> library(forecast)

> library(fUnitRoots)

> data <- gas

> # Australian monthly gas production

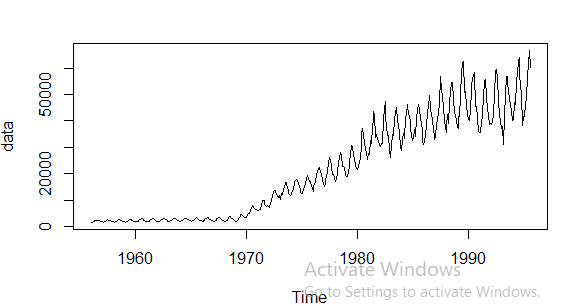
> tail(data)

Mar Apr May Jun Jul Aug

1995 46287 49013 56624 61739 66600 60054

> #plotting the data to see if there is trend

> plot(data)



> #checking if the data is stationary or non stationary by ADF test

> #if p-value > 0.05, data is non-stationary else stationary.

> #if data is non-stationary, change to stationary

> adfTest(data)

Title:

Augmented Dickey-Fuller Test

Test Results:

PARAMETER:

Lag Order: 1

STATISTIC:

Dickey-Fuller: -0.7909

P VALUE:

0.3643

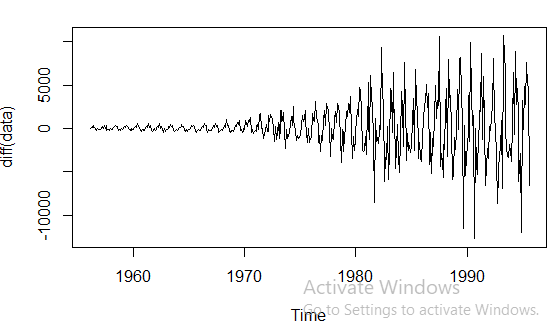
Description:

Sat Jul 04 10:19:10 2020 by user: Nabin

#changing data to stationary

#differencing the data

plot(diff(data))



> #checking again the p-value

> adfTest(diff(data))

Title:

Augmented Dickey-Fuller Test

Test Results:

PARAMETER:

Lag Order: 1

STATISTIC:

Dickey-Fuller: -10.7962

P VALUE:

0.01

Description:

Sat Jul 04 10:20:36 2020 by user: Nabin

Warning message:

In adfTest(diff(data)) : p-value smaller than printed p-value

> #running AR - I - MA test

> ARIMAfit <- auto.arima(log10(data), approximation = TRUE, trace=TRUE)

Fitting models using approximations to speed things up...

ARIMA(2,1,2)(1,1,1)[12] : -2096.181

ARIMA(0,1,0)(0,1,0)[12] : -1881.092

ARIMA(1,1,0)(1,1,0)[12] : -2013.482

ARIMA(0,1,1)(0,1,1)[12] : -2119.165

ARIMA(0,1,1)(0,1,0)[12] : -1945.621

ARIMA(0,1,1)(1,1,1)[12] : -2104.498

ARIMA(0,1,1)(0,1,2)[12] : Inf

ARIMA(0,1,1)(1,1,0)[12] : -2026.423

ARIMA(0,1,1)(1,1,2)[12] : -2109.87

ARIMA(0,1,0)(0,1,1)[12] : Inf

ARIMA(1,1,1)(0,1,1)[12] : -2116.249

ARIMA(0,1,2)(0,1,1)[12] : -2117.176

ARIMA(1,1,0)(0,1,1)[12] : -2115.881

ARIMA(1,1,2)(0,1,1)[12] : -2114.448

Now re-fitting the best model(s) without approximations...

ARIMA(0,1,1)(0,1,1)[12] : -2205.17

Best model: ARIMA(0,1,1)(0,1,1)[12]

> summary(ARIMAfit)

Series: log10(data)

ARIMA(0,1,1)(0,1,1)[12]

Coefficients:

ma1 sma1

-0.3304 -0.8256

s.e. 0.0462 0.0473

sigma^2 estimated as 0.0004815: log likelihood=1105.61

AIC=-2205.22 AICc=-2205.17 BIC=-2192.81

Training set error measures:

ME RMSE MAE MPE MAPE MASE

Training set 8.416862e-05 0.02159378 0.01546022 0.006812281 0.3798947 0.357045

ACF1

Training set -0.003186571

> #predicting the future values for three years i.e 36 month

> pred <- predict(ARIMAfit, n.ahead = 36)

> pred

$pred

Jan Feb Mar Apr May Jun Jul Aug Sep

1995 4.743679

1996 4.620859 4.627925 4.657632 4.679077 4.752725 4.788059 4.815171 4.802067 4.757308

1997 4.634489 4.641555 4.671261 4.692707 4.766355 4.801688 4.828801 4.815697 4.770938

1998 4.648118 4.655185 4.684891 4.706336 4.779985 4.815318 4.842430 4.829327

Oct Nov Dec

1995 4.712731 4.680952 4.632673

1996 4.726361 4.694582 4.646303

1997 4.739990 4.708212 4.659933

1998

$se

Jan Feb Mar Apr May Jun Jul

1995

1996 0.03667488 0.03950877 0.04215256 0.04464006 0.04699607 0.04923948 0.05138504

1997 0.06626041 0.06847058 0.07061160 0.07268958 0.07470979 0.07667680 0.07859458

1998 0.09214436 0.09425168 0.09631290 0.09833093 0.10030836 0.10224756 0.10415066

Aug Sep Oct Nov Dec

1995 0.02194228 0.02640756 0.03022011 0.03360284

1996 0.05344453 0.05656263 0.05913635 0.06160265 0.06397393

1997 0.08046668 0.08318292 0.08551137 0.08777808 0.08998770

1998 0.10601960

> #integral. showing the values

> 10^(pred$pred)

Jan Feb Mar Apr May Jun Jul Aug Sep

1995 55421.54

1996 41769.45 42454.67 45460.22 47761.38 56588.14 61384.47 65338.76 63396.75 57188.46

1997 43101.12 43808.19 46909.56 49284.08 58392.25 63341.50 67421.85 65417.93 59011.71

1998 44475.25 45204.86 48405.11 50855.33 60253.88 65360.92 69571.36 67503.55

Oct Nov Dec

1995 51609.64 47968.04 42921.35

1996 53255.03 49497.33 44289.74

1997 54952.88 51075.38 45701.77

1998

> plot(forecast(ARIMAfit, h=36))

