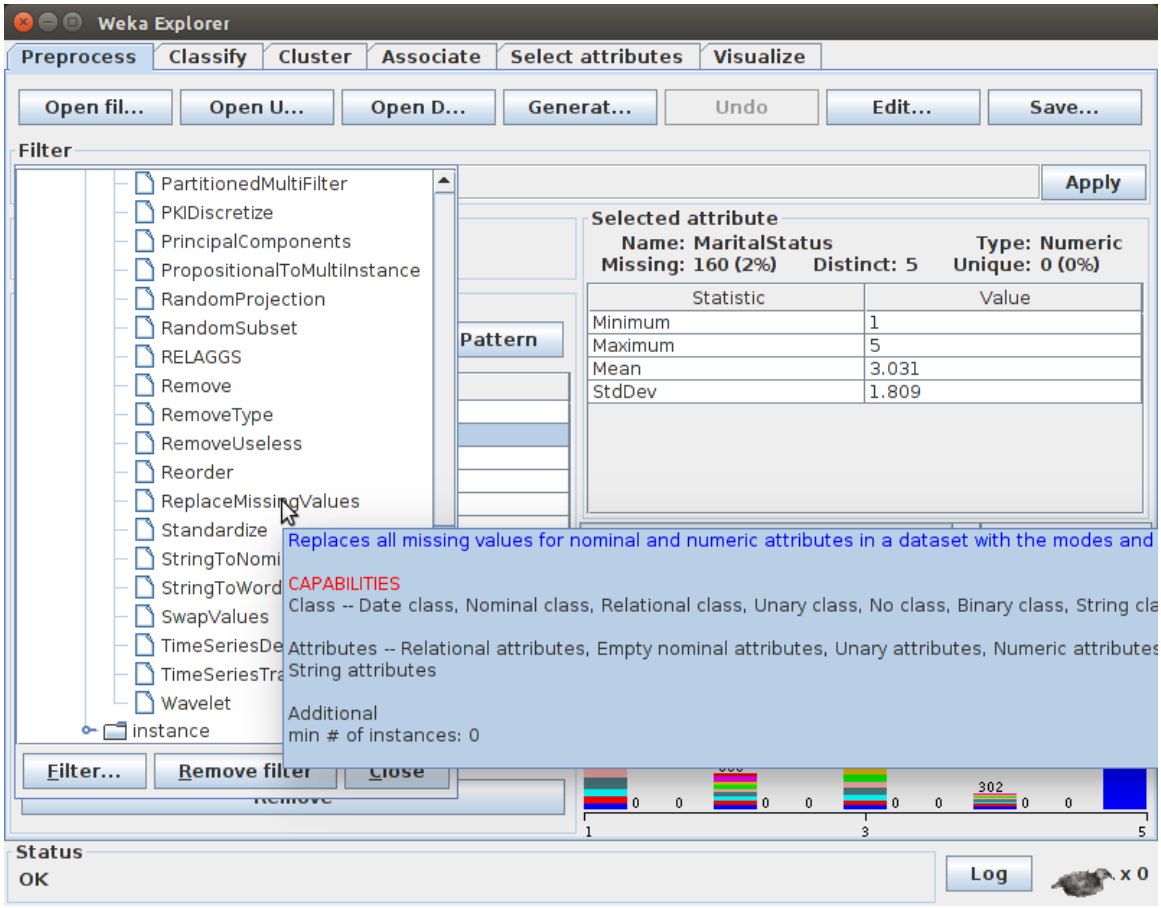
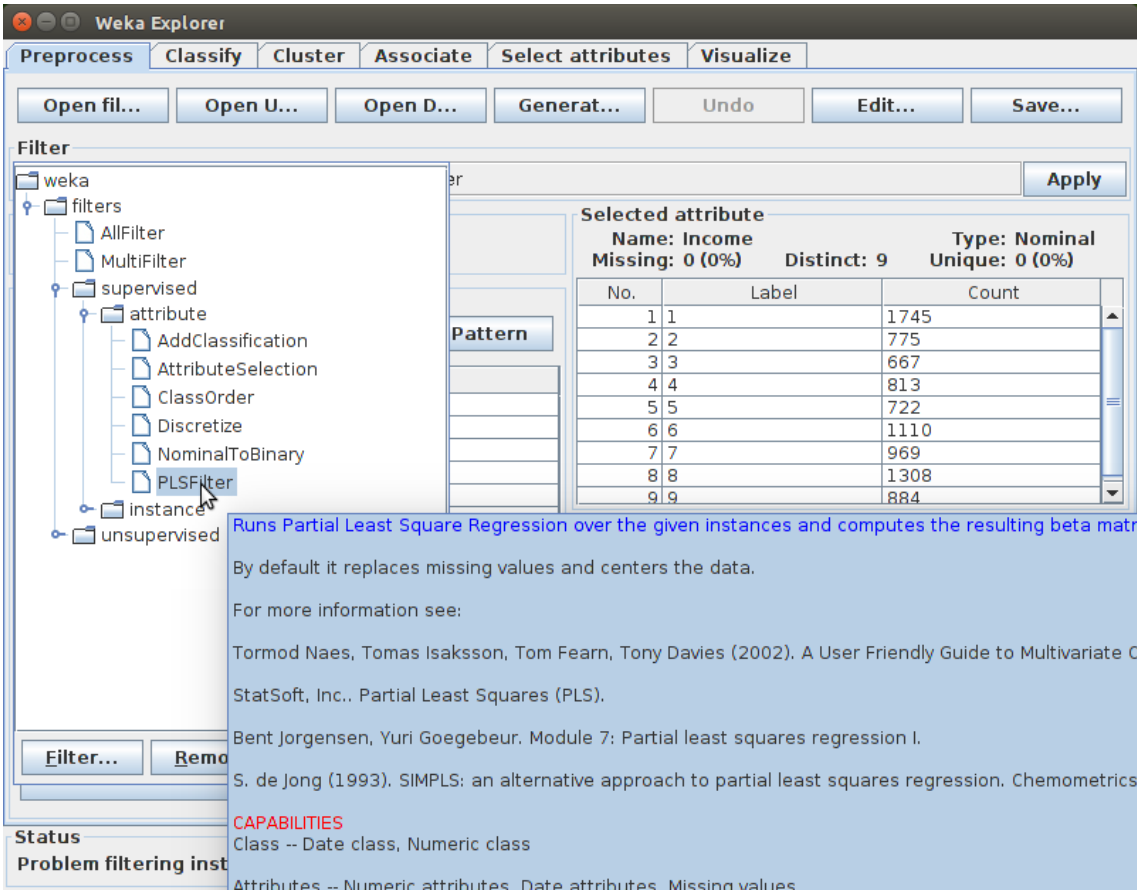
1.**Pre-Process Missing Values**

Preprocess(tab)

filter(choose)>unsupervised>attribute>**ReplaceMissingValues**



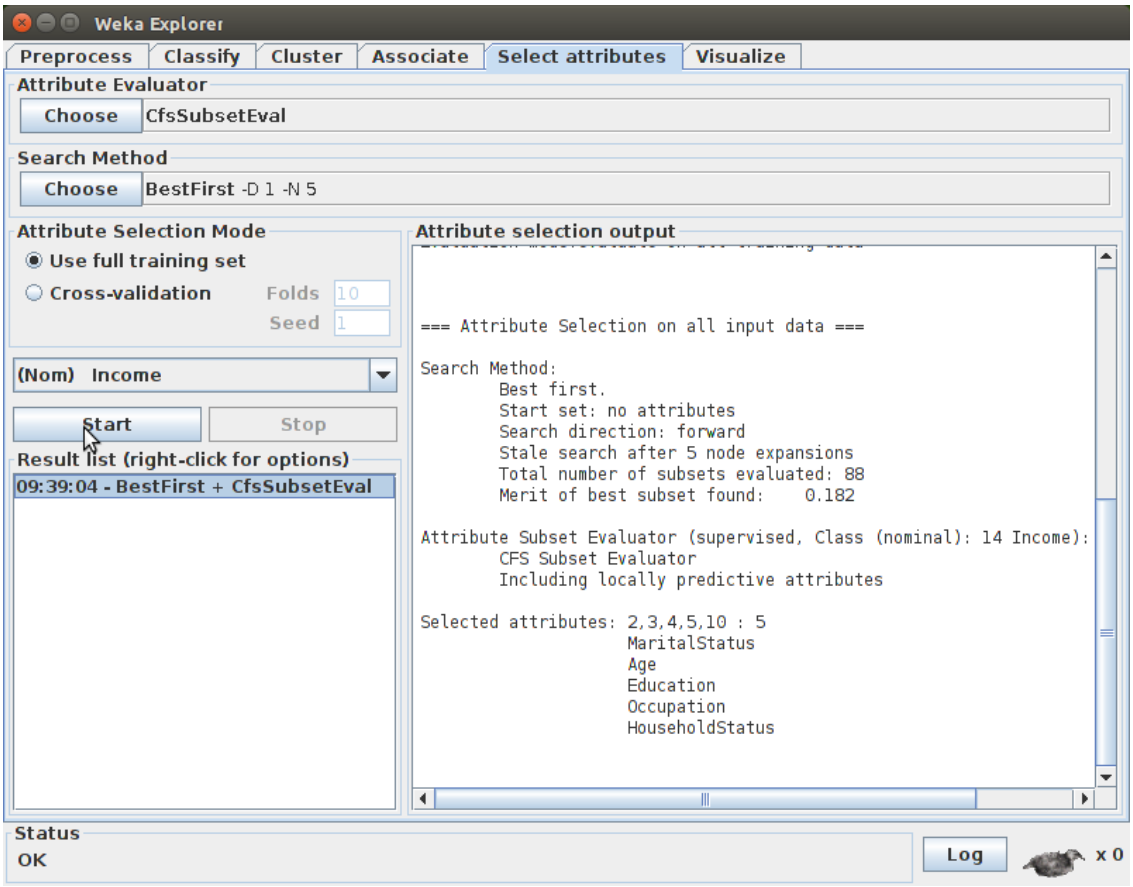
supervised>attribute>**PLSfilter**(cannot handle multi-valued nominal class)



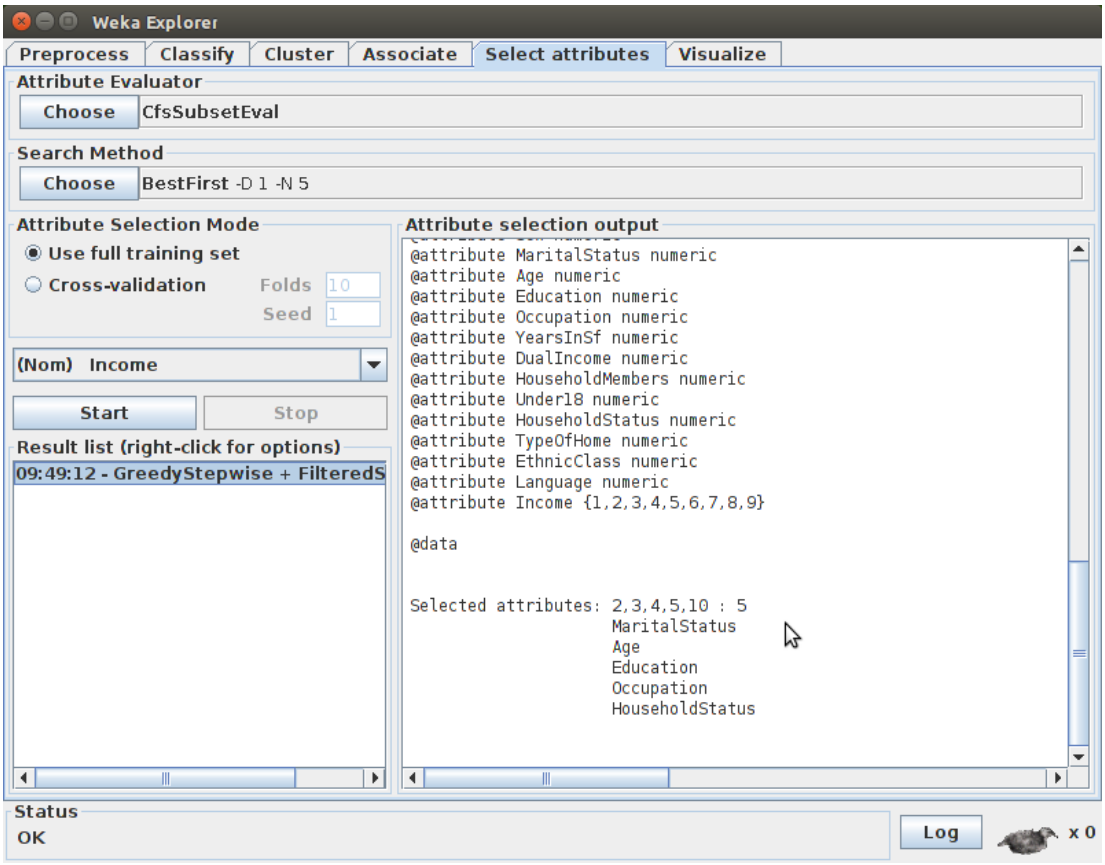
2.**Attribute Filter Option**

Select Attributes(tab)

Attribute Evaluator:CfsSubsetEval Method:BestFirst 5 selected in 14



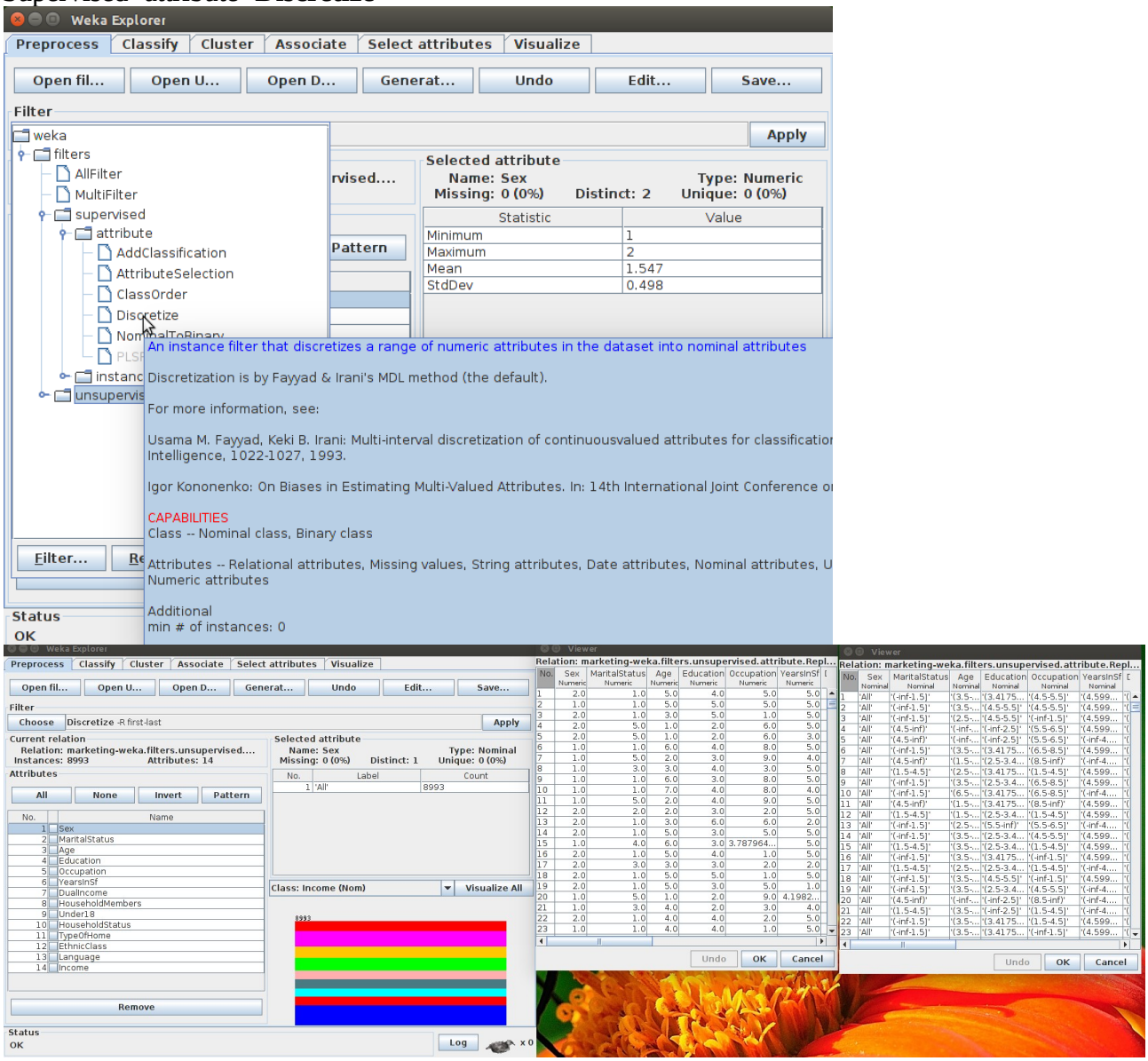
Attribute Evaluator:FilteredSubsetEval Method:GreedyStepWise 5 selected in 14



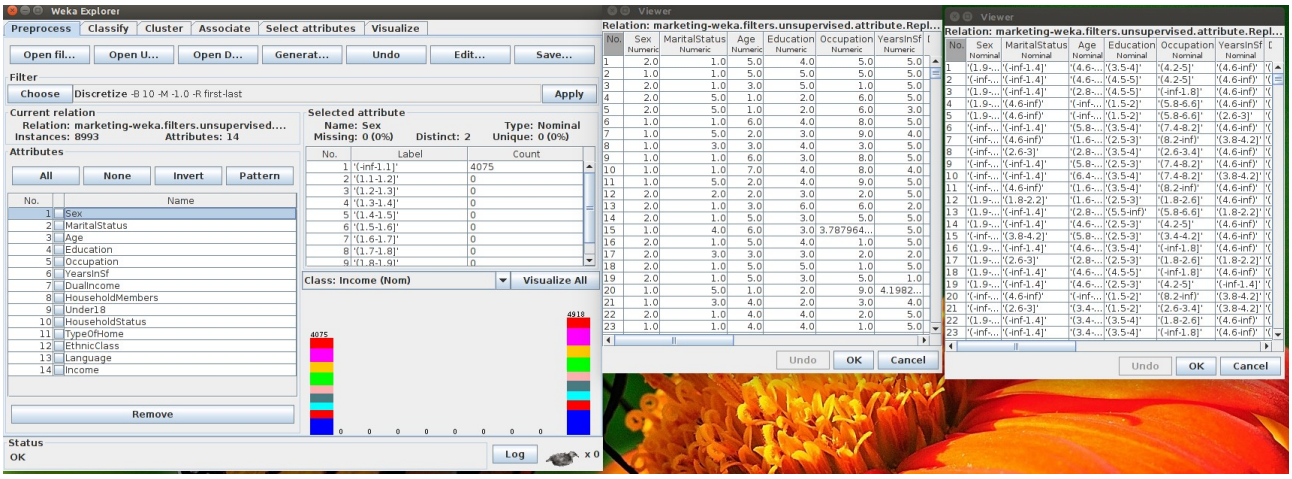
3.**Discretization**

Preprocess(tab)

Supervised>attribute>**Discretize**

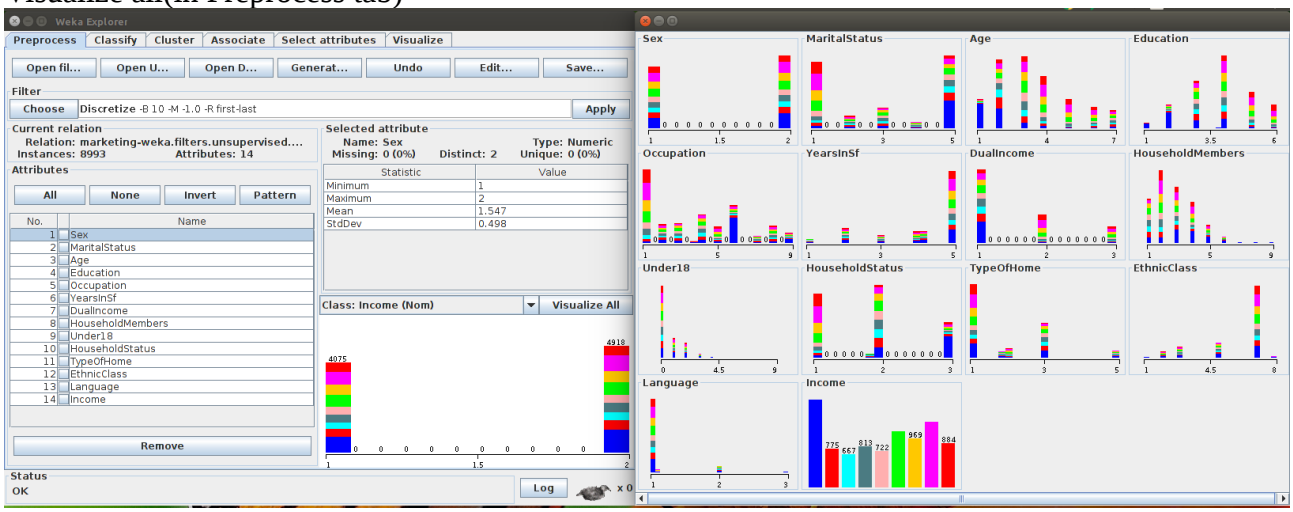


Unsupervised>attribute>**Discretize**

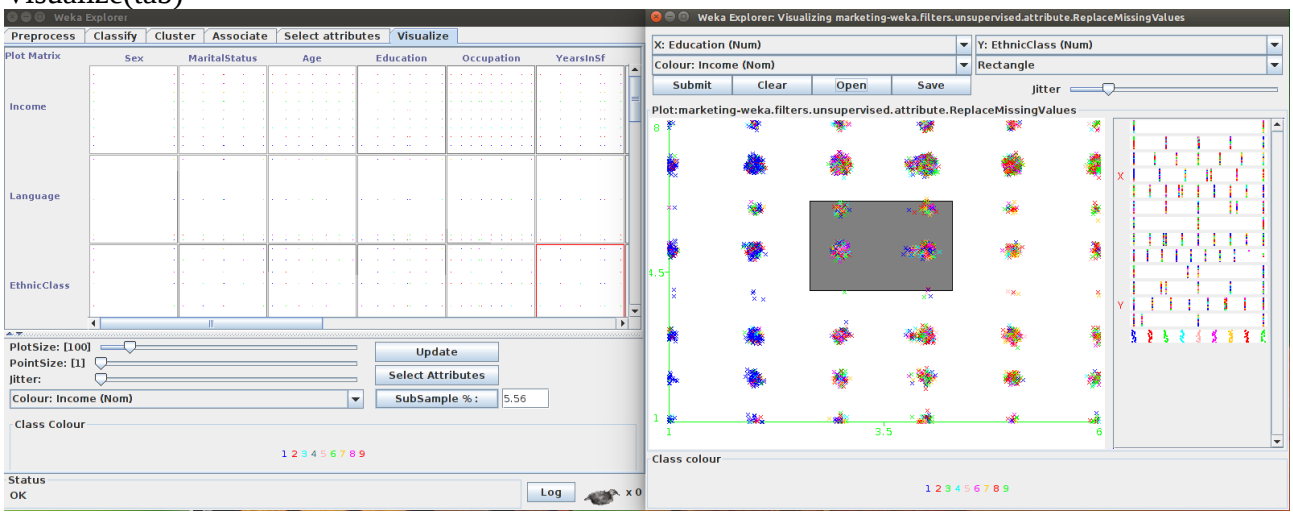


4.**Visualize**

Visualize all(in Preprocess tab)



Visualize(tab)



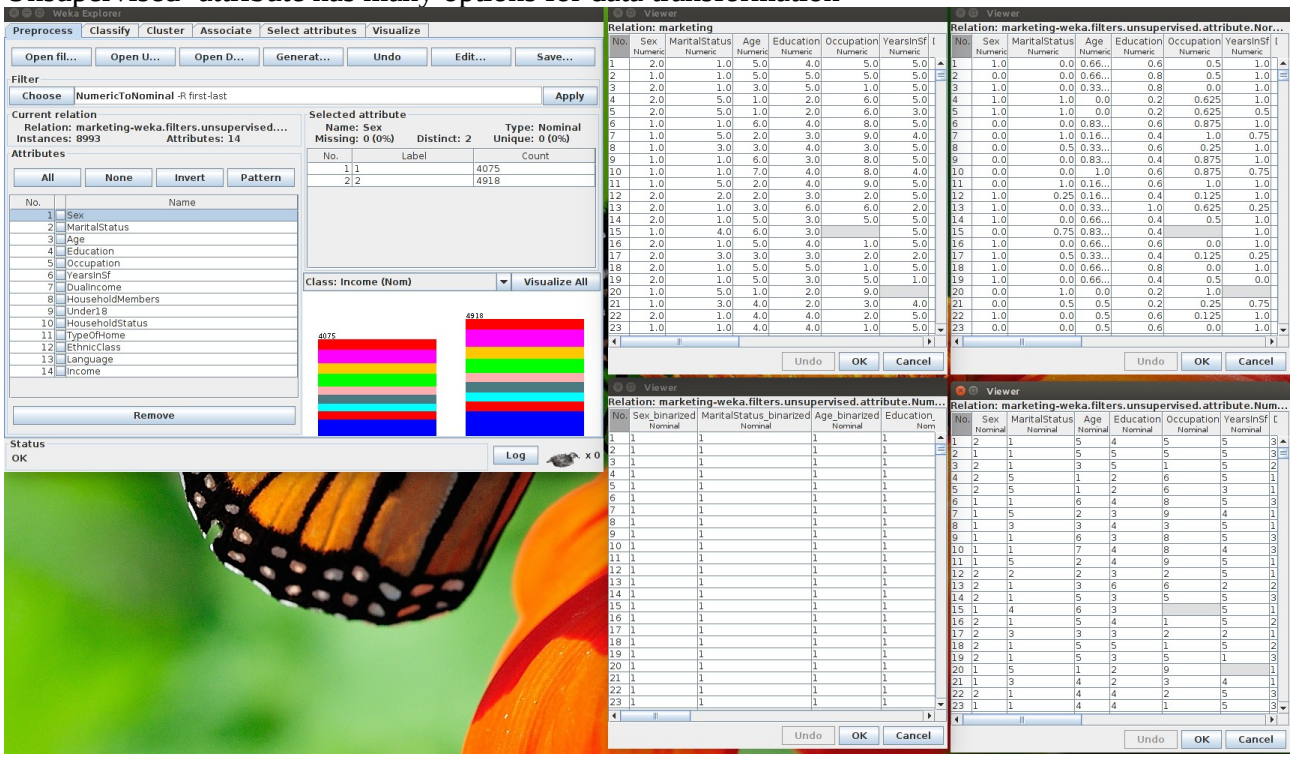
(LEFT img)A matrix of all attributes against alla ttriutes is provided(with default coloring with CLASS attribute), options to play with the visualization are provided at the bottom

(RIGHT img)Selecting any section gives a closer look onto the data, with a flexible options, like hanging X, Y axes, selecting a region(from Select Instance-(here used Rectangle))

5.**Data Transformation**

Preprocess(tab)

Unsupervised>attribute has many options for data transformation



1.Original

2.Normalize

3.NumericToBinary

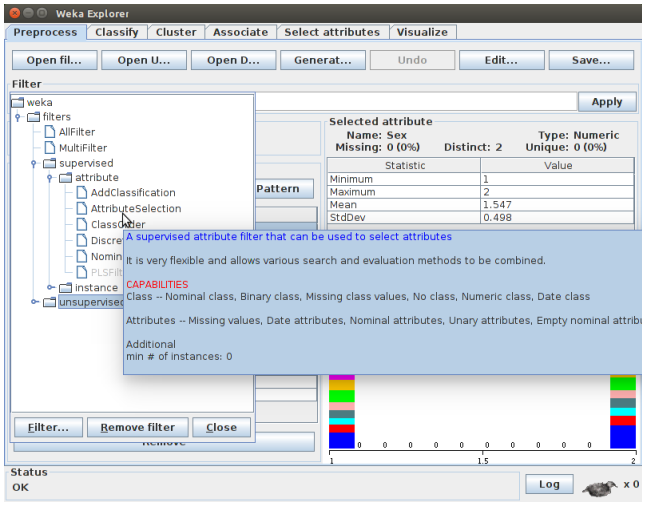
4.NumericToNominal

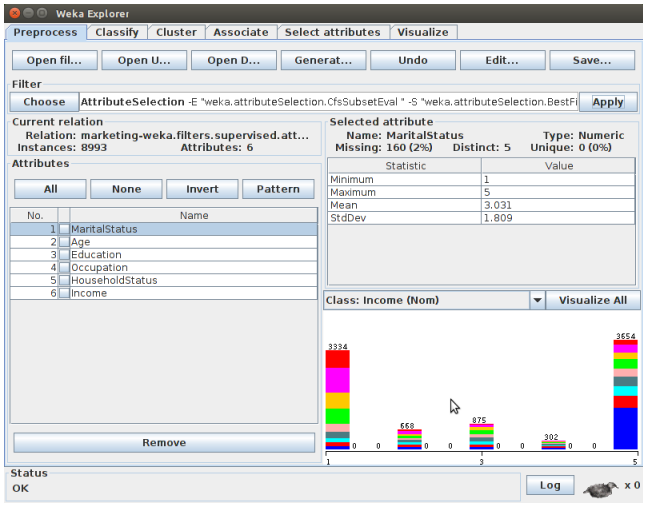
6.**Attribute Selection**

Preprocess(tab)

supervised.attribute.AttributeSelection

Default Evaluator:CfsSubsetEval method:BestFit





Final set of attributes have only **6** in number