Swift Stat

1st July 2025

Swift-Stat Instruction Manual

Contents

Introduction

Swift-Stat is a powerful software solution for analyzing numerical data from engineering and scientific experiments, manufacturing and process control. It offers a comprehensive suite of tools for data cleaning, processing, analysis, and visualization - combined with an intuitive user interface. The tools can be applied in parallel across your entire dataset with just a few clicks. In addition to saving time and reducing the risk of manual errors, most users find that Swift-Stat helps them to significantly improve the depth and quality of their data analysis by placing powerful tools at their fingertips.

The instruction manual provides a detailed explanation of all components of the software. Before reading the manual, we recommend that you start by watching the introductory tutorial available on the resource page of the Swift-stat website. The tutorial walks you through an example analysis, giving you a practical overview of how to use Swift-Stat and highlighting many key features. Once you've watched the tutorial, try exploring the software on your own using the example files found in the “tutorials” folder inside the main Swift-Stat directory. When you're ready to dig deeper, refer to the instruction manual for in-depth explanations. You can also access the relevant sections of the manual directly by clicking the Help button, available on most menus throughout the software. If you have a question that the instruction manual does not answer, please email us! We are constantly working to improve Swift-Stat so please let us know of any features you would like to see in the next release of Swift-Stat. Don’t forget, there is a free read-only version of Swift-Stat so that you can share your data analysis with anyone! “Numbers have an important story to tell. They rely on you to give them a voice” (quote by Stephen Few).

To maintain reasonable computation speeds, we recommend that individual files be limited to no more than 50,000-200,000 data points, depending on the complexity of your analysis and your system performance. However, this this is not an absolute limit. For suggestions on how to improve performance, see section 1.3. Swift-Stat stores numbers in spreadsheets at a precision of 15 decimal places. However, due to floating-point limitations, some calculations may not be accurate beyond about the 12th decimal place. As with any statistics software, this tool shouldn't be the sole basis for decisions where errors could lead to harm or significant loss.

The Swift-Stat graphical user interface (GUI) is built using PyQt, a Python binding for the Qt library. The Qt library was developed by The Qt Company. We utilize the Qt library under the terms of the GNU Lesser General Public License version 3 (LGPLv3), available at https://doc.qt.io/qt-6/lgpl.html. In accordance with the requirements of the LGPLv3, Swift-Stat dynamically links to the Qt library, which is included in the _internal directory of the Swift-Stat distribution package. You are free to re-link Swift-Stat against a different or modified version of the Qt library. To do so, substitute the files in the _internal folder with your own version of the Qt library. For assistance, contact us at support@swift-stat.com. The Qt source code can be downloaded from https://www.qt.io/download-dev. Swift-Stat has not made any modification to the Qt library.

1. Installation & System Requirements

1.1. Installation

1. Ensure that your computer's date is correct.

1. Download the license.txt file attached to your welcome email, to anywhere on your computer. There is no need to open the file.

2. Download Swift-Stat from the link contained in your welcome email.

3. Unzip Swift-Stat and place it in your preferred location e.g. C:\Users\<YourUsername>\ (the location is not important as long as you have full privileges in the folder).

4. Run Swift-Stat.exe from the Swift-Stat folder and, when prompted, select the license.txt file from the file browser (do not modify its name). If it fails, try manually moving the license.txt file to the swift-Stat folder (i.e. the same folder as the executable).

5. Read and agree to End user license agreement (EULA). Once that's done, congratulations, Swift-Stat is now ready to use!

If you run into any issues, try our troubleshooting tips bellow. If that fails, send us an email explaining the problem in sufficient detail.

Troubleshooting

- If you run into problems unzipping the zip file, download 7zip and try again or send us an email to request a non-compressed version of Swift-Stat. The non-compressed version will take longer to download but is otherwise identical.

- If you run into problems importing the license file, ensure that the license file is named license.txt. If the name is correct and you are still having issues, manual copy and license file into Swift-Stat folder (into the same folder as the executable).

- If you run into problems opening Swift-stat.exe, ensure you have placed the folder in an unrestricted location such as C:\Users\<YourUsername>\.

- If Swift-Stat.exe is blocked by your firewall, open your firewall settings, navigate to Program Permissions/ Firewall Rules and add Swift-Stat.exe to the allowed files. Swift-Stat does not require access to the internet.

1.2. System Requirements

The Recommended system settings are shown in the table below.

	Recommended system requirements
Operating System	Windows 10 or 11
CPU	x86-64 2.8 GHz or greater is recommended
RAM	8 GB or greater
Monitor	Optimized for 1920x1080 Compatible with other 16:9 or 16:10 aspect ratio resolutions

1.3. Improving Computation Time

To maintain fast computation speeds, we recommend that individual files are limited to no more than 50,000-200,000 data points, depending on the complexity of your analysis and your system performance. The following tips are intended to help you maximize performance.

- If Swift-Stat has been open for many hours or days, restart the software!
- Avoid leaving lots of empty cells in your spreadsheet by deleting unused sheets and using the crop tool (section 4.1) to remove empty rows and columns.
- Do not enable auto-averaging (section 5.2)!
- Swift-Stat makes it easy to batch process many variables at once. Don't select every variable for an analysis just because it's easy to do. Only select the ones you need!
- Delete analyses you no longer need.
- Data cleaning operations are computationally heavy. Once you have finished cleaning your data, consider flattening your data, which permanently replaces your raw data with the cleaned data and deletes the data cleaning analyses (section 4.3).
- If you don't need your row calculations to be linked (so that they automatically update when the input data is modified), unlink them in the row calculation menu.
- If you have captured data at an excessively high temporal resolution, down sample your data to reduce the quantity of data (section 4.2).
- Ensure your computer is not in power saving mode, as this could limit CPU performance. Plugging the laptop into its charger can often improve performance.
- Use a high frequency CPU such as Intel i7-13700K.

2. File Operations

2.1. Creating, saving and opening files

New File

Navigate to File > New. Select a name and location for the in the dialog window, then click save to create the new file. If a file is already open, you will be prompted to either save or discard changes before continuing. After creating a new file, you will be prompted to add it to a project. Projects are an optional feature to help you manage your files, see section 2.3 for more information about projects.

Saving a file

To save a file, either navigate to File > Save or use the keyboard shortcut ctrl + s. To save a file under a new name or location, navigate to File > Save As.

Opening a previously saved file

Navigate to File > Open and either select a recently saved file from the menu or click ‘browse computer’ to open the file browser dialogue. If an unsaved file is already open, you will be prompted to either save or discard changes to the open file before continuing.

2.2. Importing and exporting data

Import Data

Data can be imported from either Excel or CSV files. The data layout must match the Swift-Stat layout, as explained in the bullet points below. The tutorials folder inside the Swift-Stat folder contains example CSV and Excel data files.

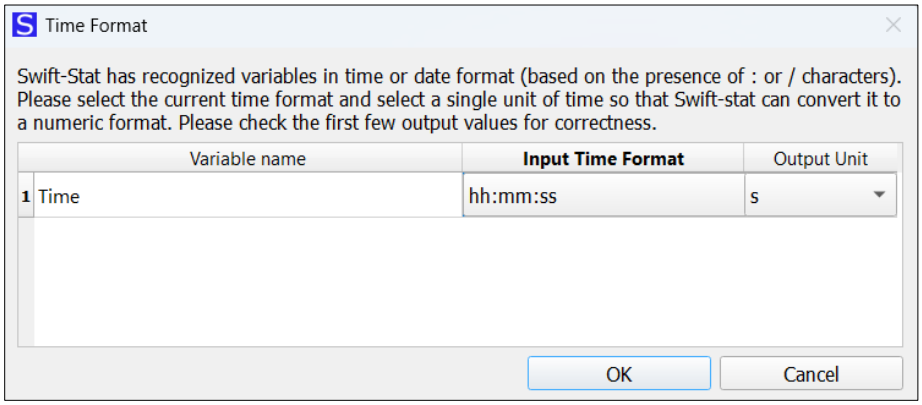
- Each column should contain the variable name on the top row, followed by the data.
- Variable names must be unique and are not case sensitive.
- If importing multiple CSV files or an Excel file with multiple sheets, the top row of each file or sheet must be identical.
- If the Swift-stat file already contains data, the variable names and their order in Excel or CSV file must match those in the Swift-stat file.

Excel: Navigate to File > Import Data > Excel and select an excel file from the file dialogue window. All sheets will be imported.

CSV: Navigate to File > Import Data > CSV and select one or more CSV files from the file dialogue window.

If your CSV or Excel files include data in time or date format, you will be prompted to define the input time format from the CSV or Excel file and the output unit to which the time will be converted, for example seconds or minutes. Swift-Stat uses this to convert the date or time into a numeric format.

The window for selecting time/date format and output unit



Exporting Data

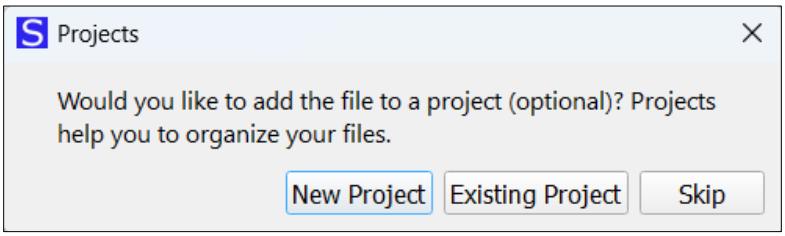
Spreadsheets can be exported to CSV or Excel format and most graphs can be exported as JPEG images either as a file or directly into an open Microsoft Word or PowerPoint document. To export the main spreadsheet, navigate to File > Export Spreadsheet and select either Excel or CSV. Select a name and location in the save dialogue and press save. Most graphs have export buttons located below or to the right of the graph.

2.3. Projects

Projects are an optional file management feature to help keep files organized. A project is simply a directory of files. You can open a project and browse or open files from within the project. When you create a new file, you will be prompted to add the file to a new or existing project, or to skip if you do not want to make use of projects. If you choose to add the file to an existing project, you will be prompted to select an existing project. Likewise, if you choose to create a new project, you will be prompted to choose a location and enter a name for the new project. A few tips about adding files to projects:

- Files and projects need not be in the same folder.
- A single file can be added to more than one project.
- You cannot add two files with the same name to a project.
- If you delete a file from a project, it will not delete the file from your hard drive.
- If you skipped adding the file to a project, you can add it later by navigating to File > Project > Add File to Project.

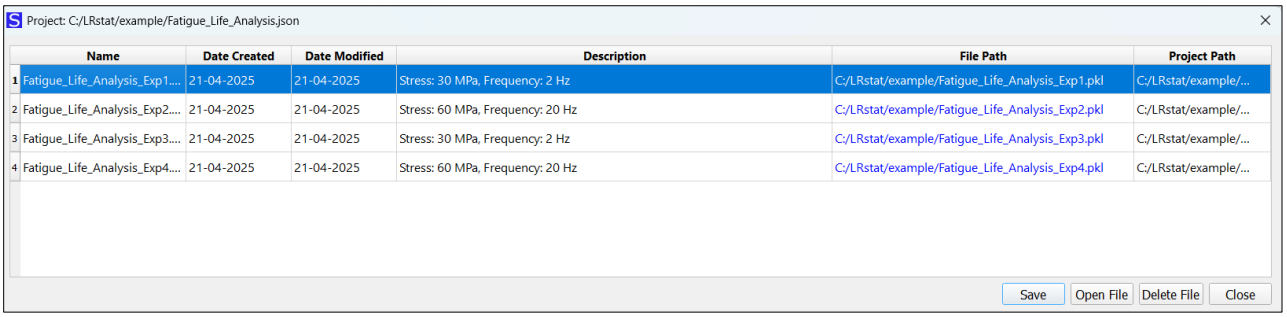
The prompt for adding the file to a project



Opening Projects

Open the project to which the current open file is assigned to by navigating to File > Project > Current Project Details. Open any other project by navigating to File > Project > Other project. You can view the files in the project from the project window and add a description for each file. To open a file or to disassociate it from the project, select the file and click either open or delete respectively.

Project viewing window

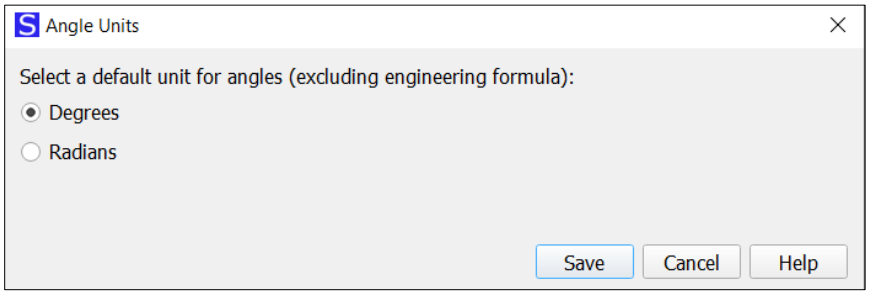


3. Settings

3.1. Changing Angle Units

Trigonometric and inverse trigonometric functions can be calculated using either degrees or radians, with the default being degrees. To change the angle units, open the 'Angle Units' menu (Settings > Angle Units) and choose either degrees or radians as the unit. The selected unit is applied both as the input unit for trigonometrical calculations and the output unit for inverse trigonometrical calculations. The selected unit is applied to in-cell calculations in the spreadsheet and scribble sheet as well as to trigonometric and inverse trigonometric row calculations, including those used in custom expressions. It is not applied to engineering formula row calculations because they inherit the units assigned to their input variables in the units menu (3.8).

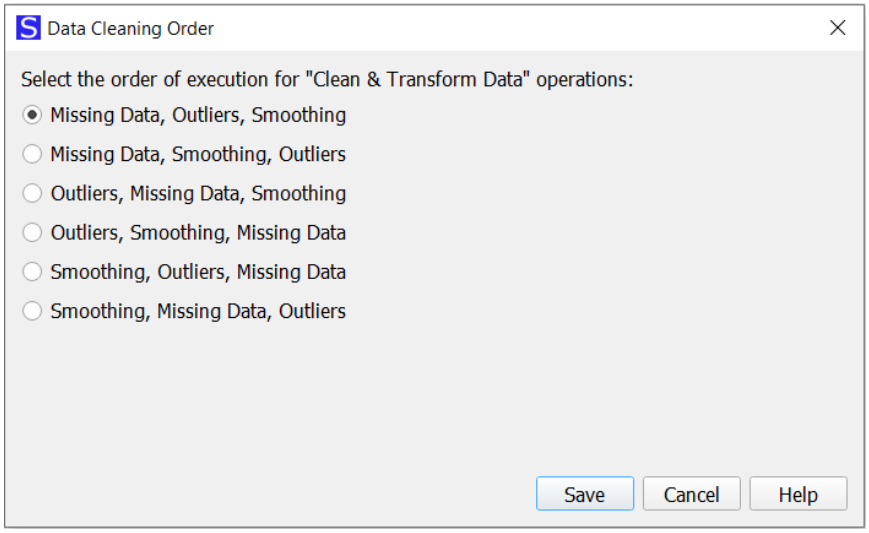
Angle Units menu



3.2. Data Cleaning Order

When applying multiple ‘Clean & Transform Data’ analyses, they are applied by default in the order "Missing Data", "Outliers" and "Smoothing & Scaling". To change their order, open the ‘Data Cleaning Order’ menu from the tools menu (Tools > Data Cleaning Order), select the desired order and press ok. If you have multiple analyses within either ‘Missing Data’, ‘Outliers’ or ‘Smoothing & Scaling’, they are performed in the order in which they were created.

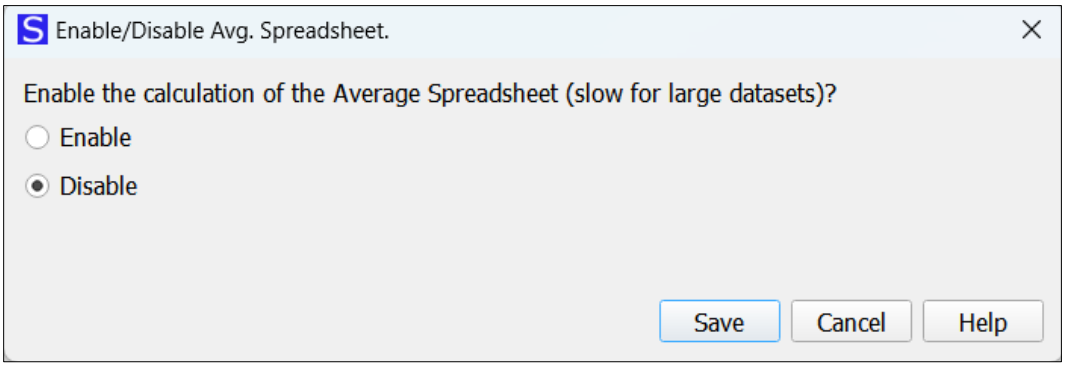
Data Cleaning Order menu



3.3. Enable or disable auto averaging

Open the Enable/Disable Auto Avg. menu from the settings (Settings > Enable/Disable Auto Avg.). If auto averaging is enabled, Swift-Stat automatically calculates the average of all sheets and if a 'multi-experiment' has been set up, the average for each experiment. The average sheets can be viewed by navigating to 'Avg. Spreadsheet' in the main menu located on the left. Auto averaging is disabled by default because it is computationally slow and reduces the overall speed of the software. To enable auto average, select enable in the 'Enable/Disable Auto Avg.' menu. It is not necessary to enable averaging in order to plot or carry out calculations on the averages.

Enable/Disable Auto Avg. menu



3.4. Formatting Decimal places

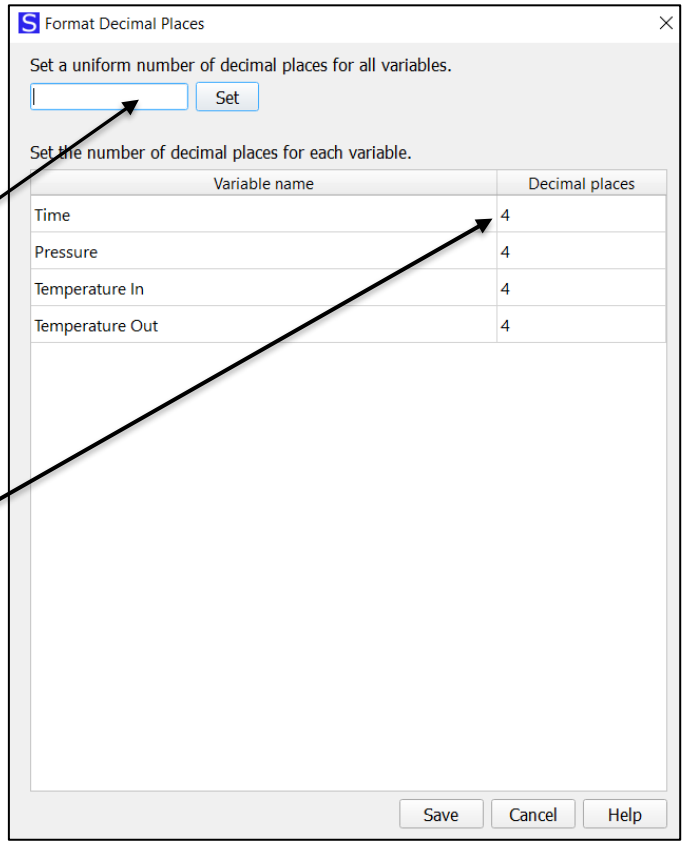
To modify the rounding of numbers, open the 'Format Decimal Places' menu from the settings (Settings > Format Decimal Places). By default, all data is rounded to 4 decimal places. To modify the number of decimal places to which a variable is rounded, locate the variable in the table on the 'Format Decimal Places' menu, enter the number of decimal places and press save. The text box at the top of the menu is a shortcut for quickly filling the table with a uniform number of decimal places for all variables. Data can be rounded to up to 15 decimal places, the maximum stored by Swift-Stat. You can view the unrounded values in the spreadsheet by double-clicking the cell containing the number. All calculations and analyses are performed using the unrounded values.

Format Decimal places menu

Enter a number and click 'Set' to assign the same number of decimal places to all variables

Or

Individually assign a number to each variable in the table



3.5. Multi-Experiment

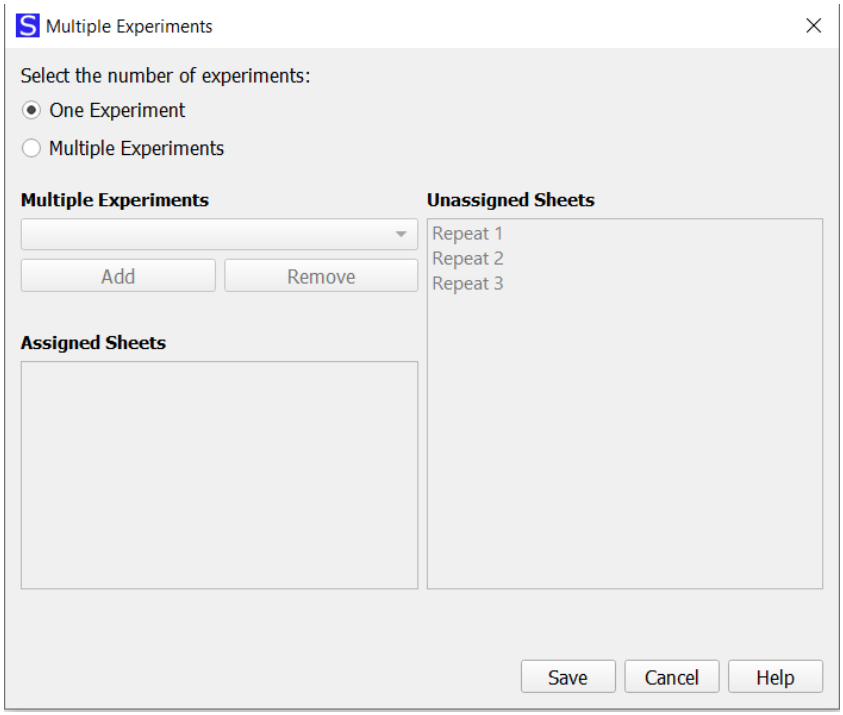
'Multi-experiment' is designed for comparing data between multiple configurations of the same experiment, where there is more than one repeat per configuration. Typically, the configurations differ by a component being swapped or a parameter being adjusted. Multi-experiment enables:

- Scatter plots of the average of each configuration, including additional data such as standard deviations and confidence intervals.
- Histograms and 3D scatter plots of the average of each configuration.
- Hypothesis tests for the null hypothesis that a variable is identical across multiple configurations of the experiment.
- Correlation studies on the averages of the configurations.
- Cross-correlation studies on the averages of the configurations.
- FFT frequency analysis studies on the averages of the configurations.

Setting up a 'multi-experiment'

1. Open the ‘Multi-experiment’ menu from the settings (Settings > multi-experiment).
2. Select to enable multi-experiment.
3. Click on the add button.
4. A popup will open, enter a name for the configuration and press OK.
5. Drag sheets for the configuration from unassigned sheets to assigned sheets.
6. Repeat for additional configurations and then press save.

Multi-Experiments menu



With 'Multi-experiment' now setup, analyses comparing experiments can be performed through the analysis builders as normal. When plotting graphs, each configuration average will now appear in the data selection panel and, where relevant, the experiment averages can be used for calculating additional statistics for the graphs. When creating a summary table or performing a hypothesis test the experiments can be selected from the 'select experiment' panel. If performing a correlation, cross-correlation or FFT analysis, select either 'averages' or 'both' from the data selection dropdown list. More detailed instructions are available from the help file for each analysis type.

Note that for multi-experiments to be useful, at least some of the same variables must exist among multiple configurations of the experiment. Swift-Stat still forces uniform variable names across all sheets even if a multi-experiment is set up. If a particular variable is not applicable to one of your configurations, simply leave the column empty in the sheets belonging to that configuration.

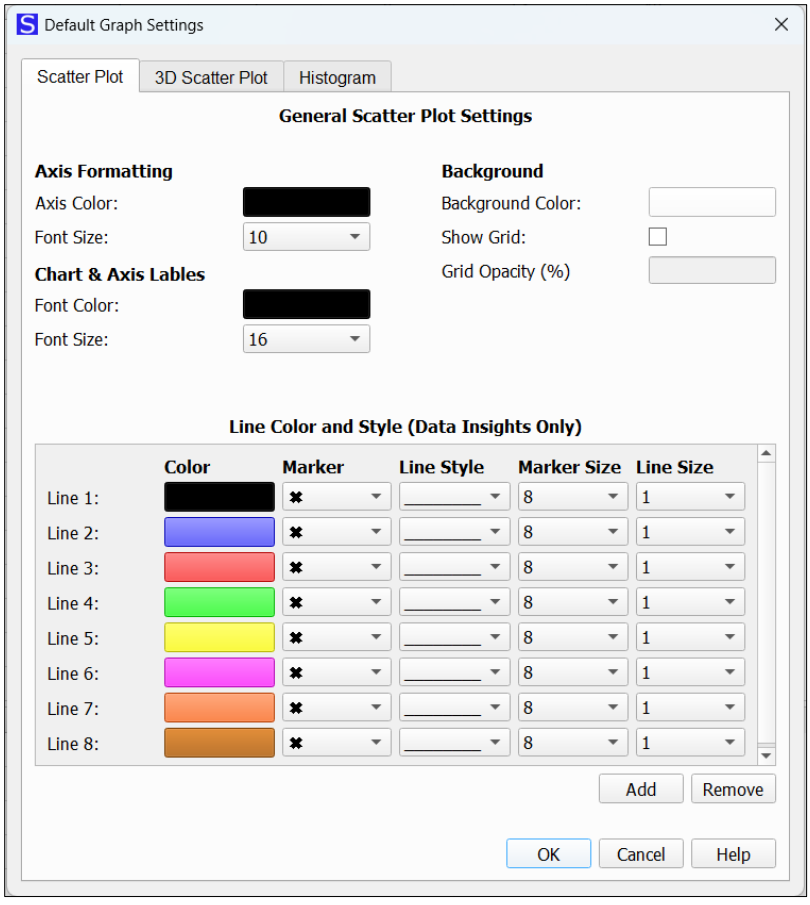
3.6. Graph Settings

The ‘Graph Settings’ menu (Settings>Graph Settings) allows you to modify the default layout for graphs. Changes made here only affect new graphs; existing graphs remain unchanged. Graph settings are specific to each file, so when a new file is created, it begins with the original default graph settings. The ‘Graph Settings’ menu contains three tabs for customizing settings for scatter plots, 3D scatter plots, and histograms.

Scatter plots

The scatter plot graph settings menu is shown below with the default settings selected.

Default scatter plot settings menu



Axis Formatting: Controls the color of the axes, the legend border (and grid, if enabled) and the font size of the axis tick marks and legend text.

Chart & Axis Labels: Controls the color of the axis tick marks, axis label text and legend text, as well as the size of the axis label text.

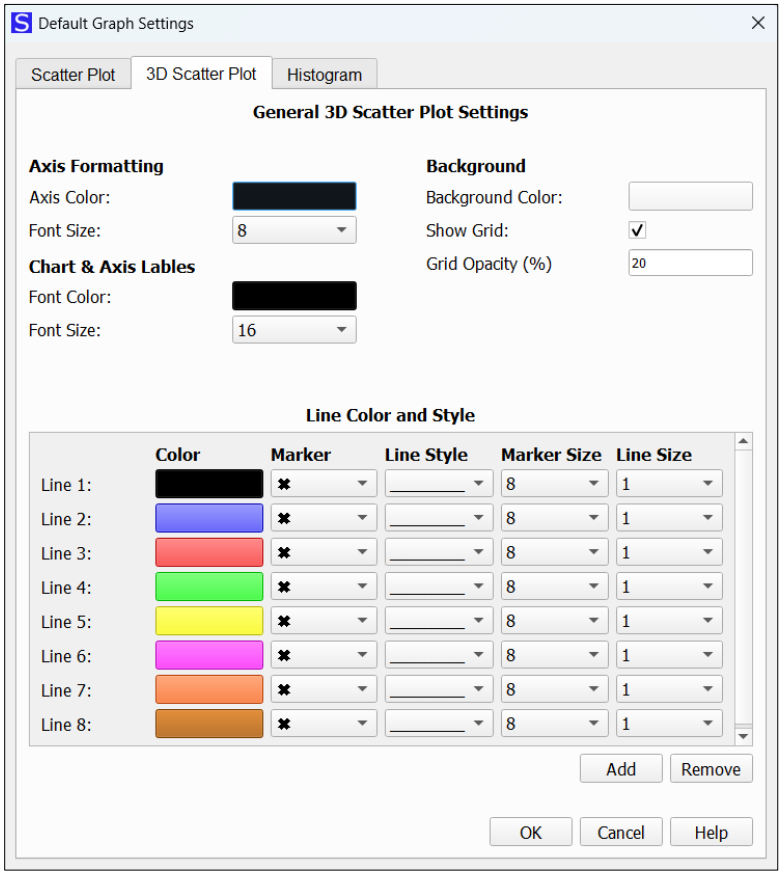
Background: Controls the background color and the grid opacity.

Line color and style: Select the plot and marker styles. Use the add button to show additional lines and the remove button to hide them.

3D Scatter plots

The 3D scatter plot graph settings menu is shown below with the default settings selected.

Default 3D scatter plot settings menu



Axis Formatting: Controls the color of the axes, the legend border (and grid, if enabled) and the font size of the axis tick marks and legend text.

Chart & Axis Labels: Controls the color of the axis tick marks, axis label text and legend text, as well as the size of the axis label text.

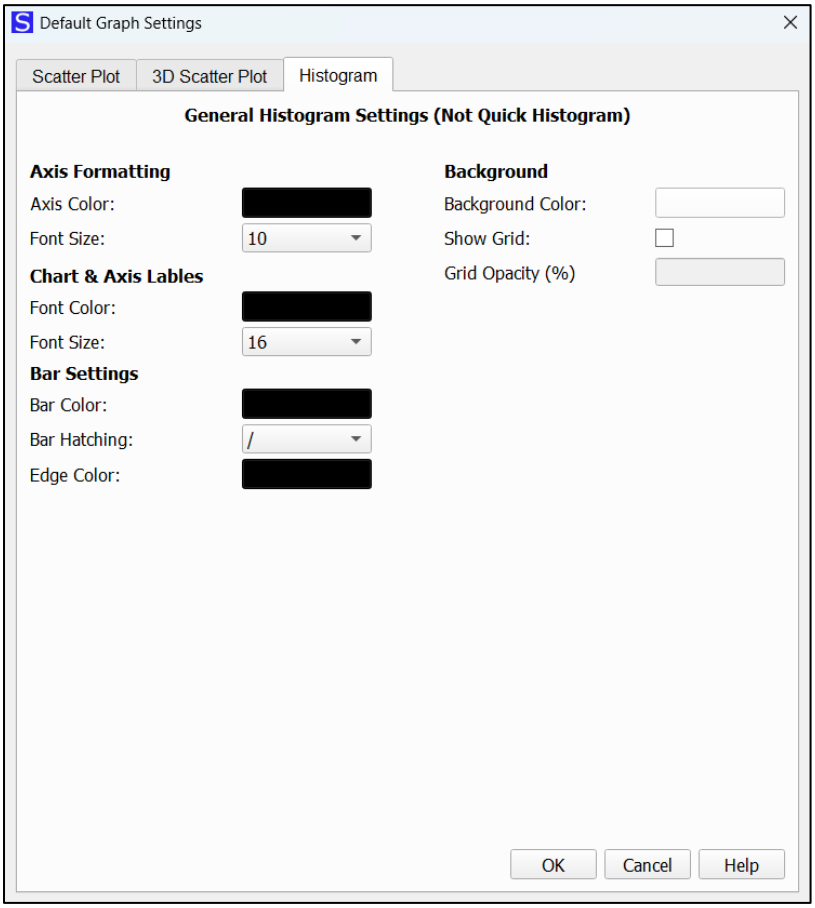
Background: Controls the background color and the grid opacity.

Line color and style: Select the plot and marker styles. Use can use the add button to show additional lines and the remove button to hide them.

Histograms

The histogram settings menu is shown below with the default settings selected.

Default histogram settings menu



Axis Formatting: Controls the color of the graph border (and the grid, if enabled) and the font of the axis tick marks.

Chart & Axis Labels: Controls the color of the axis tick marks and axis label text as well as the size of the axis label text.

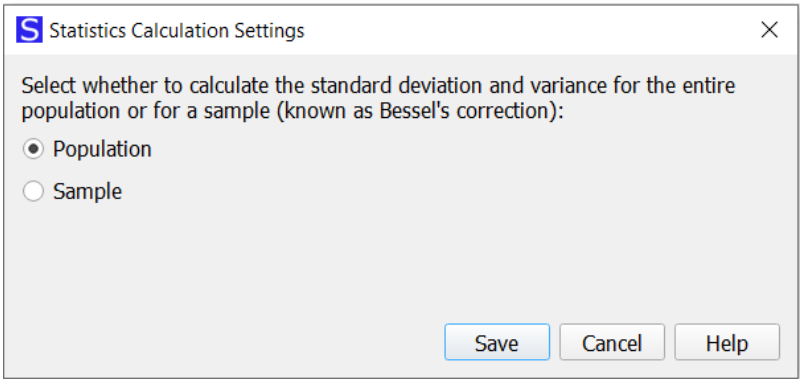
Bar Settings: Controls the border color, hatch marks and fill color of the histogram bars.

Background: Controls the background color and the grid opacity.

3.7. Statistics Settings

The ‘Statistics Settings’ menu (Settings > Units) allows you to choose if data is assumed to be a sample or a population for the purpose of calculating standard deviation and variance. If 'sample' is selected, then the data length n in subsisted with n-1, known as Bessel's correction factor. The selection is applied to all instances where standard deviation or variance are calculated, including instances where they are part of a large formula, e.g. for calculating confidence intervals. Note that the default selection is population.

Statistics Settings menu



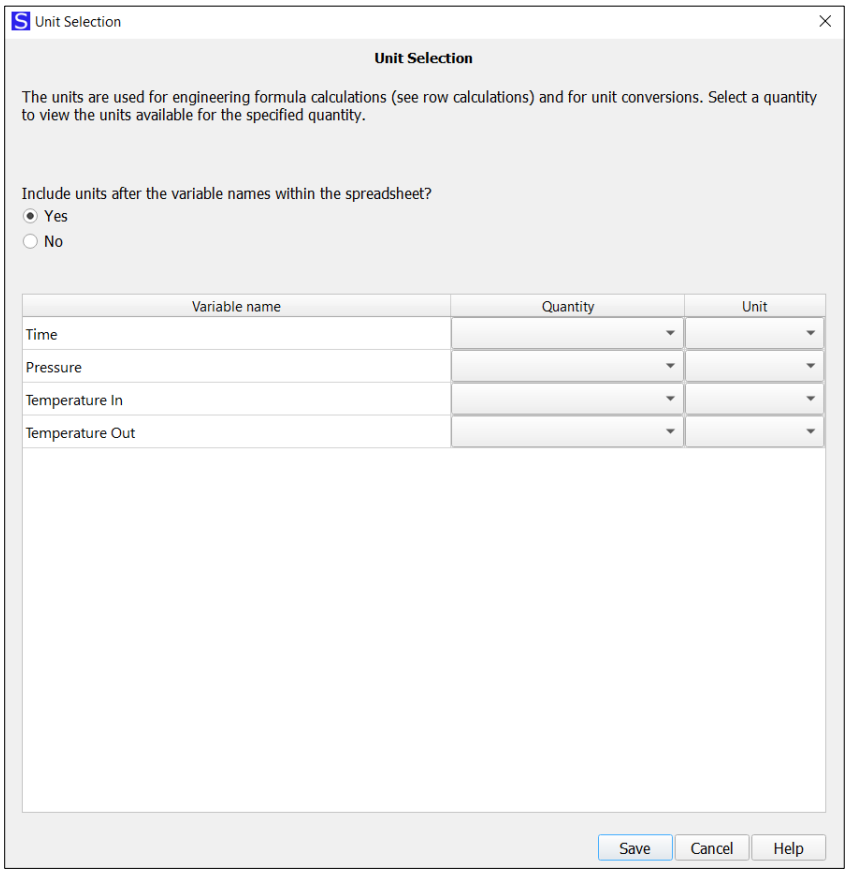
3.8. Units

Setting units for your variables enables you to convert your data between different units and to perform calculations using the built-in engineering formula (see row calculations). For help with converting units or with using engineering formula, refer to the relevant help sections.

Selecting Units

1. In edit mode, open the 'Unit Selection' menu from the settings (Settings > Units).
2. Locate the variable in the table and select a quantity in the 'Quantity' dropdown list of the same row.
3. The selection of the quantity causes the 'Unit' dropdown list to be updated with the relevant units for the chosen quantity.
4. Select a unit from the 'Unit' dropdown list.
5. Select whether to automatically include the unit names in square brackets at the end of the variable names in the spreadsheet. If selected, the units will also be included in graph legends and default graph axis titles.

Unit Selection menu



After defining units, you can convert your data into different units by right clicking inside the desired column in the spreadsheet and selecting 'Convert Units' (see section 5.1.5). To Perform a calculation with an engineering formula, open the row calculations dialogue, select 'Engineering Formula' and follow the instructions.

4. Tools

4.1. Crop Sheets

Open the ‘Crop Sheets’ menu from the tools menu (Tools > Crop Sheets). Applying 'Crop Sheets' to the spreadsheet removes empty rows and columns, improving computation speed for large datasets. Columns are only removed if they are empty on all sheets. Empty rows are removed individually from each sheet in which they are empty. The independent variable column and the row containing the variables names will not be deleted even if they are empty. Cropping can be reversed by pressing undo twice (once for rows and the second time for columns). Rows and columns can be reinserted after cropping by right clicking in the desired spreadsheet location and selecting either insert row or insert column.

Example data prior to cropping

	1	2	3	4	5	6	7	8	9	10
1	time (min)	temperature 1	temperature 2			temperature3				
2	0.0	78.1522	45.1798			42.3275				
3	1.0	75.2716	48.491			42.6844				
4	2.0	77.1323	51.2289			40.5988				
5	3.0	77.5279	57.1096			44.8172				
6	4.0	77.912	58.6838			45.1699				
7	5.0	76.7009	62.8022			42.5878				
8	6.0	77.0508	66.1707			44.1156				
9	7.0	76.0833	72.5904			41.5289				
10	8.0	76.4895	76.0406			43.9632				
11	9.0	77.1959	79.6724			42.6323				
12	10.0	77.6045	83.8264			43.1184				
13	11.0	76.4615	85.2391			45.7756				
14	12.0	78.3844	90.8651			42.4264				
15	13.0	77.0742	97.5491			43.9891				
16	14.0	77.9032	99.8855			41.4529				
17										
18										
19										
20										
21										

Example data after cropping

	1	2	3	6
1	time (min)	temperature 1	temperature 2	temperature3
2	0.0	78.1522	45.1798	42.3275
3	1.0	75.2716	48.491	42.6844
4	2.0	77.1323	51.2289	40.5988
5	3.0	77.5279	57.1096	44.8172
6	4.0	77.912	58.6838	45.1699
7	5.0	76.7009	62.8022	42.5878
8	6.0	77.0508	66.1707	44.1156
9	7.0	76.0833	72.5904	41.5289
10	8.0	76.4895	76.0406	43.9632
11	9.0	77.1959	79.6724	42.6323
12	10.0	77.6045	83.8264	43.1184
13	11.0	76.4615	85.2391	45.7756
14	12.0	78.3844	90.8651	42.4264
15	13.0	77.0742	97.5491	43.9891
16	14.0	77.9032	99.8855	41.4529

4.2. Down Sampling

Open the 'Down Sampling' menu from the tools menu (Tools > Down Sampling). Applying down sampling reduces your data's sample rate by keeping every n^{th} row of data and discarding the remaining rows. To apply down sampling, enter an integer value of 2 or more for n in the 'Down Sampling' menu and press ok. Down sampling improves computation speed for large datasets and can declutter graphs. It is best suited for data collected at excessively high sample rates. Down sampling can only be reversed with the undo button. Once the undo button is reset (e.g. by updating the project), the discarded data will be permanently lost.

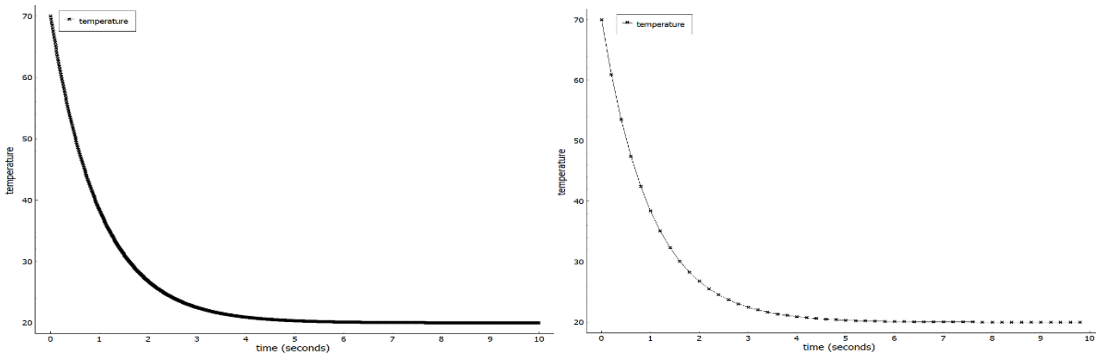
Note: Swift-Stat does not check if rows are empty during down sampling and so if empty rows are not removed prior to down sampling, the sampled data may still include the empty rows. Unlink any sequence or random distribution row calculations before down sampling.

Example data before (left) and after (right) down sampling

	1	2
1	time (seconds)	temperature
2	0.0	70.0
3	0.01	69.5025
4	0.02	69.0099
5	0.03	68.5223
6	0.04	68.0395
7	0.05	67.5615
8	0.06	67.0882
9	0.07	66.6197
10	0.08	66.1558
11	0.09	65.6966
12	0.1	65.2419
13	0.11	64.7917
14	0.12	64.346
15	0.13	63.9048
16	0.14	63.4679
17	0.15	63.0354
18	0.16	62.6072
19	0.17	62.1832
20	0.18	61.7635
21	0.19	61.348

	1	2
1	time (seconds)	temperature
2	0.0	70.0
3	0.2	60.9365
4	0.4	53.516
5	0.6	47.4406
6	0.8	42.4664
7	1.0	38.394
8	1.2	35.0597
9	1.4	32.3298
10	1.6	30.0948
11	1.8	28.2649
12	2.0	26.7668
13	2.2	25.5402
14	2.4	24.5359
15	2.6	23.7137
16	2.8	23.0405
17	3.0	22.4894
18	3.2	22.0381
19	3.4	21.6687
20	3.6	21.3662
21	3.8	21.1185

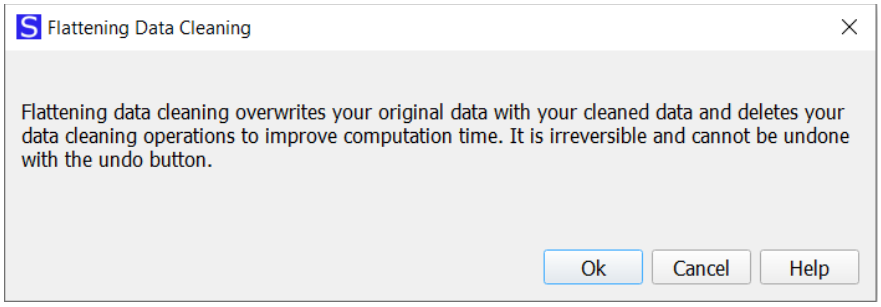
Graphs before (left) and after (right) down sampling



4.3. Flatten Data Cleaning

Open the Flatten Data Cleaning menu from the tools menu (Tools > Flatten Data Cleaning). Applying 'Flatten Data Cleaning' permanently overwrites your original date with the output of your 'Clean & Transform Data' analyses and deletes all 'Clean & Transform Data' analyses from the main menu. The 'Clean & Transform Data' analyses are among the most computationally heavy parts of Swift-Stat and so flattening the data significantly improves computation speed for large datasets. Flattening the data is irreversible and so it is advisable to save your file under a new name using 'save as' before flattening.

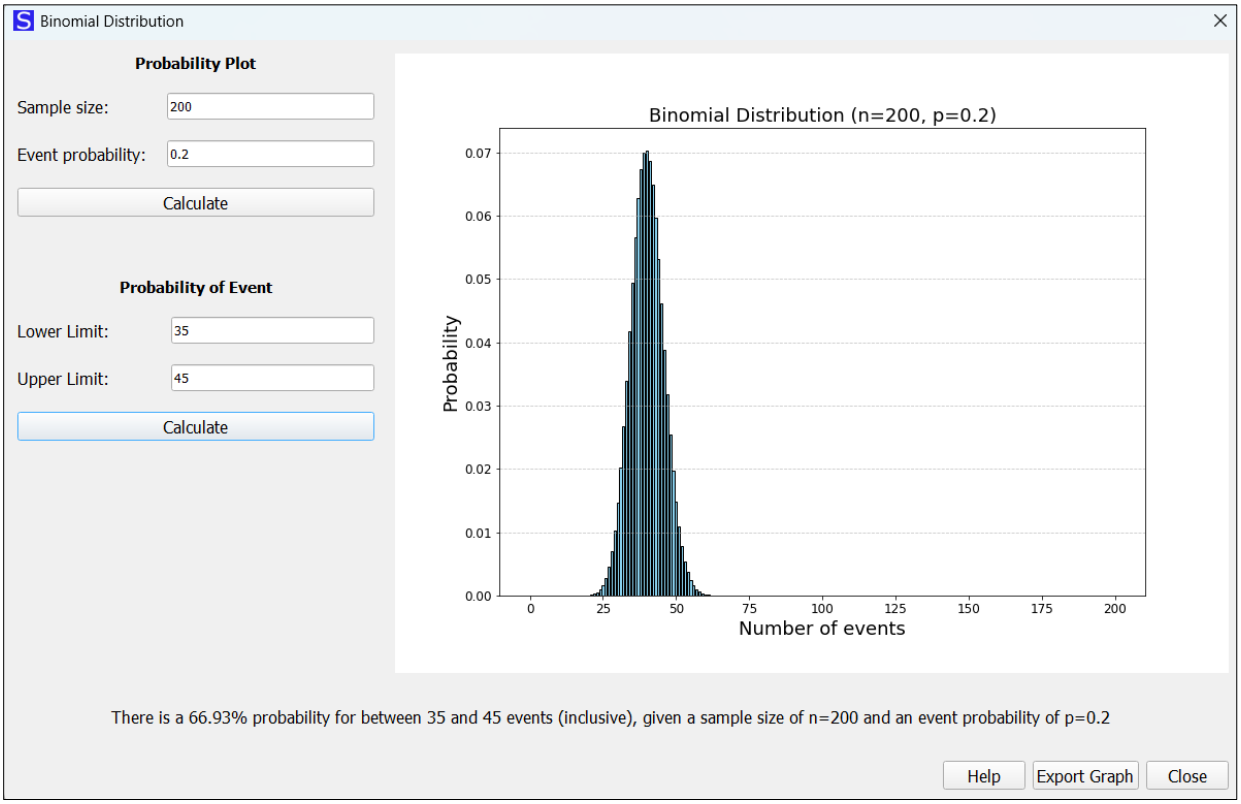
Flatten Data Cleaning menu



4.4. Binomial Probability

The Binomial probability utility (Tools > Binomial probability) enables you to plot probability plots based on the sample size and probability of an event occurring. It is used for events with two possible outcomes. You can then calculate the probability of the number of events falling within a defined range. This can be useful when performing tests with limited sample sizes as it provides an indication of the range of results that could be expected if the test were to be repeated. The tool is particularly useful for planning quality control and V&V testing.

Binomial probability menu



Sample size: The total number of tests performed

Event probability: The probability of an event occurring (between 0 and 1), i.e. the number of times it occurred divided by the sample size.

Lower limit: The lower bound for the range of events to be used in calculating the probability of the number of events falling within a defined range.

Upper limit: The upper bound for the range of events to be used in calculating the probability of the number of events falling within a defined range.

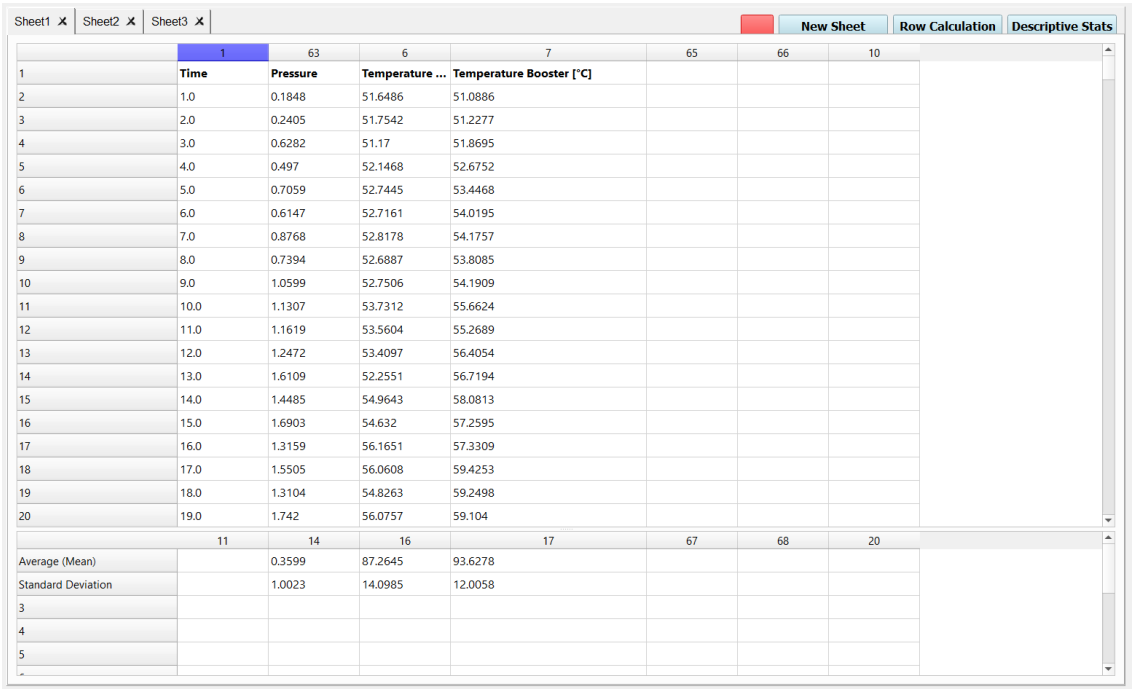
Export Graph: Save the graph as a jpeg or export it to PowerPoint or Excel.

5. Data Entry & Viewing

5.1. Spreadsheet

Each spreadsheet sheet consists of two tables, one positioned above the other. The upper table is the data table, where raw experimental data is entered. The lower table is the descriptive statistics table, where descriptive statistics are generated by clicking the Descriptive Statistics button (see section 5.1.3).

The first sheet of the spreadsheet with data table and descriptive statistics table



	1	63	6	7	65	66	10
1	**Time**	**Pressure**	**Temperature ...**	**Temperature Booster [°C]**			
2	1.0	0.1848	51.6486	51.0886			
3	2.0	0.2405	51.7542	51.2277			
4	3.0	0.6282	51.17	51.8695			
5	4.0	0.497	52.1468	52.6752			
6	5.0	0.7059	52.7445	53.4468			
7	6.0	0.6147	52.7161	54.0195			
8	7.0	0.8768	52.8178	54.1757			
9	8.0	0.7394	52.6887	53.8085			
10	9.0	1.0599	52.7506	54.1909			
11	10.0	1.1307	53.7312	55.6624			
12	11.0	1.1619	53.5604	55.2689			
13	12.0	1.2472	53.4097	56.4054			
14	13.0	1.6109	52.2551	56.7194			
15	14.0	1.4485	54.9643	58.0813			
16	15.0	1.6903	54.632	57.2595			
17	16.0	1.3159	56.1651	57.3309			
18	17.0	1.5505	56.0608	59.4253			
19	18.0	1.3104	54.8263	59.2498			
20	19.0	1.742	56.0757	59.104			

	11	14	16	17	67	68	20
Average (Mean)		0.3599	87.2645	93.6278			
Standard Deviation		1.0023	14.0985	12.0058			
3							
4							
5							

5.1.1. Adding and importing data

Data may be entered manually into the spreadsheet or imported from a CSV or Microsoft Excel file (see section 2.2).

- A file is designed to accommodate a single experiment (unless a multi-experiment is set up, see section 3.5). Each sheet within the project represents a single repeat of the experiment.

- Each column represents a variable. The variable name, which must be unique, should be entered in the first row and will automatically appear in bold. The data should be entered in the rows below the variable name. When an additional sheet is added, the variable names will be inherited automatically.

5.1.2. Core functionality

Copy, Paste, Delete and Undo

Copy: Select the data to be copied from the table, then use either the Ctrl + C keyboard shortcut or right-click and select Copy to copy the data from the current sheet.

Paste: Select the table region (or the top-left cell of the region) where you want to paste the data, then use either the Ctrl + V keyboard shortcut or right-click and select Paste. The data will be pasted only into the current sheet.

Paste (all sheets): Select the table region (or the top-left cell of the region) where you want to paste the data, then right-click and choose Paste (All Sheets). The data will be pasted into the corresponding region of every sheet.

Cut: Select the table region (or the top-left cell of the region) from which you want to cut the data, then use the Ctrl + X keyboard shortcut or right-click and choose Cut. The data will only be cut from the current sheet.

Delete: Select the data you want to delete from the table and press the Delete key on your keyboard. The data will only be deleted from the current sheet.

Undo: To undo changes, navigate to Edit > Undo. Only changes made to the spreadsheets can be reversed using the Undo button. Pressing undo reverses manual modifications to the spreadsheet data including the addition or deletion of rows and columns and the cutting or pasting of data. To remove row operations or descriptive statistics, use the Delete Column and Delete Row buttons, which are accessed with a right mouse click. Other project elements (e.g. graphs and statistical analyses) can be removed by right clicking on them in the left menu and selecting delete.

Insert & Delete Rows and Columns

Insert row: In edit mode, select one or more rows from the data table of any sheet, then right click and select 'Insert rows'. A number of rows equal to your selection will be added below the selected rows. Rows are added only in the currently selected sheet. Note: It is not possible to manually add rows to the descriptive statistics table; rows are added automatically when new descriptive statistics are generated.

Delete row: In edit mode, select one or more rows from the data table of any sheet (or the descriptive statistics table of the first sheet) then right click and select 'Delete rows'. The selected rows will be deleted. Rows can be deleted from both the data table and the descriptive statistics table. When deleting from the data table, rows will only be removed from the currently selected sheet. When deleting from the descriptive statistics table, rows will be removed from all sheets.

Insert column: In edit mode, select one or more columns from the data table of the first sheet, then right-click and choose Insert Columns. A number of columns equal to your selection will be added to the right of the selected columns. The new columns will be added to both the data table and the descriptive statistics table on all sheets. Note: Adding new columns will not affect existing row calculation references.

Delete Column: In edit mode, select one or more columns from the data table of the first sheet, then right-click and choose Delete Columns. The selected columns will be deleted from both the data table and the descriptive statistics table on all sheets. Note: Deleting columns will affect existing row calculations if a referenced column is deleted.

Add, remove and rename sheets

Adding a new Sheet: In edit mode, click the 'New Sheet' button at the top right of the window to add a new sheet. All existing column names, row calculations and descriptive statistics will automatically be added to the new sheet.

Removing a sheet: In edit mode, click the x to the right of the sheet name. When prompted, click OK to confirm deletion. The First sheet cannot be deleted.

Renaming a Sheet: In Edit mode, double click on the sheet name. Enter the new name in the dialogue window which appears and then click ok.

Quick Histogram and Quick plot

The software has a rich graphing functionality, enabling you to quickly create highly customized graphs. However, to provide a faster method of previewing data whilst in edit mode, the software includes a non-customizable 'Quick Plot' and 'Quick Histogram'. These are accessible by right clicking on the variable you want to plot and selecting either 'Quick Plot' or 'Quick Histogram'.

Quick Plot: Generates a plot of the selected variable against the independent variable. After opening the graph, user the interface below the graph to add additional data, show or hide the graph markers and line and to save the graph as an image or to export it to Word or PowerPoint.

Quick Histogram: Generates a histogram of the selected variable. The number of bins and the bar color can be customized after the histogram is created using the interface under the histogram. After modifying the number of bins or the color, press the update button for the changes to take effect. The interface under the histogram includes buttons to save the graph as an image or to export it to Word or PowerPoint.

5.1.3. Row calculations and descriptive statistics

Row calculations and descriptive statistics automate the computation of a wide range of mathematical functions including averages, standard deviations, logarithms and trigonometry across your entire dataset, with just the click of a button. They are applied on every sheet of your project, including sheets created after the calculation was added and can be updated (by pressing update above the main menu) whenever new data is added or existing data is changed.

Row calculations

Row calculations perform mathematical operations along each row e.g. a row operation to calculate the mean of 3 variables will return a column containing the average of 3 variables along each row. To add a row calculation, perform the following steps:

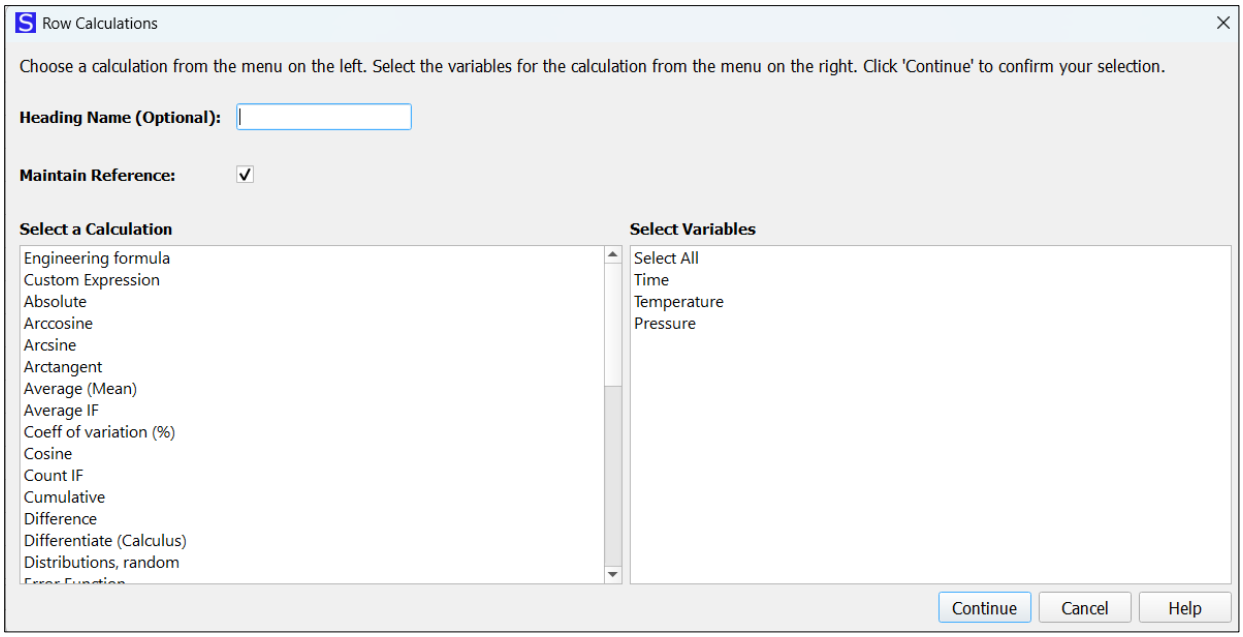
1. Navigate to the leftmost sheet and select an empty column in which to add the row calculation data.

2. Click on the blue 'Row calculation' button.

3. The Row Calculation dialogue box will open. Select a single mathematical operation from the left and at least one source variable from the right (if applicable), on which to perform the calculation.

4. Click on continue to perform the row operation. Depending on the row operation selected, an additional dialogue window may appear, prompting you for additional information.

Row calculation menu



The row calculation data will be inserted into the empty column you selected when creating the row calculation. The column will automatically be named according to the type of mathematical operation, but you can also choose a custom name within the Row Calculation dialogue box. The calculation will be performed on every sheet of the project. The calculation can be deleted by right clicking on the column and selecting delete columns. The calculation can be modified by selecting the column and clicking on the 'Row calculation' button, which will open the row calculation dialogue box with the previously chosen selections. If you make changes to your data, the row calculation data can be updated by either pressing the 'Update' button or entering analysis mode. Note that where relevant (e.g. for averages and standard deviations), empty cells and cells containing nan are ignored.

The following Row calculations are provided:

- **Absolute:** The magnitude of the number, ignoring any minus sign. Select a single source variable.
- **Arccosine:** The inverse of the cosine trigonometric function. Select a single source variable. Units are in degrees by default but can be modified in the settings.
- **Arcsine:** The inverse of the sine trigonometric function. Select a single source variable. Units are in degrees by default but can be modified in the settings.
- **Arctangent:** The inverse of the tangent trigonometric function. Select a single source variable. Units are in degrees by default but can be modified in the settings.
- **Average (Mean):** The sum of all values divided by the number of values.
- **Average IF:** The average of the values that meet a specified condition e.g. >5 (any one of <, >, >=, <=, or = are valid). See Appendix A for additional help.
- **Coefficient of variation (%):** The standard deviation as a percentage of the mean.
- **Cosine:** The Cosine (trigonometric function). Select a single source variable. Units are in degrees by default but can be modified in the settings.
- **Count IF:** The number of values that meet a specified condition. e.g. >5 (any one of <, >, >=, <=, or = are valid). See Appendix A for additional help.
- **Cumulative:** Returns a column with the cumulative sum of a single variable. Select a single source variable.
- **Custom Expression:** Build a custom expression using math symbols, basic functions and variable references. See Appendix A for additional help.
- **Difference:** Returns a column of the differences between consecutive values of the source column. For example, the 1st row of the row calculation will have the value equal to the difference between the 2nd and 1st rows of the source column.
- **Differentiate (Calculus):** Compute the differential of one variable with respect to any other. See Appendix A for additional help.

- **Distributions, random:** Generate an array of random numbers that follow a statistical distribution (Uniform, Normal, Exponential or Poisson). See Appendix A for additional help.
- **Error function:** The Gauss error function erf(x). Select a single source variable.
- **Equal:** Returns 1 if all of the source variables in the row are equal, otherwise it returns 0.
- **Exponential:** The exponential function, e^{x}. Select a single source variable.
- **Integrate (calculus):** Compute the integral of one variable with respect to any other. See Appendix A for additional help.
- **log:** The logarithm, with base 10. Select a single source variable.
- **ln:** The natural logarithm, with base e. Select a single source variable.
- **MAX:** The maximum value in the row.
- **Median:** The middle value in a sorted list of numbers.
- **MIN:** The minimum value in the row.
- **Modal Avg.:** The most frequently occurring value in the row.
- **Percentile:** The value below which the given percentage of the data in the row falls. See Appendix A for additional help.
- **Rank:** Returns the position of each value in the source column in ascending order. Select a single source variable.
- **Sequence:** Add a linear sequence of numbers. See Appendix A for additional help.
- **Sine:** The sine (trigonometric function). Select a single source variable. Units are in degrees by default but can be modified in the settings.
- **Sort (Ascending):** Sort the values of the source variable from smallest to largest. Select a single source variable.
- **Sort (Descending):** Sort the values of the source variable from largest to smallest. Select a single source variable.
- **Standard Deviation:** The measure of the average deviation of data points from the mean, indicating the spread of values in the population or sample.
- **Standard Error:** The standard deviation divided by the square root.
- **Sum:** The sum of all values in the column.
- **Sum IF:** The sum of values that meet a specified condition. e.g. >5 (any one of <, >, >=, <=, or = are valid). See Appendix A for additional help.
- **Tangent:** The tangent (trigonometric function). Select a single source variable. Units are in degrees by default but can be modified in the settings.

- **Variance:** A measure of the average squared deviation from the mean, indicating the spread of values in the population or sample.

Descriptive statistics

Descriptive statistics perform mathematical operations along each column e.g. to calculate the mean or standard deviation of all the values in a column. To add a descriptive statistic, perform the following steps:

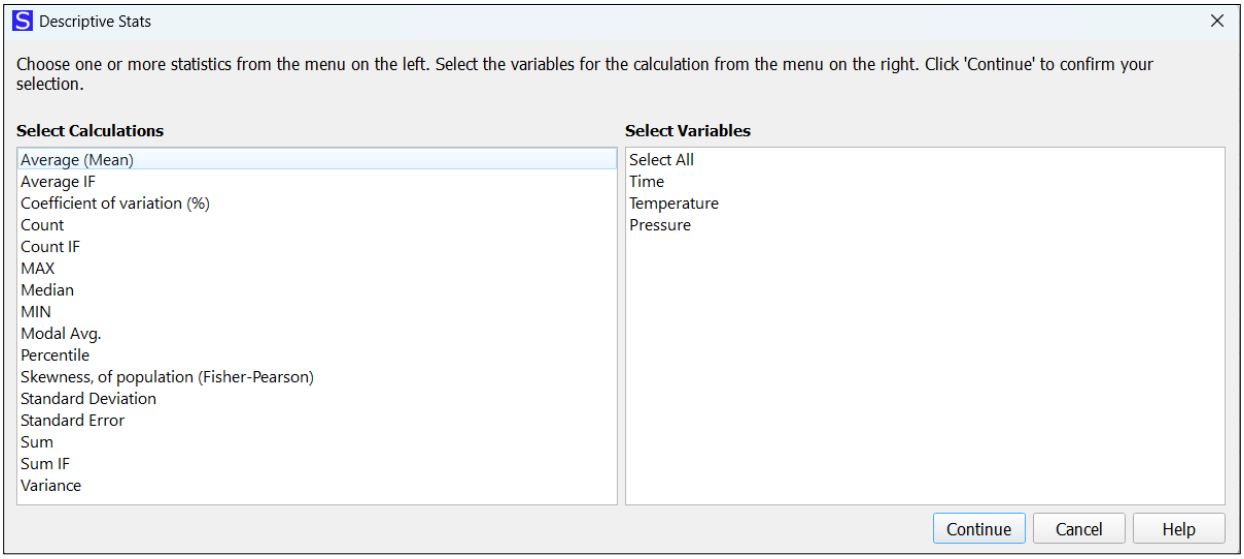
1. Navigate to the leftmost sheet and click on the 'Descriptive stats' button.

2. The descriptive statistics dialogue box will open. Select one or more mathematical operations from the left and at least one source variable from the right, on which to perform the calculation.

3. Click the continue button to perform the operation.

The descriptive statistics data will be inserted into the next available empty row in the lower table (descriptive statistics table). The calculation will be performed for every sheet of the project. The calculation can be deleted by right clicking on the row in the lower table and selecting Delete rows. The calculation cannot be directly modified and so to modify it you must delete and redo the calculation. If you make changes to your data, the calculation can be updated by either pressing the 'Update' button or entering analysis mode. Note that where relevant (e.g. for averages and standard deviations), empty cells and cells containing nan are ignored.

Descriptive statistics menu



The following Descriptive statistics are provided.

- **Average:** The sum of all values divided by the number of values.
- **Average IF:** The average of values that meet a specified condition. See Appendix A for additional help.
- **Coefficient of variation (%):** The standard deviation as a percentage of the mean.
- **Count:** The number of values, ignoring empty and nan.
- **Count IF:** The number of values that meet a specified condition. See Appendix A for additional help.
- **MAX:** The maximum value in the column.
- **Median:** The middle value in a sorted list of numbers in the column.
- **MIN:** The minimum value in the column.
- **Modal Avg.:** The most frequently occurring value in the column.
- **Percentile:** The value below which the given percentage of the data in the column falls. See Appendix A for additional help.
- **Skewness, of population (Fisher-Pearson):** Skewness quantifies the symmetry of a distribution.
- **Standard Deviation:** The measure of the average deviation of data points from the mean, indicating the spread of values in the population or sample.
- **Standard Error:** The standard deviation divided by the square root.
- **Sum:** The sum of all values in the column.
- **Sum IF:** The sum of values that meet a specified condition. See Appendix A for additional help.
- **Variance:** A measure of the average squared deviation from the mean, indicating the spread of values in the population or sample.

5.1.4. Mathematical functions in spreadsheet cells

Note: The mathematical functions are also applicable to the scribble sheet.

The spreadsheet supports the use of mathematical functions and basic arithmetic.

Basic arithmetic

Basic arithmetic can be performed in the cells of the data table by entering the '=' symbol, followed by an arithmetic calculation e.g. =(25+50+75)*3.4'. However, it is not possible to reference other cells for the calculation and only the result, not the formula, is stored by the program. To perform more complex mathematical and statistical calculations which reference cells, use row calculations (see section 5.1.3).

The following mathematical symbols can be used:

- + Addition
- - Subtraction
- * Multiplication
- / Division
- ^ Power
- ( ) Order of operations

It is also possible to calculate the sum, mean, minimum or maximum of a list of numbers in the following way:

Example of sum: =sum([2, 5, 7, 88, -55])

Example of mean: =mean([2, 5, 7, 88, -55])

Example of minimum: =min([2, 5, 7, 88, -55])

Example of maximum: =max([2, 5, 7, 88, -55])

Additional mathematics

The following additional mathematics functions are provided, where n should be replaced with the desired number. Several may be used in the same expression, and they can be combined with basic mathematic operations such as addition, subtraction or multiplication.:

deg to radian (convert):	=radians(n)
e (constant):	=e
E^x:	=exp(n)

Factorial:	=factorial(n)
Gauss Error function:	=erf(n)
Logarithm:	=log10(n)
Natural logarithm:	=log(n)
Pi (constant):	=pi
radian to deg (convert):	=degrees(n)

Trigonometry

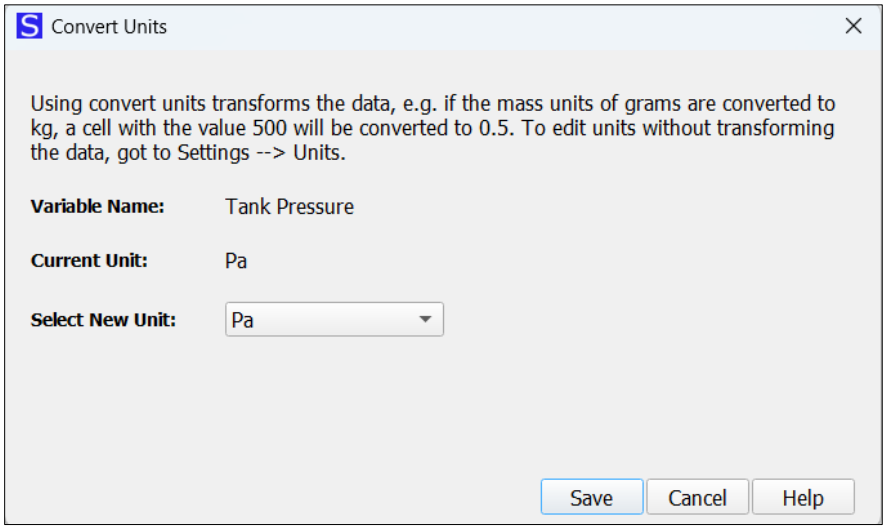
The 6 available trigonometric functions are shown below. The blue text is the command to be entered into the spreadsheet cell, where n should be replaced with the desired number. The default unit is degrees. Units can be changed to radians in the settings (see section 3.1). They cannot be combined with other functions.

Sin:	=sin(n)
Cosine:	=cos(n)
Tangent:	=tan(n)
Arcsine:	=arcsin(n)
Arccosine:	=arccos(n)
Arctangent:	=arctan(n)

5.1.5. Convert Units

Swift-Stat includes automatic unit conversions for units of many physical properties. Before converting units, initial units must be defined as explained in section 3.8. After units are defined, right click on the variable and select convert units. The convert unit's menu will show you the variable name and its current unit. Use the drop-down menu to select a new unit and press save. The data for the variable will be converted into the selected unit on all sheets.

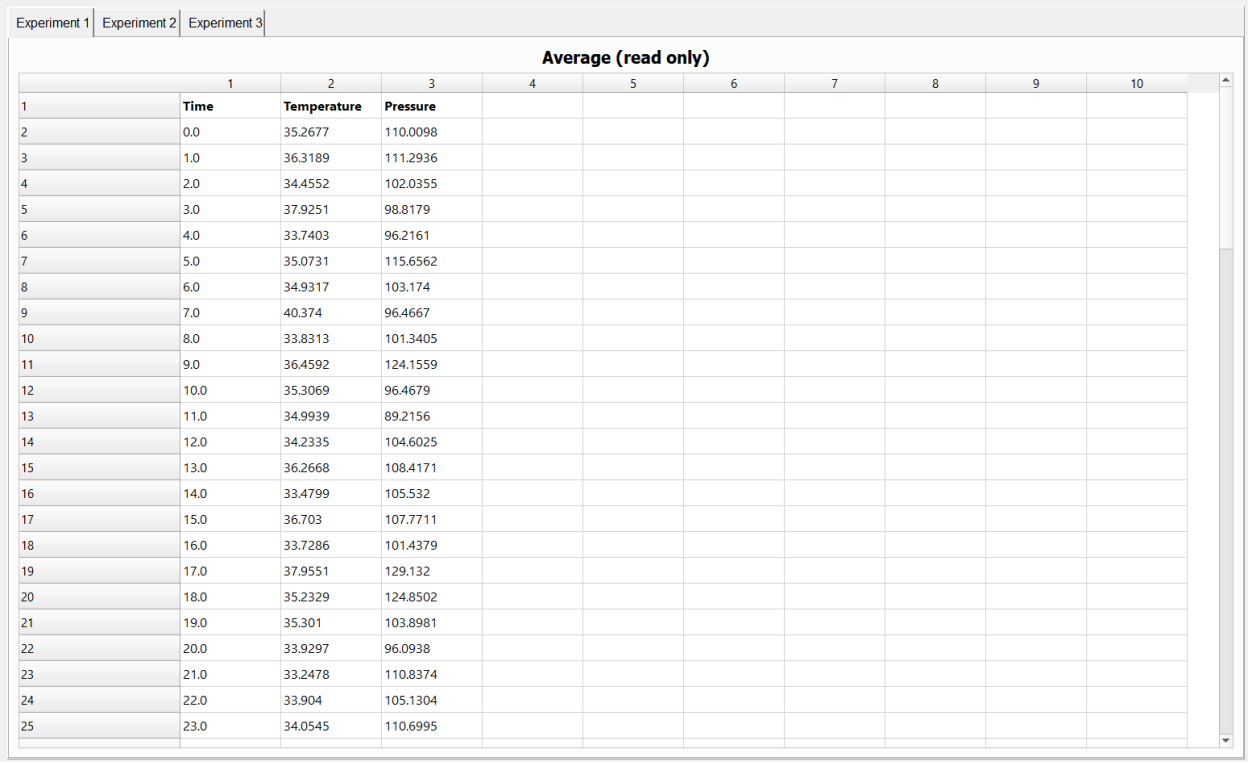
Convert unit's menu



5.2. Avg. Spreadsheet

The average spreadsheet automatically calculates the average of the data across all sheets in the spreadsheet. It is disabled by default because it can significantly slow down the software, but can be enabled in settings > Enable/disable Auto Avg. The average for each cell is calculated by summing the values from that cell across all sheets and dividing by the number of sheets. If a multi-experiment is setup, an average will be calculated for each experiment separately. It is not necessary to enable auto averaging to plot or analyze average data in your analyses. The average spreadsheet does not provide or enable any further functionality beyond showing you a spreadsheet of the average data. Note that after enabling auto averaging, you must switch to analysis mode to view the average spreadsheet.

Average spreadsheet


Experiment 1 | Experiment 2 | Experiment 3

Average (read only)

	1	2	3	4	5	6	7	8	9	10
1	Time	Temperature	Pressure							
2	0.0	35.2677	110.0098							
3	1.0	36.3189	111.2936							
4	2.0	34.4552	102.0355							
5	3.0	37.9251	98.8179							
6	4.0	33.7403	96.2161							
7	5.0	35.0731	115.6562							
8	6.0	34.9317	103.174							
9	7.0	40.374	96.4667							
10	8.0	33.8313	101.3405							
11	9.0	36.4592	124.1559							
12	10.0	35.3069	96.4679							
13	11.0	34.9939	89.2156							
14	12.0	34.2335	104.6025							
15	13.0	36.2668	108.4171							
16	14.0	33.4799	105.532							
17	15.0	36.703	107.7711							
18	16.0	33.7286	101.4379							
19	17.0	37.9551	129.132							
20	18.0	35.2329	124.8502							
21	19.0	35.301	103.8981							
22	20.0	33.9297	96.0938							
23	21.0	33.2478	110.8374							
24	22.0	33.904	105.1304							
25	23.0	34.0545	110.6995							

5.3. Scribble Sheet

The Scribble sheet is an additional spreadsheet which can be used to make notes, perform back-of-the-envelope calculations or store data without effecting your analyses. The scribble sheet does not support row calculations but does support the same in-cell formula as the main spreadsheet, as detailed below.

Basic arithmetic

Basic arithmetic can be performed in the cells of the scribble sheet by entering the '=' symbol, followed by an arithmetic calculation e.g. '=(25+50+75)*3.4'. However, it is not possible to reference other cells for the calculation and only the result, not the formula, is stored by the program.

The following mathematical symbols can be used:

- + Addition
- - Subtraction
- * Multiplication
- / Division
- ^ Power
- ( ) Order of operations

It is also possible to calculate the sum, mean, minimum or maximum of a list of numbers in the following way:

Example of sum: =sum([2, 5, 7, 88, -55])

Example of mean: =mean([2, 5, 7, 88, -55])

Example of minimum: =min([2, 5, 7, 88, -55])

Example of maximum: =max([2, 5, 7, 88, -55])

Additional mathematics

The following additional mathematics functions are provided, where n should be replaced with the desired number. Several may be used in the same expression, and they can be combined with basic mathematic operations such as addition, subtraction or multiplication.

deg to radian (convert):	=radians(n)
e (constant):	=e
E^x:	=exp(n)
Factorial:	=factorial(n)

Gauss Error function:	=erf(n)
Logarithm:	=log10(n)
Natural logarithm:	=log(n)
Pi (constant):	=pi
radian to deg (convert):	=degrees(n)

Trigonometry

The 6 available trigonometric functions are shown below. The blue text is the command to be entered into the scribble sheet cell, where n should be replaced with the desired number. The default unit is degrees. Units can be changed to radians under the settings tab (see section 3.1). They cannot be combined with other functions.

Sin:	=sin(n)
Cosine:	=cos(n)
Tangent:	=tan(n)
Arcsine:	=arcsin(n)
Arccosine:	=arccos(n)
Arctangent:	=arctan(n)

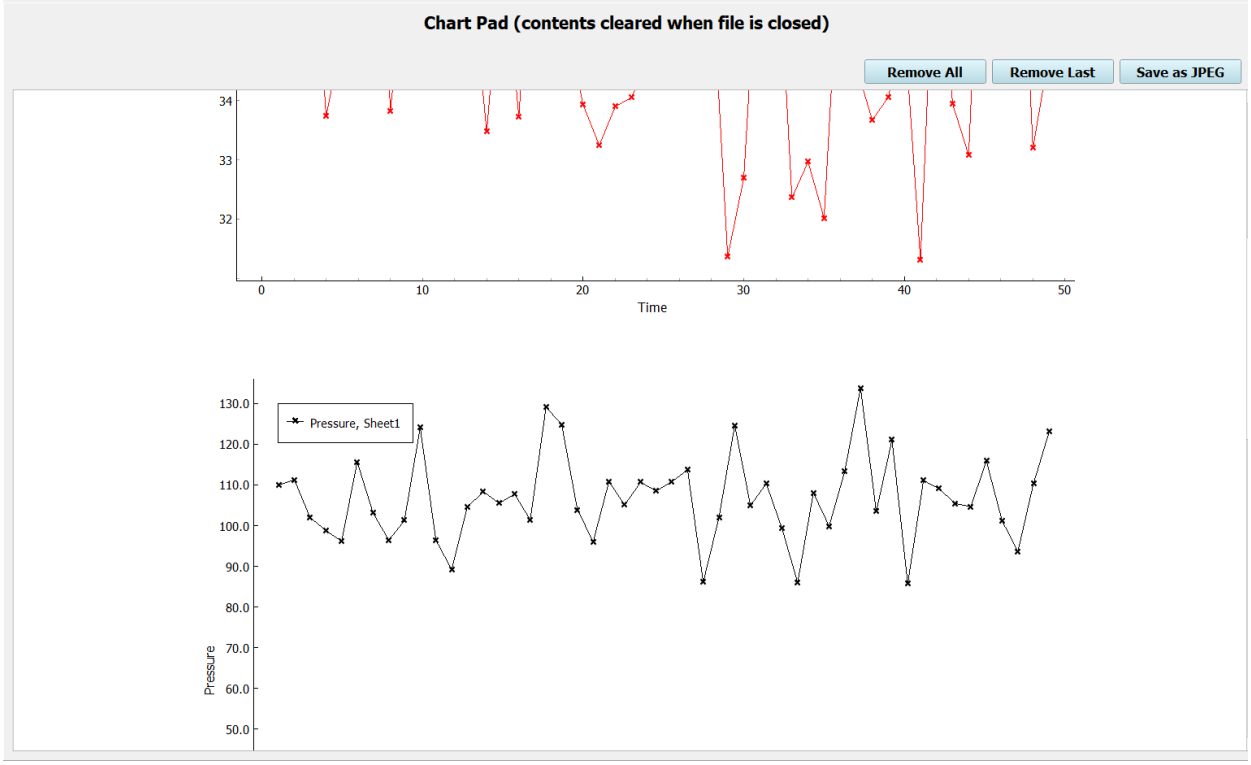
5.4. Chart Pad

Graphs from analyses and quick plots can be imported to the Chart Pad using the Add to Chart Pad button next to each graph. This allows you to view them side by side. The graphs are imported as images, capturing their state at the time of export, meaning it is not possible to zoom or scroll within the Chart Pad. To remove a graph from the Chart Pad, right-click on the graph and select Delete Graph. When you close the file, the contents of the Chart Pad will be lost. There are 3 buttons at the top right of the chart pad:

Remove All: Removes all graphs from the chart pad. The graphs within the analyses are not affected.

Remove last: Removes the last graph from the chart pad. The graph within the analysis is not affected.

Export: Saves all the graphs as image files which can be opened outside of the Swift-Stat.



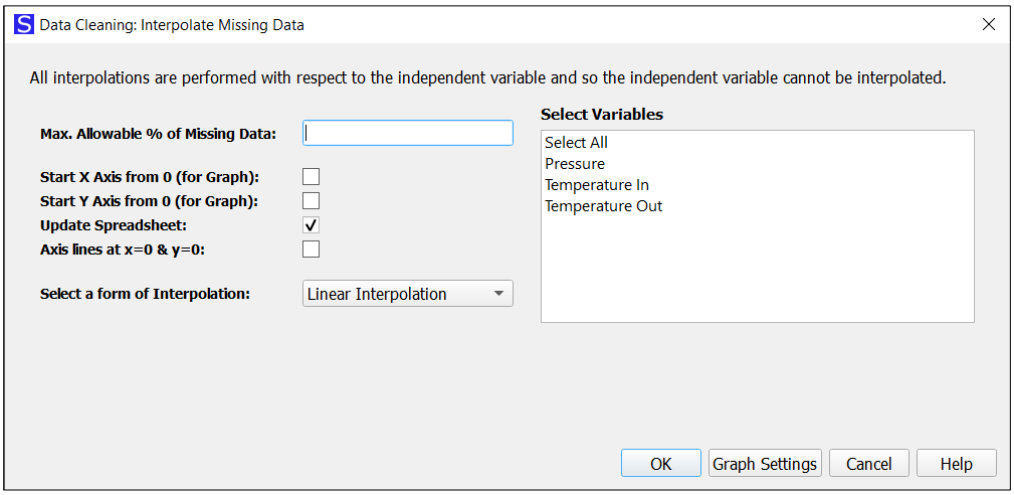
6. Clean & Transform Data

6.1. Data Cleaning: Missing Data

The missing data analysis fills in any missing values in the spreadsheet by interpolating the data. It generates graphs that display the original data, interpolation curves, and interpolated points. Additionally, it creates a table showing the proportion of missing data for each variable on every sheet. The modification to your data can be reversed by deleting the analysis (or switching to edit). If you perform the missing data analysis twice on the same variable, the 2nd analysis is applied to the output of the 1st analysis. To perform the analysis:

1. In analysis mode, open the 'Data Cleaning: Missing Data' menu (Main > Clean & Transform Data > Data Cleaning: Missing Data).
2. Select 1 of more variables from the 'Select variables' panel.
3. Choose a form of interpolation.
4. Select any of the additional settings explained below.
5. To change graph or axis settings, click on graph settings.

Data Cleaning: Missing Data Menu



Max. Allowable % of missing data: The threshold for the acceptable % of missing data. If any variable exceeds the threshold, it will be colored red in the missing data table.

Start X Axis from 0: Automatically starts the x axis from 0 even if the dataset does not cross the x axis. It is disabled if custom axis settings or logarithmic axis are selected in the graph settings.

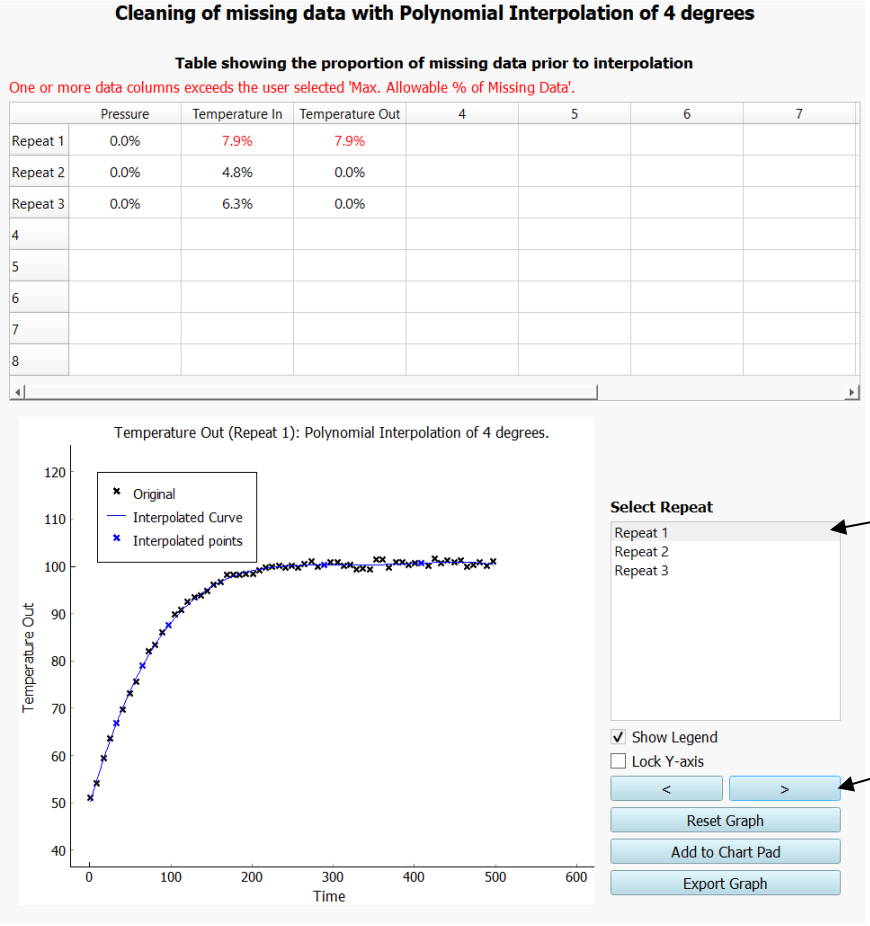
Start Y Axis from 0: Automatically starts the y axis from 0 even if the dataset does not cross the y axis. It is disabled if custom axis settings or logarithmic axis are selected in the graph settings.

Update Spreadsheet: If selected, the interpolated values are written to the spreadsheet in place of the missing values. It can be undone by either deleting the analysis or editing the analysis and unselecting 'update spreadsheet'.

Axis lines at x=0 and y=0: Adds horizontal and vertical lines to the graph at x=0 and y=0 respectively.
Select a form of interpolation: Select a form of interpolation, either linear, polynomial or exponential. If polynomial is selected, a text box will appear for selecting the polynomial order, which must be between 1 and 8. Exponential interpolation is only valid for positive numbers.

Navigating the results

An example result from a missing data analysis is shown below. The title indicates the type of interpolation applied. The accompanying table displays the proportion of missing data for each variable on each sheet. In this example, a maximum allowable missing data threshold of 7% was set and therefore the 2 cells which exceed this threshold are highlighted in red. The graph shows the original data, interpolated curve and interpolated points of the currently selected variable and repeat. Switch between repeats using the panel on the right. You can switch between repeats using the panel on the right, and navigate between variables using the two arrows in the bottom-right corner. The interactive graph allows you to scroll and zoom using your mouse, and you can hide specific plots by clicking on them in the legend.



Select a repeat of the current variable to display

Use the 2 arrows to cycle through variables

Show Legend: Toggles the visibility of the legend in the currently displayed graph.

Lock Y-Axis: Locks the Y-axis scrolling and zooming, allowing you to scroll and zoom on the X-axis without moving the Y-axis.
Reset Graph: Resets the graph view and axes scale of the currently displayed graph to its default view and scale.

Add to Chart Pad: Adds a screenshot of the currently displayed graph to the chart pad.

Export Graph: Save the graphs as jpegs or export the current graph to PowerPoint or Excel.

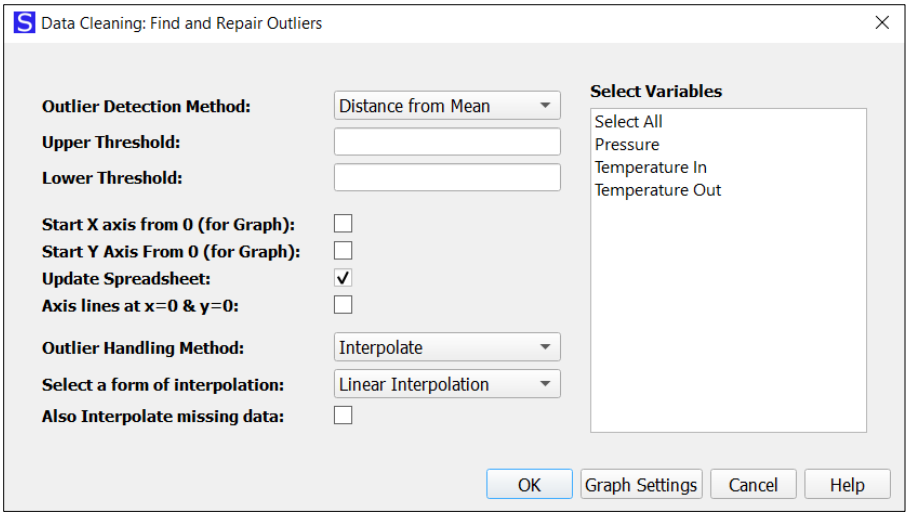
6.2. Data Cleaning: Outliers

The outlier analysis detects anomalies in the data using statistical rules and replaces them with interpolated values after excluding the outliers. It generates graphs that display the original data, the threshold used to identify outliers, the detected outliers, the interpolation curve, and the interpolated points. A table is also created to summarize the proportion of outlying data for each variable on each sheet. The modification to your data can be reversed by deleting the analysis (or switching to edit). If you perform the outlier analysis twice on the same variable, the 2nd analysis is applied to the output of the 1st analysis. To perform the analysis:

1. In analysis mode, open the 'Data Cleaning: Outliers' menu (Main > Clean & Transform Data > Data Cleaning: Outliers).
2. Select 1 of more variables from the 'Select variables' panel.
3. Select an outlier detection method.
4. Select upper and lower thresholds for outlier detection and where relevant a moving average window size.
5. Select a form of interpolation. If polynomial is selected, enter the order.
6. Select any of the additional settings explained below.
7. To change graph or axis settings, click on graph settings.

Note: The corrected datapoints are not retested after interpolation and so for complex data patterns interpolated values could hypothetically still be outliers. To retest the data after interpolation, create an additional analysis.

Data Cleaning: Outliers menu



Outlier Detection Method (see explanatory graphs at end of section)**:**

- Absolute threshold: Directly set the threshold value. If the upper and lower thresholds are defined as 15 and -4 respectively, any datapoints outside the range of -4 to 15 will be highlighted as outliers.
- Distance from mean: Calculates the mean of the data and sets the threshold as a fixed distance above and below the mean. If the mean average is -10 and the upper and lower

thresholds are chosen to be 5 and 4 respectively, any datapoints outside the range of -5 to -14 will be highlighted as outliers.

- Percentage of mean: Calculates the mean of the data and sets the threshold as a percentage of mean above and below the mean. If for example the mean average is 50 and the upper and lower thresholds are defined as 10% and 20% respectively, any datapoints outside of the range of 40 to 55 will be highlighted as outliers.
- Percentages of moving mean: Calculates the moving average of the data and sets the threshold as a percentage above and below the moving average.
- STD of mean: Calculates the mean and standard deviations of the data and sets the threshold as a user defined number of standard deviations above and below the mean. If for example the mean is 50, the standard deviation is 2 and the upper and lower thresholds are both 3, any datapoints outside of the range of 44 to 56 will be highlighted as outliers.
- STD of moving mean: Calculates the moving average and standard deviation of the data and sets the threshold as a user defined number of standard deviations above and below the moving average.
- Moving STD of moving mean: Calculates the moving average and moving standard deviation of the data and sets the threshold as a user defined number of moving standard deviations above and below the moving average.

Upper Threshold: For calculating the upper bound, above which data is anomalous. Except for absolute threshold detection, the threshold is relative to the mean.

Lower Threshold: For calculating the lower bound, below which data is anomalous. Except for absolute threshold detection, the threshold is relative to the mean and subtraction is implied.

Window size: The number of data points used to calculate each average or standard deviation when applying the moving mean or moving standard deviation. A larger window size leads to greater smoothing of the data.

Start X Axis from 0: Automatically starts the x axis from 0 even if the dataset does not cross the x axis. It is disabled if custom axis settings or logarithmic axis are selected in the graph settings.

Start Y Axis from 0: Automatically starts the y axis from 0 even if the dataset does not cross the y axis. It is disabled if custom axis settings or logarithmic axis are selected in the graph settings.

Update Spreadsheet: If selected, the interpolated values are written to the spreadsheet in place of the missing values. It can be undone by either deleting the analysis or editing the analysis and unselecting update spreadsheet.

Axis lines at x=0 and y=0: Adds horizontal and vertical lines to the graph at x=0 and y=0 respectively.

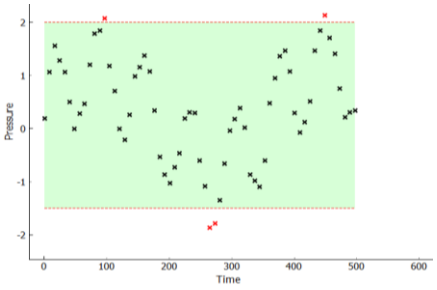
Outlier handling method: Select whether to replace outliers with interpolated values or to delete them and leave them empty

Select a form of interpolation: Select a form of interpolation, either linear, polynomial or exponential. If polynomial is selected, a text box will appear for selecting the polynomial order, which must be between 1 and 8. Exponential interpolation is only valid for positive numbers.

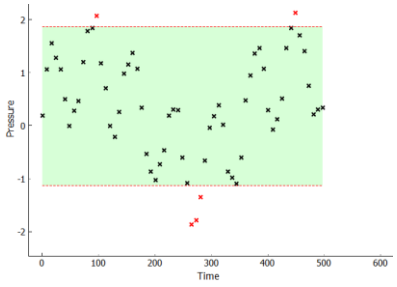
Also interpolate missing data: If selected, the interpolation will also be used to fill in any empty cells within the data series

The graphs below demonstrate each outlier detection method on a data series created by superimposing two noisy sine waves of different frequencies. It is apparent that the standard deviation-based detection methods are capable of capturing significantly more complex trends.

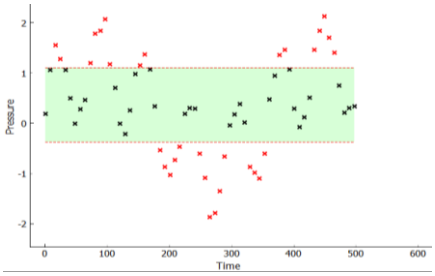
Absolute threshold



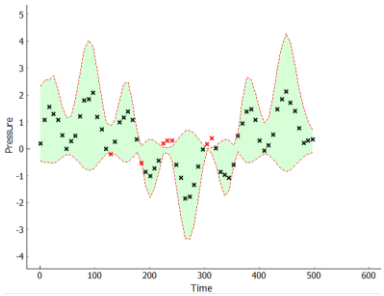
Distance from mean



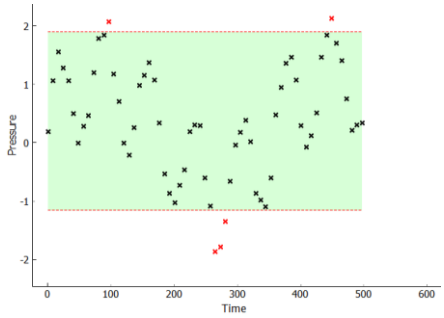
Percentage of mean



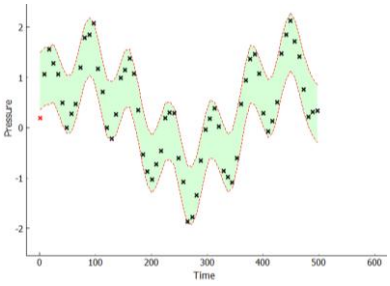
Percentage of moving mean



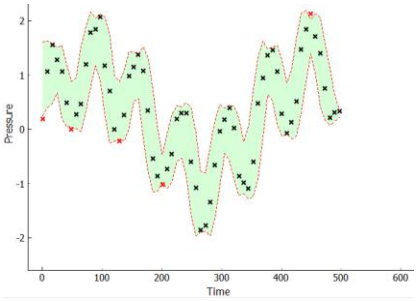
STD of mean



STD of moving mean

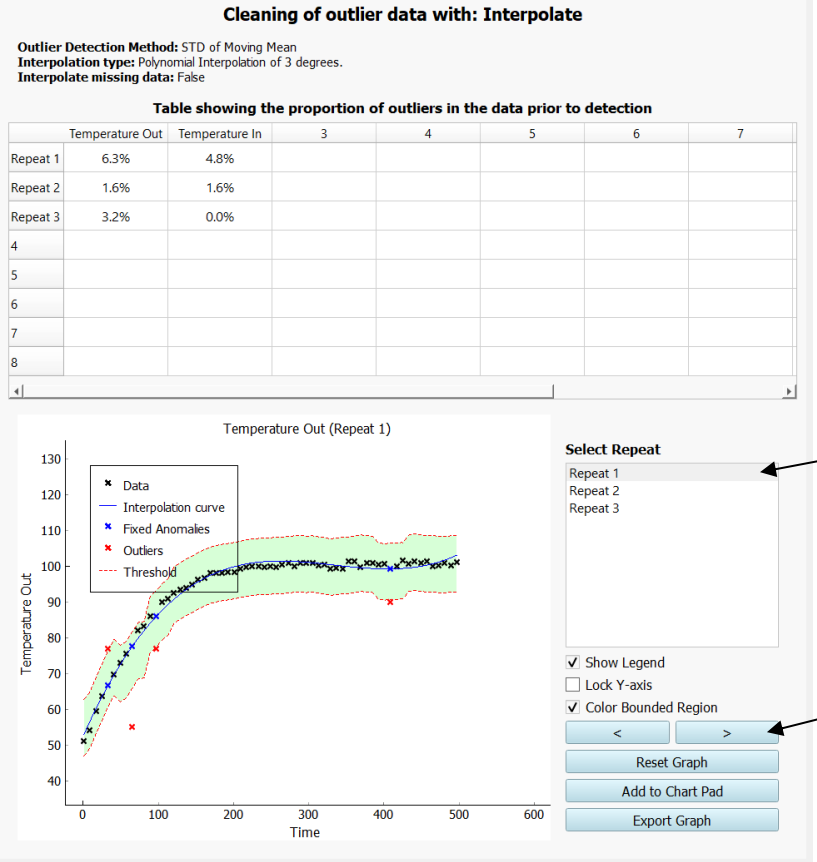


Moving STD of moving mean



Navigating the results

An example result from an outlier analysis is shown below. The heading indicates the outlier detection method and type of interpolation that were applied. The table displays the proportion of outliers for each variable on each sheet. The graph shows the original data, interpolated curve interpolated points, outliers and the outlier detection threshold of the currently selected variable and repeat. You can switch between repeats using the panel on the right, and navigate between variables using the two arrows in the bottom-right corner. The interactive graph allows you to scroll and zoom using your mouse, and you can hide specific plots by clicking on them in the legend.



Select a repeat of the current variable to display

Use the 2 arrows to cycle through variables

Show Legend: Toggles the visibility of the legend in the currently displayed graph.

Lock Y-Axis: Locks the Y-axis scrolling and zooming, allowing you to scroll and zoom on the X-axis without moving the Y-axis.

Color Bounded Region: Colors the valid region between the threshold lines in green.

Reset Graph: Resets the graph view and axes scale of the currently displayed graph to its default view and scale.

Add to Chart Pad: Adds a screenshot of the currently displayed graph to the chart pad**.**

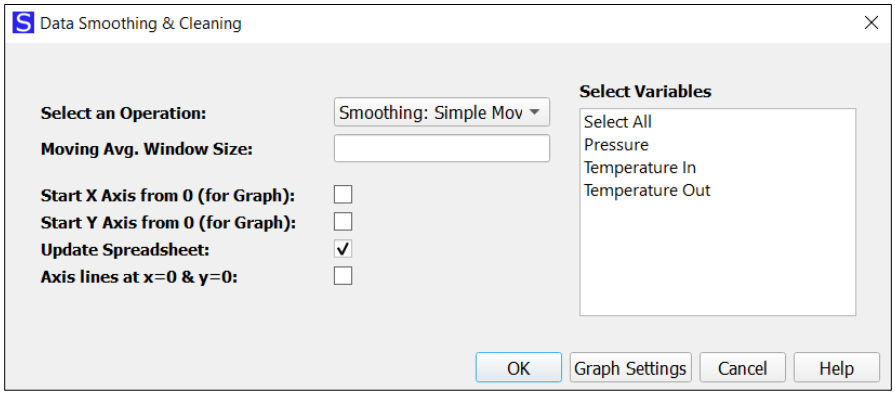
Export Graph: Save the graphs as jpegs or export the current graph to PowerPoint or Excel.

6.3. Data Smoothing & Scaling

The data smoothing & scaling analysis is used for scaling, smoothing or filtering the data in order to normalize the data, reduce noise or remove unwanted frequency components. It generates graphs showing both the original data and the processed data. The modification to your data can be reversed by deleting the analysis (or switching to edit). If you perform the smoothing & scaling analysis twice on the same variable, the 2nd analysis is applied to the output of the 1st analysis. To perform the analysis:

1. In analysis mode, open the 'Data Smoothing & Scaling' menu (Main > Clean & Transform Data > FFT frequency analysis).
2. select 1 of more variables from the 'Select Variables' panel.
3. Select a smoothing, scaling or filtering operation.
4. Enter parameters to define the selected smoothing, scaling or filtering operation.
5. Select any of the additional settings explained below.
6. To change graph or axis settings, click on graph settings.

Data Smoothing & Scaling menu



Select an operation:

- Simple Moving Average: Smooths the signal by averaging a specified number of consecutive data points. The number of points used in the averaging process is defined by the window size. For example, a window size of 5 means each point in the smoothed signal is the average of 5 consecutive data points.
- Exponential Moving Average: Smooths the signal by averaging whilst giving more weight to recent data points and exponentially decreasing weights of older data points. The degree of smoothing and sensitivity to recent data is determined by the span. A large span provides stronger smoothing and lower sensitivity to more recent data points.
- Gaussian Filter: A smoothing filter that is often more effective than a moving average at reducing noise while preserving the overall signal. The degree of smoothing is controlled by the number of standard deviations, with a higher value providing stronger smoothing.
- High-pass filter: Applies a Butterworth filter to attenuate low frequency components from the data. To apply the filter, enter values for the cutoff frequency, sample frequency and

filter order. The cutoff frequency is the frequency, in Hertz, below which the signal is attenuated. The sample frequency is the number of data points collected per second. The order of a Butterworth filter determines how sharply it attenuates signals below the cutoff frequency, with a higher order providing a stepper transition between the passband and the stopband.

- Low-pass filter: Applies a Butterworth filter to attenuate high frequency components from the data. To apply the filter, enter values for the cutoff frequency, sample frequency and filter order. The cutoff frequency is the frequency, in Hertz, above which the signal is attenuated. The sample frequency is the number of data points collected per second. The order of a Butterworth filter determines how sharply it attenuates signals above the cutoff frequency, with a higher order providing a stepper transition between the passband and the stopband.
- Max: Divides the data by its maximum absolute value. Optionally, you can also multiply the data by a fixed constant.
- Min-Max: Scales the data to between 0 and 1. Optionally, you can also multiply the data by a fixed constant.
- Standardization: Subtracts the mean of the data and divides by its standard deviation for each datapoint, thereby transforming the data to have a mean of 0 and a standard deviation of 1.
- Log: Transforms the data by replacing each value with its logarithm.
- Natural Log: Transforms the data by replacing each value with its natural logarithm.

Start X Axis from 0: Automatically starts the x axis from 0 even if the dataset does not cross the x axis. It is disabled if custom axis settings or logarithmic axis are selected in the graph settings.

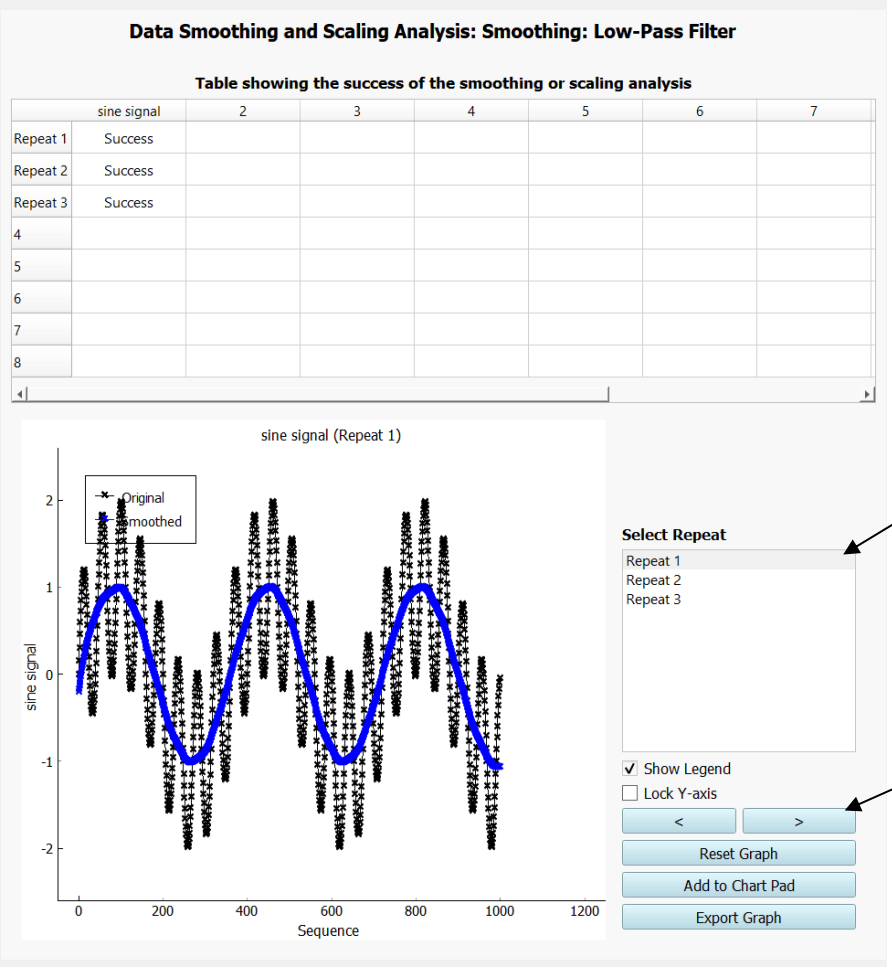
Start Y Axis from 0: Automatically starts the y axis from 0 even if the dataset does not cross the y axis. It is disabled if custom axis settings or logarithmic axis are selected in the graph settings.

Update Spreadsheet: If selected, the interpolated values are written to the spreadsheet in place of the missing values. It can be undone by either deleting the analysis or editing the analysis and unselecting ‘update spreadsheet’.

Axis lines at x=0 and y=0: Adds horizontal and vertical lines to the graph at x=0 and y=0 respectively.

Navigating the results

An example result from a data smoothing & scaling analysis is shown below. The heading indicates the type of smoothing or scaling that was applied. The accompanying table displays whether the operation was successful for each variable on each sheet. The graph shows the data before and after smoothing or scaling. Switch between repeats using the panel on the right. Switch between variables using the 2 arrows at the bottom right. The interactive graph enables you to scroll and zoom in with your mouse and to hide plots by clicking on them in the legend. In this example we have applied a low pass filter to a signal created by superimposing low frequency and high frequency sine waves. The output of the analysis is the low frequency sine wave.



Show Legend: Toggles the visibility of the legend in the currently displayed graph.

Lock Y-Axis: Locks the Y-axis scrolling and zooming, allowing you to scroll and zoom on the X-axis without moving the Y-axis.

Reset Graph: Resets the graph view and axes scale of the currently displayed graph to its default view and scale.

Add to Chart Pad: Adds a screenshot of the currently displayed graph to the chart pad.

Export Graph: Save the graphs as jpegs or export the current graph to PowerPoint or Excel.

7. Data Insights

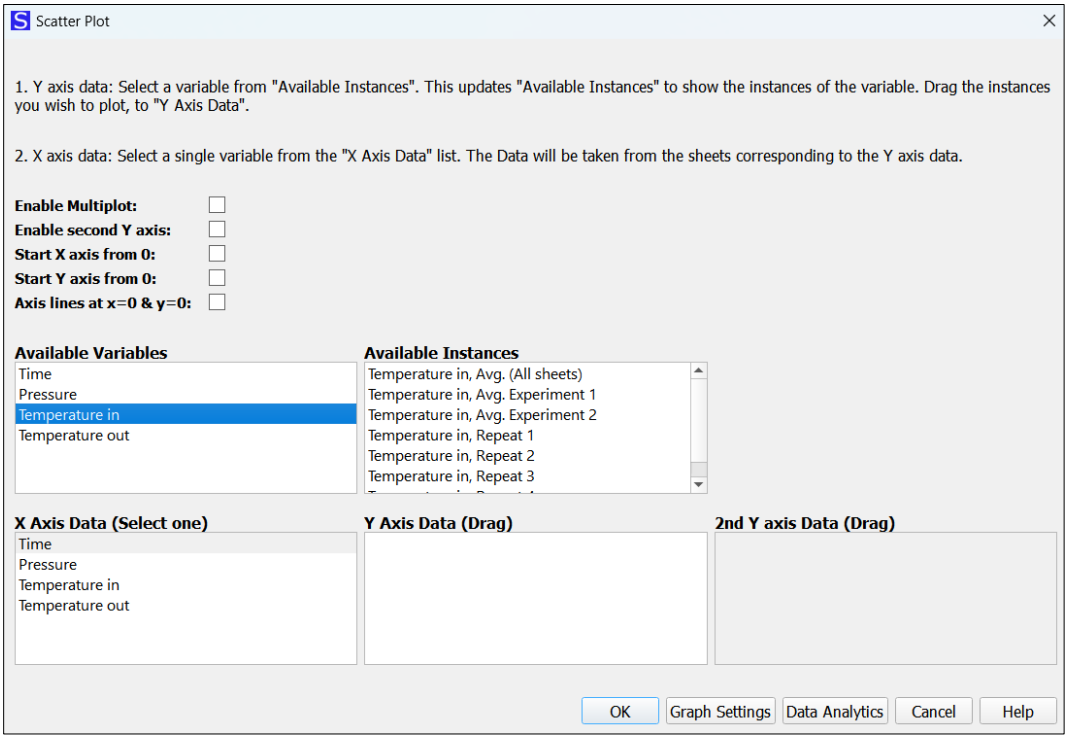
7.1. Scatter Plots

The scatter plot analysis generates interactive graphs that can be scrolled and zoomed with your mouse. There are a wide array of features including, secondary y axes, logarithmic axis scales, trend lines, errors bars and uncertainty bands utilizing statistics such as standard deviations and confidence intervals. To create a scatter plot:

1. In analysis mode, open the scatter plots menu (Main > Data Insights > Scatter Plots).
2. Select a variable from the 'Available Variables' panel to update the 'Available Data' panel with all of the instances (repeats and averages) of the variable.
3. Drag the instance you wish to plot from the 'Available Data' panel to the 'Y Axis Data' panel. The available instances include both individual repeats and averages.
4. To add additional variables to the graph, repeat steps 2 and 3.
5. Select a single Variable from the 'X axis Data' panel.
6. Select any of the additional settings explained below.
7. To modify graph or axis settings including custom axis ranges and logarithmic axes, click on graph settings.
8. To add trend lines, error bars or uncertainty bands click 'Data Analytics' (explained below).

Note: The data appears in the graph legend in the same order as it appears in the 'Y axis data' panel. Rearrange the variable instances in the 'Y axis data' panel by dragging them, which will also change their order in the graph legend. To delete data from the 'Y axis data' panel, highlight the data (e.g. Pressure, Repeat 1) and press the delete key on your keyboard.

Scatter Plot menu



Enable Multiplot: If selected, each repeat will be plotted on a separate graph. The 'Available Data' panel will show a list of all variables, instead of all repeats and averages of a single variable. Drag variables to plot to the Y axis data panel (or 2nd Y axis if enabled). Once 'OK' is pressed, a separate graph of the selected variables is created for each repeat and their average (and the average of each experiment if a 'multi-experiment' is set up).

Enable second Y axis: Plots data on a second y axis on the right of the graph. It is useful for displaying data with very different magnitudes on the same graph**.** When 'Enable second Y axis' is selected, the '2nd Y axis Data' panel becomes enabled. Drag the data for the second Y axis from 'Available Data' to '2nd Y axis Data'.

Start X Axis from 0: Automatically starts the x axis from 0 even if the dataset does not cross the x axis. It is disabled if custom axis settings are selected in the graph settings.

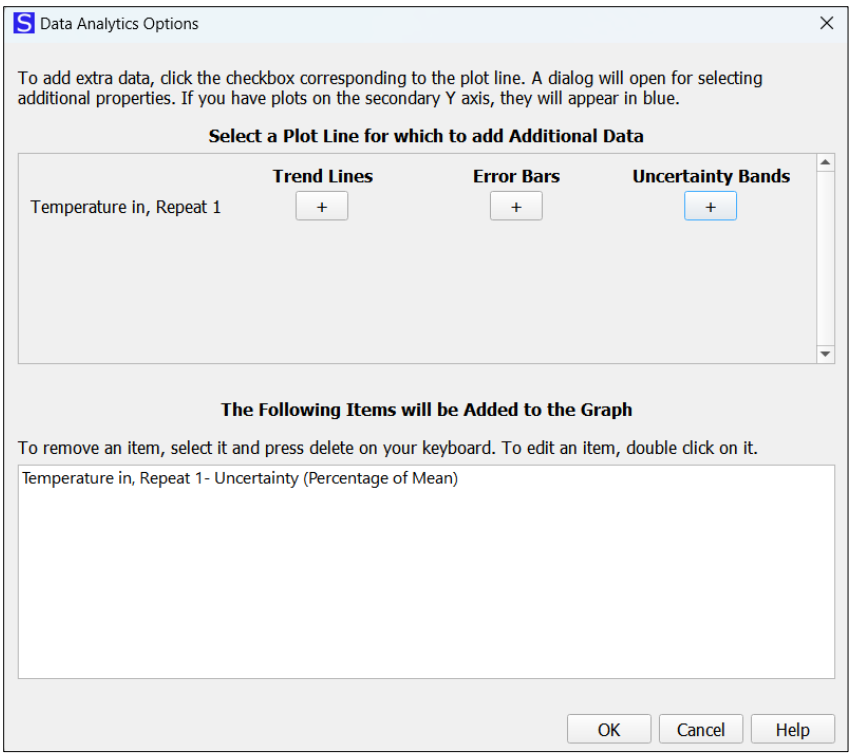
Start Y Axis from 0: Automatically starts the y axis from 0 even if the dataset does not cross the y axis. It is disabled if custom axis settings are selected in the graph settings.

Axis lines at x=0 and y=0: Adds horizontal and vertical lines to the graph at x=0 and y=0 respectively.

Data Analytics

After clicking on 'Data Analytics' in the 'Scatter Plot menu', you can add trend lines, error bars or uncertainty bands by clicking on the '+' icon corresponding to the row of the data series (i.e. variable instance) and the column of the trend line, error bar or uncertainty band. Only data series added to the plot in the main scatter plot menu will be visible in the data analytics menu.

Scatter Plot data analytics menu



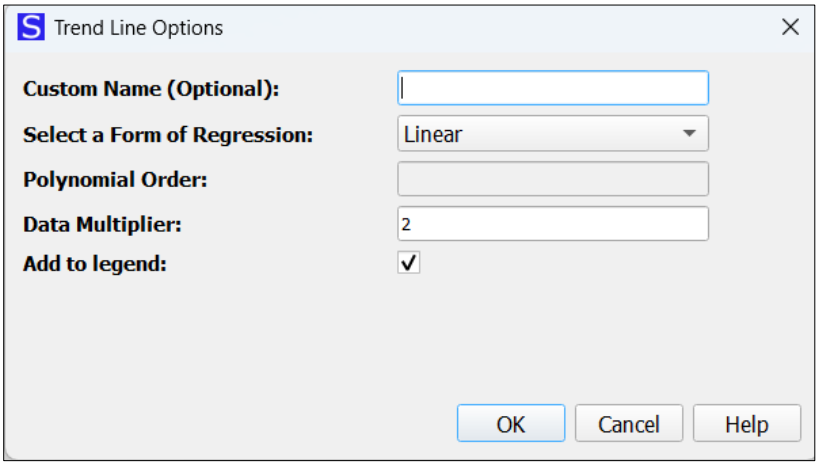
After clicking the '+' icon, a menu will open to define the settings for the data analytics element. After accepting the settings, the analytics element will be displayed in the white panel at the bottom of the scatter plot data analytics menu. All elements shown in this panel will be added to the graph.

Modifying line and marker style: The analytics element will appear in the graph settings, where its line and marker color and style can be modified.

Editing or deleting analytic elements: Analytics element can be edited by double clicking on them or deleted by selecting them and pressing delete on your keyboard.

Data analytics: Trend lines

Add trend lines to your data using linear, polynomial or exponential regression. The line and marker style of the trend line can be edited through the graph settings, accessible from the main scatter plot menu.



Custom Name (Optional): The custom name is used in the graph legend and will override the default name.

Select a form of regression: Select a form of regression, either linear, polynomial or exponential. If polynomial is selected, a text box will appear for selecting the polynomial order, which must be between 1 and 8. Exponential regression is only valid for positive numbers.

Polynomial Order: If polynomial regression is selected, enter an order between 1 and 8 for the polynomial function.

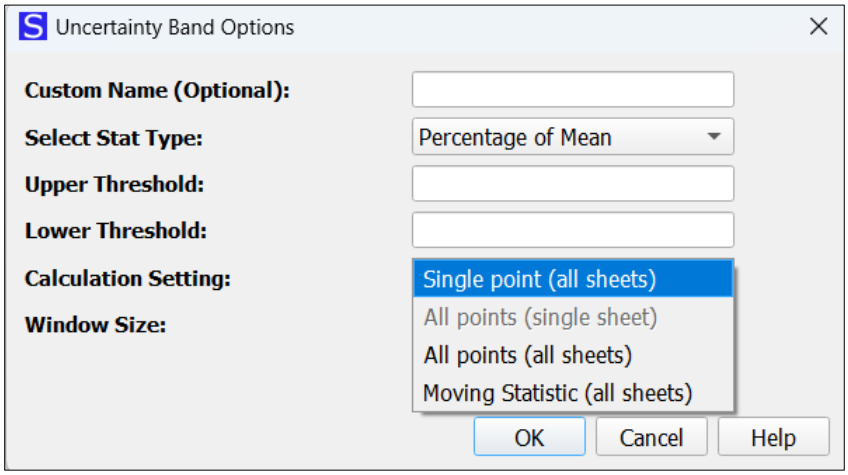
Add to legend: If enabled, the trendline will be added to the graph legend.

Data Multiplier: Select the number of datapoint markers on the trendline relative to the source data. For example, if 2 is entered, the trendline will have twice the number of datapoint markers. By default, trendline datapoint markers are hidden. They can be made visible from the graph settings.

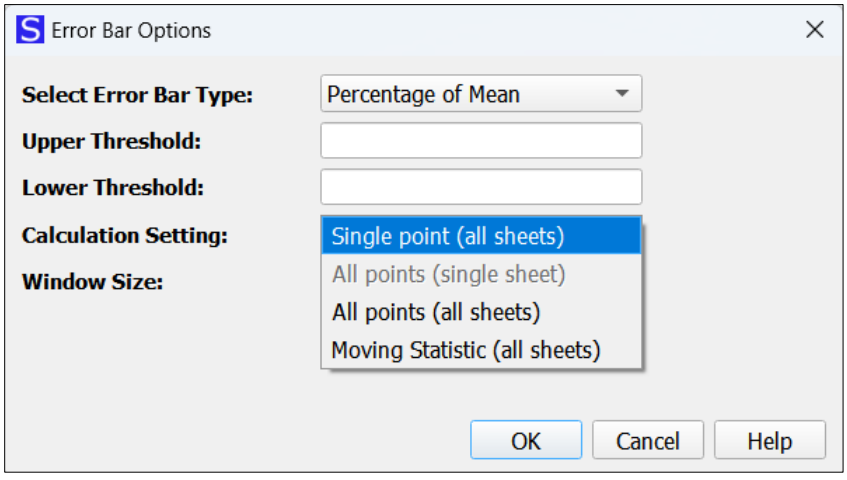
Data analytics: Error bars and Uncertainty bands

Add a wide range of statistics including percentage of mean, standard deviations, confidence intervals and prediction intervals to show the error or uncertainty in your data. The calculations for error bars and uncertainty bounds are identical, the difference being that the uncertainty bands are represented as plots that bound a region on the graph instead of as error bars. The line and marker style of the uncertainty bands can be edited through the graph settings, accessible from the scatter plot menu. Error bands inherit their color from their parent plot.

Uncertainty band menu (within the Scatter Plot data analytics menu)



Error bar menu (within the Scatter Plot data analytics menu)



Custom Name (Optional): The custom name is used in the graph legend and will override the default name.

Select Stat Type: There are 8 built-in statistic types to choose from for calculating the uncertainty bands or error bars. The 'Percentage' and 'Specify relative value' are calculated relative to each

individual datapoint on the graph since they can be applied to a single number. The other statistic types require a dataset for the calculation. The dataset is defined by the 'Calculation settings' (explained under the calculation settings heading). The 8 built-in statistic types are as follows:

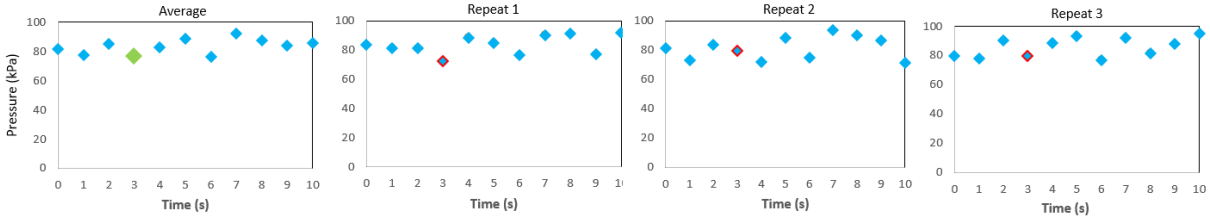
- **Confidence Interval:** Calculates the confidence interval for a user defined confidence level. This is commonly called a confidence band; it shows the expected spread of the mean.
- **Percentage:** Calculates the error or uncertainty as a fixed percentage above and below each datapoint. It is calculated relative to each individual datapoint and so does not need a calculation setting to be chosen.
- **Percentage of mean:** Calculates the error or uncertainty as the percentage of mean above and below the data.
- **Prediction interval:** Calculates the prediction interval for a user defined threshold. This is commonly called a prediction band; it shows the expected spread of the data.
- **Standard deviation:** Calculates the error or uncertainty as a fixed number of standard deviations above and below the data.
- **Standard error:** Calculates the error or uncertainty as the standard error above and below the data.
- **Specify relative value:** Calculates the error or uncertainty as a fixed distance above and below each datapoint. It is calculated relative to each individual datapoint and so does not need a calculation setting to be chosen.
- **Specify absolute value (for uncertainty bands only):** Creates horizontal lines with y axis values as specified by the upper and lower thresholds. Since it is not relative, threshold values can be negative.
- **UCL/LCL (for uncertainty bands only):** Calculates the uncertainty as 3 standard deviations above and below the data.

Calculation settings**:** The calculation settings control which data is used to calculate the statistics. They are illustrated in the graphs below which highlight the data used to calculate a statistic for the 4thdatapoint (x axis value of 3). Each row of graphs demonstrates how one of the 4 calculation settings samples the data. The enlarged green datapoint is the one for which the statistic is being calculated. The symbols with red boarders are the ones used for the calculation. Note that 'All points (single sheet)' is only applicable to individual repeats and the other options are only applicable to averages (including those of individual experiments, if a ‘multi-experiment’ is set up). Each of the 4 calculation settings uses the following data in the example illustration below:

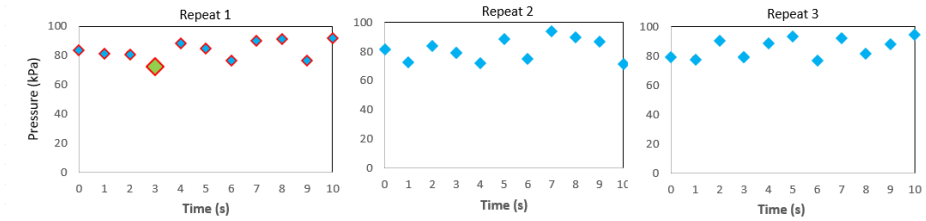
- **Single point (all sheets):** The 4th data point from each of the 3 repeats. To prevent misuse, it is only available if the source dataset is an average.
- **All points (single sheet):** All data points from repeat 1. To prevent misuse, it is only available if the source dataset is a single sheet.
- **All points (all sheets):** All data points from all repeats. To prevent misuse, it is only available if the source dataset is an average.

- **Moving Statistic (all sheets):** The 3rd to 5th datapoints of each of the 3 repeats (since the window size is 3 in the example). To prevent misuse, it is only available if the source dataset is an average.

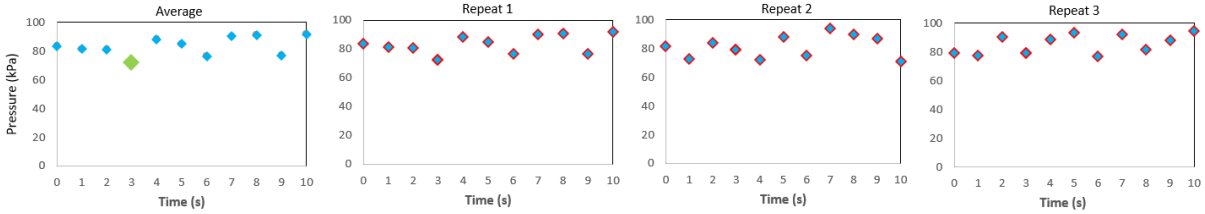
Single point (all sheets):



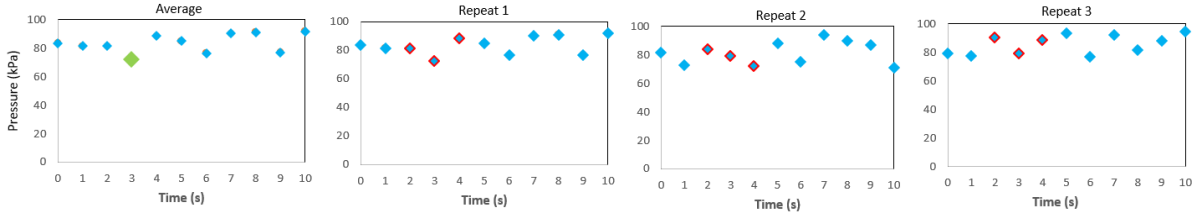
All points (single sheet)



All points (all sheets)

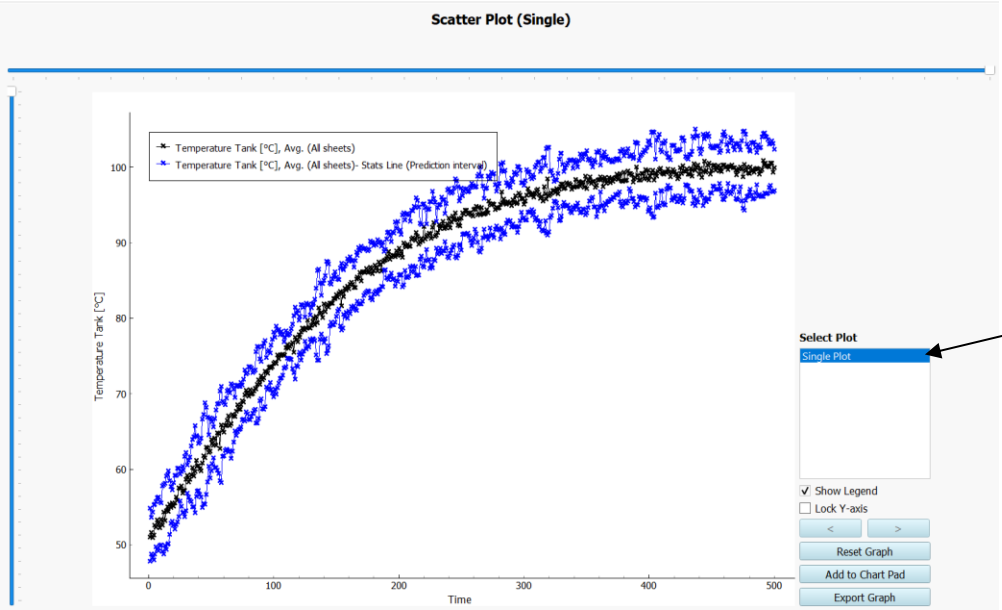


Moving Statistic (all sheets)



Navigating the results

An example scatter plot is shown below. The graph contains a single plot of the average of a variable and its prediction interval. The interactive graph allows you to scroll and zoom using your mouse, and you can hide specific plots by clicking on them in the legend. The vertical slider above the graph and horizontal slider to the left can be used to resize the graph.



Select a repeat of the current variable to display (for when Multiplot is enabled)

Reset Graph: Resets the graph view and axes scale of the currently displayed graph to its default view and scale.

Add to Chart Pad: Adds a screenshot of the currently displayed graph to the chart pad.

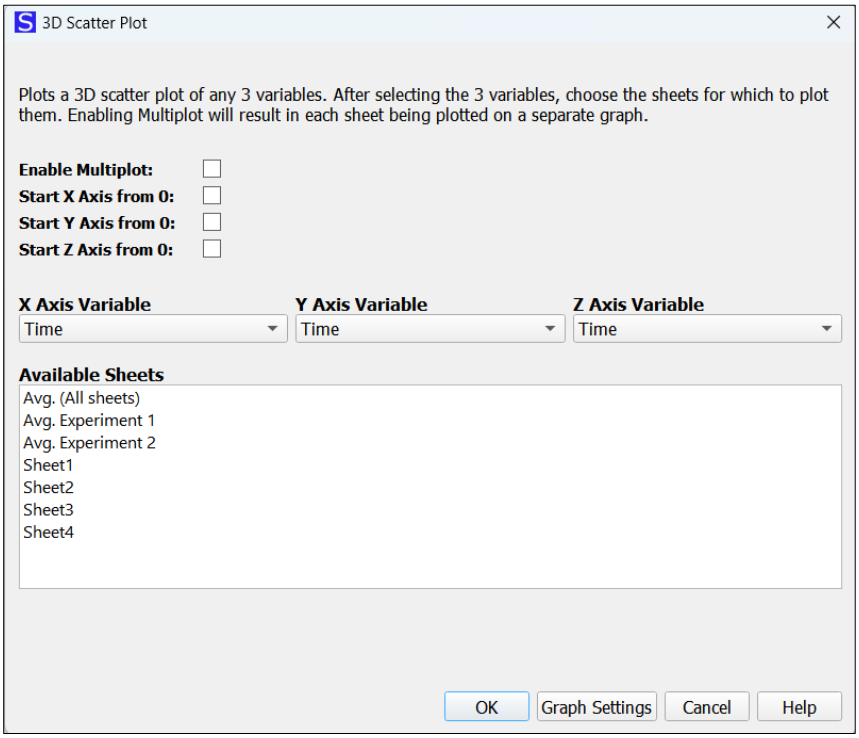
Export Graph: Save the graphs as jpegs or export the current graph to PowerPoint or Excel.

7.2. Data Insights: 3D Scatter Plots

The 3D scatter plot analysis generates an interactive 3D graph of your data, that you can rotate, scroll and zoom with your mouse. To perform the analysis:

1. In analysis mode, open the 3D Scatter Plots menu (Main > Data Insights > 3D Scatter Plots).
2. Select variables to plot for each of the x, y and z axes.
3. Choose which sheets to add to the graph. To plot each sheet on a separate graph, select 'Enable Multiplot'.
4. Select any of the additional settings explained below.
5. To change graph or axis settings, click on graph settings.

3D Scatter Plot menu



Enable Multiplot: Select to plot each sheet on a separate graph, otherwise all sheets are plotted on a single graph.

Start X Axis from 0: Automatically starts the x axis from 0 even if the dataset does not cross the x axis. It is disabled if custom axis settings are selected in the graph settings.

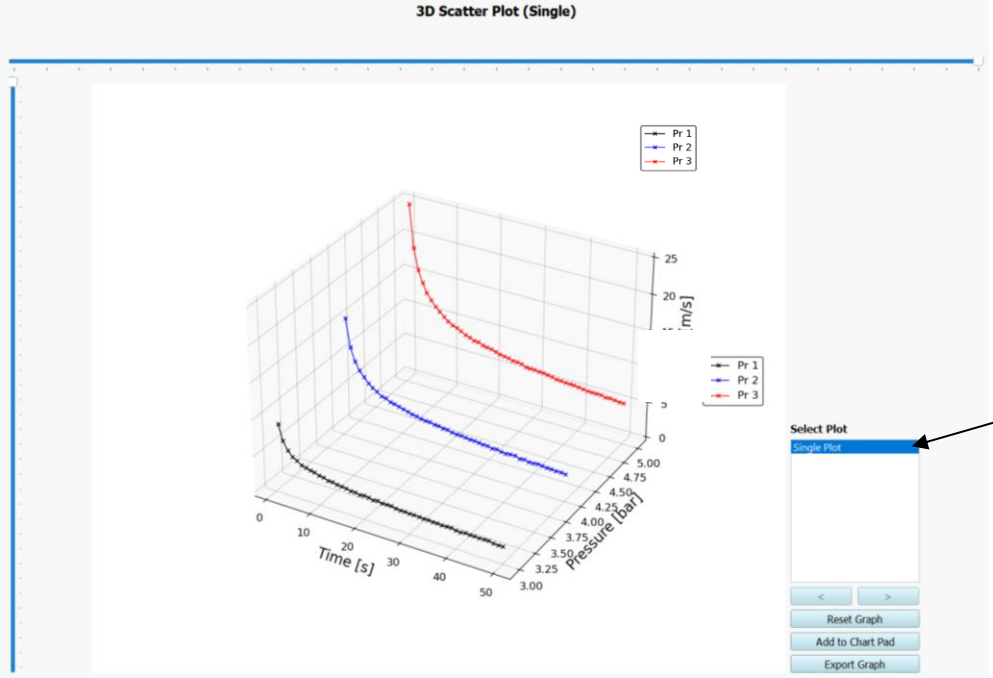
Start Y Axis from 0: Automatically starts the y axis from 0 even if the dataset does not cross the y axis. It is disabled if custom axis settings are selected in the graph settings.

Start Z Axis from 0: Automatically starts the z axis from 0 even if the dataset does not cross the z axis. It is disabled if custom axis settings are selected in the graph settings.

X, Y, Z Axis Data: Select the variable to plot on each axis.

Navigating the results

An example of a 3D Scatter plot is shown below. In the current example there are 3 plots on a single graph. Rotate the graph by clicking on it with the left mouse button and dragging. Shift the data around the graph by holding down the scroll wheel on your mouse and moving the mouse within the graph. Zoom in and out by clicking on the graph with the right mouse button and moving the mouse backwards or forwards. The vertical slider above the graph and horizontal slider to the left can be used to resize the graph.



Select a repeat of the current variable to display (if Multiplot is enabled)

Reset Graph: Resets the graph view and axes scale of the currently displayed graph to its default view and scale.

Add to Chart Pad: Adds a screenshot of the currently displayed graph to the chart pad.

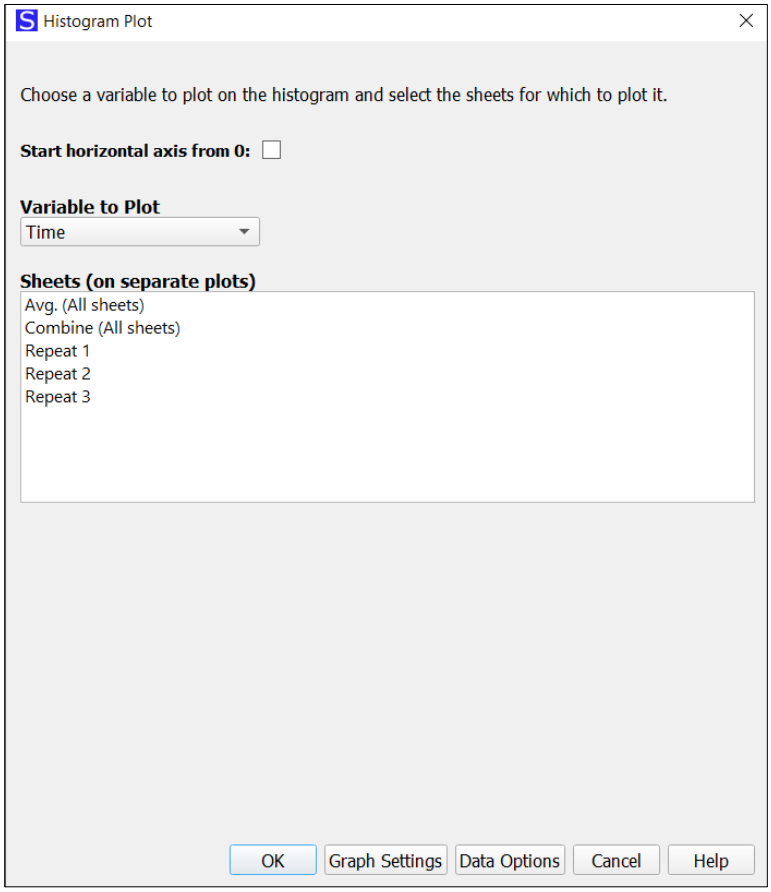
Export Graph: Save the graphs as jpegs or export the current graph to PowerPoint or Excel.

7.3. Data Insights: Histograms

The Histograms analysis generates histograms of your data. It includes the option to enhance the plot by overlaying additional statistics such as lines of mean and standard deviation, a normal curve or process capability (see full details below). To perform the analysis:

1. In analysis mode, open the histogram menu (Main > Data Insights > Histograms).
2. Select a single variable to plot.
3. Select the sheets to plot. Each sheet is plotted on a separate histogram. You can also plot the average of all sheets, the average of the sheets assigned to an experiment (if 'multi-experiment' is set up) or the combined data of all sheets or of those assigned to an experiment.
4. Select whether to start the horizontal axis from 0 (disabled if custom axis settings are selected in the graph settings).
5. To add additional statistics to the histogram, click on 'Data Options' (see below).
6. To edit graph or axis settings, click on the graph settings (see below).

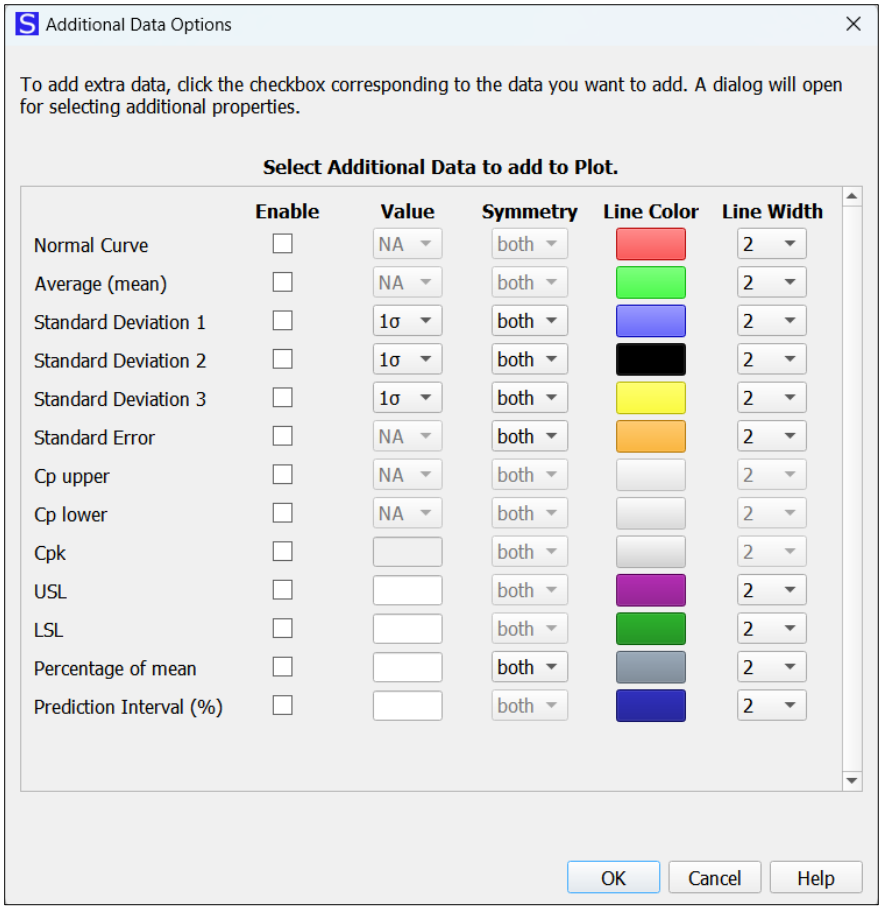
Histograms menu



Data Options

The data options menu provides statistics which can be added to the histogram, including mean and standard deviation, a normal curve and process capability. With the exception of process capability and the normal curve, these statistics are displayed as vertical lines overlaid on the histogram. To add a statistic, locate it in the data options menu (shown below) and check 'Enable'. Where applicable, select a value (e.g., the number of standard deviations or a prediction interval). For values calculated relative to the mean, use the Symmetry column to specify whether the statistic should be added to the left of the mean, the right, or both. The menu also allows you to customize the color and line width of the added statistics.

Histograms Data Options menu



Normal Curve: Overlays a curve of normal distribution fitted to the histogram data.

Average: No additional settings required.

Standard Deviation: Up to 3 lines of standard deviation can be added. Select the number of standard deviations and whether or not to add on both sides of the mean.

Standard Error: Select whether to add on both sides of the mean.

Cp upper & Cp lower: Adding the process capability upper and lower limits requires enabling and providing values for USL & LSL. 3σ is used as the divisor in the formula.

Cpk: Adding the Process Capability Index requires enabling and providing values for USL & LSL.

USL & LSL: Enter values for the upper and lower specification limits, required for calculating process capability.

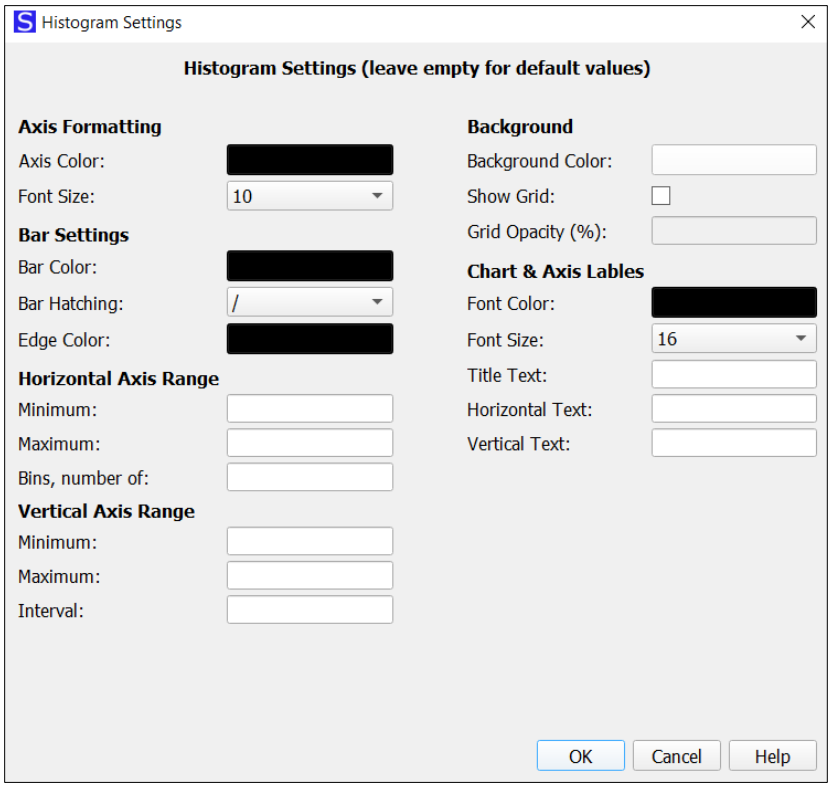
Percentage of mean: Enter the percentage for the calculation and select whether to add on both sides of the mean.

Prediction interval: Enter the prediction interval for the calculation as a percentage.

Histograms Graph Settings

Click on the 'graph settings' button from the Histogram menu to open the histograms graph settings menu (shown below). Many of the settings have default values which can be modified in settings > Graph Settings > Histogram. Once a Histogram is created, its graph settings will not be affected by future changes to the default graph settings.

Histograms Graph Settings menu



Axis Formatting: Controls the color of the graph border (and the grid, if enabled) and the font size of the axis tick mark.

Bar Settings: Controls the border color, hatch marks and fill color of the histogram bars.

Horizontal Axis Range: Controls the range and number of bins. Using a custom horizontal axis range causes the 'start the horizontal axis from 0' option from the main menu to be disabled.

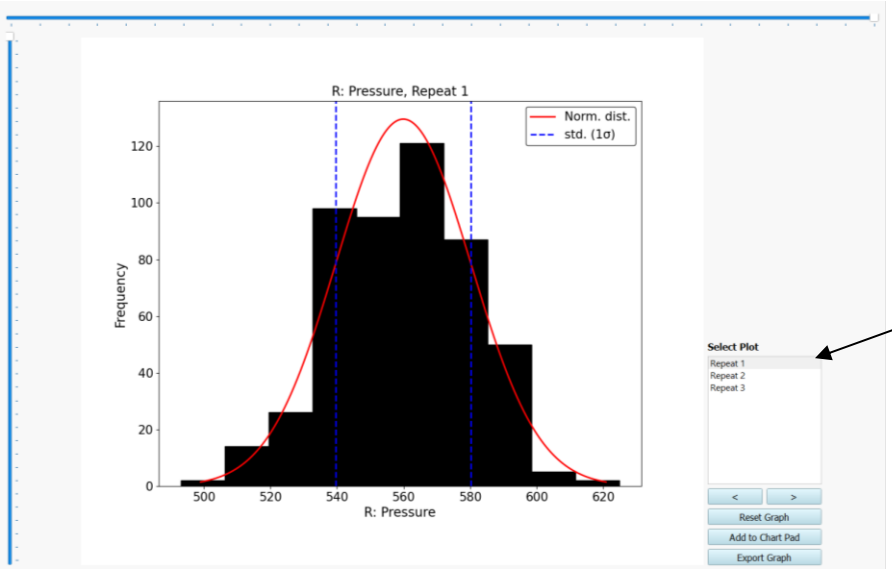
Vertical axis range: Controls the range and interval of the vertical axis ticks.

Background: Controls the background color and the grid opacity.

Chart & Axis Labels: Controls the color of the axis ticks and axis label text as well as the size of axis label text. Furthermore, custom axis and plot title texts may be entered. **To remove the default axis titles and leave them empty, enter none as the axis title.**

Navigating the results

An example histogram plot is shown below. switch between graphs using the selection panel on the right. In the current example we have added a normal curve and lines of 1 standard deviation to the plot.



Reset Graph: Resets the graph view and axes scale of the currently displayed graph to its default view and scale.

Add to Chart Pad: Adds a screenshot of the currently displayed graph to the chart pad.

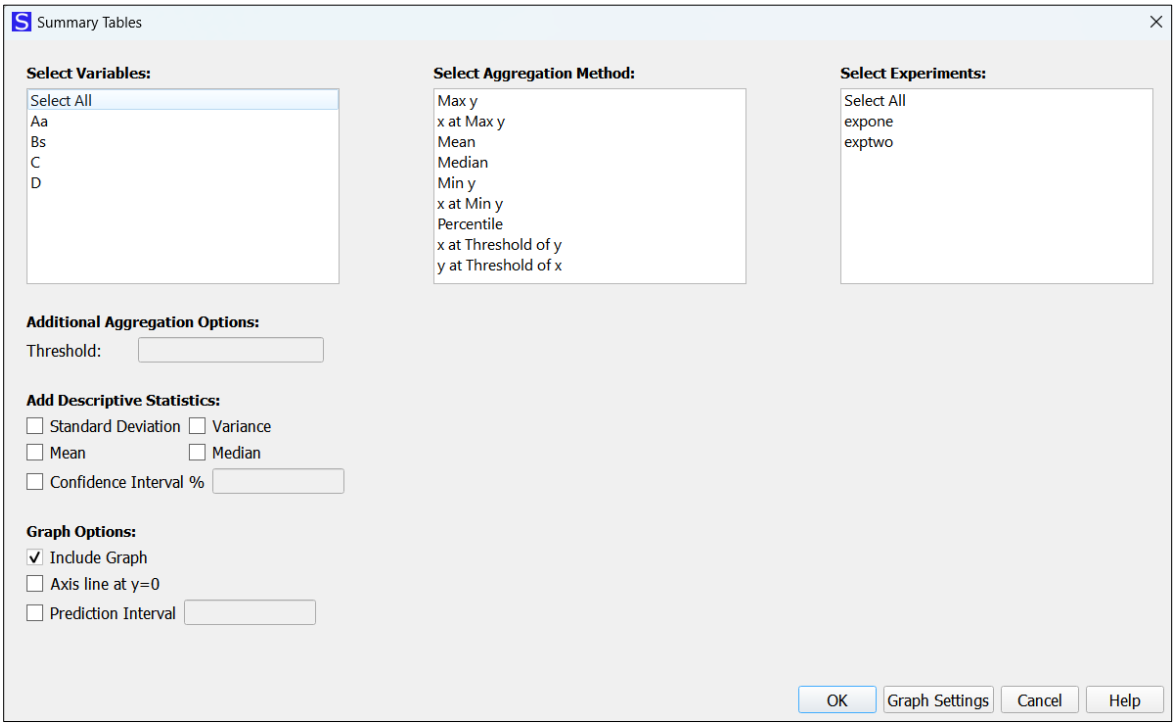
Export Graph: Save the graphs as jpegs or export the current graph to PowerPoint or Excel.

7.4. Summary Tables

The summary table analysis is used to summarize the data by aggregating each individual repeat of a variable (i.e. a single time series) into a single value (see illustration overleaf). This is useful in scenarios where you want to determine the minimum, maximum or mean value of the time series or to determine the time at which a threshold was passed. The analysis produces both a table and a graph of the summarized data. You can also add additional statistics to the table and a prediction interval to the graph. To perform the analysis:

1. In analysis mode, open the summary tables menu (Main > Statistical Tests > Summary Tables).
2. Select at least 1 variable from the 'Select Variables' panel.
3. Select an Aggregation method and if the chosen aggregation method requires it, enter a threshold value.
4. Select any of the additional settings explained below.

Summary Tables menu



Select Aggregation Method: Aggregation summarizes an individual repeat of a variable (i.e. a single time series) as a single value. Swift-Stat provides 9 aggregation methods. Note that x is the independent variable (left most column) and y is each of the variables selected for the analysis. For more detailed information see the below section, 'How the Data is Processed'.

- Max y: The maximum value of y.
- x at Max y: Finds the maximum value of y and returns the corresponding value of x from the same row.

- Mean: The sum of y values divided by the number of values.
- Median: The middle value in a sorted list of the y values.
- Min y: The minimum value of y.
- x at Min y: Finds the minimum value of y and returns the corresponding value of x from the same row.
- Percentile: Finds the value corresponding to the user selected percentile.
- x at threshold of y: Scans the y data from the beginning to find the position at which the threshold is crossed (in either direction) and returns the corresponding value of x from the same row.
- y at threshold of x: Scans the x data from the beginning to find the position at which the threshold is crossed (in either direction) and returns the corresponding value of y from the same row.

Select Experiments: If 'multi-experiment' is set up, your experiments will appear here. If you select experiments from the 'Select Experiments' panel, then instead of calculating the summary table for all repeats, a separate summary table will be calculated for each experiment, using only the repeats belonging to each experiment.

Aggregation Threshold: Enter the threshold for aggregation methods 'x at threshold of y' and 'y at threshold of x' or the percentile for the percentile aggregation method.

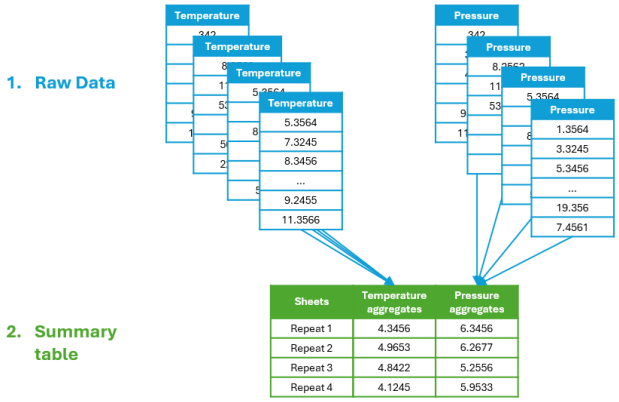
Add descriptive Stats: Add the mean, median, standard deviation, variance or a confidence interval for each variable in the summary table. If the confidence interval is enabled, enter percentage for which to calculate the interval. The confidence interval is calculated for each variable using the aggregate data from the summary table. In the example below, the 4 temperature aggregates would be used to calculate the confidence interval for the variable called temperature.

Graph Options:

- Include graph: Select to create a graph of the table data.
- Axis line at y=0: Add a horizontal line to the graph at y=0.
- Prediction Interval: Add a prediction interval to the graph. If enabled, enter a percentage for which to calculate the interval. The prediction interval is calculated for each variable using the aggregate data from the summary table. In the example overleaf, the 4 temperature aggregates would be used to calculate the prediction interval for the variable called temperature.

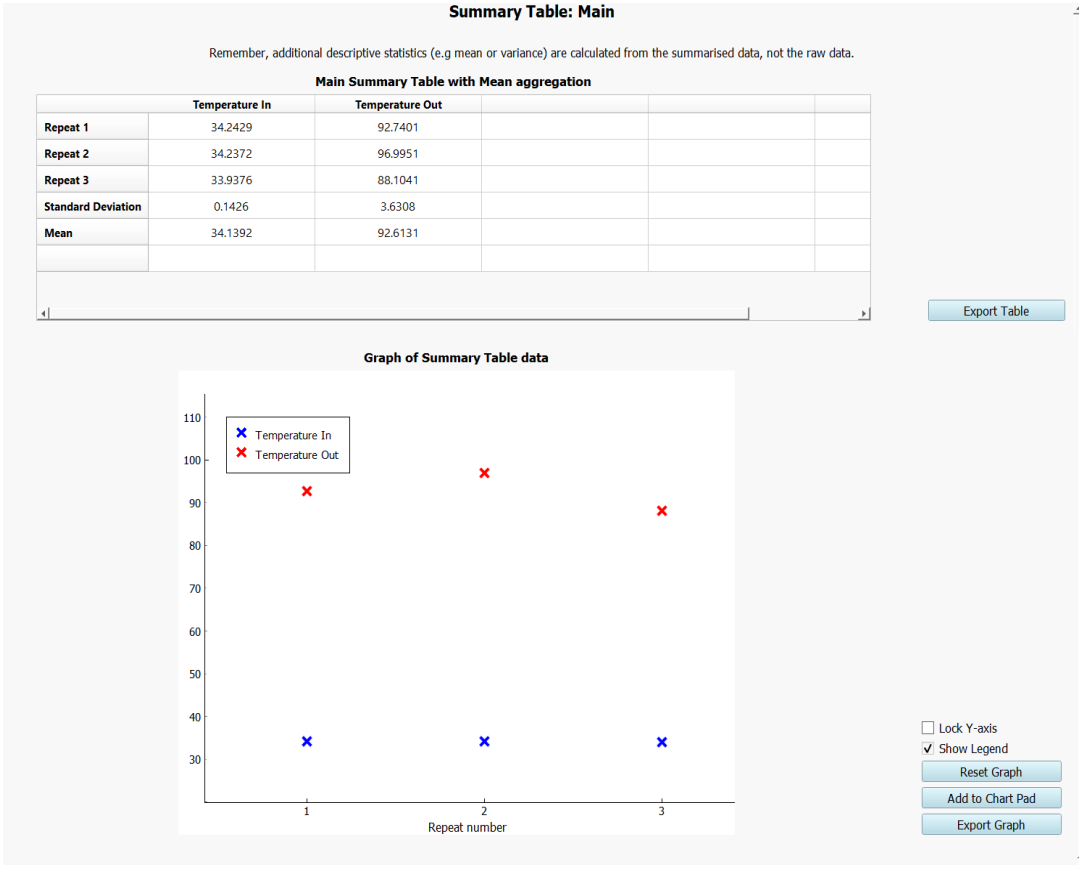
The illustration overleaf shows the process of summarizing data consisting of an experiment with 2 variables and 4 repeats.

Illustration of the process of computing a summary table



Navigating the results

An example result from a summary table analysis is shown below. The summarized data is presented in both the table and the graph. The table includes additional statistics of mean and standard deviations, which were calculated from the summarized values displayed in the table. The table's title indicates the type of aggregation that was applied. Had 'multi-experiment' been set up, the analysis would have been performed separately for each experiment using the repeats assigned for each experiment. The interactive graph enables you to scroll and zoom in with your mouse and you can hide individual plots by clicking on them in the legend.



Export Table: Export the summary table to an Excel or CSV file.

Show Legend: Toggles the visibility of the legend in the currently displayed graph.

Lock Y-Axis: Locks the Y-axis scrolling and zooming, allowing you to scroll and zoom on the X-axis without moving the Y-axis.

Reset Graph: Resets the graph view and axes scale of the currently displayed graph to its default view and scale.

Add to Chart Pad: Adds a screenshot of the currently displayed graph to the chart pad.

Export Graph: Save the graphs as jpegs or export the current graph to PowerPoint or Excel.

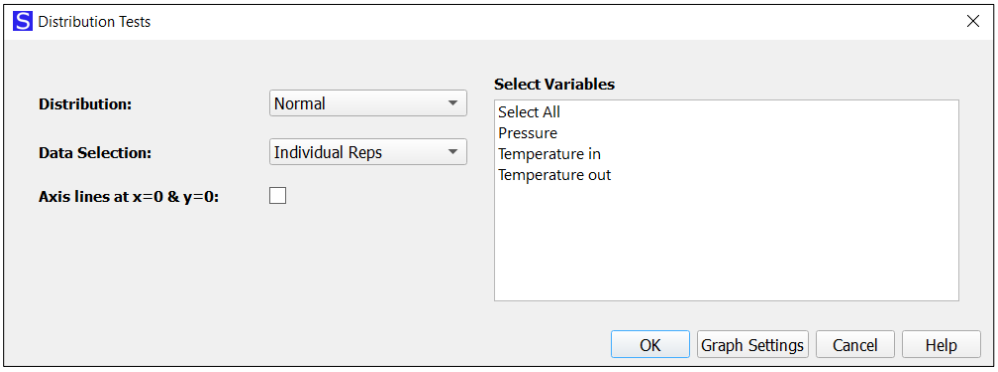
8. Statistical Tests

8.1. Distribution Tests

The distribution test determines whether the data follows a specific probability distribution, such as normal, or Poisson. Swift-Stat enables you to qualitatively determine the distribution of your data by plotting a Quantile-Quantile plot of the data against a theoretical distribution. If your data fits the theoretical distribution, the plot should be an approximately straight line, closely matching the theoretical distribution.

1. In analysis mode, open the distribution test menu (Main > Statistical Tests > Distribution Tests).
2. Select variables from the 'Select Variables' panel for your distribution test.
3. Select the theoretical distribution you want to test against.
4. Select any of the additional settings explained below.

Distribution test menu



Distribution: Select the theoretical distribution against which to test your data. For example, if you want to test whether your data is normally distributed, select 'Normal'. Swift-Stat features 6 different distributions; normal, exponential, Poisson, gamma, Weibull and uniform.

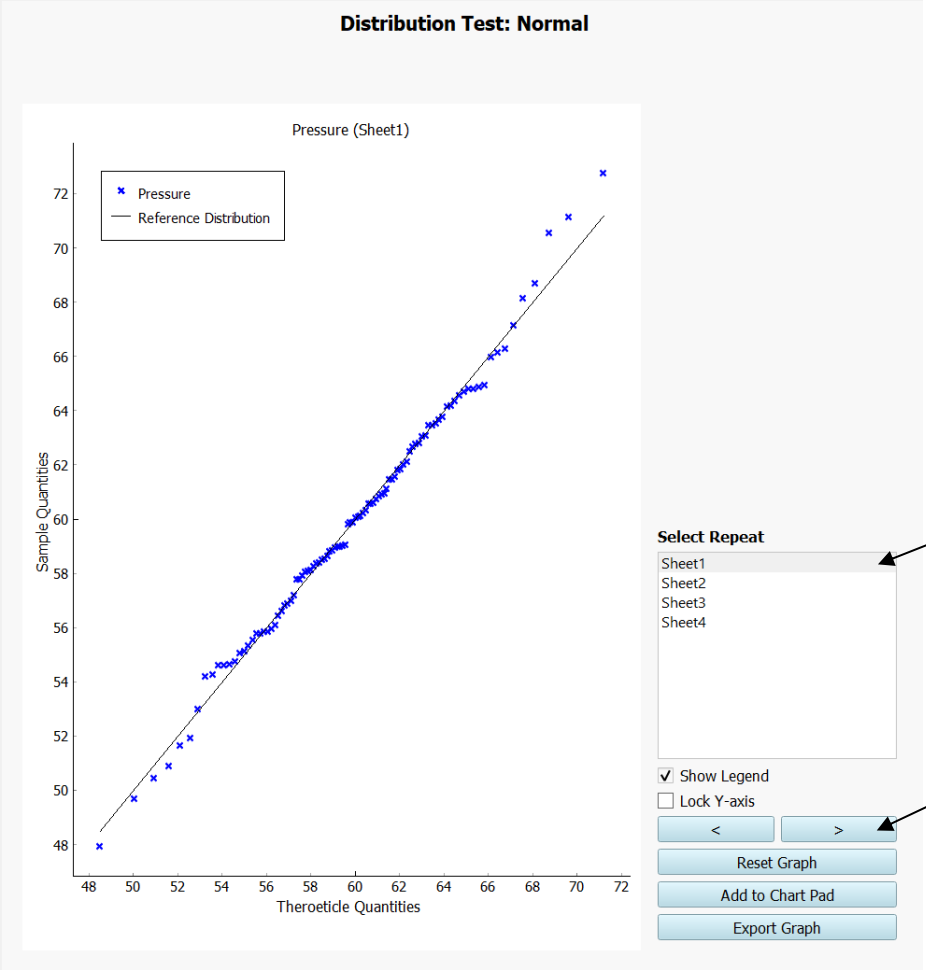
Data Selection: Choose whether to perform the analysis on individual repeats of the variables, the average of the repeats or both. **If 'multi experiment' is set up, the average will include the average for each experiment.**

Axis lines at x=0 and y=0: Adds horizontal and vertical lines to the graph at x=0 and y=0 respectively.

Navigating the results

An example result from a distribution test is shown below. The heading indicates the theoretical distribution against which the data was tested. The test output is a quartile-quartile plot showing the theoretical reference distribution and the data being tested, for each repeat of each variable. If your data fits the distribution, its plot will be an approximately straight line, closely matching the theoretical distribution. In the current example the data was tested against a theoretical normal

distribution and the quartile-quartile plot demonstrates a good fit. Switch between repeats using the selection panel on the right and switch between variables using the 2 arrows at the bottom right. The interactive graph enables you to scroll and zoom in with your mouse and to hide plots by clicking on them in the legend.



Show Legend: Toggles the visibility of the legend in the currently displayed graph.

Lock Y-Axis: Locks the Y-axis scrolling and zooming, allowing you to scroll and zoom on the X-axis without moving the Y-axis.

Reset Graph: Resets the graph view and axes scale of the currently displayed graph to its default view and scale.

Add to Chart Pad: Adds a screenshot of the currently displayed graph to the chart pad.

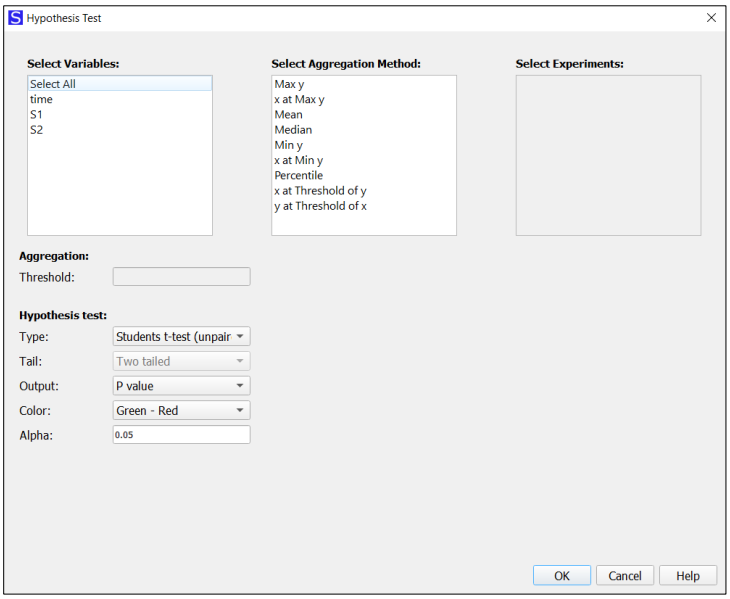
Export Graph: Save the graphs as jpegs or export the current graph to PowerPoint or Excel.

8.2. Hypothesis Tests

The hypothesis test evaluates whether differences in the data are statistically significant or a result of random chance. The test outputs a heatmap of the p-value for each variable pair from the variables selected for the analysis, thereby assessing the significance of the differences between different variables. If multi-experiment is set up (settings > Multi Experiments) the hypothesis test can also test the significance of the difference between the same variable across different experiments. To Perform the analysis:

1. In analysis mode, open the statistical tests menu (Main > Statistical Tests > Hypothesis Tests).
2. Select 2 or more variables from the 'Select Variables' panel on the left.
3. Select an aggregation method and if the chosen aggregation method requires it, enter a threshold value.
4. Select the type of hypothesis test to perform.
5. Select any of the additional settings explained below.

Correlation Analysis menu



Select Aggregation Method: Aggregation summarizes an individual repeat of a variable (i.e. a single time series) as a single value. Swift-Stat provides 9 methods to aggregate the data. Note that x is the independent variable (left most column) and y is each of the variables selected for analysis. For more detailed information see the below section, 'How the Data is Processed'.

- Max y: The maximum value of y.
- x at Max y: Finds the maximum value of y and returns the corresponding value of x from the same row.
- Mean: The sum of y values divided by the number of values.
- Median: The middle value in a sorted list of the y values.

- Min y: The minimum value of y.
- x at Min y: Finds the minimum value of y and returns the corresponding value of x from the same row.
- Percentile: Finds the value corresponding to the user selected percentile.
- x at threshold of y: Scans the y data from the beginning to find the position at which the threshold is crossed (in either direction) and returns the corresponding value of x from the same row.
- y at threshold of x: Scans the x data from the beginning to find the position at which the threshold is crossed (in either direction) and returns the corresponding value of y from the same row.

Select Experiments: If 'multi-experiment' is set up, your experiments will appear here. If you select experiments from the 'Select Experiments' panel, then instead of calculating the hypothesis test for pairs of variables across all repeats, a separate heatmap is calculated for each experiment, using only the repeats belonging to that experiment. Furthermore, for each selected variable, Swift-Stat calculates the hypothesis test between the repeats of pairs of experiments (or all experiments at once in the case of ANOVA or Kruskal-Wallis tests), thereby comparing a single variable across different experiments.

Threshold: Enter the threshold for aggregation methods 'x at threshold of y' and 'y at threshold of x' or the percentile for the percentile aggregation method.

Type: Swift-Stat features six of the most commonly used hypothesis tests, with key points on their applications summarized in the table below:

Type	Application
Student's t-test (unpaired)	Two samples of continuous data with similar variances. Each sample should be approximately normally distributed.
Student's t-test (paired)	Two samples of continuous data from the same population. Each sample should be approximately normally distributed.
Welch t-test	Two samples of continuous and independent data. Each sample should be approximately normally distributed.
Mann–Whitney U test	Two samples of independent data. It does **not** require the data to be normally distributed.
ANOVA	Two **or more** samples of continuous data with similar variances. Each sample should be approximately normally distributed.
Kruskal-Wallis	Two **or more** samples of continuous data. It does **not** require the data to be normally distributed.

Tail: Currently only two tail analyses are supported.

Output: Select whether to show the P value or the test statistic.

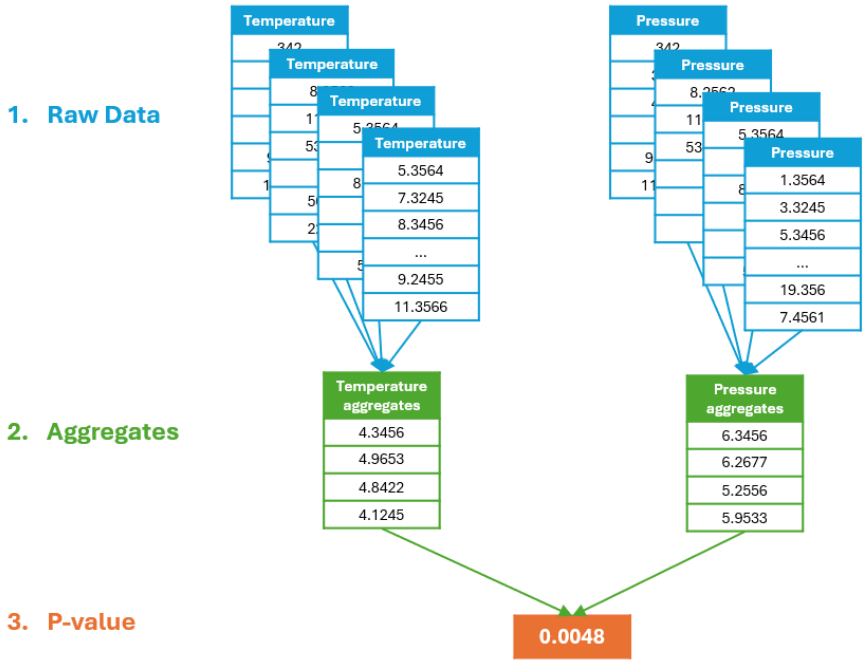
Color: Select the colormap for the heatmap. The cells are colored according to whether or not the p-value is less than alpha.

Alpha: The threshold significance level for deciding if results are statistically significant. Since a low P-value represents a higher chance of statistical significance, a low alpha results in a more stringent test.

How the Data is Processed

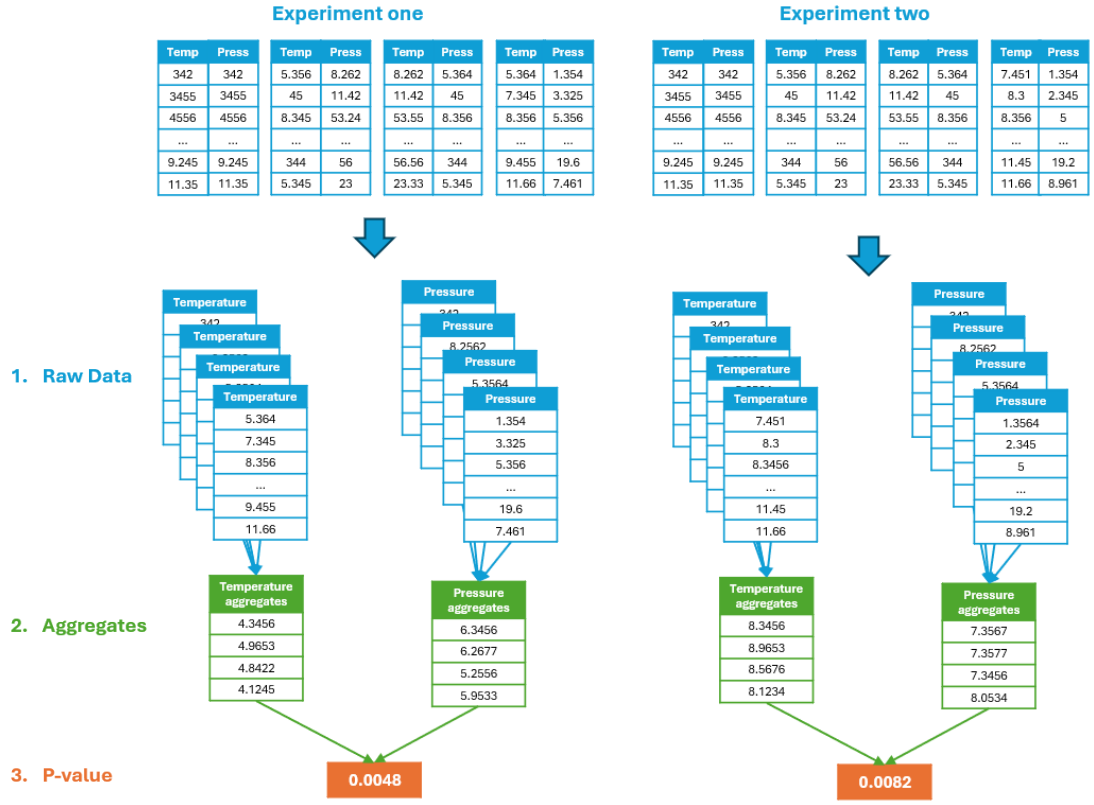
Suppose we conduct an experiment in which the temperature and pressure of a steam boiler are measured over time, resulting in a temperature time series and a pressure time series. We conducted the experiment 4 times, achieving 4 repeats for temperature and 4 for pressure. To perform the hypothesis test, Swift-Stat aggregates each individual time series into a single value giving a total of 4 temperature aggregate values and 4 pressure aggregate values. Swift-Stat then calculates the P-value of the temperature and pressure aggregates, thereby testing if the difference between the temperature aggregates and pressure aggregates is statistically significant. This is a trivial example as of course temperature will be different to the pressure, but it demonstrates the concept. The process is illustrated below and is performed for each variable pair for all of the variables selected for the hypothesis test. The results are then displayed on a heatmap. If performing ANOVA or Kruskal-Wallis tests, which can be applied to more than 2 datasets, the process is identical but extended to test all selected variables in one test.

The process of computing a Hypothesis test for a single experiment



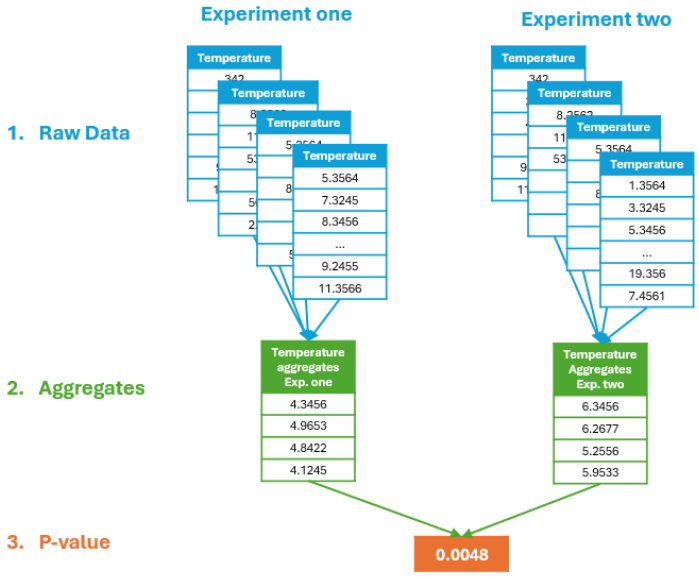
Now suppose we perform the same experiment with 4 repeats on a new steam boiler. We can set up a 'multi-experiment' in Swift-Stat and assign the 4 repeats for each steam boiler to a different experiment. Now when we run the hypothesis test, we get two results. The first result, illustrated below, is a heatmap for each experiment, showing the P values of the variable pairs, obtained using only the repeats belonging to that experiment. This provides the statistical significance of differences between variable pairs in each individual experiment.

The process of computing a Hypothesis test with multi-experiment: Comparing variables



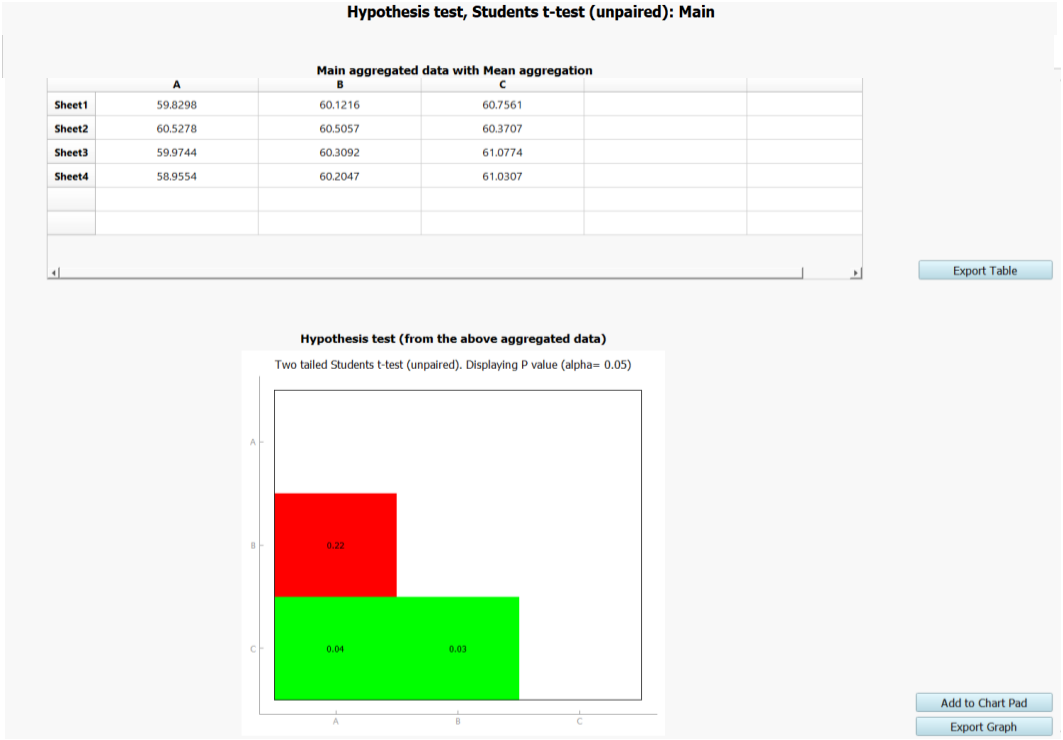
For the second part of the analysis, Swift-Stat takes the 4 repeats per experiment of a single variable and aggregates them for both experiments. The result is two lists of aggregates for a single variable, one from experiment one and the other from experiment two. The illustration below demonstrates this process for the temperature variable. Swift-Stat then calculates the P-value of the two temperature aggregates to determine the probability that there is no statistical significance in the difference between the temperature results of experiment one and two. If there are more than two experiments, this is performed for all experiment pairs, resulting in a heatmap for each variable. If performing ANOVA or Kruskal-Wallis tests, which can be applied to more than 2 datasets, the process is identical but is extended to test all experiments in a single test.

The process of computing a Hypothesis test with multi-experiment: Comparing experiments



Navigating the results

An example result from a hypothesis test is shown below. The heading indicates that the hypothesis test is for a single experiment. The table shows the aggregate values for each variable on which the hypothesis test was performed. The heatmap shows the P-value for each variable pair. Since alpha=0.05, p-values under 0.05 are colored in green and those over 0.05 in red.



Add to Chart Pad: Adds a screenshot of the currently displayed graph to the chart pad**.**

Export Graph: Save the graphs as jpegs or export the current graph to PowerPoint or Excel.

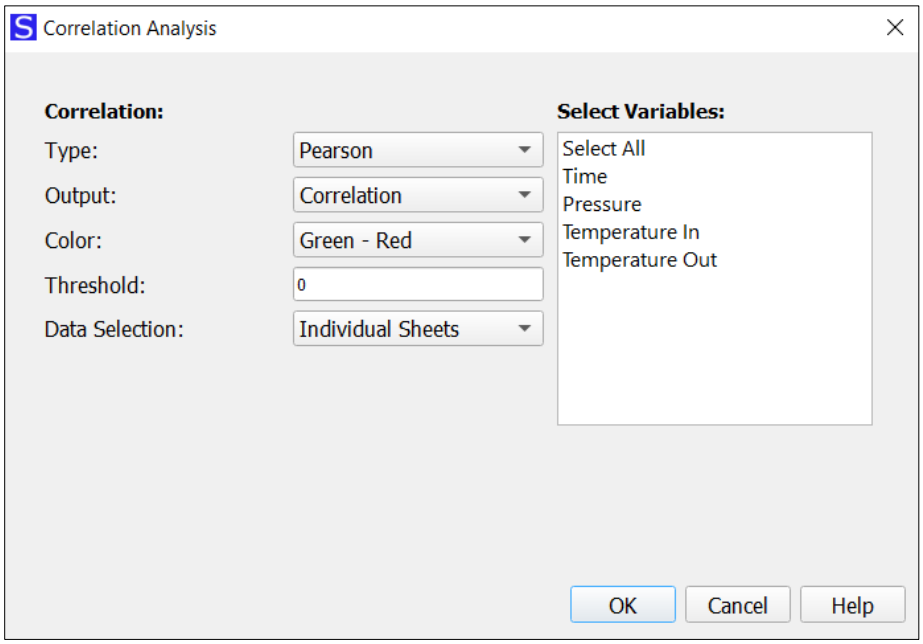
9. Correlation & Regression

9.1. Correlation

The correlation analysis measures the strength and direction of the relationship between two variables. It produces a heatmap displaying either the correlation coefficients or the p-values for each pair of selected variables. To Perform the analysis:

1. In analysis mode, open the Correlation analysis menu (Main > Correlation & Regression > Correlation).
2. Select 1 of more variables from the 'Select Variables' panel.
3. Select the type of correlation to apply.
4. Select any of the additional settings explained below.

Correlation Analysis menu



Type: Choose from three types of correlation; Pearson, Spearman Rank and Kenda's Tau. Pearson correlation measures the strength of linear relationships between two variables. For example, the relationship $y = x$ has a Pearson correlation of 1, while $y = x^2$ has a correlation of less than 1, since it is nonlinear. For nonlinear relationships, Spearman Rank and Kenda's Tau are often more appropriate because they assess whether x increases (or decreases) with y without assuming a specific relationship and so both $y=x$ and $y=x^2$ would yield a correlation of 1.

Output: Select whether to display the correlation coefficients or the p-values. The p-value indicates the probability that the observed correlation occurred by chance. A low p-value suggests that the correlation is statistically significant and unlikely to be due to random variation.

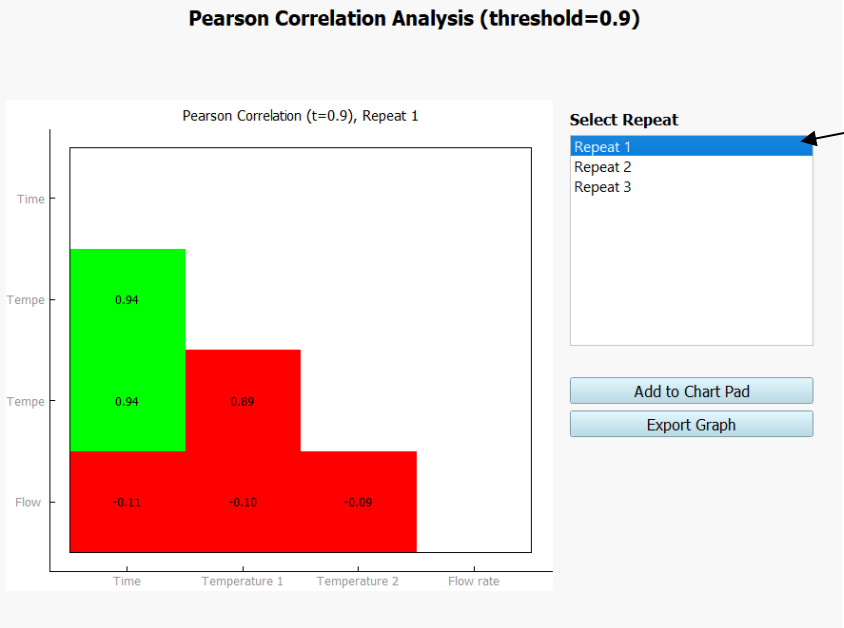
Threshold: Select a threshold for the results. values on either side of this threshold will be displayed in different colors on the heatmap to aid visual interpretation. This setting affects only the visualization and does not influence the underlying calculations.

Color: Select the colormap for the heatmap (for the threshold).

Data Selection: Choose whether to perform the analysis on individual repeats of the variables, the average of the repeats or both. **If a 'multi-experiment' is set up, the average will include the average for each experiment.**

Navigating the results

An example result from a correlation analysis is shown below. The heading indicates the type of correlation used and the threshold applied. The heatmap displays the correlation results for each pair of variables. You can switch between repeats using the panel on the right. To view the p-values instead of correlation coefficients, double-click on Analysis Results in the main menu, change the output setting to P-value, and click OK.



Add to Chart Pad: Adds a screenshot of the currently displayed graph to the chart pad.

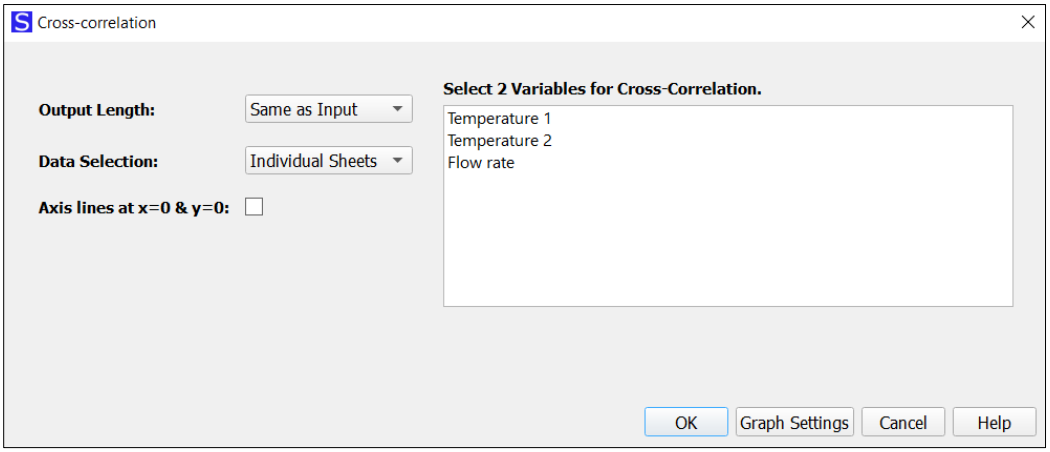
Export Graph: Save the heatmaps as jpegs or export the current heatmap to PowerPoint or Excel.

9.2. Cross-Correlation Analysis

The Cross-Correlation analysis measures the similarity of two signals as a function of the displacement of one relative to the other. When the two signals are similar but with a phase difference the analysis will determine the lag between them. The analysis generates a graph of the two signals and a graph of correlation vs lag. To Perform the analysis:

1. In analysis mode, open the Cross-Correlation analysis menu (Main > Correlation & Regression > Cross-Correlation).
2. Select 2 variables from the panel on the right for the cross-correlation analysis.
3. Select any of the additional settings explained below.

Cross-Correlation Analysis menu



Output Length: To compute the cross-correlation, one data series is shifted over the other. If both data series are the same length, the number of shifts is twice the length of the original data series. You can choose to display the full-length correlation or limit the output to the length of the original data series. If the data series are not the same length only full length may be selected.

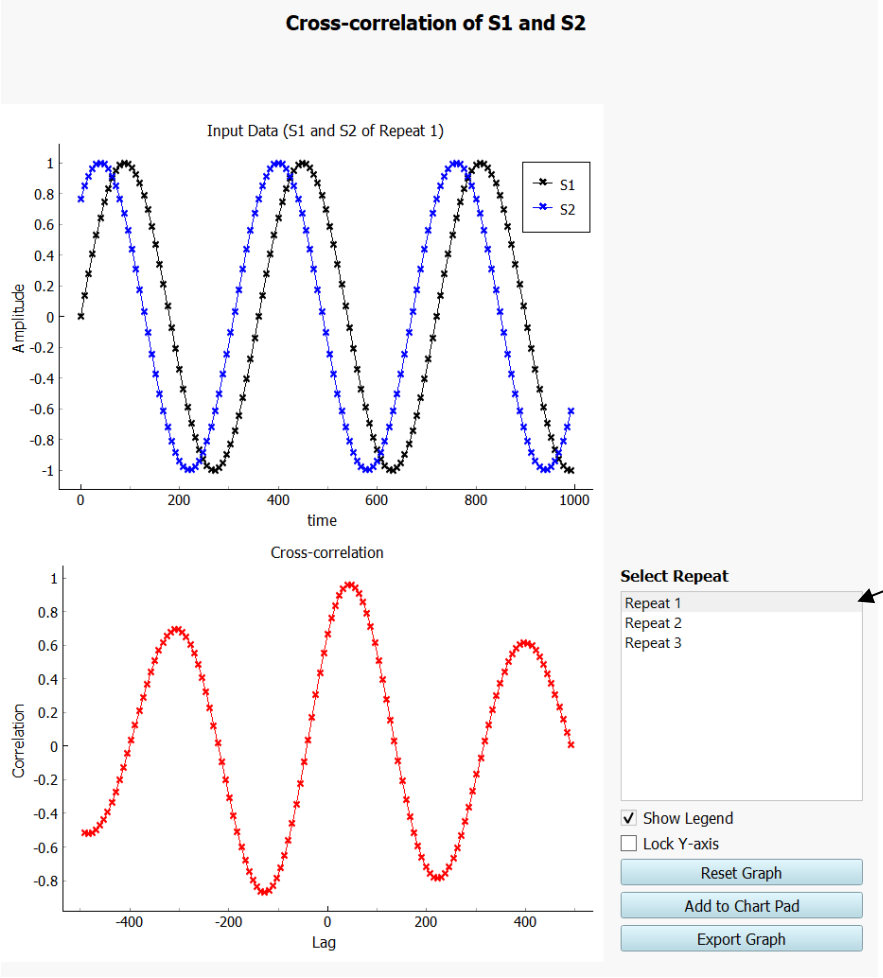
Data Selection: Choose whether to perform the analysis on individual repeats of the variables, the average of the repeats or both. **If a 'multi-experiment' is set up, the average will include the average for each experiment.**

Axis lines at x=0 and y=0: Adds horizontal and vertical lines to the graph at x=0 and y=0 respectively.

Navigating the results

An example result from a cross-correlation analysis is shown below. The heading indicates the two variables being analyzed. Two graphs are displayed: one showing the original data series, and the other showing the cross-correlation results. You can switch between repeats using the panel on the right. The interactive graphs allow you to scroll and zoom using your mouse, and you can hide individual plots by clicking on them in the legend.

In this example, cross-correlation has been applied to two out-of-phase sine waves. The analysis output shows the highest correlation at the lag corresponding to the phase shift between the two waves, with smaller correlations appearing at lags equal to plus and minus one full wavelength.



Select a repeat of the current variable to display

Show Legend: Toggles the visibility of the legend in the currently displayed graph.

Lock Y-Axis: Locks the Y-axis scrolling and zooming, allowing you to scroll and zoom on the X-axis without moving the Y-axis.

Reset Graph: Resets the graph view and axes scale of the currently displayed graph to its default view and scale.

Add to Chart Pad: Adds a screenshot of the currently displayed graph to the chart pad.

Export Graph: Save the graphs as jpegs or export the current graph to PowerPoint or Excel.

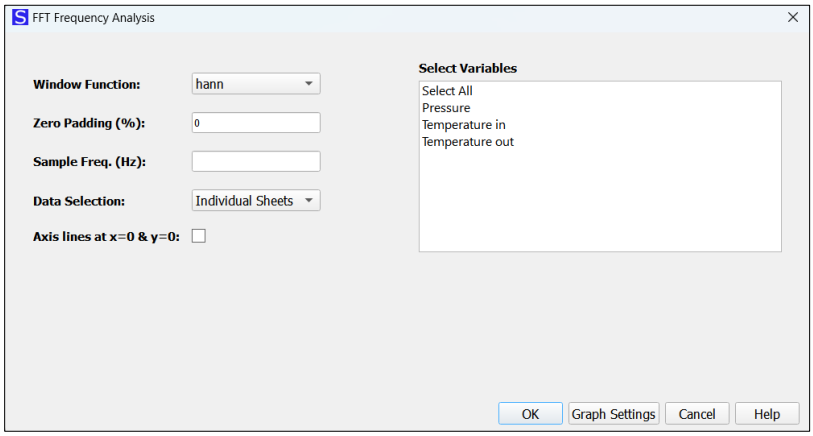
10. Other

10.1. FFT Frequency Analysis

The FFT frequency analysis converts the signal from the time domain to the frequency domain to uncover the frequency components of the signal. It generates graphs of the time and frequency domain of the data. To Perform the analysis:

1. In analysis mode, open the FFT Frequency Analysis menu (Main > Other > FFT frequency analysis).
2. select 1 of more variables from the ‘Select variables’ panel.
3. Set the sample frequency of your data.
4. Select any of the additional settings explained below.
5. To change graph or axis settings, click on graph settings.

FFT Frequency Analysis menu



Window Function: Reduce spectral leakage, which occurs when the signal being analyzed is not an integer multiple of the frequency bins, causing energy from one frequency to spread to adjacent bins. By applying a window function, the signal is tapered at the edges, minimizing leakage and improving frequency resolution. If you are uncertain which function to use, we recommend the Hann function.

Zero Padding (%): Adds extra zeros to the end of a signal before applying FFT. It does not alter the content of the signal but increases the number of FFT points, resulting in a higher frequency resolution. Zero padding is set to 0 by default.

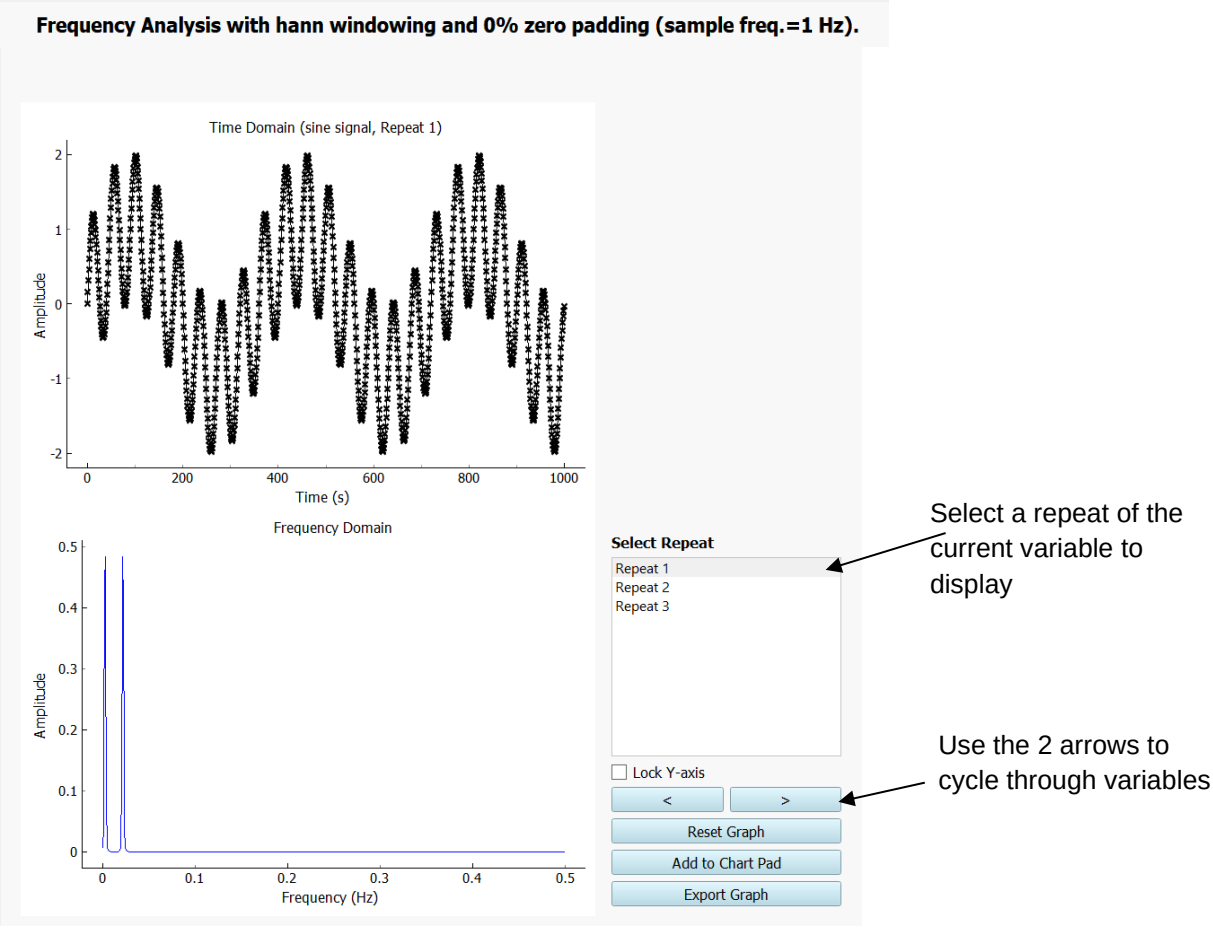
Sample Freq. (Hz): The sample frequency of your data i.e. the number of data points per second.

Data Selection: Choose whether to perform the analysis on individual repeats of the variables, the average of the repeats or both. **If a 'multi-experiment’ is set up, the average will include the average for each experiment.**

Axis lines at x=0 and y=0: Adds horizontal and vertical lines to the graph at x=0 and y=0 respectively.

Navigating the results

An example result from a FFT Frequency analysis is shown below. The heading indicates the window function, zero padding and sample frequency that were applied. There is a graph of the time domain and of frequency domain for each variable-repeat combination. Switch between repeats using the panel on the right. Switch between variables using the 2 arrows at the bottom right. The interactive graph enables you to scroll and zoom in with your mouse and to hide plots by clicking on them in the legend. In this example we have applied the FFT frequency analysis to a signal created by superimposing low frequency and high frequency sine waves. The output of the analysis shows 2 peaks in the frequency domain, corresponding to the frequencies of the 2 sine waves.



Lock Y-Axis: Locks the Y-axis scrolling and zooming, allowing you to scroll and zoom on the X-axis without moving the Y-axis.

Reset Graph: Resets the graph view and axes scale of the currently displayed graph to its default view and scale.

Add to Chart Pad: Adds a screenshot of the currently displayed graph to the chart pad.

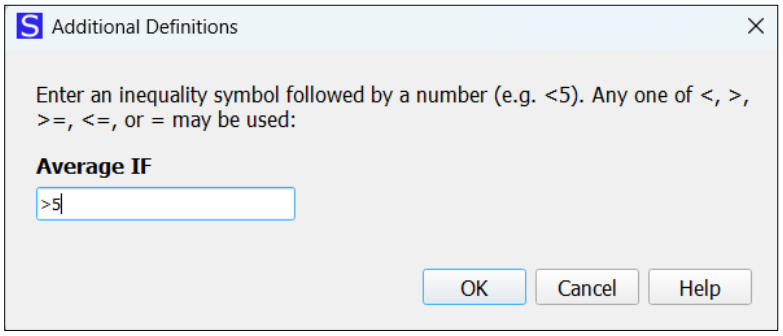
Export Graph: Save the graphs as jpegs or export the current graph to PowerPoint or Excel.

Appendix A: Row calculation and descriptive statistics submenus

'Average If',' Count If' and 'Sum If' Submenu

The 'Average If',' Count If' and 'Sum If' calculations enable you to perform calculations using only the values which meet the defined equality. 'Average If' calculates the mean average of the values which meet the equality. Count If counts the number of instances the equality is met and Sum If calculates the sum of the values that meet the equality. When used in row calculations, the calculation is performed on each row, using only cells in columns which were selected as source columns in the previous dialogue window. When used in descriptive statistics, the calculation is performed along the entire column.

The equality for the calculation is entered into the Additional definitions text box as shown below. Any of <, >, >=, <= or = may be used. For example, if a 'Sum If' calculation is created using the equality >5, then only numbers that are greater than 5 will be used for calculating the sum.



'Custom Expression' Submenu

The custom expression row calculation enables you to build your own formulae using basic arithmetic and built in functions. **Importantly, it supports referencing other columns, including those created with other row calculations.** Note that whereas other row calculations such as 'Average' and 'Standard deviation' ignore empty cells and deleted source columns, custom expressions will return nan on any row where one or more of the referenced cells is empty or a source column has been deleted.

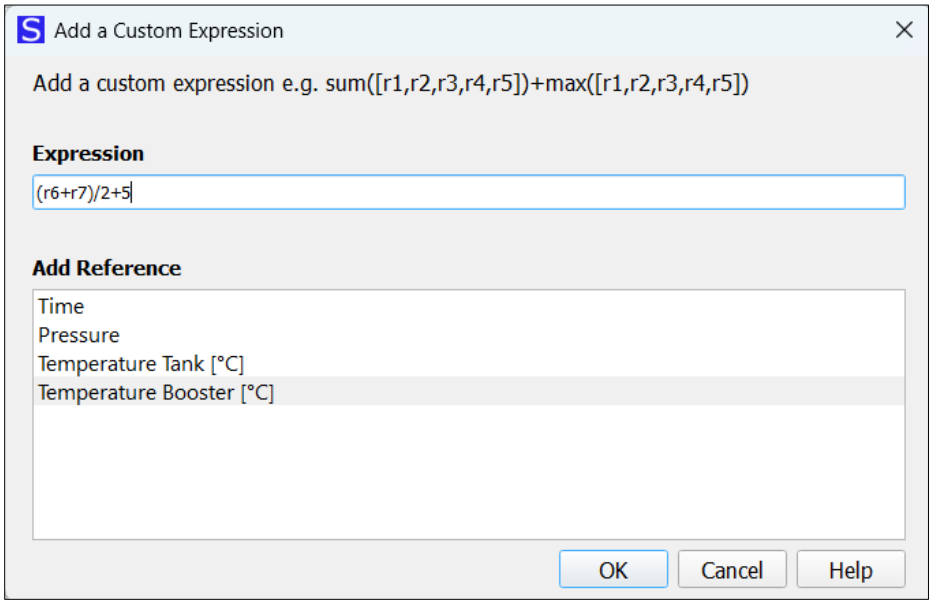
Referencing Variables

Any variables or row calculation on the sheet can be referenced in the custom expression. They are referenced by either typing the letter 'r' followed by the column ID (which appears above the variable name in the spreadsheet) or by double clicking the variable name in the 'add reference' panel on the custom expression menu.

Basic Arithmetic

The expression accepts the symbols +, -, *, /, ^ and () for addition, subtraction, multiplication, division, power and order of operation respectively. Expressions do not need to begin with an = sign, but the = sign is tolerated. The expression entered in the image below adds two variables together, divides them by 2 and then adds 5. The expression must contain at least one column reference. It cannot contain only numbers.

Custom Expression menu



Additional Functions

Custom expression row calculations may also include functions for calculating the sum, maximum, minimum, absolute, logarithm (base 10), natural logarithm, exponential function, mean average, median average and standard deviation (of population). Examples of each function are provided overleaf.

Function	Correct example	Incorrect example
Sum	sum([r2,r3,r4+7])	sum([r2,r3,r4,7])
Maximum	max([r2,r3,r4]) +7	max(r2,r3,r4)
Minimum	min([r2,r3,r4])	min([r2+r3+r4])
Absolute	abs([r2])	abs([r2,r3,r4])
Logarithm (base 10)	log10([r2])	log10([1000])
Natural logarithm	log([r2])	log(1000)
Exponential function	r1+ exp([200])	exp([200])
Mean average	mean([r2,r3,r4])	mean[r2,r3,r4,7]
Median average	median([r2,r3,r4])	median([$2,$3,$4])
Standard deviation (of population)	std([r2,r3,r4])	std=([r2,r3,r4])

General Rules

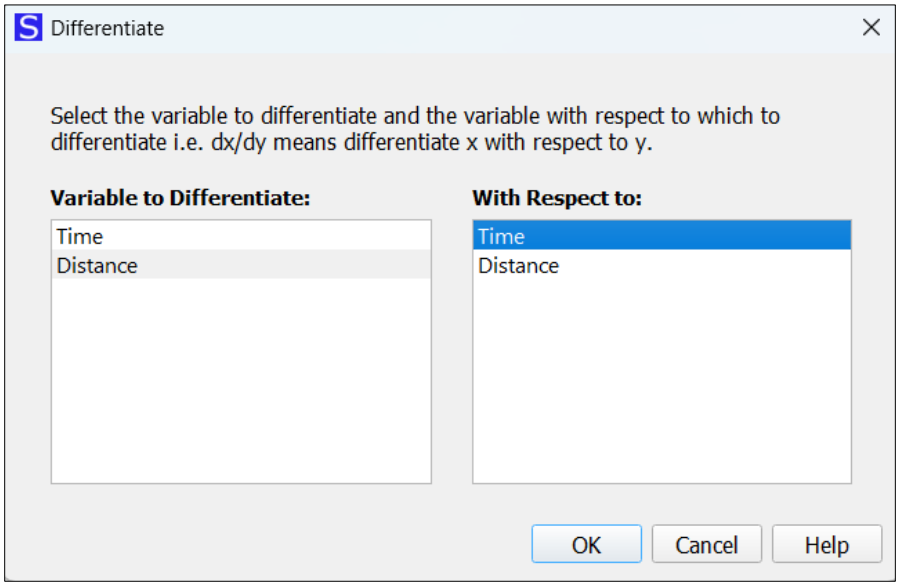
- You may reference as many columns as required, including those created by row calculations, with the exception of those created with custom expression row calculations.
- An expression must contain at least 1 column reference i.e. it cannot contain only numbers.
- Aswell as column references, you may include numbers in the expression, except for inside the sum, max, min, mean, mean and standard deviation functions.
- Note that mean([r1,r2+20]) calculates the mean of the two numbers r1 and (r2+20) NOT of the three numbers r1, r2 and 20. The same is also true for other expressions, such as standard deviation, maximum and minimum.

‘Differentiate (Calculus)’ Submenu

The ‘Differentiate (Calculus)’ row calculation numerically differentiates one variable with respect to another using the second order central differences method. For example, given the variable ‘distance’ and the variable ‘time’, you can calculate speed by selecting distance as the variable to differentiate and time as the variable with respect to which to differentiate.

Note: Numerical differentiation works best for smooth data. If the data has sharp changes or other irregularities, numerical differentiation may introduce errors. Furthermore, numerical differentiation tends to amplify noise.

Differentiation menu



‘Distributions, random’ Submenu

The ‘Distributions, random’ row calculation generates a column of random numbers that adhere to one of the following statistical distributions; exponential, normal, Poisson, uniform or uniform (integer). To create a distribution, select the desired distribution from the drop-down menu and enter the number of elements (i.e. how many rows) you want in the distribution, which must be a positive integer. Next you must define the shape of the distribution as follows:

Exponential distribution

The distribution is defined by the scale parameter, Beta (β), which is equal to the inverse of lambda (1/λ) and must be a number greater than 0. The physical meaning of Beta is the mean time between events. Examples of exponential distributions include radioactive decay and the lifetimes of electrical components.

Normal distribution

The distribution is defined by the mean, which can be any number and the standard deviation, which must be greater than 0. Examples of phenomena that are often normally distributed include sensor noise and variations in material properties e.g. tensile strength.

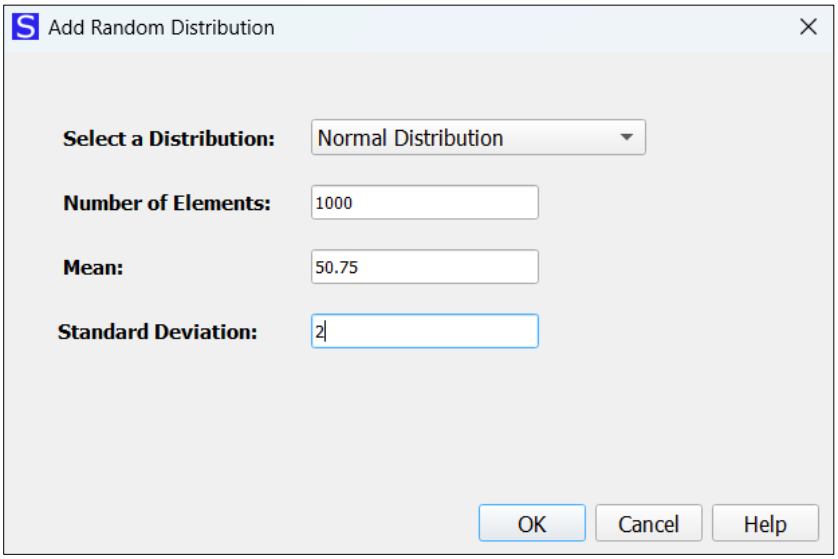
Poisson distribution

The Poisson’s distribution is defined by lambda (λ) which represents the average rate of events per unit time interval. Poisson’s distributions can be used to model defects on a production line or anomalies in an experiment.

Uniform distribution (including integer)

The distribution is defined by a minimum and maximum, where the maximum must be larger than the minimum. Examples of uniform distributions include numbers obtained from rolling a die and from playing the lottery.

Random Distribution menu

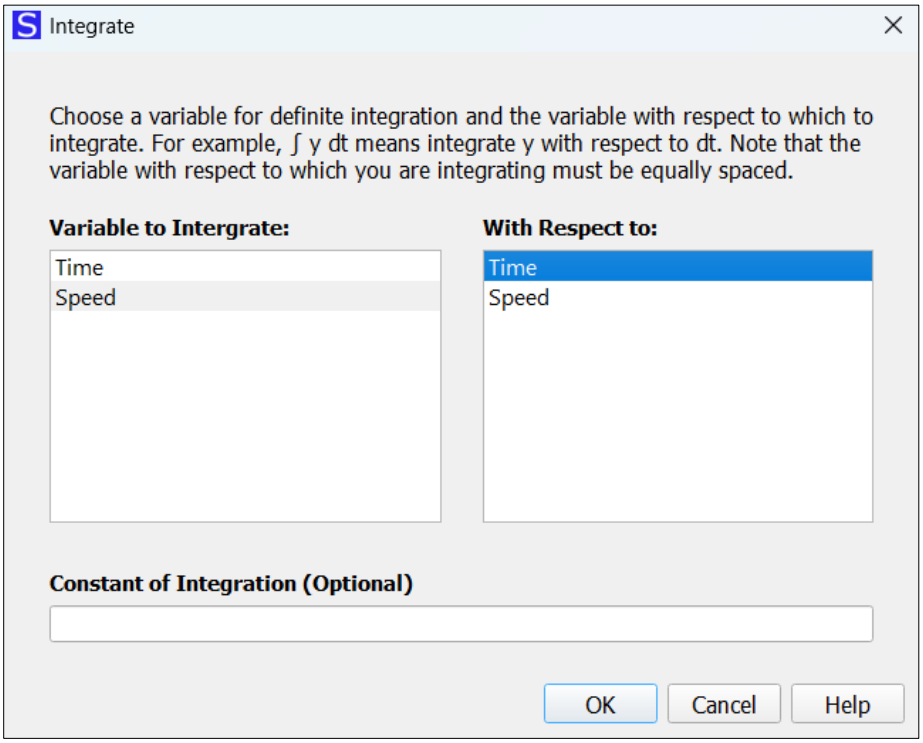


‘Integrate (Calculus)’ Submenu

The ‘Integrate (Calculus)’ row calculation integrates one variable with respect to another using the composite trapezoidal rule. For example, the variable ‘speed’ and the variable ‘time’, can be used to calculate the distance by selecting speed as the variable to integrate and time as the variable with respect to which to integrate.

Note: The ‘Integrate (Calculus)’ row calculation can only be used to integrate with respect to a uniformly spaced variable.

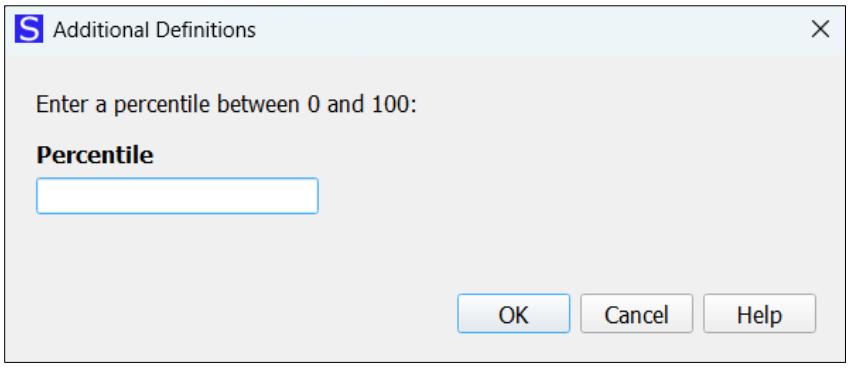
Integration menu



‘Percentile’ Submenu

The ‘Percentile’ calculation calculates the user defined percentile along each row or column of the selected source variables (for row calculations and descriptive statistics respectively). The percentile is defined as a percentage and must be between 0 and 100. A percentile of 50 is equal to the median average.

Percentile submenu



‘Sequence’ Submenu

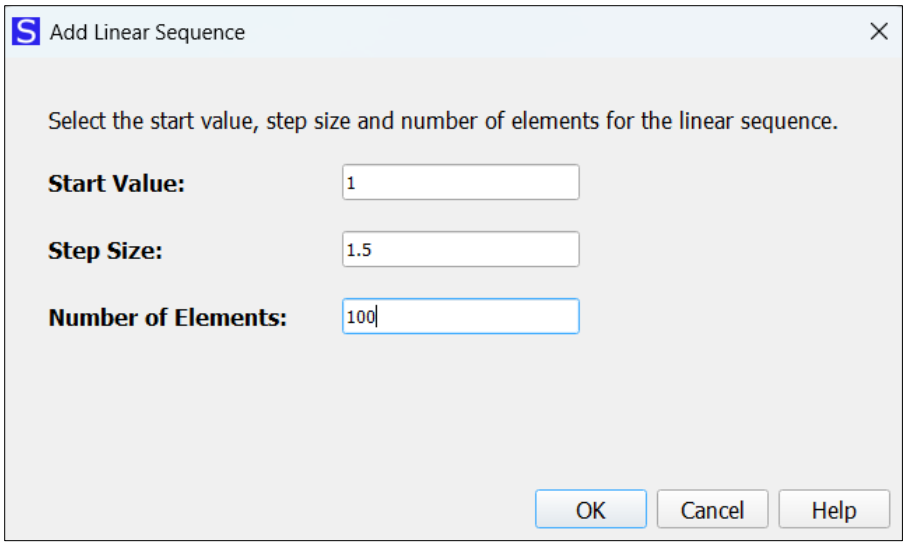
The sequence calculation creates a column of linearly increasing (or decreasing) numbers defined by a start value, a step size and the number of elements. The definition provided in the figure below will result in the sequence 1, 2.5, 4, 5.5, 7, 8.5…149.5. The rules for filling the sequence settings are as follows:

Start value: Enter the first value of the sequence. The number can be positive or negative and can include decimal places.

Step size: Enter the amount by which the sequence increases between adjacent values. The number can be positive or negative and can include decimal places.

Number of Elements: Enter the length of the sequence. The number must be a positive integer.

Sequence submenu



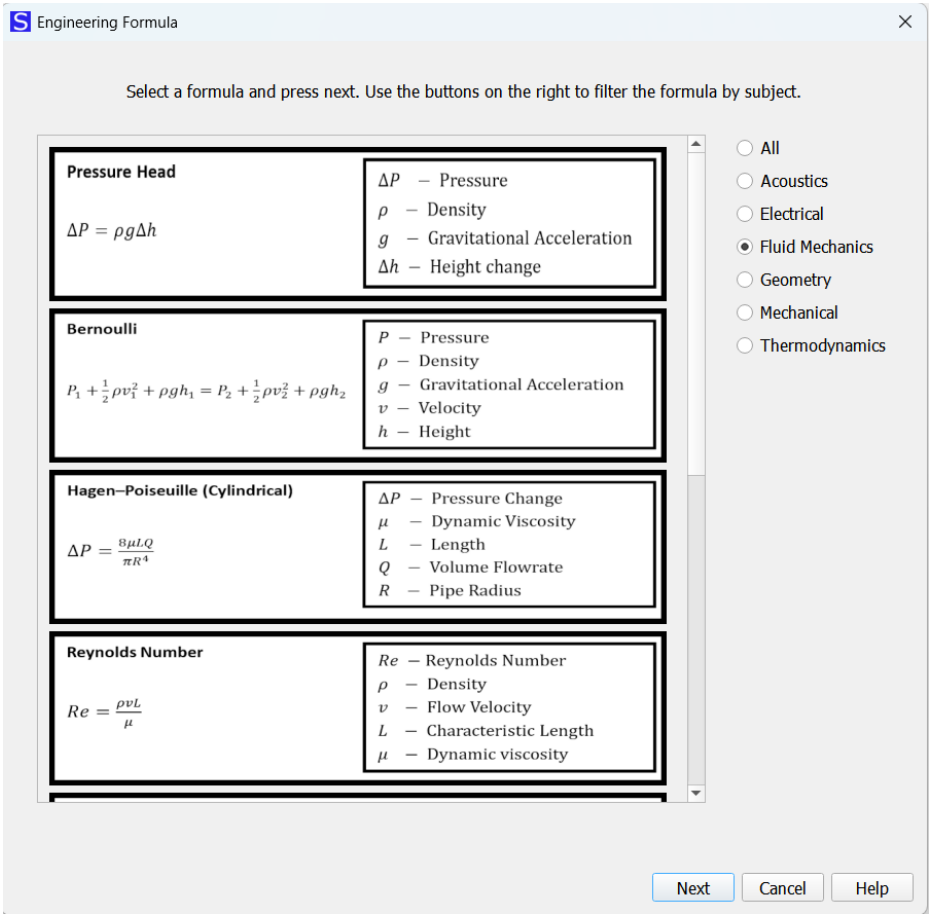
Engineering Formula

The software includes a comprehensive library of over 120 built-in engineering formulas, enabling you to perform a wide range of engineering calculations using any combination of input and output units. Before applying an engineering formula, you must first specify units for all input variables via the settings tab (see section 3.8). Once units have been configured:

1. Select an empty column in your dataset.
2. Click the Row Calculations button.
3. Choose Engineering Formula from the menu.

After clicking Continue, the Formula Selector will open, where you can browse and select the desired formula.

Formula selector



After selecting a formula and pressing ok, the formula builder will open (see next page). It should be used as follows

- Select a variable to solve for (i.e. the left-hand side of the formula) along with the desired output unit.

- For each of the input variables (i.e. the right-hand side of the formula) either choose a variable or enter a constant. Only spreadsheet variables that have been assigned a matching physical quantity will be available for each input variable. If you enter a constant, it is assumed to be in SI units.

- The Auto Fill button can be used to automatically assign variables in cases where your spreadsheet contains only one variable with the correct physical quantity for each formula input variable.

- To avoid an excessive number of physical quantities in the unit settings (section 3.8), quantities with identical units have been grouped together under a common category. For example, radius, diameter, length and distance should all be assigned the physical quantity length. Within the formula builder, if the input variable has a different name to the physical quantity that should be assigned in the settings, the physical quantity will be shown in brackets after the variable name e.g. Initial Speed (Velocity).

Formula builder