Lab 14: P10 wrapup and Exam Prep

You can use your final lab for two purposes:

1. ask TAs questions about P10
2. practice the following 10 questions to prep for the final (ask TAs if you have any questions)

The answers are here.

Practice Questions

1. what is printed?

```
from bs4 import BeautifulSoup

html = """
<b>hello</b> <i><b>cs</b> 301</i> <b>class</b>.
"""

doc = BeautifulSoup(html, "html.parser")
element = doc.find("i")
b_elements = element.find_all("b")
print(len(b_elements))
```

a. -1
b. 0
c. 1
d. 2
e. 3

2. what is a possible output?

```
import numpy as np
from numpy.random import choice
seed_val = ???? # seed is some valid integer
np.random.seed(seed_val)
x = choice(5)
np.random.seed(seed_val)
y = choice(5)
print(x, y)
```

a. 0 0
b. 1 4
c. 4
d. 5 5
e. 4 4 4

3. which of the following evaluates to "X"?

```
items = ["X", "Y"]
items.append(items)
```

a. items[1]
b. items[-1]
c. items[-1][2][0]
d. items[2][-1][0]
e. both c and d

4. what is a function that contains the yield keyword called?

a. a yielder
b. a generator function
c. a recursive function
d. an iterator

5. what do Python frames contain?

a. functions
b. variables
c. objects
d. pictures of snakes

6. what should replace ???? to get 60 in the result?

Assume the **students** table looks like this:

student	project	score
X	P1	80
Y	P1	70
X	P2	60
Y	P2	50

```
SELECT score
FROM   students
????   student = 'X' AND project = 'P2'
```

a. if
b. IF
c. WHERE
d. HAVING
e. WHEN

7. what should replace ???? to get 75 and 55 in the results?

Assume the same **students** table from the previous question.

```
SELECT AVG(score)
FROM students
????
```

a. GROUP BY student
b. GROUP BY students
c. GROUP BY project
d. GROUP BY student, project
e. GROUP BY score

8. what is printed?

```
letters = Series(["A", "B", "B", "C", "C", "C", "D", "D", "D"])
counts = letters.value_counts().value_counts() # not a typo!
print(counts[3])
```

a. 0
b. 1
c. 2
d. 3
e. 4

9. what values will be in s?

```
df = DataFrame({
  "x": [1,2,3,4],
  "y": [4,3,2,1],
  "z": [5,6,7,8],
})

s = df[(df["y"] > df["x"]) | (df["z"] > 7)]["z"]
```

a. s will be empty
b. 5, 6
c. 8
d. 5, 6, 8
e. 5, 6, 7, 8

10. which of the following is true?

Calling `df.plot.line()` on a DataFrame `df` will:

a. plot one line per row
b. plot one line per column
c. plot one line per cell
d. plot one line, with x values from the `x` column and y values from the `y` column
e. plot one line, with x values from the DataFrame's index and y values from the `y` column