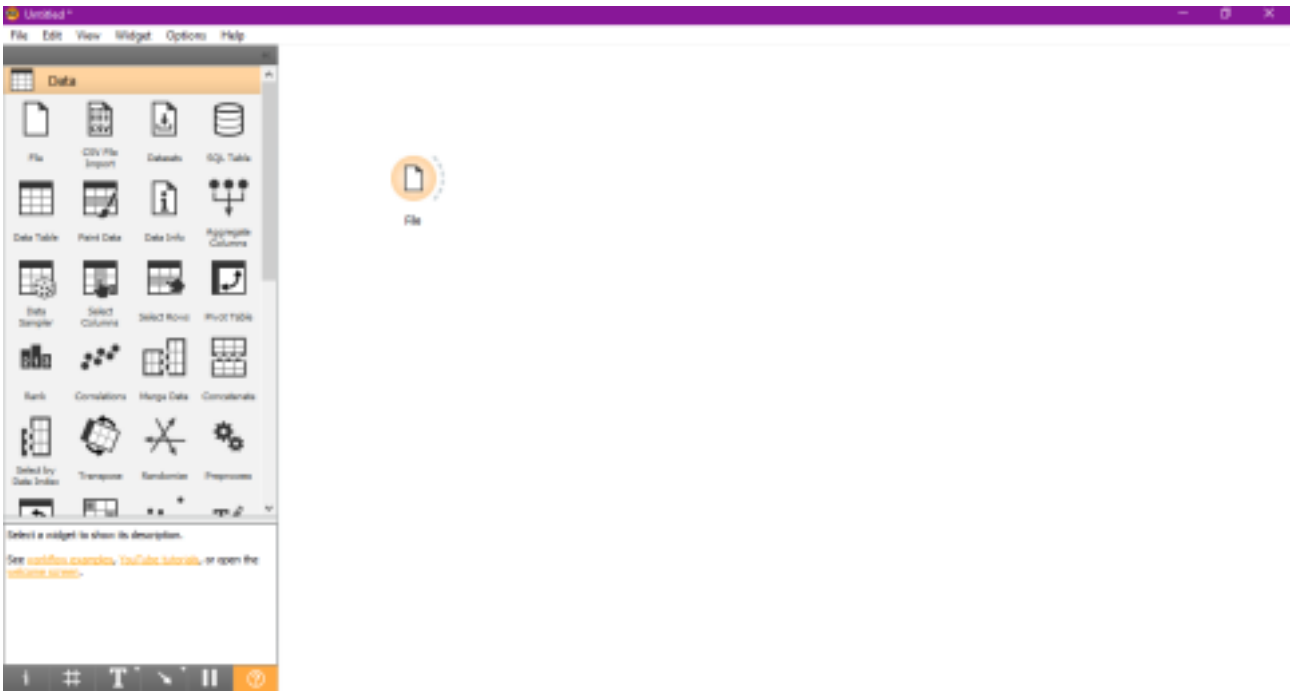
Data Warehousing and Data Mining Mini Project

Aim : Perform hierarchical clustering and classification tree in Orange3.

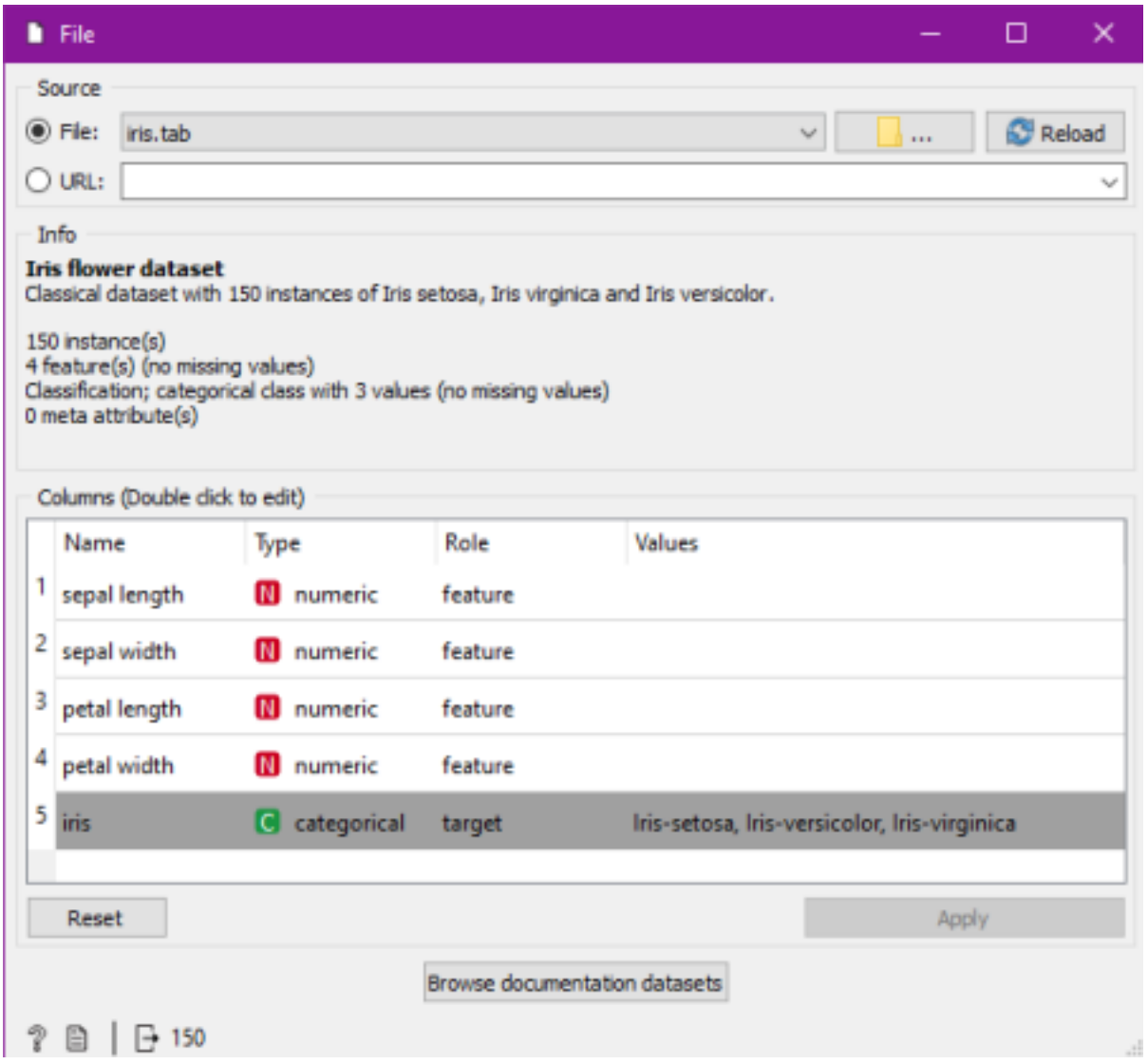
Procedure:

- Start Orange3. You should see this screen.

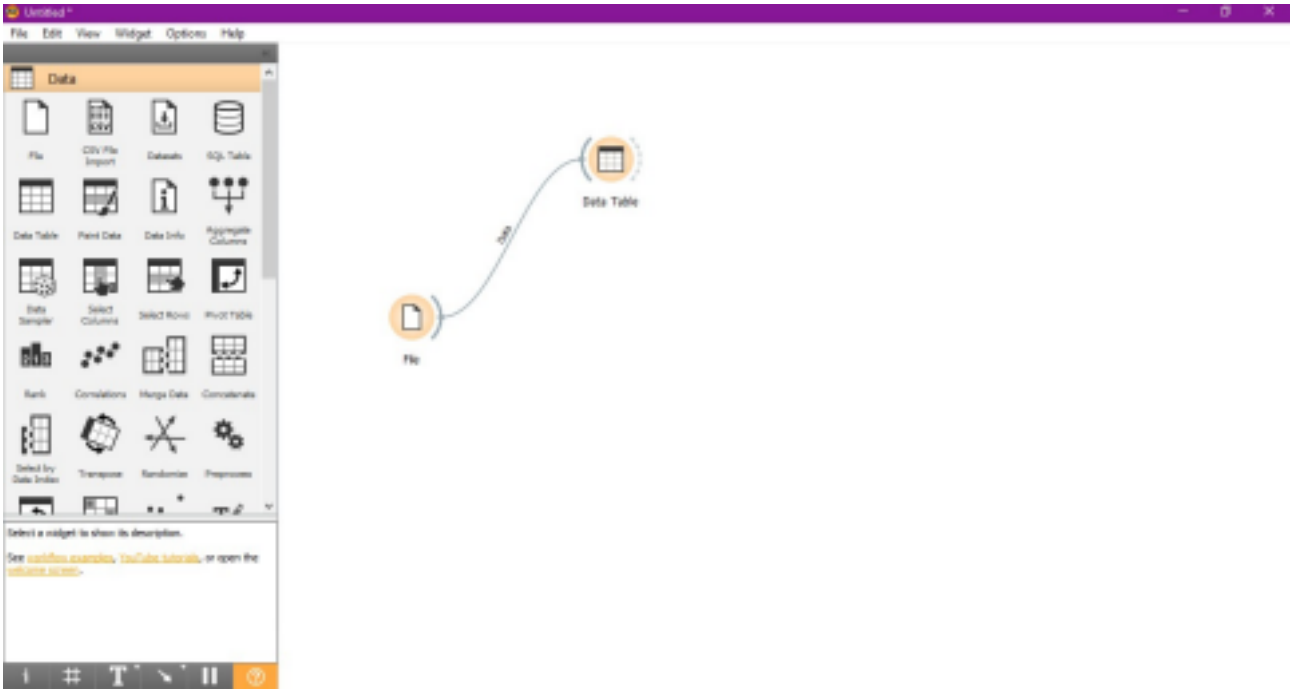
- Now, drag and drop the ‘File’ icon from the left pane onto the page.



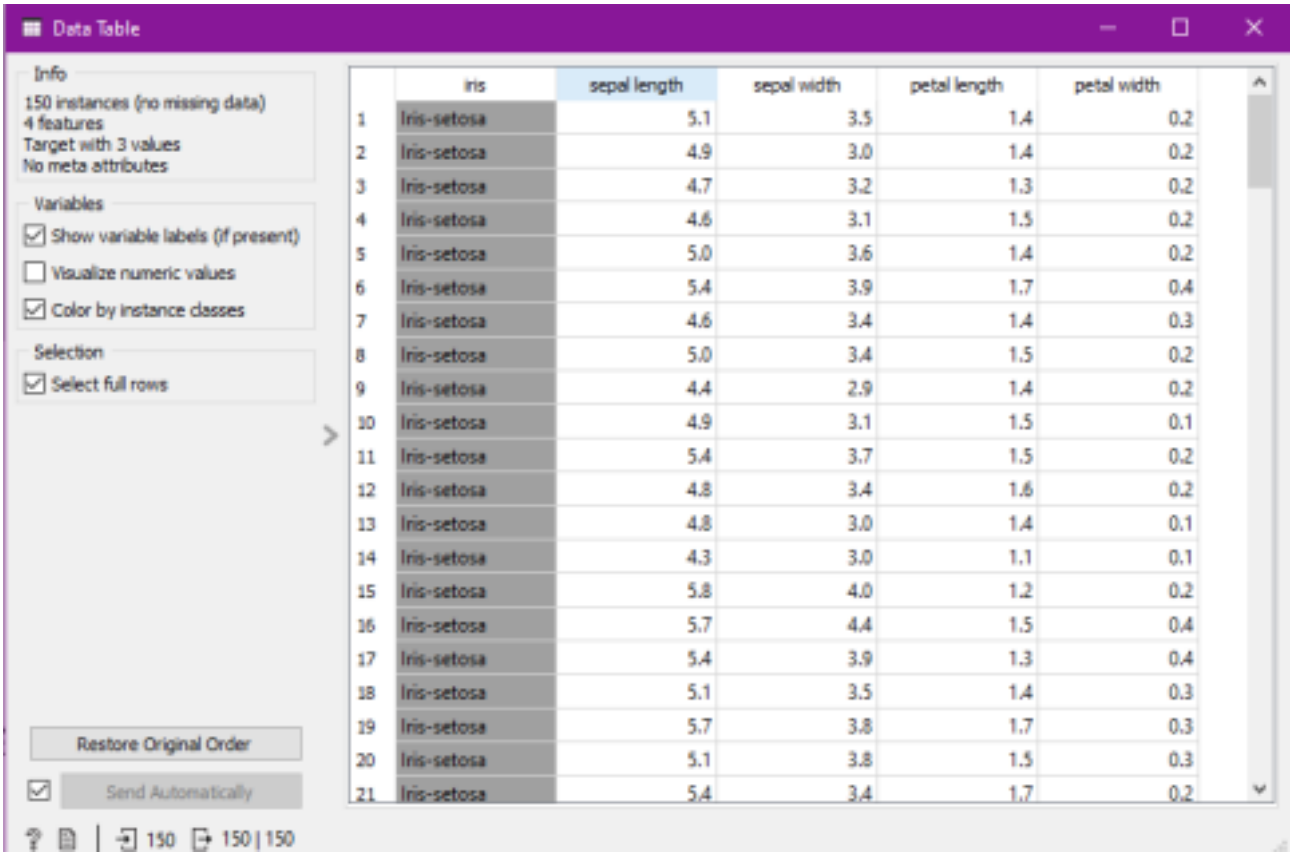
• Double click the ‘File’ widget you just dropped on the screen. You should see this window.



• Now, select the ‘Iris.tab’ dataset from the drop down list and close the window. Now, drag and drop the ‘Data Table’ widget onto the screen. Now connect the ‘File’ widget to the ‘Data Table’ widget.



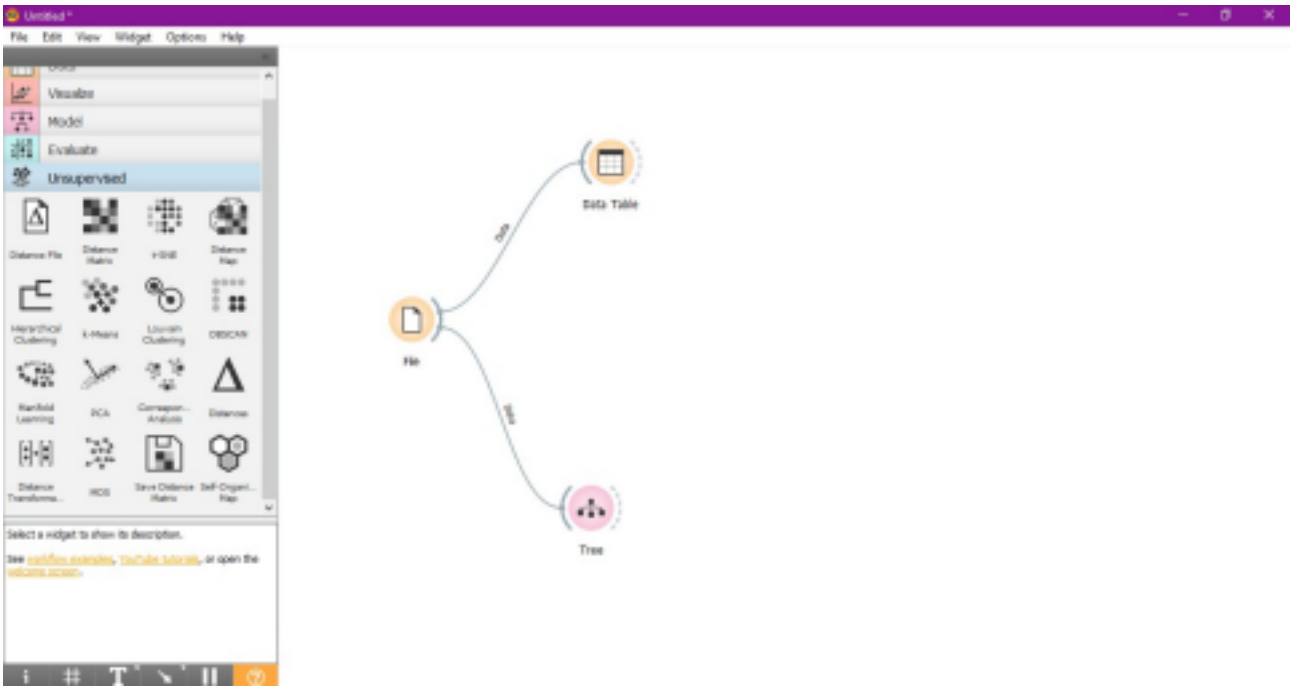
• Now, double click the ‘Data Table’ widget to generate this window.



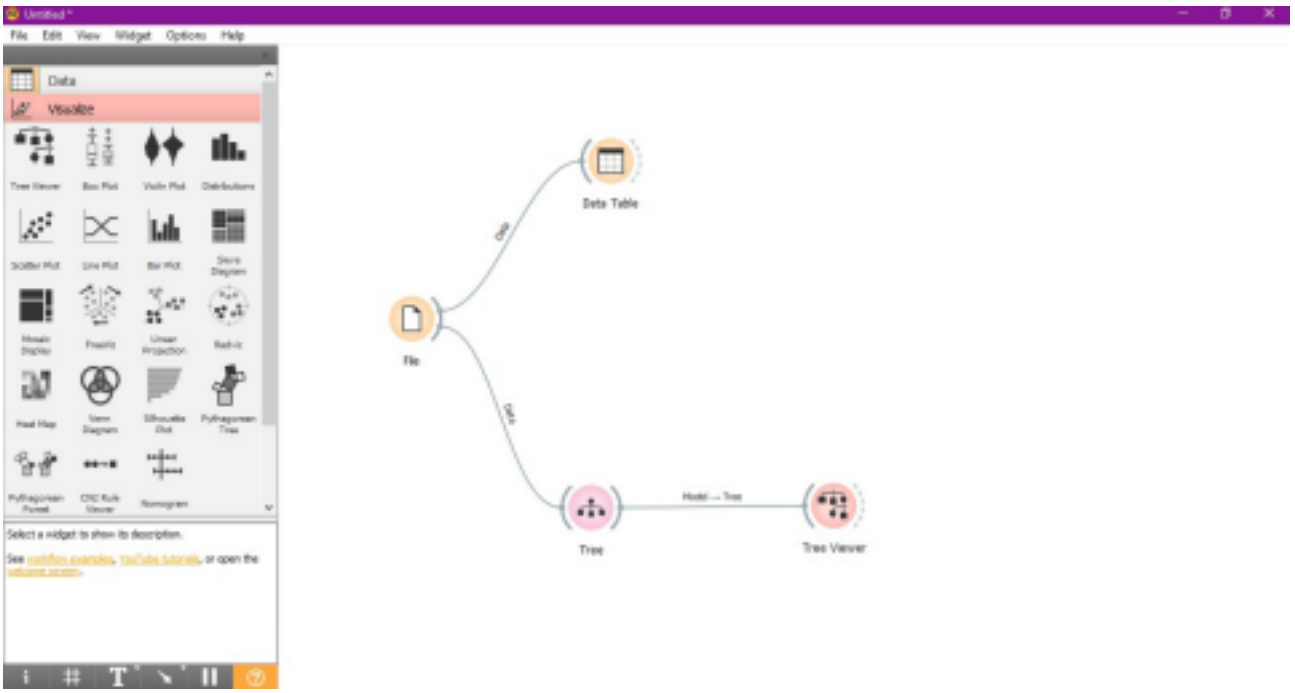
	iris	sepal length	sepal width	petal length	petal width
1	Iris-setosa	5.1	3.5	1.4	0.2
2	Iris-setosa	4.9	3.0	1.4	0.2
3	Iris-setosa	4.7	3.2	1.3	0.2
4	Iris-setosa	4.6	3.1	1.5	0.2
5	Iris-setosa	5.0	3.6	1.4	0.2
6	Iris-setosa	5.4	3.9	1.7	0.4
7	Iris-setosa	4.6	3.4	1.4	0.3
8	Iris-setosa	5.0	3.4	1.5	0.2
9	Iris-setosa	4.4	2.9	1.4	0.2
10	Iris-setosa	4.9	3.1	1.5	0.1
11	Iris-setosa	5.4	3.7	1.5	0.2
12	Iris-setosa	4.8	3.4	1.6	0.2
13	Iris-setosa	4.8	3.0	1.4	0.1
14	Iris-setosa	4.3	3.0	1.1	0.1
15	Iris-setosa	5.8	4.0	1.2	0.2
16	Iris-setosa	5.7	4.4	1.5	0.4
17	Iris-setosa	5.4	3.9	1.3	0.4
18	Iris-setosa	5.1	3.5	1.4	0.3
19	Iris-setosa	5.7	3.8	1.7	0.3
20	Iris-setosa	5.1	3.8	1.5	0.3
21	Iris-setosa	5.4	3.4	1.7	0.2

• Now, drag and drop the ‘Tree’ widget. It can be found in the ‘Model’ section. Now connect the ‘File’ widget to the ‘Tree’

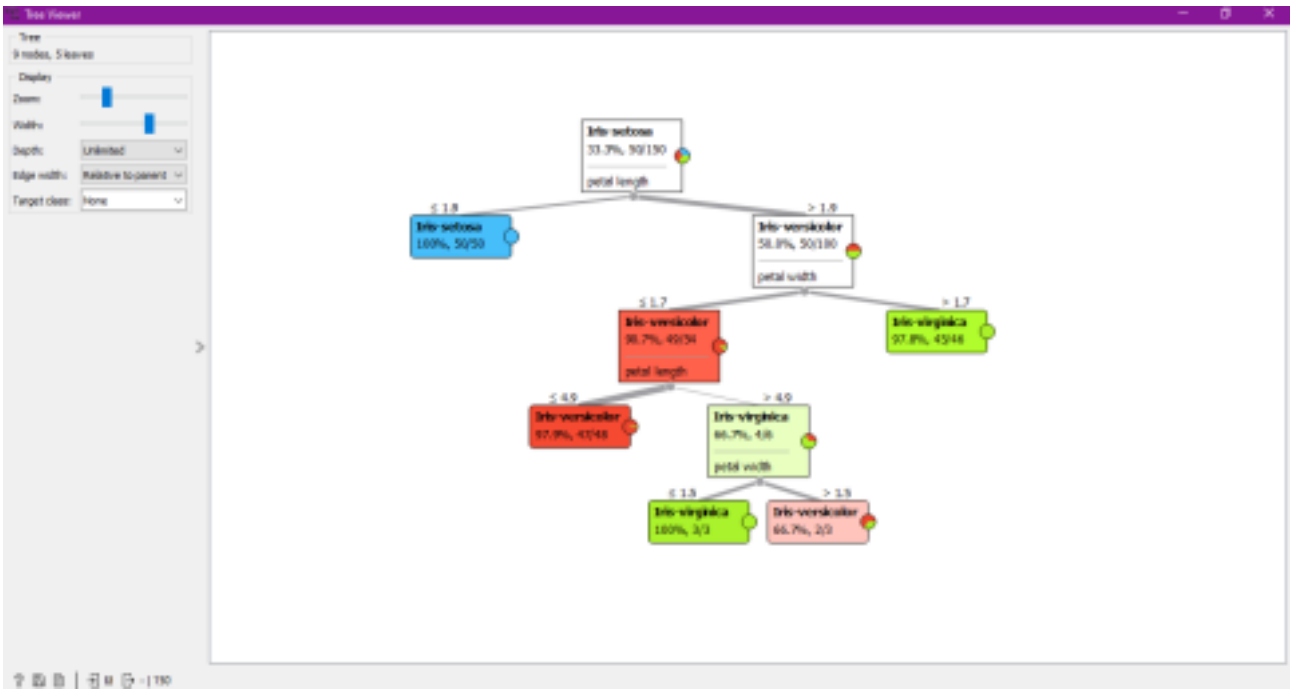
widget.



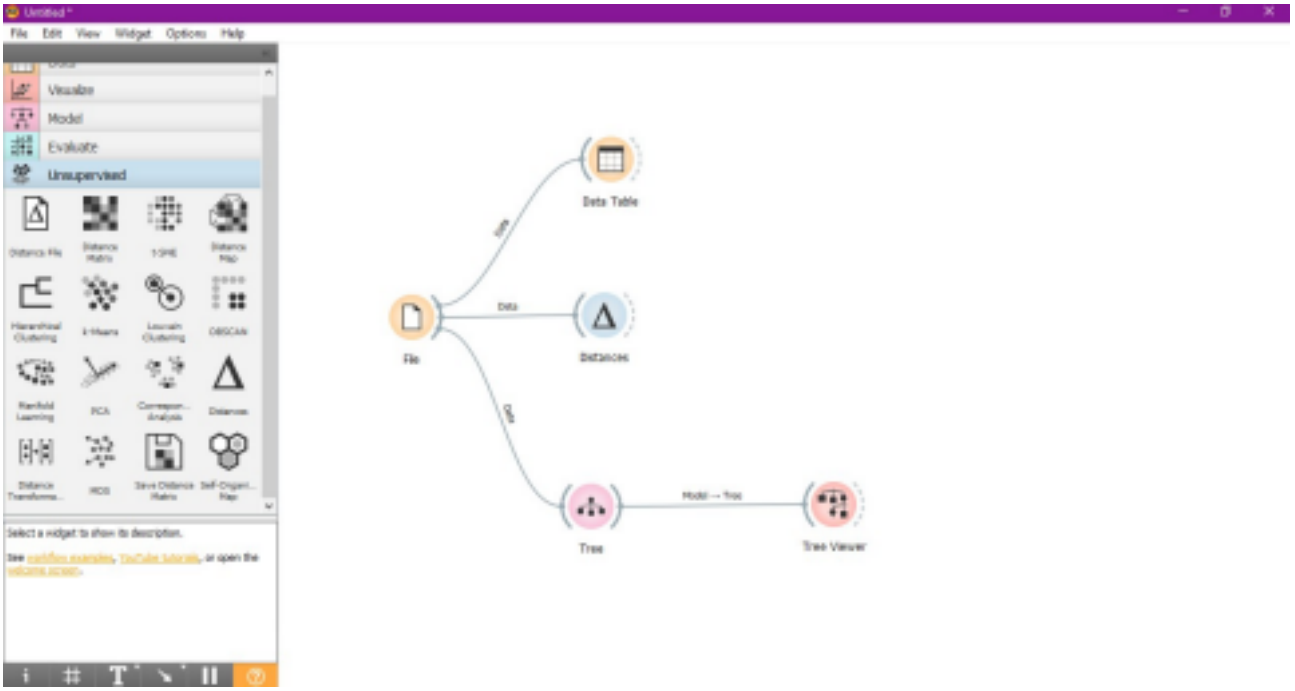
- Drag and drop the ‘Tree Viewer’ widget to the screen. It can be found in the ‘Visualize’ section. Now connect the ‘Tree’ widget to the ‘Tree Viewer’ widget.



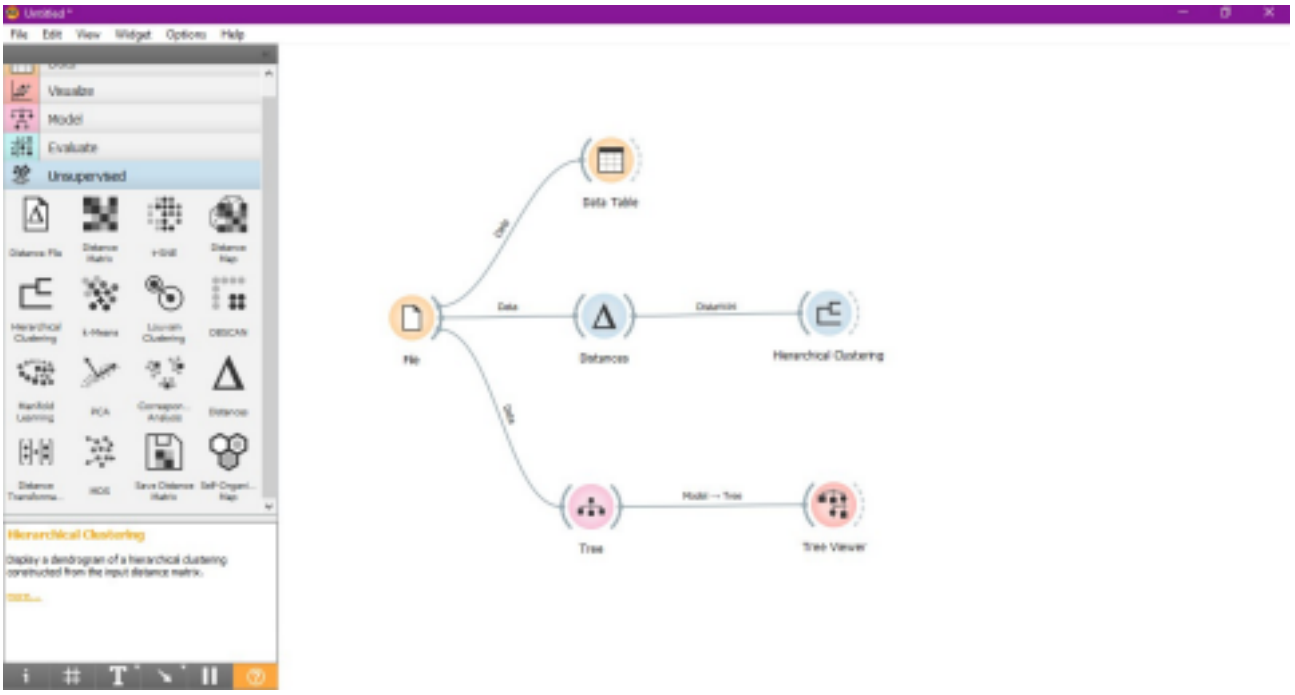
- Now, double click the ‘Tree Viewer’ widget to display the tree.



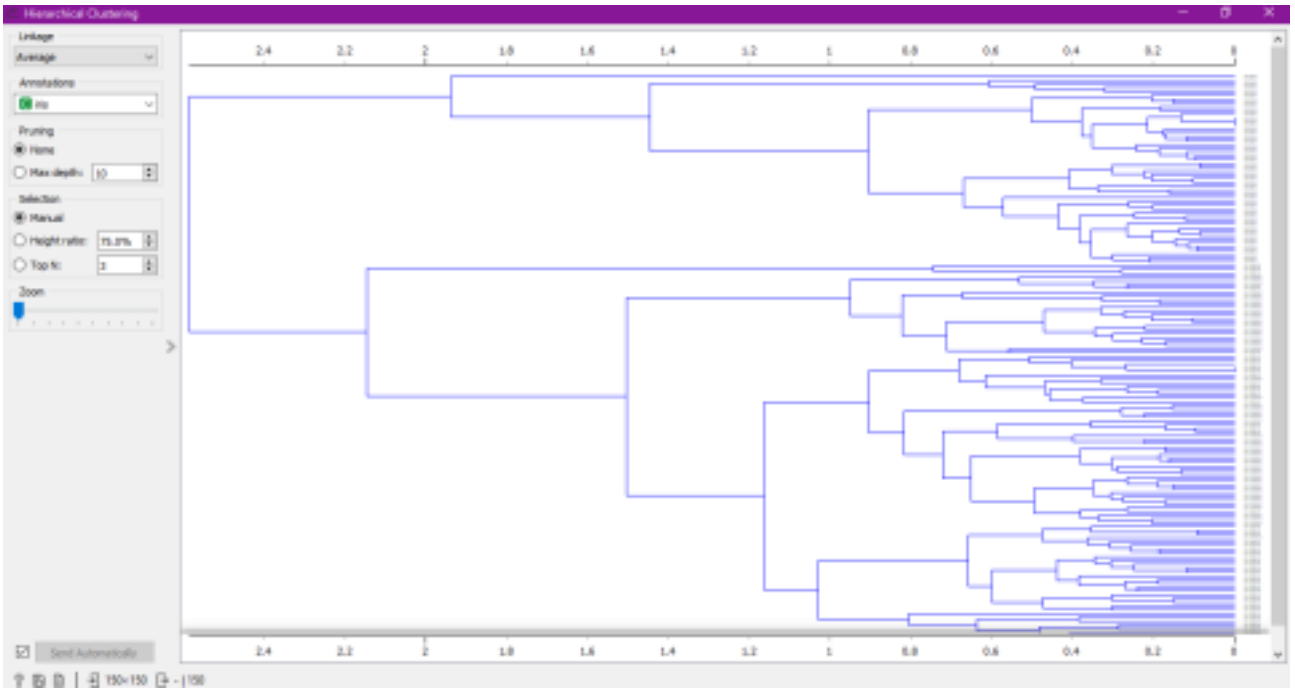
- Now, drag and drop the ‘Distances’ widget from the ‘Unsupervised’ section to the screen. Connect the ‘File’ widget to the ‘Distances’ widget.



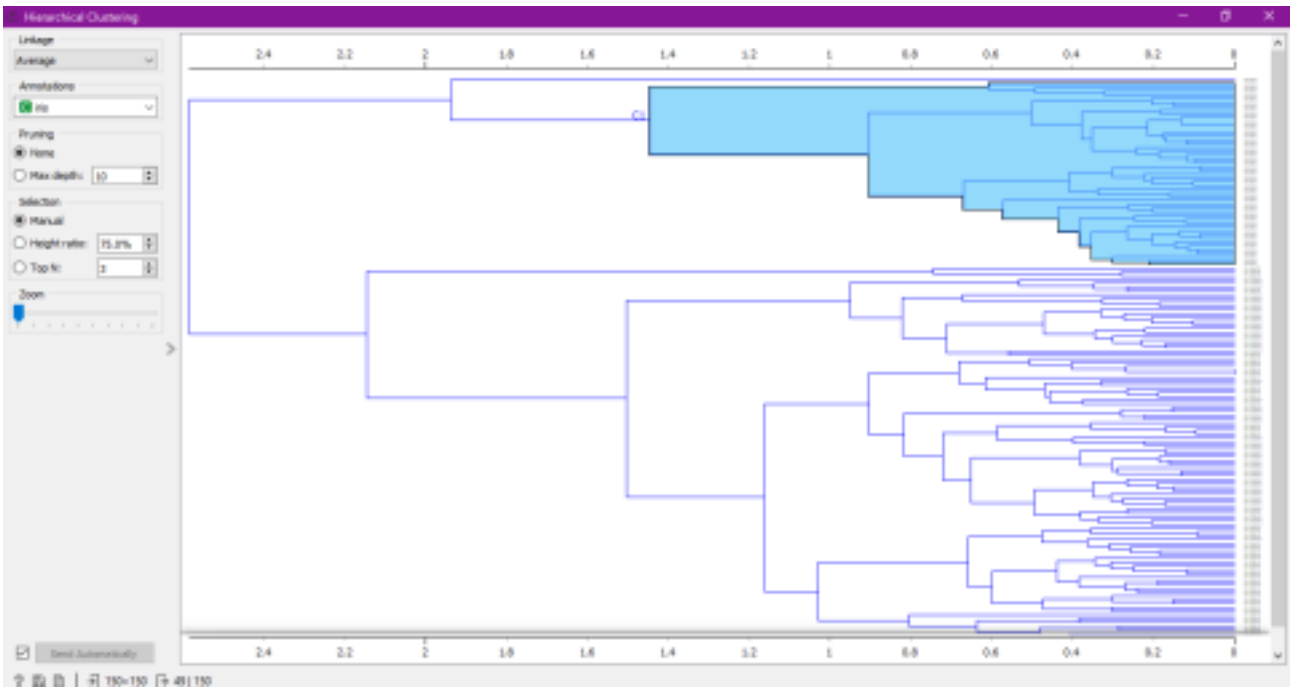
- Also drag and drop the ‘Hierarchical Clustering’ widget from the ‘Unsupervised’ section to the screen. Connect the ‘Distances’ widget to the ‘Hierarchical Clustering’ widget.



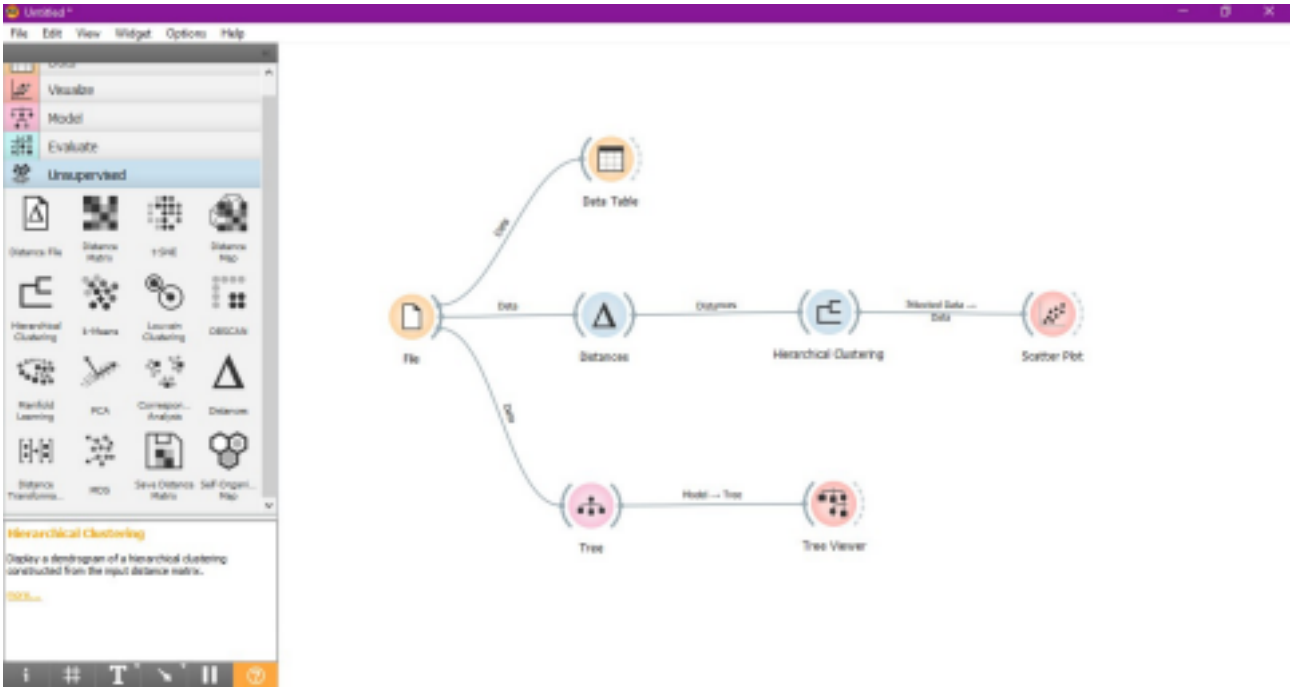
• Now, double click the ‘Hierarchical Clustering’ widget to see this window.



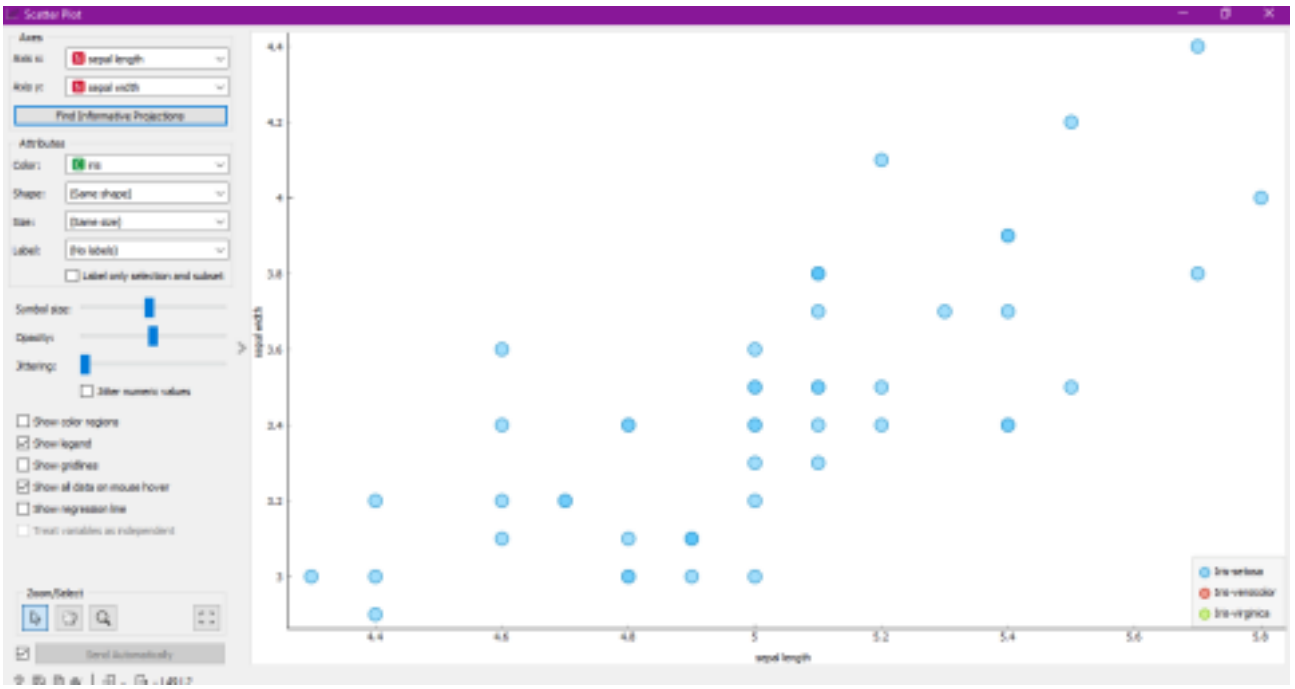
• Now, select a part of the cluster and click it to cement the selection.



- Drag and drop the ‘Scatter Plot’ widget from the ‘Visualize’ section to the screen. Connect the ‘Hierarchical Clustering’ widget to the ‘Scatter Plot’ widget.



- Now, double click the ‘Scatter Plot’ widget to view the plot.


Scatter Plot
Axes
Axis x:
sepal length
Axis y:
sepal width
Find Informative Projections
Attributes
Color:
iris
Shape:
(Same shape)
Size:
(Same size)
Label:
(No labels)
Label only selected and subset
Symbol size:
Opacity:
Jittering:
Jitter numeric values
Show color regions
Show legend
Show gridlines
Show all data on mouse hover
Show regression line
Treat variables as independent
Zoom/Select
Send Automatically
sepal width
sepal length
4.4
4.2
4
3.8
3.6
3.4
3.2
3
4.4
4.6
4.8
5
5.2
5.4
5.6
5.8
Iris-setosa
Iris-versicolor
Iris-virginica