

Ranker for Images

“Ranker” for ImageJ

1. What is Ranker?

In contrast to the common use of *ImageJ*, “Ranker” is a sorter for numerical data, that may be useful by focusing your experiments on a selected group of data showing the highest and lowest rankings according to your measured parameters.

The raw code of “Ranker” can be found at the end of this short manual, also containing the example data used here. Before you use this macro, please read the whole manual.

2. What does “Ranker”?

First, this is NO professional software/macro!

I’m not responsible for any damage that may result from use or modification of the macro.

Please use a more established program if (your) life depends on it.

It worked fine with the data samples I tested and manual sorting produced the same results as application of the macro.

3. To what kind of data this macro can be applied?

You have a group of samples and several results coming from every single sample, the data look basically like this:

Sample	Parameter 1	Parameter 2	Parameter3	Parameter 4	Parameter 5	Parameter 6
Smp(1)	Var(1)-1	Var(1)-2	Var(1)-3	Var(1)-4	Var(1)-5	Var(1)-6
Smp(2)	Var(2)-1	Var(2)-2	Var(2)-3	Var(2)-4	Var(2)-5	Var(2)-6
...						
Smp(n)	Var(n)-1	Var(n)-2	Var(n)-3	Var(n)-4	Var(n)-5	Var(n)-6

Now, you want to find the samples with the highest and lowest rank according to all of the parameters you measured in order to focus your further experiments only on those “top” and “bottom” samples.

4. How does “Ranker” work?

Basically, every parameter is (at first) considered independently. The highest values from every single parameter get the highest “rank score”. If you have 100 samples, the highest value from a single parameter gets 100 points, the lowest parameter gets 1 point.

This is repeated for every single column of parameters. At the end, the single “ranking scores” from all parameters of every single sample are summarized to identify the “winners” and “losers”.

Also, all samples are subdivided into an arbitrary (you decide) number of groups, covering the range between the lowest and the highest “ranking scores”. The sum of all groups (n) cover the whole range from the lowest to the highest rank, a single group covers the range (highest rank - lowest rank)/n).

Let’s have a practical example.

You have samples of tissue slices and get several different results per single sample.

Like **autofluorescence** (resulting from so-called AGEs, advanced glycation end products, resulting from the reaction between sugars and proteins),

protein carbonyls (the most common marker of oxidative protein modification),

3-nitrotyrosine (a typical marker of inflammation, induced by the decay products of peroxynitrite),

DNA-damage (via immunostaining of 8OH-dG),

argpyrimidine (specific immunostaining of a single AGE),

Nrf2-induction (as response to a shifted cellular redox-state),

and finally **S-glutathionylation** (a reversible posttranslational protein-modification in response to cellular redox-shifts).

After many productive days in the lab, your final results may look like this:

Sample	Autofluo.	Carbonyls	3-NT	DNA-Damage	ArgPyr	Nrf2-induction	S-gluth.
1	0.94	7.25	0.24	2.57	4.77	5.85	5.81
2	0.04	0.26	0.61	2.47	1.58	12.40	2.65
3	0.43	3.38	0.44	0.59	1.39	2.26	19.82
4	0.92	7.90	2.88	0.53	0.19	6.89	15.99
5	0.47	1.84	2.57	3.44	5.99	2.73	18.15
6	0.24	0.32	1.83	4.15	4.39	10.80	1.44
7	0.77	4.45	4.94	3.68	9.32	9.39	2.17
8	0.10	6.73	3.20	4.63	1.97	5.40	12.92
9	0.60	9.82	4.15	6.93	10.33	9.48	6.76
10	0.51	9.53	4.39	1.49	4.08	2.74	16.96
[...]	[...]	[...]	[...]	[...]	[...]	[...]	[...]
295	0.33	0.32	4.55	5.97	7.77	6.68	19.72
296	0.61	9.18	1.33	4.42	0.91	10.36	1.16
297	0.56	9.49	2.38	3.70	1.36	0.12	14.51
298	0.56	8.13	2.60	1.75	8.69	3.41	8.76
299	0.43	8.00	1.87	5.80	6.00	0.31	5.70
300	0.37	2.78	4.58	5.19	0.38	9.32	14.72

For this specific example, we used random numbers generated in Excel (“RAND()*constant”).

Now, you want to sort the samples according to their “ranks” - from the samples with the highest values in all measured classes (autofluorescence, carbonyls, 3-nitrotyrosine, DNA-damage and argpyrimidine, Nrf-2 induction and S-glutathionylation of proteins) to the ones with the lowest values in all classes, in order to focus only on those specific samples.

It is easy to see that this becomes more and more difficult to do this “manually” with an increasing amount of parameters from every single sample. While a single parameter measured can just be sorted from the lowest to the highest value, this becomes very pedestrian with dozens or even hundreds of parameters you got from every single sample.

5. Prepare your data

Just export your data into a text file with tabulator-separated columns:

The first line contains the headers/titles of every column and the first column contains the names of the sample. Due to the “rules” of arrays in ImageJ-macros, the sample names should be also numbers to get naming of the data less “problematic”. Just replace the names of your samples in the according software with numbers and replace the numbers later with the real names again.

Also you should remove empty lines from the bottom of the table. Furthermore it is important, that all cells of your table are filled with numbers. Empty cells are not processed. Thus, it is up to you, if you fill the empty cells with zeros or average values from the other samples.

After preparing your data just export them with tabulator-separated columns into a text-file.

And that's almost all, now just start the macro in ImageJ:

Plugins>Macros>Record,

Paste the macro in the opened window and click the "Create"-Button.

Run the macro in the new window.

Now, it becomes "tricky". You will not get any results (now), but a new macro is generated in the "Log"-table that is the actual macro to be executed for evaluating your data.

You may ask why? The basic idea was to create a macro that contains all the data and can be also modified in order to change data processing.

Copy the generated macro from the "Log"-Window and paste it into the (clean and empty!) macro recorder. Create a new one and start it, and THEN you will get the final results! Copy the results and paste them into your statistics software. If you use German settings on your machine, you may replace every period with a comma (do this in the editor).

This may be a little bit cumbersome, but it is still much quicker done than sorting the values for every single parameter manually and sum up the ranking-points, since you would have to produce 2 columns for every single parameter: the sample names and the measured values. Then you would have to sort both columns in order to rank the values and up the gained points in every case. If you have large tables with dozens of values from hundreds of samples you don't want to do this manually.

Modifying the parameters and final data processing:

Before you evaluate the final macro, you can change some parameters. The most important one is the number of groups your data are finally subdivided into. This number can be changed in the very first line of the final macro generated for data processing.

Copy the results from the Log-Window and paste it into software like Excel. If you have "German settings", replace the decimal separators (points) with commas using the editor before pasting the data in Excel.

Now, you should see a table like this:

Sample	Rank sum	Group (desc)	Group (asc)	Place #
1	897	4	7	232
2	500	1	10	298
3	652	2	9	288
4	1092	5	6	121
5	971	4	7	191
6	735	3	8	277
7	1250	6	5	65
8	949	4	7	208
9	1586	9	2	3
10	1188	6	5	79
[...]				

In order to sort the samples according to the "Rank Sum" or the according "Place #", just use this sample in Excel. After sorting, you should see this:

Sample	Rank sum	Group (desc)	Group (asc)	Place #
12	1863	10	1	1
216	1641	9	2	2
9	1586	9	2	3
254	1544	8	3	4
101	1543	8	3	5
163	1524	8	3	6
58	1519	8	3	7
196	1494	8	3	8
191	1480	8	3	9
181	1479	8	3	10
[...]				

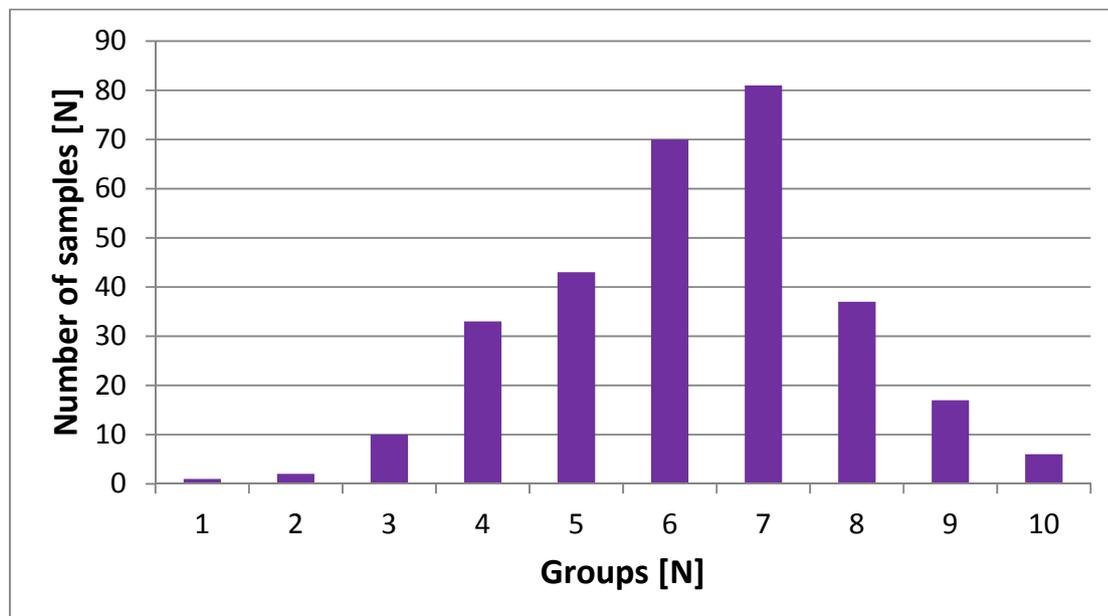
In this specific example, the samples are subdivided into 10 different groups, covering the rank-range from the highest to the lowest sample. "Group (desc)" and "Group (asc)" show the according group for every single sample, "desc", if the highest values are put into the highest groups and "asc", if the highest samples are sorted into the lowest groups.

"Rank sum" represents the overall amount of points the according sample accumulated. The highest possible "Rank sum" is the number of samples multiplied with the number of parameters. Here, we have 300 samples and 7 parameter, the maximal "Rank sum" is 2100 (if a sample "wins" in every single category, always gaining the 300 possible points), the lowest one is 7 (if a sample "loses" in every single category and gets only a single point every time).

Furthermore, you get also information about the groups:

Groupranges				
from	to	Group (asc)	Samples in Group	Group (desc)
452	593.1	1	1	10
593.1	734.2	2	2	9
734.2	875.3	3	10	8
875.3	1016.4	4	33	7
1016.4	1157.5	5	43	6
1157.5	1298.6	6	70	5
1298.6	1439.7	7	81	4
1439.7	1580.8	8	37	3
1580.8	1721.9	9	17	2
1721.9	1863	10	6	1
Groupwidth=141,1				

The overall range (in this special case) is between 1863 and 452 points. This range is subdivided into 10 different groups, every single group covers 141.1 points. "Samples in Group" shows the number of samples that was assigned to it. If we plot this (for this special example) we can see a distribution like this: most samples are found in the groups 5, 6 and 7, representing the "bulk" of the samples with average "Rank sums", while the "extreme" samples (the highest and the lowest ones) are found in the outer groups 1, 2, and 10.



Just one last thing: if you use this macro and especially if you publish your data, please mention and cite “Ranker”.

Best regards,

Tobias Jung

June 2018