Global CO_2 Emissions 1995 and Pressent

true

Abstract

Global average temperatures have increased by more than 1°C since pre-industrial time and the increase of carbon emission has played a key role in the global warming. Human emissions of carbon dioxide and other greenhouse gases are a primary driver of climate change.The rapid warming itself can have significant impacts on climate and natural systems across the world and human's greed for industrialization may have caused it.

Understanding a changing climate, and what it means for the earth's inhabitants is of growing interest to the scientific and policy community. Although, at this point in 1997 it is not entirely clear what the consequences of this growing awareness will be, in this report we present likely outcomes under "business-as-usual" scenarios. In doing so, our hope, is to establish a series of possible futures, and, as evidence, technology, and policy develop over the coming decades, that we can weigh the impacts that carbon-emission reduction efforts might take.

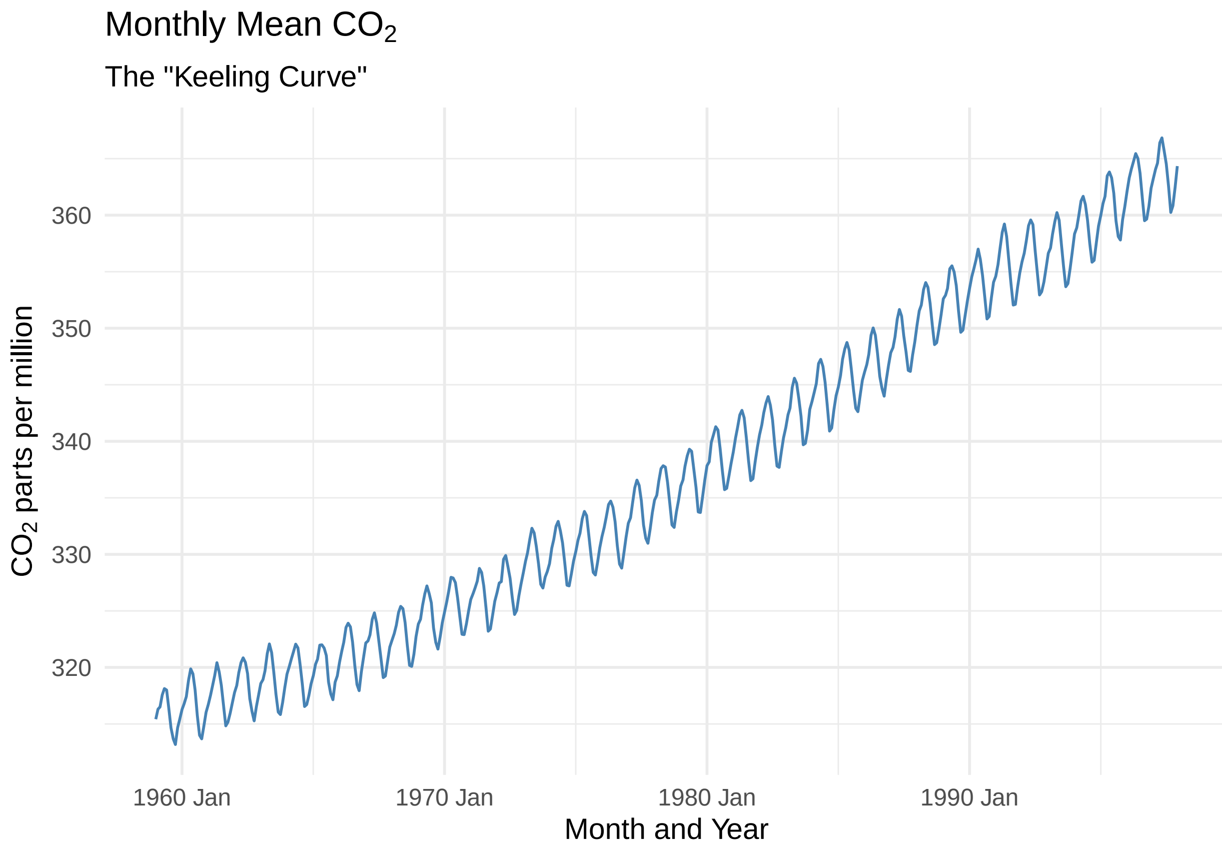
Background

Carbon Emissions

We want to understand if atmospheric CO2 levels increased since 1997 and if we anticipate that this trend will continue. High CO2 levels have been shown to negatively play into climate change. The increase in atmospheric CO2 has cascading effects from global warming to ocean acidification. Due to this, we may notice food supply chain disruptions, natural disasters of increasing intensity, and habitat disruption of all ecosystems and the animal species living there. In an economical sense, understanding these patterns is important so we can prepare for any shifts and uncertainities in supply chains that might arise for these new climate patterns. In total understanding the patterns in atmospheric CO2 concentrations is important both so that we can mitigate the existing effects of this increase and prevent continuing upwards trends in CO2. In addition if our findings are significant, this could persuade other governmental organizations to increase their CO2 monitoring efforts.

Historical Trends in Atmospheric Carbon

In 1958 Keeling began continuous monitoring of atmospheric carbon dioxide concentrations from the Mauna Loa Observatory in Hawaii and soon observed a trend increase carbon dioxide levels in addition to the seasonal cycle. He was able to attribute this trend increase to growth in global rates of fossil fuel combustion. This trend has continued to the present, and is known as the "Keeling Curve."

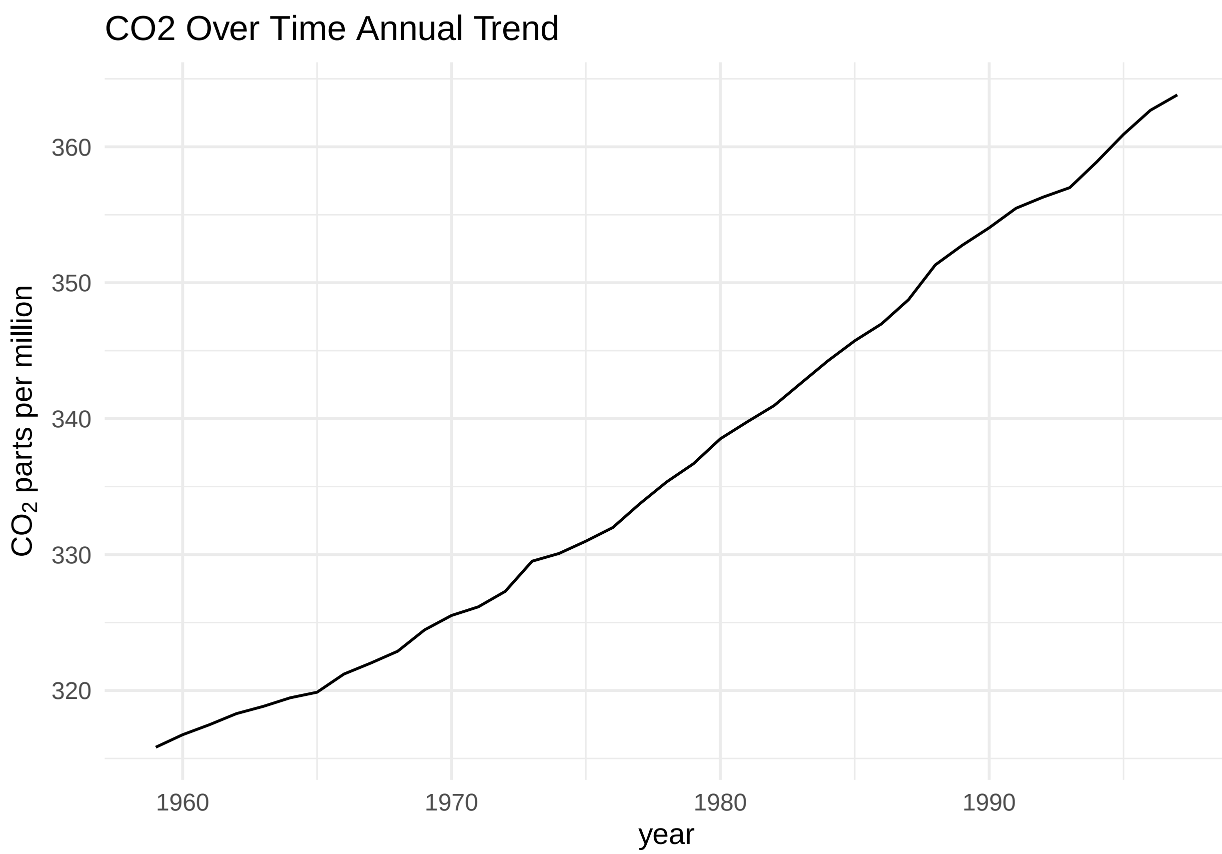


Models and Forecasts

In this section, we present evaluate two classes models the linear model and the ARIMA model to assess which time series model is most appropriate to use.

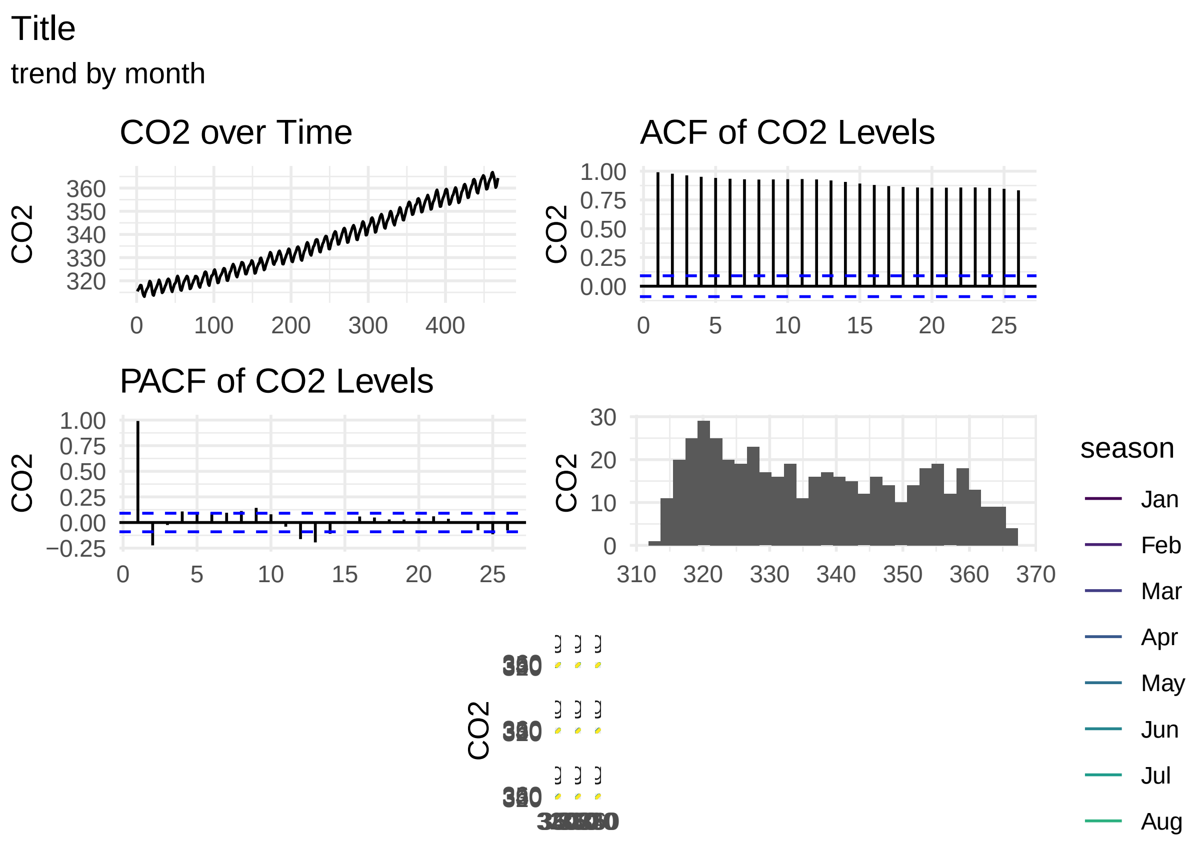
Data EDA

This data measures the mean atmospheric CO2 concentration in Mauna Loa Observatory in Hawaii. The data ranges from 313 parts per million by volume (ppmv) in March 1958 to 406ppmv in November 2018. The data was normalized to remove any influence from local contamination. Carbon dioxide measurements at the Mauna Loa Observatory in Hawaii are made with a type of infrared spectrophotometer, now known as a nondispersive infrared sensor, that is calibrated using World Meteorological Organization standards.




```
## Warning: The dot-dot notation (`..count..`) was deprecated in ggplot2 3.4.0.
## i Please use `after_stat(count)` instead.

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

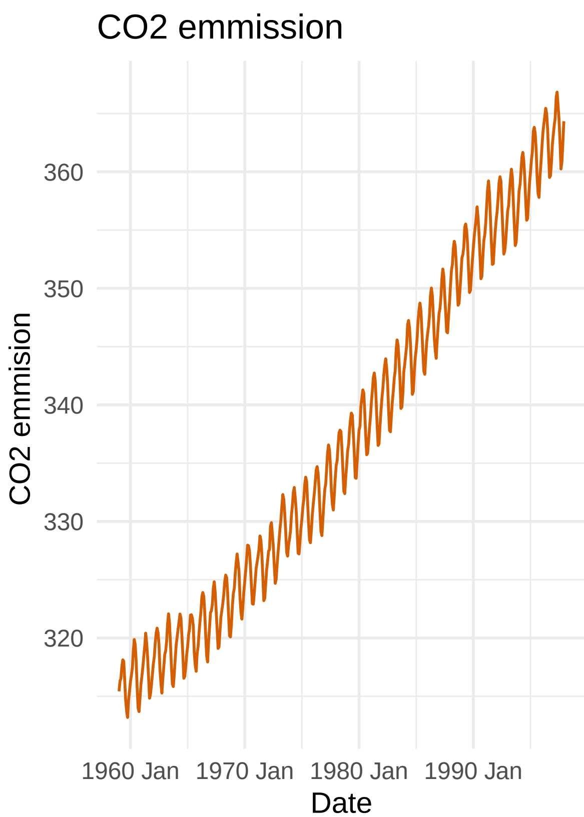


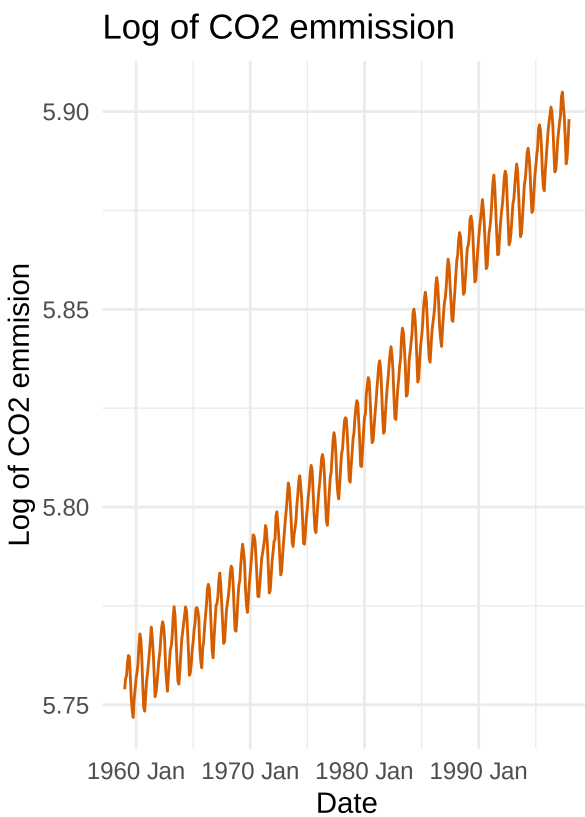
Observations: - Time Series: From the CO2 Over Time plot, there is an obvious trend of increasing levels throughout time, and we can some regular oscillations which could mean there is seasonality in the data. - ACF: There's a gradual decay in the ACF values over the lags, perhaps indicating there is a trend in the data. - PACF: PACF cuts off to zero after lag 2 and stays that way. However, at lags 12 and 13, the values become significant again. This suggests that although the model exhibits AR model-like behavior, there's some deviances in the data. - Distribution:

> Observations: - Time Series: From the yearly CO2 Over Time plot, there is an obvious trend of increasing levels throughout time, but we can see that there is no seasonality now. This points to the period of the seasonal trend being a year. - ACF: There's a gradual decay in the ACF values over the lags, indicating there is a still a trend in the data.

Linear Models

Time series decompotision





To fit linear time trend model to the `co2` series we will compare a regular time trend linear model to a quadratic time trend model. We will also fit a polynomial time trend model that incorporates seasonal dummy variables, and use this model to generate forecasts to the year 2020.

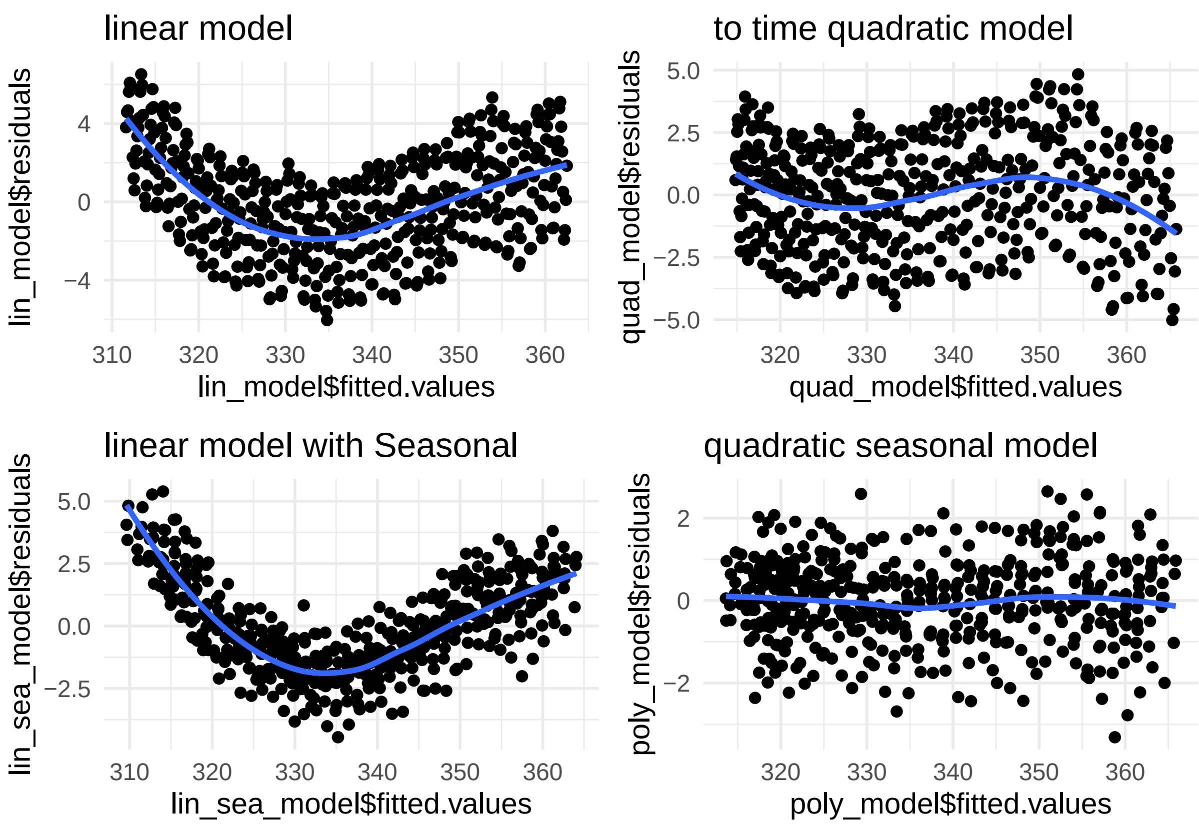
To begin, we fit a model of the form:

$$\text{CO}_2 = \phi_0 + \phi_1 + \epsilon_{eit} \tag{1}$$

```
##
## ===========================================================================================
##                                                           Dependent variable:
##                         ---------------------------------------------------------------------
##                                                                  co2
##                                   (1)                        (2)                        (3)
## -------------------------------------------------------------------------------------------
## time(co2)                 0.109*** (0.001)           0.067*** (0.003)           0.109*** (0.001)
## I(time(co2)2)                                      0.0001*** (0.00001)
## I(time(co2)3)
## spring                                                                          5.110*** (0.244)
## winter                                                                          2.690*** (0.244)
## summer                                                                          3.440*** (0.244)
## fall
```

```
## Constant                 312.000*** (0.242)          315.000*** (0.304)          309.000*** (0.230)
## -----------------------------------------------------------------------------------------------------
## Observations                     468                         468                        468
## R2                              0.969                       0.979                      0.985
## Adjusted R2                     0.969                       0.979                      0.985
## Residual Std. Error       2.620 (df = 466)            2.180 (df = 465)           1.860 (df = 463)
## F Statistic          14,795.000*** (df = 1; 466) 10,750.000*** (df = 2; 465) 7,419.000*** (df = 4; 46
## =====================================================================================================
## Note:
```

```
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
## `geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```



Model Forcast

```
## Series: co2
## Model: TSLM
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -5.02  -1.71   0.21   1.80   4.83
##
```

```
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.15e+02   3.04e-01  1035.7   <2e-16 ***
## trend()      6.74e-02   2.99e-03    22.5   <2e-16 ***
## I(trend()^2) 8.86e-05   6.18e-06    14.3   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.18 on 465 degrees of freedom
## Multiple R-squared: 0.979,   Adjusted R-squared: 0.979
## F-statistic: 1.07e+04 on 2 and 465 DF, p-value: <2e-16
```

```
## # A tibble: 1 x 3
##   .model      lb_stat lb_pvalue
##   <chr>         <dbl>     <dbl>
## 1 trend_model   3838.         0
```

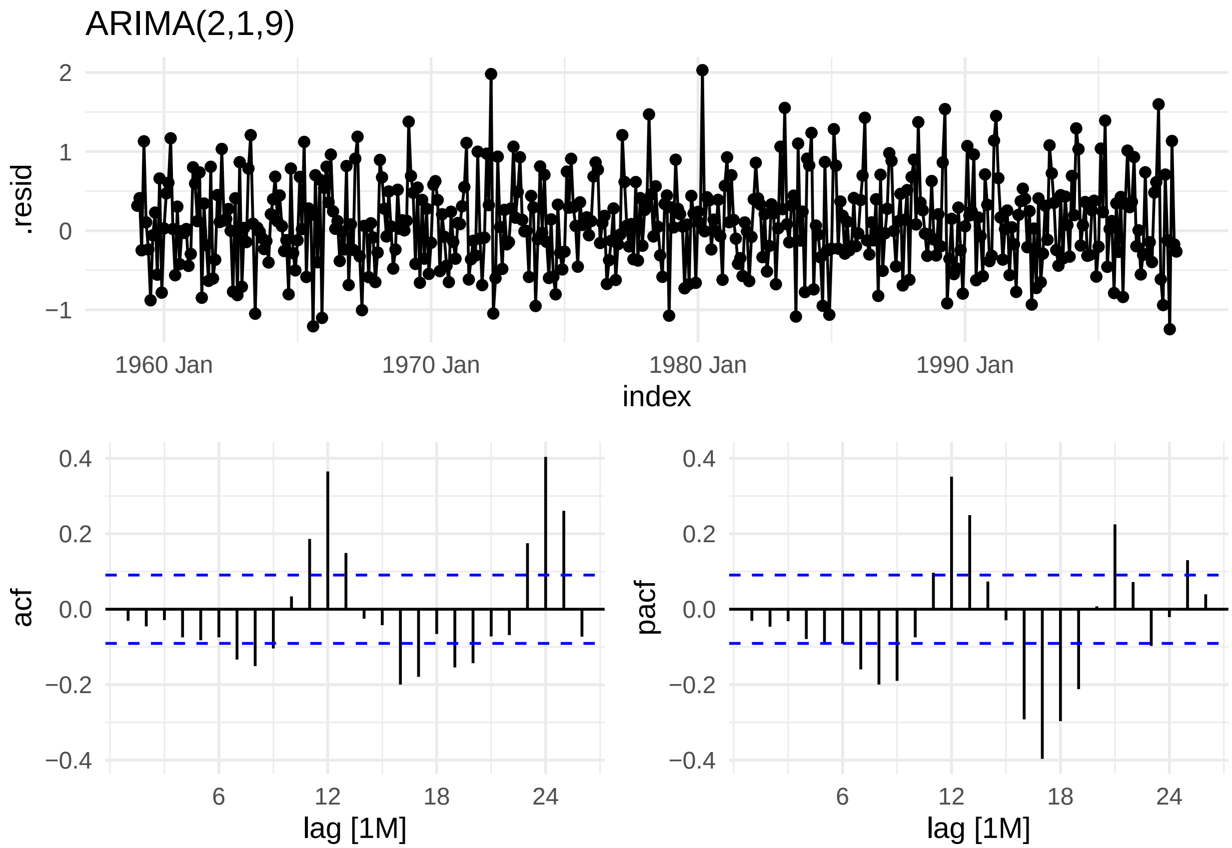
```
## # A tibble: 1 x 5
##   adj_r_squared    CV   AIC  AICc   BIC
##           <dbl> <dbl> <dbl> <dbl> <dbl>
## 1         0.979  4.79  735.  735.  752.
```

ARIMA Models

Sure we also fit some ARIMA models. And talk about them.

```
## <lst_mdl[1]>
## [1] <ARIMA(2,1,4)>
```

```
## # A tibble: 6 x 6
##   .model   term  estimate std.error statistic   p.value
##   <chr>    <chr>    <dbl>     <dbl>     <dbl>     <dbl>
## 1 ts.model ar1     1.69      0.0141   120.    0
## 2 ts.model ar2    -0.952     0.0142   -67.2   3.61e-242
## 3 ts.model ma1    -1.26      0.0448   -28.1   2.63e-102
## 4 ts.model ma2     0.111     0.0778     1.43  1.53e-  1
## 5 ts.model ma3     0.0703    0.0792     0.887 3.75e-  1
## 6 ts.model ma4     0.266     0.0464     5.72  1.90e-  8
```



Forecasts

Report from the Point of View of the Present

One of the very interesting features of Keeling and colleagues' research is that they were able to evaluate, and re-evaluate the data as new series of measurements were released.

Introduction

In this introduction, you can assume that your reader will have **just** read your 1997 report. In this introduction, **very** briefly pose the question that you are evaluating, and describe what (if anything) has changed in the data generating process between 1997 and the present.

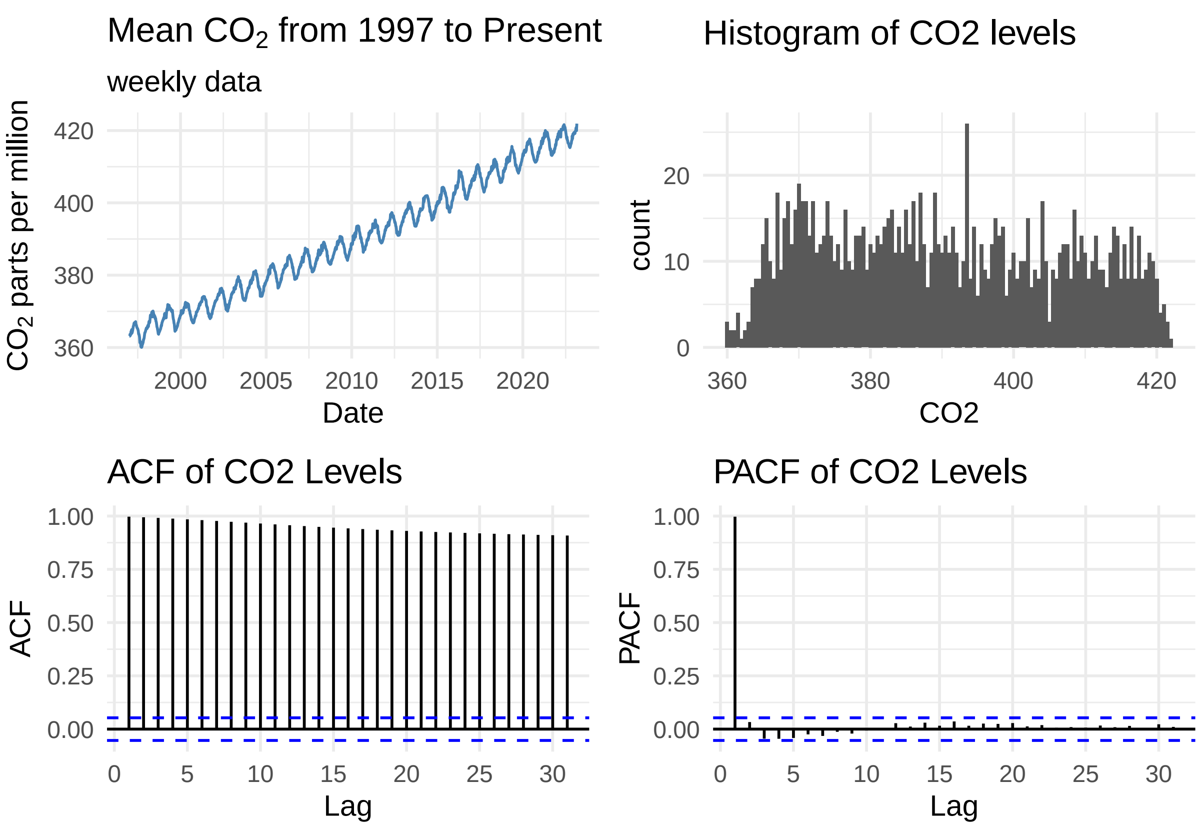
(3 points) Task 1b: Create a modern data pipeline for Mona Loa CO2 data.

The most current data is provided by the United States' National Oceanic and Atmospheric Administration, on a data page [here]. Gather the most recent weekly data from this page. (A group that is interested in even more data management might choose to work with the hourly data.)

Create a data pipeline that starts by reading from the appropriate URL, and ends by saving an object called `co2_present` that is a suitable time series object.

Conduct the same EDA on this data. Describe how the Keeling Curve evolved from 1997 to the present, noting where the series seems to be following similar trends to the series that you "evaluated in 1997" and where the series seems to be following different trends. This EDA can use the same, or very similar tools and views as you provided in your 1997 report.

```
## # A tibble: 6 x 11
##    year month   day decimal average ndays X1.year.ago X10.years.ago increase.s~1
##   <int> <int> <int>   <dbl>   <dbl> <int>       <dbl>         <dbl>        <dbl>
## 1  1997     1     5   1997.    363.     7        362.          349.         83.1
## 2  1997     1    12   1997.    363.     7        362.          349.         83.0
## 3  1997     1    19   1997.    364.     7        362.          349.         83.3
## 4  1997     1    26   1997.    363.     7        363.          349.         82.9
## 5  1997     2     2   1997.    364.     7        363.          349.         83.1
## 6  1997     2     9   1997.    364.     7        363.          348.         83.7
## # ... with 2 more variables: time_index <dttm>, month_index <mth>, and
## #   abbreviated variable name 1: increase.since.1800
```



(1 point) Task 2b: Compare linear model forecasts against realized CO2

Descriptively compare realized atmospheric CO2 levels to those predicted by your forecast from a linear time model in 1997 (i.e. "Task 2a"). (You do not need to run any formal tests for this task.)

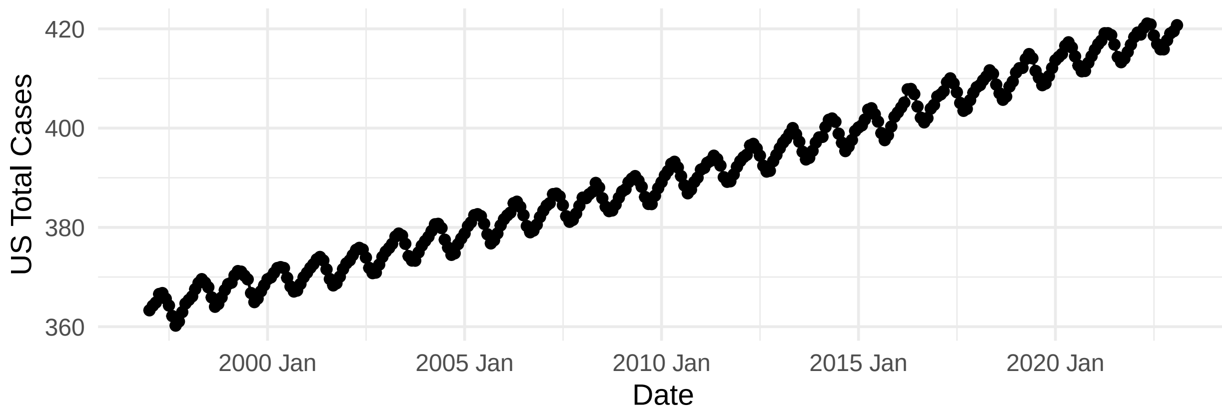
(1 point) Task 3b: Compare ARIMA models forecasts against realized CO2

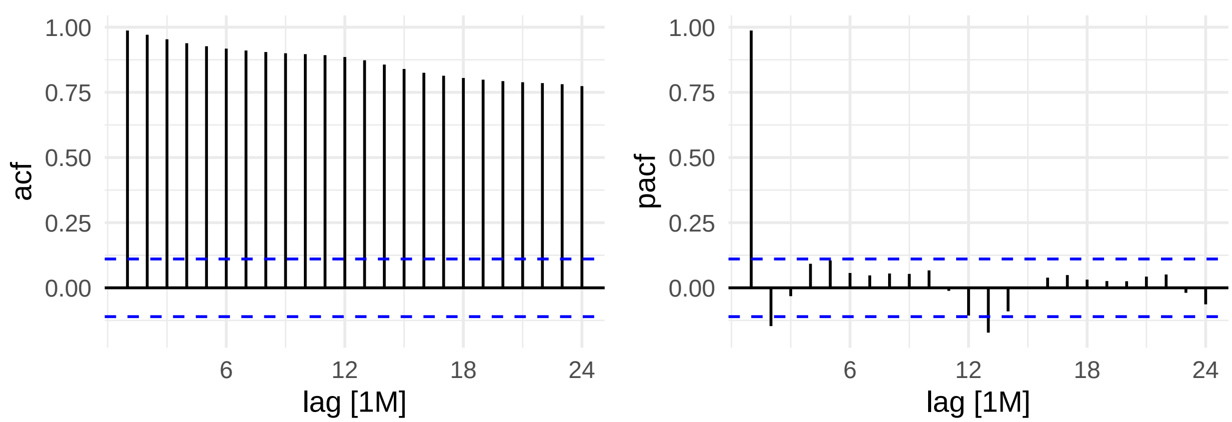
Descriptively compare realized atmospheric CO2 levels to those predicted by your forecast from the ARIMA model that you fitted in 1997 (i.e. "Task 3a"). Describe how the Keeling Curve evolved from 1997 to the present.

(3 points) Task 4b: Evaluate the performance of 1997 linear and ARIMA models

In 1997 you made predictions about the first time that CO2 would cross 420 ppm. How close were your models to the truth?

After reflecting on your performance on this threshold-prediction task, continue to use the weekly data to generate a month-average series from 1997 to the present, and compare the overall forecasting performance of your models from Parts 2a and 3b over the entire period. (You should conduct formal tests for this task.)





(4 points) Task 5b: Train best models on present data

Seasonally adjust the weekly NOAA data, and split both seasonally-adjusted (SA) and non-seasonally-adjusted (NSA) series into training and test sets, using the last two years of observations as the test sets. For both SA and NSA series, fit ARIMA models using all appropriate steps. Measure and discuss how your models perform in-sample and (psuedo-) out-of-sample, comparing candidate models and explaining your choice. In addition, fit a polynomial time-trend model to the seasonally-adjusted series and compare its performance to that of your ARIMA model.

(3 points) Task Part 6b: How bad could it get?

With the non-seasonally adjusted data series, generate predictions for when atmospheric CO2 is expected to be at 420 ppm and 500 ppm levels for the first and final times (consider prediction intervals as well as point estimates in your answer). Generate a prediction for atmospheric CO2 levels in the year 2122. How confident are you that these will be accurate predictions?

Conclusions

What to conclude is unclear.

While the most plausible model that we estimate is reported in the main, "Modeling" section, in this appendix to the article we examine alternative models. Here, our intent is to provide a skeptic that does not accept our assessment of this model as an ARIMA of order (1,2,3) an understanding of model forecasts under alternative scenarios.