edX

Course > Week 4... > Week 4... > Week 4...

Week 4 Quiz
Checkboxes

1/1 point (graded)
What are the 2 main data structures in pandas?

☑ Series ✔

☑ Data Frame ✔

☐ Atomic Vector

☐ List

☐ Array

✔

Answer
Correct:
Video: "Why pandas"
Video: "Why pandas"
Video: "Why pandas"
Video: "Why pandas"
Video: "Why pandas"

Submit

ⓘ   Answers are displayed within the problem

Multiple Choice

1/1 point (graded)

Which is NOT a way you can access data at a certain location using pandas series? As an example, let's say the series (called ser) has indices 'apple' and 'orange'.

○   ser.loc['apple']

◉   ser.loc['apple','orange'] ✔

○   ser.loc[['apple']]

○   ser.loc[['apple','orange']]

Answer

Correct: Video: "Why pandas"

[Submit]

ⓘ   Answers are displayed within the problem
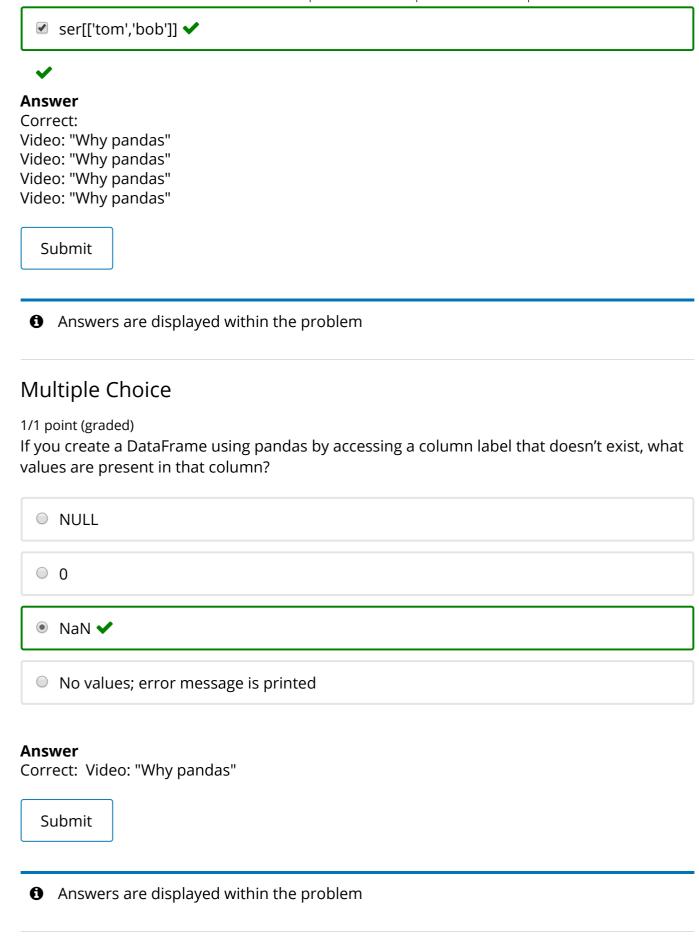

Checkboxes

1/1 point (graded)

Given the following code has run successfully,

```
import pandas as pd
ser = pd.Series([100, 200, 300, 400, 500], index = ['tom', 'bob', 'nancy', 'dan', 'eric'])
```

Which of the following calls have the same output?

☑   ser.loc[['tom','bob']] ✔

☑   ser[[0,1]] ✔

☑   ser.iloc[[0,1]] ✔

☑ ser[['tom','bob']] ✔

✔

Answer
Correct:
Video: "Why pandas"
Video: "Why pandas"
Video: "Why pandas"
Video: "Why pandas"

Submit

ℹ Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
If you create a DataFrame using pandas by accessing a column label that doesn't exist, what values are present in that column?

○ NULL

○ 0

⦿ NaN ✔

○ No values; error message is printed

Answer
Correct: Video: "Why pandas"

Submit

ℹ Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
What is the difference between deleting a column and popping a column?

○ You can store a deleted column.

◉ You can store a popped column. ✔

○ There is no difference.

Answer
Correct: Video: "Why pandas"

Submit

ℹ   Answers are displayed within the problem

True or False

1/1 point (graded)
When you import a dataset using the read_csv function in pandas, the rows of the dataset
are Series.

◉ True ✔

○ False

Answer
Correct: Video: "Data Ingestion"

Submit

ℹ   Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
What is the default number of rows that the function head() will return for a dataframe in pandas?

- ⦿ 5 ✔
- ○ 1
- ○ 10
- ○ 2

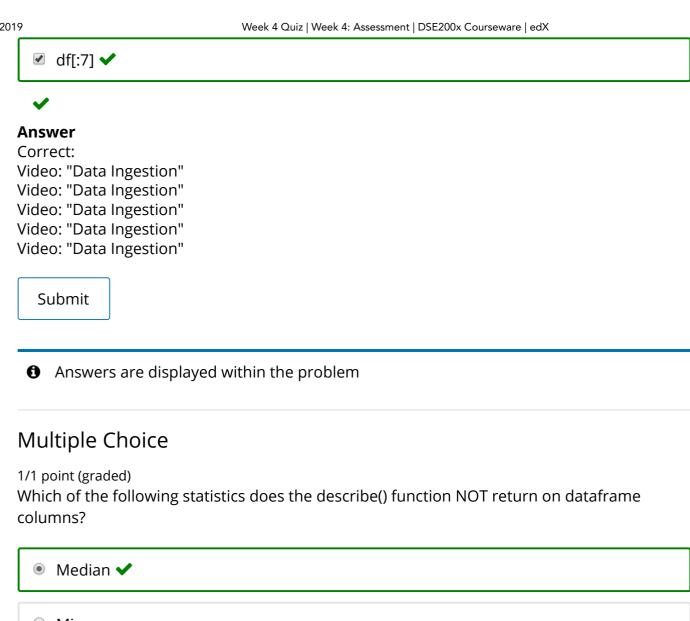**Answer**
Correct: Video: "Data Ingestion"

Submit

ⓘ   Answers are displayed within the problem

Checkboxes

1/1 point (graded)
Which of the following will display the first 7 rows in a data frame object named df?

- ☑ df.head(n=7) ✔
- ☑ df.head(7) ✔
- ☐ df.head([7])
- ☐ df.head(x=7)

☑ df[:7] ✔

✔

Answer
Correct:
Video: "Data Ingestion"
Video: "Data Ingestion"
Video: "Data Ingestion"
Video: "Data Ingestion"
Video: "Data Ingestion"

Submit

ℹ Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
Which of the following statistics does the describe() function NOT return on dataframe
columns?

◉ Median ✔

◯ Min

◯ Mean

◯ Count

Answer
Correct: Video: "Descriptive Statistics"

Submit

ℹ Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
What is the general syntax for calling the mode() function on a dataframe?

- ⦿ data_frame.mode() ✔

- ○ mode(data_frame)

- ○ data_frame[mode]

Answer
Correct: selected: Video: "Descriptive Statistics"},{unselected: Video: "Descriptive Statistics"

Submit

ⓘ   Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
What will be the output of the following lines of code?

```
d = {'one' : pd.Series([100.,200.], index=['apple','orange']),
     'two' : pd.Series([111.,211], index=['apple','orange'])}
df = pd.DataFrame(d)
filter = df['one']>100
filter.any()
```

- ⦿ True ✔

- ○ False

Answer
Correct: Video: "Descriptive Statistics"

Submit

ℹ Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
What does a negative correlation score mean?

⦿ those features in our dataset are inversely correlated ✔

○ all the values of those features in our dataset are negative

○ there is no correlation between those features in our data set

Answer
Correct: Video: "Descriptive Statistics"

Submit

ℹ Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
Which function allows you to clean a data set by dropping NA values?

⦿ dropna() ✔

○ na(FALSE)

○ na()

○ drop(na)

Answer
Correct: Video: "Data Cleaning"

Submit

ℹ Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
If you have a dataframe titled 'dat' with 5 rows and 2 columns and you run the following line of code, how many boolean values are returned?

```
dat.isnull().any()
```

⦿ 2 ✔

○ 5

○ 10

○ 1

Answer
Correct: Video: "Data Cleaning"

Submit

ℹ Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
What is the symbol that denotes Magic Functions in Jupyter?

- ⦿ % ✔
- ○ $
- ○ #
- ○ //

Answer
Correct: Video: "Data Cleaning"

[Submit]

ℹ️   Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
How do we show the histogram created with graph.hist(...)?

- ○ It will automatically show once you call the hist() function
- ○ Call "graph.show()"
- ○ Call "hist.show()"
- ⦿ Include "%matplotlib inline" ✔

Answer
Correct: Video: "Data Cleaning"

Submit

ℹ Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
Suppose you are looking at a data set consisting of how much students liked a particular class. The rows are students. There are two columns: one asks if the student has taken the class and the other asks to rate the class on a scale of 1-10. However, you notice that some students have not taken the class and therefore do not want to include their ratings. Which data filtering technique would you use to clean this dataset?

○ Slice out columns

◉ Filter out rows ✔

○ Transform data

Answer
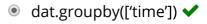Correct: Video: "Frequent Data Operations"

Submit

ℹ Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
Which function call will allow you to group a dataset (titled dat) by 'time'?

◉ dat.groupby(['time']) ✔

○ groupby(dat['time'])

○   dat.aggregate(['time'])

○   dat['time'].group()

Answer
Correct: Video: "Frequent Data Operations"

Submit

ⓘ   Answers are displayed within the problem

True or False

1/1 point (graded)
When you concatenate two dataframes using pandas concat function, the number of resulting columns will be the columns that BOTH dataframes have. As an example, if one dataframe has columns titled 'cat','dog' and another dataframe has columns titled 'dog','bunny', then the resulting dataframe columns will be 'dog'.

○   True

⦿   False ✔

Answer
Correct: Video: "Merging DataFrames"

Submit

ⓘ   Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
For the merge() function in pandas, how does the parameter "how" handle row indices when how='inner'?

~~now~~ ~~inner~~?

○ The function takes the union of the row indices

◉ The function takes the intersection of the row indices ✔

○ The function takes the complement of the row indices

Answer
Correct: Video: "Merging DataFrames"

Submit

ⓘ Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
What does the parameter expand represent in the split function?

◉ determines if the output is a dataframe ✔

○ creates every split string into a new column

○ pipes the split strings into the output

Answer
Correct: Video: "Frequent String Operations"

Submit

ⓘ Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
Given the following:

```
Regular expression: '.*\,(.*)\&.*'
String in Series: "Feuer, Eis & Dosenbier"
```

What would be the output of the extract function with this string and regular expression?

○ ,

○ &

○ Feuer

◉ Eis ✔

○ Dosenbier

Answer
Correct: Video: "Frequent String Operations"

Submit

ℹ   Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
What is an example of Unix time (int64)?

○ 2006-01-29 12:29:30

◉ 1138537770 ✔

○ 12:29 PM

○ Jan. 29, 2006

Answer
Correct: Video: "Parsing Timestamps"

Submit

ⓘ   Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
What does the parameter unit refer to in the to_datetime function?

○ The unit of the output

◉ The unit of the input ✔

Answer
Correct: Video: "Parsing Timestamps"

Submit

ⓘ   Answers are displayed within the problem

Multiple Choice

1/1 point (graded)

What is the function call to find cells in a dataframe df with timestamp on 2007-02-04, given the dataframe has a parsed time column labelled 'parsed_time'?

○ df['parsed_time'] == '2007-02-04' ✔

○ df['parsed_time'] = '2007-02-04'

○ '2007-02-04' >= df['parsed_time'] >= '2007-02-04'

○ '2007-02-05' > df['parsed_time'] > '2007-02-03'

Answer
Correct: Video: "Parsing Timestamps"

Submit

ⓘ   Answers are displayed within the problem

Multiple Choice

1/1 point (graded)
What does calling the function dropna() return from a dataframe?

○ dataframe with the rows containing the NA values

○ dataframe omitting NA values ✔

○ dataframe with the columns containing the NA values

○ dataframe with changed NA values by replacing NA values with 0

Answer
Correct: Video: "Summary of Movie Rating Notebook"

Submit

 🛈   Answers are displayed within the problem

Checkboxes

1/1 point (graded)
What are the 3 main string operations?

☑ split ✔

☑ contains ✔

☑ extract ✔

☐ delete

✔

Answer
Correct:
Video: "Summary of Movie Rating Notebook"
Video: "Summary of Movie Rating Notebook"
Video: "Summary of Movie Rating Notebook"
Video: "Summary of Movie Rating Notebook"

Submit

 🛈   Answers are displayed within the problem