

More on Spark

[继续课程](#)

9/9 得分 (100%)

[返回第 #5 周课程](#)

测验通过！



1 / 1 分

1。

Which part of SPARK is in charge of creating RDDs?



Driver Program



正确



Spark Executor



Worker Node



Storage



Local CPU



1 / 1 分

2。

How does lazy evaluation work in Spark?



Transformations are queued and executed at a certain threshold.



Actions are not executed until the transformation stage.



Transformations are not executed until the action stage.



正确

- ☐ Actions are queued and executed at a certain threshold.
-



1 / 1 分

3。

What are the consequences of lazy evaluation as mentioned in lecture?

- ☐ Hiccups within the system during queue execution.
- ☐ There are no consequences.
- ☒ Errors sometimes do not show up until the action stage.



正确



1 / 1 分

4。

What is a wide transformation?

- ☐ Transformations that take a lot of nodes to complete.
- ☐ The name for the most used transformations.
- ☐ A longer time-taking transformation compared to narrow transformations.
- ☒ A transformation that requires data shuffling across node partitions.



正确



1 / 1 分

5.

Where does the data for each worker node get sent to after a collect function is called?

- ☐ Spark SQL
- ☐ None; Stays in the Same Node
- ☐ Other Worker Nodes
- ☒ Spark Context



正确

- ☐ Spark Streaming



1 / 1 分

6.

What are DataFrames?

- ☒ A column like data format that can be read by Spark SQL.



正确

- ☐ A special type of data node that contains framework to manipulate SQL.
- ☐ A type of narrow transformation.



1 / 1 分

7.

Can RDD's be converted into DataFrames directly without manipulation?



- ☐ No: RDD's cannot be converted into DataFrames.
- ☐ Yes
- ☐ No: RDD's needed to be made relational first.
- ☒ No: lines have to be converted into row.

正确



1 / 1 分

8.

What is the function of Spark SQL as mentioned in lecture? (Choose 3)

☐

Better ability to manipulate big data.

未选择的是正确的

☐

Efficient data manipulation using SQL like structure.

未选择的是正确的

☒

Enables relational queries on Spark.

正确

☒

Connect to variety of databases.

正确

☐

Better worker node interpolation.

未选择的是正确的



Deploy business intelligence tools over Spark.



正确



1 / 1 分

9。

What is a triplet in GraphX?



A type of data to contain info on connections between vertices and edges.



正确



A type of data to contain both edge and vertex info.



A type of data to contain edge info.



A type of data to contain vertex info.

