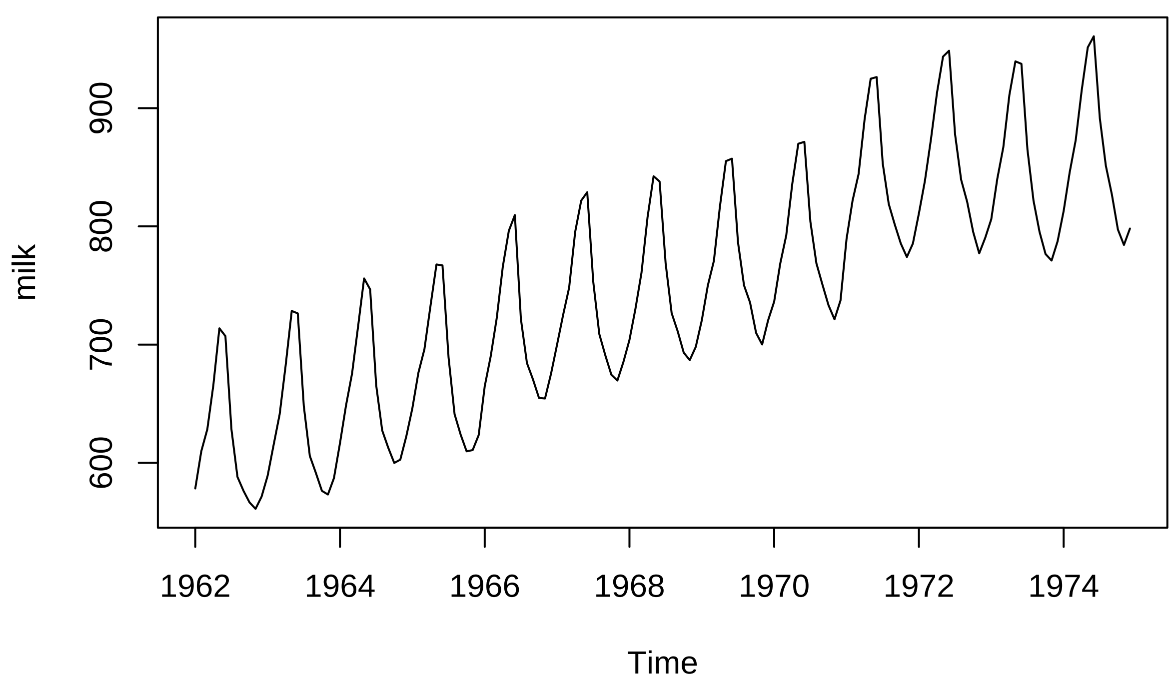
# pstat274lab5_

2023-11-03

1. We will analyze monthly milk production measured in pounds per from Jan. 1962 to Dec. 1975 from the package tsdl as Lab 4 (if you want to re-install tsdl, please refer to Lab 4). Let's denote the time series milk as Xt.

```
library(tsdl)
milk <- subset(tsdl, 12, "Agriculture")[[3]]
plot(milk)
```




(a) Explain why the series milk looks not stationary. To make series milk stationary, please difference at lag 12 and then at lag 1.

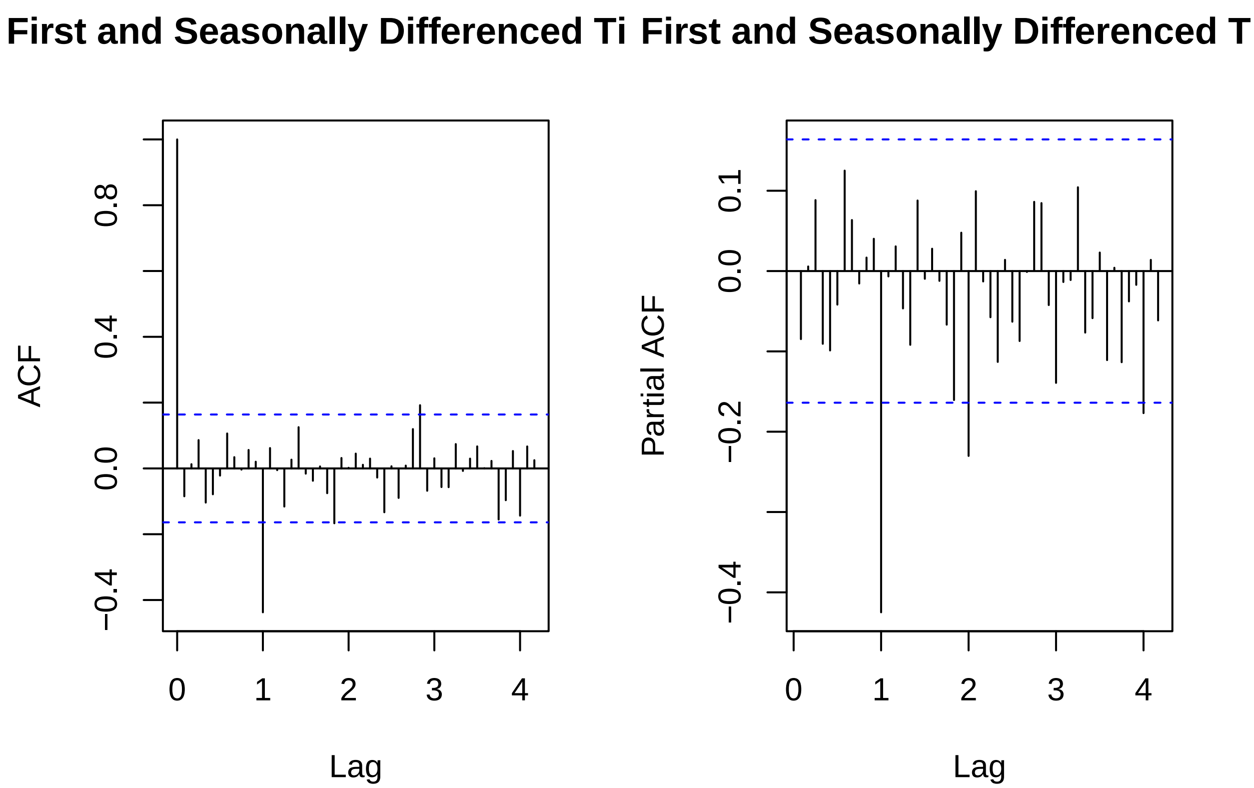
Series milk looks not stationary because the mean of series milk is not constant over time, which means there is trend for milk series.

```
# First difference at lag 12
dmilk <- diff(milk, lag = 12)
```

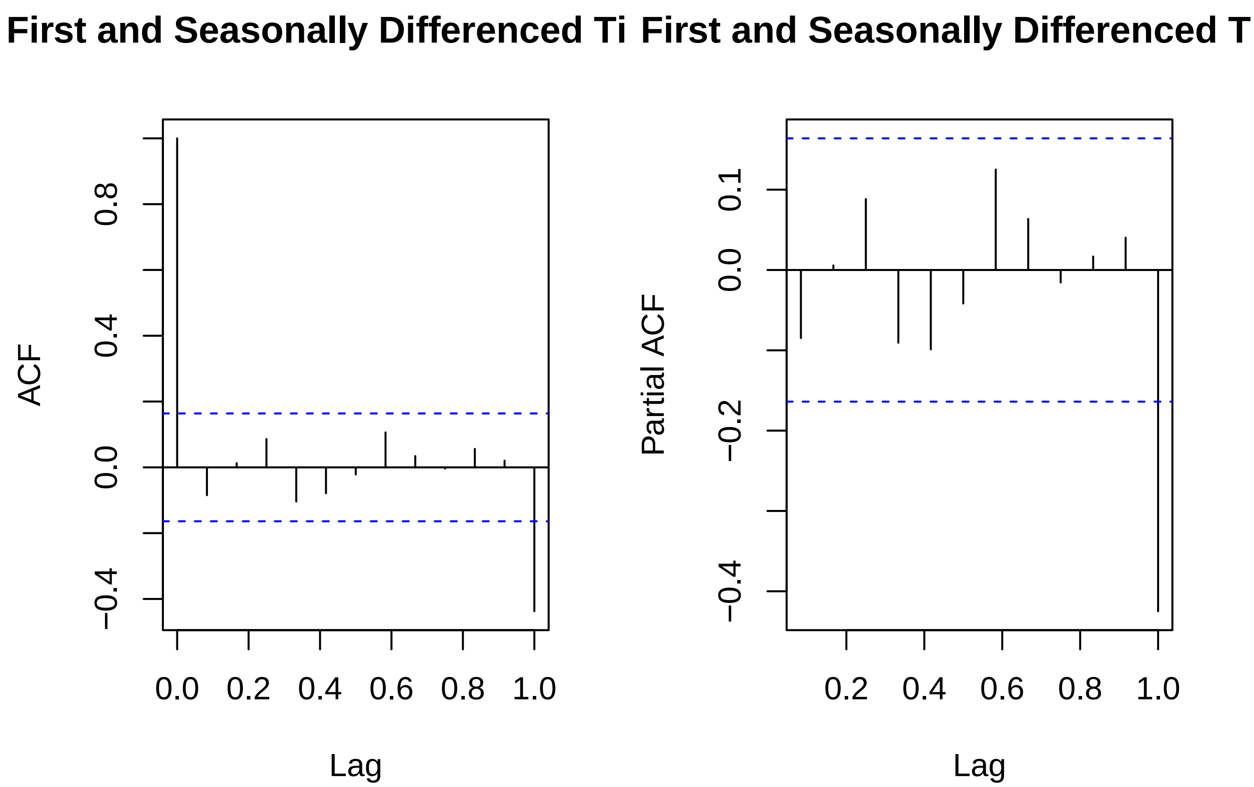
```r
# Then difference the result at lag 1
ddmilk <- diff(dmilk, lag = 1)
```

(b)

```r
op <- par(mfrow = c(1,2))
acf(ddmilk, lag.max = 50, main = "ACF: First and Seasonally Differenced Time Series")
pacf(ddmilk, lag.max = 50, main = "PACF: First and Seasonally Differenced Time Series")
```




```r
op <- par(mfrow = c(1,2))
acf(ddmilk, lag.max = 12, main = "ACF: First and Seasonally Differenced Time Series")
pacf(ddmilk, lag.max = 12, main = "PACF: First and Seasonally Differenced Time Series")
```




(c)

Modeling the seasonal part (P, D, Q): For this part, focus on the seasonal lags h = 1s, 2s, etc. We applied one seasonal differencing so D = 1 at lag s = 12. The ACF shows a strong peak at h = 1s. A good choice for the MA part could be Q=1. The PACF shows one strong peaks at h = 1s. also for 2s,4s. A good choice for the AR part could be P = 1, or 2,4. Modeling the non-seasonal part (p , d, q): In this case focus on the within season lags, h = 1,. . . ,11. We applied one differencing to remove the trend: d = 1 The ACF and PACF seems to be tailing off A good choice for the MA part could be q = 0 ,and a good choice for the AR part could be p = 0.

two possible model SARIMA(0, 1, 0) × (2, 1, 1)12 or SARIMA(0, 1, 0) * (4, 1, 1)12

(d) For SARIMA(0, 1, 0)(2, 1, 1)12

```
# install.packages("astsa")
library(astsa)
fit.i <- sarima(xdata = milk,p = 0, d = 1, q = 0,
                 P = 2 , D = 1, Q = 1, S = 12, details = F)
fit.i$fit
```

```
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
##     include.mean = !no.constant, transform.pars = trans, fixed = fixed, optim.control = list(trace =
##         REPORT = 1, reltol = tol))
```

```
##
## Coefficients:
##         sar1    sar2     sma1
##       0.0334  0.0196  -0.7017
## s.e.  0.1533  0.1209   0.1400
##
## sigma^2 estimated as 34.44:  log likelihood = -459.64,  aic = 927.28
```

Checking coefficients: From the above coefficients table,sar1 and sar2 are not significant because the confidence interval of the estimated coefficient contains 0. Therefore, we should set these coefficients to 0.

```
fit.i1 <- sarima(xdata = milk, p = 0, d = 1, q = 0,
                 P = 2 , D = 1, Q = 1, S = 12,
                 fixed=c(0, 0, NA),
                 details = F)
fit.i1$fit
```

```
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
##     include.mean = !no.constant, transform.pars = trans, fixed = fixed, optim.control = list(trace =
##         REPORT = 1, reltol = tol))
##
## Coefficients:
##       sar1  sar2     sma1
##          0     0  -0.6750
## s.e.     0     0   0.0752
##
## sigma^2 estimated as 34.47:  log likelihood = -459.66,  aic = 923.33
```

From the above output, we can write the model as: (1-B)(1-B12)Yt = (1-0.675B12)Zt

I don't why i could not us plot.roots function. here the root is outside the unit circle, representing this model is both stationary and invertible.

For SARIMA(0, 1, 0)(4, 1, 1)12

```
fit.i <- sarima(xdata = milk,p = 0, d = 1, q = 0,
                P =4 , D = 1, Q = 1, S = 12, details = F)
fit.i$fit
```

```
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
##     include.mean = !no.constant, transform.pars = trans, fixed = fixed, optim.control = list(trace =
##         REPORT = 1, reltol = tol))
##
## Coefficients:
##          sar1     sar2     sar3     sar4    sma1
##       -0.6659  -0.4356  -0.3459  -0.4243  0.0197
## s.e.   0.2086   0.1511   0.1144   0.0925  0.2348
##
## sigma^2 estimated as 30.36:  log likelihood = -454.12,  aic = 920.25
```

Checking coefficients: From the above coefficients table,sma1 is not significant because the confidence interval of the estimated coefficient contains 0. Therefore, we should set these coefficients to 0.

```
fit.i1 <- sarima(xdata = milk, p = 0, d = 1, q = 0,
                 P = 4 , D = 1, Q = 1, S = 12,
                 fixed=c(NA,NA,NA,NA,0),
                 details = F)
fit.i1$fit
```

```
##
## Call:
## arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
##     include.mean = !no.constant, transform.pars = trans, fixed = fixed, optim.control = list(trace =
##         REPORT = 1, reltol = tol))
##
## Coefficients:
##          sar1     sar2     sar3     sar4  sma1
##       -0.6497  -0.4260  -0.3412  -0.4230     0
## s.e.   0.0790   0.0979   0.0993   0.0917     0
##
## sigma^2 estimated as 30.36:  log likelihood = -454.13,  aic = 918.26
```

here is the model (1-0.6497B12-0.426B24-0.3412B36-0.423B12)Xt=Zt

There exist roots inside the circle, indicating it is not stationary.

The final fitting model would be (1-B)(1-B12)Yt = (1-0.6750B12)Zt.