ENSF 310 Project #1

Loading and Visualizing Data

Due Monday, October 10th at 11:59 pm on dropbox

In this project, you will create a visual graphic of data from a dataset of your choosing. The dataset may be of any topic, however, it must meet the following criteria:
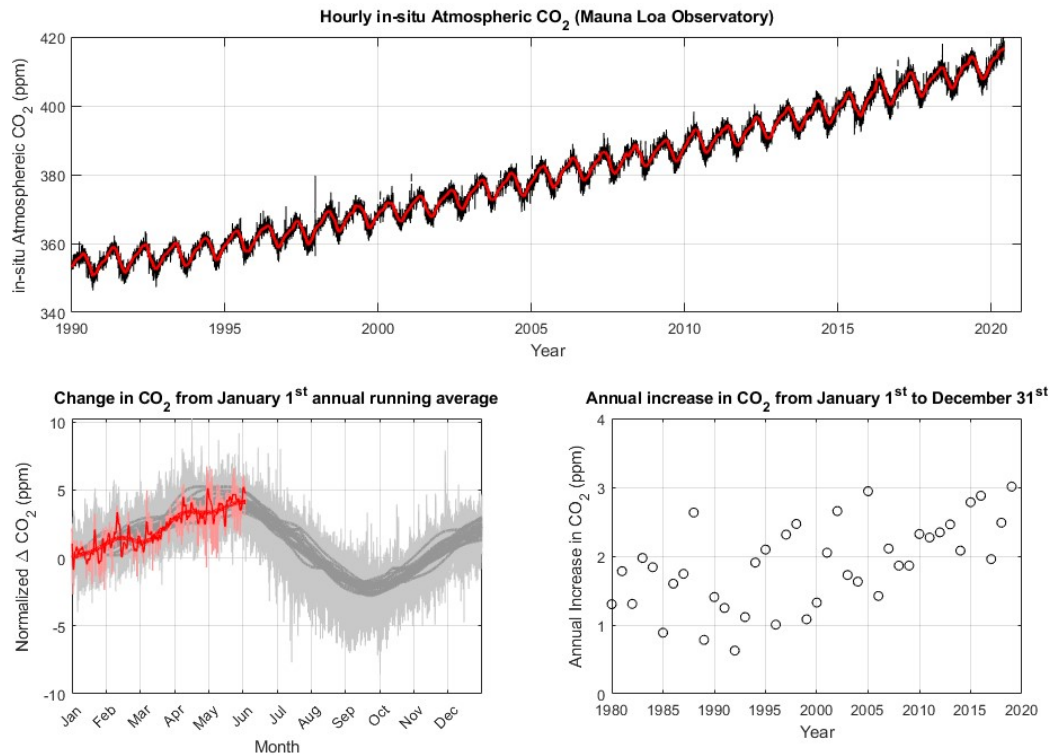
- Come from a trustworthy and verifiable public source.
- Be properly cited in your project.
- Contain at least two columns of data and consist of at least 500 rows.

With this data, you need to create a visual that summarizes the data and presents some insight. Simply plotting the data is not sufficient - you must manipulate the data to derive additional information (a running average, a trendline, annotations, etc). The visual must be easy to understand.

Examples of datasets:

- hourly CO_2 data from Mauna Loa
- hourly weather data showing seasonal and annual trends
- electricity production and demand data for Alberta from AESO
- national and/or local statistics for COVID-19

The image below is an example of a potential graphic which shows atmospheric CO_2 data since 1980 and then overlays Jan-June 2020 data over it to show the impact of the pandemic on global CO_2 levels.

Some potentially useful resources to help you get started:

- How to read a file using pandas: https://stackoverflow.com/questions/19632075/how-to-read-file-with-space-separated-values-in-pandas (https://stackoverflow.com/questions/19632075/how-to-read-file-with-space-separated-values-in-pandas)
- Filtering data in pandas: https://stackoverflow.com/questions/18172851/deleting-dataframe-row-in-pandas-based-on-column-value (https://stackoverflow.com/questions/18172851/deleting-dataframe-row-in-pandas-based-on-column-value)
- Moving average and plotting: https://towardsdatascience.com/moving-averages-in-python-16170e20f6c (https://towardsdatascience.com/moving-averages-in-python-16170e20f6c)
- Plotting data in Matplotlib: https://matplotlib.org/gallery/lines_bars_and_markers/csd_demo.html#sphx-glr-gallery-lines-bars-and-markers-csd-demo-py (https://matplotlib.org/gallery/lines_bars_and_markers/csd_demo.html#sphx-glr-gallery-lines-bars-and-markers-csd-demo-py)

Requirements

- You may use .py files, however, you should use an annotated Notebook to explain your workflow.
- Your final submission must be uploaded to dropbox.
- Your final graphic should have at least two subplots / subfigures and have at least four different data series.
- Plots should be well labeled and clear.
- You may use any online resource for inspiration, however, you must cite all sources used outside of the course notes and Think Python textbook.

In [ ]: