



Congratulations! You passed!

TO PASS 80% or higher

Keep Learning

GRADE
100%

More on Spark

LATEST SUBMISSION GRADE

100%

1. Which part of SPARK is in charge of creating RDDs?

1 / 1 point

- ☐ Worker Node
- ☒ Driver Program
- ☐ Local CPU
- ☐ Spark Executor
- ☐ Storage



Correct

2. How does lazy evaluation work in Spark?

1 / 1 point

- ☐ Actions are not executed until the transformation stage.
- ☒ Transformations are not executed until the action stage.
- ☐ Actions are queued and executed at a certain threshold.
- ☐ Transformations are queued and executed at a certain threshold.



Correct

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

1 / 1 point

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



Correct

4. What is a wide transformation?

1 / 1 point

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



Correct

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

1 / 1 point

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



Correct

6. What are DataFrames?

1 / 1 point

- ☐ A special type of data node that contains framework to manipulate SQL.
- ☐ A type of narrow transformation.
- ☒ A column like data format that can be read by Spark SQL.

✓ Correct

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

1 / 1 point

☐

No: RDD's needed to be made relational first.

☐

No: RDD's cannot be converted into DataFrames.

☐

Yes

☒

No: lines have to be converted into row.

✓ Correct

8.

1 / 1 point

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

☐

Better ability to manipulate big data.

☒

Deploy business intelligence tools over Spark.

✓ Correct

☒

Enables relational queries on Spark.

✓ Correct

☐

Efficient data manipulation using SQL like structure.

☐

Better worker node interpolation.

☒

Connect to variety of databases.

✓ Correct

9.

1 / 1 point

What is a triplet in GraphX?

☐

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.

☒

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

✓ Correct