

✔ Congratulations! You passed!

Go to next item

Grade received 100% Latest Submission Grade 100% To pass 80% or higher

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

1 / 1 point

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

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

✔ Correct

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

1 / 1 point

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

✔ Correct

4. What is a wide transformation?

1 / 1 point

- ☐ The name for the most used transformations.
- ☐ Transformations that take a lot of nodes to complete.
- ☐ A longer time-taking transformation compared to narrow 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

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

✔ Correct

6. What are DataFrames?

1 / 1 point

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

✓ Correct

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

1 / 1 point



No: lines have to be converted into row.



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



Yes



No: RDD's cannot be converted into DataFrames.

✓ Correct

8.

1 / 1 point

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



Enables relational queries on Spark.

✓ Correct



Better ability to manipulate big data.



Connect to variety of databases.

✓ Correct



Deploy business intelligence tools over Spark.

✓ Correct



Efficient data manipulation using SQL like structure.



Better worker node interpolation.

9.

1 / 1 point

What is a triplet in GraphX?



A type of data to contain edge info.



A type of data to contain vertex info.



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





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



Correct