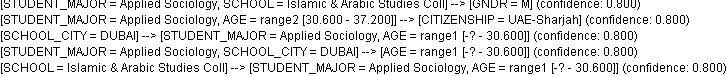
## Classification

## Clustering

## Association Rule Mining

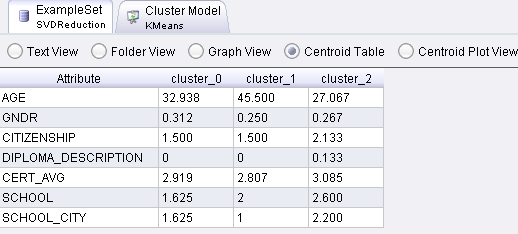


These are some rules as above figure showing:

1. If student major is "Applied Sociology" and School is "Islamic & Arabic Studies college" then there is 80% confidence that student would be male.
2. If Student major is "Applied Sociology" and age between 30-37 then there is 80% confidence that student citizenship is UAE-Sharjah
3. If Student major is "Applied Sociology", school city is "Dubai" then there is 80% confidence that student is under 30 age.
4. If School is 'Islamic & Arabic coll" then there is 80% confidence that student major is "Applied Sociology" and age is under 30

## K-Mean Clustering algorithm

K-Mean can only work for attributes having integer values, it can't be run for nominal attribute so here I converted nominal attributes to numerical before applying K-MEAN algorithm.



### Comments:

It formed 3 clusters where student's having age from 31-34 are in Cluster 0, age 32 is mean

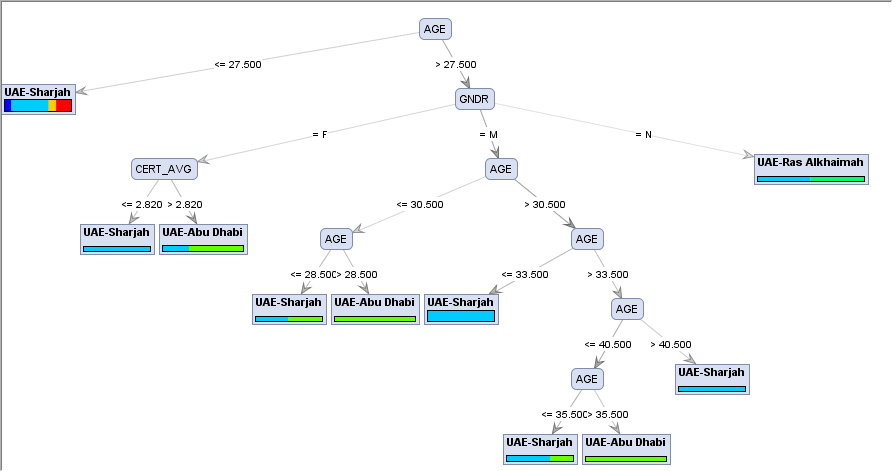
Cluster 1: mean age 45

Cluster 2: All students having range 25-30 are in Cluster 2, where 27 is mean age.

# Cluster Model

Cluster 0: 16 items  
Cluster 1: 4 items  
Cluster 2: 15 items  
Total number of items: 35

## Decision Tree Analysis



Here "CITIZEN SHIP" attributes set as Class label if you want me to set another attribute as class variable then let me know.

**First two branch analysis:**

If student age is under 27 then he belongs to UAE-SHarjah else if student age is >=27 and she is female and having CERT\_AVG either <=2.8 then she belongs to UAE\_SHARJAH

If student age is under 27 then he belongs to UAE-SHarjah else if student age is >=27 and she is female and having CERT\_AVG either > 2.8 then she belongs to UAE\_Abu Dhabi