ANALYZING AND FORECASTING CONSUMER PRICE INDEX TRENDS USING ADVANCED TIME SERIES TECHNIQUES

What is CPI?

The Consumer Price Index, or CPI, measures the change in prices of goods and services typically bought by households. It's a key indicator of inflation and the cost of living. To create the CPI, various data are collected and a range of methods are used, ensuring that the index is accurate and comparable across different regions.

Why is it important?

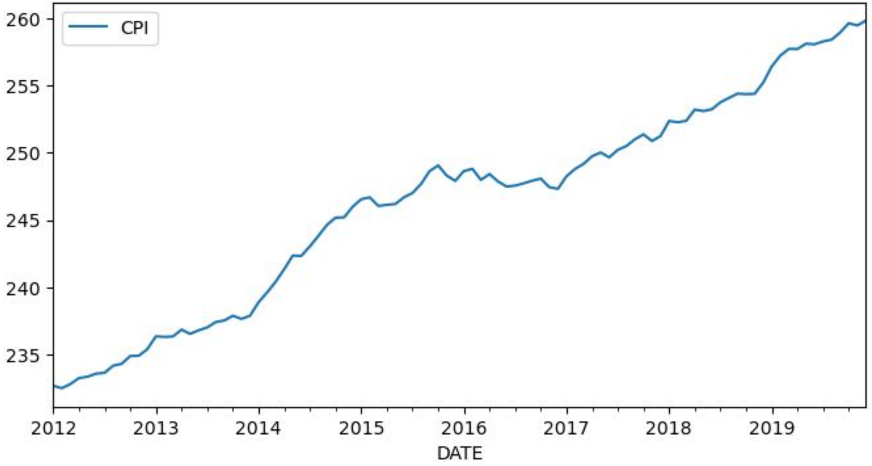
- **Inflation Tracking**: It's a principal measure of inflation, showing how prices change over time.
- **Economic Policy**: Governments use CPI to shape economic policy and adjust interest rates.
- **Wage Adjustments**: Employers may base wage increases on CPI to match living cost changes.
- **Purchasing Power**: It reflects the purchasing power of consumers' money over time.
- **Pension Updates:** Many pension plans use CPI to adjust benefits to maintain retirees' living standards.
- **Investment Insight:** Investors use CPI to inform strategies, as it can impact stock and bond markets.
- **Business Strategy:** Companies use it for pricing strategies and budget planning.
- **International Comparisons:** CPI allows for comparison of inflation rates between countries.

Raw Data

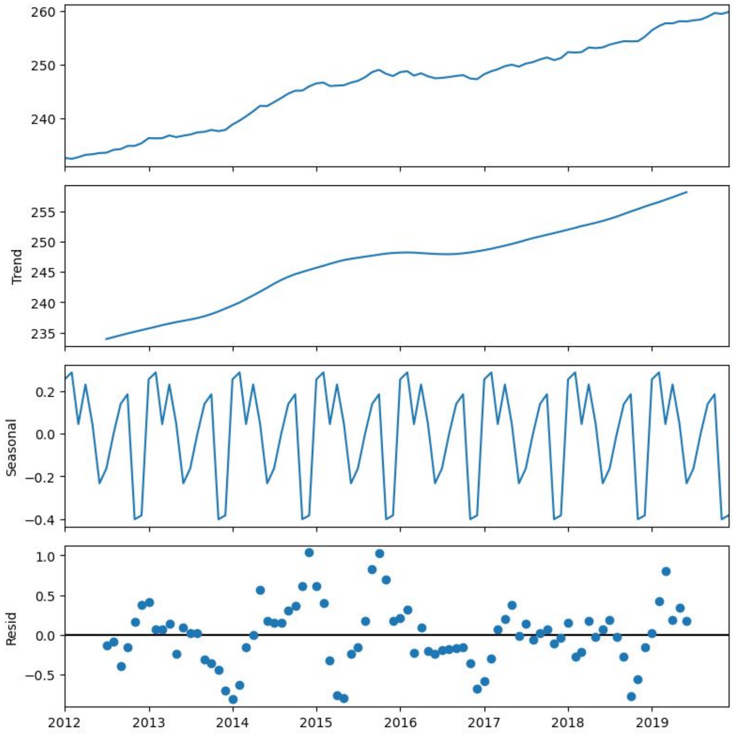
The series **CPIUFDNS** on the Federal Reserve Bank of St. Louis's FRED website represents the **Consumer Price Index for All Urban Consumers**: All Items in U.S. City Average, Not Seasonally Adjusted.

It **tracks the changes in prices** paid by urban consumers for a representative basket of goods and services over time.

Link - https://fred.stlouisfed.org/series/CPIUFDNS
Range - January 1, 2012, to December 1, 2019
Frequency: Monthly



Exploratory Analysis - Seasonal Decomposition



- **Consistent upward trend**, indicating inflationary pressures in the economy as consumer prices have generally increased.
- The seasonal component indicates that there are **predictable seasonal fluctuations** in the CPI, which could be leveraged for anticipatory economic measures.
- The residuals **do not show any apparent pattern**, suggesting that the model has effectively captured the trend and seasonality in the CPI data.

Train Test Split

95 data points representing the Consumer Price Index (CPI) was strategically divided:

1. **Training Data**: The initial **80 data points** form the training set, which corresponds to approximately **84%** of the total data.
2. **Testing Data:** The subsequent **15 data points** constitute the test set, making up about **16%** of the total data.

Examine and Forecast Using Different Detrending Approaches

1. Linear Regression
2. One-Sided Moving Average
3. Right-Sided moving average

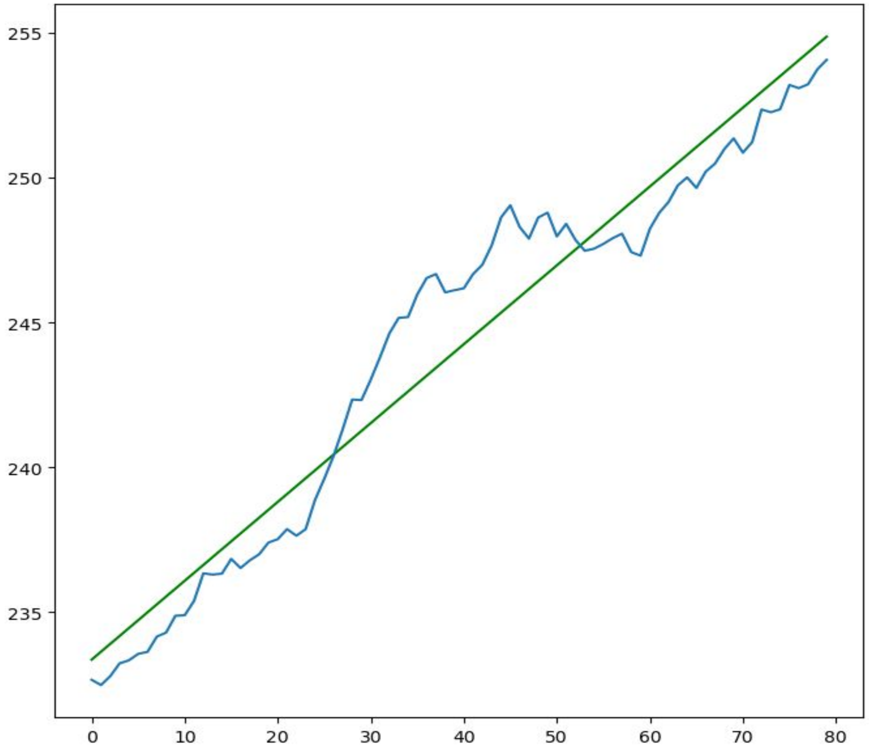
ADF Test on Original Data

```
Observations from Dickey-fuller test
Test Statistic                 -0.703888
p-value                         0.845740
#lags used                     12.000000
number of observations used    67.000000
critical value (1%)            -3.531955
critical value (5%)            -2.905755
critical value (10%)           -2.590357
dtype: float64
```

P value is 0.845 >>0.05 which means that the dataset is non-stationary

LINEAR REGRESSION

Detrending CPI Data Using Linear Regression Analysis



From the OLS regression results, we observe an **R-squared value of 0.941**, indicating that **approximately 94.1% of the variation in the CPI** can be explained by the time variable 't'.

This **high R-squared** value suggests a **strong fit of the model** to the data.

The linear regression model effectively **detrends the CPI data**, revealing a steady inflationary trend over the years.

```
train_data_df['Detrend'] = train_data_df['CPI'] - train_data_df['t'] * lr.coef_[0] - lr.intercept_
```

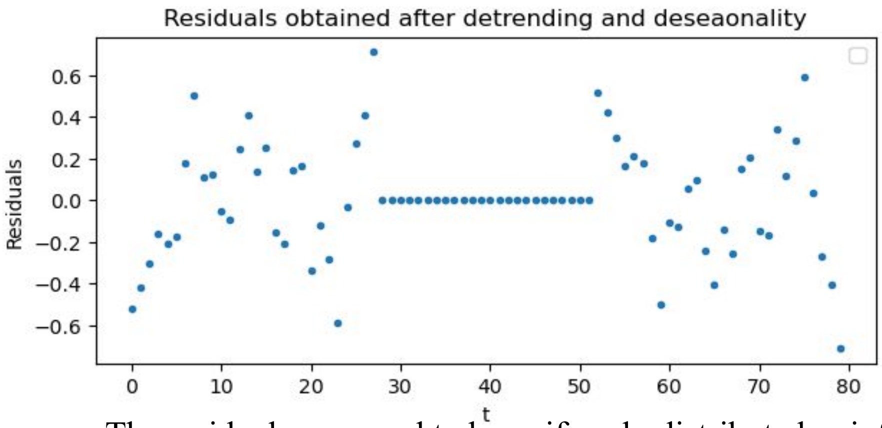
Deseasonality

The entire approach is based on the assumption that **data points exhibit a repeating seasonal pattern**.

The seasonal_mean function captures this pattern by **averaging data points from the same season across different years**.

```
seasonality = seasonal_mean(train_data_df['Detrend'])
seasonal_train = [ seasonality[i % 12] for i in range(len(train_data_df)) ]
```

Residuals and ACF Test



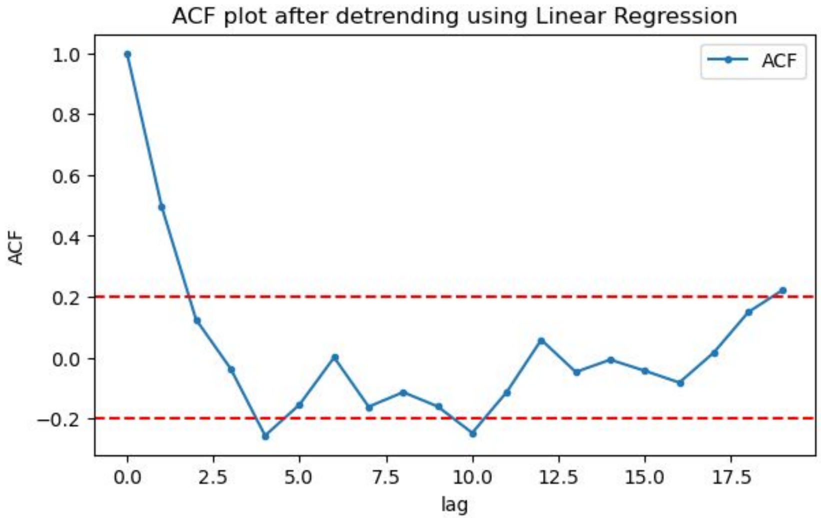

```
Observations from Dickey-fuller test
Test Statistic                  -3.683096
p-value                          0.004357
#lags used                      12.000000
number of observations used     67.000000
critical value (1%)             -3.531955
critical value (5%)             -2.905755
critical value (10%)            -2.590357
dtype: float64
```

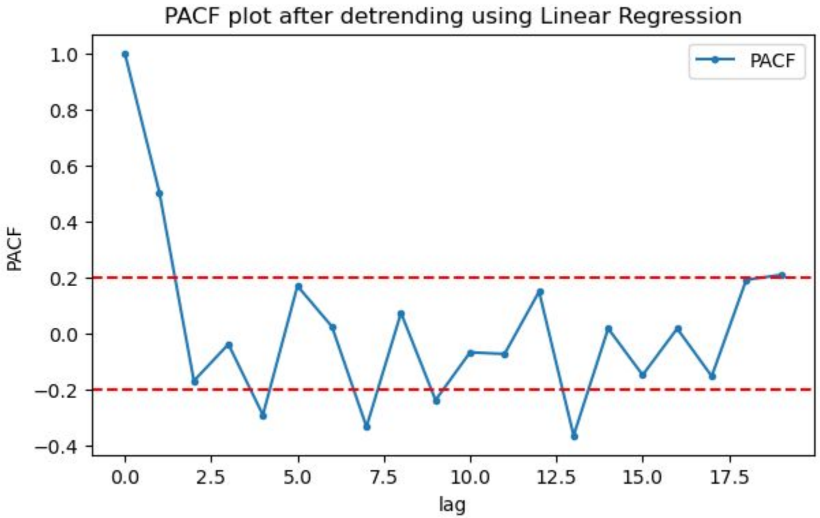
- The residuals appeared to be uniformly distributed, reinforcing the effectiveness of the detrending and deseasonalization.
- The Dickey-Fuller test results indicate that the time series data is stationary:
  - The test statistic of -3.683 is lower than the critical values for 1%, 5%, and 10% significance levels, suggesting a high level of confidence in the result.
  - A p-value of 0.0043 reinforces this confidence, showing a low probability that the data is non-stationary.

The data does not show patterns of trend or seasonality and is therefore suitable for reliable statistical analysis and forecasting.

```
train_data_df['Residual'] = train_data_df['Detrend'] - pd.Series(seasonal_train)
```

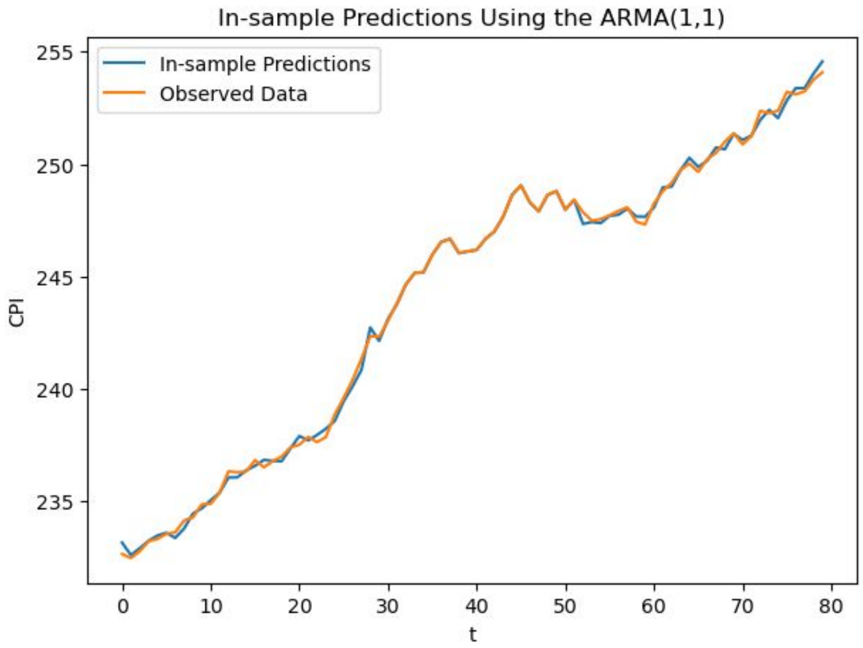
Select an ARMA Model Using ACF and PACF Plots



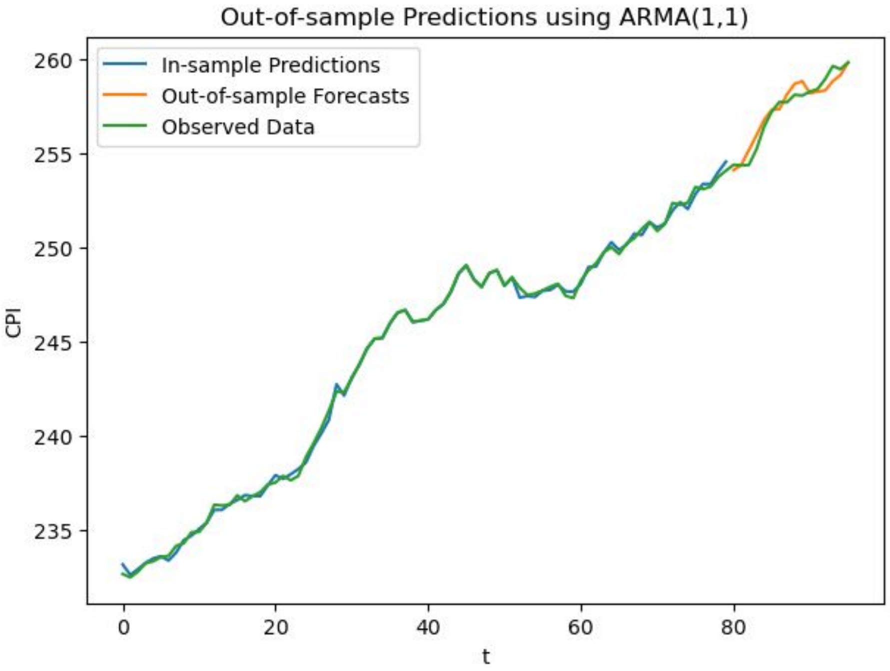


- Threshold Setting: The threshold of 0.2 helps us distinguish between significant and insignificant lags.
- Model Parameters: The strong initial spike in both plots leads us to choose an **ARMA(1,1)** model. This means the model considers one past data point and the error from the previous forecast in predicting the next value.

In-sample and Out-of-sample Predictions



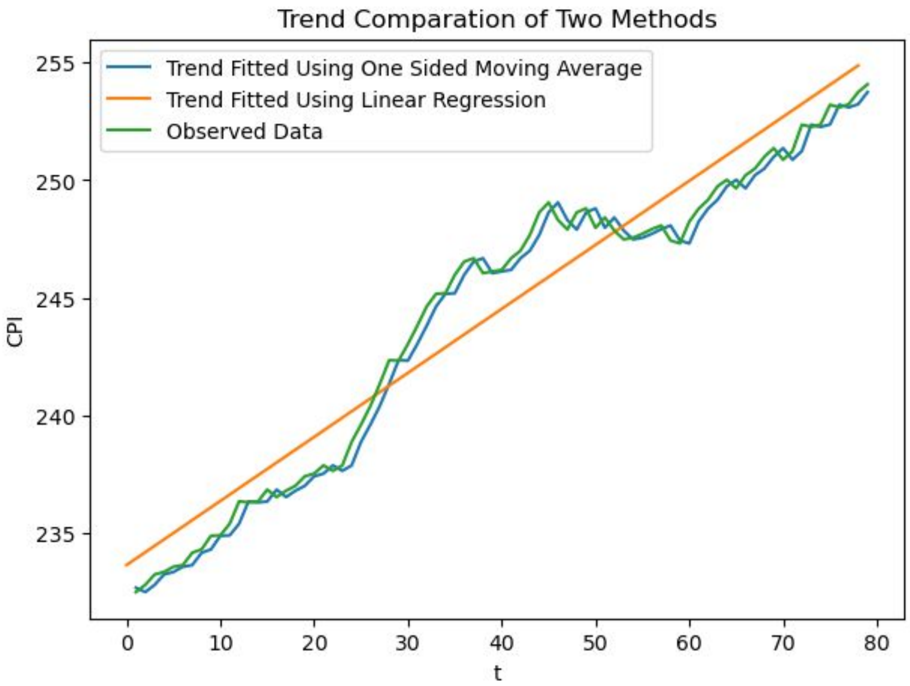
In-sample (Linear) RMSE: 0.209702639258
In-sample (Linear) MAE: 0.152691230205
In-sample (Linear) MSE: 0.043975196911

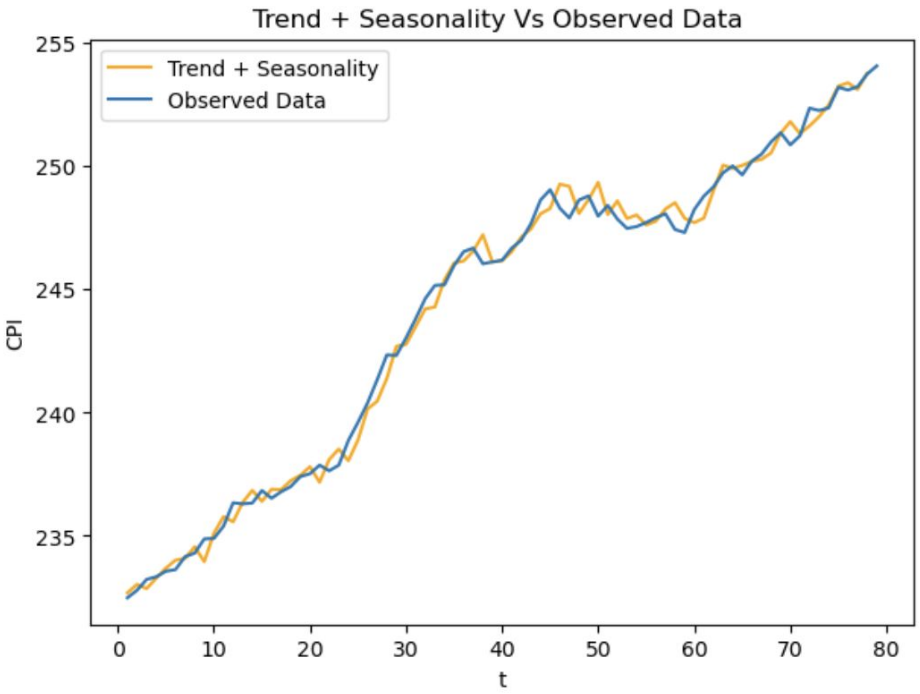


Out sample (Linear) RMSE: 0.485205840
Out sample (Linear) MAE: 0.398817482
Out sample (Linear) MSE: 0.23542470

ONE-SIDED MOVING AVERAGE

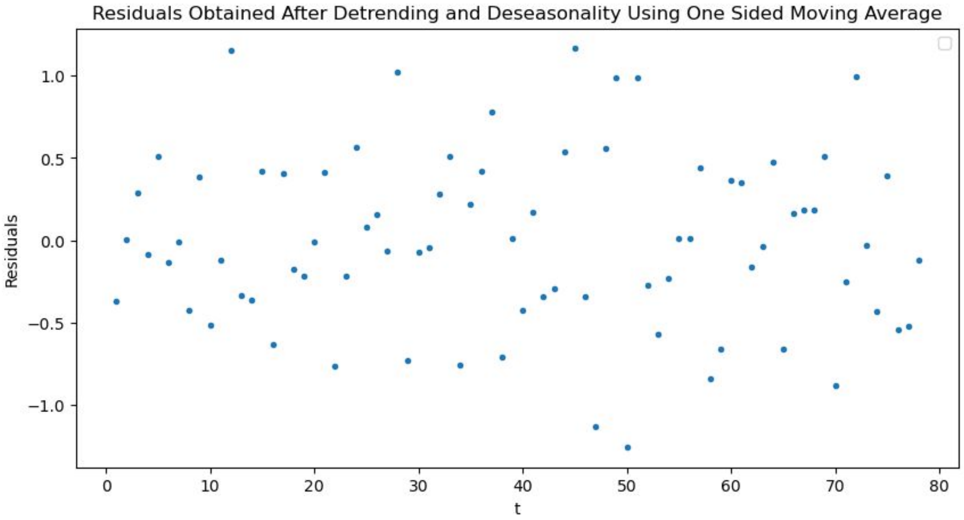
Trend Comparison





Much smoother trend using One-sided-Moving Average Approach vs Linear Regression Approach

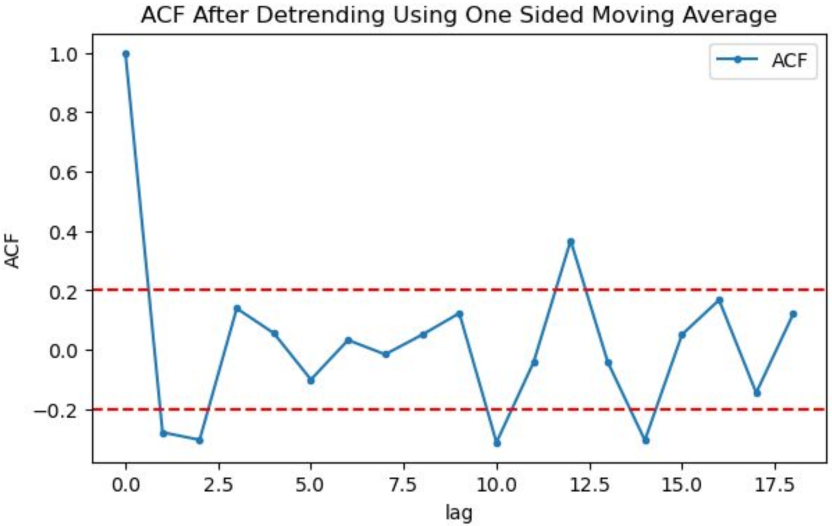
Residuals and ACF Test

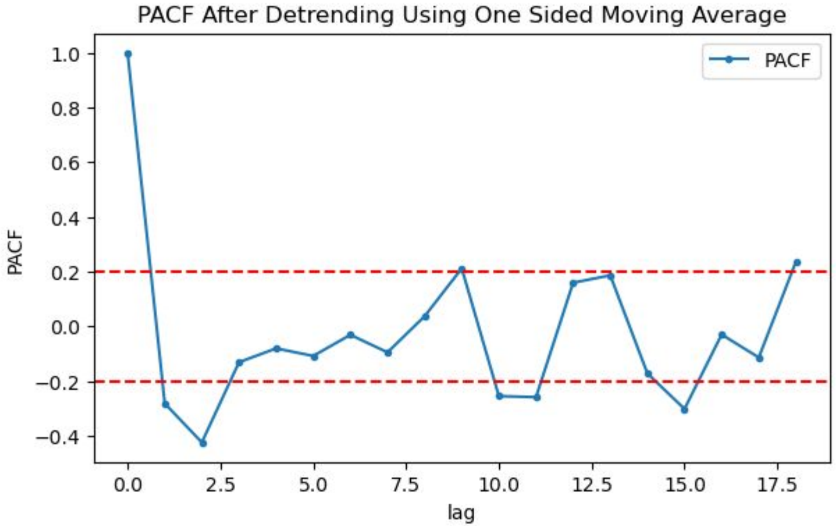



```
Observations from Dickey-fuller test
Test Statistic                 -1.064734e+01
p-value                         4.758143e-19
#lags used                      1.000000e+00
number of observations used     7.600000e+01
critical value (1%)            -3.519481e+00
critical value (5%)            -2.900395e+00
critical value (10%)           -2.587498e+00
dtype: float64
```

- Test Statistic: -10.64734, significantly **more negative**, suggesting stronger evidence of stationarity.
- p-value: Approximately 4.76e-19, **an extremely low probability** of non-stationarity.
- Lags Used: Reduced to 1, implying less complexity in achieving stationarity.

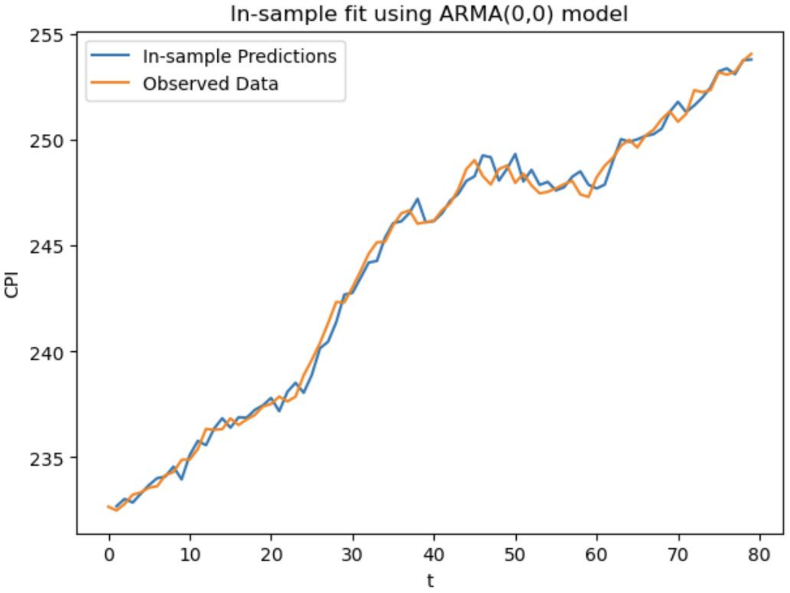
Select an ARMA Model Using ACF and PACF Plots

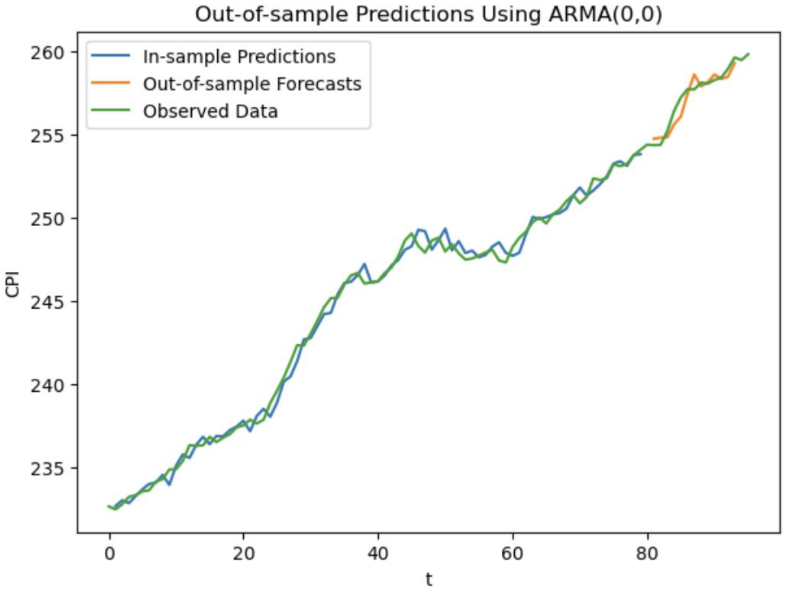




ARMA(0,0)

In-sample and Out-of-sample Predictions





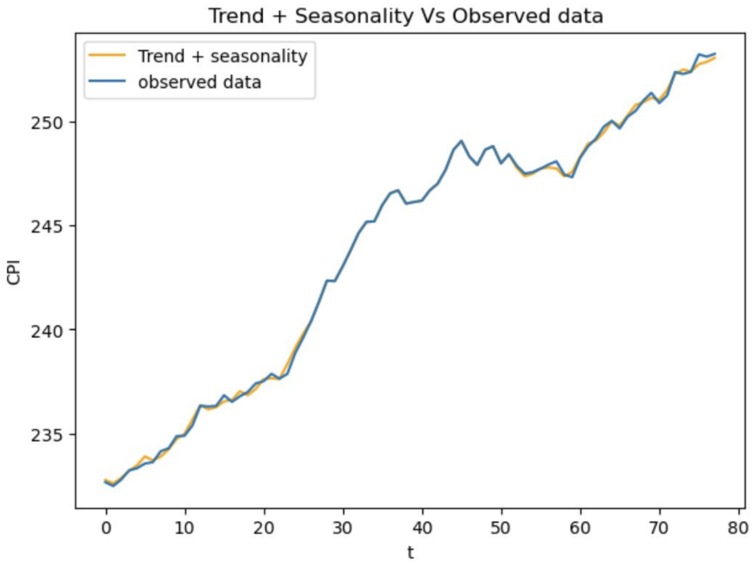

```
In-sample (One Sided Moving Average) RMSE: 0.55686
In-sample (One Sided Moving Average) MAE: 0.4390486
In-sample (One Sided Moving Average) MSE: 0.3100937
```

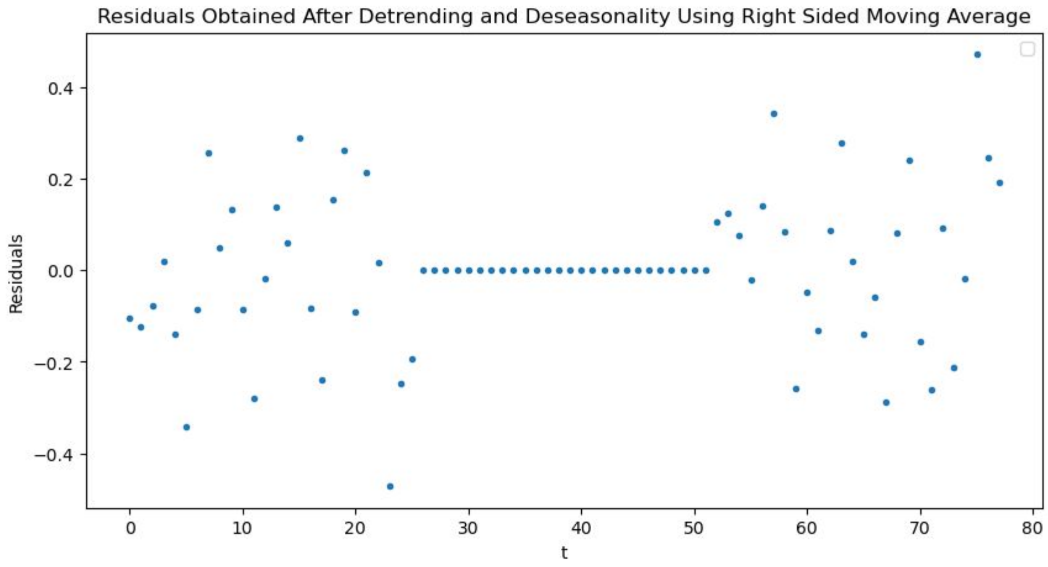
```
Out sample (One Sided Moving Average) RMSE: 0.520398
Out sample (One Sided Moving Average) MAE: 0.3986358
Out sample (One Sided Moving Average) MSE: 0.2708148
```

RIGHT-SIDED MOVING AVERAGE

Overview

- This approach uses a **window of 3 data points** to calculate a **forward-looking moving average** for the Consumer Price Index (CPI).
- Instead of using past data, this method averages each data point with the **next two future values**, offering a unique 'right-sided' perspective.
- The original CPI data is adjusted by subtracting this right-sided moving average, effectively removing the trend that includes near-future information.
- This method allows for **retrospective analysis** where **all data points are known**.
- This technique can r**eveal patterns and anomalies not easily identified** through traditional detrending methods.

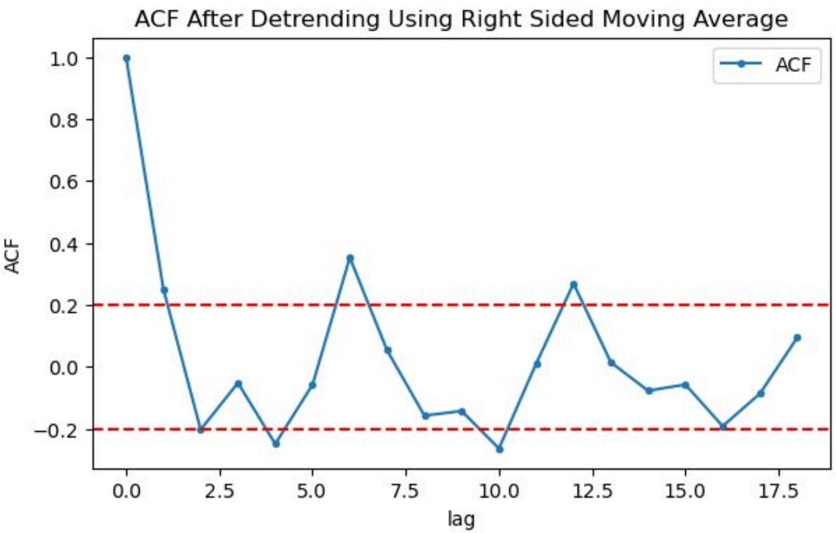


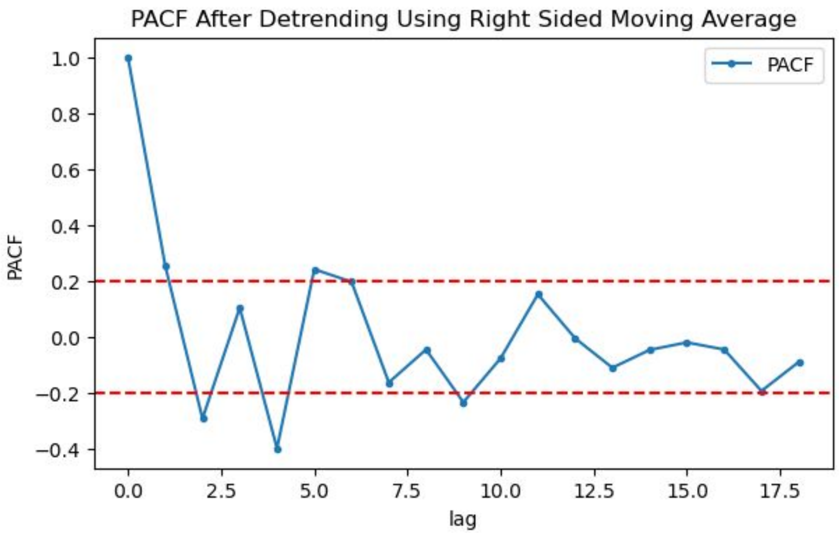


'**Trend plus Seasonality**' is runs very close to observed data.

The plot shows randomness, indicating the trend was captured and removed effectively.

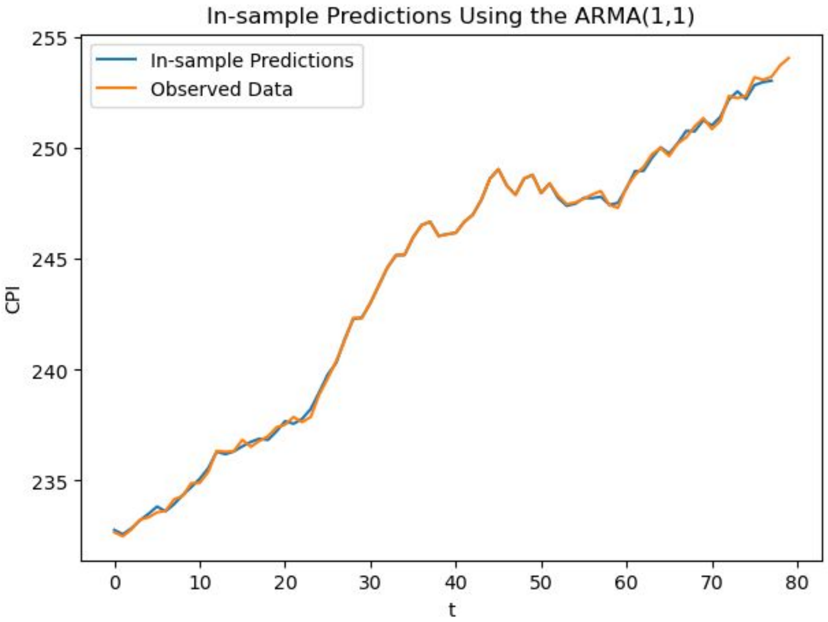
Select an ARMA Model Using ACF and PACF Plots

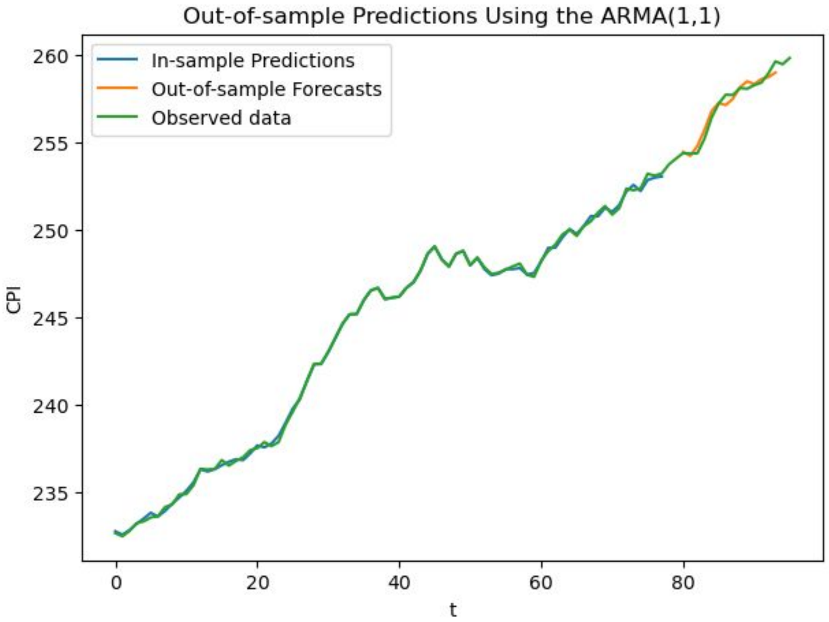




The ACF and PACF plots reveal a noticeable initial decline and subsequent fluctuations within the confidence interval, leading to the selection of an **ARMA(1,1) model.** This model, incorporating both autoregressive and moving average components, is well-suited to **represent the observed data correlation at lag one.**

In-sample and Out-of-sample Predictions






```
In-sample (Right Sided Moving Average) RMSE: 0.1431522
In-sample (Right Sided Moving Average) MAE: 0.1035195
In-sample (Right Sided Moving Average) MSE: 0.02049256
```

```
Out sample (Right Sided Moving Average) RMSE: 0.3451876
Out sample (Right Sided Moving Average) MAE: 0.27847594
Out sample (Right Sided Moving Average) MSE: 0.1191545
```

Conclusion

Here is a comprehensive approach taken to analyze the Consumer Price Index (CPI) data

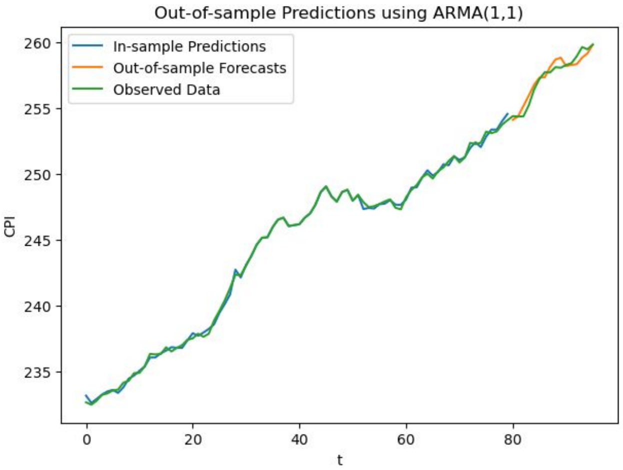
1. **Detrend the CPI Data**: Apply linear regression analysis to remove the trend, isolating the underlying pattern of the Consumer Price Index.
2. **Deseasonalize the Data**: Calculate and remove seasonal effects to focus on the non-seasonal fluctuations.
3. **Analyze Residuals**: Evaluate the residuals obtained after detrending and deseasonalizing to ensure that only random noise remains.
4. **Conduct Stationarity Testing:** Use the Dickey-Fuller test to confirm that the residual series is stationary, indicating suitability for further analysis.
5. **Select an ARMA Model**: Based on ACF and PACF plots, identify the appropriate Autoregressive Moving Average (ARMA) model parameters.

Conclusion

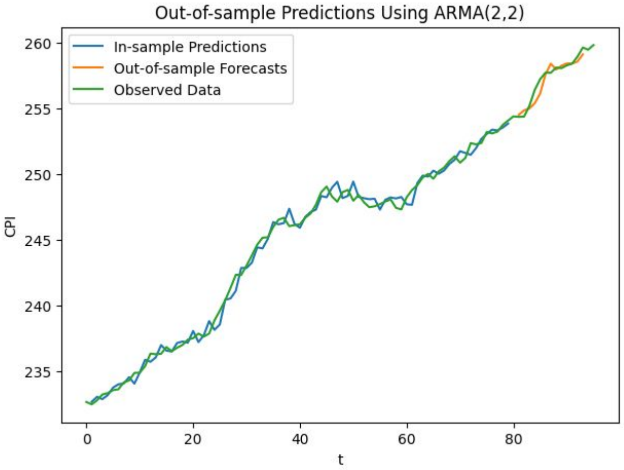
6. **Train the ARMA Model:** Fit the ARMA model to the detrended and deseasonalized data and assess its in-sample forecasting accuracy.
7. **Perform Out-of-Sample Forecasting:** Generate and evaluate out-of-sample CPI forecasts to test the model's generalizability.
8. **Improve with Alternative Detrending**: Implement a one-sided moving average approach for a different detrending method and forecast based on this model.
9. **Integrate Trend and Seasonality**: Reconstruct the CPI data by combining the trend with seasonal components and compare with the observed data for validation.
10. **Refine with Right-Sided Moving Average**: Explore the right-sided moving average as an alternative detrending technique and assess its forecasting performance.
11. **Compare the forecasts** from linear regression, one-sided, and right-sided moving average detrending methods to conclude the most effective approach for CPI analysis and forecasting.

The forecast using the right-sided moving average method turned out to be more accurate than the one-sided moving average or linear regression methods.

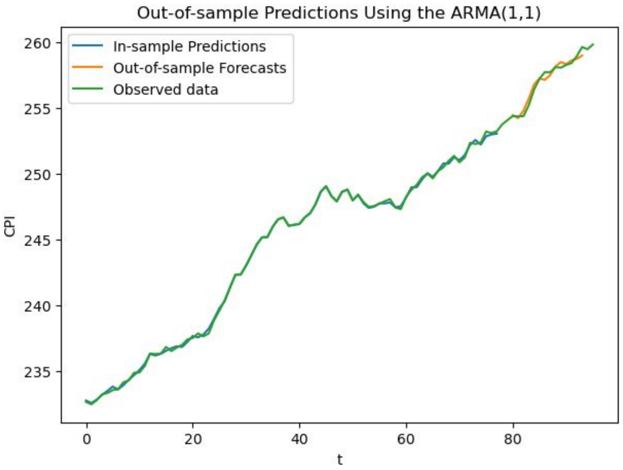
Linear Regression



One-sided Moving Average



Right-sided Moving Average



THANK YOU!