Prabhakar Reddy Chalamala

Azure Data Engineer | PySpark | Databricks | GenAI | ADF

Hyderabad, India | +91 6281646106 | reddyprabhakar303@gmail.com |
linkedin.com/in/prabhakar-reddy-chalamala-34923a232

Professional Summary

Results-oriented **Azure Data Engineer** with 3+ years of experience designing, developing, and optimizing **ETL pipelines and data workflows** using **Azure Databricks, PySpark, ADF, and SQL**. Skilled in **Medallion Architecture, data governance, automation, and GenAI tools** including GitHub Copilot, Databricks Assistant, and Microsoft 365 Copilot. Proven ability to deliver scalable, SLA-compliant, and defect-free data solutions.

Core Skills

• Languages & Frameworks: Python, SQL, PySpark, Spark SQL, Shell Scripting
• Cloud & Big Data: Azure Databricks, Delta Lake, Azure Data Lake Gen2, Azure Synapse Analytics, Unity Catalog
• ETL & Orchestration: Azure Data Factory (ADF), Databricks Workflows, AutoSys
• Data Governance: Schema Validation, Audit Logging, Data Lineage, Monitoring, Unit Testing
• GenAI & Productivity Tools: GitHub Copilot, Databricks Assistant, Microsoft 365 Copilot, Databricks Genie (Research)
• DevOps & Version Control: Git, GitHub, Azure DevOps (CI/CD)
• Collaboration & Tools: Jira, ServiceNow, iServe (Agile, Incident & Change Management)

Soft Skills

• Team Leadership: Led and managed a team of 5 members, ensuring timely delivery and quality output
• Collaboration: Worked closely with cross-functional teams and stakeholders to align project goals
• Ownership: Took full responsibility for project end-to-end delivery with minimal supervision
• Problem Solving: Proactively identified and resolved pipeline bottlenecks and data issues
• Communication: Effectively communicated technical concepts to non-technical stakeholders

Professional Experience

Infinite Computer Solutions – Hyderabad, India
Software Engineer (Azure Data Engineer) | Oct 2022 – Present

• Designed and optimized scalable ETL pipelines using PySpark and Azure Databricks for healthcare data processing through Bronze, Silver, and Gold layers.
• Built and automated workflows using Azure Data Factory (ADF) and AutoSys, reducing manual effort by 30%.
• Implemented data governance frameworks covering schema validation, audit logging, lineage tracking, monitoring, and automated unit testing.
• Migrated legacy SQL Server workloads to Azure Lakehouse architecture with Delta Lake and Unity Catalog.
• Integrated GitHub Copilot and Databricks Assistant to accelerate development of Spark jobs and SQL queries.
• Researched and evaluated Databricks Genie for AI-assisted SQL generation, notebook creation, and data exploration use cases.

• Used Microsoft 365 Copilot to automate documentation, reporting, and technical writing tasks.
• Delivered defect-free, SLA-compliant solutions with strong client recognition.

Key Projects

Chronic Condition Reporting Framework (Healthcare)
• Developed ETL pipelines in Azure Databricks to analyze member history and identify missing chronic condition records.
• Consolidated ICD code mapping and validation into reusable frameworks, reducing maintenance overhead.
• Delivered Gold layer fact tables and CSV datasets for BI reporting with automated execution via Databricks Workflows and AutoSys.
• Ensured data quality through unit testing, schema validation, lineage, and audit logging with alerting.

Monthly Data Curation Process (Healthcare)
• Created and managed ADF pipelines for ingesting data from multiple sources into Azure Blob Storage (Bronze layer).
• Designed Databricks pipelines to process data through Bronze → Silver → Gold layers for BI readiness.
• Monitored and optimized pipelines for performance and SLA adherence, reducing manual monitoring effort by 40%.

GenAI Enablement & Research
• Leveraged GitHub Copilot for accelerating development of PySpark pipelines, automation scripts, and test cases.
• Explored Databricks Genie capabilities for AI-powered SQL generation and data exploration.
• Utilized Microsoft 365 Copilot to automate meeting note summarization and documentation.

Achievements

• Performer of the Quarter (2024) – Infinite Spark Award
• Performer of the Month (2024) – Infinite Spot Award
• Improved pipeline runtime by 20% through Spark job optimizations
• Reduced manual monitoring effort by 40% through workflow automation
• Consistently delivered 100% compliant, defect-free production releases
• Recognized for independent project ownership and client satisfaction

Education

• Bachelor of Education (Mathematics), 2021
• Bachelor of Science (Computer Science), 2018

Certifications

• Databricks Certified Data Engineer Professional – 2025