

# Certification Overview

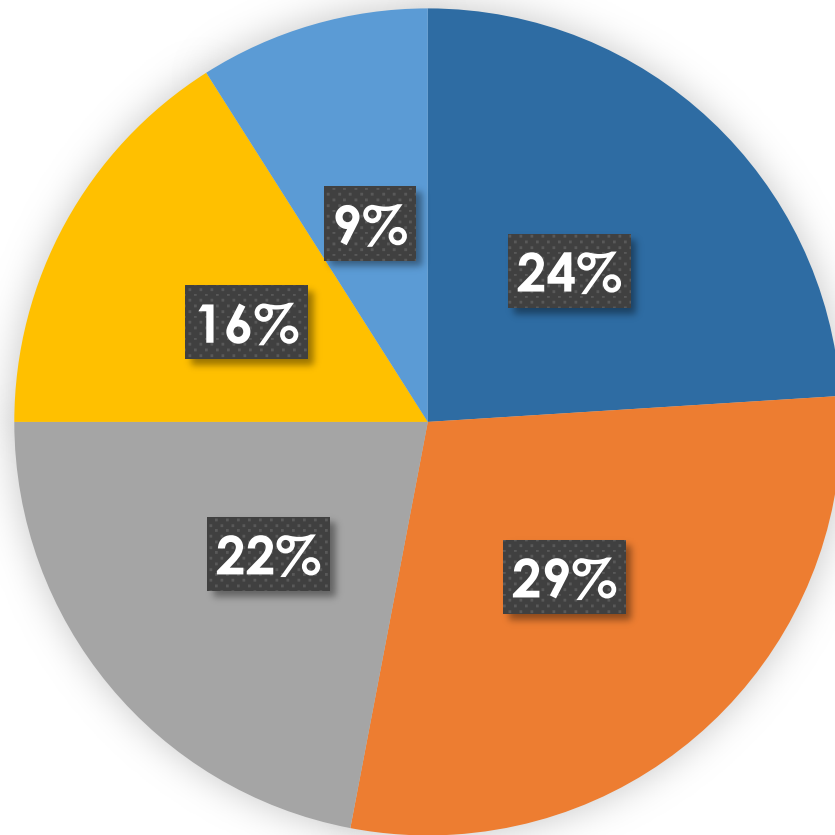
# Learning Objectives

- ▶ Format and structure of the exam.
- ▶ The topics covered in the exam.
- ▶ Types of questions provided on the exam

# Exam Details

- ▶ Time allotted to complete exam = 1.5 hours
- ▶ Number of Questions = 45
- ▶ Passing scores = At least 70% (32/45)
- ▶ Exam fee = \$200
- ▶ Exam retake fee = \$200

# Exam Questions



- Databricks Lakehouse Platform (11/45)
- ELT with Spark SQL and Python (13/45)
- Incremental Data Processing (10/45)
- Production Pipelines (7/45)
- Data Governance (4/45)

# Out-of-scope

- ▶ Apache Spark internals
- ▶ Databricks CLI and REST API
- ▶ Change Data Capture CDC/CDF
- ▶ Data modeling concepts
- ▶ GDPR/CCPA
- ▶ Monitoring and logging production jobs
- ▶ Dependency management
- ▶ Testing


# Code Examples

- ▶ Code will be in mainly in SQL
- ▶ Otherwise, Python

# Exam Platform

<https://www.webassessor.com/databricks>

[Home](#) | [Login](#) | [Forgot Password](#) | [Create New Account](#)

 **databricks**

Company | Contact | Register for an Exam

## Databricks Certification

Welcome to Databricks online assessment registration.

Follow these steps to schedule an assessment:

1. Create a New Account by clicking [here](#).
2. Log into your newly created account.
3. Click on 'Register for an Assessment' and follow the remaining steps to complete scheduling.

# Exam Proctoring

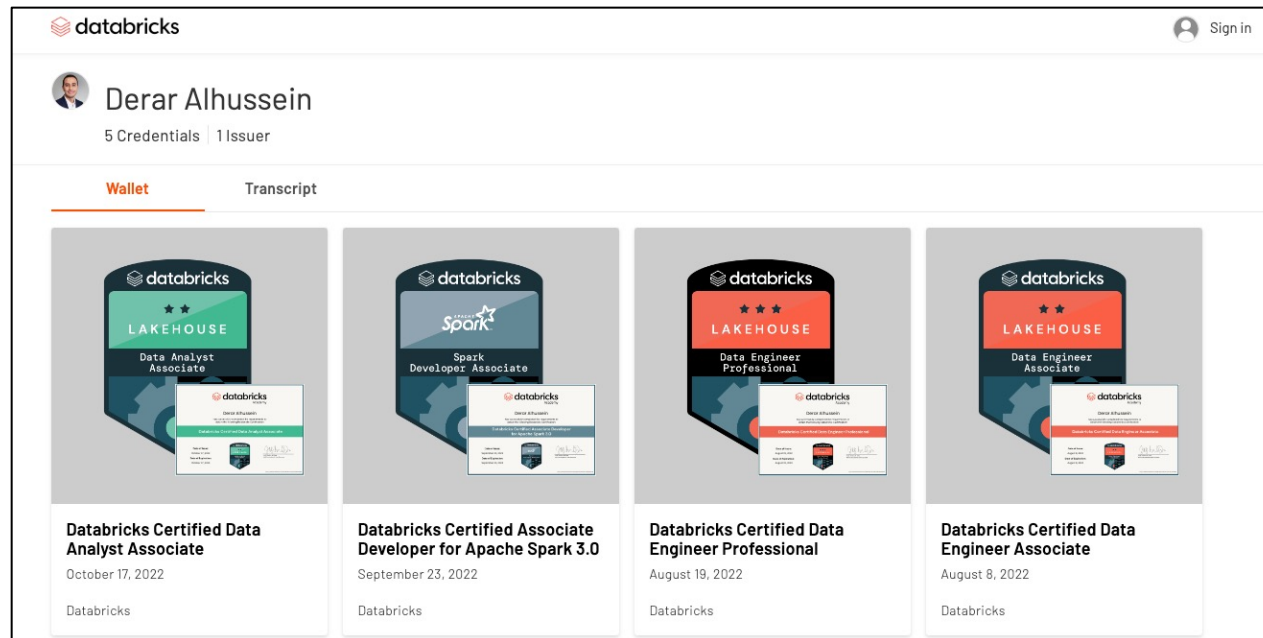
- ▶ Monitored via webcam by a Webassessor proctor.
- ▶ Asked to provide valid, photo-based identification.
- ▶ Proctor does not provide assistance on the content of the exam
- ▶ No test aids will be available during the exam



# Exam Result

- ▶ Certification exams are automatically graded.
- ▶ 24 hours to receive the badge and Certificate

<https://credentials.databricks.com>



# Questions Types

- ▶ Multiple-choice questions – Only one correct answer
- ▶ Questions Types:
  - ▶ Conceptual Questions
  - ▶ Code-Based Questions

# Conceptual Questions

- ▶ Which part of the Databricks Lakehouse Platform can data engineers use to orchestrate jobs ?
  - A. Repos
  - B. Workflows
  - C. Data Explorer
  - D. Databricks SQL
  - E. Cluster

# Conceptual Questions

- ▶ Which part of the Databricks Lakehouse Platform can data engineers use to orchestrate jobs ?
  - A. Repos
  - B. Workflows**
  - C. Data Explorer
  - D. Databricks SQL
  - E. Cluster

# Code-Based Questions

```
spark.table("sales")  
  .writeStream  
  .option("checkpointLocation", checkpointPath)  
  ._____   
  .table("new_sales")
```

If you want the query to execute a single micro-batch to process all of the available data, which of the following lines of code should you use to fill in the blank ?

- A. trigger(once=True)
- B. trigger(continuous="once")
- C. processingTime("once")
- D. trigger(processingTime="once")
- E. processingTime(1)

# Code-Based Questions

```
spark.table("sales")  
  .writeStream  
  .option("checkpointLocation", checkpointPath)  
  .  
  .table("new_sales")
```

If you want the query to execute a single micro-batch to process all of the available data, which of the following lines of code should you use to fill in the blank ?

- A. **trigger(once=True)**
- B. trigger(continuous="once")
- C. processingTime("once")
- D. trigger(processingTime="once")
- E. processingTime(1)

# Practice Exam

- ▶ Databricks official Practice test
- ▶ PDF file



# Derar Alhussein

---

[www.linkedin.com/in/DerarAlhussein](https://www.linkedin.com/in/DerarAlhussein)