

ADEEB KHAN

Data Engineer

Noida, IN | +91-8077837767 | hakhan208@gmail.com

PROFESSIONAL SUMMARY

Results-driven Data Engineer with 3+ years of experience, actively involved in 3 projects. Specializing in Databricks, Azure Data Factory, PySpark and SQL. Proven success in delivering impactful data solutions, optimizing project outcomes, and ensuring seamless integration.

WORK EXPERIENCE

Tata Consultancy Services

Noida, IN

Data Engineer – Azure

Jan 2022 - Present

- Developed data pipelines to ingest streaming/non-streaming data from sources like SAP HANA database, flat files, SharePoint, using APIs, etc. into **Azure Data Lake** ensuring data quality, reliability, and security.
- Utilizing Azure Logic Apps within **Azure Data Factory**, automated the data ingestion process, extracting data from diverse sources and sending out email notifications to stakeholders regarding its status.
- Building data models and transformation logic using **PySpark** and **Spark SQL** in **Databricks**. Also, created reporting views which are utilized by **visualization** applications like Tableau and Power BI.
- Working closely with cross-functional teams in an **agile** workspace, such as Data Scientists, Business Analysts, and Software Developers to understand their data needs and provide technical guidance.
- Monitoring data pipelines and workflows, identifying and resolving data-related issues, and **optimizing** performance and **efficiency** of data processes.
- Created Star Schema views in **Azure Databricks** for data enhancement and in **Snowflake** for analytics through Tableau and other BI tools which significantly reduced query execution time by **50%** on BI tools.
- Migrated data from SAP HANA to Azure Data Lake using Azure Data Factory and Databricks, involving the assessment of load frequency, data volume, and dependencies, while ensuring data quality and integrity throughout the process.
- Addressed database-related issues reported by end-users or clients, identifying root causes, diagnosing problems, and implementing solutions to ensure seamless database operations.

PROJECTS

Bayer Consumer Health Analytics Product

- Gathered raw data from various sources across multiple countries and refined it to extract meaningful insights.
- Generated a variety of KPIs to evaluate the performance of pharmaceutical products in both local and global markets.
- Leveraged a range of technologies, including Azure Cloud, Databricks Auto Loader (Cloud Files), Apache Spark, Python, SQL, Snowflake, and others.

SKILLS

- **Programming Languages** - C++, C#, Python, PySpark, SQL, Spark SQL, PL/SQL, JavaScript
- **Databases** - SAP HANA, SQL Server (SSMS), Snowflake, MySQL
- **Cloud** - Azure Data Factory, Azure Databricks, Azure Logic Apps, Azure SQL Database, Azure Data Lake
- **Data Warehousing** - Data Analysis, Data Modeling and Transformation, Data Ingestion and Integration
- **Other Tools/Technologies** - ETL, Hadoop, Spark, Kafka, AzureDevOps, JIRA, GitHub(CI/CD), SharePoint

CERTIFICATIONS

- **DP-203** Azure Data Engineer Associate
- **DP-900** Azure Data Fundamentals
- **DP-100** Azure Design and Implementation of Data Science Solution Training
- **AZ-900** Azure Fundamentals

EDUCATION

Amity University Noida, INDIA

B.Tech + M.Tech (Intg) (Aerospace Engineering) - GPA: 7.71 Jun 2016 - May 2021