Time Series

2024-12-21

Contents


```
library(lubridate) # year, month
```

```
## Warning: package 'lubridate' was built under R version 4.4.2
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```
library(dplyr) # %>%
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(forecast) # auto.arima
```

```
## Warning: package 'forecast' was built under R version 4.4.2
```

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

```
data = read.csv('Ch6_ridership_data_2011-2012.csv')
str(data)
```

```
## 'data.frame':    17379 obs. of  2 variables:
##  $ datetime: chr  "2011-01-01 00:00:00" "2011-01-01 01:00:00" "2011-01-01 02:00:00" "2011-01-01 03:00:00" ...
##  $ count   : int  16 40 32 13 1 1 2 3 8 14 ...
```

We can see that the data is in hourly data frame and we want to convert it into monthly data frame.

```
monthly_ride = data %>%
  group_by(year = year(datetime), month = month(datetime)) %>%
  summarise(riders = sum(count))
```

```
## `summarise()` has grouped output by 'year'. You can override using the
## `.groups` argument.
```

```
table(monthly_ride$year, monthly_ride$month)
```

```
##
##        1 2 3 4 5 6 7 8 9 10 11 12
##   2011 1 1 1 1 1 1 1 1 1  1  1  1
##   2012 1 1 1 1 1 1 1 1 1  1  1  1
```

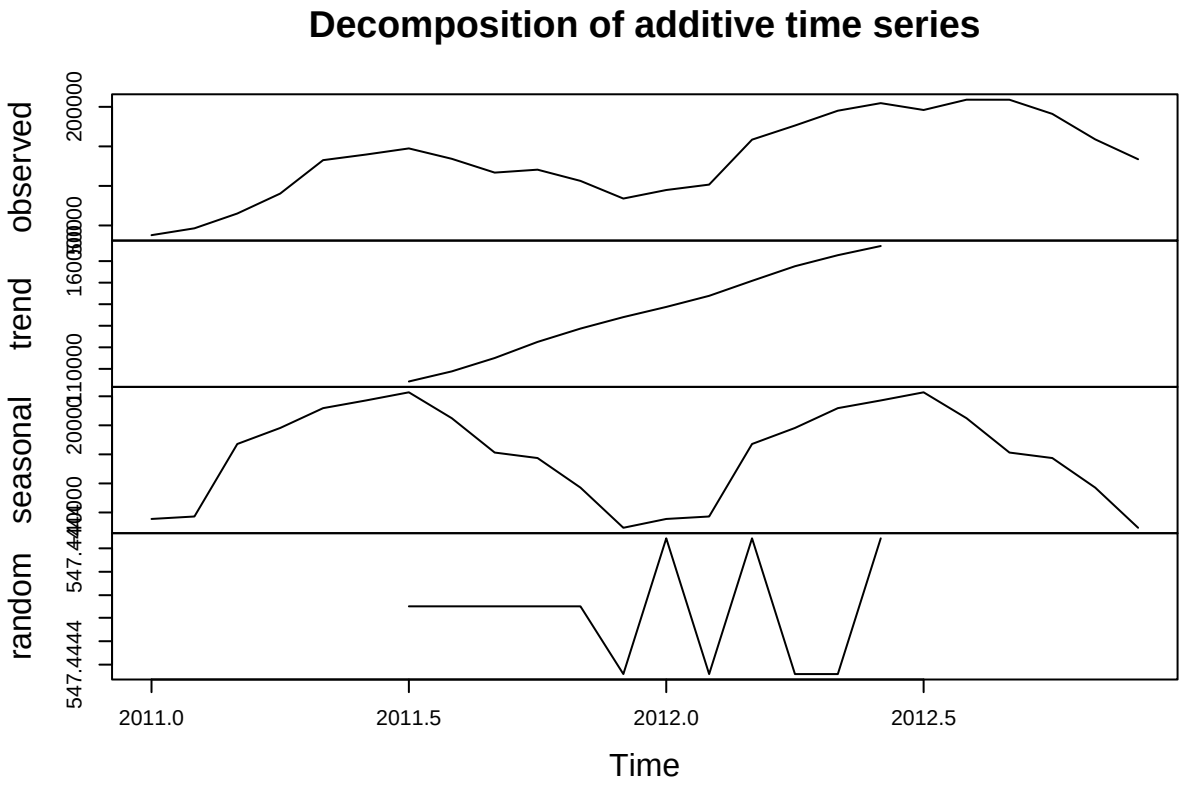
```
riders = monthly_ride[,3]
monthly = ts(riders, frequency = 12, start = c(2011,1))
class(monthly)
```

```
## [1] "ts"
```

```
monthly
```

```
##          Jan    Feb    Mar    Apr    May    Jun    Jul    Aug    Sep    Oct
## 2011  37727  46396  65109  90332 132580 139674 147426 134280 116825 120535
## 2012  94832 101668 158535 176349 195114 204683 196014 209024 208995 191108
##          Nov    Dec
## 2011 106361  84025
## 2012 158855 133735
```

```
plot(decompose(monthly))
```

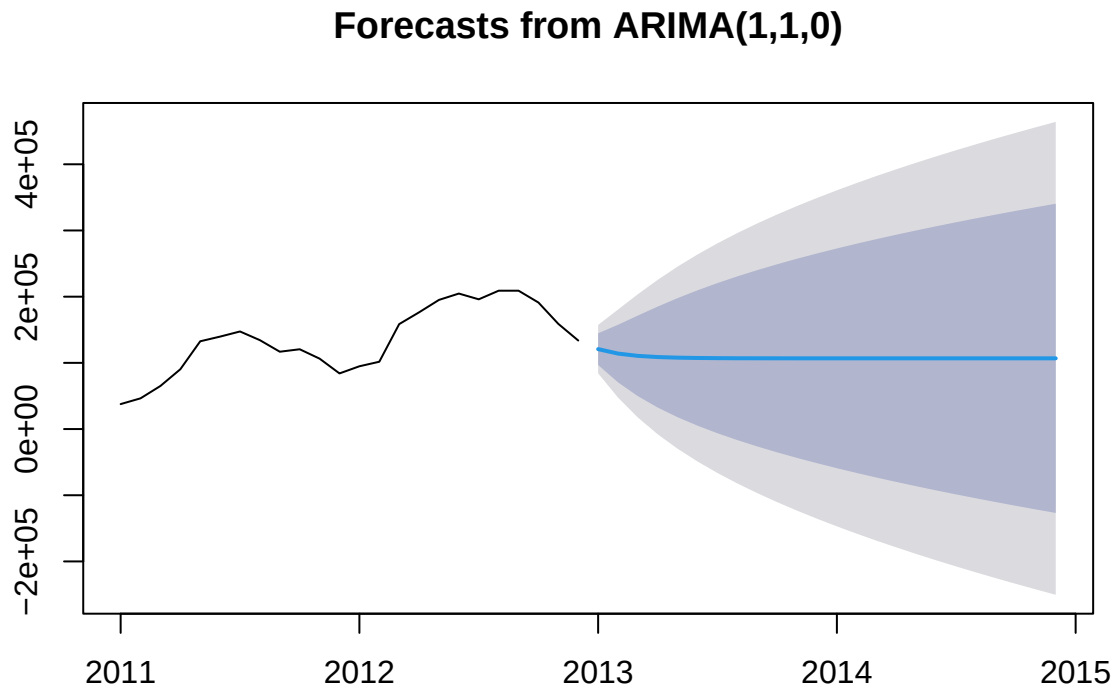



```
auto.arima(monthly)
```

```
## Series: monthly
## ARIMA(1,1,0)
##
## Coefficients:
##          ar1
##       0.5171
```

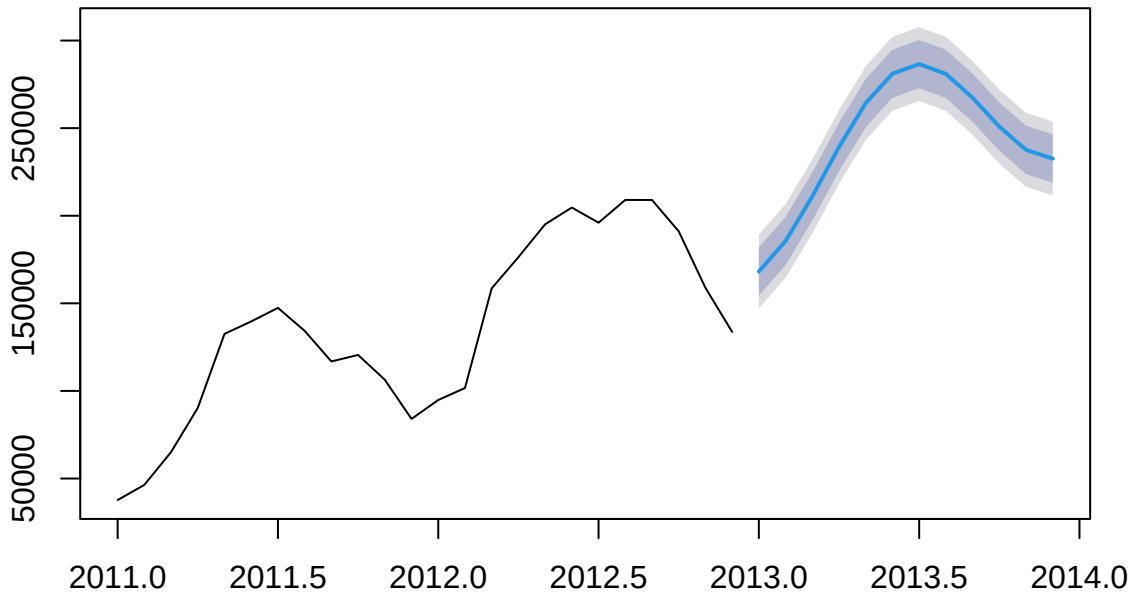
```
## s.e.  0.1772
##
## sigma^2 = 348592099:  log likelihood = -258.48
## AIC=520.96   AICc=521.56   BIC=523.23
```

```
yr_forecast = forecast(auto.arima(monthly))
plot(yr_forecast)
```



TBAT

```
year_forecast = forecast(tbats(monthly), h = 12)
plot(year_forecast)
```


Forecasts from TBATS(1, {0,0}, 1, {<12,1>})
250000
150000
50000
2011.0
2011.5
2012.0
2012.5
2013.0
2013.5
2014.0