

Using Enterprise Miner™ to Create Model Documentation And/or Reproducible Research

Rex Pruitt, SAS Institute, Indian Trail, NC

ABSTRACT

Businesses need to automate the documentation of their models and integrate the resulting documentation into a Model Risk Management process. Most model documentation processes involve interactions with too many applications in order to produce the documentation necessary for hundreds of models in production. Limited model documentation has resulted in organizations being placed on written agreement with the Federal Reserve and/or Consumer Financial Protection Bureau (CFPB). This has led to billions of dollars in fines and/or penalties.

SAS® Software provides a nice solution to this model risk governance issue. Specifically, Enterprise Miner™ will support all of the necessary functionality to support the required model documentation to satisfy governance requirements.

The key features that support Model Documentation are...

REPORTER NODE

- Uses SAS Output Delivery System to create a PDF or RTF of a process flow.
- Helps document the analysis process and facilitate results sharing.
- Document can be saved and included in SAS Enterprise Miner results packages.
- Includes image of the process flow diagram.
- User-defined notes entry.

REPRODUCIBLE RESEARCH

- Build more, better models faster.
- Provides XML diagram exchange.
- Reuse diagrams as templates for other projects or users.
- Directly load a specific data mining project or diagram, or choose from a Project Navigator tree that contains the most recent projects or diagrams.

INTRODUCTION

Businesses need to automate the documentation of their models and integrate the resulting documentation into a Model Risk Management process. Most model documentation processes involve interactions with too many applications in order to produce the documentation necessary for hundreds of models in production. Limited model documentation has resulted in organizations being placed on written agreement with the Federal Reserve and/or Consumer Financial Protection Bureau (CFPB). This has led to billions of dollars in fines and/or penalties.

FEDERAL RESERVE AND OCC GUIDANCE

According to the BOARD OF GOVERNORS - DIVISION OF BANKING SUPERVISION AND REGULATION - SR 11-7 - April 4, 2011, for Model Development, Implementation, and Use... "A sound development process includes: a clear statement of purpose to ensure that the model is developed in line with its intended use; sound design, theory, and logic underlying the model; robust model methodologies and processing components; rigorous assessment of data quality and relevance; and **appropriate documentation**."

Regarding, Governance, Policies, and Controls... "Strong governance also includes documentation of model development and validation that is sufficiently detailed to allow parties unfamiliar with a model to understand how the model operates, as well as its limitations and key assumptions."

Without **adequate documentation**, model risk assessment and management will be ineffective.

<http://www.federalreserve.gov/bankinfo/srletters/sr1107.pdf> (ref 1)

Using Enterprise Miner™ to Create Model Documentation And/or Reproducible Research Rex Pruitt, SAS Institute, Indian Trail, NC

BUSINESS IMPACT

So, why is this so important? Is it because the government says it must be done? Well, I can think of many reasons why a business would want to have robust documentation of its analytical models used for business operations. Here's a list of a few:

- Protection of intellectual property
- Disaster recovery
- Data scientist attrition
- Model reproduction support
- Model maintenance and portability
- Satisfy regulators and their audit requirements
- Answer audit question quicker
- Monitoring of model effectiveness and value
- Measurement of potential model deterioration
- Government fines connected to a breakdown in any of the above
- Damaged reputation resulting from government intervention

Just to emphasize the impact of the last two bullets, here are some examples of regulatory actions that probably could have been avoided if a more robust documentation and monitoring process had been in place:

- [Capital One Hit with Consent Order on Former Check-Cashing Business](#) (*ref 2*)
 - by [Colin Wilhelm](#) - AUG 5, 2015
 - Critical Business Issue (CBI) = Financial Crimes Enforcement Network for information on **anti-money laundering compliance** concerning "certain check casher clients" of its commercial banking business
- [HSBC's Culture Still 'Deficient' Two Years On, Monitor Says](#) (*ref 2*)
 - By [Chris Cumming](#) - April 1, 2015
 - CBI = HSBC Holdings still has serious weaknesses in its anti-laundering controls...Stubborn resistance to reform and inefficient technology have slowed improvement...efforts to comply with the...record \$1.9 billion settlement it reached with federal authorities in 2012.

Even if there were no other reason than to avoid bad publicity or millions/billions of dollars in fines, it seems a good business decision would be to leverage a robust model documentation practice.

Using Enterprise Miner™ to Create Model Documentation And/or Reproducible Research

Rex Pruitt, SAS Institute, Indian Trail, NC

MODEL GOVERNANCE MATURITY

Based on my experience, many organizations are not very mature from a model governance perspective. There are often hundreds or thousands of models that have been developed and deployed into operational decision systems. The effort to document and/or explain these models is an onerous process. Thus, regulators and executives become frustrated leading to unfavorable audit findings and reports that require resolution.

Some organizations do not even have a Model Risk Management process that ensures proper compliance across the modeling lifecycle (Figure 1).

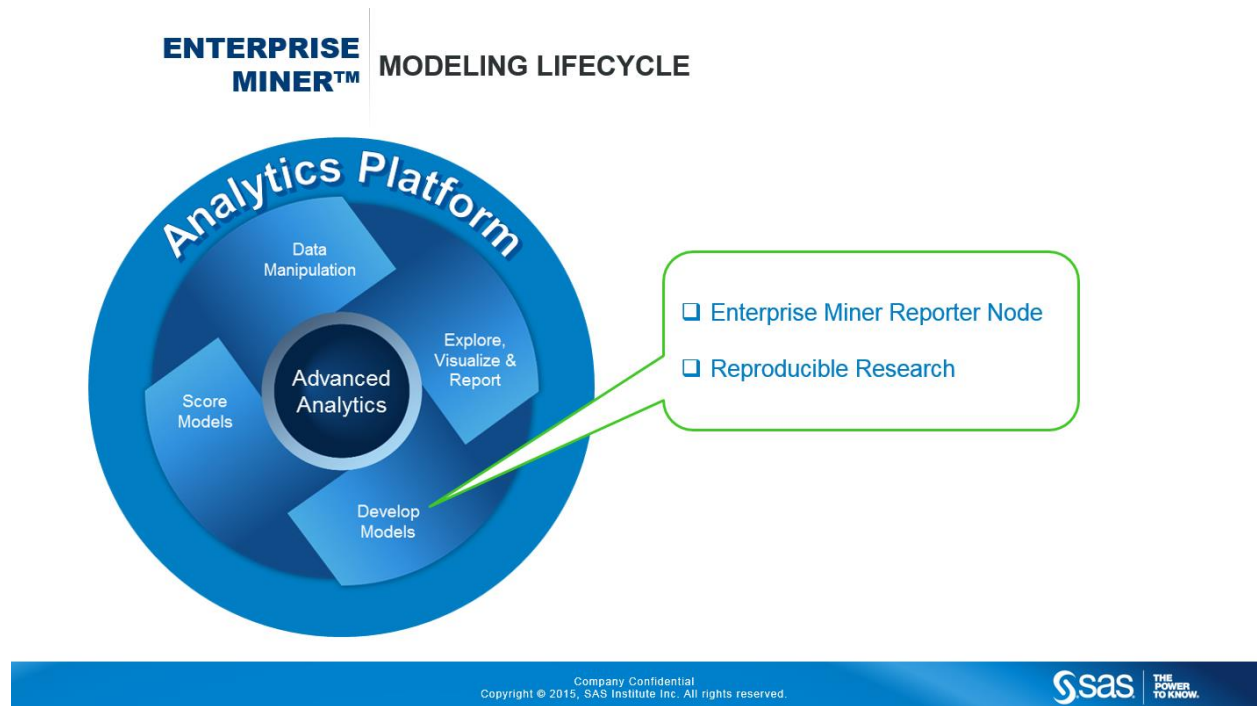


Figure 1. – Modeling Lifecycle with callout of where this functionality fits.

SAS Software provides a nice solution to this model risk governance issue. Specifically, SAS Enterprise Miner has all of the necessary functionality to support the required model documentation for satisfying governance requirements. It will also address the list of “**Business Impact**” items noted earlier.

The following illustration (Figure 2.) shows where SAS Enterprise Miner fits in the overall governance environment as supported by various SAS solutions.

Using Enterprise Miner™ to Create Model Documentation And/or Reproducible Research

Rex Pruitt, SAS Institute, Indian Trail, NC

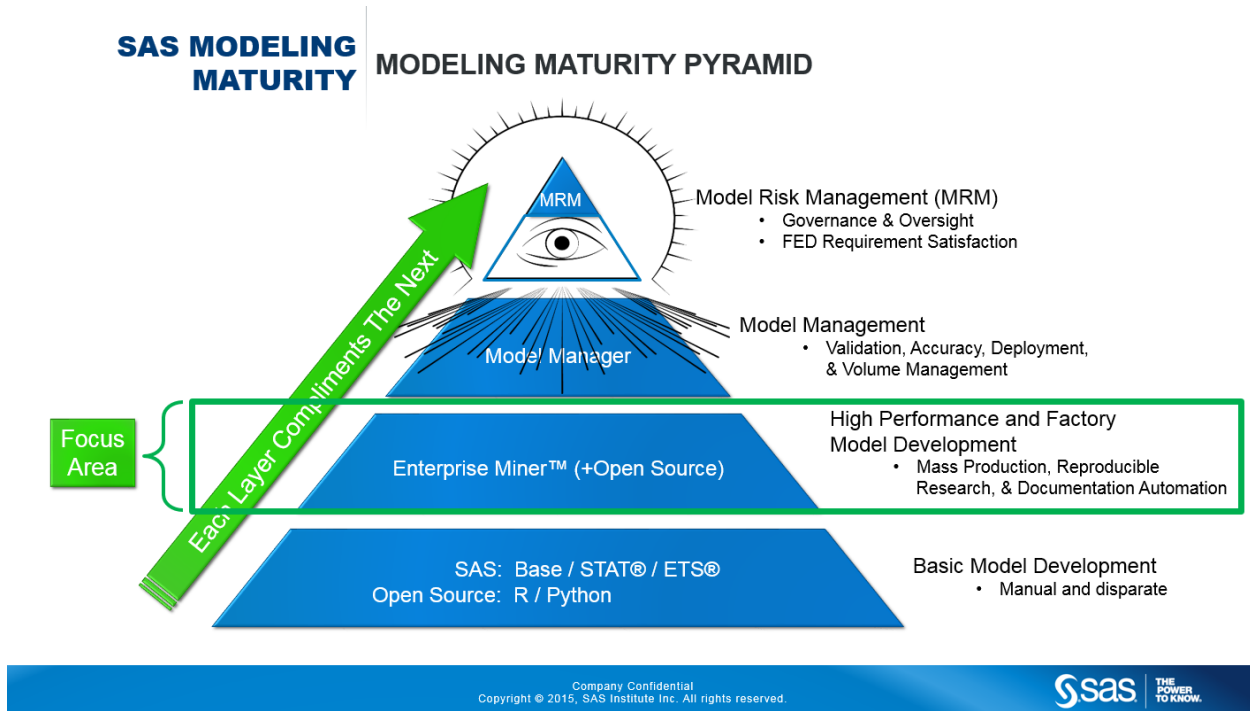


Figure 2. – Illustration of the Modeling Maturity Pyramid to Portray Overall Governance Oversight

DEMONSTRATION - MODEL DOCUMENTATION AND REPRODUCIBLE RESEARCH

The rest of this paper is focused on demonstrating documentation capabilities specifically within SAS Enterprise Miner. Key components portrayed will be centered on “Documentation Automation” and “Reproducible Research.”

Using Enterprise Miner™ to Create Model Documentation And/or Reproducible Research

Rex Pruitt, SAS Institute, Indian Trail, NC

The following two tables highlight the capabilities to be demonstrated.

ENTERPRISE MINER THE KEY FEATURES THAT SUPPORT MODEL DOCUMENTATION

Reporter node (Documentation Automation)	Reproducible Research
<ul style="list-style-type: none">✓ Uses SAS Output Delivery System to create a PDF or RTF of a process flow.✓ Helps document the analysis process and facilitate results sharing.✓ Document can be saved and included in SAS Enterprise Miner results packages.✓ Includes image of the process flow diagram.✓ User-defined notes entry.	<ul style="list-style-type: none">✓ Build more, better models faster.✓ Provides XML diagram exchange.✓ Reuse diagrams as templates for other projects or users.✓ Directly load a specific data mining project or diagram, or choose from a Project Navigator tree that contains the most recent projects or diagrams.

Company Confidential
Copyright © 2015, SAS Institute Inc. All rights reserved.



Figure 3 – SAS Enterprise Miner Capabilities that will be demonstrated

ENTERPRISE MINER DEMONSTRATION – END RESULT

Doc Demo German Credit Reporter Node.pdf
C:\Users\reprui\Documents\MWSUG 2015\Model Documentation Demo\Doc Demo German Credit Reporter Node.pdf

Export Path as SAS Program

Save

Save

Documentation Automation

Reproducible Research

Reproducible Research

Company Confidential
Copyright © 2015, SAS Institute Inc. All rights reserved.



Figure 4 – Demonstration Results of Documentation Automation and Reproducible Research

Using Enterprise Miner™ to Create Model Documentation And/or Reproducible Research

Rex Pruitt, SAS Institute, Indian Trail, NC

ENTERPRISE MINER HERE ARE THE STEPS THAT WILL BE PERFORMED

Automated Documentation:

1. Show simple migration of Base SAS Model to Enterprise Miner
2. Create the “Automated Documentation” using the Reporter Node
3. Save the Model Documentation as PDF and RTF (allows editing)
4. Show the “Notes” items where the modeler can place comments

Reproducible Research:

1. Show saving/exporting a diagram into an XML file
2. Save the model flow as a Base SAS program and execute in DMS
 - The XML file and Base SAS program will now be sharable
3. Import the “Shared” XML diagram into a new EM project



Figure 5 – Live Demonstration Steps Performed to Show How to get to the “End Result” in Figure 3

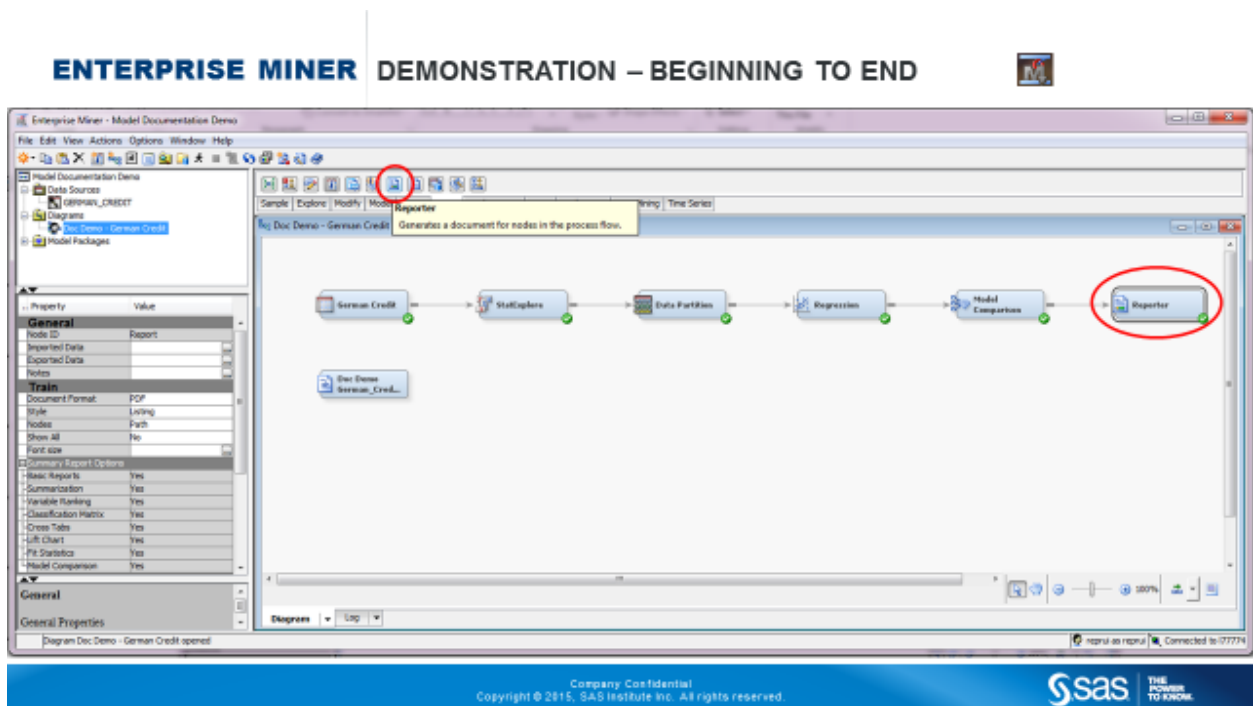


Figure 6 – SAS Enterprise Miner Interface with “Reporter Node” Highlighted

Using Enterprise Miner™ to Create Model Documentation And/or Reproducible Research

Rex Pruitt, SAS Institute, Indian Trail, NC

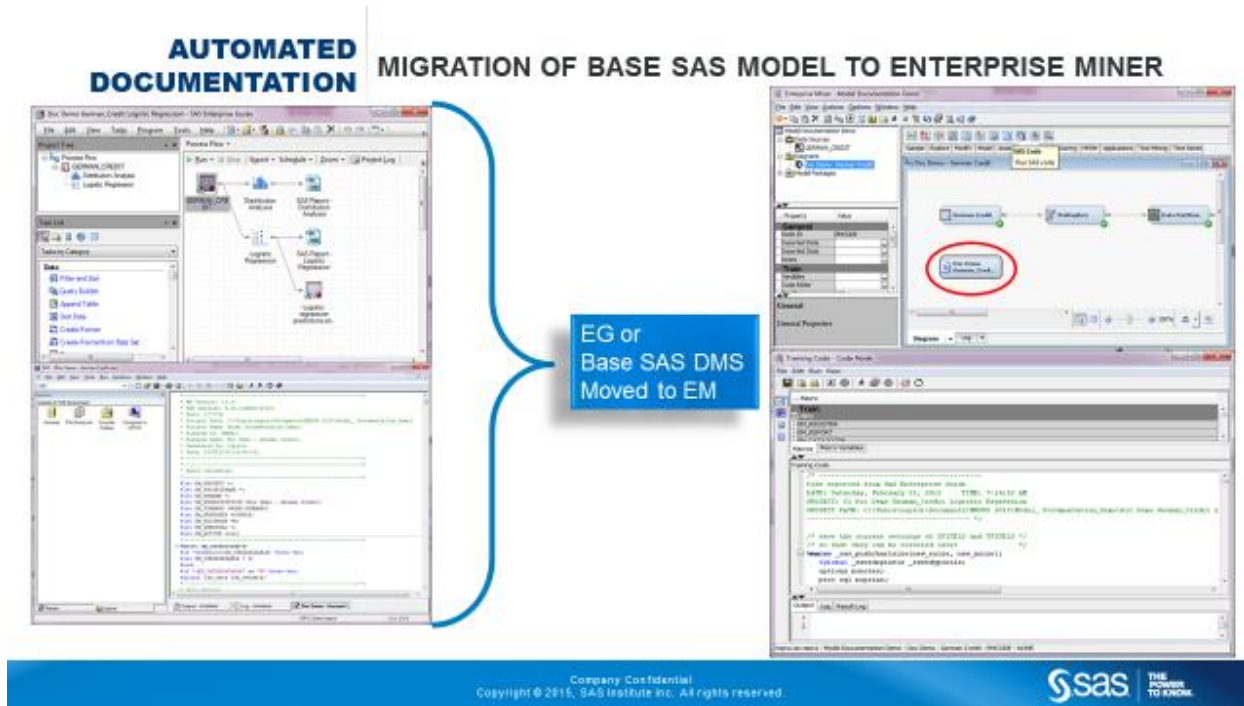


Figure 7 – Base SAS Model Migration to Enterprise Miner for “Reporter Node” Use

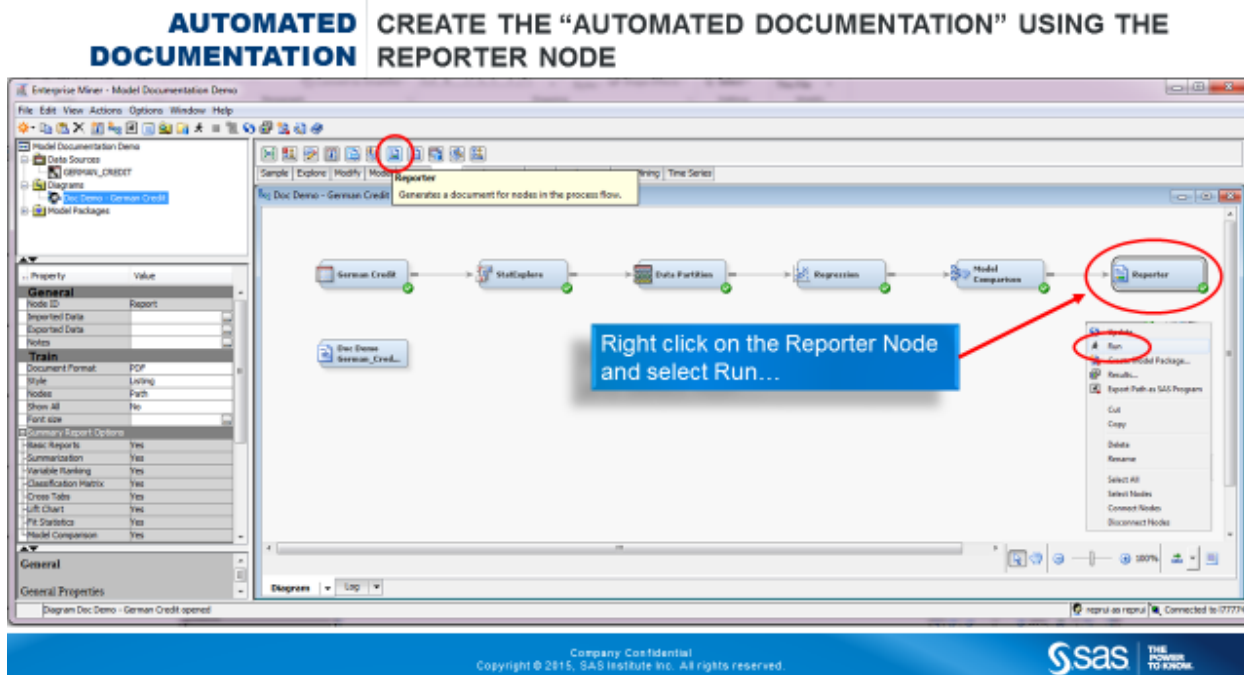


Figure 8 – Execution of the “Reporter Node” on a Migrated Base SAS Process Flow in Enterprise Miner

Using Enterprise Miner™ to Create Model Documentation And/or Reproducible Research

Rex Pruitt, SAS Institute, Indian Trail, NC

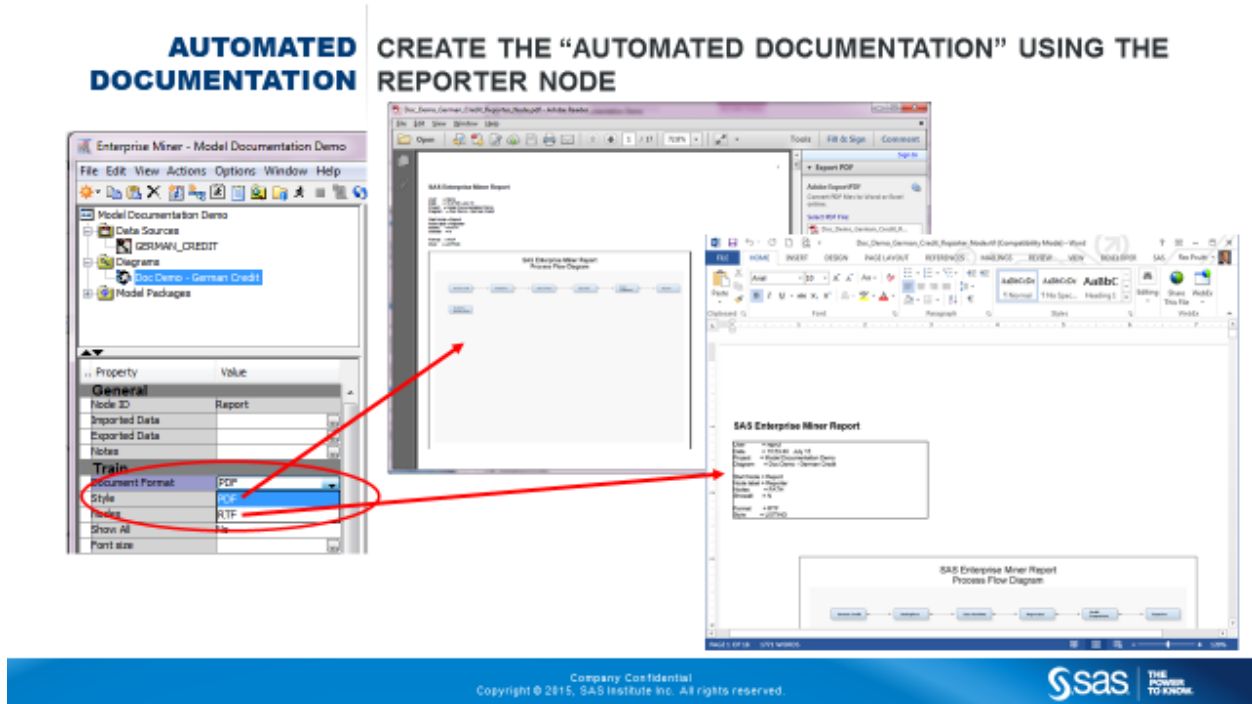


Figure 9 – Adjusting the Style Settings between PDF and RTF in the “Reporter Node” Property Window

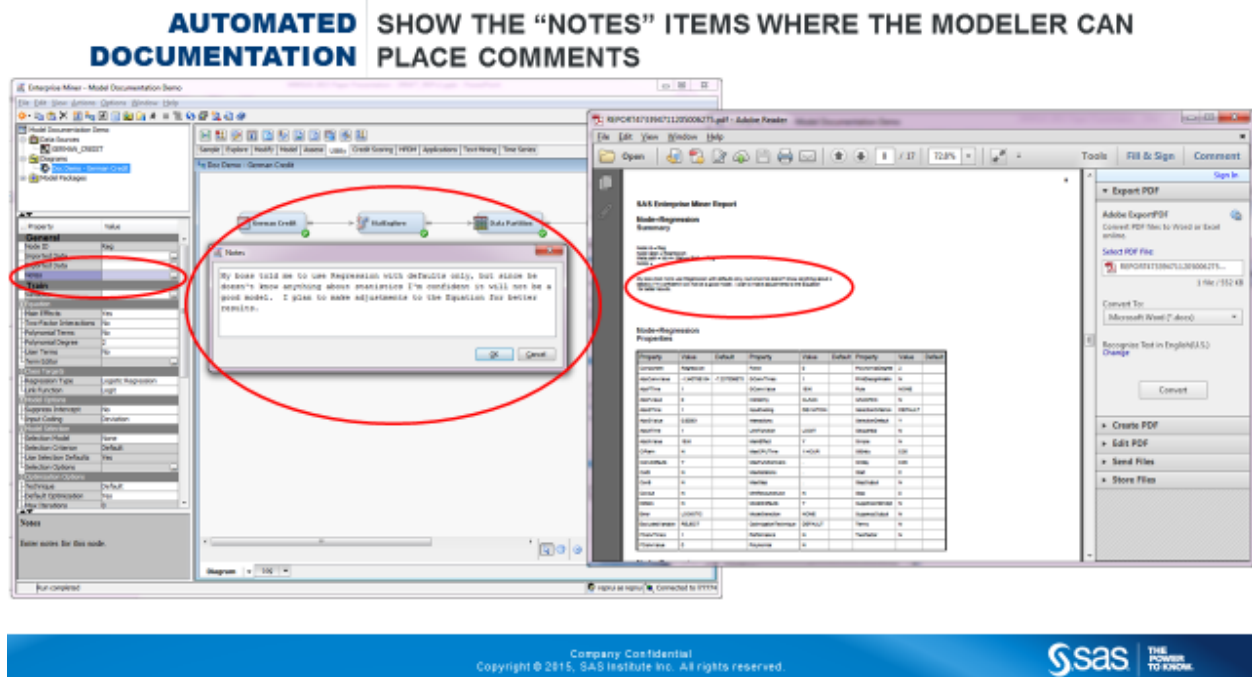


Figure 10 – Adding “Notes” to the PDF or RTF Output to Enrich the Model Documentation

Using Enterprise Miner™ to Create Model Documentation And/or Reproducible Research

Rex Pruitt, SAS Institute, Indian Trail, NC

**REPRODUCIBLE
RESEARCH**

SHOW SAVING/EXPORTING A DIAGRAM INTO AN XML FILE

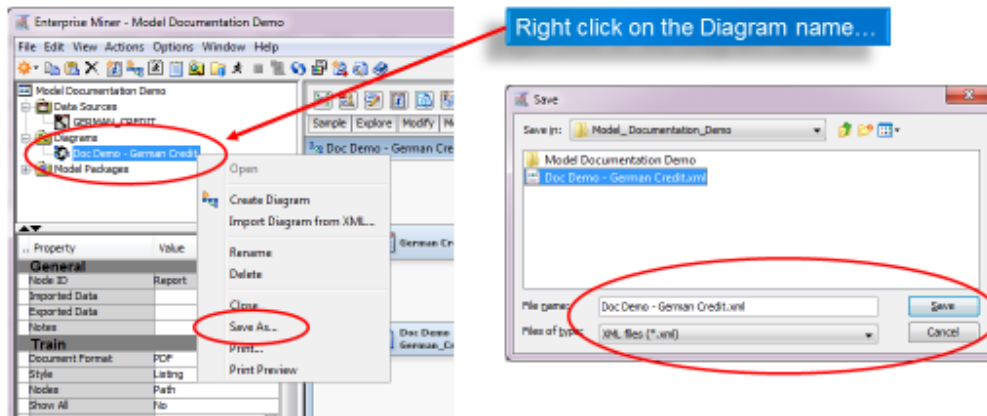


Figure 11 – Creating “Reproducible Research” using the “Save As” XML Option

**REPRODUCIBLE
RESEARCH**

SAVE THE MODEL FLOW AS A BASE SAS PROGRAM AND EXECUTE IN DMS

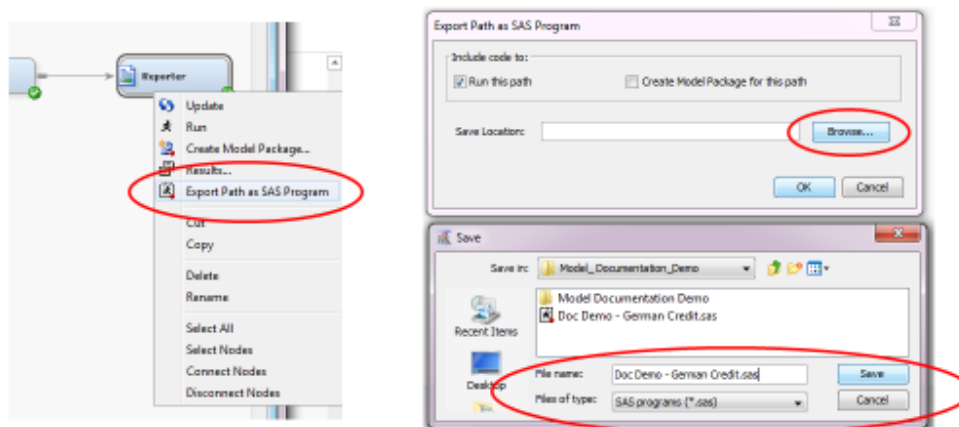


Figure 12 – Creating “Reproducible Research” using the “Export Path as SAS Program” Option

Using Enterprise Miner™ to Create Model Documentation And/or Reproducible Research

Rex Pruitt, SAS Institute, Indian Trail, NC

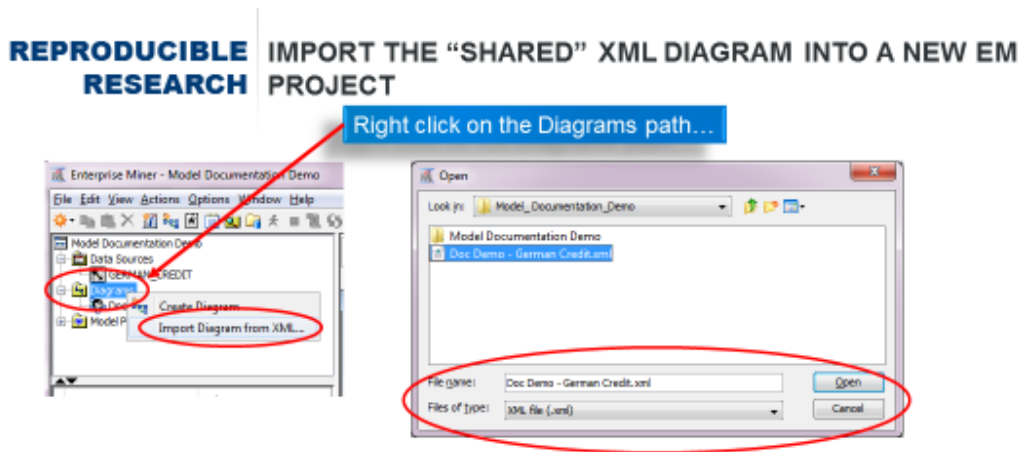


Figure 13 – Sharing the “Reproduced Research” through the “Import Diagram from XML” Option

CONCLUSION

This paper and demonstration should position your organization for the following actions:

1. If you have SAS Enterprise Miner, and you are not leveraging these documentation practices, speak to your management about implementing a representative standard of protocols in your enterprise for properly documenting your models.
2. If you do not have SAS Enterprise Miner, consider having a conversation with your SAS Account Executive for an assessment of whether SAS Enterprise Miner is right for your organization’s needs.
3. Model documentation is a critical component in the success of the enterprise. Whether proactively or reactively, an organization can avoid many pitfalls if they embrace a “Model Documentation Best Practice” using SAS Enterprise Miner for creating model documentation and reproducible research.

REFERENCES

1. BOARD OF GOVERNORS - DIVISION OF BANKING SUPERVISION AND REGULATION - SR 11-7 - April 4, 2011; <http://www.federalreserve.gov/bankinforeg/srletters/sr1107.pdf>
2. American Banker; © 2015 [SourceMedia](http://www.sourcemedia.com). All rights reserved.

**Using Enterprise Miner™ to Create Model Documentation
And/or Reproducible Research
Rex Pruitt, SAS Institute, Indian Trail, NC**

CONTACT INFORMATION

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

Name: Rex Pruitt
Enterprise: SAS Institute
Address: 3006 Sandbox Cir
City, State ZIP: Indian Trail, NC 28079
Work Phone: 605-595-1638
E-mail: trepruitt@hotmail.com
Web: www.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.

BIOGRAPHY:

Rex Pruitt works for SAS Institute as a Senior Systems Engineer. In his current role, Rex is responsible for creatively applying business analytics applications to help solve client business problems and communicate these approaches to different business audiences in the Financial Services Industry. The majority of his career has been focused on the study of corporate data and how it can be turned into a revenue source. Pruitt has been in the analytics profession since 1986 and is extremely proficient with SAS software solutions. Pruitt has spent almost three decades supporting the financial services and insurance industries. While he comes from a humble Midwestern farm background, he has had the opportunity to acquire international business acumen by traveling the world, performing accounting audits and recommending business process improvements using his analytic talents and gifts.