MAHENDRA KOLHE

Pune, India | mkolhe23@gmail.com | 7350458043 | [LinkedIn](#) | [GitHub](#)

SUMMARY

Results-driven Data Engineer with 3+ years of experience designing and implementing scalable data pipelines, automating **ETL** workflows, and enabling data-driven decision-making. Proficient in **SQL**, **Python**, **Azure Data Factory** (**ADF**), **Databricks**, and **Snowflake** with a strong foundation in real-time data processing, data quality, and performance optimization. Proven ability to deliver high-impact solutions in cloud and on-prem environments, with a strong focus on reliability, automation, and analytics.

TECHNICAL SKILLS

Programming & Scripting: SQL, Python, PySpark
Big Data Technologies: Apache Spark, Spark SQL, PySpark
Cloud & Data Platforms: Azure Data Factory (ADF), Azure Databricks, Snowflake, ADLS
ETL & Data Integration: SSIS, Custom ETL/ELT Pipelines, SQL Server Integration
DevOps & Automation: Workflow Orchestration, Job Scheduling, Data Reconciliation
Certifications: [Python(HackerRank)](#), [SQL Advanced(Udemy)](#), [SQL Advanced(HackerRank)](#)

WORK EXPERIENCE

Data Engineer, Xcaliber Infotech   **May 2023 – Present**

- Designed and implemented scalable ETL workflows using **Azure Data Factory** and **Databricks** to extract data from Azure SQL, SQL Server, and Analysis Services, staging it in **Azure Data Lake,** and loading it into **Snowflake**, reducing pipeline execution time by 40%.
- Built and maintained **real-time KPI dashboards** using SQL Stored Procedures and Tableau, improving decision-making speed and reducing report turnaround time by 70%.
- Developed a **Python-based polling service** for monitoring and triggering SQL Jobs, Python ETL, and ADLS pipelines, reducing manual monitoring and latency by 50%.
- Created and managed **Power BI dashboards** and implemented a unified data model using **Azure Analysis Services** to ensure reporting consistency across business units.
- Designed **Snowpipe**, Stored Procedures, and Views in **Snowflake** to ingest data from sources like S3/Blob Storage for reporting and analytics layers.
- Automated file validation and reconciliation processes for over **400K+ daily transactions**, achieving 99.9% accuracy and reducing data investigation time by 60% through a custom-built **data lineage dashboard**.

Software Developer, TeamLease   **Jan 2021 – Feb 2022**

- Improved HR data validation processes through dynamic scripting, enhancing overall reporting accuracy.
- Built **SSIS** pipelines to ingest flat files from email to **SQL Server**, improving operational efficiency by **30%**.
- Automated HR alerts using **SQL** stored procedures and triggers, cutting manual tracking by **50%**.
- Maintained **SQL** Server databases with >**99.5%** uptime, supporting **HRMS** data reliability.

PROJECTS

WalletWatcher - [Link](#)

- Built **WalletWatcher** using **Python** and **SQL** to automatically fetch transaction-related **emails**, parse financial details, and store data in a structured **SQL** database.
- Implemented **ETL** processes with Pandas for data **extraction**, **transformation**, and **loading**, and integrated **WhatsApp API** to send daily/monthly expense notifications as part of a prototype for a full **SaaS solution**.

EDUCATION

PG Diploma in Advanced Computing, IACSD, Pune   *, Sept 2022 - April 2023*
Bachelor of Technology (B.Tech.), Amravati University   *, July 2015 - June 2018*