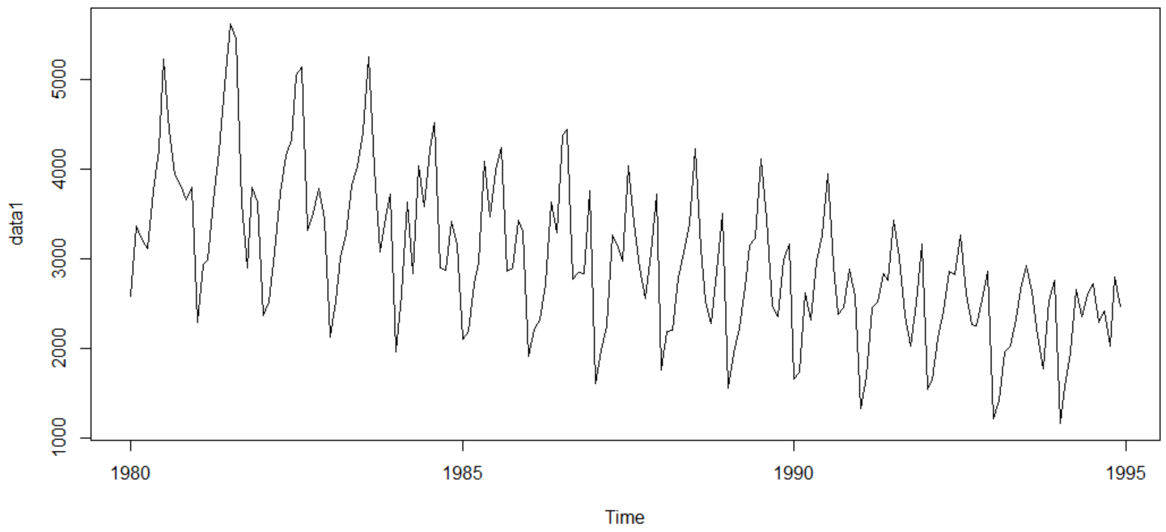
TIME SERIES ANALYSIS HANDS-ON

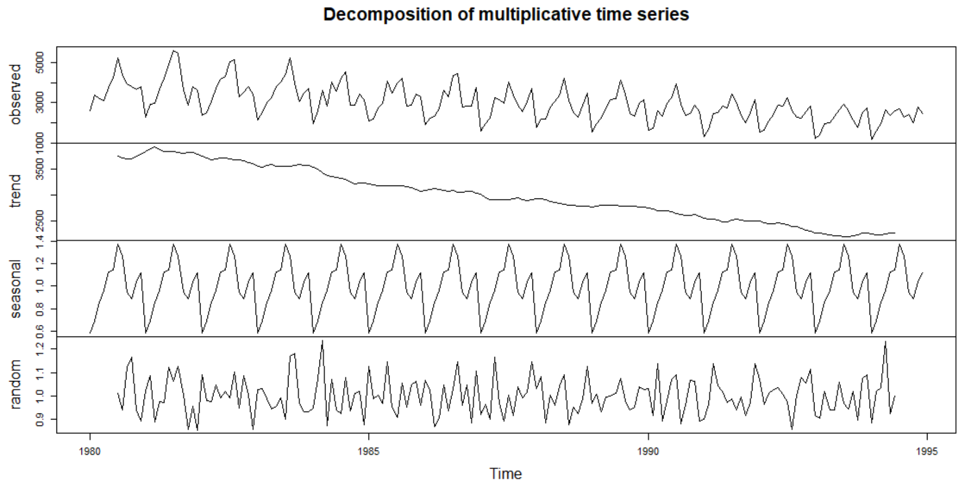
(21060641042)

TIME SERIES PLOT



Interpretation: There is decreasing trend visible along with seasonality.

DECOMPOSITION OF TIME SERIES DATA



Interpretation: Using decomposition we can confirm that there is clear decreasing trend and seasonal pattern present.

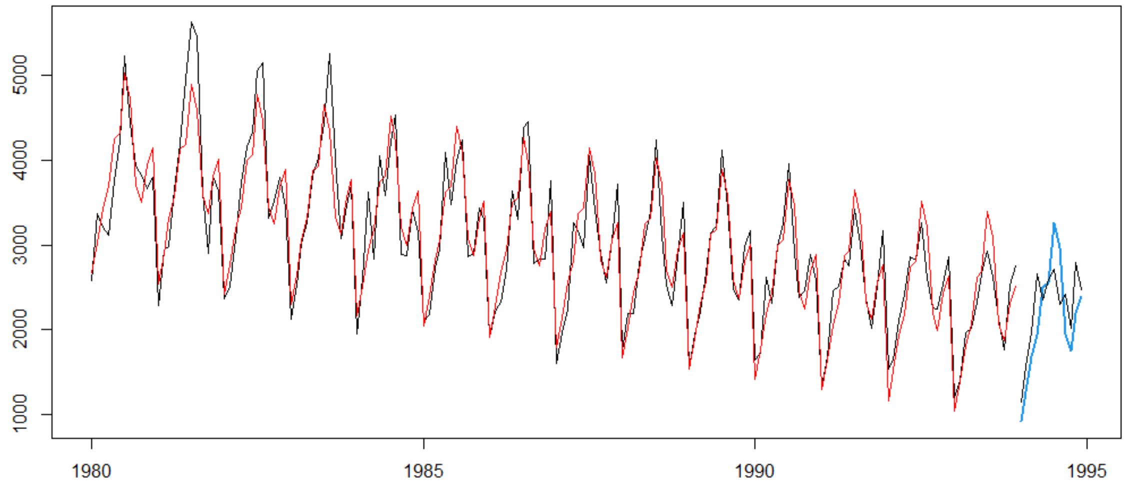
PARTITIONING DATA INTO TRAIN AND TEST

Length of train data: 168

Length of test data: 12

FITTING REGRESSION MODEL

Forecasts from Linear regression model

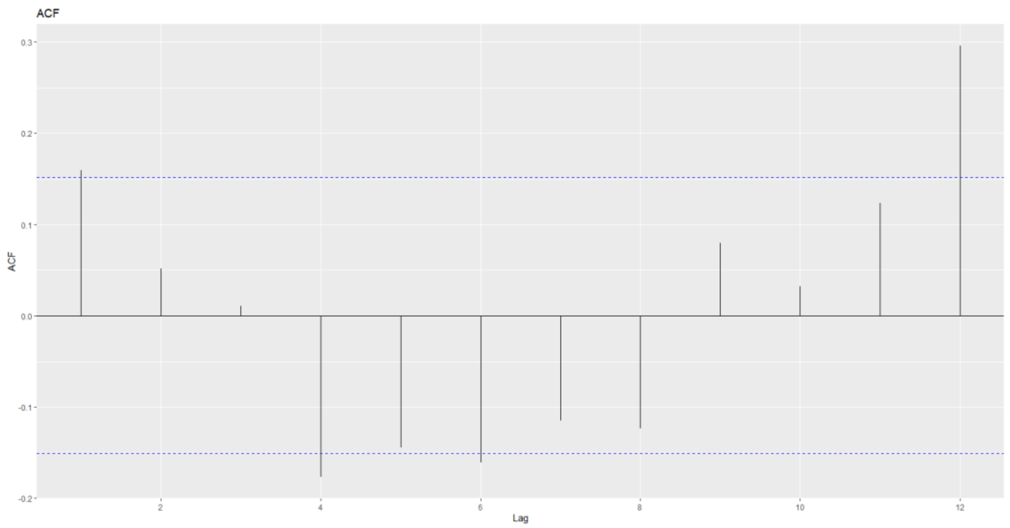


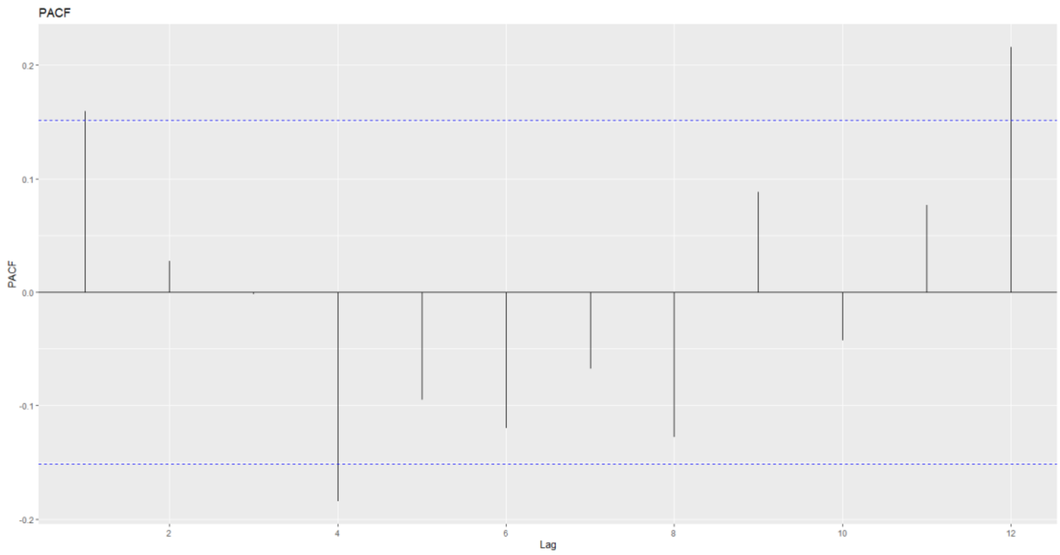
Black line: Actual values

Red line: Predicted values

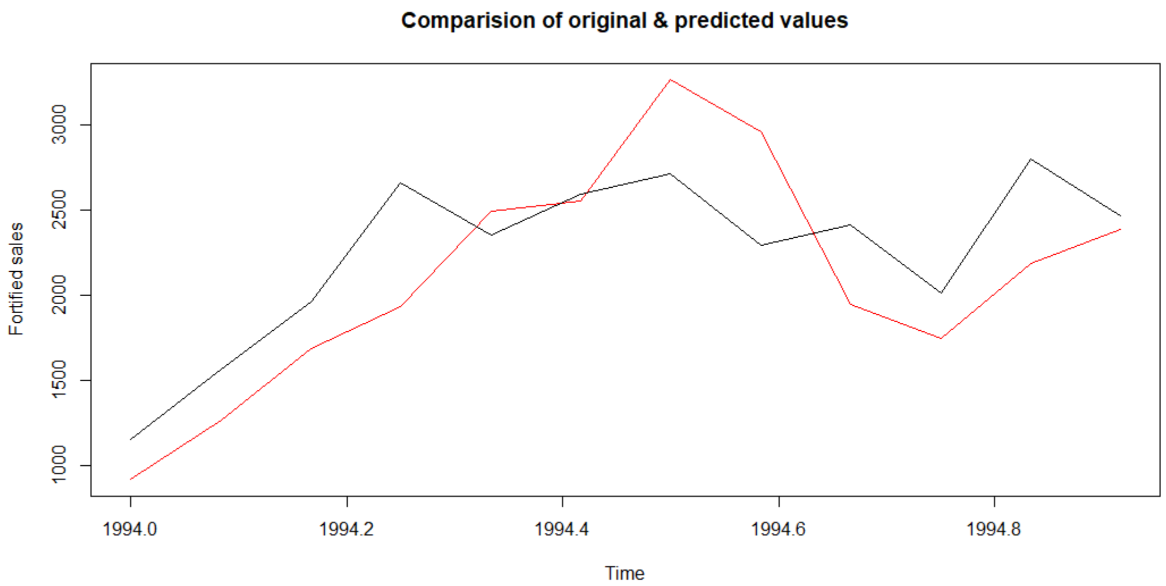
Blue line: Forecasted values

ACF PLOTS FOR RESIDUAL





ACTUAL V/S FORECAST PLOT



<u>ACCURACY</u>

	ME	RMSE	MAE	MPE	MAPE	ACF1	Theil's U
Test set	135.4991	426.5323	362.69	6.997418	**16.25365**	0.2161928	0.9317167

The MAPE value is high : 16.25365

i. Examining this plot, which of the following statements

are reasonable?

• Decembers (month 12) are not captured well by the

model.

• There is a strong correlation between sales on the

same calendar month.

• The model does not capture the seasonality well.

• We should try to fit an autoregressive model with

lag-12 to the residuals

Answer: The model does not capture the seasonality well and Decembers (month 12) are not captured well by the model.

ARIMA MODEL

<u>KPSS TEST</u>

Null Hypothesis (HO): Series is trend stationary or series has no unit root.

Alternate Hypothesis(HA): Series is non-stationary or series has a unit root.

If p-value < 0.05 – Reject Null Hypothesis(HO)

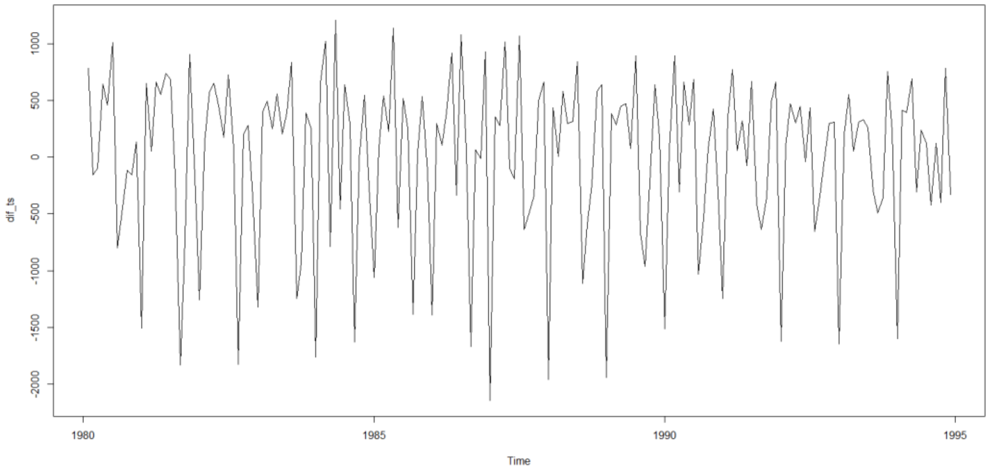
i.e., time series has a unit root, meaning the series is non stationary.

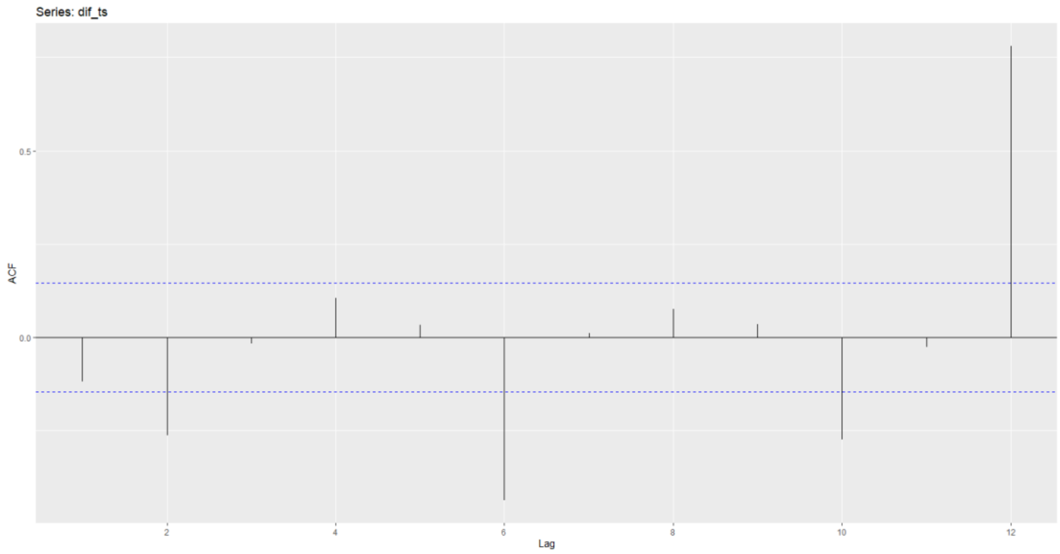
```
        KPSS Test for Level Stationarity

data:  data1
KPSS Level = 2.1912, Truncation lag parameter = 4, p-value = 0.01
```

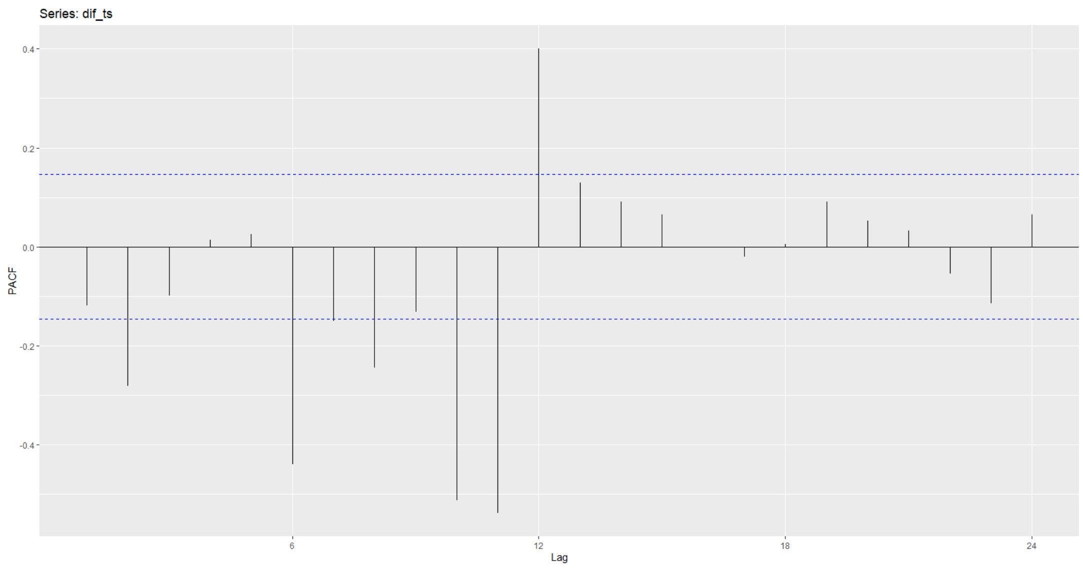
Since, p value<0.05, we reject null hypothesis.

Therefore, the time series is not stationery.

Hence, we use differencing of order 1.

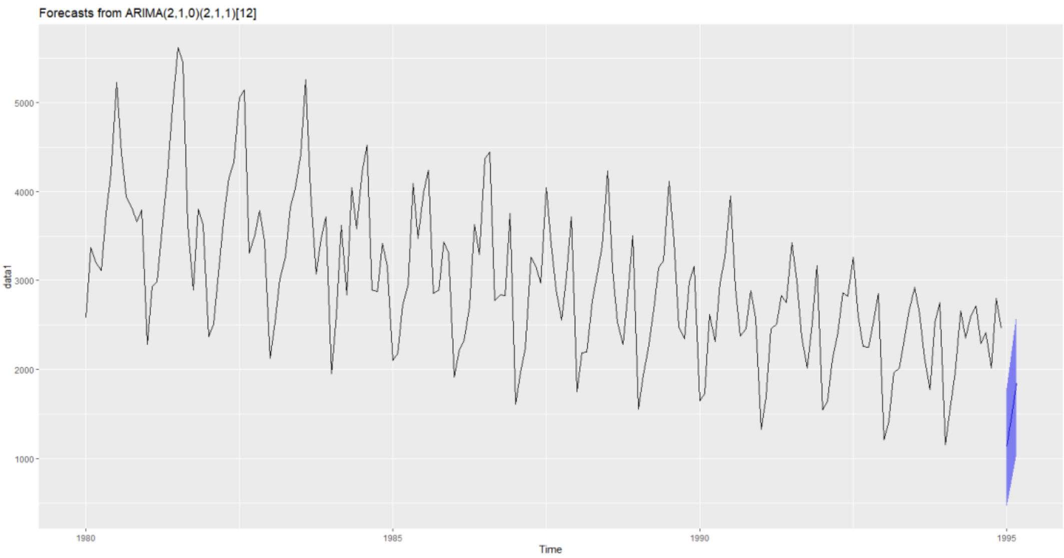






From the ACF and PACF plots, the Arima model is fitted using ARIMA(2,1,0)(2,1,1) for forecasting.

ARIMA(2,1,0)(2,1,1)



ACCURACY

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	5.140095	325.7143	246.1573	-0.175025	**8.460413**	0.4309342	-0.05251176

Auto Arima model: ARIMA(0,0,0)(2,1,1)

Series: data1

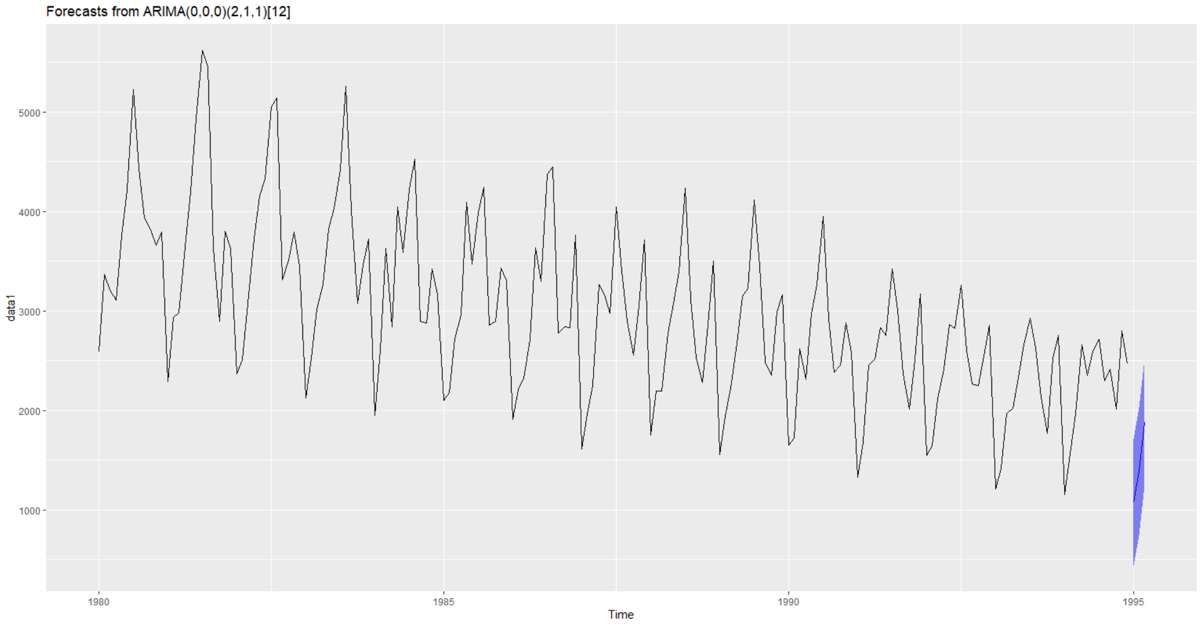
ARIMA(0,0,0)(2,1,1)[12] with drift

Coefficients:

	sar1	sar2	sma1	drift
	0.2431	0.0981	-0.7687	-9.8689
s.e.	0.1755	0.1254	0.1542	0.8464

sigma^2 = 90430: log likelihood = -1197.55

AIC=2405.1   AICc=2405.47   BIC=2420.72



ACCURACY

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	-69.38058	309.2621	228.2054	-2.762399	**7.820937**	0.3995068	0.1115797

SELECTION AND INTERPRETATION OF MODEL

RESULTS	
Model	**MAPE**
Regression	16.25365
ARIMA(2,1,0)(2,1,1)	8.460413
ARIMA(0,0,0)(2,1,1)	7.820937

Interpretation: The ARIMA(0,0,0)(2,1,1) model is better than the Regression and ARIMA(2,1,0)(2,1,1).

We choose **ARIMA(0,0,0)(2,1,1)** model for forecasting as the MAPE value is lower. Therefore, it is more accurate.