Pranjal Yadav

Delhi, India | pranjalyadav9721@gmail.com | 8840517240
LinkedIn: pranjal-yadav-33800623b

Professional Summary

Azure Data Engineer with experience building and optimizing cloud-based ETL pipelines using Azure Data Factory, Azure Databricks, and Delta Lake. Skilled in PySpark, Python, and SQL for scalable transformation, validation, and deduplication. Experienced with incremental ingestion, SCD Type 1 & 2, Auto Loader, Delta Live Tables, and Medallion Architecture. Worked with Databricks Lakeflow Jobs to orchestrate multi-notebook workflows and pass values between notebooks for automated pipelines. Strong in data ingestion, data validation, and data transformation.

Technical Skills

Cloud & Data Engineering: Azure Data Factory, Azure Databricks, ADLS Gen2, Azure SQL Database, Delta Live Tables, Lakeflow Jobs
Programming Languages: PySpark, Python, SQL
Version Control: Azure DevOps (Git Repository Management)

Experience

Data Engineer — Bx Data Solution Pvt. Ltd., Gurugram, Haryana (July 2025 - Present)
• Implemented a watermark-based incremental ingestion pipeline with a backdated refresh feature.
• Created pipelines that load only the latest file from ADLS Gen2 using Get Metadata.
• Worked with parameterized datasets to build dynamic and reusable ingestion pipelines.
• Worked with activities like Copy Data, Get Metadata, ForEach, If Condition, Lookup, Validate, and Script for orchestration.

Technical Intern — Bx Data Solution Pvt. Ltd., Gurugram, Haryana (January 2025 - June 2025)
• Learned and practiced SQL, Python, and PySpark for data processing and transformation.
• Gained foundational understanding of data pipeline development in cloud environments.

Projects

Automated Data Ingestion Pipelines using Azure Data Factory (2025) • Built ETL pipelines to ingest data from Azure SQL, REST APIs, and ADLS Gen2, on prem.
• Implemented backdated incremental ingestion using Lookup, Script, If Condition, and Copy Data.
• Automated file validation using Get Metadata, Filter, ForEach, and Copy Data.
• Designed pipeline using Medallion Architecture (Bronze → Silver → Gold).

Azure Databricks - Auto Loader, Delta Live Tables & Delta Lake Optimization (2025)
• Implemented ingestion using Auto Loader with schema evolution.
• Built SCD Type 1 & 2 using Delta Live Tables.
• Developed class-based PySpark modules for reusable transformations.
• Worked with Managed and External Delta tables.
• Used Delta Lake features: ACID, schema evolution, and time travel.

Education

Bachelor of Technology (B.Tech) in Electrical and Electronics Engineering (EEE)
ABES Engineering College, Ghaziabad, Uttar Pradesh (July 2021 - July 2025)

Certifications

• SQL Beginner to Advanced For Data Professionals | codebasics | March 2024
• Python: Beginner to Advanced For Data Professionals | codebasics | July 2024

Languages

English — Full Professional Proficiency
Hindi — Full Professional Proficiency