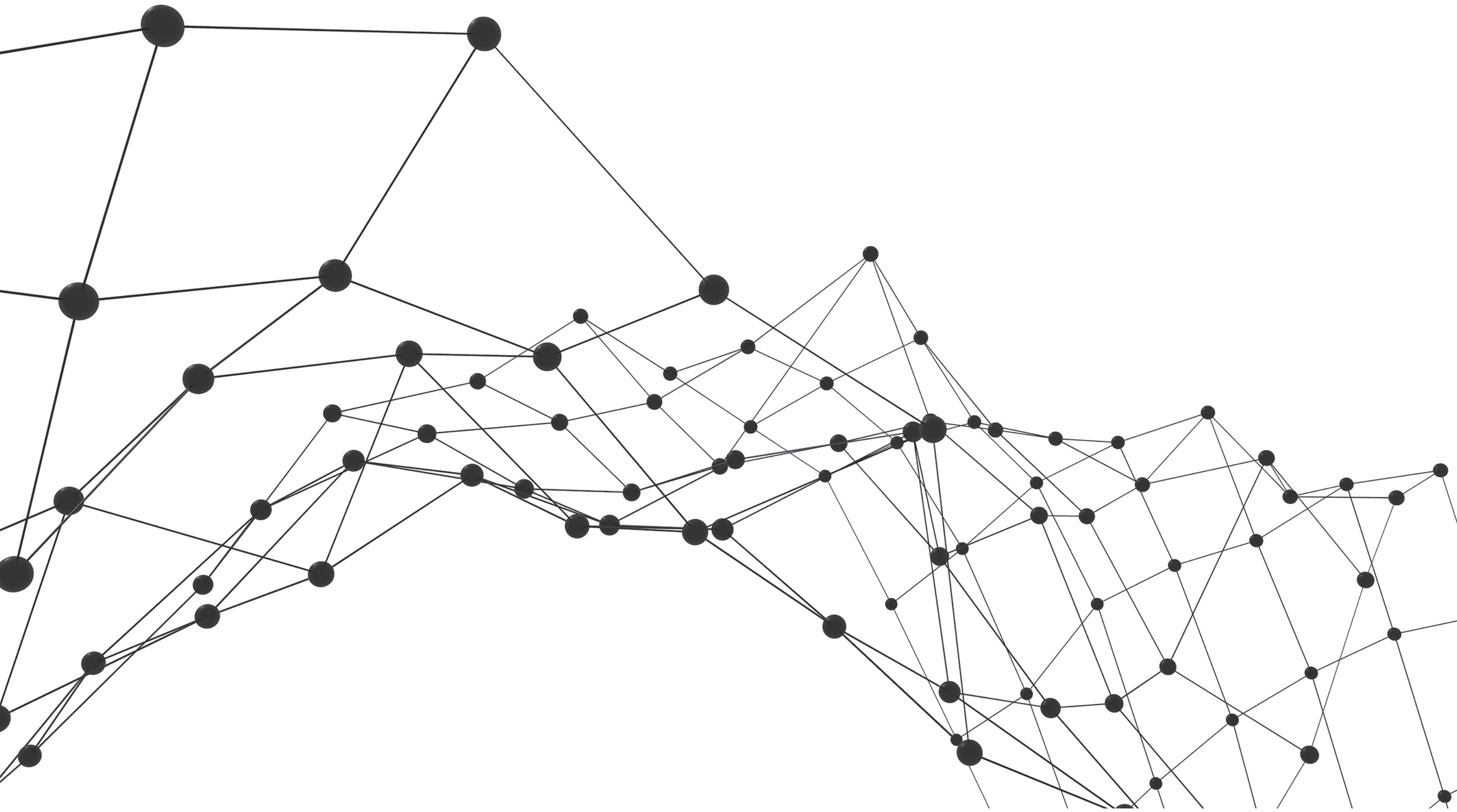


SIMULATION:

Modeling Cause & Effect for Powerful Predictive Analytics

CONTENTS

INTRODUCTION

Predictive analytics can bring the promises of big data to life, mining information about what happened in the past to help enterprises learn what is likely to happen in the future.

Any industry can use predictive analytics to improve efficiencies, reduce risks, and improve revenue. Many already are. These and other fields currently benefit from predictive analytics applications:

Healthcare	Population health management	**Manufacturing**	Preventative machine maintenance
	Optimal staffing		Demand forecast
Retail	Demand forecast	**Financial services**	Credit risk
	Better pricing		Fraud detection

There are two main approaches to implementing predictive analytics: pattern recognition and simulation. Artificial intelligence and machine learning, which generate a lot of buzz across industries, employ pattern recognition, while simulation is another, more human alternative. Simulation is a powerful approach to understanding the causes behind business problems, predicting future trends, and recommending optimum decisions. This ebook explains simulation—where it came from, where it is headed, 5 stages in developing a simulation—and most importantly, how it makes big data useful by producing actionable predictions.

WHAT IS PREDICTIVE ANALYTICS?

Simply put, **predictive analytics** generate predictions from data. Data, statistical algorithms, and machine learning techniques are used to identify the likelihood of future outcomes based on historical data. The goal of predictive analytics is to go beyond knowing what has happened and why it happened to providing guidance on what will happen in the future.

> *The goal of predictive analytics is to go beyond knowing what has happened and why it happened to providing guidance on what will happen in the future.*

Descriptive Analytics is the examination of data or content, usually performed manually, to answer the question "What happened?" or "What is happening?" It is characterized by traditional business intelligence and visualizations such as pie charts, bar charts, line graphs, tables, or generated narratives.

Diagnostic Analytics examines data or content to answer the question "Why did it happen?" Techniques include drill-down, data discovery, data mining, and correlations.

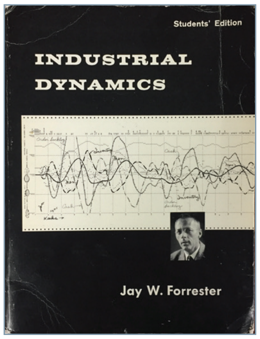
Predictive Analytics examines data or content to answer the question "What is going to happen?" or "What is likely to happen?" Techniques include regression analysis, forecasting, multivariate statistics, pattern matching, predictive modeling, and forecasting.

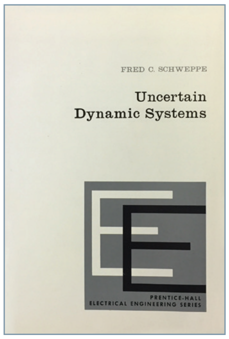
Prescriptive Analytics examines data or content to answer the question "What should be done?" or "What can we do to make ______ happen?" Techniques include graph analysis, simulation, complex event processing, neural networks, recommendation engines, heuristics, and machine learning.

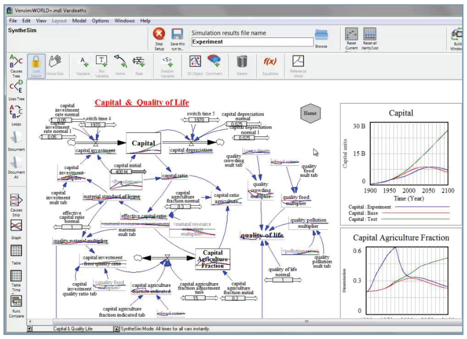
Source: Gartner IT Glossary

A LONG TIME COMING

Simulation has been around for a long time. It originated in engineering control theory and was applied to business challenges by Jay Forrester at MIT Sloan School of Management in the 1960s. Although simulation has an excellent track record, it remained relatively obscure until recently. Its biggest proponents have been academics and specialized consulting firms who have implemented predictive applications in a broad range of industries. While simulation has yet to hit the mass market, there is no doubt interest is building. More and more executives are expressing interest in simulation, while an increasing number of vendors and universities are developing applications. Current market dynamics indicate the technology may be ready to fully emerge.







Jay Forrester pioneers realistic simulation models based on expert knowledge and experience

Fred Schweppe introduces methods for testing and calibrating realistic models using realistic data

Widespread availability of simulation software with proven methodologies and industry applications

1960s

1970s

1980s and Beyond

Figure 1. Simulation for Predictive Analytics Timeline

Dimensional INSIGHT®

SIMULATION AND THE PREDICTIVE ANALYTICS WAVE

Many organizations have implemented business intelligence systems. Now, they are considering predictive analytics to help squeeze more value from their data.

Let's take a look at the broader market for predictive analytics. According to Gartner, predictive analytics is very near the peak in its 2017 Hype Cycle for Analytics and Business Intelligence. Many organizations have implemented business intelligence systems. Now, they are considering predictive analytics to help squeeze more value from their data.

Value from data begins with insights. Insights lead to decisions, and ultimately result in actions that provide organizations with a competitive edge. This work is shared between humans and software. Reports and dashboards only go so far towards impacting better decisions. Humans need to look at the reports, perform the analysis, and come up with their own course of action. Predictive analytics makes greater use of software, bringing humans closer to a decision. Prescriptive analytics goes even further, by recommending "optimized" decisions, and possibly even automating decisions without human involvement. Increasing analytics maturity is not easy and requires greater human input at the beginning stages for each progressive level.

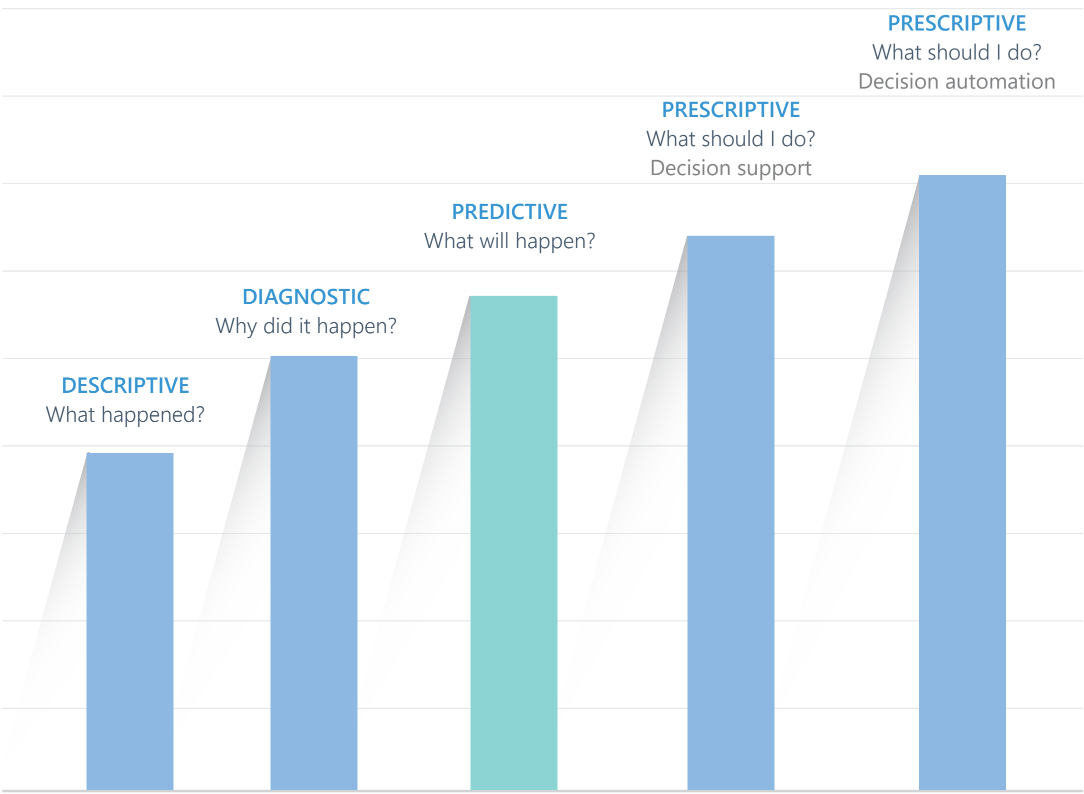


Figure 2. Progression of analytics sophistication and maturity

COMPETING PREDICTIVE METHODOLOGIES

There are two main approaches to implementing predictive analytics: pattern recognition and simulation. The fundamental difference is that pattern recognition relies on correlation while simulation relies on human knowledge of causation.

There are two main approaches to implementing predictive analytics: pattern recognition and simulation. The fundamental difference is that pattern recognition relies on **correlation** while simulation relies on human knowledge of **causation.**

Pattern recognition is inherently data-centric. You throw a bunch of data at an algorithm, it finds patterns in the data and maps future trends. Other things being equal, the larger the data set, the greater the accuracy of the predictions. Therefore, **big data** is highly desired.

Simulation, in contrast, is model-centric. You start by using **human** knowledge of cause and effect to create a model of the system in which the problem operates. You then connect the available data with that model to obtain a future projection. For example, to predict future sales, you would model its key causal factors, such as sales staff experience, product quality, various market factors, and how they all relate to one other. Other things being equal, the greater the expertise of the humans involved, the greater the accuracy of the predictions.

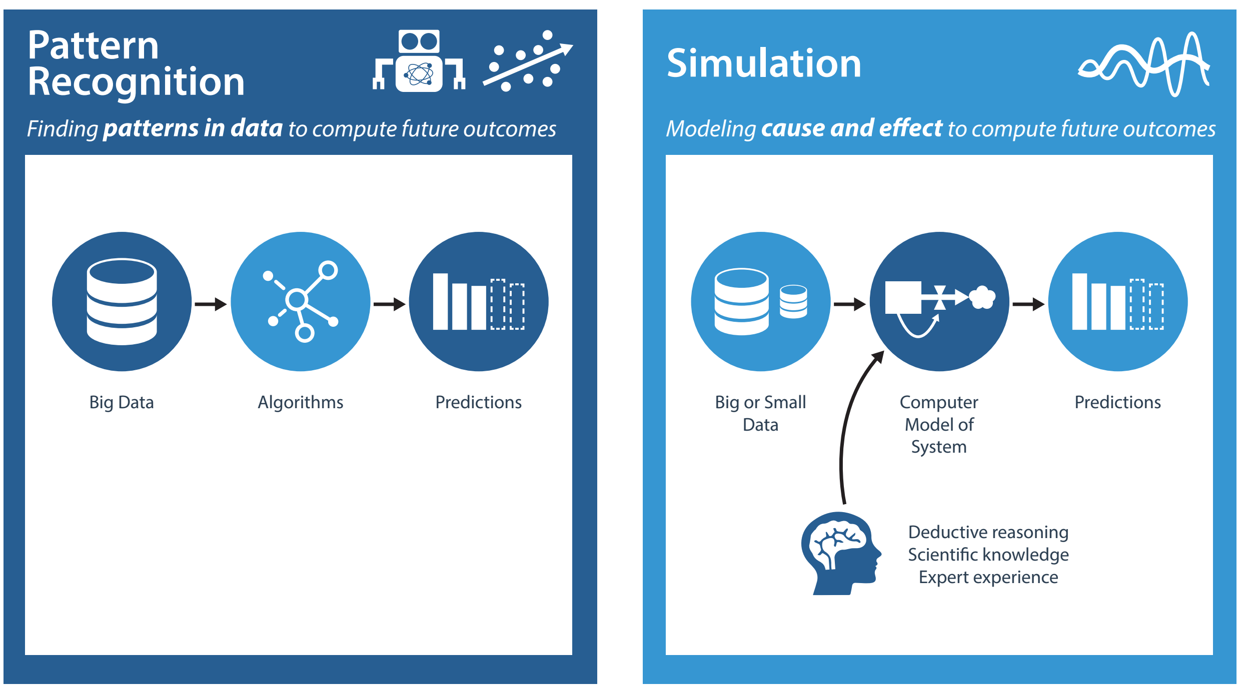


Figure 3. Pattern Recognition and Simulation Data Flows

ADVANTAGES OF SIMULATION

While both pattern recognition and simulation can be effective approaches, simulation has some advantages in predictive analytics.

Pattern Recognition

*Finding **patterns in data** to compute future outcomes*

Advantages:

- ❐ Wide range of approaches (from linear regression to deep learning)
- ❐ Minimal upfront knowledge of causal relationship required
- ❐ Potential to uncover previously unknown casual relationships
- ❐ Large pool of data scientists

Disadvantages:

- ❐ Black box algorithms that are not easily modifiable or transparent
- ❐ Expensive data processing and acquisition
- ❐ Weak reality tests (machine learning algorithms do not know about business constraints or physical law)
- ❐ Poor performance in unstable systems (where casual relationships change over time)

Simulation

*Modeling **cause and effect** to compute future outcomes*

Advantages:

- ❐ Integrates signals missing in the data
  - – Soft factors like time pressure, morale, reputation, etc…
  - – Disruptive events that are not frequent enough to produce reliable correlations
- ❐ Low data acquisition and processing costs
- ❐ Reliable predictive accuracy
  - – Avoids false predictions using reality based checks
  - – Robust over time (accounts for unstable systems where causal relationships change over time)
  - – Performs well even with limited data to extrapolate a trend

Disadvantages:

- ❐ Requires access to experts who understand the relevant cause and effect
- ❐ Small pool of simulation modelers
- ❐ Calibrating models is a slow process

Figure 4. Advantages and Disadvantages of Pattern Recognition and Simulation

Advantage 1: Simulation integrates signals missing in the data

Often, key causal factors are not present in your data. For example, soft factors, such as time pressure, morale, and reputation can have a significant effect on desired outcomes, but are rarely captured by information systems. In simulation, everything that is known about the missing factors can be included in the model, and unknown factors can be estimated. The resulting projections will take these factors into consideration, and quantify the degree of uncertainty.

Advantage 2: Simulation has relatively low data acquisition and processing costs

In contrast to pattern recognition, which relies on large volumes of high-quality data, simulation uses the data that is available and supplements it with knowledge. In addition, simulation does not require all of the data that "might be related" to the problem to look for meaningful correlations. The causes of the problem are already built into the model. Therefore, the data acquisition stage for simulation often is less time-consuming and less costly.

Advantage 3: The accuracy of simulation predictions is highly reliable

One of the challenges with pattern recognition is that correlation does not always reflect causality. Often data will contain correlations that appear to be causes, but are not. Such false correlations lead to failed predictions. Simulation starts with expert understanding of cause and effect, which is grounded in scientific knowledge, and produces reliable results. Simulation also employs a model testing and adjustment phase that both improves predictive accuracy and improves our understanding of cause and effect.

5 STAGES TO CREATING SIMULATION-BASED PREDICTIVE APPLICATIONS

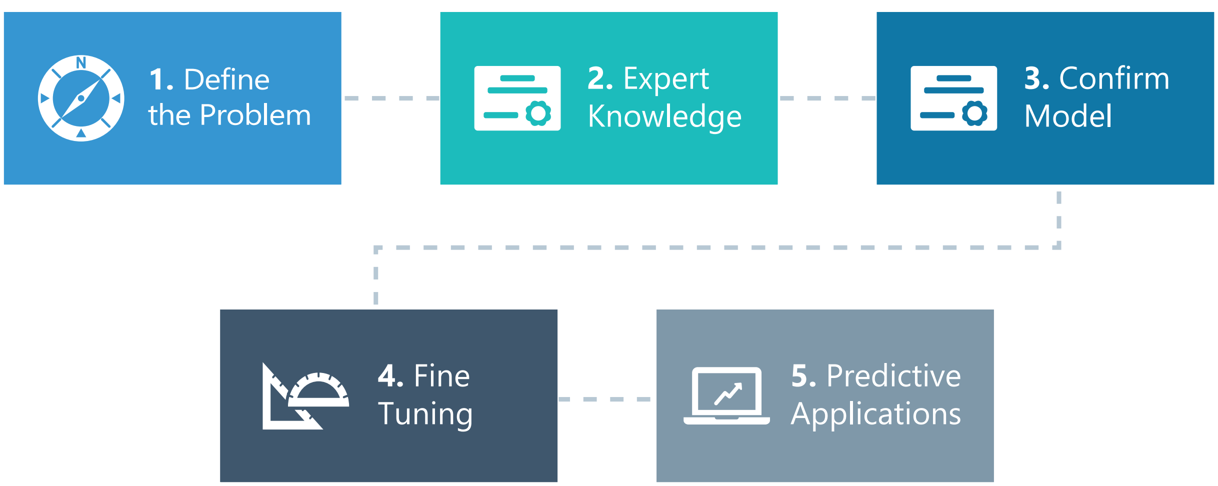


Figure 5. Five Stages in Simulation-Based Predictive Application Development

Stage 1: Define the Problem

Defining the problem is the first stage in any predictive project. Different problems may require different approaches, and for some problems you may want to consider combining multiple approaches. When defining the problem, it is wise to narrow the scope by focusing areas that will have the greatest return on investment.

For example, consider a healthcare organization striving to reduce its hospital-acquired infections. Rather than looking at all hospital-acquired infections, it might want to begin by predicting only central line associated blood stream infections (CLABSI).

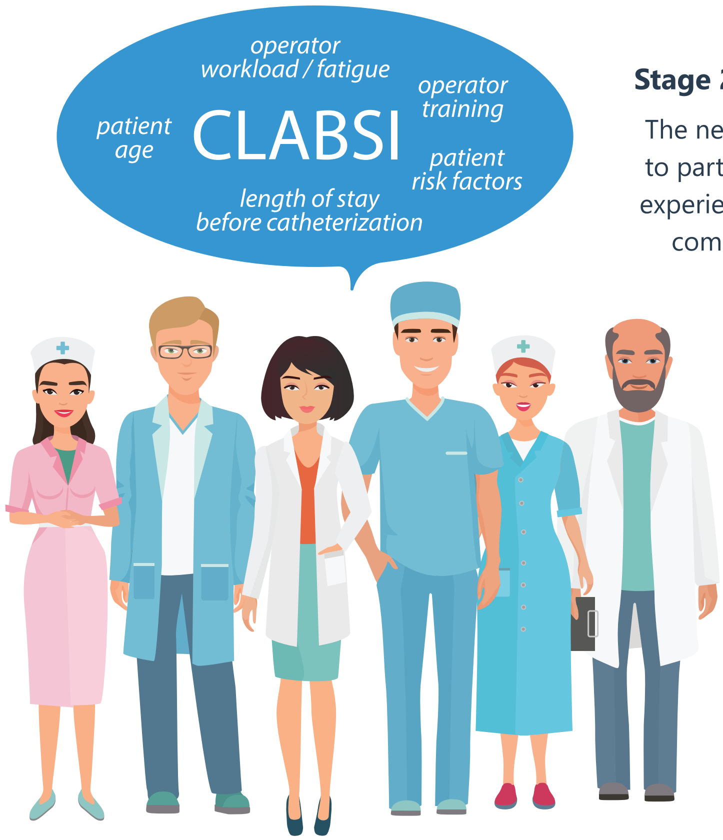


Figure 6. Stage 2—Gathering Expert Knowledge

Stage 2: Expert Knowledge

The next stage involves forming a group of "experts on the floor" to participate in the project. These individuals should have deep experience and involvement in the area of focus, so they can communicate the problem causes to the simulation consultant.

Simulation consultants are skilled at asking questions that uncover causal relationships that are often not considered by people engaged in the problem. They also contribute to the knowledge gathering by sharing their understanding of causal relationships from related projects. The combined knowledge collected in this stage is stronger than the mental models that any expert working the problem possesses or is capable of considering in the course of decision-making.

> *Simulation consultants are skilled at asking questions that uncover causal relationships that are often not considered by people engaged in the problem.*

Stage 3: Confirm Model

Next, the consultant takes the knowledge and creates system dynamics diagrams of the causal relationships. These diagrams use a human-understandable design language consisting of feedback loops, stocks and flows, and time delays. The "experts on the floor" review the diagrams together with the consultant and provide feedback.

The confirmed diagrams are the foundation for the mathematical model that is created in Stage 4 to power the simulation. The consultant uses simulation software adding parameters and equations to the causal relationships.

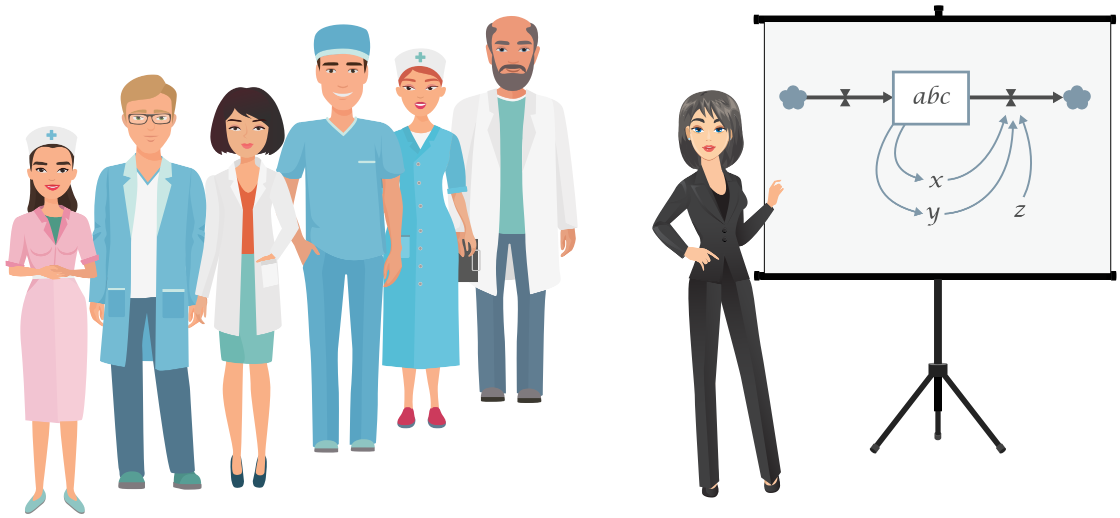


Figure 7. Stage 3—Confirming the Visual Model

Once the mathematical model is complete, you can feed data into the model and get predictions. If the project team has done a good job, the initial predictions should be valuable. However, before using these predictions in the real world, it is first important to test and tune the model.

Stage 4: Fine-Tuning

There are two primary techniques for tuning a simulation model. One technique uses reality-based checks to ensure the model behaves appropriately under extreme conditions and accounts for business constraints and physical laws. The other technique takes data from a past point in time and runs the model to see if it accurately predicts the known historic trend. During the tests, the model is adjusted to improve accuracy, and the experts involved adjust their conclusions about causal relationships.

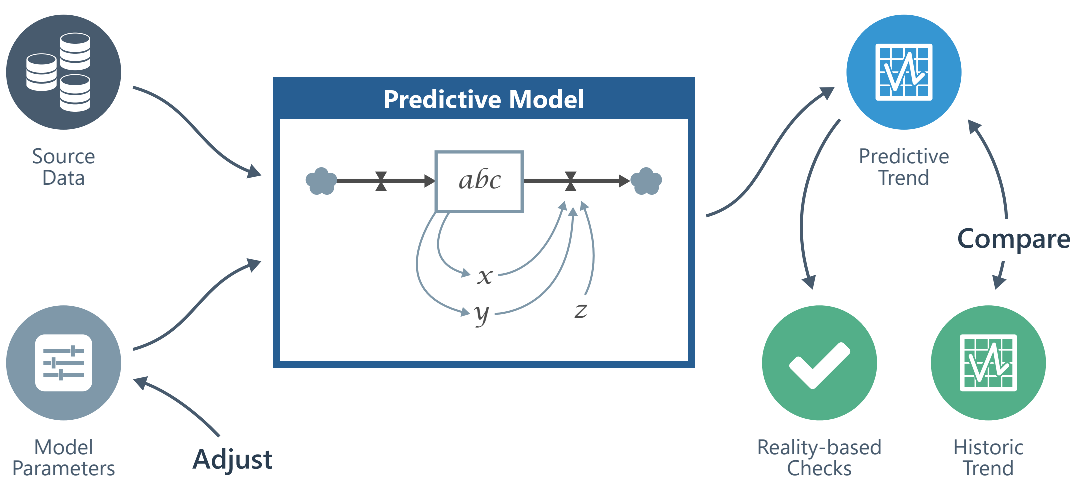


Figure 8. Stage 4—Tuning the Mathematical Model

Stage 5: Predictive Applications

After the model is tuned for a reliable degree of predictive accuracy, the project's final stage is to connect the model to a production data stream and visualize the output within reporting and decision support systems. The output might be a simple indicator embedded in an operational system indicating that a patient is high risk for acquiring a hospital-acquired infection. Or the output could be a visual dashboard showing charts and trends, with controls for users to perform "what-if" analyses on all of the possible decision options.

Since simulation models are built from causal relationships, which do not frequently change, a production system should be reliable for an indefinite period of time. If the business needs or decision parameters change over time, the model can be adjusted to reflect these changes. The model also can be expanded upon to make use of new data sources or obtain more value by addressing other problem areas. In the CLABSI example, that might be expanding the model to predict other hospital-acquired infections.

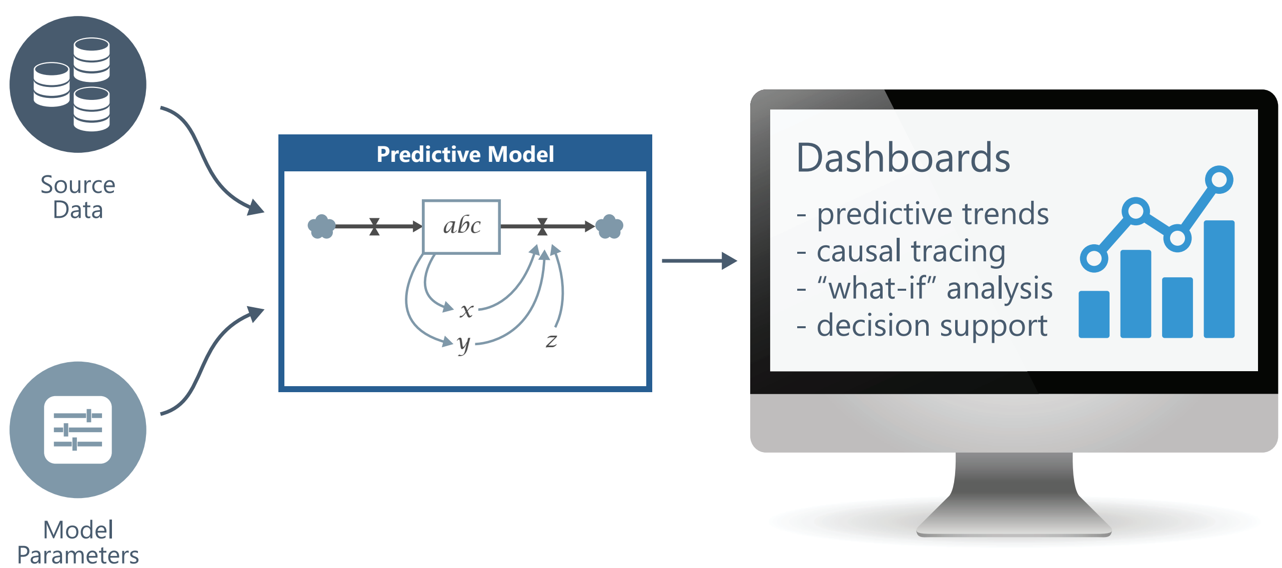


Figure 9. Stage 5—Bringing Predictive Application Online

BEYOND PREDICTION: PRESCRIPTIVE ANALYTICS AND AUTOMATED DECISION MAKING

While the benefits of helping people understand future impacts of their decisions are huge, simulation can go even further by prescribing and even automating decisions.

While the benefits of helping people understand future impacts of their decisions are huge, simulation can go even further by prescribing and even automating decisions.

In applications with decision variables, the computer can run simulations for all the possible combinations and then compare the outcomes against performance targets and recommend the "optimum" decision set. The recommended decisions can be presented to the decision makers along with information explaining the factors that make the decision optimal. For decisions that do not require human intervention, the optimal decision can be passed on to a system downstream to carry out the action.

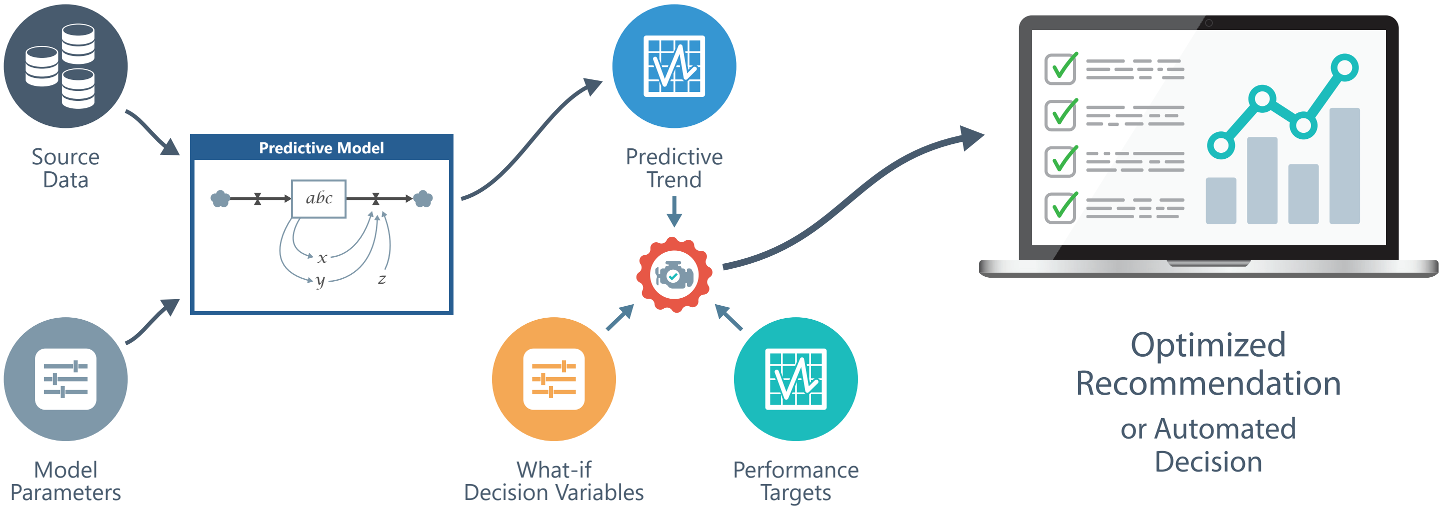


Figure 10. Optimized Recommendation/Automated Decision Engine

CLOSING THOUGHTS

It is not just hype: predictive analytics is here to stay. Data and analytics have become strategic assets for organizations, with insights relevant throughout each enterprise and across industries. Value comes from putting these insights into action. Predictive analytics offers every industry opportunities to improve revenue, reduce risk, and increase efficiencies by mining past data to forecast what is likely in the future.

Simulation is a powerful approach that can be used to understand the causes behind business problems, predict future trends, and recommend optimum decisions. In this period of intense hype around artificial intelligence an d machine learning, consider simulation, a practical and reliable *human* approach that can solve many of the same problems.

LIST OF SOURCES

1. Forrester, Jay Wright, *Industrial Dynamics.* MIT Press, 1961
2. Gartner Analytics Maturity Model, adapted from the 2016 Gartner BI Summit presentation, *Moving your Data and Analytics from Laggard to Leader*
3. *Gartner IT Glossary*
4. https://en.wikipedia.org/wiki/Jay_Wright_Forrester
5. https://en.wikipedia.org/wiki/Pattern_recognition
6. https://en.wikipedia.org/wiki/System_dynamics
7. https://healthitanalytics.com/news/93-of-payers-providers-say-predictive-analytics-is-the-future
8. https://www.gartner.com/doc/3772080/hype-cycle-analytics-business-intelligence
9. Schweppe, Fred C., *Uncertain Dynamic Systems.* Prentice-Hall, 1973
10. Sterman, John, *Business Dynamics: Systems Thinking and Modeling for the Complex World.* McGraw-Hill Education, 2010

ABOUT **DIMENSIONAL INSIGHT**

Dimensional Insight is a leading provider of analytics and data management solutions, offering a complete portfolio of capabilities ranging from data integration and modeling to sophisticated reporting, analytics, and dashboards.

Founded in 1989, Dimensional Insight has thousands of customer organizations worldwide. Dimensional Insight's Diver Platform® consistently ranks as a top performing analytics platform by customers and industry analysts in its core market segments including healthcare, manufacturing, and beverage alcohol. For more information, please visit https://www.dimins.com.

About Dimensional Insight

Dimensional Insight® is a leading provider of analytics, data management, and performance management solutions, offering a complete portfolio of capabilities ranging from data integration and modeling to sophisticated reporting, analytics, and dashboards. The company is a seven-time Best in KLAS winner in healthcare business intelligence and analytics, most recently ranking #1 in 2020. Founded in 1989, Dimensional Insight has thousands of customer organizations worldwide. Dimensional Insight consistently ranks as a top performing analytics organization by customers and industry analysts in its core market segments including healthcare, manufacturing, and beverage alcohol.

60 Mall Road
Burlington, MA 01803
t: 781.229.9111
www.dimins.com