CSPB 3022 - Craven - Introduction to Data Science Algorithms

Dashboard / My courses / 2241:CSPB 3022 / 29 January - 4 February / Math and Pandas Quiz

Started on	Friday, 2 February 2024, 10:00 PM
State	Finished
Completed on	Friday, 2 February 2024, 10:04 PM
Time taken	4 mins 22 secs
Grade	Not yet graded

Question **1**

Complete

Marked out of 6.00

Suppose you have a **biased** coin, where the **probability of heads** is p.

(a) (3 pts) You flip the coin 8 times. What is the probability that you get the sequence HTTHTTTT

(in that exact order)?

(b) (3 pts) You flip the coin 8 times. What is the probability that you get exactly three heads?

Note: If you use permutations or combinations, either calculate them fully or leave them in terms of factorials (i.e. do not leave notation such as C(10, 6) or P(3, 1) in your answer.

Part a:

We are told in the problem statement that obtaining heads is p. This means the probability of obtaining tails must be $1-p$. Using this we can then say the probability of getting the sequence HTTHTTTT is going to be

$$P = p(1-p)(1-p)p(1-p)(1-p)(1-p)(1-p) = p^2(1-p)^6.$$

Part b:

Following the logic that was found in part a for the probabilities of obtaining either heads or tails, here we are not concerned with the order in which we get three heads. So we would use combinatorics to solve this problem, precisely the binomial theorem. Namely,

$$P(x=k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}.$$

In this formula:

- n = The number of trials (in our case, the number of coin flips)
- k = The number of successes (in our case, the number of heads)
- p = The probability of success for a trial (in our case, the probability of flipping heads).

Filling in the probability for obtaining heads we would then say the probability of this happening is

$$P(x=3) = \binom{8}{3} \cdot p^3 \cdot (1-p)^{8-3} = \frac{8!}{3!(8-3)!} \cdot p^3 \cdot (1-p)^5 = \frac{8!}{3!(5)!} \cdot p^3 \cdot (1-p)^5 = 8 \cdot 7 \cdot p^3 \cdot (1-p)^5 = 56 \cdot p^3 \cdot (1-p)^5.$$

Finally, we have

$$P(x=3) = 56 \cdot p^3 \cdot (1-p)^5.$$

Question **2**

Correct

Mark 5.00 out of 5.00

Let $x_1, x_2, \ldots, x_n$ be a fixed list of numbers and let $\bar{x}$ be the mean of those numbers.

Which of the following statements are always true? (For full credit, you must select all that apply.)

Select one or more:

□ a.
$\sum_{i=1}^{n} 2 = 2\bar{x}$

□ b.
The function $f(c) = \frac{1}{n} \sum_{i=1}^{n} (x_i - c)^2$ is minimized when $c = 0$

☑ c.
The function $f(c) = \frac{1}{n} \sum_{i=1}^{n} (x_i - c)^2$ is minimized when $c = \bar{x}$

✓

□ d.
$\sum_{i=1}^{n} (x_i - \bar{x}) = n\bar{x}$

☑ e.
$\sum_{i=1}^{n} x_i = n\bar{x}$

✓

☑ f.
$\sum_{i=1}^{n} (x_i - \bar{x}) = 0$

✓

Your answer is correct.

Question **3**

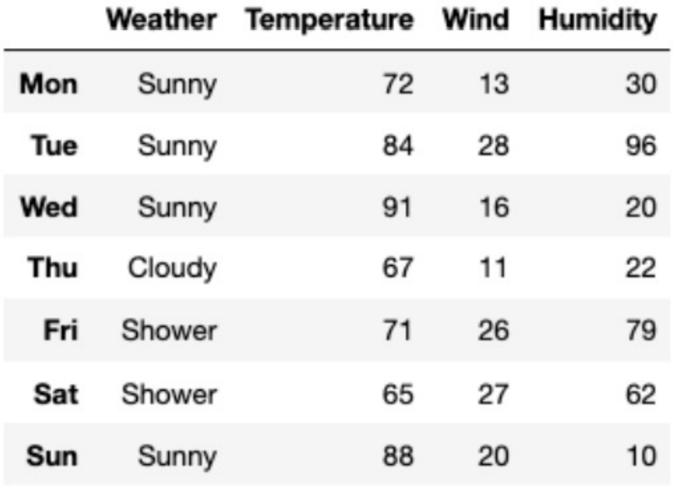
Correct

Mark 5.00 out of 5.00

The following dataframe contains data about weather for a one week period:

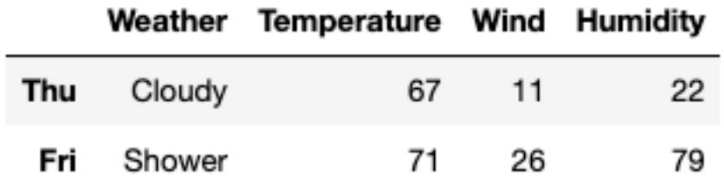
```
import pandas as pd

df=pd.read_csv("data/weather.csv", index_col=["Day"])
```



	Weather	Temperature	Wind	Humidity
Mon	Sunny	72	13	30
Tue	Sunny	84	28	96
Wed	Sunny	91	16	20
Thu	Cloudy	67	11	22
Fri	Shower	71	26	79
Sat	Shower	65	27	62
Sun	Sunny	88	20	10

Which code will return the output shown here?

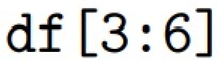


	Weather	Temperature	Wind	Humidity
Thu	Cloudy	67	11	22
Fri	Shower	71	26	79

Select all that apply.

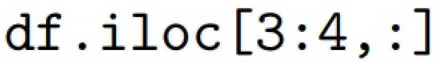
Select one or more:

□ a.



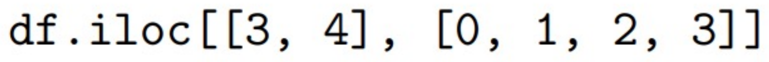
```
df[3:6]
```

□ b.



```
df.iloc[3:4,:]
```

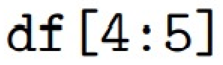
☑ c.



```
df.iloc[[3, 4], [0, 1, 2, 3]]
```

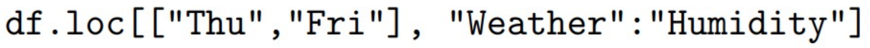
✓

□ d.



```
df[4:5]
```

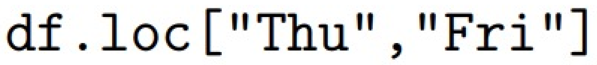
☑ e.



```
df.loc[["Thu","Fri"], "Weather":"Humidity"]
```

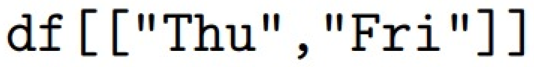
✓

□ f.



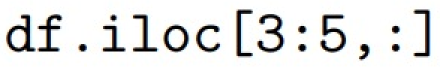
```
df.loc["Thu","Fri"]
```

□ g.



```
df[["Thu","Fri"]]
```

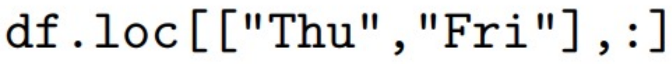
☑ h.



```
df.iloc[3:5,:]
```

✓

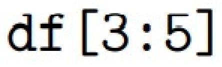
☑ i.



```
df.loc[["Thu","Fri"],:]
```

✓

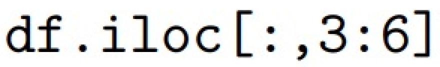
☑ j.



```
df[3:5]
```

✓

□ k.



```
df.iloc[:,3:6]
```

Your answer is correct.

Question **4**

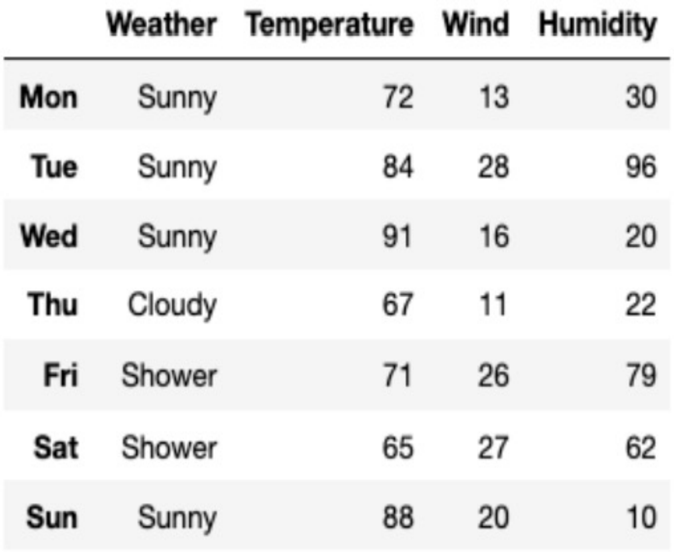
Correct

Mark 4.00 out of 4.00

The following dataframe contains data about weather for a one week period:

```
import pandas as pd

df=pd.read_csv("data/weather.csv", index_col=["Day"])
```



	Weather	Temperature	Wind	Humidity
Mon	Sunny	72	13	30
Tue	Sunny	84	28	96
Wed	Sunny	91	16	20
Thu	Cloudy	67	11	22
Fri	Shower	71	26	79
Sat	Shower	65	27	62
Sun	Sunny	88	20	10

Which code will select only the rows when it was a sunny day and the temperature was above 85? (Assume we still want to return all columns). Select all that apply.

Select one or more:

□ a.

```
df.iloc[(df["Temperature"]> 85) & (df["Weather"] == "Sunny")]
```

☑ b.

```
df[(df["Temperature"]> 85) & (df["Weather"] == "Sunny")]
```

✓

□ c.

```
df[("Temperature"> 85) & ("Weather" == "Sunny")]
```

□ d.

```
df.loc[("Temperature"> 85) & ("Weather" == "Sunny")]
```

□ e.

```
df.loc[(df["Temperature"]> 85) |  (df["Weather"] == "Sunny")]
```

☑ f.

```
df.loc[(df["Temperature"]> 85) & (df["Weather"] == "Sunny")]
```

✓

Your answer is correct.