Episode 10 Homework

Copy the file bodyfat.csv from the Data_Files folder either locally or from Github. Write a program that does the following:

Look up the formula for body-mass index (BMI). Write a function that computes the BMI for metric units (the international definition) or Imperial units. The function will need to accept weight, height, and some indicator of the units system. Use a try/except to check for a ZeroDivisionError exception when you compute the BMI. Make sure your function is a ufunc.

Test your BMI function by choosing some realistic values for height and weight and hand-computing examples. Be sure to test both units system. You may use an online calculator if you wish to compare your results, but do at least one example for each units system by hand.

When you are sure your function is working, write your `main()` function. Read the file and store the values of percent body fat, weight, and height into NumPy arrays. Note the units on the original data. Put a try/except block around your read in case there is a problem opening or reading the file.

Plot BMI versus bodyfat. Recall that this means BMI is on the vertical axis. You will need to use a scatter plot since the points are uncorrelated with one another. Do you see anything unusual? Find a way to remove the outlier point. (It is due to a probable typo in the data.) Write another function that takes the BMI, finds the locations of outliers, and returns a *mask* array consisting of Boolean values. Remember that you can use a Boolean array as an "index" to another array, so create corrected BMI and bodyfat arrays.

Use scipy to compute the regression line and correlation coefficient between BMI and bodyfat. Print the value of the correlation coefficient. Change the plot you made before to plot the corrected BMI with its regression line superimposed.