

Congratulations! You passed!

Grade received 90%

Latest Submission Grade 90%

To pass 70% or higher

Retake the assignment in 23h 50m

Go to next item

1. In the modern data ecosystem, what is one of the key challenges when stakeholders pull data from the enterprise data repository?

1 / 1 point

- ☐

Checking that all the raw data resides in a centralized repository
- ☒

Ensuring that the method of data access meets stakeholder needs
- ☐

Guaranteeing that the organizational scheme complies with security standards
- ☐

Establishing that the compiled data scales at a sufficient rate

Expand

Correct

When stakeholders pull data from the enterprise data repository, one of the key challenges is ensuring that the method of data access meets stakeholder needs.

2. In the data ecosystem, what are data scientists responsible for?

1 / 1 point

- ☐

Using the data to generate insights and identify patterns
- ☒

Creating predictive models based on the data
- ☐

Converting raw data from disparate sources into usable data
- ☐

Providing business solutions based on insights and predictions

Expand

Correct

Data scientists analyze data for actionable insights and build machine learning or deep learning models that train on past data to create predictive models.

3. Which of the following data sources contain structured data?

1 / 1 point

- ☐

Sensors
- ☐

Tweets
- ☒

Spreadsheets
- ☐

Emails

Expand

Correct

Spreadsheets are an example of structured data since data within a spreadsheet has a well-defined structure that can be represented in a tabular manner with rows and columns.

4. R is ideal for which of the following tasks?

1 / 1 point

- ☒

Developing statistical software and performing data analytics
- ☐

Automating repetitive tasks that are time-consuming to execute one line at a time
- ☐

Accessing information from relational databases
- ☐

Creating mobile and desktop web applications quickly

Expand

Correct

Widely used for developing statistical software and performing data analytics, R is especially known for its ability to create compelling visualizations, giving it an edge over some of the other languages that data professionals use.

5. Which of the following is an advantage of using relational databases?

0 / 1 point

- ☐

Easy backup and disaster recovery
- ☐

Data redundancy across tables
- ☒

Optimization for low-cost hardware
- ☐

Unlimited length of data fields

Expand

Incorrect

Refer to the RDBMS video.

6. What is the relationship between data pipelines and data integration?

1 / 1 point

- ☒

A data pipeline is a process used to perform data integration.
- ☐

Data integration is a process used to create a data pipeline.
- ☐

A data pipeline is a subprocess within data integration.
- ☐

Data integration is a process that functions separately from the data pipeline.

Expand

Correct

A data pipeline is a process used to perform data integration, while ETL is a process within data integration.

7. In a data platform architecture, the data collection, or ingestion, layer performs which of the following tasks?

1 / 1 point

- ☐

Transforms data
- ☒

Maintains metadata
- ☐

Delivers processed data
- ☐

Supports querying tools

Expand

Correct

The data collection, or ingestion, layer uses the metadata repository to maintain information about the data collected.

8. When using SQL, for what purpose should you use the STDDEV function?

1 / 1 point

- ☐

Returning a subset of specific values in a data set
- ☐

Determining the first and last values in a data set
- ☐

Calculating the rank of values in a data set
- ☒

Measuring how spread out the values are in a data set

Expand

Correct

Use the STDDEV() function to measure how spread out the numbers are in a data set. This function calculates the standard deviation.

9. What role does data erasure software play in ensuring compliance with data governance regulations?

1 / 1 point

- ☐

Edits data
- ☒

Overwrites data
- ☐

Updates data
- ☐

Deletes data

Expand

Correct

Data erasure software overwrites the data, permanently clearing the data from the system. This method for discarding data is preferable to simply deleting the data since you can still retrieve deleted data.

10. Which of the following is a mainstay for your career in data engineering?

1 / 1 point

- ☒

Broad understanding of the scope and use of data in business
- ☐

Professional certifications from respected online programs
- ☐

Extensive experience in public speaking and delivering presentations
- ☐

Master's degree in computer science or engineering

Expand

Correct

You can start a career in data engineering through various paths. However, your technical abilities will be your mainstay in this profession, along with a broad understanding of the scope and use of data in business.