EXPERIMENT -5
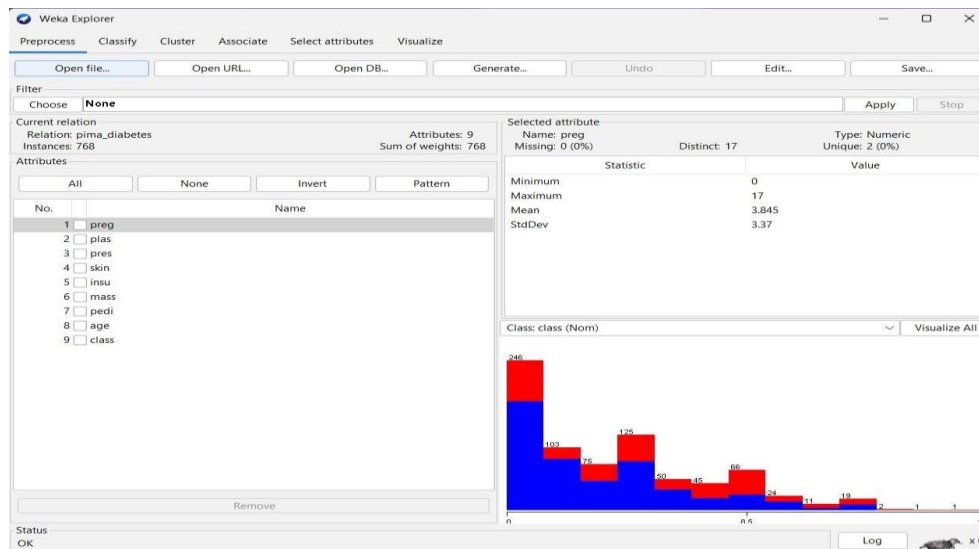
DATA PREPROCESSING AND ANALYSIS FOR DATASET USING WEKA

OUTPUT:

Preprocess   Classify   Cluster   Associate   Select attributes   Visualize

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)
18:16:47 - EM
18:19:31 - EM

Clusterer output

```
std. dev.              19.4099  12.9920  19.099

insu
  mean              173.816   71.039        0
  std. dev.         145.660   62.7051   0.0008

mass
  mean              34.6052   29.458   31.9674
  std. dev.          7.6535    8.1157    7.0137

pedi
  mean               0.642    0.4023    0.3789
  std. dev.          0.4342   0.2257    0.2266

age
  mean              36.6656   23.9583   38.9112
  std. dev.         10.7088    2.5897   12.6593

class
  tested_negative   108.1931  236.1879  158.619
  tested_positive   142.6434   19.676   108.6807
  [total]           250.8365  255.8639  267.2997


Time taken to build model (full training data) : 0.63 seconds

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

Clustered Instances

0      228 ( 30%)
1      203 ( 26%)
2      337 ( 44%)


Log likelihood: -24.97229
```

Status
OK   Log   x 0