

# INFOSOFT IT SOLUTIONS

**Training | Projects | Placements**

Revathi Apartments, Ameerpet, 1<sup>st</sup> Floor, Opposite Annapurna Block, Infosoft It solutions,  
Software Training & Development Institute, +91-9059683947|+91-9182540872

## Hive Course

### **Introduction to Apache Hive**

1. Overview of Hive
  - What is Apache Hive?
  - History of Hive
2. Architecture of Hive
  - Components of hive architecture
  - How Hive interacts with Hadoop
3. Hive vs. Traditional Databases
  - Comparison with RDBMS
  - Use cases for Hive
4. Installing and Configuring Hive
  - Prerequisites for installation
  - Steps to install Hive
5. Basic Hive Commands
  - Starting hive shell
  - Basic hive queries

### **Manipulating Data in Hive**

1. Hive Data Types
  - Primitive data types
  - Complex data types
2. Tables in Hive
  - Creating tables

- Understanding managed and external tables
- 3. Importing Data
  - Data import methods
  - Loading data from local and HDFS
- 4. Exporting Data
  - Exporting data to local and HDFS
  - Using data export tools
- 5. Modifying and Updating Data
  - Updating existing records
  - Deleting data

## **Getting Data from Hive**

1. HiveQL Select
  - Selecting columns
  - Using expressions in select
2. Sorting and Ordering Data
  - Ordering the results
  - Sorting vs ordering
3. Hive Functions
  - Built-in functions
  - Creating custom functions
4. Saving Query Results
  - Saving results to HDFS
  - Storing results in the local file system

## **Aggregating Data with Hive**

1. Using Group By
  - Syntax of the group by
  - Grouping data for aggregation
2. Hive Aggregate Functions
  - Common aggregate functions
  - Using aggregates in queries

3. Advanced Aggregations
  - Cube and rollup
  - Grouping sets
4. Having Clause
  - Filtering aggregated data
  - Syntax and use cases of having

## **Filtering Results with Hive**

1. Basic Filtering
  - Using where clause
  - Filtering data with predicates
2. Advanced Filtering
  - Working with boolean operators
  - Complex filter conditions
3. Pattern Matching
  - Using like operator
  - Regex matching in Hive
4. Handling NULL Values
  - Understanding NULL in Hive
  - Filtering with NULL values

## **Joining Tables**

1. Understanding Joins in Hive
  - Types of joins in Hive
  - Join conditions and how they work
2. Implementing Joins
  - Syntax of joins
  - Performance considerations
3. Multi-table Joins
  - Joining multiple tables
  - Handling large table joins
4. Optimizing Joins

- Map join hint
- Optimizing join strategies

## **Manipulating Data**

1. Hive Transactions
  - ACID properties in Hive
  - Managing transactions
2. Data Transformation
  - Using transform for data manipulation
  - Writing custom scripts for transformation
3. Indexing in Hive
  - Creating indexes
  - Index types and usage
4. Optimizing Hive Queries
  - Best practices for query optimization
  - Understanding explain plan
5. Hive Scripting
  - Writing hive scripts
  - Automating tasks with scripts