ima®

The Association of Accountants and Financial Professionals in Business

Statement on Management Accounting

OVERCOMING FP&A'S BIGGEST CHALLENGE: PREDICTING THE FUTURE

TABLE OF CONTENTS

ABOUT THE AUTHORS

Kip Krumwiede, Ph.D., CMA, CSCA, CPA, is the director of research for IMA. Prior to joining IMA, Kip spent 18 years as a management accounting professor and worked for two *Fortune* 500 companies in a variety of management accounting-related positions. As director of research at IMA, Kip is responsible for developing and producing research studies on a variety of accounting and business topics, especially those related to financial planning and analysis.

Lawrence Serven is an internationally recognized authority on enterprise performance management (EPM). He is the founder and former CEO of XLerant, Inc., a leading EPM software company. He is also the author of *Value Planning: The New Approach to Building Value Every Day*. Lawrence has also contributed articles to *Harvard Management Update, Journal of Corporate Accounting & Finance, CFO* magazine, *Strategic Finance, BPM* magazine, and other industry-leading publications.

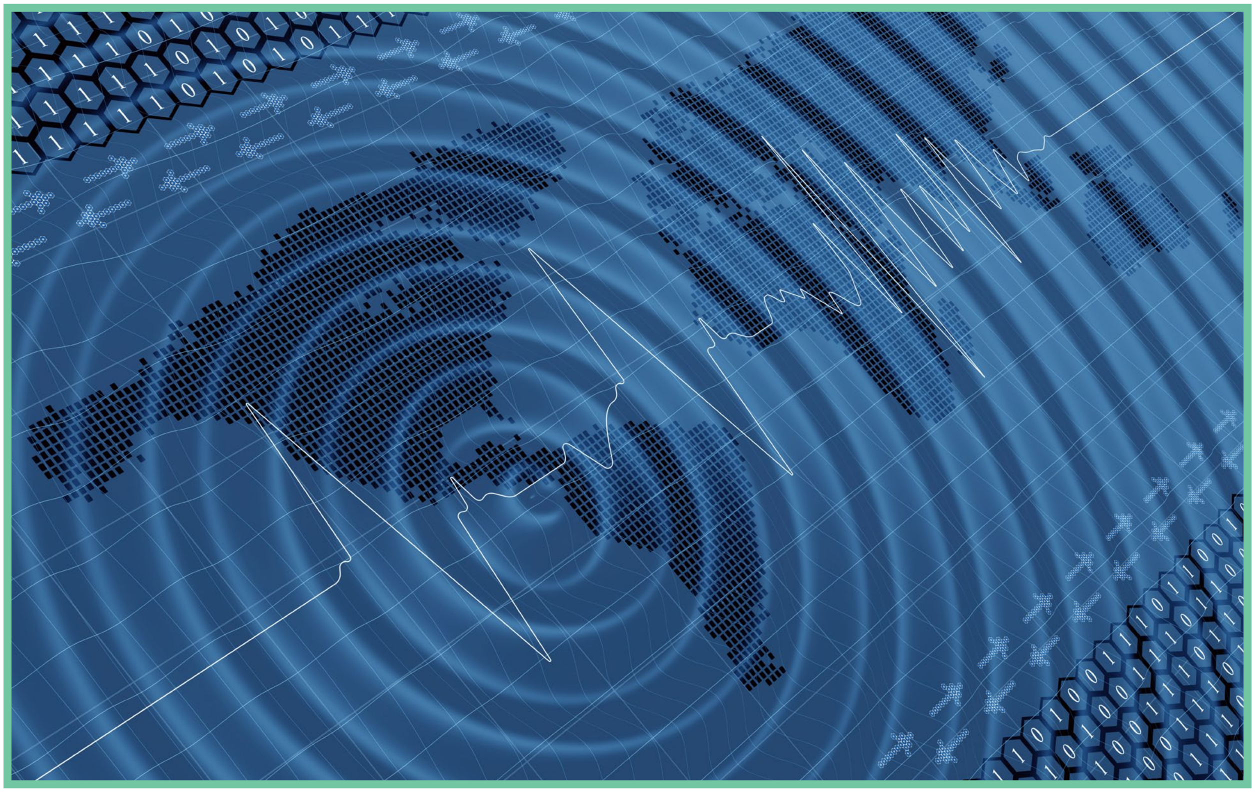
Robert Liou, CMA, CPA, is a former research manager for IMA in the China office. Prior to that, Robert was an upstream research and consulting manager for IHS Markit. He also spent five years in public accounting, most of it with a Big 4 accounting firm. Robert is a graduate of UCLA (economics) and the University of Michigan (MBA).

INTRODUCTION

Like a major earthquake, the COVID-19 pandemic has changed the landscape of the business world, and the paths forward are no longer clear. Among the many impacts of COVID-19, financial planning and analysis (FP&A) became even more challenging. Previous assumptions and trends may no longer make sense. The pandemic has accentuated prior weaknesses in FP&A processes. On the other hand, FP&A done well can be like a GPS guiding a business toward where it wants to go.

Like a major earthquake, the COVID-19 pandemic has changed the landscape of the business world, and the paths forward are no longer clear.

We take a broad view of FP&A, including financial and operational planning, key performance indicators (KPIs), variance analysis, planning and budgeting, forecasting, and financial modeling. It truly is "extended planning and analysis," or "xP&A," as Gartner often refers to it. This report is designed to serve as a guide for an organization trying to see what is ahead in uncertain markets. We provide several specific ways that all companies can use predictive analytics to improve forecasting and be able to make course corrections more quickly. The report provides the basics for using predictive analytics and shows that it does not have to be super complex or require expensive software—at least at first. The key is using it to decrease uncertainty about the future without spending more than its expected value. A future report will show how to perform more advanced predictive analytics including developing Monte Carlo simulations. •

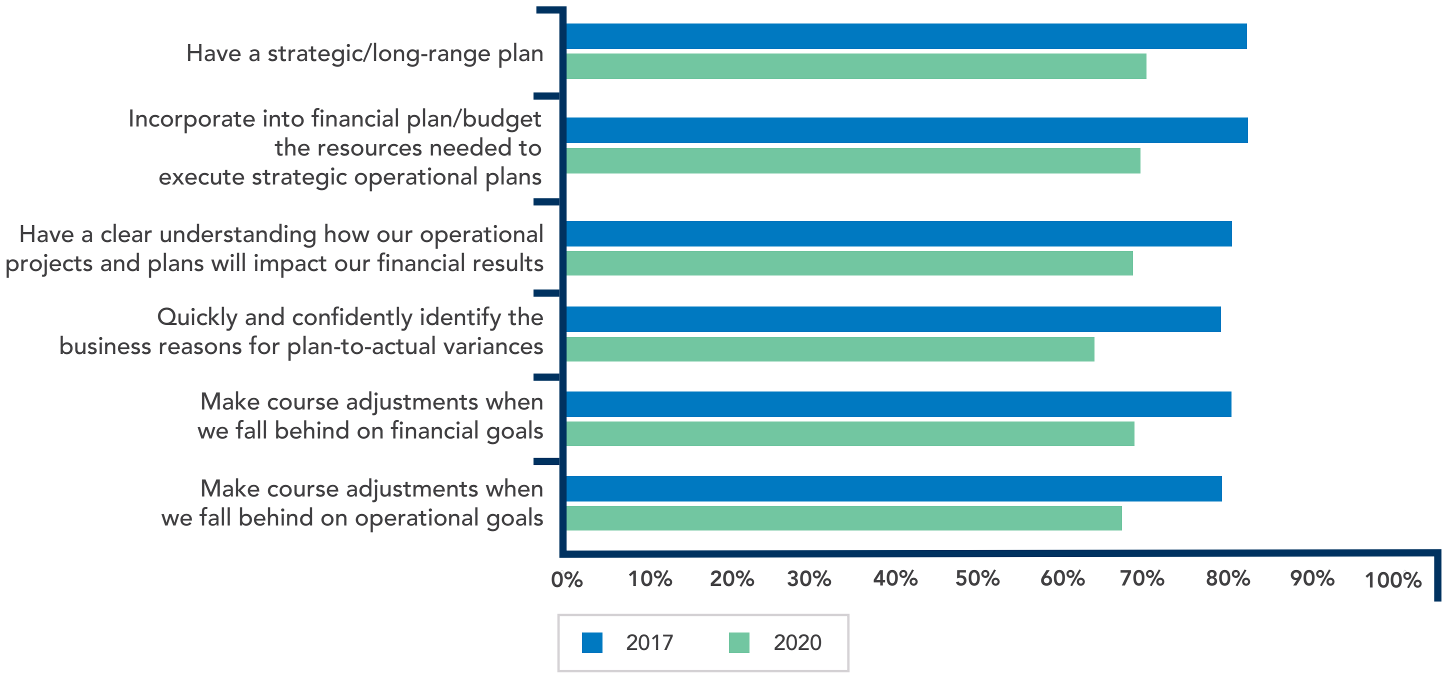


Impact of COVID-19 on FP&A

Across the world, the COVID-19 pandemic has had a devastating impact on businesses of all sizes. Some companies have fared better than others, but almost all companies have been impacted in some way. Before the pandemic, most companies already knew they needed to improve their FP&A processes. Then came the pandemic, rendering most forecasts obsolete and making other FP&A processes even less effective. In a recent survey of 245 finance and accounting professionals, the finance team's most often cited top priorities are cash forecasting and management, cost management and control, and scenario modeling.[1] Another recent survey by IMA® (Institute of Management Accountants) found that the largest increase in emphasis for finance-related areas is in risk management and cash forecasting/management. Less time is being spent on monthly close and business partnering/decision support.[2]

Four years ago, we conducted a survey of IMA members regarding their FP&A practices to determine what the best-performing companies do differently in FP&A. In 2020, we surveyed IMA members again, asking many of the same questions as in the 2017 study plus new ones relating to the impact of COVID-19 and how the more successful companies have changed and improved their FP&A processes. Figure 1 compares responses in 2020 vs. those in 2017 for several FP&A best practices. The impact of the pandemic on business is apparent as respondents were less likely to agree that their company has a strategic long-range plan, clearly understand how operational projects and plans will impact financial results, identify real business reasons for plan-to-actual variances, and make course adjustments when they do not hit their financial and operational goals.

FIGURE 1: COMPARING FP&A KEY PRACTICES 2020 VS. 2017 (% AGREE)



[1] See Appendix 1 for demographics and other results of the study.
[2] Raef Lawson, *The Impact of COVID-19 on the Finance Function*, IMA, January 2021, **bit.ly/3dWbfXp**.

A controller from a small healthcare provider lamented: "Forecasting is hard, and we don't always get everything right. I think we need to work on being okay with variances and incorrect forecasts as we spend a lot of time doing drafts and revisions on forecasts." Clearly, there is an increased need for better FP&A practices and especially forecasting.

Strong FP&A Can Impact the Bottom Line

We gauged the impact of strong FP&A practices by comparing the survey results of respondents working in best-, middle-, and worst-performing organizations. To identify best-performing organizations, we chose organizations that reported they (1) consistently meet or exceed the targets they set for themselves as a company and (2) consistently meet or exceed the results of their competitors. It is important to note that both criteria must be met for "best-performing," not just one or the other. Admittedly, this method of classification has limitations. But these criteria are good indicators of performance in relation to the quality of FP&A as well as the ultimate financial outcomes most companies are trying to achieve. The results of the research study are contained in Appendix 1.

The results consistently show that the best-performing companies tend to have stronger FP&A systems than lower-performing groups (see Table A1 in Appendix 1). We asked respondents to rate on a scale from 0 to 100 to what degree their organization realizes the full potential of its FP&A. The average response for those in the best-performing group is 69, compared to 59 and 53 for the middle- and worst-performing groups, respectively.

How can strong FP&A practices add value? Some consultants have advocated for "lowering the cost of FP&A as a percent of revenue." But that underestimates the potential value that effective FP&A can deliver. Of course, if there is no potential value, then the ultimate goal would be to simply eliminate FP&A altogether and bring the cost to zero percent of revenue. Instinctively, we know that approach is wrong. Our research found that respondents from best-performing companies that rated their overall FP&A higher had 21% average revenue growth over the past five years, compared to just 12% and 4% for the middle and worst performers, respectively. More specifically, the following benefits are types of value that better FP&A can provide:

- Driving specific initiatives, projects, and programs to execute the strategy.
- Ensuring optimal allocation of resources and coordination of initiatives, projects, and programs.
- Building organizational awareness of the strategy and each department's role in achieving it.
- Providing the mechanisms to ensure the financial and operational goals of the organization are achieved.

Bottom line: Strong FP&A systems have a positive impact on financial performance.

Hard to Predict the Future Right Now

Our study found that the top two FP&A challenges faced now by far are predicting the future and doing most of the FP&A in Excel.[3] Doing most of FP&A in Excel may not necessarily be a challenge for a small company if Excel is working well enough for them. But, in general, dedicated FP&A software can facilitate predictive analytics. Other common challenges include being too busy with financial reporting to do planning and analysis, being disconnected with strategy (or there is no real strategy), insufficient data, people working remotely, conflicting goals, and lack of analytical skills. In this report, we focus on the challenge of predicting the future. Future reports will focus on measuring and modeling what matters and speeding up the pace of FP&A. •

[3] See Figure A1 in Appendix 1.

Predictive **Analytics**

Easily the biggest challenge for the majority of survey respondents, predicting future revenues and cash flows is more difficult now. Typical comments regarding the reasons for this include "constant change" and "economy and the industry give mixed signals." Many industries, such as airlines, hotels, and countless restaurants, have been hit especially hard, seeing drastically lower revenues or, worse, closure with no revenues. Of course, there were industries thriving during the pandemic such as delivery services, media streaming, tech companies, and home improvement. But even with that success, there is a new challenge to be able to forecast what is ahead *and* meet demand. Will that success continue after the pandemic? If so, to what degree? How can we prepare for different scenarios?

Predictive analytics is the use of data, statistical algorithms, and machine learning techniques to identify the likelihood of future outcomes based on any relevant data. The goal is to go beyond knowing what has happened to decrease uncertainty about the future and associated risk.[4] Predictive analytics can be used to get clarity on what the future can look like. In its simplest form, it is the science (and a little art) of taking historical data and using it to project future results.

For example, a pharmaceutical company has relied on predictive analytics to forecast its sales and revenue using what is referred to as a "competition" of more than 50 algorithms, or models, based on data from the past 24 months. A model includes a set of potentially causal KPIs such as research and development (R&D) expense per new drug developed, number of new drugs in the development stage, number of new drugs in the regulatory review stage, pharmacy sales, time to market, etc. It runs the first model, using the first 18 months to "project" the remaining six months. It compares the projected six months the model produced with the actual results of that six-month period and calculates the variance. Next, the company does the same thing until all 50-plus models have been run and calculates the corresponding variances. The model with the lowest variance of prediction to actual "wins" the competition, and that model is then used to run sales forecast predictions.

Fast-forward to 2021. For many companies, at least some data from the last 24 months has probably lost some predictive value due to the disruption caused by the COVID-19 pandemic. How do you perform predictive analytics when past data are no longer valid for predicting the future?

Goal Is to Reduce Uncertainty

The concept of uncertainty underlies risk management and FP&A concepts. Risk can be defined as uncertainty where at least some of the possibilities involve a loss, catastrophe, or other undesirable outcome.[5] Thus, the degree of uncertainty of each aspect of the business must be considered and may determine the expected variances between actual and budget. Expected profitability for relatively stable industries, such as equipment leasing and universities, is more predictable than for other more dynamic industries, like technology and construction. Thus, there is a greater need for strong FP&A wherever there is greater uncertainty.

Keep in mind that the ultimate goal of forecasting is to reduce uncertainty and associated risk. Risks include not having enough capacity to meet demand, missing market opportunities altogether, overinvesting in new assets such as people and equipment, and so forth. Thus, any information that can be potentially useful to reduce uncertainty can add value to predictive models. In today's Internet of Things environment, there is a plethora of all types of information out there that can have predictive value. It does

[4] "Predictive Analytics: What it is and why it matters," SAS Insights, **bit.ly/2R4kkEu**.

[5] Douglas W. Hubbard, *How to Measure Anything: Finding the Value of "Intangibles" in Business*, Third Edition, Wiley, Hoboken, N.J., 2014.

not have to be perfect, or even complex, but it does need to provide some indication of what is going to happen. To be of value, it also needs to be acted upon as quickly as possible for a company to be agile. Who better than the finance team and FP&A to manage this process? And it does not necessarily require a huge amount of data to reduce uncertainty quickly. There are statistical shortcuts that can reduce uncertainty based on small samples.

First, there is the "Rule of Five," which says there is a 93.75% chance that the *median* of any population is between the smallest and largest values in a random sample of five. In his book *How to Measure Anything*, Douglas W. Hubbard provides a "mathless table" that makes estimating a 90% confidence interval of a population median quite easy. Table 1 expands on the Rule of Five. If you make eight random observations of something, the second largest and second smallest values of those observations provide a confidence interval of 93%. If you make 16 random observations, the fifth largest and fifth smallest values provide a confidence interval that exceeds 92%.[6]

TABLE 1: MATHLESS TABLE—USING THE RULE OF FIVE TO ESTIMATE THE MEDIAN OF A POPULATION

Lower bound = *n*th smallest sample; upper bound = *n*th largest sample

Sample Size	Sample Value	Confidence
5	1st	93.8%
8	2nd	93.0%
11	3rd	93.5%
13	4th	90.8%
16	5th	92.3%
18	6th	90.4%
21	7th	92.2%
23	8th	90.7%
26	9th	92.4%
28	10th	91.3%
30	11th	90.1%

Second, another option when there is a small sample size is to use the *t*-statistic approach in Excel to estimate the likelihood of the answer. When the original population's standard deviation is unknown, the *t*-statistic approach can help us to estimate the mean of a population from a sampling distribution.

Third, Bayesian inference (based on Bayes theorem) is a method for updating prior knowledge about unobservable quantities given new or known facts. It uses probability to describe the uncertainty over what the values of the unknown quantities could be. For example, say you are playing a murder-mystery game and there are seven suspects who could potentially be guilty. As you receive more clues, the list of potential suspects narrows, and the probability shifts toward the remaining suspects. Bayesian inference allows you to make a reasonable approximation of a likely outcome as new data are acquired.

Appendix 2 discusses ways to use the Rule of Five, *t*-statistic, and Bayesian inference for basic predictive analytics.

Facilitating Predictive Analytics: Nine Ways to Improve Forecasting the Future

Effective predictive analytics needs good data, tools, model building, and skills. There are nine ways to facilitate effective predictive analytics.

1 EXPAND THE DATA AVAILABLE

Good predictive analytics requires that a wide range of data be available. Ideally, any information that might be useful to reduce uncertainty would be accessible so that it can be included in the

[6] Hubbard, 2014.

model. Besides expanding information available for the finance department, management accountants should expand their view of what data are important for forecasting purposes and may even need to identify and collect such data. They can help build a "data lake" holding a vast amount of data in its native format until it is needed for management purposes, like forecasting. The data lake should support a wide range of analytics. Table 2 provides various sources of today's Big Data within organizations.

TABLE 2: BUSINESS SOURCES OF NONFINANCIAL BIG DATA INFORMATION

Business Process	NFI—Sources of Big Data Within Organizations
Property, plant, and equipment	Online databases complementing historic value
Marketing	Social media, email, Google search, website analytics, and even health data from wristband devices and smartphones
Accounts receivable	Full textual description (unstructured data) of goods or services
Purchases and sales	Radio frequency identification (RFID), GPS, and Bluetooth beacon
Cash	Mobile payment, electronic credit, and Apple Pay or Android via near-field communications
Customer service	Email, social media, and call center records
Supply chain	RFID, GPS, security video (logistics center), sensors, and greenhouse gas (GHG) data
Inventory	RFID, GPS, and security video (stocking warehouse)

Source: Association of Chartered Certified Accountants and Chartered Accountants Australia and New Zealand, *Analytics in Finance and Accountancy*, September 2020, **bit.ly/3tYcwCN**.

Data integration is central to being able to combine data from various sources and making them equally available (a process known as data extract, transform, and load). A real example of this is a CFO of a software company who uses data from the customer success system when he is looking at late payments. Rather than just sending out collection letters for payments more than a certain number of days past due, he looks at the client's level of satisfaction (through Net Promoter Scores, how satisfied they are with their implementation of the software, etc.). He approaches a happy customer who is late on a payment differently from one who is not happy. That is a smarter approach and just one example of how integrating data (in this case, from customer success and accounts receivable) can lead to better decisions and more informed actions.

Traditionally, the financial system included only data for general ledger (GL) accounts needed for reporting purposes. Today, most of a company's data is not reported in the financial statements, which are based on stakeholders' expectations of future cash flow and other concepts of value. The potential types of data needed for predictive analytics may seem foreign to many management accountants. Some may even dismiss them as not reliable or relevant. But we live in a whole new data-intensive marketplace with vastly more ways data are collected with essentially unlimited storage space. Other examples of new types of data include carbon footprint, individual credit card transactions, Twitter and Facebook statistics, and Google Trends index data. Say you are trying to assess online ad marketing channels to determine which are most profitable and leading indicators of future demand. You could set up accounts (i.e., data fields) for credit card transactions from different channels to track the volume for each. Other types of data could relate to economic trends and seasonality or other time-based variables.

It is important to distinguish between different types of accounts or data fields. There are the GL accounts needed for recurring reporting needs and, to a limited degree, cursory decision support. But there should also be "planning" accounts for other types of data to support FP&A and other complex decision making. For example, GL accounts may include revenue accounts for different customer segments. But planning accounts might include accounts or fields for different online ad channels to track volume for each. A data lake with both types of data would enable predictive analytics for different customer segments' volume through different online ad channels. Of course, you would need an information system that enables GL and planning accounts to be matched in an automated way.[7]

A common question is: How can information such as social media likes and tweets be used in a predictive model? This is part of a larger common misconception that some things are perceived "immeasurable." In accounting, we frequently debate how intangible assets such as goodwill, brand value, and innovativeness can be valued. The underlying assumption in these debates is that these things are not measurable. Hubbard, in his book *How to Measure Anything*, dispels the idea that some things cannot be measured.[8] Hubbard boils down the arguments that some things are not measurable into three categories, then dispels each one.

1. **Concept of measurement:** The definition of measurement itself is widely misunderstood.
2. **Object of measurement:** The thing being measured is not well defined.
3. **Method of measurement:** Many procedures of empirical observation are misunderstood.

Concept refers to a misunderstanding of what measurement really means. Unlike what is typically implied by the term, measurement is not an exact point value. According to Hubbard, measurement is a "quantitatively expressed reduction in uncertainty based on observation." A common perception is that if it is not in a company's enterprise resource planning, financial, or other systems, then it cannot be measured. An example would be when someone says, "We can't measure that because there is no way to put an exact point on it" or "We can't measure that because it is so uncertain." Actually, the more uncertain something is, the easier it is to measure. When there is a lot of uncertainty, even a few data points can greatly reduce it. Contrary to a common assumption that if we have a lot of uncertainty, we need a lot of data, the fact is that if we know almost nothing, almost anything will tell us something.

If we expand our view of measurement as uncertainty reduction, then we would take stock of what we know already and look for simple ways to increase what we know. For example, a CFO may want to assess the quality of the company's website in terms of users finding information they are looking for. But because she lacks other data, she may go with information that is available, such as number of clicks or number of report downloads, which do not necessarily measure website quality. Instead, why not give a few newer customers a short list of items to find and measure how long it takes to find them?

Object refers to not understanding the thing you are trying to measure. An "intangible" like goodwill or brand value may seem fuzzy and unmeasurable. The key is to think more specifically what more or less of it would look like. For example, say a company would like to evaluate its website relating to the "member experience." There are many factors contributing to the member experience on the website and the consensus is that you cannot really measure those things, so it defaults to page views and occasional focus groups. The object of measurement or its purpose has not been well defined. A better way is to identify what we would observe if the website were more effective. There might be

[7] For an excellent guide to automating the finance function, see the IMA SMA *Transforming the Finance Function with RPA* by Loreal Jiles, July 2020 **bit.ly/3bWWQcn**.
[8] Hubbard, 2014.

fewer customer emails or phone calls asking where they can find information or maybe fewer layers of website content leading to fewer clicks to get the information sought. Another question to consider is: What is the purpose of the measurement? For example, the company may be considering two potential upgrades and wants to determine which of the two would make it easier for customers to find information.

Method refers to not knowing how to measure something. Measurement does not necessarily require a large sample size. Again, the goal of measurement is to reduce uncertainty. Methods of measuring website effectiveness often get bogged down because there are so many things affecting its use. There may be a very large portfolio of products and services offered, and users come to a website for many different reasons. So which reasons do you focus on? The company could start simply by using the Rule of Five by picking a random pool of five unique users and measuring the number of clicks or time spent finding information.

So, in addition to "anything can be measured," we can add "anything can be modeled." And if the goal is to predict future sales revenue, it must link somehow to sales revenue. Here are some examples:

- Google Trends can be used to find the relative popularity of search queries related to different music genres as a way to measure the "buzz" and help predict a music company's royalty splits. (But beware of manipulated trends!)
- Number of vaccinations can be used to predict Uber and Lyft rideshare volume.
- Uber and Lyft rider volumes can be used to measure people's desire to travel, go to restaurants, go to sporting events, etc. The more people desire to travel, the greater airline passenger volume and revenues.

2 TOOLS: START SMALL AND ADD ON

Our study showed that well over 40% of the FP&A survey respondents said one of their biggest challenges is that they are still doing most of FP&A in Excel (consistent with all other surveys in this area). Of course, the upside for spreadsheets is that they are cheap, easy to use, and quite flexible. On the other hand, they quickly get unwieldy and oftentimes can be understood only by the person who created them and lack any type of workflow management (and instead rely heavily on emailing templates). It is also easy to miss a cell or reference that needs to be updated. If you want to "dip a toe in the water," start with a simple predictive model on a spreadsheet to provide the flexibility needed for the particular needs of the business. But once the simple model is validated and makes sense to everyone, it is probably time to migrate to a dedicated FP&A software program(s) to automate the update and reporting process. This is a good candidate for the "agile and scrum-based value delivery" process approach to achieve progress sooner.[9] This is essential to be able to continuously monitor results and quickly identify business reasons for variances, as we will discuss later in this report.

Basic models do not necessarily need to be run on Excel. Open-source programming languages like R and Python can be used for data analytics. R is built by statisticians and uses their specific language. Python is a general-purpose language with readable syntax and is used by many companies today. One of the big edges that programs like R and Python have over spreadsheets is that they can run iterative models very easily and quickly. Ten thousand iterations can be simulated in a flash, something Excel cannot do quickly. Large generative models allow R and Python to run Bayes theorem (for which this is a requirement), for example.

On the other end of the spectrum are enterprise packages, also known as enterprise performance management (EPM) software. This has probably been the most common migration destination and

[9] See Loreal Jiles, "Agile and Scrum-Based Value Delivery," *Strategic Finance*, January 2021, pp. 60-61, **bit.ly/3t0D70P**.

provides the enormous benefit of integrating much of the company's data. A potential downside of EPM packages is the amount of time and expense it takes to implement these systems and new modules. And the software may still not have the functionality needed—or it can be difficult to use—for the business's particular predictive analytics needs. For these reasons, many companies have not yet taken advantage of the potential value of these integrated systems and instead look for dedicated business planning software to augment other software tools.

Somewhere in between spreadsheets and EPM are dedicated business planning packages. There are many to choose from these days, each with strengths and weaknesses for a company's particular predictive analytics and other FP&A needs. It is probably impossible for a management accountant—and perhaps anyone for that matter—to be conversant with all the vendor packages available. Yet it is incumbent to be educated about them.

Getting educated on what is available will most likely reshape your requirements list and help you create a vision for what you want and what the company needs. As an analogy, a colleague needed a new car. He made a list of all the features he needed or wanted. What was missing entirely on the list was self-parking assist—because he did not know about it until he went for a test drive and discovered how useful it could be (especially for him, a person who absolutely hates parallel parking). Self-parking assist went right up to the top of his list of requirements.

We can also apply a successful practice used by tech firms in Silicon Valley. They use two important terms, minimum viable product (MVP) and whole product concept (WPC), when launching a new venture. The MVP is a product with a minimal set of features and capabilities that make it viable in the marketplace. The WPC is the full vision the company has after all funding and product development is complete. These two concepts are important because resources are limited—you cannot do it all at once. A great example is Amazon. Its MVP was selling books over the internet, and it is now just beginning to fully realize its WPC. Founder Jeff Bezos said he had great clarity on both concepts when he launched the company.

Applying this to your technology choice is easy. You want to future-proof your choice. Based on your most pressing priorities, define your MVP. Maybe that is a budgeting application with an approval workflow, an executive dashboard, or something else. Take some time to think through (and document) what your WPC is. Maybe that is fully integrated business planning from demand planning through supply chain, head count, and selling, general, and administrative budgeting, even long-term strategic planning. The right choice(s) depends on many factors, which is another reason why getting yourself educated on what features are available can be so helpful.

Think about what types of decisions your business needs to make and what information would help make those decisions. Then try to make sure that the package(s) chosen meets your anticipated analytics needs, including either native predictive analytical capability or the ability to interface with statistical software. In addition, you may want to be able to easily answer queries, mine large structured and unstructured data sets, enable complex predictive analytics models depending on the company's size and needs, run Monte Carlo simulations, and interpret and visualize the results.

You may be thinking that these models would have to be massive to capture all the key drivers in an organization, and many are. You may also be thinking of your already massive Excel workbook, and the thought of adding more types of data along with how different drivers are related makes you start to think about retiring. The good news is that predictive models should start small to include the most significant variables and approximate relationships among the variables. Once you establish a base model that you are satisfied includes all the major variables, you can start adding variables to increase its accuracy. And when the model gets sufficiently complex to make updating key assumptions cumbersome, it is probably time to find one of many FP&A software solutions available. More on that later.

Do not try to build the model that includes everything about your company. First, that is impossible. Second, it is not worthwhile. The key is to think long-term accuracy over short-term precision. By design, building predictive models is meant to reduce uncertainty about the future. Because the future is uncertain, it is not worth trying to accurately model relatively low-impact drivers.

Some of the more well-known software vendor packages dedicated to FP&A and predictive analytics include Adaptive, Board, Oracle, Python, R, Structured Query Language (SQL), Tableau, Power BI, and Anaplan, as well as many other cloud-based applications. And the list continues to grow. Each has its own strengths and limitations. For example, a recent entrant is Jirav, which aims to be an all-in-one financial forecasting package directed at small businesses with limited human or financial resources available for predictive analytics and data visualization.

3 USE SCENARIO PLANNING

Scenario planning (also known as scenario analysis) is used to stimulate the imagination and creative thinking process to better prepare for potential future scenarios and to possibly move the industry or customers in a desired direction.[10] It can be especially useful at this time of market uncertainty, both to identify opportunities, as well as potential risks, and to better prepare for them. It involves conducting research on forces that could have a big impact on the success of the company; focusing on a small number of potential scenarios; articulating each of the scenarios, including both direct and indirect effects; and developing options regarding how to succeed in each of these scenarios. Scenario planning is a good exercise to do as part of predictive analytics to model for a range of possible outcomes and estimate the likelihood of each outcome.

Scenario planning is more than just listing several potential scenarios. In fact, it is not even possible to list all of the potential scenarios, let alone predict them. Instead, it is meant to identify potential scenarios circling around one or two key focal issues, driving forces, or critical uncertainties. These might be the issues causing the biggest risks to the organization (e.g., cyber threat or customers not coming back) or a key driver of success (e.g., oil price on June 1, online traffic in May, or easing of COVID-19 restrictions).

Once these critical issues/uncertainties are identified, develop a scenario framework 2x2 matrix made up of the two issues/uncertainties with high vs. low spectrums. This leads to identifying four potential scenarios to address and model. Before doing any probability analysis, however, compose a narrative for each scenario as if it had already happened and then address the implications for the business. Include any early warning signals (i.e., leading indicators) suggesting which scenario(s) is (are) most likely. With that basis in mind, develop and model the scenarios.

Accounting textbooks typically define scenario analysis as performing three scenarios—best case, worst case, and most likely—or perhaps setting up the most likely outcome and then using sensitivity analysis by tweaking certain variables to determine the impact. But that simplified approach ignores the articulation of different scenarios. There may be a "best-case" scenario for one critical uncertainty but a "worst-case" for another. Further, it only computes simple financial outcomes void of the business reasons behind them and how to respond. Finally, it is a deterministic approach that ignores the level of uncertainty for each outcome. Instead, build a model that can include probabilities and distribution of outcomes for each of the critical variables and other business drivers that might be affected.

[10] For more on scenario planning, see IMA's SMA *Strategic Analysis—Methods for Achieving Superior and Sustainable Performance* by Mark L. Frigo and Kip Krumwiede, April 2020, **bit.ly/3heMS5I**.

For example, resorts, fairs, and beach towns have struggled with the uncertainty about pandemic restrictions and people's desire to travel.[11] Fireworks operators have been forced to provide clients with guarantees that if they are forced to cancel an event, even at the last minute, due to COVID-19 restrictions, clients will get their money back. But that causes some big critical uncertainties for the fireworks operators. It takes up to a year to prepare for a big show. Will there be high cancellations or low cancellations? And if the scheduled shows do happen, will there be high crowd turnout or low crowd turnout, a leading indicator for the number of future events? Scenario planning can help by establishing different scenarios based on the most critical unknown variables that may occur and then estimating the likelihood of the different scenarios. In this example, the fireworks operator might envision four scenarios using a 2x2 matrix as follows:

		Cancellations	
		High (negative)	Low (positive)
Crowd Turnout	Low (negative)	Big trouble!	Be patient
	High (positive)	Super-spreader events	Best case–gear up!

The planning team gives each scenario a name and composes a narrative as if it had already happened. This helps to articulate and estimate the probability of each scenario that can then be put into the model.

But scenario planning can be so much more than explaining how each scenario can occur. At a recent MIT Sloan virtual CFO conference, John Murphy, CFO of The Coca-Cola Company, described his view of scenario planning:

> *I've become less and less of a fan of the term "scenario planning" in the last nine months. I've discovered, at least in our case, it tends to be an exercise in sounding intelligent about the future. Then you get flung into the world of COVID, and it becomes a much realer topic you have to manage. I think of it more as scenario management than scenario planning. And, in that regard, we have very much sharpened our approach to having Plan B, Plan C, and Plan D available and implementable fairly quickly.*

Instead of asking: Here is a scenario; how do we explain it? Murphy recommends asking: "Here is a scenario; how are we going to manage it? These are the potential threats, and here are the decisions we will need to make by when; here are the implications for how we will allocate resources."[12]

4 ADDRESS THE KNOWING-DOING GAP

There should be a strong connection between predictive analytics, competitive strategy, and operational execution. You may want to predict customer demand using sophisticated modeling techniques, but these predictions should be evaluated and understood in light of the planned strategic actions the company is taking to grow sales. Examples of these planned actions include investing in

[11] Charles Passy, "Summer Planning Starts Now for Covid-Hit Businesses, and Forecasts Are Murky," *The Wall Street Journal*, March 5, 2021, **on.wsj.com/2QvCFe3**.

[12] Jane Thier, "Coca-Cola CFO prefers 'scenario management' to 'scenario planning,'" CFO Dive, December 1, 2020, **bit.ly/3xqsnwl**.

R&D to launch compelling new products, managing the life cycle around end-of-life for certain products, reprioritizing certain products in the portfolio and the marketing investment behind them, and doubling the number of sales associates. There should be performance measures linked to the planned actions and built into the predictive analytics model. Predictive analytics can and should assume the planned actions will occur.

What a company does or fails to do will have an impact on the business and might be the real explanation for why a predictive model "was wrong." In fact, a common reason why the future does not turn out as predicted is a failure to execute plans. Try not to overlook this point, because it may be at the heart of the variances your organization experiences. Not considering this possibility can lead to "chasing rainbows" and fine-tuning models that will continue to fail us.

The knowing-doing gap is simply the gap between saying we are going to do something and actually doing it. How do you know if a knowing-doing gap is undermining your ability to deliver valuable predictive models? If your organization consistently fails to meet its goals, that alone is a strong indication you may have a knowing-doing gap.

But let us break that down further. First, does your company measure the key metrics that drive the business? Is there an explicit understanding of how those measures impact the financial results of the company? Does your company set targets for improving those measures? Are there clearly defined initiatives or projects budgeted that are designed to achieve those targets? If the answer to those questions is "yes" but you consistently fail to meet your company goals, then you have a knowing-doing gap. You know what needs to get done; you just fail in the execution. If the answer is "no," then you have just discovered the root of your knowing-doing gap—not knowing what needs to be done.

The best answer is to combine predictive analytics thoughtfully by incorporating knowledge of the business and the company's operational plans. Monitor progress on those plans, too! (See No. 8, "Monitor results and quickly identify the business reasons behind variances.")

5 MODEL BUILDING: THINK CAUSALITY

To build a reliable predictive analytics model, think causality. Causality is a function of operations: resources, processes, and activities. A reliable predictive analytics model should be based on actual or expected causal relationships among resources, processes, customers, KPIs, external market factors, and other leading or lagging measures. The goal is long-term accuracy, not short-term precision. Remember to start small with the most critical causal factors and, once validated, add variables as warranted to improve predictive value. Using the "agile and scrum-based value delivery" process approach can be an efficient way to build a model.[13]

The process for developing a causal-based model is similar to developing a balanced scorecard. There should be a chain of cause-and-effect, leading and lagging indicators ultimately resulting in desired financial outcomes. In fact, companies that have constructed such a scorecard have essentially designed a predictive model. If designed well, it should represent the company's theory on what activities it needs to do really well on to ultimately be successful. Table 3 provides examples of leading and lagging factors that might be used in sales forecast models.

[13] See Jiles, 2021.

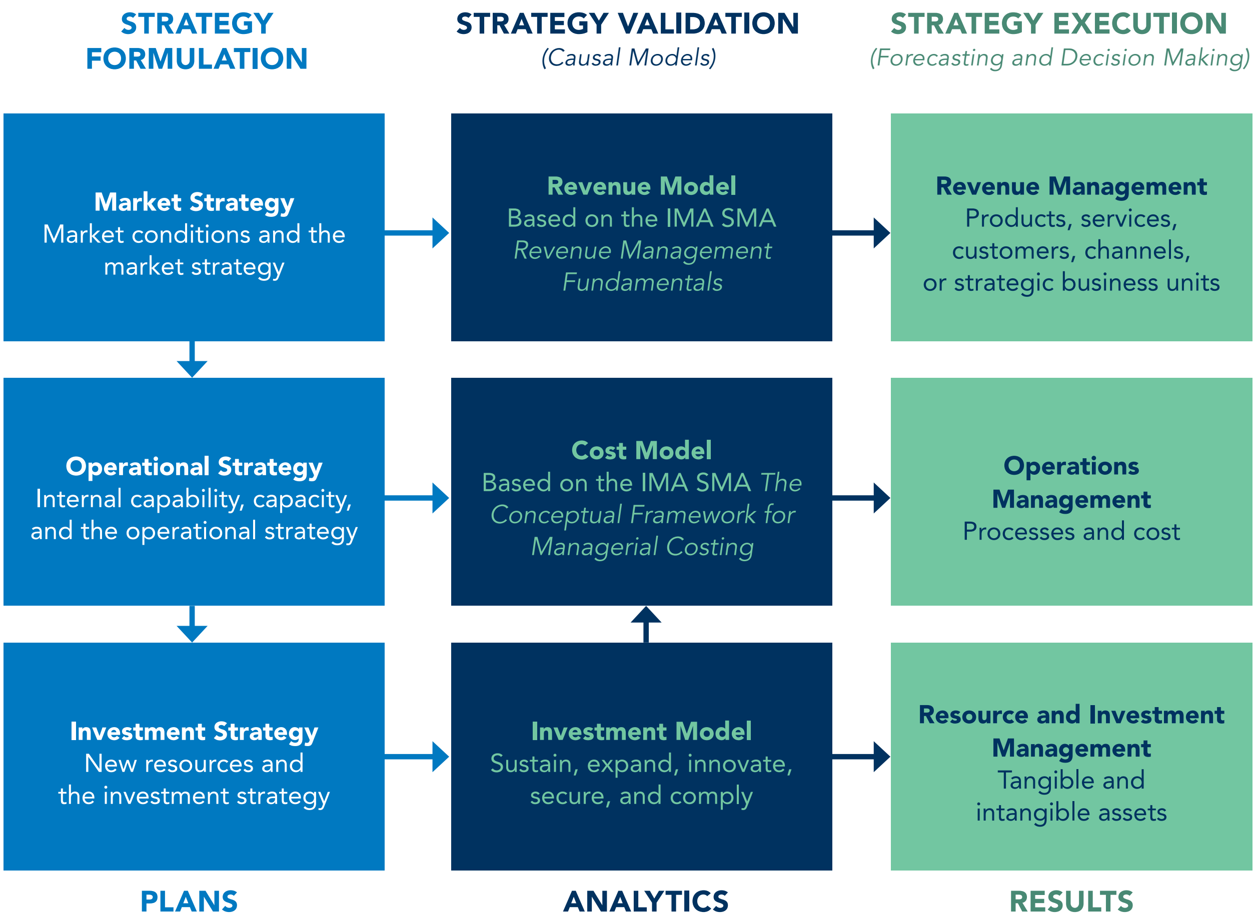
TABLE 3: EXAMPLES OF POTENTIAL CAUSAL RELATIONSHIPS IN A SALES FORECAST MODEL

Type of Company	Leading Factors	Lagging Factors
Large fast-food chain	• Price points • Product mix and duration • Unemployment rates • Local market saturation	• Traffic per store • Same-store revenues
Nonprofit professional association	• New student members • Number of virtual conferences and events	• Number of members • Membership dues
Start-up online retailer	• Price points, seasonality, market trends • Social media trends • Estimates from sales team • Prior month sales	• Total visits • Time per visit • Bounce rate
Rideshare (e.g., Uber, Lyft)	• Number of vaccinations • Hotel occupancy rates • City size	• Rideshare volume
Airline	• Uber and Lyft rider volumes • Time of year (season)	• Airline passenger volume
Health maintenance organization (e.g., Aetna, Cigna)	• Average caseload-to-staff ratio • Employer retention rate	• Employer retention rate • Market share percentage
Car dealership	• Customer acquisition cost • Number of trade-ins • New claims for unemployment	• Number of new customers
Advertising	• Facebook leads • Time it takes to close a deal • Average price of a deal • Duration of the client onboarding process • Average renewal rates, repeat business • Conversion rates at each stage of the sales process	• Number of new clients • Average revenue per client

To build a model, ask what key factors drive the majority of a desired outcome. IMA's Statement on Management Accounting (SMA) *The Profitability Analytics Framework* illustrates why and how models should go way beyond traditional budget models and take into account market, operations, and asset management factors.[14]

[14] Larry R. White, Monte R. Swain, Douglas T. Hicks, and Gary Cokins, *The Profitability Analytics Framework*, IMA, November 2020, **bit.ly/3xFn10r**.

FIGURE 2: PROFITABILITY ANALYTICS FRAMEWORK



Source: Larry R. White, Monte R. Swain, Douglas T. Hicks, and Gary Cokins, *The Profitability Analytics Framework*, IMA, November 2020, bit.ly/3xFn10r.

Internal data include a clear historical understanding of customer needs, customer buying and ordering patterns, customer paying patterns, and customer-driven costs for selling, distribution, inventory, financing, collections, and after-sale support. We need to understand where there is capacity and capability to expand, what constraints exist, where we have experienced success, and where failures have occurred. For example, revenues might drive certain costs, which may be variable or step variable, where costs increase in increments after a certain amount of additional volume is achieved. For example, say a beach hotel must add lifeguards and other staff as certain levels of tourist volume are achieved. Housekeeping wages may be increased more proportionally with tourist volume.

The market model and revenue model, based on the market strategy, describe the organization's structure and processes and campaigns to achieve market establishment and revenue goals. The operational model and cost model, based on the operational strategy, lay out the organization's structure, processes, and planned projects and initiatives to achieve operational excellence and cost goals. Key economic causality concepts help to build a model that effectively represents actual resource and cost flows in the organization.

At its core, a revenue model enables an organization to effectively assess its market and optimally segment customers, which is key to aligning resources and activities to deliver the right product at the right price at the right time to the right customer. The IMA SMA *Revenue Management Fundamentals* provides a detailed structure for revenue models that can be applied by management accountants to all types of organizations in all kinds of markets.[15] The revenue model is based on two key revenue management tools: demand management and resource management. Demand management involves aligning data and analysis to identify segmented pricing bases and allocate capacity (inventory) for optimal capture of the market value created by the organization. Resource management is focused on product configuration tailored to customer needs and duration control mechanisms that effectively manage timing of customer demand.

Revenue drivers, combined with cost drivers, become a powerful tool for modeling and analysis built on causality. For more information on designing managerial costing systems for internal decision support, refer to IMA's SMA *Developing an Effective Managerial Costing Model*.[16]

One way to determine key drivers is to start by asking straightforward questions about how your resources and processes function under various demand scenarios. Going back to the beach hotel example, you might ask questions like, "How many lifeguards do you strive to maintain based on the number of tourists?" or "How does housekeeping cost increase based on tourist volume?" Look for explanations such as, "We maintain at least three lifeguards when the beach is open and one more when we hit 50% occupancy, one more at 75% occupancy, and one more at 100% occupancy." And "We have a minimum of two housekeepers working eight-hour shifts, and then at about 50% occupancy, we increase housekeeping hours about one hour per every two rooms occupied." Of course, to be able to ask the right questions and understand the explanations, it is essential to understand the business model, who the customers are and where they come from, what value the company provides them, who does what, etc. If you do not know, ask questions until you do.

Other variables needed for a complete model include those related to external factors. What is the competition doing? What are our customers doing? What is happening in the global environment? Even though these factors are essentially beyond our control, if they can significantly impact our business, they must be included in some way. It was a lot to ask for before the pandemic, and in our post-COVID marketplace uncertainty, there may be totally new questions to ask. The business model has probably changed, and the data needed to succeed may have changed as well. New questions might include: How will COVID-19 restrictions impact our volume? How about social media and trade wars? What do we learn from call center data?

Once you have logical explanations and reasonable assumptions for various relationships, you are ready to start building (or adding to) a model. One risk to be aware of is "model creep," where the model (or accompanying spreadsheet) gets too complex to understand or maintain. We tend to have a complexity bias toward complicated rather than simple. The precision necessary depends on the complexity of the question(s) being asked, but if one formula approximates the results of several formulas, use the one formula. Remember: The goal in forecasting is long-term accuracy (and understandability) over short-term precision.

6 ESTABLISH DATA COLLECTION SYSTEMS

As described in the *Revenue Management Fundamentals* SMA, management accountants can establish data collection systems on market and sales data, including financial transaction data and nonfinancial customer/market engagement data. These data are foundational to analysis and

[15] Julie Harrison, Frederick Ng, Paul Rouse, and Monte R. Swain, *Revenue Management Fundamentals*, IMA, October 2020, **bit.ly/2R1fJmF**.
[16] IMA Managerial Costing Task Force, *Developing an Effective Managerial Costing Model*, IMA, April 2019, **bit.ly/2Xam519**.

modeling systems that provide key insights on revenue behavior causality and market demand patterns, stratified by customer segments. Combining causal market analysis with well-developed models of revenue and cost drivers will strengthen the accuracy and speed of scenario planning processes that accountants and managers use to achieve key strategic goals.

The IMA Management Accounting Competency Framework includes active participation in the data collection and management process.[17] Examples of these skills include:

- Integrating and consolidating information from multiple departments.
- Leading collaborative forecasting efforts by incorporating information from multiple internal and external expert sources and sophisticated modeling techniques.
- Transforming raw, unstructured data into a form more appropriate for analysis (e.g., data wrangling).

See Appendix 3 for more recommended competencies relating to data collection.

7 IMPROVE ASSUMPTIONS AND ESTIMATES

There is an old saying attributed to statistician George Box: "All models are wrong, but some are useful." No model can perfectly mimic the real world, and therefore models require simplifying assumptions as well as estimates about the future. Assumptions are especially important because they relate to actions by the market as well as internal resources (e.g., capacity), capabilities (e.g., productivity), and processes (e.g., quality). When a model provides results that seem counter to long-held beliefs, it is common to question the assumptions made. So, if you are going to develop a model for predictive forecasting, the assumptions and estimates should be reasonably accurate given the level of uncertainty at the time. Of course, as discussed earlier, we can decrease the uncertainty with more samples, even if it is only five.

Forecasts with faulty assumptions and estimates can do much damage. For example, a certain manufacturer's demand forecasts are typically off by 20% or more. The problem with this is that demand forecasts drive most other operational planning estimates. When those forecasts are consistently off significantly, they are not trusted and other departments go with their own estimates and make less than optimal decisions. Instead, the company could rely on artificial intelligence-assisted models, which may over time determine the systematic error and calibrate the sales forecasts for that systematic error. Adding to the challenge is that some markets are more stable than others, even pre- and post-pandemic. Calibrating estimates for stable markets is easier than for volatile markets.

So, how do you make estimates about a very uncertain global market? A common approach is for subject matter experts—like sales reps—to use industry reports to make estimates. But although they may be experts—or not—in a given field, making estimates is a very different matter. Hubbard warns to be wary of "expert estimates." They are statistically overconfident and highly inconsistent, and tend to have erroneous intuition about the math (i.e., different levels of numbers throw them off) and vary greatly in measured performance.[18] Unfortunately, companies rarely go back and check the accuracy of forecasts. It helps to ask the experts to provide their assumptions and early indicators to track the accuracy of their assumptions and estimates.

Fortunately, making estimations about the future can be calibrated. Here is a tried-and-tested calibration test approach provided by Hubbard when estimating upper and lower bounds (also known as confidence intervals, or CI) for a variable so that there is a 90% chance of the outcome falling within those bounds:

- ***Equivalent bet cheat sheet:*** Would you rather (A) win $1,000 if your CI contains the correct answer, or (B) spin a dial with a 90% chance to win $1,000? If (A), consider narrowing your range. If (B), consider widening your range. Ideally, you should be indifferent between the two.

[17] IMA Management Accounting Competency Framework, IMA, 2019, **bit.ly/3aP8K7w**.
[18] Hubbard, 2014.

- ***Do a premortem:*** Assume that a future project has failed or the worst case for some other critical factor. Next, work backward to determine what potentially could lead to the failure of the project or factor and reassess your estimate of success.
- ***Calibration process:*** A person can become calibrated for future estimates by making initial estimates, applying the equivalent bet test, doing a premortem, and making any range adjustments given previous feedback on calibration tests, i.e., are you typically 10% too high?

Making estimates for predictive models is really a test of assessing your uncertainty, whatever your level of knowledge. Although it may seem counterintuitive, it is desirable to start with a wide range for estimates as long as it realistically represents your uncertainty at that time. It is far better to see actual results within the estimated range within the model parameters than outside the range. If that variable is a key causal factor, the model would be essentially useless.

No matter how good your estimates and assumptions, a forecast model is still a plan. Reality is rarely, if ever, exactly the same as what was forecasted. But a forecast can also aid organizational learning. It is essentially a record of our assumptions at the time to learn what worked as expected, better than expected, or worse than expected. That is why it is essential to continuously monitor results and update assumptions as warranted.

8 MONITOR RESULTS AND QUICKLY IDENTIFY THE BUSINESS REASONS BEHIND VARIANCES

Successful planning and forecasting require a good understanding of the business, how operational plans should drive financial results, what signs to look for, and monitoring for progress of those plans toward financial results. In today's volatile world, it is critical to monitor results continuously and to quickly determine business reasons for any significant differences from the forecast. This ability is the heart of what it means to be an "agile" company. Here are some key practices followed by the best-run companies:[19]

- Connect operational planning with financial planning and use key performance measures to establish operational targets that will drive financial performance.
- Develop concrete initiatives to meet those goals, incorporate them explicitly in the budget/annual plan, and monitor performance along the way.
- Identify key assumptions and track whether they are being realized or not.
- Put in place early/key performance indicators to help identify the knowing-doing gap.
- Work quickly on building adaptive or mitigation plans when they fall behind on goals to get back on track.
- Encourage a culture of accountability—they hold people accountable for delivering both operational and financial goals. For example, Stryker is building these types of goals into team performance objectives and putting a governance structure in place to review and report on these goals.

Do not assume that a variance necessarily means there is something wrong with the model. Jennifer Wolfenbarger, VP and group CFO for Stryker's Global Quality and Operations Division, suggests to first assume the model is valid and look for the business reason for the variance. Remember the knowing-doing gap discussed earlier. Try hard to find the business reasons behind any variances before you take action. Remember Bayes theorem. We always have limited data, but we reduce uncertainty as we incrementally gain more data. If there is reasonably high probability that an estimate or assumption is not likely, and there is a business reason to support it, then it is time to revise the estimate. The beauty of predictive analytics coupled with continuous forecasting is that we can reasonably minimize the uncertainty about outcomes at any given time.

---

[19] For more on how best-run companies get more value from FP&A overall, see the IMA SMA *Key Principles of Effective Financial Planning and Analysis* by Lawrence Serven and Kip Krumwiede, July 2019, **bit.ly/2Y7frJP**.

Monitoring the results continuously implies that we can provide reports and visualizations that help clarify what the data are telling us. Wolfenbarger echoes the importance of data visualization: "My opinion is that it is even more important the way you visualize the information than the information itself, especially when you have a plethora of data and points to get across. The visualization can make the points extremely crisp with the right visual."

Ideally, the software tool should support the needed visualizations. Further, it is increasingly important for management accountants to have skills in data visualization. That means understanding how to best communicate results to the intended audience and choose the right type of visualizations. The following data visualization skills are especially important for forecasts:

- Communicate complex forecasts and budgets to others.
- Demonstrate an understanding of how to best communicate results with intermediate visualizations (e.g., histograms, area charts, and heat maps).
- Evaluate data visualization options and select the best approach for the intended audience.
- Demonstrate an understanding of how to best communicate results with advanced visualizations (e.g., Sankey plots, bubble charts, and network diagrams).
- Demonstrate expertise in all three aspects of data visualization: substantive, statistical, and artistic.

Appendix 3 provides a more complete list of skills and competencies helpful for data analytics and visualizations taken from the IMA Management Accounting Competency Framework. (See also IMA's upcoming SMA *Storytelling with Data Visualization*.) There should be someone in finance who has these skills and competencies. Do not rely on the IT department to provide the reports it thinks you need.

9 IMPROVE ANALYTICAL SKILLS

Strong predictive analytics skills can impact company performance. Our study found the best-performing companies' FP&A teams tend to have higher predictive analytics skills.[20] First and foremost, they are better at understanding the underlying business. Accountants who do not understand operations or the business or participate in planning with operational components of the business cannot possibly perform good predictive analytics or help influence performance. On the analytical side, they are better at drawing useful insights from the data and visually presenting and communicating data analysis results. Another recent IMA survey found that more than 80% of almost 1,500 global accounting and finance professionals have either improved or would like to improve their skills in scenario modeling/what-if analysis, cost management/control, business partnering/decision support, cash forecasting/management, and risk management.[21]

Where to Start?

As with any management improvement initiative, it is important to first decide to strengthen commitment to FP&A, especially among the executive team. Here are some ways to do this:

- **Decide to change the way you forecast.** Past results or trends are probably not as indicative of future sales as before the pandemic. Consider adding forward-looking variables to the forecast such as economic trends, consumer confidence, unemployment rate, website or other online activity, etc. Include key assumptions/metrics from the strategic plan, including nonfinancial metrics.
- **Quickly identify the business reasons.** Strive to quickly identify the business reasons behind any significant plan-to-actual financial variances so that course adjustments can be made when falling behind on financial or operational goals. These reasons should pertain to operational, market,

[20] See Figure A2 in Appendix 1.
[21] Lawson, 2021.

resource, and process elements of the business. It is not enough to generally understand the business—understand the details of the business from the perspectives of other functions. Show that you deserve to be "in the room where it happens."

- **Make the case for changing the purpose of planning and budgeting.** Wolfenbarger at Stryker illustrates a common challenge: "We often get a little wrapped around the axle on the gamesmanship of planning. At the end of the day, agreeing to and signing up for the target really sets the stage for the funding forecast for investment." Instead of using it to "control" operations and give stretch goals to sales and marketing, broaden the vision to a connection of short- and long-range financial goals with specific business initiatives, projects, plans, and resources to achieve those goals. Show evidence of positive return on investment (ROI) for improved FP&A. The best performers tend to follow this expanded vision of FP&A.[22] Caution: It may take a long time to convince management to make significant changes to how your company forecasts. Nonetheless, do all you can to show the benefits it can provide.

Consider adding forward-looking variables to the forecast such as economic trends, consumer confidence, unemployment rate, website, or other online activity.

[22] For more on how to compute ROI for FP&A improvement initiatives, see Appendix 2 in the *Key Principles of Effective Financial Planning and Analysis* SMA.

CONCLUSION

Forecasting the future is difficult but important. No matter how sound a forecast seems, reality will happen and often results will differ due to factors that simply could not be foreseen logically. A good forecast supported with good planning can be a key tool for an organization to adapt rapidly to changes in assumptions, markets, and operations.

This article has discussed nine specific ways to improve forecasting the future:

1. Expand the data available
2. Tools: Start small and add on
3. Use scenario planning
4. Address the knowing-doing gap
5. Model building: Think causality
6. Establish data collection systems
7. Improve assumptions and estimates
8. Monitor results and quickly identify the business reasons behind variances
9. Improve analytical skills

Try not to feel overwhelmed if you or your finance team are lacking the time and skills needed to improve forecasting. Every company has room for improvement in its forecasting, especially in today's tumultuous markets and economies. Assess where you are weakest and start there. Try a few steps mentioned in this report and build on that. Decide to improve! Never give up, and never surrender to the old dead-end trope of "That's how we have always done it."

A follow-up report will show how to perform more advanced predictive analytics including developing Monte Carlo simulations of multivariate models. •

APPENDIX 1: SURVEY DEMOGRAPHICS AND RESULTS

Survey Demographics

In September 2020, IMA surveyed 245 global financial executives and managers with experience in their company's FP&A practices. We eliminated 89 incomplete responses, leaving 156 usable responses from a wide variety of industries. Approximately 41% of IMA member respondents were from China, 22% from the United States, 28% from Middle East/Africa/India, 8% from Europe, and 1% from South America. They represented all major industries, and their companies ranged in sales revenue from below $100 million (47%), to $101 million to less than $1 billion (28%), to $1 billion or more (23%).

TABLE A1: PERFORMANCE GROUPS DESCRIPTIVE STATISTICS

Performance Group[a]	Frequency	Realize Potential of FP&A (0-100) (Mean)[b]	Revenue Growth Last Five Years (Mean)
Best	59	69	21%
Middle	51	59	12%
Worst	49	53	4%
Overall	159	61	12%

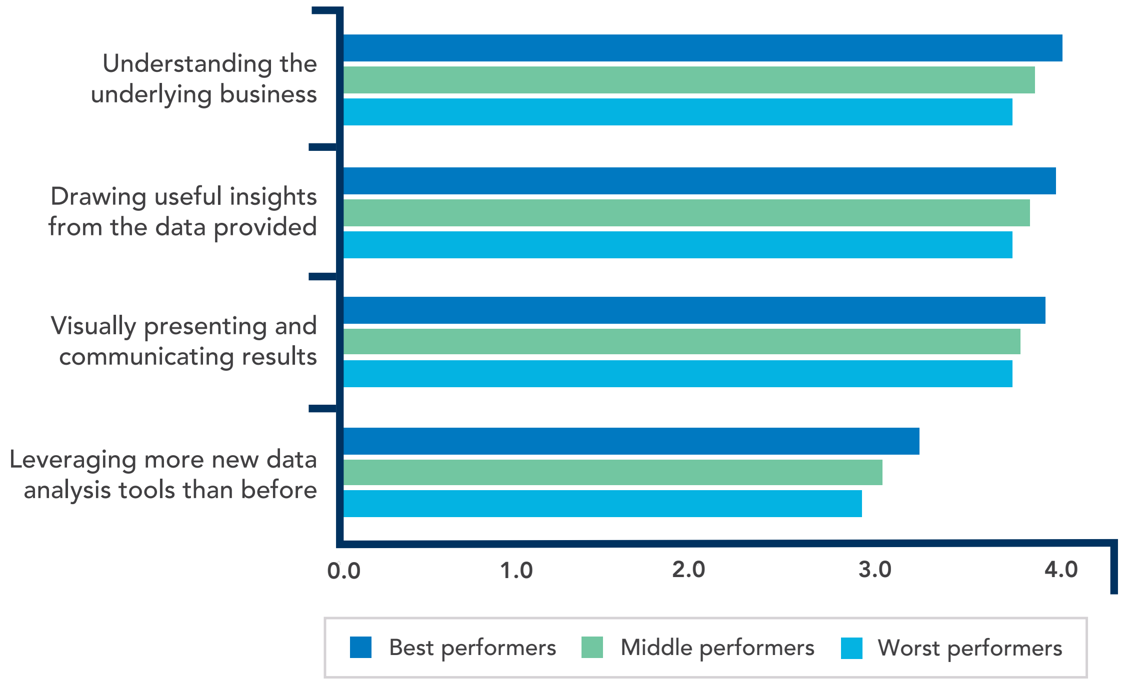
[a] Based on two questions: (1) *We consistently meet or exceed the targets we set for ourselves as a company/institution*, and (2) *We consistently meet or exceed the results of our competitors.*

[b] Based on the question, *To what degree does your organization realize the full potential of FP&A (from 0-100)?* Differences are statistically significant at the p<.002 level.

FIGURE A1: BIGGEST FP&A CHALLENGES FACED NOW

FIGURE A2: POSITIVE IMPACT OF PREDICTIVE ANALYTICS SKILLS

On a scale of 1 (lowest) to 5 (highest), how would you rate your FP&A team's ability in the following areas?



Categorizing companies as best, middle, and worst performers is based on two questions: (1) *We consistently meet or exceed the targets we set for ourselves as a company/institution*, and (2) *We consistently meet or exceed the results of our competitors*. "Best performers" are those who answered "yes" to both questions. "Middle performers" are those who answered "yes" to one question but "no" to the other. "Worst performers" are those who answered "no" to both questions.

APPENDIX 2: USING RULE OF FIVE, BAYES THEOREM, AND SOFTWARE TOOLS FOR BASIC PREDICTIVE ANALYTICS

In business, we are often presented with situations in which we have limited information or data, yet we are compelled to make a decision. At a later date, incrementally more quantitative data on the topic will be available. Are there any tools that can assist us to make a more standardized, calculated decision?

Simple Example

You are the financial analyst for MegaVolt Power Grid Company. Your company sells large-scale energy storage solutions for companies. You are tasked with analyzing customer interest in a certain target market for a new, high-cost storage system. Your company has already run a survey of existing customers in this market gauging their interest in the product (on a scale of 0 to 100) at a given price point. This is the historical data given to you: 78, 44, 42, 58, 56, 44, 71, 63, 58, 57, 60, and 48.

Given the high up-front development time and cost of building this product, management does not want to take a chance on this product unless there is sufficiently strong demand for the product by new buyers. As such, it says it will move forward only if there is a large enough customer base (at least 20%) with an interest score of 70 or above. Given the limited data set we have so far, how can we predict whether management's threshold will be met?

To solve a probability question such as this, there are different approaches we can take. One method is a shortcut known as the Rule of Five.

Using the Rule of Five

The Rule of Five states that there is an approximately 93.75% chance that the median of any population is between the smallest and largest values in a random sample of five. The following approximates the 90% confidence interval of the population median and the corresponding estimate base on the sample:

Sample Size	*n*th Largest and Smallest Sample Value		
	*n*th	Low	High
5	1st	42	78
8	2nd	44	71
11	3rd	44	63

This table shows we can use the Rule of Five to determine that after the 11th observed price point, we can be more than 93% confident that the median price is between 44 and 63. The goal of the Rule of Five is to reduce uncertainty without wasting resources by gathering lots of data. It can be a useful tool when there is little data or little time to analyze a data set. Yet the range of answers can be quite broad. For instance, in this example, the median is likely somewhere between 44 and 63 and does not help much to answer our initial question. Therefore, this rule should only be used to reduce initial uncertainty or when a "quick and dirty" low level of accuracy is acceptable.

Using Excel and *t*-stat

Excel can be a good tool for calculating uncertainty and probability distributions. One approach is to find the mean and standard deviation (SD) of a sample population and then use these amounts to quantify the likely range of values for the actual population. The SD is the amount of variation of a given set of data values. The Excel function for computing the SD is STDEV (range of values).

Step 1: Find the mean of the sample:
=average(78, 44, 42, 58, 56, 44, 71, 63, 58, 57, 60, and 48) ➔ **56.6**

Step 2: Find the standard deviation of the sample:
=stdev(78, 44, 42, 58, 56, 44, 71, 63, 58, 57, 60, and 48) ➔ **11.0**

Step 3: Compute the resulting confidence interval for the population of the mean:
At 68% (1 SD) confidence the range is between **45.6 and 67.6** [56.6 ± 11.0]
At 95% (1.96 SD) confidence the range is between **35.0 and 78.1** [56.6 ± (2)(11.0)]

For greater accuracy, we could use a random number generator between steps 2 and 3 to run simulations in Excel. Studies indicate that simulation of uncertain variables for 10,000 iterations yields a reasonable approximation of the likely outcome. But, as Excel is not built to handle very large data sets, this can be a tedious process. This is where programming languages like R and Python come in handy, which we will discuss more later.

Another option is that given that we have a small sample size, we can use the *t*-statistic approach to estimate the likelihood of the answer. When the original population's SD is unknown, the *t*-statistic approach can help us to estimate the mean of a population from a sampling distribution.

Using Excel, we can use the *t*-stat function given the following steps.

Step 1: Find the mean of the sample:
=average(78, 44, 42, 58, 56, 44, 71, 63, 58, 57, 60, and 48) ➔ **56.6**

Step 2: Find the variance of the sample:
=var(78, 44, 42, 58, 56, 44, 71, 63, 58, 57, 60, and 48) ➔ **120.6**

Step 3: Find the SD estimate of the mean:
=sqrt(mean/count(# of items)) ➔ **2.2**

Step 4: Use tinv() and calculate the 80% CI for the population mean:
=tinv(0.2,count(# of items -1)) ➔ **1.4**

Step 5: Calculate sample error:
= SD estimate of the mean * tinv() ➔ **3.0**

Step 6: Compute the resulting confidence interval of values:
=average + sample Error =56.6 ± 3.0 ➔ **53.6 to 59.6**

From our calculations, we find the probability is 80% that the original population's mean is between 53.6 and 59.6. While this method helps us reduce uncertainty at an early stage, it still does not answer our question of what percent of the population will likely score 70 and above? It can only give us a confidence interval for the mean of the real data set. To answer this question, we will need to use the Bayesian method, which will provide us more robust information than either of these Excel-based approaches.

What Is Bayes Theorem and Why Is It Useful?

Bayesian inference is a method for figuring out unobservable quantities given known facts. It uses probability to describe the uncertainty over what the values of the unknown quantities could be. It allows you to make a reasonable approximation of a likely outcome as new data are acquired.

In mathematical notation forms, the **marginal probability** of an event A happening, irrespective of other variables, is denoted as P(A). For example, the probability of rain occurring tomorrow is 40%, so P(A) = 40%. A **joint probability** of two events occurring simultaneously is P(A) x P(B), or P(A,B). For example, what is the probability that it will both rain (40%) and snow (20%)? The probability would be 40% x 20%, or 8%. The probability of one event occurring given the occurrence of another event is known as **conditional probability**, denoted as P(A|B). As in, what is the probability of P(A) occurring given P(B) has already occurred?

Here is what the Bayesian theorem looks like in mathematical terms.

Bayesian formula:

P(A|B) = (P(A) * P(B|A)) / P(B)

P(A) = Prior likelihood

P(B|A) = Probability of B given the occurrence of A

P(B) = Evidence (Note that when using computing power to create a generative model, P(B) is no longer necessary so we will leave this out of our calculation.)

Thus, the final formula is: P(A|B) = (P(A) * P(B|A))

Because our simple example is an example of a conditional probability question, Bayesian inference is an ideal method to solve it. We can use this formula to compute the conditional probability of reaching the 70% or above goal given the previous observations. The next section illustrates how to do this for our example data using the R programming language.

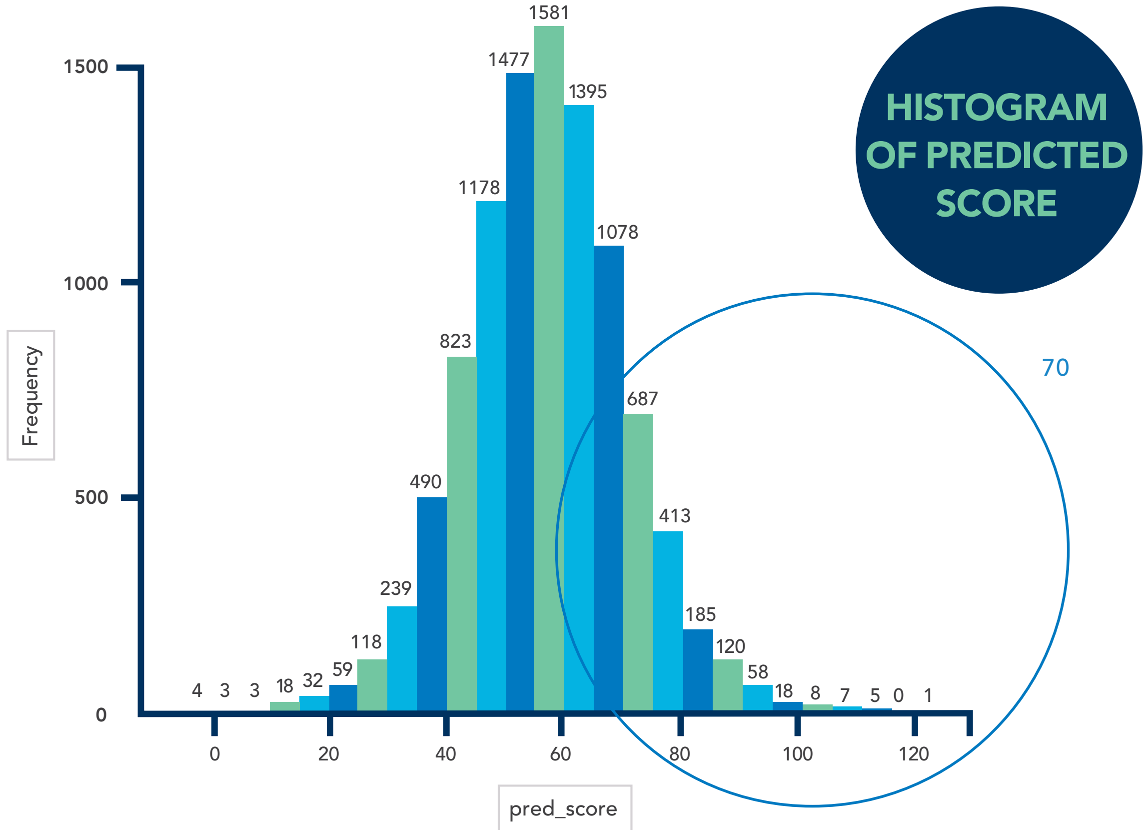
Using Bayes Theorem and R

We can code the answer using the R programming language. R is a language and free software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing. It is widely used for data analysis and is downloadable (**www.r-project.org**). Note that other more user-friendly options for doing similar analysis include SAS, SPSS, and Python but must be purchased. The use of these programs can be superior to Excel because they can more easily run large iterations of formulas that would be difficult to do in Excel. This allows us to create models that would be considered a reasonable approximation of a likely outcome (i.e., more than 10,000 iterations), which is a requirement for Bayesian inference.

In a nutshell, we begin by first establishing a prior likelihood based on the existing data. We then update this prior distribution with new data received to obtain a posterior distribution. We then condition on the posterior distribution with the metrics we are looking for to create a model that can answer our question. In this case, the question is: Are at least 20% of customers rating our proposed new system a score of 70 or above?

(Note: The following is the code the analyst writes into their R notebook. Note that "#" indicates noncode, explanatory portions.)

```
# We start by creating a "Customer_score" data frame for all of the current customer assessments.
Customer_score <- c(78, 44, 42, 58, 56, 44, 71, 63, 58, 57, 60, 48)
# We then define the size and shape of the grid to hold 10,000 data points.
pars <- expand.grid(mu = seq(0, 75, length.out = 100),
                    sigma = seq(0.1, 40, length.out = 100))
# We define and calculate variables to hold the prior density and SD for the prior model using the
functions "dnorm" and "dunif." Note that because we did not have any data prior to the customer
data, we are assuming that the distribution for the prior has a mean of 100 and follows a normal or "bell
curve" distribution. This prior model will hold 10,000 data points.
pars$mu_prior <- dnorm(pars$mu, mean = 100, sd = 100)
pars$sigma_prior <- dunif(pars$sigma, min = 0.1, max = 50)
pars$prior <- pars$mu_prior * pars$sigma_prior
# We create a for loop that will pull in the new data into the variable "likelihoods."
for(i in 1:nrow(pars)) {
  likelihoods <- dnorm(customer_iq, pars$mu[i], pars$sigma[i])
  pars$likelihood[i] <- prod(likelihoods)  }
# We calculate the factors of the posterior model by using the Bayesian formula (current likelihood given
the new factors times the prior likelihood), and then normalize the result. This data is put into a data
frame.
pars$probability <- pars$likelihood * pars$prior
pars$probability <- pars$probability/sum(pars$probability)
# At this point, we have successfully created a posterior distribution. We need to take one further step
with our posterior model. As it is not in a closed-end formula format, we need to take some samples
from it in order to create a distribution from which we can extract the relevant information and answer
our initial question.
# We create the variable "sample_indices" and select 10,000 samples from the posterior distribution.
sample_indices <- sample(nrow(pars), size = 10000,
     replace = TRUE, prob = pars$probability)
# We create the variable "pars_sample" to collect the results.
pars_sample <- pars[sample_indices, c("mu", "sigma")]
# We calculate the predicted scores.
pred_score <- rnorm(10000, mean = pars_sample$mu,
        sd = pars_sample$sigma)
# We calculate and display the quantiles of the distribution at 2.5%, 50%, and 97.5%.
quantile(pars_sample$mu, c(.025,.5,.975))
# The output of this is:
        2.5%        48.4848484848485
        50%         56.0606060606061
        97.5%       63.6363636363636
# We visualize the new distribution using the histogram function.
hist(pred_score)
```



 # We can see from the chart above that the circled region represents those with scores 70 and above. We insert a command to calculate the probability of a customer being interested (>= 70 score).

```
mean(pred_score >= 70)
```

 # The output is: **0.1443**

The output tells that only 14.43% of customers are likely to rate the new technology at the price point at a score of 70 or above. This would be the same result as adding up all of the scores 70 and above and dividing by 10,000. In other words, we can be 14.43% confident that the "credible interval" of the population is 70 or above. Note that Bayesian inference does not provide confidence intervals, because it generates an actual model to represent the original population, rather than the credible interval derived from the generated model.

Therefore, in our example, management's 20% threshold has not been met, and we should recommend to management to hold off on launching the new product for now.

Compared to the *t*-stat approach, the Bayesian inference model method provides much more useful information. A *t*-test can provide only a point estimate of parameter values, whereas the Bayesian method provides a complete range of possibilities and allows the user to sample from them. It allows the user to ask more specific follow-up questions. For example, at another given parameter, what is the likely probability? Thanks to the use of a generative model and added computing power, the Bayesian model is superior to the *t*-stat methods.

More Complex Models

We showed examples of different approaches to solve a fairly simple question. But what if the question were more complex and required many more variables, such as the price of oil one year from now? What would such a model look like and require? A follow-up report will show how to perform Monte Carlo simulations of multivariate models. •

APPENDIX 3: IMA COMPETENCY FRAMEWORK MATERIAL REGARDING PREDICTIVE ANALYTICS

Note: These skills are all covered in the CMA® (Certified Management Accountant) certification.

Budgeting and Forecasting

Applied Knowledge

- Identify and analyze the relationship between different resources and requirements of a comprehensive financial or operational forecast.
- Synthesize and interpret data from multiple sources.
- Anticipate capital requirements to support growth initiatives and optimization.
- Demonstrate and understand linkages between budget lines and perform appropriate planning (e.g., if sales go up, commissions should also go up).
- Validate assumptions made by departments.
- Develop a master budget to support the goals of a small to midsize organization or department/ division of a large organization.
- Prepare the projected income statement, balance sheet, and cash flow statement.
- Evaluate capacity constraints for budgeted activity levels.

Skilled

- Forecast in an environment of uncertainty using sensitivity analysis.
- Use statistical techniques such as regression, exponential smoothing, and confidence levels.
- Analyze and synthesize data from external sources to recognize patterns and predict customer behavior.
- Enhance the forecasting accuracy through the discovery of key and relevant trends by exploring large data sets using data analytics and data mining techniques.
- Recommend an appropriate budgeting methodology (e.g., flexible, continuous, rolling, zero-based) to use in a given business situation.
- Link the budgeting process to the strategic planning process.
- Integrate and consolidate information from multiple departments.

Expert

- Perform long-term analysis in periods of uncertainty using advanced statistical techniques.
- Lead collaborative forecasting efforts incorporating information from multiple internal and external expert sources and sophisticated modeling techniques.
- Communicate complex forecasts and budgets to others.
- Design and lead the budget and financial planning process across multiple business units in a complex organization using advanced software tools.

Data Analytics

Applied Knowledge

- Extract, transform, and query data using appropriate tools such as SQL.
- Interpret information needs and translate into actionable requests for data analysis.
- Use descriptive analytics to evaluate efficiency and effectiveness of business initiatives.

- Use simple linear regression to predict business outcomes and interpret results.
- Determine and report cause and effect using diagnostic techniques.
- Perform ad hoc exploratory data analysis using query languages.

Skilled

- Utilize specialized reporting tools (e.g., eXtensible Business Reporting Language, or XBRL), and interpret results.
- Design organizational templates for use by others.
- Mine large data sets to reveal patterns and provide insights.
- Use predictive analytics techniques to interpret results, draw insights, and make recommendations.
- Apply statistics to a data set using specialized statistical software and/or business intelligence software.
- Use multiple regression for predictive and prescriptive purposes, and interpret results.
- Transform raw, unstructured data into a form more appropriate for analysis (e.g., data wrangling).

Expert

- Implement solutions using multiple query, scripted, or interpreted languages (e.g., SQL, Python, and R).
- Build prescriptive models to optimize organizational performance (e.g., goal seeking).
- Use advanced statistical tools for exploratory data analysis to reveal patterns and discover insights to achieve business outcomes (e.g., cluster analysis, time-series analysis, and Monte Carlo analysis).

Data Visualization

Applied Knowledge

- Utilize table and graph design best practices to avoid distortion in the communication of complex information.
- Demonstrate an understanding of how to best communicate results with intermediate visualizations (e.g., histograms, area charts, and heat maps).

Skilled

- Evaluate data visualization options and select the best presentation approach for the intended audience.
- Demonstrate an understanding of how to best communicate results with advanced visualizations (e.g., Sankey plots, bubble charts, and network diagrams).
- Accelerate decision making using visualization tools and/or code packages to construct multivisual dashboards combining relevant visualizations.
- Utilize simplicity of design techniques to present results of complex data analysis in an understandable manner.
- Utilize advanced features of visualization applications.

Expert

- Interpret and communicate complex analyses to stakeholders using advanced data visualization techniques at an audience-appropriate level.
- Construct custom visualizations using JavaScript—either in website or with business intelligence platforms.
- Demonstrate expertise in all three aspects of data visualization: substantive, statistical, and artistic.

GLOSSARY OF TERMS

- **Budget (also called profit plan)** – The financial side of the annual operating plan. A statement of planned or expected revenues, expenses, assets, and liabilities. A budget provides guidelines for future operations and appraisal of performance.
- **Business planning** – Understanding where the business is today, where you want it to go, determining how to get there, monitoring progress along the way, and making course adjustments as necessary.
- **Continuous forecasting** – Forecasting on an as-needed basis.
- **Data lake** – Repository with a vast amount of data in its native format until it is needed for management purposes, like forecasting.
- **Driver-based forecasting** – Applying statistical measures (such as occupancy rate) in financial models to produce a projection of future results.
- **Forecasting** – A prediction of future events or conditions, usually based on a financial model.
- **Initiative** – Usually a longer-term program that is undertaken to meet established goals and to improve the business.
- **Key performance indicator (KPI)** – Measurable value that demonstrates how effectively a company is achieving its most important strategic objectives.
- **Machine learning** – Method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns, and make decisions with minimal human intervention.
- **Model** – A formulaic set of potentially causal KPIs or other factors that can be expected to affect the outcome.
- **Predictive analytics** – The use of data, statistical algorithms, and machine learning techniques to identify the likelihood of future outcomes based on any relevant data. The goal is to go beyond knowing what has happened to decrease uncertainty about the future and associated risk.
- **Project** – Planned set of interrelated tasks to be executed over a fixed period and within certain cost and other limitations.
- **Project plan** – Defines the activities and tasks that will be executed over a specific time period.
- **Rolling forecast** – Updating a forecast on a set schedule and looking at the same number of periods each time (such as a rolling 18-month forecast).
- **Scenario planning (also known as scenario analysis)** – Used to better prepare for the future by articulating a small number of potential scenarios involving one or two critical uncertainties and developing plans on how to succeed in each of those scenarios.
- **Variance analysis** – Investigation of the causes of the variances between actual costs and standard or budgeted cost (or actual to forecast results) with explanations for key differences.

For more information, please visit **www.imanet.org/thought_leadership**.