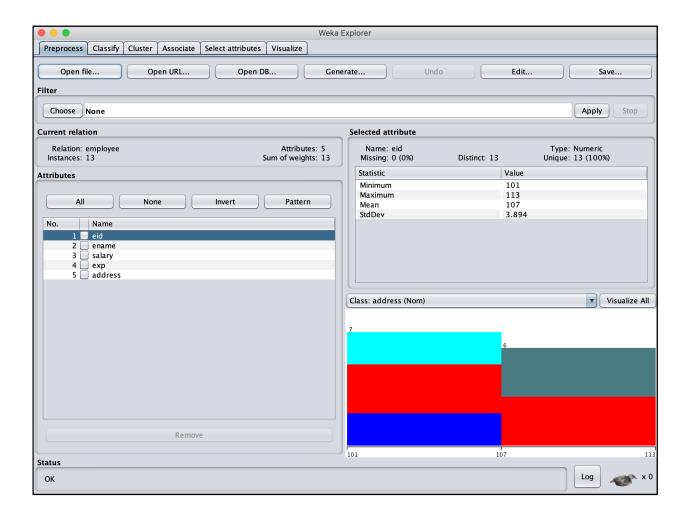# Data warehouse & Data mining lab Assignment 4a

Submitted by :
Shiv Kumar
2016UIT2563
Semester 8
IT2

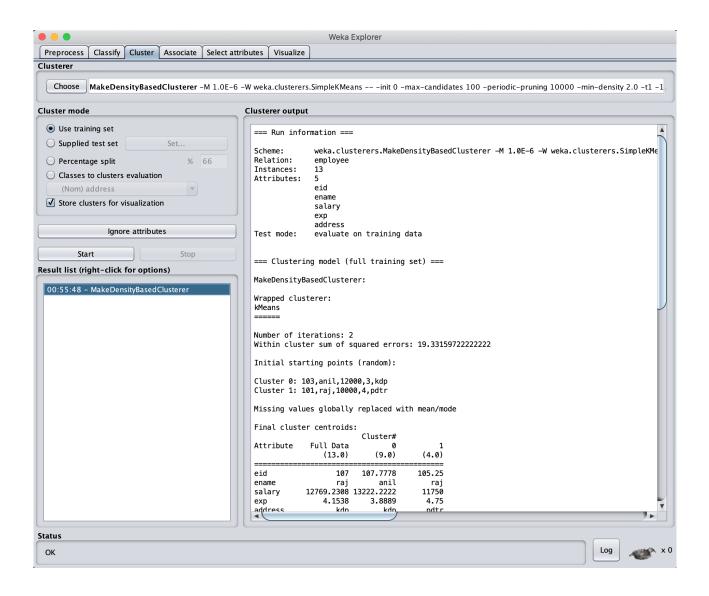# Experiment 11 : Write a procedure for Employee data using Make Density Based Cluster Algorithm.

1. Data Loading

2. Review of Loaded Data

Viewer

Relation: employee

No.	1: eid Numeric	2: ename Nominal	3: salary Numeric	4: exp Numeric	5: **address** Nominal
1	10...	raj	1000...	4.0	pdtr
2	10...	ramu	1500...	5.0	pdtr
3	10...	anil	1200...	3.0	kdp
4	10...	sunil	1300...	3.0	kdp
5	10...	rajiv	1600...	6.0	kdp
6	10...	sunitha	1500...	5.0	nlr
7	10...	kavitha	1200...	3.0	nlr
8	10...	suresh	1100...	5.0	gtr
9	10...	ravi	1200...	3.0	gtr
...	11...	ramana	1100...	5.0	gtr
...	11...	ram	1200...	3.0	kdp
...	11...	kavya	1300...	4.0	kdp
...	11...	navya	1400...	5.0	kdp

Add instance   Undo   OK   Cancel

3. Result after Make Density Based Cluster Algorithm.

```
Final cluster centroids:
                      Cluster#
Attribute     Full Data        0           1
              (13.0)        (9.0)       (4.0)
==========================================
eid               107    107.7778     105.25
ename             raj        anil        raj
salary      12769.2308  13222.2222      11750
exp            4.1538      3.8889       4.75
address          kdp         kdp        pdtr




Fitted estimators (with ML estimates of variance):

Cluster: 0 Prior probability: 0.6667

Attribute: eid
Normal Distribution. Mean = 107.7778 StdDev = 3.4247
Attribute: ename
Discrete Estimator. Counts =  1 1 2 2 2 2 2 1 2 1 2 2 2  (Total = 22)
Attribute: salary
Normal Distribution. Mean = 13222.2222 StdDev = 1396.645
Attribute: exp
Normal Distribution. Mean = 3.8889 StdDev = 1.0999
Attribute: address
Discrete Estimator. Counts =  1 7 3 2  (Total = 13)

Cluster: 1 Prior probability: 0.3333

Attribute: eid
Normal Distribution. Mean = 105.25 StdDev = 3.8324
Attribute: ename
Discrete Estimator. Counts =  2 2 1 1 1 1 1 2 1 2 1 1 1  (Total = 17)
Attribute: salary
Normal Distribution. Mean = 11750 StdDev = 1920.2864
Attribute: exp
Normal Distribution. Mean = 4.75 StdDev = 0.433
Attribute: address
Discrete Estimator. Counts =  3 1 1 3  (Total = 8)
```

```
Time taken to build model (full training data) : 0.01 seconds

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

Clustered Instances

0       9 ( 69%)
1       4 ( 31%)


Log likelihood: -16.52967
```