

Protest by City Analysis

2025-03-24

Load and Clean Data

```
ProtestByCity <- read.csv("/Users/aricaschuett/Documents/protest/ProtestByCity3-24.csv", row.names = NU

# Replace NAs with zero for selected variables
vars_to_replace <- c("AntiTrumpProtestPreGFCount", "ProtestTotal", "BlackVictimCount", "VictimCountPreG

ProtestByCity[vars_to_replace] <- ProtestByCity[vars_to_replace] %>%
  mutate(across(everything(), ~replace_na(.x, 0)))
```

Create New Variables

```
ProtestByCity <- ProtestByCity %>%
  mutate(PoliceDeathsRate = VictimCountPreGF / population,
         BlackVictimsRate = BlackVictimCountPreGF / population,
         VictimRaceRatio = BlackVictimCountPreGF / VictimCountPreGF)
```

Summaries

```
summary(ProtestByCity$BlackPovRate)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.     NA's
##    0.000   0.946    1.000   0.893   1.000    1.000    4358
```

```
table(ProtestByCity$BlackPop == 0)
```

```
##
## FALSE  TRUE
## 15391  4358
```

```
### Create Summary Tables for Cities with Protest and those without
```

```
NoProtestCities <- ProtestByCity %>%
  filter(PostGFProtestCount == 0)
```

```
summary(NoProtestCities$population)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      501      963     1967    5090    4722   617790
```

```
summary(NoProtestCities$BlackPopPct)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.00000 0.00000 0.01290 0.08319 0.06667 1.00000
```

```
summary(NoProtestCities$EduRate)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.     NA's
```

```
## 0.0000 0.1241 0.1892 0.2365 0.3020 1.0000 3
```

```
summary(NoProtestCities$CollegeStudents)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.0    25.0    63.0   250.8   193.0 29984.0
```

```
summary(NoProtestCities$AntiTrumpProtestPreGFCCount)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
## 0.0000 0.0000 0.0000 0.2003 0.0000 28.0000
```

```
summary(NoProtestCities$VictimsCount)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
## 0.00000 0.00000 0.00000 0.02595 0.00000 9.00000
```

```
ProtestCities <- ProtestByCity %>%
  filter(PostGFProtestCount > 0)
```

```
summary(ProtestCities$population)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      505    5891   15354   46648   39894 8419316
```

```
summary(ProtestCities$BlackPopPct)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
## 0.00000 0.01296 0.03965 0.10149 0.11755 0.97372
```

```
summary(ProtestCities$EduRate)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
## 0.0192 0.2000 0.2880 0.3323 0.4329 0.9356
```

```
summary(ProtestCities$CollegeStudents)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.0    207.8    729.0   3168.3   2474.2 469139.0
```

```
summary(ProtestCities$AntiTrumpProtestPreGFCCount)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
## 0.000 0.000 2.000 7.663 6.000 767.000
```

```
summary(ProtestCities$VictimsCount)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
## 0.0000 0.0000 0.0000 0.7348 1.0000 61.0000
```

Regression Models

Model 8

```
m8 <- lm(PostGFProtestCount ~ population + BlackPopPct + BlackPovRate + EduRate + AntiTrumpProtestPreGFCCount)
summary(m8)
```

```
##
## Call:
## lm(formula = PostGFProtestCount ~ population + BlackPopPct +
##      BlackPovRate + EduRate + AntiTrumpProtestPreGFCCount + PoliceDeathsRate,
```

```
##      data = ProtestByCity)
##
## Residuals:
##      Min        1Q      Median        3Q        Max
## -114.056   -0.095     0.205     0.349   222.729
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.576e-01  1.235e-01   1.276  0.20207
## population      4.666e-05  5.459e-07  85.486 < 2e-16 ***
## BlackPopPct     -2.376e-01  1.720e-01  -1.381  0.16731
## BlackPovRate    -3.285e-01  1.249e-01  -2.631  0.00852 **
## EduRate         -8.954e-01  1.833e-01  -4.884 1.05e-06 ***
## AntiTrumpProtestPreGFCCount  3.246e-01  3.997e-03  81.218 < 2e-16 ***
## PoliceDeathsRate -5.992e+02  7.059e+02  -0.849  0.39603
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.781 on 15381 degrees of freedom
## (4361 observations deleted due to missingness)
## Multiple R-squared:  0.8233, Adjusted R-squared:  0.8233
## F-statistic: 1.195e+04 on 6 and 15381 DF,  p-value: < 2.2e-16
```

Model 9

```
m9 <- lm(PostGFPProtestCount ~ population + BlackPopPct + BlackPovRate + EduRate + BlackVictimsRate, data = ProtestByCity)
summary(m9)
```

```
##
## Call:
## lm(formula = PostGFPProtestCount ~ population + BlackPopPct +
##      BlackPovRate + EduRate + BlackVictimsRate, data = ProtestByCity)
##
## Residuals:
##      Min        1Q      Median        3Q        Max
## -150.903   -0.299     0.157     0.359   237.174
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    6.974e-02  1.475e-01   0.473  0.6364
## population      8.234e-05  3.874e-07 212.547 < 2e-16 ***
## BlackPopPct     9.907e-02  2.061e-01   0.481  0.6308
## BlackPovRate    -5.963e-01  1.492e-01  -3.997 6.45e-05 ***
## EduRate         4.117e-01  2.183e-01   1.886  0.0593 .
## BlackVictimsRate 1.440e+03  2.605e+03   0.553  0.5804
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.519 on 15382 degrees of freedom
## (4361 observations deleted due to missingness)
## Multiple R-squared:  0.7476, Adjusted R-squared:  0.7475
## F-statistic: 9110 on 5 and 15382 DF,  p-value: < 2.2e-16
```

Model 10

```
m10 <- lm(PostGFPProtestCount ~ population + BlackPopPct + BlackPovRate + EduRate + VictimRaceRatio, data = ProtestByCity)
summary(m10)

##
## Call:
## lm(formula = PostGFPProtestCount ~ population + BlackPopPct +
##     BlackPovRate + EduRate + VictimRaceRatio, data = ProtestByCity)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -153.890   -2.915    0.133    2.707   232.033
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   8.797e-01  2.621e+00   0.336 0.737169
## population     8.377e-05  1.334e-06  62.784 < 2e-16 ***
## BlackPopPct    5.561e+00  2.828e+00   1.967 0.049446 *
## BlackPovRate  -7.093e+00  2.795e+00  -2.538 0.011278 *
## EduRate        1.142e+01  2.963e+00   3.852 0.000123 ***
## VictimRaceRatio 8.451e-01  1.246e+00   0.678 0.497677
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.73 on 1209 degrees of freedom
## (18534 observations deleted due to missingness)
## Multiple R-squared:  0.7715, Adjusted R-squared:  0.7706
## F-statistic: 816.5 on 5 and 1209 DF,  p-value: < 2.2e-16
```

Additional Models (11-19)

For brevity, additional models are summarized below:

```
models <- list(
  m11 = lm(PostGFPProtestCount ~ population + BlackPop + BlackPov + CollegeEdTotal + VictimCountPreGF, data = ProtestByCity)
  m12 = lm(PostGFPProtestCount ~ population + BlackPop + BlackPov + CollegeEdTotal + VictimsPreGF2020, data = ProtestByCity)
  m13 = lm(PostGFPProtestCount ~ population + BlackPop + BlackPov + CollegeEdTotal + VictimsPreGF2020Blk, data = ProtestByCity)
  m14 = lm(PostGFPProtestCount ~ population + BlackPopPct + BlackPovRate + EduRate + VictimCountPreGF, data = ProtestByCity)
  m15 = lm(PostGFPProtestCount ~ population + BlackPopPct + BlackPovRate + EduRate + BlackVictimCountPreGF, data = ProtestByCity)
  m16 = lm(PostGFPProtestCount ~ population + BlackPopPct + BlackPovRate + EduRate + PoliceDeathsRate, data = ProtestByCity)
  m17 = lm(PostGFPProtestCount ~ population + BlackPopPct + BlackPovRate + EduRate + BlackVictimsRate, data = ProtestByCity)
  m18 = lm(PostGFPProtestCount ~ population + BlackPopPct + BlackPovRate + EduRate + VictimRaceRatio, data = ProtestByCity)
  m19 = lm(PostGFPProtestCount ~ population + BlackPopPct + BlackPovRate + EduRate + AntiTrumpProtestPre, data = ProtestByCity)
)

# Display summaries
lapply(models, summary)

## $m11
##
## Call:
## lm(formula = PostGFPProtestCount ~ population + BlackPop + BlackPov +
##     CollegeEdTotal + VictimCountPreGF, data = ProtestByCity)
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -130.468   -0.167    0.014    0.057   172.062
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -6.010e-02  2.458e-02  -2.445   0.0145 *
## population    -2.562e-05  1.925e-06 -13.312 <2e-16 ***
## BlackPop       1.004e-03  1.072e-04   9.365 <2e-16 ***
## BlackPov      -9.271e-04  1.097e-04  -8.453 <2e-16 ***
## CollegeEdTotal  3.709e-04  5.990e-06  61.915 <2e-16 ***
## VictimCountPreGF -7.914e-01  3.673e-02 -21.548 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.367 on 19743 degrees of freedom
## Multiple R-squared:  0.8206, Adjusted R-squared:  0.8206
## F-statistic: 1.806e+04 on 5 and 19743 DF,  p-value: < 2.2e-16
##
##
## $m12
##
## Call:
## lm(formula = PostGFProtestCount ~ population + BlackPop + BlackPov +
##      CollegeEdTotal + VictimsPreGF2020, data = ProtestByCity)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -135.050   -0.161    0.007    0.046   167.498
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -3.544e-02  2.462e-02  -1.439   0.15
## population    -3.961e-05  1.655e-06 -23.930 < 2e-16 ***
## BlackPop       8.886e-04  1.070e-04   8.307 < 2e-16 ***
## BlackPov      -7.971e-04  1.094e-04  -7.288 3.25e-13 ***
## CollegeEdTotal  3.980e-04  5.646e-06  70.490 < 2e-16 ***
## VictimsPreGF2020 -3.103e+00  1.654e-01 -18.764 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.377 on 19743 degrees of freedom
## Multiple R-squared:  0.8196, Adjusted R-squared:  0.8195
## F-statistic: 1.794e+04 on 5 and 19743 DF,  p-value: < 2.2e-16
##
##
## $m13
##
## Call:
## lm(formula = PostGFProtestCount ~ population + BlackPop + BlackPov +
##      CollegeEdTotal + VictimsPreGF2020Blk, data = ProtestByCity)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```

## -135.693   -0.150    0.013    0.053  162.674
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -3.280e-02  2.464e-02  -1.331    0.183
## population    -4.889e-05  1.560e-06 -31.344 < 2e-16 ***
## BlackPop       9.205e-04  1.073e-04   8.579 < 2e-16 ***
## BlackPov      -8.141e-04  1.096e-04  -7.429 1.14e-13 ***
## CollegeEdTotal  4.197e-04  5.478e-06  76.613 < 2e-16 ***
## VictimsPreGF2020Blk -5.916e+00  3.247e-01 -18.220 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.378 on 19743 degrees of freedom
## Multiple R-squared:  0.8194, Adjusted R-squared:  0.8194
## F-statistic: 1.792e+04 on 5 and 19743 DF,  p-value: < 2.2e-16
##
##
## $m14
##
## Call:
## lm(formula = PostGFProtestCount ~ population + BlackPopPct +
##     BlackPovRate + EduRate + VictimCountPreGF, data = ProtestByCity)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -106.746   -0.408    0.050    0.280   232.548
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.050e-01  1.383e-01   0.760   0.4474
## population     9.967e-05  5.222e-07 190.884 < 2e-16 ***
## BlackPopPct    3.517e-01  1.927e-01   1.826   0.0679 .
## BlackPovRate   -5.587e-01  1.398e-01  -3.996 6.48e-05 ***
## EduRate        2.907e-01  2.046e-01   1.421   0.1555
## VictimCountPreGF -1.873e+00  4.057e-02 -46.161 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.236 on 15382 degrees of freedom
## (4361 observations deleted due to missingness)
## Multiple R-squared:  0.7783, Adjusted R-squared:  0.7782
## F-statistic: 1.08e+04 on 5 and 15382 DF,  p-value: < 2.2e-16
##
##
## $m15
##
## Call:
## lm(formula = PostGFProtestCount ~ population + BlackPopPct +
##     BlackPovRate + EduRate + BlackVictimCountPreGF, data = ProtestByCity)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -134.189   -0.310    0.139    0.374   237.663

```

```
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.833e-02  1.466e-01   0.261   0.7938
## population      8.732e-05  5.255e-07 166.183 < 2e-16 ***
## BlackPopPct     4.057e-01  2.054e-01   1.976   0.0482 *
## BlackPovRate    -6.061e-01  1.483e-01  -4.088 4.37e-05 ***
## EduRate         4.116e-01  2.169e-01   1.897   0.0578 .
## BlackVictimCountPreGF -1.345e+00  9.667e-02 -13.909 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.491 on 15382 degrees of freedom
## (4361 observations deleted due to missingness)
## Multiple R-squared:  0.7507, Adjusted R-squared:  0.7506
## F-statistic: 9263 on 5 and 15382 DF, p-value: < 2.2e-16
##
##
## $m16
##
## Call:
## lm(formula = PostGFProtestCount ~ population + BlackPopPct +
##     BlackPovRate + EduRate + PoliceDeathsRate, data = ProtestByCity)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -150.895   -0.298    0.160    0.363   237.180
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    6.170e-02  1.476e-01   0.418   0.676
## population      8.234e-05  3.874e-07 212.554 < 2e-16 ***
## BlackPopPct     1.016e-01  2.056e-01   0.494   0.621
## BlackPovRate    -5.937e-01  1.492e-01  -3.980 6.94e-05 ***
## EduRate         4.155e-01  2.183e-01   1.903   0.057 .
## PoliceDeathsRate 1.003e+03  8.435e+02   1.189   0.234
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.519 on 15382 degrees of freedom
## (4361 observations deleted due to missingness)
## Multiple R-squared:  0.7476, Adjusted R-squared:  0.7475
## F-statistic: 9111 on 5 and 15382 DF, p-value: < 2.2e-16
##
##
## $m17
##
## Call:
## lm(formula = PostGFProtestCount ~ population + BlackPopPct +
##     BlackPovRate + EduRate + BlackVictimsRate, data = ProtestByCity)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -150.903   -0.299    0.157    0.359   237.174
```

```
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.974e-02  1.475e-01   0.473   0.6364
## population   8.234e-05  3.874e-07 212.547 < 2e-16 ***
## BlackPopPct   9.907e-02  2.061e-01   0.481   0.6308
## BlackPovRate -5.963e-01  1.492e-01  -3.997 6.45e-05 ***
## EduRate       4.117e-01  2.183e-01   1.886   0.0593 .
## BlackVictimsRate 1.440e+03  2.605e+03   0.553   0.5804
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.519 on 15382 degrees of freedom
## (4361 observations deleted due to missingness)
## Multiple R-squared:  0.7476, Adjusted R-squared:  0.7475
## F-statistic: 9110 on 5 and 15382 DF, p-value: < 2.2e-16
##
##
## $m18
##
## Call:
## lm(formula = PostGFProtestCount ~ population + BlackPopPct +
##     BlackPovRate + EduRate + VictimRaceRatio, data = ProtestByCity)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -153.890   -2.915    0.133    2.707   232.033
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  8.797e-01  2.621e+00   0.336 0.737169
## population   8.377e-05  1.334e-06 62.784 < 2e-16 ***
## BlackPopPct   5.561e+00  2.828e+00   1.967 0.049446 *
## BlackPovRate -7.093e+00  2.795e+00  -2.538 0.011278 *
## EduRate       1.142e+01  2.963e+00   3.852 0.000123 ***
## VictimRaceRatio 8.451e-01  1.246e+00   0.678 0.497677
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.73 on 1209 degrees of freedom
## (18534 observations deleted due to missingness)
## Multiple R-squared:  0.7715, Adjusted R-squared:  0.7706
## F-statistic: 816.5 on 5 and 1209 DF, p-value: < 2.2e-16
##
##
## $m19
##
## Call:
## lm(formula = PostGFProtestCount ~ population + BlackPopPct +
##     BlackPovRate + EduRate + AntiTrumpProtestPreGFCount + VictimRaceRatio,
##     data = ProtestByCity)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```



```

## -114.353   -1.387    1.219    2.729  206.838
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.325e+00  2.151e+00   1.081   0.2800
## population        4.592e-05  1.907e-06  24.078 <2e-16 ***
## BlackPopPct       1.298e+00  2.327e+00   0.558   0.5771
## BlackPovRate      -4.961e+00  2.295e+00  -2.161   0.0309 *
## EduRate           -1.341e+00  2.488e+00  -0.539   0.5900
## AntiTrumpProtestPreGFCount 3.570e-01  1.473e-02  24.231 <2e-16 ***
## VictimRaceRatio   -8.907e-01  1.025e+00  -0.869   0.3850
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.09 on 1208 degrees of freedom
## (18534 observations deleted due to missingness)
## Multiple R-squared:  0.8463, Adjusted R-squared:  0.8455
## F-statistic: 1108 on 6 and 1208 DF,  p-value: < 2.2e-16

```
