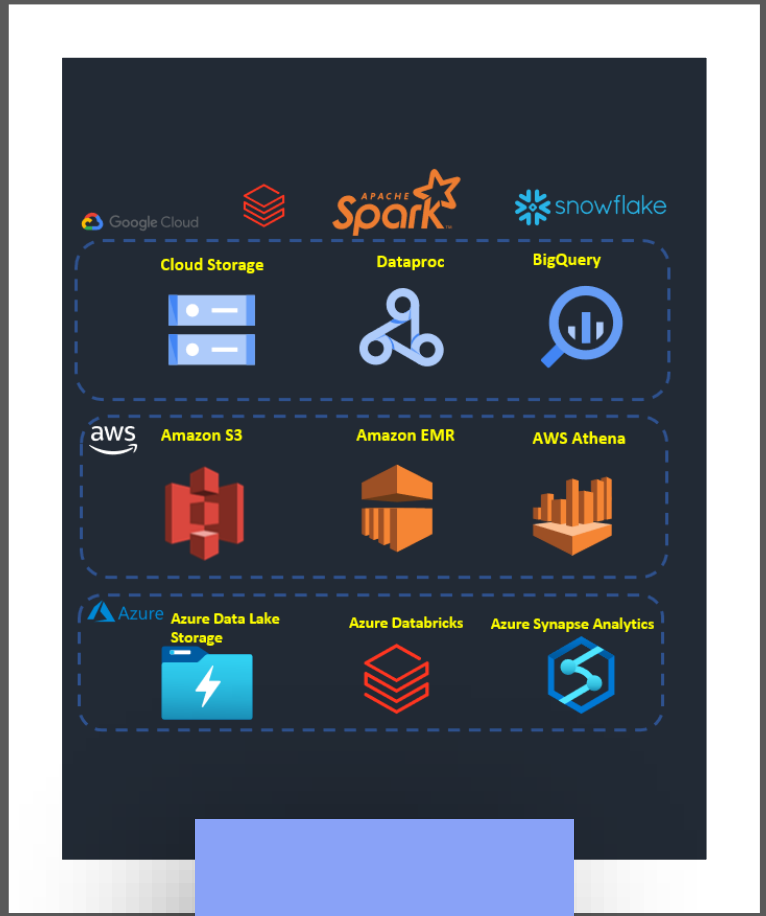
Accelerating Data Engineering Solution on Cloud

- Online Courses
- Consultancy

www.sankir.com



SanKir Technologies – Expertise in Data Engineering on Cloud

www.sankir.com

SanKir's Self-paced Online Courses

Data Engineering on Cloud

pro-Spark-aws

pro-Spark-gcp

Snowflake with pro-Spark

pro-Spark on aws

pro-Spark on gcp

Databricks on Cloud

Learn Databricks on AWS

Learn Databricks on GCP

- Quiz
- Assignment
- Bonus Lessons

For Experienced IT professionals

Contact : info@sankir.com

About SanKir

SanKir's expertise - Data Engineering on Cloud

As per Industry Trends, million+ Job opportunities are going to be created in Data Engineering on Cloud. SanKir's objective is to solve Corporate's pain point of finding the right resources for Data projects.

SanKir offers Self-paced online courses on Data Engineering on AWS and GCP Cloud platforms.
Courses involve popular platforms like AWS EMR, Dataproc, BigQuery, Databricks and Snowflake.

How can we help Corporates ?

- Exclusive batches for corporate at competitive price
- Provision for customization of course content based on corporate's immediate requirements and roadmap
- Hands-on sessions with project code
- Continuous assessment with Quiz and Assignments
- Expedite the learning process through Weekly contact sessions
- Turn around time to make a batch billable in Data Engineering projects is 4 to 6 weeks
- Digital Certificate is provided at the end of the Course

❑ Course on Data Engineering on Cloud

pro-Spark

Self-paced online course on Data Engineering

- Opportunity to experience a Data Engineering Project on Cloud with hands-on approach
- Learn Distributed Computing using Apache Spark
- Create Data Pipeline on Cloud

Download Project Code

Test, Build and Deploy Spark Application

Become proficient in Data Engineering on Cloud

pro-Spark Features

- Cloud of your Choice – GCP or AWS
- Dataset of your Choice
- Customize Business Validations
- Download the pro-Spark project code

Get Started with FREE Access to the Lesson 1 of Course

Why Data Engineering is important ?

As per LinkedIn's 2021 Emerging Jobs Report – Data Engineer Job has annual growth of 35 %

Gartner Inc - Worldwide end-user spending on public Cloud services is forecast to grow 18.4% in 2021 to total $304.9 billion, up from $257.5 billion in 2020

❑ Data Engineering on Cloud

Data Pipeline and Data Engineering

- Includes actions like Data Integration, Data Ingestion, Data Orchestration and Data Transformation
- Process of Designing Data Pipeline is Data Engineering
- Four Concepts that are necessary to become proficient in Data Engineering are
    - Data warehousing and ETL
    - Cloud computing
    - Cloud storage
    - Distributed computing

Cloud and Data Lake

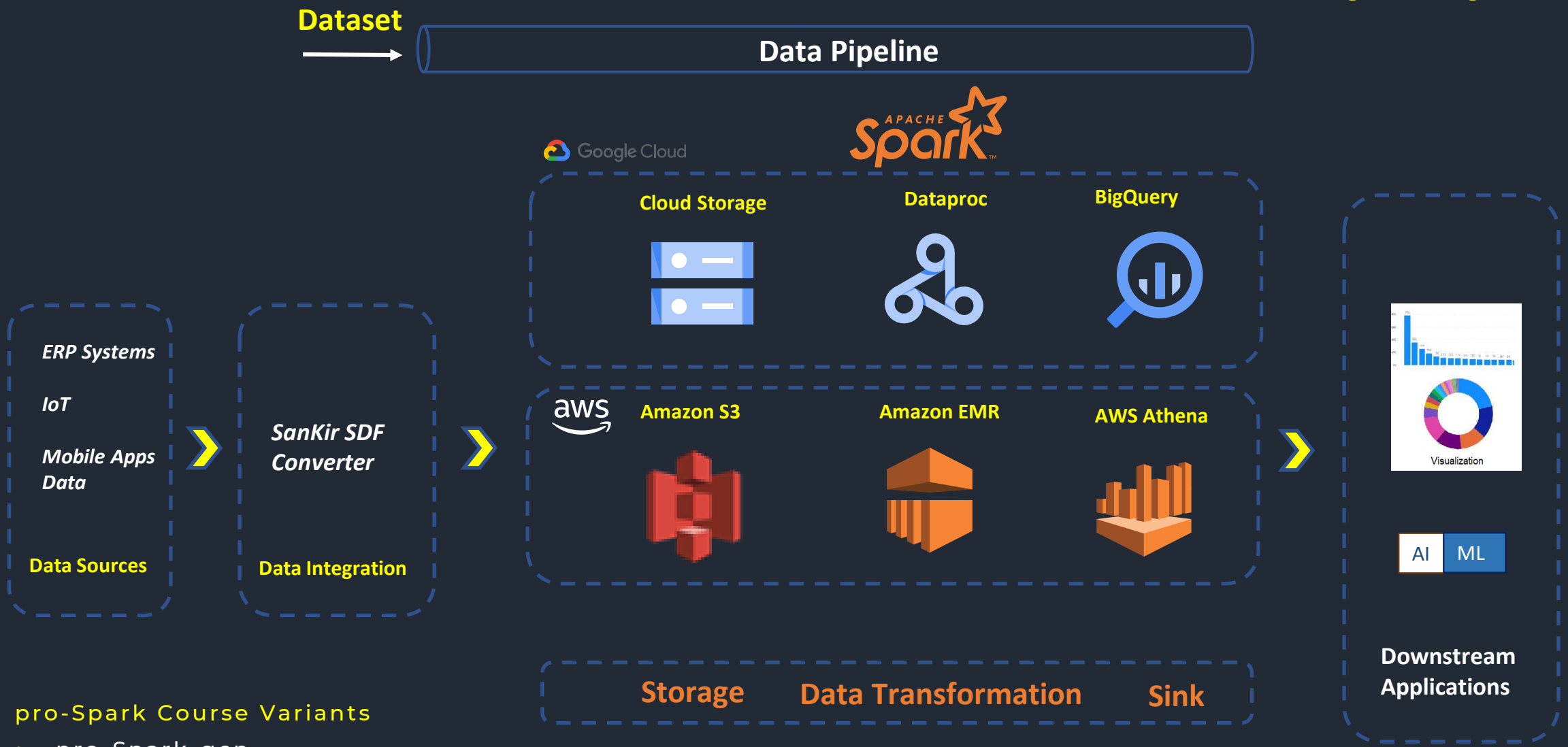
- Cloud allows Scale-out or Scale-in, or, Scale-up or Scale-down, based on the business demand
- Cloud computing is on-demand and scalable provisioning of computer resources
- Cloud Storage is Data Storage as Service and has revolutionized data storage and access
- Popular Data Lakes choices are
    - Google Cloud Storage
    - AWS S3

Distributed Computing

- Parallelism at scale
- Linking multiple computer servers over a network into a cluster
- Framework needed to coordinate work across the cluster
- Most popular Distributed Computing Framework – Apache Spark

- In-memory computation
- Agnostic to storage system
- Suitable for Performing ETL or SQL batch jobs with large data sets
- Supported distributed storage systems
    - Google storage
    - Amazon S3
    - HDFS

❑ Course on Data Engineering on Cloud



pro-Spark Course Variants

- pro-Spark-gcp
- pro-Spark-aws

pro-Spark on Cloud : End-to-End Data Pipeline

- A Data Engineering project can have multiple disparate Data sources coming from ERP, IoT and Mobile Apps.
- Next stage is Data Integration. SanKir SDF converter which acts like Data Integration tool, loads the Data into Cloud Storage such as Google Storage or Amazon S3.
- pro-Spark Application code running on Spark cluster reads the data and performs the Data Transformation. Transformed data is now written into the sink. Sink could be either a Database table or a storage such as Amazon S3.
- This transformed data is the input for any downstream AI/ML applications or consumed by Business Intelligence.
- This whole process refers to creation of Data pipeline on Cloud.

pro-Spark-AWS/GCP Course Details

❑ Course on Data Engineering on Cloud

Self-paced online course on Data Engineering on Cloud

10
Lessons

40
Topics

10
Quizzes

10
Assignments

Course Duration

Total of 8 hours of videos
8 weekends

Course Fee

13,990 INR
185 USD
140 Pounds
170 Euro / 250 AUD / 250 SGD

Course Creators – 54+ years of Industry Experience

Kiran Hiremath

Sanjay Bheemasenarao

❑ Course on Data Engineering on Cloud

pro-Spark Lessons

FREE Access

01 Introduction

02 Lifecycle of Data Engineering Project

03 Project Use Case

04 Software Installation

05 Learn Scala

06 Learn Spark

07 Cloud Platform AWS/GCP

08 Code Walkthrough

09 Build and Deploy

10 Conclusion

pro-Spark on

Google Cloud

aws

What you will be doing in pro-Spark–AWS/GCP Course

❑ Course on Data Engineering on Cloud

- Learn the nuances of Traditional Data warehousing and Cloud Data warehouse
- Explore Retail Data set in detail and define Key Performance Indicators (KPI)
- Design and Implement Data Validation and Transformation using Apache Spark

pro-Spark-AWS

- Develop connectors to Amazon S3, AWS EMR
- Design and implement EC2 and RBAC using IAM
- Design Data Lake with Amazon S3
- Create AWS EMR Spark Cluster
- Query External Tables using Amazon Athena

pro-Spark-GCP

- Develop connectors to Google Cloud Storage, BigQuery, Dataproc
- Design and implement RBAC using Service Account creation in IAM
- Design Data Lake with Google Cloud Storage
- Create Google Dataproc Spark Cluster
- Query results in BigQuery Tables

What you will be after the Course ? – Proficient Data Engineer on Cloud

❑ Course on Snowflake with pro-Spark

snowflake-pro-Spark-aws/gcp

- Add-on course to pro-Spark course
- Snowflake course is bought with pro-Spark course
  - Ex: When you buy snowflake-pro-Spark-aws, you are automatically subscribed to pro-Spark-aws
- Integrate pro-Spark Application with Snowflake
  - Learn to setup snowflake cloud data warehouse to store transformed data for downstream applications such as AI/ML

Subscription to snowflake-pro-Spark course gives access to

- pro-Spark-aws/gcp course content with 40 topics and code
- snowflake-pro-Spark-aws/gcp course content with code

16%
Datawarehousing Market Share

Cloud native Data Platform

SaaS product that enables data storage, processing, and analytic solutions that are faster and easier to use

Runs completely on Cloud Infrastructure

Deployed and managed entirely on a selected cloud platform

Build powerful and efficient Data Pipeline

Supported Cloud Platforms
- AWS
- GCP

❑ Course on Snowflake with pro-Spark

What you will be doing in Snowflake with pro-Spark–AWS/GCP Course

- Complete pro-Spark-aws Course
- Create Snowflake account for AWS/GCP
  - Setup Database, Schema, Virtual warehouse in snowflake account
- Download snowflake-pro-Spark-aws/gcp Project Code
- Integrate pro-Spark-aws/gcp Application with Snowflake
  - Code changes in configuration file, cloud connector section of pro-Spark code to connect to snowflake account

Course Duration

8 Hours of pro-Spark and 1 hour of snowflake
12 Lessons – 42 Videos
10 weekends

Course Fee

pro-Spark course : 15,990 INR
snowflake course : 3,000 INR
combo discount : 3,000 INR
Offer Price : 15,990 INR

International Currency

210 USD
160 Pounds
190 Euro / 285 AUD / 285 SGD

pro-Spark-aws/gcp course – 10 Lessons

snowflake-pro-Spark-aws/gcp course – 2 Lessons

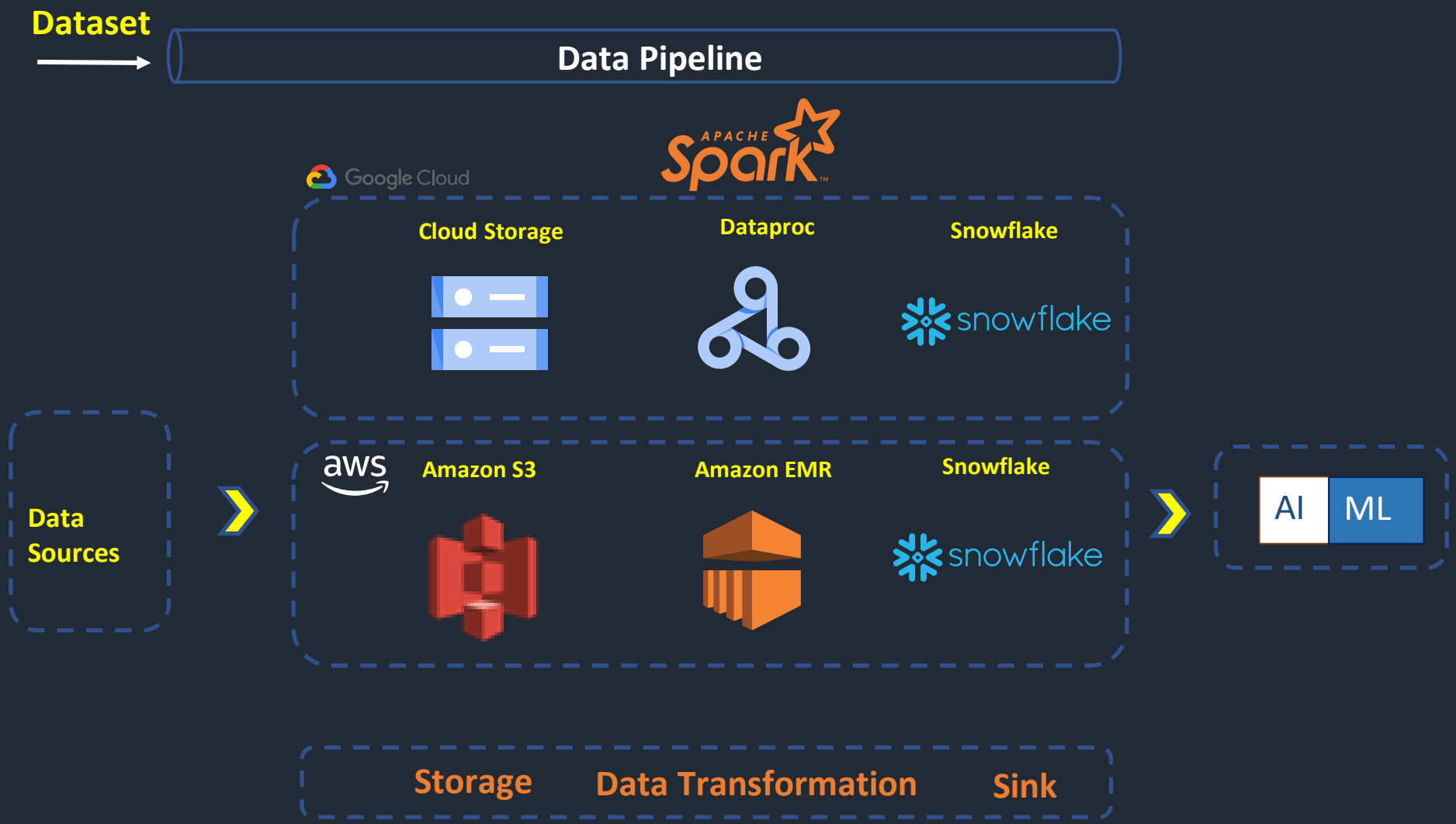
01
- Introduction
- Creation of snowflake account
- Setup Database and Warehouse

02
- Integrate pro-Spark-Application with snowflake

Course Creators – 54+ years of Industry Experience

Kiran Hiremath

Sanjay Bheemasenarao



End-to-End Data Pipeline

Snowflake with pro-Spark

How to take snowflake-pro-Spark-aws/gcp course ?

❑ Course on Snowflake with pro-Spark

Steps

01.

Execute all the topics of pro-Spark-aws/gcp course

- Download pro-Spark-aws/gcp Project code from github
- Execute the Project code with Retail dataset
- Query the Transformed data in KPI tables

02.

Integrate pro-Spark-aws/gcp Application with Snowflake

- git clone snowflake-pro-Spark-aws/gcp code and configure in IntelliJ
- Make code changes to connect to snowflake
- Test the snowflake-pro-Spark-aws/gcp locally
- Write transformed data into Snowflake tables
- Query the KPI Tables in Snowflake directly

❑ Course on Databricks on Cloud

Learn Databricks - AWS/GCP

- Highly Optimized platform from the creators of Apache Spark
- Create Apache Spark Cluster which is easy to scale
- Build and streamline ETL pipeline with ease

What you will be doing in Learn Databricks-AWS/GCP Course

- Create AWS/GCP Cloud Account and Databricks Account
- Create Databricks Workspace
- Create a Apache Spark Cluster choosing right configurations
- Upload the retail dataset on to Storage
- Create a Notebook
- Transform the data using Scala/Pyspark
- Build and streamline ETL pipeline with ease

databricks

Leader in Cloud Data warehouse

Collaborative environment for data teams to build solutions together

Per Gartner's Magic Quadrant – Leader in Cloud Database Management Systems

Provides Interactive notebook to use Apache Spark

Data engineers can Ingest, Query, Process and Transform data

Scalability: Gartner Peer Insights reviewers have praised the databricks solution's scalability. Some customers have processed petabytes of data

❑ Course on Databricks on Cloud

Learn Databricks – AWS/GCP Course

Course Duration

1 hour of videos
2 weekends

Course Fee

2,990 INR
40 USD
30 Pounds
35 Euro / 55 AUD / 55 SGD

Learn Databricks – AWS/GCP Course – 2 Lessons

01
- Introduction
- Creation of AWS/GCP account

02
- Creation of Databricks account
- Create a Apache Spark Cluster
- Create/Import Retail Notebook
- Read the data from Cloud, Transform the data using Scala/Pyspark

Course Creators – 54+ years of Industry Experience

Kiran Hiremath

Sanjay Bheemasenarao

SanKir's Self-paced Online Courses

How to Subscribe ?

- Visit https://www.sankir.com
- Choose your Course
- Pay Online (provide discount coupon if any)
- Start taking the course

Prerequisites for the course

- No prior Programming experience is required
- Laptop with min 8 GB RAM

Bonus Lessons ● Quizzes ● Assignments

Bi-Weekly Live Tech sessions

Digital Certificate

Representational Image

Reach us at info@sankir.com

https://www.facebook.com/sankirtech

https://www.linkedin.com/company/sankir

https://www.youtube.com/channel/UC7a7WWT9kiKwjWMSZDT2DQg/videos

SanKir Team

Kiran Hiremath, Director

IT Professional with 27+ years of Experience.

Expert in Data Engineering, Cloud Services and Distributed Computing using Apache Spark

- Data Engineering Pipeline – Architecture, Orchestration, Optimization and Monitoring
- Spark Cluster sizing – EMR, Dataproc
- Cloud Storage and Datawarehouse – AWS S3, Athena, GCS, BigQuery

Experience in Pre-Sales, CoE, Alliance, Software Development & Management

Developing Solutions in AWS and GCP Cloud Services

Sanjay Bheemasenrao, Director

27+ years of experience in building products and services

Technology focus areas – AWS and GCP, Data Lake and Cloud Data warehouses like Snowflake and Databricks

Experience in Product Development

Expert in Data Engineering covering all the Data Management aspects

- Data Engineering Pipeline – Architecture, Orchestration, Optimization and Monitoring
- Big Data Technologies: Hadoop, HDFS, Spark, Scala, Python, Hive
- Architecting Distributed Computing solutions using Apache Spark, Hive, BigQuery and AWS Athena.

Master Mind Group

Industry experts in distributed computing and storage

Professional Review Team

Professionals with functional and technical knowledge

www.sankir.com

Self-paced Online Courses – Pricing

	Course Name	List Price	Offer Price
01.	Data engineering on AWS : pro-Spark-aws	15990	13990
02.	Data engineering on AWS : pro-Spark-gcp	15990	13990
03.	Snowflake on AWS : snowflake-pro-Spark-aws	18990	15990
04.	Snowflake on GCP : snowflake-pro-Spark-gcp	18990	15990
05.	Data bricks on AWS	3990	2990
06.	Data bricks on GCP	3990	2990

Thank you!

www.sankir.com

Contact : info@sankir.com