Introduction to Apache Kudu

Take your knowledge to the next level
Cloudera's Introduction to Apache Kudu training teaches students the basics of Apache Kudu, a new data storage system for the Hadoop platform that is optimized for analytical queries.

Get hands-on experience
Through instructor-led discussion, as well as hands-on exercises, participants will learn topics including:

- A high-level explanation of Kudu
- How does it compares to other relevant storage systems and which use cases would be best implemented with Kudu
- Learn about Kudu's architecture as well as how to design tables that will store data for optimum performance.
- Finally, learn data management techniques on how to insert, update, or delete records from Kudu tables using Impala, as well as bulk loading methods

What to expect
This material is intended for a broad audience of students involved with either software development or data analysis. This would include software developers, data engineers, DBAs, data scientists, and data analysts.

As with our Data Analyst training course, knowledge of SQL is assumed, as is basic Linux command-line familiarity.

“The professionalism and expansive technical knowledge demonstrated by our instructor were incredible. The quality of the Cloudera training was on par with a university.”
General Dynamics
Course Contents

1. Introduction
2. Kudu Overview
   • What is Kudu?
   • Why Do We Need Kudu?
   • Kudu Use Cases
   • Strengths of Kudu
   • Kudu Platform Information
3. Kudu Tables
   • Kudu Concepts
   • Understanding Kudu Table Storage
   • Designing Kudu Tables
   • Partitioning and Performance
   • Kudu Architecture
4. Developing Kudu Applications with Impala
   • Using Impala with Kudu Tables
   • Populating Kudu Tables
   • Querying Kudu Data with Impala
   • Advanced Kudu Topics
5. Conclusion