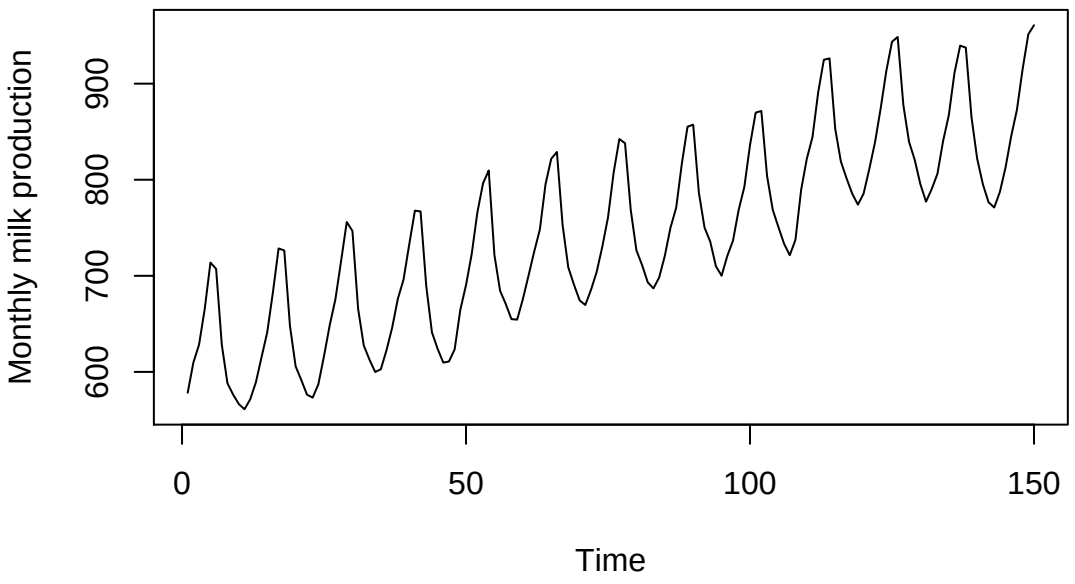
Lab Assignment 7

PSTAT 174/274

1. We will again analyze adjusted monthly milk production measured in pounds per from Jan. 1962 to Dec. 1975. And we can import the dataset from `tsdl` package as `milk` in `R`, and denote the milk time series as X_t. For comparison, we split the dataset into training set `train` and testing set `test`. The training set is used for model building, and the testing set is used for prediction verification and comparision.

```
library(tsdl)
milk <- subset(tsdl, 12, "Agriculture")[[3]]
train <- milk[1:150]
test <- milk[151:156]
ts.plot(train, ylab = "Monthly milk production")
```



From the above graph, we can conclude that X_t is non-stationary because of the upward trend and seasonality. (You can think about whether we need to transform the series or not.) To make it more stationary, we need to remove trend and seasonality with the following code:

```
dmilk <- diff(train, 12)
ddmilk <- diff(dmilk, 1)
```

Let Y_t denote the series `ddmilk`. Then, $Y_t = (1-B)(1-B^{12})X_t$. As Lab Assignment 5, we can use SARIMA$(0,1,0) \times (1,1,1)_{12}$. Now, we will conduct model diagnostic analysis.

```
library(astsa)
fit.i <- sarima(xdata=train, p=0, d=1, q=0, P=1, D=1, Q=1, S=12)
```

(a) Perform diagnostics on the chosen model fit. Do the residuals appear to be white noise? Are they normally distributed? You should conduct hypothesis testing and plot some graphs to answer this questions. (You can think about why we want to check normality of the residuals.)

(b) Forecast the next 6 observations using `sarima.for()`, and plot your predictions. And you should also add true milk prodcution points in `test`.