

Gokulraj G

Azure Data Engineer | Databricks | PySpark | Azure Data Factory

Email: XXXXX | Phone: XXXXX | LinkedIn: XXXXX

Professional Summary

Azure Data Engineer with 3+ years of experience in designing, building, and maintaining scalable data pipelines using Azure Databricks, PySpark, Azure Data Factory, and ADLS Gen2. Strong expertise in implementing Medallion architecture (Bronze, Silver, Gold) for large-scale data processing in the logistics domain. Hands-on experience working with multiple source systems including SQL Server, Cosmos DB, Salesforce (JSON), SFTP, and Google BigQuery. Proven ability to build incremental and full-load pipelines, manage data quality, and enable analytics consumption for downstream applications.

Technical Skills

- Azure Databricks (PySpark, Spark SQL, Delta Lake, Delta Live Tables, Auto Loader)
- Azure Data Factory (ADF Pipelines, Data Flows, Parameterization, ForEach, CI/CD)
- Azure Storage (ADLS Gen2, Blob Storage)
- Databases: Azure Cosmos DB, Azure PostgreSQL, SQL Server, Oracle
- Programming: Python, PySpark, SQL, PL/SQL
- Data Architecture: Medallion Architecture, Fact & Dimension Modeling, Incremental Processing
- Version Control & CI/CD: Azure DevOps, Azure Repos, Git

Professional Experience

Azure Data Engineer – Cognizant Technology Solutions

Client: UPS | Domain: Logistics | May 2025 – Present

- Designed and implemented end-to-end data pipelines using Azure Databricks and Azure Data Factory following the Bronze, Silver, and Gold architecture.
- Ingested JSON and Parquet data from multiple upstream systems into Bronze layer using Databricks Auto Loader and Delta Live Tables in streaming mode.
- Developed Silver layer transformations using incremental high-watermark logic to populate fact and dimension tables with historical tracking.
- Built Gold layer datasets by stitching order data across TMS and upstream systems to support analytics and reporting use cases.
- Integrated Azure Cosmos DB as an analytics store and supported downstream consumption through Azure AI Search by creating data sources and indexes.
- Developed ADF pipelines to ingest data from SQL Server, Salesforce, SFTP, and Google BigQuery using parameterized and parallel processing techniques.

- Implemented data masking workflows by reading and updating sensitive data within Cosmos DB using Azure Data Factory.
- Collaborated with frontend and analytics teams to ensure data availability, consistency, and performance for business-critical dashboards.

Software Engineer – HCL Technologies

Dec 2021 – Apr 2025

- Worked extensively on Oracle SQL and PL/SQL development, writing complex queries, procedures, and functions.
- Improved query performance by up to 90% through indexing strategies, query rewrites, and execution plan analysis.
- Built and maintained Azure Data Factory pipelines for data ingestion and transformation from relational sources.
- Supported data engineering workloads by assisting with Databricks-based transformations and Azure storage integrations.
- Gained strong foundation in data modeling, ETL concepts, and performance optimization techniques.

Certifications

- Databricks Data Engineer Associate – September 2025

Education

- Bachelor's Degree – Computer Science / Related Discipline

Awards & Recognition

- Star Performer Award – HCL Technologies