

PARVATHY S

Data Engineer

+91 8281519598, +91 9778728794

parumulavana25@gmail.com, Kochi, Kerala

PROFESSIONAL SUMMARY

AWS Data Engineer with 3.5+ years of experience designing and optimizing scalable data pipelines in cloud-native environments. Skilled in ETL/ELT development using PySpark, pandas, and SQL, with expertise in AWS Glue, Redshift, S3, Lambda and data warehouse design. Proven track record of reducing ETL processing times by 20% and enhancing query performance by 25%, enabling faster and more reliable analytics. Passionate about building efficient, automated data workflows that drive business insights and operational excellence.

SKILLS

- Programming Languages: SQL, Python, PySpark
- Cloud Technologies: AWS Glue, AWS Redshift, AWS S3, AWS Lambda, AWS IAM, AWS Athena, Dynamo Db, ECS Fargate, AWS Step Function, AWS Event bridge.
- Data Engineering: ETL/ELT Processes, Data Warehouse Design, Big Data, Data Lake Creation, CI/CD pipelines
- Tools & Frameworks: Databricks, Jupyter Notebook, PyCharm, VS Code, Jira, SVN, Git/Github, Terraform
- Database Technologies: PostgreSQL, MySQL
- Soft skills: Problem Solving, Analytical Thinking, Adaptability, Collaboration, Time Management

EXPERIENCE

Associate Data Engineer

Cognizant Technology, Cochin (June 2025 – Present)

- Designed and developed scalable data ingestion pipelines using PySpark on AWS Glue, transforming raw input files from multiple providers into standardised 4000-byte formatted files for mainframe processing.
- Engineered and deployed AWS Lambda functions to detect file arrivals in S3 buckets and trigger corresponding Glue jobs, enabling automated and event-driven data processing.
- Migrated Lambda-based orchestration to Amazon ECS Fargate for handling large file volumes, improving performance and reliability of the pipeline.
- Implemented DynamoDB as a metadata store to track rule IDs, provider information, and processing configurations for auditability and dynamic rule management.
- Developed a file audit framework using AWS Lambda, EventBridge and SNS to automatically detect unprocessed files in S3 input buckets and notify stakeholders, ensuring timely remediation and improved pipeline reliability.
- Optimized AWS Glue jobs to reduce processing time for a 42 GB input file from 20 minutes to just 5 minutes, achieving a 75% improvement in performance and significantly enhancing pipeline efficiency.

Data Engineer

IBS Software, Trivandrum (July 2022 – May 2025)

- Analyzed design documents provided by clients to understand requirements and specifications for ETL processes
- Developed robust ETL, code using PySpark within AWS Glue environment, adhering to best practices and design principles
- Implemented and managed end-to-end data pipelines from AWS S3 to Redshift, ensuring data security, integrity, and scalability.
- Designed and optimized over 30 AWS Glue jobs, reducing ETL processing times by 20% and improving resource utilization by 25%.
- Built and maintained scalable data lakes and data warehouses supporting 15+ critical business analytics workflows.
- Enhanced SQL query performance, resulting in a 25% faster execution time for key analytics operations
- Authored 20+ comprehensive design documents for data warehouse and ETL workflows to ensure clear communication across teams.

CLIENT-SIDE EXPERIENCE(JAPAN)

- Collaborated with the client team for 3 months during the User Acceptance Testing (UAT) phase in **Japan**.
- Ensured the quality and accuracy of deliverables by conducting hands-on testing and issue resolution in real-time.
- Gained in-depth understanding of client requirements and provided on-the-spot solutions, enhancing user satisfaction.
- Coordinated between offshore and onshore teams to streamline testing and feedback processes.

CERTIFICATIONS

- **Databricks Certified Data Engineer Associate** – Databricks.
- **AWS Certified Cloud Practitioner** – Amazon Web Services.

ACHIEVEMENTS

- Received customer-focused appreciation for proactively addressing client needs, improving project outcomes, and delivering solutions aligned with business objectives.
- Team Champ Award: Recognized for completing all assigned tasks within deadlines and resolving complex business challenges to drive project success.
- Successfully reduced ETL pipeline errors by 40% through optimized coding practices.
- Improved data processing speed by 30% by leveraging PySpark and AWS Glue optimizations.

EDUCATION

APJ Abdul Kalam Technological University

Bachelor of Technology in Computer Science and Engineering

CGPA: 8.7

Relevant courses:

- Machine Learning
- Data Mining and Warehousing
- Data Science
- Soft Computing
- Probability Distributions and Numerical Methods
- Computer Programming

LANGUAGE SKILLS

Proficient in English (exceptional reading, excellent writing, good speaking), Hindi (excellent reading/writing, good speaking), and Malayalam (exceptional reading/writing, excellent speaking).