

User manual



Table of contents

Section	Page
1 Getting started	3
About this document	4
About dataWizard	5
2 Introducing dataWizard	6
Opening dataWizard	7
Key components	8
Quick start: 4 easy steps	9
Buttons	10
3 Using dataWizard	14
Step 1 Creating a new template	15
Step 2 Importing data	20
Step 3 Validating data	21
Step 4 Cleaning up data	23
Data quality metrics	28
Activity log	29
4 Data cleanup rules explained	30
5 Template columns explained	39



Getting started

Section 1



Getting started about this document

Purpose of this user manual

This user manual contains key information on how to use dataWizard. It includes a description of the solution, its functionality and capabilities, and step-by-step instructions to access and use the solution.

Applicability of this user manual

This user manual is applicable to dataWizard v1.00. For inquiries related to dataWizard or this document, please contact us at support@soleox.com.

Tip: Verify that this user manual is applicable to the version of dataWizard you are using.





Getting started about dataWizard

What is dataWizard and what is it for?

dataWizard is a data manipulation solution that enables users to easily perform complex data import, validation and cleanup tasks in very little time and with no coding.

Tip: dataWizard is built in Microsoft Excel and is compatible with all CSV and Microsoft Excel files.

Who can use dataWizard?

dataWizard can be used by anyone who has access to Microsoft Excel (see System Requirements below for more info). dataWizard does not require any programming background.

Who should use dataWizard?

dataWizard's main target are users who have the need to compile, analyze and clean up large amounts of data from many workbooks or CSV files that are similarly formatted.

System requirements

Microsoft Windows with Microsoft Excel 2010 or newer versions*.

^{*}dataWizard will not work on Excel for the web.



Introducing dataWizard

Section 2

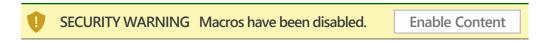


Introducing dataWizard opening dataWizard

Opening dataWizard

Download the latest version of data Wizard from https://soleox.com/datawizard and open it as you would open any other Microsoft Excel workbook.

Upon opening, you might get the following SECURITY WARNING. Make sure you click on Enable Content.



Because dataWizard relies heavily on VBA source code, macros must be enabled. If dataWizard seems dead, close it, open a blank Excel document and verify your Macro Settings are not set to disable all macros without notification. For more information on how to enable macros in Excel, refer to the following article:

https://support.office.com/en-us/article/enable-or-disable-macros-in-office-files-12b036fd-d140-4e74-b45e-16fed1a7e5c6

Private Excel instance

If any other Excel workbooks are open prior to opening dataWizard, dataWizard will create a new instance of Excel and it will reopen itself in the new instance. You do not need to do anything. Just wait a few seconds and the Home page will come up as soon as dataWizard is ready for use.

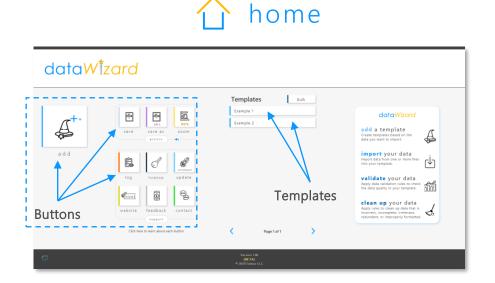
Similarly, if any workbook attempts to open itself within the instance of Excel in use by dataWizard, dataWizard will close the workbook and it will re-open it in a different instance of Excel.



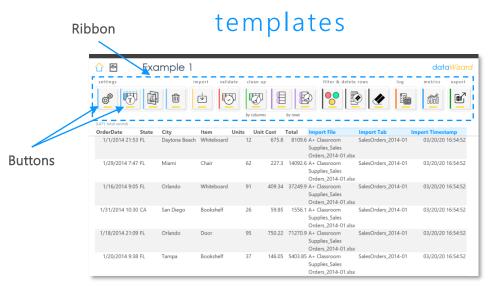
Introducing dataWizard key components

Key components

dataWizard has three (3) main components:

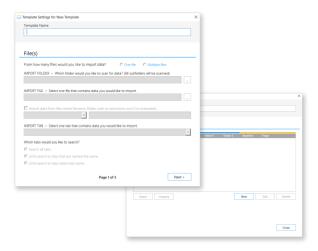


From the Home tab, you can add new templates, access existing ones, save your work, change zoom settings and much more.



You will create and use templates to import, validate and clean up data and access metrics on the quality of the data

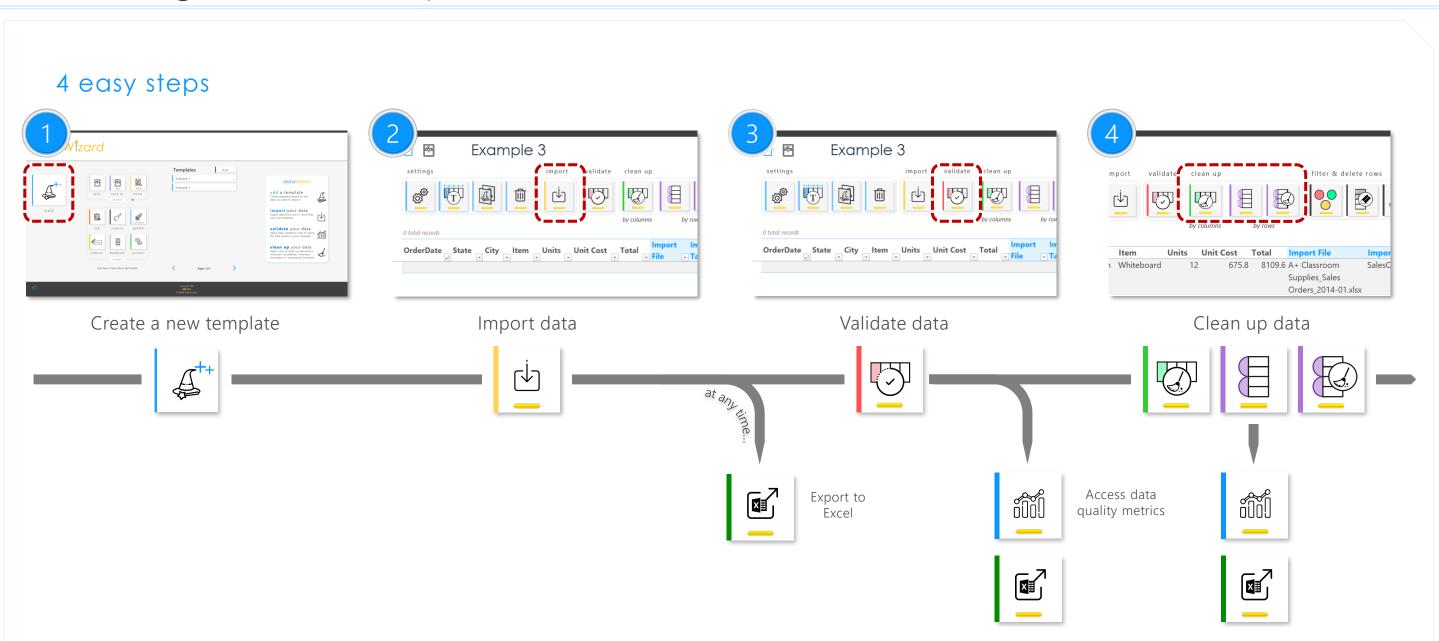
userforms



Userforms will enable you to manipulate your templates and the data in them



Introducing dataWizard quick start





Introducing dataWizard buttons

home

actions



add - Create a new template.

More information on page 15



save – Save the file that you are working on.





save as – Save the file that you are working on with a different name.



zoom – Update default zoom for all templates, Activity Log and Data Quality Metrics.

support



log – Access Activity Log, which provides an audit trail of relevant actions performed on each template and their data. It provides assurance on the integrity of the data for audit purposes.



website – Access the official dataWizard website where you can find relevant documentation, new releases, etc.



license – If you are planning on working offline or under a tight VPN connection for a while, you will need a license key to use dataWizard.



feedback – Take this short survey that will help us make dataWizard better for you and the rest of our user community.



update – When a new version of dataWizard is released, open up the new version and use this button to bring in all your templates from your previous version of dataWizard to the latest.



contact – Contact us with any questions, comments, suggestions or requests.

UPDATE



11

Introducing dataWizard buttons

template



Access/edit template settings.

T

Access template column settings (column width and data format).



Duplicate this template.



Delete the template.

Warning: Once a template is deleted it cannot be recovered.



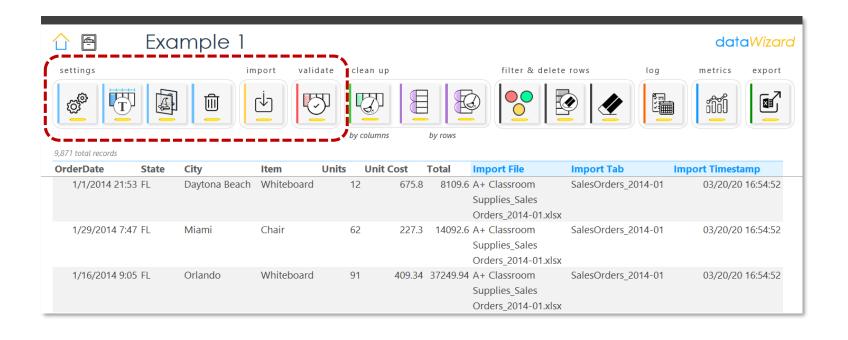
Import data from files into your template.

More information on page 20



Access data validation settings to create and apply rules to check the quality of data.

More information page 21





Introducing dataWizard buttons

template (cont.)



Access data cleanup settings to create and apply rules to clean up the data in your template.

More information on page 23



Create groupings of 2 or more rows.

More information on page 25



Access grouping settings to create and apply rules to clean your groupings.

More information on page 26





Filter template rows.

- Show rows with data validation errors.
- Show rows with no data validation errors.
- O Show all rows (i.e. unfilter template table)



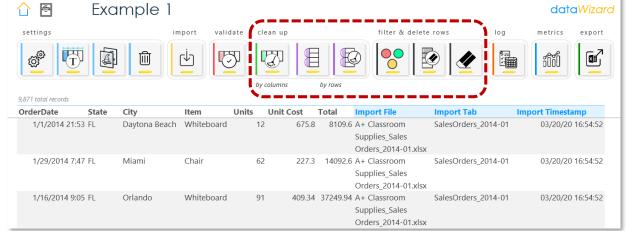
Delete all visible rows.

Warning: Once the rows are deleted they cannot be recovered.



Delete all data from the template.

Warning: Once the data is deleted it cannot be recovered.





Introducing dataWizard buttons

template (cont.)

log

Access Activity Log, which provides an audit trail of actions relevant to your template and its data.

More information on page 28

metrics



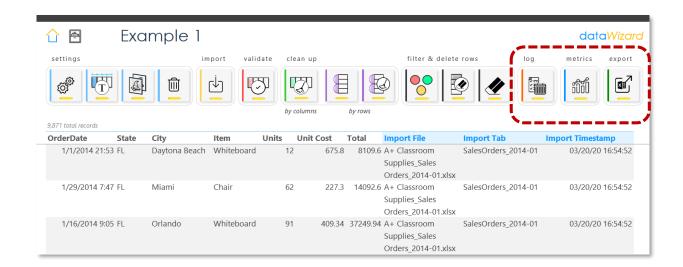
Access Data Quality Metrics to check the quality of your data.

More information on page 27

export



Export all template data to a new Excel file.







Using dataWizard

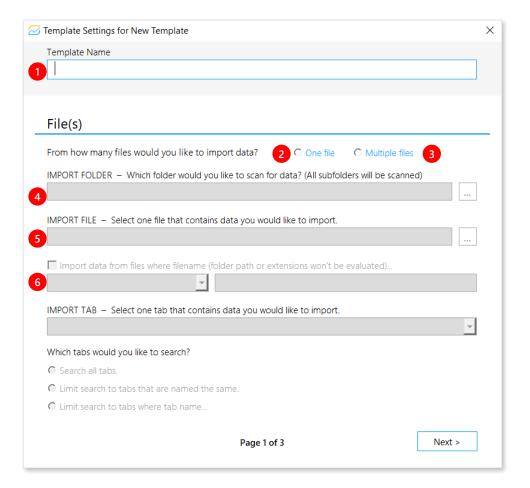
Section 3



Step 1 Creating a new template file(s)



add
create a new
template



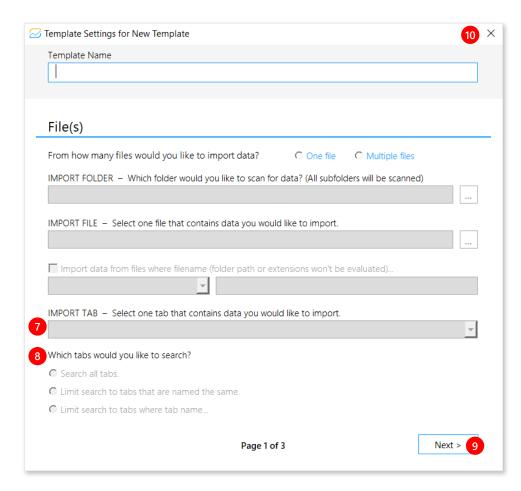
Control	Instructions	
1. Template Name	Enter a name for the template.	
2. Import data from one file	Select this option if you want to import data from just one file.	
3. Import data from multiple files	Select this option if you want to import data from multiple files.	
4. IMPORT FOLDER – Which folder would you like to scan for data? (All subfolders will be scanned)	Select the folder to be scanned for data. All folders inside the selected folder will be scanned. Tip: This control will be disabled if you choose to import data from one file.	
5. IMPORT FILE – Select one file that contains data you would like to import.	 Select the import file (file to be used as reference) for setting up the rest of the template. If the data is imported from a single file, select the file to be imported. If the data is imported from multiple files, select one of the files that contains data to be imported 	
6. Import data from files where filename (folder path extensions won't be evaluated)	Check this box to narrow down the search to files named following the specified criteria. Tip: This control will be disabled if you choose to import data from one file.	



Step 1 Creating a new template file(s)



add
create a new
template



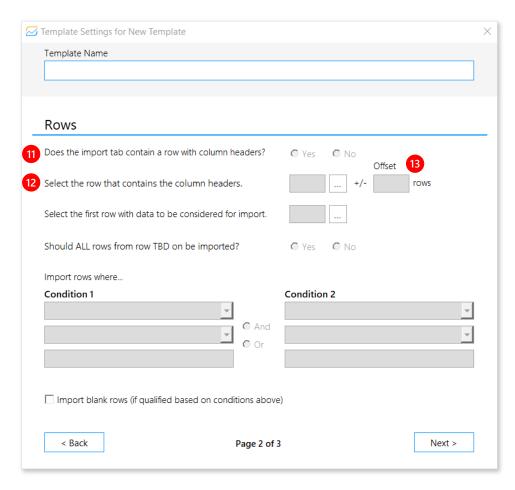
Control	Instructions		
 IMPORT TAB – Select one tab that contains data you would like to import. 	Select the import tab (tab to be used as reference) for setting up the rest of the template. Tip: This control will be disabled if the import file is a CSV file.		
8. Which tabs would you like to search?	 Search all tabs. – Select this option if the data to be imported might come from tabs that do not follow any naming convention. Tip: If you are unsure, we recommend using this option since it is the most inclusive. Limit search to tabs that are named the same. – Select this option if the data to be imported comes from tabs that are named the same as the import tab. Limit search to tabs where tab name – Select this option if the data to be imported comes from tabs that follow a naming convention. Tip: This control will be disabled if the import file is a CSV file. 		
9. Next	Click to continue to the next section.		
10. Exit	Click this button to close the Import Settings for New Template userform without saving your changes.		



Step 1 Creating a new template rows



add
create a new
template



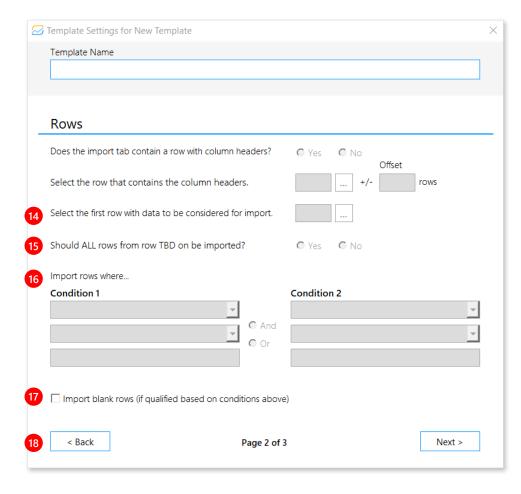
Control	Instructions
11. Does the import tab contain a row with column headers?	 Yes – Select this option if the tab to be imported contains a header row. No – Select this option if the tab to be imported does not contain a header row.
12. Select the row that contains the column headers.	Choose the column header row by double clicking it in the tab preview userform. Tip: This control will be disabled if the import tab does not contain a column header row.
13. Offset	If the column headers are always in the same row, set this value to 0. This number will determine which rows are scanned in search of the column header row. Example: If column header row is set to 5 and offset is set to 2, rows 3 (5 - 2) through 7 (5 + 2) will be evaluated in search of the column header row. Tip: This control will be disabled if the import tab does not contain a column header row.



Step 1 Creating a new template rows



add
create a new
template



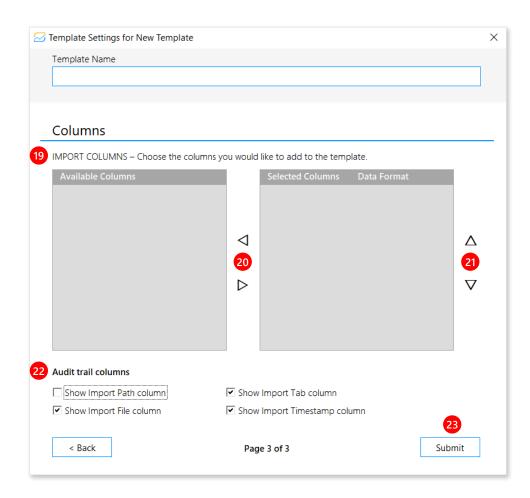
Control	Instructions
14. Select the first row with data to be considered for import.	Choose the first data row by double clicking it in the tab preview userform. By default this control will be set to the data row right after the column header row.
15. Should ALL rows from row TBD on be imported?	 Yes – Select this option if all rows (starting with the first row that contains data) should be imported. No – Select this option if you wish to narrow down the import to rows that meet specific conditions.
16. Import rows where (conditions 1 and 2)	Select the condition(s) that must be met by each row for it to be imported. Tip: These controls will be disabled if you choose to import all rows.
17. Import blank rows (if qualified based on conditions above)	Check the box if blank rows of data should be imported.
18. Back	Click this button to go back to the previous section.

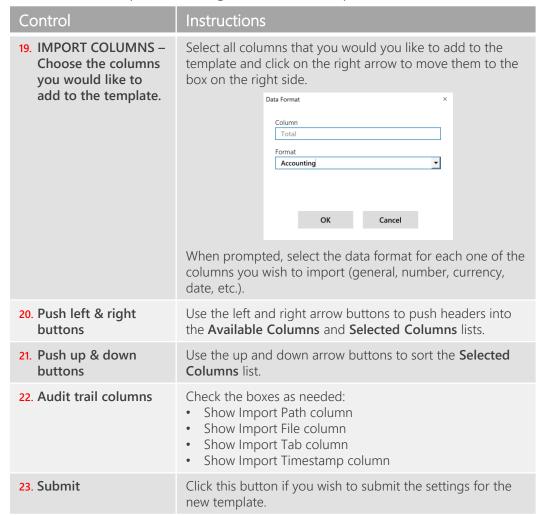


Step 1 Creating a new template columns



add
create a new
template







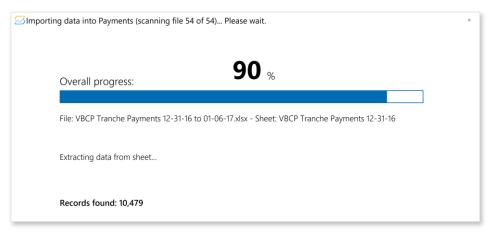
dataWİzard

Step 2 Importing data

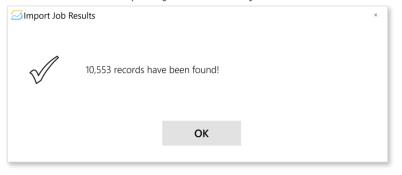


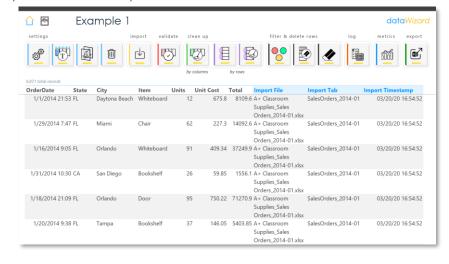
import
import data
 into the
 template

Just wait for the import to complete. You can see the progress in the Importing data progress bar...



As soon as the import job finishes, you will see the results from the import and the template will be filled out with data:





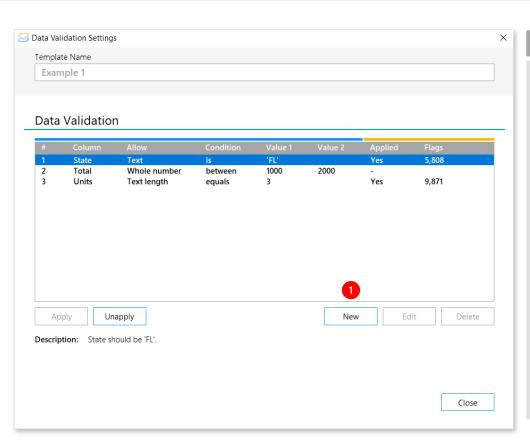


dataWizard

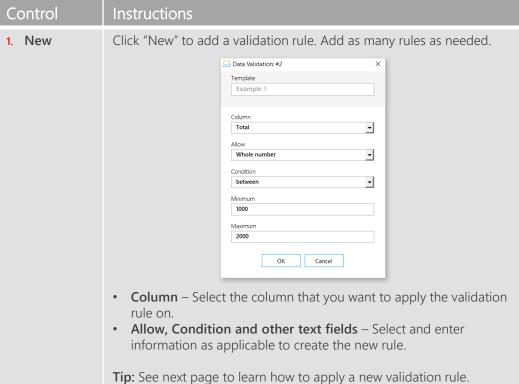
Step 3 Validating data



validate
data
validation
settings by
column



userform Data Validation Settings



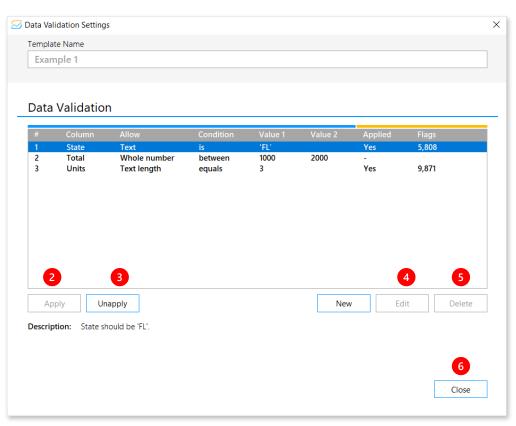


Step 3 Validating data



validate

data validation settings by column



userform Data Validation Settings

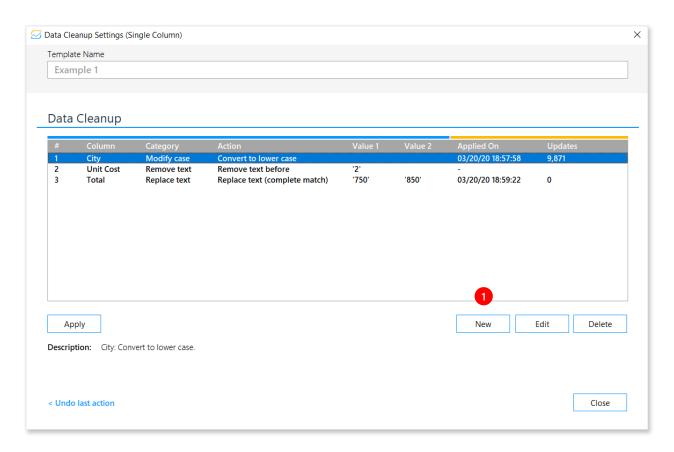
	3
Control	Instructions
2. Apply	Select one or more existing validation rules and click Apply to apply them.
3. Unapply	Select one or more existing validation rules and click Unapply to unapply the rule.
4. Edit	Select an already created validation rule and click Edit to modify its properties. Tip: In order to edit a validation rule, the validation rule must be unapplied.
5. Delete	Select an already created validation rule and click Delete to delete the rule. Tip: In order to delete a validation rule, the validation rule must be unapplied.
6. Close	Click this button to close the Data Validation Settings userform.



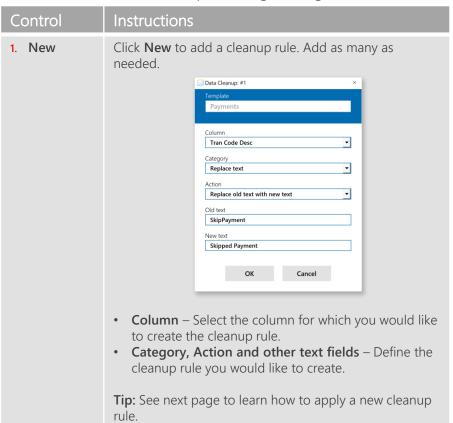
Step 4 Cleaning up data by columns



data cleanup settings by column



userform Data Cleanup Settings (Single Column)

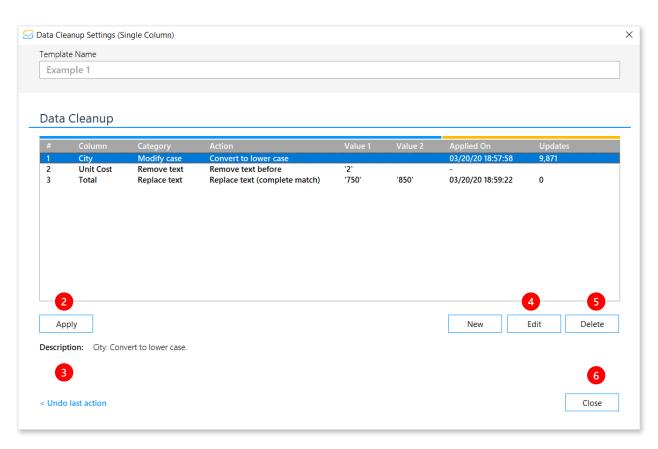




Step 4 Cleaning up data by columns



clean up data cleanup settings by column



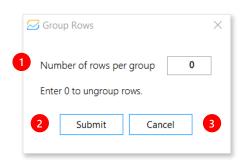
userform Data Cleanup Settings (Single Column)

Control	Instructions
2. Apply	Select one or more existing cleanup rules and click Apply to apply them.
3. Undo last action	Click this button to undo the <u>last</u> cleanup rule that has been applied.
	Tip: This button appears only after cleanup rules have been applied.
4. Edit	Select an already created cleanup rule and click Edit to modify its properties.
5. Delete	Select an already created cleanup rule and click Delete to delete the rule.
6. Close	Click this button to close the Data Cleanup Settings (Single Column) userform.



Step 4 Cleaning up data by rows



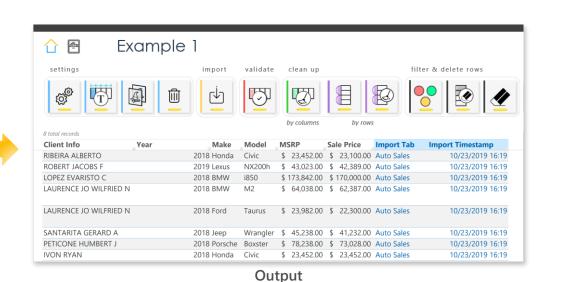


userform Group Rows

Control	Instructions
Number of rows per group	Enter the number of rows for each group.
2. Submit	Click this button to submit the grouping settings.
3. Cancel	Click this button to cancel and close the Group Rows userform.

The row grouping functionality is intended for cases where each individual record is spread out across multiple rows. This functionality lets you bulk clean up all records as needed. As an example, see the input and output of the Timesheet template below:

Client Info	VIN#					
	Year	Make	Model	Color	MSRP	Sale Price
RIBEIRA ALBERTO 6728 BETHMAN STREET	JEDFA123923892323					
LAS VEGAS, NV 89109 XXX-XX-4679	2018	Honda	Civic	Red	23,452.00	23,100.00
ROBERT JACOBS F 12 STATE BROOK CIRCLE	JEHSA748293893293					
SAN CRISTOBAL, CA 91524 XXX-XX-5708	2019	Lexus	NX200h	Black	43,023.00	42,389.00
LOPEZ EVARISTO C 1982 LAGO CANYON			BFHSA12	3923892323		
LAS VEGAS, NV 89138 XXX-XX-2198	2018	BMW	i850	Red	173,842.00	170,000.00
LAURENCE JO WILFRIED N 1284 MORNING CANYON	BFHRA987482892323					
HENDERSON, NV 89213 XXX-XX-2127	2018	BMW	M2	Dark Grey	64,038.00	62,387.00
LAURENCE JO WILFRIED N 1284 MORNING CANYON	JEHSA123923892323					
HENDERSON, NV 89213 XXX-XX-2127	2018	Ford	Taurus	Blue	23,982.00	22,300.00
SANTARITA GERARD A 1043 PIRATES LANE	JEJTH123238923239					
FLOWERVILLE, CA 95278 XXX-XX-5981	2018	Jeep	Wrangler	Red	45,238.00	41,232.00
PETICONE HUMBERT J 363 WEST MARKET STREET	UIRSA1239C2389323					
SAN JUAN, CA 95125 XXX-XX-4103	2018	Porsche	Boxster	Black	78238	73,028.00
IVON RYAN 2056 NIKON PERDIDO COURT	JEOPA123923892323					
ATWINE, CA 92901 XXX-XX-5968	2018	Honda	Civic	Red	23,452.00	23,452.00

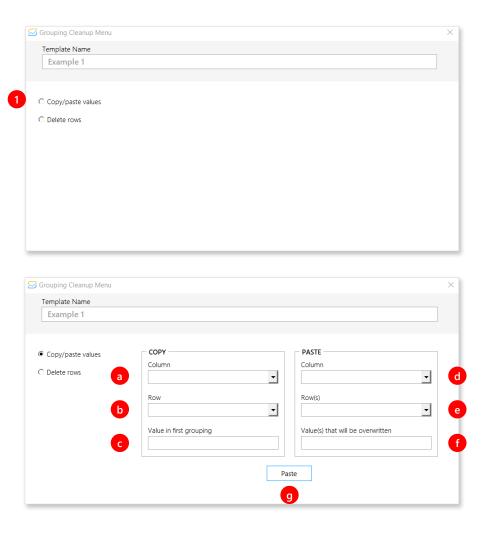


Input



Step 4 Cleaning up data by rows





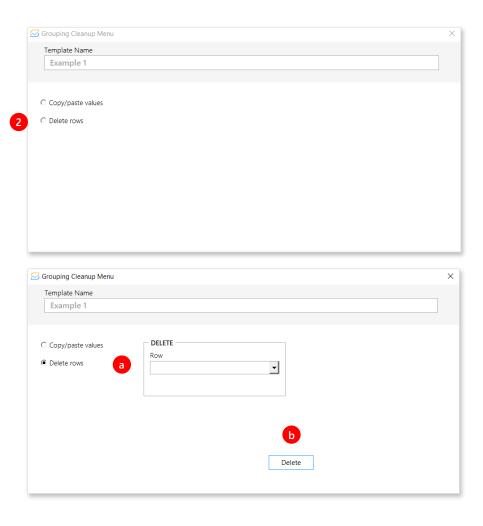
userform Grouping Cleanup Menu

Control	Instructions		
1. Copy/paste values	Select this option to copy and paste values within groupings.		
	Tip: The Copy/paste values function will take the value in the row of each grouping and it will paste it in the row(s) in the same column.		
COPY frame			
a. Column	Choose from the dropdown the column to copy.		
b. Row	Choose from the dropdown the row to copy.		
c. Value in first grouping	This box is only for reference. It displays the value to be copied from the first grouping.		
PASTE frame			
d. Column	Select from the dropdown the column to overwrite.		
e. Row(s)	Choose from the dropdown the column to overwrite.		
f. Value(s) that will be overwritten	This box is only for reference. It displays the value that will be overwritten from the first grouping.		
g. Paste	Click this button execute the copy/paste action.		



Step 4 Cleaning up data by rows





userform Grouping Cleanup Menu

Control	Instructions	
2. Delete rows	Select this option to delete rows within groupings.	
	Tip: The Delete rows function will delete the selected row from every grouping in the template.	
DELETE frame		
a. Row	Select from the dropdown the row to delete within groupings.	
b. Delete	Click this button to execute the delete action.	



Data quality metrics



The Data Quality Metrics page contains two sections:

- Overall data quality
- Data quality by column
- 1. This page has two views: chart view and table view. You can switch views by clicking on the top right buttons.

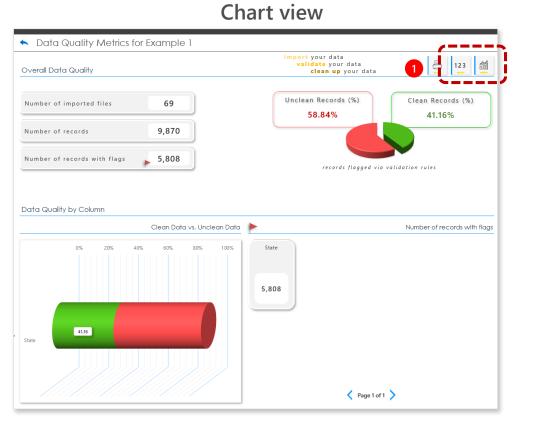


Table view

♠ Data Quality	Metrics for Example 1	
Overall Data Quality		<mark>import</mark> your data validate your data clean up your data
Number of import files Number of records Number of records with flags Clean Records (%) Unclean Records (%)	69 9,870 5,808 41.16% 58.84%	
Data Quality by Colu	nn	
Column	Number of Number of rules records Points (9 Po	
State	1 5,808 41.16% 58.84%	

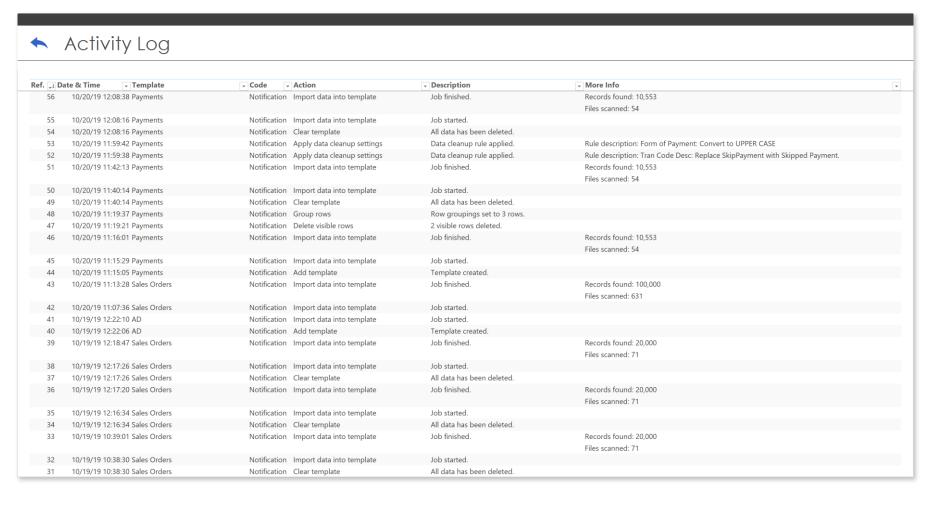


Activity log



log activity log The Activity Log is an audit trail report with a chronological sequence of actions that have been performed. It contains evidence as to the origin and integrity of the data.

You may access the Activity Log from the Home page or from a specific template.





Data cleanup rules explained

Section 4

Contact us at support@soleox.com© 2020 Soleox



Apply mathematical operation

Input

✓ A number

Action	Description	Example		
	Description	Current value	Input number	New value
Add	Adds input number to the existing value.	100	4	104
Subtract	Subtracts input number from current value.	100	4	96
Multiply	Multiplies current value by input number.	100	4	400
Divide	Divides current value by input number.	100	4	25



Convert data type

Input

- ✓ Current data type
 - Y: year*
 - M: month*
 - D: day
 - h: hour
 - m: minute
 - s: second
 - d: decisecond (0.1s)
 - c: centisecond (0.01s)
 - I: millisecond (0.001s)
 - n: nanosecond (10⁻⁹s)

Action	Description	Example		
	Description	Current value	Current data type	New value
Convert number to date	Converts current value from current data type to a date.	20200202008421	YYYYMMDDssssss	02/02/2020 02:20:21
Convert number to duration (days)	Converts current value from current data type to duration in days.	12462353	SSSSSSS	144.240196759259
Convert number to duration (hours)	Converts current value from current data type to duration in hours.	04235982	DDhhmmss	120.006111111111
Convert number to duration (minutes)	Converts current value from current data type to duration in minutes.	234804	dddddd	391.34
Convert number to duration (seconds)	Converts current value from current data type to duration in seconds.	125 <mark>99</mark> 59	hhh <mark>mm</mark> ss	455,999
Convert number to duration (deciseconds)	Converts current value from current data type to duration in deciseconds.	42 <mark>04</mark> 01	hhmmss	1,514,410
Convert number to duration (centiseconds)	Converts current value from current data type to duration in centiseconds.	10938	SSSSS	1,093,800
Convert number to duration (milliseconds)	Converts current value from current data type to duration in milliseconds.	004 <mark>02</mark>	SSSCC	4,020
Convert number to duration (nanoseconds)	Converts current value from current data type to duration in nanoseconds.	25001	ssmmm	85,000,000,000

^{*} Only applicable when Action = Convert number to date



Insert text

Input

✓ New text

Action	Description		Example		
	Description	Current value	New text	New value	
Insert text at the beginning	Inserts new text at the beginning of the current value.	0203001	INV-	INV-0203001	
Insert text at the end	Inserts new text at the end of the current value.	Newport Beach	, CA	Newport Beach, CA	

^{*} Only applicable when Action = Convert number to date



Modify case

Input

No input

Action	Description	Example		
	Description	Current value	New value the brown fox The Brown Fox	
Convert to lower case	Converts current value to lower case.	The brown fox	the brown fox	
Convert to Title Case	Converts current value to title case.	The brown fox	The Brown Fox	
Convert to UPPER CASE	Converts current value to upper case.	The brown fox	THE BROWN FOX	



Remove text

Input

✓ Text to remove

Action	Dossription	Example		
	Description	Current value	Input text	New value
Remove text before	Removes from current value any text that precedes input text.	ABCDEFGHI	DEF	DEFGHI
Remove text before and including	Removes from current value the first instance of input text and any text that precedes it.	ABCDEFGHI	DEF	GHI
Remove text after	Removes from current value any text that goes after last instance of input text.	ABCDEFGHI	DEF	ABCDEF
Remove text after and including	Removes from current value the last instance of input text and any text after it.	ABCDEFGHI	DEF	ABC



Remove number of character

Input

✓ Number of characters to remove

Action	Description	Example		
	Description	Current value	Input number	New value DEFGHI
Remove X leading characters	Removes X number of characters starting at the beginning of the current value.	ABCDEFGHI	3	DEFGHI
Remove X trailing characters	Removes X number of characters starting at the end of the current value.	ABCDEFGHI	3	ABCDEF



37

Remove unwanted characters

Input

No input

Actions

Action	Description	Example		
Action	Description	Current value	New value	
Remove leading whitespaces	Removes leading whitespaces from current value.	San Francisco	San Francisco	
Remove leading and trailing whitespaces	Remove leading and trailing whitespaces from current value.	San Francisco	San Francisco	
Remove trailing whitespaces	Removes trailing whitespaces from current value.	San Francisco	San Francisco	
Remove tabs	Removes tabulator key characters from current value.	San Francisco	SanFrancisco	
Remove line breaks	Removes all line breaks from current value.	Irvine, CA. 92618	Irvine, CA 92618	
Remove duplicate whitespaces	Removes duplicate whitespaces from current value.	SanDiego	San_Diego	
Remove all whitespaces	Removes all whitespaces from current value.	San_Diego	SanDiego	
Remove all letters	Removes all letters from current value.	123ab456cd789	123456789	
Remove all numbers	Removes all numbers from current value.	123ab456cd789	abcd	



38

Replace text

Input

✓ Old text

✓ New text

Actions

Action	Description	Example			
Action	Description	Current value	Old text	New text	New value
Replace text (complete Replaces old text v	Replaces old text with new text if current	ABCDEFGHI	DEF	XYZ	ABCDEFGHI
match)	value matches old text.	ABCDEFGHI	ABCDEFGHI	XYZ	XYZ
Replace text (complete and partial match) Replace	Dealess ald to to the second of	ABCDEFGHI	DEF	XYZ	ABCXYZGHI
	Replaces old text with new text.	ABCDEFGHI	ABCDEFGHI	XYZ	XYZ
	Replaces old text with new text if current	ABCDEFGHI	DEF	XYZ	ABCXYZGHI
	value does not fully match old text.	ABCDEFGHI	ABCDEFGHI	XYZ	ABCDEFGHI



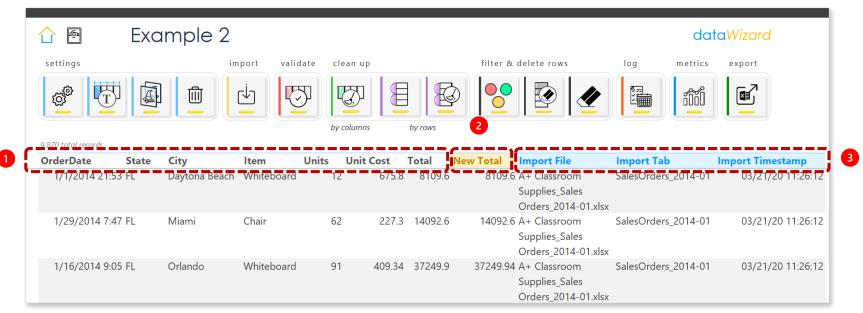
Template columns explained

Section 5

Contact us at support@soleox.com© 2020 Soleox



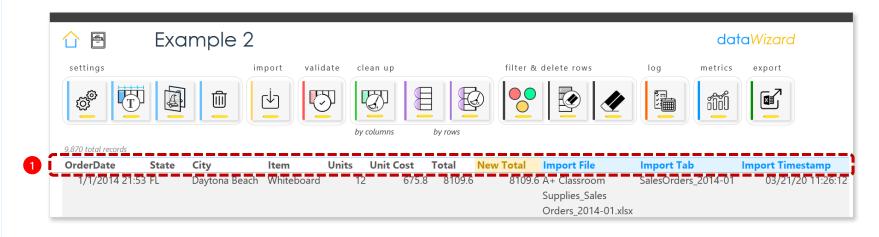
Column sources

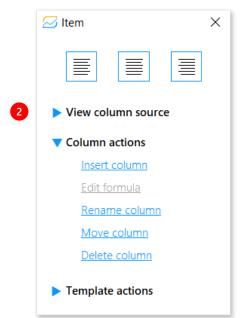


Template Columns	Color	Description	Example
1. Import columns	White	Columns imported from the template	From example above: The user imported OrderDate , State , City , Item , Units , Unit Cost and Total from import files.
2. Calculated columns	Yellow	Columns created from Import columns (for more info on how to add calculated columns, see Insert Column slide)	From example above: To verify the accuracy of the Total column, the user inserted a new column that displays the product of Units times Unit Cost.
3. Audit Trail columns	Blue	Columns used for audit trail purposes	Import Path displays the path of the file(s) imported Import File displays the name of the file(s) imported Import Tab displays the name of the tab(s) imported Import Timestamp displays the date and time when the record is imported



Column actions View column source



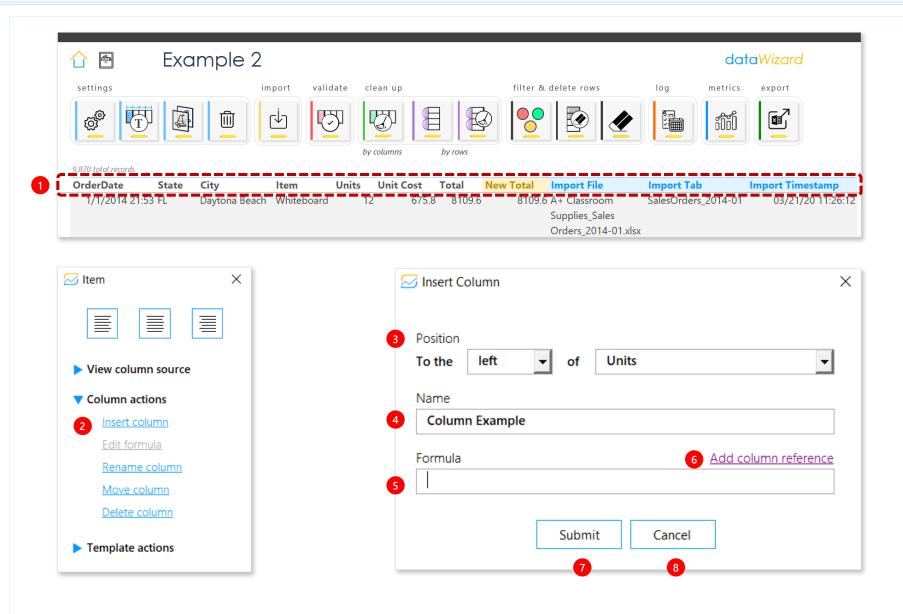


userform [name of column]

Control	Instructions
1. Column headers	Right click on any column header to activate the Column userform for that specific header.
2. View column source	Click on View column source to see the column source: Import, Calculated or Audit Trail. For more info see Column sources slide.



Column actions Insert column



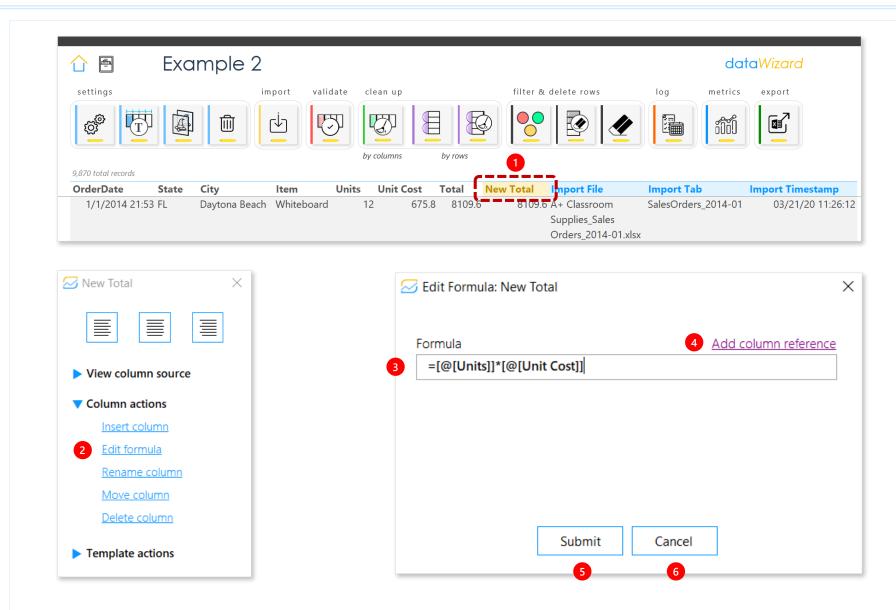
userform Insert Column

Control	Instructions
1. Column headers	Right click on any column header to activate the Column userform for that specific header.
2. Insert column	Click on Insert column to display the Insert Column userform.
3. Position	Choose from the To the and of dropdowns to select where to position the new column to be inserted.
4. Name	Enter a name for the new column.
5. Formula	Input formula to calculate the new column to be inserted.
6. Add column reference	When adding the formula in previous step, click on Add column reference to choose the column(s) that you want to add to the formula.
7. Submit	Click this button if you wish to submit the settings for the new column.
8. Cancel	Click this button to cancel and close the Insert Column userform.





Column actions Edit formula



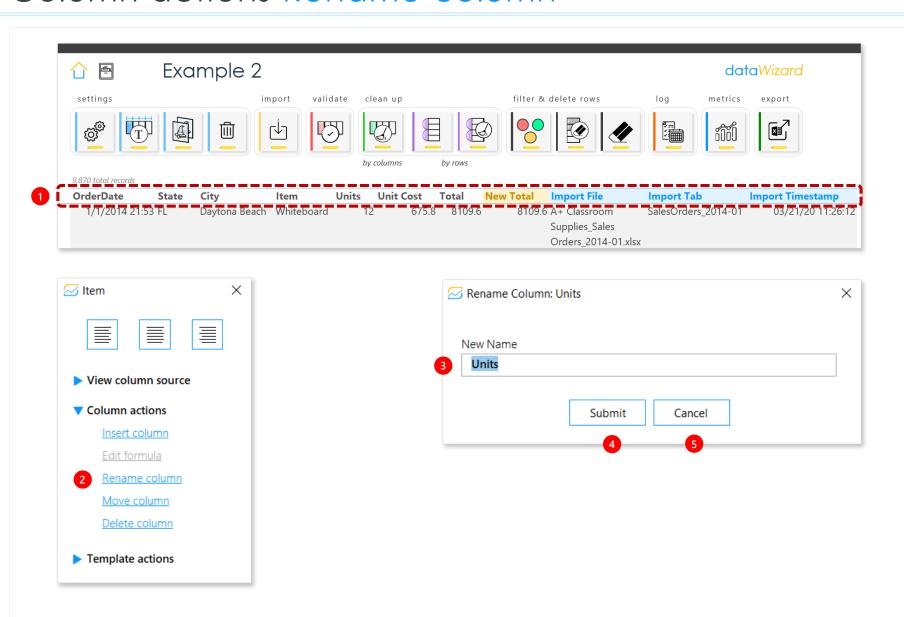
userform Edit Formula: [name of column]

Control	Instructions
1. Column headers	Right click on the header of any Calculated column to activate the Column userform for that specific header.
2. Edit formula	Click on Edit formula to display the Edit Formula : [name of column] userform.
3. Formula	Edit the formula of the Calculated column. Tip: You can only Edit formula in Calculated columns.
4. Add column reference	When editing the formula, click on Add column reference to choose the column(s) that you want to add to the formula.
5. Submit	Click this button if you wish to submit the settings for the Calculated column.
6. Cancel	Click this button to cancel and close the Edit Formula : <i>[name of column]</i> userform.





Column actions Rename column



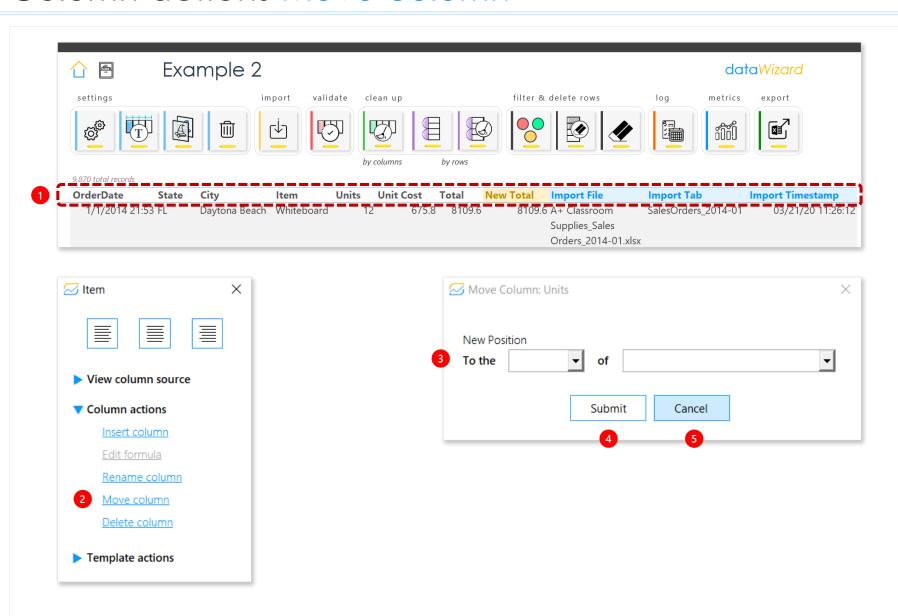
userform Rename Columns: [name of column]

Control	Instructions
1. Column headers	Right click on any column header to activate the Column userform for that specific header.
2. Rename column	Click on Rename column to display the Rename Column: [current name of column] userform. In this example, the current name of the column is "Units".
3. New Name	Enter a new name for the column. Tip: You can only Rename column in Imported and Calculated columns.
4. Submit	Click this button if you wish to submit the settings for the renamed column.
5. Cancel	Click this button to cancel and close the Rename Column: [current name of column] userform.





Column actions Move column



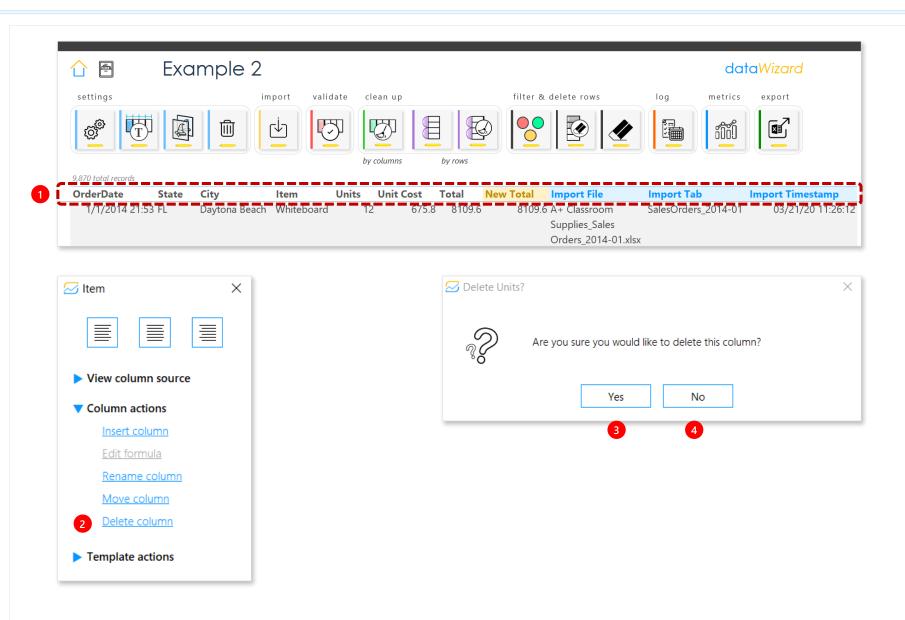
userform Move Column: [name of column]

Control	Instructions
1. Column headers	Right click on any column header to activate the Column userform for that specific header.
2. Move column	Click on Move column to display the Move Column: [name of column] userform. In this example, the name of the column we want to move is "Units".
3. New Position	Choose from the To the and of dropdowns to select the new position of the column. Tip: You can only Move column in Imported and Calculated columns.
4. Submit	Click this button if you wish to submit the settings for the column's new position.
5. Cancel	Click this button to cancel and close the Move Column: <i>[name of column]</i> userform.





Column actions Delete column

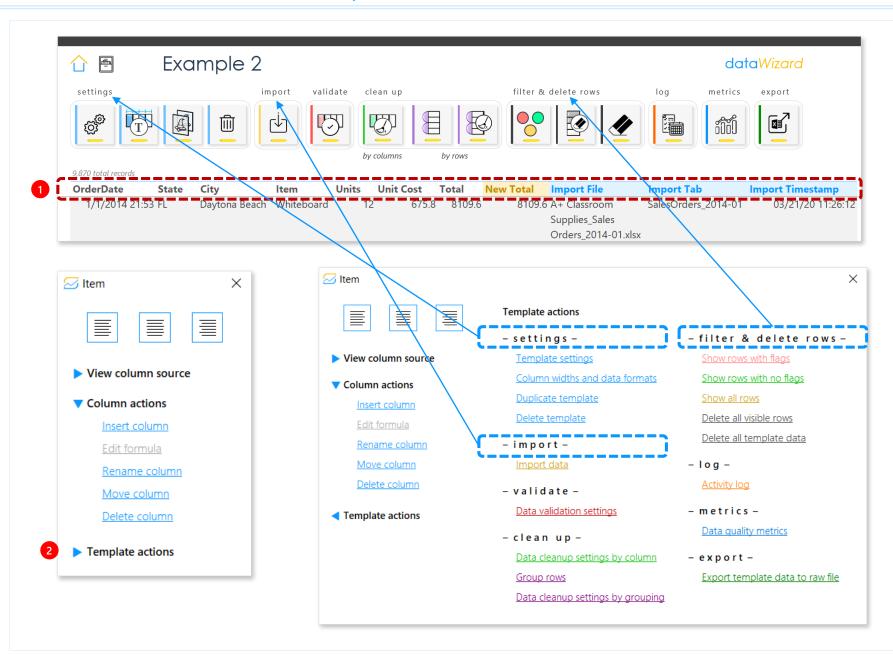


userform Delete [name of column]?

	· -
Control	Instructions
1. Column headers	Right click on any column header to activate the Column userform for that specific header.
2. Delete column	Click on Delete column to display the Delete [name of column]? userform. In this example, the name of the column we want to delete is "Units".
3. Yes	Click Yes to delete the column.
4. No	Click No to do <u>not</u> delete the column.



Column actions Template actions



userform [name of column]

Control	Instructions
1. Column headers	Right click on any column header to activate the Column userform for that specific header.
2. Template actions	Click on Template actions to display all Settings , Import , Validate , Clean up , Filter and delete rows , Log , Metrics , and Export actions.
	Tip: The Template actions match the buttons in the ribbon.

Tip: This control comes in handy when working with wide templates that may require scrolling to the right.

