

1.

What does the following filter line of code do?

1 / 1 point

df.filter(df["teamlevel"] > 1)

- ☐ Select the first two columns of the data and displays only team levels greater than 1.
- ☐ Filter each column to show only team levels larger than 1.
- ☒ Filter each row to show only team levels larger than 1.
- ☐ Select the first two columns of the data and filter each column to show only team levels larger than 1.

✔ Correct

2.

What does the following do?

1 / 1 point

df.select("userid", "teamlevel").show(5)

- ☐ Select the rows named "userid" and "teamlevel" and display first 5 rows.
- ☐ Display all columns except "userid" and "teamlevel".
- ☐ Display all rows except "userid" and "teamlevel".
- ☒ Select the columns named "userid" and "teamlevel" and display first 5 rows.

✔ Correct

3.

What does the following code do?

1 / 1 point

window("time", "10 seconds")

- ☐ Delays timestamps by 10 seconds
- ☐ Replaces the column "time" with a 10 second window
- ☒ Creates a 10 second window, using column "time" timestamps.
- ☐ Accelerates timestamps by 10 seconds

✔ Correct

4.

query = (

1 / 1 point

```
    windowed_data
    .writeStream
    .outputMode("update")
    .trigger(processingTime="10 seconds")
    .foreachBatch(lambda batch_df, epoch_id: batch_df.show(truncate=False))
    .start()
)
```

In the previous code, Which method initiates the retrieval of streaming data?

- ☐ trigger()
- ☐ outputMode()
- ☐ foreachBatch()
- ☒ start()

✔ Correct