

Peter Carlen and Mike Guzman

In order to get the data for the diseases that we were interested in we looked at the data the CDC has on the following diseases: Heart disease, diabetes, lower chronic respiratory, and Alzheimer's disease. Then we wanted to look at the health of people and how it compared to income and for that information, we looked at the IRS and the data they had on an individual's income, based on states.

In the beginning, we wanted to see how income and or wealth connected with wellness. Our first problem was that wellness was essentially too broad and we couldn't find data that just talked about the overall wellness of people so we decided to pick specific diseases. The variables we picked that dealt with diseases were chosen because they affect a large portion of the American population so we could get data that relates to a wide amount of people in the US. We downloaded the disease data into R, one csv file for each disease and one for income. We then used the `full_join()` function to combine each spreadsheet into one larger data set.

The method we are using to display this information is with a shiny app. This method would be interactive enough for a person to theoretically pick the variable they want to focus on and find the answer to whatever question they might be wondering about that relates to our findings.