

▲ Try again once you are ready
Grade received 77.50% To pass 80% or higher

Try again

Enterprise Data Storage

Total points 4

1. As data evolves during its life cycle, which of the following factors should ML pipelines address to operate properly?(check all that apply).

0.6 / 1 point

☒ Account for anomaly detection.

✔ Correct

Way to go! For example data errors must be handled in the same way as code bugs.

☒ Monitor model and data provenance.

✘ This should not be selected

Monitor provenance is an important aspect of ML pipelines, but it will not help in coping with evolving data. Please rewatch this [video](#) for clarification.

☒ Account for scalable solutions.

✔ Correct

Spot on. Production traffic will vary from day to day and thus your pipeline must scale accordingly.

☒ Use feature engineering.

✘ This should not be selected

Too bad, wrong answer. Feature engineering will not help in coping with evolving data. Please rewatch this [video](#) for clarification.

☒ Provide resilient mechanisms for disruptions.

✔ Correct

Keep it up!. For example ML pipelines should incorporate resilient mechanisms to deal with inconsistent data.

2. Many modeling problems use identical or similar features, and there is substantial value in enabling teams to share features between their own projects and for teams in different organizations to share features with each other. Which of the following storage solutions is deliberately designed to address these user cases?

0.75 / 1 point

☒ Data warehouse

✘ This should not be selected

Not really, data warehouses are repositories that aggregate data from one or more sources. Please rewatch this [video](#) for clarification.

☐ Relational database

☐ Data lake

☒ Feature Store

✔ Correct

Correct! [Feast](#) is an example of an open source feature store.

3. Which are the main advantages of using a cloud-based data warehouse?(check all that apply)

0.75 / 1 point

☒ Provides easy on-demand scalable solution

✔ Correct

Nice going! Cloud solutions are really flexible for scaling up.

☐ User needs to handle all maintenance

☐ They are cost efficient

☐ User owns and controls data governance.

You didn't select all the correct answers

4. About data lakes it's only true that:

1 / 1 point

- ☐ Handles only structured data.
- ☐ Handles only processed data
- ☐ Aggregates data from a single source only.
- ☒ Can handle both structured and unstructured data.



Correct

That's right! Data lakes are really flexible in the type of data they can handle.