our chosen independent variables have the ability to significantly influence the housing market.

Summary

Ultimately, after reflecting on all three versions of our model, including analyzing leverage points, and significance levels, we believe our data supports the validity of the majority of our original theory that there is a direct correlation between having higher education and older age and increased homeownership. However, as we rescaled the data in Models 2 and 3 we found that in the MLR there is a positive relationship between the natural log of the dependent variable and %Black Homeownership. This contradicts our hypothesis as this means that as %Black Home Ownership increases, the dependent variable increases holding all other factors constant.

Our research and results highlight several key implications about our three dependent variables in relation to the rates of homeowners in each county in Missouri. Beginning with the percentage of homeowners with a bachelor's degree or higher: The strong positive correlation between higher educational attainment and homeownership suggests that improving access to education and financial literacy can be an effective strategy for increasing homeownership rates. Moving on, we also see that there is a positive association between age and homeownership that backs claims that the older population, specifically the “Baby Boomer” and “Gen X” age groups

are more likely to be homeowners. This indicates the importance and continued need for policies aimed at younger populations, such as first-time homebuyer programs. Lastly, although the relationship between race and homeownership was positive in our final model, historical and ongoing inequalities in lending practices and housing access for Black residents emphasize the continued need to address racial disparities in homeownership, it is possible that exploring the data set using a logarithmic model on the dependent variable affected the final results, as in model one the coefficient on the MLR is negative at -60385.66151, indicating that there is a negative statistically significant linear relationship between homeownership and race (%Black), and it may be beneficial in the future to explore additional techniques.

On a managerial level, this analysis provides key takeaways that leaders in real estate and financial services can implement in firm strategy. For example, banks and credit unions can design tailored lending programs that target our highlighted demographics. Similarly, real estate developers can also take these insights to enter new markets and implement new affordable housing options. From a business standpoint, our research can provide a “win-win” scenario where firms can financially benefit while positively supporting disadvantaged communities.

There are many lessons to be learned through the methods we used, including how to effectively communicate and interpret data into actionable insights for others. This is especially true for our scatter plots and descriptive statistics, where we applied new skills to highlight complex relationships in our multivariable model.

In summation, our findings emphasize that in order to address disparities in homeownership, businesses, policymakers, and community organizations need to first tackle socioeconomic inequality. When looking at the final implications, it becomes clear that economically disadvantaged communities have high systematic barriers that reduce overall

housing rates for black citizens, those without higher education, and young people alike. However, if effectively utilized, these insights can support the development of new interventions that address said disparities in order to foster greater economic equity.

Appendix A:

a) *Table 1*: Descriptive Statistics (model 2)

Dependent Variable		*Age (35-54) * 100*		*% Black Homeowners * 100*		*% Bachelors or Higher * 100*		*Rate of Homeowners *100*	
Mean	8.71573898	Mean	30.55273025	Mean	1.632862308	Mean	24.24058699	Mean	27.487641
Standard Error	0.104244781	Standard Error	0.370841242	Standard Error	0.347813879	Standard Error	0.87172355	Standard Error	0.292988922
Median	8.591001119	Median	30.83192526	Median	0.39978678	Median	21.88926941	Median	27.55326965
Mode	7.862497197	Mode	#N/A	Mode	0	Mode	#N/A	Mode	#N/A
Standard Deviation	1.117900737	Standard Deviation	3.976829273	Standard Deviation	3.729888315	Standard Deviation	9.348193625	Standard Deviation	3.14195615
Sample Variance	1.249702058	Sample Variance	15.81517106	Sample Variance	13.91206685	Sample Variance	87.38872405	Sample Variance	9.87188845
Kurtosis	1.282039127	Kurtosis	0.082860039	Kurtosis	27.78180884	Kurtosis	2.700873402	Kurtosis	0.39427031
Skewness	0.937752986	Skewness	-0.219291812	Skewness	4.754070283	Skewness	1.559172329	Skewness	-0.386317416
Range	6.126738454	Range	22.36763594	Range	28.8844218	Range	47.40710254	Range	17.60373778
Minimum	6.426488457	Minimum	17.93007961	Minimum	0	Minimum	10.70811744	Minimum	17.0205486
Maximum	12.55322691	Maximum	40.29771555	Maximum	28.8844218	Maximum	58.11521998	Maximum	34.62428638
Sum	1002.309983	Sum	3513.563978	Sum	187.7791654	Sum	2787.667504	Sum	3161.078715
Count	115	Count	115	Count	115	Count	115	Count	115
Confidence Level(95.0%)	0.206508109	Confidence Level(95.0%)	0.734633644	Confidence Level(95.0%)	0.689016615	Confidence Level(95.0%)	1.726877639	Confidence Level(95.0%)	0.580408797

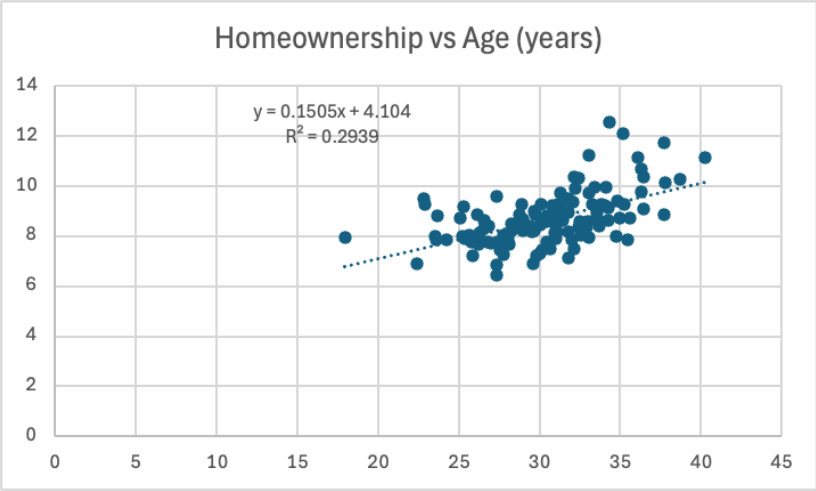
Table 2: Descriptive Statistics (model 3)

Dependent Variable (LN)	*Age (35-54) * 100*		*% Black Homeowners * 100*		*% Bachelors or Higher * 100*		*Rate of Homeowners *100*		
Mean	8.69509338	Mean	30.5075345	Mean	1.39381354	Mean	24.01787366	Mean	27.54147989
Standard Error	0.10308019	Standard Error	0.37131988	Standard Error	0.25487236	Standard Error	0.850219039	Standard Error	0.290537288
Median	8.58171989	Median	30.8305675	Median	0.39257162	Median	21.78509584	Median	27.56854634
Mode	7.8624972	Mode	#N/A	Mode	0	Mode	#N/A	Mode	#N/A
Standard Deviation	1.10059524	Standard Deviation	3.96461145	Standard Deviation	2.72129211	Standard Deviation	9.077855215	Standard Deviation	3.102089363
Sample Variance	1.21130987	Sample Variance	15.718144	Sample Variance	7.40543074	Sample Variance	82.4074553	Sample Variance	9.622958415
Kurtosis	1.48280925	Kurtosis	0.11387466	Kurtosis	14.254364	Kurtosis	3.06822416	Kurtosis	0.492735042
Skewness	0.96073569	Skewness	-0.20827749	Skewness	3.57447453	Skewness	1.595857145	Skewness	-0.388254649
Range	6.12673845	Range	22.3676359	Range	15.4713488	Range	47.40710254	Range	17.60373778
Minimum	6.42648846	Minimum	17.9300796	Minimum	0	Minimum	10.70811744	Minimum	17.0205486
Maximum	12.5532269	Maximum	40.2977155	Maximum	15.4713488	Maximum	58.11521998	Maximum	34.62428638
Sum	991.240645	Sum	3477.85893	Sum	158.894744	Sum	2738.037597	Sum	3139.728707
Count	114	Count	114	Count	114	Count	114	Count	114

Table 3: Descriptive Statistics (model 1)

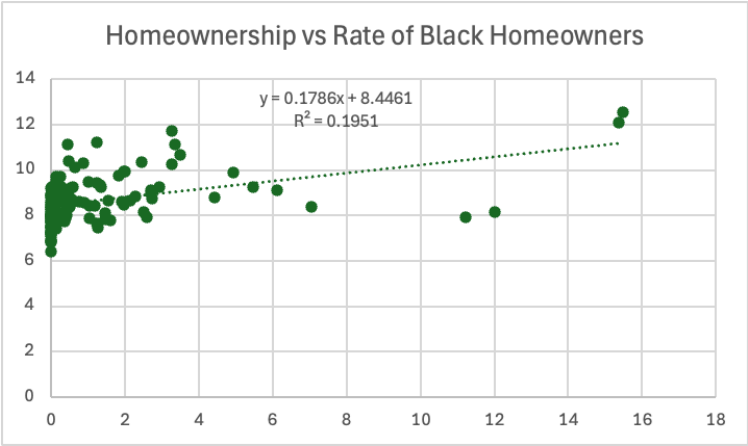
Dependent Variable		*Age (35-54)*		*% Black Homeowners*		*% Bachelors or Higher*		*Total POP*	
Mean	14450.9043	Mean	0.305527302	Mean	0.01632862	Mean	0.24240587	Mean	53368.9391
Standard Error	3194.12588	Standard Erro	0.003708412	Standard Erro	0.00347814	Standard Erro	0.00871724	Standard Erro	11639.2939
Median	5383	Median	0.308319253	Median	0.00399787	Median	0.21889269	Median	18302
Mode	2598	Mode	#N/A	Mode	0	Mode	#N/A	Mode	#N/A
Standard Deviation	34253.184	Standard Devi	0.039768293	Standard Devi	0.03729888	Standard Devi	0.09348194	Standard Devi	124817.521
Sample Variance	1173280615	Sample Variar	0.001581517	Sample Variar	0.00139121	Sample Variar	0.00873887	Sample Variar	1.5579E+10
Kurtosis	37.6243173	Kurtosis	0.082860039	Kurtosis	27.7818088	Kurtosis	2.7008734	Kurtosis	34.4929914
Skewness	5.66469002	Skewness	-0.219291812	Skewness	4.75407028	Skewness	1.55917233	Skewness	5.45633664
Range	282389	Range	0.223676359	Range	0.28884422	Range	0.47407103	Range	992192
Minimum	618	Minimum	0.179300796	Minimum	0	Minimum	0.10708117	Minimum	2013
Maximum	283007	Maximum	0.402977155	Maximum	0.28884422	Maximum	0.5811522	Maximum	994205
Sum	1661854	Sum	35.13563978	Sum	1.87779165	Sum	27.876675	Sum	6137428
Count	115	Count	115	Count	115	Count	115	Count	115
Confidence Level(95.0%)	6327.53877	Confidence Le	0.007346336	Confidence Le	0.00689017	Confidence Le	0.01726878	Confidence Le	23057.3515

b_1) *Figure 1:* Scatter Plot (Homeownership vs Age)



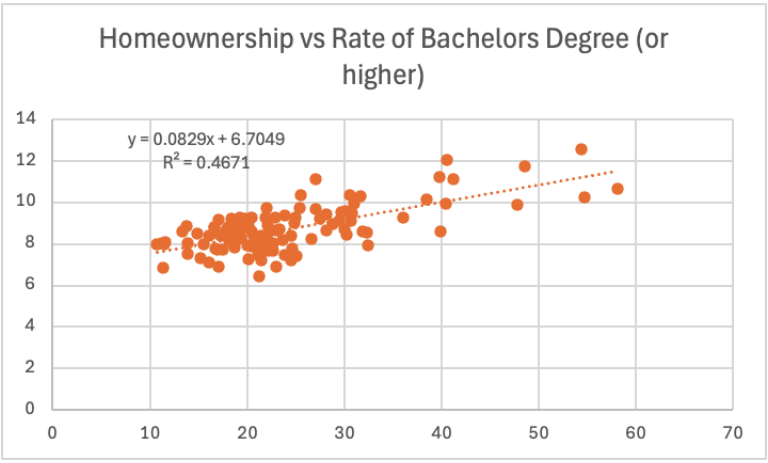
(This scatter plot illustrates the relationship between homeownership rates and median age across counties. The positive slope of the trendline, represented by the equation y=0.1505x+4.104, indicates a positive relationship between age and homeownership rates. As the median age increases, homeownership rates tend to rise. The R^2 value of 0.2939 suggests that approximately 29.39% of the variation in homeownership rates can be explained by changes in median age. While the relationship is linear, there is noticeable scatter around the trendline, implying other factors may also influence homeownership. There appear to be no significant outliers that drastically deviate from the trend)

b_2) *Figure 2:* Scatter Plot (Homeownership vs Rate of Black Homeowners)



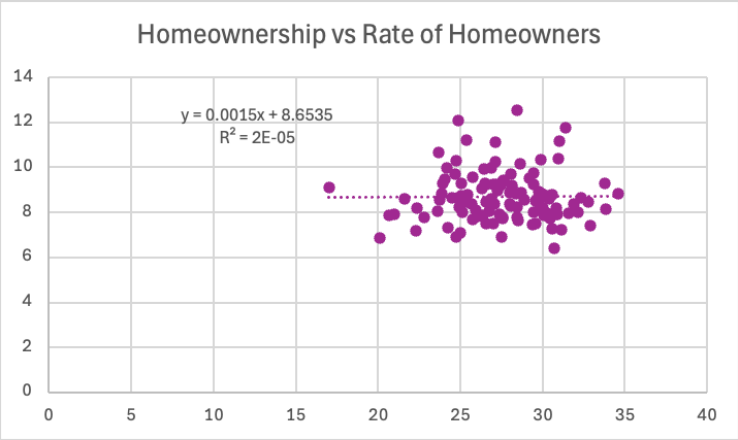
(This scatter plot shows the relationship between homeownership rates and the rate of Black homeowners across counties. The trendline, represented by y=0.1786x+8.4461, indicates a positive relationship between the two variables. As the rate of Black homeowners increases, overall homeownership rates tend to rise slightly. However, the R^2 value of 0.1951 suggests that only 19.5% of the variation in homeownership rates is explained by the rate of Black homeowners, indicating a relatively weak relationship. The data points are concentrated near the lower range of the independent variable, with a few outliers at higher rates of Black homeownership that deviate from the overall trend. These outliers warrant further investigation.)

b_3) *Figure 3:* Scatter Plot (Homeownership vs Rate of Bachelor's Degree or higher)



(This scatter plot examines the relationship between homeownership rates and the rate of individuals with a bachelor's degree or higher across counties. The trendline, given by the equation y=0.0829x+6.7049, indicates a positive relationship, where higher educational attainment is associated with increased homeownership rates. The R^2 value of 0.4671 suggests that approximately 46.7% of the variation in homeownership rates is explained by the rate of bachelor's degree holders, making this a relatively strong predictor. The data points align closely with the trendline, showing a consistent linear relationship with minimal outliers. This strong relationship highlights the significant influence of education on homeownership.)

b_4) *Figure 4:* Scatter Plot (Homeownership vs Rate of Homeowners)



(This scatter plot examines the relationship between homeownership levels and the rate of homeowners across counties. The trendline equation of y=-0.0015x + 8.6535 indicates an extremely weak and negative relationship between the variables, where an increase in the rate of homeowners corresponds to a slight decrease in overall homeownership levels. The R^2 value of less 0.001 also confirms that a relationship between the variables is essentially non-existent, with less than even 0.1% of the variation being explained by the rate of homeowners. The different data points are also pretty scattered, indicating no real trend or strong correlation. There also does not appear to be any significant outliers.)

c) *Table 4*: Correlation Table (model 3)

Correlations

		Dependent Variable	% Black Homeowners * 100	% Bachelors or Higher * 100	Rate of Homeowners *100	Age (35-54) * 100
Pearson Correlation	Dependent Variable	1.000	.442	.683	.004	.542
	% Black Homeowners * 100	.442	1.000	.429	-.296	.294
	% Bachelors or Higher * 100	.683	.429	1.000	-.121	.527
	Rate of Homeowners *100	.004	-.296	-.121	1.000	-.186
	Age (35-54) * 100	.542	.294	.527	-.186	1.000
Sig. (1-tailed)	Dependent Variable	.	<.001	<.001	.482	<.001
	% Black Homeowners * 100	.000	.	.000	.001	.001
	% Bachelors or Higher * 100	.000	.000	.	.100	.000
	Rate of Homeowners *100	.482	.001	.100	.	.024
	Age (35-54) * 100	.000	.001	.000	.024	.
N	Dependent Variable	114	114	114	114	114
	% Black Homeowners * 100	114	114	114	114	114
	% Bachelors or Higher * 100	114	114	114	114	114
	Rate of Homeowners *100	114	114	114	114	114
	Age (35-54) * 100	114	114	114	114	114

Table 5: Correlation Table (model 2)

Correlations

		Dependent Variable (LN)	Age (35-54) * 100	% Black Homeowners * 100	% Bachelors or Higher * 100	Rate of Homeowners *100
Pearson Correlation	Dependent Variable (LN)	1.000	.552	.451	.698	-.032
	Age (35-54) * 100	.552	1.000	.295	.537	-.204
	% Black Homeowners * 100	.451	.295	1.000	.477	-.338
	% Bachelors or Higher * 100	.698	.537	.477	1.000	-.162
	Rate of Homeowners *100	-.032	-.204	-.338	-.162	1.000
Sig. (1-tailed)	Dependent Variable (LN)	.	<.001	<.001	<.001	.366
	Age (35-54) * 100	.000	.	.001	.000	.014
	% Black Homeowners * 100	.000	.001	.	.000	.000
	% Bachelors or Higher * 100	.000	.000	.000	.	.042
	Rate of Homeowners *100	.366	.014	.000	.042	.
N	Dependent Variable (LN)	115	115	115	115	115
	Age (35-54) * 100	115	115	115	115	115
	% Black Homeowners * 100	115	115	115	115	115
	% Bachelors or Higher * 100	115	115	115	115	115
	Rate of Homeowners *100	115	115	115	115	115

Table 6: Correlation Table (model 1)

	Dependent Variable	*Age (35-54)*	*% Black Homeowners*	*% Bachelors or Higher*	*Total POP*
Dependent Variable	1				
Age (35-54)	0.351801944	1			
% Black Homeowners	0.557946246	0.29545491	1		
% Bachelors or Higher	0.602261006	0.53707766	0.477150707	1	
Total POP	0.995273159	0.35888398	0.603279979	0.616394076	1

d)

Table 7: MLR Model 1

SUMMARY OUTPUT

Regression Statistics		
Multiple R	0.99671497	
R Square	0.99344073	extremely high
Adjusted R Sq	0.99320221	
Standard Error	2824.13045	0.19542803
Observations	115	

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	1.3288E+11	3.3219E+10	4165.04033	4.704E-119
Residual	110	877328407	7975712.79		
Total	114	1.3375E+11			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	222.302755	2163.07897	0.10277145	0.91833149	-4064.412	4509.01756	-4064.412	4509.01756
Age (35-54)	2417.8908	7897.54378	0.30615732	0.76006357	-13233.187	18068.9689	-13233.187	18068.9689
% Black Home	-60385.662	9026.06293	-6.6901441	9.6885E-10	-78273.2	-42498.123	-78273.2	-42498.123
% Bachelors or	-3062.6208	4016.79653	-0.7624536	0.44742123	-11022.969	4897.72711	-11022.969	4897.72711
Total POP	0.28515242	0.00300869	94.7763764	2.702E-107	0.2791899	0.29111493	0.2791899	0.29111493

Table 8: MLR Model 2

SUMMARY OUTPUT									
Regression Statistics									
Multiple R	0.755459151								
R Square	0.570718528								
Adjusted R Square	0.555108293								
Standard Error	0.74564206	SE/average y=	0.085551215						
Observations	115								
ANOVA									
	df	*SS*	*MS*	*F*	*Significance F*				
Regression	4	81.30800564	20.32700141	36.56053328	2.04749E-19				
Residual	110	61.15802902	0.555982082						
Total	114	142.4660347							
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*	
Intercept	3.292290362	0.926176799	3.554710465	0.00055904	1.456825335	5.127755388	1.456825335	5.127755388	
Age (35-54) * 100	0.074434408	0.021028069	3.539764262	0.000588212	0.03276171	0.116107106	0.03276171	0.116107106	
% Black Homeowners * 100	0.058417823	0.022337876	2.615191482	0.010168736	0.014149396	0.102686249	0.014149396	0.102686249	
% Bachelors or Higher * 100	0.058610518	0.009659116	6.067896887	1.87591E-08	0.039468418	0.077752619	0.039468418	0.077752619	
Rate of Homeowners *100	0.059413341	0.023819918	2.494271487	0.014108625	0.012207856	0.106618827	0.012207856	0.106618827	

Table 9: MLR Model 3

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.74908324							
R Square	0.5611257							
Adjusted R Sq	0.54502022							
Standard Erro	0.74237558							
Observations	114							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	4	76.8057723	19.2014431	34.8406715	1.0162E-18			
Residual	109	60.0722433	0.5511215					
Total	113	136.878016						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	3.28362393	0.92214011	3.56087313	0.00054925	1.45597222	5.11127565	1.45597222	5.11127565
Age (35-54) *	0.07255504	0.02097872	3.45850607	0.00077596	0.03097589	0.11413418	0.03097589	0.11413418
% Black Home	0.08577314	0.02957104	2.9005794	0.00450638	0.02716431	0.14438197	0.02716431	0.14438197
% Bachelors o	0.05765802	0.00964071	5.98067921	2.8682E-08	0.03855044	0.0767656	0.03855044	0.0767656
Rate of Home	0.0614934	0.02376182	2.58790747	0.0109712	0.01439824	0.10858857	0.01439824	0.10858857

d) Other

Table 10, 11, 12: Coefficients + Collinearity Diagnostics + Residual Statistics

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	3.284	.922		3.561	<.001	1.456	5.111		
	% Black Homeowners * 100	.086	.030	.212	2.901	.005	.027	.144	.753	1.328
	% Bachelors or Higher * 100	.058	.010	.476	5.981	<.001	.039	.077	.637	1.570
	Rate of Homeowners *100	.061	.024	.173	2.588	.011	.014	.109	.898	1.114
	Age (35-54) * 100	.073	.021	.261	3.459	<.001	.031	.114	.705	1.418

a. Dependent Variable: Dependent Variable

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions				
				(Constant)	% Black Homeowners * 100	% Bachelors or Higher * 100	Rate of Homeowners *100	Age (35-54) * 100
1	1	4.184	1.000	.00	.01	.00	.00	.00
	2	.725	2.402	.00	.73	.00	.00	.00
	3	.074	7.514	.01	.20	.80	.02	.00
	4	.013	18.116	.00	.04	.15	.34	.54
	5	.004	33.769	.99	.02	.05	.64	.46

a. Dependent Variable: Dependent Variable

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	7.10856533	11.9894457	8.69509338	.824437445	114
Std. Predicted Value	-1.924	3.996	.000	1.000	114
Standard Error of Predicted Value	.078	.408	.144	.060	114
Adjusted Predicted Value	7.01150846	11.7454519	8.69305949	.815598427	114
Residual	-1.95132148	1.74406159	.000000000	.729117804	114
Std. Residual	-2.628	2.349	.000	.982	114
Stud. Residual	-2.657	2.406	.001	1.007	114
Deleted Residual	-1.99457097	1.82888901	.002033883	.767233256	114
Stud. Deleted Residual	-2.735	2.461	.001	1.014	114
Mahal. Distance	.241	33.141	3.965	5.128	114
Cook's Distance	.000	.175	.011	.023	114
Centered Leverage Value	.002	.293	.035	.045	114

a. Dependent Variable: Dependent Variable

Works Cited

Bayer, Patrick, et al. “What Drives Racial and Ethnic Differences in High Cost Mortgages? The Role of High Risk Lenders.” *National Bureau of Economic Research*, 1 Feb. 2016, www.nber.org/papers/w22004.

Goodman, Laurie, et al. “The Impact of Tight Credit Standards on 2009–13 Lending.” *Urban Institute*, 4 June 2016, www.urban.org/research/publication/impact-tight-credit-standards-2009-13-lending.

States, United. “Explore Census Data.” *Census.gov*, 2024, data.census.gov/table/PEPPOP2019.PEPANNRES?g=040XX00US29. Accessed 7 Dec. 2024

.

“Explore Census Data.” *Census.gov*, 2024, data.census.gov/table/ACSST5Y2022.S2502?t=Owner/Renter%20(Householder)%20Characteristics&g=040XX00US09. Accessed 7 Dec. 2024.