607 Final Project - Cause of Death by Firearm vs State Firearm Laws

Jen Abinette

2022-11-20

Research question

Are there aspects of gun policy that are predictors of firearm mortality rate?

Cases

All datasets and relevant information can be found in Github Repository - https://github.com/JAbinette/CUNY-607-Final-Project

Cause of Death by Firearm data retrieved from Centers for Disease Control & Prevention website includes 458 observations grouped by State, Year and Cause of Death Category (ICD Sub-Chapter) including firearm deaths excluding Terrorism, Legal Intervention and Operations of War. Our analysis will summarize by State (50 cases) for 2020. See Github File: '0 - wonder.cdc.gov Underlying Cause of Death cdc - Grp by State, ICD Sub-Chapter.txt'

Firearm Laws by State data retrieved from https://www.statefirearmlaws.org/resources Database containing detailed annual information on firearm-related laws in place. Our analysis will include 50 cases with 137 variables after subsetting data to only include 2020 US States. See Github File: '0 - statefirearmlaws.org Firearm Laws - DATABASE_0.xlsx'

Data Preparation

CDC 2020 Summary of Death by Firearm

```
library(tidyr)
library(readxl)
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.2.2
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union


# Load CDC Cause of Death by Firearm text file from Github
txt.url = 'https://raw.githubusercontent.com/JAbinette/CUNY-607-Final-Project/main/0%20-%20wonder.cdc.g
cod.ld <- read.delim (txt.url, header=TRUE, sep = "\t")


cod.st_yr <- subset (cod.ld, Year.Code == 2020, select = c("State", "Year.Code", "Deaths", "Population")
  group_by(State, Year.Code) %>%
  mutate(Firearm_Deaths = sum(Deaths)) %>%
  mutate (Prop.of.Pop = (Firearm_Deaths/Population) ) %>%
  ungroup()
cod.st_yr <- subset (cod.st_yr, State != "District of Columbia", select =c("State", "Year.Code", "Firea
  distinct() %>%
  na.omit()
```

CDC Reported Deaths by State in 2020

```
# Load CDC Deaths by State in 2020 text file from Github
txt.url = 'https://raw.githubusercontent.com/JAbinette/CUNY-607-Final-Project/main/0%20-%20wonder.cdc.g
cod.all <- read.delim (txt.url, header=TRUE, sep = "\t")
cod.all <- subset(cod.all, Year.Code == 2020 & State != "District of Columbia", select = c("State", "De

# Merge with Firearm Deaths data
COD <- merge( cod.st_yr, cod.all, by.x = "State", by.y = "State", all.x = TRUE)  %>%
  mutate (Prop.of.Deaths = (Firearm_Deaths/Deaths) )
```

State Firearm Laws (filtered to only include 2020)

```
# Download State Firearm Laws spreadsheet and save to your file directory (data saved in Github at http
# Set path to excel spreadsheet
path = "0 - statefirearmlaws.org Firearm Laws - DATABASE_0.xlsx"
laws_ld <- read_excel(path)


# Subset data to only include year 2020
laws_ld <- subset (laws_ld, year == 2020)
# Transform from wide to long (excluding last column which is the total number of laws)
laws_ld2 <- pivot_longer ( laws_ld, cols = 3:136, names_to ="Variable", values_to = "Indicator" )
# Remove variable records where the law was not present for that year and State
laws_ld3 <- subset ( laws_ld2, Indicator == 1 )
head(laws_ld3)


## # A tibble: 6 x 5
##   state     year lawtotal Variable        Indicator
##   <chr>    <dbl>    <dbl> <chr>               <dbl>
## 1 Alabama   2020       10 invcommitment           1
## 2 Alabama   2020       10 danger                  1
## 3 Alabama   2020       10 alcoholism              1
```

```
## 4 Alabama  2020      10 dealerh                  1
## 5 Alabama  2020      10 permitconcealed          1
## 6 Alabama  2020      10 ccbackground             1
```

Add Variable Categories to Merge with State Firearm

```r
# Load Variable Category data for State Firearm Laws data
laws.cat_ld <- read.csv('https://raw.githubusercontent.com/JAbinette/CUNY-607-Final-Project/main/0%20-%
str(laws.cat_ld)
```

```
## 'data.frame':    134 obs. of  4 variables:
##  $ Category.Code: int  1 1 1 1 1 1 1 1 1 1 ...
##  $ Category     : chr  "Dealer regulations" "Dealer regulations" "Dealer regulations" "Dealer regula
##  $ Sub.Category : chr  "Licensing" "Licensing" "Recordkeeping" "Recordkeeping" ...
##  $ Variable.Name: chr  "dealer" "dealerh" "recordsdealer" "recordsdealerh" ...
```

```r
# Replace spaces in Category with a period
laws.cat_ld$Category <- make.names (laws.cat_ld$Category)
# Merge
laws_ld4 <- merge( laws_ld3, laws.cat_ld, by.x = "Variable", by.y = "Variable.Name", all.x = TRUE)
laws_ld4 <- subset ( laws_ld4, select = c("state","year","lawtotal","Category"))

library(dplyr)
# Add count of laws by Category
laws_ld4 <- laws_ld4 %>%
  group_by(state, year, lawtotal, Category) %>%
  summarise(Category_count=n(),
            .groups = 'drop')
head(laws_ld4)
```

```
## # A tibble: 6 x 5
##   state   year lawtotal Category                             Category_co~1
##   <chr>  <dbl>    <dbl> <chr>                                        <int>
## 1 Alabama 2020       10 Concealed.carry.permitting                      4
## 2 Alabama 2020       10 Dealer.regulations                              1
## 3 Alabama 2020       10 Domestic.violence                               2
## 4 Alabama 2020       10 Prohibitions.for.high.risk.gun.possession       3
## 5 Alaska  2020        3 Buyer.regulations                               1
## 6 Alaska  2020        3 Possession.regulations                          1
## # ... with abbreviated variable name 1: Category_count
```

```r
# Transform long to wide
laws <- pivot_wider( laws_ld4, names_from = "Category", values_from = "Category_count")
# Convert NA to zero
laws[is.na(laws)] <- 0
str(laws)
```

```
## tibble [50 x 17] (S3: tbl_df/tbl/data.frame)
##  $ state    : chr [1:50] "Alabama" "Alaska" "Arizona" "Arkansas"
##  $ year     : num [1:50] 2020 2020 2020 2020 2020 2020 2020 2020 2
##  $ lawtotal : num [1:50] 10 3 8 11 111 32 89 41 30 6 ...
```

```
##  $ Concealed.carry.permitting              : int [1:50] 4 0 0 5 6 5 4 4 3 4 ...
##  $ Dealer.regulations                      : int [1:50] 1 0 0 0 14 2 12 5 2 0 ...
##  $ Domestic.violence                       : int [1:50] 2 0 3 0 18 7 16 5 1 0 ...
##  $ Prohibitions.for.high.risk.gun.possession  : int [1:50] 3 1 4 3 8 1 9 5 2 1 ...
##  $ Buyer.regulations                       : int [1:50] 0 1 0 0 14 0 8 2 6 0 ...
##  $ Possession.regulations                  : int [1:50] 0 1 1 2 8 4 5 3 9 1 ...
##  $ Stand.your.ground                       : int [1:50] 0 0 0 1 1 1 1 1 0 0 ...
##  $ Ammunition.regulations                  : int [1:50] 0 0 0 0 5 0 5 1 1 0 ...
##  $ Assault.weapons.and.large.capacity.magazines: int [1:50] 0 0 0 0 8 1 7 0 0 0 ...
##  $ Background.checks                       : int [1:50] 0 0 0 0 11 9 10 9 1 0 ...
##  $ Child.access.prevention                 : int [1:50] 0 0 0 0 9 0 6 4 3 0 ...
##  $ Gun.trafficking                         : int [1:50] 0 0 0 0 6 2 2 2 2 0 ...
##  $ Immunity                                : int [1:50] 0 0 0 0 1 0 1 0 0 0 ...
##  $ Preemption                              : int [1:50] 0 0 0 0 2 0 3 0 0 0 ...
```

```
# Merge State Laws data with Cause of Death by State and Year
df <- merge (COD, laws, by.x = c("State"), by.y = c("state"), all.x = TRUE)
df <- subset(df, select = c("State", "Prop.of.Pop", "Prop.of.Deaths", "lawtotal"))
```

Statistical Analyses

```
summary(df)
```

```
##      State             Prop.of.Pop          Prop.of.Deaths        lawtotal
##  Length:50          Min.   :0.0000334    Min.   :0.00387     Min.   :  1.00
##  Class :character   1st Qu.:0.0001093    1st Qu.:0.01078     1st Qu.:  9.25
##  Mode  :character   Median :0.0001390    Median :0.01360     Median : 20.50
##                     Mean   :0.0001473    Mean   :0.01390     Mean   : 29.44
##                     3rd Qu.:0.0001920    3rd Qu.:0.01736     3rd Qu.: 40.50
##                     Max.   :0.0002744    Max.   :0.03095     Max.   :111.00
```

```
sd(df$Prop.of.Pop)
```

```
## [1] 5.909196e-05
```

```
IQR(df$Prop.of.Pop)
```

```
## [1] 8.262369e-05
```

```
sd(df$lawtotal)
```
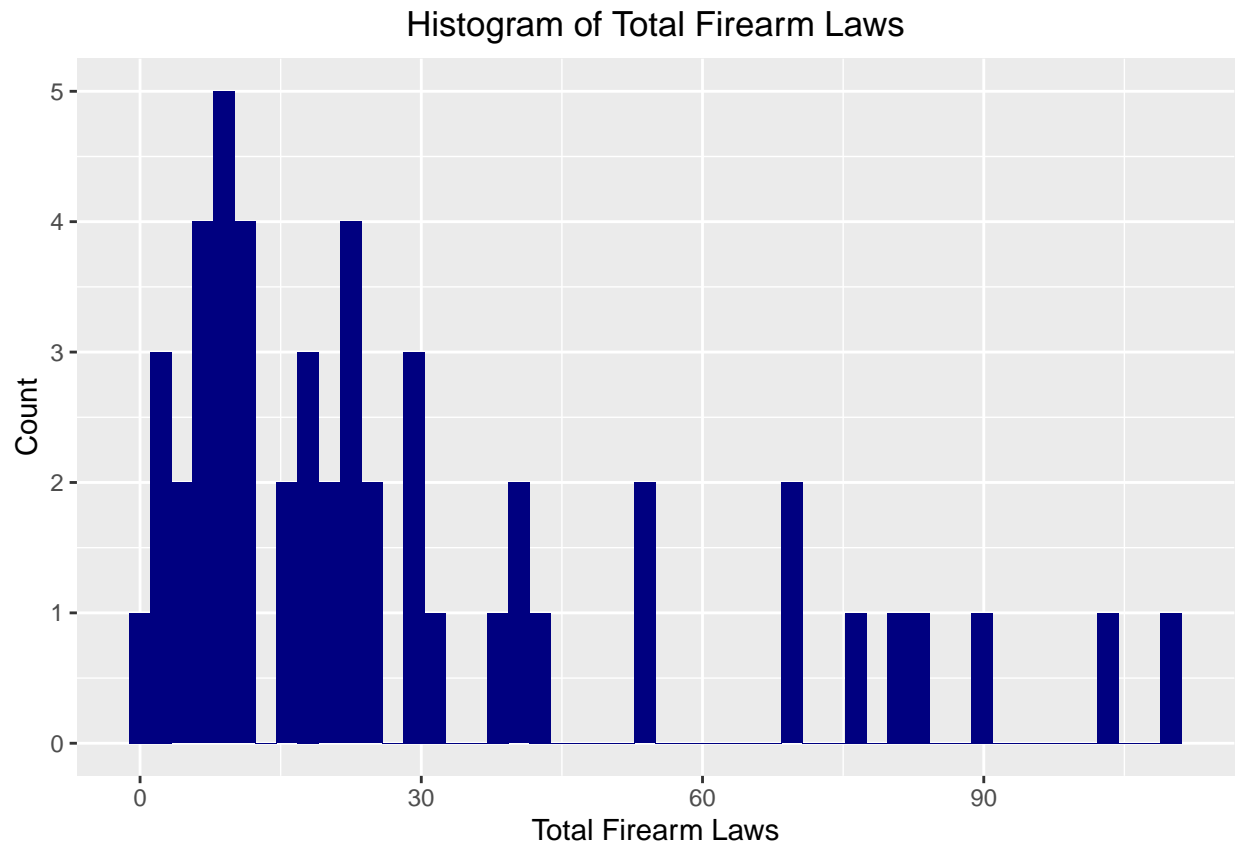
```
## [1] 28.24251
```
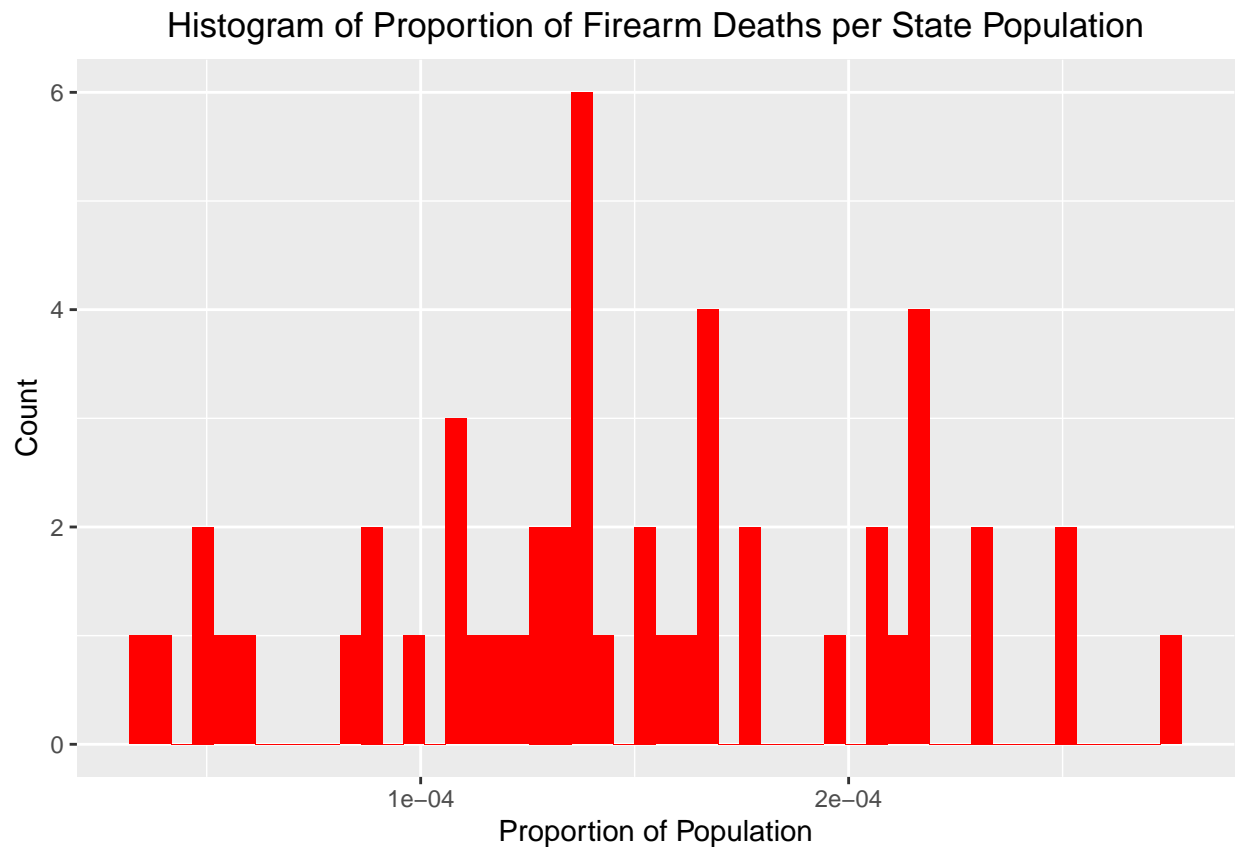
```
IQR(df$lawtotal)
```

```
## [1] 31.25
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.2.2
```

```
ggplot(data = df, aes(x=lawtotal)) + geom_histogram(fill="navy", bins = 50) + labs(x="Total Firearm Laws
```


Histogram of Total Firearm Laws

```
ggplot(data = df, aes(x=Prop.of.Pop)) + geom_histogram(fill="red", bins = 50) + labs(x="Proportion of Po
```

Histogram of Proportion of Firearm Deaths per State Population

Can the Total Number of Laws predict proportion of Firearm deaths per State Population?

```
df %>%
  summarise( cor( lawtotal, Prop.of.Pop, use = "complete.obs"))
```

```
##   cor(lawtotal, Prop.of.Pop, use = "complete.obs")
## 1                                      -0.7024303
```

```
library(ggplot2)
ggplot(data = df, aes(x = lawtotal, y = Prop.of.Pop)) +
geom_jitter(color='violetred4') + labs(x="Total Firearm Laws", y="Proportion Firearm Deaths per Populat:
```

Scatterplot of Total Laws by Proportion of Deaths per Population
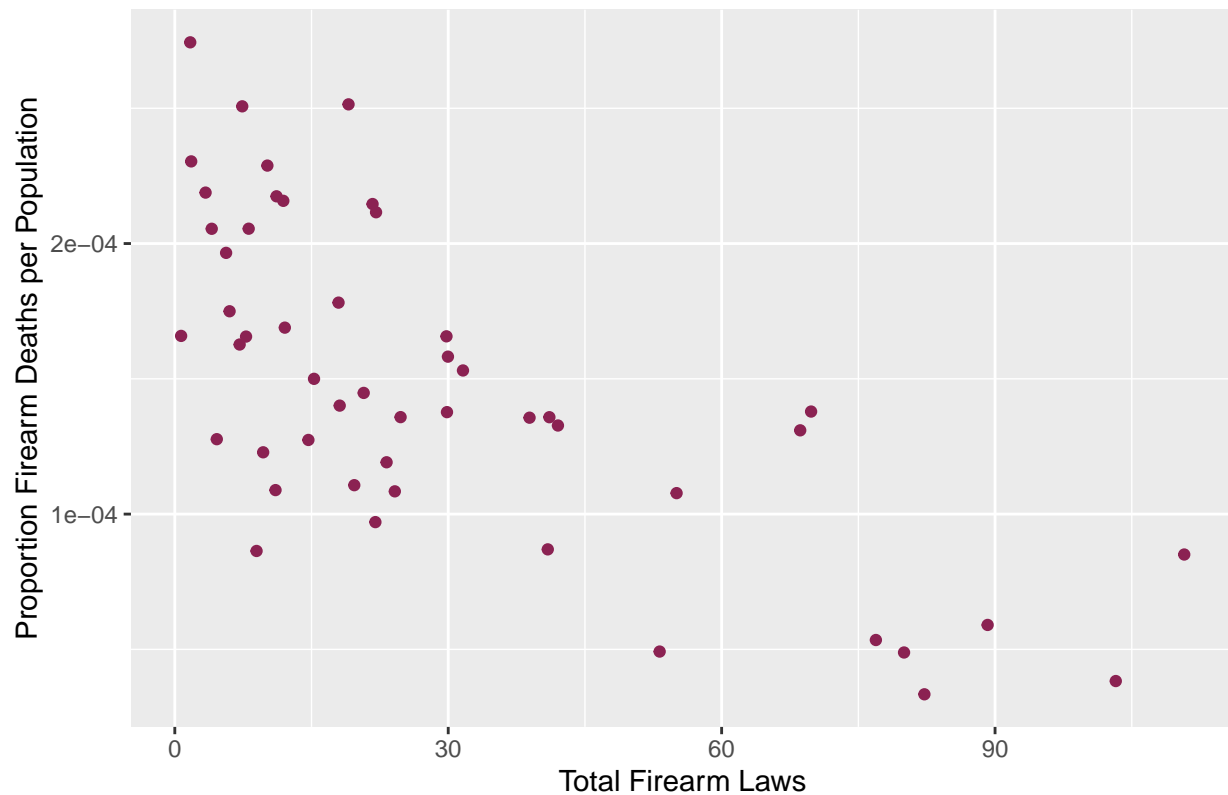

```r
# Simple Linear Regression
m.pop <- lm(Prop.of.Pop ~ lawtotal, data = df)
summary(m.pop)
```

```
##
## Call:
## lm(formula = Prop.of.Pop ~ lawtotal, data = df)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -9.100e-05 -2.413e-05 -1.467e-06  3.114e-05  8.876e-05
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.906e-04  8.727e-06   21.840  < 2e-16 ***
## lawtotal    -1.470e-06  2.149e-07   -6.837 1.31e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.249e-05 on 48 degrees of freedom
## Multiple R-squared:  0.4934, Adjusted R-squared:  0.4829
## F-statistic: 46.75 on 1 and 48 DF,  p-value: 1.305e-08
```

```r
ggplot(data = df, aes(x = lawtotal, y = Prop.of.Pop)) +
geom_jitter(color='violetred4') + geom_smooth(method = "lm") + labs(x="Total Firearm Laws", y="Proportio
```

## `geom_smooth()` using formula = 'y ~ x'

Linear Regression Model
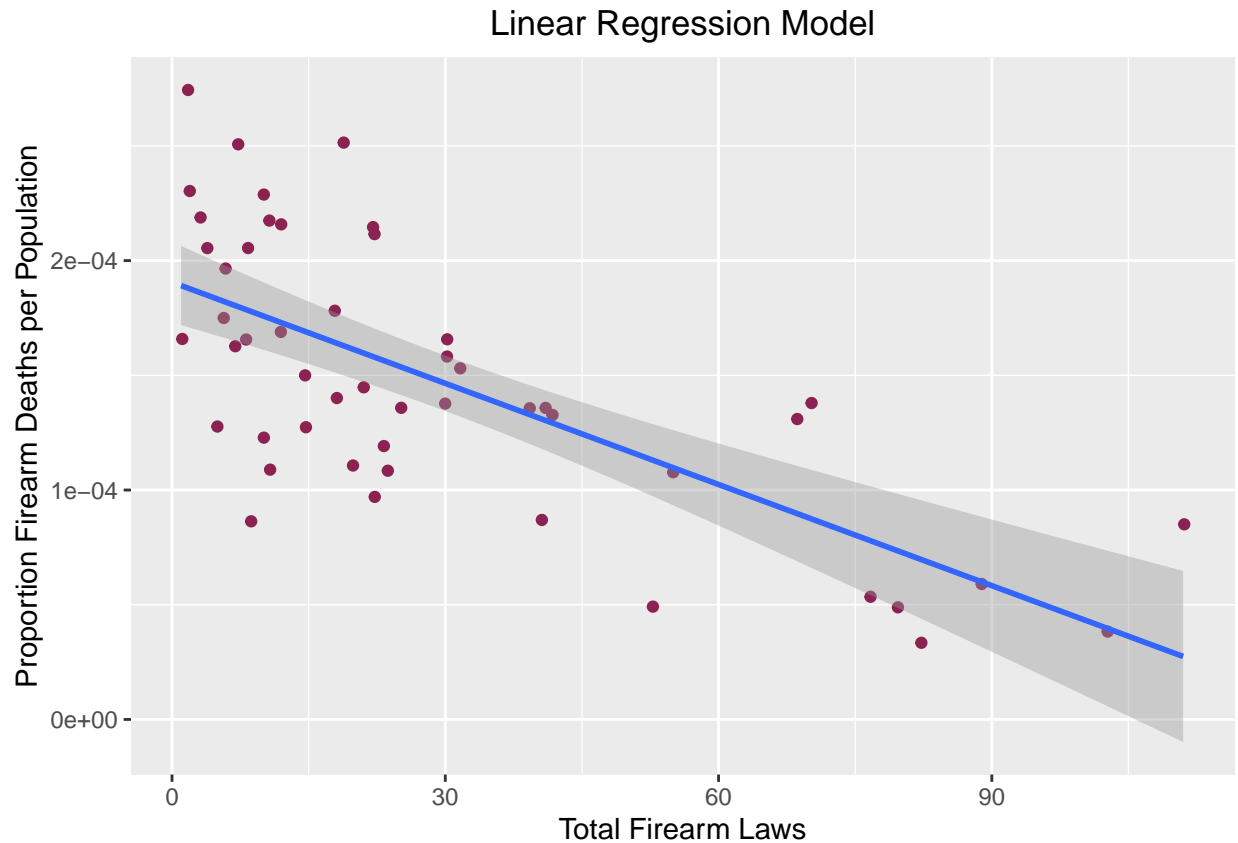
Check Model Assumptions for Simple Regression

Linearity and Constant Variability - Conditions both met as there is no apparent pattern in the residuals plot indicating there is linearity and the points are scattered around zero showing constant variability.
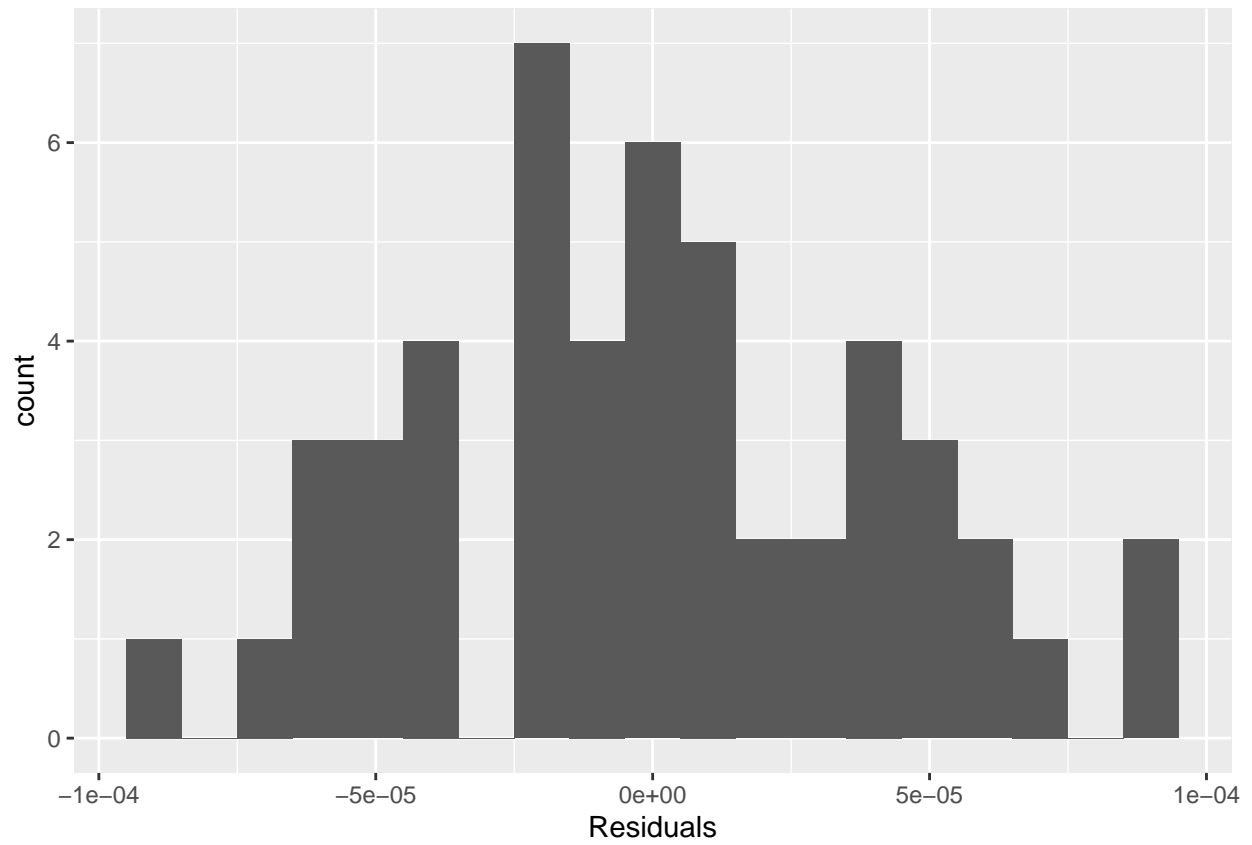
```
ggplot(data = m.pop, aes(x = .fitted, y = .resid)) +
geom_point() +
geom_hline(yintercept = 0, linetype = "dashed") +
xlab("Fitted values") +
ylab("Residuals")
```
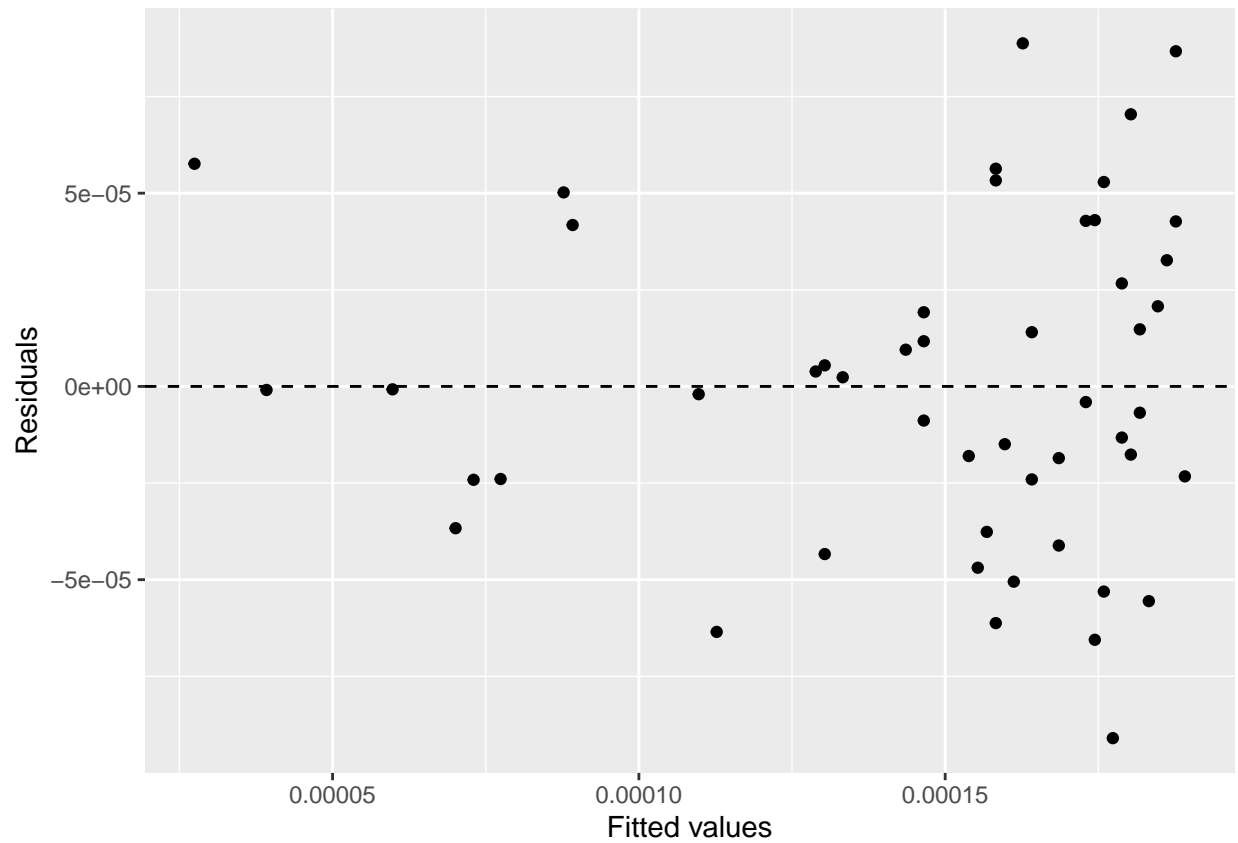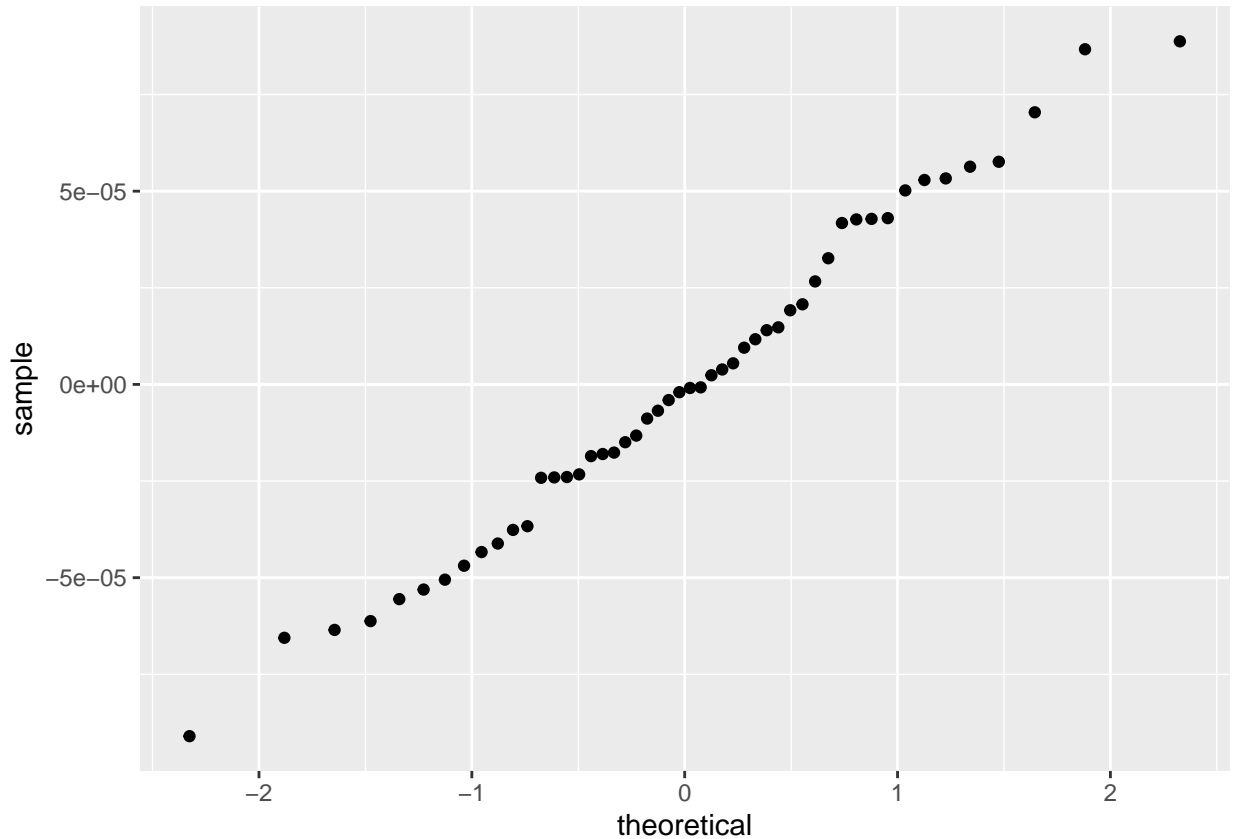
Nearly Normal Residuals - Condition is met based on below histogram and the normal probability plots.

```
ggplot(data = m.pop, aes(x = .resid)) +
geom_histogram(binwidth = .00001) +
xlab("Residuals")
```

```
ggplot(data = m.pop, aes(sample = .resid)) + stat_qq()
```

Can the Number and Category of Firearm Laws predict proportion of Firearm deaths per State Population?

```r
# Merge State Laws data with Cause of Death by State and Year
df2 <- merge (COD, laws, by.x = c("State"), by.y = c("state"), all.x = TRUE)
summary(df2)
```

```
##     State              Year.Code      Firearm_Deaths    Prop.of.Pop
##  Length:50          Min.   :2020    Min.   :  47.0    Min.   :0.0000334
##  Class :character   1st Qu.:2020    1st Qu.: 232.5    1st Qu.:0.0001093
##  Mode  :character   Median :2020    Median : 677.0    Median :0.0001390
##                     Mean   :2020    Mean   : 883.8    Mean   :0.0001473
##                     3rd Qu.:2020    3rd Qu.:1167.8    3rd Qu.:0.0001920
##                     Max.   :2020    Max.   :4113.0    Max.   :0.0002744
##      Deaths         Prop.of.Deaths          year          lawtotal
##  Min.   :  5170    Min.   :0.00387    Min.   :2020    Min.   :  1.00
##  1st Qu.: 20036    1st Qu.:0.01078    1st Qu.:2020    1st Qu.:  9.25
##  Median : 50091    Median :0.01360    Median :2020    Median : 20.50
##  Mean   : 67551    Mean   :0.01390    Mean   :2020    Mean   : 29.44
##  3rd Qu.: 79996    3rd Qu.:0.01736    3rd Qu.:2020    3rd Qu.: 40.50
##  Max.   :319808    Max.   :0.03095    Max.   :2020    Max.   :111.00
##  Concealed.carry.permitting Dealer.regulations Domestic.violence
##  Min.   :0.00               Min.   : 0.00      Min.   : 0.00
##  1st Qu.:2.25               1st Qu.: 0.00      1st Qu.: 1.00
```

```
##   Median :4.00              Median : 0.50        Median : 3.50
##   Mean   :3.38              Mean    : 2.78        Mean    : 5.64
##   3rd Qu.:5.00              3rd Qu.: 5.00         3rd Qu.: 8.50
##   Max.   :7.00              Max.    :14.00        Max.    :18.00
##   Prohibitions.for.high.risk.gun.possession Buyer.regulations
##   Min.   : 0.00                             Min.    : 0.00
##   1st Qu.: 1.00                             1st Qu.: 0.00
##   Median : 3.00                             Median : 1.00
##   Mean   : 3.26                             Mean    : 2.76
##   3rd Qu.: 5.00                             3rd Qu.: 3.75
##   Max.   :10.00                             Max.    :16.00
##   Possession.regulations Stand.your.ground Ammunition.regulations
##   Min.   : 0.00          Min.   :0.00      Min.    :0.00
##   1st Qu.: 1.00          1st Qu.:0.00      1st Qu.:0.00
##   Median : 2.50          Median :0.00      Median :0.00
##   Mean   : 3.02          Mean   :0.44      Mean    :0.72
##   3rd Qu.: 4.00          3rd Qu.:1.00      3rd Qu.:1.00
##   Max.   :11.00          Max.   :1.00      Max.    :6.00
##   Assault.weapons.and.large.capacity.magazines Background.checks
##   Min.   :0.0                                  Min.    : 0.0
##   1st Qu.:0.0                                  1st Qu.: 0.0
##   Median :0.0                                  Median : 0.0
##   Mean   :0.8                                  Mean    : 3.2
##   3rd Qu.:0.0                                  3rd Qu.: 7.0
##   Max.   :8.0                                  Max.    :11.0
##   Child.access.prevention Gun.trafficking    Immunity       Preemption
##   Min.   : 0.00           Min.   :0.00     Min.   :0.00    Min.    :0.0
##   1st Qu.: 0.00           1st Qu.:0.00     1st Qu.:0.00    1st Qu.:0.0
##   Median : 0.00           Median :0.00     Median :0.00    Median :0.0
##   Mean   : 1.90           Mean   :0.78     Mean   :0.34    Mean    :0.4
##   3rd Qu.: 3.75           3rd Qu.:2.00     3rd Qu.:1.00    3rd Qu.:0.0
##   Max.   :11.00           Max.   :6.00     Max.   :1.00    Max.    :3.0
```

```r
# Multiple Linear Regression
m_2020 <- lm(Prop.of.Pop ~ Concealed.carry.permitting + Dealer.regulations + Domestic.violence +
Prohibitions.for.high.risk.gun.possession + Buyer.regulations + Possession.regulations + Stand.your.grou
Background.checks + Child.access.prevention + Gun.trafficking + Immunity + Preemption, data = df2)
summary(m_2020)
```

```
##
## Call:
## lm(formula = Prop.of.Pop ~ Concealed.carry.permitting + Dealer.regulations +
##     Domestic.violence + Prohibitions.for.high.risk.gun.possession +
##     Buyer.regulations + Possession.regulations + Stand.your.ground +
##     Ammunition.regulations + Assault.weapons.and.large.capacity.magazines +
##     Background.checks + Child.access.prevention + Gun.trafficking +
##     Immunity + Preemption, data = df2)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -6.912e-05 -1.866e-05 -4.747e-06  2.022e-05  9.049e-05
##
## Coefficients:
##                                               Estimate Std. Error t value
```

```
## (Intercept)                                     1.749e-04  1.387e-05  12.612
## Concealed.carry.permitting                       1.298e-06  3.768e-06   0.344
## Dealer.regulations                              -5.773e-07  3.243e-06  -0.178
## Domestic.violence                               -1.357e-06  2.295e-06  -0.591
## Prohibitions.for.high.risk.gun.possession        1.717e-06  3.894e-06   0.441
## Buyer.regulations                               -4.398e-06  3.330e-06  -1.321
## Possession.regulations                           7.210e-06  4.265e-06   1.690
## Stand.your.ground                               -3.039e-05  2.141e-05  -1.420
## Ammunition.regulations                           1.221e-05  1.024e-05   1.192
## Assault.weapons.and.large.capacity.magazines     3.794e-06  5.031e-06   0.754
## Background.checks                               -7.961e-07  2.845e-06  -0.280
## Child.access.prevention                         -8.916e-06  3.997e-06  -2.231
## Gun.trafficking                                 -8.205e-06  7.430e-06  -1.104
## Immunity                                        -1.213e-06  2.057e-05  -0.059
## Preemption                                      -2.533e-05  1.217e-05  -2.082
##                                                 Pr(>|t|)
## (Intercept)                                     1.41e-14 ***
## Concealed.carry.permitting                        0.7326
## Dealer.regulations                                0.8597
## Domestic.violence                                 0.5582
## Prohibitions.for.high.risk.gun.possession         0.6620
## Buyer.regulations                                 0.1951
## Possession.regulations                            0.0998 .
## Stand.your.ground                                 0.1645
## Ammunition.regulations                            0.2414
## Assault.weapons.and.large.capacity.magazines      0.4558
## Background.checks                                 0.7813
## Child.access.prevention                           0.0322 *
## Gun.trafficking                                   0.2770
## Immunity                                          0.9533
## Preemption                                        0.0447 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.141e-05 on 35 degrees of freedom
## Multiple R-squared:  0.6493, Adjusted R-squared:  0.509
## F-statistic: 4.628 on 14 and 35 DF,  p-value: 0.0001146
```

**Fit the Best Model using backward-selection and p-value as the selection criterion**

```
# Variables removed highest p-value at a time: Immunity, Dealer.regulations, Concealed.carry.permitting
m_bestfit <- lm(Prop.of.Pop ~ Stand.your.ground + Child.access.prevention + Preemption, data = df2)
summary(m_bestfit)
```

```
##
## Call:
## lm(formula = Prop.of.Pop ~ Stand.your.ground + Child.access.prevention +
##     Preemption, data = df2)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -7.730e-05 -2.465e-05 -7.675e-06  2.893e-05  9.012e-05
```

```
## 
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)            1.842e-04  7.847e-06  23.480  < 2e-16 ***
## Stand.your.ground     -3.813e-05  1.314e-05  -2.902  0.00567 **
## Child.access.prevention -6.860e-06  2.492e-06  -2.753  0.00843 **
## Preemption            -1.777e-05  7.425e-06  -2.393  0.02083 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 3.982e-05 on 46 degrees of freedom
## Multiple R-squared:  0.5737, Adjusted R-squared:  0.5459
## F-statistic: 20.64 on 3 and 46 DF,  p-value: 1.284e-08
```

Check Model Assumptions

```
par(mfrow = c(2, 2))
plot(m_bestfit)
```
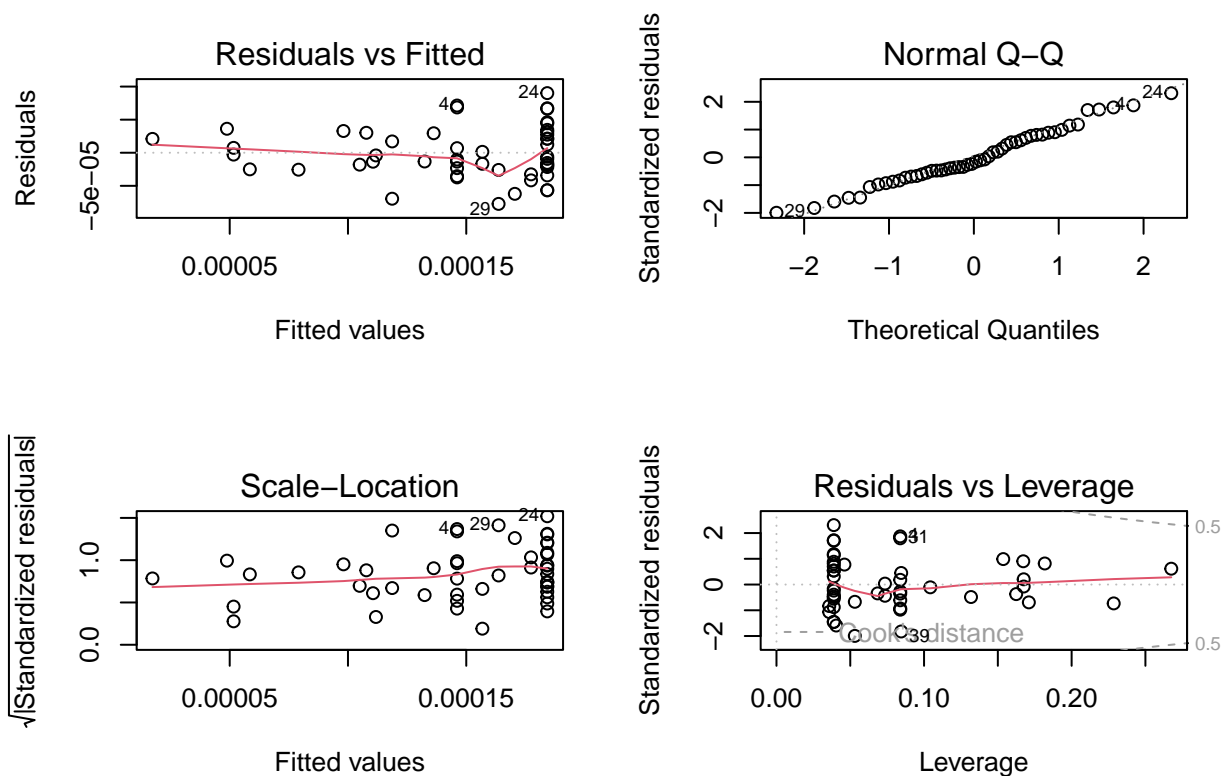

Conclusions

The total number of firearm laws is a significant predictor and accounts for 49% of the variability in Firearm Deaths as a Proportion of the State Population. In breaking down the firearm laws by category, a multi-

ple regression analysis determined that the number of Stand Your Ground, Child Access Prevention, and Preemption firearm laws can explain 55% of the variability.