Business & Technology Report

June 2026

Data Science at Rural Electric Cooperatives

Part 1: Opportunities, Foundational Requirements, and Barriers

NRECA

America's Electric Cooperatives

Business & Technology Report
June 2026

Data Science at Rural Electric Cooperatives

Part 1: Opportunities, Foundational Requirements, and Barriers

---

AUTHOR AND SUBJECT MATTER EXPERT ON THIS TOPIC:

Lisa Slaughter
Research Software Engineer and Data Scientist
Energy Research & Resilience
Business and Technology Strategies
Lisa.Slaughter@nreca.coop

NRECA
America's Electric Cooperatives

Table of Contents

Overview of Data Science

Rural electric cooperatives operate in an environment that places increasing, simultaneous demands on planning, operations, and finance. Load patterns are changing, distribution systems are becoming more dynamic, and expectations surrounding reliability and resilience remain high. Yet budgets and staffing constraints limit the ability to pursue every improvement opportunity at once. In such a setting where the best path forward can be difficult to decipher, data science offers the allure of quick, accurate insights that can inform critical decisions.

Despite its popularity (or perhaps because of it), the term "data science" is often used loosely. In some discussions, it is treated as synonymous with machine learning or artificial intelligence. In others, it is used to describe almost any quantitative analysis. This vagueness can make it difficult for cooperative leaders and technical staff to determine what the methods of data science actually achieve, what kinds of questions it can realistically answer, and what is required to apply it effectively.

For electric cooperatives, the important task is not to find out whether data science will be valuable – when properly applied, it almost always is – but to identify what resources it will take to obtain that value. Some useful questions to ask in this regard are:

1. Where can a structured analysis improve existing decision-making processes?
2. What data is required to perform such an analysis?
3. What practical constraints affect the ability to perform this analysis?

In other words, the challenge comes in clearly identifying the relevant applications, data readiness level, and implementation specifics for a particular co-op.

This paper is the first in a four-part series that examines how electric cooperatives can incorporate insights from data science into their operations through a practical, grounded approach. This series is intended to move readers through broad understanding into concrete applications. It begins with a general introduction to what data science may mean in the context of electric cooperatives, then considers how to identify worthwhile opportunities, how to assess organizational readiness, and how to think about structuring analytics efforts in a way that aligns with operational realities. The aim of the series is to support clearer judgment about where data science is most useful, what resources are needed to properly apply it, and how co-ops might approach the work realistically.

This paper offers a practical introduction to extracting insights from data at electric cooperatives. It describes what data science means in this context, identifies several areas in which data science can support decision-making, and outlines the conditions that often determine whether a project will succeed.

Defining Data Science in the Cooperative Setting

As mentioned, many of the linguistic terms related to data science tend to be used interchangeably, which causes confusion as to what benefits can be derived. To that end, it is worth distinguishing between a few related terms that one might encounter out "in the wild":

- **Data engineering** makes data usable by collecting, cleaning, structuring, integrating, and maintaining information from different systems.
- **Analytics** examines data to identify patterns, trends, relationships, and performance insights that can support understanding for decision-making.
- **Machine learning** is a subset of data science and artificial intelligence that uses algorithms to learn patterns from data and produce outputs such as predictions, classifications, and recommendations.
- **Artificial intelligence** is the broadest term, referring to systems designed to perform tasks associated with reasoning, judgment, language, and perception.
- **Data science** applies data, analytical methods, and subject matter expertise to answer decision-relevant questions in a structured way.

The relationships between these terms are roughly visualized in Figure 1.

As shown, data engineering underpins the other activities by making data usable in the first place. Analytics, machine learning, and artificial intelligence can each respectively help identify patterns, evaluate new data in context of old, and tell a story. Data science plays an integrative role by incorporating all these activities under operational context.

In a utility environment, data science can be understood as the process of combining data with analytical methods and subject matter expertise to support decision-making processes.

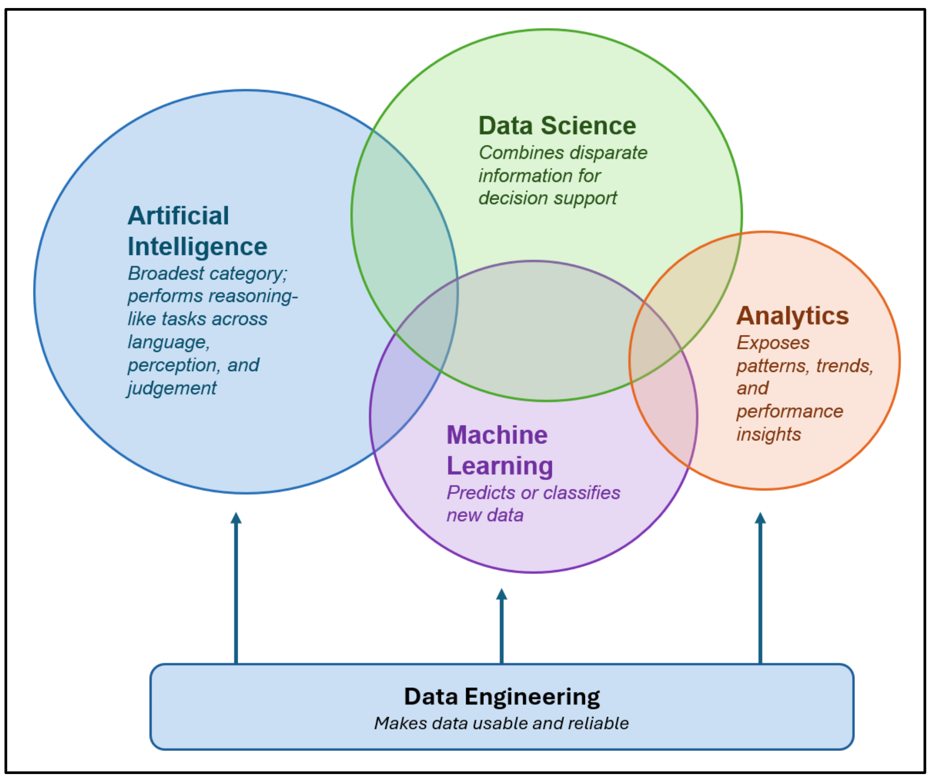

Figure 1: Conceptual diagram visualizing the relationships between the related terms data engineering, analytics, machine learning, artificial intelligence, and data science.

This definition involves a range of analytical approaches that have different levels of complexity, from simple curve-fitting to the application of artificial intelligence and machine learning methods. The decisions being supported by these analytical methods can be described in different ways:

- **Descriptive** work summarizes what has happened. Examples of outcomes are outage frequency by feeder, seasonal transformer loading behavior, or the distribution of truck rolls across the service area.

- **Diagnostic** work investigates why certain patterns appear. Insights obtained might include the reason that one class of assets fails more often than expected, or why restoration performance varies across similar circuits.

- **Predictive** work estimates what is likely to happen under comparable conditions, such as the probability of an equipment failure, the expected effects of weather on outages, or the timing of future peak demand.

- **Prescriptive** or optimization-oriented work supports choices among alternatives. Conclusions might indicate where to prioritize inspections, how to sequence certain investments, or where flexible demand might provide the greatest operational benefit.

Each of these framings involves a suite of different analytical methodologies that are used to extract insights from the data, depending on its characteristics. These analytical methodologies consist of many inventive mathematical algorithms, and it is possible for the algorithm selection process to consume a large amount of time if a data scientist does not stay focused. The selection of the analysis algorithm is indeed an important step to get right, yet it may not be the step that produces the most value. Significant insights can be obtained merely through structured data integration, visualization, or basic statistical analysis before more advanced machine learning methods become necessary.

The goal of framing the work is to match the most appropriate methodology to the problem being solved, and in many cases, the more basic the methodology, the clearer and more actionable the results. This falls well in line with the co-op decision-making approach, where the value of a grid solution is often defined less by methodological sophistication and more by the solution's clarity, reliability, relevance to operational needs, and the extent to which it clarifies priorities and supports practical action. Accurately and concisely framing the problem at hand can help ensure that the best approach is identified and that the results will be interpretable in a way that solves that problem. Data science is most useful when it is tied to a practical question, grounded in system knowledge, and interpreted by people who understand the operating environment. Without that context, even technically sound analysis can be difficult to trust or act upon.

Data science is most useful when it is tied to a practical question, grounded in system knowledge, and interpreted by people who understand the operating environment.

Areas Where Data Science Can Provide Value

The most relevant applications for electric cooperatives tend to arise in parts of the business that are already well-understood. In most cases, the value of data science comes not from novelty, but from enabling more consistent, better-informed decision-making. Here are some areas where informed analyses and data science can bring value to an electric cooperative's business.

- **Asset Management and Maintenance**

  Many cooperatives already have asset records, work history, inspection results, and failure data; the challenge is often integrating that information in a way that supports analysis and prioritization. Analytics can help systematically compare age, condition, failure experience, location, and consequences. It can also reveal incomplete or conflicting records, creating operational value before any models are even developed.

- **Reliability, Resilience, and Outage Management**

  Outage records, restoration information, patrol observations, and weather history can be analyzed to identify recurring outage patterns, high-risk equipment classes, feeders with disproportionate interruption exposure, and conditions associated with long restoration times. These insights can support maintenance prioritization, vegetation management, resilience investments, and post-event review. Even relatively simple analysis can improve prioritization where resources are limited.

- **Distribution Planning and System Operations**

  Historical load behavior, voltage data, asset loading, and switching history can help reveal emerging constraints and show where older assumptions may need to be revisited. This is especially relevant as distributed energy resources, electrification, and changing member behaviors continue to alter feeder conditions. Where system data and connectivity models are sufficiently strong, data science activities may also support switching studies, restoration planning, and evaluation of alternative operating configurations.

- **Member Programs, Demand Management, and Distributed Energy Resources**

  Member programs and distributed energy resource (DER) strategies also increasingly depend on analysis. Interval load data, participation records, and feeder-specific conditions can help identify which member segments and areas are most relevant, which loads appear most flexible, and where distributed resources may provide operational value. This is particularly important when a cooperative is considering targeted deployment rather than broad program availability. Analyses can help connect consumer-side behavior to system-side needs in support of peak management, resilience, and localized operational support.

- **Financial and Administrative Decision Support**

  Data science activities can also improve decision-making related to spending, workflow, and prioritization, including capital planning, power cost forecasting, field activity review, and repeated operational inefficiencies. In many cases, the most useful first step is a clearer understanding of where repeated work, avoidable costs, or persistent operational burdens are occurring.

Data and System Requirements

One of the most common misconceptions surrounding data science is the idea that value follows automatically from having large volumes of data. In practice, the usefulness of an analysis effort depends most on data quality, context, and structure.

Commonly relevant sources of data may include advanced metering infrastructure (AMI) data, outage records, supervisory control and data acquisition (SCADA) data, geographic information system (GIS) records, asset and work management systems, engineering models, weather data, and program participation records. However, these sources are only useful to the extent that they can be interpreted and related to one another with reasonable confidence.

To achieve this, several foundational conditions are important:

1. **Identifiers and naming conventions must be sufficiently consistent** to allow records from different systems to be connected by software. If the same device, location, or feeder appears under multiple names or incompatible IDs, for example, even basic integration becomes labor-intensive due to exceptions.

2. **Timestamps and temporal alignment affect outcomes**. Operational data drawn from different platforms are often recorded at different intervals, in different formats, and/or with inconsistent time handling. Batch data uploads from the field can impose significant delays on data updates. These issues materially affect the characteristics of the conclusions drawn.

3. **System context is usually necessary**. Feeder association, phase information, connectivity, and equipment configuration must often be documented with enough fidelity for the analysis at hand. A result that appears mathematically sound can still be operationally misleading if the underlying model of the system is inaccurate.

4. **Records must be maintained to reflect reality** closely enough to support decision-making. It is common for actual system conditions to diverge from utility records over time, especially when manually updated. This is not merely an inconvenience; In many cases, it is one of the main reasons a project becomes more difficult, more expensive, or less reliable than expected.

5. **Sufficient historical data may be required**, depending on the analysis. Some questions can be addressed with a relatively short window of recent data, but others require enough history to capture seasonal patterns, rare events, slowly changing operating conditions, or meaningful variations in asset performance and member behavior. Without an adequate historical record, the analysis may rest on patterns that are incomplete, unrepresentative, or too limited to support reliable conclusions.

These conditions are largely technical, but the barriers to data science go beyond the technical. Cooperatives may also face practical constraints related to staffing, internal capacity, and organizational adoption. In some rural areas, it may be difficult to recruit or retain staff with the right combination of data, software, utility, and communication skills needed to perform data science activities. Even when those skills are available, the value of data science still has to be socialized across the organization. Engineering, operations, member services, finance, field crews, and leadership may each understand the same data differently, and the same core story may need to be presented to each using a different lens. For this reason, readiness should be understood as both a technical condition and an organizational condition: the cooperative needs usable information, but it also needs people who can interpret and explain it in a way that connect with existing decision-making processes.

Conclusion

Data science can provide meaningful support to electric cooperatives across asset management, reliability, planning, member programs, and cost control. Its value, however, depends on the specificity of the question being asked, the quality of the underlying data, and the degree to which an analysis relates to operational realities.

For many cooperatives, the most important first step is not adopting the most advanced analytical methods available, nor is it purchasing an analysis platform. The best introductory work is in clarifying which decisions would benefit from better analysis and assessing whether the available data is sufficient to reasonably support those decisions. Where those foundations are strong, advanced methods may be useful. Where they are not, the immediate value may lie in first improving visibility, consistency, and data reliability. Either way, progress begins with the same principle: analytical work is most effective when it is connected to real operational needs and supported by trustworthy information.

Related Resource

- Data and Analytics Topic Website

Contact for Questions or Comments

Have ideas about how to gain value from your data or unsure how to get started with data science at your cooperative? Contact the author of this article:

Lisa Slaughter
Research Software Engineer and Data Scientist
Business and Technology Strategies
Lisa.Slaughter@nreca.coop
571.422.2756

To find more resources on business and technology issues for cooperatives, visit our **website**.

Legal Notice