

A man with dark hair and a beard, wearing a dark suit jacket over a light blue patterned shirt, is smiling and looking down at his smartwatch. He is in a modern, brightly lit office or cafe setting with wooden beams and large windows in the background. The overall mood is positive and professional.

Your Predictive Journey

A Practical Guide to Predictive
Analytics and Machine Learning



Introduction

- 1 Why Businesses are Moving to Predictive Analytics
- 2 The Role of Predictive in Your Analytics Strategy
- 3 Understanding Predictive Modeling
- 4 Important Considerations When Embarking on a Predictive Analytics Journey

How different would your business be today if you could have spotted equipment failures before they happened, identified unhappy employees before they quit, or known in advance which customers would up-sell and which wouldn't?

In a world that is being reshaped by data, businesses are now turning to analytic solutions to predict future outcomes and in doing so are driving maximum advantage from their data.

Since the advent of digitization and the Internet of Things, companies have been amassing data at an exponential rate. Much of this growth is due to the expansion of the internet which now collects transactional data from every facet of our economy. Another factor is the recent surge in unstructured, textual data, generated by call center logs, emails, blogs, tweets, customer comments, and customer reviews.

Many businesses have invested in Big Data initiatives in order to capture, store, and manage all this data. They are now looking further – at how to drive return on investment (ROI) from this mass of information by generating insights that add value to their business. The answer lies in predictive analytics software that uses algorithms to find patterns in large volumes of data, thus enabling predictions to be made about the future.

This eBook is aimed at helping you, as a business leader, to understand the potential competitive advantage that predictive analytics can bring to your organization. We will demystify some of the more complex areas – including how predictive analytics fits with existing analytics strategies and how predictive modeling works – and cover proven strategies for success with predictive analytics.



1 Why Businesses are Moving to Predictive Analytics

Businesses, large and small, have been using spreadsheets and business intelligence (BI) tools for years to drive insight from their data. Ten years ago, these tools were deemed 'state of the art', but technology has changed and now, with Big Data, incredible processing speeds, and predictive analytics software, a much more accurate way of predicting the future is available.

Predictive analytics is not entirely new; in financial services it underpins computer-assisted stock trading and in retail, predictions of online buyer behavior are increasingly being used to design personalized marketing campaigns. Elsewhere, companies that understand the competitive edge that it can provide, are starting to invest. A recent Forrester survey¹ of data and analytics decision makers found that 39% have data and analytics budgets of at least \$10 million.

Predictive algorithms are critical assets that give meaning to the growing mounds of data amassed by enterprises.

What Makes Predictive Analytics So Powerful?

Objective and accurate predictions: Unlike BI tools that rely on human input to infer cause and effect between results, predictive analytics uses a technique called machine learning that teaches the computer to look at

¹ Forrester: Insight Platforms Accelerate Digital Transformation

a particular outcome and then uncover the factors behind it (which could include thousands of possible causes and nonlinear relationships). The result is far more accurate predictive outcomes that can improve over time. A leading German car manufacturer is using predictive analytics in this way on the factory floor; processing 30,000 data points per second from up to 100 engine components to predict engine failure as early as possible – minimizing lost hours and resources.

Automated decision making: Algorithms are at their most powerful when used to automate decision making between two or more business processes (requiring no human input at all). A common example is a credit card algorithm that can lock an account if it spots that a purchasing pattern is abnormal for the cardholder and indicates a potential fraud. In manufacturing, where networked factory floors are being further networked into the supply chain, the adoption rate of predictive analytics is expected to expand rapidly. By digitizing analysis and decision making in this way, businesses are responding to situations in real time, achieving groundbreaking efficiencies.



Predictive analytics software is guaranteed to produce a return on investment (ROI). Even better, the ROI compounds. The more you use it, the greater the return.

New opportunities: Predictive analytics is also enabling organizations to discover new business opportunities and build digital models around them. For example, an airfare prediction firm in the United States turned its “waste” data to gold. By mining its archive of itineraries to correlate past trends with current fares, it can now find its customers the lowest possible price 95% of the time. In sales, a leading computer networking company is targeting \$1 billion in extra revenues thanks to predictive analytics providing insight into what its customers are most likely to buy.

So Why Should You Invest in Predictive Analytics?

The main reason for investing in predictive analytics is to use your data to make better decisions, which will improve your business results in a quantifiable and measurable way. It is one of the few investments that is guaranteed to produce a return on investment (ROI). Even better, the ROI of predictive analytics compounds. The more you use it, the greater the return and there are opportunities to use it in every part of your business operations where there is data. Which, in 2016, is everywhere.

Adoption of #predictive #analytics in manufacturing is expected to reach 80% (from today's 22%) over the next six years. #SAP

Source: 2016 MHI Annual Industry Report

Share
This!

What are the Steps Involved?

1

You need a business question. The more central the business question is to your profitability, the greater the return. So, for a telecom provider it could be about customer churn and penetration of multi-play offers. For every business, it's different. But once the business question has been identified, the data science team can go to work.

2

The data science team prepares the data and generates models. This involves merging lots of data sets relevant to your business question and also generating new predictive attributes such as aggregates. They then create different models and select the one that performs the best against your test data.

3

The model is deployed into day-to-day operations and employees use the output to make better decisions. At this stage, it's important to continuously monitor the model's performance and, if needed, fine-tune it to ensure peak performance.

4

The data science team can start working on another business question and then another and another. Soon you have a business with thousands of models running in parallel, helping to answer thousands of business questions more efficiently across the entire company. An organization leveraging its data to maximum effect – all powered by predictive analytics.



Predictive analytics solutions lock ROI into the daily business operations of the company.

2

The Role of Predictive in Your Analytics Strategy


If you are new to predictive analytics, one of your biggest questions may be: 'How would this fit with my existing analytics strategy?'

Terminology in the analytics world can be confusing and can make it hard to see how the component parts fit together: If predictive analytics is all about *data science* then are other descriptive forms of data analysis (such as business intelligence (BI)) not “science” but just “data art”? These analytics terms draw an artificial distinction between descriptive and predictive analytics that can create a barrier for businesses developing an analytics strategy.

Are Descriptive and Predictive Analytics that Different?

Let's take a step back. The core mission of any BI system is to analyze data, extract insights, and enable the user to make the best decisions possible. If you asked a data scientist, they would argue that predictive analytics has the exact same mission. Obviously these two analytical approaches are not replacements for each other – they are in fact ying-yang complements.

The real difference between descriptive and predictive analytics is simply how they achieve their goal of surfacing insights. BI tools make heavy use of visualization techniques and rely on human interaction to drill, slice, dice, and draw data, whereas predictive tools leverage mathematical algorithms guided by humans to look at data in ways and speeds not humanly possible.



BI tools rely on human inference to spot likely cause and effect whereas #predictive #analytics uses machine learning. #SAP

Share This!

Barriers to Predictive Analytics

Unhelpful preconceptions: The real barrier preventing many enterprises from adding predictive analytics to their existing analytics landscape is the perception that it will require a mountain of cash, a library of knowledge, and data scientists that everyone knows are scarce. We're trained to consider everything as a cost, to de-risk projects so much that they are no longer innovative, and to prove tangible value before we even start. Now is a good time to start unlearning those bad habits!

ROI versus TCO: Predictive analytics is not a TCO (total cost of ownership) discussion – it's an ROI (return on investment) one. How? Consider that you had to make a very important business decision, and a magical genie popped out of a lamp to tell you with great certainty which option would make you more money. Would you want to know how much money you could make? Or would you ask how much the lamp costs first?

Descriptive versus Predictive: That said, predictive analytics isn't a genie in a lamp – but neither is a data scientist. Even if your business employs data scientists, they likely do not have the domain knowledge for the business problems you are trying to solve. This is where that nasty barrier between descriptive analytics (BI) and predictive analytics (data science) is especially detrimental. By considering these two disciplines as separate, it is very difficult to have a coherent analytics strategy.

Adding Predictive to Your Analytics Strategy

Fortunately, integrating predictive into your existing plans doesn't have to be difficult, and best of all, it doesn't always require a data scientist. The key for determining where to start is by focusing on a defined business problem that is small, yet important. It may be a problem you don't know how to solve, or one that you are currently struggling with by using a BI solution.

For example, the question "Why are my customers canceling their service with me?" is a common problem that gets to an analyst's desk. The BI approach is to slice, dice, drill, and graph characteristics shared by those who have already canceled, to filter at-risk customers, and reach out to them before they call to cancel. We rely on the analyst's knowledge of the market, customer base, data set, and personal experience, to choose how to successively narrow the data set.

Learn how algorithms are becoming competitive assets in the SAP Digitalist blog, [Algorithms: The New Means of Production](#)

For more on how to 'train' algorithms read the SAP Digitalist blog, [Unlock Your Digital Super Power: How Digitization Helps Companies Be Live Businesses](#)

The predictive analytics approach is to "train" a predictive model by feeding it data on customers who have and haven't canceled so that the model understands each characteristic of the customer and how much each would contribute to their decision to leave. We can then use the predictive model to "score" each customer on their propensity to leave and even understand the key influencers to their potential decision. The BI analyst can then filter on the score – a weighted variable that takes all of the customer's characteristics into account, resulting in a far more accurate prediction than cascading filters in a traditional report or dashboard

Making It Happen

You may be wondering how to "train" and "score" and do all that predictive stuff without a data scientist. Automated predictive algorithms have come a long way and encode many of the steps that

data scientists currently do by hand or through scripting today. At the core are automated predictive algorithms that have been refined by data scientists over many years, both in the mathematics lab and in the real world, with hundreds of customers who rely on them for mission-critical decisions every day.

While it may sound like automated predictive analytics is only for the non-data scientists, it's quite the opposite. Automated predictive technologies make quick work of the boring, tedious, and error-prone tasks that make up "data science." A skilled data scientist can tweak and optimize automated algorithms to make them even better, and at a fraction of the time it would take them to do even a basic predictive model by hand. This frees up the data scientist's precious time to work on more projects and provide value to more parts of the business.

Automated predictive algorithms are like a "Data Scientist-in-a-Box."

Regardless of where you are in your analytics journey and whether you have data scientists on staff or just really smart BI experts, there's a 100% chance that predictive technologies can either remove some analytical obstacles – or at least make them much easier to conquer.

Predictive modeling is complex and the logic behind the predictions it produces can be opaque, so in this chapter we explain some core principles to help you navigate a predictive analytics implementation.

The Desire to Understand

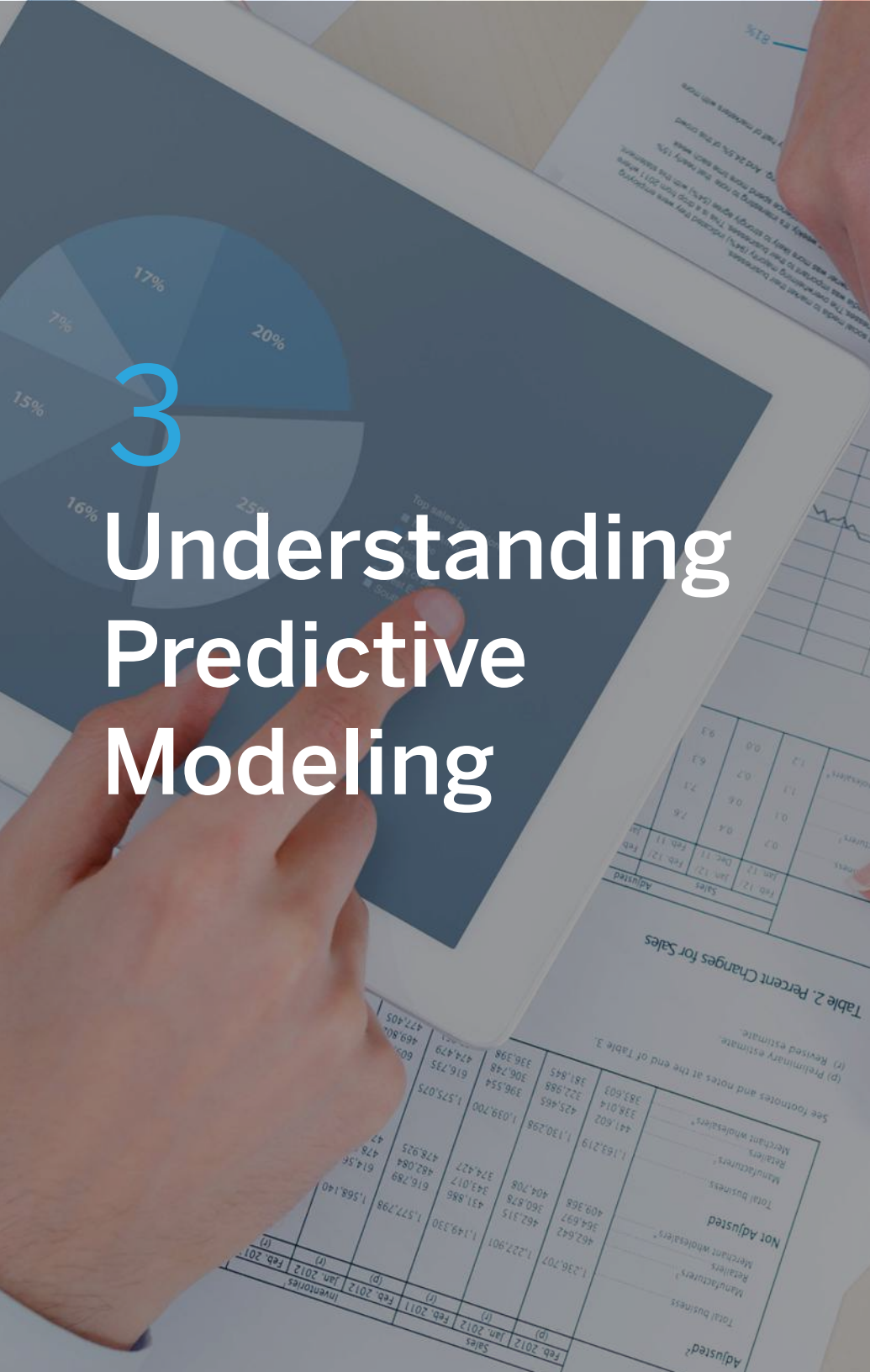
When embarking on a predictive project you should ask yourself: “How important is it that I understand how the predictive model works?” In other words, how crucial is it for you to know why your predictive algorithm makes the decisions that it does? The answer will help determine the most appropriate algorithm and the best predictors to build into your analytical data set.

For example, in B2C organizations, predictive algorithms are often integrated into marketing and sales systems (like campaign or customer relationship management) that focus on cross-sell or up-sell campaigns. These systems use probabilities and business rules to optimize offers for customers and prospects, and often the customer has no idea that a predictive model is behind their personalized offer. In these situations, how important is it that the business ‘owner’ of the algorithm (sales person or marketing manager, for example) knows how or why the algorithm makes the decisions that it does?

Your response to the question about understanding the predictive model may be that you simply don’t care how the model works and just want the most predictive model possible. Or, you may feel it’s crucial for you to understand how the model works, and to identify the model’s key business drivers (for example, the most important variables, and why they are predictive).

3

Understanding Predictive Modeling



Wanting to understand how the model works is perfectly normal, however it may be challenging to explain, for two reasons:

- You may have difficulty gaining this kind of insight from the data you have available at the time of model building.
- The decisions made by predictive algorithms are based on correlation rather than causation.

Correlation Versus Causation

Algorithms use machine learning to identify and measure which predictors are *highly correlated* with the targets that you are trying to predict. In other words, they try to identify predictors that appear to have a strong relationship or connection with outcomes. This is not the same as identifying predictors as actually being the cause of specific outcomes. However, as mentioned earlier, humans have a tendency to want to make an inferential leap to *causation*.

Proving causation is especially difficult when what we are trying to predict is human behavior, because homo sapiens are flawed information processors. Psychological and behavioral economic research is full of examples of people behaving in ways that aren't in their own best interest. In psychology, we're taught the following axiomatic principle: The best predictor of future behavior is past behavior.

Inevitably, it is past behavioral predictors that are the most important fields in predictive models, and this is why your transactional and interactive data is so valuable even if it can't provide a 100% understanding of cause and effect.

Deployment: Knowledge Versus Embedded in a Process

As a final point on predictive models, it's worth mentioning that the best return on investment often comes when models are deployed directly into operational systems. Classification models, for example, generate probabilities which, when combined with rules, are used to make decisions. The previous examples of algorithms that help to personalize marketing offers are a good example of this. The consumers of the probabilities are your operational systems. Once the models are up and running and driving solid ROI for your business, your interest in understanding the models will likely move to the back burner.

Some Practical Tips

- Get guidance on what is achievable, given what you are trying to predict and the data domains that are available.
- "All models are wrong—some are useful," said George Box, an internationally renowned statistician. A model is simply a mathematical representation of the way the world works, and there is always going to be error. Ultimately this means you will need to have a certain tolerance for ambiguity.
- Predictive models don't override the natural laws of science (like gravity), but they can supplement our understanding of how people/objects/systems function in the real world under different conditions and constraints.
- Causation is great, but correlation ain't half bad!

Success with predictive analytics relies on more than just designing models and training algorithms. It also requires a focus on the bigger picture challenges of demonstrating benefit and ensuring integration across the organization. The following points will help you to get these aspects right:

1. Be Sure to Measure Business Value and Return on Investment

A predictive analytics solution must make a measurable impact. If not, the solution won't get noticed, never mind be appreciated. This holds even truer if the ROI can't be realized as a significant opportunity to drive business growth or new market opportunities.

Take the example of a marketing campaign. The ROI is in having the intelligence to target the customers who are likely (if persuaded) to buy your product rather than finding customers who would have bought the product without any marketing required.

A predictive analytics solution will be short-lived if it creates a "wow" effect, but nothing else.

Your #predictive #analytics solution must generate recurrent value, revenue, and business opportunities. #SAP



4 Important Considerations When Embarking on a Predictive Analytics Journey



An enterprise-wide vision for predictive analytics is crucial for proper discipline and governance.

2. Avoid Silos

Predictive analytics can drive value to all of your business functions, be it marketing, finance, human resources, and so on. However, enterprise functions often want to embed predictive analytics into their business workflow and embark on predictive analytics initiatives in silos. Though there is value in doing so, the results can be underwhelming.

This is because they're adopting various technologies, methodologies, and practices to address use cases that might exist – but without an enterprise-wide vision for predictive analytics. Therefore, walls rather than bridges are built between business functions.

The problem becomes self-perpetuating. With increasing adoption of predictive analytics solutions in various business units, the business as a whole finds it difficult to consolidate all the activities into a central initiative.

The solution is to create the vision and execute it across all functions – even if the pilot starts from one or two activities. The functions must agree that predictive analytics is an enterprise-wide mission. Leadership must

demonstrate belief in an analytics-driven business that is going to provide competitive advantage. In this way, predictive analytics becomes a true company asset.

3. Build Predictive Analytics into Your Business Strategy

Predictive analytics is not just another technology project. If considered to be a technology project, the business understands only the technical feasibility and not its business impact.

As mentioned, a predictive analytics initiative is the means by which a business gains a competitive advantage. It follows that outcomes provide the data to help make well-informed decisions. A lesser or confined approach is a step in the wrong direction. There is no ROI associated with technology-only thinking, because no tangible results are expected as an outcome. An initiative to embrace predictive analytics must be inseparable from your business strategy.

“The combination of some data and an aching desire for an answer does not ensure that a reasonable answer can be extracted from a given body of data.”

The American Statistician, Volume 40 (1), Sunset Salvo

4. Always Start with a Business Question

Big Data is not equal to high quality insight. A traditional business approach is to think, "We've captured huge amounts of data, but how do we make sense of it?" This is a wrong start. The right approach is to start with a business question in mind. That way, you can ask if the data that you have is sufficient enough to provide the answer.

These are several pieces of the puzzle that need to be put together for you to find meaningful, actionable insights from your data. This is, after all, the quest that you have embarked on.

If you have implemented Big Data technologies, predictive analytics is the next step towards achieving true ROI on your investment. It will help you gain actionable insights right across your business as long as you implement it with an enterprise-wide vision and have a clear understanding of how it fits with your existing analytics strategy. By streamlining operations, automating decision making, and uncovering new business opportunities, predictive analytics will help your business gain the competitive edge it is after.



Find out how SAP can help you optimize resources and improve business margins with predictive analytics.

This eBook is produced by SAP, and includes the insight and thought-leadership perspectives from a number of SAP's specialists in predictive analytics and data science:

Doug Freud is Associate Vice President, Data Science and Strategy. Doug is responsible for developing strategies and solutions around predictive analytics and Big Data for the SAP HANA Global Center of Excellence.

Richard Mooney is the Lead Product Manager with the Predictive Analytics team. He has spent 20 years in the software industry in a wide variety of roles including development, sales, marketing, and product management.

Ashish C. Morzaria is Director of Advanced Analytics at SAP and currently drives the vision and commercialization strategy for SAP BusinessObjects Predictive Analytics. He has been with SAP for twelve years in various product management roles across the entire analytics portfolio.

Pierre Leroux is Global Director of SAP BusinessObjects Predictive Analytics product marketing. His experience spans development, consulting, product management, and product marketing roles in small to very large companies.

Paul Pallath is the Chief Data Scientist and Director with the Advanced Analytics organization at SAP. With over 20 years of experience in machine learning, Paul has several research articles published in the fields of machine learning and data mining.



+

“Your Predictive Journey” eBook Contributors

The following are definitions of the key predictive analytics terms used in this eBook.

Predictive Analytics: Predictive analytics is software that uses the technique of machine learning to find patterns in large volumes of data, which in turn enable predictions to be made about the future.

Business Intelligence (BI): BI software enables the acquisition and transformation of raw data into meaningful and useful information for business analysis purposes. BI tools make heavy use of visualization techniques and rely on human interaction to drill, slice, and dice the data and infer correlation between the different factors (historical data) and the outcome (eg sales).

Data Science: Data science is an interdisciplinary field about the processes and systems used to extract knowledge or insights from data in various forms, either structured or unstructured.

Algorithm: An algorithm is a step-by-step procedure, like a map or flowchart, designed to perform an operation, which will lead to the sought result if followed correctly. Algorithms are suitable for solving structured problems (i.e. amenable to sequential analysis) but are unsuitable for problems where value judgments are required. Predictive analytics uses business algorithms to provide guidance or insight derived from data.

Machine Learning: The field of machine learning is concerned with the question of how to construct computer programs that automatically improve with experience. Predictive analytics uses machine learning to teach programs to look at a particular outcome and then uncover the factors behind it (which could include thousands of possible causes and nonlinear relationships).



+ Predictive Analytics Glossary of Terms

A close-up photograph of a man with dark, wavy hair and a beard, smiling warmly. He is wearing a dark grey suit jacket over a light blue patterned shirt. He is looking down at a smartwatch on his left wrist, which is being held by another person's hand. The background is softly blurred, showing what appears to be an office or meeting room with wooden beams and windows.

© 2017 SAP SE or an SAP affiliate company. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of SAP SE or an SAP affiliate company.

SAP and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP SE (or an SAP affiliate company) in Germany and other countries. Please see <http://www.sap.com/corporate-en/legal/copyright/index.epx#trademark> for additional trademark information and notices. Some software products marketed by SAP SE and its distributors contain proprietary software components of other software vendors.

National product specifications may vary.

These materials are provided by SAP SE or an SAP affiliate company for informational purposes only, without representation or warranty of any kind, and SAP SE or its affiliated companies shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP SE or SAP affiliate company products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

In particular, SAP SE or its affiliated companies have no obligation to pursue any course of business outlined in this document or any related presentation, or to develop or release any functionality mentioned therein. This document, or any related presentation, and SAP SE's or its affiliated companies' strategy and possible future developments, products, and/or platform directions and functionality are all subject to change and may be changed by SAP SE or its affiliated companies at any time for any reason without notice. The information in this document is not a commitment, promise, or legal obligation to deliver any material, code, or functionality. All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements, which speak only as of their dates, and they should not be relied upon in making purchasing decisions.