Quiz 2 - DSC 10, Fall 2024

Full Name:

PID:

Quiz Time: ◯ 4PM ◯ 4:30PM ◯ 5PM ◯ 5:30PM ◯ 7PM

Instructions:

- This quiz consists of 6 questions. You have a total of 20 minutes to complete it.
- Please write **clearly** in the provided answer boxes; we will not grade work that appears elsewhere. Completely fill in bubbles and square boxes; if we cannot tell which option(s) you selected, you may lose points.

  ◯ A bubble means that you should only **select one choice**.

  ☐ A square box means you should **select all that apply**.
- If your answer is a string, make sure to put it in quotes. If your answer is a float, make sure to include a decimal point.
- No aids are allowed (no notes, no calculators, and no computers).

By signing below, you are agreeing that you will behave honestly and fairly during and after this quiz.

Signature:

Version A

Please do not open your quiz until instructed to do so.

The DataFrame `uc` has information about the number of degrees awarded by each University of California campus, for each of the past six academic years (`"2018-2019"` through `"2023-2024"`). Since there are ten UC campuses and six years of data, there are 60 rows. The columns are `"Campus" (str)`, `"Year" (str)`, and `"Degrees" (int)`.

The first three rows of `uc` are shown here.

	Campus	Year	Degrees
0	San Diego	2018-2019	10285
1	San Diego	2019-2020	10914
2	San Diego	2020-2021	11453

Question 1

Select the correct way to fill in the blank such that the code below evaluates to a DataFrame with a single row for each UC campus, where the `"Degrees"` column contains the average number of degrees awarded by that campus per academic year (over the past six years).

```
uc.get(["Campus", "Degrees"]).groupby(______).mean()
```

◯ `"Year"` ◯ `"Campus"` ◯ `"Degrees"`

◯ `["Campus", "Year"]` ◯ `["Campus", "Degrees"]` ◯ `["Year", "Degrees"]`

Question 2

Fill in the blank so that the expression below evaluates to the number of years (out of the past six years) in which UC San Diego awarded more than 12,000 degrees.

```
uc[______].shape[0]
```

Question 3

Suppose we plot a density histogram showing the distribution of the `"Degrees"` column in `uc`, using the argument `bins=np.arange(0, uc.get("Degrees").max() + 3000, 2000)` .

Suppose there are 12 rows of `uc` corresponding to campuses and years in which fewer than 2000 degrees were awarded. How tall will the first bar in the histogram be? Give your answer as an **exact decimal or simplified fraction**.

PID: _____

Question 4

a) Typically, graduation ceremonies happen only at the end of an academic year. For example, students who earn a degree in the 2024-2025 academic year celebrate with a graduation ceremony in 2025.

The function `ceremony_year` takes as input a string formatted like those in the `"Year"` column of `uc`, and **returns an `int` corresponding to the year of the graduation ceremony** for that academic year. For example, `ceremony_year("2024-2025")` should return `2025`. Fill in the return statement of this function below.

```
def ceremony_year(academic_year):
    return int(academic_year.split("-")[1])
```

b) What does the following expression evaluate to? Write your answer exactly how the output would appear in Python.

```
uc.get("Year").apply(ceremony_year).min()
```

Question 5

The DataFrame `mascots` has 10 rows, one for each UC school, and only two columns, `"school"` and `"mascot"`. Assuming that the entries in the `"school"` column are formatted exactly how they appear in the `"Campus"` column of `uc`, how many rows does the following merged DataFrame have?

```
uc.merge(mascots, left_on="Campus", right_on="school")
```

◯ `10` ◯ `60` ◯ `70` ◯ `600` ◯ None of these.

Question 6

Define the function `puzzle` as below.

```
def puzzle(word):
    if not(len(word) > 2):
        return False
    elif word in "UC San Diego":
        return True
    else:
        return not("UC" in word)
```

Select all the function calls that evaluate to `True`.

- ☐ `puzzle("UC San")`
- ☐ `puzzle("UC")`
- ☐ `puzzle("UCB")`
- ☐ `puzzle("HDSI")`

Reminders:

- **Write your PID** on the front page and on the top right corner of page 3.
- If your answer is a string, make sure to put it in **quotes**.
- If your answer is a float, make sure to include a **decimal point**.
- Stay in your seat until the quiz is over. No leaving early.