2025-11-28_In Class Exercise

Htet Khant Linn

2025-11-29

#Library Installation

```
library(zoo)
```

```
##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

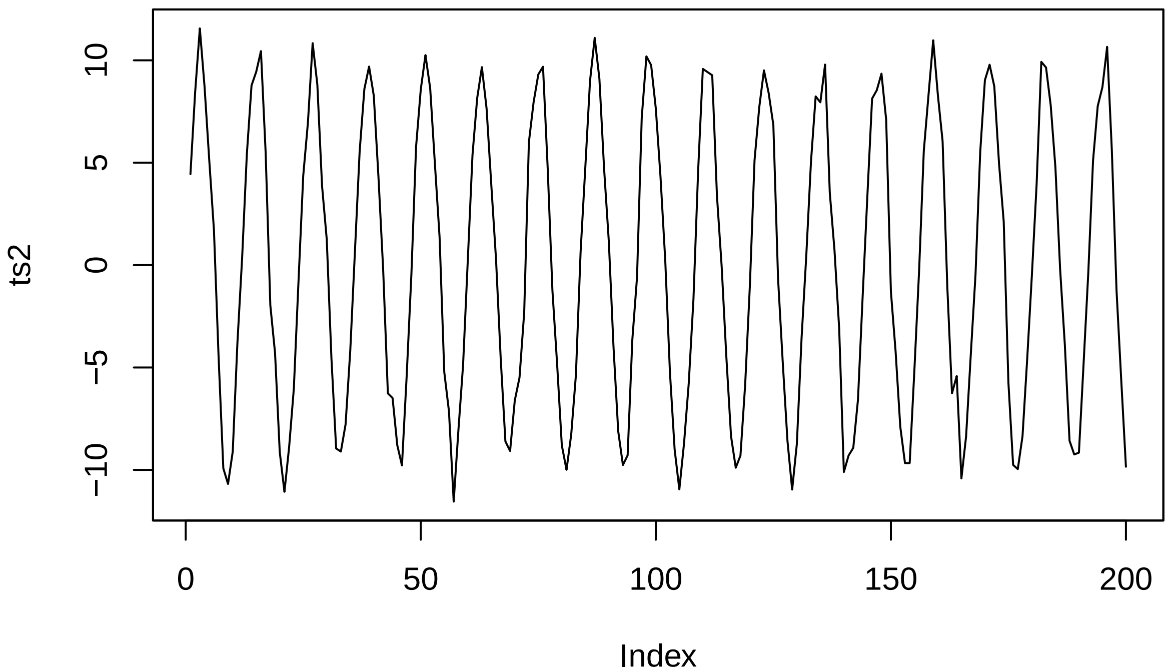
```
library(forecast)
```

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

```
library(ggplot2)
library(tseries)
```

```
 time <- 1:200
set.seed(123)
 seasonal <- 10*sin(2*pi*time/12) + rnorm(200)

 ts2 = zoo(seasonal, time)
 plot(ts2)
```



#Perform ADF test

```
adf.test(ts2)
```

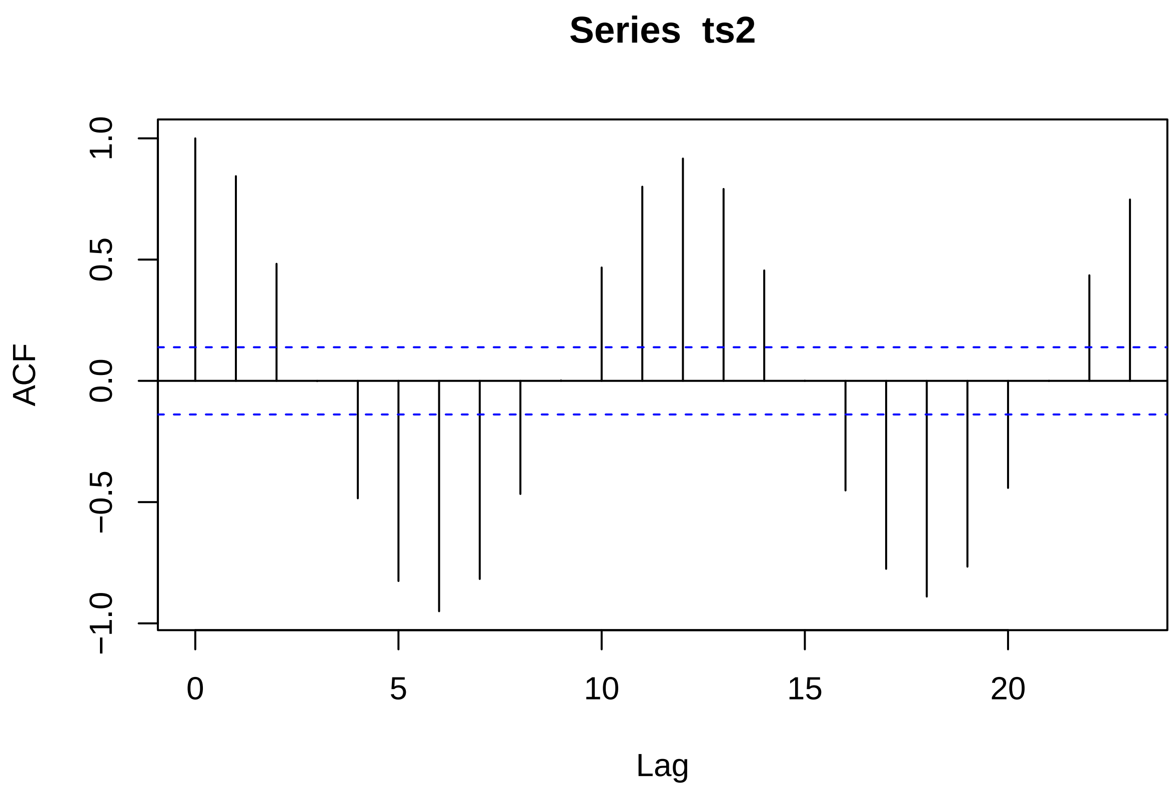
```
## Warning in adf.test(ts2): p-value smaller than printed p-value
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  ts2
## Dickey-Fuller = -16.426, Lag order = 5, p-value = 0.01
## alternative hypothesis: stationary
```

The test yields a p-value of 0.01, which supports the conclusion that the time series is stationary and suitable for modeling.

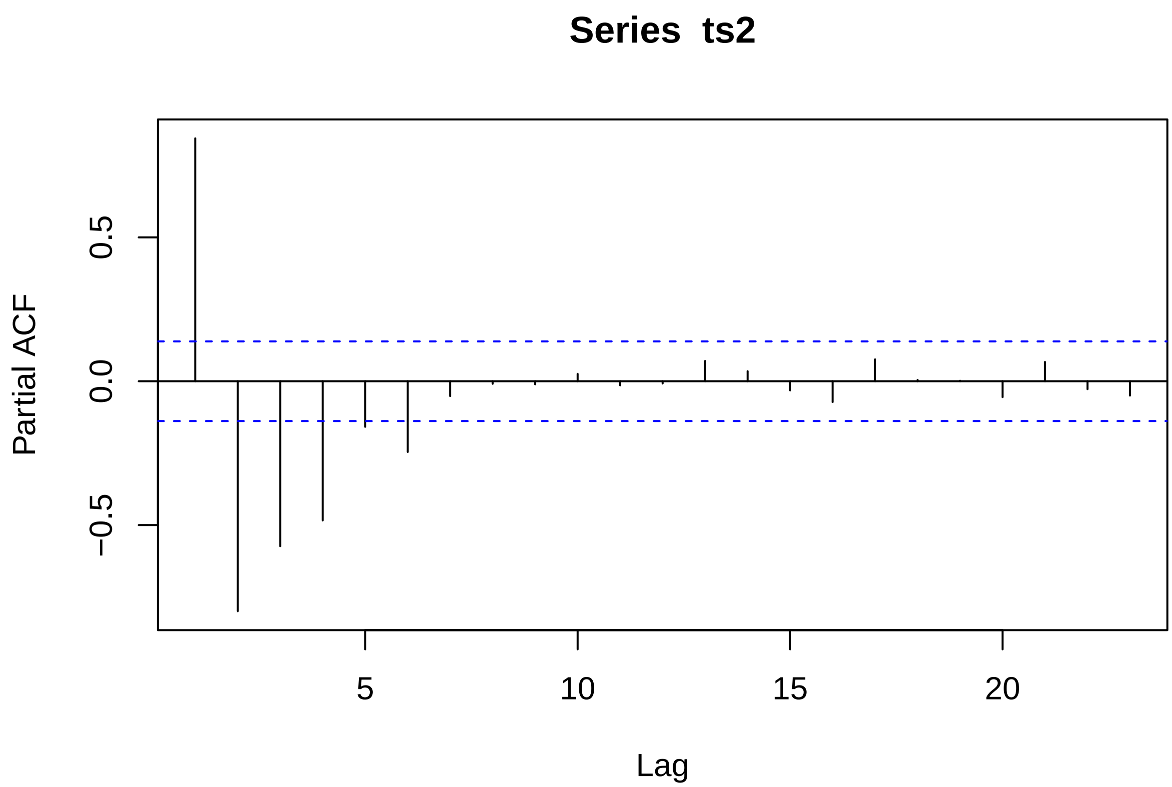
#Perform ACF test

```
acf(ts2)
```



The plot displays a distinct wave-like pattern, serving as evidence of strong seasonality within the data. #Perform PACF test

```
pacf(ts2)
```



Significant spikes at the initial lags indicate that immediate past observations are strong predictors of current values.

#Fit an automated ARIMA model

```
model = auto.arima(ts2,trace=TRUE, seasonal= TRUE)
```

```
##
##  Fitting models using approximations to speed things up...
##
##  ARIMA(2,0,2) with non-zero mean : Inf
##  ARIMA(0,0,0) with non-zero mean : 1355.511
##  ARIMA(1,0,0) with non-zero mean : 1102.86
##  ARIMA(0,0,1) with non-zero mean : 1143.273
##  ARIMA(0,0,0) with zero mean     : 1353.518
##  ARIMA(2,0,0) with non-zero mean : 865.6897
##  ARIMA(3,0,0) with non-zero mean : 770.9281
##  ARIMA(4,0,0) with non-zero mean : Inf
##  ARIMA(3,0,1) with non-zero mean : Inf
##  ARIMA(2,0,1) with non-zero mean : Inf
##  ARIMA(4,0,1) with non-zero mean : Inf
##  ARIMA(3,0,0) with zero mean     : 768.8365
##  ARIMA(2,0,0) with zero mean     : 863.6135
##  ARIMA(4,0,0) with zero mean     : Inf
##  ARIMA(3,0,1) with zero mean     : Inf
##  ARIMA(2,0,1) with zero mean     : Inf
##  ARIMA(4,0,1) with zero mean     : Inf
```

```
##
##  Now re-fitting the best model(s) without approximations...
##
##  ARIMA(3,0,0) with zero mean     : 773.954
##
##  Best model: ARIMA(3,0,0) with zero mean
```

The automated algorithm identified an ARIMA(3,0,0) model as the best fit, utilizing the last three data points for prediction.

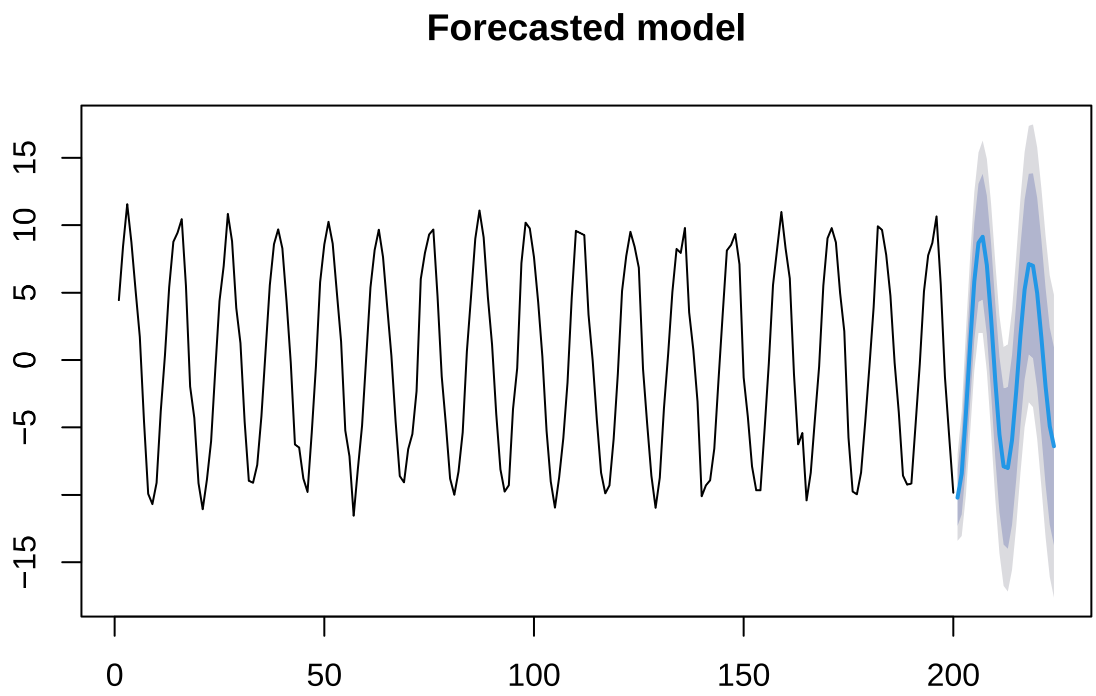
#Model Summary

```r
summary(model)
```

```
## Series: ts2
## ARIMA(3,0,0) with zero mean
##
## Coefficients:
##          ar1     ar2      ar3
##       1.0435  0.1267  -0.6186
## s.e.  0.0551  0.0918   0.0554
##
## sigma^2 = 2.665:  log likelihood = -382.87
## AIC=773.75   AICc=773.95   BIC=786.94
##
## Training set error measures:
##                      ME     RMSE      MAE      MPE    MAPE      MASE       ACF1
## Training set 0.01073615 1.620109 1.278655 4.819774 54.4755 0.3752211 -0.3860561
```

#Forecasting for 2 years

```r
forecast_future= forecast(model, h=24)
plot(forecast_future, main="Forecasted model")
```



#Ljung_Box test for model validation

```
Box.test(forecast_future$residuals,type="Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  forecast_future$residuals
## X-squared = 30.257, df = 1, p-value = 3.784e-08
```

The resulting p-value is extremely low, indicating that significant patterns remain in the residuals and the model fit is not yet optimal.