

# User guide for ExNormalizeMets:

For this guide, a windows machine is being used. No prior software other than Microsoft Excel and a web browser are required for installation. The guide will explain how to install R to be able to use the NormalizeMets R package through its ExNormalizeMets Excel interface (no direct interaction with R or command line functions is needed).

A detailed guide with illustrations for using the software is also provided, showing its main functionalities in a ‘walk through’ guide of an example project (both the edited excel workbook and data used for the guide are provided).

For support contact: [g.olshansky@student.unimelb.edu.au](mailto:g.olshansky@student.unimelb.edu.au).

## Getting the files ready:

To get the required files, go to <https://github.com/metabolomicstats/ExNormalizeMets> and download the latest version of the *ExNormalizeMet*:

metabolomicstats / ExNormalizeMets

Code Issues 0 Pull requests 0 Projects 0 Insights

Excel edition integrating the NormalizeMets R package

1 commit 1 branch 0 releases 0 contributors

Branch: master New pull request Find file Clone or download

metabolomicstats Submission version

Name	Type	Last modified
ExampleData	Submission version	18/01/2018 3:14 PM
Installation	Submission version	18/01/2018 3:14 PM
.DS_Store	Submission version	18/01/2018 3:14 PM
ExNormalizeMets_0.24.xlsx	Submission version	16/01/2018 7:22 AM
General_Report_example.html	Submission version	16/01/2018 7:22 AM
MyFirstNormalizeMetsProject_example.xlsx	Submission version	16/01/2018 7:22 AM
R-3.4.3-win.exe	Submission version	16/01/2018 7:22 AM

Clone with HTTPS <https://github.com/metabolomicstats/ExNormalizeMets>

Open in Desktop Download ZIP

After the download is complete, unzip the folder, it should contain the following files:

Organize Open with Adobe Acrobat Reader DC Share with Print E-mail New folder

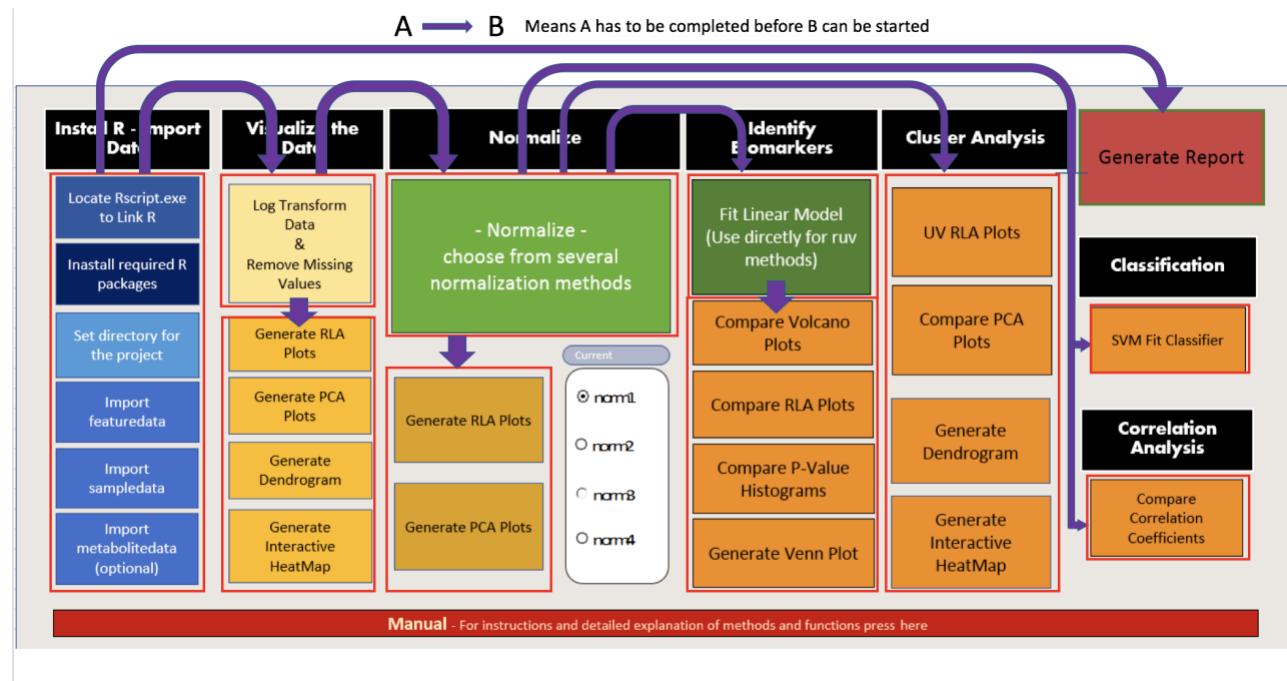
Documents library ExNormalizeMets-master

Name	Date modified	Type	Size
ExampleData	18/01/2018 3:14 PM	File folder	
Installation	18/01/2018 3:14 PM	File folder	
ExNormalizeMets_0.24	16/01/2018 7:22 AM	Microsoft Excel M...	794 KB
ExNormalizeMets_manual	18/01/2018 3:04 PM	Adobe Acrobat D...	4,438 KB
General_Report_example	16/01/2018 7:22 AM	Firefox HTML Doc...	16,771 KB
MyFirstNormalizeMetsProject_example	16/01/2018 7:22 AM	Microsoft Excel M...	2,325 KB
R-3.4.3-win	16/01/2018 7:22 AM	Application	80,445 KB

PDF ExNormalizeMets\_manual Date modified: 18/01/2018 3:04 PM Date created: 18/01/2018 3:46 PM Size: 4.33 MB

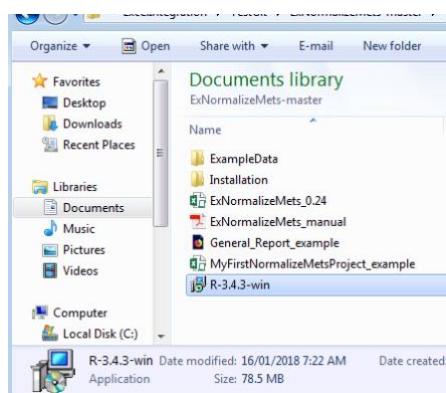
They are all the files needed for installing R and loading NormalizeMets through excel.

## General Workflow:

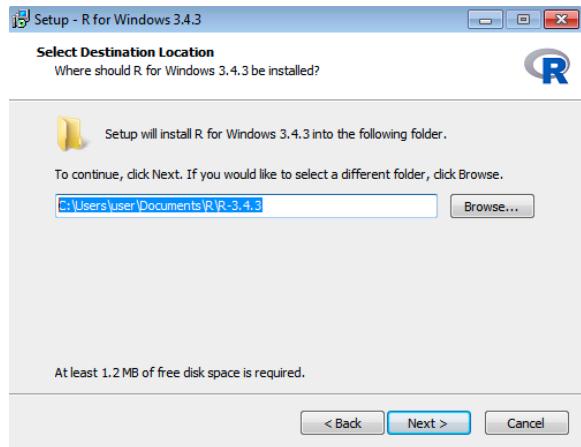


## Installing R:

1. Install R from the extracted ExNormalizeMetsSetup folder by running the *R-3.4.3-win* file.

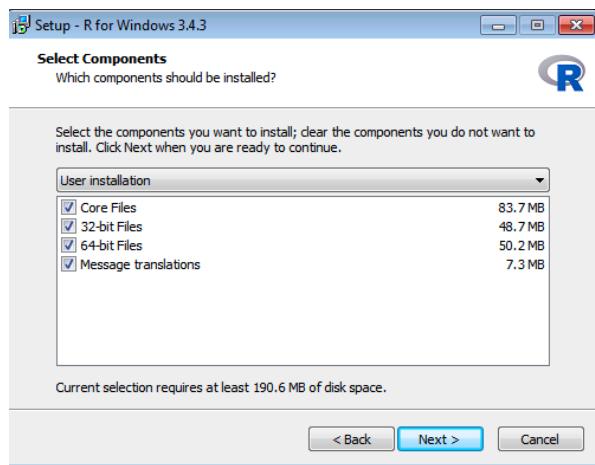


2. Follow the installation instructions and choose the location where R should be installed.

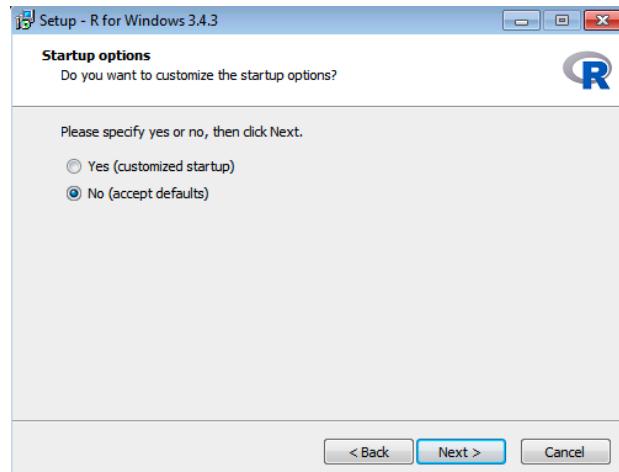


Click *Next* for default location (Recommended) or choose location manually. Make sure you **remember where R is installed** as you will need to locate this folder later to link R to Excel.

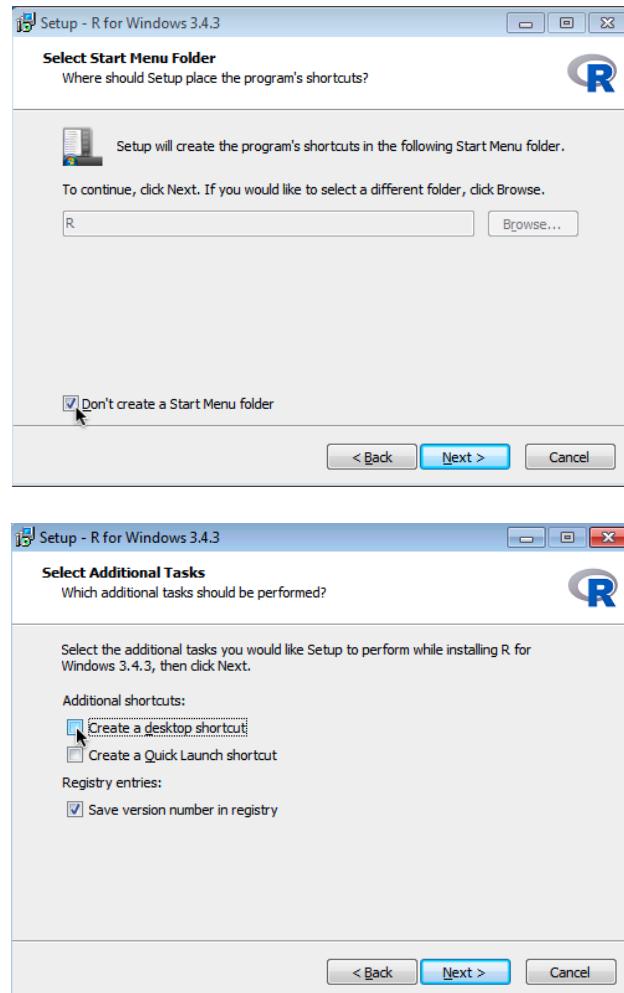
3. Click *Next* to install with the different required settings:



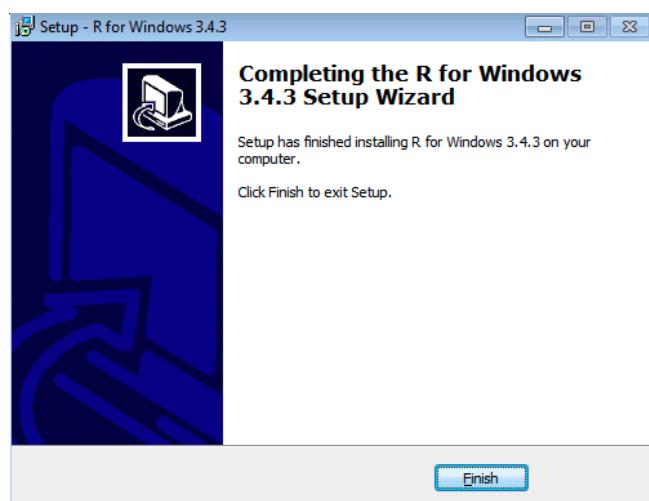
4. Click *Next* to install with default settings:



5. In the next screen check the *Don't create a Start Menu folder* icon if you don't intend to use R by itself and click next. Also uncheck the *Create a desktop shortcut* in the next screen:

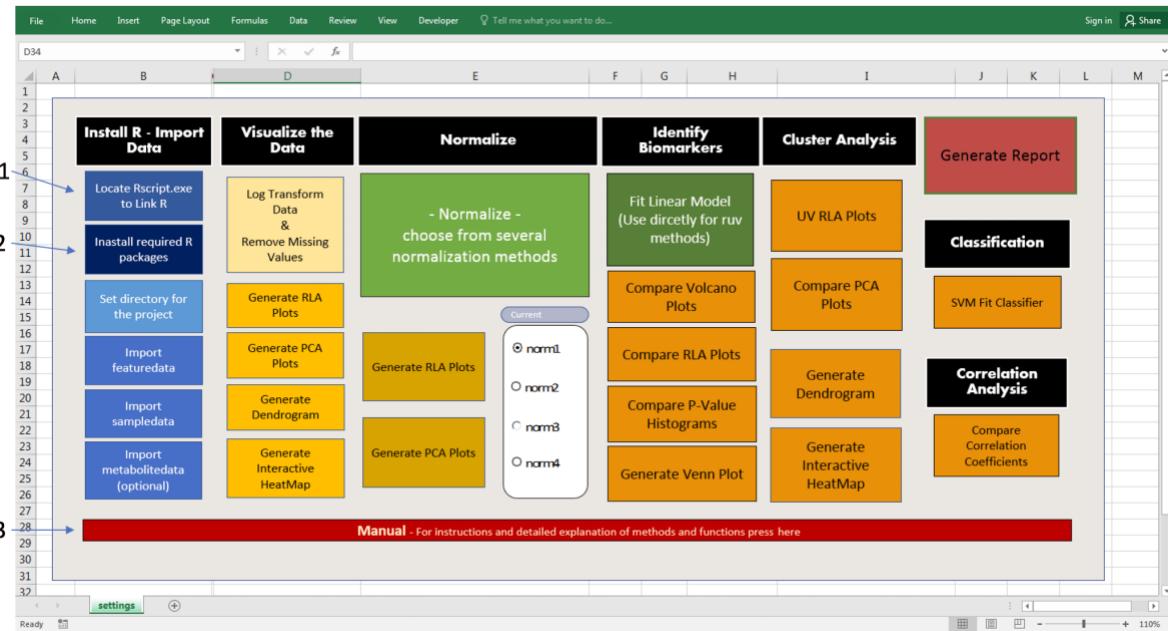


6. Wait until installation is done, the following screen should appear



## Linking R to excel and installing the required packages (first use only):

After installing R open the excel file ExNormalizeMets.xlsx, this will open the Excel interface onto the settings sheet:

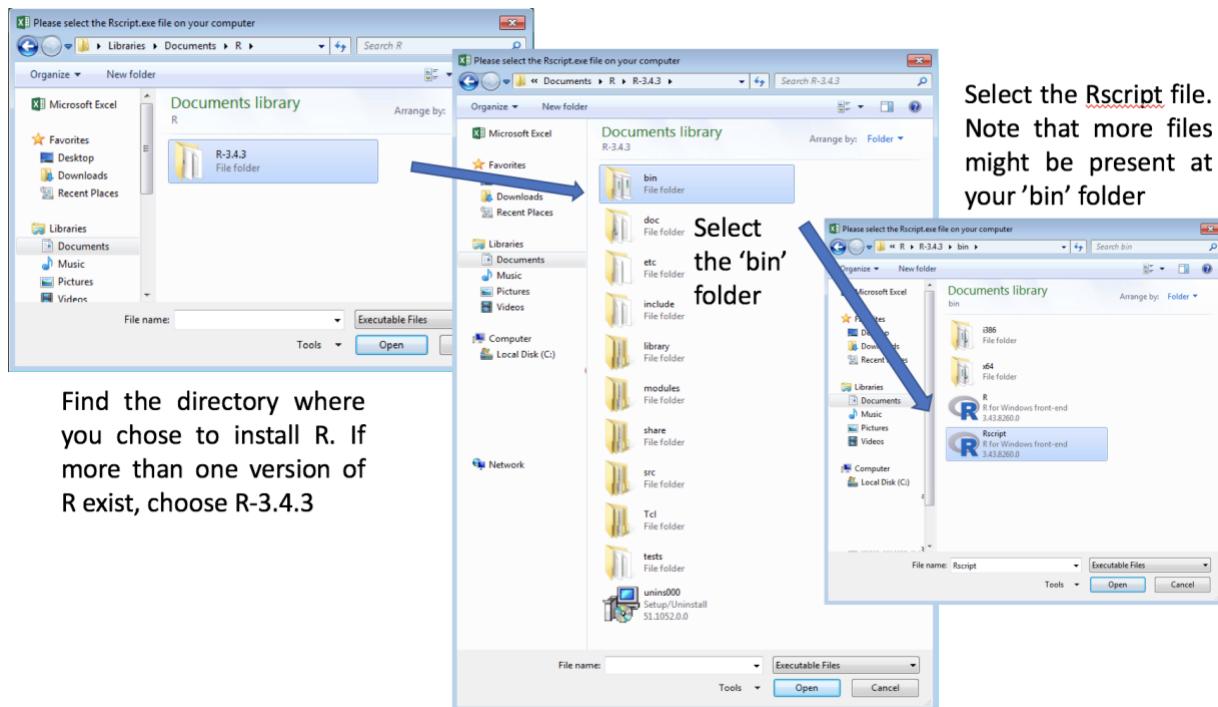


The *settings* sheet is your ‘**Control sheet**’, any function you want to run, from importing data, Normalizing, viewing results and opening the manual can be done from here.

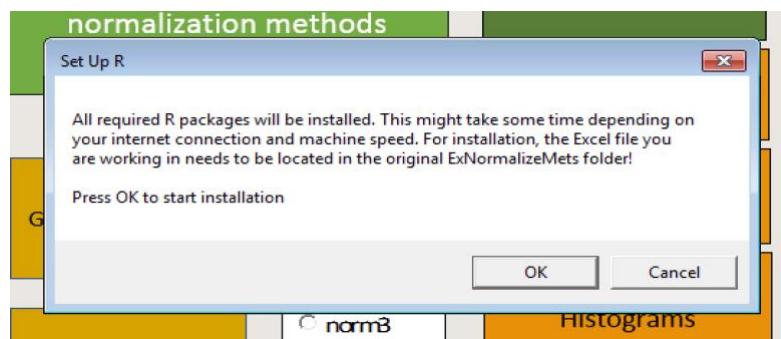
This manual can be accessed at any stage by clicking on (3) but more on this later. For now, first locate the Rscripts.exe file so that excel will know how to run R commands it generates.

### **Locating Rscript:**

Press (1. ‘*Locate Rscript.exe to Link R*’) to locate the Rscript file in the window that opens up. Make sure to select the *Rscript* file in the *bin* folder of your R installation:



After selecting the file, to install all the required R packages and set up the needed dependencies, press (2. ‘Install required R packages’).



Pressing ok will start the installation, this might take a few minutes if you are using R for the first time as many of the base packages will need to be installed.

The installation window looks like this:

```
downloaded 45 KB
trying URL 'http://cran.us.r-project.org/bin/windows/contrib/3.4/cran_0.0.20.zip'
Content type 'application/zip' length 246551 bytes <240 KB>
=====
downloaded 240 KB

trying URL 'http://cran.us.r-project.org/bin/windows/contrib/3.4/knitr_1.18.zip'
Content type 'application/zip' length 913875 bytes <892 KB>
=====
downloaded 892 KB

trying URL 'http://cran.us.r-project.org/bin/windows/contrib/3.4/rmarkdown_1.8.zip'
Content type 'application/zip' length 2291306 bytes <2.2 MB>
=====
downloaded 2.2 MB

package 'GGally' successfully unpacked and MD5 sums checked
package 'plotly' successfully unpacked and MD5 sums checked
package 'ggplot2' successfully unpacked and MD5 sums checked
package 'htmlwidgets' successfully unpacked and MD5 sums checked
```

When the installation is done, a window with the message *Done!* will appear.

Now that the installation is complete, NormalizeMets is ready for use!

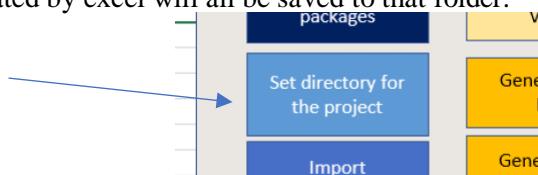
## Using NormalizeMets

Data used in the following examples is provided with NormalizeMets (alldata\_eg in R), it is located in the *ExampleData* folder in your downloaded *ExNormalizeMetsSetup* file. Future references in this guide refer to this data by default.

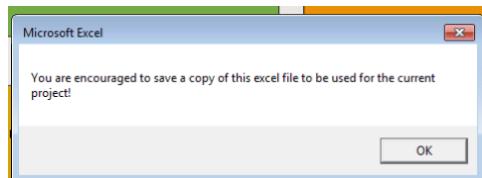
An example excel document containing the data used for the tutorial with all settings identical to those in the tutorial is provided (MyFirstNormalizeMetsProject\_example.xlsxm), if using this document, make sure to set the Rscript location and working directory for your machine.

### Starting a new project:

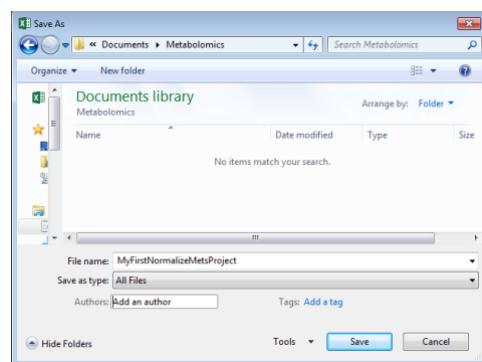
To start a new project, you will need to set up or use an existing directory where the project will ‘live’, data and plots generated by excel will all be saved to that folder.



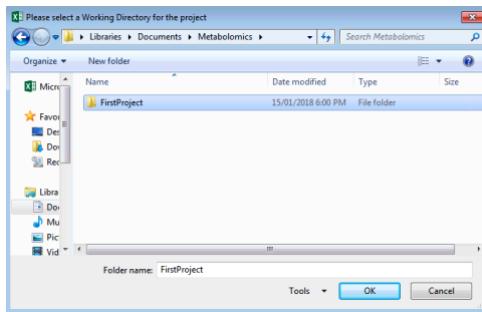
Before selecting the folder, you will be prompted to save a copy of the current version of the excel file, it is recommended to save it with a new name to make sure a ‘clean’ version always stays in your ExNormalizeMetsSetup folder.



After saving the workbook under the name of your choice:

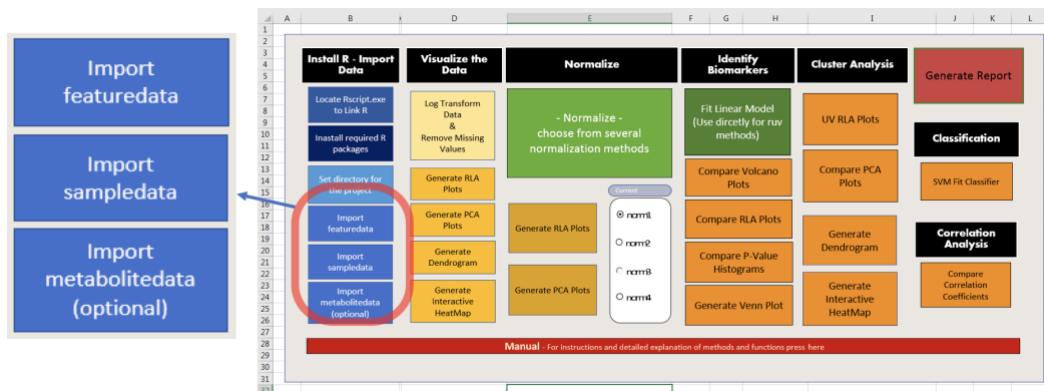


Select the working directory where all files should be generated. We recommend making a new folder for each project.



## Loading data:

To load data for the project, in turn click on the following to load the relevant data:



Loaded data needs to be in .csv format. After clicking on the required file, it will open a new sheet, showing the imported data. Select the setting sheet import more data and get back to the options.

For *featuredata*, set metabolites in columns and samples in rows. Unique sample names should be provided as row names.

MyMetaboliteMetProject - Excel													
A	B	C	D	E	F	G	H	I	J	K	L	M	
1	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9	m_10	m_11	m_12	
2	10485.86719	33220.5625	1112.079492	1408.639648	455.7529297	1100.402344	122.3631592	3855.804688	1637.482422	519.292969	28793.65625	5200.7	
3	8960.46375	20995.51563	036.3890035	520.0440119	873.6009375	2301.246004	172.0802003	5000.6367119	2011.509077	8862.859375	31582.09375	6183.9	
4	10160.44531	28559.64063	1230.333008	1306.320313	1027.507813	2066.591797	269.7998047	4483.00625	1644.607422	776.335938	4494.59375	10505.1	
5	8794.476563	27593.75	901.762207	1800.083004	675.0395508	1675.849609	287.9992676	4969.261719	1508.802734	8405.171875	37030.402625	10726.4	
6	8956.921875	28161.76563	975.1723633	818.0366211	904.0253906	1245.991211	167.413301	4302.033594	1460.012699	6976.429688	29579.53125	5421.9	
7	9092.257813	31685.3125	636.6899414	478.3811038	980.4233398	1259.563477	97.5625	4406.34375	2349.587891	7090.136719	39835.71875	5057.8	
8	2271.044922	7692.175781	143.1688221	109.83900404	1387.791992	149.0966797	1543.40625	961.7314453	2356.685547	20212.703113	4483.1		
9	7850.402344	26462.79683	822.5991211	1341.21582	583.5830078	2164.126953	227.7696533	5030.6875	1604.274414	6241.457031	30873.14063	8130.1	
10	10969.0625	21605.14063	408.0004883	1499.479492	452.2375488	915.5283203	234.4975586	4657.453906	1847.6875	6606.15625	34395.96875	7142.3	
11	1743.932617	7647.645451	218.4199219	127.4294434	408.2912598	999.5717773	74.82695533	1510.011719	895.9853516	2407.103516	18256.14063	2358.1	
12	1.11	1284.412109	4819.99609	61.62762451	15.51902009	394.8364258	1112.143555	28.55443115	1163.742188	628.7016602	2646.07070313	10544.94531	1728.
13	1.12	2272.923828	6813.679683	243.9420166	280.7983398	434.3505859	1167.960938	86.25360107	1665.297852	1009.748047	3927.849609	21808.45313	6546.9
14	1.13	9938.0625	28498.1875	577.8486318	1302.723633	350.9599609	1200.636719	134.2883301	4406.34375	1777.508583	8467.085938	30871.17188	6432.0
15	1.14	6795.753906	24264.64063	745.7470703	465.0424805	990.1313477	1794.898438	119.0358276	4313.539063	1480.68652	6845.367188	32930.875	8029.3
16	1.15	1836.59375	6009.238281	142.5959473	37.6514282	491.6550293	1753.573242	941.4536131	3346.365234	20822.03125	2188.4		
17	1.16	1699.289063	6458.503906	96.57451771	200.644165	310.729248	774.0219727	183.9802248	1690.515625	904.3242188	3143.683594	14266.88281	1877.8
18	1.17	1497.520508	5609.605496	77.26855469	169.138916	226.6845703	567.8823242	79.0146484	1362.202148	830.5599609	2024.548828	913.3125	1730.2
19	1.18	9033.570313	21372.0625	1250.067383	1564.307617	892.3813477	1589.58867	252.1553953	4251.226563	1510.931641	6424.839844	38623.5625	7712.2

*sampeldata* should have sample information matching *featuredata* (samples in rows).

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1		batch	gender	Age	bmi											
2	s_1	Batch 2	code_1	58.7	22.2											
3	s_2	Batch 2	code_0	76.7	23.7											
4	s_3	Batch 2	code_0	56.2	28.2											
5	s_4	Batch 2	code_0	77.2	26.2											
6	s_5	Batch 2	code_0	74.3	26.4											
7	s_6	Batch 2	code_1	66.8	26.4											
8	s_7	Batch 1	code_0	65	29											
9	s_8	Batch 2	code_1	66.5	26.1											
10	s_9	Batch 2	code_0	70.2	27.3											
11	s_10	Batch 1	code_1	55	25											
12	s_11	Batch 1	code_0	51	28											
13	s_12	Batch 1	code_1	55	29											
14	s_13	Batch 2	code_0	79.2	30.7											
15	s_14	Batch 2	code_1	52.7	25.4											
16	s_15	Batch 1	code_0	42	32											
			featuredata	sampleddata	metabolitedata											

Optional *metabolitedata* should have metabolite information matching featuredata with metabolite names in rows.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	names	IS	neg_controls	pos_controls	gender										
2	m_1	0	1	0											
3	m_2	0	1	0											
4	m_3	0	1	0											
5	m_4	0	0	0											
6	m_5	0	0	0											
7	m_6	0	0	0											
8	m_7	0	0	0											
9	m_8	0	1	0											
10	m_9	0	1	0											
11	m_10	0	1	0											
12	m_11	0	0	0											
13	m_12	0	0	0											
14	m_13	1	1	0											
15	m_14	0	0	0											
			featuredata	sampleddata	metabolitedata	settings									

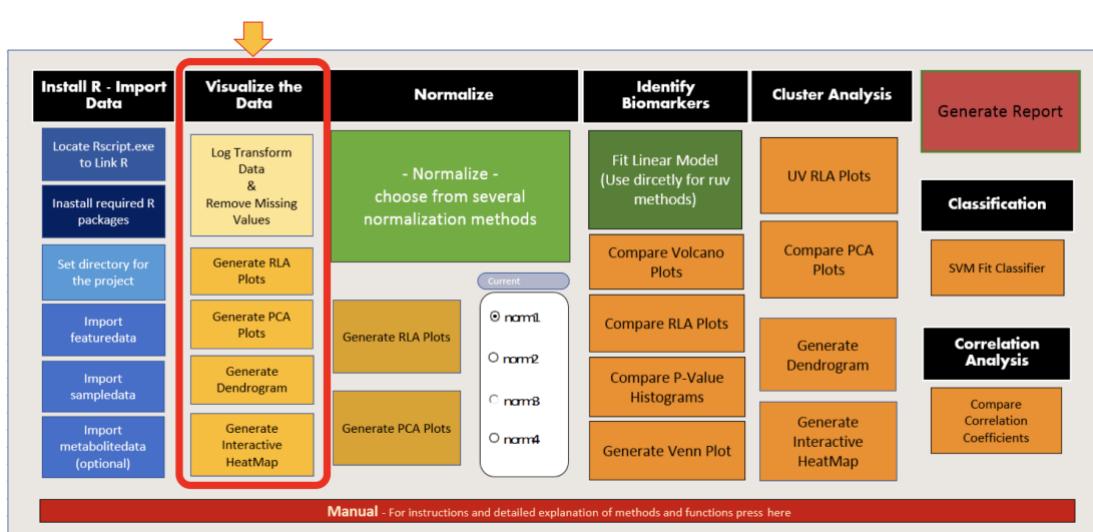
*metabolitedata* can include any metabolite information such as grouping structures, internal standard metabolites, negative control and positive control metabolites.

After the data is loaded, you are ready to proceed to analyse the data!

## NormalizeMets Workflow:

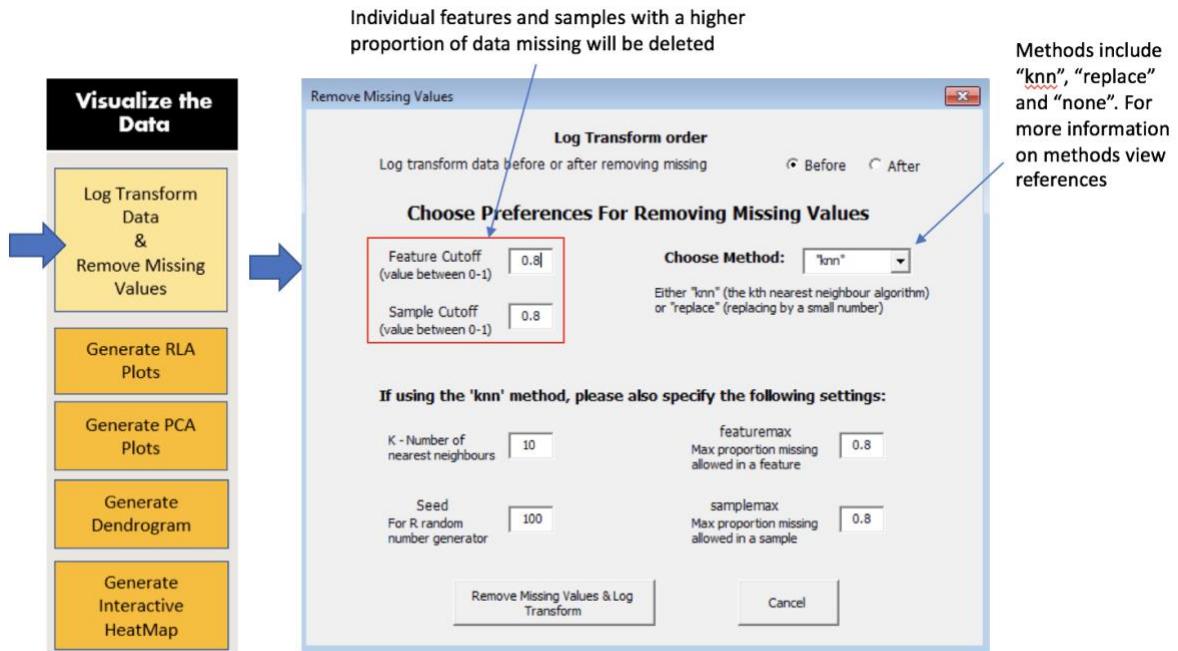
### Visualize the Data

The following section refers to the visualize part:



## Log Transforming the data and removing missing value (mandatory):

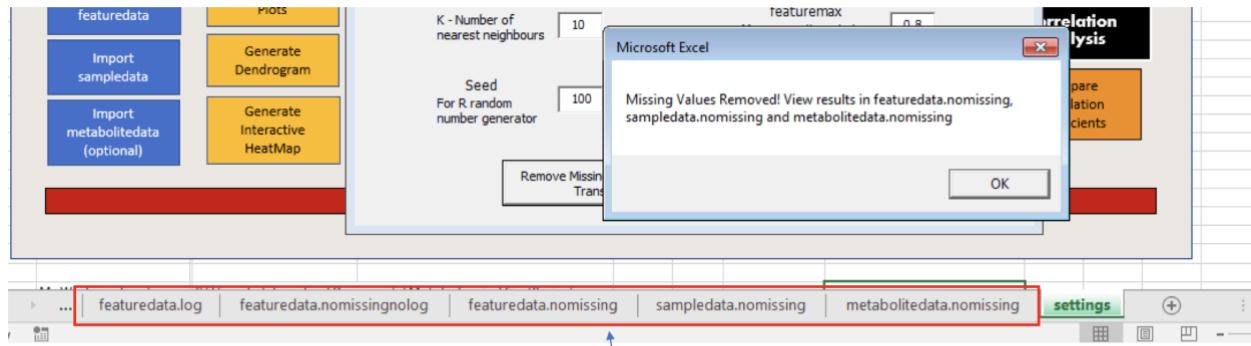
A frequent issue in metabolomics data sets is the occurrence of missing values. It is important to reduce the number of missing values as much as possible by using an effective pre-processing procedure. For example, a secondary peak picking method can be used for LC-MS data to fill in missing peaks which are not detected and aligned.



“knn” – use k nearest neighbours method to replace missing values.

“replace” – replaces missing values by half the minimum value in featuredata.

Clicking ‘Remove Missing Values & Log Transforming’ the following appears:



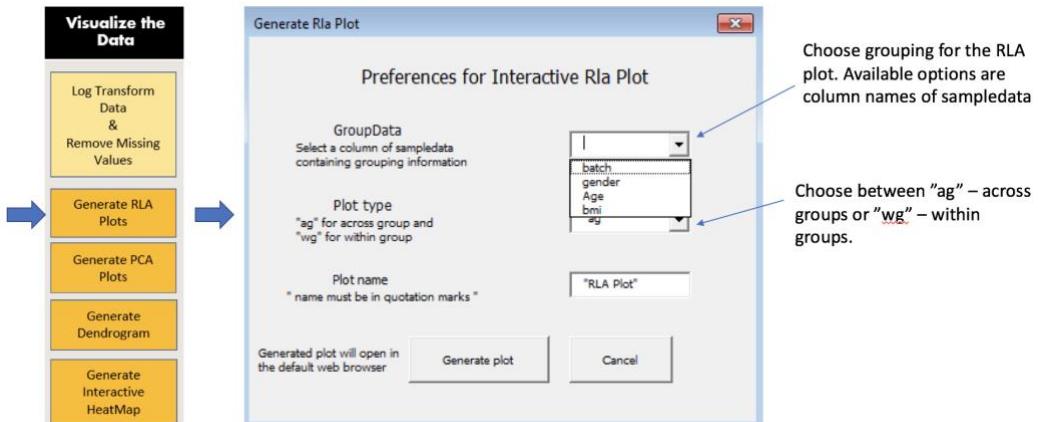
Note the new sheets that appeared, they have respectively the regular log transformed *featuredata*, the log transformed *featuredata* with missing values removed, *featuredata* with missing values removed and without log transformation, *sampledata* with rows removed corresponding to *featuredata.nomissing*, *metabolitedata* with rows removed corresponding to *featuredata.nomissing*.

Unless you are interested to view or copy any of this data, those sheets are only going to be used for further internal functions.

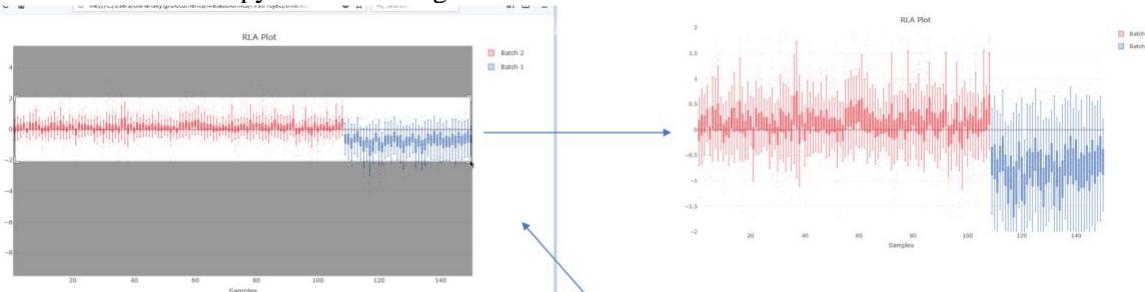
Now the plots in *Visualize the Data* can be generated! The data generated is also going to be used for the Normalization section.

## RLA plots

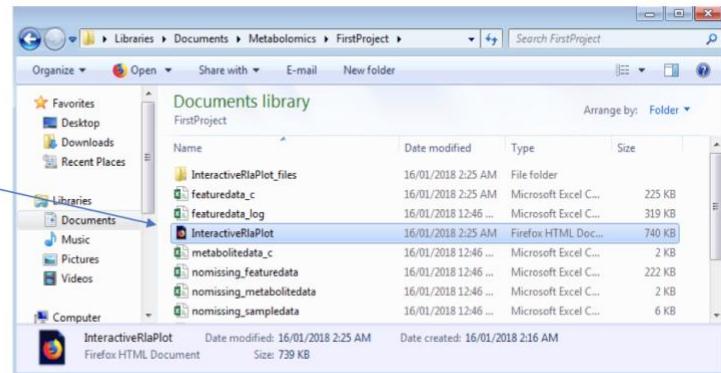
One way of visualising the log transformed metabolomics data is the use of *across group* or *within group* relative log abundance (RLA) plots (De Livera et al. 2012 De Livera et al. (2015)).



Setting *groupdata* to *batch* and selecting *Generate plot* opens the interactive plot in the default web browser and saves a copy in the working folder.

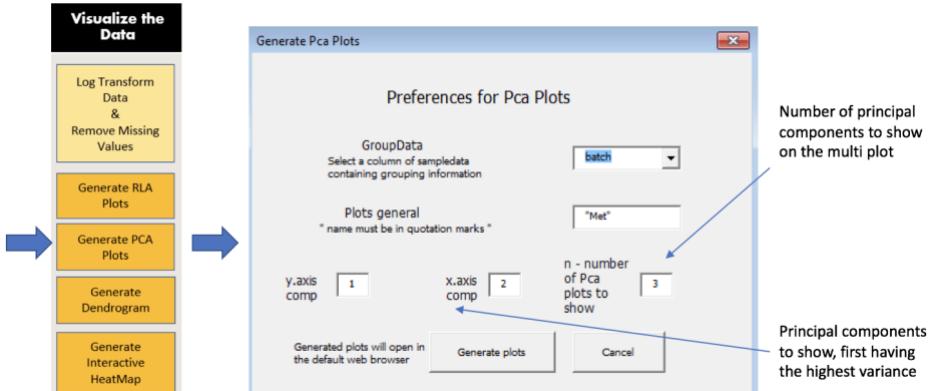


Copy of all plots together with all other generated files are stored in the working directory



## PCA plots

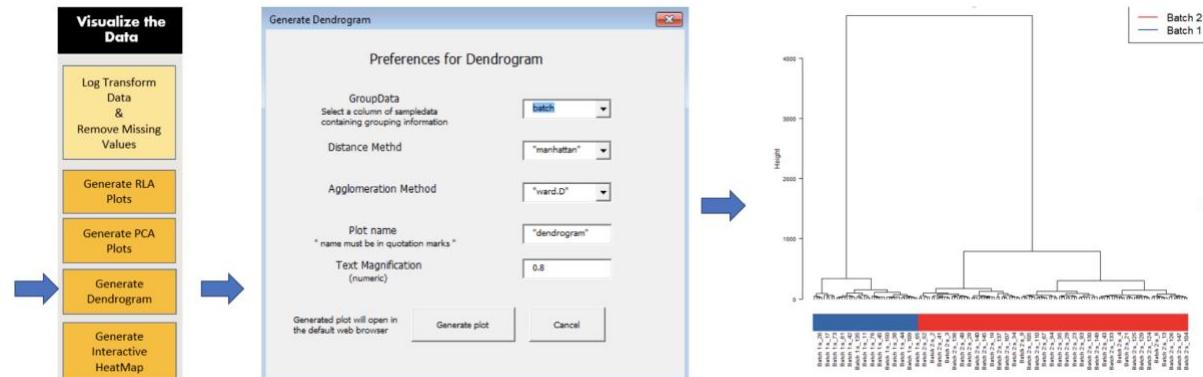
The following function can be used to obtain multiple plots for exploration of the principal components of the *featuredata* matrix: a bar plot indicating the variance explained by each principal component, scores and loading plots with specified axes (interactive and non-interactive), and a pairs plot of the first  $n$  principal components. These plots are useful in identifying any outlying samples and getting a preliminary understanding of the structure of the data.



All produced plots are stored in the working directory, with interactive plots opened in the browser and static plots located in the new *plots* sheet.

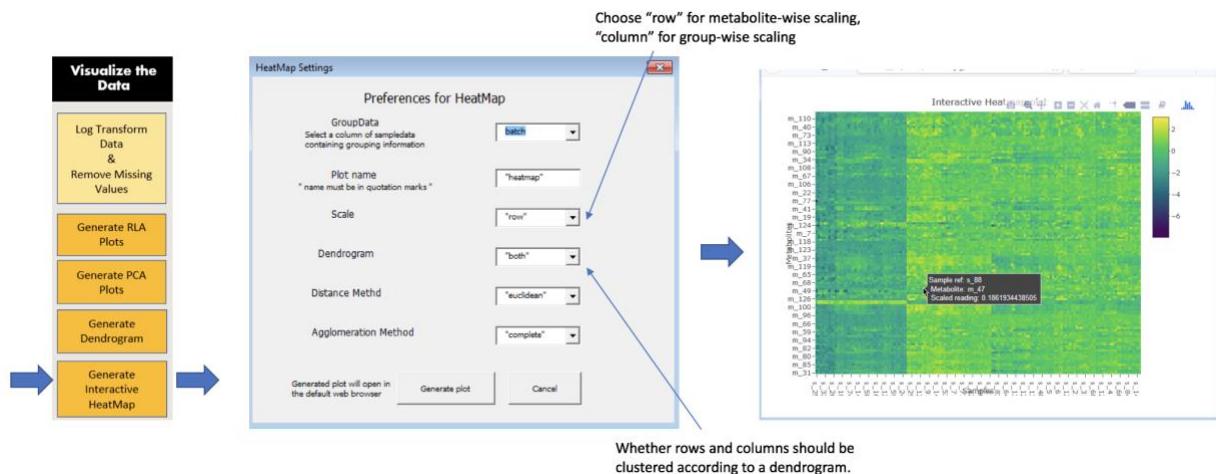
## Dendrogram

Generates a dendrogram to visualise clustering structures in the data, many different methods are available.



## HeatMap

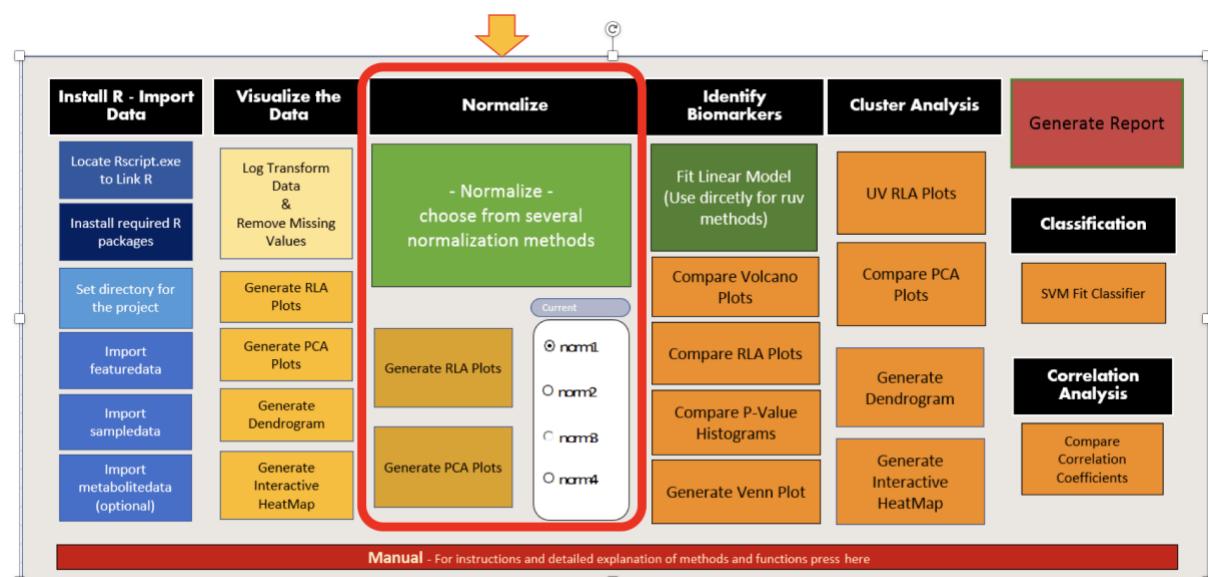
The HeatMap produced can reveal interesting structures in the data.



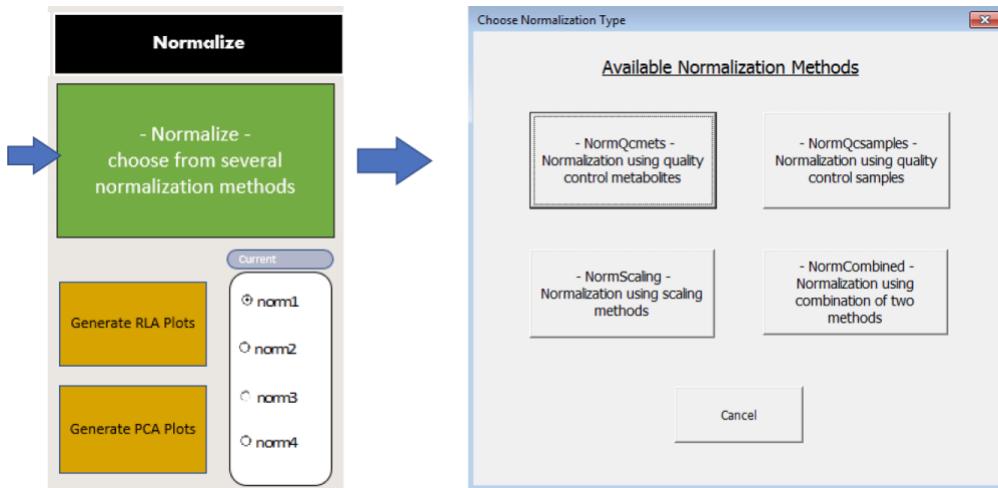
## Normalization

Normalization methods presented in this package are divided into four categories, as those which use (i) internal, external standards and other quality control metabolites (*NormQcmets*) (Sysi-Aho et al. 2007, Redestig et al. (2009), De Livera et al. (2012), De Livera et al. (2015), Gullberg et al. (2004)) (ii) quality control samples (*NormQcsamples*) (Dunn et al. 2011), (iii) scaling methods (*NormScaling*) (Scholz et al. 2004, Wang et al. (2003)), and (iv) combined methods (*NormCombined*) (Kirwan and Broadhurst (2013)).

The normalization methods are accessible in the following section:



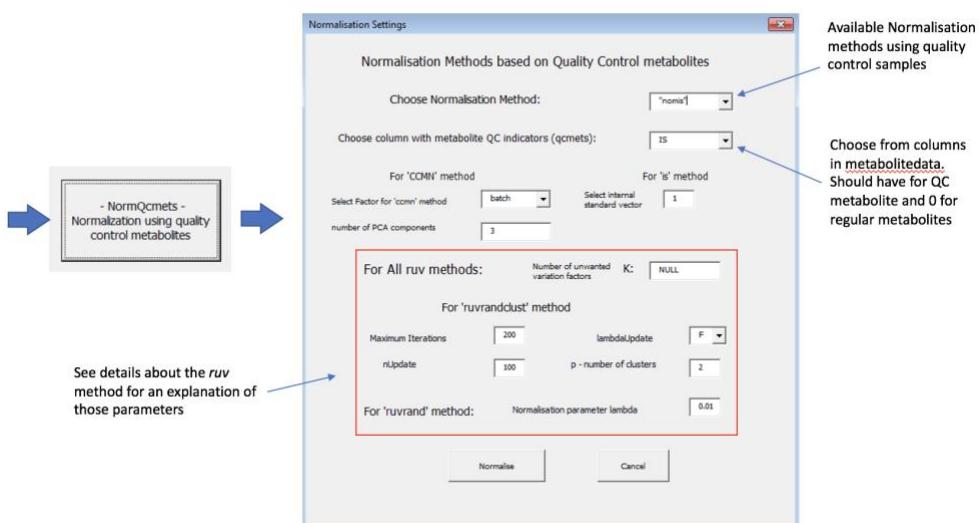
Clicking on the Normalize button opens the following menu enabling the choice of different normalization methods.



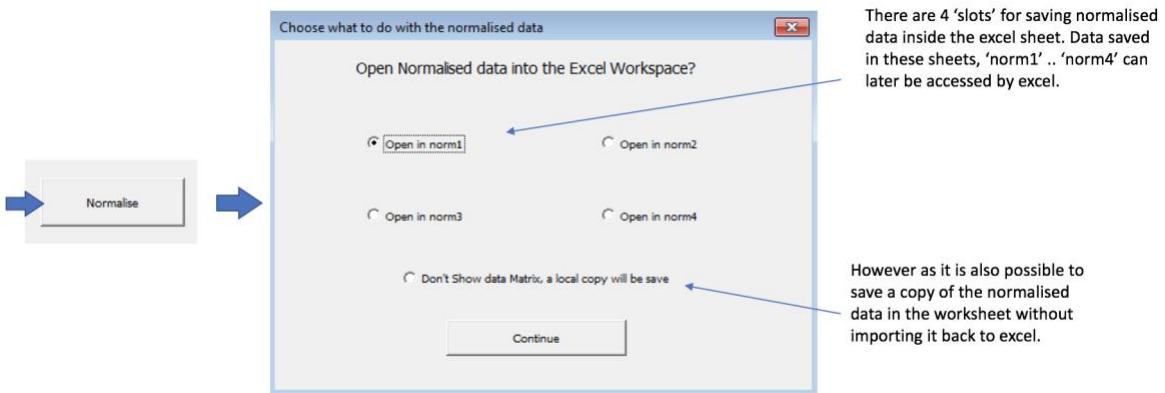
## NormQcmets

The approaches in *NormQcmets* use internal, external standards and other quality control metabolites. These include the *is* method which uses a single standard (Gullberg et al. 2004), the *ccmn* (cross contribution compensating multiple internal standard) method (Redestig et al. 2009), the *nomis* (normalization using optimal selection of multiple internal standards) method (Sysi-Aho et al. 2007), and the remove unwanted variation methods (J. A. Gagnon-Bartsch, Jacob, and Speed 2014) as applied to metabolomics using “ruv2” (De Livera et al. 2012), “ruvrnd” and “ruvrndclust” (De Livera et al. 2015). Note that *ruv2* is an application specific method designed for identifying biomarkers using a linear model that adjusts for the unwanted variation component.

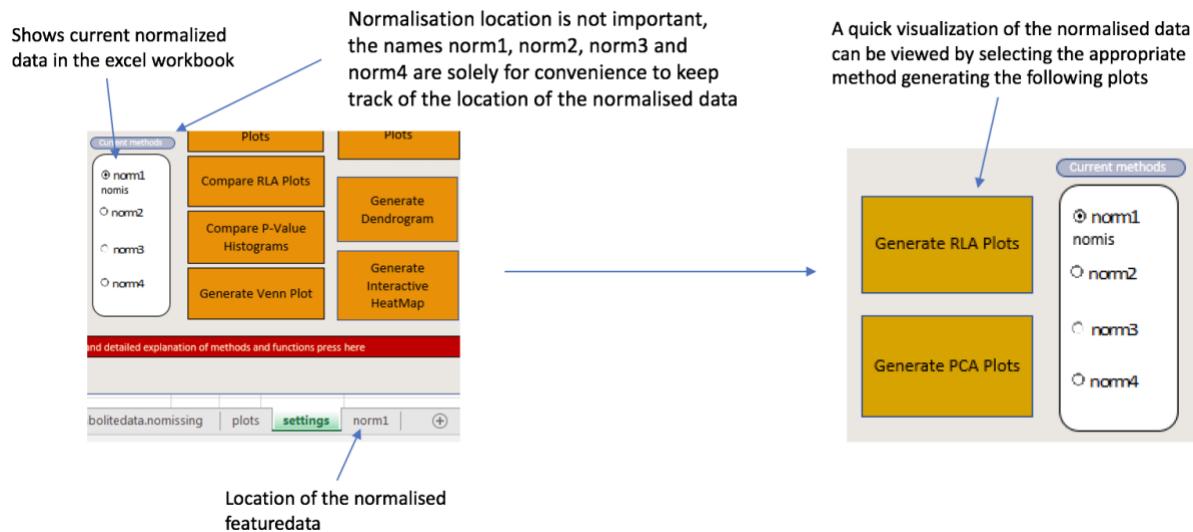
To Normalize:



After Clicking the Normalise button the screen asking you where the normalized data is to be saved appears.



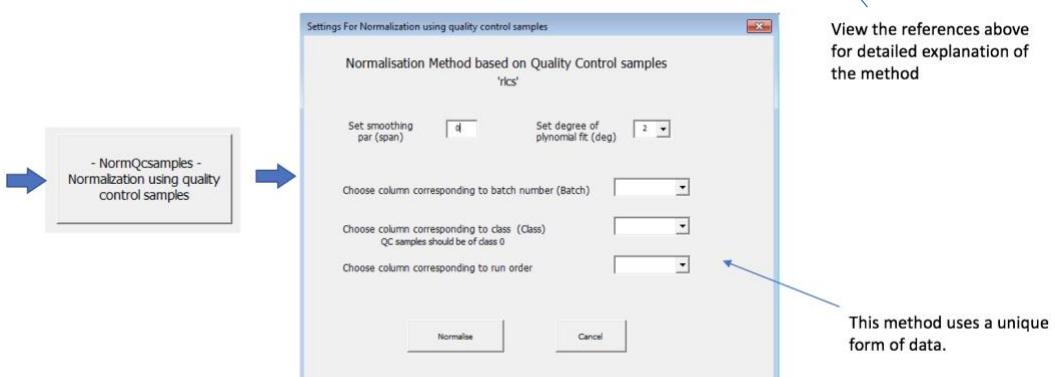
Upon clicking continue, you will return back to the settings sheet but you can notice some changes:



## NormQcsamples

This function is based on the quality control sample based robust LOESS (locally estimated scatterplot smoothing) signal correction (QC-RLSC) method as described by Dunn et al. (2011) and implemented statTarget (Luan 2017)

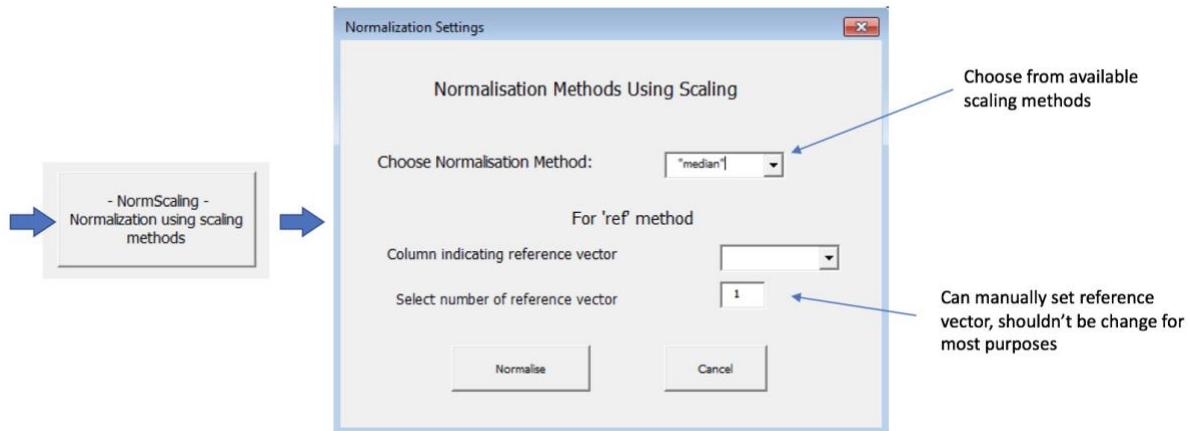
To Normalize:



## NormScaling

The scaling normalization methods (Scholz et al. 2004, Wang et al. (2003)) included in the package are normalization to a total sum, normalisation by the median or mean of each sample, and are denoted by *sum*, *median*, and *mean* respectively. The method *ref* normalises the metabolite abundances to a specific reference vector such as the sample weight or volume.

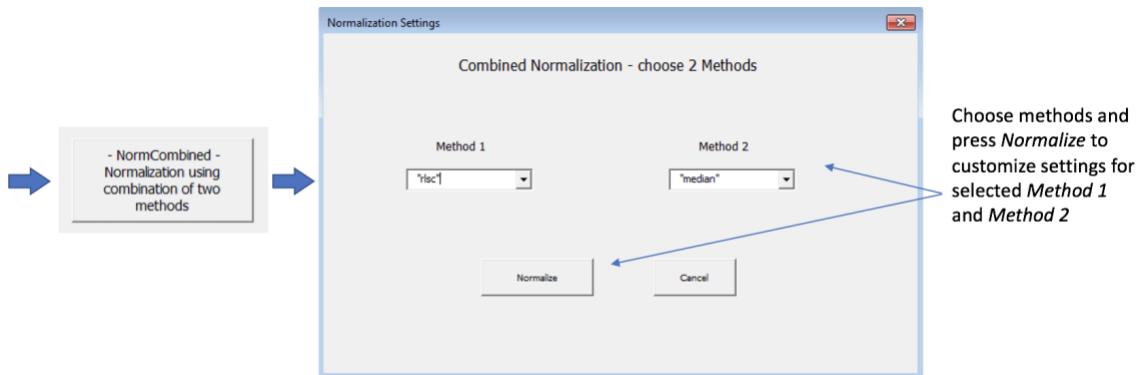
To Normalize:



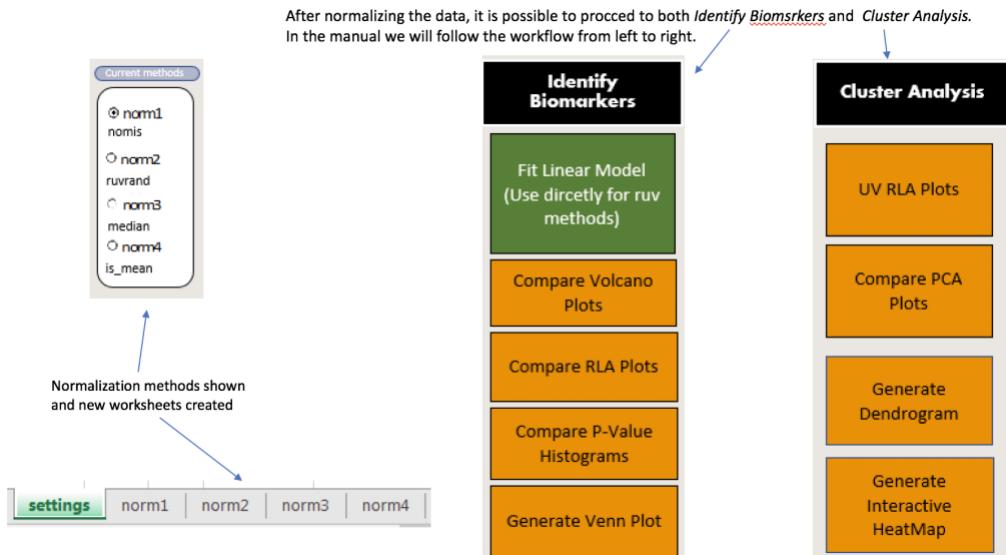
## NormCombined

In some circumstances, researchers use a combination of the above normalizations (i.e., one method followed by another). This can be achieved using the *NormCombined* function. The function defaults to employing 'rlsc' approach followed by the 'median'.

To Normalize:



Note that normalizing the data is not necessary to proceed to fitting a linear model although it is highly recommended to try a few normalization methods when analysis data.

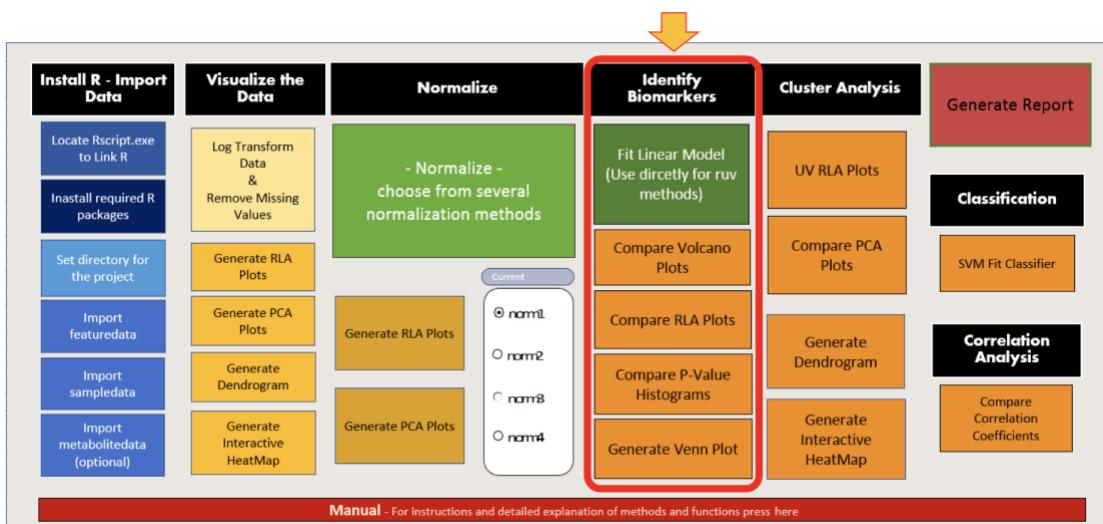


## Assessing and choosing normalization methods

The criteria for assessing and choosing a normalization method implemented NormlizeMets have been described in detail by De Livera et al. (2012), De Livera et al. (2015) and J. A. Gagnon-Bartsch, Jacob, and Speed (2014).

## Identifying Biomarkers

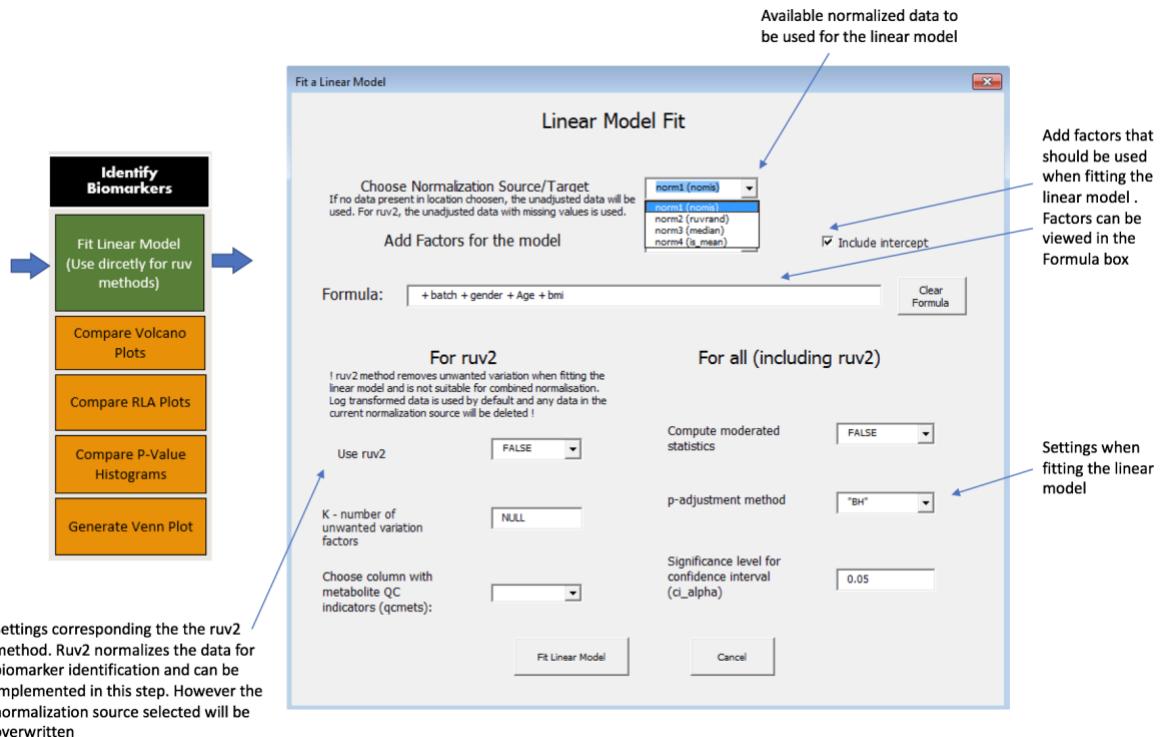
To view and compare the biomarkers identified, first a linear model has to be fitted to the data.



## Fit Linear Model

A linear model has to be fitted for every Normalization method that is to be used down the line for Biomarker identification. Setting from one ‘run’ of the Fit Linear model will be saved for the next.

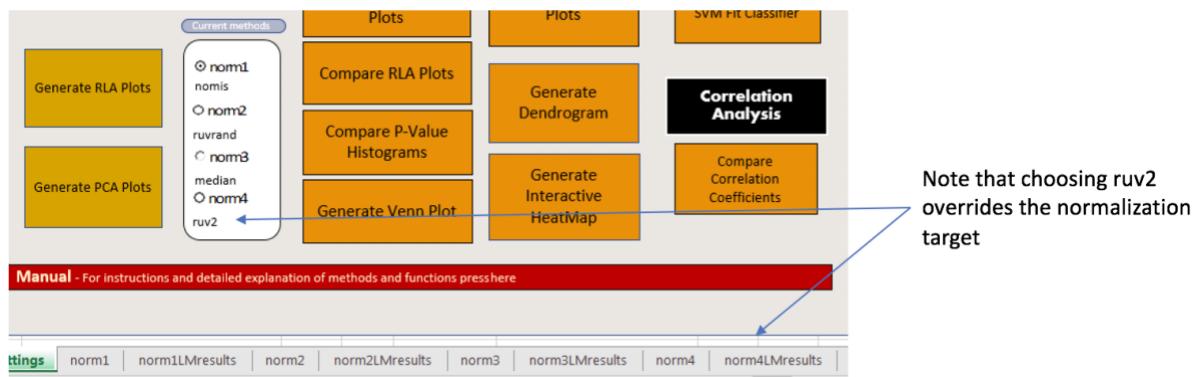
To Fit Linear Model:



Settings corresponding to the rvu2 method. Rvu2 normalizes the data for biomarker identification and can be implemented in this step. However the normalization source selected will be overwritten

	A	B	C	D	E	F	G	H	I	J	K	L	M	
	F stat	P value	Adjusted F p value	cooff (intercept)	cooff batchBatch 2	cooff genderfemale_1	cooff Age	cooff bmi	t stat (intercept)	t stat batchbatch2	t stat genderfemale_1	t stat Age	t stat bmi	
1	m_1	367.0640508	1.83e-40	8.956e-01	8.053e-137	0.0000000000000000	0.0000000000000000	0.0000000000000000	42.8505000	0.5900000000000000	0.5900000000000000	0.0000000000000000	0.0000000000000000	
2	m_2	367.0640508	1.83e-40	8.956e-01	8.053e-137	0.0000000000000000	0.0000000000000000	0.0000000000000000	42.8505000	0.5900000000000000	0.5900000000000000	0.0000000000000000	0.0000000000000000	
3	m_3	367.0640508	1.83e-40	8.956e-01	8.053e-137	0.0000000000000000	0.0000000000000000	0.0000000000000000	42.8505000	0.5900000000000000	0.5900000000000000	0.0000000000000000	0.0000000000000000	
4	m_4	39.81815528	1.29e-25	1.67e-25	6.793188000	0.020513445	0.300187081	-0.01077832	13.55311266	0.181800463	3.705513605	-0.672314342		
5	m_5	14.54096231	1.53e-11	1.76e-11	4.98533120	0.00452139	0.287969155	0.01046794	7.735175731	0.39048884	2.724893926	1.77603426		
6	m_6	74.81961081	2.05e-58	2.32e-58	6.6677930	0.124584715	0.065071851	0.010069860	18.9129260	1.321700023	1.321700023	-1.1479006	-3.33111804	
7	m_7	15.47639232	1.23e-25	2.53e-25	2.53e-25	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.330397000	0.330397000	0.330397000	0.330397000	0.330397000	
8	m_8	1462.397870	4.02e-122	2.47e-120	8.10725599	0.0488054	-0.06572878	0.0212321	8.650757264	0.426092827	-0.705953192	0.454894642		
9	m_9	1226.58959	1.05e-16	3.23e-15	7.21770872	0.009903118	-0.004356857	-0.001835148	85.50233518	0.49416064	-0.78447664	0.18047732		
10	m_10	1226.58959	1.05e-16	3.23e-15	7.21770872	0.009903118	-0.004356857	-0.001835148	78.30533970	0.48008164	-0.2919364	-0.107033345		
11	m_11	369.103156	1.93e-75	8.79e-75	10.000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	42.8505000	0.5900000000000000	0.5900000000000000	0.0000000000000000	0.0000000000000000	
12	m_12	369.103156	1.93e-75	8.79e-75	10.000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	39.15273904	1.25269134	1.843111314	-0.467172766		
13	m_13	89.17162338	1.84e-42	3.23e-42	7.64516519	0.039796591	0.080723464	0.05668386	0.02113052	20.89799548	0.848660154	1.35572297	1.74875421	
14	m_14	88.20131195	2.13e-42	3.68e-42	8.13817720	0.035334734	0.126133374	0.000418627	-0.02077858	20.9552018	0.670748264	2.096502074	-1.320504835	
15	m_15	45.54797628	4.31e-27	1.37e-26	5.251160527	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000
16	m_16	45.54797628	4.31e-28	5.96e-28	5.251160527	0.09770374	0.040021189	0.000191219	0.007444165	14.05952391	1.17053969	0.714540203	1.004895076	
17	m_17	149.098252	2.60e-55	5.81e-55	7.47379830	0.022929771	0.174416195	0.000191219	0.007444165	27.01482123	0.472646389	8.317067487	-0.047635338	
18	m_18	32.90009500	2.29e-22	3.85e-22	4.95504007	0.008160398	0.044135203	0.002846402	0.03401139	12.64600705	0.91341199	0.493808483	-0.80550207	
19	m_19	32.90009500	2.29e-22	3.85e-22	4.95504007	0.008160398	0.044135203	0.002846402	0.03401139	12.64600705	0.91341199	0.493808483	-0.80550207	
20	m_20	18.10842903	4.09e-41	4.79e-41	5.9933842	-0.057434947	0.114407096	0.004116869	-0.020884242	20.38594503	-0.899593121	2.385459558	1.693119137	
21	m_21	13.12848513	1.51e-10	7.71e-10	2.81292127	0.174571041	0.051311434	0.02053416	0.05851318	6.11175200	1.71110084	-0.299121519	0.591837814	
22	m_22	261.150608	1.34e-70	4.64e-70	7.69309715	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	1.139411818	1.139411818	
23	m_24	190.9438527	1.81e-66	4.8933097	-0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	
24	m_24	190.9438527	6.23e-42	1.59e-41	6.416501638	0.051411949	0.20307016	0.019395348	29.7046369	1.064710241	7.445599934	2.41040648		
25	m_25	367.24408011	2.00e-40	1.02e-39	9.71953441	-0.009912387	0.104940607	-0.015994997	-0.040939575	42.44545436	-0.195044779	2.810682881	-0.752922428	
26	m_26	192.1110341	4.24e-42	1.11e-41	8.78784813	0.030552211	0.190005358	0.000620403	-0.007441415	30.2026805	-0.553124672	4.091900082	-0.240494302	
27	m_27	26.18120597	2.05e-29	5.83e-29	1.22e-28	0.0000000000000000	0.0000000000000000	0.0000000000000000	0.0000000000000000	38.16649315	1.874042328	4.433512088	-0.133511176	
28	m_28	65.18120598	1.13e-33	1.66e-33	8.949488788	0.296430414	-0.238164862	-0.00053037	0.078147492	18.16649313	2.269644663	-2.4671454	-0.133511176	

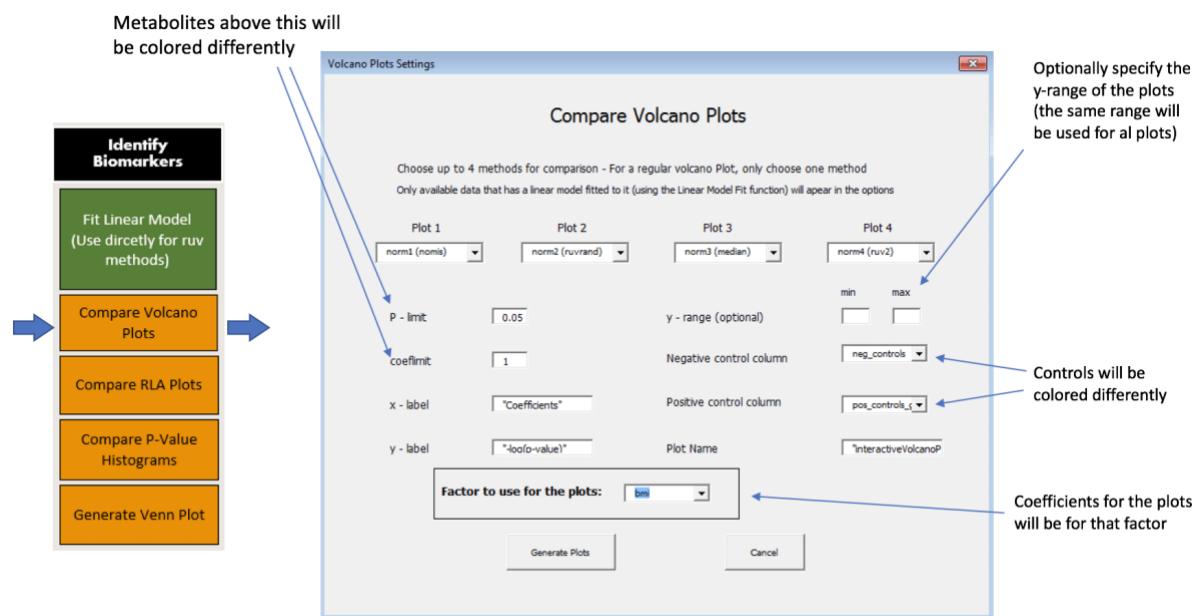
A new sheet is created, it stores the linear model output. This data should not be altered manually as other functions will need to access it.



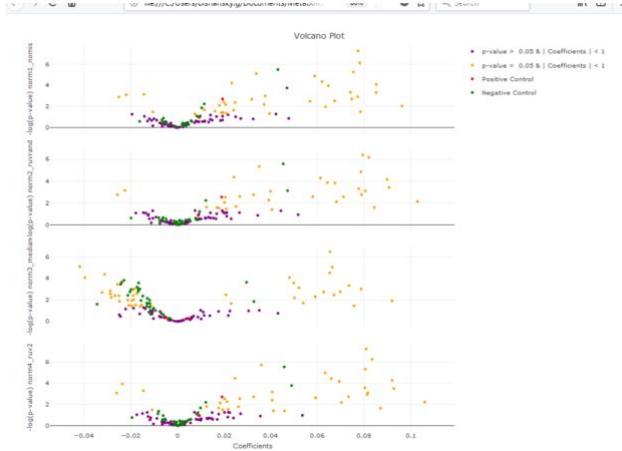
## Volcano Plots

Volcano plots are useful in identifying biomarkers and generally assessing the normalization.

Compare Volcano Plots:



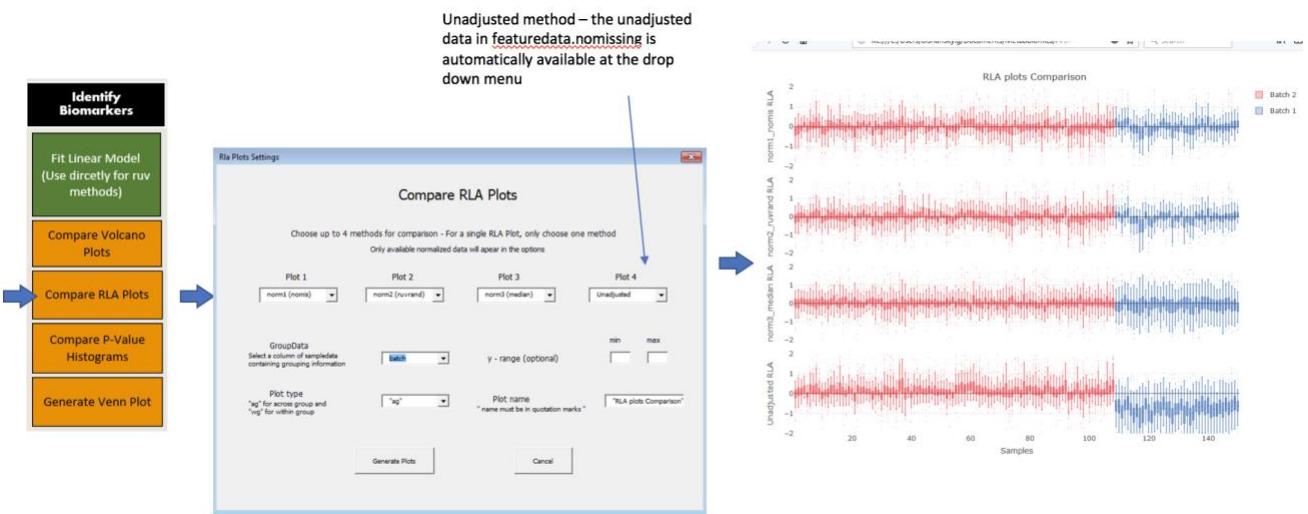
The plot is saved in the working directory and opens in the default browser.



## Compare RLA plots

Used to assess normalization by comparing relative log abundance plots, similar input to the *Generate RLA* plots function

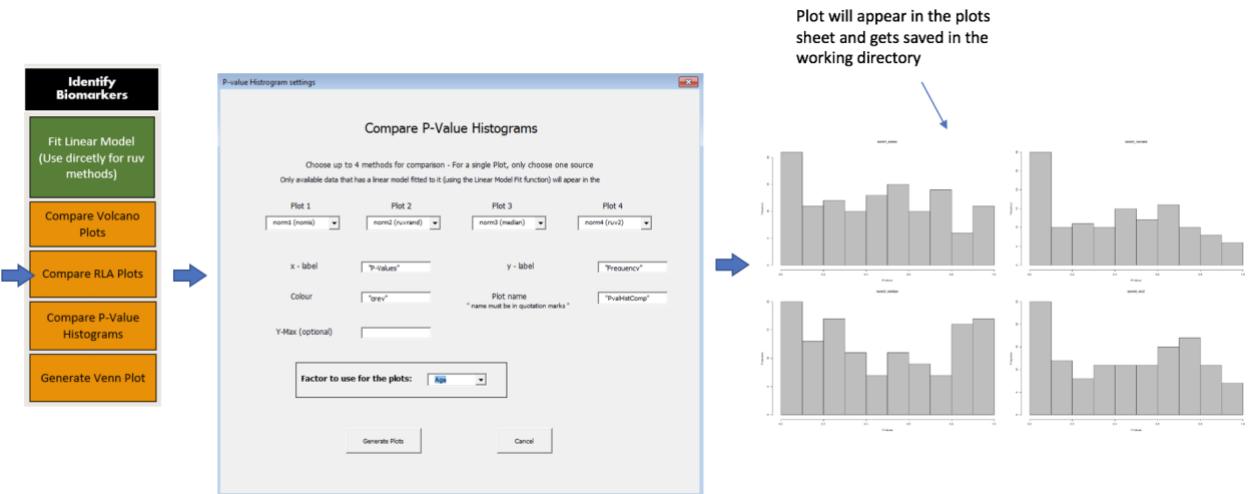
Compare Rla Plots:



## Compare P-Value Histograms

Compare histograms of the coefficient's p-values. The distribution of the p-values should be used to assess the success of the normalization.

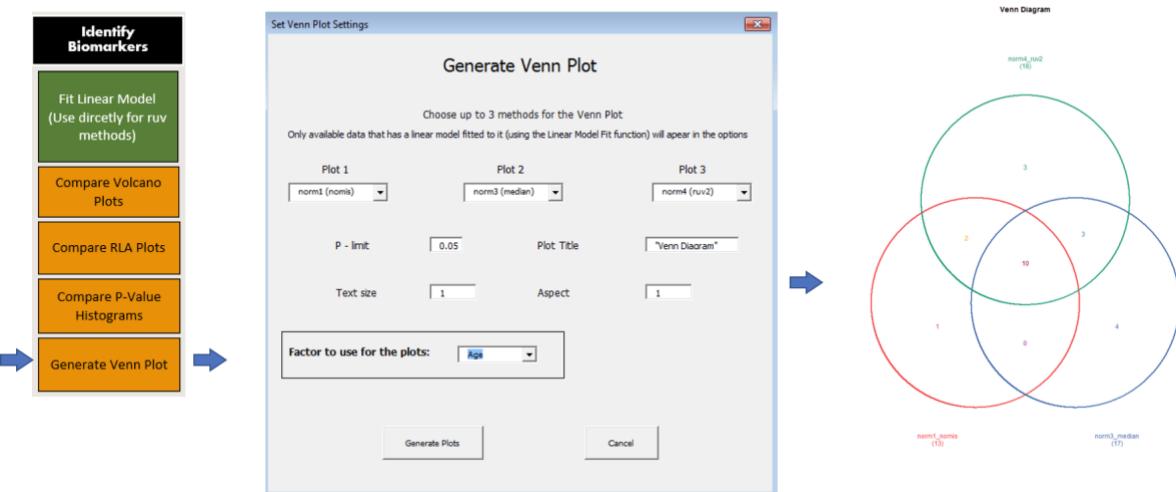
Compare P-Value histograms:



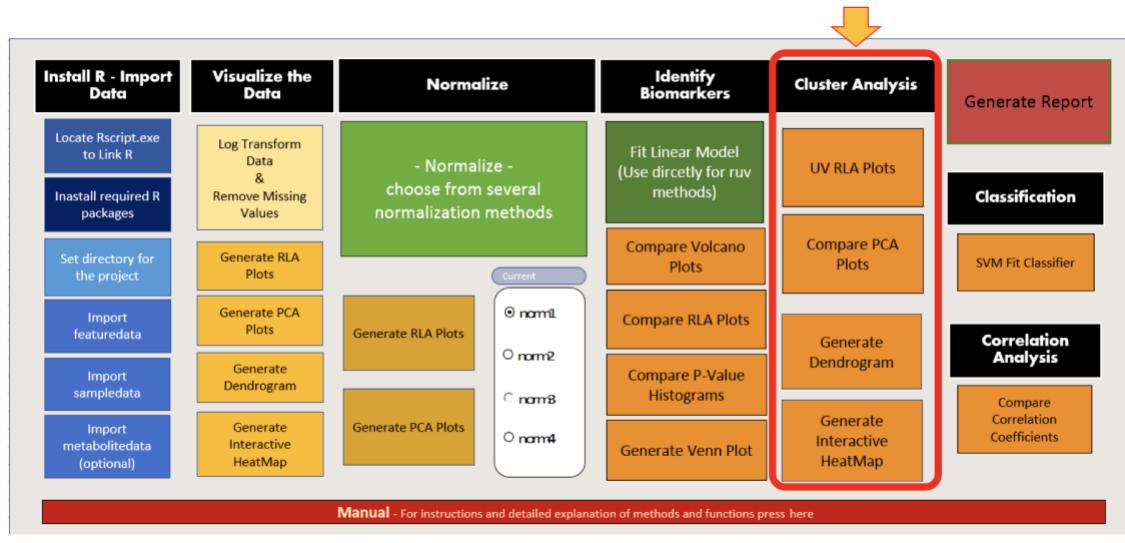
## Generate Venn Plot

Generates a Venn plot that compares the biomarkers identified by the different normalization methods.

Generate Venn Plot:

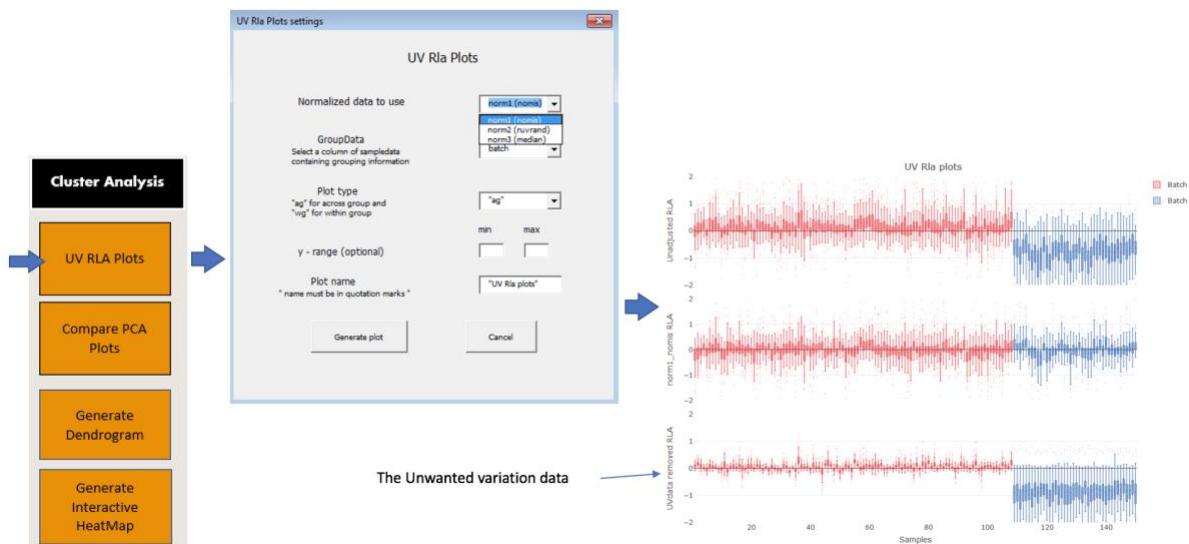


## Cluster Analysis



## UV RLA plots

Unwanted Variation relative log abundance plots enable visualisation of the unwanted variation removed by each normalization method.



## Compare PCA Plots

Compare principal component multi-plots for differed normalization methods.



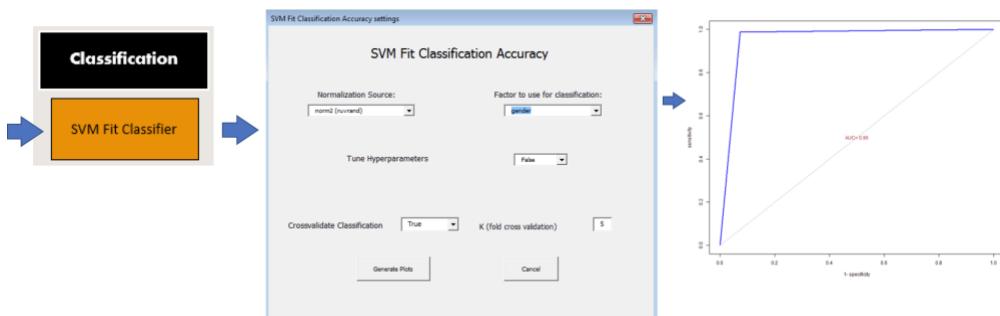
**Generate Dendrogram** and **Generate Interactive HeatMap** are identical to those discussed in the *Visualize the Data* section. The user has to choose the normalized data to be used.

## Classification

Classification accuracy is a good way to assess the success of a given normalization method

### SVM-Fit

The Support Vector Machines method is used to classify the data, the classification accuracy is then assessed based on the factor specified.



## Correlation Analysis

It is important to look at correlation coefficients when normalizing data.

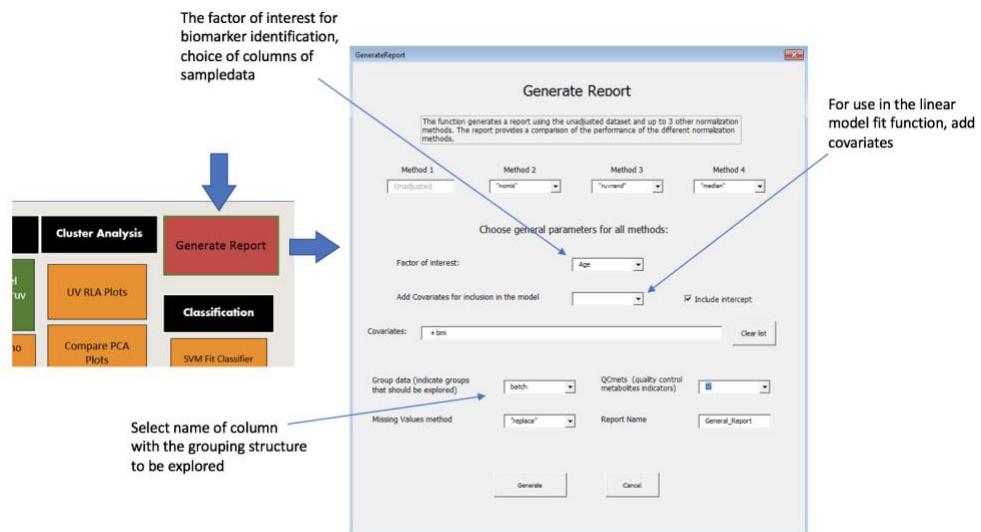
### Compare Correlation coefficients



## Generate Report

The *Generate Report* function generates an interactive report based on basic user input. There is a choice of up to 3 normalization methods to be included together with the unadjusted data. The report includes various plots and diagnostic to assess the normalization. Guidance on interpretation of the various plots, together with notes of what the user should look for when assessing the results is provided in the generated document.

### Generate Report:



An example report is available in your downloaded ExNormalizeMets folder.

### Full Package Vignette:

For package vignette with detailed explanations of methods and workflow, follow the link to:  
[https://github.com/metabolomicstats/NormalizeMets/blob/master/NormalizeMets\\_vignette.html](https://github.com/metabolomicstats/NormalizeMets/blob/master/NormalizeMets_vignette.html)