

Common data analysis tools do not meet the needs of test organizations

A comparison of Verifide's data management and analysis capabilities against some other industry leading tools.

Rev 0.2 (10/4/2017)

Neil Baliga

Verifide Technologies, Inc.

Executive Summary

Manufacturing organizations are having to contend with devices that are getting more complex. Along with shrinking schedules and higher cost structures, they are facing rising challenges with gleaning useful information from the vast troves of test data being stored. These difficulties contribute significantly to operating and capital costs, and in the worst of cases product failures that should have been caught in the factory. Test data needs fall in 6 categories as follows:

- Storage and Warehousing
- Verification and Validation
- Test-Level Analysis and Reports
- Trending and Visualization
- Analytics and Statistics
- Data Security

The industry leading tools for “data analysis” are often highly capable, but focused on a small subset of the overall needs of data from the above categories. This leaves organizations to handle the remaining needs themselves with messy in-house code or Excel scripts that require effort from specialized and senior engineers. Even when implemented, iteration in test systems and devices is frequent over a few years – resulting in painful recurring costs.

Verifide’ software is a complete test data solution that serves your test data needs throughout the entire life-cycle from design and manufacturing to maintenance. It empowers all the role players in the test process – from engineers troubleshooting on the test floor to systems engineers selling off programs, as well as managers looking for analytics and insightful ways to be more efficient.

	Verifide	Tableau	JMP	Minitab
Storage and Warehousing	Yes	No	No	No
Verification and Validation	Yes	No	No	No
Test-Level Analysis/Reports	Yes	No	No	No
Trending and Visualization	Yes	Yes*	Yes*	Yes*
Analytics and Statistics	Yes	Yes*	Yes*	Yes*
Enterprise Automation	Next Release	No**	No**	No**
Data Security	Yes	Partially	No	No

* Only performs analysis on neat structured row/column data, user must import/convert data.

** Can create reports on demand, but requires custom coding to automate and email.

Introduction

The handling of test data in manufacturing organizations can be segmented in 6 categories each of which poses challenges. Verifide's software provides a solution for each of these to offer a comprehensive data management and analysis platform; This platform serves all handlers of test data from test engineers troubleshooting on the floor, to systems engineers selling off projects, to managers using analytics to estimate project costs better and get insights to improve efficiency.

The section after this in the document compares Verifide's solution against 4 other common data analysis tools including Excel, Tableau, JMP, and Minitab.

Storage and Warehousing

Warehousing data is one of the biggest impediments to productive data analysis – this is because data that is scattered across files without relational connections makes higher level analysis difficult.

Organizations have two options with regards to their test data: Either 1) store their data in a custom and rigid database that is expensive to design, develop and maintain (and iterate) or, 2) store their data in test data files with various formats (xml, csv etc.) because it provides more flexibility.

When data is stored in a database, it requires a specific schema and structure. The code to store and analyze the data is subject to frequent iterations making it expensive and difficult to support multiple schemas and code bases.

When data is stored in files, the ability to analyze the data is limited; data in files are scattered and not “structured” in a neat row/column table, relationally mapped or indexed in any way. Getting information from files requires lots of manual sorting and filtering to reduce the data set to analyze. Keeping track of test runs and re-runs is difficult during analysis because files are not indexed in any way.

Test organizations, also, struggle with data analysis because the software and systems developed to analyze their data (files or databases) cannot be reused across product types, tests, and systems, and often cannot be reused on next generations of the same product.

Verifide's software solves this problem with our **Dynamic Database** technology that allows test data to be stored in a relational database without any database level design, implementation, coding, or iteration. The software builds the database structures dynamically as data is fed to it and it provides an easy to use Application Programmatic Interface (API) for applications to store and retrieve data. Our analytical tools build upon this design to offer a generic way to slice and analyze the data as well.

Verification and Validation

Tests usually perform limit checking to provide a pass/fail disposition. Setting limits is a delicate art. If you set them too wide, then there is a risk that you are not testing thoroughly. If you set them too tight, it means more failed tests and therefore more senior engineers are spending time on the factory floor troubleshooting and dispositioning results. The potential for unnecessary costs is present in both cases, so smarter limits provide an opportunity for efficiency that is often overlooked.

To achieve better limit setting in your process, the data management system needs to support the ability to process and re-process old test data against new limits. This will allow engineers to perform hypothesis testing and allow them to predict failure rates by trying different combinations of limits. It also provides the opportunity to improve production throughput by using wider limits on the factory systems (less down-time) which can then be tightened in post-testing analysis (offline) before sell-off.

Verifide's software solves this problem with the **Limit Checker** in our data analysis application. Users can create and load their own custom limit sets, and do hypotheticals to anticipate the resulting yield across different limit sets. Optionally, with elevated permissions, the user can update the results in the database with the new disposition so that future analysis can reflect the updated limit dispositions. (eg. yield, metrics, trends)

Test-Level Analysis/Reports

While higher level analysis is necessary to get insights into processes and trends, day-to-day production depends upon access to the raw test data at the individual test level. This is typical when anomalies are detected and engineers need to figure out the root cause and disposition the test result.

When a test fails on the floor, it results in down-time because the test must be reviewed, troubleshooted, and often re-tested. Furthermore, when higher level analysis tools expose an anomaly in a trend plot, for example, it becomes necessary to drill-down into the individual test results to identify the root cause of the problem.

Test data is typically composed of multiple data fields and reports are a visual indication of the test result with plots, tables, and data fields etc. These reports often serve as the official record of compliance during sell-off.

However, test reports suffer the same problems as data analysis in general because reports are hard-coded to the data sets, to the test data schema, and most often also tied to test code and logic. Iterating reports can be an inconspicuous contributor to software configuration management problems because it fragments the test system code base.

Verifide's software can retrieve the original data objects stored in the database. It allows the user to view the raw data for tests, and to drill down directly from trend tables into raw results. Our metadata

visualization tools allow the user to go deeper into the raw data by zooming into plots, viewing image data, filtering and grouping group tables, and more.

Our report generation software allows engineers to create report templates that are then filled in with data from the database. This provides a zero-coding reporting system whereby reports can be generated after-the-fact with new templates or augmented data with review notes etc.

Trending and Visualization

Trending and Visualizations provide a subjective view of information that a human can use to assess the behavior of the data. Visualizations are important because patterns in test data are difficult to identify with automated software or artificial intelligence technologies that exist today – not to mention the astronomical costs associated with developing and implementing such models to attempt to do so. In most cases a human can evaluate data patterns more efficiently than automated decision-making software.

A difficulty with visualizations is that there are numerous permutations of ways to visualize test data. An engineer, for example, may want to look at trend data from the last month, then want to look at the trend data for the last 3 months instead, but only for data from test station 1 etc. In other words, one visualization can trigger the need for another view of the data.

When tools like MS Excel are used, such iteration becomes difficult because the worksheets, references, and source data that is imported are all meshed together and need to be reworked; they are also error-prone.

Test data often stores two-dimensional data in arrays and tables such as for power sweeps. This data for a single test is visualized as a plot. For correlation, engineers need to be able to see hundreds or thousands of plots from multiple units or test runs overlaid on top of each other to detect in-family correlation. This is often problematic in most systems as they don't deal with unstructured and multi-dimensional data like arrays and sub-tables.

Another challenge with visualizations is that it requires manual effort from engineers, to filter, find, and extract this data from large volumes of files and folders into a structured form that can be processed by these other tools.

Yet another issue with visualization is schema dependency – when new test data structures are stored, or data formats are changed, the software that was developed for the visualizations (plotting, analysis etc.) also needs to be iterated.

Verifide's data analysis tools build upon our Dynamic Database; they can extract data fields across different tests, devices, and even across databases. These tools operate directly on a dynamic data structure, so there is no hard-coded implementation to view reports and plots.

Users can then create filters and can search and group data to view the data in multiple ways, back-to-back, with no setups or coding needed between them. Our tools allow plotting, pivoting, and correlating data with an easy to use interface.

Verifide also provides a built-in tool for 3D trending – to overlay 2D plots across multiple test runs. Such correlation analysis is useful for visualizing two-dimensional data correlations and detecting out-of-family test runs.

Analytics and Statistics

Analytics and statistics provide mathematical computations based on theoretical models of data. This field of mathematics is vast and contains many different analyses and measurements like process capability and repeatability that are used in various industries. A key benefit of analytics and statistics are that they are based on known scientific models of data and probability that have been proven to be reliable measurements over time.

Analytics and statistics are different from visualization, in that they provide deterministic measurements with algorithms. i.e. they are objective pieces of information, versus visualizations which are more subjective and require human interpretation.

One of the challenges with getting analytics is, again, the problem of non-structured (messy) data and schema dependencies. To process data through algorithms and calculations, it requires reducing the large sets of files and folders into structured results that can be analyzed. Also, when data schemas are changed, such analysis software or worksheets need to also be iterated and are error-prone.

Verifide's software, through our Dynamic Database technology, can reduce large unstructured test data sets into structured tables for analysis. Our software provides statistical analysis capabilities for six-sigma and process capability measurements and we are always adding new capabilities based on our customers' needs.

Enterprise Automation

While using a tool to visualize and analyze data provides great benefit, it can be taken one step further with automation. A manager, for example, may want to see a weekly report that shows the utilization of all the testers. A system engineer, may want a trigger alert when the standard deviation of the temperature readings exceeds a certain value.

With enterprise automation, it reduces the time to get the desired information. It also allows for historical tracking of reports without needing to repeat manual steps to create a report.

Furthermore, dashboards are also necessary to be able to create simplified user interfaces to view just the pertinent information at a glance and these dashboards need to be tailored to each user because there are numerous role players in the test process (e.g. operator, systems engineer, test director, manager etc.)

Verifide is in-process of implementing these enterprise capabilities for our next release, due to be released in December 2017. Users will be able to take a manually created report and configure it to be sent on fixed intervals and/or on trigger events. These reports will support emailing reports via PDF, and the option to archive all reports within the user's profile to view later.

Data Security

Security is becoming a major concern with test systems that are storing mountains of data. Even for non-classified uses, test data can expose lots of proprietary information about a device if it falls into the wrong hands.

One of the major drawbacks to file based test data is that it is vulnerable. Without a container to protect it a single data file can be emailed, copied onto thumb drives etc. File systems also are hard to administer permissions for and can require heavy IT involvement daily to move data around a factory. For organizations that support multiple programs and projects, it may also be necessary to only provide access to certain data to certain users. Again, this involves IT to configure directory level permissions on folders on share drives to different users etc.

For commercial tools like JMP, Tableau, and Minitab that can operate on data from a database, they operate while directly connected to the database – which exposes a connection into the database server. These credentials need to be handed out to each user and administered by a database administrator, making the risk profile for vulnerabilities higher.

Verifide's Dynamic Database is an abstract layer on top of a Microsoft SQL Server database. It implements internal security and **NEVER** exposes a direct connection to the database. Users, roles, and permissions are abstracted and are not implemented at the database level, so they can never log in to the database directly with their credentials – they can only login through the software applications and API and can only perform supported operations. The database can be hardened with encryption to the highest security standards.

Verifide also implements permissions that can be tailored at the database level – allowing certain users access to only databases for projects they are working on, for example.

For reports, Verifide supports security features built into the PDF standard including encryption with password protection, as well as options to prevent printing etc.

Head-To-Head Comparisons

Microsoft Excel

Publisher: Microsoft (<https://products.office.com/en-US/excel>)

Microsoft Excel is by far the most commonly used data analysis tool for most test organizations. This is because it works well with structured data in tables and allows engineers to aggregate data across worksheets and to perform calculations with formulas and additional scripting and macros. Plotting data is easy in Excel – the user can select the data to plot and view many different chart types.

Excel is a “free” product for most users since they likely have subscriptions to Microsoft Office and have this tool available to them. When performing ad-hoc analysis, Excel really shines because it allows the user to manipulate the data in many ways to see deltas, and to even apply transforms and complex math to see different perspectives of data.

For multi-dimensional and relational data, excel provides pivot tables, though it can be a bit difficult to setup if you need to pivot across multiple variables and fields and to repeat the analysis multiple times.

Excel is not, however, designed to handle large data sets or data that is unstructured (not in a nice table format). When trying to perform analysis on thousands or millions of data sets, it becomes difficult to manage that test data in worksheets. Also, if the number of analyses/reports are high, then it becomes difficult to repeat formulas and worksheets because they take up a lot of manual labor; often, this is expensive labor because it requires highly skilled engineers to perform this task.

Data in Excel is analyzed by importing data into it. It can import from databases and other file formats, but it only handles data that is in a row/column format. This makes it hard to work with multiple file formats like XML, or other file formats where data may reside. Source data is also scattered around in multiple folders on disk, so it is a difficult task for the user to hand-select the data of interest to import for analysis.

To use Excel with all these data files, there needs to be some other data converter utility that parses data files and imports them into Excel for processing. Sometimes, this is doable within Excel itself using scripts and macros. Though this system works, it does not tolerate iteration well. For example, when a batch of test data is re-tested, it is hard to find the correct data in the existing worksheets to overwrite so that the reports are updated. A well-developed Excel sheet would have problems if the data format being imported had new fields. The repetitive copy/paste into Excel is also highly error-prone.

Excel data analysis is more vulnerable to security lapses because data is freely imported and moved around and there is zero traceability of what data was used to perform the analysis and reports. There is no user login enforcement or limitation on who can access what data.

Pros: Excel is a great tool for crunching structured data like financial data and specific measurement tables.

Cons: Not a good way to warehouse large amounts of data or perform more iterative data analysis and visualizations. Not a good way to protect secure data.

Table 1- Comparison of test data capabilities between Verifide and Excel

	Verifide	MS Excel
Storage and Warehousing	Data is warehoused in Dynamic Database built on MS SQL Server. Schema-free database.	Does not store or manage test data, user must compose the applicable subset of data as the data source. User must import data into excel worksheets manually and keep track of re-runs etc. Need to convert/parse data in flat files (xml etc.)
Verification and Validation	Limit checking and Yield evaluation tools included. Can post updated yield results back to database for sell-off auditing.	User can setup formulas, but for large test volumes, this is not practical in Excel.
Test-Level Analysis/Reports	Access to raw test data and test-level reports.	Possible, but for larger test volumes, this is not practical to keep all test data in an Excel file and in memory.
Trending and Visualization	Tools to extract data fields across large data sets, pivot tables and charts are easy with drag and drop. Data can be visualized in graphs, histograms, pie charts, stacked bars etc. Supports 3D trending – overlaying plots from multiple tests. Customizable dashboards in next release.	Easy to plot data in worksheets. Can apply formulas and create derivative data to plot. Pivot tables are difficult to setup, charts are not interactive and zoomable etc. Changing data ranges or filter criteria is not user friendly. 3D trending possible but requires specific chart setup, not flexible.
Analytics and Statistics	Common six-sigma and process capability statistics. New statistics and analytics being added driven by customer needs.	User must implement own code and scripts for formulas and algorithms. Must iterate multiple times.
Enterprise Automation	Automated reports and triggers coming in our next release.	None.
Data Security	Verifide provides abstracted access to the database so no connections directly to the database are allowed. Users and roles are abstracted. Databases can be hardened and encrypted to high security standards.	None. Data is even more vulnerable because data is often copy/pasted into Excel and moved around in files.

Tableau

Publisher: Tableau (<https://www.tableau.com/>)

Tableau is a data visualization product. It has both a desktop and server edition that allows access to data over a web browser. Tableau offers a very modern and nice user interface with easy to use visualizations. It provides drag-and-drop fields to view plots and to pivot on data sets.

Tableau supports creation of custom dashboards and allows some sharing and integration of those dashboards if you also have the Server edition of the software. Tableau only has a minimal set of statistical capabilities that sometimes require scripting – it is not designed to compete with JMP and Minitab as they market it as a visualization/business intelligence tool, not a pure analytical tool. The user can get some basic statistical control lines, but cannot view process capability or predictive or statistical analytics.

For multi-dimensional and relational data, Tableau allows you to drag and drop to pivot your data and view graphs directly.

Tableau is only sold on a subscription basis. The price points are relatively low and offer a smaller entry point to get data visualization capabilities out of the box. Some features of the product require a constant connection to the network, but offline license activation and validation can be done monthly or yearly for systems that are not connected to the internet.

To share reports or dashboards requires a separate license for the Tableau Server edition. This allows users to view reports across a web browser, but does not allow the user to change the analysis parameters or create new reports. To create analyses and dashboards still requires the desktop edition.

Tableau only works with structured row/column data sets. It does not warehouse or store data, so all data to analyze must be obtained from somewhere else. It can connect to databases, and unlike JMP or Minitab, it can operate on live databases (without needing to import data into memory). They do recommend importing data to avoid performance issues. They support a wide array of data sources, but regardless of source, they still can only process structured table data.

Like Excel, it is up to the user to search and comb through the troves of test data files to get the data reduced to a neat row/column format that the Tableau software can process. This is not feasible for test data that is inherently unstructured and messy.

The Server edition of Tableau provides many security guidelines and role based authentication including integration with Active Directory systems. However, the desktop edition is the primary application used to interface with data and publish dashboards and reports to the server. If connecting to a database, the desktop edition requires direct access to the database to connect and pull in data. This can leave vulnerabilities by exposing users and connections to the database. If converting unstructured data files into structured tables, that process is also vulnerable to exposure if data is handled manually at various stages.

Pros: Easy to use and nice modern interface. Supports pivoting results and customizable dashboards.

Cons: Does not handle test data well because it is unstructured. Does not have common statistical analysis capabilities. May be a hassle for systems not networked due to subscription license model. No security model for data. Possible security vulnerabilities due to exposed DB connections.

Table 2- Comparison of test data capabilities between Verifide and Tableau

	Verifide	Tableau
Storage and Warehousing	Data is warehoused in Dynamic Database built on MS SQL Server. Schema-free database.	Does not store or manage test data, user must compose the applicable subset of data as the data source. It requires structured data, so cannot deal with unstructured data as is typical in test. Can connect to multiple data sources and databases (structured table only). Does allow working directly with live DB connection (don't have to import into memory)
Verification and Validation	Limit checking and Yield evaluation tools included. Can post updated yield results back to database for sell-off auditing.	Does not provide for limits or hypothesis testing of data.
Test-Level Analysis/Reports	Access to raw test data and test-level reports.	None. Only does higher level data analysis.
Trending and Visualization	Tools to extract data fields across large data sets, pivot tables and charts are easy with drag and drop. Data can be visualized in graphs, histograms, pie charts, stacked bars etc. Supports 3D trending – overlaying plots from multiple tests. Customizable dashboards in next release.	Lot of visualizations and simple to use once it has a neat structured data set. Tableau is primarily a visualization tool and does it well.
Analytics and Statistics	Common six-sigma and process capability statistics. New statistics and analytics being added driven by customer needs.	Minimal statistics capabilities out of the box. The tool is primarily for visualization and does not have advanced statistical computation capabilities like JMP or Mini-tab. Tableau has its own scripting language with formulas and user can implement code in there to do additional analysis.
Enterprise Automation	Automated reports and triggers coming in our next release.	Only available by adding Server edition; user must still must automate with custom script code. Tableau Server

		edition supports more centralized reporting, but does not support periodic emails out of the box – that must be manually added with other software.
Data Security	Verifide provides abstracted access to the database so no connections directly to the database are allowed. Users and roles are abstracted. Databases can be hardened and encrypted to high security standards.	Tableau Server edition has security models for accessing Tableau’s services over the web or intranet– they do not protect access to the data source per se. Desktop edition exposes direct connections to the database.

JMP

Publisher: SAS (https://www.jmp.com/en_us/home.html)

JMP is a “statistical discovery” tool by SAS and has the largest catalog of statistical and predictive analytics in a tool. The Pro version of JMP has advanced predictive analytical capabilities such as neural network analysis. They also have separate products tailored for clinical and genomics applications.

For pure statistical and higher complexity analytics (genetics etc.), JMP offers a very compelling solution with lots of statistical calculations and plots for a variety of industries.

The user interface for JMP is more complex than Tableau or Minitab, and requires specialized knowledge to use the application. Getting additional functionality and creating reports requires writing code in their proprietary scripting language JSL.

JMP only works on imported data in memory. It only works with structured row/column data sets. It does not warehouse or store data, so all data to analyze must be obtained from somewhere else. It can connect to databases, but unlike Tableau, it does NOT operate on live databases. It only uses the database connection to import the data into memory. For larger datasets, memory can become a problem.

Like Tableau, and Minitab, JMP allows importing data from databases. This can leave vulnerabilities by exposing connections to the database. If converting unstructured data files into structured tables, that process is also vulnerable to exposure if data is handled manually at various stages.

Pros: Extremely sophisticated analyses. Supports extensions with their scripting language.

Cons: Not very easy to use – requires specific skills. Does not pivot on data easily, cannot handle unstructured test data well, possible memory issues with large data sets.

Table 3- Comparison of test data capabilities between Verifide and JMP

Verifide	JMP
----------	-----

<p>Storage and Warehousing</p>	<p>Data is warehoused in Dynamic Database built on MS SQL Server. Schema-free database.</p>	<p>Does not store or manage test data, user must compose the applicable subset of data as the data source. It requires structured data, so cannot deal with unstructured data as is typical in test. Can connect to multiple data sources and databases (structured table only). Can connect to database sources, but data must be imported into memory. User must write SQL query to pull data in.</p>
<p>Verification and Validation</p>	<p>Limit checking and Yield evaluation tools included. Can post updated yield results back to database for sell-off auditing.</p>	<p>Does not provide for limits or hypothesis testing of data.</p>
<p>Test-Level Analysis/Reports</p>	<p>Access to raw test data and test-level reports.</p>	<p>None. Only does higher level data analysis.</p>
<p>Trending and Visualization</p>	<p>Tools to extract data fields across large data sets, pivot tables and charts are easy with drag and drop. Data can be visualized in graphs, histograms, pie charts, stacked bars etc. Supports 3D trending – overlaying plots from multiple tests. Customizable dashboards in next release.</p>	<p>A complex and highly capable tool. Takes a lot of know-how to setup and use it, and requires learning JSL Scripting to get most out of the product. Has customizable dashboards that run within the application.</p>
<p>Analytics and Statistics</p>	<p>Common six-sigma and process capability statistics. New statistics and analytics being added driven by customer needs.</p>	<p>Extensive statistical analysis and reports. Has large amounts of capabilities in crunching data and showing analytical plots.</p>
<p>Enterprise Automation</p>	<p>Automated reports and triggers coming in our next release.</p>	<p>Supports this via scripting. User must write code scripts in JMP’s JSL language to automate things. Configurations are not centralized, so user must have the predefined .jmp files physically on the computer itself to run the analysis and create the reports. Does not email reports, that must be automated separately.</p>
<p>Data Security</p>	<p>Verifide provides abstracted access to the database so no connections directly to the database are allowed. Users and roles are abstracted. Databases can be hardened and encrypted to high security standards.</p>	<p>JMP software can connect to the database to import data. This exposes direct connections to the database.</p>

Minitab

Minitab is a “smart data analysis” application. It provides an Excel like worksheet look and feel. Data is imported into worksheets and analysis can be done from there. It offers a comprehensive set of capabilities for six-sigma and process control metrics.

Minitab is very Excel centric and supports excel like workflows. The software has a very easy to use interface and supports working directly with data such as being able to edit data values in the worksheet before performing analysis. It allows importing data and then using formulas to add computed data that can be visualized and processed.

Minitab is available in single-seat licenses and volume licenses. No server or web access to data or dashboards is available, though there are “companion” type products that have dashboard type capabilities which are focused on specific business objectives (eg. tools to optimize Cpk). It is not clear how these companion products are related or not to Minitab itself and whether they can be integrated together.

Minitab provides multiple statistical analysis capabilities from Statistical Process Control, Gauge Study, Analysis of Variance (ANOVA) and much more. These are designed to perform parametric and non-parametric analysis of data. Minitab also has plotting capabilities to plot and view data in multiple chart types along with various regression analysis lines.

Minitab does not have native ability to pivot data for analyzing multi-dimensional data. It is possible with scripting and with multiple steps in setting up the data, but it is not user friendly.

Minitab only works on imported data in memory. It only works with structured row/column data sets. It does not warehouse or store data, so all data to analyze must be obtained from somewhere else. It can connect to databases, but unlike Tableau, it does NOT operate on live databases. It only uses the database connection to import the data into memory. For larger datasets, memory can become a problem.

Like Tableau, and JMP, Minitab allows importing data from databases. This can leave vulnerabilities by exposing connections to the database. If converting unstructured data files into structured tables, that process is also vulnerable to exposure if data is handled manually at various stages.

Pros: Easy to use, excel like interface. Lots of standard reports, graphs, and statistics. A six-sigma focused tool.

Cons: Does not pivot on data easily, cannot handle unstructured test data well, possible memory issues with large data sets.

Table 4- Comparison of capabilities between Verifide and Minitab

	Verifide	Minitab
Storage and Warehousing	Data is warehoused in Dynamic Database built on MS SQL Server. Schema-free database.	Does not store data, user must specify data source. It requires structured data, so cannot deal with unstructured data. Can connect to multiple data sources and databases (structured table only), but data must be imported into memory. User must write SQL query to pull data in.
Verification and Validation	Limit checking and Yield evaluation tools included. Can post updated yield results back to database for sell-off auditing.	Does not provide for limits or hypothesis testing of data.
Test-Level Analysis/Reports	Access to raw test data and test-level reports.	None. Only does higher level data analysis.
Trending and Visualization	Tools to extract data fields across large data sets, pivot tables and charts are easy with drag and drop. Data can be visualized in graphs, histograms, pie charts, stacked bars etc. Supports 3D trending – overlaying plots from multiple tests. Customizable dashboards in next release.	Simple to use tool. It has many standard graphs and reports that are used in various industries. Does not support pivoting on data natively, but there are long workarounds to get some summaries.
Analytics and Statistics	Common six-sigma and process capability statistics. New statistics and analytics being added driven by customer needs.	Extensive statistical analysis and reports. Has large amounts of capabilities in crunching data and showing analytical plots.
Enterprise Automation	Automated reports and triggers coming in our next release.	No automated reports, but it supports macros that can save reports to image file. Custom software would be needed to automate and catalog reports.
Data Security	Verifide provides abstracted access to the database so no connections directly to the database are allowed. Users and roles are abstracted. Databases can be hardened and encrypted to high security standards.	Minitab software can connect to the database to import data. This exposes direct connections to the database. Also, data can be copy/pasted for analysis, leaving it vulnerable during handling.

Conclusion

Commercial “data analysis” tools are capable and sophisticated in providing specialized analyses of data. These tools operate on structured data and can compute numerous statistics and analytics on that data.

The problem is that TEST data is messy; it is unstructured, hierarchical, yet it is relational. To process data in these commercial tools requires considerable effort from users to gather and select data files and to convert and import test data into the structured form that is needed by these tools. Iterative and repetitive analysis makes this task very difficult.

Furthermore, data handling in test goes beyond the statistics and analytics – it requires handling at all levels of the product life-cycle from design, manufacturing, and maintenance. This includes acquisition, troubleshooting, management and warehousing, disposition/review/audit, yield optimization, and access to raw data for troubleshooting and root-cause disposition.

All these data requirements are based on the same set of test data, so a tool that only provides a solution for a single aspect of data handling does not meet the needs of test organizations.

Verifide’s software provides a comprehensive platform for managing and analyzing test data, catering to all the role players. It warehouses the data, provides visualizations, analytics and insights, but also supports the day-to-day data handling needs that bear upon schedules and cost.

Contact

For more information or to discuss anything please contact us:

Neil Baliga
neil.baliga@verifide.com
408.834.8713

Verifide Technologies, Inc.
6050 Hellyer Ave #150
San Jose, CA 95138
<http://www.verifide.com>