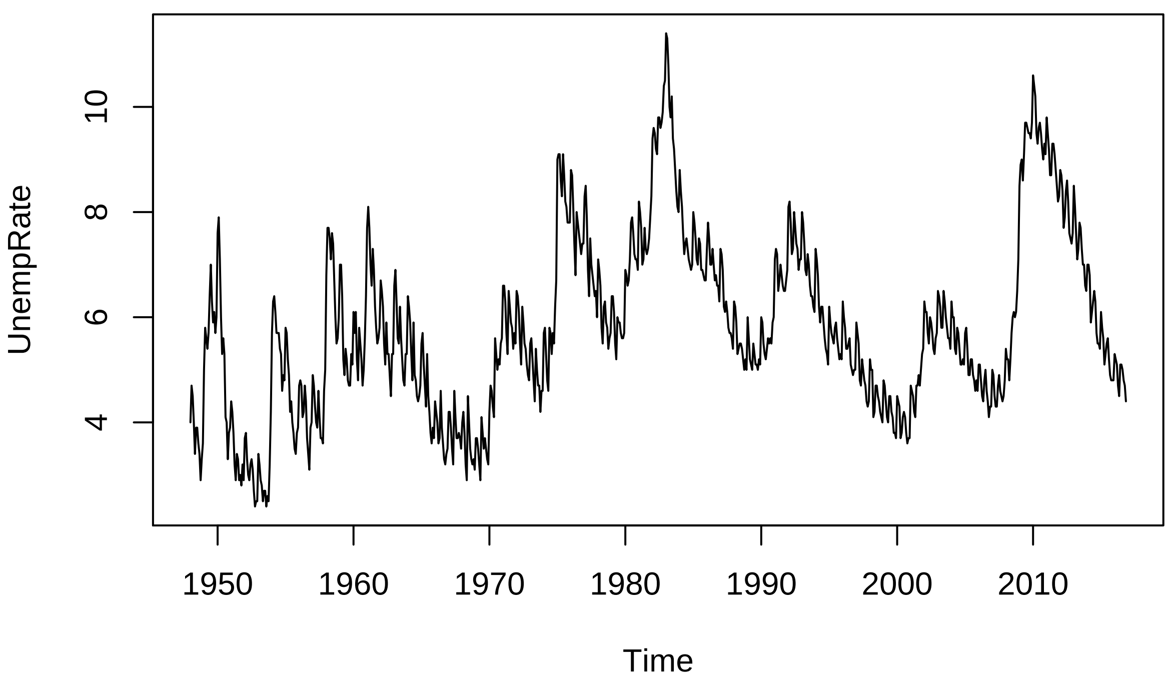
A4-Q3

Zhaoqi Li

2023/4/9

Step 1: Load the data and plot transformation

```
library("astsa")
plot(UnempRate)
```



The series is not stationary because it shows a seasonal pattern. The trend is slightly upward.

Step 2: Check how many differences are required to make the series stationary

```
library("forecast")
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo

##
```

```
## Attaching package: 'forecast'

## The following object is masked from 'package:astsa':
##
##     gas
```

```r
ndiffs(UnempRate)
```

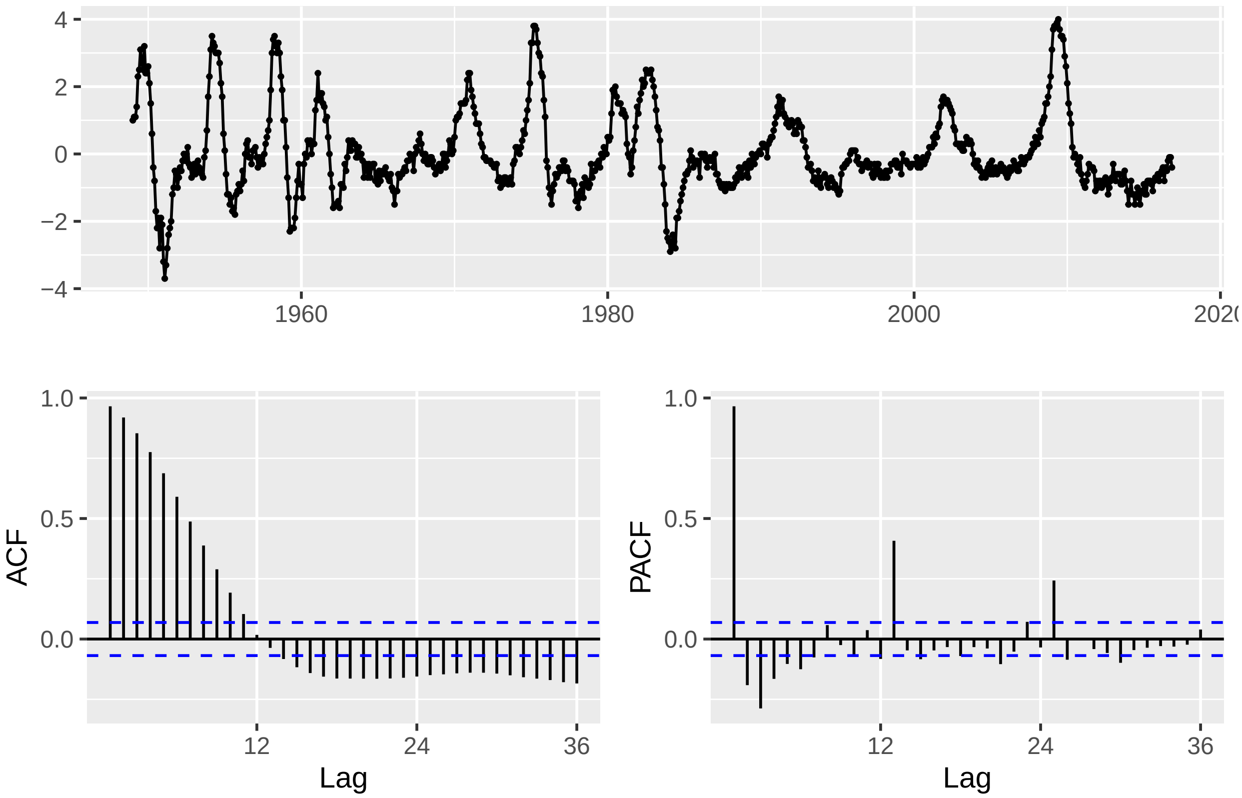
```
## [1] 1
```

Step 3: Take the first seasonal difference series and examine the plots of the series, ACF, and PACF

```r
library("tidyverse")
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.0     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.1     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.1
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

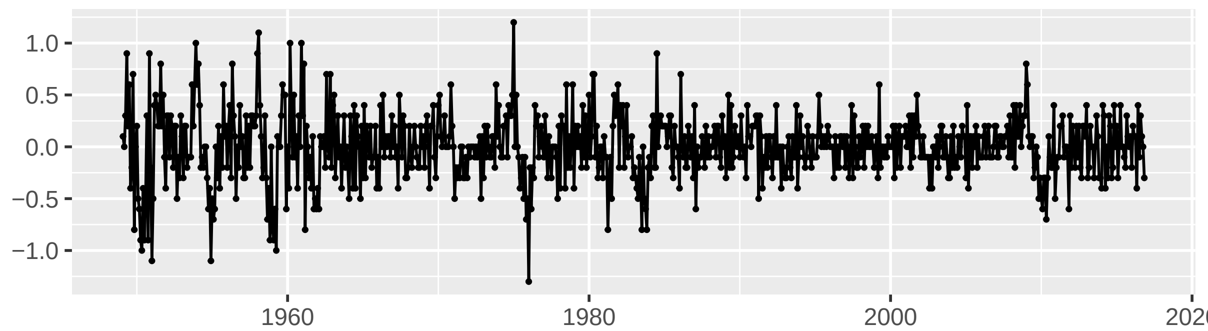
```r
UnempRate %>% diff(lag=12) %>% ggtsdisplay()
```

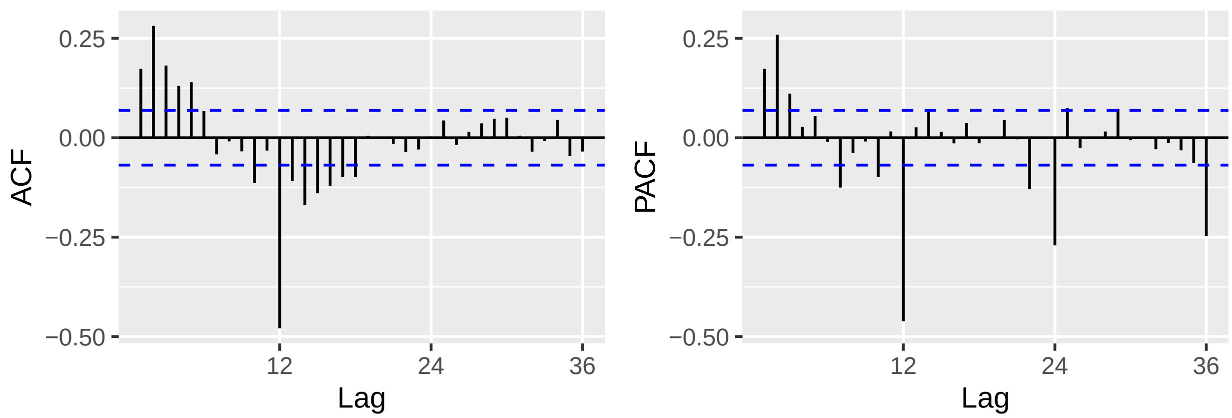


According to the plot, the ACF decrease, so we need non-seasonal difference.

Step 4: Take another first seasonal difference series and examine the plots of the series, ACF, and PACF

```r
UnempRate %>% diff(lag=12) %>% diff() %>%
ggtsdisplay()
```

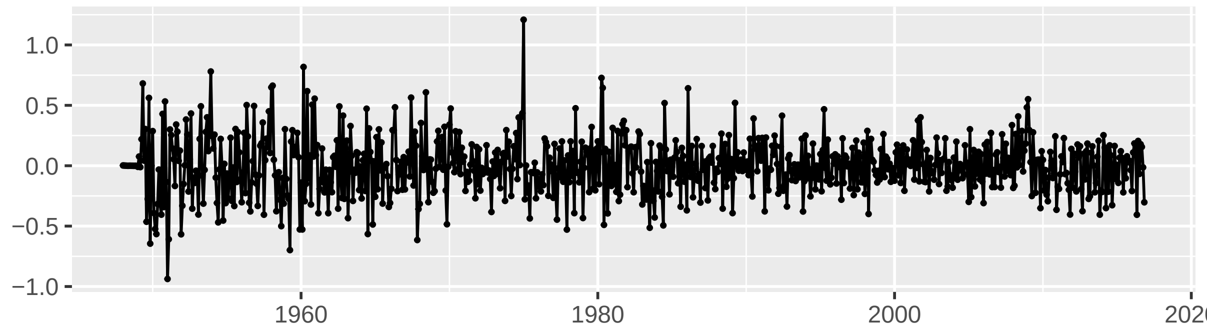


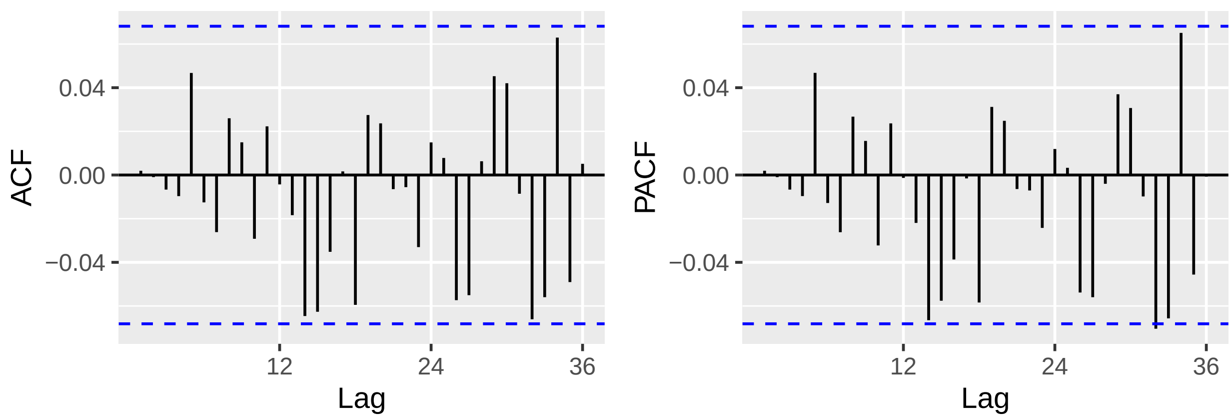


According to the plots, there are two spikes at lags 1, 2, 3 in PACF, so we suggest a non-seasonal AR(3). According to the ACF plot, there are significant spikes at lag 1, 2, 3, 4, 5 and lag 12 , so we suggest a non-seasonal MA(5) and a seasonal SMA(1). Thus, our tentative model is

$$ARIMA(3, 1, 5)\ddot{O}(0, 1, 1)_{12}$$

```r
UnempRate %>%
  Arima(order=c(3, 1, 5), seasonal = c(0, 1, 1)) %>%
  residuals() %>%
  ggtsdisplay()
```






```
fit <- Arima(UnempRate, order=c(3, 1, 5), seasonal = c(0, 1, 1))
fit$aic
```

```
## [1] -21.30229
```

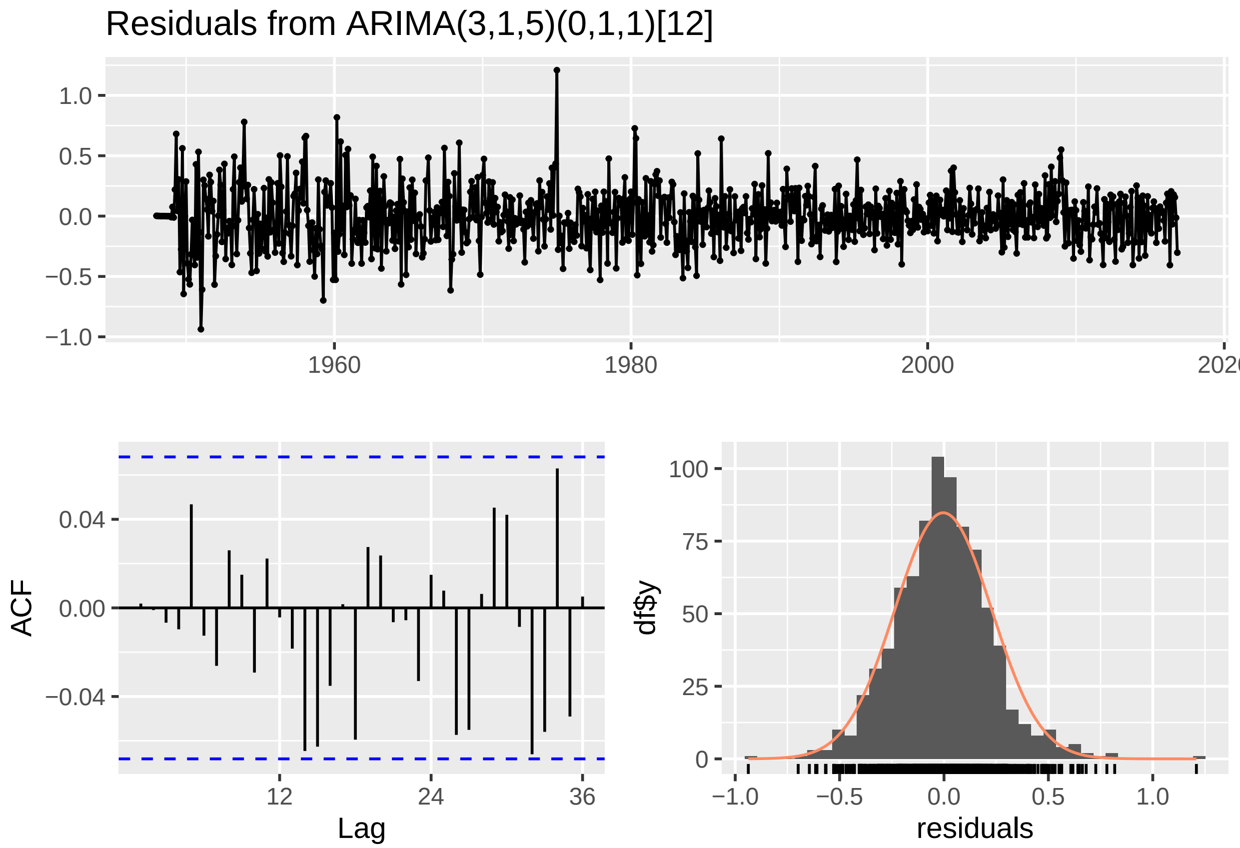
```
fit$aicc
```

```
## [1] -21.02831
```

```
fit$bic
```

```
## [1] 25.71732
```

```
checkresiduals(fit)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(3,1,5)(0,1,1)[12]
## Q* = 18.01, df = 15, p-value = 0.2621
##
## Model df: 9.   Total lags used: 24
```

According to the value of the BIC, we choose this model to fit the log series. According to the diagnostic plots of the residual and the Ljung-Box portmanteau test statistic, the p-value is 0.2621 ,so we indicate the residuals are white noise but the fitted model is not very good because the p-value is not very samll.

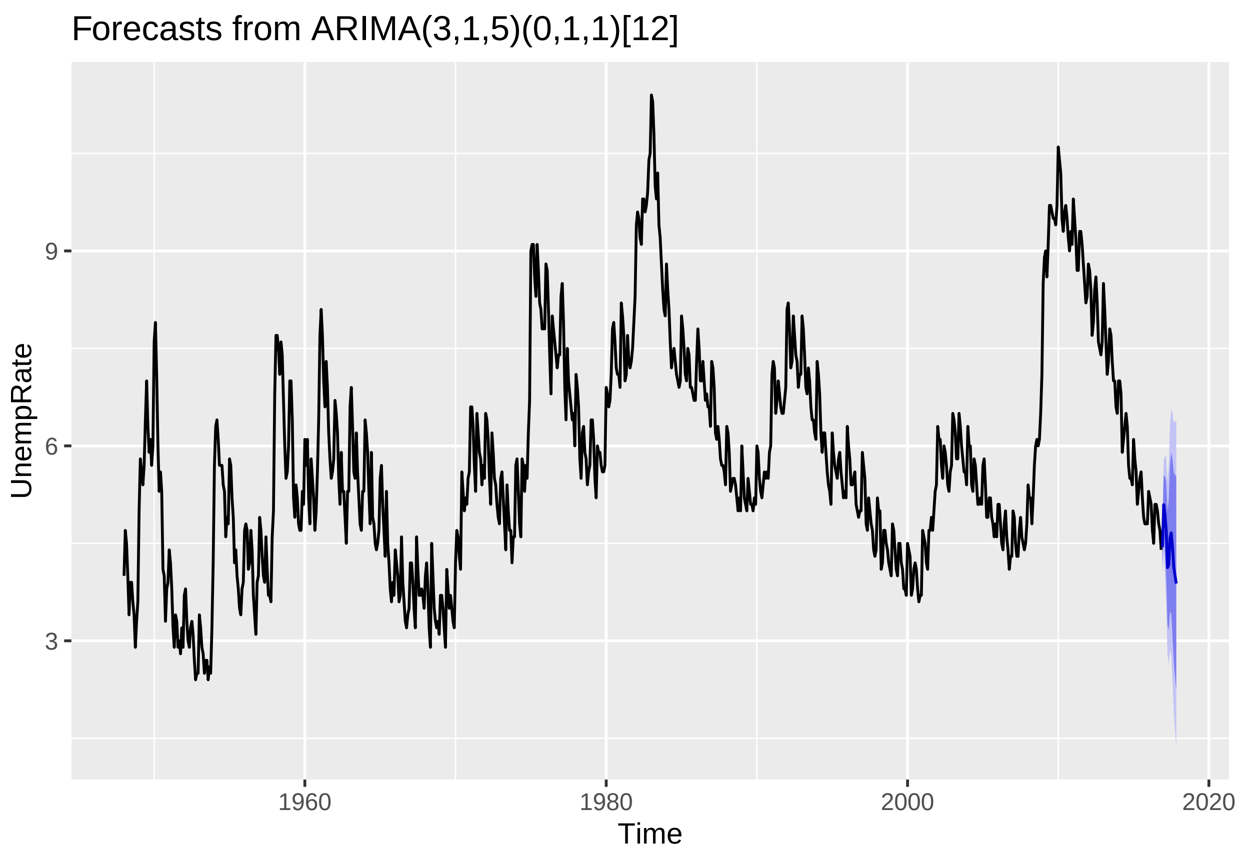
Step 5: Forecast the next 12 months

```
forecast(fit, h=12)
```

```
##          Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## Dec 2016       4.440720 4.139138 4.742302 3.979491 4.901950
## Jan 2017       5.095999 4.648426 5.543572 4.411495 5.780503
## Feb 2017       4.937205 4.342327 5.532084 4.027417 5.846993
## Mar 2017       4.702013 3.967341 5.436686 3.578429 5.825598
## Apr 2017       4.126218 3.255698 4.996739 2.794872 5.457564
## May 2017       4.166100 3.164105 5.168094 2.633681 5.698518
## Jun 2017       4.580188 3.453405 5.706972 2.856922 6.303455
## Jul 2017       4.661219 3.418226 5.904211 2.760225 6.562212
## Aug 2017       4.437242 3.086579 5.787906 2.371580 6.502904
## Sep 2017       4.131342 2.679588 5.583096 1.911075 6.351608
```

```
## Oct 2017       4.006854 2.457943 5.555766 1.637999 6.375710
## Nov 2017       3.881358 2.237379 5.525338 1.367108 6.395609
```

```
fit %>%
  forecast(h=12)%>%
  autoplot()
```

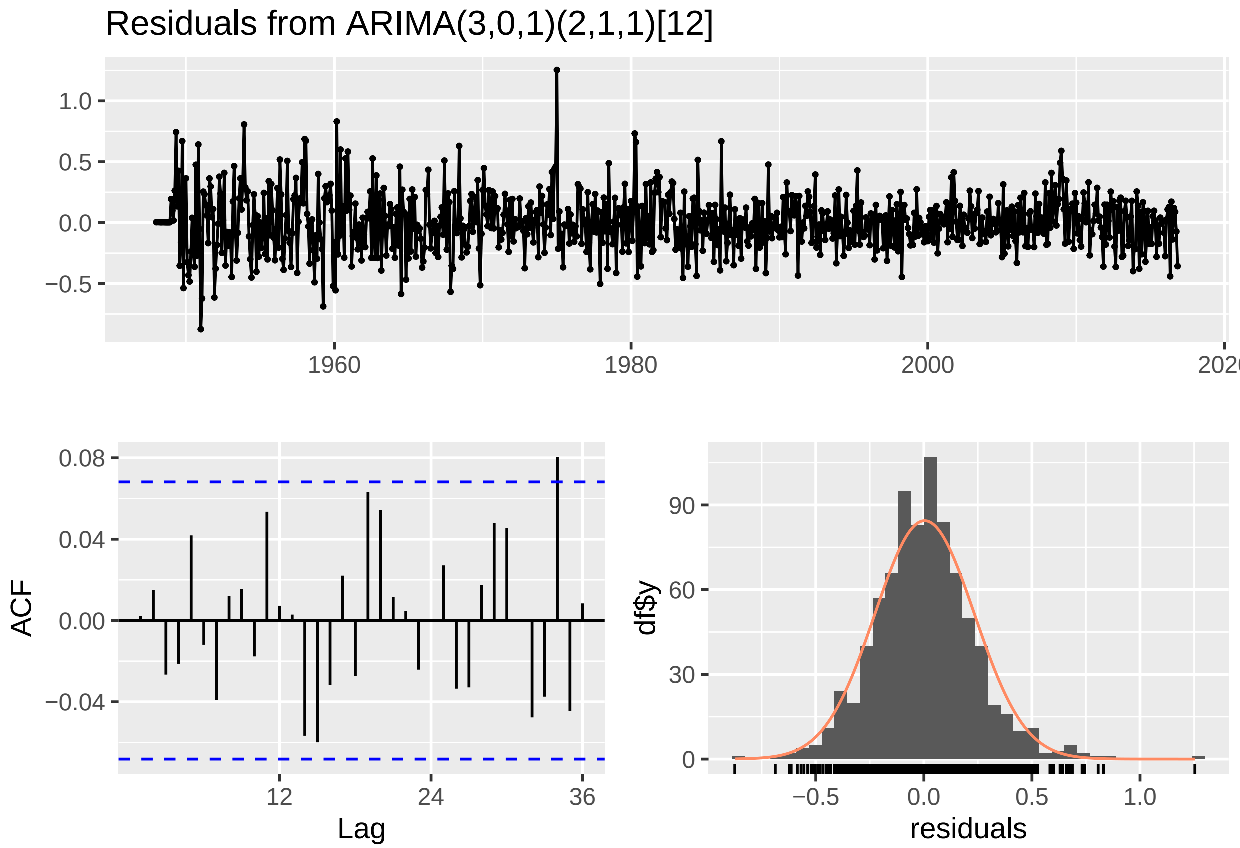


Compared with the perivous model, the new model with the function auto.arima() is better because the vaule of AIC, AICC, and BIC is samller.

```
fit.12 <- auto.arima(UnempRate)
fit.12
```

```
## Series: UnempRate
## ARIMA(3,0,1)(2,1,1)[12]
##
## Coefficients:
##          ar1      ar2      ar3      ma1    sar1    sar2     sma1
##       1.7057  -0.6043  -0.1124  -0.6292  0.0245  0.0286  -0.7769
## s.e.  0.0656   0.1289   0.0661   0.0606  0.1493  0.0947   0.0814
##
## sigma^2 = 0.05459:  log likelihood = 25.43
## AIC=-34.86   AICc=-34.68   BIC=2.77
```

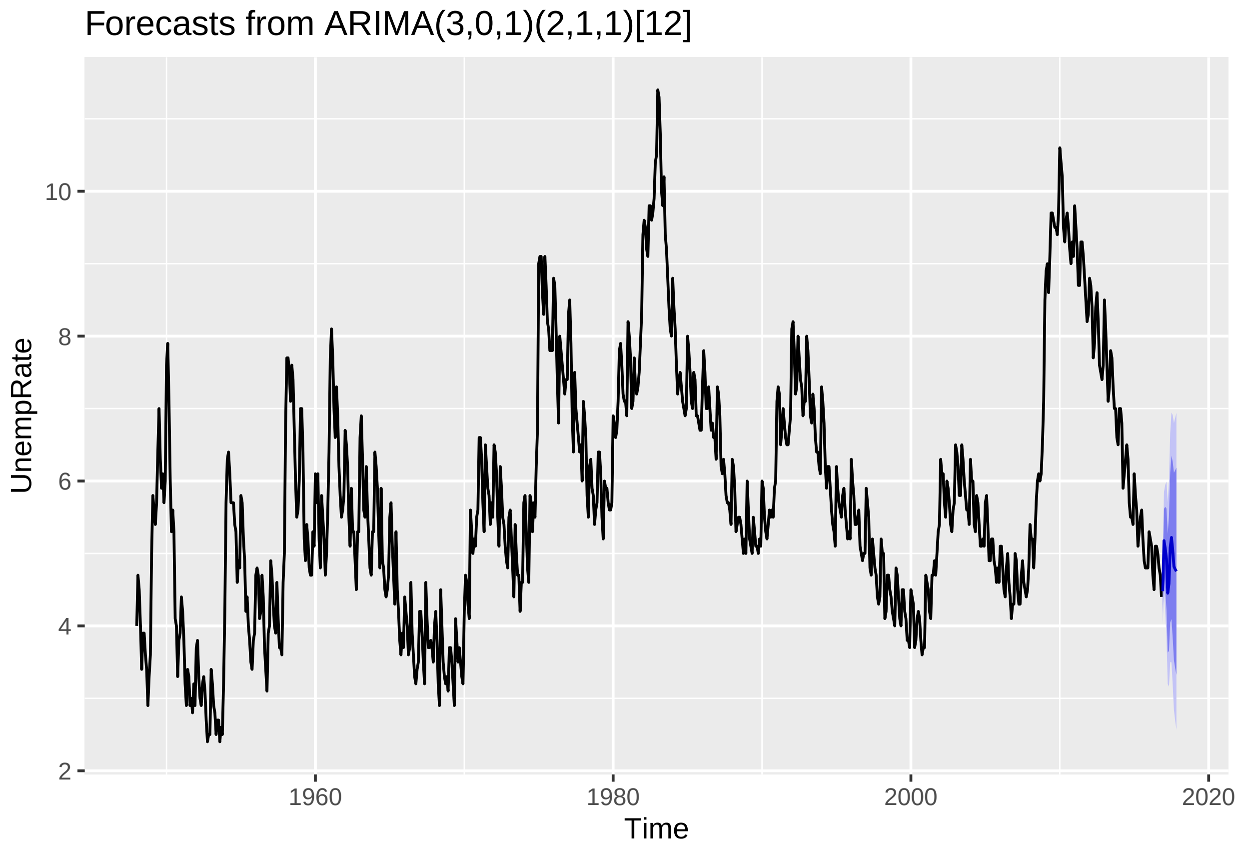
```
checkresiduals(fit.12)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(3,0,1)(2,1,1)[12]
## Q* = 21.239, df = 17, p-value = 0.2158
##
## Model df: 7.   Total lags used: 24
```

According to the plots, both models are available.

```
fit.12 %>%
  forecast(h=12)%>%
  autoplot()
```


Forecasts from ARIMA(3,0,1)(2,1,1)[12]
UnempRate
10
8
6
4
2
1960
1980
2000
2020
Time