Kaustav Kumar Das

Senior Data Engineer
Bengaluru, KA, India
+91 8906358974
kaustavkumardas1992@outlook.com
LinkedIn Github

Professional Summary

Senior Data Engineer with 10+ years experience architecting and building scalable ELT/ETL pipelines(using dbt,Apache Spark(pyspark),AWS Glue,Informatica Powercenter,IBM Datastage), building and optimizing data warehousing solutions(using Snowflake,Databricks,Oracle), and ensuring data reliability and data quality. Skilled in distributed systems, orchestration frameworks, and cloud data platforms. Strong analytical and problem-solving abilities, with proven success in reducing costs and improving performance. Adept at clear documentation and collaborative teamwork in Agile environments.

Work Experience

Senior Data Engineer

Smith and Nephew, Pune, India
06/2023 – Present
Project: Data Analytics and Enterprise Architecture

- Involved with Product Owners,Business Engagement Leads,BRMs from discovery phase of a data project to planning and designing all the tools/tech stacks/methods required for data ingestion,storage and organization,transformation and processing,validation and quality,serving and consumption,governance, security and monitoring. Also helped with data platform governance.
- Designed and built ingestion pipelines, converting full loads to incremental models, saving **80% of Snowflake cost**. Additionally designed custom warehousing solution which saved another **40% of Snowflake cost** in return helping to enhance refresh rate of warehouse from once a day to thrice a day with same budget.

- Developed and maintained **300+ automated ELT pipelines** using **dbt, AWS S3, Apache Spark(pyspark), AWS Glue, AWS Lambda,AWS Step Functions, and Snowflake/Databricks.**

- Built **Unified Data Platform (UDP)**, centralized and homogenized source of truth to serve enterprise-wide data needs.

- Helped in data migration of older data warehouse solutions to the Snowflake data warehouse.
- Optimized queries and pipelines, improving performance and reliability.

- Documented **30+ data sources, flows, and infrastructure** for knowledge sharing.

- Automated deployments with Python scripts, GitLab CI/CD, and Snowflake scripting.

- Standardized release management using **Git, Confluence, Jira, and DevOps practices.**

- Ensured data quality through reconciliation, auditing, and rigorous testing.

- Led team of data engineers, business analysts having size of **20(mix of FTE and vendors)** to deliver on data products.

Data Engineer

IBM India Pvt Ltd, Pune, India
04/2021 – 06/2023
Project: USAA CFO Treasury Team

- Led a **5-member team** to develop and maintain **20+ ETL/ELT pipelines** using IBM Datastage, dbt, and Snowflake.

- Integrated data from **10+ sources** (Oracle, Snowflake, CSV, JSON, APIs).

- Built Tableau dashboards for CFO Treasury reporting (Collateral, Bank Product Reporting).

- Migrated Datastage from v11.5 to v11.7.

- Automated credential management with Python (CyberArk) and enhanced pipelines with Unix shell scripts.

- Managed Control-M jobs, Git repositories, and deployments via Urban Code Deploy.

- Created test cases in Qtest and conducted code reviews in SmartBear.

- Optimized ETL workflows, improving stability and performance.

ETL Developer

Tata Consultancy Services, Kolkata, India
10/2015 – 04/2021
Projects: PepsiCo LATAM & Mattel (USA)

- Developed ETL processes using **IBM DataStage (8.5, 11.5, 11.7)** and **Informatica Power-Center (9.1, 10.2).**

- Integrated SAP systems with legacy applications for data warehouse loading.

- Monitored and tuned DataStage/Informatica jobs for performance.

- Automated workflows with Control-M, Windows Task Scheduler, and RPA tools (Athus, Power Automate).

- Built dashboards with **Power BI and IBM Cognos**.

- Led offshore teams in Agile environments (PSM1 certified).

- Migrated Informatica jobs from v9.1 to v10.2.

- Designed and documented ETL architecture, mappings, and technical specifications.

- Implemented multiple process improvements, reducing operational overhead.

Education

Bachelor of Technology
West Bengal University of Technology – Asansol
2011 – 2015

Certifications

- SnowPro Associate: Platform Certification (2026)
- AWS Certified Cloud Practitioner (2022)

- Microsoft Certified: Azure Data Engineer Associate (2021)

- Microsoft Certified: Azure Administrator Associate (2021)

- Microsoft Certified: Data Analyst Associate (2021)

- Microsoft Certified: Azure Fundamentals & Data Fundamentals (2021)

- IBM Certified Advocate – Cloud v1 (2021)

- Salesforce Certified Administrator (2020)

- Professional Scrum Master I (2020)

Technical Skills

- **Data Platforms:** Snowflake, dbt, AWS Glue, Apache Spark(pyspark), Databricks, IBM Datastage, Informatica PowerCenter, TIBCO BW

- **Programming & Tools:** Python, SQL, Git, Control-M, Urban Code Deploy, Globalscape EFT

- **Visualization:** Tableau, Power BI, Cognos

- **Cloud & Architecture:** AWS (S3, Lambda, Step Functions), Apache Airflow.

- **Other:** Data Modeling, Data Warehousing, Orchestration Frameworks

---

Languages

- English

Extracurricular

- Singing: Featured in international, national, and regional channels.

Disclaimer

I hereby acknowledge that the information furnished above is true to the best of my knowledge. Open to relocation opportunities.