TIME SERIES ANALYSIS FINAL PROJECT

Akshat Vijayvargia

S3826627

AKSHAT VIJAYVARGIA(S3826627)

INTRODUCTION

India is democratic country with dedicated Government body for all walks of life. There is one body for protecting rights for children - National Commission for Protection of Child Rights (NCPCR), but are the crimes against children are falling or not?

The purpose of this investigation is to understand the trajectory of crimes happen against children and predict the numbers for 10 years specifically in Rajasthan, India.

OBJECTIVE

The objective of this analysis is to analyze the number of crimes happen against children by using time series models and choosing the best possible model for our forecast for next 10 years.

METHODOLOGY

I briefly have breakdown the steps to have the essence of the report-

Data Preprocessing- In this step, I have imported the dataset and named the rows from 2001 to 2012. I have converted this dataset into time series format as we have to predict the future.

Model Selection- This is the most important section, where I will analyze Linear Trend Model, Quadratic Trend Model, and ARIMA Model using diagnostic checks. Then will evaluate them using CSS(Conditional Sum of Squares), ML(Maximum Likelihood Estimation), and BIC-AIC values.

Prediction- In this section, I will select the best model and predict next 10 years of crimes against children in Rajasthan on the same.

OUTPUTS AND INTERPRETATION

DESCRIPTIVE ANALYSIS

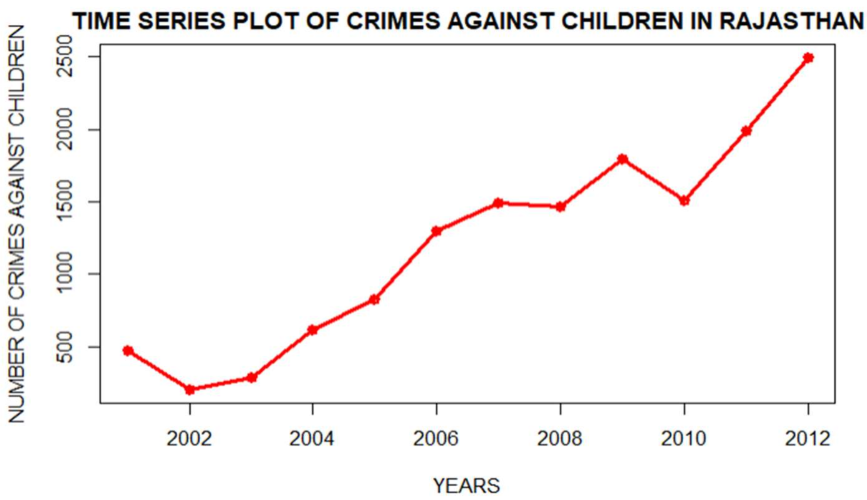


Figure 1- Time series plot of crimes against children in Rajasthan

AKSHAT VIJAYVARGIA(S3826627)

From above time series plot, we observe-

1) <u>Trend</u> – Minor dips, but the overall trend of the model is in upward fashion.
2) <u>Seasonality</u>- No obvious seasonality
3) <u>Intervention</u>- Though the movements are upward-downward, there might be an intervention because after year 2010 there is sharp elevation.
4) <u>Variance</u>- Not a major change in variance
5) <u>Behavior</u>- Moving average behavior due to succeeding datapoints.

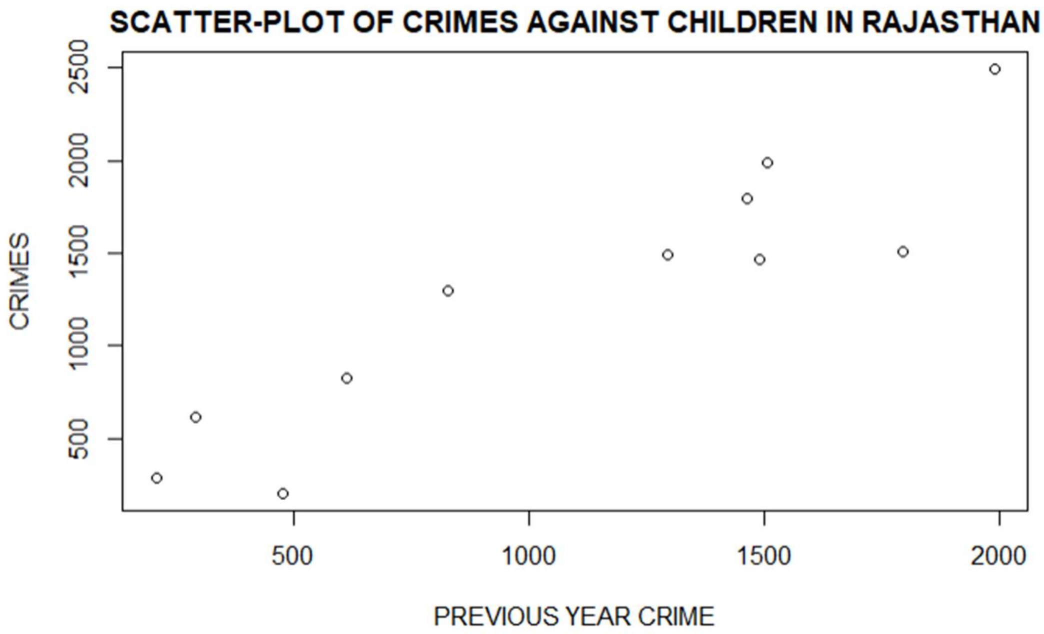


Figure 2-Scatterplot of crimes against children in Rajasthan

```
```{r}
cor(TS_CRIME[2:length(zlag(TS_CRIME))],zlag(TS_CRIME)[2:length(zlag(TS_CRIME))])
```

[1] 0.9204083
```

The correlation factor of 92%(approx) and scatterplot shows high positive correlation between suceeding points.

AKSHAT VIJAYVARGIA(S3826627)

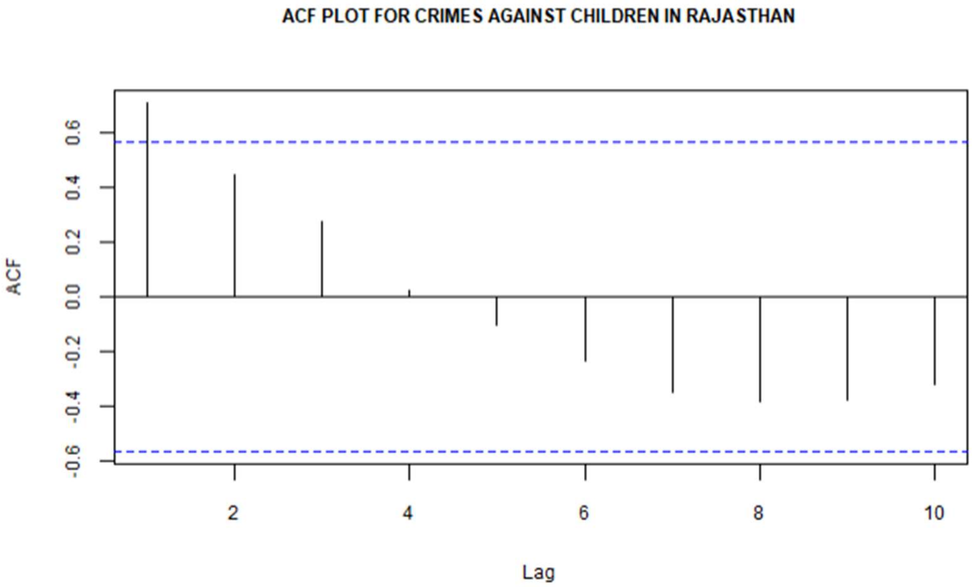


Figure 3- ACF plot for crimes against children in Rajasthan

ACF plot has a wavelike slow decaying pattern and significant autocorrelation in the very first lag, but there is no negative autocorrelation.

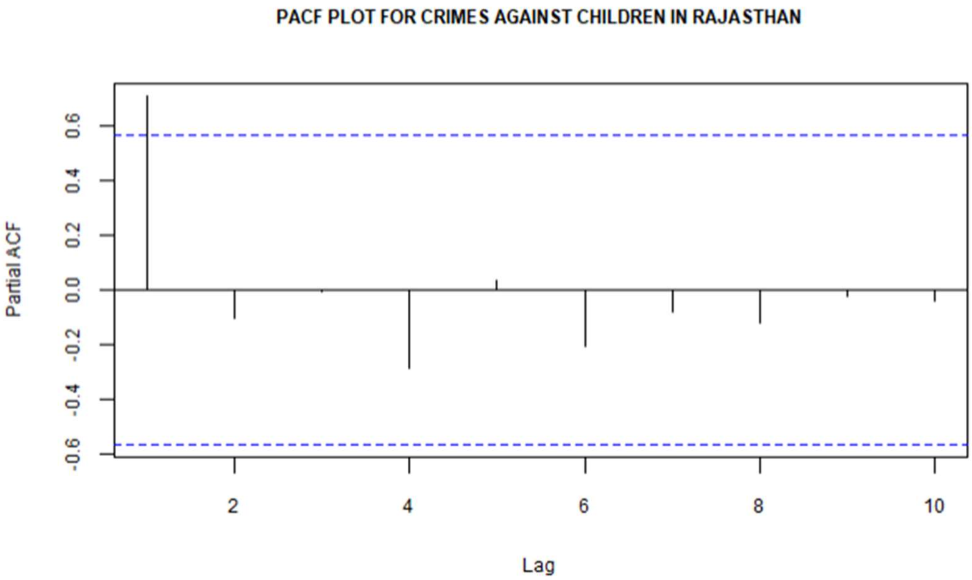


Figure 4-Partial ACF plot of crimes against children in Rajasthan

In partial ACF, we can again observe significant autocorrelation at lag 1, and no negative autocorrelation.

```
        Augmented Dickey-Fuller Test

data:  TS_CRIME
Dickey-Fuller = -1.9936, Lag order = 2, p-value = 0.5748
alternative hypothesis: stationary
```

Augmented Dickey-Fuller Test

H0- The series is not stationary.

HA- The series is stationary.

Since, our p-value is greater than 0.05, we do not have enough evidence to reject null hypothesis. Hence, the series is non-stationary. I will use data transformation to convert the data into stationary.

DATA TRANSFORMATION

To transform the data, I have used BoxCox transformation with yule-walker method.

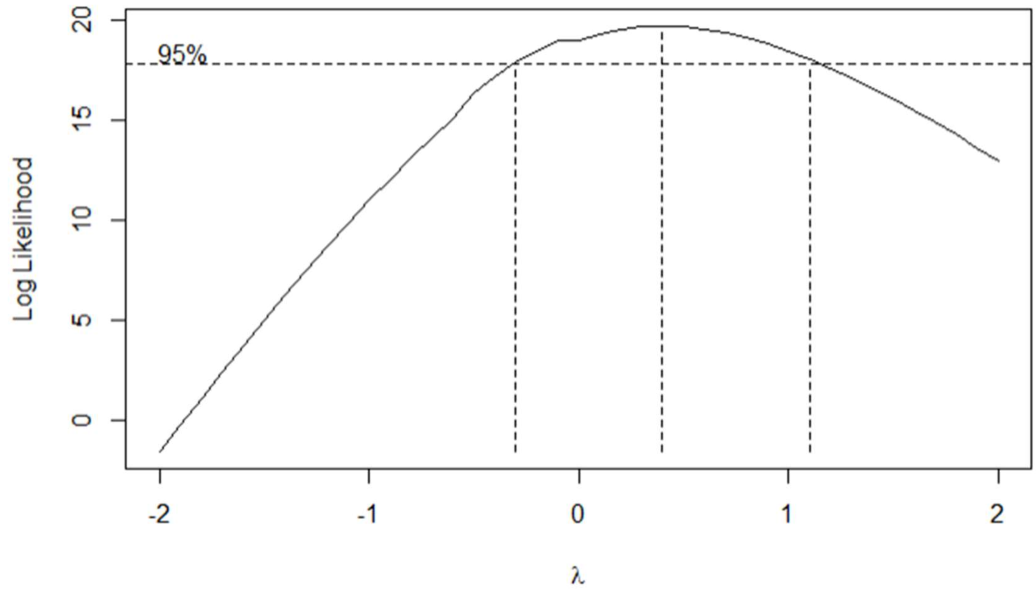


Figure 5-BoxCox transformation

```
BC_CRIME$ci

## [1] -0.3  1.1

lambda= 0.4
```

The intervals are – (-0.3,1.1), so we will go for the midpoint of these intervals i.e. 0.4.

AKSHAT VIJAYVARGIA(S3826627)

Let us, check the normality of the transformed data-

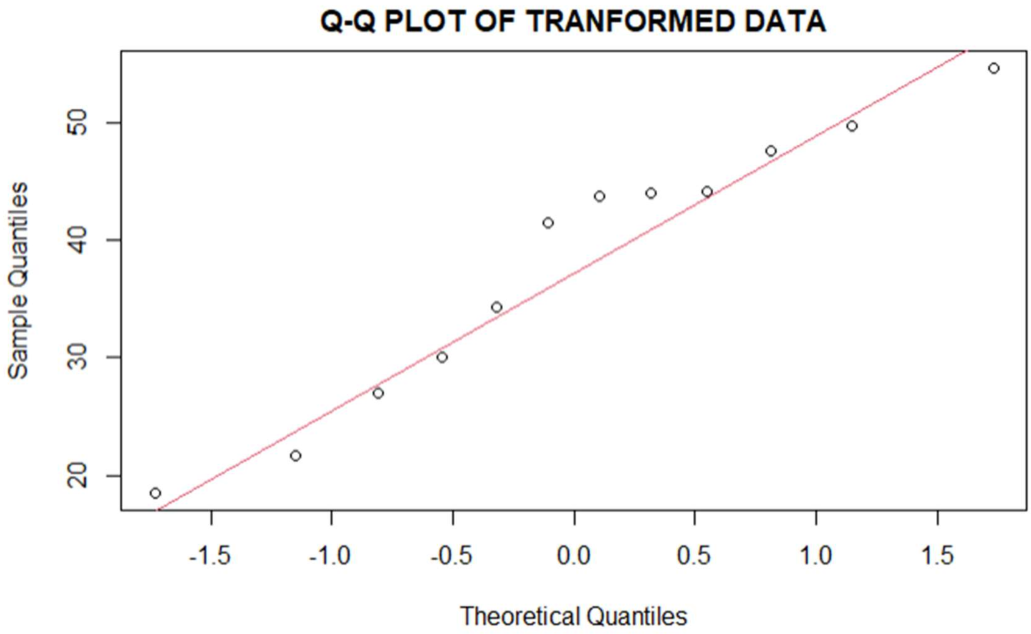


Figure 6-Q-Q Plot of BoxCox transformed data

```
        Shapiro-Wilk normality test

data:  BC_CRIME
W = 0.93919, p-value = 0.4876
```

Shapiro-Wilk Normality Test

H0- The series is normally distributed.

HA- The series is not normally distributed.

Since, our p-value is greater than 0.05, we fail to reject null hypothesis and hence, can conclude that the series is normally distributed.

From Fig6, we see that most of the dots are aligned with the red line.

DIFFERENCING

1st Differencing

```
DIFF_CRIME=diff(BC_CRIME,difference=1)
ADFTEST1=adf.test(DIFF_CRIME,alternative = 'stationary')
ADFTEST1

##
##  Augmented Dickey-Fuller Test
##
## data:  DIFF_CRIME
```

```
## Dickey-Fuller = -0.6368, Lag order = 2, p-value = 0.9636
## alternative hypothesis: stationary
```

Since, our p-value is greater than 0.05, we do not have enough evidence to reject null hypothesis. Hence, the series is non-stationary.

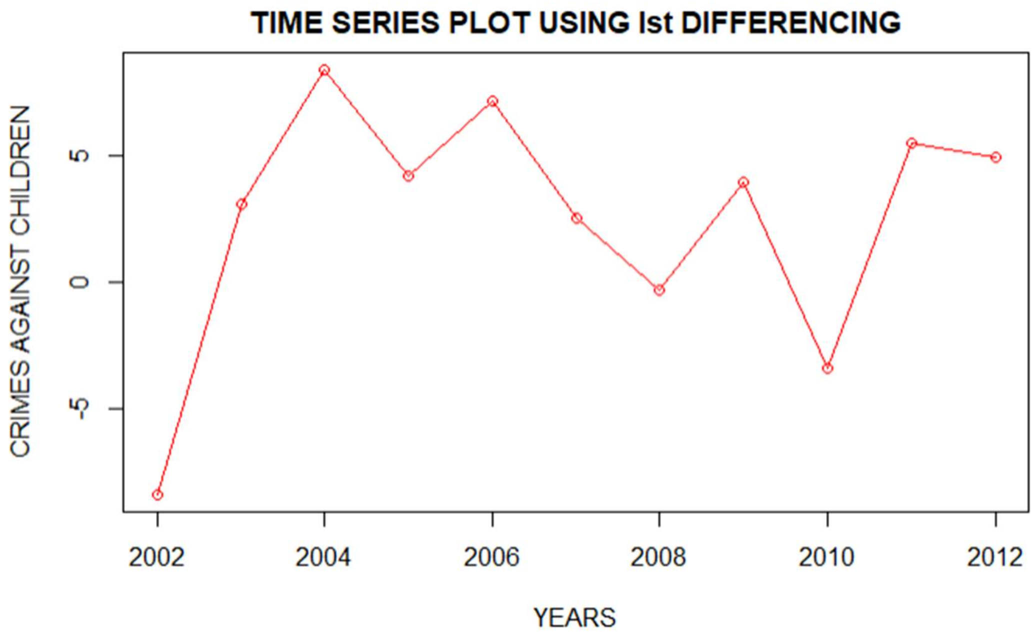


Figure 7-First Differencing

Trend is still present in our series.

2nd Differencing

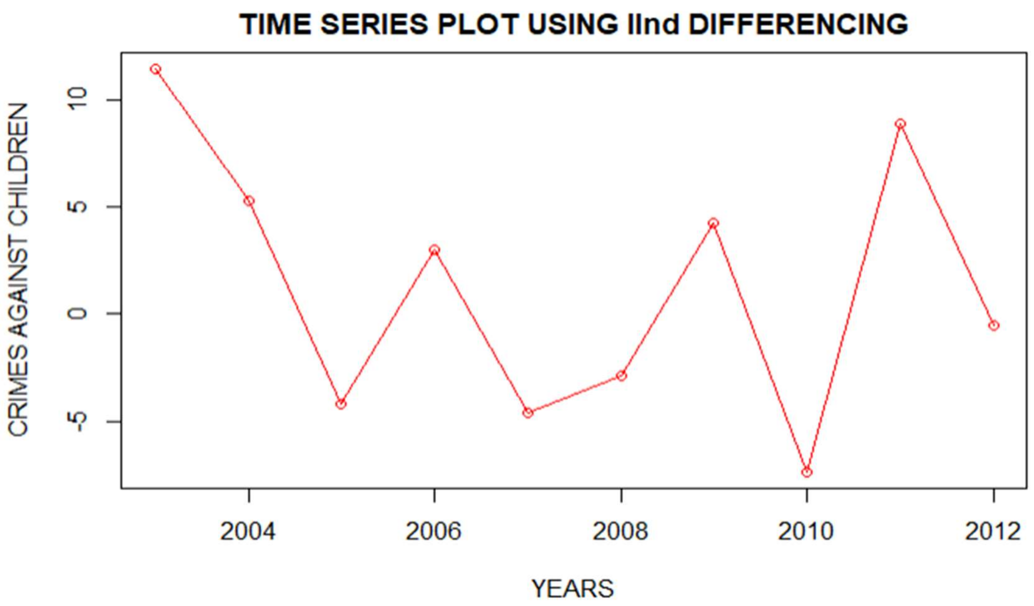


Figure 8-Second Differencing

AKSHAT VIJAYVARGIA(S3826627)

There seems to be no trend or seasonality existing, we can further verify this by ADF test for stationarity.

```
DIFF_CRIME2=diff(BC_CRIME,difference=2)
ADFTEST2=adf.test(DIFF_CRIME2,alternative = 'stationary')

ADFTEST2

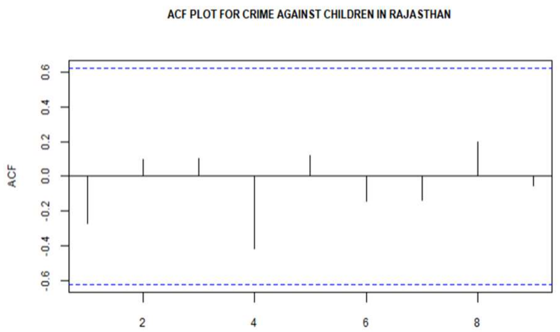
##
##  Augmented Dickey-Fuller Test
##
## data:  DIFF_CRIME2
## Dickey-Fuller = -8.4575, Lag order = 2, p-value = 0.01
## alternative hypothesis: stationary
```

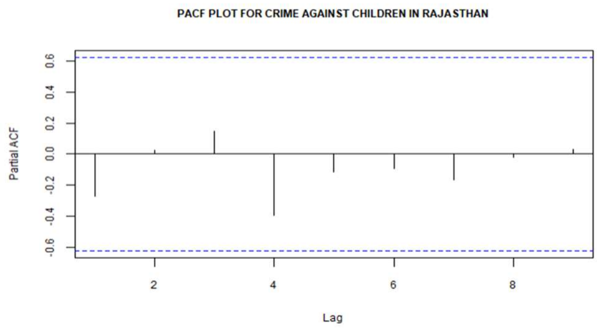
Since, our p-value is less than 0.05, we have enough evidence to reject null hypothesis. Hence, the series is now stationary.

MODEL SELECTION

POSSIBLE CANDIDATE MODELS

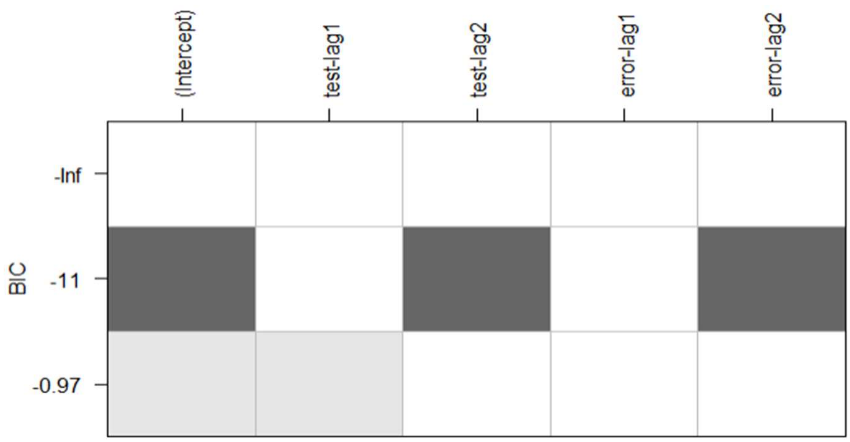
We will plot ACF, Partial ACF, EACF and BIC plot for determining possible ARIMA models.





In ACF and Partial ACF, there is no significant lag, so we cannot figure out values for p and q from them.

```
AR/MA
  0 1 2
0 o o o
1 o o o
2 o o o
```



AKSHAT VIJAYVARGIA(S3826627)

From BIC and EACF, possible ARIMA models are-

ARIMA(0,2,1), ARIMA(0,2,2),ARIMA(1,2,0),ARIMA(1,2,1),

ARIMA(1,2,2), ARIMA(2,2,0),ARIMA(2,2,1),ARIMA(2,2,2)

PARAMETER ESTIMATION

For estimation, I have used CSS (Conditional Sum Of Squares) and ML(Maximum Likelihood Estimation) on each model to figure whether they are significant or not.

1) ARIMA (0,2,1)

```
MODEL.021_ML<-arima(BC_CRIME,order=c(0,2,1),method='ML')
coeftest(MODEL.021_ML)

##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ma1 -0.99999    0.36346 -2.7513 0.005935 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

MODEL.021_CSS<-arima(BC_CRIME,order=c(0,2,1),method='CSS')
coeftest(MODEL.021_CSS)

##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ma1 -0.22861    0.30456 -0.7506   0.4529
```

2) ARIMA (0,2,2)

```
MODEL.022_ML<-arima(BC_CRIME,order=c(0,2,2),method='ML')
coeftest(MODEL.022_ML)

##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ma1 -0.82545    0.63467 -1.3006   0.1934
## ma2 -0.17453    0.52441 -0.3328   0.7393

MODEL.022_CSS<-arima(BC_CRIME,order=c(0,2,2),method='CSS')
coeftest(MODEL.022_CSS)

##
## z test of coefficients:
##
```

```
##     Estimate Std. Error z value Pr(>|z|)
## ma1 -0.20398    0.30382 -0.6714   0.5020
## ma2  0.23328    0.48300  0.4830   0.6291
```

3) ARIMA (1,2,0)

```
MODEL.120_ML<-arima(BC_CRIME,order=c(1,2,0),method='ML')
coeftest(MODEL.120_ML)

##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1 -0.32431    0.34503 -0.9399   0.3472

MODEL.120_CSS<-arima(BC_CRIME,order=c(1,2,0),method='CSS')
coeftest(MODEL.120_CSS)

##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1 -0.24145    0.24196 -0.9979   0.3183
```

4) ARIMA (1,2,1)

```
MODEL.121_ML<-arima(BC_CRIME,order=c(1,2,1),method='ML')
coeftest(MODEL.121_ML)

##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1  0.12682    0.48398  0.2620 0.793294
## ma1 -0.99996    0.36662 -2.7275 0.006381 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

MODEL.121_CSS<-arima(BC_CRIME,order=c(1,2,1),method='CSS')

coeftest(MODEL.121_CSS)

##
## z test of coefficients:
##
##       Estimate Std. Error  z value  Pr(>|z|)
## ar1  0.2475166  0.0025477   97.154 < 2.2e-16 ***
## ma1 -1.9719187  0.0196100 -100.557 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

5) ARIMA (1,2,2)

```
MODEL.122_ML<-arima(BC_CRIME,order=c(1,2,2),method='ML')
coeftest(MODEL.122_ML)

##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1 -0.56840    0.48057 -1.1828  0.23690
## ma1 -0.16982    0.52166 -0.3255  0.74477
## ma2 -0.83007    0.48545 -1.7099  0.08728 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

MODEL.122_CSS<-arima(BC_CRIME,order=c(1,2,2),method='CSS')
coeftest(MODEL.122_CSS)

##
## z test of coefficients:
##
##      Estimate Std. Error z value  Pr(>|z|)
## ar1  0.299458   0.079728  3.7560 0.0001727 ***
## ma1 -0.827743   0.206879 -4.0011 6.305e-05 ***
## ma2  1.225334   0.168102  7.2892 3.117e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

6) ARIMA (2,2,0)

```
MODEL.220_ML<-arima(BC_CRIME,order=c(2,2,0),method='ML')
coeftest(MODEL.220_ML)

##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1 -0.49429    0.47857 -1.0329   0.3017
## ar2 -0.27308    0.52772 -0.5175   0.6048

MODEL.220_CSS<-arima(BC_CRIME,order=c(2,2,0),method='CSS')
coeftest(MODEL.220_CSS)

##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1 -0.67609    0.22767 -2.9695 0.002982 **
## ar2 -0.12054    0.20618 -0.5846 0.558788
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

7) ARIMA (2,2,1)

```
MODEL.221_ML<-arima(BC_CRIME,order=c(2,2,1),method='ML')

coeftest(MODEL.221_ML)

##
## z test of coefficients:
##
##       Estimate Std. Error z value Pr(>|z|)
## ar1  0.075891   0.485368  0.1564 0.875751
## ar2 -0.158238   0.431130 -0.3670 0.713596
## ma1 -1.000000   0.351740 -2.8430 0.004469 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

MODEL.221_CSS<-arima(BC_CRIME,order=c(2,2,1),method='CSS')
coeftest(MODEL.221_CSS)

##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ar1 -0.61764    0.39303 -1.5715   0.1161
## ar2 -0.10447    0.21087 -0.4954   0.6203
## ma1 -0.11484    0.60842 -0.1888   0.8503
```

8) ARIMA(2,2,2)

```
MODEL.222_ML<-arima(BC_CRIME,order=c(2,2,2),method='ML')
coeftest(MODEL.222_ML)

##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ar1 -0.66688    0.48027 -1.3886  0.16497
## ar2 -0.25328    0.49503 -0.5116  0.60890
## ma1 -0.22483    0.51163 -0.4394  0.66034
## ma2 -0.77517    0.46329 -1.6732  0.09429 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

MODEL.222_CSS<-arima(BC_CRIME,order=c(2,2,2),method='CSS')

coeftest(MODEL.222_CSS)

##
## z test of coefficients:
##
```

```
##       Estimate Std. Error  z value  Pr(>|z|)
## ar1 -0.7656581  0.0307618 -24.8899 < 2.2e-16 ***
## ar2  0.0608094  0.0037129  16.3777 < 2.2e-16 ***
## ma1 -0.9489376  0.1080291  -8.7841 < 2.2e-16 ***
## ma2  4.0307264  0.3323938  12.1264 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ARIMA(1,2,1)-CSS, ARIMA(1,2,2)-CSS AND ARIMA(2,2,2) have all coefficients significant. MA(1) is significant in ML method.

Let us sort all models according to AIC and BIC values to facilitate the decision for choosing the best model.

AIC scores-

```
##              df      AIC
## MODEL.021_ML  2 66.49612
## MODEL.120_ML  2 67.63244
## MODEL.022_ML  3 68.39235
## MODEL.121_ML  3 68.42523
## MODEL.220_ML  3 69.38852
## MODEL.122_ML  4 69.89451
## MODEL.221_ML  4 70.29962
## MODEL.222_ML  5 71.67504
```

BIC scores-

```
##              df      BIC
## MODEL.021_ML  2 67.10129
## MODEL.120_ML  2 68.23761
## MODEL.022_ML  3 69.30010
## MODEL.121_ML  3 69.33298
## MODEL.220_ML  3 70.29627
## MODEL.122_ML  4 71.10485
## MODEL.221_ML  4 71.50996
## MODEL.222_ML  5 73.18796
```

Now, the model with least scores will the best model for our prediction. Here, our best model is Model(0,2,1).

AKSHAT VIJAYVARGIA(S3826627)

DIAGNOSTIC CHECKS

For diagnostics, I will perform residual analysis on the best model, which is Model(0,2,1). In this section, I will analysis the behavior of ACF plot, Parameter estimate and likelihood plot and residual histogram to understand the normality and lastly Ljung Box test for independence of residuals.

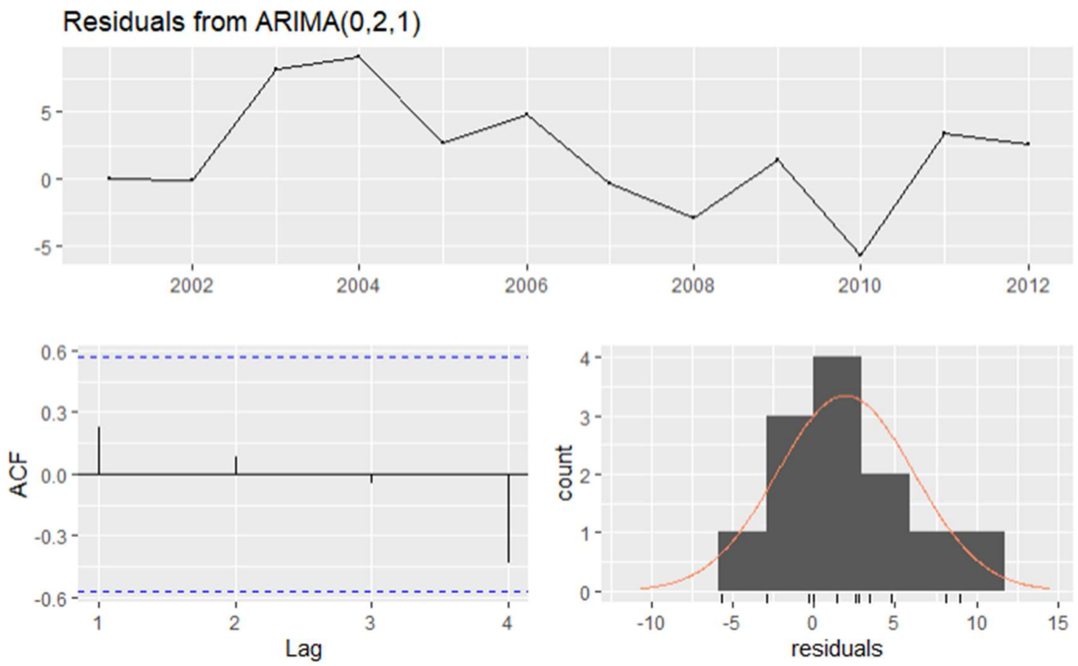


Figure 9-Residual Analysis of Model(0,2,1)

From above plots, we can vaguely say that our model meets the assumptions and might be fit for our data. But let us dig in-

In the ACF plot, we observe that there are no significant lags, which is a very good sign because it means that residuals are independent.

The Histogram of Residuals shows the normality of residuals, here, we can say that they are approximately normally distributed.

Now let us perform Ljung Box Test to authenticate our observation.

```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(0,2,1)
## Q* = 4.8622, df = 3, p-value = 0.1822
##
## Model df: 1.   Total lags used: 4
```

Ljung Box Test

H0- The residuals are independently distributed.

HA- The residuals are not independently distributed.

Since, our p-value is greater than 0.05, we do not have enough evidence to reject the null hypothesis and can conclude that the residuals for our timeseries model are independent.

FORECASTING CRIMES AGAINST CHILDREN IN RAJASTHAN FOR NEXT 10 YEARS

```
##      Point Forecast    Lo 80     Hi 80     Lo 95     Hi 95
## 2013       2780.050 2045.261  3653.969 1709.7984  4175.602
## 2014       3082.661 1974.700  4498.281 1503.8487  5380.536
## 2015       3404.221 1952.757  5362.493 1371.5441  6620.968
## 2016       3745.109 1952.736  6275.246 1272.1610  7941.070
## 2017       4105.699 1965.288  7247.780 1191.4475  9358.439
## 2018       4486.355 1986.039  8286.312 1122.7290 10883.001
## 2019       4887.437 2012.600  9395.029 1062.3603 12521.569
## 2020       5309.297 2043.536 10577.125 1008.1459 14279.424
## 2021       5752.283 2077.922 11835.241  958.6697 16160.995
## 2022       6216.735 2115.133 13171.685  912.9700 18170.185
```

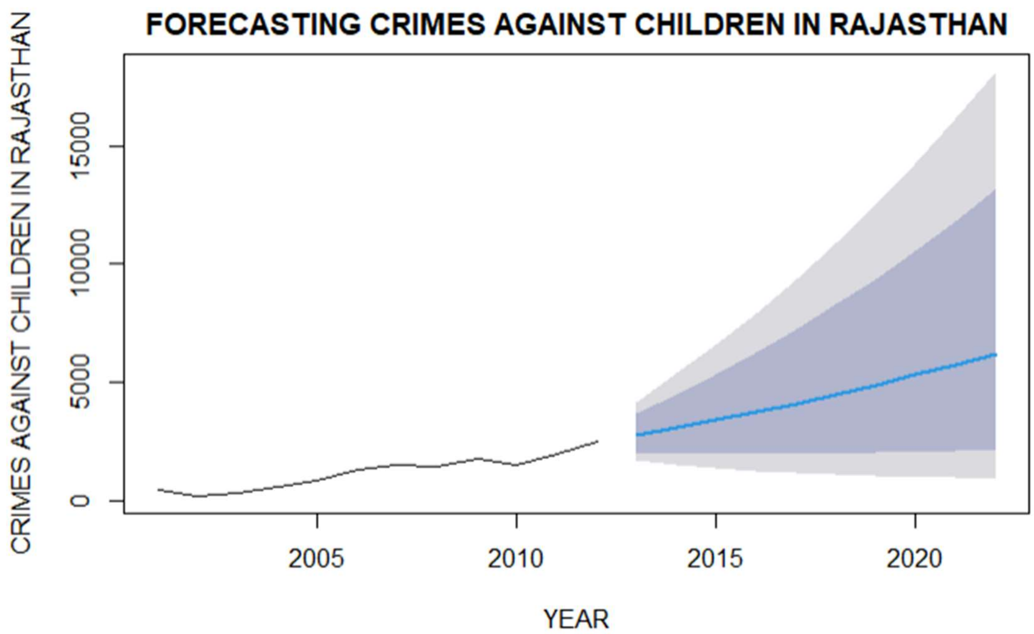


Figure 10- 10 years ahead forecast of crimes against children in Rajasthan

As per the forecast of Model(0,2,1), the cases of kidnapping, abduction, infanticide etc. are going to rise against children in coming years.

AKSHAT VIJAYVARGIA(S3826627)

CONCLUSION

After considering EACF, ACF, PACF, BIC plot, AIC-BIC values, and residual checks, we conclude that ARIMA (0,2,1) is the promising model for our prediction. If we observe Figure 10, we can clearly see the elevation in the number of crimes against children in Rajasthan.

The strength of this investigation is that the dataset is retrieved from Data.world, which is the data repository. Moreover, this dataset is also available at Open Government Data [Open Government Data (OGD) Platform India] So we are pretty sure that the data is authentic and pure.

The limitation of this investigation is that we are restricted and cannot introduce other features to better understand the situation and reasons behind this increase.

REFERNCES

1) RMIT Study Materials. Available at - Course modules: Time Series Analysis (2110) (instructure.com) [Accessed 2021-6-2]

2) Dataset of Crimes Against Children. Data.World. Available at - Crime rate against Children-India 2001-2012 - dataset by bhavnachawla | data.world [Accessed 2021-5-29]

AKSHAT VIJAYVARGIA(S3826627)

APPENDIX

```
library(readxl)
library(tseries)

library(lmtest)

library(TSA)

library(fUnitRoots)

library(forecast)

CRIME <- read_excel("C:/Users/61422/Desktop/CRIME AGAINST CHILDREN IN RAJASTH
AN.xlsx")

rownames(CRIME)<-seq(from=2001, to=2012)

TS_CRIME=ts(as.vector(CRIME['CRIMES']),start=2001,end=2012)
class(TS_CRIME)

plot(TS_CRIME,ylab='NUMBER OF CRIMES AGAINST CHILDREN',xlab='YEARS',type='o',
col = c("red"),lwd=3,main = 'TIME SERIES PLOT OF CRIMES AGAINST CHILDREN IN R
AJASTHAN')

hist(CRIME$CRIMES,xlab = 'CRIMES AGAINST CHILDREN',main ='HISTOGRAM OF CRIMES
AGAINST CHILDREN IN RAJASTHAN' ,ylim=c(0,3.5))

plot(y=TS_CRIME,x=zlag(TS_CRIME),ylab='CRIMES', xlab='PREVIOUS YEAR CRIME',ma
in ='SCATTER-PLOT OF CRIMES AGAINST CHILDREN IN RAJASTHAN')

par(mar=c(5,4,5,4),mfrow=c(1,1),cex.main=0.75, cex.lab=0.75, cex.axis=0.75)

acf(TS_CRIME, main = 'ACF PLOT FOR CRIMES AGAINST CHILDREN IN RAJASTHAN')

pacf(TS_CRIME, main = 'PACF PLOT FOR CRIMES AGAINST CHILDREN IN RAJASTHAN')

cor(TS_CRIME[2:length(zlag(TS_CRIME))],zlag(TS_CRIME)[2:length(zlag(TS_CRIME)
)])

par(mar=c(5,4,5,4),mfrow=c(1,2),cex.main=0.75, cex.lab=0.75, cex.axis=0.75)

acf(TS_CRIME, main = 'ACF PLOT FOR CRIME AGAINST CHILDREN IN RAJASTHAN')
pacf(TS_CRIME, main = 'PACF PLOT FOR CRIME AGAINST CHILDREN IN RAJASTHAN')

adf.test(TS_CRIME,alternative = 'stationary')

BC_CRIME=BoxCox.ar(TS_CRIME,method = 'yule-walker')

BC_CRIME$ci

lambda= 0.4
BC_CRIME=(TS_CRIME^lambda-1)/lambda
```

AKSHAT VIJAYVARGIA(S3826627)

```
qqnorm(BC_CRIME)
qqline(BC_CRIME,col=2)

DIFF_CRIME=diff(BC_CRIME,difference=1)
ADFTEST1=adf.test(DIFF_CRIME,alternative = 'stationary')
ADFTEST1

##
##  Augmented Dickey-Fuller Test
##
## data:  DIFF_CRIME
## Dickey-Fuller = -0.6368, Lag order = 2, p-value = 0.9636
## alternative hypothesis: stationary

plot(DIFF_CRIME,type='o',ylab='CRIMES AGAINST CHILDREN',col='red',xlab='YEARS
',main='TIME SERIES PLOT USING Ist DIFFERENCING')

DIFF_CRIME2=diff(BC_CRIME,difference=2)
ADFTEST2=adf.test(DIFF_CRIME2,alternative = 'stationary')

ADFTEST2

##
##  Augmented Dickey-Fuller Test
##
## data:  DIFF_CRIME2
## Dickey-Fuller = -8.4575, Lag order = 2, p-value = 0.01
## alternative hypothesis: stationary

plot(DIFF_CRIME2,type='o',ylab='CRIMES AGAINST CHILDREN',col='red',xlab='YEAR
S',main='TIME SERIES PLOT USING IInd DIFFERENCING')

par(mar=c(5,4,5,4),mfrow=c(1,1),cex.main=0.75, cex.lab=0.75, cex.axis=0.75)

acf(DIFF_CRIME2, main = 'ACF PLOT FOR CRIME AGAINST CHILDREN IN RAJASTHAN')

pacf(DIFF_CRIME2, main = 'PACF PLOT FOR CRIME AGAINST CHILDREN IN RAJASTHAN')

eacf(DIFF_CRIME2, ar.max = 2, ma.max =2)

## AR/MA
##   0 1 2
## 0 o o o
## 1 o o o
## 2 o o o

ARMA_CRIME<-armasubsets(y=DIFF_CRIME2,nar=2,nma=2,y.name='test',ar.method='ol
s')

plot(ARMA_CRIME)

MODEL.022_ML<-arima(BC_CRIME,order=c(0,2,2),method='ML')
coeftest(MODEL.022_ML)
```

```
##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ma1 -0.82545    0.63467 -1.3006   0.1934
## ma2 -0.17453    0.52441 -0.3328   0.7393
```

```
MODEL.022_CSS<-arima(BC_CRIME,order=c(0,2,2),method='CSS')
coeftest(MODEL.022_CSS)
```

```
##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ma1 -0.20398    0.30382 -0.6714   0.5020
## ma2  0.23328    0.48300  0.4830   0.6291
```

```
MODEL.222_ML<-arima(BC_CRIME,order=c(2,2,2),method='ML')
coeftest(MODEL.222_ML)
```

```
##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1 -0.66688    0.48027 -1.3886  0.16497
## ar2 -0.25328    0.49503 -0.5116  0.60890
## ma1 -0.22483    0.51163 -0.4394  0.66034
## ma2 -0.77517    0.46329 -1.6732  0.09429 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
MODEL.222_CSS<-arima(BC_CRIME,order=c(2,2,2),method='CSS')

coeftest(MODEL.222_CSS)
```

```
##
## z test of coefficients:
##
##       Estimate Std. Error  z value  Pr(>|z|)
## ar1 -0.7656581  0.0307618 -24.8899 < 2.2e-16 ***
## ar2  0.0608094  0.0037129  16.3777 < 2.2e-16 ***
## ma1 -0.9489376  0.1080291  -8.7841 < 2.2e-16 ***
## ma2  4.0307264  0.3323938  12.1264 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
MODEL.021_ML<-arima(BC_CRIME,order=c(0,2,1),method='ML')
coeftest(MODEL.021_ML)
```

```
##
## z test of coefficients:
##
```

AKSHAT VIJAYVARGIA(S3826627)

```
##     Estimate Std. Error z value Pr(>|z|)
## ma1 -0.99999    0.36346 -2.7513 0.005935 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
MODEL.021_CSS<-arima(BC_CRIME,order=c(0,2,1),method='CSS')
coeftest(MODEL.021_CSS)
```

```
##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ma1 -0.22861    0.30456 -0.7506   0.4529
```

```
MODEL.120_ML<-arima(BC_CRIME,order=c(1,2,0),method='ML')
coeftest(MODEL.120_ML)
```

```
##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1 -0.32431    0.34503 -0.9399   0.3472
```

```
MODEL.120_CSS<-arima(BC_CRIME,order=c(1,2,0),method='CSS')
coeftest(MODEL.120_CSS)
```

```
##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1 -0.24145    0.24196 -0.9979   0.3183
```

```
MODEL.122_ML<-arima(BC_CRIME,order=c(1,2,2),method='ML')
coeftest(MODEL.122_ML)
```

```
##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1 -0.56840    0.48057 -1.1828  0.23690
## ma1 -0.16982    0.52166 -0.3255  0.74477
## ma2 -0.83007    0.48545 -1.7099  0.08728 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
MODEL.122_CSS<-arima(BC_CRIME,order=c(1,2,2),method='CSS')
coeftest(MODEL.122_CSS)
```

```
##
## z test of coefficients:
##
##      Estimate Std. Error z value  Pr(>|z|)
```

```
## ar1  0.299458   0.079728  3.7560 0.0001727 ***
## ma1 -0.827743   0.206879 -4.0011 6.305e-05 ***
## ma2  1.225334   0.168102  7.2892 3.117e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

MODEL.121_ML<-arima(BC_CRIME,order=c(1,2,1),method='ML')
coeftest(MODEL.121_ML)

##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1  0.12682    0.48398  0.2620 0.793294
## ma1 -0.99996    0.36662 -2.7275 0.006381 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

MODEL.121_CSS<-arima(BC_CRIME,order=c(1,2,1),method='CSS')

coeftest(MODEL.121_CSS)

##
## z test of coefficients:
##
##       Estimate Std. Error   z value  Pr(>|z|)
## ar1  0.2475166  0.0025477   97.154 < 2.2e-16 ***
## ma1 -1.9719187  0.0196100 -100.557 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

MODEL.220_ML<-arima(BC_CRIME,order=c(2,2,0),method='ML')
coeftest(MODEL.220_ML)

##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1 -0.49429    0.47857 -1.0329   0.3017
## ar2 -0.27308    0.52772 -0.5175   0.6048

MODEL.220_CSS<-arima(BC_CRIME,order=c(2,2,0),method='CSS')
coeftest(MODEL.220_CSS)

##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1 -0.67609    0.22767 -2.9695 0.002982 **
## ar2 -0.12054    0.20618 -0.5846 0.558788
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

AKSHAT VIJAYVARGIA(S3826627)

```
MODEL.221_ML<-arima(BC_CRIME,order=c(2,2,1),method='ML')
```

```
## Warning in log(s2): NaNs produced
```

```
coeftest(MODEL.221_ML)
```

```
##
## z test of coefficients:
##
##      Estimate Std. Error z value Pr(>|z|)
## ar1  0.075891   0.485368  0.1564 0.875751
## ar2 -0.158238   0.431130 -0.3670 0.713596
## ma1 -1.000000   0.351740 -2.8430 0.004469 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
MODEL.221_CSS<-arima(BC_CRIME,order=c(2,2,1),method='CSS')
coeftest(MODEL.221_CSS)
```

```
##
## z test of coefficients:
##
##     Estimate Std. Error z value Pr(>|z|)
## ar1 -0.61764    0.39303 -1.5715   0.1161
## ar2 -0.10447    0.21087 -0.4954   0.6203
## ma1 -0.11484    0.60842 -0.1888   0.8503
```

```
sort.score <- function(x, score = c("bic", "aic")){
  if (score == "aic"){
    x[with(x, order(AIC)),]
  } else if (score == "bic") {
    x[with(x, order(BIC)),]
  } else {
    warning('score = "x" only accepts valid arguments ("aic","bic")')
  }
}
```

```
sort.score(AIC(MODEL.021_ML,MODEL.022_ML,MODEL.120_ML,MODEL.121_ML,MODEL.122_
ML,MODEL.220_ML,MODEL.221_ML,MODEL.222_ML), score = "aic")
```

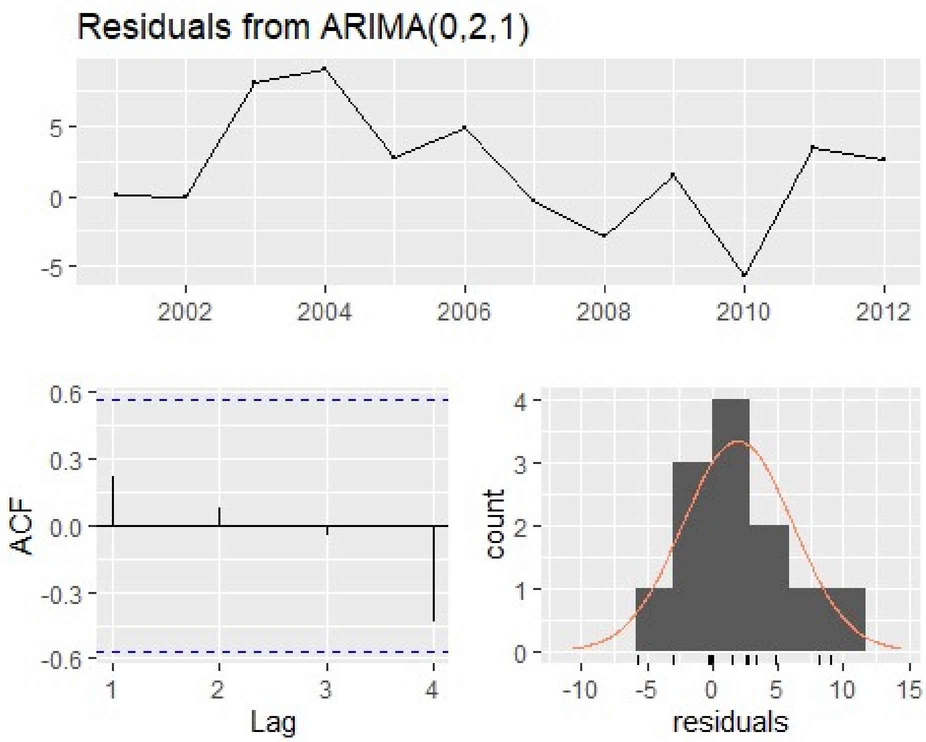
```
##                df      AIC
## MODEL.021_ML  2 66.49612
## MODEL.120_ML  2 67.63244
## MODEL.022_ML  3 68.39235
## MODEL.121_ML  3 68.42523
## MODEL.220_ML  3 69.38852
## MODEL.122_ML  4 69.89451
## MODEL.221_ML  4 70.29962
## MODEL.222_ML  5 71.67504
```

```
sort.score(BIC(MODEL.021_ML,MODEL.022_ML,MODEL.120_ML,MODEL.121_ML,MODEL.122_
ML,MODEL.220_ML,MODEL.221_ML,MODEL.222_ML), score = "bic")
```

AKSHAT VIJAYVARGIA(S3826627)

```
##                df      BIC
## MODEL.021_ML  2 67.10129
## MODEL.120_ML  2 68.23761
## MODEL.022_ML  3 69.30010
## MODEL.121_ML  3 69.33298
## MODEL.220_ML  3 70.29627
## MODEL.122_ML  4 71.10485
## MODEL.221_ML  4 71.50996
## MODEL.222_ML  5 73.18796
```

```
checkresiduals(MODEL.021_ML)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(0,2,1)
## Q* = 4.8622, df = 3, p-value = 0.1822
##
## Model df: 1.   Total lags used: 4
```

```
FORECAST_MODEL=Arima(TS_CRIME,c(0,2,1),lambda=.4)
forecast(FORECAST_MODEL,h=10)
```

```
##      Point Forecast    Lo 80    Hi 80     Lo 95    Hi 95
## 2013       2780.050 2045.261 3653.969 1709.7984 4175.602
## 2014       3082.661 1974.700 4498.281 1503.8487 5380.536
## 2015       3404.221 1952.757 5362.493 1371.5441 6620.968
```

AKSHAT VIJAYVARGIA(S3826627)

```
## 2016       3745.109 1952.736  6275.246 1272.1610  7941.070
## 2017       4105.699 1965.288  7247.780 1191.4475  9358.439
## 2018       4486.355 1986.039  8286.312 1122.7290 10883.001
## 2019       4887.437 2012.600  9395.029 1062.3603 12521.569
## 2020       5309.297 2043.536 10577.125 1008.1459 14279.424
## 2021       5752.283 2077.922 11835.241  958.6697 16160.995
## 2022       6216.735 2115.133 13171.685  912.9700 18170.185
```

```
plot(forecast(FORECAST_MODEL,h=10),xlab='YEAR',ylab='CRIMES AGAINST CHILDREN
IN RAJASTHAN',main='FORECASTING CRIMES AGAINST CHILDREN IN RAJASTHAN')
```