Document Generated: 11/05/2026

Learning Style: On Demand

Technology:

Difficulty: Intermediate

Course Duration: 15 Hours

From 0 to 1: Hive for Processing Big Data



Taught by a 4 person team including 2 Stanford-educated, ex-Googlers and 2 ex-Flipkart Lead Analysts. This team has decades of practical experience in working with large-scale data.

Hive is like a new friend with an old face (SQL). This course is an end-to-end, practical guide to using Hive for Big Data processing.

Let's parse that

A new friend with an old face: Hive helps you leverage the power of Distributed computing and Hadoop for Analytical processing. It's interface is like an old friend : the very SQL like HiveQL. This course will fill in all the gaps between SQL and what you need to use Hive.

End-to-End: The course is an end-to-end guide for using Hive: whether you are analyst who wants to process data or an Engineer who needs to build custom functionality or optimize performance - everything you'll need is right here. New to SQL? No need to look elsewhere. The course has a primer on all the basic SQL constructs, .

Practical: Everything is taught using real-life examples, working queries and code .

Audience:

- Analysts who want to write complex analytical queries on large scale data
- Engineers who want to know more about managing Hive as their data warehousing solution

Course Objective:

Analytical Processing: Joins, Subqueries, Views, Table
Generating Functions, Explode, Lateral View, Windowing and more

Tuning Hive for better functionality: Partitioning, Bucketing, Join Optimizations, Map Side Joins, Indexes, Writing custom User Defined functions in Java. UDF, UDAF, GenericUDF, GenericUDTF, Custom functions in Python, Implementation of MapReduce for Select, Group by and Join

For SQL Newbies: SQL In Great Depth

Prerequisites:

Hive requires knowledge of SQL. The course includes and SQL primer at the end. Please do that first if you don't know SQL. You'll need to know Java if you want to follow the sections on custom functions.

Course Outline:

- You, Us & This Course
- Introducing Hive
- Hadoop and Hive Install
- Hadoop and HDFS Overview
- Hive Basics
- Built-in Functions
- Sub-Queries
- Partitioning
- Bucketing

- Windowing
- Understanding MapReduce
- MapReduce logic for queries: Behind the scenes
- Join Optimizations in Hive
- Custom Functions in Python
- Custom functions in Java
- SQL Primer - Select Statemets
- SQL Primer - Group By, Order By and Having
- SQL Primer - Joins
- Appendix

Credly Badge:



Display your Completion Badge And Get The Recognition You Deserve.

Add a completion and readiness badge to your Linkedin profile, Facebook page, or Twitter account to validate your professional and technical expertise. With badges issued and validated by Credly, you can:

- Let anyone verify your completion and achievement by clicking on the badge
- Display your hard work and validate your expertise
- Display each badge's details about specific skills you developed.

Badges are issued by QuickStart and verified through Credly.

Find Out More or See List Of Badges