STA 137 - Final Project

Falak Shah (SID: 914663151)

12/5/2019

INTRODUCTION: The dataset consists of data on monthly handgun sales and firearm related deaths in California from 1980-1998.This report aims to analyze the presence of trends and/or seasonal components present in each of the two variables and accordingly fit a suitable model for each variable. Furthermore,this report also aims to construct a model relating the firearm guns and handgun sales to determine the presence of any relationship between the two variables. It is a time series data because both variables are measured monthly across various time points,i.e. from 1980-1998.It is important to analyze this data because the data points taken over time might have an internal structure such as autocorrelation,trend or seasonal variation which need to be accounted for when selecting the correct model,which could be used to understand the past as well as predict the future.

1 ANALYSIS AND RESULTS [GUN SALES DATA]:

Observing figure 1.1, we can see that the gun sales varies across the different months of measurement. There are large trends in the data. Majorly, there are mixed trends of increase and decrease in Sales, which is followed by a big decreasing trend towards the end of the measurement period. To analyze further details about the data, which cannot be viewed from Figure 1, we move onto observing the acf and pacf plots of the original, untransformed Gun Sales series.

Referring to figure 1.2, we can observe from the acf plot that the ACF ordinates are large and decay slowly. Furthermore,the PACF has one large peak at lag 1. Therefore, the time series is not stationary. As it is non stationary and the acf depicts slow linear decay, differencing is required and a remedy might be to look at the log transformed series.

Looking at figure 1.3, we can observe that log differncing has removed the trend in the series. There still seems to exist an apparent trend component, which is confirmed from the acf and pacf plots of the first differnced log series in figure 1.4. We can observe that acf decays slowly at lags multiple of 12 and the pacf has one large peak at the first multiple of 12, lag 12. Therefore, this suggests lag 12 differncing in order to remove the seasonality.

Figure 1.5 depicts the lag 12 differnced time series, the mean looks centered around zero and the series looks smoother compared to the one before. Furthermore,the series looks stationary after the aforementioned transformations. Close analysis of the acf and pacf plots in Figure 1.6 confirms these observations and helps us determine the candidate models. On first observation, the ACF seems to cut off after lag 2 and PACF seems to tail off, whereas the acf at lags multiple of 12 also seems to cut off after the second multiple of 12, lag 24, whereas PACF at lags multiple of 12 seems to tail off. Thus,first model suggested is SARIMA(0,1,2)x(0,1,2)[12].

Another way of interpreting the acf and pacf plots is that the ACF cuts off after lag 2,PACF tails off, whereas the ACF at lags multiple of 12 seems to tail off and the PACF at lags multiple of 12 seems to cut off after the third lag. Thus, second candidate model is SARIMA(0,1,2)x(3,1,0)[12]

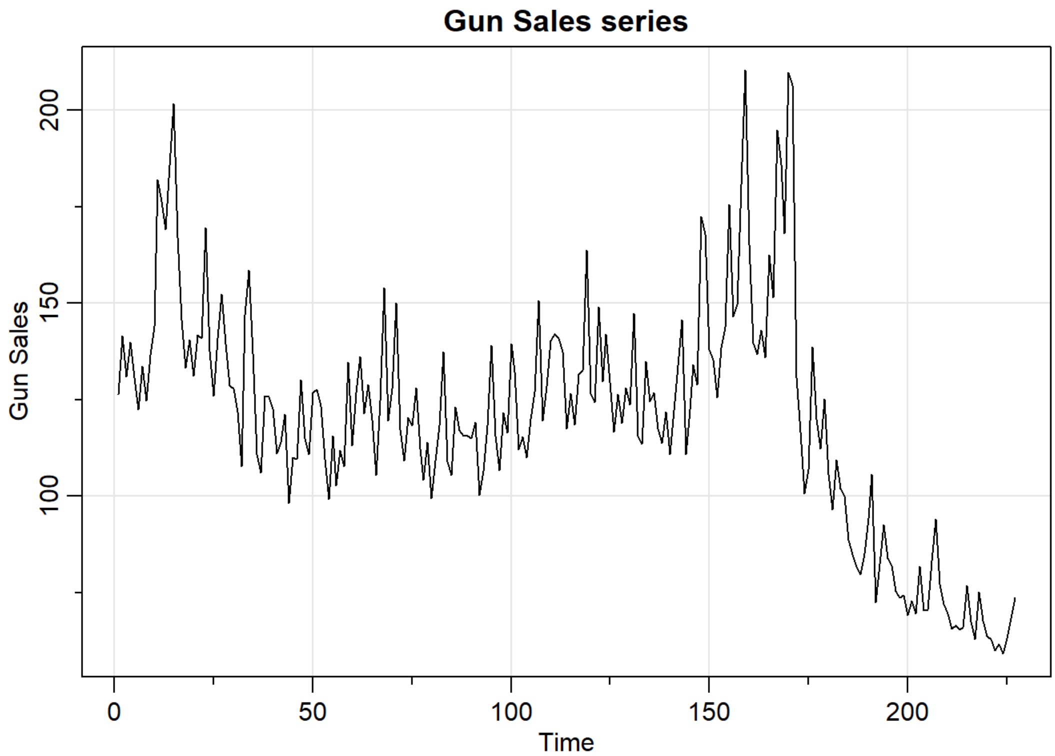


Figure 1.1: Plots for Gun Sales data

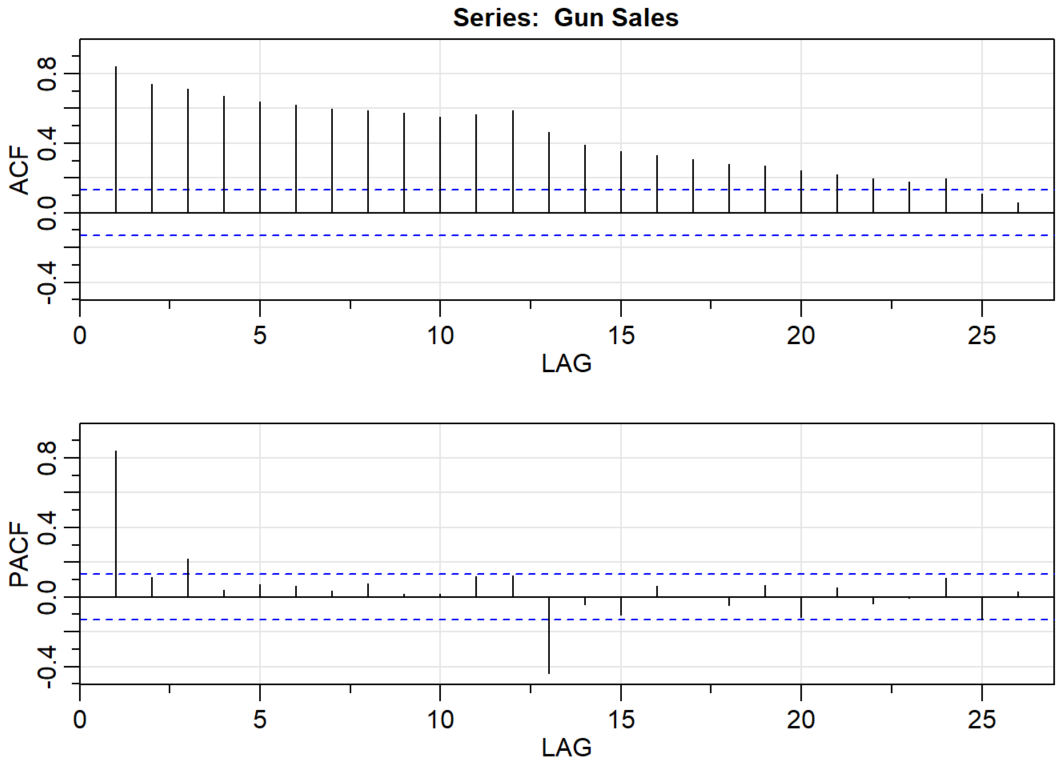


Figure 1.2: Plots for Gun Sales data

```
##        ACF  PACF
##  [1,] 0.84  0.84
##  [2,] 0.74  0.11
##  [3,] 0.71  0.22
##  [4,] 0.67  0.04
##  [5,] 0.64  0.07
##  [6,] 0.62  0.06
##  [7,] 0.60  0.03
##  [8,] 0.59  0.08
##  [9,] 0.57  0.02
## [10,] 0.55  0.02
## [11,] 0.56  0.12
## [12,] 0.59  0.12
## [13,] 0.46 -0.44
## [14,] 0.39 -0.04
## [15,] 0.35 -0.10
## [16,] 0.33  0.06
## [17,] 0.31  0.00
## [18,] 0.28 -0.05
## [19,] 0.27  0.07
## [20,] 0.24 -0.12
## [21,] 0.22  0.05
## [22,] 0.20 -0.04
## [23,] 0.18  0.00
## [24,] 0.19  0.11
## [25,] 0.11 -0.13
## [26,] 0.06  0.03
```

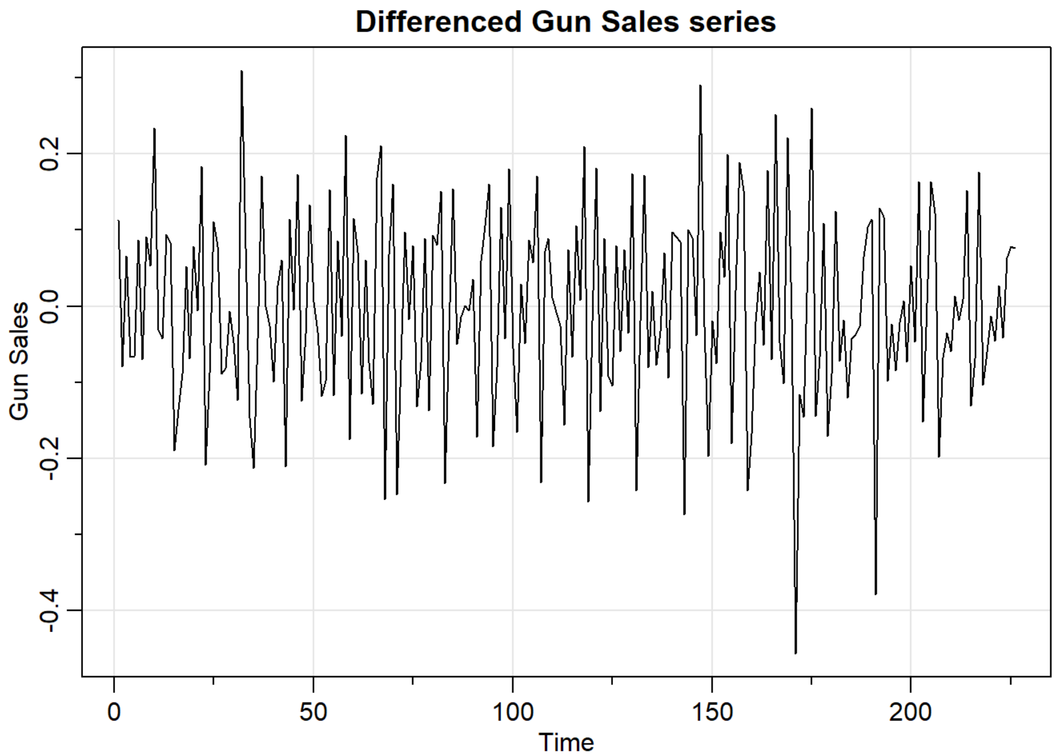


Figure 1.3: Plots for Gun Sales data

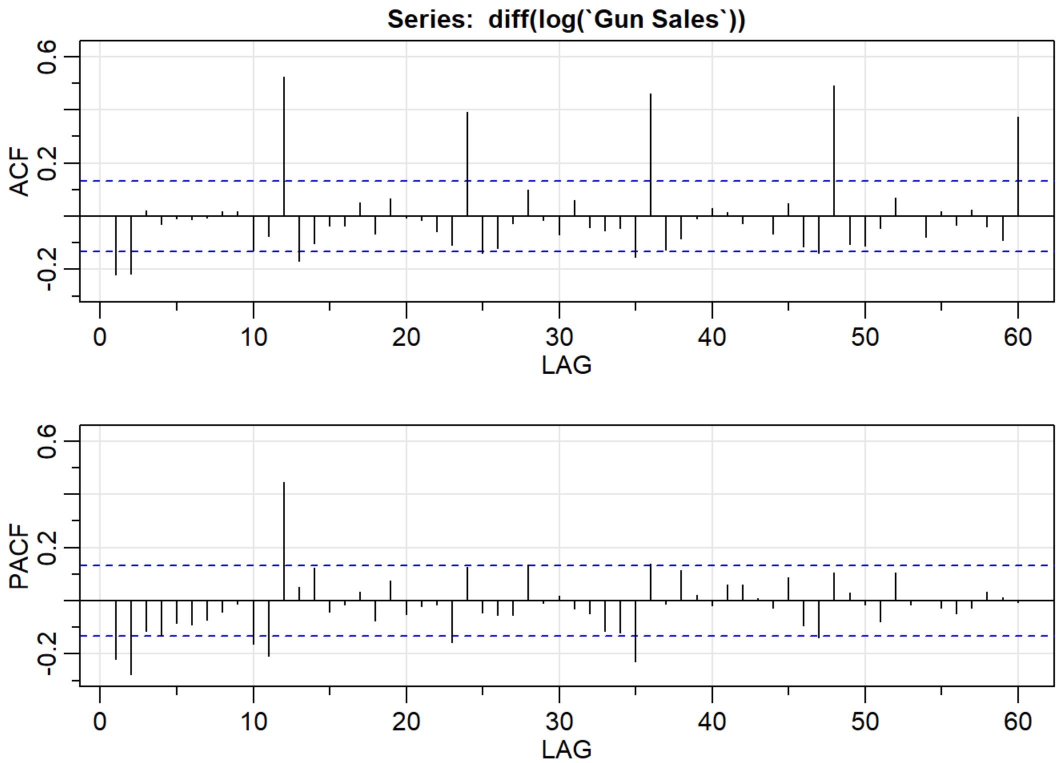


Figure 1.4: Plots for Gun Sales data

```
##         ACF  PACF
##  [1,] -0.22 -0.22
##  [2,] -0.22 -0.28
##  [3,]  0.02 -0.11
##  [4,] -0.03 -0.13
##  [5,] -0.01 -0.08
##  [6,] -0.01 -0.09
##  [7,] -0.01 -0.07
##  [8,]  0.02 -0.04
##  [9,]  0.02 -0.01
## [10,] -0.13 -0.16
## [11,] -0.08 -0.21
## [12,]  0.52  0.45
## [13,] -0.17  0.05
## [14,] -0.10  0.12
## [15,] -0.04 -0.04
## [16,] -0.04 -0.02
## [17,]  0.05  0.03
## [18,] -0.07 -0.08
## [19,]  0.07  0.08
## [20,] -0.01 -0.05
## [21,] -0.01 -0.02
## [22,] -0.06 -0.02
## [23,] -0.11 -0.16
## [24,]  0.39  0.13
## [25,] -0.14 -0.05
## [26,] -0.12 -0.05
## [27,] -0.03 -0.05
## [28,]  0.10  0.14
## [29,] -0.01 -0.01
## [30,] -0.07  0.02
## [31,]  0.06 -0.03
## [32,] -0.04 -0.05
## [33,] -0.05 -0.11
## [34,] -0.05 -0.12
## [35,] -0.15 -0.23
## [36,]  0.46  0.14
## [37,] -0.13 -0.01
## [38,] -0.08  0.12
## [39,] -0.01  0.02
## [40,]  0.03 -0.02
## [41,]  0.02  0.06
## [42,] -0.03  0.06
## [43,]  0.00  0.01
## [44,] -0.07 -0.03
## [45,]  0.05  0.09
## [46,] -0.12 -0.09
## [47,] -0.14 -0.14
## [48,]  0.49  0.11
## [49,] -0.10  0.03
## [50,] -0.11 -0.02
## [51,] -0.05 -0.08
## [52,]  0.07  0.10
## [53,]  0.00 -0.01
## [54,] -0.08  0.00
## [55,]  0.02 -0.03
## [56,] -0.03 -0.05
## [57,]  0.02 -0.03
## [58,] -0.04  0.03
## [59,] -0.09  0.01
## [60,]  0.38 -0.01
```

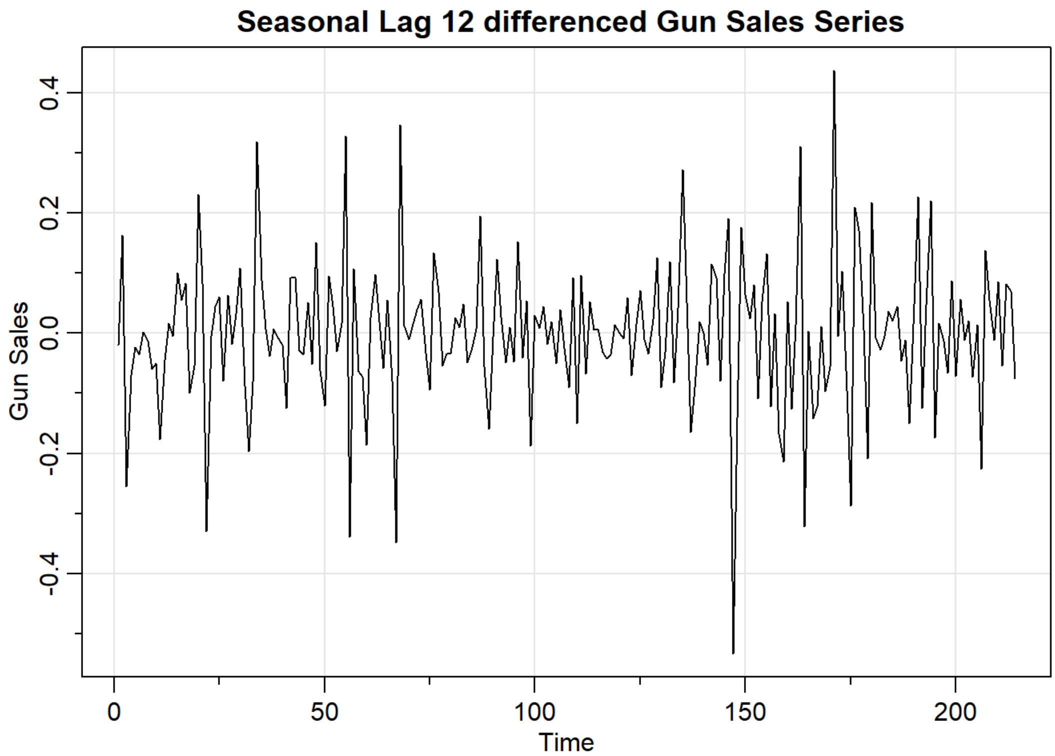


Figure 1.5: Plots for Gun Sales data

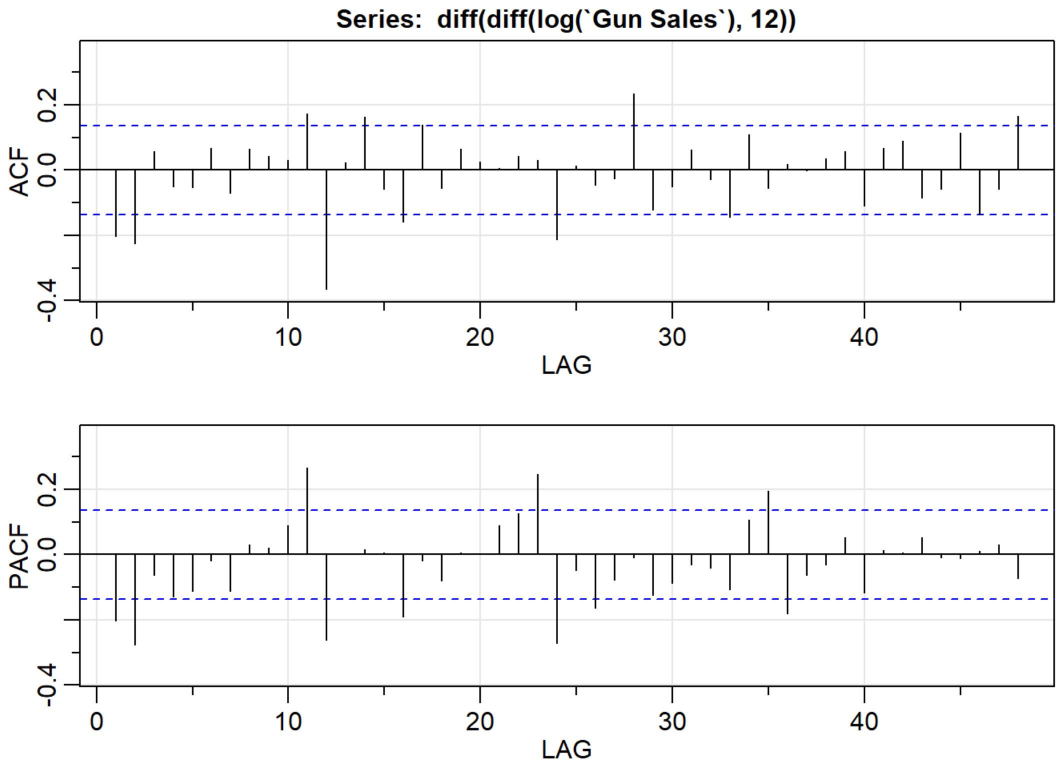
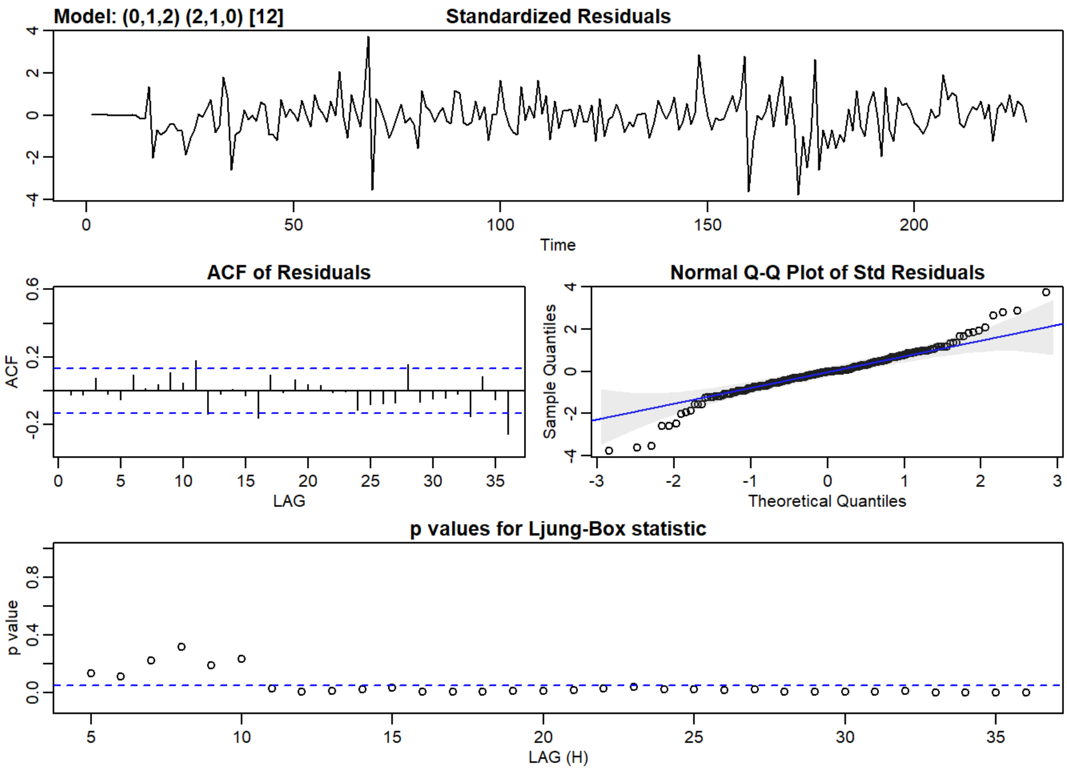


Figure 1.6: Plots for Gun Sales data

```
##          ACF  PACF
##  [1,] -0.20 -0.20
##  [2,] -0.22 -0.28
##  [3,]  0.06 -0.06
##  [4,] -0.05 -0.13
##  [5,] -0.05 -0.11
##  [6,]  0.07 -0.02
##  [7,] -0.07 -0.11
##  [8,]  0.06  0.03
##  [9,]  0.04  0.02
## [10,]  0.03  0.09
## [11,]  0.17  0.27
## [12,] -0.37 -0.26
## [13,]  0.02  0.00
## [14,]  0.16  0.02
## [15,] -0.06  0.01
## [16,] -0.16 -0.19
## [17,]  0.14 -0.02
## [18,] -0.06 -0.08
## [19,]  0.07  0.01
## [20,]  0.03  0.00
## [21,]  0.01  0.09
## [22,]  0.04  0.13
## [23,]  0.03  0.25
## [24,] -0.21 -0.27
## [25,]  0.01 -0.05
## [26,] -0.05 -0.16
## [27,] -0.03 -0.08
## [28,]  0.24 -0.01
## [29,] -0.12 -0.12
## [30,] -0.05 -0.09
## [31,]  0.06 -0.03
## [32,] -0.03 -0.04
## [33,] -0.14 -0.11
## [34,]  0.11  0.11
## [35,] -0.06  0.20
## [36,]  0.02 -0.18
## [37,]  0.00 -0.06
## [38,]  0.04 -0.03
## [39,]  0.06  0.05
## [40,] -0.11 -0.12
## [41,]  0.07  0.01
## [42,]  0.09  0.01
## [43,] -0.08  0.05
## [44,] -0.06 -0.01
## [45,]  0.11 -0.01
## [46,] -0.14  0.01
## [47,] -0.06  0.03
## [48,]  0.17 -0.07
```

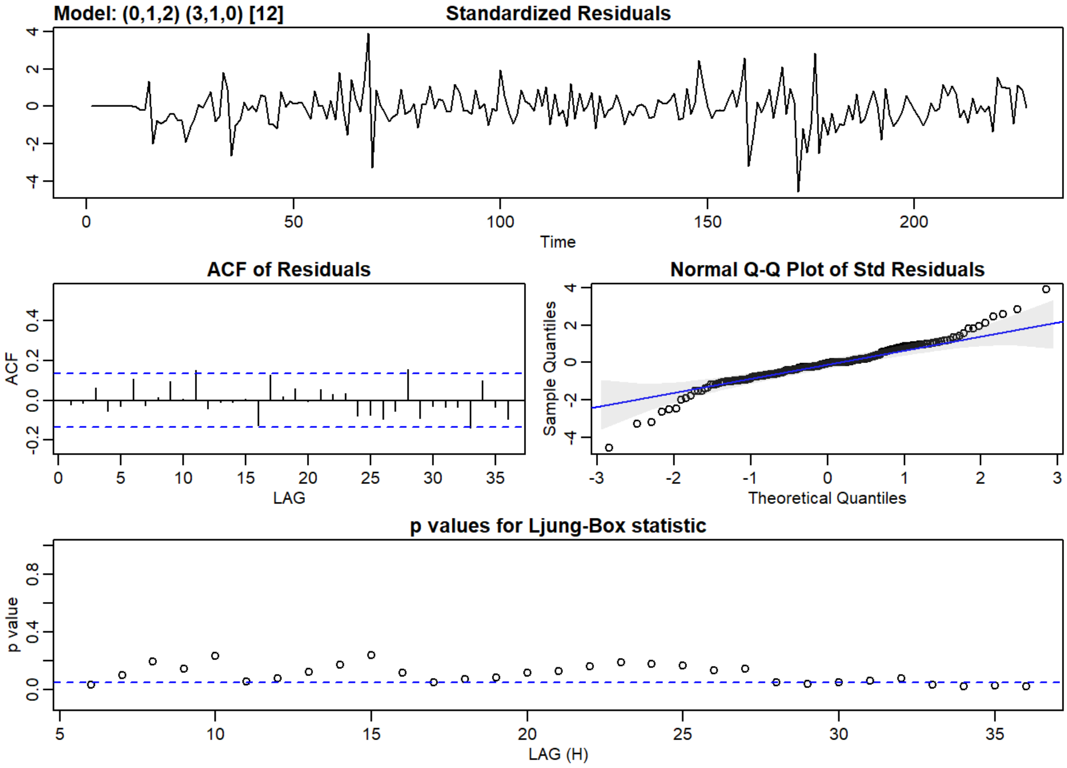
```
# MODEL 1
t1 = sarima(log(`Gun Sales`),p=0,d=1,q=2,P=2,D=1,Q=0,S=12)
```

```
## initial  value -2.100429
## iter   2 value -2.338615
## iter   3 value -2.365540
## iter   4 value -2.375811
## iter   5 value -2.376173
## iter   6 value -2.376198
## iter   7 value -2.376209
## iter   8 value -2.376209
## iter   8 value -2.376209
## iter   8 value -2.376209
## final  value -2.376209
## converged
## initial  value -2.347249
## iter   2 value -2.347511
## iter   3 value -2.348096
## iter   4 value -2.348104
## iter   5 value -2.348131
## iter   5 value -2.348131
## iter   5 value -2.348131
## final  value -2.348131
## converged
```




```
# MODEL 2
t2 = sarima(log(`Gun Sales`),p=0,d=1,q=2,P=3,D=1,Q=0,S=12)
```

```
## initial  value -2.102196
## iter   2 value -2.318815
## iter   3 value -2.363324
## iter   4 value -2.419340
## iter   5 value -2.427416
## iter   6 value -2.427837
## iter   7 value -2.427862
## iter   8 value -2.427912
## iter   9 value -2.427913
## iter   9 value -2.427913
## iter   9 value -2.427913
## final  value -2.427913
## converged
## initial  value -2.404253
## iter   2 value -2.405129
## iter   3 value -2.405718
## iter   4 value -2.405773
## iter   5 value -2.405806
## iter   6 value -2.405806
## iter   6 value -2.405806
## iter   6 value -2.405806
## final  value -2.405806
## converged
```




```
# MODEL SELECTION
t1$ttable
```

```
##       Estimate     SE t.value p.value
## ma1    -0.2380 0.0686 -3.4673   6e-04
## ma2    -0.2854 0.0663 -4.3028   0e+00
## sar1   -0.5238 0.0638 -8.2150   0e+00
## sar2   -0.4608 0.0618 -7.4575   0e+00
```

```
t2$ttable
```

```
##       Estimate     SE  t.value p.value
## ma1    -0.2556 0.0695  -3.6787   3e-04
## ma2    -0.2603 0.0686  -3.7947   2e-04
## sar1   -0.6763 0.0663 -10.1938   0e+00
## sar2   -0.6331 0.0661  -9.5751   0e+00
## sar3   -0.3445 0.0657  -5.2418   0e+00
```

```
# COMPARING INFORMATION CRITERIONS
c(t1$AIC,t1$AICc,t1$BIC)
```

```
## [1] -1.723086 -1.722278 -1.648286
```

```
c(t2$AIC,t2$AICc,t2$BIC)
```

```
## [1] -1.823908 -1.822690 -1.734148
```

According to the diagnostic plots and the ttable, all estimated are significant for both the models. There is no apparent trend or pattern in the plot of standardized residuals for either of the model.ACF is significant for a couple of lags for the FIRST model, Q statistic is also significant at many lags. For the second model, normality assumption holds apart from a couple of outliers, Q statistic is not significant at most of the lags and ACF shows no apparent dependant structure. Furthermore, all three information criterions, AIC,AICc and BIC suggest the second model. Thus, the second model fits best for Gun Sales Data.

2 ANALYSIS AND RESULTS [DEATHS DATA]

In the initial time series plot in figure 2.1 shows an obvious trend, although mixed. Decrease in death followed by increase and followed by sizeable decrease again. The series does not look stationary. The acf and pacf plots in figure 2.2 verify these observations. There is a slow linear decay depicted by acf and the ordinates are large. There is one large peak in the pacf at lag 1. Therefore, due to these features,log differencing is a usual transformation.

The log differencing has removed the trend which can be observed in figure 2.3. The transformed series looks a lot closer to a stationary series, but there still exists a seasonal component in the transformed series,which is confirmed by the acf and pacf plots in figure 2.4. The acf decays slowly at lags multiple of 12 and the pacf also seems to depict the same linear decay at lags multiple of 12 behaviour. Thus, this suggests lag 12 differencing in order to remove this seasonal behaviour.

The transformed lag 12 differnced series is stationary with constant and centered around zero(figure 2.5). The slow linear decay at lags multiple of 12 behavior has also been removed which can be verified from the acf and pacf plots in figure 2.6. The ACF seems to cut off after lag 1 and the PACF seems to tail off, whereas the acf at lags multiple of 12 seems to cut off after the first multiple of 12(lag 12), and the PACF at lags multiple of 12 seems to be tailing off. Thus, the model suggested is SARIMA(0,1,1)X(0,1,1)[12]. Furthermore, I also fitted SARIMA(1,1,1)X(0,1,1)[12] to check which fits best.

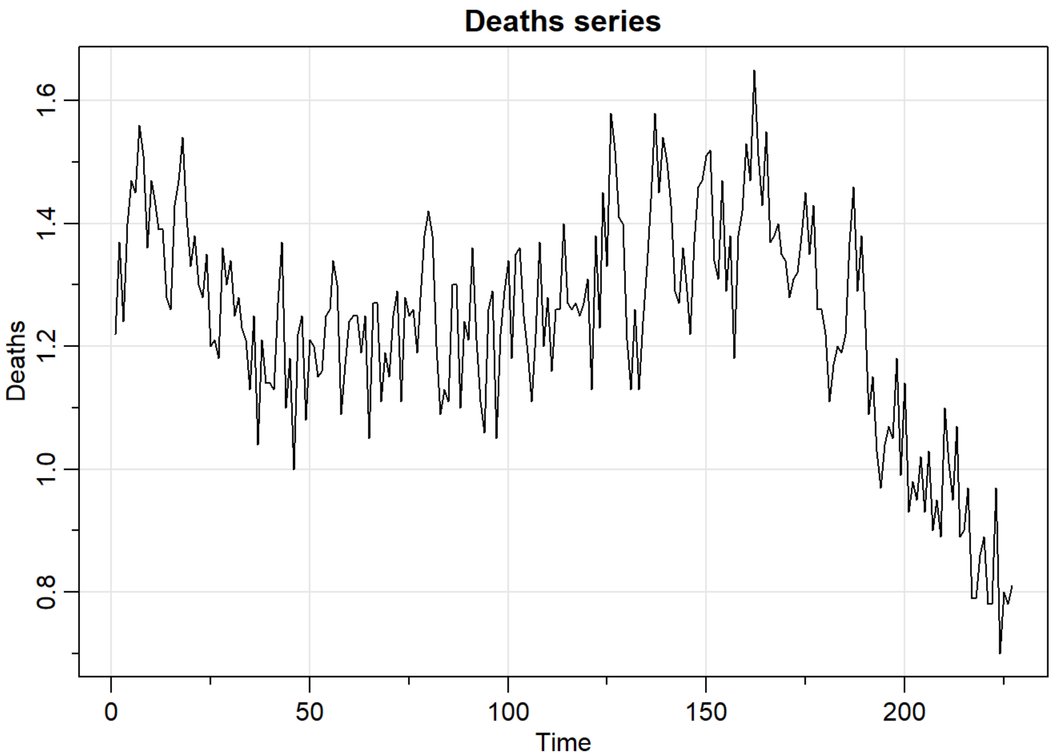


Figure 2.1: Plots for Deaths Data

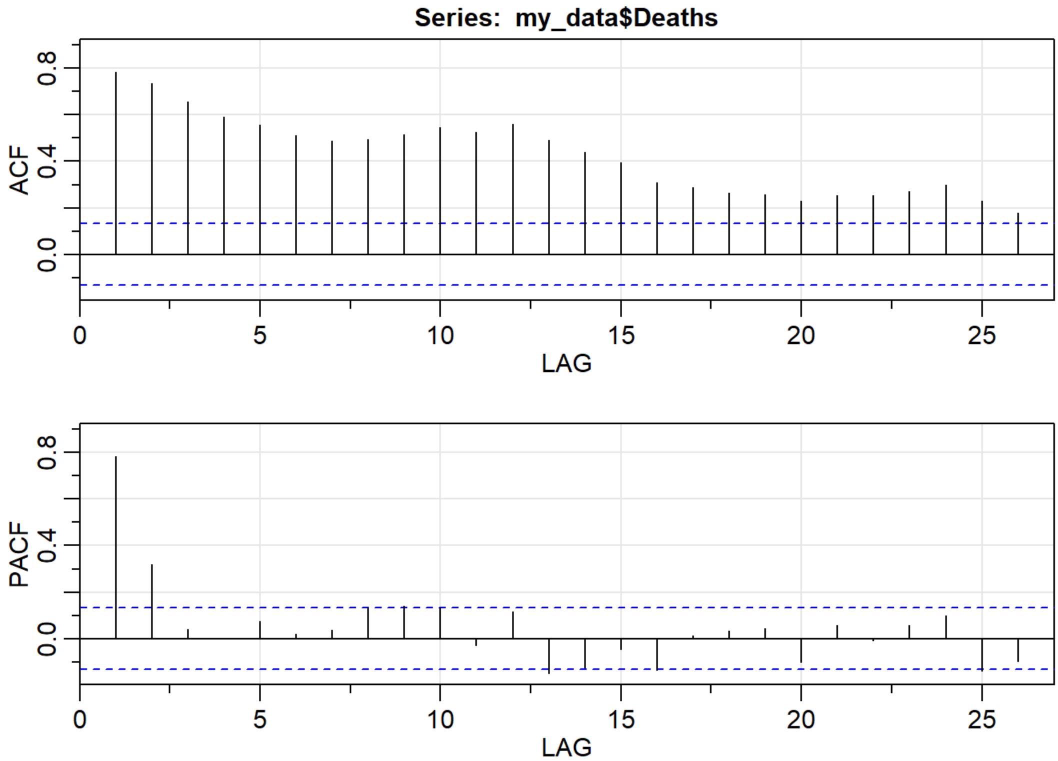


Figure 2.2: Plots for Deaths Data

```
##        ACF  PACF
##  [1,] 0.78  0.78
##  [2,] 0.74  0.32
##  [3,] 0.66  0.04
##  [4,] 0.59  0.00
##  [5,] 0.56  0.07
##  [6,] 0.51  0.02
##  [7,] 0.49  0.04
##  [8,] 0.50  0.13
##  [9,] 0.51  0.14
## [10,] 0.54  0.13
## [11,] 0.53 -0.03
## [12,] 0.56  0.12
## [13,] 0.49 -0.15
## [14,] 0.44 -0.13
## [15,] 0.39 -0.05
## [16,] 0.31 -0.14
## [17,] 0.29  0.01
## [18,] 0.26  0.03
## [19,] 0.26  0.04
## [20,] 0.23 -0.10
## [21,] 0.25  0.06
## [22,] 0.25 -0.01
## [23,] 0.27  0.06
## [24,] 0.30  0.10
## [25,] 0.23 -0.14
## [26,] 0.18 -0.10
```

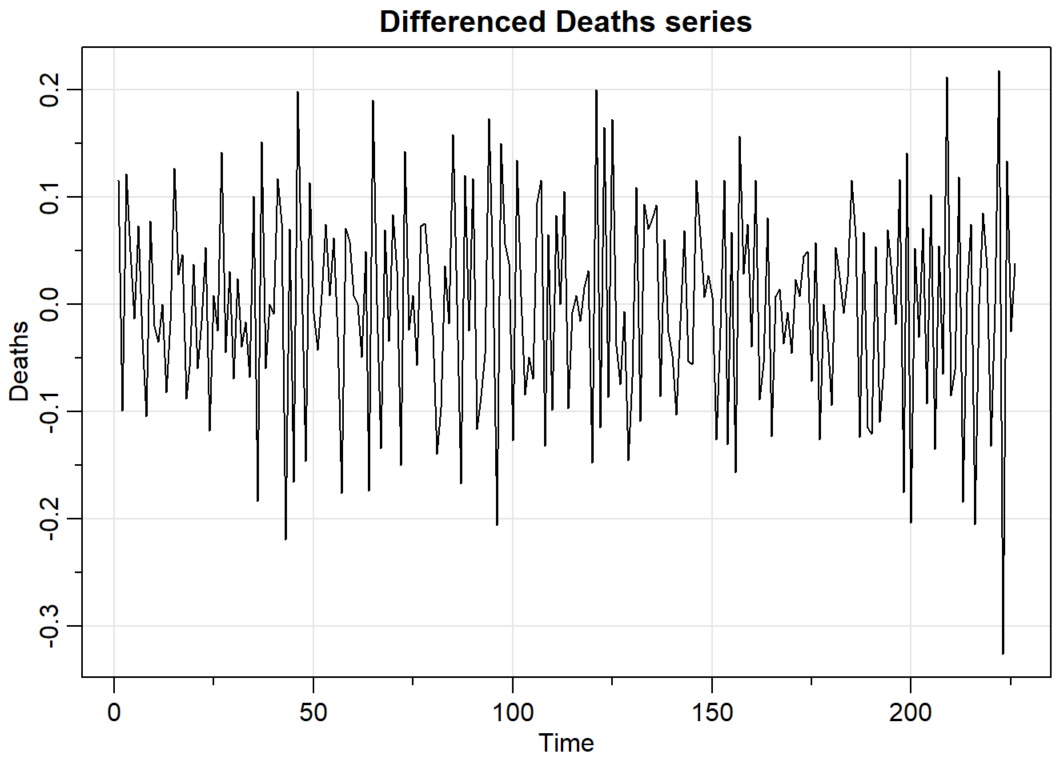


Figure 2.3: Plots for Deaths Data

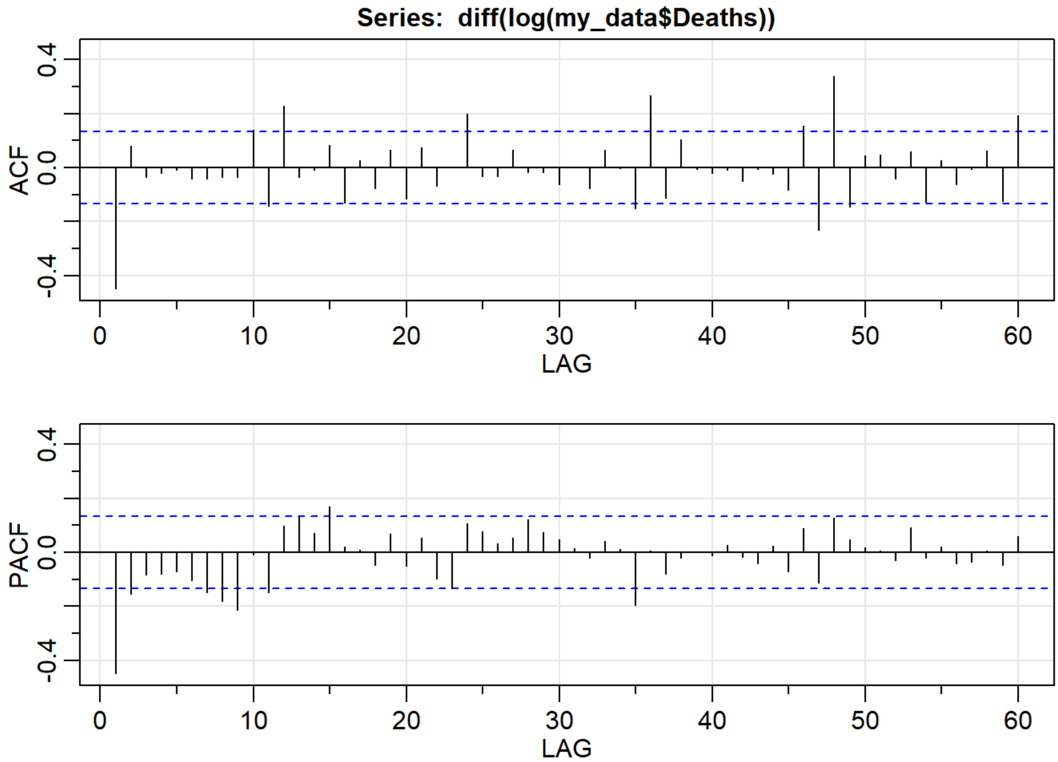


Figure 2.4: Plots for Deaths Data

```
##         ACF  PACF
##  [1,] -0.45 -0.45
##  [2,]  0.08 -0.15
##  [3,] -0.04 -0.08
##  [4,] -0.02 -0.08
##  [5,] -0.01 -0.07
##  [6,] -0.04 -0.10
##  [7,] -0.04 -0.15
##  [8,] -0.04 -0.18
##  [9,] -0.04 -0.21
## [10,]  0.14 -0.01
## [11,] -0.14 -0.15
## [12,]  0.23  0.10
## [13,] -0.03  0.14
## [14,] -0.01  0.07
## [15,]  0.08  0.17
## [16,] -0.13  0.02
## [17,]  0.03  0.01
## [18,] -0.08 -0.05
## [19,]  0.06  0.07
## [20,] -0.11 -0.05
## [21,]  0.07  0.05
## [22,] -0.07 -0.10
## [23,]  0.00 -0.13
## [24,]  0.20  0.11
## [25,] -0.03  0.08
## [26,] -0.03  0.03
## [27,]  0.07  0.05
## [28,] -0.02  0.12
## [29,] -0.02  0.08
## [30,] -0.06  0.05
## [31,]  0.00  0.01
## [32,] -0.08 -0.02
## [33,]  0.07  0.04
## [34,]  0.00  0.01
## [35,] -0.15 -0.20
## [36,]  0.27  0.00
## [37,] -0.11 -0.08
## [38,]  0.10 -0.02
## [39,] -0.01  0.00
## [40,] -0.02 -0.01
## [41,] -0.01  0.03
## [42,] -0.05 -0.02
## [43,]  0.00 -0.04
## [44,] -0.02  0.02
## [45,] -0.08 -0.07
## [46,]  0.16  0.09
## [47,] -0.23 -0.11
## [48,]  0.34  0.13
## [49,] -0.15  0.05
## [50,]  0.04  0.02
## [51,]  0.05  0.01
## [52,] -0.04 -0.03
## [53,]  0.06  0.09
## [54,] -0.13 -0.02
## [55,]  0.03  0.02
## [56,] -0.06 -0.04
## [57,] -0.01 -0.04
## [58,]  0.06  0.01
## [59,] -0.12 -0.05
## [60,]  0.19  0.06
```

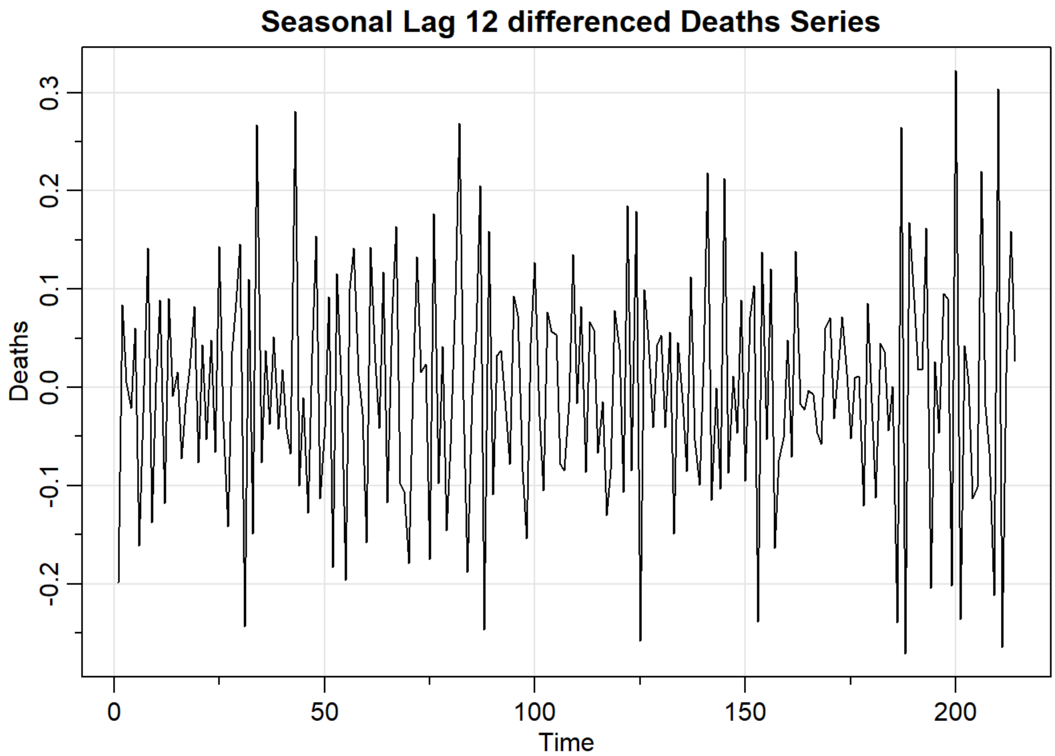


Figure 2.5: Plots for Deaths Data

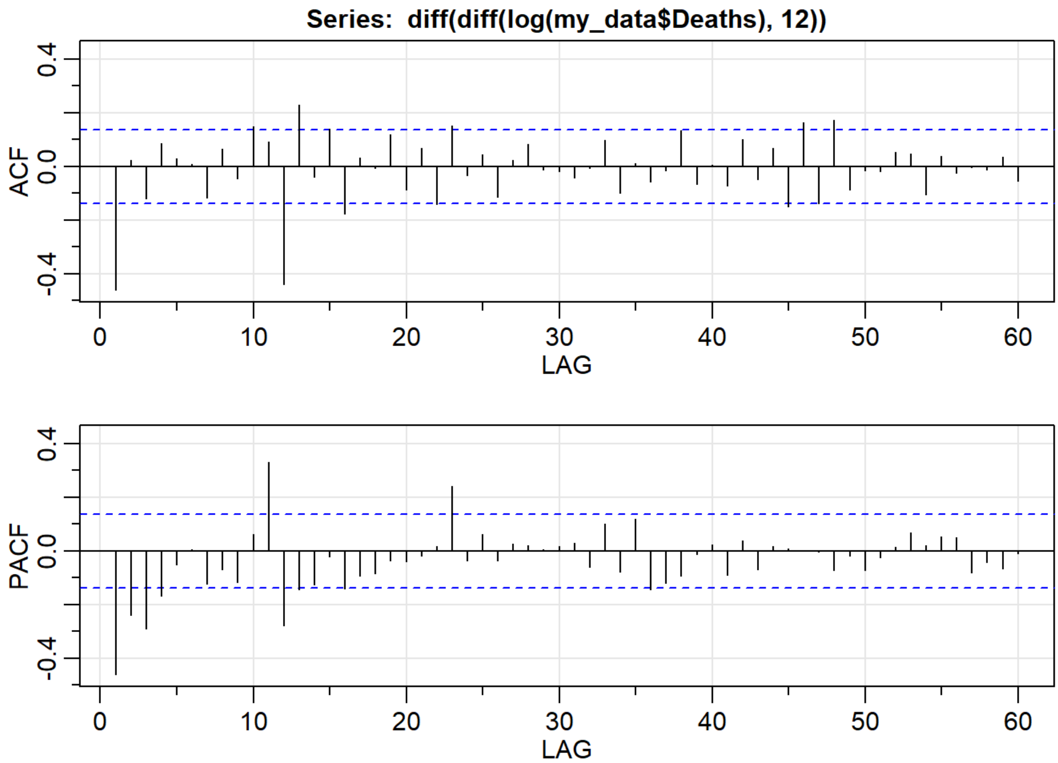
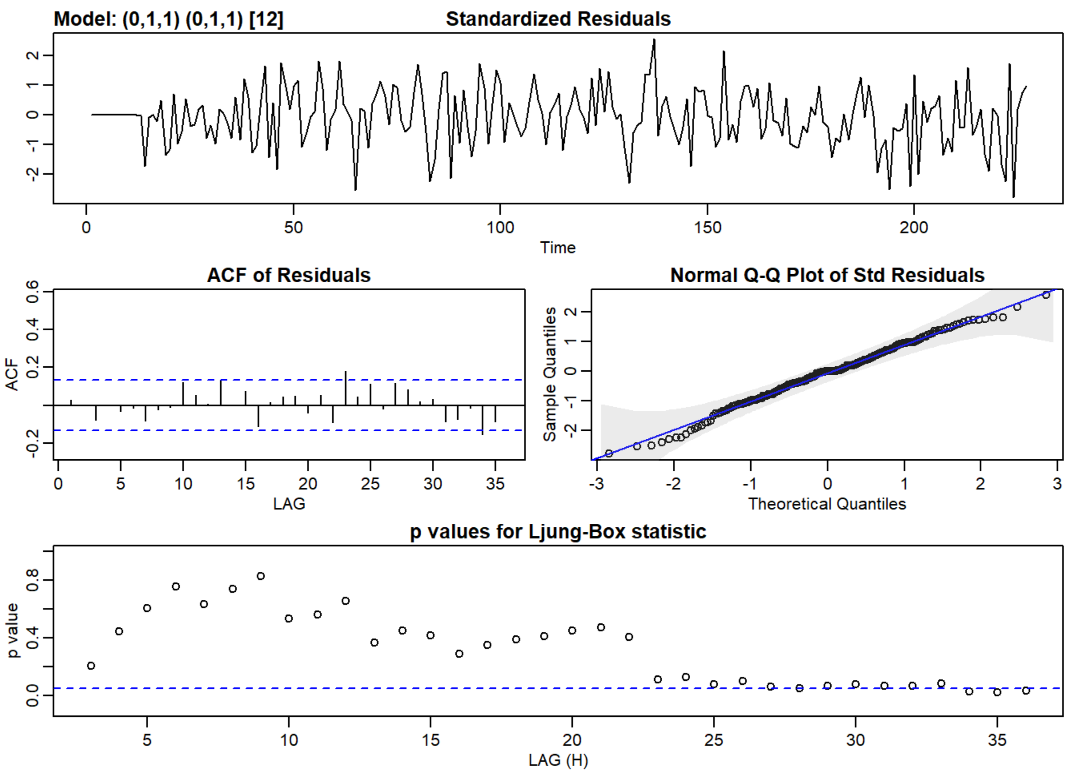


Figure 2.6: Plots for Deaths Data

```
##         ACF  PACF
##  [1,] -0.46 -0.46
##  [2,]  0.02 -0.24
##  [3,] -0.12 -0.29
##  [4,]  0.09 -0.17
##  [5,]  0.03 -0.05
##  [6,]  0.01  0.00
##  [7,] -0.12 -0.12
##  [8,]  0.06 -0.07
##  [9,] -0.05 -0.12
## [10,]  0.15  0.06
## [11,]  0.09  0.33
## [12,] -0.44 -0.28
## [13,]  0.23 -0.14
## [14,] -0.04 -0.13
## [15,]  0.14 -0.02
## [16,] -0.18 -0.14
## [17,]  0.03 -0.09
## [18,] -0.01 -0.08
## [19,]  0.12 -0.04
## [20,] -0.09 -0.04
## [21,]  0.07 -0.02
## [22,] -0.14  0.02
## [23,]  0.15  0.24
## [24,] -0.03 -0.04
## [25,]  0.04  0.06
## [26,] -0.11 -0.04
## [27,]  0.02  0.03
## [28,]  0.08  0.02
## [29,] -0.01  0.01
## [30,] -0.02  0.02
## [31,] -0.04  0.03
## [32,] -0.01 -0.06
## [33,]  0.10  0.10
## [34,] -0.10 -0.08
## [35,]  0.01  0.12
## [36,] -0.06 -0.14
## [37,] -0.01 -0.12
## [38,]  0.13 -0.09
## [39,] -0.07 -0.01
## [40,]  0.01  0.02
## [41,] -0.07 -0.09
## [42,]  0.10  0.04
## [43,] -0.05 -0.07
## [44,]  0.07  0.02
## [45,] -0.15  0.01
## [46,]  0.16  0.00
## [47,] -0.14  0.00
## [48,]  0.17 -0.07
## [49,] -0.09 -0.02
## [50,] -0.02 -0.07
## [51,] -0.02 -0.02
## [52,]  0.05  0.02
## [53,]  0.05  0.07
## [54,] -0.10  0.02
## [55,]  0.04  0.05
## [56,] -0.02  0.05
## [57,]  0.00 -0.08
## [58,] -0.01 -0.04
## [59,]  0.04 -0.07
## [60,] -0.05 -0.01
```

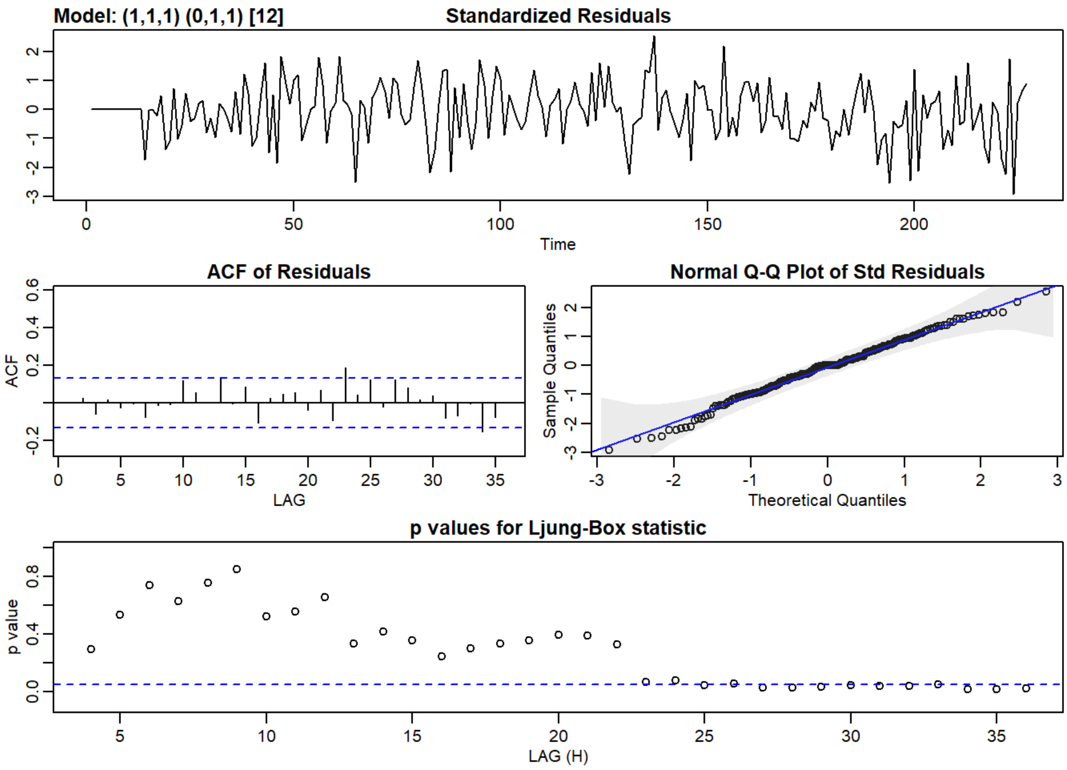
```
# MODEL 1
w1 = sarima(log(my_data$Deaths),p=0,d=1,q=1,P=0,D=1,Q=1,S=12)
```

```
## initial  value -2.154772 
## iter   2 value -2.478337
## iter   3 value -2.540191
## iter   4 value -2.556627
## iter   5 value -2.559159
## iter   6 value -2.566006
## iter   7 value -2.566593
## iter   8 value -2.567831
## iter   9 value -2.567891
## iter  10 value -2.567893
## iter  11 value -2.567894
## iter  11 value -2.567894
## final  value -2.567894 
## converged
## initial  value -2.587001 
## iter   2 value -2.596230
## iter   3 value -2.600277
## iter   4 value -2.600383
## iter   5 value -2.600414
## iter   6 value -2.600418
## iter   7 value -2.600418
## iter   8 value -2.600418
## iter   8 value -2.600418
## iter   8 value -2.600418
## final  value -2.600418 
## converged
```




```
# MODEL 2
w2 = sarima(log(my_data$Deaths),p=1,d=1,q=1,P=0,D=1,Q=1,S=12)
```

```
## initial  value -2.159321 
## iter   2 value -2.470073
## iter   3 value -2.544202
## iter   4 value -2.556757
## iter   5 value -2.562590
## iter   6 value -2.564163
## iter   7 value -2.567692
## iter   8 value -2.569172
## iter   9 value -2.569416
## iter  10 value -2.569625
## iter  11 value -2.569678
## iter  12 value -2.569681
## iter  13 value -2.569681
## iter  13 value -2.569681
## iter  13 value -2.569681
## final  value -2.569681 
## converged
## initial  value -2.581007 
## iter   2 value -2.595381
## iter   3 value -2.600802
## iter   4 value -2.601087
## iter   5 value -2.601350
## iter   6 value -2.601354
## iter   7 value -2.601355
## iter   8 value -2.601355
## iter   8 value -2.601355
## final  value -2.601355 
## converged
```




```
w1$ttable
```

```
##      Estimate     SE  t.value p.value
## ma1   -0.6510 0.0561 -11.6056       0
## sma1  -0.9998 0.1519  -6.5827       0
```

```
w2$ttable
```

```
##      Estimate     SE t.value p.value
## ar1    0.0669 0.1047  0.6392  0.5234
## ma1   -0.6908 0.0766 -9.0220  0.0000
## sma1  -0.9999 0.1551 -6.4483  0.0000
```

```
# COMPARING INFORMATION CRITERIONS
c(w1$AIC,w1$AICc,w1$BIC)
```

```
## [1] -2.220771 -2.220530 -2.175891
```

```
c(w2$AIC,w2$AICc,w2$BIC)
```

```
## [1] -2.213663 -2.213180 -2.153824
```

From the diagnostic plots of both the candidate models, we can draw the following conclusions: There is no apparent trend or pattern in the plot of the standardized residuals;There are no outliers with residuals exceeding three standard deviation in magniture;ACF shows no apparent significant dependence structure;Q statistics is not significant at most of the lags;Normality assumption seems to be appropriate with the exception of outliers. But estimated parameter for ar1 in the second model is not significant as the p value is not close to 0, thus we drop it from the model and choose model 1, which is (0,1,1)X(0,1,1)[12]. All the three information criterions, AIC, AICc, BIC also suggest the first model.

3 Regression with Autocorrelated Errors

As the ccf,in figure 3.1, does not show any pattern of leading and lagging between the two time series, we examine regression with autocorrelated errors to find out the presence of linear relationship between firearm deaths and handgun sales. The time series plot of the residuals,figure 3.2, shows a slight increasing trend followed by a decreasing trend towards the end. The existence of trend and seasonality can be seen from the acf and pacf plots in figure 3.3. The plot is also not stationary. A first difference is taken to remove trend, but there still exists slow linear decay in acf at lags multiple of 12 and large peak of pacf at lag 1, which can be observed from figure 3.4. Thus,lag 12 difference is taken to remove this behavior and make the series stationary. We can fit a model for the residuals based on the acf and pacf plots in figure 3.5. The acf cuts off after lag 1, and the pacf tails off. At lags multiple of 12, the acf seems to be tailing off and the PACF sees to cut off after second multiple of 12, lag =24. Thus, the candidate model is SARIMA(0,1,1)X(2,1,0)[12]

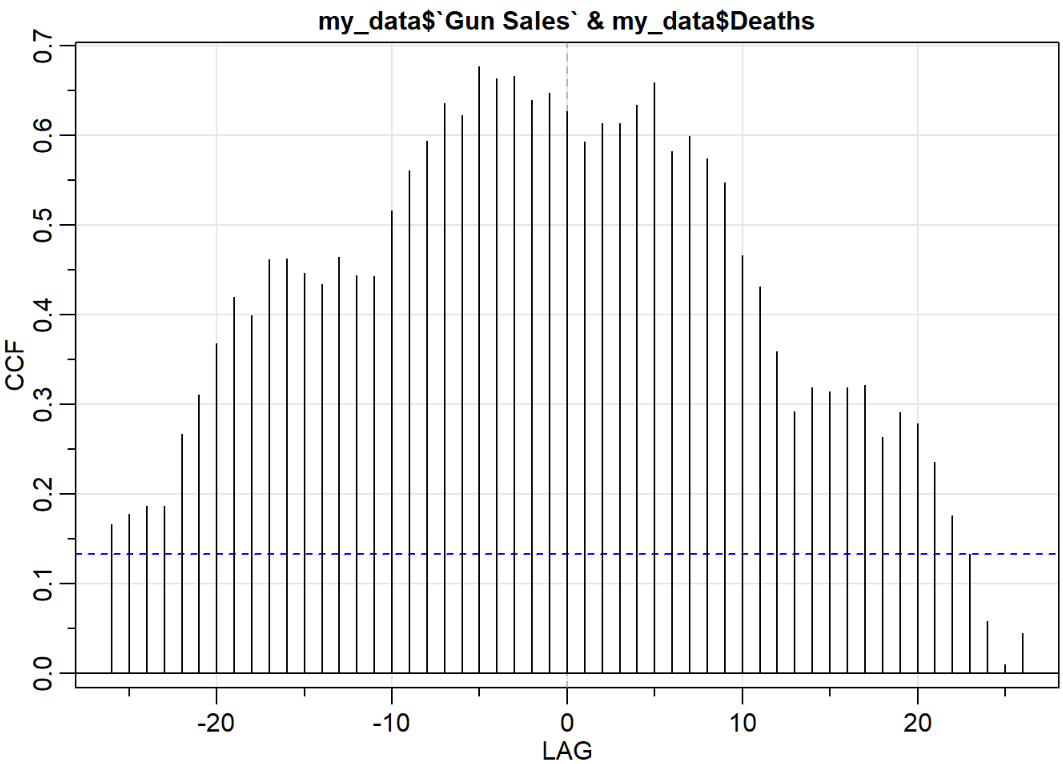


Figure 3.1: ccf and Reg. plots

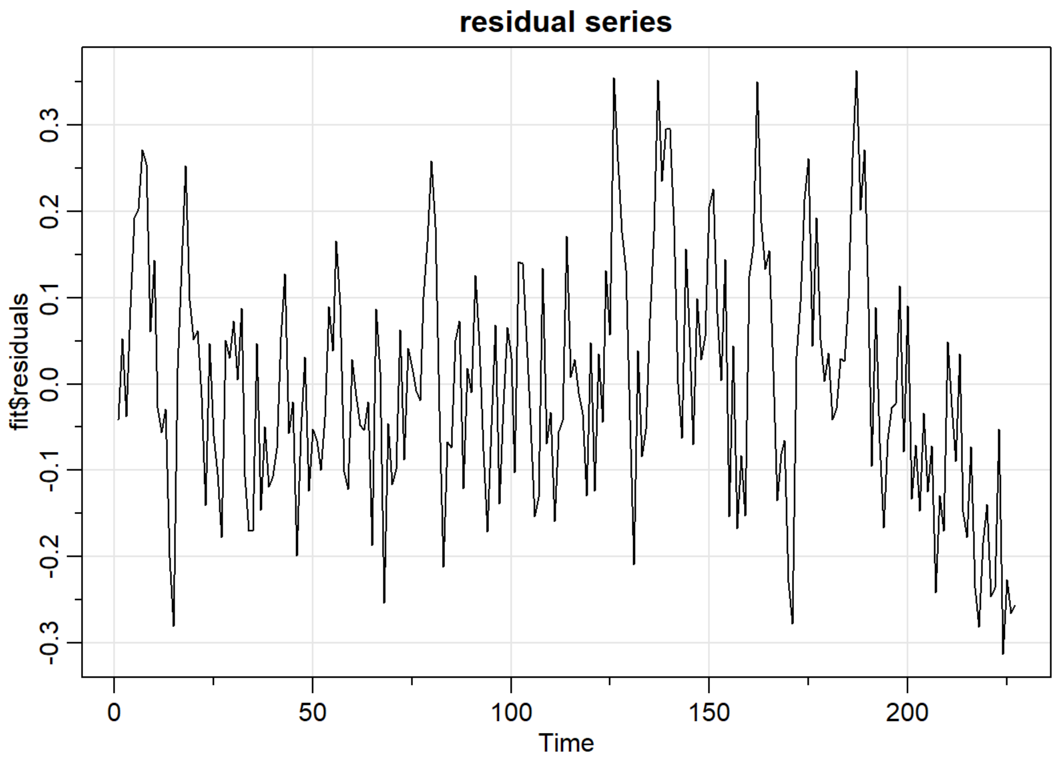


Figure 3.2: ccf and Reg. plots

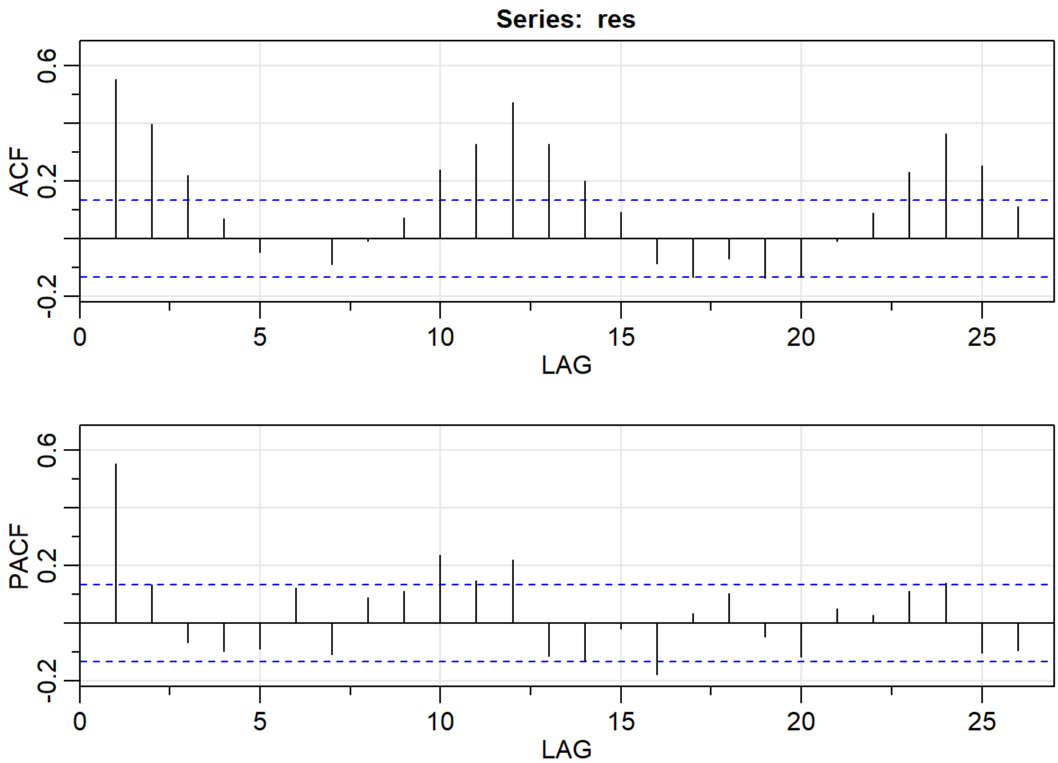


Figure 3.3: ccf and Reg. plots

```
##          ACF  PACF
##  [1,]  0.55  0.55
##  [2,]  0.40  0.13
##  [3,]  0.22 -0.07
##  [4,]  0.07 -0.10
##  [5,] -0.05 -0.09
##  [6,]  0.00  0.12
##  [7,] -0.09 -0.11
##  [8,] -0.01  0.09
##  [9,]  0.07  0.11
## [10,]  0.24  0.24
## [11,]  0.33  0.15
## [12,]  0.47  0.22
## [13,]  0.33 -0.11
## [14,]  0.20 -0.13
## [15,]  0.09 -0.02
## [16,] -0.09 -0.18
## [17,] -0.13  0.03
## [18,] -0.07  0.10
## [19,] -0.14 -0.05
## [20,] -0.13 -0.12
## [21,] -0.01  0.05
## [22,]  0.09  0.03
## [23,]  0.23  0.11
## [24,]  0.36  0.14
## [25,]  0.25 -0.10
## [26,]  0.11 -0.09
```

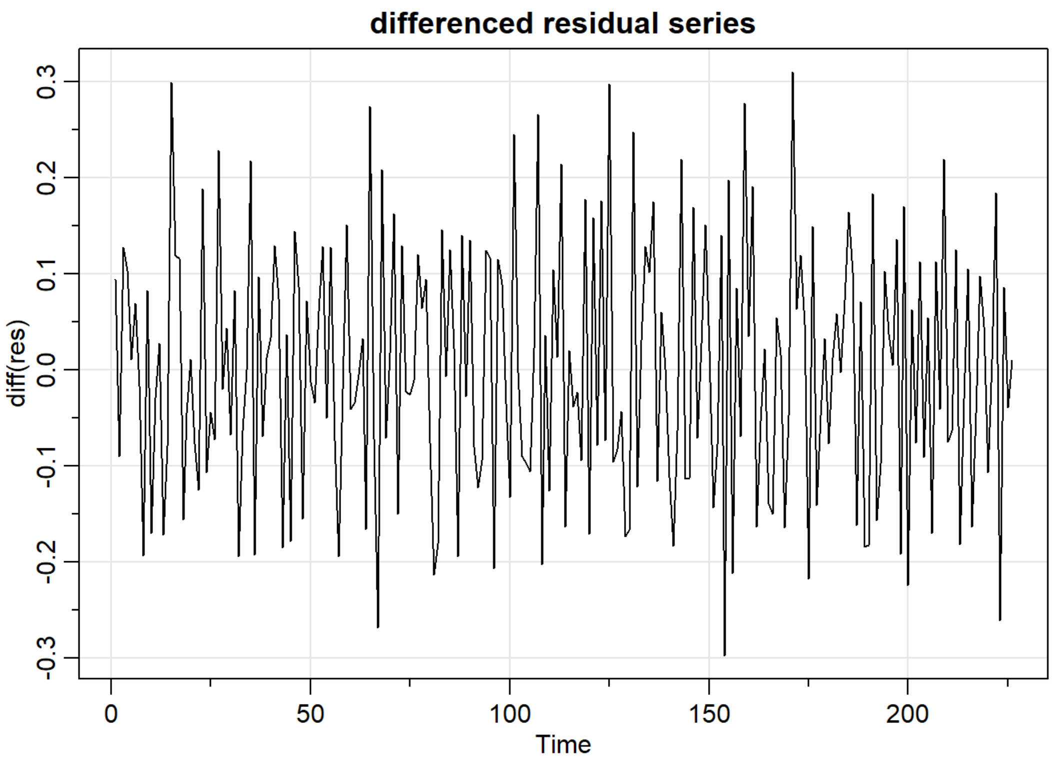


Figure 3.4: ccf and Reg. plots

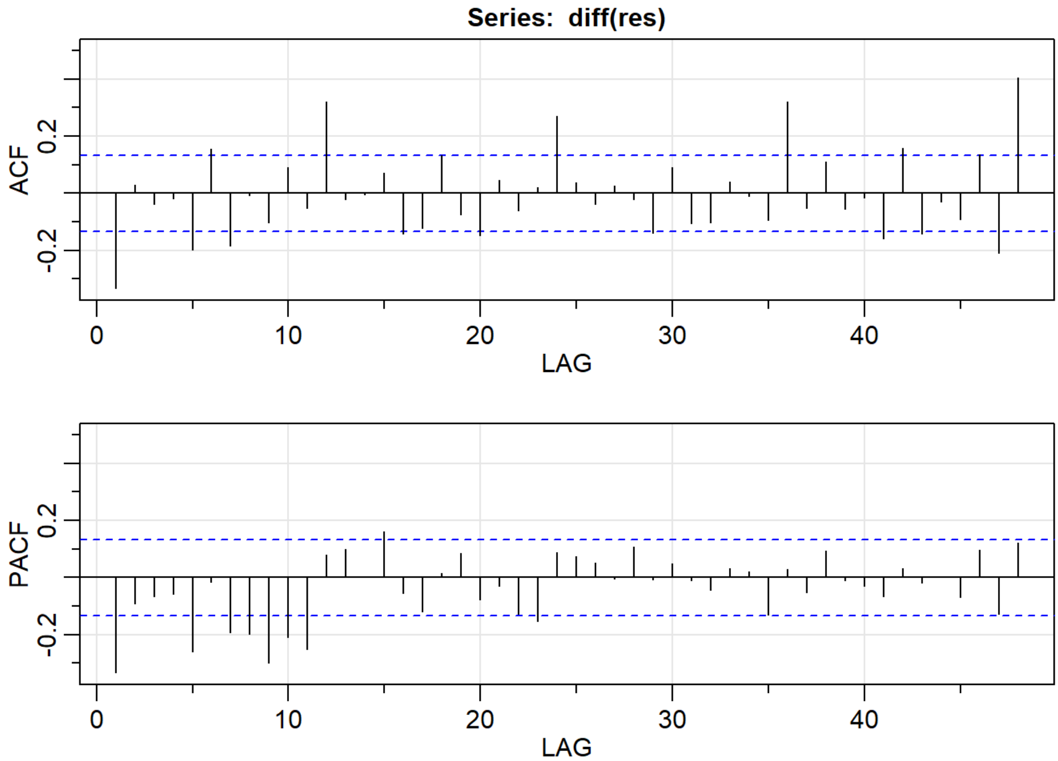


Figure 3.5: ccf and Reg. plots

```
##         ACF  PACF
##  [1,] -0.33 -0.33
##  [2,]  0.03 -0.09
##  [3,] -0.04 -0.07
##  [4,] -0.02 -0.06
##  [5,] -0.20 -0.26
##  [6,]  0.15 -0.02
##  [7,] -0.18 -0.19
##  [8,] -0.01 -0.20
##  [9,] -0.10 -0.30
## [10,]  0.09 -0.21
## [11,] -0.05 -0.25
## [12,]  0.32  0.08
## [13,] -0.02  0.10
## [14,] -0.01  0.00
## [15,]  0.07  0.16
## [16,] -0.14 -0.06
## [17,] -0.12 -0.12
## [18,]  0.13  0.01
## [19,] -0.08  0.09
## [20,] -0.15 -0.08
## [21,]  0.05 -0.03
## [22,] -0.06 -0.13
## [23,]  0.02 -0.15
## [24,]  0.27  0.09
## [25,]  0.04  0.07
## [26,] -0.04  0.05
## [27,]  0.02  0.00
## [28,] -0.02  0.11
## [29,] -0.14 -0.01
## [30,]  0.09  0.05
## [31,] -0.11 -0.01
## [32,] -0.10 -0.04
## [33,]  0.04  0.03
## [34,] -0.01  0.02
## [35,] -0.10 -0.13
## [36,]  0.32  0.03
## [37,] -0.05 -0.05
## [38,]  0.11  0.09
## [39,] -0.06 -0.01
## [40,] -0.02 -0.03
## [41,] -0.16 -0.07
## [42,]  0.16  0.03
## [43,] -0.14 -0.02
## [44,] -0.03  0.00
## [45,] -0.09 -0.07
## [46,]  0.13  0.10
## [47,] -0.21 -0.13
## [48,]  0.41  0.12
```

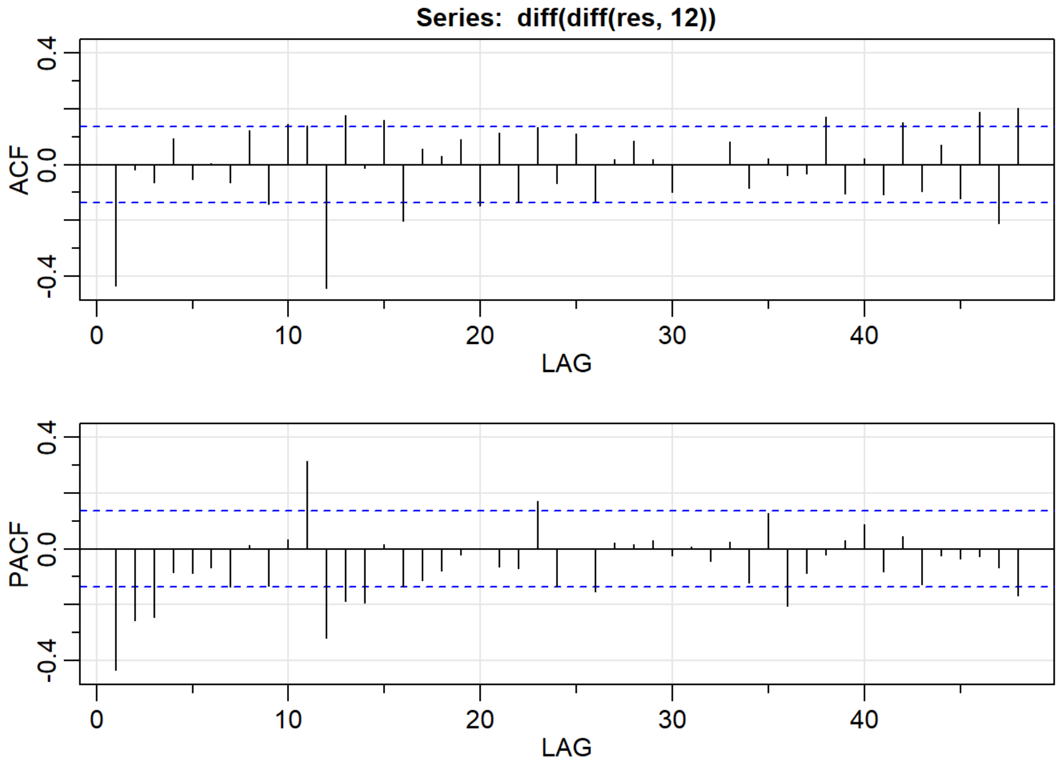
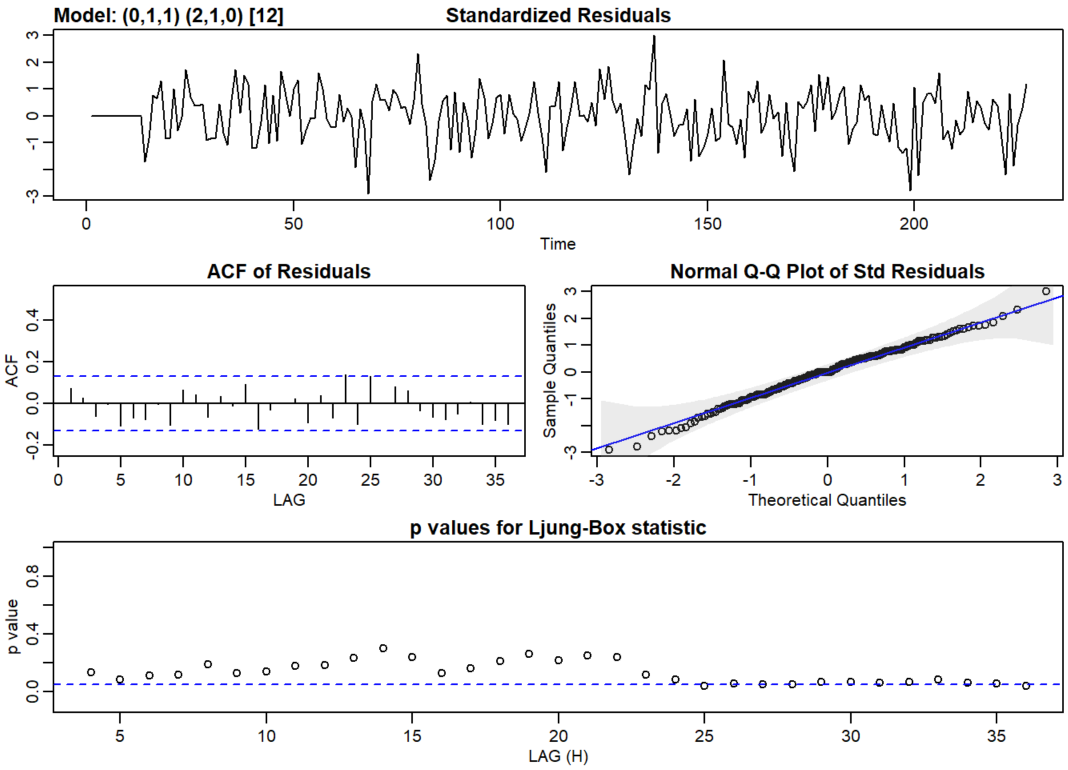


Figure 3.6: ccf and Reg. plots

```
##         ACF  PACF
##  [1,] -0.44 -0.44
##  [2,] -0.02 -0.26
##  [3,] -0.06 -0.25
##  [4,]  0.09 -0.08
##  [5,] -0.05 -0.09
##  [6,]  0.01 -0.07
##  [7,] -0.06 -0.14
##  [8,]  0.12  0.01
##  [9,] -0.14 -0.13
## [10,]  0.14  0.03
## [11,]  0.14  0.32
## [12,] -0.44 -0.32
## [13,]  0.18 -0.19
## [14,] -0.01 -0.19
## [15,]  0.16  0.02
## [16,] -0.20 -0.14
## [17,]  0.06 -0.11
## [18,]  0.03 -0.08
## [19,]  0.09 -0.02
## [20,] -0.15  0.00
## [21,]  0.11 -0.07
## [22,] -0.14 -0.07
## [23,]  0.13  0.17
## [24,] -0.07 -0.14
## [25,]  0.11  0.00
## [26,] -0.13 -0.15
## [27,]  0.02  0.02
## [28,]  0.08  0.02
## [29,]  0.02  0.03
## [30,] -0.10 -0.03
## [31,]  0.00  0.01
## [32,]  0.00 -0.04
## [33,]  0.08  0.03
## [34,] -0.08 -0.12
## [35,]  0.02  0.13
## [36,] -0.04 -0.21
## [37,] -0.03 -0.09
## [38,]  0.17 -0.02
## [39,] -0.10  0.03
## [40,]  0.02  0.09
## [41,] -0.11 -0.08
## [42,]  0.15  0.04
## [43,] -0.10 -0.13
## [44,]  0.07 -0.02
## [45,] -0.12 -0.04
## [46,]  0.19 -0.03
## [47,] -0.21 -0.07
## [48,]  0.20 -0.17
```

```
# MODEL FITTING FOR RESIDUALS
sarima(res,p=0,d=1,q=1,P=2,D=1,Q=0,S=12)
```

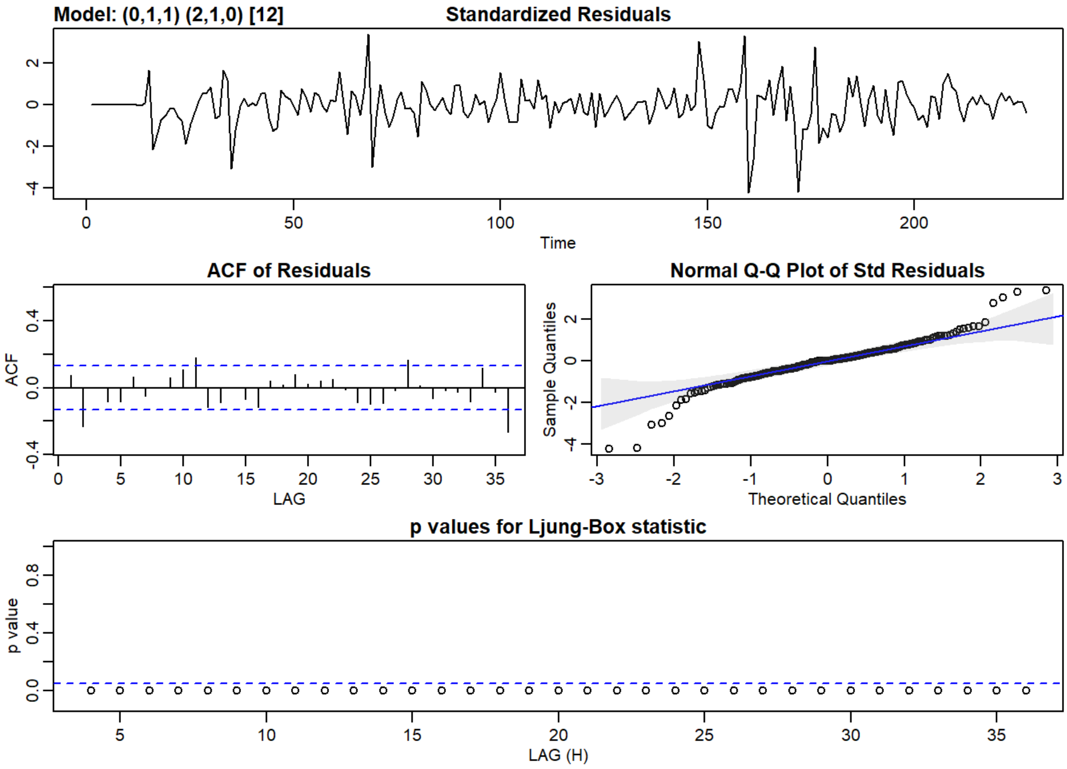
```
## initial  value -1.876382
## iter   2 value -2.189632
## iter   3 value -2.232210
## iter   4 value -2.257917
## iter   5 value -2.262970
## iter   6 value -2.263771
## iter   7 value -2.263851
## iter   8 value -2.263852
## iter   8 value -2.263852
## final  value -2.263852
## converged
## initial  value -2.257514
## iter   2 value -2.257573
## iter   3 value -2.257576
## iter   3 value -2.257576
## iter   3 value -2.257576
## final  value -2.257576
## converged
```



```
## $fit
##
## Call:
## stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,
##     Q), period = S), include.mean = !no.constant, transform.pars = trans, fixed = fixed,
##     optim.control = list(trace = trc, REPORT = 1, reltol = tol))
##
## Coefficients:
##           ma1      sar1      sar2
##       -0.7323   -0.6572   -0.2910
## s.e.   0.0626    0.0675    0.0689
##
## sigma^2 estimated as 0.01061:  log likelihood = 179.47,  aic = -350.94
##
## $degrees_of_freedom
## [1] 211
##
## $ttable
##      Estimate     SE  t.value p.value
## ma1   -0.7323 0.0626 -11.7070       0
## sar1  -0.6572 0.0675  -9.7394       0
## sar2  -0.2910 0.0689  -4.2252       0
##
## $AIC
## [1] -1.559718
##
## $AICc
## [1] -1.559236
##
## $BIC
## [1] -1.499879
```

```r
# CHECKING THE SAME MODEL FOR GUN SALES
sarima((my_data$`Gun Sales`),p=0,d=1,q=1,P=2,D=1,Q=0,S=12)
```

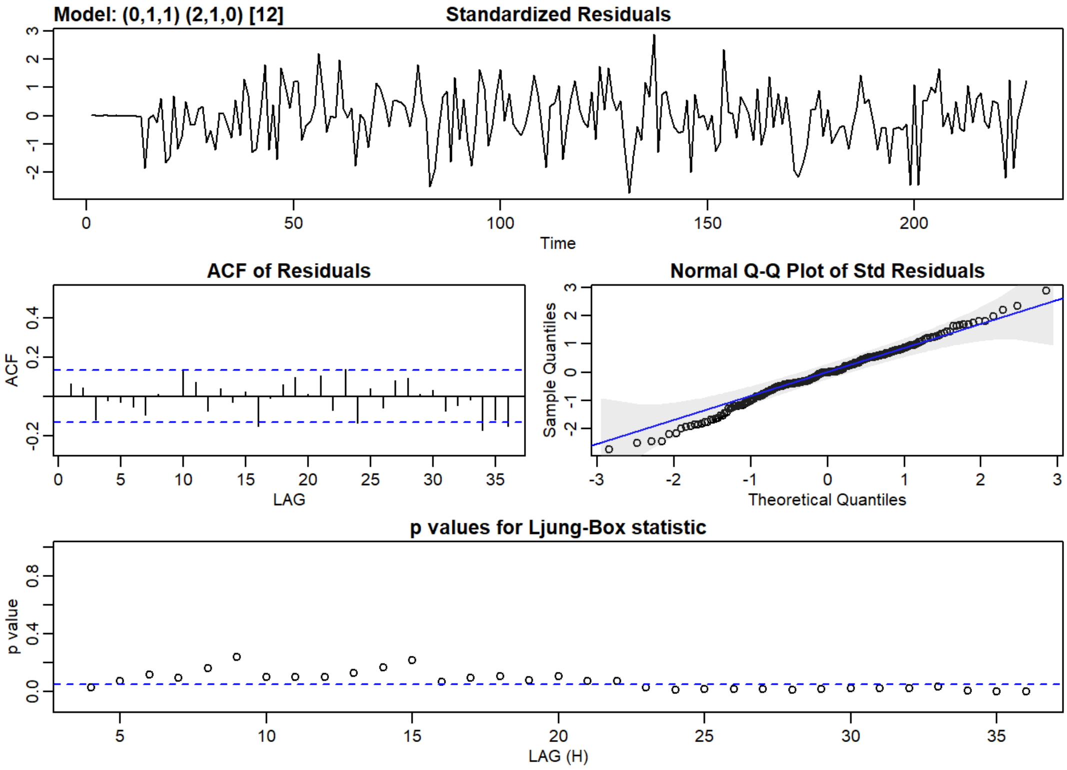
```
## initial  value 2.807938
## iter   2 value 2.627894
## iter   3 value 2.618715
## iter   4 value 2.602064
## iter   5 value 2.601500
## iter   6 value 2.601337
## iter   6 value 2.601337
## iter   6 value 2.601337
## final  value 2.601337
## converged
## initial  value 2.642894
## iter   2 value 2.641903
## iter   3 value 2.641812
## iter   4 value 2.641775
## iter   5 value 2.641774
## iter   5 value 2.641774
## iter   5 value 2.641774
## final  value 2.641774
## converged
```




```
## $fit
##
## Call:
## stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,
##     Q), period = S), include.mean = !no.constant, transform.pars = trans, fixed = fixed,
##     optim.control = list(trace = trc, REPORT = 1, reltol = tol))
##
## Coefficients:
##           ma1     sar1     sar2
##       -0.3299  -0.4713  -0.4258
## s.e.   0.0897   0.0638   0.0625
##
## sigma^2 estimated as 191.3:  log likelihood = -868.99,  aic = 1745.98
##
## $degrees_of_freedom
## [1] 211
##
## $ttable
##      Estimate     SE t.value p.value
## ma1   -0.3299 0.0897 -3.6792   3e-04
## sar1  -0.4713 0.0638 -7.3839   0e+00
## sar2  -0.4258 0.0625 -6.8111   0e+00
##
## $AIC
## [1] 7.759932
##
## $AICc
## [1] 7.760415
##
## $BIC
## [1] 7.819772
```

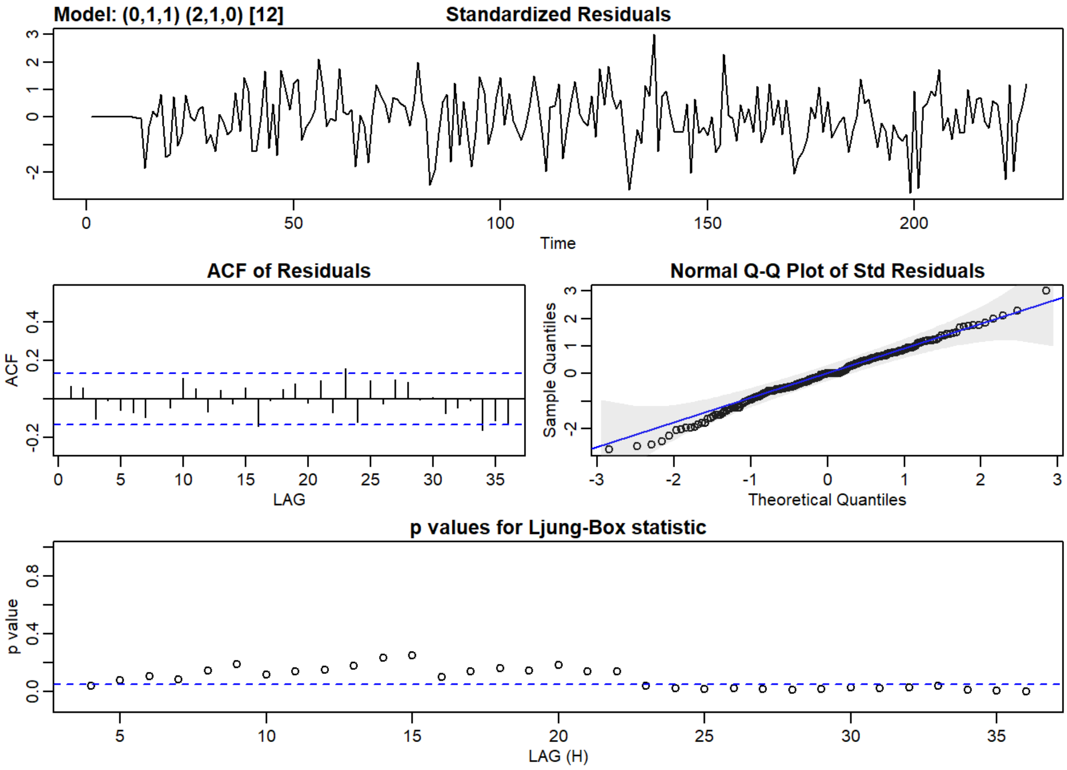
```
# CHECKING SAME MODEL FOR DEATHS
w3 = sarima((my_data$Deaths),p=0,d=1,q=1,P=2,D=1,Q=0,S=12)
```

```
## initial  value -1.951232
## iter   2 value -2.257666
## iter   3 value -2.299601
## iter   4 value -2.322615
## iter   5 value -2.331960
## iter   6 value -2.334601
## iter   7 value -2.335509
## iter   8 value -2.335641
## iter   9 value -2.335644
## iter  10 value -2.335646
## iter  11 value -2.335646
## iter  11 value -2.335646
## iter  11 value -2.335646
## final  value -2.335646
## converged
## initial  value -2.345482
## iter   2 value -2.345745
## iter   3 value -2.345763
## iter   4 value -2.345763
## iter   5 value -2.345763
## iter   5 value -2.345763
## iter   5 value -2.345763
## final  value -2.345763
## converged
```




```
# Running weighted least squares on the regression model for autocorrelated errors using the model specified
# for residuals
sarima((my_data$Deaths),p=0,d=1,q=1,P=2,D=1,Q=0,S=12,xreg = cbind(my_data$`Gun Sales`))
```

```
## initial  value -1.951636
## iter   2 value -2.261456
## iter   3 value -2.306136
## iter   4 value -2.328635
## iter   5 value -2.341509
## iter   6 value -2.345732
## iter   7 value -2.347628
## iter   8 value -2.349130
## iter   9 value -2.349390
## iter  10 value -2.349412
## iter  11 value -2.349412
## iter  12 value -2.349412
## iter  12 value -2.349412
## iter  12 value -2.349412
## final  value -2.349412
## converged
## initial  value -2.358714
## iter   2 value -2.359162
## iter   3 value -2.359186
## iter   4 value -2.359194
## iter   5 value -2.359194
## iter   5 value -2.359194
## iter   5 value -2.359194
## final  value -2.359194
## converged
```



```
## $fit
##
## Call:
## stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,
##     Q), period = S), xreg = xreg, transform.pars = trans, fixed = fixed, optim.control = list(trace = tr
c,
##     REPORT = 1, reltol = tol))
##
## Coefficients:
##           ma1     sar1     sar2   xreg
##       -0.7684  -0.6591  -0.2553  1e-03
## s.e.   0.0574   0.0686   0.0710  4e-04
##
## sigma^2 estimated as 0.008665:  log likelihood = 201.21,  aic = -392.43
##
## $degrees_of_freedom
## [1] 210
##
## $ttable
##      Estimate     SE  t.value p.value
## ma1   -0.7684 0.0574 -13.3889  0.0000
## sar1  -0.6591 0.0686  -9.6075  0.0000
## sar2  -0.2553 0.0710  -3.5960  0.0004
## xreg   0.0010 0.0004   2.4484  0.0152
##
## $AIC
## [1] -1.744131
##
## $AICc
## [1] -1.743323
##
## $BIC
## [1] -1.669331
```

```
# Comparing information criterions of  original fitted model for deaths along with this model
c(w2$AIC,w2$AICc,w2$BIC)
```

```
## [1] -2.213663 -2.213180 -2.153824
```

```
c(w3$AIC,w3$AICc,w3$BIC)
```

```
## [1] -1.727471 -1.726988 -1.667631
```

From the diagnostic output for the residuals, all the requirements are satisfied.There is no apparent trend or pattern in the plot of the standardized residualsThere are no outliers with residuals exceeding three standard deviation in magniture;ACF shows no apparent significant dependence structure;Q statistics are not significant significant at most lags;Normality assumption seems to be appropriate with the exception of outliers. Thus, SARIMA(0,1,1)X(2,1,0)[12] fits the reiduals well. The same model does not fit the gun sales data as Q statistic is significant at all lags. This model also does not fit well for the deaths data as Q statistic is insignificant at more lags for the originally fitted model as compared to this one. The information criterions also suggest the same.

On running the weighted least squares method on the regression model with autocorrelated errors using the residuals model, we find that all the estimated parameters are significant and the regression coefficient for x is positive and equal to 0.001, which signifies a positive linear relationship between handgun sales and deaths. That linear model is Y~ 0.001x + Wt, where Wt is white noise ~(0,sigma^2)

4 CONCLUSION:

1. The best model for handgun sales is SARIMA(0,1,2)x(3,1,0)[12]
2. The best model for deaths is SARIMA(0,1,1)x(0,1,1)[12]
3. The best model for residuals is SARIMA(0,1,1)X(2,1,0)[12]
4. Handgun Sales and Deaths are positively related. The linear model is Y~ 0.001x + Wt, where Wt is white noise ~(0,sigma^2)

4.0.1 Appendix

```r
my_data = read.table("C:/Users/Falak/Desktop/137/GD.dat.txt",header = FALSE)
library(astsa)
colnames(my_data) = c("Gun Sales","Deaths")
attach(my_data)
# GUN SALES DATA
# INITIAL CHECKING
tsplot(`Gun Sales`,ylab = "Gun Sales",main = "Gun Sales series")
acf2(`Gun Sales`)
# FIRST DIFFERNCED SERIES
tsplot(diff(log(`Gun Sales`)),ylab = "Gun Sales",main = "Differenced Gun Sales series" )
acf2(diff(log(`Gun Sales`)),max.lag = 60)
# LAG 12 DIFFERENCED SERIES
tsplot(diff(diff(log(`Gun Sales`),12)),ylab = "Gun Sales",main = " Seasonal Lag 12 differenced Gun Sales Ser
ies")
acf2(diff(diff(log(`Gun Sales`),12)),max.lag = 48)
# MODEL 1
t1 = sarima(log(`Gun Sales`),p=0,d=1,q=2,P=2,D=1,Q=0,S=12)
# MODEL 2
t2 = sarima(log(`Gun Sales`),p=0,d=1,q=2,P=3,D=1,Q=0,S=12)
# MODEL SELECTION
t1$ttable
t2$ttable
# COMPARING INFORMATION CRITERIONS
c(t1$AIC,t1$AICc,t1$BIC)
c(t2$AIC,t2$AICc,t2$BIC)
# DEATHS DATA
# INITIAL CHECKING
tsplot(my_data$Deaths,ylab = "Deaths",main = "Deaths series")
acf2(my_data$Deaths)
# FIRST DIFFERENCED SERIES
tsplot(diff(log(my_data$Deaths)),ylab = "Deaths",main = "Differenced Deaths series" )
acf2(diff(log(my_data$Deaths)),max.lag = 60)
# LAG 12 DIFFERENCED SERIES
tsplot(diff(diff(log(my_data$Deaths),12)),ylab = "Deaths",main = " Seasonal Lag 12 differenced Deaths Series
")
acf2(diff(diff(log(my_data$Deaths),12)),max.lag = 60)
# MODEL 1
w1 = sarima(log(my_data$Deaths),p=0,d=1,q=1,P=0,D=1,Q=1,S=12)
# MODEL 2
w2 = sarima(log(my_data$Deaths),p=1,d=1,q=1,P=0,D=1,Q=1,S=12)
w1$ttable
w2$ttable
# COMPARING INFORMATION CRITERIONS
c(w1$AIC,w1$AICc,w1$BIC)
c(w2$AIC,w2$AICc,w2$BIC)
ccf2(my_data$`Gun Sales`,my_data$Deaths)
fit = lm(my_data$Deaths~my_data$`Gun Sales`)
tsplot(fit$residuals,main = "residual series")
res = fit$residuals
acf2(res)
tsplot(diff(res),main = "differenced residual series")
acf2(diff(res),max.lag = 48)
acf2(diff(diff(res,12)),max.lag = 48)
# MODEL FITTING FOR RESIDUALS
sarima(res,p=0,d=1,q=1,P=2,D=1,Q=0,S=12)

# CHECKING THE SAME MODEL FOR GUN SALES
sarima((my_data$`Gun Sales`),p=0,d=1,q=1,P=2,D=1,Q=0,S=12)
# CHECKING SAME MODEL FOR DEATHS
w3 = sarima((my_data$Deaths),p=0,d=1,q=1,P=2,D=1,Q=0,S=12)

# Running weighted least squares on the regression model for autocorrelated errors using the model specified
# for residuals
sarima((my_data$Deaths),p=0,d=1,q=1,P=2,D=1,Q=0,S=12,xreg = cbind(my_data$`Gun Sales`))

# Comparing information criterions of  original fitted model for deaths along with this model
c(w2$AIC,w2$AICc,w2$BIC)
c(w3$AIC,w3$AICc,w3$BIC)
```