Poverty Data Analyses in California and Minnesota

Brian Morales, Kevin Dang

December 7, 2021

0.0.1 Introduction/Background

Poverty is one of the greatest concerns that the United States is currently faced with. This issue is often defined as a household or community lacking the necessary finances, resources, or even shelter. However, there are many underlying factors that add further complexity to the topic at hand. With this project, we aimed to distinguish the ethnic and social groups that are below the poverty line in each of our home states over time, California and Minnesota. The motivation of this paper and data stems from our own communities, backgrounds, and experiences as children. Our asperities and empirical accounts in underprivileged communities inspired us to analyze this data from an objective, wide-scale perspective.

With the definition of poverty often varying from person to person, we will be following the U. S Census Bureau's measurement of poverty which establishes a threshold using household income through earnings, child support, educational assistance, and others. In the table below, the thresholds are the same throughout the United States but the measurement of needs varies depending on the family's size and the age of their members. Additionally, to help understand the history of poverty measures, further information can be found on the Census Bureau's website.

Family Size	Avg Threshold 2016	Avg 2017	Avg 2018	Avg 2019
One person	12228	12228	12784	13011
Under age 65	12486	12752	13064	13300
Aged 65 and older	11511	11756	12043	12261
Two people	15569	15877	16247	16521
Household under 65	16151	16493	16889	17196
Household 65 and older	14522	14828	15193	15468
Three people	19105	19515	19985	20335
Four people	24563	25094	25701	26172
Five people	29111	29714	30459	31021
Six people	32928	33618	34533	35129
Seven People	37458	38172	39194	40016
Eight people	41781	42648	43602	44461
Nine people and more	49721	50681	51393	52875

Our data were obtained from the U.S Census Bureau, the principal agency of the U.S Federal Statistical System which is responsible for producing data about the American people, and OpentIntro, a non-profit open-source organization that the American Institute of Mathematics has approved. OpenIntro consists of government sources from the Census Bureau to the Economic Research Service. We believe that the data that we used is observational since it has been collected to measure or survey income levels, the demographics, employment rates, etc. by county for each state. Furthermore, the observations and variables range from population size from 2000 to 2019 to county

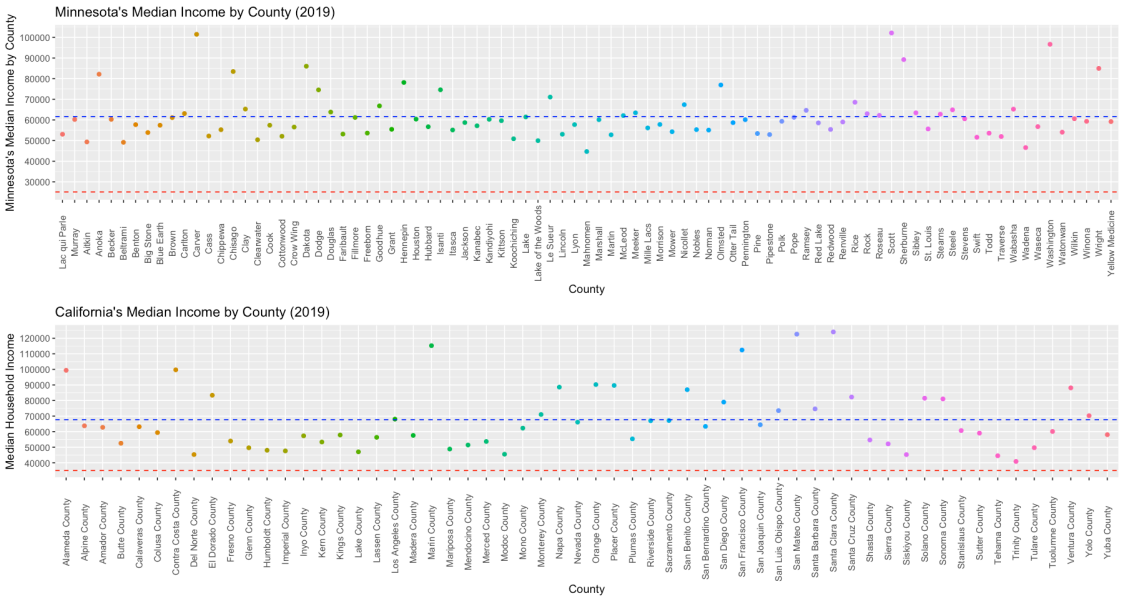
income levels and the percentage of ethnic groups in each county.

0.0.2 Method and Results

Following the collection process, most of the project was spent cleaning the data and deciding which variables to use before the analysis. This is because when the data was downloaded as a .xls file the format and structure of the data were not consistent so there were problems loading the files in R. The cleaning process consisted of merging datasets, fixing structural errors, removing irrelevant data, and amending variables in order to ensure the datasets are consistent with each other. During this process, there were inconsistencies in the U.S Consensus dataset from the years 2010 to 2015. Thus, it was decided to work with data from the following time period, 2016 to 2019. Nevertheless, there were missing values in the data set for 2016 to 2019. The U.S. consensus data set had values for all county populations, on the other hand, the yearly OpenIntro data set had the corresponding variables we wanted to use. To polish our datasets we worked with both the U.S Consensus and OpenIntro's data set to calculate correct percentage values for citizens under a certain age, in a certain ethnic group, and with a particular education level.

The analysis most appropriate to resolve what ethnic and social groups are under poverty was exploratory and time period analysis. Our goal in the analysis was to focus on the different types of population characteristics that are below the poverty line in California and Minnesota. Before discussing our results, we want to keep in mind when looking at the graphs, the population in California is vastly larger than in Minnesota. In the first section, we were interested in looking at income inequality in the most recent year in our data. This led to our curiosity about the percentage of people and children living below poverty throughout both states in 2019. In the next few sections, we will give an explanation on how we obtained our data, what calculations we applied to correctly get the percentage values for each county in California and Minnesota, and compare one another, and examine the difference. The following piece of information displays the different age groups living in poverty from 2016 to 2019. It is important to note this is different from the previous people under 18 living in poverty. The different age groups living in poverty are all accumulated and calculated under one total population. The next bit exhibits each ethnic group living in poverty from 2016 to 2019. The last segment presents the poverty rate for each household compared to the level of education.

0.0.3 Graphs for Median Household in 2019

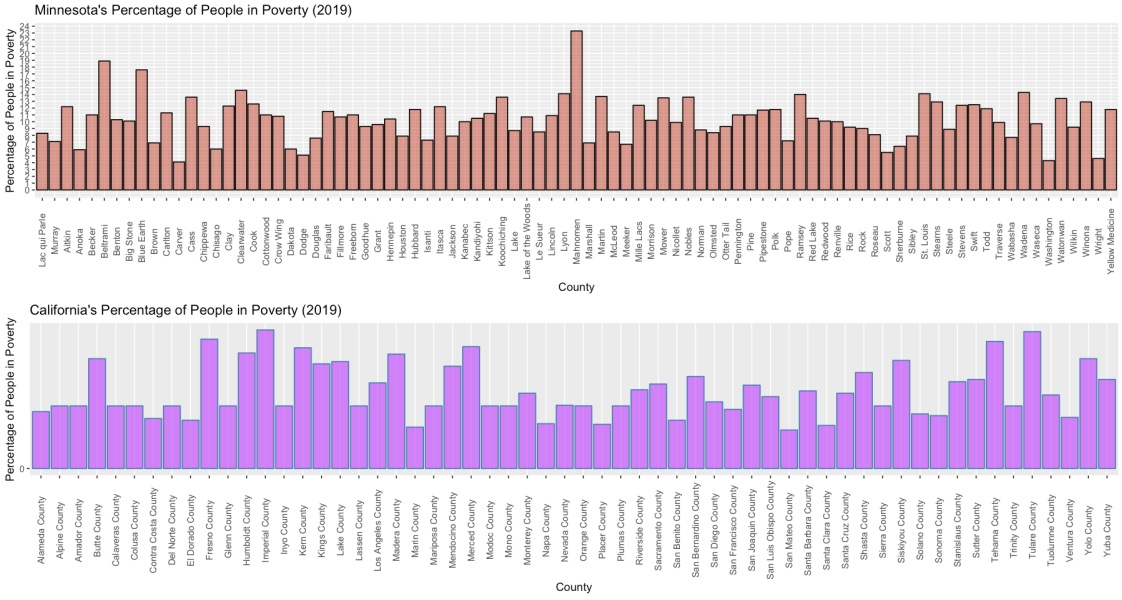


The two graphs above show the median household income for each county in 2019. To achieve our results we gathered variables from the U.S. Consensus that held information on the median household income, the county, and the number of family members in each household. We defined our household size to be four. How we calculated the household size was by looking at the person per household variable in the U.S Consensus dataset and averaging them all. California's and Minnesota's household sizes had a small difference, however, the differences were not significant enough to decide a different household size for each state. We chose median household income to avoid any outliers that could possibly skew our data. The graphs above show each county's median household income. The blue horizontal line represents the mean of the median household income in each corresponding state in California that was roughly $\$67,713$ and for Minnesota, it was \$61656. The red horizontal line represents the threshold to be considered under poverty.

Minnesota's median household income is well above the threshold with Mahnomen county being closest to the threshold. That being said, Mahnomen's population is small compared to Hennepin county which is one of the largest counties in Minnesota.

Looking at California, the median household income is relatively close to the poverty threshold. California has multiple counties near the threshold value with Trinity county being the closest. It can be argued that California's median household income for each county is closer to the threshold because of population density. Nevertheless, the cost of living in California is very expensive. From our observed data, it seems that Minnesota has more people living above poverty than California, however, we do not have enough evidence to support that claim.

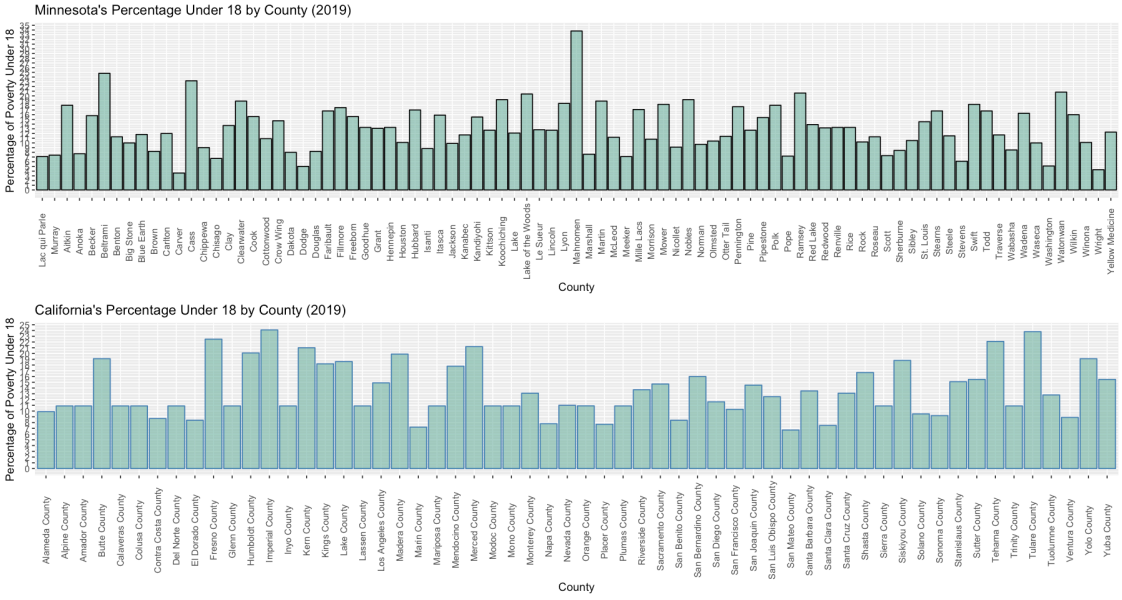
0.0.4 Percentage of People in Poverty



In the beginning, we were primarily interested in looking at the overall poverty rates between California and Minnesota, for which we obtained approximately 13.51% and 9.67% respectively. Furthermore, we wanted to take a closer look at the highest and lowest poverty rates by county.

When looking at the graphs at glance, they tell us that the largest percentage of people living in poverty is roughly 24% for Imperial, California, and Mahnomen, Minnesota. On the other hand, Carver, Minnesota, and San Mateo, California have the lowest rate of poverty. Here, we can see the large difference when comparing counties to the state-level poverty rate. As we mentioned earlier, we want to keep in mind that there is a large difference in populations between the two states and we want to note that each county is diverse regarding the size. In summary, the poverty rates in California are much higher than in Minnesota when we look at the side of the graph by side.

0.0.5 Percentage Under 18 in Poverty

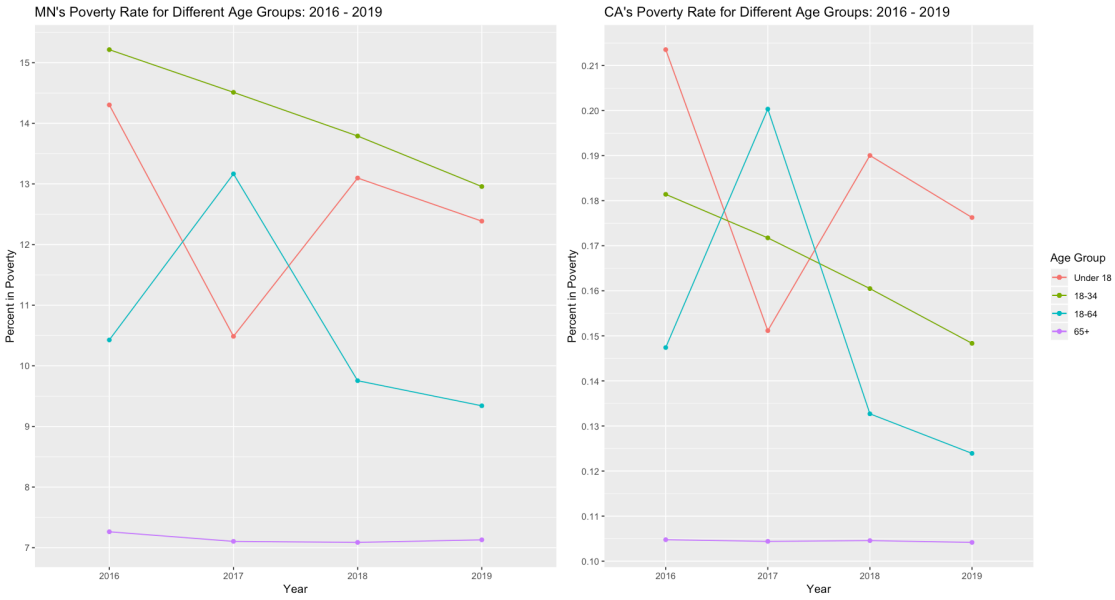


Section five represents the amount of people under 18 in each county that live in poverty. The measurement of a person under 18 living in poverty is different from an adult living in poverty. When observing a family of four one can be an adult and the rest children. Therefore, one adult's median income will place the children in poverty due to the fact that it is not a sufficient amount of money to support three kids.

To obtain the above data we utilized poverty under 19 in 2019 variable in the U.S Consensus data set. Minnesota's data set had all the required values to graph each county's percentage, however, California's had missing values. To compensate for the missing percentages we calculated the mean of all counties percentages and replaced the missing values.

Comparing both histograms we notice that Minnesota's histogram fluctuates more than California's. California does not fluctuate as much because of the amount of counties that had missing percentages. The mean percentage results in a more stable histogram. Regardless, California has a larger percentage of underage people living in poverty.

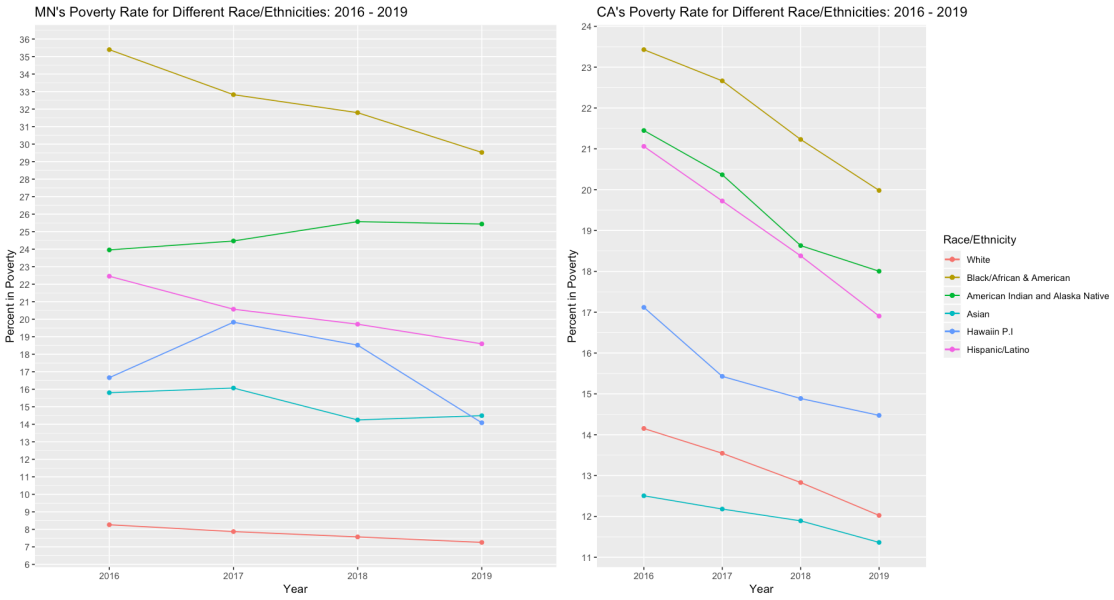
0.0.6 Different Age groups in poverty from 2016, 17, 18, 19



Section 6, we observe the level of poverty from 2016 to 2019 for each age group. In this section, we applied OpenIntro's dataset to extract ages under 5, under 18, ages between 18 to 64, and ages 65 and older. The variables had to be amended because the age group under 18 and ages between 18 to 64 were split into different variables in the data set. Additionally, we had to scale the percentage to correctly match the population size. Hence, in the above graphs, we employed both the U.S Consensus and OpenIntro to complete our analysis. The percentages were taken from OpenIntro and the correct population was used from the U.S. Consensus. Having both datasets we managed to scale the percentage according to each state's population size.

Intriguingly enough, age's 18 and under have the highest percentage of poverty for California, and age's 16 to 34 were the highest in Minnesota for 2016. Notice for the following year both California and Minnesota's underage poverty decreased significantly. Contrarily, the rate for ages 18 to 64 notably increased in CA and MN. That is interesting since the poverty rate for ages 18 to 34 consistently decreased for both states. Overall, we notice that the poverty rate is at a decreasing trend from 2016 to 2019 except for the ages 65 and older. In both plots, CA's and MN's poverty rates for ages 65 and older are relatively low, especially CA's with a low of 0.105.

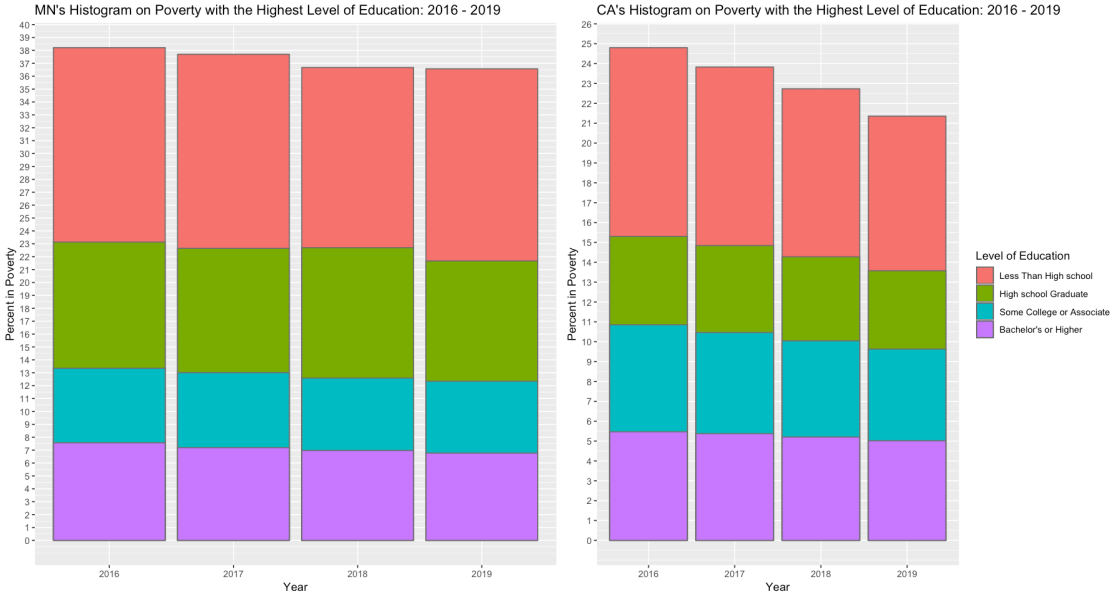
0.0.7 Different Ethnicities in poverty for 2016, 17, 18, 19



Section 7 plots the different races/ethnicities under poverty. Similar to the previous graph, to plot and analyze this graph we used both datasets to complete our analysis. We looked at White, Black, American Indian/Alaskan Native, Pacific Islander, and Latinos with their correlating poverty rate. Similarly to Section 6, we scaled their percentages according to each culture's population size for CA and MN. We made sure to sum all values for each race's poverty rate and divide the sum by the total population to its agreed state.

Dismally, Blacks had the highest poverty rate with MN's at a high of 34% in 2016. Unfortunately, it was a stat we expected since so much research and statistics is done to aid and provide poor black communities. A shocking result was Asian's are the least poor out of all groups throughout the whole measured time period in CA. We honestly conjectured that Whites would have the lowest poverty rate, especially in CA. MN's had a singular trend flip as the poverty rate for Asians increased to be above the poverty rate for Pacific Islanders in 2019. Overall, CA's poverty rate decreased for all ethnic groups and most of MN's ethnic groups decreased except American Indian/Alaska Native.

0.0.8 Poverty level compared to the highest level of education in 2016, 17, 18, 19



We often hear that education plays a significant role in an individual's income, which leads to our interest in looking at the percentage of people living in poverty with the highest level of education from 2016 to 2019. Moreover, we grouped the educational attainment by individuals that didn't complete high school, completed high school, some college or obtained an associate's degree, and have obtained a bachelor's degree or higher.

For both states, there is a noticeable trend in both states that the lower the education level the higher the poverty rate. Overall, California has much lower poverty rates based on education levels when compared to Minnesota over the years. We were surprised by this fact since California's population is much higher, so we assumed before that Minnesota would have a lower poverty rate regarding education levels.

1 Conclusion

After analyzing our selected data, examining the median household income in 2019, the percentage of poverty rate for each county in 2019, the poverty for underage individuals, the different age groups in poverty, the ethnic groups that are needful, and the percentage of poverty compared to education level, we want to say that there is a decreasing trend in poverty, however, we do not have enough evidence to support that. If we were to solely focus on our data then our concluding statement is that the poverty rate had a decreasing trend from 2016 to 2019. We learned that the U.S Consensus determined poverty from each age group differently, specifically, underage children, we noticed that from 2016 to 2019 there was a downwarding trend in ages and ethnicities living in poverty, and people with higher education have less poverty rate.

In the future, we would like to extend our research by considering more time periods such as 2010 to 2015 or more recent times as in 2020 and 2021. We would also like to extend this research during the pandemic and investigate what affects did COVID-19 has in poor communities. We think it

would be imperative to know if the poverty rate increased and if so by what proportion or if the poverty rate decreased? Other thoughts are comparing the different poverty levels from the 2010s to the 1990s; if we have improved in decreasing the poverty rate. This research not only showed us how to utilize our statistical analysis, but, confirm our unsettled knowledge.