Deliver better customer experiences & more efficient supply chain with Databricks

Omnichannel consumer trends have accelerated, with global e-commerce sales increasing their share of sales in 10 weeks what had previously taken 10 years to achieve.

Retailers need to be quickly respond to fast-changing consumer behaviors, leveraging new data and analytic approaches to build a more resilient supply chain.

To survive and thrive, retailers will need to break down data silos, and derive actionable insights from structured, semi-structured, and unstructured data.

Databricks is your Unified Data Analytics Platform for big data and analytics from data ingest to data query.

Leading innovators powered by Databricks

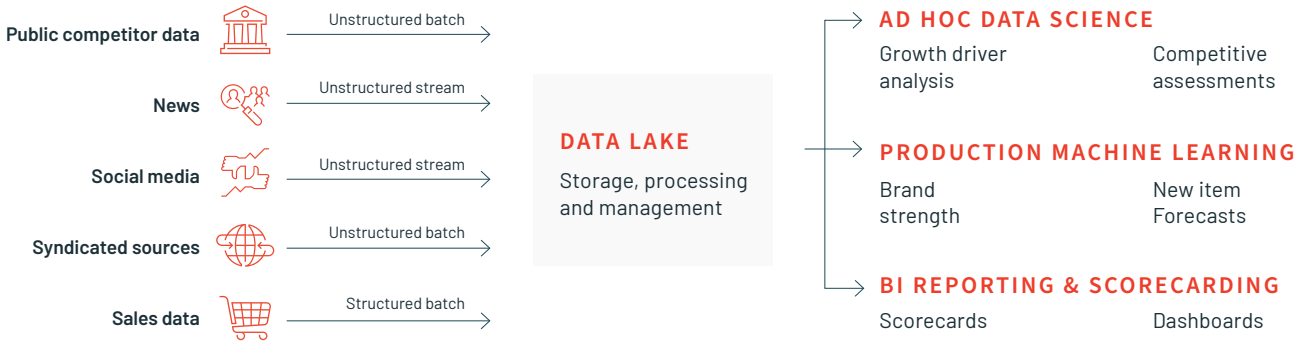
Starbucks forecasts demand at scale with Databricks

H&M streamlines supply chain operations with AI

CVS delivers personalized experiences to 8 million customers per day

Unlock the value of data lakes for BI and ML

Databricks provides a unified data analytics platform that helps retail & consumer brands simplify data and AI - accelerating omnichannel innovation.



DATA CHALLENGE	DATABRICKS SOLUTION
DATA INGEST: Processing batch and streaming data can be slow and error-prone, impacting downstream analytics.	Connect real-time inventory data with real-time customer experience data.
DATA LAKE MANAGEMENT: Data silos can limit ability to gain a complete view of the customer.	Easily handle large volumes of data from multiple sources (clickstream, PoS, social, etc) built on a strong privacy foundation.
DATA QUERY: Fragmented, siloed and inconsistent data sources for BI and data science.	Ability to rapidly and inexpensively experiment, manage, and push out at scale from a single platform.

Databricks customers in Retail & Consumer Goods

Use cases

From in-store to mobile, consumers are expecting personalized omnichannel experiences. With Databricks, retailers and consumer brands can truly support a 360 understanding of their customers and have a real-time view of their supply chain.

DEMAND FORECASTING & INVENTORY

Improve accuracy in inventory predictions by understanding customer demand, enabling you to reduce excess inventory while avoiding lost sales.

Supply Chain Control Tower	**Safety stock analysis**	**Causal forecasting**
Time series forecasting	**On-shelf availability (OSA)**	**SKU rationalization**

PERSONALIZATION

Drive incremental revenue through enhanced segmentation based on deeper behavioural insights.

Consumer segmentation	**Propensity to buy**	**Survival analysis & churn**
Customer lifetime value	**Customer retention**	**Real-time recommendations**

PRICING & PROMOTION OPTIMIZATION

Optimize pricing during key moments of the customer lifecycle or season

Dynamic pricing	**Promotion optimization**
Price optimization	**Promotion effectiveness**

Why Databricks for Retail & CPG is Proven

DELIVER HIGH-IMPACT ANALYTICS IN TIGHT SLAs

Databricks enables customers to deliver the most demanding of analytics to the front-line in your service windows. Need to allocate inventory or predict on-shelf availability for every store and SKU today? It's not a problem.

POWER PERSONALIZED EXPERIENCES WITH REAL-TIME DATA

Real-time awareness drives higher relevancy in recommendations, leading to higher incrementality and stronger customer satisfaction. Databricks enables companies to incorporate batch and real-time data of all types to power your e-commerce and mobile experiences.

GREATER AGILITY LEADS TO IMPROVED RESILIENCY

In this era of volatility, retailers and consumer goods companies need a platform that enables them to respond to changes in real-time. Deliver new insights and analytics in days and weeks, not months with your traditional warehouse.

The Databricks Impact

Databricks helps companies automate infrastructure management, increase ETL performance at scale, and accelerate machine learning and analytics initiatives.

12x faster
ETL pipelines

Impact:
Faster time-to-market of new analytics insights and models.

+25%
gains in productivity

Impact:
More productive data scientists results in more AI innovation.

+47%
Overall cost savings

Impact:
Lower infrastructure costs boosts operational margins

To learn more visit us at **dbricks.co/retail**

databricks

Case Study: H&M

Revolutionizing Fashion with AI

Leveraging Data-Driven Decision Making To Improve Supply Chain Operations

As a major disruptor and innovator in the fashion and retail industry, H&M relies on data as the core for everything they do. With stores opening up globally at a rapid pace, they needed to improve supply chain and forecasting operations to streamline costs and maximize revenues. But their on-premise Hadoop system crippled their ability to ingest and analyze data generated by millions of customers needed to power predictive models. Understanding they had reached their scalability ceiling, H&M moved to Databricks Unified Data Analytics Platform to simplify infrastructure management, enable performant data pipelines at scale, and simplify the machine learning lifecycle — allowing them to make data-driven decisions that accelerate business growth.

INDUSTRY:
Retail

VERTIAL USE CASE:
Improve supply chain and forecasting

INDUSTRY:
Data ingest and ETL, machine learning

Legacy Architecture Unable To Support Company Growth

In order to improve supply chain efficiencies, they chose to utilize data and AI to improve decisioning and operations. However, their legacy Hadoop based architecture was inefficient and wasn't able to scale to meet their rapid business requirements.

- Massive volumes of data from over 5,000 stores in over 70 markets, with millions of customers every day.
- Data engineering was challenged with fixed size clusters, complex infrastructure that was resource intensive and costly to scale, and data security issues.
- Struggled to scale operations to support data science efforts against all of this data coming from various siloed data sources.
- Time-to-market suffered because of significant DevOps delays, which impacted the ability for their data scientists to build, train, and deploy models quickly. It would take a whole year to go from ideation to productionization.

70%

REDUCTION IN OPERATIONAL COSTS

databricks

Case Study: H&M

Simplifying Data Operations Boosts MI Innovations

Databricks provides H&M with a Unified Data Analytics Platform on Azure that has fostered a scalable and collaborative environment across data science and engineering, allowing data engineers and scientists to focus on the entire data lifecycle instead of managing clusters, to train and operationalize models rapidly with the goal of accelerating supply chain decisions for the business.

- Fully managed platform with automated cluster management simplifies infrastructure management and operations at scale.
- Collaborative notebook environment with support for multiple languages (SQL, Scala, Python, R) enables a diverse team of users to work together in their preferred language — creating a unified cross-team environment to fuel productivity.
- Integrated Databricks platform with Azure and other technologies like Apache Airflow and Kubernetes, so elastic model training at massive scale can be achieved.

Smarter Decisioning, Dramatic Cost Savings

At H&M, even a 0.1% improvement in accuracy of a single model has a huge impact on the business. With Databricks, H&M is making data more accessible for each and every decision maker, making business grow faster and more relevant.

- **IMPROVED OPERATIONAL EFFICIENCY:** Features such as auto-scaling clusters has improved operations from data ingest to managing the entire machine learning lifecycle — reducing operational costs by 70%.
- **BETTER CROSS-TEAM COLLABORATION:** Unified analytics environment for both data scientists and engineers has dramatically reduced the number of components needed to go in production with easy setup and management.
- **HUGE BUSINESS IMPACT WITH FASTER TIME-TO-INSIGHT:** The ability to be more granular in decision making has allowed them to improve strategic decisioning and business forecasting.

> “Databricks is the core of our data business, it’s the place we go for insights.”

ERROL KOOLMEISTER

Head of AI Technology and Architecture, H&M

databricks

Case Study: Starbucks

Brewing Data and AI at Scale

Starbucks Serves Up Omni-Channel Experiences Across 30,000 Stores With Databricks

Starbucks serves millions of coffee lovers every day all across the globe. In order to maintain the highest level of customer service, they focus on building lasting customer connections, creating product innovations, and accelerating the digital experience for customers.

WATCH VIDEO

INDUSTRY:
Retail and CPG

SOLUTION
Demand forecasting, personalization, product innovation

TECHNICAL USE CASE:
Machine learning, data ingest & ETL

The key to their success lives within their data and can impact multiple domains, from improving inventory management to personalizing the digital experience. With Databricks, they have put a unified data and analytics infrastructure in place that can be leveraged enterprise-wide, building fast data pipelines at petabytes-scale that allows them to rapidly build ML models that improve inventory management, and unlock new product and service innovations.

Building An Architecture To Support Petabyte-Scale Data and MI

Data is crucial at Starbucks. Across 30,000+ stores, they generate billions of transactional data points that can be used to fuel data-driven innovations and operational improvements. Their data strategy and guiding principles are built on three pillars: 1. A single version of the truth; 2. Data and analytics enablement; and 3. Trusted data. However, extracting value from their data was the first and foremost challenge.

With petabytes of data to be ingested for downstream machine learning and analytics, their architecture struggled to handle the scale. They also dealt with a variety of structured and unstructured data that was fast-changing and fragmented across various systems, making it difficult to gain a complete view of the customer and business.

With a huge variety of data sources and types, data reliability and governance was of utmost importance, but difficult to achieve. They needed a way to build out their historical data and live aggregations together to ensure they were delivering real-time, accurate insights to their stores and partners.

They also struggled to provision clusters to support their data needs. Data engineering was often overwhelmed with spinning up and maintaining clusters. "Our engineering services were not optimal," explained Vishwanath Subramanian, Director of Data Engineering and Analytics at Starbucks. "We struggled to scale compute in a timely manner, often taking over 30 minutes to scale clusters."

Once the data did make it downstream to the data science and analytics teams, the lack of a unified user experience acted as an impediment to innovation, blocking exploration, experimentations, and reproducibility. To truly create meaningful connections with their customers, they needed to remove these barriers to innovation.

1000+
DATA PIPELINES

50-100X
FASTER DATA PROCESSING

15 minutes
TO DEPLOY ML MODELS

databricks

Case Study: Starbucks

A Single Source Of Truth To Brew Up New MI Use Cases

To address these challenges, Starbucks' developed BrewKit, a zero-friction analytics framework, built on top of Azure Databricks. Their goals were to ensure the democratization of data by creating a single source of truth, while creating an environment that fosters cross-team collaboration to unlock the possibilities of machine learning at scale.

"We wanted to make sure the smallest of teams at Starbucks had the ability to do data science and data engineering at scale," said Subramanian. "The only way to enable that was to empower them with a massively scalable unified analytics platform."

With Azure Databricks and Delta Lake, their data engineers are able to build pipelines that support batch and real-time workloads on the same platform. This has enabled their data science teams to blend various data sets to train new models that improve the customer experience. Most importantly, data processing performance has improved dramatically, allowing them to deploy environments and deliver insights in minutes.

From a data science perspective, the interactive notebooks have enabled users to onboard quickly and collaborate more efficiently and more easily manage various use cases. Once models are developed, MLflow allows them to easily experiment and test models in a rapid fashion. "From a data team collaboration productivity standpoint, this has been huge. The tooling has been collaborative. We also now foster a culture of experimentation and self-service, and maintain shared responsibility across environments," said Subramanian.

Databricks Is Helping Starbucks' Build a Data-Driven Future

With a unified data analytics platform at the core of their data strategy, Starbucks' entire data strategy has been transformed. Data can now flow seamlessly through their pipelines and models, allowing for new ideas and solutions to flourish. The processing power of Databricks and Delta Lake paired with Azure services has increased performance 50-100x, giving data science and analytics teams the data they need faster.

Delta Lake provides a trusted, persistent storage layer that securely delivers quality data that enables downstream data analytics. This allows them to explore many analytics use cases across the board such as tour operations, quality of service analysis, demand forecasting and inventory management, personalized shopping experiences, and much more.

"With Databricks, we can now take a strategic view into data analytics," expressed Subramanian. "So much so, that our teams can now focus on business problems up the value chain rather than simply moving data from point A to point B."

As Starbucks continues to focus on providing world-class customer experiences, Subramanian is excited about the impact Databricks will continue to have in achieving their mission. "At Starbucks, we are elevating customer connections through the convergence of data and AI," concluded Subramanian. "As we extend our channels for delivery and adjust to the new norms in today's new era, data will play an extremely crucial role in this effort."

> "With Databricks, we can now take a strategic view into data analytics. Our teams can spend time focusing on business problems up the value chain, rather than simply moving data from point A to point B."
>
> **VISHWANATH SUBRAMANIAN**
> Director of Data Engineering and Analytics, Starbucks

databricks

Case Study: CVS Health

Personalizing The Pharmacy Experience To Enable Better Outcomes

CVS Health integrates pharmacy and shopping customer journeys across 10,000 locations with Databricks

With over 80 million customers passing through their pharmacies every day, CVS Health is always striving to provide more meaningful interactions that put customers on a path to better health. In 2018, they embarked on a journey to personalize experiences through machine learning on Hadoop environment, but the complexity and scale of the diverse data set prohibited understanding the behaviors of a large number of micro-segments of customers. With Databricks, CVS Health was able to analyze their customer data to implement different experiences, experiments, and a variety of segments for personalization at scale.

CVS Health

INDUSTRY:
Healthcare

SOLUTION
Patient-centric healthcare

TECHNICAL USE CASE:
Ingest and ETL, Data analytics, Machine learning, Data teams

Understanding Varied and Unpredictable Customer Behavior

In 2018, CVS Health was ready to focus on personalization, but with 10,000 stores across the United States and a large number of microsegments full of unpredictable behaviors, personalization was easier said than done.

"We're not dealing with a typical grocery store customer," explained Raghu Nakka, Sr. Analytics Advisor at CVS Health. "It's really hard to predict behavior because a customer could go to a convenience store just because he forgot the milk, or he could go to a convenience store like CVS to pick up his prescription and stop to grab some candy. So the predictor of the behavior could be unpredictable, which is why our data dimensionality leads to overfitting issues for any kind of machine learning model."

CVS Health kicked off its personalization journey in Hadoop. They built out an environment and launched their first personalized campaign to 1% of customers within a few months, but ran into roadblocks when they tried to scale from 1% to 5% because of a lack of processing power and physical data storage. "There's an actual constraint around building additional hardware to support the scale that we wanted to get to," said Michelle Un, Director of Enterprise Analytics at CVS Health.

In need of a more robust platform to achieve the level of personalization they wanted, the CVS Health team once again began setting goals and exploring their options. "We wanted to understand which channel the customer would respond to better, whether that's a text message or a phone call, or an in-store offer, so optimizing our process was a really important goal," added Raghu. "And the timing was just as important — like let's say someone is up for a prescription renewal. That could be the right time to send an offer to the customer, rather than the customer who has already filled a prescription."

1.6%

INCREASE IN MEDICATION ADHERENCE (TAKING MEDICATIONS AS PRESCRIBED)

databricks

Case Study: CVS Health

A Cloud-Based Environment Powers Personalization At Scale

CVS Health ultimately decided to transition to a cloud-based environment using Azure Databricks, which allowed them to expand the number of use cases that they were able to support and eventually scale to offer personalized offers to pharmacy customers.

"Through Azure Databricks we have the flexibility to spin up clusters that meet our unique business and various business use cases," said Michelle. "We're also not restricted by any more physical hardware constraints."

Databricks also lent the teams some much-needed agility, which enabled them to focus on building and testing experimentation frameworks, rapidly iterate on different experiences, and expand their ability to utilize machine learning. And with centralized assets and interactive notebooks, any issues related to disparate data sources, such as limited data access and resource-intensive pipelines, were a thing of the past.

In their new, collaborative environment, all teams were finally able to work together: data engineering on building faster data pipelines, data scientists on training ML models more efficiently, and data analysts on visualizing financial and operational metrics for smarter internal decision making through Tableau.

> "Databricks helps us to predict customer behavior and then make our offers more relevant and increasingly personal."
>
> **RAGHU NAKKA**
> Sr. Analytics Advisor at CVS Health

Personalized Communications Power A Healthier Day-To-Day

With Databricks, CVS Health has a much better idea of who their customers are and what their needs are in the moment. They can measure the probability of buying a particular product, or when to remind a patient to fill or pick up their medication, and even identify potential side effects the patient may be experiencing.

"That's really the core of our strategy," said Raghu. "We predict the behavior of the customer and then make our offers more relevant and increasingly personal."

The results of boosting personalization at CVS Health include a 1.6% improvement in medication adherence, meaning an increasing number of patients are now taking their medication on time and as directed.

With Databricks as their data analytics foundation, CVS Health is able to further patient-centric healthcare through actionable data insights and innovative models focused on improving long-term customer health and well being.