Pig for wrangling Bigdata

Modality: On Demand

Duration: 6 Hours

About the course:

Instructed by a group that incorporates 2 Stanford-taught, 2 ex-Flipkart, and ex-Googlers Lead Analysts. This group has many years of down to earth involvement with working with enormous scope information handling works.

Pig is relevantly named, will devour any information that you toss at it and make a few bucks!

Let's parse that

Omnivorous: Pig functions with unstructured information. It has numerous tasks that are very SQL-like yet Pig can play out this procedure on informational collections that have no fixed blueprint. Pig is extraordinary at wrestling information into a structure that is perfect and can be put away in an information stockroom for analysis and reporting.

Make home a few bucks: Pig permits you to change information such that makes is organized, unsurprising and helpful, prepared for utilization.

Audience:

Designers who need to parse and extricate helpful data from enormous datasets

Experts who need to wrangle enormous, unstructured information into shape

Course Objective:

Pig Basics: Complex and Scalar information types (Maps, Bags, Tuples), fundamental transformations, for example, Foreach, Filter, Dump, Store, Load, Limit, Distinct, Order by and other inherent capacities.

Advanced-Data Optimizations and Transformations: The brain-twisting Nested Foreach, Links and their enhancements utilizing "merge", "parallel", "replicated" and different Co-gatherings, keywords, and Semi-joins, troubleshooting utilizing Illustrate and Explain directions

True model: Clean up server logs utilizing Pig

Prerequisites:

Working with Pig requires some essential information on the SQL question language, a concise comprehension of the MapReduce and Hadoop eco-framework.

Course Outline:

You, This Course and Us

Where does Pig fit in?

Pig Basics

Pig Operations And Data Transformations

Advanced Data Transformations

Optimizing Data Transformations

A real-world example

Installing Hadoop in a Local Environment