SOCIETY OF ACTUARIES
MODELING SECTION

THE MODELING PLATFORM
JUNE 2020

Models That Will Make the Cut

By Nate Worrell

I have a confession to make. I watch—and enjoy—reality fashion competition shows. I'm watching "Making the Cut" with my daughter on Amazon. It's part of our pandemic stay-at-home sanity program.

Speaking of pandemics, have you noticed that models seem to be popping up everywhere? Actuaries, university researchers, epidemiologists and governments all seem to be trying to make forecasts. What's interesting is seeing which of these forecasts end up being widely dispersed, featured in headline news stories or sitting in front of the president at the White House. In some ways, it's a real-life modeling competition.

The competition is fierce, and the verdicts can be cruel and sudden. As Heidi Klum put it in "Project Runway," "One day you're in; the next day you're out."

Being an actuary, my first instinct to understand the risks of the COVID-19 virus was to turn to other actuaries. They did not disappoint.

- Stuart McDonald first started throwing out predictions of the number of UK cases on Twitter, and his model turned out to be pretty accurate. More than that, he included lots of good discussions about shortcomings and unknowns, and he seemed to be constantly dispelling invalid conclusions from the media. Lastly, he updated his model to show the effects of interventions, showcasing how things would vary from his original forecast. McDonald is also part of the COVID-19 Actuaries Response Group formed on LinkedIn.[1]

- The Society of Actuaries (SOA) Research Team put out a solid package[2] that not only took a look at how the virus progressed, along with ample discussions of limitations, but also brought in context to give a comprehensive view that linked disease spread to hospitalization rates to economic impacts.

But I didn't see these actuarial models making it to the front page. Instead, universities seemed to lead the way. Johns Hopkins,[3] the University of Washington[4] and others got a bit more press. Why? Certainly the models from actuaries were as strong as, if not better than, any of the others. Perhaps it is a question of popularity, with other organizations enjoying more prominence among reporters.

In this case, perhaps the best competitive strategy is to form an alliance. The SOA package uses Johns Hopkins data. But is Johns Hopkins looking to the SOA? Actuaries should be in the game, but it will require some assertiveness. If you see a model in your local newspaper or on your favorite website, why not reach out to that author? Maybe there can be a collaboration in the future.

But beyond prestige, is there something else? How do we make a model mass media–worthy? Tyra Banks asked contestants of "America's Next Top Model," "Do you want to be on top?"

What if this were your orienting question as you built and summarized your model? What would you need to do differently? For mass public appeal, three traits seem to matter: aesthetics, story and punchline:

- The results of the model have to be presented in a compelling way.
- The output must make sense to a nonexpert.
- You must be able to summarize the results with a newsworthy headline.

Note that these criteria are all about the presentation of results and less about the model architecture itself.

Yes, the model should be accurate, fast and transparent. Let's not minimize these key qualities—we have a whole Actuarial Standard of Practice (ASOP)[5] about them. A model must be reliable, but it will be worthless if it takes a herculean effort to extract conclusions from it.

Compromising on accuracy or not evaluating the data thoroughly could result in a lack of trust. However, there seems to be a certain tolerance for error in models. The people who put together weather predictions, election forecasts, Super Bowl odds or buy/sell recommendations for stocks seem to keep their day jobs despite having a history of (sometimes significant) misses. And that doesn't even touch on your latest horoscope. In a world with high uncertainty, any prediction can serve as a lifeline, a light in the fog. So the main purpose and value of a predictive model may not be in the actual prediction, but rather in having a voice that sets a journey in motion. Set a course now and refine later, turning in a new direction if you must. To be clear, I'm not suggesting that we be satisfied with models that are "good enough" or simply directionally accurate. Rather, being bold with a flawed model may yield more exposure and attention than hesitating and waiting for a more perfect model.

Let's elaborate on the criteria of aesthetic, story and punchline from a tactical perspective.

MAKE IT PRETTY

The good news is that there are lots of software options that can do this for you. Time matters. The first to show results has an edge over any followers, even if the newcomers are more accurate. Building graphs in Excel may be a great way to kill time on the weekend; it is not a winning strategy.

Take the time to learn how to use data visualization tools, how to build something interactive (new SOA research[6] is moving in this direction), and how to understand design principles.

MAKE IT MATTER

What makes a great story? Certainly the writing matters, and a clever concept counts, but at the end of the day it comes down to emotion. Do you see yourself in the story?

A useful element that can add depth to any chart is some sort of comparison mechanism. How does this result compare to something I know and can relate to? Think of COVID-19 vs. influenza.

A story is more impactful if it is easy to understand. The harder readers or observers have to work to digest the information, the more energy they spend on logic, with less remaining to spend on emotion. Things like numbers with lots of different digits and percentages are inherently complex.

- It is easier to absorb $3.5M than $3,489,750.
- One in 5 is easier to imagine than 20 percent.
- Be clear about absolute vs. relative risk. The former tends to be more useful.

I've been a judge and mentor for the Actuarial Foundation's Model the Future Challenge for the past couple of years. The winning teams use strong models, but a lot of strong models get turned away because the message gets lost. Strong communication elements take models to the winner's circle.

How do the dots connect? What is the takeaway message?

MAKE IT NEWSWORTHY

Nuance doesn't matter. Caveats are worthless. Spin a sensational statement instead.

My fellow actuaries may find this approach blasphemous and reckless. It goes against the fabric of the mathematical mind. Our default position and strongest statement is usually "It depends."

Is that the answer you hoped for when you asked your spouse to marry you? When you wanted to know if you'd make it out of surgery alive? There needs to be something bolder, something more definitive to latch onto.

Leave the dissection and debate about details to academic journals and Twitter threads.

To add more drama to the claims, consider a couple of key questions. Who wins, and who loses? What are the fears and hopes of the people who will see this output?

Imagine you were presented with a month-end result titled "Excess Lapses Challenge Company Profits—Just a Blip or Long-Term Trend?" It may seem a bit dramatic or hyperbolic, but that's the point. I'll concede that being an alarmist may not be a welcome feature in your day-to-day job, but in the world of public consumption, there needs to be a "wow" moment.

As a bonus, I want to throw out the idea to make it yours. What is your data output style? Do you like a certain color palate, a certain font? If I look at a series of charts, can I pick yours out from the rest?

THE REAL PURPOSE OF MODELS

Models are more than just mathematical applications: they are communication devices. To really serve members of the public—to get our objective, disciplined work into their hands—we need to appreciate that message delivery is as important as message content. Model output needs a powerful package.

Otherwise, we'll hear Heidi Klum say, "Auf Wiedersehen," as she blows us a kiss good-bye. ■

Nate Worrell, FSA, is an associate director at Moody's Analytics. He can be contacted at *nathan.worrell@moodys.com.*

ENDNOTES

1 COVID-19 Actuaries Response Group, *LinkedIn*, 2020, *https://www.linkedin.com/company/covid-19-actuaries-response-group/*.

2 Society of Actuaries COVID-19 Research, 2020, *https://www.soa.org/resources/research-reports/2020/impact-coronavirus/*.

3 "COVID-19 Dashboard by the Center for Systems Science and Engineering," Johns Hopkins Medicine & Science, 2020, *https://coronavirus.jhu.edu/map.html*.

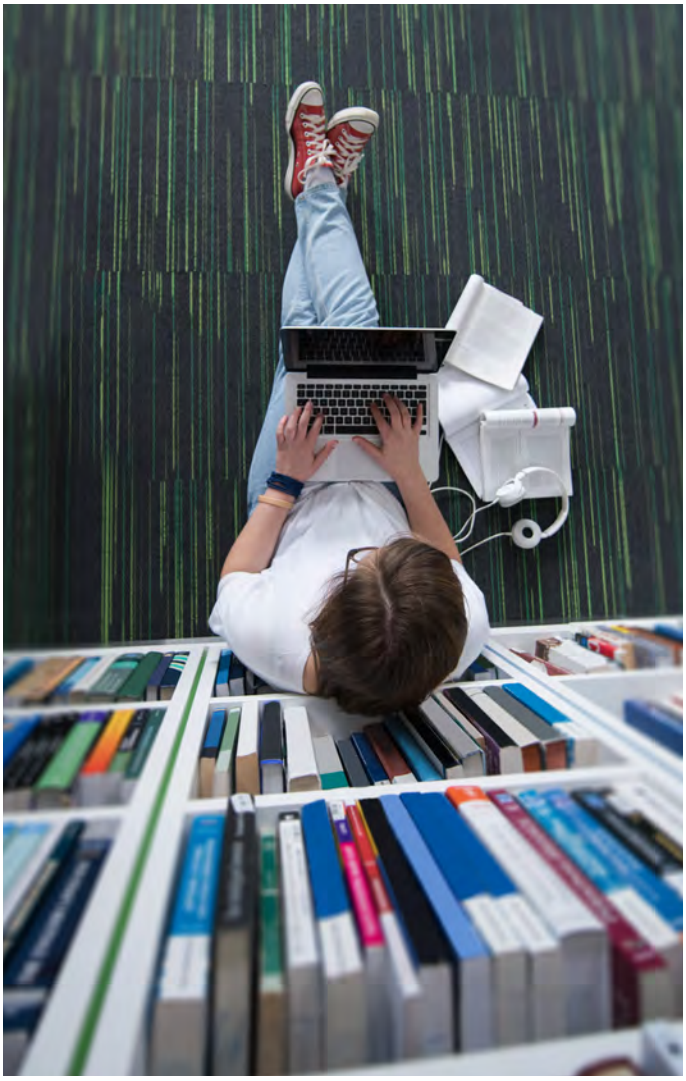
4 "COVID-19 Projections," Institute for Health Metrics and Evaluation at the University of Washington, 2020, *https://covid19.healthdata.org/*.

5 "Modeling," Actuarial Standards Board, June 2013, *http://www.actuarialstandardsboard.org/asops/modeling/*.

6 "Fifty States, Fifty Stories: A Decade of Health Care Reform Under the Affordable Care Act," Society of Actuaries, March 2020, *https://www.soa.org/resources/research-reports/2020/50-states-50-stories/*.

Modeling: An Actuarial Student's Perspective

By Justin Serebro



The purpose of this article is to share some important modeling lessons that I have learned and to offer some advice for pre-ASA and new ASA actuaries doing modeling work. I am an ASA with a total of four years' experience and still study for FSA exams while working full time. Therefore, I consider myself an actuarial student.

Entering my first actuarial internship, I knew that actuaries did a lot of modeling but I did not fully understand what modeling was. It turns out that actuarial modeling is very different from the mainstream depiction in "Zoolander." As Derek Zoolander has his signature "blue steel," core actuarial modeling relies on Excel. In this article, I would like to share answers to common questions about modeling such as:

- What kind of modeling do actuaries do?
- How do actuaries know if their model is right?
- How do you become "good" at modeling?
- What is the most important skill for effective modeling?

WHAT KIND OF MODELING DO ACTUARIES DO?

The wonderful (and scary) part of life is that there is always risk involved. This risk may result from a decision that must be made, or it may be inevitable. Risk occurs because we do not know what the outcomes will be ahead of time (if you do, then please feel free to skip the rest of this article). Modeling allows actuaries to quantify risks and determine whether these risks are within the company's risk tolerance. Depending on the practice area and the product being modeled, actuaries may be more concerned with certain risks. It is important to be able to model these material risks properly. For example, a life actuary pricing a yearly renewable term life product will be more concerned about mortality risk than interest rate risk, due to the short-term nature of the product and the significant financial losses that will occur if actual mortality is higher than the pricing assumption.

HOW DO ACTUARIES KNOW IF THEIR MODEL IS RIGHT?

The scary and ubiquitous truth is that no model is perfect. All models are simplified versions of reality. Again, if every outcome and its timing were already known with certainty, then there would be no need for models. Actuaries regularly adjust and fine-tune models as changes in regulations, economic environment and technology occur. The key is to make sure you understand any limitations, approximations and weaknesses of the model so decision makers can take these into account. Conducting sensitivity testing, developing ranges around estimates, and comparing model results with actual outcomes can be useful tools to get more comfortable with the results produced by models.

HOW CAN YOU BECOME "GOOD" AT MODELING?

Like any discipline, modeling takes time, practice and dedication. Through studying for the FSA exams and experience on the job, you will get more exposure to best practices for actuarial modeling. Nearly every form of actuarial work involves using a model or interpreting model results to advise management regarding business decisions. Therefore, as an actuary supporting your organization's decision-making processes, you need to understand how your model works and its intended purpose. I have found it useful to try to replicate model calculations using Excel. Validation exercises are also done within many modeling platforms. Additionally, detailed documentation regarding calculations performed in the model can be a valuable starting point in building your simplified model. It may not be reasonable to expect an entry-level actuarial student to replicate the model under every scenario or throughout all time periods; however, the idea is to try to understand the fundamental calculations that are going on in your model.

WHAT IS THE MOST IMPORTANT SKILL FOR EFFECTIVE MODELING?

Clear communication is essential if you want to become an effective modeler. However, communication might be the most overlooked aspect of modeling. Anytime you are assigned work, it is important to understand why you are doing that task and clearly confirm what it is that others need you to do. This is a very important consideration because the model's use needs to be aligned with its intended purpose. Also, understanding why you are doing the work can provide insight into management's concerns. When you receive a request, try to create a visual representation of your interpretation of the assignment. This could be a simplified Excel workbook or drawing that illustrates your understanding. By conveying to others how you interpret their request, you give them the chance to agree or disagree with your version. If there has been some misinterpretation, then you need to set up a meeting to get everyone on the same page.

I hope these insights are helpful and that you can apply some of them to your work. ■

Justin Serebro, ASA, is a risk management actuary at Pacific Life. He can be reached at *justin.serebro@pacificlife.com.*

SOCIETY OF ACTUARIES
MODELING SECTION

The Importance of Centralization of Actuarial Modeling Functions, Part 3

Implementing Model as Data Using Moody's Axis

By Bryon Robidoux

In part 1 of this article, I used software engineering principles to show that decentralizing models come with very high costs. I showed that centralization of a modeling department is a step in the right direction, but it isn't enough. The key to running a smaller, better, faster and cheaper modeling department is to focus on modularity and work product[1] reuse according to software engineering principles. Part 2 introduced the reader to the major components of DevOps and how it is the basis for actuarial modernization. But up to this point, there haven't been any practical examples of how to implement DevOps or what the end result would look like.

This article will focus on how to transform a Moody's Axis model that suffers greatly from the monolithic-system problem[2] to a model that has maximized data reuse and promotes the use of DevOps. This will be defined as a Model as Data (MAD), a model that has completely data-driven processes that will speed up the throughput of enhancements, improve testing, simplify production processes and make ad hoc runs easier.

AN INTRODUCTION TO MOODY'S AXIS

Non-Axis users may need a frame of reference for its two major components: E-Link and the dataset. E-Link's main goal is to manage the collection of the organization's models and orchestrate their execution. It has a very Windows Explorer feel. E-Link can be automated with scripts to externally manipulate datasets and customize orchestration using Axis Jobs and E-Link scripts, respectively. One of the most important enhancements to E-Link in the last year or so is Formula Link. This extension builds reusable libraries that can be shared among multiple models and E-Link scripts.

The dataset is the model itself. This is where 98 percent of the work is done. From E-Link's point of view, the dataset is like a big zip file full of Axis proprietary and user-created files. From within the dataset, the dataset's interface has its own subcomponents, such as batches, remote tables, datalinks and datalink tables. Batches instruct Axis to do calculations and other operations needed to manipulate the dataset. The remote tables point to data that live outside the model. Datalink batches transform the remote tables into datalink tables. Datalink tables are the internal tables the model uses for seriatim policyholder information and other data. The cell tables and projection tables are used to describe how a particular insurance plan or rider will be calculated.

IT'S ALL ABOUT THAT DATA

The dataset can be split into two fundamental pieces: data and code. This article will be focused on the data portion, while the

code will be addressed in the next article. One of Axis' powerful features is its import/export batch. On export it creates a turnaround document, which is a file that can be exported in one step and then reloaded into the program at a future step to populate with new data. This allows a user to dynamically create Axis objects without manual intervention and consequently means that a large portion of the dataset can be created on the fly. Examples of objects that can be exported are datalink tables, cell tables and projection tables, which means they can be treated like data. This will be defined as Calculation as Data (CAD).

Data Transformations

In the worst-case scenario, there are up to three places where data can potentially be transformed. The first is within the databases inside the insurance organization. The second is the SQL server instances that sit in the Moody's cloud. The third is within the dataset using datalink. If your organization is using AWS or Azure, they both remove the Moody's cloud layer. As data are transformed and moved, each layer in the data architecture increases the controls because the actuary has to validate that the transformations were successful. (When a production run breaks down over the weekend, there is nothing better than having to dig through three different locations and multiple people to figure out what went wrong. Those are the weekends I look forward to the most!)

All the data should be located in one place along with its transformations. This location should be outside the dataset to create a one-stop-shop for diagnosing data issues, to simplify debugging and to reduce the monolithic-system problem by maximizing data reuse. The data should be transformed until they can be loaded one-to-one with a resulting datalink table, cells table or projection table that they will inevitably populate. Never read directly from a database table. Always read from a view or materialized view to promote good data encapsulation protocols.

Using Files

The actuary should not make the dataset a dumping ground. Even though Axis allows it, no files of any kind should be stored within the dataset. (This is a huge source of potential error if people use outdated files, and it can cause confusion during reporting cycles.) Loading files from the dataset requires manual intervention and makes the data-driven processes clunky at best or moot at worst. If external Excel model results are used to populate the dataset, they should first be stored in the corporate database along with all other data so the results can be reused easily by multiple parties and be included in the automated data assembly process. This removes the potential of putting production data in email and using the incorrect version for a production process. How many times has this caused problems in your organization?

CONTINUOUS DATA INTEGRATION

Continuous data integration allows the same hydration routines to be reused to dynamically create models for production calculations, attribution analysis and ad hoc analysis. The fully data-driven, flexible model will encapsulate and abstract away the details so these calculations can be uniformly handled, thereby eliminating the manual setup usually required.

To implement continuous data integration, automation is required. Automation might include a simple interface—external to the dataset—to select data for any activity, such as a testing or production run. Behind the interface would be all the mechanics and metadata required to assemble the correct inputs, hydrate the model and execute a run. All interfaces should be configurable and script driven, so they can be parameterized and executed repeatedly without human intervention.

The valuation team should never have to touch the dataset during a production cycle. Better yet, there should be no manual processes required during any reporting cycle to produce results from execution to final report. If such processes are required, it implies that at least one manual touchpoint is likely to produce errors, so controls will also have to be created and productivity will be murdered. All required materials should be staged and ready to go **before** that day occurs. Any processes that pull the latest market data or the like should be viewed as part of the model and automated accordingly. Errors cause the model to be rerun, deadlines to be missed, weekends to be lost and the model execution's cost to rise.

CONTINUOUS DEPLOYMENT

To continuously deploy changes to a model, the data need to be broken down between different environments such as development, quality assurance and production. It is critically important that database architecture be nearly identical for all three types of environment so that only the remote tables have to be updated to make all processes work. The process of deployment consists of nothing more than appending the turnaround documents created for the new features to the staging environment's turnaround documents. If any GUI is required for deployment, the productivity and robustness will suffer immensely.

The ultimate goal of a MAD is to create an environment in which all runs and the hydration of their data can be orchestrated through E-Link scripts. These scripts should depend heavily on Formula Link libraries for building reusable components with maximum logic reuse. For ultimate automation, these libraries should be data-driven so every aspect of running the model can be dynamic and configurable.

CONTINUOUS TESTING

The testing team will want data categorized for different types of tests, such as positive[3] and negative.[4] There is a tendency to use only positive testing in a modeling department, especially when the dataset has a million manual processes to import data

A get-it-done attitude no longer works. … The focus needs to change to a get-it-done-better attitude.

for use. No one will want to go through the time, effort or trouble to load different sets of data. Or worse yet, teams are afraid to use testing data because such data may accidently slip into the production model! It is important to realize that DevOps promotes using production-like environments for testing and development, but that doesn't imply that only production data should be used in all environments. This leaves many potential errors uncovered. (There is nothing worse than a vague or confusing error message in production that sends the valuation team on a wild goose chase with delivery timelines quickly approaching—especially when the issue is a simple input error that an intelligent error message could have easily pointed out.)

Continuous testing is only possible with continuous integration and continuous deployment. With these capabilities, the testing team can hydrate the model with any data they need to verify that the model is fit for production. Tiny models that are specifically targeted at the enhancement or fix being implemented can be created on the fly. This immensely improves the testing that can be performed by reducing errors, allowing the testing teams to receive feedback in minutes instead of hours or days and allowing features to be implemented within the time frame of a single Agile sprint. There is no possibility of the wrong data being used because the data are controlled by the well-tested environment. This means all kinds of tests can be run to verify that the model will handle bad data gracefully and good data properly.

TELEMETRY

As the data are assembled and transformed so they can be placed directly in the model, logs, error checking and efficiency metrics need to be captured and retained. They will provide immediate feedback on the health of the system, and they can help actuaries and IT find places in the processes where weaknesses and bottlenecks can be improved. All this enhances the turnaround time for results, so decisions can be made faster with better information.

BACK TO LIFE! BACK TO REALITY!

Now that the ultimate modeling platform has been described, where does reality set in and thwart the vision? It all falls apart with the batches! Why? you might ask. As of July 2019, the last time I used Axis, the batches cannot be exported from the dataset so that all the fields of the batches can be externally manipulated. In other words, Axis has no turnaround documents for the batches. The only fields that can be manipulated externally through Axis Jobs and E-Link scripts are in the dataset parameters, but fields such as report location and override set are not located in the dataset parameters. This means that from the user's perspective batches cannot be updated without using the dataset interface. This leads right down the manual processes' and manual controls' rabbit hole that should be avoided. **It is important to be able to export all batches so the model's behavior can also be data.** Thus, the dataset's behavior can be completely and dynamically manipulated. This is defined as Behavior as Data (BAD).

INFRASTRUCTURE AS CODE

Having Axis objects and batches as data—that is, CAD + BAD—is required to effectively create a MAD-compliant model. MAD is synonymous with the infrastructure-as-code (IAC) concept within DevOps. IAC means that any part of an application's environment and its versioned controlled components can be provisioned and set up in an automated fashion with no manual intervention. To generalize this to actuaries, the spreadsheets, Alteryx scripts and so on should also be contained in source control. Source control, such as GitHub, does not allow comparison or track changes of binary files such as spreadsheets. This is another reason to discourage the use of spreadsheets, unless the organization wants the added complexity of coordinating Incisive with GitHub.

CONCLUSION

This article has focused on reducing the monolithic-system problem by making the model and its processes data-driven so they are DevOps complaint. Model as Data makes the production, audit and controls teams more confident in the model they receive because it is designed specifically for continuous integration, continuous deployment and infrastructure as code. Therefore, it allows all aspects of production to be tested *before* the production cycle.

Developing a data-driven model that is compliant with MAD vision is no simple task. It requires a massive amount of infrastructure, which in turn requires a massive amount of coordination between IT and actuaries. To make this successful, it is important to remove the silos between these professional groups and "cross-pollinate" training and ideas. This will require cultural changes in how actuaries and IT work together.

At some point, the organization needs to decide where the complexity is going to reside. There is no such thing as a free lunch, and the complexity has to reside somewhere. Will it be in the automation of the processes or in the operation and coordination of people and their manual touch points? The latter may seem a comfortable choice because the complexities are all surface level and observable, but it will kill the opportunities to create economies of scale. Every new task will require a new person, eventually leading to more managers orchestrating the processes. The organization will get so large

and hard to move that it will be crushed by its own weight. As regulators and auditors expect more and faster results, organizations will beg regulators to kick the can down the road because they are overwhelmed with change. A get-it-done attitude no longer works because all the late nights and weekends are already being consumed with the current processes. As people are pushed harder, the technical debt will grow faster and more abundant, which is actually counterproductive.

The focus needs to change to a get-it-done-**better** attitude. This means constantly analyzing the processes for bottlenecks and brittleness, learning better approaches from other disciplines, finding the similarities and differences between processes and exploiting these similarities to create economies of scale. This can only be done if the automation of processes choice is made. Actuaries need to realize that the only way forward is to embrace technology and use it to its full extent. Just producing numbers may have been acceptable in the past, but now it is just as important to understand how to build the most efficient and maintainable processes possible. If you are not willing to do it for your organization, then do it for the health and viability of your profession!

If organizations and actuaries are willing to brave this new world, the modeling and valuation departments will become much faster. Their development methods and processes will become compliant with IT approaches to development. Development, testing, production and special analysis will become much easier. Late nights and herculean efforts should subside, which will increase morale, productivity, the confidence in the model's results and the organization's efficiency. ■

Bryon Robidoux, FSA, CERA, is a lead and corporate actuary, Actuarial Transformation, at The Standard. He can be reached at *bryon.robidoux@standard.com.*

ENDNOTES

1 A work product is any logic, data or data transformation that has the potential of being used in multiple places within the organization.

2 The monolithic-system issue was defined in part 1 of this article. It comprises the problems that exist from locking away all the logic, data and data transformation in the model so they can't be reused elsewhere. This is counter to sound software engineering issues and leads to duplication of effort on many fronts.

3 Positive tests use data that are within valid ranges to verify the system works as expected.

4 Negative tests use data that are outside valid ranges to verify the systems will fail where expected and have meaningful error messages.

MAKE THE MOST OF YOUR SECTION MEMBERSHIP

Stay engaged with your community

SECTION COMMUNITY

Looking for a venue for discussion around how assumptions are tested, approved, documented, stored, implemented and used? The Society of Actuaries (SOA) Modeling Section's Assumption Development and Governance subgroup is open to everyone (including non-SOA members and non–Modeling Section members). Not yet a member? It's free to join! Sign up to ensure you're aware of all the subgroup's activities.

PROFESSIONAL DEVELOPMENT

Market risk benefit (MRB) is a new financial statement line item to be measured at fair value under long duration targeted improvements (LDTI). MRBs principally include GMxBs on variable and fixed product and other market-sensitive benefits. Join us for the "LDTI Implementation Series: Modeling Considerations for MRB" webcast on June 23 from 12:00 to 1:15 p.m. ET to learn about key modeling considerations and issues for MRBs in the implementation of LDTI.

Listen to our latest podcast "Data Warehouse: Data Lake or Data Swamp?" with Tom Peplow, MSc. In this episode, April Shen, FSA, CERA, CFA, MAAA, interviews Tom Peplow, principal actuary with Milliman, on the interaction between actuarial modeling and big data. Tom discusses the basic concepts of data warehouses, as well as other tools and resources actuaries can use in the era of big data.

Get access to more info at ***SOA.org/sections/modeling***