Your grade: 100%

Your latest: **100%** • Your highest: **100%** • To pass you need at least 70%. We keep your highest score.

Next item

1. ETL process consists of Extract > Transform > Load. Which of the three processes is also known as data wrangling? 1 / 1 point

- (●) Transform
- ( ) Data wrangling is a term for another data warehouse process.
- ( ) Extraction
- ( ) Load

Correct
Correct! This process wrangles the data into the format suitable for destination and use.

2. The ELT process has no information loss. What is the main reason for this benefit? 1 / 1 point

- ( ) Data source integration
- ( ) It separates the data pipeline from processing.
- ( ) There is a separation between moving and processing data.
- (●) Data is acquired and directly loaded, as-is, into its destination environment.

Correct
Feedback: Correct! ELT provides a replica of the source data, and with that, no information loss occurs.

3. Which of the following in an ELT process best compares to the “Staging area” in the ETL process? 1 / 1 point

- ( ) Database servers
- (●) Data lake in ELT process
- ( ) Storage for source data in the ELT process
- ( ) Transformed data storage

Correct
Correct! The staging area fits the description of a data lake, which is a modern self-serve repository for storing and manipulating raw data.

4. Which of the following pain points does ELT address? 1 / 1 point

- ( ) Request for fixed processes
- ( ) Cost-effectiveness
- (●) Challenges imposed by Big Data
- ( ) Transformation in data pipeline

Correct
Correct! Challenges like scalability imposed by Big Data are addressed.

5. There are many techniques for extracting data. What does the choice of technique depend on? 1 / 1 point

- ( ) Type of client
- ( ) Optical or analog
- ( ) Operating system
- (●) Kind of data source and intended use

Correct
Correct! It depends on the kind of data source and intended use of the data.

6. Extracting data from IoT devices involves large volumes of redundant data. What is used to decrease the data volume of redundant data and only extract features of interest from raw data? 1 / 1 point

- (•) Edge computing
- ( ) APIs
- ( ) Biometric sensors
- ( ) SQL languages

Correct
Correct! Edge computing reduces the data volumes of redundant data by extracting features of interest from the raw data.

7. ETL uses the schema-on-write approach. What is the biggest disadvantage of this approach? 1 / 1 point

- ( ) Stability
- ( ) More data access
- ( ) Consistency
- (•) Limited versatility

Correct
Correct! The ETL approach limits the versatility whereas the ELT approach is versatile since it obtains multiple views of the same source data with ad-hoc schemas.

8. Why is there no information loss in ELT unlike ETL where there is loss of information? 1 / 1 point

- ( ) Because ELT uses edge computing
- ( ) Because ELT involves lossy data compression
- ( ) Because ETL uses aggregation of data but ELT doesn't
- (•) Because the data is copied as is

Correct
Correct! All the original information content is left intact as the data is copied as is.

9. Which of these is most useful for incremental loading strategy? 1 / 1 point

- ( ) Only batch loading
- ( ) File partitioning
- (•) Both batch and stream loading
- ( ) Only stream loading

Correct
Correct! Stream loading is useful for incremental loading and should be real-time. Batch loading is useful when the incremental loading can be in batches.

10. Which of the following loading techniques can split a single file into smaller chunks? 1 / 1 point

- (•) Parallel loading
- ( ) Scheduled loading
- ( ) Stream loading
- ( ) Batch loading

Correct

Correct! Parallel loading technique splits single files into small chunks and loads them simultaneously.