

NITHISH KUMAR A.K

AWS DATA ENGINEER

 9629752346 |  nithish216216@gmail.com

PROFILE SUMMARY

AWS Data Engineer with 2 years of experience delivering scalable cloud-based data engineering solutions for investment banking and financial services environments. Strong expertise in SQL-driven development, Python programming, and PySpark-based distributed data processing using Apache Spark. Hands-on experience with AWS services including S3, Glue, Lambda, EC2, IAM, Redshift, and CloudWatch. Proven experience in building reporting data marts using Amazon Redshift and enterprise data warehousing using Snowflake. Skilled in Apache Airflow for workflow orchestration, data validation, schema design, and performance optimization.

CORE SKILLS

SQL, Advanced SQL, Python, PySpark, Apache Spark, AWS S3, AWS Glue, AWS Lambda, AWS EC2, AWS IAM, Amazon Redshift, Snowflake Data Warehouse, Apache Airflow, ETL, ELT, Data Pipelines, Data Warehousing, Reporting Data Marts, Data Modeling, Investment Banking Data, Financial Data Processing

PROFESSIONAL EXPERIENCE

- Designed and developed scalable AWS data pipelines for investment banking reporting platforms
- Built SQL-centric ETL workflows using AWS Glue and PySpark
- Implemented Amazon S3 data lake architecture for raw and curated datasets
- Developed Python and PySpark transformations for cleansing and enrichment
- Created reporting data models in Amazon Redshift
- Orchestrated workflows using Apache Airflow
- Collaborated with analytics and compliance teams

PROJECT 1:

Company: Ernest & Young

Enterprise Investment Banking Reporting Data Platform on AWS

Role: AWS Data Engineer responsible for SQL-based ETL development, PySpark transformations, Redshift reporting models, and Airflow orchestration for investment banking datasets.

Designed and implemented a cloud-native reporting data platform for investment banking datasets using AWS services.

The solution focused on processing structured financial data from upstream systems and transforming it into analytics-ready formats.

Implemented ETL pipelines using AWS Glue and PySpark to perform data cleansing, normalization, and enrichment.

Utilized Amazon S3 as a centralized data lake and Amazon Redshift for building reporting data marts consumed by analytics and compliance teams.

Integrated Apache Airflow for workflow orchestration, scheduling, and dependency management. Ensured data quality, schema consistency, and query performance optimization to meet enterprise reporting requirements.

PROJECT 2:

Company: Ernest & Young

Legacy Investment Banking Data Warehouse Migration to AWS & Snowflake

Role: AWS Data Engineer responsible for legacy data migration, Snowflake schema design, cloud data ingestion, validation frameworks, and post-migration performance optimization.

Executed migration of legacy on-premise investment banking data warehouse workloads to AWS cloud and Snowflake. Focused on data extraction, schema conversion, and cloud-native storage strategies.

Designed Python-based ingestion pipelines to move historical and incremental datasets into Amazon S3.

Implemented Snowflake as the target enterprise data warehouse to support scalable analytics and reporting workloads.

Performed data reconciliation, validation, and performance tuning to ensure migration accuracy and stability. Supported post-migration optimization and operational handover.

EDUCATION

Bachelor of Engineering – Electronics and Communication Engineering

St. Joseph's College of Engineering | 2019 – 2023

Higher Secondary Education – SRV boys Higher Secondary School | 2019

Secondary School (SSLC) – St. Anne's Matric Higher Secondary School | 2017

TOOLS & METHODOLOGIES

Git, Version Control, Agile Development, Data Quality Validation, Schema Enforcement, Cloud Security Best Practices.