Electricity Demand Forecasting Results

_____

Some General colors used in every graph:
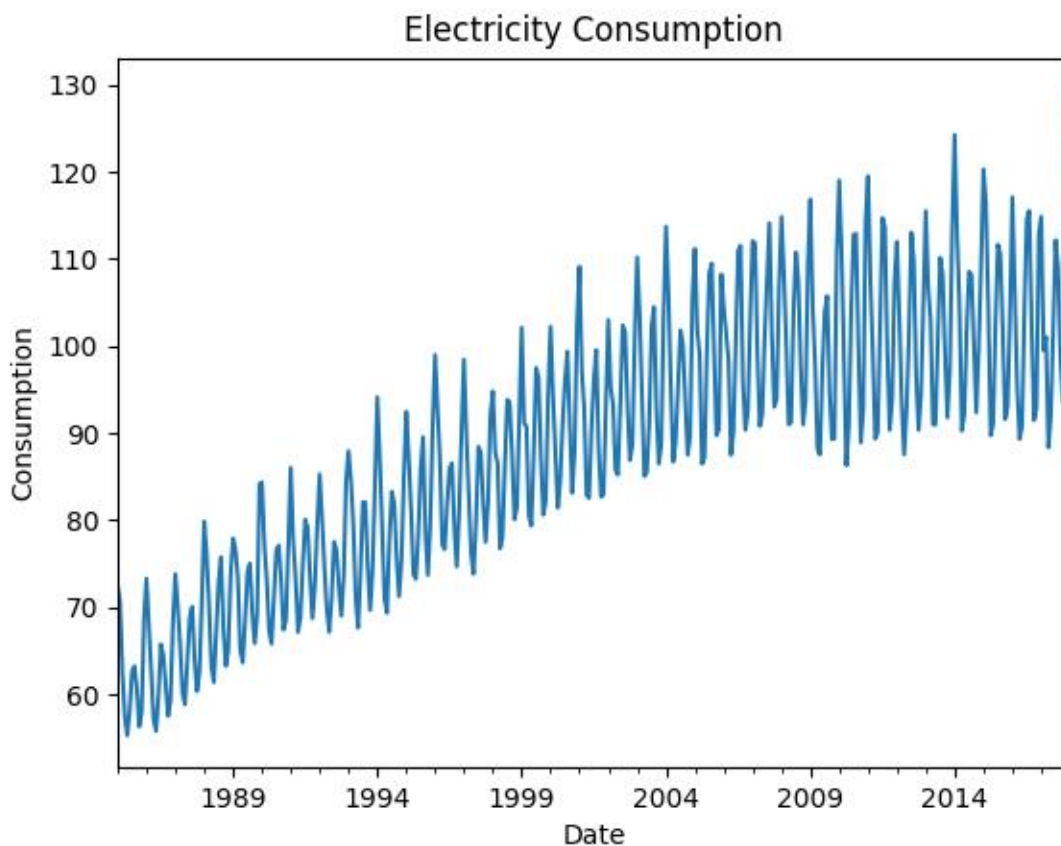
- ❖ Red: Trend
- ❖ Blue: Training data
- ❖ Green: Future predictions

_____

1. Data Analyses

Data Source: https://www.kaggle.com/code/nageshsingh/predict-electricity-consumption/data

We have monthly electricity consumption data from **1985** to **2017.**

Data Plot:

Forecasting Approach used: For Time series forecasting we can use different models and methods. Given electricity data has **seasonality**, so my approach for forecasting is to separate out seasonal component and then using **ARIMA** model for forecasting.
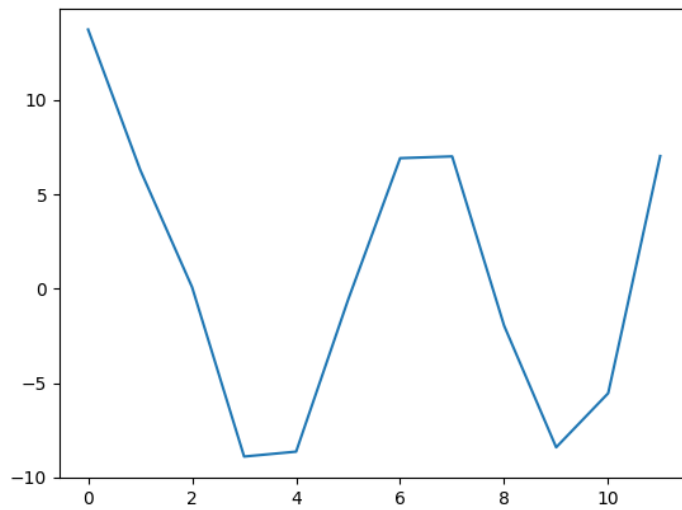
Actual data= Trend + Seasonality + Residual error

Clearly Data has year seasonality.

Trend: Data has seasonality with **frequency=12,** So trend can be found by calculating Moving Average 12

(MA-12). MA(12) for data is as follows(MA-12 is in red):
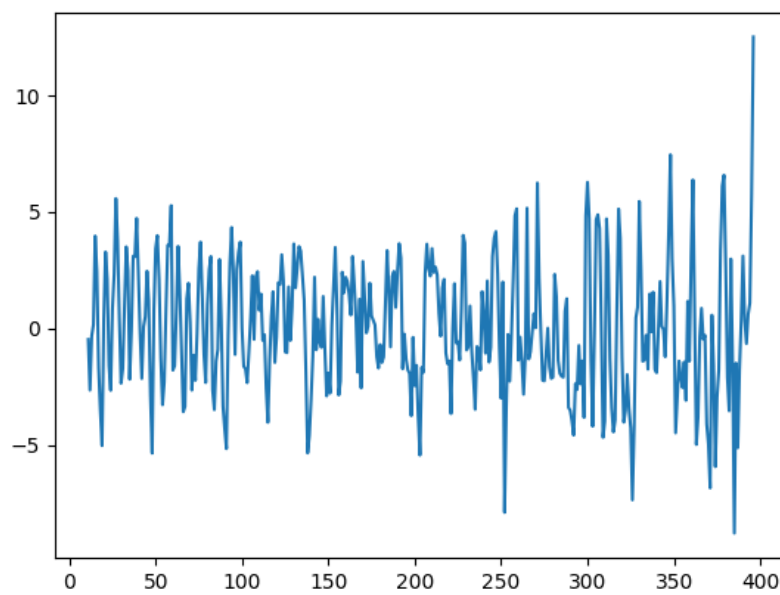

Rolling Mean of Data
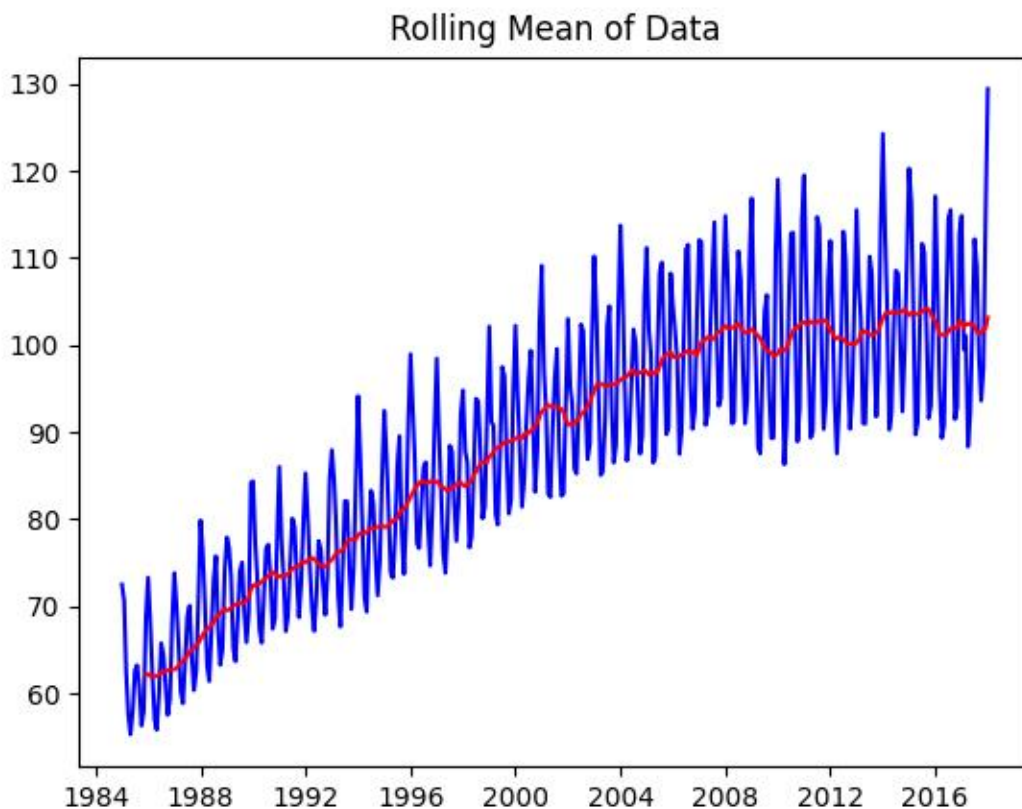
Seasonality: Data has seasonality with **frequency=12**. Graph for one cycle is as follows:

Residual Errors:

Residual Error plot for data:
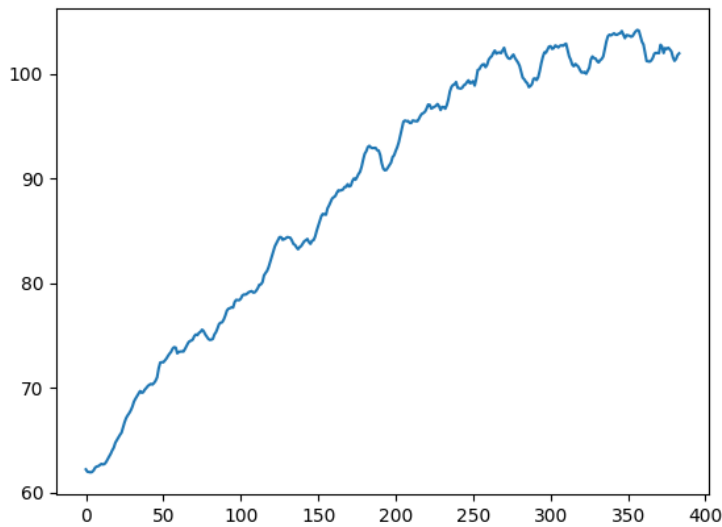
ADF Test on residual errors:

```
Results of dickey fuller test
--------------------------------
- For Residual Errors
ADF Statistic: -4.169110
p-value: 0.000743
Critical Values:
        1%: -3.448
        5%: -2.869
        10%: -2.571
```

Residual errors data is stationary.

2. Model Fitting

- **Model Fitting for Trend**:

  Data plot for trend:

  Applying ADF Test on difference 1 and difference 2 data for finding d value for ARIMA model.

  Finally we get d=1. ADF Test results are on next page:

ADF Test Results:

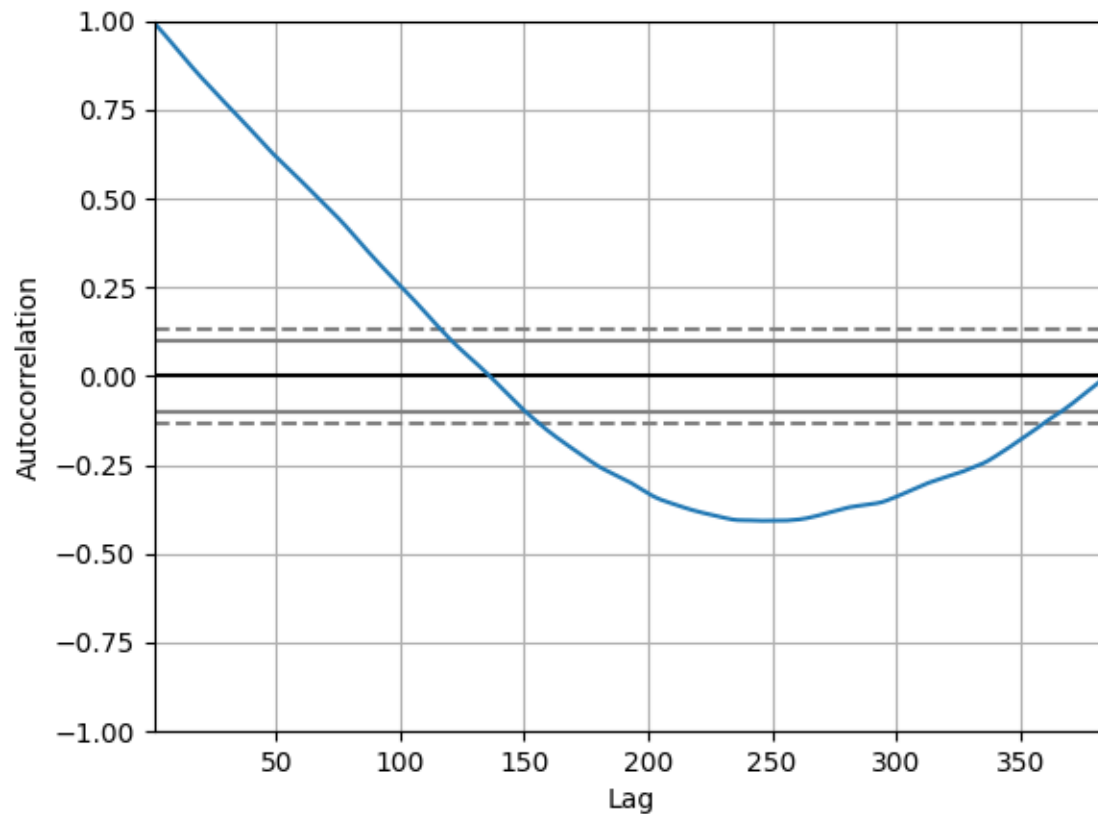```

- For d=1
ADF Statistic: -5.667091
p-value: 0.000001
Critical Values:
        1%: -3.448
        5%: -2.869
        10%: -2.571

- For d=2
ADF Statistic: -8.030912
p-value: 0.000000
Critical Values:
        1%: -3.448
        5%: -2.869
        10%: -2.571
```

PACF Plot for trend:



Partial Autocorrelation
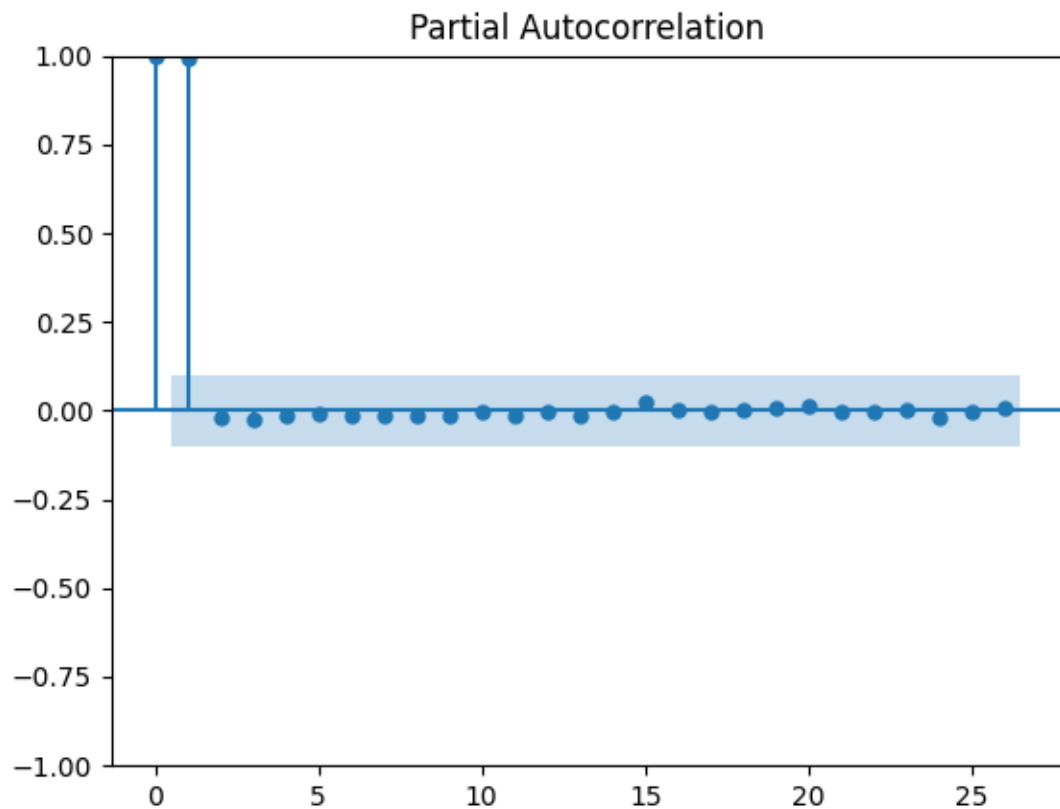
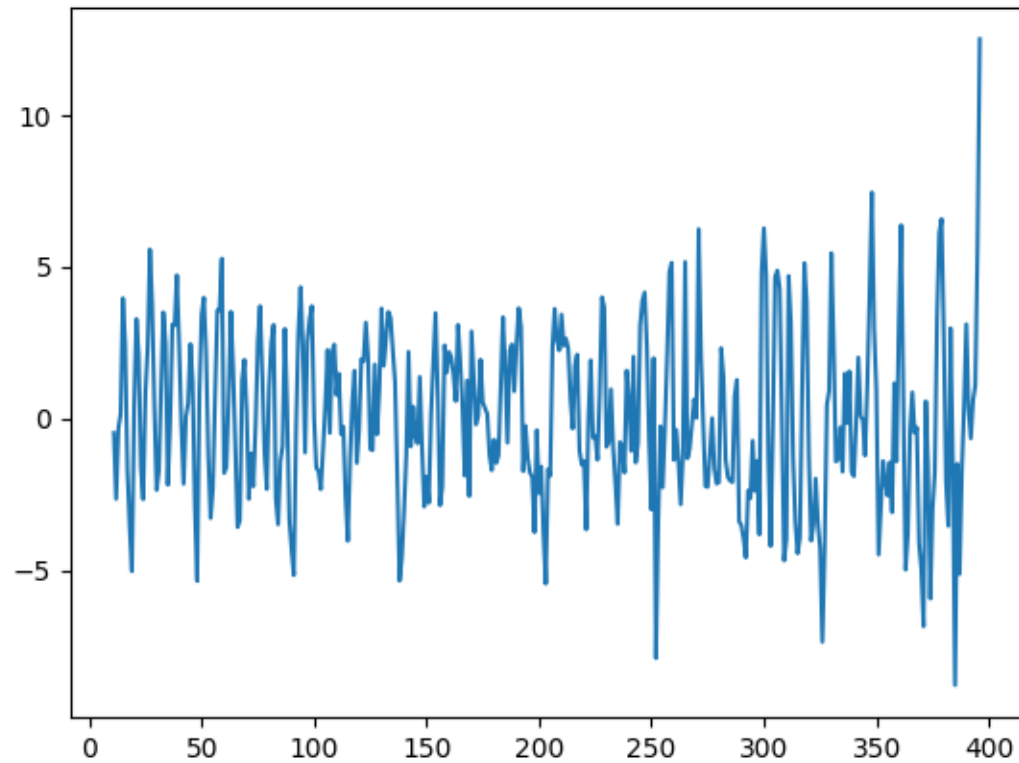So, q=2.

Autocorrelation plot for Trend:



p=136 from autocorrelation plot. But this is very high value program will take about 1 day to predict values or even more than this. So we tested for p=0 to 7. Results are as follows:

```
ARIMA(0, 1, 2) RMSE=0.287
ARIMA(0, 2, 2) RMSE=0.295
ARIMA(1, 1, 2) RMSE=0.288
ARIMA(1, 2, 2) RMSE=0.290
ARIMA(2, 1, 2) RMSE=0.288
ARIMA(2, 2, 2) RMSE=0.286
ARIMA(3, 1, 2) RMSE=0.289
ARIMA(3, 2, 2) RMSE=0.289
ARIMA(4, 1, 2) RMSE=0.288
ARIMA(4, 2, 2) RMSE=0.286
ARIMA(5, 1, 2) RMSE=0.279
ARIMA(5, 2, 2) RMSE=0.286
ARIMA(6, 1, 2) RMSE=0.293
ARIMA(6, 2, 2) RMSE=0.288
ARIMA(7, 1, 2) RMSE=0.289
ARIMA(7, 2, 2) RMSE=0.290
Best ARIMA(5, 1, 2) RMSE=0.279
```
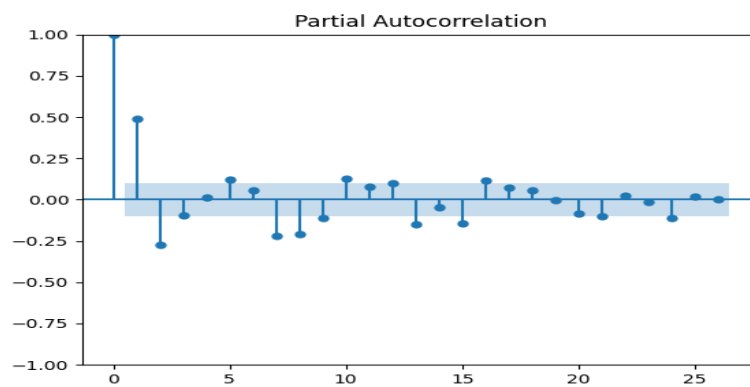
But we can't finalize the model now.

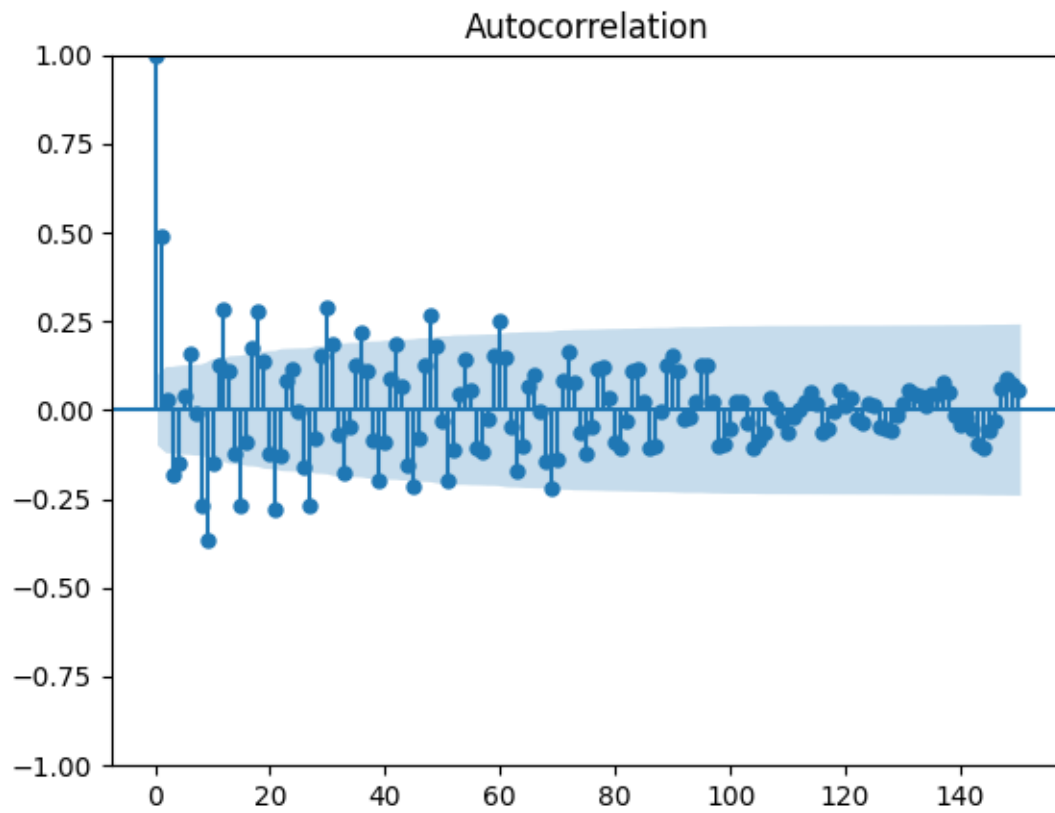- ## **Model Fitting for Residual Error:**
  Data Plot:



Data is already stationary as checked by ADF Test. So, d=0.

PACF Plot:

So, q=2.

Autocorrelation plot :



p=3 from the ACF plot.

# FINAL FORECASTING MODEL Training:

Final forecasting results:

ARIMA(0,1,2) for trend and ARIMA(4,0,2) for Residual Errors fitted best with

**RMSE: 3.3994531027484345.**

**Table 1:** from **Jan-2017** to **Dec-2017**

| Actual Value | Prediction | % Error |
|--------------|------------|---------|
| 114.85 | 118.444 | 3.13 |
| 99.49 | 110.137 | 10.7 |
| 101.04 | 102.176 | 1.12 |
| 88.353 | 92.104 | 4.24 |
| 92.49 | 93.175 | 1.19 |
| 102.153 | 103.04 | 0.87 |
| 112.154 | 111.711 | -0.395 |
| 108.931 | 111.182 | 2.07 |
| 98.615 | 100.435 | 1.84 |
| 93.614 | 92.802 | -0.87 |
| 97.336 | 96.293 | -1.07 |
| 114.721 | 110.572 | -3.62 |

**Graph: Predictions and Analysis:**
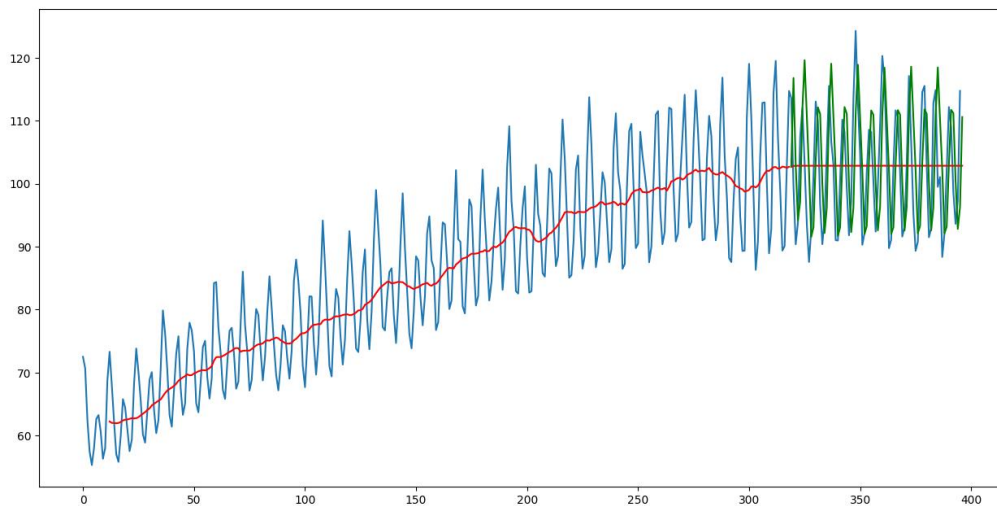
## Future Predictions with finalized model: