

SPARK Syllabus

A)Scala

- Intro to Scala
- Base Object Oriented Programming in Scala
- Case Objects and Classes

B)Spark Overview

- Spark Opportunity and Solution
- Capabilities and Ecosystem
- Spark Components vs Hadoop

C)RDD Fundamentals

- Purpose and Structure of RDDs
- Transformations, Actions, and DAG
- RDD programming API

D)Spark SQL / Dataframes

- Spark SQL and DataFrame Uses
- DataFrame / SQL APIs
- Catalyst Query Optimization

E)Spark Job Execution

- Jobs, Stages, and Tasks
- Partitions and Shuffles

- Data Locality
- Job Performance

F) Spark Streaming

- Streaming Sources and Tasks
- DStream APIs and Stateful Streams
- Reliability and Fault Recovery



SevenMentor
PVT.LTD