

Sk Sayan Islam

sayanislam777 | 6290719800 | LinkedIn | GitHub

WORK EXPERIENCE

ZS Associates, Gurgaon - Decision Analytics Associate

Jul 2025 – Present

- Managed distributed data processing and operational analytics workflows for a US pharmaceutical client across 10+ recurring deliverables and reporting systems.
- Optimized ETL pipelines across 5+ datasets, reducing reporting runtime from 12+ hours to less than 1.5 hours while improving reliability and execution efficiency.
- Converted and modernized legacy Alteryx workflows to PySpark on Databricks, improving execution runtime by 75% and ensuring consistency through regression testing and validation.
- Implemented automated Databricks Workflows and orchestration pipelines with Delta Lake, reducing runtime from 12 hours to 4 hours while improving monitoring and delivery timelines.
- Designed QC and validation frameworks improving discrepancy detection by 40% and reducing reporting errors by 30%.
- Built and launched 3+ Agentic AI solutions (using Databricks GENIE) enabling self-serve analytics, automated dashboards, and insights from sales, call, and sampling data.
- Led UAT, troubleshooting, and issue-resolution activities for dashboards and pipelines, identifying root causes and ensuring reliable outputs.
- Automated SLA-driven reporting workflows using Excel macros and operational tooling, reducing turnaround time from 6 hours to 2 hours.
- Collaborated with global cross-functional stakeholders to support operational reporting, process improvements, and scalable analytics delivery.

ZS Associates, Gurgaon - Decision Analytics Associate Intern

Jan 2025 – Jun 2025

- Performed data validation and QC processes, reducing reporting errors by 30%.
- Collaborated with global teams to improve analytics workflows and delivery efficiency.
- Developed domain understanding of healthcare data and business processes.
- Conducted knowledge transfer sessions to support onboarding and scalability.

PROJECTS

HIVE Agent (HIV Insights and Visualization Engine)

Apr 2024 – Present

- Developed a supervisor AI agent coordinating multiple Databricks GENIE agents across Sales, Calls Execution, Workload, and Alignment Change domains, enabling scalable analytics across large distributed datasets.
- Reduced complex analysis turnaround time from multiple days to a few hours while decreasing ad-hoc analysis workload by 40% through automated QC validation and analytics workflows.

Legacy Pipeline Migration & Performance Optimization

Mar 2026 – May 2026

- Migrated legacy Alteryx workflows to PySpark-based Databricks architecture, modernizing distributed data processing pipelines and improving maintainability.
- Optimized execution using parallelization and skew-data handling, reducing runtime from 3+ hours to under 1 hour and saving 100+ engineering hours annually.

Automated Data Pipeline & Performance Optimization

Mar 2026 – Apr 2026

- Performed large-scale cohort-based analysis across HCP segments using sales, calls, and patient datasets to improve targeting effectiveness and operational insights.
- Increased targeting effectiveness by 20% through iterative experimentation, automated analysis workflows, and scalable validation processes across 5–8 cohorts.

TECHNICAL AND OTHER SKILLS

Programming Languages: Python, SQL, C++, JavaScript

Analytics: A/B Testing, KPI Tracking, Data Visualization, Business Insights, Quantitative Analysis

Tools: Databricks, SparkSQL, Alteryx, MS Excel, MS Powerpoint, Tableau, Streamlit, Git, JIRA

Data Engineering: ETL Pipelines, PySpark, AWS Athena, AWS S3, Data Processing, Automation

Reliability & Operations UAT, Regression Testing, Data Validation, QC Frameworks, Troubleshooting, Root Cause Analysis, SLA/TAT Management

Soft Skills: Stakeholder Management, Cross-functional Collaboration, Problem Solving

EDUCATION

Thapar University, Patiala, Punjab

Graduated Jul 2025

B.E. in Electrical and Computer Engineering

CGPA: 8.82

Relevant Coursework: Data Structures and Algorithms, Operating Systems, DBMS, Software Engineering, AI, ML, Cloud Computing