Paper #SAS004-2014

SAS® Predictive Asset Maintenance: Find out why before it's too late!

George S. Habek, M.S., SAS Institute Inc.

ABSTRACT

Are you wondering what is causing valuable machine assets to fail? What could the drivers be, and what is the likelihood of those failures? Do you want to be proactive rather than reactive? These are all areas in which SAS® Predictive Asset Maintenance can help immensely. This solution provides an analytical framework to reduce the amount of unscheduled downtime and optimize maintenance cycles and costs. Key aspects of this paper include the following:

- Key business drivers for and capabilities of SAS Predictive Asset Maintenance
- Detailed analysis of the solution:
  - Data model
  - Explorations
  - Data Selections
  - Path I: Analysis Workbench---maintenance analysis and stability monitoring
  - Path II: Analysis Workbench---JMP®, SAS® Enterprise Guide®, and SAS® Enterprise Miner™
  - Analytical case development using SAS Enterprise Miner, SAS® Model Manager, and SAS® Data Integration Studio
  - Portlet for reports within SAS Predictive Asset Maintenance.

A realistic business example in the oil and gas industry is used in this paper.

BUSINESS SCENARIO

Suppose there is an oil and gas company ("ABC") that includes off-shore countries with specific locations. ABC has water injection systems injecting steam into 85 wells (measured by a concept known as "conformity"). There are two groups of wells, as follows:

- Wells that are over-injected
- Wells that are under-injected.

The goal is to inject the "right" amount of steam into the wells and achieve conformity of 90 percent. The location used in this paper is off the Persian Gulf shore of Kuwait. Figure 1 depicts the problem that ABC is challenged with solving. This figure depicts a snapshot of January – May 2012, which clearly indicates that the wells are out of conformance, with an especially rapid decline towards the end of May. The historical time frame of the data to be analyzed is calendar year 2012.

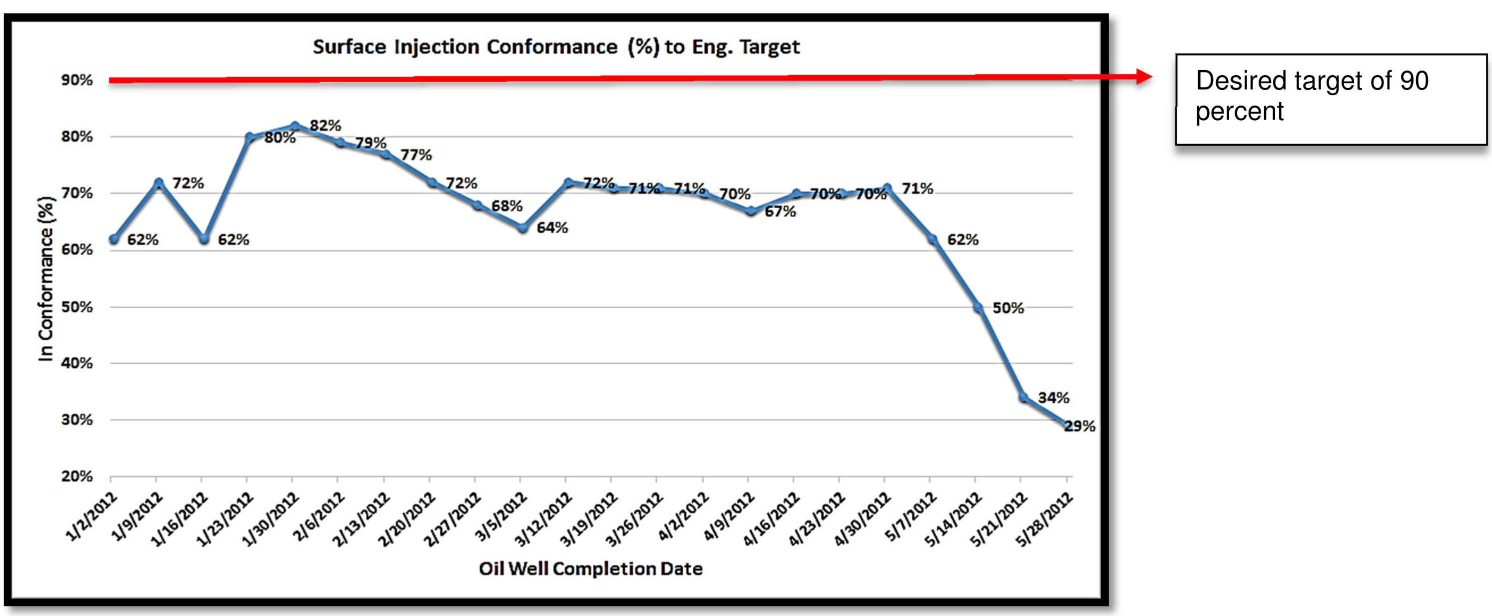

Figure {1}. ABC Business Challenge (BUSINESS SCENARIO)

In order to gain a better understanding of the SAS Predictive Asset Maintenance framework and how it should be applied, Figure 2 depicts the business scenario process flow. One of the key tasks is to accurately define the modeling statement based on the business challenge.

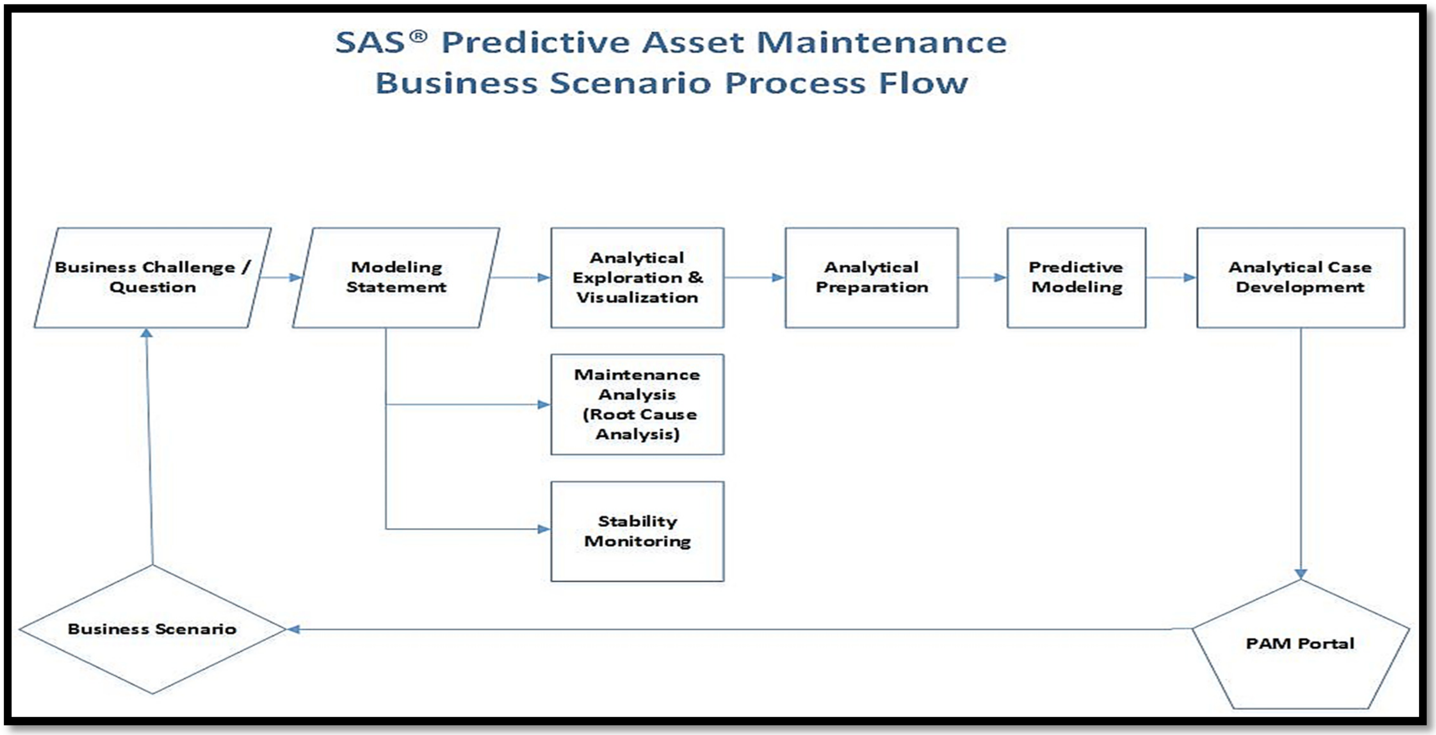

Figure {2}. Business Scenario Process Flow (BUSINESS SCENARIO)

Figure 3 is an illustration of this example for ABC's oil wells in Kuwait.

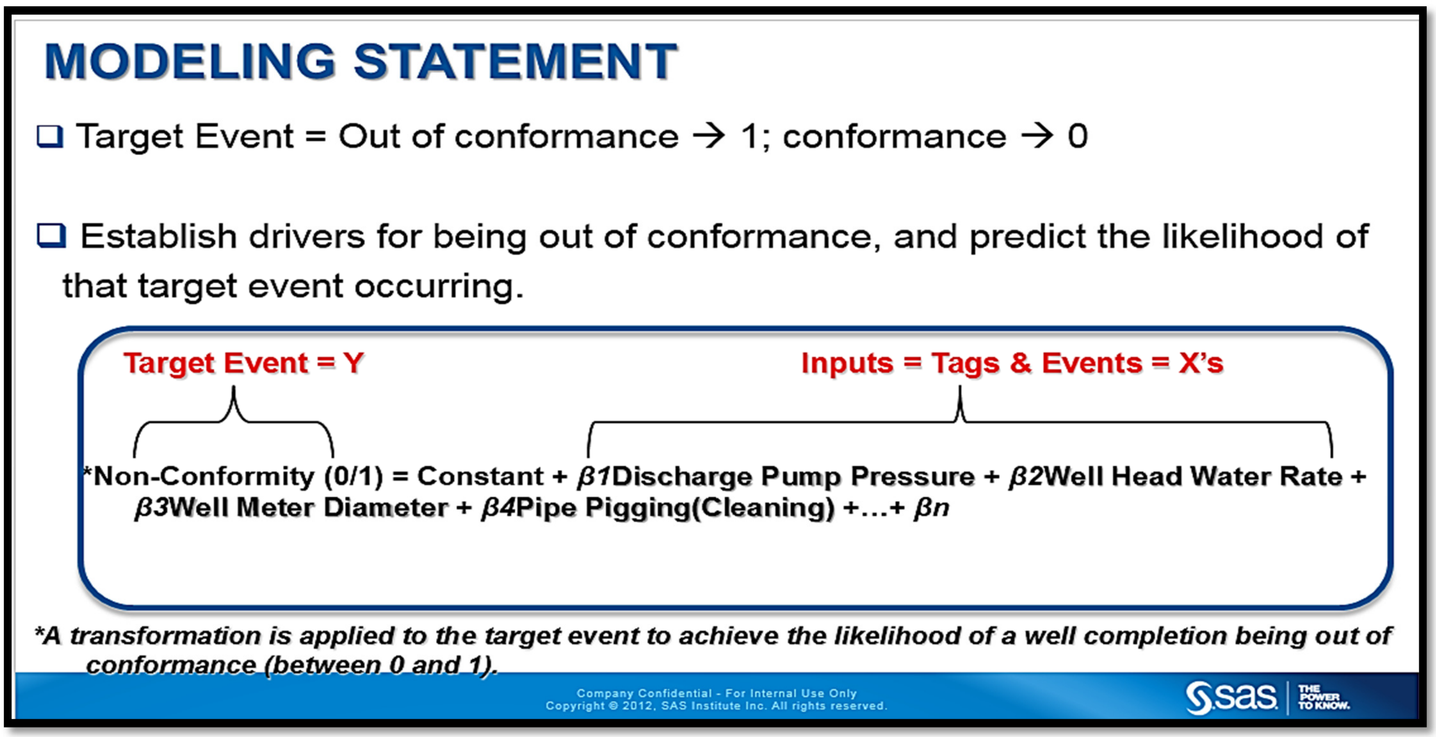

Figure {3}. ABC Modeling Statement (BUSINESS SCENARIO)

DATA

A large amount of data is produced from the oil wells within a given period. Figure 4 shows color-coded information from the asset based on two types of data, as follows:

- Tag information from sensors
- Event information.

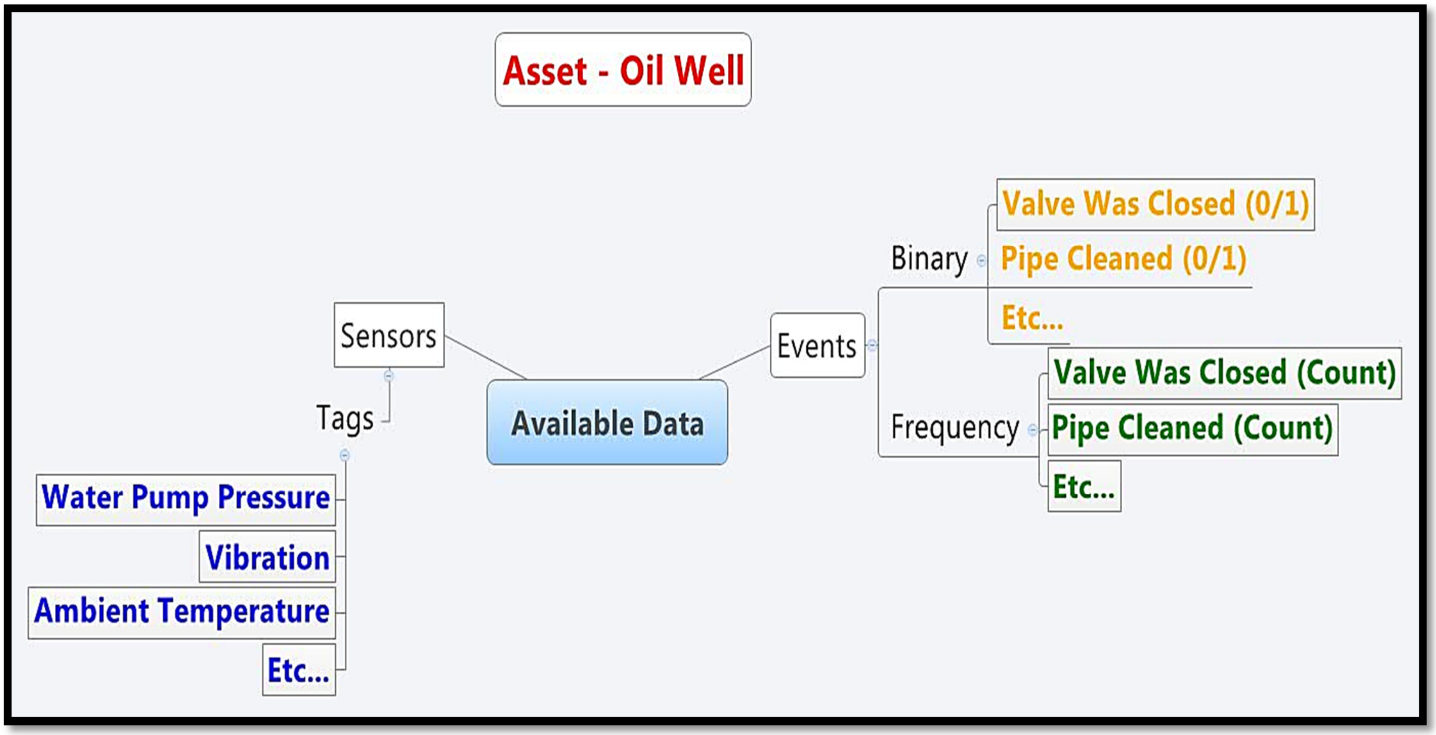

Figure {4}. Tag Information from Sensors and Event Information (DATA)

The colors indicate different types of data, as follows:

- **Blue** represents continuous measurements known as “tags” produced from the sensors
- **Yellow** represents event information in binary form (no = “0;” yes = “1”)
- **Green** represents information in the form of frequency counts for a given event.

There are other types of data that can be used within the SAS Predictive Asset Maintenance framework. Examples include the following:

- Financial
- Operations
- Test cell and reliability
- Operator and compliance
- Inspection
- Technician and engineer
- External.

The minimum information required is depicted in Figure 4.

DATA MODEL

The SAS Predictive Asset Maintenance framework contains a pre-built structure to contain data from desired assets. Organizations have two technology choices for storing solution data in the data model, as follows:

- SAS® Scalable Performance Data Server
- Hadoop.

The starting point for using the data model is data originating from the organization's source system(s). Next, site-specific ETL is applied to populate the data warehouse, which is essentially a staging area. Finally, the data mart is populated using ETL provided with SAS Predictive Asset Maintenance. The data warehouse and data mart are designed to handle big data quickly and easily. Again, there are two areas (tag measurements and events) for population. Figure 5 depicts the data model flow.

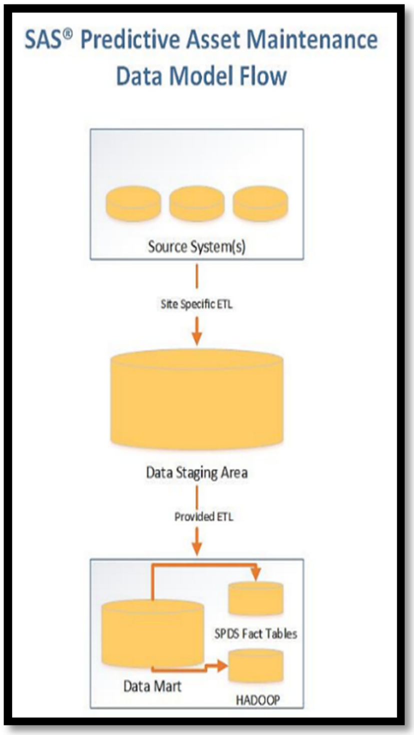

Figure {5}. Data Model Flow (DATA MODEL)

The solution interface consists of three main sections, as follows:

- Explorations
- Data Selections
- Analysis Workbench.

In addition to these three main sections, there is also an "Administration" section. This section of the interface within SAS Predictive Asset Maintenance involves diagnosing certain aspects of the framework, separated into the following components:

- Manage data selections
- Manage maintenance analysis
- Configuration
- Logs
- Manage defaults.

EXPLORATIONS

The "Explorations" section of the interface within SAS Predictive Asset Maintenance allows users to visually explore both tag and event information within specific historical time frames in order to detect anomalies within the data. In this example, the Kuwait region contains some tags and events for review for the entire calendar year of 2012. There are six tags of interest for up to five different oil wells, and there are also some events to be investigated, as shown in Figure 6 (entire calendar year 2012). This figure shows that there are two areas (May and August) where potential anomalies may be occurring in the data. One important question may revolve around what is causing the first dip in May to occur. This is further investigated throughout this paper.

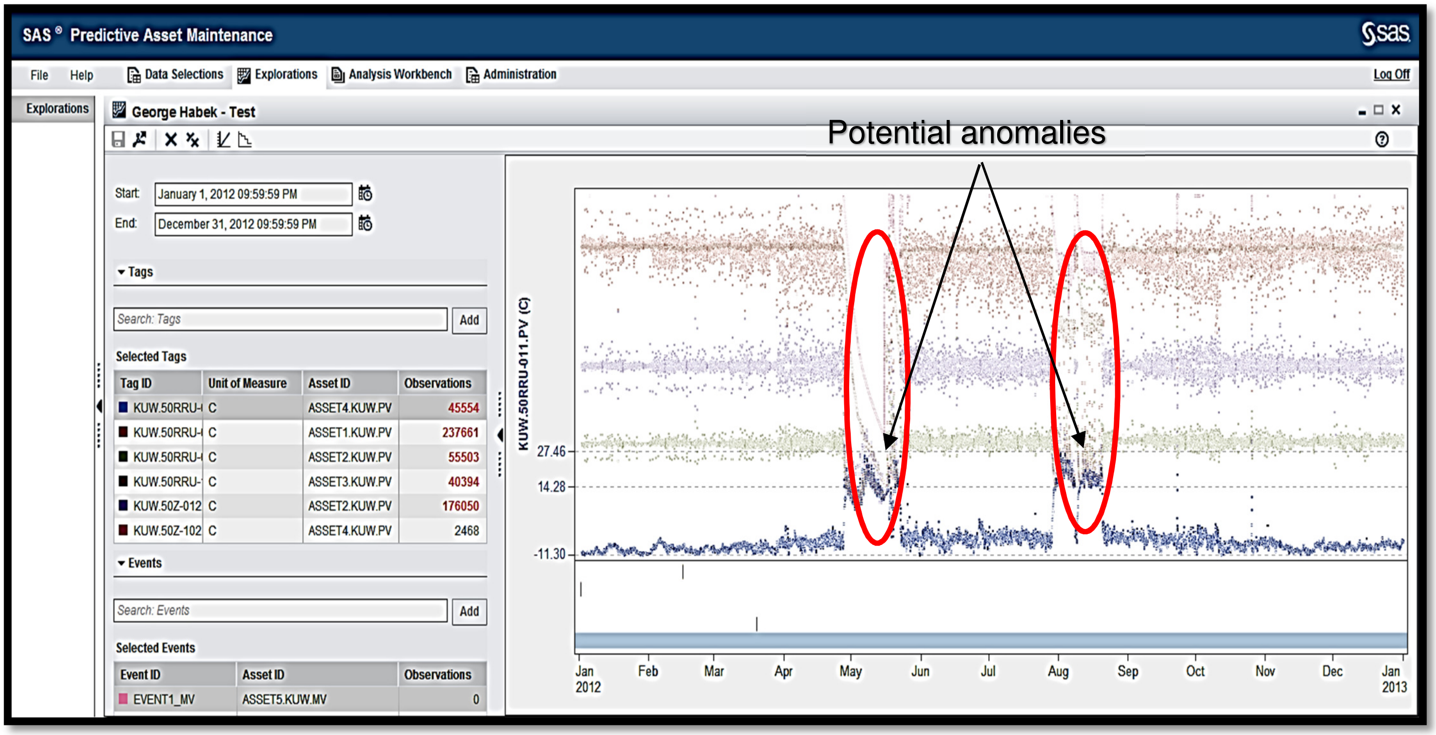

Figure {6}. Potential Anomalies in Data (EXPLORATIONS)

DATA SELECTIONS

The next logical step from the "Explorations" section is to create a candidate filtered subset for analytics based on insight discovered from the visual assessment. The "Data Selections" section is basically a set of questions that users respond to, depending on the responses selected. These questions are reviewed and applied to this particular business example, as follows:

1. Identification (provide a name and description for the data selection):

   Name: ABC Oil Company Data Selection

   Description: This data selection creates a candidate filtered subset.

2. Select a subject area (select either tags or events, or both, along with a date and time for the historical data:

   Tags and events are selected along with a historical time frame from January 1, 2012 – May 31, 2012 in order to investigate the first dip/anomaly observed from the "Explorations" graph.

3. Select dimensions:

   a. Location code (treatment location):

      The location of interest is Kuwait.

   b. Location type:

      Only one location type ("KSC") exists.

4. Select an asset ID (select the assets):

   All five assets for Kuwait are selected.

5. Select tags (select the continuous measurements):

   All possible continuous measurements are selected.

6. Select events (select binary or frequency events):

   All possible events are selected.

7. Subset data (filter on specific tags to include or exclude specific ranges):

   No tags are filtered.

8. Modify data (apply options such as transposition, interpolation, periodicity, and aggregation methods for tags and events):

   All options are applied for the candidate filtered subset. This step of the Data Selection Wizard demonstrates one of the key values offered by the SAS Predictive Asset Maintenance framework, which is that massive amounts of transactional data can be automatically prepared for downstream analytics using these key methods. Typically, this step would take days or even weeks to complete outside of the solution framework.

If an organization desires different and/or additional questions than the ones provided, the Data Selection Wizard can be customized to fit specific business needs. The Analysis Workbench within the SAS Predictive Asset Maintenance framework is discussed in the subsequent section. There are two potential paths:

- One that utilizes the solution interface directly
- One that utilizes other SAS software to perform advanced analytics.

PATH I: ANALYSIS WORKBENCH---MAINTENANCE ANALYSIS AND STABILITY MONITORING

MAINTENANCE ANALYSIS

This area involves performing a set of 12 SAS® Stored Processes known as “root cause analysis” (RCA). The required data set comes from the Data Selection Wizard. The following criteria must be in place for the candidate filtered subset in order to execute an RCA:

- It must contain tags only (no events)
- It must not be transposed.

The steps are as follows:

0. IDENTIFICATION:

   A name, description, and SAS Stored Process folder for the analysis are provided.

1. INITIALIZATION:

   This SAS Stored Process does not require input parameters. Output is generated by clicking on “Run and Save Step.” The corresponding output is the data selections available for the RCA. The log can be viewed, if necessary.

2. SPECIFY DATA SELECTION:

   The actual data set of interest and the type of report (that is, detail or summary) are selected. It is recommended that detail reports be selected, as they provide an abundance of output required for future steps:

   - Data table metrics
   - Tags in the data selection
   - Tag counts by year
   - Tag counts by month and year
   - Date and value ranges for tag data included for the analysis
   - Time series plots of the tags
   - Histograms of the tags.

3. FILTER DATA:

   The beginning and end dates for the analysis are selected. In this example, the range from January 31, 2012 – May 31, 2012 was selected based on the exploration results observed in Figure 6. Tags must also be selected, and the six tags explored in this figure are used in this example. The output includes metrics for the filtered data table for the date range specified and the tags selected.

4. EXPAND AND BUILD STATISTICS:

The interval (in minutes) is entered for the analysis, and the default value of 20 is used in this example. The cut-off date must also be entered. In this analysis, a validation data set is used for verification of the results, which contains data from the cut-off date to the end date. This is sometimes viewed as a “hold-out” data set. In this example, a week prior to the end date of May 23, 2012 is used. The final set of information required for this step is the minimum and maximum band values, as well as how missing values should be handled for each of the tags selected in Step 3. This information is revealed from investigating the histogram results of the tags from Step 2. In this example, six tags must be observed for the bands. The main goals in determining the bands are to identify where the majority of the data points lie for a uniformly normal distribution and to observe any outliers in the left and right tails of the histogram distribution. The first tag in this example is shown in Figure 7. It is quite from clear from this figure that the uniform distribution rests between -10 and 0; the remainder of the bars indicate possible anomalies within the data. Thus, “-10” is selected as the value for the minimum band, and “0” is selected as the value for the maximum band. These bands produce some key event variables from the tag variable in order for the analysis to be able to identify potential root causes of a target event. Specifically, an event variable is created for the minimum band in the following manner:

If the value is < -10 then the event variable is a 1; else it is a 0.

The same holds true for the maximum band value of 0. Several other variables pertaining to the different percentiles that may be of interest are created from these event variables. The “Set Missing” option in this step pertains to handling missing values, with the two options being “Previous” and “Average.” “Previous” is selected in this example, as it relates to the dead band and replaces any missing values with the last present value for the tag.

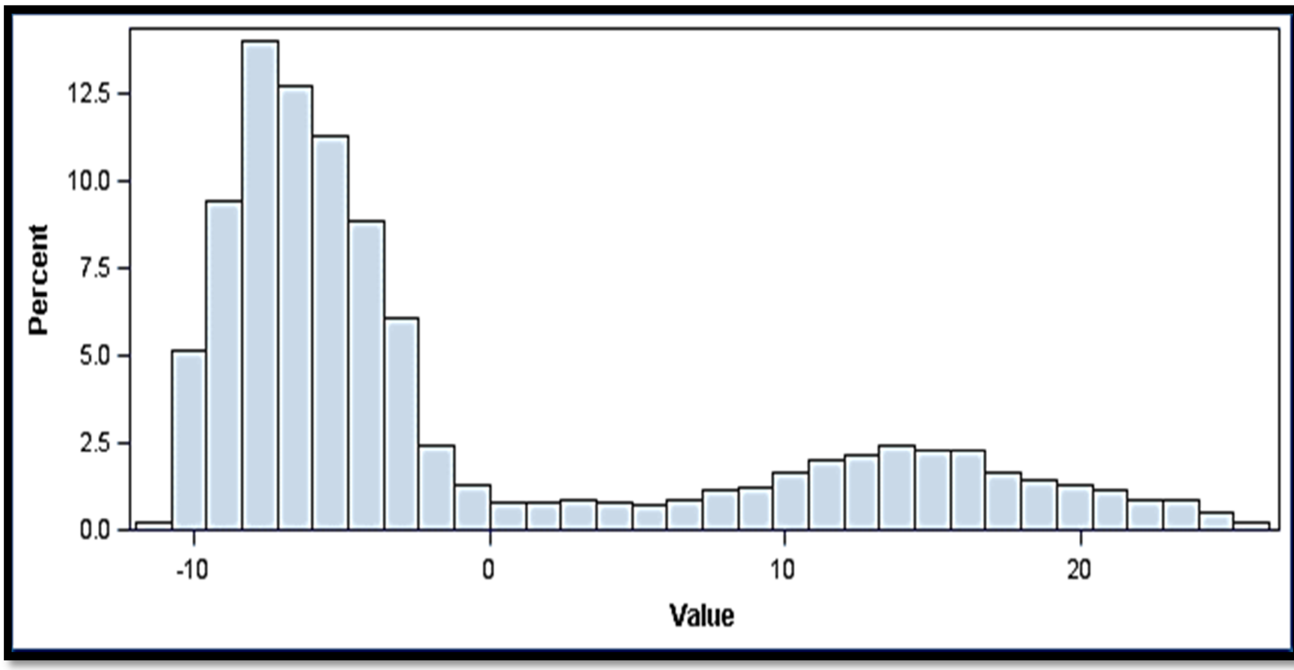

Figure {7}. Histogram of Values for Tag KUW.50RRU-011.PV (EXPAND AND BUILD STATISTICS)

5. CREATE DETAILED TAG ANALYSIS:

The event date and days prior to the event are provided. In this example, April 27, 2012 is selected as the event date, again based on subjective inspection of the results in Step 2. This date seems to be when anomalies began. It is desirable to have approximately two weeks of advance notice or an alert prior to the target event occurring, so 14 days prior to the event is selected, and the output displays a dashboard of tag measurements for 14 days prior to the event date of April 27, 2012. This dashboard represents the time series plots using the minimum and maximum bands for the six tags selected.

6. DEFINE RULES FOR TARGET EVENT:

The target event name is defined. In this example, the issue is a “malfunction.” Up to two tags can be examined with corresponding statistics. Only the “KUW.50RRU-011.PV” tag is used in this example, and it is selected above the fixed value set in the value band (over band). Determination of these values is done based on Step 5. At the end, the default number of lags between measures of three remain (Figure 8).

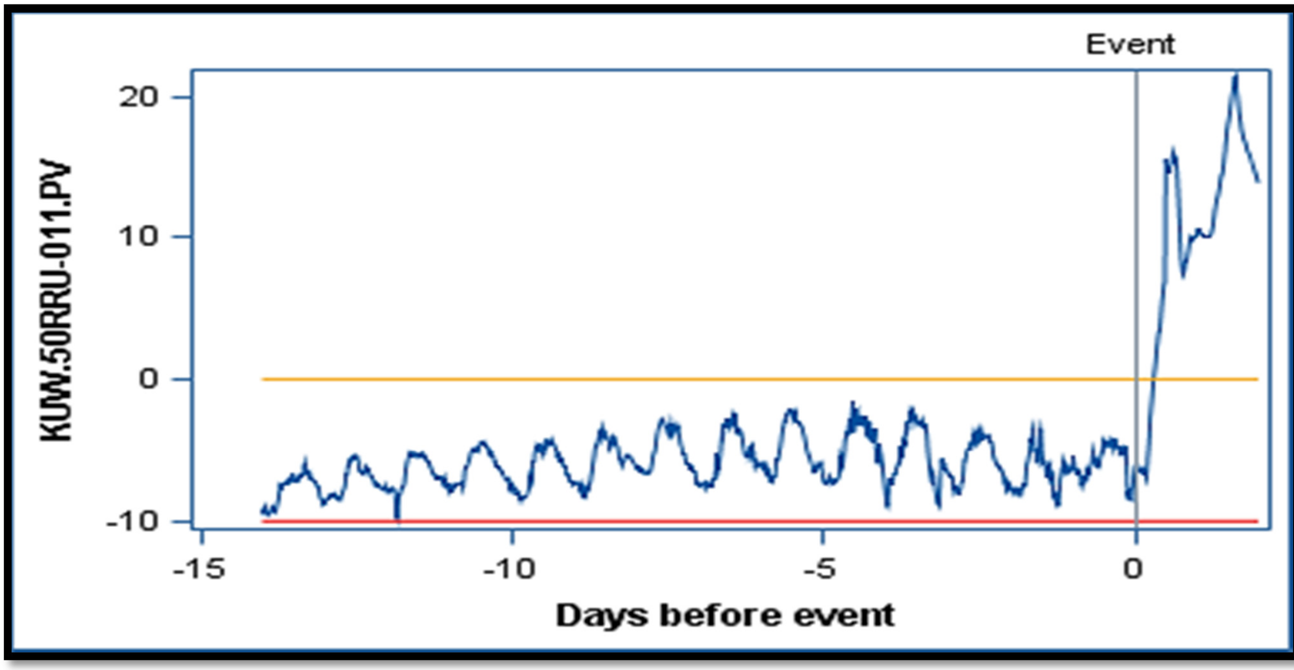

Figure {8}. Number of Lags (DEFINE RULES FOR TARGET EVENT)

7. PREPARE COUNT DATA:

The frequency (in hours) of the count data set is selected, and the default value of “4” is used. The output is event count data summary statistics by event, with the frequency of counts equal to four hours.

8. DEFINE AND ANALYZE EVENTS:

There are two entries to be made here:

- Threshold for masking small variation, which pertains to the dead band value
- Time (in minutes) prior to each target event analysis.

The default values of 0.15 and 240 minutes, respectively, will remain for both entries. The output is a chart that displays the derived event variables created in Step 4 for all four events. Figure 9 (top of subsequent page) depicts two events as an example. This figure shows a stacked bar chart of the event variables for the tags where the weight against the time to event appears, with “0” being the point in time at which the target event (“MALFUNCTION”) actually occurs. This provides a visual indication of which event variables give early warning alerts for the target event (“MALFUNCTION”).

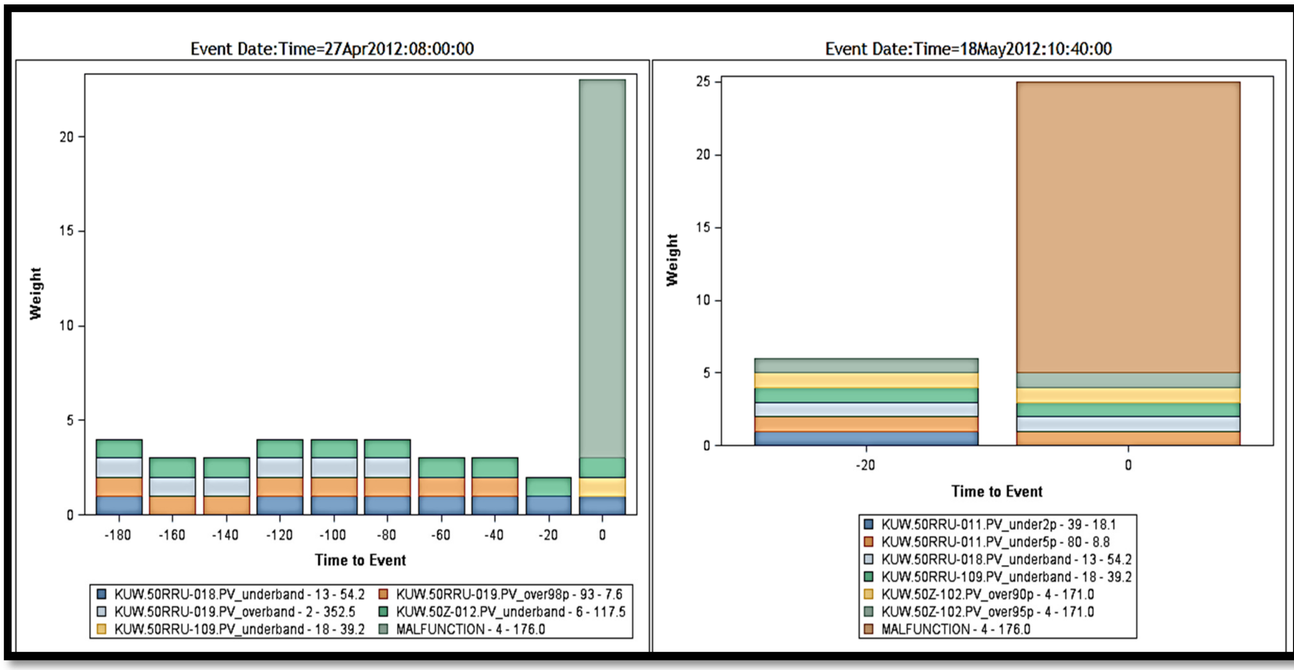

Figure {9}. Two Examples of Events (DEFINE AND ANALYZE EVENTS)

9. CREATE ROOT CAUSE REPORTS:

   This SAS Stored Process does not require input parameters. In order to generate output, users click on the "Run and Save Step" button. The corresponding output is shown in Figures 10 and 11 (latter at top of subsequent page). These figures indicate which of the event variables tend to lead to the target event ("MALFUNCTION"), as well as which events provide early warning alerts based on minutes prior to the first occurrence of the target event ("MALFUNCTION").

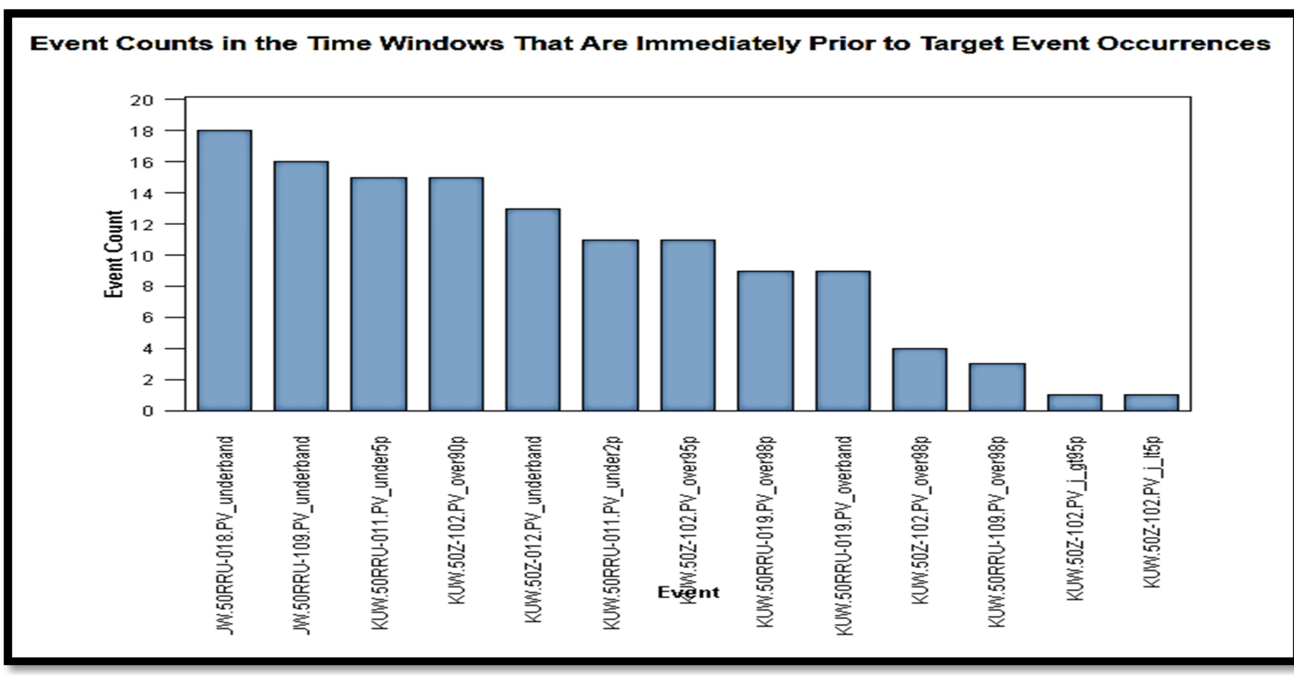

Figure {10}. Event Counts in Time Windows Immediately Prior to Target Event ("MALFUNCTION") (CREATE ROOT CAUSE REPORTS)

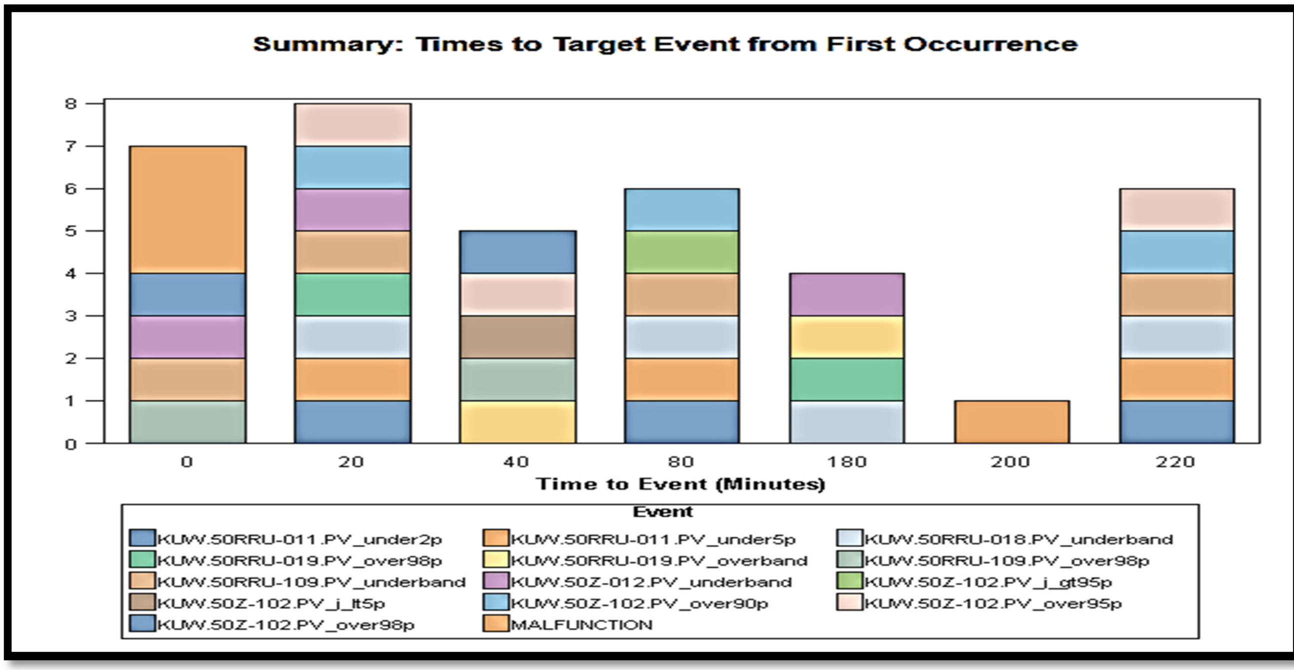

Figure {11}. Times to Target Event ("MALFUNCTION") from First Occurrence (CREATE ROOT CAUSE REPORTS)

10. PREPARE ASSOCIATION DATA FROM SAS® ENTERPRISE MINER™:

This SAS Stored Process does not require input parameters. In order to generate output, users click on the "Run and Save Step" button. The output displays a list of single- and double-sided rules leading to the target event ("MALFUNCTION"). Figure 12 depicts the single-sided listing. This figure displays association rules directly produced from SAS Enterprise Miner. The left-hand side of the rule indicates which specific event variable from the tag tends to lead to occurrence of the target event ("MALFUNCTION"). An important measure of predictive power is the first column (Lift), which indicates how many times more likely the target event ("MALFUNCTION") is to occur given that the event variable is present. One rule of thumb for lift is that a minimum of "2" indicates good predictive ability. As can be seen in this example, all the lifts are well above that minimum threshold, with the lowest being 21.86.

List of Single-Sided Rules Leading to MALFUNCTION

Lift	Transaction Count	Rule	Left Hand of Rule	Right Hand of Rule	Rule Item 1	Rule Item 3	Rule Index (Observation Number in ALL2)	Expected Confidence(%)	Confidence(%)	Support(%)
75.84	6.00	KUW.50Z-102.PV_over90p ==> MALFUNCTION	KUW.50Z-102.PV_over90p	MALFUNCTION	KUW.50Z-102.PV_over90p	MALFUNCTION	10	0.99	75.00	0.74
75.84	6.00	KUW.50Z-102.PV_over95p ==> MALFUNCTION	KUW.50Z-102.PV_over95p	MALFUNCTION	KUW.50Z-102.PV_over95p	MALFUNCTION	8	0.99	75.00	0.74
67.42	4.00	KUW.50Z-102.PV_over98p ==> MALFUNCTION	KUW.50Z-102.PV_over98p	MALFUNCTION	KUW.50Z-102.PV_over98p	MALFUNCTION	12	0.99	66.67	0.49
67.42	4.00	KUW.50RRU-019.PV_overband ==> MALFUNCTION	KUW.50RRU-019.PV_overband	MALFUNCTION	KUW.50RRU-019.PV_overband	MALFUNCTION	14	0.99	66.67	0.49
54.45	7.00	KUW.50Z-012.PV_underband ==> MALFUNCTION	KUW.50Z-012.PV_underband	MALFUNCTION	KUW.50Z-012.PV_underband	MALFUNCTION	21	0.99	53.85	0.87
25.28	7.00	KUW.50RRU-018.PV_underband ==> MALFUNCTION	KUW.50RRU-018.PV_underband	MALFUNCTION	KUW.50RRU-018.PV_underband	MALFUNCTION	31	0.99	25.00	0.87
21.86	8.00	KUW.50RRU-109.PV_underband ==> MALFUNCTION	KUW.50RRU-109.PV_underband	MALFUNCTION	KUW.50RRU-109.PV_underband	MALFUNCTION	37	0.99	21.62	0.99

Figure {12}. Single-Sided Rules Leading to Target Event ("MALFUNCTION") (PREPARE ASSOCATION DATA FROM SAS® ENTERPRISE MINER™)

11. ANALYZE TAG ASSOCIATIONS:

This SAS Stored Process does not require input parameters. In order to generate output, users click on the "Run and Save Step" button. The output displays a "constellation chart" for each association rule (both single- and two-sided). Figure 13 depicts the single-sided chart as an example. This figure displays the association chart for the rules based on the event variables of the tags and their relation to the target event ("MALFUNCTION"). The sizes of the lines between the nodes is based on the lift for that specific rule, while the size of the nodes themselves is based on the maximum number of occurrences for that node in that specific rule.

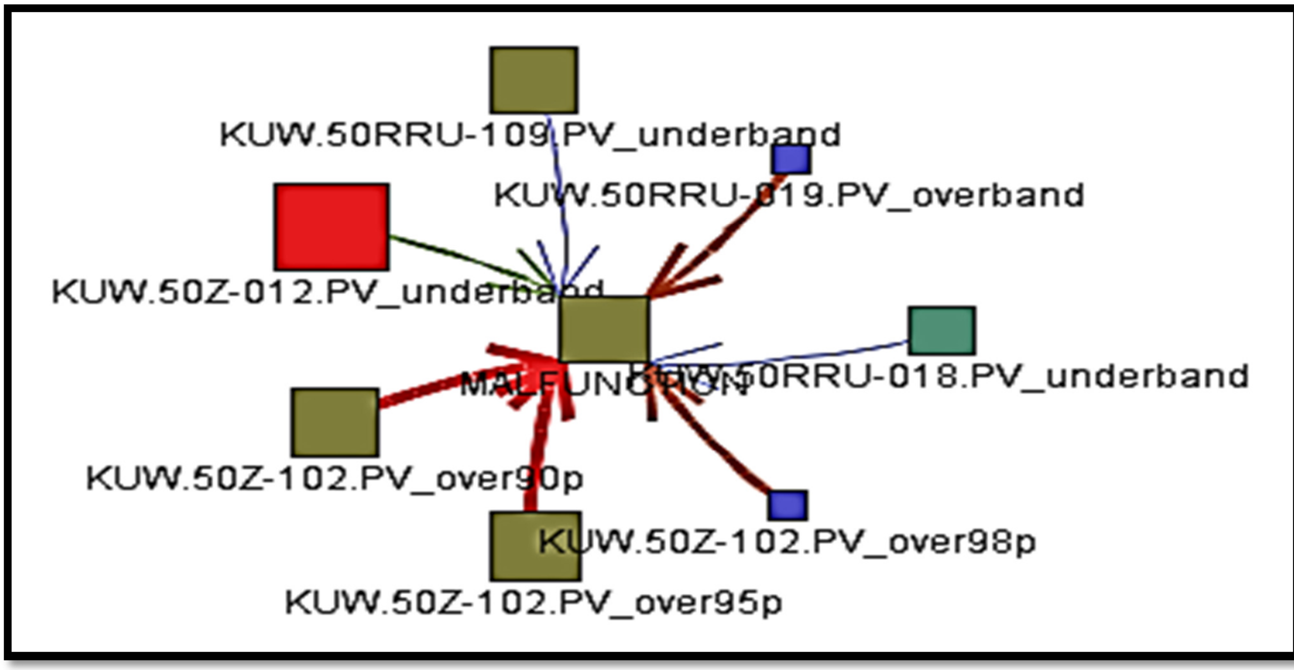

Figure {13}. Example Single-Sided Chart for Event-Based Rules (ANALYZE TAG ASSOCIATIONS)

12. VALIDATE RULES:

The last step uses the validation data created from the cut-off date set earlier as May 23, 2012 in Step 4. There are two entries required here:

- Number of unique analytical tags present when the event occurred
- Number of unique analytical tag rules that resulted from validation rules.

In this example, six unique tags and seven rules are entered using the one-sided association rules from Steps 10 and 11.

STABILITY MONITORING

This area involves the application of predictive modeling techniques to a specific continuous measurement (tag) from the historical data. The main goal is to understand the drivers for that measurement, which becomes a target (*y*) in the model and produces predictions for future scoring efforts. The steps are as follows:

0. IDENTIFICATION:

A name and a description for the analysis are provided.

1. SELECT TRAINING DATA:

The time period for the training data (in this example, January 1, 2012 – May 31, 2012) is selected. The target variable (*y*) of interest is the "KUW.50RRU-011.PV" tag, which refers to a temperature reading (in Celsius) within the oil well completion. One or more stable data periods must also be selected. Because the dip occurred in May, January 2, 2012 – April 27, 2012 is selected as the stable period (Figure 14; top of subsequent page).

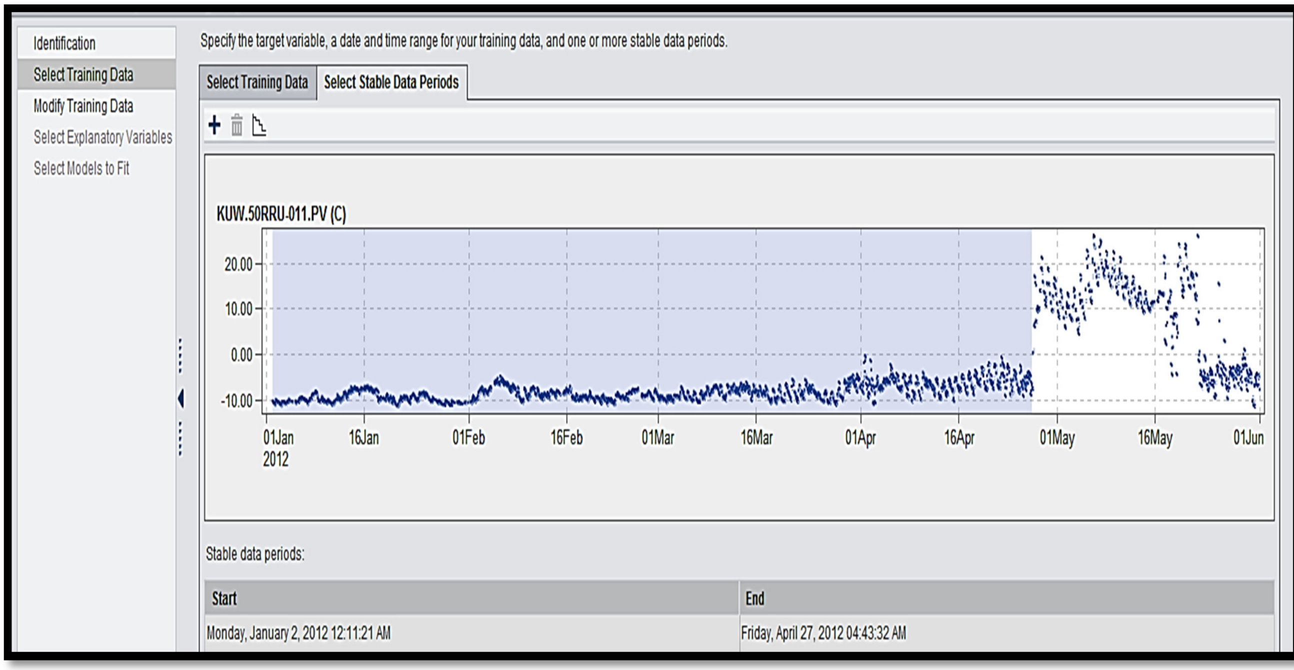

Figure {14}. Selection of Stable Period (SELECT TRAINING DATA)

2. MODIFY TRAINING DATA:

   This step involves modifying the data for model estimation. The selected data set is transposed and tag values are interpolated using the step function. The periodicity selected is one hour. The mean aggregation method is used and the option to apply the natural log transformation exists, but it is not used in this example.

3. SELECT EXPLANATORY VARIABLES:

   This step involves selection of any additional tags and events to use as explanatory variables (*x*) during the fitting process. All possible tags with observations available to enter into the model are selected. Events can be selected, but in this example tags are used.

4. SELECT MODELS TO FIT:

   This step allows selection of up to two models for estimation, as follows:

   - Regression analysis model (REG)
   - Autoregressive integrated moving average model (ARIMA).

   In this example, both models will be selected (Figures 15 and 16; both at top of subsequent page). Figure 15 displays the “champion” model selected by the stability monitoring process. In this case, the ARIMA model had a better fit for monitoring the stability of the target variable for the oil well completion. One important statistic is the ADJRSQ (adjusted R^2), which is a measure of the amount of variability explained by the model accounting for the inclusion of other variables (*x*). In this example, this value is higher at 93 percent for the ARIMA model, as opposed to 85 percent for the REG model. Another important statistic is the mean absolute percentage error, which measures the accuracy of the historical fit. Typically, this value should be around 10 percent or less (that is, 90 percent or more accurate). In this example, the accuracy for the ARIMA model is approximately 95 percent, while the REG model is at approximately 92 percent accuracy. Ideally, then, the ARIMA model would be selected for scoring a new set of data, but that can be overridden and the REG model can be selected instead. Figure 16 displays the selected fit for the ARIMA model. For the most part, the fit of the historical data using the ARIMA model is acceptable, with a couple of notable exceptions. The main “lack of fit” deals with the key time period of May 2012 revealed by the “Explorations” section. The main goal of this exercise is to find the appropriate algorithm to monitor the stability of the selected target (*y*) in helping to solve the oil well completion issue of non-conformance.

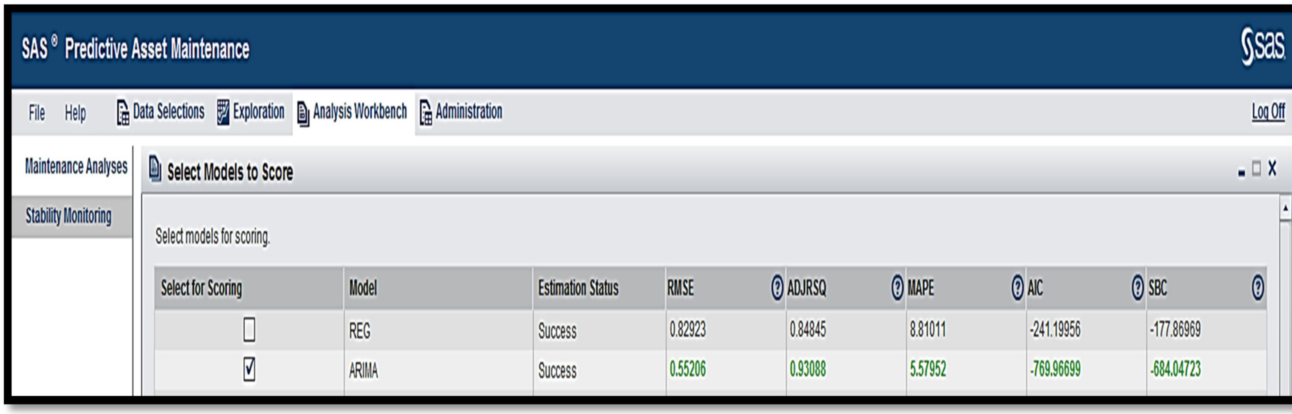

Figure {15}. Champion Model Selected by Stability Monitoring Process (SELECT MODELS TO FIT)

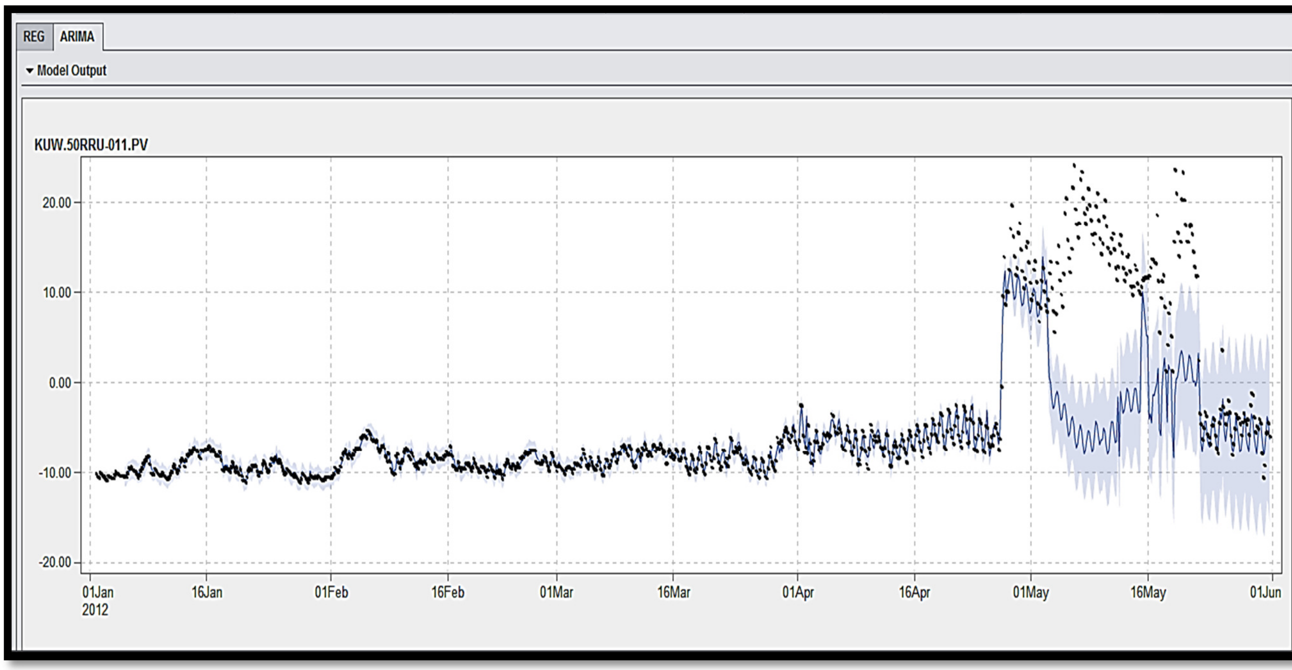

Figure {16}. Selected Fit for ARIMA Model (SELECT MODELS TO FIT)

PATH II: ANALYSIS WORKBENCH---JMP®, SAS® ENTERPRISE GUIDE®, AND SAS® ENTERPRISE MINER™

The next area begins with using a candidate filtered subset from the “Data Selections” portion of the interface within SAS Predictive Asset Maintenance. The data set can be directly opened via three different SAS applications, as follows:

- JMP
- SAS Enterprise Guide
- SAS Enterprise Miner.

The Kuwait example described as the business scenario is utilized in this paper.

VISUAL EXPLORATION WITH JMP®

The first tool, JMP, is known for powerful data visualization and exploration that is very interactive. The distribution of some of the variables is assessed within the candidate filtered subset already created. Some of the tags already explored in the “Data Selections” section are seen here, in addition to others. Figure 17 depicts the normality of the variables while also paying attention to possible anomalies within the data. Basic descriptive statistics that can be customized from a list of many options are also displayed.

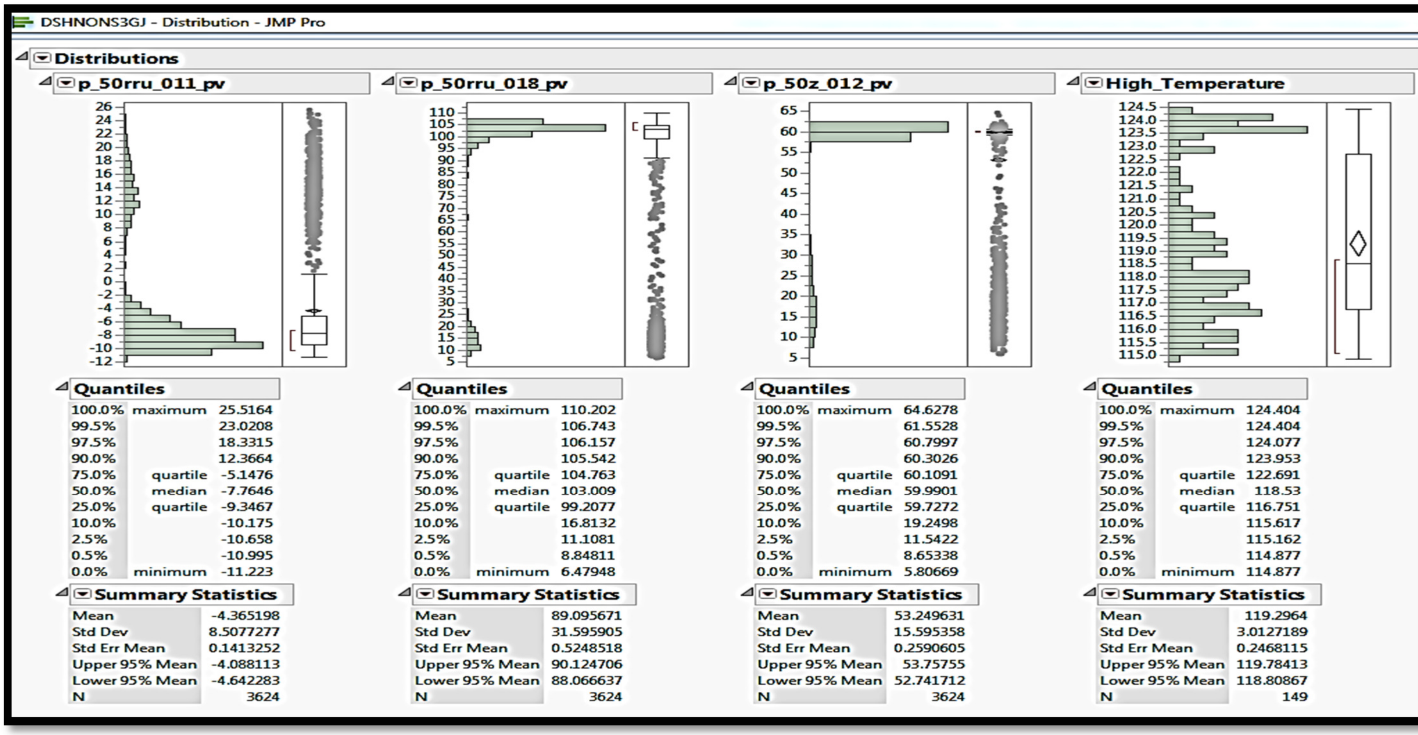

Figure {17}. Normality of Variables (Including Anomalies) (VISUAL EXPLORATION WITH JMP®)

ANALYTICAL PREPARATION WITH SAS® ENTERPRISE GUIDE®

SAS Enterprise Guide is an excellent tool for analytically preparing and querying data in order to produce final modeling data sets. Based on visualization of the data within JMP, some rules may be set up in order to define a target event (*y*), which is critical for building predictive models. A comprehensive method within SAS Enterprise Guide for profiling the entire data set is the “Characterize Data” task within the “Describe” menu option for tasks. This task has goals with respect to categorical variables (*x*), as follows:

- High cardinality: create buckets or new variables (expand the original variables)
- 50/50 split (when there are two levels)
- Dominance of a missing level
- Dominance of a specific level.

All of these issues introduce bias within the data mining and predictive modeling process, and thus interfere with successfully answering the business question(s) being asked. After completing a visual exploration within JMP and analytically assessing data within SAS Enterprise Guide, it is determined that one variable is key for the target event (“MALFUNCTION”). The rule is as follows:

IF P_50rru_018_pv > 90 then MALFUNCTION = 1; else MALFUNCTION = 0.

This creates two populations within the data, as follows:

- One for events where the target event (“MALFUNCTION”) did not occur (0)
- One for events where the target event (“MALFUNCTION”) did occur (1).

One best practice is to have a frequency distribution of the 0s and 1s be approximately 80 percent for 0s and 20 percent for 1s; this has ensured very strong predictive models in several past instances. There may be several variables that define the rule for the target event creation, although in this particular example there is only one. Another important ingredient to have within the data set is a unique identifier (ID) for each record, in this example a unique ID for each well completion of the different assets. In order to achieve this, the “asset_id” and “date_time” fields must be concatenated. Figure 18 depicts the SAS Enterprise Guide flow displaying all the steps described here. The distribution of 0s and 1s for the target event (*y*) is 83 percent for 0s and 17 percent for 1s, which is theoretically quite sound for predictive modeling.

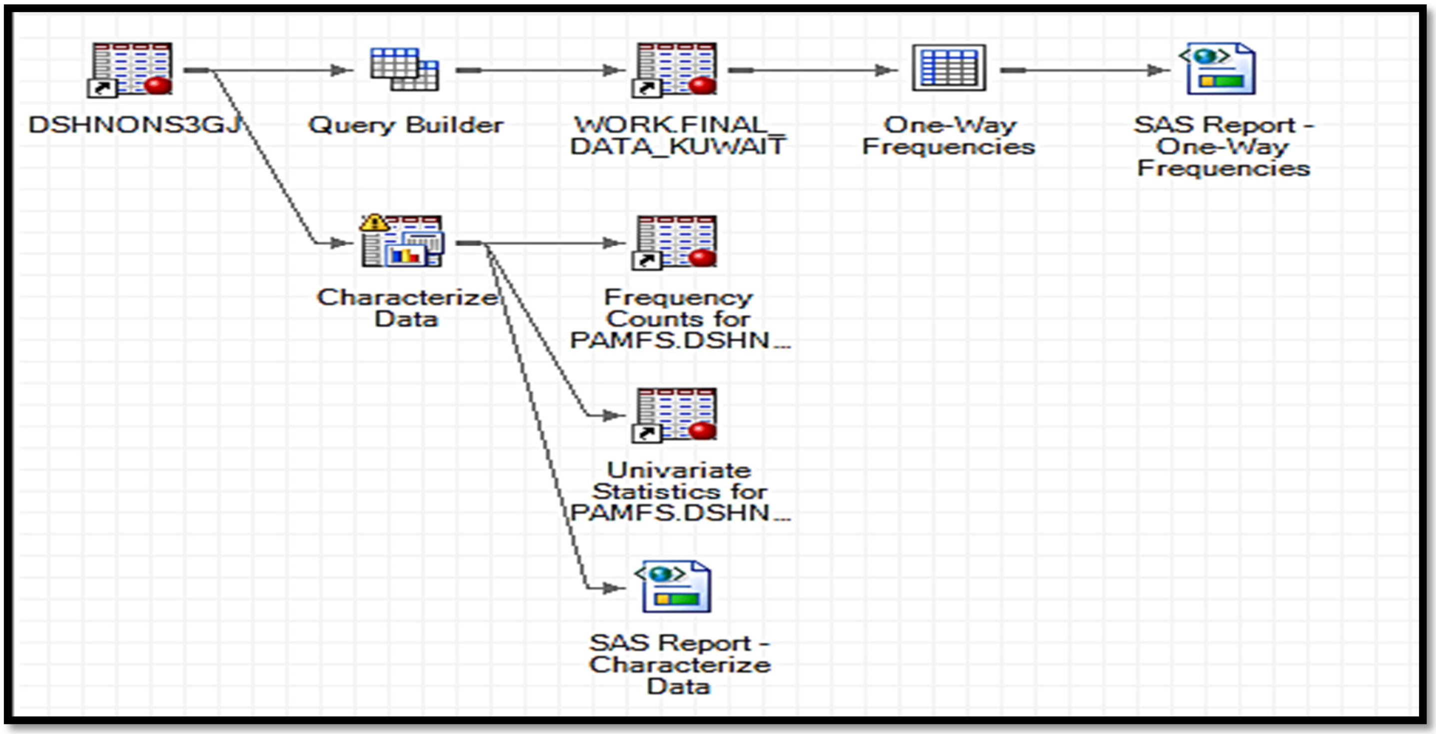

Figure {18}. SAS® Enterprise Guide® Flow (ANALYTICAL PREPARATION WITH SAS® ENTERPRISE GUIDE®)

DATA MINING AND PREDICTIVE MODELING WITH SAS® ENTERPRISE MINER™

Once a data set has been analytically prepared, the data mining and predictive modeling processes can occur in order to solve the business question(s) being asked. The data mining process used within SAS Enterprise Miner is known as “SEMMA,” where the data is **S**ampled **E**xplored **M**odified **M**odeled and **A**ssessed (to complete the acronym). Two important aspects of SEMMA are as follows:

- A candidate filtered subset created from the “Data Selections” area within SAS Predictive Asset Maintenance does *not* automatically create a predictive modeling target (*y*)
- A unique ID must be created for each record.

Both issues can be solved with the “Analysis Workbench” area within SAS Predictive Asset Maintenance or with SAS Enterprise Miner, but a best practice is to conduct data preparation activities in JMP and/or SAS Enterprise Guide. Figure 19 (top of subsequent page) depicts how a typical SAS Enterprise Miner predictive model flow might appear. This figure depicts three models being compared, as follows:

- Decision tree
- Regression
- Neural network.

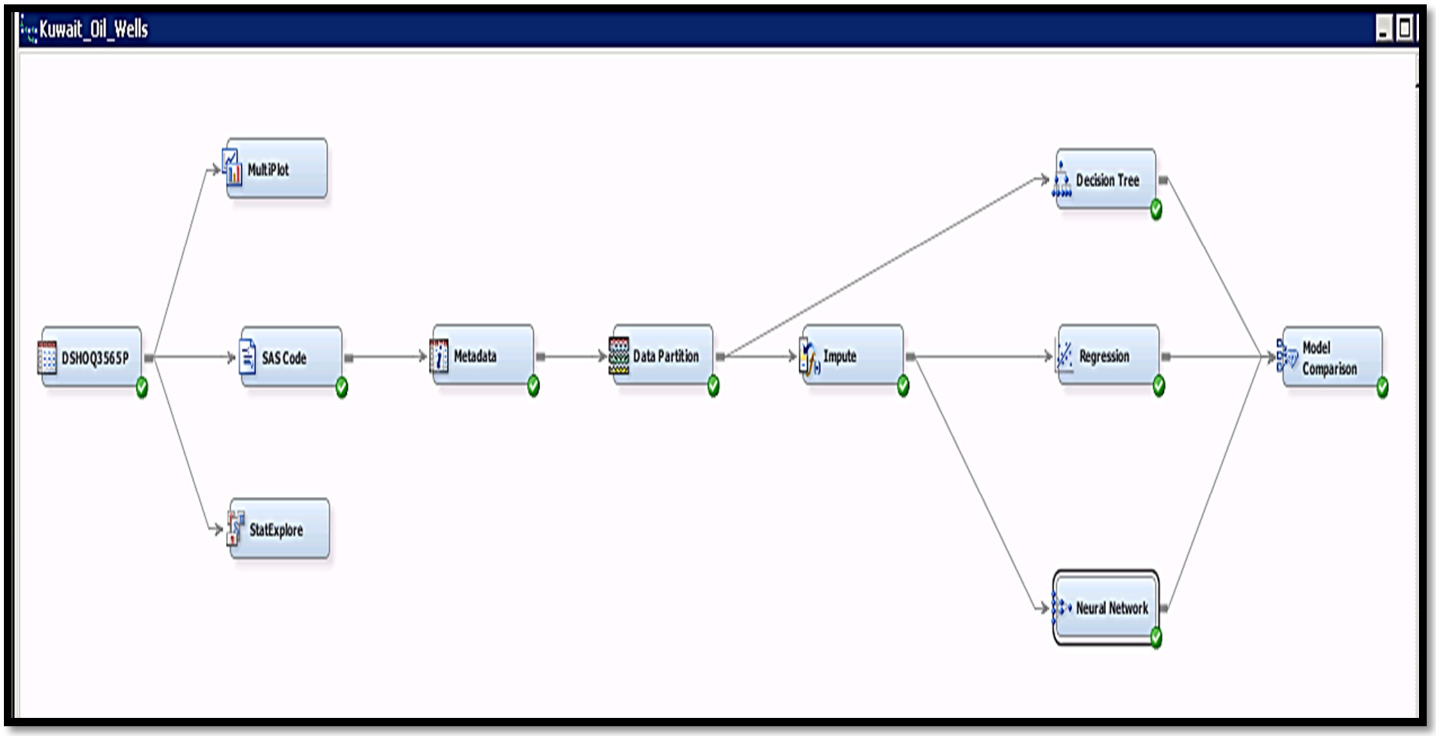

Figure {19}. SAS® Enterprise Miner™ Flow (DATA MINING AND PREDICTIVE MODELING WITH SAS® ENTERPRISE MINER™)

In this example, the regression model for developing the analytical case previously described is selected. The predictive drivers from the model are as follows:

- HIGH_FLOW_RATE_N
- IMP_P_23VXI_001_PV
- IMP_P_50SZO_033_PV
- IMP_P_50TX_033_MV
- IMP_P_50TX_033_PV
- IMP_P_50Z_012_PV.

These drivers relate to two main aspects of the oil well completions, as follows:

- Higher flow rate of steam being injected leads to non-conformance
- Higher amount of chemical ethanol leads to non-conformance.

ANALYTICAL CASE DEVELOPMENT

After data has been loaded into a data mart and compiled into a candidate filtered subset, it is ready for query and analysis. Analytical cases that use data to help monitor assets and make business decisions based on findings can be developed, as follows:

- Consists of four parts:
    - Candidate filtered subsets as input
    - SAS code that processes the data for analysis
    - One or more performance indicators as output
    - One or more analysis results as output.

1. CANDIDATE FILTERED SUBSETS AS INPUT:

   Physical data set that results from subsetting SAS Predictive Asset Maintenance data mart with criteria in “Data Selections” area

2. SAS CODE THAT PROCESSES THE DATA FOR ANALYSIS:

   - Simple business rules or more complex code produced from SAS/STAT® software, SAS Enterprise Guide, or SAS Enterprise Miner
   - Performs additional tasks:
     - Specifies performance indicators as output
     - Can specify analysis results as output

3. ONE OR MORE PERFORMANCE INDICATORS AS OUTPUT:

   - Performance indicators could be simple (e.g., counts) or analytical (e.g., predicted likelihoods)
   - Based on the indicators, thresholds can be defined into categories:
     - *Alarm* – immediate action
     - *Alert* – moderate action
     - *Normal* – no action

4. ONE OR MORE ANALYSIS RESULTS AS OUTPUT:

   Optional additional numeric measures without thresholds.

In this example, the focus is on one or more performance indicators as output. After a predictive model has been developed (for example, with SAS Enterprise Miner), the likelihood scores can be used to create performance indicators based on those scores. For example, values >= 0.90 could be classified as alarms (red/highly out of conformance), values from 0.60 – 0.90 could be classified as alerts (orange/moderately out of conformance), and values <= 0.60 could be classified as normal (green/in conformance). The development of analytical cases can involve the following:

- Using SAS Model Manager to organize analytical cases
- Using SAS Data Integration Studio to deploy models of either type (i.e., SAS program or SAS DATA Step):
  - Setting up SAS Model Manager:
    - Create a folder, project, and version in SAS Model Manager
    - Publish (register) a model in SAS® Metadata Repository:
      - Models that are SAS programs (e.g., from SAS Enterprise Guide) (ScoreCodeType = SAS program)
      - Models from SAS Enterprise Miner (ScoreCodeType = SAS DATA Step); the example in this paper focuses on this method
      - Models from SAS® Business Rules Manager (ScoreCodeType = SAS DATA Step).

Indicators are surfaced in the SAS Predictive Asset Maintenance portlet, which is discussed in the subsequent section.

SAS® PREDICTIVE ASSET MAINTENANCE PORTLET

Once an analytical case has been developed through the application of performance indicators, the SAS Predictive Asset Maintenance portal can be invoked to create a portlet that manages the fleet of oil well completions (assets). Figure 20 (top of subsequent page) is an illustration of the SAS Predictive Asset Maintenance portlet. This figure depicts color-coded bars to indicate the number of performance indicators having statuses of **Alarm**, **Alert**, and **Normal**. Users can filter data by date and category, and each cell in the grid can be selected to display performance indicator values and associated reports. Three dimensions are used to display the data, as follows:

- One dimension is used for filters
- One dimension is used for columns in the grid
- One dimension is used for rows in the grid.

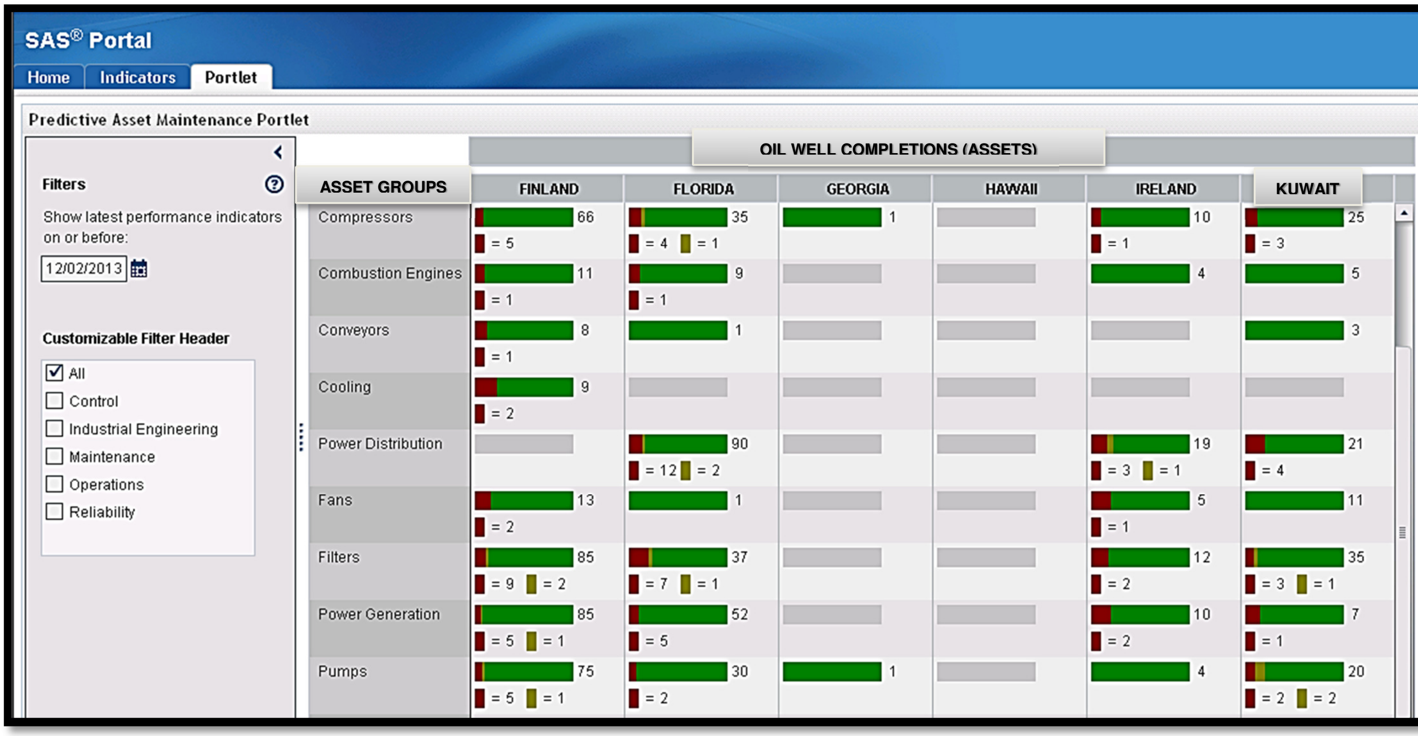

Figure {20}. SAS® Predictive Asset Maintenance Portlet

Colors represent the percentage of performance indicators requiring attention, as follows:

- *Alarm (red)* – immediate action
- *Alert (yellow)* – moderate action
- *Normal (green)* – no action.

The numbers indicate the total number of performance indicators, as well as a breakdown of each type. Clicking on a bar in a cell navigates to a window that shows the performance indicators and their values, along with a list of reports available for that cell. For example, the bar for the Kuwait column at the far right crossed with the asset groups of pumps shows two alarms and two alerts, so the bar is 10 percent red and 10 percent yellow. A gray bar can be seen when the cooling intersection for Kuwait is examined, which indicates that there are no performance indicators for that cell.

CONCLUSION

SAS Predictive Asset Maintenance provides a superior framework for advanced analytics and reporting. Utilizing this solution enables companies such as the one discussed in this paper to achieve the following:

- Improved asset availability
- Minimized unscheduled outages
- Avoidance of performance degradations and catastrophic failures

while at the same time achieving the following:

- Contained cost of maintenance
- Optimization of operating costs.

Remember the following scenario: If there is a valuable asset or machine that is unexpectedly failing, the downtime results in a loss of production throughout the organization. The goal to be proactive rather than reactive, establish the drivers as to what is causing that asset to fail, and develop the likelihood of failure is often critical. Don’t hesitate to find out before it’s too late!

POSTSCRIPT

An all-new version (SAS Predictive Asset Maintenance 6.1, slated for release in early 2014) is being offered from SAS Research and Development. Although JMP® will not be included within this new version of the solution, users will be able to utilize JMP if they are already licensed to do so.

ACKNOWLEDGEMENTS

My thanks for the excellent feedback and suggestions from the following:

- Evan Fuchs, Technical Consultant, SAS Institute Inc.
- Doug Grossman, Senior Technical Editor, SAS Institute Inc.
- Reinhard Höne, Senior Product Manager, SAS Institute Inc.
- Laura Jackson, Software Developer, SAS Institute Inc.
- Gokhan Sevik, Senior Manager for Software Development, SAS Institute Inc.

CONTACT INFORMATION

Comments and questions are valued and encouraged. Contact the author at George.Habek@SAS.com.