Maarij Rizvi

CSCI-1070-02

12/13/24

<div align="center">Analyzing Income Inequality in Urban vs. Rural Areas Across the U.S.</div>

Introduction

This report investigates the disparity in income inequality between urban and rural counties within the United States. The research question driving this analysis is: Is income inequality greater in urban areas compared to rural areas in the U.S.? This question is explored by examining the Gini index, a measure of income distribution, alongside median income and population density, across a selection of counties from different regions of the U.S. and urban-rural classifications. Data from the U.S. Census Bureau is used to calculate the Gini index and generate visualizations to show comparisons between the counties and classifications.

This study utilizes data from select counties in New York and Pennsylvania (representing the East Coast), Illinois and Michigan (representing the Midwest), and California and Washington (representing the West Coast). Within each state, both urban and rural counties were chosen through their high and lower population densities respectively. This approach allows for comparing income inequality through the urban-rural classification.

The Gini index serves as the dependent variable (Y) in this analysis, indicating income inequality. The independent variables (X) include the area type (urban/rural) and population density. By comparing Gini indices across these counties, the study aims to determine whether the hypothesized relationship between urban counties and higher inequality is true. The examination of median income and population density serves as further evidence and context for understanding the factors that contribute to income inequality differences.

Background

Income inequality is a complex societal issue with significant implications, influenced by various factors such as education levels, access to resources and opportunities, local job markets, historical economic trends, governmental policies, etc. Understanding its factors is essential to developing effective strategies to combat the negative effects that income inequality brings.

Urban areas, typically characterized by higher population densities and large infrastructure, often experience greater income disparities. This can be attributed to factors such as the concentration of high-paying jobs in certain areas and neighborhoods along with limited access to these opportunities by lower-income residents. This creates a wider gap between the wealthy and the poor compared to rural areas. Rural areas, while potentially exhibiting lower overall income levels, may demonstrate different patterns of income inequality. Factors such as limited job opportunities and geographic isolation can contribute to an income standstill and inequality.

The Gini index, a statistical measure of income distribution, quantifies inequality on a scale of zero to one. A Gini index of zero represents perfect equality, where all individuals have the same income, while a value of 1 represents absolute inequality, where a single individual holds all of the wealth. This index provides a standardized method for comparing income inequality for different populations and areas, and in this case, urban and rural counties.

Data and Methods

Data for this study were sourced from the U.S. Census Bureau website. The data included median household income, population density, and income distribution data which were broken down into earning brackets. These brackets were crucial in calculating the Gini index as

population density and median income are not sufficient enough to show income inequality, as numbers of people in certain income brackets are used to show how wealth is spread out through a region.

Data collection involved gathering data on median income, population density, and income distribution across different income brackets from the US Census Bureau website for the selected counties. The Gini index calculation relies on the Lorenz curve, which graphically represents the distribution of income within a population. The Lorenz curve plots the cumulative share of income against the cumulative share of the population, ranked from lowest to highest income. A perfectly equal income distribution would result in a straight diagonal line (the line of equality). The Gini index is calculated as the ratio of the area between the line of equality and the Lorenz curve to the total area under the line of equality. In this study, the Gini index was calculated for each county using a Python function that implements the trapezoidal rule for numerical integration. This method approximates the area under the Lorenz curve by dividing it into trapezoids. Visualizations were generated using the matplotlib library in Python which includes a bar chart for comparing median income between urban and rural counties, a scatter plot to explore the relationship between population density and median income, and a line plot displaying the Gini index for each county.

The choice of visualization methods was aimed at presenting the data in a clear and informative manner to show comparisons and emphasize trends. The bar chart allowed for a direct comparison of median incomes, while the scatter plot allowed for an examination of the relationship between population density and median income. The line plot of Gini indices provides a straightforward visualization of income inequality differences between urban and rural counties. These methods combined with the calculated Gini indices provided a

comprehensive approach to examining the income inequality between these urban and rural counties.
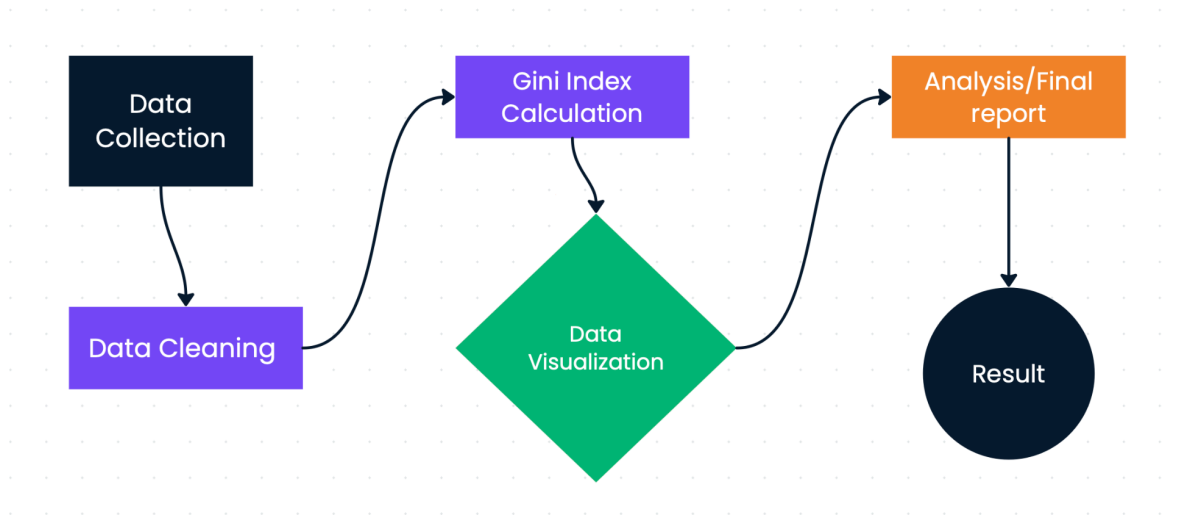
Results

The analysis, based on the calculated Gini indices and visualized in the included plots, indicates a general trend of higher income inequality in urban counties compared to rural counties. The line graph directly comparing Gini indices clearly shows this difference, with urban counties exhibiting higher values. The bar chart comparing median incomes reveals the expected trend of higher median incomes in urban counties; however, this isn't uniform, and the scatter plot further emphasizes this point, highlighting the outlier of Alpine County, California. Alpine County presents an anomaly, being a rural county with a lower Gini index than anticipated. This suggests the presence of unique local factors influencing income distribution that may need further investigation. It can be assumed that Alpine County is made up of particularly wealthy individuals who prefer living away from urban areas, which may be the cause of the high median income.

Conclusion

      The results, using the Gini index as a measure of income inequality, generally support the hypothesis that urban areas experience greater income inequality than rural areas. While the trends aren't completely uniform, the differences in Gini indices between urban and rural counties suggest a correlation between urbanization and income disparity. The Gini index is a reasonable measure for this study as it provides a single, standardized metric for comparison across different geographic areas. The unexpected result for Alpine County highlights the complexity of income inequality and suggests that localized factors can play a significant role. Expanding the sample size to include more counties and potentially incorporating more detailed socio-economic data would strengthen the analysis and provide further insight into the nuances of income distribution across the U.S.