Australian beer production forecasting

Manish Kumar

8/14/2018

Case Study

Analyze the historical beer production data and use time series forecasting techniques to forecast future beer production.

load packages

```
library(fpp2)
library(astsa)
library(DT)
library(dygraphs)
```

load data

```
beer <- read.csv("data/monthly-beer-production-australia.csv")
head(beer)
```

```
##     Month Monthly.beer.production.in.Australia
## 1 1956-01                                 93.2
## 2 1956-02                                 96.0
## 3 1956-03                                 95.2
## 4 1956-04                                 77.1
## 5 1956-05                                 70.9
## 6 1956-06                                 64.8
```

```
tail(beer)
```

```
##       Month Monthly.beer.production.in.Australia
## 471 1995-03                                 152
## 472 1995-04                                 127
## 473 1995-05                                 151
## 474 1995-06                                 130
## 475 1995-07                                 119
## 476 1995-08                                 153
```

```
summary(beer)
```

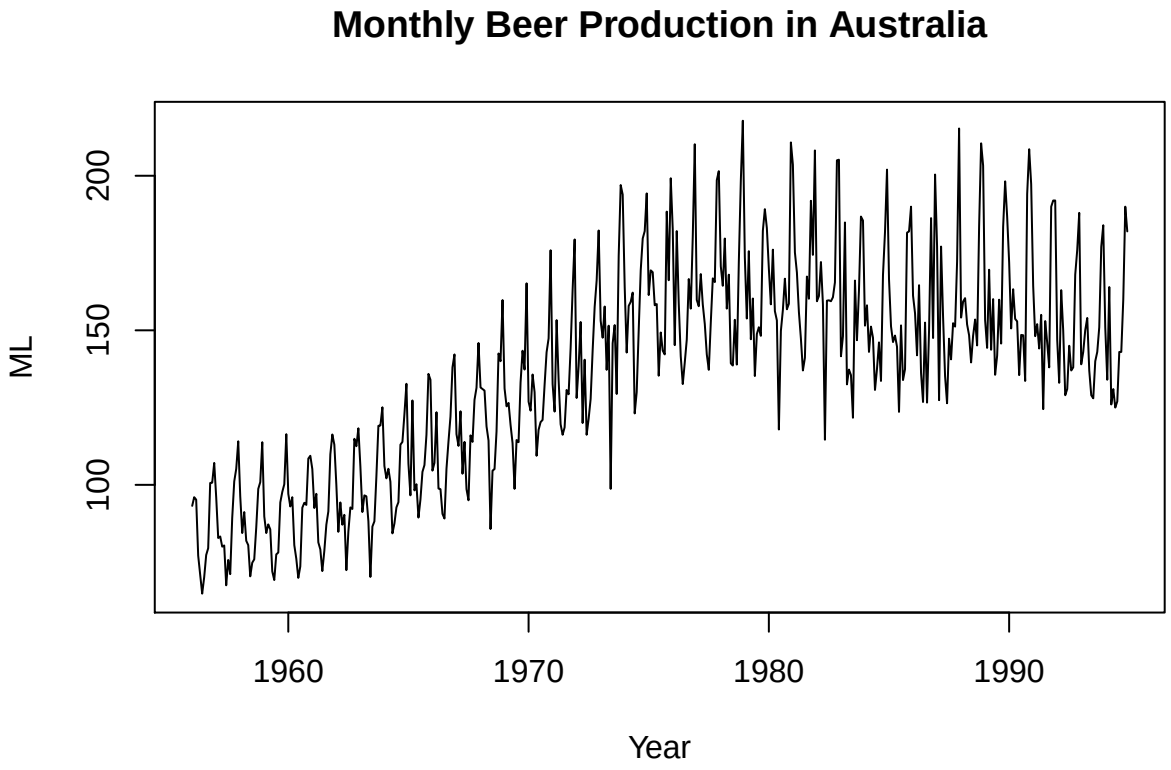
```
##      Month     Monthly.beer.production.in.Australia
##  1956-01:  1   Min.   : 64.8
##  1956-02:  1   1st Qu.:112.9
##  1956-03:  1   Median :139.2
##  1956-04:  1   Mean   :136.4
##  1956-05:  1   3rd Qu.:158.8
##  1956-06:  1   Max.   :217.8
##  (Other):470
```

```r
beer.ts <- ts(beer, frequency = 12, start = c(1956,1), end = c(1994,12))
```

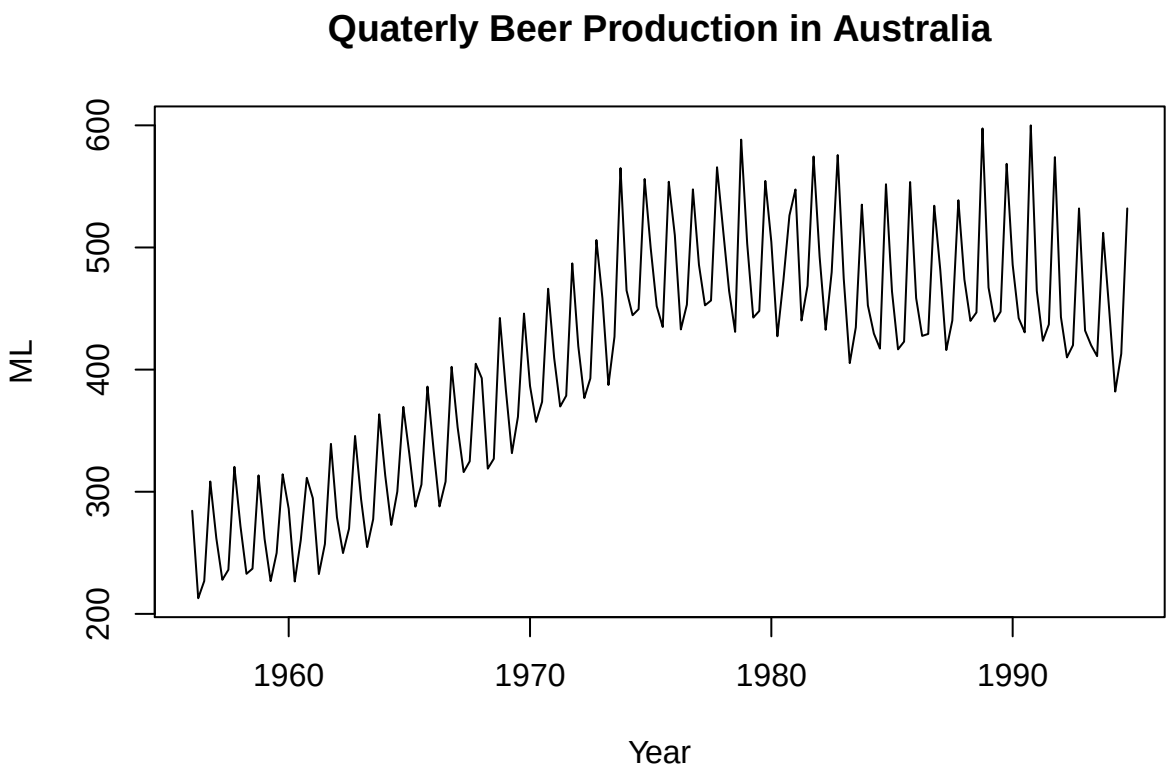
```r
head(beer.ts)
```

```
##          Month Monthly.beer.production.in.Australia
## Jan 1956     1                                 93.2
## Feb 1956     2                                 96.0
## Mar 1956     3                                 95.2
## Apr 1956     4                                 77.1
## May 1956     5                                 70.9
## Jun 1956     6                                 64.8
```

```r
plot.ts(beer.ts[,2], main = "Monthly Beer Production in Australia", xlab = "Year", ylab = "ML")
```

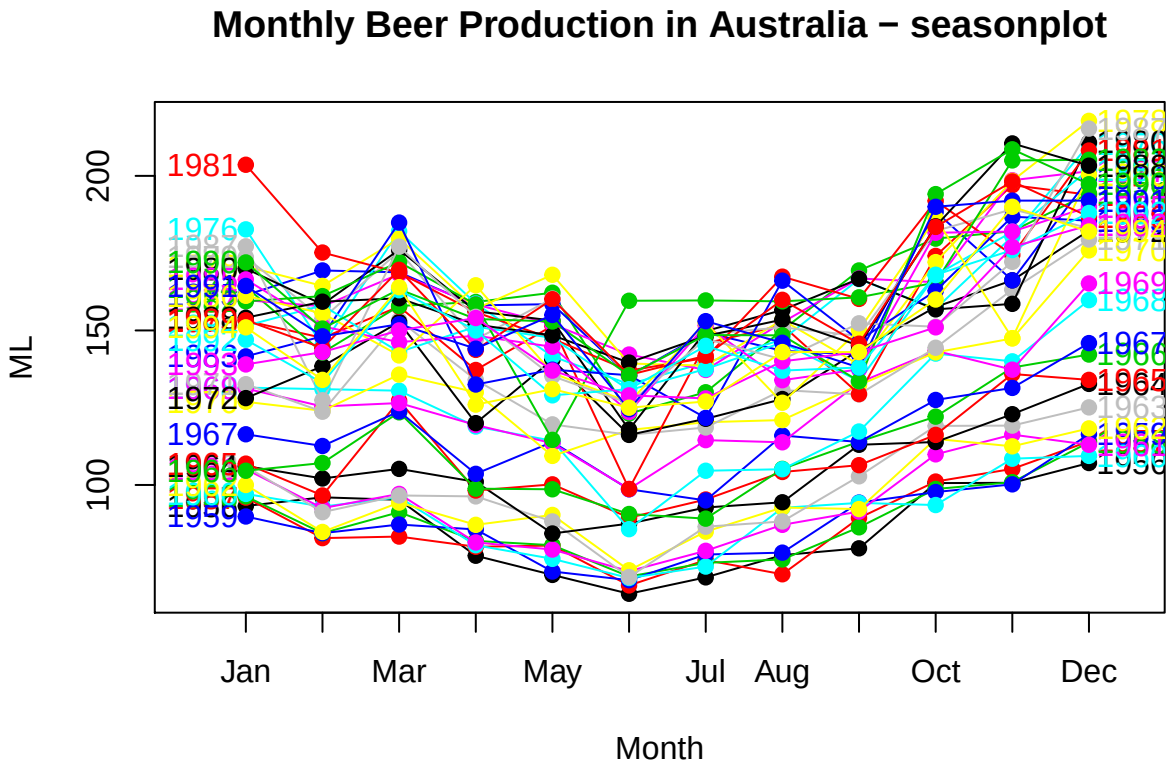



```r
# aggregared on quater level
beer.ts.qtr <- aggregate(beer.ts, nfrequency=4)
plot.ts(beer.ts.qtr[,2], main = "Quaterly Beer Production in Australia", xlab = "Year", ylab = "ML")
```

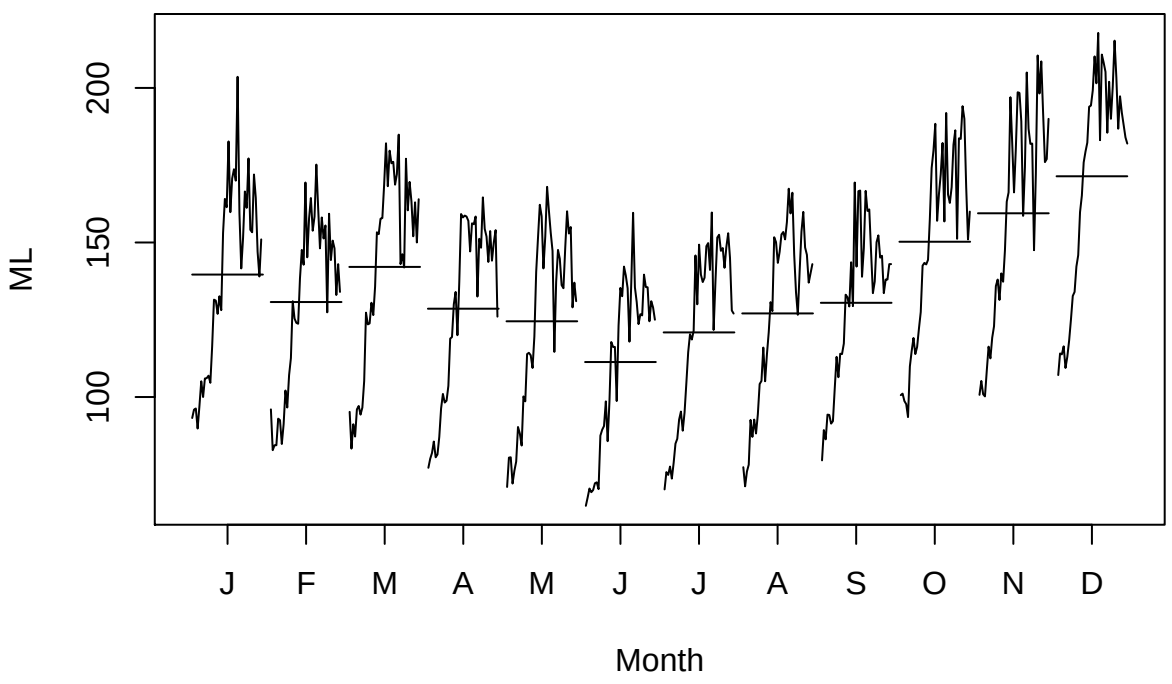



```
# aggregared on year level
beer.ts.yr <- aggregate(beer.ts, nfrequency=1)
plot.ts(beer.ts.yr[,2], main = "Yearly Beer Production in Australia", xlab = "Year", ylab = "ML")
```

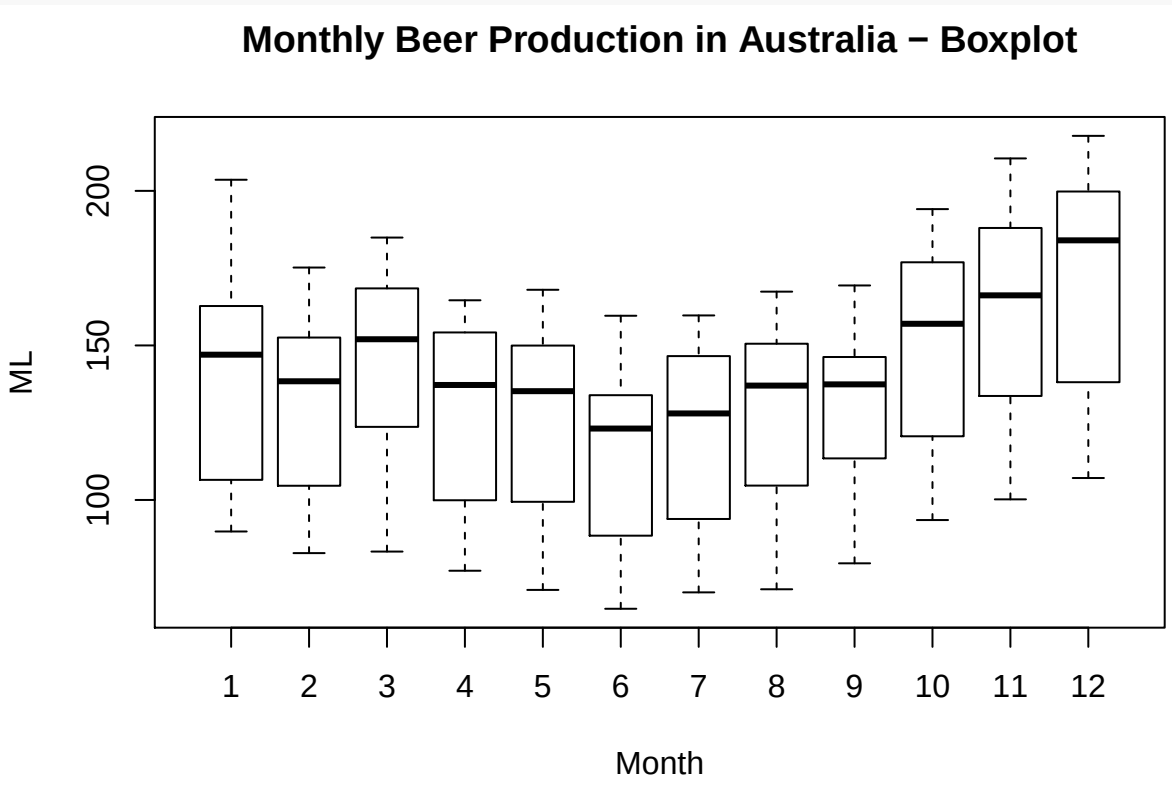
```
seasonplot(beer.ts[,2], year.labels = TRUE, year.labels.left=TRUE, col=1:40, pch=19, main = "Monthly Bee
```



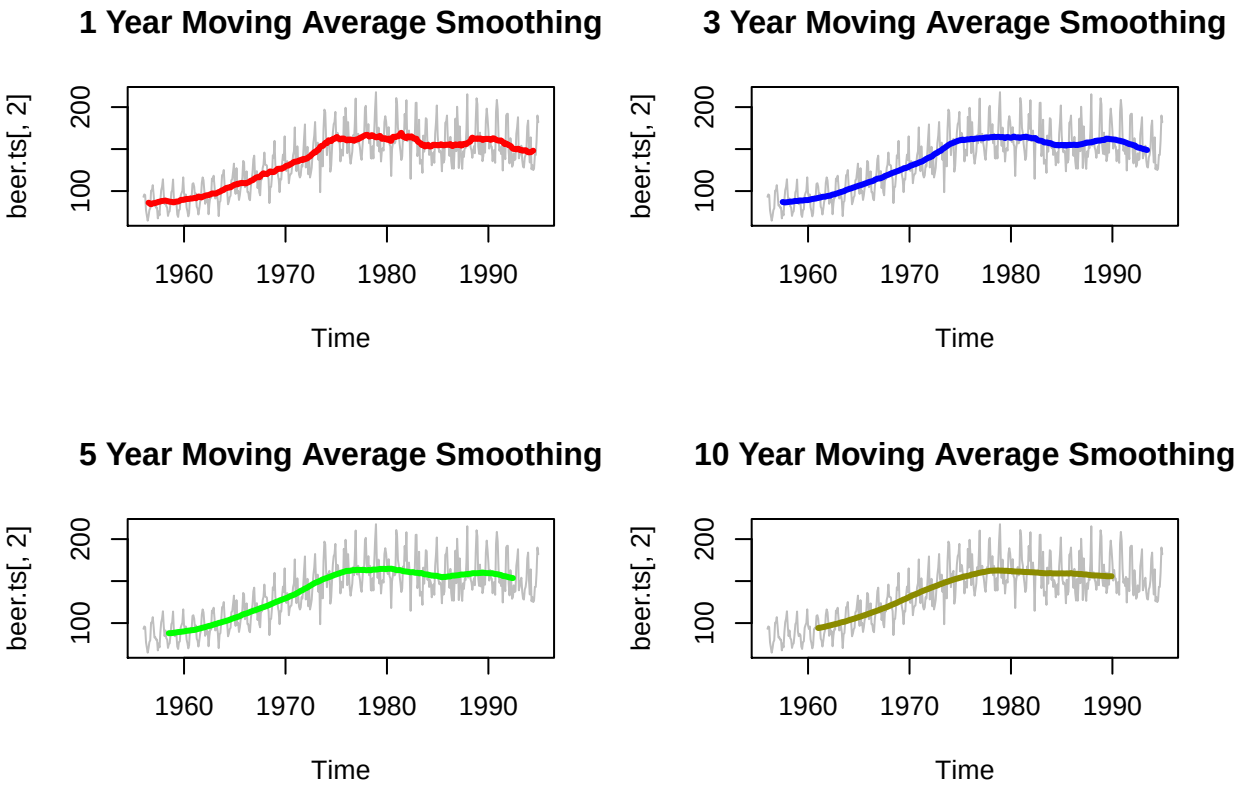

```
# monthly plot aggreated the data of all year for montly analysis
# each month plots show the variation for entire data in each month
monthplot(beer.ts[,2], main = "Monthly Beer Production in Australia - monthplot", xlab = "Month", ylab =
```



```
boxplot(beer.ts[,2] ~ cycle(beer.ts[,2]), xlab = "Month", ylab = "ML", main = "Monthly Beer Production
```




```
# moving average is useful when we need to analyse trend for the underlying data
# here, we see the moving average for 1 year, 3 year 5 year and 10 year
# Note : If there is not trend, average is good enough for the analysis
par(mfrow = c(2,2))
plot(beer.ts[,2], col="gray", main = "1 Year Moving Average Smoothing")
lines(ma(beer.ts[,2], order = 12), col = "red", lwd=3)
plot(beer.ts[,2], col="gray", main = "3 Year Moving Average Smoothing")
lines(ma(beer.ts[,2], order = 36), col = "blue", lwd=3)
plot(beer.ts[,2], col="gray", main = "5 Year Moving Average Smoothing")
lines(ma(beer.ts[,2], order = 60), col = "green", lwd=3)
plot(beer.ts[,2], col="gray", main = "10 Year Moving Average Smoothing")
lines(ma(beer.ts[,2], order = 120), col = "yellow4", lwd=3)
```


1 Year Moving Average Smoothing
3 Year Moving Average Smoothing
5 Year Moving Average Smoothing
10 Year Moving Average Smoothing
beer.ts[, 2]
100
200
1960
1970
1980
1990
Time