

GAUTAM PRAKASH

Bengaluru

+91 7204609155

gautamprakash17@gmail.com

SUMMARY

Accomplished Data Engineer with a robust background in leveraging Python, SQL, Spark, and ETL techniques to enhance data management and analytics. With over two years of experience, I have demonstrated expertise in Snowflake, Databricks, and AWS cloud services, making significant contributions to data migration projects, outlier detection systems, and data model optimizations. My proficiency extends to developing end-to-end ETL pipelines using Apache Spark.

EXPERIENCE

Data Engineer, Elint Data

January 2023 -Present

- Implemented multiple data pipeline that gather data from multiple sources like API, SFTP server, RDS and performed transformation and aggregation and finally, it will create excel reports and send to the client.
- Data Extraction, aggregations and consolidation of the data within AWS Glue using PySpark.
- Developed ETL workflow involving multiple crawlers, jobs and triggers using AWS GlueWorkflow.
- Design and Developed ETL process in AWS glue to migrate data from external source like S3, Parquet/txt Files into snowflake.
- Implemented Incremental Data Processing using AWS SQS for ETL Data in AWS.
- Migrated data warehouse from RDS to Snowflake, enhancing data storage structures for better performance, scalability, and cost-efficiency.
- Bulk loading and unloading data into Snowflake tables using copy command.
- Created DWH, Databases schemas, Tables, write SQL queries against snowflake.
- Integrated and automated data workloads to snowflake warehouse.
- Implemented CI/CD pipelines using GIT and CodeCommit.

Software Engineer, Bharat Electronics Limited

October 10, 2021-August 15, 2022

- Designed, deployed, and supported an automated, scalable ETL pipeline using Python for Batch Processing. Integrated diverse data sources into the data warehouse, ensuring efficiency and maintaining robustness throughout the data life cycle.
- Used AWS glue catalog with crawler to get the data from S3 and perform SQL query operations using AWS Athena.
- Implemented Data quality check tools using Lambda, step function and SNS.
- Migrated on Premises ETL process to AWS cloud.
- Responsible for analyzing and data cleaning using PySpark and SQL queries.
- Exploring with spark various modules of Spark and working with dataframes, RDD and Spark Context.
- Database design and development including tables, Primary and Foreign key, Indexes and Stored Procedure.

EDUCATION

Bachelor of Engineering (Computer Science and Engineering)

SDM College Of Engineering and Technology

SKILLS

Languages: Python, SQL, PySpark

Database: PostgreSQL, MS SQL, MongoDB, Snowflake

Big Data Technologies- Spark

Control System- GIT, Jira

AWS Cloud Solutions: Lambda, S3, Glue, EC2, CloudWatch, Athena, RDS