Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

- Name of Data Matrix: Select Data
- Name of Variable Matrix: Select Variables
- Name of Filename Matrix: Select Filenames
- Name of Totalcounts Matrix: Select Totalcounts
- Name of Samplenames Matrix: Select Samples

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

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

<table>
<thead>
<tr>
<th>Name of Data Matrix</th>
<th>Name of Variable Matrix</th>
<th>Name of Filename Matrix</th>
<th>Name of Totalcounts Matrix</th>
<th>Name of Samplenames Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>ndatat</td>
<td>exactmass</td>
<td>filenames</td>
<td>totalcounts</td>
<td>samplenames</td>
</tr>
</tbody>
</table>

This tutorial will cover how to run PCA on your data, create Scores and Loadings plots, Save the plots, and Export the Scores and Loadings data.
To run PCA on the data:

From the 'MVA' menu choose -> 'Run PCA'.
Data Selection Panel
These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop-down menus to select the data and information you want to use in your analysis.

Name of Data Matrix: ndatat
Name of Variable Matrix: exactmass
Name of Filename Matrix: filenames
Name of Totalcounts Matrix: totalcounts
Name of Samplenames Matrix: samplenames

Input Information
Make sure the data selected above is the data you want to analyze and then press the "Load Selected Data" button.

PCA Panel
After making sure the correct data is selected above. Press the 'Load Selected Data' button.

Data: None
Samples: None
Variables: None

Scaling Method
Choose one

Show Eigenvalue Scree Plot

PCA Summary
PC# %Var %Vartotal

Close Panel
Data Selection Panel
These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

- Name of Data Matrix: ndat
- Name of Variable Matrix: exactmass
- Name of Filename Matrix: filenames
- Name of Totalcounts Matrix: totalcounts
- Name of Samplenames Matrix: samplenames

Input Information
Make sure the data selected above is the data you want to analyze and then press the "Load Selected Data" button.

PCA Panel
Choose the pre-processing.scaling you want to use from the drop down menu.
Here we will choose 'SQRT-MeanCenter'.
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

Name of Data Matrix: ndatat
Name of Variable Matrix: exactmass
Name of Filename Matrix: filenames
Name of Totalcounts Matrix: totalcounts
Name of Samplenames Matrix: samplenames

Input Information

Make sure the data selected above is the data you want to analyze and then press the "Load Selected Data" button.

PCA Panel

Load Selected Data

Data: ndatat
Samples: samplenames
Variables: exactmass

Scaling Method: SQRT-MeanCen...

Run PCA

Then Press the 'Run PCA' button.

Show Eigenvalue Scree Plot

PCA Summary

PC#  %Var  %Vartotal

Close Panel
A new panel appears that allows you to quickly browse the scores and loadings plots. These plots can be saved by pressing the 'Ext Plot' button and saving the files as desired.

This new panel is for saving the current PCA model to the Matlab workspace.
A summary of the PCA model is provided here showing the % variance captured for each PC (%Var) and the cumulative %variance for each sequential PC (%Vartotal).
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop-down menus to select the data and information you want to use in your analysis.

- Name of Data Matrix: ndatat
- Name of Variable Matrix: exactmass
- Name of Filename Matrix: filenames
- Name of Totalcounts Matrix: totalcounts
- Name of Samplenames Matrix: samplenames

Input Information

Make sure the data selected above is the data you want to analyze and then press the "Load Selected Data" button.

- Data: ndatat
- Samples: samplenames
- Variables: exactmass

Scaling Method: SQRT-MeanCentering

Run PCA

Press the 'Show Eigenvalue Scree Plot' to see a plot of the eigenvalue vs PC number.

PCA Panel

Plot PCA Scores and Loadings

These plots are for a quick check of the results and cannot be saved. Loading plots default to m/z vs PC#.

X-Axis

PC#

Plot Scores

Y-Axis

PC#

Plot Loads

Save PCA Data to Workspace

Number of PCs to keep

Scores output name

Loadings output name

Percent variance output name

Model output name

Save To Workspace

Close Panel
This is a plot of the eigenvalue vs PC number. This plot is often used to decide how many PCs to keep in a model.

The "rule of thumb" is to look for an inflection point in the line and keep the PCs before that inflection point.

In this data set it looks like there is an inflection point at PC 6 or 8.
Close this panel to get back to the rest of the Main PCA panel window.
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

Input Information

Make sure the data selected above is the data you want to analyze and then press the "Load Selected Data" button.

PCA Panel

Plot PCA Scores and Loadings

These plots are for a quick check of the results and cannot be saved. Loading plots default to m/z vs PC#.

X-Axis

Y-Axis

To create a plot, choose which axes you want to plot from the drop down menus.

Scores and loadings plots are most easily interpreted when plotting 1 PC at a time.

Here we will use the sample number for the x axis.
And plot PC1 on the y axis.
Press the 'Plot Scores' button to see a scores plot.
Press the 'Plot Loads' button to see a Loadings plot.
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop-down menus to select the data and information you want to use in your analysis.

- Name of Data Matrix: ndatat
- Name of Variable Matrix: exactmass
- Name of Filename Matrix: filenames
- Name of Totalcounts Matrix: totalcounts
- Name of Samplenames Matrix: samplenames

Input Information

Make sure the data selected above is the data you want to analyze and then press the "Load Selected Data" button.

PCA Panel

Plot PCA Scores and Loadings

- X Axis
- Y Axis

The scores plot is shown here.

PCA Summary

<table>
<thead>
<tr>
<th>PC#</th>
<th>%Var</th>
<th>%Vartotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31.4</td>
<td>31.4</td>
</tr>
<tr>
<td>2</td>
<td>22.5</td>
<td>53.9</td>
</tr>
<tr>
<td>3</td>
<td>18.8</td>
<td>72.7</td>
</tr>
<tr>
<td>4</td>
<td>12.3</td>
<td>85</td>
</tr>
<tr>
<td>5</td>
<td>8.4</td>
<td>93.4</td>
</tr>
<tr>
<td>6</td>
<td>6.1</td>
<td>99.5</td>
</tr>
<tr>
<td>7</td>
<td>4.1</td>
<td>99.9</td>
</tr>
<tr>
<td>8</td>
<td>0.1</td>
<td>99.9</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>
Press the 'EXT Plot' button to open the figure in an external Matlab figure.
This figure can be modified and saved as desired.
Next we'll plot PC1 vs PC2.
Select PC1 from the x-axis menu.
Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop-down menus to select the data and information you want to use in your analysis.

- Name of Data Matrix: ndat
- Name of Variable Matrix: exactmass
- Name of Filename Matrix: filenames
- Name of Totalcounts Matrix: totalcounts
- Name of Sample Names Matrix: samenames

Input Information

Make sure the data selected above is the data you want to analyze and then press the "Load Selected Data" button.

- Data: ndat
- Samples: samenames
- Variables: exactmass

Scaling Method: SQRT-MeanCenter

Run PCA

Show Eigenvalue Scree Plot

PCA Summary

<table>
<thead>
<tr>
<th>PC #</th>
<th>%Var</th>
<th>%VarTotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31.4</td>
<td>31.4</td>
</tr>
<tr>
<td>2</td>
<td>22.5</td>
<td>53.9</td>
</tr>
<tr>
<td>3</td>
<td>18.8</td>
<td>72.7</td>
</tr>
<tr>
<td>4</td>
<td>12.3</td>
<td>85</td>
</tr>
<tr>
<td>5</td>
<td>8.4</td>
<td>93.4</td>
</tr>
<tr>
<td>6</td>
<td>6.1</td>
<td>99.5</td>
</tr>
<tr>
<td>7</td>
<td>4.4</td>
<td>99.9</td>
</tr>
<tr>
<td>8</td>
<td>1.1</td>
<td>99.9</td>
</tr>
<tr>
<td>9</td>
<td>0.1</td>
<td>99.9</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

Plot PCA Scores and Loadings

These plots are for a quick check of the results and cannot be saved. Loading plots default to m/z vs PC#

- X-Axis PC# 1
- Y-Axis PC# 2

And PC2 from the y-axis menu.

Save PCA Data to Workspace

Number of PCs to keep
- Scores output name
- Loadings output name
- Percent variance output name
- Model output name

Save To Workspace

Close Panel
Press the 'Plot Scores' button to update the scores plot.
Once you have browsed the PCA scores and loadings and decided you want to save the results, you must first choose how many PCs to save.

The number of PC you keep can be based off:
- the point of inflection in the scree plot
- the PC number where the % variance captured does not change significantly
- Where you see only noise in the scores plot

Enter the information requested in the boxes and press the 'Save To Workspace' button.

Scores output name: scores
Loadings output name: loads
Percent variance output name: var
Model output name: model

Close Panel
The data is saved to the workspace using the names provided.
Data Selection Panel
These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

Input Information
Make sure the data selected above is the data you want to analyze and then press the "Load Selected Data" button.

PCA Panel
Plot PCA Scores and Loadings
These plots are for a quick check of the results and cannot be saved. Loading plots default to m/z vs PC#.

X-Axis
1

Y-Axis
2

Scores output name

Loadings output name

Percent variance output name

Model output name

Press the 'Close Panel' button to close the panel.

PCA Summary

<table>
<thead>
<tr>
<th>PC#</th>
<th>%Var</th>
<th>%Vartotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31.4</td>
<td>31.4</td>
</tr>
<tr>
<td>2</td>
<td>22.5</td>
<td>53.9</td>
</tr>
<tr>
<td>3</td>
<td>18.8</td>
<td>72.7</td>
</tr>
<tr>
<td>4</td>
<td>12.3</td>
<td>85</td>
</tr>
<tr>
<td>5</td>
<td>8.4</td>
<td>93.4</td>
</tr>
<tr>
<td>6</td>
<td>6.1</td>
<td>99.5</td>
</tr>
<tr>
<td>7</td>
<td>0.4</td>
<td>99.9</td>
</tr>
<tr>
<td>8</td>
<td>0.1</td>
<td>99.9</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>99.9</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>
### Data Selection Panel

These are the main input data that will be used in further analysis unless you specify otherwise. Use the drop down menus to select the data and information you want to use in your analysis.

<table>
<thead>
<tr>
<th>Name of Data Matrix</th>
<th>Name of Variable Matrix</th>
<th>Name of Filename Matrix</th>
<th>Name of Totalcounts Matrix</th>
<th>Name of Samplenames Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>ndatat</td>
<td>exactmass</td>
<td>filenames</td>
<td>totalcounts</td>
<td>samplenames</td>
</tr>
</tbody>
</table>

---

That's it for this tutorial.

Press the green button on the left to go back to the previous step. Press the button the right to go back to the beginning of the tutorial.