GAUTAM PRAKASH

Bengaluru

+91 7204609155 gautamprakash17[@gmail.com](mailto:roh8raj@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 Glue Workflow.
* Design and Developed ETL process in AWS glue to migrate data from external source like S3, Parquet/txt Files into snowflake.
* Migrated data warehouse from RDS to Snowflake, enhancing data storage structures for better performance, scalability, and cost-efficiency.

# **Software** Engineer,**Bharat Electronics Limited**

October 10,2021-August15,2022

* 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.
* Created Delta tables on the existing parquet and delta formatted files.
* Converted Data Lake history to Delta Lake.
* Ingest existing data sources using autoloader in Databricks.
* Created target Delta tables and transformation code in Databricks: bronze, silver, gold.
* Implemented Delta live tables pipeline in Databricks.

# 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, Databricks

**Control System**-GIT, Jira

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