

# Exploration of Poverty and Murder By State

2024-03-01

Description Using the crime\_data data, I will look into states with the highest murder rates, as well as the variables to describe a possible relationship between being impoverished and murder rates. So the variables I will be looking at and using are 'murder', 'poverty', and 'state'. And the question is, what is more prevalent: murder in a specific state or within poverty? I will also attempt to find any relationship between poverty and/or state. I will find the average poverty rate, as well as the average crime rate. I will compare the two on a graph, and then see if most states have a higher mean crime rate or poverty rate. So the variables of interest are state, crime, and poverty. Using several graphs, I will compare the variables against each other as well as the states.

#Univariate Exploration. The following code was used to set and create specific variables as well compare them to one another while exploring them. Such as telling RStudio what variables mean what, and looking at the variables poverty and murder, as well as creating their mean counterparts for later coding. We will also summarize to further familiarize ourselves with said variables and the data set as a whole.

```
library(ggplot2)
library(readxl)
crime_data<-read_xlsx("/Users/zacharydiblasia/Downloads/Crime_Data.xlsx")

state<-(crime_data$state)
murder <-(crime_data$murder)
murder_rate<-(mean(murder))
poverty_rate<-(mean(crime_data$poverty))
poverty<-(crime_data$poverty)
```

#Summary of poverty

```
summary(crime_data$murder)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  1.600   3.900   6.800   8.727 10.350   78.500
```

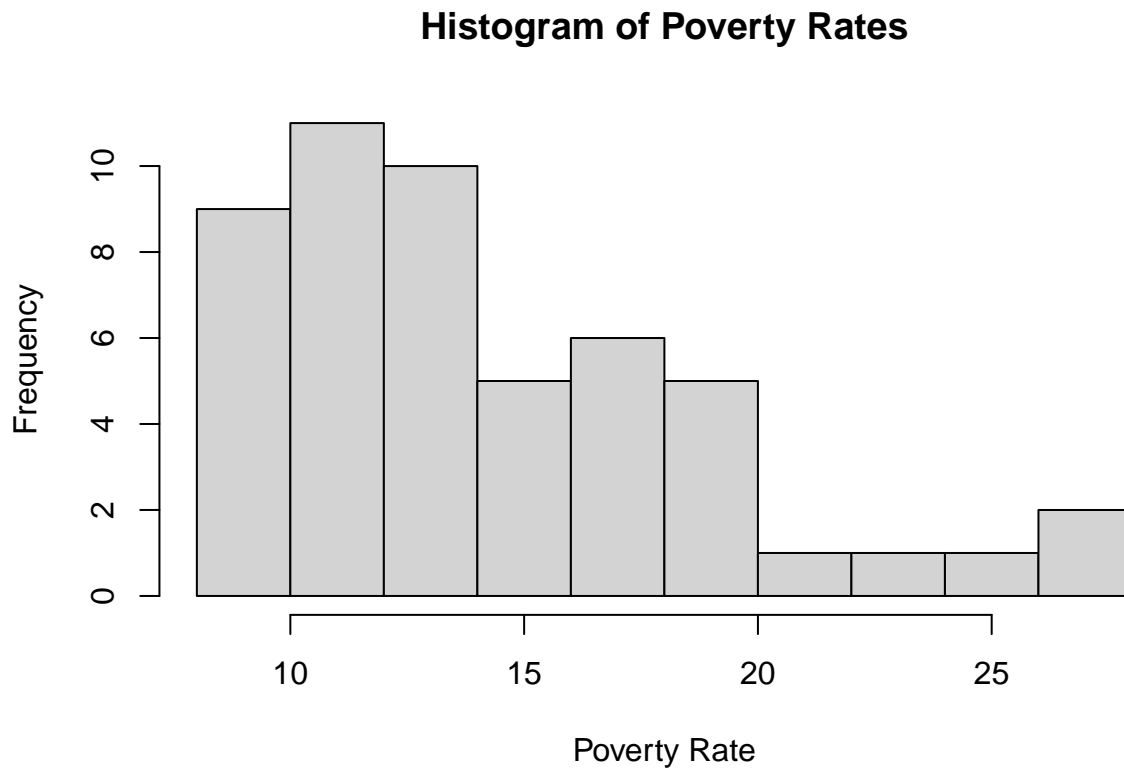
#Summary of Murder

```
summary(crime_data$poverty)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   8.00   10.70   13.10   14.26   17.40   26.40
```

## Histogram of poverty rates

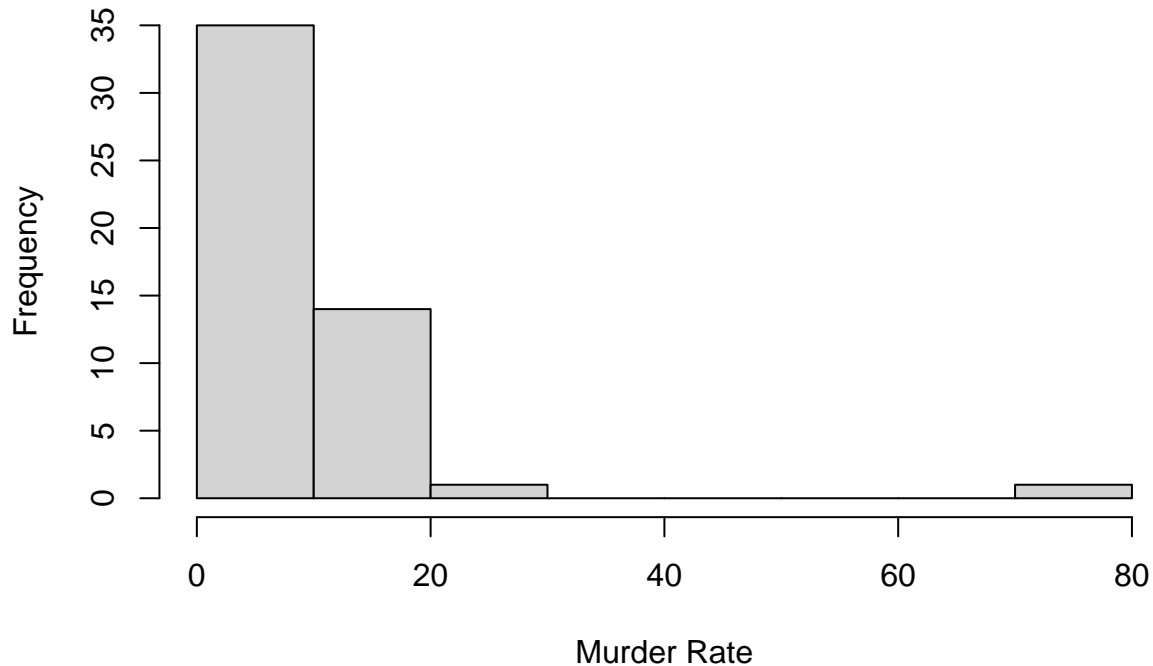
```
hist(crime_data$poverty, main = "Histogram of Poverty Rates", xlab = "Poverty Rate", ylab = "Frequency")
```



### Average Murder Rate by Histogram

```
hist(crime_data$murder, main = "Histogram of Murder Rates", xlab = "Murder Rate", ylab = "Frequency")
```

## Histogram of Murder Rates



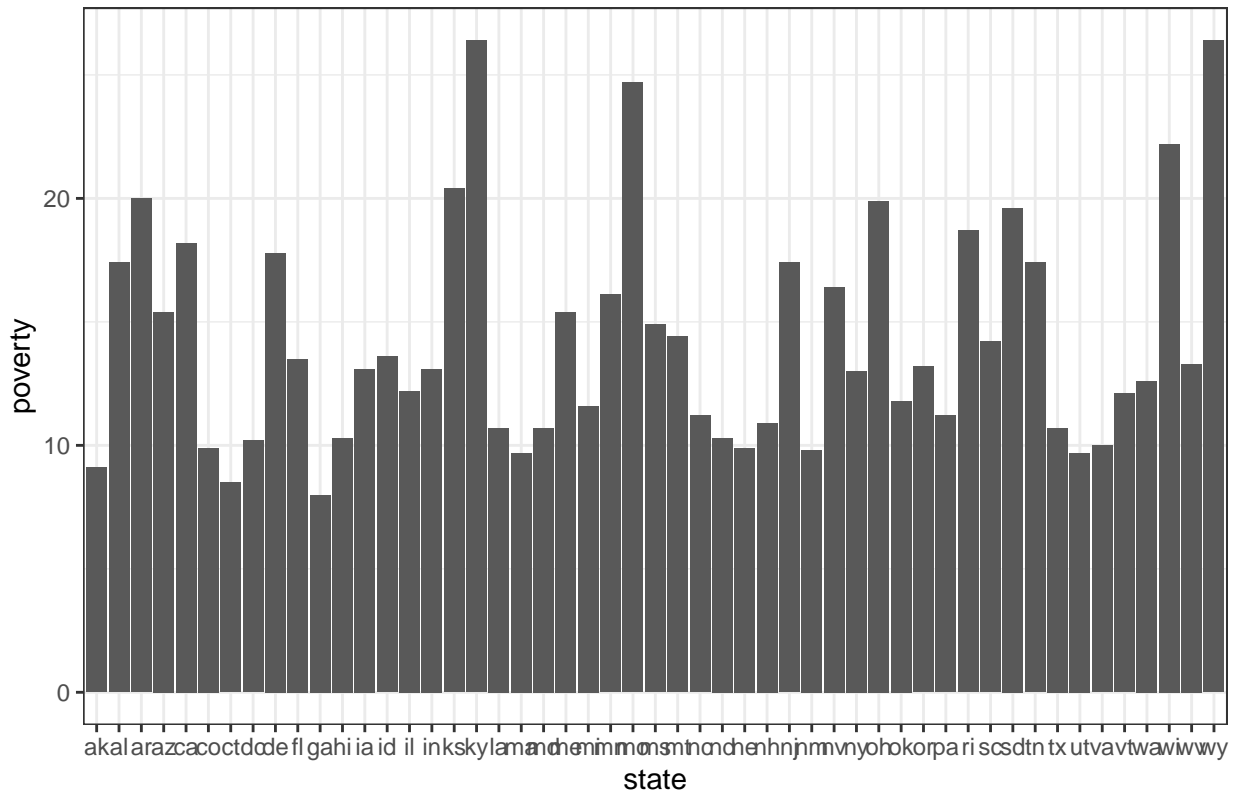
#Bivariate Exploration Now, we will explore established relationships regarding murder and poverty within different states as well as between each other. The data will be visualized using coded graphs, which could show a correlation or lack thereof. We will use bar graphs to represent poverty and murder by state respectively, and a scatter plot to examine the potential relationship between poverty and murder. This approach will allow us to determine if the graphs exhibit similarities or differences, allowing us to draw descriptive conclusions

#Poverty of State

```
ggplot(crime_data, aes(x=state)) + theme_bw() +  
  geom_bar(aes(y =..poverty..)) + ggtitle("Poverty By State")
```

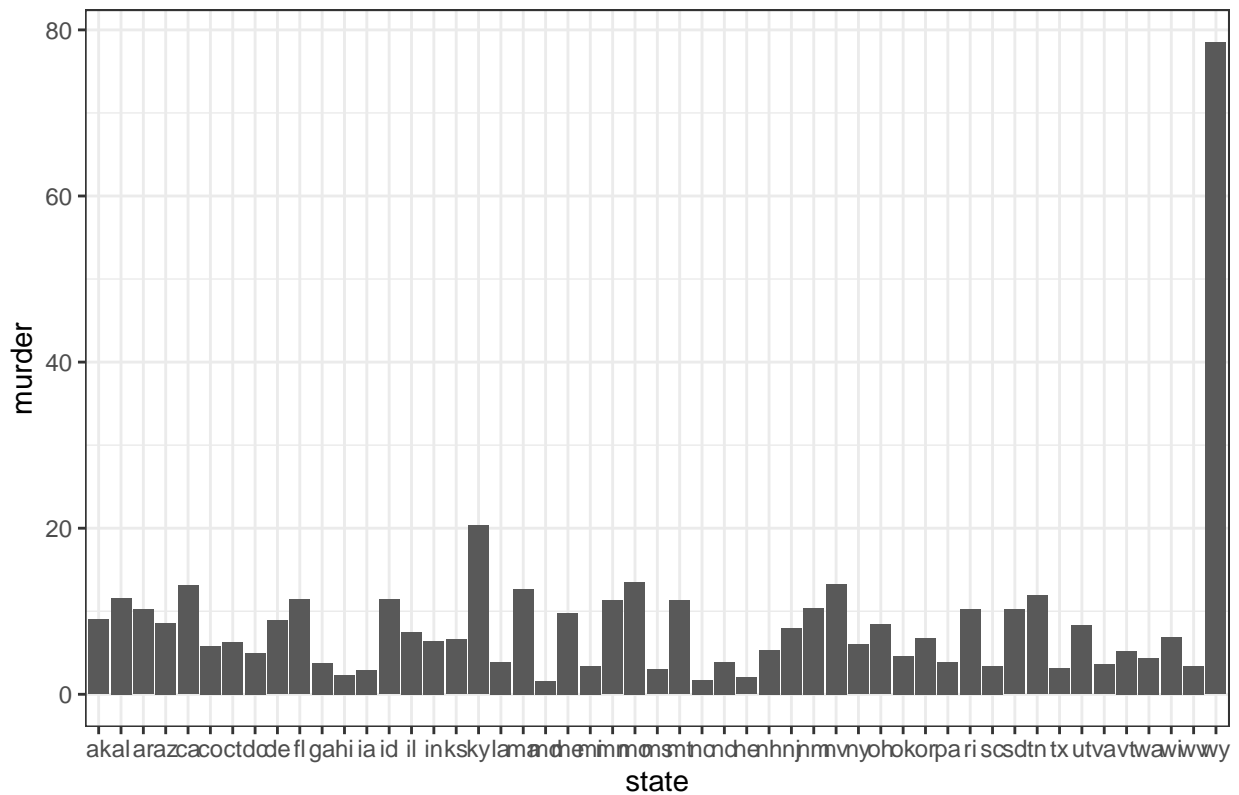
```
## Warning: The dot-dot notation ('..poverty..') was deprecated in ggplot2 3.4.0.  
## i Please use 'after_stat(poverty)' instead.  
## This warning is displayed once every 8 hours.  
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was  
## generated.
```

# Poverty By State



}} # Murder by State

## Murders By State



## Bivariate Exploration

### Grouped summary statistics - mean murder rates by poverty levels

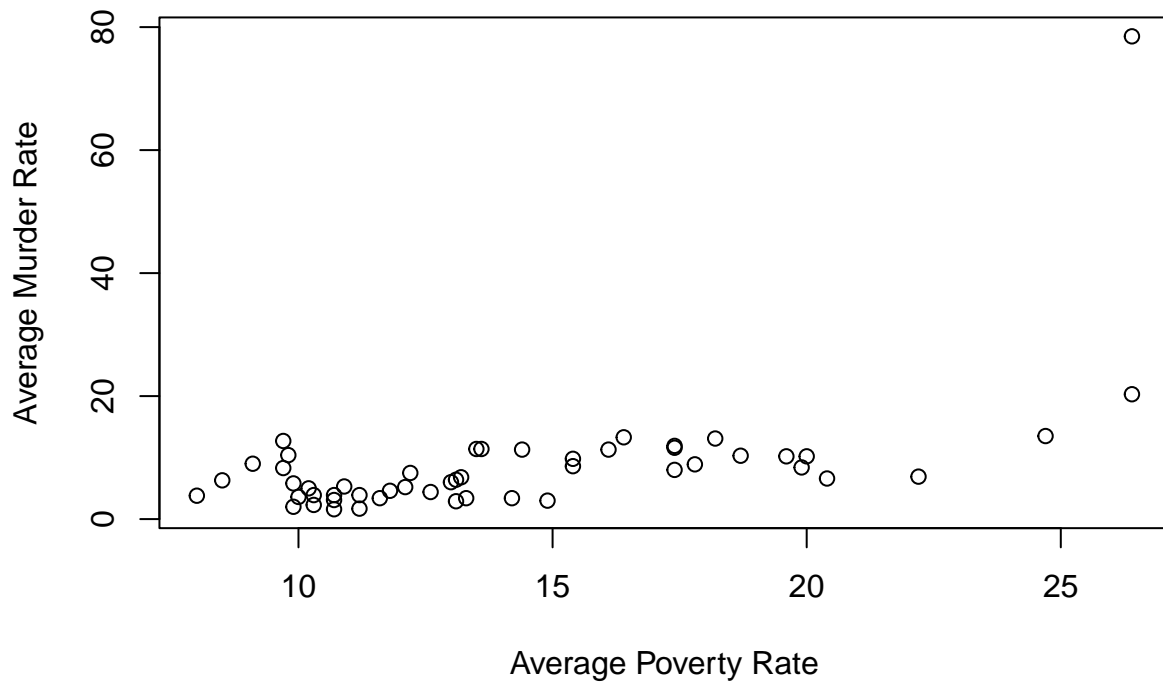
```
aggregate(murder ~ cut(poverty, breaks = 4), data = crime_data, FUN = mean)
```

```
## cut(poverty, breaks = 4) murder
## 1 (7.98,12.6] 5.15
## 2 (12.6,17.2] 7.56
## 3 (17.2,21.8] 9.92
## 4 (21.8,26.4] 29.80
```

```
#Scatterplot of Avg Murder vs Avg Poverty Rates
```

```
plot(poverty, murder,
     xlab = "Average Poverty Rate",
     ylab = "Average Murder Rate",
     main = "Average Murder Rate vs. Average Poverty Rate by State")
```

## Average Murder Rate vs. Average Poverty Rate by State



#Avg Crime Vs Avg Poverty

```
mean_poverty_rate <- mean(poverty)
mean_murder_rate <- mean(murder)

cat("Mean Poverty Rate:", mean_poverty_rate, "\n")
```

## Mean Poverty Rate: 14.25882

```
cat("Mean Murder Rate:", mean_murder_rate, "\n")
```

## Mean Murder Rate: 8.727451

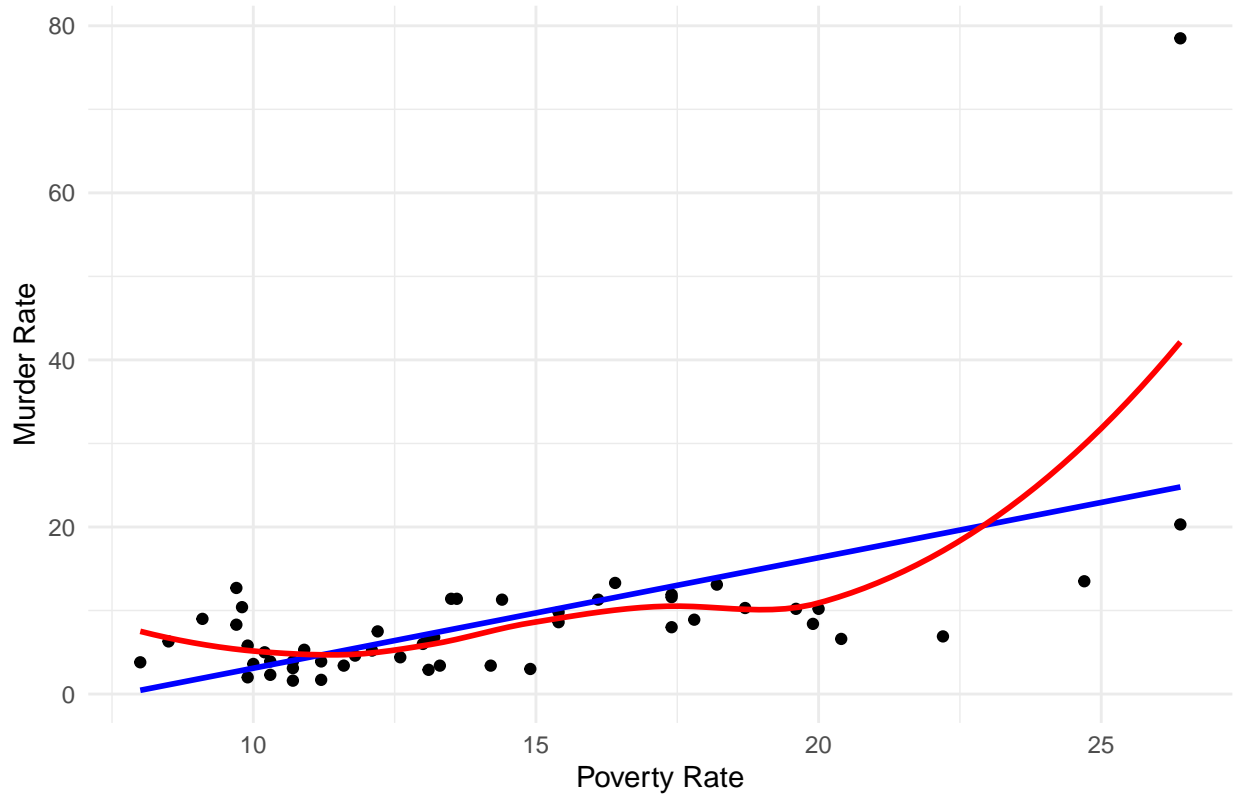
## Scatter plot of murder rates vs. poverty rates

```
ggplot(crime_data, aes(x = poverty, y = murder)) +
  geom_point() +
  labs(title = "Murder Rates vs. Poverty Rates", x = "Poverty Rate", y = "Murder Rate") +
  theme_minimal() +
  geom_smooth(se=FALSE, method="lm", color="blue") +
  geom_smooth(se=FALSE, color="red")
```

## 'geom\_smooth()' using formula = 'y ~ x'

## 'geom\_smooth()' using method = 'loess' and formula = 'y ~ x'

## Murder Rates vs. Poverty Rates



#Conclusion From the analysis, we observe a wide range of murder rates and poverty rates that truly differ across each state. The histogram we created using poverty and murder rates, along with the summary statistics, provide an overview of the distribution of these variables. The scatter plot we generated shows a potential positive relationship between murder rates and poverty rates, essentially suggesting that higher poverty rates may correlate with higher murder rates, and vice versa. However, further analysis would be necessary to confirm this relationship, and more data availability would be beneficial in determining our validity. A key observation is the outlier in both murder and poverty rates respectively, which is D.C., exhibiting high poverty and murder rates.