Certification Overview

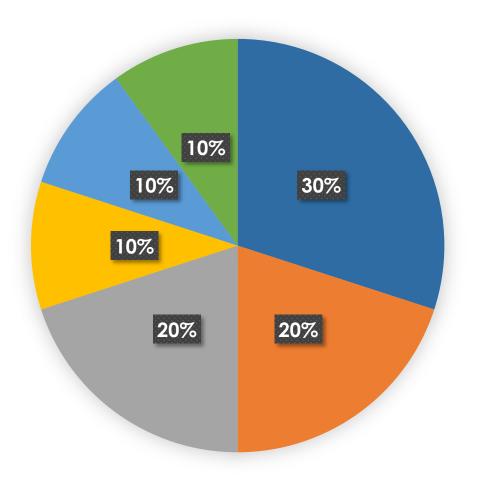
Learning Objectives

- ▶ 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 = 120 minutes
- Number of Questions = 60
- ▶ Passing scores = At least 70% (42/60)
- ► Exam fee = \$200
- Exam retake fee = \$200

Exam Questions



- Data Processing (18/60)
- Data Modeling (12/60)
- Databricks Tooling (12/60)
- Security and Governance (6/60)
- Testing and Deployment (6/60)
- Monitoring and Logging (6/60)

Out-of-scope

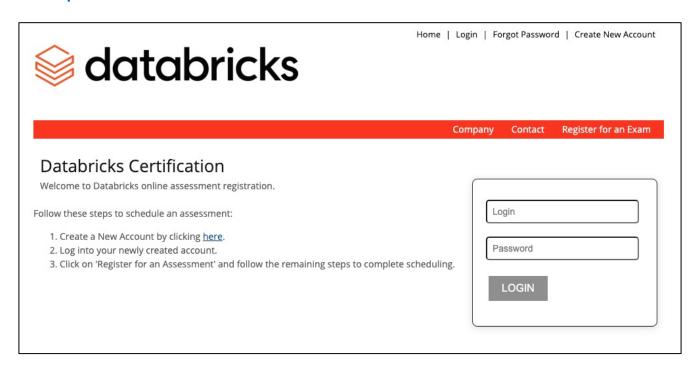
- Delta Live Tables
- Scala
- Orchestration tools (Airflow, ADF, etc)
- Pub/Sub systems configurations (Kafka, ...)
- Automation servers (Jenkins, Azure DevOps, etc)
- Managed CI/CD
- Gitflow
- Cloud-specific security and integrations
- Infrastructure as code (Terraform, ...)

Code Examples

- ► Code will be in mainly in Python
- ▶ Delta Lake functionality in SQL

Exam Platform

https://www.webassessor.com/databricks



Exam Proctoring

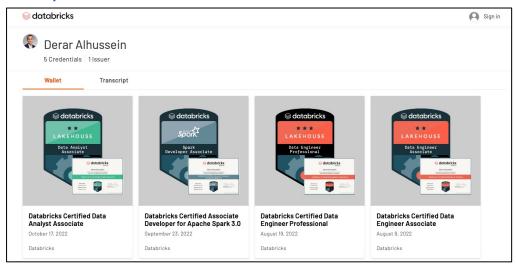
- Monitored via webcam by a Webassessor proctor.
- ▶ Asked to provide valid, photo-based identification.

- 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)

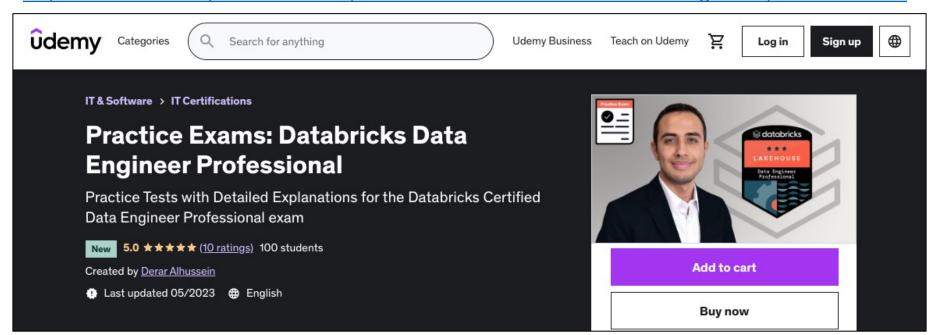
Practice Exam

▶ No official Practice test from Databricks

Practice Exam

Practice exams course on Udemy

https://www.udemy.com/course/practice-exams-databricks-data-engineer-professional-k/





Derar Alhussein

www.linkedin.com/in/DerarAlhussein