

# Mastering MapReduce

## Module 01 : Introduction to MapReduce

- What is MapReduce?
- Why is MapReduce important?
- MapReduce vs traditional data processing

## Module 02 : Understanding MapReduce Architecture

- How MapReduce works
- The MapReduce runtime environment
- The role of the JobTracker and TaskTracker

## Module 03 : Developing MapReduce Applications

- Setting up the development environment
- Writing MapReduce programs in Python
- Writing MapReduce programs in Java(Optional)

## Module 04 : Basic MapReduce Concepts

- Input and output formats
- Mapper and reducer functions
- Partitioning and shuffling

## Module 05 : Advanced MapReduce Concepts

- Combiners and their role
- Joining datasets in MapReduce
- MapReduce best practices

## **Module 06 : MapReduce Performance Tuning**

- Common performance bottlenecks
- Tuning MapReduce for performance
- Monitoring and profiling MapReduce jobs