Final Quiz

Quiz • 1h 40m

Submit your assignment

Due Jul 16, 11:59 PM EDT **Attempts** 3 every 24 hours

Receive grade

To Pass 80% or higher

Try again

Your grade
65%

View Feedback
We keep your highest score

Like **Dislike** **Report an issue**

1. What is meant by the term "data transformation"? **1 point**

- (•) Processing data to make it conform to requirements.
- ( ) Configuring access to the data and reading it into an application
- ( ) Making data readily available for ingestion by analytics applications so that end users can gain value from it.
- ( ) Writing data to some new destination environment

Expand

2. Which of the following best describes a use case for ELT? **1 point**

- ( ) When a very large amount of information is either already recorded or being generated, but is not yet captured, or accessible.
- ( ) When moving data is usually more of a bottleneck than processing it.
- ( ) When data is structured and relational rather than unstructured.
- ( ) When data is being live streamed from a single location.

Expand

3. Which of the following best describes a driving factor of the evolution from ETL to ELT? **1 point**

- ( ) The demand to release raw data to a wider user base.
- ( ) The need to decouple transformations from the data pipeline.
- ( ) The need to build a single dashboard from multiple data sources.
- ( ) The need to transfer data from a legacy system to an updated one.

Expand

4. What are two examples of raw data sources? **1 point**

- ( ) Analytics and human genomes data
- ( ) Paper documents and weather station networks.
- ( ) Calculations and web pages
- ( ) Social media and artificial intelligence

Expand

5. What is batch loading? **1 point**

- ( ) Batch loading refers to loading data in real time.
- ( ) Batch loading refers to loading a small window of recent data for an imn
- ( ) Batch loading refers to loading an initial history into a database.
- ( ) Batch loading refers to loading data in chunks defined by time windows

Expand

6. How is shell scripting used to implement an ETL pipeline? **1 point**

- ( ) By using an API
- ( ) By building a basic Bash script
- ( ) By using Python to build the shell
- ( ) By using the crontab editor

Expand

7. What is the definition of a data pipeline? **1 point**

- ( ) Data pipelines are systems that specifically move or modify data.
- ( ) Data pipelines are processes in chains.
- ( ) Data pipelines are any series of connected processes.
- ( ) Data pipelines are units of data queued for ingestion.

Expand

8. Which of these fall into the category of data pipeline monitoring? **1 point**

- ( ) Extraction and ingestion
- ( ) Latency and throughput
- ( ) Loading and scheduling
- ( ) Scheduling and maintenance

Expand

9. Higher quality output is usually a trade-off with what other pipeline feature? **1 point**

- ( ) Real-time streaming
- ( ) Batch processing
- ( ) Latency
- ( ) Throughput

Expand

10. Which of the following is a stream-processing technology? **1 point**

- ( ) Panoply
- ( ) Alteryx
- ( ) Apache Spark
- ( ) Pandas

Expand

11. Which of the following are the four principles Apache Airflow is built upon?" **1 point**

- ( ) Robust, scalable, effective, dynamic
- ( ) Effective, simple to use, scalable, agile
- ( ) Scalable, dynamic, extensible, lean
- ( ) Sustainable, competitive, agile, simple to use

Expand

12. Which statement is true about directed acyclic graphs (DAGs)? **1 point**

- ( ) All DAGs are trees but not all trees are DAGs.
- ( ) Some DAGs have loops
- ( ) All DAGs have loops.
- ( ) All trees are DAGs but not all DAGs are trees.

Expand

13. Which statement is true about the Apache Airflow UI? **1 point**

- ( ) The interface flags graphs that contain loops in red.
- ( ) DAGs are always displayed as trees.
- ( ) Clicking the task duration button allows you to see a graph representing
- ( ) Clicking on the Code button allows you to see the Python source code t

Expand

14. In the Apache Airflow DAG, which code block contains the nodes of the DAG? **1 point**

- ( ) DAG argument specification
- ( ) DAG definition
- ( ) Task pipeline
- ( ) Task definitions

Expand

15. In Airflow, which is an example of a counter metric? **1 point**

- ( ) The number of failed tasks.
- ( ) The latency of a task.
- ( ) The number of currently running tasks.
- ( ) The amount of time it takes to complete a task.

Expand

16. Select the correct statement regarding events. **1 point**

- ( ) Events describe a clearly observable entity, such as a car or a hospital p
- ( ) An event is something that no one notices, but it is happening.
- ( ) Events describe an entity's observable state updates over time.
- ( ) The temperature at which water freezes is an example of an event.

Expand

17. Which statement best describes Kafka brokers? **1 point**

- ( ) A Kafka broker is a communication protocol to exchange data between
- ( ) A Kafka broker is a collection of shell scripts to communicate with Kafku
- ( ) A Kafka broker is a server cluster acting to receive, store, and distribute
- ( ) A Kafka broker is a system that ensures events streams are handled in a

Expand

18. Select two common components of an event streaming platform (ESP). **1 point**

- [ ] Event storage
- [ ] Event analysis
- [ ] Event transportation
- [ ] Event pipeline

Expand

19. What is the function of the ZooKeeper in Kafka? **1 point**

- ( ) To log events
- ( ) To process events
- ( ) To synchronize and manage brokers
- ( ) To manage an event stream

Expand

20. Which of the following explains what an ad hoc data processor does? **1 point**

- ( ) Filter raw data based on a condition
- ( ) Send processed data to a producer
- ( ) Send data to a dashboard
- ( ) Analyze data stored in a Kafka topic.

Expand

Coursera Honor Code Learn more

[ ] I, **Smit Shah**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.

Submit **Save draft**

Like **Dislike** **Report an issue**