**Homework 6 Jupyter Notebooks and CO2 Data**

Let’s use your container from lab 4 and pull in some “Live” data. You will want to start your data science container and go into Jupyter in the browser. Code will be in red to make it easy to recognize.

(If you have forgotten how to set up your container, just look back at the last lab or try some suggestions below)

*The easiest way to start a container is to go into the docker desktop program and start the existing one your created . Then remember to double click on the container name in desktop to see the command line and get the 127…. url to long on.*

*You can also start the container in command line by typing “docker start - i “ and the name of your container. You can check the name of your containers with “docker container ls –a” or looking in the docker app on your computer.*

If you really get stuck with the container setup, this lab can be done using the online jupyter notebooks.

<https://jupyter.org/try> (Use the classic Notebook)

1. Now let’s pull in some data. We are going to start pretty simple with some csv formatted data. CSV is a file with the data separated by commas. For example you can save an excel file as csv data.
   1. First let’s import the libraries we will need for working with this data. Libraries just add in some additional tools to Jupyter for our program.

import requests

#This library helps us grab the data from the usgs site

#Run the code

import pandas as pd

#This library helps us read the file we grab from the usgs site

#Run the code

import matplotlib.pyplot as plt

#This library helps us plot the data in a chart

#Run the code

* 1. Now let’s pull in the Carbon Dioxide data from Mana Loa in Hawaii and create a chart. Do a screen shot of your chart for the submission. (Use the submission document)

dataCO2 =pd.read\_csv('https://pkgstore.datahub.io/core/co2-ppm/co2-mm-mlo\_csv/data/5d0e2bc9cc70b8dd8590a31d7ab4526b/co2-mm-mlo\_csv.csv')

print(dataCO2)

#This code grabs the csv file for CO2 data

#Then it prints that data for us to look at.

#Run the Code

dataCO2 =pd.read\_csv('https://pkgstore.datahub.io/core/co2-ppm/co2-mm-mlo\_csv/data/5d0e2bc9cc70b8dd8590a31d7ab4526b/co2-mm-mlo\_csv.csv', usecols = ['Interpolated'])

#This grabs just the CO2 column

plt.plot(dataCO2)

#plots the data

* 1. Great! Now we can analyze the chart. What is the trend of Carbon Dioxide over the last 50 years or so?
  2. Notice that Carbon Dioxide has a high point every year and a low point. This up and down variation is labeled as regular variability. When is the high point during the year and when is the low point during the year and why? (To answer the variability question you may want to watch my climate change videos and also examine the data more closely) (Answer the questions on the submission document.)
  3. Use the following code snippets below if you need to examine the data more closely to answer the variability question.

pd.set\_option('display.max\_rows', dataCO2.shape[0]+1)

#this allows us to see all the data

dataCO2 =pd.read\_csv('https://pkgstore.datahub.io/core/co2-ppm/co2-mm-mlo\_csv/data/5d0e2bc9cc70b8dd8590a31d7ab4526b/co2-mm-mlo\_csv.csv')

print(dataCO2)

#prints all the data now