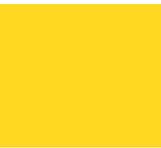


CASE STUDY

Achieved
17% Cost Savings
with **FinOps** on
Azure Databricks :

A Global
Pharmaceutical **GCC**
Transformation Story



Client Overview

The client is the **Global Capability Center (GCC) of a leading global pharmaceutical company**, responsible for managing enterprise data and analytics platforms that support research, manufacturing, regulatory, and commercial operations across multiple regions.

To enable these capabilities, the GCC operates a large-scale analytics ecosystem built on **Azure Databricks**, powering data engineering and analytics workloads across multiple business domains. The platform manages over **150,000 datasets** and more than **40,000 data pipelines**, supporting critical use cases ranging from clinical research analytics to supply chain optimization and commercial insights.

As the analytics ecosystem rapidly expanded, the growing scale of datasets, pipelines, and workloads introduced challenges related to **cloud cost visibility, governance, and operational efficiency**.

To address these challenges, the client partnered with **Polestar Analytics to implement a FinOps-driven governance framework on Azure Databricks**. The objective was to enable centralized monitoring of cloud costs, strengthen workload governance, improve data lineage visibility, and drive sustainable cost optimization across large-scale data engineering operations managed by the GCC.

Key Challenges

As the platform matured and adoption grew across teams and domains, the GCC began to face several operational and governance challenges in managing the expanding data ecosystem.

Limited Visibility into Cloud Spend

Annual Azure Databricks spend exceeded **USD 10 million (post discounts)**, but the GCC lacked granular visibility into costs across regions, business domains, and projects.

Rapid Data and Pipeline Proliferation

With **150K+ datasets and 40K+ pipelines**, analytics adoption grew rapidly. This led to duplication of datasets, scripts, and pipelines, increasing infrastructure costs and operational complexity.

Fragmented Cost Monitoring

Cost monitoring practices varied across teams, making it difficult to accurately attribute spend across resource groups, projects, and workloads.

Lack of End-to-End Data Lineage

Data moved across **Source Systems → Azure Data Factory → Databricks → Power BI**, but the absence of unified lineage limited traceability, impacting compliance, auditing, and troubleshooting.

Inconsistent Governance Practices

Workload deployment practices varied across teams, with no centralized governance framework to standardize dataset reuse, pipeline development, and cluster configurations.

Solutions Implemented

Polestar Analytics designed and implemented a centralized **FinOps and governance framework** to bring transparency, operational efficiency, and cost optimization to the Azure Databricks environment. The solution enabled the GCC to centrally monitor platform usage, enforce governance standards, and optimize workloads across the enterprise data ecosystem.

01

FinOps Dashboards and Cost Monitoring

- * Built custom dashboards to track Databricks spend across Resource Groups, Business Domains, Projects, and Workloads.
- * Enabled granular cost attribution to improve cost accountability across teams and domains.
- * Provided real-time visibility into consumption patterns to support proactive cost management.

02

End-to-End Data Lineage Framework

- * Developed a lineage tracker mapping the complete data flow across **Source Systems** → **Azure Data Factory** → **Databricks** → **Power BI**.
- * Enabled full traceability of data transformations across the analytics stack.
- * Improved transparency, troubleshooting, and audit readiness for regulatory compliance.

03

Governance Framework and Workload Optimization

- * Established standardized practices for workload deployment, dataset reuse, and pipeline development across the platform.
- * Identified underutilized clusters, optimized autoscaling configurations, and improved resource utilization.
- * Rationalized duplicate datasets and pipelines to reduce redundancy and improve operational efficiency.

04

FinOps Adoption and Enablement

- * Conducted enablement workshops for Data Engineering and BI teams across the organization.
- * Introduced best practices for monitoring and managing cloud costs.
- * Promoted a FinOps culture to encourage cost-aware development and platform usage.



Business Impact

✔ **17%** **reduction in Azure Databricks spend** through cluster and workload optimization.

✔ **100%** **visibility into cloud costs** across domains, projects, and workloads.

✔ **150K+** **datasets streamlined** by eliminating redundant pipelines and scripts.

✔ **End-to-end lineage implemented** from Source Systems → ADF → Databricks → Power BI.

✔ **40K+** **pipelines governed** through standardized deployment practices.

✔ **Foundation built for predictive FinOps** and proactive cloud cost management.

About

Polestar Analytics

is a leader in **Data, Analytics, AI, and Enterprise Planning** helping organizations to unlock intelligent outcomes through our proprietary products like **IPlatform, accelerators, and services**. Our expertise spans data engineering, data science, agentic and generative **AI**, and advanced planning for **CPG/Retail, Pharmaceuticals, Manufacturing, IT/ITeS, and Financial Services**.

Reach out to us today!