

# What Drives Rent?

## A Deep Dive Into People, Place, and Price

### Description of the Data Source

Data sources from three publicly available sources were used in this analysis to investigate how demographic and socioeconomic factors relate to rental prices.

1. **American Community Survey (ACS)** The U.S. Census Bureau conducts the ACS annually to provide detailed demographic, social, economic, and housing data. Approximately 3.5 million addresses are randomly sampled each year, with participation required by law, ensuring a high response rate and robust data quality. The unit of observation in this dataset is the ZIP Code. We used the 2023 ACS S0601 dataset, which includes variables such as median income, educational attainment, racial composition, and housing characteristics, which are core indicators for examining disparities across regions.
2. **Fair Market Rent (FMR)** The U.S. Department of Housing and Urban Development (HUD), through its Office of Policy Development and Research, publishes annual FMR for use in federal housing programs such as Section 8. These rent estimates are calculated using a combination of the Census Bureau's five-year ACS data (as a base), HUD-defined geographic areas, and adjustments based on the Bureau of Labor Statistics' Consumer Price Index (CPI). The unit of observation here is also at the ZIP-level, aligning with the ACS data for comparison. FMR serves as a standardized estimate of local rental costs, allowing us to assess rent relative to regional socioeconomic conditions.
3. **HUD-USPS ZIP-TRACT Crosswalk File** To ensure geographic precision and data relevance, we used the 2023 Q2 ZIP-TRACT Crosswalk file developed by HUD in collaboration with the U.S. Postal Service. This dataset links USPS ZIP Codes with Census geographic areas and includes quarterly USPS residential vacancy indicators. We used it to filter out ZIP Codes with low residential ratios or primarily commercial/industrial land use, focusing our analysis on areas where rental data reflects meaningful residential patterns.

To ensure balanced sampling across geographic regions, the dataset was stratified into five regions (Midwest, Northeast, Southeast, Southwest, and Northwest) and 1,000 ZIP Codes were randomly selected per region, resulting in a final dataset of 5,000 rows and 37 columns. This consolidated dataframe served as the single source of truth for all subsequent modeling and visualization steps.

### Linear Regression Models

To examine how demographic and socioeconomic factors relate to rent prices across the United States, five separate linear regressions were conducted—one for each region: Northeast (NE), Northwest (NW), Southeast (SE), Southwest (SW), and Midwest (MW). Each model used the SAFMR 1BR variable (Small Area Fair Market Rent for one-bedroom units) as the dependent variable, with demographic indicators such as race, marital status, educational attainment, income, and language use as predictors.

The tables display the full regression results with robust standard errors (HC1) to address potential heteroskedasticity. In the following sections, most notable results from each region are summarized, focusing on statistically significant coefficients ( $p < 0.05$ ), the direction and magnitude of associations, and any unique patterns that emerge across geographies.

## Regression Results for the Northeastern Region

In the Northeast, higher percentages of Black residents and Hispanic/Latino populations are associated with significantly lower rent prices, with coefficients of -80.33 and -67.40 respectively (both  $p < 0.01$ ). Conversely, a higher percentage of Asian residents is associated with higher rents (+67.64,  $p < 0.05$ ). Median income also shows a strong positive effect on rent (+0.13 per \$1,000,  $p < 0.01$ ). Educational attainment, marital status, and other demographic variables showed no significant associations in this region.

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac@gmail.com % Date and time: Fri, Aug 01, 2025 - 6:49:22 PM

Table 1: Regression Results for NE Region

	SAFMR 1BR
Median.Age	7.164*** (1.633)
Percent.Males	-5.386*** (1.961)
Percent.Females	
Percent.Race..White	-4.889 (4.041)
Percent.Race..Black	-2.228 (4.056)
Percent.Race..Native.or.Alaskan	-8.252 (3.764)
Percent.Race..Asian	4.143 (5.375)
Percent.Race..Native.Hawaiian.or.Pacific.Islander	7.844** (15.724)
Percent.Race..Hispanic.or.Latino	1.073 (3.075)
Percent.speak.another.language.at.home	14.108 (2.051)
Percent.never.married	-181.941 (165.379)
Percent.married	-189.016 (165.259)
Percent.divorced	-191.922 (165.273)
Percent.widowed	-189.405*** (165.456)
Percent.Race..Two.or.more.races	-0.307 (4.404)
Percent.less.than.high.school	22.413 (149.084)
Percent.high.school.graduate	20.622 (148.990)
Percent.some.college.or.associates	18.740 (148.864)
Percent.college.graduate	24.708 (148.889)
Percent.graduate.degree	22.854*** (149.157)
Median.income.in.past.year	0.015 (0.001)
Percent.born.in.state.of.residence	0.844 (0.900)
Percent..100.poverty.status	-5.520 (1.683)
Constant	17,257.550 (22,233.270)
Observations	1,000
R <sup>2</sup>	0.700
Adjusted R <sup>2</sup>	0.693
Residual Std. Error	299.599 (df = 977)
F Statistic	103.376*** (df = 22; 977)
Note:	*p<0.1; **p<0.05; ***p<0.01

## Regression Results for the Northwestern Region

In the Northwest, rents rise significantly with Median Age (+11.99,  $p < 0.01$ ) and Median Income (+0.07 per \$1,000,  $p < 0.01$ ). Black (%) and Hispanic/Latino (%) show significant negative associations with rent (-77.70 and -72.97 respectively,  $p < 0.01$ ). Interestingly, Percent Married has a large positive effect on rent (+74.51,  $p < 0.01$ ), suggesting a link between household structure and rental prices in the NW region.

% Table created by stargazer v.5.2.3 by Marek Hlavac, Social Policy Institute. E-mail: marek.hlavac@gmail.com % Date and time: Fri, Aug 01, 2025 - 6:49:22 PM

Table 2: Regression Results for NW Region

	SAFMR 1BR
Median.Age	8.938*** (1.909)
Percent.Males	-3.025 (3.871)
Percent.Females	
Percent.Race..White	-8.434** (2.719)
Percent.Race..Black	-5.347 (3.517)
Percent.Race..Native.or.Alaskan	-11.039*** (3.445)
Percent.Race..Asian	-2.639 (3.226)
Percent.Race..Native.Hawaiian.or.Pacific.Islander	-20.897*** (7.530)
Percent.Race..Hispanic.or.Latino	-6.072*** (2.101)
Percent.speak.another.language.at.home	9.846 (1.678)
Percent.never.married	92.085 (192.930)
Percent.married	83.915 (192.942)
Percent.divorced	84.925 (192.930)
Percent.widowed	91.270*** (192.514)
Percent.Race..Two.or.more.races	4.383 (2.669)
Percent.less.than.high.school	144.992 (175.139)
Percent.high.school.graduate	147.699 (175.384)
Percent.some.college.or.associates	153.272 (175.434)
Percent.college.graduate	161.903 (175.258)
Percent.graduate.degree	159.108*** (175.134)
Median.income.in.past.year	0.013 (0.002)
Percent.born.in.state.of.residence	2.043 (0.949)
Percent..100.poverty.status	-1.952 (2.622)
Constant	-22,993.970 (25,606.000)
Observations	1,000
R <sup>2</sup>	0.664
Adjusted R <sup>2</sup>	0.656
Residual Std. Error	358.123 (df = 977)
F Statistic	87.653*** (df = 22; 977)
Note:	*p<0.1; **p<0.05; ***p<0.01