Australian beer

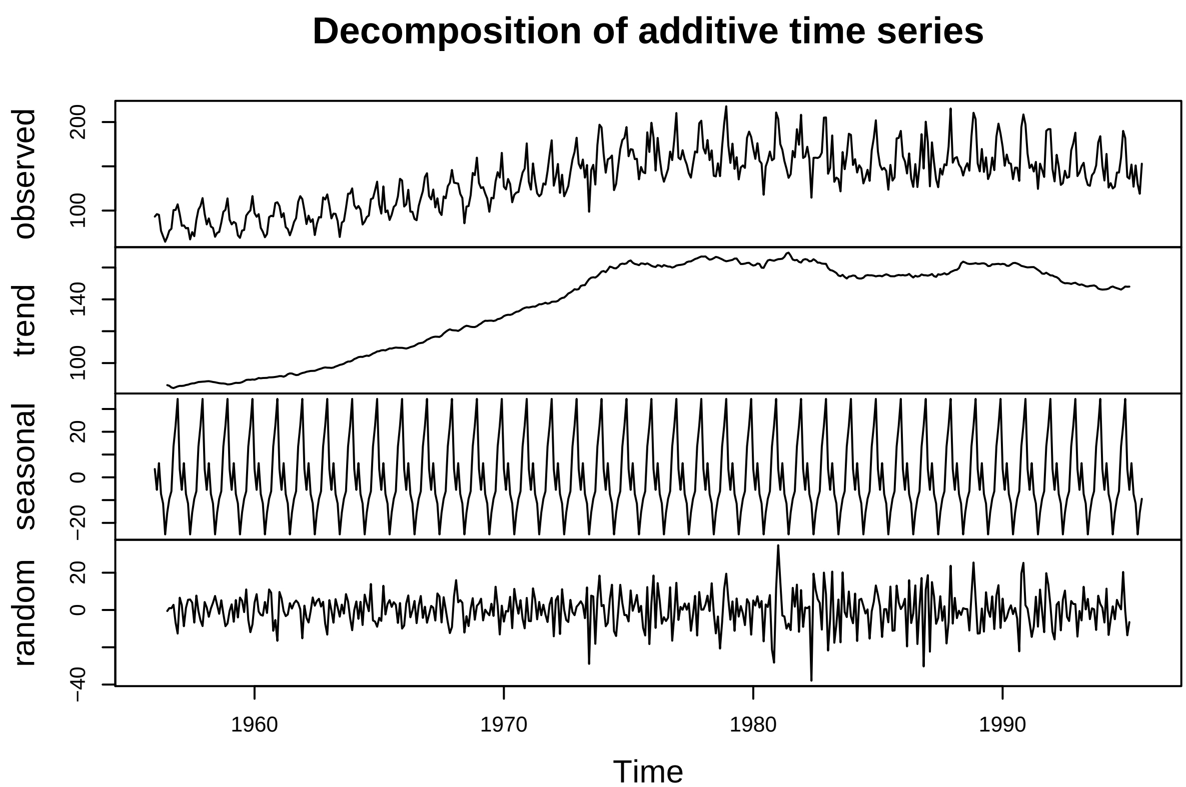
Claudius Taylor

10/5/2018

```
data2 = ("/Users/taylormade/Desktop/Time Series/Australian Beer.txt")
beer = read.table(file = data2, stringsAsFactors = FALSE, sep = ",", header = TRUE)
class(beer)
```

```
## [1] "data.frame"
```

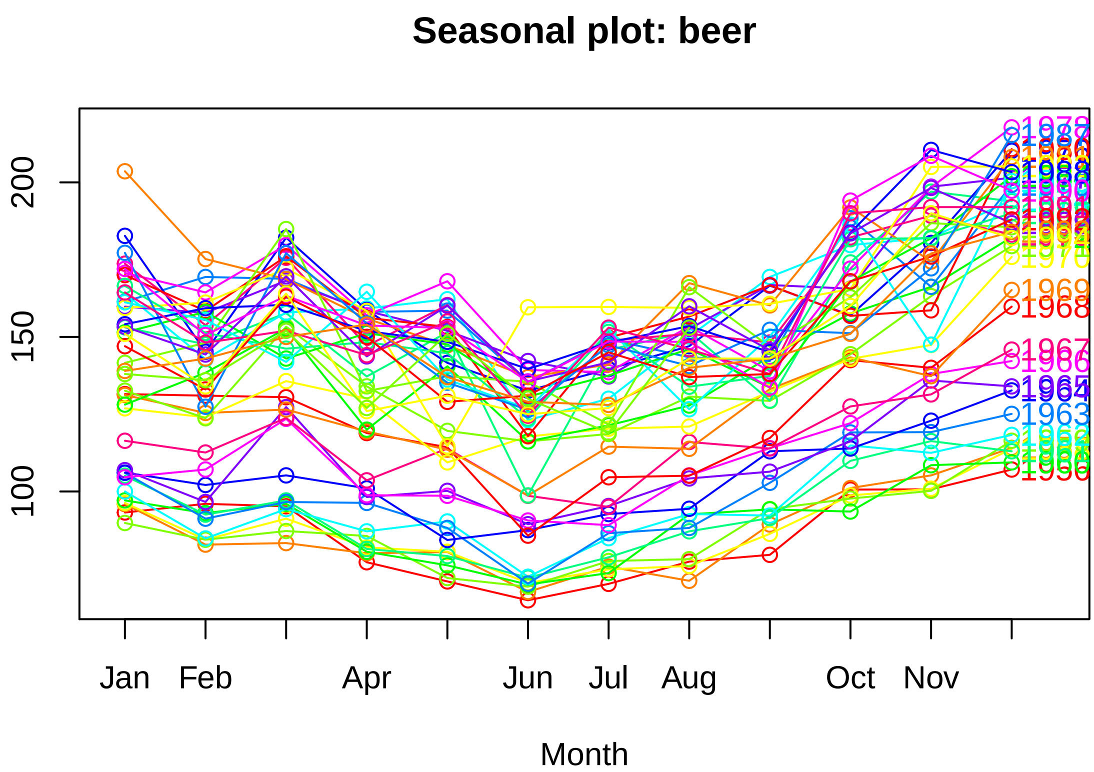
```
beer = ts(beer$Production, start = 1956, frequency = 12) # convert to Time Series
decomp.beer = decompose(beer, type="additive") # decompose the data
plot(decomp.beer)
```



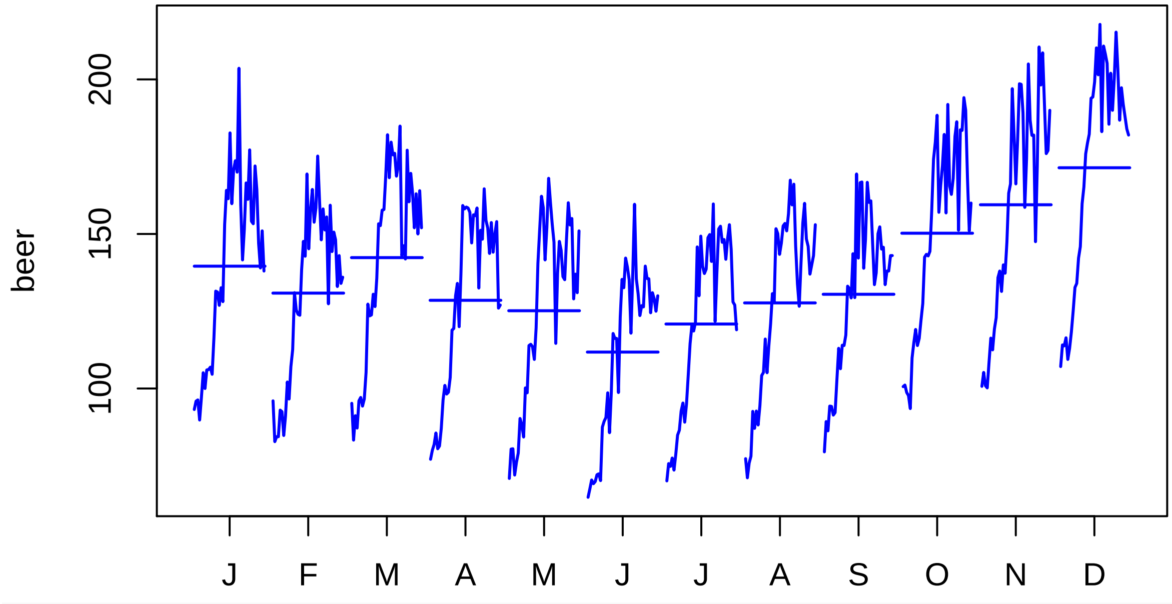

```
decomp.beer$figure # seasonal effects for each quarter
```

```
##  [1]   3.643172  -5.500204   6.215633  -7.319126 -11.503226 -25.051691
##  [7] -15.491764  -9.438772  -6.113986  13.536121  22.602360  34.421484
```

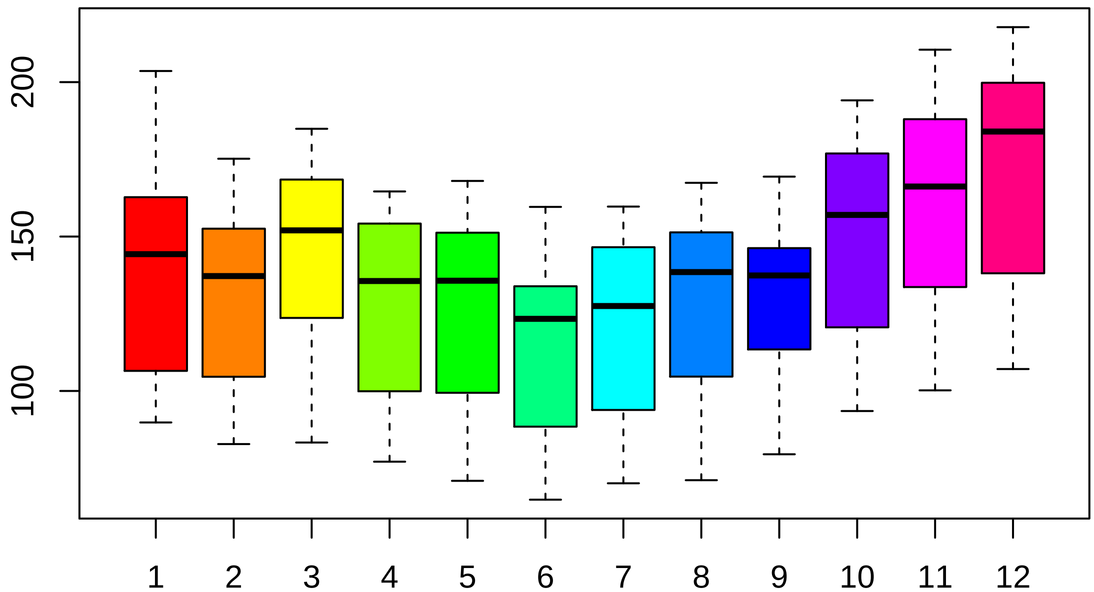
```
seasonplot(beer, col=rainbow(12), year.labels=TRUE)
```




```
monthplot(beer, col = 4, lwd = 1.5)
```

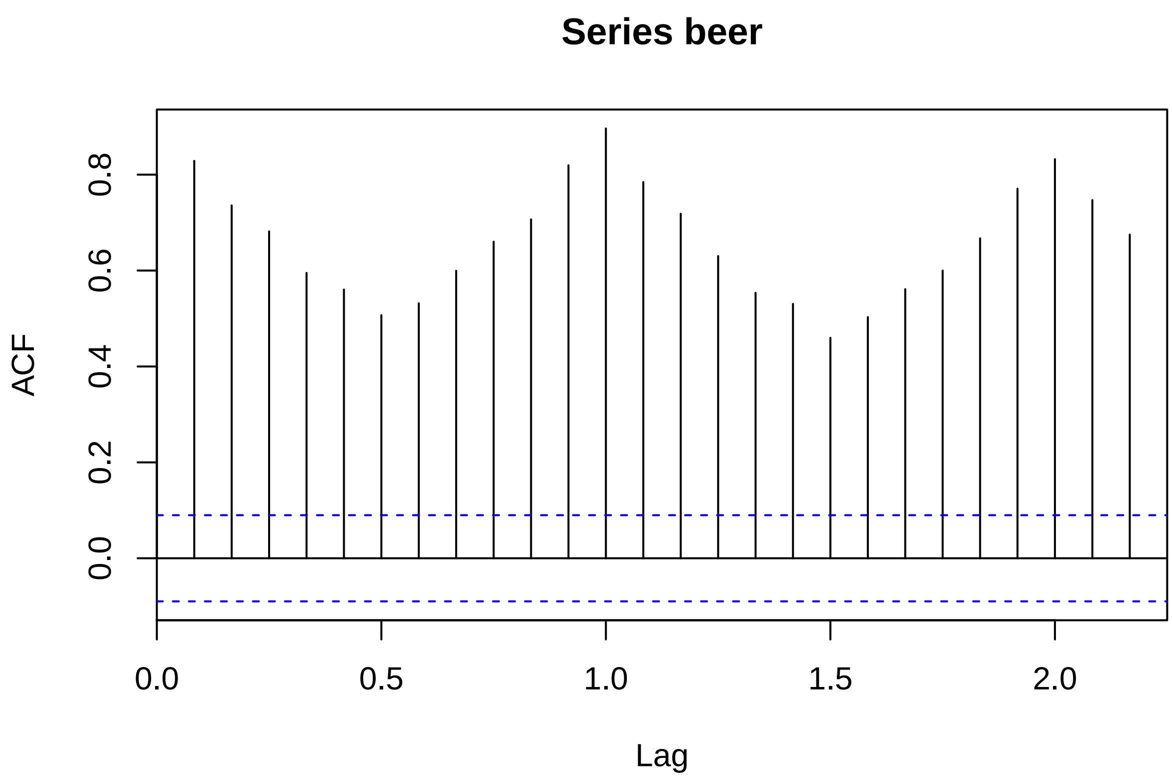



```
boxplot(beer ~ cycle(beer), col=rainbow(12))
```

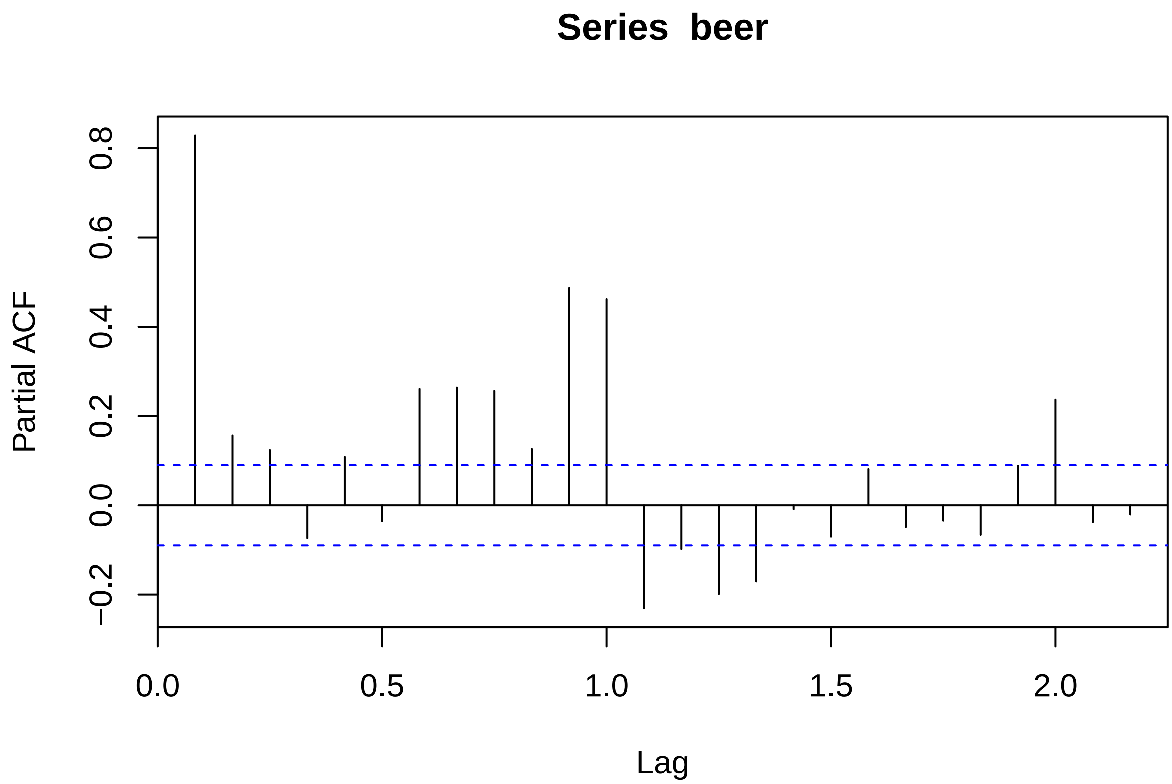


Plotting the data shows a seasonal pattern oscillating with spikes and valleys. The dataset is probably not stationary.

```r
# plot the ACF and PACF on the same graph
acf(beer)
```

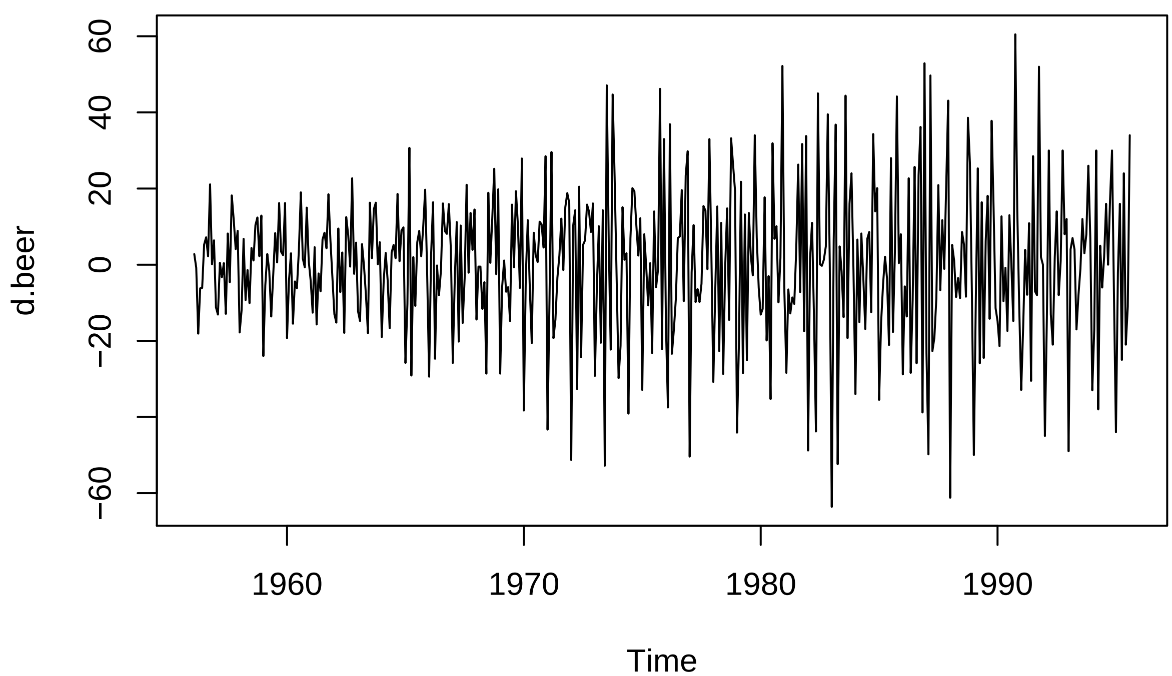



```r
pacf(beer)
```

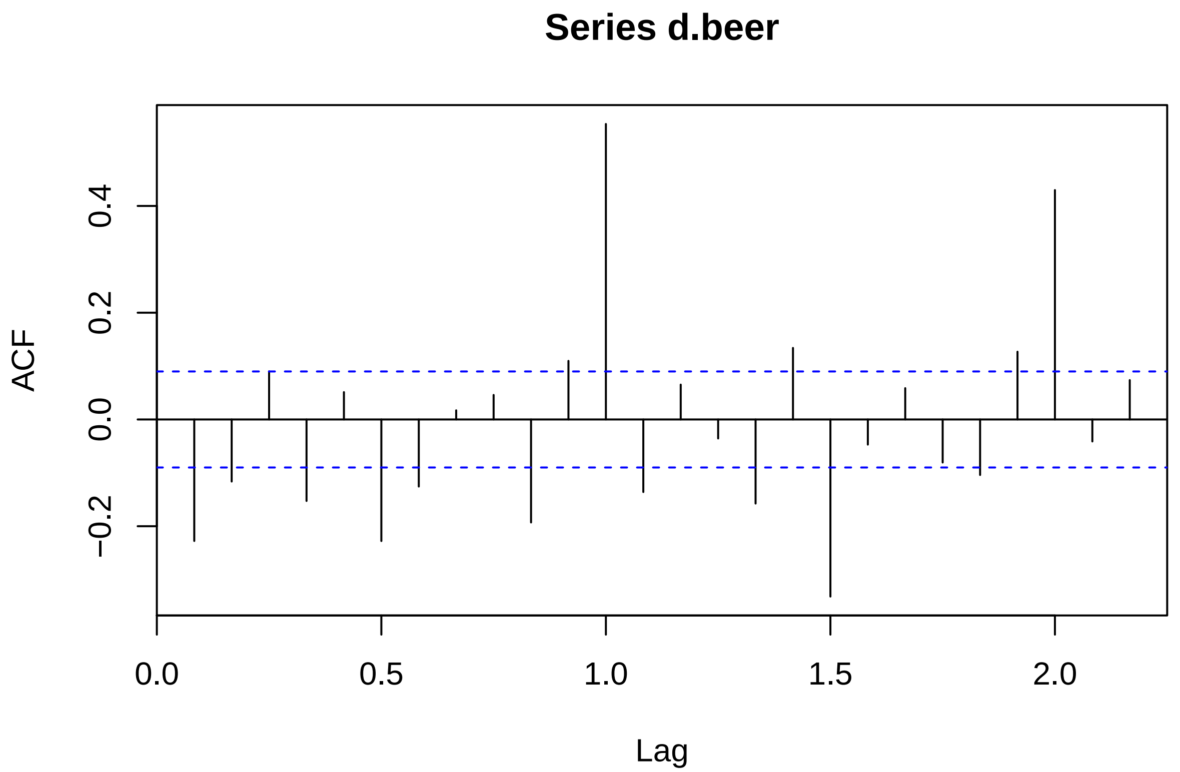


Based on the graph we see that the ACF does not tail off quickly hence it is not stationary, so we need to do some diffencing to get it to become stationary.

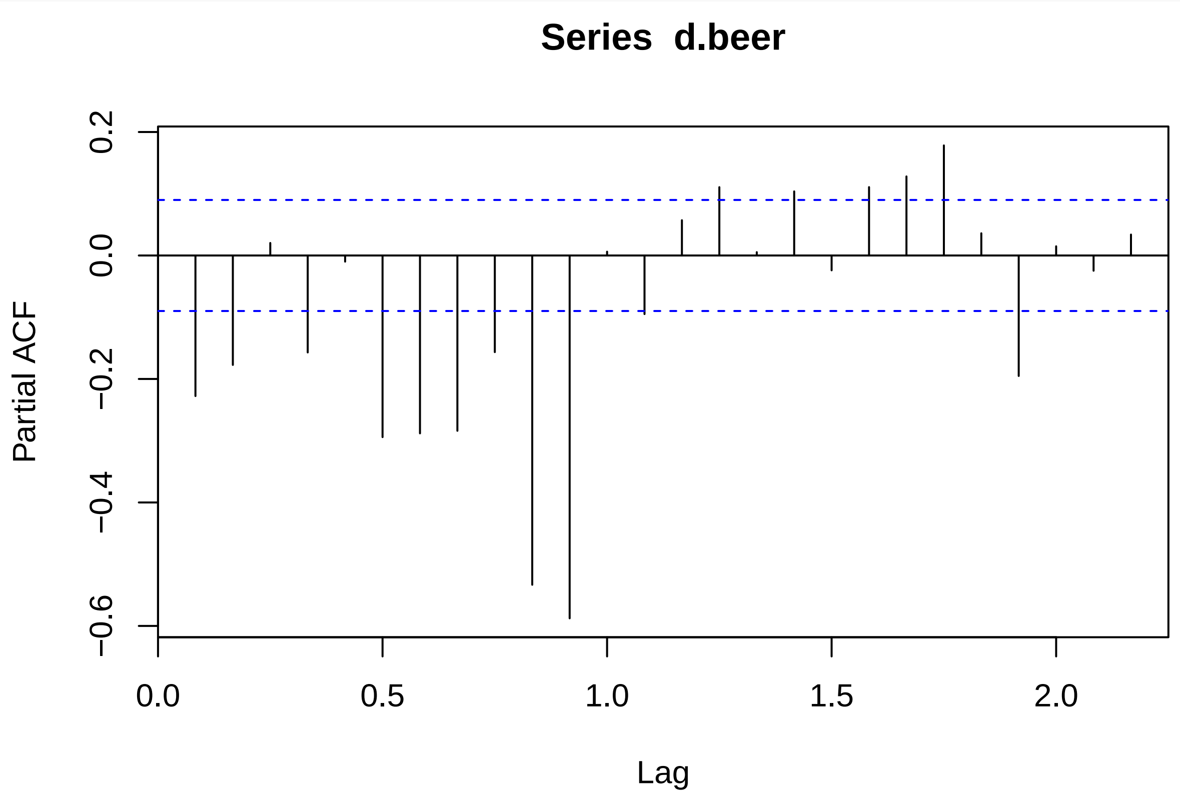
```
# differencing the series and plot
d.beer = diff(beer)
plot(d.beer)
```



```r
# par(mfrow=c(1,2))
acf(d.beer)
```




```r
pacf(d.beer)
```

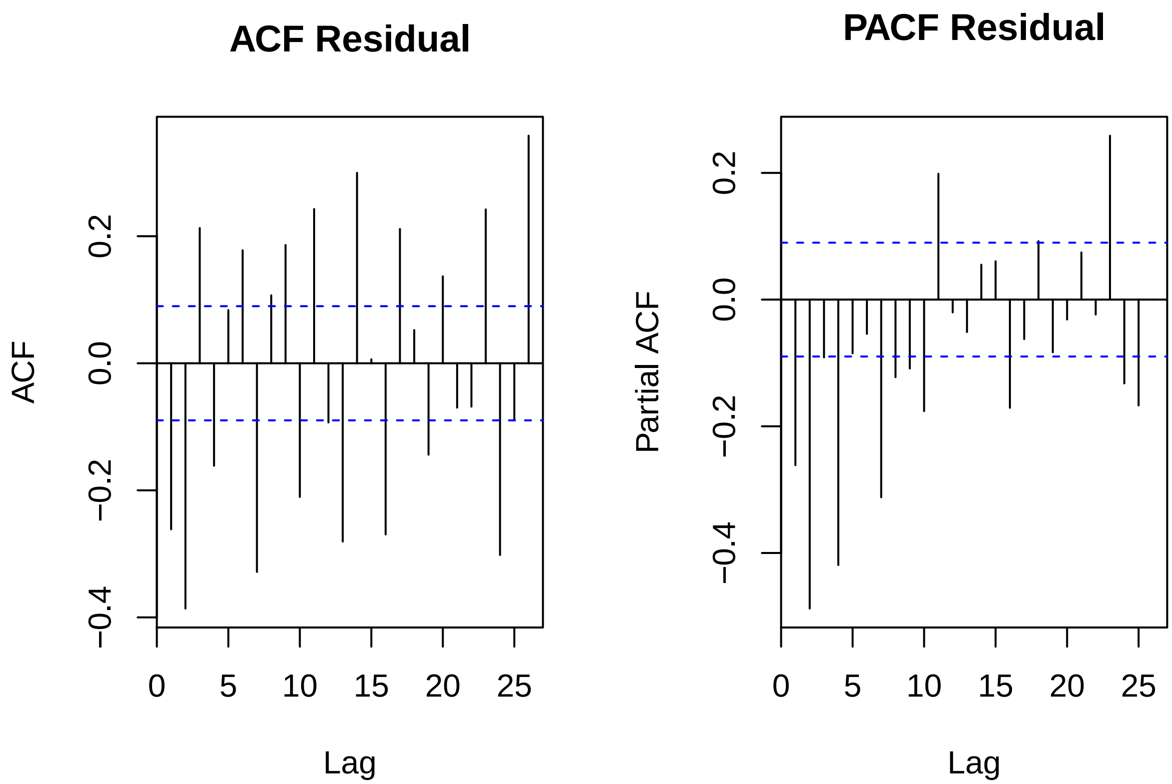


Now we see the graph no longer has a trend to it. However we can see these oscillating spikes keep getting larger as we move from left to right, which means there is heterostadaticity or non-constant variance. The ACF tails off quickly and there are regular spikes at every 12th period which corresponds to this being monthly data. So each 12th period is correlated with the previous 12. So annually there's a periodicity to it.

```r
library(forecast)
# fit ARIMA model
beer.fit = arima(beer, order = c(1,1,0),
                 seasonal = list(order = c(1,1,0), period = 12),
                 include.mean = FALSE)
beer.fit
```

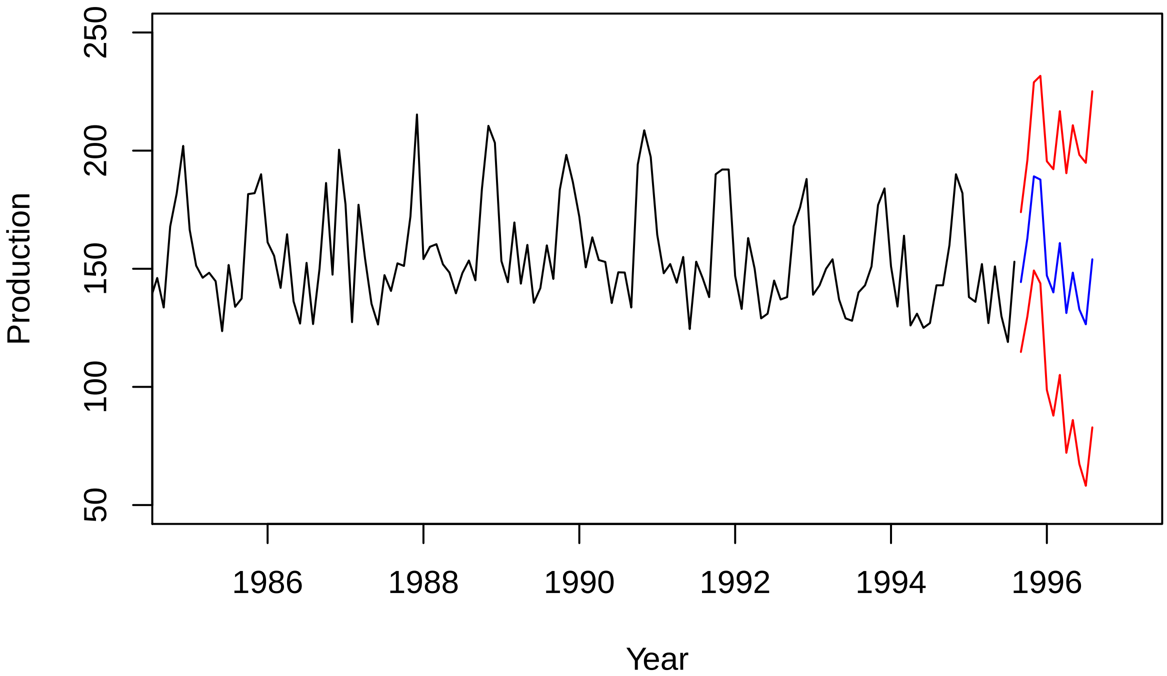
```
##
## Call:
## arima(x = beer, order = c(1, 1, 0), seasonal = list(order = c(1, 1, 0), period = 12),
##     include.mean = FALSE)
##
## Coefficients:
##           ar1     sar1
##       -0.4993  -0.3611
## s.e.   0.0403   0.0431
##
## sigma^2 estimated as 218.7:  log likelihood = -1905.25,  aic = 3814.51
```

```r
# residuals analysis
par(mfrow=c(1,2))
acf(ts(beer.fit$residuals),main='ACF Residual')
pacf(ts(beer.fit$residuals),main='PACF Residual')
```



Since there are spikes outside the insignificant zone for both ACF and PACF plots we can conclude that there are Moving average and Autoregressive pattern in the residuals and thus the residuals is NOT random in nature which is not desirable outcome.

```
# generate predictions
beer.pred = predict(beer.fit, n.ahead = 12) # predict 12 periods ahead
plot(beer, type = "l", xlim=c(1985,1997), ylim=c(50,250), xlab = "Year", ylab = "Production")
lines(beer.pred$pred, col = "blue") # mean of predictions into future (actual)
lines(beer.pred$pred + 2*beer.pred$se, col = "red") # prediction bounds +1.96
lines(beer.pred$pred - 2*beer.pred$se, col = "red") # prediction bounds -1.96
```



Blue line is the actual predictions, red lines the prediction bounds where the uncertainty predictions will fall for my future values. The blue looks like it's mimicking what's occurring in the previous time periods, 12 periods back. Prediction bounds also have zigzag in it reflecting that seasonality shape