

hw3

2022-09-21

We are gonna explore xxx dataset from fpp ##### package loading

```
library(fpp)
```

```
## Loading required package: forecast
## Registered S3 method overwritten by 'quantmod':
##   method             from
##   as.zoo.data.frame zoo
## Loading required package: fma
## Loading required package: expsmooth
## Loading required package: lmtest
## Loading required package: zoo
##
## Attaching package: 'zoo'
## The following objects are masked from 'package:base':
##
##   as.Date, as.Date.numeric
## Loading required package: tseries
```

```
library(fpp2)
```

```
## -- Attaching packages ----- fpp2 2.4 --
## v ggplot2 3.3.6
##
##
## Attaching package: 'fpp2'
## The following objects are masked from 'package:fpp':
##
##   ausair, ausbeer, austa, austourists, debitcards, departures,
##   elecequip, euretail, guinearice, oil, sunspotarea, usmelec
####showing exesisting dataset.
```

```
data()
```

```
####Explore the dataset of UKgas.
```

```
window(UKgas)
```

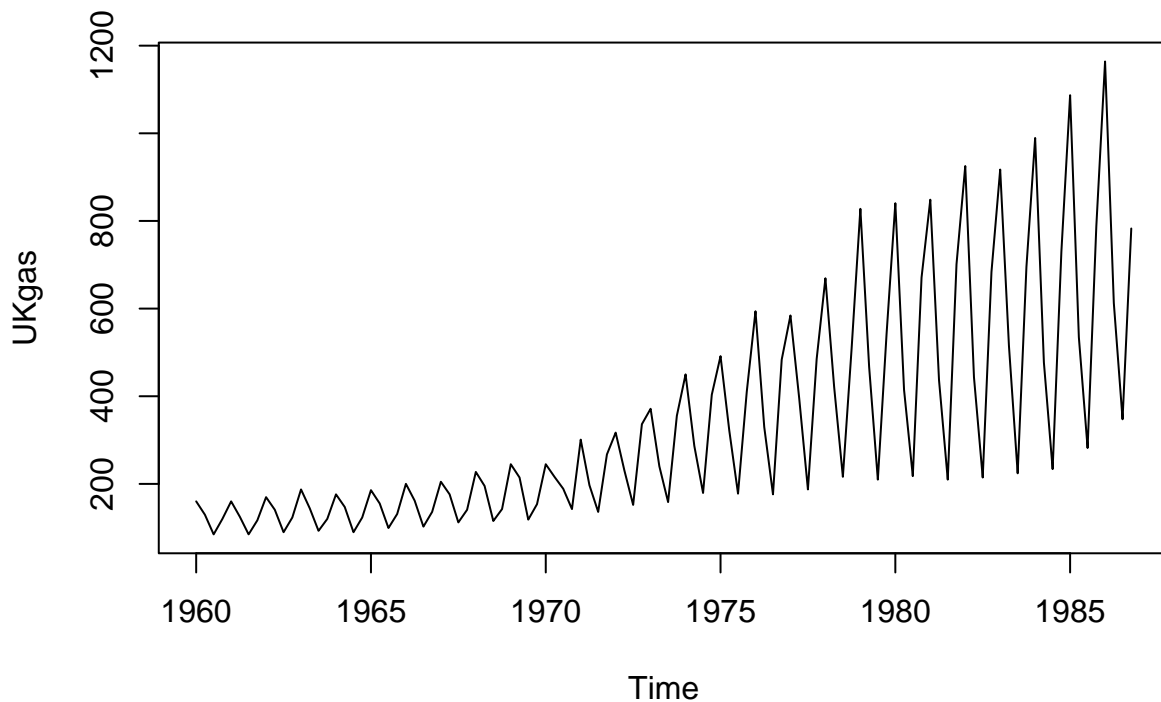
```
##           Qtr1    Qtr2    Qtr3    Qtr4
## 1960    160.1    129.7     84.8    120.1
```

```
## 1961 160.1 124.9 84.8 116.9
## 1962 169.7 140.9 89.7 123.3
## 1963 187.3 144.1 92.9 120.1
## 1964 176.1 147.3 89.7 123.3
## 1965 185.7 155.3 99.3 131.3
## 1966 200.1 161.7 102.5 136.1
## 1967 204.9 176.1 112.1 140.9
## 1968 227.3 195.3 115.3 142.5
## 1969 244.9 214.5 118.5 153.7
## 1970 244.9 216.1 188.9 142.5
## 1971 301.0 196.9 136.1 267.3
## 1972 317.0 230.5 152.1 336.2
## 1973 371.4 240.1 158.5 355.4
## 1974 449.9 286.6 179.3 403.4
## 1975 491.5 321.8 177.7 409.8
## 1976 593.9 329.8 176.1 483.5
## 1977 584.3 395.4 187.3 485.1
## 1978 669.2 421.0 216.1 509.1
## 1979 827.7 467.5 209.7 542.7
## 1980 840.5 414.6 217.7 670.8
## 1981 848.5 437.0 209.7 701.2
## 1982 925.3 443.4 214.5 683.6
## 1983 917.3 515.5 224.1 694.8
## 1984 989.4 477.1 233.7 730.0
## 1985 1087.0 534.7 281.8 787.6
## 1986 1163.9 613.1 347.4 782.8
```

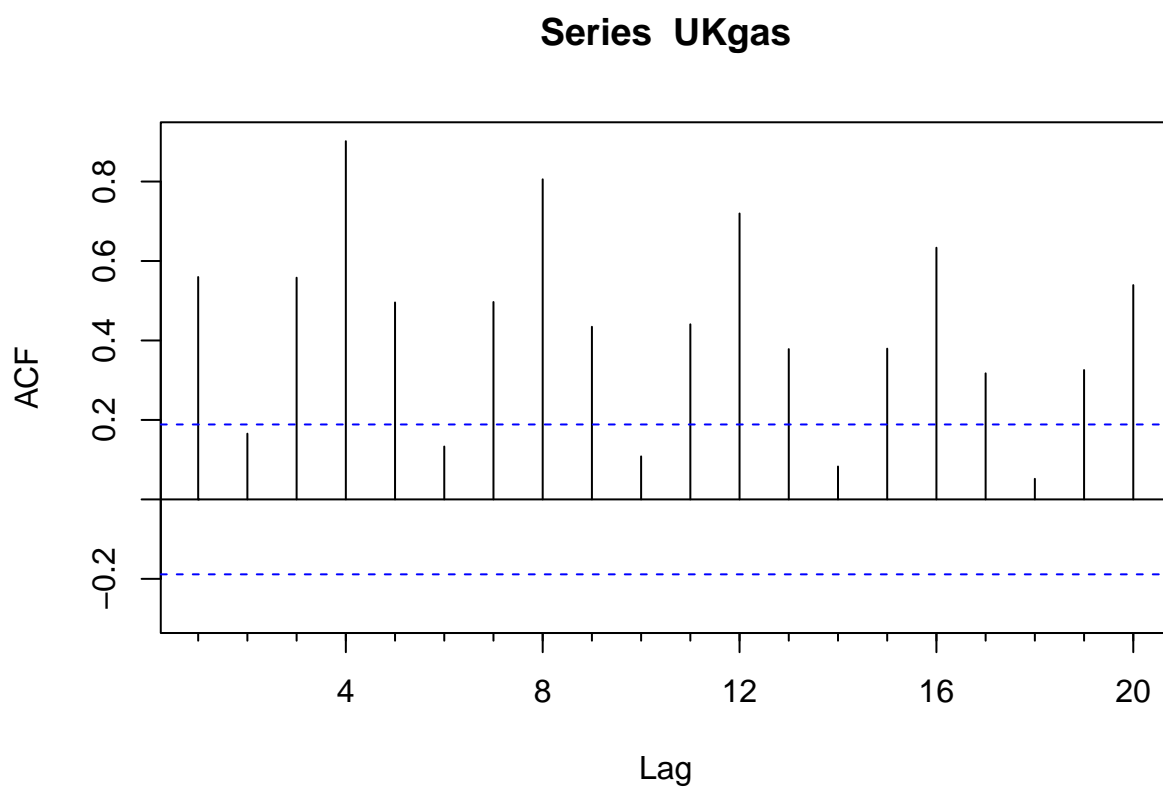
```
str(UKgas)
```

```
## Time-Series [1:108] from 1960 to 1987: 160.1 129.7 84.8 120.1 160.1 ...
```

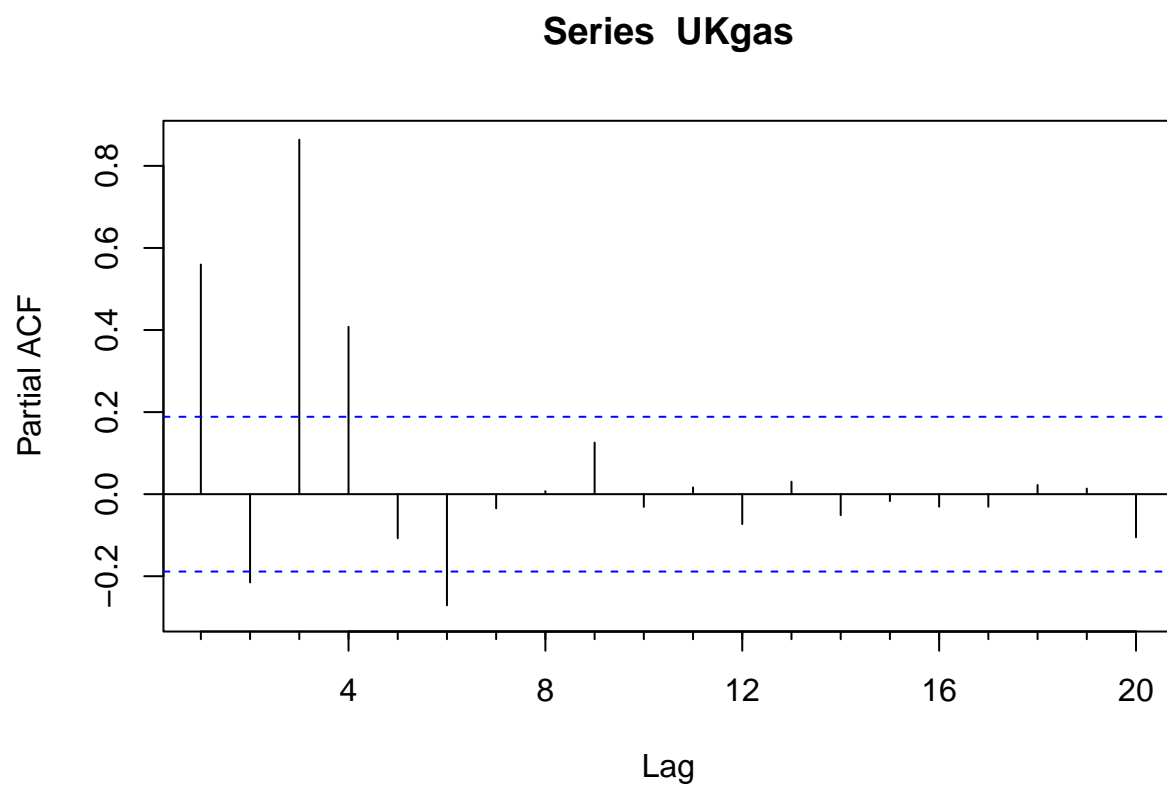
```
plot(UKgas)
```



Acf(UKgas)



Pacf(UKgas)

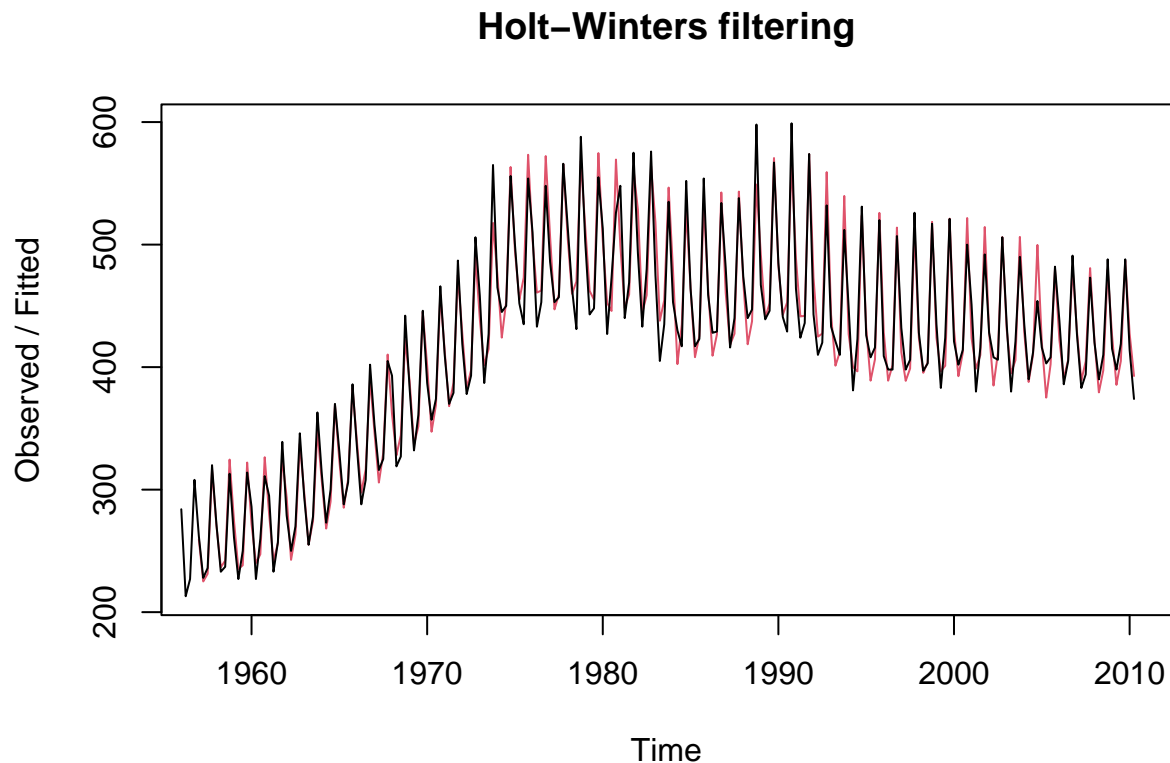


In figure1 we can see that all the lines showing a nice trending with no sudden drop and increase over time means that we can use the whole data set instead shorten its periodicity. We can see that in the UK gas usage between 1960 and 1987 auto correlation in figure2, there is a strong coefficient in the gap of 4 quarters(a year). And since it is all positive and dropping, we can say that it has a likely growing seasonal trend. However, this still need further exploration.

####Models As the figure 2 and figure 3 showed above, the coefficient are significant at lag of 4 in ACF and tail off in PACF, we should use MA models.

```
UKgasHotltWinters<- HoltWinters(ausbeer)
attributes(UKgasHotltWinters)
```

```
## $names
## [1] "fitted"      "x"           "alpha"       "beta"        "gamma"
## [6] "coefficients" "seasonal"    "SSE"         "call"
##
## $class
## [1] "HoltWinters"
plot(UKgasHotltWinters)
```



##we can see that the predictions have generally small errors.

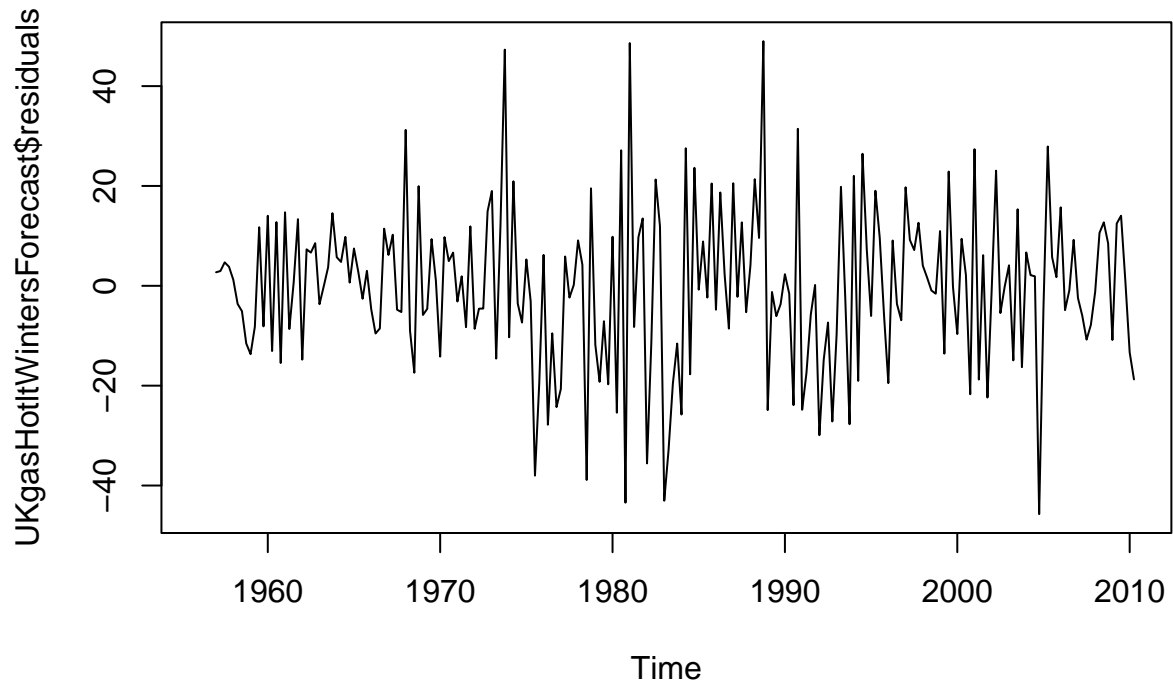
```
UKgasHotltWintersForecast <- forecast(UKgasHotltWinters)
attributes(UKgasHotltWintersForecast)
```

```
## $names
## [1] "method"      "model"       "level"       "mean"        "lower"       "upper"
## [7] "x"           "series"      "fitted"      "residuals"
##
## $class
## [1] "forecast"
```

```
head(UKgasHotltWintersForecast$residuals,20)
```

```
##           Qtr1           Qtr2           Qtr3           Qtr4
## 1956         NA            NA            NA            NA
## 1957    2.712500    2.957591    4.695659    3.809655
## 1958    1.202940   -3.629963   -5.079526  -11.545379
## 1959  -13.688612   -8.196576   11.733310   -8.128276
## 1960   14.026249  -13.066206   12.763914  -15.464205
```

```
## "guess": the first row of residuals empty because the forecast use the 1956 data as a start point to
plot(UKgasHotltWintersForecast$residuals)
```

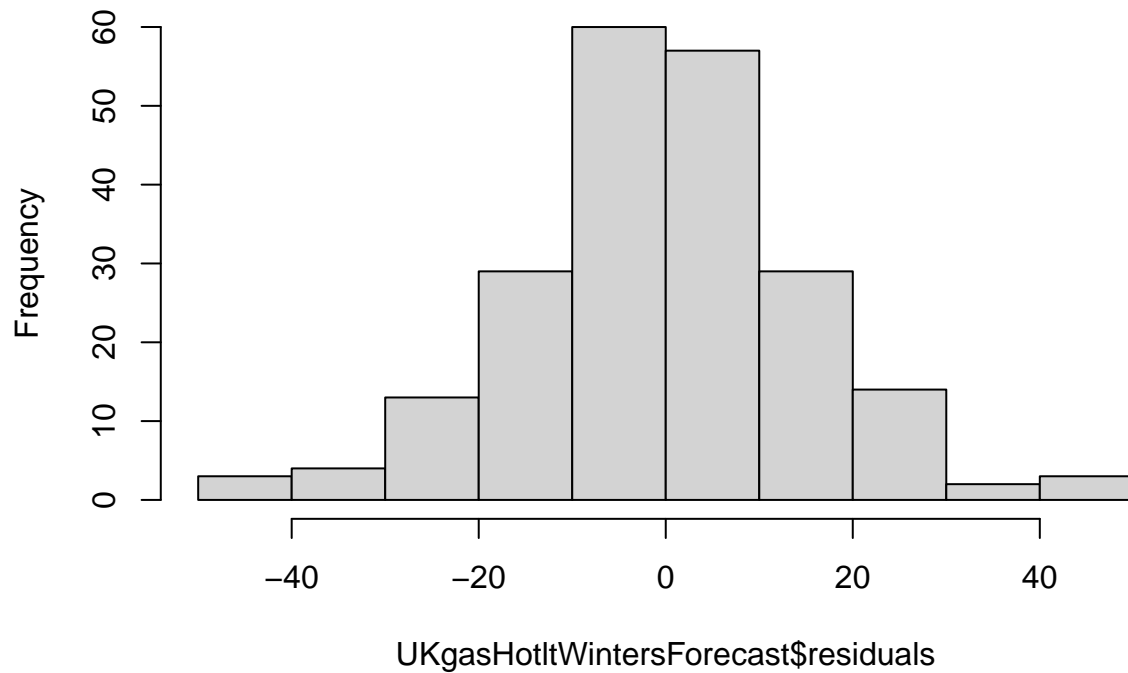


```
#####Display the Histogram of prediction's residuals
```

we can see that the distribution of the residuals has a mean close to 0, completely random, and since the graph doesn't have severe skewed so that it have a constant std deviation.

```
hist(UKgasHotltWintersForecast$residuals)
```

Histogram of UKgasHotltWintersForecast\$residuals



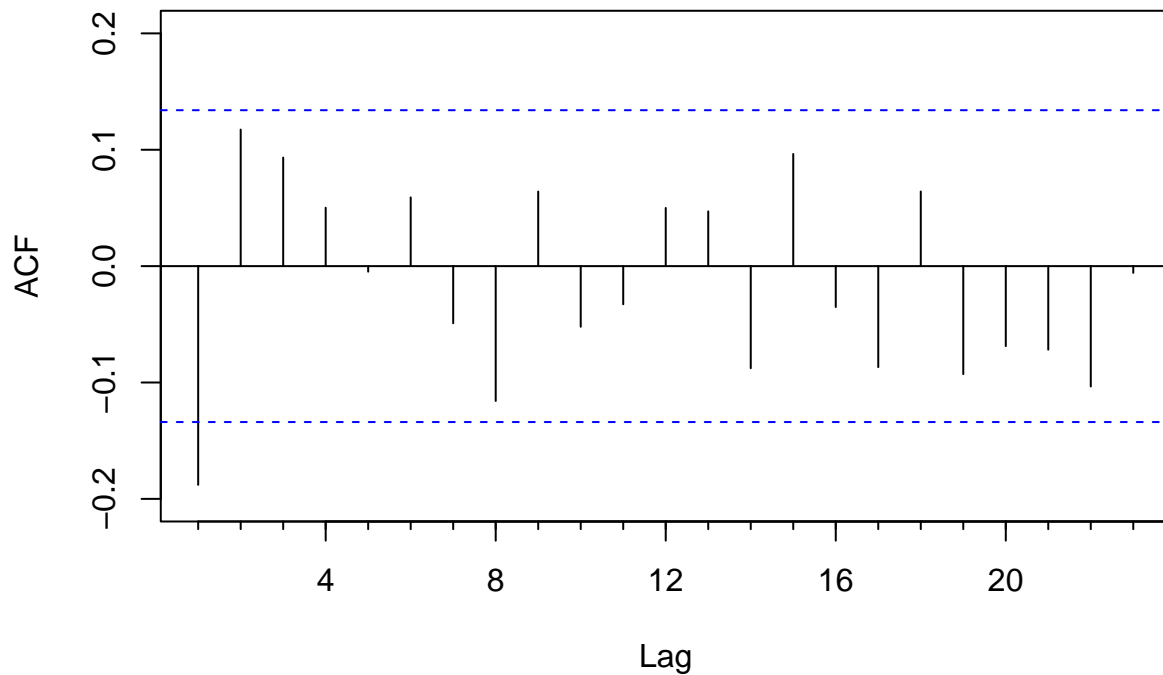
```
summary(UKgasHotltWintersForecast$residuals)
```

```
##      Min.   1st Qu.   Median     Mean  3rd Qu.    Max.     NA's  
## -45.7208  -8.8918   -0.1887   -0.3702   9.5377   48.9959      4
```

```
####Residuals' pattern Now we should find that our prediction's residuals have a pattern or not.
```

```
Acf(UKgasHotltWintersForecast$residuals)
```

Series UKgasHotltWintersForecast\$residuals



We can see that our residual don't have seasonality.

####Showing the Accuracies. Personaly, I think MAPE and RMSE are in resonable range.

```
accuracy(UKgasHotltWintersForecast)
```

```
##              ME      RMSE      MAE      MPE      MAPE      MASE
## Training set -0.3702353 15.86279 12.15598 -0.1241059 2.874404 0.7837843
##              ACF1
## Training set -0.1879524
```