CFAMILY IT Solutions
cfamily IT
Learn from Experts & Feel the difference.
AKKEM SREENIVASULU
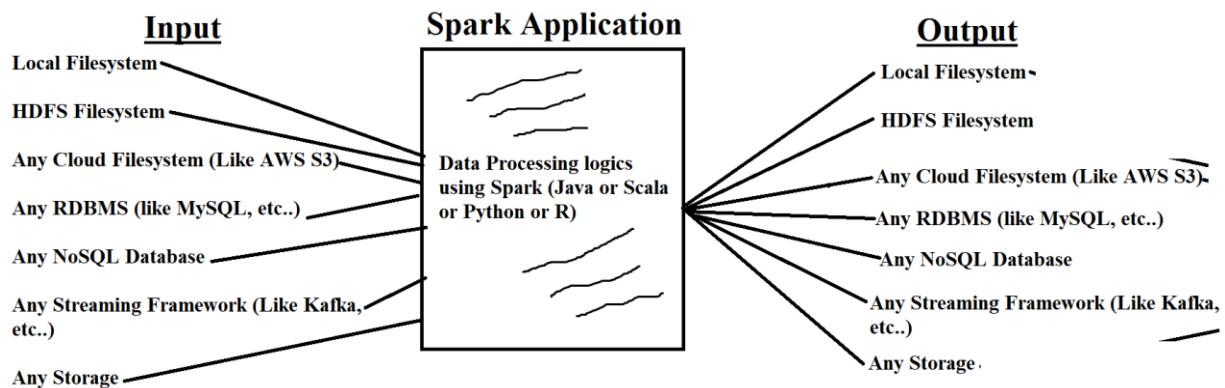Akkem Sreenivasulu
Data Engineering

PySpark Data Engineering with AWS

Data Engineering

"Data engineering refers to the building of systems to enable the collection and usage of data. This data is usually used to enable subsequent analysis and data science; which often involves machine learning.**"**

Apache Spark:-

→ Is a Data Processing Framework. We can implement Spark Applications using Java or Scala or Python or R Language.

→ It has only Data Processing. It does not have storage.

→ It can process any data.

Apache Data Engineering Components:-

1. Apache Hadoop – Storage (HDFS)
2. Apache Hive – Data Warehouse
3. Apache Spark (Java or Scala or Python (**PySpark)** or R (SparkR)) – Data Processing
4. Apache Kafka – Streaming Framework
5. Apache Cassandra, Apache HBase – NoSQL
6. Apache Oozie (less Usage) , Apache Airflow (Mostly Used) – Workflow, Schedule

Data Engineering Servers on,

On-Premise (Our Own Server) or **Any Cloud Computing (AWS or Azure or GCP)**

Now Every Cloud Vendor has Their Own Data Engineering Technologies.

AWS has its own Data Engineering Components.

PySpark Data Engineering with AWS

PySpark Data Engineering with AWS

Pre-Requisites

1. Python for Data Engineering
2. SQL for Data Engineering (SQL with Any RDBMS (MySQL, Oracle, PostgreSQL, M.S SQL Server, IBM DB2, etc..)
3. Linux Essentials for Data Engineering

PySpark Data Engineering with AWS

1. Data Engineering Pre-Requisites
2. AWS Services for Data Engineering
   AWS Foundation (AWS Intro, AWS Account Creation, Free-tier, AWS Global Infrastructure, etc..)
   AWS EC2
   AWS IAM
3. PySpark
   PySpark Foundation
   PySpark Core Programming
   PySpark SQL Programming
   PySpark Streaming Programming
4. AWS S3
5. AWS RDS
6. AWS EMR
7. AWS Glue
8. AWS Athena
9. AWS Kinesis
10. AWS Lambda
11. AWS Redshift
12. AWS Steps Functions
13. Apache Airflow
14. End-to-End Flow Application

Duration: 45 to 50Days (Daily 1:30Mins)
Fee : 20K/-- Offer Price 15K
Monday : 19th December, 2022 -- 7PM IST & 9PM IST
Contact : +91-9133161144
Monday to Friday
Optional: Sat
Sunday – No Class
Material + Workspace + Videos Access from GDrive

Website: www.cfamilyit.com   **eMail: info@cfamilyit.com**

PySpark Data Engineering with AWS

By

Akkem Sreenivasulu – Founder of CFAMILY IT Solutions
Website: www.cfamilyit.com
eMail: info@cfamilyit.com
Contact/Msg/WhatsApp: +91-9133151144, +91-9133161144
Telegram Channel: https://t.me/cfamilyit
YouTube: www.youtube.com/@cfamilyit
Facebook: https://www.facebook.com/cfamilyit/
Instagram: https://www.instagram.com/cfamilyit/
LinkedIn: https://www.linkedin.com/in/cfamilyit
Twitter: https://twitter.com/cfamilyit
GitHub: https://github.com/cfamilyit/

GCP (Google) Data Engineering:-

GCP Basic Services
GCP Regions & Zones
GCP Services Introduction
IAM
Provisioning Compute Engine
Deploy Application on App Engine
etc..

GCP Cloud Storage
Google Cloud Storage
Cloud Block Storage
etc..

GCP Database
Google Cloud SQL
Cloud Filestore
Cloud Datastore
Cloud Memorystore
Cloud Bigtable
etc..

GCP Data Processing
Google Cloud Pub/Sub
Google Bigquery
Cloud DataFlow
Cloud DataProc
Cloud Data Fusion
etc..

PySpark Data Engineering with AWS

AWS Data Engineering

1. Introduction
2. Python for Data Engineering
3. Linux Essentials for Data Engineering
4. SQL for Data Engineering
5. AWS Services for Data Engineering
6. AWS S3
7. PySpark Cluster Setup on AWS
8. PySpark Development
   PySpark Core Programming
   PySpark SQL Programming
9. AWS Kenesis
10. PySpark Streaming
11. AWS Redshift
12. PySpark Integrations
    PySpark Integration with AWS S3
    PySpark Integration with AWS Kenesis
    PySpark Integration with AWS Redshift
13. Databricks with AWS Introduction
14. AWS Glue Introduction
15. AWS Athena Introduction
    Etc..

By

Akkem Sreenivasulu – Founder of CFAMILY IT Solutions

Website: www.cfamilyit.com

eMail: info@cfamilyit.com

Contact/Msg/WhatsApp: +91-9133151144, +91-9133161144

Telegram Channel: https://t.me/cfamilyit

PySpark Data Engineering with AWS

YouTube: www.youtube.com/@cfamilyit

Facebook: https://www.facebook.com/cfamilyit/

Instagram: https://www.instagram.com/cfamilyit/

LinkedIn: https://www.linkedin.com/in/cfamilyit

Twitter: https://twitter.com/cfamilyit

GitHub: https://github.com/cfamilyit/