Quiz 3 - DSC 10, Spring 2026

Full Name:

PID:

Quiz Time: ◯ 3PM ◯ 3:30PM ◯ 4PM ◯ 4:30PM

Instructions:

- This quiz consists of 3 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.
- You may use one handwritten sheet of notes. No calculators and no computers.
- Assume we have already run `import babypandas as bpd` and `import numpy as np`.

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.

Regal jumping spiders (*Phidippus regius*) are gaining popularity as pets, and are known for their ability to jump many times more than their body length. Sofia's sister has a pet jumping spider named Elise, whose jumps we will study in this quiz.

	distance	time	humidity
0	142	morning	62
1	68	night	57
2	139	afternoon	60
3	120	morning	72
4	138	afternoon	50

The DataFrame `elise` contains 400 rows (i.e. recorded jumps) and has the following columns:

- `"distance"` (int): the horizontal distance, in millimeters, of Elise's jump
- `"time"` (str): the time of day during which the jump occurred—either `"morning"`, `"afternoon"`, or `"night"`
- `"humidity"` (int): the percent humidity of Elise's enclosure, at the time of the jump

Question 1

Jumping spiders prefer humidity to be between 50 and 70 percent. Suppose the `"humidity"` column has mean 60, and the range $[48, 72]$ is the smallest range guaranteed to contain at least 75% of the values in `"humidity"`. What is the standard deviation of `"humidity"`?

Question 2

a) Elise tends to be most active during the afternoon. Fill in the blanks below to produce 10,000 bootstrapped estimates of Elise's mean jumping distance during the afternoon. Store these estimates in the array `aft_means`.

```
aft = elise[__(a)__]
aft_means = np.array([])

for i in np.arange(10000):
    resample = __(b)__
    estimate = resample.get("distance").mean()
    aft_means = __(c)__
```

(a):

(b):

(c):

b) Write code to compute the endpoints of a 70% bootstrapped confidence interval for Elise's mean afternoon jumping distance, using the values in `aft_means`.

```
left = __(d)__
right = __(e)__
```

(d):

(e):

Question 3

A jump is considered a long jump if it is at least 100 millimeters in length. Sofia creates a new DataFrame called `long`, containing only the rows in `elise` that correspond to long jumps. Suppose the distribution of the `"distance"` column in `long` is approximately normal, with mean 144 and variance 36.

a) Fill in the blank to add a new column to `long` called `"distance_su"`, containing the same information as the `"distance"` column but converted to standard units.

```
standard_units = __(a)__
long = long.assign(distance_su = standard_units)
```

(a):

b) Is `"distance_su"` approximately normally distributed? ◯ Yes ◯ No

c) Assume we have run `from scipy import stats`. Using `stats.norm.cdf`, write a Python expression that evaluates to the approximate proportion of distances in `long` that are at least 140 millimeters.

d) The data in `long` can be thought of a sample from the population of all long jumps that Elise takes. Suppose we use the data in `long` to create a 95% CLT-based confidence interval for the mean distance of all long jumps by Elise (a population parameter). The resulting interval is wider than we'd like, so we decide to collect a new sample that is just large enough to guarantee that the width of the resulting 95% confidence interval is no more than $\frac{3}{2}$, assuming the new sample has the same standard deviation as the original. How large of a sample should we take? Give your answer as an integer.

Reminders:

- **Write your PID** on the front page and next to the PID line in the header on pages where it appears.
- 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.