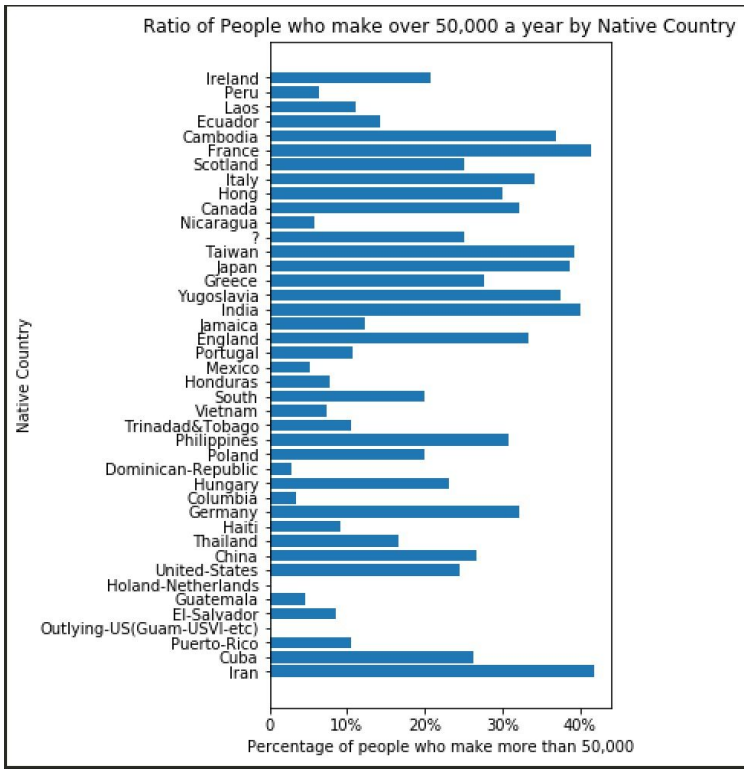
Nathan Struhs
CS 434
Homework 1

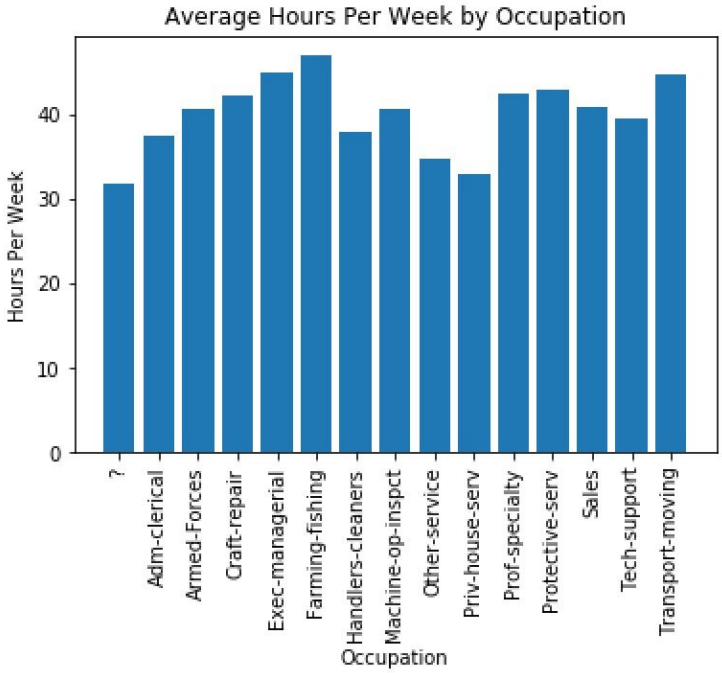
Exploring Data through Visualization

Ratio of people who make greater than 50k/year by native country



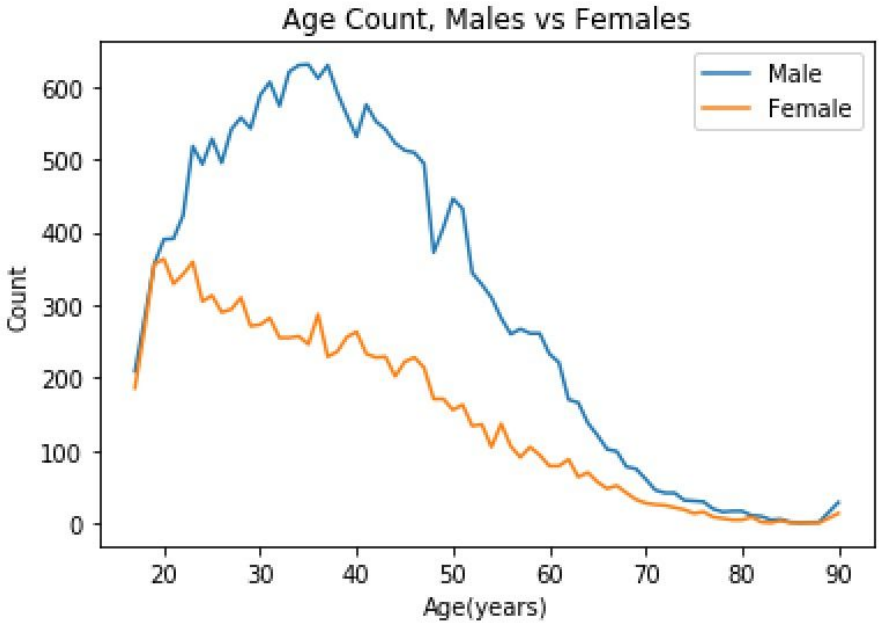
In this experiment I organized the data to find the most monetarily successful group by native country. I expected that the people originally from the United States to be the most successful since they would be established in the country and probably have a safety net of family around them. Though, what I found was that people from the United States were a little better than average, and people from France, Iran, and India do quite well. I found the trends by organizing the data by native country and then cross referencing the wealth standard column (‘<=50k, ‘>50k’). Once I found how many people in each category made over 50k, I could divide by the total and get the ratio. This data could be skewed due to the fact that some immigrants may be undocumented and did not participate in the census.

Average hours worked per week by occupation



The figure above shows each occupation in the census data and shows the respective average hours per week worked. I was expecting to see that farming and fishing workers would work the most which turned out to be true. Transport and moving professionals and executive/managerial came close behind. It can also be seen that private house services work the least amount in a week. These trends were found by organizing the data by occupation and creating an index on each. Then by cross referencing with hours per week and taking the mean, I could find the average hours worked per week.

Age count of males vs females



Here I was comparing the age count of men vs women in the census. I was expecting the trend of both sexes to be roughly the same while there being a discrepancy in the total count of male vs females. I found that even though the count of men in their 30s was higher than women, then trend was much higher as well. I found these trends by organizing the data into groups of males and females and then counted each by age.