CSE 465/565
Spring 2023
Homework #6
100 points

*For this homework, you will work independently to solve a problem in Python 3. Your program file will be called **hw6.py**.*

*Part of the grading for this assignment will be based on your usage of Python 3 features and **data structures** that are appropriate to the problem. You should utilize language features that make your program shorter, more readable, and/or more reliable. You will be asked to submit a PDF called **hw6.pdf** in which you explain the language features you have chosen and why. Your explanation should be between half a page and one page long.*

*Submit a **zip archive** that contains your two files: **hw6.py** and **hw6.pdf**.*

Note: The next homework assignment (HW 7) will ask you to solve the same problem in C#.

Mail merge program. Your program will be invoked on ceclnx01 as shown here:

```
> python3 hw6.py a.tsv a.tmp
```

The first command line argument is a **tab separated** file of records -- **a.tsv** (tsv stands for "tab separated values"). The second one is a **form letter** with embedded codes -- **a.tmp**. Sample .tsv and .tmp files are provided in the archive TestFiles.zip. Matching files have matching names: b.tsv and b.tmp; c.tsv and c.tmp, etc. Make sure that your program works for all pairs of files with matching names, not just on the pair a.tsv a.tmp.

Here are the contents of the **a.tsv** file:

NAME	ID	COURSE	DUE	SUBMITTED	MINUTESLATE
Steve Smith	smiths	465	2/5/2016 23:59:00	2/6/2016 0:45:00	46
Mark Watson	watsonm	465	2/5/2016 23:59:00	2/6/2016 3:47:00	228
Kenny Briddle	briddlek	465	2/5/2016 23:59:00	ontime	0
Bill Haygood	haygoodb	565	2/5/2016 23:59:00	ontime	0

(continuation)

LATEDEDUCTION	P1	P1COMMENTS	P2	P2COMMENTS	SUBTOTAL	TOTAL
1	35	Excellent work	65	Nice	100	99
5	30	Test case 10 provides incorrect output; otherwise, nicely done.	65	Well done	95	90
0	15	Many test cases missing.	55	Missing the report.	70	70
0	0	Not attempted	65	Nice	65	65

Here are the contents of the **a.tmp** file:

```
Name: <<NAME>>   (<<COURSE>>)
ID: <<ID>>
Total: <<TOTAL>>/100  Subtotal: <<SUBTOTAL>>           Total deductions:
<<LATEDEDUCTION>>

Time due: <<DUE>>
Submitted: <<SUBMITTED>>
Late minutes: <<MINUTESLATE>>
Late deduction: <<LATEDEDUCTION>>

Problem 1: <<P1>>/35
<<P1COMMENTS>>

Problem 2: <<P2>>/65
<<P2COMMENTS>>
```

When run, your program should produce one output file for each record in the file. The file should be named using the ID column. In this case, the four files should be smiths.txt, watsonm.txt, briddkek.txt, and haygoodb.txt. Here is one of the output files:

```
Name: Steve Smith     (465)
ID: smiths
Total: 99/100   Subtotal: 100         Total deductions: 1

Time due: 2/5/2016 23:59:00
Submitted: 2/6/2016 0:45:00
Late minutes: 46
Late deduction: 1

Problem 1: 35/35
Excellent work

Problem 2: 65/65
Nice
```

Notes:

- All columns in the tsv file will have a unique name.
- One of the tsv columns will have the name ID.
- Any string inside the tmp file having the form <<*letters*+>> is considered a tag.
- The field values in the tsv file may contain << and >>. These values are to be treated literally and not to be substituted as a tag.

Grading:

- Implementation (hw6.py) – 90 points
  - 10 points: correct basic implementation (e.g., reading input filenames from the console)
  - 10 points: correct output filenames
  - 10 points for each test case (a – f) for correct output results = 60 points
  - 10 points: appropriate usage of Python features
- Report and rationale for features used (hw.pdf) – 10 points