Present and Future Global CO_2 Emissions

true true true true

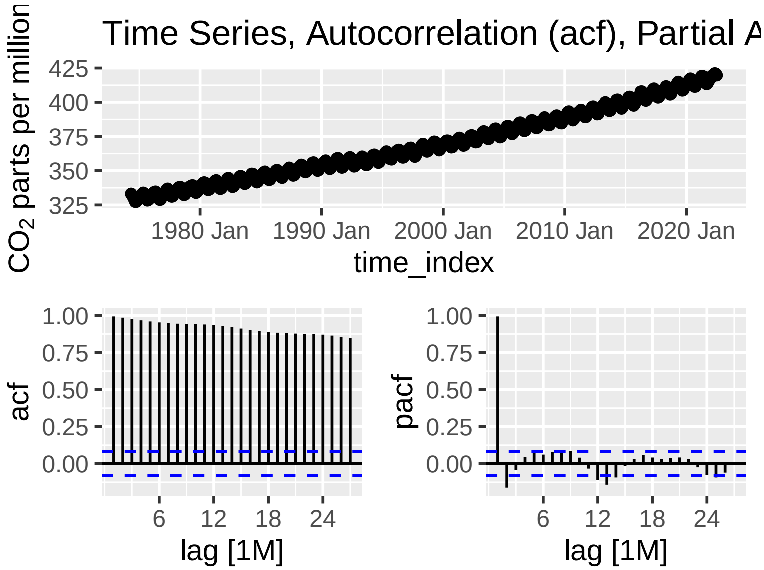
Abstract

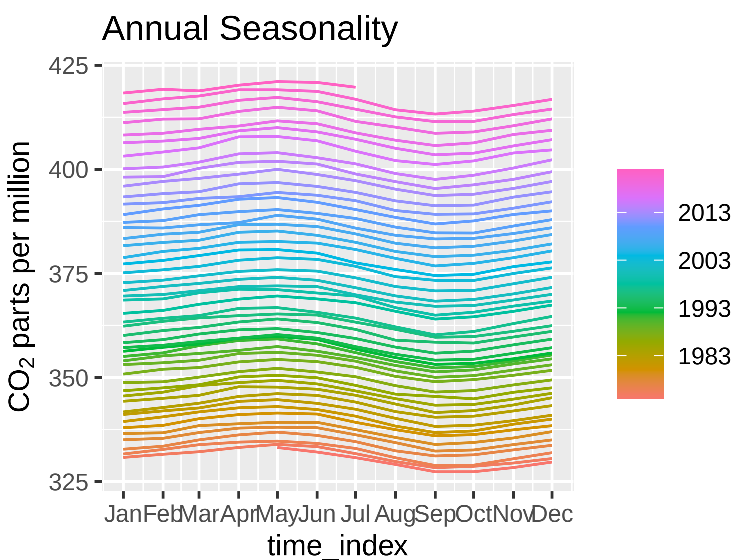
The message from the latest Intergovernmental Panel on Climate Change (IPCC) report released in April 2022 on the topic of climate change mitigation is clear: urgent and drastic action is needed if we are to limit global warming to 1.5°C. Fossil fuel based Bitcoin Mining, electricity usage for cellphone and computer, industrial pollution, water contamination, and others play a critical role in the global rise in temperature. IPCC has been examining these trends for more than sixty years. The incessant use of electric gadgets and the rising attraction towards mining digital currencues enforce us to solve the problems in newer ways like a circular economy, carbon capture and storage and greener cities. This report adds new findings and alarming predictions for the next decade.

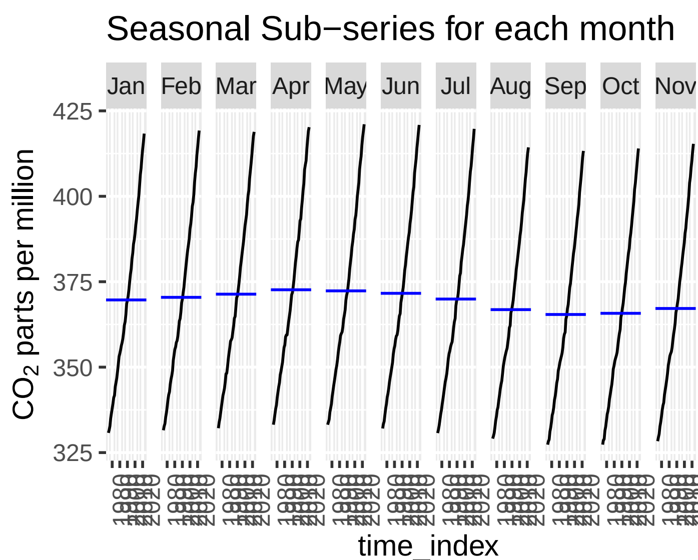
Introduction

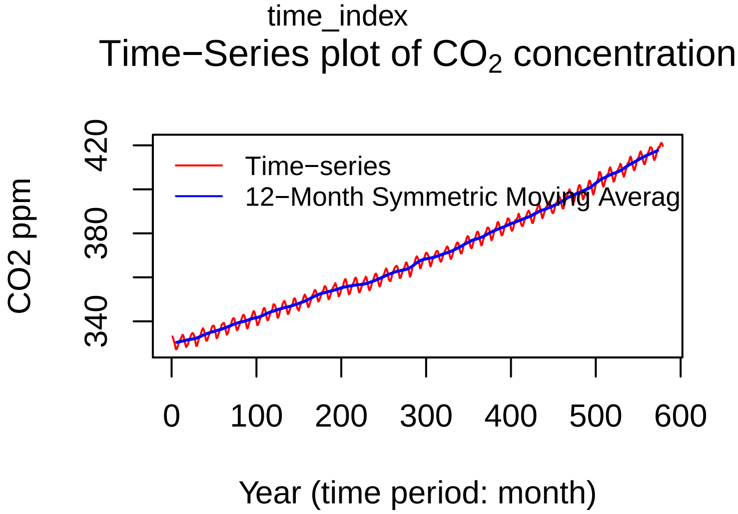
The data set is pulled from the NOAA Global Monitoring Laboratory website for the monthly average CO2 levels in Mauna Loa. Data from March 1958 through April 1974 have been obtained by C.David Keeling of the Scripps Insitiution of Oceanography. Weekly CO2 values are constructed from daily mean values. NOAA has "confidence that the CO2 measurements made at the Mauna Loa observatory reflect truth about our global atmosphere," given that the observatory is at the summit of Mauna Loa at an altitude of 3400 meters and can measure air masses that are representative of large areas, the measurements are frequently and rigorously calibrated, and ongoing comparisons are made to ensure data accuracy. The NOAA GML measures the "mole fraction" of CO2 in dry air, the number of CO2 molecules in a given number of molecules of air after the removal of water vapor. This process is done given that the dry mole fraction reflects the addition and removal of the gas given that dry air does not change when air expands upon heating or ascending to a higher altitude where the pressure is lower. The data is missing values for one week in 1954, and that value has been replaced by the average of its surrounding values on each side. Monthly averages are constructed by taking simple averages of values across each month of the time series.

From plotting the time series of the data, we see a strong upward trend and likely aspects of seasonality given the oscillating pattern of the time series throughout the trend. From the ACF of the CO2, we see high (and statistically significant) correlation for many lags (100 lags shown on the above plot). Significance is maintained through ~250 lags. This high level of autocorrelation is indicative of trended data. Observing the PACF plot, we see no statistically significant values past the first lag.

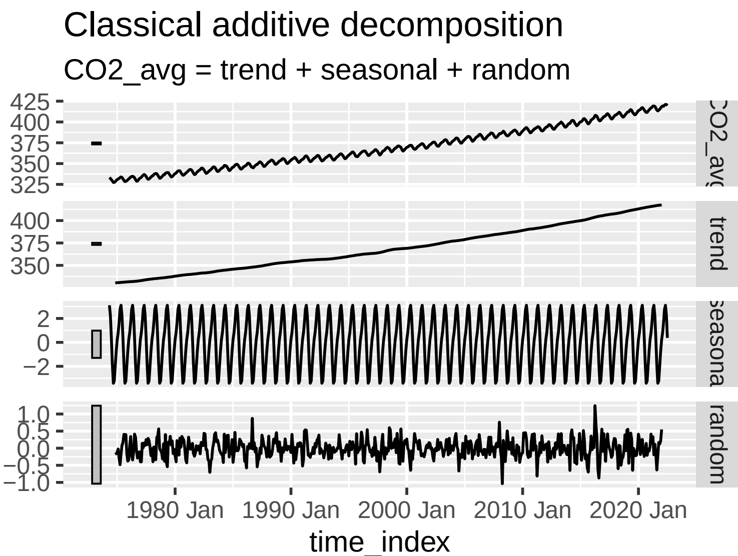




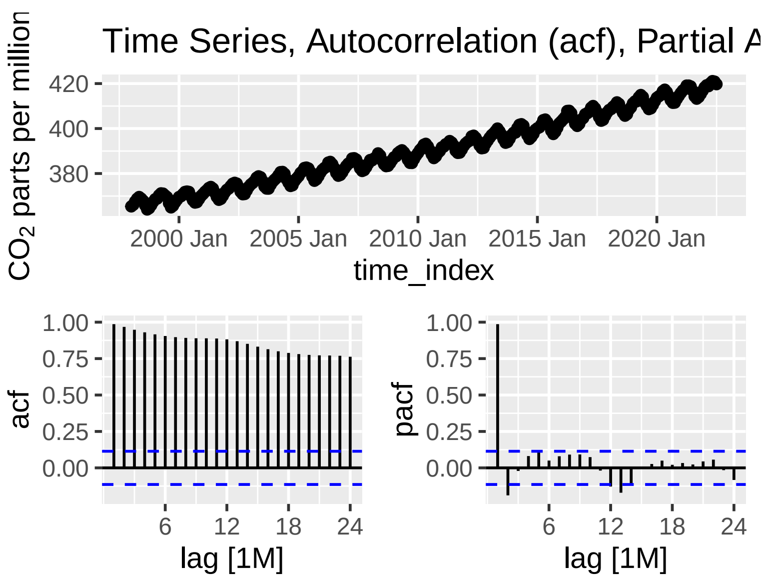




From the above plots, we observe a strong seasonal trend, persistent across years.



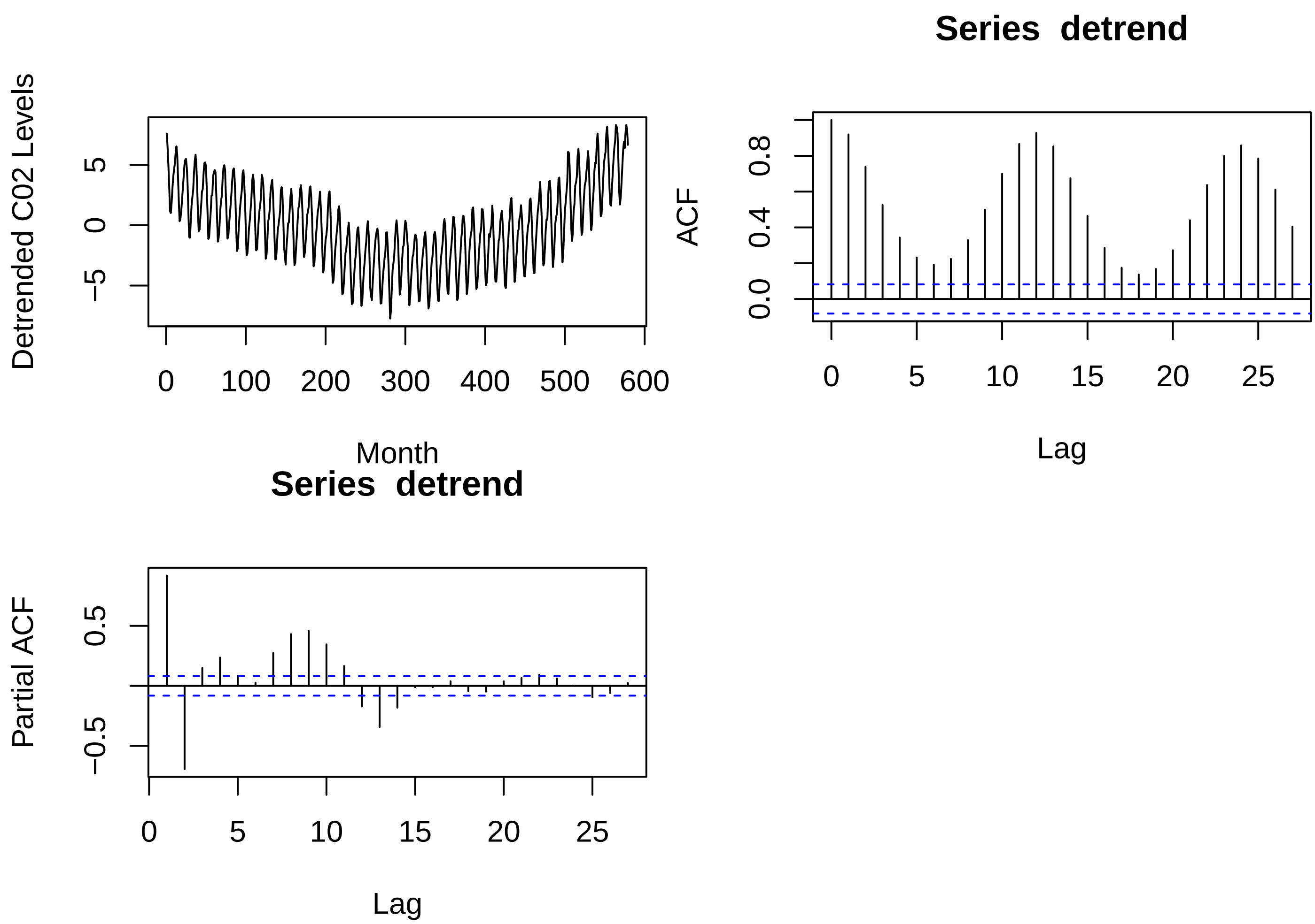
Following the observation of strong seasonality, we can decompose the time series and observe the trend of the data as well as the seasonality.



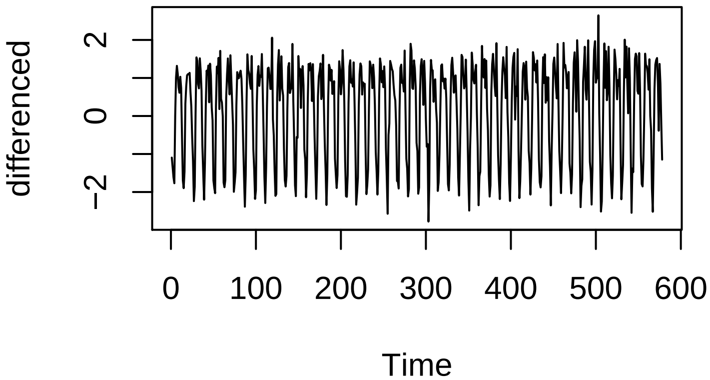
Examining just the years after 1979, we see a fluctuation of autocorrelations on the ACF plot, dipping below statistical significant at ~75 lags. The PACF plot shows statistically significant values ocurring for the first two lags and then at additional lags further out. From the linear model of the co2 levels on the time index, we see a highly statistically significant linear relationship between the two variables, as expected from the trends observed in the time series plot and the results of the ACF plot. To examine the detrending data, we will look at the residuals of the trend model.

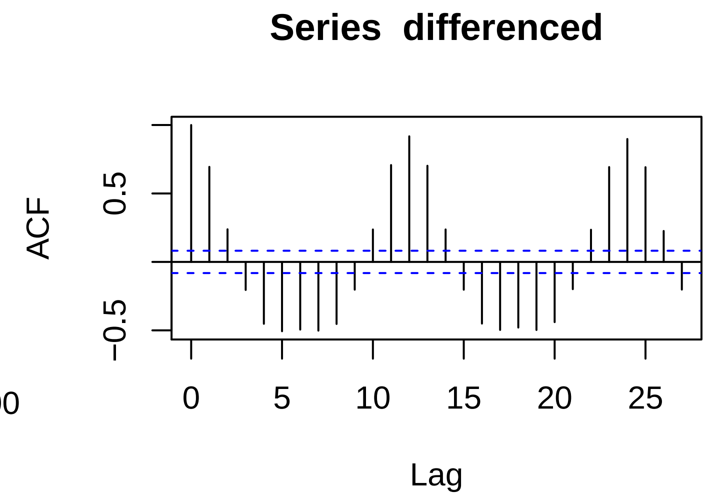
```
##
## Call:
## lm(formula = co2_present$CO2_avg ~ co2_present$time_index)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -7.74  -2.40  -0.20   2.31   8.31
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)             3.18e+02   3.16e-01    1006   <2e-16 ***
```

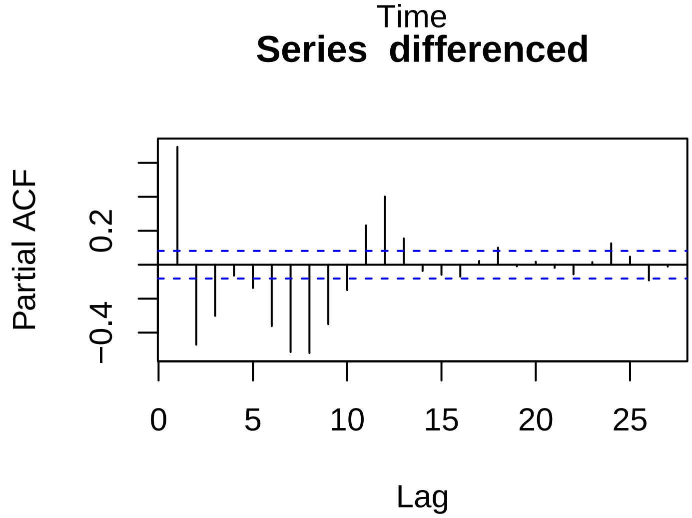
```
## co2_present$time_index 4.97e-03   2.73e-05     182   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.34 on 577 degrees of freedom
## Multiple R-squared:  0.983,  Adjusted R-squared:  0.983
## F-statistic: 3.32e+04 on 1 and 577 DF,  p-value: <2e-16
```



Detrending the data we still see the seasonal pattern and somewhat of a quadratic form. We still see highly statistically significant autocorrelation in the ACF plot, and PACF values that remain statistically significant.







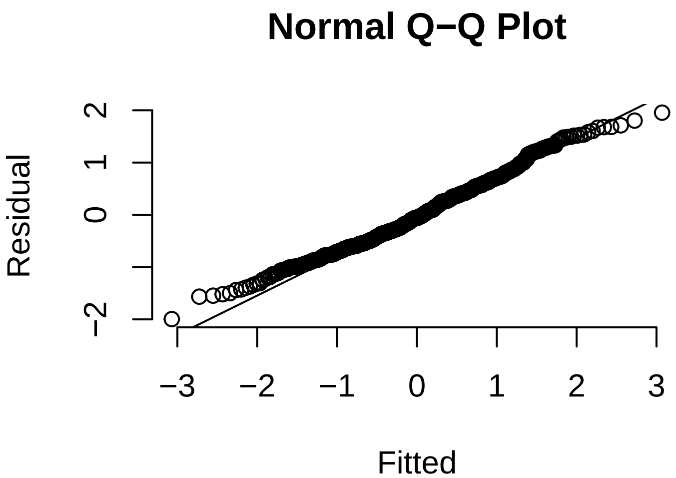

```
##
##  Box-Ljung test
##
## data:  differenced
## X-squared = 280, df = 1, p-value <2e-16
```

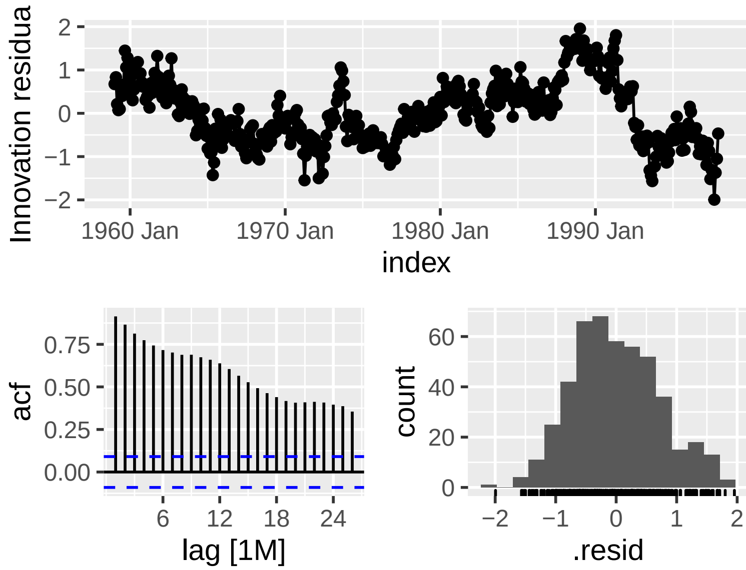
The differenced time series appears more like a white noise model. Even in the differenced data, we still see statistically significant autocorrelations, which is not indicative of a white noise model. The Ljung-Box test has a very small p value, indicating that we should reject the null hypothesis that the residuals of our time series model are independent.

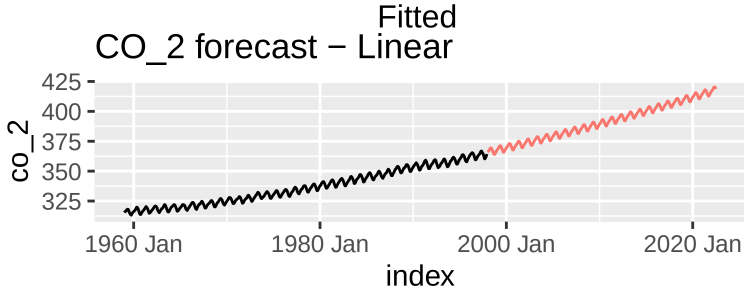
Task 2b: Compare linear model forecasts against realized CO2

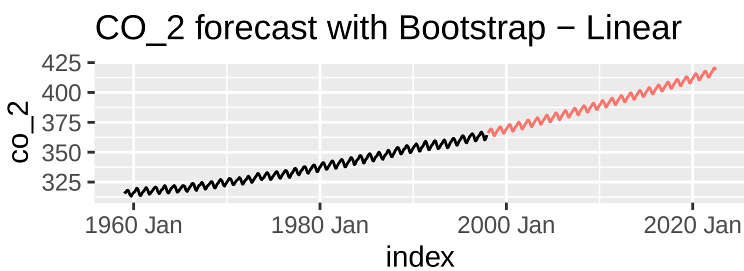
```
## [1] "Innnovation Residual Mean"
```

```
## [1] 1.21e-15
```










```
## [1] "Accuracy of trend + season"

##                 ME  RMSE   MAE     MPE  MAPE
## Test set -0.0403 0.796 0.641 -0.0151 0.163

## [1] "Accuracy of trend + season with Bootstrap"

##                 ME  RMSE   MAE     MPE  MAPE
## Test set -0.0309 0.801 0.643 -0.0127 0.163

## # A tibble: 1 x 3
##   .model                                          .type winkler
##   <chr>                                           <chr>   <dbl>
## 1 TSLM(co2_ppm ~ trend() + I(trend()^2) + season()) Test     2.68

## # A tibble: 1 x 3
##   .model                                          .type winkler
##   <chr>                                           <chr>   <dbl>
## 1 TSLM(co2_ppm ~ trend() + I(trend()^2) + season()) Test     3.01

## # A tibble: 1 x 3
```

```
##   .model                                          .type winkler
##   <chr>                                           <chr>   <dbl>
## 1 TSLM(co2_ppm ~ trend() + I(trend()^2) + season()) Test     3.28
```

```
## # A tibble: 1 x 3
##   .model                                          .type winkler
##   <chr>                                           <chr>   <dbl>
## 1 TSLM(co2_ppm ~ trend() + I(trend()^2) + season()) Test     4.07
```

```
## [1] 418 418 420 421 420 419
```

```
## [1] "RMSE of Last 36 months"
```

```
## [1] 1.18
```

```
## [1] "RMSE of First 36 months"
```

```
## [1] 0.621
```

Linear Model Explanation

To begin with let's take a look at the residual of the fitted linear model. A good forecasting method will yield innovation residuals with the following properties:

1) The innovation residuals are uncorrelated. If there are correlations between innovation residuals, then there is information left in the residuals which should be used in computing forecasts.
2) The innovation residuals have zero mean. If they have a mean other than zero, then the forecasts are biased.The innovation residuals also fails to show us the white noise. Thus we anticipate the forecast-ed values to not reflect the full absorption of trend and seasonality.
3) The innovation residuals have constant variance. This is known as "homoscedasticity".
4) The innovation residuals are normally distributed.

Looking at the ACF, there is a strong auto-correlation between the preset values of the series and the lagged values. So we conclude the residuals are correlated. The Innovation mean is close to zero. The innovation residuals have near constant variance. The innovation residuals are normally distributed in the histogram plot.

As the residual still have trend/seasonality left in it, while forecasting, the team explored both with **Residual Re-sampling** via bootstrap residuals and without boot-straping residuals to explore how well the model behaves.

As expected the forecast-ed model have a few positives and pitfalls.

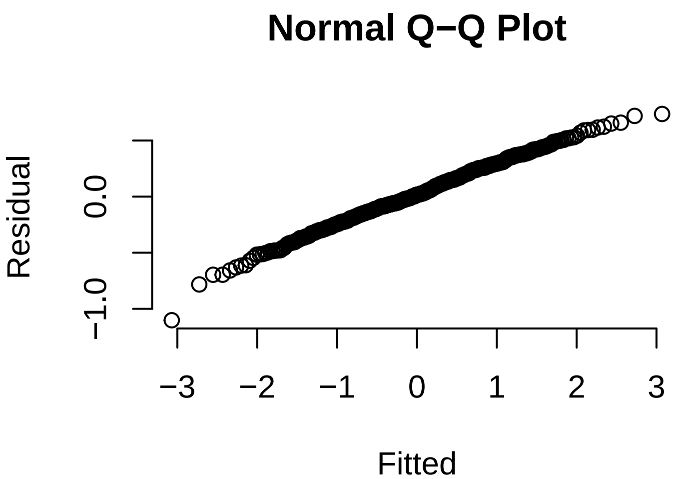
1) The Seasonality is consistently under estimated. This is evident from the height of the seasonality smaller than the training data in both boot-strapped and non-boot-strapped forecasts.
2) The RMSE and other error metrics are reasonably small.
3) To evaluate the distributional forecast accuracy our team employed **Winkler Score** test. The Winkler score can be interpreted like an absolute error. For observations that fall within the interval, the Winkler score is simply the length of the interval. So low scores are associated with narrow intervals. However, if the observation falls outside the interval, the penalty applies, with the penalty proportional to how far the observation is outside the interval. A score of 2.6 for 80% confidence interval and 3.28 for 95% confidence interval shows the forecast-ed values are not way off.
4) The trend is not captured well. This is evident from the **90.6% increase** in the RMSE values for the last 36 months compared to the first 36 months - 1.18415 vs 0.6213753. Thus, the small unobserved trend in the residuals for the initial forecast-ed values create a large effect in the far future forecasts - **the butterfly effect** is the sensitive dependence on initial conditions in which a small change in one state of a deterministic nonlinear system can result in large differences in a later state.

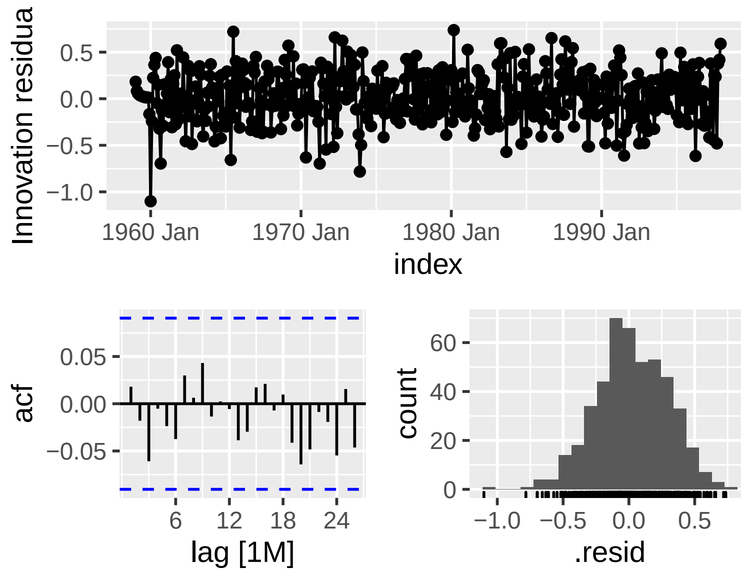
Based on the above observations, **the best linear model consistently under predicts**.

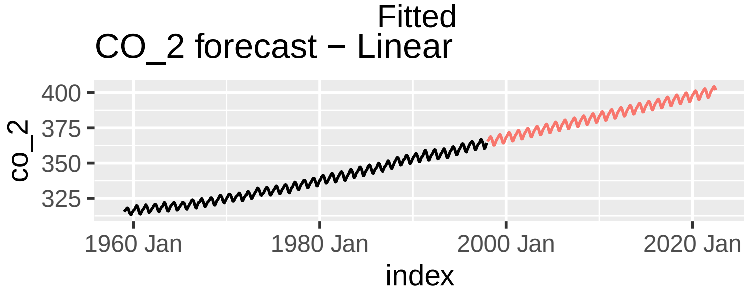
Task 3b: Compare ARIMA models forecasts against realized CO2

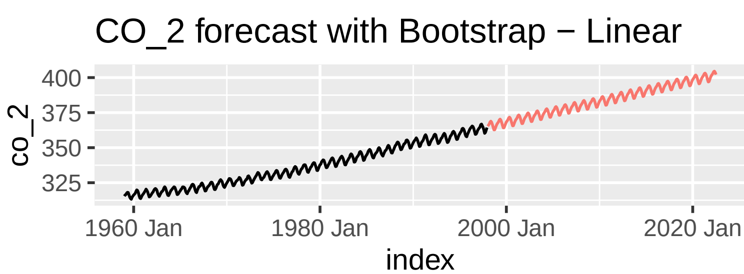
```
## [1] "Innnovation Residual Mean"

## [1] 0.0221
```










```
## [1] "Accuracy of trend + season"

##              ME RMSE  MAE  MPE MAPE
## Test set 7.29 8.89 7.29 1.82 1.82

## [1] "Accuracy of trend + season with Bootstrap"

##              ME RMSE  MAE  MPE MAPE
## Test set 7.11 8.64 7.11 1.77 1.77

## # A tibble: 1 x 3
##   .model                                                          .type winkler
##   <chr>                                                           <chr>   <dbl>
## 1 "ARIMA(co2_ppm ~ 0 + pdq(7, 1, 8) + PDQ(1, 1, 1), stepwise = FA~ Test     37.3
```

```
## # A tibble: 1 x 3
##   .model                                                       .type winkler
##   <chr>                                                        <chr>   <dbl>
## 1 "ARIMA(co2_ppm ~ 0 + pdq(7, 1, 8) + PDQ(1, 1, 1), stepwise = FA~ Test     38.0
```

```
## # A tibble: 1 x 3
##   .model                                                       .type winkler
##   <chr>                                                        <chr>   <dbl>
## 1 "ARIMA(co2_ppm ~ 0 + pdq(7, 1, 8) + PDQ(1, 1, 1), stepwise = FA~ Test     47.1
```

```
## # A tibble: 1 x 3
##   .model                                                       .type winkler
##   <chr>                                                        <chr>   <dbl>
## 1 "ARIMA(co2_ppm ~ 0 + pdq(7, 1, 8) + PDQ(1, 1, 1), stepwise = FA~ Test     68.2
```

```
## [1] "RMSE of Last 36 months"
```

```
## [1] 16.2
```

```
## [1] "RMSE of First 36 months"
```

```
## [1] 1.19
```

ARIMA Model Explanation

To begin with let's take a look at the residual of the fitted linear model. A good forecasting method will yield the above mentioned innovation residual properties.

ACF strongly reflects white noise. So we conclude the residuals are not correlated. The Innovation mean is close to zero, but substantially higher than the linear model. The innovation residuals have near constant variance. The innovation residuals are normally distributed in the histogram plot.

The forecast-ed model is not exhibiting substantial difference to the Linear model.

1) The Seasonality is estimated better than the Linear model. This is evident from the height of the seasonality relatively closer to the training data in both boot-strapped and non-boot-strapped forecasts.
2) The RMSE and other error metrics are high. As a matter of fact, all the error metrics are higher than the Linear model.
3) To evaluate the distributional forecast accuracy our team employed **Winkler Score** test. A score of 37.3 for 80% confidence interval and 47.13 for 95% confidence interval shows the forecast-ed values are way off.
4) The trend is not captured well. This is evident from the **1200%** increase in the RMSE values for the last 36 months compared to the first 36 months - 16.19685 vs 1.187938. Thus, the small unobserved trend in the residuals for the initial forecast-ed values create a large effect in the far future forecasts - **the butterfly effect**.

Based on the above observations, **the ARIMA model consistently under predicts** and no better than the linear model predictions.

???Describe how the Keeling Curve evolved from 1997 to the present.???

Task 4b: Evaluate the performance of 1997 linear and ARIMA models

```
## [1] 402 403 404 404 403 402
```

```
## [1] 418 418 420 421 420 419
```

```
## New names:
## * NA -> ...3
## * NA -> ...4
```

```
## [1] "Residual t-test Result"

##
##  Welch Two Sample t-test
##
## data:  forecast.resi$linear.residual and forecast.resi$arima.residual
## t = -24, df = 308, p-value <2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -7.92 -6.74
## sample estimates:
## mean of x mean of y
##   -0.0403    7.2916

## [1] "RMSE of Last 36 months - ARIMA"

## [1] 16.2

## [1] "RMSE of First 36 months - ARIMA"

## [1] 1.19

## [1] "RMSE of Last 36 months - Linear"

## [1] 1.18

## [1] "RMSE of First 36 months - Linear"

## [1] 0.621

## [1] "Accuracy of Linear Model"

##               ME  RMSE   MAE     MPE  MAPE
## Test set -0.0403 0.796 0.641 -0.0151 0.163

## [1] "Accuracy of ARIMA Model"

##            ME RMSE  MAE  MPE MAPE
## Test set 7.29 8.89 7.29 1.82 1.82

##
##  Box-Ljung test
##
## data:  forecast.resi$linear.residual
## X-squared = 215, df = 1, p-value <2e-16

##
##  Box-Ljung test
##
## data:  forecast.resi$arima.residual
## X-squared = 290, df = 1, p-value <2e-16
```

Explanation

???Explain how the model meets 420ppm???

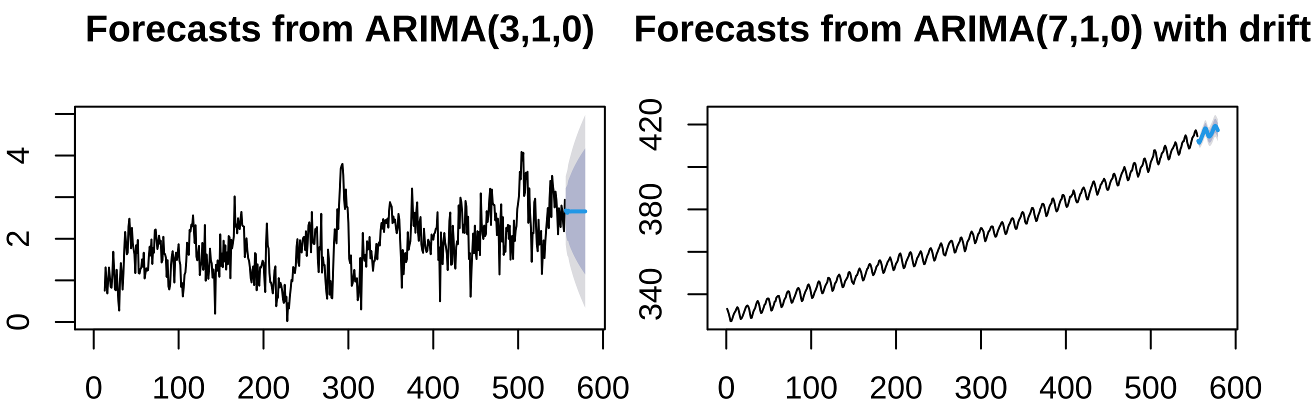
The forecasting performance of the models are assessed by a test and two metrics.

1) **t-test** and **Box test** - t-test determines whether the means of two groups are equal to each other. True difference in means is not equal to 0 for the residuals of Linear and Arima tells us that the models have different means ttest H0 => Mean value of 2 residuals are same Ha=> it is not. The test shows that the low p-value. mean value of linear residual is lower linear prediction seems to do a better job. two residuals are statistically different.

Box test is telling the residuals are not independent There is no independence in the residuals.

2) **The butterfly effect** - It is the sensitive dependence on initial conditions in which a small change in one state of a deterministic nonlinear system can result in large differences in a later state. A **90.6% increase** in the RMSE values for the last 36 months compared to the first 36 months - 1.18415 vs 0.6213753 of the linear model proves the trend is not fully captured. Thus, the small unobserved trend in the residuals for the initial forecast-ed values create a large effect in the far future forecasts. Also, a **1200%** increase in the RMSE values for the last 36 months compared to the first 36 months - 16.19685 vs 1.187938 for the ARIMA model echoes the same.

3) **Error Metrics** - ME, RMSE, MAE, MPE, MAPE: ???




```
## # A tibble: 1 x 10
##   .model                    .type    ME  RMSE   MAE   MPE  MAPE  MASE RMSSE  ACF1
##   <chr>                     <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 TSLM(CO2_avg ~ trend() ~ Test   6.34  6.36  6.34  1.52  1.52   NaN   NaN 0.347
```

```
## # A tibble: 1 x 10
##   .model                   .type      ME  RMSE   MAE   MPE  MAPE  MASE RMSSE  ACF1
##   <chr>                    <chr>   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 TSLM(seasonally_adjust~ Test  -0.214 0.501 0.346 -15.8  20.6   NaN   NaN 0.434
```

Question 2.5

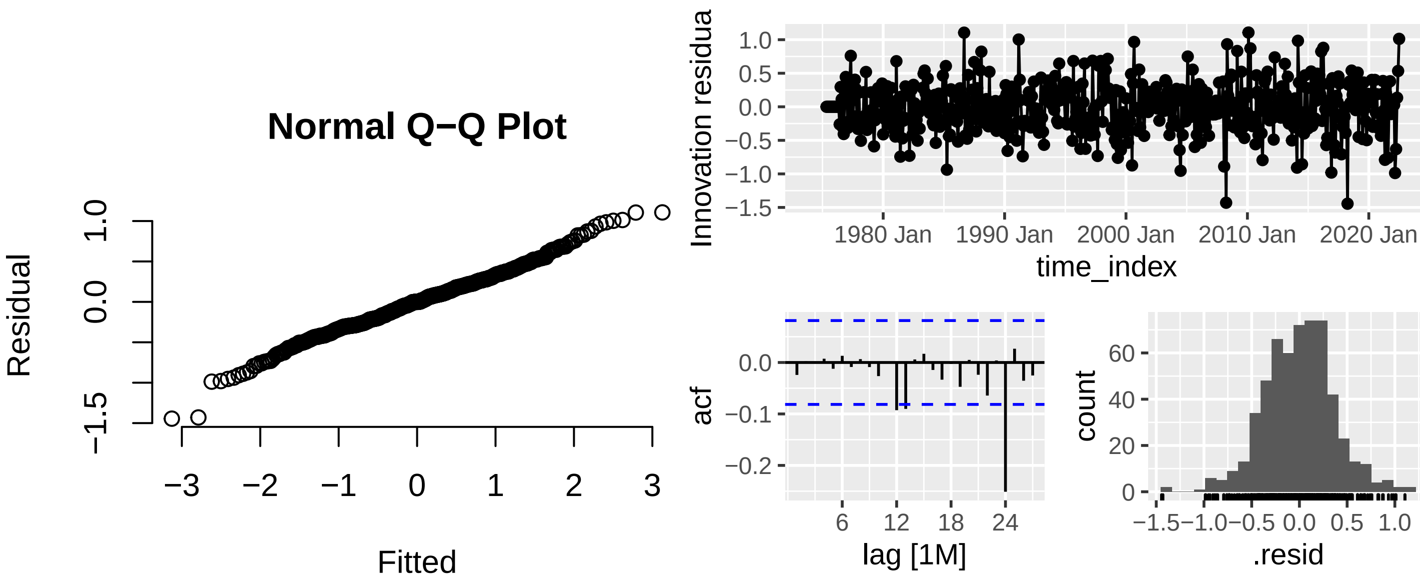
First, we split the seasonally adjusted and non-seasonally adjusted time series into training and test sets. following, we will find the top ARIMA models by evaluating on AIC, fitting the models both for in sample testing (inclusive of the entire time series) and out of sample testing (fitting only on the training set).

```
##    p_seq d_seq q_seq aic_seq
## 38    12     1     0     557
## 10     3     0     0     609
## 7      2     0     0     611
## 13     4     0     0     612
## 11     3     1     0     616
## 8      2     1     0     617
```

```
##    p_seq d_seq q_seq aic_seq
## 38    12     1     0     590
## 10     3     0     0     640
## 13     4     0     0     642
## 7      2     0     0     642
## 11     3     1     0     647
## 8      2     1     0     649
```

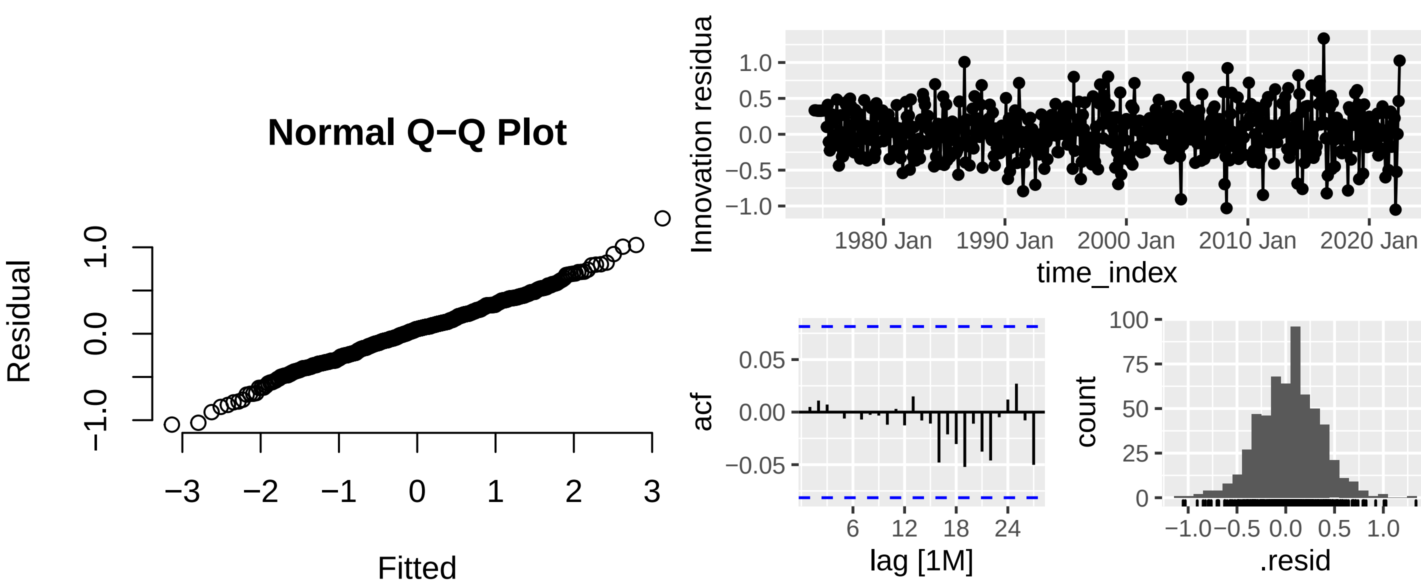
Following the model fitting, we will evaluate the performance of the in-sample models.

```
## [1] NA
```



Using the top model from our AIC criterion, for seasonally adjusted in sample data, we observe a fit of 0.133. We also observe that the residuals appear akin to a whitenoise pattern in their ts plot, as well as roughly normally distributed in their histogram. However, we observe a statistically significant value at lag 12 and 24 in the ACF plot.

```
## [1] 0.0377
```



Using the top model from our AIC criterion, for seasonally adjusted in sample data, we observe a fit of 0.104. We also observe that the residuals appear akin to a whitenoise pattern in their ts plot, as well as roughly normally distributed in their histogram. The ACF plot shows statistically insignificant values for up through 24 lags.

Next we will be looking at out of sample performance of our models.

```
##    p_seq d_seq q_seq aic_seq   mse
## 38    12     1     0     557 0.212
## 10     3     0     0     609 0.657
## 7      2     0     0     611 0.823
```

```
## 13      4      0      0      612 0.657
## 11      3      1      0      616 0.619
## 8       2      1      0      617 0.668
```

```
##      p_seq d_seq q_seq aic_seq   mse
## 240     9     0     5     422 0.204
## 241     9     0     6     434 0.198
## 296    11     0     9     440 0.200
## 319    12     0     6     452 0.188
## 308    11     1     8     453 0.201
## 295    11     0     8     455 0.201
```

Evaluating the out of sample performance of the top 6 ARIMA models fit to the seasonally adjusted data, we see that the ARIMA model (12, 1, 0) produces the best out of sample performance, with the lowest mean squared error. This model also produces the lowest AIC.

Evaluating the out of sample performance of the top 6 ARIMA models fit to the seasonally adjusted data, we see that the ARIMA model (11, 0, 9) produces the best out of sample performance, with the lowest mean squared error. This model produced the third highest AIC, but the difference in AIC score from the top ranked model and the third is relatively small.

Finally we will fit several polynomial models and evaluate performance. As a first step, we will fit a linear polynomial trend model

```
## Series: seasonally_adjusted_co2
## Model: TSLM
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.681 -0.396 -0.024  0.418  1.952
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.183542   0.053780    22.0   <2e-16 ***
## trend()     0.002269   0.000166    13.7   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.606 on 541 degrees of freedom
## Multiple R-squared: 0.257,   Adjusted R-squared: 0.256
## F-statistic:  187 on 1 and 541 DF, p-value: <2e-16
```

We see statistically significant coefficients at a very low p-value and an r-squared of 0.257, which means that only about a quarter of the variance in the data is explained by the model. The MSE is 0.249 which is higher than the lowest MSE ARIMA out of sample performance of 1.88.

```
## Series: seasonally_adjusted_co2
## Model: TSLM
## Transformation: log(seasonally_adjusted_co2)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -4.35  -0.17   0.06   0.28   0.81
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.13895    0.03891    3.57  0.00039 ***
```

```
## trend()      0.00131    0.00012   10.92  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.438 on 541 degrees of freedom
## Multiple R-squared: 0.181,   Adjusted R-squared: 0.179
## F-statistic:  119 on 1 and 541 DF, p-value: <2e-16
```

We see statistically significant coefficients at a very low p-value and an r-squared of 0.181, which means that only about a fifth of the variance in the data is explained by the model, worse than the linear model. The MSE is 3.522 which is higher than the lowest MSE ARIMA out of sample performance of 1.88.

```
## Series: seasonally_adjusted_co2
## Model: TSLM
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.588 -0.371 -0.032  0.394  2.068
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.45e+00   8.42e-02   17.23  < 2e-16 ***
## trend()      -4.33e-04   6.82e-04   -0.63     0.53
## I(trend()^2)  4.76e-06   1.17e-06    4.08  5.2e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.597 on 540 degrees of freedom
## Multiple R-squared: 0.279,   Adjusted R-squared: 0.277
## F-statistic:  105 on 2 and 540 DF, p-value: <2e-16
```

We see statistically significant coefficients for the quadratic term at a very low p-value and an r-squared of 0.181, which means that only about a fifth of the variance in the data is explained by the model, worse than the linear model. The MSE is 0.441 which is lower than the lowest MSE ARIMA out of sample performance of 1.88, indicating that this model may be more accurate at predicting future values than the ARIMA model.

How bad could it get?