

Population by Race and Ethnicity

Tejas Patel *Department of Informatics University of Massachusetts Amherst* Amherst, Massachusetts, USA trpatel@umass.edu

Tejas was responsible for all code, more specifically: data preprocessing, analysis, and visualization

I. INTRODUCTION

The goal of this data investigation is to obtain a better knowledge of the racial and ethnic diversity of various geographic locations in the United States, from the national level down to Economic Development Districts (EDDs). The data collection being worked with includes race and ethnicity population estimates for the United States, states, counties, metros, micros, and EDDs from 2000 to 2019.

This analysis has significance for several reasons. First of all, it gives us a visual representation of the national distribution of individuals of various races as well as the evolution of this distribution over time. This can offer insightful information about demographic trends and patterns, which can help shape a variety of programs and policies, including housing, education, and healthcare, as well as urban planning.

Second, we may pinpoint regions with high concentrations of a certain racial or ethnic group as well as areas of racial variety by looking at the racial makeup of each district. This can assist in highlighting locations that have the potential to be rich in culture as well as those that could require focused community support or intervention.

Finally, by analyzing how the population of each race has changed over time in each state or district, we may get a sense of the national demographic trends. This can provide insight into migratory patterns, changes in birth and death rates, and other variables influencing these trends.

The main questions being answered through this analysis are:

1. What is the racial composition of each district?
2. How has the population of each race changed over the years in each state or district?

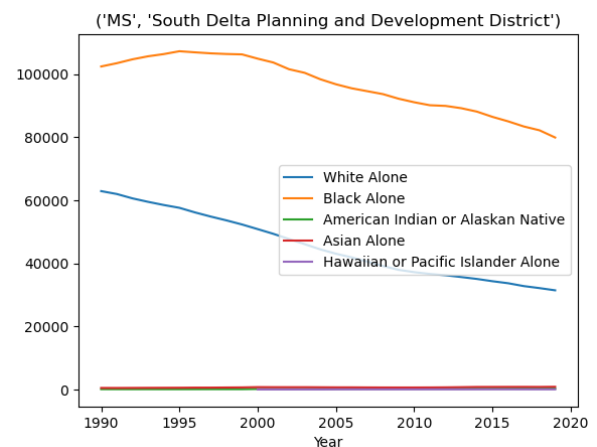
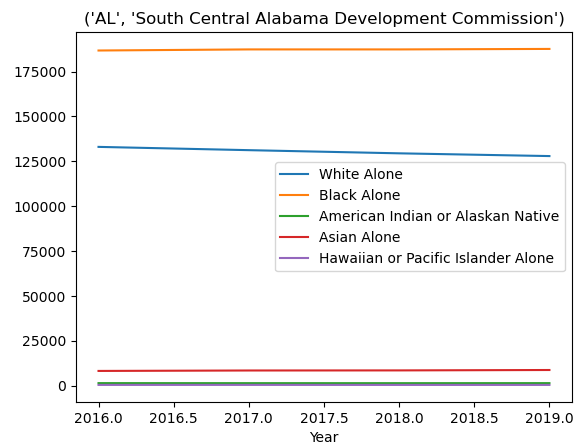
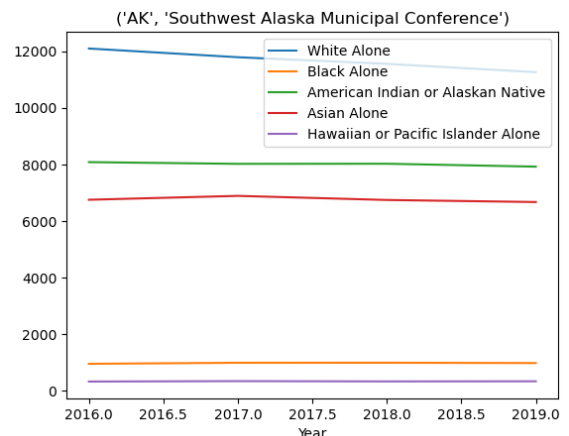
By answering these questions, the aim is to provide a comprehensive picture of racial and ethnic demographics in the U.S. and contribute to a better understanding of our diverse nation. This, in turn, can help to foster a more inclusive and equitable society.

II. DATA and PREPROCESSING

The data provided offers a comprehensive snapshot of the racial and ethnic composition across various districts in the United States over a period of time. It includes a combination of three separate CSV files into one data frame, which was then cleaned of most of its N/A values. Leaving the final cleaned data frame with columns for the total population of each district, along with the population count of various racial and ethnic groups such as White Alone, Black Alone, American Indian or Alaskan Native, Asian Alone, Hawaiian or Pacific Islander Alone, Two or More Races, Not Hispanic, and Hispanic. Additionally, the data has been processed to provide the proportion of each racial and ethnic group to the total population of each district. This temporal data allows for an analysis of trends over time. The data also reveals variations in the proportions of other racial and ethnic groups in these districts over the years. This information is crucial for understanding the diverse and evolving demographics of the United States, providing a detailed analysis of demographic trends and patterns. It can be used to gain insights into the changing racial and ethnic composition of these areas over time.

III. ANALYSIS TASK 1

The first task involved filtering out districts where the population of the 'White Alone' racial group exceeded a certain threshold. This was done to concentrate on places with a more diversified racial makeup. We may observe a visualized racial makeup over time after identifying these districts. This allows us to track trends and changes in the racial mix of these districts over time, revealing demographic movements and patterns.



III. ANALYSIS TASK 1- FINDINGS

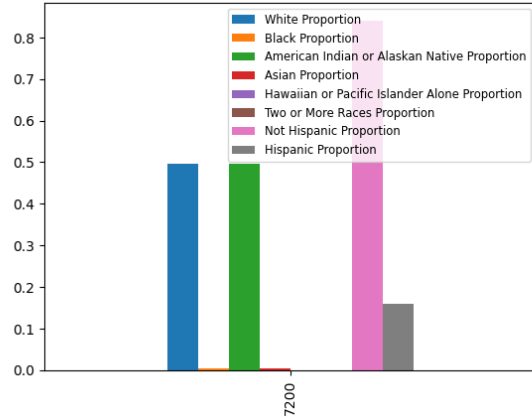
By setting a threshold to filter out the Districts with a high number of 'White Alone' population, we can better see which districts are more diverse and potentially see how the population of an ethnic group either increases or decreases throughout this time period.

Further, we can see that the Districts of 'South Central Alabama Development Commission' and 'South Delta Planning and Development District' have a more prominent 'Black Alone' population than a 'White Alone' population. Also, we can see in the 'Southwest Alaska Municipal Conference' District that from the years of 2016 to 2019, the 'White Alone' population is steadily decreasing. While the 'American Indian or Alaskan Native' and 'Asian Alone' populations stayed relatively neutral/steady.

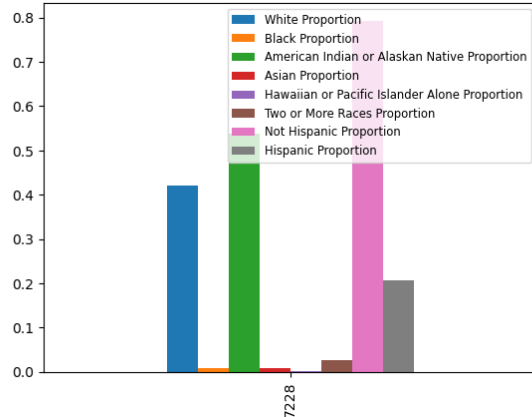
IV. ANALYSIS TASK 2

The second task was comparing the racial makeup of the Northwest New Mexico Council of Governments, a particular district, between the years 1990 and 2018. Through a comparison of the racial composition between these two years, any noteworthy shifts or patterns in the racial mix of this district over this time frame might be found.

Racial Composition of Northwest New Mexico Council of Governments in 1990



Racial Composition of Northwest New Mexico Council of Governments in 2018

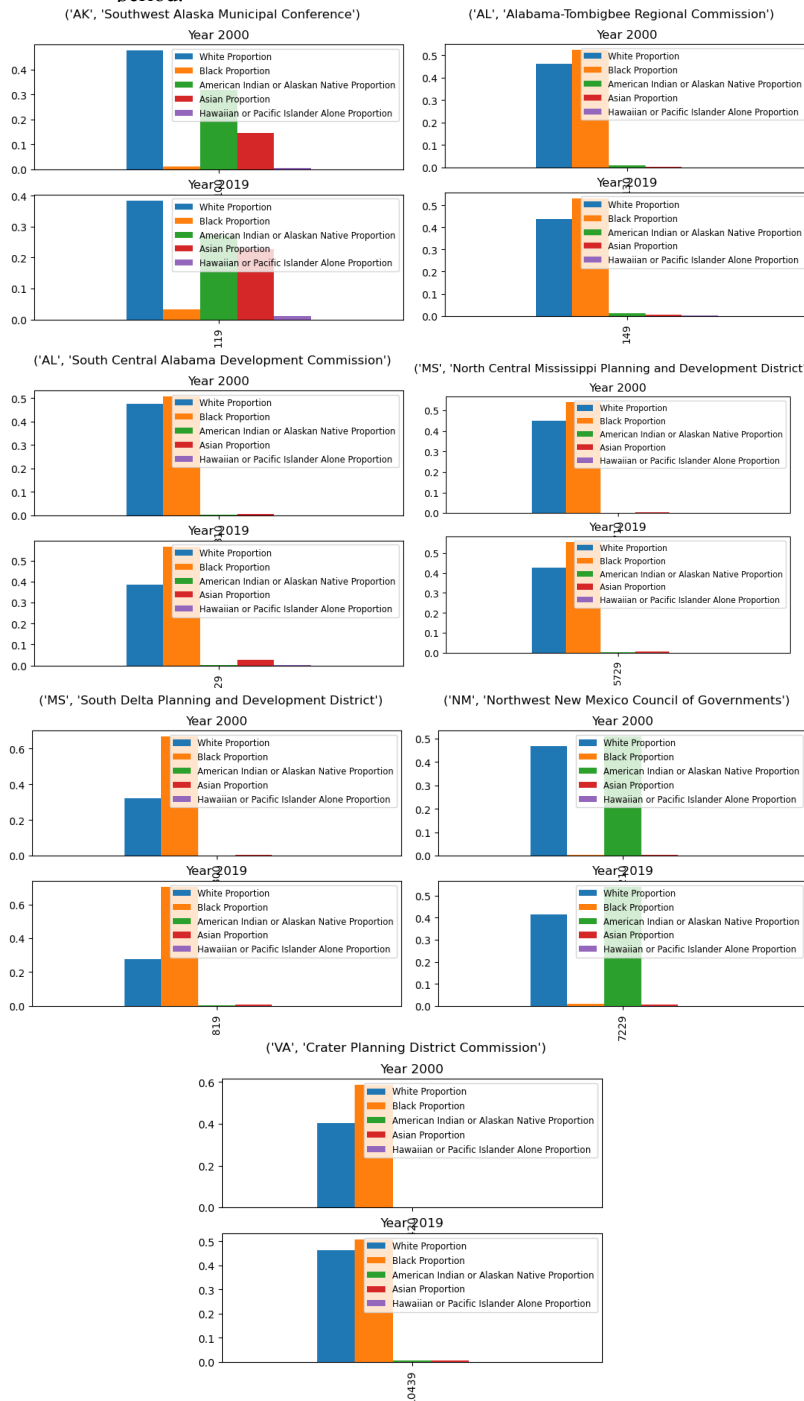


IV. ANALYSIS TASK 2- FINDINGS

Through this visualization of 'Racial Composition of Northwest New Mexico Council of Governments' from 1990 to 2018, we can see that the proportion of almost all ethnic groups changed. We can see that the proportion of 'White Alone' decreased, the proportion of 'Black Alone' slightly increased, the proportion of 'American Indian or Alaskan Native' increased slightly, the proportion of 'Asian Alone' increased slightly, the proportion of 'Hawaiian or Pacific Islander Alone' increased very slightly, the proportion of 'Two or More Races' increased slightly, the proportion of 'Not Hispanic' decreased, and the proportion of 'Hispanic Alone' increased slightly.

V. ANALYSIS TASK 3

The comparison study is broadened in the third task to cover all districts in two specific years, 1991 and 2019. This enables us to identify wider patterns and changes in racial makeup across all districts throughout this time period.



VI. ANALYSIS TASK 4

The fourth task involved identifying the district with the highest total population in a specific year, 2000. This information can be useful for understanding population distribution and identifying areas of high population density.

VI. ANALYSIS TASK 4- FINDINGS

The district with the highest total population in the year 2000 is the 'Kenai Peninsula Economic Development District' in Alaska (AK), with a total population of 49,618. The racial composition of this district in 2000 is also provided, with the White Alone population making up approximately 87.37% of the total population.

VII. ANALYSIS TASK 5

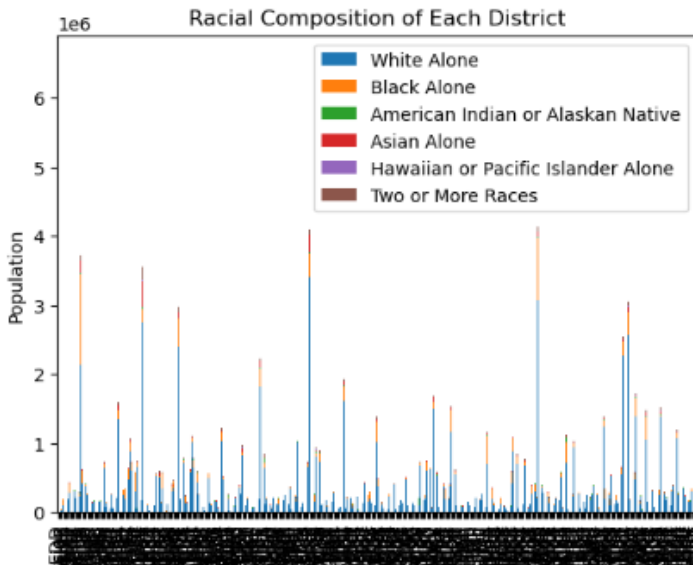
The final task was to identify the state with the greatest change in the 'White Alone' population over time. This research can give insights into state-level demographic transitions, emphasizing places where the 'White Alone' population has increased or decreased significantly over time. This data might help us comprehend wider demographic trends and patterns at the state level.

VII. ANALYSIS TASK 5- FINDINGS

Utah has had the greatest percentage shift in the 'White Alone' population over the years. The Wasatch Front Regional Council/Wasatch Front Economic Development District in Utah was affected by the change. The percentage change is around 20.03. This suggests that the 'White Alone' population within this Utah district has increased significantly.

VIII. ANALYSIS TASK 6

The code calculates the mean population of each racial group (White Alone, Black Alone, American Indian or Alaskan Native, Asian Alone, Hawaiian or Pacific Islander Alone, Two or More Races) for each district. This is done by grouping the data by 'District Name' and then calculating the mean for each racial group. The result is then visualized using a stacked bar plot, with each color in the bar representing a different racial group. The x-axis represents the districts, and the y-axis represents the population. This visualization provides a clear picture of the racial composition of each district.

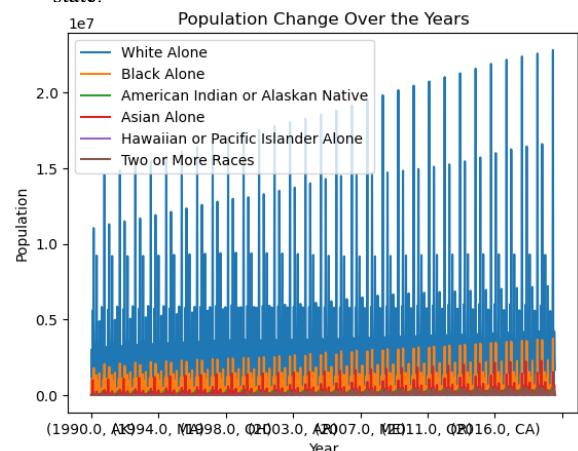


VIII. ANALYSIS TASK 6- FINDINGS

The stacked bar graph titled 'Racial Composition of Each District' reveals a predominant trend: a significant proportion of the population in most districts identifies as 'White Alone'. This graph also helps to answer my question "What is the racial composition of each district?" because the graph shows some diversity but most of the populations are largely made up of people who identify as 'White Alone.'

IX. ANALYSIS TASK 7

The population of each racial group for each year and each state is summed up by the code. To achieve this, first group the data according to "Year" and "State," and then figure out the total for each racial category. The outcome is then shown using a line plot, in which the population is represented by the y-axis and the years by the x-axis. This graphic gives an accurate depiction of how each racial group's population has evolved in each state.



IX. ANALYSIS TASK 7- FINDINGS

According to the plot, the numbers of the "White Alone" and "Black Alone" groupings have stayed mostly unchanged throughout time, whilst the numbers of the "Asian Alone" and "Two or More Races" groups have increased. Which answers my second question "How has the population of each race changed over the years in each state or district?"

This image illustrates how diversity has increased over time and offers insightful information on the demographic changes taking place in the US.

X. CONCLUSION

In conclusion, research on the racial and ethnic makeup of US districts has shed light on a number of important demographic trends and patterns. Using the right data preparation methods helped to overcome the difficulties this project experienced, such as handling missing data. Observing how various racial and ethnic groupings are distributed around the nation and tracking changes in these distributions over time are among the insights gathered from this approach.

X. CONCLUSION CONT.

For future exploration, this analysis opens up several interesting avenues. It would be worthwhile to investigate the factors contributing to the observed changes in racial composition, such as migration patterns or changes in birth and death rates. Additionally, correlating these demographic shifts with other socioeconomic indicators could provide a more holistic understanding of the changing demographics in the United States. This could potentially inform a wide range of policies and initiatives, from urban planning and housing to education and healthcare. Overall, this project underscores the importance of demographic analysis in understanding and addressing the needs of our diverse society.