

# Spark Streaming

1. What will happen if you run 'country\_counts.show()'? Why?

1 point

*Mark only one oval.*

- ☐ The contents of the dataframe will be printed as usual and contain whatever data already received.
- ☐ We get an error complaining that the query must be executed with start()
- ☐ it depends if we called show() before or after the call to start()

2. change OFFSETS\_PER\_TRIGGER to 100. How does it affect the processing?

1 point

*Mark only one oval.*

- ☐ The only effect is that there will be more calls to process the batches
- ☐ an error due to "batch too small"
- ☐ the worker node may crash because out of memory

3. replace outputMode("complete") with "append".  
Is it working? why?

1 point

*Mark only one oval.*

- ☐ It works as expected. Spark knows how to handle the difference from last call.
- ☐ There is an error because if there is no watermark, Spark cannot know when it is safe to append.
- ☐ There is an error because the specific aggregation we used in the code is not supported for this mode.

4. When using **memory** sink, (the memory of) which node is used?

1 point

*Mark only one oval.*

- ☐ The memory of the worker node where the executor is currently running
- ☐ worker nodes share their memory, so several nodes are used
- ☐ the driver node

5. what happens if you subscribe to nonexistent topic in Kafka?

1 point

*Mark only one oval.*

- ☐ An error during the creation of the readstream
- ☐ It's ok to subscribe but we will not get any data
- ☐ An error when applying a transformation of a dataframe read from the stream

6. What happens if the microbatch processing time is longer than the input data rate?

1 point

*Mark only one oval.*

- ☐ Spark will raise an error or let the user retry, depending on configuration
- ☐ the space used for this topic and consumer in the kafka server will grow
- ☐ processing time cannot be longer than input data rate because data rate is always a fraction of the processing rate.

---

This content is neither created nor endorsed by Google.

Google Forms

