

# Ajay Jayaraj

+91-8220937538 | [ajaysri2016@gmail.com](mailto:ajaysri2016@gmail.com) | [www.linkedin.com/in/ajayj2000/](https://www.linkedin.com/in/ajayj2000/) | [github.com/AjayAJ2000](https://github.com/AjayAJ2000) | [ajayaj2000.github.io/portfolio/](https://ajayaj2000.github.io/portfolio/)

## Professional Summary

- Microsoft Certified Azure Data Engineer and AI Engineer with **3 years of experience** in designing, developing, and implementing robust data solutions across various industries.
- Developed and maintained **near real-time data ingestion frameworks** for operational analytics using spark streaming and databricks.
- Led a cross-functional team in the successful implementation of **Agile methodologies**, resulting in a 40% increase in project delivery speed.
- Proven expertise in data engineering with firsthand experience in creating high-performance data pipelines, data transformation, and data integration using **Azure Data Factory, Azure Databricks, and PySpark**.
- Solid knowledge of **Data Warehousing and Dimensional Modeling** concepts, with the ability to build and optimize complex data models and data products.
- Strong proficiency in **Python**, and **Spark SQL**, including advanced query writing using CTEs, joins, grouping, aggregation, and nested subqueries for efficient data transformations.
- Proficient in Python programming for data processing, **ETL** development, and automation. Skilled in using libraries for efficient data manipulation and analysis. Strong experience with Python's OOP principles, functional programming, and error-handling mechanisms to build robust, reusable code.
- Skilled in building end-to-end ETL solutions, leveraging **cloud technologies** to automate processes and optimize workflows, contributing to a **67% reduction in manual tasks and improving query performance by 40%**.
- Experienced in implementing **Data Governance** practices and ensuring compliance through **data cleansing** and **data profiling** techniques with Azure Data Factory and Databricks.
- Expertise in using **Spark for large-scale data processing and optimization** of internal infrastructure, leading to improved data processing performance and scalability.
- Demonstrated **leadership skills** by mentoring and training junior associates, managing project timelines, coordinating with cross-functional teams, and directly interfacing with clients to ensure timely delivery and alignment with business needs.
- Collaborative team player with a strong focus on driving team success through **effective communication**, attention to detail, and a proactive approach to **problem-solving**.
- Highly motivated to continue growing and advancing in the field of **Data Engineering**, with a passion for continuous learning and applying new technologies to solve complex business challenges.

## EDUCATION

Bachelor of Technology | SASTRA UNIVERSITY

Jun 2018 - Apr 2022

## SKILLS

**Data Engineering:** Azure Cloud [ Azure Databricks, Azure Data Factory, Azure Data Lake Storage, Azure Blob Storage, Azure Functions, Azure Logic apps], Snowflake, Python, SQL, Apache Spark (Spark SQL)

**Data Analysis:** Root Cause Analysis, Data Warehousing, Data Modelling.

**AI/ML Tools:** Retrieved Augmented Generation, Prompt engineering, AI Engineering, N8N.

**Version Control/CI-CD:** GitHub, Git, Github Actions [Git Strategy]

**Project Management:** Agile, Jira

## CERTIFICATIONS

- Microsoft Certified Azure Data Engineer Associate - DP-203
- Microsoft Certified Azure AI Engineer Associate - AI-102
- Databricks Accredited Azure Platform Architect
- Gen AI and LLM on Databricks
- SnowPro Associate: Platform Certification

## EXPERIENCE

### ZS Associates

#### Senior Data Engineer

Apr 2025 - Present

- Created an **End-to-End Data & Platform Architecture** using Azure cloud and databricks for data ingestion and data processing for a critical time constraint project.

- Recognized for **developing scalable and reusable frameworks** that address recurring challenges in data engineering projects cutting down 40% of project turnaround time.
- Took on project management **responsibilities**, including creating business presentations, managing **delivery timelines**, and directly consulting with clients and the onshore team.
- Demonstrated leadership skills by **mentoring and training junior associates**, managing project timelines, coordinating with cross-functional teams, and directly **interfacing with clients** to ensure timely delivery and alignment with business needs.
- Collaborative **team player with a strong focus on driving team success** through project management methodologies like Agile, effective communication, attention to detail, and a proactive approach to problem-solving.

## Data Engineer

Sep 2022 - Apr 2025

- Developed and managed **automated data workflows** with Azure Databricks, enabling **validated data ingestion pipelines** from multiple [15+] sources to production environments ensuring scalable, secure data delivery via Azure Data Factory.
- Established and **enforced data governance and security** in Azure Databricks achieving 100% compliance with regulatory standards for PHI/HIPAA data.
- Architected automated data pipelines using Azure Data Factory to **ingest and transform** provider solutions and Pharma data, optimizing Snowflake for seamless processing.
- Programmed **complex and efficient SQL queries** to power dashboards tracking \$500M contracts, purchasing behaviors, and rebate metrics, **enhancing transparency and operational efficiency**.

---

## PROJECTS

### InspiroGene Product - Data & Platform Engineering

*Client: Fortune 500 Healthcare Provider in the US & Canada*

- **Architected and optimized** the core data pipeline and platform infrastructure powering the InspiroGene product's [Cell and Gene Therapies] to track the end-to-end workflow of patients undergoing gene therapy, presenting financial, logistics, and therapy-related data to Pharma customers via analytics dashboard, resulting in a 40% improvement in report generation speed and enabling near real-time insights for key stakeholders
- **Spearheaded data and platform engineering** for InspiroGene, architecting data pipelines and infrastructure for near real-time analytics. Drove performance improvements, ensured scalability, and collaborated across teams for seamless product integration performance for more than 5 service lines.

### Customer Self-Serve Portal

*Client: Fortune 500 Healthcare Provider in the US*

- Designed and implemented a Salesforce (SFDC) web portal for Pharma customers, providing financial insights into purchases, discounts, and rebates.
- Developed report tiles in SFDC, integrating data from Snowflake and SAP BW to drive business insights.
- Utilized Databricks, Snowflake, and ADF for data ingestion, orchestration, and reporting, leveraging Tableau for data visualization.
- Built eight high-performance data pipelines and complex Spark SQL queries for data transformation, supporting frameworks like Row-Level Security (RLS).
- Took on project management responsibilities, including creating business presentations, managing delivery timelines, and directly consulting with clients and the onshore team.

### Enterprise Financial Planning and Analysis - Data and Gen AI Engineering

*Client: Fortune 500 Healthcare Provider in the US & Canada*

- Designed and implemented the data infrastructure underpinning 4+ FP&A Data Projects, enabling accurate and timely reporting of critical **financial metrics**. Leveraged **SCDs** and time-based aggregations to provide comprehensive historical analysis and support **strategic decision-making for multiple projects**.
- Engineered a robust document chunking solution utilizing LLMs and Vector Databases, enabling a streamlined search and retrieval process for rebate details from GPO contracts; improved query response efficiency by 40%.

---

## AWARDS and ACHIEVEMENTS

- Lauded within cross-functional teams at ZS as the '**Flawless Finishers**' after implementing refined protocols which increased accuracy on deliverables up to 40%, reflecting superior discipline and focus on achieving industry-leading results throughout all phases of development.

---

## Hobbies

Writing Data Engineering Blogs, Attending Tech Events