Manual. Shiny R App

Augusta University | 1120 15th Street, Augusta, GA 30912

Calcium imaging data analysis using shiny r app

Taraskin, Aleksandr

2023

Contents

[1. Opening the program for calcium imaging data analysis 2](#_Toc144461557)

[2. TAB: Preliminary analysis 2](#_Toc144461558)

[Uploading excel files 3](#_Toc144461559)

[How to handle several sheets in a file 4](#_Toc144461560)

[Custom sheet names 5](#_Toc144461561)

[Correct “Time” column 6](#_Toc144461562)

[Change column names 9](#_Toc144461563)

[Display option 12](#_Toc144461564)

[Calculate basic statistics 12](#_Toc144461565)

[Save descriptive statistics table as excel file 15](#_Toc144461566)

[Creating plots and excluding bad traces 15](#_Toc144461567)

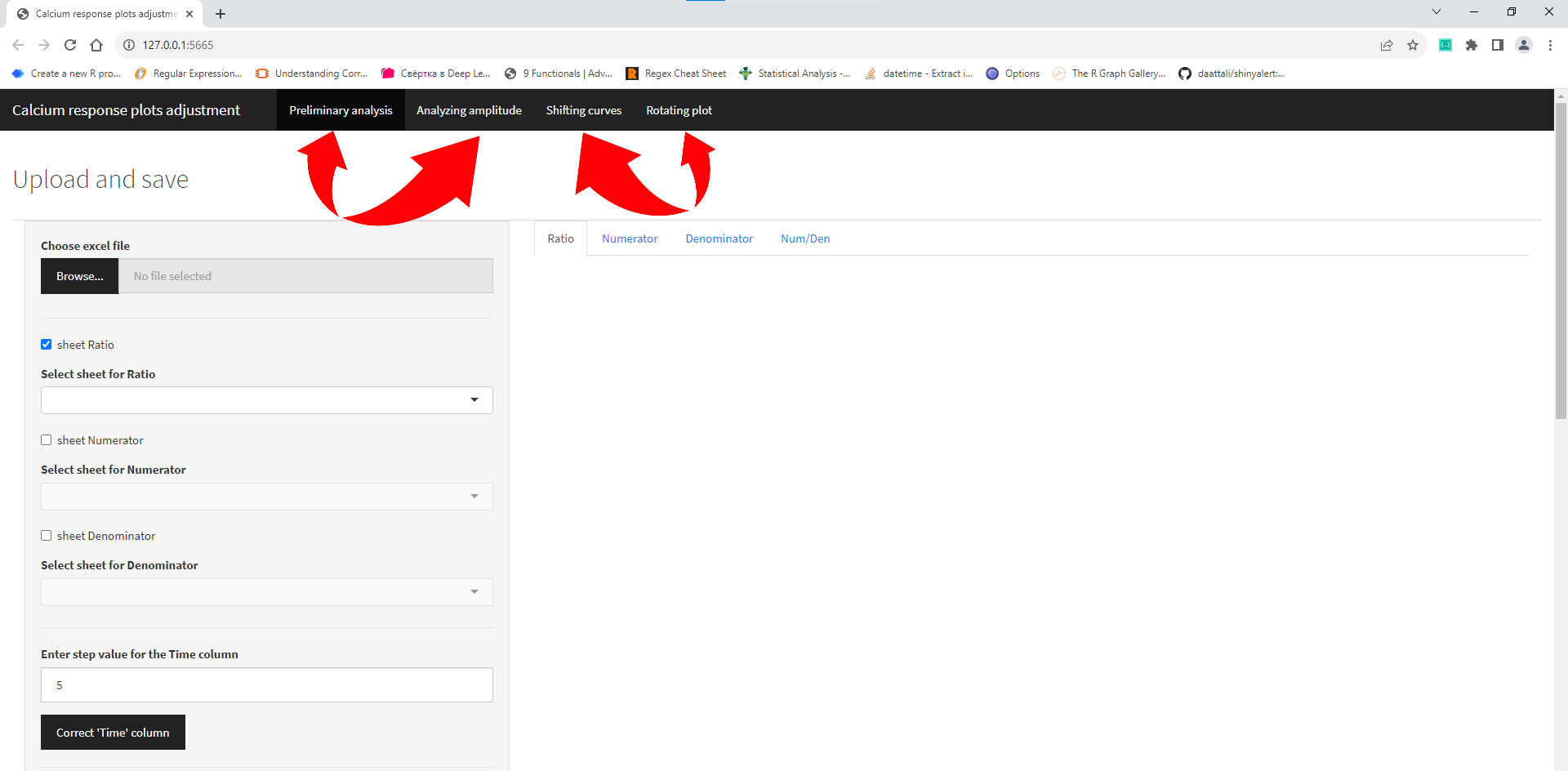
[Excluding traces 17](#_Toc144461568)

[Discarding unnecessary traces 19](#_Toc144461569)

Calcium imaging

# Opening the program for calcium imaging data analysis

The upper part of the program’s start screen look like this:



Altogether there are 4 tabs:

* Preliminary analysis (loading raw excel files, correct Time column, change column names, exclude traces, etc.)
* Analyzing amplitude (finding amplitude for curves that have 1, 2 or 3 major maximums)
* Shifting curves (shift plots automatically to the left using two algorithms)
* Rotating plot (rotating plots and their parts)

You can start using program after choosing any of these 4 (Preliminary analysis is chosen by default).

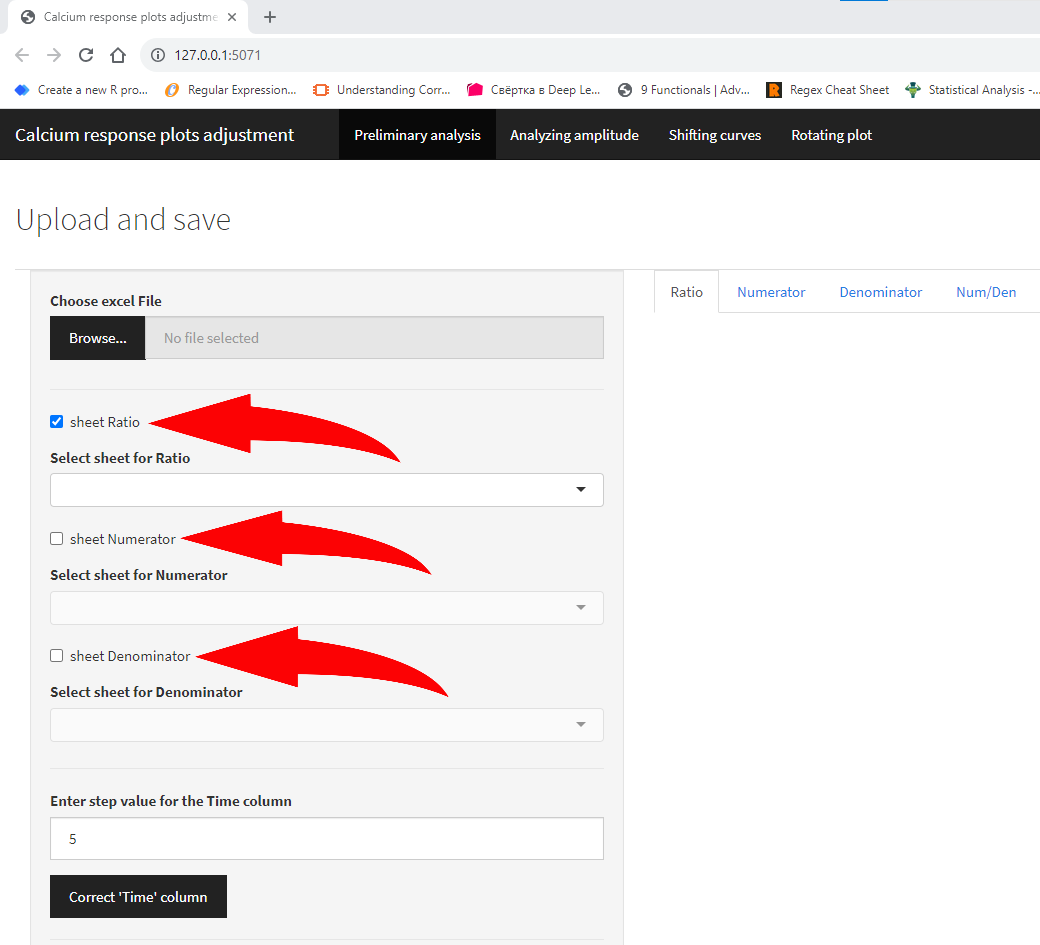
# TAB: Preliminary analysis

On this tab the main preview window (to the right from a sidebar) contains 4 internal tabs by default:

* Ratio – (fluorescence in channel 340/380)
* Numerator – (fluorescence in channel 340)
* Denominator – (fluorescence in channel 380)
* Num/Denominator – (custom ratio: values from Numerator are divided by values from Denominator)

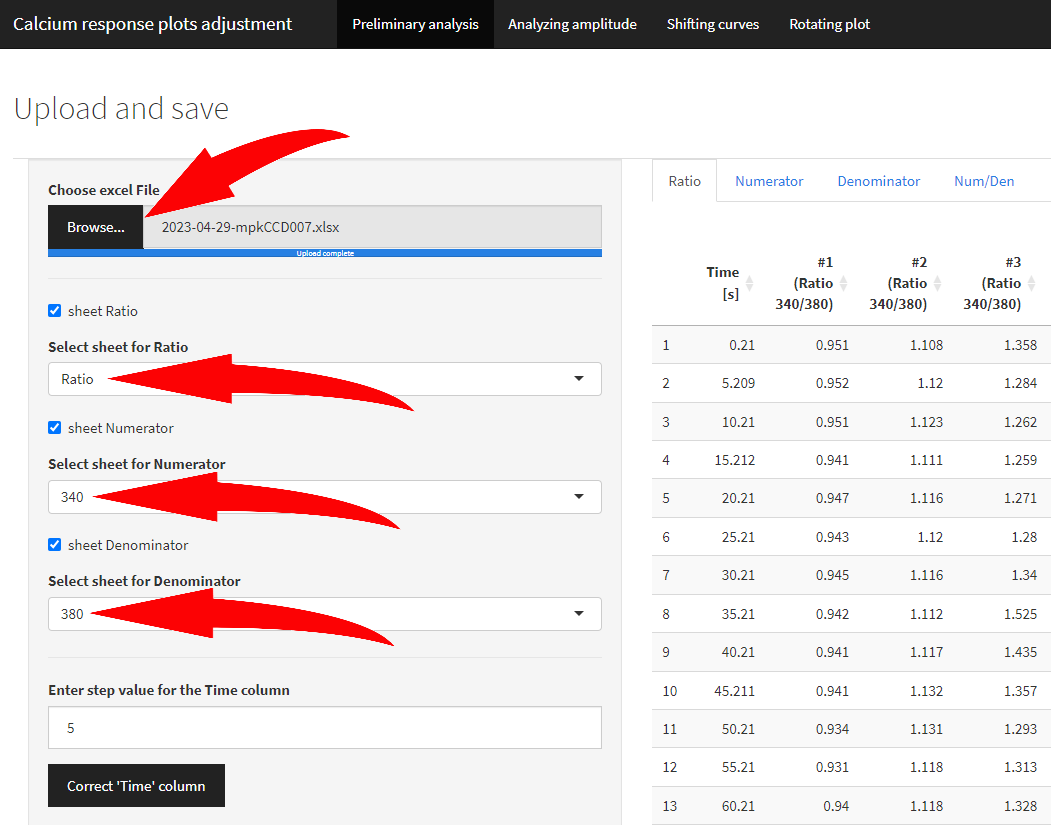
## Uploading excel files

In order to load your file and read sheets that contain information about Ratio, 340 or 380 channels you need to check necessary boxes on the left (user is allowed to do that either before the file was chosen and loaded or after):



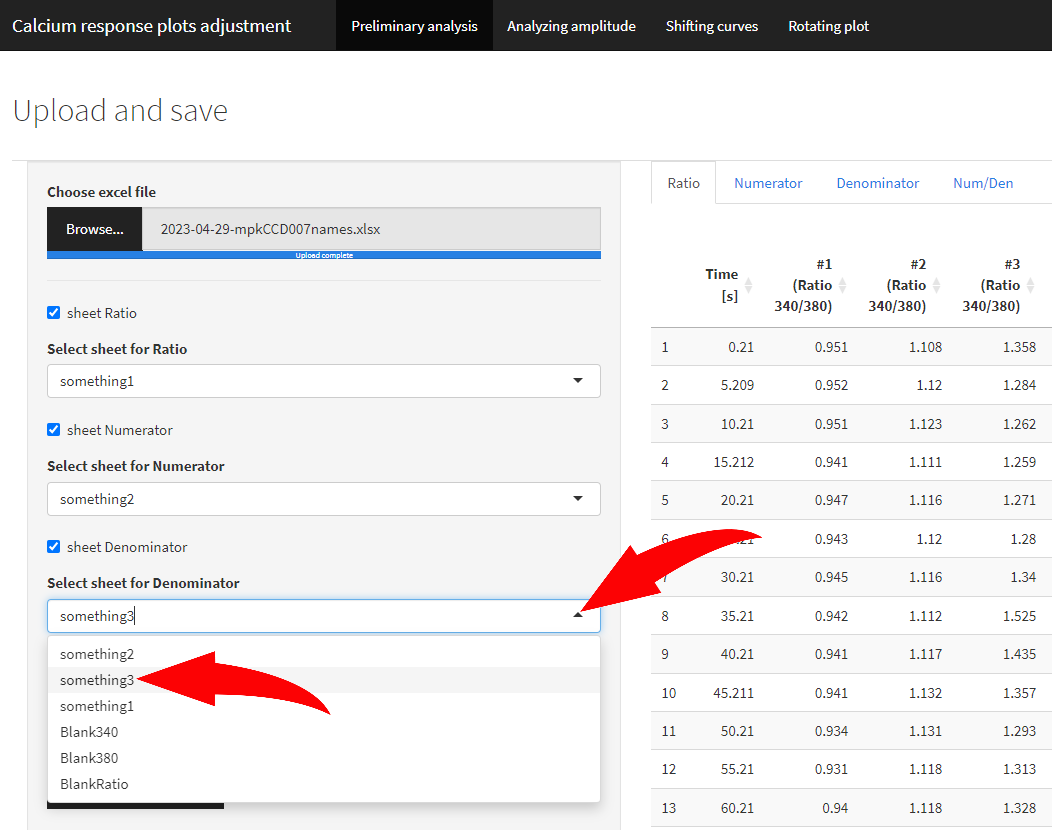
## How to handle several sheets in a file

An excel file should contain at least one sheet if you checked the one box or more. After pressing “Browse” button and choosing an excel file: if the file has sheets with names “Ratio” or “ratio”, “340” or “380” – they will be automatically recognized and reflected in “Select sheet for …” boxes:



## Custom sheet names

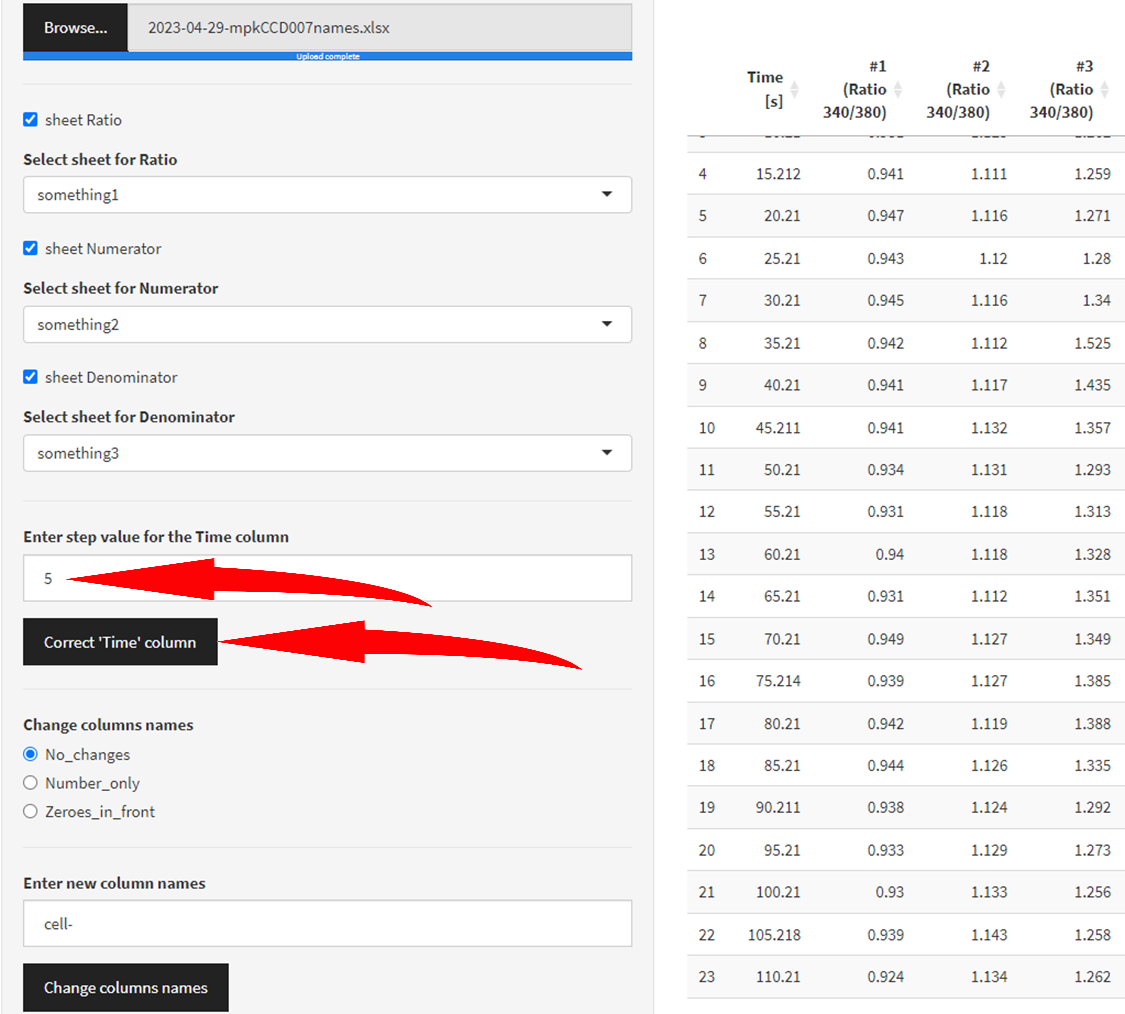
If names are different from reference, user can manually set any sheet name that corresponds to Ratio/Numerator/Denominator:



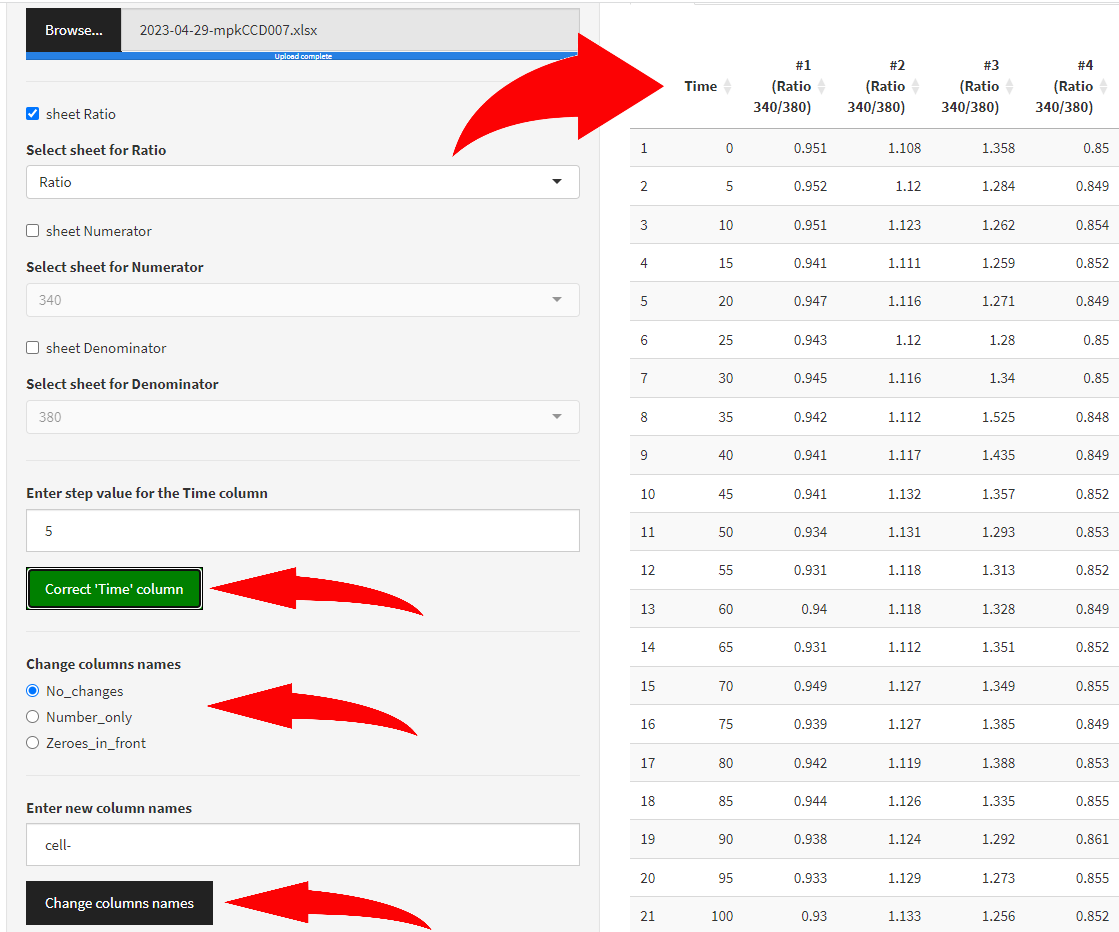
## Correct “Time” column

Next step allows you to correct the “Time” column and column names if necessary (but recommended).

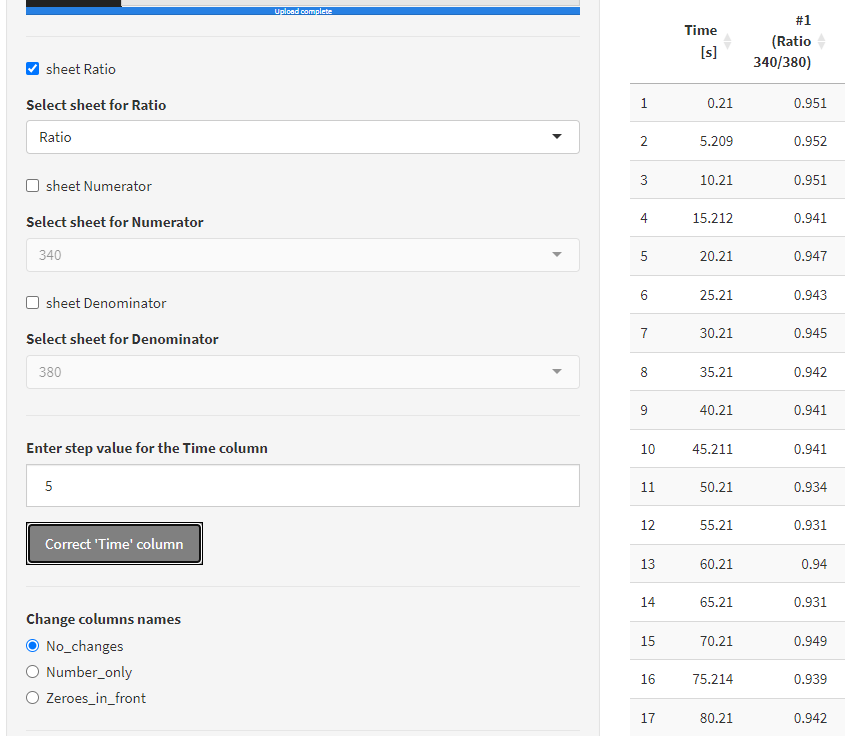
To replace “Time [s]” column to “Time” and fill it with values with custom step (by default 5 sec), enter the step value into the box below (if different from default) and press “Correct ‘Time’ column”.



After pressing the “Correct ‘Time’ column” button, it turns green indicating that action has been taken.

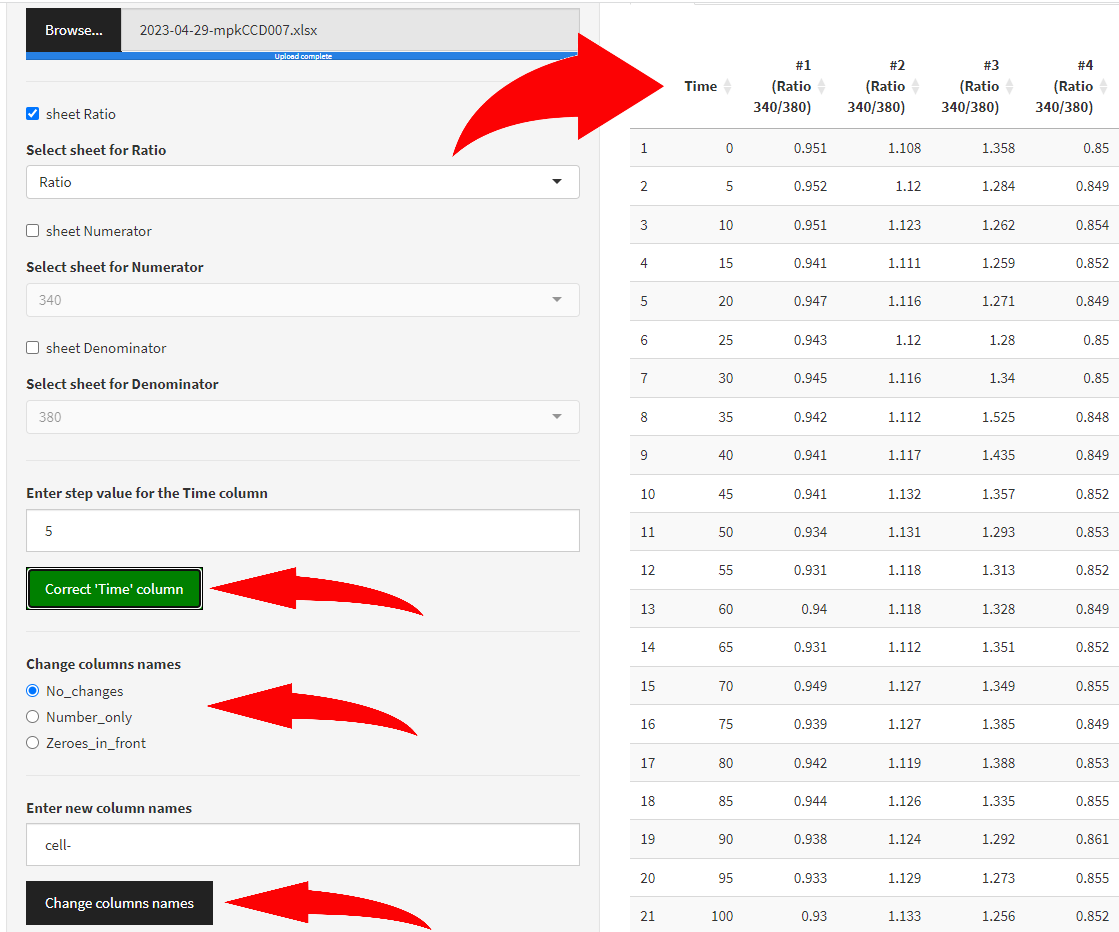


By pressing it again it turns grey and deny changes that have been made to the “Time” column.

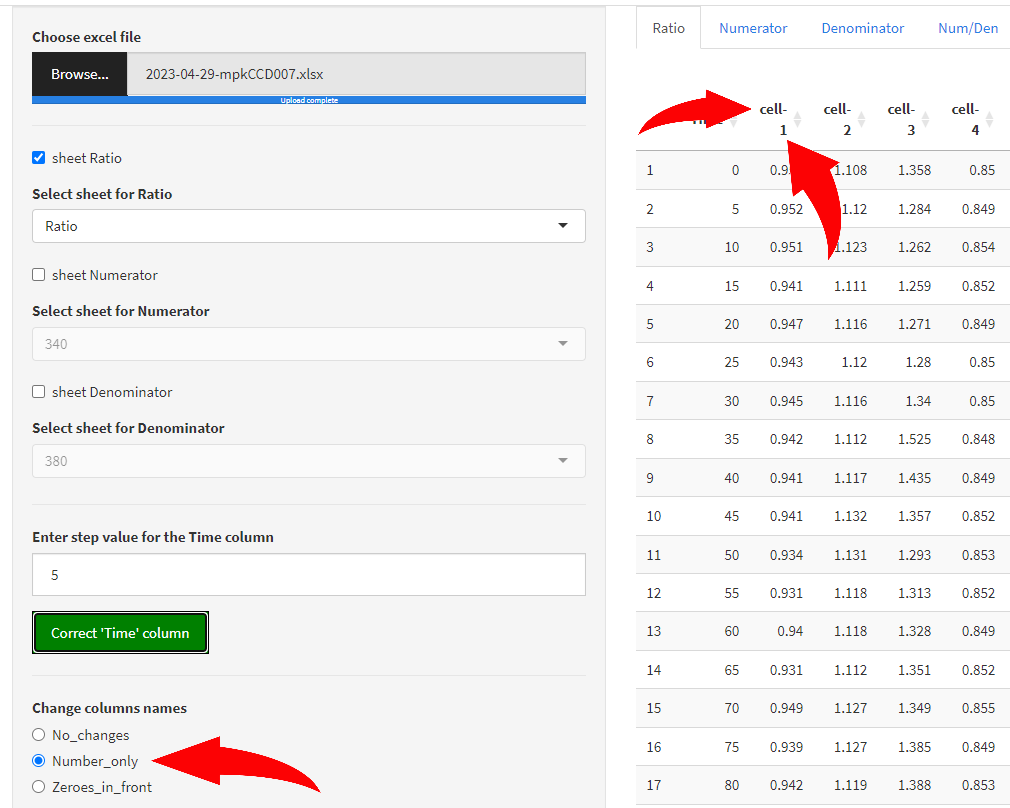


## Change *column names*

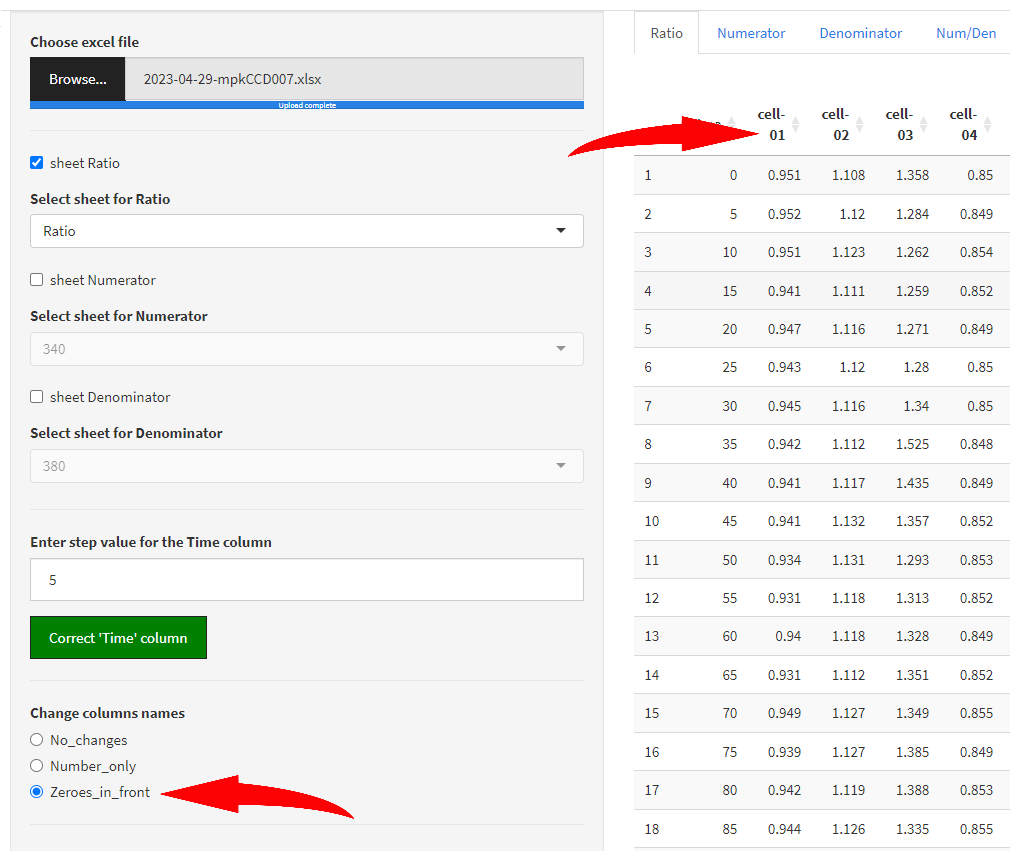
To change other column names there are two options for number conversion and option for custom prefix.



By choosing “Number\_only” column, any number in a column name will be captured and column will be renamed using the following pattern: column name prefix (by default “cell-”) + captured number.

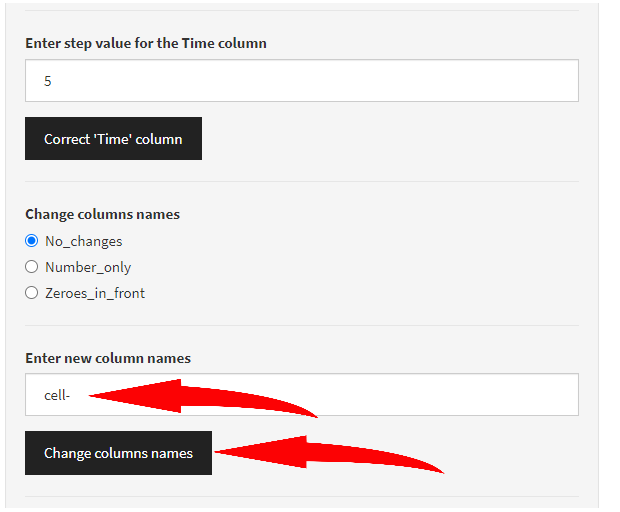


By choosing “Zeroes\_in\_front” option to any captured number a certain amount of zeroes will be added so the result number has fixed amount of all digits (this can help with sorting column names).



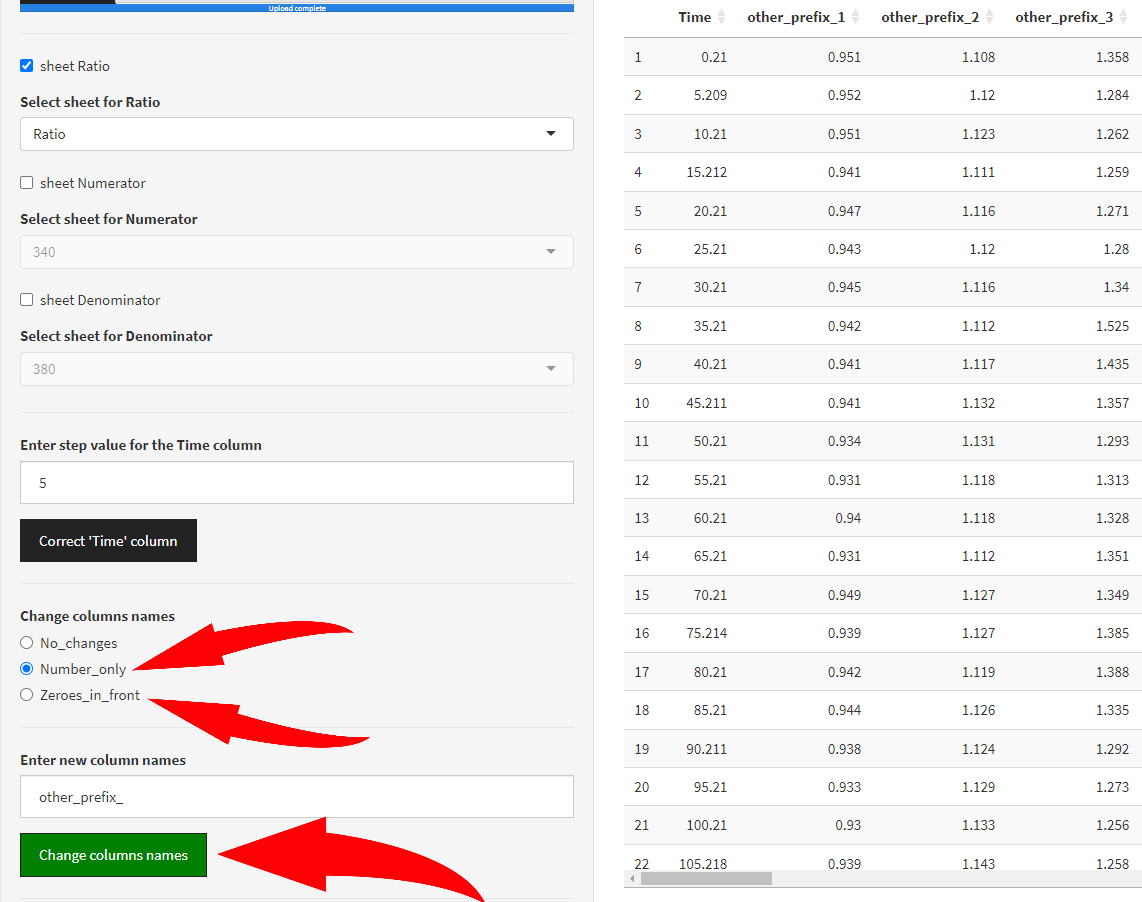
The amount of zeroes depends on the maximum number captured (maximum 9999, i.e. cell-9999). If maximum number has 4 digits and minimum number – only one, therefore the pattern is the following: “cell-0001, cell-0013, cell-0020, cell-0149, cell-1328”. The 2 digits maximum case is presented above.

If user wants to change the prefix from “cell-” to custom prefix, it can be done by entering the prefix in the “Enter new column names” box:



In this case if “No\_changes” radio button is chosen “Change column names” button won’t cause any changes and just will turn green one time indicating that the button was pressed at least one time. As long as this is an action button there are no “grey” state for it, “green” only. User can discard changes by choosing “No\_changes” radio button any time.

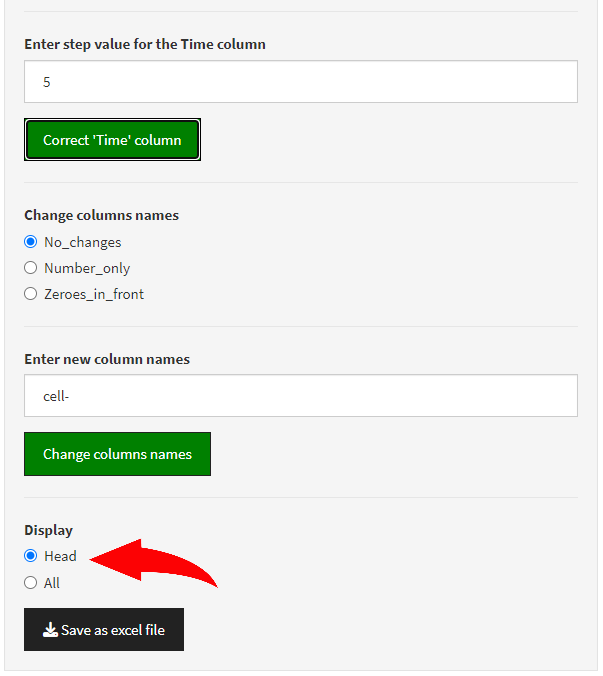
For changes to take effect user needs to choose first the column names style above (“Number\_only” or “Zeroes\_in\_front”) – entered prefix will be used momentarily to change column names. For the custom prefix AFTER this action has taken place user can enter custom prefix value and press “Change column names” button to change the prefix in column names.



After choosing column names style or changing column names prefix “Time [s]” column also changes its name automatically to “Time” only (if you haven’t done it earlier) but no automatic changes to its values are applied though.

## Display option

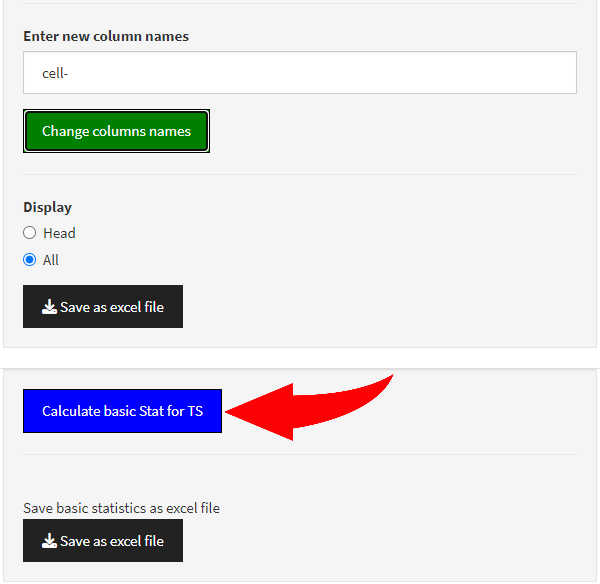
The last “Display” option in this section allows you to just show several rows only in a preview table. It is relevant only for amount of traces > 500 to speed up the process of loading and previewing data.



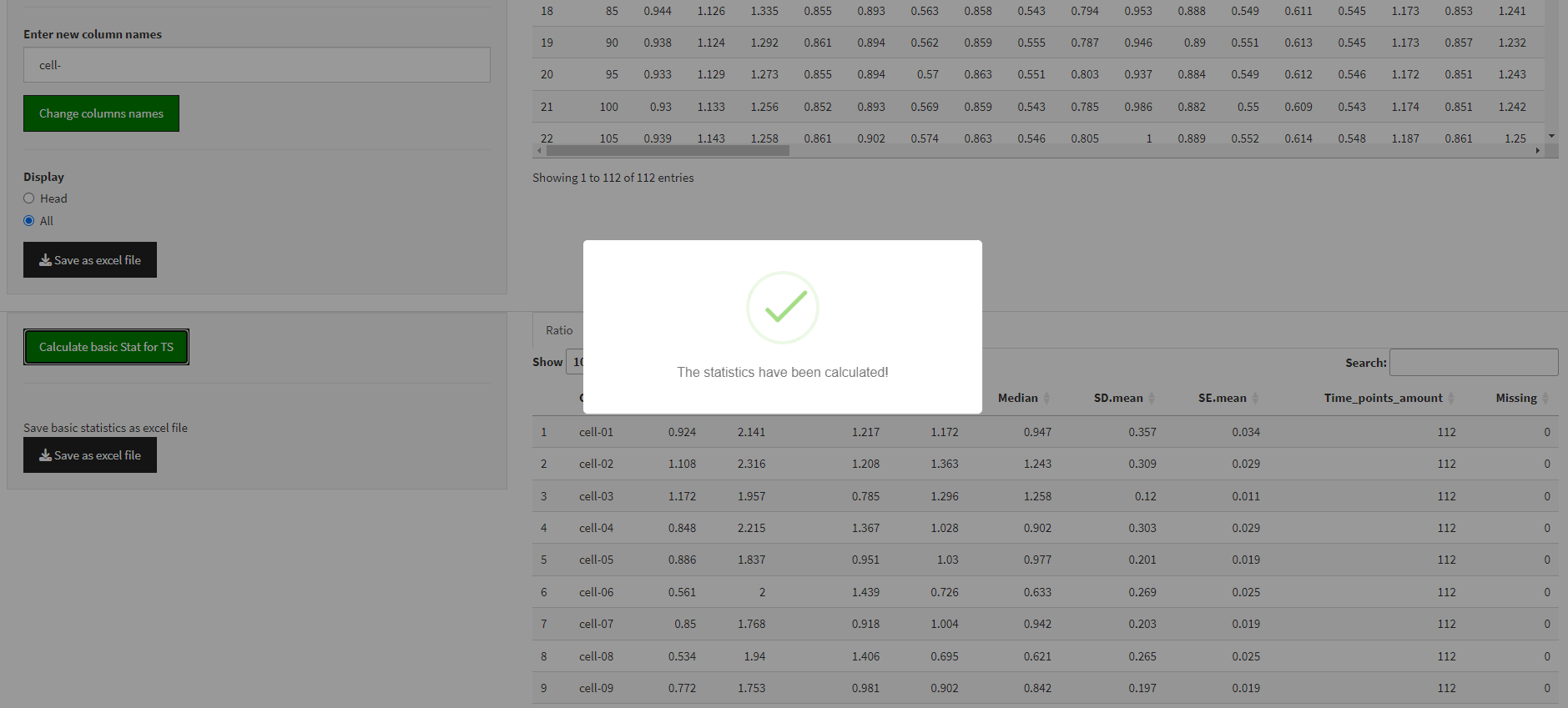
Finally, “Save as excel file” button allows to save an excel file (**{*initial name of the file}-*ProcessedTable.xlsx**) with data that match the preview’s format.

## Calculate basic statistics

This step is mandatory to proceed further with the preliminary data analysis (tab: Preliminary analysis). All the mandatory buttons that are required to be pressed in order to move further are in blue. After pressing it, button turns green. If data was processed a descriptive statistic table will appear at the right to the sidebar.

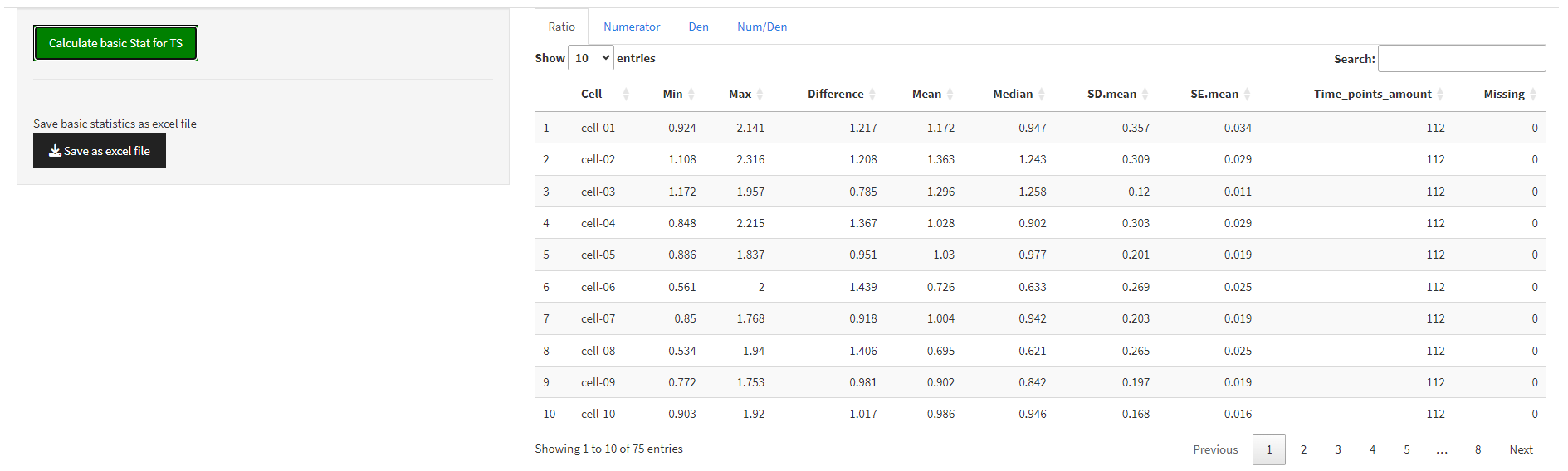


After pressing the button “Calculate basic Stat for TS” user needs to wait until the notification with the green check mark appears:



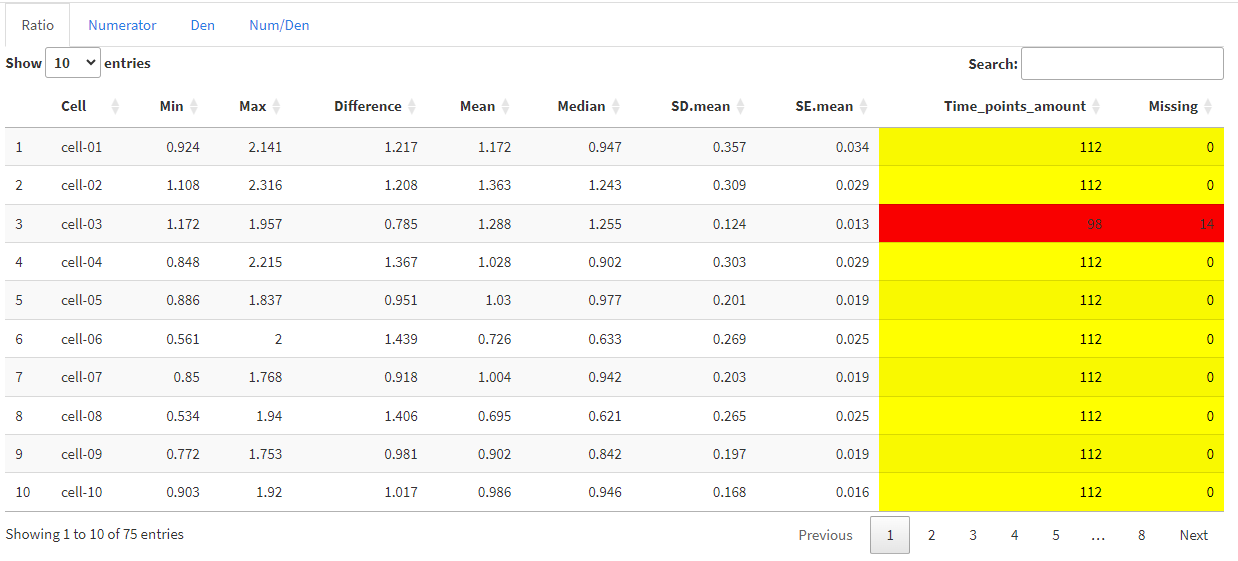
The obtained table with major descriptive statistics parameters have the following columns:

“Cell”, “Min”, “Max”, “Difference”, “Mean”, “Median”, “SD.mean”, “SE.mean”, “Time\_points\_amount”, “Missing”.



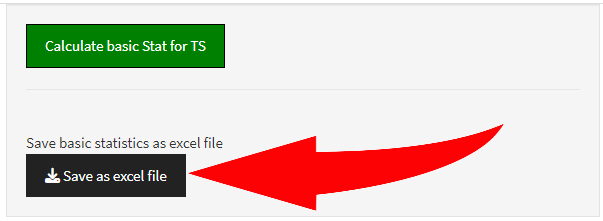
* “Cell”– the name of the initial column. All the parameters below are calculated for all the values in each column of the initial table.
* “Min” – minimum value in the specific column
* “Max” – maximum value in the specific column
* “Difference” – difference between Max and Min
* “Mean” – mean value in the specific column
* “Median” – median in the specific column
* “SD.mean” – standard deviation for values in the specific column
* “SE.mean” – standard error of the mean in the specific column
* “Time\_points\_amount” – the amount of time points in each column
* “Missing” – if some values are absent in the specific column this value will be > 0 indicating their amount

**If there are some missing values last two columns of this table will be marked as yellow and the trace with missing values (row) will be in red (otherwise nothing is highlighted):**



## Save descriptive statistics table as excel file

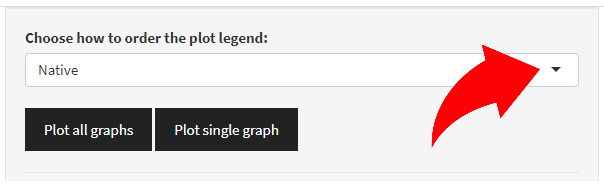
The table with descriptive statistics can be saved as an excel file (**{*initial name of the file}*-BasicStatisticsTable.xlsx**):



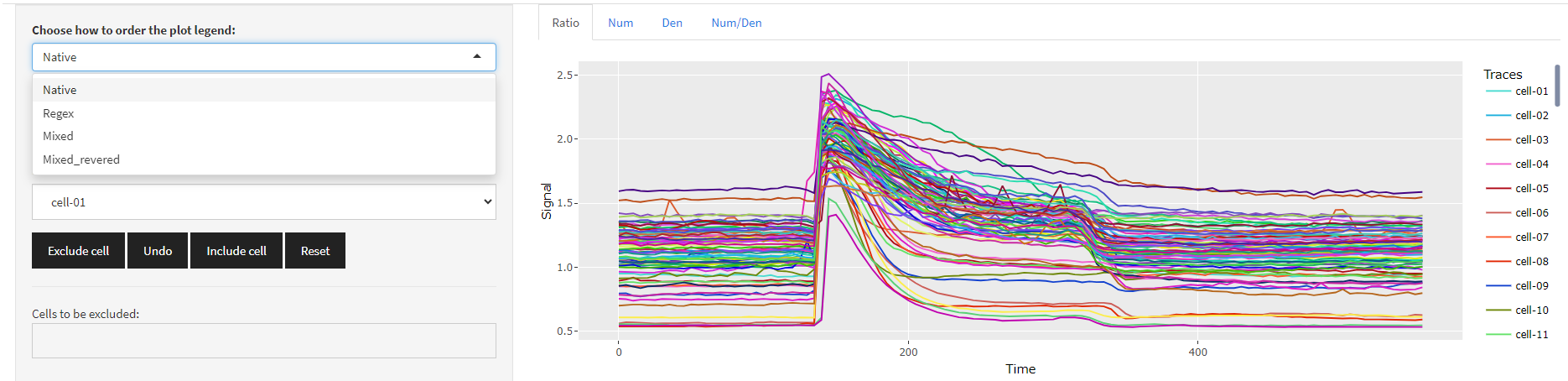
## Creating plots and excluding bad traces

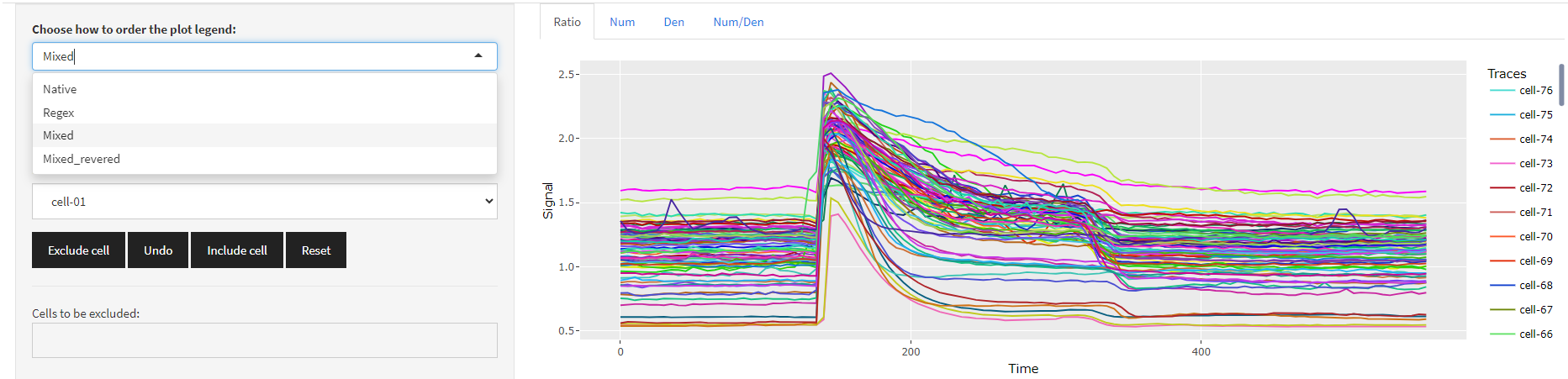
The final step for the “Preliminary analysis” tab allows user to plot all the graphs at once as well as each of them one by one to analyze them and exclude bad traces.

At first user can choose the way to order column names in the table uploaded from the excel file on the first step (the initial one).



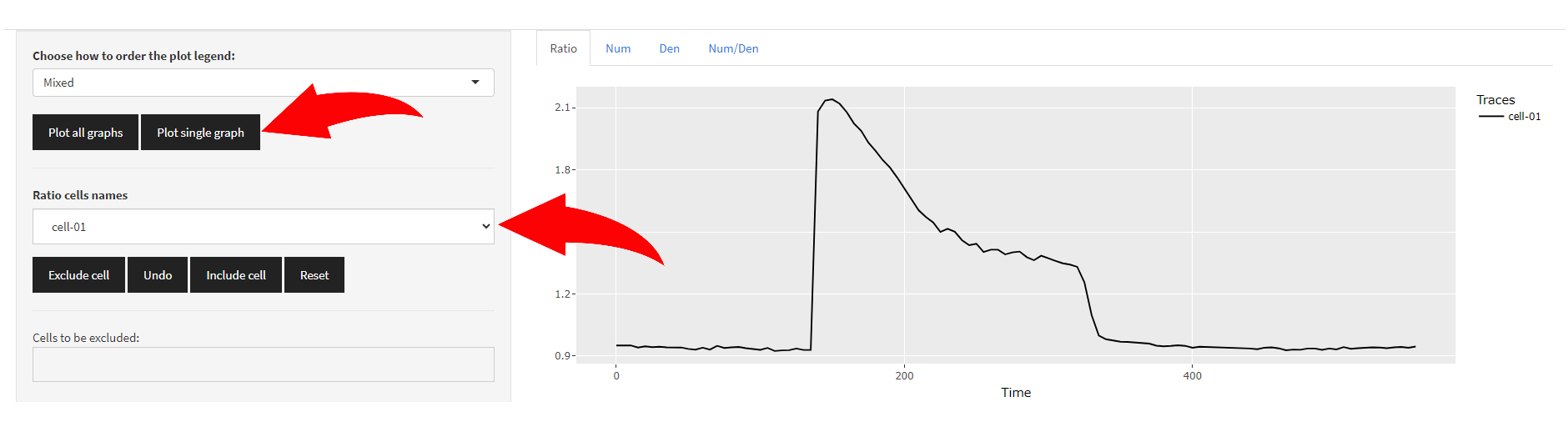
By default the sorting algorithm is “Native” which means that columns are be ordered the same way like in the initial excel file. If for some reasons they were unordered or mixed it is possible to reorder columns which will affect the legend of the plot (when pressing “Plot all graphs”). By looking at the legend when all traces are plotted at the same graph user can define the sorting method of choice:



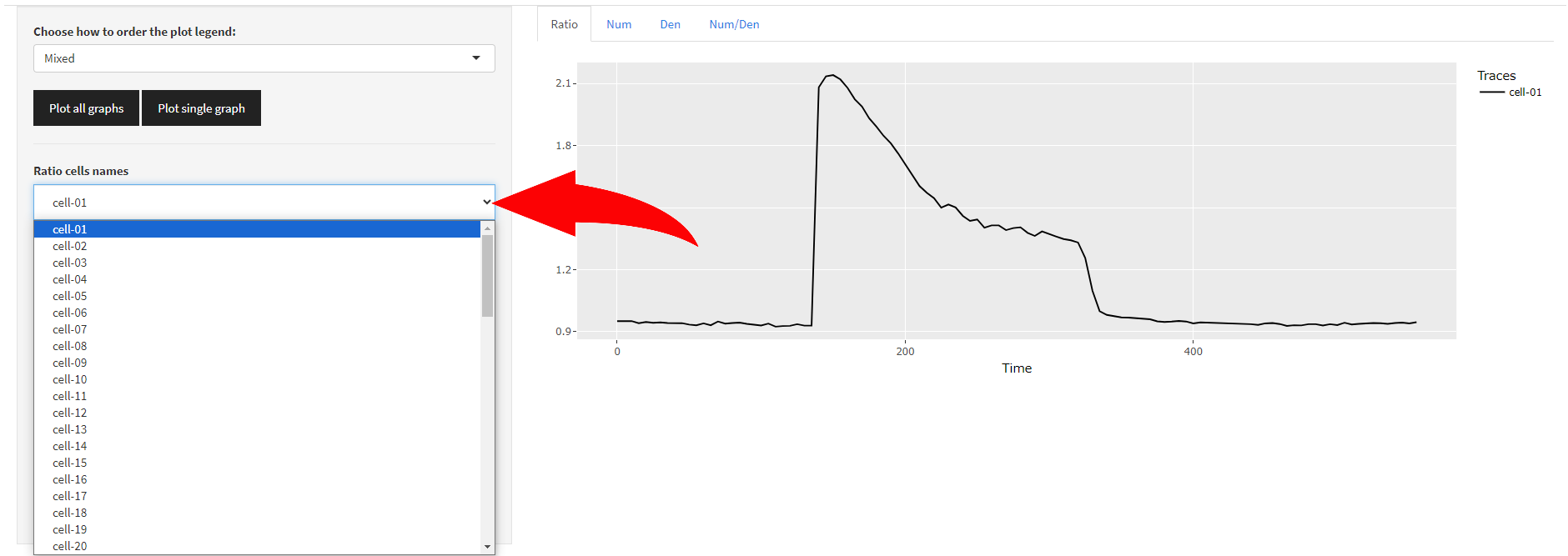


The algorithm to choose depends on a prefix or a column name style and should be manually adjusted if columns are unordered or a custom prefix was chosen.

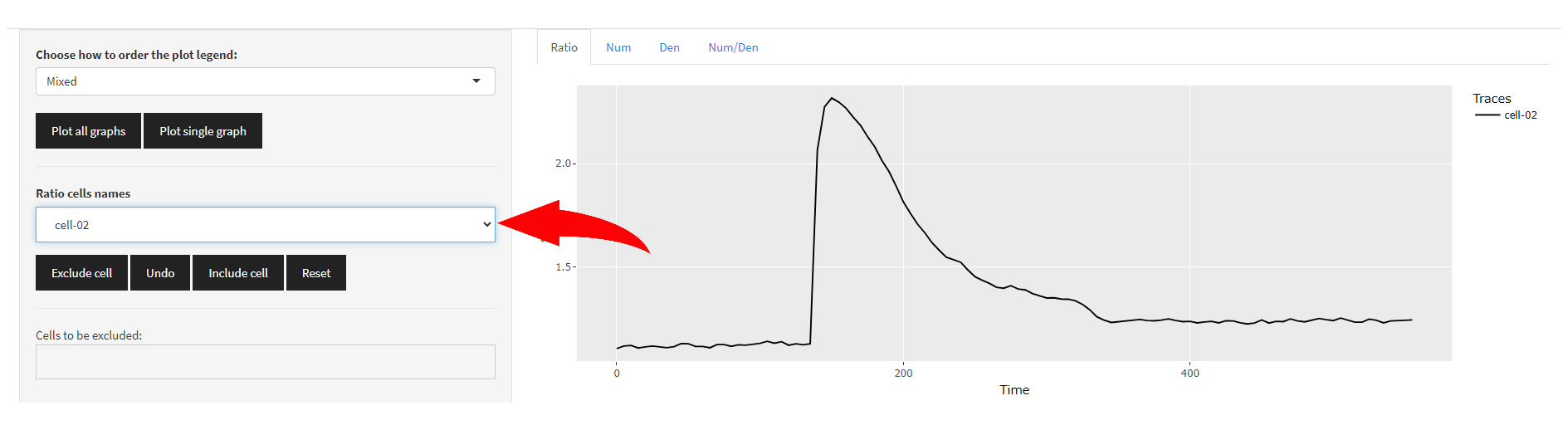
By pressing “Plot single graph” a single plot can be rendered. To choose a specific name of the traces use the “Ratio cells names” box:



To switch traces using keyboard’s arrows (up and down) it is necessary to press ones on the “Ratio cells names” box and choose any trace at first:

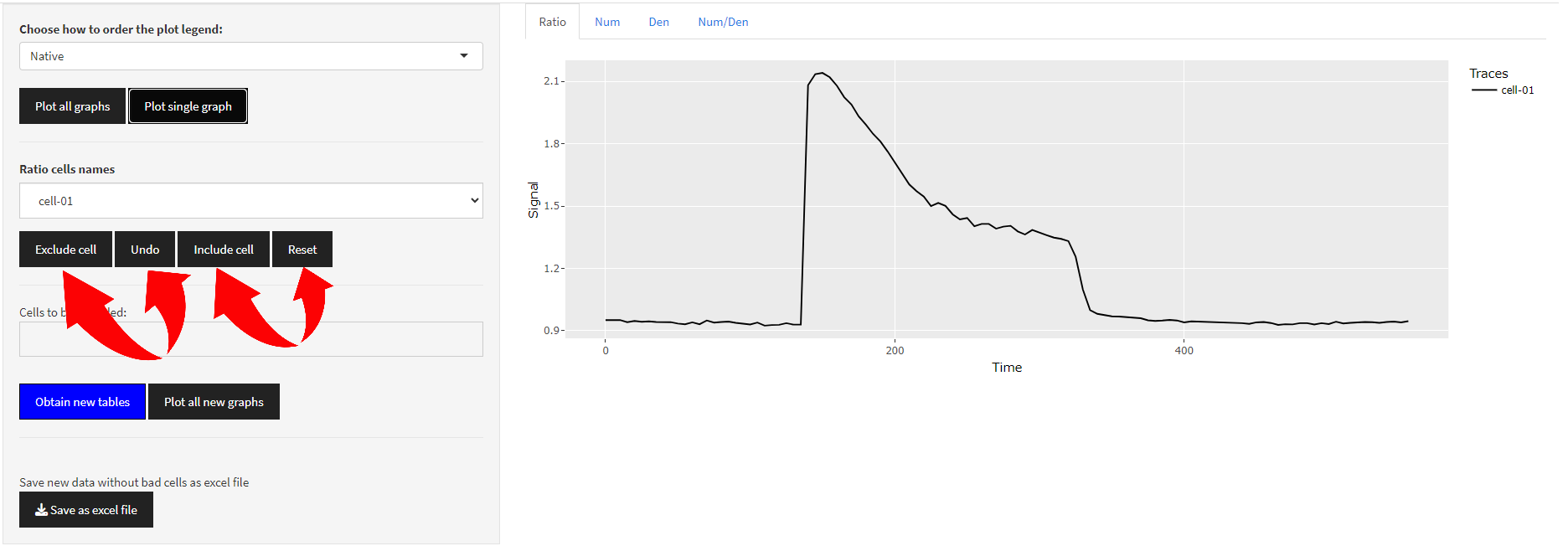


After choosing any trace it will be displayed to the right from the sidebar and user can switch traces (their plots) by pressing arrows buttons (up and down only) for convenience. When this mode is active (switching mode) the box is glowing in blue:

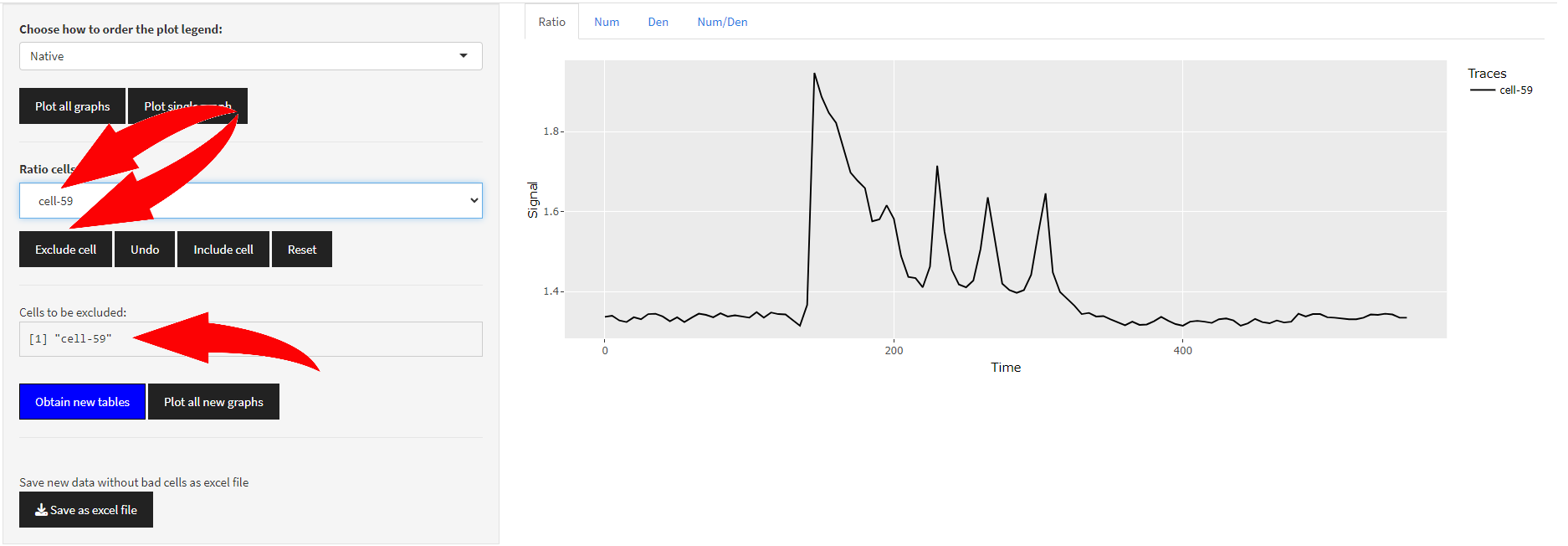


## Excluding traces

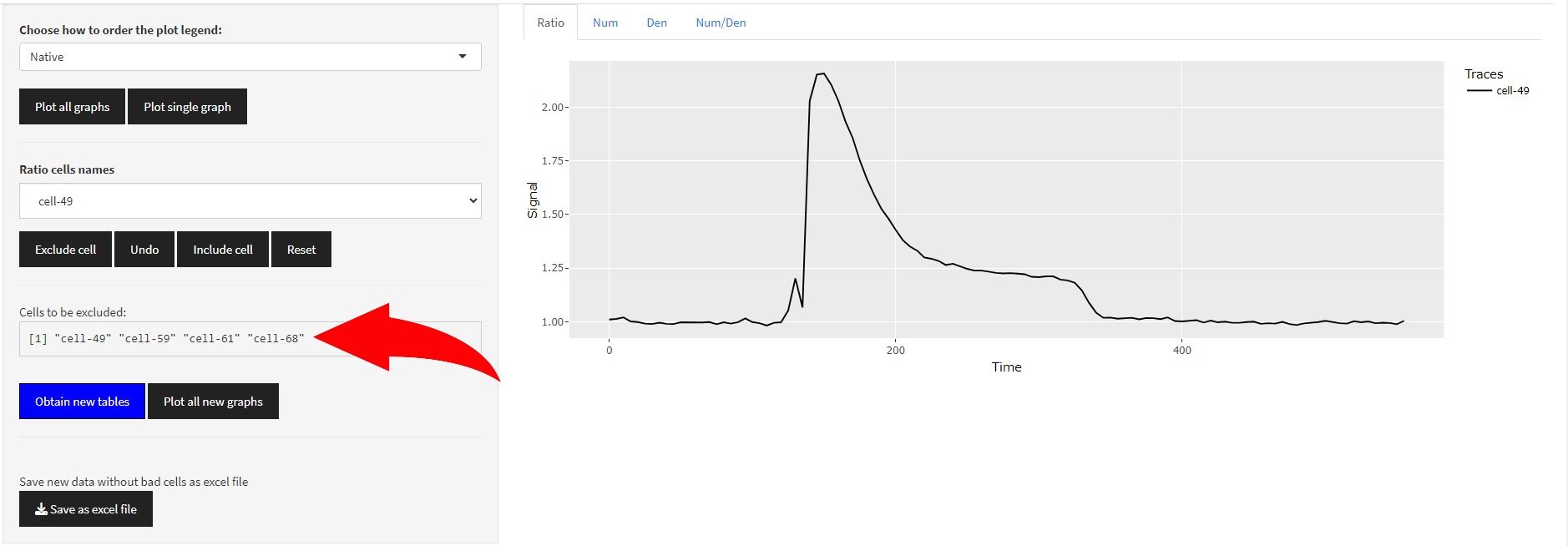
To exclude traces that seem to be outliers there are four buttons that allow user to manipulate data:



To exclude traces one by one user needs to choose the trace in the “Ratio cells names” box and press **“Exclude cell” button**:

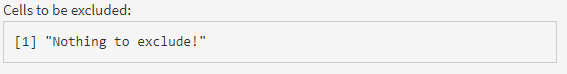


The “Cells to be excluded:” box indicates the name of the excluded trace or a sequence of traces if exclude traces sequentially:

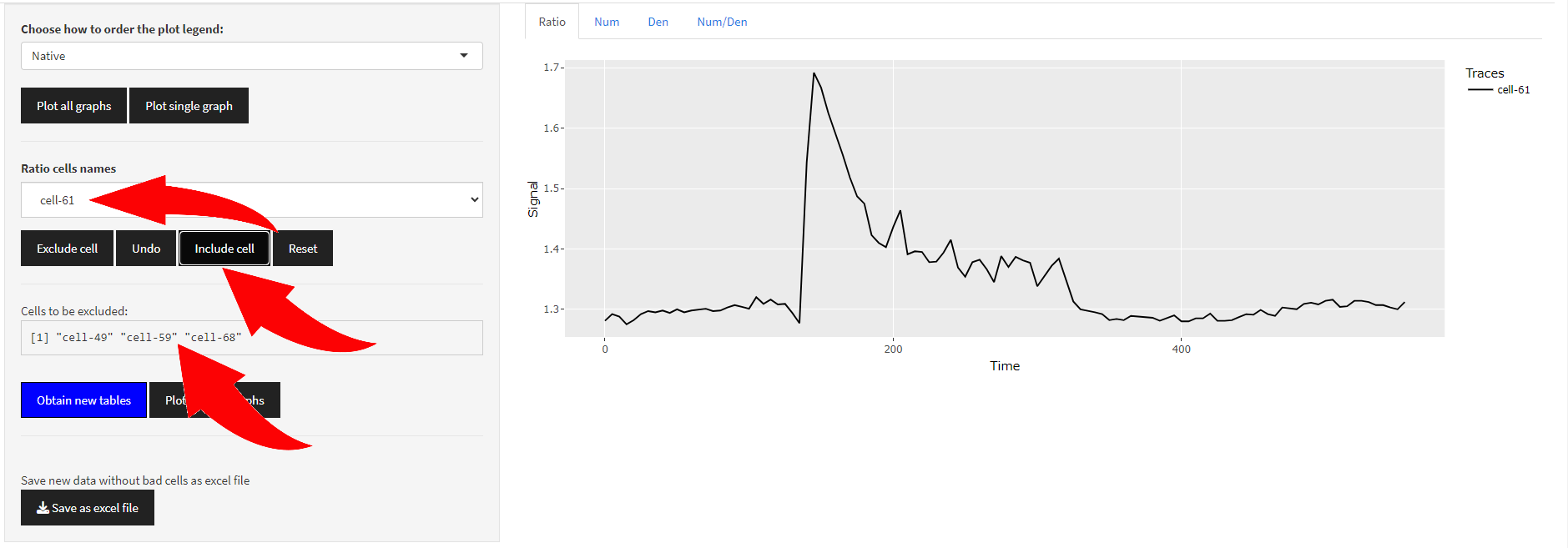


Sequence of trace’s names will be ordered using “Mixed” algorithm, if names were changed to have prefix “cell-” the order will be increasing and for raw names decreasing otherwise.

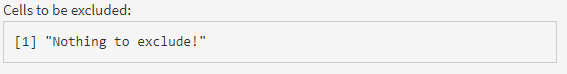
**Button “Undo”** cancel the LAST action. If all actions have been canceled manually user get this message:



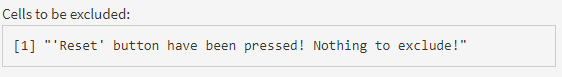
**Button “Include cell”** allows to include in the result table (and exclude from the “Cells to be excluded:” box) any trace of choice. But user is required to pick the name of the trace first in the “Ratio cells names” box.



The same message will appear if all traces have been included in the result table manually:

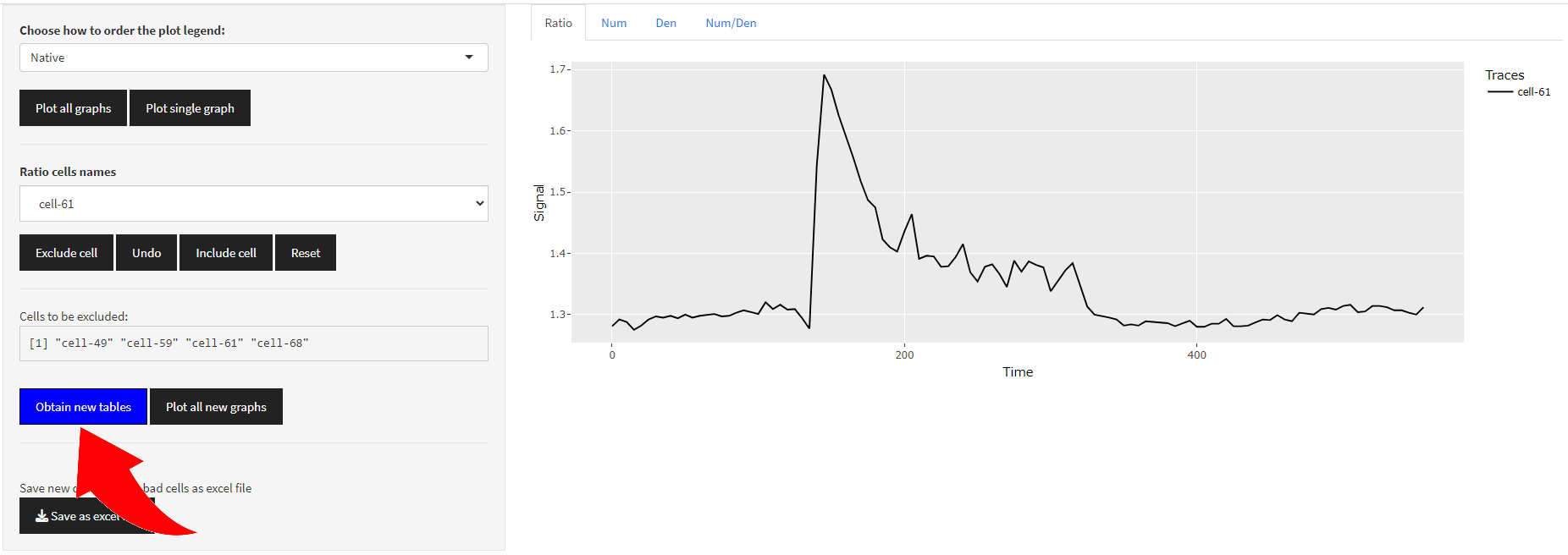


**Button “Reset”** clears the list of excluded traces with the following message:



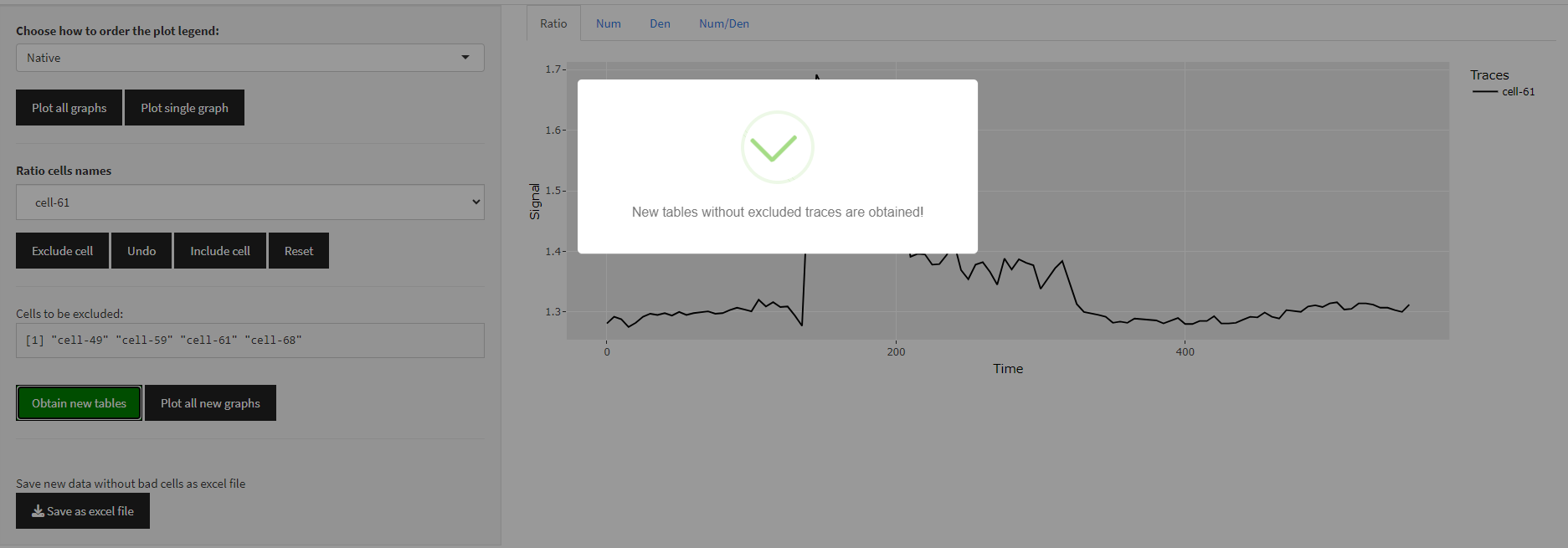
## Discarding unnecessary traces

When all the “bad” traces were chosen in order to save changes the “Obtain new tables” button (in blue, indicating mandatory action) should be pressed.

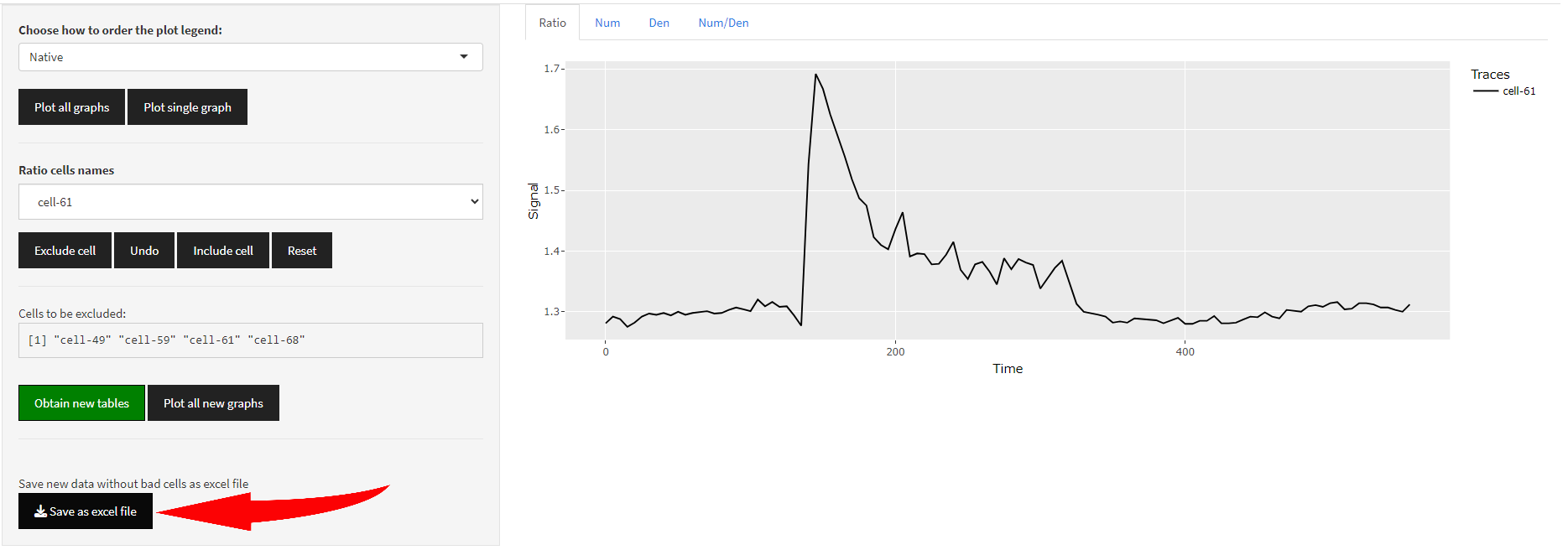


In order to finish the process, user needs to wait for the notification with the green check mark to appear.

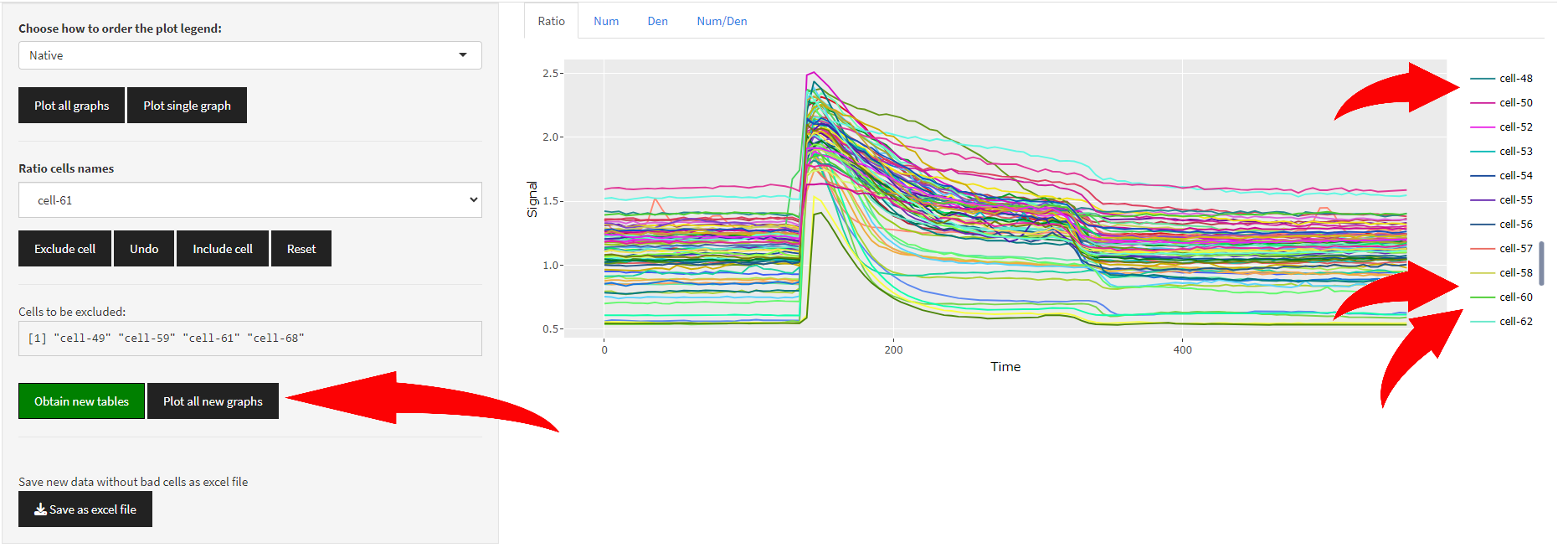
After pressing the “Obtain new tables” button, it turns green indicating that action has been taken.



After that it is possible to save the new table as excel file (**{*initial name of the file}*-CleanTable.xlsx**).



Or plot all new graphs without excluded traces.



Even after that it is possible to look through the whole set of initial traces using the options above and pressing **“Plot single graph” button**. Excluded traces can be included again in the result table by choosing them in **“Ratio cells names” box** and pressing **“Include cell” button**. **After that it is necessary to press “Obtain new tables” button again to save changes!**

# TAB: Analyzing amplitude

# TAB: Shifting curves

## Explaining algorithms

This tab allows user to adjust curves for every trace in order to dismiss possible delay between cells in the response to any compound introduced during the experiment. Altogether, there are two possible algorithms to adjust curves:

* **CCF (cross correlation function)**
* **Alignment of maximums**

**CCF (cross correlation function)** — in this case for each pair of traces a cross-correlation matrix is found. In order to find such matrix each trace is shifted from another one to the right and to the left (the value for that is called “maximum lag for CCF”). If, for example,

“maximum lag for CCF” = 2,

two columns of sample data (two sequences): (2, 4, 6, 5, 8, 9, 10) and (21, 43, 63, 58, 82, 92, 102) will be shifted to the right and to the left for 2 steps:

(2, 4, 6, 5, 8, 9, 10)  
(21, 43, 63, 58, 82, 92, 102)

(2, 4, 6, 5, 8, 9, 10)  
(21, 43, 63, 58, 82, 92, 102)

(2, 4, 6, 5, 8, 9, 10)  
(21, 43, 63, 58, 82, 92, 102)

(2, 4, 6, 5, 8, 9, 10)  
(21, 43, 63, 58, 82, 92, 102)

(2, 4, 6, 5, 8, 9, 10)  
 (21, 43, 63, 58, 82, 92, 102)

For every intersection of values, like in the last example:

X = (6, 5, 8, 9, 10)  
 Y = (21, 43, 63, 58, 82)

the following parameter is found:

= ,

Where and are mean values of each sequence, *N* – amount of values in each intersection of sequences (*X* and *Y* in the example).

Maximum acf value represents the higher correlation between two curves. This value corresponds to the lag value to apply to the second curve in order to get aligned curves. These lag values are put in the correlation matrix. Having this matrix allows to find a trace with the earliest response and shift every other trace to the left in order to get maximum correlation between curves.

**This algorithm is very slow but much more precise.** It is better to use it only if the amount of traces are lower than 300 and “maximum lag for CCF” lower than 60. Otherwise, it can take a long time to calculate.

**Alignment of maximums** — finds local maximums that are more likely related to a response to any temporary compound introduction. It means that a **response on the plot should be represented only as a local maximum**, i.e. values before and after maximum should be lower, monotonous increase/decrease areas won’t be taken into account.

In order to find such value algorithm uses the following approach:

For every trace a moving average and moving maximum values are calculated. The window for moving values is **150 seconds** by default (“response time” value, that can be changed in a box). It is necessary because any random outliers (random local maximums related to noise) or invalid values can lead to incorrect search for response-specific local maximum. In this case the second derivative for moving average that equals “-2” represents the beginning index of local maximum area and moving maximum for this index represents the local maximum that is more likely related to response-specific one.

Each index and related time point of every trace’s local maximum are stored in a list where the trace with earliest local maximum can be found. Using the list’s information algorithm finds lag values for each trace to shift them to the left to match the earliest local maximum.

**This algorithm is very fast but less precise especially if a local maximum of any trace is represented by two close peaks. What’s more important, using this algorithm before launching CCF algorithm won’t increase the CCF’s performance.** The longest stage for CCF algorithm is calculating the correlation matrix with recommended lag values and it will be calculated in any case.