

Rohan Srivastwa,

Pune, India

Summary

Professional experience in Data Engineering/ETL working on Azure Tech Stack with expertise in PySpark, Python and SQL. Developed Data pipelines from various sources like SQL Server, Oracle DB, SFTP to ingest data into ADLS Gen2 based on Medallion Architecture.

Area of expertise

<ul style="list-style-type: none">• Azure Databricks (ADB)• PySpark• Spark SQL (SQL)• Python• Excel (advanced)	<ul style="list-style-type: none">• Azure Data Factory (ADF)• Azure Synapse• Azure Data Lake Storage (ADLS)• Azure SQL Database• Modern Data Warehouse• Snowflake	<ul style="list-style-type: none">• PowerBI• Delta tables• Azure Log Analytics• Azure DevOps• GitHub• Unity Catalog
---	--	--

Key Clients

1. AB-InBev

- Worked as data engineer for refactoring tables from legacy system to modern system based on Medallion architecture using Azure Databricks, Databricks workflows for orchestration.
- Created feature branches to develop code from GitHub repos, committed the changes and merged the PR using GitHub.
- Created silver and gold procedures using Databricks for assigned tables using PySpark, Spark SQL based on standard and user defined libraries.
- Validated completeness and equivalency check for each table migrated to Unity Catalog of databricks.
- Migrated tables to snowflake from Unity Catalog as an alternate Data Warehouse.
- **Tools used** – Azure Databricks, GitHub, Azure DevOps, VS code, PySpark, Python, Spark SQL

2. Kelly Services

- Worked on Data Ingestion Framework to ingest data from multiple sources (SQL Server, Oracle, SFTP, SAP) to ADLS Gen2 using code driven framework built on databricks.
- Created delta table for each table ingested to ADLS Gen2
- Worked on creating Data Quality Framework to ensure data ingested has no discrepancy.
- Implemented logging & monitoring in the code for auditing and performance monitoring.
- Email alert sending system in case of any code failure and a final data quality report.
- Orchestrated the pipeline using Databricks workflows.
- Used Azure DevOps to create PR to deploy code to production.
- **Tools used** – Azure Databricks, Azure Log analytics – App insights, Azure DevOps.

3. Rich's Product

- Worked as a data engineer to build pipelines to ingest data from API sources – Facebook Graph API to ADLS gen2 based on Medallion model – Bronze, Silver & Gold layer.
- Responsible for creating Solution architecture for Data Engineering workflow.
- Performed API testing for data using POSTMAN.
- To facilitate Data Science team with Solution Architecture and ad-hoc requirements related to data engineering. Ex- Databricks setup, Cluster configuration etc.
- Spearheaded Client calls for requirements and updates with Weekly Presentations
- **Tools used** – Azure Databricks, Azure Synapse analytics, Azure data factory.

4. Mark and Spencer

- Data Engineer at Mark and Spencer
- Designed, built, and maintained data pipelines for seamless data flow across departments.
- Developed and maintained a robust data warehouse architecture for efficient data storage and retrieval.
- Implemented ETL processes to extract, transform, and load data from diverse sources into the data warehouse.
- Ensured data quality and governance through data validation routines and error handling mechanisms.
- Monitored and optimized data systems and pipelines for enhanced performance and reduced downtime.

5. Bank of America

- Expertise in querying and analyzing large databases for financial insights, risk assessment, and compliance reporting.
- Skilled in data manipulation and transformation using Excel functions, PivotTables, and Macros.
- Experience in building financial models and conducting scenario analysis to evaluate investment opportunities and assess risk.
- Track record of creating automated Excel reports and dashboards to monitor performance metrics and support data-driven decision-making.
- Ability to identify cross-selling opportunities and provide data-driven insights for optimizing lending strategies and credit risk assessment.
- Familiarity with regulatory requirements and compliance reporting in the banking and finance industry.
- Strong attention to detail, accuracy, and the ability to present complex financial information in a clear and concise manner.
- Collaborative team player with excellent communication and problem-solving skills.

Education

- Bachelor of Engineering, SIT, Tumkur, Karnataka.

Certifications

- Azure DP 203 certified.
- Azure DP 900 certified.
- Associate Data Engineer - Databricks certified.