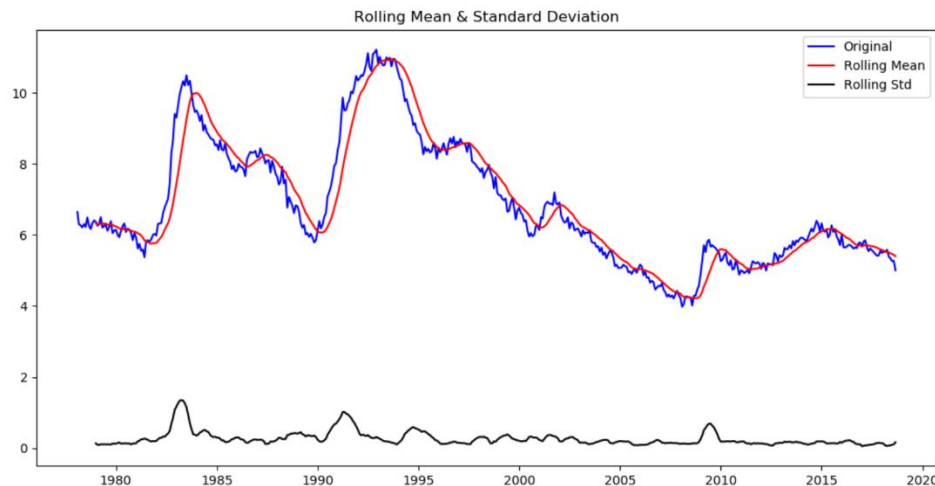


Harmonized Unemployment Rate

The dataset is from the [Organization for Economic Co-operation and Development \(OECD\)](#) hosted by the Federal Reserve Economic Database (FRED).

Harmonized Unemployment Rate for Australia

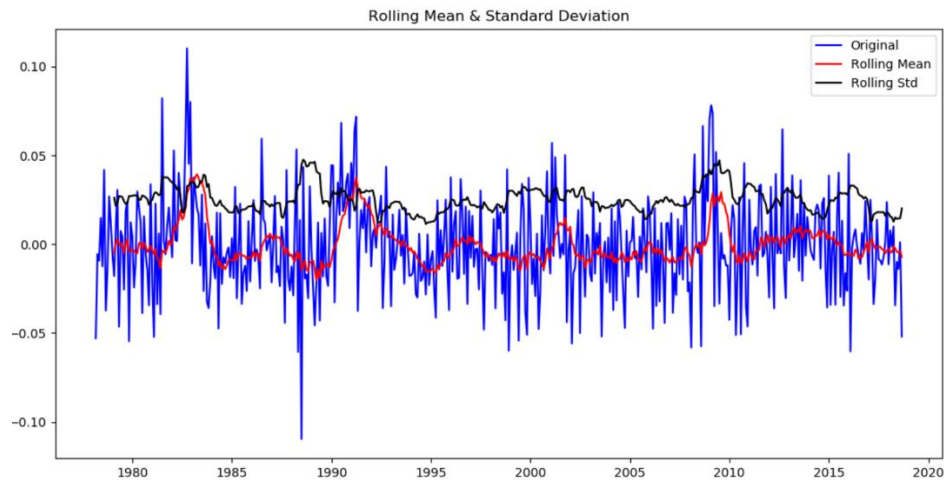
- **Checked for stationarity** – Plotted the graph of the original series and of the rolling mean and standard deviation. As we can see that the mean is not constant, so the series is not stationary.



```
Results of Dickey-Fuller Test:
Test Statistic          -1.797214
p-value                  0.381846
#Lags Used               12.000000
Number of Observations Used 475.000000
Critical Value (1%)      -3.444192
Critical Value (5%)      -2.867644
Critical Value (10%)     -2.570021
dtype: float64
```

The p value from the Dickey-Fuller Test is 0.3818 so we cannot reject the null hypothesis of non-stationarity and conclude that the series is not stationary.

- **Made the Series Stationary:** There are several approaches to make the series stationary, I used difference technique to make the series stationary and observed that the series was stationary after 1 difference.

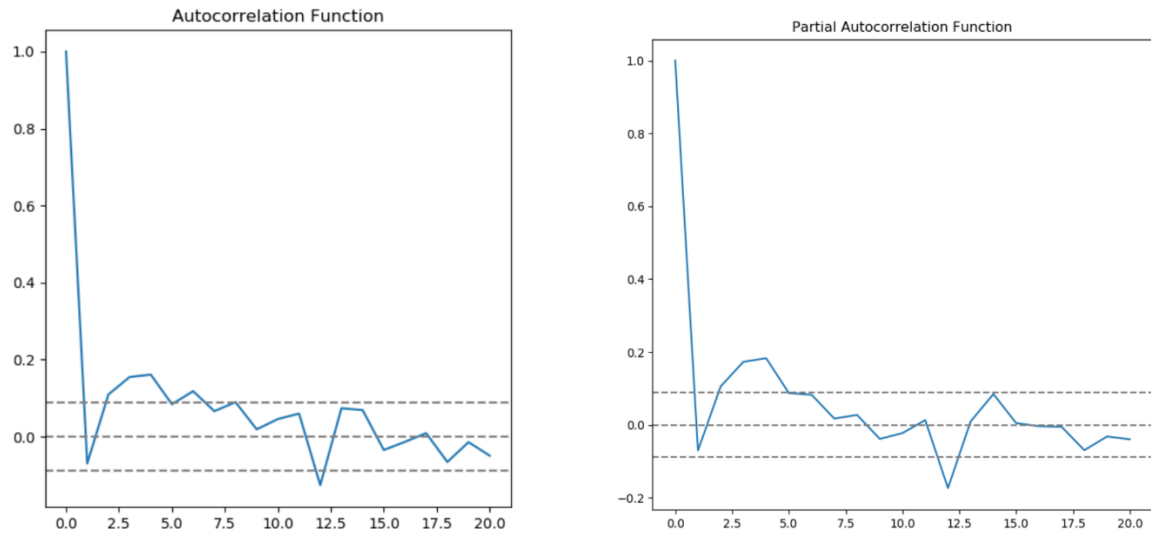


```
Results of Dickey-Fuller Test:
Test Statistic          -5.655912e+00
p-value                 9.622260e-07
#Lags Used              1.100000e+01
Number of Observations Used  4.750000e+02
Critical Value (1%)     -3.444192e+00
Critical Value (5%)     -2.867644e+00
Critical Value (10%)    -2.570021e+00
dtype: float64
```

As we can see from the graph and the p value for Dickey-Fuller test we can easily say the series is stationary.

- **Checked the Number of lags to be included in the ARIMA model:**

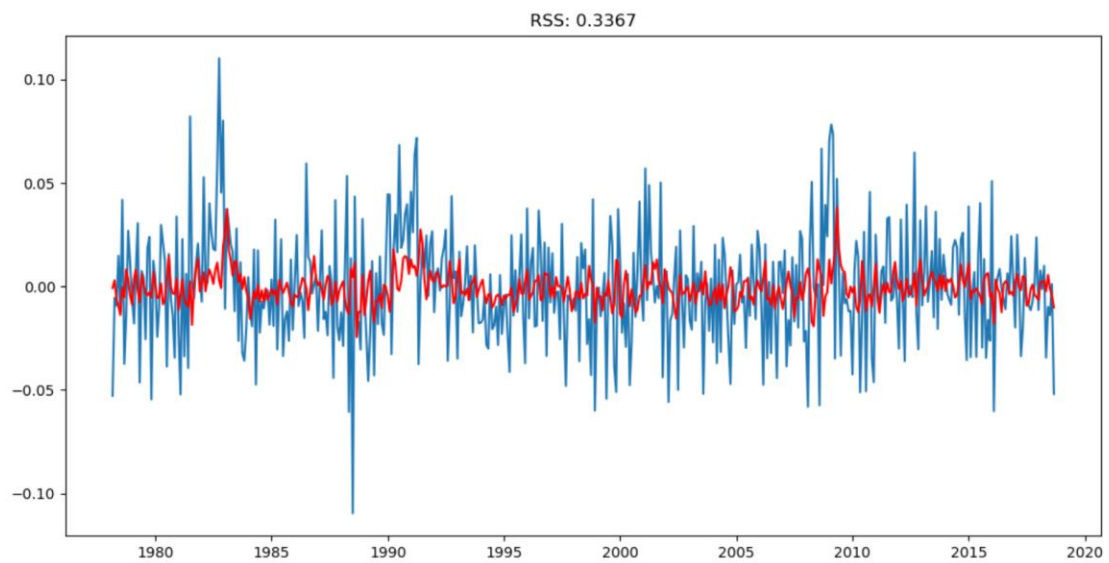
ACF & PACF



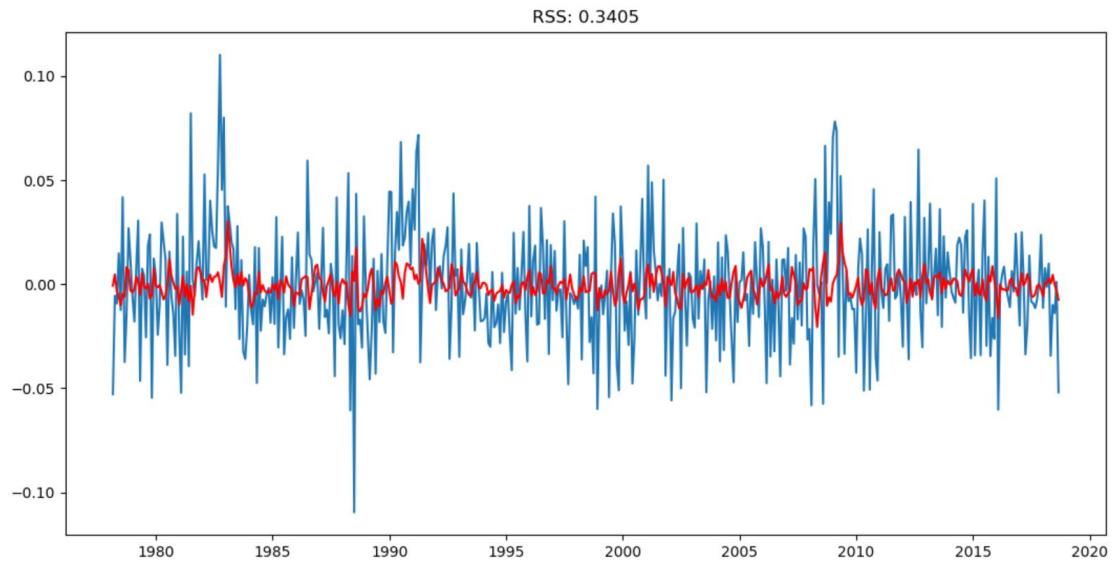
From ACF we find that the number of lags required are 5 and from PACF we get 4 lags.

- **ARIMA Models:**

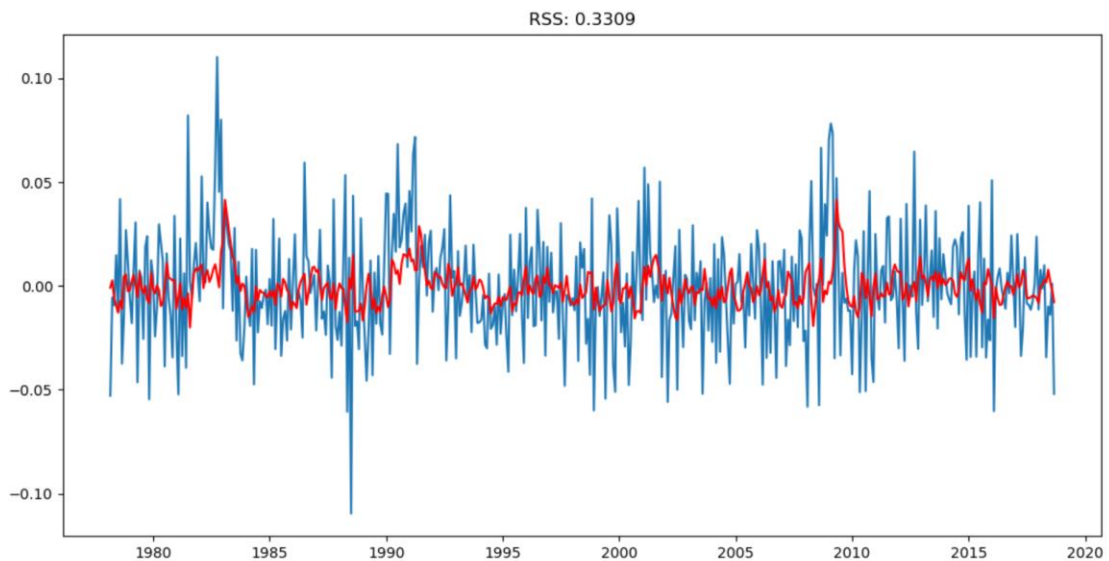
ARIMA (4,1,0)



ARIMA (0,1,5)



ARIMA (4,1,5)



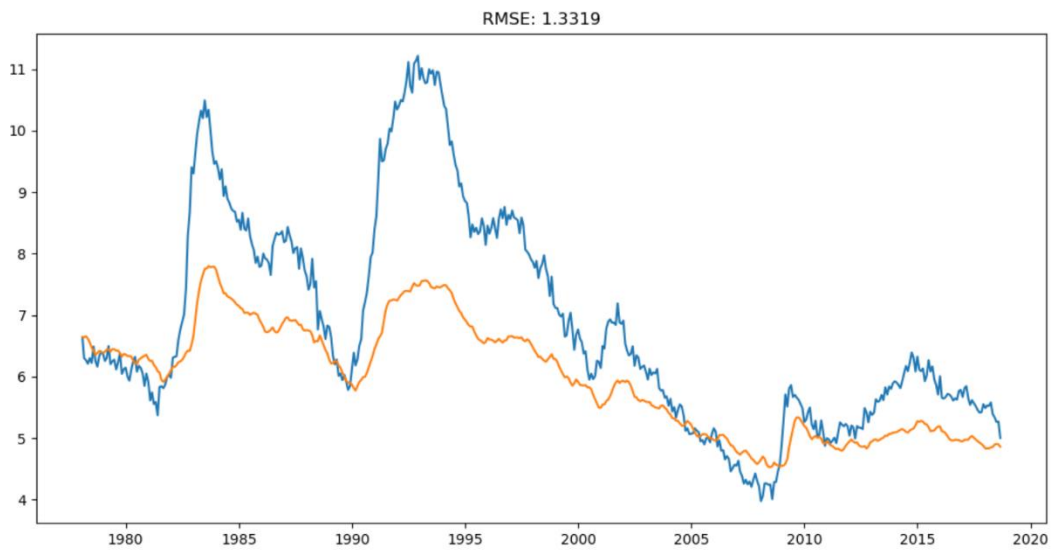
We see that model ARIMA (4,1,5) has the lowest RSS value so we will choose this model for our predictions.

- **Predictions and Accuracy**

Please find below some predicted values from the model:

```
date
1978-02-01    1.893833
1978-03-01    1.893026
1978-04-01    1.895635
1978-05-01    1.889727
1978-06-01    1.880194
dtype: float64
```

Below is the graph for comparison between the original and predicted values:



We see that the value of RMSE is 1.3319 this shows that the predictions are quite accurate, and our model is good.