CS 620–Introduction to Data Science and Analytics, HW2

For this problem you will use part of the actual customer dataset called "custdata.csv" which contains 1000 records of customer data including the income, marital status, and other information.

Download and start with the skeleton code provided. Warning: **DO NOT CHANGE** the driver code. You are free to experiment with the driver code, but don't forget to revert it back to the original form before the submission.

Note that each "for loop" used in the solution will incur -3pt penalty. Instead, use DataFrame techniques using pandas. Hint: *subset = df[(df.columnA == "test") & (df.columnB >= 100)]*

I. **(40 pts)** By using the data from "custdata.csv", do the following.

    a. Extract a subset of customers that are <u>Married</u> and has an <u>income</u> greater than $40,000.
    b. How many of them are <u>renters</u> <u>younger</u> than 45?
    c. How many of them have <u>recently moved</u> to <u>Texas</u>?

II. **(60 pts)** By using the subset calculated from part (I(b)) above, do the following.

    a. Assume that all customers are married filing jointly for their tax. If the standard deduction for Married Filing Jointly is $24,800, write a function to calculate the taxable income of a customer.

    b. Using the part (a) taxable income, write a function to calculate the total tax that a customer is supposed to pay for the IRS.

    You'll be using the following tax brackets for your calculations. Hint: *if-conditions*

Taxable Income Brackets	Tax Rate	Tax Owed from each Bracket
0-20,000	0 if income is less than $20,000. Otherwise 10%	2,000
20,001-80,000	12%	7,200
80,001-160,000	22%	17,600
160,001-300,000	24%	33,600
300,001-400,000	32%	32,000
400,001-600,000	35%	70,000
600,001 or more	37%	

    For example, if your taxable income is $350,000 then you'll use the following equation to generate the tax amount, where S is the Start of the income bracket ($300,001), R is the tax rate (32%) and A is the total tax amount owes from previous brackets ($2000+ $7,200+$$17,600+$33,600 = $60,400).

$$tax = (income - S) * R + A$$
$$tax = (350,000 - 300,001) \times 32\% + 60,400 = \$76,399.68$$

    c. Use the functions from part (a) and (b) to calculate the tax amount owed by each customer and display the custid, income, tax_owed, and state_of_res

What to turn in (to Blackboard): Due: Sunday, Feb. 18, 11.59pm

      Lastname-hw2.py should contain the following information at the top:

CS620
HW2
@author