

Databricks Course Contents (5 days)

By Dr. Vishwanath Rao

Learning Objectives

- Apache Spark (Overview)
- Real-world uses for Databricks
- How Databricks and Apache Spark fit together
- Getting started with Databricks
- Uploading, preparing and transforming data
- Creating Notebooks
- Creating Clusters
- Running Jobs
- Delta Lake and Delta Tables

Overview of Apache Spark and Databricks

- How do we define Big Data?
- Spark languages - Scala, Python, R, Java, SQL
- Databricks Community Edition
- Databricks Architecture
- Defining Data Analytics
- Defining Machine Learning
- Azure implementation
- AWS implementation
- Databricks
- Creating Databricks Workspace UI
- The Azure Databricks Service Blade
- The Databricks Dashboard
- Databricks Cluster Creation UI
- Databricks File System (DBFS)
- Databricks Integration with Data Lake

Databricks benefits

- Integrating into Pipelines
- Databricks Developer Experience
- Development Environments
- Which Databricks-Supported Language Should I Use?
- Notebook Runtime Flavor Configuration
- The Notebook UI
- Creating Tables
- Create a New Table UI

- Creating a Table from a DBFS File

Uploading data

- Connecting to a Spark data source
- Previewing your Table
- Columns and Datatypes basics

Bringing your data into your Notebook

- Writing the initial SQL query to import
- View aggregates
- Perform Joins

- Creating Your Table Visually with Databricks UI (The Preview Screen)
- Querying a Databricks Table using SQL
- Spark and Databricks
- Real-time Transformations
- Databricks Machine Learning (ML)
- The Cost of Doing Business on Databricks

Visualisations & DataFrames

- Datatypes
- Structured Streaming DataFrames
- Plots
- Choosing Chart types
- Chart Toolbar
- Layout and styling considerations
- Machine Learning visualisations
- A Data Profile Visualization Example
- Performing Exploratory Data Analysis (EDA) with Data Charts

Databricks Jobs

- View Jobs and Job details
- Automation Jobs
- Running your first Job
- Viewing completed jobs
- Setting up Alerts

Delta Lake and Delta Tables

- Getting data into Delta Lake
- Delete, update, merge
- Overview of Delta Engine

