An Introduction to WEKA Explorer

In part from: Yizhou Sun

2008

What is WEKA?

- **W**aikato **E**nvironment for **K**nowledge **A**nalysis
  - A data mining/machine learning tool developed by Department of Computer Science, University of Waikato, New Zealand.
  - Weka is also a bird found only on the islands of New Zealand.

How does it works?

- First, you select a dataset and a Machine learning algorithm
- You can manipulate the dataset in several ways, as we will see.
- When datset is ready, you select a ML algorithm from the list, and adjust learning parameters, as we will see
- When you run a ML algorithm, the system will:
  1. Split the data set into training and testing subsets;
  2. Learn a classification function C(x) based on examples in the training set;
  3. Classify instances x in the test set based on the learned function C(x);
  4. Measure the performances by comparing the generated classifications with the "ground truth" in the test set.

Download and Install WEKA

- Website:
  http://www.cs.waikato.ac.nz/~ml/weka/index.html
- Support multiple platforms (written in java):
  - Windows, Mac OS X and Linux

Main Features

- 49 data preprocessing tools
- 76 classification/regression algorithms
- 8 clustering algorithms
- 3 algorithms for finding association rules
- 15 attribute/subset evaluators + 10 search algorithms for feature selection

Main GUI

- Three graphical user interfaces
  - "The Explorer" (exploratory data analysis)
  - "The Experimenter" (experimental environment)
  - "The KnowledgeFlow" (new process model inspired interface)
  - Simple CLI- provides users without a graphic interface option the ability to execute commands from a terminal window

Explorer

- The Explorer:
  - Preprocess data
  - Classification
  - Clustering
  - Association Rules
  - Attribute Selection
  - Data Visualization
- References and Resources

Explorer: pre-processing the data

- Data can be imported from a file in various formats: ARFF, CSV, C4.5, binary
- Data can also be read from a URL or from an SQL database (using JDBC **Java DataBase Connectivity**)
- Pre-processing tools in WEKA are called "filters"
- WEKA contains filters for:
  - Discretization, normalization, resampling, attribute selection, transforming and combining attributes, …

WEKA only deals with "flat" files

@relation heart-disease-simplified

@attribute age numeric
@attribute sex { female, male}
@attribute chest_pain_type { typ_angina, asympt, non_anginal, atyp_angina}
@attribute cholesterol numeric
@attribute exercise_induced_angina { no, yes}
@attribute class { present, not_present}

@data
63,male,typ_angina,233,no,not_present
67,male,asympt,286,yes,present
67,male,asympt,229,yes,present
38,female,non_anginal,?,no,not_present
…

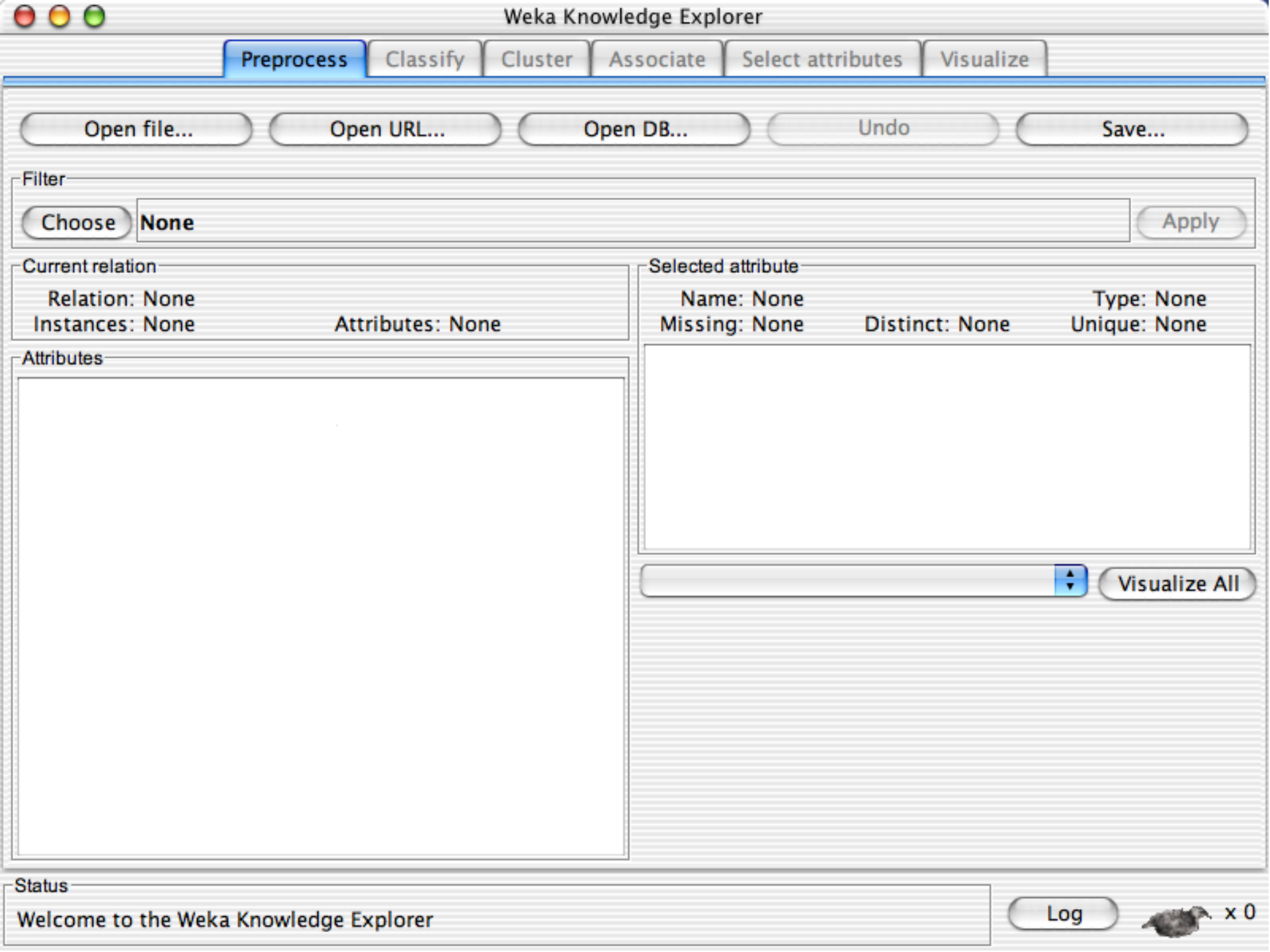
Flat file in ARFF format

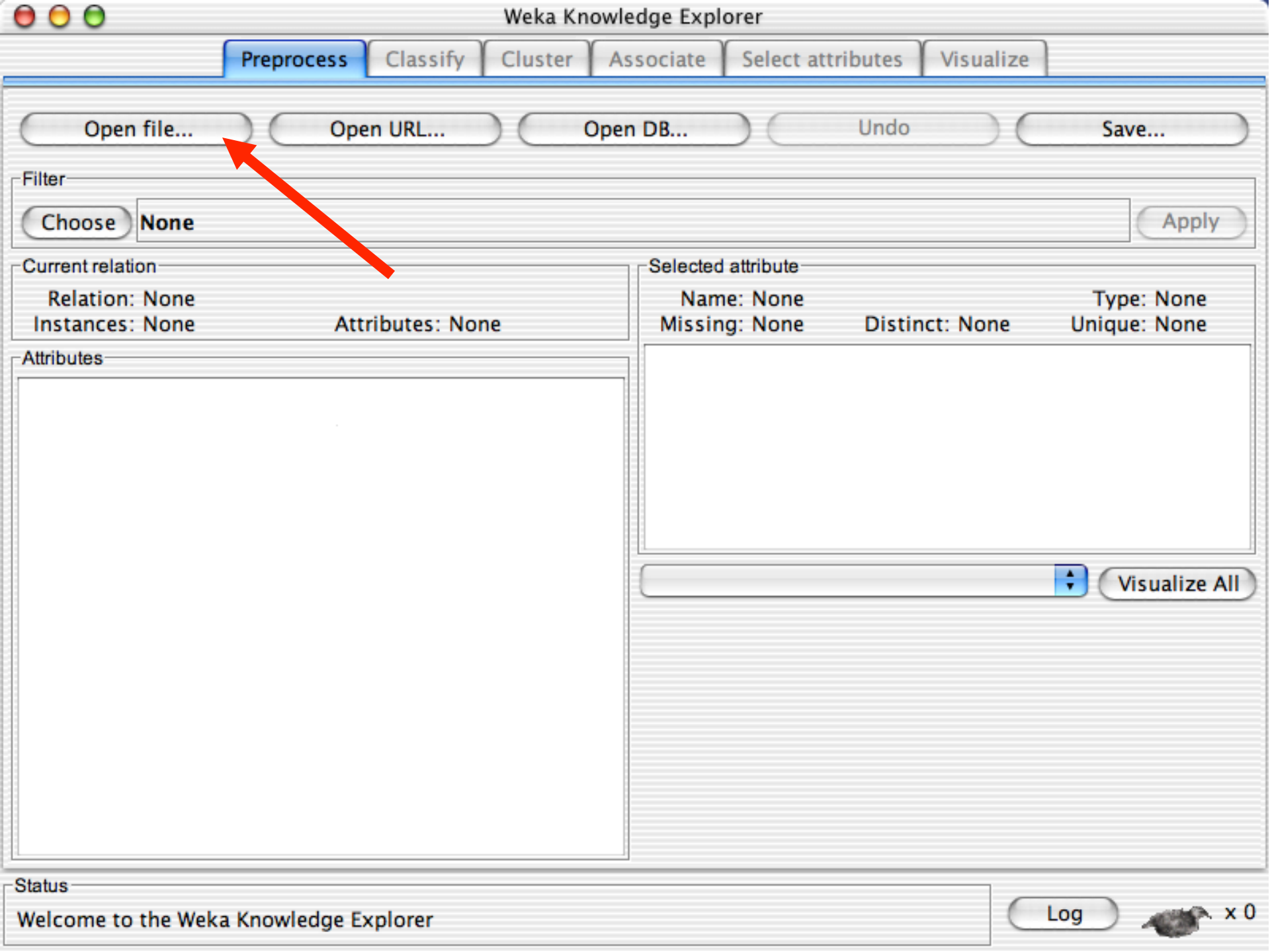
WEKA only deals with "flat" files

@relation heart-disease-simplified

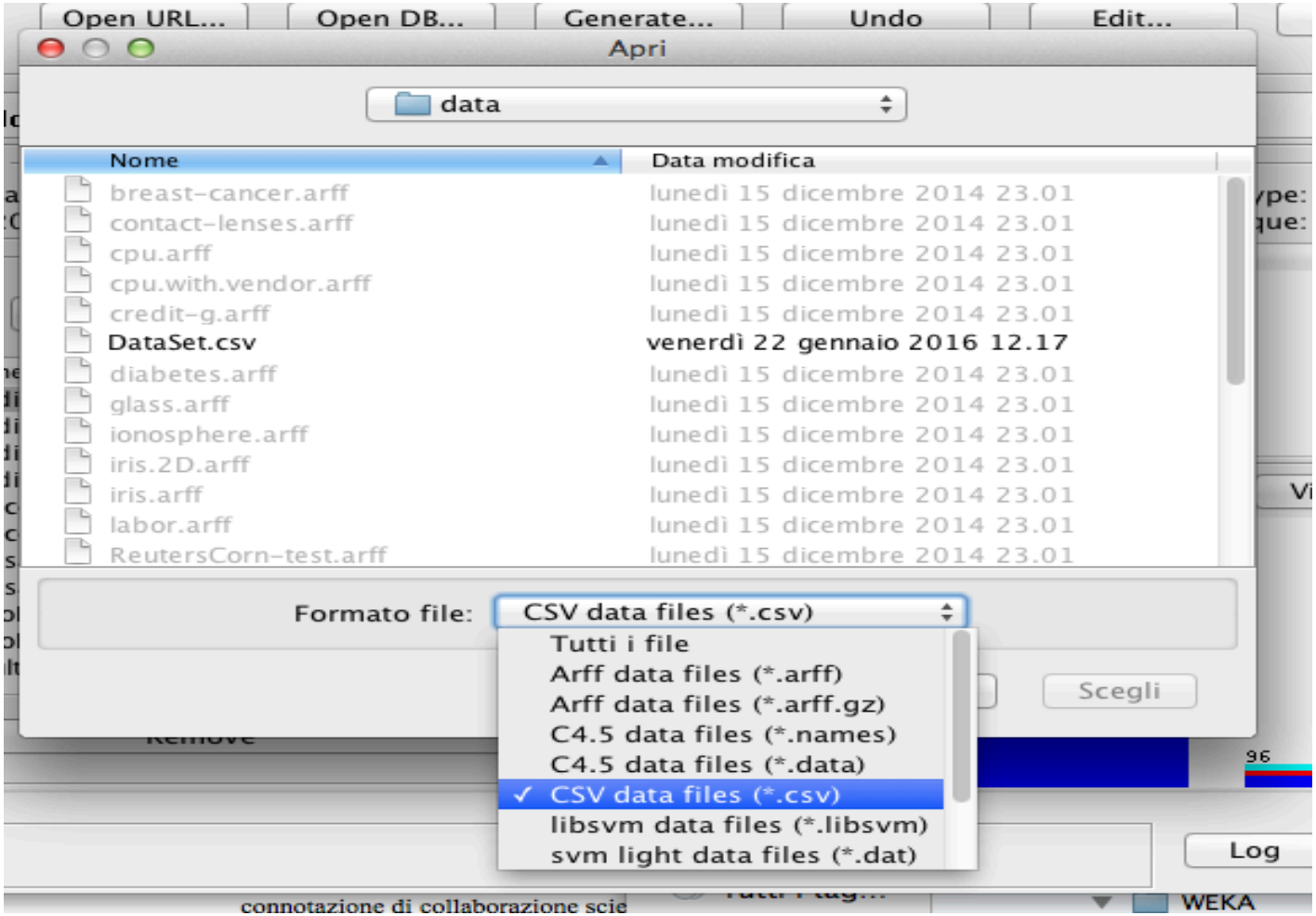
@attribute age numeric
@attribute sex { female, male}
@attribute chest_pain_type { typ_angina, asympt, non_anginal, atyp_angina}
@attribute cholesterol numeric
@attribute exercise_induced_angina { no, yes}
@attribute class { present, not_present}

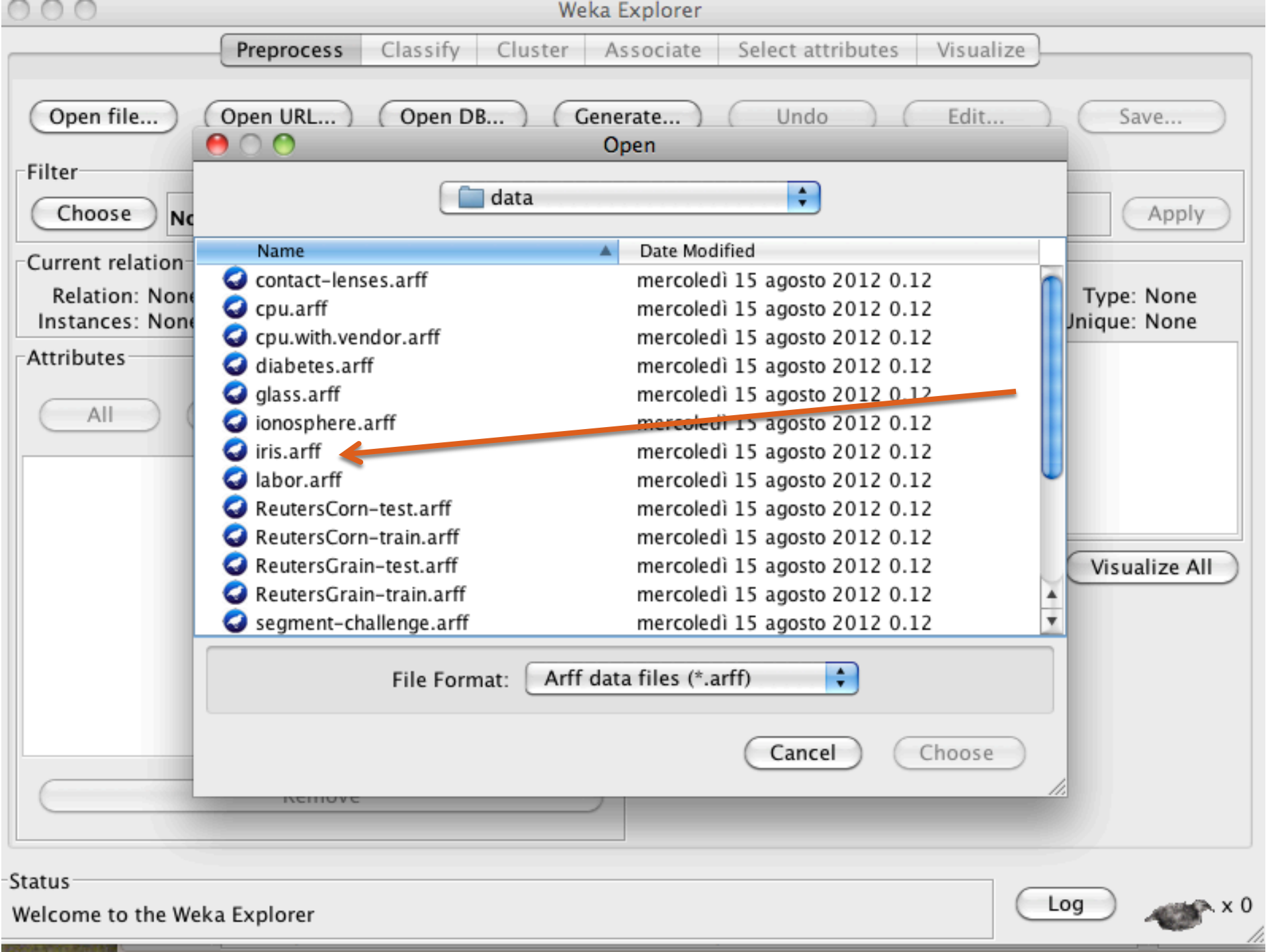
@data
63,male,typ_angina,233,no,not_present
67,male,asympt,286,yes,present
67,male,asympt,229,yes,present
38,female,non_anginal,?,no,not_present
…

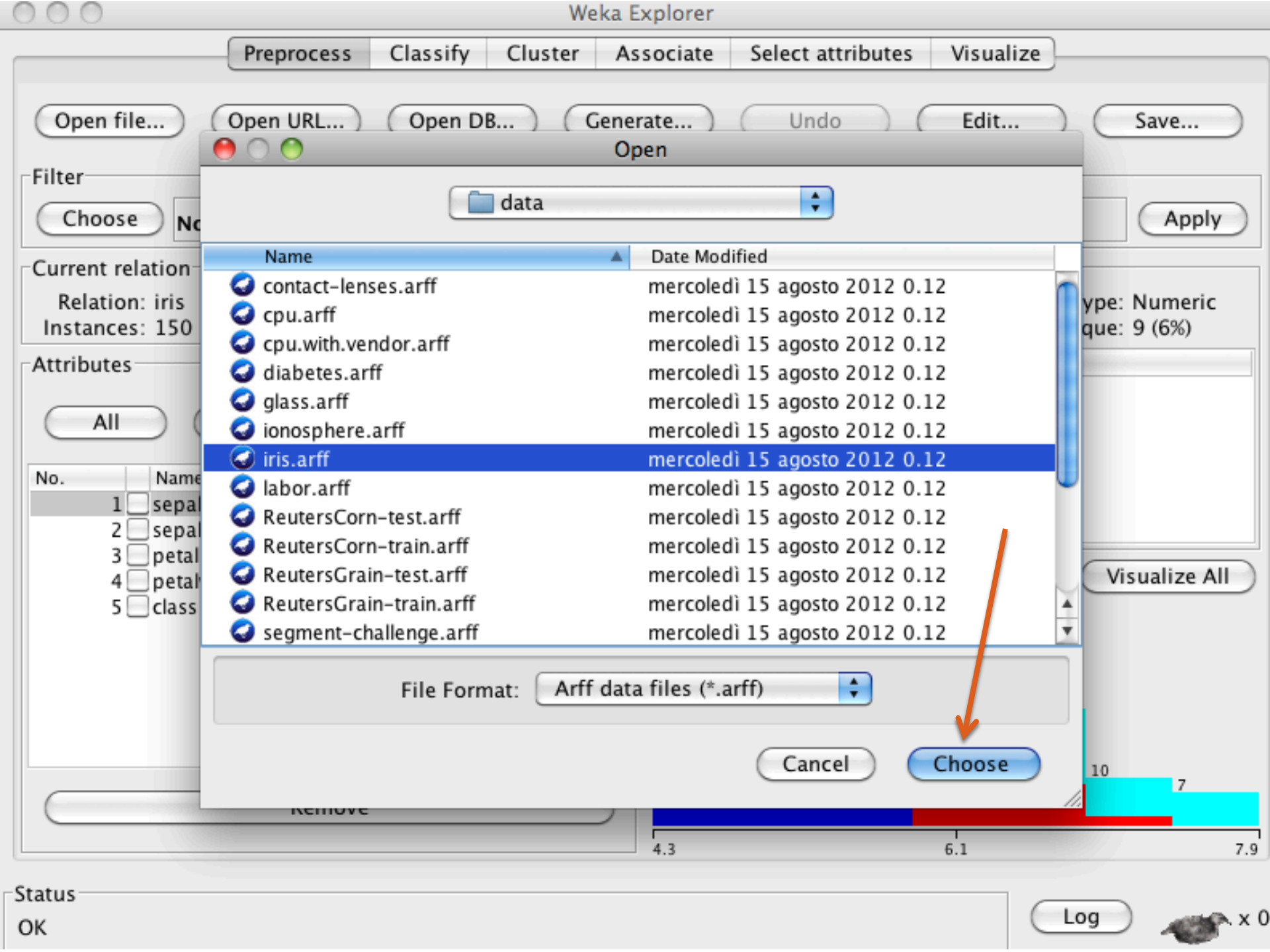

Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
Choose
None
Apply
Current relation
Relation: None
Instances: None
Attributes: None
Selected attribute
Name: None
Type: None
Missing: None
Distinct: None
Unique: None
Attributes
Visualize All
Status
Welcome to the Weka Knowledge Explorer
Log
x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
Choose
None
Apply
Current relation
Relation: None
Instances: None
Attributes: None
Selected attribute
Name: None
Type: None
Missing: None
Distinct: None
Unique: None
Attributes
Visualize All
Status
Welcome to the Weka Knowledge Explorer
Log
x 0

You can either open arff file or convert from other formats

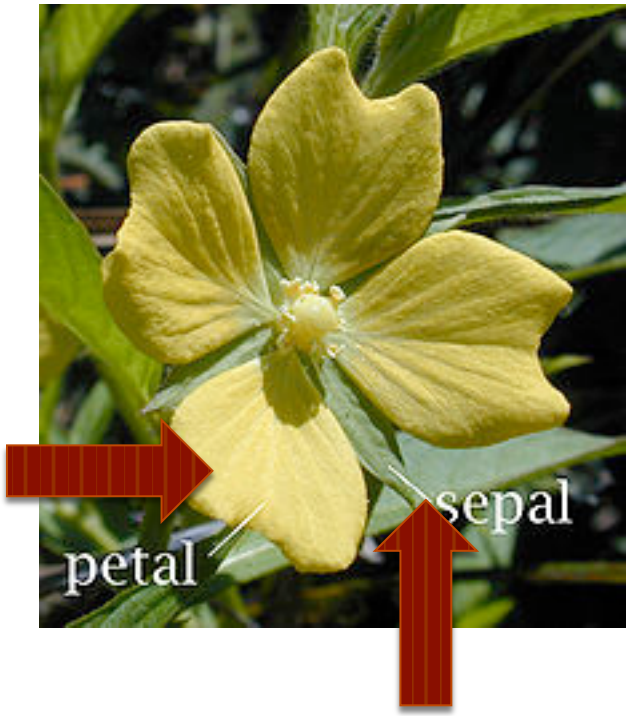



Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Generate...
Undo
Edit...
Save...
Filter
Choose
Apply
Current relation
Relation: None
Instances: None
Type: None
Unique: None
Attributes
All
Visualize All
Open
data
Name
Date Modified
contact-lenses.arff
cpu.arff
cpu.with.vendor.arff
diabetes.arff
glass.arff
ionosphere.arff
iris.arff
labor.arff
ReutersCorn-test.arff
ReutersCorn-train.arff
ReutersGrain-test.arff
ReutersGrain-train.arff
segment-challenge.arff
mercoledì 15 agosto 2012 0.12
File Format:
Arff data files (*.arff)
Cancel
Choose
Status
Welcome to the Weka Explorer
Log
x 0


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Generate...
Undo
Edit...
Save...
Filter
Choose
Apply
Current relation
Relation: iris
Instances: 150
Attributes
All
No.
Name
1 sepal
2 sepal
3 petal
4 petal
5 class
Visualize All
Remove
Open
data
Name
Date Modified
contact-lenses.arff mercoledì 15 agosto 2012 0.12
cpu.arff mercoledì 15 agosto 2012 0.12
cpu.with.vendor.arff mercoledì 15 agosto 2012 0.12
diabetes.arff mercoledì 15 agosto 2012 0.12
glass.arff mercoledì 15 agosto 2012 0.12
ionosphere.arff mercoledì 15 agosto 2012 0.12
iris.arff mercoledì 15 agosto 2012 0.12
labor.arff mercoledì 15 agosto 2012 0.12
ReutersCorn-test.arff mercoledì 15 agosto 2012 0.12
ReutersCorn-train.arff mercoledì 15 agosto 2012 0.12
ReutersGrain-test.arff mercoledì 15 agosto 2012 0.12
ReutersGrain-train.arff mercoledì 15 agosto 2012 0.12
segment-challenge.arff mercoledì 15 agosto 2012 0.12
File Format:
Arff data files (*.arff)
Cancel
Choose
10
7
4.3
6.1
7.9
Status
OK
Log
x 0

IRIS dataset

- 150 instances of IRIS (a flower)
- 5 attributes, one is the classification c(x)
- 3 classes: iris setosa, iris versicolor, iris virginica



Weka Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open file... | Open URL... | Open DB... | Generate... | Undo | Edit... | Save...

Filter

Choose | None | Apply

For any selected attribute you can get statistics

Current relation

Relation: iris

Instances: 150 Attributes: 5

Attributes

All | None | Invert | Pattern

No.	Name
1	sepallength
2	sepalwidth
3	petallength
4	petalwidth
5	class

Remove

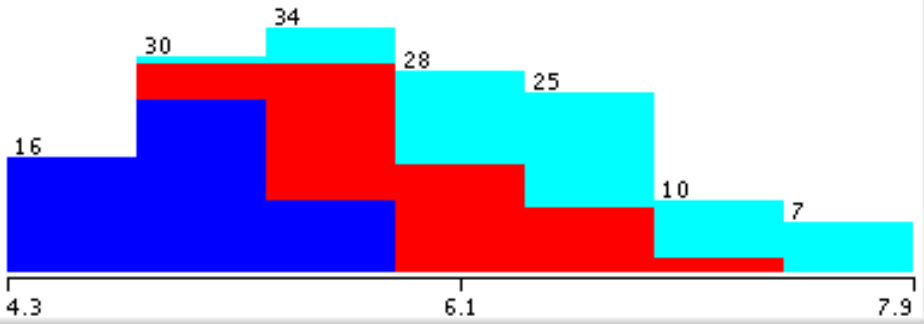
Selected attribute

Name: sepallength Type: Numeric

Missing: 0 (0%) Distinct: 35 Unique: 9 (6%)

Statistic	Value
Minimum	4.3
Maximum	7.9
Mean	5.843
StdDev	0.828

Class: class (Nom) | Visualize All



Status

OK

Log

x 0

Attribute data

- Min, max and average value of attributes
- distribution of values :number of items for which:

$$a_i = v_j \mid a_i \in A, v_j \in V$$

- class: distribution of attribute values in the classes
- The class (e.g. C(x), the classification function to be learned) is by default THE LAST ATTRIBUTE of the list.

Weka Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open file... | Open URL... | Open DB... | Generate... | Undo | Edit... | Save...

Filter

Choose | None | Apply

Current relation

Relation: iris
Instances: 150 Attributes: 5

Attributes

All | None | Invert | Pattern

No.		Name
1	☐	sepallength
2	☐	sepalwidth
3	☐	petallength
4	☐	petalwidth
5	☐	class

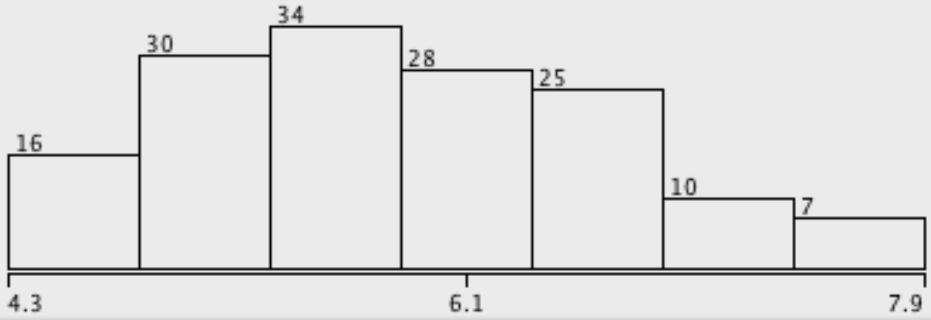
Remove

Selected attribute

Name: sepallength Type: Numeric
Missing: 0 (0%) Distinct: 35 Unique: 9 (6%)

Statistic	Value
Minimum	4.3
Maximum	7.9
Mean	5.843
StdDev	0.828

Class: sepallength (Num) | Visualize All



Status

OK

Log | x 0

Weka Knowledge Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open file... | Open URL... | Open DB... | Undo | Save...

Filter

Choose **None** | Apply

Current relation

Relation: iris
Instances: 150 Attributes: 5

Attributes

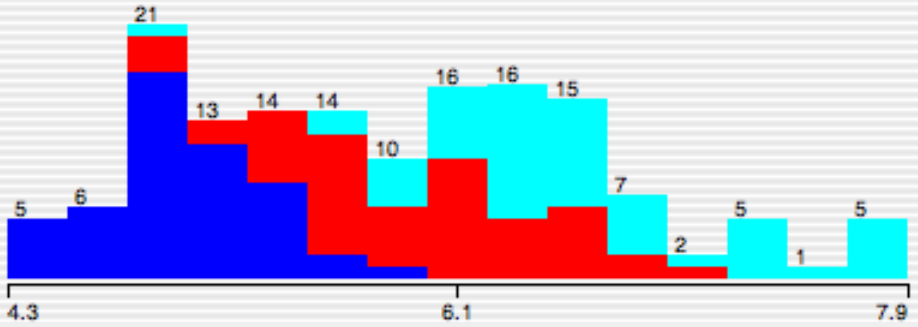
No.	Name
1	sepallength
2	sepalwidth
3	petallength
4	petalwidth
5	class

Selected attribute

Name: sepallength Type: Numeric
Missing: 0 (0%) Distinct: 35 Unique: 9 (6%)

Statistic	Value
Minimum	4.3
Maximum	7.9
Mean	5.843
StdDev	0.828

Colour: class (Nom) | Visualize All



Status

OK

Log x 0

Weka Knowledge Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open file... | Open URL... | Open DB... | Undo | Save...

Filter

Choose **None** | Apply

Current relation

Relation: iris
Instances: 150 Attributes: 5

Attributes

No.	Name
1	sepallength
2	sepalwidth
3	petallength
4	petalwidth
5	class

Selected attribute

Name: class Type: Nominal
Missing: 0 (0%) Distinct: 3 Unique: 0 (0%)

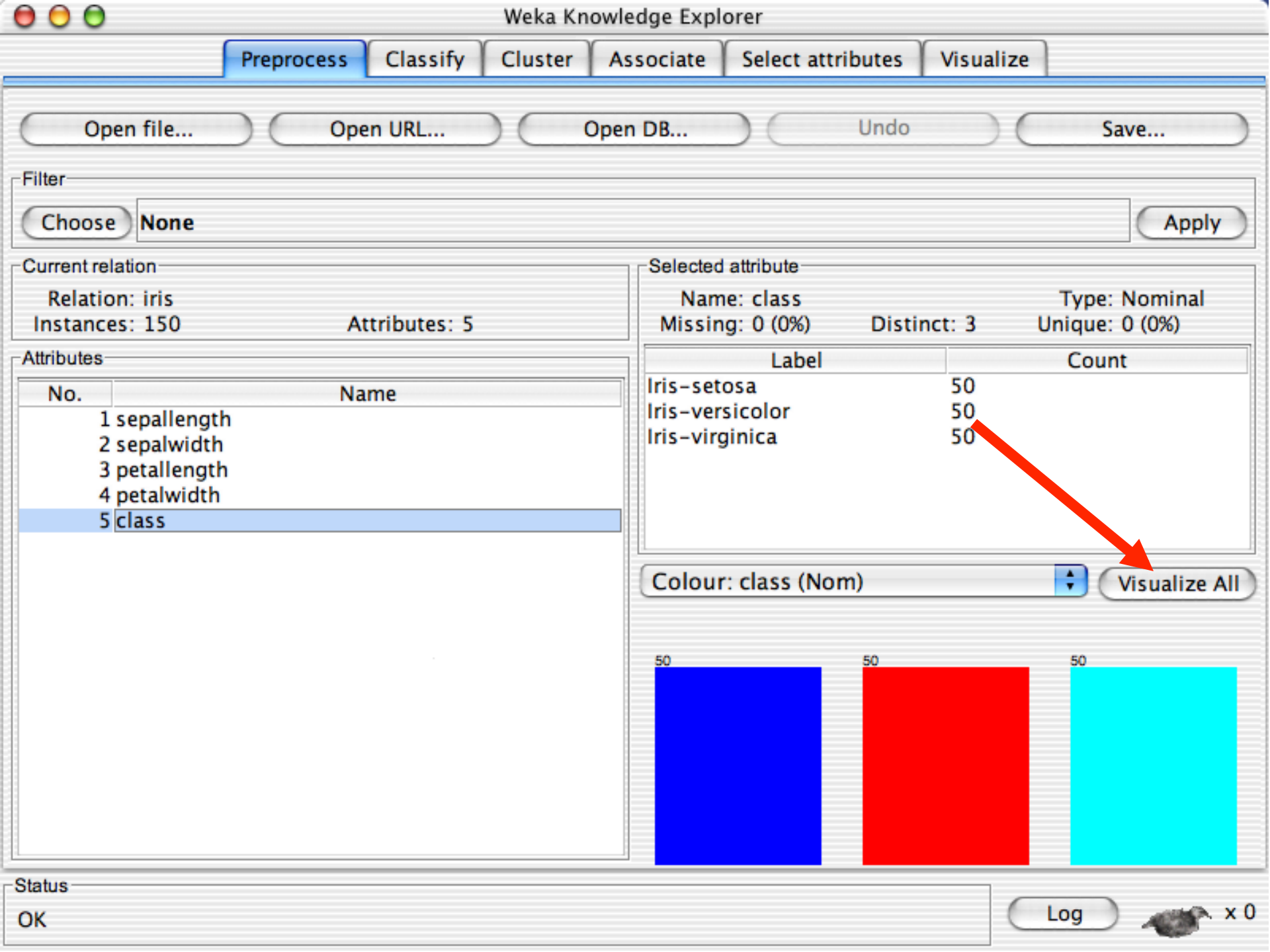
Label	Count
Iris-setosa	50
Iris-versicolor	50
Iris-virginica	50

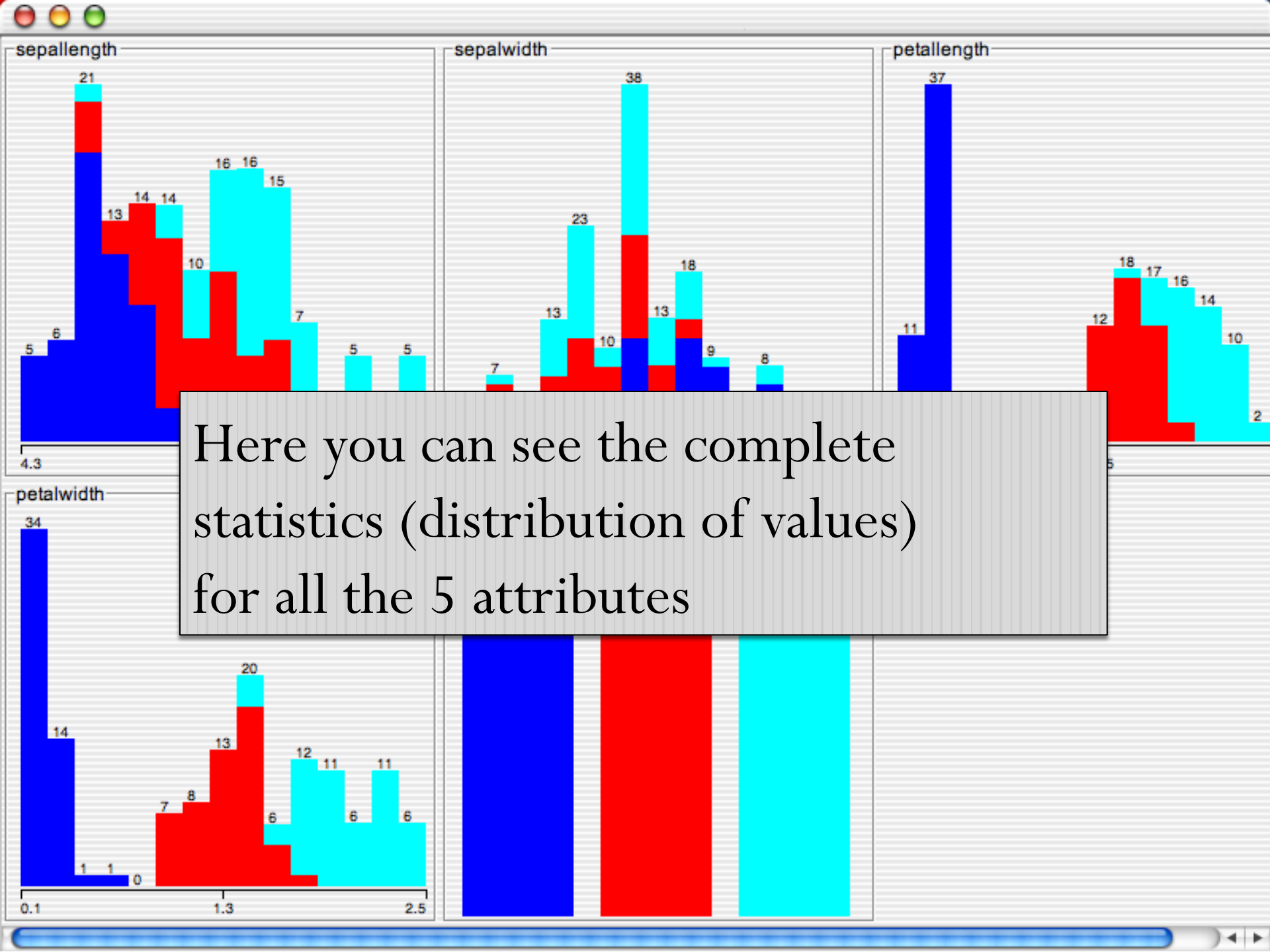
Colour: class (Nom) | Visualize All

Status

OK

Log x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
Choose
None
Apply
Current relation
Relation: iris
Instances: 150
Attributes: 5
Attributes
No.
Name
1 sepallength
2 sepalwidth
3 petallength
4 petalwidth
5 class
Selected attribute
Name: class
Type: Nominal
Missing: 0 (0%)
Distinct: 3
Unique: 0 (0%)
Label
Count
Iris-setosa
50
Iris-versicolor
50
Iris-virginica
50
Colour: class (Nom)
Visualize All
50
50
50
Status
OK
Log
x 0


sepallength
21
16 16
14 14
13
15
10
7
6
5
5
5
4.3
sepalwidth
38
23
18
13
13
10
9
8
7
petallength
37
18
17
16
14
12
11
10
2
petalwidth
34
14
20
13
12
11
11
8
7
6
6
6
1
1
0
0.1
1.3
2.5
Here you can see the complete statistics (distribution of values) for all the 5 attributes

Filtering attributes

- Once the initial data has been selected and loaded the user can select options for **refining the experimental data**.
- The options in the preprocess window include selection of **optional filters** to apply and the user can **select or remove different attributes** of the data set as necessary to identify specific information (or even write a regex in Perl).
- The user can modify the attribute selection and change the relationship among the different attributes by deselecting different choices from the original data set.
- There are many different filtering options available within the preprocessing window and the user can select the different options based on need and type of data present.

Weka Knowledge Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open file... | Open URL... | Open DB... | Undo | Save...

Filter

Choose **None** | Apply

Current relation

Relation: iris
Instances: 150 Attributes: 5

Attributes

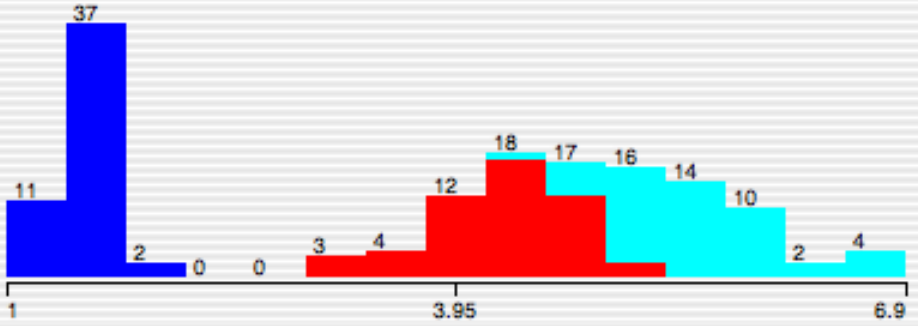
No.	Name
1	sepallength
2	sepalwidth
3	petallength
4	petalwidth
5	class

Selected attribute

Name: petallength Type: Numeric
Missing: 0 (0%) Distinct: 43 Unique: 10 (7%)

Statistic	Value
Minimum	1
Maximum	6.9
Mean	3.759
StdDev	1.764

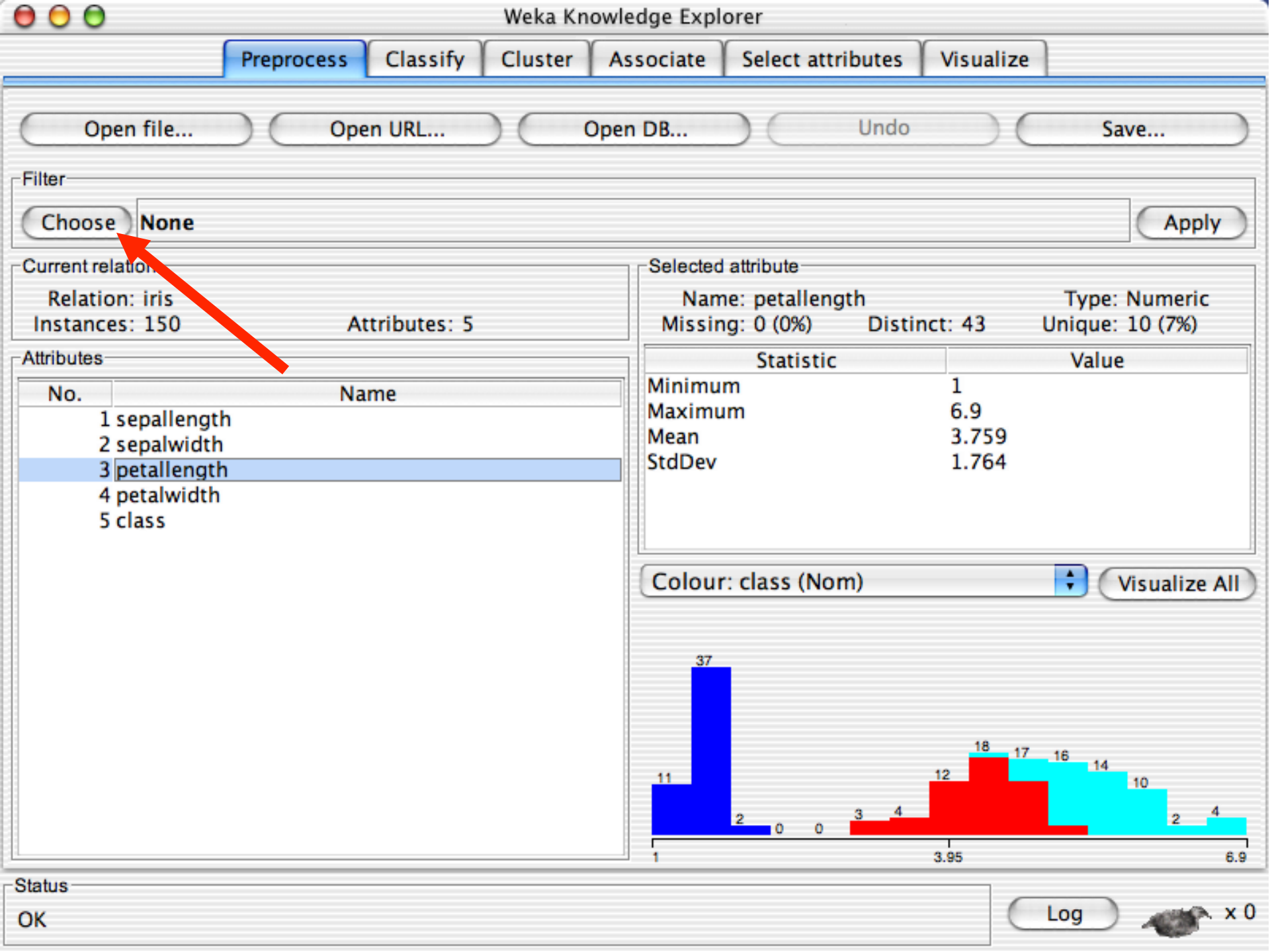
Colour: class (Nom) | Visualize All

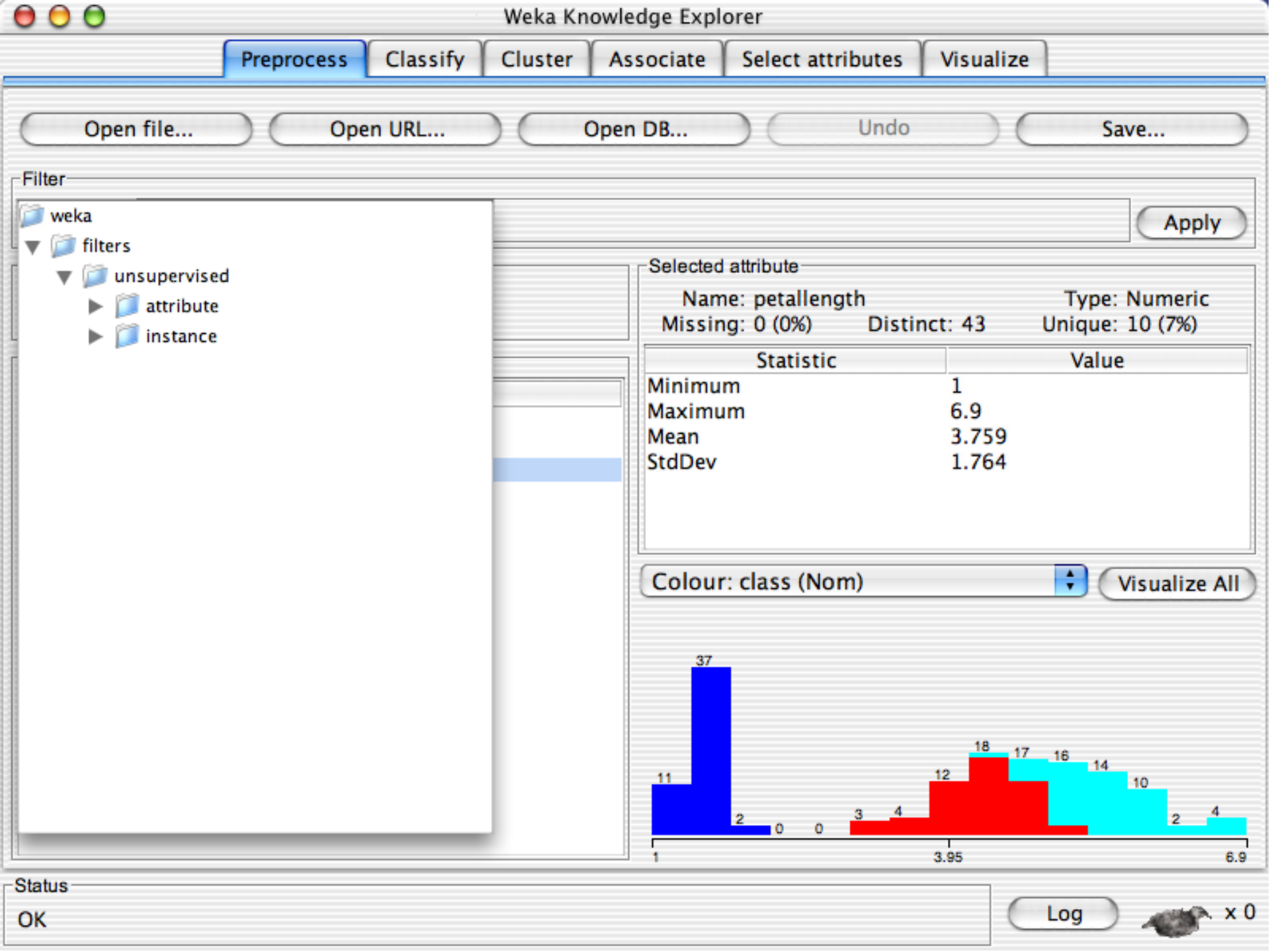


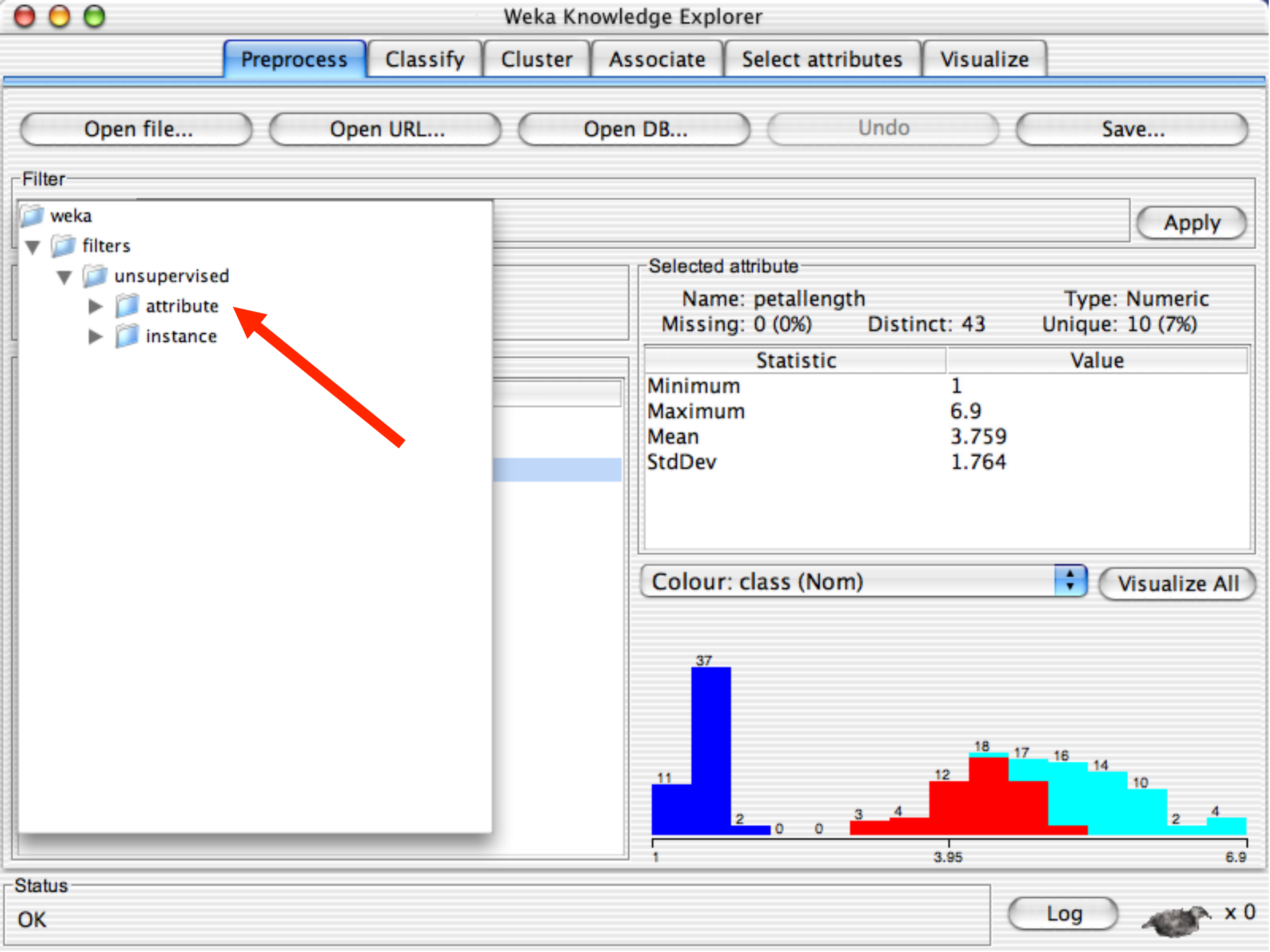
Status

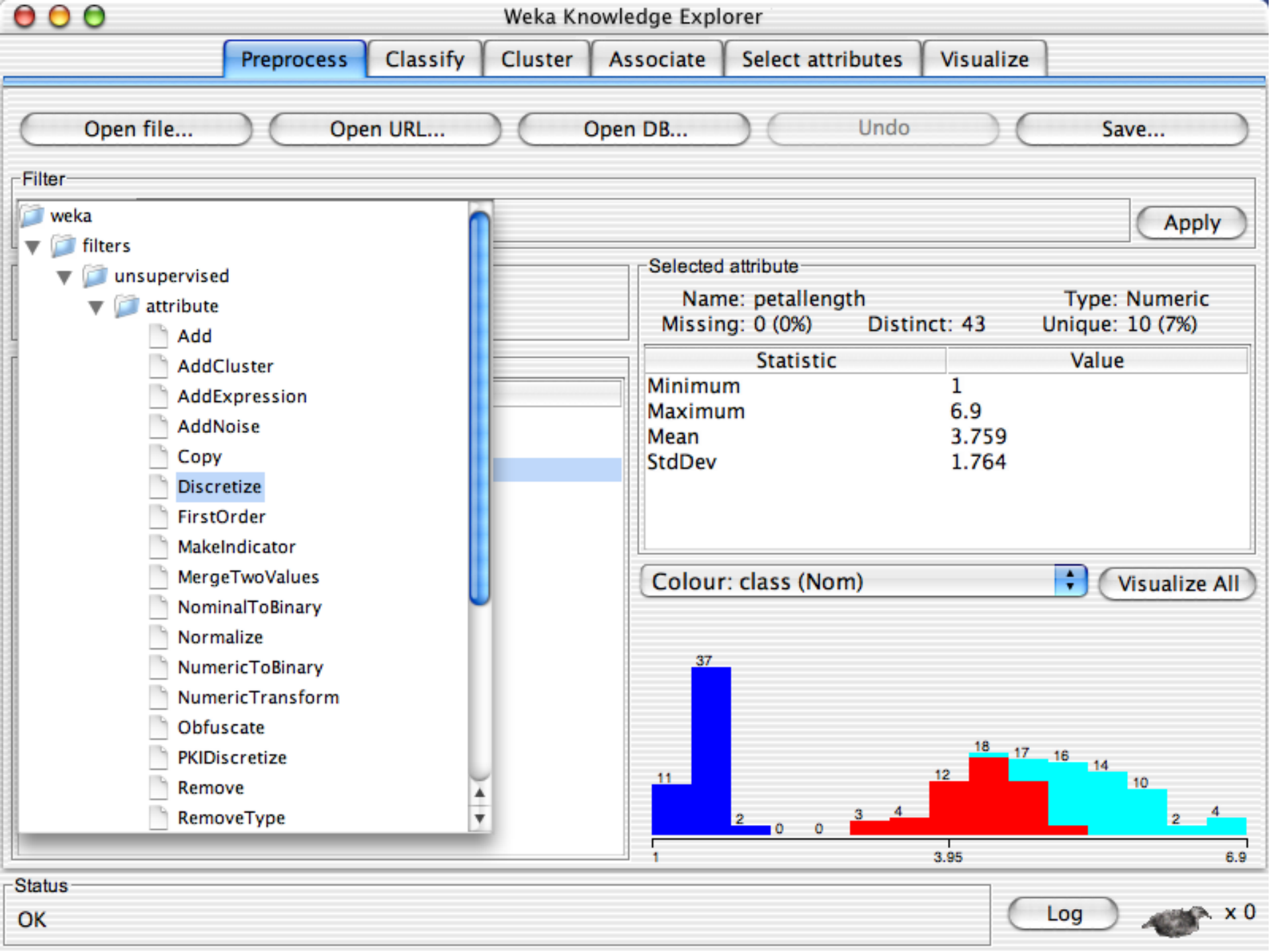
OK

Log | x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
Choose
None
Apply
Current relation
Relation: iris
Instances: 150
Attributes: 5
Attributes
No.
Name
1 sepallength
2 sepalwidth
3 petallength
4 petalwidth
5 class
Selected attribute
Name: petallength
Type: Numeric
Missing: 0 (0%)
Distinct: 43
Unique: 10 (7%)
Statistic
Value
Minimum
1
Maximum
6.9
Mean
3.759
StdDev
1.764
Colour: class (Nom)
Visualize All
37
11
2
0
0
3
4
12
18
17
16
14
10
2
4
1
3.95
6.9
Status
OK
Log
x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
weka
filters
unsupervised
attribute
instance
Apply
Selected attribute
Name: petallength
Type: Numeric
Missing: 0 (0%)
Distinct: 43
Unique: 10 (7%)
Statistic
Value
Minimum
1
Maximum
6.9
Mean
3.759
StdDev
1.764
Colour: class (Nom)
Visualize All
37
11
2
0
0
3
4
12
18
17
16
14
10
2
4
1
3.95
6.9
Status
OK
Log
x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
Apply
weka
filters
unsupervised
attribute
instance
Selected attribute
Name: petallength
Type: Numeric
Missing: 0 (0%)
Distinct: 43
Unique: 10 (7%)
Statistic
Value
Minimum
1
Maximum
6.9
Mean
3.759
StdDev
1.764
Colour: class (Nom)
Visualize All
37
11
2
0
0
3
4
12
18
17
16
14
10
2
4
1
3.95
6.9
Status
OK
Log
x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
weka
filters
unsupervised
attribute
Add
AddCluster
AddExpression
AddNoise
Copy
Discretize
FirstOrder
MakeIndicator
MergeTwoValues
NominalToBinary
Normalize
NumericToBinary
NumericTransform
Obfuscate
PKIDiscretize
Remove
RemoveType
Apply
Selected attribute
Name: petallength
Type: Numeric
Missing: 0 (0%)
Distinct: 43
Unique: 10 (7%)
Statistic
Value
Minimum
1
Maximum
6.9
Mean
3.759
StdDev
1.764
Colour: class (Nom)
Visualize All
37
11
2
0
0
3
4
12
18
17
16
14
10
2
4
1
3.95
6.9
Status
OK
Log
x 0

Weka Knowledge Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open file... | Open URL... | Open DB... | Undo | Save...

Filter

Choose **Discretize** -B 10 -R first-last | Apply

Current relation

Relation: iris
Instances: 150 Attributes: 5

Attributes

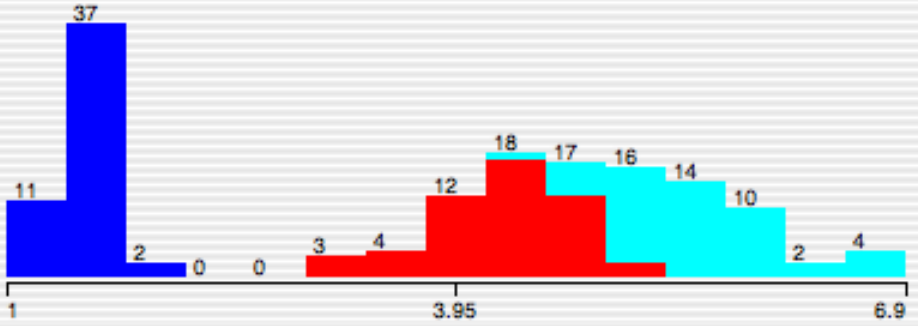
No.	Name
1	sepallength
2	sepalwidth
3	petallength
4	petalwidth
5	class

Selected attribute

Name: petallength Type: Numeric
Missing: 0 (0%) Distinct: 43 Unique: 10 (7%)

Statistic	Value
Minimum	1
Maximum	6.9
Mean	3.759
StdDev	1.764

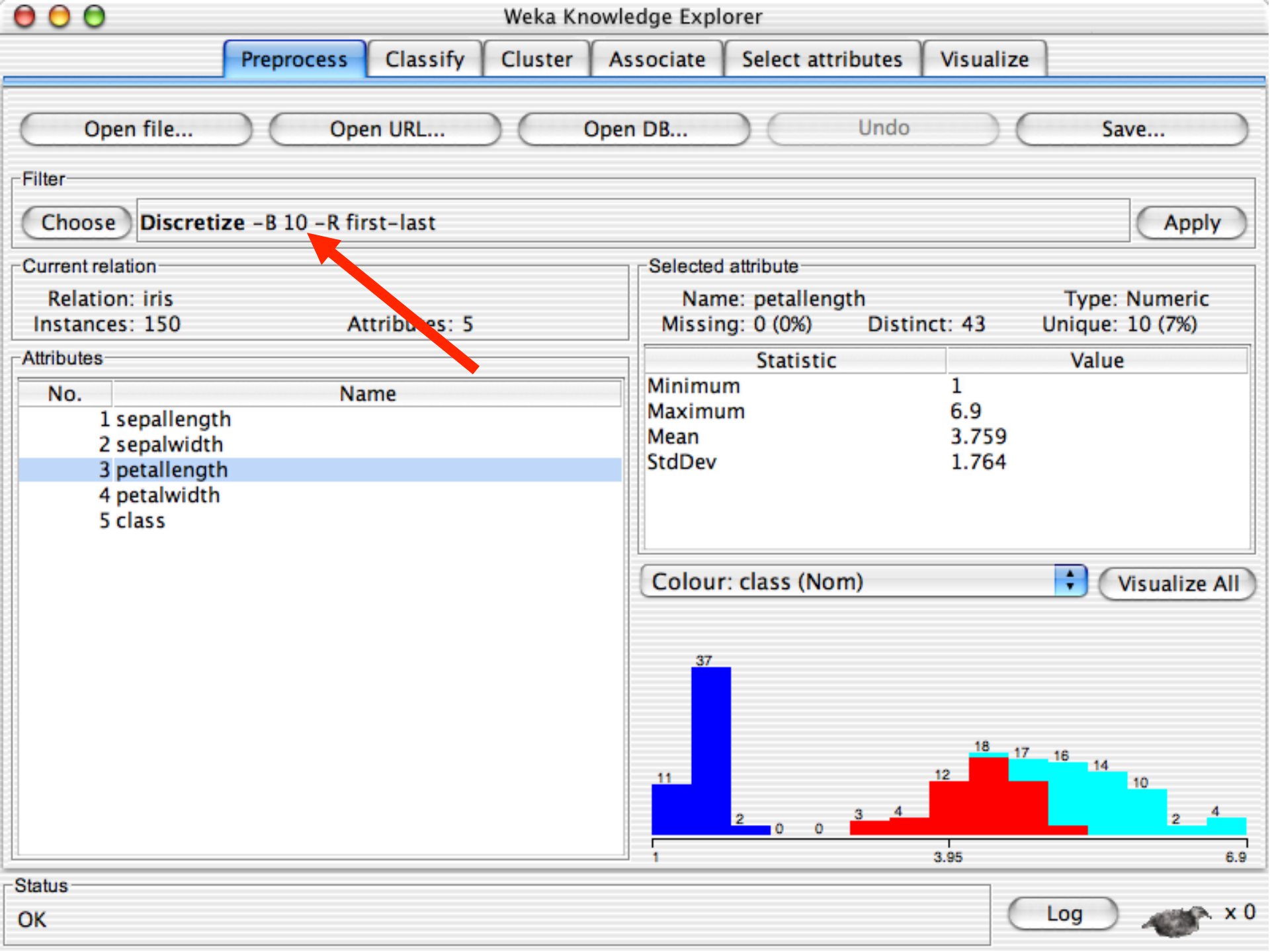
Colour: class (Nom) | Visualize All

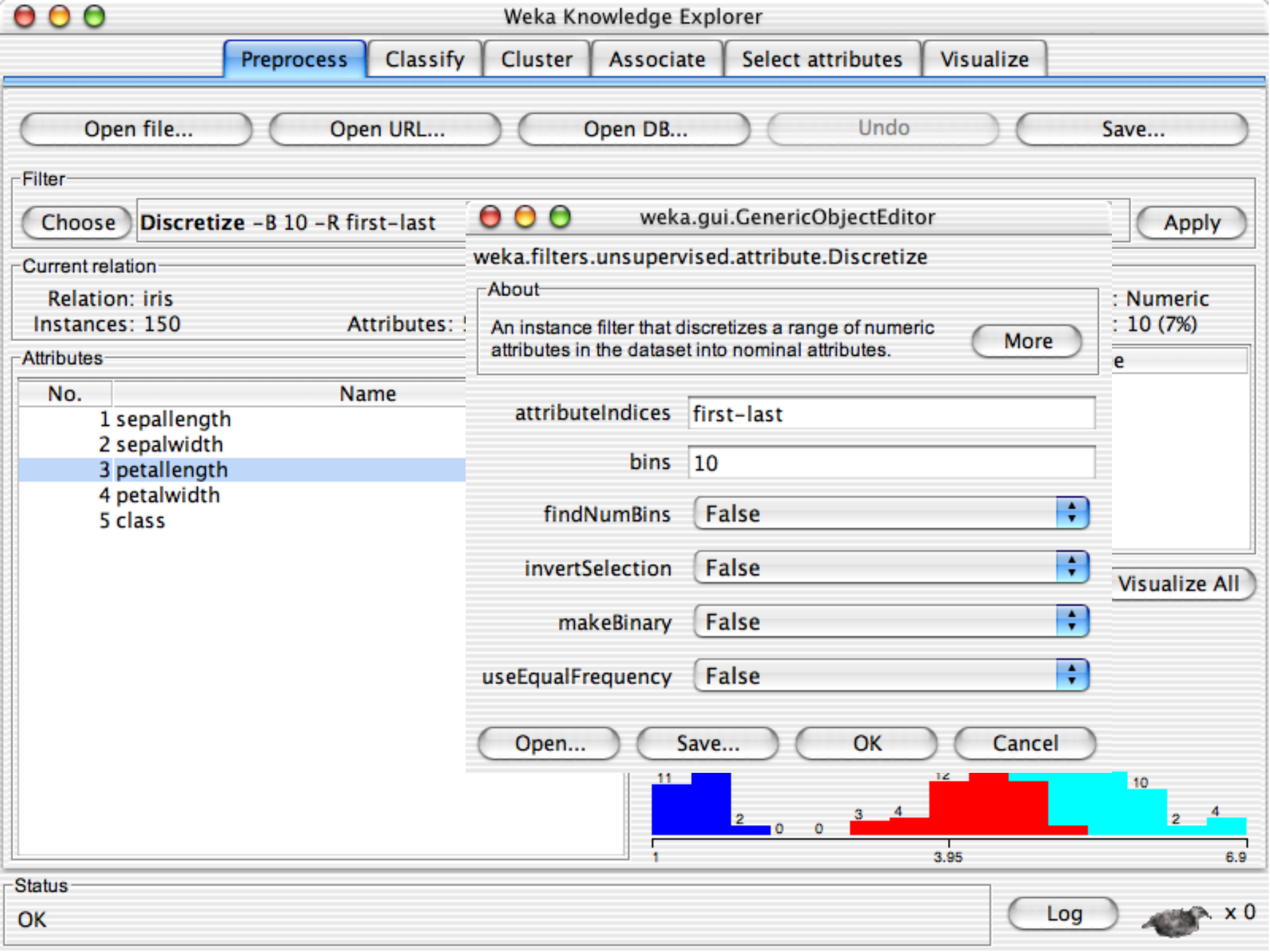


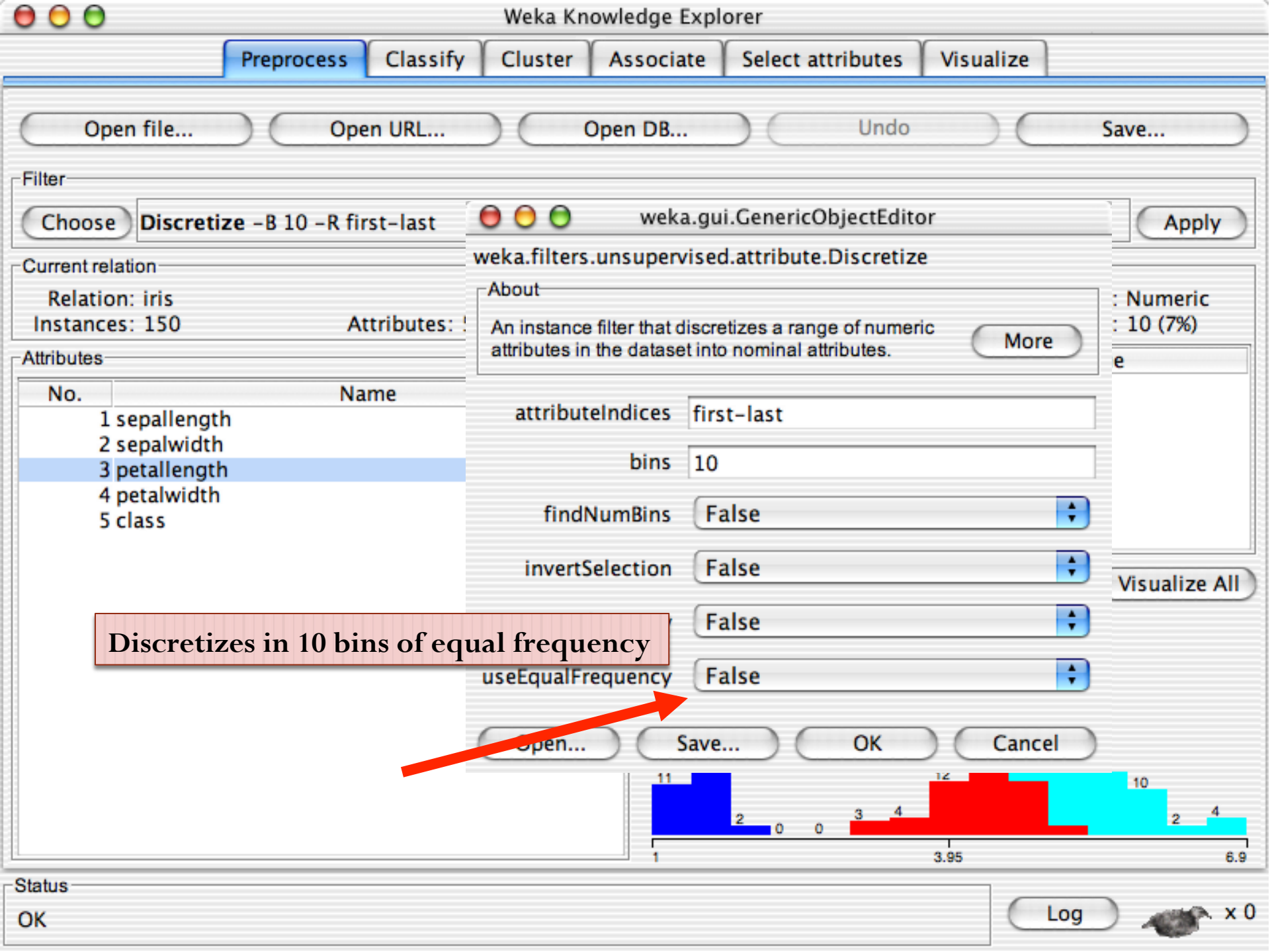
Status

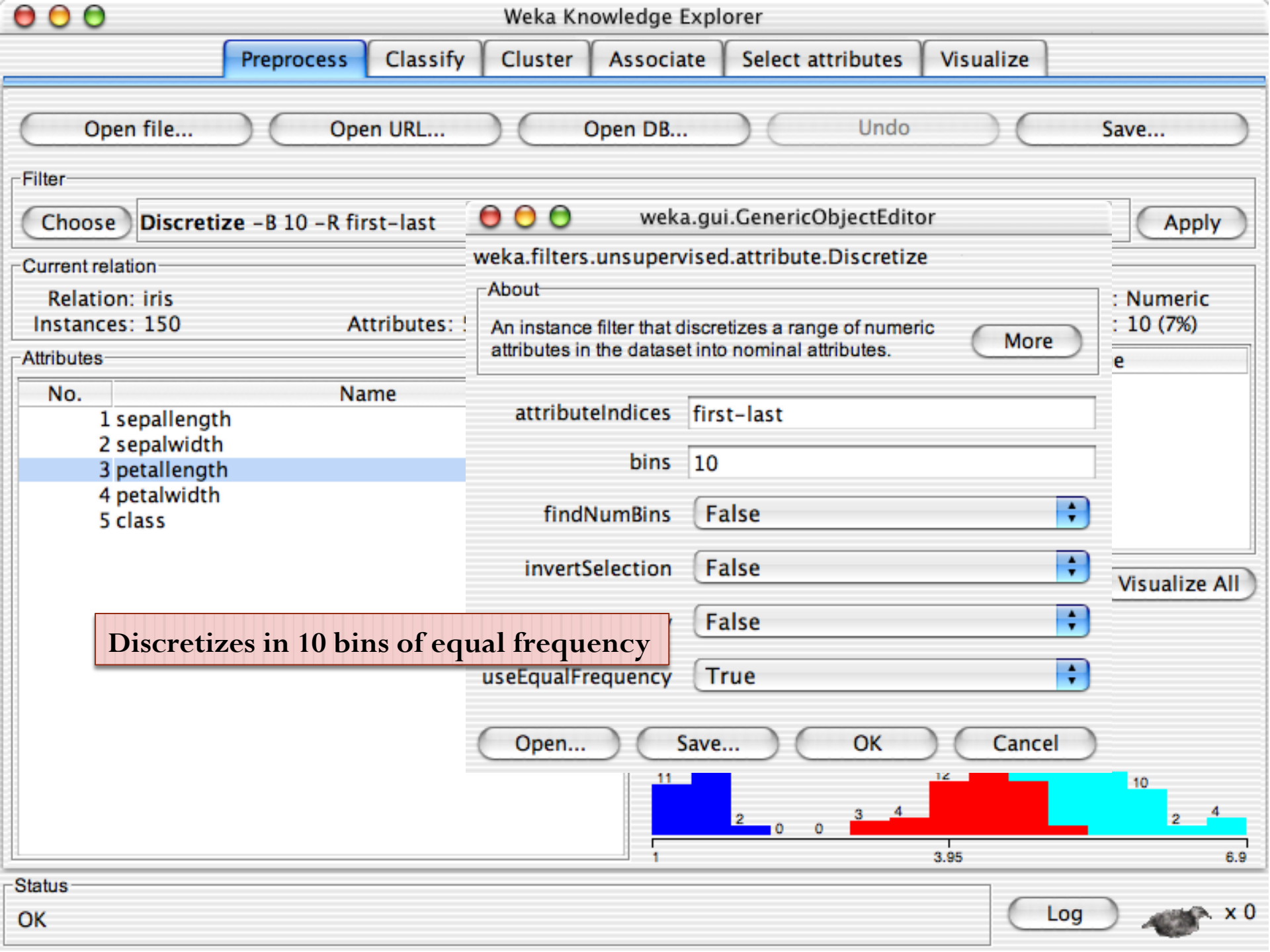
OK

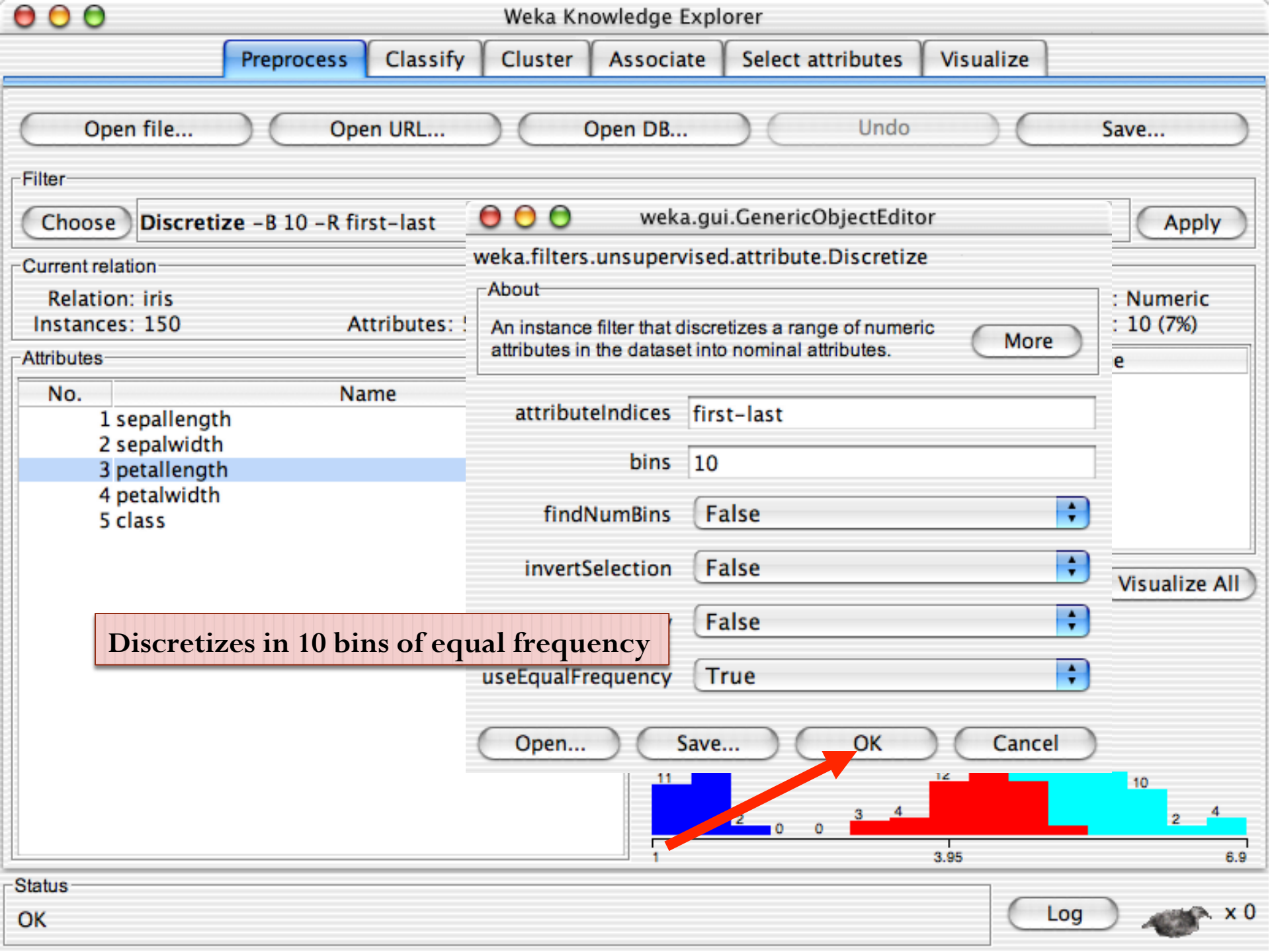
Log | x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
Choose
Discretize -B 10 -R first-last
Apply
Current relation
Relation: iris
Instances: 150
Attributes: 5
Attributes
No.
Name
1 sepallength
2 sepalwidth
3 petallength
4 petalwidth
5 class
Selected attribute
Name: petallength
Type: Numeric
Missing: 0 (0%)
Distinct: 43
Unique: 10 (7%)
Statistic
Value
Minimum
1
Maximum
6.9
Mean
3.759
StdDev
1.764
Colour: class (Nom)
Visualize All
37
11
2
0
0
3
4
12
18
17
16
14
10
2
4
1
3.95
6.9
Status
OK
Log
x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
Choose
Discretize -B 10 -R first-last
Apply
Current relation
Relation: iris
Instances: 150
Attributes:
Attributes
No.
Name
1 sepallength
2 sepalwidth
3 petallength
4 petalwidth
5 class
weka.gui.GenericObjectEditor
weka.filters.unsupervised.attribute.Discretize
About
An instance filter that discretizes a range of numeric attributes in the dataset into nominal attributes.
More
attributeIndices
first-last
bins
10
findNumBins
False
invertSelection
False
makeBinary
False
useEqualFrequency
False
Open...
Save...
OK
Cancel
: Numeric
: 10 (7%)
Visualize All
11
2
0
0
3
4
12
10
2
4
1
3.95
6.9
Status
OK
Log
x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
Choose
Discretize -B 10 -R first-last
Apply
Current relation
Relation: iris
Instances: 150
Attributes:
Attributes
No.
Name
1 sepallength
2 sepalwidth
3 petallength
4 petalwidth
5 class
weka.gui.GenericObjectEditor
weka.filters.unsupervised.attribute.Discretize
About
An instance filter that discretizes a range of numeric attributes in the dataset into nominal attributes.
More
attributeIndices
first-last
bins
10
findNumBins
False
invertSelection
False
False
useEqualFrequency
False
Open...
Save...
OK
Cancel
: Numeric
: 10 (7%)
Visualize All
Discretizes in 10 bins of equal frequency
11
2
0
0
3
4
12
10
2
4
1
3.95
6.9
Status
OK
Log
x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
Choose
Discretize -B 10 -R first-last
Apply
Current relation
Relation: iris
Instances: 150
Attributes:
Attributes
No.
Name
1 sepallength
2 sepalwidth
3 petallength
4 petalwidth
5 class
weka.gui.GenericObjectEditor
weka.filters.unsupervised.attribute.Discretize
About
An instance filter that discretizes a range of numeric attributes in the dataset into nominal attributes.
More
attributeIndices
first-last
bins
10
findNumBins
False
invertSelection
False
False
useEqualFrequency
True
Open...
Save...
OK
Cancel
: Numeric
: 10 (7%)
Visualize All
Discretizes in 10 bins of equal frequency
11
2
0
0
3
4
12
10
2
4
1
3.95
6.9
Status
OK
Log
x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
Choose
Discretize -B 10 -R first-last
Apply
Current relation
Relation: iris
Instances: 150
Attributes:
Attributes
No.
Name
1 sepallength
2 sepalwidth
3 petallength
4 petalwidth
5 class
weka.gui.GenericObjectEditor
weka.filters.unsupervised.attribute.Discretize
About
An instance filter that discretizes a range of numeric attributes in the dataset into nominal attributes.
More
attributeIndices
first-last
bins
10
findNumBins
False
invertSelection
False
False
useEqualFrequency
True
Open...
Save...
OK
Cancel
Visualize All
Discretizes in 10 bins of equal frequency
11
2
0
0
3
4
12
10
2
4
1
3.95
6.9
Status
OK
Log
x 0

Weka Knowledge Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open file... | Open URL... | Open DB... | Undo | Save...

Filter

Choose **Discretize** -F -B 10 -R first-last | Apply

Current relation

Relation: iris
Instances: 150 Attributes: 5

Attributes

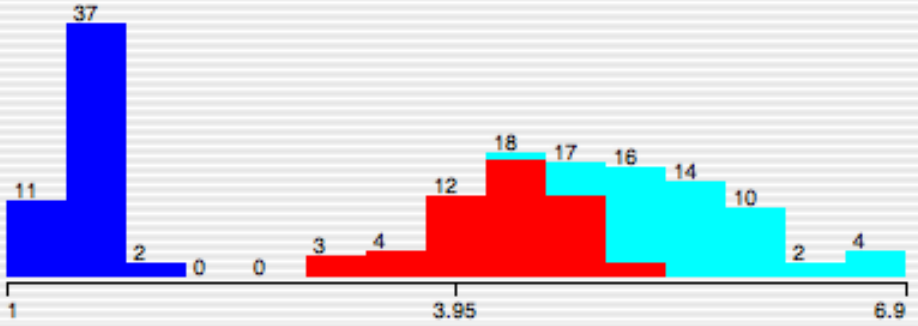
No.	Name
1	sepallength
2	sepalwidth
3	petallength
4	petalwidth
5	class

Selected attribute

Name: petallength Type: Numeric
Missing: 0 (0%) Distinct: 43 Unique: 10 (7%)

Statistic	Value
Minimum	1
Maximum	6.9
Mean	3.759
StdDev	1.764

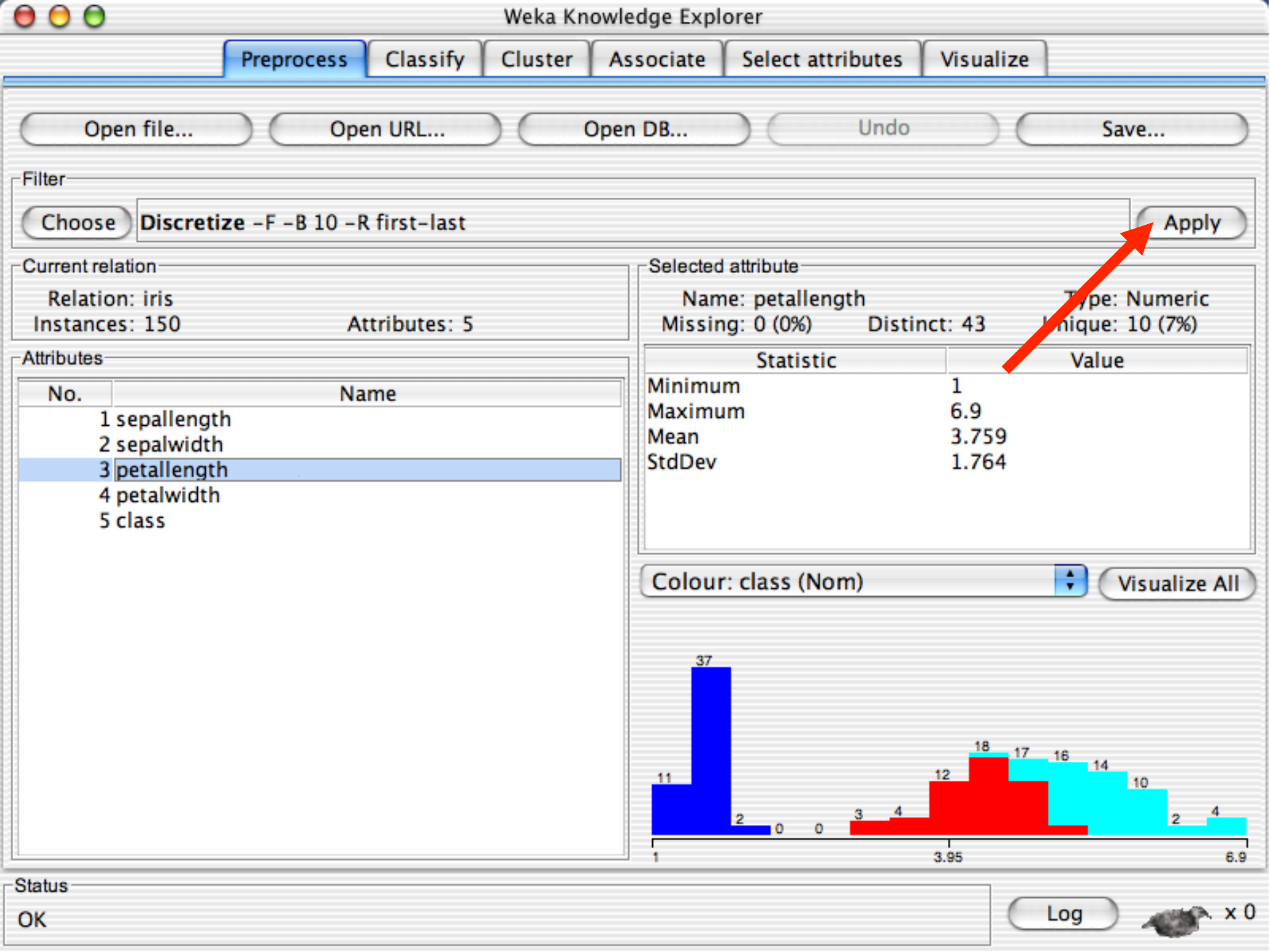
Colour: class (Nom) | Visualize All



Status

OK

Log | x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
Choose
Discretize -F -B 10 -R first-last
Apply
Current relation
Relation: iris
Instances: 150
Attributes: 5
Selected attribute
Name: petallength
Type: Numeric
Missing: 0 (0%)
Distinct: 43
Unique: 10 (7%)
Attributes
No.
Name
1 sepallength
2 sepalwidth
3 petallength
4 petalwidth
5 class
Statistic
Value
Minimum
1
Maximum
6.9
Mean
3.759
StdDev
1.764
Colour: class (Nom)
Visualize All
37
11
2
0
0
3
4
12
18
17
16
14
10
2
4
1
3.95
6.9
Status
OK
Log
x 0

Weka Knowledge Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open file... | Open URL... | Open DB... | Undo | Save...

Filter

Choose **Discretize** -F -B 10 -R first-last | Apply

Current relation

Relation: iris-weka.filters.unsupervised.attribute.Disc...
Instances: 150 Attributes: 5

Attributes

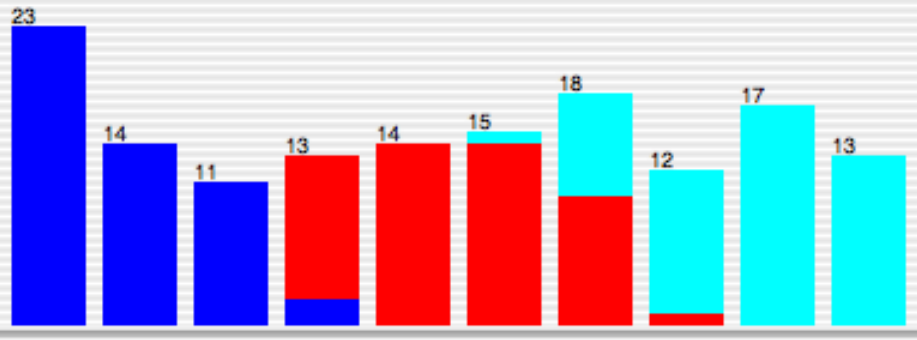
No.	Name
1	sepallength
2	sepalwidth
3	petallength
4	petalwidth
5	class

Selected attribute

Name: petallength Type: Nominal
Missing: 0 (0%) Distinct: 10 Unique: 0 (0%)

Label	Count
'(-inf-1.45]'	23
'(1.45-1.55]'	14
'(1.55-1.8]'	11
'(1.8-3.95]'	13
'(3.95-4.35]'	14
'(4.35-4.65]'	15
'(4.65-5.05]'	18

Colour: class (Nom) | Visualize All



Status

OK

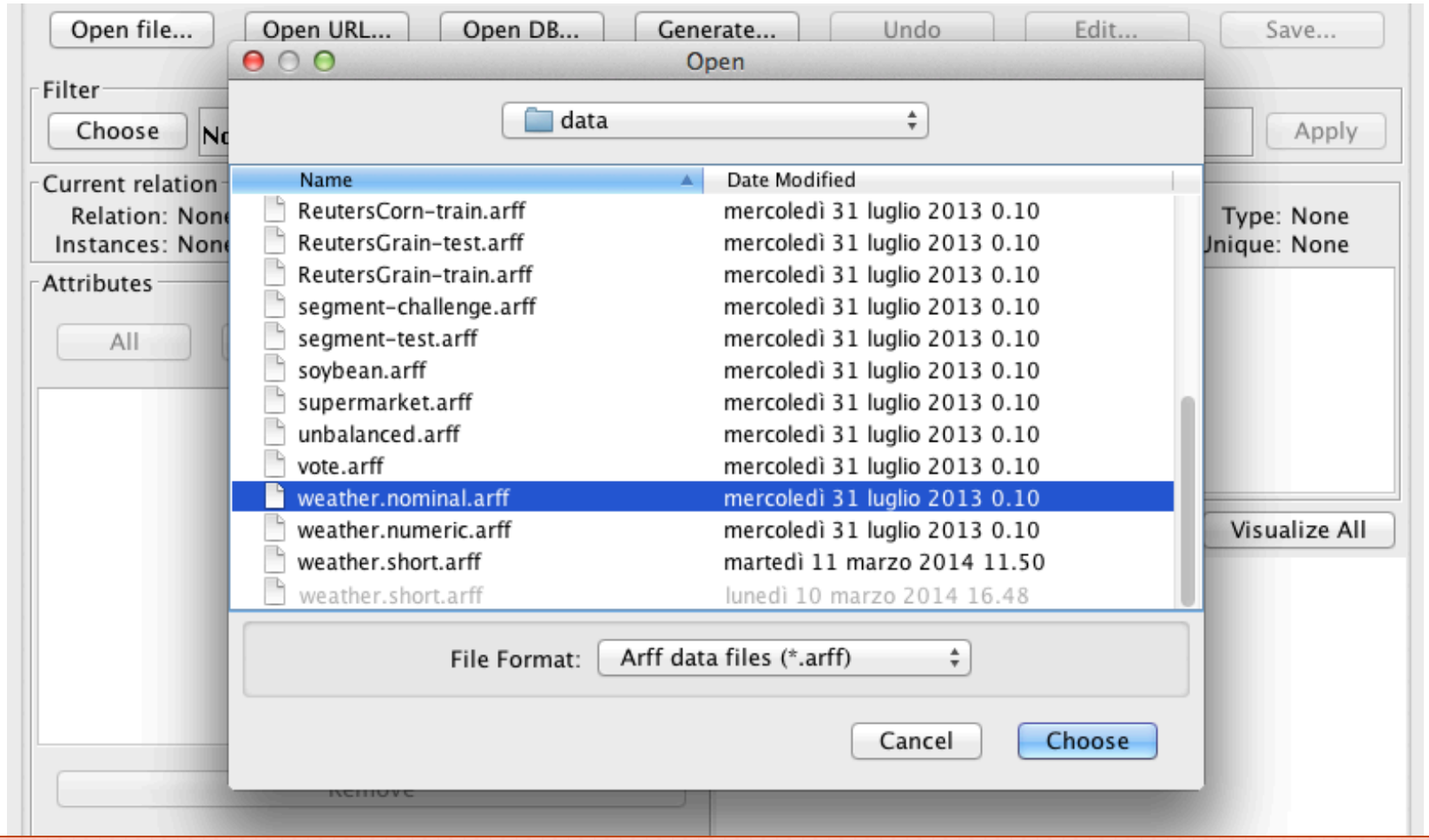
Log | x 0

WILL SEE MORE ON FILTERING DURING FIRST LAB!!

Explorer: building "classifiers"

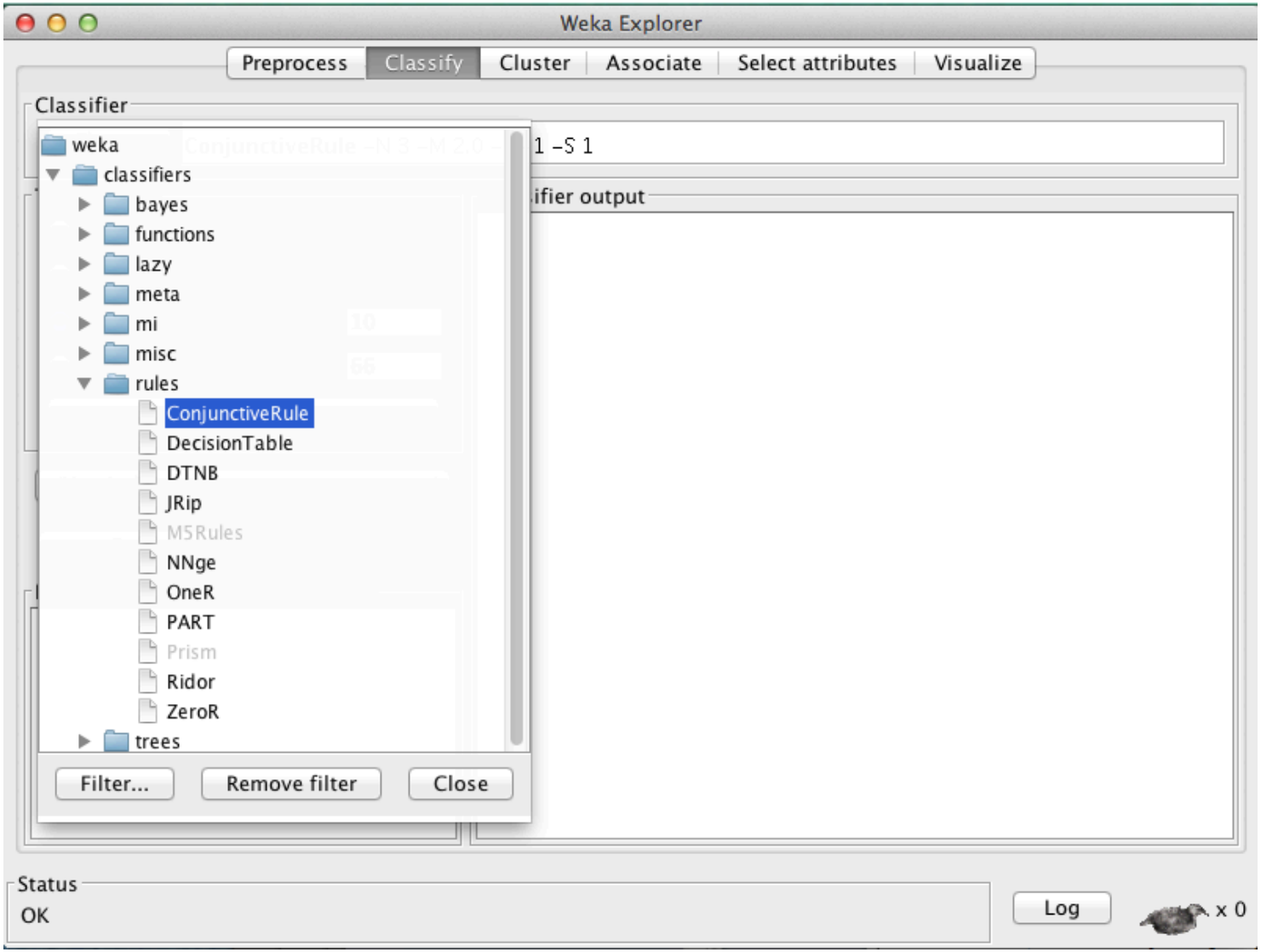
- "Classifiers" in WEKA are machine learning algorithms for predicting nominal or numeric values of a selected attribute (e.g. the CLASS attribute in the IRIS file)
- Implemented learning algorithms include:
  - Conjunctive rules, decision trees and lists, instance-based classifiers, support vector machines, multi-layer perceptrons, logistic regression, Bayes' nets, …
  - Most, but not all, the algorithms that we will present in this course (e.g. no genetic or reinforcement algorithms)

Explore Conjunctive Rules learner

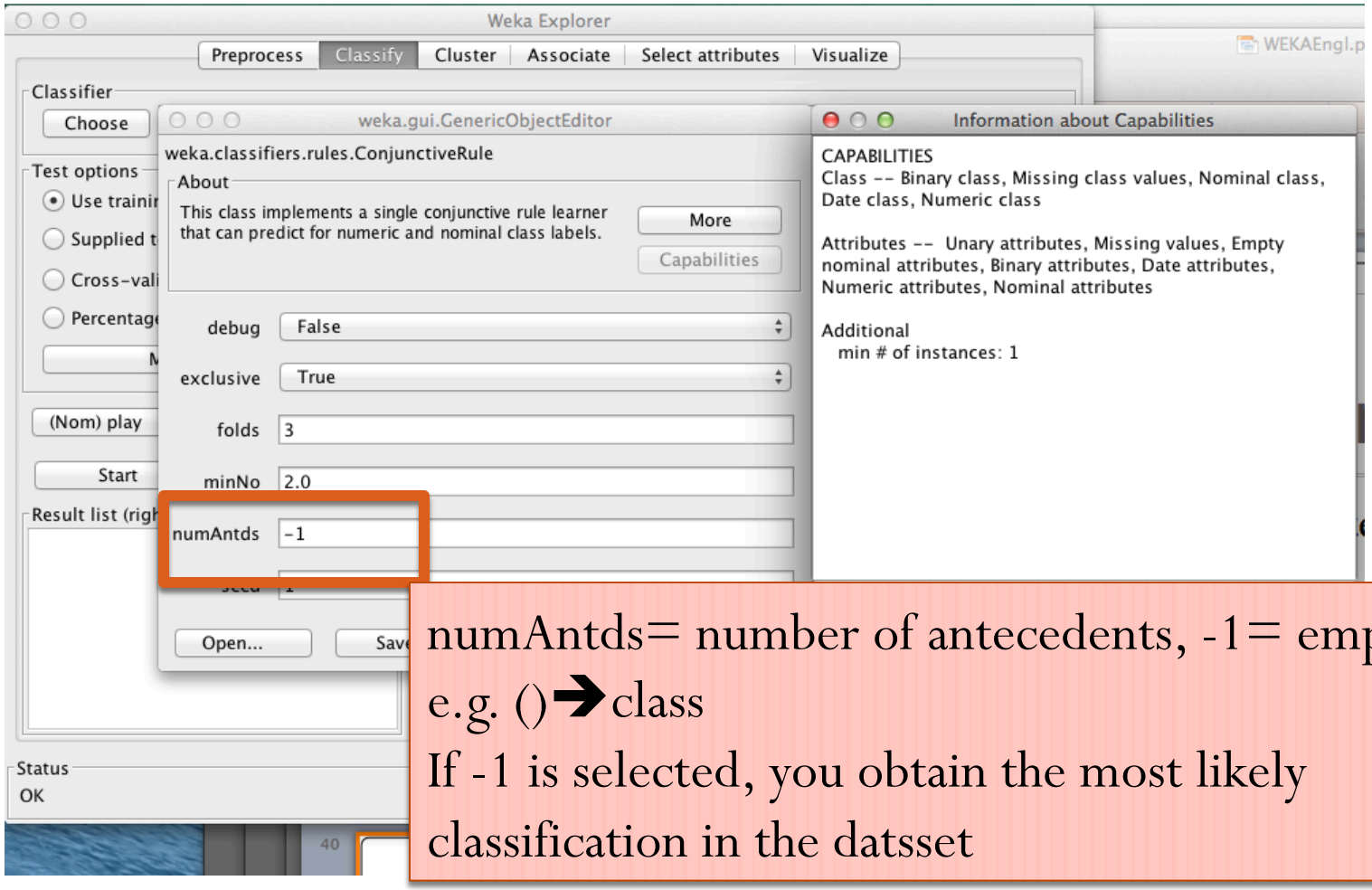


Need a simple dataset with few attributes , let's select the weather dataset

Select a Classifier

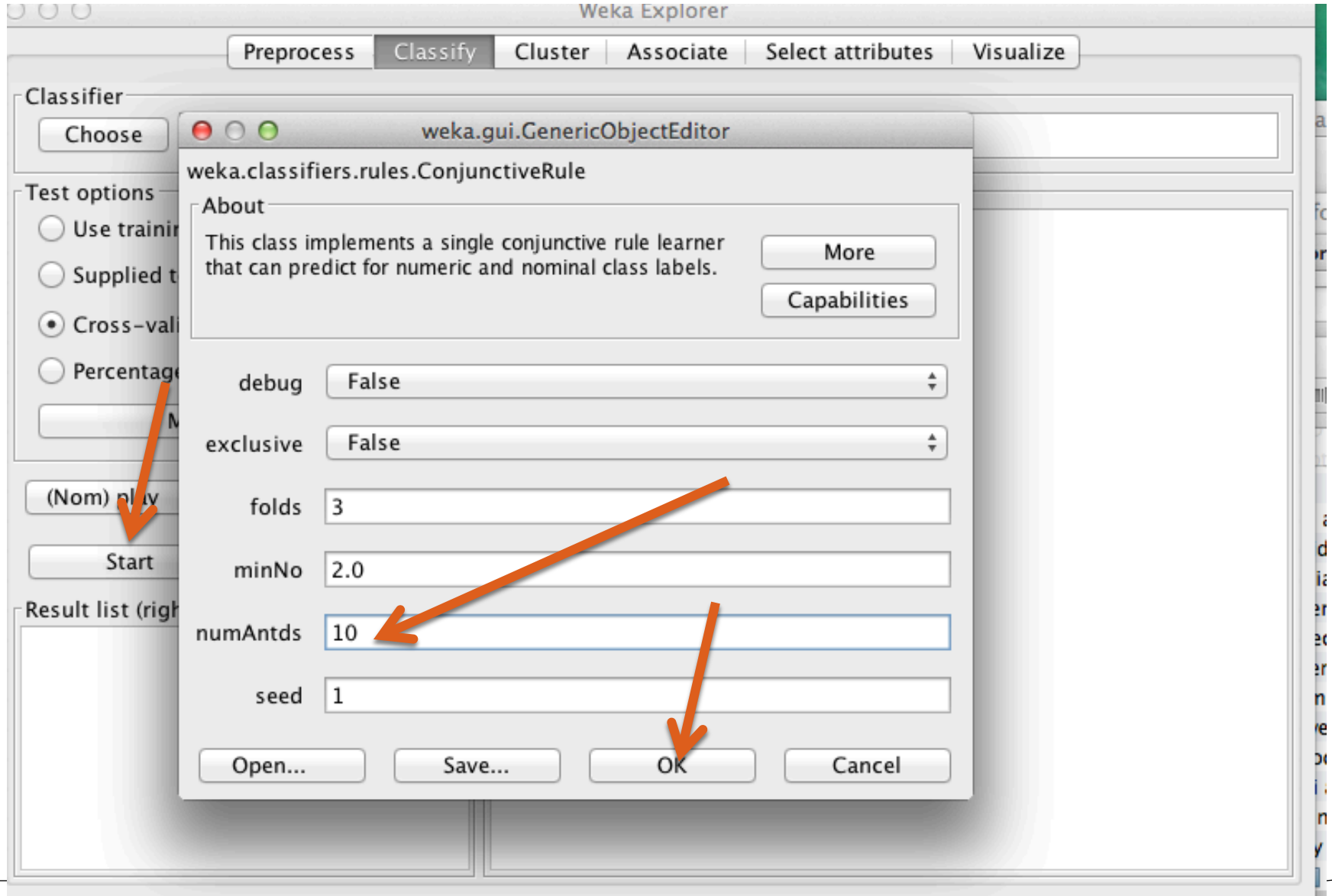


Right-click to select parameters

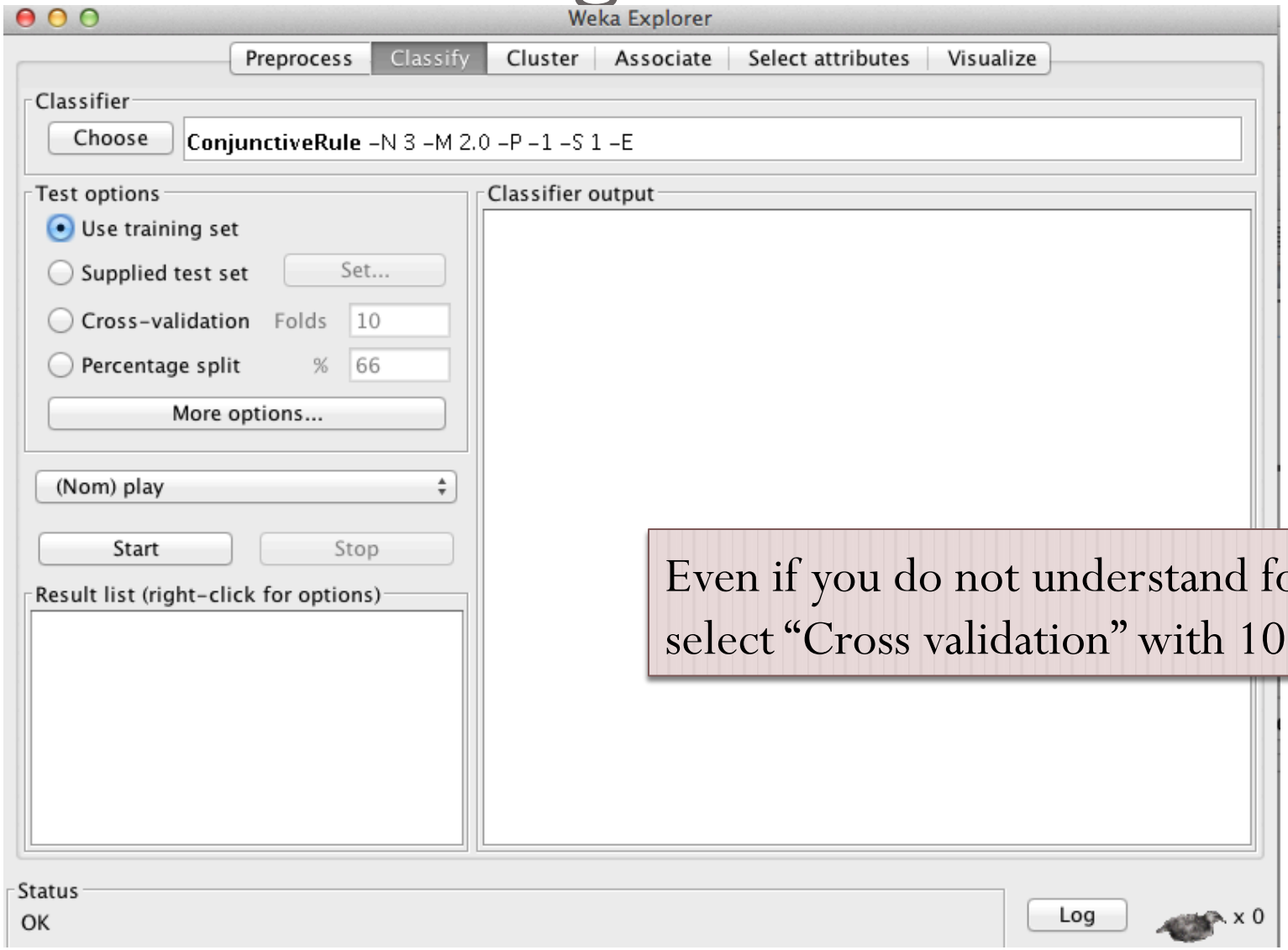


numAntds= number of antecedents, -1= empty rule
e.g. ()➔class
If -1 is selected, you obtain the most likely classification in the datsset

Select numAntds=10

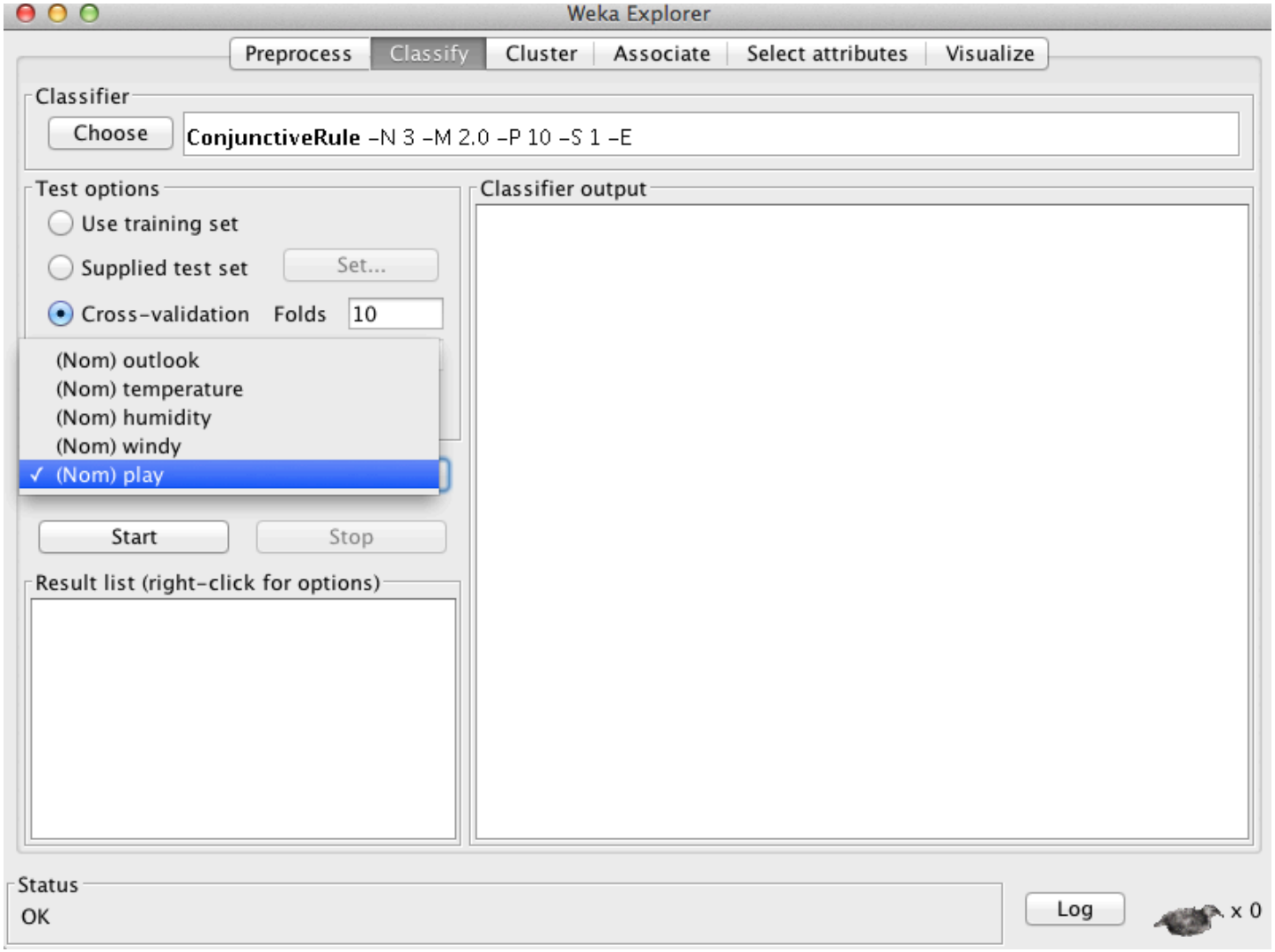


Select training method

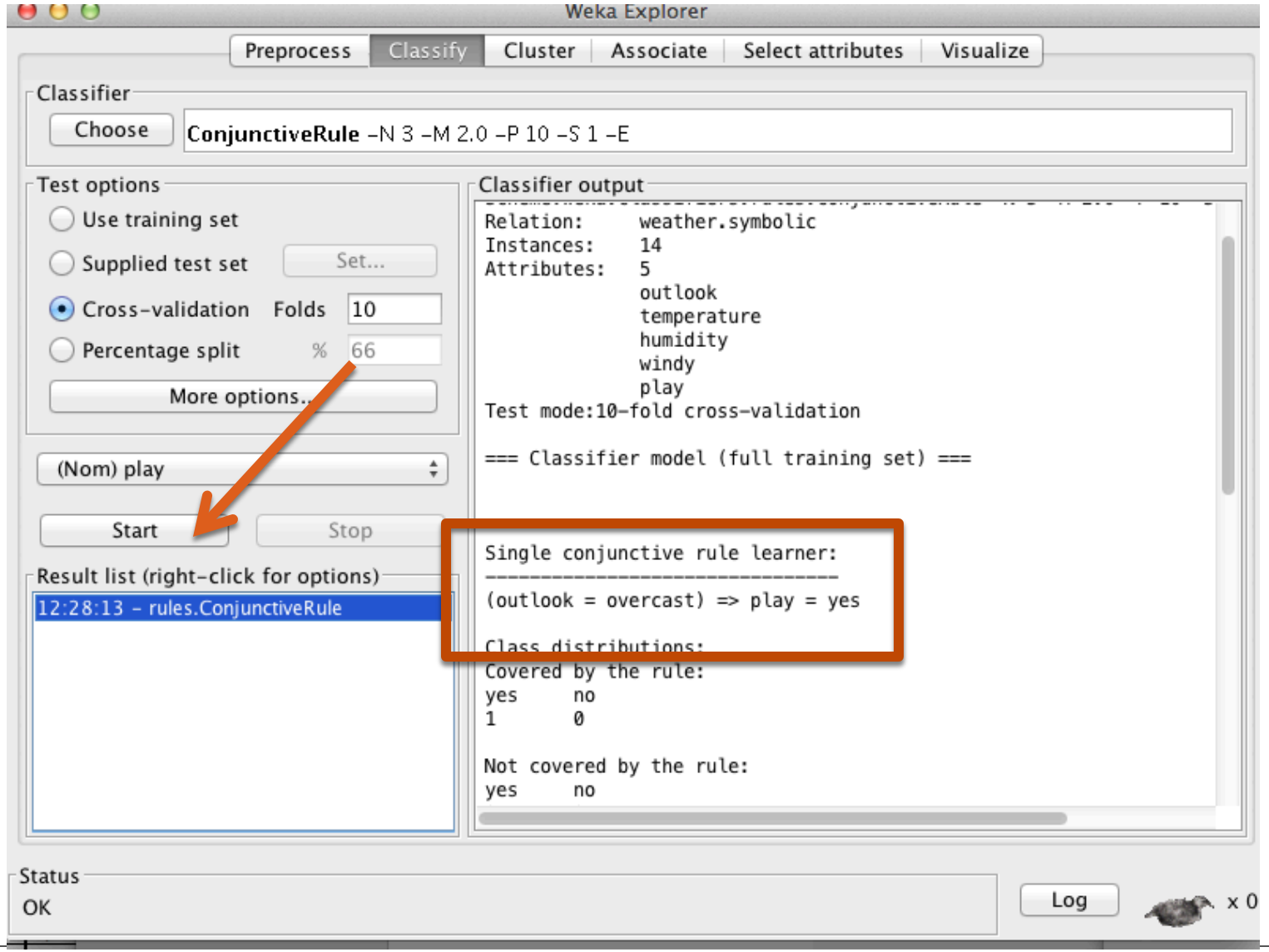


Even if you do not understand for now, select "Cross validation" with 10 Folds.

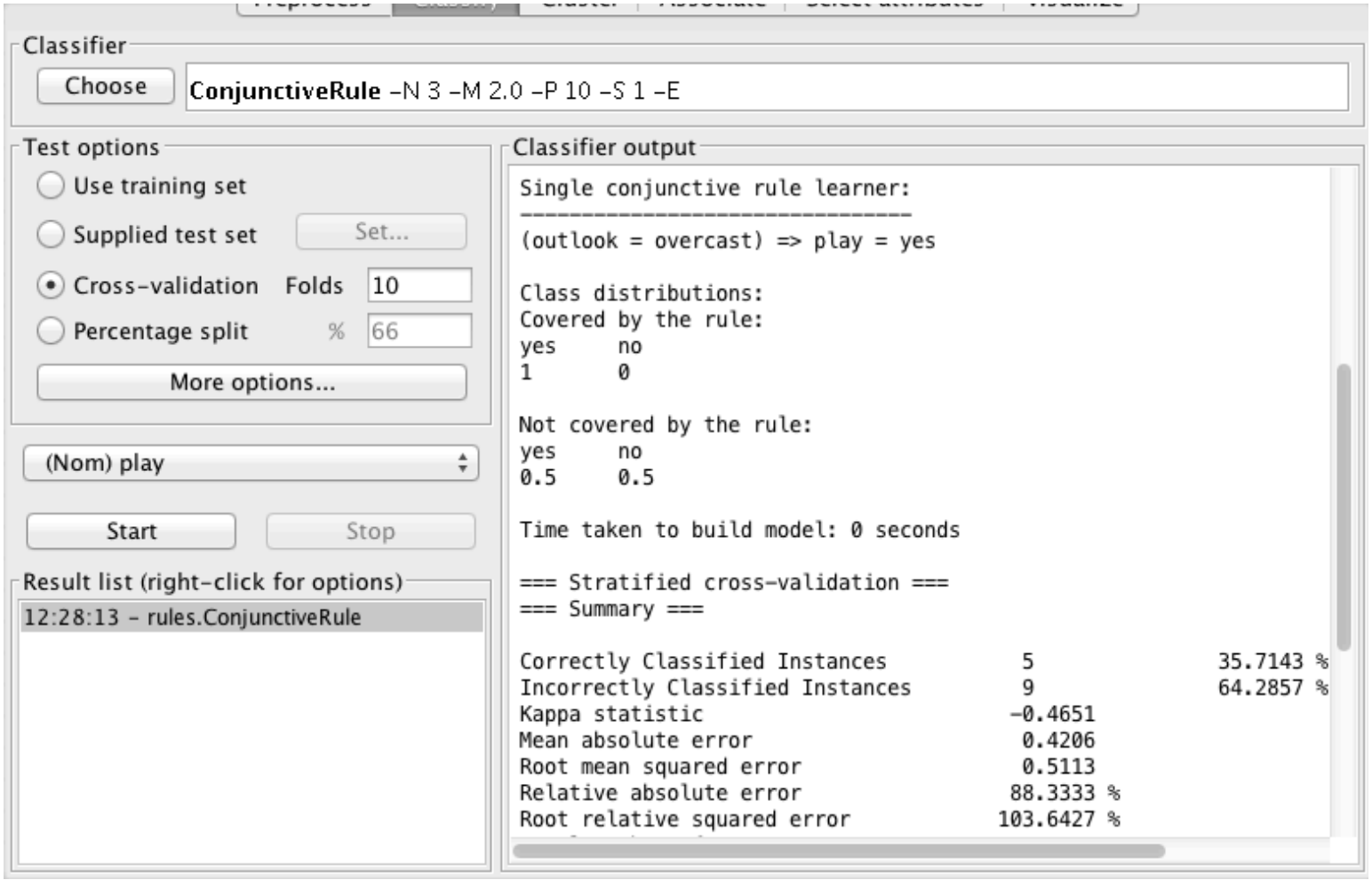
Select the right hand side of the rule (the classification function)



Run the algorithm



Performance data



Error measures in WEKA

Mean absolute error is:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}|\hat{\theta}_i - \theta_i|$$

Root mean square error is:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i}^{N}\left(\hat{\theta}_i - \theta_i\right)^2}$$

Where θ_i is the "true" classification for x_i and θ_i ^ is the classification produced by the ML algorithm (for binary classifiers θ is either 0 or 1 and | θ_i ^- θ_i | is either 0 or 1.

$$\sum_{i=1}^{N}|\theta - \theta_i|$$

where $\bar{\theta}$ is a mean value of θ.

Root relative squared error:

$$RRSE = \sqrt{\frac{\sum_{i=1}^{N}\left(\hat{\theta}_i - \theta_i\right)^2}{\sum_{i=1}^{N}\left(\bar{\theta} - \theta_i\right)^2}}$$

Confusion Matrix

	System classified as a	System classified as b
Truly classified as a	# of instances that system classifies a, ground truth is a	# of instances that system classifies b, ground truth is a
Truly classified as b	# of instances that system classifies a, ground truth is b	# of instances that system classifies b, ground truth is b

```
=== Confusion Matrix ===

 a b   <-- classified as
 5 4 | a = yes
 5 0 | b = no
```

Cells (1,1) and (2,2) represent "good" classifications. The others are wrong. In fact, we are told that there are 5 correctly classified instances and 9 errors.