Descriptive Statistics of US States along Sociological Metrics

By Ayo Akinnawonu

**Overview**

The dataset is about several sociological metrics of data about the people of the United States separated by the 50 states. I believe that it is important to track the quality of life of people from across the country and to this end we should look at a variety of factors that can have an influence on that quality of life. The whole dataset was created and posted at https://data.world/kevinnayar/us-states-sociological-metrics.

**About the Data**

> US\_States<-read.csv(file.choose())

> View(US\_States)

The dataset contains data about the 50 states of the United States along sociological metrics. The first column, Percent.Educational.Attainment, represents the percentage of the population that have completed higher education, this is defined as completion of a bachelor’s degree or higher. The information for this data set was sourced from the American Community Survey. The second column, Percent.Peace.Index, is sourced from visonfromhummanity.org and the index is compiled from data from the levels of homicide, violent crime, incarceration and the availability of fire arms. The higher the score, the safer the state is projected to be. The third column, Percent.Above.Poverty.Rate, has the percentage of households within each state that are above the poverty line. The data is originally from the US Census Bureau. The final column, Percent.Non.religious, represents the percentage of people that do not report themselves as highly religious. The information comes from the Pew Research Center.

**Analysis**

Column Means

> colMeans(US\_States[,-1])

Percent.Educational.Attainment Percent.Peace.Index

27.2040 49.4392

Percent.Above.Poverty.Rate Percent.Non.religious

85.1980 45.3000

Correlation

|  |
| --- |
| > cor(US\_States[,-1])  Percent.Educational.Attainment Percent.Peace.Index  Percent.Educational.Attainment 1.0000000 0.3579645  Percent.Peace.Index 0.3579645 1.0000000  Percent.Above.Poverty.Rate 0.6783820 0.6059370  Percent.Non.religious 0.6375694 0.6125116  Percent.Above.Poverty.Rate Percent.Non.religious  Percent.Educational.Attainment 0.6783820 0.6375694  Percent.Peace.Index 0.6059370 0.6125116  Percent.Above.Poverty.Rate 1.0000000 0.6706585  Percent.Non.religious 0.6706585 1.0000000 |
|  |
| There are no strong correlations (>0.8) between the any of the sociological metrics. There are moderate correlations between poverty and the other metrics as well as religion versus and the other three. The piece of information that stands out the most is that the education metric and the peace index has the lowest correlation. This demonstrates that more education in a population does not make people safer at the same extent that it perhaps keeps people out of poverty.  Covariance  cov(US\_States[,-1])  Percent.Educational.Attainment Percent.Peace.Index  Percent.Educational.Attainment 22.998759 21.11445  Percent.Peace.Index 21.114452 151.27783  Percent.Above.Poverty.Rate 9.983682 22.87069  Percent.Non.religious 32.845714 80.92841  Percent.Above.Poverty.Rate Percent.Non.religious  Percent.Educational.Attainment 9.983682 32.84571  Percent.Peace.Index 22.870692 80.92841  Percent.Above.Poverty.Rate 9.417343 22.10878  Percent.Non.religious 22.108776 115.39796   |  | | --- | | ggplot(US\_States, aes(x=Percent.Educational.Attainment, y=Percent.Peace.Index, color=State)) + geom\_point() + geom\_text(aes(label=State),size=2,hjust=0,vjust=0) + labs(title="Education vs. Peace") | |
| The peace index is based on multiple factors outside the control of an individual as opposed to pursing higher education which can be viewed as the pursuits of an individual. Higher education can lead people to more opportunities which perhaps lead people to not have to consider illegal activities. This graph is important because it represents the weakest correlation of all the sociological metrics, and it demonstrates the education or at least the opportunity of pursuing it does not necessarily raise the standards of peace. The correlation is likely as weak as it is because the outliers are inconsistent to the trend of the data. For example, Louisiana has the lowest peace index but not the lowest educational level although it is near the bottom. Maine on the other hand has the highest peace index but its education level in comfortably in the middle. Another outlier, West Virginia, has a relatively high score in the peace index of 56, which is 7 points above the colMeans but, the state has the lowest education attainment with a score of only 17, a full 10 points below that column’s average.  cor\_US<-cor(US\_States[,-1])  corrplot.mixed(cor\_US, lower="number", upper="circle", tl.pos=c("d"), tl.cex=.4)    The matrix demonstrates that there is a positive correlation between all the metrics. The column Percent.Non.religious has the most consistent correlation with all three correlations between 0.6-0.7. I believe that education is the metrics that is influence the other three. A more educated populace may have the opportunities to not have to commit crimes or live in areas where one would be more likely to be a victim of violent crime. Higher education is widely believed to lead to greater economic potential and beneficiaries of it may have been exposed to different perspectives on life and this can increase the likelihood of critical thinking in previous accepted beliefs. Such reflection can be the reason that there is a distinct correlation between religious belief and educational attainment.  ggplot(US\_States, aes(x=Percent.Educational.Attainment, y=Percent.Above.Poverty.Rate, color=State)) + geom\_point() + geom\_text(aes(label=State),size=2,hjust=0,vjust=0) + labs(title="Education vs. Poverty")    Education is a signal of opportunity in a community. Although it is not necessary in order to succeed in America, the opportunity to pursue higher education is indicative of a state having a higher percentage of individuals who are financially solvent. As compared to the plot of “Education vs Peace,” the above plot as a tighter and more consistent trend. The states with the lowest educational attainment also have the highest rates of poverty. The state with the lowest level of poverty, New Hampshire does have an educational attainment that is over 30 which would be above average. I do not see a clear distinction of region of the US with regards to the data. The lowest performing states are southern and southeastern while the highest performing states are from across country with northern, southern and west states.  ggplot(US\_States, aes(x=Percent.Non.religious)) + geom\_histogram(binwidth = 1, color="black") + labs(title="Religion in America")    The number of people in the United States who profess their faith for an organized religion appears to be declining. The colMeans of the Percent.Non.religious is about 45 which is slightly lower than the mode, which is 46. This demonstrates that around half of the population of the US does not consider themselves to be overtly religious. Out of 50 states, only 19 have a score that is below the average. That means that 19 states have percentages of their populace that are greater than average. The States with the lowest score in this column are mostly congregated in the south while the four highest scores all come from New England where 2/3 of population appear to be non-religious according to the data.  Bibliogrpahy  Nayar, Kevin. “Us-States-Sociological-Metrics - Dataset by Kevinnayar.” *Data.world*, 11 Mar. 2018, data.world/kevinnayar/us-states-  sociological-metrics. |
|  |