Final Project

Erin & Tracy

2019/12/13

```
install.packages("readr")
```

```
## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.6'
## (as 'lib' is unspecified)
```

```
install.packages("ggplot2")
```

```
## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.6'
## (as 'lib' is unspecified)
```

```
install.packages("sandwich")
```

```
## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.6'
## (as 'lib' is unspecified)
```

```
install.packages("lmtest")
```

```
## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.6'
## (as 'lib' is unspecified)
```

```
install.packages("dplyr")
```

```
## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.6'
## (as 'lib' is unspecified)
```

```
install.packages("psych")
```

```
## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/3.6'
## (as 'lib' is unspecified)
```

```
library(readr)
library(ggplot2)
library(sandwich)
library(lmtest)
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
library(dplyr)
```

```
##
## 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
```

```
library(psych)
```

```
##
## Attaching package: 'psych'
```

```
## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha
```

```
Crime <- read_csv("Final project Data.csv")
```

```
## Parsed with column specification:
## cols(
##   Year = col_double(),
##   Population = col_double(),
##   ViolentCrimeRate = col_double(),
##   PropertyCrimeRate = col_double()
## )
```

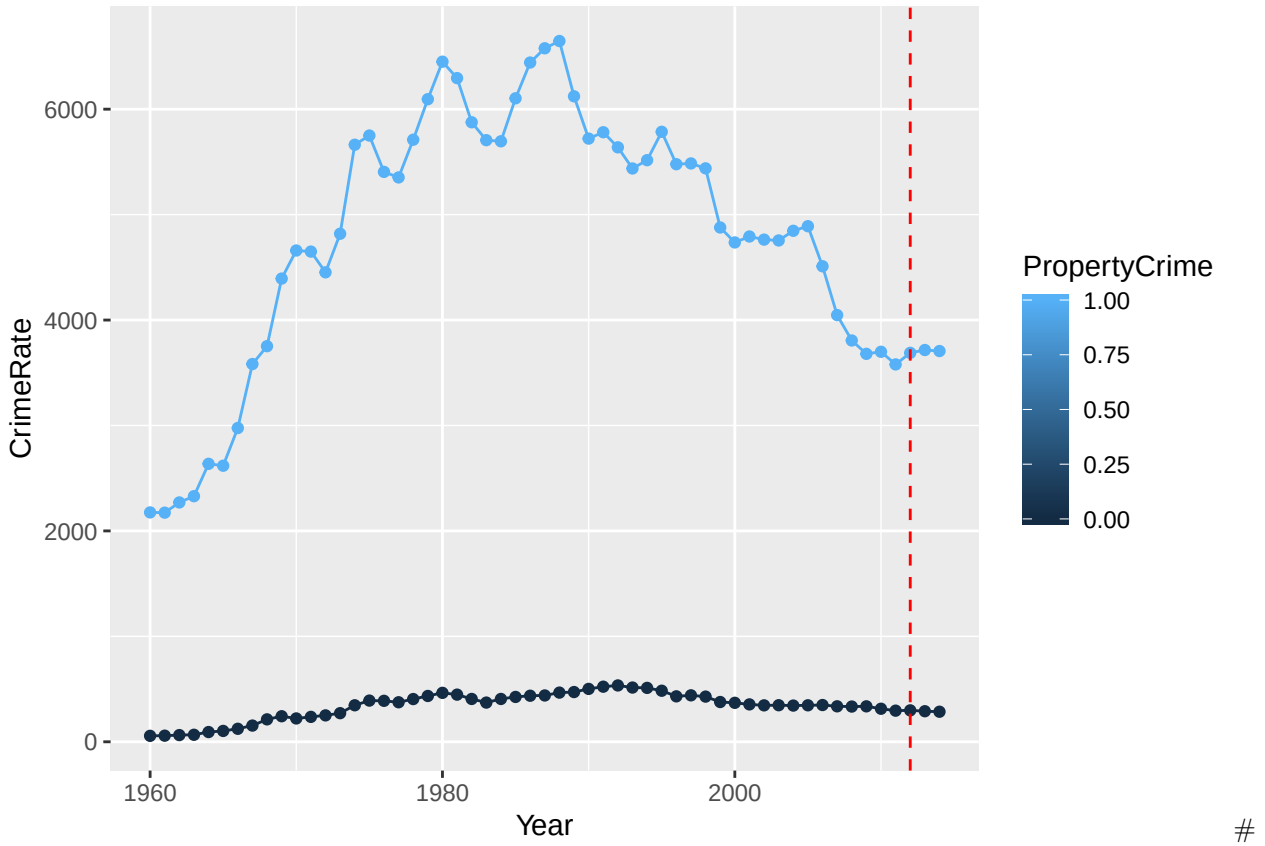
```
Crime2 <- read_csv("Final project Data2.csv")
```

```
## Parsed with column specification:
## cols(
##   Year = col_double(),
##   Population = col_double(),
##   CrimeRate = col_double(),
##   PropertyCrime = col_double()
## )
```

```
Crime$crimerate <- Crime$ViolentCrimeRate+Crime$PropertyCrimeRate
```

Graphs

```
ggplot(data=Crime2, aes(x=Year, y=CrimeRate, group=PropertyCrime, color=PropertyCrime)) + geom_line() +
```



#

Data Summary

```
summary(Crime)
```

```
##       Year        Population      ViolentCrimeRate PropertyCrimeRate
##  Min.   :1960   Min.   :2853214   Min.   : 56.6    Min.   :2173
##  1st Qu.:1974   1st Qu.:3510000   1st Qu.:284.8    1st Qu.:3694
##  Median :1989   Median :4761000   Median :346.3    Median :4792
##  Mean   :1989   Mean   :4923329   Mean   :334.3    Mean   :4658
##  3rd Qu.:2004   3rd Qu.:6169172   3rd Qu.:429.9    3rd Qu.:5701
##  Max.   :2018   Max.   :7535591   Max.   :534.5    Max.   :6647
##    crimerate
##  Min.   :2231
##  1st Qu.:3989
##  Median :5108
##  Mean   :4993
##  3rd Qu.:6109
##  Max.   :7113
```

```
describe(Crime)
```

```
##                    vars  n       mean         sd    median    trimmed        mad
## Year                  1 59    1989.00      17.18    1989.0    1989.00      22.24
## Population            2 59 4923328.86 1447243.46 4761000.0 4883393.57 1945171.20
## ViolentCrimeRate      3 59     334.26     125.32     346.3     342.84     117.27
## PropertyCrimeRate     4 59    4658.33    1249.70    4792.5    4711.48    1466.29
## crimerate             5 59    4992.59    1360.50    5108.4    5060.84    1624.63
##                       min       max     range  skew kurtosis        se
## Year                1960.0    2018.0      58.0  0.00    -1.26      2.24
```

```
## Population        2853214.0 7535591.0 4682377.0  0.17    -1.36 188415.05
## ViolentCrimeRate       56.6     534.5     477.9 -0.67    -0.23     16.32
## PropertyCrimeRate    2173.2    6646.6    4473.4 -0.37    -0.97    162.70
## crimerate            2231.3    7113.0    4881.7 -0.41    -0.91    177.12
```

General Population Model:

$$CrimeRate_t = \beta_0 + \beta_1 Population_t + \beta_2 Legalization_t + \beta_3 Crime_t + u_t$$

```
Crime$Legalization <- ifelse(Crime$Year > 2012, 1, 0)
```

Violent Crimes

Population Model:

$$ViolentCrimeRate_t = \beta_0 + \beta_1 Population_t + \beta_2 Legalization_t + \beta_3 Year + \beta_4 ViolentCrimeRate_{t-1} + u_t$$

```
model1 <- lm(ViolentCrimeRate ~ Population + lag(ViolentCrimeRate,1) + Year + Legalization, data = Crime
summary(model1)
```

```
##
## Call:
## lm(formula = ViolentCrimeRate ~ Population + lag(ViolentCrimeRate,
##     1) + Year + Legalization, data = Crime)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -42.685 -10.277   0.259  12.704  44.286
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)               -1.436e+04  4.721e+03  -3.042  0.00365 **
## Population                -9.207e-05  2.774e-05  -3.319  0.00164 **
## lag(ViolentCrimeRate, 1)   8.979e-01  3.621e-02  24.797  < 2e-16 ***
## Year                       7.467e+00  2.446e+00   3.052  0.00355 **
## Legalization               1.088e+01  1.162e+01   0.936  0.35338
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 20.34 on 53 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.9736, Adjusted R-squared:  0.9717
## F-statistic: 489.6 on 4 and 53 DF,  p-value: < 2.2e-16
```

```
coeftest(model1, vcov = vcovHC(model1, type="HC1"))
```

```
##
## t test of coefficients:
##
##                              Estimate  Std. Error t value  Pr(>|t|)
## (Intercept)               -1.4362e+04  4.7814e+03 -3.0038  0.004064 **
## Population                -9.2071e-05  2.8590e-05 -3.2204  0.002188 **
## lag(ViolentCrimeRate, 1)   8.9793e-01  3.6587e-02 24.5423 < 2.2e-16 ***
```

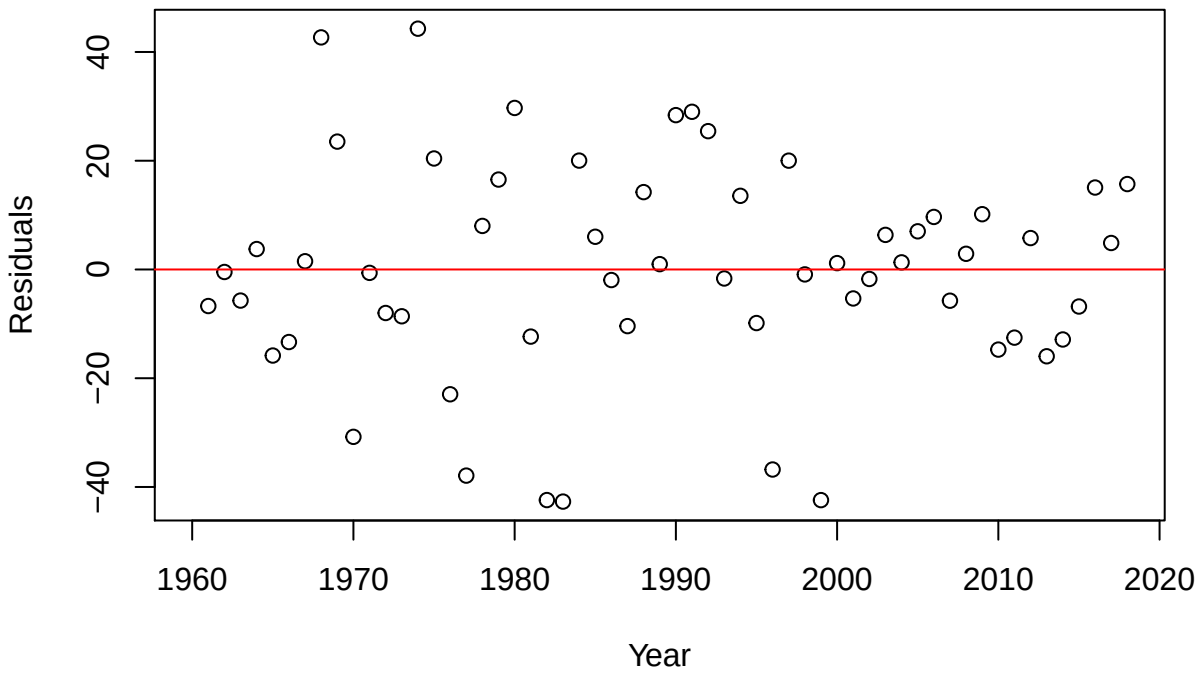
```
## Year                        7.4673e+00  2.4791e+00  3.0121  0.003971 **
## Legalization                1.0879e+01  7.3894e+00  1.4722  0.146867
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

##Graph

```
plot(Crime$Year,Crime$ViolentCrimeRate)
abline(v=2012, col="blue")
```

##Test for Heteroscedasticity and autocorrelation

```
Crime$e_1 <- c(NA, residuals(model1))
plot(Crime$Year, Crime$e_1, main = " ", xlab="Year", ylab="Residuals")
abline(lm(Crime$e_1 ~ Crime$Year), col="red")
```




```
bgtest(model1, order=1, type="F")
```

```
##
##  Breusch-Godfrey test for serial correlation of order up to 1
##
## data:  model1
## LM test = 2.2126, df1 = 1, df2 = 52, p-value = 0.1429
```

```
bptest(model1)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  model1
## BP = 8.4154, df = 4, p-value = 0.07749
```

Property Crimes

Population Model:

$PropertyCrimeRate_t = \beta_0 + \beta_1 Population_t + \beta_2 Legalization_t + \beta_3 Year + \beta_4 PropertyCrimeRate_{t-1} + u_t$

```
model2 <- lm(PropertyCrimeRate ~ Population + lag(PropertyCrimeRate,1) + Year + Legalization, data = Cri
summary(model2)
```

```
##
## Call:
## lm(formula = PropertyCrimeRate ~ Population + lag(PropertyCrimeRate,
##     1) + Year + Legalization, data = Crime)
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -520.82 -123.05   -7.43  143.73  568.39
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)               -1.517e+05  6.631e+04  -2.287   0.0262 *
## Population                -1.013e-03  4.011e-04  -2.525   0.0146 *
## lag(PropertyCrimeRate, 1)  8.710e-01  4.748e-02  18.343   <2e-16 ***
## Year                       7.908e+01  3.442e+01   2.298   0.0256 *
## Legalization              -3.182e+01  1.416e+02  -0.225   0.8230
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 249.2 on 53 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.961,  Adjusted R-squared:  0.958
## F-statistic: 326.3 on 4 and 53 DF,  p-value: < 2.2e-16
```

```
coeftest(model2, vcov = vcovHC(model1, type="HC1"))
```

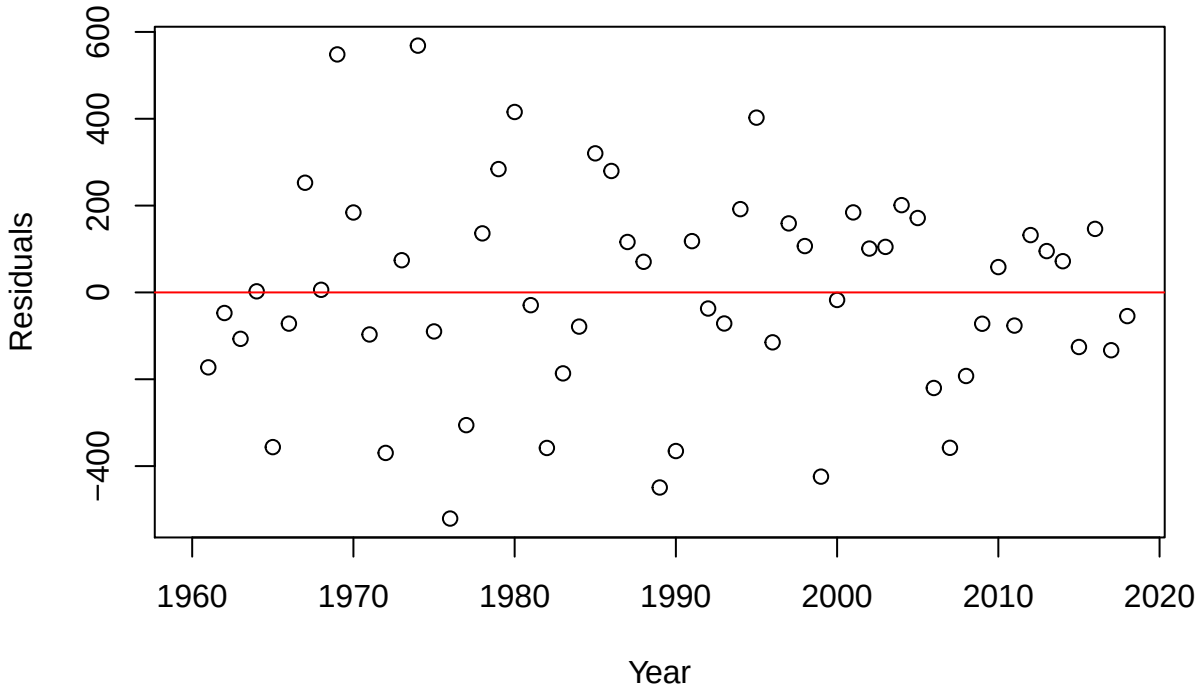
```
##
## t test of coefficients:
##
##                 Estimate  Std. Error  t value  Pr(>|t|)
## (Intercept)  -1.5168e+05  4.7814e+03 -31.7232 < 2.2e-16 ***
## Population   -1.0127e-03  2.8590e-05 -35.4227 < 2.2e-16 ***
## Year          7.9077e+01  2.4791e+00  31.8971 < 2.2e-16 ***
## Legalization -3.1824e+01  7.3894e+00  -4.3067 7.196e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

##Graph

```
plot(Crime$Year,Crime$PropertyCrimeRate)
abline(v=2012, col="blue")
```

##Test for Heteroscedasticity and autocorrelation

```
Crime$e_2 <- c(NA, residuals(model2))
plot(Crime$Year, Crime$e_2, main = " ", xlab="Year", ylab="Residuals")
abline(lm(Crime$e_2 ~ Crime$Year), col="red")
```




```
bgtest(model2, order=1, type="F")
```

```
##
##  Breusch-Godfrey test for serial correlation of order up to 1
```

```
##
## data:  model2
## LM test = 4.5131, df1 = 1, df2 = 52, p-value = 0.03841
```

```
bptest(model2)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  model2
## BP = 8.0039, df = 4, p-value = 0.09143
```

Total Crimes Rate

Population Model:

$$CrimeRate_t = \beta_0 + \beta_1 Population_t + \beta_2 Legalization_t + \beta_3 Year + \beta_4 CrimeRate_{t-1} + u_t$$

```
model3 <- lm(crimerate ~ Population + lag(crimerate,1) + Year + Legalization, data = Crime)
summary(model3)
```

```
##
## Call:
## lm(formula = crimerate ~ Population + lag(crimerate, 1) + Year +
##     Legalization, data = Crime)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -543.12 -131.90   -1.58  150.89  613.79
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -1.639e+05  6.852e+04  -2.392   0.0203 *
## Population        -1.091e-03  4.132e-04  -2.641   0.0108 *
## lag(crimerate, 1)  8.748e-01  4.508e-02  19.407   <2e-16 ***
## Year               8.544e+01  3.556e+01   2.403   0.0198 *
## Legalization      -1.947e+01  1.475e+02  -0.132   0.8955
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 259.1 on 53 degrees of freedom
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.9643, Adjusted R-squared:  0.9616
## F-statistic: 357.6 on 4 and 53 DF,  p-value: < 2.2e-16
```

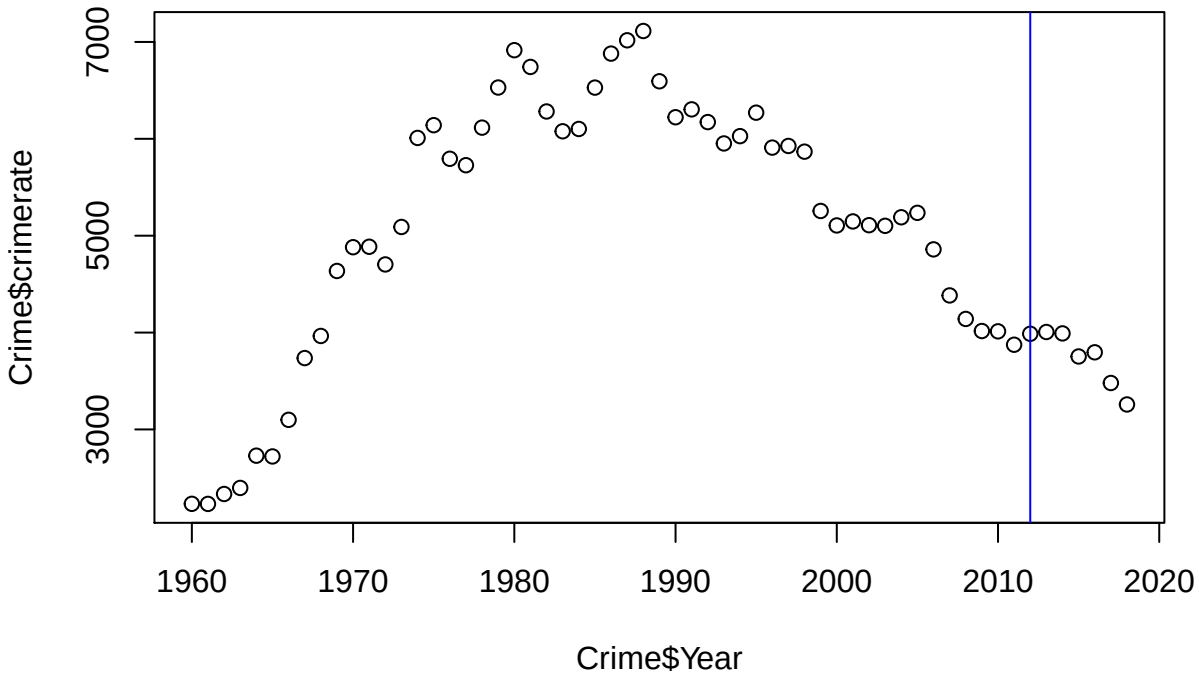
```
coeftest(model3, vcov = vcovHC(model1, type="HC1"))
```

```
##
## t test of coefficients:
##
##                 Estimate  Std. Error  t value Pr(>|t|)
## (Intercept)  -1.6393e+05  4.7814e+03 -34.2843  < 2e-16 ***
## Population   -1.0913e-03  2.8590e-05 -38.1710  < 2e-16 ***
## Year          8.5441e+01  2.4791e+00  34.4644  < 2e-16 ***
```

```
## Legalization -1.9469e+01  7.3894e+00  -2.6347  0.01102 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

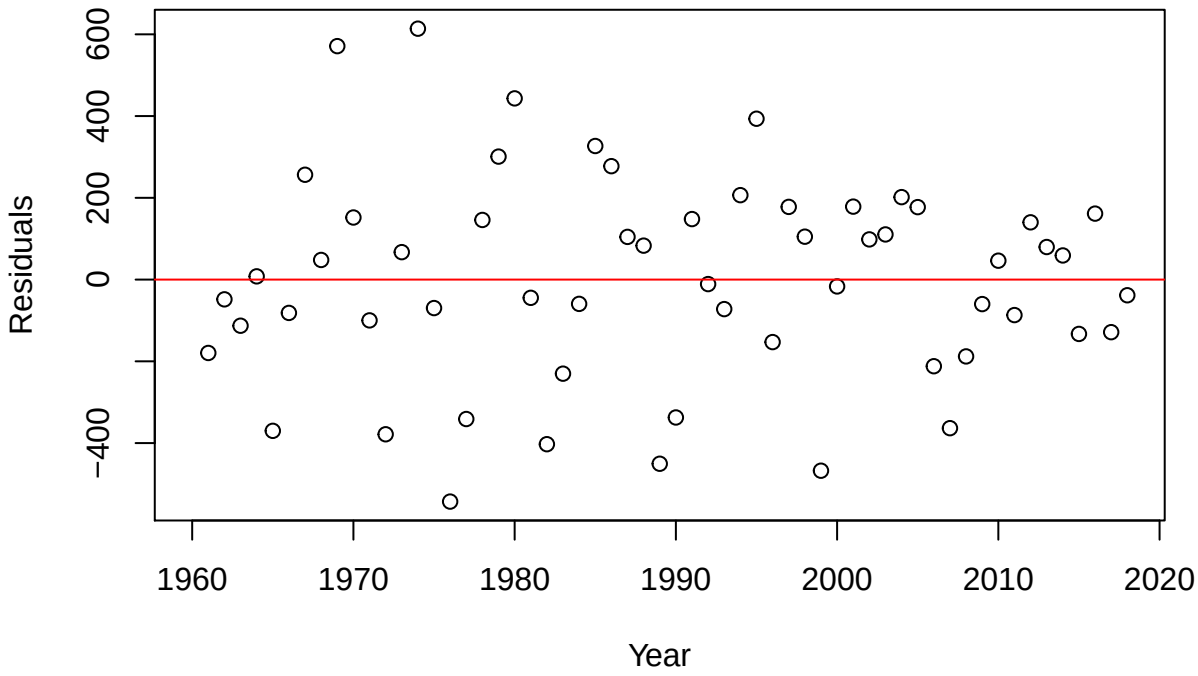
##Graph

```
plot(Crime$Year,Crime$crimerate)
abline(v=2012, col="blue")
```



##Test for Heteroscedasticity and autocorrelation

```
Crime$e_3 <- c(NA, residuals(model3))
plot(Crime$Year, Crime$e_3, main = " ", xlab="Year", ylab="Residuals")
abline(lm(Crime$e_3 ~ Crime$Year), col="red")
```




```
bgtest(model3, order=1, type="F")
```

```
##
##  Breusch-Godfrey test for serial correlation of order up to 1
##
## data:  model3
## LM test = 4.1099, df1 = 1, df2 = 52, p-value = 0.04777
```

```
bptest(model3)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  model3
## BP = 8.2584, df = 4, p-value = 0.08256
```