Time Series Analysis
MA641

Title: Time series forecasting using monthly car sales in Quebec

Guide: Prof. Hadi Safari Author: Siddhesh More

Abstract/Introduction:
In this project we investigated seasonal ARIMA model that provides a way to model time series whose seasonal tendencies are not as regular as we would have with a deterministic model. We then proceeded with the approach of time series modelling:

- Model specification
- Residual approach
- Model fitting
- Model diagnostics
- Prediction for next 10 years based on the model we drawn

Data Description:
The dataset used in this project is from Kaggle and the source of the dataset is here. This dataset is about car sales in Quebec, Canada. Data consists of three columns mainly, year, month, and sales. Year column shows the year in which the cars are sold. Similarly, month column shows the number of particular month (such as 1-> Jan, 2-> Feb and 12-> Dec etc.) in which the sales are made. The last column Sales shows the total number of cars sold. Dataset consists of the records from 1960 to 1968. This is a comparatively small dataset. There are total 108 observations.

```
cars <- read.csv("cars.csv",header=TRUE)
head(cars)
```

Insight of the dataset:

	Year <int>	Month <int>	Sales <int>
1	1960	1	6550
2	1960	2	8728
3	1960	3	12026
4	1960	4	14395
5	1960	5	14587
6	1960	6	13791

6 rows

Methodology, Analysis, and results:
As we convert the data frame into time series data and generate a plot to draw conclusions regarding trends, seasonality, and behavior.

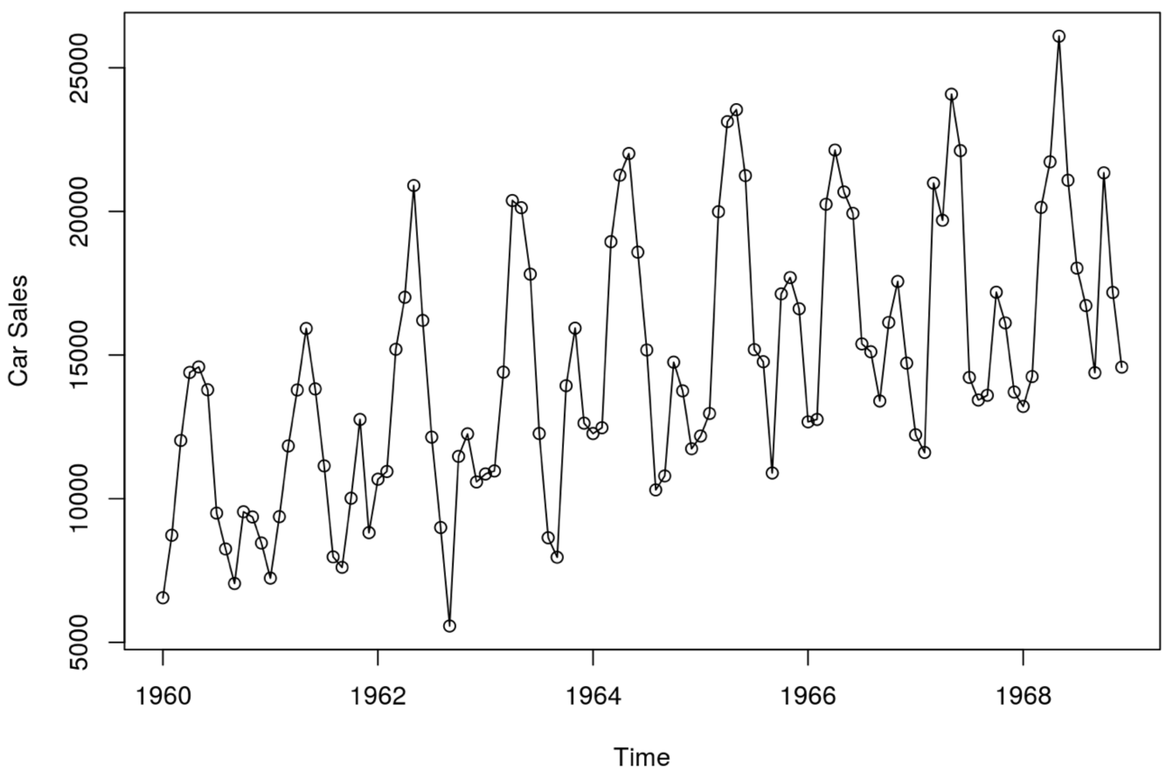
-Trend: There is strong upward trend.

-Seasonality: A seasonal pattern is rise and fall in data values that repeats after regular intervals. We can visualize it from the plot.

As we observe non stationarity due to upwards trend, we then plot ACF & PACF respectively.

```
cars.ts <- matrix(cars$Sales,nrow=108,ncol=1)

cars.ts<- as.vector(t(cars.ts))

cars.ts <- ts(cars.ts,start=c(1960,1), end=c(1968,12), frequency=12)

plot(cars.ts,type='o',ylab='Car Sales')
```

Time series plot of monthly cars sales from 1960-1968:

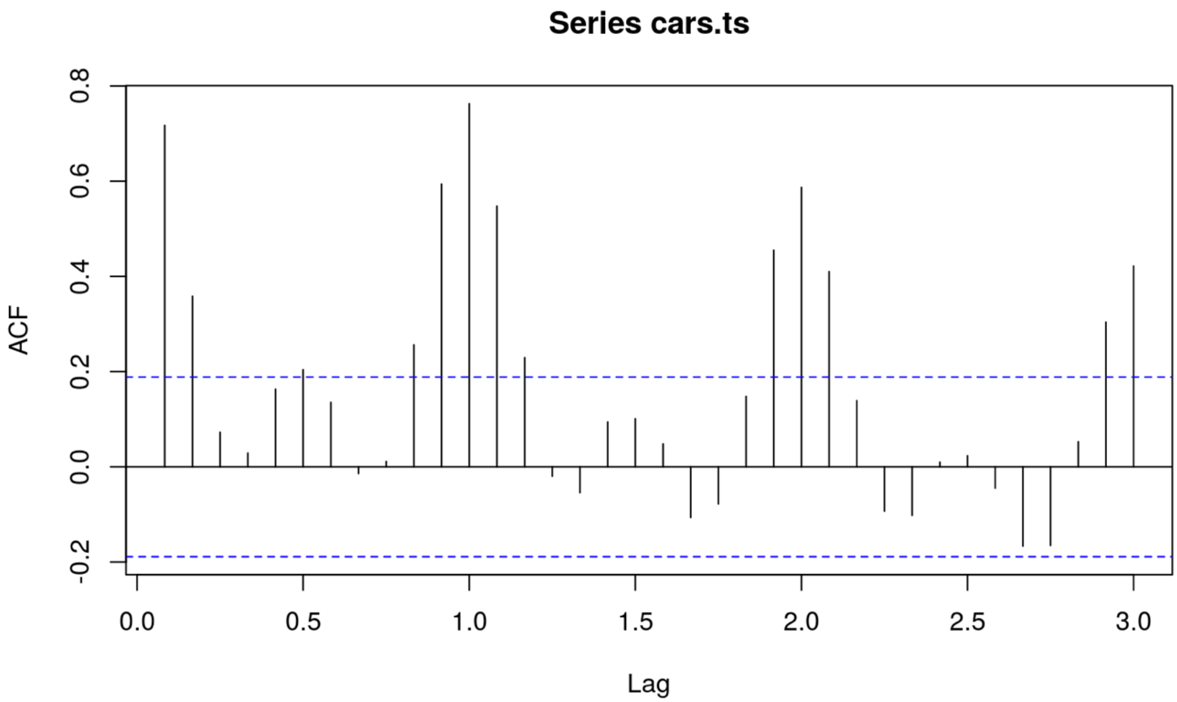


ACF plots shows correlation between elements of the time series.Lags are the correlation for time series observations with their previous time stamps.

Because we have strong correlation at lags 12,24,36 and so on we consider the existence of seasonal auto correlation relationship.

```
acf(cars.ts,lag.max = 36)
```

ACF plot for the data

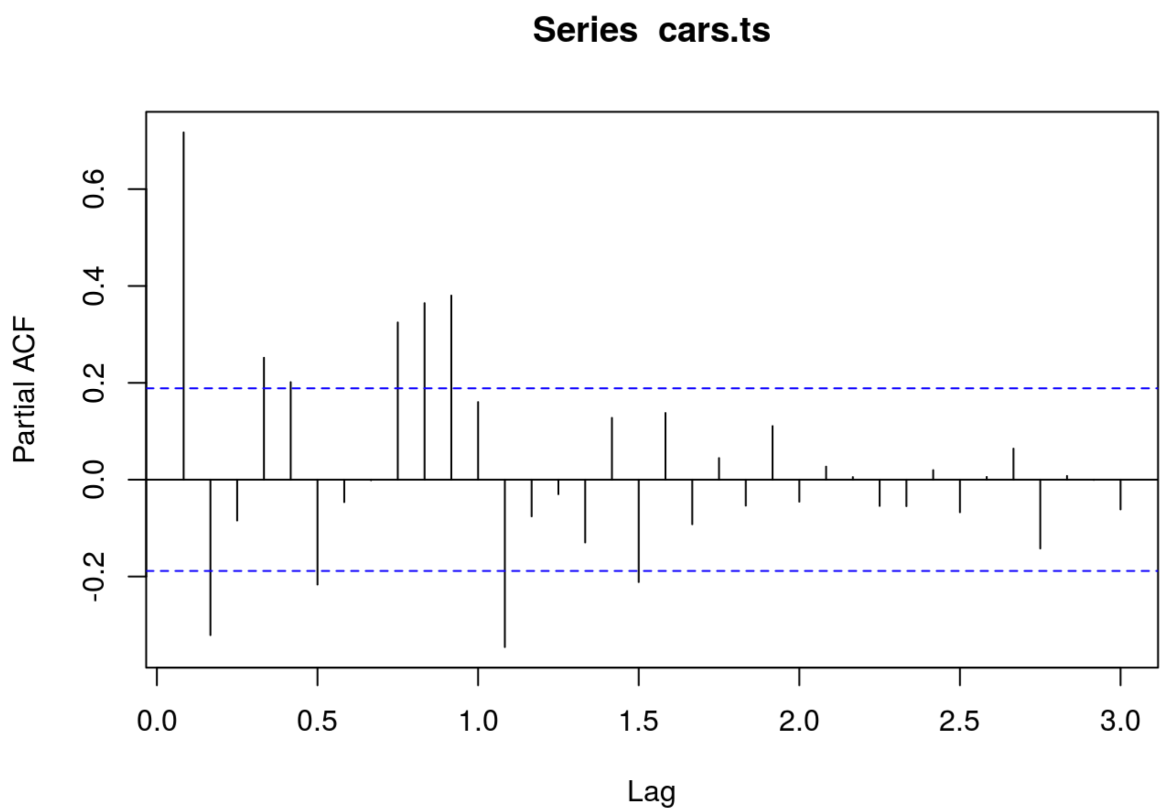


PACF is the amount of correlation between a variable and a lag of itself that is not explained by correlation at all lower-order-lags.

The PACF below shows 1 seasonal lag at 12.

```
pacf(cars.ts, lag.max = 36)
```

PACF plot of the data



Seasonality:

- Seasonality in time series is a regular pattern of change that repeats over S time periods, where s is defines the number of time periods until the pattern repeats again.
- In seasonal ARIMA model, seasonal AR and MA terms predict X_t using data values and errors at times with lags that are multiple of S (i.e. the span of seasonality).

Residual Approach:

Differencing:

It is necessary to examine differences data when we have seasonality. Seasonality usually causes series to be non-stationary because the average value at some particular time within the seasonal span may be different than the average values at other times.

Seasonal ARIMA differencing:

the seasonal ARIMA model incorporates both non-seasonal and seasonal factors in multiplicative model. One shorthand notation is :

$$ARIMA(p, d, q)x(P, D, Q)_s$$

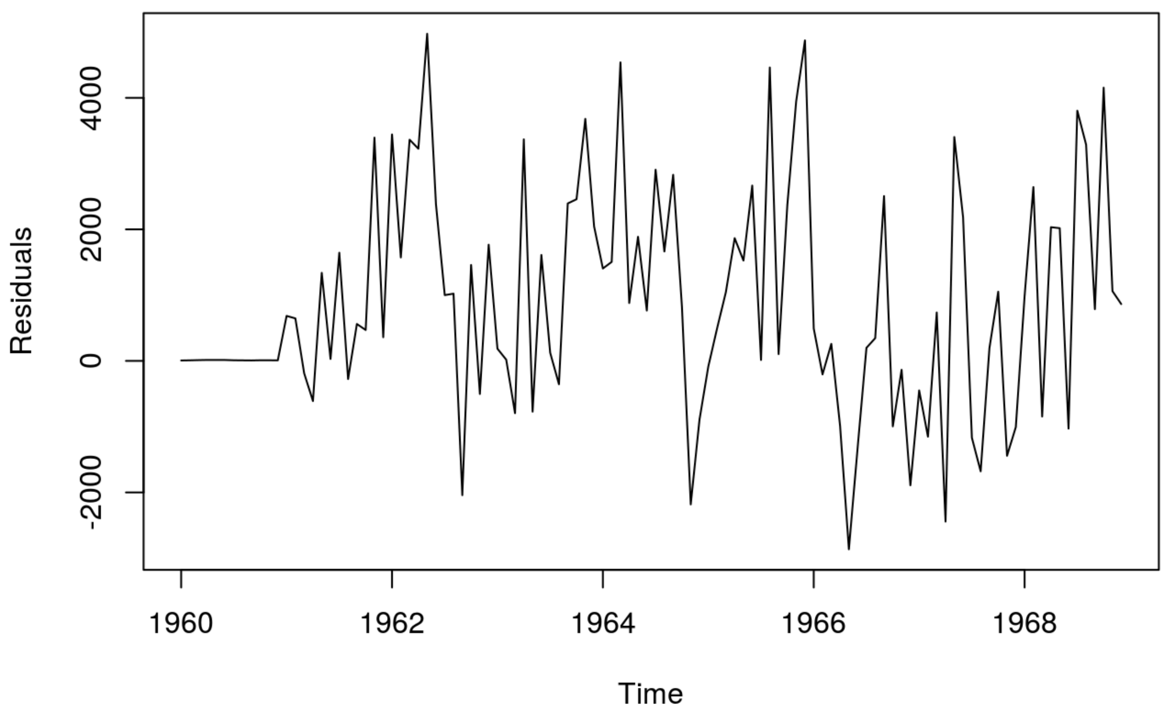
with p = non-seasonal AR order, d = non-seasonal differencing, q = non-seasonal MA order and P = seasonal AR order, D = seasonal differencing & Q = seasonal MA order and s is time span of repeating seasonal pattern.

Specification of seasonal part (D=1)

- Firstly, we do seasonal differencing to get rid of seasonal trend and fitting a plain model until time series and ACF/PACF plots of residuals show no sign of seasonality.
- Then we determine the orders of P & Q of the seasonal part based on final ACF/PACF plots of residuals.
- We do so by fitting the ARIMA(0, 0, 0) x (0, 1, 0) model and plotting the graphs.
- Although the general upward trend is resolved, we plot the ACF & PACF.

```
m1.cars = arima(cars.ts,order=c(0,0,0),seasonal=list(order=c(0,1,0), period=1
2))
res.m1 = residuals(m1.cars);
par(mfrow=c(1,1))
plot(res.m1,xlab='Time',ylab='Residuals')
```

Time series plot of 1st seasonal difference



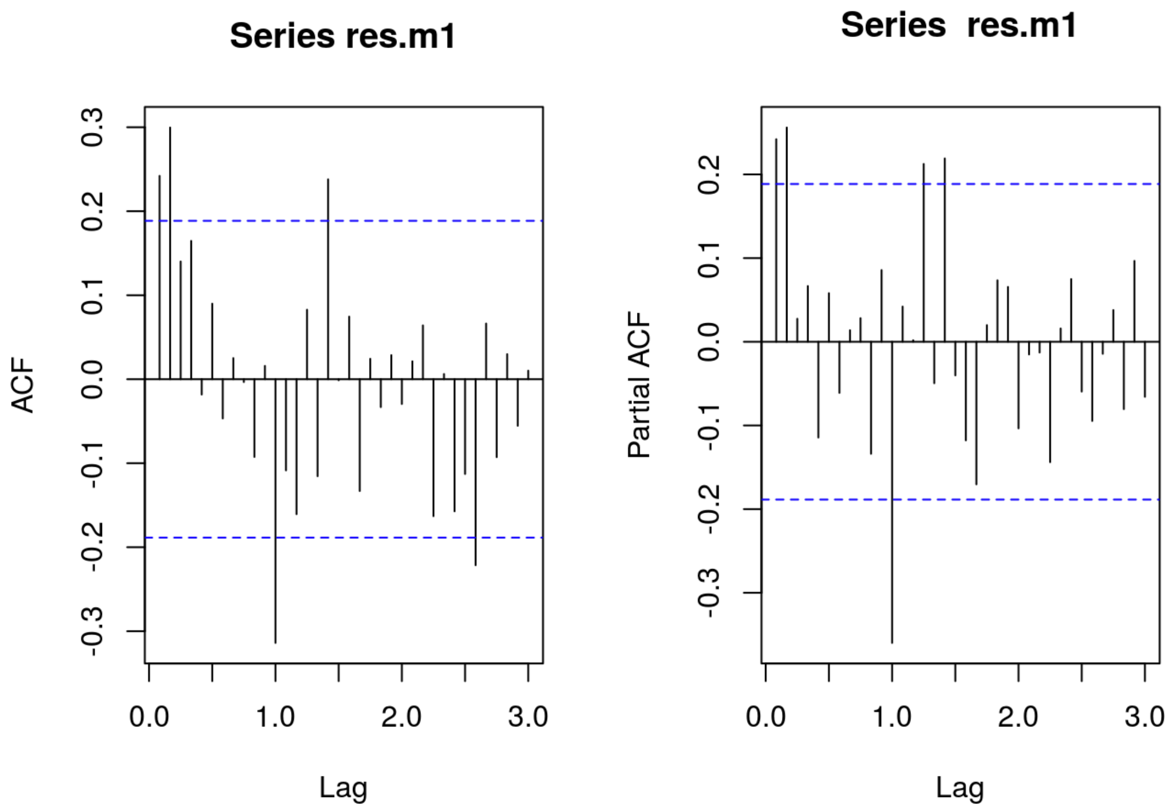
ACF & PACF plot

The ACF/PACF plots show seasonal trend is filtered out.

ACF & PACF shows 1 seasonal lag , indicates SARMA(1,1)

```
par(mfrow=c(1,2))

acf(res.m1, lag.max = 36)

pacf(res.m1, lag.max = 36)
```

ACF/PACF plot of the residuals

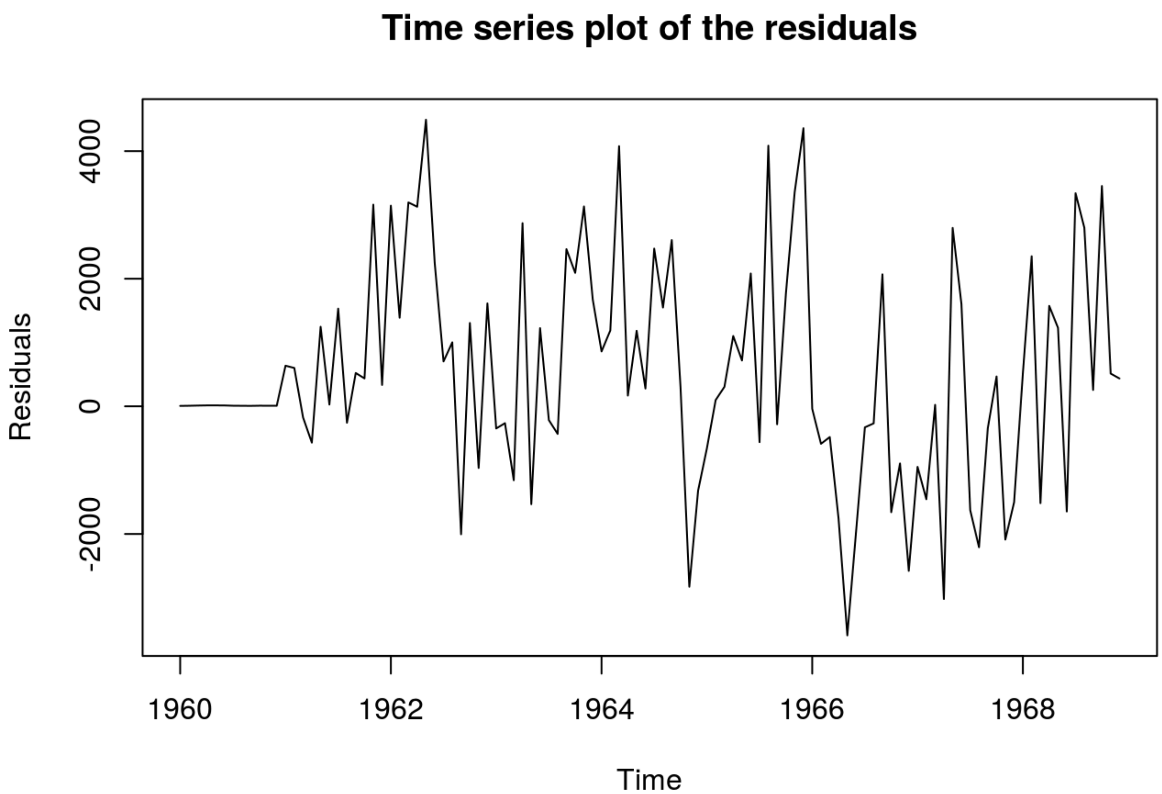


Specification of seasonal part (D=1, P=1, Q=1)

The general upward trend is no longer seen. We plot the residuals ACF & PACF plots.

```
m2.cars = arima(cars.ts,order=c(0,0,0),seasonal=list(order=c(1,1,1), period=1
2))

res.m2 = residuals(m2.cars);

par(mfrow=c(1,1))

plot(res.m2,xlab='Time',ylab='Residuals',main="Time series plot of the residu
als")
```

Time series plot with seasonal AR & MA coefficient (P=1,Q=1)

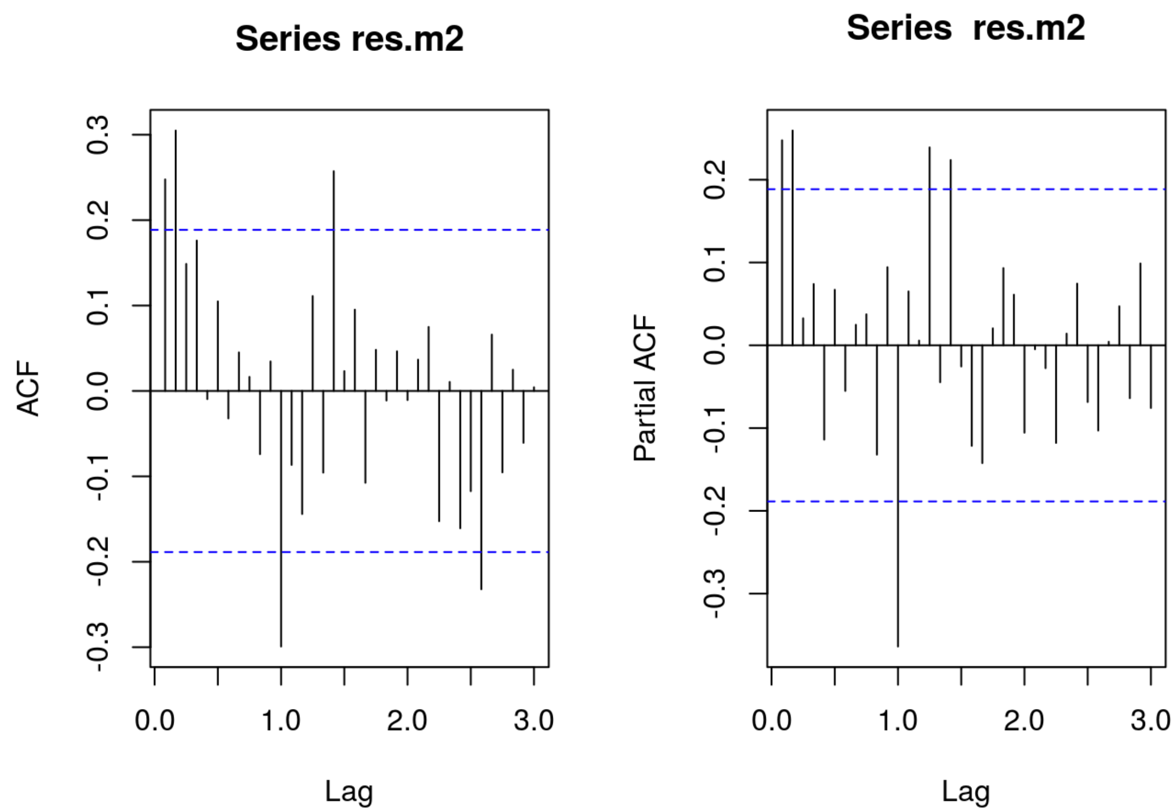


ACF & PACF plots

Auto correlation is still present on the seasonal lags. Therefore we need to repeat the process with higher number of Q until filtering out seasonality.

```
par(mfrow=c(1,2))
acf(res.m2, lag.max = 36)
pacf(res.m2, lag.max = 36)
```

ACF/PACF plot of the residuals with AR & MA coefficient (P=1,D=1,Q=1)



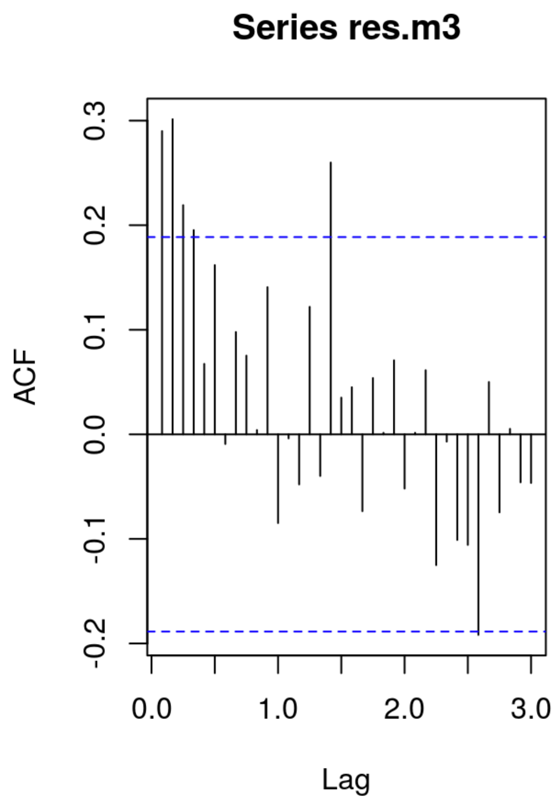
Specification of seasonal part (D=1, P=1, Q=2)

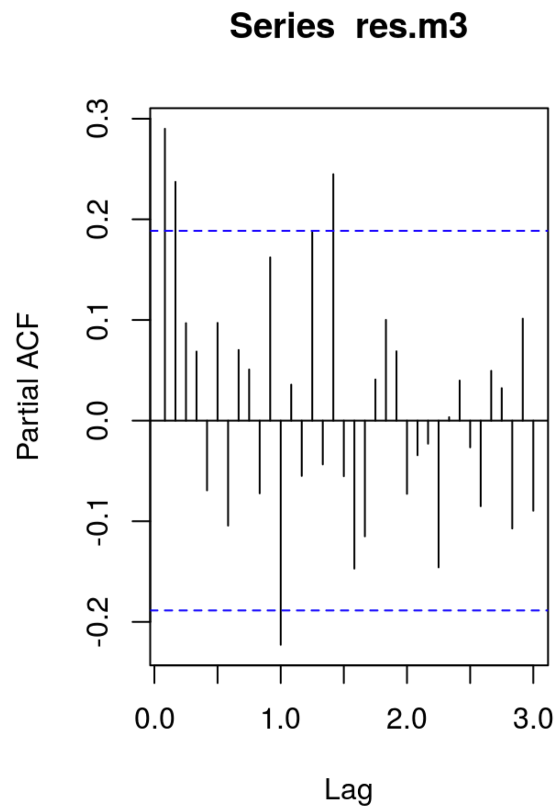
Now no seasonal lags are present. Hence seasonality specification completes at P=1, D=1, Q=2. We also see that there are multiple lags in ordinary pane (i.e. lags before 1st seasonal lag).

As we do not see any trend, we first apply transformation and check for the significant lags.

```
m3.cars = arima(cars.ts,order=c(0,0,0),seasonal=list(order=c(1,1,2), period=1
2))

res.m3 = residuals(m3.cars)

par(mfrow=c(1,2))

acf(res.m3, lag.max = 36)

pacf(res.m3, lag.max = 36)
```

ACF/PACF plot of the residuals with P=1,D=1,Q=2





Transformation

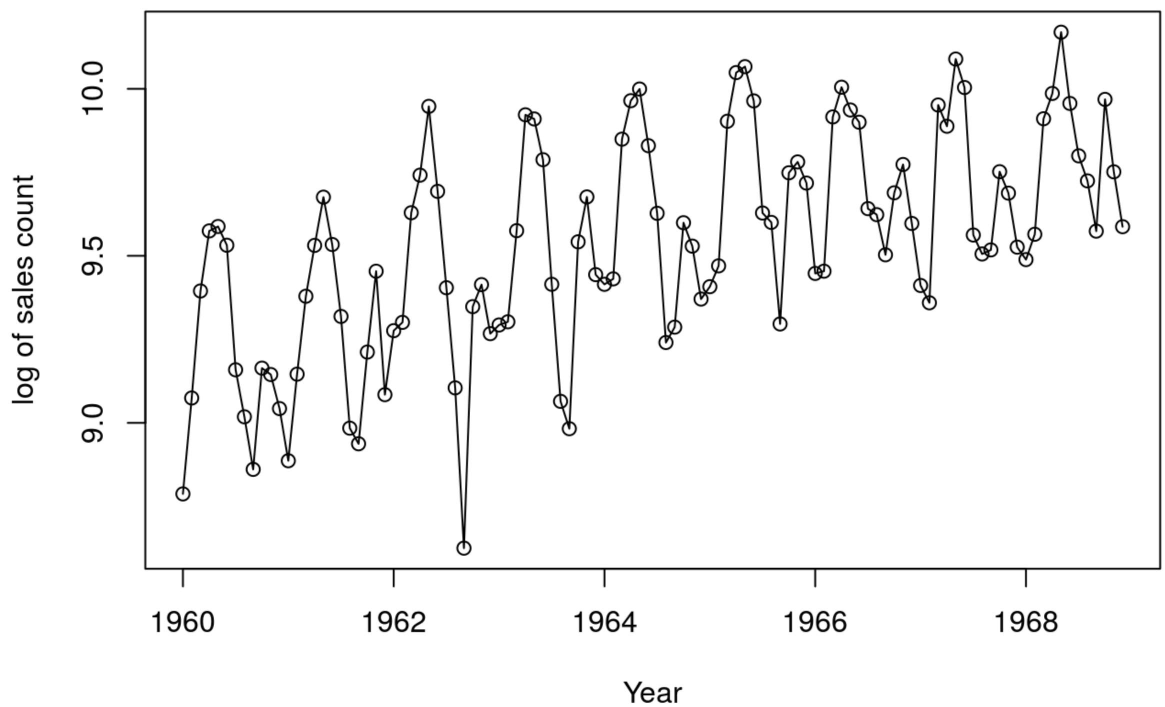
Data transformation are important tools for proper statistical analysis.

We have applied log transformation on the time series data and plotted it on the graph.

We see an intervention point in the TS plot. Apart from it the time series look as same with upward trend.

```
log.cars.ts = log(cars.ts)
par(mfrow=c(1,1))
plot(log.cars.ts,ylab='log of sales count',xlab='Year',type='o')
```

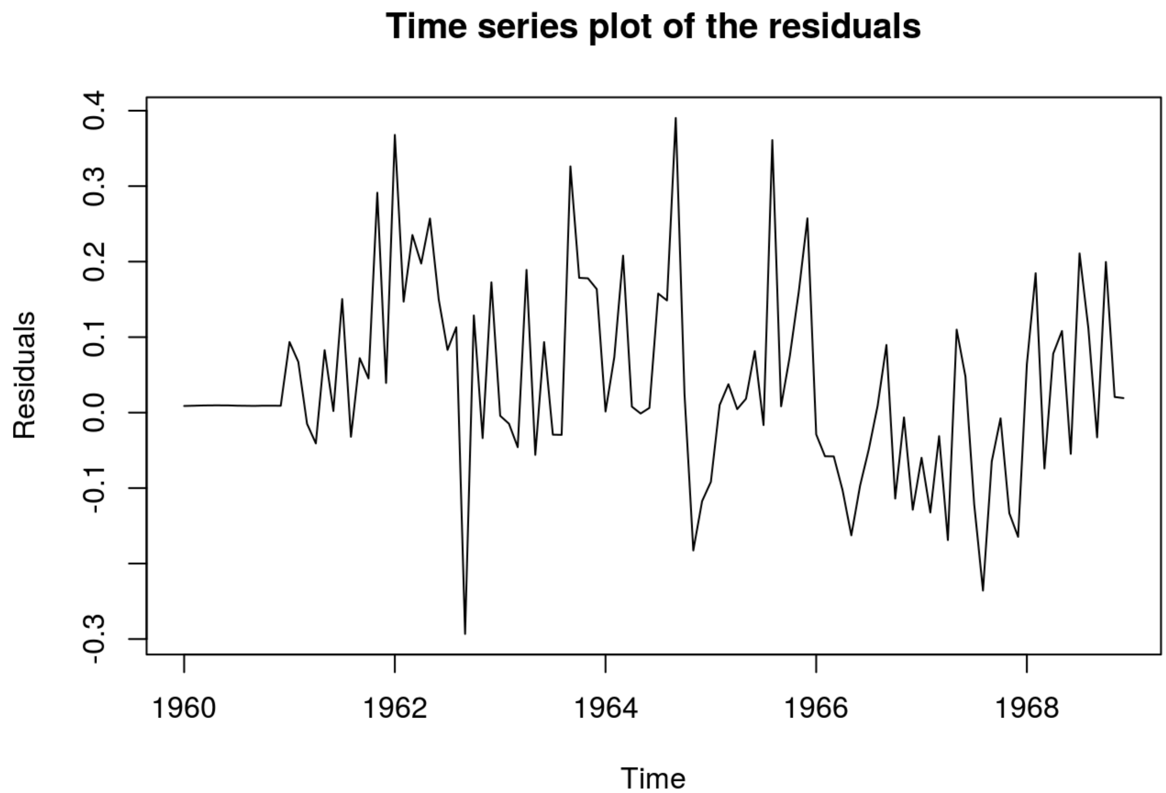
Time series plot with transformed data



We fit the model with log transformed data and draw conclusions from sample ACF & PACF's plot.

```
m4.cars = arima(log.cars.ts,order=c(0,0,0),seasonal=list(order=c(1,1,2), peri
od=12))

res.m4 = residuals(m4.cars)

plot(res.m4,xlab='Time',ylab='Residuals',main="Time series plot of the residu
als")
```

Time series plot of the residuals after transformation

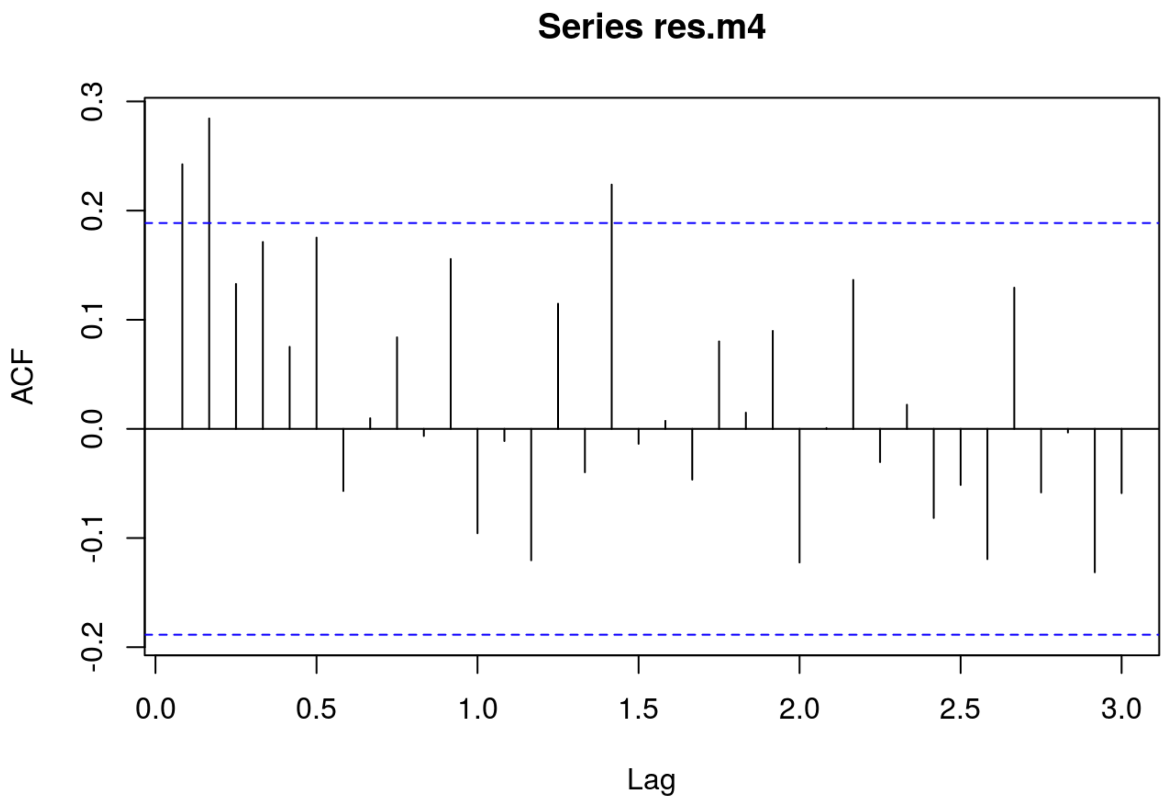


The ACF plot shows two significant lags before 1st seasonal lag.

PACF shows 2 significant lags.

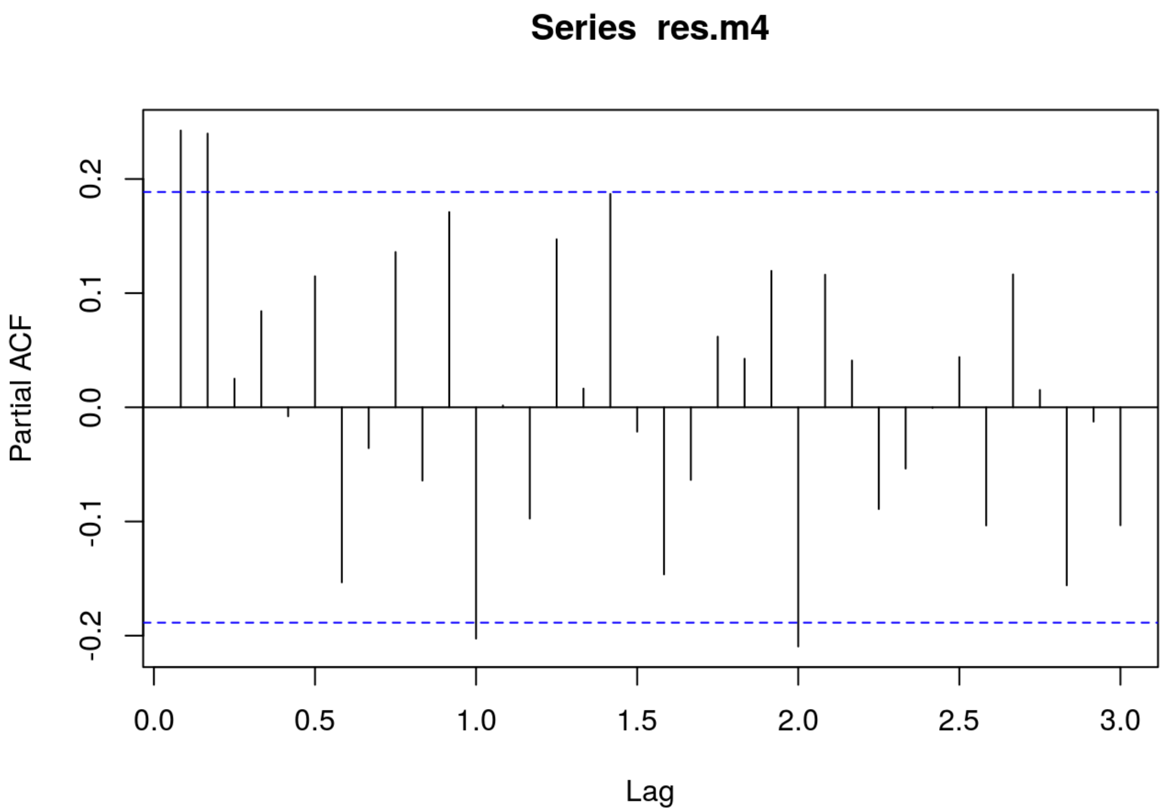
```
acf(res.m4, lag.max = 36)
```

ACF plot of the residuals after transformation




```
pacf(res.m4, lag.max = 36)
```

PACF plot of the residuals after transformation



Non-seasonal Differencing

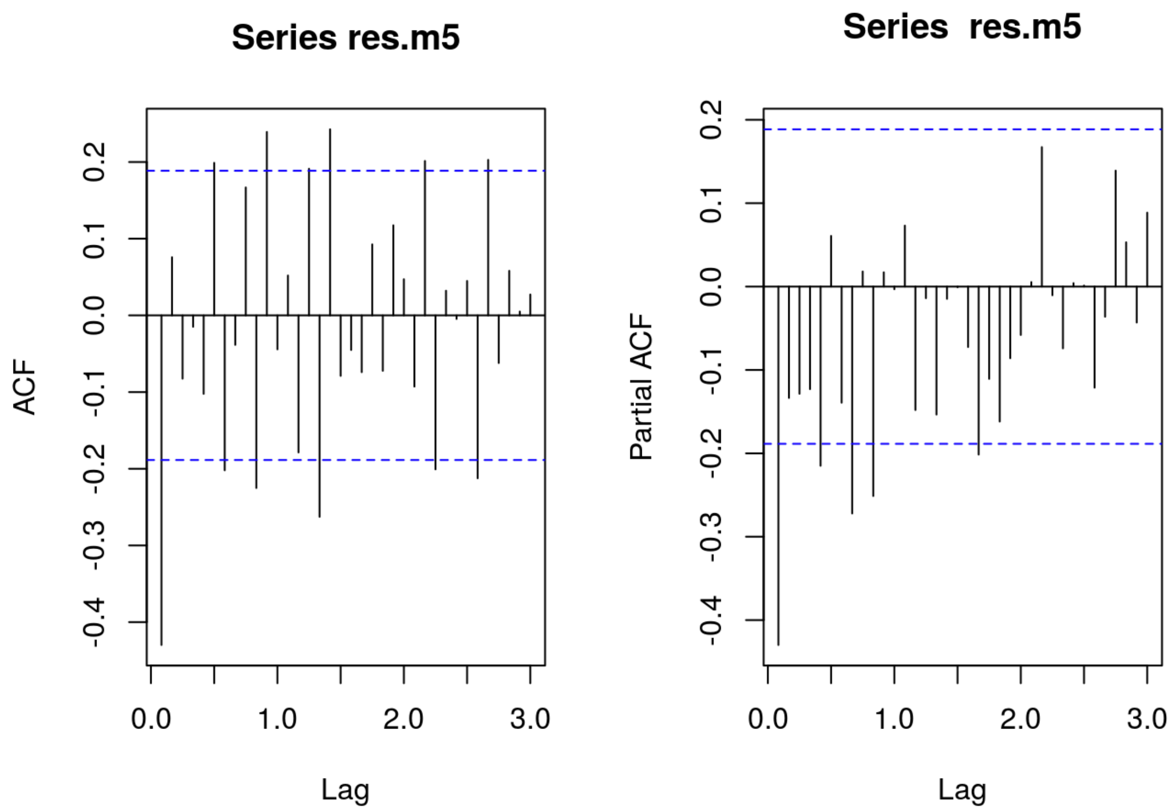
Here main aim is to come up with set of possible models. So let's start with d=1.

As we don't see any decaying pattern, we start with ordinary differencing d=1, to get rid of remaining trend and correlation in ACF/PACF plots.

Differencing gives us something. It helps us come up with a model.

```
# SARIMA(0,1,0)x(1,1,2)
m5.cars = arima(log.cars.ts,order=c(0,1,0),seasonal=list(order=c(1,1,2), peri
od=12))
res.m5 = residuals(m5.cars)
par(mfrow=c(1,2))
acf(res.m5, lag.max = 36)
pacf(res.m5, lag.max = 36)
```

ACF/PACF plot of the residuals with ordinary differencing



We get a high correlation at 1st lag and also we observe few significant lags. All this is due to the intervention point.

We come up with **MA(3 or 4)**and **AR(3)** from ACF & PACF plots.

Also there is no evidence for an ordinary trend. we can still apply adf test on residuals to make sure.

```
adf.test(res.m5)
```

```
## Warning in adf.test(res.m5): p-value smaller than printed p-value
##  Augmented Dickey-Fuller Test
## data:  res.m5
## Dickey-Fuller = -6.9914, Lag order = 4, p-value = 0.01
## alternative hypothesis: stationary
```

EACF

EACF is basically another tool just like ACF/PACF, for identifying order of ARIMA model. In the table of EACF's with MA coefficient listed across top and AR coefficient listed down the side, the top-left most EACF that is less than absolute value of 2 times standard error of EACF is at position that indicates good choice for orders of model.

Now we use EACF on the residuals of the previous step (i.e. res.m5), to see information about AR and MA parts lefts in residuals. Our candidates for ARMA part are ARMA(1,2), ARMA(2,2) & ARMA (2,1)

```
eacf(res.m5)
```

```
## AR/MA
##   0 1 2 3 4 5 6 7 8 9 10 11 12 13
## 0 x o o o o x x o o x x  o  o  o
## 1 x x o o o o x o o o x  o  o  o
## 2 x o o o o o o o o o x  o  o  o
## 3 x o o o o o o o o o o  o  o  o
## 4 x o o o o o o o o o x  o  o  o
## 5 x x x x o o o o o o o  o  o  o
## 6 x o x o x o o o o o o  o  o  o
## 7 x x x o x o x o o o o  o  o  o
```

The tentative models are specified as:

- SARIMA (0,1,4)x(1,1,2) by ACF if we consider pacf is tailing off so we take only MA coefficient.
- SARIMA (0,1,3)x(1,1,2) by ACF if we consider pacf is tailing off so we take only MA coefficient.
- SARIMA (3,1,4)x(1,1,2) by ACF/PACF
- SARIMA (2,1,1)x(1,1,2) by EACF
- SARIMA (2,1,2)x(1,1,2) by EACF
- SARIMA (3,1,2)x(1,1,2) for over fitting SARIMA (2,1,2)x(1,1,2)

Model Fitting

Parameter Estimations

From the given set of models we start fitting one by one and see it any of those are .

- We fit these models on the original series to obtain residuals.
- Then we do residual analysis if we have white noise residuals.

1.SARIMA (0,1,3)x(1,1,2)

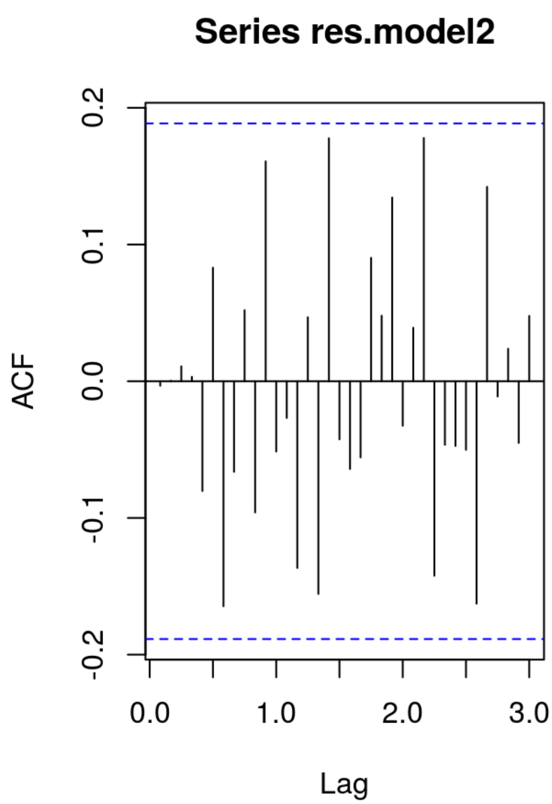
```r
model2.cars = arima(log.cars.ts,order=c(0,1,3),seasonal=list(order=c(1,1,2),
period=12),method = "ML")

coeftest(model2.cars)
```

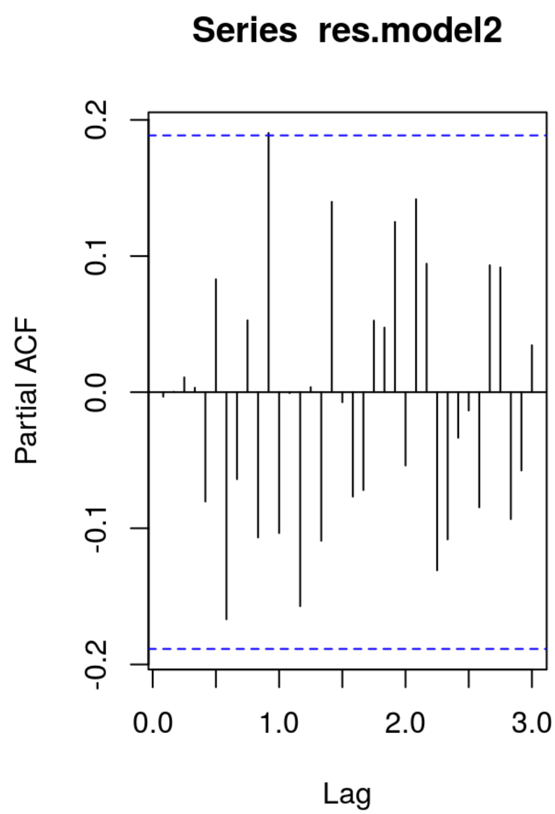
```
## z test of coefficients:
##        Estimate Std. Error z value  Pr(>|z|)
## ma1  -0.750678   0.103762 -7.2346 4.668e-13 ***
## ma2   0.074003   0.137082  0.5398  0.589307
## ma3  -0.217749   0.108075 -2.0148  0.043927 *
## sar1  0.480677   0.334715  1.4361  0.150979
## sma1 -1.084238   0.393826 -2.7531  0.005904 **
## sma2  0.084263   0.325977  0.2585  0.796027
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ACF/PACF shows presence of white noise and this indeed is one of the good model.we further fit other models and compare at the end.

```r
res.model2 = residuals(model2.cars)
par(mfrow=c(1,2))
acf(res.model2, lag.max = 36)
pacf(res.model2, lag.max = 36)
```

ACF/PACF plot of the residuals for ARIMA(0,1,3)x(1,1,2)





SARIMA(0,1,4)x(1,1,2)

```
model1.cars = arima(log.cars.ts,order=c(0,1,4),seasonal=list(order=c(1,1,2),
period=12),method = "ML")

coeftest(model1.cars)
```

```
## z test of coefficients:
##        Estimate Std. Error z value  Pr(>|z|)
## ma1  -0.716539   0.096911 -7.3938 1.427e-13 ***
## ma2   0.070897   0.122981  0.5765   0.56429
## ma3  -0.205801   0.119286 -1.7253   0.08448 .
## ma4  -0.041378   0.083517 -0.4954   0.62028
## sar1 -0.197613         NA      NA        NA
## sma1 -0.290238         NA      NA        NA
## sma2 -0.336851         NA      NA        NA
## ---
```

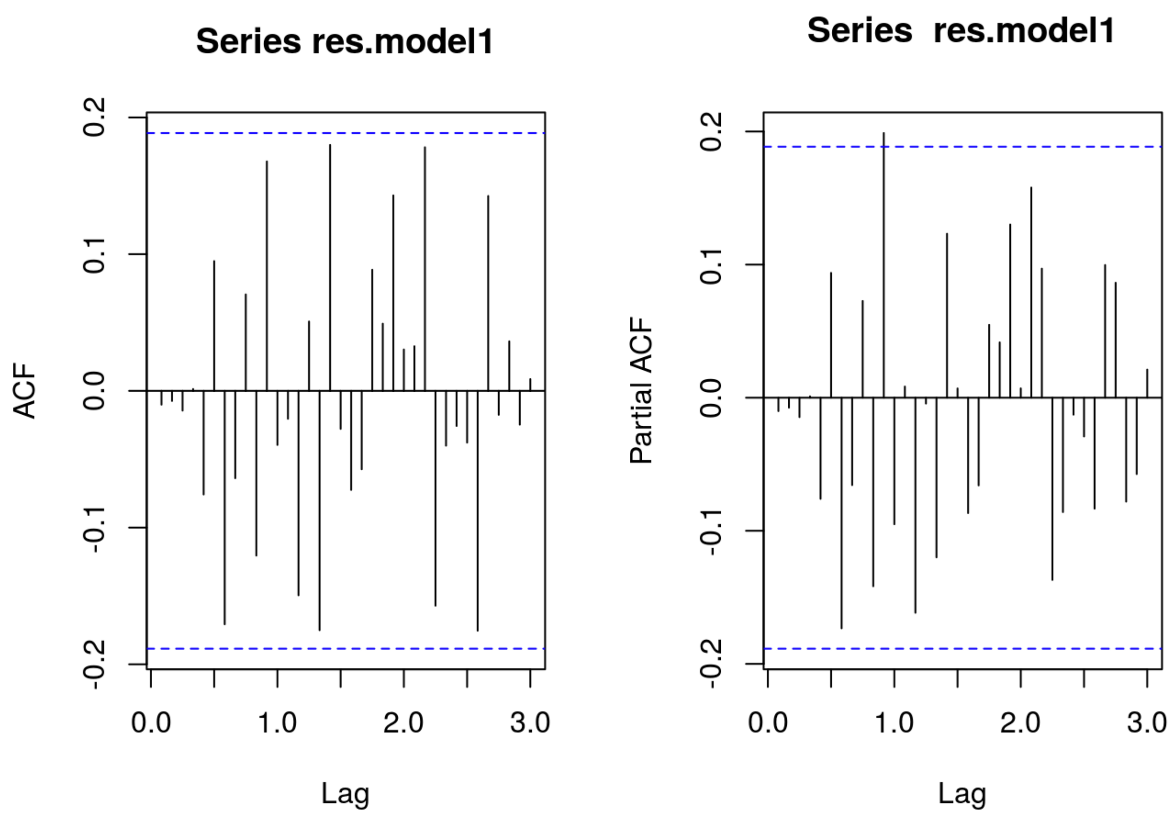
```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The coefficient test shows that MA(2),MA(4) are insignificant while MA(3) is slightly significant. Hence we cannot use this model for further analysis.

Also in PACF, we see some slightly significant lag.

```
res.model1 = residuals(model1.cars);

par(mfrow=c(1,2))

acf(res.model1, lag.max = 36)

pacf(res.model1, lag.max = 36)
```

ACF/PACF plot of the residuals for ARIMA(0,1,4)x(1,1,2)



SARIMA(3,1,4)x(1,1,2)

We go with the bigger model from ACF/PACF

```
model3.cars = arima(log.cars.ts,order=c(3,1,4),seasonal=list(order=c(1,1,2),
period=12),method = "ML")
```

```
coeftest(model3.cars)
```

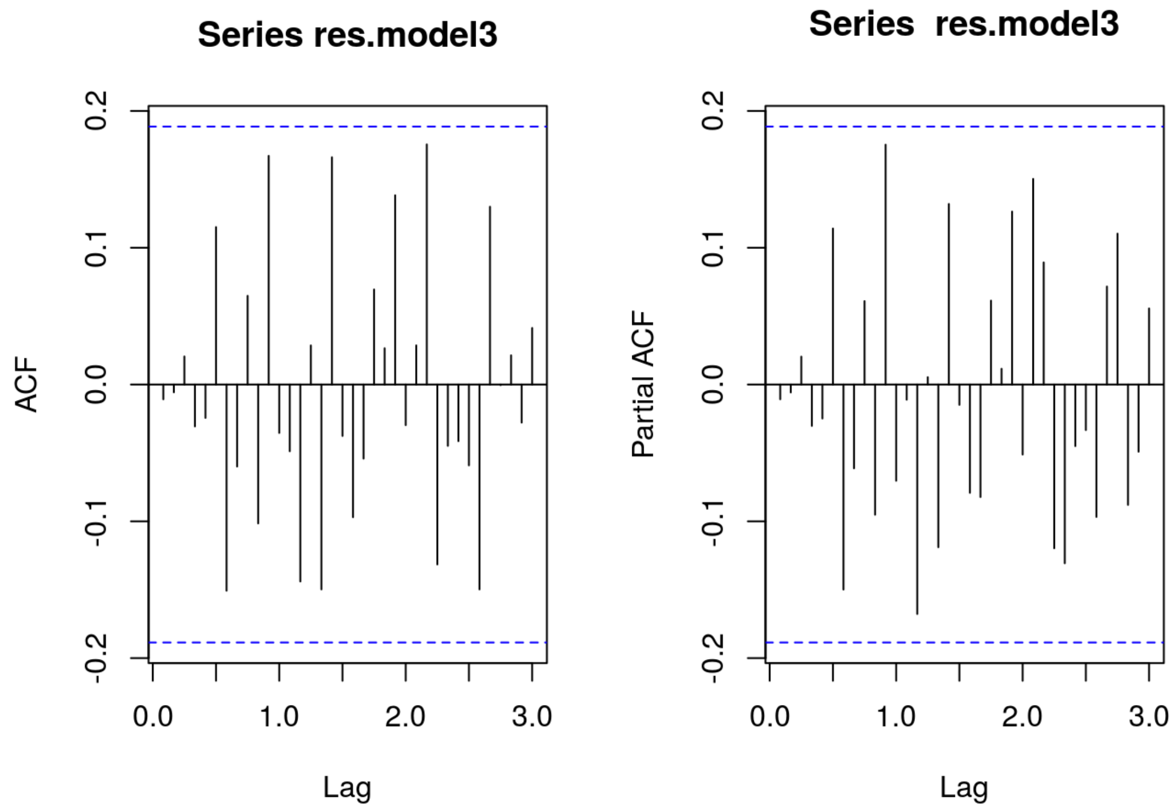
```
## z test of coefficients:
##       Estimate Std. Error z value Pr(>|z|)
## ar1   0.726370   0.912899  0.7957  0.42622
## ar2   0.538303   0.349953  1.5382  0.12400
## ar3  -0.472199   0.411453 -1.1476  0.25112
## ma1  -1.458156   0.908939 -1.6042  0.10866
## ma2   0.047285   0.723648  0.0653  0.94790
## ma3   0.608440   0.565129  1.0766  0.28164
## ma4  -0.170793   0.238443 -0.7163  0.47382
## sar1  0.438596   0.384708  1.1401  0.25426
## sma1 -1.047716   0.546545 -1.9170  0.05524 .
## sma2  0.072712   0.371171  0.1959  0.84469
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The coefficient test do not show any significant values but we see the presence of white noise from the residuals.

Hence we later check if this model is suitable or not.

```
res.model3 = residuals(model3.cars)
par(mfrow=c(1,2))
acf(res.model3, lag.max = 36)
pacf(res.model3, lag.max = 36)
```

ACF/PACF plot of the residuals for ARIMA(3,1,4)x(1,1,2)



SARIMA(2,1,1)x(1,1,2)

```
model4.cars = arima(log.cars.ts,order=c(2,1,1),seasonal=list(order=c(1,1,2),
period=12),method = "ML")

coeftest(model4.cars)
```

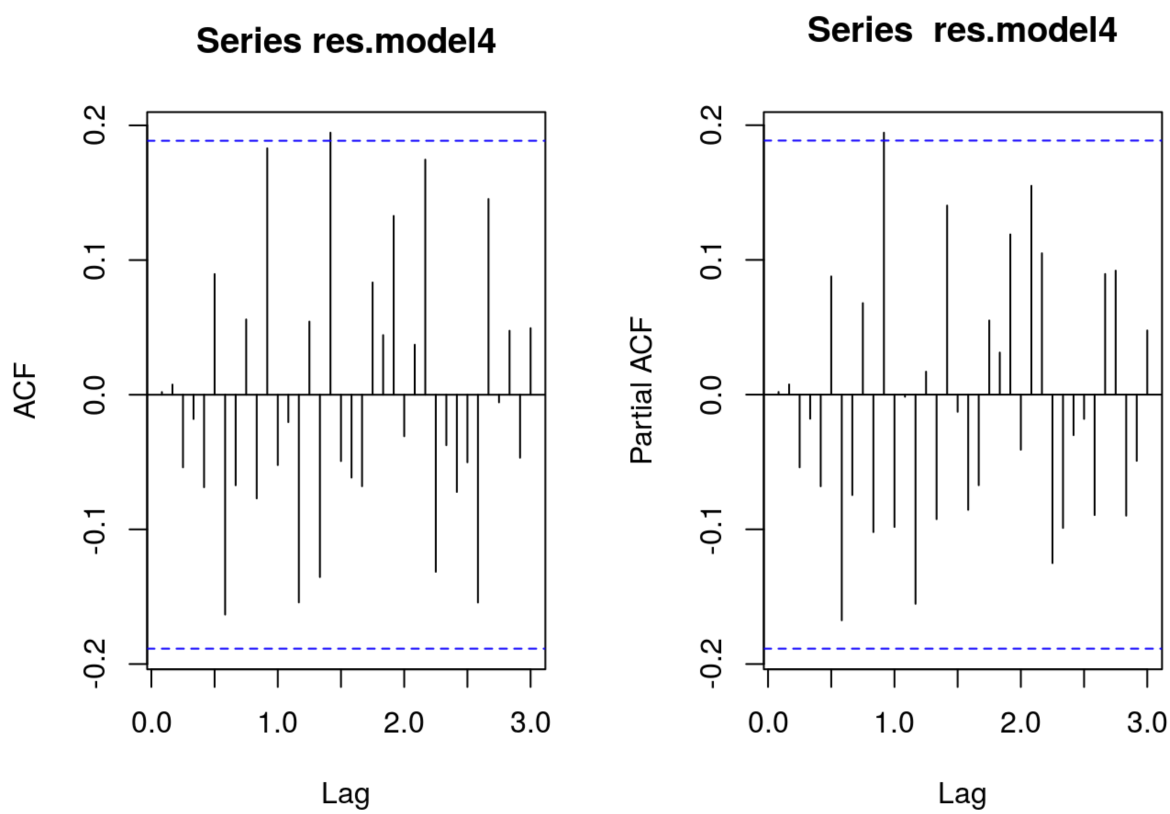
```
## z test of coefficients:
##        Estimate Std. Error  z value  Pr(>|z|)
## ar1    0.182219   0.111112   1.6400  0.101014
## ar2    0.207476   0.116201   1.7855  0.074180 .
## ma1   -0.943115   0.051848 -18.1901 < 2.2e-16 ***
## sar1   0.511863   0.312659   1.6371  0.101604
## sma1  -1.102519   0.380088  -2.9007  0.003723 **
## sma2   0.102524   0.307102   0.3338  0.738497
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

MA(1) coefficient shows significant values but AR component doesn't.

Father ACF/PACF plots shows slightly significant lags.

```
res.model4 = residuals(model4.cars)

par(mfrow=c(1,2))

acf(res.model4, lag.max = 36)

pacf(res.model4, lag.max = 36)
```

ACF/PACF plot of the residuals for ARIMA(2,1,1)x(1,1,2)



SARIMA(2,1,2)x(1,1,2)

```
model5.cars = arima(log.cars.ts,order=c(2,1,2),seasonal=list(order=c(1,1,2),
period=12),method = "ML")

coeftest(model5.cars)
```

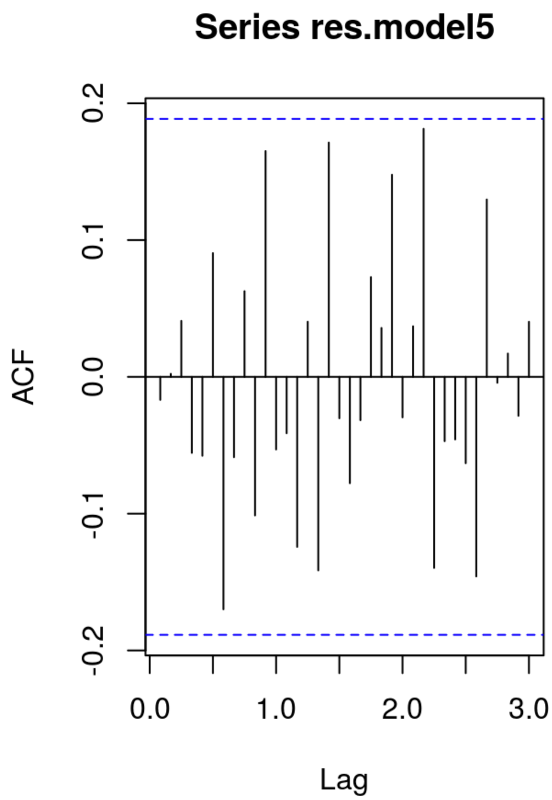
```
## z test of coefficients:
##       Estimate Std. Error z value Pr(>|z|)
## ar1  -0.309047   0.378002 -0.8176 0.413597
## ar2   0.299031   0.119764  2.4968 0.012531 *
## ma1  -0.423538   0.392802 -1.0782 0.280924
## ma2  -0.476045   0.347773 -1.3688 0.171049
## sar1  0.457113   0.333679  1.3699 0.170712
## sma1 -1.045249   0.393968 -2.6531 0.007975 **
## sma2  0.045396   0.329495  0.1378 0.890417
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

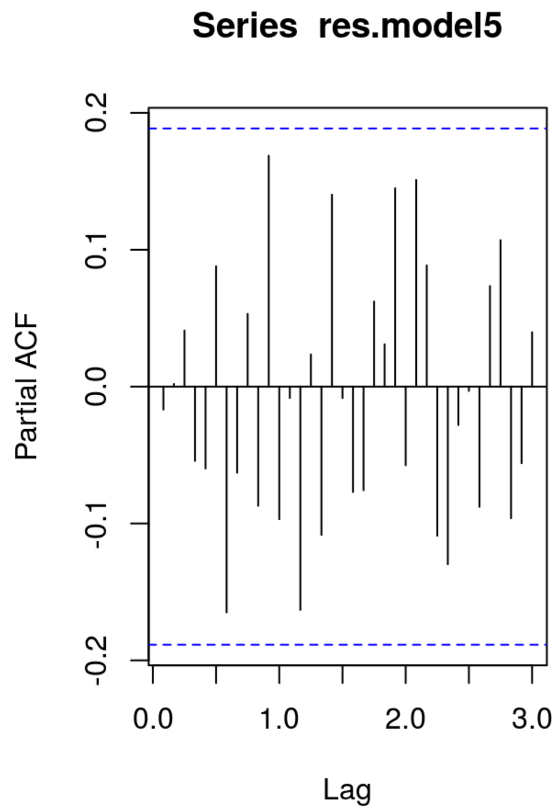
Only AR(2) is significant.

ACF/PACF shows presence of white noise.

```
res.model5 = residuals(model5.cars)
par(mfrow=c(1,2))
acf(res.model5, lag.max = 36)
pacf(res.model5, lag.max = 36)
```

ACF/PACF plot of the residuals for ARIMA(2,1,2)x(1,1,2)





AIC & BIC

Above we have fitted all the set of possible models and computed the coefficient test.To select one best model we can go with AIC/BIC function.

```
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")')
  }
}
```

```
sc.AIC=AIC(model1.cars,model2.cars,model3.cars,model4.cars,model5.cars)
```

```
sc.BIC=AIC(model1.cars,model2.cars,model3.cars,model4.cars,model5.cars,k=log(
length(cars.ts)))
```

AIC gives us model2 i.e. SARIMA(0,1,3)x(1,1,2) as best model.

```
sort.score(sc.AIC, score = "aic")
```

	df <dbl>	AIC <dbl>
model2.cars	7	-129.3019
model4.cars	7	-129.1018
model5.cars	8	-127.7664
model1.cars	8	-126.5273
model3.cars	11	-122.4308

5 rows

BIC also gives us the same model2.cars i.e. SARIMA(0,1,3)x(1,1,2) as best one.

```
sort.score(sc.BIC, score = "aic")
```

	df <dbl>	AIC <dbl>
model2.cars	7	-110.52694
model4.cars	7	-110.32686
model5.cars	8	-106.30931
model1.cars	8	-105.07023
model3.cars	11	-92.92739

5 rows

And as we always go for one with white noise residuals meaning model whose residuals have presence of white noise.

We further check on specified model by residual analysis.

Model Diagnostics

Residual Analysis

Residuals in time series are what is left over after fitting a model.For most of the time series models, residuals are equal to difference between the observation and corresponding fitted values.

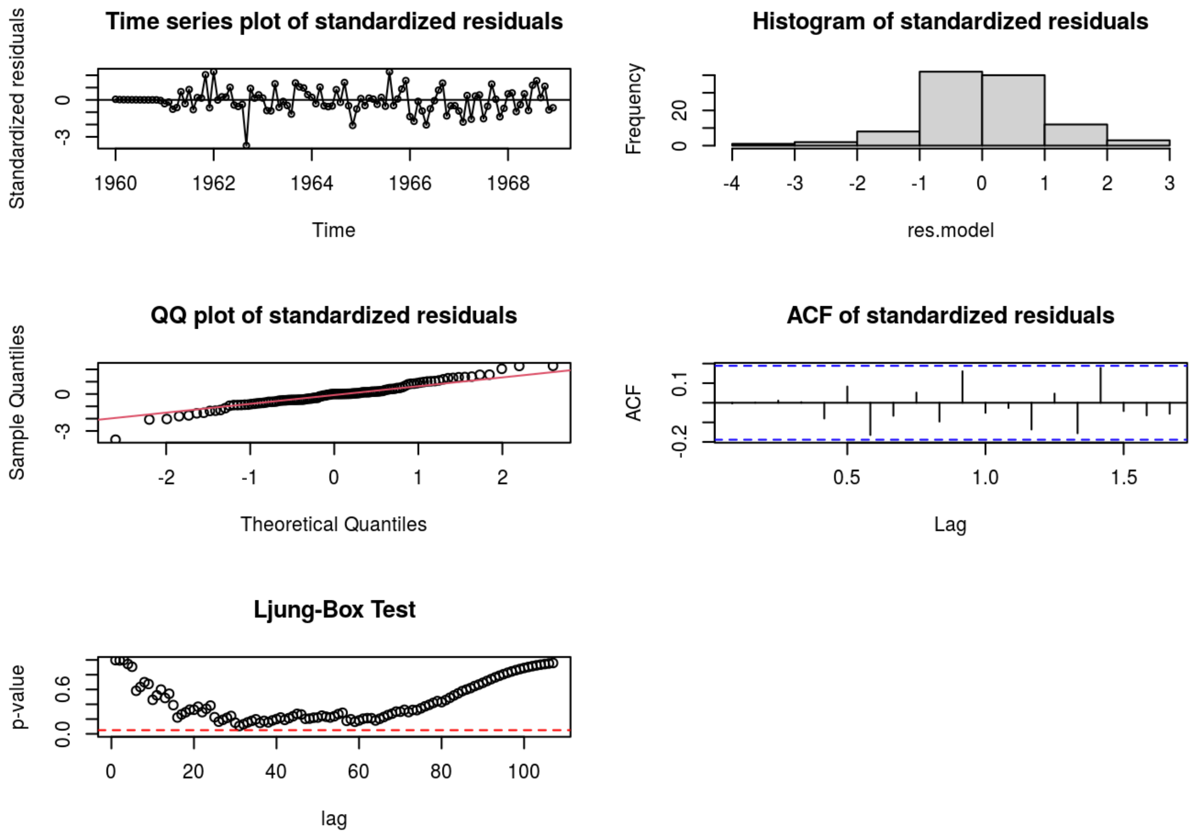
SARIMA(0,1,3)x(1,1,2) will be used for parameter estimation, testing the significance of parameters and forecasting.

Also we need to check if the residuals are normal (on top of being presence of white noise)

```
residual.analysis <- function(model, std = TRUE){
  library(TSA)
  library(FitAR)
  if (std == TRUE){
    res.model = rstandard(model)
  }else{
    res.model = residuals(model)
  }
  par(mfrow=c(3,2))
  plot(res.model,type='o',ylab='Standardized residuals', main="Time series pl
ot of standardized residuals")
  abline(h=0)
  hist(res.model,main="Histogram of standardized residuals")
  qqnorm(res.model,main="QQ plot of standardized residuals")
  qqline(res.model, col = 2)
  acf(res.model,main="ACF of standardized residuals")
  print(shapiro.test(res.model))
  k=0
  LBQPlot(res.model, lag.max = length(model$residuals)-1 , StartLag = k + 1,
k = 0, SquaredQ = FALSE)
}
```

```
residual.analysis(model=model2.cars)
```

```
##  Shapiro-Wilk normality test
## data:  res.model
## W = 0.97161, p-value = 0.02062
```



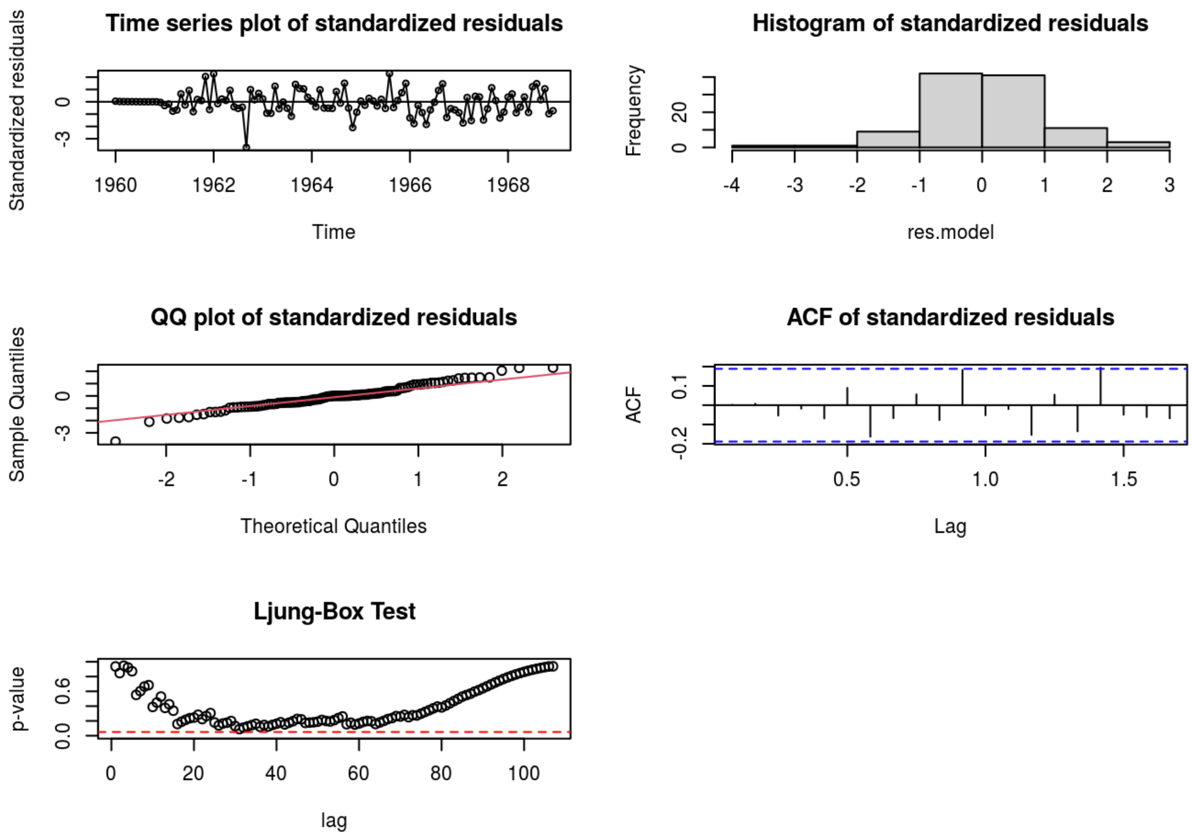
ARIMA(0,1,3)x(1,1,2)

From the residual analysis of SARIMA (0,1,3)x(1,1,2) model we draw the following conclusions:

1. Histogram shows normal distribution of the residuals.
2. ACF plot shows presence of white noise . 3.Ljung Box test shows good residuals.
3. QQplot shows normality with 1 outlier at the head.
4. TS plots do not show any trend(as seen earlier) except an intervention point.

```
residual.analysis(model=model4.cars)
```

```
##   Shapiro-Wilk normality test
## data:  res.model
## W = 0.97345, p-value = 0.02925
```



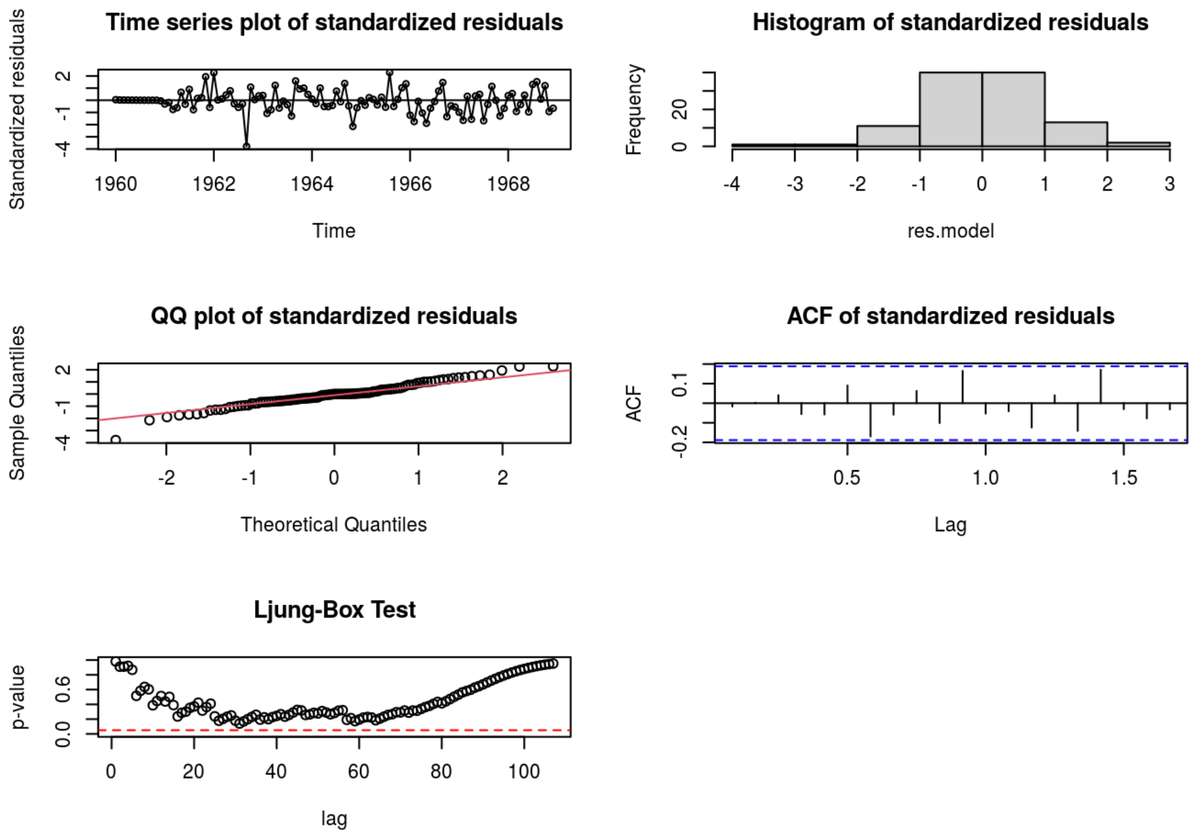
ARIMA(2,1,1)x(1,1,2)

From the residual analysis of SARIMA (2,1,1)x(1,1,2) model we draw the following conclusions:

1. Histogram shows normal distribution of the residuals.
2. ACF plot shows slightly significant lag or near to white noise residuals . 3.Ljung Box test shows few points on and below the threshold line.
3. QQplot shows normality with some points deviating away from the line.
4. TS plots do not show any trend(as seen earlier) except an intervention point.

```
residual.analysis(model=model5.cars)
```

```
##  Shapiro-Wilk normality test
## data:  res.model
## W = 0.97214, p-value = 0.02278
```



ARIMA(2,1,2)x(1,1,2)

From the residual analysis of SARIMA (2,1,2)x(1,1,2) model we draw the following conclusions:

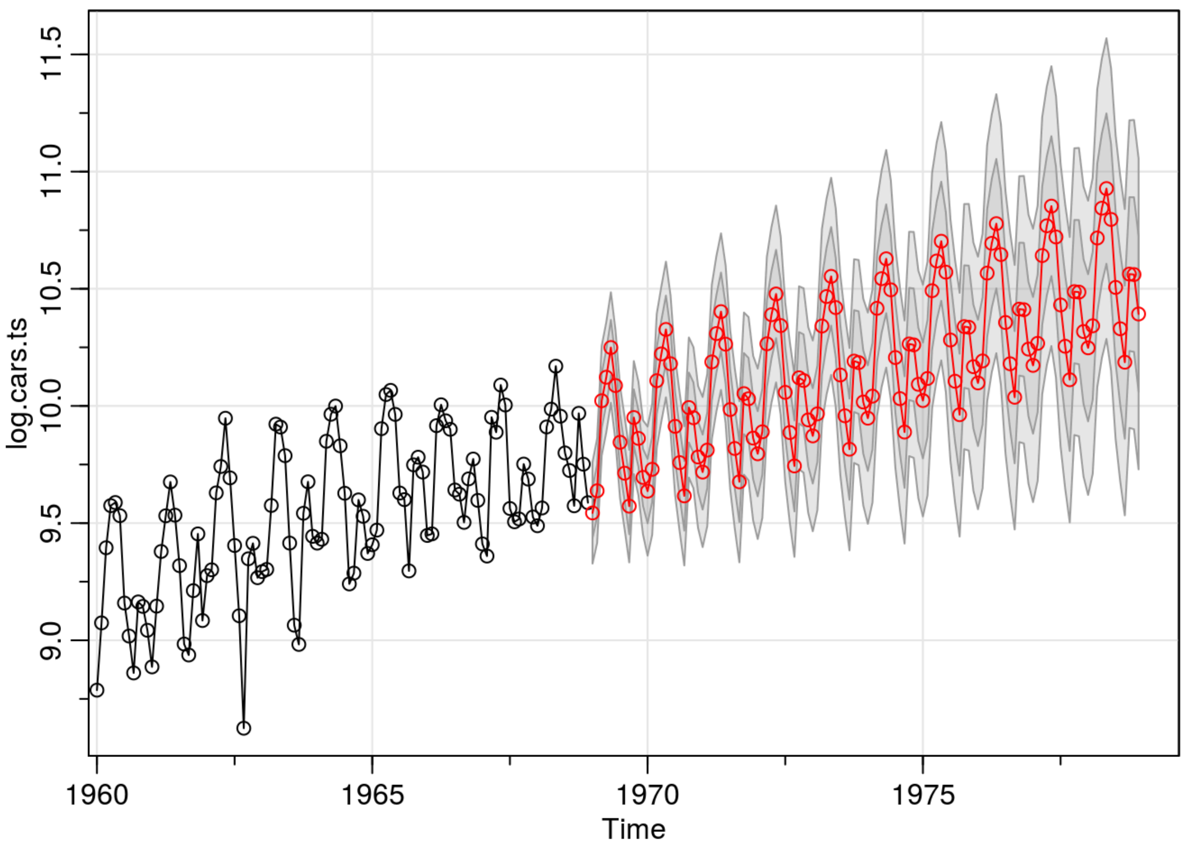
1. Histogram shows normal distribution of the residuals.
2. ACF plot shows presence of white noise residuals .
3. Ljung Box test shows all points above the threshold line.
4. QQplot shows normality.
5. TS plots do not show any trend(as seen earlier) except an intervention point.

Forecasting

After getting one best model {ARIMA (0,1,3)x(1,1,2)} we will consider it for prediction of future realizations of time series. Below figure shows forecasts for a lead time of 10 years for the model that we fit.

the order of parameters in the command is name of data series, the number of times for which we want forecast (10 years in our case) followed by the parameters of ARIMA model.

```
sarima.for(log.cars.ts,120,0,1,3,1,1,2,12)
```




```
## $pred
##            Jan       Feb       Mar       Apr       May       Jun       Jul
## 1969  9.543435  9.637878 10.022233 10.122532 10.249291 10.087102  9.844820
## 1970  9.636433  9.729600 10.107755 10.221978 10.326415 10.180118  9.912970
## 1971  9.717803  9.811430 10.187752 10.308669 10.402376 10.263716  9.984616
## 1972  9.795804  9.889652 10.265094 10.389227 10.477777 10.342789 10.057944
## 1973  9.872186  9.966140 10.341158 10.466838 10.552909 10.419686 10.132079
## 1974  9.947790 10.041794 10.416610 10.543033 10.627912 10.495537 10.206602
## 1975 10.023019 10.117048 10.491766 10.618546 10.702852 10.570885 10.281313
## 1976 10.098069 10.192110 10.566780 10.693732 10.777763 10.645992 10.356113
## 1977 10.173032 10.267078 10.641726 10.768761 10.852659 10.720983 10.430956
## 1978 10.247953 10.342003 10.716639 10.843714 10.927549 10.795917 10.505820
##            Aug       Sep       Oct       Nov       Dec
## 1969  9.713065  9.572391  9.950444  9.861546  9.695025
## 1970  9.758004  9.616215  9.993182  9.949526  9.782053
## 1971  9.818494  9.676168 10.052614 10.030704  9.862774
```

```
## 1972  9.886458  9.743875 10.120070 10.108613  9.940464
## 1973  9.958015  9.815308 10.191383 10.184951 10.016696
## 1974 10.031300  9.888533 10.264550 10.260533 10.092227
## 1975 10.105414  9.962619 10.338608 10.335752 10.167422
## 1976 10.179928 10.037119 10.413094 10.410797 10.242455
## 1977 10.254634 10.111818 10.487787 10.485758 10.317410
## 1978 10.329431 10.186613 10.562578 10.560678 10.392328
##
## $se
##            Jan       Feb       Mar       Apr       May       Jun       Jul
## 1969 0.1080666 0.1113838 0.1167439 0.1172965 0.1178464 0.1183938 0.1189386
## 1970 0.1376082 0.1397741 0.1426344 0.1436233 0.1446056 0.1455811 0.1465502
## 1971 0.1600609 0.1620482 0.1644269 0.1656688 0.1669015 0.1681252 0.1693400
## 1972 0.1817234 0.1836356 0.1858122 0.1871863 0.1885504 0.1899047 0.1912494
## 1973 0.2034206 0.2052975 0.2073807 0.2088220 0.2102535 0.2116752 0.2130875
## 1974 0.2252750 0.2271334 0.2291702 0.2306461 0.2321125 0.2335697 0.2350179
## 1975 0.2472752 0.2491227 0.2511345 0.2526286 0.2541139 0.2555906 0.2570588
## 1976 0.2693913 0.2712317 0.2732284 0.2747326 0.2762286 0.2777166 0.2791967
## 1977 0.2915975 0.2934324 0.2954190 0.2969291 0.2984317 0.2999267 0.3014143
## 1978 0.3138739 0.3157044 0.3176833 0.3191972 0.3207040 0.3222037 0.3236965
##            Aug       Sep       Oct       Nov       Dec
## 1969 0.1194810 0.1200209 0.1205584 0.1211269 0.1216856
## 1970 0.1475130 0.1484695 0.1494198 0.1504400 0.1514344
## 1971 0.1705461 0.1717438 0.1729332 0.1742281 0.1754827
## 1972 0.1925847 0.1939108 0.1952280 0.1966802 0.1980808
## 1973 0.2144905 0.2158844 0.2172693 0.2188131 0.2202967
## 1974 0.2364573 0.2378879 0.2393100 0.2409100 0.2424433
## 1975 0.2585186 0.2599703 0.2614138 0.2630510 0.2646162
## 1976 0.2806689 0.2821334 0.2835904 0.2852538 0.2868412
## 1977 0.3028946 0.3043677 0.3058337 0.3075169 0.3091206
## 1978 0.3251824 0.3266615 0.3281340 0.3298330 0.3314496
```

Conclusion:
We have finally forecasted the cars sales for next 10 years using the model we found out to be the best out of set of possible models. The forecast shows upward trend which indeed sounds true as cars sales possibly goes high every year due to advancement in technologies, population growth and numerous reasons.

References:

1. Guidance: Prof. Hadi Safari
2. Dataset: https://www.kaggle.com/datasets/dinirimameev/monthly-car-sales-in-quebec-1960?select=monthly-car-sales-in-quebec-1960.csv