Naman Gupta

guptanaman0555@gmail.com | +91 8007670555 | GitHub | LinkedIn

Summary

Data Engineer with more than **5 years** of experience in **Data Modeling, ETL (Extract, Transform, Load)**, **Data Warehousing**, **Data Governance**, and **Data Security**. Delivered 10+ scalable, cloud-native solutions using Python, PySpark, SQL, AWS, Databricks, and Apache Spark, which enabled enterprise analytics and accelerated decision-making. Expert in **CI/CD automation** and performance tuning, achieving up to **80% memory reduction** and **10x faster deployments**. Skilled in building **batch & streaming pipelines** using **Apache Kafka**, **Apache Airflow** and **Serverless Architectures**.

Certifications

- AWS Certified Developer Associate
- Databricks Data Engineer Associate
- ITIL IT Service Management

Technical Skills

Programming Languages: Python, PySpark, SQL, YAML, Shell Scripting
Big Data & Cloud Technologies: Databricks, Apache Spark, Delta Lake, Unity Catalog, AWS Glue, AWS Lambda, AWS S3, Apache Kafka, Apache Airflow
Data Warehousing & Modeling: AWS Lakehouse, Amazon Redshift, Star Schema, Snowflake Schema
Databases: Amazon RDS, PostgreSQL, Amazon Redshift, Athena
Data Integration & Sources: CSV, Excel, Parquet, Delta Tables, Flat Files, SAP ECC & IBP, C3ME (Oracle)
Pipelines: Streaming pipelines, Batch pipelines
DevOps & Tools: GitHub, GitHub Actions, Docker, Postman, JIRA, CI/CD Pipelines

Work Experience

Deloitte USI   *Feb 2022 – Present*
Data Engineer

- Architected & delivered **scalable, secure data pipelines** & reusable ETL frameworks for structured, semi-structured, unstructured data using AWS S3, AWS Lambda, PySpark, Databricks, RDS (Aurora PostgreSQL) in AWS S3 Lake environments.
- Directed design of **batch & real-time processing** with Apache Spark & Databricks, enabling enterprise analytics and data-driven insights.
- **Automated infrastructure provisioning with parameterized YAML-based IaC scripts**, handling **table creation**, **job creation**, **task based cluster setup**, and **Job trigger setups**, with branch-based approval strategy for PRs & Github Action based Deployments, reducing code promotion time from **2 weeks to 30 minutes**.
- Changed from complex SQL queries to optimized PySpark scripts, reducing memory usage by **80%**.
- Led PoV evaluations for AWS analytics services, influencing **cloud modernization** strategies.
- Mentored developers, enforced code quality, and ensured Agile delivery excellence.

NICE Systems   *Aug 2020 – Jan 2022*
Associate Engineer

- Automated troubleshooting & issue resolution on remote servers using Python scripts, reducing manual intervention by **40%**.
- Built automation for SQL queries & real-time reporting, improving speed & accuracy.
- Streamlined workflows to meet SLAs & reduce operational load.

Projects

- **Full-Stack Book Reader Web App**
- **Excel Automation Tool**
- **Real-Time Chat Web App**
- **Spotify API Integration Tool**

Awards

- **Outstanding Performance Award** – Led **SAP migration** with zero downtime.
- **Spot Award** – Delivered **critical development** under tight deadlines.
- **Applause Award (x2)** – Managed **two concurrent projects** and drove **Unity Catalog migration** with reusable authentication scripts.

Education

B.E. (Electronics & Communication Engineering)
International Institute of Information Technology, Pune
CGPA: 8.41   2016 – 2020