Covid_Infection_Modeling_R

Mitchell Nelson

2023-08-10

Packages

```
library(tidyverse)
```

```
## Warning: package 'tibble' was built under R version 4.2.3
```

```
## Warning: package 'dplyr' was built under R version 4.2.3
```

```
## Warning: package 'stringr' was built under R version 4.2.3
```

```
## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## v dplyr     1.1.1     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.1     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.1
## -- Conflicts ----------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(ggplot2)
library(faraway)
library(leaps)
```

```
## Warning: package 'leaps' was built under R version 4.2.3
```

```
library(car)
```

```
## Warning: package 'car' was built under R version 4.2.3
```

```
## Loading required package: carData
##
## Attaching package: 'car'
##
## The following objects are masked from 'package:faraway':
##
```

```
##      logit, vif
##
## The following object is masked from 'package:dplyr':
##
##      recode
##
## The following object is masked from 'package:purrr':
##
##      some
```

```
library(nlme)
```

```
##
## Attaching package: 'nlme'
##
## The following object is masked from 'package:dplyr':
##
##      collapse
```

```
library(corrplot)
```

```
## Warning: package 'corrplot' was built under R version 4.2.3
```

```
## corrplot 0.92 loaded
```

Brief Data Cleaning and Updating with More Currents Values (according to the time of this project's completion)

```
covid <- read.csv("COVID19_state.csv", sep = ",")

#Removed date of school closure
covid <- covid[, -26]

#Removed DC
covid <- covid[-8, ]

#Updating counts
DeathsCurr <- c(1209, 19541, 11366, 29852, 89947, 12273,10826, 2896, 73789, 36217, 1404, 9510,
                4921, 37849, 23535,8524, 15322, 17220, 20219, 14422, 2280, 35935, 12752, 20192,
                12423, 3262, 23373, 2304, 4195, 2472, 33371, 7438, 10684, 67548, 38360, 14219,
                7478, 44578, 3532, 17734, 2901, 25846, 88146, 4740, 20124, 629, 12703, 14402,
                6823, 1807)

InfCurr <- c(250000,1300000, 835000, 2000000, 9180000, 1380000, 751000, 261000, 5900000,
             2440000, 239000, 761000, 445000,3120000,1700000, 773000, 1330000, 1240000,
             1740000,1030000,  242000, 2410000, 1440000, 1430000, 797000, 274000, 2650000,
             241000, 497000,307000 , 2240000, 521000, 717000, 5090000, 2690000, 1040000,
             714000, 2800000,368000, 1470000, 238000 , 1980000,6730000, 931000,
             1690000,122000, 1480000, 1600000 , 500000, 156000  )
```

```
covid["DeathsCurr"] <- DeathsCurr
covid["InfCurr"] <- InfCurr

#Removing old counts
covid <- covid[, -2]
covid <- covid[, -2]
covid <- covid[, -2]
```

Creating percentages for deaths and infections

```
DeathInf <- c()
InfPop <- c()
DeathPop = c()

for(i in 1:50)
{
InfPop[i] <- covid$InfCurr[i]/covid$Population[i] *100
DeathInf[i] <- covid$DeathsCurr[i]/ covid$InfCurr[i] *100
DeathPop[i] <- covid$DeathsCurr[i]/covid$Population[i] *100
}

covid["InfPop"] <- InfPop
covid["DeathInf"] <- DeathInf
covid["DeathPop"] <- DeathPop
```

Exploratory Data Analysis and Data Visualizations

Summary Statistics

```
summary(covid)
```

```
##     State              Population         Pop.Density            Gini
## Length:50          Min.   :  567025   Min.   :   1.286   Min.   :0.4063
## Class :character   1st Qu.: 1857760   1st Qu.:  47.707   1st Qu.:0.4520
## Mode  :character   Median : 4572438   Median : 107.784   Median :0.4673
##                    Mean   : 6611966   Mean   : 203.901   Mean   :0.4646
##                    3rd Qu.: 7692445   3rd Qu.: 219.566   3rd Qu.:0.4789
##                    Max.   :39937489   Max.   :1215.199   Max.   :0.5229
##     ICU.Beds         Income           GDP          Unemployment
## Min.   :  94    Min.   :37994    Min.   :37948    Min.   :2.200
## 1st Qu.: 365    1st Qu.:45908    1st Qu.:50885    1st Qu.:2.825
## Median :1146    Median :49663    Median :57224    Median :3.400
## Mean   :1489    Mean   :51684    Mean   :58722    Mean   :3.482
## 3rd Qu.:1851    3rd Qu.:56728    3rd Qu.:64530    3rd Qu.:3.800
## Max.   :7338    Max.   :74561    Max.   :85746    Max.   :5.800
##    Sex.Ratio       Smoking.Rate     Flu.Deaths    Respiratory.Deaths
## Min.   :0.9231   Min.   : 8.90   Min.   : 9.60   Min.   :19.60
```

```
##    1st Qu.:0.9430    1st Qu.:15.03    1st Qu.:13.10    1st Qu.:34.95
##    Median :0.9596    Median :17.15    Median :15.05    Median :42.80
##    Mean   :0.9652    Mean   :17.33    Mean   :15.32    Mean   :42.79
##    3rd Qu.:0.9831    3rd Qu.:19.30    3rd Qu.:17.05    3rd Qu.:48.67
##    Max.   :1.0547    Max.   :26.00    Max.   :26.10    Max.   :64.30
##      Physicians        Hospitals      Health.Spending    Pollution
##    Min.   :  1172    Min.   :  7.00   Min.   : 5982    Min.   : 4.400
##    1st Qu.:  5642    1st Qu.: 46.25   1st Qu.: 7381    1st Qu.: 6.625
##    Median : 12526    Median : 90.50   Median : 8092    Median : 7.400
##    Mean   : 19959    Mean   :103.76   Mean   : 8260    Mean   : 7.366
##    3rd Qu.: 24334    3rd Qu.:130.75   3rd Qu.: 8918    3rd Qu.: 8.100
##    Max.   :112906    Max.   :523.00   Max.   :11064    Max.   :12.800
##    Med.Large.Airports Temperature      Urban           Age.0.25
##    Min.   :0.00       Min.   :26.60    Min.   :38.70    Min.   :0.260
##    1st Qu.:0.00       1st Qu.:45.25    1st Qu.:65.10    1st Qu.:0.310
##    Median :1.00       Median :51.20    Median :73.75    Median :0.325
##    Mean   :1.24       Mean   :51.95    Mean   :73.59    Mean   :0.324
##    3rd Qu.:1.00       3rd Qu.:58.65    3rd Qu.:86.95    3rd Qu.:0.340
##    Max.   :9.00       Max.   :70.70    Max.   :95.00    Max.   :0.420
##     Age.26.54         Age.55.          DeathsCurr       InfCurr
##    Min.   :0.3500    Min.   :0.2100    Min.   :  629    Min.   : 122000
##    1st Qu.:0.3700    1st Qu.:0.2900    1st Qu.: 4785    1st Qu.: 497750
##    Median :0.3700    Median :0.3000    Median :12728    Median :1140000
##    Mean   :0.3744    Mean   :0.3006    Mean   :19662    Mean   :1600800
##    3rd Qu.:0.3800    3rd Qu.:0.3100    3rd Qu.:23495    3rd Qu.:1920000
##    Max.   :0.4000    Max.   :0.3700    Max.   :89947    Max.   :9180000
##      InfPop           DeathInf          DeathPop
##    Min.   :16.60    Min.   :0.4836    Min.   :0.09939
##    1st Qu.:22.85    1st Qu.:0.9570    1st Qu.:0.22724
##    Median :25.13    Median :1.2098    Median :0.30316
##    Mean   :24.85    Mean   :1.1675    Mean   :0.28978
##    3rd Qu.:26.79    3rd Qu.:1.3973    3rd Qu.:0.34764
##    Max.   :34.84    Max.   :1.5921    Max.   :0.41559
```

Correlation Plots

```r
covidcor <- covid[,-1]
full_covid_cor <- cor(covidcor)
corrplot(full_covid_cor, tl.col = "black", tl.cex = .75, order = "hclust")
```
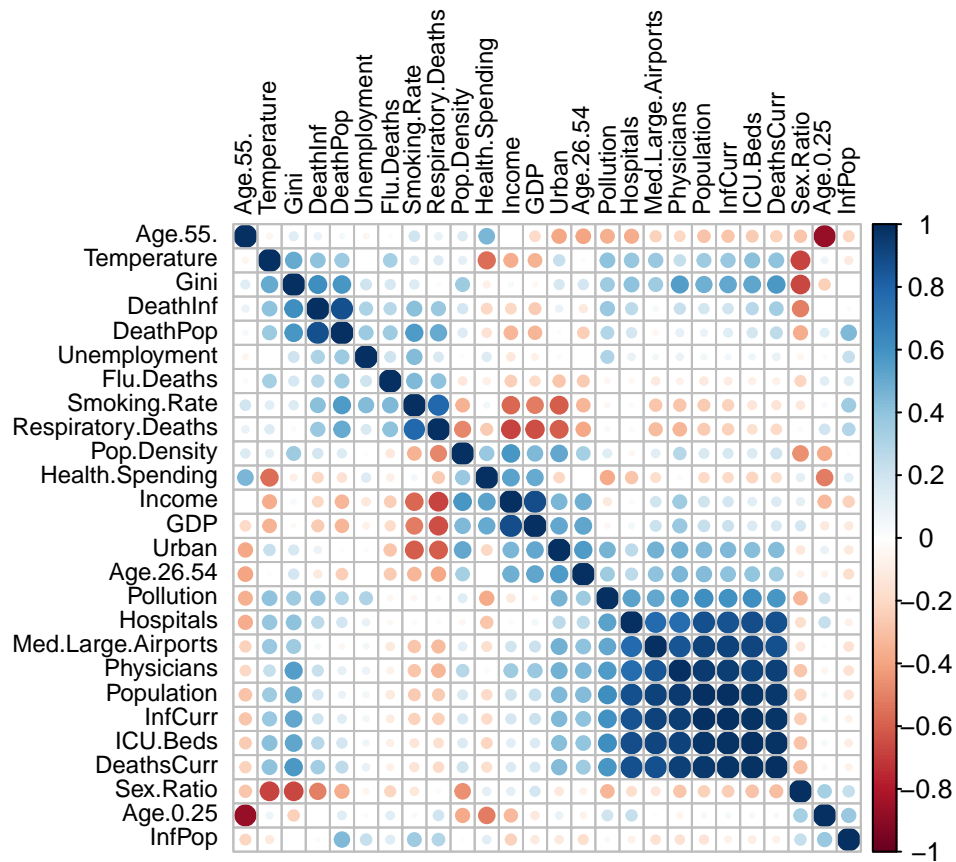
```
covid_cor_test <- cor.mtest(covidcor, conf.level = 0.95)
corrplot(full_covid_cor, p.mat = covid_cor_test$p, method = 'circle', insig='blank',
         number.cex = 0.5, tl.col = "black", tl.cex = .75, order = "hclust")
```
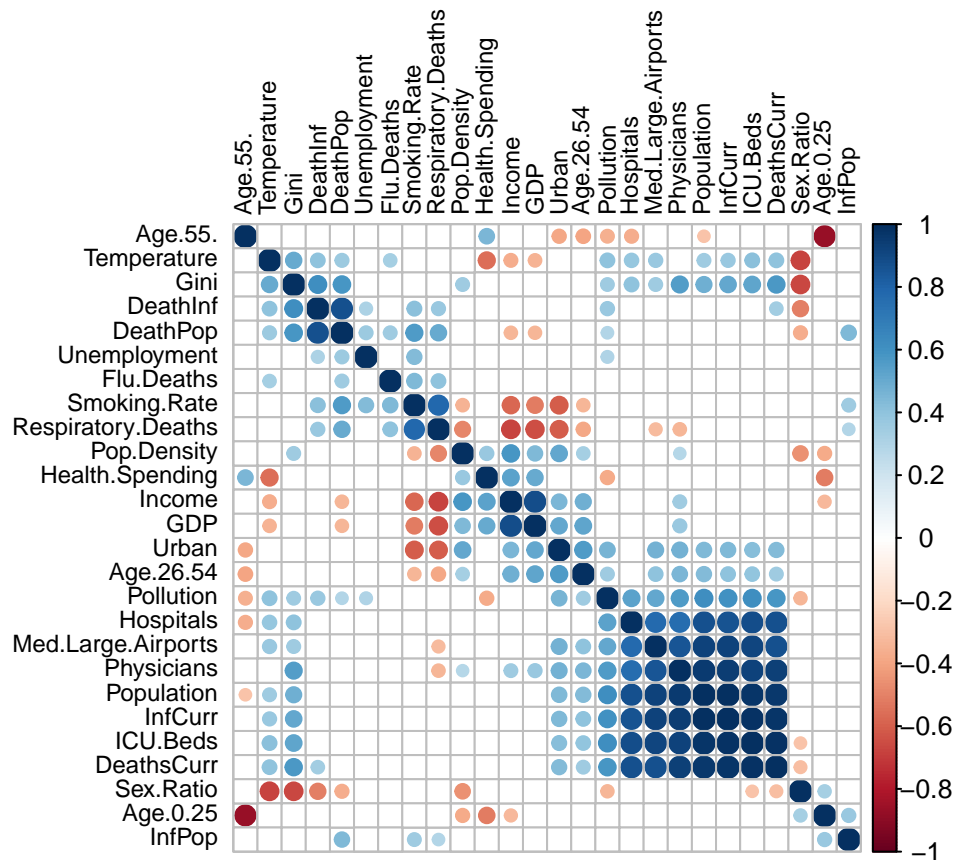
Deaths and Infections Boxplots and Scatterplots

```r
Pop.D.Quart <- ntile(covid$Pop.Density, 4)
covid.plot <- covid
covid.plot["Pop.D.Quart"] <- Pop.D.Quart

outliers <- function(x) {
  return(x < quantile(x, .25) - 1.5*IQR(x) | x > quantile(x, .75) + 1.5*IQR(x))
}


covid.plot <- covid.plot %>%
        group_by(Pop.D.Quart) %>%
        mutate(outlierDth = ifelse(outliers(DeathsCurr), State, NA))

covid.plot <- covid.plot %>%
        group_by(Pop.D.Quart) %>%
        mutate(outlierInf = ifelse(outliers(InfCurr), State, NA))


ggplot(data = covid.plot, aes(y = DeathsCurr, x = Pop.D.Quart, group = Pop.D.Quart)) + geom_boxplot(fill
  stat_summary(fun.y = mean, color = "red")+
 scale_y_continuous(name = "Deaths by covid") +
```

```
scale_x_continuous(name = "Population Density Grouped by Quartiles") +
    geom_text(aes(label=outlierDth), na.rm=TRUE, vjust=-0.6)
```

```
## Warning: The `fun.y` argument of `stat_summary()` is deprecated as of ggplot2 3.3.0.
## i Please use the `fun` argument instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

```
## Warning: Removed 4 rows containing missing values (`geom_segment()`).
```
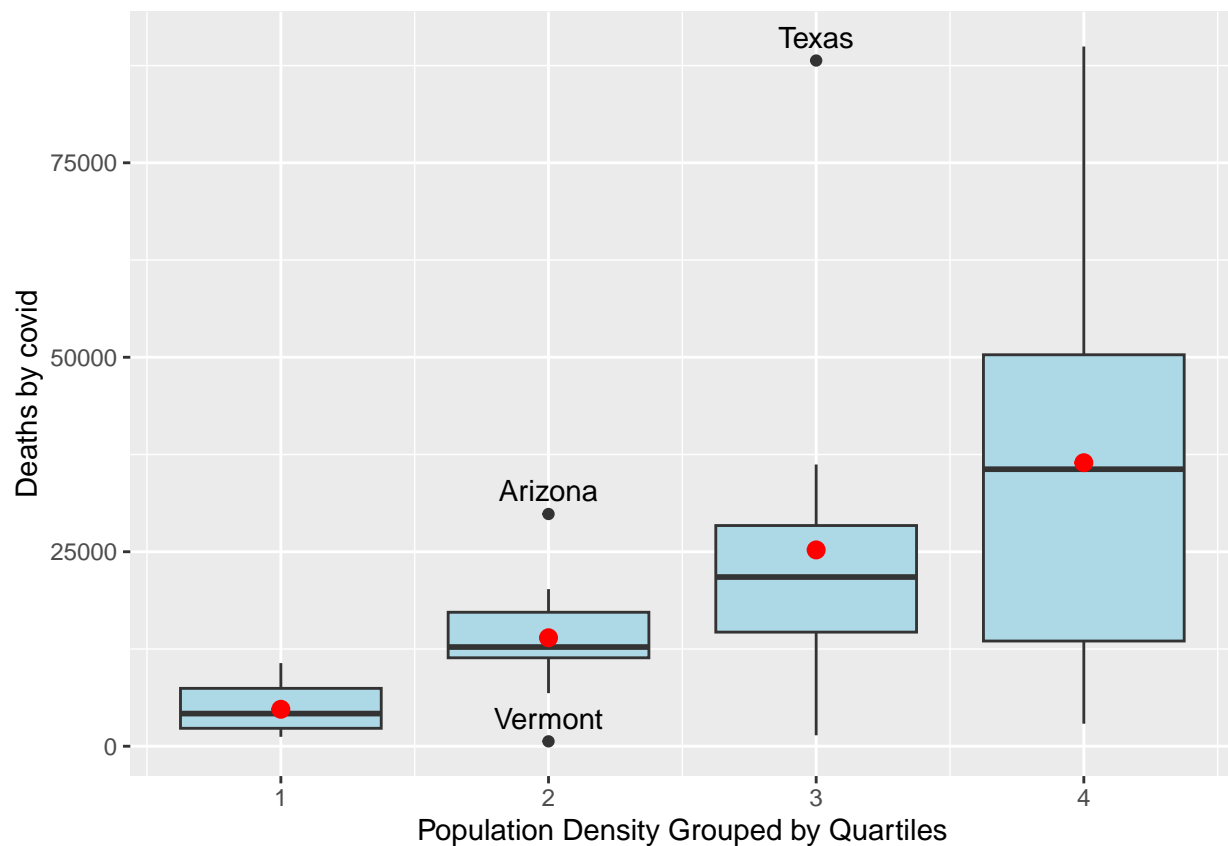


```
ggplot(data = covid.plot, aes(y = InfCurr, x = Pop.D.Quart, group = Pop.D.Quart)) + geom_boxplot(fill =
  stat_summary(fun.y = mean, color = "red")+
 scale_y_continuous(name = "Covid Infections") +
 scale_x_continuous(name = "Population Density Grouped by Quartiles") +
  geom_text(aes(label=outlierInf), na.rm=TRUE, vjust=1.2)
```

```
## Warning: Removed 4 rows containing missing values (`geom_segment()`).
```
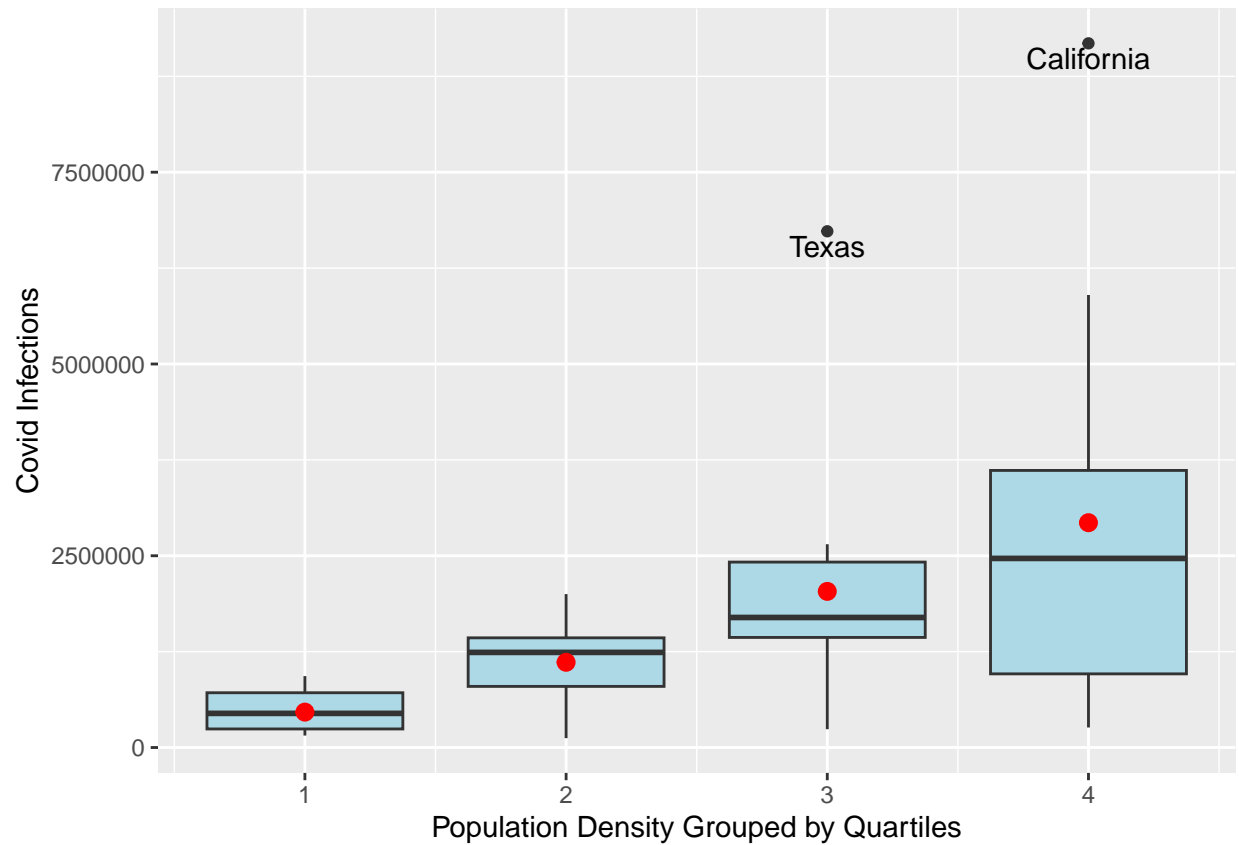
```
ggplot(data = covid, aes(x = Population, y = DeathsCurr)) + geom_point() +
 scale_y_continuous(name = "Deaths by covid") +
  geom_smooth(method = "lm")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```
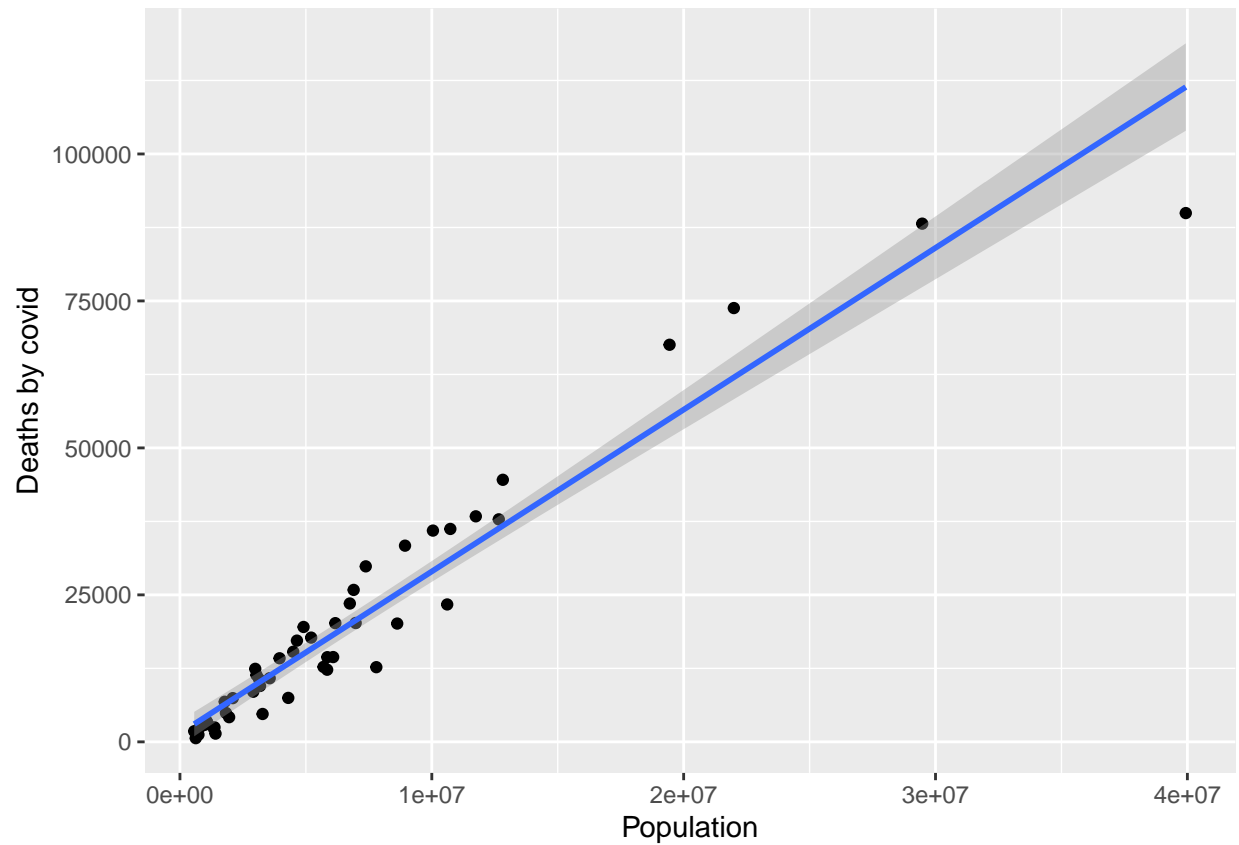
```
ggplot(data = covid, aes(x = Population, y = InfCurr)) + geom_point() +
 scale_y_continuous(name = "Covid Infections") +
  geom_smooth(method = "lm")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```
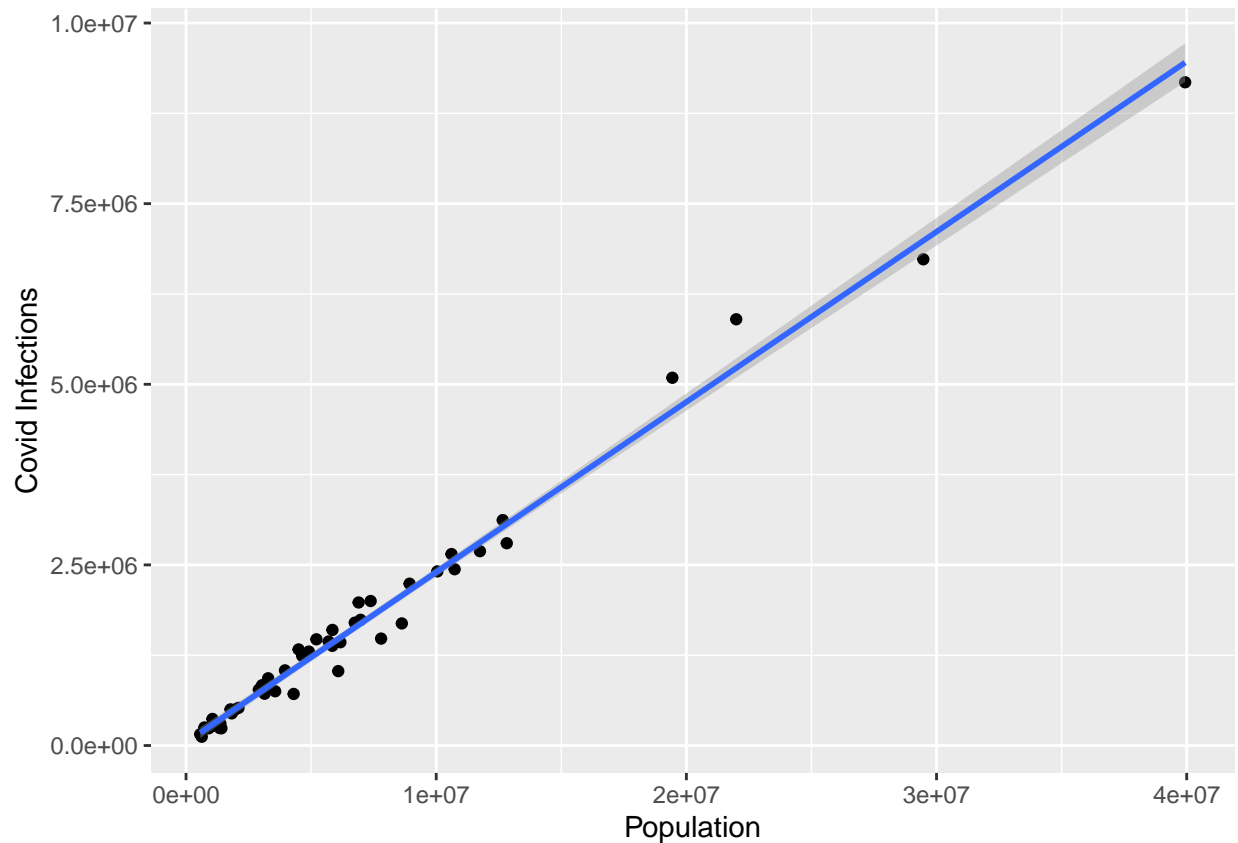
As population is naturally correlated to both values, it will be excluded from our regressions to get a better look at other contributing factors.

Histograms showing the distributions of States across various factors

```
ggplot(covid, aes(x = DeathInf)) + geom_histogram(bins = 10, fill = "light blue", color = "black")  +
 scale_y_continuous(name = "Number of States") +
 scale_x_continuous(name = "Percentage Dead after Infected")
```
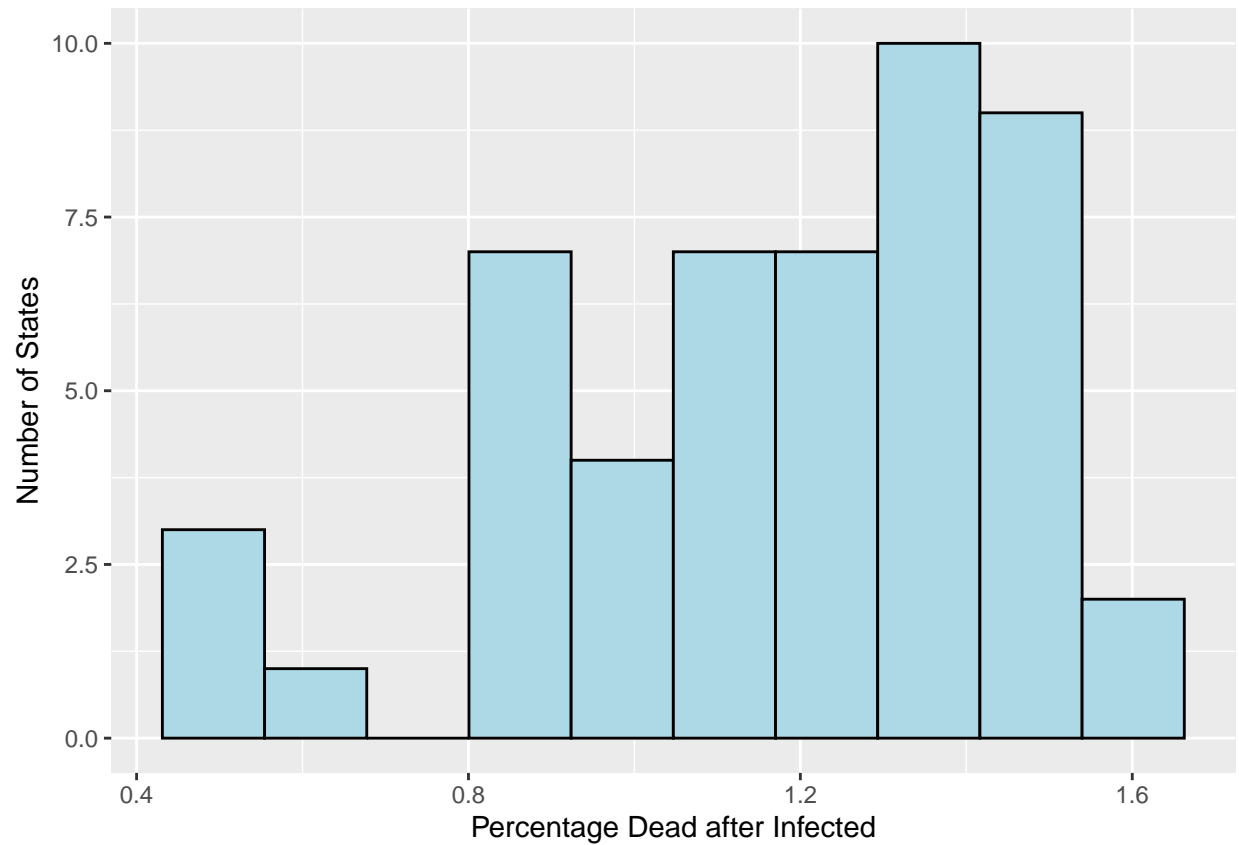
```
ggplot(covid, aes(x = DeathsCurr)) + geom_histogram(bins = 10, fill = "light blue", color = "black")
 scale_y_continuous(name = "Number of States") +
 scale_x_continuous(name = "Current Deaths from Covid")
```
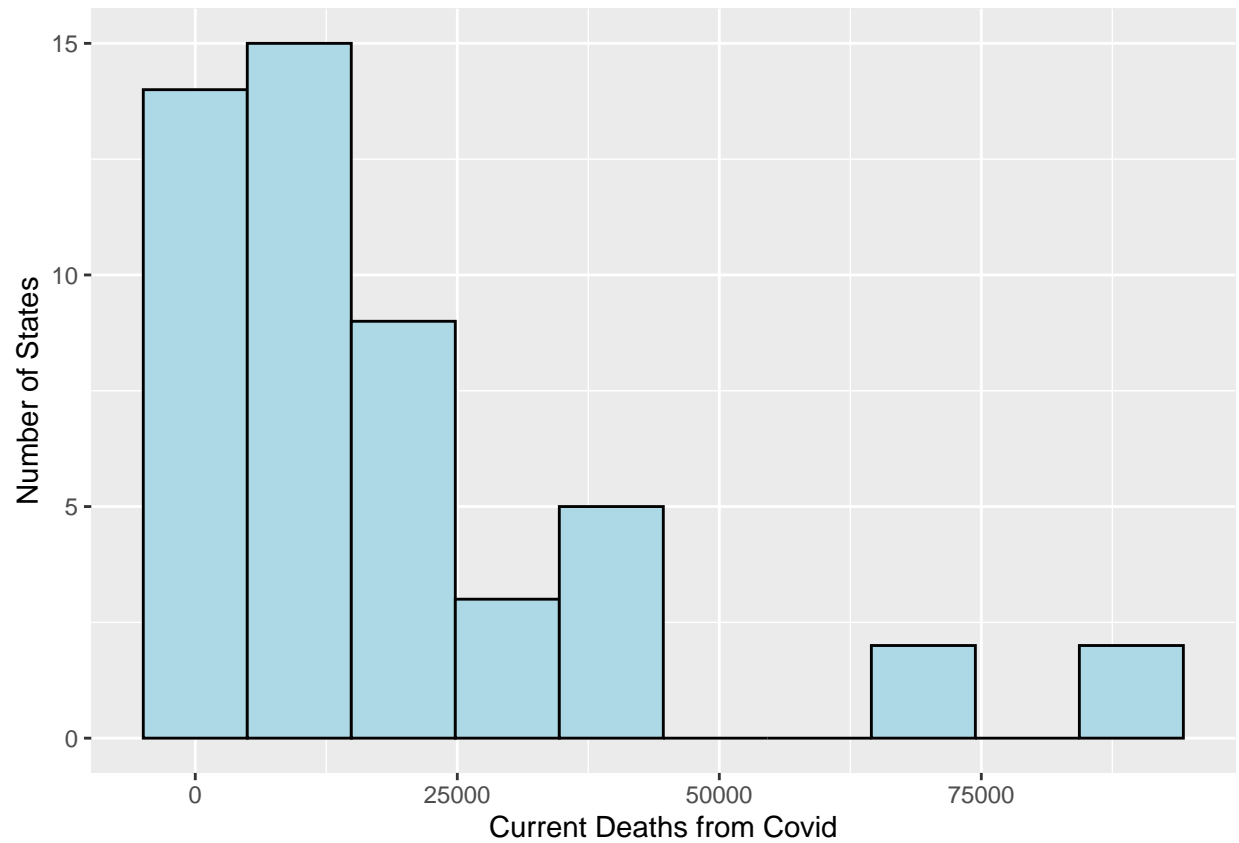
```
ggplot(covid, aes(x = InfCurr)) + geom_histogram(bins = 10, fill = "light blue", color = "black")    +
 scale_y_continuous(name = "Number of States") +
 scale_x_continuous(name = "Current Infections from Covid")
```
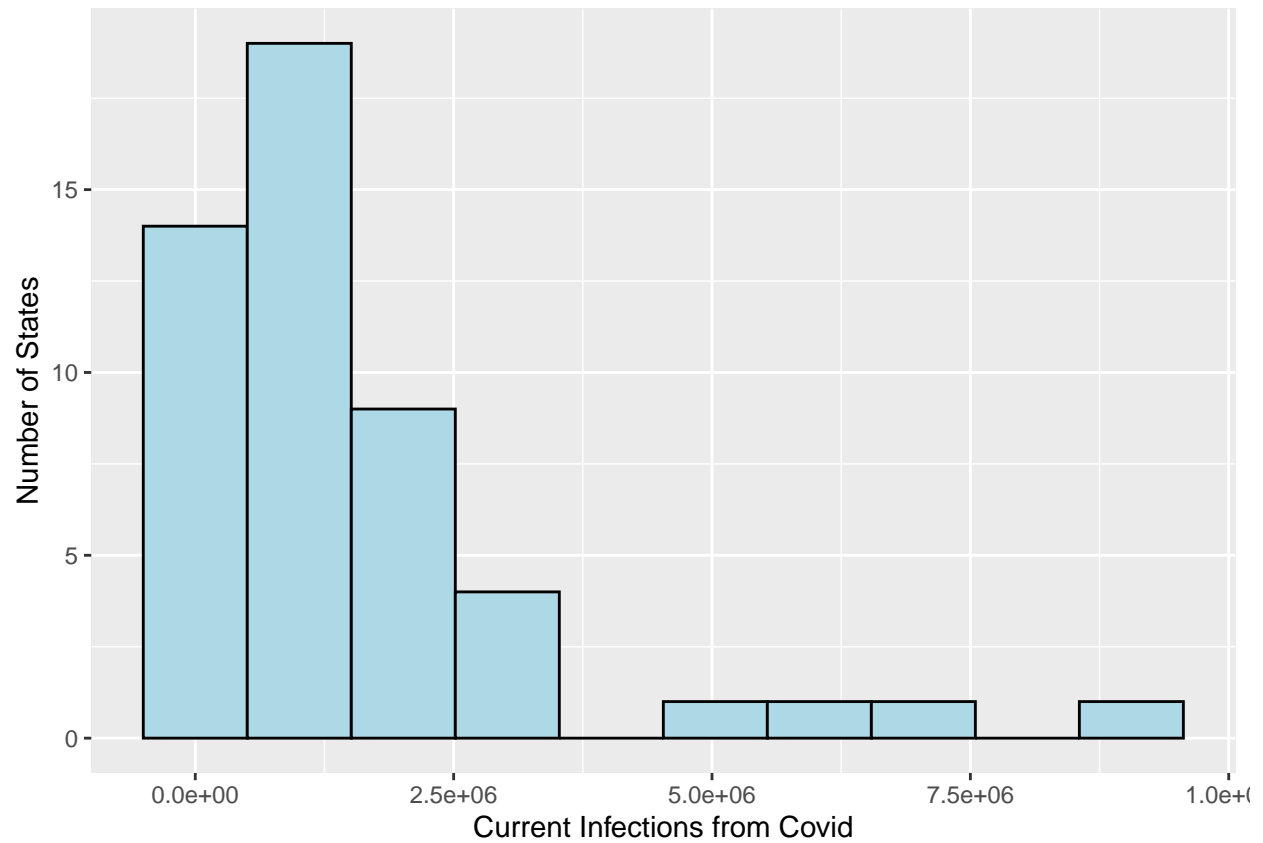
```
ggplot(covid, aes(x = Pop.Density)) + geom_histogram(bins = 10, fill = "light blue", color = "black")
 scale_y_continuous(name = "Number of States") +
 scale_x_continuous(name = "Population Density")
```
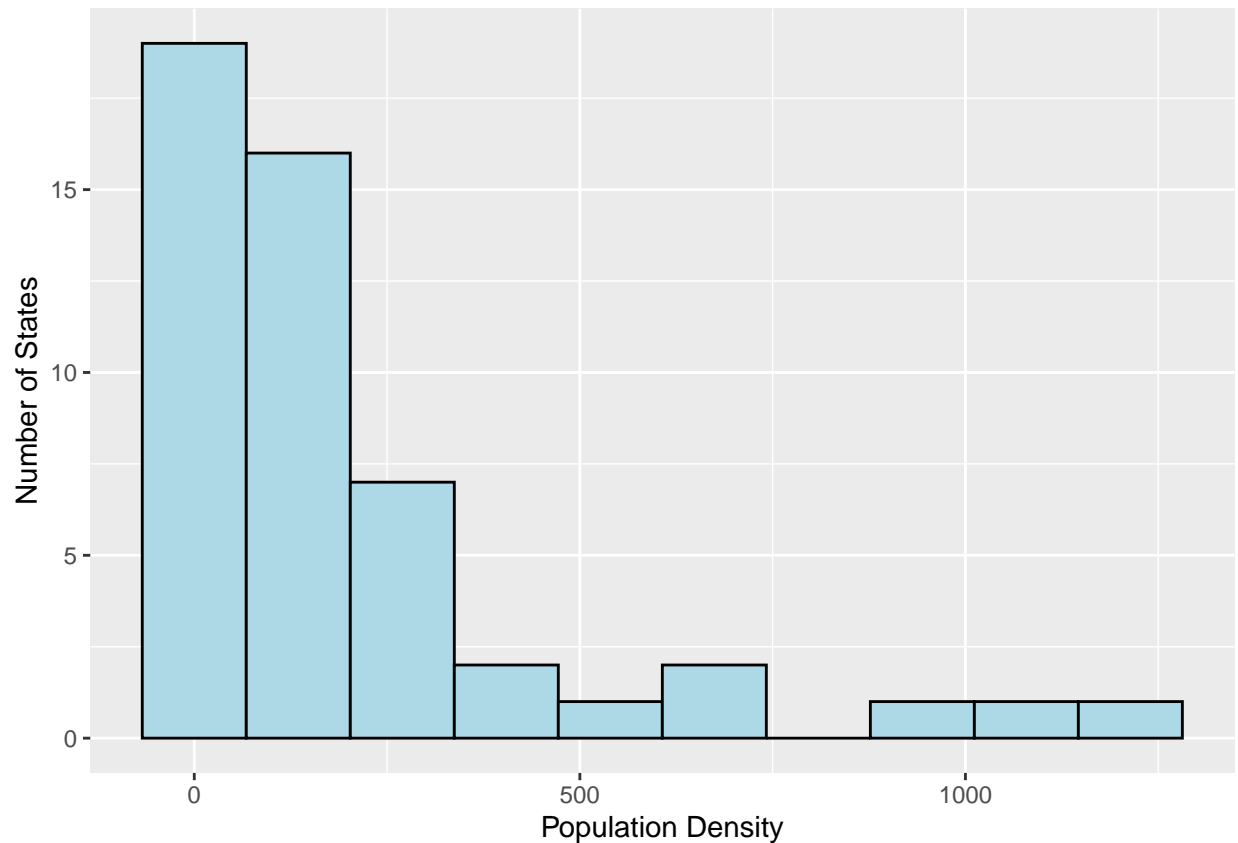
Deaths per Infection Regression Analysis

```
# Defining our selected variables and determining optimal number
# based on Adjusted R-Squared, Cp, and BIC.

covid.DI <- subset(covid, select = -c(Population, Pop.Density, DeathsCurr, InfCurr, InfPop, DeathPop, S

reg.full.DI <- regsubsets(DeathInf ~., data = covid.DI)
reg.full.DI.sum <- summary(reg.full.DI)

which.max(reg.full.DI.sum$adjr2)
```

```
## [1] 8
```

```
which.min(reg.full.DI.sum$cp)
```

```
## [1] 7
```

```
which.min(reg.full.DI.sum$bic)
```

```
## [1] 6
```

```
# Adjusted R-Squared against Number of Predictors
plot(reg.full.DI.sum$adjr2, xlab = "Number of Predictors", ylab = "Adjusted R-Squared", type = "l")
points(8, reg.full.DI.sum$adjr2[8], col = "red", cex = 2,
pch = 20)
```
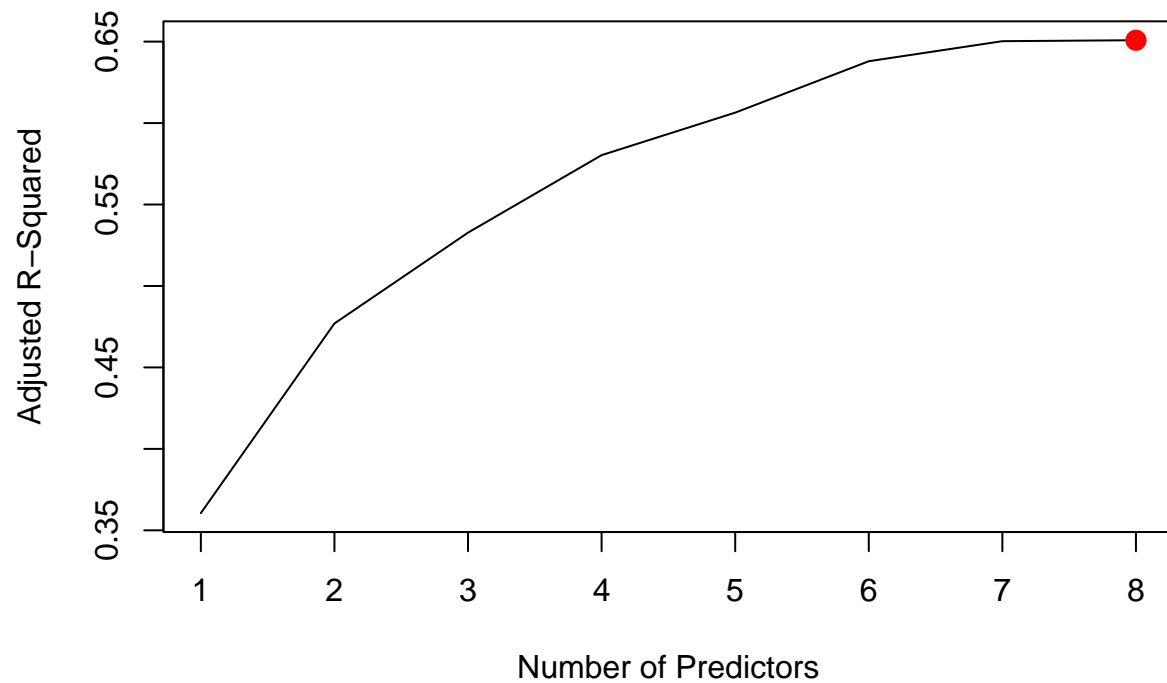


```
# Cp against Number of Predictors
plot(reg.full.DI.sum$cp, xlab = "Number of Predictors", ylab = "Mallow CP", type = "l")
points(7, reg.full.DI.sum$cp[7], col = "red", cex = 2,
pch = 20)
```
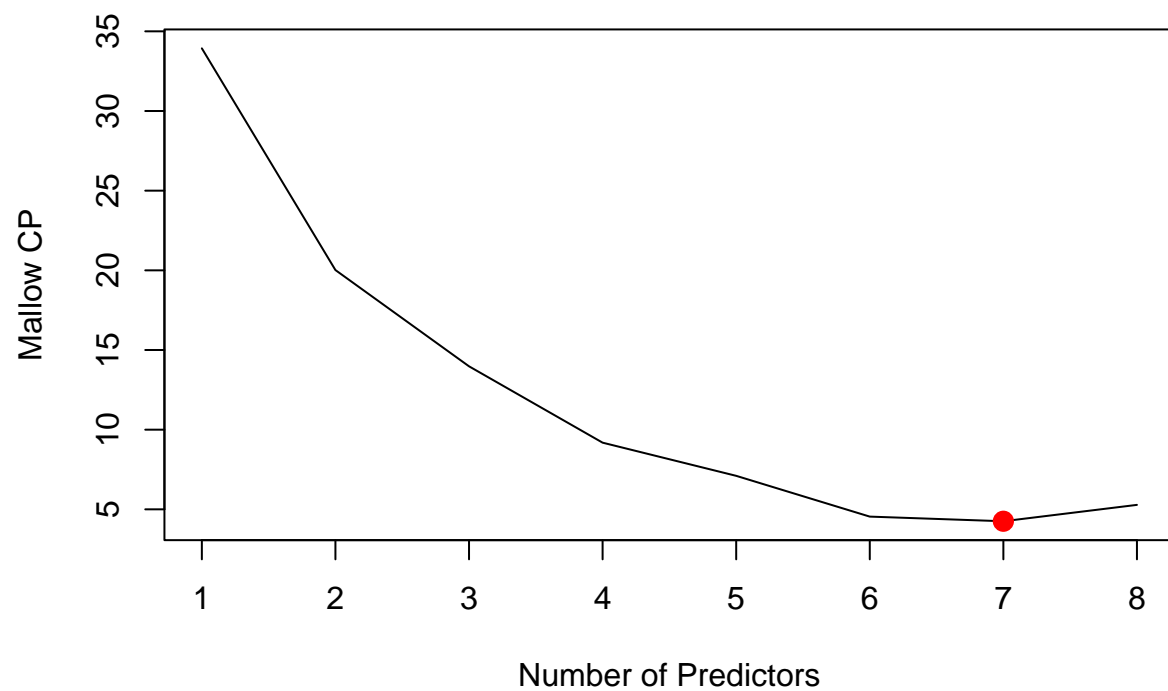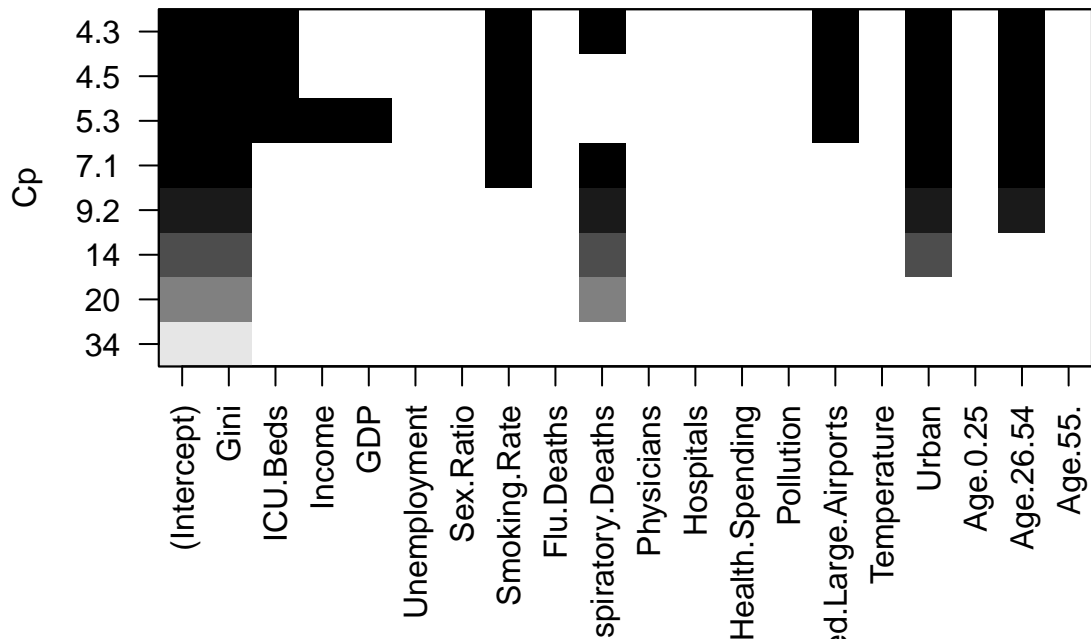
```
# BIC against Number of Predictors
plot(reg.full.DI.sum$bic, xlab = "Number of Predictors", ylab = "BIC", type = "l")
points(6, reg.full.DI.sum$bic[6], col = "red", cex = 2,
pch = 20)
```
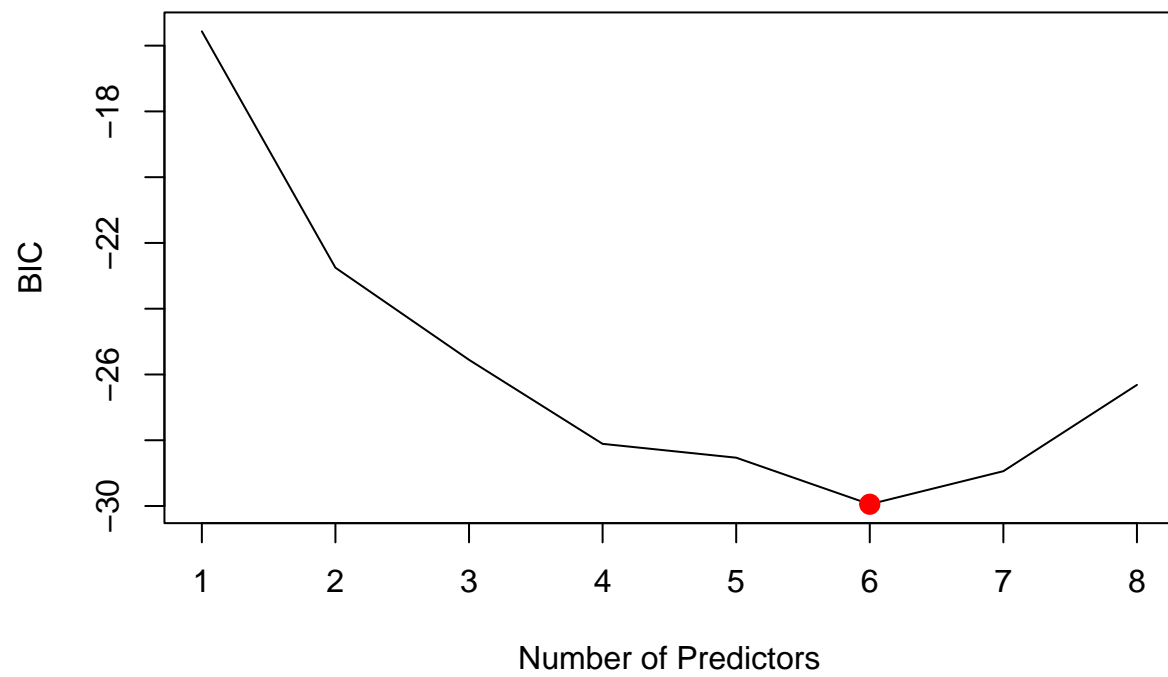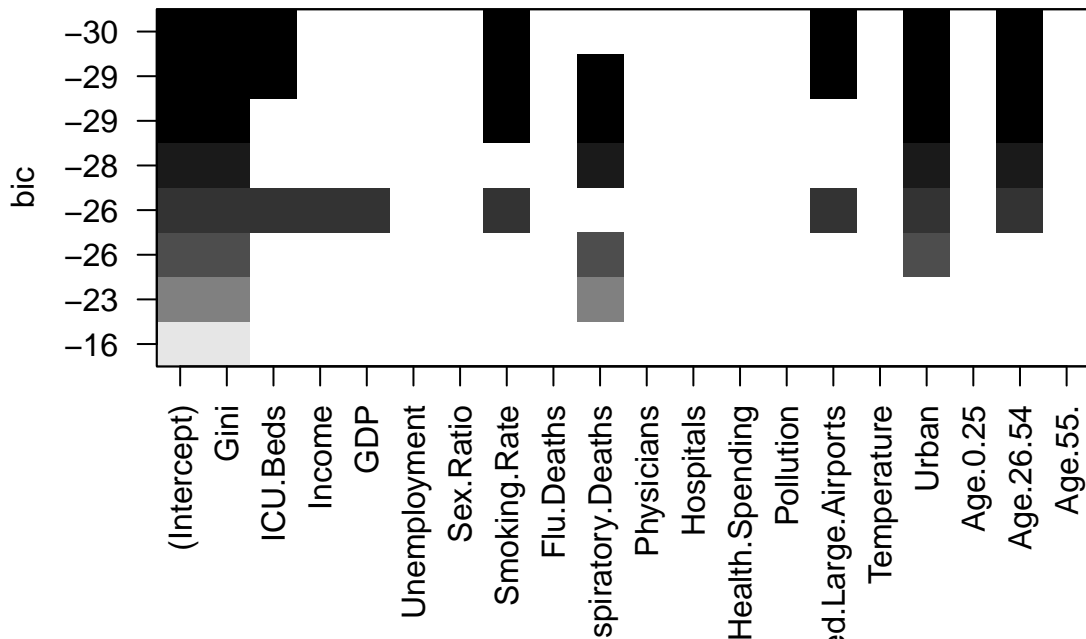
```r
# Best variable selections for Adjusted R-Squared
plot(reg.full.DI, scale = "adjr2")
```
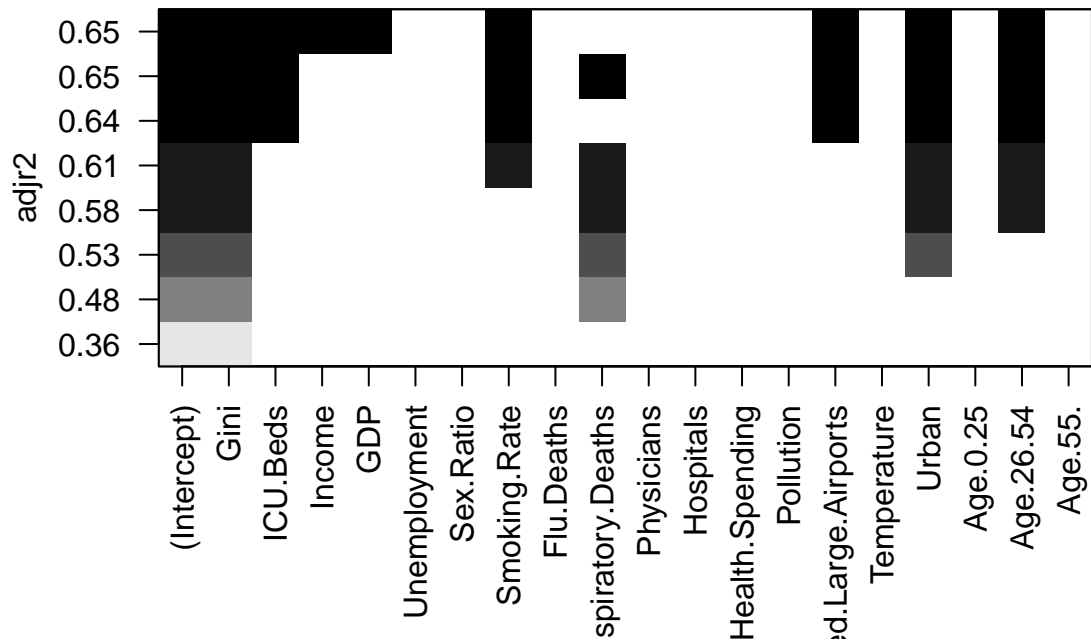
```r
# Best variable selections for Cp
plot(reg.full.DI, scale = "Cp")
```

```r
# Best variable selections for BIC
plot(reg.full.DI, scale = "bic")
```

```r
# Final Coefficients
coef(reg.full.DI, 6)
```

```
##        (Intercept)                Gini             ICU.Beds          Smoking.Rate
##       -0.2641526410        5.5471943547        0.0001329616         0.0377828873
## Med.Large.Airports               Urban            Age.26.54
##       -0.1251790745        0.0106097959       -7.0091894078
```
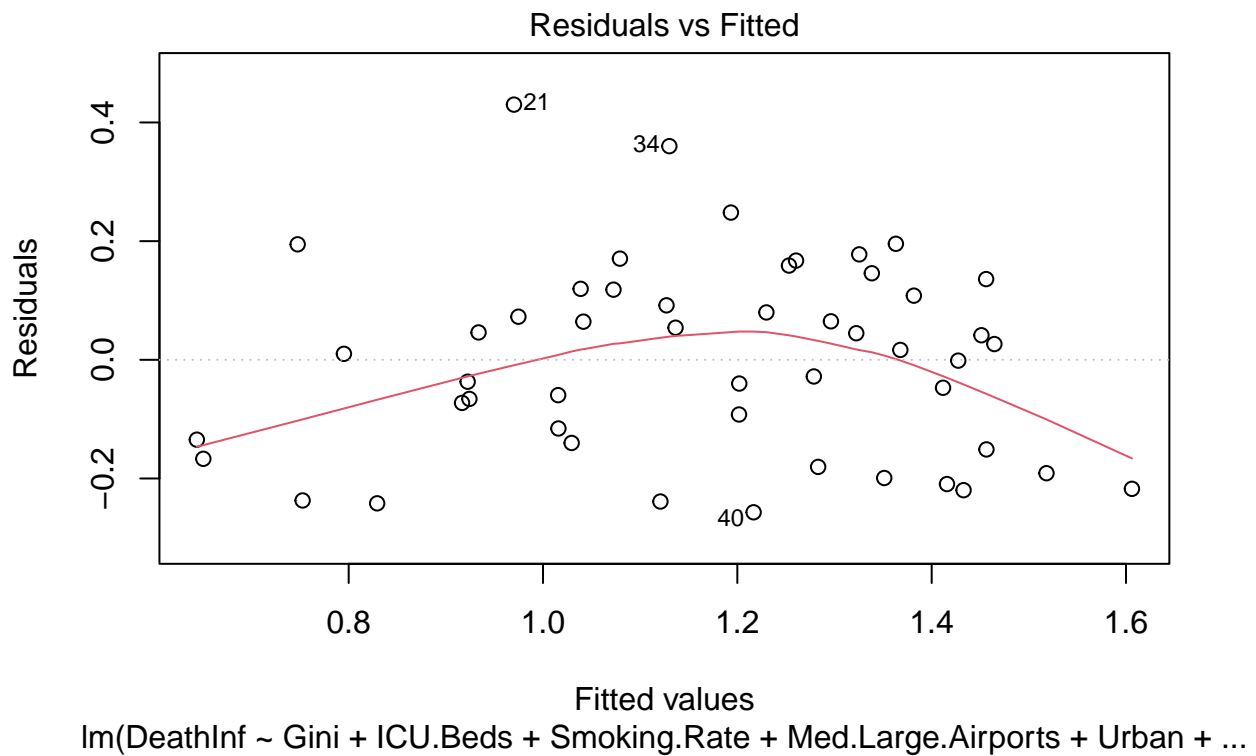
Final model for Deaths Per Infected and Diagnostics
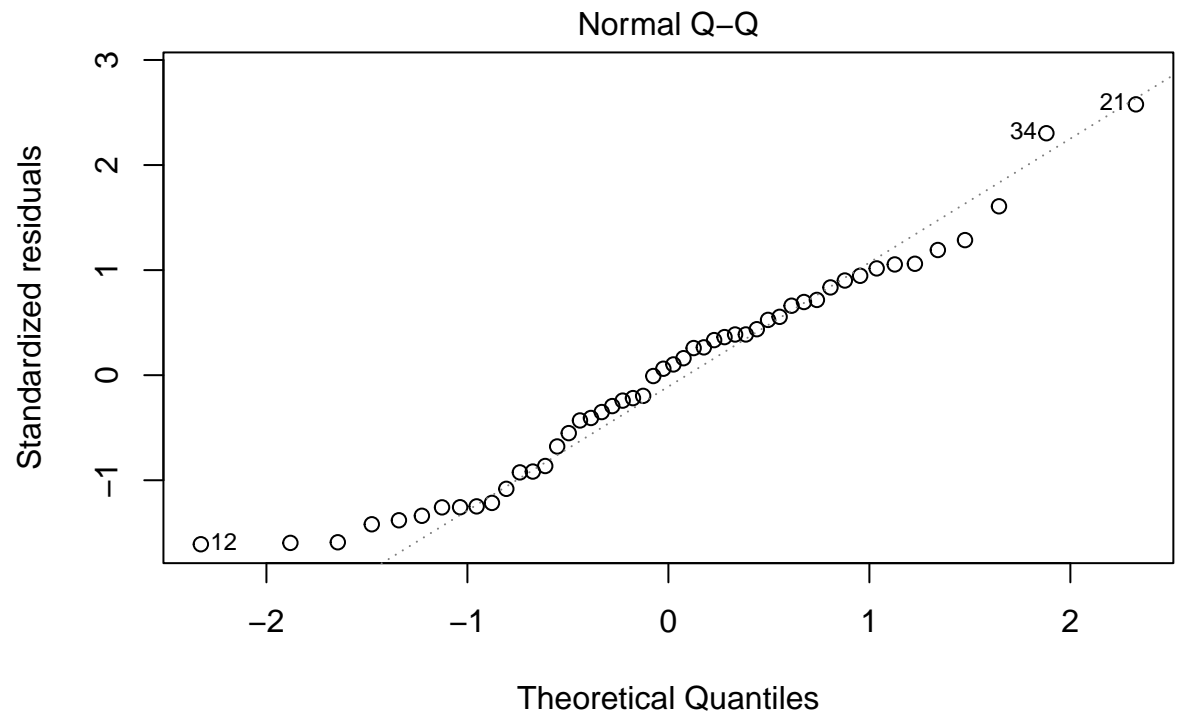
```r
DI.best <- lm(data = covid.DI, DeathInf ~ Gini + ICU.Beds + Smoking.Rate +  Med.Large.Airports + Urban
summary(DI.best)
```

```
##
## Call:
## lm(formula = DeathInf ~ Gini + ICU.Beds + Smoking.Rate + Med.Large.Airports +
##     Urban + Age.26.54, data = covid.DI)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.25696 -0.13874  0.01339  0.11569  0.43007
##
## Coefficients:
```

```
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -2.642e-01  1.064e+00  -0.248 0.805108
## Gini                5.547e+00  1.518e+00   3.655 0.000696 ***
## ICU.Beds            1.330e-04  4.743e-05   2.803 0.007563 **
## Smoking.Rate        3.778e-02  9.694e-03   3.897 0.000335 ***
## Med.Large.Airports -1.252e-01  3.956e-02  -3.164 0.002854 **
## Urban               1.061e-02  2.603e-03   4.077 0.000194 ***
## Age.26.54          -7.009e+00  2.353e+00  -2.979 0.004744 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1738 on 43 degrees of freedom
## Multiple R-squared:  0.6823, Adjusted R-squared:  0.638
## F-statistic: 15.39 on 6 and 43 DF,  p-value: 2.523e-09
```

```
plot(DI.best)
```


Residuals vs Fitted

lm(DeathInf ~ Gini + ICU.Beds + Smoking.Rate + Med.Large.Airports + Urban + ...

Normal Q–Q

Standardized residuals

Theoretical Quantiles
lm(DeathInf ~ Gini + ICU.Beds + Smoking.Rate + Med.Large.Airports + Urban + ...

Scale−Location

lm(DeathInf ~ Gini + ICU.Beds + Smoking.Rate + Med.Large.Airports + Urban + ...

Residuals vs Leverage

lm(DeathInf ~ Gini + ICU.Beds + Smoking.Rate + Med.Large.Airports + Urban + ...
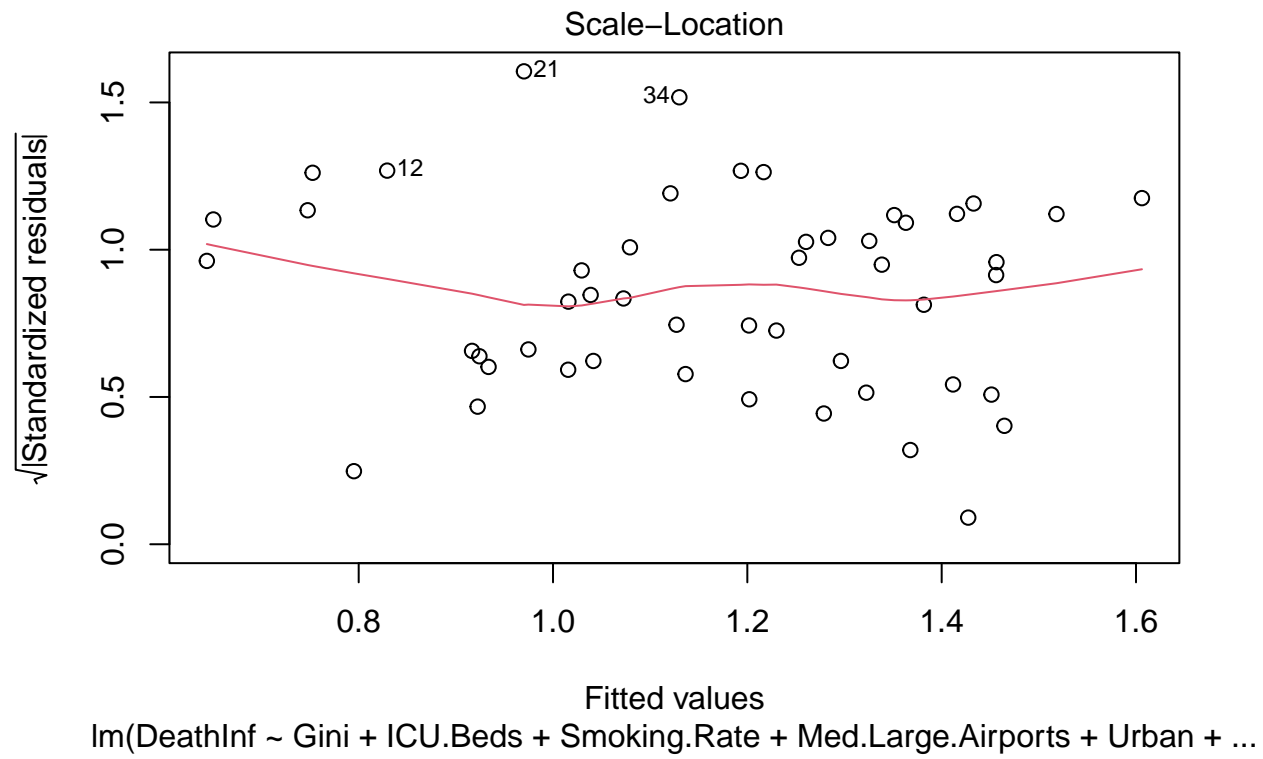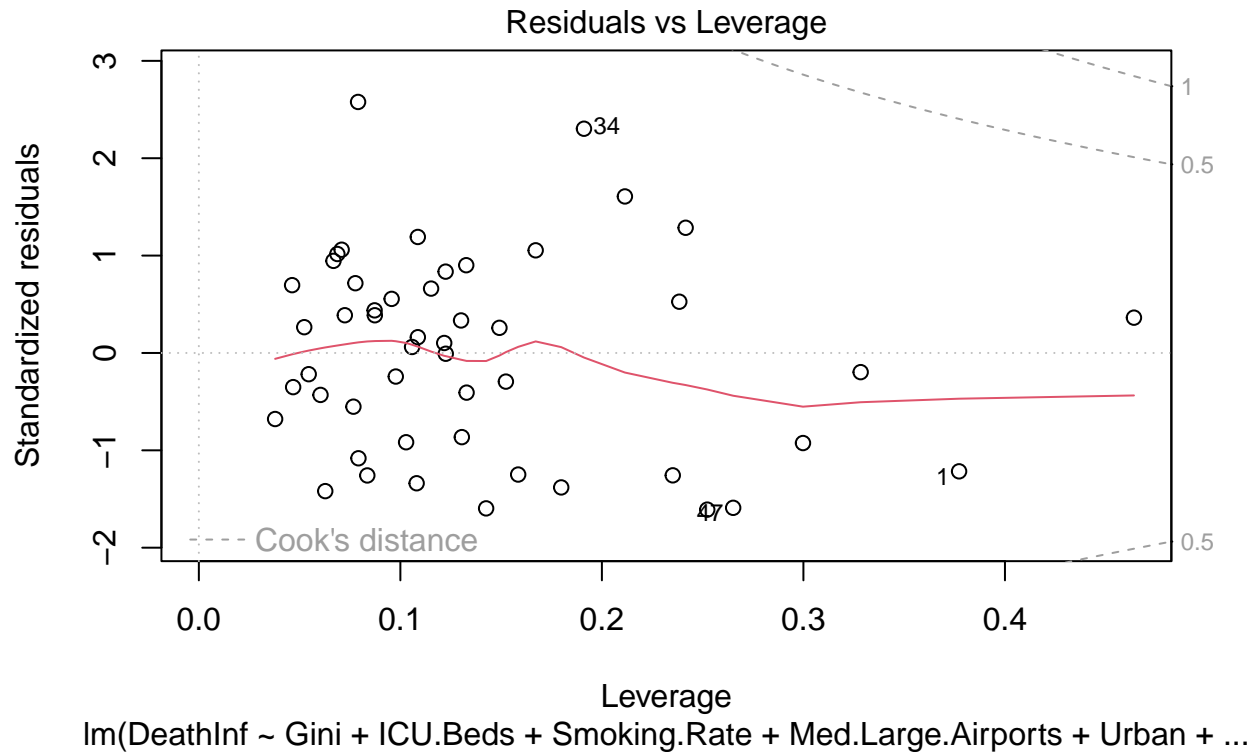
Infected Population Regression Analysis

```
# Defining our selected variables and determining optimal number
# based on Adjusted R-Squared, Cp, and BIC.

covid.IP <- subset(covid, select = -c(Population, Pop.Density, DeathsCurr, InfCurr, DeathInf, DeathPop,

reg.full.IP <- regsubsets(InfPop ~., data = covid.IP)
reg.full.IP.sum <- summary(reg.full.IP)

which.max(reg.full.IP.sum$adjr2)
```

```
## [1] 8
```

```
which.min(reg.full.IP.sum$cp)
```

```
## [1] 8
```
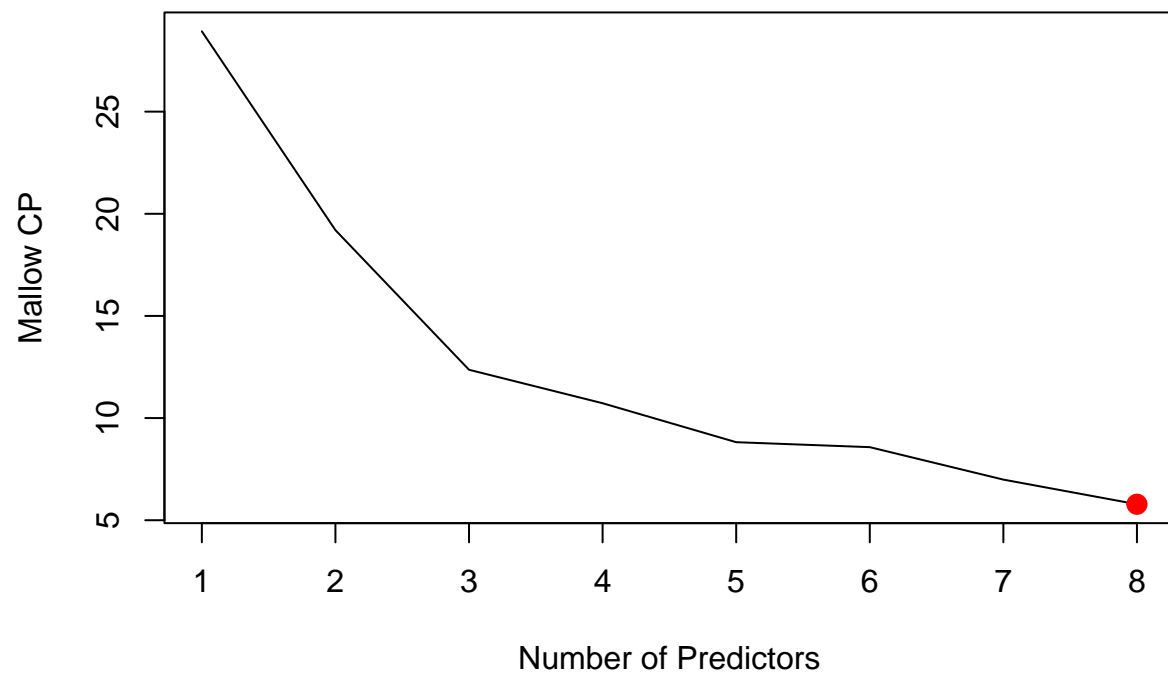
```
which.min(reg.full.IP.sum$bic)
```

```
## [1] 3
```

```
# Adjusted R-Squared against Number of Predictors
plot(reg.full.IP.sum$adjr2, xlab = "Number of Predictors", ylab = "Adjusted R-Squared", type = "l")
points(8, reg.full.IP.sum$adjr2[8], col = "red", cex = 2,
pch = 20)
```
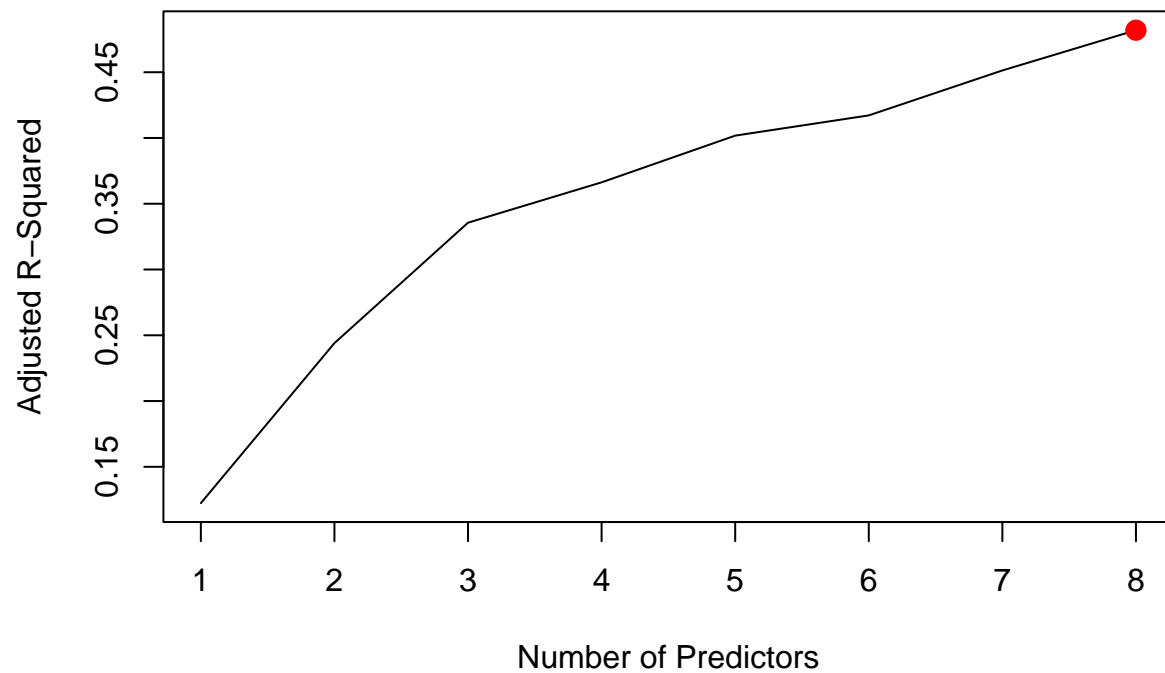


```
# Cp against Number of Predictors
plot(reg.full.IP.sum$cp, xlab = "Number of Predictors", ylab = "Mallow CP", type = "l")
points(8, reg.full.IP.sum$cp[8], col = "red", cex = 2,
pch = 20)
```
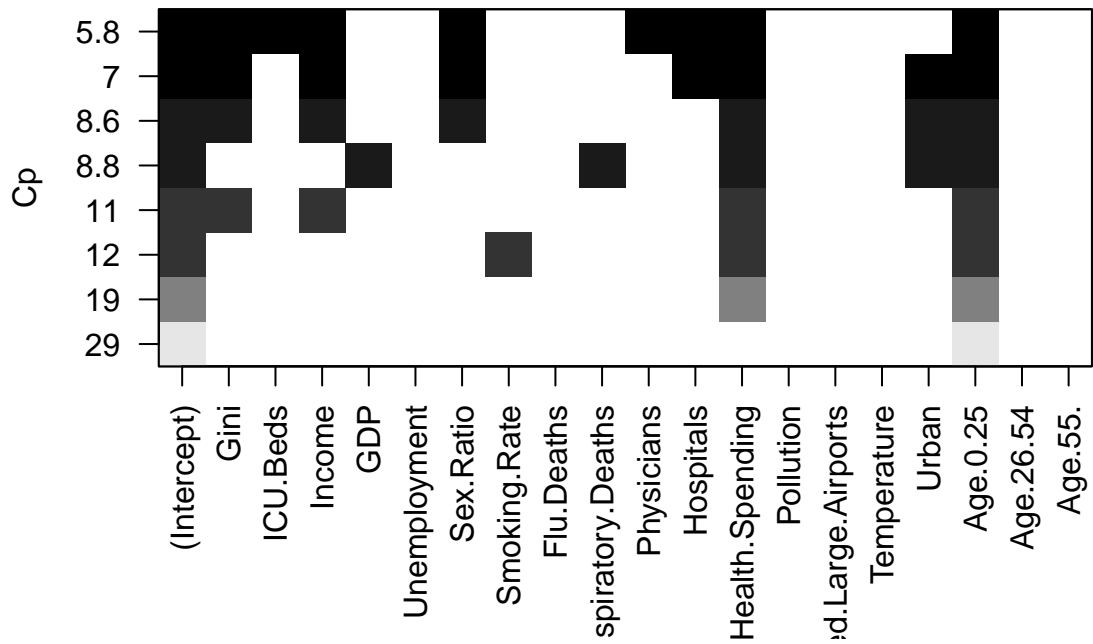
```
# BIC against Number of Predictors
plot(reg.full.IP.sum$bic, xlab = "Number of Predictors", ylab = "BIC", type = "l")
points(3, reg.full.IP.sum$bic[3], col = "red", cex = 2,
pch = 20)
```
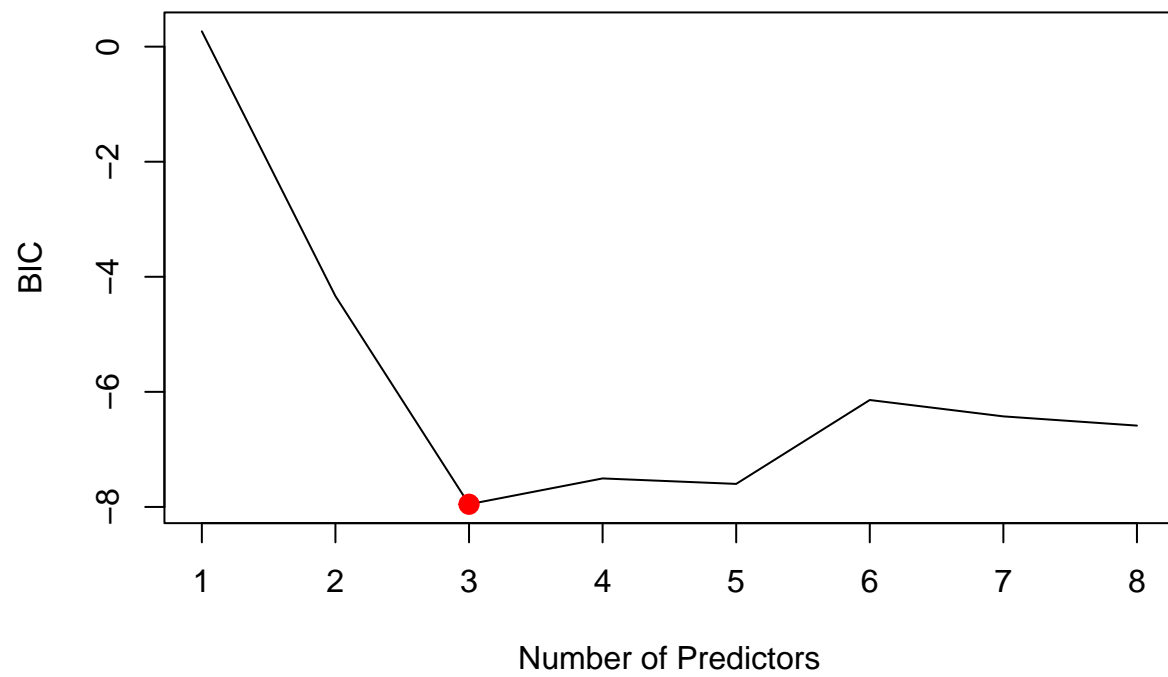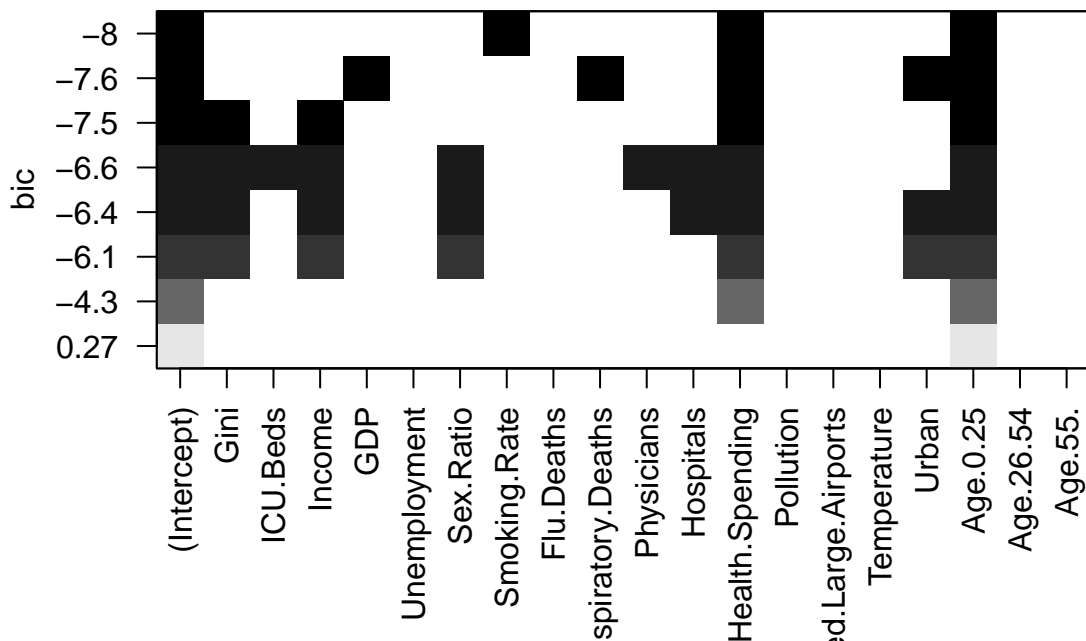
```
# Best variable selections for Adjusted R-Squared
plot(reg.full.IP, scale = "adjr2")
```
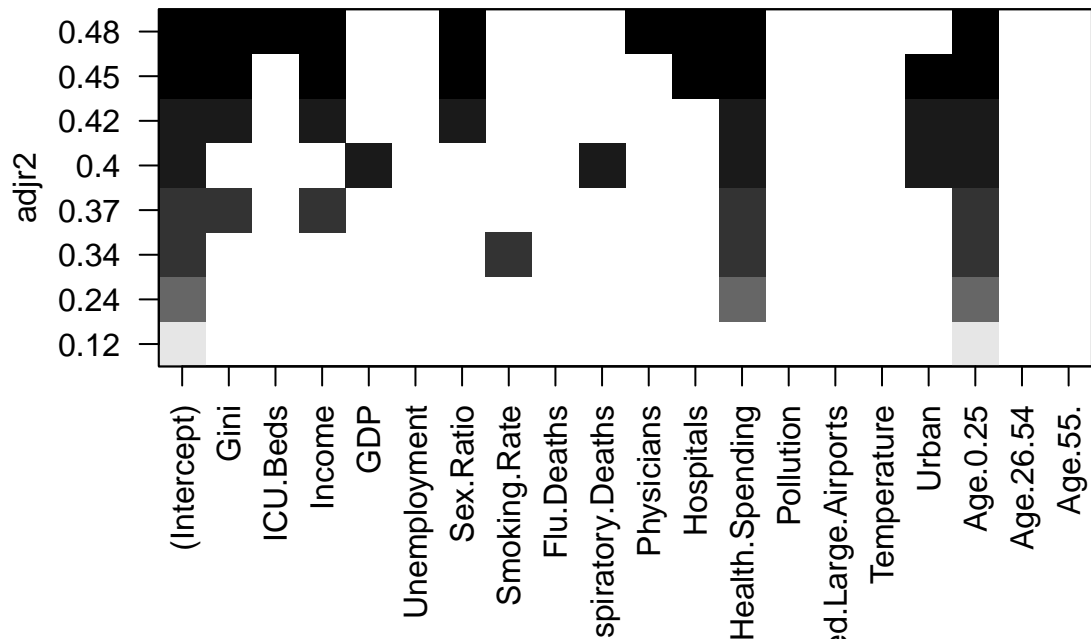
```
# Best variable selections for Cp
plot(reg.full.IP, scale = "Cp")
```

```
# Best variable selections for BIC
plot(reg.full.IP, scale = "bic")
```

```
# Final Coefficients
coef(reg.full.IP, 8)
```

```
##     (Intercept)           Gini        ICU.Beds          Income        Sex.Ratio
## -1.008641e+02   9.381372e+01    3.105573e-03   -1.628671e-04    3.931816e+01
##       Physicians       Hospitals Health.Spending        Age.0.25
## -1.306106e-04   -3.225497e-02    2.444950e-03    1.040697e+02
```
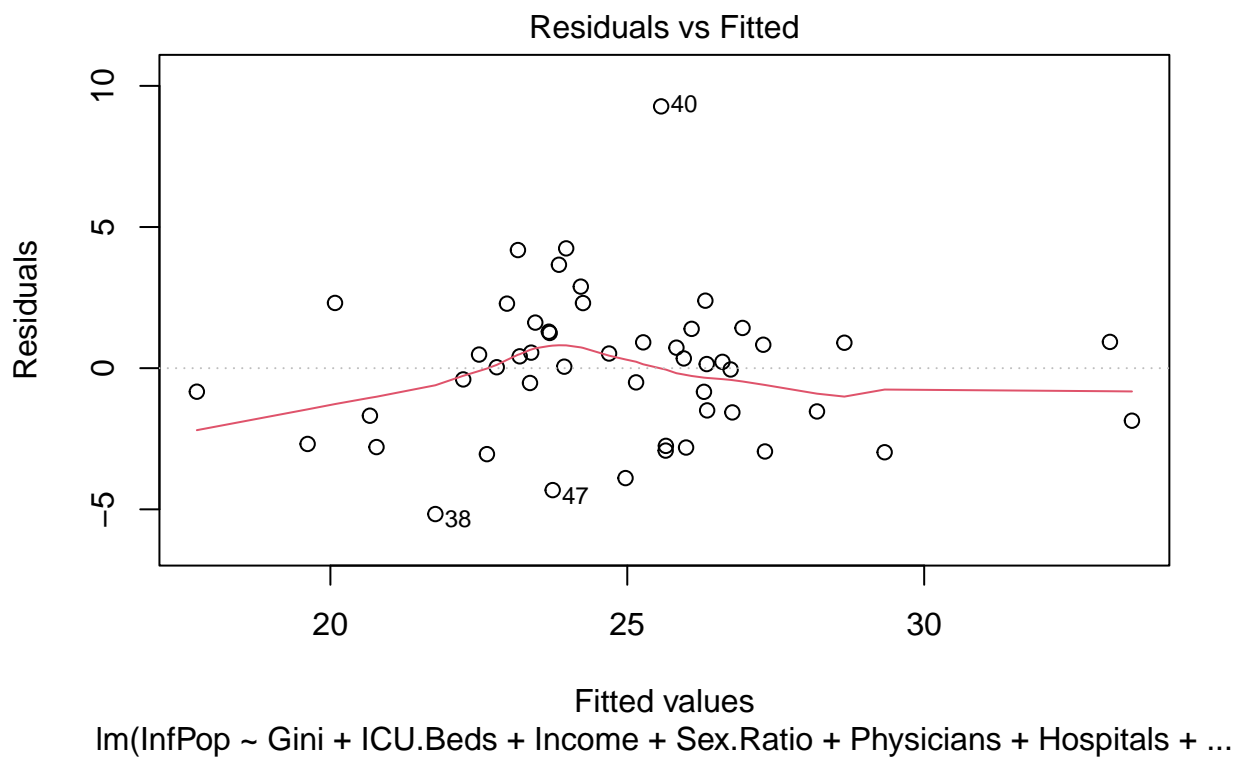
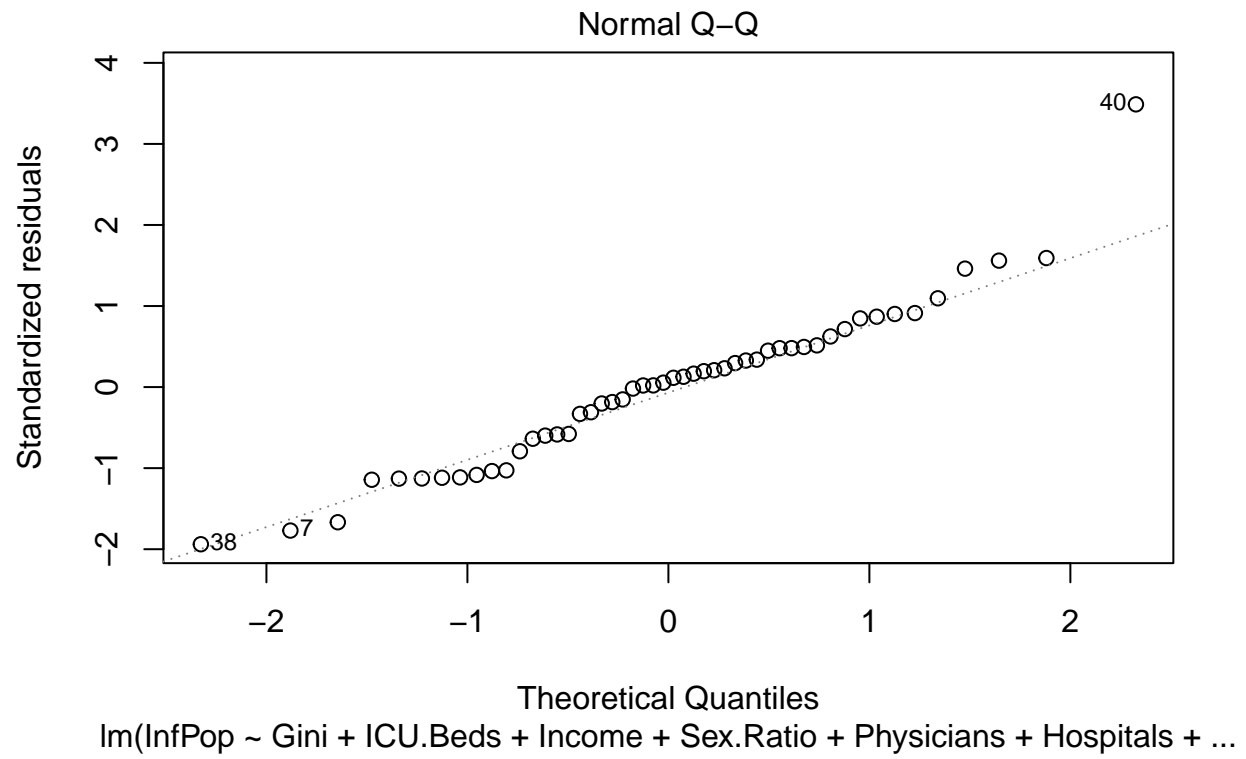## Final model for Infections per Population and Diagnostics

```
IP.best <- lm(data = covid.IP, InfPop ~ Gini + ICU.Beds + Income + Sex.Ratio + Physicians + Hospitals +
summary(IP.best)
```
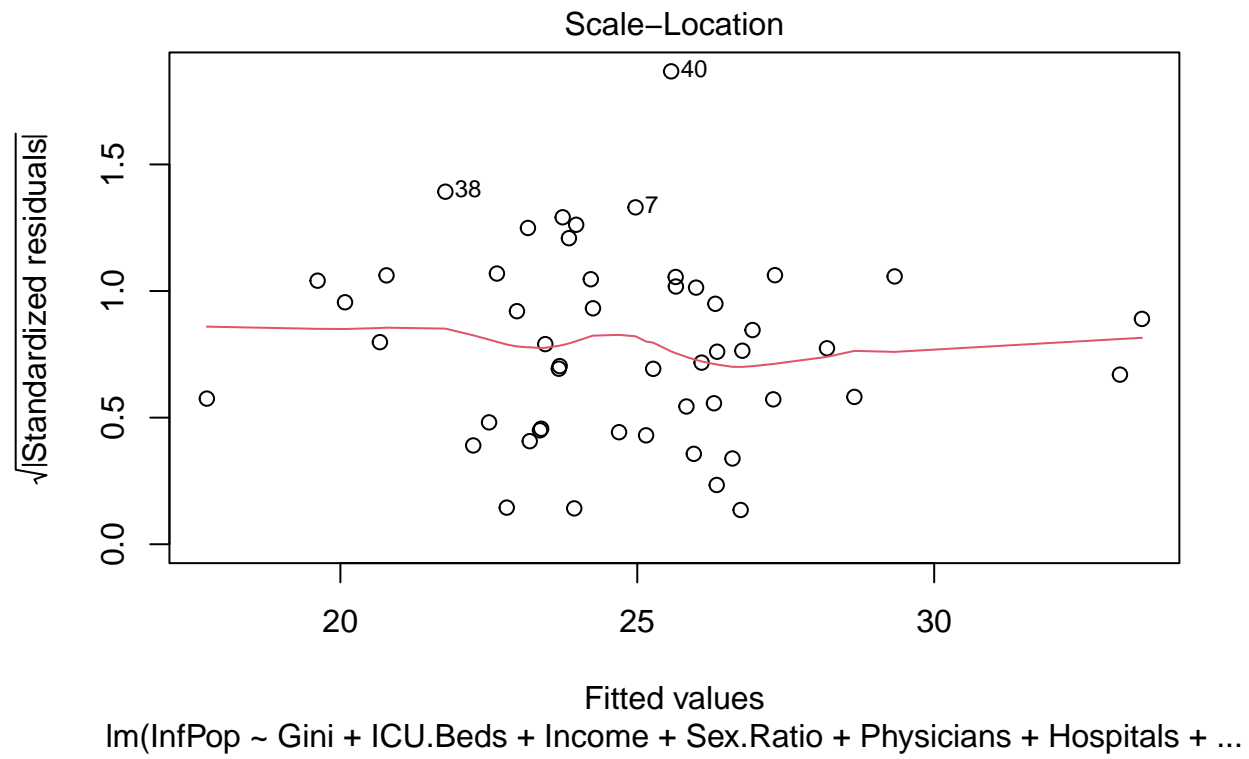
```
##
## Call:
## lm(formula = InfPop ~ Gini + ICU.Beds + Income + Sex.Ratio +
##     Physicians + Hospitals + Health.Spending + Age.0.25, data = covid.IP)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.1671 -1.6548  0.1852  1.2813  9.2729
##
## Coefficients:
```
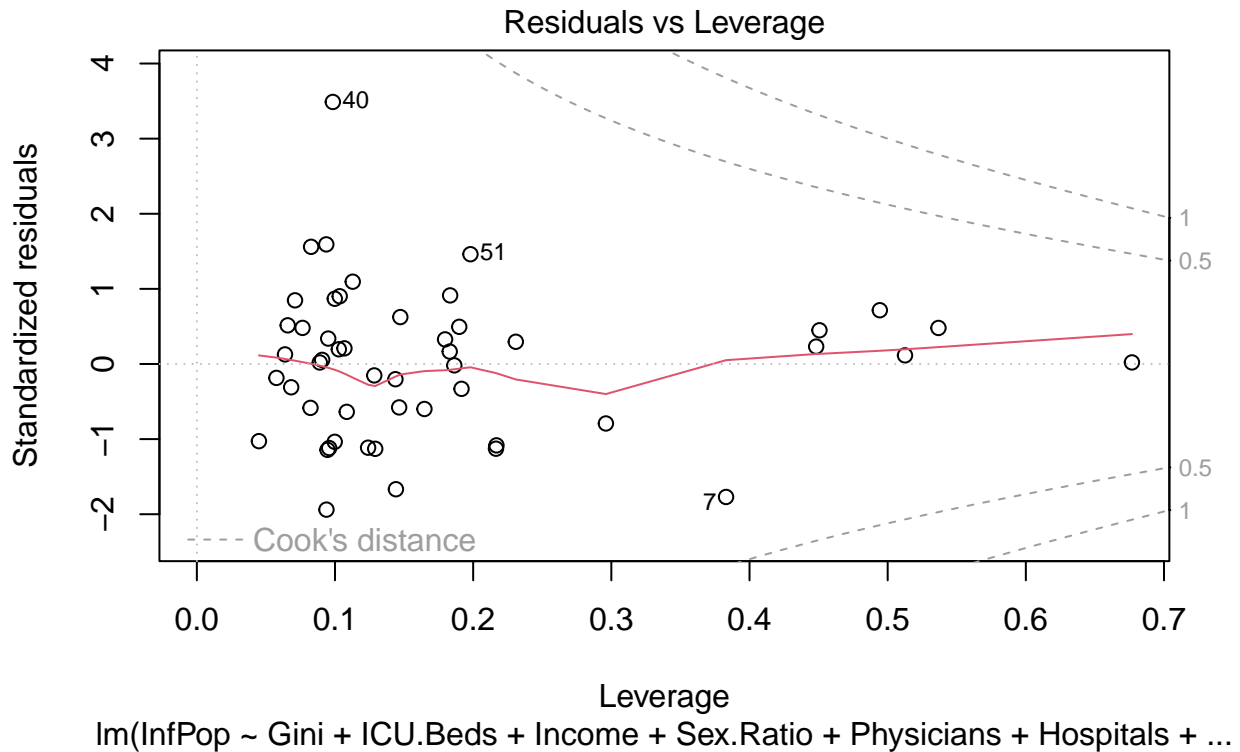
```
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -1.009e+02  2.874e+01  -3.510  0.00111 **
## Gini             9.381e+01  2.999e+01   3.129  0.00323 **
## ICU.Beds         3.106e-03  1.143e-03   2.718  0.00958 **
## Income          -1.629e-04  7.310e-05  -2.228  0.03142 *
## Sex.Ratio        3.932e+01  1.936e+01   2.031  0.04874 *
## Physicians      -1.306e-04  6.630e-05  -1.970  0.05562 .
## Hospitals       -3.225e-02  1.137e-02  -2.838  0.00703 **
## Health.Spending  2.445e-03  4.998e-04   4.891 1.59e-05 ***
## Age.0.25         1.041e+02  2.078e+01   5.009 1.09e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.8 on 41 degrees of freedom
## Multiple R-squared:  0.5665, Adjusted R-squared:  0.4819
## F-statistic: 6.698 on 8 and 41 DF,  p-value: 1.387e-05
```

```
plot(IP.best)
```


Residuals vs Fitted

lm(InfPop ~ Gini + ICU.Beds + Income + Sex.Ratio + Physicians + Hospitals + ...

Normal Q–Q

Standardized residuals

Theoretical Quantiles
lm(InfPop ~ Gini + ICU.Beds + Income + Sex.Ratio + Physicians + Hospitals + ...

Scale−Location

Fitted values
lm(InfPop ~ Gini + ICU.Beds + Income + Sex.Ratio + Physicians + Hospitals + ...

Residuals vs Leverage

lm(InfPop ~ Gini + ICU.Beds + Income + Sex.Ratio + Physicians + Hospitals + ...

Dead Population Regression Analysis

```r
# Defining our selected variables and determining optimal number
# based on Adjusted R-Squared, Cp, and BIC.

covid.DP <- subset(covid, select = -c(Population, Pop.Density, DeathsCurr, InfCurr, InfPop, DeathInf, S


reg.full.DP <- regsubsets(DeathPop ~., data = covid.DP)
reg.full.DP.sum <- summary(reg.full.DP)

which.max(reg.full.DP.sum$adjr2)
```
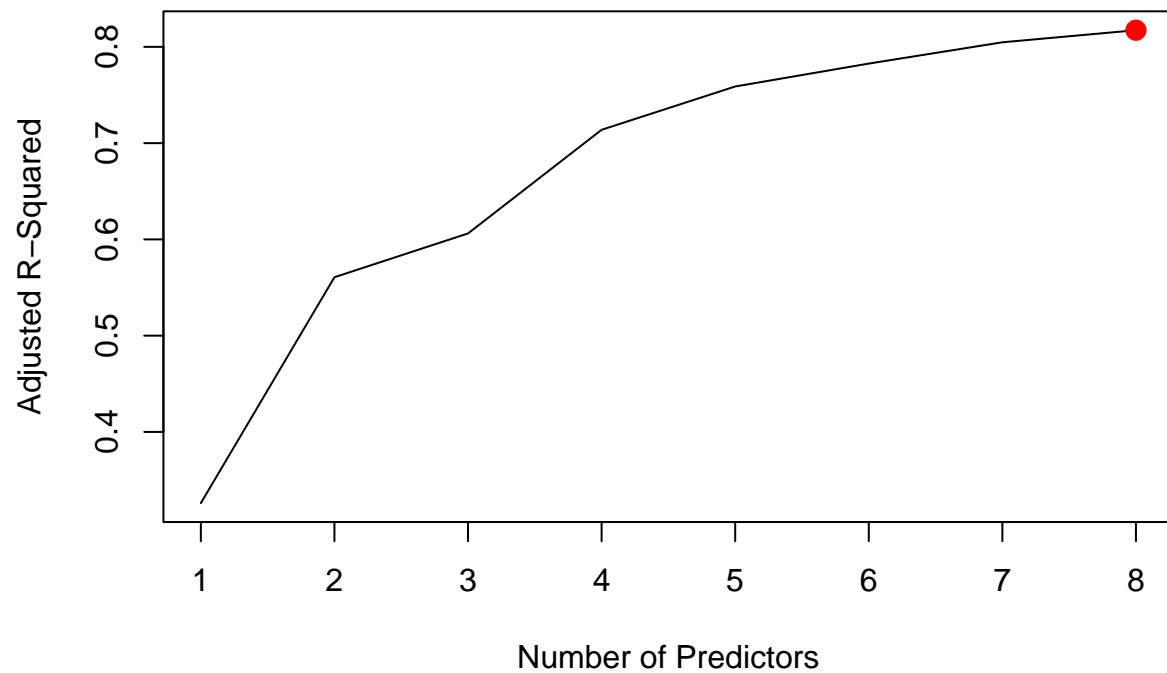
```
## [1] 8
```

```r
which.min(reg.full.DP.sum$cp)
```
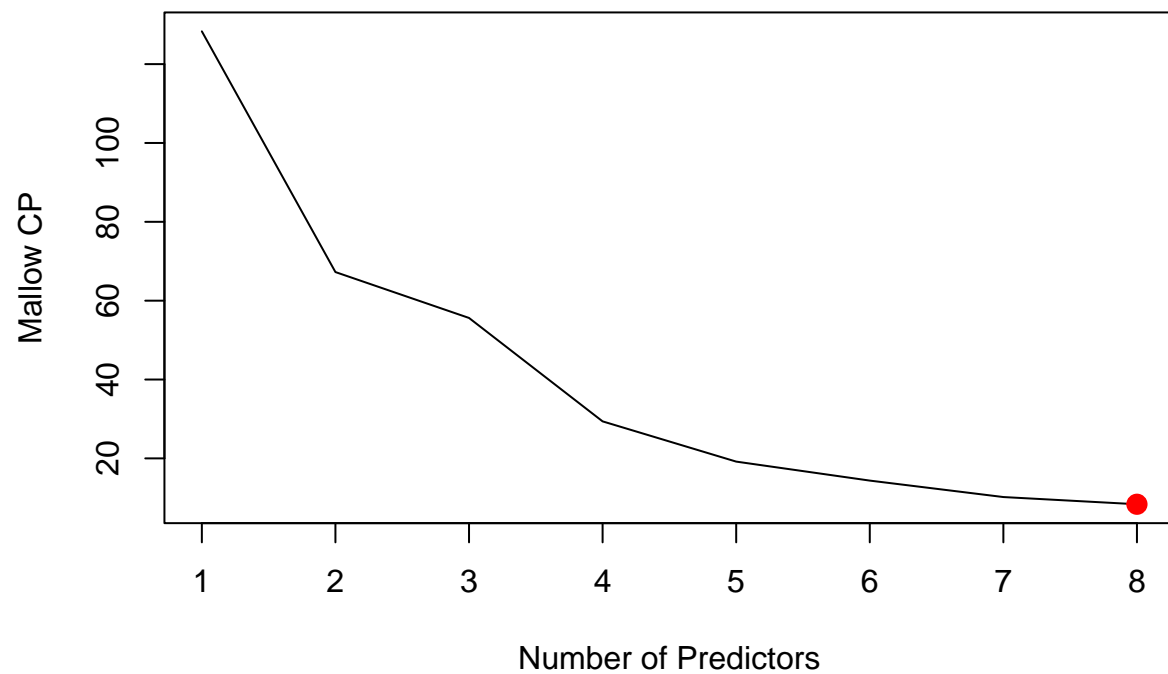
```
## [1] 8
```

```r
which.min(reg.full.DP.sum$bic)
```

```
## [1] 8
```

```
# Adjusted R-Squared against Number of Predictors
plot(reg.full.DP.sum$adjr2, xlab = "Number of Predictors", ylab = "Adjusted R-Squared", type = "l")
points(8, reg.full.DP.sum$adjr2[8], col = "red", cex = 2,
pch = 20)
```



```
# Cp against Number of Predictors
plot(reg.full.DP.sum$cp, xlab = "Number of Predictors", ylab = "Mallow CP", type = "l")
points(8, reg.full.DP.sum$cp[8], col = "red", cex = 2,
pch = 20)
```
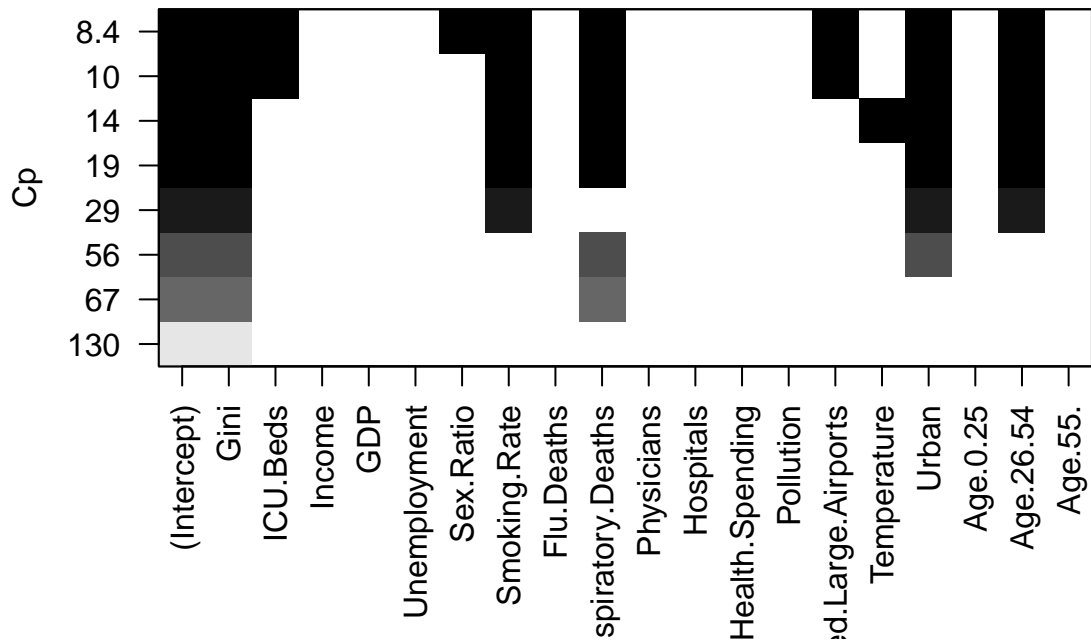
```
# BIC against Number of Predictors
plot(reg.full.DP.sum$bic, xlab = "Number of Predictors", ylab = "BIC", type = "l")
points(8, reg.full.DP.sum$bic[8], col = "red", cex = 2,
pch = 20)
```
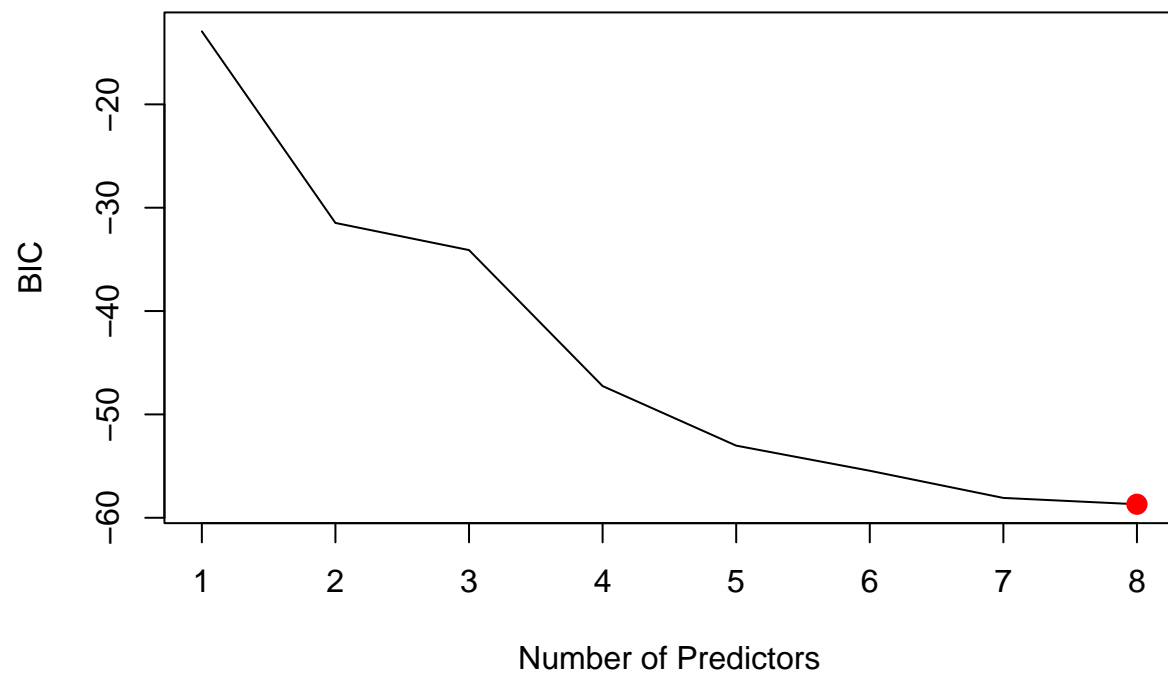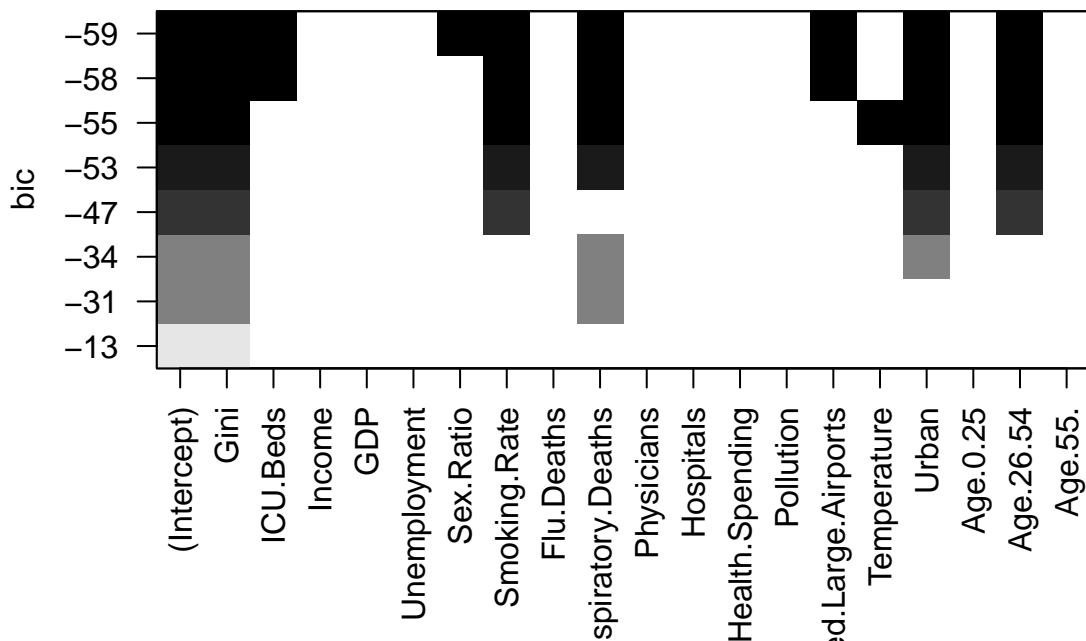
```r
# Best variable selections for Adjusted R-Squared
plot(reg.full.DP, scale = "adjr2")
```
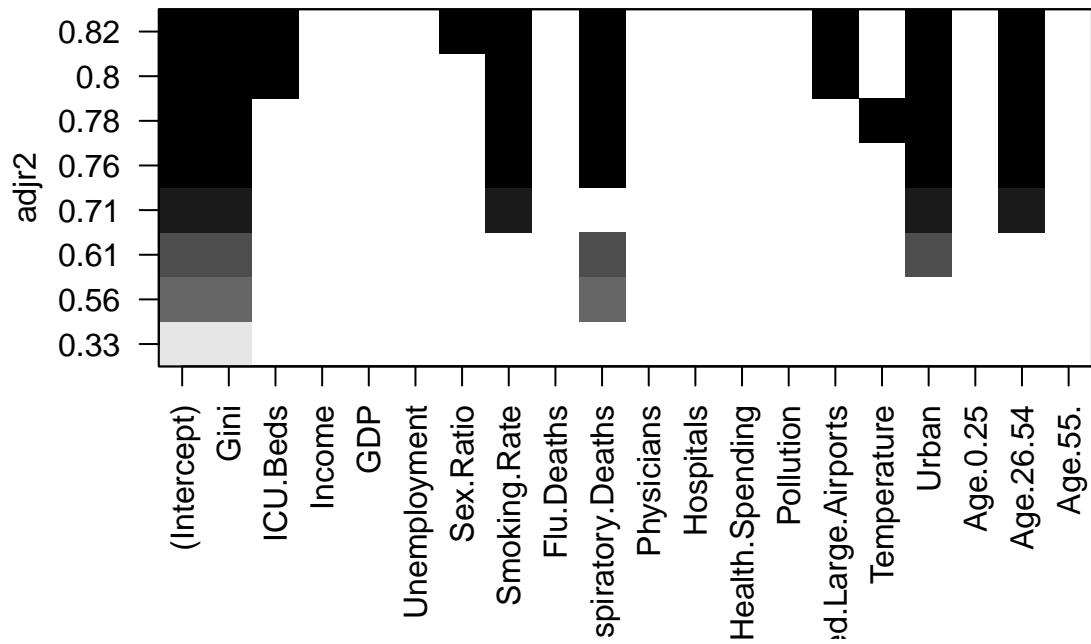
```
# Best variable selections for Cp
plot(reg.full.DP, scale = "Cp")
```

```
# Best variable selections for BIC
plot(reg.full.DP, scale = "bic")
```

```r
# Final Coefficients
coef(reg.full.DP, 8)
```

```
##       (Intercept)              Gini            ICU.Beds           Sex.Ratio
##     -7.010129e-01      2.026635e+00        2.898127e-05        4.405600e-01
##      Smoking.Rate Respiratory.Deaths Med.Large.Airports               Urban
##      1.011537e-02      2.049785e-03       -3.106051e-02        3.579588e-03
##         Age.26.54
##     -2.423072e+00
```

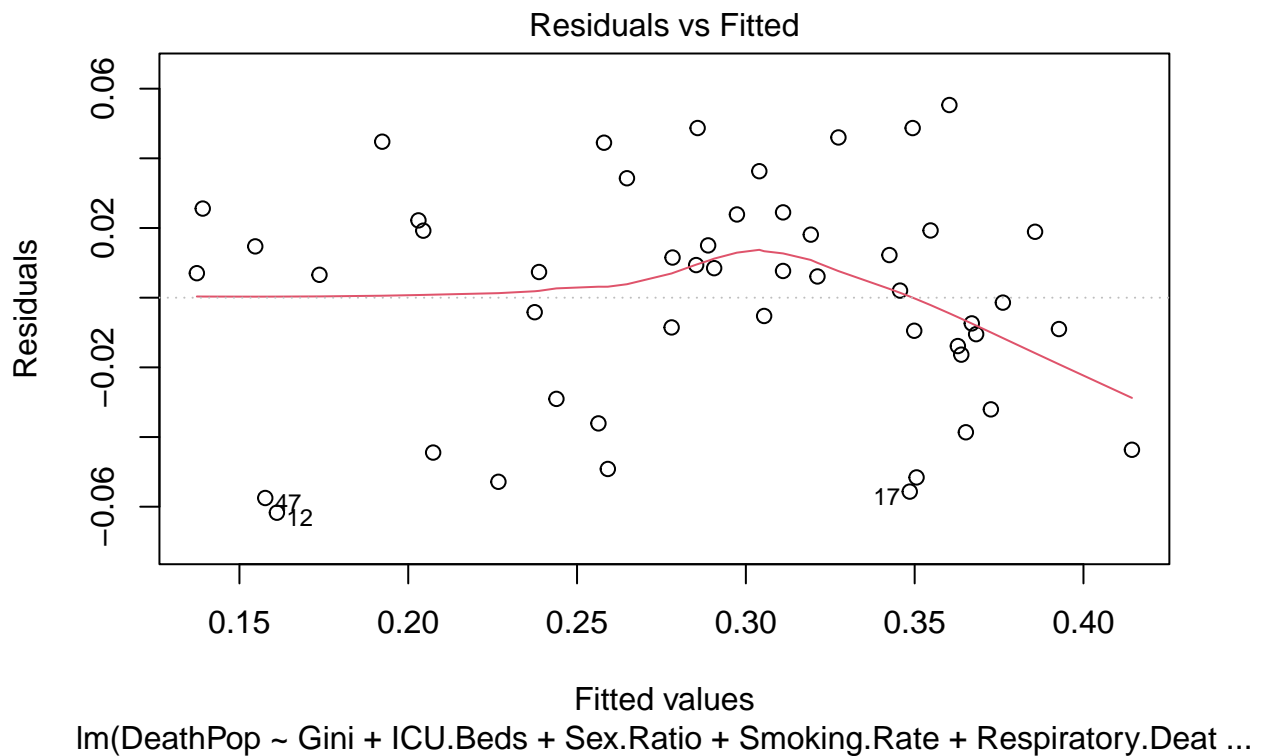## Final model for Deaths Per Population and Diagnostics

```r
DP.best <- lm(data = covid.DP, DeathPop ~ Gini + ICU.Beds + Sex.Ratio + Smoking.Rate + Respiratory.Deat
summary(DP.best)
```
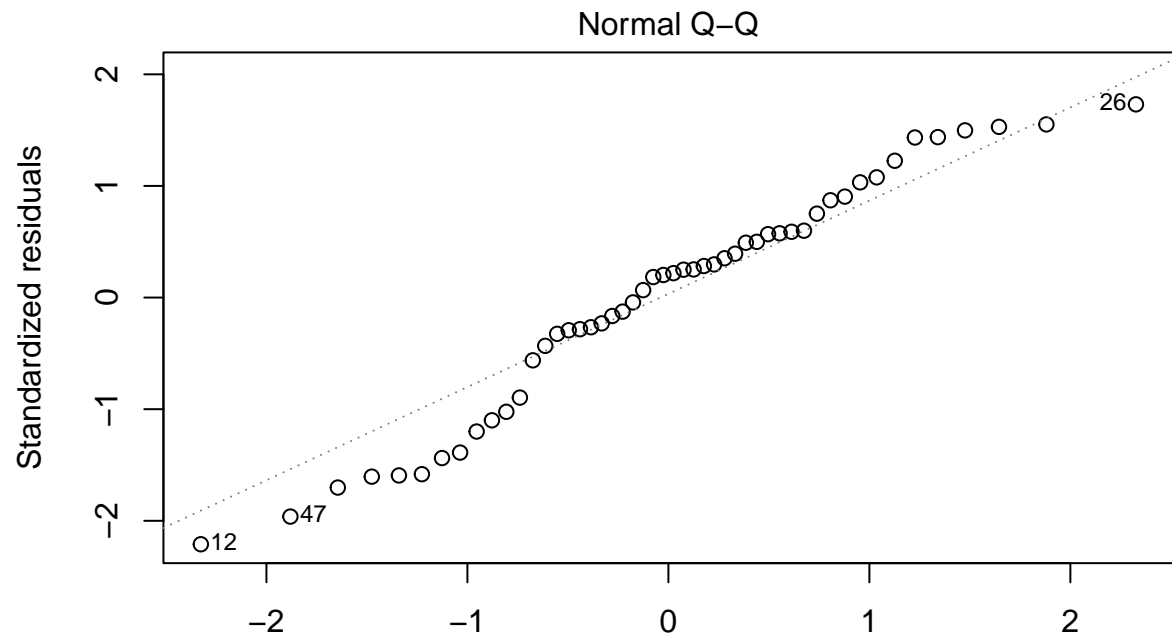
```
##
## Call:
## lm(formula = DeathPop ~ Gini + ICU.Beds + Sex.Ratio + Smoking.Rate +
##     Respiratory.Deaths + Med.Large.Airports + Urban + Age.26.54,
##     data = covid.DP)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -0.061717 -0.015734  0.006799  0.019285  0.055303
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -7.010e-01  3.724e-01  -1.882 0.066916 .
## Gini                2.027e+00  3.700e-01   5.478 2.39e-06 ***
## ICU.Beds            2.898e-05  9.802e-06   2.957 0.005137 **
## Sex.Ratio           4.406e-01  2.236e-01   1.970 0.055586 .
## Smoking.Rate        1.012e-02  2.482e-03   4.076 0.000205 ***
## Respiratory.Deaths  2.050e-03  8.217e-04   2.495 0.016735 *
## Med.Large.Airports -3.106e-02  8.224e-03  -3.777 0.000505 ***
## Urban               3.580e-03  5.268e-04   6.795 3.21e-08 ***
## Age.26.54          -2.423e+00  4.714e-01  -5.140 7.14e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03451 on 41 degrees of freedom
## Multiple R-squared:  0.8471, Adjusted R-squared:  0.8173
## F-statistic: 28.39 on 8 and 41 DF,  p-value: 2.251e-14
```
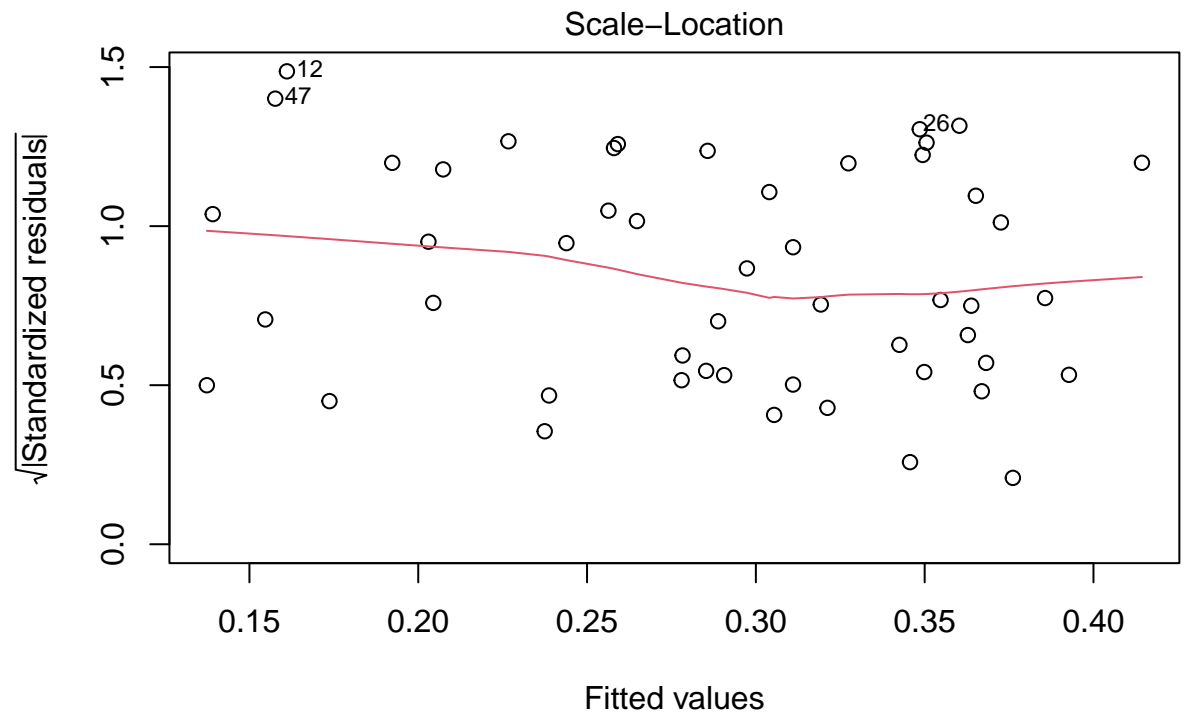
```
plot(DP.best)
```


Residuals vs Fitted

Fitted values
lm(DeathPop ~ Gini + ICU.Beds + Sex.Ratio + Smoking.Rate + Respiratory.Deat ...

Normal Q–Q

lm(DeathPop ~ Gini + ICU.Beds + Sex.Ratio + Smoking.Rate + Respiratory.Deat ...

Scale–Location

Fitted values
lm(DeathPop ~ Gini + ICU.Beds + Sex.Ratio + Smoking.Rate + Respiratory.Deat ...

Residuals vs Leverage

lm(DeathPop ~ Gini + ICU.Beds + Sex.Ratio + Smoking.Rate + Respiratory.Deat ...