ISSUE 37 | OCTOBER 2010

SOCIETY OF ACTUARIES Technology Section

CompAct

ELECTRONIC NEWSLETTER

CONTENTS

Share   Print Article   Search Back issues

QUICK LINKS

Technology Section Web site

LETTER FROM THE CHAIR

by Tim Deitz

The year is wrapping up, so this will be my last "Letter from the Chair." It has been a pleasure serving the Technology Section as chairman. I am pleased with what we have accomplished this year, and excited about the initiatives underway. We have a strong team with our new council and many dedicated volunteers that are entering 2011 with enthusiasm.

[full article]

EDITOR'S NOTES

by Howard Callif and J. Eddie Smith

As the days start getting shorter and the air crisper, we once again find ourselves trying to figure out where the year went. 2010 was another busy year for the Technology Section and a milestone year for CompAct. In January, we debuted our new Web-based newsletter.

[full article]

OUR EXPERIENCES WITH AGILE

by Robert Ream and Justin Bozonier

MG-ALFA is financial projection software developed and supported by Milliman and used by leading life insurance and financial firms worldwide to perform financial projections. Among other things, our product is known for consistently delivering quarterly software releases.

Council

Links of Interest

Fiction Contest

Howard Callif, Editor

SOA Staff
Meg Weber, Staff Partner

Sue Martz,
Section Specialist

Sam Phillips, Staff Editor

[full article]

EXCEL 2010–AN ANALYST'S PERSPECTIVE

by Andrew Chan

I have used Microsoft Excel for more than 15 years and it has been a great product that's getting better and better with every upgrade. However, I have never felt so enthusiastic about an upgrade as I did about Excel 2010; I can't stop talking to people about it and encouraging them to upgrade. Why is Excel 2010 so special? Excel 2010 allows the user to:

[full article]

ACTUARIES LOOKING TO THE SCAI FOR ANSWERS

by Van Beach

When running stochastic projections, the actuaries' need for computing power is insatiable. A multi-tier Scalable Cloud Actuarial Infrastructure (SCAI) may hold the solution.

[full article]

R CORNER–MEMORY MANAGEMENT

by Steve Craighead

Even though I'm still looking for all of my marbles, I'm going to talk about how to manage memory within R.

[full article]

NUMBER PUZZLE

by Marc Whinston

A challenging, fun puzzle for the readers of CompAct. Answers are due to Marc Whinston, at mwhinston@tiaa-cref.org, by November 22.

[full article]

LAST ISSUE'S NUMBER PUZZLE SOLVED

by Marc Whinston

The solver of last issue's puzzle is: Lee Michelson (Office of the Insurance Commissioner, Washington).

[full article]

THE 9TH ANNUAL SPECULATIVE FICTION CONTEST

by Gary Lange

Some of us dream about entering a blue alien body, complete with a tail, on the moon Pandora, while others hope their hot tub turns into a time machine, and even a few of us want to don the outfit that would make us Iron Man. The rest of you though, have more practical ideas about the future, and many of them revolve around our favorite topic, actuarial science!

[full article]

ARTICLES NEEDED

by CompAct Newsletter

Your help and participation are needed: Articles, book reviews and other information on technology useful to actuaries are needed!

[full article]

Actuaries
Risk is Opportunity.®

475 North Martingale Road, Suite 600 Schaumburg, Illinois 60173
Phone: 847.706.3500 Fax: 847.706.3599 www.soa.org

ISSUE 37 | OCTOBER 2010

SOCIETY OF ACTUARIES   Technology Section

CompAct

ELECTRONIC NEWSLETTER

QUICK LINKS

Technology Section Web site

Share   Print Article   Search Back issues

LETTER FROM THE CHAIR

The year is wrapping up, so this will be my last "Letter from the Chair." It has been a pleasure serving the Technology Section as chairman. I am pleased with what we have accomplished this year, and excited about the initiatives underway. We have a strong team with our new council and many dedicated volunteers that are entering 2011 with enthusiasm. I want to thank our retiring council members: Holly Loberg, Carl Desrochers and Tim Pauza. It has been a pleasure working with them and I wish them each the very best as they move on to other opportunities. Jeff Pomerantz, our new chair, will bring his skills, ideas and enthusiasm to the section leadership. Please feel free to contact him, or any of the other council members, with your ideas for how we can make the Technology Section valuable to you, the members.

I want to mention a few highlights from the past year that will continue to be major initiatives for 2011. The LinkedIn group for section members, "Society of Actuaries Technology Section," is continuing to grow. This online forum for networking and information sharing will focus on the challenges and interests of our members. As the technology demands grow for actuaries in an ever-more complex and challenging economic environment, we believe LinkedIn can provide us an invaluable opportunity to keep abreast of the latest developments in arenas where technology is playing a major role in the future of our industry. Addressing how this forum, as well as others, such as Twitter, can best serve the members will continue to be a priority for the 2011 council. There is a lot of potential here that we trying to understand and utilize. If you have

SOA Staff
Meg Weber, Staff Partner

Sue Martz,
Section Specialist

Sam Phillips, Staff Editor

not already done so, join LinkedIn (www.linkedin.com), join our group, see what others are up to and contribute to the discussions.

Another item I would like to mention is a collaboration we have developed over the past three years with Gartner Group (www.gartner.com). We published a survey in 2008 and 2009 to investigate the effectiveness of the collaboration between actuarial departments and information technology departments in the industry. The results of these surveys were presented in sessions at the last two years' annual meetings. The response we received from those sessions indicated a high level of interest in this important topic, so we once again repeated the survey for 2010. Hopefully, you will be able to attend our session in New York this year (Session #137), where we will be discussing observations from the survey results, including trends from our findings of last year.

As I close, I want to thank all of our volunteers for their willingness to get involved, and say thanks to all of our council members and friends of the council who work hard to keep our section going. Let us know if you have any ideas for the section or want to get involved. Please feel free to contact any of the council members listed inside the front cover of this newsletter with your ideas or to let us know if you would like to get involved.

Tim Deitz
2010 Section Council Chair

Actuaries
Risk is Opportunity.®

475 North Martingale Road, Suite 600 Schaumburg, Illinois 60173
Phone: 847.706.3500 Fax: 847.706.3599 www.soa.org

ISSUE 37 | OCTOBER 2010

SOCIETY OF ACTUARIES   Technology Section

CompAct
ELECTRONIC NEWSLETTER

QUICK LINKS

Technology Section Web site

Share   Print Article   Search Back issues

EDITORS' NOTES

by J. Eddie Smith, IV, FSA, MAAA; Howard Callif, ASA, MAAA

As the days start getting shorter and the air crisper, we once again find ourselves trying to figure out where the year went. 2010 was another busy year for the Technology Section and a milestone year for CompAct. In January, we debuted our new Web-based newsletter.

The October issue of CompAct is special because we still offer a print version, but we encourage you to check out the Web-based version. The HTML newsletter offers many more features than the traditional print and PDF versions. For starters, in the Web-based format, we have three useful tools at the top of the page:

- Sharing

- Print

- Search back issues

The last two are fairly self-explanatory and simply add a level of convenience that we did not have before. However, the first item, Sharing, is particularly useful. When you aim your mouse pointer at the Sharing icon, you'll be presented with a concise menu of popular social networking and social bookmarking options, such asTwitter, LinkedIn and more.

Sharing articles via social networks is different from traditional

Council

SOA Staff
Meg Weber, Staff Partner

Sue Martz,
Section Specialist

Sam Phillips, Staff Editor

sharing methods like e-mail. With social networks, you don't necessarily know (or need to know) with whom you're sharing. This is actually a good thing–both for you and others. A regular habit of sharing information has at least three key benefits:

- Network building. Even before the advent of online social networks, most people would agree that a well-developed and actively maintained network is one of your best assets as a professional. It's true whether you're an actuary or a comedian. Maybe some of us unintentionally wear both hats some days? Over time, sharing interesting articles attracts others that have interests similar to yours.

- Education. Sharing helps inform others, even those that you do not know.

- Promotion. Sharing SOA content online promotes not only your brand but also the SOA. You are building an awareness with existing members, future members and within non-actuarial disciplines that could complement actuarial work.

In particular, LinkedIn is one of the best professional social networking sites around. Sharing articles within LinkedIn groups is a great idea. Most SOA sections now have LinkedIn groups. If you find an article in CompAct that interests you, I would encourage you to share it in the Technology Section LinkedIn group (and others). Even better, share it and offer your own thoughts and comments. Doing so may lead to additional discussion on the topic and incorporate the ideas of others.

In summary, it's now easier than ever to facilitate conversation around technology topics important to you and other members of the Technology Section because sharing is step one, and you can now do that with a simple click of the mouse.

Our thanks to all of the authors contributing to the newsletter. Also a special thanks to Richard Junker, who has offered to help edit articles, and followed through with comments and suggestions. We encourage you to share this resource with people you think may be interested, and to provide us (or the authors) with feedback. For example, the sharing link was suggested by Andrew Chan–thanks!

Actuaries

475 North Martingale Road, Suite 600 Schaumburg, Illinois 60173

Phone: 847.706.3500 Fax: 847.706.3599 www.soa.org

ISSUE 37 | OCTOBER 2010

SOCIETY OF ACTUARIES   Technology Section

CompAct

ELECTRONIC NEWSLETTER

CONTENTS

QUICK LINKS

Technology Section Web site

Share   Print Article   Search Back issues

OUR EXPERIENCES WITH AGILE

By Robert Ream and Justin Bozonier

MG-ALFA is financial projection software developed and supported by Milliman and used by leading life insurance and financial firms worldwide to perform financial projections. Among other things, our product is known for consistently delivering quarterly software releases.

Historically these regular, content-rich releases were delivered with a small team of very talented developers. We were an agile development team, but it was by accident. We were not aware of why we were so effective, so as we tried to expand, our team struggled to maintain the agility. We decided to evaluate our development process from the ground up, and allow for the possibility of radical change. After several team members and our stakeholders read *The Art of Agile Development* by James Shore, we decided to bring James in to help us understand how to regain our effectiveness with our larger team. Often, moving to an Agile approach is difficult to sell to the stakeholders, but because the core principles and expected outcomes of Agile software development were consistent with the lost expectations of the stakeholders, it was an easy decision.

In today's software development world there are few buzz words that are more overloaded and overused than "Agile." Coined by a group of 17 thought leaders in "lightweight" software development methodologies, this "Agile Alliance" authored what became known as

the "The Agile Manifesto." It is a very succinct statement, so we have included it here:

"We are uncovering better ways of developing software by doing it and helping others do it. Through this work we have come to value:

> Individuals and interactions over processes and tools Working software over comprehensive documentation Customer collaboration over contract negotiation Responding to change over following a plan

That is, while there is value in the items on the right, we value the items on the left more."

–The Agile Alliance

Our Solution

Before James would commit to coming in, he talked with every person individually to assess his or her comfort level with changing core processes and communication styles. Taking on this challenge took a tremendous act of faith, courage and hard work on everyone's part, from the engineers, eager to provide the required flow of software, all the way up to our stakeholders, anxious to the regain our stride and take advantage of emerging market opportunities.

The solution we committed to as a team was to restructure how we were working and communicating together. James helped us learn how to rapidly respond to change and keep our software in a constantly releasable state, and how to communicate as a single indivisible unit.

The first step in our transformative process was to construct a vision with our stakeholders. Vision statements are often broad and banal, and consequently are also often ignored. The vision we created was tangible and focused on the next six months of product development. It also focused on the actual challenges we were facing; it was specific. Simply writing down this direct and honest statement did something our team hadn't experienced in too long: It aligned our individual goals and gave us a common purpose.

Now we all knew where we were heading over the next six months, but how were we going to get there? The vision that we laid out was concrete enough to ensure we all knew what would make the next six months a success, but was not specific enough to be

independently actionable.

The vision was further divided up into Minimally Marketable Features (MMF). These are the smallest features we could possibly deliver to our customers while still providing value. These features formed the road map to our destination. As you can imagine, there is a large list of MMFs, but with our vision clearly articulated, we can more easily prioritize them and identify how each will contribute to achieving our vision. While this seems simple, we realized that prior to implementing this process we had been assuming a common vision and understanding. That had been true when our team was small and had a long history, but we had lost that during our growth without realizing it had happened, or the consequences.

Even though we now had features clearly defined and their value established, they were still rather large chunks of work. How would we know when these were done? How could we give our stakeholders visibility into our progress while we worked on these features that might take six to eight weeks to complete? Even more important, how can we ensure that we are delivering what our customers want?

For this James Shore introduced us to "Stories" and "embedded customers." Stories are ideally the smallest possible body of work that has any real value. It acts as a sort of visible checkpoint so that your customers have a real and testable mark for where you are. In an ideal world, every Story would be an MMF. To ensure we are delivering what our customers want, our team includes proxies for our customers–embedded customers. Having users included in the team room with the programmers provides subject matter expertise and feedback in real time. This has significantly reduced rework and developer downtime related to waiting for input from the business experts.

Many times we found that we had created more stories than we really needed to finish an MMF. This allowed us to "Maximize Work Not Done" and focus like a laser on the value we were actually delivering to our customer base. Previous to our transition to Agile development, our checkpoints had such a long interval between them and access to business experts was less convenient, so decisions on value were often made by the software engineers. The issue here, of course, is that developers don't always have the same context as the business customers, and therefore may make suboptimal decisions.

Prioritizing which MMFs to work on was and still is a team decision

for us. Our embedded customers ultimately have the final say, but the cost estimates that the development team provides are taken into consideration and allow the stakeholders to make fully informed decisions. If a story will take three weeks to be completed and there is only one week left in a release cycle, it is in our best interest to prioritize some smaller stories to fill the gap.

Providing these estimates was a challenge, but we have become more adept and predictable during our transition process. Our developers estimate the stories in "Ideal Pair Days." In doing this we get a consistent idea of how long it would take a pair of engineers pair programming to finish the story at hand. These estimates also enable our business to predict how much work we will get done and enable a good amount of risk management.

To recap: We had a vision, a set of features that would help us achieve our vision upon completion, and a set of embedded customer oriented checkpoints that would help to guide us along our way. The only thing left was for us to plan out the individual tasks for the software engineers on the team. You might think that this is where the embedded customers can check out and go about their day, but that would be dead wrong.

While the customers did work on other tasks while the planning was going, they remained immediately available for questions from our engineers. Where debates would occur regarding importance of UI look and feel, the customers could quickly squash them by making a value decision. Where engineers may fret over concerns about performance, customers can provide guidance for acceptable performance levels with engineer feedback regarding the costs of solutions.

Finally, we got to work. As the weeks of the release cycle ticked by, we noted how much value we had delivered to our embedded customers with what we call "velocity." The value of velocity is driven by the estimates given by the programmers ,and how it is used can be somewhat confusing. For example, just because your velocity is steadily increasing does NOT mean your team is improving in any way. One of the key ideas behind these Agile methods is that they are not a silver bullet. James warned us that just following the "rules" by rote memorization wouldn't solve our problems. He encouraged us to always ask "why?" What are some reasons other than team improvement that might cause our velocity to go up? Perhaps we scheduled a number of stories with a high variance in the same iteration. Or maybe we were overly pessimistic in our estimates. On the other hand, what are some of the reasons velocity might

decrease, other than degradation in team performance? Since we are measuring velocity each week, absences for illness or vacations can lead to a drop in velocity.

What did we do when key developers were out sick or away on vacation? We mitigated this risk by having all of our developers pair program on all production code that is written. Pair programming was the one practice about which our stakeholders were most wary. Why would we pay for two people to do the work of one? However, it did not take long for everyone to appreciate the benefits of pair programming–including reduced cost.

Not only does pair programming eliminate the risk of knowledge silos as mentioned above, it also produces better designed and higher quality code. Many of our business experts are pairing with each other as they work on client projects, presentations and other non-development tasks. Everyone has recognized that working collaboratively to solve problems, whether through writing code or writing a report, is productive and ultimately creates a higher quality product.

Another technique we use to ensure a higher quality product is Test Driven Development (TDD). TDD centers around the idea of writing a failing test prior to writing code. In this way we express what the expectations of our next change are, then hypothesize and implement the small isolated changes to the system that we think are necessary to get the test to pass. If our hypothesis is verified and the code passes the test, we refactor the code to make it as clean as possible and then move on to the next test. We made a concerted effort to develop all new code in this manner. Furthermore, we refactored legacy code to allow for this style of development when possible.

In fact, what we found is that programmers who were less familiar in a particular code base asked questions that quickly led to "aha" moments for those more experienced in those areas. You can almost always find that someone with fresh eyes for the problem at hand provides valuable insights. To maximize the advantages of pair programming, we gave our whole team equal code ownership. We have no architect and no software designer, just a team of passionate engineers who are encouraged to discuss new system designs and architectures together to formulate the future direction of our product.

Looking Back

Looking back on our journey until now, it would be best characterized as always questioning what "done" looked like. Our stakeholders had a vision and they communicated what it would like when we achieved it in a real and measurable way. We worked with our business team and described the stories that would need to be done to have a complete feature. Everything, down to each line of code where a failing test suddenly passes, shows us that we've completed yet another step in our journey.

This is how Agile software development works for us. It's a process of continuous change and improvement. We must always make sure we have time to experiment and never assume we should not change the process. And most importantly, we must remain focused on business value. We've taken what James Shore has taught us and made it our own with our own successes to show for it. That is really the heart of the lesson in the end. Find what works best for you; keep searching with the belief that you can always do better, and that focusing on continuous improvement is the best investment you can make for individual growth, team growth and business growth.

Robert Ream is a Senior Software Simian for Milliman's MG-ALFA team, and can be contacted at robert.ream@milliman.com

Justin Bozonier is a Code Samurai for Milliman's MG-ALFA team, and can be contacted at justin.bozonier@milliman.com

Excel 2010—An Analyst's Perspective

by Andrew Chan

I have used Microsoft Excel for more than 15 years and it has been a great product that's getting better and better with every upgrade. However, I have never felt so enthusiastic about an upgrade as I did about Excel 2010; I can't stop talking to people about it and encouraging them to upgrade. Why is Excel 2010 so special? Excel 2010 allows the user to:

- Make use of more physical memory (64-bit Excel)
- Create data charts in single cells (Sparklines)
- Quickly filter in pivot tables (Slicer)
- Perform more powerful analysis (PowerPivot)

I am not going to discuss every new feature in Excel 2010; instead I will focus on the analytical features. I believe you will also feel excited after your trial with Excel 2010.

64-bit Excel

If we have more than 4 GB of memory and a 64-bit Windows operating system, then we may consider upgrading to the 2010 64-bit edition, which allows us to create bigger Excel spreadsheets (more than 2 GB).

I normally don't recommend that Excel users create huge workbooks because it can cause problems. Instead, I encourage them to move large data sets to a database. However, I understand some organizations may not have a database server and use Excel worksheets as data store.

With the 64-bit edition, the 2 GB limit is gone, and we can also do a pivot table analysis on large database record sets. Before proceeding, we should review our VBA code and make sure it will work with 64-bit Excel. We will also need to acquire 64-bit versions of any ActiveX controls, COM Add-ins or XLLs if these are used in our solutions.

Sparklines

A graph is an effective way to present a large amount of data. However, data-driven graphs can also be quite large in older versions of Excel, so it is not easy to post multiple graphs on one worksheet. Microsoft Excel 2010 added sparklines to solve this problem. Sparklines are "intense, simple, word-sized graphics." Their inventor, Edward Tufte, describes them in his book *Beautiful Evidence.*

Each sparkline takes up a cell. It does not float on the grid like chart does. It does not give numeric details, but it does add context and meaning to the number next to it. Below, we see a few more examples of how sparklines can present extra information.

Creating Sparklines
- First, we insert a new column where we want to place our sparklines.

All Books			01/01/2010	02/01/2010	03/01/2010	04/01/2010
Art & Photography	$17,705.03		$726.26	$207.25	$786.05	$987.71
Children's books	$12,989.33		$60.62	$539.33	$141.11	$133.72
Compter & Internet	$16,490.16		$952.15	$706.86	$407.52	$78.62
History	$16,883.79		$734.82	$124.14	$361.57	$412.66
Mystery & Thrillers	$13,636.45		$606.37	$110.86	$498.76	$511.64
Romance	$15,838.32		$363.05	$841.39	$86.35	$125.40
Science Fiction & Fantasy	$17,417.54		$435.48	$358.26	$843.39	$868.30
Sports	$15,463.65		$769.24	$764.18	$440.59	$725.07
Total	$126,424.27		$4,647.98	$3,652.26	$3,565.34	$3,843.10

- Then we go to the insert tab on the ribbon and choose the type of **Sparklines** that we want to inset into our column.

Line   Column   Win/Loss
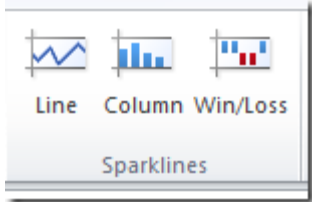
Sparklines

- We define the **Data Range** and **Location Range** in the **Create Sparklines** dialog box and click the **OK** button.

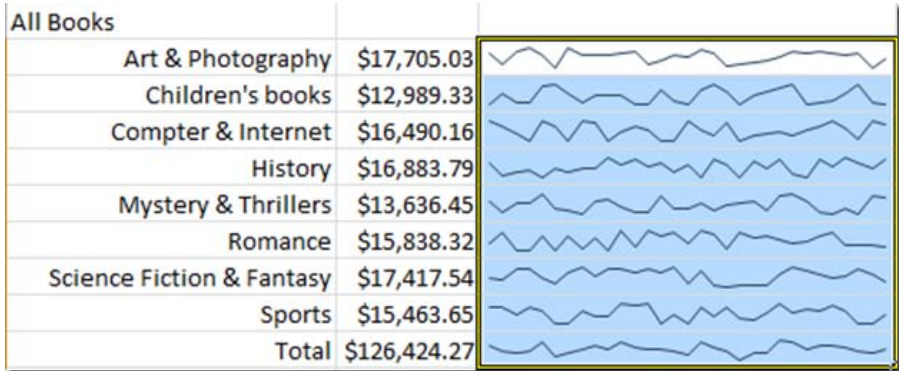| All Books			01/01/2010	02/01/2010	03/01/2010
Art & Photography	$17,705.03		$726.26	$207.25	$786.05
Children's books	$12,989.33		$60.62	$539.33	$141.11
Compter & Internet	$16,490.16		$952.15	$706.86	$407.52
History	$16,883.79				$361.57
Mystery & Thrillers	$13,636.45				$498.76
Romance	$15,838.32				$86.35
Science Fiction & Fantasy	$17,417.54				$843.39
Sports	$15,463.65				$440.59
Total	$126,424.27				$3,565.34

Create Sparklines

Choose the data that you want

Data Range: D2:AH10

Choose where you want the sparklines to be placed

Location Range: $C$2:$C$10

OK   Cancel

- Now we have our first sparklines.

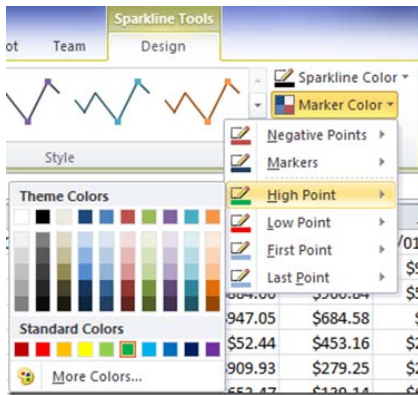| All Books		
Art & Photography	$17,705.03	
Children's books	$12,989.33	
Compter & Internet	$16,490.16	
History	$16,883.79	
Mystery & Thrillers	$13,636.45	
Romance	$15,838.32	
Science Fiction & Fantasy	$17,417.54	
Sports	$15,463.65	
Total	$126,424.27	

- We can format sparklines (e.g., highlight the high and low points on the sparklines). When we select the sparklines, we would see a sparkline contextual tab in the ribbon.

- We can highlight high and/or low points by specifying their color.


- Our sparklines would then have the new format.

All Books		
Art & Photography	$17,705.03	
Children's books	$12,989.33	
Compter & Internet	$16,490.16	
History	$16,883.79	
Mystery & Thrillers	$13,636.45	
Romance	$15,838.32	
Science Fiction & Fantasy	$17,417.54	
Sports	$15,463.65	
Total	$126,424.27	

Slicer

Slicer allows us to slice-and-dice a PivotTable. In older versions of Excel, if we wanted to use a PivotTable to do an income analysis by province and product type, we would use a report filter to select corresponding items in province and product type. But report filters do not give us a good visualization, whereas slicers show us what we have chosen.
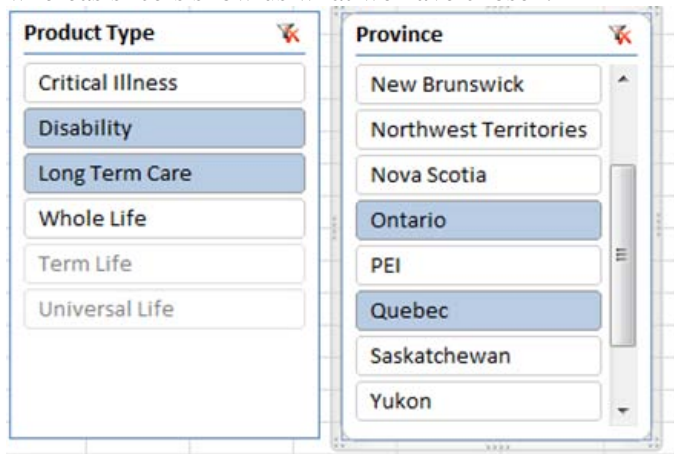
Creating a Slicer in an Existing PivotTable

- Click anywhere in the PivotTable report for which we want to create a slicer. This displays the PivotTable Tools, adding an **Options** and a **Design** tab on the ribbon.
- On the **Options** tab, in the **Sort & Filter** group, click **Insert Slicer**.

- In the **Insert Slicers** dialog box, select the check box of the PivotTable fields for which we want to create a slicer.
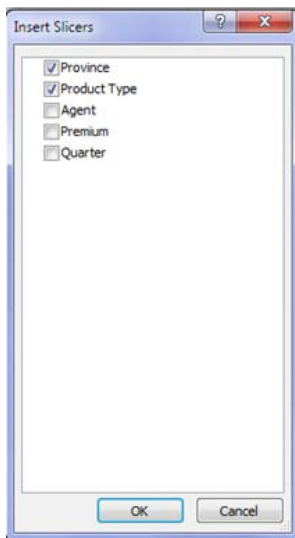
- Click the **OK** button. Now we can use the slicers to choose our provinces and product types.

The slicer tells us what we have and have not included in our PivotTable.

PowerPivot

If PowerPivot were the only new feature in Microsoft Excel 2010, I would still upgrade to Excel 2010. It is especially valuable for anyone who does not have business intelligence (BI) or an online analytical processing (OLAP) tool.

PowerPivot is a personal BI tool that enables Excel users to perform powerful data analysis using the skills they already have. It empowers advanced Excel users to transform enormous amount of data into meaningful information instantly.

It also allows us to consolidate data from many different sources. We can easily handle 10 years of expenses analysis and drill down to investigate different expense categories and territories, or even compare the income from different industries.

Here a few things that make PowerPivot shine:

- **Row-and-column limitation in Excel**—We can have more than 1 million rows of data.
- **Lightning fast**—It handles tens or hundreds of millions of rows around the same time as Excel handles tens or hundreds of thousands of row. It uses the new In-Memory BI (IMBI) database and a great compression algorithm in order to significantly reduce the size of the overall data stored.
- **Data from multiple sources**—We can load data from databases (Access, SQL Server, Oracle, Terradata, DB2 and Sybase), Excel, text files, OLAP Cube, reports and even data feed—virtually anything that we can think of.
- **Data relationship**—PowerPivot provides a data relationship layer that lets us integrate data from different sources so we can analyze all our data as it is from a single source; no more VLookup!
- **Powerful new analytical capabilities**—Data Analysis Expressions (DAX) is designed to work with multiple tables of data. DAX also adds new functionality that allows us to create dynamic aggregations. This makes DAX formulas smart about calculating values in a PivotTable.
- **Time series calculations in pivots**—There are about 40 DAX functions to handle time (e.g. Year Over Year Growth, Previous Period and more).

After we install PowerPivot, a new item appears on the Excel 2010 ribbon:
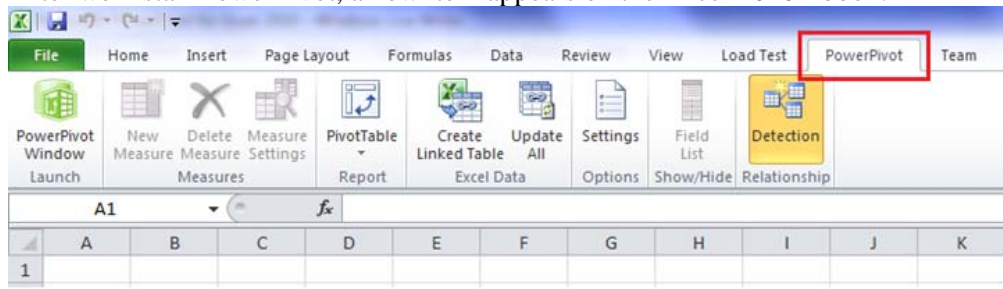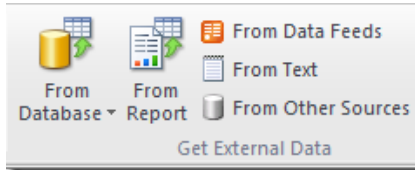

After we click on the PowerPivot Window Launch button, we would see new window come up.
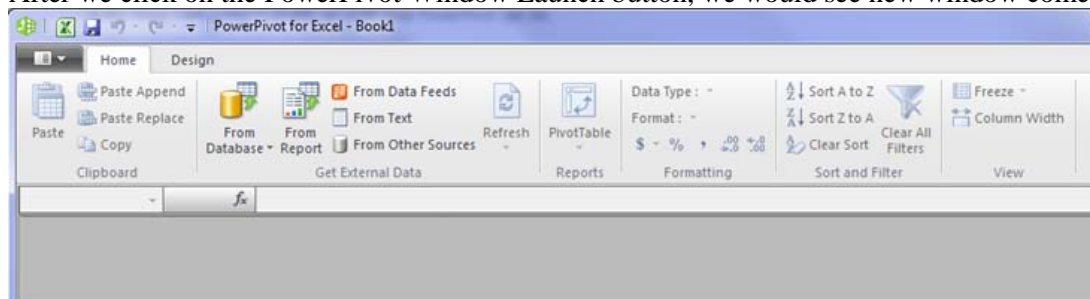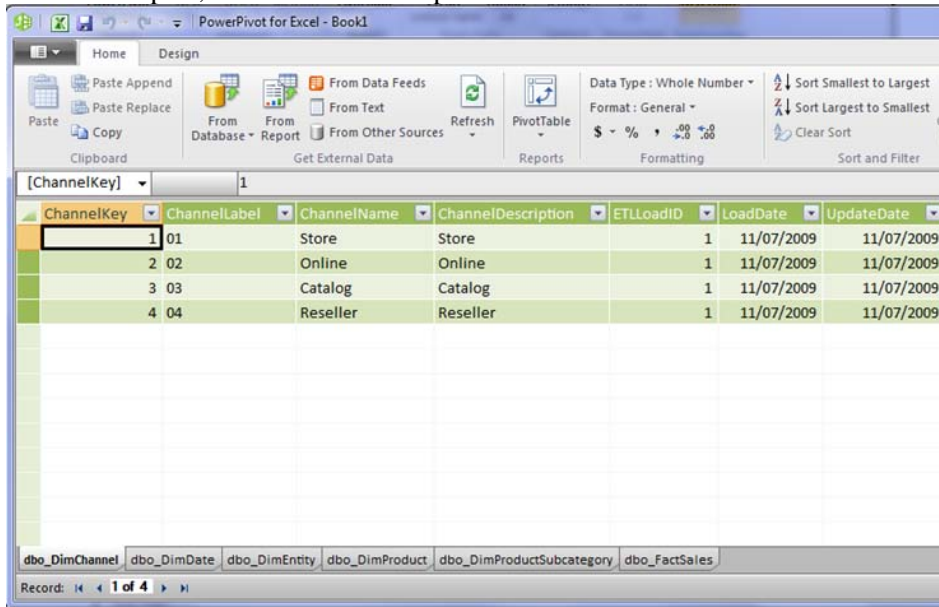

Importing Data

We may have data in many different places, such as database, Excel, reports or even from the internet. Microsoft PowerPivot offers five different ways to get external data:

- Database
- Reports
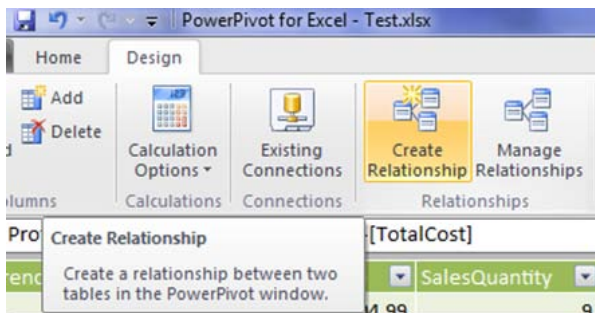- Data Feeds
- Text Files
- Other Sources

After the import, each table shows up as a tab in the PowerPivot client window.
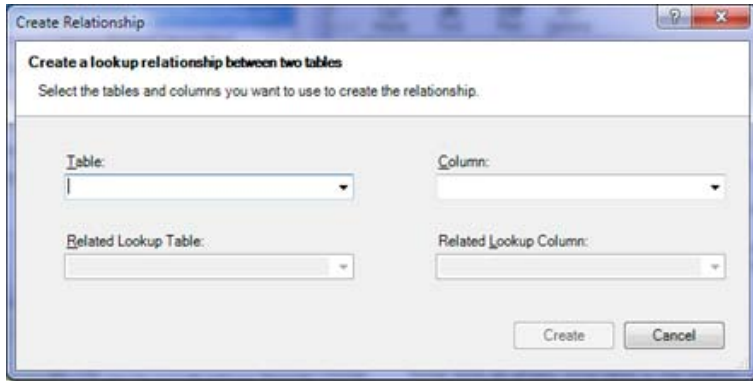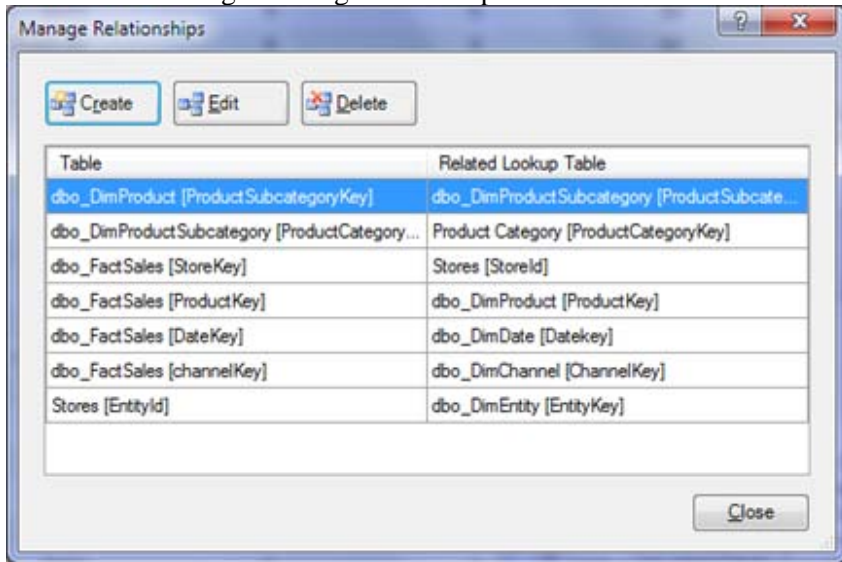

Managing Relationships

After data are imported, the next step is to define the relationships among tables. We can create a relationship and give it a meaningful name.

After we click the **Create Relationship** button, we can specify the relationship in the **Create Relationship** dialog box.

We can also manage existing relationships.

Presenting Information

Once our data are ready and relationships have been properly established, we can create a PivotTable to do analysis or a PivotChart to present information. This is an environment with which most of us are very familiar. PowerPivot enhances the Excel experience by providing quick templates for frequently used layouts:
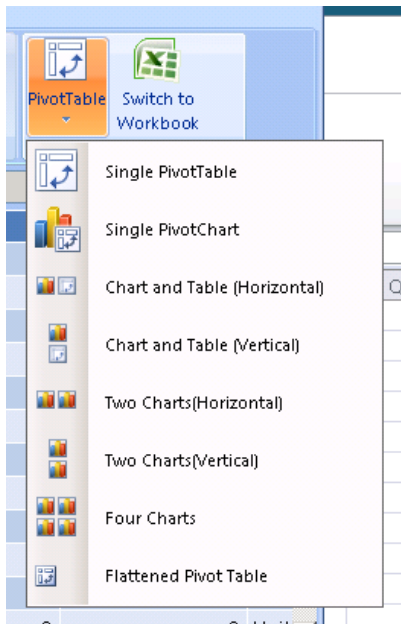
However, since our data are stored in PowerPivot, we must start the PivotTable within the PowerPivot. PowerPivot overrides the default PivotTable Field List by providing its own Task Pane:
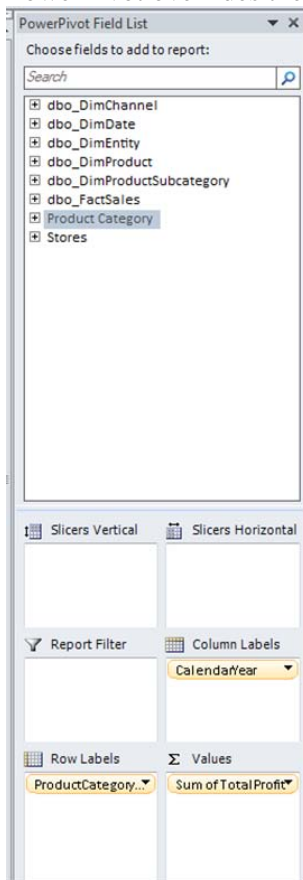
PivotTable

We probably can't tell if the following PivotTable is from Excel or PowerPivot.

| Profit by Category	Calendar Year			
Category	2007	2008	2009	Grand Total
AUDIO	17,328,091	30,042,134	38,457,341	85,827,566
CAMERAS & CAMCORDERS	656,344,901	481,290,164	374,034,900	1,511,669,965
CELL PHONES	199,930,763	138,824,926	150,147,401	488,903,090
COMPUTERS	642,423,219	553,828,147	603,940,141	1,800,191,506
GAMES & TOYS	24,327,480	22,357,334	33,997,881	80,682,695
HOME APPLIANCES	759,248,495	780,214,564	633,114,078	2,172,577,136
MUSIC, MOVIES & AUDIO BOOKS	45,253,575	32,096,888	21,994,970	99,345,433
TV & VIDEO	223,938,776	254,995,837	251,862,862	730,797,476
Grand Total	2,568,795,300	2,293,649,994	2,107,549,574	6,969,994,867

PivotChart
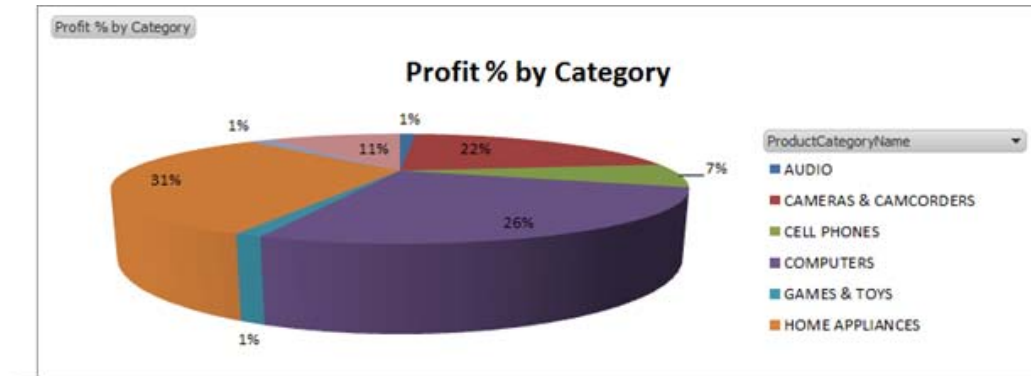
Slicer can be used with PowerPivot as well.
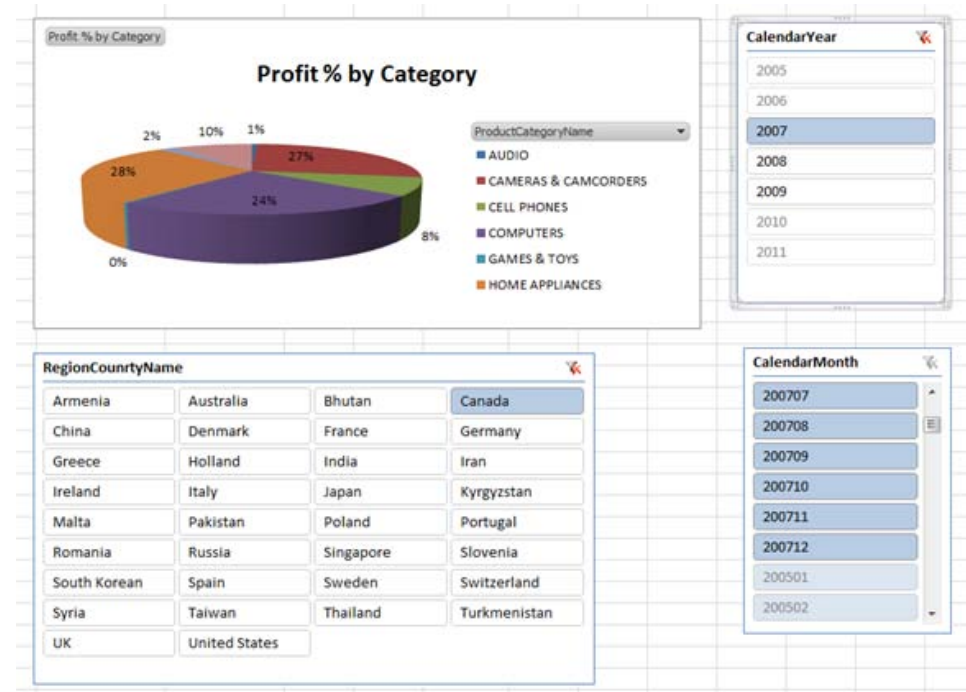
Conclusion

I only showed some of the most outstanding new analytical features in Excel 2010. There are also enhancements to many existing analytical features in Excel: new icons and enhanced data bars in

conditional formatting, new search filter for PivotTable and a new genetic Evolutionary Solver. I will cover these in a future article.

ISSUE 37 | OCTOBER 2010

SOCIETY OF ACTUARIES — Technology Section

CompAct
ELECTRONIC NEWSLETTER

Share | Print Article | Search Back issues

ACTUARIES LOOKING TO THE SCAI FOR ANSWERS
by Van Beach, FSA, MAAA

When running stochastic projections, the actuaries' need for computing power is insatiable. A multi-tier Scalable Cloud Actuarial Infrastructure (SCAI) may hold the solution.

Imagine the following scenario: Results are needed immediately, but the in-house compute cluster (i.e., the grid of servers used to parallelize computations) is already fully utilized and there is a queue of jobs that will keep the resources occupied for days. A job could get bumped, or priorities reordered, but even with all 200 computing resources (cores) in the cluster available, running another 1,000 scenarios will require more than 10 hours to complete.

The 10-hour run time is certainly a dramatic improvement over running on a single computer (which was the situation before the compute cluster was installed), but is still far from ideal.

The drivers are well-known–greater reliance on actuarial models for risk management and reporting, more complex models, stochastic methodologies, etc. The same situation seems to play out every quarter-end–everyone needs the cluster at the same time. There is no indication the problem will go away.

A better solution is needed.

Instead of 200 cores available 365 days a year, a more effective distribution might be 2,000 cores 36.5 days a year. At certain times,

20,000 cores for 3.65 days would be the optimal configuration. If computing capacity were marginal cost, these would all be equivalent. In addition to marginal cost, the ideal solution would also provide infinite, on demand, high performance resources. While still nirvana, several Scalable Cloud Actuarial Infrastructure (SCAI) options make this vision more than just a daydream.

Clouds and SCAI

Before proceeding further, some additional definitions are required. While there are many definitions of a "cloud," I prefer an inclusive version, which suggests that a cloud is a computing resource–server, storage, software, etc.–that is available via a network, often an Internet connection. Typically, the cloud operates on an on demand, pay-per-use business model and provides the illusion of being an infinite resource. The cloud can be internal or external (i.e., provided by a third party).

I'm defining a "SCAI" (pronounced like "sky") to be a subset of the resources available in the cloud. First, a SCAI is an infrastructure (servers, storage, network, etc.), rather than a software application. The software can be initiated from the desktop, but is executed in the SCAI. Second, the SCAI can be used to parallelize and distribute actuarial projections. Third, the SCAI resources are effective in supporting actuarial calculations (not all infrastructures can handle the unique performance requirements of actuarial models).

In summary, a cloud infrastructure that can be leveraged for parallelizing actuarial calculations is a SCAI. By this definition, an in-house cluster of computers is a SCAI: a bank of servers, centrally available via a network, used to distribute calculations for actuarial models.

I'm going to discuss four computing options that I believe will play key roles in providing the flexibility and capacity to meet actuarial demands as part of a SCAI:

1. Multi-core computers

2. In-house compute clusters

3. Third-party clusters

4. Cloud platforms

Multi-core Computers

The power packed into a desktop computer continues to spiral

upward, largely due to additional processors and cores. These create the potential for a self-contained "cluster" where actuarial projections can be parallelized across the cores. No sharing or queuing is required when the cluster is your desktop, and it is often much easier and more efficient to run a projection locally rather than trying to schedule time on a shared cluster of computers.

If budgets don't allow each user to have a multi-core computer, a common practice is to have a bank of shared "work" machines that are available via Windows Remote Desktop, essentially creating a local cloud of work machines. With this arrangement, users log into the more powerful work machines when projections need to be run. If the data reside on a local network, moving the job execution from your personal computer to a shared work computer should be a snap.

In-house Compute Clusters

In recent years, the prevalence of in-house compute clusters to support actuarial calculations has continued to increase, providing a computational "backbone" for actuarial projection work. The key innovation that triggered this expansion was the introduction of Microsoft HPC: a cost-effective, centralized operating system that turns a bank of servers into a scalable resource that can be made broadly available.

The introduction of Microsoft HPC effectively removed the cost hurdle for implementing industrial-strength distribution, scheduling, and cluster management software. While hardware costs and IT support are still issues that must be addressed, many companies have been able to implement small clusters to greatly alleviate computing bottlenecks for their actuaries. For most companies, an in-house cluster–a fixed set of dedicated resources, maintained internally, and managed via Microsoft HPC–is their primary source of actuarial computing capacity.

Third-party Clusters

The concept of providing a shared compute cluster that many companies can leverage during periods of peak demand has a lot of intuitive appeal–until met with the practicalities of actuarial computing requirements. The biggest challenges to this business model have been the unique performance profiles of actuarial projections and the correlation of peak demand across all potential customers; everyone needs the extra capacity during the same periods during the year, so there is no way to fill in "valleys" with

another company's "peaks."

However, there are examples of vendors who are proving to be quite successful in providing their infrastructure as a service offering to insurers. An example is R Systems NA, Inc., in Champaign, Ill., which highlights several key traits required to be a viable alternative in this market.

First, they are able to provide flexible scheduling to address peak demand periods. While this has been a stumbling block for others, R Systems provides their services to multiple verticals, thus allowing insurance peaks and valleys to be offset by other industries' uncorrelated usage patterns. Second, R Systems' hardware is current, turns over often and is tailored for high performance computing. Third, they have gained an understanding of the unique performance characteristics of actuarial models and are quite skilled in tuning their environment for high performance. Finally, their costs create a solid value proposition, as compared with in-house computing resources.

Using a third-party comes with risks–counterparty, availability, security, etc.–but the potential to augment an in-house cluster with additional capacity for peak periods has tremendous value for many insurers.

Cloud Platforms

Taking the infrastructure as a service concept one step further, a cloud platform (a platform as a service) provides the hardware, storage, network, availability, support and more, all in a pay-as-you-use package. For actuarial computing needs, however, not all platforms are created equal. The definition of a cloud infrastructure suggests the illusion of being infinite. The reality is that not all clouds can maintain this illusion– especially as the resource demands increase.

When considering the computational demands of actuarial projections–thousands of scenarios per projection, multiple projections per user, several users per company across potentially hundreds of companies, it is clear that only the largest cloud platforms will be viable options. Microsoft Azure is one cloud platform that has potential to be part of the actuarial computing capacity solution. It is still in its relative infancy, but Microsoft's commitment to providing this service and the trajectory of resources being brought online suggests Azure will be a viable for the long term.

Further, the potential to provide (the illusion of) infinite, on demand, marginal cost resources is clearly a game changer when trying to meet actuarial computing demands.

Putting It All Together

The supposition might be that a company should choose one of the above options for their SCAI. What I would suggest, however, is that the SCAI should consist of all four: multi-core work computers, an in-house cluster, third-party cluster(s), and a cloud platform. Each provides a different value proposition and meets a different need. Simplistically, the usage might be as follows:

- Day-to-day model development and testing is done on shared work computers.

- The in-house cluster is available for on demand ad hoc model runs, larger tests, benchmarks, etc., during off-peak periods. During peak periods, the in-house cluster is used for the mission-critical runs and generally as the first tier of capacity

- The third-party cluster might be arranged to meet 80 to 90 percent of peak demand and scheduled only during peak periods. The third-party cluster provides overflow capacity needed above the steady-state levels met by the in-house cluster. Given there are often better marginal rates with advance notice, this need should be planned as much as possible.

- The cloud platform is then used for unscheduled or exceptionally heavy periods of capacity demands. It provides the final tier of capacity when other sources have been exhausted.

Going back to our original challenge:

- 200 cores available 365 days a year could be met with an in-house cluster,

- 2,000 cores 36.5 days a year could be met with a third-party cluster, and

- 20,000 cores for 3.65 days could be provided by a cloud platform.

In the case of the cloud platform, the resources are marginal cost, (provide the illusion of being) infinite, and on-demand. If high-performance resources are required, a third-party cluster can provide these. Yes, this is more than just a daydream. A SCAI that consists

of all four options provides flexibility and capacity to meet essentially any flavor of computing demand.

The Role of the Application

Each of the above–multi-core computers, in-house clusters, third-party clusters and cloud platforms–provide powerful options for meeting the actuarial capacity challenge. However, not all options will necessarily be available (or effective) for a given actuarial projection software application. There are at least two levels to consider: integration and optimization.

First, an application needs to provide an integration layer to make the computing option available. For example, there needs to be an integration layer between the application and Microsoft HPC for the Microsoft compute cluster to be available as a compute option. Third-party cluster providers might be able to use multiple integration options, but the integration layer is still required. Cloud platforms again require specific development to be able to be utilized as part of a SCAI.

It is a different question to ask whether the application has been optimized for the given computing option. For example, an application might be able to take advantage of all cores in a multi-core computer with the use of Microsoft HPC software. However, this introduces a complexity and overhead that would not be required if the application could automatically detect and utilize the available cores. Working with a cloud platform like Azure also requires a tremendous amount of effort and expertise on the part of the application vendor to provision and utilize the cloud resources optimally.

In short, the application needs to be integrated and optimized for the given computing option. Each application will have a different level of integration, optimization, and value for the given computing option. So the optimal SCAI configuration will be application specific, in addition to company specific.

Smart Modeling

Lest we forget amid all this infrastructure discussion that modeling techniques also play a role in addressing the capacity challenge. Cluster modeling and scenario reduction techniques are two examples of methodologies that are very effective in reducing computation time while maintaining a high level of accuracy.

Summary

The need for actuarial computing capacity is not going to recede. To meet regulatory and risk management requirements, companies will need to find solutions to this capacity challenge. There are multiple options–multi-core computers, in-house clusters, third-party clusters, and cloud platforms–that can be leveraged, assuming your actuarial projection software can support the option. The question is no longer "if" additional capacity is needed: The question is "how" the capacity can most effectively be provided. When considering this question, actuaries should look to the SCAI for answers.

Van Beach, FSA, MAAA, is the MG-ALFA product manager at Milliman and can be reached at van.beach@milliman.com.

Actuaries
Risk is Opportunity.®

475 North Martingale Road, Suite 600 Schaumburg, Illinois 60173
Phone: 847.706.3500 Fax: 847.706.3599 www.soa.org

ISSUE 37 | OCTOBER 2010

SOCIETY OF ACTUARIES — Technology Section

CompAct
ELECTRONIC NEWSLETTER

Share   Print Article   Search Back issues

R CORNER–MEMORY MANAGEMENT[1]

By Steve Craighead

Editor's note: R Corner is a series by Steve Craighead introducing readers to the "R" language used for statistics and modeling of data. See footnote 4 at the end of the article for information on installing and learning more about R.

Even though I'm still looking for all of my marbles, I'm going to talk about how to manage memory within R.

R has been built to run in 32-bit and 64-bit environments. Most of the time, the limits to memory are related to how R is built, but memory in the Windows' 32-bit R depends on what operating system is being used.

Usually, you know you have a memory issue when you get an error message like this: "Cannot allocate vector of ... "

Old Technology

Traditionally, I have managed the memory in R by using the two embedded properties "memory.size" and "memory.limit."

You use the memory.size command to determine how much current memory is available and how much can be made available.

For instance, "memory.size()" returns the current memory limit in that workbook, while "memory.size(TRUE)" returns the maximum size that the operating system can supply.

For instance, on my 32-bit Windows machine I get the following:

```
> memory.size()
[1] 20.47
> memory.size(TRUE)
[1] 23.44
>
```

Now, "memory.limit()" returns the memory limit, while "memory.limit(nnnn)," where "nnnn" is the number of megabtyes requested as a new limit. This is limited to 4096 Mb on 32-bit R builds.

On my machine, I get the following results:

```
>memory.limit()
[1] 1535
> memory.limit(3000)
[1] 3000
> memory.limit()
[1] 3000
```

So, now my memory limit has been increased to 3-Gig. When the memory.limit exceeds the max memory.size, R will still run, but the extra memory will be paged off and onto your hard drive. This means that you can use a large dataset, but processing time will go up dramatically since your hard drive doesn't run as fast as your computer's memory.

If you do use Windows, you may find that you can increase the operating system limit from 2Gig to 3Gig by using the "/3GB" switch in your "Boot.Ini" file. Please visit this Web address for more information.

I have used all of the above methods to great success. I was able to load million-record files into R memory through these techniques. If you review my earlier article on how to get data into R, you can also use the RODBC package to create SQL queries against large databases to extract what you need for your study.

New Technology–bigmemory[2]

Michael J. Kane and John W. Emerson developed a new matrix class called "big.matrix." The packages bigmemory, biganalytics,

synchronicity, bigalgebra, biglm and bigtabulate use this new class and allow you to manipulate larger and larger datasets within R. These packages are designed to run in parallel processes as well.

It took R about five seconds to create a five million row, three column double matrix with this command[3]:

```
Mymat
<- big.matrix(5000000,3,type='double',init=0,dimnames=list
(NULL,,c("first","second","third")))
```

It took less than a second to put 5 million random normal samples in the first column by this command:

```
Mymat[,"first"]<-rnorm(5000000,0,2)
```

It took a little bit longer to store 15 million random samples in the entire matrix by this command:

```
Mymat[,] <- rnorm(15000000,0,2)
```

You can read and write to big matrices as well. For instance, I'll first write Mymat out to a file "test.txt" and then I'll read it back in.

I was able to write this out in less than 30 seconds with this command:

```
write.big.matrix(Mymat,"test.txt")
```

The output file "test.txt" was comma delimited with three columns and 5 million rows.

Reading took less than 30 seconds when using this command:

```
MyMat2 <- read.big.matrix("test.txt",type="double")
```

There are many other commands within bigmemory that will allow you to share the same big.matrix object across multiple processors on your computer or across a cluster. I won't go into further detail regarding bigmemory, but if you are interested, please refer to Bigmemory.org for more information.

I also have loaded the biganalytics and the biglm packages to analyze Mymat.

Below are some summary statistics on my 15 million random normal draws:

```
> mean(Mymat)
[1] 0.0006737719
> colmean(Mymat)
```

first	second	third
0.0001679818	0.0021683848	-0.0003150509

```
> summary(Mymat)
```

	min	max	mean	NAs
first	-1.093267e+01	9.627050e+00	1.679818e-04	0.000000e+00
second	-1.088475e+01	9.889715e+00	2.168385e-03	0.000000e+00
third	-1.026170e+01	1.038852e+01	-3.150509e-04	0.000000e+00

I don't recommend this, but if you want the row averages you could run this command:

```
apply[Mymat,1,mean)
```

You can use the biglm function from the biglm package to determine a linear regression model on a big.matrix. For instance:

```
Mymat.biglm <- biglm.big.matrix(first ~ second + third,
data=Mymat)
```

This command ran in under 20 seconds and the results are:

Large data regression model: biglm(formula = formula, data = data, ...)

```
Sample size = 5000000
```

	Coef	(95%	CI)	SE	p
(Intercept)	2e-04	-0.0016	0.0020	9e-04	0.8500
second	-5e-04	-0.0014	0.0004	4e-04	0.3051
third	5e-04	-0.0004	0.0014	4e-04	0.2240

The above model is not at all significant, as was expected.

According to Kane and Emerson, you can expand these various packages to include other analytics.

I hope that you can find a use for big.matrix in your future work!

Footnotes

[1] Parts of this article are based on the "Memory Limits in R" R documentation. This documentation can be obtained in R by using the command: help("Memory-limits")

[2] Parts of this section are based on the R Documentation on bigmemory. Please use the command: help("bigmemory") to learn more about theses packages.

[3] If you copy these and subsequent commands that contain double quotes (") into R, you will need revise them to match the double quotes in R. This also applies to commands with single quotes.

[4] The first column of R Corner was published in the October 2008 issue, and explains how to download and install the package, as well as providing a basic introduction to the language. Refer to each CompAct issue since then for additional articles in the series. The introductory article can be found on p. 24 of the October 2008 issue on the SOA Web site. The R language is a very popular open source statistical modeling environment, which you can obtain from the webpage http://cran.r-project.org. If you are interested in building the toy model in this article, you will need to download the base R Binary for the operating system of your choice, (Windows, Linux or MAC). Also, you will need to download the contributed copula package. You can do this initially when you download the base system, or you can use the package installation facility when you start R.

ISSUE 37 | OCTOBER 2010

SOCIETY OF ACTUARIES — Technology Section

CompAct
ELECTRONIC NEWSLETTER

QUICK LINKS

Technology Section Web site

Share   Print Article   Search Back issues

THE 9TH SPECULATIVE FICTION CONTEST

by Gary Lange

Some of us dream about entering a blue alien body, complete with a tail, on the moon Pandora, while others hope their hot tub turns into a time machine, and even a few of us want to don the outfit that would make us Iron Man. The rest of you though, have more practical ideas about the future, and many of them revolve around our favorite topic, actuarial science! Even as you read this, you may not be aware that you have many different thoughts on various aspects of actuarial science that you should share. Stop reading for a minute, and if you have an actuarial thought, jot it down. Now that you are back, read the rest of this page to see what you can do with your idea, and how it might even make you famous, at least among actuaries.

First, you want to put your idea in the form of an exciting story, so that once a reader starts reading, the reader won't be able to put down your story until he or she has digested all of the details. Second, with such creativity bubbling off the page, you should probably send me a copy of the story, so I can share it with the world and help the future arrive a lot sooner! Third, you should encourage all of your friends to share their ideas through this contest, too!

So for all of the Avatar fans, time travelers, heroes and other actuaries who speculate about the future in their daydreams, now is the time to use that imagination to take a shot at describing the future of our profession, express an emotion about the changes in our jobs, tell a story, or just make that definite point that needs to be made.

Council

Links of Interest

Fiction Contest

Howard Callif, Editor

SOA Staff
Meg Weber, Staff Partner

Sue Martz,
Section Specialist

Sam Phillips, Staff Editor

Enter the 9th Actuarial Speculative Fiction contest, proudly sponsored again this year by the Society of Actuaries' Technology Section, the Actuary of the Future Section, and the Forecasting & Futurism Section, the three Sections that assist us in preparing for the future! The future comes faster and faster every day, so think ahead now and imagine what might be coming.

The rules for this writing contest are very simple, as they should be when someone decides to put speculative thoughts on paper:

The story must include some sort of actuarial topic.
The author must have passed at least one actuarial exam.
The story should be fewer than 6,000 words (if you have to count the words, you are using too many).

For those looking for incentives other than sharing their wonderful, eye-opening ideas with the rest of the world, there are prizes in this contest!

Best overall story	$200
Second best overall story	$100
Most unique use of technology in a story	$ 50
Best story emphasizing forecasting or futurism methods	$ 50
The most creative actuarial career of the future	$ 50
Readers' favorite story	A technological surprise!

The top two best entries will be determined by Dr. Bob Mielke, professor of English at Truman State University, with additional comments by your contest organizer, Gary Lange. The awards presented by each Section will be chosen by that Section Council, so be smart and read up on each Section at soa.org.

The contest begins with this announcement and ends **Jan. 31**, 2011. Winners will be announced on April 1, 2011. Submit all entries to me at garyslange@yahoo.com. Each entry will be published online at the Society of Actuaries' website.

Tell the actuaries what their world will soon look like! Start writing now!

Actuaries
Risk is Opportunity.®

475 North Martingale Road, Suite 600 Schaumburg, Illinois 60173
Phone: 847.706.3500 Fax: 847.706.3599 www.soa.org

ISSUE 37 | OCTOBER 2010

SOCIETY OF ACTUARIES   Technology Section

CompAct

ELECTRONIC NEWSLETTER

Share   Print Article   Search Back issues

ARTICLES NEEDED FOR COMPACT NEWSLETTER

Your help and participation are needed: Articles, book reviews and other information on technology useful to actuaries are needed! All articles will include a byline to give you full credit for your effort. An annual prize is awarded for the best article of the year. If you would like to submit or discuss an idea for an article, please contact Howard Callif or J. Eddie Smith.

Submissions for future editions are also welcome!

Submission deadline is December 1 for the January issue.

Thank you for your ideas!

Council

Links of Interest

Fiction Contest

Howard Callif, Editor

SOA Staff
Meg Weber, Staff Partner

Sue Martz,
Section Specialist

Sam Phillips, Staff Editor