Pig, Hive & Impala

Apache Pig, Hive, and Impala are important components of the Big Data Hadoop ecosystem that enable data analytics and transformations using various user-defined functions, joins and filters from other technologies. Pig is a procedural language which helps in easy programming of complex MapReduce programs. Hive is a declarative language like SQL that is used by data analysts for creating reports This Koenig course will help big data analysts make the most of their deployments and utilize some of the vital tools and programming languages for advanced processing. Impala is a high performance SQL engine which is used for processing huge amounts of data.

- Introduction to Pig
- Pig programming and commands
- Pig interactive shell
- Analyzing complex data using Pig
- Using Pig functions
- Introduction to Hive
- Hive Metastore
- Writing queries using Hive
- Joins and grouping
- Impala architecture
- Impala Vs. Hive
- Queries using Impala

Course #	Duration
Pig, Hive & Impala	**6 days**