Apache Druid

Description:

Apache Druid is a real-time database to power modern analytics applications. It is a modern technology which has made itself indispensable in the world of big data and data analytics. It comes up with cutting-edge features like easy integration in data pipelines, fast consistent queries, high concurrency support, support in both on-prem and cloud infrastructure, and many more.

Start Date:

Doubt Clear Time:

Course Time:

Features:

Course material

Course resources

On demand recorded videos

Practical exercises

Quizzes

Assignments

Course completion certificate

What we learn:

Introduces Apache Druid

Apache Druid Installation

Walkthrough of Apache Druid Console

Data Transformation while ingestion

Data filtering while ingestion

Nested data parsing while ingestion

Data Rollup while ingestion

Introduction to Data Segments

Data Deletion in Druid

Other ways of loading data into Druid

Querying data in Druid

Druid Internals

Druid integration with Superset

Druid at Netflix

Requirements:

System with Internet Connection

Interest to learn

Dedication

Instructor:

Name:

Shruti Mantri

Description:

Shruti Mantri is a well-known software architect, instructor and mentor in the industry. She has 10+ years of experience in the software industry, and has worked with different organizations like Oracle, Flipkart, Amazon, Myntra and Twitter. She is known for her expertise in the data engineering field, and has a sound knowledge on the latest technologies in this domain. She has helped develop data platform at organizations, and guided several mentees in understanding data engineering and how to get better at it.

>Introduces Apache Druid:

>>An Introduction to Apache Druid

>>Course Objectives

>>Where to use & where not to use Druid

>Apache Druid Installation:

>>Single Server Installation

>>Overview of the Druid Console

>Understanding Data Load:

>>Load data from file using console

>>Other ways to load data from file

>>Transformations over data

>>Applying filters over data

>>Parsing nested data

>>Rollup Data

>Other Data Loads:

>>Load data from Kafka

>>Load data from Https

>>Load ORC data format

>Querying:

>>Druid SQL Queries

>>Native Queries

>>Querying Data via Http

>Druid Internals:

>>Druid Architecture

>Use Cases:

>>Integrating Superset with Druid

>>Industry Use-case: Druid at Netflix