Paper 4704-2020

Language Lessons for Data-Driven Decisions

Achieving Data Literacy

Emma Warrillow, Data Insight Group Inc.

ABSTRACT

As they collect more and more data, organizations everywhere are pushing for more data-driven decision-making. They are hiring data scientists, data engineers, and data analysts to help them harness that data and reach this goal. However, while there is no question that technical programming and statistical skills are critical inputs in this endeavor, they are not enough. When you travel abroad, a phrase book or Google Translate® can help you navigate—as can learning a few key phrases—but you would not confuse this with knowing the language. The language of data literacy is no different.

Speaking the language of data is a skill you need to cultivate like any other. Whether you are the traveler (aka an analytical or data professional) trying to make yourself understood, or the person trying to help those visiting, learning the language is critical. And it is especially so when you are trying to make key business decisions; getting it wrong will result in more than just the wrong sandwich or a detour.

Gartner describes this ability to derive meaningful information from data as Data Literacy. This paper aims to explain why Data Literacy matters now and what leaders should do about it.

INTRODUCTION

Language. It's a powerful tool. It's how we communicate with each other, share ideas, and tell stories. It allows us to ask questions, solve problems, and think about how we interact with the world. It is shared, but it is also varied. Two people, speaking different languages, may be able to get by, but they lose the depth and clarity that comes from mutual understanding and vocabulary.

In business, as in life, language is important. Having employees, leaders, and executives speak the same language is critical to an organization's success.

Today, it is the language of data that matters most.

Driven by rapid technological change and the seemingly endless quantities of data that organizations have access to, data literacy is now a requirement for all employees. It is no longer enough for a few Data Scientists, working under the IT department, to know how to interpret and utilize data. Going forward, everyone, in all departments and functional areas, needs to be confident incorporating data into their decision making.

This paper will introduce and define the concept of data literacy and demonstrate why now is the time for senior executives to focus on improving data literacy rates across the entire organization. It will then discuss how organizations can facilitate this learning so that employees at all levels and skillsets can use, interpret, understand, and communicate with data.

DEFINING DATA LITERACY

Data literacy is the ability to collect, manage, evaluate, and apply data in a critical matter.[1] Data literate individuals can ask and answer relevant questions using data, interpret, understand, and question the results of data analysis, and put these results into the context of the organization's larger strategy and objectives.

"Data literacy is the ability to collect, manage, evaluate, and apply data in a critical matter."

More than being able to read the numbers themselves, data literacy is about finding the meaning behind the numbers. This requires an individual to dive deeper into how and why the data was collected, the problem that it was attempting to solve, how it was analyzed, and how it was visualized and presented. Data literacy also means that individuals can effectively share information with others in the organization, identifying the most important and relevant insights to drive stronger, more data-driven decision making.

While a relatively new term, organizations have been quick to identify that data literacy is a critical competency. However, it is also an area where most organizations and their employees fail to excel – in part because the education system has not caught up to the demand (and data literacy has rarely been taught as a matter of course). This paper will establish some of the reasons for its growth in importance and look ahead to how organizations can successfully navigate the barriers and build up the data literacy levels in their employees to best position themselves for the future.

WHY DATA LITERACY MATTERS NOW

The ability to interpret information and make decisions based on all the information available has always been central to an organization's success. Today, however, the type of information available is rapidly changing. Whereas in the past much of it was experience and a few key metrics, today it must be pried out of vast data stores. As data plays a larger role, not only in business, but in the world, the time to start building data literacy capabilities is now. Several factors have propelled data literacy to its position as a critical capability.

THE SPECIALIZATION AND CENTRALIZATION OF ANALYTICS

Historically, data analysts have been scattered across the organization supporting a variety of business needs and functions. They worked closely with business leaders when providing reports or building models, and they brought a varied range of skills and abilities to their role.

In recent years, however, organizations have began shifting away from this approach, centralizing their resources into an Analytic Centre of Excellence that supports all business units. According to IDC, 93 percent of executives in the US say their enterprise is leveraging some form of centre of excellence to drive artificial intelligence (AI) and data science initiatives, [2] while Gartner states that creating a centre of excellence is a best practice as organizations develop their data capabilities.[3]

"According to IDC, 93 percent of executives in the US say their enterprise is leveraging some form of analytics centre of excellence."

While this approach does offer many advantages, it can also serve to isolate the deep analytics expertise from the business itself. As data scientists become removed from business units, they risk losing the context of how their models align with the overall strategy.

In addition, within these centres of excellence, individuals are specializing in specific areas of analytics. Rather than the generalist analytic teams in the decentralized model, organizations now have experts in data visualization, text mining, machine learning, AI, and data engineering, and these experts sometimes even further specialize in a specific type of data or technique.

And as they become more specialized and build deeper technical skills, they increasingly need to rely business translators to define the problem and interpret and extrapolate their results.

THE DEMOCRATIZATION OF DATA

In parallel to this trend of analytic specialization, data is simultaneously being made available to more people than ever before. Self-service platforms, portals, and other business intelligence tools have made data accessible to a wide variety of users. Today, each department has its own metrics, key performance indicators (KPIs), and insights that drive decision making and put more information into the hands of people who need it.

However, this democratization of data puts the onus on the employee to be comfortable using and understanding it. Unfortunately, the reality is that many people lack the data literacy skills required to work with so much information. In a recent survey, a mere 25 percent of respondents said they feel prepared to make use of data, while only 37 percent think their decisions are made better with data. Most tellingly, 74 percent feel overwhelmed when working with data in any capacity.[4]

"74 percent of employees feel overwhelmed when working with data in any capacity."

Though the rollout of self-service tools has made data more visual and easier to read, most have done little to boost the data literacy of the users who are viewing this information. As a result, users either ignore the data in their decision making, or worse, misunderstand or misinterpret data and make decisions that go against the organization's interests.

Perhaps not surprisingly, much of the data literacy education movement has been led by data visualization vendors who provide the software to support the self-service solutions.

In addition to the rollout of self-service tools, some companies are going further to deliver insights in new and better formats. PepsiCo, for example, has shifted away from PowerPoint decks and Excel spreadsheets in favor of a system of apps, experiential activations, and quizzes. The goal is to help users understand and digest large amounts of data through visualization and storytelling, and to foster data literacy at all levels by making learning more fun and data more inspiring and exciting.[5]

THE DATA DELUGE

A recent IDC report predicts that the Global Datasphere will grow from 33 zettabytes in 2018 to 175 zettabytes by 2025.[6] With so much data, it's no surprise that organizations are

bombarded with more than they can possibly handle. Forrester showed that between 60 and 73 percent of all enterprise data is never even analyzed.[7]

"Forrester showed that between 60 and 73 percent of all enterprise data is never even analyzed."

Despite this deluge, the majority of organizations believe their data is an asset that they can use to their advantage.[8] In most cases, the technology exists to help them handle the volume of data. What's missing are the skills needed to drive the organization forward. Gartner reported that "poor data literacy" was the second biggest internal roadblock to success, behind the cultural challenges that come with making the large scale changes necessary to become a data-driven enterprise.[9] Improving employees' confidence and comfort using data addresses both of these barriers.

THE RISKS OF GETTING DATA WRONG

Data is only as useful as the ability to understand and use it, and the best models mean nothing if the insights cannot be incorporated and implemented. As more decisions become influenced by data, and organizations become more reliant on these decisions, the risks of getting it wrong become more significant. Corporations may decide to invest in expensive programs, divest of key assets, or realign a workforce on the basis analytic insights; so, it is crucial that these insights are correct.

The data literacy of the decision makers and their analytic resources is the key to mitigating this risk.

For example, data literate people understand the difference between causation and correlation. Perhaps a marketer notices that members of their loyalty program purchase more than most customers. This could mean that the loyalty program is stimulating more purchases, or it could reflect the fact that those that sign up for the loyalty program were *already* committed customers. Or, it could even mean that, because the loyalty card is what ties purchases together, the model is attributing more to those customers. Before committing more budget to the loyalty program, it will be critical to know which of these conclusions is correct.

Another area that can be addressed through improved data literacy is bias. One of the most common forms of bias is confirmation bias, where people tend to accept information that agrees with their pre-existing view and discredit information that conflicts with it. Other forms of bias include an overreliance on outliers, selection bias, or availability bias, all of which have the potential to influence how data is presented and interpreted. By being aware of the possibility of bias the user will ask more questions to ensure the results are reliable.

Finally, data usage must be viewed with an ethical lens too. Too often, data is thought of as an abstract, with models just a technical challenge or puzzle to be solved. In reality, there can be a lot of personal and sensitive information built into the data. Data literacy empowers people to ask how the data was collected, whether it should be used for certain purposes, and how it should be handled and stored to avoid ethical or regulatory issues. This is becoming a bigger concern for more sophisticated organizations, with 60 percent of companies with more than 20 data scientists expected to require a professional code of conduct that incorporates the ethical use of data by 2023. [10]

BUILDING DATA LITERACY

Like any new language, becoming fluent in data will not happen overnight, nor will everyone reach the same level of proficiency within an organization. Data scientists and data analysts who build models and deal with data every day will require the highest levels of data fluency, while business leaders and front-line employees will focus more on how they can apply the insights and results into their day-to-day roles. There are several steps organizations should take to build more data literate workforces.

FACILITATE LEARNING AND EVALUATE PERFORMANCE

Though data is playing a larger role in many different areas, academic institutions have lagged at teaching data literacy skills. As a result, the responsibility for learning has shifted from these schools to employers and the employees themselves. [11]

It is critical that everyone in an organization understands the need for data literacy and takes steps to improve their own fluency. This should happen both through organization development programs and self-guided learning, with the goal of increasing the organization's overall data literacy.

Training and development opportunities with the purpose of upskilling employees have become priorities at leading organizations who see the value in improving data literacy. In 2018, Gartner predicted that 80 percent of organizations would acknowledge their deficiency and initiate deliberate competency development in the field of data literacy by 2020. [12]

"Gartner predicted that 80 percent of organizations would acknowledge their deficiency and initiate deliberate competency development in the field of data literacy by 2020."

Beyond creating their own in-house learning programs, organizations are also supporting employee learning through tuition reimbursements or other incentives. Fortunately, universities and colleges are finally adding this competency to their programs to close the gap.

Once these learning opportunities exist, organizations can track the performance of employees over time by developing metrics and measuring their improvements. From this, they can identify gaps and adjust their training programs to benefit employees at all levels.

One large insurance company, for example, made significant investments in learning and talent development programs to prepare its workers for what it sees as the future digital economy. Along with creating a consolidated data lake and making data accessible to everyone in the organization who required it, they built online, in-person, and blended training programs focused on data science, data analytics tools, and predictive analytics. The increase in proficiency allowed Data Scientists to work closely with other professionals throughout the organization, enabling the company to capture more business and better assess risk when selling new policies.[13]

ESTABLISH A COMMON LANGUAGE AND VOCABULARY

Like learning a new language, data literacy can only be possible if there is a shared meaning and understanding of key terminology. Establishing this common language is therefore a necessity for any organization building its data literacy capabilities. Each organization will

develop its own dialect, so to speak, as words take on meanings that reflect the company's industry, purpose, and business strategy. In general, however, there are three common areas that should be articulated when creating a common vocabulary.

Statistical and Analytic Concepts

These are terms based in mathematics that relate to how data is collected, used, and interpreted. They are often thought of as the domain of data analysts and data scientists, but everyone in the organization must be able to use and understand these terms to effectively communicate with data. Examples include knowing the difference between mean, median, and mode, or understanding correlation and causation.

Domain and Industry Knowledge

This type of language is typically considered to be relevant to business units and decision makers, but just as everyone should know statistical concepts, Data Scientists should understand the context of their analysis within the larger business. This type of language may include terms specific to functional areas, such as marketing or finance terms, or common industry terminology.

Organizational Dialect

These are terms used specifically by an organization to define and communicate KPIs and metrics. It includes, for example, how to differentiate between a Marketing Qualified Lead and a Sales Qualified Lead or how an organization measures customer attrition rates or employee turnover. Understanding these terms ensures everyone is driving toward the same objectives and allows the organization to define what's important based on their unique business goals.

While it is often true that newly trained graduates may be armed with a stronger understanding of the analytic concepts, longer-tenured employees are typically much better versed in the dialect. Showcasing the value of all three components helps build mutual respect along this journey.

LEAD BY EXAMPLE AND BUILD A DATA DRIVEN CULTURE

In one survey, more than 85 percent of respondents reported that their firms had started programs to create data-driven cultures, but only 37 percent reported any success. Management understanding, organizational alignment, and a general organizational resistance to change were the primary reasons cited for this.[14]
Improving data literacy and becoming a truly data-driven organization means employees must change their behaviours and ways of thinking. People are inherently resistant to change, especially if the benefits of using data are not immediately clear or if they aren't incentivized to do so.

To overcome these challenges, organizations need someone who can lead by example and serve as a Data Visionary. This is a senior leader who sees the value of data and can look outside the organization for inspiration on how to use it. They should act as a champion for the use of data and constantly be communicating the importance of data literacy to employees.

In many cases, this role falls on the Chief Data Officer (CDO). In 2012, just 12 percent of Fortune 1000 companies had a CDO. By 2018, 68 percent of surveyed firms reported having a CDO.[15]

To embed data into the organization's culture, executive leaders must walk the walk. Daily conversations, board meetings, and team calls should discuss data and its role in the organization. Decision makers should constantly think about the data they have and the data they need when solving a strategic problem. And new initiatives should be justified and explained through the data and insights that led to them. Without reinforcement and encouragement from the top, organizations will find that enthusiasm around data literacy will fade and dissipate as people return to the way things have always been done.

CONCLUSION

As the world heads unceasingly toward a data-driven future, organizations must ensure they have the skills and capabilities necessary to effectively collect, analyze, interpret, use, and communicate data. Like any new language, improving data literacy rates amongst employees will take time, effort, and practice, but the benefits of doing so far outweigh the costs of initiating these programs and upskilling employees.

Data literacy ensures that employees at all levels are confident and comfortable using the data that is already available to them through self-service platforms and dashboards. It allows Data Scientists to get the most value from data and to work closely with business units and algin their models with the organization's strategic goals; and it empowers leaders to incorporate the right insights into their decision-making process. Through it all, organizations must encourage and facilitate this improvement and build a culture that rewards the use of data in all aspects of the business.

All employees don't need to become Data Scientists. They don't need to be experts in programming languages or know how to build models and select variables. They do need to be able to ask informed questions about insights, communicate results to their team, and identify information that is most relevant to their specific problem. Improved data literacy benefits everyone, and organizations must ensure their employees have the skills needed to operate in a data-driven world.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Emma Warrillow
Data Insight Group Inc. (DiG)
(416)699-5730 x 101
emma@datainsightgroup.ca
www.datainsightgroup.ca

REFERENCES

[1] http://dataliteracy.ca/wp-content/uploads/2016/04/Strategies-and-Best-Practices-for-Data-Literacy-Education.pdf
[2] https://www.cio.com/article/3391579/analytics-success-starts-with-a-center-of-excellence.html
[3] https://www.gartner.com/en/newsroom/press-releases/2019-07-15-gartner-survey-reveals-leading-organizations-expect-t
[4] https://thedataliteracyproject.org/files/downloads/Qlik_Accenture_Human_Impact_of_Data_Literacy.pdf
[5] https://digiday.com/marketing/pepsico-sweetens-consumer-insights/
[6] https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf
[7] https://www.accenture.com/us-en/insights/technology/closing-data-value-gap
[8] https://thedataliteracyproject.org/files/downloads/Qlik_Accenture_Human_Impact_of_Data_Literacy.pdf
[9] https://www.gartner.com/en/doc/3860965-fostering-data-literacy-and-information-as-a-second-language-a-gartner-trend-insight-report
[10] https://www.gartner.com/en/doc/3860965-fostering-data-literacy-and-information-as-a-second-language-a-gartner-trend-insight-report
[11] https://hbr.org/2020/02/boost-your-teams-data-literacy
[12] https://www.gartner.com/smarterwithgartner/gartner-keynote-do-you-speak-data/
[13] https://joshbersin.com/wp-content/uploads/2019/10/Build_vs_buy_Bersin_1.0.pdf
[14] https://www.techrepublic.com/article/85-of-big-data-projects-fail-but-your-developers-can-help-yours-succeed/
[15] https://newvantage.com/wp-content/uploads/2018/12/Big-Data-Executive-Survey-2019-Findings-Updated-010219-1.pdf