

HW4 is due on Oct 20th, at 1pm.

On Brightspace, you should find two corresponding files:

1. HW4_Patient_FBG has 3 fasting blood glucose (FBG) measures across 3 years for 200 patients. It is almost identical to the FBG file you all used in HW3 with one important difference; There are now missing values in the file (denoted as 'NA').
2. HW4_Patient_Sex has the same set of patients and their reported Sex

In this assignment you are expected to:

1. Determine the diabetic status for each patient.
2. Find how many of each status there are and break it down by Sex. Write out your answer to either a '.txt' file or a '.csv' file.
3. Use at least one function of your own. This function should be able to accept parameter(s), perform some operation(s) and return result(s) when necessary. We would like everyone to practice on this topic, so please try to write meaningful functions. Below are examples you may or may not want to follow:

a. Bad examples:

```
def print_msg():  
    print('Help message')
```

```
def write_output(file_name):  
    with open(file_name) as fh:  
        fh.write('output')
```

b. A good example:

```
def average(my_list):  
    total = 0  
    for value in my_list:  
        total = total + float(value)  
    return total/len(my_list)
```

4. Import modules to read and write csv are not allowed, such as csv, pandas, etc.

Rules for determining the diabetic status

There are three possible statuses: Non diabetic, Diabetic, and NA.

1. Non diabetic:
 - If a subject has only fasting blood glucose measure of 125 and below
2. Diabetic:
 - If a subject has ANY fasting blood glucose measure of 126 and above
3. NA: If a subject has no measures at all (only 'NA')

Examples:

	Year_1	Year_2	Year_3	
Patient_1	121	110	94	<--- Non diabetic
Patient_2	150	90	121	<--- Diabetic
Patient_3	98	160	121	<--- Diabetic
Patient_4	NA	NA	106	<--- Non diabetic
Patient_5	101	NA	153	<--- Diabetic
Patient_6	NA	NA	NA	<--- NA

Output format

Your output should look like something like this (though feel free to format it however you want):

Sex	Non diabetic	Diabetic	NA
Male	34	24	4
Female	43	41	1

OR

Sex,Non diabetic,Diabetic,NA

Male,34,24,4

Female,43,41,1

Notes and tips:

- The Patient_Sex.csv file is not ordered by the patient number
- I recommend that your code reads through the Patient_sex file first and that you store that information somewhere (perhaps a Dictionary or a list might be useful...)
- You do not need to send to us the diabetic status for each patient; All we need to is the total #s for each category (see example answer)
- Your code will have to check if a value is 'NA' or not, an easy way to do this is to use the ".isnumeric()" function. This function returns true if your value is a number (even if it's technically a string) and false if the value is not a number. A simple if statement checking if the value is 'NA' would also work just fine.
- Making a function which accepts a series of fbg values and returns diabetic status could make things easier for you, but feel free to approach this problem however you want

Good luck!