Taming the Rule

Charlotte Crain, Chris Upton, SAS Institute Inc.

ABSTRACT

When business rules are deployed and executed--whether a rule is fired or not—if the rule-fire outcomes are not monitored or investigated for validation or challenged, over time unintended business impacts can occur because of changing data profiles or characteristics of the input data for the rules. Comparing scenarios using modified rules and visually comparing how they might impact your business can aide you in meeting compliance regulations, knowing your customers, and staying relevant or accurate in your particular business context. Visual analysis of rules outcomes is a powerful way to validate what is expected or to compare potential impact that could lead to further investigation and refinement of existing rules. This paper will show how to use SAS® Visual Analytics and other visualizations to perform various types of meaningful and useful rule-fire outcome analysis with rules that have been created in SAS® Business Rules Manager. Using visual graphical capabilities can give organizations or businesses a straightforward way to validate, monitor, and keep rules from running wild.

INTRODUCTION

Business rules represent the business logic that drives decisions in companies. These business rules are often found in government policies, regulations, tax law, financial guidelines, accounting policies, call center policies, federal and local laws, and in every business process with varying complexity and application. Moreover, business rules exist and can remain the same for months to years depending on the overall business process or resources available internally for making changes. Business rules can change frequently (for example, in marketing and customer relationship applications). And the business should be able to respond quickly as needed, which may require modifying and validating rules and then quickly deploying or redeploying. The investigation of validating existing rules or examining what rules could or should change is more easily done if visual graphing capabilities are used. The recommended updates should be versioned for reporting and auditing purposes. Because business rules are integral to overall decision making processes and one decision could be hundreds or thousands of rules, having a way to analyze potentially large or varying outcomes with relative ease and speed can provide the necessary or desired monitoring an organization desires.

This paper will present some graphical explorations and reports for analyzing business rules that have been created or authored using the SAS Business Rules Manager and executed against an input table. The goals are to examine rule-fire outcomes, provide an analysis of findings and overall recommendations, and also to briefly examine the rules without outcomes. These visualizations are examples and are not meant to be a comprehensive scope for all analyses possible. You will gain insight and understanding of how this can be done using SAS Visual Analytics and SAS Business Rules Manager.

ANALYZING BUSINESS RULES WITH OUTCOMES

Assessing or examining a business rule either as an object by itself without outcomes or as the rule-fire outcomes generated during execution can help you understand why results were produced; if the correct rules are being applied; or if some of the rules should be challenged and ultimately changed. In any analysis, understanding depends on domain expertise, what we are looking for in terms of relationships, and what do we want to know or should know in terms of behavior of our data. For analytics, we perform exploratory analysis and have expert domain knowledge and build models. Business rules that are separate from the analytical model rules (but could interact or be linked with analytical models) are typically written by business analysts for other business use cases such as cost allocation or revenue attribution, data tagging or enrichment, or additional information that is not conducive to being part of a model.

EXAMPLE – WHICH CARD TO ISSUE

A credit card business is currently running a set of business rules that helps them make a decision about which credit card to offer their customers. The group under investigation consists of customers who make less than $40,000 and have good or excellent credit. These customers are given a recommended action to issue a Standard credit card. Some visual analyses of current outcomes can provide insight into how you could improve your rules by leveraging your business acumen and expert knowledge you already have about your customer or rules application.

Customer characteristics used in the current rules include the following: total income, credit history, credit utilization, total credit card balance, home ownership, do they have a loan, how many months into their loan, have they maxed out their credit card, and do they have rolling accounts. These are the criteria used selectively within a particular business rule. Table1 below shows a partial list of the 44 total rules used in this example. The rule names seen in Table 1 are shortened for brevity and are arbitrary (R1, R2, R3, and so on). All of these rules are being used in the Which_Card_To_Issue rule set. Your rules and rule sets would be named something meaningful to you. There are several conditions where the Standard action is recommended and rule one, R1: Total_Income < 40000 and Credit_History in ('Good','Excellent'), will be the main rule discussed in this example and identified by the rule set name and rule name (that is, Which_Card_To_Issue_R1 as shown in the blue outlined box in Table 1).

Which_Card_to_Issue: Version 1.6 Rules Logic (44 Rules)

#	Rule Name	Condition	Action
1	R1	If (Total_Income < 40000) AND (Credit_History in ('Good' , 'Excellent'))	then Action_To_Take = 'Standard'
2	R2	Else if (Total_Income < 40000) AND (Credit_History = 'Fair')	then Action_To_Take = 'Refer'
3	R3	Else if (Total_Income >= 40000 and Total_Income < 60000) AND (Credit_History = 'Fair') AND (Credi..	then Action_To_Take = 'Refer'

Table 1. Rules Decision Table for Which_Card_To_Issue

Table 2 below shows the input table that has 3404 rows which the rule set will run against. There is additional information in the input table such as whether a customer is profitable. Profitable was derived using an external analysis and was not available on the input table when the original rules were written. Profitable is not currently being used in any of the existing rules. Credit utilization is being used for some of the Standard action classifications of the Which_Card_To_Issue rules, but is not being used in the Which_Card_To_Issue_R1 rule. The four columns available on the input table used in visual analysis are the following: Profitable, credit utilization (Credit_Util), credit history (CreditHist) and total income (Tot_Income). Because the column Profitable is not being used in current rule logic anywhere but is desired on the output table, the profitable input column is passed through and available on the output table shown in Table 3.

Tests	Results											
Test:	WhichCardToIssue_Test_BASELINE		Version: Current									

	ID	Credit_Util	CreditHist	CreditMax	Tot_Card_...	Tot_Income	HomeOwn...	RollingAcct	Profitable	CollegeSt...	DEBTINC	VALUE
Input T... Output... Rules ...	ID_000000000004	3	Excellent	0	494	48556	1	1	Y			
Plot	ID_000000000005	13	Good	0	5685	31989	1	1	Y			11
Code	ID_000000000006	19	Good	0	2300	19413	1	0	N	Y	37.113613558	4
Log	ID_000000000007	2	Good	0	11302	14512	0	1	Y	Y		5
Propert...	ID_000000000011	13	Good	1	2140	80009	1	1	Y			
Terms	ID_000000000012	1	Excellent	0	3436	87741	1	0	Y			2
Prepro...	ID_000000000013	0	Excellent	0	8256	206302	1	1	Y			5

Table 2. Input Table for Which_Card_To_Issue Rule Set

Analyzing Rules by Leveraging the Outcomes

After running the Which_Card_To_Issue rule set against the input table above (Table 2) within the SAS Business Rules Manager test interface, several types of results are available including the output table. The output table is shown in Table 3 below. A bar chart plot of the number of rules that fired is shown in Figure 1. When you know how many rules fired, a baseline for analysis and comparison purposes can be set. The Which_Card_To_Issue_R1 fired 1016 times as shown in Figure 1.

Table 3. Output Table Results Generated by Which_Card_To_Issue Rule Set
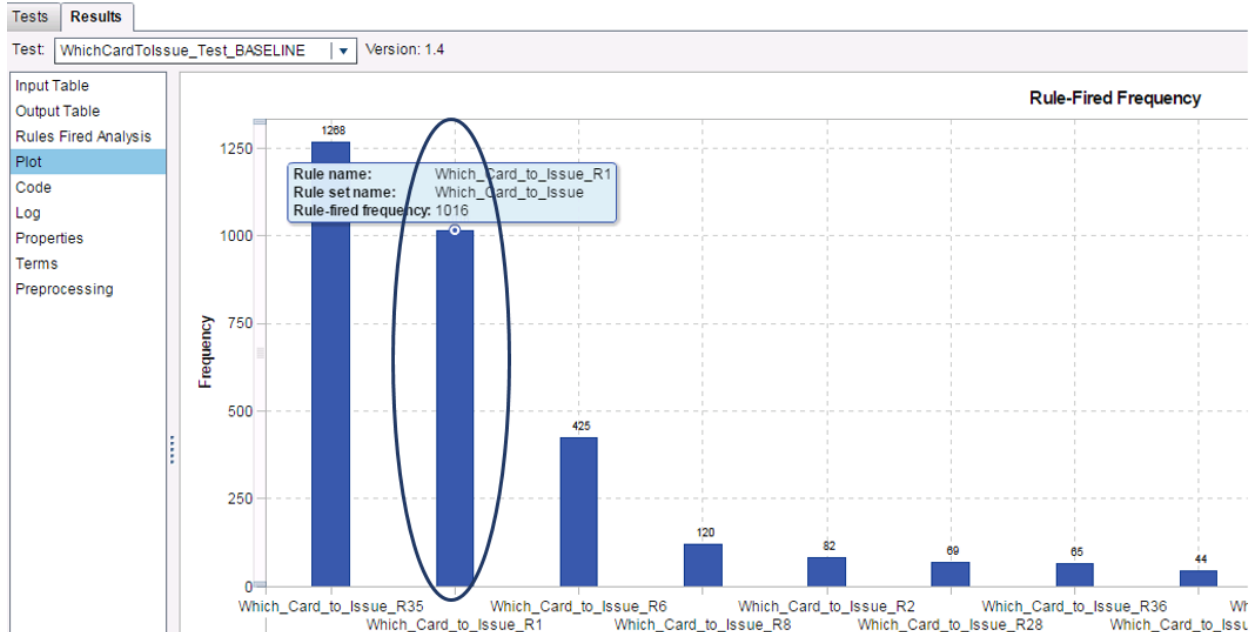

Figure 1. Which_Card_To_Issue_R1 Rules Fired in the SAS Business Rules Manager Testing Interface

Finding Rule Actions That Might Be Challenged

You can graphically examine the output table that was generated by running the Which_Card_To_Issue rule set in SAS Business Rules Manager in SAS Visual Analytics. This output table can be easily uploaded to SAS Visual Analytics whether SAS Business Rules Manager is local to the SAS Visual Analytics machine or not. If not immediately local to the SAS Visual Analytics machine, then a file share or SFTP or a registered table in SAS metadata is needed to be accessible by SAS Visual Analytics. The steps used for this example to upload the rule fire output table to SAS Visual Analytics is shown below in Figure 2. For this example, the SAS Business Rules Manager output is on a file share because the SAS Business Rules Manager server is a different machine than the SAS Visual Analytics application server. There are several ways to upload data to SAS Visual Analytics, which are out of scope for this paper.
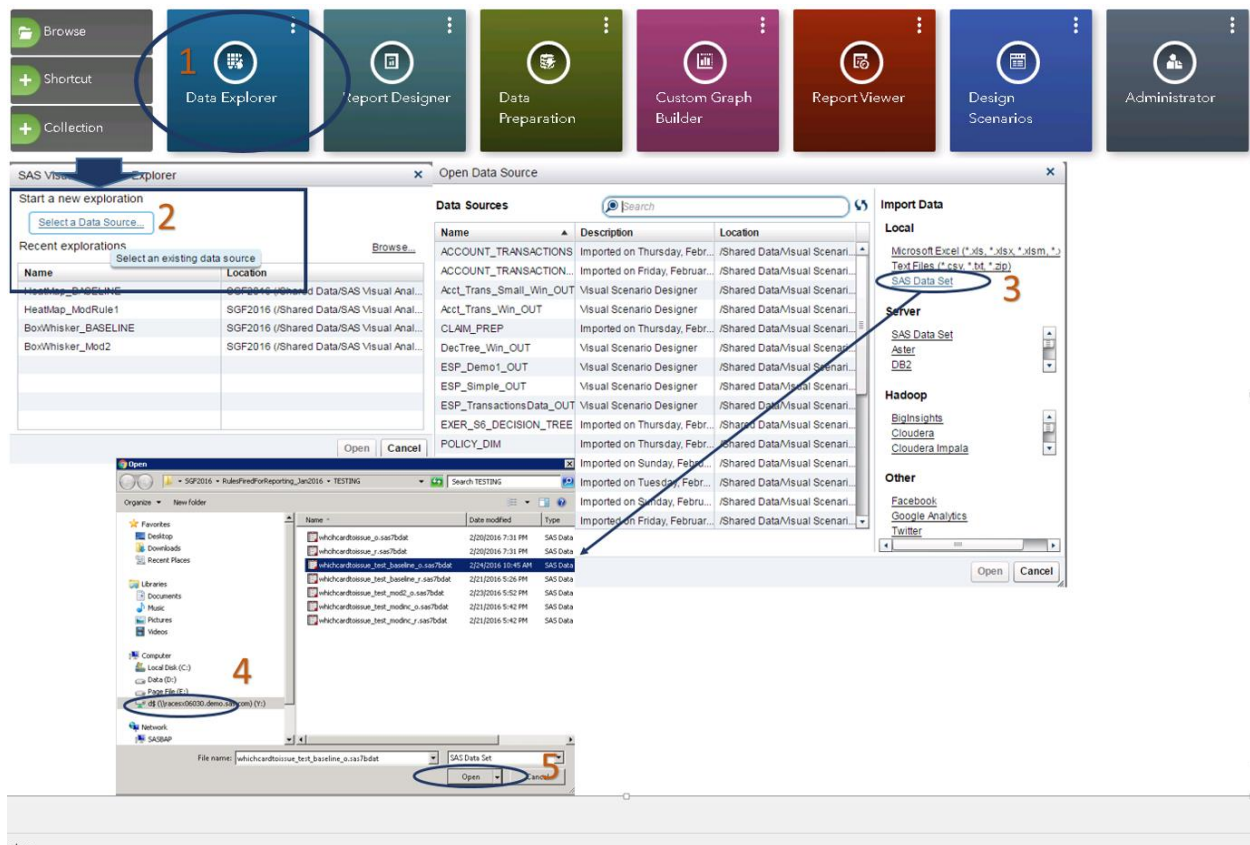
3

Figure 2. Steps to Upload a SAS Data Set or Table to SAS Visual Analytics

One of the first type of analyses you may want to do is to look at the distribution of rule-fire actions using a box-and-whisker plot in SAS Visual Analytics. A box-and-whisker plot is designed to give a picture of the extent of the data as well as its distributional features. You can see (red bracket for emphasis) that for credit utilization, there is a minimum of 3 for the lower whisker or the 25th percentile and a maximum of 38 for the upper whisker or 75th percentile. It is apparent that the variation or spread of Standard action relative to the median or mean (average) for credit utilization is greater than all of the other Actions of Refer, Gold, Platinum, or Diamond as shown below in Figure 3. This shows that perhaps the Standard action population requires additional investigation and indicates an opportunity for further segmentation.
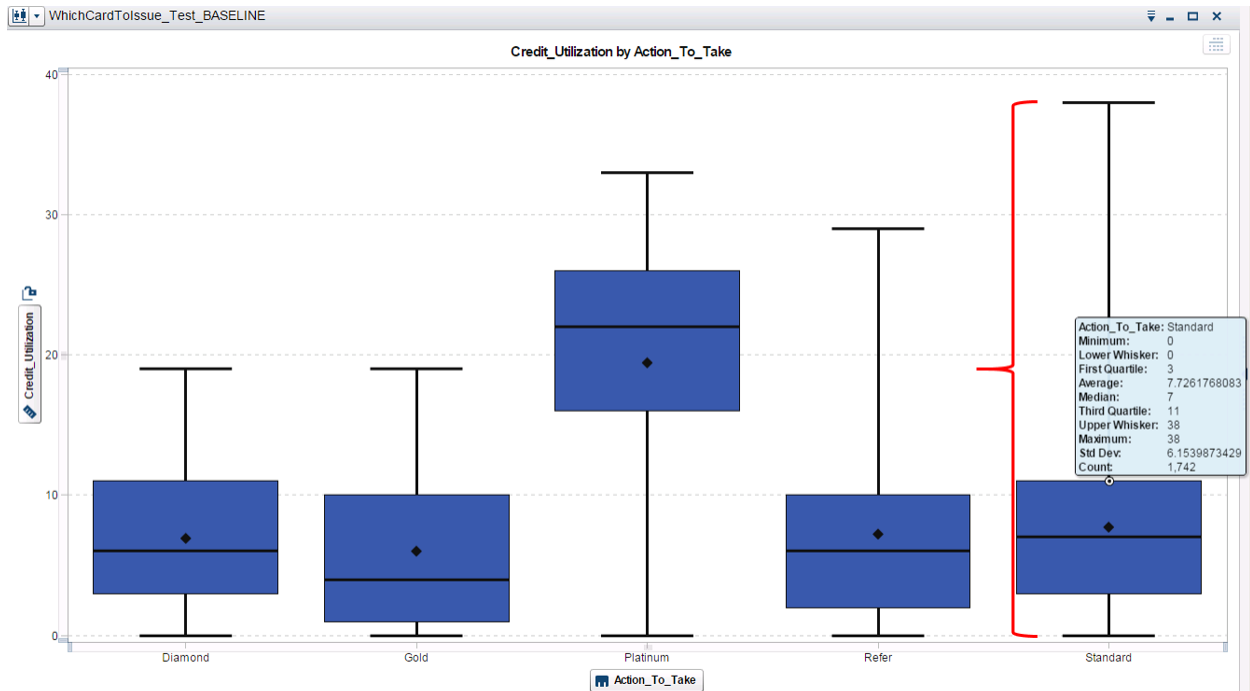
Figure 3. Box-and-Whisker Plot for Credit_Utilization by Action_To_Take

The focus of the spread of the Standard level shown above in the box-and-whisker plot in Figure 2 supports further segmentation. Because the variable profitable is available from input and is not used in any rule for the Standard card recommendation but credit utilization is used in Which_Card_To_Issue_R1, a decision tree analysis can be one type of graphical display used to quickly see if the overall business rule logic coincides with your current business strategy toward your credit card customers. In this example, part of the overall business strategy entails re-evaluating profitable customer groups. The tree was filtered on Standard and Refer based on the box-and-whisker analysis. The decision tree is using credit utilization and profitable as predictors as seen in Figure 4. By looking at the thickest line path of logic that indicates where most of the rules have been fired, you can see that out of 1742 customers who have been issued a Standard credit card with credit utilization less than 19 (which is typically considered an acceptable utilization ratio) 88.37% or 912 are not considered profitable. Is it normal to expect this or would further analysis help provide insight to potentially validate or challenge the current rule logic? The decision tree shows branching when credit utilization is less than 19 or greater than or equal to 19, which might challenge existing rules that use credit utilization. Tree models normally have a lot more leaves unless you have limited the data or depth which is the case for this example. However, being able to see analysis for branching at credit utilization and the profitable relationship is useful. Further investigation is desired.
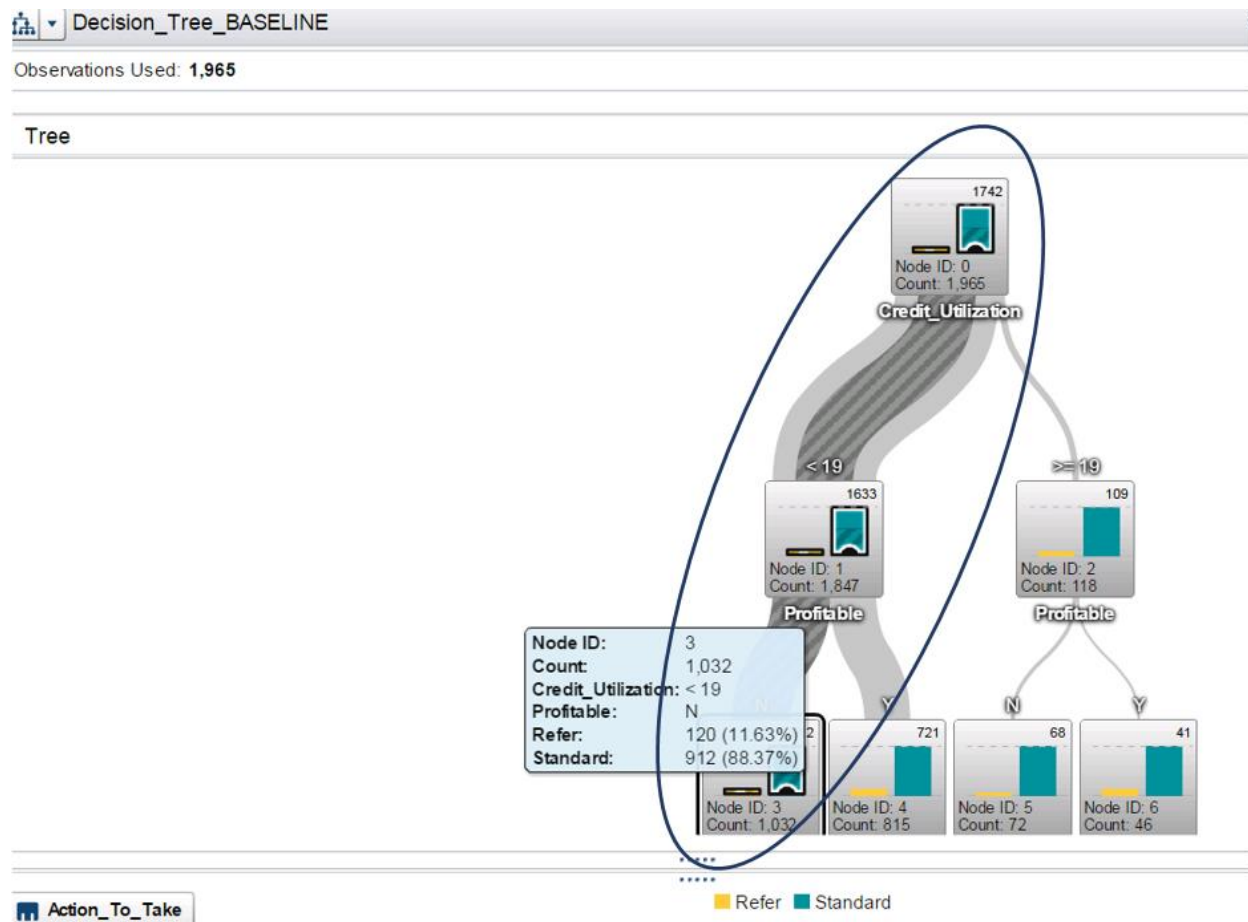
5

Figure 4. Decision Tree from SAS Visual Analytics Exploration

Continuing on with the visual investigation, a heat map can be a useful way to look at rule fire output relationships seen in Figure 5 below. The white boxes show there are no actions recommended by the rules for Diamond, Platinum, and Refer in the respective category and is expected behavior. However, for all of the Standard actions (seen in the gray outlined box), credit utilization is less than the mean of 14.93 (circled in gray). By using "average" credit utilization, the variance or spread of the credit utilization scores is lost information, and no information is drawn from the heat map that would help determine if any rule change is needed. It would be difficult to see this in a table or even by comparing several univariate statistics such as frequency, maximum, minimum, and so on. It would be interesting to investigate profitable as well as it relates to the decision tree, but profitable is not included in any rule yet and the heat map was used for the variables that are used in rules.
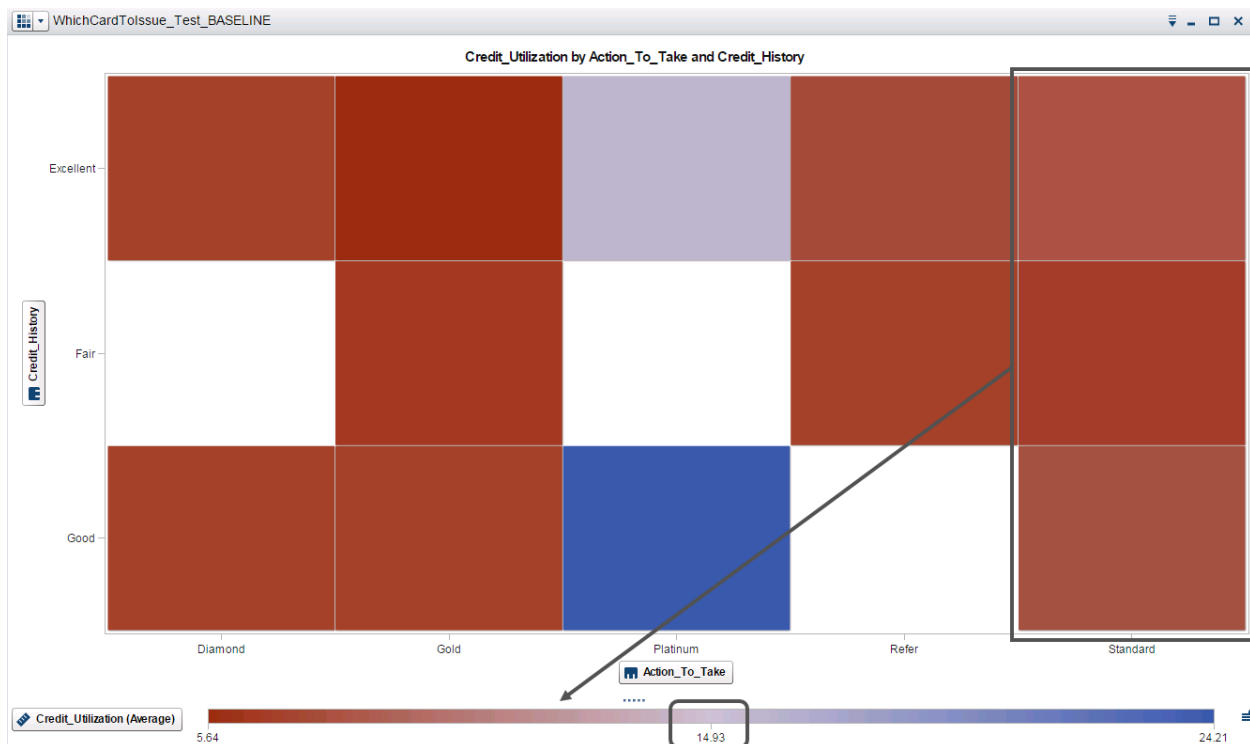
Figure 5. Heat Map of Credit Utilization and Credit History by Action to Take

Using a heat map again for the Standard action and total income (averaged), when credit history is good or excellent, the lower average total income is 28,234.58 in Figure 6. This provides information for potentially modifying Rule 1 where income < 40,000 and to consider other income boundaries.
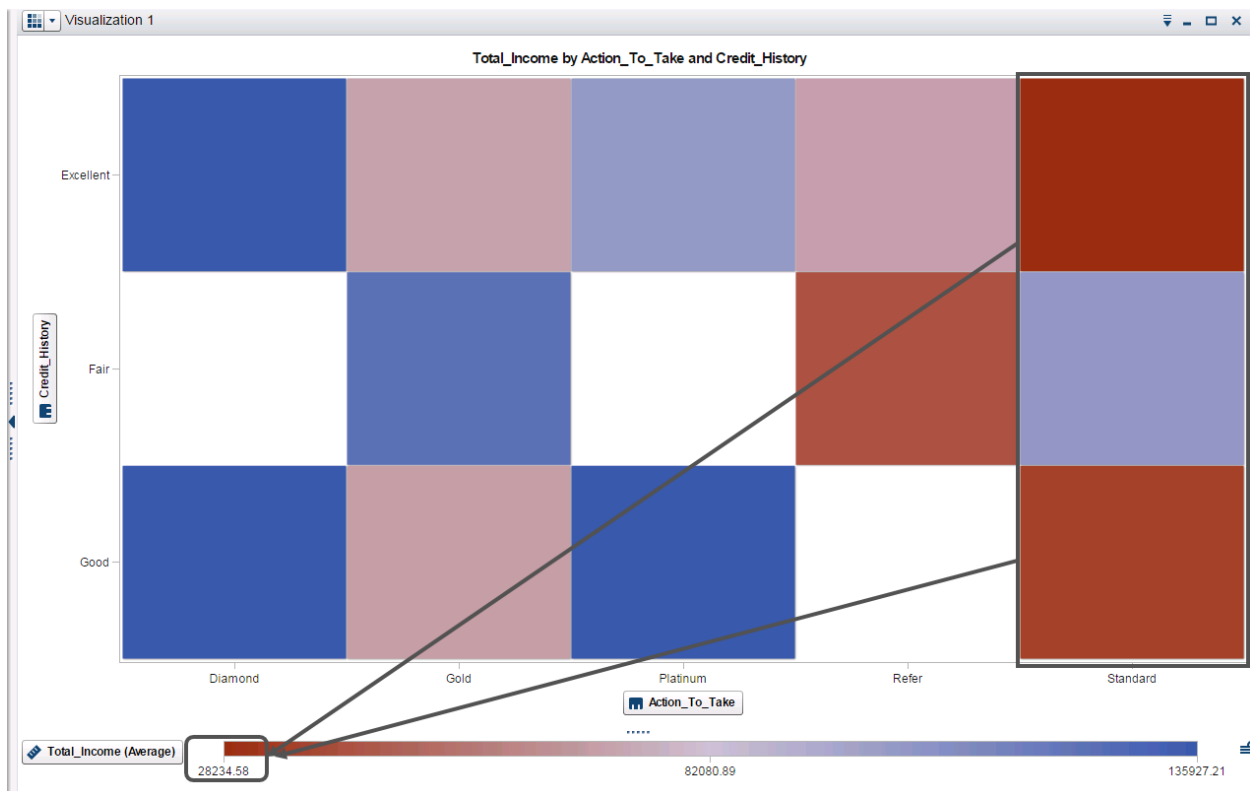
Figure 6. Heat Map of Total Income by Credit History and Action to Take

Re-examining a Component of the Data Population

The visual analysis of the current Which_Card_To_Issue rule set based on the box-and-whisker plot (Figure 4), the Decision Tree (Figure 3), and the Heat Map (Figure 6) present information for potential rule modification. Since credit utilization for customers with a total income less than $40,000 and credit history as good or excellent has not yet been considered in the current rule set logic, one change will be made to total income for Which_Card_To_Issue_R1. The modified rule set will be executed again on the same input table (Figure 2) and outcomes reviewed with some plausible conclusions for next steps.

The Which_Card_To_Issue_R1 rule can be broken down further. Instead of classifying all of the customers whose total income is less than $40,000, the rule is modified as seen below in Figure 7.

			Condition Term				Action Term
#			⑫ Total_Income	⚠ Credit_History	⑫ Credit_Utiliz...	‖ Is_Cre	⚠ Action_To_...
1	If	▼	>20000 and < 40000	in ('Good','Excellent')			'Standard'
2	Else	▼	< 40000	= 'Fair'			'Refer'
3	Else	▼	>= 40000 and < 60000	= 'Fair'	>=15		'Refer'
4	Else	▼	>= 40000 and < 60000	= 'Fair'	< 15		'Standard'
5	Else	▼	>= 40000 and < 60000	in ('Good','Excellent')	>= 30		'Refer'

Figure 7. Change to Which_Card_To_Issue_R1

You could now rerun the updated rule set on the same input data set as before in the SAS Business Rules Manager testing interface and compare the number of times that Rule 1 fired. The modified Rule 1 fires 480 times instead of 1016. (See Figure 1). Figure 8 below shows the results.
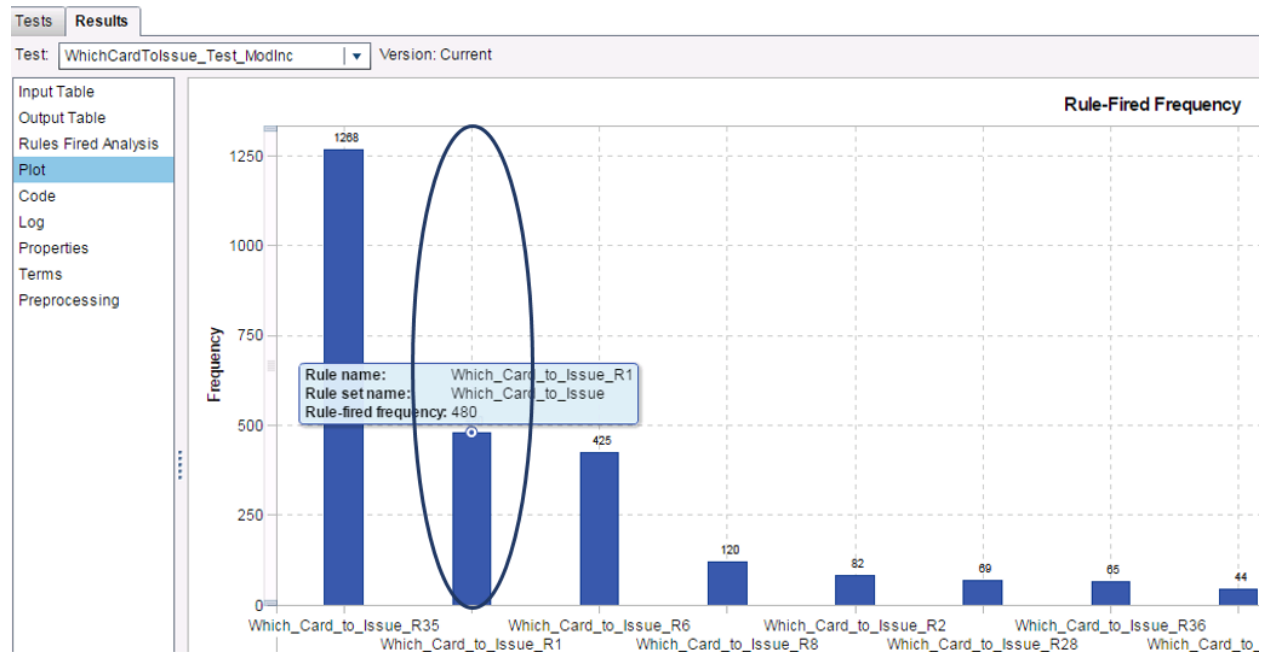

Figure 8. Which_Card_To_Issue_R1 Rule Fired Output after Modifying Rule 1

You can also view the rule fired data for both the baseline and modified income for Which_Card_To_Issue_R1 side by side using SAS Visual Analytics as shown in Figure 9 below. Outlined with the blue box, there were 1,016 records where R1 fired. R1: Total_Income < 40000 and

Credit_History in ('Good','Excellent'), but only 480 records had R1 fire after modifying R1 as: Total_Income > 20000 and Total_Income < 40000 and Credit_History in ('Good','Excellent'). (See Figure 7.) The side-by-side bar chart comparison in Figure 9 for R1 rule fires shows a substantial impact when making only one modification by adding a lower bound of $20,000 for total income versus zero as it was in the original R1.
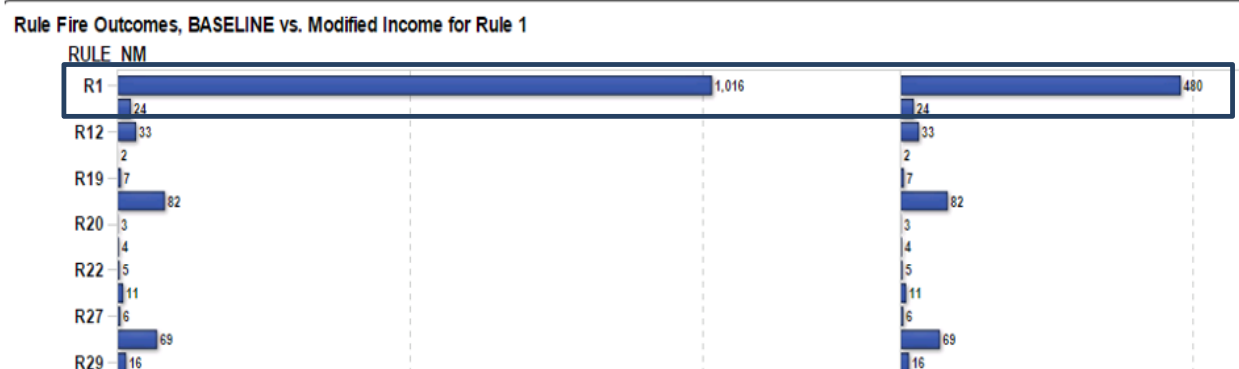

Figure 9. Bar Chart Showing the BASELINE and Modified Income Rule Fire Outcomes for Rule 1

The credit utilization ratio is a key component of your credit score. A high credit utilization ratio can lower your score, while a low credit utilization ratio can raise your score. Credit-scoring formulas assume that consumers who use more of their available credit are riskier borrowers than those who use less of their available credit. From the graphical analyses on the resulting data set after running the Which_Card_To_Issue business rule set, the recommended actions for many Standard card holders with good or excellent credit and a lower credit utilization score might be to upgrade to a Gold card, for example.

Modifying the total income range for the Which_Card_To_Issue_R1 rule might more accurately describe customer reality and pave the way for additional analysis and other rule modifications or additions. For all Standard customers whose credit history is good or excellent, taking into consideration total income and perhaps whether a given customer is profitable, might reflect more realistically what a rule set should be. Figure 10 shows the percentage of Standard actions that are profitable. Additional rules could be created with a different action for customers who were not analyzed specifically in this paper but might have good or excellent credit history and whose incomes are less than $20,000. Perhaps this group consists of responsible college students. A special card could be created or offered to them. Additional analysis and perhaps modeling would be appropriate for a comprehensive examination of this particular group.

In the original Which_Card_To_Issue rule set (Table 1), for customers with a total income less than $40,000, credit utilization has not been used in any rule logic. From what was learned in the exploratory data analyses, credit utilization should be taken more seriously as a criterion for the Standard card holders. In practice, upon agreement with business strategy and marketing, the rule logic would be modified and the entire rule set would be versioned.
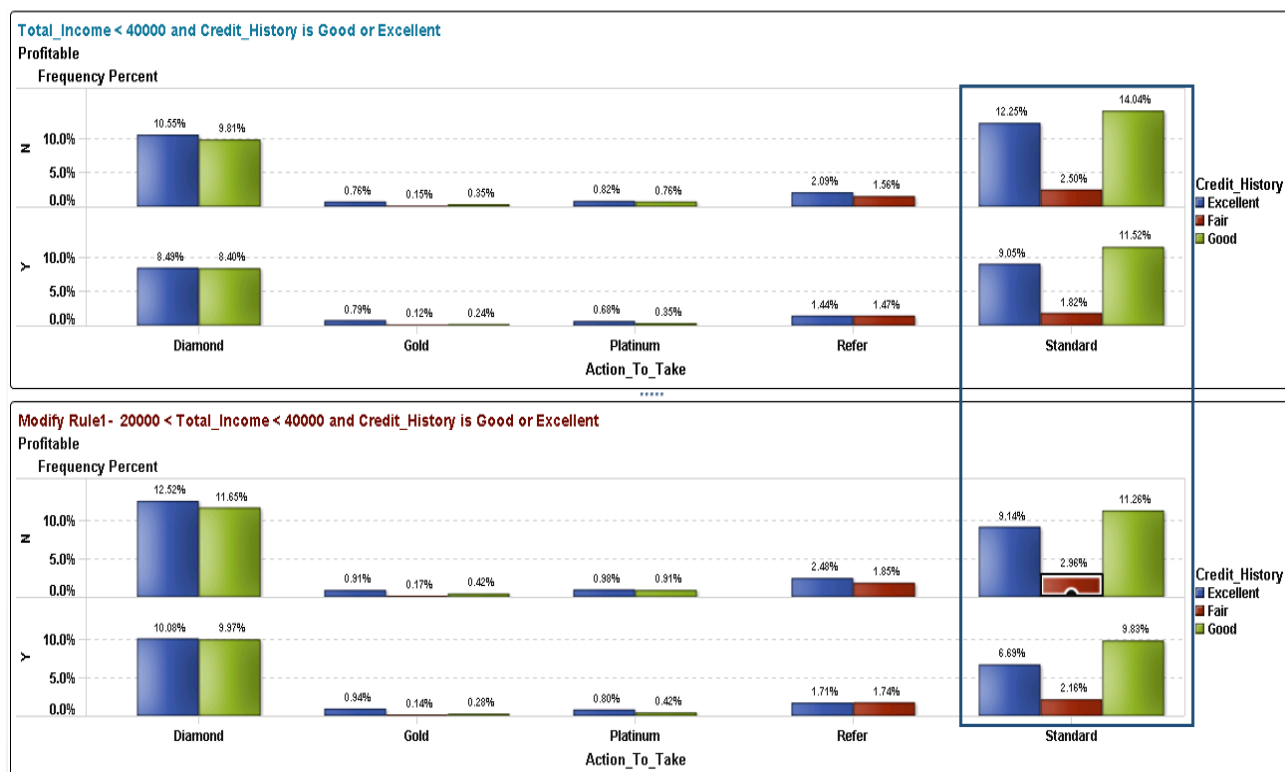
Figure 10. Profitable as a Percentage by Action_To_Take for Standard in Blue Outlined Box

ANALYZING RULES WITHOUT OUTCOMES

Sometimes after authoring rules and before publishing a rule set for production, it can be interesting to see how the rules flow in relationship to each other. This can show a high-level view of a rule set and enable you to visually validate the logic including a rule that was inadvertently missed. Figure 11 shows a heat map of Which_Card_To_Issue and the 44 rules within that rule set with a count by action and condition. For example, Credit_History='Good' is a condition that is found frequently up to 17 times for the Standard action (Action_To_Take='Standard'). It would be difficult to see this in tabular or bar chart form. The heat map can be an easy way to visually inspect overall business rule logic or to detect a rule that should exist and does not. For example, in the white areas (circled in orange) for Action_To_Take=Standard, there are no rules for when credit utilization is less than 15 or greater than or equal to 15. What are other factors in the Standard action when credit utilization is Excellent and no rules are fired? We saw in the Decision Tree (Figure 3) for customers issued a Standard card with credit utilization branched (split) at 19 and where credit utilization < 19 that 1,032 customers are not considered profitable and yet 721 customers are considered profitable. Having a way to visually inspect or assess your overall rule representation between conditions and actions could provide significant information for rule revision and where to look first. The graphic could be used as part of your documentation as a baseline to compare future validations giving a high-level view to monitor your rules. SAS Visual Analytics enables you to filter in many different ways so that if you have thousands of rules, you can apply your filters to look at a given segment of your conditions and actions for the entire set all at once.
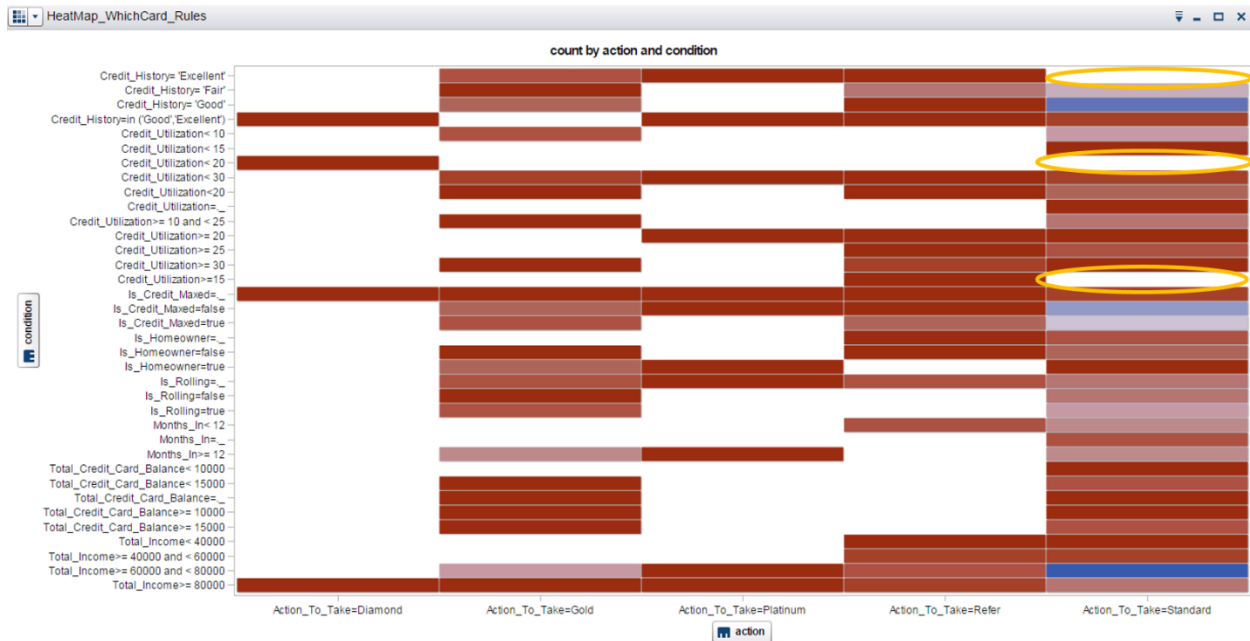
Figure 11. Heat Map of Which_Card_To_Issue Conditions and Actions Count

AUDITING TRAIL

In many industries, particularly health care and banking, there is a set of regulation and compliance policies that must be followed. Violation of these policies can result in substantial fines. The question of "how did you come to a decision or action" is the proof required by the examiners. You might want to know for each input record that was run against a rule set, what was the conditional logic and action and which rules were applied to a given input record to give a particular outcome record. By executing the SAS Business Rules Manager rules using the Business Rules transform inside of SAS® Data Integration Studio, OUTPUT and the DCM_RULE_ACTION_FIRE tables are generated by default. By using join keys _recordCorrelationKey and entity_primary_key (identified as a primary join key from the input table) on the OUTPUT and DCM_RULE_ACTION_FIRE tables, you can create a join table with the input table. Specific queries to the SAS Business Rules Manager repository database (PostgreSQL or Oracle) can be made to create a final table of rule set conditions and action expressions together in syntax form (Figure 12). Each record is a one-to-one map of the input, rules processed and the output or final result. A simple SAS Visual Analytics list report shows several rows of conditions and actions in the Which_Card_To_Issue rule set as shown below in Figure 12, which can help explain auditor questions.

Audit Trail for Which Card To Issue

Customer_ID	Condition	Action_To_Take
ID_000000000004	(Credit_History in ('Good','Excellent')) and (Credit_Utilization < 30) and (Total_Income >= 40000 and < 60000)	Standard
ID_000000000005	(Credit_History in ('Good','Excellent')) and (Credit_Utilization) and (Is_Credit_Maxed) and (Is_Homeowner) and (Is_Rolling) and (Months_In) and (Total_Credit_...	Standard
ID_000000000006	(Credit_History in ('Good','Excellent')) and (Credit_Utilization) and (Is_Credit_Maxed) and (Is_Homeowner) and (Is_Rolling) and (Months_In) and (Total_Credit_...	Standard
ID_000000000007	(Credit_History in ('Good','Excellent')) and (Credit_Utilization) and (Is_Credit_Maxed) and (Is_Homeowner) and (Is_Rolling) and (Months_In) and (Total_Credit_...	Standard
ID_000000000011	(Credit_History in ('Good','Excellent')) and (Credit_Utilization < 20) and (Total_Income >= 80000)	Diamond
ID_000000000012	(Credit_History in ('Good','Excellent')) and (Credit_Utilization < 20) and (Total_Income >= 80000)	Diamond
ID_000000000013	(Credit_History in ('Good','Excellent')) and (Credit_Utilization < 20) and (Total_Income >= 80000)	Diamond
ID_000000000014	(Credit_History = 'Fair') and (Credit_Utilization < 15) and (Total_Income >= 40000 and < 60000)	Standard
ID_000000000018	(Credit_History in ('Good','Excellent')) and (Credit_Utilization < 20) and (Total_Income >= 80000)	Diamond
ID_000000000019	(Credit_History in ('Good','Excellent')) and (Credit_Utilization) and (Is_Credit_Maxed) and (Is_Homeowner) and (Is_Rolling) and (Months_In) and (Total_Credit_...	Standard
ID_000000000020	(Credit_History in ('Good','Excellent')) and (Credit_Utilization < 20) and (Total_Income >= 80000)	Diamond
ID_000000000021	(Credit_History in ('Good','Excellent')) and (Credit_Utilization < 20) and (Total_Income >= 80000)	Diamond
ID_000000000025	(Credit_History in ('Good','Excellent')) and (Credit_Utilization < 20) and (Total_Income >= 80000)	Diamond
ID_000000000026	(Credit_History in ('Good','Excellent')) and (Credit_Utilization) and (Is_Credit_Maxed) and (Is_Homeowner) and (Is_Rolling) and (Months_In) and (Total_Credit_...	Standard
ID_000000000027	(Credit_History in ('Good','Excellent')) and (Credit_Utilization) and (Is_Credit_Maxed) and (Is_Homeowner) and (Is_Rolling) and (Months_In) and (Total_Credit_...	Standard
ID_000000000028	(Credit_History in ('Good','Excellent')) and (Credit_Utilization < 20) and (Total_Income >= 80000)	Diamond
ID_000000000032	(Credit_History in ('Good','Excellent')) and (Credit_Utilization) and (Is_Credit_Maxed) and (Is_Homeowner) and (Is_Rolling) and (Months_In) and (Total_Credit_...	Standard
ID_000000000033	(Credit_History = 'Excellent') and (Credit_Utilization < 30) and (Is_Credit_Maxed FALSE) and (Months_In < 12) and (Total_Income >= 60000 and < 80000)	Refer
ID_000000000034	(Credit_History = 'Fair') and (Total_Income < 40000)	Refer
ID_000000000035	(Credit_History in ('Good','Excellent')) and (Credit_Utilization < 20) and (Total_Income >= 80000)	Diamond
ID_000000000039	(Credit_History in ('Good','Excellent')) and (Credit_Utilization < 20) and (Total_Income >= 80000)	Diamond
ID_000000000040	(Credit_History = 'Good') and (Credit_Utilization >= 10 and < 25) and (Is_Credit_Maxed FALSE) and (Is_Rolling TRUE) and (Total_Credit_Card_Balance >= 15000)...	Standard
ID_000000000041	(Credit_History = 'Good') and (Credit_Utilization < 10) and (Is_Credit_Maxed FALSE) and (Months_In < 12) and (Total_Income >= 60000 and < 80000)	Standard
ID_000000000042	(Credit_History in ('Good','Excellent')) and (Credit_Utilization) and (Is_Credit_Maxed) and (Is_Homeowner) and (Is_Rolling) and (Months_In) and (Total_Credit_...	Standard
ID_000000000046	(Credit_History = 'Excellent') and (Credit_Utilization < 30) and (Is_Credit_Maxed TRUE) and (Total_Income >= 60000 and < 80000)	Gold

Figure 12. Conditions and Actions for Each Input Record for Which_Card_To_Issue Rule Set

CONCLUSION

The value of business rules analyses with SAS Business Rules Manager and SAS Visual Analytics is to provide a fast and relatively easy way to visually understand the impact of rule outcomes as well as graphically describe the rules without outcomes. Using SAS Visual Analytics as your visual graphics analysis tool enables you to see dimensions of your rule conditions, actions, or outcomes in ways not possible with tabular or simple statistics. Loading your data to SAS Visual Analytics and then freely exploring relationships between available input columns that may not be used in rules and the conditions and actions in your rules gives you extended creativity for exploring rule outcomes. You can change graphics interactively for different analyses until you find the best graphic to display the information you need to explain outcomes and potentially improve decisions. You can create reports that support governance for mandatory auditing and compliance regulations for some industries. The reporting intervals can be adapted for whatever frequency of monitoring or analyses needed if the underlying data is available.

The initial deployment or redeployment, of rules after testing and viewing outcomes can be implemented quickly on a unified SAS framework. All rule sets are versioned and for any changes and redeployment. Rule sets are automatically versioned up from the previous version as changes are made, which provides a level of governance, accuracy, and transparency expected from a rules manager.

Data preparation detail for the analysis exploration and reporting as well as graphing details will be provided from the authors upon request.

REFERENCES

ACKNOWLEDGMENTS

RECOMMENDED READING

Duling, David. 2015. "Make Better Decisions with Optimization." *Proceedings of the SAS Global 2015 Conference.* Cary, NC. SAS Institute Inc. Available http://support.sas.com/resources/papers/proceedings15/SAS1785-2015.pdf.

Crain, Charlotte, and Steve Sparano. 2015. "Next Generation: Using SAS® Decision Manager to Modernize and Improve Your Operational Decisions." *Proceedings of the SAS Global 2015 Conference.* Cary, NC. SAS Institute Inc. Available http://support.sas.com/resources/papers/proceedings15/SAS1912-2015.pdf.

Duling, David, Steve Sparano, and Charlotte Crain. 2014. "Introducing SAS Decision Manager." *Proceedings of the SAS Global 2014 Conference.* Cary, NC. SAS Institute Inc. Available http://support.sas.com/resources/papers/proceedings14/SAS276-2014.pdf.

SAS Visual Analytics 'Try interactive demos'. Available http://www.sas.com/en_us/software/business-intelligence/visual-analytics.html.

SAS Visual Analytics Documentation. http://support.sas.com/documentation/onlinedoc/va/index.html.

CONTACT INFORMATION

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

Charlotte Crain
Charlotte.crain@sas.com

Chris Upton
Chris.Upton@sas.com