[REGEX]*
Software Services- Learn To Create

Registered Office:
B-23, Krishi Nagar II,
Taru Chhaya Nagar,
Taro Ki koot, near B-2 Bypass,
Jaipur, Rajasthan 302029

BigData Analytics & Hadoop

PREREQUISITE:

- Basic understanding of any programming language.

DURATION: 7 Days [15 Hours]

DAY-1	• Big Data Analytics & Hadoop:
	  ○ Introduction to the Term Big Data
	  ○ Work of Big Data Analytics
	  ○ Journey to Start as Data Scientist
	  ○ 6 V's of Big Data
	  ○ Configuring Hadoop Environment in your Local System
	  ○ Cloudera Installation
	  ○ Introduction to Hadoop
	• HDFS [Hadoop Distributed File System]
	  ○ Components of HDFS - NameNode & DataNode
	  ○ Hadoop Daemons
	  ○ Hadoop Admin & Commands

DAY-2	• Map-Reduce [Data Processing]
	  ○ Introduction to Map-Reduce
	  ○ Map-reduce V1 vs YARN
	  ○ Introduction to YARN
	  ○ FS Image & Secondary Namenode
	  ■ Shuffle, Sorting & Partitioning
	  ■ Map reduce Word-Count Problem - Example

DAY-3	• HIVE
	  ○ Introduction to Hive
	  ○ Difference between SQL & Hive
	  ○ Database & Table Creation
	  ○ Internal vs External Table
	  ○ Functions
	  ■ Date & String Functions
	  ○ PARTITIONING & BUCKETING
	  ■ Static vs Dynamic Partitioning
	  ■ Buckets in Hive

.[REGEX]*
Software Services- Learn To Create

Registered Office:
B-23, Krishi Nagar II,
Taru Chhaya Nagar,
Taro Ki koot, near B-2 Bypass,
Jaipur, Rajasthan 302029

DAY-4 **&** **DAY-5**	• Apache Spark on Azure DataBricks ○ Introduction to Apache Spark ○ Usage & workflow of Spark ○ **Trick** - Account creation on Azure DataBricks ○ RDD - Resilient Distributed DataSet ■ Transformation & Action [Operation] ○ RDD Vs DataFrame ○ DataFrame - ■ Creating DataFrame with several file formats ■ Benefits of using Dataframes ■ Manipulating Data Frame ■ Group By operation on Data Frame ○ Introduction to MLLib ■ Linear Regression - Case study
DAY-6	• Neo4j Graph Analytics & NoSQL DataBase ○ Graph Analytics - Introduction ○ Understanding of Graphs ○ Running & Installing Neo4j ○ HBase - Introduction ○ SQL VS NoSQL [Which is better] ○ NoSQL - Introduction ■ Creating Table & column family ■ Create, retrieve, update & delete operation ○ Hands on learning with NoSQL from scratch
DAY-7	• Amazon EMR ○ Introduction to Cloud ○ Running Hadoop Eco-system on cloud ○ Creating 100 Notes cluster within seconds • Capstone Project