EBOOK

Manage your Data as a Business

Contents

Context	1
Managing your Data as a Business	2
Grow Revenue	3
Optimize Cost of Goods Sold	5
Streamline Operating Expenses	7
Monitor Performance	8
Get Started	10

This paper provides a perspective on the importance of thinking of your data as a business asset, and managing your data like a business operation. In a fast changing world where AI holds immense potential, and requires careful safeguarding, it is more important than ever to get your data business in order. Doing so requires a balanced focus on growing top line returns to the business with data whilst also improving the bottom line cost to serve the business by modernizing data foundations, optimizing resources, transforming culture, skills, and the experience of using data across the business. The paper takes a financial management lens toward how CDOs should manage data transformation to engage with business stakeholders on the business value of data, and create the foundations for AI at scale.

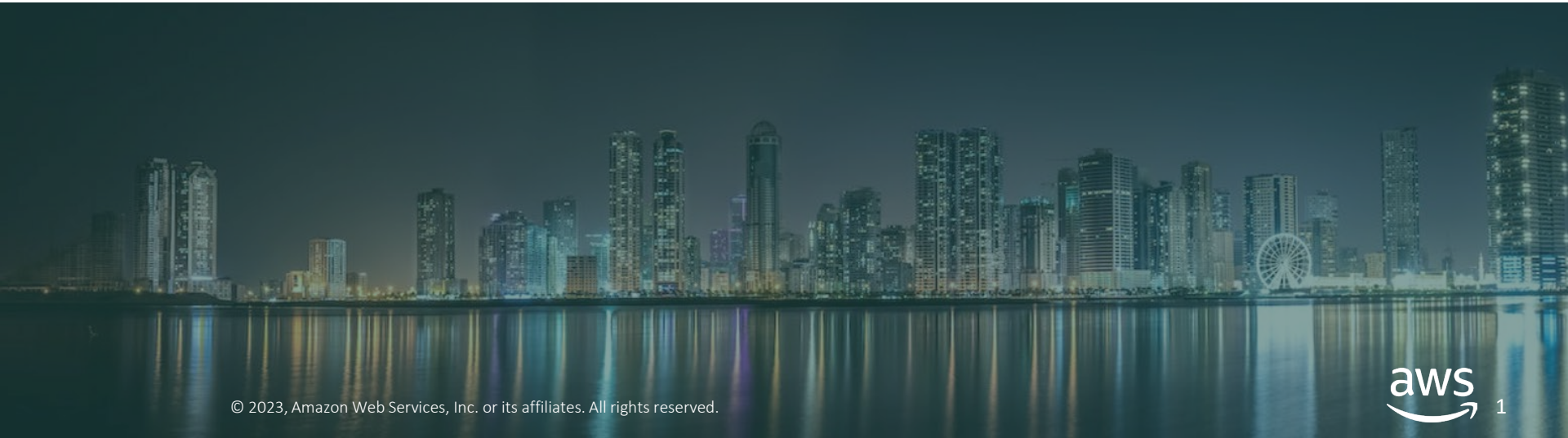
Context

The importance of data for organizations is widely acknowledged, but often, its potential is poorly leveraged. A 2023 Wavestone survey finds the Chief Data Officer (CDO) role is at critical mass with 82% of organizations now having appointed a CDO, and 93.9% of organizations increasing investments in data in 2023. Whilst organizations place data as a top strategic priority, many still fall short of delivering value with only 23.9% categorising themselves as data driven, harnessing data as an asset, driving sustained innovation, and creating actionable intelligence.

The sheer volume, velocity, and variety of data generated through customer interactions, operational processes, and external sources hold immense potential to uncover valuable insights, drive new innovation with AI, enhance operational efficiency, and fuel growth. Organizations that harness the power of data and leverage it strategically gain a significant advantage. According to Forrester, advanced data-driven businesses are 8.5 times more likely than beginners to report at least 20% revenue growth.

At AWS we see customers like The Very Group deliver 9.9% improvement in SKU management worth £110 million by leveraging machine learning to optimize their demand forecasting. Engie realised 1% of savings on $500m renewable energy assets using data as they transform to a sustainable energy business. Lyell, a next- generation cancer treatment pioneer, uses data to accelerate cell therapy manufacturing. Moody's realized a 193% increase in revenue using machine learning to predict global markets.

To create sustained business value from data, and harness the potential power of AI, requires a mindset shift to think about managing your data as a business. According to Wavestone only 39% of organizations currently think of data as a business asset. To do so requires a focus on growing top line returns in value to the business. It also requires improving the bottom line by lowering the cost of creating and managing data, unlocking business access, and speeding innovation, whilst continuing to monitor performance and adjust priorities, much like you would do when running the operations of a business.



Manage your data as a business

The dynamics of running data as a business can be compared to the efficient and effective running of any other revenue generating business operation.

If we take an example from the Retail industry, retailers need to balance growing their top line (Revenue) whilst improving their bottom line by optimizing Cost of Goods Sold (COGS), and streamlining Selling, General, and Administrative expenses (SG&A). For a retailer, COGS comprises sourcing and procuring the best input wholesale materials, optimizing how these are manufactured to finished goods, and ensuring they are stored and inventoried accurately for distribution close to where customers make a purchase, whether online or in store. Revenue consists of selling the finished merchandise to customers and SG&A covers all of the expense relating to talent, branding, marketing, and customer services required to raise awareness of the brand, engage with customers, and create a great experience.

The better a retail business can manage this balance, the more they can reap returns that can be reinvested into the growth and ongoing optimisation of the business as it expands.

Contrasting this with the business of data, the same balance is required. Data business owners (The CDO) need to continuously grow their top line, showing value to the business by facilitating the delivery of use cases that meet with priorities to either boost business revenue, optimize cost in business operations, mitigate risk, or power differentiating customer experiences.

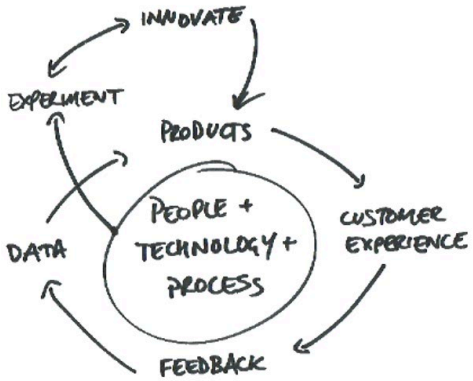
Alongside this the CDO needs to streamline the data bottom line by improving their Data COGS. This includes sourcing relevant data, investing in lowering the cost of creating and managing data across the business, and unlocking business access to data products that are trusted, reliable, well defined, and secure. MartinFowler defines a data product as a high-quality, ready-to-use set of data that people across an organization can easily access and apply to different business challenges. High-quality data products are also essential to train, deploy, and operate AI/ML models optimally and avoid inaccurate, biased, or unethical results. Additionally, the CDO needs to focus on streamlining SG&A expense by developing a data driven culture, building talent across the organization, marketing the data products to the business, and ensuring a compelling experience for business customers.

Retail business

Data business

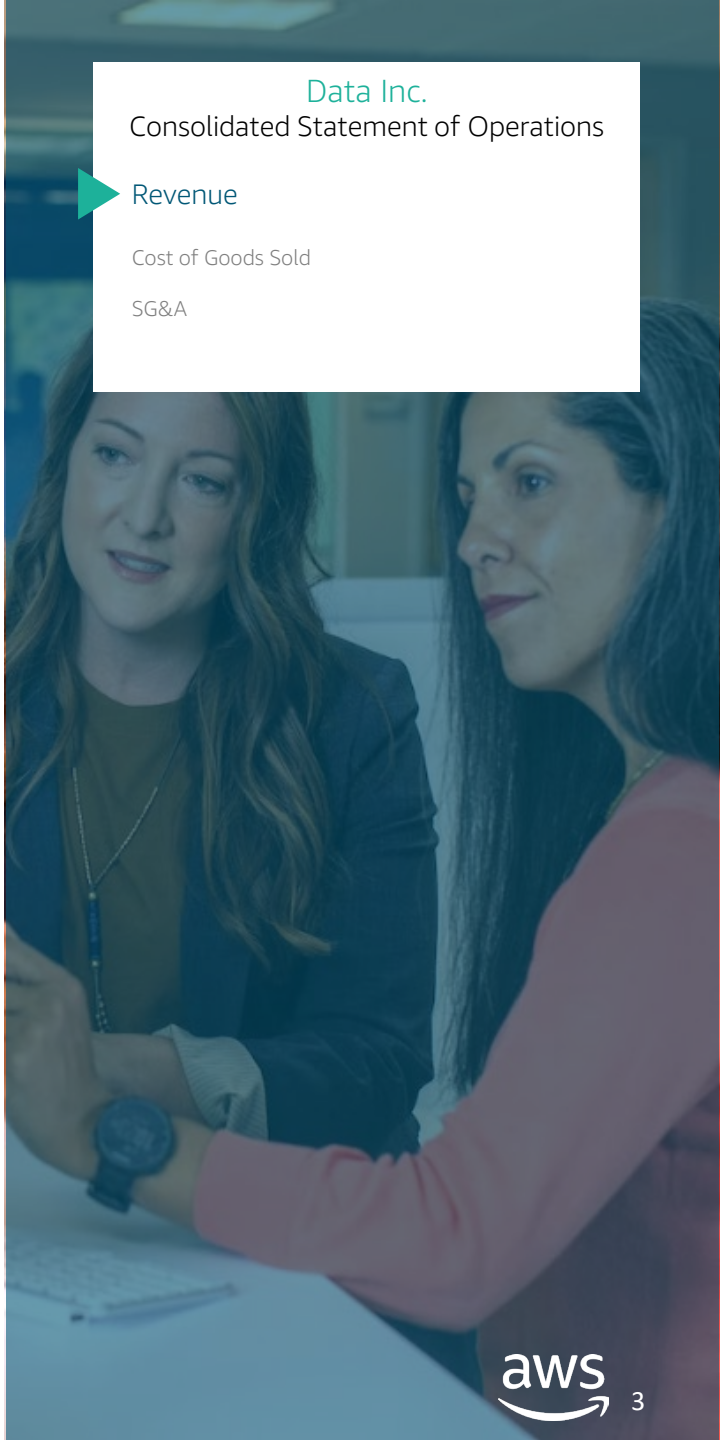
Grow Revenue

Growing data revenue requires a focus on business and customer priorities and using these to inform the right analytics or AI/ML use cases, and supporting data products in which to invest in. This is similar to the process a retailer might leverage to identify customer demand and market opportunities, and then create or acquire products that meet the specific customer need. At Amazon we have a process called working backwards. This is all about starting with a focus on customer priorities and working backwards from these to the right solution. When we think about data we apply the same concept. Rather than spending months or years building a datalake to then figure out what value to derive from it, struggling to prove value, and then having the business disengage, we see that the best practice for scaling data value is to drive a flywheel approach working backwards from priorities in the business. In the current rapid adoption phase of AI, this remains true, organizations need to continue to focus on business value to work backwards to the right AI solutions supported by modern data foundations rather than investing in expensive training cycles and model deployments without connecting these to business priorities..



By working backwards from the customer and continuously collecting feedback, this ensures a focus on data products that add value. This allows organizations to *think big* about what they want to achieve, but *start small* so they can experiment, learn, and realize value while driving faster innovation.

United Airlines took a flywheel approach to driving value across their business. Beginning with a focus on taking cost out of operating their 36,000 ground service assets, they leveraged data, IoT, and machine learning to optimise asset usage and save $120 million dollars from a $10 million dollar investment. They reinvested the benefits from this first use case to scale further use cases such as predictive scheduling to improve maintenance staff schedules saving 1.3 million hours, and to reduce carbon emissions across their airport operations by 7%.



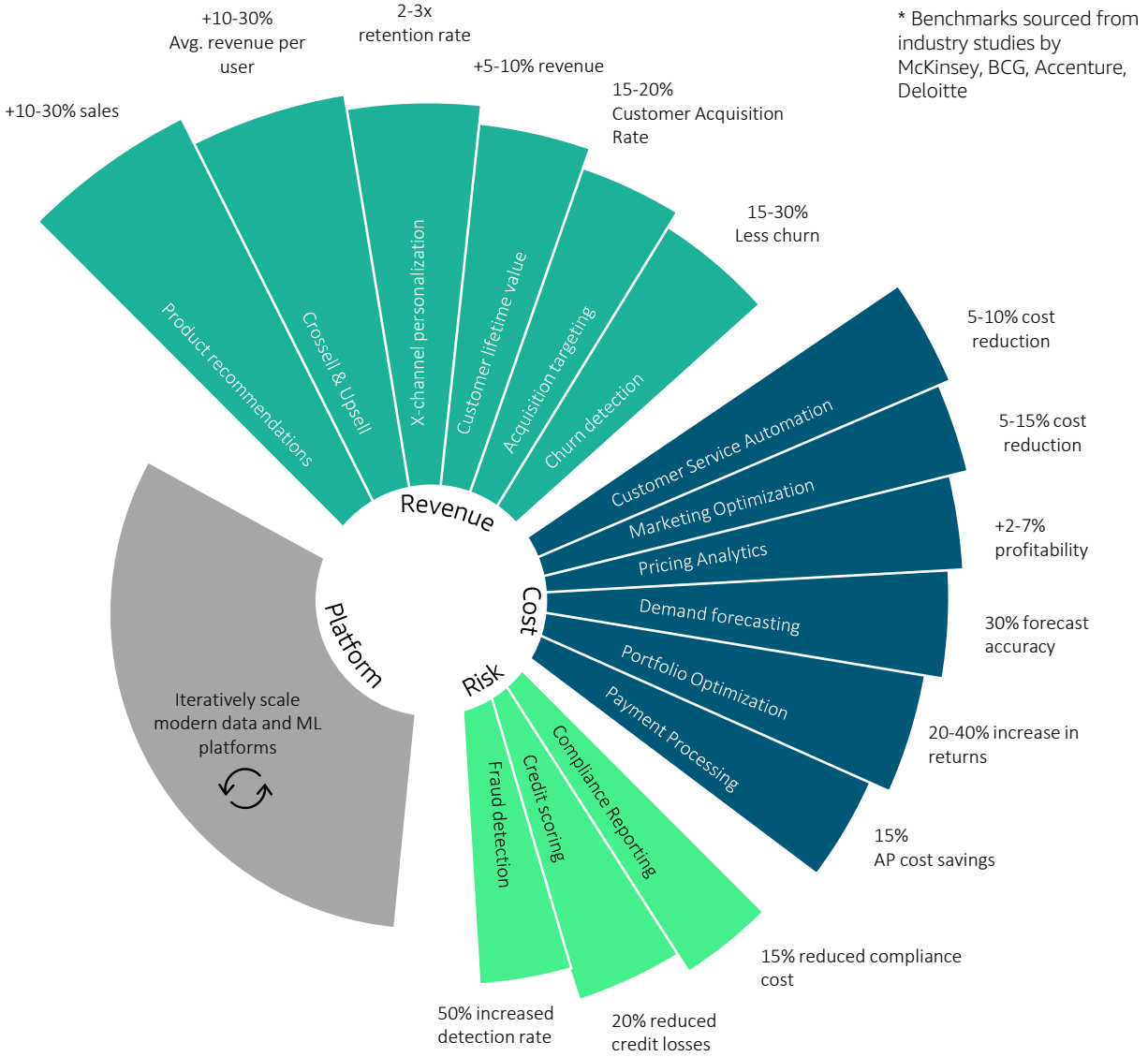
Grow Revenue: an example for B2C focused businesses

An illustration from consumer-facing businesses (B2C) shows many opportunities to grow the data top line and generate business value with AI, analytics, and data. This could begin with a focus on growing business revenue, optimize cost, or mitigating risk, depending on where the organizations priorities lie.

When beginning, it is important to not get stuck in elongated periods of use case identification and value estimation resulting in analysis paralysis. Business priorities continually shift and where there are high value, low risk opportunities, it pays to move fast and show value rather than risk disenfranchising the business by over architecting solutions.

As the flywheel spins you can iteratively scale foundational data and ML platforms that allow for more and more use cases to be leveraged at a lower cost, driving value into the business faster, and creating momentum in returns to invest in new data bottom line capabilities that unlock further value. The flywheel becomes a self-funding mechanism for data leaders to invest and innovate in the business.

This flywheel can be applied to monetizing your data externally as well. When considering external data monetization, data product owners must identify the addressable market segment who would benefit from their data, analytics, or ML models. Product owners need to quantify market segment share and price to determine the revenue contribution over time, and monetized data products should be subject to regular profitability evaluation with direct and indirect expenses assigned.



Optimize the Cost of Goods Sold

To increase capacity to scale business value requires the CDO to optimize the cost of data and develop modern data foundations that provide increased economies of scale as use cases continue to be delivered to the business. There are 7 business management levers for COGS which can readily be applied to data.

1. Supplier Management and Sourcing → Data modernization
The CDO should invest in streamlining the cost of data generation and acquisition. Data can often be trapped in legacy systems, with poorly understood definitions or ownership, and slow, costly mechanisms to access and extract data. Much like a modern supply chain the CDO needs to modernize where and how data is sourced. Cloud is an enabler for this, and GLG Brightpaper research shows migrating and modernizing legacy databases results in 13% lower licensing cost, 68% faster query execution, and 53% less DB administration .

2. Product design and standards → Data product design
To allow for uniform, interoperable use of data use across the business there needs to be universal standards that are defined for common data products in the same way that you need standards for retail stock keeping units (SKUs). Data products such as customer data, stock data, or pricing data require a single threaded team to define the standards for how these should be described and structured.

For example a Customer Service team might steward the standards for capturing mandatory attributes of the customer like unique identifiers and Personal Identifiable Information (PII). They would not own all customer data throughout the organization, they would own the standard. Anyone who produces customer data across the organisation then leverages the template as the universal standard for customer data. Producers of customer data can also append to this template with further attributes or fields specific to the context for how they generate customer data.

3. Inventory management → Data cataloguing and ML model management
As data products are produced across the organization it is critical to have good inventory management practices in place to minimize redundancy and waste, and build trust. A common catalogue where all data is registered with business as well as technical metadata allows for consumers across the business to easily discover, understand, and access data products. Business consumers can leverage the catalogue to understand the quality levels, reliability, and source of the data product before deciding if it is the right one for them.

In the same way a common Machine Learning feature store and model management platform are required to eliminate inconsistent use of features in different models, and reduce duplication of models by uniquely identifying deployed models, managing model versions, and tracking where they are being applied in the business.

Optimize the Cost of Goods Sold

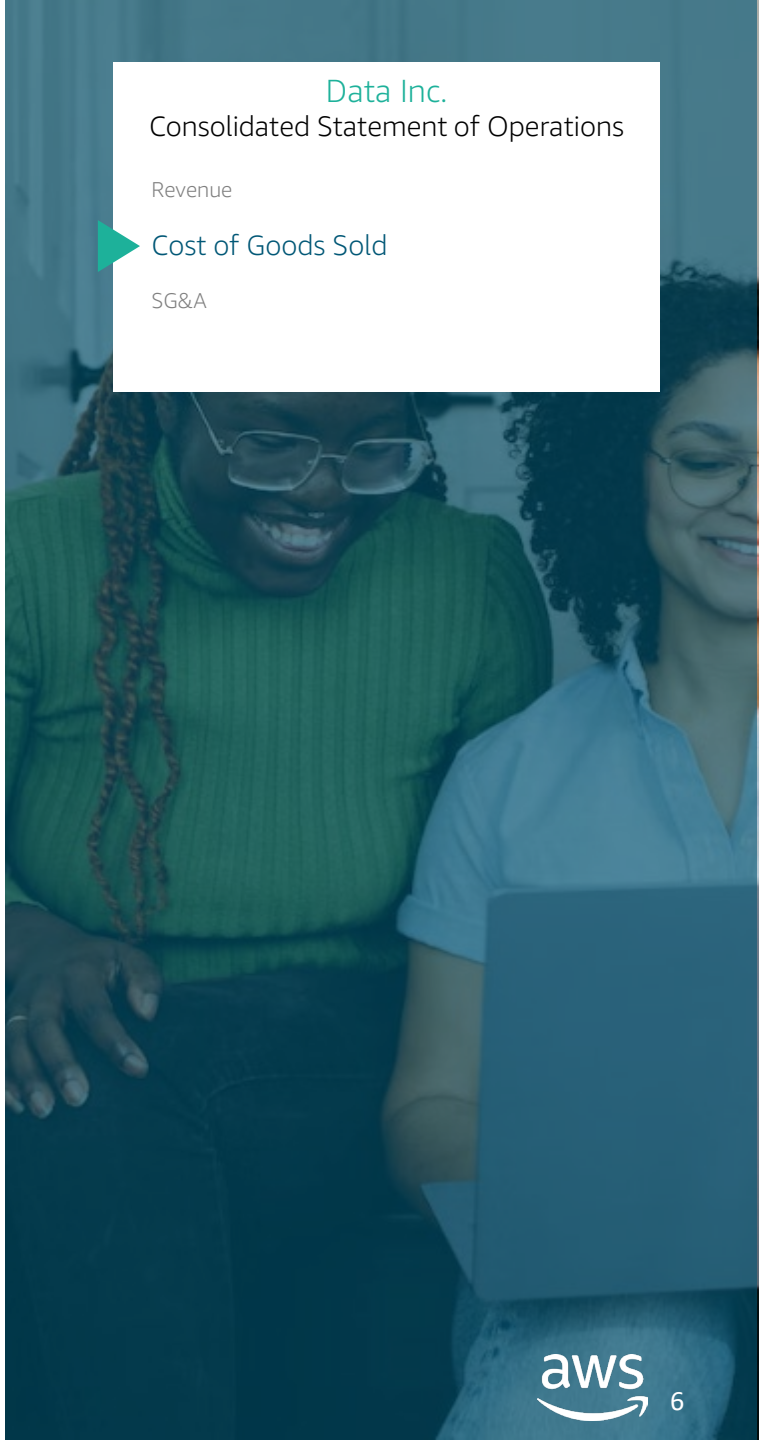
4. Last Mile Delivery → Data Consumption
Like a retailer, last mile delivery of data is important to deliver data products to the consumer in the right place, at the right time, and in the right form. Data consumers should be given the choice to consume data products to the endpoint that matches their needs, whether this be in a descriptive insight tool, ad-hoc analysis, predictive modelling, or through API consumption. Consumers should also have the ability to setup ongoing subscriptions with automatic refreshes and updates on the data product.

5. Operational processes → Data Platforms and tools
The CDO can leverage lean manufacturing principles to drive data platform efficiencies, and minimize the cost of connecting producers and consumers of data across the organization to share data, and meet business needs.

To do so requires a continuous effort to identify bottlenecks where data is not moving fast enough from the teams who produce it, and delivering automation to remove manual effort in areas like data quality management, data classification, and detection of sensitive information. This also requires implementing common Data and ML DevOps platforms to standardize the deployment and operations of new data solutions across the business. Finally, investing in templates for business consumers to leverage like reporting templates or standard query scripts removes heavy lifting from teams across the business wanting to extract business value from data products.

6. Cost control → Data usage monitoring and compliance
As the usage of data proliferates across the organisation, it is important to continue to optimise the cost of redundant storage or services, monitor for improper use, and ensure compliance to legal and regulatory requirements. Doing this does not require a huge manual effort, and customers like JPMC and Roche leverage tools like Amazon CloudTrail to automate monitoring of usage without creating bureaucracy.

7. Continuous improvement
As the data platform continues to scale, the CDO should foster a culture of continuous improvement across the organization. Incentivising teams to identify and implement process improvements, cost-saving initiatives, and efficiency measures. Data teams should also regularly review performance metrics, conduct root cause analyses, and implement corrective actions to drive ongoing cost optimization.



Streamline Operating Expenses (SG&A)

Change culture to reduce barriers to adopting data
Fostering a data culture requires a shift from executive support to executive engagement, where top management is seen leading the change and demonstrating visible progress. The data team needs to also collaborate closely with operational business teams to solve problems linked to strategic and operational priorities to build change momentum bottom up. Delivering value quickly, and collaboratively with the business shifts mindset, reduces barriers to adopting data, and converts skeptics to advocates.

Grow and democratize data talent across the business
Unlocking data value requires an on-going investment in improving data literacy and creating a common data vocabulary to increase autonomous use of data across the business. Doing so increases economies of scale in data use and lowers the cost of inefficiencies in trying to serve the business from a single central data organization. National Australian Bank trained 4,500 people across the business to increase data literacy and engineering skills, improving their ability to meet customer expectations and their speed of innovation.

Shift the data economy
Early on, data value creation often starts with technical and data teams proposing ways the business can better use the organization's data to improve outcomes. We call this a "supply-side" data economy. As more business skeptics become advocates, and more talent is comfortable using data, an organization will begin to pivot to a "demand-side" data economy. This is a clear sign that a culture shift is occurring, leading to broader use of data and greater economies of scale from data investments. At Amazon, there are about 90,000 employees innovating with data and more than 70% are in non-technical roles.

Evangelize and market success
Evangelize data products and available use cases in the same way that you would create marketing campaigns in a retail business. Celebrating wins early and often will keep the organization energized about data. Provide guidance on how consumers can use data products to solve business challenges. Cross-pollinate key learnings of successes and failures by creating cross-organizational communities, and nurturing growth and engagement.

Create a compelling customer experience
Building data products that provide insight in the moments that matter creates a compelling customer experience, whether for external customers or those inside your organization. Reduce the mental strain from context switching and look to integrate insights into workflows and decision-making processes. Discovery Sports Events delivers near-real time stats and ML predictions to create an immersive fan experience at its Eurosport UCI Track Championship races. This is delivered at the velodrome, on broadcast screen and in its app to grow its fan base and deepen engagement. These "in the moment" insights can be created for internal data customers and integrated into operational applications to create a frictionless experience when using data-driven insights.

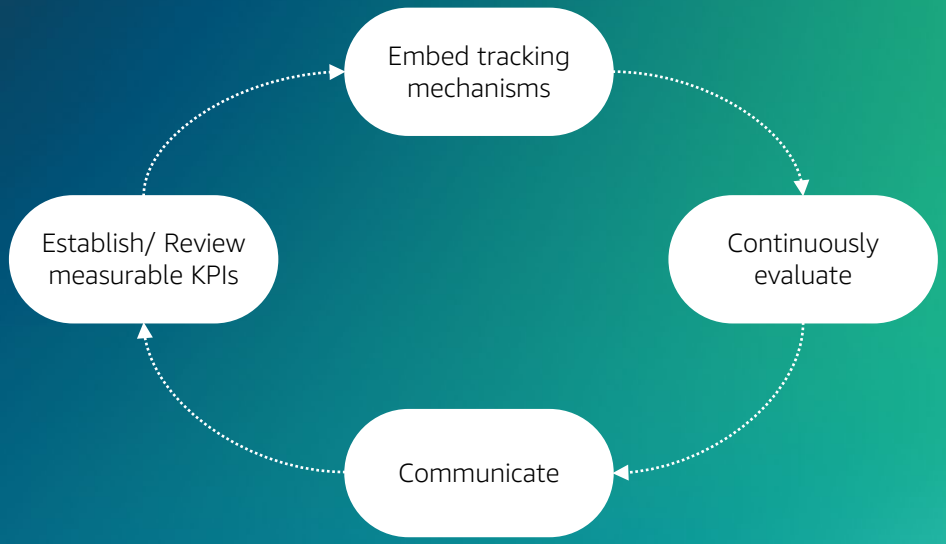
Monitor Performance

Regular monitoring and assessment of business asset performance involves tracking key performance indicators (KPIs) such as availability, reliability, utilization, and efficiency to ensure assets are performing as intended. Monitoring helps identify potential issues, plan maintenance activities, and make decisions to improve asset performance. Tracking and reporting on the value of data in a business is similarly crucial for understanding the impact and return on investment of data.

Doing so requires Establishing measurable KPIs that align with your strategic business objectives, such as improved operational efficiency, increased revenue, or enhanced customer satisfaction, and then embedding mechanisms in the business to track these KPIs.

Continuously evaluate the impact of data, analytics, and ML initiatives. Assess the effectiveness of models and algorithms, measure the business value created, and identify areas for improvement. Regularly review and update your KPIs to ensure they remain aligned with evolving business priorities.

Finally communicate the value effectively. Ensure that the insights and findings derived from data are effectively communicated across the organization. Use storytelling techniques to convey the impact in a compelling and relatable manner. Emphasize the tangible benefits, such as cost savings, revenue growth, or improved customer experiences. Communicate success stories and share best practices to inspire further adoption and support from stakeholders.

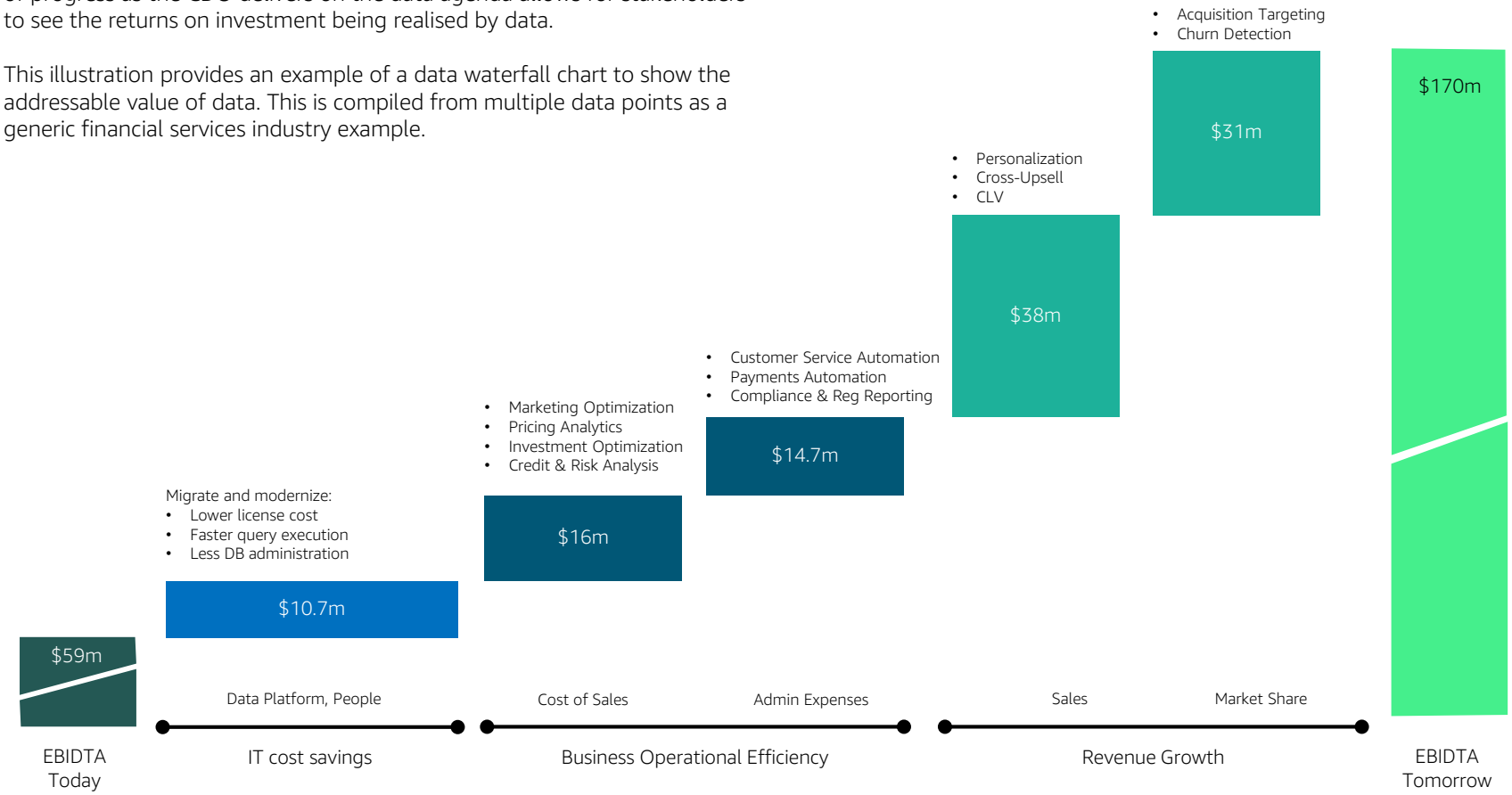


An industry example:

The Value of data in Financial Services for a $600m revenue generating business

Communicating the value of data in the format and language of the business can be incredibly powerful in helping stakeholders to understand the case for change and addressable value upfront. Continued monitoring of progress as the CDO delivers on the data agenda allows for stakeholders to see the returns on investment being realised by data.

This illustration provides an example of a data waterfall chart to show the addressable value of data. This is compiled from multiple data points as a generic financial services industry example.

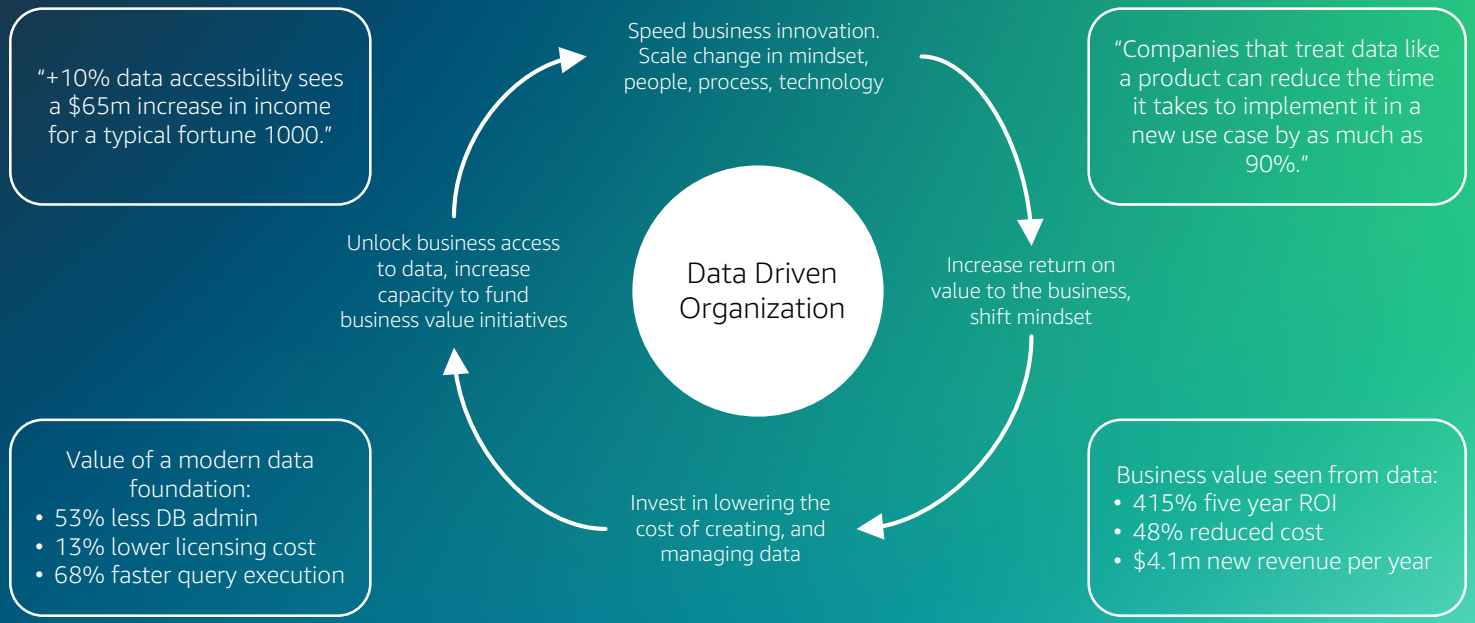


The Data Business Flywheel: Getting Started

Thinking of your data as a business asset, and managing your data like a business operation, enables sustainable scaling of end to end data capabilities, and business stakeholder buy in with clear articulation of the value gains versus the required data investments. There is a flywheel effect to mobilizing change and transforming how data is managed through the business. This requires frequent delivery of value to the business. Doing so builds enthusiasm and drives the need for more data and information, which further enables insights. This in turn defines new data products to be delivered that scale across the enterprise. An IDC study shows that businesses who build data, analytics, and ML solutions on AWS see on average a 415% ROI, 48% reduced cost, and $4.1m new revenue each year.

The value generated from starting with a focus on key business priorities can be invested to lower the costs of creating and managing data by migrating priority legacy DBs and applications to the cloud, and creating a modern data foundation. A recent survey of AWS customers revealed that a modern data foundation delivers 53% reductions in DB administrative effort, 13% savings on licensing costs, and 68% faster query execution. With this optimization, IT teams increase capacity, and can focus on unlocking business access to Data as a Product with the right cataloguing, marketplace and intelligence tools. A survey by Forrester shows that a typical fortune 1000 customer will see a $65 million increase in income by increasing data accessibility by 10%.

A McKinsey study shows that companies that treat data like a product reduce the time to implement new use cases by 90%. The more the business has access to Data as a Product the more they can scale innovation with AI, change mindset, grow appetite for more, and turnover business value faster and faster. Your flywheel begins to turn and gain self sustaining momentum, and your journey to being data driven is well on its way.



Dykes, Brent. The Four Key Pillars to Fostering a Data-Driven Culture, 2019
McKinsey & Co. 2022
IDC, The Business Value of AWS Data Lakes, Analytics, and ML Services, June, 2020

How AWS can help build a modern data strategy

Need peer-level executive guidance?

aws Data Strategy

- ✓ Mental models and strategies based on the first-hand experience of former CXOs
- ✓ Get a peer-level sounding board and sparring partner

Inspire and accelerate your data transformation

Want to build a data vision and strategy?

aws Data Driven Everything

- ✓ Create an organizational vision for innovation with data to drive business outcomes
- ✓ Define the first pilot, learn, and build

Jump-start the data flywheel

Want to modernize your data foundation?

aws Re:Imagine data

- ✓ define the migrate & modernization strategy for a future data foundation
- ✓ Lower cost, increase capacity, and unlock business access

Migrate & Modernize data

Visit Data Driven Everything for further detail and speak to your account manager to get your data flywheel turning.

aws

About AWS

For over 15 years, Amazon Web Services has been the world's most comprehensive and broadly adopted cloud offering. AWS has been continually expanding its services to support virtually any cloud workload, and it now has more than 200 fully featured services for compute, storage, databases, networking, analytics, machine learning and artificial intelligence (AI), Internet of Things (IoT), mobile, security, hybrid, virtual and augmented reality (VR and AR), media, and application development, deployment, and management from 81 Availability Zones within 25 geographic regions, with announced plans for 21 more Availability Zones and seven more AWS Regions in Australia, India, Indonesia, Israel, Spain, Switzerland, and the United Arab Emirates. Millions of customers—including the fastest- growing startups, largest enterprises, and leading government agencies— trust AWS to power their infrastructure, become more agile, and lower costs. To learn more about AWS, visit aws.amazon.com.

Contributing Authors

Craig Suckling
AWS Data Strategy lead

Shari Elia
AWS Data Migration and Modernization lead

Kathy Koontz
Data Strategist

Daniel Brock
Data Migration and Modernization specialist

Karen Hensley
Data Strategist

Timothy Geiger
Data Migration and Modernization specialist