Mach 1 ARIMA

```
library(readxl)
Tng_Ctr_Hour2 <- read_excel("Tng_Ctr_Hour2.xlsx")
View(Tng_Ctr_Hour2)
```

```
library(data.table)
library(ggplot2)
library(stringr)
library(TTR)
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 package: 'fpp2'

## The following objects are masked from 'package:fpp':
##
##     ausair, ausbeer, austa, austourists, debitcards, departures,
##     elecequip, euretail, guinearice, oil, sunspotarea, usmelec
```

```
library(fpp3)
```

```
## -- Attaching packages --------------------------------------- fpp3 0.4.0 --

## v tibble      3.1.5     v tsibble     1.1.0
## v dplyr       1.0.7     v tsibbledata 0.3.0
## v tidyr       1.1.4     v feasts      0.2.2
## v lubridate   1.8.0     v fable       0.3.1

## -- Conflicts ------------------------------------------------ fpp3_conflicts --
## x dplyr::between()       masks data.table::between()
## x lubridate::date()      masks base::date()
## x dplyr::filter()        masks stats::filter()
## x dplyr::first()         masks data.table::first()
## x fabletools::forecast() masks forecast::forecast()
## x lubridate::hour()      masks data.table::hour()
## x tsibble::index()       masks zoo::index()
## x tsibble::intersect()   masks base::intersect()
## x tsibble::interval()    masks lubridate::interval()
## x lubridate::isoweek()   masks data.table::isoweek()
## x tsibble::key()         masks data.table::key()
## x dplyr::lag()           masks stats::lag()
## x dplyr::last()          masks data.table::last()
## x lubridate::mday()      masks data.table::mday()
## x lubridate::minute()    masks data.table::minute()
## x lubridate::month()     masks data.table::month()
## x lubridate::quarter()   masks data.table::quarter()
## x lubridate::second()    masks data.table::second()
## x tsibble::setdiff()     masks base::setdiff()
## x tsibble::union()       masks base::union()
## x lubridate::wday()      masks data.table::wday()
## x lubridate::week()      masks data.table::week()
## x lubridate::yday()      masks data.table::yday()
## x lubridate::year()      masks data.table::year()

##
## Attaching package: 'fpp3'

## The following object is masked from 'package:fpp2':
##
##     insurance

## The following object is masked from 'package:fpp':
##
##     insurance
```

```
library(ggplot2)
library(stats)
library(dplyr)
library(graphics)
library(ggfortify)
```

```
## Registered S3 methods overwritten by 'ggfortify':
##   method                 from
##   autoplot.Arima         forecast
##   autoplot.acf           forecast
##   autoplot.ar            forecast
##   autoplot.bats          forecast
##   autoplot.decomposed.ts forecast
##   autoplot.ets           forecast
##   autoplot.forecast      forecast
##   autoplot.stl           forecast
##   autoplot.ts            forecast
##   fitted.ar              forecast
##   fortify.ts             forecast
##   residuals.ar           forecast
```

```
summary(Tng_Ctr_Hour2)
```

```
##      Year                Yr              Quarter            Month
##  Length:81          Length:81          Length:81          Length:81
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##    Device_Hrs     DH_Prev_Year       DH_YoY_Change      DH_YoY_Ch_Per
##  Min.   : 222.8   Length:81          Length:81          Length:81
##  1st Qu.: 899.0   Class :character   Class :character   Class :character
##  Median :1008.0   Mode  :character   Mode  :character   Mode  :character
##  Mean   : 990.1
##  3rd Qu.:1101.7
##  Max.   :1519.9
##
##  Total_Inst_Hrs   Total_Inst_Hrs_Prev_Year Inst_Hrs_YoY_Change
##  Min.   : 504.6   Length:81                Length:81
##  1st Qu.:1937.3   Class :character         Class :character
##  Median :2203.2   Mode  :character         Mode  :character
##  Mean   :2165.7
##  3rd Qu.:2446.8
##  Max.   :3084.1
##
##  Total_Inst_Hrs_YoY_Change_Per2   Cons_Sent          NJURN
##  Length:81                      Min.   : 70.30   Min.   : 2.900
##  Class :character               1st Qu.: 89.00   1st Qu.: 4.100
##  Mode  :character               Median : 93.80   Median : 4.900
##                                 Mean   : 91.49   Mean   : 5.615
##                                 3rd Qu.: 97.90   3rd Qu.: 6.200
##                                 Max.   :101.40   Max.   :16.600
##
##       RPM               CPIUrban       CPIMedian
##  Min.   :  2908236   Min.   :234.7   Min.   :0.9755
##  1st Qu.: 68459347   1st Qu.:241.2   1st Qu.:2.1551
##  Median : 77115921   Median :250.8   Median :2.5922
##  Mean   : 70822495   Mean   :250.2   Mean   :2.5862
```

```
##  3rd Qu.: 85326186   3rd Qu.:257.4   3rd Qu.:2.9557
##  Max.   :101794185   Max.   :274.1   Max.   :5.5690
##  NA's   :1
```

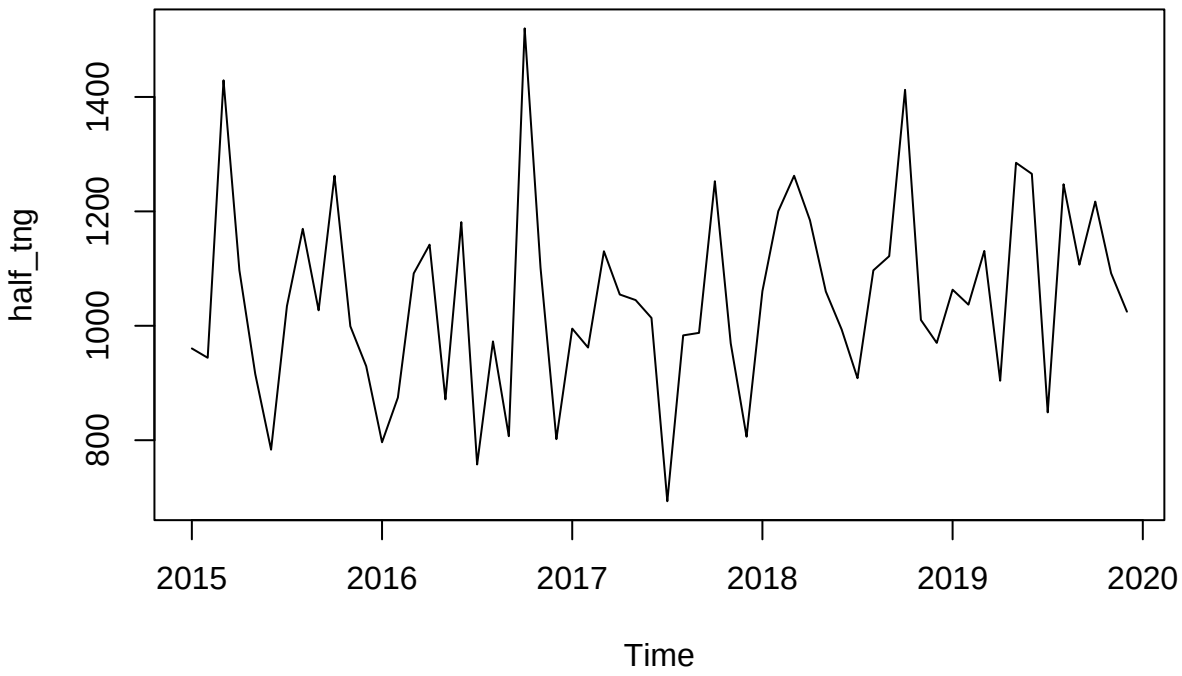
```
df_Tng = Tng_Ctr_Hour2[,c(5)]
df_Tng
```

```
## # A tibble: 81 x 1
##    Device_Hrs
##         <dbl>
##  1       960.
##  2       944.
##  3      1429.
##  4      1097
##  5       916.
##  6       783.
##  7      1035.
##  8      1170.
##  9      1027.
## 10      1262.
## # ... with 71 more rows
```

```
half_tng = ts(data = df_Tng,frequency = 12,start = c(2015, 1),end = c(2019,12))
half_tng
```

```
##          Jan     Feb     Mar     Apr     May     Jun     Jul     Aug     Sep
## 2015  960.42  944.08 1429.12 1097.00  915.85  783.45 1034.52 1169.50 1027.08
## 2016  796.42  874.55 1091.55 1141.84  871.36 1181.21  757.59  972.73  807.02
## 2017  995.09  962.00 1130.24 1054.71 1044.95 1013.73  693.33  983.25  987.64
## 2018 1060.57 1200.25 1262.25 1184.45 1059.92  993.55  908.37 1096.93 1121.75
## 2019 1063.13 1036.95 1130.87  903.97 1284.95 1265.56  848.64 1247.40 1106.84
##          Oct     Nov     Dec
## 2015 1262.32  999.25  929.42
## 2016 1519.92 1101.67  801.83
## 2017 1252.69  969.31  806.10
## 2018 1412.47 1010.25  970.12
## 2019 1217.08 1091.84 1024.67
```

```
plot(half_tng)
```




```
adf.test(half_tng)
```

```
## Warning in adf.test(half_tng): p-value smaller than printed p-value
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  half_tng
## Dickey-Fuller = -6.2576, Lag order = 3, p-value = 0.01
## alternative hypothesis: stationary
```

P-Value is less than 0.05, differences are not needed

```
kpss.test(half_tng)
```

```
##
##  KPSS Test for Level Stationarity
##
## data:  half_tng
## KPSS Level = 0.37795, Truncation lag parameter = 3, p-value = 0.08666
```

P- value is greater than 0.05, so differences is not needed.

```
nsdiffs(half_tng)
```

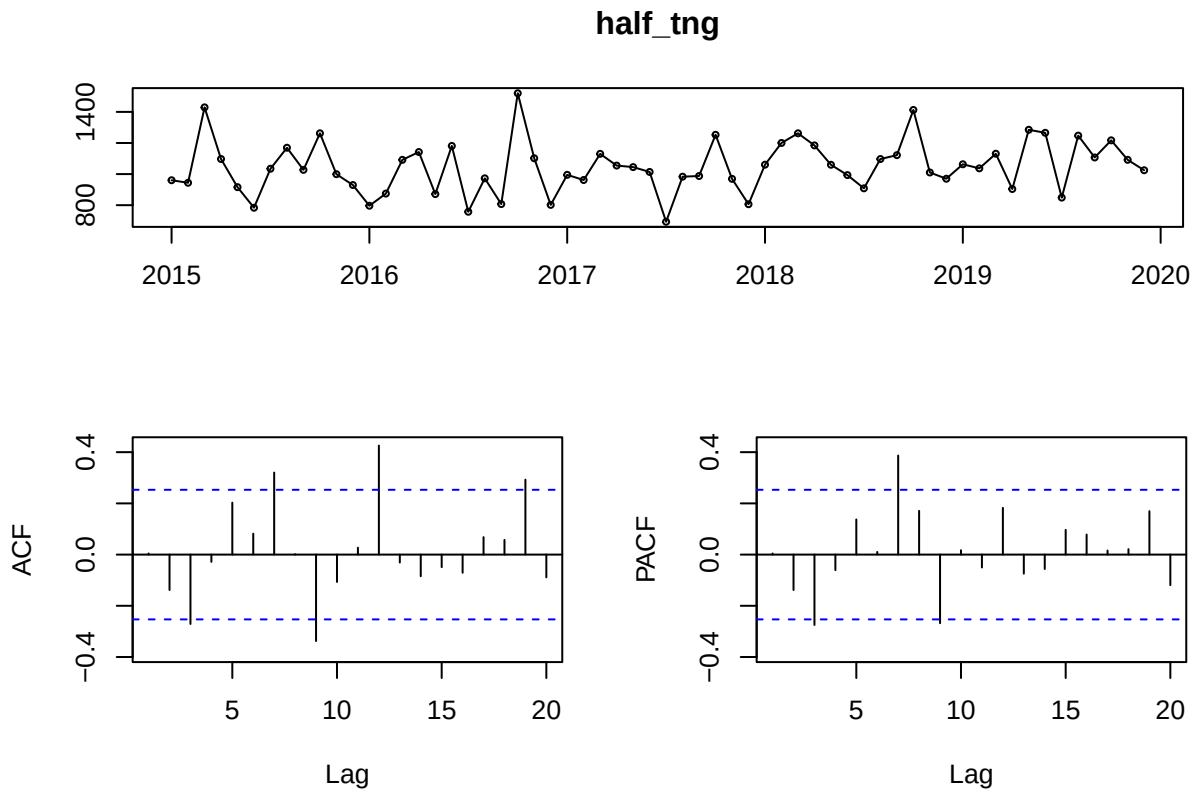
```
## [1] 0
```

1 Diff needed for seasonal data. Our data has shown to be seasonal.

```
TngDiff = ndiffs(half_tng)
TngDiff
```

```
## [1] 0
```

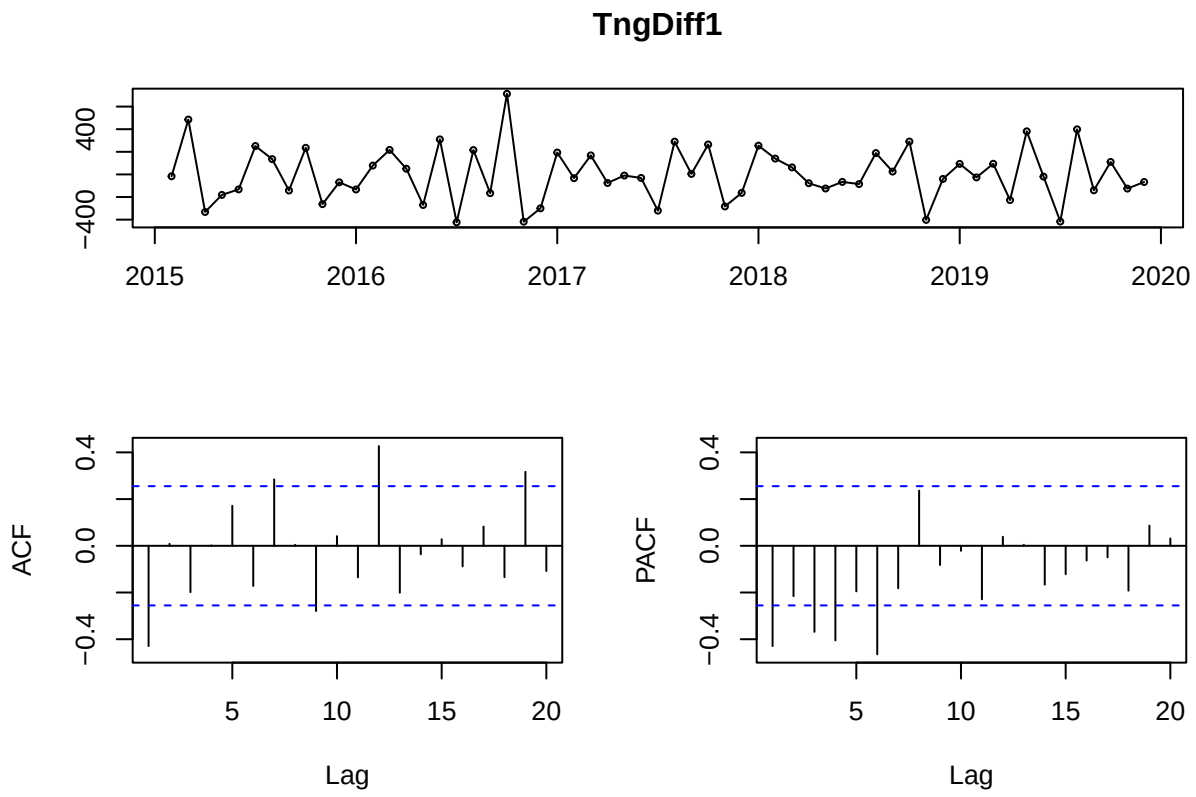
```
#tsdisplay plots ACF,PACF and timeseries plot together. How cool!
tsdisplay(half_tng)
```



See some auto-correlation at 3, 7, 9, and 12 months. See some PACF at 3,7,9 months.

```
# take first order differences as indicated by ndiffs function
#diff function can help you do this
TngDiff1 <- diff(half_tng, differences=1)
```

```
tsdisplay(TngDiff1)
```

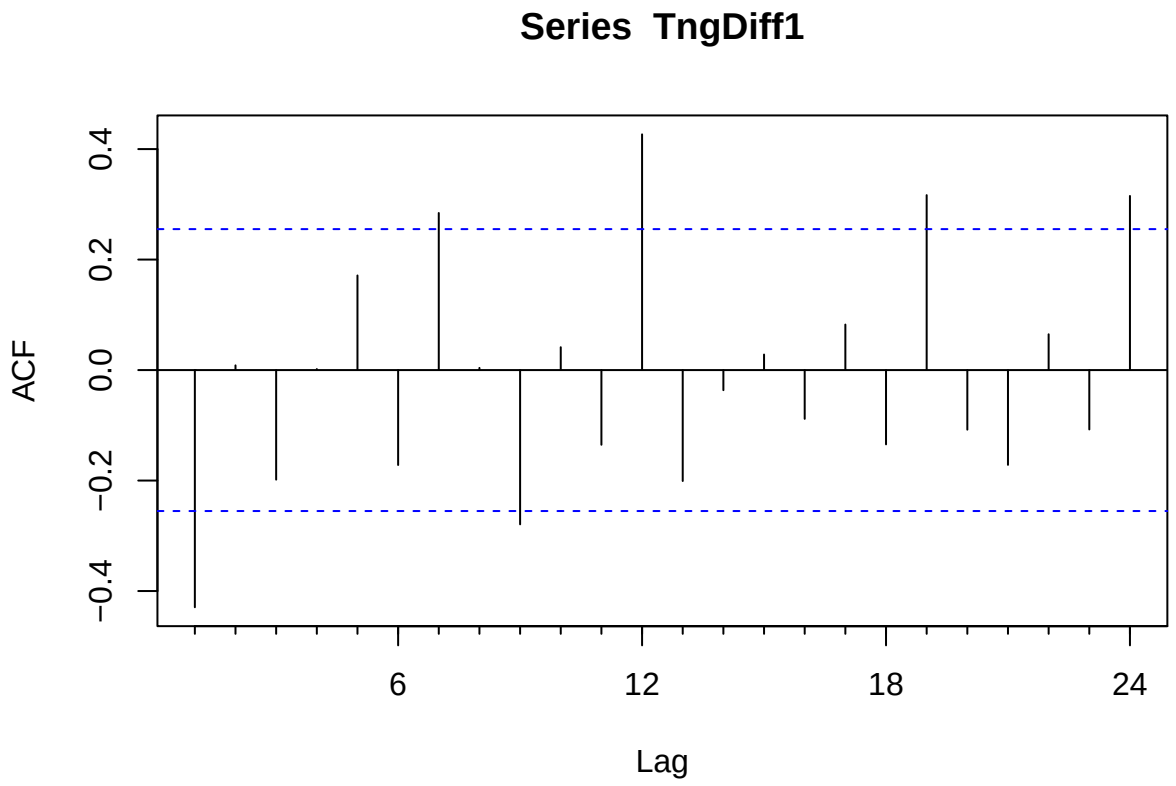


We have increased frequency of correlation

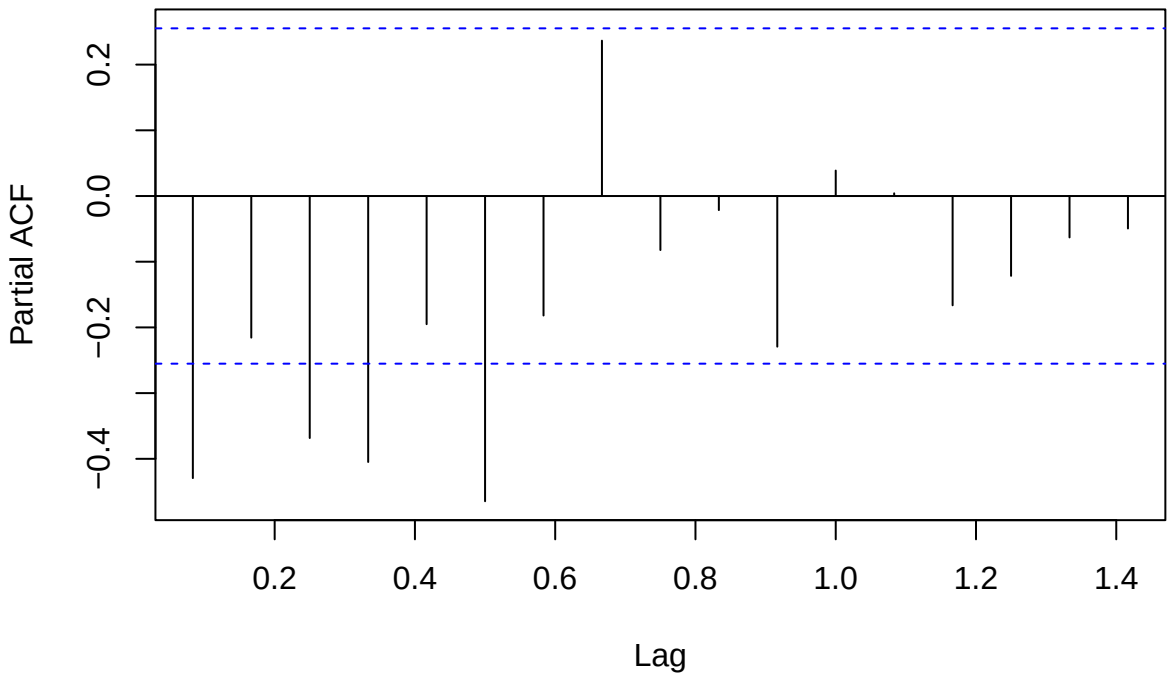
```
ndiffs(TngDiff1)
```

```
## [1] 0
```

```
Acf(TngDiff1)
```




```
pacf(TngDiff1)
```




```
Acf(TngDiff1, lag.max=20,plot=FALSE)
```

```
##
## Autocorrelations of series 'TngDiff1', by lag
##
##      0      1      2      3      4      5      6      7      8      9     10
##  1.000 -0.429  0.009 -0.198  0.002  0.171 -0.172  0.284  0.004 -0.279  0.041
##     11     12     13     14     15     16     17     18     19     20
## -0.135  0.427 -0.201 -0.036  0.028 -0.088  0.082 -0.135  0.317 -0.108
```

```
auto.arima(half_tng)
```

```
## Series: half_tng
## ARIMA(0,0,0)(1,0,0)[12] with non-zero mean
##
## Coefficients:
##         sar1      mean
##       0.5000  1055.231
## s.e.  0.1131    31.467
##
## sigma^2 estimated as 21383:  log likelihood=-384.96
## AIC=775.91   AICc=776.34   BIC=782.2
```

Look at all of the possible models

```
# or save the model. BIC and AIC is also given as values.
auto_fit <- auto.arima(half_tng, trace=TRUE, stepwise=FALSE)
```

```
##
##  ARIMA(0,0,0)            with zero mean     : 1008.383
##  ARIMA(0,0,0)            with non-zero mean : 788.9207
##  ARIMA(0,0,0)(0,0,1)[12] with zero mean     : Inf
##  ARIMA(0,0,0)(0,0,1)[12] with non-zero mean : 779.2159
##  ARIMA(0,0,0)(1,0,0)[12] with zero mean     : Inf
##  ARIMA(0,0,0)(1,0,0)[12] with non-zero mean : 776.3417
##  ARIMA(0,0,0)(1,0,1)[12] with zero mean     : Inf
##  ARIMA(0,0,0)(1,0,1)[12] with non-zero mean : 777.4751
##  ARIMA(0,0,1)            with zero mean     : 952.3381
##  ARIMA(0,0,1)            with non-zero mean : 791.1371
##  ARIMA(0,0,1)(0,0,1)[12] with zero mean     : Inf
##  ARIMA(0,0,1)(0,0,1)[12] with non-zero mean : 781.4647
##  ARIMA(0,0,1)(1,0,0)[12] with zero mean     : Inf
##  ARIMA(0,0,1)(1,0,0)[12] with non-zero mean : 778.6208
##  ARIMA(0,0,1)(1,0,1)[12] with zero mean     : Inf
##  ARIMA(0,0,1)(1,0,1)[12] with non-zero mean : 779.7131
##  ARIMA(0,0,2)            with zero mean     : 907.2159
##  ARIMA(0,0,2)            with non-zero mean : 791.9724
##  ARIMA(0,0,2)(0,0,1)[12] with zero mean     : Inf
##  ARIMA(0,0,2)(0,0,1)[12] with non-zero mean : 782.9171
##  ARIMA(0,0,2)(1,0,0)[12] with zero mean     : Inf
##  ARIMA(0,0,2)(1,0,0)[12] with non-zero mean : 780.6271
##  ARIMA(0,0,2)(1,0,1)[12] with zero mean     : Inf
##  ARIMA(0,0,2)(1,0,1)[12] with non-zero mean : 781.8547
##  ARIMA(0,0,3)            with zero mean     : 894.2347
##  ARIMA(0,0,3)            with non-zero mean : 790.644
##  ARIMA(0,0,3)(0,0,1)[12] with zero mean     : Inf
##  ARIMA(0,0,3)(0,0,1)[12] with non-zero mean : 782.7132
##  ARIMA(0,0,3)(1,0,0)[12] with zero mean     : Inf
##  ARIMA(0,0,3)(1,0,0)[12] with non-zero mean : 780.7328
##  ARIMA(0,0,3)(1,0,1)[12] with zero mean     : Inf
##  ARIMA(0,0,3)(1,0,1)[12] with non-zero mean : 782.8134
##  ARIMA(0,0,4)            with zero mean     : Inf
##  ARIMA(0,0,4)            with non-zero mean : 790.0069
##  ARIMA(0,0,4)(0,0,1)[12] with zero mean     : 858.5537
##  ARIMA(0,0,4)(0,0,1)[12] with non-zero mean : 784.3248
##  ARIMA(0,0,4)(1,0,0)[12] with zero mean     : Inf
##  ARIMA(0,0,4)(1,0,0)[12] with non-zero mean : 782.8422
##  ARIMA(0,0,5)            with zero mean     : Inf
##  ARIMA(0,0,5)            with non-zero mean : 787.4745
##  ARIMA(1,0,0)            with zero mean     : 833.2195
##  ARIMA(1,0,0)            with non-zero mean : 791.1376
##  ARIMA(1,0,0)(0,0,1)[12] with zero mean     : 824.8455
##  ARIMA(1,0,0)(0,0,1)[12] with non-zero mean : 781.4753
##  ARIMA(1,0,0)(1,0,0)[12] with zero mean     : Inf
##  ARIMA(1,0,0)(1,0,0)[12] with non-zero mean : 778.6238
```

```
##  ARIMA(1,0,0)(1,0,1)[12] with zero mean     : Inf
##  ARIMA(1,0,0)(1,0,1)[12] with non-zero mean : 779.7377
##  ARIMA(1,0,1)            with zero mean     : Inf
##  ARIMA(1,0,1)            with non-zero mean : 792.5145
##  ARIMA(1,0,1)(0,0,1)[12] with zero mean     : Inf
##  ARIMA(1,0,1)(0,0,1)[12] with non-zero mean : 782.803
##  ARIMA(1,0,1)(1,0,0)[12] with zero mean     : Inf
##  ARIMA(1,0,1)(1,0,0)[12] with non-zero mean : 780.7184
##  ARIMA(1,0,1)(1,0,1)[12] with zero mean     : Inf
##  ARIMA(1,0,1)(1,0,1)[12] with non-zero mean : 782.291
##  ARIMA(1,0,2)            with zero mean     : Inf
##  ARIMA(1,0,2)            with non-zero mean : 793.6809
##  ARIMA(1,0,2)(0,0,1)[12] with zero mean     : Inf
##  ARIMA(1,0,2)(0,0,1)[12] with non-zero mean : 784.6528
##  ARIMA(1,0,2)(1,0,0)[12] with zero mean     : Inf
##  ARIMA(1,0,2)(1,0,0)[12] with non-zero mean : 782.5433
##  ARIMA(1,0,2)(1,0,1)[12] with zero mean     : Inf
##  ARIMA(1,0,2)(1,0,1)[12] with non-zero mean : 784.137
##  ARIMA(1,0,3) with zero mean     : Inf
##  ARIMA(1,0,3)            with non-zero mean : 792.4993
##  ARIMA(1,0,3)(0,0,1)[12] with zero mean     : Inf
##  ARIMA(1,0,3)(0,0,1)[12] with non-zero mean : 785.0136
##  ARIMA(1,0,3)(1,0,0)[12] with zero mean     : Inf
##  ARIMA(1,0,3)(1,0,0)[12] with non-zero mean : 783.1365
##  ARIMA(1,0,4)            with zero mean     : Inf
##  ARIMA(1,0,4)            with non-zero mean : 788.8028
##  ARIMA(2,0,0)            with zero mean     : 824.0601
##  ARIMA(2,0,0)            with non-zero mean : 792.3004
##  ARIMA(2,0,0)(0,0,1)[12] with zero mean     : 814.1226
##  ARIMA(2,0,0)(0,0,1)[12] with non-zero mean : 783.2925
##  ARIMA(2,0,0)(1,0,0)[12] with zero mean     : Inf
##  ARIMA(2,0,0)(1,0,0)[12] with non-zero mean : 780.6224
##  ARIMA(2,0,0)(1,0,1)[12] with zero mean     : 809.0495
##  ARIMA(2,0,0)(1,0,1)[12] with non-zero mean : 781.6862
##  ARIMA(2,0,1) with zero mean     : Inf
##  ARIMA(2,0,1)            with non-zero mean : 792.7482
##  ARIMA(2,0,1)(0,0,1)[12] with zero mean     : Inf
##  ARIMA(2,0,1)(0,0,1)[12] with non-zero mean : Inf
##  ARIMA(2,0,1)(1,0,0)[12] with zero mean     : Inf
##  ARIMA(2,0,1)(1,0,0)[12] with non-zero mean : 782.2036
##  ARIMA(2,0,1)(1,0,1)[12] with zero mean     : Inf
##  ARIMA(2,0,1)(1,0,1)[12] with non-zero mean : 783.8318
##  ARIMA(2,0,2) with zero mean     : Inf
##  ARIMA(2,0,2)            with non-zero mean : Inf
##  ARIMA(2,0,2)(0,0,1)[12] with zero mean     : Inf
##  ARIMA(2,0,2)(0,0,1)[12] with non-zero mean : 782.2635
##  ARIMA(2,0,2)(1,0,0)[12] with zero mean     : Inf
##  ARIMA(2,0,2)(1,0,0)[12] with non-zero mean : Inf
##  ARIMA(2,0,3) with zero mean     : Inf
##  ARIMA(2,0,3)            with non-zero mean : 792.2954
##  ARIMA(3,0,0)            with zero mean     : Inf
##  ARIMA(3,0,0)            with non-zero mean : 789.6905
##  ARIMA(3,0,0)(0,0,1)[12] with zero mean     : 813.255
##  ARIMA(3,0,0)(0,0,1)[12] with non-zero mean : 782.134
```

```
##  ARIMA(3,0,0)(1,0,0)[12] with zero mean     : 808.3923
##  ARIMA(3,0,0)(1,0,0)[12] with non-zero mean : 780.3886
##  ARIMA(3,0,0)(1,0,1)[12] with zero mean     : Inf
##  ARIMA(3,0,0)(1,0,1)[12] with non-zero mean : 782.5533
##  ARIMA(3,0,1)            with zero mean     : Inf
##  ARIMA(3,0,1)            with non-zero mean : 792.0619
##  ARIMA(3,0,1)(0,0,1)[12] with zero mean     : Inf
##  ARIMA(3,0,1)(0,0,1)[12] with non-zero mean : 784.5655
##  ARIMA(3,0,1)(1,0,0)[12] with zero mean     : Inf
##  ARIMA(3,0,1)(1,0,0)[12] with non-zero mean : 782.9557
##  ARIMA(3,0,2)            with zero mean     : Inf
##  ARIMA(3,0,2)            with non-zero mean : Inf
##  ARIMA(4,0,0)            with zero mean     : Inf
##  ARIMA(4,0,0)            with non-zero mean : 791.9508
##  ARIMA(4,0,0)(0,0,1)[12] with zero mean     : Inf
##  ARIMA(4,0,0)(0,0,1)[12] with non-zero mean : 784.4907
##  ARIMA(4,0,0)(1,0,0)[12] with zero mean     : Inf
##  ARIMA(4,0,0)(1,0,0)[12] with non-zero mean : 782.9545
##  ARIMA(4,0,1)            with zero mean     : Inf
##  ARIMA(4,0,1)            with non-zero mean : 793.4599
##  ARIMA(5,0,0)            with zero mean     : Inf
##  ARIMA(5,0,0)            with non-zero mean : 793.1683
##
##
##
##  Best model: ARIMA(0,0,0)(1,0,0)[12] with non-zero mean
```

```
auto_fit
```

```
## Series: half_tng
## ARIMA(0,0,0)(1,0,0)[12] with non-zero mean
##
## Coefficients:
##         sar1      mean
##       0.5000  1055.231
## s.e.  0.1131    31.467
##
## sigma^2 estimated as 21383:  log likelihood=-384.96
## AIC=775.91   AICc=776.34   BIC=782.2
```

Gives us all options for ARIMA, with the lowes AIC.

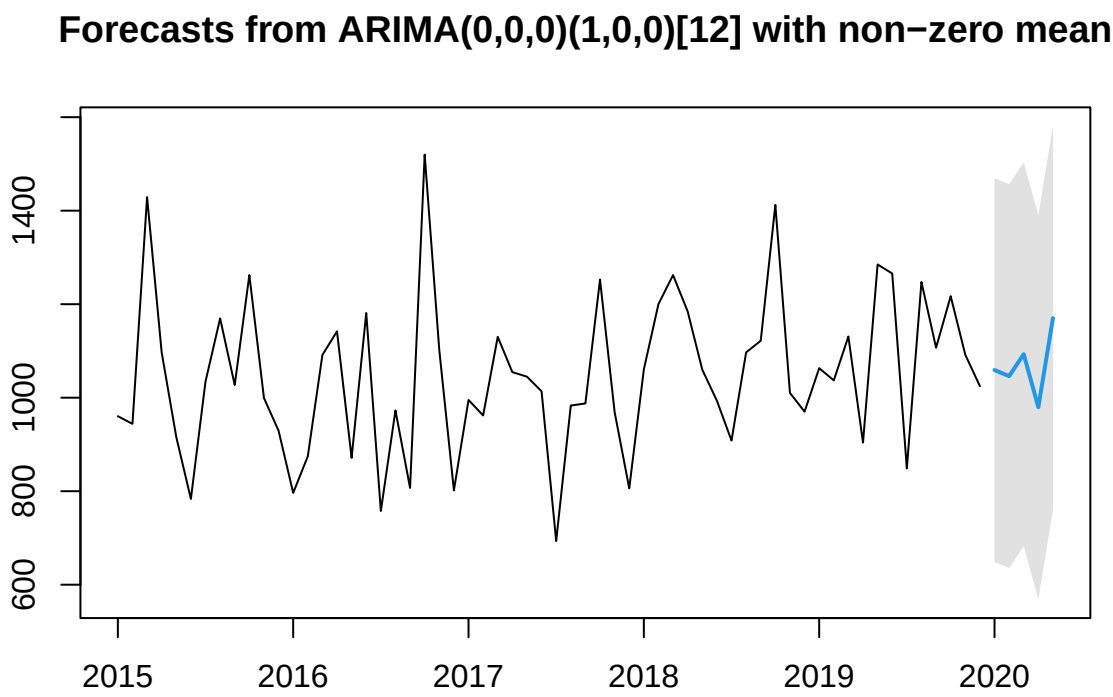
Best Model has an ACF of 0, Diff of 1, and PACF of 0. Or ACF of 1, Diff of 1, and PACF of 0

```
forecast(auto_fit,h=5,level=c(99.5))
```

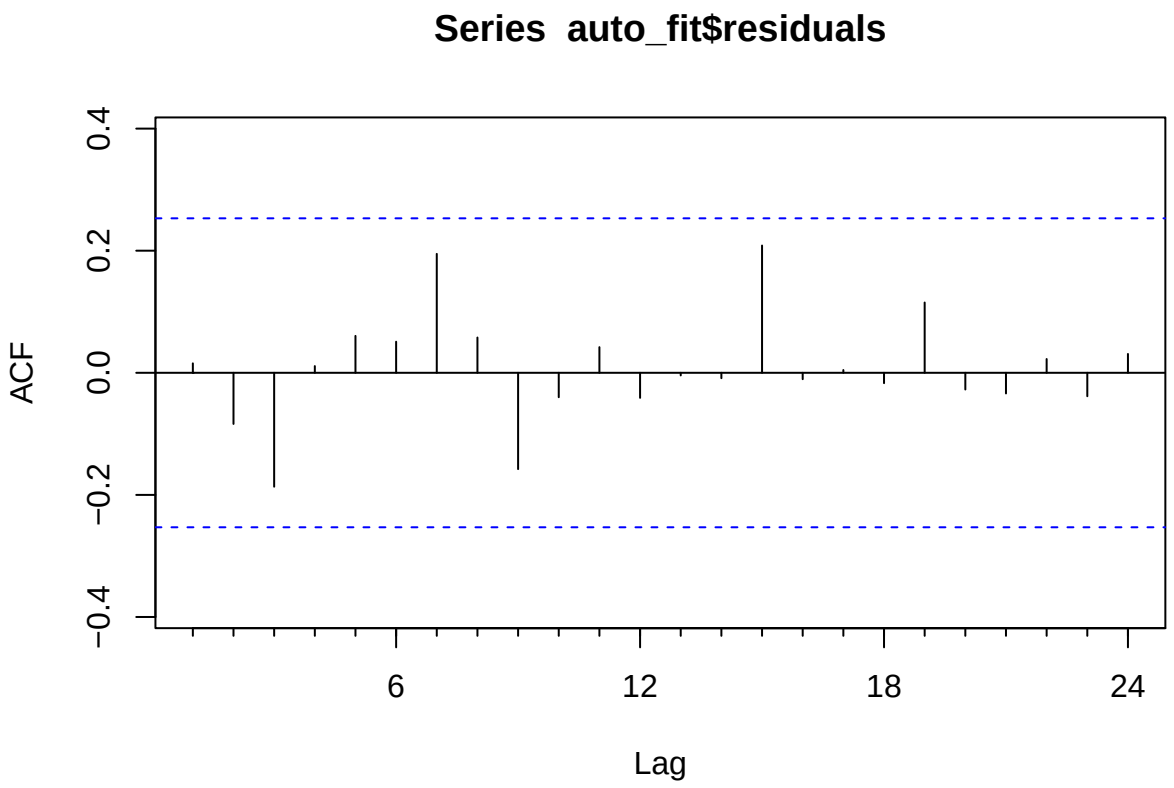
```
##          Point Forecast  Lo 99.5  Hi 99.5
## Jan 2020      1059.1806 648.7058 1469.655
## Feb 2020      1046.0905 635.6157 1456.565
## Mar 2020      1093.0509 682.5761 1503.526
## Apr 2020       979.5999 569.1251 1390.075
## May 2020      1170.0915 759.6168 1580.566
```

Shows we have a 99.5% certainty that Jan 2020 data will fall between 648.7 and 1469.7.

```
plot(forecast(auto_fit,h=5,level=c(99.5)))
```

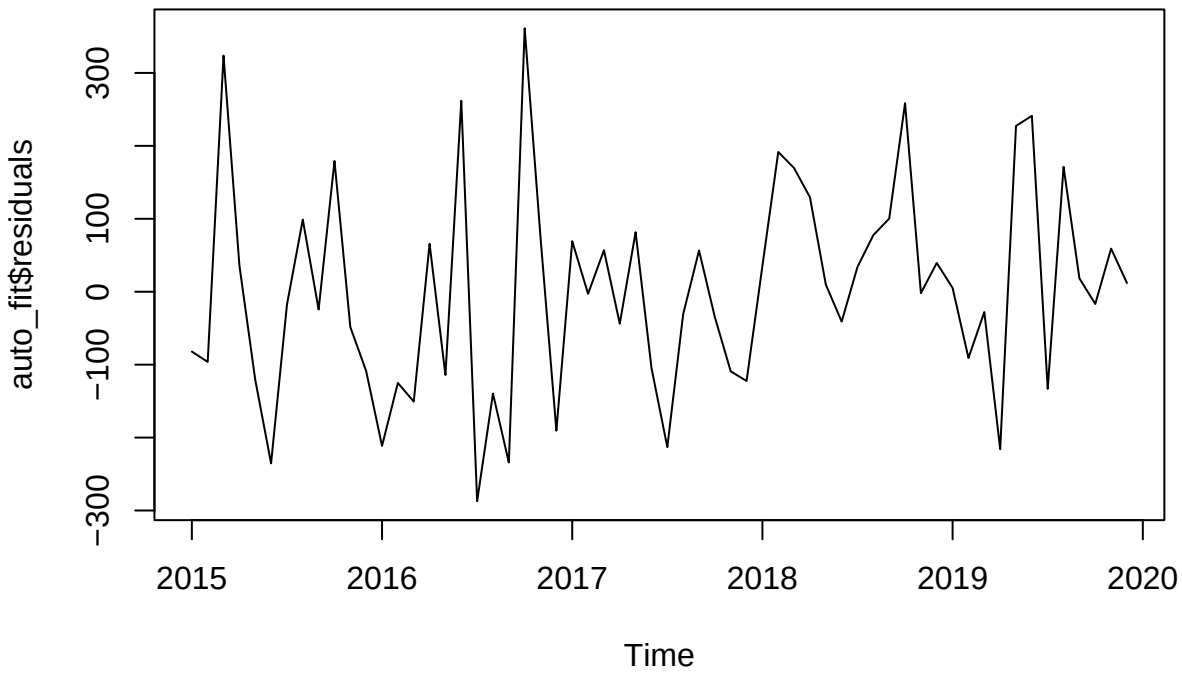



```
Acf(auto_fit$residuals)
```

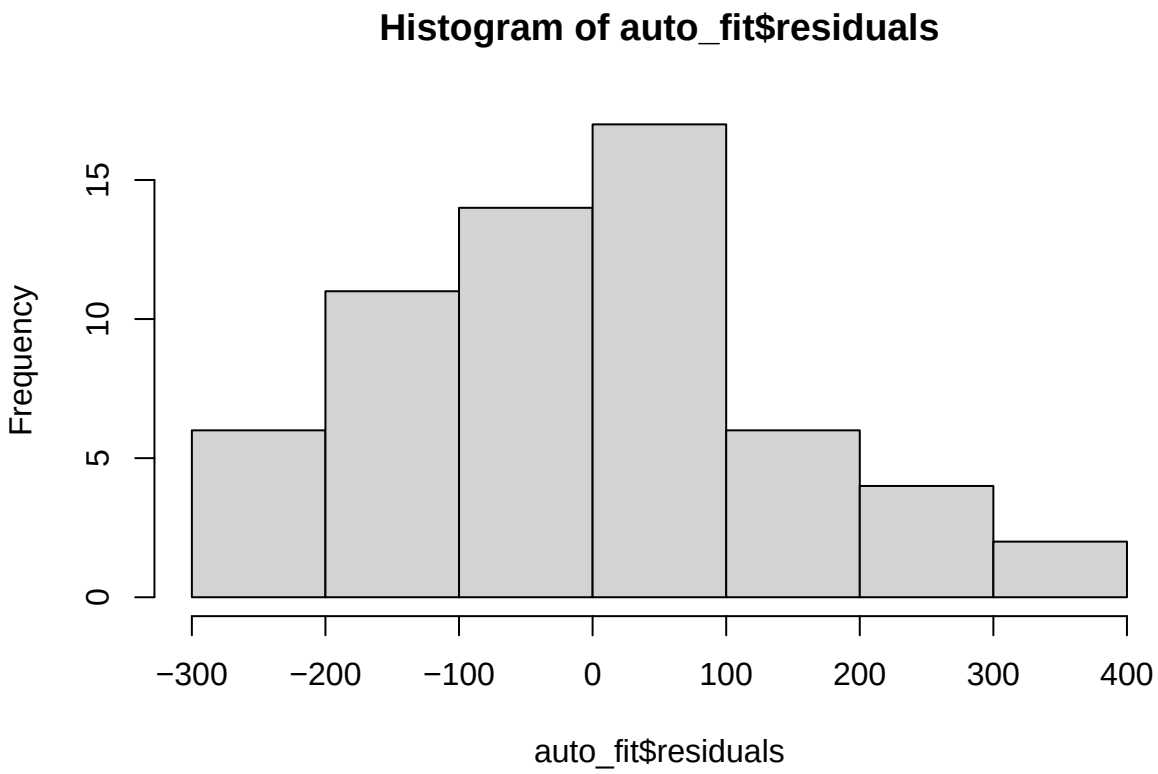


No significance in the residuals. Looks Good!

```
plot(auto_fit$residuals)
```

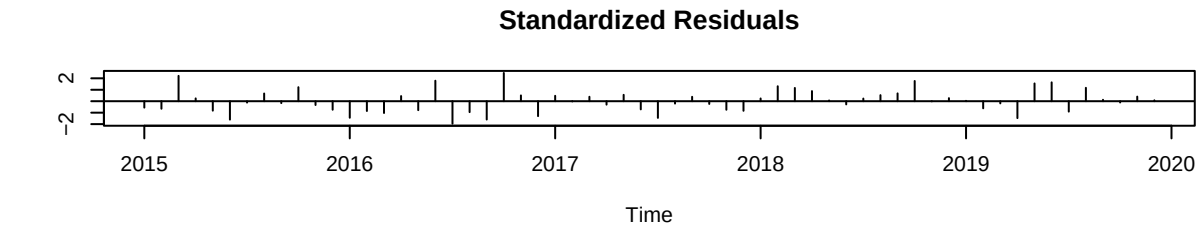


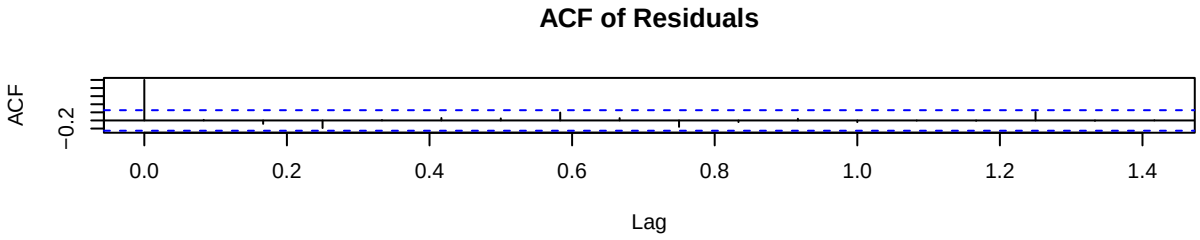

```
hist(auto_fit$residuals)
```

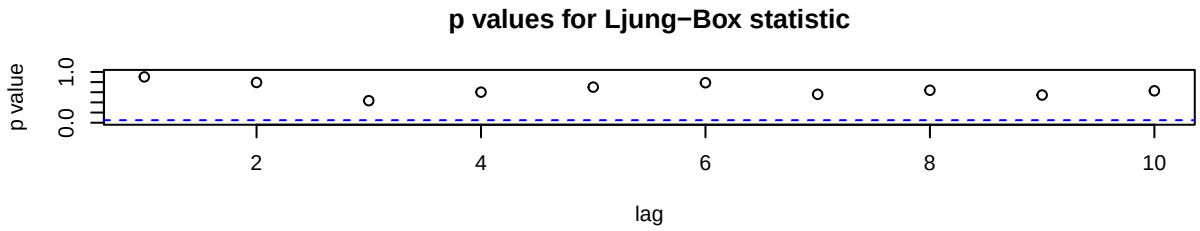


Skewed Lefts so not completely normal.

```
tsdiag(auto_fit)
```


Standardized Residuals
2
−2
2015
2016
2017
2018
2019
2020
Time


ACF of Residuals
ACF
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
Lag


p values for Ljung−Box statistic
p value
0.0
1.0
2
4
6
8
10
lag