Group #3

Effect of Demographic Variables on Median House Prices Across the USA

Our team

Jose Traboulsi
The Process Guy

Melanie Liao
The Insights Lady

Bansri Patel
The API Guru

ChatGPT (?)
Emotional Support

Agenda

1. Introduction
2. Research Questions
3. Process
4. Analysis of Variables
5. Conclusion

1 Introduction

Objective: Our project aims to analyze the demographic factors driving changes in median house prices across various counties in the United States.

Key Variables
We focus on four demographic metrics:

- **Education Rates** (% of population with a bachelor's degree).
- **Median Household Income**
- **Employment Rates**
- **Poverty Rates** (% of population living in poverty)

Research Questions

2 Key Research Questions

1) How have median house prices and our variables changed over time across the USA?

2) How do our selected demographic variables correlate with median house prices across counties in the USA?

3) Are median house prices higher in counties with a more educated population?

4) What is the distribution of housing prices in high-poverty vs. low-poverty areas?

5) How do median house prices and poverty rates compare across urban and rural areas in the USA?

Process

3 Data Analysis Process

Data Collection:
Sourced county-level data from the U.S. Census Bureau's American Community Survey (ACS) using the Census API (ACS1). The data is consistently in 2013 inflation-adjusted values.

Data Cleaning:
Handled missing values and ensured data consistency. Filtered the data to include only relevant records.

Exploratory Data Analysis (EDA):
Conducted initial analysis to understand data distribution and detect outliers. Created summary statistics and visualizations to identify trends.

Statistical Analysis:
Applied statistical models to analyze the relationship between education rates, median household income, employment rate, poverty rates, and median house prices. Assessed the impact of these variables on price fluctuations.

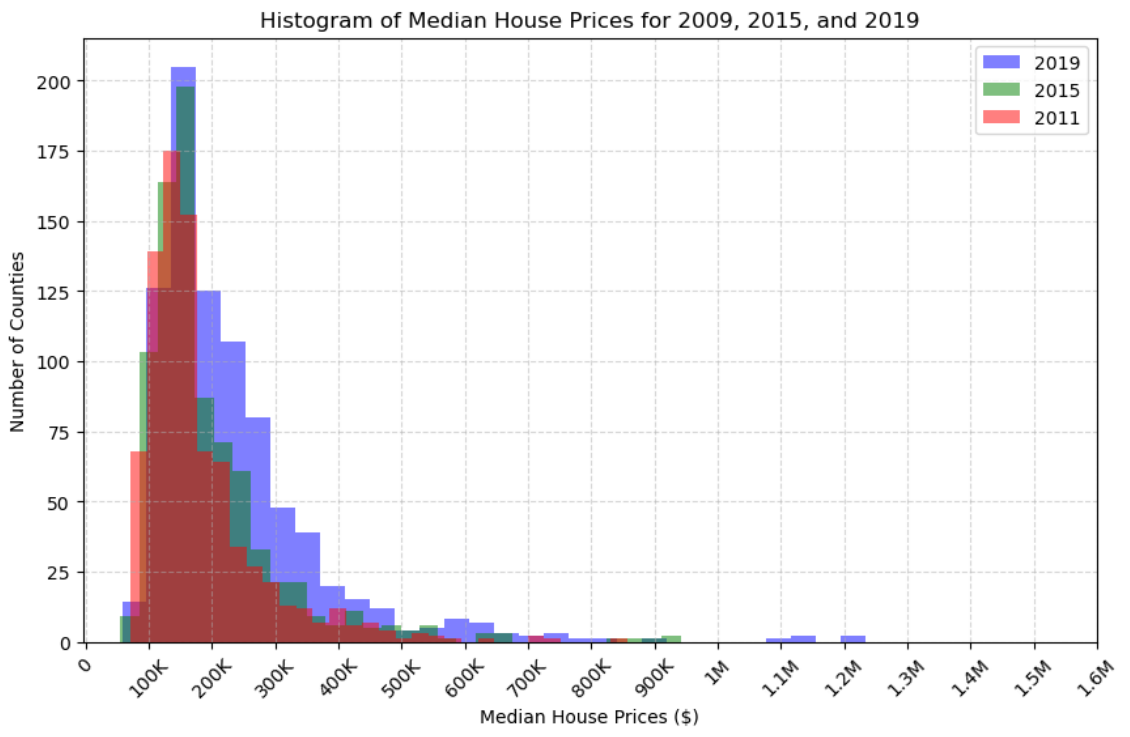
Visualization:
Utilized tools such as Matplotlib, SciPy, and Pandas to create histograms, heatmaps, and scatter plots to visualize data trends and relationships.

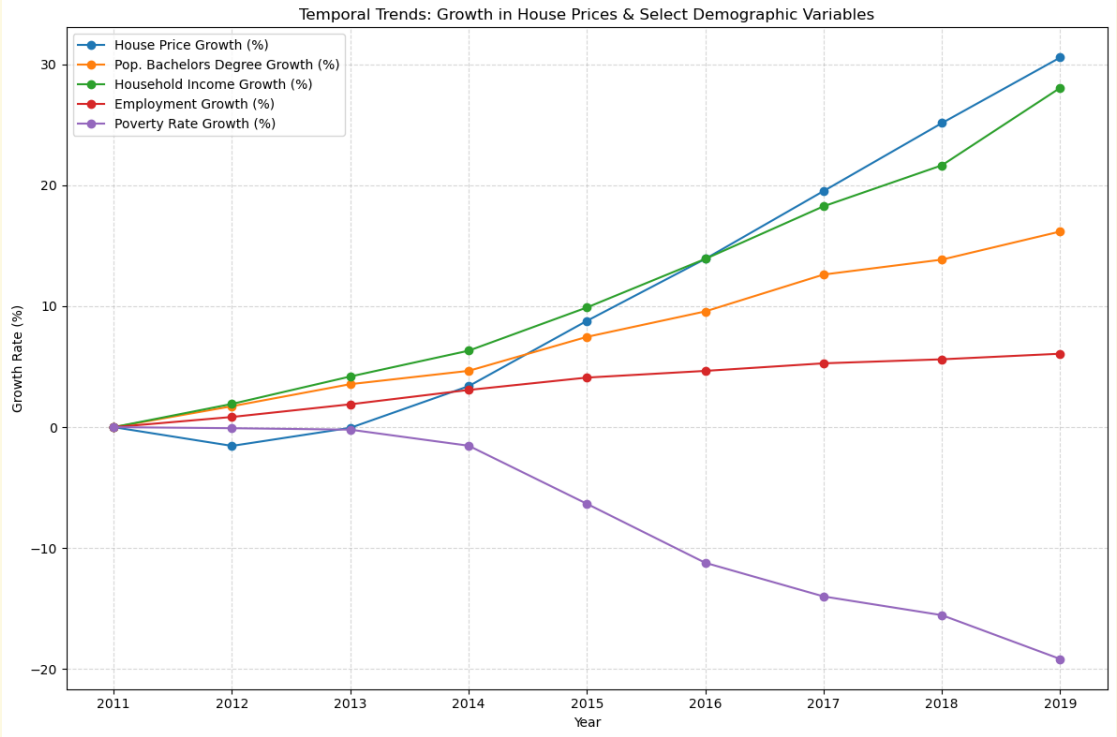
Interpretation:
Interpreted the results to understand how education, income, employment, and poverty influence median house prices. Identified key trends and regional disparities.

Analysis

Investigating Our Key Research Questions

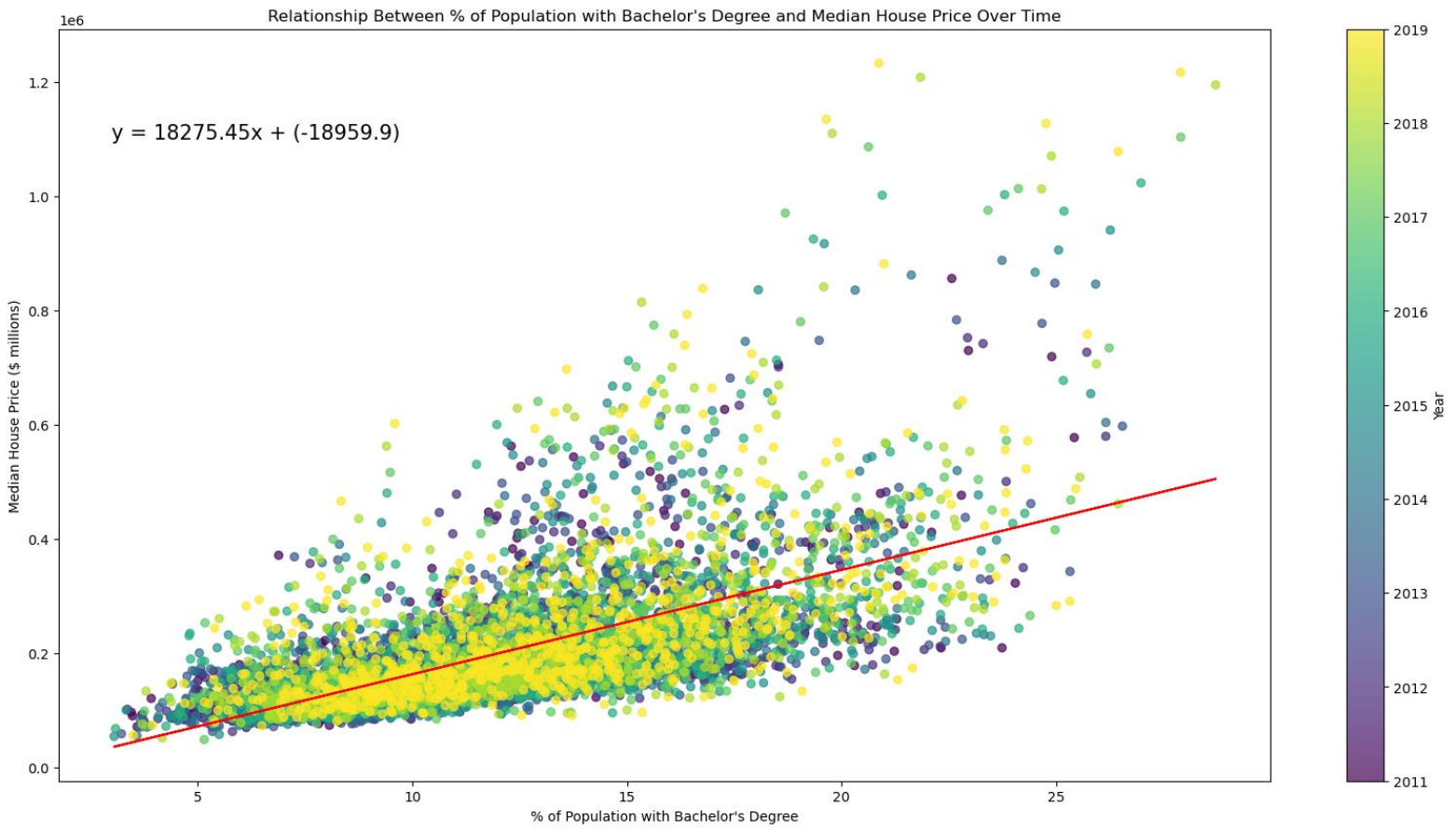
1 How have median house prices and our variables changed over time?





2 How do our selected demographic variables correlate with median house prices across counties in the USA?

Education Rates



Linear Regression Statistics:

Slope: 18275.45

Intercept: -18959.9

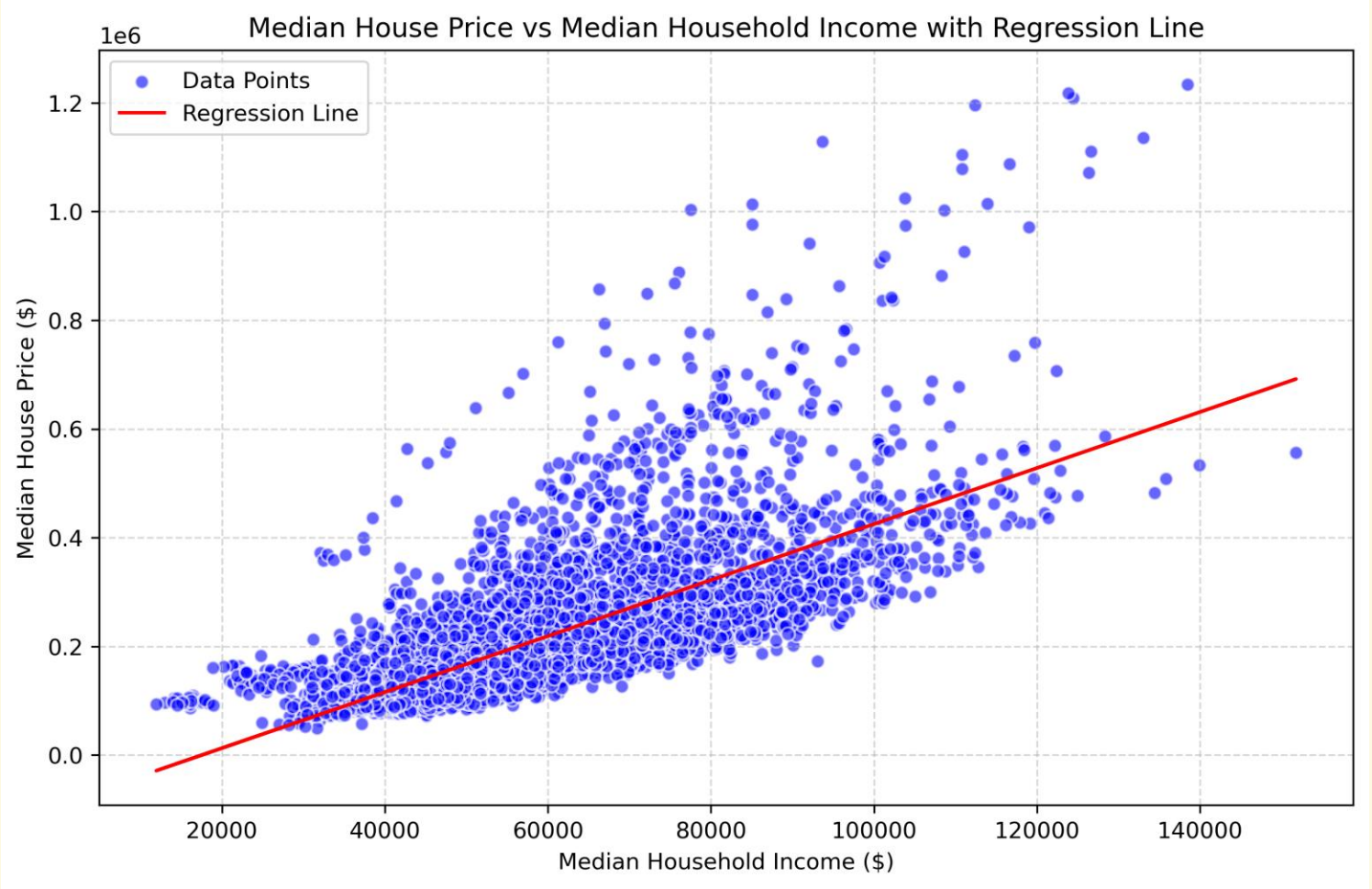
R: 0.63

R^2: 00.3998

P-value: 0.00e+00

2 How do our selected demographic variables correlate with median house prices across counties in the USA?

Median Household Income



Linear Regression Statistics:
Slope: 5.15
Intercept: -90380.97
R: 0.71
R^2: 0.51
P-value: 0.00e+00
Standard Error: 0.06

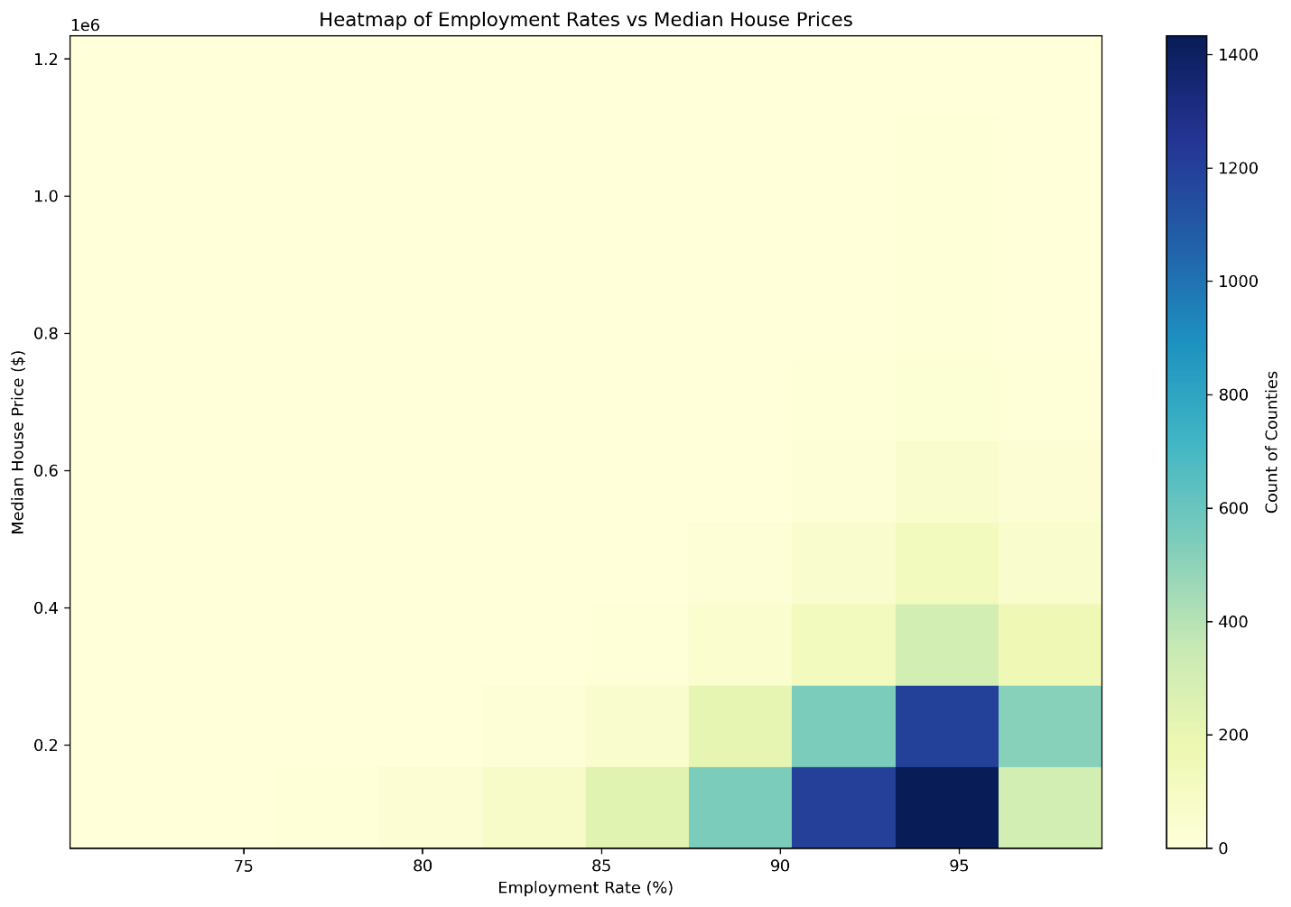
2 How do our selected demographic variables correlate with median house prices across counties in the USA?

Poverty Rates

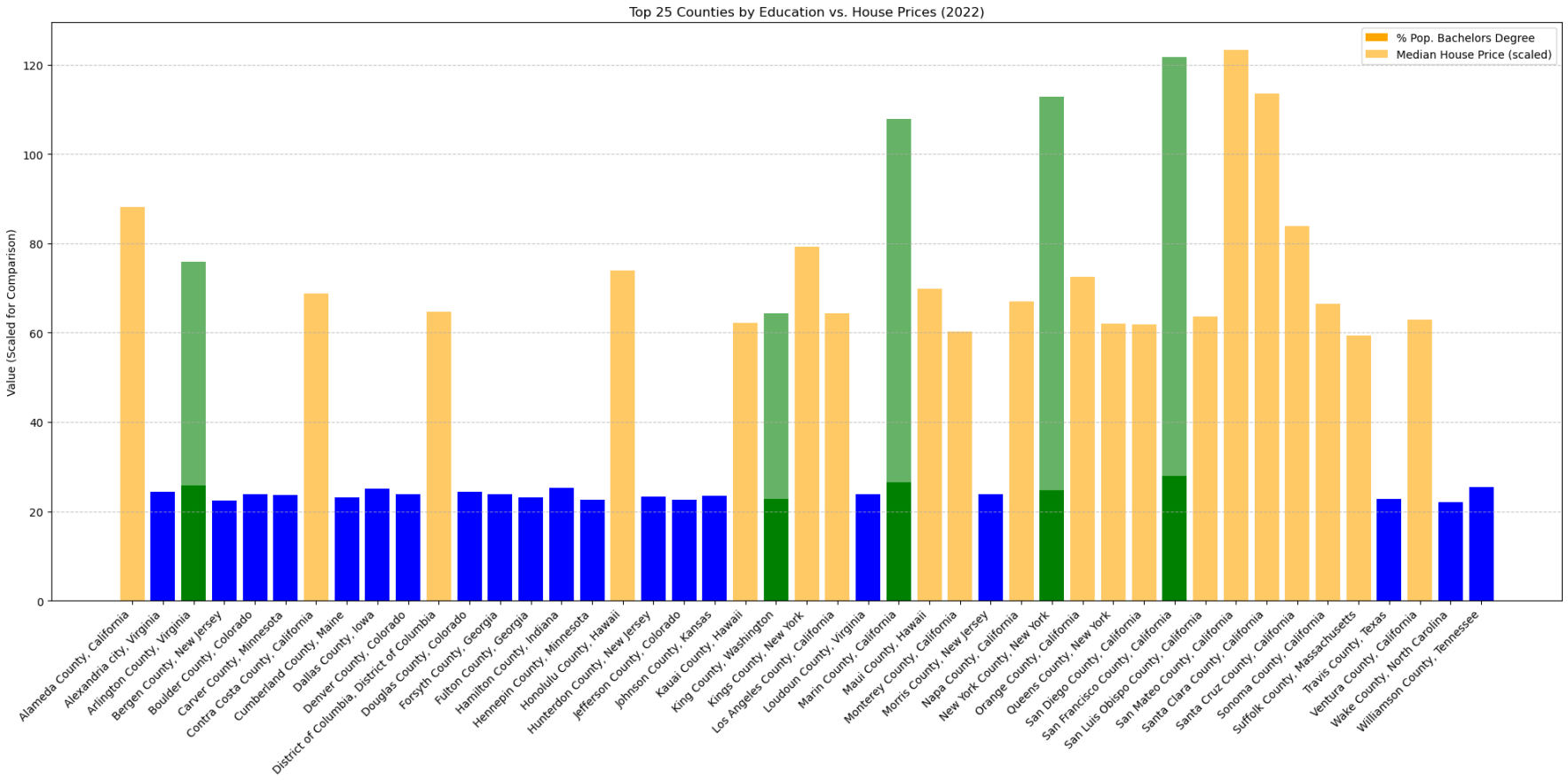
Log-Linear Regression Statistics:
R (Correlation Coefficient): -0.44
R^2 (R-squared): 0.19
P-value: 0.0000

2 How do our selected demographic variables correlate with median house prices across counties in the USA?

Employment Rates



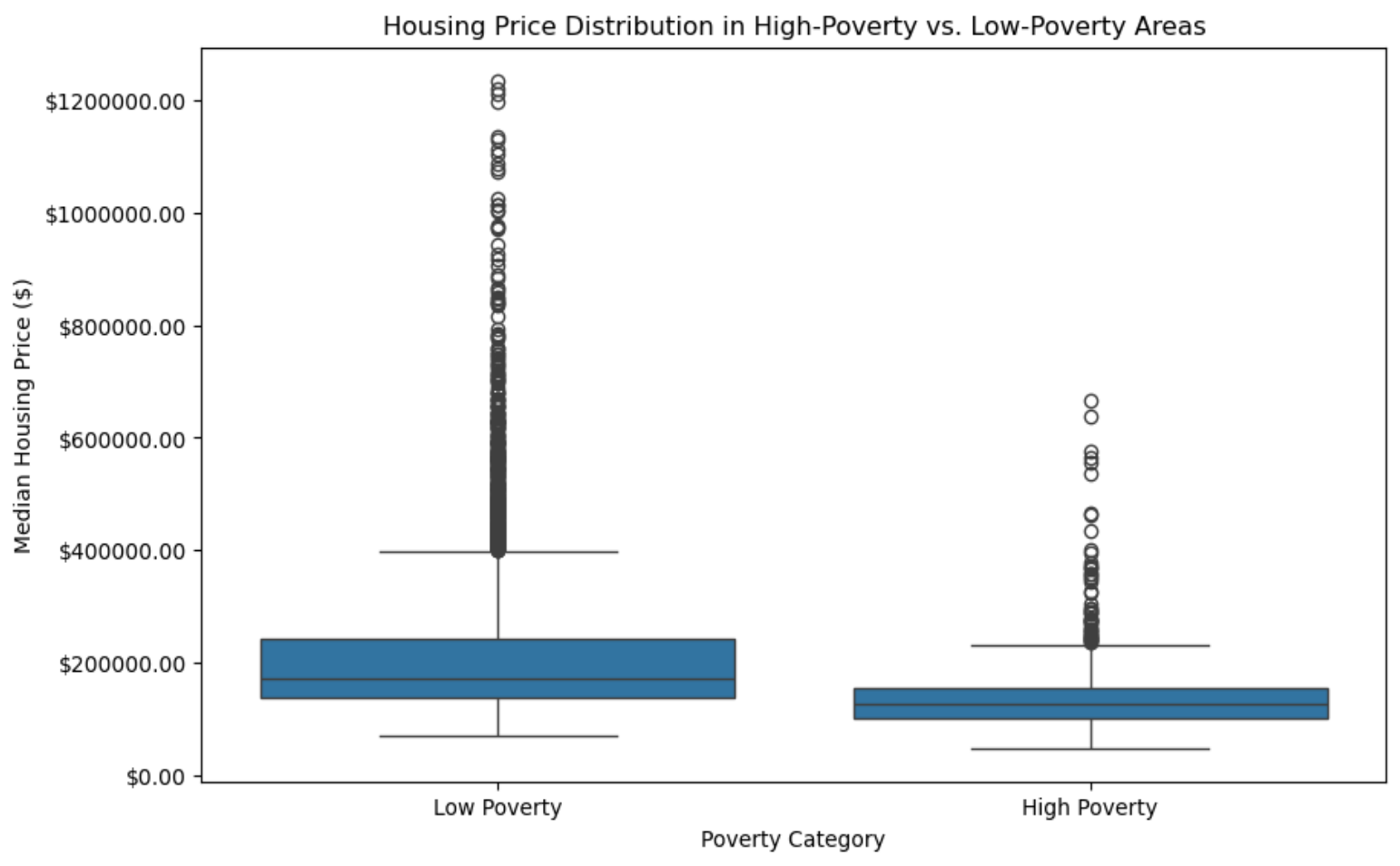
3 Are median house prices higher in counties with a more educated population?



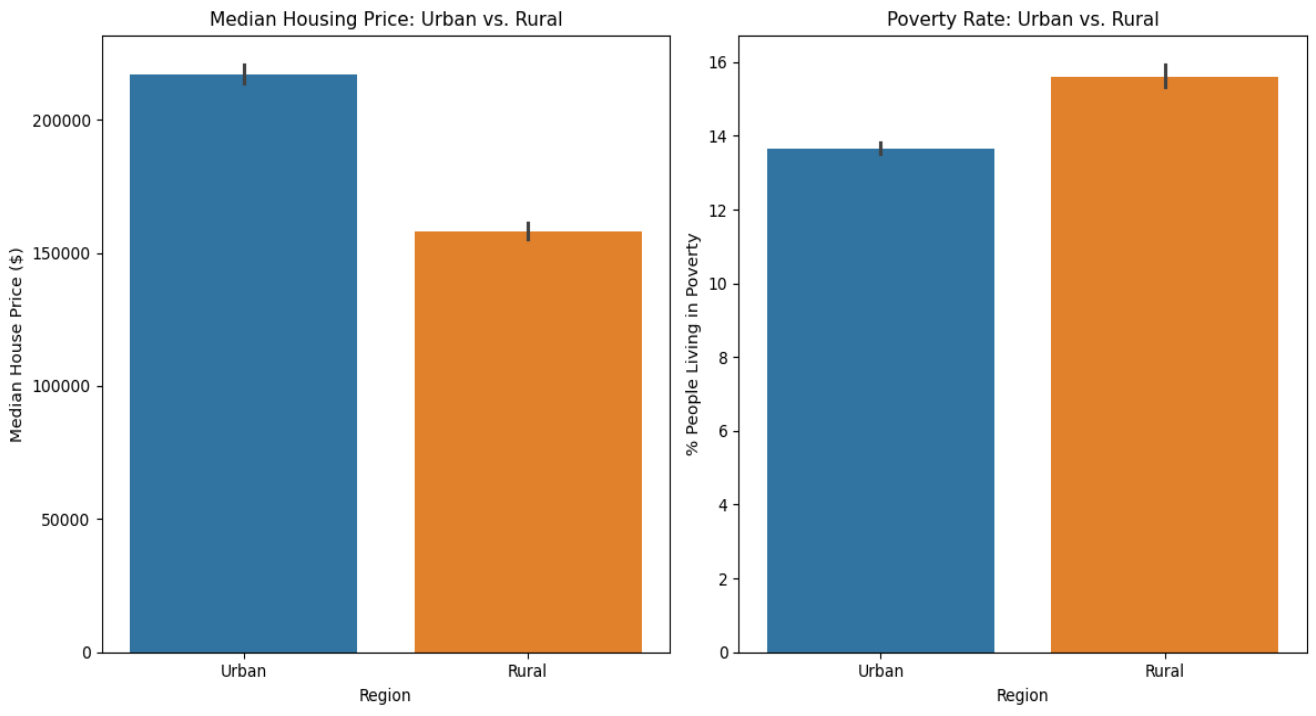
RESULTS OF A T TEST CONDUCTED ON THE DIFFERENCE BETWEEN MEAN HOUSE PRICES FOR TOP AND BOTTOM 25% COUNTIES

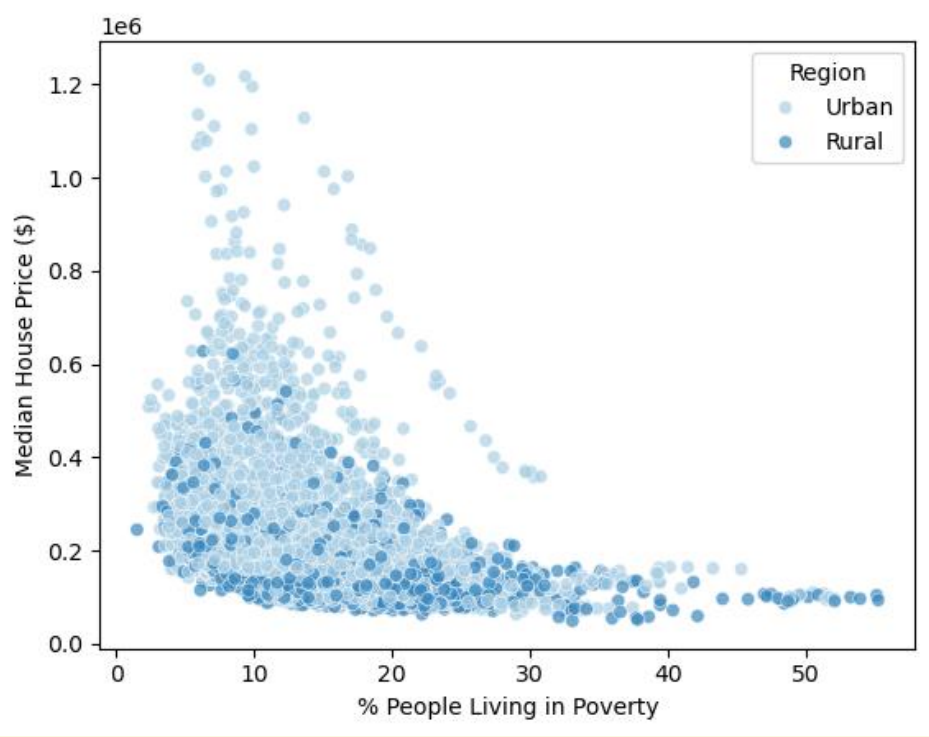
- T-statistic: 15.7639
- P-value: 2.85146e-44

4 What is the distribution of housing prices in high-poverty vs. low-poverty areas?



5 How do median house prices and poverty rates compare across urban and rural areas in the USA?





Conclusion

Key Findings

5 Key Findings

- Median house prices and median household, share a positive relationship
- Employment growth, while steady, does not directly correlate
- Educational attainment and house prices are related, but with mid-strength, this is not the only variable at play.
- Higher poverty rates are associated with lower median house prices, indicating an inverse relationship between these variables.

Thank you

Happy Holidays!

See you later!

No-code vacay!