12/14/18 EC 245 Professor Grubaugh

Boston House Price Index

A Time Series Analysis

Andrew Jaroncyk

INTRODUCTION

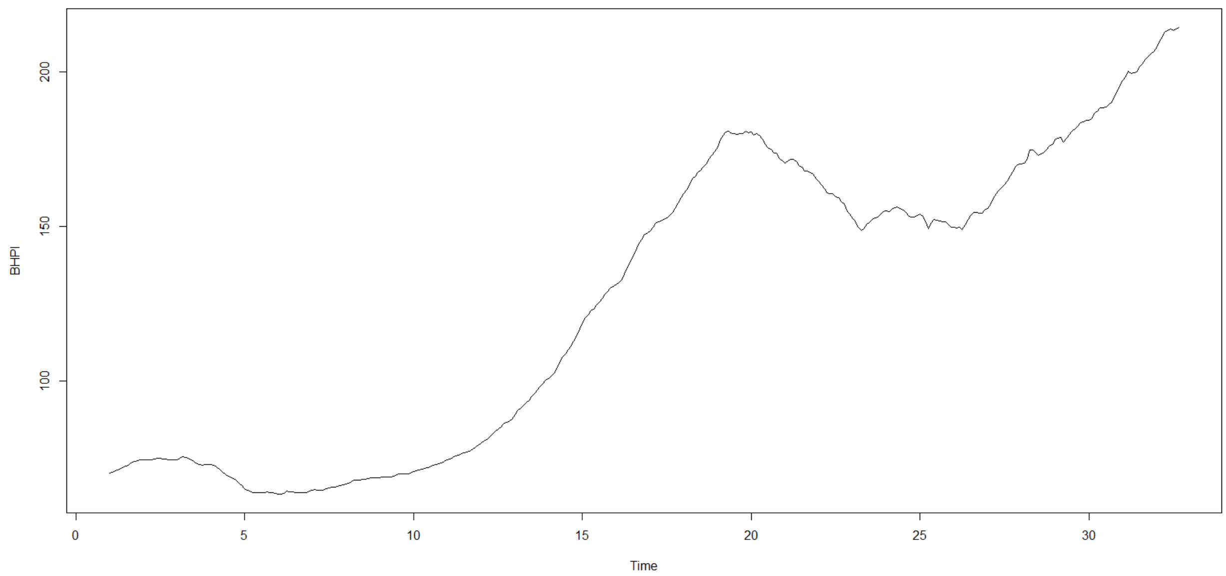
This report intends to analyze the relationship between the Standard & Poor (S&P) CoreLogic Case-Shiller Boston Home Price NSA Index measure over the time periods of January 1, 1987 to September 1, 2018. In addition, the relationship of the S&P CoreLogic Case-Shiller Home Price Index Series, an index measuring the Home Price Index across all single-family housing stock across the United States, will also be analyzed within this specific timeframe. A time-series analysis will be conducted to further understand how the Boston Home Price Index and the United States Home Price Index have changed over the course of time through using the open-source statistical scripting language R. Data visualization will be used extensively to further analyze the Indexes over the course of time, and moving average and autoregressive models will be contrasted to accurately explain the relationship of these Indexes over time.

BACKGROUND INFORMATION

The Standard & Poor (S&P) CoreLogic Case-Shiller Home Price NSA Index is a measure of the average monthly change in the value of residential real estate in a given location assuming a constant level of quality. The area of interest that will be considered in this report is the Boston, Massachusetts area. Assume that for the analysis portion the specified Confidence Level is 95%, so $\alpha = 0.05$.

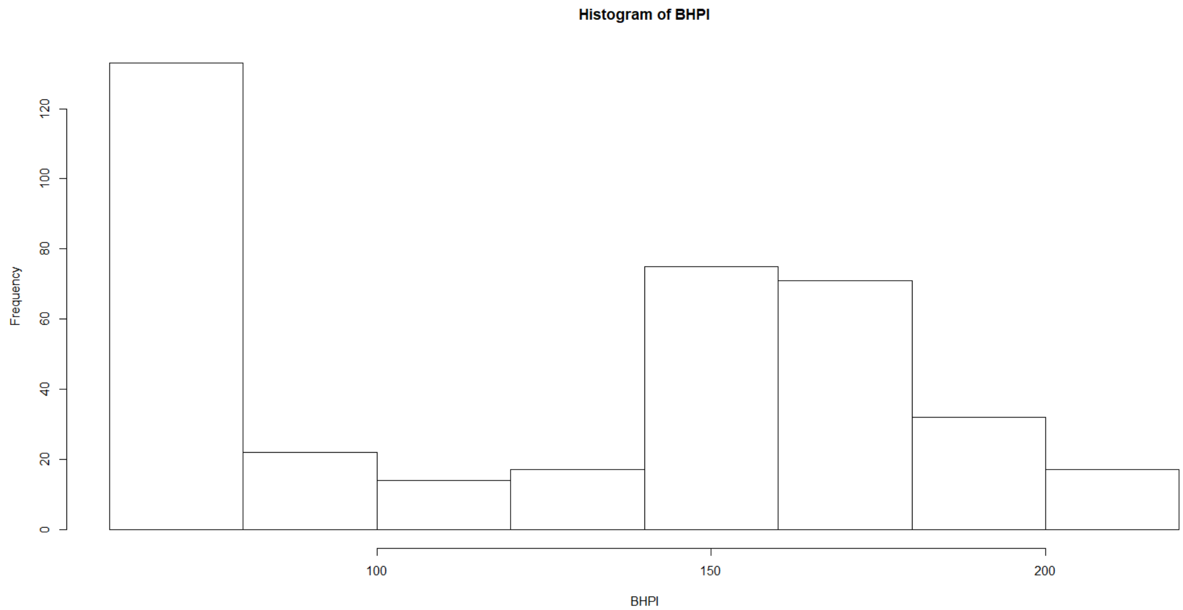
ANALYSIS

A time-series plot of the Boston House Price Index is shown below, where the values of the Boston House Price Index are measured on a monthly basis from January 1987 to September 2018.



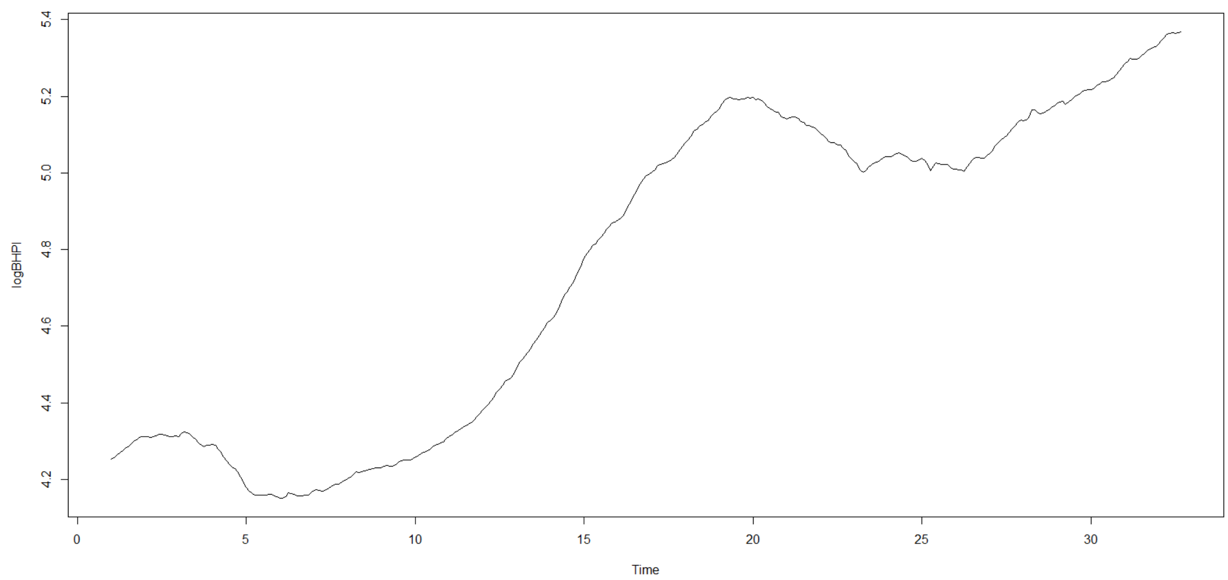
An initial analysis of the time-series plot suggests that over the course of time between January 1987 and September 2018, the Boston House Price Index has steadily increased, with a drop off in the value of the Boston House Price Index occurring somewhere approximately in the 1995 – 1998 range. During the 1990s in Boston, there appeared to be an increase in price volatility and sharp differences in prices behavior, resulting in a very large housing price boom in the Boston housing market. To add to the housing price boom, the United States economy was in a recession period during the 1990s. Once the 1990s concluded and Year 2K commenced, the Boston House Price Index stabilized and has increased since, signifying price restoration and reduced price volatility in the Boston housing market.

A histogram of the Boston House Price Index values is also provided to gain a better sense of the distribution.



The histogram above indicates that the data is very much negatively skewed, with the vast majority of its values lying in the 0 to 100 range. Transformations to the data must be performed in order to normalize the data and make it more suitable for further analysis.

After taking the natural logarithm (ln) of the Boston House Price Index data, the data is plotted again. Note that the shape of the graph remains exactly the same as the time-series plot of the original data, but the scale on the y-axis has been altered to reflect the values of the Boston House Price Index on the natural logarithm's domain, extending from 0 to positive infinity (+∞).



A quick observation of the adjusted time-series plot above corroborates that over the course of time in the Boston housing market, the Boston House Price Index has steadily increased, with a drop off occurring in the 1990s due to local housing market and macroeconomic conditions.

The first difference of the natural logarithm of the Boston House Price Index must be taken in order to account for in order to remove the effects of lags that occur in the dataset provided. In essence, the first difference of the natural logarithm of the Boston House Price Index produces the growth rate of the Boston House Price Index and expresses this rate in decimal format.

Summary statistics and the standard deviation of the growth rate are provided.

```
> summary(diffBHPI)
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.0167246 -0.0009038  0.0034877  0.0029352  0.0074436  0.0188182
> sd(diffBHPI)
[1] 0.006328069
```

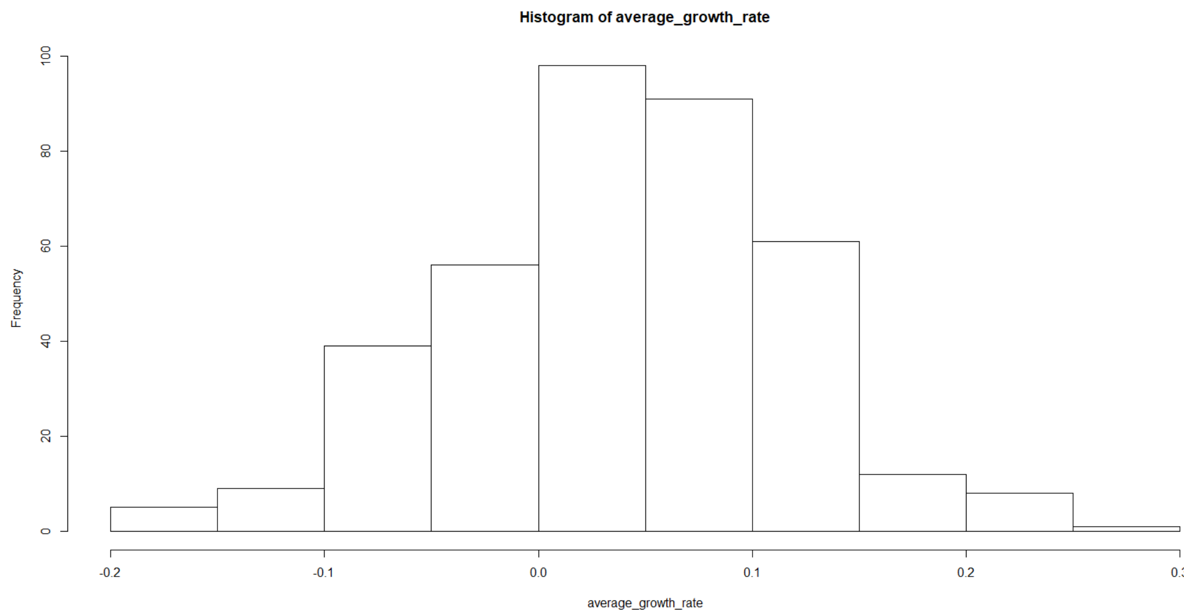
With the growth rate over this period of time averaging out to 0.0029, or about 0.29%, this shows that the Boston House Price Index that, on average, over this course of time from January 1978 to September 2018, the average growth in the Boston House Price Index is 0.29%. The standard deviation of 0.0063, or about 0.63%, suggests that the growth rate of the Boston House Price Index deviates from the mean value of 0.29% about 0.63 percentage points, which is a relatively small level of variability over this length of time, signifying security in the Boston housing market.

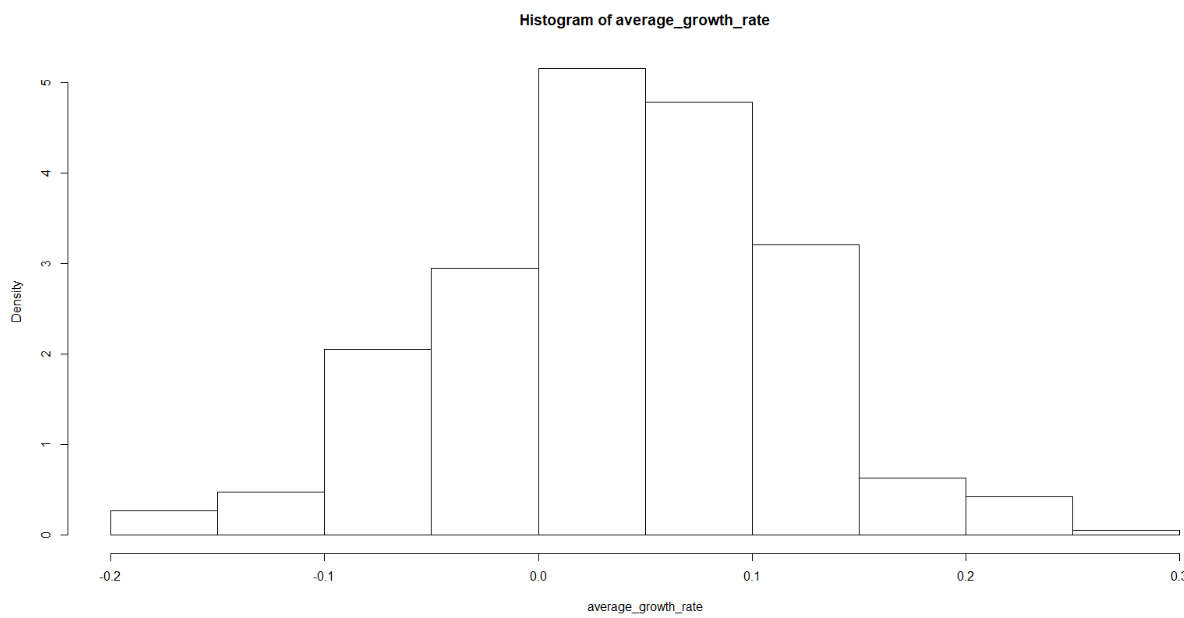
A more accurate representation of the growth rate over this extensive period of time is to determine the average growth rate, which can be found using the following formula:

$$\text{Average Growth Rate} = (1 + \ln BHPI)^{12} - 1$$

Summary statistics for the average growth rate of the Boston House Price Index is provided along with a histogram and density histogram analyzing the distribution of the Boston House Price Index.

```
> summary(average_growth_rate)
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max. 
-0.18323 -0.01079  0.04266  0.03849  0.09307  0.25072 
```





Analyzing the summary statistics, the average growth rate of the Boston House Price Index over the course of time from January 1978 to September 2018 is 0.038, or about 3.8%, with a standard deviation of 0.078, or about 7.8%. The histograms show that the distribution of the data appears to be approximately Normal, but the argument can be made that the data is skewed negatively ever so slightly. The average growth rate appears to fall somewhere between 0.00 and 0.15%.

Now, a 1-Sample t-Test must be conducted to determine whether the average growth rate is different from the value 0:

$$H_0: \mu = 0$$

$$H_A: \mu \neq 0$$

```
> t.test(average_growth_rate, mu = 0)

        One Sample t-test

data:  average_growth_rate
t = 9.6441, df = 379, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.03064617 0.04634269
sample estimates:
 mean of x
0.03849443
```

Observing the results from the 1-Sample t-Test, the p-value of $2.2e^{-16}$ suggests that at the 95% Confidence level (which will be the Confidence Level that is assumed going forward) that the average growth rate differs from 0.

Say we would like to determine the average growth rate between January 2012 and January 2018, which is more recent data, and would also like to conduct another 1-Sample t-Test to determine if the average growth rate during this 6-year span differs from the value of 0.

```
> t.test(recent_average_growth_rate, mu = 0)

        One Sample t-test

data:  recent_average_growth_rate
t = 8.8584, df = 80, p-value = 1.668e-13
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.04323940 0.06829634
sample estimates:
 mean of x
0.05576787
```

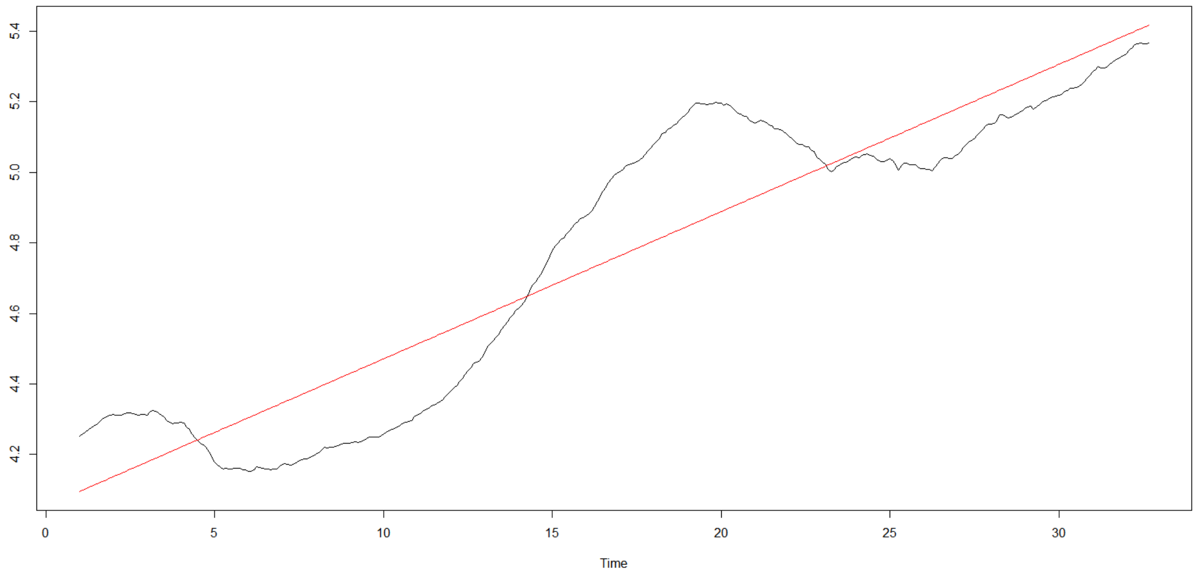
An analysis of this particular 1-Sample t-Tests concludes that the average growth rate during this 6-year time period differs from 0.

For future analysis of this time-series variable, the proper R libraries must be loaded to allow for such analysis.

Now we wish to produce a model that predicts the value of BHPI using the fact that the data has a trend associated to it. A trend in time-series analysis is defined to be a smooth and slow evaluation of the data that appears to be headed in one direction over the course of time. The time series plot of the Boston House Price Index suggests that from January 1978 to September 2018, there appears to be a positive trend of the data, whereas time progresses, the value of the Boston House Price Index increases over time.

```
> summary(model)

Time series regression with "ts" data:
Start = 1(1), End = 32(9)

Call:
dynlm(formula = BHPI ~ trend(BHPI))

Residuals:
   Min     1Q Median     3Q    Max 
-23.61 -15.66  -5.82  14.55  42.52 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  48.7273     1.8694   26.07   <2e-16
trend(BHPI)   4.8679     0.1018   47.83   <2e-16

(Intercept) ***
trend(BHPI) ***
---
Signif. codes:  
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 18.21 on 379 degrees of freedom
Multiple R-squared:  0.8579,	Adjusted R-squared:  0.8575 
F-statistic:  2287 on 1 and 379 DF,  p-value: < 2.2e-16
```

The output above shows that the trend coefficient has a value of 4.8679, meaning that for a one-year increase in the trend(BHPI) variable, on average, the BHPI increases by 4.8679 points. The trend(BHPI) coefficient is statistically significant at the 5% significance level, since its p-value of $2.2e^{-16}$ is significantly less than the significance level α of 0.05. Its high test statistic of 47.83 further corroborates the claim that trend(BHPI) is a statistically significant variable. In fact, the intercept also is a statistically significant coefficient in predicting future BHPI values, since its p-value is $2.2e^{-16}$, which is significantly less than the specified significance level α. It can be concluded from an overarching perspective that there is a long-term trend with the data. A linear model attempting to fit this data is provided.



A Durbin-Watson Test shall now be performed on the model that utilizes the trend of the natural logarithm of BHPI to assess whether the true autocorrelation values are greater than 0. The following hypotheses are tested:

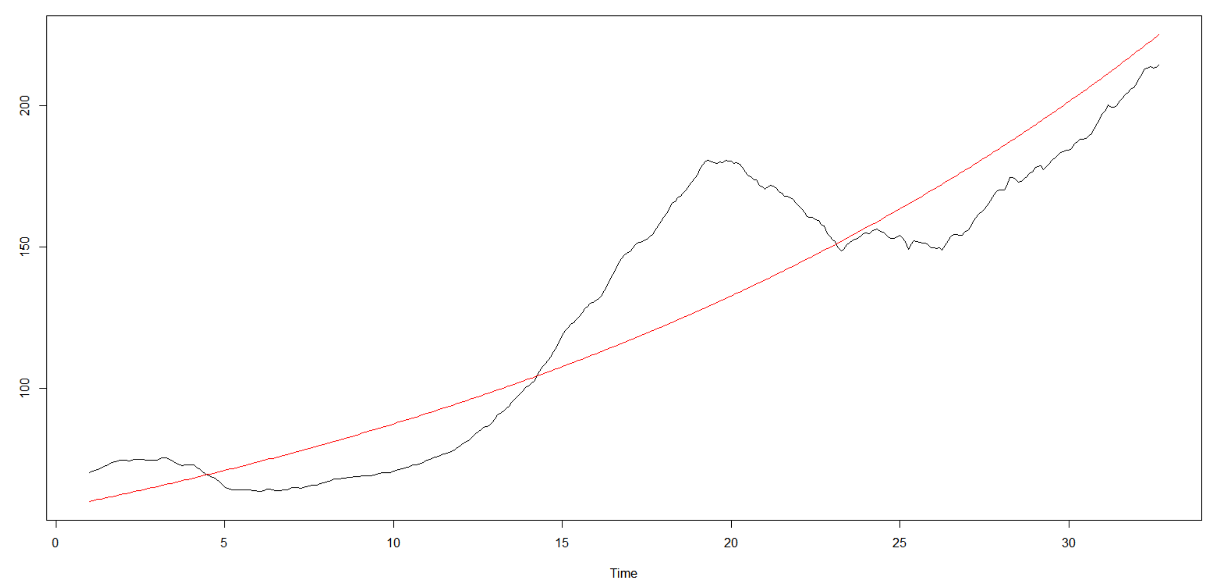
$$H_0: \rho = 0$$

$$H_A: \rho > 0$$

```
> dwtest(logmodel)

        Durbin-Watson test

data:  logmodel
DW = 0.0015516, p-value < 2.2e-16
alternative hypothesis: true autocorrelation is greater than 0
```

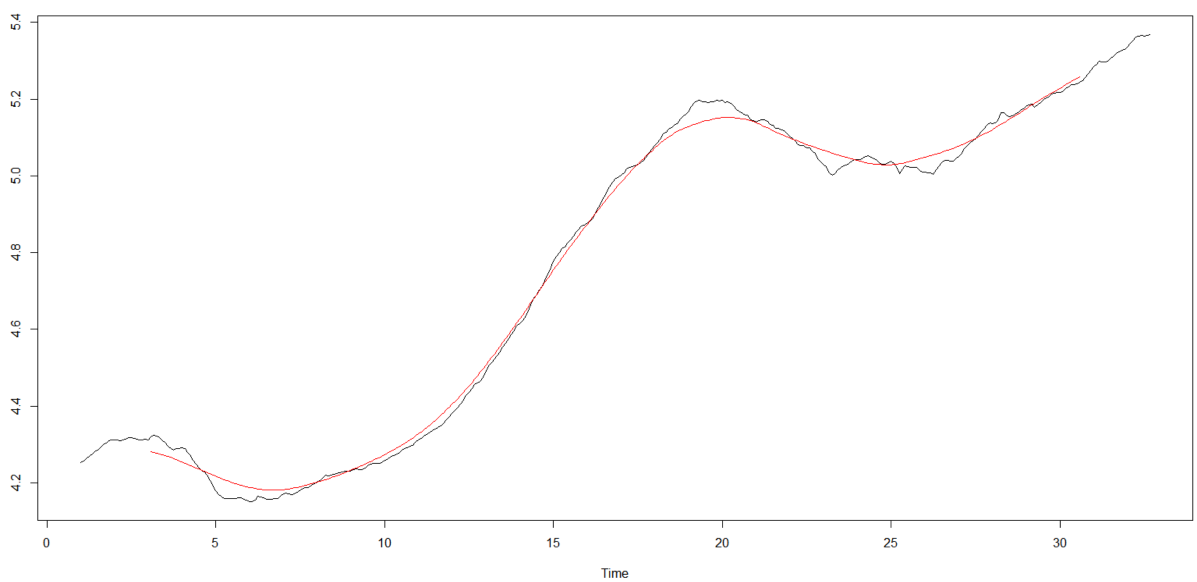
The output above shows that the p-value of $2.2e^{-16}$ is once again significantly less than α, which provides enough evidence to reject the Null Hypothesis and conclude that the true autocorrelation value is greater than 0 and that there appears to be autocorrelation amongst the residuals, which is a violation of one of the major assumption of regression. With this in mind, the analysis will carefully continue, keeping in mind that the residuals are highly correlated amongst each other.

Using a linear model to fit the growth and the trend found in the natural logarithm of BHPI is unreasonable here since the growth of this variable does not appear to be linear. The growth rate of the ln(BHPI) appears to be exponential, so a model named GrowthFit will be the exponential model used to determine this growth rate for the response variable BHPI. A plot of the GrowthFit model with respect to the ln(BHPI) is provided.



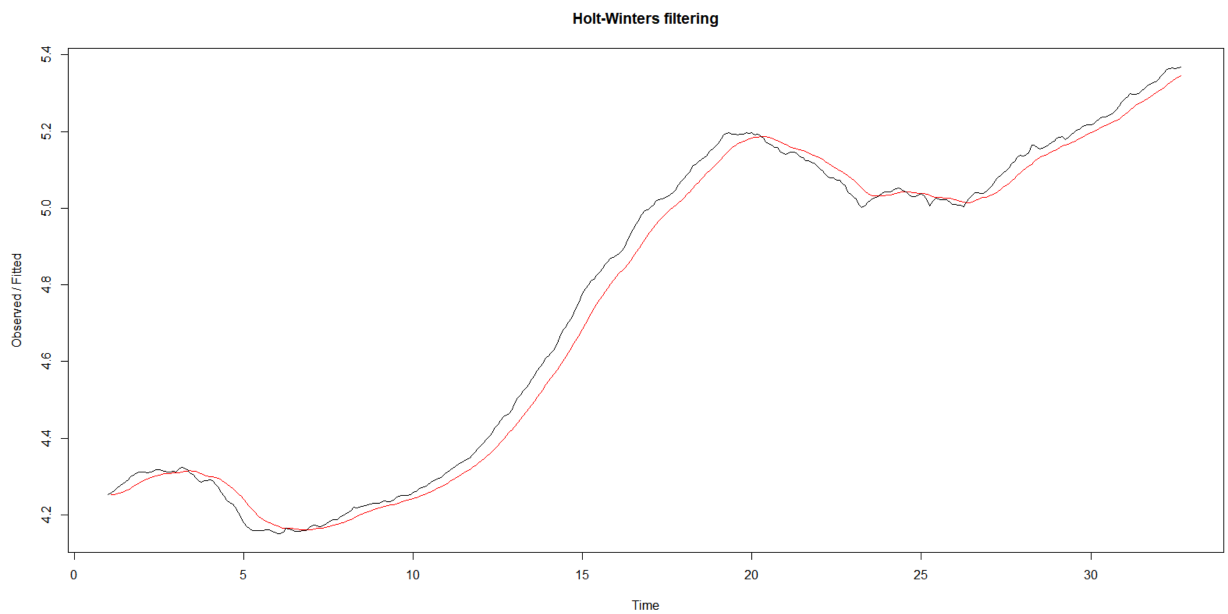
Although the exponential growth model that GrowthFit provides fits the data slightly better, it still remains a very rough estimate in an attempt to plot the average growth rate of ln(BHPI) over time while accounting for this trend. A more accurate exponential model must be constructed to better fit the data.

The mean of the BHPI over the January 1978 to September 2018 timeframe is 128.2073, which rounded to three decimals as the original data provides, is 128.207 points. However, time must be considered in the average value of BHPI, so an appropriate Moving Average must be determined. The process for selecting the appropriate Moving Average term consists of trial-and-error, and although tedious, a Moving Average Model of Order 50 appears to be the appropriate Moving Average term. A plot of this Moving Average term and the ln(BHPI) ensues. The red curve is the Moving Average term, while the black curve is the ln(BHPI).

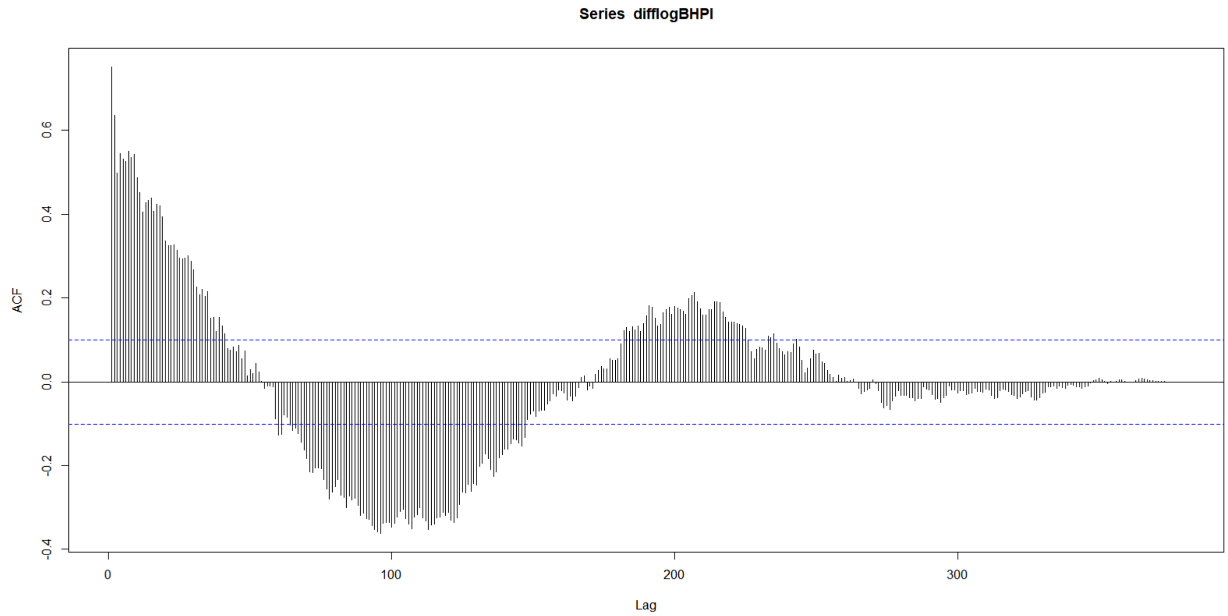


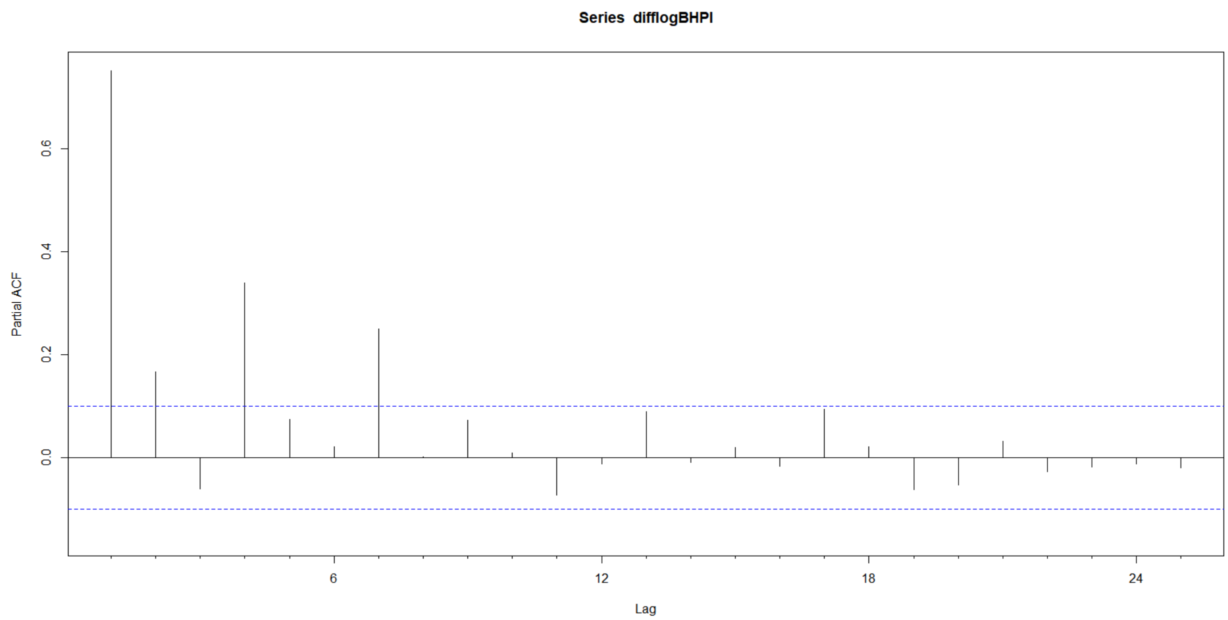
Suppose we wish to construct a Holt-Winters Filtering Model that predicts the ln(BHPI) trend better than the Moving Average Model of Order 50. We will see later in the report that a Moving

Average of Order 50 is actually not as great of an estimate as it initially appears to be, so the Holt-Winters Filtering Model seeks to further improve the accuracy behind predicting the exponential growth that appears in the ln(BHPI). Using $\alpha = 0.15$, the Holt-Winters Filtering Model is provided, with the red curve indicating the Holt-Winters Filtering Model and the black curve once again representing the ln(BHPI).



We now advance to analyzing the autocorrelation and Partial Autocorrelation functions of the first difference in the ln(BHPI). Autocorrelation and Partial Autocorrelation plots are provided.





Looking at the Autocorrelation plot first, the terms do exponentially approach to 0, but do so after about 300 lags, which suggests that a Moving Average Model of 300 might be the statistically more accurate Moving Average model to predict the values of the first difference of ln(BHPI).

Analyzing the Partial Autocorrelation plot, which requires significantly less lags than the Autocorrelation plot, it too does exponentially decline to 0, but this plot suggests that the model is an Autoregressive Model of Order 7. The data is not stationary until the first difference of the ln(BHPI) is taken, which then it becomes stationary.

A Box Test on monthly data shall now be performed, which tests the following hypotheses:

$$H_0 = Residuals\ are\ White\ Noise$$

$$H_A = Residuals\ are\ NOT\ White\ Noise$$

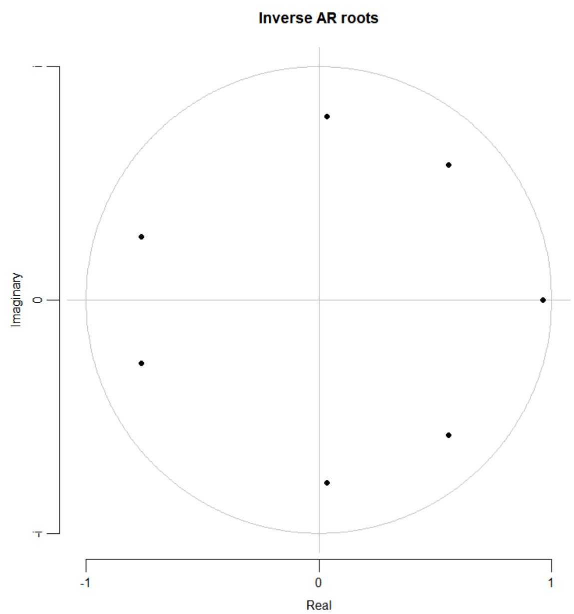
We wish to prove that the residuals are white noise, so in an unusual turn of events, we wish to prove the Null Hypothesis correct. The Box Test output is as follows:

```
> Box.test(difflogBHPI, 12, "Lj")

        Box-Ljung test

data:  difflogBHPI
X-squared = 1382.9, df = 12, p-value <
2.2e-16
```

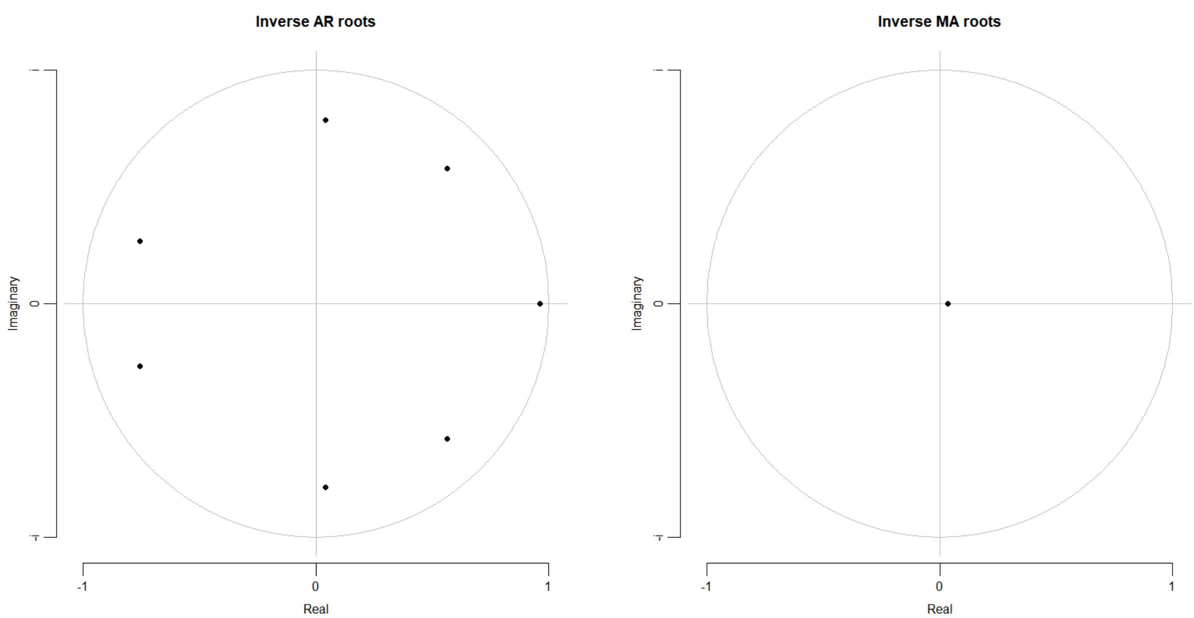
The Ljung, or Chi-Squared test statistic (χ^2) of 1382.9 and a p-value of $2.2e^{-16}$ leads us to reject the Null Hypothesis and conclude that the residuals are not white noise, meaning the model we estimated thus far has been a poor estimation.

Another consideration to note is whether or not the data is weakly stationary. Three separate ARIMA models have been generated, named Arima1, Arima2, and Arima3. Arima1 consists of the 7 Autoregressive terms the Partial Autocorrelation plot suggested the model should incorporate, Arima2 has the 7 Autoregressive terms and 1 Moving Average term, and Arima3 consists of 7 Autoregressive terms and 2 Moving Average terms. To determine whether or not these ARIMA models are weakly stationary, their plots must be provided.

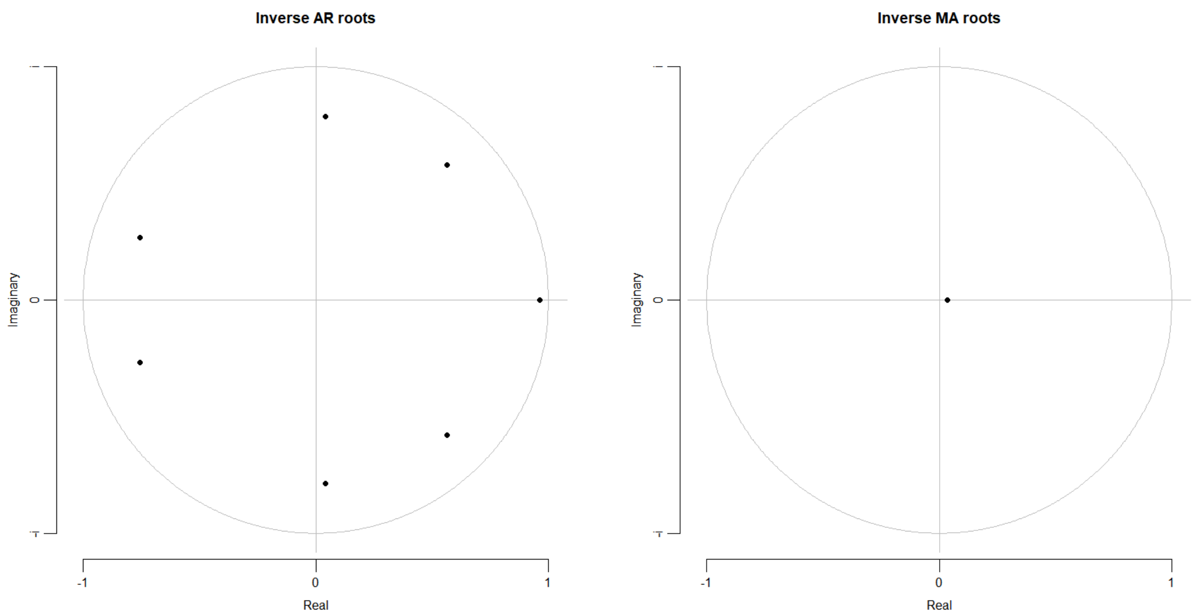
The plot for Arima1 shows that the data is weakly stationary since all of the autoregressive terms, represented by the black points, lie within the circle:



The plot for Arima2 also shows that the data is weakly stationary since the Moving Average and Autoregressive terms, both represented by black points, lie within the circle:



Similar to Arima1 and Arima2, the model Arima3 is also weakly stationary, with its Moving Average and Autoregressive terms lying within their respective circles:



We now wish to individually analyze the ARIMA models constructed. Starting with Arima1, its values are reported below.

```
> summary(Arima1)
Series: difflogBHPI
ARIMA(7,0,0) with non-zero mean

Coefficients:
         ar1     ar2      ar3     ar4     ar5
      0.6233  0.1336  -0.3537  0.3567  0.0229
s.e.  0.0495  0.0584   0.0590  0.0588  0.0588
          ar6     ar7    mean
      -0.1362  0.2509  0.0031
s.e.   0.0585  0.0494  0.0018

sigma^2 estimated as 1.413e-05:  log likelihood=1585.69
AIC=-3153.38   AICc=-3152.89   BIC=-3117.92

Training set error measures:
                        ME        RMSE
Training set -2.573939e-05 0.003719271
                     MAE       MPE     MAPE
Training set 0.002795314 -4.322606 216.0714
                  MASE         ACF1
Training set 0.5102566 0.0003329842
```

Note that the Akaike Information Criteria (AIC) value is equivalent to -3153.38 and the Bayseian Information Criteria (BIC) value is equivalent to -3117.92 for the Arima1 model. Its Root Mean Square Error (RMSE) is 0.0037, and its Mean Absolute Error is 0.0028. Outputs for the summaries of Arima2 and Arima3 are also provided for comparison.

```
> summary(Arima2)
Series: difflogBHPI
ARIMA(7,0,1) with non-zero mean

Coefficients:
         ar1     ar2      ar3     ar4     ar5
      0.6570  0.1124  -0.3586  0.3661  0.0132
s.e.  0.2915  0.1900   0.0732  0.0997  0.1014
          ar6     ar7     ma1    mean
      -0.1382  0.2501  -0.036  0.0031
s.e.   0.0614  0.0502   0.307  0.0018

sigma^2 estimated as 1.417e-05:  log likelihood=1585.7
AIC=-3151.39   AICc=-3150.79   BIC=-3111.99

Training set error measures:
                        ME        RMSE         MAE
Training set -2.609142e-05 0.003719205 0.002796465
                   MPE     MAPE      MASE
Training set -5.173804 216.6741 0.5104667
                    ACF1
Training set 0.001981283
```

For Arima2, its AIC, BIC, RMSE, and MAE values are -3151.39, -3111.99, 0.0037, and 0.028, respectively

```
> summary(Arima3)
Series: difflogBHPI
ARIMA(7,0,2) with non-zero mean

Coefficients:
         ar1     ar2      ar3     ar4     ar5
      0.6231  0.3152  -0.4685  0.3286  0.0747
s.e.  0.1847  0.2187   0.1194  0.0879  0.0931
          ar6     ar7      ma1      ma2    mean
      -0.1897  0.2391  -0.0015  -0.1949  0.0031
s.e.   0.0737  0.0537   0.1900   0.1786  0.0019

sigma^2 estimated as 1.416e-05:  log likelihood=1586.33
AIC=-3150.67   AICc=-3149.95   BIC=-3107.32

Training set error measures:
                        ME        RMSE         MAE
Training set -2.394793e-05 0.003712872 0.002793625
                   MPE     MAPE      MASE
Training set -3.198315 214.0546 0.5099483
                    ACF1
Training set 0.001417487
>
```

For Arima3, its AIC, BIC, RMSE, and MAE values are -3150.67, -3107.32, 0.0037, and 0.0028, respectively.

A summary table for Arima1, Arima2, and Arima3 is provided to view these values in a more concise and convenient manner.

Model	**AR**	**MA**	**AIC**	**BIC**	**RMSE**	**MAE**
Arima1	7	0	-3153.38	-3117.72	0.00372	0.0028
Arima2	7	1	-3151.39	-3111.99	0.00372	0.0028
Arima3	7	2	-3150.67	-3107.32	0.00371	0.0028

From this data table, two of these models shall be selected for further comparison and to run forecasts on them. The two models to use for further analysis would be Arima1 and Arima2, since their AIC and BIC values each are better than Arima3's, despite these three ARIMA models having extremely similar RMSEs and MAEs.

With the models Arima1 and Arima2 selected for further analysis, it is time to prepare the forecasts for both these models. Before forecasting, however, it is necessary to divide the data into training and testing datasets, where the training dataset will consist of about 90%, or 344, of the 380 observations and the remaining 10%, or 36 observations, of the dataset being in the testing dataset. The training dataset serves to build the ARIMA models and the testing dataset serves to validate the accuracy of the ARIMA models

With the data now appropriately divided into the training and testing datasets, accurate forecasts can be produced. However, certain Hypothesis Tests must be conducted to ensure that the residuals are 0 and are white-noise, and determine whether the value within the forecast of the ARIMA model are efficient.

The first test to perform is a t-Test evaluation of Arima1's residuals and Arima2's residuals to determine whether or not these residuals are equivalent to 0. The following hypotheses are tested:

$$H_0 = Residuals = 0$$

$$H_A = Residuals \neq 0$$

The outputs of the t-Tests for Arima1's and Arima2's residuals are provided.

```
> t.test(Arima1$residuals)

        One Sample t-test

data:  Arima1$residuals
t = -0.13473, df = 379, p-value = 0.8929
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.0004013735  0.0003498948
sample estimates:
    mean of x
-2.573939e-05
```

```
> t.test(Arima2$residuals)

        One Sample t-test

data:  Arima2$residuals
t = -0.13658, df = 379, p-value = 0.8914
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.0004017186  0.0003495358
sample estimates:
    mean of x
-2.609142e-05
```

Evaluating the results of these t-Tests for residuals of Arima1 and Arima2, their p-values of 0.8929 and 0.8914, respectively, show that their respective residuals do indeed equal to 0, which is what is desired for this test.

The second test to run before forecasting is the Box Test, which intends to analyze whether or not the residuals are white-noise. The following hypotheses are tested in the Box Test:

$$H_0 = Residuals\ are\ White\ Noise$$

$$H_A = Residuals\ are\ NOT\ White\ Noise$$

The output for these Box Tests for Arima1's and Arima2's residuals is provided below.

```
> Box.test(Arima1$residuals, 12, "Lj")

        Box-Ljung test

data:  Arima1$residuals
X-squared = 4.7135, df = 12, p-value = 0.9669
```

```
> Box.test(Arima2$residuals, 12, "Lj")

        Box-Ljung test

data:  Arima2$residuals
X-squared = 4.7617, df = 12, p-value = 0.9655
```

Analyzing the Box Test results of Arima1's and Arima2's residuals, their respective p-values of 0.9669 and 0.9655 show that their respective residuals are white-noise, which is what we want to conclude from this particular Hypothesis test.

Now summaries of the Arima1 and Arima2 testing datasets will be performed, with their outputs provided below.

```
> summary(Arima1Testing)
Series: difflogBHPI[344:380]
ARIMA(7,0,0) with non-zero mean

Coefficients:
         ar1     ar2      ar3      ar4      ar5
      0.2890  0.2044  -0.3778  -0.0223  -0.1041
s.e.  0.1594  0.1647   0.1716   0.1819   0.1727
          ar6     ar7    mean
      -0.1634  0.1985  0.0047
s.e.   0.1725  0.1700  0.0005

sigma^2 estimated as 9.634e-06:  log likelihood=165.05
AIC=-312.1   AICc=-305.43   BIC=-297.6

Training set error measures:
                        ME        RMSE         MAE
Training set -4.735215e-05 0.002747906 0.002212701
                   MPE     MAPE   MASE       ACF1
Training set -4.711436 108.6978 0.7118 0.06736798
```

```
> summary(Arima2Testing)
Series: difflogBHPI[344:380]
ARIMA(7,0,1) with non-zero mean

Coefficients:
          ar1     ar2      ar3      ar4      ar5
      -0.0068  0.2906  -0.3223  -0.1285  -0.1351
s.e.   0.4214  0.1884   0.1799   0.2343   0.1696
          ar6     ar7     ma1    mean
      -0.1959  0.1932  0.3158  0.0047
s.e.   0.1677  0.1853  0.4118  0.0005

sigma^2 estimated as 9.798e-06:  log likelihood=165.31
AIC=-310.62   AICc=-302.16   BIC=-294.51

Training set error measures:
                        ME        RMSE         MAE
Training set -5.190489e-05 0.002722935 0.002199702
                   MPE     MAPE      MASE      ACF1
Training set -5.400628 105.1835 0.7076185 0.0427824
```

Given these outputs, we find that the RMSE values of the Arima1 and Arima2 testing datasets are both 0.0027 and the MAE values of the Arima1 and Arima2 testing datasets are both 0.0022.

The Efficiency Test must now be performed on the Arima1 and Arima2 models to effectively determine whether the parameter Arima1$fitted or Arima2$fitted is 0 and therefore efficient. The hypotheses to consider for this particular test are as follows:

$$H_0 = The\ parameter\ Arima1\$fitted\ or\ Arima2\$fitted\ is\ Efficient$$

$$H_A = The\ parameter\ Arima1\$fitted\ or\ Arima2\$fitted\ is\ NOT\ Efficient$$

The outputs for Efficiency Test for Arima1’s and Arima2’s residuals are provided below.

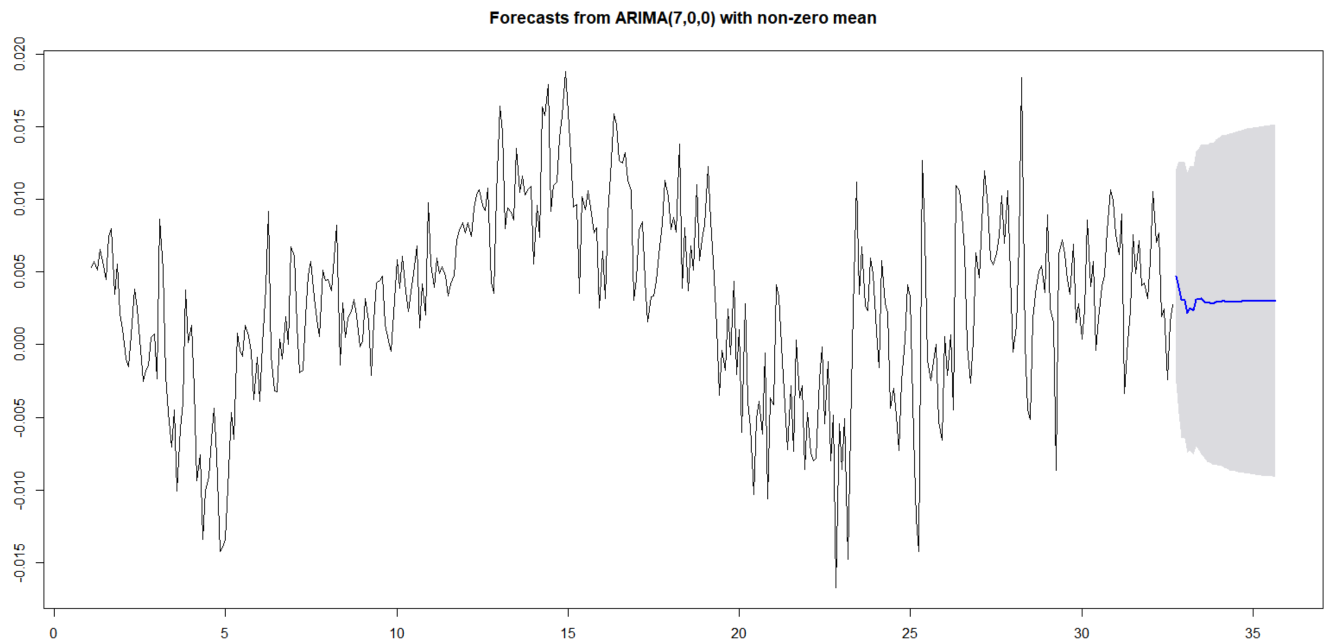
```
> summary(EfficiencyTest1)

Call:
lm(formula = Arima1$residuals ~ Arima1$fitted)

Residuals:
       Min         1Q     Median         3Q
-0.0124276 -0.0022073 -0.0000096  0.0023109
       Max
 0.0201672

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -3.918e-05  2.213e-04  -0.177    0.860
Arima1$fitted  4.538e-03  3.761e-02   0.121    0.904

Residual standard error: 0.003729 on 378 degrees of freedom
Multiple R-squared:  3.851e-05, Adjusted R-squared:  -0.002607
F-statistic: 0.01456 on 1 and 378 DF,  p-value: 0.904
```

```
> summary(EfficiencyTest2)

Call:
lm(formula = Arima2$residuals ~ Arima2$fitted)

Residuals:
       Min         1Q     Median         3Q
-0.0124497 -0.0021910  0.0000226  0.0023053
       Max
 0.0202007

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -3.966e-05  2.213e-04  -0.179    0.858
Arima2$fitted  4.581e-03  3.761e-02   0.122    0.903

Residual standard error: 0.003729 on 378 degrees of freedom
Multiple R-squared:  3.925e-05, Adjusted R-squared:  -0.002606
F-statistic: 0.01484 on 1 and 378 DF,  p-value: 0.9031
```

Examining the outputs provided, the Efficiency Tests for Arima1's and Arima2's residuals yield p-values of 0.904 and 0.903, respectively. Since we fail to reject the Null Hypothesis in this test, we can conclude that the value of their parameters (Arima1$fitted and Arima2$fitted) are nonzero and are therefore efficient.

It is officially the time to run and plot the forecasts for 36 time periods ahead (36 months, or 3 years) at the 80% and 95% Confidence Levels for Arima1 and Arima2. The plots are provided, as well as their forecasted values.

The Arima1 forecast at 95% Confidence and its forecasted values at 80% and 95% Confidence:

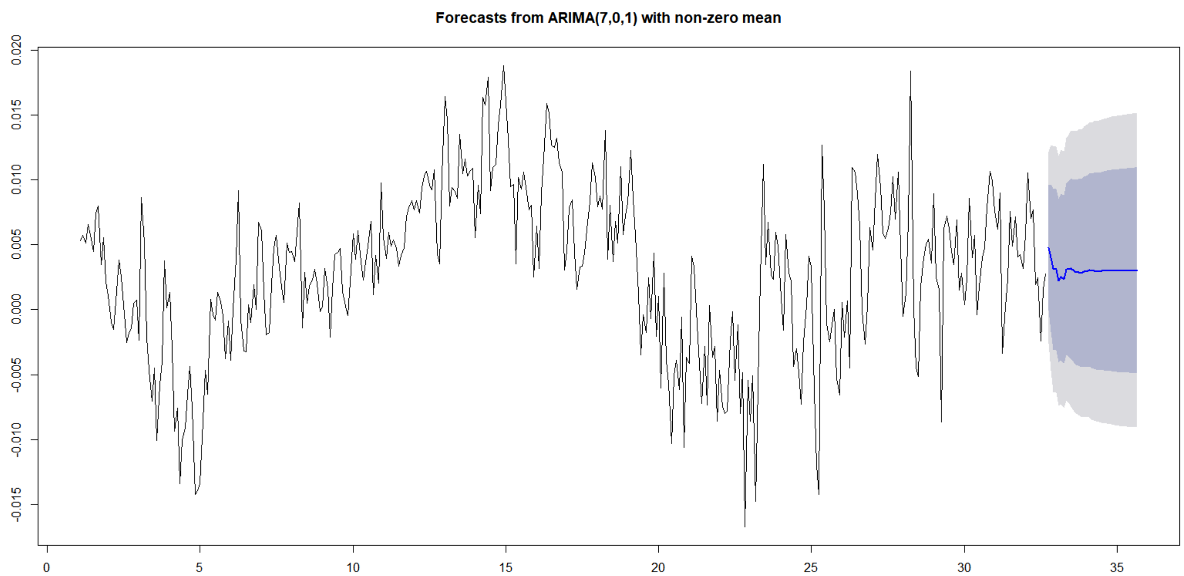

Forecasts from ARIMA(7,0,0) with non-zero mean
0.020
0.015
0.010
0.005
0.000
-0.005
-0.010
-0.015
0
5
10
15
20
25
30
35

The forecasted values for Arima1 is provided.

```
> print(Forecast1)
       Point Forecast         Lo 80       Hi 80        Lo 95      Hi 95
Oct 32    0.004729217 -8.820011e-05 0.009546635 -0.002638387 0.01209682
Nov 32    0.003952656 -1.723975e-03 0.009629287 -0.004729002 0.01263431
Dec 32    0.003097629 -3.111260e-03 0.009306519 -0.006398048 0.01259331
Jan 33    0.003064165 -3.150382e-03 0.009278712 -0.006440165 0.01256849
Feb 33    0.002179572 -4.141890e-03 0.008501034 -0.007488269 0.01184741
Mar 33    0.002527220 -3.880648e-03 0.008935087 -0.007272768 0.01232721
Apr 33    0.002337215 -4.151986e-03 0.008826416 -0.007587162 0.01226159
May 33    0.003152056 -3.470225e-03 0.009774337 -0.006975849 0.01327996
Jun 33    0.003116923 -3.655717e-03 0.009889564 -0.007240937 0.01347478
Jul 33    0.003164860 -3.816528e-03 0.010146247 -0.007512251 0.01384197
Aug 33    0.002954141 -4.109787e-03 0.010018069 -0.007849205 0.01375749
Sep 33    0.002858612 -4.281146e-03 0.009998371 -0.008060706 0.01377793
Oct 33    0.002873182 -4.317178e-03 0.010063543 -0.008123525 0.01386989
Nov 33    0.002801644 -4.458274e-03 0.010061561 -0.008301442 0.01390473
Dec 33    0.002927973 -4.391598e-03 0.010247543 -0.008266344 0.01412229
Jan 34    0.002937767 -4.441175e-03 0.010316709 -0.008347351 0.01422288
Feb 34    0.003029795 -4.411114e-03 0.010470703 -0.008350093 0.01440968
Mar 34    0.002978738 -4.516185e-03 0.010473660 -0.008483757 0.01444123
Apr 34    0.002973214 -4.570602e-03 0.010517030 -0.008564057 0.01451049
May 34    0.002950187 -4.630545e-03 0.010530919 -0.008643542 0.01454392
Jun 34    0.002951044 -4.666735e-03 0.010568823 -0.008699343 0.01460143
Jul 34    0.002964713 -4.686392e-03 0.010615817 -0.008736642 0.01466607
Aug 34    0.002968275 -4.715784e-03 0.010652334 -0.008783479 0.01472003
Sep 34    0.002993726 -4.720123e-03 0.010707576 -0.008803588 0.01479104
Oct 34    0.002992952 -4.748925e-03 0.010734829 -0.008847227 0.01483313
Nov 34    0.003001255 -4.766563e-03 0.010769073 -0.008878597 0.01488111
Dec 34    0.002993014 -4.797860e-03 0.010783889 -0.008922099 0.01490813
Jan 35    0.002996774 -4.815235e-03 0.010808783 -0.008950663 0.01494421
Feb 35    0.002998330 -4.832739e-03 0.010829400 -0.008978257 0.01497492
Mar 35    0.003003088 -4.846077e-03 0.010852254 -0.009001175 0.01500735
Apr 35    0.003008675 -4.857063e-03 0.010874412 -0.009020933 0.01503828
May 35    0.003012069 -4.869130e-03 0.010893268 -0.009041184 0.01506532
Jun 35    0.003017096 -4.878274e-03 0.010912465 -0.009057829 0.01509202
Jul 35    0.003017859 -4.890593e-03 0.010926312 -0.009077075 0.01511279
Aug 35    0.003020639 -4.899824e-03 0.010941102 -0.009092663 0.01513394
Sep 35    0.003021777 -4.909678e-03 0.010953232 -0.009108337 0.01515189
```

The output provided here gives the values of the first difference of the ln(BHPI) projected 36 time periods ahead, which is 36 months, or 3 years. Further analysis of this table suggests that at time progresses past September 2018, the value of the first difference of the ln(BHPI) approaches 0.003.

The Arima2 forecast at 95% Confidence and its forecasted values at 80% and 95% Confidence:



Note how similar the forecast for Arima2 is compared to Arima1; in fact, these plots appear to reveal the exact same forecasts. Due to the interval chosen to forecast on, which was 36 months, or 3 years, the forecast itself can be slightly difficult to see and practically interpret from. Both plots, however, on the 80% (colored as dark blue shaded area) and the 95% Confidence Level (the area that includes the 80% dark blue area in addition to the gray shaded area) suggest that the forecasted values of the first difference of ln(BHPI) projected 36 months ahead (3 years ahead) will be decreasing and eventually leveling off at around 0.003.

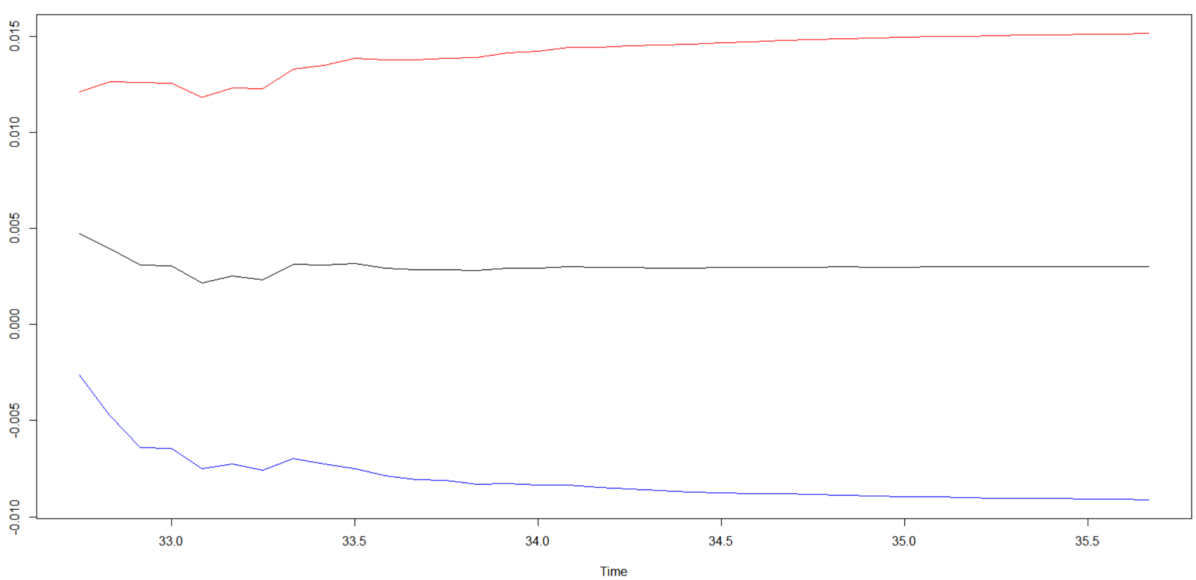
The forecasted values for Arima2 are also printed at the 80% and 95% Confidence Levels.

```
> print(Forecast2)
       Point Forecast         Lo 80       Hi 80        Lo 95      Hi 95
Oct 32    0.004767123 -5.669674e-05 0.009590943 -0.002610273 0.01214452
Nov 32    0.004001279 -1.676966e-03 0.009679524 -0.004682847 0.01268541
Dec 32    0.003171014 -3.037401e-03 0.009379429 -0.006323937 0.01266596
Jan 33    0.003101200 -3.112490e-03 0.009314890 -0.006401819 0.01260422
Feb 33    0.002229872 -4.087902e-03 0.008547646 -0.007432329 0.01189207
Mar 33    0.002520989 -3.881683e-03 0.008923662 -0.007271053 0.01231303
Apr 33    0.002337522 -4.144539e-03 0.008819582 -0.007575935 0.01225098
May 33    0.003139094 -3.470338e-03 0.009748527 -0.006969160 0.01324735
Jun 33    0.003143919 -3.614917e-03 0.009902755 -0.007192829 0.01348067
Jul 33    0.003200030 -3.769090e-03 0.010169150 -0.007458319 0.01385838
Aug 33    0.002989584 -4.066160e-03 0.010045328 -0.007801245 0.01378041
Sep 33    0.002888744 -4.242725e-03 0.010020214 -0.008017897 0.01379539
Oct 33    0.002889254 -4.292712e-03 0.010071220 -0.008094616 0.01387312
Nov 33    0.002817670 -4.432895e-03 0.010068235 -0.008271113 0.01390645
Dec 33    0.002930404 -4.379721e-03 0.010240530 -0.008249468 0.01411028
Jan 34    0.002949987 -4.419252e-03 0.010319227 -0.008320291 0.01422027
Feb 34    0.003043172 -4.388009e-03 0.010474353 -0.008321838 0.01440818
Mar 34    0.003001255 -4.485165e-03 0.010487676 -0.008448237 0.01445075
Apr 34    0.002992201 -4.544120e-03 0.010528523 -0.008533608 0.01451801
May 34    0.002966800 -4.607345e-03 0.010540945 -0.008616855 0.01455046
Jun 34    0.002965016 -4.646320e-03 0.010576351 -0.008675518 0.01460555
Jul 34    0.002975615 -4.669455e-03 0.010620685 -0.008716511 0.01466774
Aug 34    0.002979641 -4.698732e-03 0.010658013 -0.008763417 0.01472270
Sep 34    0.003003800 -4.704866e-03 0.010712465 -0.008785587 0.01479319
Oct 34    0.003006099 -4.731119e-03 0.010743318 -0.008826956 0.01483915
Nov 34    0.003013985 -4.749785e-03 0.010777755 -0.008859677 0.01488765
Dec 34    0.003006266 -4.781240e-03 0.010793772 -0.008903697 0.01491623
Jan 35    0.003008244 -4.800930e-03 0.010817418 -0.008934858 0.01495134
Feb 35    0.003009103 -4.819629e-03 0.010837836 -0.008963909 0.01498212
Mar 35    0.003013245 -4.834006e-03 0.010860495 -0.008988090 0.01501458
Apr 35    0.003018357 -4.845944e-03 0.010882657 -0.009009053 0.01504577
May 35    0.003021980 -4.858237e-03 0.010902196 -0.009029771 0.01507373
Jun 35    0.003026829 -4.868023e-03 0.010921682 -0.009047305 0.01510096
Jul 35    0.003027913 -4.880477e-03 0.010936303 -0.009066926 0.01512275
Aug 35    0.003030173 -4.890669e-03 0.010951015 -0.009083709 0.01514406
Sep 35    0.003031077 -4.901177e-03 0.010963331 -0.009100258 0.01516241
```

Analyzing the forecasted values for Arima2, over the course of the 36 months, the predicted first difference value of the ln(BHPI) seems to be approaching 0.003, further corroborating the findings on the time-series plot of the forecasted values of Arima2.

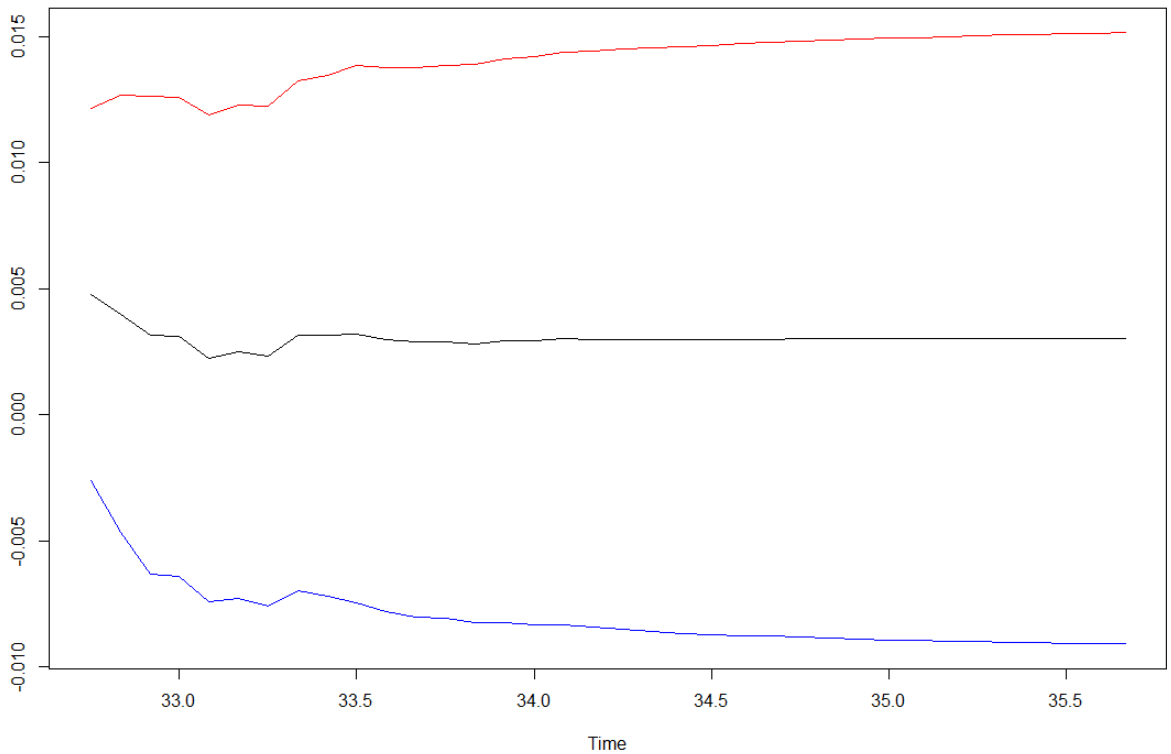
Note that the forecasted values are extremely similar between Arima1 and Arima2; the only term differentiating them from each other is the addition of a Moving Average term found in Arima2, and it seems to have very little impact on the forecasted values.

Now the mean, lower, and upper forecasted values for Arima1 at 80% Confidence levels will be plotted, which is R's default Confidence Level to plot these variables. The black curve represents the time-series plot of the mean 80% Confidence values for Arima1, the blue curve represents the time-series plot of the lower 80% Confidence values for Arima1, and the red represents the time-series plot of the upper 80% Confidence values for Arima1.



Analyzing this plot, we see that the lower 80% of the Confidence Level projects that for the 36 time periods projected, the value of the first difference of the ln(BHPI) contains values that lie between -0.0025 and -0.008. The upper 80% of the Confidence Level projects that for the 36 time periods projected, the values of the first difference of the ln(BHPI) range between 0.013 and 0.015. The mean, or average, 80% Confidence Level projects that these values of the first difference of ln(BHPI) forecasted 36 time periods ahead have values that lie between approximately 0.0035 and 0.005.

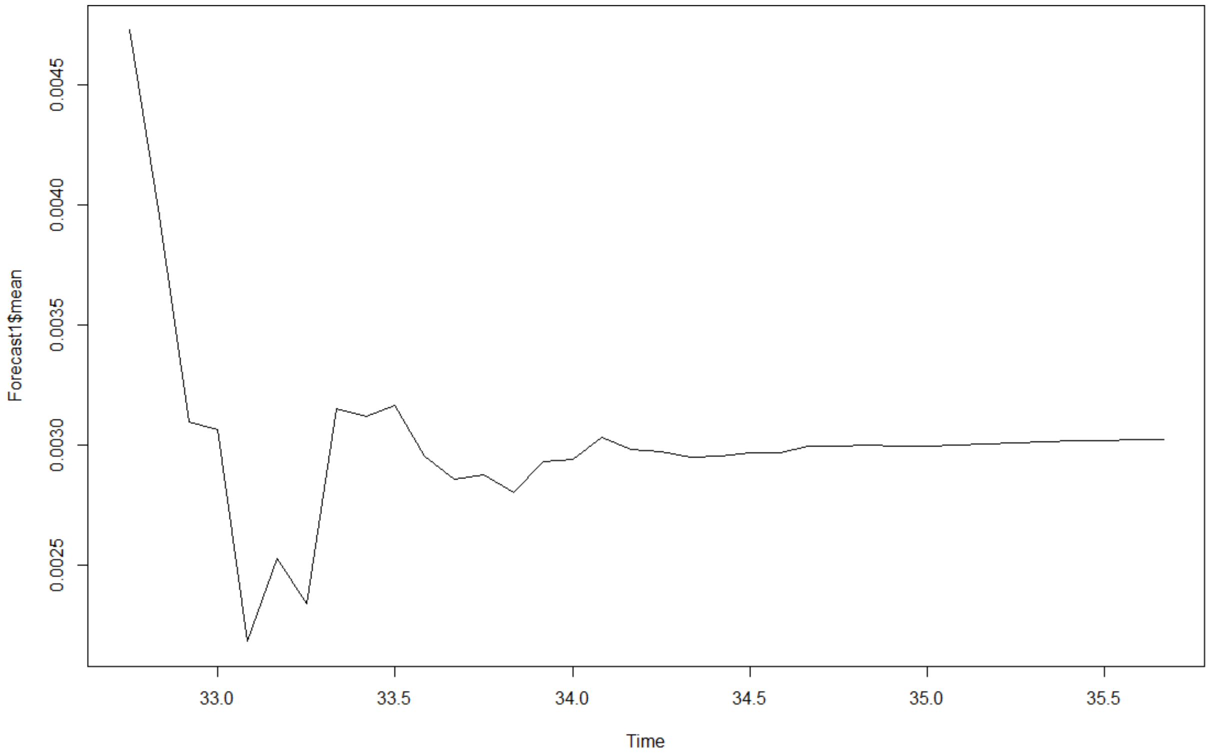
Arima2 has a very similar plot to the one above, which is also shown below for convenience. The black curve represents the time-series plot of the mean 80% Confidence values for Arima1, the blue curve represents the time-series plot of the lower 80% Confidence values for Arima1, and the red represents the time-series plot of the upper 80% Confidence values for Arima1.



Similar to the Arima1 plot, we see that the lower 80% of the Confidence Level projects that for the 36 time periods projected, the value of the first difference of the ln(BHPI) contains values that lie between -0.0025 and -0.008. The upper 80% of the Confidence Level projects that for the 36 time periods projected, the values of the first difference of the ln(BHPI) range between 0.013 and 0.015. The mean, or average, 80% Confidence Level projects that these values of the first difference of ln(BHPI) forecasted 36 time periods ahead have values that lie between approximately 0.0035 and 0.005.

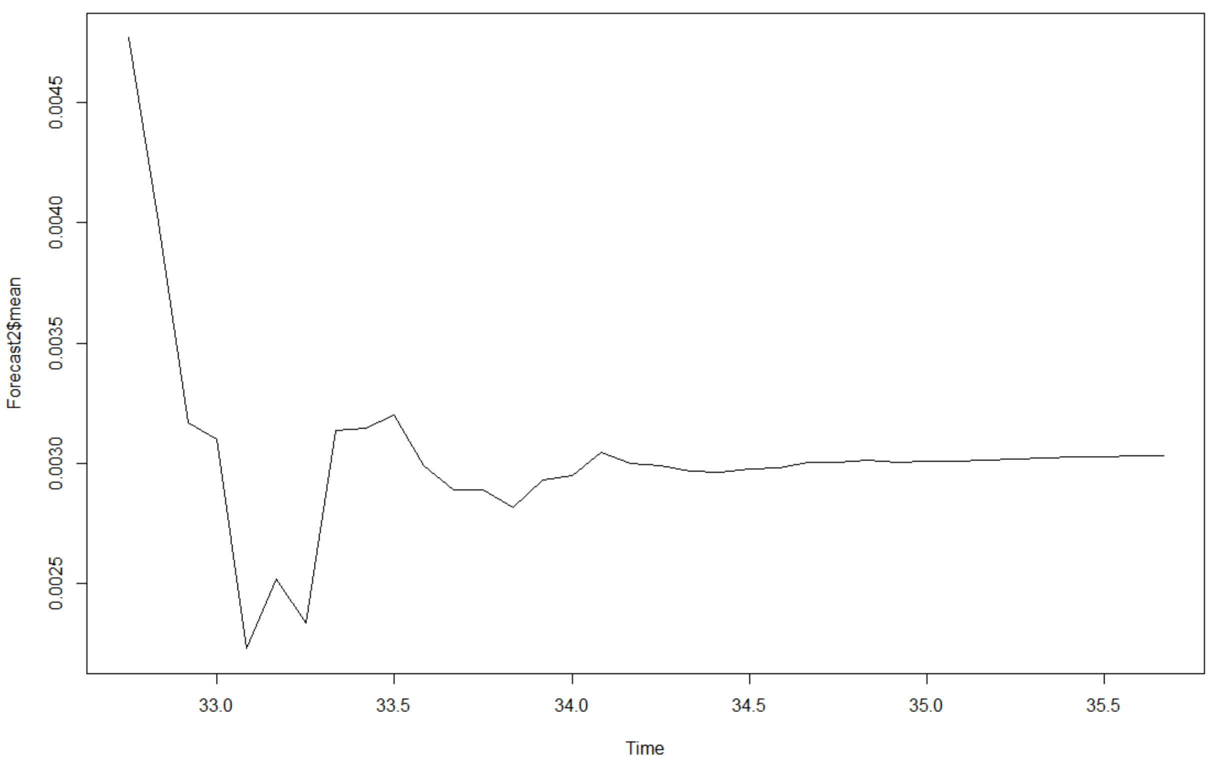
Individual plots of the forecasted lower, mean, and upper 80% and 95% Confidence Level values for the forecast of Arima1 and Arima2 in order to better understand the values the forecasts of Arima1 and Arima2 take over this projection of 36 time periods.

This particular plot displays the relationship shared between the number of time periods projected into the future (36) and the Arima1's forecasted mean values at the 80% Confidence Level of the first difference of the ln(BHPI) it takes over this time period projection of 36 months.



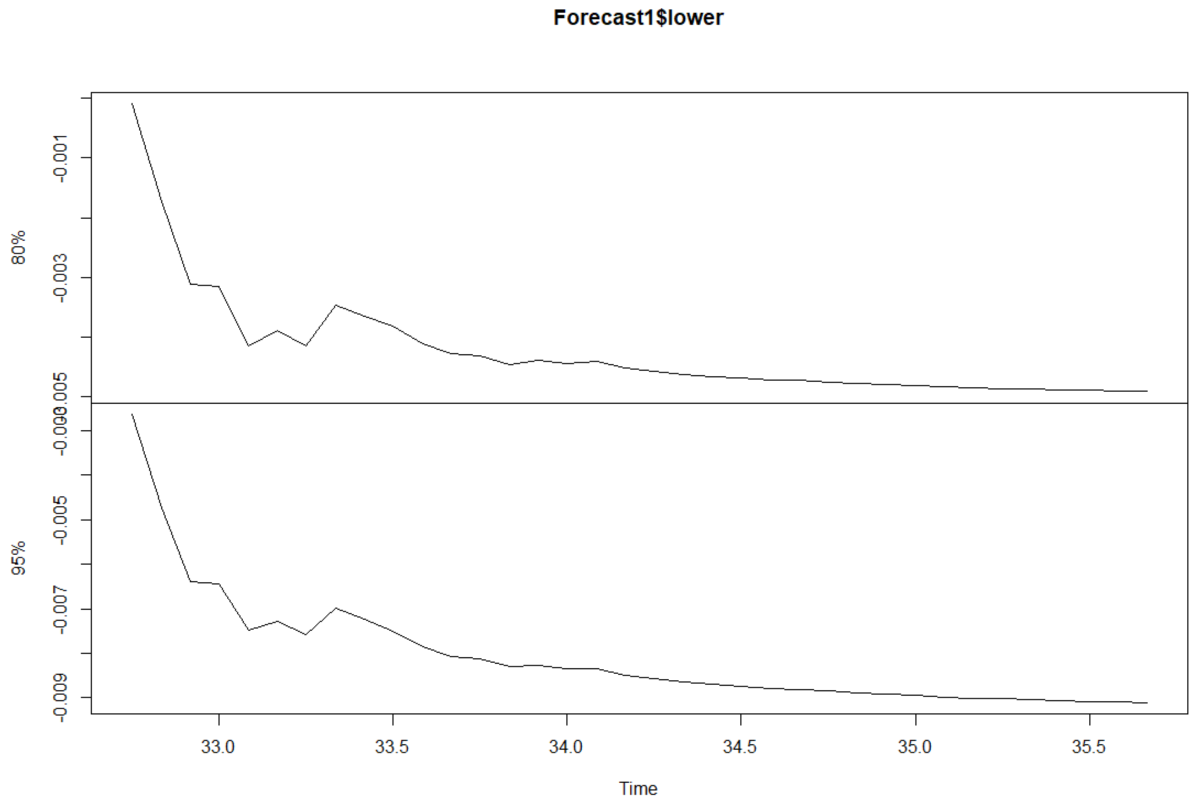
Notice that initially, it takes values that exceed over 0.045, but as time progresses monthly, the first difference of the ln(BHPI) appears to approach to the value of around 0.0030, corroborating the findings made in the plot for Arima1's forecasted values of the first difference of the ln(BHPI), which also begins to approach 0.0030 as the number of time periods projected into the future on a monthly basis increases.

Similar findings can be made for Arima2's mean forecasted values at the 80% Confidence Level for the first difference of the ln(BHPI). The plot of these particular values is displayed below.



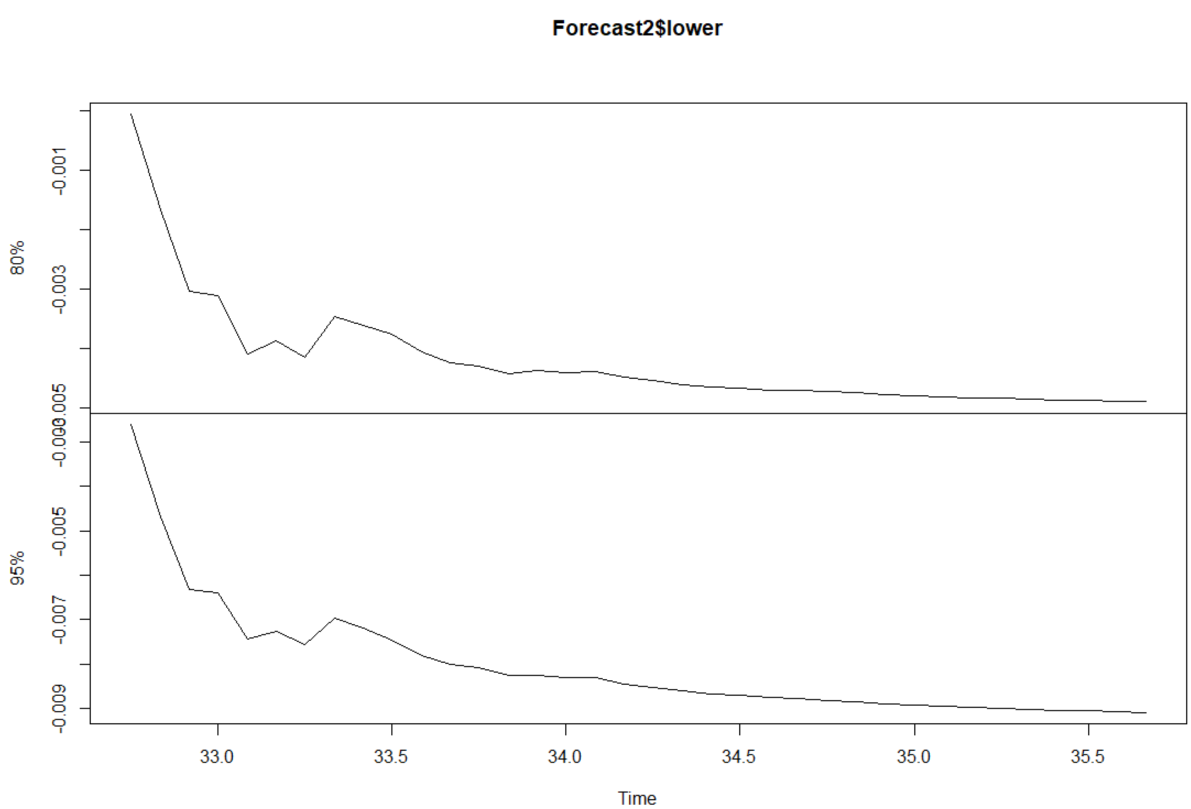
Just like Arima1's forecasted mean values at the 80% Confidence Level, the first difference of the ln(BHPI) appears to approach the value of 0.0030 as the number of time periods projected into the future. This also supports the findings in the Arima2 plot that includes the 80% Confidence Level mean forecasted values introduced earlier, where it also approaches 0.0030 as the number of time periods projected into the future on a monthly basis increases.

A more interesting plot is introduced here. With respect to the mean forecasted values of the Arima1 model, the lower 80% and 95% Confidence Level values are plotted over the time projection of 36 months.



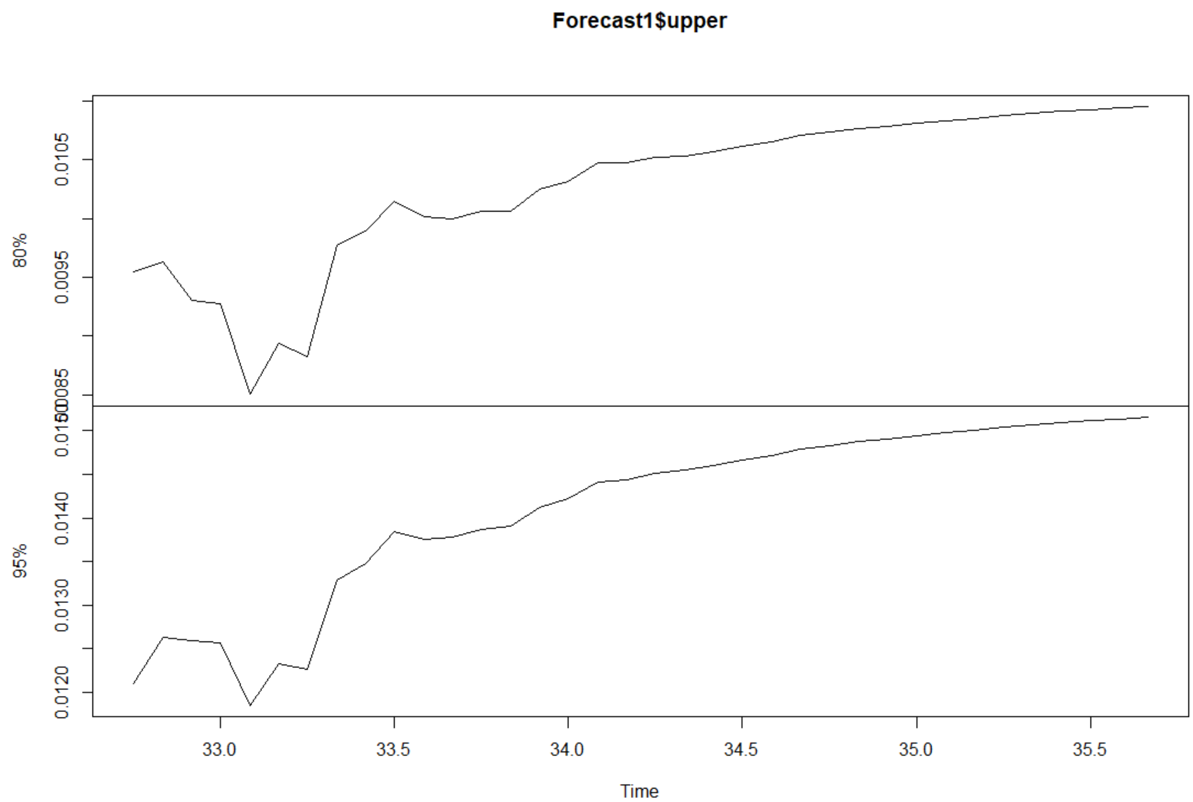
Beginning the analysis at the Lower 80% Confidence Level values for the forecasted values of Arima1, it appears that as the number of months projected increases, these forecasted values begin to approach -0.005 after initially starting at levels that exceed the negative values. Similar conclusions can be made for the Arima1 forecasted values at the Lower 95% Confidence Level, where the values here also begin as positive, but the Lower 95% Confidence Level values begin to approach -0.010. This is due to the fact that an increase in the Confidence Level results in a direct increase in the interval these forecasted values at this given Confidence Level can take.

Now the same plot analyzing the Lower 80% and 95% Confidence Levels with respect to the mean forecasted values of the Arima2 model projected over 36 months is below.



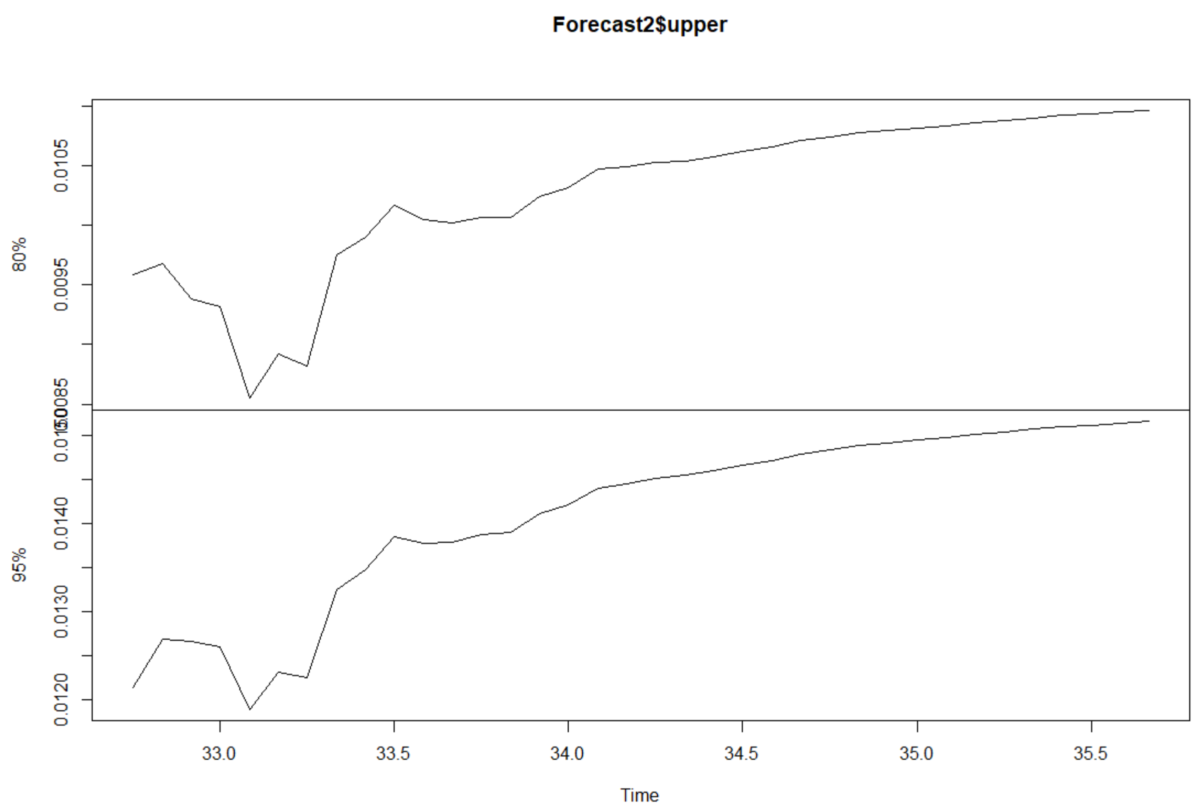
According to the Lower 80% Confidence Level values for the forecasted values of Arima2, as the number of months projected increases, these forecasted values begin to approach -0.005 after initially starting at levels that are positive, very similar to the findings found in the Arima1 forecasted values at the Lower 80% Confidence Level. And just like in the Arima1 forecasted values at the Lower 95% Confidence Level, the values here also begin as positive, but the Lower 95% Confidence Level values begin to approach -0.010. The similarity in the values between the Arima1 forecasted Lower 80% and 95% Confidence Levels and the Arima2 forecasted Lower 80% and 95% Confidence Levels is due to how similar the models are with respect to the number of Autoregressive terms incorporated into them, with only 1 Moving Average term differentiating Arima1 from Arima2.

We also have plots that analyze the Upper 80% and 95% Confidence Level forecasted values for both Arima1 and Arima2, which has been forecasted 36 time periods ahead. The plot below analyzes Arima1's forecasted values within the Upper 80% and 95% Confidence Levels.



Examining the Upper 80% Confidence Level values for the forecasted values of Arima1's forecasted values, it appears that as the number of months projected increases, these forecasted values initially drop from around 0.010 to as low as 0.0085 for about 2.5 time periods before increasing rather rapidly and then approaching a value that exceeds 0.0105. At the Upper 95% Confidence Level for Arima1's forecasted values projected 36 time periods into the future, values appear to be very volatile over the first 2.5 time periods before rapidly increasing and approaching a value exceeding 0.0105. These Upper 80% and 95% Confidence Levels contain forecasted values from Arima1 that are exceptionally higher than the average value of these forecasted values for Arima1, as Upper ends of Confidence Intervals tend to contain values that significantly exceed average or above average values observed.

The focus will now shift to Arima2's forecasted values projected 36 time periods into the future that fall within the Upper 80% and 95% Confidence Levels. A plot exploring this relationship is provided.



Analyzing the Upper 80% Confidence Level values for the forecasted values of Arima2's forecasted values, it also appears to have a large amount of volatility in the first 2.5 time periods projected, with these forecasted values initially dropping from around 0.010 to as low as 0.0085 before increasing rapidly and then exceeding 0.0105. At the Upper 95% Confidence Level for Arima2's forecasted values projected 36 time periods into the future, values also appear to be very volatile over the first 2.5 time periods before rapidly increasing and approaching a value exceeding 0.0105. Similar to the findings in the Arima1 forecasted values projected 36 time periods ahead that fall within the Upper 80% and 95% Confidence Level, volatility plagues the forecasted values of the first difference of the ln(BHPI) for a small portion of time before these values rapidly increase and begin to exceed values of 0.0105.

Suppose we wish to construct another model named Model3 that analyzed just the ln(BHPI) and was composed of 5 Autoregressive terms. We are also interested in forecasting 36 time periods ahead, or 36 months ahead, the values of ln(BHPI). Running summary statistics on Model yielded the following output produced by R.

```
> summary(Model3)
Series: logBHPI
ARIMA(5,0,0) with non-zero mean

Coefficients:
         ar1      ar2      ar3     ar4      ar5
      1.6707  -0.5386  -0.4056  0.6211  -0.3479
s.e.  0.0397   0.0434   0.0309  0.0358      NaN
        mean
      5.2457
s.e.     NaN

sigma^2 estimated as 1.54e-05:  log likelihood=1575.36
AIC=-3136.72   AICc=-3136.42   BIC=-3109.12

Training set error measures:
                       ME        RMSE         MAE
Training set 0.0001272785 0.003893406 0.002922875
                     MPE      MAPE       MASE
Training set 0.002387765 0.0612294 0.05036346
                    ACF1
Training set -0.04198778
```

There appears to be an error in the Standard Error of the Model, since it outputs a "NaN" in R, which stands for "Not a Number" This value occurred in computing the fifth Autoregressive Term's standard error, thereby affecting the Standard Error of the Model. During the calculation of the Standard Error of the Model, it might have encountered a negative number inside a radical, division by 0, or some other computational error that occurred in the calculation of the Standard Error of the Model. We are not too concerned with this measure, but it is worth pointing out.

The AIC, BIC, RMSE, and MAE of Model 3 is -3136.72, -3109.12, 0.00399, and 0.00292, respectively. With a new model constructed, it is necessary to add the values of this model into the data table containing the ARIMA models. Note that the Arima3 and Model3 models are **NOT** the same.

Model	AR	MA	AIC	BIC	RMSE	MAE
Arima1	7	0	-3153.38	-3117.72	0.00372	0.0028
Arima2	7	1	-3151.39	-3111.99	0.00372	0.0028
Arima3	7	2	-3150.67	-3107.32	0.00371	0.0028
Model3	5	0	-3136.72	-3109.12	0.00389	0.0029

The next step in analyzing Model3 is to run the Box Test on it, which tests whether or not the residuals around Model3 are considered to be white-noise. The following hypotheses are tested in this Box Test of Model3:

$$H_0 = Residuals\ are\ White\ Noise$$

$$H_A = Residuals\ are\ NOT\ White\ Noise$$

The output of the Box Test for Model3 is provided.

```
> Box.test(Model3$residuals, 12, "Lj", 6)

        Box-Ljung test

data:  Model3$residuals
X-squared = 29.892, df = 6, p-value = 4.121e-05
```

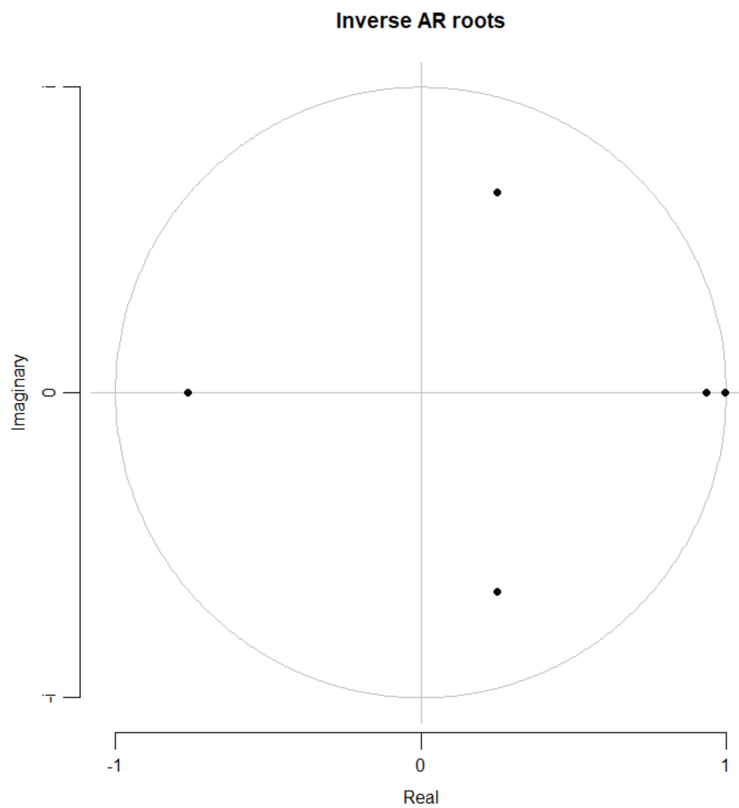
With a p-value of $4.121e^{-5}$, and a Significance Level α = 0.05, the p-value is significantly less than the Significance Level α, which means the Null Hypothesis can be rejected and it can be concluded that the residuals are white-noise.

Another Hypothesis Tests to run for Model3 is the Coefficient Test, which is a Hypothesis Test that seeks to determine which Autoregressive and/or Moving Average components of a model are statistically significant coefficients. Running the Coefficient Test in R for Model3 using the command "coeftest(Model3)," the following output is produced.

```
> coeftest(Model3)

z test of coefficients:

           Estimate Std. Error z value  Pr(>|z|)    
ar1        1.670729   0.039659  42.128 < 2.2e-16 ***
ar2       -0.538620   0.043379 -12.417 < 2.2e-16 ***
ar3       -0.405648   0.030896 -13.129 < 2.2e-16 ***
ar4        0.621058   0.035815  17.340 < 2.2e-16 ***
ar5       -0.347929         NA      NA        NA    
intercept  5.245695         NA      NA        NA    
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Running the Coefficient Test for Model3, the Autoregressive terms ar1, ar2, ar3, and ar4 are all deemed to be statistically significant coefficients because all of their p-values, equivalent to $2.2e^{-16}$ are all significantly less than the Significance Level α = 0.05, which at that level can all be rejected. The Autoregressive terms ar5 is not a statistically significant variable because its value according to R's calculations is "NA," or "Not Available." As stated earlier, the fifth Autoregressive term cannot be calculated due to a mathematical calculation error. Analysis of ar5, therefore, is ignored, and the variable ar5 should be removed from the model. For future analysis, Model3 will only have 4 Autoregressive components to it: ar1, ar2, ar3, and ar4, with ar5 being ignored from future analysis.

A plot of Model3 is provided to determine whether or not the data appears to be weakly stationary:



Except for the point that lies on the edge of the circle on the far right (that point is the ar5 term we are ignoring), the data appears to be weakly stationary.

The analysis continues through the construction of another ARIMA model, which shall be named Model4. Model4 uses the command "auto.arima," allowing for R to automatically determine the best ARIMA model automatically. Summary statistics of Model 4 are provided, and the information criterion the best ARIMA model R provides is based off the Akaike Information Criteria (AIC).

```
> summary(Model4)
Series: logBHPI
ARIMA(1,2,3)

Coefficients:
          ar1     ma1      ma2
      -0.4248  0.0787  -0.2197
s.e.   0.0865  0.0766   0.0492
          ma3
      -0.5112
s.e.   0.0463

sigma^2 estimated as 1.441e-05:  log likelihood=1577.4
AIC=-3144.8   AICc=-3144.64   BIC=-3125.11

Training set error measures:
                        ME        RMSE
Training set -2.887823e-05 0.003765907
                     MAE           MPE
Training set 0.002795668 -0.0003022996
                   MAPE       MASE
Training set 0.05862185 0.04817159
                    ACF1
Training set -0.01697067
```

The best ARIMA model R has generated given the AIC information criteria to base its best ARIMA off of has 1 Autoregressive term, 3 Moving Average terms, and requires 2 degrees of differencing. The AIC of Model4 is -3126.01, the BIC value is -3106.32, the RMSE is equivalent to 0.0038, and the MAE value is 0.0028.

Another model will be generated through the R command “auto.arima,” but instead it will **NOT** incorporate backward stepwise procedures to produce the best model and will **NOT** take seasonal adjustments of the data into account either. Summary statistics of this model, named Model5, have been provided for interpretation.

```
> summary(Model5)
Series: logBHPI
ARIMA(0,2,4)

Coefficients:
          ma1      ma2      ma3     ma4
      -0.3548   -0.127  -0.4663  0.2179
s.e.   0.0503    0.048   0.0453  0.0528

sigma^2 estimated as 1.439e-05:  log likelihood=1577.71
AIC=-3145.42   AICc=-3145.25   BIC=-3125.73

Training set error measures:
                         ME        RMSE         MAE
Training set -3.083036e-05 0.003762888 0.002806321
                        MPE       MAPE        MASE
Training set -0.0003540192 0.05879487 0.04835515
                    ACF1
Training set -0.01054443
```

Model5, through R's "auto.arima" command, has incorporated no Autoregressive terms, 4 Moving Average terms, and requires 2 degrees of differencing. The AIC of Model5 is -3145.42, its BIC is -3125.73, its RMSE is 0.0038, and its MAE is 0.0028.

It is now necessary to run Box Tests on Model4 and Model5 in order to test whether their residuals are white-noise. The hypotheses being tested are as follows:

$$H_0 = Residuals\ are\ White\ Noise$$

$$H_A = Residuals\ are\ NOT\ White\ Noise$$

The Box Test outputs for Model4 and Model 5 are below for further analysis.

```
> Box.test(Model4$residuals, 12, "Lj", 5)

        Box-Ljung test

data:  Model4$residuals
X-squared = 8.8469, df = 7, p-value = 0.2638

> Box.test(Model5$residuals, 12, "Lj", 5)

        Box-Ljung test

data:  Model5$residuals
X-squared = 8.3702, df = 7, p-value = 0.3011
```

Examining the output of the Box Tests for Model4 and Model5, the p-value of the Box Test for Model4 is 0.2638 and the p-value of 0.3011. Model4's and Model5's p-values are higher than the specified Significance Level $\alpha = 0.05$, meaning that for both Model4 and Model5, we fail to reject the Null Hypothesis and conclude that the residuals in Model4 and Model5 are not white-noise, meaning that other models should be considered and that the command "auto.arima" failed to produce an adequate model of consideration.

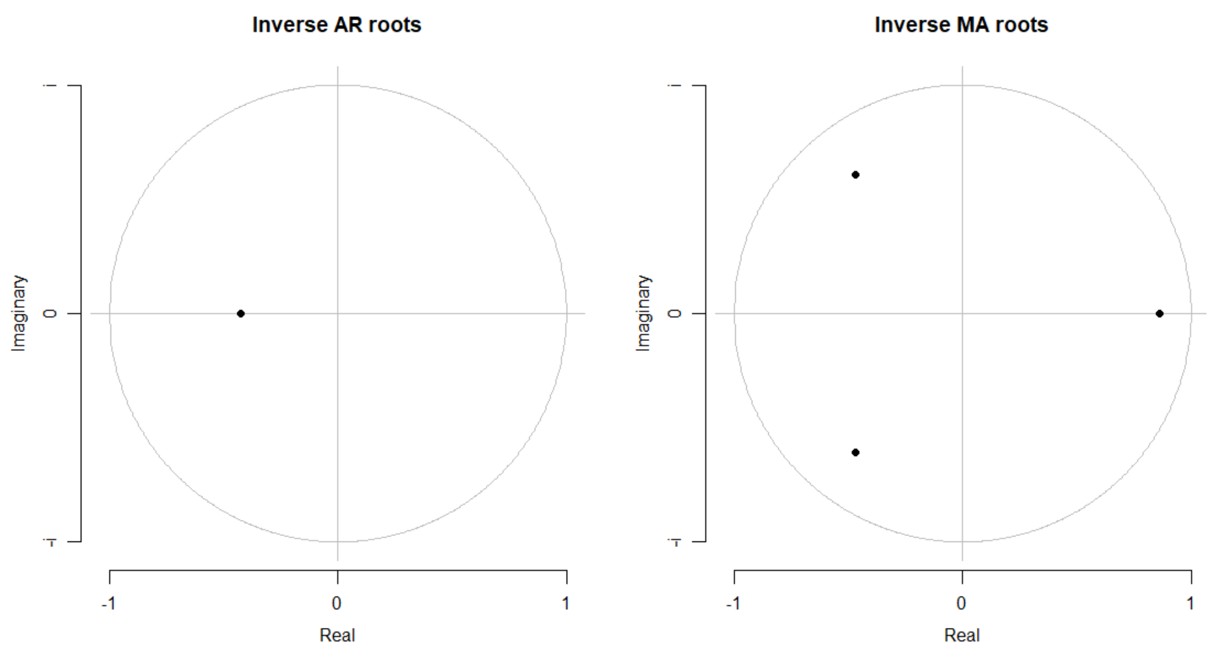
Now it is time to execute Coefficient Tests on Model4 and Model5 to determine which coefficients are statistically significant at the 5% Significance Level. The outputs for the Coefficient Tests of Model4 and Model5 are provided.

```
> coeftest(Model4)

z test of coefficients:

     Estimate Std. Error  z value  Pr(>|z|)    
ar1 -0.424801   0.086476  -4.9124 8.999e-07 ***
ma1  0.078683   0.076581   1.0274    0.3042    
ma2 -0.219749   0.049220  -4.4646 8.022e-06 ***
ma3 -0.511214   0.046336 -11.0329 < 2.2e-16 ***
---
Signif. codes:  
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> coeftest(Model5)

z test of coefficients:

     Estimate Std. Error  z value  Pr(>|z|)    
ma1 -0.354767   0.050278  -7.0560 1.713e-12 ***
ma2 -0.127044   0.048041  -2.6445  0.008181 ** 
ma3 -0.466266   0.045276 -10.2983 < 2.2e-16 ***
ma4  0.217867   0.052832   4.1237 3.728e-05 ***
---
Signif. codes:  
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

With respect to Model4, ar1, ma2, and ma3 are the statistically significant coefficients in the model, with their p-values being $8.999e^{-7}$, $8.022e^{-6}$, and $2.2e^{-16}$, all of which are significantly less than the Significance Level α = 0.05. Because these p-values are so small and are clearly less than α, the Null Hypothesis can be rejected for these coefficients in particular and it can be concluded that ar1, ma2, and ma3 are statistically significant coefficients. One of the terms, ma1, has a p-value of 0.3042, which is greater than α, and because the p-value is greater than α, the Null Hypothesis cannot be rejected for ma1 and it can be concluded that ma1 is **NOT** a statistically significant coefficient.

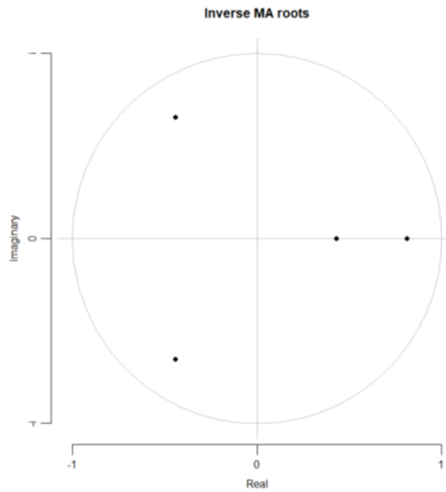
With respect to Model5, all the Moving Average terms (ma1, ma2, ma3, and ma4) have p-values that are less than the Significance Level α. The p-values for ma1, ma2, ma3, and ma4 are $1.713e^{-12}, 0.008181, 2.2e^{-16}$, and $3.728e^{-5}$, respectively. Because all of these coefficients have p-values less than α, the null hypothesis can be rejected and it can be concluded that the terms ma1, ma2, ma3, and ma4 are all statistically significant variables. Therefore, it appears that Model5 is a statistically significant model.

Now, Model4 and Model5 will be plotted to assess whether the data is weakly stationary. The first plot is the plot of Model4, and the second plot is the plot of Model5.



The plot of Model4 suggests that the data is weakly stationary, as all the Autoregressive and Moving Average components that are represented by the black dots lie within their respective circles.

The plot of Model5 is also provided.



Since all of the Moving Average terms represented by black dots lie within the circle, the data is considered to be weakly stationary.

Now let us go back to the data table and add Model4's and Model5's values to this data table for further analysis. Note that for Model3, the AR section has now been altered to 4 to reflect the accurate number of Autoregressive terms.

Model	**AR**	**MA**	**AIC**	**BIC**	**RMSE**	**MAE**
Arima1	7	0	-3153.38	-3117.72	0.00372	0.0028
Arima2	7	1	-3151.39	-3111.99	0.00372	0.0028
Arima3	7	2	-3150.67	-3107.32	0.00371	0.0028
Model3	4	0	-3136.72	-3109.12	0.00389	0.0029
Model4	1	3	-3144.64	-3125.11	0.00377	0.0028
Model5	0	4	-3145.42	-3125.73	0.00377	0.0028

Using this data table, we must select one Model to use for forecasting. The model I chose to select for this forecasting procedure is Arima1. This model yielded both the lowest AIC and BIC values, as well as maintaining a relatively low RMSE and MAE compared to the other models. Unfortunately, Model4 and Model5 failed the Box Test, meaning that their residuals were not white-noise and that Model4 and Model5 as a result should be reevaluated. However, I will also need Arima2 later on in the analysis, so the models I chose for forecasting and predicting are Arima1 and Arima2.

The first step that now must be performed is to divide Arima1 into a Training Dataset that consists of about 90% of the observations, or about 344 observations, and a Testing Dataset that contains the remaining 10%, or 36 observations. The Training Dataset is used to develop the model and the Testing Dataset validates the effectiveness of the model.

After partitioning the data into Training and Testing Datasets appropriately named Model1Training and Model1Forecast, respectively, summaries must be run on Model1Training and Model1Forecast to determine their summary statistics. Their outputs are below.

```
> summary(Model1Training)
Series: logBHPI[1:343]
ARIMA(7,0,0) with non-zero mean

Coefficients:
         ar1      ar2      ar3     ar4      ar5      ar6      ar7    mean
      1.6399  -0.5049  -0.4157  0.5951  -0.2656  -0.0194  -0.0298  5.1025
s.e.  0.0399   0.0773   0.0807  0.0834   0.1031   0.1057   0.0544     NaN

sigma^2 estimated as 1.577e-05:  log likelihood=1414.78
AIC=-2811.55   AICc=-2811.01   BIC=-2777.01

Training set error measures:
                       ME        RMSE        MAE         MPE       MAPE      MASE        ACF1
Training set 9.705347e-05 0.003924351 0.00294006 0.001859389 0.06209839 0.5058765 -0.01253786
```

```
> summary(Model1Forecast)
Series: logBHPI[344:381]
ARIMA(7,0,0) with non-zero mean

Coefficients:
         ar1     ar2      ar3     ar4     ar5      ar6     ar7    mean
      0.6233  0.1336  -0.3537  0.3567  0.0229  -0.1362  0.2509  0.0031
s.e.  0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000  0.0000

sigma^2 estimated as 1.413e-05:  log likelihood=-46.4
AIC=94.79   AICc=94.91   BIC=96.43

Training set error measures:
                    ME      RMSE       MAE     MPE    MAPE     MASE      ACF1
Training set 0.6774851 0.8007456 0.6774851 12.8526 12.8526 141.1862 0.2968443
```

With respect to Model1Training, its AIC is -2811.55, its BIC is -2777.01, its RMSE is 0.0039, and its MAE is 0.0029. Since it has a vast majority of the dataset, its AIC and BIC values should be reasonably large.

With respect to Model1Forecast, its AIC is 94.79, its BIC is 96.43, its RMSE is 0.8008, and its MAE is 0.6775. Since the Testing Dataset here has a very small number of observations, we expect AIC and BIC to be very low and expect RMSE and MAE to be relatively large due to the small record count.

A 1-Sample t-Test will now be run on Model1Forecast's residuals to ensure that the value of these residuals is not 0. The hypotheses being tested here are as follows:

$$H_0\colon Residuals = 0$$

$$H_A\colon Residuals \neq 0$$

The output for this 1-Sample t-Test for Model1Forecast$residuals is provided for further analysis.

```
> t.test(Model1Forecast$residuals)

        One Sample t-test

data:  Model1Forecast$residuals
t = 9.6542, df = 37, p-value = 1.187e-11
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.5352971 0.8196731
sample estimates:
mean of x
0.6774851
```

With a p-value of $1.187e^{-11}$, it is significantly less than α, so the Null Hypothesis can be rejected and it can be concluded that the true mean of the residuals is not equivalent to 0.

Now a Box Test must be performed on Model1Forecast's residuals to ensure that the residuals are white-noise. Given this, the hypotheses being tested here are the following:

$$H_0 = Residuals\ are\ White\ Noise$$

$$H_A = Residuals\ are\ NOT\ White\ Noise$$

The output for the Box Test on the residuals for Model1Forecast is incorporated below.

```
> Box.test(Model1Forecast$residuals, 12, "Lj")

        Box-Ljung test

data:  Model1Forecast$residuals
X-squared = 8.2292, df = 12, p-value = 0.767
```

The p-value of 0.767 is significantly greater than the Significance Level α, which means we fail to reject the Null Hypothesis and conclude that the residuals are **NOT** white noise. Although this is not promising in terms of the analysis, I would like to see the results of the Efficiency Test Model1Forecast. It should also be noted that due to the very small sample size in Model1Forecast that these conclusions are arising.

The Efficiency Test for Model1Forecast will now be conducted, which means a dependent variable Y and independent variable X must be identified. Let Y, or the dependent variable, be defined as the Model1Forecast residuals, known by variable name in R Model1Forecast$residuals. Let X, or the independent variable, be defined as the Model1Forecast fitted values, known by the variable name in R Model1Forecast$fitted. Running the Efficiency Test for Model1Forecast with Y being Model1Forecast$residuals, and X being Model1Forecast$fitted, the following output is produced.

```
> EfficiencyModel1 <- lm(Model1Forecast$residuals~Model1Forecast$fitted)
> summary(EfficiencyModel1)

Call:
lm(formula = Model1Forecast$residuals ~ Model1Forecast$fitted)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.062264 -0.046680  0.002199  0.038343  0.069794 

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)    
(Intercept)            5.00469    0.07910   63.27   <2e-16 ***
Model1Forecast$fitted -0.93897    0.01708  -54.97   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.04759 on 36 degrees of freedom
Multiple R-squared:  0.9882,	Adjusted R-squared:  0.9879 
F-statistic:  3021 on 1 and 36 DF,  p-value: < 2.2e-16
```

The predictor Model1Forecast$fitted has a p-value of $2e^{-16}$, which is significantly less than α, thereby suggesting to reject the Null Hypothesis and conclude that the predictor Model1Forecast$fitted is a statistically significant predictor in predicting Model1Forecast's residuals.

A Linear Hypothesis Test, known in R as a "lht," must be performed on Model1Forecast$fitted to determine if the restricted model is better in predicting the residuals in Model1Forecast than the model that incorporates the use of the Model1Forecast$fitted variable. The following hypotheses are tested:

H_0 = Restricted Model Predicts Model1Forecast$residuals Better than Full Model

H_A =Restricted Model Does **NOT** Predict Model1Forecast$residuals Better than Full Model

The output for the Linear Hypothesis Test is as follows.

```
> lht(EfficiencyModel1, "(Intercept)=Model1Forecast$fitted")
Linear hypothesis test

Hypothesis:
(Intercept) - Model1Forecast$fitted = 0

Model 1: restricted model
Model 2: Model1Forecast$residuals ~ Model1Forecast$fitted

  Res.Df    RSS Df Sum of Sq      F    Pr(>F)
1     37 8.7414
2     36 0.0815  1    8.6599 3823.9 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’
1
```

With a p=value of $2.2e^{-16}$, it is clearly less than the Significance Level α and it can be concluded that model that should be selected to predict Model1Forecast's residuals is the model that incorporates Model1Forecast's fitted values.

Now Arima2 will be used to forecast and predict just like Arima1. A Training and Testing Dataset must be defined, with about 90% of the observations, or about 344 observations, entering into a Training Dataset and the remaining 10%, or about 36, of the observations shall be developed into a Testing Dataset. The Training Dataset is used to develop the model and the Testing Dataset validates the effectiveness of the model.

After partitioning the data into Training and Testing Datasets appropriately named Model2Training and Model2Forecast, respectively, summaries must be run on Model2Training and Model2Forecast to determine their summary statistics. Their outputs are below.

```
> summary(Model2Training)
Series: logBHPI[1:343]
ARIMA(7,0,1) with non-zero mean

Coefficients:
         ar1     ar2      ar3     ar4     ar5      ar6
      0.7536  0.9888  -0.9251  0.2505  0.3321  -0.3985
s.e.  0.0219  0.0181   0.0139  0.0507  0.0541   0.0413
          ar7     ma1    mean
      -0.0023  0.9175  4.8485
s.e.   0.0548  0.0008     NaN

sigma^2 estimated as 1.542e-05:  log likelihood=1417.81
AIC=-2815.61   AICc=-2814.95   BIC=-2777.23

Training set error measures:
                       ME        RMSE         MAE
Training set 0.0002193735 0.003874946 0.002925965
                     MPE       MAPE      MASE         ACF1
Training set 0.004480872 0.06189668 0.5034512 -0.009545217


> summary(Model2Forecast)
Series: logBHPI[344:381]
ARIMA(7,0,1) with non-zero mean

Coefficients:
        ar1     ar2      ar3     ar4     ar5      ar6
      0.657  0.1124  -0.3586  0.3661  0.0132  -0.1382
s.e.  0.000  0.0000   0.0000  0.0000  0.0000   0.0000
         ar7     ma1    mean
      0.2501  -0.036  0.0031
s.e.  0.0000   0.000  0.0000

sigma^2 estimated as 1.417e-05:  log likelihood=-46.29
AIC=94.59   AICc=94.7   BIC=96.22

Training set error measures:
                   ME      RMSE      MAE      MPE     MAPE
Training set 0.674155 0.7985695 0.674155 12.78988 12.78988
                 MASE      ACF1
Training set 140.4922 0.2992321
```

With respect to Model2Training, its AIC is -2815.61, its BIC is -2777.23, its RMSE is 0.0039, and its MAE is 0.0029. Since it has a vast majority of the dataset, its AIC and BIC values should be reasonably large.

With respect to Model2Forecast, its AIC is 94.59, its BIC is 96.22, its RMSE is 0.7986, and its MAE is 0.6742. Since the Testing Dataset here has a very small number of observations, we expect AIC and BIC to be very low and expect RMSE and MAE to be relatively large due to the small record count.

A 1-Sample t-Test will now be run on Model2Forecast's residuals to ensure that the value of these residuals is not 0. The hypotheses being tested here are as follows:

$$H_0: Residuals = 0$$

$$H_A: Residuals \neq 0$$

The output for this 1-Sample t-Test for Model1Forecast$residuals is provided for further analysis.

```
> t.test(Model2Forecast$residuals)

        One Sample t-test

data:  Model2Forecast$residuals
t = 9.58, df = 37, p-value = 1.458e-11
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.5315692 0.8167407
sample estimates:
mean of x
 0.674155
```

With a p-value of $1.458e^{-11}$, it is significantly less than α, so the Null Hypothesis can be rejected and it can be concluded that the true mean of the residuals is not equivalent to 0.

Now a Box Test must be performed on Model2Forecast's residuals to ensure that the residuals are white-noise. Given this, the hypotheses being tested here are the following:

$$H_0 = Residuals\ are\ White\ Noise$$

$$H_A = Residuals\ are\ NOT\ White\ Noise$$

The output for the Box Test on the residuals for Model1Forecast is incorporated below.

```
> Box.test(Model2Forecast$residuals, 12, "Lj")

        Box-Ljung test

data:  Model2Forecast$residuals
X-squared = 8.4043, df = 12, p-value = 0.7528
```

The p-value of 0.75287 is significantly greater than the Significance Level α, which means we fail to reject the Null Hypothesis and conclude that the residuals are **NOT** white noise. Although this is not promising in terms of the analysis, I would like to see the results of the Efficiency Test Model2Forecast. It should also be noted that due to the very small sample size in Model2Forecast that these conclusions are arising.

The Efficiency Test for Model2Forecast will now be conducted, which means a dependent variable Y and independent variable X must be identified. Let Y, or the dependent variable, be defined as the Model2Forecast residuals, known by variable name in R Model2Forecast$residuals. Let X, or the independent variable, be defined as the Model2Forecast fitted values, known by the variable name in R Model2Forecast$fitted. Running the Efficiency Test for Model2Forecast with Y being Model2Forecast$residuals, and X being Model2Forecast$fitted, the following output is produced.

```
> EfficiencyModel2 <- lm(Model2Forecast$residuals~Model2Forecast$fitted)
> summary(EfficiencyModel2)

Call:
lm(formula = Model2Forecast$residuals ~ Model2Forecast$fitted)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.062020 -0.046638  0.002133  0.038270  0.070121 

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)            5.00436    0.07884   63.48   <2e-16
Model2Forecast$fitted -0.93895    0.01701  -55.19   <2e-16
                         
(Intercept)           ***
Model2Forecast$fitted ***
---
Signif. codes:  
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04753 on 36 degrees of freedom
Multiple R-squared:  0.9883,	Adjusted R-squared:  0.988 
F-statistic:  3046 on 1 and 36 DF,  p-value: < 2.2e-16
```

The predictor Model2Forecast$fitted has a p-value of $2e^{-16}$, which is significantly less than α, thereby suggesting to reject the Null Hypothesis and conclude that the predictor Model2Forecast$fitted is a statistically significant predictor in predicting Model2Forecast's residuals.

A Linear Hypothesis Test must be performed on Model2Forecast$fitted to determine if the restricted model is better in predicting the residuals in Model2Forecast than the model that incorporates the use of the Model2Forecast$fitted variable. The following hypotheses are tested:

H_0 = Restricted Model Predicts Model2Forecast$residuals Better than Full Model

H_A =Restricted Model Does **NOT** Predict Model2Forecast$residuals Better than Full Model

The output for the Linear Hypothesis Test is as follows.

```
> lht(EfficiencyModel2, "(Intercept)=Model2Forecast$fitted")
Linear hypothesis test

Hypothesis:
(Intercept) - Model2Forecast$fitted = 0

Model 1: restricted model
Model 2: Model2Forecast$residuals ~ Model2Forecast$fitted

  Res.Df    RSS Df Sum of Sq    F    Pr(>F)
1     37 8.7794
2     36 0.0813  1    8.6981 3850 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

With a p=value of $2.2e^{-16}$, it is clearly less than the Significance Level α and it can be concluded that model that should be selected to predict Model2Forecast’s residuals is the model that incorporates Model2Forecast’s fitted values.

Now it is time to shift the focus in analyzing the residuals. Suppose the variable “e1” is defined in R to be Model1Forecast$residuals and “e2” is defined in R to be Model2Forecast$residuals. Then the square of the residuals of these models would simply inherit the names “e1sq” and “e2sq,” respectively.

The mean value of the residuals in Model1Forecast is 0.6775, with its square being 0.6412. The mean value of the residuals in Model2Forecast is 0.6742, with its square being 0.6377.

Suppose we wish to run a 1-Sample t-Test on the residuals of Model1Forecast and Model2Forecast e1 and e2, respectively. Conducting the t-Tests produces the following output:

```
> t.test(e1)

        One Sample t-test

data:  e1
t = 9.6542, df = 37, p-value = 1.187e-11
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.5352971 0.8196731
sample estimates:
mean of x
0.6774851

> t.test(e2)

        One Sample t-test

data:  e2
t = 9.58, df = 37, p-value = 1.458e-11
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.5315692 0.8167407
sample estimates:
mean of x
 0.674155
```

According to the t-Test on Model1Forecast$residuals, the p-value of $1.187e^{-11}$ is significantly less than α, meaning that we can reject the Null Hypothesis and conclude that the value of the residuals in Model1Forecast differs from 0.

According to the t-Test on Model2Forecast$residuals, the p-value of $1.458e^{-11}$ is significantly less than α, meaning that we can reject the Null Hypothesis and conclude that the value of the residuals in Model2Forecast also differs from 0.

Now suppose that we wish to execute a Box Test of the residuals found in Model1Forecast and Model2Forecast but wanted to determine whether or not the residuals were considered to be white-noise. The Box Test aims to clarify whether these residuals are white-noise or not, with the following hypotheses being tested:

$$H_0 = Residuals\ are\ White\ Noise$$

$$H_A = Residuals\ are\ NOT\ White\ Noise$$

The Box Tests for the Model1Forecast$resiuals and Model2Forecast$residuals, or e1 and e2, respectively, have been conducted and have the following outputs.

```
> Box.test(e1, 12, "Lj")

        Box-Ljung test

data:  e1
X-squared = 8.2292, df = 12, p-value = 0.767

> Box.test(e2, 12, "Lj")

        Box-Ljung test

data:  e2
X-squared = 8.4043, df = 12, p-value = 0.7528
```

Looking at the Box Tests, the p-value for the Model1Forecast$residual is 0.767 and the p-value for the Model2Forecast$residual is 0.7528. Both p-values are above the significance level α, and because of this, we fail to reject the Null Hypothesis and conclude that the residuals in both Model1Forecast and Model2Forecast are white-noise.

The Efficiency Test for both the residuals found in Model1Forecast and Model2Forecast will now be conducted, meaning that a dependent variable Y and independent variable X must be identified. Let Y, or the dependent variable, be defined as the Model1Forecast residuals or the Model2Forecast residuals, defined as either Model1Forecast$residuals or Model2Forecast$residuals, respectively. Let X, or the independent variable, be defined as the Model1Forecast fitted values or Model2Forecast fitted values, known by the variable name in R Model1Forecast$fitted and Model2Forecast$fitted, respectively. Running the Efficiency Test for both Model1Forecast and Model2Forecast with Y being either Model1Forecast$resiudals or Model2Forecast$residuals, pending on what number model selected is analyzed, and X being either Model1Forecast$fitted or Model2Forecast$fitted, pending on the number of the model selected. The output is available for analysis.

```
> eft <- lm(e1~Model1Forecast$fitted)
> eft2 <- lm(e2~Model2Forecast$fitted)
> summary(eft)

Call:
lm(formula = e1 ~ Model1Forecast$fitted)

Residuals:
      Min        1Q    Median        3Q       Max
-0.062264 -0.046680  0.002199  0.038343  0.069794

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)            5.00469    0.07910   63.27   <2e-16 ***
Model1Forecast$fitted -0.93897    0.01708  -54.97   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.04759 on 36 degrees of freedom
Multiple R-squared:  0.9882,    Adjusted R-squared:  0.9879
F-statistic:  3021 on 1 and 36 DF,  p-value: < 2.2e-16

> summary(eft2)

Call:
lm(formula = e2 ~ Model2Forecast$fitted)

Residuals:
      Min        1Q    Median        3Q       Max
-0.062020 -0.046638  0.002133  0.038270  0.070121

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)            5.00436    0.07884   63.48   <2e-16 ***
Model2Forecast$fitted -0.93895    0.01701  -55.19   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.04753 on 36 degrees of freedom
Multiple R-squared:  0.9883,    Adjusted R-squared:  0.988
F-statistic:  3046 on 1 and 36 DF,  p-value: < 2.2e-16
```

Analyzing the output, the Efficiency Test for Model1Forecast$residuals suggests that the model that incorporates the use of the predictor Model1Forecast$fitted is the stronger model to predict Model1Forecast$residuals because its p-value of $2e^{-16}$ is far lower than the significance level α, suggesting that the model incorporating the term Model1Forecast$fitted is a statistically stronger model to predict Model1Forecast$residuals than the restricted model.

Analyzing the output, the Efficiency Test for Model2Forecast$residuals suggests that the model that incorporates the use of the predictor Model2Forecast$fitted is the stronger model to predict Model2Forecast$residuals because its p-value of $2e^{-16}$ is far lower than the significance level α, suggesting that the model incorporating the term Model2Forecast$fitted is a statistically stronger model to predict Model2Forecast$residuals than the restricted model.

A Linear Hypothesis Test must be performed on Model1Forecast$residuals, or e1, to determine if the restricted model is better in predicting the residuals in Model1Forecast than the model that incorporates the use of the Model1Forecast$fitted variable. The following hypotheses are tested:

$H_0 =$ Restricted Model Predicts Model1Forecast$residuals Better than Full Model
$H_A =$Restricted Model Does **NOT** Predict Model1Forecast$residuals Better than Full Model

Output is provided for further analysis.

```
> lht(eft, c("(Intercept)", "Model1Forecast$fitted"))
Linear hypothesis test

Hypothesis:
(Intercept) = 0
Model1Forecast$fitted = 0

Model 1: restricted model
Model 2: e1 ~ Model1Forecast$fitted

  Res.Df     RSS Df Sum of Sq      F    Pr(>F)
1     38 24.3654
2     36  0.0815  2    24.284 5361.5 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

With a p=value of $2.2e^{-16}$, it is clearly less than the Significance Level α and it can be concluded that model that should be selected to predict Model1Forecast’s residuals is the model that incorporates Model1Forecast’s fitted values.

Assume that there is a linear relationship shared between the residuals of Model1Forecast and Model2Forecast. The variable name “linear” takes the difference between Model1Forecast’s residuals and Model2Forecast’s residuals and establishes a linear relationship between the two. A 1-Sample t-Test will now be conducted from “linear” to determine if the difference between the residuals found in Model1Forecast and Model2Forecast differs from the value of 0:

$$H_0 = e1 - e2 = 0$$

$$H_A = e1 - e2 \neq 0$$

The output of the 1-Sample t-Test is provided.

```
> t.test(linear)

	One Sample t-test

data:  linear
t = 9.7474, df = 37, p-value = 9.172e-12
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.002637905 0.004022379
sample estimates:
  mean of x
0.003330142
```

The 1-Sample t-Test shows that there is a difference between the residuals of Model1Forecast and Model2Forecast since the p-value of $9.172e^{-12}$ is far lower than the specified Significance Level α, leading us to reject the Null Hypothesis and conclude that the difference between the residuals of Model1Forecsat and Model2Forecast is nonzero.

Another consideration comes to mind when the residuals are squared and then differenced from each other. Using the variable name "quad" to identify the difference of the squares of the residuals between Model1Forecast and Model2Forecast, we wish to run a 1-Sample t-Test to determine if the difference between the squares of the residuals of Model1Forecast and Model2Forecast are nonzero:

$$H_0 = e1sq - e2sq = 0$$

$$H_A = e1sq - e2sq \neq 0$$

The output for the 1-Sample t-Test is provided.

```
> quad <- e1sq - e2sq
> t.test(quad)

        One Sample t-test

data:  quad
t = 7.6718, df = 37, p-value = 3.679e-09
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.002561088 0.004399420
sample estimates:
  mean of x
0.003480254
```

Examining the 1-Sample t-Test, the p-value of $3.697e^{-9}$ is much lower than the specified Significance Level α, allowing us to reject the Null Hypothesis and to conclude that the difference of the squares of the residuals between Model1Forecast and Model2Forecast is nonzero.

Before running the Unit Root Dickey Fuller Test, the Efficiency Test concerning the Testing Dataset of the ln(BHPI) must be executed to determine whether the restricted model or the model incorporating the use of the term Model2Forecast$fitted predicts the ln(BHPI) Testing Dataset better:

$$H_0 = Restricted\ Model\ Predicts\ \ln(BHPI)$$

$$H_A = Model\ with\ Predictor\ Model2Forecasted\$fitted\ Predicts\ \ln(BHPI)$$

Running the Efficiency Test for this variable in R provides the following output:

```
> eff <- lm(logBHPI[344:381]~Model2Forecast$fitted)
> summary(eff)

Call:
lm(formula = logBHPI[344:381] ~ Model2Forecast$fitted)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.062020 -0.046638  0.002133  0.038270  0.070121 

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)    
(Intercept)            5.00436    0.07884  63.475  < 2e-16 ***
Model2Forecast$fitted  0.06105    0.01701   3.589 0.000982 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.04753 on 36 degrees of freedom
Multiple R-squared:  0.2635,	Adjusted R-squared:  0.243 
F-statistic: 12.88 on 1 and 36 DF,  p-value: 0.000982

> lht(eff, c("(Intercept)", "Model2Forecast$fitted"))
Linear hypothesis test

Hypothesis:
(Intercept) = 0
Model2Forecast$fitted = 0

Model 1: restricted model
Model 2: logBHPI[344:381] ~ Model2Forecast$fitted

  Res.Df     RSS Df Sum of Sq      F    Pr(>F)    
1     38 1061.87                                  
2     36    0.08  2    1061.8 234984 < 2.2e-16 ***
```

The output above suggests that the inclusion of the predictor Model2Forecasted$fitted predicts values of ln(BHPI) better than the restricted model, as the low p-value of $2.2e^{-16}$ suggests that because it clearly is less than the Significance Level α, we can reject the Null Hypothesis and conclude that the inclusion of the predictor Model2Forecast$fitted predicts values of ln(BHPI) better than the restricted model.

It is now time to execute a Unit Root Dickey Fuller Test on the variable ln(BHPI), assuming that there is a trend in the data and AIC is the selection criterion to utilize. The following hypotheses are tested:

$$H_0 = \text{There is a Unit Root}$$

H_A =There is **NOT** a Unit Root

Running the Unit Root Dickey Fuller Test through R yields the following results.

```
> library(urca)
> t1 <- ur.df(logBHPI, "trend", 12, "AIC")
> summary(t1)

###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0126162 -0.0021383 -0.0001102  0.0022280  0.0195926
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  1.279e-02  5.074e-03   2.521   0.0122 *  
z.lag.1     -3.108e-03  1.244e-03  -2.498   0.0129 *  
tt           1.171e-05  4.812e-06   2.433   0.0155 *  
z.diff.lag1  6.157e-01  5.075e-02  12.133  < 2e-16 ***
z.diff.lag2  1.323e-01  5.983e-02   2.211   0.0277 *  
z.diff.lag3 -3.620e-01  6.025e-02  -6.008 4.63e-09 ***
z.diff.lag4  3.660e-01  6.015e-02   6.084 3.01e-09 ***
z.diff.lag5  2.499e-02  6.029e-02   0.414   0.6788    
z.diff.lag6 -1.394e-01  5.989e-02  -2.327   0.0205 *  
z.diff.lag7  2.695e-01  5.085e-02   5.299 2.04e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.003774 on 358 degrees of freedom
Multiple R-squared:  0.6618,	Adjusted R-squared:  0.6533 
F-statistic: 77.83 on 9 and 358 DF,  p-value: < 2.2e-16


Value of test-statistic is: -2.498 2.6219 3.1765 

Critical values for test statistics: 
      1pct  5pct 10pct
tau3 -3.98 -3.42 -3.13
phi2  6.15  4.71  4.05
phi3  8.34  6.30  5.36
```

Examining the output, the test statistic is -2.498, and at the 5% Significance Level, tau3 is -3.42. Because our test statistic is not as extreme than the tau3 test statistic at 5% Significance, we fail to reject the Null Hypothesis and conclude that there is a unit root, meaning that this relationship is at least I(1).

Now we take the first difference of ln(BHPI) and adjust the “trend” argument in the “ur.df” command to “drift,” which includes a constant term and plotting the first difference of the ln(BHPI) would show no trend anyway.

Running the Unit Root Dickey-Fuller Test again in R yields us the following output.

```
> t1 <- ur.df(diff(logBHPI), "drift", 12, "AIC")
> summary(t1)

###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression drift


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0123962 -0.0022122 -0.0000532  0.0022803  0.0200991

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.0002808  0.0002266   1.239   0.2161
z.lag.1     -0.0967527  0.0380213  -2.545   0.0114 *
z.diff.lag1 -0.2779694  0.0580113  -4.792 2.43e-06 ***
z.diff.lag2 -0.1438044  0.0587004  -2.450   0.0148 *
z.diff.lag3 -0.5052682  0.0579051  -8.726  < 2e-16 ***
z.diff.lag4 -0.1407680  0.0553523  -2.543   0.0114 *
z.diff.lag5 -0.1154694  0.0543653  -2.124   0.0344 *
z.diff.lag6 -0.2575282  0.0510355  -5.046 7.18e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.003801 on 359 degrees of freedom
Multiple R-squared:  0.3069,    Adjusted R-squared:  0.2934
F-statistic: 22.71 on 7 and 359 DF,  p-value: < 2.2e-16


Value of test-statistic is: -2.5447 3.2378

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.44 -2.87 -2.57
phi1  6.47  4.61  3.79
```

Examining the output, the test statistic is -2.5447 and at the 5% Significance Level, tau2 is -2.87. Because our test statistic is not as extreme as the tau2 test statistic at 5% Significance shows that we fail to reject the Null Hypothesis and conclude that there is another unit root, meaning that this relationship is now at least I(2).

Another Unit Root Dickey-Fuller Test shall be conducted to see if the second difference of the ln(BHPI) has a Unit Root. The following output is produced.

```
> t1 <- ur.df(diff(diff(logBHPI)), "drift", 12, "AIC")

> summary(t1)

###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression drift


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
      Min        1Q    Median        3Q       Max
-0.012872 -0.002102 -0.000097  0.002280  0.020769

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.176e-05  2.002e-04   0.059    0.953
z.lag.1     -2.745e+00  1.987e-01 -13.811  < 2e-16 ***
z.diff.lag1  1.391e+00  1.764e-01   7.884 3.85e-14 ***
z.diff.lag2  1.179e+00  1.473e-01   8.006 1.66e-14 ***
z.diff.lag3  6.126e-01  1.137e-01   5.387 1.30e-07 ***
z.diff.lag4  4.322e-01  8.464e-02   5.106 5.34e-07 ***
z.diff.lag5  2.816e-01  5.067e-02   5.559 5.31e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.003831 on 359 degrees of freedom
Multiple R-squared:  0.7229,    Adjusted R-squared:  0.7183
F-statistic: 156.1 on 6 and 359 DF,  p-value: < 2.2e-16


Value of test-statistic is: -13.8106 95.3683

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.44 -2.87 -2.57
phi1  6.47  4.61  3.79
```

Examining the output, the test statistic is -13.8106 and at the 5% Significance Level, tau2 is equivalent to -2.87. Because the test statistic is more extreme than the tau2 test statistic at 5% Significance, we can reject the Null Hypothesis and conclude that there no other Unit Root and the model we have currently (the second differenced model of ln(BHPI)) is stationary and the model is I(2).

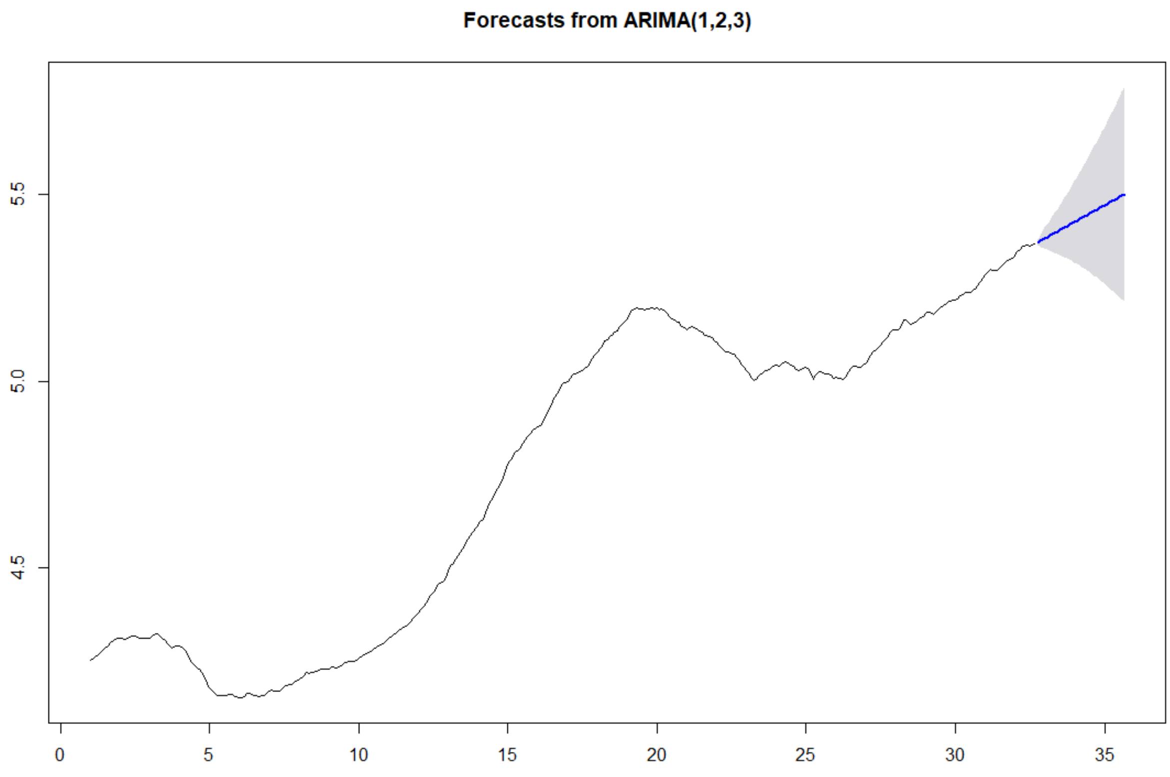
Taking a look at the other data table of all the models produced so far, this is what we have.

Model	**AR**	**MA**	**AIC**	**BIC**	**RMSE**	**MAE**
Arima1	7	0	-3153.38	-3117.72	0.00372	0.0028
Arima2	7	1	-3151.39	-3111.99	0.00372	0.0028
Arima3	7	2	-3150.67	-3107.32	0.00371	0.0028
Model3	4	0	-3136.72	-3109.12	0.00389	0.0029
Model4	1	3	-3144.64	-3125.11	0.00377	0.0028
Model5	0	4	-3145.42	-3125.73	0.00377	0.0028

Using strictly the AIC as the selection criterion from this table, I select Model4 because it has the lowest AIC amongst all the other models in the table while providing a low RMSE and MAE (rounded to 3 decimals).

From this model, we wish to run an ARIMA but include 2 integrated terms, meaning that the ARIMA model is integrated twice. In R, Model4 would now be an ARIMA model where it uses the ln(BHPI) but has 1 Autoregressive term, 3 Moving Average terms, and is integrated twice.

From here, we would like to forecast Model4 36 time periods ahead (3 years) at 95% confidence. Plotting the forecast, labeled as Forecast4 in R, reveals the following plot:



The plot here identifies that over a forecast that projects ln(BHPI) 36 time periods ahead, the data appears to be headed in a positive trend that looks like it is exceeding the value of 5.5. The gray shaded area around the blue line (identifying the forecast) indicates the 95% Confidence Interval of the potential values the forecast can take.

Now it is time to make a dataframe named "data_frame" containing the variables the first difference of the ln(BHPI) known in R as "difflog(BHPI)" and the first difference of the ln(US), known in R as "difflog(US)."

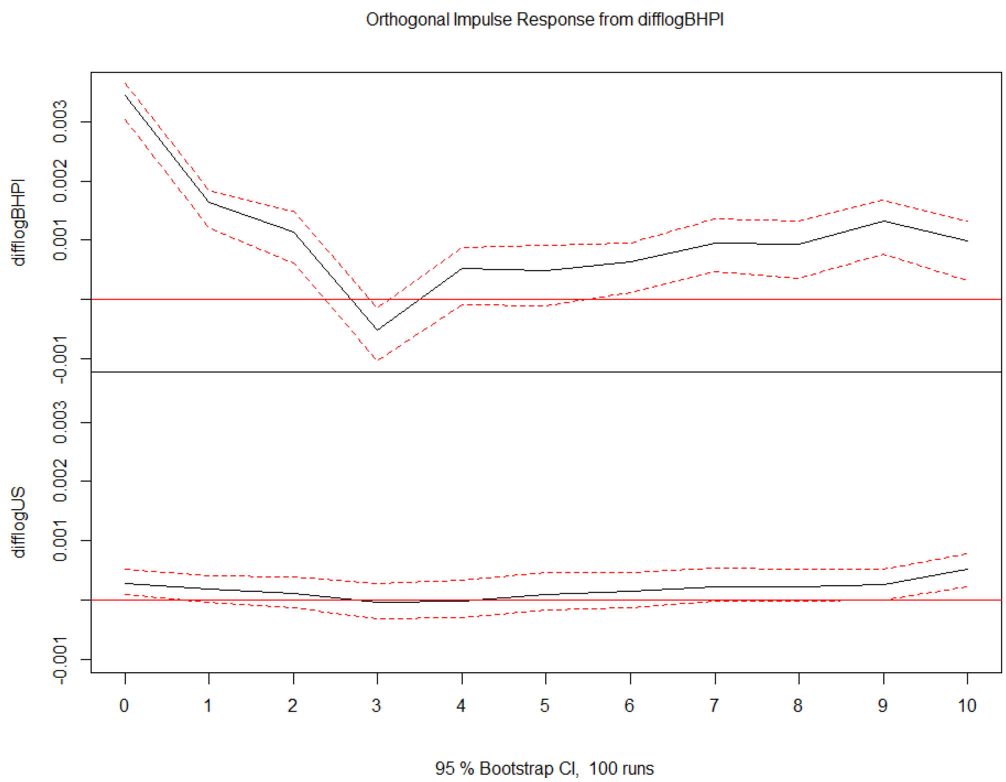
Form here, we wish to determine the appropriate number of lags necessary for future analysis. To determine the correct number of lags, the command “VARselect” from R prompts the dataframe to determine the appropriate number of lags given a selection of identification criteria such as AIC, BIC, and the like. For the sake of this report, the AIC criterion will be used to determine the number of lags appropriate for the model.

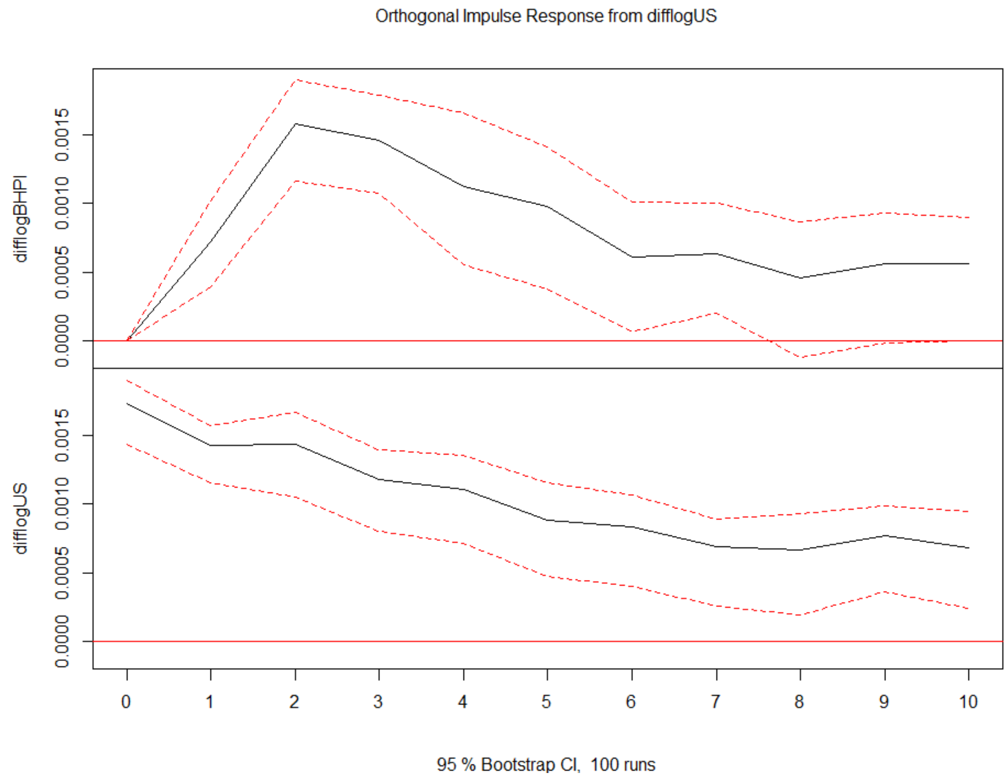
Running the command and focusing on the AIC criterion identifies that the correct lag length is 12 for a Vector Autoregressive Model, or a VAR. From here, it is imperative to run VAR for Model 4 assuming 12 lags as the AIC Criterion had determined from the “VARselect” command.

```
> US <- ts(BostonHousePriceIndex$US, c(1987, 1), frequency = 12)
> difflogUS <- diff(log(US))
> data_frame <- data.frame(difflogBHPI, difflogUS)
> VARselect(data_frame, 12)
$`selection`
AIC(n)  HQ(n)  SC(n) FPE(n)
    12      7      4     12

$criteria
                   1             2             3             4
AIC(n) -2.367638e+01 -2.367976e+01 -2.370070e+01 -2.382589e+01
HQ(n)  -2.365107e+01 -2.363756e+01 -2.364163e+01 -2.374995e+01
SC(n)  -2.361266e+01 -2.357356e+01 -2.355203e+01 -2.363474e+01
FPE(n)  5.217689e-11  5.200129e-11  5.092365e-11  4.493190e-11
                   5             6             7             8
AIC(n) -2.382798e+01 -2.383298e+01 -2.389648e+01 -2.389670e+01
HQ(n)  -2.373516e+01 -2.372328e+01 -2.376991e+01 -2.375325e+01
SC(n)  -2.359434e+01 -2.355687e+01 -2.357789e+01 -2.353562e+01
FPE(n)  4.483900e-11  4.461623e-11  4.187254e-11  4.186528e-11
                   9            10            11            12
AIC(n) -2.390714e+01 -2.391874e+01 -2.391930e+01 -2.393686e+01
HQ(n)  -2.374681e+01 -2.374154e+01 -2.372522e+01 -2.372591e+01
SC(n)  -2.350359e+01 -2.347271e+01 -2.343079e+01 -2.340587e+01
FPE(n)  4.143242e-11  4.095716e-11  4.093755e-11  4.022861e-11
```

With the proper number of lags selected according to the AIC Criterion, it is time to commence the Impulse Response Function (IRF) and Granger Causality Tests.

The IRF plots are provided.





According to the Impulse Response Function (IRF) plots, the BHPI does not seem to have an effect on the US Average HPI. Furthermore, the IRF plots show that the BHPI is affected by the US Average HPI, as would be anticipated. It happens with a lag of 3 periods yet still has a positive effect well over 6 time periods, showing that this effect lasts over several time periods. We need to confirm the findings produced from the IRF plots with the Granger Causality Test with respect to the first difference of the ln(US) variable.

The output for the Granger Causality test with respect to the variable difflog(US) is provided.

```
> causality(Model4, "difflogUS")
$`Granger`

        Granger causality H0: difflogUS do
        not Granger-cause difflogBHPI

data:  VAR object Model4
F-Test = 8.6646, df1 = 12, df2 = 686,
p-value = 1.776e-15


$Instant

        H0: No instantaneous causality
        between: difflogUS and difflogBHPI

data:  VAR object Model4
Chi-squared = 9.6534, df = 1, p-value
= 0.00189
```

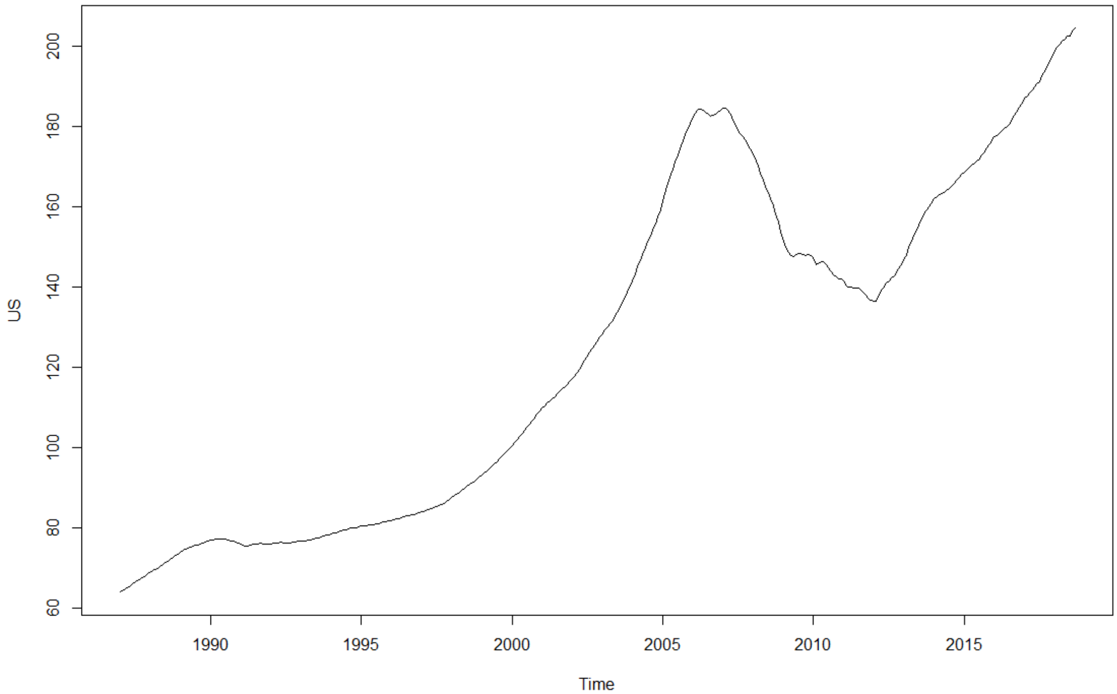
With the p-value equivalent to $1.776e^{-15}$^-15, this provides enough evidence at the specified Significance Level α to reject the Null Hypothesis and conclude that the US Average Housing Price Index has a Granger-Cause Effect on Boston Housing Price Index. In other words, the lags of US HPI is statistically significant in explaining the BHPI.

Now it is necessary to run the Granger Causality Test once again, but with the variable difflogBHPI, whose output is provided below.

```
> causality(Model4, "difflogBHPI")
$`Granger`

        Granger causality H0: difflogBHPI do not Granger-cause difflogUS

data:  VAR object Model4
F-Test = 0.82091, df1 = 12, df2 = 686, p-value = 0.629


$Instant

        H0: No instantaneous causality between: difflogBHPI and difflogUS

data:  VAR object Model4
Chi-squared = 9.6534, df = 1, p-value = 0.00189
```

The output above identifies that the p-value is 0.629, which fails to be less than the Significance Level α, which means we fail to reject the Null Hypothesis and conclude that the lags of BHPI are not statistically significant in explaining US HPI.

Next, an Augmented Dickey Fuller Test will be conducted for the variable difflogUS. Before the teset commences, we must plot the US HPI to confirm the supposed trend in the data, and the plot is provided below.



Indeed, there appears to be a positive trend in the data between January 1978 and 2018. With that being identified, the following hypotheses are being tested:

$$H_0 = There\ is\ a\ Unit\ Root$$

$$H_A = There\ is\ \boldsymbol{NOT}\ a\ Unit\ Root$$

Again, the output for this Augmented Dickey Fuller Test is provided for further analysis.

```
> t2 <- ur.df(logUS, "trend", 12, "AIC")
> summary(t2)

###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression trend


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
       Min         1Q     Median         3Q
-0.0086574 -0.0007367 -0.0000828  0.0006485
       Max
 0.0108649

Coefficients:
              Estimate Std. Error t value
(Intercept)  1.117e-02  3.302e-03   3.384
z.lag.1     -2.664e-03  7.911e-04  -3.367
tt           8.355e-06  2.563e-06   3.260
z.diff.lag1  8.163e-01  5.274e-02  15.477
z.diff.lag2  1.029e-01  6.876e-02   1.496
z.diff.lag3 -1.143e-01  6.970e-02  -1.640
z.diff.lag4  8.727e-02  6.998e-02   1.247
z.diff.lag5 -4.322e-02  7.010e-02  -0.617
z.diff.lag6  3.389e-02  6.999e-02   0.484
z.diff.lag7 -3.710e-02  6.976e-02  -0.532
z.diff.lag8  3.989e-02  6.958e-02   0.573
z.diff.lag9  8.261e-02  5.361e-02   1.541
            Pr(>|t|)
(Intercept) 0.000794 ***
z.lag.1     0.000842 ***
tt          0.001221 **
z.diff.lag1  < 2e-16 ***
z.diff.lag2 0.135558
z.diff.lag3 0.101831
z.diff.lag4 0.213141
z.diff.lag5 0.537862
```

```
z.diff.lag6 0.628540
z.diff.lag7 0.595149
z.diff.lag8 0.566838
z.diff.lag9 0.124214
---
Signif. codes:
  0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
  0.1 ' ' 1

Residual standard error: 0.001769 on 356 degrees of freedom
Multiple R-squared:  0.8791,    Adjusted R-squared:  0.8754
F-statistic: 235.4 on 11 and 356 DF,  p-value: < 2.2e-16


Value of test-statistic is: -3.3673 4.2219 5.7064

Critical values for test statistics:
      1pct  5pct 10pct
tau3 -3.98 -3.42 -3.13
phi2  6.15  4.71  4.05
phi3  8.34  6.30  5.36
```

The test statistic of -3.3673 is less than the tau3 Critical Value at 5% Significance of -3.42., leading us to fail to reject the Null Hypothesis and claim that it has a Unit Root that is at least I(1).

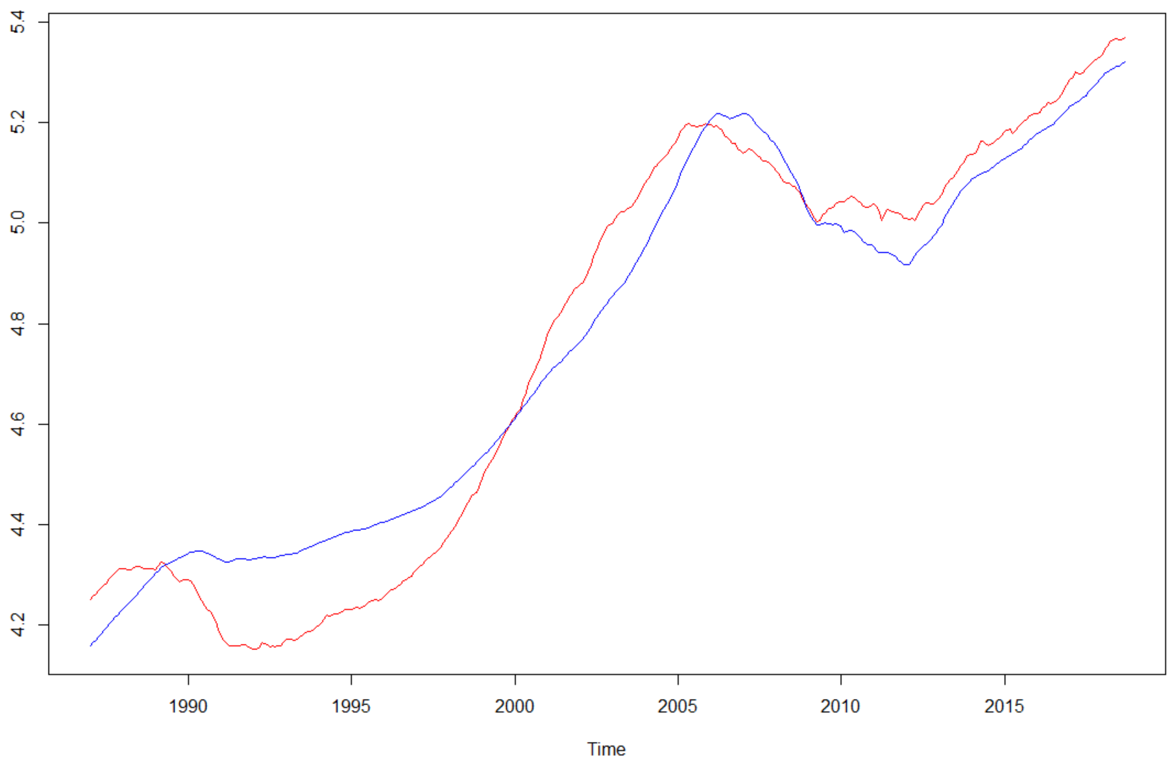
Another Augmented Dickey Fuller Test shall be conducted, but now with the first difference of the ln(US) and the argument “trend” changes to “drift.” The output for this test is provided below.

```
> t2 <- ur.df(difflogUS, "drift", 12, "AIC")
```

```
> summary(t2)

###############################################
# Augmented Dickey-Fuller Test Unit Root Test #
###############################################

Test regression drift


Call:
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0101685 -0.0006939  0.0000024  0.0006999  0.0109339

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.0001647  0.0001053   1.564  0.11865
z.lag.1      -0.0575057  0.0196065  -2.933  0.00358 **
z.diff.lag1  -0.0667485  0.0514865  -1.296  0.19567
z.diff.lag2   0.0146250  0.0517683   0.283  0.77772
z.diff.lag3  -0.1328964  0.0518545  -2.563  0.01080 *
z.diff.lag4  -0.0096937  0.0522527  -0.186  0.85293
z.diff.lag5  -0.0370486  0.0520765  -0.711  0.47729
z.diff.lag6  -0.0219982  0.0517994  -0.425  0.67133
z.diff.lag7  -0.0480723  0.0517319  -0.929  0.35339
z.diff.lag8  -0.0275941  0.0515890  -0.535  0.59307
z.diff.lag9   0.1131135  0.0514884   2.197  0.02868 *
z.diff.lag10  0.0101444  0.0508045   0.200  0.84185
z.diff.lag11 -0.0321525  0.0507525  -0.634  0.52681
z.diff.lag12  0.3606204  0.0503162   7.167 4.51e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.001678 on 353 degrees of freedom
Multiple R-squared:  0.1932,    Adjusted R-squared:  0.1635
F-statistic: 6.504 on 13 and 353 DF,  p-value: 4.119e-11


Value of test-statistic is: -2.933 4.3043

Critical values for test statistics:
      1pct  5pct 10pct
tau2 -3.44 -2.87 -2.57
phi1  6.47  4.61  3.79
```

Evaluating the output, the test statistic here is -2.933, with the tau2 5% Significance Level Critical Value being -2.87. Because the test statistic is more extreme than the Critical Value, we can reject the Null Hypothesis and conclude that it does not have a second unit root I(2) and only has a first unit root of I(1). In other words, after the first difference, logUS is stationary.

A plot examining ln(US) and ln(BHPI) is provided. The red curve represents the ln(BHPI) and the blue curve represents the ln(US).



The plot of ln(US) and ln(BHPI) seems to show that ln(US) and ln(BHPI) appear to move together and on top of each other. This is not surprising because they are both indexes and are set to the same level.

From here, we must run VARselect from the dataframe containing the logBHPI and logUS and through the AIC criterion, select 12 as the appropriate lag. Output is provided to better understand how 12 was selected as the appropriate lag.

```
> data_frame <- data.frame(logBHPI, logUS)
> VARselect(data_frame, 12)
$`selection`
AIC(n)  HQ(n)  SC(n) FPE(n) 
    12      8      5     12 

$criteria
                   1             2             3             4             5             6
AIC(n) -2.144061e+01 -2.368764e+01 -2.369236e+01 -2.371249e+01 -2.383927e+01 -2.384058e+01
HQ(n)  -2.141535e+01 -2.364554e+01 -2.363342e+01 -2.363670e+01 -2.374664e+01 -2.373112e+01
SC(n)  -2.137702e+01 -2.358166e+01 -2.354399e+01 -2.352172e+01 -2.360610e+01 -2.356502e+01
FPE(n)  4.880484e-10  5.159292e-11  5.135008e-11  5.032758e-11  4.433571e-11  4.427848e-11
                   7             8             9            10            11            12
AIC(n) -2.384726e+01 -2.391844e+01 -2.392158e+01 -2.393953e+01 -2.394851e+01 -2.395246e+01
HQ(n)  -2.372096e+01 -2.377530e+01 -2.376159e+01 -2.376270e+01 -2.375484e+01 -2.374196e+01
SC(n)  -2.352931e+01 -2.355810e+01 -2.351884e+01 -2.349440e+01 -2.346099e+01 -2.342255e+01
FPE(n)  4.398492e-11  4.096456e-11  4.083862e-11  4.011456e-11  3.975887e-11  3.960566e-11
```

From this step, the Johansen test must be performed on the variables ln(BHPI) and ln(US). The R code to run the test and its output is provided.

```
> jt <- ca.jo(data_frame, type = "trace", ecdet = "none", K = 12)
> summary(jt)

######################
# Johansen-Procedure #
######################

Test type: trace statistic , with linear trend

Eigenvalues (lambda):
[1] 0.049954369 0.001119209

Values of teststatistic and critical values of test:

          test 10pct  5pct  1pct
r <= 1 |  0.41  6.50  8.18 11.65
r = 0  | 19.32 15.66 17.95 23.52

Eigenvectors, normalised to first column:
(These are the cointegration relations)

            logBHPI.l12  logUS.l12
logBHPI.l12    1.000000  1.0000000
logUS.l12     -1.201198 -0.1965106

Weights W:
(This is the loading matrix)

           logBHPI.l12     logUS.l12
logBHPI.d -0.009260880 -0.0002121079
logUS.d    0.003022974 -0.0001469450
```

Looking specifically at r = 0, the test statistic is 19.32, and the 5% Critical Value is 17.95. Because our test statistic exceeds this Critical Value, we can reject the Null Hypothesis and conclude that ln(BHPI) and ln(US) are cointegrated.

Finally, we must Run Vector Error Correction Model (VE) through R, which its code and output are provided.

```
> ve <- cajorls(jt, 1)
> summary(ve$rlm)
Response logBHPI.d :

Call:
lm(formula = logBHPI.d ~ ect1 + constant + logBHPI.dl1 + logUS.dl1 +
    logBHPI.dl2 + logUS.dl2 + logBHPI.dl3 + logUS.dl3 + logBHPI.dl4 +
    logUS.dl4 + logBHPI.dl5 + logUS.dl5 + logBHPI.dl6 + logUS.dl6 +
    logBHPI.dl7 + logUS.dl7 + logBHPI.dl8 + logUS.dl8 + logBHPI.dl9 +
    logUS.dl9 + logBHPI.dl10 + logUS.dl10 + logBHPI.dl11 + logUS.dl11 -
    1, data = data.mat)

Residuals:
       Min         1Q     Median         3Q
-0.0086122 -0.0020358 -0.0000048  0.0020877
       Max
 0.0135653

Coefficients:
              Estimate Std. Error t value
ect1         -0.009261   0.002865  -3.233
constant     -0.008761   0.002842  -3.083
logBHPI.dl1   0.400099   0.054532   7.337
logUS.dl1     0.470088   0.106623   4.409
logBHPI.dl2   0.042412   0.058620   0.724
logUS.dl2     0.383100   0.136204   2.813
logBHPI.dl3  -0.366334   0.058199  -6.295
logUS.dl3    -0.183498   0.138253  -1.327
logBHPI.dl4   0.338190   0.061145   5.531
logUS.dl4     0.011775   0.138562   0.085
logBHPI.dl5   0.012584   0.061814   0.204
logUS.dl5     0.103928   0.138628   0.750
logBHPI.dl6  -0.073974   0.061288  -1.207
logUS.dl6    -0.295320   0.138637  -2.130
logBHPI.dl7   0.287682   0.061418   4.684
logUS.dl7     0.065491   0.139645   0.469
logBHPI.dl8  -0.014663   0.060734  -0.241
logUS.dl8    -0.058184   0.139000  -0.419
logBHPI.dl9   0.137509   0.057504   2.391
logUS.dl9    -0.135073   0.138855  -0.973
logBHPI.dl10  0.109469   0.057622   1.900
logUS.dl10   -0.005256   0.138634  -0.038
```

```
logBHPI.dl11 -0.061504   0.052461  -1.172
logUS.dl11   -0.219398   0.112203  -1.955
             Pr(>|t|)
ect1          0.00134 **
constant      0.00221 **
logBHPI.dl1  1.58e-12 ***
logUS.dl1    1.39e-05 ***
logBHPI.dl2   0.46985
logUS.dl2     0.00519 **
logBHPI.dl3  9.36e-10 ***
logUS.dl3     0.18530
logBHPI.dl4  6.30e-08 ***
logUS.dl4     0.93233
logBHPI.dl5   0.83881
logUS.dl5     0.45395
logBHPI.dl6   0.22827
logUS.dl6     0.03387 *
logBHPI.dl7  4.05e-06 ***
logUS.dl7     0.63938
logBHPI.dl8   0.80936
logUS.dl8     0.67577
logBHPI.dl9   0.01732 *
logUS.dl9     0.33135
logBHPI.dl10  0.05829 .
logUS.dl10    0.96978
logBHPI.dl11  0.24185
logUS.dl11    0.05135 .
---
Signif. codes:
  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
  ‘ ’ 1

Residual standard error: 0.003417 on 345 degrees of freedom
Multiple R-squared:  0.7775,	Adjusted R-squared:  0.762
F-statistic: 50.22 on 24 and 345 DF,  p-value: < 2.2e-16
```

```
Response logUS.d :

Call:
lm(formula = logUS.d ~ ect1 + constant + logBHPI.dl1 + logUS.dl1 +
    logBHPI.dl2 + logUS.dl2 + logBHPI.dl3 + logUS.dl3 + logBHPI.dl4 +
    logUS.dl4 + logBHPI.dl5 + logUS.dl5 + logBHPI.dl6 + logUS.dl6 +
    logBHPI.dl7 + logUS.dl7 + logBHPI.dl8 + logUS.dl8 + logBHPI.dl9 +
    logUS.dl9 + logBHPI.dl10 + logUS.dl10 + logBHPI.dl11 + logUS.dl11 -
    1, data = data.mat)

Residuals:
       Min         1Q     Median         3Q
-0.0088852 -0.0008579  0.0000396  0.0007116
       Max
 0.0101047

Coefficients:
              Estimate Std. Error t value
ect1          0.003023   0.001474   2.051
constant      0.003126   0.001462   2.138
logBHPI.dl1  -0.005534   0.028062  -0.197
logUS.dl1     0.815283   0.054868  14.859
logBHPI.dl2  -0.013654   0.030165  -0.453
logUS.dl2     0.114249   0.070090   1.630
logBHPI.dl3  -0.030098   0.029949  -1.005
logUS.dl3    -0.099436   0.071144  -1.398
logBHPI.dl4   0.015848   0.031465   0.504
logUS.dl4     0.085572   0.071303   1.200
logBHPI.dl5   0.023582   0.031809   0.741
logUS.dl5    -0.057395   0.071337  -0.805
logBHPI.dl6  -0.005029   0.031539  -0.159
logUS.dl6     0.032963   0.071342   0.462
logBHPI.dl7   0.031358   0.031605   0.992
logUS.dl7    -0.057942   0.071861  -0.806
logBHPI.dl8  -0.015247   0.031253  -0.488
logUS.dl8     0.037906   0.071528   0.530
logBHPI.dl9   0.014245   0.029591   0.481
logUS.dl9     0.067966   0.071454   0.951
logBHPI.dl10  0.095530   0.029652   3.222
logUS.dl10   -0.074408   0.071340  -1.043
logBHPI.dl11 -0.041668   0.026996  -1.544
logUS.dl11   -0.008481   0.057739  -0.147
```

```
 -
               Pr(>|t|)
ect1             0.0410 *
constant         0.0332 *
logBHPI.dl1      0.8438
logUS.dl1        <2e-16 ***
logBHPI.dl2      0.6511
logUS.dl2        0.1040
logBHPI.dl3      0.3156
logUS.dl3        0.1631
logBHPI.dl4      0.6148
logUS.dl4        0.2309
logBHPI.dl5      0.4590
logUS.dl5        0.4216
logBHPI.dl6      0.8734
logUS.dl6        0.6443
logBHPI.dl7      0.3218
logUS.dl7        0.4206
logBHPI.dl8      0.6260
logUS.dl8        0.5965
logBHPI.dl9      0.6305
logUS.dl9        0.3422
logBHPI.dl10     0.0014 **
logUS.dl10       0.2977
logBHPI.dl11     0.1236
logUS.dl11       0.8833
---
Signif. codes:
  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
  ' ' 1

Residual standard error: 0.001758 on 345 degrees of freedom
Multiple R-squared:  0.9144,    Adjusted R-squared:  0.9084
F-statistic: 153.5 on 24 and 345 DF,  p-value: < 2.2e-16
```

Although there is a ton of output that could be analyzed, there are two particular values we are interested in. For the response logBHPI.d, the ect1 p-value is 0.00134, which is less than the specified Significance Level α and provides enough evidence to reject the Null Hypothesis and conclude that the y (the BHPI) is directly affected by the difference between the BHPI and the US HPI. For the response logUS.d, the ect1 p-value is 0.0410, which is less than the specified

Significance Level α and provides enough evidence to reject the Null Hypothesis and conclude that y (US HPI) is directly affected by the difference between the BHPI and the US HPI.

CONCLUSION

Through a time-series approach to the data, we were successfully able to analyze how both the BHPI and the US Average HPI has changed over the period of time in terms of their growth rates, first differences of the growth rates, and so forth. Multiple data visualizations showed the distribution of the variables and transformations were made to these variables in an attempt to normalize them. Numerous forms of Hypothesis Tests were conducted to explore relationships between the variables and their transformations, as well as to predict future values of these indexes and to forecast these values over a select timeframe.

Although the Analysis portion of the report does the bulk of the explanation in terms of explaining how time has affected the value of both the BHPI and US Average HPI, it should be taken into consideration that this data is currently collected on a monthly basis and is an ever-changing dataset. New BHPIs and US Average HPIs for the months of October, November, and potentially even December of this year are being produced, further adding to the practicality of analyzing this dataset through a time-series lens. It is only a matter of time to really understand, to predict accurately, and to witness where the Boston, and the United States, housing markets are headed, and time-series data of these indices serve as a tool to utilize past time-series data to predict and to prepare for the evolving housing markets of the future.