

Select by Data Index

Match instances by index from data subset.

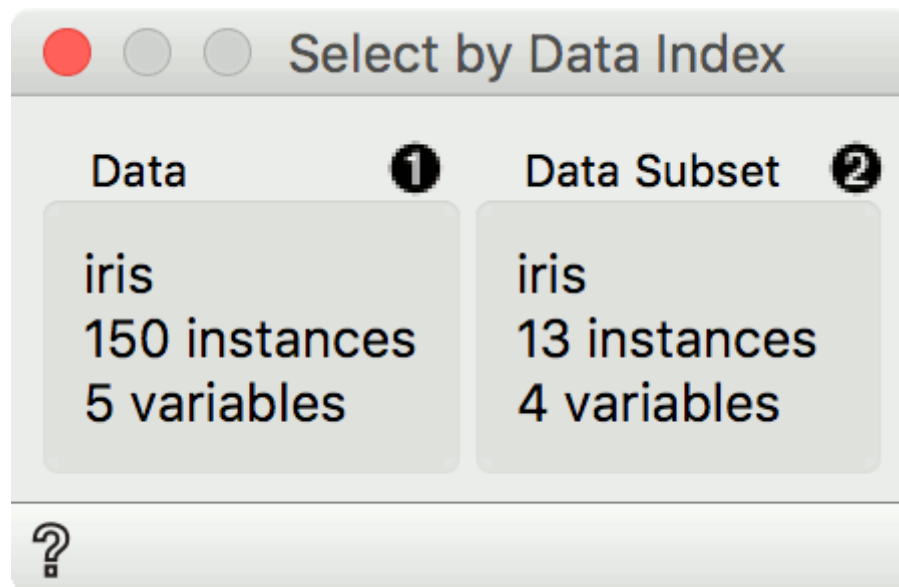
Inputs

- Data: reference data set
- Data Subset: subset to match

Outputs

- Matching data: subset from reference data set that matches indices from subset data
- Unmatched data: subset from reference data set that does not match indices from subset data
- Annotated data: reference data set with an additional column defining matches

Select by Data Index enables matching the data by indices. Each row in a data set has an index and given a subset, this widget can match these indices to indices from the reference data. Most often it is used to retrieve the original data from the transformed data (say, from PCA space).



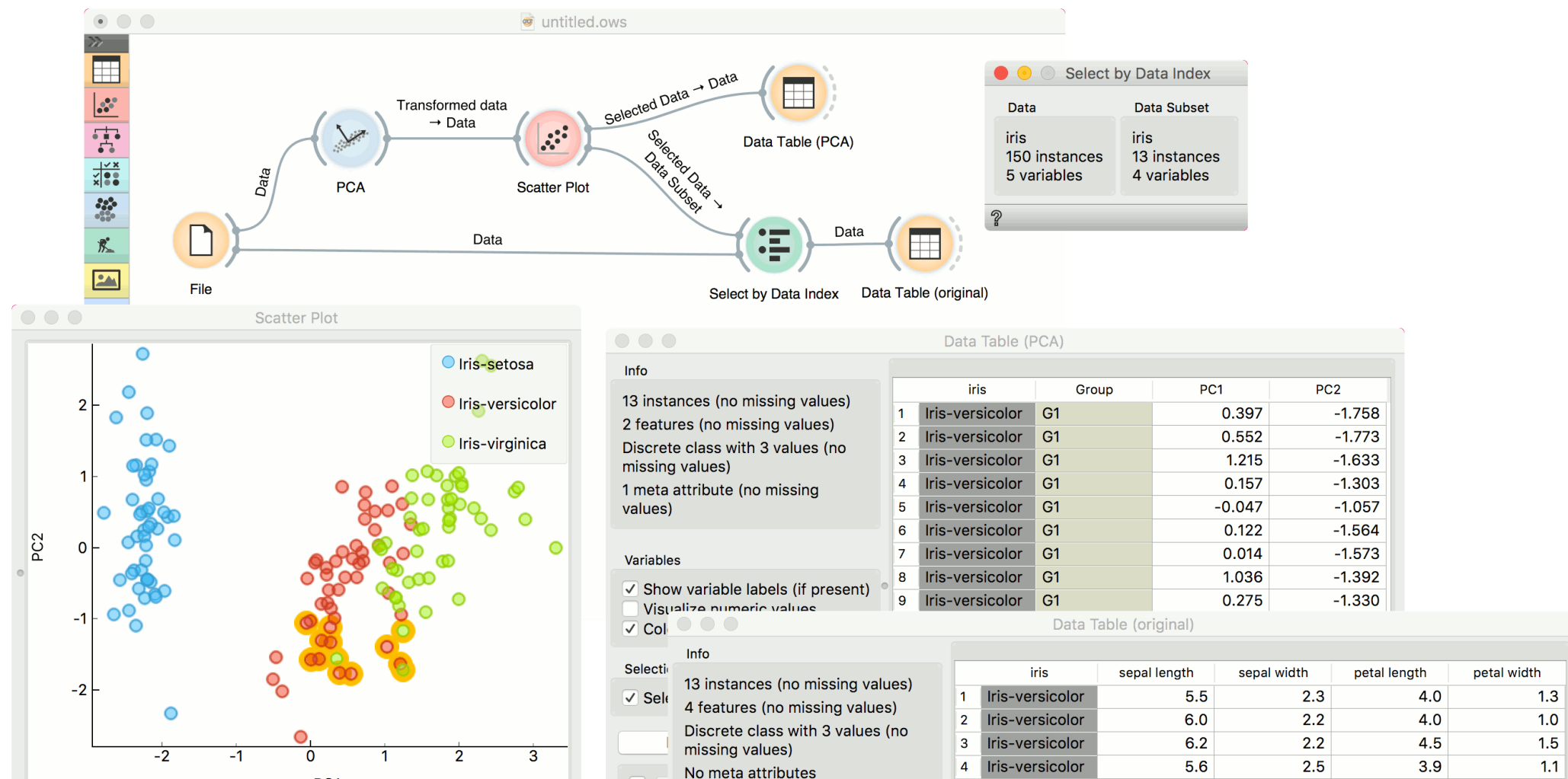
1. Information on the reference data set. This data is used as index reference.
2. Information on the data subset. The indices of this data set are used to find matching data in the reference data set. Matching data are on the output by default.

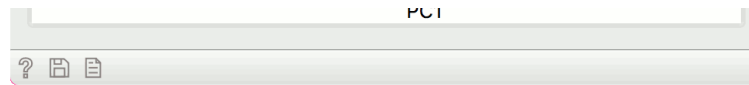
Example

A typical use of **Select by Data Index** is to retrieve the original data after a transformation. We will load *iris.tab* data in the **File** widget. Then we will transform this data with **PCA**. We can project the transformed data in a **Scatter Plot**, where we can only see PCA components and not the original features.

Now we will select an interesting subset (we could also select the entire data set). If we observe it in a **Data Table**, we can see that the data is transformed. If we would like to see this data with the original features, we will have to retrieve them with **Select by Data Index**.

Connect the original data and the subset from **Scatter Plot** to **Select by Data Index**. The widget will match the indices of the subset with the indices of the reference (original) data and output the matching reference data. A final inspection in another **Data Table** confirms the data on the output is from the original data space.





Orange Data Mining - Select by Data Index

☒ Variables

- ☒ Show variable labels (if present)
- ☐ Visualize numeric values
- ☒ Color by instance classes

☒ Selection

☒ Select full rows

Restore Original Order

☒ Send Automatically

5	Iris-versicolor	5.7	2.6	3.5	1.0
6	Iris-versicolor	5.5	2.4	3.8	1.1
7	Iris-versicolor	5.5	2.4	3.7	1.0
8	Iris-versicolor	6.3	2.3	4.4	1.3
9	Iris-versicolor	5.5	2.5	4.0	1.3
10	Iris-versicolor	5.5	2.6	4.4	1.2
11	Iris-virginica	4.9	2.5	4.5	1.7
12	Iris-virginica	5.7	2.5	5.0	2.0
13	Iris-virginica	6.0	2.2	5.0	1.5