

# Final Project

Peter Koenen

2022-09-19

## #Data Description

We will be looking at Regional data on crime and murder rates. However, the variables we will be focusing on is poverty and crime by state. Is there a correlation between poverty and crime/murder among certain United States Regions?

```
head(crime)
```

```
## # A tibble: 6 x 11
##   sid state region division  crime murder pctme~1 pctwh~2 pcths poverty single
##   <dbl> <chr> <chr> <chr>    <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1     1 ak    West  Pacific    761     9    41.8   75.2  86.6   9.10  14.3
## 2     2 al    South East Sou~    780  11.6   67.4   73.5  66.9  17.4  11.5
## 3     3 ar    South West Sou~    593  10.2   44.7   82.9  66.3   20   10.7
## 4     4 az    West  Mountain    715   8.60   84.7   88.6  78.7  15.4  12.1
## 5     5 ca    West  Pacific   1078  13.1   96.7   79.3  76.2  18.2  12.5
## 6     6 co    West  Mountain    567   5.80   81.8   92.5  84.4   9.90  12.1
## # ... with abbreviated variable names 1: pctmetro, 2: pctwhite
```

## #Univariate Description

The first variable being analyzed is poverty. To begin, the summary will give a general idea of the poverty rates by each state. The line graph represents the relationship between poverty and each state. With density representing the proportion of people impoverished by state.

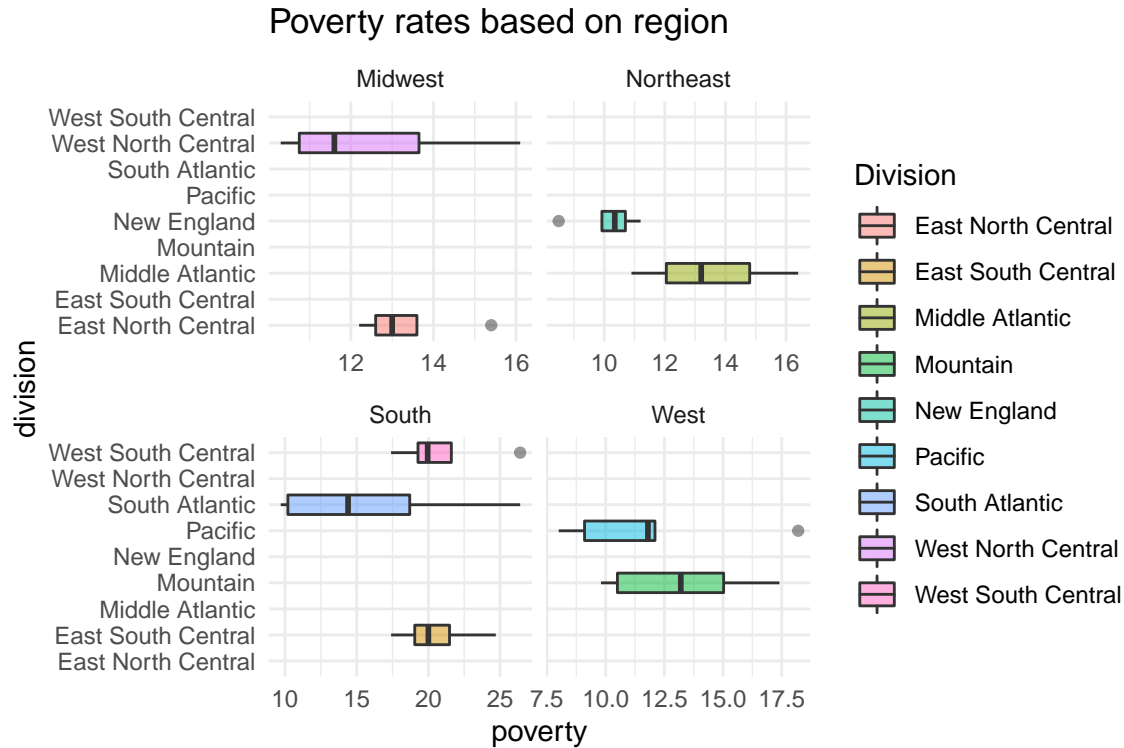
```
table(crime$poverty)
```

```
##
##           8           8.5 9.100000381 9.699999809 9.800000191 9.899999619
##           1           1           1           2           1           2
##          10 10.19999981 10.30000019 10.69999981 10.89999962 11.19999981
##           1           1           2           3           1           2
## 11.60000038 11.80000019 12.10000038 12.19999981 12.60000038           13
##           1           1           1           1           1           1
## 13.10000038 13.19999981 13.30000019           13.5 13.60000038 14.19999981
##           2           1           1           1           1           1
## 14.39999962 14.89999962 15.39999962 16.10000038 16.39999962 17.39999962
##           1           1           2           1           1           3
## 17.79999924 18.20000076 18.70000076 19.60000038 19.89999962           20
##           1           1           1           1           1           1
## 20.39999962 22.20000076 24.70000076 26.39999962
##           1           1           1           2
```

```
summary(crime$poverty)
```

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

```
ggplot(crime, aes(x=division, fill=division,y=poverty))+geom_boxplot(alpha=.5)+facet_wrap(~region,scales=
```



This histogram graph shows that among the regions with the highest amount of poverty are the West South Central, East South Central, Middle Atlantic, East North Central, and the Mountain Region. The next variable is murder. The bar graph compares murder rates among regions.

```
table(crime$murder)
```

```
##
## 1.600000024 1.700000048          2 2.299999952 2.900000095          3
##          1          1          1          1          1          1
## 3.099999905 3.400000095 3.599999905 3.799999952 3.900000095 4.400000095
##          1          3          1          1          3          1
## 4.599999905          5 5.199999809 5.300000191 5.800000191          6
##          1          1          1          1          1          1
## 6.300000191 6.400000095 6.599999905 6.800000191 6.900000095          7.5
##          1          1          1          1          1          1
##          8 8.300000191 8.399999619 8.600000381 8.899999619          9
##          1          1          1          1          1          1
## 9.800000191 10.19999981 10.30000019 10.39999962 11.30000019 11.39999962
##          1          2          1          1          2          2
## 11.60000038 11.89999962 12.69999981 13.10000038 13.30000019          13.5
##          1          1          1          1          1          1
```

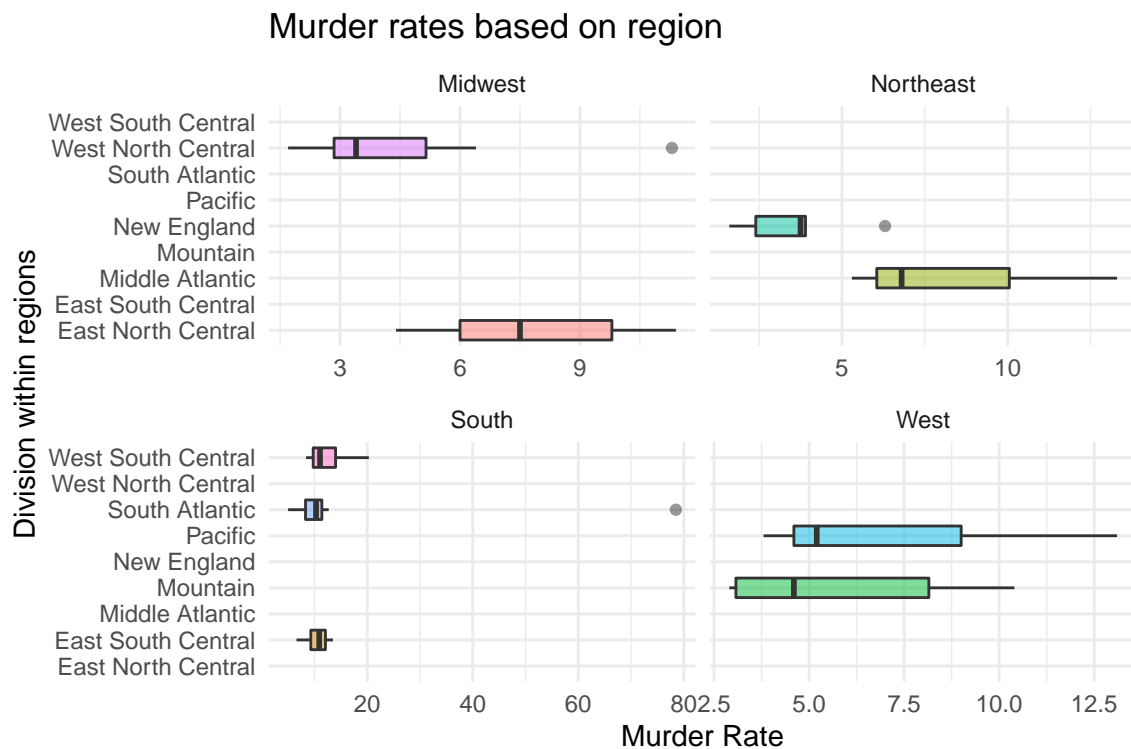
```
## 20.29999924      78.5
##           1          1
```

```
summary(crime$murder)
```

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

```
ggplot(crime, aes(x=division, fill=division,y=murder))+geom_boxplot(alpha=.5)+facet_wrap(~region,scales=
```

```
## Warning: It is deprecated to specify 'guide = FALSE' to remove a guide. Please
## use 'guide = "none"' instead.
```



As seen by this graph, murder numbers are highest in the South Region, with the West South Central, West North Central, and East South Central having the highest median murder numbers. With South Atlantic having a max murder number of 78.5 due to Washington D.C.

#### #Conclusion

This data shows that regions with high poverty such as East South Central, West South Central have high levels of poverty along with high murder rates. However, the region of West North Central has a moderate rate of poverty but is one of the highest ranking regions for murder rate. This data shows that although poverty may be correlated to murder rate, it may not be a conclusive factor.