Opeyemi Adesina, PhD
Assistant Professor
School of Computing
University of the Fraser Valley
Office: C2435, Abbotsford Campus
Tel: (604) 504-7441 (ext: 4931)
opeyemi.adesina@ufv.ca

Assignment 3 — Dictionaries & File Processing

COMP 150 : Introduction to Programming

(100 points)

When Due : April 4, 2022 – 23:59:00 (PDT) [Submission via Blackboard]

Brief Description

This assignment amounts to **10%** of the entire course grade. In particular, whatever your obtains as a score will be scaled to this value for final grade computation. You are required to work **ALONE**. No late submission will be permitted (see deadline above).

The goal of this assignment is to assess your knowledge and skills on Control structures, while developing skills to map requirements to program code. You will find a grading scheme at the end of this document – to guide you on instructor's expectations while preparing your submission.

Program Requirements

1. **Province of Newfoundland and Labrador's (NL) Income Tax Calculator:** Employees within the province of NL are taxed at different levels - Federal and Provincial. Other deductions like Canadian Pension Plan (CPP), Employment Insurance, and Tax amounts are computed according to salary brackets. (See Payroll Deduction for NL Province – Effective January 1, 2022)

   NL's (i.e., provincial) tax rates and thresholds are given in table 2 while that of federal is given in table 1. Other deductions include CPP (Canada Pension Plan) and EI (Employment Insurance). While CPP is 5.7% of the gross income (but to a maximum amount of $3,499.80), EI is 1.58% of the gross income (but to a maximum amount of $952.74).

Annual Taxable Income (range of incomes - $)	Federal Tax Rate (%)
0.00 ... 50,197.00	15.0
50,197.01 ... 100,392.00	20.5
100,392.01 ... 155,625.00	26.0
155,625.01 ... 221,708.00	29.0
221,708.01 (and above)	33.0

Table 1: 2022 federal tax rates and income thresholds

Annual Taxable Income (range of incomes - $)	Provincial Tax Rate (%)
0 ... 39,147.00	8.7
39,147.01 ... 78,294.00	14.5
78,294.01 ... 139,780.00	15.8
139,780.01 ... 195,693.00	17.8
195,693.00 ... 250,000.00	19.8
250,000.01 ... 500,000.00	20.0
500,000.01 ... 1,000,000.00	21.3
1,000,000.01 (and above)	21.8

Table 2: 2022 NL's tax rates and income thresholds

You are to note that the computations for provincial and federal taxes have been revised in comparison with the first assignment. For example, if gross income is $92,486.79 the taxes are computed below:

- **Provincial Tax: $11,234.56** = $(3,405.79 + 5,676.31 + 2,242.46)
  - 1st category $39,147 is tax at the rate of 8.7% $\implies$ $3,405.79
  - 2nd category ($78,294 - $39,147.01) is taxed at the rate of 14.5% $\implies$ $5,676.31
  - Remainder (i.e., $92,486.79 - $78,294.01) is taxed at the rate of 15.8% $\implies$ $2,242.46
- **Federal Tax: $16,198.95** = $(7,529.55 + 8,669.40)
  - 1st category $50,197.00 is tax at the rate of 15.0% $\implies$ $7,529.55
  - Remainder (i.e., $92,486.79 - $50,197.01) is taxed at the rate of 20.5% $\implies$ $8,669.40

In addition to the deductions above assuming NL's employees are charged health premiums and that the premiums are based on the following rules:

(a) If the taxable income is less than or equal to $22,000.00, the premium is $0.00.

(b) If the taxable income is greater than $22,000.00 and less than or equal to $38,000.00, the premium is equal to the lesser of: (i) $300.00 and (ii) 6% of (taxable income - $22,000.00);

(c) If the taxable income is greater than $38,000.00 and less than or equal to $50,000.00, the premium is equal to the lesser of: (i) $450.00 and (ii) 300 + 6% of (taxable income - $38,000.00);

(d) If the taxable income is greater than $50,000.00 and less than or equal to $74,000.00, the premium is equal to the lesser of: (i) $600.00 and (ii) 450 + 25% of (taxable income - $50,000.00);

(e) If the taxable income is greater than $74,000.00 and less than or equal to $202,000.00, the premium is equal to the lesser of: (i) $750.00 and (ii) 600 + 25% of (taxable income - $74,000.00);

(f) If the taxable income is greater than $202,000.00, the premium is equal to the lesser of: (i) $900.00 and (ii) 750 + 25% of (taxable income - $202,000.00).

Write a program that:

- **take gross income of N employees from an excel file (*.xls or *.xslx format), convert the file to its JSON equivalent (for interoperability purposes);**
- **read gross income of N employees from the generated JSON file and do the following:**
  - **compute their net incomes (i.e., after deducting appropriate provincial and federal level taxes);**
  - **compute and display population variance, and median for the entire net incomes;**
  - **display means (or averages) of all gross and net incomes; and**

– **average of all modes (use Dictionary data structure – don't copy online codes) for the entire net incomes.**

Things to note:

- **Median** – Computations differ for even and odd lengths of any given list.
  - Even length: let $lst = [1, 2, 8, 5, 6, 3]$, $lst = sort(lst) \implies median(lst) = (3 + 5)/2 = 4$
  - Odd length: let $lst = [1, 2, 8, 5, 6]$, $lst = sort(lst) \implies median(lst) = 5$
- **Mode** – By definition, it is the net income with the highest frequency. But since there may be more than one net income with same but highest frequency – you are to compute the list of all net incomes with highest number of occurrence.
  For example, let $lst = [5, 4, 3, 2, 1, 3, 2, 1] \implies modes(lst) = [3, 2, 1]$
- **Excel File** – Two files of data (randomly generated) in excel formats. The files contain exactly the same data – you may use any of them in your program. If you decide to try both of them out – they should give the same output.
- **Variance** – This is the square of standard deviation, that is $Var([x_0, x_1, \ldots, x_{N-1}]) = \sigma^2$.

$$\sigma^2 = Var([x_0, x_1, \ldots, x_{N-1}]) = \frac{\sum_{i=0}^{N-1}(x_i - k)^2}{N}, \tag{1}$$

such that:
$$k = \mu([x_0, x_1, \ldots, x_{N-1}]) = \frac{\sum_{i=0}^{N-1} x_i}{N} \tag{2}$$

  HINT: $\sum_{i=1}^{n} x_i$ with $x_i \in [1, 2, 3, 4, 5, 6] \implies 1 + 2 + 3 + 4 + 5 + 6 = 21$. You are to re-use the standard deviation function you developed in assignment 2.
- **Reading Excel and Generating JSON** – In the attached stubs, you will find a working code to generate JSON and extract a list of of gross incomes from the input excel file. Your task is to engineer these functions in order to produce a list of gross incomes to compute median and mode.
- You may need to update your codes for mean and standard deviation to handle edge cases for the tests to pass.

An example test run of the program is given in figure 1. Thus, your program should produce the following output (using the same format and set of taxable incomes) if everything is right:

```
C:\Users\adesino\Documents\UFV\COMP 150\W22\Assignments\A3\codes>main.py
JSON File Successfully Generated!!!

************* Required Information *************

Average gross-income for employees: $123,099.00 (CAD)
Average net-income for employees: $76,332.30 (CAD)
Average of all modes for employees: $75,676.08 (CAD)
Population variance for employees: $1,261,432,584.82 (CAD)
Median for employees: $75,923.61 (CAD)

//.............Running testcases for the operations.............//
Test 0, Q1 passed. [Alloted point(s): 1.0]
Test 1, Q2 passed. [Alloted point(s): 1.0]
Test 2, Q3 passed. [Alloted point(s): 1.0]
Test 3, Q4 passed. [Alloted point(s): 1.0]
Test 4, Q5 passed. [Alloted point(s): 1.0]
Test 5, Q6 passed. [Alloted point(s): 1.0]
Test 6, Q7 passed. [Alloted point(s): 1.0]
Test 7, Q8 passed. [Alloted point(s): 1.0]
Test 8, Q9 passed. [Alloted point(s): 17.0]
Test 9, Q10 passed. [Alloted point(s): 5.0]
Test 10, Q11 passed. [Alloted point(s): 20.0]
All tests are successful.
Total points earned in this assignment: 50.0 (out of) 50.0 points
//.............................End.............................//
Time taken for execution: 1.44 seconds.
```

Figure 1: Expected Output - if your program works correctly.

Please note that the program has been developed such that the entry-point for program execution is ***main.py***. This is the coordinating file - by executing this file, you are indirectly executing other files (in order to realize the program objectives). We have relied on the power of *modules* to integrate the system.

Grading Scheme

The following scheme will be used to grade your submission. Therefore, you may also use it as a guide in preparing your deliverable.

Grade Item	Weight
A syntactically and semantically correct program.	**50**
A program with detailed program design - see examples in my video lectures (it was embedded as comments in the video). I am enforcing program design this time.	**15**
A flowchart of your program must be included. This must represent your program. You can use pen and paper for drawings. You may also decompose the problem into separate components, draw them each and show how the components will be coupled to realize your code.	**20**
Program efficiency. That is, using efficient but correct control structures.	**5**
A program with detailed program documentation and uses sensible variable names. Your program's file name and other files should be zipped and named in the following format - [firstName_lastName_studentID]	**10**
Total	**100**