This tutorial contains navigation buttons that enable you to move throughout the tutorial.

Please use the navigation buttons and not the page up/page down or arrow keys to navigate through the tutorials.

This is the 'Next' button. It takes you to the next frame or stop point.

This is the 'Previous' button. It takes you to the previous frame or stop point.

This is the 'Go to frame' button. It takes you to a specified frame.

This is the 'Go to URL' button. It takes you to a website link.

Press the 'Next' button below to start this tutorial.
## Data Selection Panel

<table>
<thead>
<tr>
<th>Name of Image Matrix</th>
<th>Name of Variable Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>imagedata_dan01</td>
<td>exactmass_dan01</td>
</tr>
</tbody>
</table>

This tutorial will cover how to run PCA on an image data set.

If desired the data needs to be normalized before running PCA.
From the 'MVA' menu choose:

MVA -> PCA
Data Selection Panel

Name of Image Matrix: imagedata_DAN...
Name of Variable Matrix: exactmass_DAN01

Load Selected Data

Image: None
Variables: None

Data Preprocessing
Choose an option below

Run PCA

This opens the Image PCA panel.

PC Summary
PC#  %Var  %Vartotal

Name for scores matrix

Name for loadings matrix

Name for Variance matrix

# of PCs to keep in model

Save PCA Data

Close Panel
Press the 'Load Selected Data' button to load the data selected in the drop down menus above into the image PCA panel.
Choose a preprocessing option from the 'Data Preprocessing' drop down menu.

These option are included to allow the user to explore different data preprocessing methods. Their inclusion herein does not denote a recommendation or suggestion on how to process SIMS image data.

The options include:
None - do not carry out any preprocessing

Autoscale - autoscale the data. This mean centers the data and then divides by the variance of each variable. Autoscaling is not recommended for image data.

Mean Center - Mean centers the data by subtracting the variable means from each variable.

Squareroot mean center - takes the squareroot of the data and then mean centers the data.

Poisson scaling - scales the data according to the method developed by Keenan and Kotula (Appl. Surf. Sci. 231-232 (2004) 240-244)

Mass scale - scales each peak area by multiplying by the mass of the peak.

Mass^2 scale - scales each peak area by the mass of the peak squared.

Many of the above methods also include an option to mean center after doing the respective scaling or transformation.
Here we will choose 'Poisson Scaling & Mean Center'
Data Selection Panel

Name of Image Matrix  Name of Variable Matrix
imagedata_DAN01   exactmass_DAN01

Load Selected Data
Image: imagedata_DAN01 Variables: exactmass_DAN01

Data Preprocessing
Choose an option below

Run PCA
Press the 'Run PCA' button to carry out PCA on the selected data.

PCA Summary
PC#  %Var  %Vartotal

Name for scores matrix
Name for loadings matrix
Name for Variance matrix

# of PCs to keep in model

Save PCA Data

Close Panel
A summary of the variance captured per PC is shown in the table. New boxes appear that enable looking at the scores and loadings.
To see a score and loading plot choose the desired PC number from the drop down menu.
And press the 'Plot Scores and Loads' button.
The scores and loadings for the selected variable are shown in the plots on the right.
This button brings up windows showing the positive and negative scores and loadings plotted separately.

This can be useful to more clearly see what areas within the image correspond with the positive and negative scores.
Here is an example of what the positive and negative scores and loadings look like.

In these plots the negative scores and loadings have been multiplied by -1 in order to display them on a typical color scale starting at zero.

This transformation is only done for the data display and does not affect the original scores or loadings.

Such transformations are okay as long as you always multiply the scores and loadings by the same value.
Data Selection Panel

Name of Image Matrix: imagedata_DAN01
Name of Variable Matrix: exactmass_DAN01

Load Selected Data

Image: imagedata_DAN01
Variables: exactmass_DAN01

Data Preprocessing
Poisson Scaling & Mean Center

Run PCA

PC Summary
PC# %Var %Vartotal
1 35.8 35.8
2 6.1 41.9
3 4.6 46.5
4 3.8 50.3
5 3.7 54
6 3.6 57.6
7 3.5 61.1
8 3.5 64.6
9 3.4 68
10 3.4 71.4

Name for scores matrix
Name for loadings matrix
Name for Variance matrix

# of PCs to keep in model

Save PCA Data

One can also browse through all available PCs by pressing the 'Browse all PCA plots' button.
This brings up an additional window. By left clicking anywhere in the window it will automatically cycle through all available PC scores and loading plots.
Each mouse click updates the window with the next PC.
You can close this external window at any time.
You can also plot the scores in a grid showing multiple score images. First enter the number of PCs you want displayed in the grid.
Then press the 'Plot Multiple PCs in Grid' button. A figure will open up in a separate window showing the requested score images.
Here you can see the grid plot of the 4 requested PCs.

The number of PCs selected will affect the initial size of the images within the grid. The grid plot window can be resized in order to make the images larger.
To save the PCA model data, enter appropriate names in each of the fields to the left, specify how many PCs to keep in the model, and press the 'Save PCA Data' button.

You must save the PCA model data in order to use the functions to plot the scores and loadings in the 'Data Display' menu.
The panel can be closed by pressing the 'Close Panel' button.
### Data Selection Panel

<table>
<thead>
<tr>
<th>Name of Image Matrix</th>
<th>Name of Variable Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>imagedata_dan01</td>
<td>exactmass_dan01</td>
</tr>
</tbody>
</table>

That ends this tutorial. Press the button on the left to go back to the previous step. Press the button on the right to start the tutorial over.