

ISSUE 41 | OCTOBER 2011

SOCIETY OF ACTUARIES

Technology
Section

CompAct

ELECTRONIC NEWSLETTER



CONTENTS

[Table of Contents](#)[Letter from the Chair](#)[Editors' Notes](#)[Actuarial Software Wins](#)[Microsoft Innovation](#)[Award](#)[Touch the SCAI](#)[One Actuary's Journey](#)[Through Technology](#)[Keys to Successful](#)[Reports](#)[Excel Formula Rogue's](#)[Gallery 2: Electric](#)[Bugaloo](#)[Speaking "Data" Properly](#)[Part 2](#)[Articles Needed](#)

QUICK LINKS

[Technology Section](#)[Web site](#)[Council](#)

LinkedIn

Forward
to Colleague

Print

Search
Back Issues

KEYS TO SUCCESSFUL REPORTS

by Jeremy Benson

In my 15 years as an actuary, I have had many opportunities to produce reports for many different people. The actuarial reports could be anything from simple loss ratio reports to a full pricing analysis. I have also produced reports for marketing, underwriting, claims, external clients and others.

Today, I am responsible for approximately 20 different reports. Some of these were built by my team and me, while some were built by others and transferred to us. Because of this, we are going through an extensive process improvement project.

Based on these experiences, I have found that customers, quality and efficiency are the three key components to successful reports.

And that implies that:

- Reports that are not used by your customer are useless
- Reports that have bad information are useless
- Reports that are late are useless

In addition to the above, technology enables you to create the reports, and product knowledge enables you to understand the information in the reports.

Customers

First, let's define what a customer is. A customer is anyone who uses

[Links of Interest](#)

[Fiction Contest](#)

[Richard Junker, Co-editor](#)

[Andrew Chan, Co-editor](#)

SOA Staff

[Meg Weber, Staff Partner](#)

[Sue Martz](#)

[Section Specialist](#)

[Sam Phillips, Staff Editor](#)

your product. In this case your product is the report. Part of a successful reporting process is to find out who your customers are for each report. Once you have done that, the next step is to determine how your customers are using the report. A tool used in Six Sigma is called Voice of Customer (VoC). VoC is used to understand what is critical to the quality of the process according to your customer. It can be used to gather the needs and issues of the customers and requires a dialogue between both parties.

For those who are using the reports, discuss with them ways that you can improve the product. Is there additional data they would like to be included? Would they like it in a different format? Would they like it more or less frequently? Would they like an alternative way to access the report? Would they like canned reports or dynamic?

If there are people that you thought were your customers, but they aren't using the reports—find out why. Is the report too complicated? Does it have the wrong information? Is it too complex to use (i.e., is it in the format of a pivot table and the customer doesn't know how to use pivot tables)? Does the report not reconcile with other information?

If no one is using the reports, you should also consider whether the report is worth producing. It does not make sense to create a report that no one is using.

After talking to your customers, you will start turning their needs into requirements. These requirements will then need to be prioritized with all other requirements for all reports. The needs of the business along with the complexity and time required to complete the requirements should determine the priority.

Quality

Ideally, every report that you produce will have information that is 100 percent accurate. In the real world, this is not always possible. However, you should strive to produce reports with as high a quality of information as possible.

The first issue is the quality of the data that is used for the report. This is the classic "garbage-in garbage-out" situation. If your data that you use for the report is bad, then the information in the report will also be bad.

Before you start testing for quality of the data, you need to understand what is in the database. It is ideal to have metadata, or data about the data. Metadata can be pretty extensive, so it might be

too time consuming to capture every piece of information about each field. At a minimum you should have an operational definition, possible values and any business rules associated with each field in the report.

There are two main approaches that I have used on testing data quality on input data. First, a thorough data profiling exercise on the data is important. Data profiling is an exercise in exploring and learning about data. The main types of information that you should capture during this stage are:

1. Fill Rates - This is the percentage of the field that does not contain missing or null records for a particular field. It will help you identify whether data even exists or not. Care should be taken, however. There could be some fields that may have missing values by design. It is important to compare these results to the expected domain (desired or valid values) in the metadata.
2. Frequency Distributions - This is a count of how often each value occurs in the dataset for a given attribute (field). When comparing to expectations, you should be able to identify instances where information is not valid.
3. Two-way profiling - This is basically extending the two methods above to a second dimension. For example, you may want to see how an attribute's fill rate has changed over time, so you calculate it by year. Two-way frequency distributions are useful in looking at relationships between two attributes. If you know that two values shouldn't be related, but they occur for the same record, this signals a quality issue.
4. Basic statistics - Statistics like mean, minimum, maximum, standard deviation should be compared to the expected values. They could identify outliers or data that is unexpected.

A second way to identify data quality errors is more direct. It is through the testing of business rules. Since they are tests of the data already in the database, some call them data validation rules, but they are usually synonymous with business rules used in application development.

These rules are tested against the data and any violations are flagged and investigated. If the data is determined to be in error, then there are two ways to fix the error.

1. First, the actual data should be cleaned. Either it can be

cleaned at the source or in the database used for reporting (e.g., data warehouse).

2. Second, the process that caused the error should be fixed so the error doesn't occur again.

It is very important to analyze the data to understand the root cause of any data quality issues. Only through an understanding of the root cause can the underlying issue be fixed. Otherwise, even if you clean the data, more errors will occur in the future.

There are several ways to examine the root cause. A qualitative way used in Six Sigma is a Cause and Effect Diagram which is sometimes called a Fishbone because of the way it looks. It organizes possible causes for a specific problem by displaying them in increasing detail. A quantitative way to examine root causes with data is data lineage. Data lineage is the history of a data item, including its acquisition, transformation and integration. Knowing where the data came from can help you find the root cause of the error.

Even if the underlying data is correct, the report could have information that is not. Therefore, the person producing the reports should perform additional quality tests on the reports themselves before they are released to the customer.

There are two main tests that we do. They are called delta reports and control totals. The delta reports show what has changed from one report to the next. If what has changed is different from expected, then the reason should be investigated. Delta reports are part art, part science. It requires someone knowledgeable enough about the business to know what is expected; therefore, a subject matter expert can be referred to in such cases.

Control totals are basically a reconciliation of the report results with a source that either should be equal to the report totals (or subtotals), or the differences can be explained. If the report cannot be reconciled, then further investigation is required. The results of the investigation might show an issue with the process that creates the report, or it may be a data quality issue.

One important lesson on data quality when doing reporting is to get to know the people who use the data and the people who own the data. The people who use the data will be able to help you answer questions about unusual things you see in the data. They also will help you identify whether rule violations are errors or not. The people who own the data will many times be the ones you will have to work

with to fix the process that is causing the data quality error.

One other way to find out about data and information quality errors is to listen to your customer. This can be conducted formally through surveys, focus groups or interviews or done informally. In talking with your customers, many times you will discover information that is very difficult to glean from the data itself.

Successful reporting provides information that is used by customers, is of a high quality, and arrives in a timely manner. Potential misuse is still possible (if not probable) when users misunderstand or misinterpret the information. Three ways to minimize this are to:

1. provide clear definitions for the fields in the reports
2. identify key takeaways from each report for the user
3. train report users on how to interpret the information in the report

Efficiency

There are many ways that the process used to produce the report can be inefficient. The main ones that I have witnessed are lack of automation, useless information, redundant information, wasteful process steps and a lack of documentation.

Lack of Automation

When a report is first created, manual steps are important in order to understand the process. However, once the report is in production, manual steps can add significant time to the process. In addition, manual steps increase the probability of human error.

One of the dangers of automation is that a change in the underlying data might go unrecognized. That is why you must still be diligent in testing the quality of the data and reports as discussed above.

It is also important to make sure the process continues to run smoothly. Capturing the time it takes to run each step and building Statistical Control Charts are one way to do this. Whenever a step takes significantly more time, it will be outside of the upper control level (UCL). When this happens, the cause of the variation should be investigated.

Useless Information

Extra information not used by the customer is useless. Having this information slows down processing time. It also slows down the time to test the quality of the data. If your customers are not using that information, then you should consider removing it from the report.

One thing to keep in mind is that you might be a customer of your own report, so if you need the information, keep it.

Removing information from your report does not mean removing the data from your database. You should probably keep more data in your database than in the reports, in case you need to fulfill requests for reports using that data.

Redundant Information

When you have different customers asking for similar information, but in different formats, there is a tendency to create different reports and thus have a different process for each of them. When this happens, you could meet with each of your customers and determine if there is a way to combine the reports.

The decision to do this also depends on how long it takes to create a single report. For example, if the report can be created easily from a data mart with a dynamic interface, then the time saved may not be significant. So make sure that the benefits of the time saved by combining the reports outweigh the benefits that your customers may have had with separate reports.

Wasteful Process Steps

Some steps don't add any value to the process. Either in the development stage or when the reports are in production, it is important to determine if each step leads to the ultimate goal. If it doesn't, then the process should be redesigned.

Lean Six Sigma has a process improvement method called MUDA, which is a Japanese term for eliminating waste. It identifies activities that add no value to the overall process, but add time and/or cost. These concepts come from manufacturing, but can be used in data management and reporting also. There are seven categories of waste in lean:

1. Overproduction - this usually happens when we give the customer more information than they actually need (such as too many fields).
2. Unnecessary transportation - this happens when data moves from one system to another unnecessarily. The process should minimize the number of movements of data from source to final report.
3. Waiting/queuing - this is waste caused by inactivity in the process. It is usually caused by manual processes in which you rely on someone else for information.

4. Extra processing - this is additional work needed after the product is completed that adds no value to the customer. One example in reporting is continually cleaning data instead of fixing the process.
5. Motion - this is movement of people that doesn't add value. One example is having different people generating similar reports.
6. Inventory - this is having more reports on hand than are actually needed.
7. Defects - this is reports that are produced with errors.

Lack of Documentation

When a reporting process is put into production, the process should be documented. Lack of documentation becomes an issue when:

- Frequency of reports is low - A person tends to forget the steps if there is significant time between creation of the reports.
- Turnover - When a process has to be transferred from one person to another, lack of documentation makes it difficult to train the new person on the process.
- Manual Steps - If there are any steps that are manual, a process with documentation takes less time and is more efficient.

Each process should have standard operating procedures. In addition, every attribute (column) in the report should be defined so that the user of the report does not misinterpret it.

Technology

Technology is an enabler to the process. However, the choice of which technology we use does not change the customer requirement or the quality of the input data. Certain technologies can help you be more efficient, but process efficiencies go far beyond the type of technology that is being used.

Too many times we have seen companies implement new systems and talk about all of the great things they can do. But in my experience, too much of the focus is on the front-end of the system, and as a result the back-end data and the reporting suffer. When building new systems, customer reporting requirements, data quality and how efficiently you can create reports should be part of the main considerations.

Product Knowledge

The reporting team needs to have knowledge of the products that reports are based on. However, it should not be assumed that they know everything. There will always be people inside of the business that know a lot more (called subject matter experts, SMEs). You should work with these people to help tackle issues that are outside your area of expertise.

Conclusion

In order to have a successful reporting process, it is essential to understand customer needs, produce high quality reports and produce them in an efficient manner. Technology and product knowledge will also enable you to have successful reports.

Jeremy Benson, FSA, FCAS, MAAA is senior pricing actuary for the Medical Expense Group of Swiss Re. He can be reached at Jeremy_Benson@swissre.com.



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