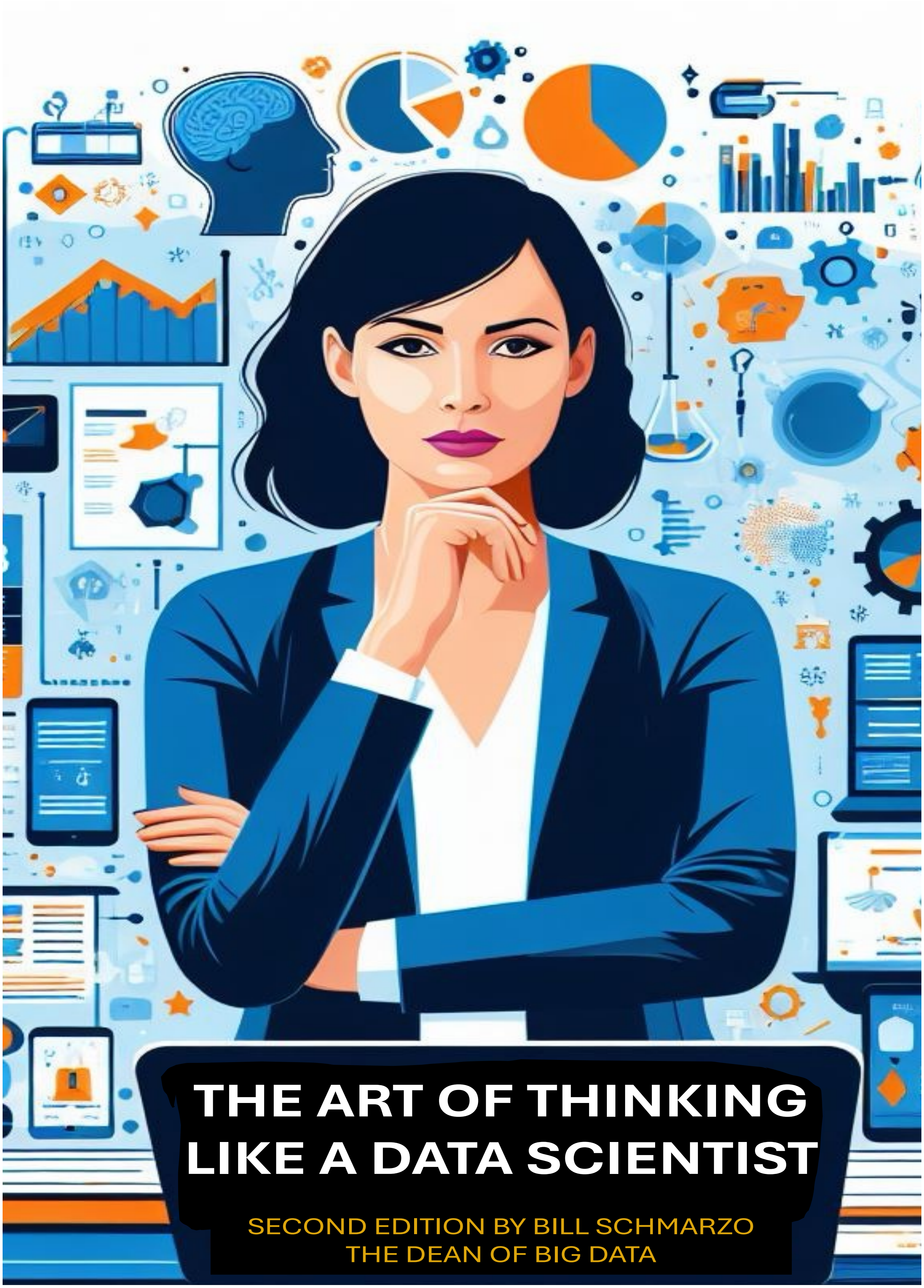

THE ART OF THINKING
LIKE A DATA SCIENTIST
SECOND EDITION BY BILL SCHMARZO
THE DEAN OF BIG DATA

"The Art of Thinking Like a Data Scientist" Second Edition
By Bill Schmarzo, The Dean of Big Data

How can we become more effective at leveraging data and analytics to power our business and operational models?

In my many years in the data and analytics industry, I've been confronted with this question countless times. Organizations lacked a methodology that drove the collaboration and alignment between essential business and data science teams to identify, design, develop, deploy, and manage the data and analytic economic assets that will be the sources for value creation in the 21st century. I created the "Thinking Like a Data Scientist" (TLADS) methodology to address that challenge and have gone on to teach this methodology at several universities and many organizations. While teaching this methodology, I have identified several tweaks and modifications that make it more effective in leveraging advanced analytics like artificial intelligence (AI) to help organizations get value from their data. This book reflects those learnings[1].

This book will focus on updating the TLADS methodology based on the experience gained in teaching and deploying this methodology. Part 1 represents an update to the original "The Art of Thinking Like a Data Scientist," which can be downloaded from www.DeanofBigData.com. The methodology steps have been refined and expanded, and the supporting design thinking canvases have been updated to complement the refined, expanded, and new TLADS steps.

📝 **Note:** A forthcoming supplemental workbook will integrate the TLADS methodology with a Generative AI (GenAI) tool such as OpenAI ChatGPT, Microsoft Copilot, or Google Gemini. The GenAI tool will provide an "exploratory sandbox" to uncover, validate, and flesh out the organizational, stakeholder, data, and analytic requirements required to produce more meaningful, relevant, responsible, and ethical AI outcomes.

A GenAI tool can be a powerful companion for anyone trying to explore and master the TLADS methodology. By providing detailed content, the GenAI tool can yield more accurate and relevant responses and provide the foundation for an exploratory and creative dialogue with the GenAI tool (think 6 Socratic Questions). I will share some helpful techniques for creating effective GenAI prompts leveraging the TLADS methodology. While these GenAI tools can improve TLADS productivity, their most significant benefit is that they stimulate our natural curiosity and imagination, ultimately enhancing the relevance of the TLADS methodology.

I hope the second edition is even more relevant and effective in empowering an organization of "Citizens of Data Science" collaborating to identify, define, design, and deploy AI models that deliver more relevant, meaningful, responsible, and ethical business and operational outcomes.

I'd like to express my gratitude towards Dan Blake and Jean-Philippe Martin for their meticulous review and constructive feedback on this book. They went through every concept, explanation, and word in the book and even spotted typos on images that had been there for years. Thanks to their guidance and valuable suggestions, the flow and clarity of the book have significantly improved.

Table of Contents

Chapter 1: Thinking Like a Data Scientist (TLADS) Introduction

This book is intended for anyone who needs to work with data, analytics (such as Artificial Intelligence, Machine Learning, and Deep Learning)[2], and data science teams to make the most out of data and analytics and deliver better business and operational outcomes that are both responsible and ethical. While data science teams can benefit from having a shared methodology to foster collaboration with business and operational stakeholders, the real value of this book will be for non-scientists who want to become "citizens of data science" and ensure that data and analytics are used to deliver outcomes that are more relevant, meaningful, responsible, and ethical.

The Thinking Like a Data Scientist (TLADS) process is a value-based problem-solving and data-driven decision-making framework. It helps organizations identify where and how data science can leverage the organization’s wealth of data to generate new sources of customer, product, and operational value.

The eight-step TLADS methodology blends data science, design thinking, and economic concepts to improve the organization’s effectiveness at leveraging data and analytics to deliver meaningful, relevant, responsible, and ethical business and operational outcomes. The methodology accomplishes this by:

- **Thoroughly assessing a targeted business initiative** or opportunity, including objectives, desired outcomes, benefits, impediments, execution risks, potential unintended consequences, and KPIs/metrics against which to measure initiative success.
- **Identifying internal and external stakeholders** and understanding their desired outcomes, critical decisions, and KPSs/metrics against which they will measure decision and outcome effectiveness.
- Grouping stakeholder decisions and KPIs/metrics into **business or operational use cases** that deliver the desired business and operational outcomes.
- **Driving cross-stakeholder alignment and consensus** on prioritizing use cases and creating an actionable, ROI-driven use case roadmap.
- **Defining the data science requirements**, including business entities, predictive analytic scores, ML features, and a feedback loop that ensures that the AI / ML models can continuously learn and adapt.
- **Creating a culture** of empowered “Citizens of Data Science” based on a framework of AI and data Literacy.

Some of the most significant benefits to an organization adopting the updated TLADS methodology are:

[2] Throughout the book, I will use the generic term "AI" to cover the full range of advanced analytic capabilities, including machine learning, deep learning, and neural networks. One could write a collection of books explaining the nuanced differences between the growing family of advanced analytic algorithms.

- **Aligns its AI and data science projects with its strategic business goals and value drivers** by identifying and prioritizing the most relevant and impactful business initiatives and supporting use cases.
- **Engages and empowers its business stakeholders** by understanding their needs, desired outcomes, critical decisions, business and operational benefits, and potential execution impediments, providing them with prescriptive recommendations that optimize their critical decisions and deliver improved business and operational outcomes.
- **Leverages the organization's data and analytic assets more effectively and efficiently** by identifying the most relevant business entities, capturing their analytic insights or predictive propensities, and selecting and applying the most appropriate analytic algorithms.
- **Fosters a culture of continuous improvement and learning** by monitoring the effectiveness of the analytic models and feeding those learnings back into them so that they continuously learn and adapt.

Thinking Like a Data Scientist 8-step Process

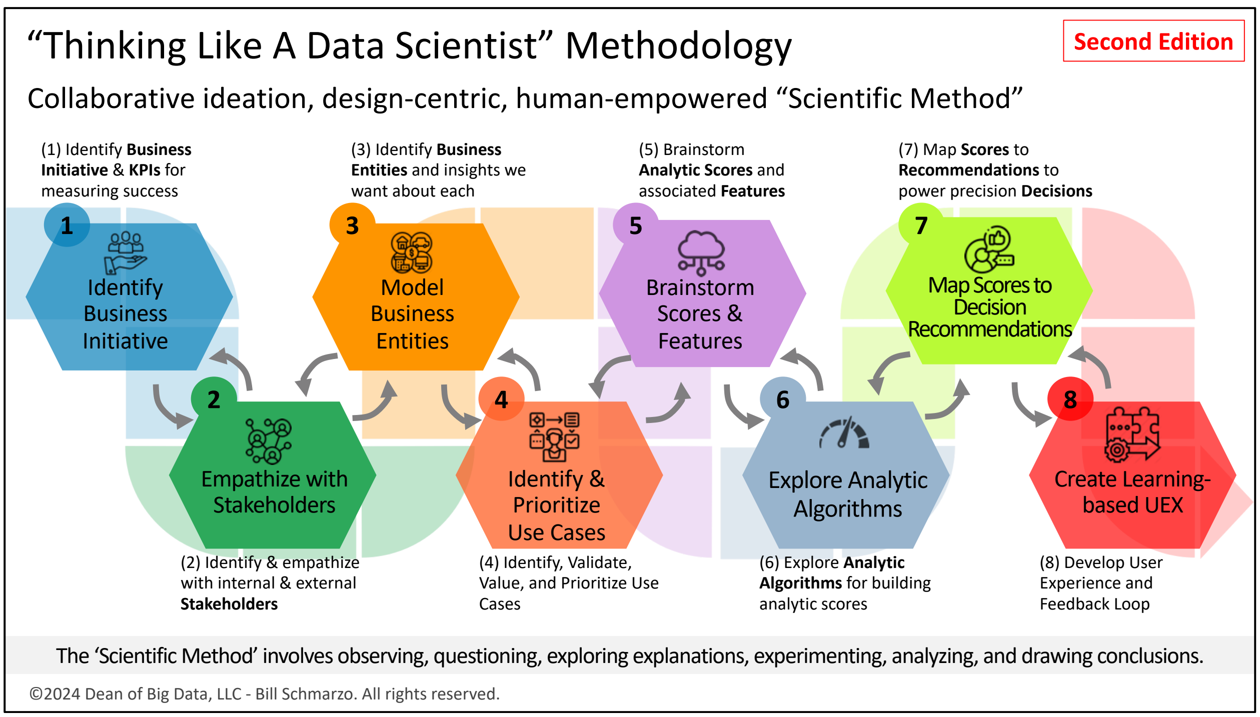


Figure 1: The Art of Thinking Like a Data Scientist **(TLADS) – Second Edition**

Figure 1 illustrates the blending of design thinking's user-centered approach with the analytical accuracy of data science, which establishes the basis of the "Thinking Like a Data Scientist" (TLADS) methodology. The objective of TLADS is to develop solutions that merge human creativity with real-world data analytics to create technically robust outputs that connect with users, leading to more efficient and relevant solutions.

Each step of the TLADS methodology comes with several design canvases to aid business and data science team collaboration. Here is an overview of the eight steps that comprise the TLADS methodology:

- **Step 1: Identify Target Business Initiative.** Step 1 involves defining the key business initiative or challenge you want to address with data and analytics, such as increasing customer retention, reducing operational costs, or improving product quality. We will specify the initiative's desired outcomes and success criteria, such as increasing customer loyalty by 10%, reducing waste by 15%, or improving defect detection by 20%.
- **Step 2: Identify Business Stakeholders.** Step 2 involves identifying and understanding the internal and external stakeholders involved in or affected by the target business initiative, such as customers, employees, managers, suppliers, or regulators. We will define the stakeholders' objectives, potential impediments, desired outcomes, key decisions, and the KPIs / metrics they might use to measure decision and outcome effectiveness.
- **Step 3: Model Business Entities.** Step 3 involves identifying the human and device entities relevant to the targeted business initiative and modeling their predictive behavioral and performance propensities. These are the human and device entities around which you will create and capture analytic insights to optimize their performance, behavior, or experience. We will define the potential predictive features and attributes of these entities, such as customer demographics, product features, machine specifications, or performance metrics.
- **Step 4: Identify & Prioritize Use Cases.** Step 4 involves grouping the stakeholders' desired outcomes, decisions, and KPIs / metrics into common business and operational use cases, such as customer segmentation, demand forecasting, predictive maintenance, or anomaly detection. We will prioritize these use cases based on their relative business value and implementation feasibility over the next 12 to 18 months.
- **Step 5: Brainstorm Scores & Features.** Step 5 involves defining the analytic scores and supporting Machine Learning (ML) features necessary to generate the prescriptive recommendations to help stakeholders optimize their prioritized use cases. Analytic scores are numerical values that measure or predict some aspect of the business entities, such as customer churn probability, demand elasticity, machine failure risk, or anomaly score. Supporting features are variables that influence or explain the analytic scores, such as customer behavior patterns, market trends, machine operating conditions, or sensor signals.
- **Step 6: Explore Analytic Algorithms.** Step 6 involves selecting and applying the appropriate analytic algorithms to build the analytic scores and supporting features based on the available data and the desired outcomes. Analytic algorithms are data science techniques that can extract insights and patterns from data, such as machine learning, deep learning, natural language processing, or computer vision. We will provide the rationale for choosing these algorithms.

- **Step 7: Map Scores to Decision Recommendations.** Step 7 involves mapping the analytic scores to the prescriptive recommendations the stakeholders will use to support their use case decisions. Prescriptive recommendations are actionable suggestions that can optimize the outcomes and value of the decisions, such as offering a personalized incentive to a customer at risk of churn, adjusting the price of a product based on demand elasticity, scheduling a maintenance service for a machine with high failure risk, or alerting an operator of an anomaly in a sensor signal.
- **Step 8: Create Learning-based UEX.** Step 8 involves creating a learning-based user experience (UEX) by defining how you will monitor and evaluate the effectiveness of the prescriptive recommendations and their impact on the outcomes and value of the decisions for the prioritized use cases. We will define a feedback loop and UEX testing methods for capturing and applying the feedback and learnings to update the analytic models that support the analytic scores.

Thinking Like a Data Scientist Organizational Benefits

There are several reasons why your organization should embrace the Thinking Like a Data Scientist methodology, including:

- **Fosters data-driven decision-making:** The methodology emphasizes using data to drive business outcomes, which can help companies make more informed decisions and reduce the risk of relying on intuition or anecdotal evidence.
- **Drives cross-organizational collaboration:** The methodology encourages collaboration between data scientists and business stakeholders, which can help to ensure that data insights are relevant and actionable for the business.
- **Focuses on delivering meaningful, relevant, responsible, and ethical outcomes:** The methodology strongly emphasizes defining and solving specific business problems, which can help companies achieve their goals and objectives.
- **Increases organizational and process efficiency:** By streamlining the data analysis process and focusing on the most crucial business problems, companies can improve efficiency and reduce the time and resources required to generate insights.
- **Improves communication:** The methodology encourages data scientists to communicate their findings clearly and actionably, which can help ensure that business stakeholders understand and act on insights.
- **Drives cultural transformation:** The methodology provides the foundation for educating everyone in the organization on their roles, responsibilities, and rights in leveraging data and analytics in defining, developing, and delivering analytics that deliver meaningful, relevant, responsible, and ethical outcomes.

The bottom line is that methodology seeks to answer one simple but powerful question:

How can we become more effective at leveraging data and analytics to power our business and operational models?

Chapter 2: Data Economics: The Power of Nanoeconomics

Before we discuss the details of the TLADS methodology, we need to review some critical data economic concepts that will be leveraged throughout the methodology. We must embrace an economic mindset to fully exploit the business, operational, and societal benefits derived from TLADS.

> 💡 ***Economics*** *is the branch of knowledge concerned with producing, consuming, and transferring wealth or value.*

The economic mindset is about understanding how economies work and making decisions that create and distribute value for customers, stakeholders, and society. However, do not confuse an economic mindset with a financial mindset.

- A **financial mindset** is a way of thinking that focuses on achieving financial goals.
- An **economic mindset** is a way of thinking that focuses on creating and distributing value for customers, stakeholders, and society.

An economic mindset empowers us to mold our future, adapt to ever-changing conditions, and respond effectively to market fluctuations and obstacles, facilitating the creation and distribution of value. The benefits of an economic mindset include:

- **Discover** unique competitive advantage through value-creation creativity and innovation.
- **Improve** organizational performance by optimizing and re-engineering the value-creation resources and processes.
- **Enhance** social well-being by balancing profits, operational excellence, environmental impact, society improvements, and ethical treatment.

> 📝 **Note:** Traditional economics is based on making decisions to allocate scarce resources optimally. However, data is not a scarce resource. Instead, data economics is propelled by the abundance of data. Every organization sits atop a treasure trove of data. The challenge isn't scarcity but harnessing this data abundance. By understanding the data we collect (awareness), making it easily available (access), and leveraging it (application) to create new sources of customer, product, service, and operational value, we unlock the unprecedented economic potential of data.

Let's jump into some critical economic and data economic concepts.

Economic Value Curve

Understanding the power of data economics starts by understanding the challenges associated with an organization's economic value curve.

> 💡 *The* ***Economic Value Curve*** *measures the relationship between a dependent and independent variable to achieve a particular outcome, such as retaining customers, increasing operational uptime, or optimizing inventory.*

The economic value curve illustrates how a change in an independent variable affects a targeted dependent variable. For instance, let's say you want to improve operational uptime, which is the dependent variable. In that case, you may have to increase maintenance expenditures, such as work hours, overtime, maintenance, engineering, parts inventory, consumables inventory, maintenance tools, and mechanic and operator training, which are all independent variables. Using the economic value curve, you can determine the optimal level of each independent variable to maximize the desired outcome of your dependent variable (Figure 2).

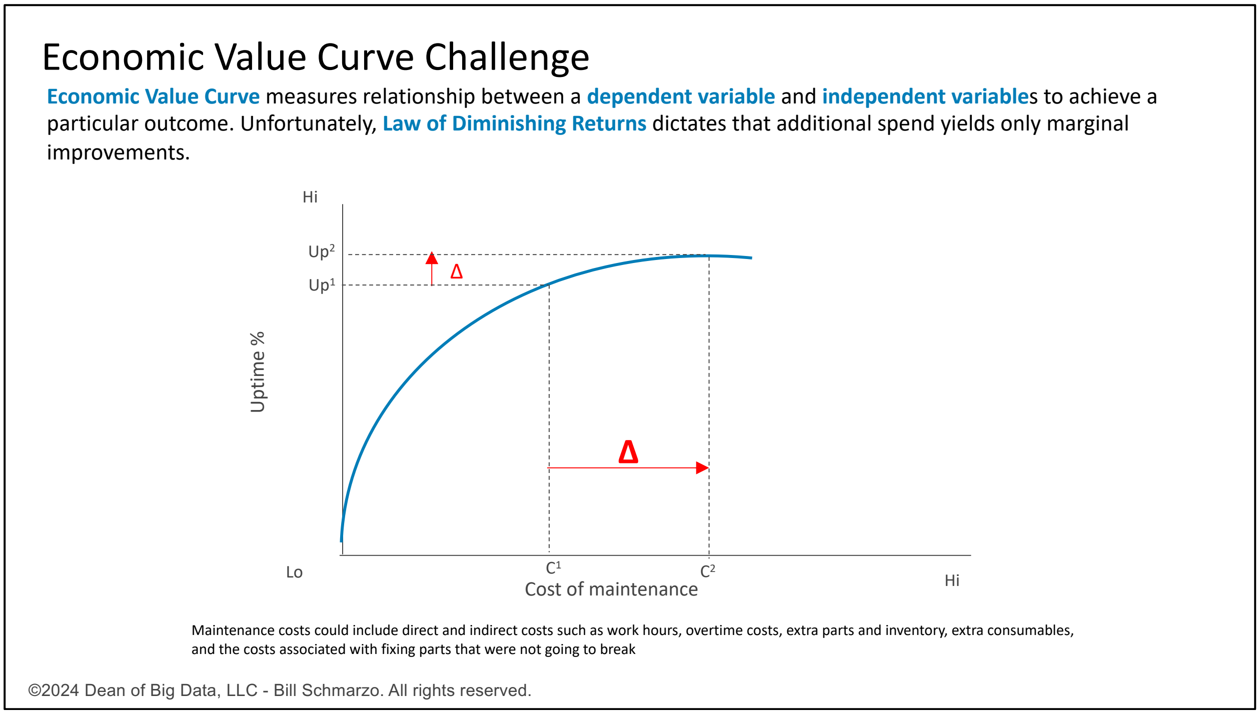


Figure 2: Economic Value Curve Challenge

Unfortunately, the economic value curve is subject to the **law of diminishing returns**, which means that beyond a certain point, increasing the investment in the independent variables will have less and less impact on the dependent variable outcomes.

> *The **Law of Diminishing Returns** describes how the increase in economic value gained from additional inputs diminishes as the quantity or quality of those inputs continues to rise (assuming other factors remain constant).*

To overcome the law of diminishing returns, we can embrace a new economic concept – **nanoeconomics** – enabled by the granular entity-level data associated with **Big Data**.

> ***Big Data** refers to extremely large data sets that can be analyzed computationally to reveal patterns, trends, and associations relating to human and device behavior and performance.*

We will leverage detailed transaction and engagement data at the individual entity level, whether that entity is a human (e.g., customer, patient, doctor, student, teacher, athlete) or a device (e.g., car, train, vacuum cleaner, compressor, air conditioner, engine) to drive the precision decisions necessary to overcome the law of diminishing returns.

Nanoeconomics

What made Big Data so valuable wasn't the volume of data; it was its granularity.

The value-generation capabilities of an organization are limited by its economic value curve and the associated challenge of the law of diminishing returns. This is where the vital data economic concept of **nanoeconomics** comes into play.

> 💡 ***Nanoeconomics** is the* economics *of individual entity's (human or machine) predicted behavioral and performance propensities.*

Nanoeconomics focuses on uncovering and codifying individual human and entity predictive behavioral and performance tendencies, also called as "**predictive propensities**." These predictive propensities are driven by inclinations, preferences, patterns, trends, and relationships at the individual level rather than at a broader group or macro level. Nanoeconomics analyzes granular or detailed data to uncover and codify individual entities' predicted propensities that can be used to deliver personalized insights, inform precise decision-making, and optimize critical organizational use cases (Figure 3).

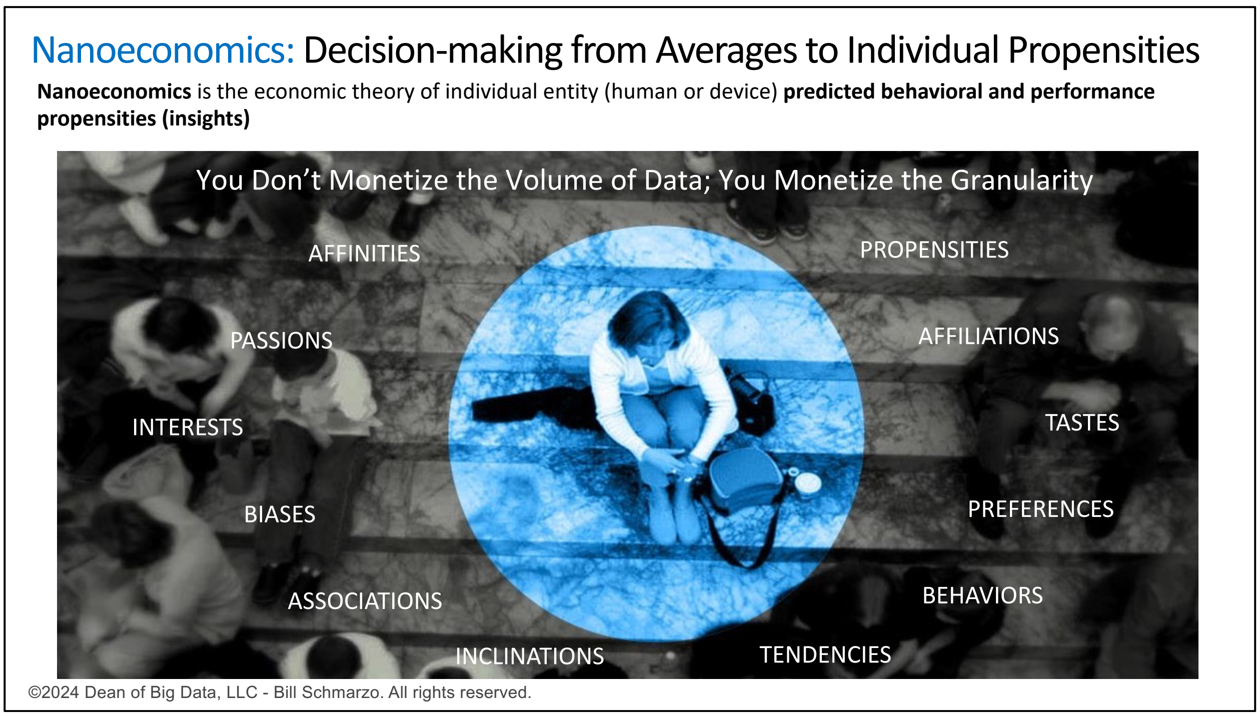


Figure 3: Nanoeconomics: Decision-making from Averages to Individual Propensities

When I was the Vice President of Advertiser Analytics at Yahoo, we leveraged nanoeconomics to analyze the browsing behaviors of hundreds of millions of web users to determine and quantify (using analytic scores) products or services in which they were most interested. We then used those scores to determine which ads to serve to them with the objective of driving clicks and conversion rates.

Similarly, we can leverage nanoeconomics to power the precision decisions that can transform our economic value curve, yielding higher returns for each unit of input (Figure 4).

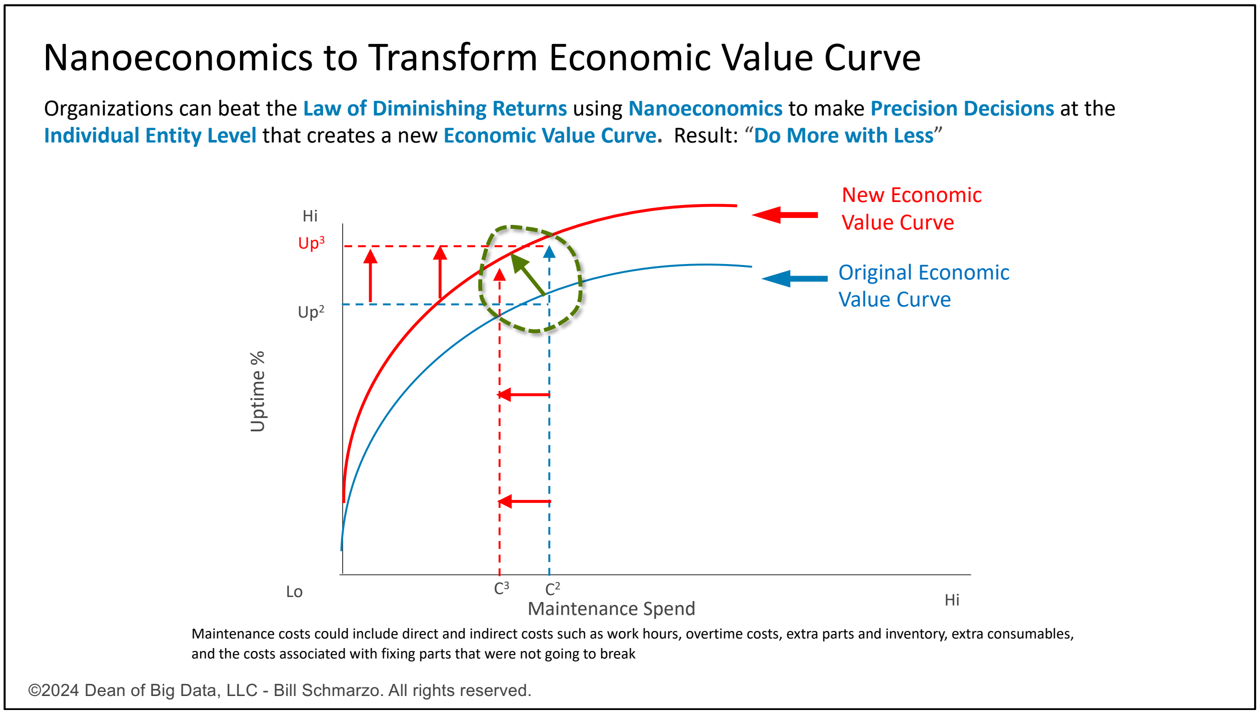


Figure 4: Nanoeconomics to Transform Economic Value Curve

Analytic Scores

Nanoeconomics is the economic theorem for deriving value from an individual entity's predictive behavioral and performance propensities derived from granular data. By focusing on individual behaviors and propensities, Nanoeconomics provides a more precise understanding of individual entities, allowing for more relevant and actionable predictive models. This approach enables organizations to make more informed decisions, improving and optimizing customer interactions and operational efficiencies. The results of this approach are captured as **analytical scores**.

> *An **Analytical Score** is a normalized, mathematically generated number that predicts the likelihood (or propensity) of a particular outcome or action for an individual human or device entity.*

Take, for example, the retail industry, where nanoeconomics is leveraged to predict individual consumer purchasing behaviors. Retailers can accurately forecast buying patterns by analyzing granular data, such as previous purchases, browsing history, and social media interactions. This enables personalized marketing strategies, optimized inventory management, and a more satisfying customer shopping experience.

In a similar vein, the medical field stands to benefit enormously from the application of nanoeconomics. Consider a healthcare provider that utilizes nanoeconomics to predict patient-specific health outcomes based on a combination of electronic health records, wearable technology data, and genetic information. This predictive capability allows personalized treatment plans that anticipate potential health issues and tailor interventions to the individual's unique physiological makeup. Such an approach could transform chronic disease management, improve patient outcomes, and significantly reduce healthcare costs by focusing on prevention and early intervention tailored to each patient's specific risk factors and needs.

These entity-level analytic scores enable our stakeholders to drive more precise decisions that can transform our economic value curve by optimizing the decisions supporting our critical business and operational use cases (Figure 5).

What are Analytic "Scores"?

***Scores** predict the likelihood that a person or event will perform in a predictable manner.*

Analytic "**Scores**" are dynamic ratings or grades normalized to aid in comparisons, performance tracking, and decision-making. Scores predict the (probabilistic prediction) likelihood of certain actions or outcomes. Scores are typically normalized on a 0 to 100 scale (0 = unlikely outcome and 100 = inevitable outcome).

Financial	**Credit Cards**	**Manufacturing**	**Casino**
Credit (FICO) Retirement Readiness Investment Risk	Attrition Risk Fraud Risk Product Preferences	Maintenance Risk Supplier Reliability Supplier Quality	Customer LTV Lifestage Influence Gaming Preferences
Education	**Healthcare**	**Utilities**	**Pro Sports**
Graduation Readiness Cohorts Influence	Wellness Condition Stress Risk	Energy Efficiency Conservation Effectiveness	Fatigue Factor Motivation Factor

Figure 5: What are Analytic Scores?

It is important to note that an analytic score needs to be presented in an easy-to-understand, easier-to-act format for our stakeholders. For example, we can define and present analytic "thresholds" that facilitate the interpretation and actionability of an analytic score (e.g., "Yes / No / Maybe" or "High / Medium /Low" or "Green / Yellow / Red").

For example, the credit score, a version provided by multiple vendors such as Fair Isaac, Experian, and Acxiom, is probably the most familiar analytic score. Your credit score predicts your *likelihood* that you will repay a loan, which can impact whether you get a loan, your interest rate, payment schedule, the loan's terms, and more. The Fair Isaac FICO® Credit Meter provides a clear visualization of a person's credit score, simplifying the determination of associated credit risk (Figure 6).

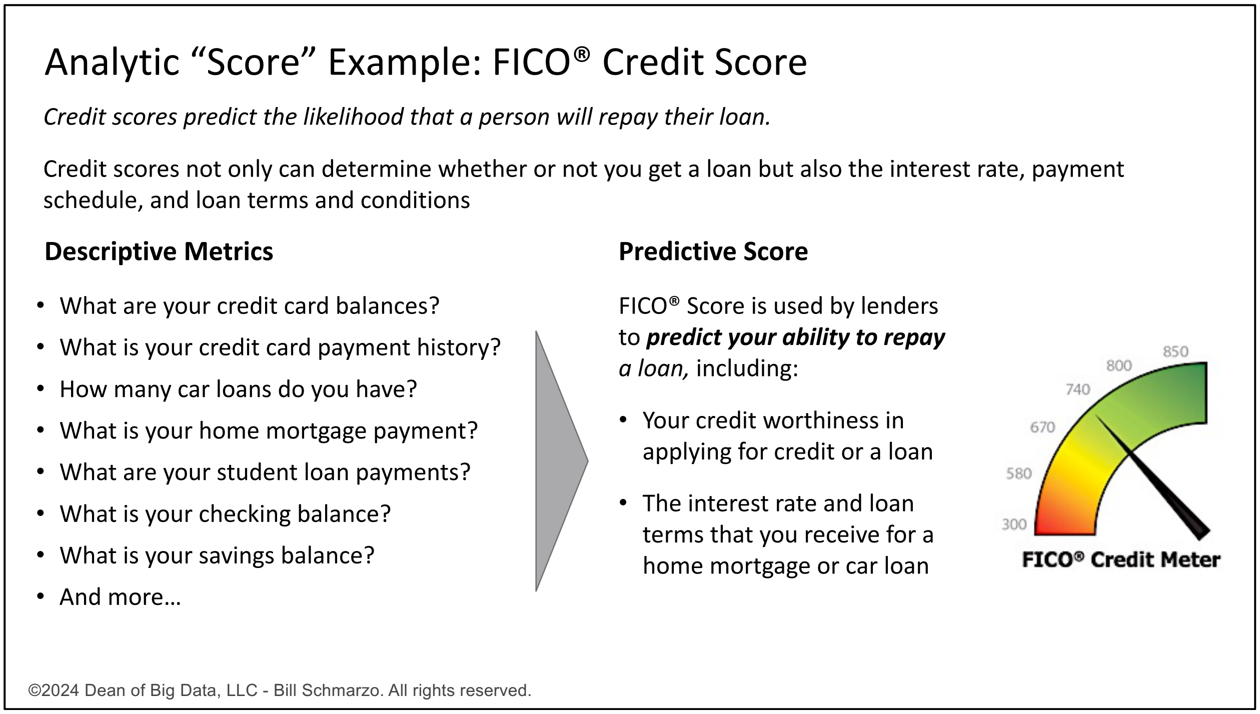


Figure 6: FICO® Credit Score Components

Analytic (Digital) Profiles

The resulting entity-level analytic scores are stored in an analytic (digital) profile, enabling the management, monitoring, and continuous refinement of the individual entity analytic scores.

> 💡 *An **Analytic (Digital) Profile** is a comprehensive framework that encapsulates the key characteristics, behaviors, and propensities of entities such as customers, products, or machines derived from data analysis to inform strategic decisions and actions.*

An Analytic Profile is a framework that encapsulates the key predictive characteristics, behaviors, and patterns of individuals, processes, devices, or any entity of interest within a structured data model. It is a foundational element in data science and analytics, enabling the systematic analysis and understanding of the individual entity it represents (Figure 7).

Analytic Profiles: Codifying and Sharing Asset Predicted Propensities

Analytic Profiles codify, share, re-use and continuously-refine the predicted behavioral and performance propensities (**Analytic Scores**) for the organization's key human and device assets

Patient Care Data

- Demographic
- Behavioral Demographics
- Psychographics
- Patient care / treatment history
- Patient vital stats history
- Physician / Nurse care notes
- Patient comments
- Pharmacy/Prescriptions
- Others…

Schmarzo Patient Healthcare Profile	NCE Score	Variance	Trend
Health Score	92	1.89	↑
Wellness Score	92	1.85	↔
Diet Score	67	3.25	↔
Exercise Score	82	2.25	↑
Stress Score	65	1.90	↓
COVID19 At-Risk Score	22	2.35	↓
Cancer At-Risk Score	14	1.74	↑
Pulmonary At-Risk Score	02	1.15	↔
Oncology At-Risk Score	08	1.20	↓
Heart Attack At-Risk Score	09	1.25	↔
Stroke At-Risk Score	06	1.10	↔
….			

External Patient Data

- Diet History (DietPlanner, MyFitnessPal)
- Physical Exercise History (Apple Watch, FitBit)
- Mental Acuity History (Lumosity, CogniFit)
- Stress History (Stress Doctor, Happify)
- Emotional History (Text, Social)
- Vices History
- Vacation / Relaxation History
- Others…

Figure 7: Analytic (Digital) Profiles

An Analytic Profile is a multidimensional representation of an entity. It enables data scientists to systematically uncover and codify the intricate patterns and behaviors that define the entity, thereby uncovering actionable insights and opportunities for value creation. Analytic Profiles are dynamic and designed to evolve as new data becomes available, ensuring their insights and predictions remain accurate and reflect the entity's current state or behavior. This dynamic nature allows for more effective and targeted decision-making, strategic planning, and operational improvements.

📝 **Note:** Color-coded analytic scores and trend arrows are used in Figure 7 to simplify the interpretation and application of individual entities' analytic profiles. This color-coding and trend representation facilitates at-a-glance assessment, allowing healthcare professionals to identify areas of concern or improvement quickly. It's a dynamic tool for monitoring and managing patient health, emphasizing proactive and preventive care through continuous refinement of predictive analytics.

Applying Analytic Scores to Optimize Use Cases

We then apply these entity-level analytic scores to optimize business and operational use cases, delivering more meaningful, relevant, responsible, and ethical business and operational outcomes (Figure 8).

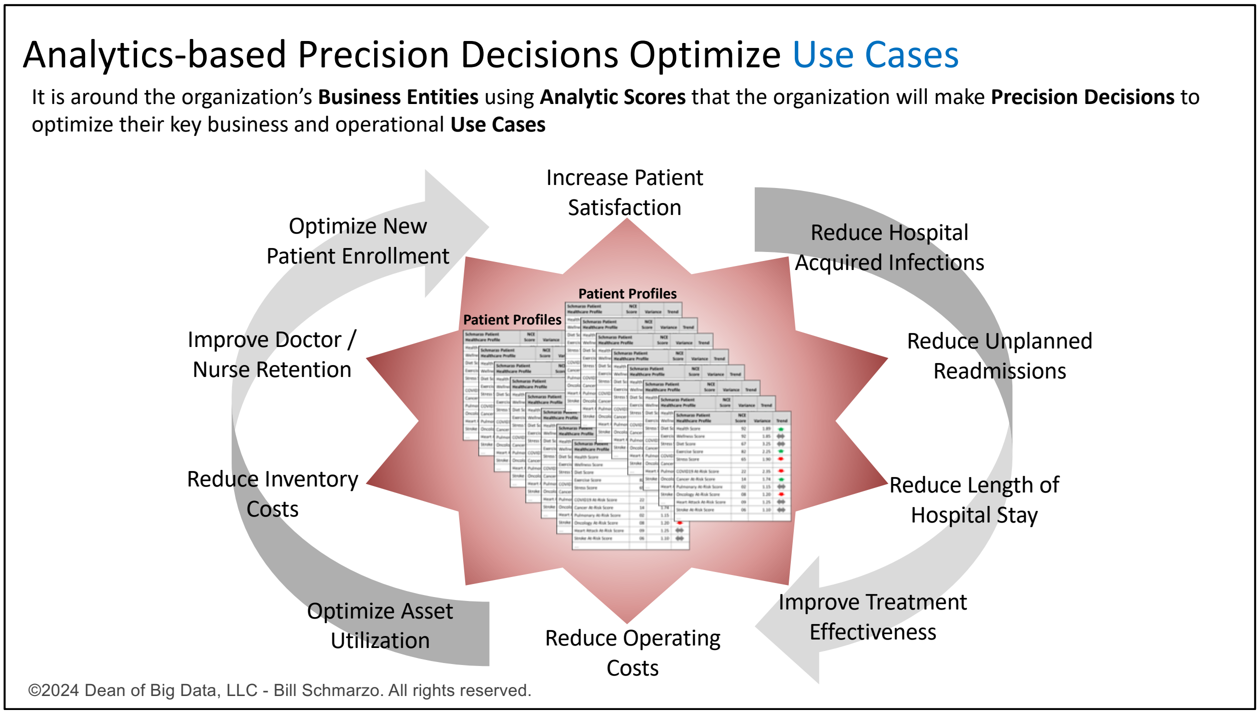


Figure 8: Analytics-based Precision Decisions Optimize Use Cases

Using entity-level analytic scores, we can improve and optimize business and operational outcomes by ensuring our actions and decisions are more relevant, meaningful, responsible, and ethical.

Data Economics Value Chain

It's not technology advancements that are the game-changers. The game-changer is how those technological advancements are leveraged to economically transform industries and society.

The Data Economics value concept combines nanoeconomics, analytics scores, analytic profiles, and use cases to deliver game-changing levels of performance and innovation that can transform the organization's economic value curve (Figure 9).

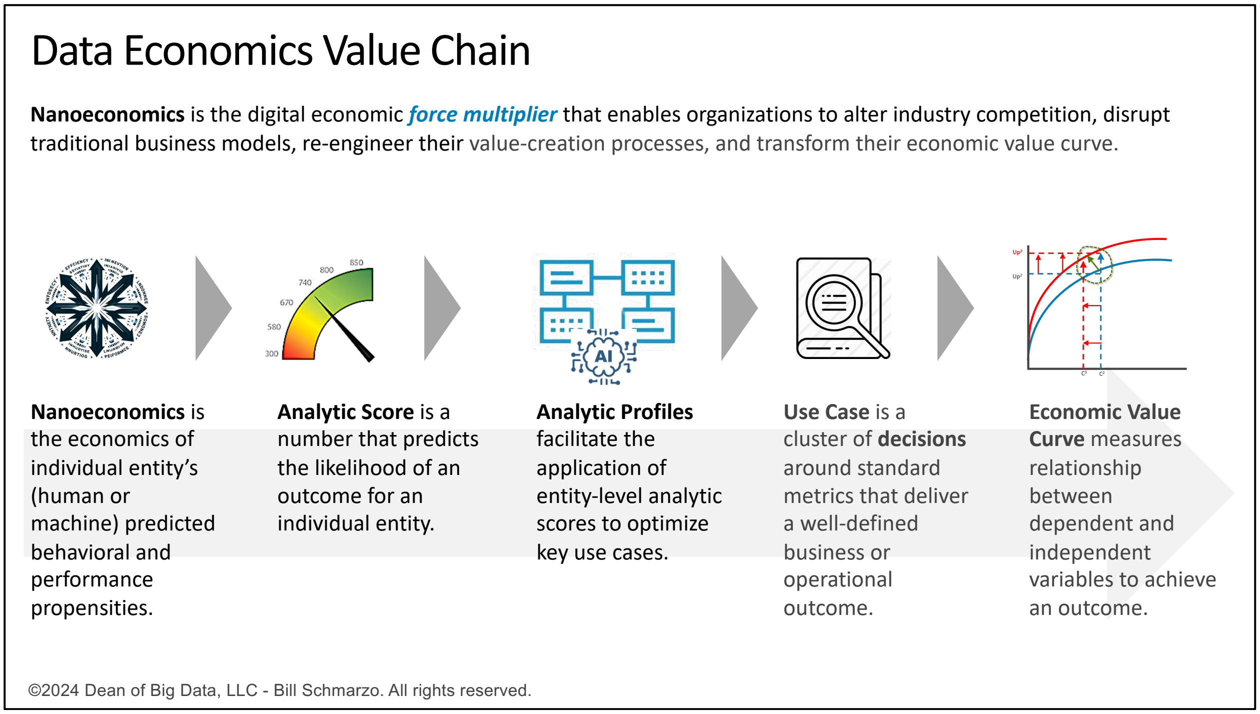


Figure 9: Data Economics Value Chain

Nanoeconomics is the Data Economics Value Chain *force multiplier* enabling organizations to alter industry competition, disrupt traditional business models, re-engineer their value-creation processes, and transform their economic value curve.

> 💡 ***Force Multiplier,*** *from an economics perspective, refers to a factor or element that dramatically increases (multiplies) the effectiveness of an effort or investment, leading to significantly greater output or results in an economic system.*

Summary: The Power of Nanoeconomics

Nanoeconomics leverages entity-level analytics to enhance an organization's economic value-creation capabilities. Companies can change how they compete in their industry by implementing nanoeconomics, disrupting traditional business models, improving their value-creation processes, and transforming their economic value curve. Applying nanoeconomics can significantly benefit various domains and industries, generating differentiated and tangible customer, product, service, and operational value (see Table 1).

Industry	Business / Operational Outcomes	Nanoeconomics Drivers
Healthcare	Personalized treatment plans increasing patient satisfaction and health outcomes.	Detailed patient data analytics driving personalized care, focusing on genetics, lifestyle, and patient preferences. Integration of historical health data and predictive analytics to customize treatment and therapy plans.

Education	Customized learning paths enhancing student engagement and achievement.	Use of detailed insights on student abilities and interests to tailor educational content and methods to individual learning styles, tutoring, and goals.
Manufacturing	Optimized production processes and maintenance schedules improving efficiency and reducing costs.	Real-time performance and condition monitoring of machines to optimize each machine's production process and maintenance schedule based on its performance, condition, usage, and environment.
Retail	Personalized customer recommendations and optimized pricing, promotion, and product placement leading to increased sales and customer loyalty.	Analysis of customer detailed shopping history, preferences, and budget to deliver highly-personalized recommendations and shopping. Nanoeconomics can also enable dynamic pricing and product placement strategies driven by market demand, supply, and competition analysis.
Transportation	Improved efficiency, safety, and sustainability of mobility systems through optimized traffic management and customized travel options.	Analysis of granular transportation data to improve the efficiency, safety, and sustainability to optimize the traffic flow, routing, congestion, and emissions of vehicles. Nanoeconomics can also create customized travel solutions based on traveler needs and behaviors, with incentives for sustainable choices.
Energy	Balanced electricity supply and demand, optimized renewable energy sources, and reduced carbon footprint.	Real-time, granular smart meters, sensors, and devices data to manage electricity supply and demand of electricity and optimize energy consumption, generation, and storage. Nanoeconomics can also help optimize renewable energy sources and reduce the energy sector's carbon footprint

Table 1: Nanoeconomic: Industry Transformation Examples

The business and operational benefits of nanoeconomics include:

- **Improve** the efficiency, effectiveness, and quality of operational and policy decisions by reducing the costs and increasing the benefits of each action.
- **Enhance** customer satisfaction, loyalty, and retention by providing tailored products, services, and experiences that meet their specific needs and expectations.
- **Foster a culture** of continuous exploration, learning, and adaptation by using data and analytics as a source of innovation and differentiation.

Nanoeconomics is the foundation for driving value creation and industry transformation. The TLADS methodology embraces nanoeconomics and the economic mindset as foundational principles for creating new sources of customer, product, service, and operational value.

Chapter 3: Design Science

***Data science** is about identifying those variables and metrics that __might__ be better predictors of performance.*

Data science is the art and science of extracting meaningful and actionable customer, product, service, and operational insights or predictive propensities buried in the data. It is the practical combination of art (human creativity) and science (mathematical algorithms).

- Data science uses advanced algorithms to analyze vast amounts of granular data, uncover patterns and trends, and make informed decisions to solve complex problems.
- Data science enables organizations to adopt a forward-thinking approach that uses predictive analytics to simulate future events based on different scenarios. It predicts likely outcomes and recommends actions based on those predictions.

> 💡 ***Data science** is the systematic study of data to extract knowledge and insight with the intent to discover better indicators and predictors of performance.*

Data science enables informed and precise decision-making by blending technical skills (software development, statistics, advanced mathematics) with professional experience (domain knowledge) and creativity (design thinking, user experience design). See Figure 10.

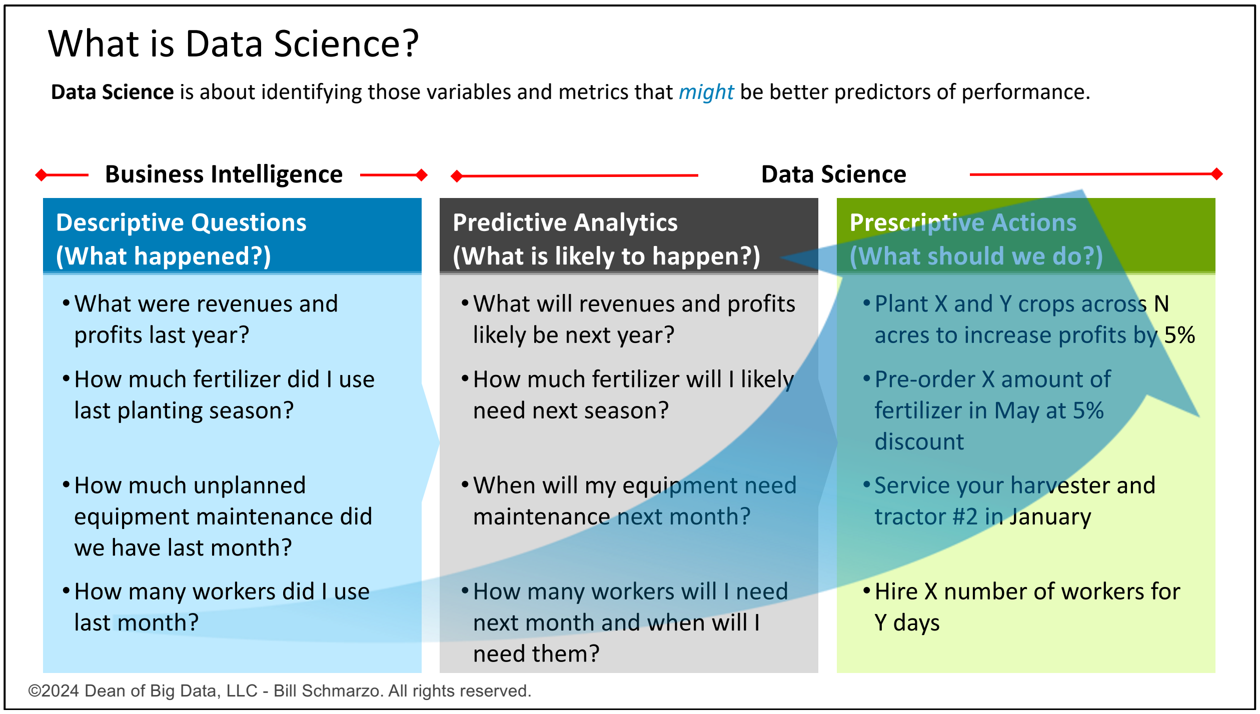


Figure 10: What is Data Science?

However, data science is more than the technical abilities required to analyze massive amounts of granular data. At its core, data science is about the economic value it generates. By transforming raw data into actionable insights, data science can help businesses improve their strategies, enhance customer experiences, drive innovation, and uncover new sources of customer, product, service, and operational value.

The Data Science Process

Software development defines the criteria for success; data science discovers them.

The data science development process demands curiosity, creativity, and willingness to explore a wide variety of analytic algorithms, feature exploration, data enrichment transformations, and synthesized composite metrics[3] to discover actionable customer, product, service, and operational insights (predictive propensities) buried in the data. The data science team will use various techniques and algorithms to identify and codify entity and event-based patterns, trends, associations, and relationships buried in the data that can be used to optimize key business and operational use cases (Figure 11).

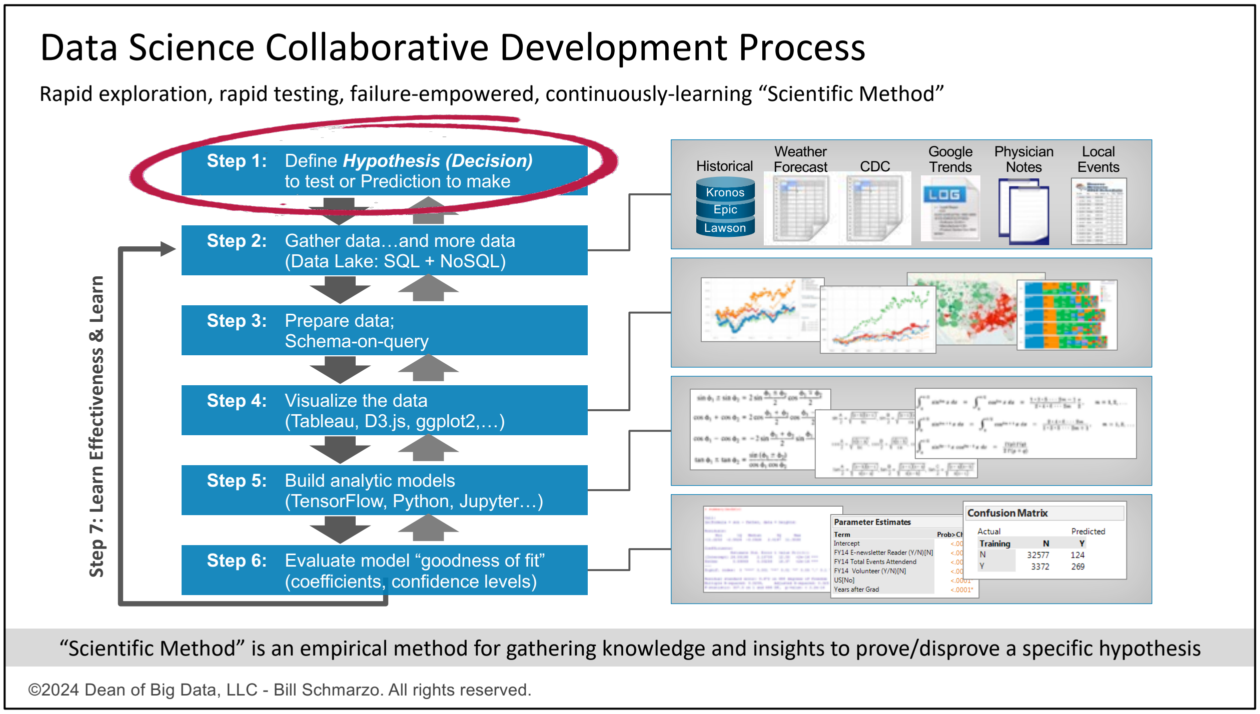


Figure 11: Iterative, Collaborative Data Science Development Process

The highly iterative, exploration-centric data science process involves the following steps:

- **Step 1: Define Decision.** The data science development process starts by identifying and thoroughly documenting the hypothesis (or decision) to be optimized. The hypothesis should be specific, measurable, achievable, relevant, and time-bound. The data science development process requires close collaboration with the business stakeholders to understand the key sources of business differentiation (e.g., how the organization delivers value) and then brainstorm the variables and metrics (ML features) that might yield better predictors of performance. The Hypothesis Development Canvas, discussed in Chapter 14, helps frame the problem and guide the data science discovery effort.

[3] A composite metric is an aggregated indicator synthesizing multiple data points into a singular value to provide a multi-dimensional assessment of a subject or performance.

- **Step 2: Gather Data.** The data scientists (in collaboration with data engineers, business stakeholders, and Subject Matter Experts) identify and gather relevant data from multiple internal and external sources. This will include structured transactional data (e.g., Point of Sales, Electronic Health Records, Enterprise Resource Planning, Customer Relationship Management), semi-structured (logs, RSS feeds, JSON, XML), and unstructured (PDFs, websites, social media posts, images, audio, videos). The data science team needs a unified (virtualized) data platform to access and integrate data from multiple internal and external sources.
- **Step 3: Prepare Data.** Given the targeted hypothesis, the data science team collaborates with business Subject Matter Experts to assess the data's level of readiness across multiple dimensions, such as reliability, completeness, coverage, consistency, timeliness, uniqueness, and validity. This step involves several data engineering tasks to transform the data into a usable format, including removing or filling in missing values, handling outliers and anomalies, resolving inconsistencies and errors, encoding categorical variables, scaling numerical variables, and applying feature engineering techniques to create new variables and reduce dimensionality. This step can be time-consuming and monotonous, but it is critical to ensure the quality and usability of the data and the accuracy of subsequent analytics. It's important to remember the principle of GIGO: Garbage In → Garbage Out.
- **Step 4: Explore Data.** The data scientist uses a variety of data visualization and descriptive statistics techniques to explore the data sources in search of patterns, trends, correlations, and outliers of interest buried in the data. This step helps the data scientist gain insights into the data, understand its distribution characteristics, identify potential relationships and associations, and generate hypotheses for further testing. This step helps the data scientist check the assumptions and limitations of the data (e.g., granularity, latency, completeness, metadata, quality) and the analytic methods.
- **Step 5: Build and Refine Analytic Models.** This is where the real data science work begins. The data scientist uses various tools and analytic algorithm libraries to codify cause and effect. They will explore combinations of analytic techniques and algorithms to create a predictive model that balances the predictive accuracy versus the cost to generate and manage the analytic models. The choice of analytic techniques and algorithms depends on the problem's type and complexity, the data's nature and availability, performance metrics, and the desired outcome. The data scientist will use cross-validation, regularization, and hyperparameter tuning to optimize the models to prevent overfitting or underfitting.
- **Step 6: Evaluate Model Goodness of Fit.** The data scientist must assess the performance and reliability of the analytical model through a comprehensive evaluation of its "goodness of fit." This involves conducting a series of statistical tests and leveraging various metrics to understand how well the model captures the underlying patterns of the dataset upon which it was trained and how accurately it can predict outcomes on new, unseen data. Key performance indicators such as accuracy, precision, recall, the F1-score, ROC curves, AUC (Area Under the Curve), and confusion matrices

provide a holistic view of the model's effectiveness.

During the “goodness of fit” evaluation, one of the nuanced challenges is determining when “good enough” is actually “good enough." This decision is not solely technical but also financial, requiring a careful consideration of the cost-benefit trade-off associated with further model refinement. For example, if you can get 80% accuracy using basic statistical techniques for $10, that saves $5,000; there is no point in improving accuracy to 82% if it costs $2,000 but only saves a net $500

Finally, the “goodness of fit” evaluation process is further complicated by the varying impact of false positives and false negatives, which must be weighed against each other in light of the specific business goals, constraints, and desired outcomes.

- **Step 7: Measure Effectiveness and Learn.** Step 7 of the data science process involves measuring the effectiveness of the analytic model. This is done by learning from the results and feeding those learnings back into the model to continuously learn and adapt. Once the model is deployed in a production environment, its performance must be monitored to ensure it delivers the desired outcomes. The data scientist will use systematic monitoring and end-user feedback to evaluate the model's impact on the business outcomes. This feedback and the results obtained will be used to refine the model, update the data, and improve the data science process.

Data Science Saddle Back Challenge

It's important to note that the Data Science Development process is often a series of steps that involve progress and setbacks. This is because new insights may be discovered along the way, or requirements may be clarified. For example, if a data quality issue is discovered during Step 3 of the Data Science Development process, it may be necessary to return to Step 2 to gather higher-quality data from stakeholders.

Each phase aims to bolster the model's predictive accuracy and overall efficacy in the iterative data science journey. Yet, a pivotal moment invariably arises—a "saddle point" in the development process.

> 💡 ***Saddle Point** refers to a phase in analytic model development where further iterations fail to significantly improve the model's performance, signaling a plateau in predictive accuracy and effectiveness.*

As a data scientist, it's crucial to understand a saddle point, a stage in the analytics development process where performance and accuracy improvements start to level off. This could indicate that you need to change your data science approach. You may need to pivot your development approach by scrutinizing the fundamental assumptions and methodologies that support your analytic models. This involves exploring various data engineering techniques, new data transformations, and a different orchestration of analytic algorithms (Figure 12).

Data Science “Saddle Point" Challenge

Optimization paths (A and B) may top out sooner with lower “Saddle Points” than optimal path (C). Data scientist must be prepared to jettison their current path in hopes of finding more predictive path.

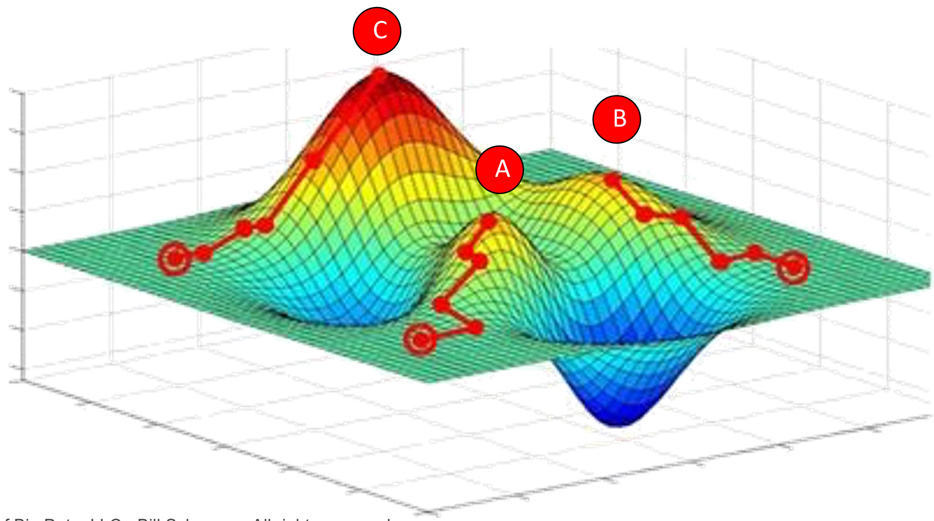


Figure 12: Data Science “Saddle Point” Challenge

This stage of reflection and exploration is crucial since it evaluates the worth of further improvements. To make this assessment, it is essential to consider the resources required, the law of diminishing returns, and the potential risk of overfitting. Before proceeding with model optimization, balancing the pursuit of minor improvements with the practical goal of developing a robust and comprehensive model is vital. Essentially, navigating through the saddle point requires a combination of analytical rigor, creative problem-solving, and strategic foresight - all of which are fundamental to pushing the boundaries of data science to new heights.

Analytic Algorithms Maturity Model

The data science team can choose from various categories of analytic algorithms. The team will likely need to assemble or orchestrate multiple algorithms to create an analytic model that provides the best results while considering the costs associated with false positives and false negatives (Figure 13).

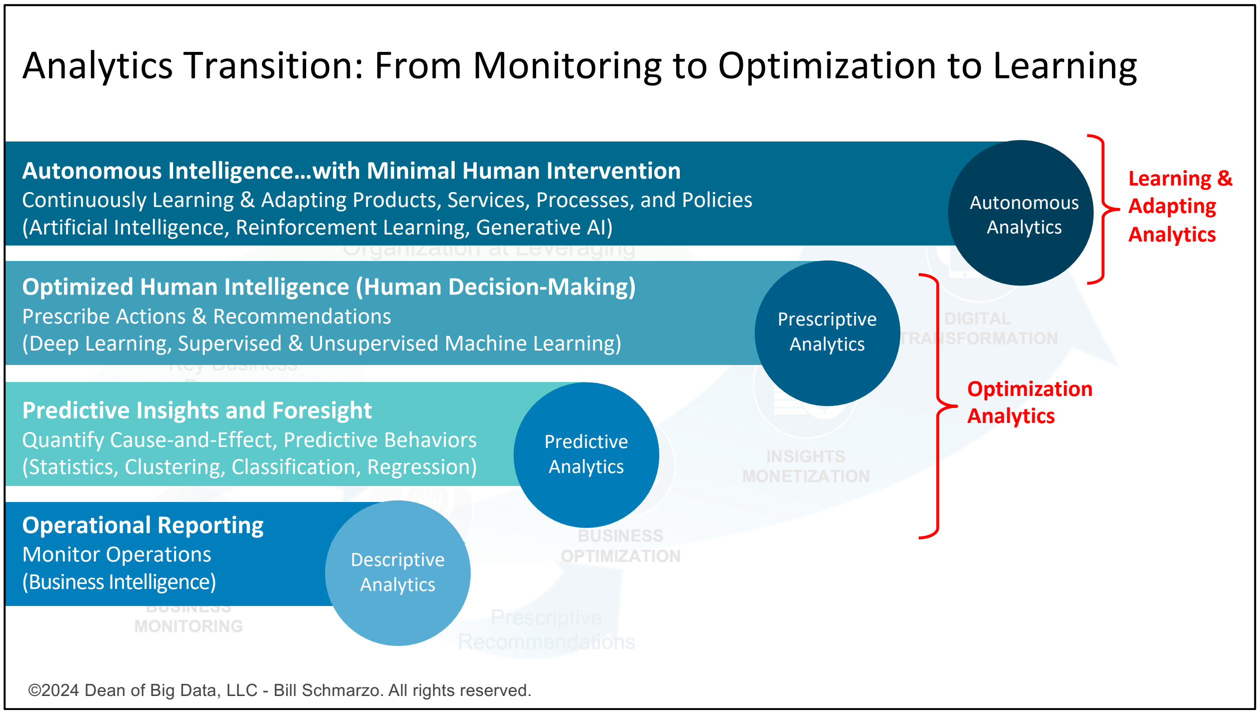


Figure 13: Analytic Algorithms Maturity Levels

The different levels of analytics maturity in Figure 13 can guide data science teams to develop their analytic models and ensure that they deliver optimal outcomes for their organizations. Each level indicates a distinct degree of sophistication and value in using data and analytics to solve problems and make informed decisions.

- **Level 1: Operational Reporting.** Level 1 analytics is all about operational reporting, which keeps track of the business's state. It uses historical data to create management reports and operational dashboards. Level 1 analytics identifies anomalies and generates management and operational alerts by comparing current and past performance. It's like having a report card for the business's operations.
- **Level 2: Predictive Insights and Foresight**. Level 2 analytics aims to identify and decipher data trends, patterns, and relationships. This level of analytics involves statistical analysis, data mining, and exploratory analytics techniques such as clustering, classification, and regression. These techniques are used to uncover the relationships between data elements, assess the cause and effect, determine the confidence levels, and measure the Goodness of Fit to determine the strength of these trends, patterns, and relationships.

- **Level 3: Optimized Human Intelligence.** Level 3 of advanced analytics aims to enhance human decision-making effectiveness by predicting future events and recommending the next best actions. This level builds on the trends, patterns, and relationships identified in Level 2 to generate prescriptive recommendations and actions. The decision-maker combines the insights from analytics with their domain expertise, critical thinking, qualitative research, and observations. Categories of analytics that fall under Level 3 include neural networks (or deep learning), regression analysis (supervised machine learning), recommendation engines, and federated learning.
- **Level 4: Autonomous Analytics.** Level 4 analytics aims to develop autonomous or semi-autonomous analytics-driven components such as products, devices, policies, procedures, and processes that learn and adapt continuously as they interact with their surroundings. Level 4 analytics differ significantly from the previous three levels since they are focused on learning and adapting with minimal human intervention. This is a matter of concern for regulators and politicians due to the capability of level 4 analytics to learn and adapt more rapidly than laws and regulations can be created to govern their ethical and responsible use.

The use of artificial intelligence (AI) is growing rapidly, but there are concerns about the need for proper regulations. To address this issue, regulators worldwide are working on developing regulatory frameworks. China has taken a fragmented approach by publishing numerous documents, while the European Union (EU) has taken a comprehensive approach with the EU Artificial Intelligence Act. Data scientists must consider legal and political factors when implementing AI.

By understanding and applying the different levels of analytics maturity, a data science team can create analytic models that deliver more relevant, meaningful, responsible, and ethical outcomes.

Chapter 4: Design Thinking Humanizes Data Science

Design thinking is the key to unlocking and leveraging the organization's institutional knowledge and insights, typically at the frontlines of customer engagement and operational execution. It is a powerful approach that can help create a culture where everyone is empowered to identify where and how data science can leverage its data with AI to generate new sources of customer, product, service, and operational value.

> 💡 ***Design thinking** is a human-centered approach to problem-solving that seeks to solve problems by thoroughly understanding, validating, and prioritizing consumers' needs and desired outcomes.*

Design thinking can facilitate collaboration between business and data science teams, enabling them to create analytic-driven products and services that provide value to their customers and constituents. This approach can help reveal hidden assumptions, biases, and predictive features that can accelerate the data science development process. When design thinking is combined with data science, it can result in more innovative and impactful outcomes for customers, products, services, and operations.

Design thinking accomplishes this by:

- **Empathizing with customers** to understand their tendencies, propensities, inclinations, and behaviors.
- **Diverging ideation before converging** in brainstorming new ideas (where all ideas are worthy of consideration) before converging on a best few.
- **Rapid experimentation** by building prototypes (minimum viable products) to validate learnings and expand idea exploration (quick and dirty works well here).
- **Collaborative learning** by sharing what you've learned with the people for whom you are designing (frequent feedback is the best feedback).
- **Continuous learning and adaption** from launching, monitoring, and learning from customer experiences and success in using the product and service in real-world use cases.

Stanford's Design School (d.school) is famous for creating and integrating design thinking tools and mindsets across Stanford's diverse schools (e.g., business, computer science, electrical engineering, mechanical engineering, health care). Many folks, including me, believe that design thinking is the secret sauce to Stanford's success in cultivating the entrepreneurial spirit of its students and faculty (Figure 14).

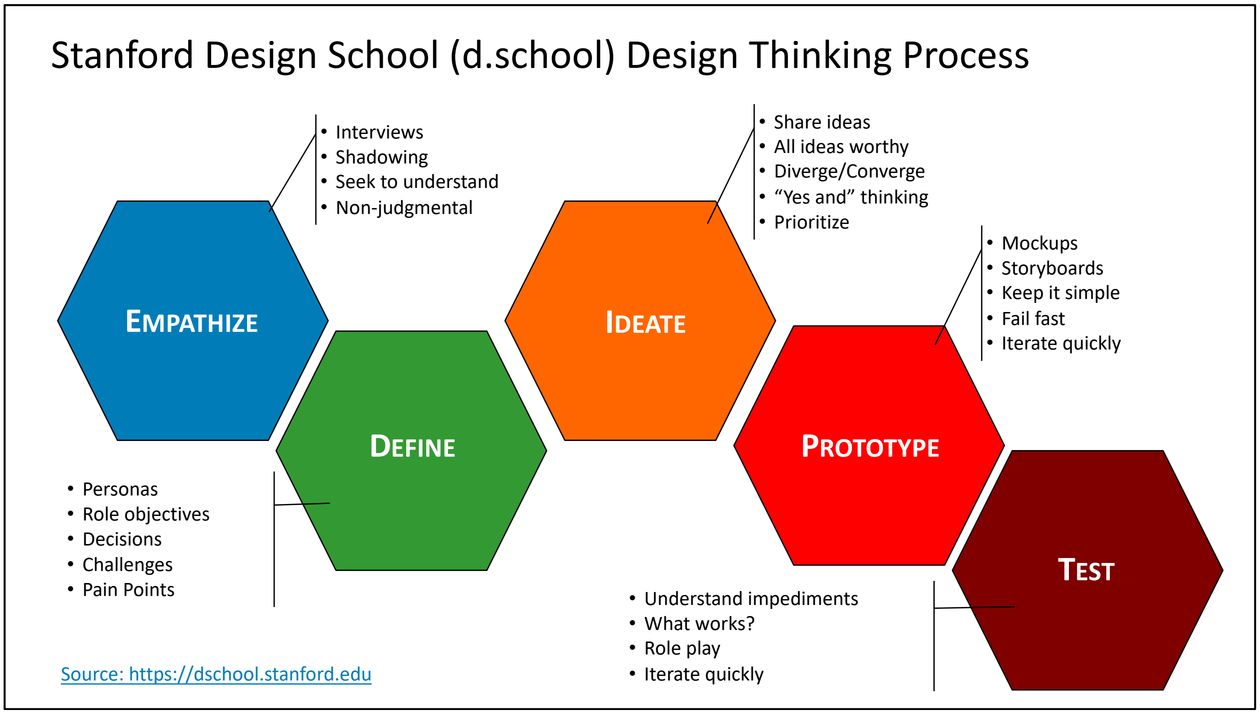


Figure 14: Stanford University d.school Design Thinking Process

The five stages of the design thinking process are:

- **Empathize**: This stage involves understanding the people's needs, empathizing with their experience, and reframing the challenge to generate out-of-the-box solutions. It relies heavily on qualitative research methods such as interviews, observations, and immersive experiences to gather deep insights into user behaviors and emotions. For example, a team designing a new healthcare app might spend time observing and interviewing patients with chronic illnesses to understand their daily challenges and emotional experiences.
- **Define**: This stage involves synthesizing the insights from the empathy stage and defining a clear and actionable problem statement. It focuses on identifying the core issues that need addressing and ensures that all team members have a unified understanding of these issues to target their efforts effectively. In our example, after gathering patient insights, the team might define the problem as "Patients with chronic conditions need a way to track symptoms and medications effectively to improve their daily management of the illness."
- **Ideate**: This stage involves generating many possible solutions using brainstorming, mind mapping, sketching, etc. It encourages free thinking and the exploration of a wide array of ideas, fostering creativity and innovation without the constraints of feasibility or practicality initially. For our example, the team might brainstorm various features for the app, such as medication reminders, symptom trackers, or communication tools for consulting with caregivers.
- **Prototype**: This stage involves creating low-fidelity versions of the solutions, such as mock-ups, storyboards, models, etc., and getting feedback from the users. This iterative

process helps translate abstract ideas into tangible expressions, allowing designers to visualize and refine the concepts' practicality. For our example, they might then create a simple clickable prototype of the app, incorporating the most promising features identified during the ideation phase to see how users interact with the design and navigate its functionalities.

- **Test**: This stage involves refining the prototypes based on the feedback and testing them with real users to evaluate their effectiveness and desirability. This phase is crucial for understanding how well the solution meets the user's needs in a real-world context and identifying any areas for further improvement. For our example, this prototype could be tested with a small group of users, collecting feedback on its usability and the relevance of its features, which would then be used to refine the app further before a wider release.

Design Thinking versus UEX/UI Design

Design thinking is distinct from User Experience (UX) Design and User Interface (UI) Design, although all three frameworks focus on understanding and empathizing with the user. They share an iterative process that involves prototyping and testing, and both seek collaboration and feedback to refine and improve outcomes.

Although design thinking and UEX/UI Design may seem similar, they differ in terms of their approach to creating more appropriate and valuable offerings for users.

- **Design thinking** is a broad problem-solving process that can be applied to various product, service, and organizational challenges. It fuels human curiosity, exploration, imagination, and innovation.
- **UEX/UI design**, on the other hand, is more tactical and focused on the specifics of product interaction and presentation. It combines usability and aesthetics principles to enhance the user's interaction with the product.

Table 2 highlights the key differences and similarities between design thinking and UEX/UI design:

Aspect	**Design Thinking**	**UEX/UI Design**
Focus	Solving complex problems through a human-centered approach.	Enhancing user satisfaction through usability, accessibility, and the visual appeal of the product.
Process	An iterative process involving understanding the user, challenging assumptions, redefining problems, and creating innovative solutions to prototype and test.	Combines research, prototyping, visual design, usability testing, and iteration based on user feedback to create seamless and aesthetically pleasing products.

Key Principles	Empathy, Ideation, Prototyping, Testing, Iteration.	Usability, Accessibility, Efficiency, Aesthetics, Visual Design Principles, Brand Consistency.
Outcomes	Innovative solutions that deeply consider user needs and behaviors.	A product that is not only easy to use and meets users' needs but is also visually appealing and provides a seamless journey from start to finish.
Tools & Methods	Workshops, Stakeholder Interviews, Brainstorming, Prototyping.	User Research, Personas, Journey Mapping, Wireframing, Prototyping, Usability Testing, Sketch, Adobe XD, Figma, Color Theory, Typography.
Goal	To explore and identify innovative solutions to problems.	To ensure the product is both aesthetically pleasing and provides an optimal experience for the user.

Table 2: Comparing Design Thinking and UEX/UI Design

We will spend the rest of this chapter exploring design thinking to enable the organization to unleash its tribal knowledge, resulting in more relevant, meaningful, responsible, and ethical AI and data-driven business and operational outcomes.

Primary Design Thinking Tools

Design thinking provides effective tools to assist you in understanding your stakeholders' perspective, gaining a thorough understanding of what they value and the obstacles that impede achieving that value, and identifying areas where analytics can enhance positive experiences and eliminate negative ones. Below, I have shared a few of my favorite design thinking tools.

The Persona Map

Personas are an effective tool for comprehending and recording the daily activities of our primary stakeholders. This includes their desired outcomes, important decision-making requirements, queries to assist those decisions, areas of difficulty, and the metrics they use to gauge the effectiveness of their decisions and outcomes (Figure 15).

[Store Manager] Persona Map

PERSONA: Store Manager
BUSINESS INITIATIVE: Increase Same Store Sales

- Description: Typically, 2 to 4 years of Chipotle experience; promoted through the ranks to store manager
- Quote: "I love Chipotle, so working here only seemed natural"
- Desired Outcomes: 1) Achieve monthly revenue and profit numbers to get bonuses, 2) create an enjoyable, friendly work environment, 3) get promoted to regional manager
- Outcomes Impediments: 1) lack visibility into local events and how to leverage them, 2) employee attrition, 3) customer satisfaction

Key Decisions	Supporting Business Questions	Decision Pain Points	KPIs	Predictive Variables
Staffing	•What's the likely store traffic throughout the day? •When and how long are local events?	Lacks real-time model that allows him to update changing demand variables	% employee absence, % late employees, employee satisfaction, employee overtime	Employee hours worked, Employee absence, Employee tenure, Employee unused vacation, Employee complaints
Inventory	•Given store traffic estimates, how much additional inventory? •What local events that could impact inventory?	No insights into neighboring stores inventory to fill temporary inventory gaps	% out of stock, % product waste, current inventory levels, % supplier on-time deliveries, % supplier product quality	Daily inventory levels, Daily traffic forecasts, Product supplier reliability, Product supplier quality
Customer Satisfaction	•What's the best comp to give unsatisfied customer? •How important is customer to me?	Doesn't know "how valuable" that particular customer to the Chipotle	Customer LTV, Customer Visit Frequency, Return Customers, Customer Satisfaction, NPS, LTR	Customer visits, Frequency of customer visits, customer spend, customer social advocacy, size of party
Production	•Shift staff to produce more materials given store traffic? •Which products are most popular for local events?	Doesn't have insights into local events that impact production needs	Average product throughput, average prep time, % employee absence,	Staff tenure, staff certification, store traffic forecasts, inventory levels, local events forecast, supplier reliability & quality

Figure 15: Store Manager "Persona Map" Canvas

The "Persona Map" provides a deeper understanding of the success criteria for our key stakeholders, including:

- **Key Decisions**: These are the key decisions the stakeholders need to make to support the targeted business initiative. For example, the Store Manager must make staffing, inventory, customer satisfaction, and production management decisions.
- **Supporting Business Questions**: These are the queries that stakeholders might ask to inform their decision-making. For instance, a store manager might ask questions about staffing decisions, such as the expected foot traffic during the day or the timings and duration of local events.
- **Decision Pain Points**: These are the challenges or pain points the stakeholder currently faces in making their key decisions. For example, a Store Manager's pain point might be the lack of real-time insights that allow them to optimize staffing decisions.
- **Key Performance Indicators (KPIs):** These are the KPIs and metrics that the stakeholder might use to measure the effectiveness of their decision-making. For the Store Manager, that could include employee absences, percentage of time employees are late to work, employee satisfaction, and employee overtime hours.
- **Predictive Variables**: These variables and metrics (features) might influence the stakeholder's key decision. The predictive variables affecting the Store Manager's staffing decisions might include employee hours worked, absence, tenure, unused vacation, and complaints.

Stakeholder Journey Maps

Stakeholder Journey Maps are a way of capturing the decisions, critical measures of decision effectiveness, the gains, and the pains experienced by a stakeholder who is trying to achieve a specific outcome, such as buying a house, fixing a wind turbine, going on vacation, or increasing same store sales. The stakeholder journey is divided into five stages, as shown in Figure 16, which tracks the customer's progress through each stage:

- **Awareness** identifies the actions and metrics that make the persona aware that they have a need, problem, or opportunity.
- **Consideration** considers the options or alternatives the persona could take or make to address the problem or opportunity.
- **Execution** manages the execution or acquisition of the selected solution, including resolving conflicts that arise during the implementation.
- **Review** assesses the solution's success or effectiveness, resolves any outstanding or unresolved issues or problems, and plans the next steps in the solution's execution.
- **Sustain / Retirement** determines ongoing support requirements, measures ongoing success, and decides when to retire the solution when appropriate.

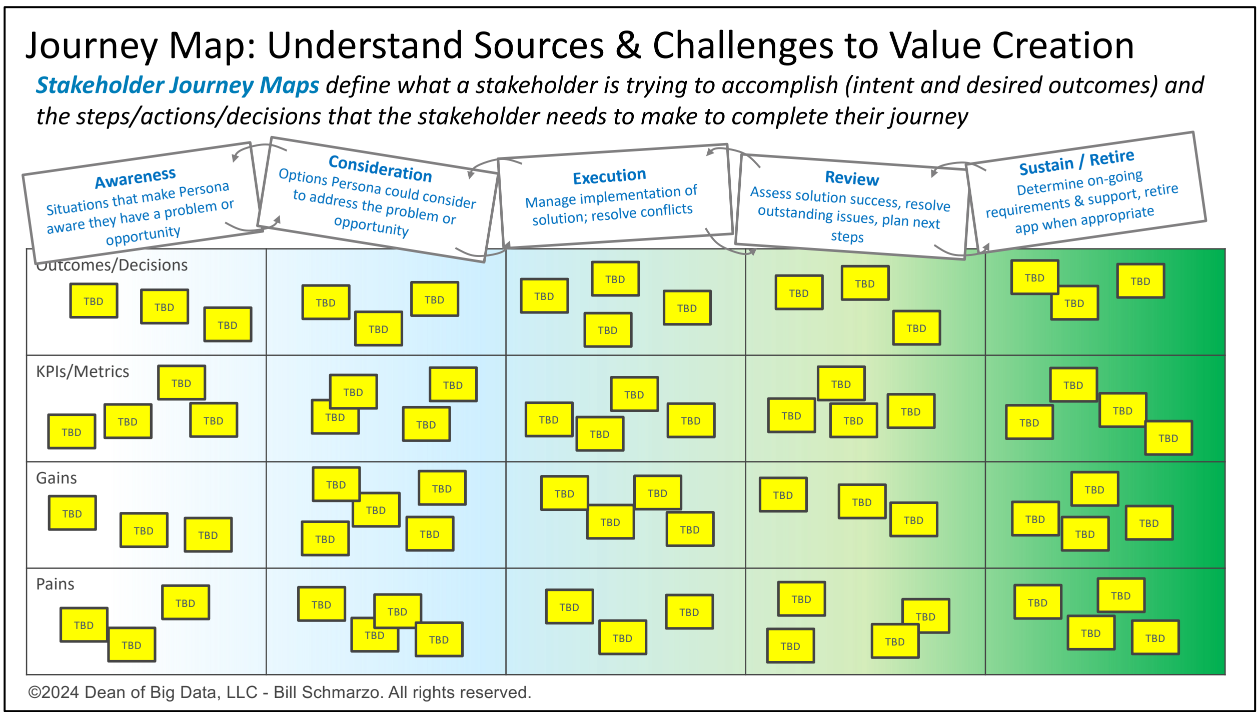


Figure 16: Stakeholder Journey Map

Figure 17 below is an example of the Store Manager "Journey Map" canvas corresponding to the Store Manager "Persona Map" created in Figure 15.

[Store Manager] Journey Map

PERSONA: Store Manager

BUSINESS INITIATIVE: Increase Same Store Sales

- Desired Outcomes: 1) Achieve monthly revenue and profit numbers to get bonuses, 2) create an enjoyable, friendly work environment, 3) get promoted to regional manager
- Outcomes Impediments: 1) lack of visibility into local events and how to leverage them, 2) employee attrition, 3) customer satisfaction

	Awareness	Consideration	Execution	Review/Evaluate	Sustain/Retire
Decisions	Identify key value propositions that differentiate Chipotle. Decide on the channels (social media, online ads, in-store promotions) to promote awareness	Decide on strategies to showcase Chipotle's unique offerings, such as quality of ingredients or sustainability practices.	Implement strategies to streamline the ordering process, enhance customer service, and ensure food quality.	Collect and analyze customer feedback on their dining experience. Decide on metrics for evaluating customer satisfaction.	Assess the effectiveness of loyalty programs and decide on whether to continue, enhance, or retire certain marketing strategies.
KPIs	Social media engagement rates, website traffic, ad impressions.	Conversion rates from marketing campaigns, time spent on website/menu pages.	Average order value, order accuracy rate, service speed.	Customer satisfaction scores, repeat customer rate, online reviews rating.	Loyalty program enrollment and redemption rates, cost-benefit analysis of promotional strategies.
Gains	Increased brand recognition and customer interest.	Customers appreciate the brand's value and uniqueness.	Improved customer satisfaction and loyalty.	Insight into customer preferences and areas for improvement.	Long-term customer retention and brand loyalty.
Pains	Cost of advertising and marketing campaigns.	Resource allocation for content and engagement strategies.	Training staff and implementing new technologies.	Managing negative feedback and operational adjustments.	Costs associated with loyalty programs and promotions.

Figure 17: Store Manager "Journey Map" Canvas

Additional Design Thinking Tools

Below is a list of additional design thinking tools and concepts that can be used to understand where and how AI and data can drive new sources of stakeholder value.

- **Empathy Maps** help designers empathize with users by capturing what they say, think, do, and feel. For example, a team developing a fitness app uses empathy maps to uncover that their users feel motivated yet intimidated by advanced workouts.
- **Prototyping** is creating a preliminary model or sample of a product, allowing designers to explore ideas and test functionalities before final production. For example, smartwatch designers can create a foam model to explore the ergonomics and wearability of their design before moving to more sophisticated physical prototypes.
- **Storyboarding** is a sequential visual representation of how a user interacts with a product, illustrating the user's experience and emotions. For example, a mobile banking storyboard might explore how a user receives a payment notification and splits the bill with friends, emphasizing ease of use and social connectivity.
- **Usability Testing** is an observation technique that evaluates a product or service by testing it with representative users to improve its usability. For example, a team testing a new e-commerce website might observe users struggling to find the checkout button, prompting a redesign for better visibility and accessibility.
- **Service Blueprints** are detailed diagrams that visually map the service process and its touchpoints, showing the relationship between different service components—people, props (physical or digital evidence), and processes. For example, a restaurant used a

service blueprint to identify necessary backend processes for online ordering, like order management and delivery logistics.

- **Affinity Diagrams** are used to organize data and ideas into groups based on their natural relationships for review and analysis. For example, public transit app developers could use an affinity diagram to group user comments and feedback into categories like navigation, payment, and accessibility.
- **Feedback Loops** are systems used to capture users' feedback about their experience with a product or service, allowing for continuous improvement. For example, an educational tech company might create feedback loops by regularly surveying students using their software, leading to continuous updates based on user suggestions and needs.

These tools and concepts are central to the design thinking process, fostering a user-centric approach to problem-solving and encouraging innovative solutions.

Design Thinking Enhances Data Science

The key difference between design thinking and data science is that design thinking focuses on human learning, while data science focuses on algorithmic learning.

Design thinking enables data science teams to generate more relevant and meaningful outcomes. As previously mentioned, incorporating a human-centered approach in the data science development process can lead to more precise and accurate results. Successful design thinking and data science cultures share similar mindsets and characteristics, such as:

- An open environment for sharing and building upon the work of others.
- An inclusive ethos where “All ideas are worthy of consideration.”
- A learning philosophy nurtured through experimentation (and learning through failure).
- A willingness to unlearn old methods that one has held as the gospel truth.

One of the most intriguing cultural similarities between design thinking and data science is the mindset that if you don't have enough "might" moments - ideas that could potentially have a significant impact - you'll never have any "breakthrough" moments. Figure 18 illustrates the complementary aspects of design thinking and data science.

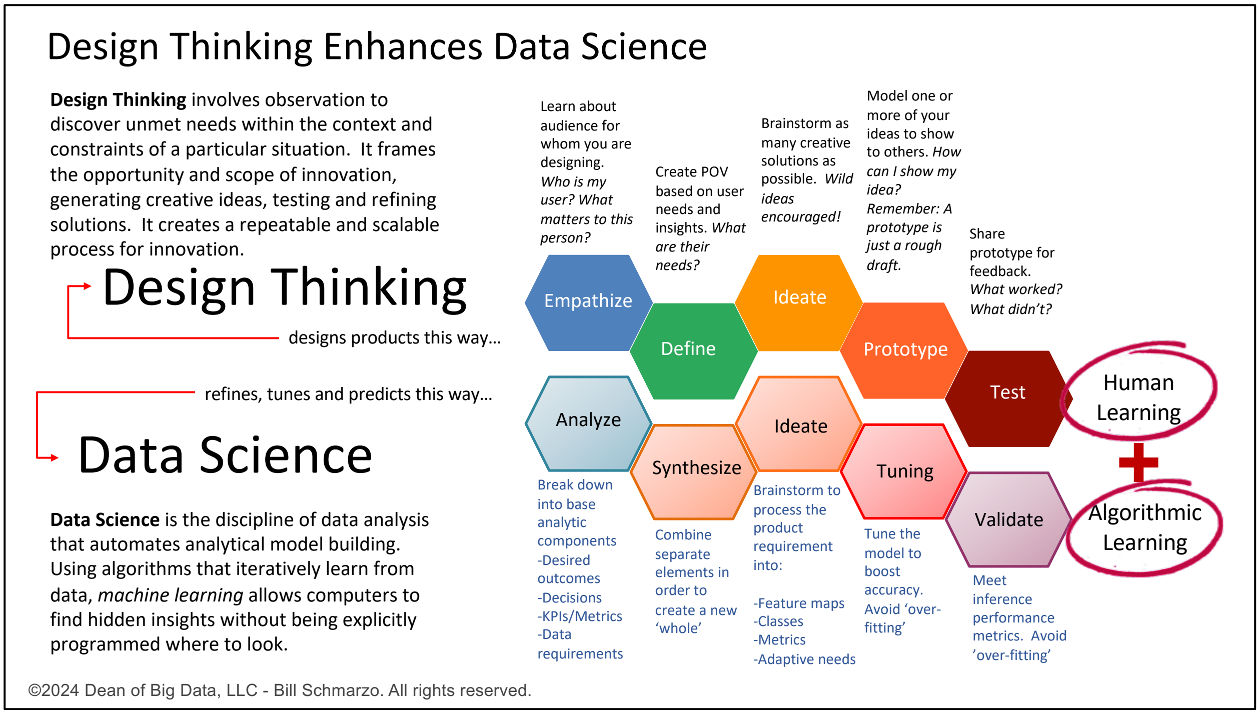


Figure 18: Design Thinking Enhances Data Science

The Figure 18 diagram illustrates the cyclical and iterative process inherent to design thinking and data science. It emphasizes the non-linear progression of tasks, where insights obtained at any stage may prompt revisiting and refining previous steps, ensuring that the solution continually evolves. This dynamic feedback loop between stages leads to more profound insights and more effective solutions over time as the process iteratively converges on the most suitable outcomes.

The only significant difference between design thinking and data science is that design thinking seeks to unleash human learning, while data science seeks to unleash algorithmic learning.

Let's drill into the similarities of the five stages of the blended design thinking and data science processes:

- **Step 1: Empathize and Analyze**. The "Emphasize and Analyze" step seeks to understand your users and build empathy for the challenges and constraints that get in the way of their desired outcomes: Who is my user? What matters to this person? What are they trying to accomplish? What are their impediments to success? What frustrates them today?
- **Step 2: Define and Synthesize.** The "Define and Synthesize" step seeks to understand the user's Point of View (POV) regarding their needs: What are their needs? What capabilities will the user need? In what type of environment will the user be working? What is likely to impede the execution of their job? What is likely to hinder adoption?
- **Step 3: Ideate and, well, ideate some more**. The "Ideate and Ideate" step focuses on brainstorming creative solutions to address the users' needs and obstacles. Consider questions such as: What are the most important decisions? What operational requirements need to be met? Does everyone understand the challenges and opportunities? Has

everyone had a chance to contribute? Encourage participants to share even the wildest ideas because exploring ideas beyond the current status quo can lead to breakthrough solutions.

- **Step 4: Prototype and Tune.** We develop the initial prototype incorporating analytics in the "Prototype and Tune" stage. This is the phase where we transform ideas into product or service prototypes, such as user interface mockups, storyboards, and example outputs. This helps us to validate and establish usage patterns and identify the metrics and KPIs that would be used to measure operational effectiveness. We need to answer questions such as: What metrics must be captured? What's our instrumentation and data strategy? And how can we present the analytic insights in an actionable way for the users?
- **Step 5: Test and Validate.** The “Test and Validate” step involves testing and validating the designed solution and analytics. The objective is to make the design and analytics operational. This step is crucial as it initiates the continuous improvement from the user experience and analytic model fine-tuning perspectives. It becomes necessary to instrument or tag the product or solution to continuously monitor its usage. By doing so, we can gather insights such as which features are used most, what paths are the most common, and whether any usage patterns indicate confusion among users. Additionally, we can identify usage paths from which users eject and never return, which can help us improve the product or solution.

Blending Design Thinking and Data Science

Combining design thinking and data science is a powerful approach that helps us create technically sound, relevant, useful, and engaging AI models and analytic scores for the users. By leveraging the data science process, we can test and validate our hypotheses and use design thinking to empathize with the users and iterate on the user experience. We can incorporate design thinking concepts to generate new and creative ideas to improve the relevance and effectiveness of our AI models. By blending design thinking and data science, we can create data products that are both intelligent and user-friendly.

Table 3 highlights the complementary nature of design thinking and data science.

	Design Thinking	**Data Science**
Step 1: Empathize and Analyze	Understand and capture the user’s intent, desired outcomes, operational requirements, and impediments to success; learn as much as possible about the users for whom you are designing.	Capture and prioritize the user’s key decisions; ideate the variables and metrics that might be better predictors of those decisions.
Step 2: Define and Synthesize	Define, document, and validate your understanding of the user’s intent, desired outcomes, operational requirements, and	Synthesize your understanding of the decisions (e.g., decision latency, granularity, frequency, governance, sequencing) to flesh out the potential

	potential impediments. Don't be afraid of being wrong.	variables and metrics (ML features) and assess potential analytic algorithms and approaches.
Step 3: Ideate and uh… Ideate	Brainstorm as many potential solutions as possible. Diverge in your brainstorming ("all ideas are worthy of consideration") before you converge (priority those best ideas based upon potential business and customer value and implementation feasibility).	Pilot potential analytic models and algorithms with small sample data sets to see what types of insights and relationships are buried in the data. Capture and refine the hypotheses as you test and learn.
Step 4: Prototype and Tune	Create one or more interactive mockups with which your key constituents can "play". Study users' interactions with the mockups to see what works and where they struggle. Identify what additional design guides and/or analytics insights could be provided to improve the user experience.	Identify where analytic insights or recommendations are needed – and what additional data can be captured – as the users' "play" with the mockups. Explore opportunities to delivery real-time actionable insights to help "guide" the user experience. Fail fast but learn faster! Embrace the "Art of Failure."
Step 5: Test and Validate	Monitor usage and navigational metrics to determine the effectiveness of the product or solution. Create a continuous improvement environment where usage and performance feedback can be acted upon quickly to continuously improve the product's design.	Exploit the role of "Recommendations" to improve or guide the user experience. Capture, validate, and prioritize user usage feedback and success. Leverage the "wisdom of crowds" to continuously fine-tune and re-tune the supporting analytic models predictive and prescriptive effectiveness.

Table 3: Design Thinking and Data Science Common Similarities

Creating and Nurturing a Culture of Empowerment

Design thinking is a powerful discipline that can help organizations create a culture of empowerment by involving all stakeholders in the ideation and problem-solving process. This approach is beneficial as it fosters inclusivity, innovation, and adaptability – something that I refer to as "Organizational Improvisation."

> 💡 ***Organizational Improvisation*** *(Improv) refers to an organization's ability to quickly innovate and adapt to customer and market changes by leveraging human creativity and AI-based learning models.*

Organizational Improvisation is facilitated by creating "organizational swirls" that promote collaboration and innovation within and across an organization. These swirls function like a well-coordinated jazz quartet or a soccer team, allowing for dynamic interactions and collaborations across different parts of the organization. This fosters the free flow of ideas and insights, catalyzing new initiatives and improvements. The main objective is to break down silos and encourage a culture where data, analytic insights, and customer, product, service, and market opportunities can intersect in creative ways. By doing so, this approach leads to innovative solutions and enhances organizational agility by actively involving diverse teams in the problem-solving process, allowing the organization to quickly adapt to new challenges and opportunities.

Some of the key tenets of design thinking that can help create a culture of empowerment include:

- **Empathy and inclusion** are crucial to design thinking. This approach emphasizes understanding and considering different perspectives, which creates an environment where everyone's voice is heard, not just those in leadership or specific departments. Encouraging employees from various levels and departments to share their insights and experiences enriches the understanding of the challenges and opportunities facing the organization.
- Design thinking is a process that enables all members of an organization to participate in **ideation and innovation**. It fosters an open and collaborative environment where brainstorming and creative thinking are highly valued. Workshops and ideation sessions are typical activities where employees can contribute their ideas without fearing criticism. This open invitation for participation enhances team morale and nurtures a sense of ownership and engagement among team members.
- One crucial aspect of design thinking is accepting **failure** as a necessary step toward learning and success. This approach fosters a culture where experimentation and risk-taking are vital for discovering and innovating. By reframing failures as learning opportunities, organizations can reduce the negative perception of making mistakes.
- In a culture that embodies design thinking, all ideas are **worthy of consideration**, regardless of their source. This approach eliminates hierarchical obstacles to innovation and encourages participation from individuals at all levels. When everyone believes their ideas are appreciated, it can foster a wider range of diverse perspectives, essential for developing groundbreaking innovations and tackling intricate challenges.
- "**Diverge Before You Converge**" is a fundamental concept that promotes creativity and facilitates an inclusive approach to problem-solving. The first step in this process is divergent thinking, where all team members are encouraged to generate a broad range of ideas and solutions without any restrictions or criticism. This approach allows organizations to explore a wider range of solutions and perspectives, discovering novel opportunities and insights that may have gone unnoticed otherwise (Figure 19).

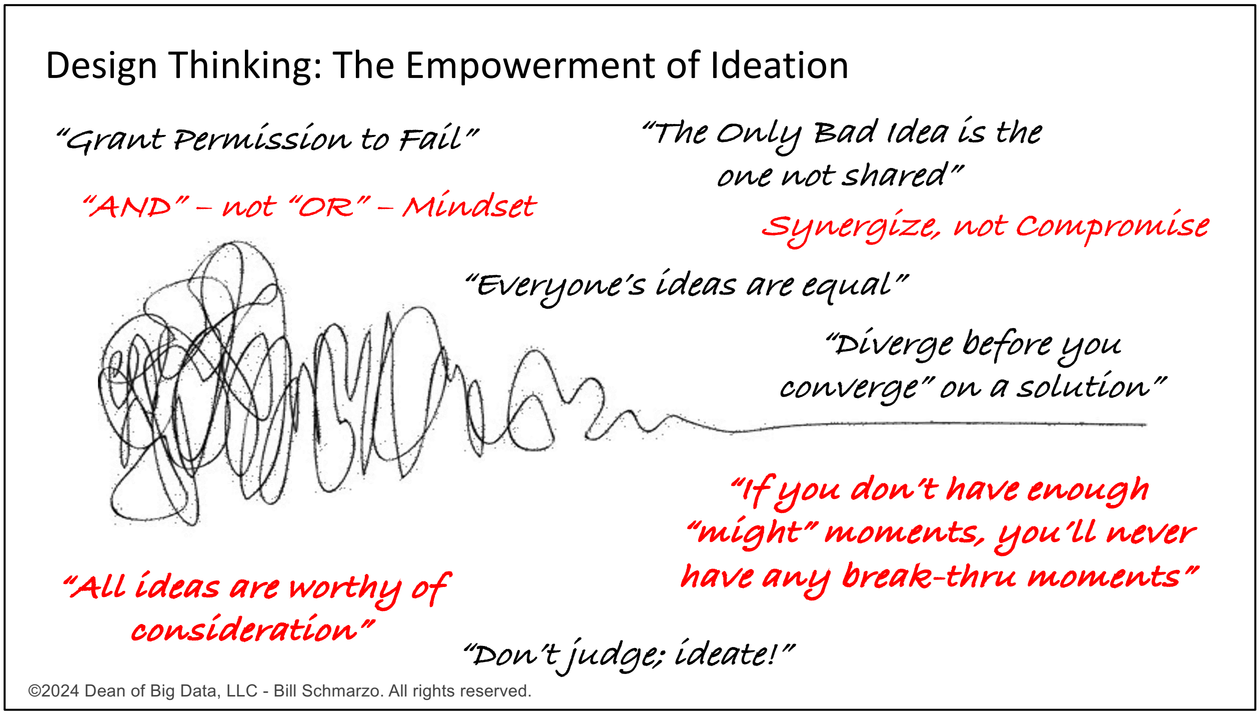
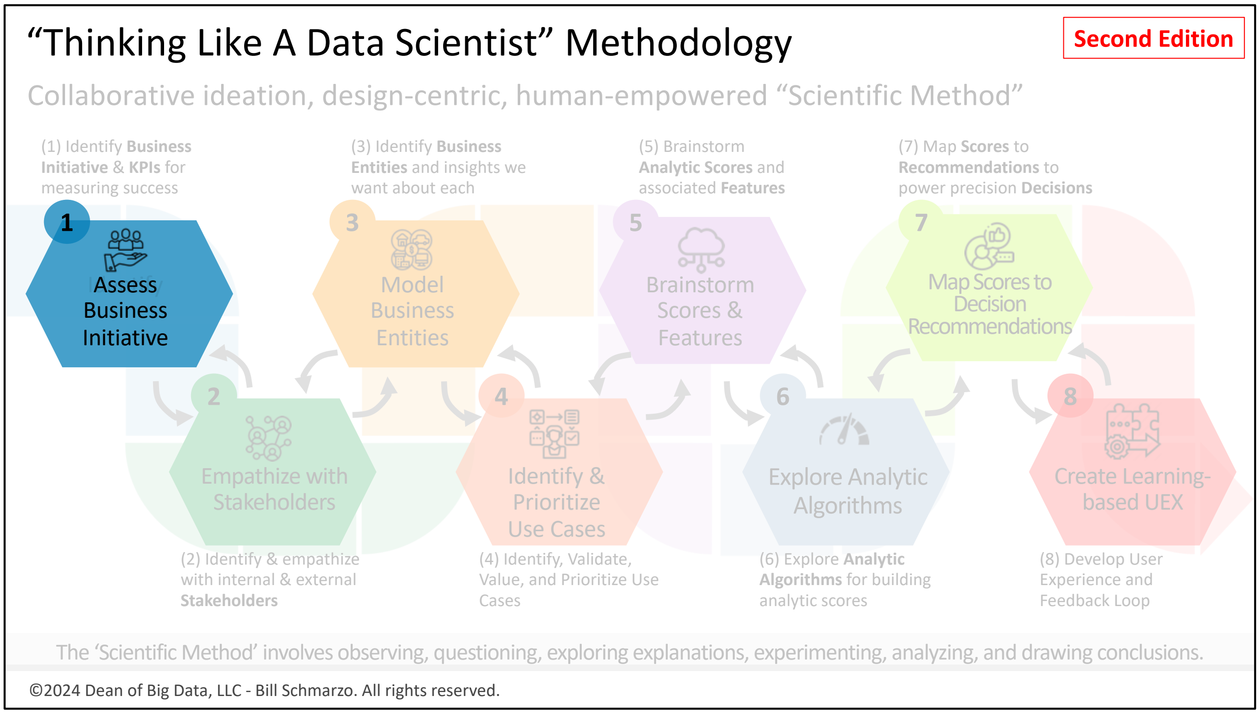


Figure 19: Design Thinking: The Empowerment of Ideation

Design thinking is not only about improving products or services. It has the potential to significantly shift an organization's culture. By creating a work environment that promotes open sharing of ideas, learning from mistakes, valuing all contributions, and promoting empathy, design thinking can encourage employees to be more innovative. This can help organizations become more resilient and better prepared to flourish in an ever-changing business landscape.

We have prepared ourselves to combine data science, design thinking, and economics to provide more relevant, meaningful, responsible, and ethical outcomes. We are ready to apply the "Thinking Like a Data Scientist" methodology!

Chapter 5: TLADS Step 1 – Assess Business Initiative



Step 1 of the TLADS methodology identifies and assesses the targeted business initiative an organization wants to address in the next 12 to 18 months. During this stage, the focus is on defining the strategic objectives, stating the desired outcomes, and prioritizing the metrics to measure the initiative's success. This approach helps ensure that the chosen initiatives align with the organization's long-term vision and market dynamics, setting a strong foundation for the subsequent steps of the TLADS methodology. Data assets and analytic capabilities will be systematically aligned with these business imperatives in the following stages.

What is a Business Initiative?

> *A **business initiative** is a cross-functional effort that supports the organization's objectives and delivers significant value. It has clearly defined and measurable goals, a specific time frame, and a senior executive who takes ownership of the initiative. A business initiative can address either an opportunity or a challenge the organization faces in its market or industry.*

A **Business Initiative** is characterized by the following:

- **Business opportunity:** It addresses a potential or existing business or operational problem or opportunity the organization faces in its market or industry.
- **Creates value:** It delivers a significant, compelling, and distinguishable financial, organizational, customer, operational, or competitive advantage that enhances its market position and improves operational performance.

- **Critical importance:** It is vital for the organization's immediate performance (over the next 12 to 18 months).
- **Documented and communicated:** It is communicated internally or publicly to drive cross-organizational alignment, focus, and transparency.
- **Cross-functional:** It involves more than one business function, requiring close collaboration across different teams and departments.
- **Senior management champion:** It is advocated by a senior business executive who is accountable for its success and has the authority to allocate resources and resolve issues.
- **Measurable:** It has well-defined and quantifiable goals, KPIs, and metrics against which to track its progress and outcomes.
- **Time-bound:** It has a clear and realistic delivery time frame that is agreed upon by all stakeholders.

Focusing on a business initiative to start your AI and data journey aligns your data, analytics, and AI strategy with your organization's core business strategy. By framing your data and analytics strategy around business initiatives, you can prioritize the most valuable business and operational use cases, allocate your data and analytic resources and budget effectively, and track and communicate your AI outcomes and benefits.

📝 **Note:** It is not advisable to begin your AI and data journey by just inventorying your organization's use cases. The use cases identified may not be significant enough to warrant the time and resources required for successful data science and AI implementation from a customer, product, service, or operational value creation perspective. Starting with a business initiative is one way to ensure you identify interrelated use cases of significant value.

Focusing on a business initiative can also help overcome the organizational challenges of AI adoption. A clear business case and senior executive champions can secure stakeholder support and buy-in and foster a culture of collaboration and innovation across your organization.

"Assess Business Initiative" Canvas

In the TLADS methodology, Step 1 involves completing the "Assess Business Initiative" design canvas. This canvas evaluates the factors contributing to successfully executing the targeted business initiative (see Figure 20).

Step 1: Assess Business Initiative

Template 1 of 9

Business Initiative:	Chipotle: Increase "Same Store Sales" by 7.1% over the next 12 months
KPIs:	Average Sales per Visit, Store Traffic per Hour / Day, Sales per Employee, Average Line Wait Time, % On-line Cart Abandonment, % Mobile Orders, Positive Social Media Mentions, Likelihood to Recommend, Table Turns, % of supplies locally-sourced, % organic ingredients, % compostable Waste, % Recyclable Waste, Minority Hiring %, Minority Management %, Charity Giving, Community Programs
Desired Outcomes	• Increase Store Traffic • Increase Shopping Bag Revenues • Acquire Net New Customers • Increase Customer Repeat Visits • Increase Supplier Reliability & Quality • Increase Customer Satisfaction • Increase Positive Social Media mentions • Reduce carbon footprint • Increase locally-supplied ingredients • Increase employee satisfaction • Increase workforce diversity • Increase community satisfaction

Benefits	Potential Impediments
• Increase store sales and profitability drives more promotion opportunities for employees of all diversities • Increase store sales and profitability increases co-opt marketing funds • Increase in organic, locally-supplied food supplies stays supports Chipotle's mission • Reduction in carbon waste (and increase in compostable and recyclable waste) supports Chipotle's mission • Increase in community involvement and charity giving increases relevance of Chipotle brand	• Can't find enough locally-sourced organic food suppliers • Can't hire enough of the right skilled people to support store sales and profitability goals • Local competitors are more appealing to the local demographics • Dearth of successful new product introductions • Supply chain volatility • Pricing pressure from suppliers

Failure Ramifications	Unintended Consequences Ramifications
• Limited promotional opportunities for employees • Lay employees off, reduce hours, miss bonuses • Close store - negatively impacts employees, customers, suppliers, community	• Hiring & promotion programs limit highly-qualified non-minority candidates • Non-sponsored local events offend those participants & community • Sponsorship of questionable events leads to store boycotts • Increase in product demand requires use of non-organic products

Version 2.0

Figure 20: TLADS Step 1: Assess Business Initiative

The Assess Business Initiative canvas drives collaboration between the business and data science teams in understanding and assessing an organization's targeted business initiative. The Assess Business Initiative canvas captures the following business initiative specifications:

- **Business Initiative**: This panel captures the business initiative's key objectives, which are what the business is trying to achieve financially and operationally with this business initiative over the next 12 to 18 months. For the Chipotle example, that focus will be on increasing same store sales by 7.1% over the next 12 months.
- **Key Performance Indicators (KPIs)**: Here, you determine the KPIs and metrics that will be used to measure the effectiveness, track progress, and ultimately achieve the business initiative's success. For the Chipotle example, we might capture and track operational effectiveness, customer experience, employee satisfaction, and supplier quality and reliability metrics.
- **Desired Outcomes**: In this panel, you identify the desired outcomes from financial, operational, customer, employee, community, environmental, and ethical perspectives. For the Chipotle example, we might have the desired outcomes of increasing repeat customer visits, employee retention, and more effective local marketing campaigns.
- **Benefits**: This panel focuses on identifying the potential business and operational benefits expected from the successful execution of the business initiative. For the Chipotle example, benefits could include increasing employee promotional opportunities and increasing the financial viability of the local organic farmers.
- **Potential Impediments**: In this panel, you identify potential obstacles or challenges that may hinder the successful implementation of the targeted business initiative. For Chipotle,

for example, impediments might include employee training and development, supply chain concerns, and the actions of local competitors.

- **Failure Ramifications**: Here, you consider the financial, operational, customer, employee, community, environmental, and ethical ramifications that might result from this business initiative's failure. For Chipotle, failure ramifications might include decreased employee bonuses, funding for local community events, and even the restaurant closing.
- **Unintended Consequences Ramifications**: This panel explores the unintended consequences of this business initiative's success, including the potential second—and third-order ramifications. For the Chipotle example, unexpected success might create supply chain problems for organic food and decrease customer satisfaction from increased store wait times.

📝 **Note:** In the updated version of the "Assess Business Initiative" canvas, I separated failure ramifications from unintended consequences ramifications. Identifying and implementing the KPIs and metrics to monitor and avoid unintended consequences is a significant risk area for AI models that can make decisions millions of times faster than humans.

Identifying / Brainstorming KPIs and Metrics Process

It is important to determine and discuss the Key Performance Indicators (KPIs) and metrics that will be used to evaluate the advancement and success of a business initiative. Think of the KPIs/metrics panel of the "Assess Business Initiatives" panel as a bucket in which to gather the different KPIs and metrics that we will identify through the "Assess Business Initiative" process:

- **Step 1: Brainstorm Desired Outcomes** and the KPIs and metrics against which outcomes' effectiveness will be measured.
- **Step 2: Identify Benefits** and the KPIs and metrics against which their effectiveness will be measured.
- **Step 3: Brainstorm Potential Impediments** and identify the KPIs and metrics against which potential impediments will be identified and tracked.
- **Step 4: Brainstorm Failure Ramifications** and identify the KPIs and metrics against which failure ramifications will be identified and tracked.
- **Step 5: Brainstorm Unintended Consequences Ramifications** and identify the KPIs and metrics against which potential unintended consequences will be identified and tracked.

Here is another example of a completed "Assess Business Initiative" canvas from my students at Iowa State University, this time focused on the John Deere business initiative of growing its EV/Hybrid equipment customer base (Figure 21).

Step 1: Assess Business Initiative

Template 1 of 9

Business Initiative: John Deere: Grow sales of Hybrid/Electric powered machines by 9%

KPIs: Hybrid-Electric Sales / Total Sales, Gross Profit / Hybrid-Electric Equipment, Electric/Hybrid Efficiency, Customers Willing to Switch, % of New Customers, % of Market Share, % of Eclectic Market Share, Hybrid-Electric Production / Total Production, Early (Quarterly) Total Greenhouse Gas Emissions, Greenhouse Gas Emissions Per Equipment, % Completion Towards Sustainability Goals, Total Incentives Received / Total Sales, % Revenue from Electric Products, Customer Satisfaction, Likelihood to Recommend, Customer Satisfaction with Electric Products, % of Customers with the Ability to Buy, Average Production Time Per Unit, Average Wait Time Per Customer

Desired Outcomes

- Increase market share in the electric equipment market
- Expand production capabilities of electric equipment
- Reduce emissions/promote sustainability
- Become a leader in ag. technology/differentiate from competition
- Deliver on sustainability/revenue goals set to stakeholders
- Increased industry competitiveness and innovation from new technologies
- Reduced environmental and health impacts by promoting recycling / reuse of batteries and components

Benefits

- Possible Government and Tax Incentives.
- New Revenue Streams and Product Offerings.
- Environmental Improvements (reduced emissions and increased re-usable energy)
- Improved Customer and Media Reputation
- Reduced fuel and maintenance costs for customers
- Enhanced operator comfort and safety with lower noise and vibration levels
- Increased customer loyalty and retention with innovative and eco-friendly solutions
- Improved brand image and reputation as a leader in sustainability and technology

Potential Impediments

- Reluctance of customers to change to battery power
- Customer dissatisfaction with charge time.
- Financial struggles regarding buying new equipment.
- Supply Chain issues concerning industry movement towards electric/hybrid.
- Higher initial investment and operating costs for electric vs conventional equipment
- Limited availability of charging infrastructure and battery replacement services
- Technical challenges & uncertainties regarding battery performance, durability, safety
- Stakeholder resistance about the feasibility and profitability of electric equipment

Failure Ramifications

- Legal and reputational damages due to accidents, injuries, or environmental impacts
- Regulatory and compliance challenges due to changing standards, policies, or incentives
- Customer dissatisfaction and complaints due to technical issues, breakdowns, or recalls
- Potential financial Fees or Fines for not meeting Carbon emission cutbacks

Unintended Consequences Ramifications

- Social and ethical implications of battery disposal, recycling, or reuse
- Trade-offs and conflicts between environmental, economic, and social goals and values
- Increased demand for electricity strains grid during peak hours or extreme weather events
- Reduced tax revenue from gasoline sales affects roads and public transportation funding

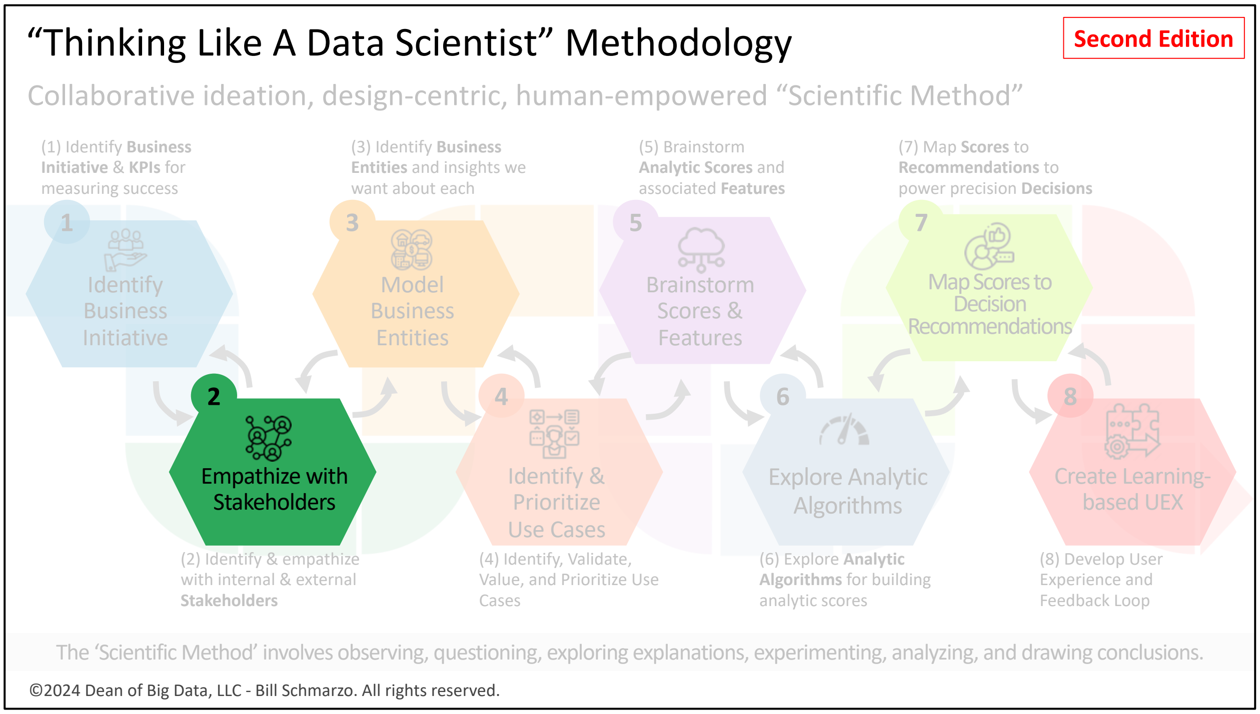
Version 2.0

Figure 21: John Deere: Grow sales of Hybrid/Electric powered machines by 9%

Step 1: "Assess Business Initiative" Summary

Chapter 5 – "Step 1: Assess Business Initiative" – explains how to align organizational strategies with an actionable data science, data, and AI strategy. The first step identifies and evaluates a targeted business initiative the organization needs to address within 12 to 18 months. This involves setting clear strategic objectives, defining desired outcomes, and identifying the KPIs and metrics to guide the initiative's success. The process uses the "Assess Business Initiative" canvas to drive business and data science team collaboration to capture and analyze critical success criteria, such as the business initiative's goals, desired outcomes, benefits, potential impediments, and risks. This foundational step ensures that the subsequent phases of the TLADS methodology are strategically aligned to support the organization's overarching goals.

Chapter 6: TLADS Step 2 - Empathize with Stakeholders



Step 2 of the TLADS process significantly emphasizes understanding the perspectives, pain points, and priorities of all stakeholders involved. This step fosters a culture of **empathy** that is essential for crafting solutions that are effective, inclusive, and well-received. This step ensures that subsequent strategies and actions are deeply rooted in the needs and expectations of those who will be most affected, paving the way for more engaged collaboration, stakeholder adoption, and successful outcomes.

> ***Empathy*** *is the capacity to understand and share the feelings, thoughts, and experiences of another from one's perspective rather than one's own.*

Empathizing with our business stakeholders is a multifaceted endeavor that goes beyond merely understanding their challenges and needs; it requires emotional investment to connect with their experiences and desired outcomes on a human level. This deep empathy, cultivated through design thinking, acknowledges the diverse and sometimes conflicting interests of all stakeholders affected by our targeted business initiative, appreciating the subtleties of their experiences and engaging with them to create practical and emotionally resonant solutions.

> ***Business Stakeholders*** *are those business users or functions (Sales, Marketing, Finance, Logistics, etc.) that either impact or are impacted by the targeted business initiative.*

Design thinking provides the framework, concepts, and tools necessary to achieve this understanding. We capture a comprehensive picture of user intentions, desired outcomes, critical decisions, gains (personal benefits), and pains (personal impediments) by employing empathy maps, personas, user interviews, ideation workshops, and observations. Additionally, customer journey maps allow us to visualize the complete spectrum of user interactions, identifying pain points and opportunities for improvement. These techniques enable us to gain an insightful, empathetic perspective informing each design stage, ensuring our solutions are tailored to our stakeholders' genuine needs.

The "Empathize with Stakeholders" canvas captures critical information about the organization's key internal and external stakeholders who will be impacted by the targeted business initiative (Figure 22).

Steps 2: Empathize with Stakeholders

Template 2 of 9

Business Initiative: Chipotle: Increase "Same Store Sales" by 7.1% over the next 12 months

(2) Stakeholders

Who are the key stakeholders who impact or are impacted by business initiative and why is initiative important to them?

Stakeholder	Importance to Stakeholder	Potential Impediments	Key Decisions	Desired Outcomes	KPIs / Metrics
Customer	Enjoyable experience; high-quality product at a fair price; eat healthy while helping the environment	• Budget / Cost • Time availability • Convenience	• Eat out / order in • Where to go eat • What to order	• Enjoyable restaurant experience • Great tasting, affordable meal • Fast and efficient service	Customer Sat, Likelihood to Recommend, # Repeat Visits, Recency of Last Visit
Store Manager	Achieve monthly and quarterly performance numbers, getting bonuses, and promotion	• Lack timely performance insights • Data not actionable	• Store Staff Scheduling • Inventory Management • Production Management • Customer Satisfaction	• Happy, repeat customers • Happy employees • Strong word of mouth recommendations	Revenue #'s, Customer Sat, Employee Sat, Waste & Shrinkage
Field Marketing	Run successful local marketing events that reflect their own creativity	• Identifying local marketing opportunities • Assessing performance	• Local Events Marketing • Social Media Marketing	• Successful campaigns • Strong word of mouth recommendations	Local Events Effectiveness, First Time Try, Store Traffic
Corporate Marketing	Successful marketing campaigns that deliver ROI, leads, and customer satisfaction	• Prioritize spend • Lack visibility into local marketing campaigns	• Marketing Campaigns • New Product Introductions	• Strong customer social sentiment • Successful campaigns • WOM recommendations	Campaign Effectiveness, New Product Intro Effectiveness, Channel Effectiveness
Procurement	Predictability of locally sourcing high-quality food ingredients at a fair price	• Finding reliable suppliers • Monitoring supplier performance	• Inventory Management • Supplier Management • Pricing	• Highly reliable demand forecasts • Reliable, high-quality suppliers • Cost competitive suppliers	Vendor Quality, Vendor Reliability, Vendor Financial Rating, Diversity Mix

Version 2.0

Figure 22: TLADS Step 2: Empathize with Stakeholders

We will capture the following information in the "Empathize with Stakeholders" canvas:

- **Business Initiative**: This refers to the name and key objectives of the business initiative that we are targeting. In our example, this is to increase same store sales by 7.1% over the next 12 months.
- **Stakeholder**: This refers to the list of key individuals or groups the targeted business initiative will impact. Examples could include customers, store managers, field marketing, and procurement.
- **Importance to Stakeholder**: This explains why the targeted business initiative is important to the stakeholders and their jobs. For Store Managers, this might mean achieving monthly and quarterly performance numbers, getting bonuses, and getting promoted.
- **Potential Impediments**: This refers to a list of potential impediments the stakeholders may face that could hinder their ability to participate effectively in the targeted business

initiative. For Store Managers, potential impediments might be the lack of timely performance insights and non-actionable reports and dashboards.

- **Key Decisions**: This refers to a list of the key decisions the stakeholders need to make to support the targeted business initiative. For Store Managers, key decisions might include store staffing, inventory control, and production management.
- **Desired Outcomes**: This refers to a list of the desired outcomes the stakeholders wish to achieve through the targeted business initiative. For Store Managers, these might include repeat customers and satisfied employees.
- **KPIs /Metrics**: This refers to a list of the KPIs and metrics against which the stakeholders will measure the effectiveness of the decisions and outcomes associated with the targeted business initiative. For Store Managers, critical KPIs and metrics might include Revenue, Customer Likelihood to Recommend, Employee Satisfaction Ratings, and Percentage Waste and Shrinkage numbers.

Here is an example of a completed "Empathize with Stakeholders" canvas from my students at Iowa State University, focusing on John Deere's business initiative to grow its EV/Hybrid equipment customer base (Figure 23).

Steps 2: Empathize with Stakeholders

Template 2 of 9

Business Initiative: John Deere: Grow sales of Hybrid/Electric powered machines by 9%

Stakeholder	Why Important to Them?	Impediments	Outcomes	Decisions	KPIs
Customer (Farmer)	Reduce costs, improve environment, enhance productivity and profitability	Reluctance, high costs, limited infrastructure, technical challenges	Reliable, efficient, and cost-effective electric equipment that meets needs, reduces emissions, and increases competitiveness and innovation	Buy/lease/rent, models/features, finance/operate, infrastructure/services, integrate/technologies	$ and maintenance savings, emission reductions, productivity and profitability improvements, satisfaction and loyalty, innovation and differentiation
Dealership	Increase sales, revenue, and market share, improve service and retention, differentiate from competitors	Resistance, financial struggles, supply chain issues, technical issues, regulatory challenges	Loyal, satisfied, and profitable customers, increased demand and sales, reduced costs and risks, improved brand and reputation, competitive advantage and leadership	Market and sell, finance and lease, infrastructure and services, train and support, partner with stakeholders	Sales, revenue, and market share, service and retention, costs and risks, brand and reputation, competitive advantage and leadership
R&D Team	Develop and deliver innovative and sustainable solutions, meet customer needs and expectations, comply with standards	Technical challenges, limited resources, changing demands, complex requirements	High-quality, reliable, and safe electric equipment that provides superior performance and efficiency, reduces emissions and impact, and creates value	Technologies and solutions, design and test, optimize and integrate, collaborate and communicate	Performance and efficiency, emission and impact, quality and reliability, safety and security, innovation and value
Shareholders	Increase value and growth, ensure sustainability and profitability, align with values	Competition, volatility, churn, damages, challenges	Increased sales, revenue, and market share, reduced costs and risks, improved brand and reputation, enhanced competitiveness and innovation, positive impact	Invest or divest, how much, evaluate and monitor, engage and influence	Sales, revenue, and market share, costs and risks, brand and reputation, competitiveness and innovation, impact
Suppliers	Maintain and expand relationship, provide quality and timely products and services, adapt to demands	Disruptions, issues, negotiations, challenges, pressure	Stable, long-term, and profitable contracts, satisfied and loyal customers, reduced costs and risks, improved efficiency and productivity, competitive edge and opportunities	Products and services, price and delivery, quality and performance, collaborate and communicate, innovate and differentiate	Contracts, revenue, and profit, satisfaction and loyalty, costs and risks, efficiency and productivity, innovation and differentiation

Version 2.0

Figure 23: John Deere "Empathize with Stakeholders" Canvas

Ensure that you assess a broad and diverse community of stakeholders to capture the success criteria and involve the key folks who will ultimately be the recipients and users of the resulting analytics. The best way to ensure the adoption of the resulting analytics is to involve those who will be using those analytics in the definition, development, and ongoing management of those analytics.

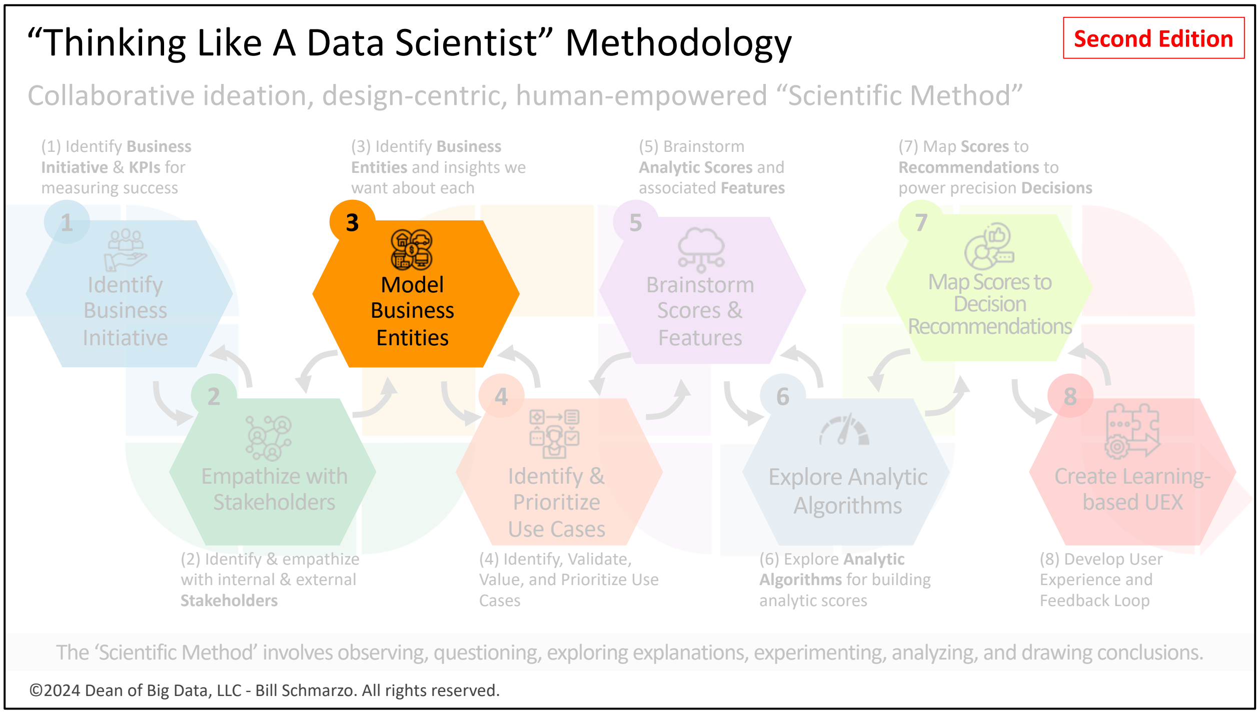
One common mistake in Step 2 is considering only a narrow range of internal stakeholders. Understanding your internal and external stakeholders is critical to your data science success. Not only might they have valuable input that helps guide the success of the analytic model development, but it also helps to ensure that the necessary implementation and operational requirements are captured (which we explore further in TLADS steps 7 and 8).

Note: The stakeholders' desired outcomes, critical decisions, and KPIs and metrics against which they will measure decision and outcome effectiveness become critical in Step 4 when we brainstorm the potential use cases and again in Step 5 when we ideate the analytic scores and associated ML features to help optimize our use cases. Again, do not shortchange this step!

Step 2: "Empathize with Stakeholders" Summary

TLADS Step 2, "Empathize with Stakeholders," is about understanding the stakeholders' desired outcomes, the critical decisions they need to make, and the metrics they'll use to measure the success and effectiveness of the targeted business initiative. By actively collaborating with the key stakeholders, data science teams can create more effective and relevant analytic models that support and advance the targeted business initiative. It's equally important to involve stakeholders in defining, designing, and deploying analytics, as this increases the likelihood of them using the results of the models in their daily operations. After all, what's the point of investing time and money in building advanced analytic models if the stakeholders aren't invested in using the insights they generate?

Chapter 7: TLADS Step 3 - Model Business Entities



Step 3 of the TLADS methodology emphasizes meticulously cataloging and analyzing key business entities' predictive behaviors, inclinations, and tendencies. This process leads to creating granular analytic models that catalyze optimizing our prioritized use cases. An in-depth understanding of the business entities' predictive behavioral and performance propensities is crucial for exploiting nanoeconomics and driving the precision decisions necessary to create new sources of customer, product, service, and operational value.

> 💡 *A **Business Entity** is a person or device that relates to an organization's key business initiative and has measurable behaviors or performance that enable the creation of analytic scores.*

Examples of business entities include (Figure 24):

- **Humans**: Customers, patients, nurses, doctors, students, teachers, technicians, operators, baristas, architects, employees, parolees, and police officers.
- **Devices**: Trucks, cars, trains, turbines, jet engines, motors, compressors, chillers, presses, clutches, brakes, transmissions, air conditioners, furnaces, and outlets.

Business Entities

Business Entities *are humans or devices that organizations seek to uncover and quantify their individual predicted behavioral and performance propensities (**analytic insights**).*

We apply the **analytic insights (or predictive propensities)** of the individual **Entities** to yield individualized **Precision Decisions** to optimize business and operational **Use Cases**

Human Entities
Customers, patients, nurses, doctors, students, teachers, technicians, operators, baristas, architects, employees, parolees, police officers

Device Entities
Trucks, cars, trains, turbines, jet engines, motors, compressors, chillers, presses, clutches, brakes, transmissions, air conditioners, furnaces, stores

Figure 24: Business Entities

Note: In the process of empathizing with stakeholders in Step 1, we may capture some information that will affect the modeling of business entities. This applies particularly when a human stakeholder, such as a patient, nurse, doctor, or technician, is also considered a key business entity. However, it is important to note that not all business entities are human, as they can also include treatments, procedures, CT scan machines, MRI machines, the ER, the ICU, and other non-human entities.

Template #3: "Model Business Entities" Canvas

In TLADS Step 3, we want to brainstorm and capture relevant information about each business entity to help us predict their behaviors or optimize their performance. This important business entity information will be captured in the "Model Business Entities" canvas (Figure 25).

Step 3: Model Business Entities

Template 3 of 8

Business Initiative: Chipotle: Increase "Same Store Sales" by 7.1% over the next 12 months

(3) Business Entities

What are the business initiative's key Business Entities (human or device/thing) and what insights would we want to know about each?

Entity	Entity Description	Predictive Insights
Customers	Purchase and consume products and services	Product preferences, location preferences, typical visit time and day of week, frequency of visits, recency of visit, ordering method, promotion responsiveness, weekly restaurant eating, organic importance, sustainability importance, social issues importance
Stores	Physical locations where customers order, pick up, or dine in	Location, store traffic patterns, local demographics, nearby local venues and schools, product preferences, nearby transportation, last remodel date, next remodel date, store capacity
Employees	Work in restaurants to produce products	work tenure, performance ratings, wages history, promotions history, demographics, education level, income levels, training certifications, management potential
Local Events	Activities relevant to Chipotle's customers or communities	Dates of events, times of events, event duration, type of event, frequency of events, number of participants, age of participants, local or national event
Suppliers	Provide ingredients and materials for products	Financial history, financial viability, supplier performance ratings, product quality, delivery reliability, minority-owned, conservation efforts
Competitors	Offer competing products to Chipotle's customers	Location, size of store, last remodel date, store capacity, products, price points, customer demographics, category rank, local advertising spend, national advertising spend
Tortilla Press	press and heat tortillas for burritos and tacos	Next Service Date, Estimated cost to repair, Estimated cost to replace, Remaining Useful Life, Operating Effectiveness, Salvage Value, Hourly capacity

Version 2.0

Figure 25: "Model Business Entities" Canvas

Our goal is to capture the following information for each business entity on the "Model Business Entities" canvas:

- **Entity:** This is the name of the human and device entity around which we will build our analytic models. Our Chipotle example would include Customers, Stores, Suppliers, Employees, and Competitors.
- **Entity Description.** We want to briefly describe the entity and its role in support of the targeted business initiative. For example, Chipotle Stores are the physical locations where customers order, pick up food, or dine in.
- **Predictive Insights.** Next, we want to capture the predictive insights or behaviors we might want to know about each entity. We will leverage these entity-level insights to guide the define the analytic scores in Step 5. For example, we might want to capture and codify the following predictive propensities for our customers: Product preferences, location preferences, typical visit time and day of the week, frequency of visits, recency of visit, ordering method, promotion responsiveness, weekly restaurant eating, organic food importance, sustainability importance, and social issues importance.

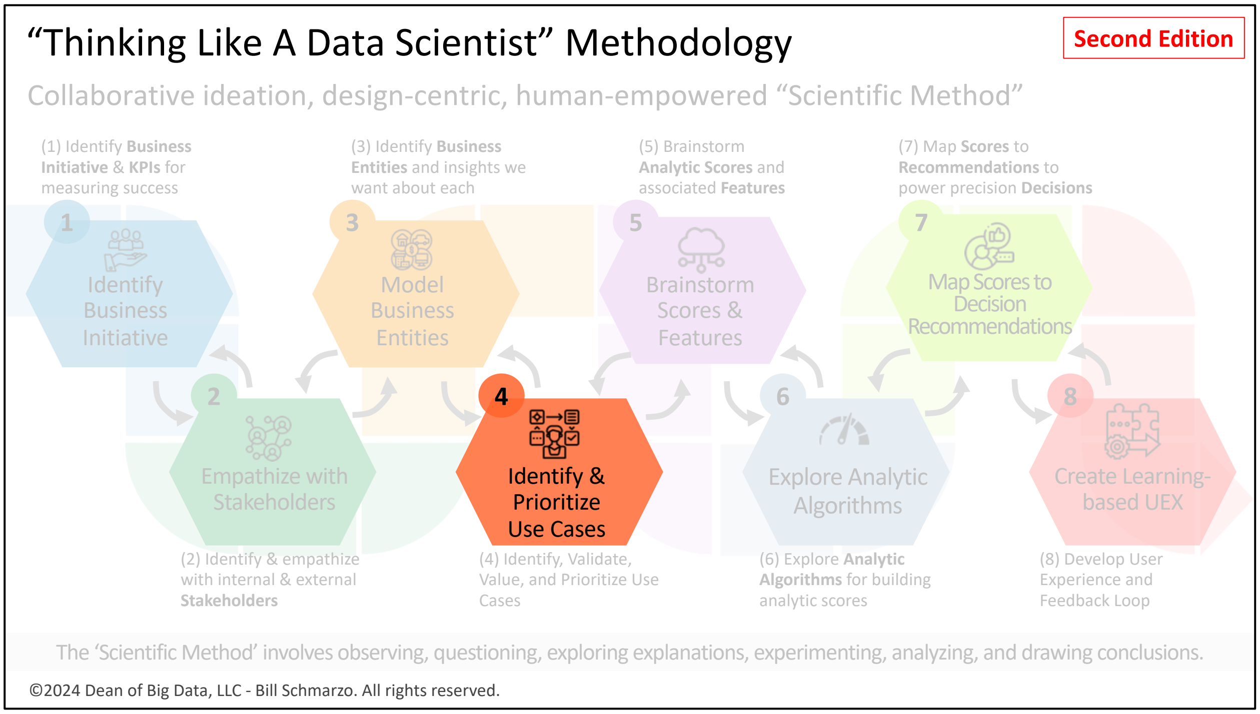
Step 3: "Model Business Entities" Summary

Step 5 of the TLADS methodology leverages nanoeconomics to create advanced AI models that analyze large datasets and provide actionable analytic scores. These analytic scores help predict the behavior and performance of key business entities. By utilizing this information, we can anticipate customer buying patterns, prevent employee turnover, predict health-related events more accurately, forecast outcomes in sporting events, predict device maintenance and reduce unplanned operational downtime.

The process of generating entity-level analytic scores is essential for making **precise decisions**. This term refers to decision-making that is accurate and relevant to the context. Such targeted decisions are crucial in fine-tuning a business's operational levers, enhancing efficiency, customer satisfaction, risk management, and overall organizational agility. Doing so can improve current operational outcomes and strategically align our operations with future business objectives and market demands.

The TLADS methodology aims to generate analytical scores for individual entities. These scores are used to make precise decisions to optimize and re-engineer our key business and operational use cases. To achieve this goal, it is essential to thoroughly understand the key entities involved in the business initiative and model them accordingly.

Chapter 8: TLADS Step 4a – Identify & Assess Use Cases



📝 **Note:** Step 4 is actually divided into two sub-steps: 4a), which aims to identify potential use cases from the organizational and stakeholder information captured in TLADS Steps 1 and 2, and 4b), which aims to prioritize those use cases based on their organizational value and feasibility of implementation.

In the TLADS methodology, Step 4a involves categorizing the desired outcomes, decisions, and metrics identified in Step 2 into common subject areas or use cases that support the targeted business initiative. This process helps to identify the main focus areas and prioritize data science efforts toward achieving specific business or operational goals. For instance, some sample use cases for an Optimize Customer Lifecycle Management business initiative might include improving customer acquisition, cross-selling and upselling, retention, and advocacy.

> 💡 ***Use Case** is a narrative or blueprint outlining a business or operational scenario, detailing decisions, desired outcomes, and success metrics to streamline and enhance specific actions or series of actions in alignment with an organization's strategic business initiative.*

Use cases drive collaboration between business and data teams by:

- Creating a **common vocabulary** encapsulates the organization's business, operational goals, and desired outcomes.
- Developing an **in-depth understanding** of the targeted business initiative that guides the application of data and analytics to enhance business and operational performance.

- Ensuring strategic focus and optimal resource allocation by identifying, valuing, and prioritizing the use cases that advance the targeted business initiative.
- Fueling ideation with business stakeholders to identify and refine the variables and metrics (ML features) most effectively predict and optimize use case success (Figure 26).

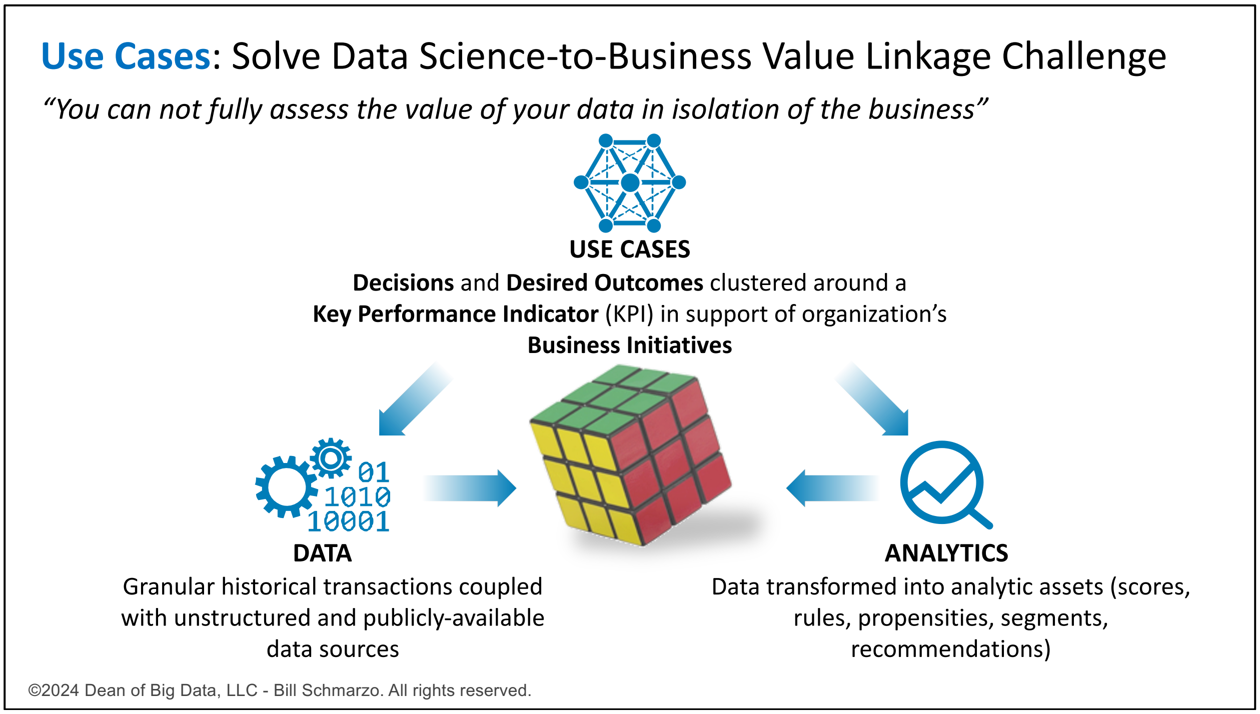


Figure 26: Use Cases Solve the Technology-to-Business Linkage Challenge

Identifying Potential Use Cases Process

Through TLADS Steps 1 and 2, we have gathered a wealth of organizational and stakeholder information with respect to our targeted business initiative. In Step 4, we will use that information to identify potential use cases that support our targeted business initiative. Here's the process for how we'll identify the potential use cases:

- **Step 1: Consolidate Organization and Stakeholder Information**: Consolidate the desired outcomes, key decisions, and KPIs/metrics from "Step 1: Assess Business Initiative" and "Step 2: Empathize with Stakeholders." For example, desired outcomes could include Improving Supplier Reliability, Improving Supplier Quality, Acquiring Cost-Competitive Suppliers, Increasing Store Traffic, Improving Successful Marketing Campaigns, and Improving Customer Social Sentiment.
- **Step 2: Categorize or Cluster Information**: Group the collected desired outcomes, decisions, and KPIs/Metrics into similar categories. For example, the desired outcomes of Improving Supplier Reliability, Improving Supplier Quality, and Acquiring Cost-Competitive Suppliers could cluster into an "Optimize Supplier Management" use case. The desired outcomes of Increasing Store Traffic, Improving Successful Marketing Campaigns, and Improving Customer Social Sentiment could cluster into an "Improve Marketing Effectiveness" use case.

- **Step 3: Assess Use Case Relationships**: To validate the use case clusters, evaluate how each use case interacts with the other use cases in supporting the targeted business initiative. For example, what is the potential interplay between the "Optimizing Supplier Management" and "Improve Marketing Effectiveness" use cases? For example, dramatically improving marketing effectiveness could surface challenges in supplier product quality and delivery reliability.
- **Step 4: Expand Use Case KPIs/Metrics**: Brainstorm additional KPIs and metrics that might support and optimize the potential use case. For example, we might want to consider additional metrics supporting our "Optimizing Supplier Management," such as supplier financial viability, supplier diversity policies, and supplier regulatory compliance.

Identifying Potential Use Cases – Chipotle Example

Now, let's deep dive into the Chipotle "Increase Same Store Sales" business initiative. We have collected desired outcomes, key decisions, and KPIs / metrics from TLADS Template 1: "Assess Business Initiative" canvas and Template 2: "Empathize with Stakeholders" canvas supporting our "Increase Same Store Sales" business initiative at Chipotle, including:

- **Desired Outcomes:** Increase Store Traffic, Increase Shopping Bag Revenues, Acquire Net New Customers, Increase Customer Repeat Visits, Increase Supplier Reliability, Increase Supplier Quality, Increase Customer Satisfaction, Increase Positive Social Media mentions, Reduce carbon footprint, Increase locally supplied food ingredients, Increase employee satisfaction, Increase workforce diversity, Increase community satisfaction, Increase Employee Satisfaction, Increase Positive word of mouth recommendations, Achieve Strong customer social sentiment, Improve demand forecasts reliability, Retain high-quality suppliers, and Acquire Cost-competitive suppliers
- **Decisions**: store staff scheduling, hiring, firing, inventory management, production management, customer satisfaction, local events marketing, social media marketing, social media engagement, local advertising, local promotions, new product introductions, pricing, and supplier management.
- **KPIs / Metrics**: Average Sales per Visit, Store Traffic per Hour / Day, Sales per Employee, Average Line Wait Time, % of online cart Abandonment, % Mobile Orders, Positive Social Media Mentions, Likelihood to Recommend, Table Turns, % of supplies locally-sourced, % organic ingredients, % compostable Waste, % Recyclable Waste, Minority Hiring %, Minority Management %, Charity Giving, and # of Community Programs.

By analyzing relationships, overlaps, and similarities amongst these desired outcomes, key decisions, and KPIs/Metrics, we can identify some potential use cases, such as:

- **Sales Performance & Efficiency**: This use case focuses on maximizing revenue through enhanced service speed and customer turnover. It uses metrics such as average sales per visit and table turns to optimize staff scheduling and store layout. By analyzing store traffic and sales per employee, Chipotle can refine operational efficiency, ensuring that first-time customers become regular patrons.

- **Digital Engagement & Online Ordering**: This use case examines the rate of online cart abandonment and mobile order percentages to improve the online ordering system's usability and streamline the digital customer journey. It also measures the effectiveness of marketing campaigns and the introduction of new products to boost digital sales and customer engagement.
- **Customer Experience & Satisfaction**: This use case seeks to create a superior dining experience by reducing line wait times and enhancing overall customer satisfaction. Metrics like the likelihood to recommend, positive social media mentions, and repeat visits are measures of success.
- **Sustainability & Ethical Sourcing**: This use case focuses on Chipotle's commitment to sustainability by increasing the use of locally sourced and organic ingredients and minimizing waste. The compostable and recyclable waste percentage is measured to ensure the brand's operational practices align with its eco-friendly ethos.
- **Diversity & Community Engagement**: This use case emphasizes fostering an inclusive workforce and engaging with the community, with metrics tracking minority hiring rates and management representation. The effectiveness of charity programs and community events reflects Chipotle's commitment to social responsibility and local engagement.
- **Operational Efficiency & Waste Management**: This use case seeks to streamline processes by leveraging data on waste and shrinkage, alongside on-time delivery rates, to minimize resource loss and optimize supply chain operations. Tracking product damage and operational waste contributes to developing strategies that enhance efficiency and sustainability.
- **Vendor & Supply Chain Management**: This use case supports the goal of sourcing high-quality ingredients reliably, which is vital for product consistency and customer trust. By monitoring vendor quality, reliability, and financial stability, Chipotle can ensure a resilient and ethical supply chain.
- **Employee Engagement & Satisfaction**: This use case centers on understanding and improving employee morale, which is crucial for maintaining high-quality customer service. By assessing the effectiveness of local events and employee satisfaction surveys, Chipotle can create a better work environment, contributing to overall operational success.
- **Local Events Marketing**: This use case aims to integrate Chipotle's brand into the community by leveraging local events and festivals as marketing platforms to drive store traffic and engagement. By aligning promotional activities with local events and measuring the impact on store visits and sales, Chipotle can capitalize on community presence to enhance visibility and attract new customers, fostering a sense of local inclusion and brand loyalty.

📝 **Note:** In the GenAI supplement, we'll explore how to use a GenAI tool to assist us in analyzing and clustering the desired outcomes, key decisions, and KPIs/metrics into potential use cases that support our targeted business initiative.

Make sure to label your use cases in an **actionable format**. Start each use case label with an Action Verb such as increase, reduce, optimize, improve, mitigate, rationalize, consolidate, enhance, accelerate, or maximize. Additionally, end each use case label with the generic "by X%" to indicate the need to measure the effectiveness of improving or optimizing that use case.

For example, if our targeted business initiative is "Reduce Unplanned Operational Downtime," the following might be the supporting use cases written in an actionable format:

- [Optimize] [equipment maintenance schedules] [by X%]
- [Improve] [predictive maintenance effectiveness] [by X%]
- [Enhance] [real-time equipment performance monitoring] [by X%]
- [Improve] [first time fix effectiveness] [by X%]
- [Reduce] [obsolete and excessive inventory] [by X%]
- [Improve] [in-stock consumables] [by X%]
- [Accelerate] [real-time training effectiveness] [by X%]
- [Consolidate] [inventory depots to reduce inventory costs] [by X%]

Template #4: "Identify & Assess Use Cases" Canvas

We will capture the key information about each potential use case in the "Identify & Assess Use Cases" canvas (Figure 27).

Step 4a: Identify & Assess Use Cases

Template 4 of 9

Business Initiative: Chipotle: Increase "Same Store Sales" by 7.1% over the next 12 months

Use Case	Desired Outcomes	Key Decisions	KPIs and Metrics
Increase shopping bag revenue by X%	• Increase average order value and customer lifetime value • Increase customer satisfaction • Increase product sales & profits	• Decide promotional objectives • Decide promotional plans (up-selling, cross-sell). • Decide how to evaluate promotional effectiveness	• Shopping bag revenue • Shopping bag profits • Average order size • Customer Lifetime Value • Customer satisfaction score
Increase Store Traffic via local events by X%	• Increase brand awareness • Increase customer acquisition and retention • Increase customer satisfaction • Increase customer referrals	• Decide which local events • Decide promotion type • Decide sponsorship level and investment • Decide social media investment	• Number of participants • Promotion effectiveness • Store foot traffic increase • Sales lift • Social media positive mentions
Improve supplier effectiveness by X%	• Increase supplier quality and reliability • Increase supplier sustainability and innovation • Increase supplier collaboration and trust • Increase supplier satisfaction and loyalty	• Decide quality criteria and standards • Monitor supplier quality and reliability performance • Decide Supplier improvement plans • Decide supplier remediation actions	• Supplier quality score • Supplier reliability score • Supplier sustainability score • Supplier innovation score
Increase Store Traffic via Customer Loyalty program by X%	• Increase customer retention • Increase customer referrals • Increase social media engagement • Increase customer satisfaction	• Decide Loyalty program rewards • Decide Loyalty program promotion plan • Decide loyalty program effectiveness	• Customer retention rate • Customer lifetime value • Customer referral rate • Customer feedback score • Social media engagement score
Improve Product Intro Effectiveness by X%	• Increase customer loyalty • Increase sales and revenue • Increase customer satisfaction	• Decide new product line targets • Decide promotional plans • Decide promotional effectiveness	• Product demand score • Product success score • Product quality score

Version 2.0

Figure 27: "Identify & Assess Use Cases" Canvas

For each potential use case, we will capture the following information in the "Identify & Assess Use Cases" canvas:

- **Use Case Name:** The use case name is labeled in an actionable format, such as “Increase Customer Satisfaction X%,” “Reduce Excessive and Obsolete Inventory by X%,” or “Reduce Unplanned Hospital Readmissions by X%.”
- **Desired Outcomes:** List the desired outcomes the organization and stakeholders seek in supporting that particular use case. For example, for the “Increase Store Traffic via local events by X%” use case, the key desired outcomes might include increasing local brand awareness, increasing customer acquisition and retention, increasing customer satisfaction, and increasing customer referrals.
- **Key Decisions:** List the critical decisions the stakeholders must make to execute or optimize that use case. For the “Increase Store Traffic via local events by X%” use case, we might capture key decisions such as deciding which local events to support, the promotion type, sponsorship level, and the amount of social media investment.
- **KPIs and Metrics:** List the KPIs and metrics against which the organization and stakeholders will measure outcomes and decision effectiveness. For the “Increase Store Traffic via local events by X%” use case, the KPIs and metrics against which we might want to measure decision and desired outcomes effectiveness might include the number of participants, promotion effectiveness, store foot traffic increase, sales lift, and social media positive mentions.

Here is an example of a completed “Identify & Assess Use Cases” canvas from my students at Iowa State University, focusing on John Deere's business initiative to grow its EV/Hybrid equipment customer base (Figure 28).

Step 4a: Identify & Assess Use Cases

Template 4 of 9

Business Initiative: John Deere: Grow sales of Hybrid/Electric powered machines by 9%

(4) Use Cases

Case Name	Key Stakeholders	Key Decisions	Desired Outcomes	KPIs and Metrics
Increase Market Share by X%	Sales, Marketing, Product Management	Market and sell strategies, partner with stakeholders	Increase market share in the electric equipment market	% of new customers, % of market share & % of electric market share, sales growth rate, customer acquisition cost
Expand Production Capabilities by X%	Operations, Supply Chain	Technologies and solutions, design and test	Expand production capabilities of electric equipment	Hybrid-Electric Production / Total Production, capacity utilization rate, production efficiency
Reduce Emissions by X%	Sustainability, Operations	Invest in green technologies, evaluate and monitor	Reduce emissions, promote sustainability	Early (quarterly) total greenhouse gas emissions, Greenhouse gas emissions per equipment, carbon footprint reduction
Enhance Customer Satisfaction by X%	Sales, Customer Service	Train and support, customer feedback systems	Loyal, satisfied, and profitable customers	Customer satisfaction, likelihood to recommend, net promoter score (NPS), customer loyalty index
Enhance Product Efficiency by X%	R&D, Operations	Optimize and integrate technologies, design efficiency	Reliable, efficient, and cost-effective electric equipment	Electric/hybrid efficiency, gross profit per hybrid-electric equipment, energy consumption per unit
Reduce Purchase Impediments by X%	All Stakeholders	Address customer reluctance, supply chain optimization	Overcome potential impediments to electric equipment adoption	% of customers unwilling to switch, average wait time per customer, supply chain efficiency, inventory turnover rate
Increase Equipment Reliability by X%	Operations, Customer Service	Product and service quality evaluations, feedback systems	High-quality, reliable, and safe electric equipment	Quality and reliability metrics, service and retention rates, mean time between failures (MTBF), warranty claim rate
Increase Product Performance by X%	R&D, Marketing	Optimize product features, market feedback analysis	Better performance and efficiency in electric equipment	Performance and efficiency metrics, customer satisfaction ratings, product feature utilization rate, operational performance improvement

Version 2.0

Figure 28: John Deere “Identify & Assess Use Cases” Canvas

Leading versus Lagging Indicators

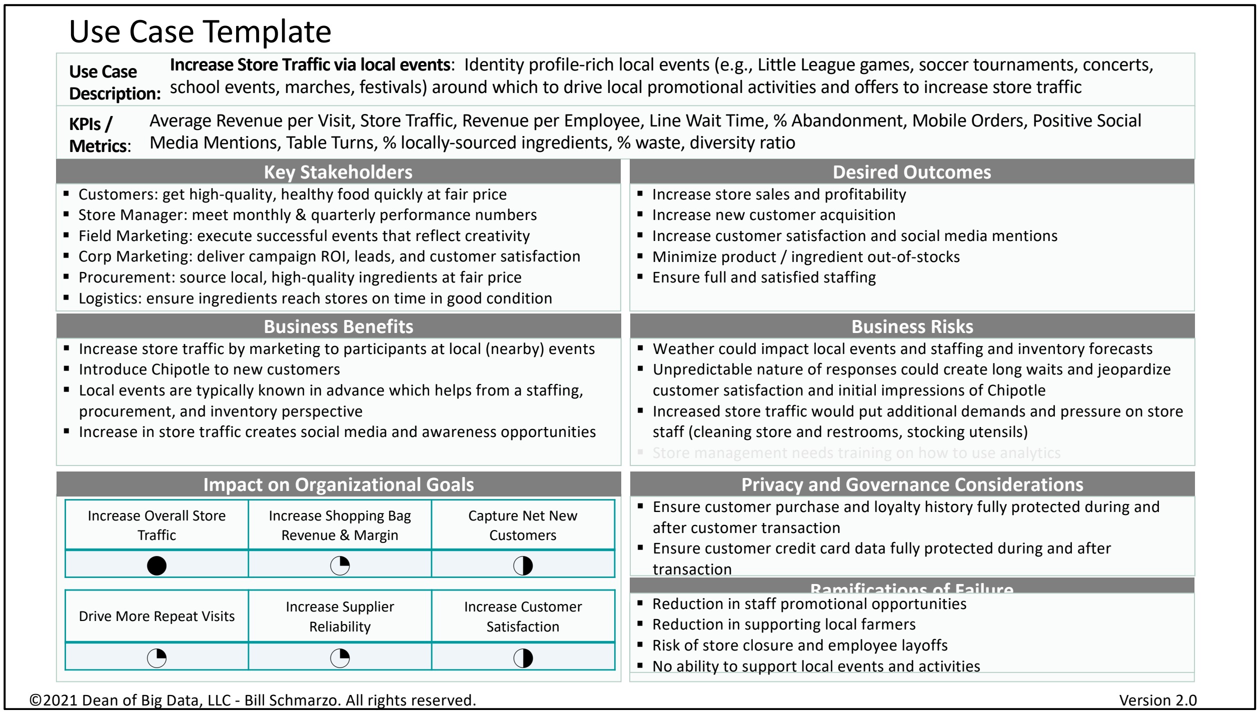
When brainstorming the KPIs and metrics that measure the effectiveness of executing or optimizing a use case, it is critical to distinguish between leading and lagging indicators.

- **Lagging Indicators** represent the aftermath of actions or behaviors, serving as historical benchmarks. These indicators are descriptive and reactive and outline past performance, providing insight into the effectiveness of past actions through metrics such as revenue, profit, and market share. For example, at Chipotle, lagging indicators might include the total quarterly sales revenue or the number of stores opened last year. These offer solid figures on what has been accomplished but don’t suggest actions for future improvement.
- **Leading Indicators**, on the other hand, are forward-looking metrics that anticipate future performance by measuring current activities or behaviors likely to influence desired outcomes. These indicators are predictive, proactive, and instrumental in guiding strategic decisions to foster improvement and goal attainment. Customer satisfaction, employee engagement, and product quality exemplify such forward-thinking metrics. For instance, the number of new followers on Chipotle's social media platforms can serve as a leading indicator, as it may correlate with raising brand awareness and potentially increasing future sales. Similarly, measuring the frequency of repeat Chipotle customers within a month might indicate customer loyalty, a key predictor of stable revenue streams.

Leading indicators act as an early warning mechanism, offering valuable insights to enhance decision-making and improve the optimization of use cases. They help organizations predict changes, minimize risks, and capitalize on opportunities, leading the business toward better performance and competitive advantage.

Optional: Use Case Template

We can also use the Use Case template (Figure 29) to capture additional use case information that might be useful in the use case prioritization process that is part of Step 4b.



Use Case Template

Use Case Description: **Increase Store Traffic via local events**: Identity profile-rich local events (e.g., Little League games, soccer tournaments, concerts, school events, marches, festivals) around which to drive local promotional activities and offers to increase store traffic

KPIs / Metrics: Average Revenue per Visit, Store Traffic, Revenue per Employee, Line Wait Time, % Abandonment, Mobile Orders, Positive Social Media Mentions, Table Turns, % locally-sourced ingredients, % waste, diversity ratio

Key Stakeholders
- Customers: get high-quality, healthy food quickly at fair price
- Store Manager: meet monthly & quarterly performance numbers
- Field Marketing: execute successful events that reflect creativity
- Corp Marketing: deliver campaign ROI, leads, and customer satisfaction
- Procurement: source local, high-quality ingredients at fair price
- Logistics: ensure ingredients reach stores on time in good condition

Desired Outcomes
- Increase store sales and profitability
- Increase new customer acquisition
- Increase customer satisfaction and social media mentions
- Minimize product / ingredient out-of-stocks
- Ensure full and satisfied staffing

Business Benefits
- Increase store traffic by marketing to participants at local (nearby) events
- Introduce Chipotle to new customers
- Local events are typically known in advance which helps from a staffing, procurement, and inventory perspective
- Increase in store traffic creates social media and awareness opportunities

Business Risks
- Weather could impact local events and staffing and inventory forecasts
- Unpredictable nature of responses could create long waits and jeopardize customer satisfaction and initial impressions of Chipotle
- Increased store traffic would put additional demands and pressure on store staff (cleaning store and restrooms, stocking utensils)
- Store management needs training on how to use analytics

Impact on Organizational Goals

Increase Overall Store Traffic	Increase Shopping Bag Revenue & Margin	Capture Net New Customers
Drive More Repeat Visits	Increase Supplier Reliability	Increase Customer Satisfaction

Privacy and Governance Considerations
- Ensure customer purchase and loyalty history fully protected during and after customer transaction
- Ensure customer credit card data fully protected during and after transaction

Ramifications of Failure
- Reduction in staff promotional opportunities
- Reduction in supporting local farmers
- Risk of store closure and employee layoffs
- No ability to support local events and activities

 Version 2.0

Figure 29: Use Case Design Canvas

Here's a brief description of each panel in the Use Case template:

- **Use Case Description**: This section provides a detailed narrative of the use case, explaining what the initiative is about and what it aims to achieve.
- **Key Performance Indicators (KPIs)**: Here, you list and define the KPIs/metrics that will be used to measure the use case's success, indicating how progress will be tracked against the stated objectives.
- **Key Stakeholders**: This section identifies key internal and external stakeholders impacted by the use case and what each wants to gain from it.
- **Desired Outcomes**: This section captures the use case's desired outcomes from the perspective of the different key stakeholders.
- **Business Benefits**: Summarizes the advantages and positive impacts expected from successfully executing the use case for the organization and its stakeholders.
- **Impact on Organizational Goals**: This section connects the dots between the use case and the organization's broader strategic objectives, showing how the initiative aligns with and supports the organization's overall goals.
- **Privacy and Governance Considerations**: This section addresses the necessary policies and regulations that must be adhered to, ensuring that the use case is compliant with data protection laws and governance standards.
- **Business Risks**: This section identifies the potential financial, operational, and reputational hazards that could jeopardize the use case's success.

- **Unintended Consequences Ramifications**: This section explores possible unexpected effects or secondary impacts that might arise from implementing the use case, especially those unintended consequences that result from its success.

Each panel captures critical aspects of the use case, from its business and operational objectives and expected benefits to potential risks and alignment with wider organizational goals, forming a strategic blueprint for the proposed business use case.

Linking Use Cases to Business Initiatives

Starting with a business initiative and then identifying its supporting use cases helps ensure that the identified use cases are related to something meaningful to the organization. Rather than creating a list of everyone's favorite use cases, we can identify a family of use cases valuable to the business stakeholders in supporting an organization's key business initiative (Figure 30).

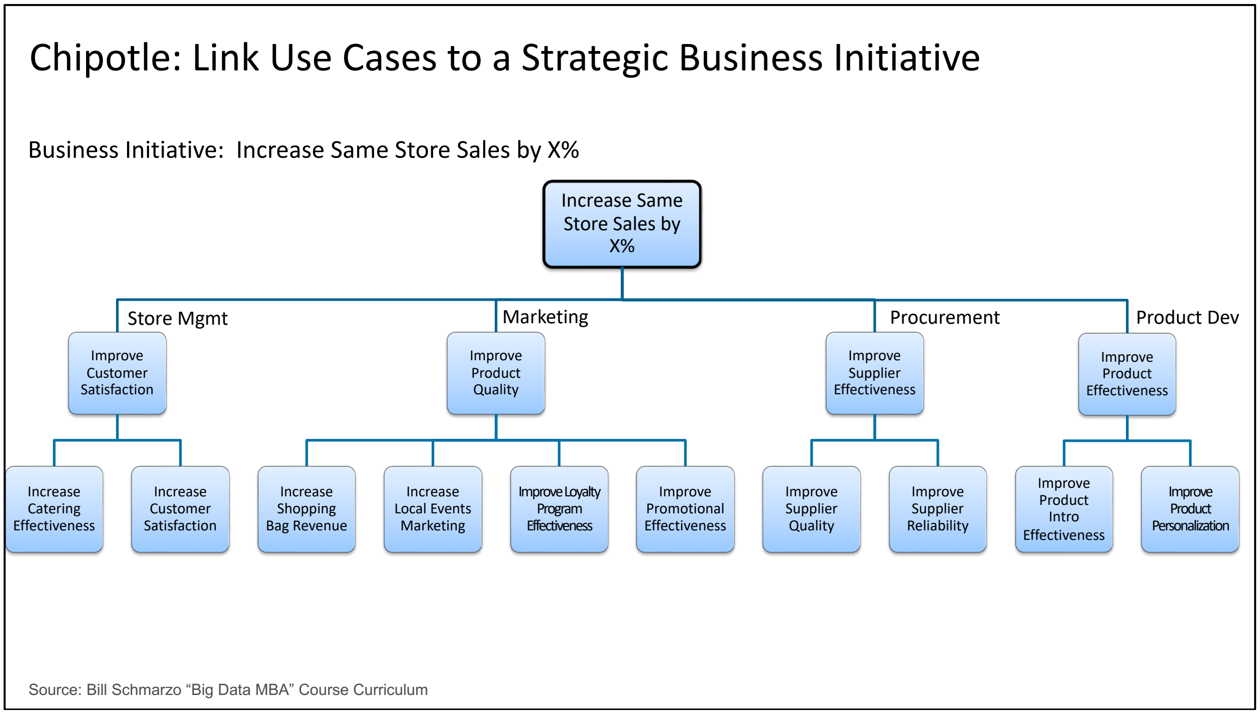


Figure 30: Linking Use Cases to Business Initiatives

Another benefit of the Business Initiative-to-Use Case decomposition approach is that the use cases that support the targeted business initiative likely share some of the same data. This allows the organization to activate the **Data Economic Multiplier Effect.**

> 💡 ***Data Economic Multiplier Effect*** *is the accumulation of attributable and quantifiable "**value**" from applying a data set against multiple use cases.*

This sets the stage for developing a roadmap of interrelated, high-value use cases that can accelerate time-to-value and reduce implementation risks because they share much of the same data.

Think small, but think small and strategic, not small and random.

Step 4a: "Identify & Assess Use Cases" Summary

Chapter 8 of the TLADS methodology, "Step 4a – Identify & Assess Use Cases", elaborates on the process of identifying potential use cases by utilizing the organizational and stakeholder information gathered in Steps 1 and 2. The step involves consolidating outcomes, decisions, and KPIs/metrics from these previous steps to identify main subject areas or use cases that align with the business initiative. The chapter details a structured approach to categorize these elements into similar use case clusters, assess their interrelationships, and explore additional metrics to support each use case. This systematic categorization helps to refine and prioritize data science efforts, ensuring they are strategically aligned with business goals and capable of enhancing operational performance. The emphasis is on creating a common vocabulary and shared understanding between business and data teams, fostering strategic focus and efficient resource allocation towards high-value use cases.

Chapter 9: TLADS Step 4b – Prioritize Use Cases

Step 4b, "Prioritize Use Cases", outlines an inclusive process to drive collaboration among key stakeholders in determining and prioritizing the use cases that will drive the organization's AI and data journey. To achieve this, we will use the Prioritization Matrix to align and reach consensus across stakeholders on which use cases to prioritize. We will consider the value of each use case in relation to its feasibility of implementation and how it supports our targeted business initiative.

Establishing a transparent and formal process to prioritize use cases and align data science efforts with organizational goals is crucial. However, it is equally important to collaborate with key stakeholders to ensure everyone has a voice in deciding where and how to allocate the organization's critical data and analytic resources. The success of data science (AI and Big Data) projects depends not solely on technology but also organizational alignment and agreement. Organizations must leverage their collective intelligence to define, develop, deploy, and manage use cases to achieve desired outcomes and improve decision-making effectiveness.

The Prioritization Process

The Prioritization Matrix is indispensable in driving cross-functional stakeholder alignment and consensus on the high-priority / high-feasibility use cases that should spearhead the organization's AI and data journey. This critical convergence sets the strategic direction and fosters organizational coherence and unity of purpose, ensuring that the precious data and analytic resources are channeled effectively toward shared goals.

Organizations can dramatically improve the project's success using the Prioritization Matrix to identify the "Sweet Spot" use cases that balance high-value potential with practical execution. Sometimes called the "Goldilocks Zone," this optimal intersection is where actionable insights and strategic investments coalesce, moving the organization towards transformative outcomes that are attainable, impactful, and aligned with the organization’s overarching vision and mission (Figure 31).

Prioritization Matrix Process

- Use cases are placed on the Prioritization Matrix vis-à-vis other use cases based on Value and Implementation Feasibility.
- The heart of the Prioritization process is the discussion and agreement between stakeholders about the relative placement of each use case.
- Beware the Prioritization Traps!
  - "Zone of Mismanaged Expectations," those use cases with significant business value but little chance of successful execution (Use Case [A])
  - "Zone of User Disillusionment," those use cases that are easy to execute but provide little business value (Use Case [C])
  - "Zone of Career Limiting Moves," those use cases of little value and low feasibility

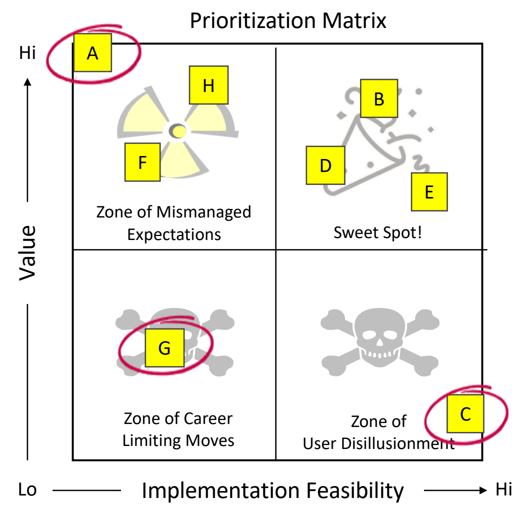


Figure 31: Prioritization Matrix Process

However, as you navigate the Prioritization Matrix, beware of these prioritization traps:

- The "**Zone of Mismanaged Expectations"** (High Value/Low Feasibility) refers to those use cases that hold significant business value but have little chance of successful execution (as described in Use Case [A] in Figure 31). The challenge with these use cases is that, although an executive might own that use case, the organization may not have the necessary data, analytics, and people capabilities. Additionally, senior management support for a project that may take several years to realize the fruits of the AI and data effort may not be feasible.
- The "**Zone of User Disillusionment**" (Low Value/High Feasibility) refers to use cases that are easy to execute but provide little business value (Use Case [C] in Figure 31). The challenge with these use cases is that while they may be relatively easy for the AI and data team to deliver, no one in the business cares as they provide little direct, measurable value to them (e.g., data center migrations, moving to the cloud).
- The "**Zone of Career Limiting Moves**" (Low Value/Low Feasibility) is the zone of Career-Limiting Moves (CLM), where one settles on a use case that has little value and is hard to implement. Some senior executives' pet projects can sway organizations into starting in this dangerous zone.

Understanding, being aware, and opening the organization's eyes to the characteristics of Prioritization Matrix traps to navigate around them is crucial. Remember, forewarned is forearmed!

Template #5: Prioritize Use Cases

To begin the prioritization process, arrange the identified use cases on the Prioritization Matrix relative to other use cases based on value versus implementation feasibility. The value factors should include financial, customer, employee, partner, product, operational, environmental, societal, and ethical factors. The implementation feasibility factors could include data accessibility, data quality, data governance, analytic skills maturity, culture, and management support (Figure 32).

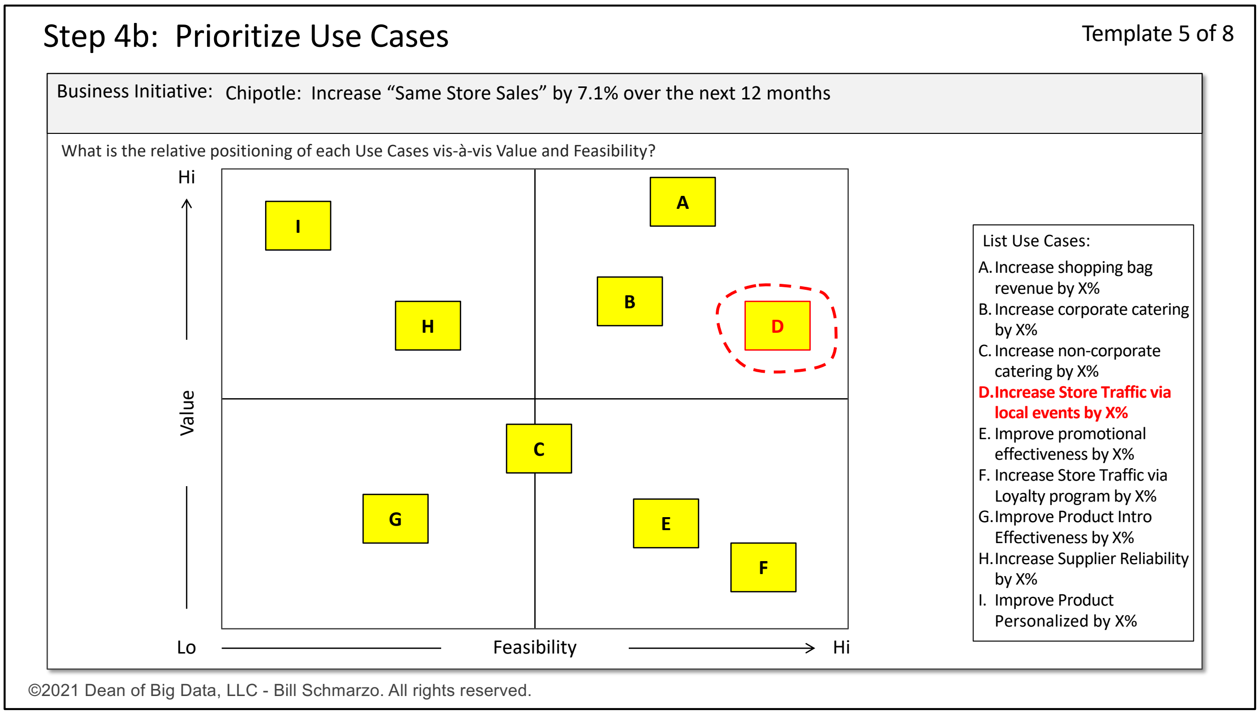


Figure 32: "Prioritize Use Cases" Canvas

It is important to note that this is just the starting point. The heart of the prioritization process involves discussing and debating the relative placement of each use case compared to the others. For instance, you may debate "Why is use case [B] more valuable than use case [E]?" or "Why is use case [E] more feasible than use case [B]?".

The Prioritize Use Case canvas is a tool used to capture the agreed-upon final version of the Prioritization Matrix by all stakeholders. This matrix will serve as the foundation for creating a roadmap of use cases that are highly relevant in supporting our targeted business initiative.

Note: The incremental implementation approach can manage the gradual development of each use case's value and feasibility. For instance, the use cases that fall into the "Zone of Mismanaged Expectations" category may become more feasible as the data and analytics architecture and capabilities are built from the deployment of the initial use cases. Similarly, the use cases in the "Zone of User Disillusionment" may become more valuable as they build upon the value created from the initial use cases. Therefore, it's necessary to revisit the prioritization process after implementing the first few use cases.

Use Case Value and Feasibility Assessment Templates

We can leverage additional design templates to help us better understand and quantify the Prioritization Matrix value and feasibility dimensions. The first Prioritization Matrix aid is the Template 4.1 "Use Case-to-Value Drivers Assessment" canvas (Figure 33).

Template 4.1: [Chipotle] Use Case-to-Value Drivers Assessment

	Use Cases	Value Drivers (Desired Outcomes)						
		Increase Overall Store Traffic	Increase Shopping Bag Revenues	Acquire Net New Customers	Increase Customer Repeat Visits	Increase Supplier Reliability	Increase Customer Satisfaction	Overall Score
A	Increase Shopping Bag revenue and margin	◕	○	○	●	●	◕	2.50
B	Increase corporate catering	◑	◑	◕	◔	◕	◕	2.50
C	Increase non-corporate catering	◕	○	○	◕	◕	◕	2.00
D	Increase Store Traffic via local events	◔	◕	◕	○	◔	◑	1.75
E	Increase Store Traffic via Loyalty program	◔	◔	◑	○	◑	◑	1.25
F	Improve Product Intro Effectiveness	●	○	○	◕	●	●	2.50
G	Improve Supplier Reliability	◑	○	◔	◑	◕	◑	1.50
H	Improve Product Personalized	◔	○	○	◑	○	◔	0.50

○ Low ● High | 0 Low 4 High

Figure 33: "Use Case-to-Value Drivers Assessment" Canvas

The "Use Case-to-Value Drivers Assessment" canvas is a tool that evaluates the potential of each use case and its desired outcomes to support the targeted business initiative. The canvas rates each use case on a scale of 0 to 4 and calculates an "Overall Score" based on the weighted sum of its value drivers. To provide a quick visual assessment of the most important use cases vis-à-vis the value drivers, I like to use Harvey Balls to visually display the rating of each use case's value potential.

The second Prioritization Matrix aid is the Template 4.2 "Use Case-to-Data Source Mapping" canvas (Figure 34).

Template 4.2: [Chipotle] Use Case-to-Data Source Mapping

Data Mapping

	Use Cases	Point of Sales	Market Baskets	Store Demo	Local Competition	Store Manager Demo	Consumer Comments	Social Media	Local Events	Weather	Vehicle Traffic
A	Increase Shopping Bag revenue and margin	●	●	◑	◕	◑	◑	◕	◑	●	◕
B	Increase corporate catering	◑	○	●	●	◕	◑	◔	○	◕	◑
C	Increase non-corporate catering	◑	○	●	●	●	◕	●	○	◑	◔
D	Increase Store Traffic via local events	●	●	◑	◕	◑	◔	◕	●	●	◔
E	Increase Store Traffic via Loyalty program	●	●	◕	◕	◔	◔	◕	○	◔	◔
F	Improve Product Intro Effectiveness	●	●	◑	◕	◔	◔	●	○	◔	◔
G	Improve Supplier Reliability	◑	◑	◑	○	◑	○	○	◑	◔	○
H	Improve Product Personalized	●	●	◑	◔	◔	◕	◕	○	○	○

○ Low ● High | 0 Low 4 High

Figure 34: "Use Case-to-Data Source Mapping" Canvas

The "Use Case-to-Data Source Mapping" canvas systematically evaluates and ranks the pertinence of various data sources, on a scale from 0 to 4, in relation to their utility and criticality for each identified use case. This mapping provides a barometer for assessing the data issues that might be integral to assessing the feasibility of the use cases.

The "Use Case-to-Implementation Feasibility Assessment" canvas (Template 4-3) is a tool that helps evaluate the feasibility of a use case based on several critical factors. These factors include data availability, analytic complexity, analytic skill set requirements, and organizational preparedness. By assessing these factors, the canvas provides a quantified evaluation, on a scale from 0 to 4, of the potential implementation risks associated with the use case. This can help identify potential challenges and enable a risk management strategy tailored to the unique demands of the use case (Figure 35).

Template 4.3: [Chipotle] Use Case-to-Implementation Feasibility Assessment

		Implementation Risks						
	Use Cases	Business Process Integration	Organization Readiness	Analytic Skills	Data Quality & Accessibility	Data Governance	Scalable Infrastructure	Overall
A	Increase Shopping Bag revenue and margin	◕	◕	◑	◕	◔	◔	2.00
B	Increase corporate catering	◑	◑	◕	◑	◑	◑	2.00
C	Increase non-corporate catering	◕	◑	◕	◕	◕	◔	2.25
D	Increase Store Traffic via local events	◑	◔	◔	◑	◑	◔	1.50
E	Increase Store Traffic via Loyalty program	◑	◔	◑	◔	◔	◔	1.50
F	Improve Product Intro Effectiveness	◑	◔	◕	◕	◑	◑	2.50
G	Improve Supplier Reliability	◕	◑	◕	◕	◑	◑	2.50
H	Improve Product Personalized	◔	◔	◔	◔	○	○	1.00

○ Low ● High | 0 Low 4 High

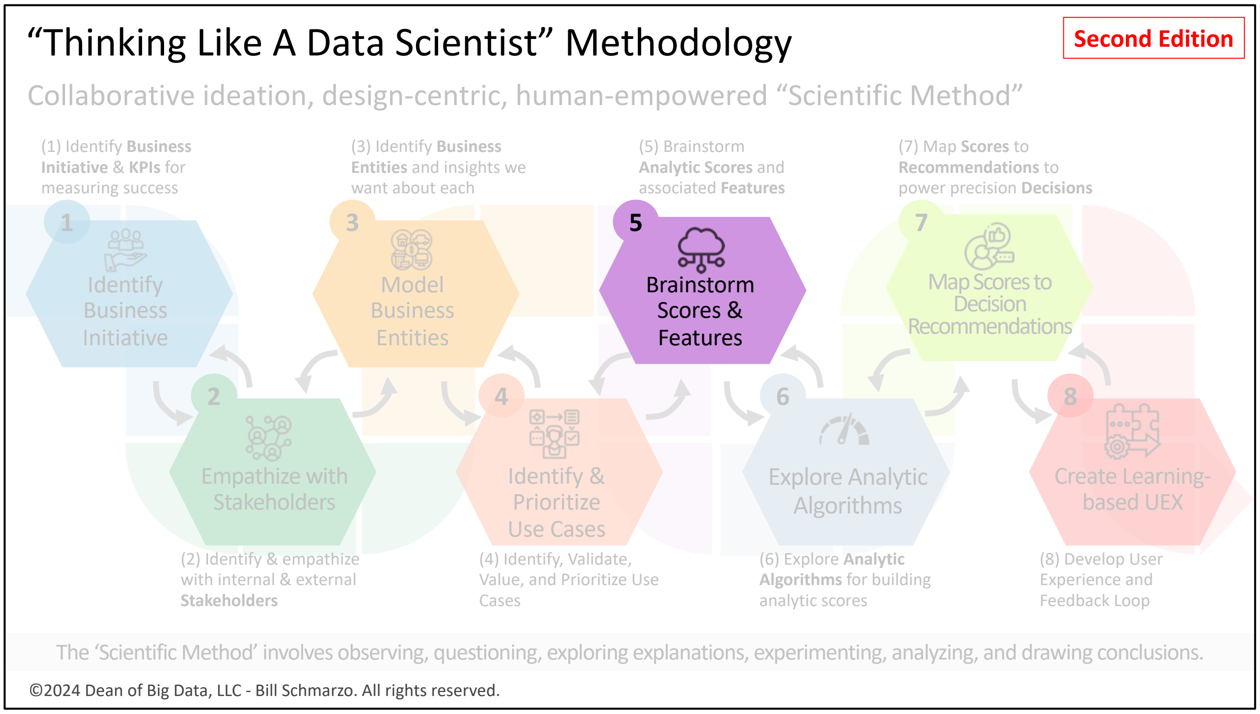
Figure 35: "Use Case-to-Implementation Feasibility Assessment" Canvas

Templates 4.1, 4.2, and 4.3 are useful tools that can assist you in prioritizing your use cases. These templates allow you to assess the value drivers, map the data sources, and determine the feasibility of implementation for each use case. However, validating and refining these factors by conducting user interviews and workshops before moving forward with the TLADS methodology is important. By taking a scientific approach, you can make more informed decisions about aligning and allocating your business and data science resources to support your targeted business initiative.

Step 4b: "Prioritize Use Cases" Summary

Chapter 9 of the TLADS methodology, "Step 4b – Prioritize Use Cases," outlines a structured process for ranking use cases by utilizing the Prioritization Matrix, which assesses the value of each use case against its implementation feasibility. This step is crucial for aligning cross-functional teams and ensuring that data science efforts focus on use cases that offer the greatest strategic impact while practically feasible. The chapter emphasizes the need for transparency and broad stakeholder involvement in the prioritization process to build consensus and support for the selected use cases. This process helps set the direction for the organization's AI and data journey and fosters organizational unity by focusing resources on shared goals.

Chapter 10: TLADS Step 5 – Brainstorm Scores & Features



TLADS Step 5 is called "Brainstorm Scores & Features." This step requires collaboration between data scientists and business stakeholders to identify and engineer the analytical scores and machine learning (ML) features necessary to generate recommendations for the prioritized use case’s critical decisions.

The foundation of Step 5 is the concept of Analytic Scores.

> 💡 ***Analytic Scores*** *are algorithmically derived numerical estimations that quantify the probability of specific outcomes or behaviors. Typically, these scores are standardized on a scale from 0 to 100, signifying a spectrum where 0 represents a minimal chance of occurrence, and 100 indicates a near certainty.*

Developing precise and pertinent analytic scores demands a rigorous process of selecting and curating the most indicative ML features. This ensures that each score is built upon a robust comprehensive and contextually appropriate data foundation.

> 💡 ***Machine Learning (ML) Features** are variables processed by algorithms during the training and operational stages of analytic model development. They encompass quantitative and qualitative data, offering nuanced predictive insights into prospective behaviors or actions. These features encapsulate a wide range of data characteristics, from linear to categorical, each contributing unique, valuable information for predictive model accuracy and efficacy.*

Categorical features, which classify data into distinct categories such as 'yes' or 'no,' 'type A' or 'type B,' are often encoded or transformed to facilitate their use in mathematical models. This enables the nuanced interpretation of data that isn't inherently numerical but is critical for accurate predictions.

Let's further explore the concept of Analytic Scores and ML Features.

Analytic Scores

Analytic scores guide stakeholders' key decisions and actions across various business domains. These scores measure and evaluate different aspects of a business, such as operational performance, customer satisfaction, and financial stability. By analyzing and interpreting these scores, decision-makers can make informed choices and devise strategies to improve the business's overall performance. Table 2 presents sample analytic scores used in different industries to assess and optimize performance.

Financial	**Credit Cards**	**Manufacturing**	**Casino**
Credit (FICO) Retirement Readiness Investment Risk	Attrition Risk Fraud Risk Product Preferences	Maintenance Risk Supplier Reliability Supplier Quality	Customer LTV Lifestage Influence Gaming Preferences
Education	**Healthcare**	**Utilities**	**Pro Sports**
Graduation Readiness Cohorts Influence	Wellness Condition Stress Risk	Energy Efficiency Conservation Effectiveness	Fatigue Factor Motivation Factor

Table 2: Analytic Scores by Business Domain

One of the main goals of TLADS Step 5 is to encourage cross-organizational brainstorming and ideation to identify potential analytic scores that can help optimize our prioritized use case. For example, let's take our prioritized Chipotle use case - Improving the Effectiveness of Local Events Marketing - and brainstorm potential analytic scores to help us reach our objectives.

- **Local Event Fit Score:** This score evaluates an event's alignment with Chipotle's brand values, target demographic engagement, and potential for sales uplift. It factors in historical data on customer demographics, past event successes, and regional food preferences to determine the event's compatibility with Chipotle's marketing objectives.
- **Local Event Impact Score:** Beyond assessing the immediate effects on brand metrics, this score would also forecast the long-term value of customer relationships initiated or strengthened at the event. It would incorporate advanced sentiment analysis from post-

event surveys and online chatter to gauge deeper brand perception changes, integrating these with direct measures like lead conversion rates and incremental sales data.

- **Local Event Promotion Effectiveness Score:** This predictive metric would track customer footfall and retention post-event and model the nuanced customer journey that leads from event awareness to purchase. Using data from various touchpoints, including special offer redemption rates and subsequent visit frequency, would provide a multi-faceted view of promotional campaigns' ROI.
- **Social Media Engagement Score:** This score would aggregate interactions across platforms, weighting them by potential reach and influence to provide a holistic view of engagement levels. It would also consider the virality of content, audience growth, and engagement trends over time, offering insights into the ebb and flow of customer attention in relation to specific campaigns or general brand activity.
- **Social Media Propensity Score:** Designed to pinpoint where to allocate resources for maximum effect, this score would measure raw activity and engagement levels and the propensity of users on each platform to engage with Chipotle's content. It would use predictive modeling to assess how different platforms contribute to the progression of the marketing funnel, from awareness to advocacy.

Next, let's understand the role of ML features in building more effective, relevant, and accurate analytic scores.

Machine Learning (ML) Features

Machine Learning (ML) features are crucial data points and characteristics extracted and generated from raw data. These features inform the predictive capabilities of our analytic scores and can range from simple numerical inputs to complex text and image data. They are transformed into a format that algorithms can understand. Examples of features include consumer behavior metrics, consumer engagement levels, consumer purchase history, and customer shopping images used to improve model accuracy and the predictive strength of the scores generated.

📝 **Note:** In the supplement workbook, we will explore how to use GenAI to analyze large datasets and discover hidden patterns to create new ML features that can enhance the effectiveness of the analytic scores.

A **Decision Journey Map** is a design tool that can be used to brainstorm the features that the data science team can investigate to generate more relevant and meaningful analytic scores. The Decision Journey Map enables us to explore the stakeholders' key decisions to determine which variables or factors could impact their decision-making process.

For example, Figure 36 represents a Decision Journey Map for our top priority Chipotle use case: "Increase Store Traffic via local Events Marketing" from the perspective of the Store Manager stakeholder.

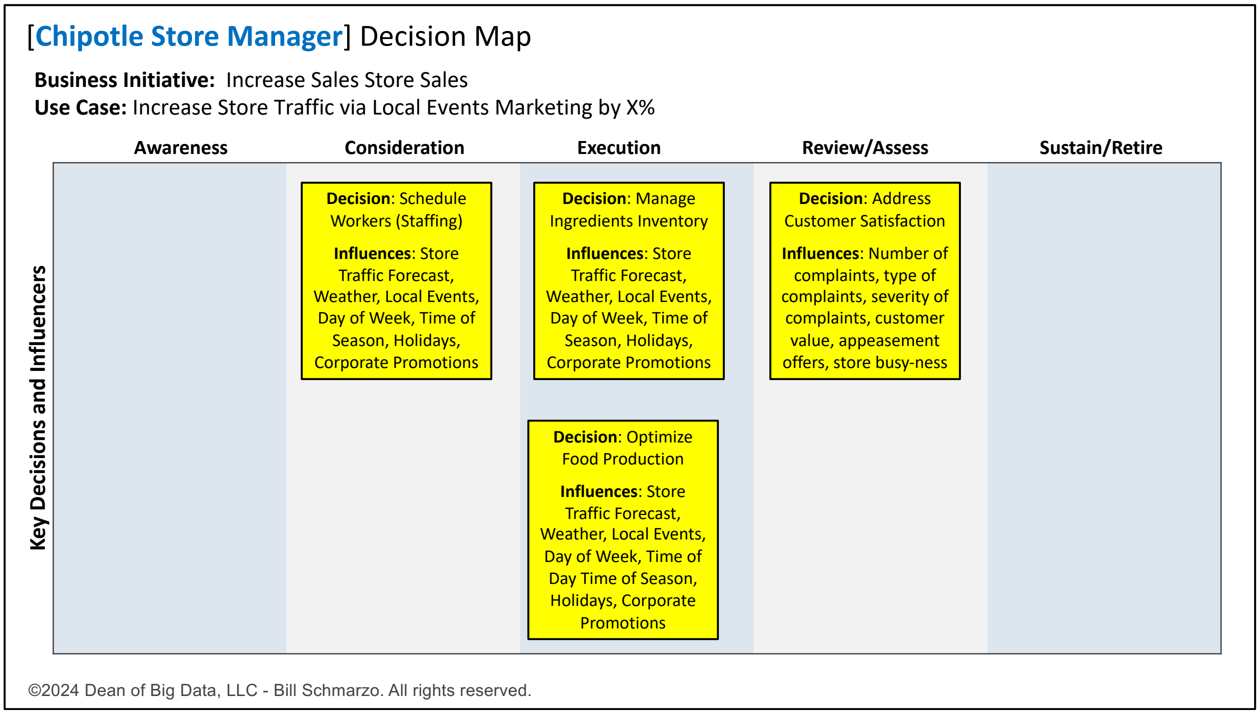


Figure 36: Stakeholder Decision Journey Map

Continuing our Store Manager Decision Journey Map exercise, we can identify various ML features that can support our potential analytic scores, including:

- **Local Event Fit** score potential features could include location, date, duration, event type, event theme, historical attendance, demographic match, past sponsorship outcomes, local competitor activity, event goal, and event cost.
- **Local Event Impact** score potential features could include the estimated number of attendees, interactions, social media metrics (impressions, leads, conversions, referrals, feedback, reviews), conversion rates, customer feedback metrics, brand sentiment metrics, and media coverage.
- **Local Event Promotion Effectiveness** score potential features could include promotion type, promotion value, frequency, eligibility terms, satisfaction, loyalty, redemption rates, foot traffic, sales lift, and cost per acquisition.
- **Social Media Engagement** score potential features could include interaction and attention Chipotle receives on its social media platforms. Potential features include the number of followers, likes, comments, shares, mentions, engagement rate, content reach, follower growth, and click-through rates.
- **Social Media Propensity** score potential features could include the number of active users, content type, level of interaction, advertising, platform activity levels, user interaction patterns, content virality, user demographics, and interaction length.

For our local marketing effectiveness use case, identifying, exploring, and validating the predictive nature of these potential ML features can help us create better analytic scores that provide valuable insights into the effectiveness of local event marketing campaigns and overall audience engagement on social media.

Template #6: "Brainstorm Scores & Features" Canvas

We will use the "Brainstorm Scores & Features" canvas in Figure 37 to capture details about the analytic scores and supporting ML features that support our prioritized "Increase Store Traffic via local Events Marketing" use case.

Step 5: Brainstorm Scores & Features

Template 6 of 9

Business Initiative: Chipotle: Increase "Same Store Sales" by 7.1% over the next 12 months

Prioritized Use Case: Use Case: Increase Store Traffic via Local Events Marketing by X%
Key Decisions: Local Events, Promotions, Sponsorships, Social Media

Key Decisions	Stakeholders	Potential Analytic Scores	Potential ML Features
Decide in which local events to participate	•Customers •Employees •Local Events Mgmt	**Local Event Fit** score measures how suitable and effective a local event is for Chipotle to sponsor.	location, date, duration, event type, event theme, historical attendance, demo match, event cost
		Local Event Impact score measures the positive impact of a local event on brand awareness.	# of attendees,, social media metrics, conversion rates, customer feedback metrics, brand sentiment metrics, and media coverage
Decide what type of promotion for local event	•Customers •Employees	**Local Event Promotion Effectiveness** score measures how effectively a promotion attracts and retains customers.	promotion type, promotion value, frequency, eligibility terms, satisfaction, loyalty, redemption rates, sales lift, cost per acquisition
		Promotion Optimization score measures how optimal a promotion is for maximizing revenue and profit.	Promotion cost, benefit, margin, return on investment, and break-even point
Decide how much to spend on sponsorship	•Customers •Local Events Mgmt •Competitors	**Sponsorship Level** score measures how appropriate and beneficial a sponsorship level is for Chipotle's visibility and reputation.	Sponsorship type, category, tier, package, logo placement, media exposure, recognition.
		Sponsorship Budget score measures how affordable and reasonable a sponsorship budget is for Chipotle's financial resources and objectives.	Sponsorship cost, benefit, margin, return on investment, break-even point.
Decide how much time and money to invest in social media	•Customers •Local Events Mgmt •Competitors •Social Media Marketing	**Social Media Engagement** score measures how much interaction and attention Chipotle receives on its social media platforms.	# of followers, likes, comments, shares, mentions, engagement rate, content reach, follower growth, and click-through rates.
		Social Media Conversion score measures how many customers Chipotle acquires or retains through its social media platforms.	Number of clicks, leads, conversions, referrals, reviews, and ratings.
		Social Media Propensity score measures level of activity and engagement on each specific social media platform	# active users, content type, level of interaction, advertising, content virality, user demographics, length of Interaction

Version 2.0

Figure 37: "Brainstorm Scores & Features" Canvas

For our prioritized "Increase Store Traffic via local Events Marketing" use case, we want to capture the following information:

- **Key Decisions**: This is a list of the critical decisions our key stakeholders need to make to support our prioritized use case (captured in Template #3: "Empathize with Stakeholders" canvas). For our prioritized use case, these decisions would involve determining which events to sponsor, the sponsorship level, and what type of promotions to pair with the sponsorship.
- **Stakeholders**: A list of the key internal and external stakeholders impacted by the prioritized use case (also captured in Template #3: "Empathize with Stakeholders" canvas). For our use case, that would include Customers, Employees, Local Events Management, and Social Media Marketing.
- **Potential Analytic Scores**: The potential analytic scores that might need to be created to support our key decisions. For example, those analytic scores could include Local Event Fitness, Local Event Impact, Local Event Promotion Effectiveness, and Social Media Engagement.

- **Potential ML Features:** The potential ML features (variables and metrics) needed to construct our analytic score. For instance, the analytic score for Local Event Promotion Effectiveness may include features such as promotion type, promotion value, frequency, eligibility terms, satisfaction, loyalty, redemption rates, sales lift, and cost per acquisition."

📝 **Note:** The data science team needs to work closely with business stakeholders and subject matter experts to develop an effective and relevant analytic score. Together, they will assess and refine the analytic scores and explore additional ML features. This iterative process involves experimenting with different ML feature weights and combinations to generate an analytic score that effectively balances cost-effectiveness and strategic significance. Working together as a team ensures that the final score is robust and tailored to meet the critical stakeholder decisions for the prioritized use case.

Preparing for Step 6: "Exploring Analytic Algorithms"

Step 6, "Exploring Analytic Algorithms," may initially appear daunting to those without a background in data science, but understanding this step is crucial for two key reasons:

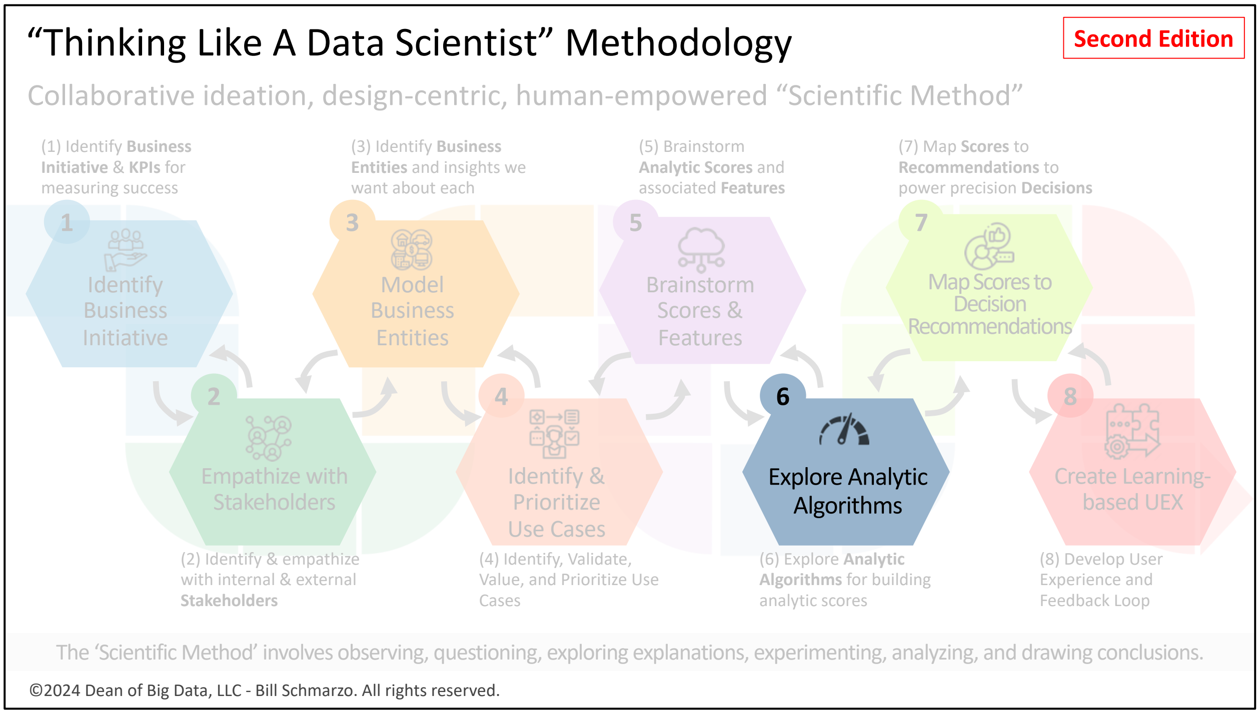
1. **Enhanced Collaboration**: Gaining a foundational understanding of the various analytic algorithms allows non-data scientists to contribute more effectively to the ML features brainstorming and exploration process. This knowledge helps them understand how certain features can enhance the predictive power and accuracy of analytic scores, fostering a more productive dialogue between them and the data science team.
2. **Leveraging Generative AI (GenAI)**: GenAI is set to revolutionize how we select and utilize analytic algorithms and ML features. By automating complex tasks, GenAI makes it possible for individuals without extensive data science expertise to make informed decisions and achieve desired outcomes more efficiently. For instance, current GenAI tools can generate pseudo-codes for constructing scores, and future advancements will these tools to develop analytic scores with minimal coding requirements. I will delve deeper into these developments in the forthcoming GenAI book supplement.

By understanding the mechanics of Step 6, all stakeholders, regardless of their technical proficiency, can contribute more meaningfully to the development of analytic-driven business, operational, and societal solutions, ensuring that they are both innovative and aligned with business and social objectives.

Step 5: "Brainstorm Scores & Features" Summary

Chapter 10, "TLADS Step 5 – Brainstorm Scores & Features," focused on the collaborative effort between data scientists and business stakeholders to identify and engineer the analytical scores and ML features needed to generate more relevant and meaningful recommendations. This process is fundamental for enhancing decision-making related to prioritized use cases. This step involves deep dives into the data to uncover insights leading to more informed, effective business decisions and strategies.

Chapter 11: TLADS Step 6 – Explore Analytic Algorithms



In the sixth step of the TLADS process, our focus is on developing the analytic scores that support the optimization of our prioritized use case. To achieve this, we use the data science collaboration development process, which drives cross-functional collaboration, gains insights, and ensures the analytic scores are accurate and robust. This collaborative approach makes sure that the analytic scores are responsive to diverse perspectives and leverage the collective expertise of the entire organization.

> 💡 ***Data science** is about identifying those variables and metrics that might be better predictors of performance.*

This step involves testing and refining algorithms to ensure that they closely align with the requirements of the use case and the objectives of the business. We use the data science collaboration development process to refine and validate the selection of algorithms. By employing a combination of traditional and advanced techniques, this process guarantees that the analytic scores are both accurate and highly relevant. This, in turn, supports informed and strategic decision-making.

> 💡 ***Analytic algorithms** are data science techniques that can extract insights and patterns from data, such as machine learning, deep learning, natural language processing, or computer vision.*

The importance of this step extends beyond algorithm selection; it is about ensuring that these scores are valid, reliable, and precise enough to effectively inform and influence important strategic decisions. It serves as a bridge between raw data and actionable insights, embodying the transition from data processing to extracting business value. In the broader context of the data science development process, this step ensures that the analytics framework is not only aligned with the organization's strategic goals but also scalable and adaptable to evolving business needs. Proper execution of this step can significantly enhance the decision-making process, improving operational efficiency, customer satisfaction, and competitive advantage.

To become a "Citizen of Data Science," it is essential to have a fundamental understanding of the various types of algorithms utilized to generate analytical scores. This knowledge enables individuals to confidently participate in defining, designing, developing, implementing, and continuously improving the analytical scores.

High-level Analytic Algorithms Overview

Table 3 provides a summarized view of the different analytic algorithm classes for those not deeply versed in data science. It outlines various classes of algorithms, provides tangible examples of each, and associates them with real-world applications, illustrating the breadth and impact of these computational tools.

Algorithm Class	Algorithm Examples	Example Use Cases
Supervised Machine Learning	- Linear Regression - Logistic Regression - Decision Trees - Random Forest - Support Vector Machines (SVM) - K-Nearest Neighbor (KNN) - Naive Bayes	- Predicting house prices based on features like square footage and location. - Customer churn prediction in subscription services. - Credit risk assessment for loan approvals. - Customer segmentation for targeted marketing. - Image classification and fraud detection. - Email spam filtering.
Unsupervised Machine Learning	- K-Means Clustering - Hierarchical Clustering - Principal Component Analysis (PCA)	- Market segmentation based on customer behavior. - Identifying similar products in e-commerce. - Dimensionality reduction for feature extraction.
Deep Learning	- Neural Networks - Convolutional Neural Networks (CNN) - Recurrent Neural Networks (RNN)	- Natural language processing (NLP) tasks like sentiment analysis. - Image recognition and object detection. - Time series forecasting and sequence modeling.
Generative AI	- Generative Adversarial Networks (GANs) - Variational Autoencoders (VAEs)	- Creating realistic images or generating synthetic data. - Anomaly detection and data augmentation.
Reinforcement Learning	- Q-Learning - Deep Q Networks (DQNs) - Policy Gradient Methods	- Optimizing supply chain logistics. - Game playing (e.g., chess, Go). - Personalized recommendation systems.

Table 3: High-level Analytic Algorithms Overview

Each algorithm serves a different purpose, like the tools in your toolbox. The choice of algorithm depends on the specific business problem you are trying to solve. Understanding the strengths and limitations of each algorithm will guide effective decision-making in designing and developing your analytic models and analytic scores.

Appendix A provides a more detailed breakdown of the different analytic algorithms and the use cases against which each is most effective.

Template #7: "Explore Analytic Algorithm" Canvas

For "Citizens of Data Science," understanding the "Explore Analytic Algorithms" step is important for effectively collaborating with the data science team. Understanding how different algorithms like regression or cluster analysis can unveil customer patterns and behaviors enables non-data scientists to contribute to developing relevant and meaningful analytic models. Their involvement ensures that the analytical models built are technically robust and tightly aligned with the organization's strategic objectives, fostering a collaborative environment where everyone has a clearly defined role in the data science journey.

For the prioritized use case, the data science team will complete the "Explore Analytic Algorithms" canvas (Figure 38).

Step 6: Explore Analytic Algorithms

Template 7 of 9

Business Initiative: Chipotle: Increase "Same Store Sales" by 7.1% over the next 12 months

Prioritized Use Case: Use Case: Increase Store Traffic via Local Events Marketing by X%
Key Decisions: Local Events, Promotions, Sponsorships, Social Media

Analytic Score	Score Explanation	Analytics Algorithm	Algorithm Rationale
Local events promotion effectiveness score	Measures of how well a local event promotion attracts new customers	Regression analysis	identify the most influential factors that affect the customer response and quantify their impact on the local event promotion effectiveness score
		Cluster analysis	identify the different customer groups that respond differently to the local event promotion and tailor the marketing strategies accordingly
		Sentiment analysis	measure the customer satisfaction and loyalty related to the local event promotion and identify the strengths and weaknesses of the marketing campaign
Promotion effectiveness score	Measures how effective a promotion is for maximizing revenue and profit	Regression analysis	Estimate the impact of each independent variable on the promotion optimization score and identify the most significant and influential factors
		Cluster analysis	segment customers or events based on their similarities and differences and assign them different promotion optimization scores
		Decision tree analysis	Evaluate the trade-offs and risks involved in choosing a promotion type or an event type for a given customer segment or event scenario

Version 2.0

Figure 38: "Explore Analytic Algorithms" Canvas

To complete the "Step 6: Explore Analytic Algorithm" canvas, we will follow a structured approach that aligns use case business and operational goals with the corresponding algorithms and the rationale for selecting that particular algorithm.

Here's how to approach each section of the canvas, with examples specific to the "Increase Store Traffic via Local Events Marketing" use case for Chipotle:

- **Analytic Score** defines the analytic scores relevant to our prioritized use case. For example, the Local Event Promotion Effectiveness Score would measure the increase in in-store traffic attributable to local event promotions.
- **Score Explanation** describes what each analytic score measures and what insights it seeks to provide to optimize the prioritized use case. For example, the Local Event Promotion Effectiveness score quantifies the success of event promotions by measuring the uptick in customer visits and transactions during and after local events.
- **Analytics Algorithm** lists the specific algorithms that will be used to build the AI models that power the analytic score. The Local Event Promotion Effectiveness score might use the following algorithms:
  - Regression Analysis to model and predict store traffic increases based on variables such as event size, type, and promotion spend.
  - Cluster Analysis to segment local events based on characteristics that correlate with increased store visits.
- **Algorithm Rationale** captures why the algorithm is best suited for developing the analytic score. For example, the rationale for the analytic algorithms that will be used to create the Local Event Promotion Effectiveness Score might include the following:
  - The regression algorithm can identify which event characteristics (e.g., duration, type, proximity to the store) correlate strongly with store traffic increases.
  - The cluster algorithm rationale can categorize events into clusters based on similar features and past performance, identifying which clusters are most likely to lead to increased store visits.

The rationale should reflect how the algorithm's output will improve decision-making for our prioritized use case.

Why Understanding Analytic Algorithms is Important for Non-Data Scientists

Understanding the analytic algorithms used to construct AI models that underpin the analytic scores is useful for non-data scientists for several reasons:

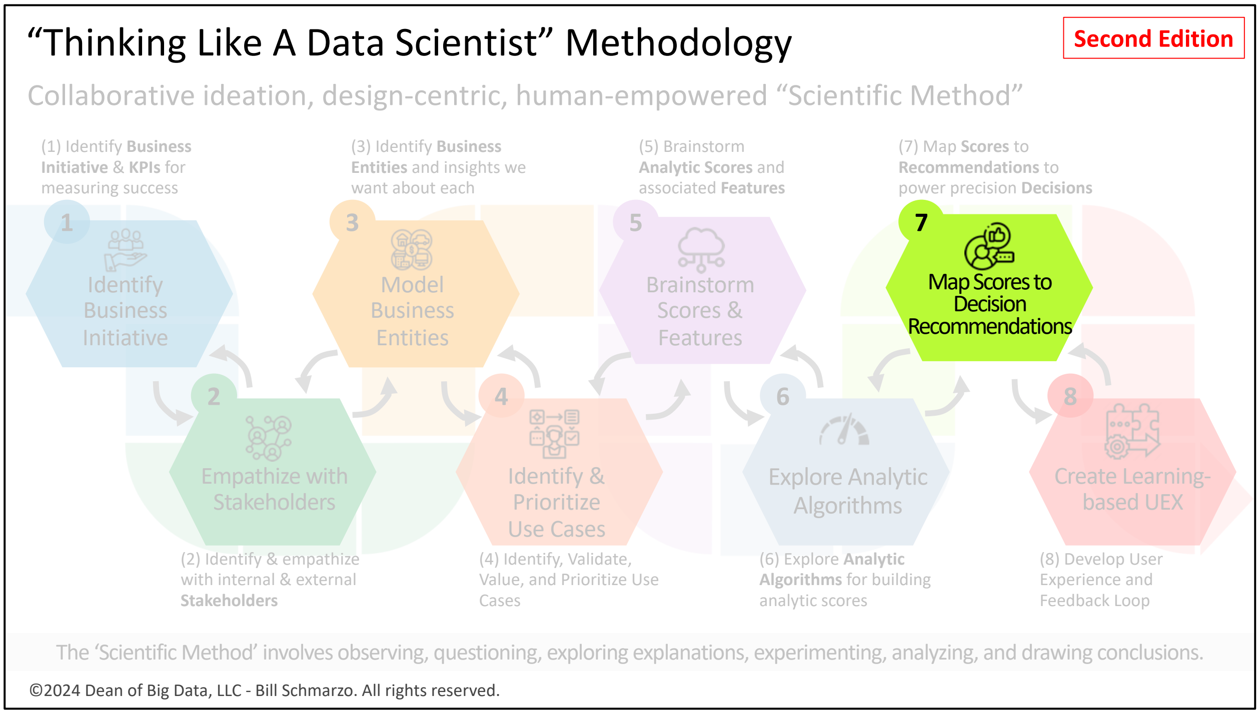
- Firstly, fostering a broad understanding of analytic algorithms equips non-data scientists with the ability to engage in strategic discussions about use case development. It allows them to appreciate different algorithms' strengths, limitations, and applications.
- Secondly, a well-informed team can provide valuable context and insights that could lead to more effective feature selection and algorithm tuning. When non-data scientists grasp the fundamental mechanics of how algorithms process and analyze data, they can contribute to refining the input (ML features) based on their domain expertise.
- Finally, when non-data scientists understand the rationale behind selecting certain algorithms, they are better positioned to anticipate the model's potential impact and interpret the outcomes. This becomes particularly relevant when communicating findings to stakeholders or making decisions based on the model's outputs.

As AI increasingly drives the business landscape, organizations must cultivate a data and analytics-knowledgeable culture where every team member, regardless of technical background, feels comfortable questioning, suggesting, and innovating. Step 6 is not merely a technical step in a larger methodology; it's a crucial paradigm shift in empowering the entire organization to think critically about the data, AI models, and analytic scores that are becoming the backbone of modern business decision-making.

Step 6: “Explore Analytics Algorithms” Summary

TLADS Step 6 – “Explore Analytic Algorithms" – details the selection and application of the relevant analytic algorithms essential for building the analytic scores that support critical decisions and desired outcomes. It emphasizes the need for a basic understanding of various data science techniques like machine learning, deep learning, and natural language processing, which are instrumental in extracting insights from data. The chapter introduces algorithm classes and examples and discusses the importance of choosing the right algorithm based on the specific business problem to develop effective and robust analytic models.

Chapter 12: TLADS Step 7 – Map Scores to Decision Recommendations



Step 7 is the pivotal moment in the TLADS methodology, where we leverage analytic scores (defined and developed in Steps 5 and 6) to generate prescriptive recommendations (identified in Step 3). This facilitates informed decision-making for stakeholders (highlighted in Step 2) and thus enhances the effectiveness of the prioritized use case (outlined in Step 4) that supports our targeted business initiative (assessed in Step 1).

To effectively map analytics to actions in Step 7, it is crucial to understand several key steps:

- **Identify** the key decisions the stakeholders must address to optimize the use case, pinpointing where analytics can have the greatest impact.
- **Clarify** the stakeholders' desired outcomes from these key decisions, setting a baseline for measuring decisions' progress and success.
- **Identify** which KPIs and metrics will best gauge the efficacy and consequences of these decisions, providing a quantifiable measure of progress.
- **Understand** the role of analytic scores in informing these decisions, ensuring that data-driven insights are leveraged to their full potential.

With this foundational understanding, we can devise a strategy for presenting analytic scores and prescriptive recommendations to our key stakeholders in an understandable and actionable format, ensuring that data insights lead to meaningful action.

Template #8: "Map Scores to Decision Recommendations" Canvas

The "Map Scores to Decision Recommendations" canvas provides a structured framework for documenting and visualizing how each component interconnects. It facilitates clear communication of how data-driven insights can directly inform and optimize stakeholder decision-making.

For example, in the "Improve Local Events Marketing Effectiveness" prioritized use case, we would use the "Map Scores to Decision Recommendations" canvas to determine how the scores can inform critical decisions related to marketing strategy, budget allocation, and local event selection (Figure 39).

Step 7: Map Scores to Decision Recommendations

Template 8 of 9

Business Initiative: Chipotle: Increase "Same Store Sales" by 7.1% over the next 12 months

Prioritized Use Case: Use Case: Increase Store Traffic via Local Events Marketing by X%
Key Decisions: Local Events, Promotions, Sponsorships, Social Media

Key Decisions	Stakeholder	Analytic Score
Decide in which local events to participate	Store Management: recommend which local events are relevant in their area Field Marketing: recommend which local events are suitable and effective to sponsor Corporate Marketing: confirm local events are aligned with brand and fund them Local Community: recommend local events to sponsor and type of participation	• Local Event Score measures how suitable and effective a local event is to sponsor • Local Event Impact Score measures positive impact a local event has on brand awareness, and customer acquisition
Decide what type of promotion for local event	Store Management: recommend the type of promotion to increase store visibility & reputation Field Marketing: recommend the type of promotion to attract target audience Corporate Marketing: recommend the type of promotion to build brand image & awareness	• Promotion Effectiveness Score measures effectiveness in attracting customers. • Promotion Optimization Score measures effectiveness for maximizing revenue and profit
Decide what level of sponsorship and how much to spend on that sponsorship	Store Management: recommend the level of sponsorship to increase store visibility & reputation Field Marketing: recommend the level of sponsorship to attract target audience Corporate Management: recommend the level of sponsorship to build brand image & awareness	• Sponsorship Level Score measures sponsorship effectiveness is for visibility and reputation. • Sponsorship Budget Score measures sponsorship budget affordability given objectives.
Decide how much time and money to invest in social media to promote local event involvement	Store Management: recommend social media investment to promote store & local event involvement Field Management: recommend social media investment to attract the target audience Corporate Management: recommend social media investment to build brand image & awareness	• Social Media Engagement Score measures social media engagement effectiveness • Social Media Conversion Score measures social media-driven customer acquisition effectiveness

Version 2.0

Figure 39: "Map Scores to Decision Recommendations" Canvas

The "Map Scores to Decision Recommendations" canvas captures the following critical and actionable information:

- **Key Decisions**: This column lists the crucial choices that stakeholders need to make to optimize the use case. For "Improve Local Events Marketing Effectiveness," decisions could range from selecting which local events to sponsor to determining the marketing message and promotional offers.
- **Stakeholder**: This section identifies who is responsible for each decision. It could include roles such as the Marketing Director, Event Coordinator, or Social Media Manager, specifying who will act on the recommendations provided by the analytic scores.
- **Analytic Score**: This pivotal column links specific data-driven insights to each decision. For example, an analytic score might indicate the propensity of event attendees to visit a Chipotle store after an event, which would inform the Marketing team on budget allocations and event selections.

The “Map Scores to Decision Recommendations” canvas allows stakeholders to see a clear path from data insights to pragmatic actions.

Step 7 Organizational Benefits

The "Map Scores to Decision Recommendations" canvas bridges the gap between data scientists' analytics development and business stakeholders' practical, day-to-day decisions. The importance of the canvas lies in several areas, including:

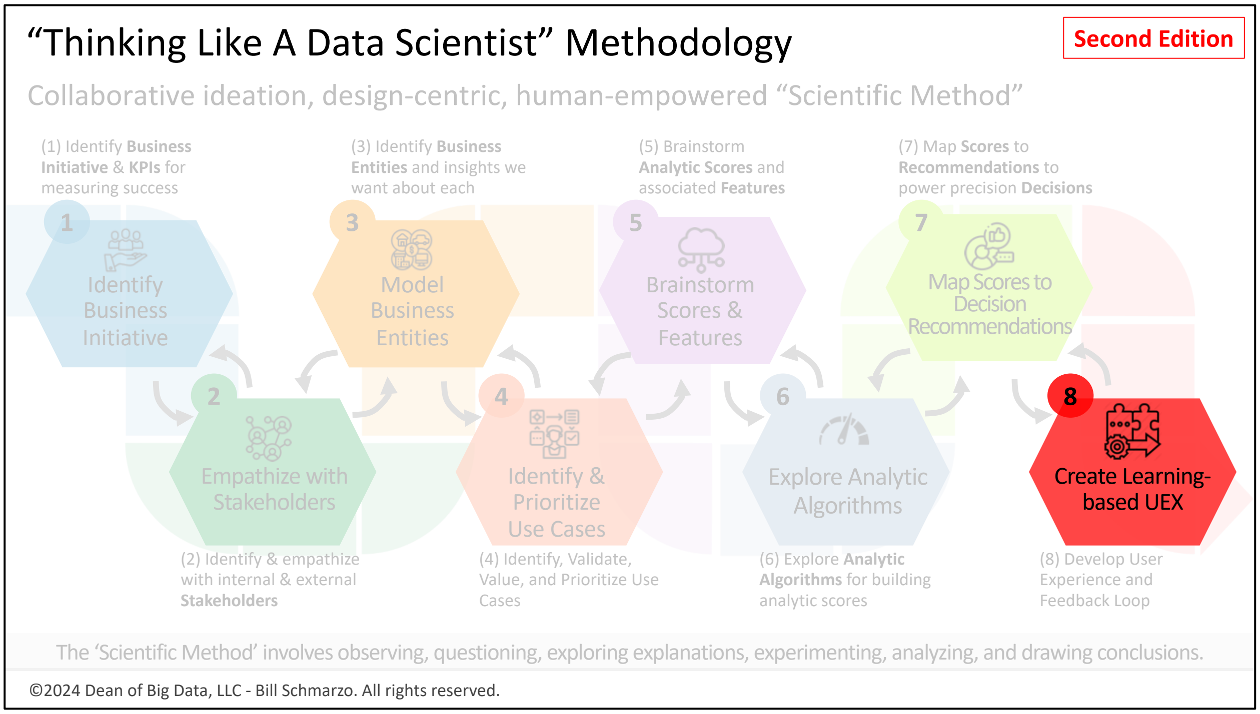
- **Translating data into action**: The canvas acts as a translation layer, converting data and analytic models (scores) into understandable, actionable insights for decision-makers.
- **Ensuring strategic alignment**: It aligns the organization's data and analytics capabilities with its strategic objectives, ensuring that every decision is supported by empirical evidence to drive specific business outcomes.
- **Enabling collaborative insight**: The canvas's visual nature fosters collaboration, enabling stakeholders from various parts of the organization to engage with and understand the data and analytic insights necessary to inform their decision-making.
- **Fueling data and AI literacy**: It encourages a culture of data and AI literacy within the organization by exposing non-technical stakeholders to data-driven thinking and analytics-powered rationale.
- **Ensuring accountability**: The canvas clarifies accountability and encourages ownership of the outcomes by documenting who is responsible for each decision.
- **Guiding resource allocation**: It helps prioritize and allocate precious data and analytic resources by showing which decisions will have the most significant business and operational impact.
- **Supporting continuous learning and improvement**: The canvas is a dynamic tool that evolves as new data comes in, supporting an iterative approach to analytics development and stakeholder decision-making.

In summary, this design canvas is a strategic asset that empowers an organization to operationalize data insights, ensuring that the power of data analytics is fully harnessed to drive critical business and operational outcomes.

Step 7: “Map Scores to Decision Recommendations” Summary

TLADS Step 7 – “Map Scores to Decision Recommendations" – integrates analytic scores with actionable recommendations to support informed decision-making. This step translates data-driven insights into practical, strategic actions. The process includes identifying key decisions, defining the desired outcomes, selecting appropriate KPIs and metrics, and understanding how analytic scores can guide stakeholder decisions.

Chapter 13: TLADS Step 8 – Create Learning-based UEX



The TLADS methodology final step is Step 8: "Create Learning-based UEX." This step involves integrating analytics with human insights to develop a continuously evolving user experience. At this stage, AI models become dynamic and interactive systems that learn from human feedback and interactions. They move beyond simple automation tools to become adaptive systems that increase in complexity with each user interaction.

To create a learning-based User Experience (UEX), it is essential to have a framework that can continuously refine the AI models. Similar to how Generative AI models evolve using Reinforcement Learning with Human Feedback (RLHF), we also want to leverage user feedback to ensure that the AI models and analytic scores keep improving.

> 💡 ***Reinforcement Learning with Human Feedback*** *(RLHF) is a machine learning technique where an AI model is trained to perform tasks by iteratively receiving feedback from human interactions to adjust and improve its responses or actions.*

To unlock the potential of Reinforcement Learning with Human Feedback (RLHF), it is essential to grasp critical analytical concepts.

- **Continuous learning and adaptation**: RLHF is an established methodology for assisting AI systems in learning and evolving through iterative feedback loops with human interactions. This approach is advantageous for enhancing models based on real-world use and feedback, guaranteeing that the models remain pertinent and effective as conditions change.
- **Leveraging user feedback**: It is important to incorporate user feedback to improve AI-based applications and services. By doing so, the AI models can be fine-tuned to better

meet users' needs and expectations in a dynamic business, economic, and social environment.

- **Human-machine collaboration**: This concept is crucial for developing reliable AI systems. It involves blending human expertise with machine learning to mitigate the risks of model drift, biases, and overfitting. This collaboration is essential because it allows diverse perspectives and corrective feedback to be incorporated into the model training process.
- **Self-improving systems**: A forward-thinking goal is to create AI systems requiring less frequent human intervention while improving precision and relevance. Such systems are designed to be sustainable and efficient, reducing the long-term maintenance burden and improving scalability.
- **Ethical considerations**: It is increasingly important to ensure that AI systems are precise, relevant, and ethically sound. Ethical AI involves careful design to avoid biases, ensure fairness, and maintain transparency, which are essential for user trust and legal compliance.

Step 8 emphasizes the blending of human expertise with algorithmic learning. It seeks to constantly improve AI model accuracy and relevance within constantly changing business, economic, and societal environments, leading to self-improving AI systems that reduce the need for frequent human intervention and create more precise, relevant, and ethically sound AI-driven products and services.

To make the "Learning-based UEX" a reality, Step 8 is split into Part 1) Developing an Analytics Feedback Loop and Part 2) Considering User Experience Learning Implications.

Part 1: Creating AI Model Feedback Loop

A feedback loop is essential to creating a learning-based user experience. It creates an environment where analytics can independently enhance their ability to produce more accurate and relevant recommendations and actions. The feedback loop measures the effectiveness of the AI model and provides feedback to it, allowing for iterative improvements to the precision and reliability of the model (Figure 40).

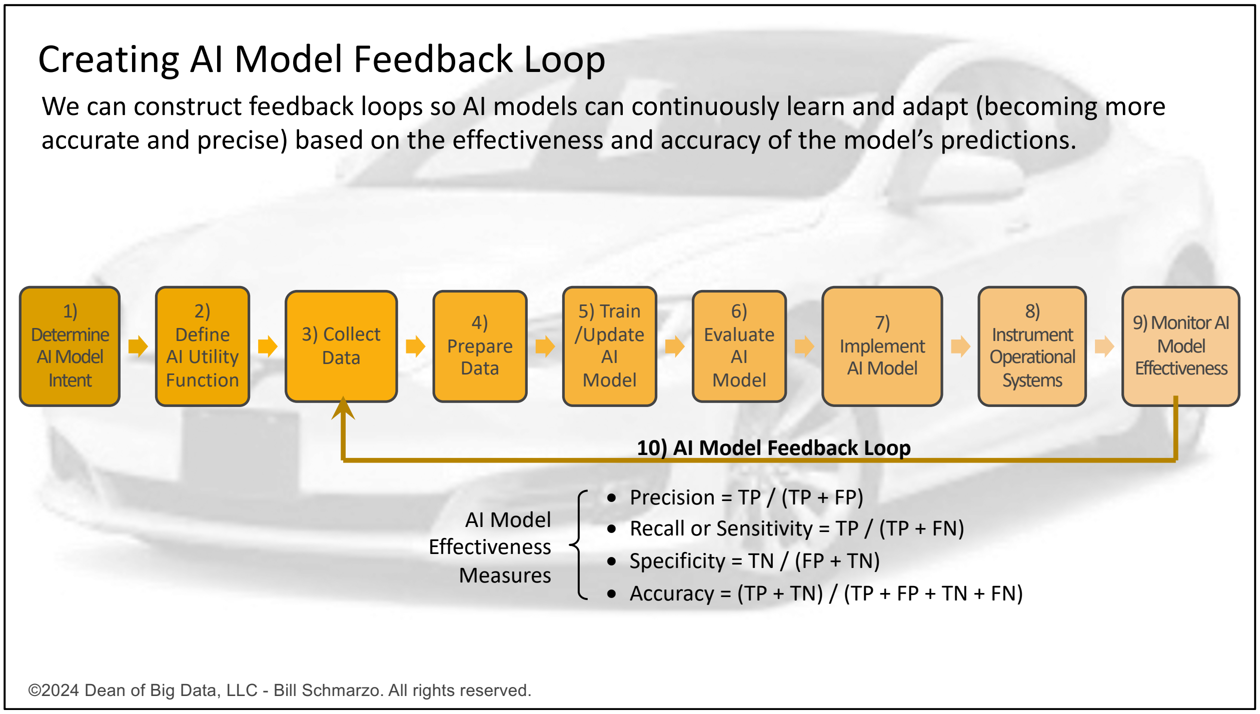


Figure 40: Creating an AI Model Feedback Loop

The process of creating a robust AI model feedback loop includes the following steps:

- **Step 1: Determine AI Model Intent**: Establish the AI model's purpose, targeting specific problems, desired outcomes, the decisions it will influence, and the metrics for assessing decision quality and impact. For instance, a retailer might define their AI model's intent to forecast inventory needs, aiming to minimize overstock and stockouts, serving store managers and supply chain analysts.
- **Step 2: Define AI Utility Function**: Craft a mathematical function that captures the value of the model's predictions, ensuring alignment with business goals and stakeholder desired outcomes incorporating considerations for ethical AI practices. For example, the AI utility function for a loan approval AI system could weigh the importance of loan amount, risk of default, and customer relationship tenure to balance the bank's long-term profitability and customer satisfaction
- **Step 3: Collect Data**: Aggregate the data that powers the AI Utility Function, emphasizing the data's relevance and quality given the model's intent and desired outcomes. In the context of a healthcare provider, this might involve gathering patient records, treatment histories, and outcome data to more accurately predict patient readmission risks.
- **Step 4: Prepare Data**: Rigidly prepare data to enhance its quality and completeness, eliminating errors and biases, thus providing a solid foundation for AI model training. For a marketing campaign, data preparation would involve cleaning customer demographic data, purchase history, and engagement metrics to create a high-quality dataset that improves targeting precision.

- **Step 5: Train/Update AI Model**: This step involves identifying and codifying data patterns and relationships to train the AI model. It is important to periodically update and validate the model to improve its performance and accuracy and eliminate any biases. For instance, an e-commerce platform may use historical sales data, customer reviews, and browsing behaviors to train a recommendation engine to suggest products.
- **Step 6: Evaluate AI Model**: Rigorously test the AI model's performance on new datasets before full deployment, pinpointing potential shortcomings that need resolution. For example, a manufacturing plant could test its AI model's prediction of equipment failure against a set of data from sensors not used during the training phase to ensure its predictive maintenance is effective.
- **Step 7: Implement AI Model**: Integrate the AI model into the operational systems, ensuring compatibility with existing infrastructure and adherence to security protocols. For example, a financial institution may integrate a fraud detection AI model within its transaction processing system to flag potentially fraudulent activities in real-time.
- **Step 8: Instrument Operational Systems**: Install monitoring tools to capture the model's output and performance data continuously. For example, a logistics company could implement GPS and telematics sensors to capture real-time data on vehicle performance for its AI-powered route optimization tool.
- **Step 9: Monitor AI Model Effectiveness**: Continuously assess the model against intended outcomes and adjust as market conditions and customer behaviors evolve. For example, an online streaming service would track viewer engagement and subscription retention rates after implementing a new content recommendation model.
- **Step 10: AI Model Feedback Loop**: Improving the accuracy of an AI model involves utilizing systematic and user feedback. This is done by recalibrating new data, adjusting algorithms, and verifying the model's enhancements against the established utility function. For instance, if a social media platform's content moderation AI shows an increase in false positives, the feedback loop would involve retraining with recent user report data and adjusting sensitivity thresholds to improve accuracy.

By following these steps, we can create a feedback loop that enhances the analytical model with each iteration. This ensures that its decisions align with evolving business objectives and ethical considerations.

Part 2: User Experience Considerations

Integrating mechanisms for user feedback is critical in creating a learning-based user experience. This approach prioritizes the user's needs and desired outcomes by capturing feedback at different touchpoints, which helps refine and personalize AI-driven interactions. It also ensures that all improvements are sensitive to ethical standards. By basing these advancements on user-focused design and ethical considerations, a learning-based user experience enhances functional aspects, creating a secure and respectful environment for users to interact and learn with AI systems.

Here are some effective user experience feedback mechanisms that we can use to gather user insights that improve AI model effectiveness:

- **A/B Testing**: A straightforward and powerful method for UEX research, A/B testing involves comparing two versions of a single element (like a webpage or app feature) to see which performs better against predefined metrics such as conversion or click-through rates. Users are randomly divided into two groups to test each version, allowing developers to make data-driven decisions based on direct user feedback and optimize the user experience through empirical evidence.
- **Multivariate Testing (MVT):** Building on A/B testing, MVT allows the simultaneous testing of multiple variables to evaluate how different elements interact and affect user behavior. This can be particularly insightful for understanding how combinations of changes like text, images, and colors impact user engagement on a website, helping to pinpoint the most effective design configurations.
- **Sequential Testing:** This method tests different versions sequentially over distinct time periods to consider external influences such as seasonal variations in user behavior. Sequential testing is beneficial for understanding long-term user interactions and trends but may be more susceptible to external variables than simultaneous testing methods like A/B testing.
- **Reinforcement Learning with Human Feedback (RLHF):** RLHF integrates human input directly into the training loop of an AI model, allowing it to adapt and refine its algorithms based on qualitative feedback from actual user interactions. This method enhances the AI's understanding and performance by iteratively adjusting its responses based on specific, nuanced feedback, leading to a more user-centric and effective model that aligns closely with user expectations and preferences.

Figure 41 provides a quick summary of some of the different User Experience (UEX) testing methods.

User Experience (UEX) Testing Methods

UEX testing is essential for developing a process that captures user feedback, enabling analytics to continuously adapt and improve based on real-world interactions and needs

Testing Method	Benefits	Key Differences
A/B Testing	• Simple and effective • Direct feedback for specific changes	• Tests only two versions at a time • Best for straightforward, binary testing scenarios
Multivariate Testing (MVT)	• Tests multiple variables simultaneously • Provides insights on variable interactions	• More complex, requires more traffic to achieve statistical significance • Suitable for exploring how different elements interact within a single test environment
Sequential Testing	• Useful for observing long-term effects and external influences • Can adjust for seasonal or environmental changes	• Tests are conducted one after another, not simultaneously • Potential for time-related bias affecting the results
Reinforcement Learning (RL)	• Continuously learns and adapts from operational interactions • Personalizes user experience based on ongoing data	• Based on a model of rewards and penalties • Highly dynamic, suited for environments where user behavior or conditions are continuously evolving

Figure 41: User Experience (UEX) Testing Methods

These techniques provide robust frameworks for enhancing user experience and achieving the users' intent and desired outcomes by integrating direct user feedback into AI system design and functionality, ensuring continuous improvement and alignment with user needs.

Template #9: "Analytic Score Feedback" Canvas

The "Analytic Score Feedback" canvas has been created to help identify the feedback needs for the analytic scores that support each decision in our prioritized use case. This canvas can help the business and data science teams collaborate better and ensure that the engineered analytic scores deliver measurable value. It also enables the scores to continuously learn and adapt to changes in the operational environment (Figure 42).

Step 8: Create Learning-based UEX

Template 9 of 9

Business Initiative: Chipotle: Increase "Same Store Sales" by 7.1% over the next 12 months

Prioritized Use Case: Use Case: Increase Store Traffic via Local Events Marketing by X%
Key Use Case Decisions: Key Decisions: Local Events, Promotions, Sponsorships, Social Media

Decision	Analytic Scores	Stakeholders (Decision Makers)	Presentation Medium	Decision Effectiveness Metrics	UEX Testing Methods
Which Local Events to Sponsor	•Local Events Goodness of Fit Score	•Store Manager •Field Marketing	Store Operations Dashboard	•Customer Satisfaction •Event Awareness	•A/B Testing of Different Events •Sales Team Feedback
Which Local Events to run promotions	•Local Events Goodness of Fit Score	•Store Manager •Field Marketing	Email with Recommendation	•Projected Store Traffic Estimate •Conversion Rate	•Simulation
What promotion types and customer value	•Promotional Effectiveness Score •Pricing Effectiveness Score	•Field Marketing	Mobile app	•Promotion Redemption Estimate •Campaign ROI	•Promotion Spend Optimization
Frequency of Social Media Messaging	•Social Media Sentiment Score	•Field Marketing	Store Operations Dashboard	•Social Media Sentiment	•Machine Learning Sentiment Tracking

Version 2.0

Figure 42: "Create Learning-based UEX" Canvas

The "Create Learning-based UEX" canvas is a crucial tool for the ongoing improvements in the effectiveness and relevance of the analytic scores that support our prioritized use case. The key components of this canvas are:

- **Critical Decisions**: These are the pivotal choices necessary to support the targeted use case, as identified in **Step 4**. Understanding these decisions is essential for tailoring analytics effectively.
- **Analytic Scores**: These scores are crucial in guiding key stakeholders to make informed decisions. Derived from the analytics process (as outlined in **Step 5**), they provide valuable, actionable insights that can drive more relevant, accurate, and ethical decisions.
- **Stakeholders**: These are the key individuals who rely on the analytic scores to guide their decision-making. These stakeholders were identified during **Step 2** and play a central role in the success of the use case.
- **Presentation Mode**: This aspect defines how the analytic scores and their associated recommendations will be presented to the critical stakeholders. Whether through reports, dashboards, or interactive visualizations, the chosen mode impacts communication and feedback effectiveness.
- **Decision Effectiveness Metrics**: These metrics serve as yardsticks for evaluating the impact of analytic scores and their prescriptive recommendations. By defining these metrics, we establish clear criteria for measuring progress and success.

- **Measurement Method**: This component outlines the specific metrics used to assess the effectiveness of the analytic model and details how these metrics will be captured and tracked over time.

The “Create Learning-based UEX” canvas ensures that the analytics align with critical decisions, empower stakeholders, and contribute to data-driven success for the prioritized use case.

Step 8: “Create Learning-based UEX” Summary

Chapter 13 of the TLADS methodology focuses on Step 8, which involves creating user experiences that integrate continuous learning mechanisms. This step emphasizes the importance of using AI to evolve through feedback from real-world use. By creating systems that adapt and improve over time, they become more personalized and effective. The chapter highlights using tools such as the "Create Learning-based UEX" canvas and the "Analytic Score Feedback" canvas to facilitate this integration. These tools help map out how feedback mechanisms enhance the utility of AI models, ensuring they remain relevant and effectively address user needs. The ultimate goal is to foster a dynamic environment where human insights continually refine AI functionality, leading to more intelligent systems that are attuned to the evolving conditions and preferences of their users.

Chapter 14: TLADS Hypothesis Development Canvas

The Hypothesis Development Canvas is a useful tool that consolidates all the necessary data and analytical requirements from the "Thinking Like a Data Scientist" methodology into a single document. This document helps guide the data science team in developing AI models that are optimized for the targeted use case. It serves as the starting point for the data science development process, as discussed earlier in Figure 11.

The Hypothesis Development Canvas is a structured framework that streamlines the initial stages of model development. It helps identify and document key hypotheses and variables, while also fostering collaboration across various departments. By involving stakeholders from business, operations, and data science from the outset, the canvas ensures that the insights generated are relevant and actionable. This cross-functional engagement is crucial for aligning the project’s objectives with the strategic goals of the organization, maximizing the impact and relevance of the data science initiatives undertaken.

The Hypothesis Development Canvas ensures that the data science work directly supports the organization’s business and operational strategies and that the organization’s precious data science and business leadership resources are focused on the organization’s most important business initiatives.

Hypothesis Development Canvas Benefits

The benefits of the Hypothesis Development Canvas as a summary of the Thinking Like a Data Scientist methodology are (Figure 43):

- The Hypothesis Development Canvas provides a clear and concise way for the business stakeholders to understand the data science work and its value proposition. It helps them to see how the data science activities are aligned with the business initiatives and goals and how they can support critical business decisions. It also helps them provide feedback and guidance to the data science team and monitor and evaluate the outcomes and impacts of the data science work.
- For the data science team, the Hypothesis Development Canvas provides a structured and organized way to frame and prioritize the data science ideas and use cases. It helps them define and refine the data and analytics requirements for each use case or hypothesis and validate them with actual data and analytics. It also helps them communicate and collaborate with the business stakeholders and decision-makers and demonstrate the value and results of their data science work.

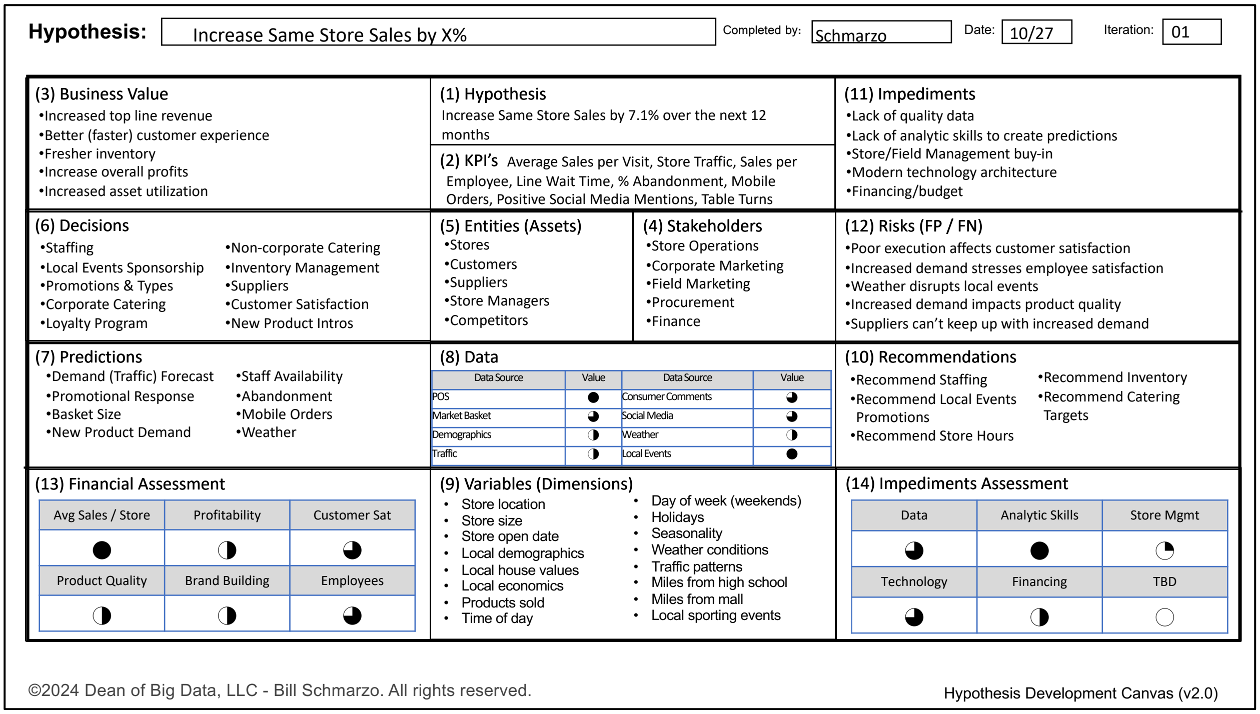


Figure 43: Hypothesis Development Canvas

Here are the instructions for the completion of each of the numbered panels on the Hypothesis Development Canvas:

1. What hypothesis or **use case** are you trying to improve, reduce, or optimize (e.g., increase customer retention, reduce obsolete and excessive inventory, optimize marketing spend, reduce unplanned operational downtime)?
2. What are key performance indicators (KPIs) or metrics against which hypothesis progress and success will be measured?
3. What is the business value of the hypothesis from the financial, operational, and customer perspectives? That is, what are the benefits of the successful execution of the hypothesis on the finances of the business (e.g., increase sales, reduce costs, optimize spend), the operations of the business (e.g., improve operational effectiveness, reduce unnecessary inventory, reduce overtime, reduce logistics, provisioning, and inventory costs), and upon customers (Customer Satisfaction, Likelihood to Recommend, Net Promoter Scores)?
4. Who are the business stakeholders, functions, or constituents who either impact or are impacted by the Hypothesis?
5. What are the entities – human or device/machine – around which to build the analytics (e.g., customers, patients, students, turbines, chillers, compressors)?
6. What are the most critical business and operational decisions that need to be made to support the hypothesis?

7. What predictions might be required to support the business or operational decisions (e.g., predict at-risk customers, successful campaigns, potentially failing parts, component wear-and-tear)?
8. What data sources might be necessary to support the predictions? That is, "What predictions are you trying to make, and *what data might be useful in making that prediction*?"
9. What variables and metrics might yield better predictors of performance? How can I leverage the key stakeholders' subject matter expertise to ideate other variables and metrics that might drive more accurate and relevant predictions?

> **Note**: The "**By Analysis**" technique utilizes business stakeholders' subject matter experience (tribal knowledge) to brainstorm additional metrics, measures, and KPIs that might yield better predictions. For example, a store manager might leverage the "By Analysis" technique to brainstorm: "I want to make performance predictions **by** store, store remodel date, most popular products, local demographics, and customer segment."

10. What business and operational recommendations are required to support the decisions? What should I do, where should I do it, who should do it, when should they do it, what will they need to do it, etc.?
11. What are the potential technical, data, and organizational impediments to success (e.g., data availability, data quality, technical skills, architecture, management support)?
12. What are the costs or risks of the wrong predictive model (False Positives and False Negatives)? What are the financial, operational, and customer costs and liabilities associated with the model yielding False Positives (predicting a condition to exist when it does not exist) and False Negatives (not predicting a condition when the condition does exist)?
13. What is this use case's estimated impact level (rough order estimate from 0 to 4) on critical financial metrics (e.g., sales, profitability, customer and employee satisfaction, product quality, and brand building)?
14. What items and their level of impact (rough order estimate from 0 to 4) hinder the organization's ability to implement the use case (e.g., data quality and accessibility, analytic skills, operational management, technology debit, financing, employee adoption)?

Remember that, like the other canvases, the Hypothesis Development Canvas is a living document and should be updated (and version controlled) as the business and data teams learn more about the use case success criteria (including use case desired outcomes, critical decisions, and the KPIs/metrics against which use case success and effectiveness will be measured).

Chapter 15: Thinking Like a Data Scientist Summary

The Thinking Like a Data Scientist methodology is a process that helps organizations utilize data and analytics to enhance their business and operational models. It comprises eight steps that guide stakeholders from defining the target business initiative to creating the analytic scores and features, mapping the scores to decision recommendations, measuring the effectiveness of the prescriptive recommendations, and defining the feedback and learning loop that enables the AI models and analytic scores to continuously learn and adapt with minimal human intervention.

The TLADS methodology helps organizations by:

- **Aligning their data science efforts** with their strategic goals and objectives, ensuring they solve real problems that matter to their customers, employees, communities, suppliers, regulators, and other key constituents.
- **Optimizing key business and operational processes**, such as customer retention, operational costs, product quality, demand forecasting, predictive maintenance, or anomaly detection.
- **Mitigating compliance, regulatory, and security risks** by ensuring their data science models are ethical, transparent, accountable, and trustworthy.
- **Identifying new innovation opportunities** to uncover net new sources of customer, product, and operational value.
- **Creating a more compelling customer experience** by delivering personalized and relevant solutions that meet or exceed customer expectations
- **Fostering a culture of AI and data literacy** among employees by providing essential tools, concepts, techniques, and vocabulary for understanding data science concepts and techniques.
- **Continuously improve and adapt** their data and AI models by capturing and applying feedback and lessons from the business's ongoing operations.

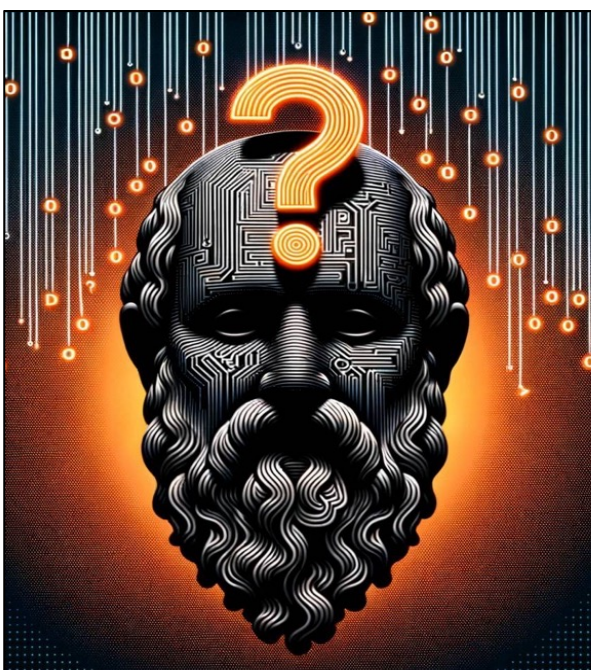
Organizations can better prepare for Thinking Like a Data Scientist by following these steps:

- **Focusing on the appropriate and actionable KPIs**. D&A solutions should be centered around questions and key performance indicators (KPIs) that are relevant, measurable, and aligned with the organization's strategic goals and objectives.
- **Promoting organization-wide decision-making**. Delivering business value in the cheapest and quickest possible ways often requires decentralized decision-making. This means empowering data scientists and business analysts to make recommendations based on data and analytics rather than relying on top-down directives or centralized reports.
- **Building and executing on the use case road map**. A clear and comprehensive list of use cases that data and analytics can address can help prioritize the most impactful and

feasible projects. A use-case library can also serve as a reference for future opportunities and innovations.

- **Delivering solutions and stakeholder value incrementally**. Rather than building a one-time solution that may not meet the changing needs or expectations of the stakeholders, it is better to deliver solutions in small batches that can be tested, validated, and improved over time.
- **Operationalizing the "Economies of Learning"**. This approach harnesses continuous learning from data to refine operations, enabling organizations to improve efficiency and decision-making while fostering a knowledge-sharing culture that drives ongoing operational excellence.
- **Inspiring "Citizens of Data Science."** Data science is more than just technical skills. It also requires strong communication and collaboration skills. To achieve this, data scientists and business stakeholders must work together to define their problems, intent, desired outcomes, preferences, and feedback process fully and accurately. They should present clear explanations, visualizations, and prescriptive recommendations that are easily understandable and actionable.

We are about to embark on an exciting and educational journey into the transformative realm of Generative AI (GenAI). The "Thinking Like a Data Scientist with GenAI" workbook is designed to integrate a GenAI tool such as OpenAI ChatGPT, Microsoft Copilot or Google Gemini into the "Thinking Like a Data Scientist" methodology. Our journey will harness the capabilities of GenAI to fuel our inherent curiosity and guide us in leveraging the GenAI tool in collaboration with the TLADS approach to drive the development of more relevant, meaningful, responsible, and ethical AI outcomes. Get ready to channel your inner Socrates!



GenAI Prompt Engineering:
The Socratic Method

1. Questions for clarification
2. Questions that probe assumptions
3. Questions that probe reasons, rationale, and evidence
4. Questions about viewpoints and perspectives
5. Questions that probe potential implications and consequences
6. Questions about the question
7. Questions to validate information sources

Appendix A: Understanding the Different Analytic Algorithms

Your data science team has access to a wide range of analytic algorithms for constructing analytic scores. Knowing what combinations of algorithms to use in what situations takes experience, patience, and curiosity to find variables and metrics that might be better predictors of performance.

Here are some of those analytic algorithms. I expect new algorithms will be discovered even before this book is completed!

Supervised machine learning algorithms learn from labeled data and make predictions based on the learned patterns. Some examples of supervised machine learning algorithms are:

- **Linear regression**: This algorithm quantifies the relationship between one or more predictor variables and a response variable and can be used to estimate the effect of an exposure on an outcome, predict an outcome using known factors, balance dissimilar groups, etc.
- **Logistic regression**: This algorithm categorizes whether an instance of an input variable either fits within a category or not and can be used for answering yes or no questions, such as detecting fraud, approving loans, or identifying spam emails.
- **Support vector machines**: These algorithms find the optimal hyperplane that separates the data into different classes and can be used for classification or regression tasks, such as recognizing handwritten digits, diagnosing diseases, or predicting stock prices.
- **Decision trees**: These algorithms split the data into branches based on specific criteria and can be used for classification or regression tasks, such as identifying customer segments, predicting sales, or diagnosing faults.
- **Random forests**: These algorithms combine multiple decision trees and use voting or averaging to improve the accuracy and reduce the variance. It can be used for classification or regression tasks, such as detecting malware, predicting customer churn, or estimating house prices.
- **Support Vector Machines (SV)**: A supervised learning algorithm that finds the hyperplane that best separates different classes in the feature space, maximizing the margin between data points of those classes.
- **K-nearest neighbors**: These algorithms classify or regress a new data point based on the similarity or distance to its k-nearest neighbors in the training data and can be used for classification or regression tasks, such as recommending products, predicting ratings, or detecting outliers.

Unsupervised machine learning algorithms: These algorithms learn from unlabeled data and find hidden patterns or structures within the data. Some examples of unsupervised machine learning algorithms are:

- **Clustering**: This technique groups data points based on specific similarities and can be used for recommender systems, anomaly detection, genetics, image segmentation, etc.

- **Association analysis**: This technique identifies patterns or relationships within data that may not be immediately apparent and can be used for tasks such as market basket analysis, customer segmentation, fraud detection, etc.
- **Principal component analysis**: This technique reduces the dimensionality of the data by finding the linear combinations of features that capture the most variance and can be used for tasks such as data compression, visualization, feature extraction, etc.
- **Independent component analysis**: This technique separates a multivariate signal into statistically independent additive subcomponents that can be used for tasks such as blind source separation, feature extraction, noise reduction, etc.

Neural networks are inspired by the structure and function of the human brain and consist of layers of interconnected nodes that process information and adjust their weights based on feedback. Neural networks can handle high-dimensional and nonlinear data and are often used for image recognition, natural language processing, and speech synthesis tasks. Some examples of neural networks are:

- **Convolutional neural networks**: These neural networks use convolutional layers to extract features from images and can be used for tasks such as image recognition, face detection, object detection, etc.
- **Recurrent neural networks**: These neural networks use recurrent layers to process sequential data and can be used for natural language processing, speech recognition, text generation, etc.
- **Long short-term memory networks**: These are a type of recurrent neural networks that use memory cells to store and retrieve information over long time periods. They can be used for tasks such as machine translation, sentiment analysis, video captioning, etc.
- **Generative adversarial networks**: These are networks that consist of two competing models: a generator that tries to create realistic data and a discriminator that tries to distinguish between real and fake data. They can be used for tasks such as image synthesis, style transfer, super-resolution, etc.
- **Transformers:** This deep learning algorithm uses attention mechanisms to process sequential data, such as natural language or speech. Transformers can perform complex tasks such as generating realistic images, captions, and text or creating complex applications like AlphaFold or ChatGPT4.

Other analytic algorithms: These are algorithms that do not fit neatly into machine learning or deep learning categories but are still useful for data analysis and problem-solving. Some examples of other types of analytic algorithms are:

- **Sentiment analysis**: This technique analyzes the emotions or opinions expressed in text and can be used for customer feedback, social media monitoring, product reviews, etc.
- **Monte Carlo simulation**: This technique uses repeated random sampling to estimate the probability of specific outcomes and can be used for risk analysis, optimization, forecasting, etc.

- **Factor analysis**: This technique reduces the dimensionality of the data by finding the latent factors that explain the correlations among the observed variables. It can also be used for tasks such as psychometrics, marketing, genetics, etc.
- **Cohort analysis**: This technique segments data into groups of individuals that share a common characteristic or experience within a defined period and can be used for tasks such as customer retention, user behavior, product performance, etc.
- **Time series analysis**: This is a technique that analyzes data that is collected over time and can be used for tasks such as trend analysis, forecasting, anomaly detection, etc.
- **Reinforcement learning**: A type of machine learning that trains an agent to interact with an environment and maximize a reward through trial and error, and can be used for tasks such as self-driving cars, ad recommendations, personalized chatbot responses, etc.

Here is a matrix that maps the different analytic algorithms to the types of use cases for which that algorithm is most applicable.

Algorithm	Use Case	Rationale
Linear regression	Estimating the effect of an exposure on an outcome, predicting an outcome using known factors, balancing dissimilar groups, etc.	Linear regression quantifies the relationship between one or more predictor variables and a response variable, and test hypotheses about the effects
Logistic regression	Answering yes or no questions, such as detecting fraud, approving loans, or identifying spam emails	Logistic regression categorizes whether an instance of an input variable fits within a category and estimates the probability of the outcome
Support vector machines	Recognizing handwritten digits, diagnosing diseases, or predicting stock prices	Support vector machines find the optimal hyperplane that separates the data into different classes, and handles nonlinear and high-dimensional data
Decision trees	Identifying customer segments, predicting sales, or diagnosing faults	Decision trees split the data into branches based on certain criteria and provides a simple and intuitive way to classify data
Random forests	Detecting malware, predicting customer churn, or estimating house prices	Random forests combine multiple decision trees and use voting or averaging to improve the accuracy and reduce the variance, and handle missing values and outliers
K-nearest neighbors	Recommending products, predicting ratings, or detecting outliers	K-nearest neighbors classifies a new data point based on the similarity or distance to its k nearest neighbors in the training data, and adapt to the local structure of the data
Clustering	Recommender systems, anomaly detection, genetics, image segmentation, etc.	Clustering groups data points based on certain similarities, and reveal the underlying structure or categories of the data
Association analysis	Market basket analysis, customer segmentation, fraud detection, etc.	Association analysis identifies patterns or relationships within data that may not be immediately apparent, and discovers rules

		that describe the associations between different items
Principal component analysis	Data compression, visualization, feature extraction, etc.	Principal component analysis reduces the data dimensionality by finding the linear combinations of features that capture the most variance, and removes noise and redundancy
Independent component analysis	Blind source separation, feature extraction, noise reduction, etc.	Independent component analysis separates a multivariate signal into additive subcomponents that are statistically independent, and recover the original sources from the mixed signals
Convolutional neural networks	Image recognition, face detection, object detection, etc.	Convolutional neural networks use convolutional layers to extract features from images, and handle spatial and translation invariance
Recurrent neural networks	Natural language processing, speech recognition, text generation, etc.	Recurrent neural networks use recurrent layers to process sequential data, and handle variable-length inputs and outputs
Long short-term memory networks	Machine translation, sentiment analysis, video captioning, etc.	Long short-term memory networks use memory cells to store and retrieve information over long time periods, and handle long-term dependencies and vanishing or exploding gradients
Generative adversarial networks	Image synthesis, style transfer, super-resolution, etc.	Generative adversarial networks use two competing models: a generator that tries to create realistic data, and a discriminator that tries to distinguish between real and fake data, and generate novel and diverse data
Sentiment analysis	Customer feedback, social media monitoring, product reviews, etc.	Sentiment analysis analyzes the emotions or opinions expressed in text, and provide insights into the attitudes and preferences of the customers or users
Monte Carlo simulation	Risk analysis, optimization, forecasting, etc.	Monte Carlo simulation repeats random sampling to estimate the probability of certain outcomes, and handle uncertainty and variability
Factor analysis	Psychometrics, marketing, genetics, etc.	Factor analysis reduces data dimensionality by finding the latent factors that explain the correlations among the observed variables, and provide a parsimonious and interpretable representation of the data
Cohort analysis	Customer retention, user behavior, product performance, etc.	Cohort analysis segments data into groups of individuals that share a common characteristic within a defined time period, and provides insights into the trends and patterns of the cohorts

Time series analysis	Trend analysis, forecasting, anomaly detection, etc.	Time series analysis analyzes data that is collected over time, and model the temporal dependencies and seasonality of the data
Reinforcement learning	Self-driving cars, ad recommendations, personalized chatbot responses, etc.	Reinforcement learning train an agent to interact with an environment and maximize a reward through trial and error, and learn optimal policies for complex sequential decision-making tasks
Transformers	Natural language processing, speech recognition, image synthesis, etc.	Transformers use attention mechanisms to learn the context and relevance of each element in a sequence, and handle variable-length inputs and outputs

Appendix B: Blank TLADS Templates

Step 1: Assess Business Initiative

Template 1 of 9

Business Initiative:	
KPIs:	
Desired Outcomes	
Benefits	Potential Impediments
Failure Ramifications	Unintended Consequences Ramifications

Version 2.0

Steps 2: Empathize with Stakeholders

Template 2 of 9

Business Initiative:

(2) Stakeholders

Who are the key stakeholders who impact or are impacted by business initiative and why is initiative important to them?

Stakeholder	Importance to Stakeholder	Potential Impediments	Key Decisions	Desired Outcomes	KPIs / Metrics

Version 2.0

Step 3: Model Business Entities

Template 3 of 8

Business Initiative:

(3) Business Entities

What are the business initiative's key Business Entities (human or device/thing) and what insights would we want to know about each?

Entity	Entity Description	Predictive Insights

Version 2.0

Step 4a: Identify & Assess Use Cases

Template 4 of 9

Business Initiative:

(4) Use Cases

Case Name	Key Stakeholders	Key Decisions	Desired Outcomes	KPIs and Metrics

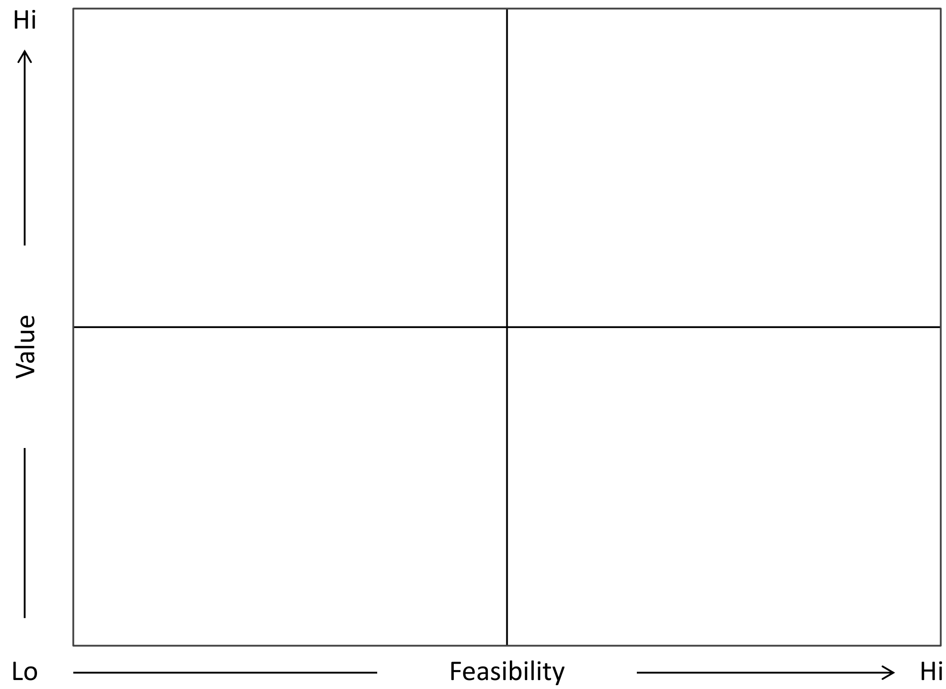
Version 2.0

Step 4b: Prioritize Use Cases

Template 5 of 9

Business Initiative:

What is the relative positioning of each Use Cases vis-à-vis Value and Feasibility?



List Use Cases:

A.

B.

C.

D.

E.

F.

G.

H.

Version 2.0

Step 5: Brainstorm Scores & Features

Template 6 of 9

Business Initiative:

Prioritized Use Case:

Key Decisions	Stakeholders	Potential Analytic Scores	Potential ML Features

Version 2.0

Step 6: Explore Analytic Algorithms

Template 7 of 9

Business Initiative:

Prioritized Use Case:

Analytic Score	Score Explanation	Analytics Algorithm	Algorithm Rationale

Version 2.0

Step 7: Map Scores to Decision Recommendations

Template 8 of 9

Business Initiative:

Prioritized Use Case:

Key Decisions	Stakeholder	Analytic Score

Version 2.0

Step 8: Create Learning-based UEX

Template 9 of 9

Business Initiative: ____

Prioritized Use Case: ____

Key Use Case Decisions: ____

Decision	Analytic Scores	Stakeholders (Decision Makers)	Presentation Medium	Decision Effectiveness Metrics	UEX Testing Methods

Version 2.0

Template 4.1: Use Case-to-Value Drivers Assessment

Value Drivers (Desired Outcomes)

Use Cases								
1								
2								
3								
4								
5								
6								
7								
8								
9								

○ Low ● High | **0** Low **4** High

Template 4.2: Use Case-to-Data Source Mapping

Data Source Mapping

Use Cases											
1											
2											
3											
4											
5											
6											
7											
8											
9											

○ Low ● High

0 Low **4** High

Template 4.3: Use Case-to-Implementation Feasibility Assessment

		Implementation Risks						
	Use Cases							Overall
1								
2								
3								
4								
5								
6								
7								
8								
9								

○ Low ● High

0 Low 4 High

Hypothesis: Completed by: Date: Iteration:

(3) Business Value

(1) Hypothesis

(2) KPI's

(11) Impediments

(6) Decisions

(5) Entities (Assets)

(4) Stakeholders

(12) Risks (FP / FN)

(7) Predictions

(8) Data

Data Source	Value	Data Source	Value

(10) Recommendations

(13) Financial Assessment

(9) Variables (Dimensions)

(14) Impediments Assessment

Hypothesis Development Canvas (v2.0)