

Pratik Patra

Data Engineer | +91 9776569069 | pratikpatra09@gmail.com | [LinkedIn](#)

Professional Summary

Data Engineer specializing in AWS and Big Data, with expertise in Spark, Snowflake, and Python-based ETL automation. Experienced in cloud migrations, cost optimization, and CI/CD pipeline integration to deliver high-performance, production-ready solutions.

Experience

Business Technology Solutions Consultant – ZS [10th Nov 2025 – Present]

Project 1 | Industry: Healthcare | Role: Senior Data Engineer

- Building a FHIR healthcare data platform to convert data into FHIR-compliant standards using the Medallion (Bronze–Silver–Gold) architecture.
- Using ADF, Databricks, Kafka, PySpark, Python, and SQL to ingest and process data from SQL Server, flat files, JSON, and Kafka-based streaming sources.

Software Engineer III - HashedIn by Deloitte [10th Jan 2022 – 31st Oct 2025]

Project 1 | Industry: Capital Markets | Role: Data Engineer

- Led a 5-engineer squad to modernize Cost Basis statement generation; designed cloud architecture and delivered production rollout replacing legacy batch jobs. (Tech: AWS Glue, Lambda, Redshift, IAM, S3, CloudFormation).
- Built automated Glue/Spark pipelines with cross-account access, SNS/SQS alerts and error handling, supporting continuous data ingestion and reducing end-to-end manual efforts by ~600hours/year.
- Engineered Python comparison tools for legacy vs cloud outputs, reducing manual validation by ~90% and accelerating production sign-off.
- Proposed and implemented decoupling of Lambda from Glue to optimize job runtimes and reduce pipeline cost; tracked cost improvement and reliability gains.
- Delivered 10+ production-grade enhancements in cost basis jobs, including custom filtration logic, custodian comment handling, and mutual fund validation - dramatically increasing accuracy and transparency.
- Integrated SonarQube and CI/CD checks into pipeline deployments to enforce code quality and reduce post-deployment defects.
- Ensured flawless production deployments by coordinating with IAM, Prod Support, Tech Approvers, and external teams, while handling network-level Glue connector issues under pressure.

Project 2 | Industry: Health Care | Role: Data Engineer

- Migrated approximately 300 tables from Hadoop to Snowflake to enhance data accessibility, scalability, and security for research advancements.
- Developed ETL pipelines using **Python, Spark, and AWS services (Lambda, Glue, Step Functions)** for **data extraction, transformation, and loading**.
- Implemented tokenization and encryption functions for PHI/PII data to adhere to client data security policies.
- Utilized **Python scripts** for generating tokenized files and partitioned **SQL scripts**, reducing manual efforts by 90%.
- Collaborated with senior stakeholders to resolve critical data migration blockers, enabling on-time delivery of healthcare research pipelines.

Project 3 | Industry: Health Care | Role: Data Engineer

- Created **SSIS** packages for data extraction, transformation, and loading, reducing processing time by 30%.
- Enhanced existing features and incorporated new features for claims settlement solutions.
- Developed complex **T-SQL** queries and stored procedures for efficient data retrieval and manipulation.
- Designed and maintained data models for improved organization and accessibility.
- Implemented error handling and logging mechanisms to ensure data integrity.
- Conducted data quality assessments and cleansing processes to enhance data accuracy.

Release (DevOps) Engineer - ADP [16th Aug 2021 – 5th Jan 2022]

Project 1 | Industry: HCM | Role: Software Engineer

- Optimized database rollover process, saving 50% of resources for client training.
- Improved reverse proxy for faster routing and loading mapping files.
- Managed product owner environments daily.
- Worked on build & release processes and modules for new client onboarding and user management.

Systems Engineer (MSBI Developer) - Infosys [30th Dec 2019 – 11th Aug 2021]

Project 1 | Industry: Energy | Role: Software Engineer

- Developed ETL packages in **SSIS** for legacy data conversion automation.
- Prepared **SQL** queries and stored procedures for Business Intelligence tasks.
- Configured and deployed parameterized **SSRS & Power BI** reports for data visualization.
- Identified bottlenecks in **SSIS** solutions, tuned **SQL** queries, and provided post-deployment support to clients.
- Administered and enhanced **Power BI & SSRS** reports as per user requests.

Technical Skills

- **Big Data & Processing:** Spark, PySpark, ETL/ELT Pipelines, Data Migration, Data Modeling (Dimensional & Relational)
- **Cloud & Data Platforms:** AWS (Glue, Lambda, Step Functions, S3, Redshift, CloudFormation, IAM, RDS), Snowflake, Hadoop
- **Programming & Scripting:** Python, SQL Performance Tuning/Optimization, Shell Scripting
- **Databases:** Oracle, MS SQL Server, PostgreSQL
- **Data Integration & ETL Tools:** SSIS, AWS Glue, REST API Integration
- **DevOps & CI/CD:** CloudFormation, Git, Jenkins, SonarQube, Docker, GitHub Actions
- **Security & Compliance:** PHI/PII Tokenization, Data Encryption, IAM
- **Tools & Others:** Postman, JIRA, Agile/Scrum

Education

- **B.Tech in Electrical Engineering,** IGIT Sarang (BPUT Odisha) — 2019 | CGPA: 7.65/10

- **Higher Secondary (Science – PCM)**, U.N. College (CHSE Odisha) — 2015 | 71%
- **Secondary School**, G.B. High School (BSE Odisha) — 2013 | 81.17%