

LIMS software

Data analytics custom dashboards in SampleManager LIMS software

Common use case challenges and their solutions using Statistical Scripts

Statistical Scripts in the laboratory

The core Statistical Script functionality provides Thermo Scientific™ SampleManager™ LIMS software with the additional flexibility to interact with the diverse and vast amount of data it contains.

SampleManager software can now utilize R or Python through the Statistical Scripts to take advantage of functions, packages, and saved models from the most popular data science programming languages.

These new features open up the possibility for hundreds of applications, without leaving SampleManager software and whilst maintaining data governance.

Applications using Statistical Scripts functionality

The following applications can use the Statistical Script functionality:

- Performing complex calculations and plots.
- Executing custom codes or libraries created by the laboratory or the community.
- Augmenting SampleManager LIMS information with external sources.
- Management of large amounts of data for analytics.
- Laboratory automation.
- Application of statistics and machine learning algorithms.

You will find more details about these applications and what can be achieved to help you solve your business challenges in this brochure. Our Thermo Fisher Scientific team are here to support you as you create applications specific to your needs.



Typical use case challenges and solutions

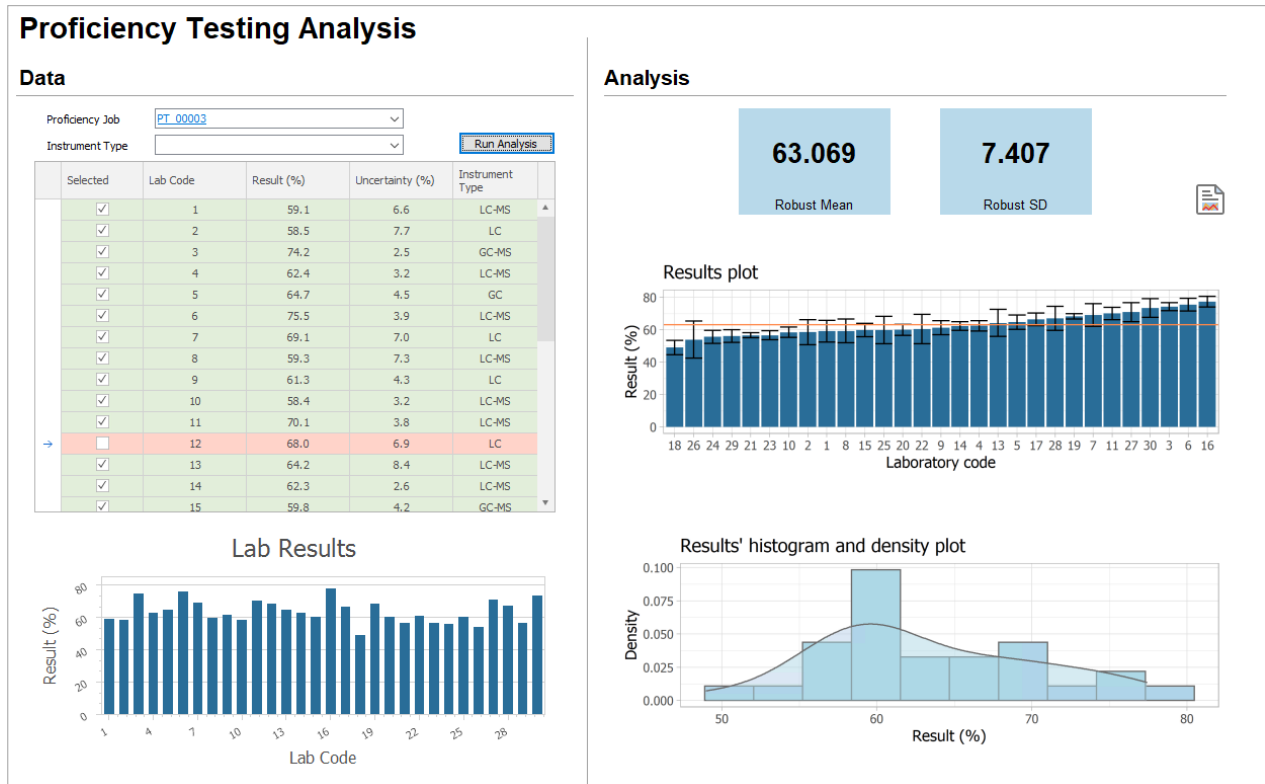
Proficiency testing

Challenge

Organizations apply Algorithm A described in ISO-13528-2015 to view and understand variances in laboratory and instrument performances. The algorithm requires complex and iterative calculations that are difficult to implement.

Solution

Our team sends a set of inputs from SampleManager LIMS software, such as “Laboratory”, “Equipment” and Sample Information to an R script, where the robust parameters are calculated. Both inputs and outputs are displayed in a friendly UI, with additional plots included to aid the decision-making process of the analyst.



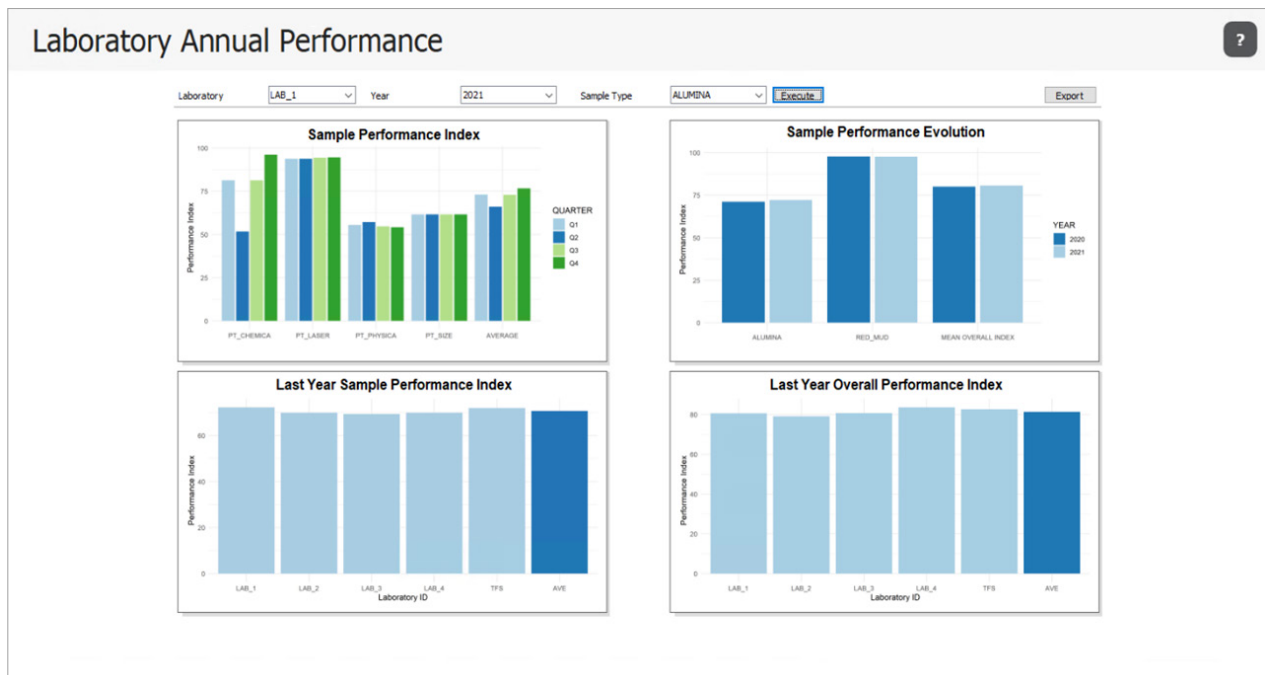
Laboratory annual performance

Challenge

Many organizations use manual steps and programs to undertake in-house annual proficiency testing calculations. Some laboratories are reliant on Microsoft Excel macros to perform statistical calculations such as z-scores, and Microsoft Word to generate reports. Most of the time, they are unable to apply algorithms across their data and create reports in a consistent, automated, and reliable manner with minimum user interaction and data transfers.

Solution

The transfer of in-house statistical calculations and management of data at various levels is possible in SampleManager LIMS software. A set of inputs such as “Laboratory”, “Year”, “Sample Type” and their results are sent to an R script where the relevant statistics are calculated together with the plots related to them. Both inputs and outputs are displayed via a friendly user interface.



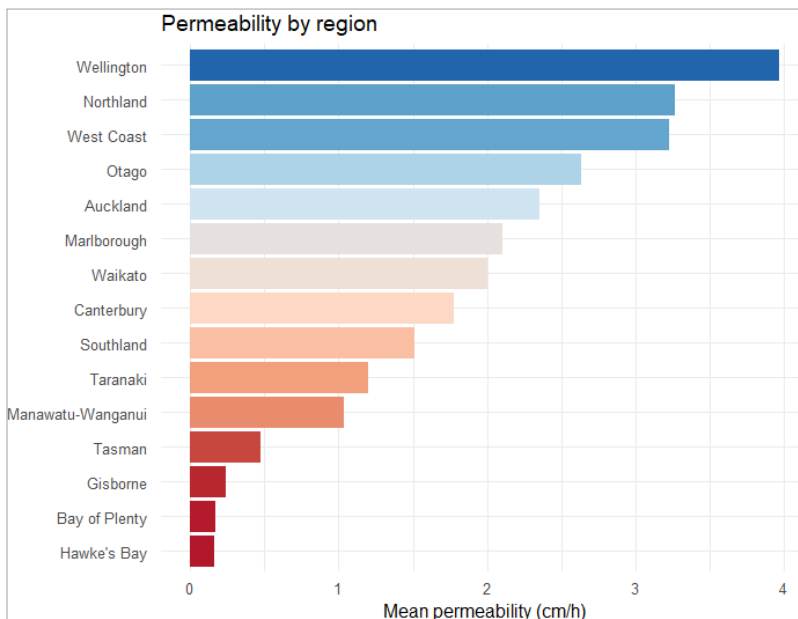
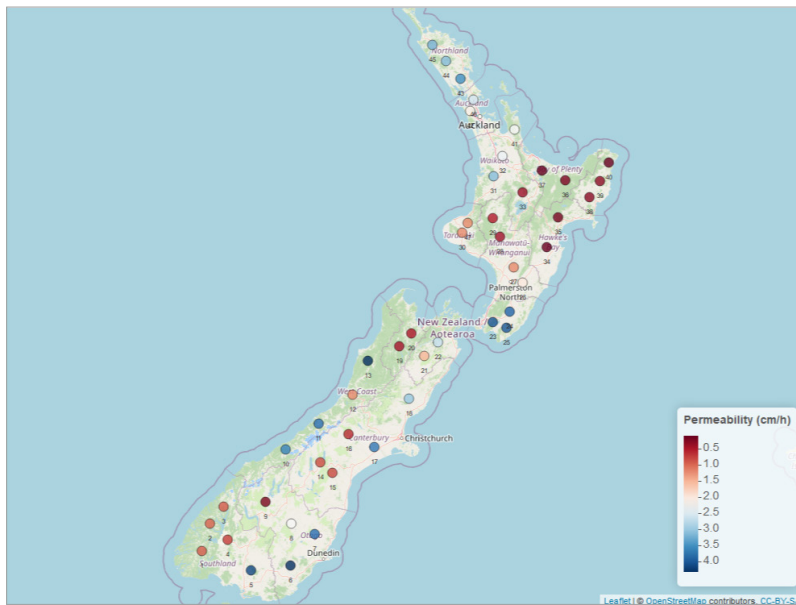
Soil geographical information

Challenge

An organization wants to generate map plots of various sampling points that will be included in reports and augment the information with geographically explicit external data sources. However, LIMS systems are often not intended to apply geographical operations on data or save geographic information files such as “shapefiles”. Specific Geographical Information Systems (GIS) software or libraries are needed to perform those calculations.

Solution

The Statistical Script capabilities in SampleManager LIMS software are triggered via workflows on jobs that will access external data (“shapefiles”) to extract information for each of the sample points, generate various maps and store the results at the sample level for further analysis.



Genomic scoring

Challenge

An organization wants to use a custom machine learning model to generate predictions on SampleManager LIMS data without leaving their LIMS software, avoiding any manual steps and human errors. However, they have disassociated data sources, processing, and results storage. Their custom software is not programmed to interact with a LIMS, and integrating such operations is time-consuming and expensive.

Solution

SampleManager LIMS software uses the Statistical Script capabilities to pick up the relevant information selected by the user to process it and execute the machine learning model. When the results are ready, the software displays the most relevant information and saves the results to the sample for further analysis. Everything is done within the SampleManager LIMS interface.

Mel38 melanoma genomic score

Sample sheet definition

Scan:
Lane:
Sample ID:
Cartridge:
Select RCC files:

Sample sheet to analyze

Scan	Lane	SampleID	RCCFilename	Cartridge	RCCFilePath
209103000002	2	2	20201002_2091026401...	20201003	C:\Windows\TEMP\lib86...
209103000003	3	3	20201002_2091026401...	20201004	C:\Windows\TEMP\lib86...
209103000004	4	4	20201002_2091026401...	20201005	C:\Windows\TEMP\lib86...
209103000005	5	5	20201002_2091026401...	20201006	C:\Windows\TEMP\lib86...

Results

Melanoma genomic score calculation

RCCFilename	sample.id	CartridgeBarc...	Sample.Mean	Sample.SD	svm_score
20201002_209...	1	209102640122	49.5085	216.3763	2.90775200040...
20201002_209...	6	209102640122	26.0223	53.5944	3.61724035152...
20201002_209...	2	209102640122	41.8200	66.5958	2.76037884626...
20201002_209...	3	209102640122	55.6960	107.6839	2.96382168648...
20201002_209...	4	209102640122	40.3306	101.0783	3.36333850855...

Diagnosis based on Mel38 score

Diagnosis	Mel38 score
1	~2.8
2	~2.8
3	~2.8
4	~3.2
5	~2.8
6	~3.5

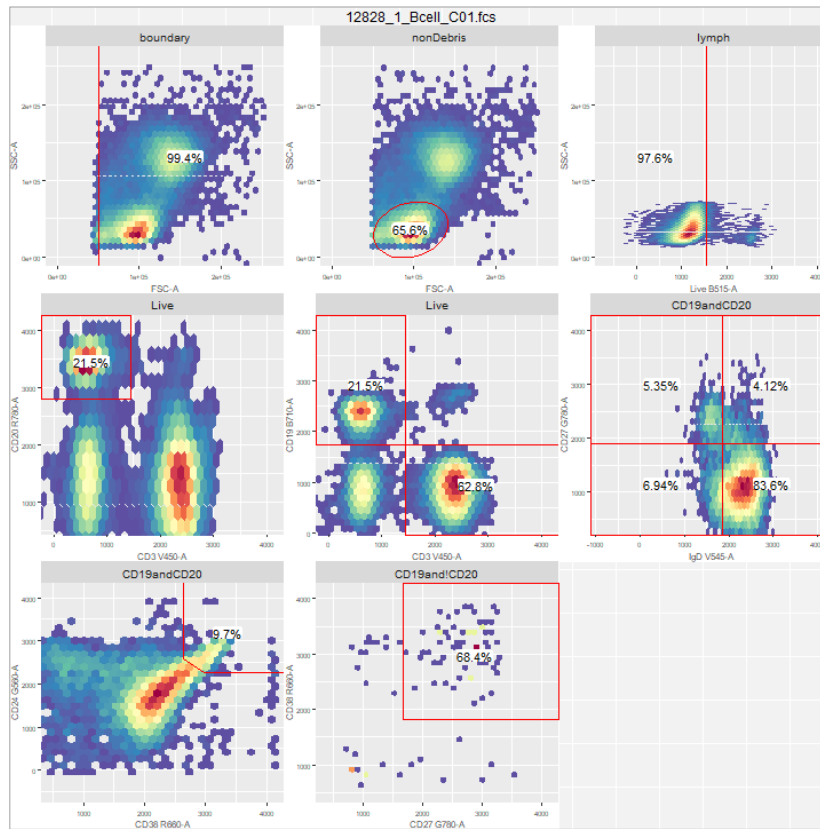
Flow cytometry

Challenge

A LIMS often does not provide the specialized calculations and plots needed by organizations to perform flow cytometry analysis from a “.wsp FlowJo” file. Such procedures can be possible with the use of specialized library repositories outside a LIMS, however, organizations struggle with the synchronization of the data management, processing, and results storing with the sample life cycle. Normally implementing a harmonious integration of such operations is time-consuming and expensive.

Solution

The Statistical Script capability in SampleManager LIMS selects the relevant information and, after processing the data, sends it to the specialized software packages accessed via Statistical Scripts. When the results are ready, SampleManager software displays the more relevant information and saves the results to the sample for further analysis.



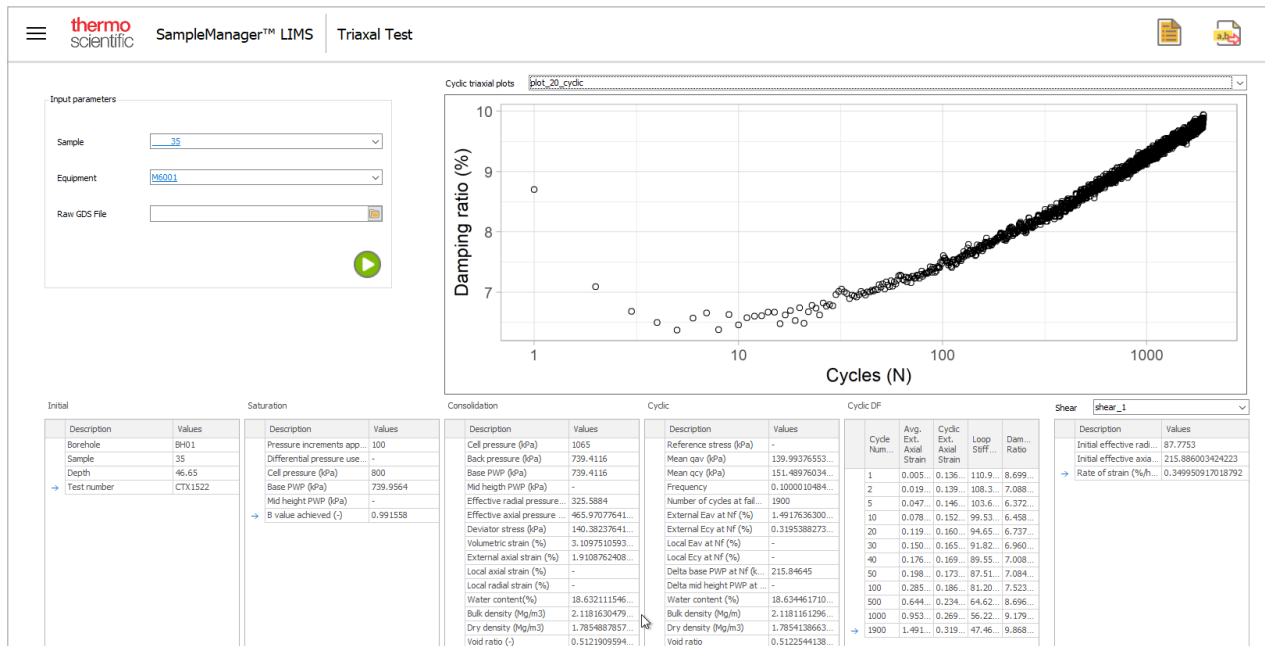
Soil analysis large data processing and automation

Challenge

Specialized calculations and visualizations of hundreds of thousands of data rows are difficult to implement in traditional LIMS infrastructure. An organization needs to transform its traditional soil analysis process that includes data retrieval, processing in large Microsoft Excel files, and storing the results to a more efficient, reliable, scalable, and traceable format without leaving the LIMS.

Solution

The Statistical Script capabilities in SampleManager LIMS software reads the relevant information selected by the user. After processing the data, it sends it to an R script that includes all the calculations in the original Excel files. When the results are ready, SampleManager software displays the most relevant information and saves the results to the sample for further analysis. The capability presents various plots and allows the user to make a report with all the relevant results and visualizations. Besides assuring traceability and augmenting reliability, the capability speeds up the process more than 20x the original duration, freeing resources and optimizing laboratory operations.



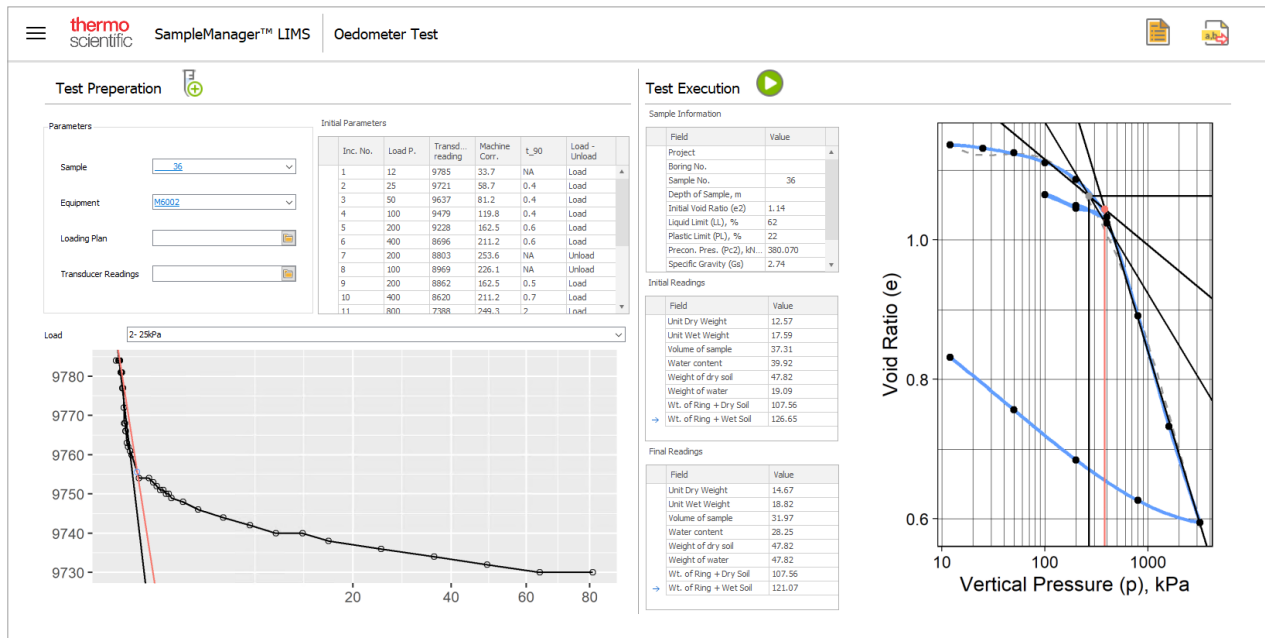
Soil analysis automation

Challenge

Companies undertaking soil analysis want to apply statistical algorithms to soil testing, to identify key points that previously required a visual inspection or manual calculation, to read and process various files attached to a sample, to create visualizations and store relevant results in a LIMS. To automate this process, complex algorithms are needed to mimic an analyst's visual inspection. The data source, processing, and results storing are often dissociated, and the integration of such operations can be time-consuming and expensive.

Solution

The Statistical Script capability in SampleManager LIMS software is used to extract the data and with the aid of open-source libraries and custom programming, automatically detects trends and other graphical cues that were previously done by the analysts. When the results are ready, the software displays the more relevant information and saves the results to the sample for further analysis.



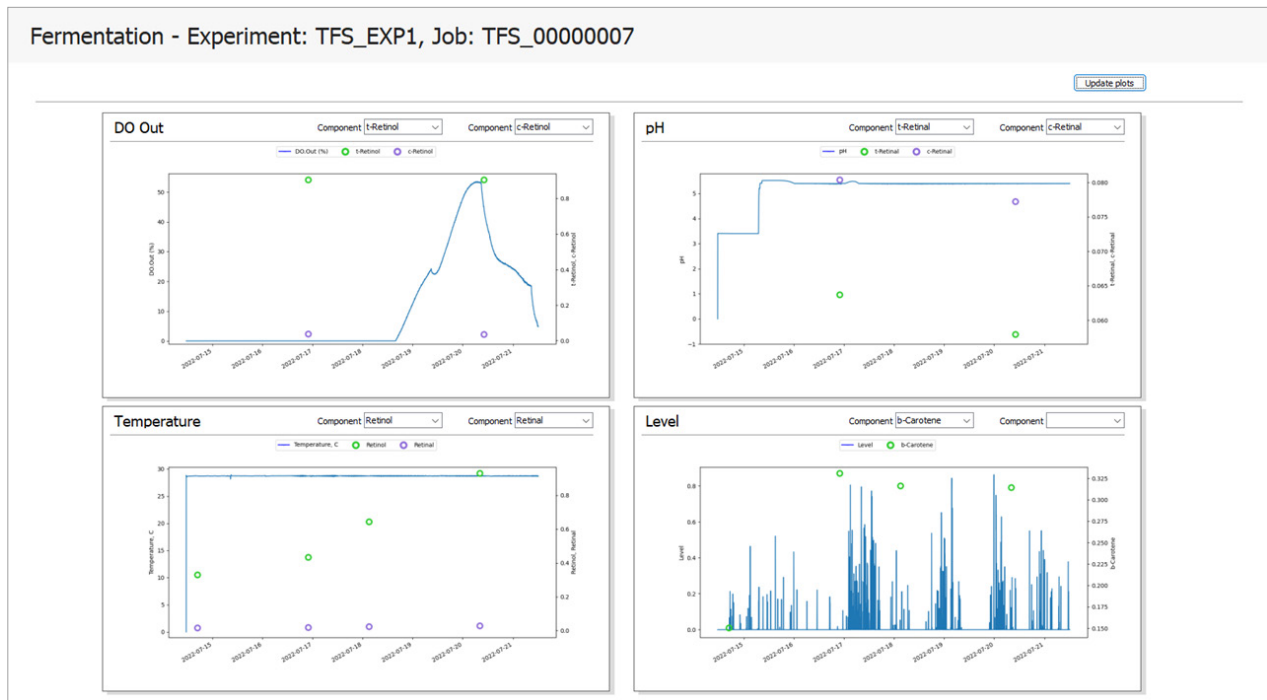
Large data visualization

Challenge

Native LIMS capabilities do not efficiently plot various data sources or large files linked to laboratory entities. Storing thousands of data points per job is often not optimal for the data infrastructure of the LIMS. Instead, companies need to be able to extract, transform, and load data from various sources for effective and customized visualizations.

Solution

The Statistical Script capability in SampleManager LIMS software reads the channel information of samples inside a job and packages the sample test results together with the user-selected configuration, sending it to a Python script that processes the information. If a user wants to plot the files natively in SampleManager software, the Python script outputs an easy-to-read CSV file that will be used to plot the information via the software's functionalities. If not, four plots are plotted as png files and sent to SampleManager software's form.



Sequence visualization

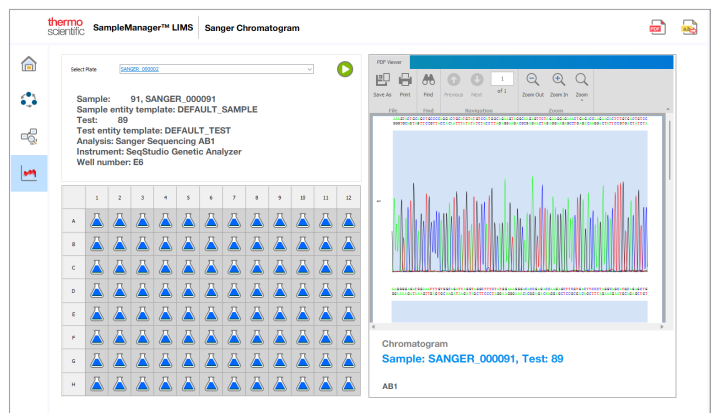
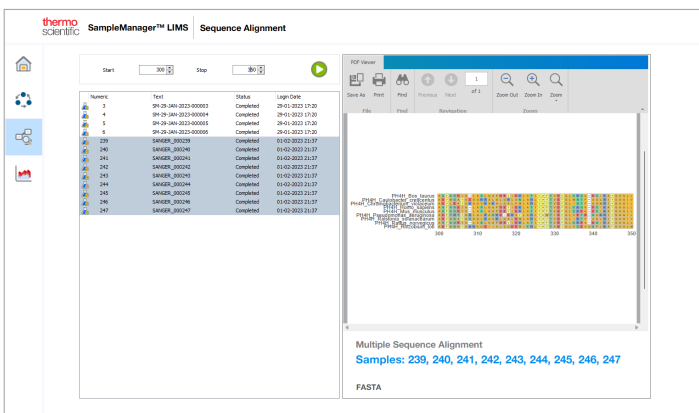
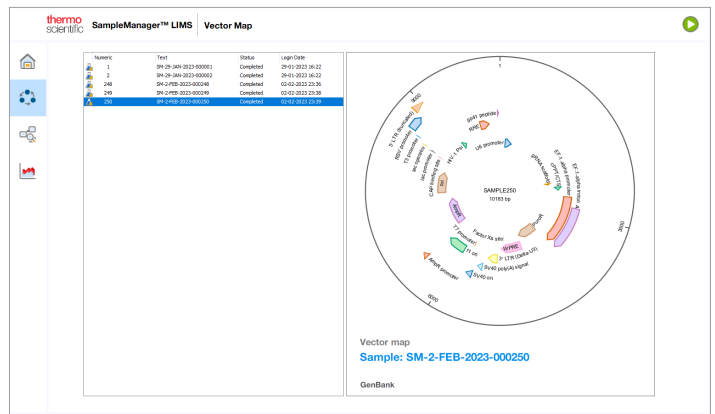
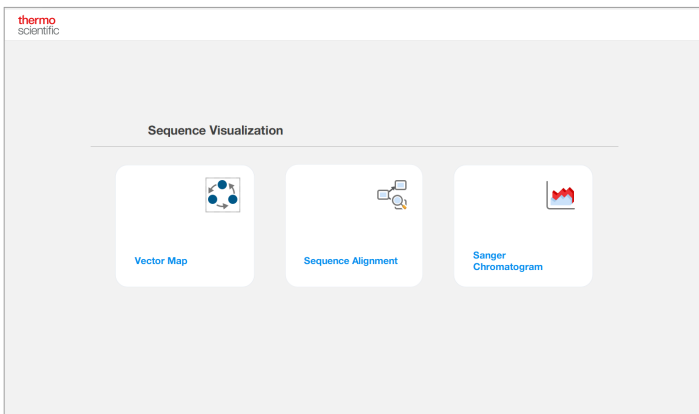
Challenge

Companies need to perform complex calculations and plots, execute custom codes or libraries created by their laboratory or community, manage large amounts of data for analytics, and extract, transform, and load data from non-conventional data sources for specialized visualizations. Nonetheless, native LIMS capabilities can fail to provide the specialized visualizations needed for sequencing visualization, making it difficult to manage various data sources and types.

Solution

The Statistical Scripts capabilities in SampleManager LIMS software can be used to receive and process relevant data associated with the samples. Specialized packages from the scientific community were accessed via R and were used to create the plots and display them on dashboards. The dashboards include:

- Visualization of a vector map from a GenBank file attached to a sample.
- Visualization of Multiple Sequence Alignment (MSA) by selecting multiple samples where each have a FASTA file attachment.
- Visualization of a Sanger sequencing Chromatogram by selecting a sample (from a plate UI), which has an AB1 file attached to it.



Data access via OData APIs

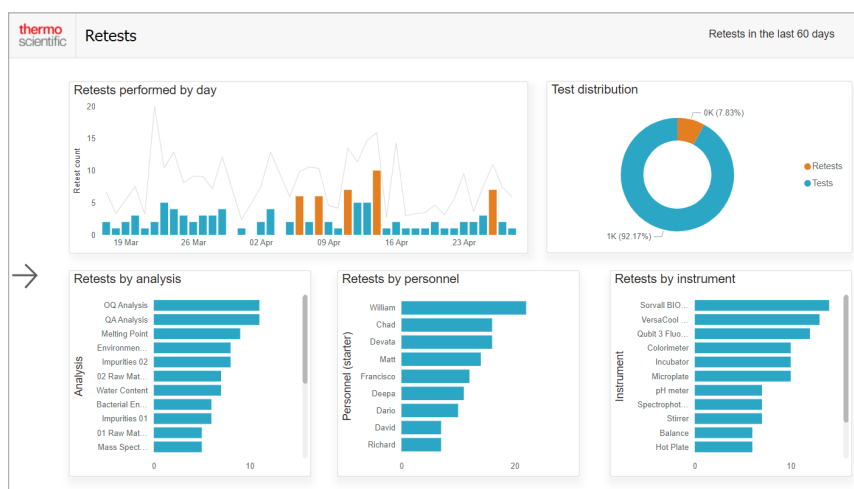
SampleManager LIMS software provides an OData API which allows business intelligence tools such as Power BI, Tableau and Spotfire to connect. The advantage of the OData API is that it maintains user access control, so users connect to the OData API through the reporting application using their SampleManager LIMS login credentials. Users only see the data they have the correct privileges to access – typically by group membership.

As with the Statistical Scripts, customers are free to setup and use the OData for their own reports which requires minimal configuration.

The Data Analytics Solution dashboards use the same approach, when creating and optimizing reports using tools such as Power BI and Tableau.

Further support can be provided by utilizing Data Analytics Service hours to help as required.

Here are two examples that show dashboards from the Data Analytics Solution as Power BI Reports, the Incidents and Retest dashboards.



Learn more at thermofisher.com/dataanalytics