SIT384 Cyber security analytics

Credit Task 3.2C: Process data using Pandas

Task description:

You are given a student result data file (result_withoutTotal.csv). It has columns:

ID: student id

Ass1 ~ Ass4: assignment scores (out of 100); weight of ass1, ass2, ass3 and ass4 is 5%, 15%, 5%, and 15%, respectively.

Exam: examination score (**out of 120**); weight is **60%**.

ID	Ass1	Ass2	Ass3	Ass4	Exam
1	89.1	50	85	88.9	65
2	95.1	82.5	90.5	94.5	52
3	74.3	54.4	63	63.9	31
4	89.8	81.3	82	90.4	37
5	91.3	98.8	92.5	95.9	79
…	…	…	…	…	…

(The above data is for demonstration purposes only. Please download the full version of result_withoutTotal.csv.)

Total score can be calculated using formula:

Total = 5%*(ass1+ass3) + 15%*(ass2+ass4) + 50%*exam (as exam is out of 120)

Read students' result data from file result_withoutTotal.csv,

add:

- Total column: Total = 5%*(ass1+ass3) + 15%*(ass2+ass4) + 50%*exam.
- Final column: Final = Total score rounded to the nearest integer.
    - To pass the unit, a student must achieve at least 50 of the Total and 40% of Exam which is 48 out of 120 (or Total >= 50 and Exam >= 48).
    - If a student failed the hurdle (Exam >= 48), the **max** Final is 44. No change to Final score if Final <44 already.
- Grade column: N (Final <=49.45), P (49.45 < Final <=59.45), C (59.45 < Final <=69.45), D (69.45 < Final <=79.45) and HD (79.45 < Final). Border values are as follows:

HD	
D	79.45
C	69.45
P	59.45
N	49.45

save:

- the result data file with the 3 new columns to a file called result_updated.csv.
- the students' records with exam score < 48 to a file called failedhurdle.csv.

display:

- the result data file with the 3 new columns
- the students with exam score < 48 (these who failed the hurdle)
- the students with exam score > 100

(Hints: import pandas, use DataFrame, DataFrame.loc and display)

(Sample output as shown in the following figure is for demonstration purposes only.)

result_updated:

ID	Ass1	Ass2	Ass3	Ass4	Exam	Total	Final	Grade
1	89.1	50.0	85.0	88.9	65	62.040	62	C
2	95.1	82.5	90.5	94.5	52	61.830	62	C
3	74.3	54.4	63.0	63.9	31	40.110	40	N
4	89.8	81.3	82.0	90.4	37	52.845	44	N
5	91.3	98.8	92.5	95.9	79	77.895	78	D
6	83.9	82.5	89.0	98.6	68	69.810	70	D
7	81.9	50.0	68.5	95.4	59	58.830	59	P
8	50.0	54.9	50.0	87.7	51	51.890	52	P
9	90.5	65.9	50.0	72.2	63	59.240	59	P
10	89.0	89.9	94.0	90.3	84	78.180	78	D
11	96.6	100.0	98.0	97.3	102	90.325	90	HD

failedhurdle.csv:

ID	Ass1	Ass2	Ass3	Ass4	Exam	Total	Final	Grade
3	74.3	54.4	63	63.9	31	40.11	40	N
4	89.8	81.3	82	90.4	37	52.845	44	N
15	66.3	53.7	53	81.9	30	41.305	41	N
24	57.7	76.3	71	87.7	35	48.535	44	N
25	84.7	65	73	88.9	34	47.97	44	N
26	84.7	53.8	75	78.1	36	45.77	44	N
33	64.2	50	18	0	0	11.61	12	N
42	81.5	43.8	0	0	0	10.645	11	N
44	71.9	61.3	76	94.5	38	49.765	44	N
47	50	71.3	56	93.8	34	47.065	44	N
54	76.1	50	50	50	33	37.805	38	N
60	73.9	53.2	74	95.9	34	46.76	44	N
78	52.9	53.2	50	50	36	38.625	39	N

Display output:

```
students with exam score < 48
    Ass1  Ass2  Ass3  Ass4  Exam   Total  Final Grade
ID
3   74.3  54.4  63.0  63.9    31  40.110     40     N
4   89.8  81.3  82.0  90.4    37  52.845     44     N
15  66.3  53.7  53.0  81.9    30  41.305     41     N
24  57.7  76.3  71.0  87.7    35  48.535     44     N
25  84.7  65.0  73.0  88.9    34  47.970     44     N
26  84.7  53.8  75.0  78.1    36  45.770     44     N
33  64.2  50.0  18.0   0.0     0  11.610     12     N
42  81.5  43.8   0.0   0.0     0  10.645     11     N
44  71.9  61.3  76.0  94.5    38  49.765     44     N
47  50.0  71.3  56.0  93.8    34  47.065     44     N
54  76.1  50.0  50.0  50.0    33  37.805     38     N
60  73.9  53.2  74.0  95.9    34  46.760     44     N
78  52.9  53.2  50.0  50.0    36  38.625     39     N

students with exam score > 100
    Ass1   Ass2  Ass3   Ass4  Exam   Total  Final Grade
ID
11  96.6  100.0  98.0   97.3   102  90.325     90    HD
84  93.6  100.0  96.0  100.0   106  92.480     92    HD
```

Submission:

Submit the following files to OnTrack:

1. Your program source code (e.g. task3_2.py)
2. result_updated.csv/**txt** (save result_updated.csv **as .pdf** if .csv is not supported by OnTrack) with the 3 newly added columns Total, Final and Grade
3. failedhurdle.csv/**txt** (save failedhurdle.csv **as .pdf** if .csv is not supported by OnTrack) generated by your code
4. A screen shot of your program running (only “exam <48” and “exam >100” output required)

Check the following things before submitting:

1. Add proper comments to your code