Paper 4120-2016

Prescriptive Analytics – Providing the Instruction to Do What's Right

Tho Nguyen, Teradata Corporation Fiona McNeill, SAS Institute Inc.

ABSTRACT

Automation of everyday activities holds the promise of consistency, accuracy and relevancy. When applied to business operations, the additional benefits of governance, adaptability and risk avoidance are realized. Prescriptive analytic empowers both systems and front-line workers to take the desired company action – each and every time. And with data streaming from transactional systems, from IoT, and any other source – doing the right thing with exceptional processing speed embodies the responsive necessity that customers depend on. This talk will describe how SAS® and Teradata® are enabling prescriptive analytics – in current business environments and in the emerging IoT.

INTRODUCTION

Being an analytically-driven organization means basing decisions and actions on data, rather than gut instinct. As more organizations recognize the competitive advantages of using analytics, the impact can wane as competitors build this same capability. To cross this innovation chasm and sustain the competitive advances that come from analytical adoption, organizations continually test and expand data sources, improve algorithms and evolve the application of analytics to every day activity.

Predictive algorithms describe a specific scenario, and using historic knowledge increase awareness of what comes next. But knowing what is most likely to happen, and what needs to be done about it are two different things. That's where prescriptive analytics comes in. Prescriptive analytics answers the question of what to do, providing decision option(s) even based on predicted future scenarios.

Seldom (if ever) do events happen in isolation. It's through their interconnections that we develop the detailed understanding of what needs to be done to change future trajectories. The richness of this understanding, in turn, also determines the usefulness of the predictive models (Pinheiro & McNeill, 2014). Just as the best medicine is prescribed based on thorough examination of patient history, existing symptoms and alike – so are the best prescriptive actions, founded in well understood scenario context. And just as some medicines can react with one another - with one medicine not be as effective when it's in the presence of another, so can decisions and corresponding actions taken from analytics – which in turn can impact the outcome of future scenarios.

As you'd expect, under different scenarios – you'd have different predictions. When conditions change, the associated prediction for that same data event can also change. When you apply one treatment, you affect another, changing the scenario. Actions that are taken not only create a new basis for historical context, but also create new data that may not have been considered by the original model specification. In fact, the point of building predictive models is to understand future conditions in order to change them. Once you modify the conditions and associated event behavior, you change the nature of the data. As a result, models tend to degrade over time, requiring updates to ensure accuracy to the current data, scenario, and new predicted future context.

Well-understood scenarios are fed by data. The more data you have to draw from to examine dependencies and relationships that impact the event being predicted, the better the prediction will likely be. This is where the value of big data comes in... as big data is more data with finer detail, and greater context richness. Big data offers details not historically available that explain the conditions under which events happen, or in other words, the context of events, activities and behaviors. Big data analytics allows us, like never before, to assess context – from a variety of data, and in detail. And when that big data is also fast data (on the order of thousands of events per second), it's a stream of events. When we bridge big data analytics with event streams, as generated in the IoT - we have the power to write more

timely and relevant business prescriptions that are much harder for competitors to mimic.

BUILDING PRESCRIPTIVE ANALYTICS

Prescriptive analytics define the instructions for actions based on both analytic models and the business rules that trigger the models. Combined by Boolean logic, SAS[®] Decision Manager provides an intuitive interface to build decision logic, associating the models and the business rules with the appropriate conditions, as illustrated in Figure 1.

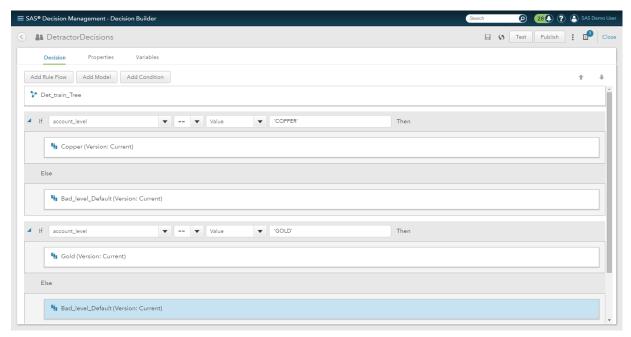


Figure 1. Operational decisions are built by combining business rules (e.g. account_level = "COPPER") with analytical models (e.g. Bad_level_Default) using conditional logic in SAS Decision Manager.

The instruction, as defined SAS Decision Manager's decision logic, encapsulates the conditions under which a particular model is valid – and when it should trigger to deliver results. Scoring data is then reserved for when the appropriate conditions are met, i.e. specific to the model design scenario, avoiding unnecessary data processing.

Typically, business analysts are the decision designers. They are often tasked with working through the logic of what actions need to be taken under different operational scenarios – whether they are product related decisions, customer actions, service requirements or other types of day-to-day business activities. These analysts draw upon the work of others, namely those analytical experts and data scientists who have built the models, as well as reaching into data, like that from Teradata[®] Unified Data Architecture™, which has been vetted and validated by IT.

Building decisions therefore requires the foundation of models that have already been developed, tested and validated using applications like SAS® Analytics for Teradata®. Decisions also require predetermined business rules to be defined to the system. Management of both business rules and analytical models is necessary, particularly given the expanse of users who often benefit from a formalized decision management environment.

BUSINESS RULE MANAGEMENT

Business analysts themselves may have access to all the governing policies, regulatory rules, constraints, best practices and other and business logic necessary to define business rules. More often than not, however, the business knowledge that's needed is retained across different divisions of the organization – like compliance, finance, sales, marketing and others. Thus, the need to have a centralized and well-manage environment for defining business rules, business logic and terminology helps to eliminate debates between different divisions of the organization. It also promotes consistency in the use and application of business rules to operations.

Within SAS Decision Manager, SAS[®] Business Rules Manager provides a centralized and managed repository for rules. Individual rules are joined together using a wizard, which define the specific scenario conditions as rule flows (as illustrated in Figure 2). Rules can be defined, tested, validated against data, and even discovered using analytic methods - all from within the same environment. When rule flows are published for execution in operations, the published rule flow is automatically locked down – to secure it from additional testing and modification. Authorizations and defined workflows ensure that changes are documented, approved and authorized by the appropriate personnel.

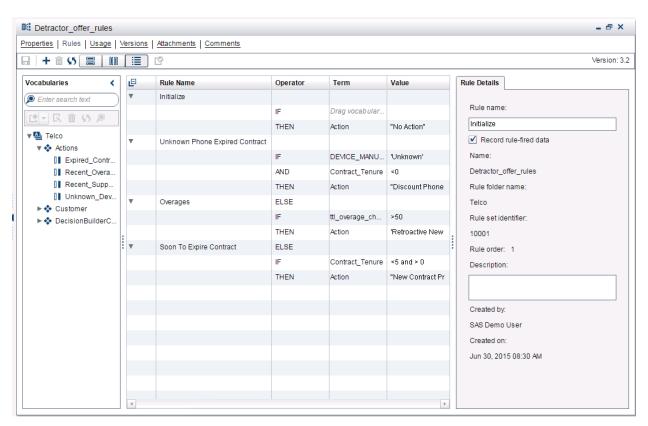


Figure 2: Wizard edit environment for creating, editing and managing business rule flows.

The collection of terms used to build rules is foundational to the common language that communicates the objectives and responsibilities of the business, appropriately described as a vocabulary. You can import pre-existing vocabularies (from .CSV files), edit them, reuse ones extracted from physical tables and share vocabularies across rule sets. SAS Business Rules Manager allows multiple and authorized users to contribute to rule definition, facilitates change management control, retains audit details, empowers validation by subject matter experts, and governs rule elements. When business rules are designed in this environment, they are safe from the risk of undocumented tribal knowledge and become a corporate asset.

ANALYTICAL MODEL MANAGEMENT

Just as business rules are the domain of experts who understand the business, analytical models are the domain of data scientists, statisticians and data miners alike. SAS Decision Manager includes SAS[®] Model Manger, which manages the inventory of models developed in in SAS[®] Factory Miner, SAS/STAT[®], SAS/ETS[®], SAS[®] Enterprise Miner™, PMML, generic R models, code snippets from other code bases¹, as well as from SAS[®] High-Performance Data Mining. Having forecasts, predictions and other models registered as a comprehensive collection (as shown in Figure 3) allows organizations to monitor for signs of degradation as scenario context changes, manage versioning, authorship, workflow, usage tracing, and provides detailed visibility into production quality.

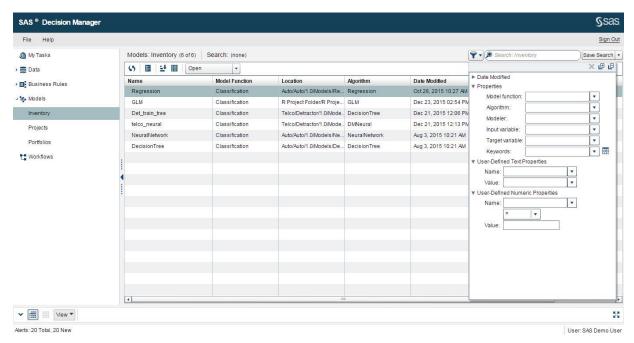


Figure 3: Collections of analytical algorithms are centralized in one, governed environment.

Business analysts, who are focused on creating complete decisions, select the appropriate model as designated by the analytic expert. Taking the guesswork out of which model is most appropriate to a particular scenario and streamlining the often tedious tasks of understanding model definitions and data input needs. The business analyst readily has the business context of the model, explicitly defined and in a recognizable, intuitive format. Building complete decision flows therefore becomes an exercise of defining the rule flows in conjunction with the prescribed model – associating them together by the appropriate conditional logic – all from the same, simplified interface (as was illustrated in Figure 1). Moreover, the logic used, definitions and ownership of each element of the decision flow is retained – so that when it comes to deploying models into production, IT has a complete perspective of who, what, why and how these decisions are defined, the testing done and how to apply to business operations for prescriptive actions.

_

¹ Other code bases, such as C, C++, Java, Python, etc.

VALUE OF PRESCRIPTIVE ANALYTICS

Prescriptive analytics provides the instruction of what to do – and as importantly – what not to do when analytical models are deployed into production environments. Defined as decisions, they are applied to scenarios where there are too many options, variables, constraints and data for a person to evaluate without assistance from technology. These prescriptive decisions are presented to the front-line worker – providing the answer they seek, and accounting for the detailed aspects of the scenario that they find themselves in. For example, call center personnel often rely on prescriptive analytics to know the appropriate options, amount, and under what conditions, a prospective customer can be extended varying levels of credit.

Prescriptive analytics also provides organizations with the ability to automate actions, based on these codified decisions. Every organization has simple, day-to-day decisions that occur hundreds to thousands of times (or more), and which don't have to require human intervention. For example, the identification and placement of a targeted advertisement based on a web shopper's session activity is popular in the retail industry. In such cases, prescriptive analytics are used to ingest, define and take the optimal action (for example, place the most relevant ad) based on scenario conditions (in our example, what has been viewed and clicked on during the session). What is optimal, for the purposes of this paper, is defined as an action that best meets the business rule definitions and associated predicted likelihoods. What is optimal can also refer to a mathematically optimized solution, as Duling (2015) has previously described.

Scoring data with a model typically involves IT. Sent in an email, or some of other notification, IT is presented with an equation and the data inputs needed. What is often very lacking is the business rationale, context and a translation of terminology into IT terms. As such, IT will ask all the necessary questions, often recode the model – run tests and validate output, and then, after applying any specific business policies, and/or regulatory rules – will put the model into 'production', aka operationalize the model so it can generate results.

While in some organizations these steps may not all be done by IT, they still happen. As illustrated in Figure 4, each step – even after the model is developed - adds time to implementing the model, and cashing in on the business benefits. In many organizations the latency associated with model deployment to business action is weeks, if not months. As a result, by the time a model is ready to generate results in a production context, it's often too late – and either the opportunity to impact is gone or conditions have changed to the point where the model is no longer relevant.

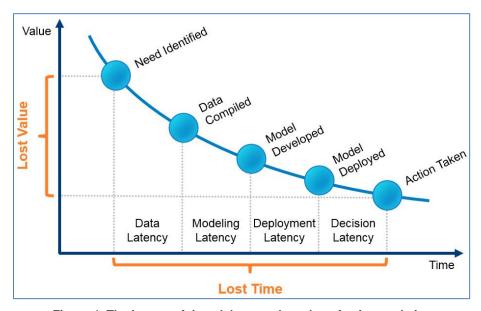


Figure 4: The impact of time delays on the value of using analytics

Prescriptive analytics defined using SAS Decision Manager reduces this latency. Streamlining the time from when a model is developed to when actions are taken. Furthermore, the context of the model is explicit, defined by the business rules - to the point that impact assessments across any point of the decision flow is transparent (as illustrated in Figure 5). And because of this explicit decision definition, changes and adjustments to new models, rules, conditions, data or combinations of any of these dynamics are readily done – tracked as part of version control and documented for the purview of auditors and alike. Analytical model deployment and usage becomes part of a governed, managed environment, reducing the risk associated with incorrect definitions, poor market timing and regulatory non-compliance.

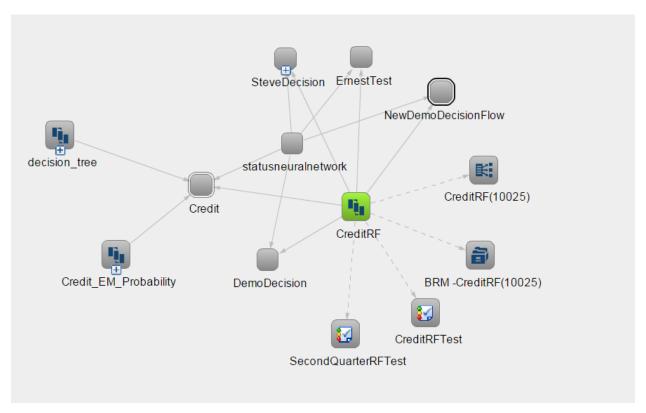


Figure 5: Lineage across decisions can be examined as part of impact assessments

Prescriptive analytics have the benefit of automating instructions and best suggested options that are acted upon by a person. Prescriptive analytics is also be used to directly automate actions, for more mundane tasks, doing so consistently and accurately. In both cases, relevancy to the current scenario is assured in this managed environment and is the product of the vetted, tested and detailed decision flow (as was illustrated in Figure 1). As data volume, variety and velocity are only set to increase, and as technology continues to develop to process more data, faster – the trend to automating actions taken from analytics will correspondingly rise.

The business need to automate prescriptive analytics stems from companies that demand real-time responses from data-driven decisions. It is obvious that every company will increasing become inundated with data and that data needs to be analyzed. The reality is that organizations simply don't have enough people to analyze all the data – even if they could comprehend all the scenario details and volumes, to make all decisions in a timely manner. Prescriptive analytics – defined in SAS Decision Manager have the benefit of being:

- Relevant, consist and accurate decisions
- Easily automated for human instructions and downstream application/system actions
- Explicit of the business context
- Tested, vetted and documented decisions
- Adjustable to changing scenarios
- Timely deployed actions
- Governed in a single environment, providing an unequivocal source of truth
- Assets, by encapsulating intellectual property and managing lifecycle degradation.

OPERATIONALIZING PRESCRIPTIVE ANALYTICS IN TERADATA

SAS Decision Manager has integrated with Teradata to further extend the benefits of prescriptive analytics – by moving analytics deployment to the data, eliminating the burden on network resources and reducing the latency of time to action. By applying prescriptive analytics to where the data reside, the process is significantly streamlined - because data movement and redundancy is eliminated. In addition, this improves data integrity since there is no copying of data or moving data to a different, silo server.

HOW IT WORKS

SAS and Teradata are well integrated to deliver complete, data-driven decisions. The Teradata database can be leveraged to handle the heavy processing of data analytics. Teradata offers a powerful and scalable architecture to enable massively parallelize processing (MPP). This MPP architecture is a "shared nothing" environment and can disseminate large queries across nodes for simultaneous processing. It is capable of high data consumption rates through parallelized data movement which completes any task in a fraction of the time. The end-to-end process can be executed inside the Teradata platform to improve performance, economics and governance, as illustrated in Figure 6.

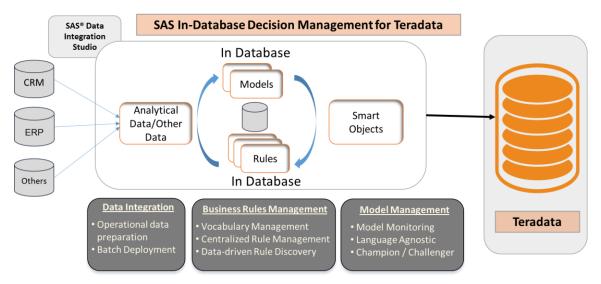


Figure 6: End-to-end decision processing with SAS and Teradata

Complete decisions, which include data definitions, business rules and analytical models² are recognized within SAS[®] Data Integration Studio. Treated as a single decision flow, SAS Data Integration Studio generates SAS DS2 code which can be run within Teradata to inherently leverage the highly scalable environment for processing big data. This is enabled by an embeddable processing technology, the SAS Threaded Kernel (TK) within the Teradata platform. This embedded process generates work using units of parallelism scheduled on each AMP of the Teradata platform. Teradata's workload manager manages the SAS embedded process as a standard Teradata workload.

Analytic models can also be built in-database using SAS® Analytics Accelerator. The SAS® Analytics Accelerator for Teradata contains specialized vendor defined functions for Teradata that enable indatabase processing for a collection of modeling and data mining algorithms. For model building, the SAS Scoring Accelerator for Teradata transforms models created by multiple SAS/STAT or Enterprise Miner for scoring inside the database using the SAS embedded process technology.

Decision that are deployed, and "published" to Teradata directly from SAS Data Integration Studio, as SAS macros, or if only model scoring is desired, they can be published using SAS Model Manager. SAS Decision Manager includes both SAS Data Integration Studio as well as SAS Model Manager, providing options which optimize analytically-based processing in-database with Teradata (as shown in Figure 7). The metadata about models, rules and logic are all encapsulated within decisions – helping organize production deployment.

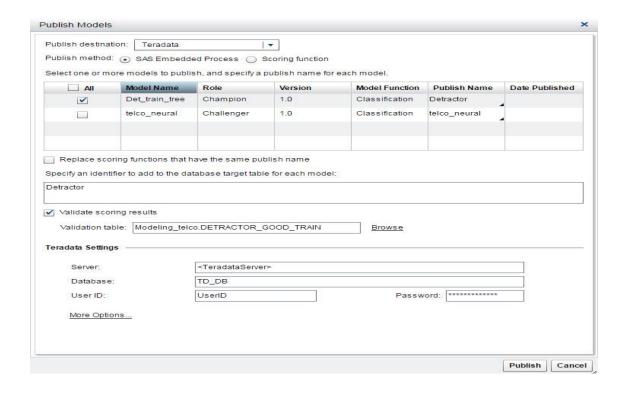


Figure 7: Publish models using SAS Model Manager (included with SAS Decision Manager) and Teradata

In some organizations, prescriptive decisions are deployed into operational data streams. There may also be instances that only business rules, without analytic models, are needed for the appropriate action. For example, internal organizational accounting often requires a distribution of revenue (aka revenue

_

² The SAS[®] Code Accelerator for Teradata and SAS[®] Analytics Accelerator for Teradata push SAS executable code to process directly inside the Teradata data warehouse.

attribution) across business divisions and functions – based on corporate policies or governance measures. Business rules defined within SAS Business Rules Manager can be pushed down and directly executed inside the database without any recoding or redefinition³. For the in-database business rule execution inside Teradata, the processing tasks are further streamlined and are fully scalable without data replication. It also has the advantage of being highly amenable to commonly required changes in business rule definitions – with organizational and product changes, acquisitions, mergers and business policy dynamics.

DATA EXPLORATION

Once data sources are gathered in Teradata, you can begin to explore it using your preferred data exploration tool, like SAS® Visual Analytics (which is also enabled on the Teradata Appliance for SAS). Data exploration is a process that examines and explores data, often discovering or extracting new knowledge. Typically performed by a business analyst, exploring what the data looks like, the scenario at hand, and what variables are in the data set – evaluates the relationships and patterns necessary to understand decision conditions.

This initial exploration of the data helps explain common inquiries and is a productive way to become more familiar and intimate with the data that defines the scenario. One of the best practices is to explore all your data directly in the database, so data is well understood before identifying the key factors for conditional logic and rule definitions, while eliminating redundancy and removing irrelevant data. This same exploration capability is also a powerful and flexible way to monitor business rule execution and retrospectively decision actions – as a dashboard or reports. And it's not just for the business analyst. The ability to quickly extract knowledge from large complex data sets also provides the data scientist, statistician and data miner alike with this same advantage of dynamically exploring data as part of the model development process.

Prescriptive analytics require the right process, skilled personnel and scalable technology. With SAS and Teradata, prescriptive analytics is streamlined, effective and efficient – from the perspectives of both IT and the business. These integrated technologies deliver data-driven decision options and even automated actions, helping organizations take advantage of future opportunities and alleviating potential risks each time a decision is made.

LEVERAGING INTERNET OF THINGS (IOT)

The Internet of Things (IoT) can mean different things for many people and works in conjunction with big data. It is a system of physical objects—devices, vehicles, buildings, machines and others — that are embedded with electronics, software, sensors, and network connectivity so that these objects can communicate through the exchanged of data. IoT is, and will continue to generate a lot of data. Data transmitted by objects provides entirely new opportunities to measure, collect and act upon an ever-increasing variety event activity.

If we isolate to just consider sensor data, say in the transportation industry – use cases abound in ways to identify potential equipment defects in planes, trains and automobiles. Going beyond collection of data for exploration, and even analysis - prescriptive analytics will not only uncover patterns in events as they occur, they are used to take automated actions to prevent unnecessary outages and costs. By sending alerts, notifications, updating situational war room dashboards, and even providing instructive action to other objects, the need for real-time actions has never been greater.

Sensor data can be in the form of structured and semi-structured data. Sensor data can be integrated with other data sources, as lookup tables while it's still in motion, and after it's landed in a big data repository/warehouse, And while many organizations are simply streaming sensor data, storing it to be analyzed retrospectively, prescriptive analytics – embedded within sensor (and other event) streams holds the promise of consistent reaction and even proactive intervention. For example, retail sales activity

_

³ The SAS Code Accelerator for Teradata is used to execute SAS Business Rules Manager code in the Teradata data warehouse.

based on marketing campaign effectiveness is often associated with a targeted list of loyal customers. By collecting web clicks in real-time, and past purchases, prescriptive analytics could indicate that they have a high likelihood of purchasing shoes with the pants they are viewing – prompting a pop-up savings coupon for pants.

Use cases leveraging prescriptive analytics in IoT applications abound. Everything from analyzing social media watch by collecting tweets, blogs and posts to determine what consumers are recommending as a service/product to security and surveillance of login sessions and data access for data security breaches – and all else in between.

CONCLUSION

In the eyes of customers, in both business-to-business and business-to consumer industries, purchase choices can be summarized as being dependent on product quality, service and support excellence, and the ability to appropriately fulfill the purchase need. As such, ensuring product health, the responsiveness of fulfillment, and understanding the full context of the purchase decision is paramount to being the selected candidate. For day-to-day decisions, prescriptive analytics fulfills that need - giving organizations the ability to accurately decipher the scenario context and to take the appropriate action in a manner that's consistent and relevant. With SAS® In-Database Decision Management for Teradata®, you can:

- Be more responsive, proactive and reliant on data-driven operational decisions for new opportunities.
- Improve performance and minimize time previously spent moving or duplicating data and code between systems.
- Increase security and compliance of data in one integrated, highly governed environment.

Taking prescriptive analytics to the data and running in-database extends the benefits of relevant, timely instructions and actions without having to move data. Model and business rule deployment – as complete, documented and vetted decisions – become part of job processing, for even the biggest of big data. With SAS and Teradata, the integrated portfolio of solutions enables you to explore all options, determine the appropriate approach, execute the action and evaluate/improve the business decision.

REFERENCES

C. Pinheiro, F. McNeill 2014 *Heuristics in Analytics: A Practical Perspective of What Influences Our Analytical World.* NJ: John Wiley and Sons.

Duling, D. 2015 "Make Better Decisions with Optimization" SAS Global Forum 2015 Proceedings, Paper SAS1785-2015, Dallas, TX. Available at:

http://support.sas.com/resources/papers/proceedings15/SAS1785-2015.pdf

SAS and Teradata Partnership

www.teradata.com/sas

SAS and Teradata In-Database Decision Management for Teradata

http://www.teradata.com/partners/SAS/SAS-In-Database-Decision-Management-for-Teradata-Advantage-Program/

RECOMMENDED READING

IIA Research Brief, Prescriptive Analytics: Just What the Doctor Ordered, 2014

http://epictechpage.com/sms/sas/wp-content/uploads/2015/01/iia-prescriptive-analytics-Just-What-Dr-Ordered.pdf

CONTACT INFORMATION

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

Tho Nguyen Teradata tho.nguyen@teradata.com

Fiona McNeill SAS fiona.mcneill@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.