Name: Solutions

SA433 · Data Wrangling and Visualization   Fall 2024 · Uhan

Quiz 3 – 10/23/2024

Instructions. You have 15 minutes to complete this quiz. You may <u>not</u> use any outside materials. This quiz is closed computer, closed book, closed notes, and closed internet. No collaboration allowed.

The problems in this quiz are based on the two DataFrames below, `first` (on the left) and `second` (on the right).

	a	b	c	d
0	A	4	3	-2.1
1	C	-3	6	3.3
2	B	1	7	7.4
3	A	8	-2	5.6

	a	b	c	d
3	B	-2	9	4.7
2	C	7	8	8.3
1	B	4	3	0.1
0	A	5	5	4.5

Problem 1. What is the result of the code snippet below?

```
first['b'] * first['c']
```

Your answer:

B

A.

```
0    7
1    3
2    8
3    6
dtype: int64
```

B.

```
0     12
1    -18
2      7
3    -16
dtype: int64
```

C.

```
0     1
1    -9
2    -6
3    10
dtype: int64
```

D. None of the above.

Problem 2. What is the result of the code snippet below?

```
first['b'] + second['b']
```

Your answer:

A

A.

```
0    9
1    1
2    8
3    6
Name: b, dtype: int64
```

B.

```
0    -1
1    -7
2    -6
3    10
Name: b, dtype: int64
```

Recall from Lesson 12 that Pandas automatically aligns the data by the index label when performing arithmetic operations.

C.

```
0     2
1     4
2     5
3    13
Name: b, dtype: int64
```

D.   None of the above.

Problem 3. What is the result of the code snippet below?

Your answer:

```
first.query('a not in ["B", "C"]')
```

B

A.

	a	b	c	d
1	C	-3	6	3.3
2	B	1	7	7.4

B.

	a	b	c	d
0	A	4	3	-2.1
3	A	8	-2	5.6

C.

	a	b	c	d
2	B	1	7	7.4

D. None of the above.

Problem 4. What is the result of the code snippet below?

Your answer:

```
second.query('c < 1')
```

D

A.

	a	b	c	d
3	B	-2	9	4.7
2	C	7	8	8.3
1	B	4	3	0.1
0	A	5	5	4.5

B.

	a	b	c	d
1	B	4	3	0.1

C.

	a	b	c	d
3	B	-2	9	4.7

D. None of the above.

Problem 5. Suppose you have a list called desired defined as follows:

```
desired = ["A", "C", "F"]
```

Using .query(), write valid code that keeps the rows of the DataFrame second whose value in the column a is in the list desired. You must refer to desired <u>directly</u> (that is, you may not write out the list ["A", "C", "F"] in .query()).

```
second.query('a in @desired')
```

Recall from Lesson 13 that we can refer to variables from the Python environment in .query() by prefixing them with @.

Problem 6. What is the result of the code snippet below?

```
first.sort_values(['a', 'c'])
```

Your answer:

C

A.

	a	b	c	d
0	A	4	3	-2.1
3	A	8	-2	5.6
2	B	1	7	7.4
1	C	-3	6	3.3

B.

	a	b	c	d
3	A	8	-2	5.6
0	A	4	3	-2.1
1	C	-3	6	3.3
2	B	1	7	7.4

C.

	a	b	c	d
3	A	8	-2	5.6
0	A	4	3	-2.1
2	B	1	7	7.4
1	C	-3	6	3.3

D. None of the above.

Problem 7. Write valid code that sorts the rows of the DataFrame second in descending order of the values in column b.

```
second.sort_values('b', ascending=False)
```

Problem 8. What is the result of the code snippet below?

```
second[['c', 'a']]
```

Your answer:

A

A.

	c	a
3	9	B
2	8	C
1	3	B
0	5	A

B.

	a	c
3	B	9
2	C	8
1	B	3
0	A	5

C.

	a	b	c	d
2	C	7	8	8.3
0	A	5	5	4.5

D. None of the above.

Problem 9. What is the result of the code snippet below?

```
first[['b', 'c', 'd']].sum(axis='rows')
```

A.

```
0      4.9
1      6.3
2     15.4
3     11.6
dtype: float64
```

B.

```
b    10.0
c    14.0
d    14.2
dtype: float64
```

C.

```
b    2.50
c    3.50
d    3.55
dtype: float64
```

D. None of the above.

Problem 10. What is the result of the code snippet below?

```
first.assign(
    c_minus_d=lambda x: x['c'] - x['d']
)
```

A.

	a	b	c	d	c_minus_d
0	A	4	3	-2.1	5.1
1	C	-3	6	3.3	2.7
2	B	1	7	7.4	-0.4
3	A	8	-2	5.6	-7.6

B.

```
0     5.1
1     2.7
2    -0.4
3    -7.6
dtype: float64
```

C.

	a	b	c	d	c_minus_d
3	B	-2	9	4.7	4.3
2	C	7	8	8.3	-0.3
1	B	4	3	0.1	2.9
0	A	5	5	4.5	0.5

D. None of the above.