Time\_Series\_Assignment

library(tseries)

## Registered S3 method overwritten by 'xts':  
## method from  
## as.zoo.xts zoo

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

library(forecast)

## Registered S3 methods overwritten by 'ggplot2':  
## method from   
## [.quosures rlang  
## c.quosures rlang  
## print.quosures rlang

## Registered S3 methods overwritten by 'forecast':  
## method from   
## fitted.fracdiff fracdiff  
## residuals.fracdiff fracdiff

class(gas)

## [1] "ts"

class of dataset is “ts”, which refer to timeseries

frequency(gas)

## [1] 12

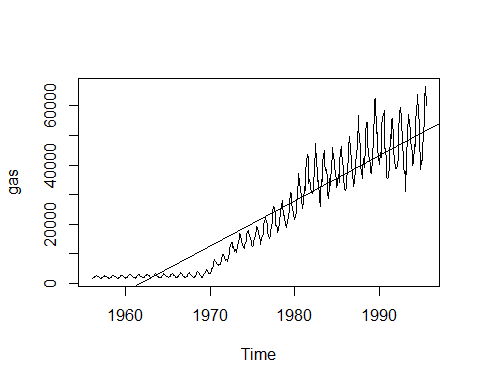
The Periodicity of Dataset is 12(Monthly)

gas

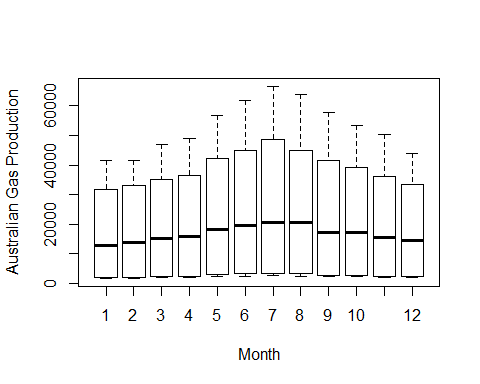
## Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov  
## 1956 1709 1646 1794 1878 2173 2321 2468 2416 2184 2121 1962  
## 1957 1751 1688 1920 1941 2311 2279 2638 2448 2279 2163 1941  
## 1958 1773 1688 1783 1984 2290 2511 2712 2522 2342 2195 1931  
## 1959 1730 1688 1899 1994 2342 2553 2712 2627 2363 2311 2026  
## 1960 1762 1815 2005 2089 2617 2828 2965 2891 2532 2363 2216  
## 1961 1804 1773 2015 2089 2627 2712 3007 2880 2490 2237 2205  
## 1962 1868 1815 2047 2142 2743 2775 3028 2965 2501 2501 2131  
## 1963 1910 1868 2121 2268 2690 2933 3218 3028 2659 2406 2258  
## 1964 1889 1984 2110 2311 2785 3039 3229 3070 2659 2543 2237  
## 1965 1962 1910 2216 2437 2817 3123 3345 3112 2659 2469 2332  
## 1966 1910 1941 2216 2342 2923 3229 3513 3355 2849 2680 2395  
## 1967 1994 1952 2290 2395 2965 3239 3608 3524 3018 2648 2363  
## 1968 1994 1941 2258 2332 3323 3608 3957 3672 3155 2933 2585  
## 1969 2057 2100 2458 2638 3292 3724 4652 4379 4231 3756 3429  
## 1970 3345 4220 4874 5064 5951 6774 7997 7523 7438 6879 6489  
## 1971 5919 6183 6594 6489 8040 9715 9714 9756 8595 7861 7753  
## 1972 7778 7402 8903 9742 11372 12741 13733 13691 12239 12502 11241  
## 1973 11569 10397 12493 11962 13974 14945 16805 16587 14225 14157 13016  
## 1974 11704 12275 13695 14082 16555 17339 17777 17592 16194 15336 14208  
## 1975 12354 12682 14141 14989 16159 18276 19157 18737 17109 17094 15418  
## 1976 13260 14990 15975 16770 19819 20983 22001 22337 20750 19969 17293  
## 1977 15117 16058 18137 18471 21398 23854 26025 25479 22804 19619 19627  
## 1978 17243 18284 20226 20903 23768 26323 28038 26776 22886 22813 22404  
## 1979 18839 18892 20823 22212 25076 26884 30611 30228 26762 25885 23328  
## 1980 21433 22369 24503 25905 30605 34984 37060 34502 31793 29275 28305  
## 1981 27730 27424 32684 31366 37459 41060 43558 42398 33827 34962 33480  
## 1982 30715 30400 31451 31306 40592 44133 47387 41310 37913 34355 34607  
## 1983 26138 30745 35018 34549 40980 42869 45022 40387 38180 38608 35308  
## 1984 28801 33034 35294 33181 40797 42355 46098 42430 41851 39331 37328  
## 1985 32494 33308 36805 34221 41020 44350 46173 44435 40943 39269 35901  
## 1986 31239 32261 34951 38109 43168 45547 49568 45387 41805 41281 36068  
## 1987 32791 34206 39128 40249 43519 46137 56709 52306 49397 45500 39857  
## 1988 35567 37696 42319 39137 47062 50610 54457 54435 48516 43225 42155  
## 1989 37541 37277 41778 41666 49616 57793 61884 62400 50820 51116 45731  
## 1990 40459 40295 44147 42697 52561 56572 56858 58363 45627 45622 41304  
## 1991 35592 35677 39864 41761 50380 49129 55066 55671 49058 44503 42145  
## 1992 38963 38690 39792 42545 50145 58164 59035 59408 55988 47321 42269  
## 1993 37059 37963 31043 41712 50366 56977 56807 54634 51367 48073 46251  
## 1994 39975 40478 46895 46147 55011 57799 62450 63896 57784 53231 50354  
## 1995 41600 41471 46287 49013 56624 61739 66600 60054   
## Dec  
## 1956 1825  
## 1957 1878  
## 1958 1910  
## 1959 1910  
## 1960 2026  
## 1961 1984  
## 1962 2015  
## 1963 2057  
## 1964 2142  
## 1965 2110  
## 1966 2205  
## 1967 2247  
## 1968 2384  
## 1969 3461  
## 1970 6288  
## 1971 8154  
## 1972 10829  
## 1973 12253  
## 1974 13116  
## 1975 14312  
## 1976 16498  
## 1977 18488  
## 1978 19795  
## 1979 21930  
## 1980 25248  
## 1981 32445  
## 1982 28729  
## 1983 30234  
## 1984 34514  
## 1985 32142  
## 1986 34879  
## 1987 37958  
## 1988 39995  
## 1989 42528  
## 1990 36016  
## 1991 38698  
## 1992 39606  
## 1993 43736  
## 1994 38410  
## 1995

Timeseries Data is available from January-1956 to August-1995

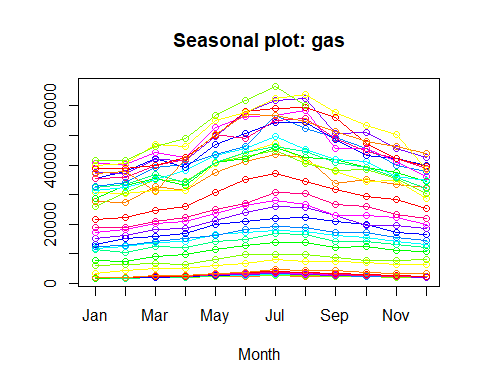
plot(gas)  
abline(reg = lm(gas~time(gas)))



boxplot(gas~cycle(gas),xlab = "Month",ylab = "Australian Gas Production")



library(forecast)  
seasonplot(gas,col = rainbow(12))



###Observation:

The Gas production increase over time with each year which may be indicative of an increasing linear trend.

In the boxplot there is more production of Gas during months May to August with higher means and higher variances than the other months, indicating seasonality with a apparent cycle of 12 months.

###Data Stationarity Check

The Australian Gas Production dataset in not stationary, becasue the abline indicate bothe the mean and variance are not constant. To validate this, KPSS test can be Performed

H0: Series is Stationary Ha: Series is not stationary

kpss.test(gas)

## Warning in kpss.test(gas): p-value smaller than printed p-value

##   
## KPSS Test for Level Stationarity  
##   
## data: gas  
## KPSS Level = 7.7018, Truncation lag parameter = 5, p-value = 0.01

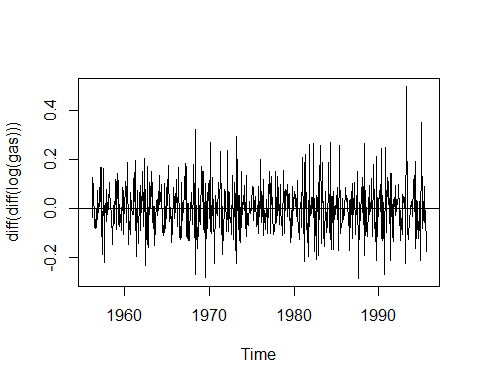
The P-Value from this test is 0.01 and we reject Null hypothesis, which indicates that the series is not stationary.

To make the series stationry,some transformation functions are used

Log transformation to variance constant

Differencing can be used to make mean as constant over time.

#plot(diff(log(gas)))  
plot(diff(diff(log(gas))))  
abline(h=0)



kpss.test(diff(diff(log(gas))))

## Warning in kpss.test(diff(diff(log(gas)))): p-value greater than printed p-  
## value

##   
## KPSS Test for Level Stationarity  
##   
## data: diff(diff(log(gas)))  
## KPSS Level = 0.0061339, Truncation lag parameter = 5, p-value =  
## 0.1

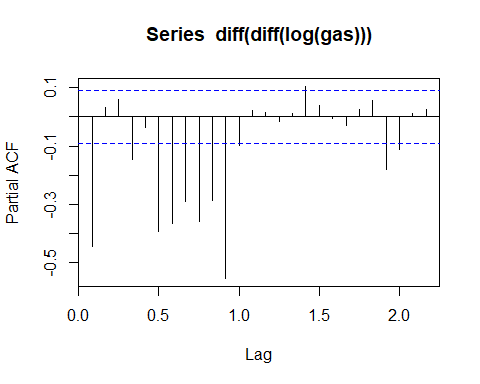
The P-Value from this test is 0.1 and we accept Null hypothesis, which indicates that the series is stationary.

### ARIMA Model

####PACF Plot

To determine how many previous periods required to forecast next peiods

pacf(diff(diff(log(gas))))

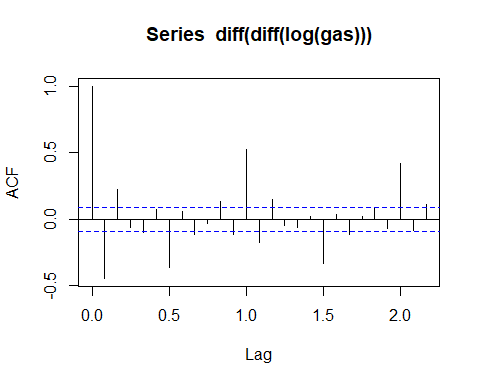


P=1

### ACF plot

How many previous errors are impacting the value to be forecasted.

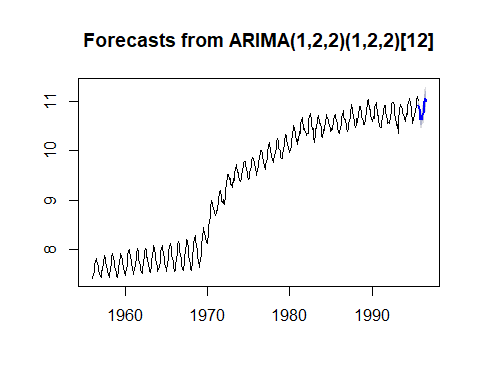
acf(diff(diff(log(gas))))

 q=2

arima\_model=arima(log(gas),c(1,2,2),seasonal = list(order=c(1,2,2),period=12))  
summary(arima\_model)

##   
## Call:  
## arima(x = log(gas), order = c(1, 2, 2), seasonal = list(order = c(1, 2, 2),   
## period = 12))  
##   
## Coefficients:  
## ar1 ma1 ma2 sar1 sma1 sma2  
## 0.0971 -1.4255 0.4429 0.1733 -1.8952 0.8953  
## s.e. 0.1533 0.1385 0.1395 0.0694 0.1281 0.1204  
##   
## sigma^2 estimated as 0.002521: log likelihood = 651.33, aic = -1288.65  
##   
## Training set error measures:  
## ME RMSE MAE MPE MAPE  
## Training set -0.0003511113 0.04883609 0.03481337 -0.004339459 0.3669012  
## MASE ACF1  
## Training set 0.4189708 -0.004583663

arima\_pred<-forecast(arima\_model,12)  
plot(forecast(arima\_pred,12))



accuracy(arima\_pred)

## ME RMSE MAE MPE MAPE  
## Training set -0.0003511113 0.04883609 0.03481337 -0.004339459 0.3669012  
## MASE ACF1  
## Training set 0.3491706 -0.004583663