



**SOCIETY OF
ACTUARIES**

**TECHNOLOGY
SECTION**

CompAct

ISSUE NUMBER 58 • APRIL 2017



Power BI: Reporting and Dashboards taken to the Next Level

By Tim Heng

Page 4

- 3 Letter from the Editors**
By Mark Africa and
Sean Hayward
- 4 Power BI: Reporting and
Dashboards taken to the
Next Level**
By Tim Heng
- 10 Dealing with Large CSV Files
in R**
By Jeff Heaton
- 12 Drivers for the Digitalization
of Insurance**
By Hartmut Schroth
- 16 Coding Standards and the
Efficient Model**
By Brody Lipperman
- 18 Wearable Wellness: Five
Quick Takeaways from RGA's
Fitness Tracker Study**
By Chris Falkous

CompAct

2017
SECTION
LEADERSHIP

Issue Number 58 • April 2017

Published twice a year by the
Technology Section of the Society of
Actuaries.

475 N. Martingale Road, Suite 600
Schaumburg, Ill 60173-2226
Phone: 847.706.3500 Fax: 847.706.3599
www.soa.org

This newsletter is free to section mem-
bers. Current issues are available on the
SOA website (www.soa.org).

To join the section, SOA members and
non-members can locate a membership
form on the Technology Section web-
page at [http://www.soa.org/sections/
technology/technology-landing/](http://www.soa.org/sections/technology/technology-landing/).

This publication is provided for informa-
tional and educational purposes only.
Neither the Society of Actuaries nor the
respective authors' employers make any
endorsement, representation or guar-
antee with regard to any content, and
disclaim any liability in connection with
the use or misuse of any information
provided herein. This publication should
not be construed as professional or
financial advice. Statements of fact and
opinions expressed herein are those of
the individual authors and are not neces-
sarily those of the Society of Actuaries or
the respective authors' employers.

Copyright © 2017 Society of Actuaries.
All rights reserved.

Publication Schedule

Publication Month: October
Articles Due: 7/12/17

Officers

Paul Ramirez, FSA, MAAA, Chairperson
Mark Africa, ASA, MAAA, Vice Chairperson
Frank Reynolds, FSA, FCIA, MAAA, Secretary/Treasurer

Council Members

Robert Hrischenko, FSA, MAAA
Rich Junker, FSA, MAAA
Jason Kehrborg, FSA, MAAA
Kyle Nobbe, FSA, MAAA
Wendy Liang, FSA, CERA, MAAA
Sean Hayward, FSA, MAAA

Newsletter Editor

Mark Africa, ASA, MAAA
mark.africa@aig.com

Sean Hayward, FSA, MAAA
Sean.Hayward@fisglobal.com

Program Committee Coordinators

Kyle Nobbe, FSA, MAAA
2017 Valuation Actuary Symposium Coordinator

Wendy Liang, FSA, CERA, MAAA
2017 Life & Annuity Symposium Coordinators

Paul Ramirez, FSA, MAAA
2017 Health Spring Meeting Coordinators

Robert Hrischenko, FSA, MAAA
Wendy Liang, FSA, CERA, MAAA
2017 SOA Annual Meeting & Exhibit Coordinators

SOA Staff

David Schraub, Staff Partner
dschraub@soa.org

Jane Lesch, Section Specialist
jlesch@soa.org

Julia Anderson Bauer, Publications Manager
jandersonbauer@soa.org

Sam Phillips, Staff Editor
sphillips@soa.org

Erin Pierce, Graphic Designer
epierce@soa.org

Letter from the Editors

By Mark Africa and Sean Hayward

Welcome fellow members of the Society of Actuaries Technology Section to the first edition of *CompAct* in 2017. As you can tell, this edition arrived in your mailbox, not your inbox like last year. While it may seem counter-intuitive for a technology section to move back to paper newsletters, the feedback we've received from our members is that you are more likely to read the newsletters if they are paper vs. digital. As noted in the last issue of 2016, we are moving to two issues a year, approximately six months apart. This allows our contributors time to continue with their high quality articles, without overburdening them while they do their day jobs.

In addition to our excellent contributors, Mark and I would like to extend a thank you to Paul Ramirez for his work on *CompAct* over the past year. Paul has served his term as editor, helping to round-up articles and working to ensure that our members receive valuable, insightful content. We also congratulate Paul on the role as Technology Section Chair and look forward to his leadership this year.

Mark and I welcome feedback on *CompAct* from all of our members, both related to content as well as the changes to our delivery mechanism and timing. We encourage any of you to reach out to either of us with ideas for potential articles, or if you have an interest in authoring an article. We can be reached at mark.africa@aig.com or sean.hayward@fisglobal.com.

In this latest issue of *CompAct*, we have five articles. One is a continuation of a recurring series, and we have four new contributors.

POWER BI

Following up on last issue's Power Query article, Tim Heng's latest article delves into the capabilities of Power BI, a suite of business intelligence tools that can be used to create live dashboards and drill down capabilities into data without the need to be a database expert or to constantly export results manually to spreadsheets.

DEALING WITH LARGE CSV FILES IN R

Adding to the recent theme of R related articles, Jeff Heaton from RGA has provided an insightful article on how R can be used to manage large datasets in a CSV file, including tips and tricks for navigating some of the potential memory limitations that can arise from such use.

DRIVERS FOR THE DIGITALIZATION OF INSURANCE

Harmut Schroth from SAS joins the list of new contributors with an article outlining some of the considerations involved in maximizing the use of data collected by insurers through the use of digitalization. Bringing together the myriad of data gathered across an insurer to better understand customers and risks provides insurers with a great opportunity, but requires additional consideration around data management practices that are highlighted nicely in this article.

CODING STANDARDS AND THE EFFICIENT MODEL

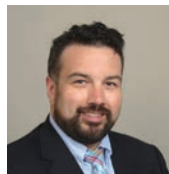
Brody Lipperman from FIS has contributed a well-articulated article outlining the need to consider more than pure speed when assessing the efficiency of a financial model. Given the significant investment insurers make in the human capital associated with their financial modeling platforms, the author states transparency, ease-of-use and maintainability need to carry significant weight as well.

WEARABLE WELLNESS: FIVE QUICK TAKEAWAYS FROM RGA'S FITNESS TRACKER STUDY

Finally, Chris Falkous from RGA has provided an interesting summary of the latest trends and considerations around the use of wearables in the insurance industry. Both drivers for usage as well as potential limitations around their inclusion in a wellness program are considered. ■



Mark Africa, ASA, MAAA, is an IT actuary at AIG. He can be reached at mark.africa@aig.com.



Sean Hayward, CFA, FSA, MAAA, is a software development actuary with FIS. He can be reached at sean.hayward@fisglobal.com.

Power BI: Reporting and Dashboards Taken to the Next Level

By Tim Heng

BUSINESS INTELLIGENCE AT YOUR FINGERTIPS

There's a certain irony about having something "at your fingertips." While the corporate jargon implies that it's always accessible and easy to use, let's be honest, how much control do you really have on something with just your fingertips?

With that said, let's get it out of the way—Power BI is business intelligence at your fingertips; at least, those fingertips that are gently balanced on your computer mouse. Those fingertips that click and drag and point at things on a screen. Or the ones that tap on your phone or tablet touch screen. If a seven-year-old can use it to build a reporting dashboard, I'm pretty sure that we can justifiably say that it's easy to access and easy to use.

This is the third and final part of the series on the Power BI suite of tools (the first two articles on Power Pivot and Power Query were in the last two newsletters, available online).

WHAT IS POWER BI?

BI stands for business intelligence and Power BI is a business intelligence tool designed to quickly take data from a range of sources, rearrange and transform the data if necessary, generate reports and distribute them to stakeholders. Historically, the challenge to do this effectively has been in overcoming the wide range of sources, data transformation tools and reporting capabilities that businesses have had at their disposal. (See Figure 1 below)

MONTH-END REPORTING

Consider a traditional business reporting framework: business data is stored across several SQL databases. These are either worked on by IT teams to create a single output to be exported into Excel or several files are exported and combined in Excel instead. There may even be an interim Access database to perform transformation and linking in between.

Once in Excel, an analyst will run calculations over the data, arrange it into tables and export these back into Hyperion or some other cube software, which may have a bespoke or other third-party add-on interface that allows management to look at the underlying report numbers.

Alternatively, charts are created in Excel that are then saved as PDFs, which are either emailed to management, or printed and put into a reporting pack, completing the gruelling two-week

Figure 1

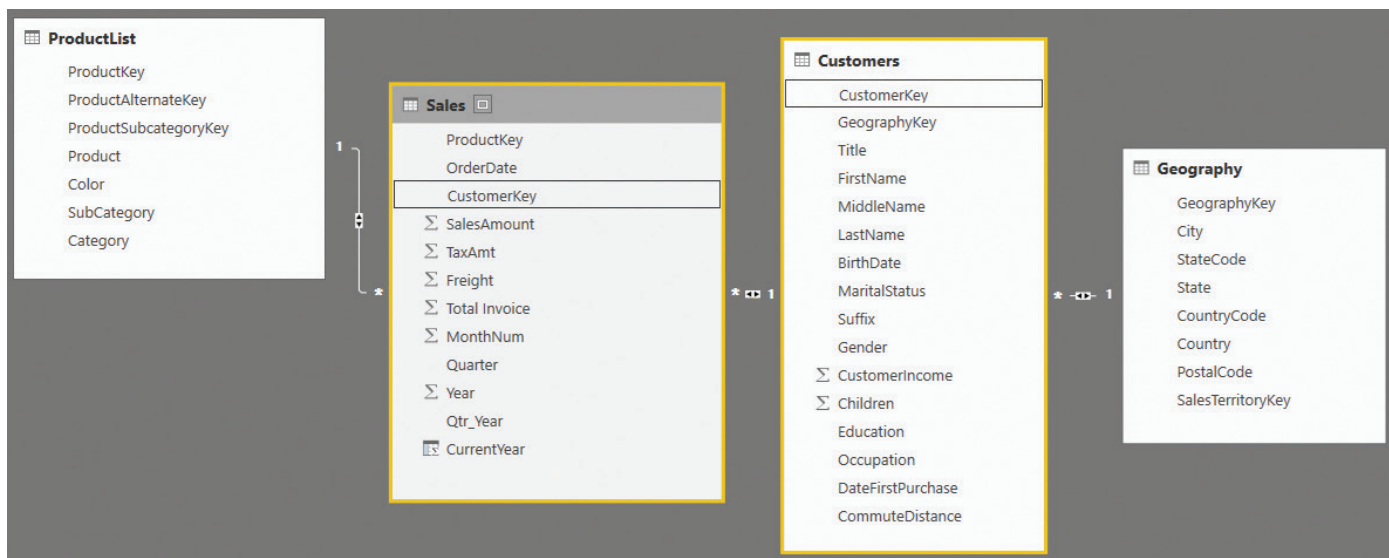
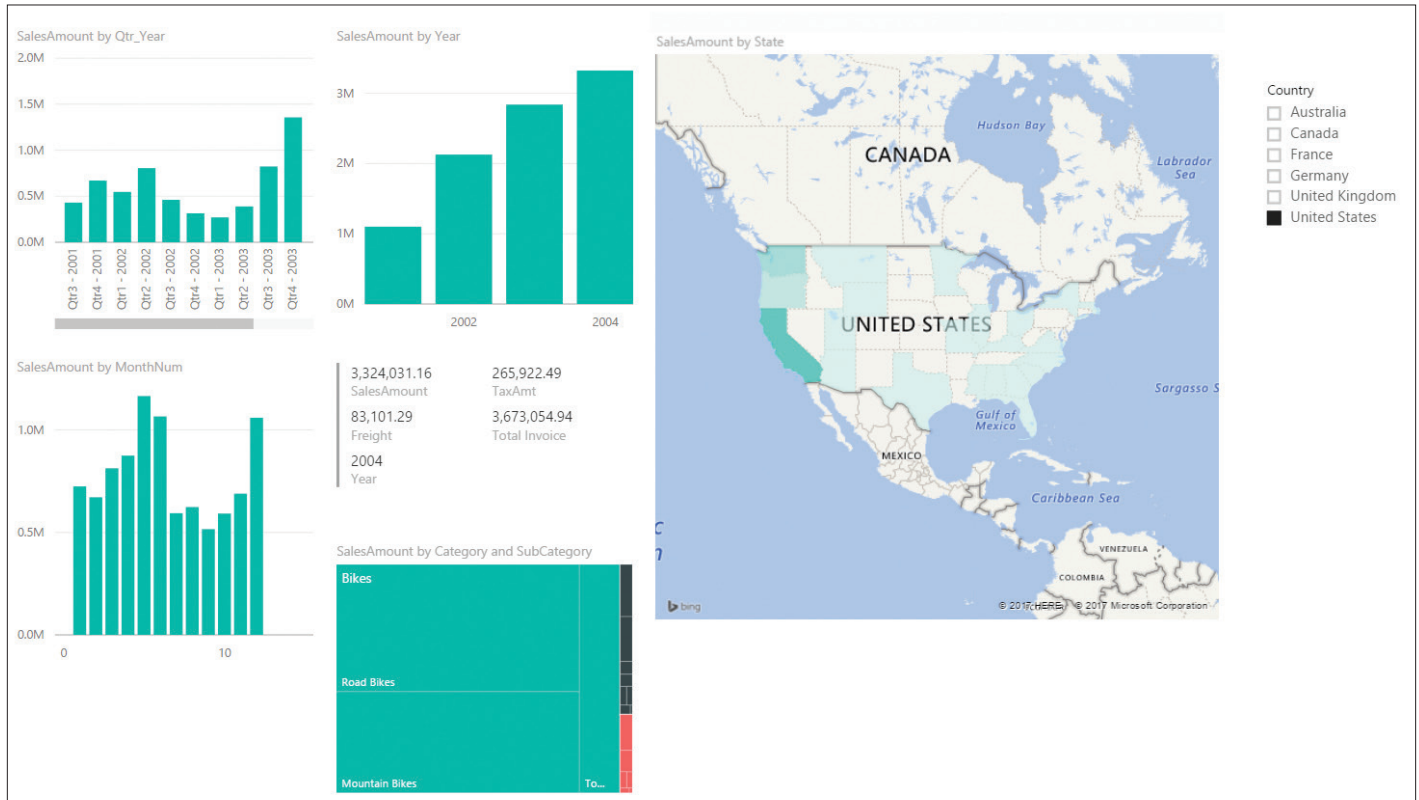


Figure 2



reporting process. Finally, this reporting process is repeated two weeks later, at the start of the next month.

KISS 2: KEEPING IT STREAMLINED, SMART

The challenge to streamline reporting processes is generally focused around removing the human interaction around reporting. By minimizing what people have to do, you're

improving not only the speed and timeliness of producing the reports, but you're also reducing the risk of errors. Generally speaking, the human parts of the reporting process have always been around getting data from A to B (moving data between systems) and processing dynamic and changing datasets using static spreadsheets that need to be updated and adjusted each month as new data comes in.

Figure 3
Quick Insights for Clean Power BI Data





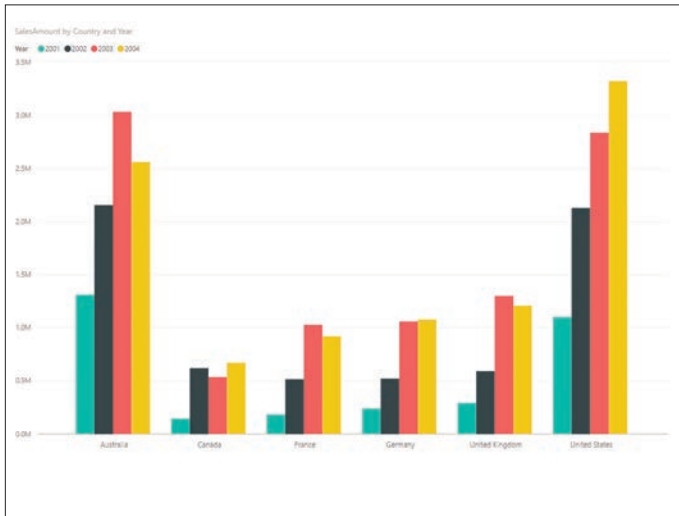
Now, how does Power BI help this process? Well, firstly, it doesn't need the specialized IT skills that database work usually requires to extract and/or combine the data you need. Secondly, by establishing a direct connection that can be scheduled to refresh, we limit the need to "export to Excel" that slows down the process. Finally, by having calculations and reports that automatically scale to size and don't require people-interaction to create or distribute, this drastically reduces the time from data being available to information being consumed by stakeholders. In fact, due to the real-time information flow, there is nothing stopping businesses from developing on-demand dashboards, rather than historical monthly dashboards. (See Figure 2, pg 5)

CLOUD ANALYTICS

The bonus to all of this is the ability to load your data up to PowerBI.com and let Microsoft's servers at it. There are two game-changing tools Power BI has at its disposal: the first is Quick Insights and the second is Q&A.

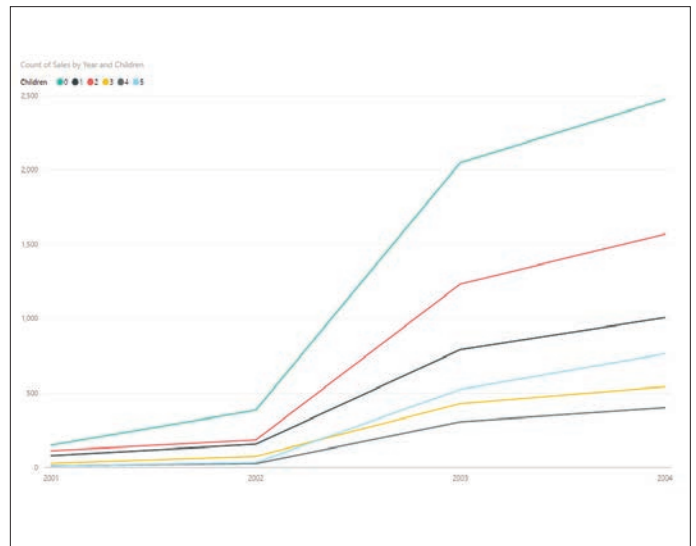
Quick Insights is an easy way to scan your data for insights and automatically displays charts and dashboard outputs. It can search for major factors, category outliers, time series outliers, determine trends and seasonality, and generally detect correlations between fields in your dataset. Once these have been generated, it is a point-and-click matter to upload these to a dashboard for ongoing monitoring and reference. (See Figure 3, pg 5)

Figure 4



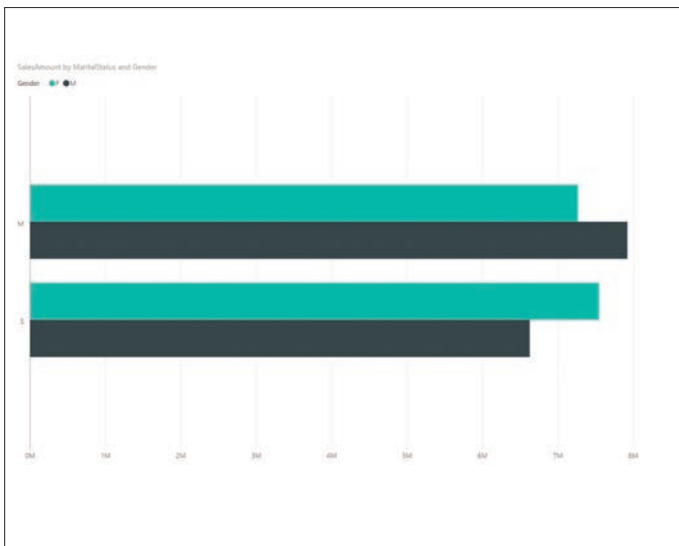
The other tool that people are raving about is Q&A, Power BI's feature to take in plain-language questions and respond

Figure 6



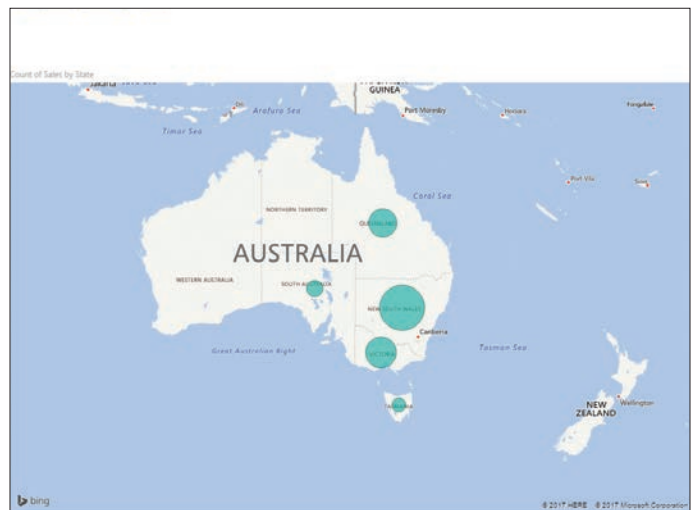
Do you have some customer data on your hands? Let's take a look at total sales by gender and marital status. (See Figure 5)

Figure 5



How about the number of sales in Australia, broken down by the number of children in the household? (See Figure 6)

Figure 7



with charts and datasets that answer the questions for you. Here, the power goes as far as your data and your imagination can take you. Let's start simple with our sample dataset—let's see what sales are, by country, over time (sorted by year). (See Figure 4, above)

Staying on the Australian front, what if we want to map sales in Australia by state? (See Figure 7)

The other benefit of Power BI is the ability to access your dashboards and reporting across a range of devices.

There are businesses out there that are now using Power BI to generate all of their reporting and dashboard solutions. Rather than exporting the data to Excel and needing to present the data nicely in tables and rows, simply asking nicely will get the answers that you want.

DASHBOARDS ANYWHERE AND ANY WAY YOU LIKE

Not only can you create these great dashboards and distribute them to users, there are a number of ways to enhance the value of these. The first is row-level security; you can restrict data access to users by writing a query script that will allow only certain rows to be accessible across specified assigned user groups. You might use this to restrict a team manager to access only data and KPIs relating to his or her team, for example.

The other benefit of Power BI is the ability to access your dashboards and reporting across a range of devices. While you may use Power BI Desktop to view your reports, or log into Power BI (online) to see dashboards, you can also download the Power BI Mobile app on your iPhone, Android or even your Windows device.

HOW DO WE GET IT?

At the time of writing, there are two main options to get access to Power BI. The simplest and cheapest approach is to sign up to the Power BI plan, which gives you access to 1GB of data capacity, and schedule refreshes on a daily basis. By cheapest, at the moment, I actually mean absolutely free. The only catch is that anyone with whom you share your dashboards will also need to have a Microsoft account as well.

The Power BI Pro plan is currently US\$9.99 per user per month. This enhances your data capacity to 10GB per user, gives you the ability to connect to your live data sources, and allows you to refresh your dataset up to eight times per day. Although this gives you an enhanced ability to use Power BI, any dashboards you create will not be able to be shared with users who do not also have a Power BI Pro plan, so you can't get away with having a single Pro license doing the grunt work and distributing insights to an entire company using the free licenses.

FINAL THOUGHTS

People keep predicting the death of Excel. Power BI is perhaps the closest thing that can lead to a decreased reliance on Excel, with connections directly to source databases negating the need to have Export to Excel as the intermediary step for reporting. However, Power BI is continuing to build up integration with Excel, including the ability to publish Excel data sets to Power BI directly from Excel 2016. There is also the likelihood that, as the Power Pivot and Power Query tools within Power BI Desktop are better understood, people will be more likely to use the same tools within Excel.

Finally, it's important to remember—what are all our tools actually designed for? Excel was not necessarily intended to be a monthly reporting tool. Excel is at its best when it's used as an agile tool that allows for dynamic and ad-hoc analysis. Tools like Power BI that can reduce the time taken to complete the reporting cycle will allow analysts to move away from day-to-day mundane tasks, giving them more time and freedom to do real analysis. ■



Tim Heng is a director at SumProduct, specializing in Excel modeling and training services. He can be contacted at tim.heng@sumproduct.com.



Be a part of the premier event for health actuaries and other health finance professions.

June 12-14, 2017
The Diplomat Resort & Spa
Hollywood, FL

2017 SOA
**Health
Meeting**

Learn from expert presenters and gain insight on the most relevant issues facing the health industry today.

Register by May 15 and save \$200.

HealthMeeting.SOA.org

Dealing with Large CSV Files in R

By Jeff Heaton

The R programming language is becoming a common tool for actuaries and data scientists to examine and model a variety of different data types. A number of useful functions are provided to load data into memory, process the dataset, and then write results to another file. Unfortunately, an additional complication can enter the picture when these files become large. If R is commanded to read a CSV that is larger than the computer's memory, an error will be returned. If you experience this, don't worry, there are a number of solutions available.

The most obvious solution is to obtain more memory. This could be more physical memory. This could also mean using part of the hard drive as virtual memory.

The most obvious solution is to obtain more memory. This could be more physical memory. This could also mean using part of the hard drive as virtual memory. These are certainly viable solutions. There are also a host of "Big Data" solutions. A multi-node Hadoop or Spark solution could be installed that allows many computers to work together to process the file. Again, this is certainly a solution, but it might not be necessary. There definitely is a class of data that are so large that "Big Data" technology is required to process them in any reasonable amount of time. However, this is often neither necessary nor economical.

Most R functions simply load an entire file into memory. This is the simplest way of handling the file, but it means that the

program can only handle data up to a certain size. Consider a simple example that illustrates the process. A large data file has been provided that contains the premium payments for customers over a potentially large span of time. Such a file might appear as follows.

```
policy_number,product,premium,month,year
J10234,term15,110,1,2010
Z10400,term10,100,1,2010
J10523,term15,110,1,2010
Z10624,term10,100,1,2010
J10234,term15,110,2,2010
Z10400,term10,100,2,2010
J10523,term15,110,2,2010
Z10624,term10,100,2,2010
...
```

To see how to handle a file of any length, consider a simple example where it is necessary to bin/roll up the premium amount by product and month. This would produce a result file similar to the following:

```
product,premium
term15,220
term10,200
...
```

A simple R program to perform this task is provided:

```
data <- read.csv("c:\\test\\sample.csv ")
result <- aggregate(data$premium,
  by=list(product=data$product, month=-
  data$month, year=data$year),
  FUN=sum)
write.csv(result, c:\\test\\sample_output.
csv ")
```

This program begins by reading the entire file "sample.csv" into the variable named "data." If this file fits into memory, everything works well and the output file is written. However, if the file does not fit into memory, an error occurs and there is no output. An alternative approach is to read the file line by line and perform the aggregation by the program. This approach is a bit more complex, but it will work on very large files.

```

# Hold all of the bins
bins <- list()

# Open the file.
fp <- file("c:\\test\\sample.csv", open =
"r")

# Skip header
readLines(fp, n = 1)

# Loop over entire file
while (length(line <- readLines(fp, n = 1))
> 0) {

  # Read a single line from the file
  line <- unlist((strsplit(line, ",")))

  # Extract the columns we care about
  product <- line[2]
  premium <- as.numeric(line[3])
  month <- as.integer(line[4])
  year <- as.integer(line[5])

  # Produce a key that holds all values we
  want to "group by"

  # Is this the first time we've seen this
  combination of month/year/product?
  key <- paste(product,month,year)
  if (key %in% names(bins)) {
    # Add to our running premium bin
    binprem <- as.numeric(bins[[key]][4])
    bins[[key]] <- list( product, year, month,
premium + binprem )
  } else {
    # Create a new premium bin
    bins[[key]] <- list( product, year, month,
premium )
  }
}
close(fp)

# Transform the bin's list into a dataframe
for output
bins <- as.data.frame(matrix(unlist(bins),
nrow=4, byrow = T))
colnames(bins) <- c('product',
'year', 'month', 'premium')

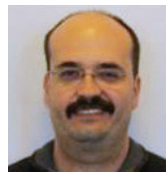
bins

```



The above code uses named lists, called “bins” to hold the value of each of our bins that aggregate product, year and month. A key is created to find the correct bin. This key is nothing more than a string, such as “term15 2010 1” to represent the bin for January 2010’s term15 premiums. Comments are provided to demonstrate the process. This short program could be a great starting point for any other situations where it is necessary to iterate over a very large file. Similar techniques can be very useful for other types of files, such as XML, JSON or even raw text.

In conclusion, each of the solutions outlined above should be reviewed within the common context of any problems solving activity, including money, time and available human and hardware resource capacity. I hope this discussion provides some meaningful alternatives in the increasing landscape of widespread utilization of R within the financial services vertical industry. ■



Jeff Heaton is a senior data scientist for RGA. He can be contacted at jheaton@rgare.com.

Drivers for the Digitalization of Insurance

By Hartmut Schroth

The insurance industry is becoming increasingly focused on the digitalization of its business processes. There are many factors driving digitalization, but a reliable and meaningful data architecture is the basic prerequisite to a successful digitalization strategy.

Digitalization is not just the “buzzword of the day.” A number of trends and changes in the insurance business environment are forcing companies to prioritize investment in digitalization.

Digitalization is not just the “buzzword of the day.”

KNOWLEDGE ABOUT CUSTOMER BEHAVIOR IS BECOMING MORE IMPORTANT THAN EVER

Many insurance companies have already invested in IT solutions for customer intelligence and customer analytics in recent years. In the future, the identification of current customer needs will be essential for a successful sales strategy and creating long-term customer loyalty. How has the customer (possibly including his peer group) behaved in the past? What were the motives? How will the customer behave in the future?

Today, a 360-degree customer profile includes all available customer information: data from party, core, debt collection and claims management systems and business analytics data warehouses as well as available external information (social media, Google maps, blogs and more).

NEW DISTRIBUTION CHANNELS ARE REPLACING TRADITIONAL ONES

Customers expect their insurance companies to be accessible at all times (online, call centers, local personnel). Thus, online sales and traditional distribution channels (agency, broker,

sales department) needs to be seamless; but omnichannel management creates difficult challenges for insurance IT.

CHANGING CONDITIONS OF COMPETITION

New digital business models (InsurTechs) are pushing into the insurance market with innovative offers:

- Online broker and comparison portals are gaining market shares in new business.
- Peer-to-peer (P2P) insurers are trying to replace traditional insurance models. For example, Friendsurance is taking a social community approach.

COST AND CUSTOMER REQUIREMENTS ARE DRIVING PRODUCT INNOVATION

Falling interest rates and changing customer demands require the development of new and innovative insurance products:

- Customizable insurance policies, e.g., a life insurance policy with flexible investment options for the portion of the premium that goes into savings.
- Linking the internet of things (IoT) and insurance products will grow and is already prevalent today in auto, home, life and health insurance.

COST OPTIMIZATION IS REQUIRED TO REDUCE LOSS RATIOS AND ADMINISTRATIVE COSTS

Altered basic conditions (e.g., increasing damage rates, declining interest rates) are forcing cost reduction:

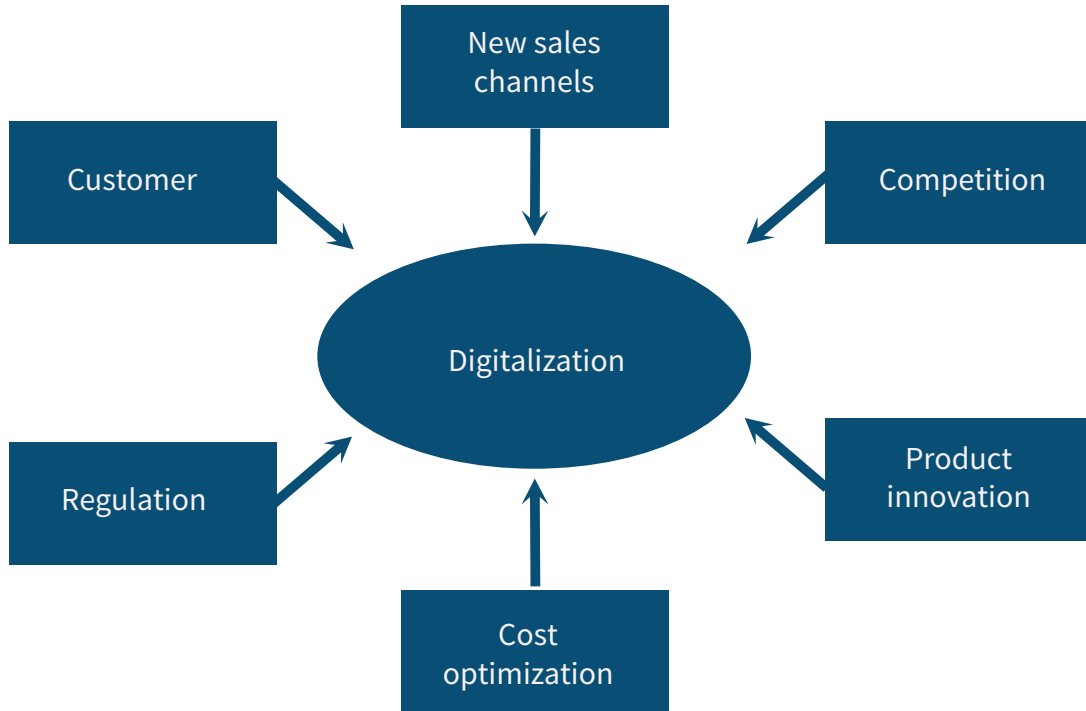
- Claims predictions using predictive modeling for property insurance, ideally in combination with IoT solutions for claim avoidance.
- Improved fraud detection for all lines of business using analytics tools and methods.
- Optimization of management functions by automating business processes (for example, in application examination).

COMPLIANCE REQUIREMENTS DEMAND A FUNDAMENTAL MODERNIZATION OF INSURANCE IT PROCESSES

Regulatory and legal requirements are defining new standards for the IT systems of an insurance company:

- Solvency II includes stricter requirements for transparency of IT processes.
- The introduction of the EU General Data Protection Regulation requires fundamental adjustments to data management processes related to personal data.

Figure 1



All of the digitalization drivers described above represent substantial data management challenges for IT departments throughout the insurance industry. So what's an insurer to do to stay competitive?

1. IMPLEMENT 360-DEGREE CUSTOMER AND PARTY PROFILES

Most insurance companies have implemented “party systems,” which uniquely identify a business party, regardless of the role he or she plays in the insurance business process (for example, customer/policyholder, contributor, intermediary/producer, external service provider).

In practice, however, many companies still have duplicate and incorrect information on parties due to data quality problems. Therefore, data cleansing is a must.

A further challenge is using not just internally stored customer information, but also to evaluate externally available data, and storing the insights from it in suitable structures. This includes geo-information (for example from Google Maps), as well as information from social networks and blogs. Since this information is generally not well-structured, new data management concepts are required (for example, based on Hadoop), as well as access mechanisms that must be incorporated into ETL processes and analytical evaluations.

2. EMPLOY OMNICHANNEL MANAGEMENT

Until now, it was tolerated that internal departments, brokers and agencies had different databases available for their sales activities. Now insurers are realizing that these data silos are killing an omnichannel strategy.

New concepts such as a customer decision hub, which provides all channels with complete customer history in a consolidated, quality-assured version, will be the norm. The customer decision hub may be linked to a business-analytics data warehouse thus giving a 360-degree view of all party information.

3. LEARN FROM NEW COMPETITORS

New competitors are not only a threat to traditional insurers, but also an opportunity.

- P2P products will likely secure larger risks through cooperation partners (i.e., insurance companies).
- Online comparison portals prefer high-value products, which can also be offered by established insurers.

In both cases, it's a great advantage for insurance companies if they're able to exchange data with insurtechs' IT systems via standard interfaces. Insurance companies that have already cooperated with insurance brokers will have an advantage over



companies with exclusive distribution channels. Furthermore, companies that have already implemented standard structures like ACORD in their data systems will be leaders.

4. KEEP INTRODUCING PRODUCT INNOVATIONS

Customizable insurance products, which in some cases also require a link to raw data and/or scoring results from IoT applications, will require extended and comprehensive access to policy and risk information. This relates to core and rate-making systems, as well as to already existing dispositive systems.

Without extensive modernization of the data systems in question, insurance companies will find it very difficult to introduce product innovations successfully.

5. OPTIMIZE COSTS THROUGH CLAIM PREDICTION AND IMPROVED FRAUD DETECTION

IT optimization in the context of claims leads to the topic of data:

- Claims prediction using analytical models is only as good as the underlying database. If the data of historical claims

cannot be evaluated correctly, the analytical models will not yield accurate scoring values. Data quality is an important prerequisite.

- Fraud detection by analytical methods requires the unmistakable identification of all involved persons as well as all claims objects over a claim's history. This requires a solid data management strategy.

6. MODERNIZE IT TO COMPLY WITH NEW REGULATIONS

The increasing regulatory and legal requirements for insurance compels companies to fundamentally renew their data processes. As a rule, the requirements cannot be served by a simple expansion or adaptation of software programs and IT infrastructures, especially ones that are piecemeal and outdated. A fundamental modernization of data management concepts is required:

- Solvency II requires that the underlying rules and all source information (data fields and source systems) be extensively documented. The resulting efforts to change and extend proprietary developments is immense and requirements are generally not met by existing IT solutions. Therefore, new data management solutions are necessary—optimally automated and supported by a metadata solution and a business glossary.
- With EU General Data Protection Regulation, insurance companies must ensure that personal data in the information systems is no longer displayed to all users. The necessary anonymization or pseudonymization requires either complex adaptations in existing ETL programs or the use of data management solutions designed for this purpose. In many cases, it is a challenge to recognize where all personal data is stored in the branched data landscape of an insurance company. Again, an intelligent data management solution may help.

SUMMARY

It becomes clear that the requirements for data management processes, including the underlying data structures, are in increasing demand to implement a digitalization strategy. This is why it is becoming crucial for insurance companies to implement new, powerful and flexible data management concepts. ■



Hartmut Schroth is a business advisor for data strategies in insurance at SAS. He is based in Frankfurt, Germany. For further discussions, connect with him on LinkedIn.



2017 SOA
**Valuation Actuary
Symposium**

**Aug. 28-29, 2017
San Antonio, TX**

The premier event for the financial reporting actuary.

Don't miss the 33rd offering of the Valuation Actuary Symposium. This is where you need to be to prepare for valuation and financial reporting challenges and opportunities in your field.

More information coming soon.

SOA.org/ValAct

Coding Standards and the Efficient Model

By Brody Lipperman

For most actuarial programmers, a lot of thought and effort is put into making their models run faster. Gains in speed are easy to measure, look good on reports, and either save money by requiring cheaper hardware to run, or gain value by allowing the actuaries to run more models and get more data. Run time, however, is not the only way to make your models more efficient.

In the normal life cycle of an actuarial model, there are thousands of hours put into developing, enhancing, testing, explaining, documenting and validating. The total cost of the human capital used for these models vastly outstrips the cost of the hardware required to complete a model run in an acceptable amount of time. Usually, very little attention is paid to any gains in these areas because they are very hard to quantify. If I spend 10 hours cleaning up code and documenting my model, does that save 10 hours of time down the road as various other people have an easier time of understanding the model? While the benefits of this type of efficient model are often hard to quantify, they are without a doubt just as valuable, if not more so, than the benefits of run time improvements.

One of the best ways to improve a model's overall efficiency is to develop a set of detailed coding standards. ...

One of the best ways to improve a model's overall efficiency is to develop a set of detailed coding standards for your modeling team. These standards should address stylistic considerations, function use, documentation rules, and any other aspects of model coding. If done properly, this should allow your model to be easier to read and understand, decrease the amount of time required to make changes, reduce key-man risk, and reduce coding errors. It will require a change in the mind-set of the team, and buy in for a shift in personal responsibilities.

A good set of coding standards should address most issues that will arise when writing code. While bracket placement and indentation length might not seem overly important at first, uniformity in the presentation of code allows users to focus on the content of the code, rather than being distracted by stylistic differences. The standards should also address naming conventions. Today's languages have typically removed size constraints, so developers should strive to avoid abbreviations when possible, but there should be rules on what abbreviations to use if necessary (so that users will understand that ANN is short for annual, not annuity). In setting up the standards, it is important to set rules for most coding situations, even if they are arbitrary (for example, a function should be no longer than 30 lines).

While exceptions to the standards can be approved by model owners, they should be well thought out and documented. If a code change will decrease run time by 10 percent, but increase the complexity of that section of code, attention should be paid to both the benefit and the downside. If the model is run overnight, would the 10 percent speed increase have any noticeable benefit? Or if weighed against the fact that the more complex code would take more time to explain to end users, could only be modified by a specific set of coders, and increases the risk of future errors, is it worth it? This type of decision-making should be brought up for each exception to the coding standards.

Each model (or subset of the model, for more complex models) should have a model owner. While the model owner will have the final say over content, they should not be the only person that validates that the coding standards are maintained through the model. Each user should have shared ownership and responsibility for the entire model. Frequent code review meetings are a fantastic way to foster this responsibility. Members of the team present code that they have been working on, and get feedback from other members of the team. This allows all members of the team to scrutinize the code and ensure that it conforms with the coding standards. It also allows less experienced developers to learn from their more tenured counterparts. Knowledge is shared across the team, both in coding methods and content of the model. This will reduce key-man risk, as each developer should understand any new pieces of code in the model, as they will be reviewed in these meetings.

It is important to encourage all developers to provide feedback, as there can be a tendency to have "experts" in various areas of the model. These experts will receive less scrutiny with their code reviews, and as a result, will typically have more mistakes in their final product. Other developers will also tend to defer to the expert when questions about their areas of code arise. This can reintroduce key-man risk, and potentially bottleneck future change requests.



Coding standards should also cover reusability aspects of the code. Any formula that is repeated in the model, should be converted to a function. By centralizing the code, developers can reduce the amount of time it takes to make any future changes. They can also reduce future errors that would be caused by changing a calculation in one section of the model without modifying the same calculation in another section. Having well-defined function names can also make reading the code easier for end users. If a user wanted to understand what all is included in an AV calculation, they can look at the code and see the calculation includes COI charges. They don't necessarily care that the COI function calculates a NAR after premium and loan interest is taken out. Functions allow users to absorb as much detail about the code as they want, while still being able to dive into each function if they need more information.

The final aspect of the coding standards should be rules around types and goals of documentation. Since all modern languages resemble English, the model code should be self-documenting. Variable names should be descriptive enough to be easily understood without any reference, and most users should be able to follow the basic logic constructs (if then, for loops, etc.). Code should be written in such a way as to reduce the complexity of each section of code as well. If the developer needs to document end points for If Then statements or for

Loops, then they should attempt to break the code into smaller, more digestible blocks of code. The goal of documentation, then, should be to explain why the code does what it is doing, instead of what the code is doing. This will allow future developers and end users to understand the choices the developers have made, and allow them to follow the whole model easier. Any documentation included in the code should be short and concise, anything longer than two sentences should be moved into a more formal document.

A strong set of coding standards and the proper team mindset can greatly reduce the amount of manpower required to maintain a model. By setting up rules and guidelines, developers are forced to consider methods that are easier to understand for future developers or end users and can create a better overall product. These standards can also help spread knowledge, responsibility, and ownership throughout the team, leading to a stronger, more flexible organization. The end result should be a model that is easy to maintain, easy to understand, easy to validate, and easy to manage. ■

Brody Lipperman, FSA, CERA, MAAA. He is a Lead Actuarial Developer with FIS. He can be reached at brody.lipperman@fisglobal.com

Wearable Wellness: Five Quick Takeaways from RGA's Fitness Tracker Study

By Chris Falkous

Using wearable fitness tech to develop insurance wellness programs has been the subject of an industry-wide discussion for years. And as the technology improves, the conversation grows. At RGA, we wanted to gather meaningful data and gain experience over theory in our understanding of insurance wellness initiatives.

In 2016, RGA conducted an anonymous study among its employees and their friends and family to explore wearable fitness trackers' potential application for insurance product development. The study included around 1,000 participants from 23 countries and was conducted over 12 weeks using five tracking devices.

Key insights for insurers included these five takeaways:

ACCURACY REMAINS AN ISSUE

Not all devices are created equal. For improved plan performance, insurers should consider mandating high-quality trackers, retrieving data from multiple devices, or limiting member benefits based on the quality of the device, particularly for those plans in which higher activity levels trigger additional benefits.

Key findings:

- Between the study's two main devices—both wrist-based—steps recorded on one were, on average, around 8 percent higher than those recorded by the other.
- Anecdotally, one participant recorded approximately 19,000 steps on his smartphone while simultaneously recording approximately 13,000 steps on a wrist-based device.
- Manufactured steps—via swinging an arm, for example—were identifiable as they occurred at activity levels not normally seen and at times of day when activity did not normally occur. This suggests that fraudulent individuals



will need to be quite sophisticated if they want to hide their “cheating.”

- Notably, a rather surprising result was that fraudulent steps, such as swinging an arm, sometimes raised heart rate to a level above that which would have been expected, so was this cheating at all? Significant further work is required to generate an accurate and robust solution for identifying fraudulent activity, but the initial signs are promising.

Questions the industry should be asking:

- How can insurers easily evaluate the accuracy and reliability of wearable devices?
- How should insurers translate these differences to ensure fairness among customers?
- How do insurers build a multi-device, multi-location solution?
- How can insurers identify and limit fraudulent activity?

STYLE AND COMFORT COUNT

Although it is important to evaluate and approve devices based on accuracy and reliability, insurers should also consider allowing participants as much choice as possible in what they wear. The more choices among approved devices, the better the user engagement and persistency.

Key findings:

- Attractiveness, unobtrusiveness, and ease of setup were listed as the most desirable features in a device.
- Reliability mattered as well; one device produced an unacceptable failure rate of 5-10 percent.
- Many participants also found it uncomfortable to sleep wearing their devices.

Key questions to consider moving forward:

- What is the human real estate wearables will eventually claim?
- With the proliferation and wide variety of wearable tech devices, how can insurers both allow for consumer choice and confirm device accuracy?
- Some of the latest devices do address style and comfort concerns. When will these devices be more affordable and widely available?

PRIVACY IS A PRIORITY

The increasing popularity of wearables is undeniable, yet many choose to remain non-users, mainly due to lack of interest or privacy concerns. Insurers need to factor this in as they develop wellness plans.

Key findings:

- Our survey of non-participants, which received hundreds of responses, provided insight into reasons why employees chose not to participate in the study. The figure below shows the results of this survey:

Reason for Not Participating	Results
Not interested in wearables	28%
Privacy concerns	25%
Not enough incentives	13%
Missed the deadline	13%
Misunderstood the criteria	9%
Own device not supported	6%
Other	6%

These results elicited two clear questions:

- How can insurers generate interest among those apathetic toward wearable tech?
- What can be done to allay privacy concerns?

DEMOGRAPHICS BRING DIFFERENCES

When designing wellness plans, it is important for insurers to identify demographic differentiators and consider support and incentives to appeal to each.

Key findings:

- When analyzing participation by region, we found very little variation in participation as the study progressed.
- On the contrary, participation levels by age did show significant differences. The over-50 age group had the highest participation level and the under-30 group the lowest.

Demographic questions to consider:

- To which target markets should wellness plans be tailored?
- What works in one region for one demographic is unlikely to be directly transferrable to other regions and other demographics. How can insurers best adjust programs accordingly?

DATA IS THE DRIVER

The wearables study was RGA's first step at gaining experience over theory and provides a starting platform from which we can advise our clients and launch deeper, more interesting work. We have initiated a program of gathering more hard data to support meaningful conclusions in the wellness space and are now working on additional pilot projects and seeking partners with more sophisticated metrics and modelling frameworks.

Much remains to be investigated before wearable wellness initiatives can fully progress from promising explorations to practical, long-term solutions. Continued real-world studies to collect and analyze real data must drive that process.

To read a more robust version of this article, please view the report on RGA's Knowledge Center. To learn more about RGA's wearables study and to view additional results and analysis, contact RGA. ■



Chris Falkous is a biometric research actuary in RGA's Global Research and Data Analytics team. He can be contacted at cfalkous@rgare.com.



SOCIETY OF ACTUARIES

475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
p: 847.706.3500 f: 847.706.3599
w: www.soa.org

NONPROFIT
ORGANIZATION
U.S. POSTAGE
PAID
SAINT JOSEPH, MI
PERMIT NO. 263

