

# **SAS® Visual Investigator and SAS® Visual Analytics: Bridging the Gap between Subject Matter Experts**

**Emily Chapman-McQuiston, SAS Institute Inc., Cary, NC**

## **ABSTRACT**

Health care payers constantly work in the world of big data. Because of the insights that big data can provide, payers are increasingly required to use it to improve outcomes for their patients, lower costs, and prevent fraud, waste, and abuse (FWA) in their networks. This enormous task requires contributors from diverse areas of the organization such as strategically minded executives, clinicians, auditors, and data scientists. Together, SAS® Visual Investigator and SAS® Visual Analytics provide a toolset that enables users to bridge the gaps between these different roles. This paper examines how advanced analytics, coupled with visualizations, make complex analytics readily available to subject matter experts across the health care industry.

## **INTRODUCTION**

As the health care payer industry has grown and billing systems have become more complex, the roles of analysts and subject matter experts (SMEs) have grown increasingly vital. Traditionally, analysts and SMEs have interacted through static reports and presentations. The issue with this for payers is that enrollment and claims information change daily. This means that without automated systems to update the information and analytical results, the information can become stale relatively quickly, and the process must be repeated.

Many organizations use dashboards to provide quick, automated views of summary statistics and cost measures. These summaries are extremely important to a payer's daily operations. This paper details several use cases to examine how the traditional method of analytical reports can be combined with technologies such as SAS Visual Analytics and SAS Visual Investigator to provide not only summary information in a dashboard, but also predictive information about an automated, regularly updated cycle. This paper also discusses visualization techniques that can be used to bridge the gap between the clinical and policy language of a health care SME with predictive and accurate information that can be provided using advanced analytical techniques.

## **SAS VISUAL ANALYTICS: INTEGRATING ADVANCED ANALYTICS INTO SUMMARY DASHBOARDS**

### **USING WATERFALL CHARTS TO DISPLAY MODEL EFFECTS**

Health care analysts and epidemiologists have a long history of using regression analysis to understand the causes of improvement in health care quality and the long-term health of patients (Mihaylova, Briggs, O'Hagan, & Thompson, 2010). For health care payers, the need to understand these concepts on large data sets can be tremendously important.

This use case examines acute inpatient stays as billed to commercial insurance. Most payers use a payment system for inpatient stays that is based on the Centers for Medicare & Medicaid Services' (CMS) inpatient prospective payment system (CMS, 2017). The payment for a stay is based on a diagnosis related group (DRG) code of which the case was assigned using an algorithm known as a grouper, which takes information such as diagnoses codes, procedures, and demographics to assign a code. This code is the basis of the payment for that stay. Although CMS uses Medicare Severity-DRGs (MS-DRGs), which are designed for the Medicare population, Medicaid and commercial payers tend to use different groupers to consider a wider range of causes of inpatient stays, especially those incurred by the pediatric population.

In the following use case, SAS built a model to understand the causes of longer or shorter lengths of stay (LOS) for different patients for the same DRG code. SAS used the medical system utilization history

based on claims and demographic information of a patient as a proxy for that patient's health status. The target in this case was the LOS, and the input variables were the utilization history and demographics.

Within the SAS platform, SAS built a model using PROC GLIMMIX to examine how a patient's claims history could be used to understand the causes of a longer or shorter LOS. Using a mixed model allows the modeler to account for the correlation between repeated stays of the same patient. Although this has been done for single DRG codes, an automated system was created so that this could be repeated on a regular basis for more than 100 DRG codes.

The output of this system comprises hundreds of models and thousands of effects. This is where visualization of the data, confidence intervals, and effects in an interface allow SMEs to interact with the models and gain clinical understanding.

Figure 1 presents a dashboard that allows users to filter by DRG code. The table in the top-left table shows demographic and enrollment information about specific patients. The top right table shows information about the facility that submitted the inpatient stay. The line chart in the middle shows the upper and lower confidence intervals as predicted by the model (gray), the predicted LOS (blue), and the actual LOS (green). Users can quickly identify stays where the actual LOS was outside the predicted confidence interval. Clicking on a stay in the line chart filters the top tables and the bottom chart down to that stay.

The bottom chart, shows waterfall charts in SAS Visual Analytics that enable clinicians and experts to not only understand the causes, but also see exactly what predicted a longer or shorter LOS for specific patients. The red portion of the waterfall chart indicates the effects that are expected to decrease the LOS, and the green boxes are effects that increase the LOS. In Figure 1, utilization of office visits and going to a facility in a well-populated county are the two effects that decreased the length of stay of the patient, whereas patient age greater than 65 and a high utilization of facility claims (inpatient and outpatient services) increased the LOS. Although these results are not exceedingly surprising, the third effect that decreased LOS was a previous diagnosis of hypertension, which could lead to some insight for this individual patient.

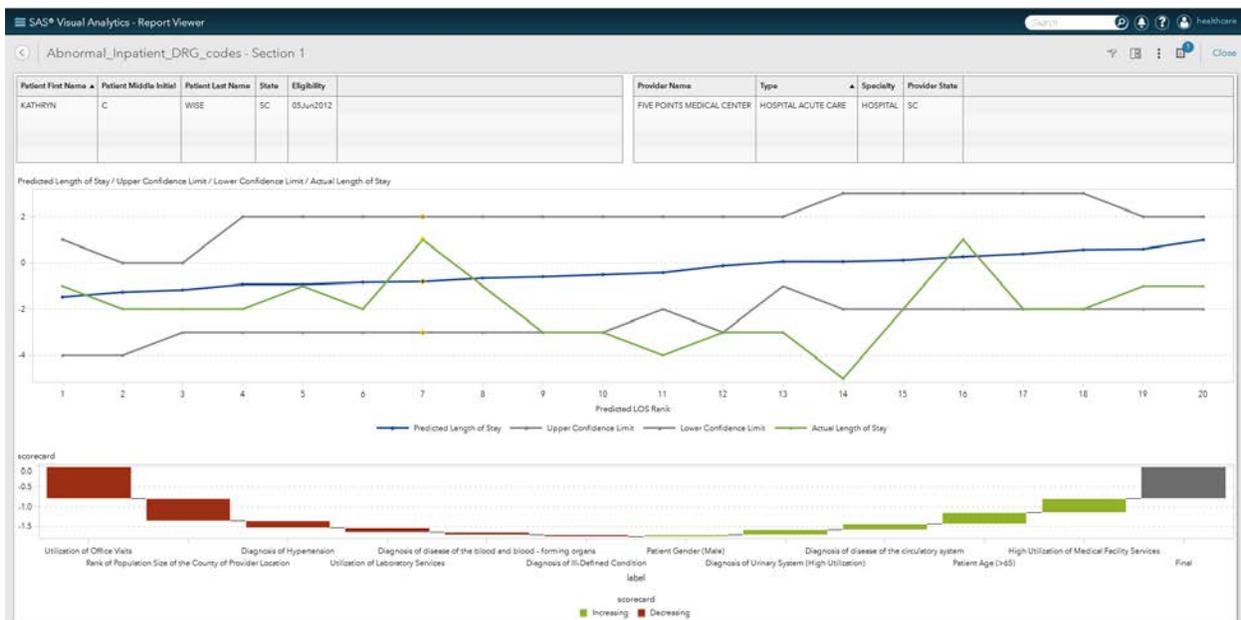


Figure 1. SAS Visual Analytics Dashboard – Results from a Mixed Model to Predict a Patient's LOS

This analysis could aid in prioritizing audit candidates when a facility provider is suspected of upcoding DRG codes to maximize pay. Another example of using this type of technique is to change the DRG code to admit diagnosis. This type of analysis could be re-examined to allow hospitals to predict LOS upon patient admissions and give explanations about the components of the patient's history that might contribute to complications.

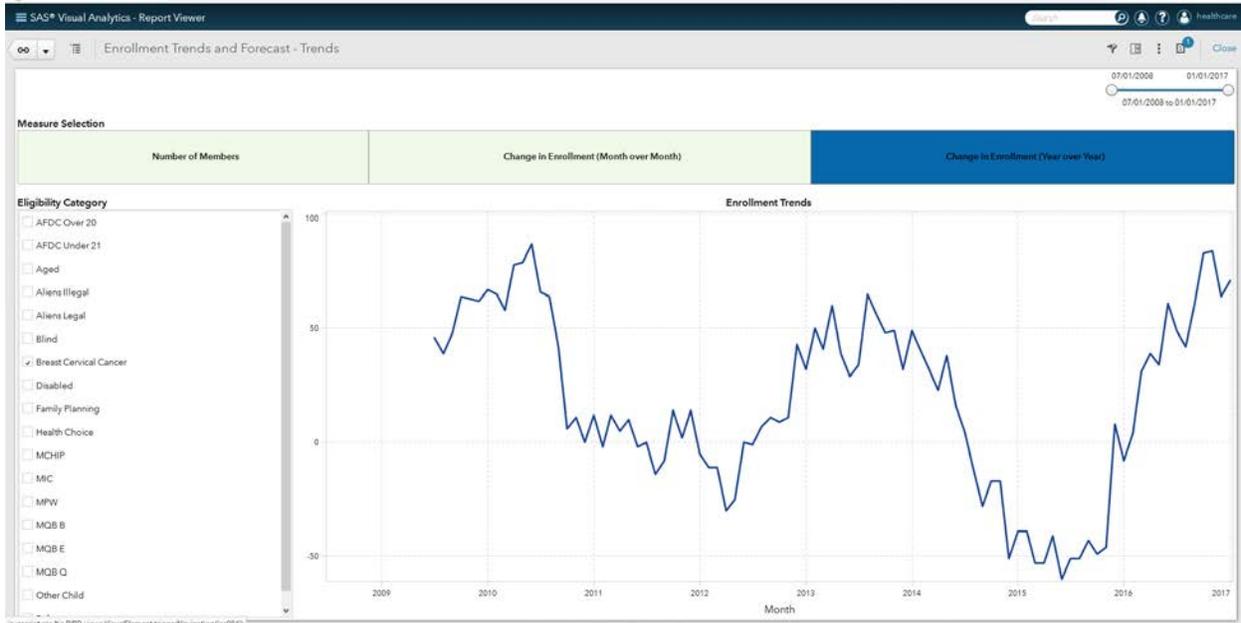
## **INTEGRATING INTERACTIVE TRENDS AND FORECASTS WITH DASHBOARDS**

Health care payers spend significant resources on actuarial departments to forecast health care expenditures (CMS, 2018). Each year, CMS presents extensive reports that project changes in health care spending. In this section, a simple forecasting use case examines the enablement possibilities of technology in this process. One of the most basic needs is understanding projected plan membership, or, in the case of state Medicaid agencies, projected enrollment in different eligibility categories. This greatly affects the budgeting and planning for health care costs in the upcoming year. Traditional dashboards give historical trends along with current summary numbers, while this use case shows how to incorporate forward-looking information.

This example is based on Medicaid eligibility and the data components and issues that Medicaid state agencies face. The following steps are necessary for creating a dashboard that is automated to both understand eligibility trends and provide interactive forecasts:

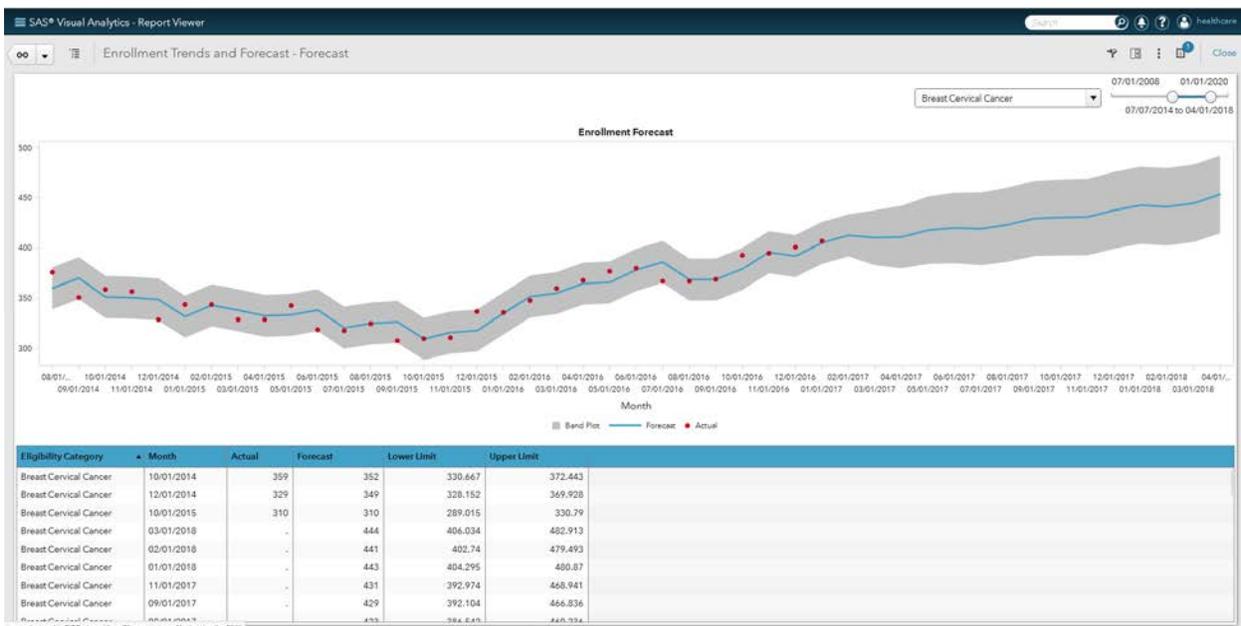
1. Data preparation and automated data rules for data reconciliation that account for the following:
  - a. Demographic information
  - b. Reconciliation of Medicaid eligibility categories
  - c. Month-to-month changes in the individuals' enrollment
2. Identification of events that can greatly affect eligibility forecasts—in this case, changes associated with the Affordable Care Act policy that need to be monitored and uploaded.
3. Forecast creation and automatic refresh and retuning at the individual eligibility category level using SAS® Forecast Studio.
4. Data loading to a pre-developed SAS Visual Analytics dashboard.

The steps above allow users to present the enrollment information in an attractive way in which policy SMEs can account for current enrollment trends for their state at the individual eligibility category. This process leverages the powerful forecasting power of SAS Forecast Studio along with the user-friendly visualizations / dashboards of SAS Visual Analytics in a fully automated manner. Figure 2 shows a trends dashboard that allows SMEs to explore their data. Users can examine the total number of members, month-over-month change, and year-over-year change. The check boxes and options enable users to drill down into individual eligibility categories.



**Figure 2. Year-over-Year Changes in Enrollment**

Although the trend dashboard is useful for understanding the past, the advantage of analytics is giving insight into the future. The second part of the process as described above is providing interactive exploration of both the historical trend measures and the automated forecasts. Figure 3 shows the automated forecast output integrated with SAS Visual Analytics. The forecasts were performed at the individual eligibility category, allowing for model selection to be performed for each eligibility category. In this dashboard, historical trend data is combined with the forecasts along with upper and lower limits to provide insight into expected eligibility trends.



**Figure 3. Historical Trend Combined with the Forecasted Trend for the Individual Eligibility Group (top); Numbers for Use in Other Calculations (bottom)**

A drop-down menu enables users to specify an individual eligibility category forecast and a time filter for highlighting certain parts of the trend or forecast. The combination of the simple forecast visual, the ability to drill down to the enrollment numbers for different eligibility categories, and the ability to interactively examine the historical trends along with the automated forecasts can aid in upcoming budget decisions. One of the strengths of this integration with SAS Forecast Studio and SAS Visual Analytics is that changes can be made to forecasts with minimal effort. For example, if a new eligibility category or new event variable (for example, a policy change) are introduced, you can change the SAS Forecast Studio project, and everything downstream automatically updates accordingly.

## **USING DECISION TREE VISUALIZATIONS FOR CLINICAL TRANSLATIONS OF PATIENT RISK**

A difficult problem facing claims-based editing systems is identifying claims that are abnormal based on a patient's history at the point of submission. In the following use case, SAS created a system to automate the process of detecting risky claims based on a patient's unique history (Enck, Chapman-McQuiston, & Kelly, 2015). First, every procedure, drug, and inpatient and outpatient service is aggregated into a table detailing patient history. A decision tree is then created for every procedure code, primary and secondary diagnoses code, drug code, revenue code, and DRG code entering the health care claims payment system. Each tree is further broken down by patient demographics and length of eligibility to ensure that each tree accounts for the population in the model.

The claims data is aggregated so that each claim contributes to that patient's medical claim history. Using the procedure PROC ARBOR that is provided with a SAS® Enterprise Miner™ license, you can automate the training and creation of tens of thousands of decision trees. Each model is subset by the code of study (for example, procedure, drug, diagnosis, and revenue), and the population is sampled into patients who had claims with that code and patients who did not as the final target. Further information about the sampling and training process can be found in the patent referenced above. Each claim or claim line is given a risk score that is based on events in a patient's history that occurred before the current claim being submitted.

An example is a patient who is prescribed an opioid. It would be expected that within the last 30 days, the patient would have had a visit to the emergency room, an MRI or X-ray, a surgery, or another procedure. In other words, the patient should have had some sort of a medical event or diagnosis that would have necessitated the prescription. Although this model is designed to score claims at the point of submission, the output assigns a risk score to every claim, and the output can be presented to understand if there are any providers who, at an aggregate level, continually submit more risky claims.

Figure 4 shows the aggregated risk of high-value claims on a map. The dashboard provides a view around which counties' patients are having high-risk claims submitted.

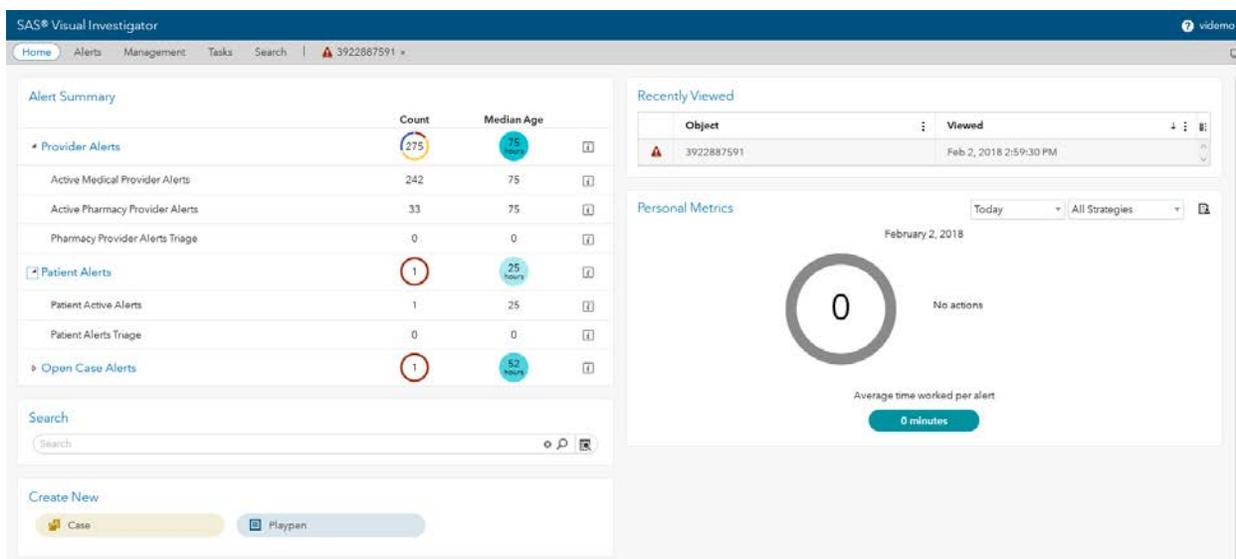


In SAS Visual Analytics, users can link out from a dashboard to a URL. The URL can then be used to kick off a stored process on the SAS stored process server. This visualization is created by using the SAS stored process facility, SAS code, and open-source graphics package D3 (<https://d3js.org/>). In the example in Figure 5, a clinical SME can follow the light-blue line through a patient’s history. It is shown that this patient received a physical therapy re-evaluation that was determined to be a risky claim. The patient had a diagnosis of spondylosis, and had therapeutic procedures within 30 days but did not have any between 30 and 180 days. Looking further into the claims data for this patient in the last 180 days, you can see that this physical therapy center is coding physical therapy re-evaluations on a biweekly basis. For this payer, policy limits the frequency of re-evaluations to five months; it is not expected to find a dramatic change in patient progress when they completed 30 days of therapy without a new presenting problem.

## SAS VISUAL INVESTIGATOR: AUDITS AND INVESTIGATIONS

### INTRODUCTION TO SAS VISUAL INVESTIGATOR

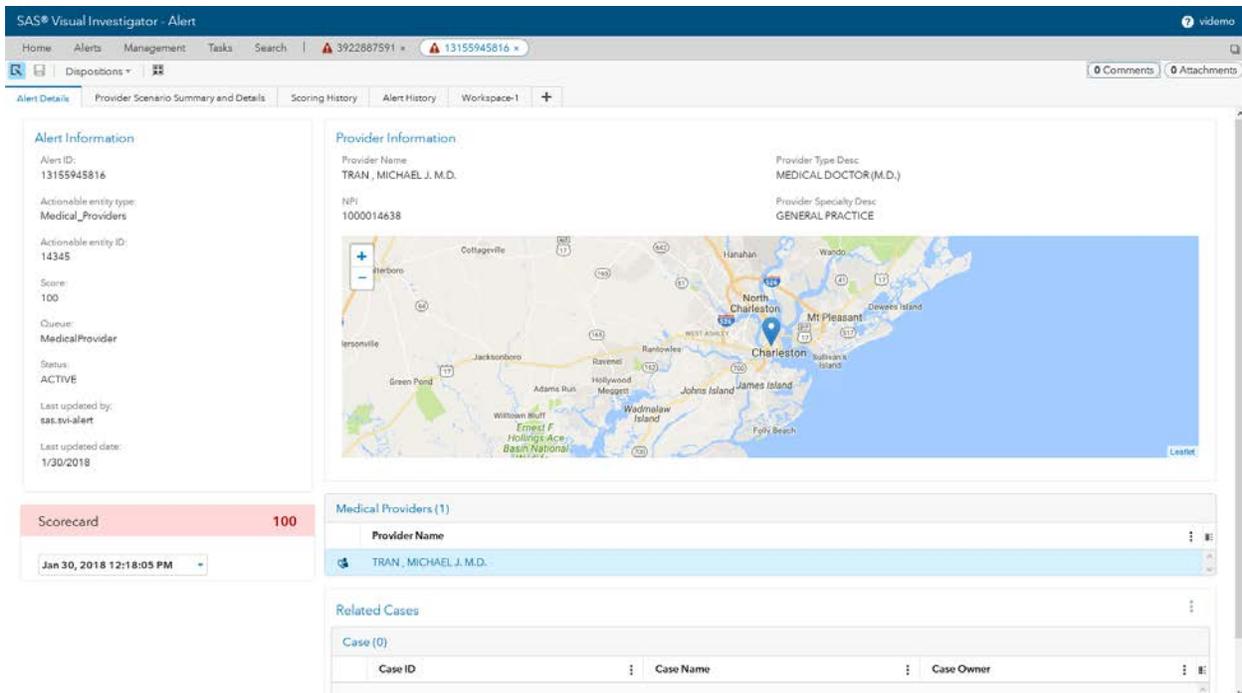
SAS Visual Investigator provides an attractive graphical user interface to enable the process of audits and investigations (SAS Institute Inc., 2017). Figure 6 shows the introductory interface of SAS Visual Investigator designed for health care payers. This summary screen allows users to drill down and investigate medical providers, pharmacies, or specific patients. Note that the investigation metrics provided on this screen in the system will automatically track alert age and the investigator’s personal metrics.



**Figure 6. Introductory Interface for SAS Visual Investigator Configured for Health Care Payers**

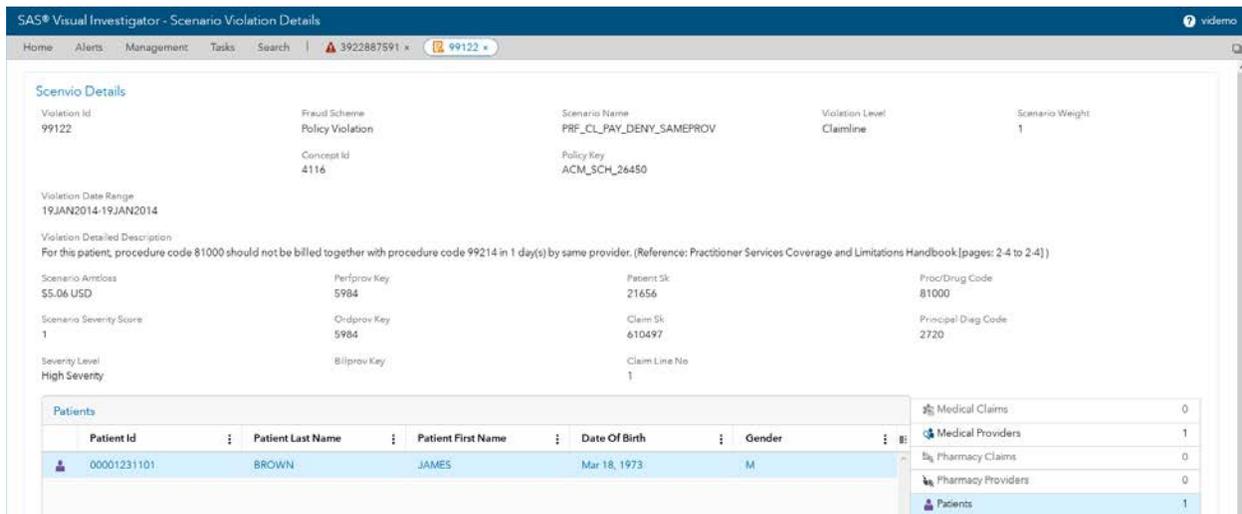
The alerts presented in Figure 6 are prioritized using a mix of anomaly detection, rules, and risk models. Each piece of logic that contributes to determining the risk score of a provider is known as a “scenario.” The difficulty is combining these different forms of analysis or scenarios in an automated fashion to identify where and how a provider is manipulating the billing system or even putting patients at risk of harm.

To aid this process, once the alerts are prioritized, users can drill down to a screen that distinctly summarizes information about the provider and can present practice location information on a map (see Figure 7).



**Figure 7. Summary Information about an Alerted Provider**

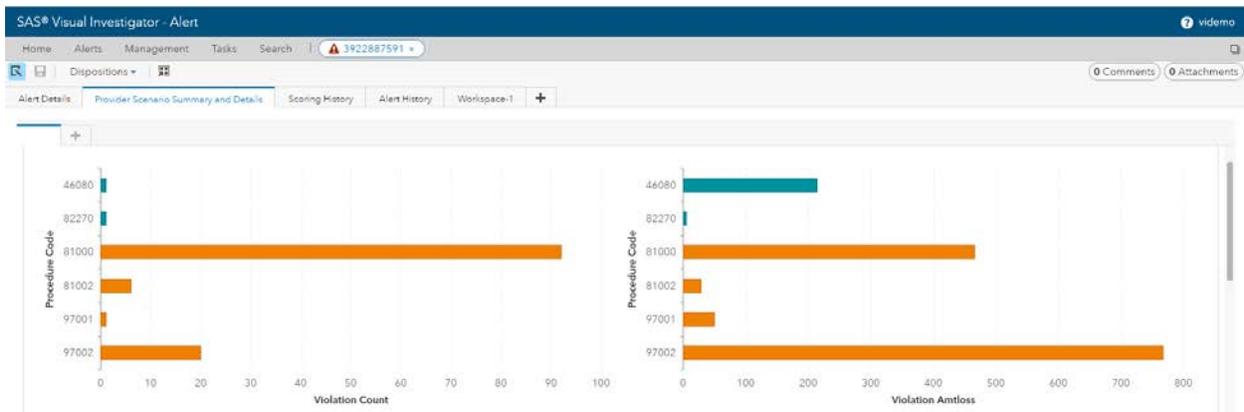
After the initial summary screen, an investigator can develop an understanding of the behavior that caused the automated alert. In the scenario complexity spectrum, rules are the easiest to understand and generally do not need any type of visualization. A rule can be detailed in language with references to the policy manual or bulletin (see Figure 8).



**Figure 8. A Scenario Violation That Occurred Based on a Rule Derived from Policy**

Moving farther down the complexity spectrum, it is less effective to use words to convey a concept. For example, when an investigator is trying to understand which billing area is an issue, it is one thing to use the words, "This provider is an anomaly at the number of billing violations for ordering procedure code 81000, 'Urinalysis, by dip stick or tablet reagent for bilirubin, glucose, hemoglobin, ketones, leukocytes, nitrate, pH, protein, specific gravity, urobilinogen, any number of these constituents; non-automated, with microscopy' when compared to similar providers." Figure 9 conveys this information quickly and shows

that the two procedure codes with the most violations are 81000 and 97002. This indicates that the provider's behavior in billing these areas should be examined.



**Figure 9. Provider Scenario and Procedure Codes**

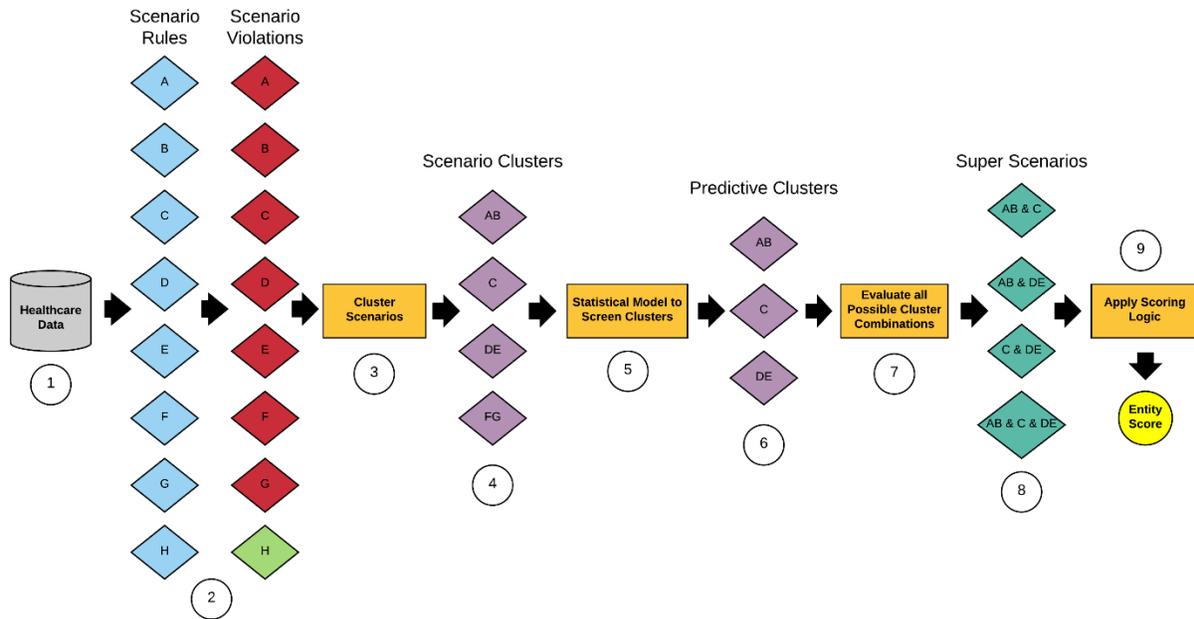
Rules provide quick monetary recoveries but can ignore significant abuses of the system by missing behaviors like services billed that were never rendered, which cannot be found by a policy violation but by more advanced techniques. It is also important to have scenarios at different levels such as claim, claim line scenarios, servicing, referring, and billing provider behavior scenarios. The next section discusses combining scenarios along this spectrum to detect abusive behavior along with a custom visualization to allow straightforward investigation.

## DEVELOPING CUSTOM CONTROLS TO PROVIDE INSIGHT INTO RISK

SAS Visual Investigator has investigational interfaces and views with an easy point-and-click interface. The product also provides the ability to build custom visualizations to give the flexibility to visualize complex analysis.

Identifying health care FWA among most providers who provide good care is extremely difficult. Even with a provider who is severely manipulating the billing system, 80% of their claims can still be complete and accurate. There are many other challenges in finding providers that are manipulating the system, including the size and complexity of health care coding as well as the fact that abusive providers are able to look like “good” providers (in other words, providers who provide good care) in the data.

To successfully identify providers for investigation or audit, complex sets of analyses need to be combined into one score. As previously discussed, SAS uses a mix of rules, anomaly detection, and models as scenarios, and combines these to result in one score. One method to do this involves the process depicted in Figure 10. Scenarios of all complexities are run, resulting in scenario violations that are claims, claim lines, or providers outside the set of logic. The hundreds of scenarios are clustered so that highly correlated scenarios are not double counted, and known outcomes are used to identify combinations of scenarios that lead to providers who are likely to result in recoveries during audits and investigations.



**Figure 10. Sets of Logic (Scenarios) Run Against Health Care Data**

The results of this process are a combination of behaviors. An example of a simple output of this algorithm in a table looks like the output shown in Table 1.

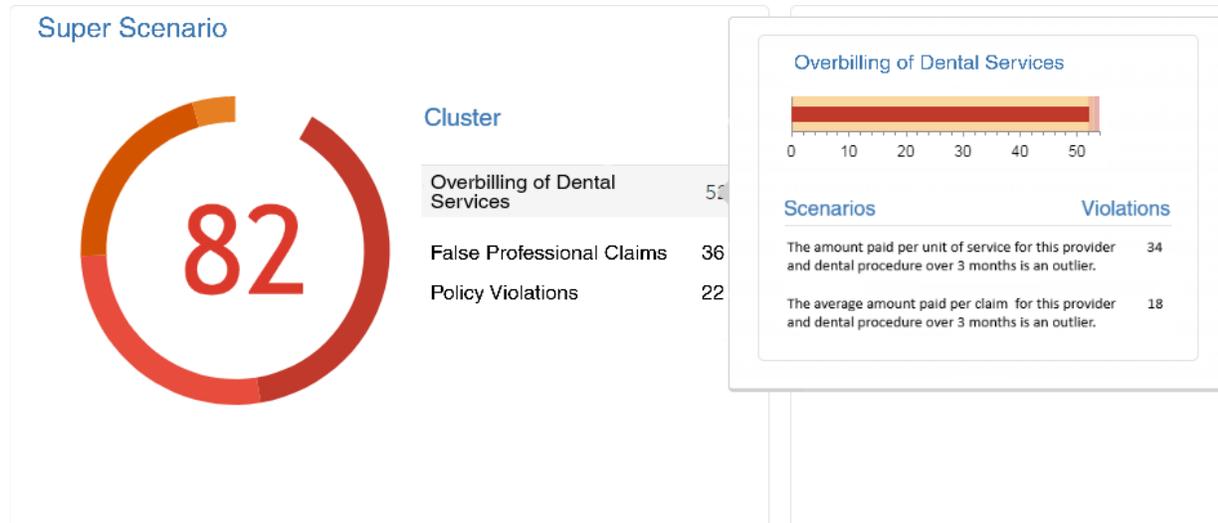
Complex Scenario	Behavior	Scenario Cluster	Scenario Description
A	Overbilling of Dental Services	259	The average amount paid per dental claim for this servicing provider and procedure code in the dental claims over the past three months versus the provider's peer is an outlier.
A	Overbilling of Dental Services	259	The average amount paid per unit of service for this servicing provider and procedure code in the dental claims over the past three months vs. the provider's peer is an outlier.
A	False Professional Claims	258	Two providers submitted duplicate professional claims for this member on the same day of service.
A	Policy Violations	109	This dental procedure code has been billed more frequent than is accepted by the policy manual.
A	Policy Violations	109	The dental procedure code policy violation has been billed for this provider many more times than the provider's peers.

**Table 1. Output of the Process Described in Figure 10**

To provide this information to investigators in a way that is consumable, you can use a feature in SAS Visual Investigator known as the custom control. This feature allows developers to create custom visualizations to display analytical results within the user interface (SAS Institute Inc., 2017).

Figure 11 is an example of a custom control that provides insight into the behaviors that are triggering these super scenarios. Users can also drill down into the scenarios and reasons that a provider has been triggered for audit or investigation. In Figure 11, the number 82 indicates that, of the known outcomes that were provided, 82% of those providers ended in a viable candidate for audit or investigation. The table to the right of the pie chart shows the list of behaviors that need to be triggered for a provider to trigger a

super scenario along with the total number of scenario violations the provider has accumulated in each behavior cluster. The example provider shown below is an oral surgeon who has both dental and professional claims. This provider has been shown to trigger scenarios that indicate a tendency to overbill dental services, scenarios that indicate possible false professional claims, and submits claims that are against this payer's policy but do not necessarily get triggered in the editing system.



**Figure 11. Custom Control Developed within SAS Visual Investigator**

The custom control also provides drill-down capability for an auditor to understand which scenarios have been triggered as part of this behavior cluster. The example shows a drill down on “Overbilling of Dental Services,” and you can see that this provider is an outlier when it comes to the amount paid per claim and unit of service. This allows an investigator to narrow in on dental claims that might be manipulated to get additional money per claim outside of standard billing practice.

## CONCLUSION

The use cases discussed in this paper detail how technology, analysts, and SMEs can combine visually appealing graphics and analytical techniques in an automated fashion to provide internal business users and executives with up-to-date, predictive information. The complexity of the world of health care payers exacerbates the necessity for SMEs and analysts to be able to work with updated information and analytical techniques. SAS Visual Analytics and SAS Visual Investigator provide unique, front-end interfaces that can be combined with the analytical strength of the SAS platform to create analytical visualizations on an enterprise platform.

## REFERENCES

Centers for Medicare & Medicaid Services. 2017. “Acute Inpatient PPS.” Accessed February 19, 2018. <https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/index.html>.

Centers for Medicare & Medicaid Services, 2018. “Projected.” Accessed February 19, 2018. <https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/nationalhealthexpenddata/nationalhealthaccountsprojected.html>.

Enck, S. W., Chapman-McQuiston, E., & Kelly, D. 2015. *United States Patent No. US9646258B2*. Cary, NC: United States Patent and Trademark Office.

Mihaylova, B., Briggs, A., O’Hagan, A., & Thompson, S. G. 2010. “Review of Statistical Methods for Analysing Healthcare Resources and Costs.” *Health Economics*, 20(8):897–916.

SAS Institute Inc. 2017. *SAS Visual Investigator: Admin Guide*. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2017. *SAS Visual Investigator: Fact Sheet*. Cary, NC: SAS Institute Inc.

## ACKNOWLEDGMENTS

The author would like to thank the lengthy list of analysts who have contributed ideas and hours of work on these health care analytics solutions in the Global Hosting & US Professional Services Division including Sarah Mohamed, Ruth Baldasaro, Sarah Rawls, Dan Kelly, Steven Enck, Jay King, Chas. Cavalier, Diane Emerton, William Nadolski, Zoya Asgari, Katherine Grimshaw, Fusun Meric, Eduardo Brasileiro, Bong Choi, Susan Adams, Patrick Williams, Mauricio Alvarez, Eddie Rowe, Suliko Ayvazov.

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Emily Chapman-McQuiston  
SAS Institute Inc.  
emily.mcquiston@sas.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.