---

1. A cloud data analyst is validating a dataset. They find errors in the data. What are two validation rules that the cloud data analyst can use to handle these errors?

Select two answers.

- Stream the data.
- ✅ **Flag the data.**
  Correct: Flagging helps identify problematic records for review or correction.
- Move the data.
- Batch the data.
- ✅ **Discard the data.**
  Correct: Discarding removes invalid or unusable data to maintain data quality.

---

2. A cloud data analyst validating data, checks the uniqueness of the data. What validation technique are they using?

- Format validation
- Type validation
- ✅ **Duplicate validation**
  Correct: Checking uniqueness is part of duplicate validation, ensuring no repeated records.
- Null validation

---

3. A cloud data analyst is ready to start the loading phase of their pipeline. As a next step, they need to prepare the storage to receive the data. What is an action they should perform before starting to load the data?

- ✅ **Create new tables and/or directories.**
  Correct: Storage must be structured and ready to receive data, typically by creating tables or directories.
- Create new data and queries.
- Design the extract phase.
- Design the transformation phase.

---

4. A cloud data analyst is using batch load to load the data into their final destination. What could be a consequence of this type of load?

- Slow streaming of the data.
- ✅ **Overload of the destination storage.**
  Correct: Batch loads can push large volumes of data at once, potentially overwhelming storage systems.
- Slow transformation of the data.
- Reformat of the destination storage.