

# PS6<sub>Reinsch</sub>

Emma Reinsch

March 2021

## 1 Data Cleaning Steps

1 Visualization 1: I obtained data from the FRED website. I obtained income overtime for African Americans in the United States. I filtered the data to only include years ranging from 1990 to March 2021. I also transformed the data into a data frame.

2 Visualization 2: I obtained my second form of data from the FRED website. I obtained income overtime for White/Caucasian individuals in the United States. I filtered the data to only include years ranging from 1990 to March 2021. Once again, I transformed the data set into a data.frame.

3 Visualization 3: I obtained my third form of data from the IPUMS USA website. I used American Community Survey data (ACS) and gathered several variables. My objective is to regress African Americans on income wage to see the residual wage gap. I mutated the "RACBLK" variable coded 1 for no (not black) and 2 for yes into a dummy variable named "black". After filtering income wage to include values from 10000 dollars to 300000 dollars and total income from 50000 dollars to 350000 dollars, I obtained the count of African American individuals as opposed to non-African American individuals. I intend to repeat the same for the "RACWHT" variable (coded 1 for non-white and 2 for white).

I also took the natural logarithm of the income variable. I intended to create a graph with race as the independent variable and log(incomewage) as the dependent variable, but ran into some issues. I will be making further adjustments to the data set.