Neighbors

Compute nearest neighbors in data according to reference.

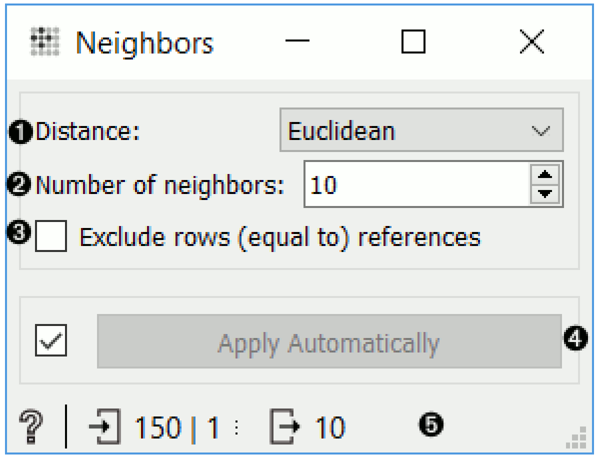
Inputs

- Data: An input data set.
- Reference: A reference data for neighbor computation.

Outputs

- Neighbors: A data table of nearest neighbors according to reference.

The **Neighbors** widget computes nearest neighbors for a given reference and for a given distance measure. The reference can be either one instance or more instances. In the case with one reference widget outputs closest `n` instances from data where `n` is set by the **Number of neighbors** option in the widget. When reference contains more instances widget computes the combined distance for each data instance as a minimum of distances to each reference. Widget outputs `n` data instances with lowest combined distance.


Neighbors

1 Distance: Euclidean

2 Number of neighbors: 10

3 ☐ Exclude rows (equal to) references

☑ Apply Automatically 4

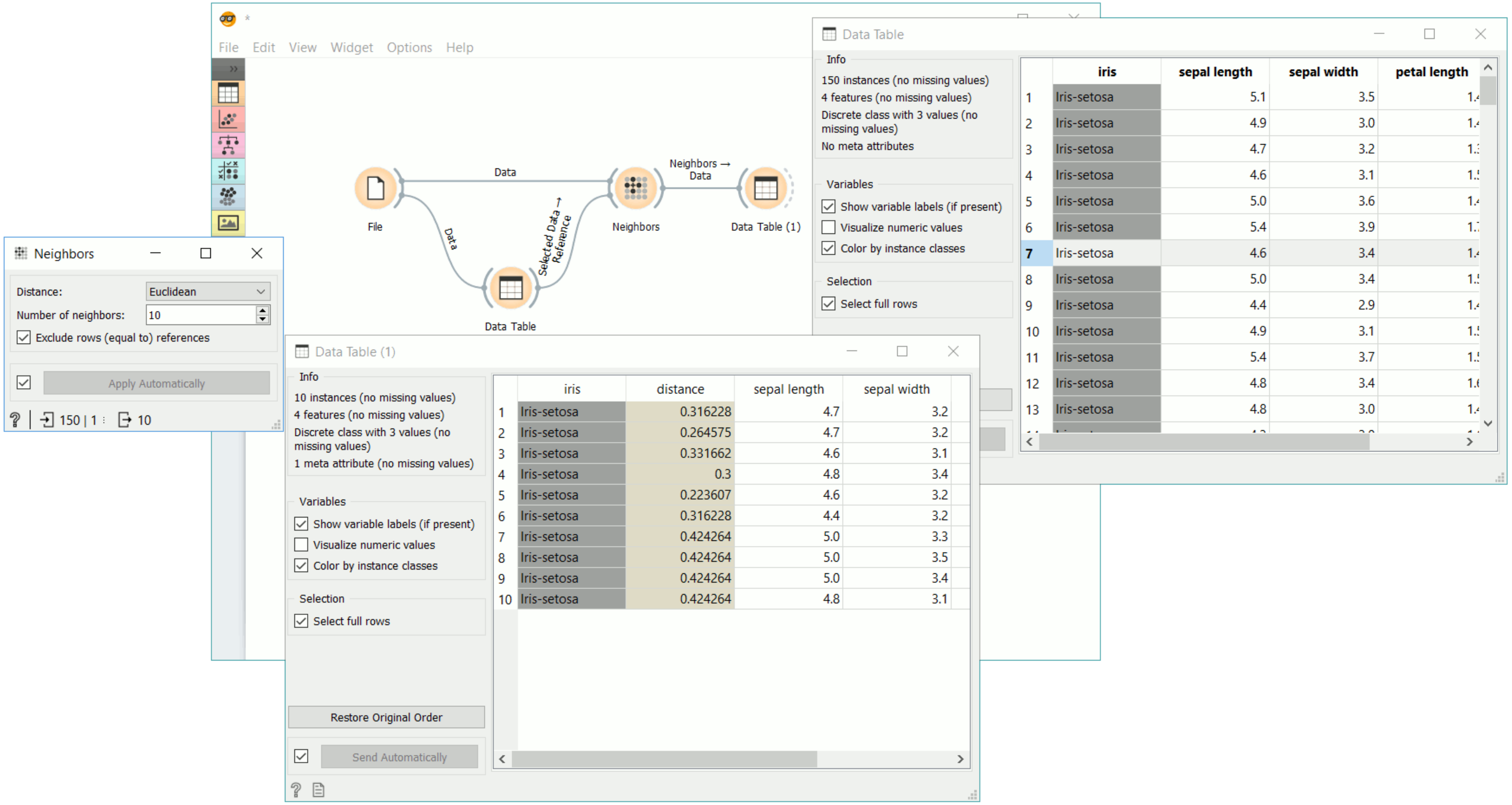
150 | 1 10 5

1. Distance measure for computing neighbors. Supported measures are: Euclidean, Manhattan, Mahalanobis, Cosine, Jaccard, Spearman, absolute Spearman, Pearson, absolute Pearson.
2. Number of neighbors on the output.
3. If *Exclude rows (equal to) references* is ticked, data instances that are highly similar to the reference (distance < 1e-5), will be excluded.
4. Click *Apply* to commit the changes. To communicate changes automatically tick *Apply Automatically*.
5. Status bar with access to widget help and information on the input and output data.

Examples

In the first example, we used *iris* data and passed it to **Neighbors** and to Data Table. In **Data Table**, we selected an instance of iris, that will serve as our reference, meaning we wish to retrieve 10 closest examples to the select data instance. We connect **Data Table** to **Neighbors** as well.

We can observe the results of neighbor computation in **Data Table (1)**, where we can see 10 closest images to our selected iris flower.



Neighbors

Distance: Euclidean
Number of neighbors: 10
☑ Exclude rows (equal to) references
☑ Apply Automatically
150 | 1 → 10

Data Table

Info
150 instances (no missing values)
4 features (no missing values)
Discrete class with 3 values (no missing values)
No meta attributes

Variables
☑ Show variable labels (if present)
☐ Visualize numeric values
☑ Color by instance classes

Selection
☑ Select full rows

	iris	sepal length	sepal width	petal length
1	Iris-setosa	5.1	3.5	1.4
2	Iris-setosa	4.9	3.0	1.4
3	Iris-setosa	4.7	3.2	1.3
4	Iris-setosa	4.6	3.1	1.5
5	Iris-setosa	5.0	3.6	1.4
6	Iris-setosa	5.4	3.9	1.7
7	Iris-setosa	4.6	3.4	1.4
8	Iris-setosa	5.0	3.4	1.5
9	Iris-setosa	4.4	2.9	1.4
10	Iris-setosa	4.9	3.1	1.5
11	Iris-setosa	5.4	3.7	1.5
12	Iris-setosa	4.8	3.4	1.6
13	Iris-setosa	4.8	3.0	1.4

Data Table (1)

Info
10 instances (no missing values)
4 features (no missing values)
Discrete class with 3 values (no missing values)
1 meta attribute (no missing values)

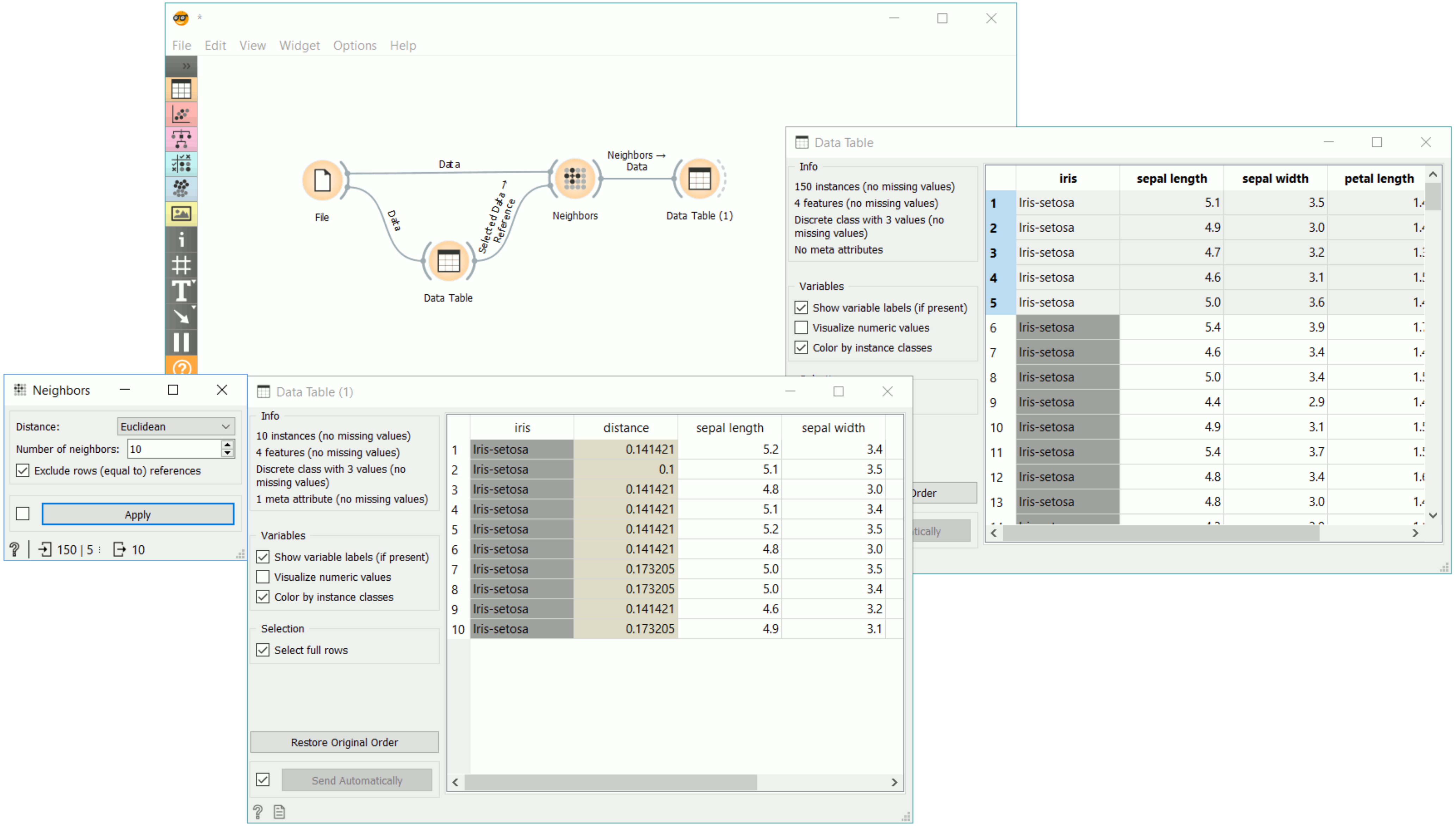
Variables
☑ Show variable labels (if present)
☐ Visualize numeric values
☑ Color by instance classes

Selection
☑ Select full rows

Restore Original Order
☑ Send Automatically

	iris	distance	sepal length	sepal width
1	Iris-setosa	0.316228	4.7	3.2
2	Iris-setosa	0.264575	4.7	3.2
3	Iris-setosa	0.331662	4.6	3.1
4	Iris-setosa	0.3	4.8	3.4
5	Iris-setosa	0.223607	4.6	3.2
6	Iris-setosa	0.316228	4.4	3.2
7	Iris-setosa	0.424264	5.0	3.3
8	Iris-setosa	0.424264	5.0	3.5
9	Iris-setosa	0.424264	5.0	3.4
10	Iris-setosa	0.424264	4.8	3.1

Now change the selection **Data Table** to multiple examples. As a result, we get instances with closest combined distances to the references. The method computes the combined distance as a minimum of distances to each reference.



Data Table

Info

150 instances (no missing values)
4 features (no missing values)
Discrete class with 3 values (no missing values)
No meta attributes

Variables

- [x] Show variable labels (if present)
- [ ] Visualize numeric values
- [x] Color by instance classes

	iris	sepal length	sepal width	petal length
1	Iris-setosa	5.1	3.5	1.4
2	Iris-setosa	4.9	3.0	1.4
3	Iris-setosa	4.7	3.2	1.3
4	Iris-setosa	4.6	3.1	1.5
5	Iris-setosa	5.0	3.6	1.4
6	Iris-setosa	5.4	3.9	1.7
7	Iris-setosa	4.6	3.4	1.4
8	Iris-setosa	5.0	3.4	1.5
9	Iris-setosa	4.4	2.9	1.4
10	Iris-setosa	4.9	3.1	1.5
11	Iris-setosa	5.4	3.7	1.5
12	Iris-setosa	4.8	3.4	1.6
13	Iris-setosa	4.8	3.0	1.4

Neighbors

Distance: Euclidean
Number of neighbors: 10
- [x] Exclude rows (equal to) references

- [ ] Apply

150 | 5   10

Data Table (1)

Info

10 instances (no missing values)
4 features (no missing values)
Discrete class with 3 values (no missing values)
1 meta attribute (no missing values)

Variables

- [x] Show variable labels (if present)
- [ ] Visualize numeric values
- [x] Color by instance classes

Selection

- [x] Select full rows

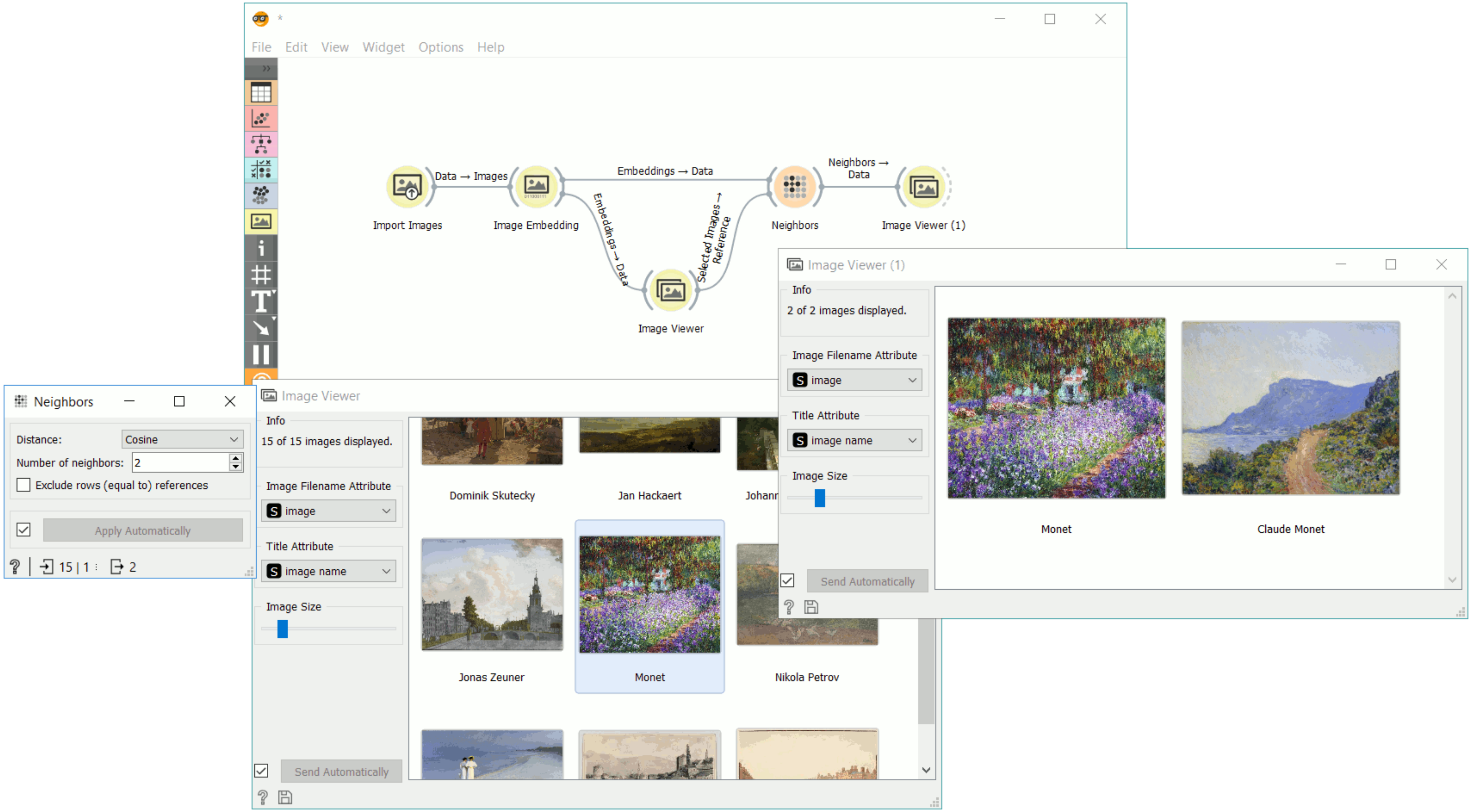
Restore Original Order

- [x] Send Automatically

	iris	distance	sepal length	sepal width
1	Iris-setosa	0.141421	5.2	3.4
2	Iris-setosa	0.1	5.1	3.5
3	Iris-setosa	0.141421	4.8	3.0
4	Iris-setosa	0.141421	5.1	3.4
5	Iris-setosa	0.141421	5.2	3.5
6	Iris-setosa	0.141421	4.8	3.0
7	Iris-setosa	0.173205	5.0	3.5
8	Iris-setosa	0.173205	5.0	3.4
9	Iris-setosa	0.141421	4.6	3.2
10	Iris-setosa	0.173205	4.9	3.1

Another example requires the installation of Image Analytics add-on. We loaded 15 paintings from famous painters with **Import Images** widget and passed them to **Image Embedding**, where we selected *Painters* embedder.

Then the procedure is the same as above. We passed embedded images to **Image Viewer** and selected a painting from Monet to serve as our reference image. We passed the image to **Neighbors**, where we set the distance measure to *cosine*, ticked off *Exclude reference*and set the neighbors to 2. This allows us to find the actual closest neighbor to a reference painting and observe them side by side in **Image Viewer (1)**.



Neighbors

Distance: Cosine
Number of neighbors: 2
☐ Exclude rows (equal to) references
☑ Apply Automatically
15 | 1 → 2

Image Viewer

Info
15 of 15 images displayed.
Image Filename Attribute: image
Title Attribute: image name
Image Size

Image Viewer (1)

Info
2 of 2 images displayed.
Image Filename Attribute: image
Title Attribute: image name
Image Size