Nejma Ahmed

Date ; 16/05/2023

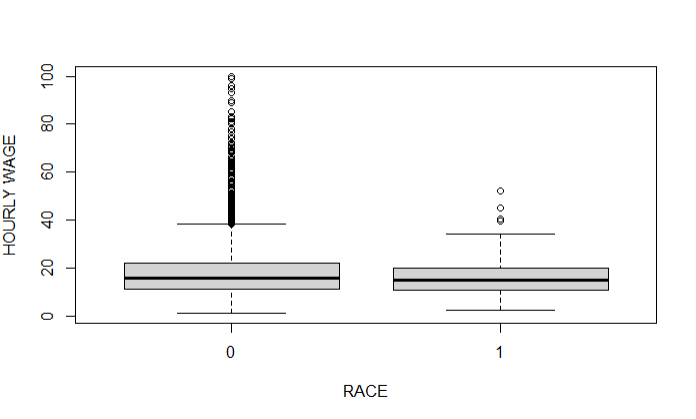
Report on Data Fest 2

Exploring Income and Wealth Disparities among American Indian and Alaskan (AIAN) Communities

American Indian and Alaska Native (AIAN) communities in the United States face major economic issues as a result of huge income and wealth discrepancies with the general community. The major goal of this study was to investigate the association between hourly wages and race in the context of AIAN individuals. The investigation is based on data from the Current Population Survey (CPS) IPUMS database, which provides useful variables for analyzing the economic position of AIAN individuals and households.

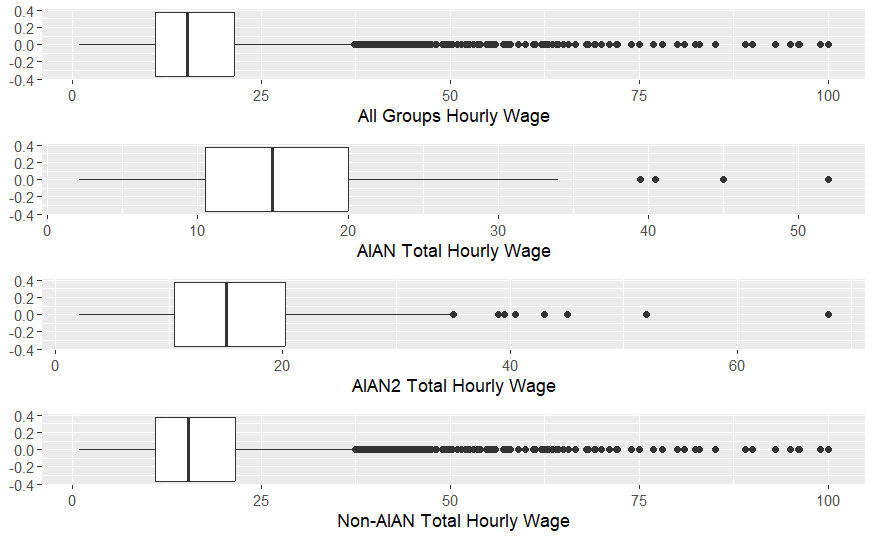
We wanted to look into economic differences between AIAN and non-AIAN people, in particular when race, age, and family income were taken into consideration. We were interested in getting insight into the magnitude of the income differences encountered by AIAN groups, as well as the potential underlying factors leading to these disparities, by studying these variables. Understanding these processes is critical for creating targeted interventions and policy actions to address and reduce economic inequality among AIAN populations and promote equitable economic opportunities for all.

To begin the analysis, we created a boxplot to visualize the association between hourly wage and race. The x-axis in this box plot depicts the racial category, with 0 being non-AIAN and 1 representing AIAN. The y-axis represents the hourly wage. We can gain insight into any potential discrepancies or inequities by analyzing the distribution of hourly wages across the two race categories.

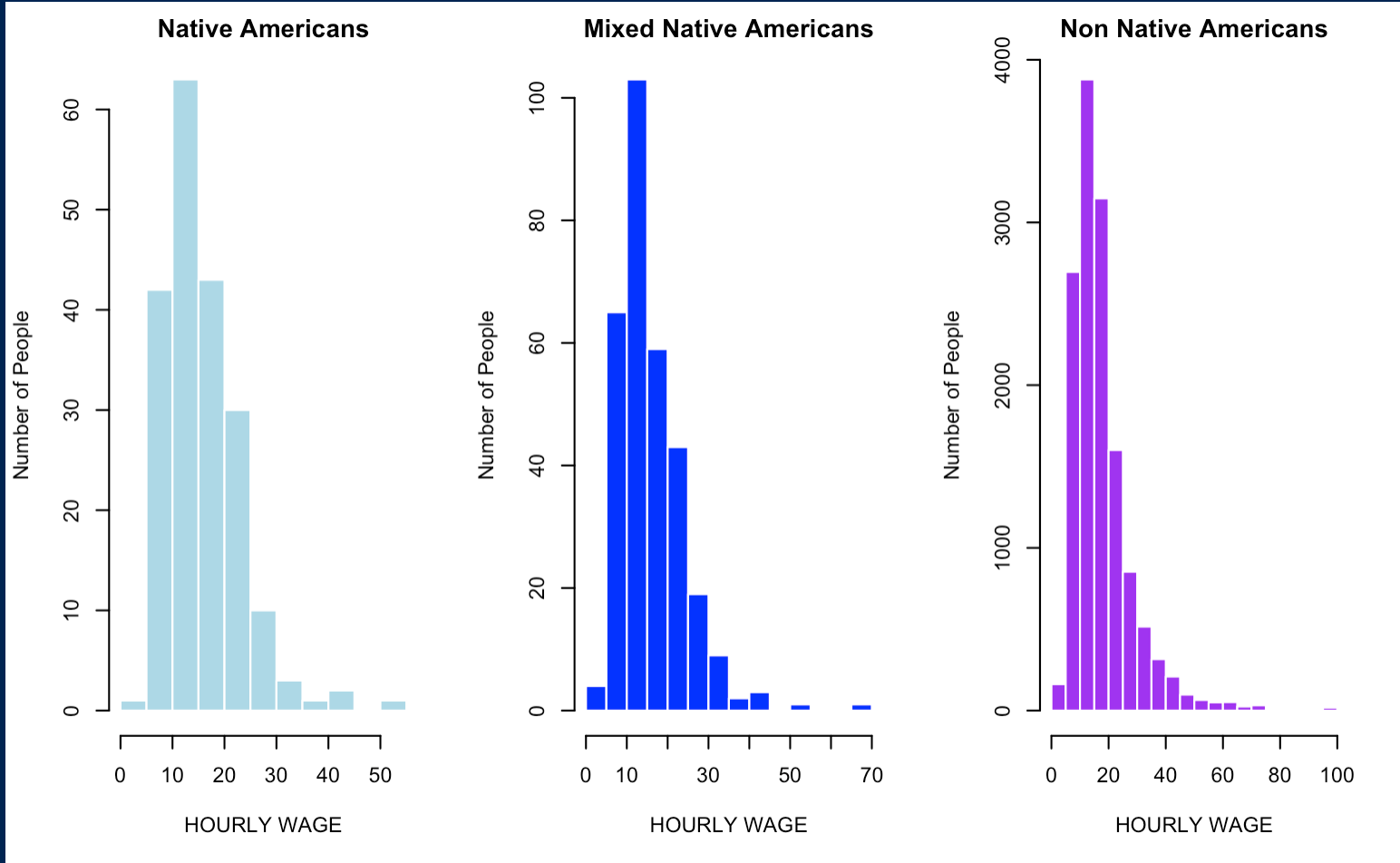


The boxplot above shows that the non-AIAN category earns slightly more per hour than the AIAN category. This is shown by the median line, which shows that the median hourly wage for non-AIANs looks to be slightly greater than for AIANs. In addition, the upper quartile (75th percentile) of hourly pay for non-AIANs is slightly greater than for AIANs. However, it is important to note that there is overlap between the two groups, indicating that individuals from both categories earn a broad range of hourly salaries. Further study and statistical testing would be required to assess the significance of this distinction and to understand the underlying reasons that contribute to the observed gap in hourly pay between the two groups.

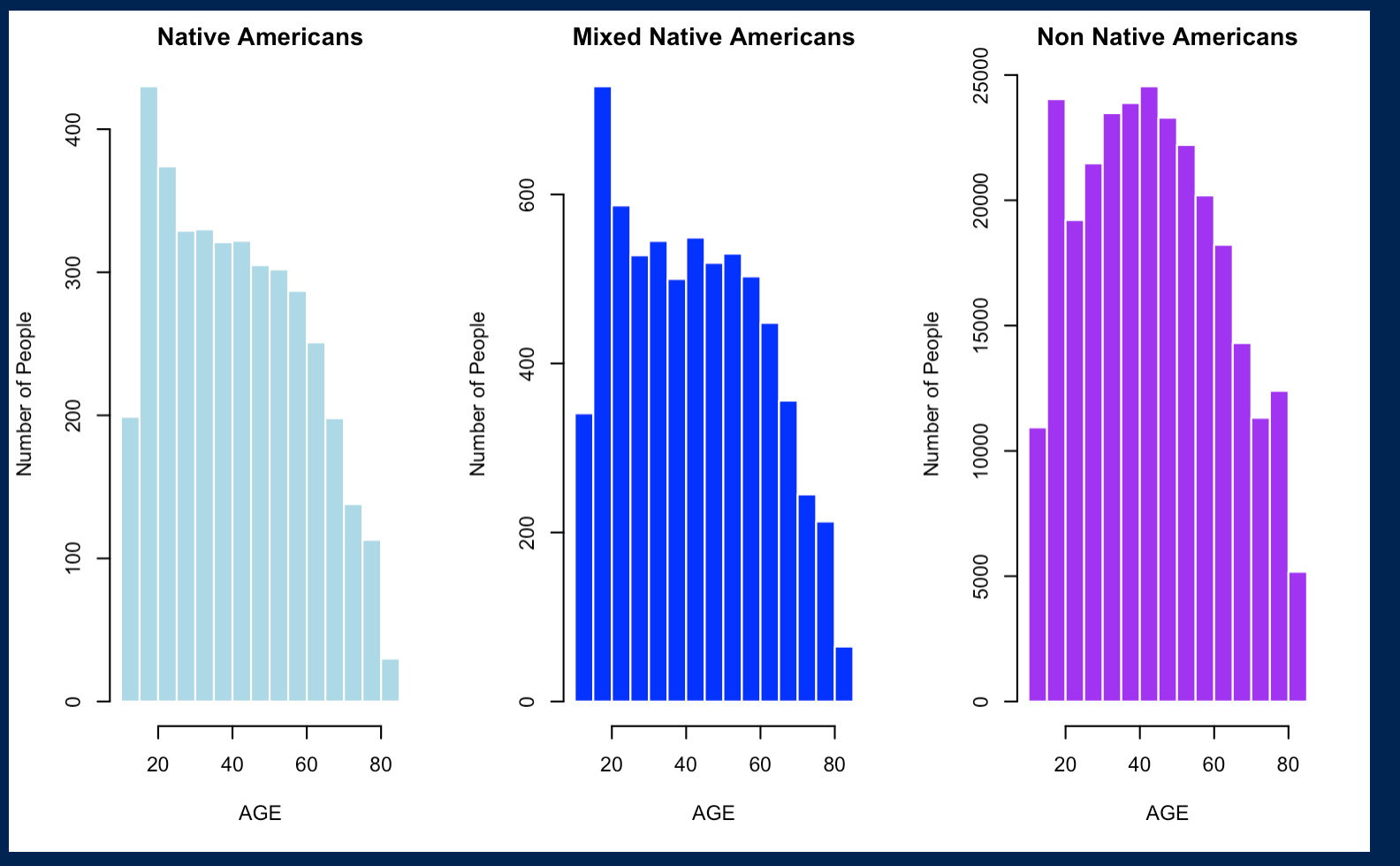
To continue this investigation, we also ran another box lot representing age and race, and the results showed no difference between non-AIAN and AIAN.



To further our investigation, we conducted an analysis of the hourly wages among different groups, particularly focusing on AIAN and AIAN2, which represent individuals with mixed AIAN heritage. When examining the comprehensive chart, we observed that all groups had an hourly wage reaching up to $100. However, upon closer inspection of the AIAN-only boxplot, a distinct pattern emerged, revealing that their hourly wage peaked at $50. Similarly, AIAN2 demonstrated a slightly higher wage compared to AIAN, highlighting a significant difference between all the groups. Notably, the non-AIAN group's hourly wage was positioned at the very bottom of the chart, exhibiting a substantial disparity between AIAN and AIAN2. In fact, it appeared to be double the amount, reaching up to $100, whereas AIAN and AIAN2 only reached approximately $50.

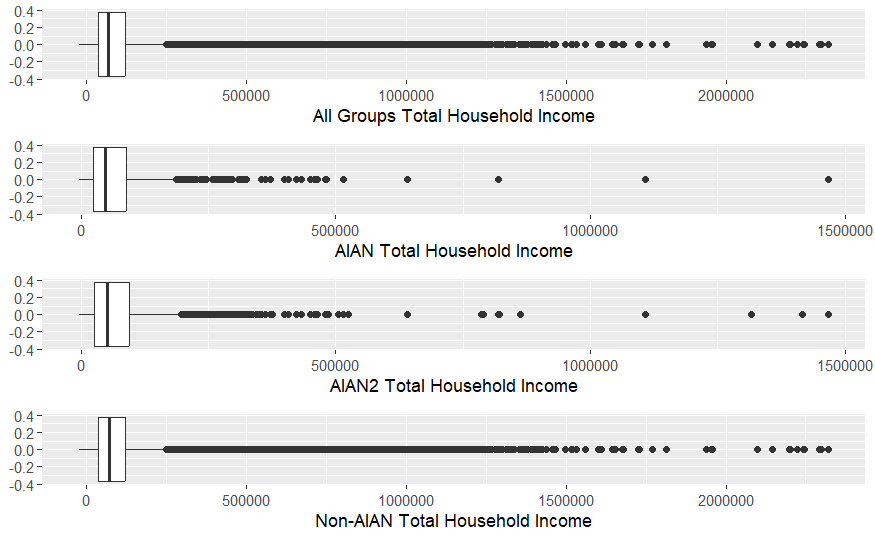
We also created a histogram that depicts the same pattern for better visualization.

Histogram depicting native American mixed native American and non-native American age of employment.



In addition, we conducted an investigation to determine if age contributed to the observed differences illustrated in the previous chart. Analyzing the histogram provided above, it is evident that there are no significant variations in age among the groups. However, an interesting trend emerges when examining Native Americans and individuals of mixed Native American and non-Native American heritage. All three groups display a distinct right-skewed pattern, indicating a higher concentration of individuals in older age ranges.

Boxplot of income comparison of native american households , mixed native american households , non- native american households

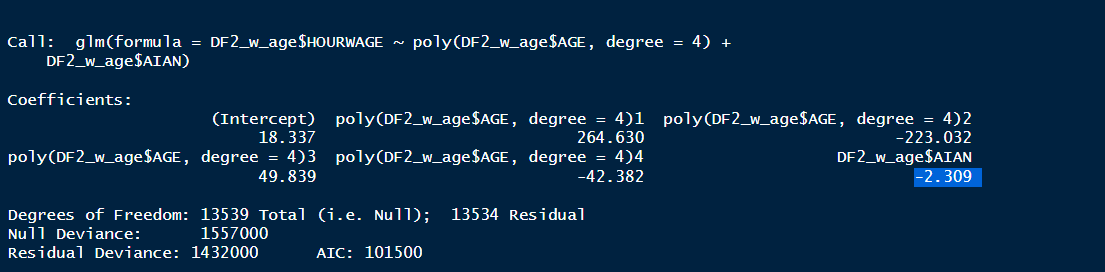


Furthermore, the box plot above provides a comprehensive analysis of household incomes among Native Americans, non-Native Americans, and individuals with mixed Native American backgrounds. We observe a consistent trend akin to the hourly wage findings, further emphasizing the disparities. Notably, the household incomes of non-Native Americans stand out with a significant difference compared to those of Native Americans and individuals with mixed Native American backgrounds.

Upon closer examination, the box plot reveals that the household incomes of Native Americans (Aian) and those with mixed Native American backgrounds (Aian2) range from approximately $50,000 to almost $100,000. This range represents the middle 50% of the income distribution within these groups. The median household income, represented by the central line in the box, falls somewhere within this range.

In contrast, the non-Native American group's household incomes exhibit a much wider range, spanning from approximately $50,000 to $200,000. This expanded range indicates greater income variation within this demographic. The median household income for non-Native Americans is positioned at a higher value than that of Native Americans and individuals with mixed Native American backgrounds, reflecting a higher midpoint of the income distribution.

These findings highlight a notable disparity in household incomes, where non-Native Americans enjoy a significantly higher income range compared to both Native Americans and individuals with mixed Native American backgrounds. This discrepancy in income distribution underscores the importance of further examining the factors contributing to these disparities and addressing potential underlying socioeconomic factors.



In analyzing the household income box plot, we employed an orthogonal polynomial to account for the influence of age on income disparities. This approach is crucial, as younger individuals generally tend to earn less than their older counterparts. By controlling for age, we aimed to mitigate the confounding effect and obtain more accurate results.

Another important consideration in regression analyses is multicollinearity, which refers to a high degree of linear intercorrelation among explanatory variables in a multiple logistic regression model. Multicollinearity can lead to erroneous outcomes and increase the risk of overfitting, where the model becomes too complex and loses generalizability.

Taking age and multicollinearity into account, we found that being an AIAN (American Indian/Alaska Native) individual was associated with a reduction in hourly wage of approximately $2.31 (p-value = 0.00218, 95% confidence interval [0.8632416, 3.7567584]). This statistically significant finding suggests that, even after adjusting for age, AIAN individuals tend to earn less compared to other demographic groups.

These findings shed light on the complex relationship between ethnicity, age, and income disparities. Further exploration is warranted to understand the underlying factors contributing to these differences and to develop strategies that address and rectify the income gaps experienced by AIAN individuals.

Conclusion:

In conclusion, the results presented above indicate the presence of disparities in household income among different ethnic groups. However, it is important to note that this investigation would benefit from further comprehensive analysis. Due to the tight time frame, we were unable to conduct extensive research to delve deeper into the underlying factors contributing to these disparities.

To enhance the analysis in future studies, it would be valuable to consider additional variables that may influence household income, such as education level, occupation, and geographic location. By including these factors, a more nuanced understanding of income disparities could be achieved.

Furthermore, expanding the sample size and incorporating longitudinal data would provide a more robust and accurate depiction of income trends over time. This would allow for a more comprehensive examination of the dynamics and potential changes in income disparities among different ethnic groups.

Given the limitations of the current investigation, it is crucial to recognize that these results provide a preliminary snapshot of the situation. A more thorough and in-depth analysis is necessary to fully comprehend the complex dynamics underlying household income disparities.

Personal Reflection

After working with my group during the Data Fest, I am happy of several achievements that have significantly contributed to my personal and professional progress. Witnessing the effectiveness of teamwork and collaboration in moving project tasks forward was a great experience. Although our performance did not meet our initial expectations, the experience itself was essential in terms of learning and personal development.

In addition, One of the key challenges we faced was the high amount of missing data and the related issue of successfully imputing it. Despite our best efforts, we were unable to resolve this issue, which had an impact on the overall analysis and interpretation of our results. We also encountered a significant challenge in appropriately selecting and refining our models. Despite putting in a lot of time and effort, we ran into problems choosing the best models for our data and struggling with how to interpret their conclusions. These problems, however, provided important learning opportunities, emphasizing the significance of strong data pretreatment approaches as well as the need for a more in-depth understanding of model selection and interpretation.

In conclusion, our data analysis provided clarification on the income and wealth discrepancies that Native Americans experience. We have provided insights into the challenges and inequalities that this community faces by researching numerous socioeconomic variables.

However, in order to acquire a more full picture of the problem, future investigations should include other variables. One important factor to examine is occupation, which has a substantial impact on income and wealth. Policymakers can learn about occupational segregation, salary disparities, and challenges to career progression by investigating the types of occupations held by Native Americans.