

IT 166 Lab 9

pandas

Objectives

- Be able to use pandas to solve problems in Python.

Preparation

- Launch the Jupyter notebook.
- Rename the notebook page as “lab9”.
- Solution to one problem should occupy one cell.

Please provide solutions to the problems below.

Problem 1

ID	Name	Age	Midterm_1	Midterm_2
1	Alice	19	100	80
2	Bob	21	85	99
3	Caro	20	90	85
4	David	22	77	75

Given the above table:

- 1) Create a DataFrame, with ID as the index. The columns should be in the exact order as they are in the table.
- 2) Compute the average of both midterm scores, and save them into a new column, named “Average”.
- 3) Add one more column, named “Major”, set Alice and David’s major to IT; set Caro’s major to Mathematics; set Bob’s major to NaN.
- 4) Replace the index with the names. (You will need to get rid of the Name column after it has been used as index.)

Expect outcomes:

	Name	Age	Midterm_1	Midterm_2	Average	Major
1	Alice	19	100	80	90.0	IT
2	Bob	21	85	99	92.0	NaN
3	Caro	20	90	85	87.5	Mathematics
4	David	22	77	75	76.0	IT

	Age	Midterm_1	Midterm_2	Average	Major
Name					
Alice	19	100	80	90.0	IT
Bob	21	85	99	92.0	NaN
Caro	20	90	85	87.5	Mathematics
David	22	77	75	76.0	IT

Problem 2

Use NumPy's random number generator (randn) and pandas to create a DataFrame that has a shape of 5 by 5. Use only pandas or NumPy functions to solve the following problems:

- 1) Compute the sum of all the positive numbers and the sum of all the negative numbers.
- 2) Select all rows having a value exceeding 1.8 or -1.8
- 3) Drop a row and drop a column.

Expect outcomes:

```

      0      1      2      3      4
0 -0.551642 -0.917440  2.192903 -1.403199  0.819914
1 -1.317458  0.372976 -0.037259  1.046262  1.831721
2 -0.210886  0.869276  1.851818  1.465877  0.642305
3 -1.132067  0.157274 -0.793022 -0.088709  0.323303
4 -0.557636 -0.340715 -0.204136  0.749714  1.260035

```

```

The sum of all the positive values: 13.583378195698685
The sum of all the negative values: -7.5541695948812695

```

Rows that meet the requirement:

```

      0      1      2      3      4
0 -0.551642 -0.917440  2.192903 -1.403199  0.819914
1 -1.317458  0.372976 -0.037259  1.046262  1.831721
2 -0.210886  0.869276  1.851818  1.465877  0.642305

```

After dropping the 3rd row:

```

      0      1      2      3      4
0 -0.551642 -0.917440  2.192903 -1.403199  0.819914
1 -1.317458  0.372976 -0.037259  1.046262  1.831721
3 -1.132067  0.157274 -0.793022 -0.088709  0.323303
4 -0.557636 -0.340715 -0.204136  0.749714  1.260035

```

After dropping the 2nd column:

```

      0      2      3      4
0 -0.551642  2.192903 -1.403199  0.819914
1 -1.317458 -0.037259  1.046262  1.831721
2 -0.210886  1.851818  1.465877  0.642305
3 -1.132067 -0.793022 -0.088709  0.323303
4 -0.557636 -0.204136  0.749714  1.260035

```