 ✓ Congratulations! You passed! Grade Latest Submission To pass 80% or 	next item
Grade Latest Submission To pass 80% or received 100% Grade 100% higher 1. Which of the following Apache Spark benefits helps manage big data processing?	1/1 point
In-memory processing Python Scala Unified framework	
∠ [¬] Expand	
 Correct Correct Correct. Apache Spark creates a comprehensive, unified framework to manage big data processing. The three Apache Spark components are data storage, compute interface, and cluster management framework. Which order does the data flow through these components? 	1 / 1 point
Which order does the data flow through these components? Data flows from API into different nodes for parallel tasks, and then into a Hadoop file system. Data flows from Hadoop file system, into compute interface and then into different nodes for distributed. Data flows from compute interface to various nodes for distributed tasks and then goes to the Hadoop file system.	
Data flows from a Hadoop file system into different nodes for distributed task, but then flows to the APIs.	
Correct Correct. the data from a Hadoop file system flows into the compute interface or API, which then flows into different nodes to perform distributed/parallel tasks.	
3. Which of the following is true of datasets? APIs are available in Scala and Java as well as other languages.	1 / 1 point
 Datasets are strongly typed and therefore provide compile-time type safety. DataFrames are built on top of datasets. Datasets compute more slowly than RDDs. 	
 ✓ Correct Correct. Compile-time type safety means that Spark can detect syntax and semantic errors in production 	
applications before deployment. 4. Which of these is one of the four phases of Catalyst query optimization?	1 / 1 point
Logical planningPhysical optimizationCode analysisAnalysis	
∠ [™] Expand	
Correct Correct. The other three stages are logical optimization, physical planning, and code generation. 5. How does IBM Spectrum Conductor help avoid downtime when running Spark?	1 / 1 point
 Cluster resources divided dynamically Shares cluster resources Deploy multiple versions 	
Automatic troubleshooting Expand	
Correct Correct. this avoids downtime.	
O SUI What is the name of the Spark unified interface? SIII	1/1 point
SUI	
Correct Correct. The spark-submit script is found in the 'bin/' directory.	
 7. Which command specifies the number of executor cores for a Spark standalone cluster for the application? Use the command 'total-executor-cores' followed by the number of cores. Use the command 'appexecutor-cores' followed by the number of cores. 	1 / 1 point
Use the command '—apptotalexecutor-cores' followed by the number of cores. Use the command '—apptotal-executor-cores' followed by the number of cores	
8. If a task fails due to a dependency problem, what is the best way to identify the issue? © Examining the event log for stack trace errors Checking APIs	1 / 1 point
Checking required data files for corruption Cataloging the libraries on the system	
Correct Correct. These identify which libraries the application loaded.	
 Select the answers that describe the relationship between Big Data and today's personal assistants including Google, Alexa Siri and others. Personal assistants also rely on unstructured data sources including personal data in the form of photos, videos, and text that people send to each other as the bulk of data collected by consumer goods companies. 	1 / 1 point
 Correct Yes! Personal assistants use unstructured data sources including personal data in the form of photos, videos, and texts that people send to each other as the bulk of data collected by consumer goods companies. 	
 □ Assistants base their answers solely on structured data sources. □ Personal assistants use data sources including location tracking, and historical shopping data to help provide predictive answers based on personal preferences. ✓ Correct Yes! Assistants combine data from a multitude of sources, apply algorithms, and AI to 	
provide users with what the user will deem to be a correct answer. Assistants take questions and provide answers via some of the most advanced neural networks that exist. Correct Yes! Advanced neural networks process the user's words and even voice tone when	
creating responses to questions and requests. Expand	
Correct Great, you got all the right answers. 10. Which of the following is true of Big Data?	1/1 point
 It is complex and requires specialized software to interpret and make it available for human interpretation. It can be stored on private servers. It refers to just large volumes of data. 	
○ It is only generated by certain specialized sensors and devices. ∠ Expand	
Correct. Big Data arrives at a massive volume and with little or no structure. 11. What is Data Scaling?	1/1 point
 Data scaling divides workloads to run in parallel. Data scaling is a technique to manage, store, and process the overflow of data. Data scaling is only applicable within cloud environments. 	
O Data scaling is the process of transforming data values for end use. ∠ Expand	
Yes! This answer is correct! 12. Semi-structured data	1/1 point
 Includes some metadata that identifies certain characteristics. Includes databases and spreadsheets. Has a pre-structured data model. 	
○ Includes sensor data from Internet of Things devices. ∠ Expand	
Correct Correct. Semi-structured data combines unstructured and structured data. 13. Which of the following Hadoop core components prepares the RAM and CPU for Hadoop to run data in batch,	1/1 point
stream, interactive, and graph processing? HDFS Hadoop Common	
 MapReduce ● YARN Expand	
Correct Correct. Yarn is short for "yet another resource negotiator" and it's one of the most important components because it prepares the RAM and CPU for Hadoop to run data in these types of processing.	
14. Increasing Executors and cores necessitates dividing jobs into tasks transforms data partitions	1 / 1 point
transforms data partitions increases cluster parallelism requires a shuffle	
 ∠ Expand ✓ Correct Correct. Tasks run in separate threads until all cores are used. 	
✓ Correct	1/1 point
 ✓ Correct Correct. Tasks run in separate threads until all cores are used. 15. Which configuration method enables you to adjust settings on a per-machine basis? 	1/1 point
Correct Correct. Tasks run in separate threads until all cores are used. 15. Which configuration method enables you to adjust settings on a per-machine basis? Properties Logging Environment variables Manual Correct	1/1 point
Correct Correct. Tasks run in separate threads until all cores are used. 15. Which configuration method enables you to adjust settings on a per-machine basis? Properties Logging Environment variables Manual Correct Correct. Feedback about why this answer is correct	1/1 point
 ✓ correct Correct. Tasks run in separate threads until all cores are used. 15. Which configuration method enables you to adjust settings on a per-machine basis? Properties Logging Environment variables Manual ✓ correct Correct. Feedback about why this answer is correct 	
Correct Correct. Tasks run in separate threads until all cores are used. 15. Which configuration method enables you to adjust settings on a per-machine basis? Properties Logging Environment variables Manual Correct Correct. Feedback about why this answer is correct 16. Kubernetes runs containerized applications on a cluster. What else is true about it? It cannot be deployed automatically. It only runs in the cloud. It's portable. It cannot be run on a single machine.	
Correct Correct. Tasks run in separate threads until all cores are used. 15. Which configuration method enables you to adjust settings on a per-machine basis? Properties Logging Environment variables Manual Carrect Correct. Feedback about why this answer is correct 16. Kubernetes runs containerized applications on a cluster. What else is true about it? It cannot be deployed automatically. It only runs in the cloud. It is portable. It cannot be run on a single machine. Expand Correct Correct. Kubernetes can be run in the cloud or on-premises.	
 ✓ Correct Correct. Tasks run in separate threads until all cores are used. 15. Which configuration method enables you to adjust settings on a per-machine basis? Properties Logging Environment variables Manual ✓ Correct Correct. Feedback about why this answer is correct 16. Kubernetes runs containerized applications on a cluster. What else is true about it? It cannot be deployed automatically. It only runs in the cloud. It's portable. It cannot be run on a single machine. ✓ Correct Correct Correct. Kubernetes can be run in the cloud or on-premises. 	1/1 point
Correct Correct. Tasks run in separate threads until all cores are used. 15. Which configuration method enables you to adjust settings on a per-machine basis? Properties Logging Environment variables Manual Correct Correct. Feedback about why this answer is correct 16. Kubernetes runs containerized applications on a cluster. What else is true about it? It cannot be deployed automatically. It only runs in the cloud. It represents the run on a single machine. Correct Correct. Kubernetes can be run in the cloud or on-premises. 17. Which of the following is NOT true of open-source software? It can only be changed by a designated organization. It is the industry standard for servers worldwide. It is profitable to use. It works very well for large, complex projects.	1/1 point
Correct. Tasks run in separate threads until all cores are used. 15. Which configuration method enables you to adjust settings on a per-machine basis? Properties Logging Environment variables Manual Correct Correct. Feedback about why this answer is correct 16. Kubernetes runs containerized applications on a cluster. What else is true about it? It cannot be deployed automatically. It only runs in the cloud. It's portable. It cannot be run on a single machine. Correct Correct. Kubernetes can be run in the cloud or on-premises. 17. Which of the following is NOT true of open-source software? It can only be changed by a designated organization. It's the industry standard for servers worldwide. It is profitable to use. It works very well for large, complex projects.	1/1 point
Correct. Tasks run in separate threads until all cores are used. 15. Which configuration method enables you to adjust settings on a per-machine basis? Properties	1/1point 1/1point
Correct. Tasks run in separate threads until all cores are used. 15. Which configuration method enables you to adjust settings on a per-machine basis? Properties Logging Environment variables Manual 16. Kubernetes runs containerized applications on a cluster. What else is true about it? It cannot be deployed automatically. It cannot be deployed automatically. It enry runs in the cloud. It cannot be nun on a single machine. 2 Expand 2 Correct. Correct. Kubernetes can be run in the cloud or on-premises. 17. Which of the following is NOT true of open source software? It can only be changed by a designated organization. It she inclustly sunatural for servers workforkle. It is profitable to use. It works very well for large, complex projects. 18. Which of the following happens in the map task of MapReduce? Prouces data into key value pans Produce a final report Aggregate as est of results Give consistent names to pieces of data	1/1point 1/1point
Correct. Tasks run in separate threads until all cores are used. 15. Which configuration method enables you to adjust settings on a per-machine basis? Properties	1/1point 1/1point
Correct Correct. Takes run in separate threads until all cores are used. 15. Which configuration method enables you to adjust settings on a per-machine basis? Preporties Logaling	1/1 point
Correct. Tables run in separate threads until all cores are used. 13. Which configuration method enables you to adjust settings on a per-machine basis? Proporties Logging	1/1 point
Cerrent Cerrent Tasks run in separate threeds until all cores are used. 15. Which configuration method enables you to adjust settings on a per machine basis? Proporties	1/1 point 1/1 point
Cerrect. Tests run in separate threeds until all cores are used. 12. Which configuration method enables you to adjust settings on a per machine basis? Properties Usgins	1/1 point
Cerrent Cerrent Tables rou in separate threads until all cores are axed. 1.1. Which configuration method enables you to adjust settings on a per-machine basis? Progenes copying	1/1 point 1/1 point