Extracting benefits with text mining

Drilling equipment maker FMC Technologies uses JMP® to optimize processes

Manufacturing engineers spend a lot of time crunching numbers to make sure production and efficiency levels remain at their peak.

Now, managers at a Houston-based oil equipment company are discovering that words as well as numbers can be used to analyze problems and make the manufacturing process even better.

“We are always looking to improve our operations, to get better, to be more efficient,” says Dan Fortune, Global Business Excellence Manager for FMC Technologies Inc. “And we have found that text mining can help us do that in a significant way.”

Text mining involves sorting words and phrases into categories based on commonalities. At FMC Technologies, the words come from manufacturing reports. They are coded and then analyzed using JMP statistical discovery software from SAS.

FMC Technologies is a global leader in equipment for the oil and gas industry, including special systems used in subsea oil extraction, surface wellheads and marine loading operations. Some of the equipment manufactured at FMC Technologies is made to be used undersea at depths of nearly 2 miles.

The ability to accurately analyze engineering and performance is an ongoing need for any manufacturer, as designs are adjusted during setup and often become variable over time. At FMC Technologies, ongoing analysis and troubleshooting takes place at 30 production facilities in 16 countries.

“There are quite a few manufacturers out there, but we are the leaders in subsea equipment,” Fortune says.

Using statistical and graphical analysis in JMP over the past several years has made FMC Technologies’ manufacturing processes more efficient and provided a significant cost savings, Fortune says.

“If I wanted to look at all the manufacturing defect areas for 2013, it would probably take two people an entire week to do that task and do it right,” he says. “With JMP, I can do it in 15 minutes, then do it over and over, play with it, and look at it a lot of different ways.”
Text mining:
An added advantage

Working with numbers comes naturally to engineers, but adding words to the analytical mix provides additional information that improves prioritization and problem-solving tasks, says Gokhan Sarpkaya, Subsea Systems Western Region Planning System Manager for FMC Technologies.

“Our goal is reducing or eliminating variations,” Sarpkaya says. “Text data provides an advantage. It allows us to do a deeper analysis and help decision makers make better decisions.”

The text used for the analyses is taken from written descriptions that are entered into a company database to denote problems, irregularities or other facets of the production process.

“There is often a very accurate written description of the problem at the time it occurs,” Fortune says. “Text mining allows us to group similar problems, even if they are written in different ways. So it helps me to be able to pick out certain projects where we aren’t performing as well as we should, which has the potential to make a significant difference.”

Fortune can select text data from JMP and – using a custom JMP script developed by consulting firm Adsurg – send the data to the open-source statistical

CHALLENGE
Analyze manufacturing performance data to enhance efficiency and prioritize tasks for the design and production of oil and natural gas extraction equipment. Adopt a convenient and flexible data analysis program for Six Sigma and other company functions requiring analytics.

SOLUTION
JMP helps FMC Technologies analyze text as well as numeric data to produce a more thorough and detailed picture of its operations. JMP also is used for quality control training, cost studies and other functions.

RESULTS
A custom JMP application supports a creative integration of text mining and JMP analytics to give this manufacturer the precise information it needs to stay on target with worldwide operational goals.

To group and analyze content from manufacturing reports, FMC Technologies pairs a custom JMP script (right) with the open-source statistical and graphical software known as R to perform appropriate text-mining functions. The results are then returned to JMP (above) for further analysis and visualization.
and graphical software known as R to perform appropriate text-mining functions. The results are then returned to JMP for further analysis and visualization. The script lets users customize how JMP data is processed by R – without the need for coding.

The text-mining techniques address both “stemming” and “stop words” to make the analysis more accurate, Fortune says.

To address stemming, words are broken down to their roots, regardless of what forms they take in a particular sentence. For example, “walking,” “walked” and “walks” would all be grouped together by the stem word “walk.”

“The stems have different endings, but they are all really expressing the same thought,” Fortune explains.

Stop words are removed from the text before analysis because they do not contribute to the content of the sentence. Examples include “the,” “which,” “at” and “on.”

“We use R to reduce the complexity of the text data, turning it into a matrix of numbers that can be used for follow-up analysis in JMP. This technique is known as singular value decomposition,” Fortune says. “R sends data back to JMP, and all the analysis is done in JMP.”

Once these functions are carried out, the similar concepts that emerge can be grouped and added to the numeric data typically used for analysis.

Text analysis can help mitigate errors that often pop up in numeric systems.

“Quite often the code doesn’t match the text,” Fortune says.

Errors can occur when numeric coding categories are too general and fail to permit detailed diagnosis of a problem, he says. Problems also crop up when codes are overly specific, causing confusion in determining the appropriate designation for a given entry.

“Too many codes make it difficult for people putting in the data to enter it correctly,” Fortune says.

Analyzing text and data graphically allows managers to identify trends in recurring operations, such as discrepancies between anticipated and actual start and completion dates or the output quality for an individual activity. At FMC Technologies, this can seem a formidable task.

“If I have five projects with 15,000 activities, that gives me 75,000 lines of activities,” Fortune explains.

“For one human being to be able to go in there and combine them where different texts are expressing the same thought and do an analysis, it may not be possible. But with text mining and a cluster analysis, it takes a while, but very few hours compared to practically impossible.”

**Switching to JMP**

JMP also is the company’s go-to tool for project prioritization, an important step
in staying ahead of the competition. It is also now used by training leaders in the Six Sigma program to help ensure quality in manufacturing for reduced costs and increased customer satisfaction.

Sarpkaya says he used Minitab for a decade to train new users in the Six Sigma system before switching to JMP last year.

“I’m very happy with the switch,” he says. “JMP is more customizable. You can write your own scripts. It’s just much more flexible.”

Fortune says he championed JMP for the Six Sigma program at FMC Technologies because he had used it at a previous company and thought it was a superior product.

A project carried out by Fortune led to a major cost savings for FMC Technologies and helped convince company leaders to go with JMP.

Energy industry equipment is expensive and large, with some parts weighing 80 tons. Parts that are deemed defective or unreliable must be scrapped, costing the company $250,000 or more per item in some cases, Fortune says.

After coming on board as a quality leader at FMC Technologies in 2011, Fortune had questions about the testing measure used to determine the viability of the equipment, so he did an analysis using JMP.

“I used the graphical capabilities of JMP to demonstrate the results to many people in the company who had no concept of statistical techniques,” he says. “I was able to show them that there were enough questions that we should have an evaluation by a third party.”

The outside evaluation revealed that some equipment that had been designated for scrap was indeed reliable and could be brought back into use. Two of those items had a combined value of $400,000, thus saving the company nearly a half-million dollars.

But he says the biggest advantage of JMP is the way it enhances FMC Technologies’ ability to “meet customer expectations.”

“Almost all of our subsea equipment are unique designs,” he explains. “We also make surface equipment, where there’s an incredible amount of competition in the market. It’s important to do a really good job of forecasting to make sure we meet customers’ needs at all times.”