Document Generated: 11/24/2026

Learning Style: On Demand

Technology:

Difficulty: Intermediate

Course Duration: 15 Hours

From 0 to 1: Hive for Processing Big Data



About this course:

This course is taught by an incredible team, which consists of two Stanford graduates, a former Google employee, and two former Flipkart Lead Analysts. This team has decades of formal experience both individually and together.

Hive is based off on SQL but it comes with a modern twist. This course will walk you through the complete process of using Hive for the processing of Big Data.

Hive combines the wonders of Distributed Computing and Analytical processing using Hadoop on one single platform. However, even though HiveQL is a fairly new technology it is very much like SQL. Also, many operations are pretty similar as well. This course will help you bridge the gaps between the two.

This course provides a comprehensive end to end guide on using Hive. Both Analysts who work on data or Engineers who want to customize functions for optimal performance can leverage the course and use it to their needs.

Also, each topic in this course has real-time examples included in them.

Audience:

With this course, we target the following audience:

- Data Analysts who are responsible for complex analytical queries on massive amounts of data
- Engineers who want to gain an idea about Hive as a data warehouse solution

Course Objective:

- For Data Analysts: Joins, View, Explode, Table Generating Functions, Sub-queries, Windowing, Lateral View, and other basics.
- For Hive Engineers: Bucketing, Map Side Joins, User-defined functions using Java, UDF and UDAF (GenericUDF and GenericUDFT), Partitioning, Map-Reduce including Select, Join, and Group, and much more
- For SQL Beginners: SQL in all its glory

Prerequisites:

Students must have some basic concepts of SQL, however, it is completely fine if you don’t know it yet. This course comes with a SQL primer at the end, refer to it before taking the course.

Other than that, all candidates shall have some basic Java understanding, however, it is not compulsory.

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