For my final project I used a data set of information provided by the Center for Disease Control, having originated from the National Center for Health Statistics. The data set is entitled “Death rates and life expectancy at birth”. This data set contains five columns, or variables: Year, Race, Sex, Average Life Expectancy (in years), and Age-adjusted Death Rate. The question that I sought to answer revolved around the Average Life Expectancy variable, and the correlation between that expectancy and the Year. As a society we have come to understand that life expectancy has steadily increased over the years, most drastically in comparison to many (i.e. 100) years in the past. The outcome of my exploratory analysis on this data set does communication a relevant linear relationship between the Year variable and the Average Life Expectancy variable. Therefore, our understanding that later years mean a higher life expectancy is supported by these data. I feel that the main piece missing from my data wad additional variables. I fear that I may have misunderstood when I first began this project and selected a data set that has five variables but not all these variables have the ability to be statistically analyzed for the purposes of answering my intended question. With a larger number of variables, a deeper analysis could have been worked in order to obtain the best possible comprehension of the data. I would be interested in additional variables such as perhaps age of death of a parent or health status or another variable with the potential to influence life expectancy. I did face some challenges with my data set initially but I believe that I was making things more difficult than they needed to be. I was attempting to run functions essentially “from scratch” before I realized that I could utilize some of the pre-built functions provided in our textbook repository. This was incredibly helpful for me and allowed me to complete the project. However, I would like to revisit some of these concepts (for example, hypothesis testing) to try to gain a deeper understanding of how to do this, particularly when it concerns a more complex data set.