Name: Heet Dhanuka
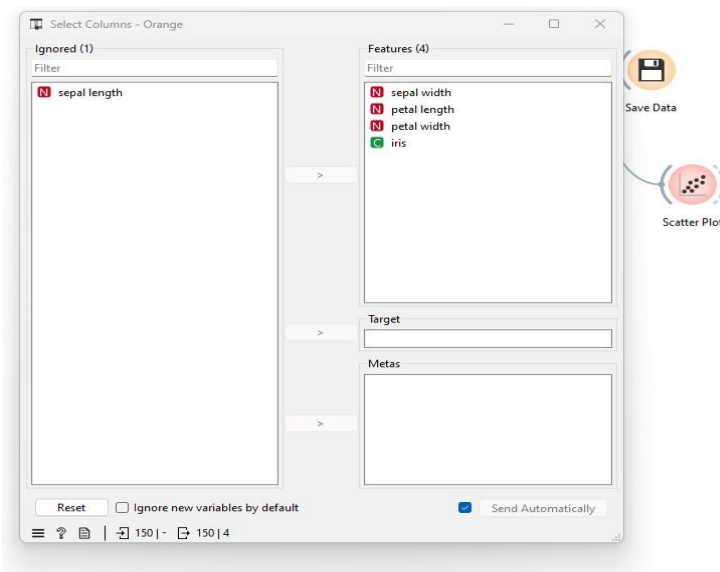
Roll No.: 34

Batch: B2

<u>DMDW Practical 5</u>

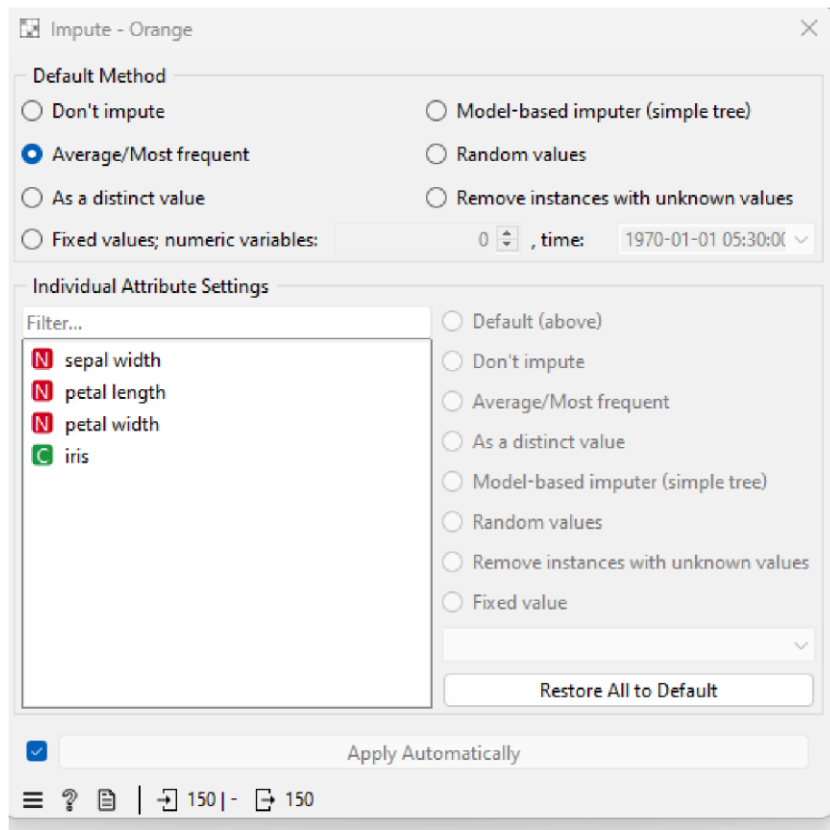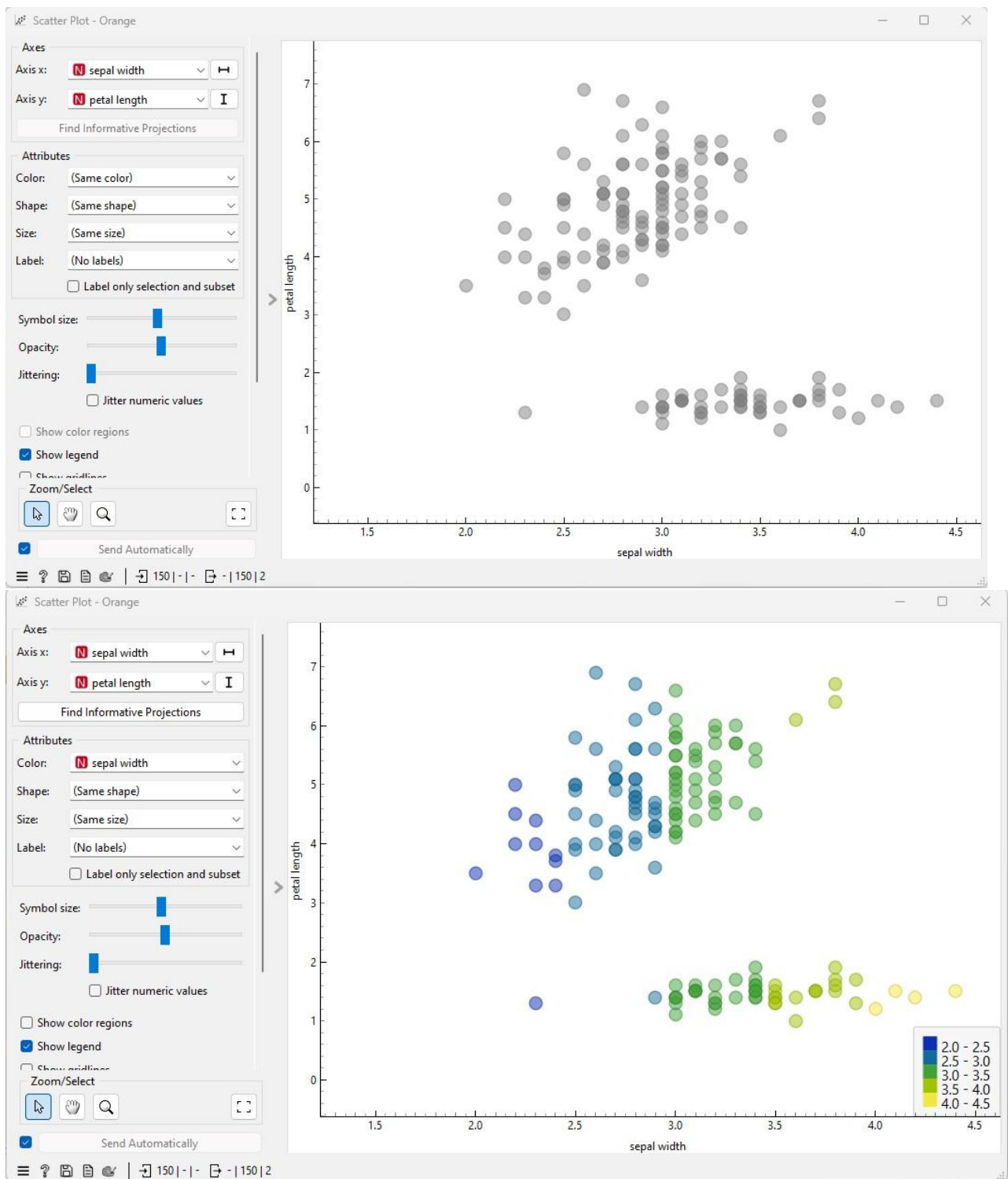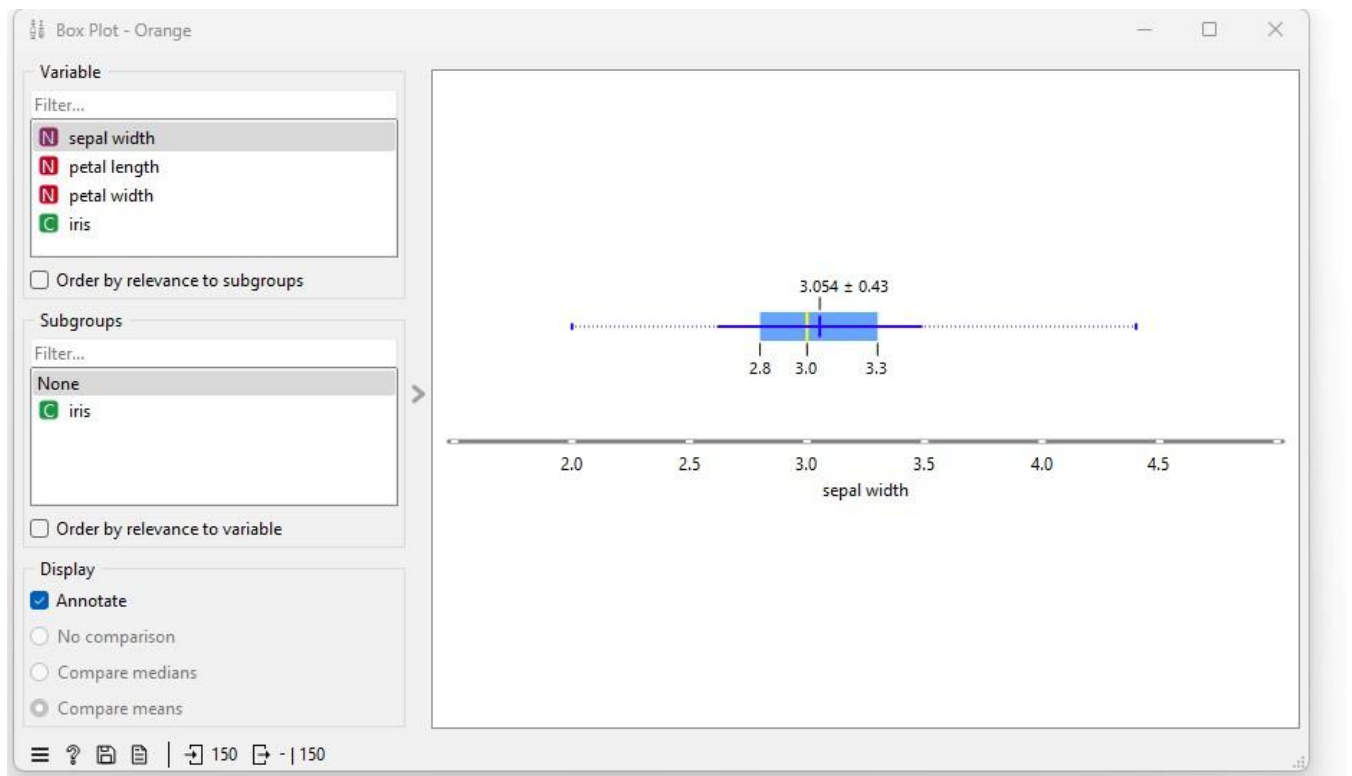
ETL Operations:
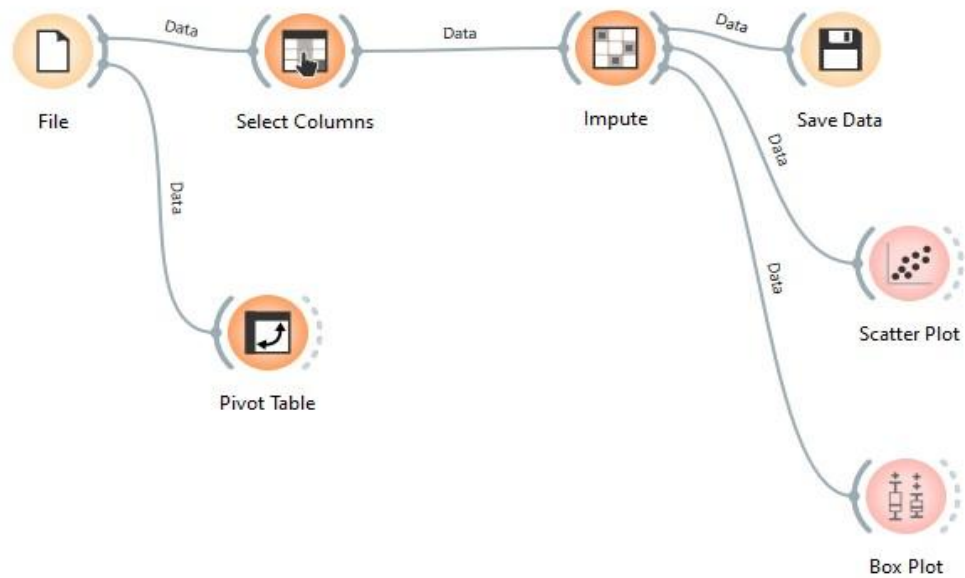
Selecting Columns:
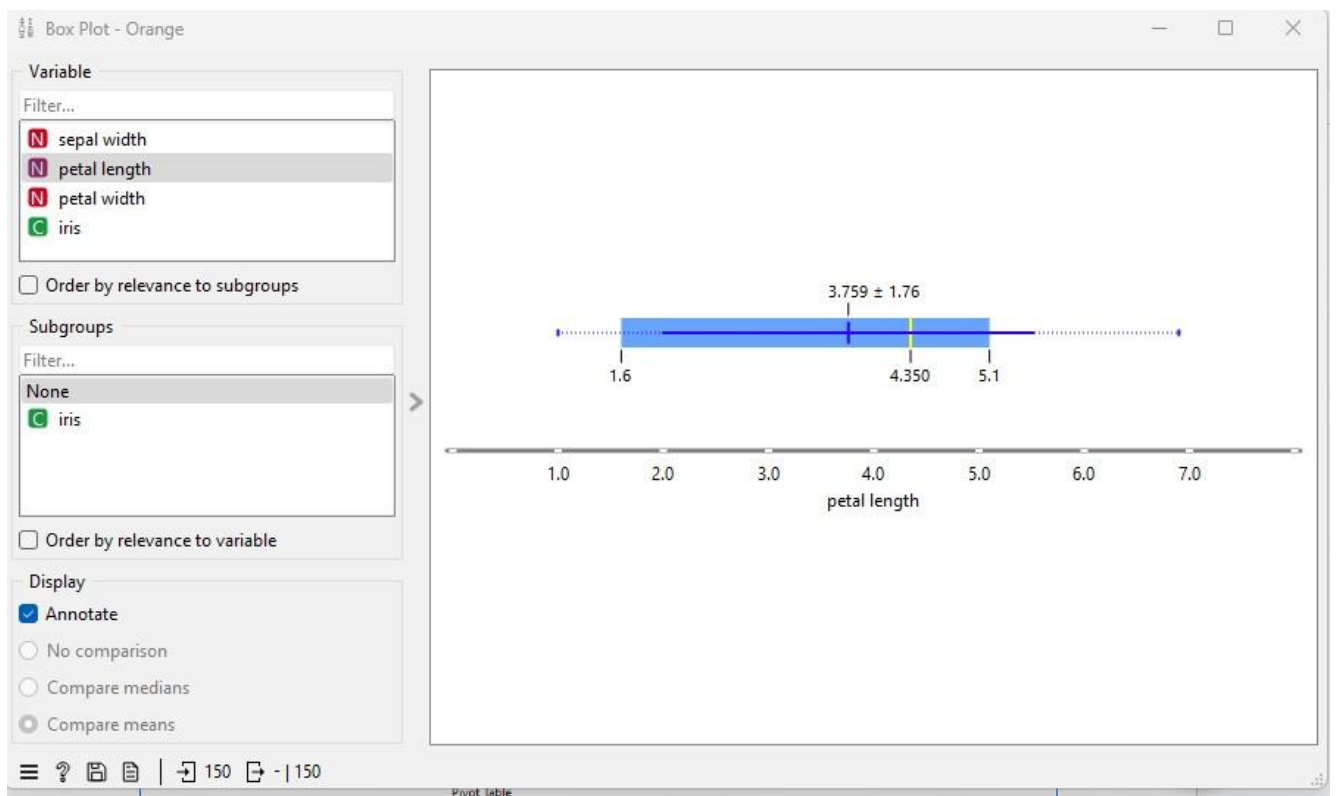
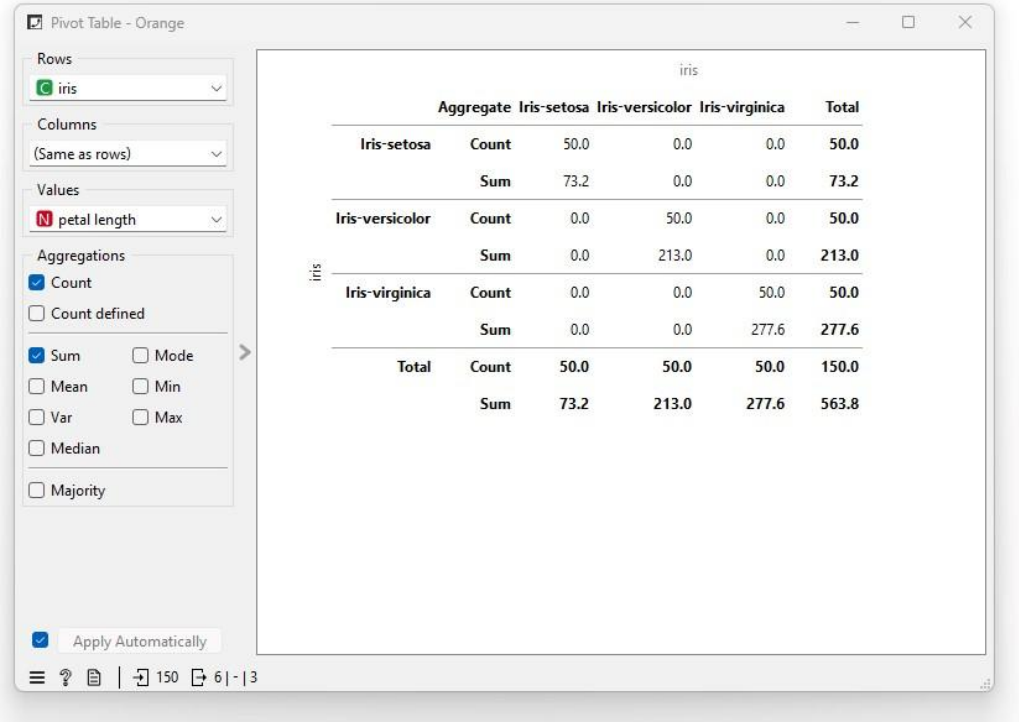Impute: Filling missing values with Average values
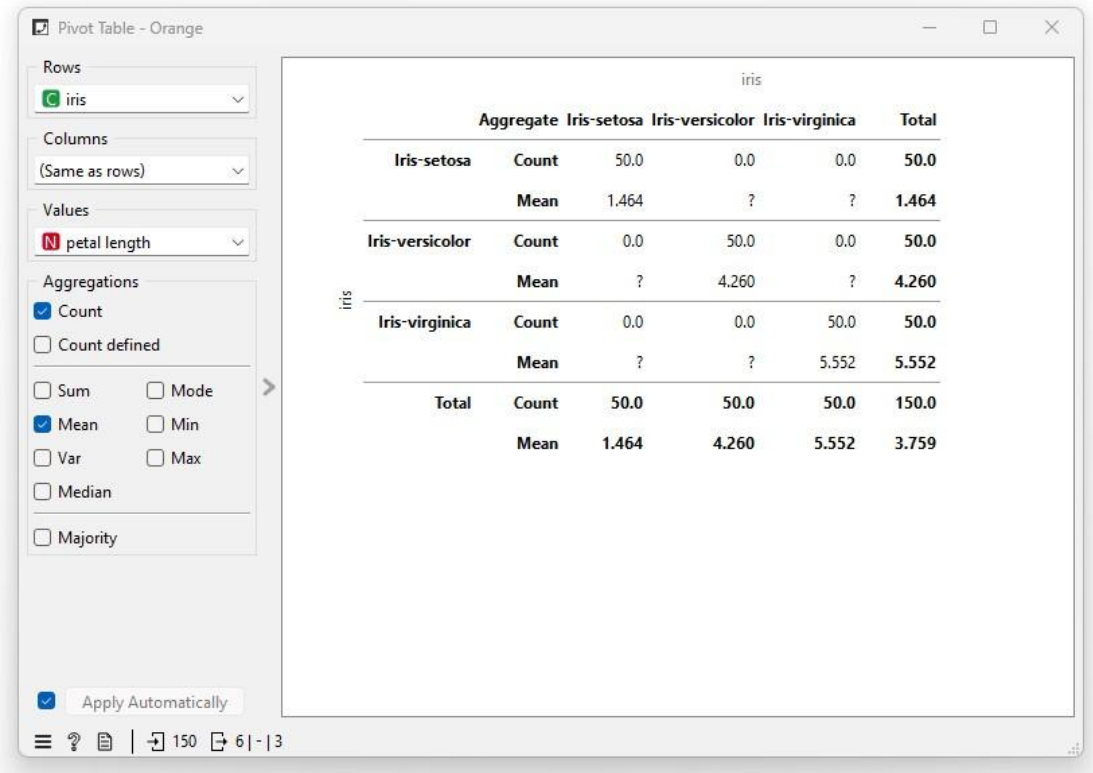

Scatterplot:

Boxplot for sepal width:

Boxplot for petal length:

Pivot Table:
Sum:

The Pivot Table - Orange window shows the following configuration and table:

Settings (left panel):
- Rows: iris
- Columns: (Same as rows)
- Values: petal length
- Aggregations: Count ✓, Count defined, Sum ✓, Mode, Mean, Min, Var, Max, Median, Majority
- Apply Automatically ✓

Table (iris):

	Aggregate	Iris-setosa	Iris-versicolor	Iris-virginica	Total
Iris-setosa	Count	50.0	0.0	0.0	50.0
	Sum	73.2	0.0	0.0	73.2
Iris-versicolor	Count	0.0	50.0	0.0	50.0
	Sum	0.0	213.0	0.0	213.0
Iris-virginica	Count	0.0	0.0	50.0	50.0
	Sum	0.0	0.0	277.6	277.6
Total	Count	50.0	50.0	50.0	150.0
	Sum	73.2	213.0	277.6	563.8

Mean:

The Pivot Table - Orange window shows the following configuration and table:

Settings (left panel):
- Rows: iris
- Columns: (Same as rows)
- Values: petal length
- Aggregations: Count ✓, Count defined, Sum, Mode, Mean ✓, Min, Var, Max, Median, Majority
- Apply Automatically ✓

Table (iris):

	Aggregate	Iris-setosa	Iris-versicolor	Iris-virginica	Total
Iris-setosa	Count	50.0	0.0	0.0	50.0
	Mean	1.464	?	?	1.464
Iris-versicolor	Count	0.0	50.0	0.0	50.0
	Mean	?	4.260	?	4.260
Iris-virginica	Count	0.0	0.0	50.0	50.0
	Mean	?	?	5.552	5.552
Total	Count	50.0	50.0	50.0	150.0
	Mean	1.464	4.260	5.552	3.759

Min and Max:

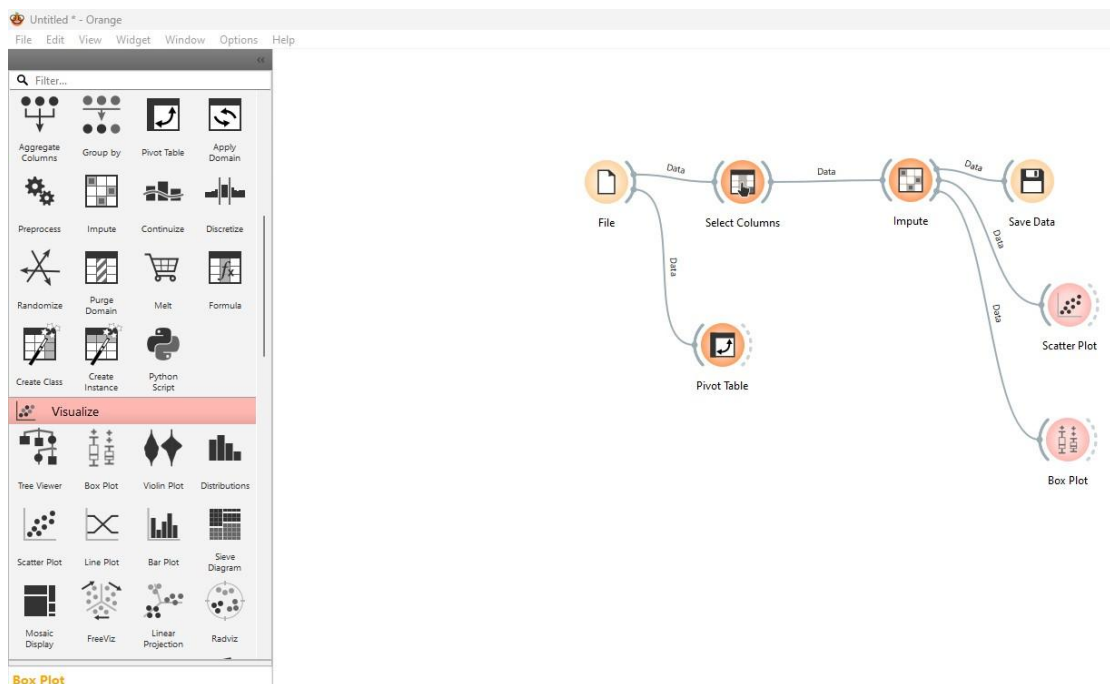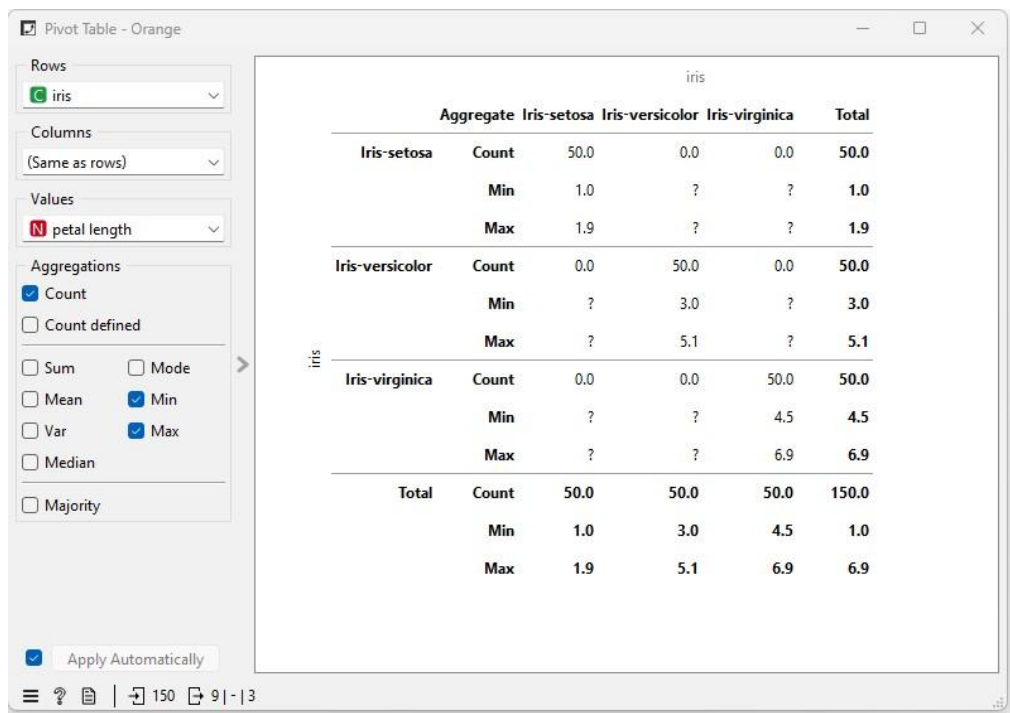|		Aggregate	Iris-setosa	Iris-versicolor	Iris-virginica	Total
Iris-setosa	Count		50.0	0.0	0.0	50.0
	Min		1.0	?	?	1.0
	Max		1.9	?	?	1.9
Iris-versicolor	Count		0.0	50.0	0.0	50.0
	Min		?	3.0	?	3.0
	Max		?	5.1	?	5.1
Iris-virginica	Count		0.0	0.0	50.0	50.0
	Min		?	?	4.5	4.5
	Max		?	?	6.9	6.9
Total	Count		50.0	50.0	50.0	150.0
	Min		1.0	3.0	4.5	1.0
	Max		1.9	5.1	6.9	6.9

In conclusion, the ETL operations and cube analysis conducted on the Iris dataset using Orange offered valuable insights into data transformation and multi-dimensional analysis methods.