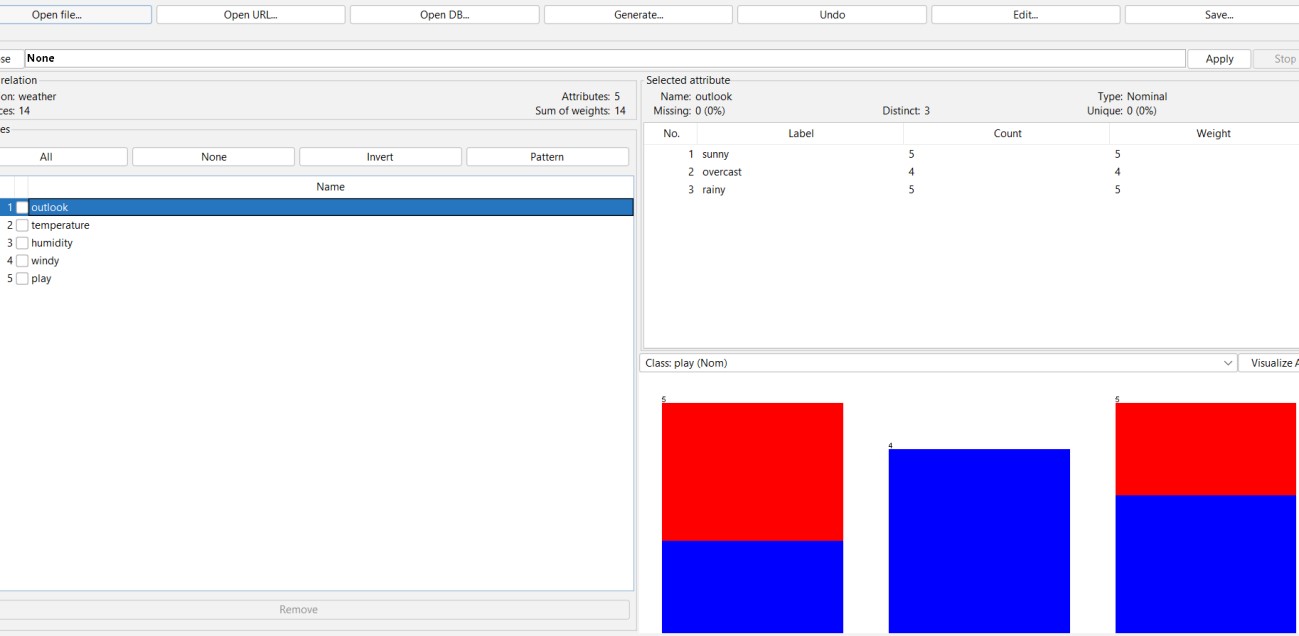
Dataware house data miniming for scientific applications

Lab experiments

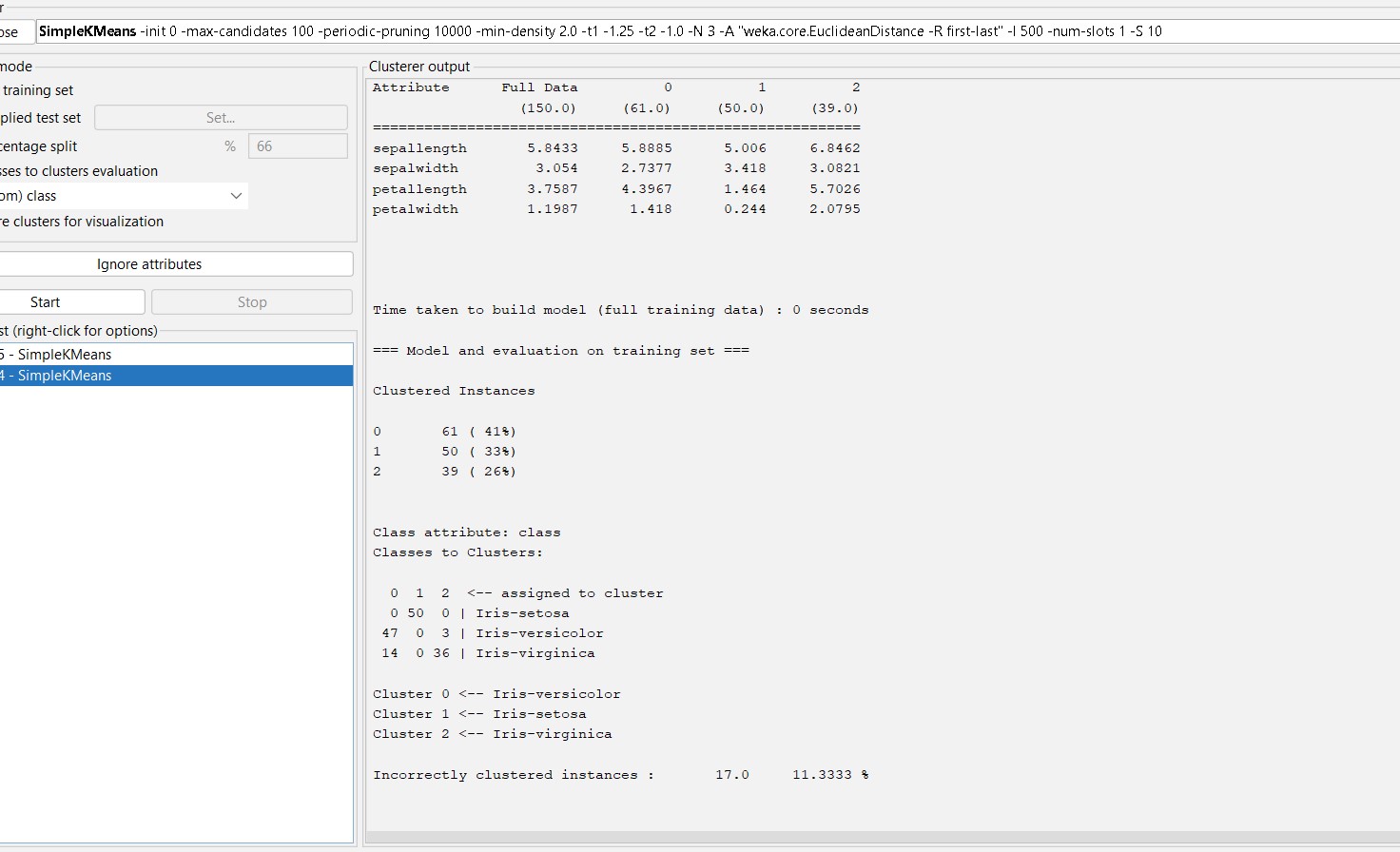
192011265

VR THARUN KUMAR REDDY

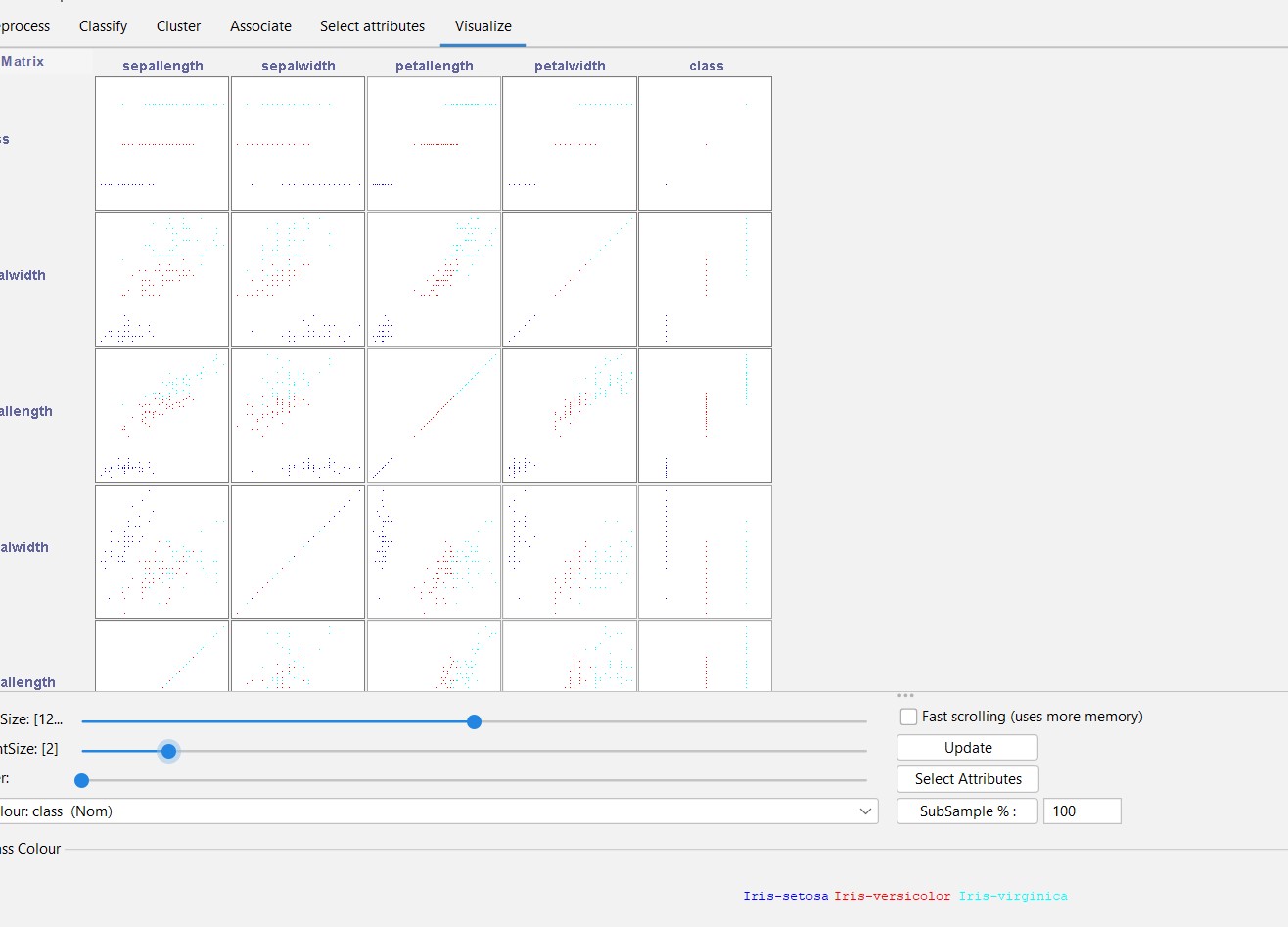
1. PREPROCESSING OF DATA USING WEKA



1. K-MEANS CLUSTERING BY WEKA

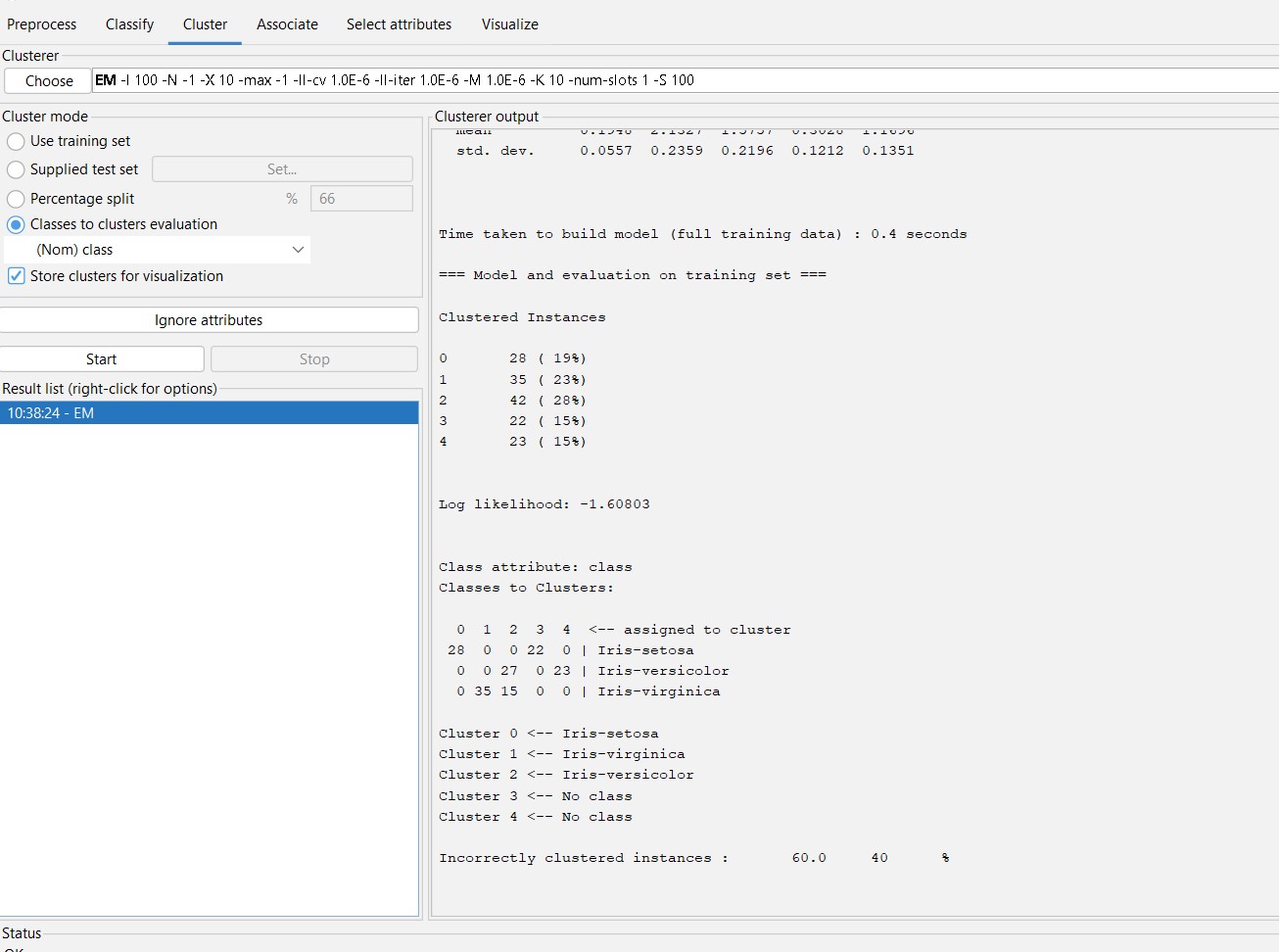


OUTPUT INFORMATION OF SIMPLE K MEANS CLUSTERING –CLUSTERER OUTPUT

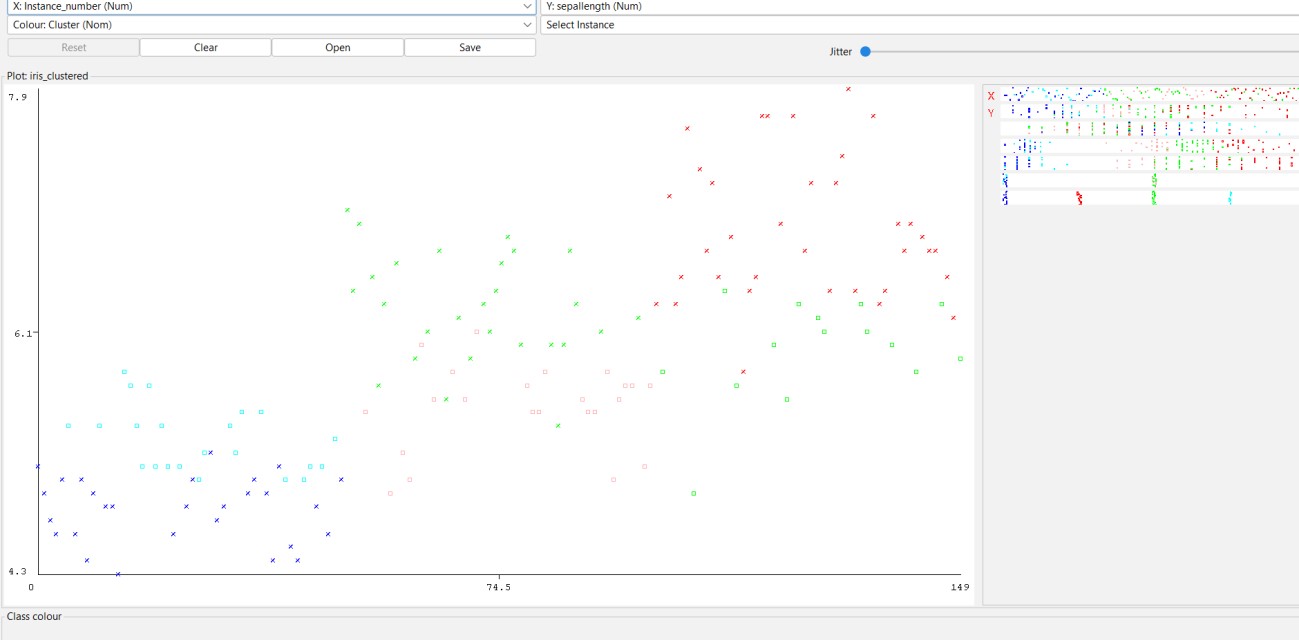


K MEANS CLUSTERING IN WEKAA USING IRIS INFORMATION – PLOTTING ..

9. DATA ANALYSIS BY EXPECTION MAXIMISATION ALGORITHM USING WEKA

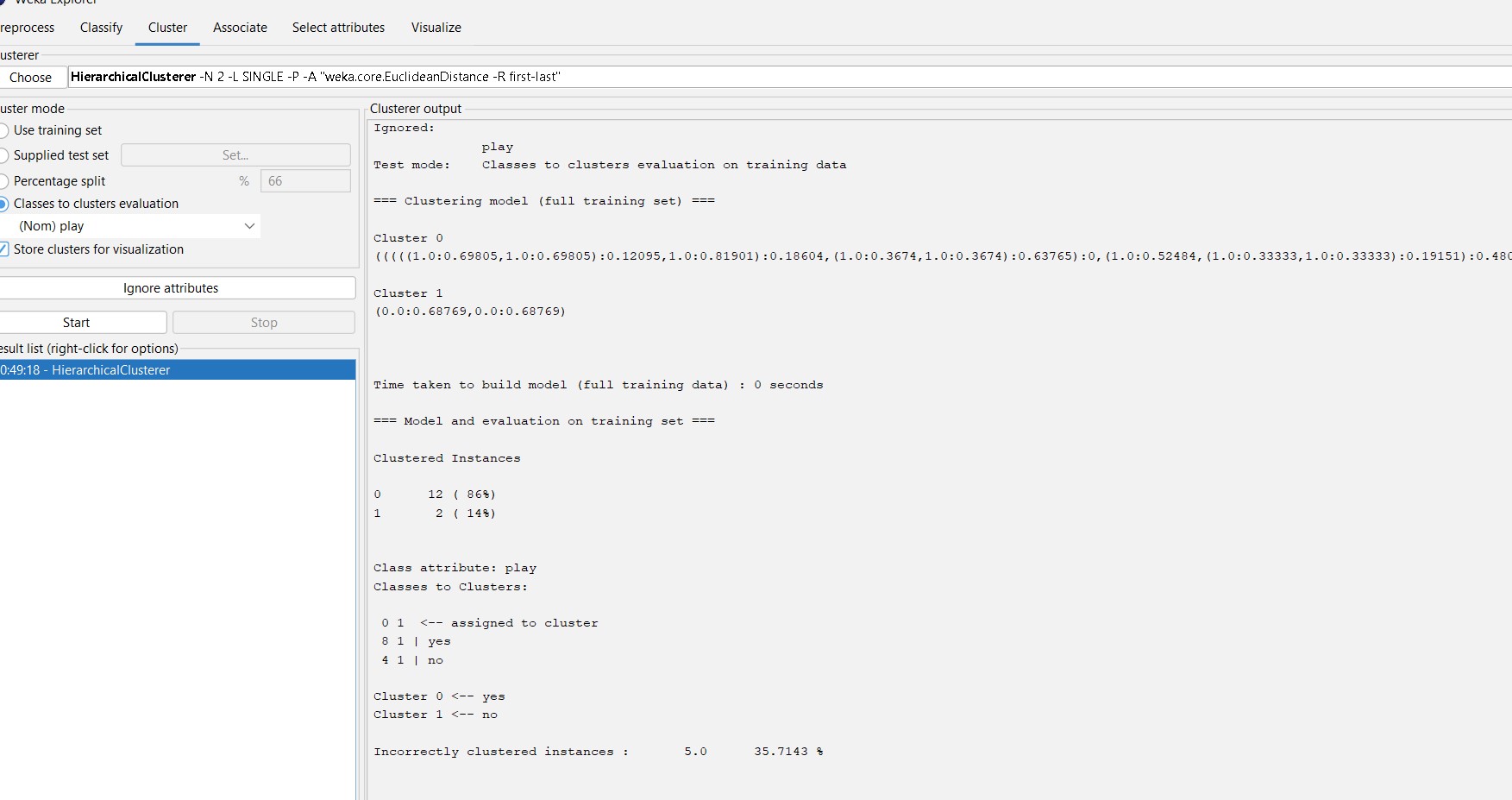


DATA ANALYSIS BY EXPECTATION MAXIMISATION ALGORITHM – OUTPUT INNFO OF IRIS

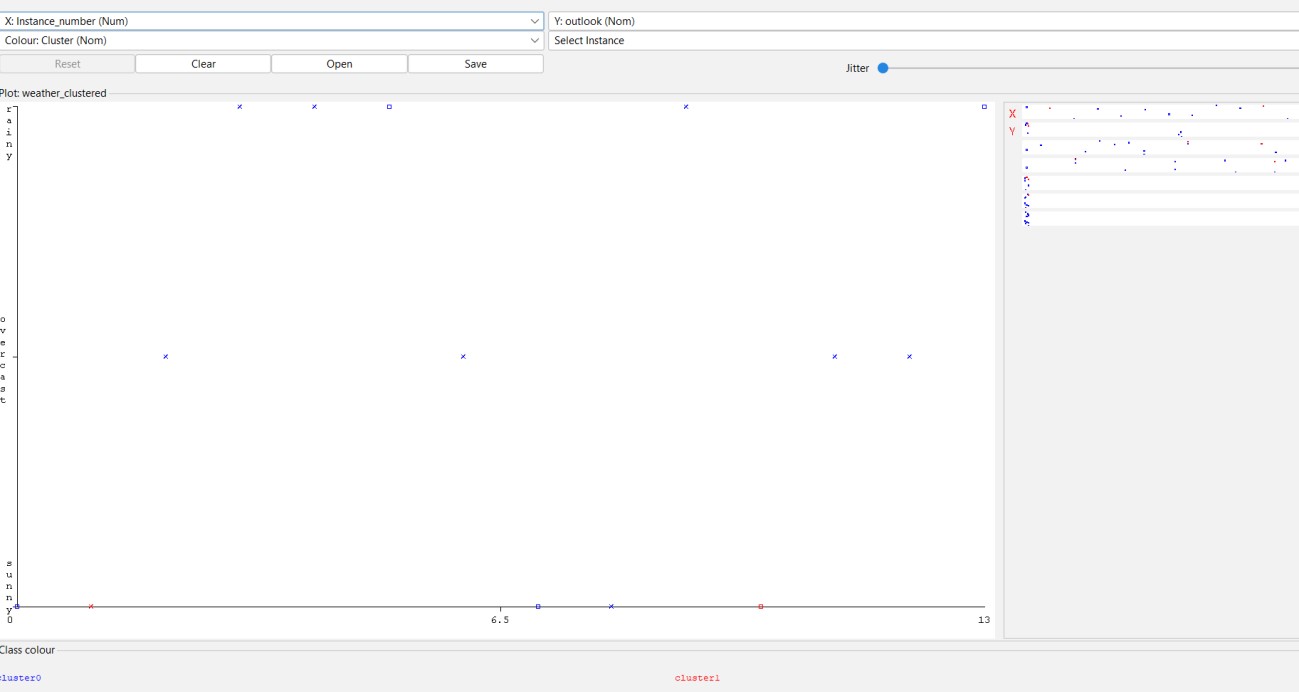


DATA ANALYSIS BY EXPECTATION MAXIMISATION ALGORITHM—PLOT DIAGRAM OF IRIS INFO

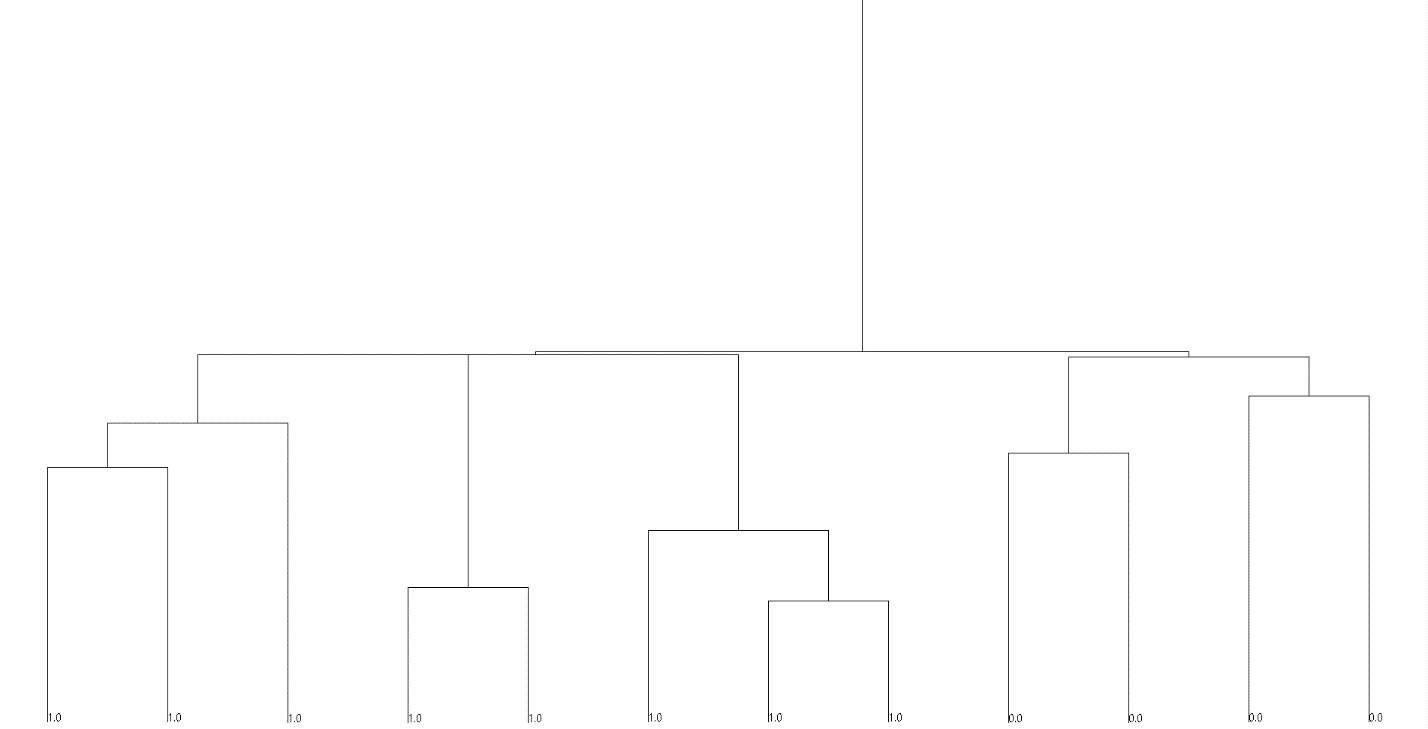
10.DATA ANALYSIS BY HIERARCHICAL CLUSTERING. IN WEKA



OUTPUT INFORMATION

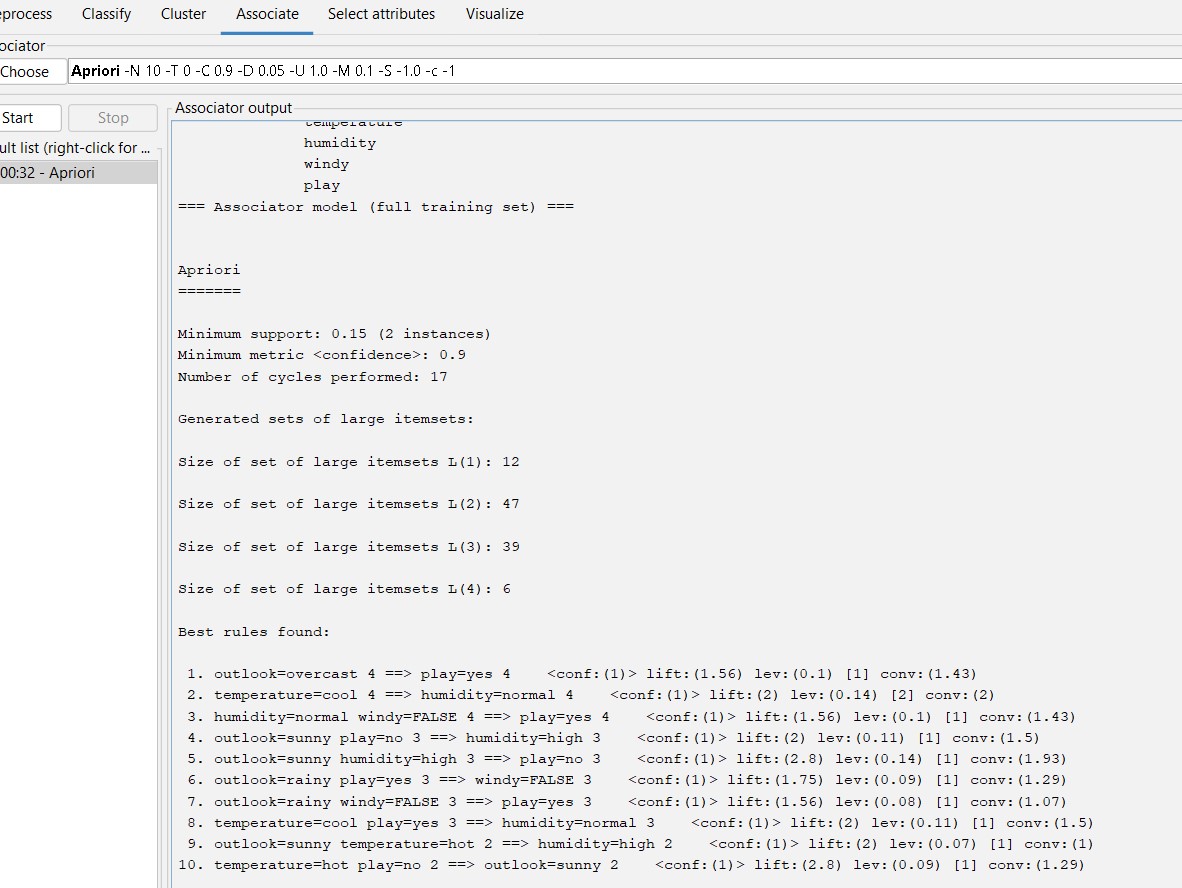


VISUALISING THE PLOT GRAPH OF CLUSTER ASSIGNMENTS



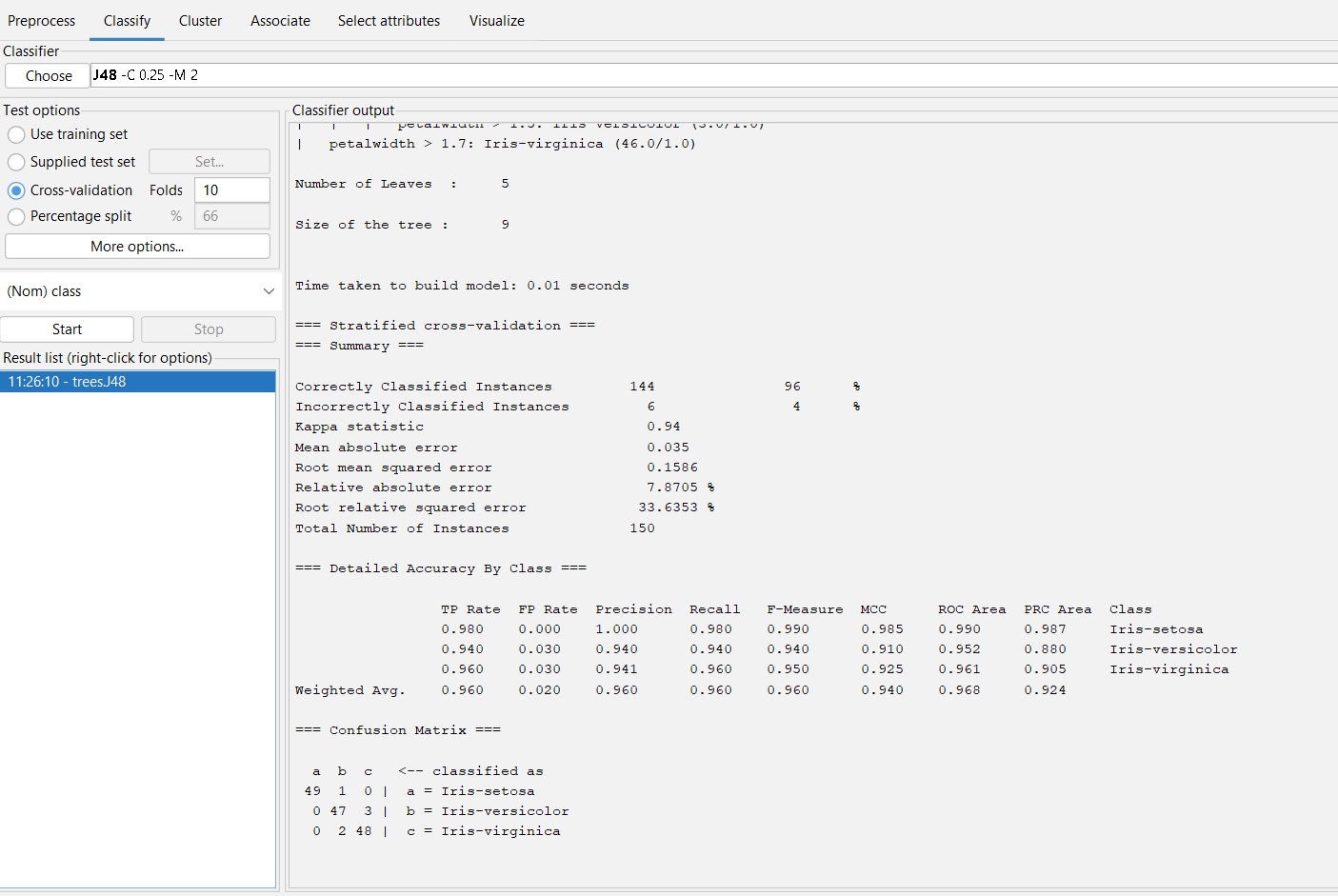
HIERARCHICAL TREE FOR DATA ANALYSIS.

11.KNOWLEDGE MINING FOR ASSOCIATION RULES USING WEKA

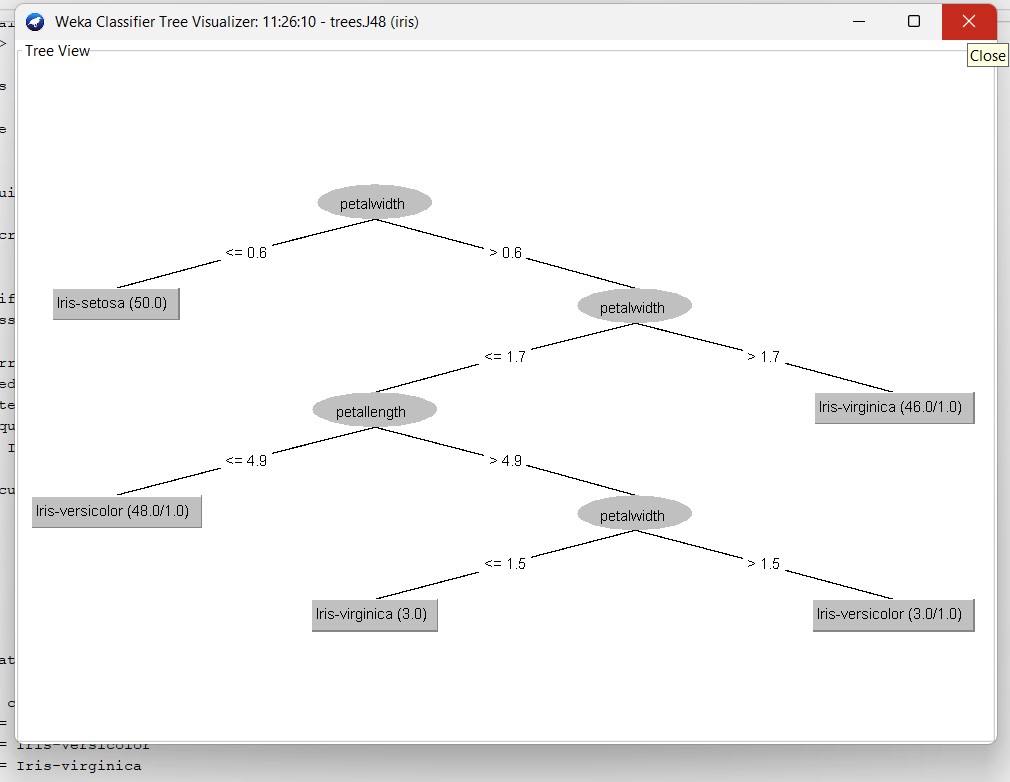


OUTPUT INFORMATION USING WEATHER NOMINAL DATA

15. EVALUATING THE ACCURACY OF THE CLASSIFIERS



OUTPUT INFORMAATION OF IRIS IN ACCURACY FINDING.

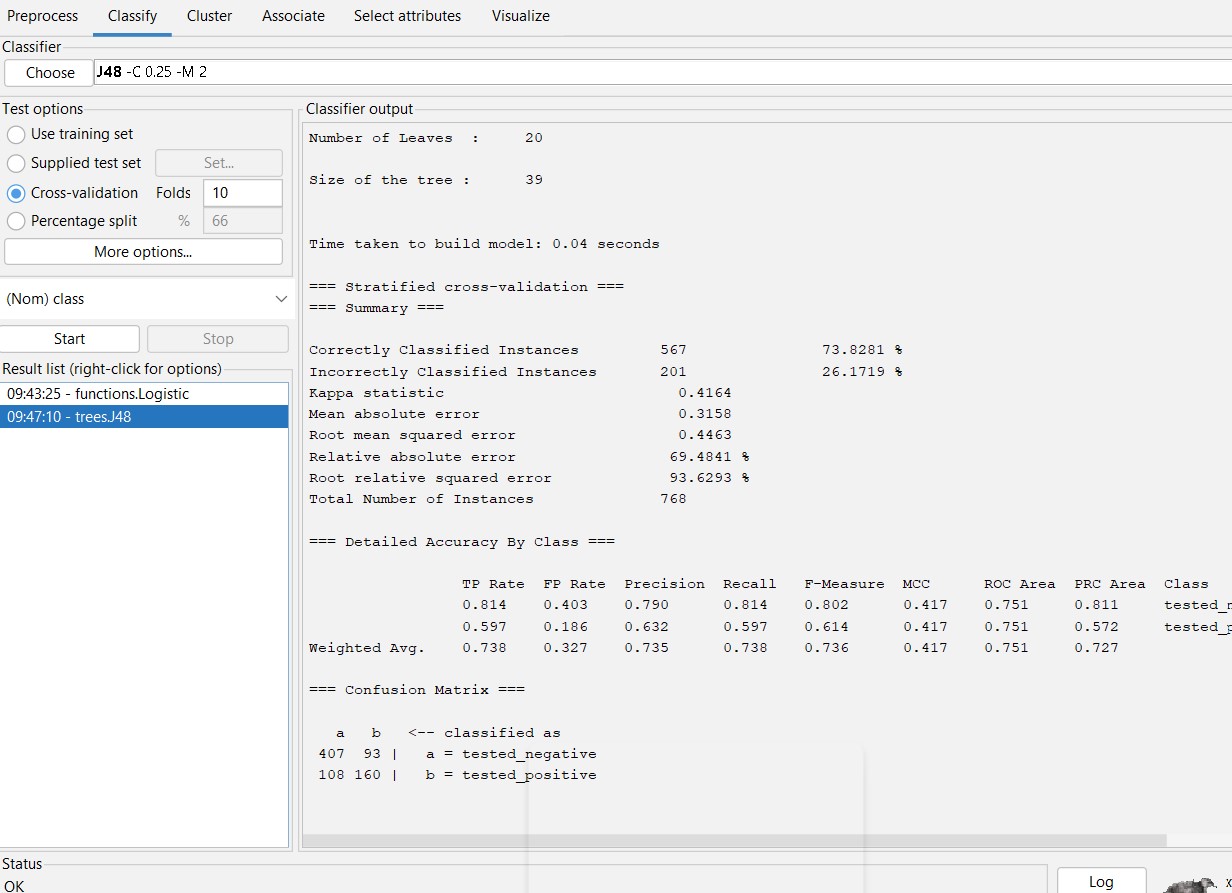


TREE DIAGRAM OF ACCURACY WITH IRIS INFO IN WEKA

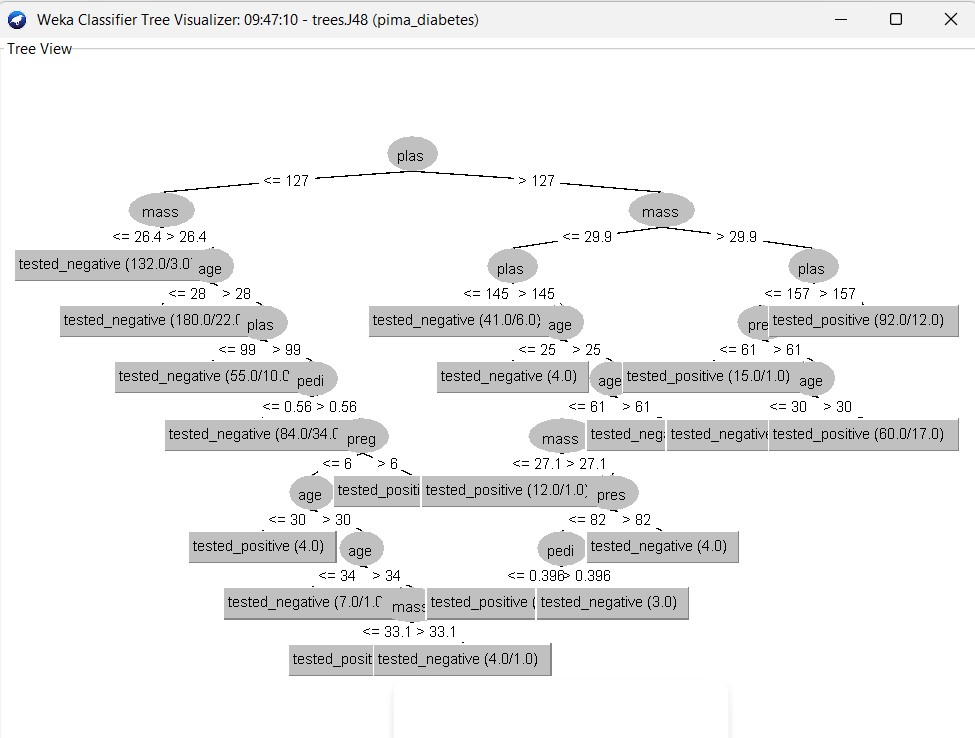


ERROR ACCURACY RATE OF IRIS IN WEKA.

1. **CREATE THE ARFF FILE FOR THE DIABETES DATABASE AND PERFORM THE RULE BASED CLASSIFICATION.**

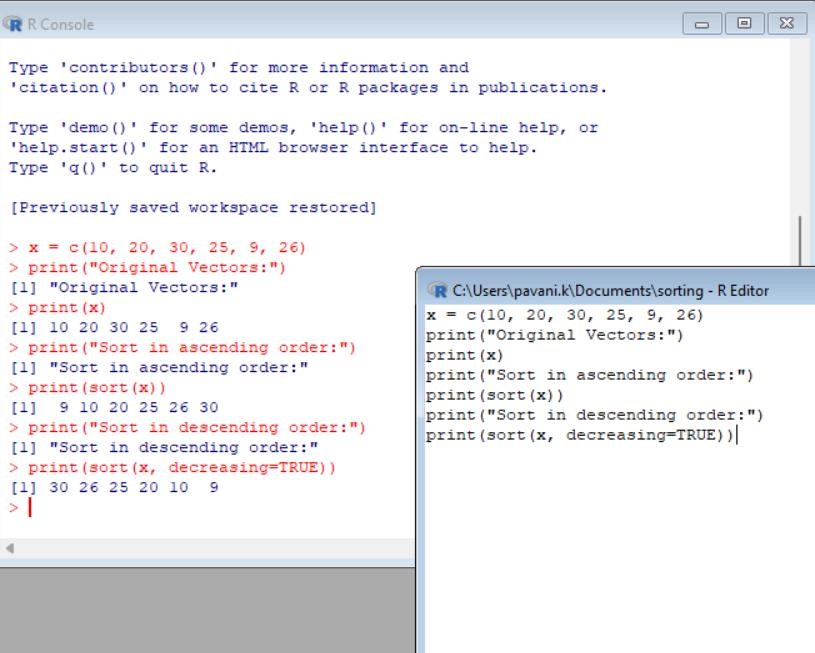


**Output info of diabetes by rule based classification.**

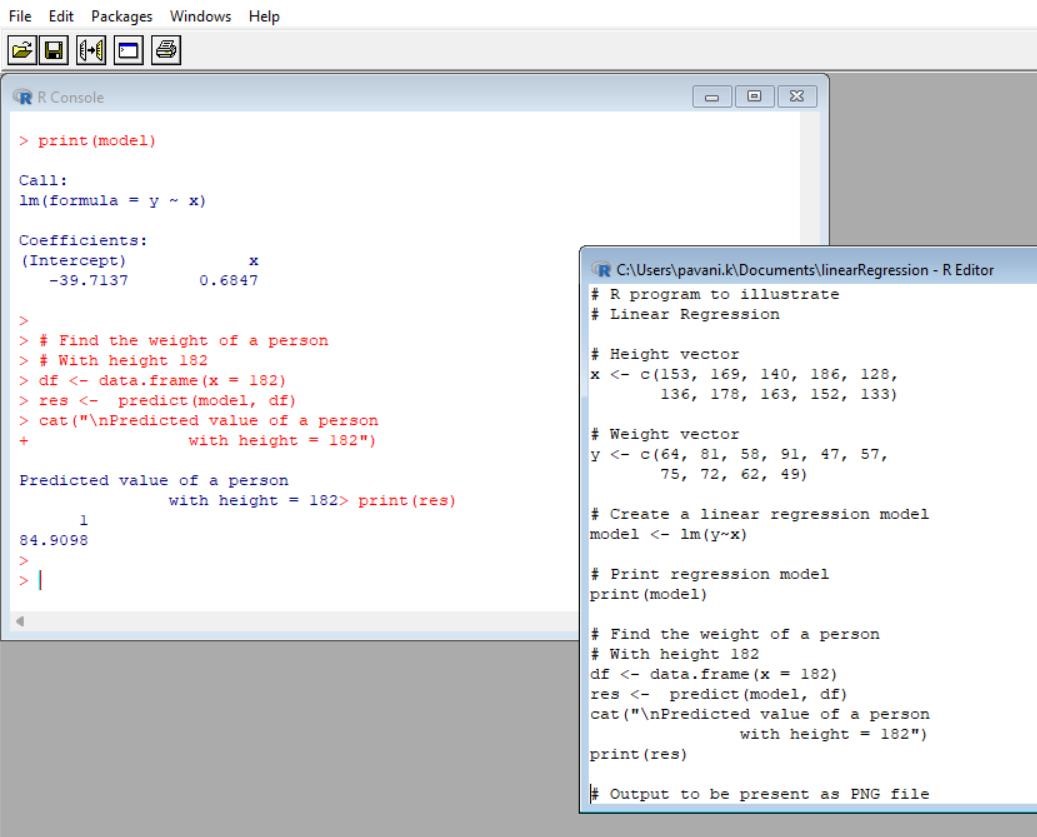


**Tree view of diabetes on rule based classification.**

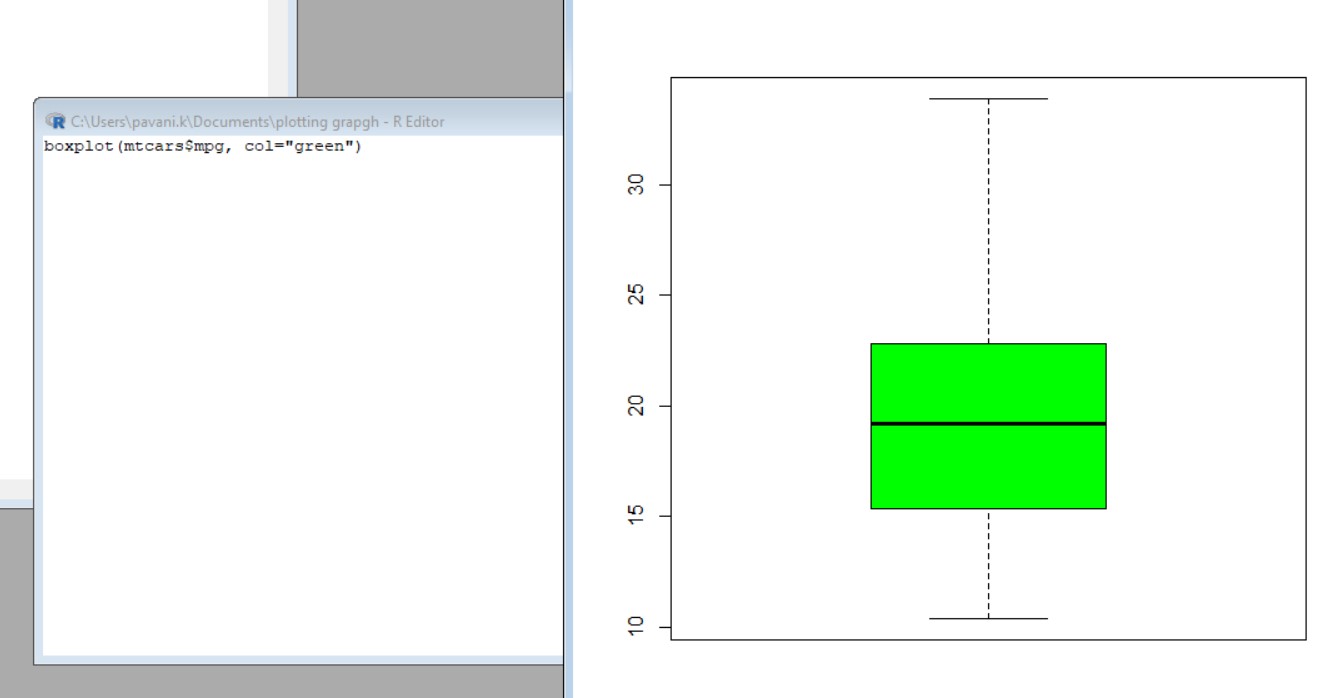
\*. R PROGRAM FOR SORTING VECTOR



2.R PROGRAM FOR LINEAR REGRESSION.

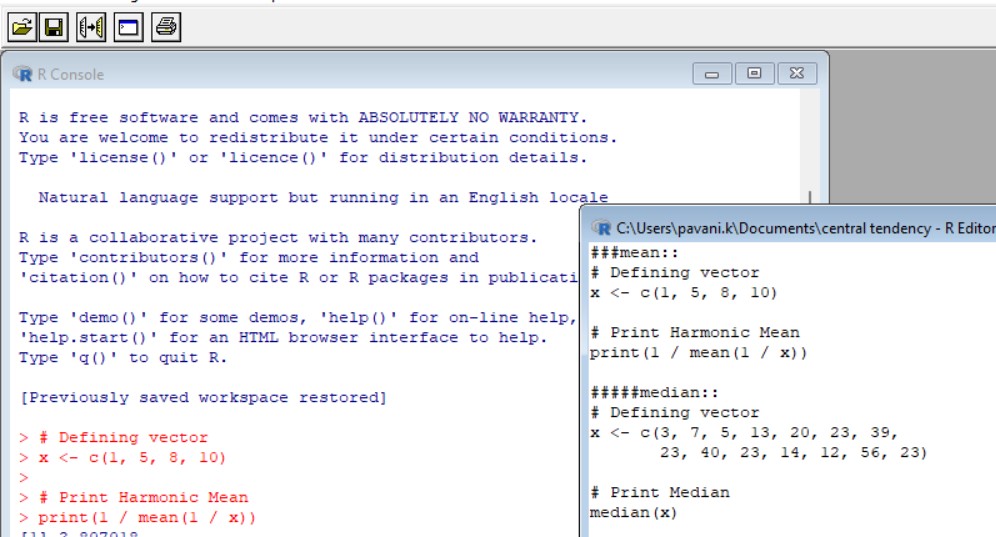


3.PLOTTING GRAPH

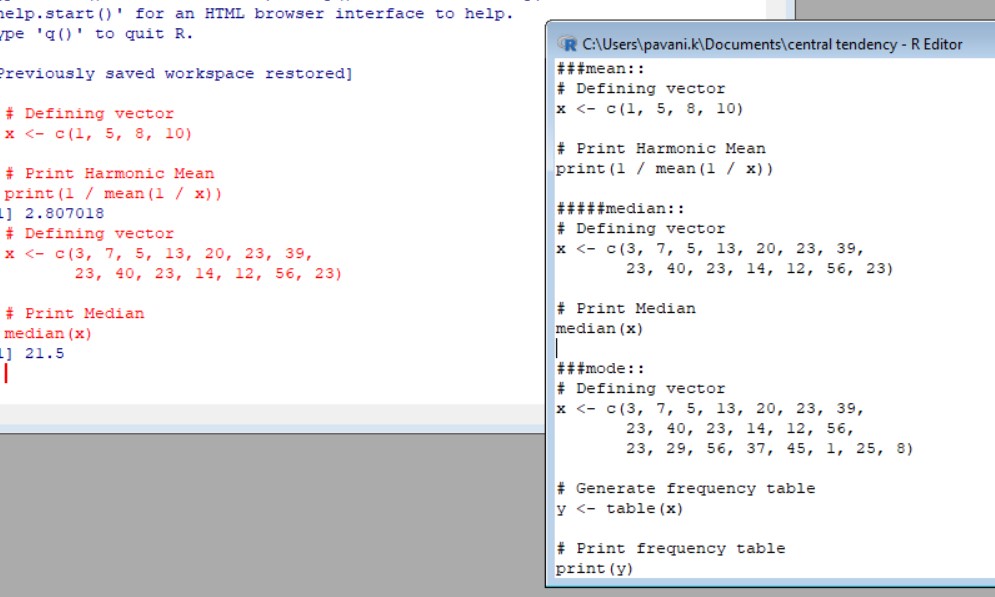


4.

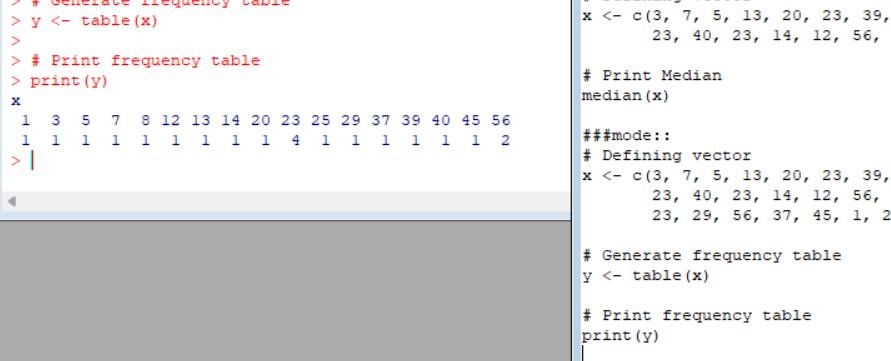
1. CENTRAL TENDENCY—MEAN:



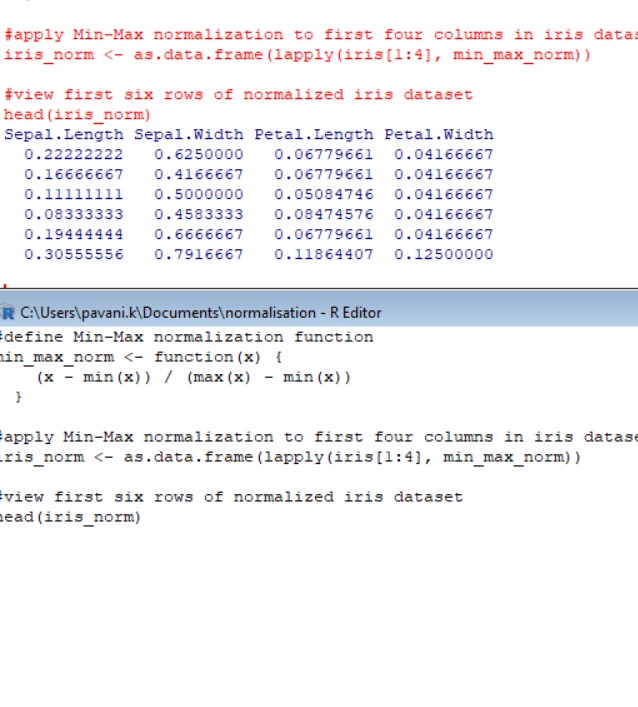
1. CENTRAL TENDENCY –MEDIAN:



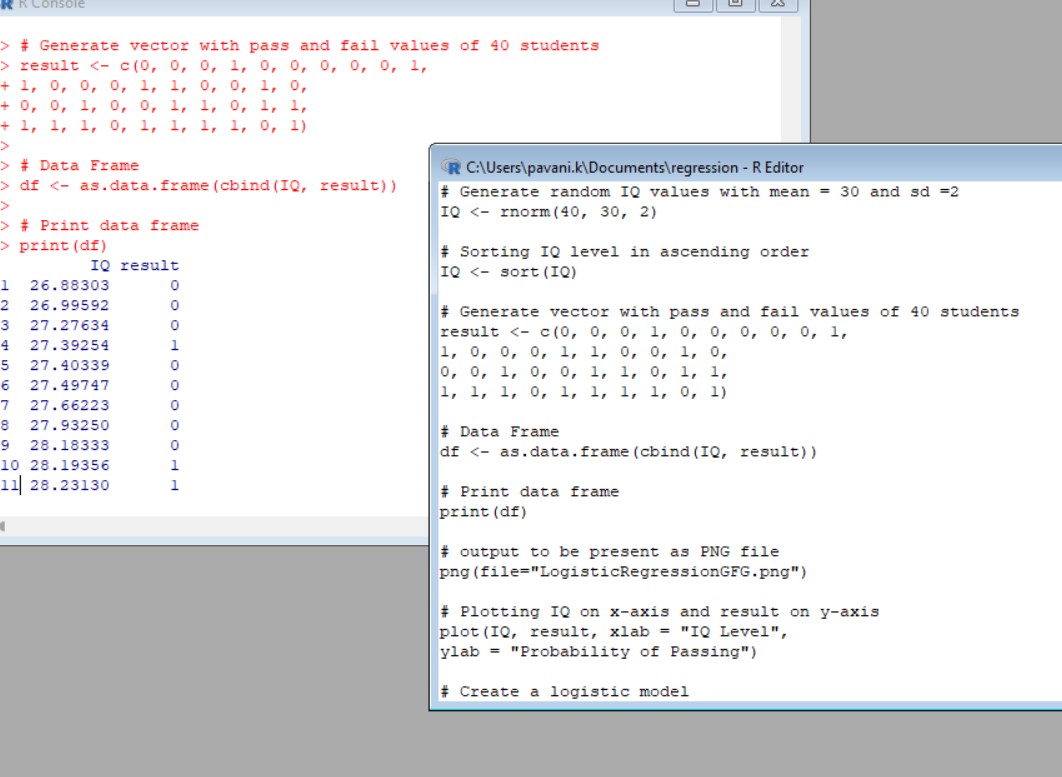
1. CENTRAL TENDENCY—MODE:



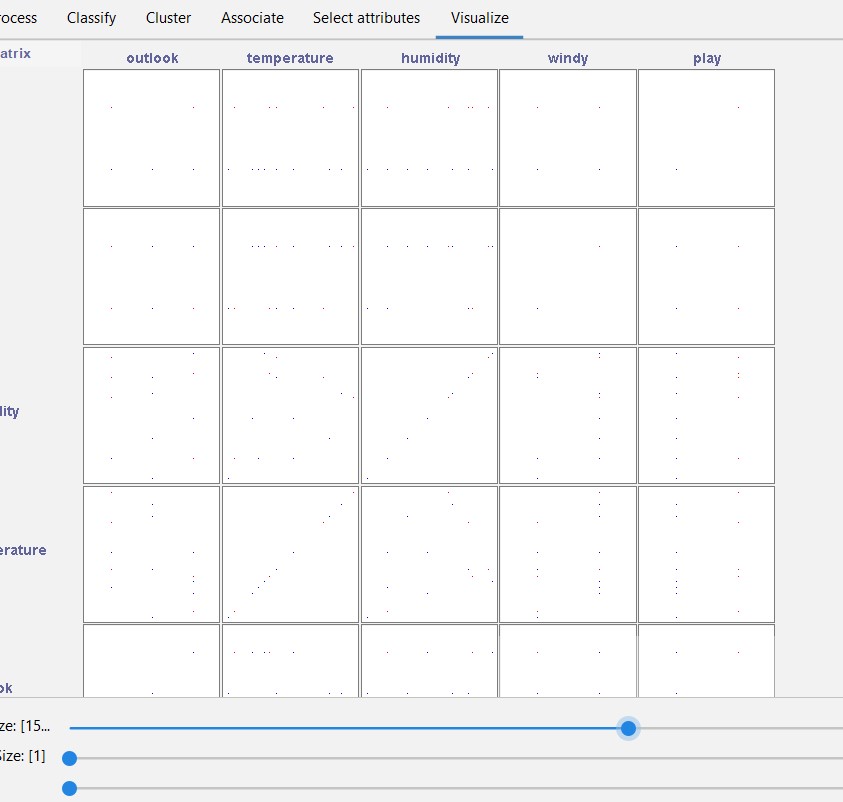
5.NORMALISATION AND ANALYSIS:



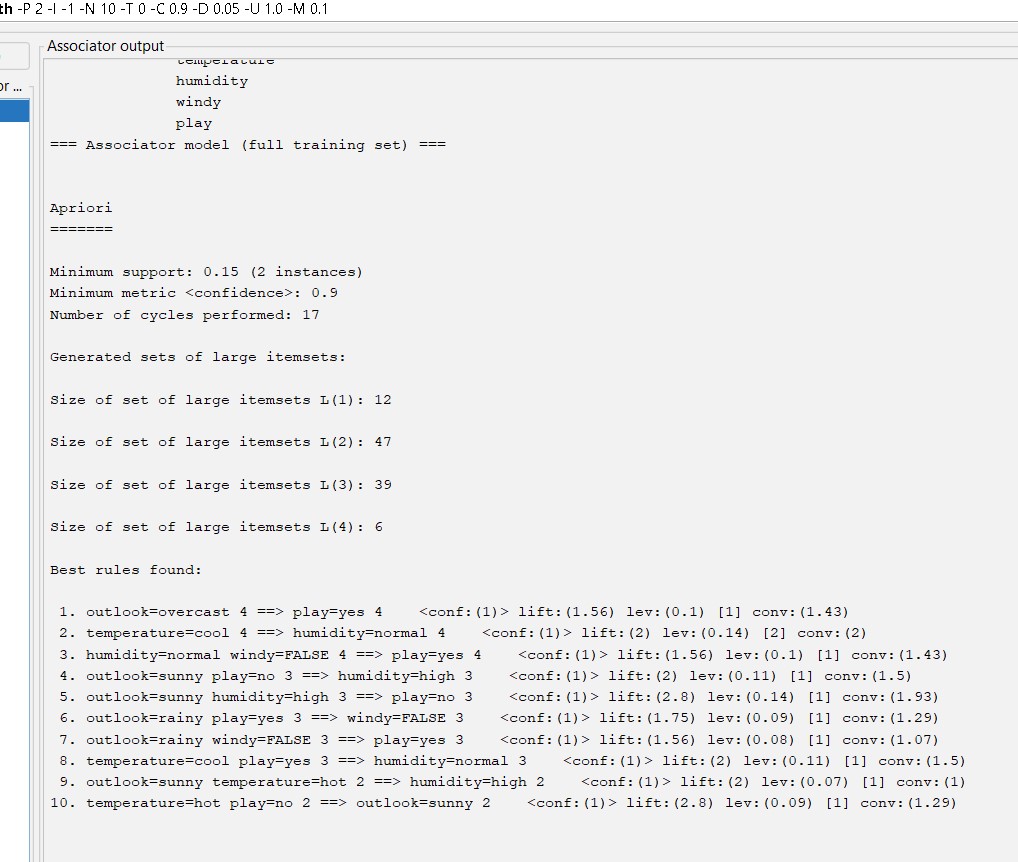
6. REGRESSION:



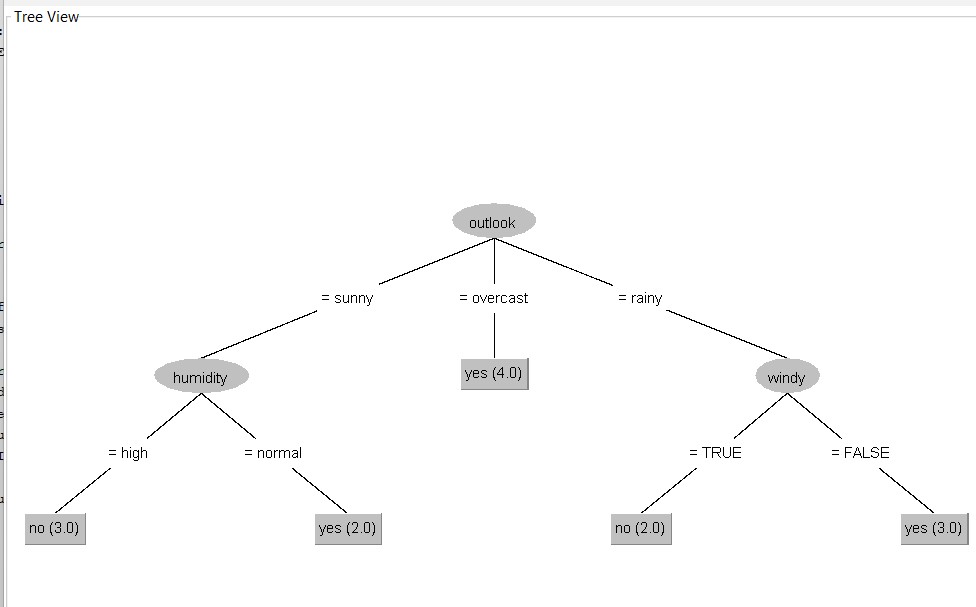
12.FP GROWTH



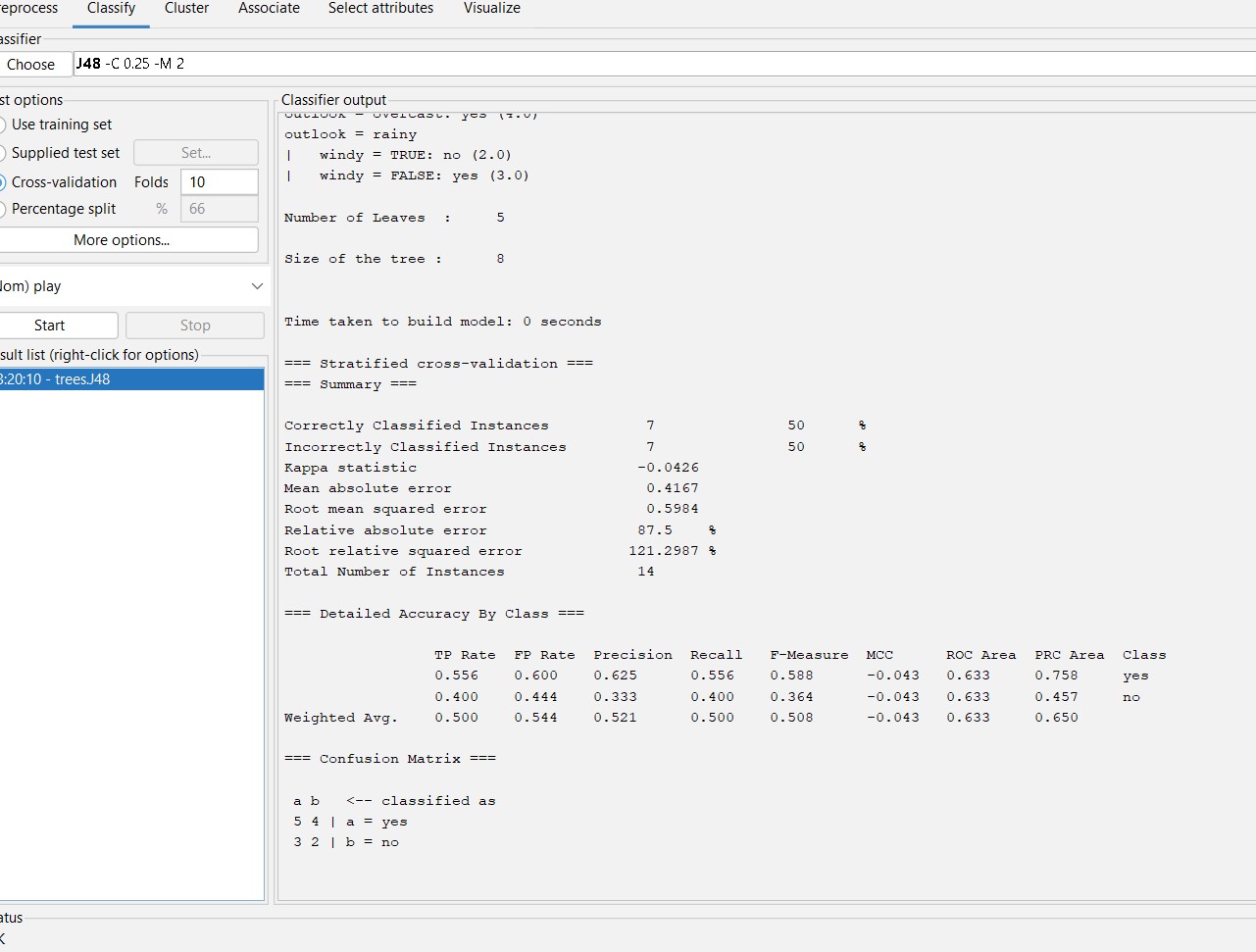
FP GROWTH VISUALISATION



13.DECISION TREE



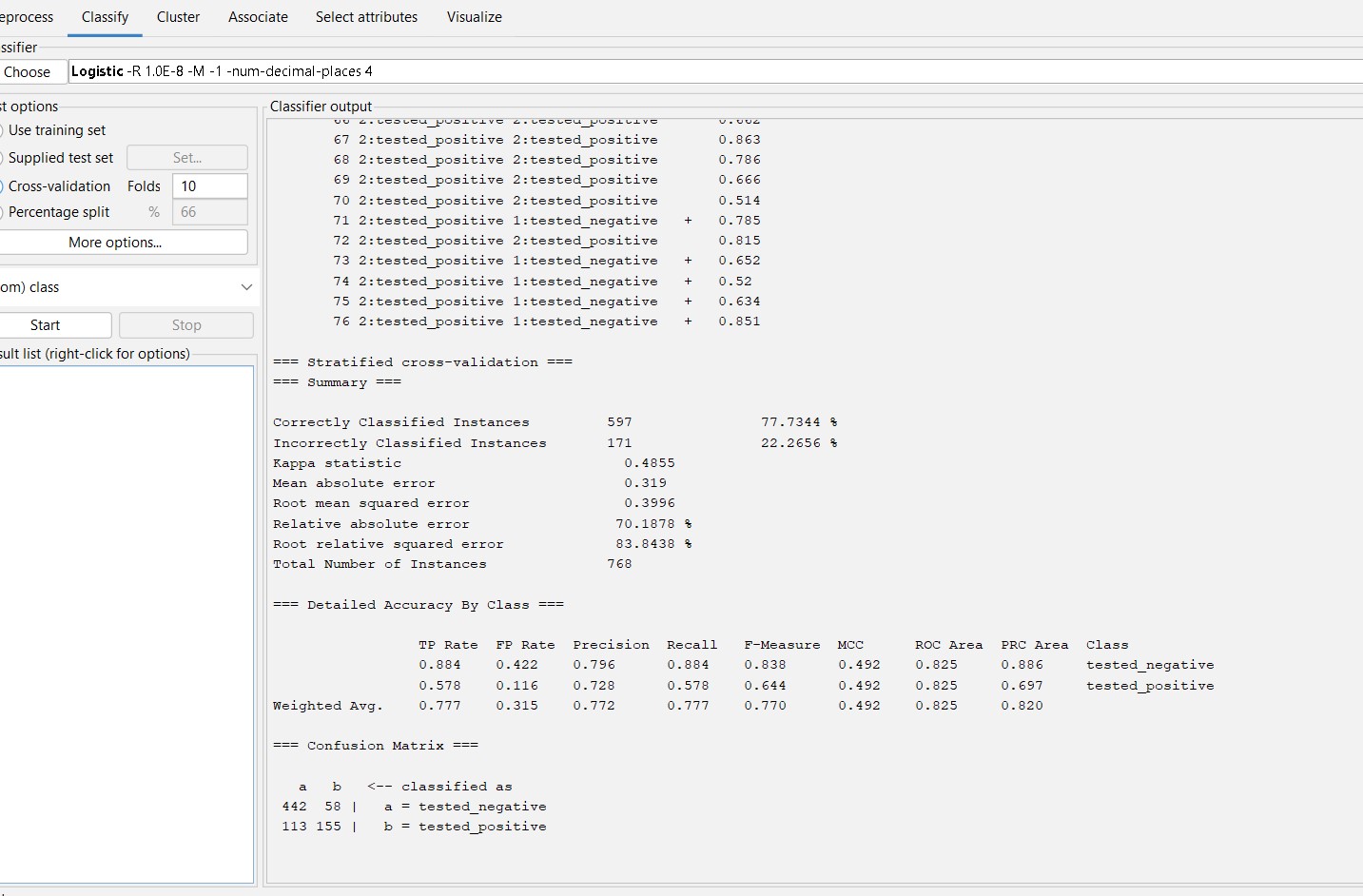
TREE VISUALISATION



OUTPUT INFO.

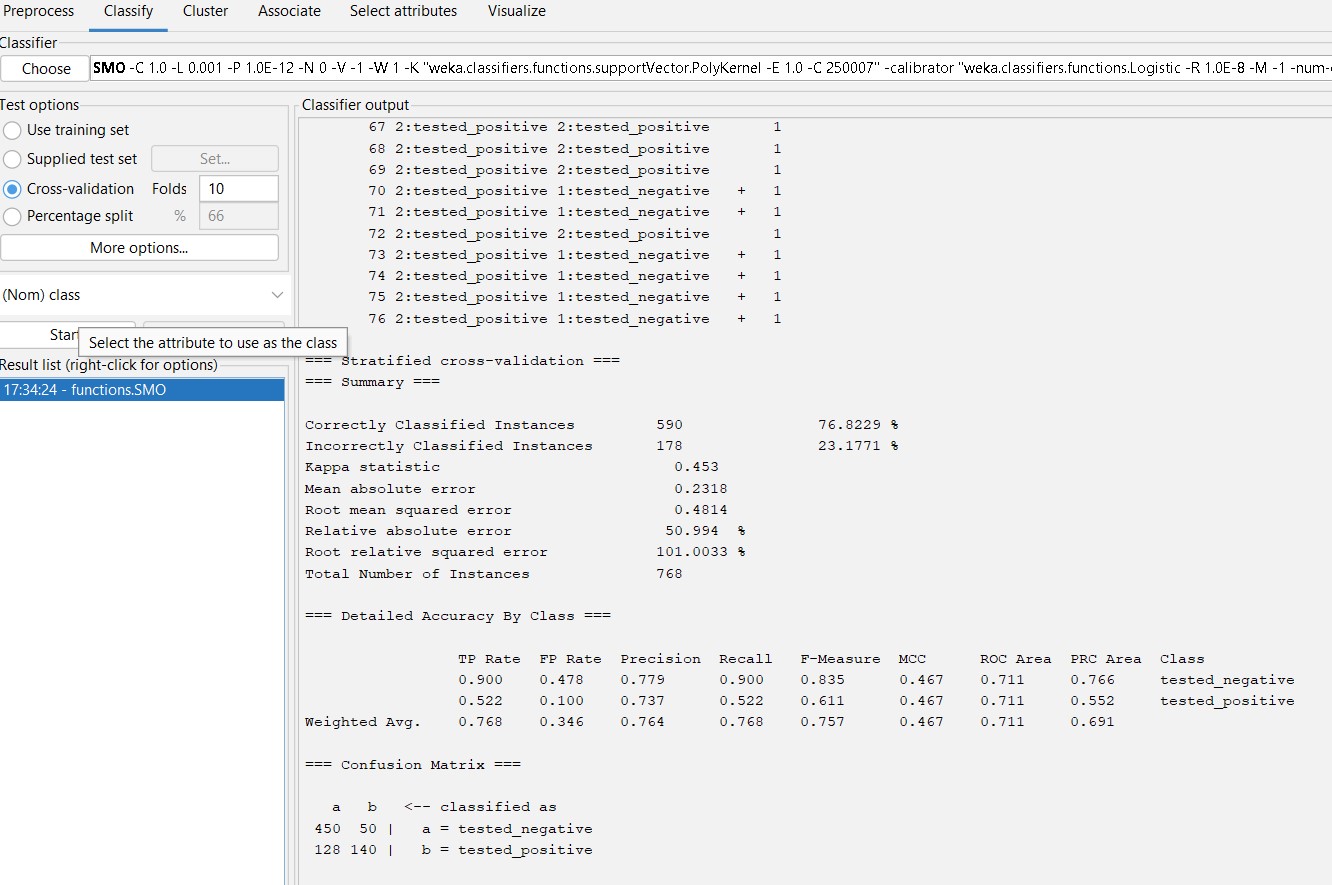
16. **CREATE THE ARFF FILE FOR THE DIABETES DATABASE AND**

**PERFORM SVM BASED CLASSIFICATION**



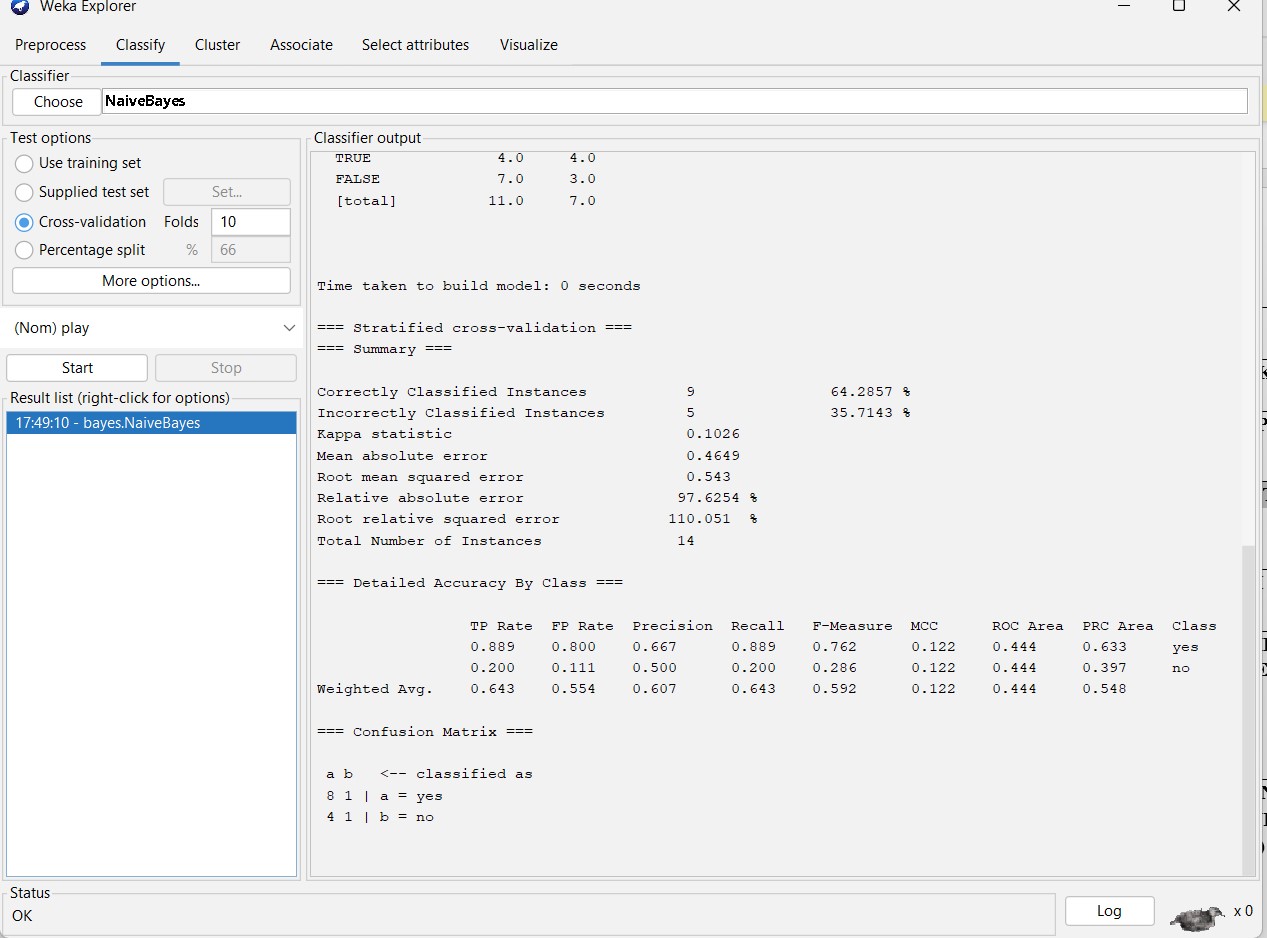
Output information of data diabetes with logistics.

14. **PREDICTION OF CATEGORICAL DATA USING SMO ALGORITHM USING WEKA.**



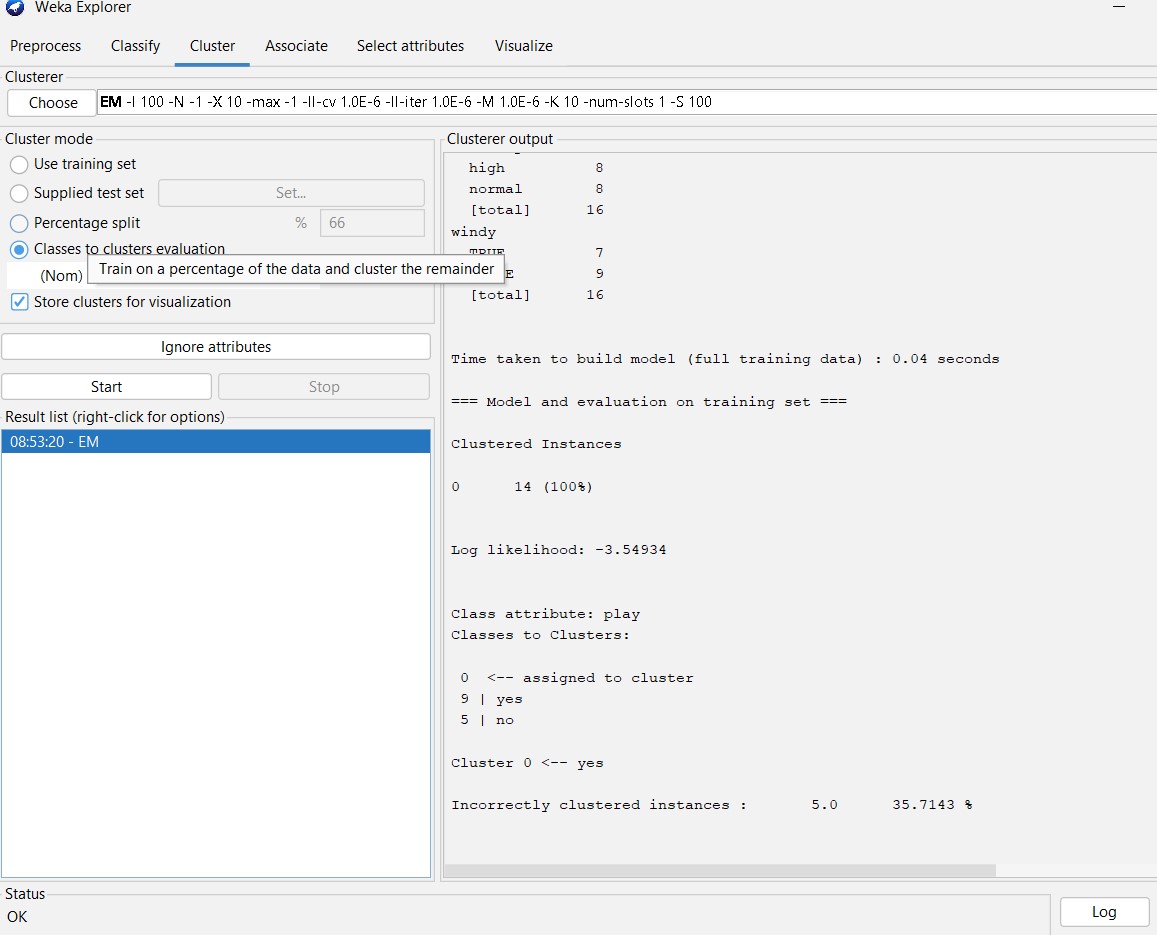
The output information of diabetes data using smo..

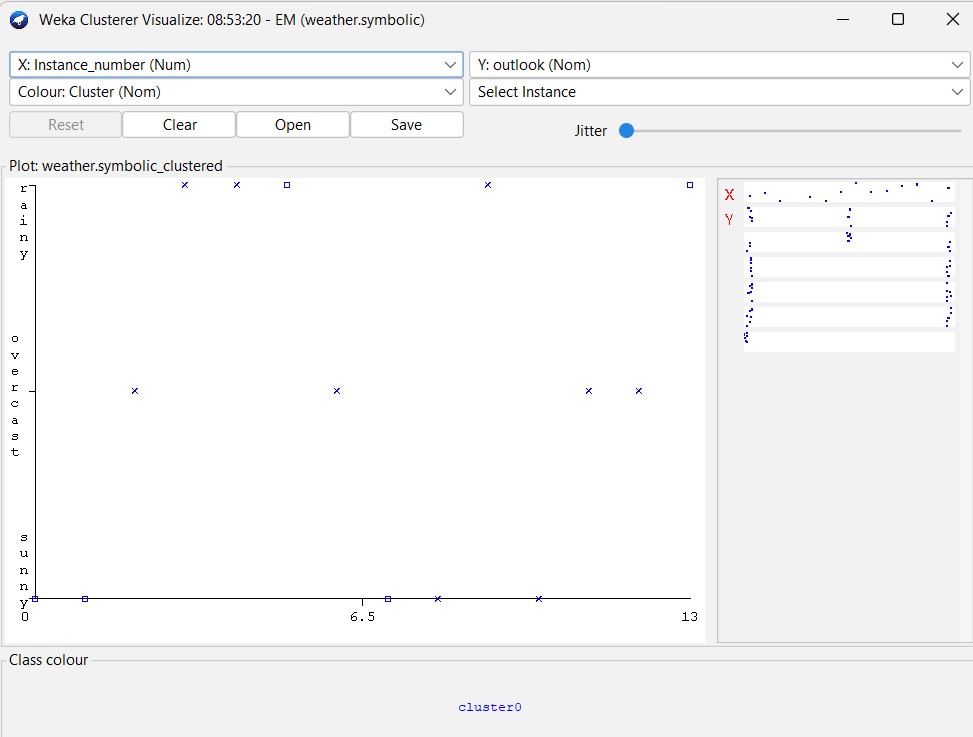
1. **PREDICTION OF CATEGORICAL DATA USING BAYESIAN ALGORITHM USING WEKA**



OUTPUT INFORMATION WEATHER NUMERIC BASED ON NAVIES BAYES RULE..

1. **DATA ANALYSIS BY DENSITY BASED CLUSTERING ALGORITHM USING WEKA.**

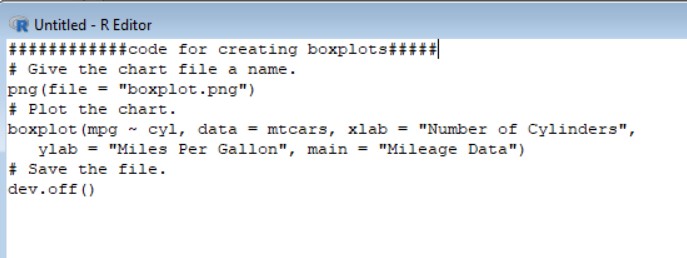




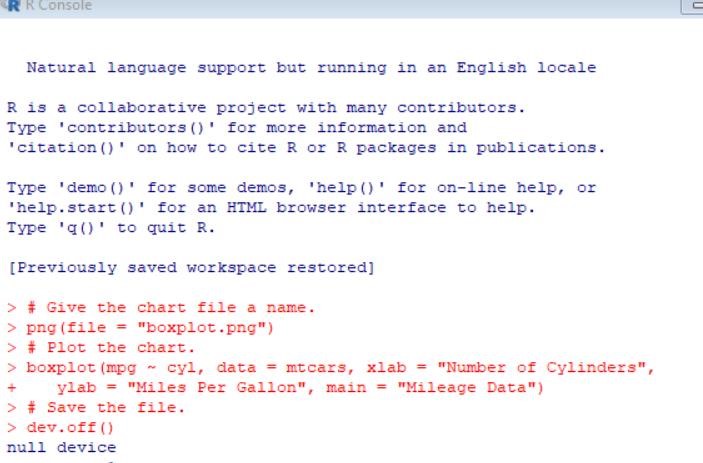
Output information…

1. **CREATE A BOXPLOT GRAPH FOR THE RELATION BETWEEN "MPG"(MILES PER**

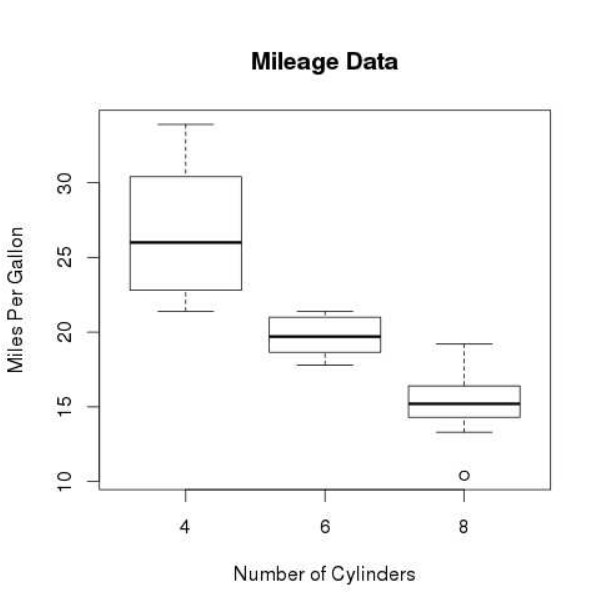
**GALLOON) AND "CYL"(NUMBER OF CYLINDERS) FOR THE DATASET "MTCARS" AVAILABLE IN R ENVIRONMENT**



Code for creating boxplots in r programing.



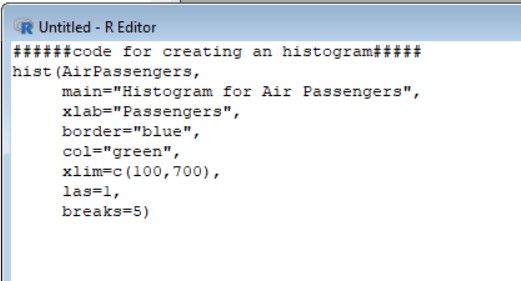
Output result of created boxplot..



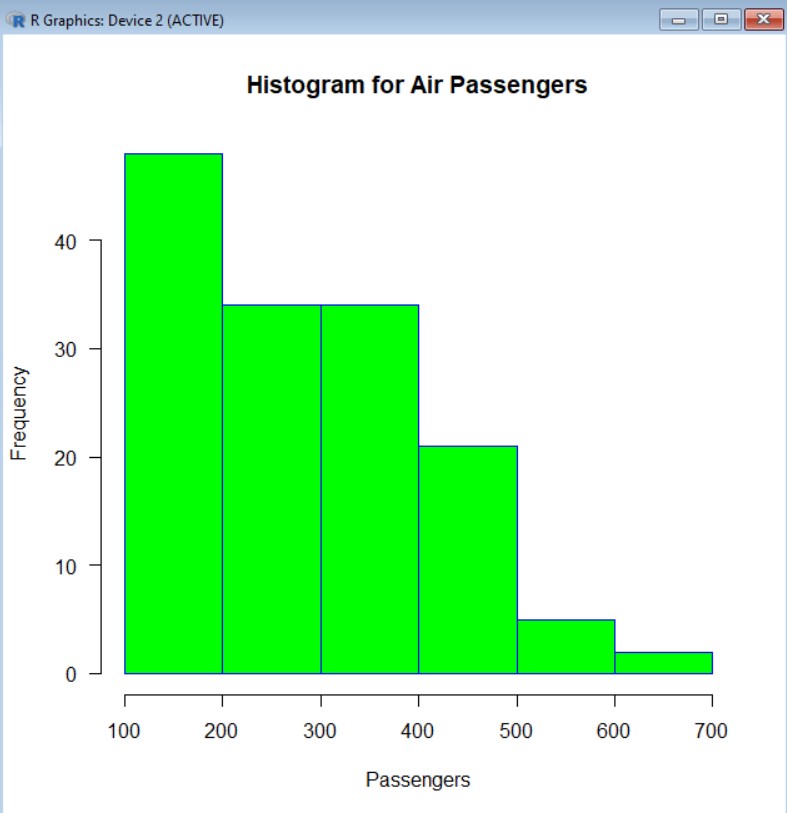
Output graph of boxplot..

1. **USING R PROGRAM MAKE A HISTOGRAM FOR THE**

**“AIRPASSENGERS “DATASET, START AT 100 ON THE X-AXIS, AND FROM VALUES 200 TO 700, MAKE THE BINS 150 WIDE**



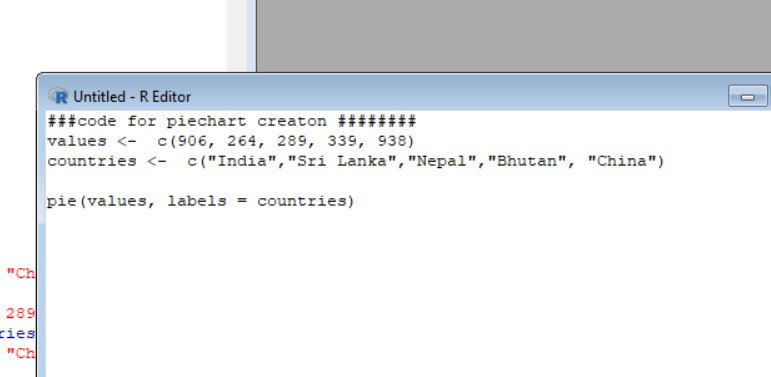
Code for creating an histogram..



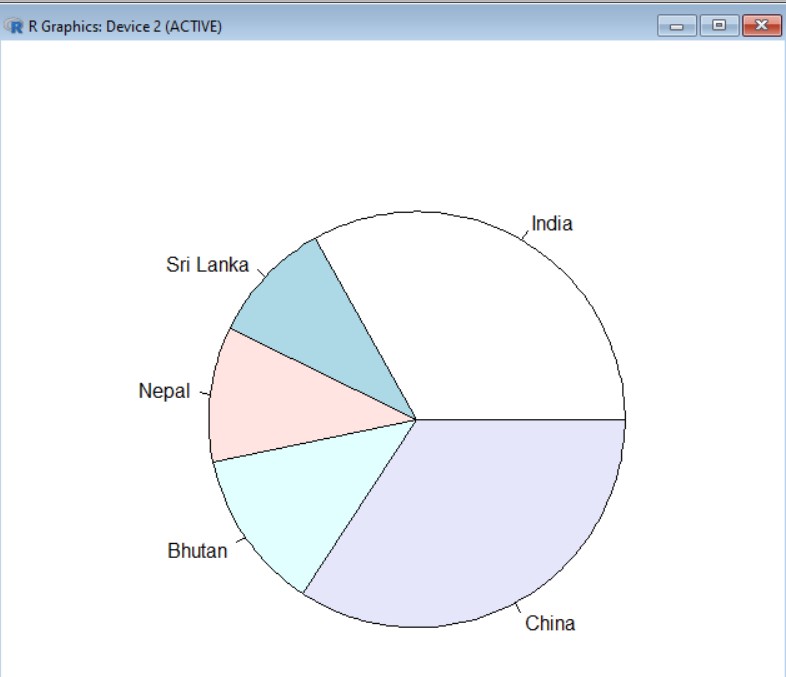
Output of histogram graphical representation..

1. **USING R PROGRAM CREATE A 3D PIE CHART FOR THE DATASET**

**“POLITICAL KNOWLEDGE” WITH SUITABLE LABELS AND COLOURS.**



**Code for creating the pie charts**

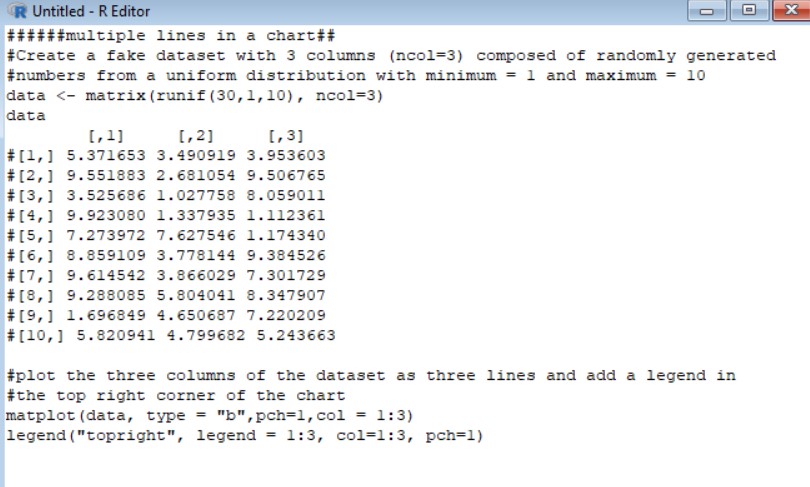


**Output of created pie chart.**

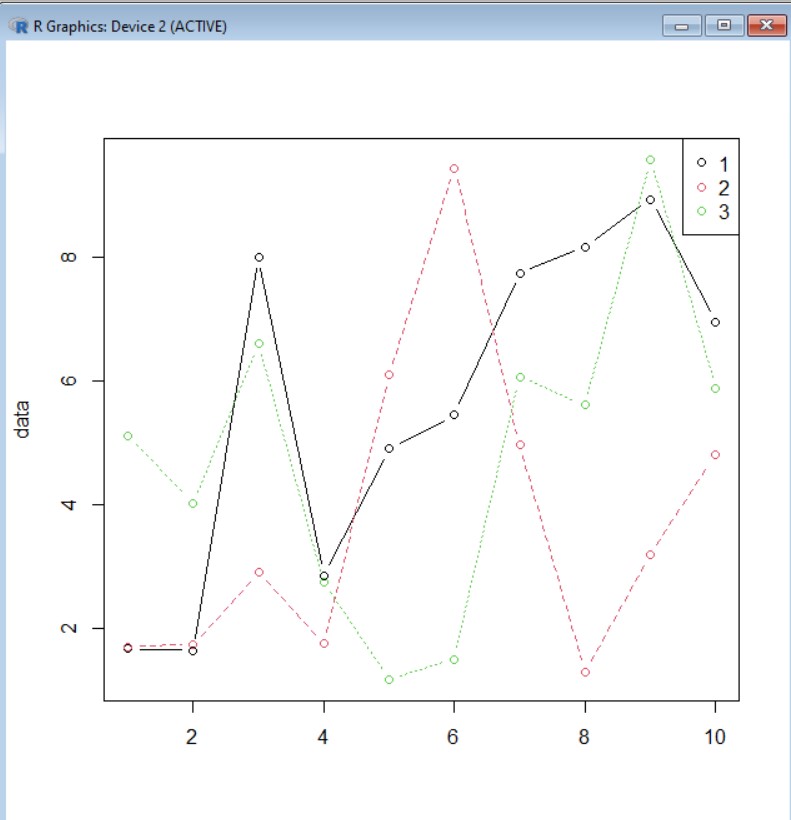
1. **OBTAIN MULTIPLE LINES IN LINE CHART USING A SINGLE PLOT**

**FUNCTION IN R.USE ATTRIBUTES “MPG”AND “QSEC” OF THE DATASET**

**“MTCARS”**



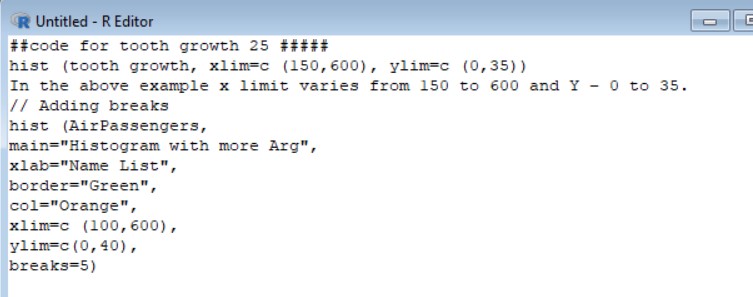
**code for getting multiple lines in chart.**



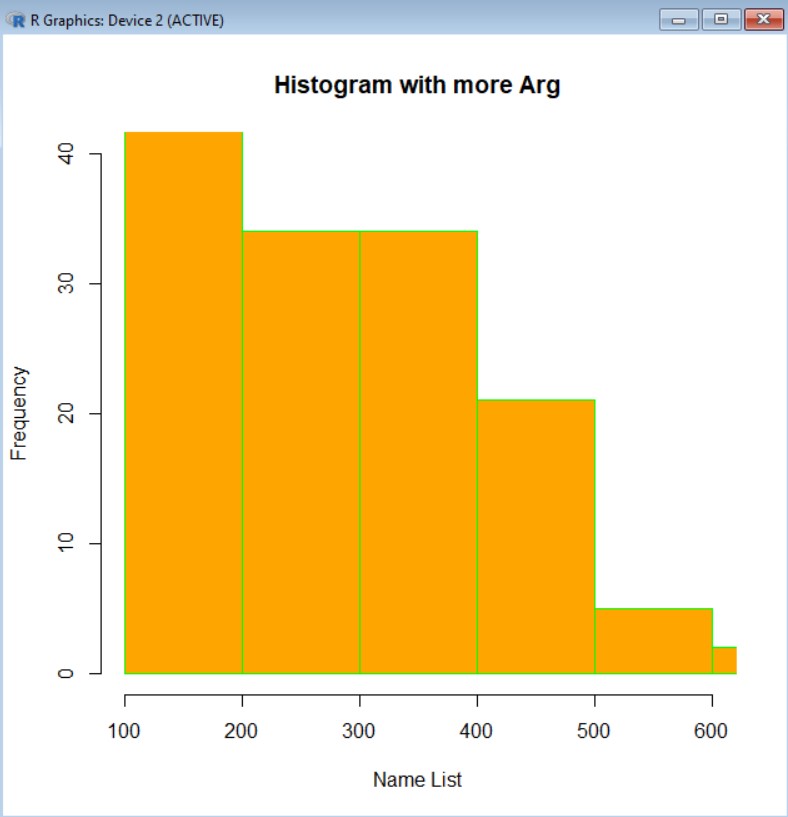
**Output of multiple lines in chart.**

1. **USING R PROGRAM MAKE A HISTOGRAM FOR THE “TOOTHGROWTH”**

**DATASET, START AT 100 ON THE X-AXIS, AND FROM VALUES 200 TO 700, MAKE THE BINS 150 WIDE**



**Code for creating a histogram for tooth growth.**



**Output of histogram, for tooth growth.**