Data-driven organizations demystified

SQUADRA ANALYTICS WHITEPAPER

Squadra
Data-Driven Value Creation

Many organizations seem to recognize the potential value of data and analytics but struggle to adapt to the required change. Inspired by lean thinking, this whitepaper proposes five interventions to overcome some of the typical pitfalls.

Introduction

A decade has passed since Thomas Davenport and DJ Patil (2012) argued data scientists had the sexiest job of the 21st century. In their words, a data scientist is "a high-ranking professional with the training and curiosity to make discoveries in the world of big data". A lot has changed since. No longer a curiosity that can be studied at length, investments in data require a tangible business benefit for most organizations nowadays. The popularity of Big Data as a search term has passed its peak, according to Google Trends. And data scientists are no longer jacks of all trades, but usually collaborate in teams with data engineers, data architects, visualization experts, and analytics translators (Henke, 2018).

According to a recent survey by NewVantage Partners (2022), nearly 75% of organizations have appointed a Chief Data or Analytics Officer in 2022, up from 12% when this survey was first conducted in 2012. At the same time, only 26.5% of organizations report having established a data-driven organization, according to the same survey. If these numbers are any indication, it appears many organizations recognize the potential benefits of data & analytics but seem to struggle to adapt to the required change.

Evolution of the data-driven organization

The term data-driven organization can be defined in multiple ways. This article adapts the following definition from Michelle Knight (2021).

The term data-driven describes a business state where data is used to power decision-making and other related activities efficiently, in real-time.

The definition above consists of three distinct elements with a specific implication:

- Data is used: the application of data (only) has value when you put the derived insights into use.
- Decision-making: putting insights into use involves a method of evaluating which action to trigger, if any.
- Real-time: a data-driven organization requires a systematic, highly automated approach toward decision-making, as humans are unlikely to be able to make instant, sound decisions without errors.

The systematic computational analysis of data or statistics is often referred to as analytics and typically encompasses three distinct phases:

- Descriptive analytics: answers the question "What has happened". As such, it provides insight into historical patterns of behaviors and performance.
- Predictive analytics: answers the question "What will happen". Instead of only identifying historical patterns, predictive models use these patterns to make predictions about future outcomes.
- Prescriptive analytics: answers the question "What should happen". By taking advantage of the results of descriptive and predictive analytics, prescriptive analytics suggests decision options.

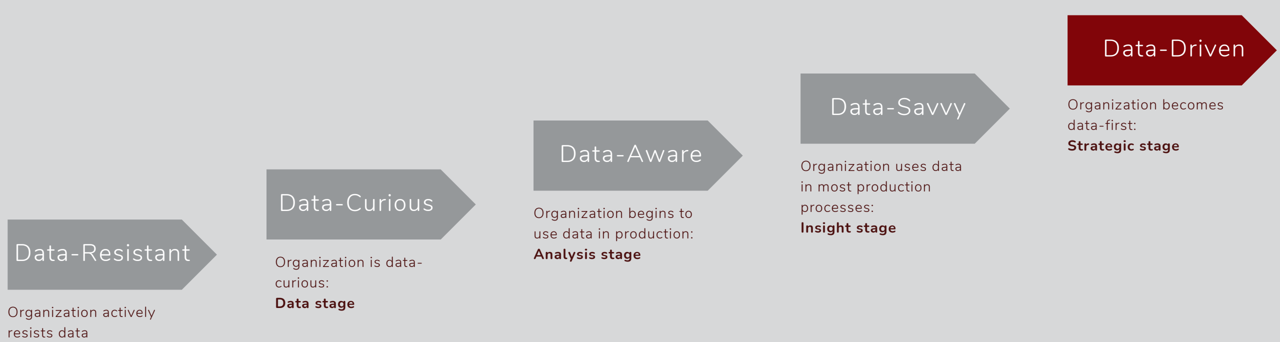
As the three phases of analytics suggest, the application of data & analytics is an evolution, where each phase builds upon the previous one. This can be taken quite literally, according to a survey by Anaconda (2020).

On average, data scientists spend 45% of their time on data preparation, which includes loading and cleansing. The time spent on modeling, including selection, training, and deployment sums up to about 34%. As the application of analytics becomes more advanced, so does the implied value. Put in the context of organizations, Christopher Penn (2019) recognizes five different stages of maturity when plotting value against adoption (see Figure 1). The ultimate maturity phase is the data-driven organization.

Introduction to required capabilities

The shift to a data-driven organization is first and foremost a cultural change. The survey of NewVantage Partners mentioned earlier shows that over 90% of respondents indicate cultural challenges represent the greatest impediment to becoming data driven.

Figure 1: Maturity stages of analytics adoption (adapted from Christopher Penn)



In its simplest form, organizational culture can be characterized as a set of shared assumptions that guide behaviors, according to Davide Ravasi and Majken Schultz (2006).

On a more practical level, Bernard Rosauer (2013) defines organizational culture as an emergence - an extremely complex incalculable state that results from the combination of a few manageable levers:

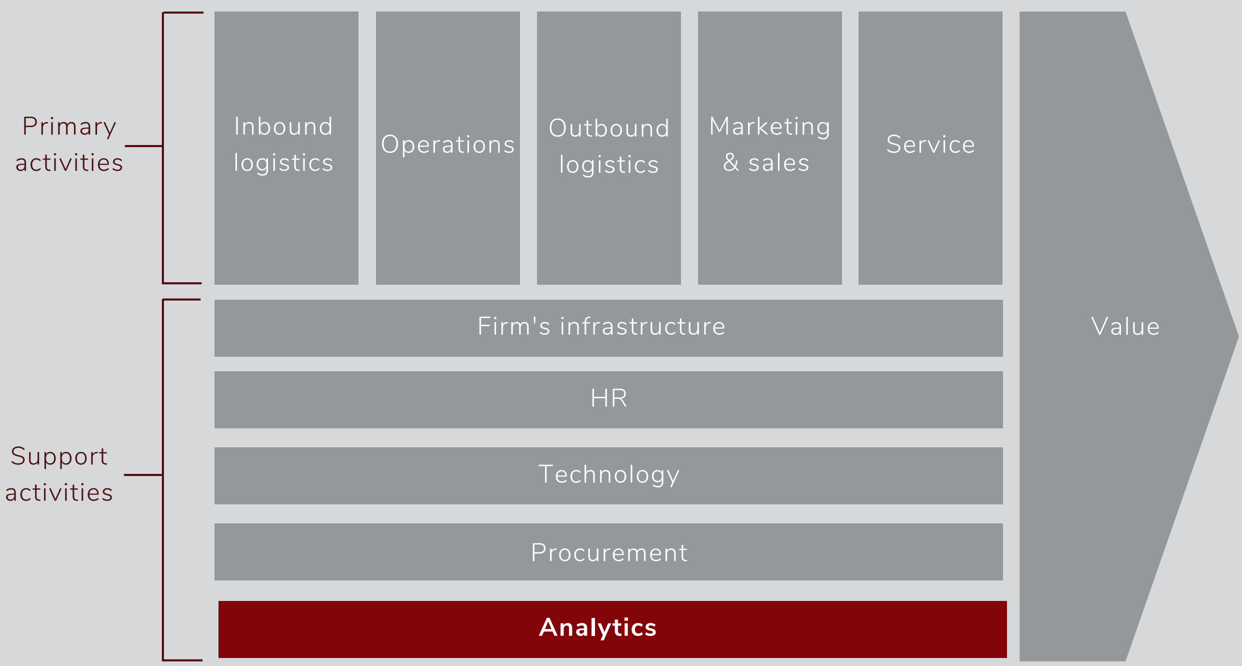
- Employee: focus on engagement
- Work: focus on eliminating waste to increase value
- Customer: focus on the likelihood of referral

To put Rosauer's levers into context, two concepts require further explanation: the value chain and lean thinking.

According to Michael Porter (1985), an organization can be disaggregated into strategically relevant activities that ultimately add value for its customers. Successful organizations either have a cost advantage or differentiate themselves from their competitors. An organization's value chain is typically part of a larger value system that includes companies either upstream (suppliers) or downstream (distribution channels). See Figure 2 for an illustration.

Lean thinking, introduced by James Womack and Daniel Jones (1996), focuses on increasing value for the customers by eliminating waste. In essence, value is anything the customer is willing to pay for, and waste is a manifestation of all non-value-adding activities to a product or service unless required. Examples of

Figure 2: Analytics as supportive activity in an organizations value chain (adapted from Michael Porter)



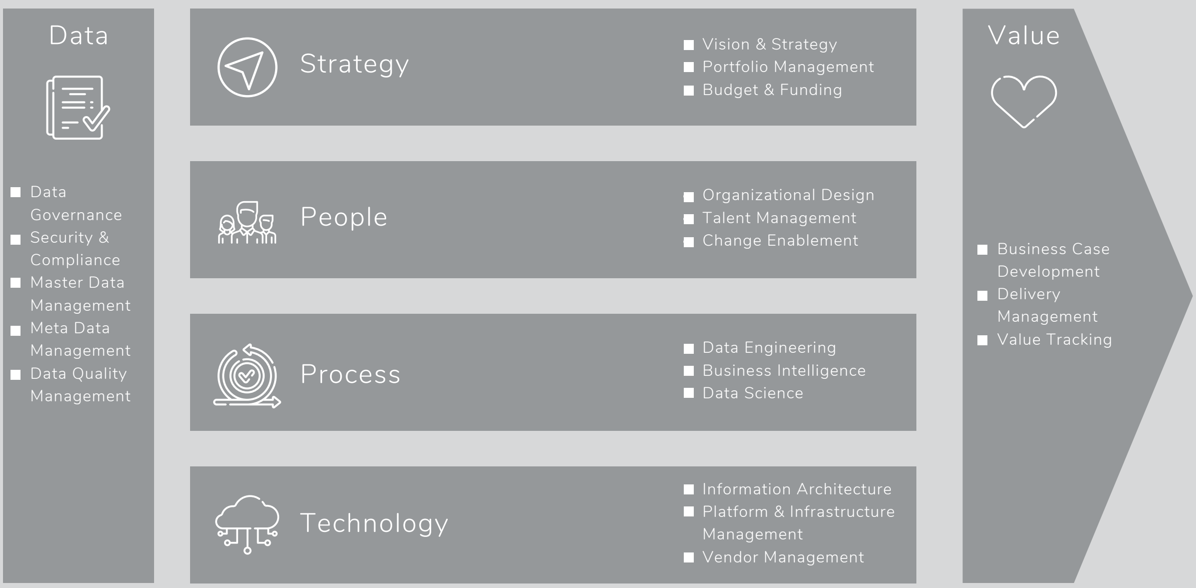
waste are waiting time, defects, and unnecessary transport. In the context of the value chain, analytics is a supportive activity that, using data as input, produces actionable insights to add or increase value. Building upon the work of Harold Leavitt (1964) and Bruce Schneier (2013), the analytics activities themselves can be viewed through the lenses of people, process, and technology.

Finally, to drive and coordinate cultural change, a clear strategic vision is required, as pointed out by Thomas Cummings and Christopher Worley (2004). The high-level analytics framework puts these perspectives together (see Figure 3).

Common pitfalls on the journey to becoming data-driven

The high-level analytics capability framework allows us to review the current analytics capabilities from different perspectives. A well-known framework in the field of data management is DAMA-DMBOK (DAMA International 2017). The presented analytics capability framework has some overlap with DAMA-DMBOK but has a different focus. Where DAMA-DMBOK focuses on governing data as an (enterprise) asset, the analytics capability framework emphasizes on adding value by putting data into effective use through analytics. By studying the

Figure 3: High-level analytics framework



different lenses of the analytics capability framework, we can now explore some typical pitfalls holding back organizations to become data driven.

The following list is by no means exhaustive. An in-depth assessment of an organization's analytics capabilities is beyond the scope of this whitepaper. However, based on nearly twenty years of experience, the following pitfalls appear to be common themes. The next paragraphs explore these five pitfalls and introduce re-commendations on how to overcome them.

- Strategy & Value: Focus on the means instead of the end.
- People: Organizing teams in silos.
- Process: Lingering in the lab approach.
- Technology: Failing to enforce technology standards.
- Data: Improving data quality on the fly.

Pitfall 1 - Focus on the means instead of the end

Especially in the earlier stages of data maturity, such as data-resistant or data-curious organizations, the organization's leadership might be wary of analytics initiatives. Analytics brings a more rational view that might conflict with personal agendas. The organization might also have a track record of lengthy and costly analytics projects, resulting in decision-makers being skeptical of new initiatives. New trends and technology might be exciting and inspiring for technology enthusiasts and analytics professionals, but hardly address the concerns of executives managing a tight budget, or who have ambitious growth targets.

To reach common ground, the key directive is to focus on adding value, as taught by Porter's value chain. By understanding the organization's vision and strategy, one should identify the key business objectives where analytics can contribute the most value. When done as a joint exercise with business leads, the outcome should outline a tangible, actionable analytics strategy and roadmap. Popular formats to capture the essence on a one-pager are VSEM (Vision, Strategy, Execution, and Metrics) and OGSM (Objective, Goals, Strategies, and Measures).

Pitfall 2 - Organizing teams in silos

Analytics requires all kinds of (new) disciplines, such as data engineering, data science, and data visualization. People with the right skills and experience are in high demand and can usually choose from multiple job offers, making it difficult to attract talent. Perhaps an even bigger challenge is to retain these professionals. Research from Robert Half (2022) indicates employees are not motivated by compensation alone. Instead, they seek fulfillment and recognition. A lack of impact, data sets being unavailable, or outdated tools can all contribute to dissatisfaction and finally resignation.

Drawing from lean thinking, a key cause of waste is handover. Examples might be waiting for a data source to become available for data ingestion, or a reporting specialist waiting for a new attribute to be added to an existing dataset. The traditional functional organization inadvertently strengthens these barriers. Recognizing this effect in software development, the Agile approach has gained significant momentum in the industry. One of its key principles is to put trust in self-organized teams. Once the high-level objectives and business priorities have been established, analytics teams should be empowered to drive the initiatives themselves. Scrum is a popular approach to organizing effective development teams. Its practitioners' guide (2020) suggests teams should consist of no more than ten people. A Scrum team is a cohesive unit of cross-functional professionals, meaning the members jointly have all the skills necessary to create value. Linking these teams to prioritized initiatives of the analytics roadmap ensures focus on adding value.

Pitfall 3 - Lingering in the lab approach

During exploration and experimentation, it is encouraged to start small and to use prototyping. A lab approach often works well with small datasets and the limited computing power provided by a typical workstation. However, such an approach often scales poorly. Machine learning models deployed to production must adhere to security standards, be resilient, and be maintainable, which includes proper documentation and version control. Metrics that are part of a self-service analytics environment might suddenly drift over time due to unanticipated changes to the source data (e.g., slowly changing dimensions).

Fortunately, data ingestion, data processing, and workflow orchestration can be highly standardized and automated. By exploiting the fact that such steps use for example SQL, Java, or Python, one can manage the source code in a version control system. Adding steps to automatically test the individual steps will eventually put the guardrails in place to trigger automated deployments for each code revision. The so-called continuous integration/ continuous deployment (CI/CD) cycle is a key enabler for automated, scalable data products. Lastly, advancements in software containers and container orchestrators allow for the scalable deployment of machine-learning models, with uninterrupted upgrades in production.

Pitfall 4 - Failing to enforce technology standards

Technology advances at a tremendous pace in the field of analytics tools and technology. Not long ago, distributed file systems based on Hadoop were considered state-of-the-art. Nowadays, all major cloud providers offer software-as-a-service (SaaS) for virtually unlimited amounts of cloud storage. Right sizing an on-premises Enterprise Data Warehouse used to be a key aspect of the design and implementation of such systems, while nowadays the only limit is the size of your annual budget (and proper cost control). In the past, several tools focused on providing integrated Extraction, Transformation, and Load (ETL) functionality, whereas today many organizations opt for separation of Extraction & Load from Transformation (ELT).

Organizations with multiple local analytics initiatives, or organizations that have grown through multiple mergers and acquisitions, might see different technology stacks and multiple standards in various stages of maturity. Fragmentation makes it more difficult to maintain and service the various analytics applications. It also hinders collaboration. Investing in standardization, whether it is in discoverability, information exchange standards, or preferred tooling reduces the amount of technical debt, thus freeing up resources to work on development instead of maintenance.

Pitfall 5 - Improving data quality on the fly

As established earlier, data is the key input for providing analytics services and products. Although it is tempting to cleanse and improve data on the go, in the end, it is more beneficial to improve the data directly in the source. DAMA-DMBOK provides an exhaustive framework with common vocabulary and best practices to actively govern data. One of the recommendations is to assign data ownership to business leaders, who should directly benefit from improving the data. In this approach, data owners are helped by data stewards and data quality specialists, who actively identify areas of concern and provide recommendations to improve data quality.

Assigning ownership sounds simple enough but is a key driver for business adoption and growth in data maturity. Once more, it helps to quantify problem areas and to turn them into business opportunities. Thomas Redman (2017) estimates that poor data quality costs between 15% to 25% of revenue for most organizations. Like other analytics applications, data quality specialists require access to collected and refined data. Once made available, they can deep-dive into potential causes of data quality issues and suggest recommendations to address them.

How to move forward

Many organizations seem to recognize the potential value of data and analytics but struggle to adapt to the required change. By unveiling the various capabilities of an analytics organization, we have demystified the various levers that influence the data maturity of the organization. Becoming a data-driven organization should not be a goal on its own but should be seen as a journey that organizations can follow to turn data and analytics into a profit center.

Squadra Analytics has proven experience in helping organizations accelerate their journey to become data driven. By reviewing the organization's objectives, we usually set the analytics ambition by jointly studying and selecting relevant value cases. Using a fact-based approach, we then carefully assess the organization's analytics capabilities and identify key improvement areas needed to realize the identified objectives. Finally, we jointly shape an action plan to move forward. Learning from successful innovators, we advocate the principle of "think big, start small, learn fast". Reach out to the author of this whitepaper to discuss how Squadra can help you.

About the author

Mark Dumay is a partner of Squadra Analytics, a boutique consulting firm based in the Netherlands. With nearly 20 years of experience, Mark has helped numerous clients in Utilities, Telecom, and Wholesale to get more value from data and analytics. You can reach him at mark.dumay@squadra.nl.

About Squadra

Squadra helps organizations to innovate efficiently using data to achieve their goals. We do this in an entrepreneurial and relationship-driven way, by supporting the establishment of a data foundation, creating value from this data, and organizing leadership.

References

Anaconda. 2020. "The State of Data Science 2020. Moving from hype toward maturity." https://www.anaconda.com/state-of-data-science-2020.

Cummings, T. and Worley, C. 2004. Organizational development and change. Cincinnati, OH: South-Western.

DAMA International. 2017. DAMA-DMBOK. Data management body of knowledge (Second edition). Basking Ridge, New Jersey: Technics Publications.

Davenport, T. H. and D. J. Patil, (2012, 90 10). 2012, 90 10. "Data Scientist: The Sexiest Job of the 21st Century." Harvard Business Review 70–76, 128.

Half, R. 2022. The Robert Half Blog. 10 July. https://www.roberthalf.com/blog/management-tips/effective-employee-retention-strategies.

Henke, N. and J. Levine, P. McInerney. 2018, February 5. "You Don't Have to Be a Data Scientist to Fill This Must-Have Analytics Role." Harvard Business Review.

Knight, M. 2021. What Is Data-Driven? 31 March . https://www.dataversity.net/what-is-data-driven/.

Leavitt, H. J. (1964). 1964. "Applied Organization Change in Industry: structural, technical and human approaches." In New Perspectives in Organization Research, by H. J. Leavitt & M. W. I. Shelly (Eds.) W. Cooper, 55-71. New York: John Wiley.

NewVantage Partners. 2022. "Data and AI Leadership Executive Survey 2022." https://www.newvantage.com/_files/ugd/e5361a_ad5a8b3da8254a71807d2dccdb0844be.pdf.

Penn, C. S. 2019. The Evolution of the Data-Driven Company. 8 August. https://www.christopherspenn.com/2019/08/the-evolution-of-the-data-driven-company/.

Porter, M. E. 1985. Competitive advantage. Creating and sustaining superior advantage. New York: The Free Press.

Ravasi, D. and Schultz, M. 2006, 49 3. "Responding to organizational identity threats: Exploring the role of organizational culture." Academy of Management Journal 433–458.

Redman, T. C. 2017, November 27. "Seizing Opportunity in Data Quality." MIT Sloan Management Review.

Rosauer, B. L. 2013, January. "Three bell curves. Business culture decoded."

Schneier, B. 2013. People, Process, and Technology. 30 January . https://www.schneier.com/blog/archives/2013/01/people_process.html.

Schwaber, K. and J. Sutherland. 2020, November. The Scrum Guide. The Definitive Guide to Scrum: The Rules of the Game.

Womack, J.P. and D.T. Jones. 1996. Lean Thinking - Banish Waste and Create Wealth in your Corporation. New York: Free Press.

Squadra

Data-Driven Value Creation