Here's a detailed outline for a comprehensive Hadoop course:

Week 1: Introduction to Hadoop

- **Overview of Hadoop:** History, ecosystem, and architecture.
- **Hadoop Components:** HDFS (Hadoop Distributed File System), MapReduce, YARN (Yet Another Resource Negotiator).
- **Hadoop Installation:** Setting up a single-node Hadoop cluster.

Week 2: Hadoop Distributed File System (HDFS)

- **HDFS Architecture:** NameNode, DataNode, Secondary NameNode.
- **Data Storage and Replication:** How data is stored and replicated in HDFS.
- **HDFS Commands:** Basic commands for file operations in HDFS.

Week 3: MapReduce Fundamentals

- **MapReduce Programming Model:** Mapper and Reducer concepts.
- **Writing MapReduce Jobs:** Writing, compiling, and running MapReduce jobs in Java.
- **Debugging and Optimizing MapReduce Jobs:** Techniques for performance improvement.

Week 4: YARN and Resource Management

- **YARN Architecture:** ResourceManager, NodeManager, ApplicationMaster.
- **Job Scheduling:** How YARN schedules and manages resources.
- **Monitoring and Troubleshooting YARN Jobs:** Tools and techniques.

Week 5: Hadoop Ecosystem Tools

- **Apache Hive:** Data warehousing, HiveQL, and data querying.
- **Apache HBase:** NoSQL database, architecture, and CRUD operations.
- **Apache Pig:** Scripting language for data processing, Pig Latin basics.

Week 6: Data Ingestion and Integration

- **Apache Sqoop:** Importing and exporting data between Hadoop and relational databases.

- **Apache Flume:** Collecting, aggregating, and moving large amounts of log data.

Week 7: Data Processing with Apache Spark

- **Introduction to Apache Spark:** Spark architecture and RDDs (Resilient Distributed Datasets).
- **Spark SQL and DataFrames:** Querying data and performing transformations.
- **Spark Streaming:** Real-time data processing with Spark.

Week 8: Data Management and Governance

- **Data Management:** Best practices for managing large datasets.
- **Data Governance:** Policies, data lineage, and auditing.
- **Security in Hadoop:** Authentication, authorization, and encryption.

Week 9: Advanced Topics

- **Hadoop Ecosystem Integration:** Integrating Hadoop with other tools and technologies.
- **Performance Tuning:** Optimizing Hadoop cluster performance.
- **Scaling Hadoop Clusters:** Strategies for scaling and managing large clusters.

Week 10: Real-World Projects and Case Studies

- **Case Studies:** Real-world applications and case studies.
- **Project Work:** Implementing a project using Hadoop technologies (e.g., data processing, analytics, or ETL).

Week 11: Future Trends and Technologies

- **Emerging Trends:** New developments and future directions in Hadoop and Big Data.
- **New Tools and Technologies:** Exploring additional tools and frameworks in the Hadoop ecosystem.

Week 12: Review and Exam Preparation

- **Review:** Recap of key concepts and tools.
- **Exam Preparation:** Practice questions and review of important topics.