



DATA SCIENCE CENTER
of **EXCELLENCE ASSESSMENT**
MOSAIC DATA SCIENCE | CASE STUDY

MACHINE LEARNING CASE STUDY



Data Science Center of Excellence Assessment

We designed and delivered a custom, extensive data science assessment to help this energy firm modernize.



BACKGROUND

Many large corporations have recognized that data science creates significant revenue-enhancement or cost-reduction opportunities, and yet do not fully understand what data science is. Historically, data science work has been the purview of experts such as industrial engineers, clinical researchers, market researchers, advanced business intelligence analysts, or stock market modelers. Recently, corporate firms have started hiring chief data officers to implement the strategic, scientific use of data. Many of these same firms are creating internal, general-purpose data science teams.

Standing up a successful data science team is a complex and error-prone undertaking. Data science differs substantially from ordinary IT or business intelligence work. The most effective data science practitioners are trained applied scientists – scientists with the training and intuition to reach practical business outcomes using scientific methods. Analytics are no longer a hope and a wish; all businesses need to adapt to the advantages data-driven decision making brings, or they risk being left behind. As seen in the figure below, advanced analytics are more challenging, but they also bring greater value.

Analytics Are No Longer a Nice to Have

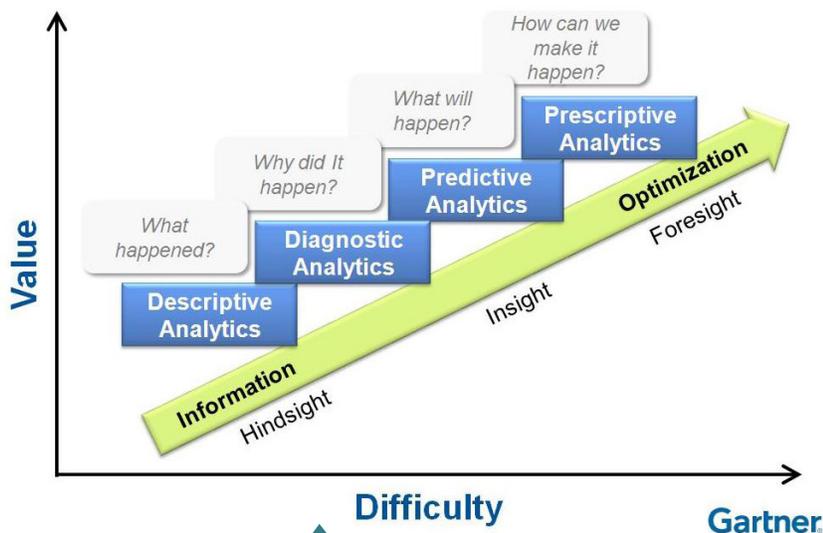


Figure 1
Gartner Analytics
Maturity Curve

A prominent energy generation and transmission firm was seeking to understand how data science and advanced analytics could help their business, as well as the level of investment and organizational change that might be necessary to take advantage of those opportunities.

The public utility wanted to focus on utilizing internal data for improved business decision making, optimizing their data analytics Center of Excellence (CoE) team structure, matching analytics technology with organizational fit, and convincing business stakeholders of the value and possibilities of advanced analytics.

Mosaic Data Science (Mosaic) was contacted to perform this enterprise-wide assessment. Mosaic brings over a decade of advanced analytics consulting experience and has helped many businesses ascend the analytics maturity curve. Due to some of the organizational challenges inherent with this far reaching assessment, Mosaic tapped one of its partners, [Evans Incorporated](#) (Evans). Evans brought 25 years of experience helping firms elevate their performance through strategic technology investments utilizing organizational change management and human factors engineering.

The Mosaic team designed and delivered an intensive, custom data science assessment keying in on the energy firm's current organizational capabilities and activities related to advanced analytics, identifying gaps and bottlenecks, and selecting/prioritizing projects that could generate business value with the appropriate level of investment.

ASSESSMENT FRAMEWORK

Mosaic decided a four-day onsite kick-off session at the energy company's headquarters was an appropriate start. Mosaic asked the company to invite stakeholders from the different business units and leaders from the analytics group to provide a baseline understanding of the objectives, opportunities, challenges and gaps as a foundation for the assessment.

After completing the initial kickoff session, the Mosaic team highlighted the following areas to deliver maximum value to the energy company:

- Best Practices for Data Analytics
- Summary of Expected Use Cases
- Data Analytics CoE Structure and Process
 - CoE overview and mission
 - CoE Roles, Responsibilities, Staffing and Required Skillsets
 - CoE Interaction with Business and IT
- Technology Selection and Design
 - Data Infrastructure Design (Data Lake/Warehouse)
 - Data Analytics Platform
 - Compute and Storage Infrastructure (Capacity Estimates)
 - Technology Platform Administration and Support
- Timeline and Cost Estimation for CoE implementation

ASSESSMENT EXECUTION

Team Mosaic recommended the energy firm implement a repeatable data analytics process. A workshop was designed and delivered to focus on best practices around analytics. This session involved interviewing the different business unit leaders, IT and analytics leaders as well as an overview of what data science is, how it is optimally applied, and a high-level review of technologies/processes best suited to meet those objectives. Team Mosaic proposed implementing the Cross Industry Standard Process for Data Mining (CRISP-DM, as seen in Figure 2). This process breaks the analytics solution development process into six steps; business understanding, data understanding, data preparation, modeling, evaluation, and deployment. CRISP-DM emphasizes the iterative nature of analytics solution development with all development rooted in a deep understanding of the business requirements and objectives.

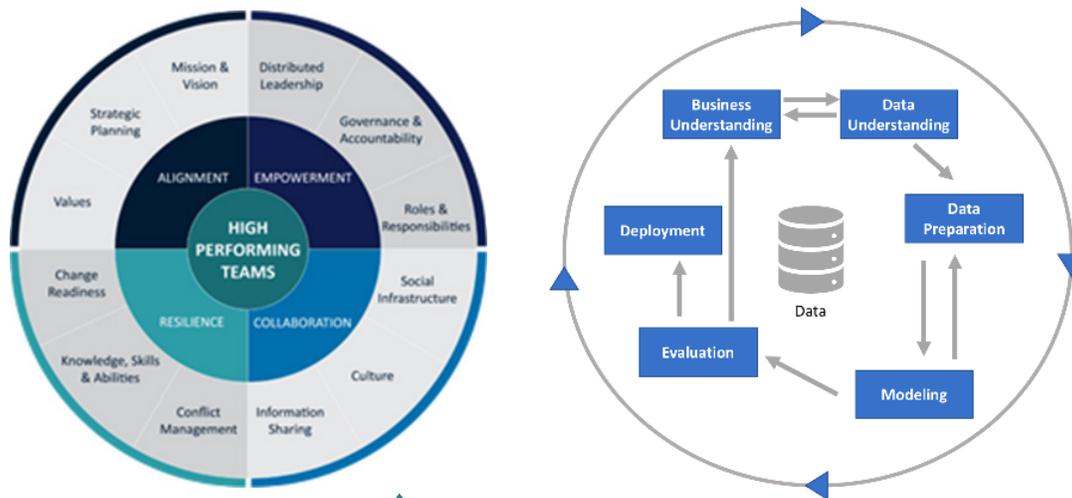
Armed with knowledge and business insights from the preceding phases, Team Mosaic established a set of guiding principles and built an optimal roadmap for people, process technology, and projects. The report delivered from this section mapped known advanced analytics skillsets and capabilities throughout the organization, cataloged current analytics toolsets in use, identified gaps with respect to realizing value from data science, highlighted recommendations and a roadmap for addressing gaps, and analyzed current feasibility of analytics use cases.

Following initial stakeholder interviews, Team Mosaic assisted the client in assembling a core team to serve as the conduit for continuing the discovery process, interacting with stakeholders, and collaboratively developing a data analytics strategy supported by stakeholders. This involved designing and delivering recommendations for a hybrid data management organization to include a CoE that would function in collaboration with the lines of business (LoB).

To ensure the CoE could provide centralized support for LoB analytics, the strategy needed to include a structure that enabled CoE staff and LoB analytics representatives to build strong relationships and collaborate effectively in support of achieving organizational goals. Team Mosaic drew from a proprietary method, High Performing Teams (HPT) (Figure 2) a model that focuses on four foundational elements:

- Alignment to strategy
- Empowerment to deliver results
- Collaboration to forge relationships
- Resilience to accommodate change

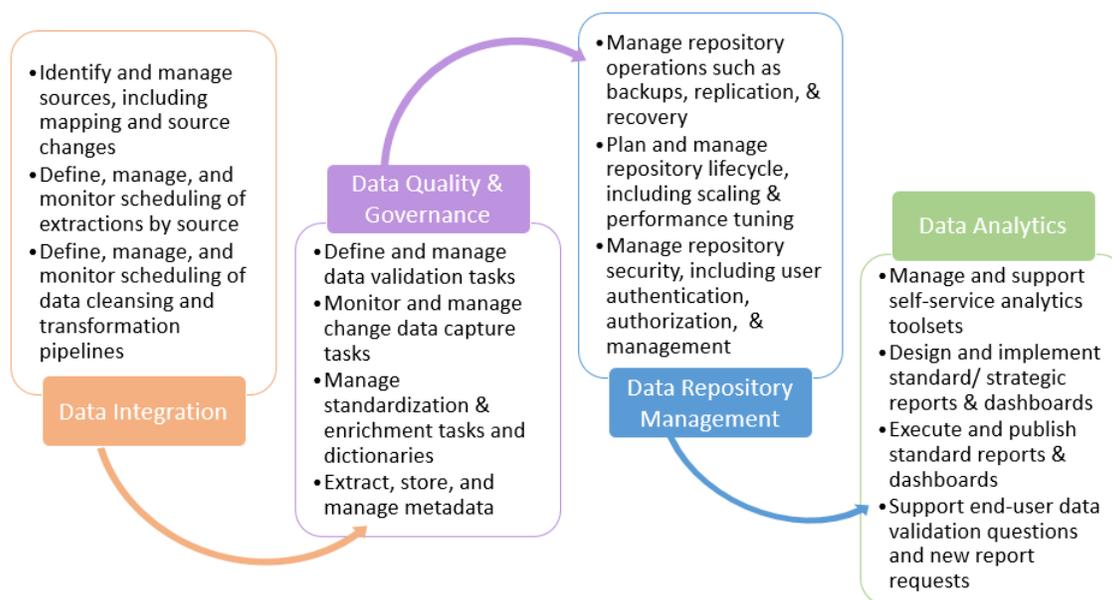
Most prominently, Team Mosaic built a strategy based on the Data Management Lifecycle (Figure 3). By leveraging this framework, the strategy not only focused on the functions and roles in scope for the CoE, but also defined those outside of the CoE, to address current pain points and ensure the business was mature in all data management phases necessary to enable advanced analytics.



*Figure 2
The pillars of a High
Performing Team and the
CRISP-DM Process Flow*

Recurring stakeholder sessions were held to collaborate on how to optimally structure the CoE and how CoE members could achieve maximum analytics value while interacting with the LoBs.

In parallel with defining the CoE structure and services, Team Mosaic also developed recommendations for technology to support the CoE.



The Mosaic team was able to provide an objective third-party view of the technology market since neither firm is tied to any specific platform. Mosaic champions open source technologies such as R and Python for firms just starting out in their analytical journey. R and Python are also great tools for mature organizations and can easily be utilized with almost any data platform.

The Mosaic team did not come in with a preconceived plan, technology package or recommendations. The approach was to collaborate with the business stakeholders from across the business, synthesize the information gathered, and then develop recommendations that would work within the constraints and budget of the business. The detailed implementation plan covered prioritized activities, implications/assumptions, timeline, risks, and cost estimates (both people and technology).

RESULTS

Upon completion of the rigorous data science assessment, the utility had a comprehensive plan to leverage data analytics to inform decision making across the business. They have a clear path forward on how to build a CoE, staff it with the optimal number and type of resources, and provide those resources with cutting-edge technology. Team Mosaic was able to draw upon its deep expertise, maintain an objective view of the utility's enterprise landscape, and deliver a powerful, return-on-investment focused engagement.

Figure 3

The different skillsets needed to satisfy each phase of the Data Management Lifecycle

FOR MORE INFORMATION

Want to learn more?

Please contact info@mosaicdatascience.com





ABOUT MOSAIC DATA SCIENCE

We provide innovative machine learning, AI and analytics consulting across organizations.

Mosaic is a leading data science consulting company focused on helping organizations build and deploy actionable analytics solutions. Our customers are as varied as the techniques we use — some just starting their first predictive analytics project; others with deep in-house machine learning expertise.

HOW WE WORK WITH YOU

We work in a highly collaborative partnership with our customers to ensure you get only the best results to consistently drive business value.



MACHINE LEARNING

We design and deploy predictive algorithms to solve the most challenging problems facing businesses today



ARTIFICIAL INTELLIGENCE

We bring a wealth of knowledge on how to tune AI models to deliver the maximum business value



BUSINESS ANALYTICS

We leverage technical expertise and experience across a swath of industries, bringing fresh approaches to challenging problems



mosaicdatascience.com



info@mosaicdatascience.com



(866) 202-8600