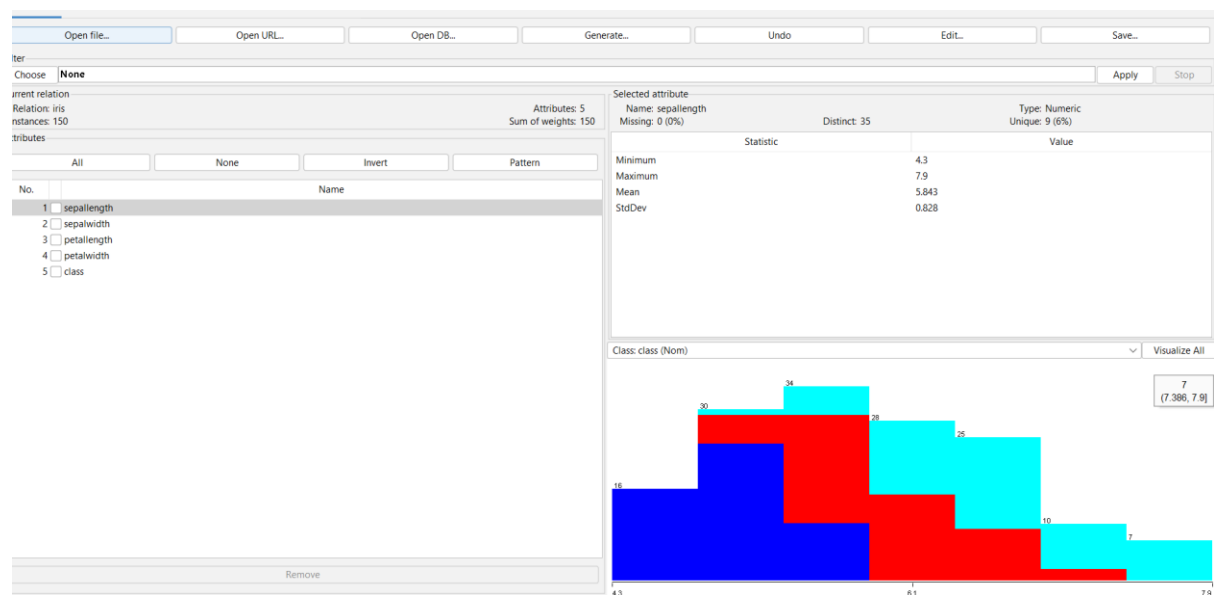# Data analysis by expectation maximization algorithm



## Cluster  EM

Clusterer

Choose   **EM** -I 100 -N -1 -X 10 -max -1 -ll-cv 1.0E-6 -ll-iter 1.0E-6 -M 1.0E-6 -K 10 -num-slots 1 -S 100

Cluster mode

- ● Use training set
- ○ Supplied test set          Set...
- ○ Percentage split        %   66
- ○ Classes to clusters evaluation
  - (Nom) class                    ∨
- ☑ Store clusters for visualization

Ignore attributes

Start              Stop

Result list (right-click for options)
13:26:51 - EM
13:32:36 - SimpleKMeans
13:38:51 - EM

Clusterer output

```
sepallength
  mean                5.897    5.006   6.9426  6.1304
  std. dev.          0.5279   0.3489    0.498  0.2943

sepalwidth
  mean               2.7519    3.418   3.1103  2.8088
  std. dev.          0.3103   0.3772   0.2952  0.2361

petallength
  mean               4.2267    1.464   5.8559  5.0993
  std. dev.           0.445   0.1718   0.4626  0.2462

petalwidth
  mean               1.3134    0.244   2.1495  1.8254
  std. dev.          0.1864   0.1061    0.232  0.2152

class
  Iris-setosa             1       51        1       1
  Iris-versicolor   48.1125        1   1.0182  3.8693
  Iris-virginica     2.0983        1  31.0375 19.8641
  [total]           51.2108       53  33.0557 24.7335


Time taken to build model (full training data) : 0.19 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      48 ( 32%)
1      50 ( 33%)
2      29 ( 19%)
3      23 ( 15%)


Log likelihood: -2.03504
```

Status
Ok                                                    Log     x 0