

Data Engineer with 2+ years of experience designing and optimizing scalable ETL pipelines using Azure Data Factory, Databricks, and PySpark. Proficient in SQL, Python, and Big Data technologies with a track record of success in improving data quality, automating workflows, and enhancing performance. Focused on transforming complex datasets into meaningful insights that power strategic business decisions.

EXPERIENCE

9thPixel Technologies India Pvt.Ltd	03/2021 - 06/2023
Data Engineer	Hyderabad, India
<ul style="list-style-type: none">Designed and implemented scalable ETL pipelines in Azure Data Factory (ADF), automating data integration across multiple sources and improving data consistency, reducing manual data handling by 20%.Accelerated data processing speeds by 35% by optimizing ADF activities, including copy, lookup, and metadata processes, and integrating stored procedures, ensuring reliable data flow.Developed reusable, parameterized ADF pipelines for scalable workflows, cutting future development time by 40%.Built PySpark transformations in Databricks to streamline large data processing and accelerate BI reporting.Collaborated with stakeholders to deliver real-time analytics solutions, driving a 5% increase in revenue by supporting data-driven decision-making.Spearheaded the migration of legacy data systems to Azure Synapse, resulting in a 25% improvement in query performance for downstream analytics teams.Engineered logging and error-handling in ADF pipelines, improving monitoring and cutting resolution time by 50%.Led team training on PySpark and Azure best practices, boosting productivity and maintainability.	

PROJECTS

Databricks End-to-End ETL Pipeline Databricks, PySpark, Delta Lake, DLT, ADLS, SQL Personal Project	Git Hub
<ul style="list-style-type: none">Implemented dimensional data models, governing customer and product information using SCD Type 1 & 2, and enhanced The FactOrders table, leading to a 20% improvement in data retrieval speeds.Orchestrated ingestion to transformation using Databricks Workflows and parameterized notebooks.Fashioned core data warehouse components, applying SCD Type 1 & 2 to DimCustomer and DimProduct, and built FactOrders tables, completing 100% of assigned data modeling tasks ahead of schedule.Processed 100K+ records daily using Delta Live Tables (DLT) for real-time and incremental ETL.Transformed raw data into a clean, deduplicated, and surrogate-keyed data warehouse, accelerating analytics and BI reporting cycles by 40% and enabling faster decision-making across departments.	

TECHNICAL SKILLS

- Big Data & ETL Tools:** Databricks, Apache Spark (DataFrame, Spark SQL), Apache Hadoop (HDFS), Snowflake, Apache Airflow (Basic), Azure Data Factory, Unity Catalog
- Cloud Platforms:** Microsoft Azure (ADF, Databricks, Synapse, Blob Storage), AWS Basic (S3), GCP Basic (Big Query)
- Programming & Scripting:** Python (Pandas, NumPy, PySpark), SQL (Advanced), T-SQL
- Data Modeling & Warehousing:** Dimensional Modeling, Star Schema, Snowflake Schema, Azure Synapse, Data Warehousing
- Data Pipelines & Integration:** ETL/ELT workflows using ADF, PySpark, SQL; Real-time and batch processing
- CI/CD & Version Control:** Git, GitHub Actions (Basic), Azure DevOps (Basic), SDLC Understanding
- Databases:** MySQL, SQL Server, Azure SQL Database, NoSQL (Familiar)
- Data Visualization:** Power BI (Basic), Looker (Familiar), MS Excel (Advanced)
- Collaboration:** Cross-functional Communication, Agile (Scrum), Training, Documentation

EDUCATION

Florida Atlantic University	10/2023 - 05/2025
Candidate for Master of Science in Information Technology Management	
GPA: 4.0/4.0 Dean’s List	

CERTIFICATIONS

Big Data Analytics, Florida Atlantic University	05/2025
MS Excel - Advanced, Cengage	04/2024