Mindluster Platform

# About the course Apache Spark

What is Apache Spark used for? Apache Spark is an open-source, distributed processing system used for big data workloads. It utilizes in-memory caching and optimized query execution for fast queries against data of any size. Simply put, Spark is a fast and general engine for large-scale data processing.Is Apache Spark a database? Apache Spark can process data from a variety of data repositories, including the Hadoop Distributed File System (HDFS), NoSQL databases and relational data stores, such as Apache Hive. ... The Spark Core engine uses the resilient distributed data set, or RDD, as its basic data type.Is Apache Spark dying? The hype has died down for Apache Spark, but Spark is still being modded/improved, pull-forked on GitHub D-A-I-L-Y so its demand is still out there, it's just not as hyped up like it used to be in 2016. However, I'm surprised that most have not really jumped on the Flink bandwagon yet.

**Programming Category's Courses**

# Course Lesson(50)


Lesson 1 : **Apache Spark Introduction**

Lesson 2 : **Big Data Overview**

Lesson 3 : **Types of Data Under Big data**

Lesson 4 : **Big Data Technologies**

Lesson 5 : **Five V s in Big Data**

Lesson 6 : **What is Hadoop**

Lesson 7 : **Brief History on Hadoop**

Lesson 8 : **Hadoop Ecosystem**

Lesson 9 : **Introduction to Environment Setup**

Lesson 10 : **Installing Virtual Box Setup Machine**

Lesson 11 : **Installing Linux Environment**

Lesson 12 : **Installing JDK and JRE for Hadoop**

Lesson 13 : **Installing Hadoop for Single Node Cluster**

Lesson 14 : **Installing Hadoop on Pseudo Distributed Mode**

Lesson 15 : **Installing Scala into Linux System**

Lesson 16 : **Installing Spark into Linux System**

Lesson 17 : **Installation of Eclipse IDE**

Lesson 18 : **Install Scala Plugin for Eclipse**

Lesson 19 : **Spark Introduction**

Lesson 20 : **Spark Overview**

Lesson 21 : **History of Apache Spark**

Lesson 22 : **Spark Ecosystem**

Lesson 23 : **Features of Apache Spark**

Lesson 24 : **Limitations of Apache Spark**

Lesson 25 : **Use Cases of Apache Spark**

Lesson 26 : **Spark RDD Introduction**

Lesson 27 : **RDD in Spark**

Lesson 28 : **Features of RDDs**

Lesson 29 : **RDD Operation Transformations**

Lesson 30 : **RDD Operation Actions**

Lesson 31 : **Spark Map and FlatMap**

Lesson 32 : **Spark Core Programming Introduction**

Lesson 33 : **Spark Commands**

Lesson 34 : **Spark Transformation Functions**

Lesson 35 : **Spark Action Functions**

Lesson 36 : **Spark Wordcount Example**

Lesson 37 : **Spark SQL Introduction**

Lesson 38 : **Spark SQL Overview**

Lesson 39 : **Spark DataFrames Datasets**

Lesson 40 : **Spark SQL Advantages Disadvantages**

Lesson 41 : **Spark SQL Architecture**

Lesson 42 : **Spark Streaming Introduction**

Lesson 43 : **Spark Streaming Overview**

Lesson 44 : **Spark Streaming Architecture**

Lesson 45 : **DStream in Spark Streaming**

Lesson 46 : **DStream Operations in Spark**

Lesson 47 : **Tutorix Simply Easy Learning Steps**

Lesson 48 : **Tutorix Brings Simply Easy Learning**

Lesson 49 : **Tutorix Brings Simply Easy Learning**

Lesson 50 : **Tutorix Brings Simply Easy Learning**

# Related courses

**React Projects Workshops**

**Node js back end JavaScript**

**Python programming language**

**Social Network Theme UI With Sass**

**Full Stack React Django**

**Build a Node js App With Sequelize**



for Business Contact
business@mindluster.com