

MICROSOFT OFFICIAL CURRICULUM

Optimizing Apache Spark

Databricks Academy-aligned program for Optimizing Apache Spark.

Databricks

Program aligned

Data

Advanced

Databricks Data Intelligence Platform

DURATION
1 to 2 days

LEVEL
Advanced

FORMAT
Virtual, On-site, or Hybrid

CERTIFICATION
**Databricks Academy
learning plan**

AUDIENCE PROFILE

Built for these roles

Built for Data Engineer learners adopting Optimizing Apache Spark.

OVERVIEW

Executive overview

For those familiar with Apache Spark programming, this is the best data engineering course to learn how to mitigate bottlenecks with the Spark UI. This course explores five major performance problems for Apache Spark applications in production: Skew, Spill, Shuffle, Storage and Serialization. We'll work with 1 TB+ data sets to diagnose these issues and discuss mitigation strategies. We'll also explore optimization techniques for data ingestion, including managing Spark-partition sizes, disk-partitioning, bucketing, Z-Ordering and more. Students can also expect the course to cover performance c

PROGRAM OUTCOMES

Capabilities your teams will gain

- Strengthen capability in apache spark scenarios
- Strengthen capability in databricks data intelligence platform scenarios

ENTERPRISE CUSTOMIZATION

Enterprise customization

Tailor this program to your organization's priorities: Develops Databricks lakehouse delivery capability for apache spark teams through official Academy-aligned training.

Customization options

- Align labs to your production environment and platform priorities
- Add readiness reviews and instructor-led practice sessions
- Extend into project-specific architecture or delivery coaching

CURRICULUM

Curriculum roadmap

1

MODULE 1

Apache Spark

Cover apache spark skills and implementation practices aligned to Optimizing Apache Spark.

- Cover apache spark skills
- implementation practices aligned to Optimizing Apache Spark

2

MODULE 2

Databricks Data Intelligence Platform

Cover databricks data intelligence platform skills and implementation practices aligned to Optimizing Apache Spark.

- Cover databricks data intelligence platform skills
- implementation practices aligned to Optimizing Apache Spark

PRIVATE DELIVERY

Plan a private cohort for your team

Scope delivery format, role mix, and rollout timeline with our solutions team.

Plan Private Delivery

Book advisory call

vnodeites.com

TALK TO US

-  info@vnodeites.com
-  +91 9779 91 4792
-  WhatsApp · +91 9779 91 4792
-  +91 9419 11 4792

OFFICE

DLF Cyber City, Gurugram
Haryana — India
Azure · AI · Data · Power Platform