Paper SAS2900-2016

The Six Tenets of a Better Decision

Dan Jahn, SAS Institute Inc., Cary, NC

ABSTRACT

SAS® helps people make better decisions. But what makes a decision better? How can we make sure we are not making worse decisions? There are six tenets to follow to ensure we are making better decisions. Decisions are better when they are: (1) Aligned with your mission; (2) Complete; (3) Faster; (4) Accurate; (5) Accessible; and (6) Recurring, ongoing, or productionalized. By combining all of these aspects of making a decision, you can have confidence that you are making a better decision. The breadth of SAS software is examined to understand how it can be applied toward these tenets. Scorecards are used to ensure that your business stays aligned with goals. Data Management is used to bring together all of the data you have, to provide complete information. SAS® Visual Analytics offerings are unparalleled in their speed to enable you to make faster decisions. Exhaustive testing verifies accuracy. Modern, easy-to-use user interfaces are adapted for multiple languages and designed for a variety of users to ensure accessibility. And the powerful SAS data flow architecture is built for ongoing support of decisions. Several examples from the SAS Solutions OnDemand group are used as case studies in support of these tenets.

INTRODUCTION

Instinctually, humans are inclined to want to help each other. This drive is so strong that it is not uncommon for people to neglect themselves in order to improve the lives of others. A data scientist helps people make better decisions about a huge variety of topics. It is always important to keep in mind who is making the decisions and what decisions they are making. For example:

- 29,000 criminal justice professionals in North Carolina use the CJLEADS application to make better decisions because they have a single place to look for offender data from more than 100 data sources. When police officers, judges, attorneys, sheriffs, magistrates, state troopers, investigators, et al. make better decisions, everyone is safer, and that can save lives.
- Technicians who maintain large equipment such as engines, turbines, trucks, and oil wells make better decisions when software can almost instantly look at the censor data created by that equipment. When maintenance technicians make better decisions, their equipment runs better, they save money, and we are safer. Think aircraft engines and nuclear power plants.
- Service contract negotiators decide what price to place on a contract by using software that assembles the historical cost of the contract—that is, the cost over the life of the contract.
- Inventory managers decide how much inventory to have on hand and when to have it. Railroads run more efficiently, planes are safer, and you are more likely to get to your destination on time.
- Fraud investigators make better decisions about who to investigate and what to examine. Fraud is rampant—in part because it is difficult to detect. Discovering fraud saves money for insurance companies, state governments, federal governments, and ultimately you.
- Doctors make better decisions about which treatments are most effective for a given set of symptoms. Why is it that everyone who has a headache should take two aspirin? What is the most effective dosage? And for that matter, maybe ibuprofen or acetaminophen would be more effective. What is the best treatment for your case of Kaposi's sarcoma, and is that treatment different for someone else with different genetics? These better decisions that doctors make can save your life.
- Quality engineers make better decisions about the materials they purchase that are used to build railroad tracks across the country. Engineers can determine which materials are not up to standard, which materials to reject, and who the best suppliers are.
- Social workers determine which children are most likely to encounter abuse and neglect, giving them an opportunity to intervene earlier.

This list is not anywhere near the complete list of decisions that people make. That complete list won't be finalized until all the data in the world and beyond is collected. This list is just a few of the projects SAS

Solutions OnDemand has developed over the last few years. The work done on these projects has helped define the six tenets of a better decision.

There are two levels of people who consume SAS software—those who create applications and implement projects using the software and those who make decisions using the applications.

It is very important to ensure that we are helping people make better decisions. Likewise, it would be very bad to help someone make a worse decision.

There are six tenets to follow to ensure that you are helping people make better decisions. These tenets can be used as a guide to not only make sure you are helping people make better decisions but to help people understand the gamut of help you can provide.

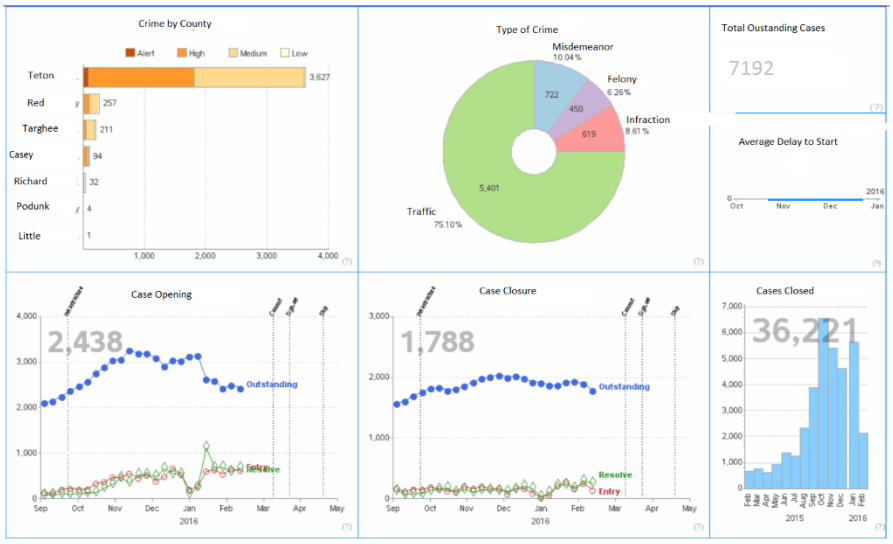
TENET 1: ALIGNED WITH YOUR MISSION

A good decision must be aligned with the mission of an organization. If an organization builds bicycles and someone decides to start selling computer programs, then that is probably not a good decision. It is important for everyone in an organization to know what the mission is and to know how well the organization is progressing toward that mission. Many organizations define key performance indicators (KPIs) to identify their objectives, to communicate those objectives, and to measure progress toward those objectives. Ideally, everyone in the organization (not just the executives) has access to the dashboard to help communicate the objectives as well as to incentivize people.

SAS helps people make decisions aligned with organizational goals by helping to communicate those goals and progress in dashboards. After gaining an understanding of the business and available data, a set of key performance indicators is defined. This set of KPIs can be displayed in a dashboard. It might also make sense to create a roll-up of several KPIs. An example is reliability measures and the Risk Priority Number (RPN) when looking at specific entities to help users assess the overall risk of failure (such as an offender to recidivate, a machine to break down, or a person to get sick again). It is useful to drill down in a dashboard to explore the drivers of the summary that the dashboard displays.

Dashboards can be created using the Output Delivery System (ODS) and be displayed as stored processes, or they can be created as reports in SAS Visual Analytics.

Figure 1. A Typical Dashboard



TENET 2: COMPLETE

Better decisions are based on more information. The more data that is used to make a decision, the more informed the decision is. Bringing disparate data sources together can often yield new insights and information that were never considered when the original systems were created.

For example, in the CJLEADS application, when the concealed handgun permit data was combined with the court data, it was discovered that many people still had a permit who also had a felony. (Being convicted of a felony should have revoked the permit.)

GET ALL THE DATA AND BRING IT TOGETHER

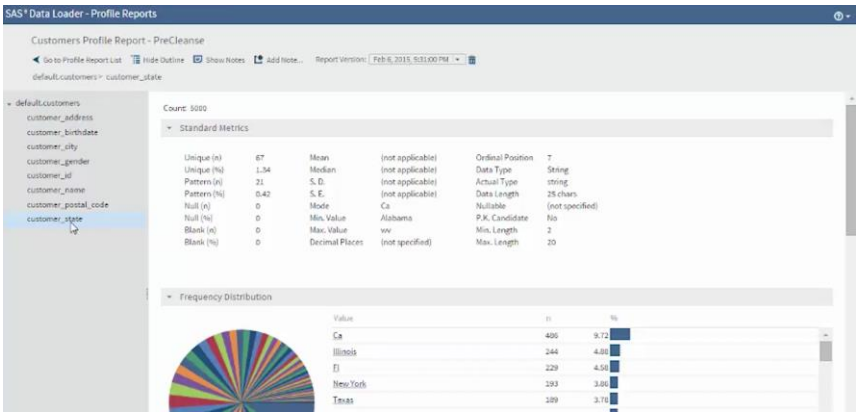
Data management is the art of preparing data for viewing or analyzing. Data management is a part of data governance, which defines policies to ensure consistently good and secure data. In data management, there is a series of typical steps used to prepare data for visualization and analysis. The steps start with getting access to the data—that is, data intake. The next step is to gain an understanding of the data by looking at a data profile. Using the data profile, business rules are defined for cleaning the data. One of the key steps is entity resolution, which is also known as matching, merging, clustering, or record linkage. The goal of entity resolution is to combine records when there might not be a common key or ID. Finally, the data is loaded into a system for analysis.

Figure 2. Steps of Data Management

SAS provides numerous tools for performing the tasks of data management:

1. Data intake can be performed by either extracting the data using the source system or by accessing the data in the source system using the SAS/ACCESS® products.
2. A data profile can be created in several ways. A simple profile can be created using Base SAS® and the FREQUENCY procedure. SAS® Master Data Management and SAS® Data Loader provide more advanced data profiling as illustrated in Figure 3.

Figure 3. Data Profile from SAS Data Loader for Hadoop



3. SAS match codes (available through DATA step functions such as DQMATCH or web services in SAS® Data Quality Server) are extremely useful for data cleansing. The match code can make

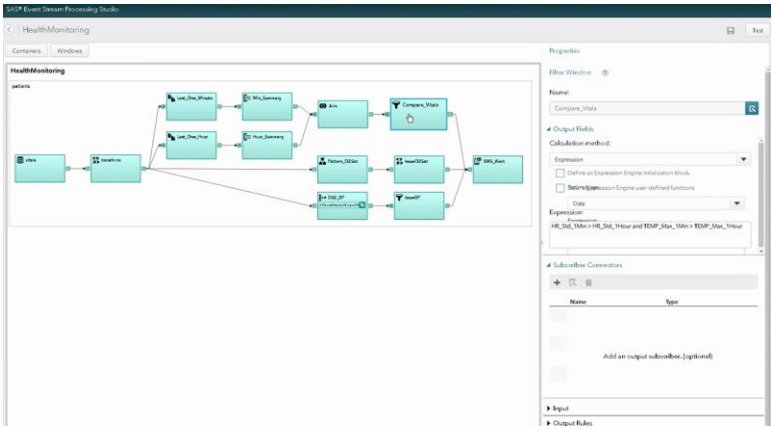
names and addresses that humans know are the same look the same to a computer. For example, “Bill” and “William” would resolve to the same match code. SAS also provides a variety of functions that help with parsing data (for example, addresses using the DQPARSE function).

4. There are a variety of options for performing entity resolution. SAS code can be used (typically in conjunction with match codes) to create either rule-based or probability-based matching. SAS Master Data Management has a user interface that can help with entity resolution; SAS® Data Integration Studio provides a variety of tools to help with entity resolution; and the HPENG procedure can be used to perform entity resolution in the high-performance server.
5. Clean data with resolved entities can be loaded into a server for analysis. Typical approaches are to use a state vector or a temporal vector, but the actual model of data depends on the type of analysis to be performed.

DEAL WITH HUGE VOLUMES OF DATA—SAS® EVENT STREAM PROCESSING

Getting all of the data doesn’t mean all the data must be kept. In the highly connected world we live in, the amount of data that is constantly flowing from and between devices is astronomical. The name of the game for dealing with large amounts of data is to find the meaningful data in real time. SAS Event Stream Processing is excellent at handling large amounts of data (millions of records per second). It can look at every record and create real-time alerts or actions based on customized rules. For maximum flexibility in the rules, use the SAS DS2 language.

Figure 4. SAS Event Stream Processing



KNOW WHAT YOU DON’T KNOW

Systems will never have all the data that can impact a decision. It is important to notify users what data a system does not have. For example, if a system brings together all of the data about a person but doesn’t have any data from foreign countries, then users should be notified that the system doesn’t have any data from foreign countries. In this way, users can know what they don’t know.

TENET 3: FASTER

Making a faster decision can result in an organization outperforming competitors. A faster decision could also save lives (for example, if a doctor is making a diagnosis, a police officer is preventing a crime, or an engineer is fixing an aircraft).

There are two perspectives on making a faster decision: one is the amount of time it takes from the beginning of a project to the availability of the information (reducing time to value). The other perspective is the amount of time it takes every time new data is received.

REDUCING TIME TO VALUE

Time to value is the amount of time it takes from the inception of a project to when the decision makers are making better decisions. SAS Solutions OnDemand specializes in reducing that time to value by providing hosting in the cloud as well as a team of professionals who work closely together to bring projects to life. This team includes:

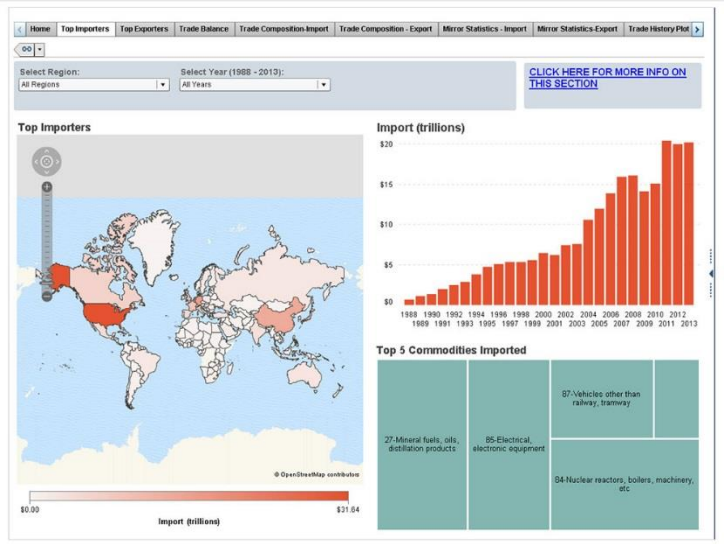
- Sales—works with people to create the vision and set the expectations of how data can be transformed into information and help people make better decisions.
- Project managers—orchestrate the efforts of all team members to track and streamline progress on a project while coordinating with customers on goals.
- IT administrators—configure machines and networks that are well proven at working with data.
- Database administrators—configure databases for optimal analytical use.
- SAS administrators—install and configure SAS software tuned for each project.
- Technical architects—design hardware and software systems to meet project goals.
- Data management engineers—implement the data management steps outlined earlier.
- Tech leads—make the technical decisions for the project in consultation with the appropriate team members.
- Business intelligence engineers—generate dashboards and reports that help people understand information contained in the data.
- Analytical engineers—implement statistical algorithms and understand how to translate numerical algorithms that run against data into information that can help people make better decisions.
- Quality assurance engineers—validate that the entire system is meeting the needs of decision makers.
- Operational engineers—monitor running systems and provide ongoing system support.

The team brings a breadth of knowledge and skills to a project, while each team member brings a specialized depth of knowledge.

THE ADVANTAGE OF FAST SOFTWARE

Reducing the amount of time it takes to make a decision based on additional data results in better decisions. SAS provides many tools and techniques for speeding up data processing. One of the most interesting is SAS Visual Analytics, which uses in-memory processing to allow massive amounts of data to be processed in short periods of time. It is not uncommon for tasks that took 10–12 hours to be reduced to minutes. Data mining, correlation, frequencies, and many other tasks are very suitable to in-memory processing.

Figure 5. SAS Visual Analytics



TENET 4: ACCURATE

If inaccurate information is presented, then people might be misled into making worse decisions rather than better ones. Inaccuracy could come from not properly performing entity resolution, miscoding an algorithm, or a variety of other reasons. The best decisions are based on accurate information. It could be tragic to tell someone the answer is 3 when it is actually 2.

SAS software undergoes a great deal of testing rigor during its development. Here are a few of the testing techniques:

- Cross-checking—compares results of an existing validated procedure with a procedure that has not yet been validated.
- Published results—are used from publications and journals such as the *International Journal of Forecasting* and compare the results of a SAS procedure against what is published in a paper.
- Independent validation using SAS/IML®—rewrites the algorithm in another language.
- Regression testing—verifies things continue to work as they always have.

Throughout the processes, testers must be aware of the nuances of machine precision differences. The key to quality is following processes—knowing what the changes are, what the scope of impact is, and when they happened are all important. One key process to follow is to make sure someone other than the original author of the code is responsible for testing it.

The bottom line is that to make reliably better decisions, you must use software that has been thoroughly tested.

TENET 5: ACCESSIBLE

There are three perspectives on making information accessible: putting it in a place where people can find it (such as the web); following guidelines so that all people can read it (such as the Web Content Accessibility Guidelines or WCAG); and presenting the information in ways that people can comprehend. People need all three perspectives in order to understand the information and make better decisions.

Putting information on the web makes that information widely available to many people across an organization. Health organizations in poorer countries are now using mobile phones (in place of posting

notices or even going door-to-door) to distribute information because most people, regardless of financial status, have a cell phone. Businesses make their dashboards available on intranets to help everyone understand their progress on organizational objectives.

There are many techniques that can be used to make information available to people with reduced sight. These include using colors that are visible to people with color blindness or using touch or sound to convey the content of the information. For more information, see "Using SAS/GRAPH® to Create Visualizations That Also Support Tactile and Auditory Interaction" available at https://support.sas.com/resources/papers/proceedings12/279-2012.pdf.

It is also important to present information in a way the decision maker can understand it. For example:

- "The usable level of BTEX compounds, alkanes, cycloalkanes, and asphaltenes has been reduced to 1.43223 with a standard deviation of 0.000323."
- "You are almost out of gas."

The presentation might need to be different for a petroleum engineer versus an average driver of a car. When presenting information, it is always important to make it accessible to that user by presenting information in a way the decision maker understands.

TENET 6: RECURRING, ONGOING, OR PRODUCTIONALIZED

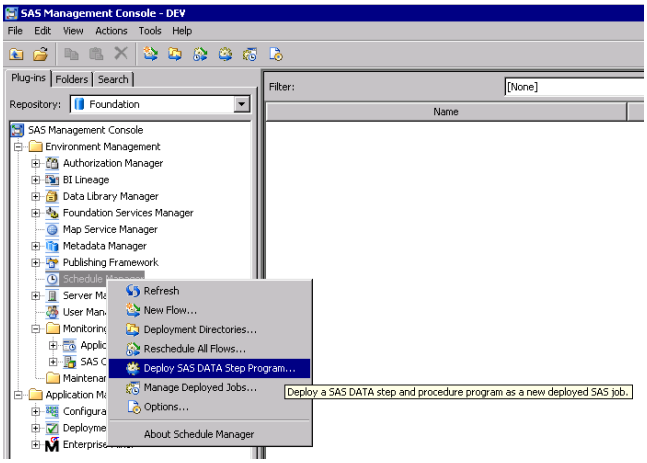
At SAS Global Forum 2015, a keynote presentation by Jeff Ma demonstrated the difference between making the right decision and getting the right outcome. Statistics does not guarantee an outcome—it only helps us understand the likelihood of possibilities. Over time, if you continue to make the right decisions, you will cumulatively get it right more often than not. This is why it is important to productionalize a project. Productionalization is making a data analysis process run over and over again automatically and reliably, which enables users to continue to get the latest data and use the latest data to make better decisions.

The continuous updating of data can be done in different ways:

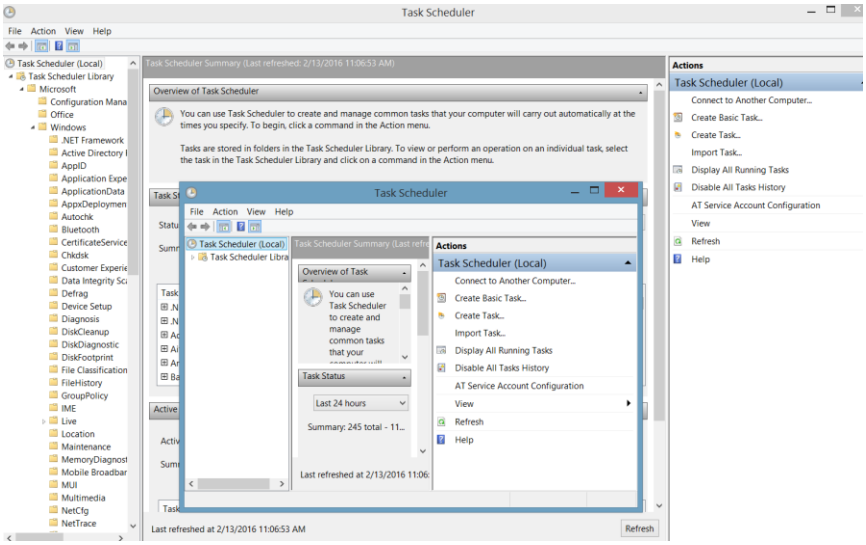
- Data updates can be made as data arrives. This is frequently called "real time" (although real time could also mean updates occur as frequently as you need them to, not as frequently as they possibly could.) SAS BI Web Services is a great tool for this because it can trigger SAS code to run as soon as the data arrives and can give the caller a response. Message queues are another option. Note that it might be necessary to have a transactional database (such as Postgres) to support making updates to data while applications are using it.
- Data updates can be made at defined intervals, such as daily or monthly. The benefit of this approach is that it allows more complex algorithms to process all of the data as opposed to just the part that updated.

The SAS programming language is a fantastic tool for processing periodic updates of data. There are several options for getting SAS code to run periodically:

- SAS provides the SAS Schedule Manager as part of SAS® Management Console:



- Systems based on Microsoft Windows have the Windows Task Scheduler:



- Systems based on UNIX (such as Linux) provide the cron command, which allows users to run a command at any time.

Both Windows and UNIX provide the “at” command, but it is useful for running a job only once, so it is generally not used for periodic updates.

When productionalizing a process, there are many questions to consider:

- What happens if data updates are not received before the scheduled time?
- Who gets notified if there’s an error?
- What ID does the job run as?
- What needs to happen on weekends and holidays?
- Is the project using Type 2 Slowly Changing Dimensions (SCDs)?
- What if the processing job takes longer than 24 hours and needs to run daily?
- Can processes handle skipping a day?

- How are real-time updates processed (and do your users need real-time updates)?

A key tenet to making better decisions is to continuously update the data that information is derived from.

CONCLUSION

Making better decisions can save and improve lives, save and improve equipment, and save and make money. Helping people achieve great outcomes by helping them make better decisions is very rewarding.

There is no tenet that is more important than any other. They all have an important role in helping people get the best outcomes. It is not necessary to immediately follow all six tenets at the same time. Most successful projects start by following a small subset and organically growing into following more tenets as time goes on. Projects typically follow the order of tenets as they are presented here, but this is not necessary.

SAS provides many more products that have not been reviewed here that help fulfill the six tenets and help people make better decisions.

While the focus has been on following the six tenets in a data system, these six tenets can be applied to any decision. Better decisions can be applied to a wide variety of situations: personal health and nutrition, personal finance, organizational goals, government programs, and many more. Of course, the best decision you can make is to use “The Six Tenets of a Better Decision” to guide and improve your decision-making processes.

REFERENCES

Summers, E.; J. Langston; R. Allison; and J. Cowley. “Using SAS/GRAPH® to Create Visualizations That Also Support Tactile and Auditory Interaction.” *Proceedings of the SAS Global Forum 2012 Conference*. Cary, NC: SAS Institute Inc. Available https://support.sas.com/resources/papers/proceedings12/279-2012.pdf.

ACKNOWLEDGMENTS

The author would like to acknowledge the fantastic team at SAS Solutions OnDemand for the wonderful job they do helping people make better decisions and for teaching the author what that means.

CONTACT INFORMATION

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

Dan Jahn
100 SAS Campus Drive
Cary, NC 27513
SAS Institute Inc.
Dan.Jahn@sas.com
http://www.sas.com