Module 1: Climate change

```r
library(tidyverse)
```

Team members, etc

List your team member(s): Nikhil Shah

Include a link to your forked GH repository: https://github.com/nikhilshah7/CLES131-module1-climate#

Include a link to your .qmd file: https://github.com/nikhilshah7/CLES131-module1-climate/blob/main/module1.qmd

CO2 trends from Mauna Loa

We will begin at the global scale by plotting the famed Mauna Loa CO_2 curve at monthly intervals, with the twin objectives of (1) making layered plots with 'ggplot2', a package within the 'tidyverse' collection widely adopted by the R data science community, and (2) refreshing understanding of the seasonal cycle in atmospheric [CO_2].

With data you've collected yourself, it is more common to store locally as a csv. However, we can use tidyverse tools to read directly from files stored online and will do so here to get the latest data.

```r
co2 <-read_table("https://gml.noaa.gov/webdata/ccgg/trends/co2/co2_mm_mlo.txt",
                 comment="#",
                 col_names = c("year", "month", "decimal_date", "average",
                               "interpolated", "trend", "days"),
                 na = c("-1", "-99.99"))
```

You can examine these data in several ways. One option is to use your mouse in the Environment tab and click on the dataset name. Alternatively, you can type `View(co2)` in your console to achieve the same outcome. Below are some other common options to understand the data structure at a glance.

```r
str(co2)
```

```
spc_tbl_ [810 x 7] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
 $ year        : num [1:810] 1958 1958 1958 1958 1958 ...
 $ month       : num [1:810] 3 4 5 6 7 8 9 10 11 12 ...
 $ decimal_date: num [1:810] 1958 1958 1958 1958 1959 ...
 $ average     : num [1:810] 316 317 318 317 316 ...
 $ interpolated: num [1:810] 314 315 315 315 315 ...
 $ trend       : num [1:810] NA NA NA NA NA NA NA NA NA NA ...
 $ days        : num [1:810] -9.99 -9.99 -9.99 -9.99 -9.99 -9.99 -9.99 -9.99 -9.99 -9.99 ...
 - attr(*, "problems")= tibble [810 x 5] (S3: tbl_df/tbl/data.frame)
  ..$ row     : int [1:810] 1 2 3 4 5 6 7 8 9 10 ...
  ..$ col     : chr [1:810] NA NA NA NA ...
  ..$ expected: chr [1:810] "7 columns" "7 columns" "7 columns" "7 columns" ...
  ..$ actual  : chr [1:810] "8 columns" "8 columns" "8 columns" "8 columns" ...
  ..$ file    : chr [1:810] "'https://gml.noaa.gov/webdata/ccgg/trends/co2/co2_mm_mlo.txt'" 
 - attr(*, "spec")=
  .. cols(
  ..   year = col_double(),
  ..   month = col_double(),
  ..   decimal_date = col_double(),
  ..   average = col_double(),
  ..   interpolated = col_double(),
  ..   trend = col_double(),
  ..   days = col_double()
  .. )
```

```r
head(co2)
```

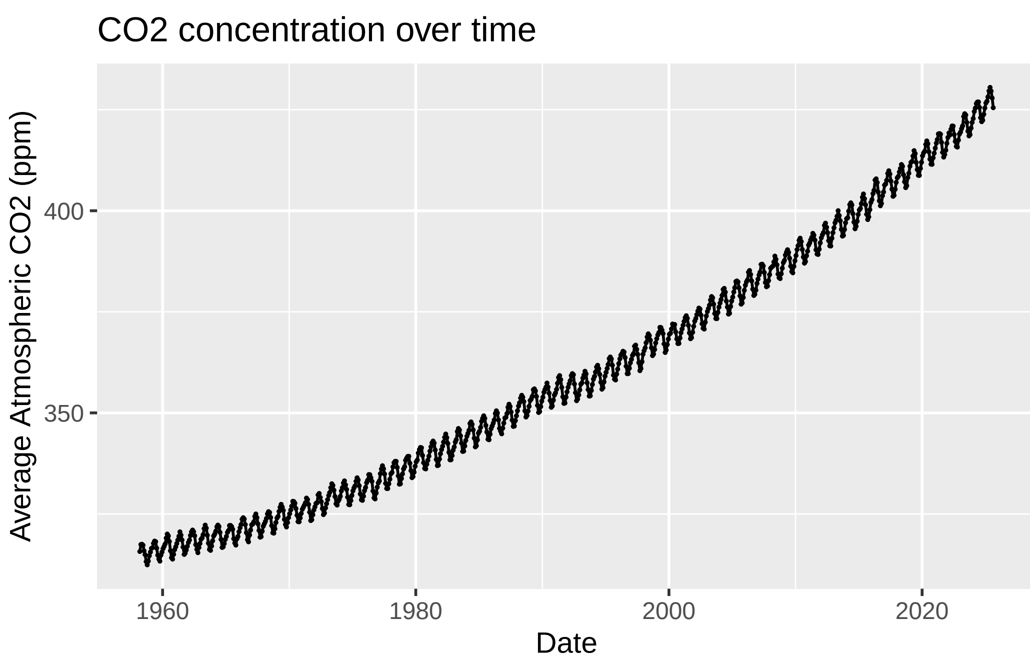
```
# A tibble: 6 x 7
   year month decimal_date average interpolated trend  days
  <dbl> <dbl>        <dbl>   <dbl>        <dbl> <dbl> <dbl>
1  1958     3        1958.    316.         314.    NA -9.99
2  1958     4        1958.    317.         315.    NA -9.99
3  1958     5        1958.    318.         315.    NA -9.99
4  1958     6        1958.    317.         315.    NA -9.99
5  1958     7        1959.    316.         315.    NA -9.99
6  1958     8        1959.    315.         316.    NA -9.99
```

Q1 (1 point)

Create a timeseries plot of average [CO_2]. Use `as.Date()` to create a formatted date for the x-axis. Label both the y-axis and add a line to connect the points.

```
co2 <- mutate(co2, date = as.Date(str_c(year, month, '15', sep = '-'),
                                  format = "%Y-%m-%d"))

ggplot(co2, aes(x = date,
                y = average)) +
  geom_point(size = 0.2) +
  geom_line() +
  labs(x = 'Date',
       y = 'Average Atmospheric CO2 (ppm)',
       title = 'CO2 concentration over time')
```

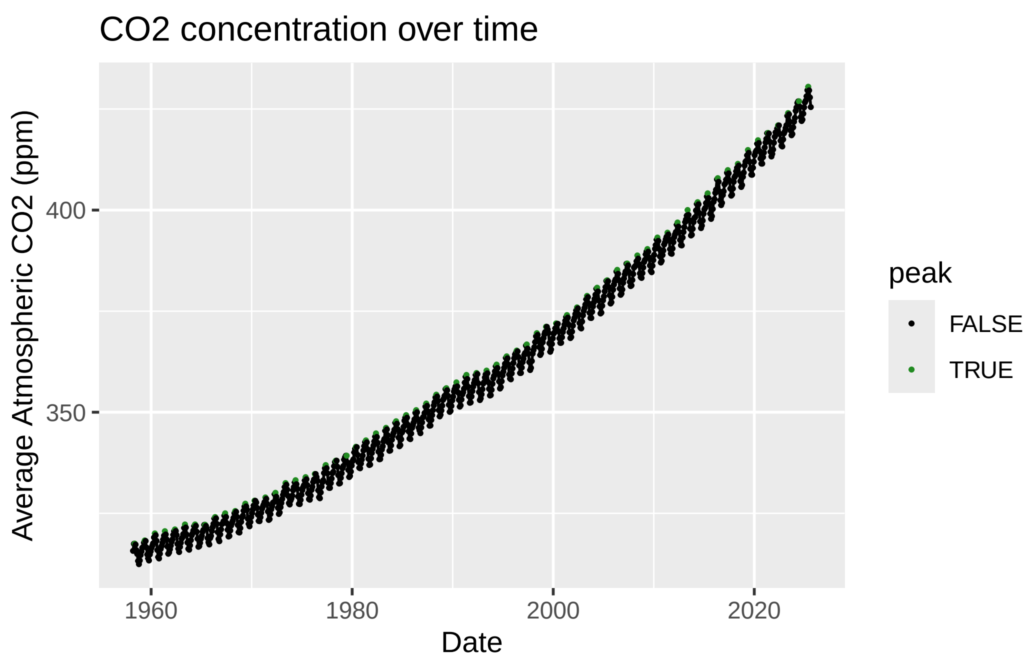


Q2 (1 point)

What months are the [CO_2] values at the maximum? Create another dataframe with these months selected and create a plot like above, but with the maximum months colored in "forest-green".

```r
co2 <- co2 |>
  group_by(year) |>
  mutate(peak = average == max(average)) |>
  ungroup()

ggplot(co2, aes(x = date,
                y = average)) +
  geom_line() +
  geom_point(aes(color = peak),
             size = 0.4) +
  scale_color_manual(values = c('black', 'forestgreen')) +
  labs(x = 'Date',
       y = 'Average Atmospheric CO2 (ppm)',
       title = 'CO2 concentration over time')
```



Q3 (1 point)

What explains the seasonal variation in [CO_2]? After reading Lewandowsky et al. 2016, how might you go about separating the long term trend and the seasonal variation in [CO_2]?

- [CO_2] is affected by photosynthesis, which does not occur at a constant rate year-round. Fluctuations in photosynthetic activity would cause correlated fluctuations in the amount of CO_2 in the atmosphere.

Global temperature

Next, we turn to a global temperature anomaly dataset from NASA GISS, similar to the GMST used by Lewandowsky et al. 2016.

Q4 (1 point)

Read some background as well as the simplified documentation for this dataset.

How are the measurements made? Describe each column in the dataset and its units. What are the resolution of the data?

Q5 (1 point)

Construct code to import dataset in a format that can be used for transforming and visualization.

Q6 (1 point)

Plot the trend in global mean temperature over time. Describe what you see and how you interpret the patterns you observe.

Q7 (1 point)

Consider the "pause" analyzed in Lewandowsky et al. 2016. Do you see evidence for such a pause?

Bonus 1 (1 point)

Recreate the analysis from Lewandowsky's Fig. 1.

Long term records of CO2

The data from Mauna Loa go back as far as modern instrumentation. To put these values into geological perspective, we must go back much further than humans have been monitoring [CO_2]. Ice sheets contain layers of snow with trapped air bubbles and can be cored and dated like tree rings.

Previously hosted by Oak Ridge National Lab, the data associated with the famed Vostok Ice Core is now archived on Zenodo. Skim Barnola et al. 1987 as needed.

After downloading this dataset, create a folder entitled `data` and save the file within the folder with a .txt extension. Read this file into R, making sure you understand the column names and units.

Q8 (1 point)

After downloading this dataset, create a folder entitled `data` and save the file within the folder with a .txt extension. Read this file into R. What are the columns and units? How were these data obtained?

Q9 (1 point)

Transform the data as needed to create a plot in chronological order. Add at least two different smoothing window averages to your plot.

Q10 (1 point)

Join your data with the Mauna Loa data and plot together, considering different colors for the different datasets.

Bonus 2 (1 point)

Write a catchy title and short description in the style of a NYtimes infographic, contrasting the rate of anthropogenic vs. background climate change for a general public audience.