

# Similarity matrix

Similarity matrix task is only available on bulk RNA-seq count matrix data node. It is used to compute the correlation of every sample/or feature vs every other sample/or feature. The result is a matrix with the same set of samples/or features on rows and columns, the value in the matrix is correlation coefficient --r.

Click on **Similarity matrix** task in *Correlation* section on the menu (Figure 1)

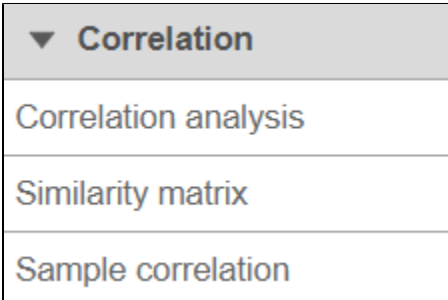


Figure 1. Similarity matrix menu

When the dialog opens, you will be asked to select whether the calculation samples or features (Figure 2). There are three correlation method options:

$$r_{xy} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}}$$

Pearson: linear correlation:

$$r_s = \frac{\sum_i (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_i (R_i - \bar{R})^2} \sqrt{\sum_i (S_i - \bar{S})^2}}$$

Spearman: rank correlation:

$$\tau = \frac{concordant - discordant}{\sqrt{concordant + discordant + extraY} \sqrt{concordant + discordant + extraX}}$$

Kendal: rank correlation:

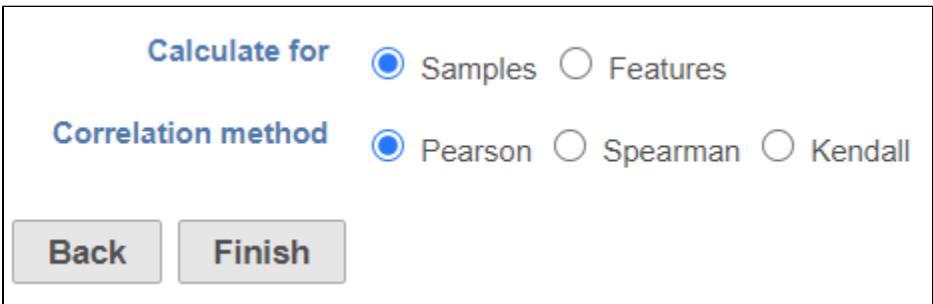


Figure 2. Similarity matrix dialog

Click **Finish** to run the task. The output report of this task can be displayed in heatmap and/or table in the data viewer.

## Additional Assistance

If you need additional assistance, please visit [our support page](#) to submit a help ticket or find phone numbers for regional support.



Your Rating:  Results:  12 rates