Data Analysis & Knowledge Management

Understanding the distinct difference between data analysis and data management is essential for leveraging the power of data and its influence on mission accomplishment. This understanding will help to transform raw data into actionable knowledge, thereby supporting decision-making processes.

Definitions and Differences

Data Management involves the development and execution of architectures, policies, practices, and procedures to manage the full data lifecycle in an organization. This lifecycle includes data creation, storage, maintenance, use, and deletion. Data Management also supports the content management principles of making knowledge products visible, accessible, understandable, and reliable.[1]

Conversely, Data Analysis involves examining and modeling raw data to discover useful information, draw conclusions, and support decision-making. Data analysis can range from simple descriptive statistics to complex machine learning algorithms, but the core goal is always to extract insights from data.[1]

Data Management and Data Analysis have some overlapping aspects such as cleaning, transforming, and modeling, but they differ in focus, objectives, and processes. Data Management is centered on the control, governance, and optimization of data and information within an organization. Data Analysis is focused on using specific methods and techniques to interpret data and extract insights, draw conclusions, or make informed decisions. It's about deriving meaning and value from raw data.

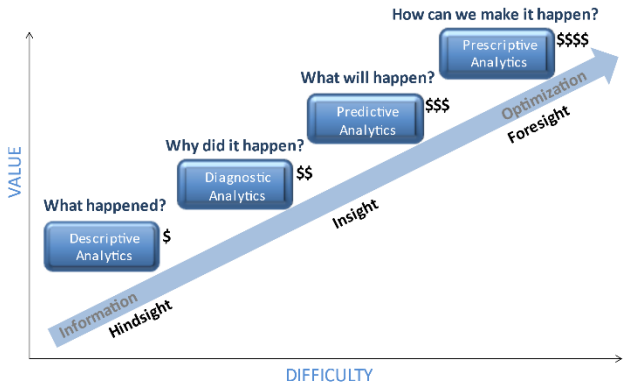
Benefits and Stages of Data Analysis

Data Analysis can drive strategic and tactical decisions, reveal patterns, and provide competitive advantage through the use of analytics. The four levels of data analytics are:

Descriptive Analytics: It answers, "What happened?" For example, reviewing last year's recruitment numbers to determine if there were any changes.

Diagnostic Analytics: It answers, "Why did it happen?" For instance, understanding why a particular training program has a high dropout rate.

Predictive Analytics: It answers, "What will happen?" For example, forecasting future recruitment trends based on historical data.



Prescriptive Analytics: It answers, "How can we make it happen?" For instance, recommending changes to improve a training program's retention rate. The more we are able to conduct predictive and prescriptive analytics as an organization, the more we can realize value from the outcomes.

Training and skills that are typically required

Data analysis requires a combination of trained personnel and skills. In terms of personnel, the size of the team required will depend on the scale and complexity of the data analysis needed. Small projects might be handled by one or two data analysts, while larger ones might require a team of analysts, possibly including data scientists, data engineers, statisticians, and business analysts.[1]

Statistical Knowledge/Mathematical Abilities: A strong foundation in statistics is key. This includes understanding probability, distributions, statistical tests, regressions, and more. Additionally, linear algebra, calculus, and other areas of mathematics are often used in advanced data analytics. For instance, probability is used to determine the likelihood of an action occurring.

Data Wrangling/Visualization Skills: Real-world data is often messy. Therefore, skills in data cleaning, data transformation, and data preprocessing are necessary. The ability to create clear, impactful visualizations is essential to communicate the results of data analysis. Knowledge in tools like Matplotlib, Seaborn, or Tableau can be beneficial.

Programming Skills/Machine Learning: Knowledge in programming languages is useful in some data analyses. For example, Structured Query Language (SQL) is commonly used for database interaction. , SAS and Statistical Product and Service Solutions (SPSS) have proprietary syntax for performing analysis. For more complex data analysis, a grounding in machine learning methods, both supervised and unsupervised, are very useful. Python and R are commonly used for machine learning, The ability to analyze complex data more efficiently and accurately can lead to quicker, more informed decisions, and better predictive analytics.

Critical Thinking: The ability to formulate the right questions and interpret results correctly is key to good data analysis.

Domain Knowledge: Understanding the context in which the data exists can be just as important as technical skills. This might require training or experience in a specific industry or field.

Training to become a data analyst can be acquired through formal education like a degree program in data science, statistics, mathematics, or computer science. Many universities also offer specialized postgraduate programs in data analysis or related fields. There are also many online platforms offering courses in data analysis, such as Coursera and Udemy, which provide comprehensive learning paths. Certification programs from professional organizations can also be a way to gain recognition and skills in this area.

Knowledge Management (KM) Role in Data Analysis

The role of KM in data analysis is to empower knowledge workers and leadership and provide them with the knowledge and understanding to make decisions.[2] A KM program plays a pivotal role in facilitating the flow from data to information and ultimately into actionable knowledge by:

Data Collection and Organization: Oversee the collection, organization, and categorization of data and information. This might include structuring data in a way that facilitates effective data analysis.

Data Quality: Ensure that the data being collected and stored in the organization's systems is of high quality, accurate, and relevant. This is crucial as the quality of data directly impacts the results of data analysis.

Facilitating Access to Data: Ensure that relevant data and information are readily available to those who need it, including data analysts. This can involve creating systems for data storage and retrieval, setting access permissions, and making sure data is accessible in a usable format for data analysis.

Knowledge Sharing: Responsible for disseminating the results of data analysis throughout the organization. This can involve creating reports, dashboards, or presentations and ensuring that insights derived from data analysis are communicated effectively.

Cross-Functional Collaboration: Work with other staff and teams to help facilitate collaborations and ensure that all relevant parties have access to the data they need.

Training and Guidance: Help train and guide staff in how to use data systems and tools. This can include training data analysts or other staff members in how to access and interpret data.

Conclusion

Having sound data management practices in place is important for an organization, but it is data analysis that allows an organization to maximize the utility of its data, turning it into actionable knowledge that supports strategic decisions.

This TRADOC OCKO publication was written by Nathan Truckenbrod. Want to learn more? Respond to this article or access related articles, blogs, media presentations, and more at https://www.milsuite.mil/book/groups/tradoc-km.

Notes:

1. Kelley, Karin, *What is Data Analysis? Methods, Process and Types Explained*, May 9, 2023, simplilearn.com, Mastering Data Analysis: Process, Types, Methods, and Techniques (simplilearn.com).
2. Army Techniques Publication 6-01.1, *Techniques for Effective Knowledge Management*, March 6, 2015, https://armypubs.army.mil/ProductMaps/PubForm/Details.aspx?PUB_ID=105059.