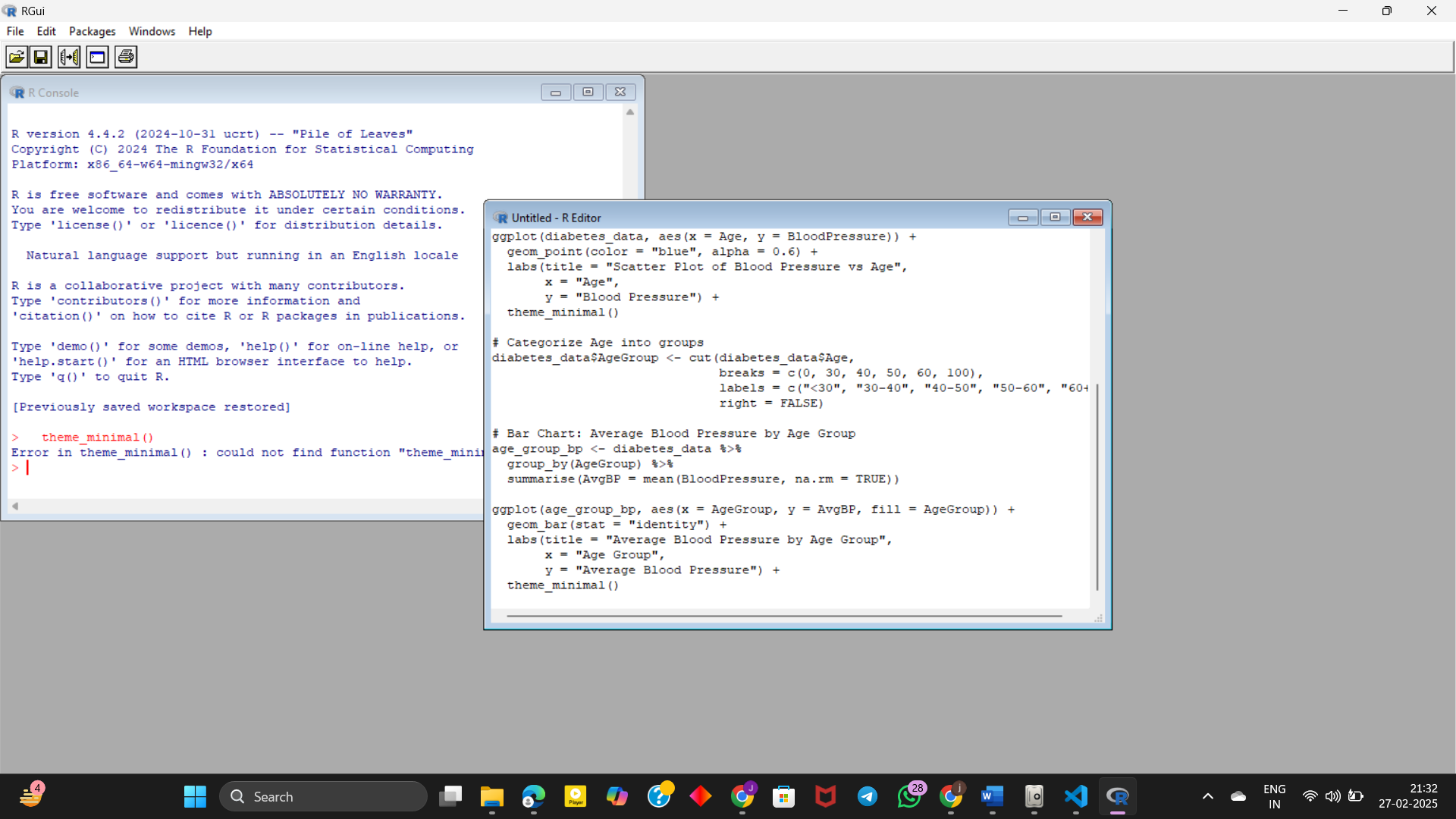
DWDM R PROGRAMMING-PRACTICALS

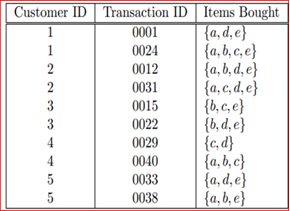
CH,JASWANTH

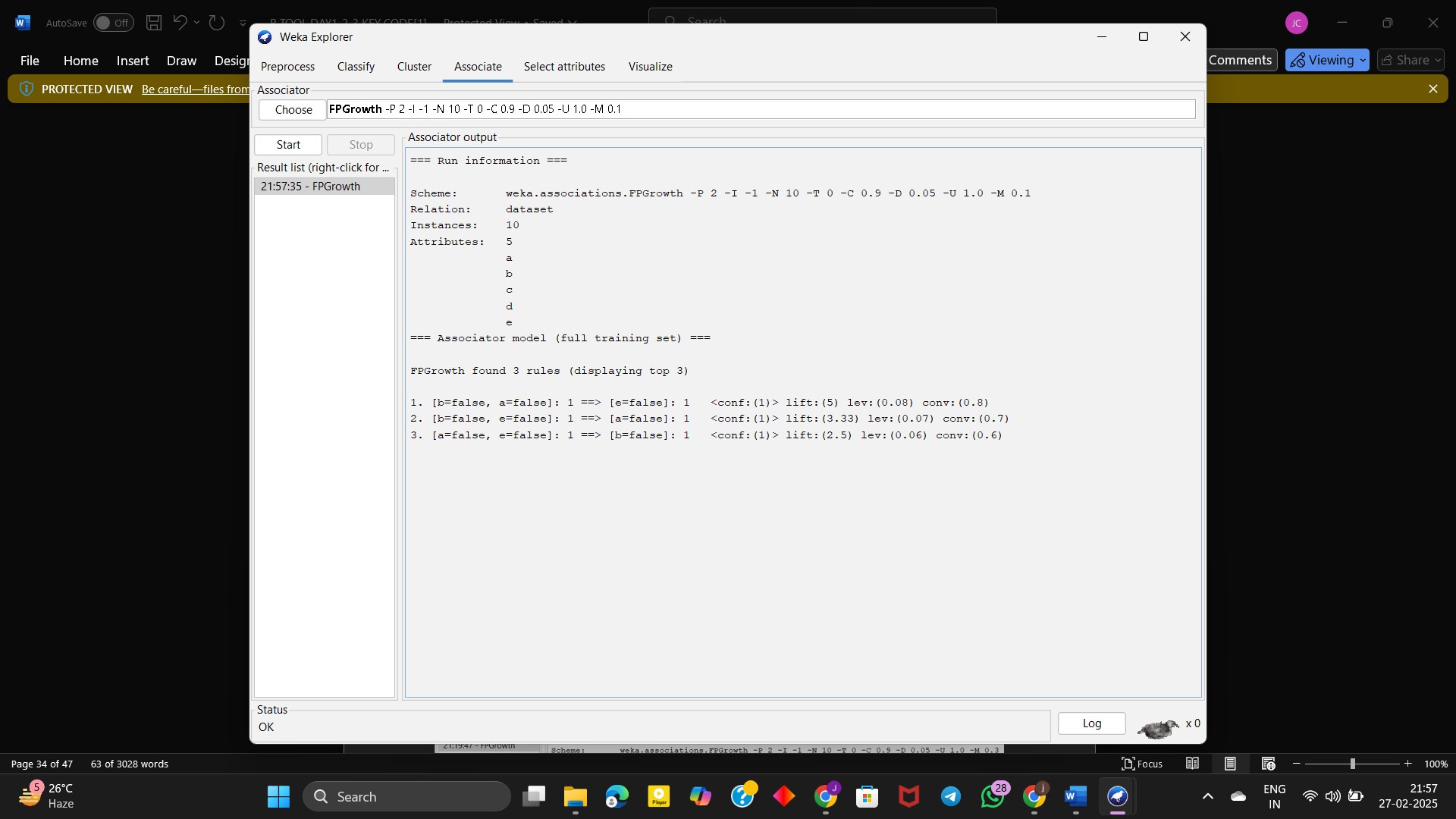
192372199

21.Implement using R language in which age group of people are affected by blood pressure based on the diabetes dataset show it using scatterplot and bar chart (that is Blood Pressure vs Age using dataset “diabetes.csv”)



22.Consider the data set and perform the Apriori Algorithm and FP algorithm support:3 and confidence=50%





23.Consider the data set and perform the Apriori Algorithm and FP algorithm support:3 and confidence=50%

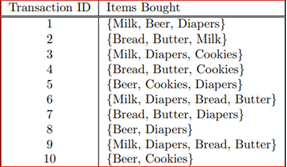
Consider the market basket transactions shown in the above table.

(a) What is the maximum number of association rules that can be extracted

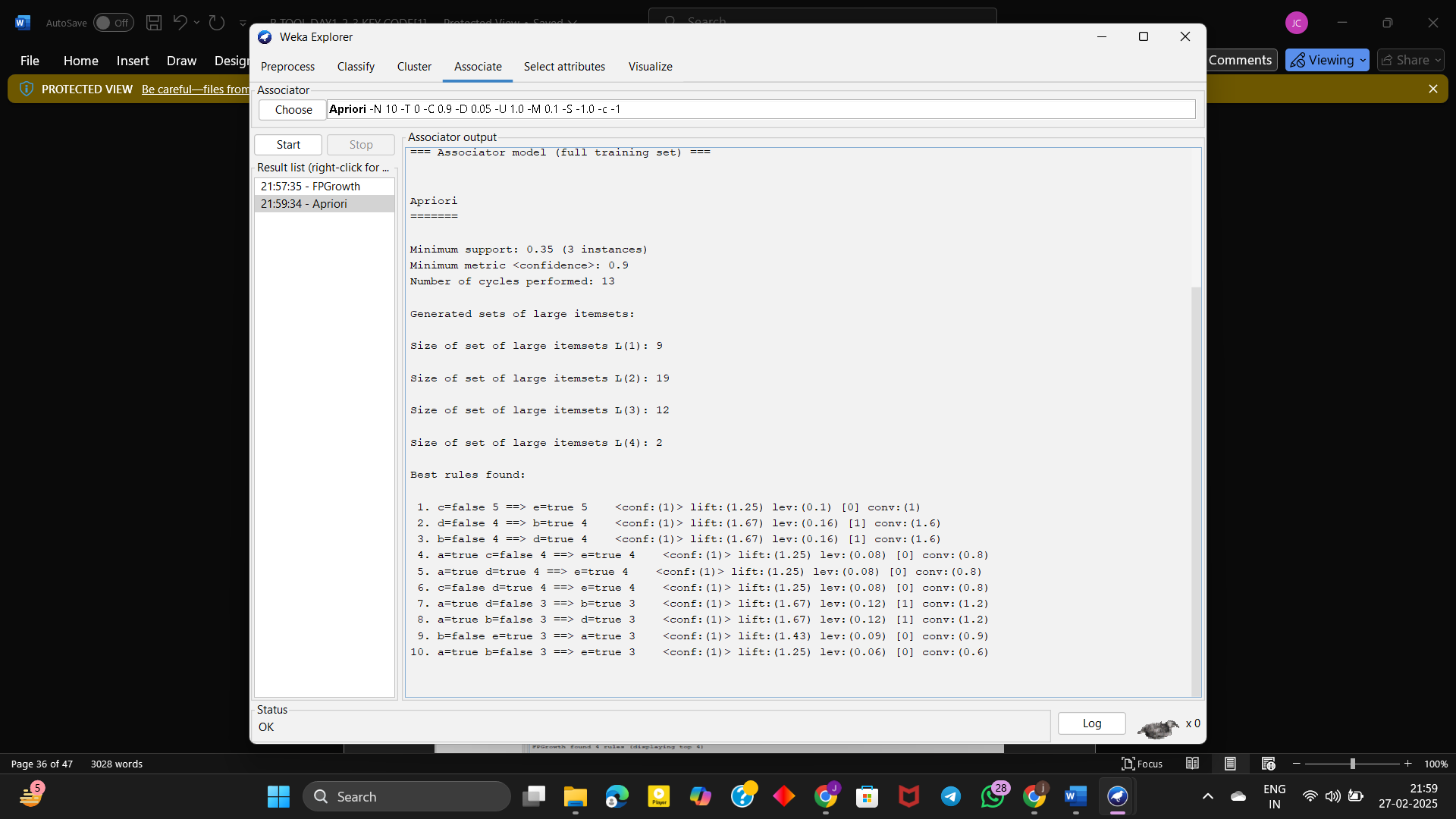
from this data (including rules that have zero support)?

(b) What is the maximum size of frequent itemsets that can be extracted

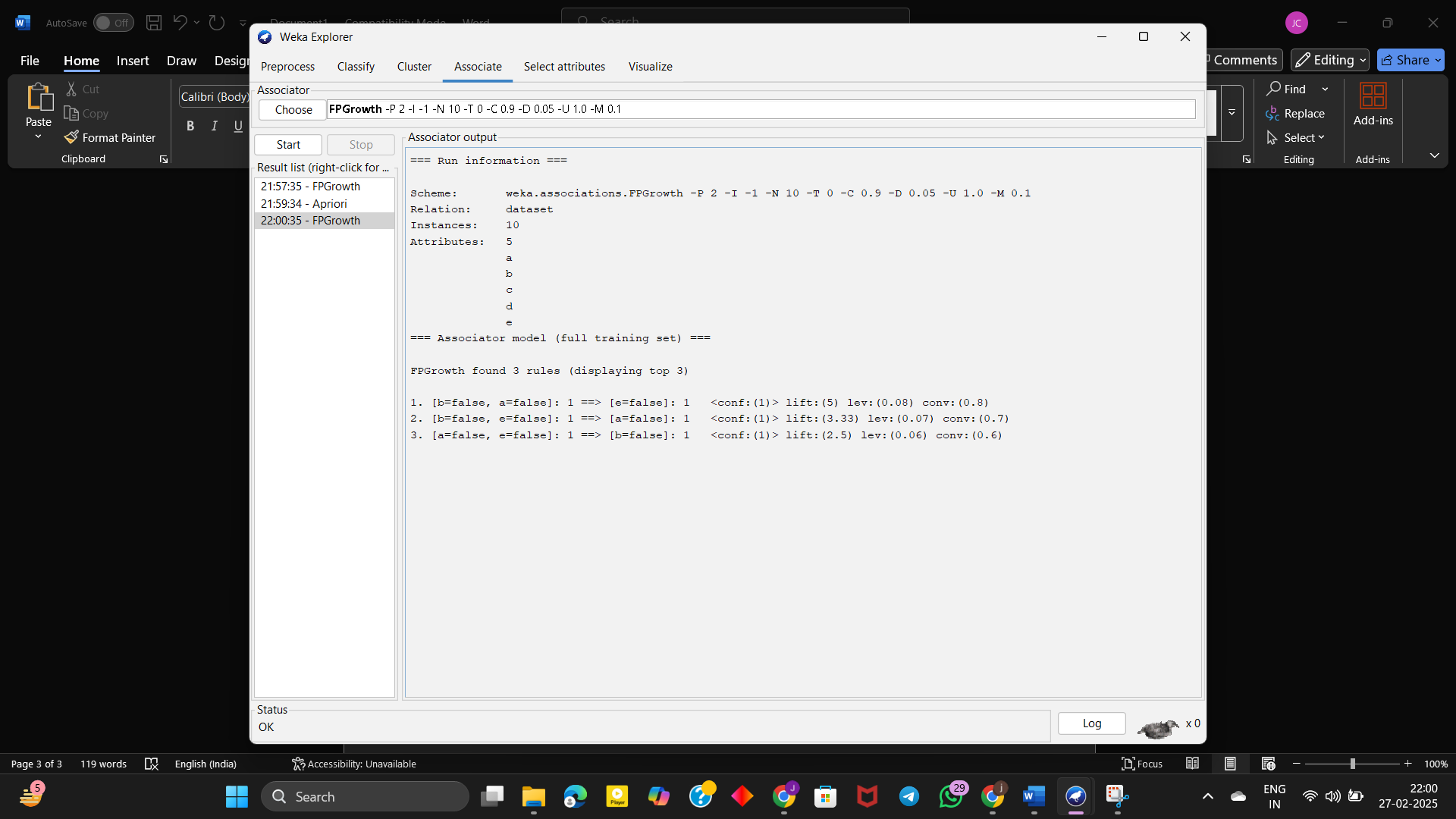
(assuming minsup > 0)?



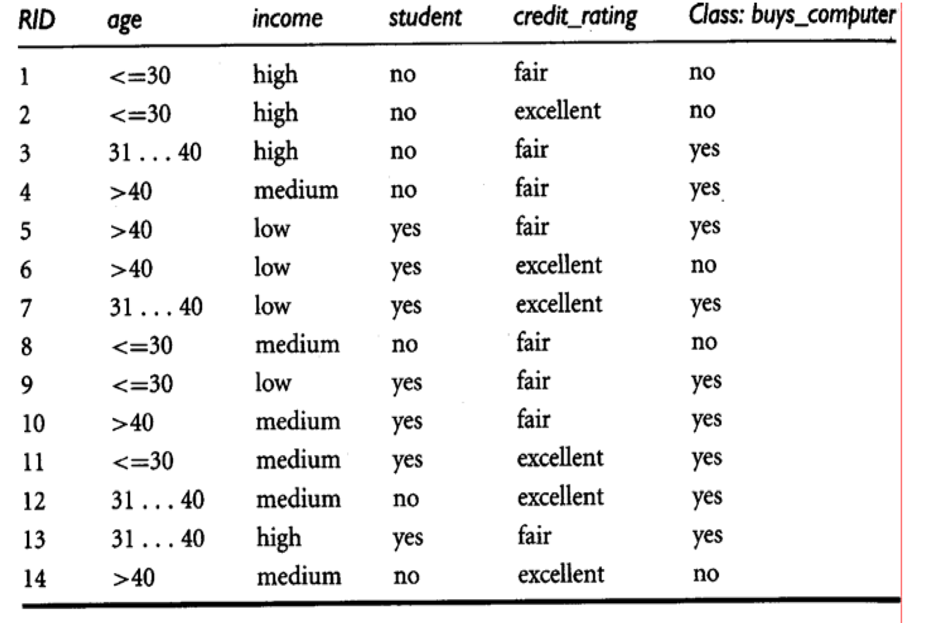
APRIORI



FG



24.Bayes classification and descion tree (using training and test data)



Input:

@relation decision\_tree

@attribute age{young,middle,old}

@attribute income{low,medium,high}

@attribute student{yes,no}

@attribute Creit\_rating{fair,excellent}

@attribute class{yes,no}

@data

young high no fair no

young high no excellent no

middle high no fair yes

old medium no fair yes

old low yes fair yes

old low yes excellent no

middle low yes excellent yes

young medium no fair no

young low yes fair yes

old medium yes fair yes

young medium yes excellent yes

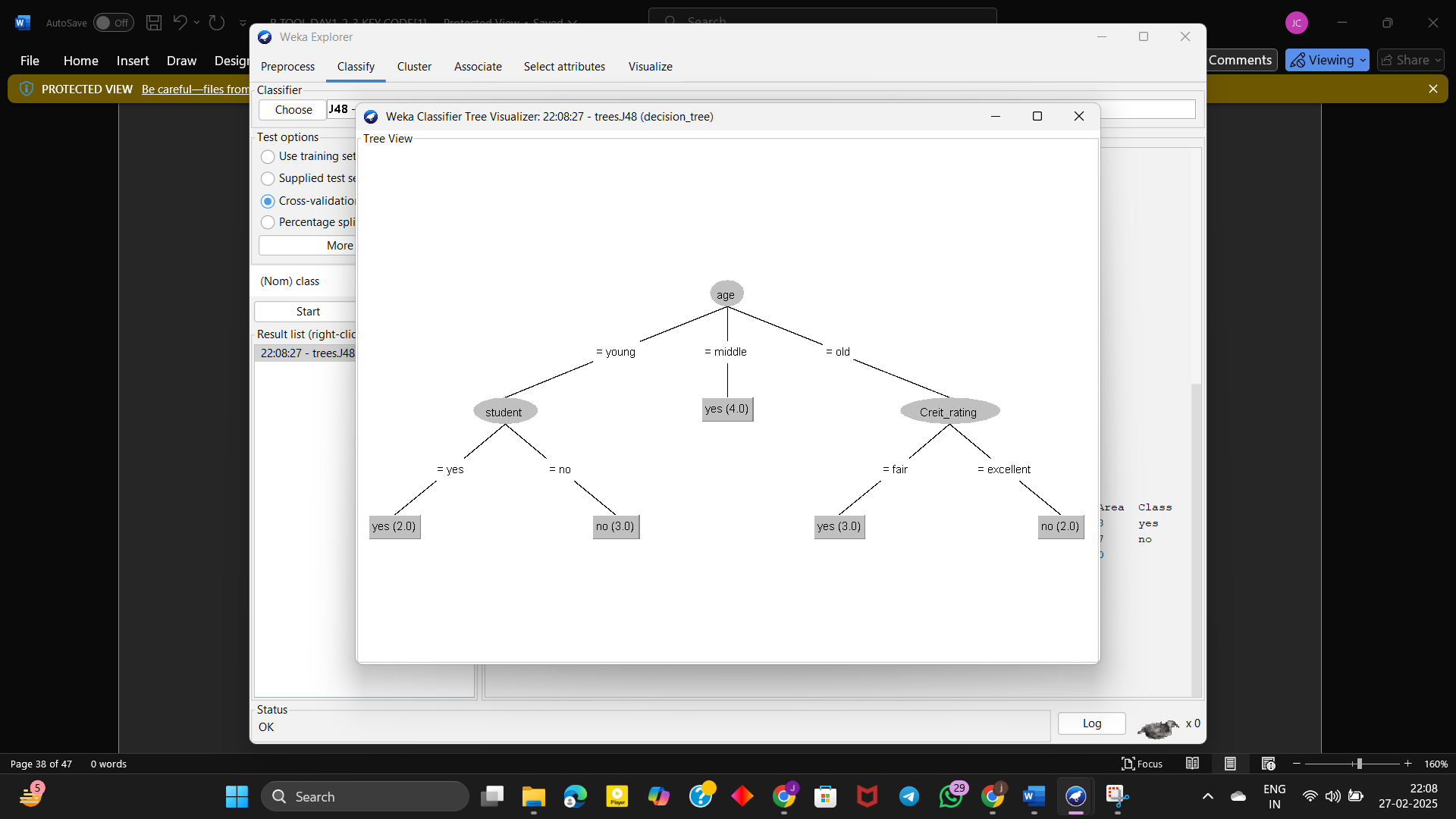
middle medium no excellent yes

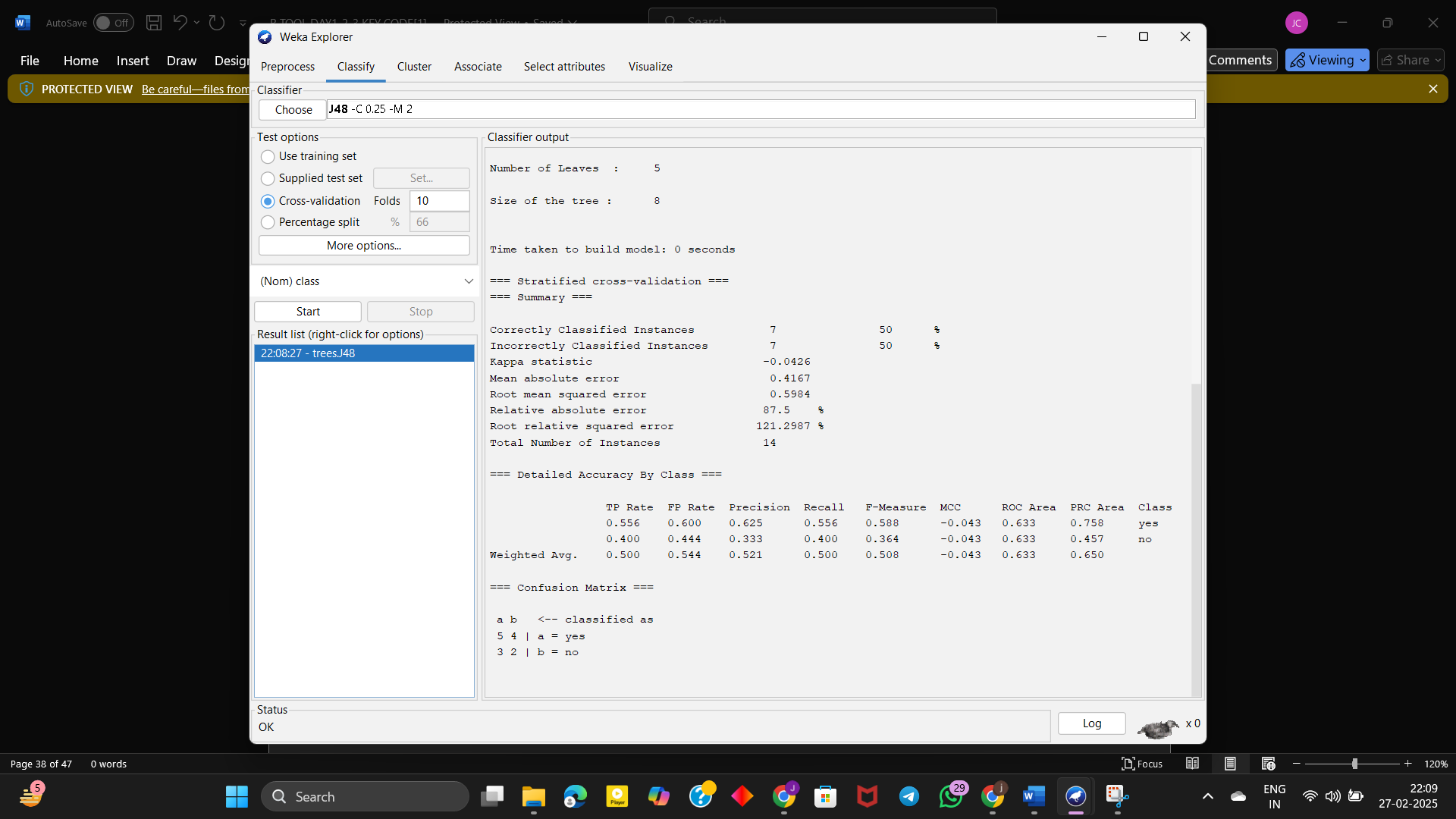
middle high yes fair yes

old medium no  excellent no

output:

tree:





25.Analysis the dataset “diabetes. csv” how the diabetes trend is for different age people, using linear regression and multiple regression.

Input:

data<-read.csv("C:/Users/Hari Naidu/Desktop/POM/download papers/diabetes.csv")

data

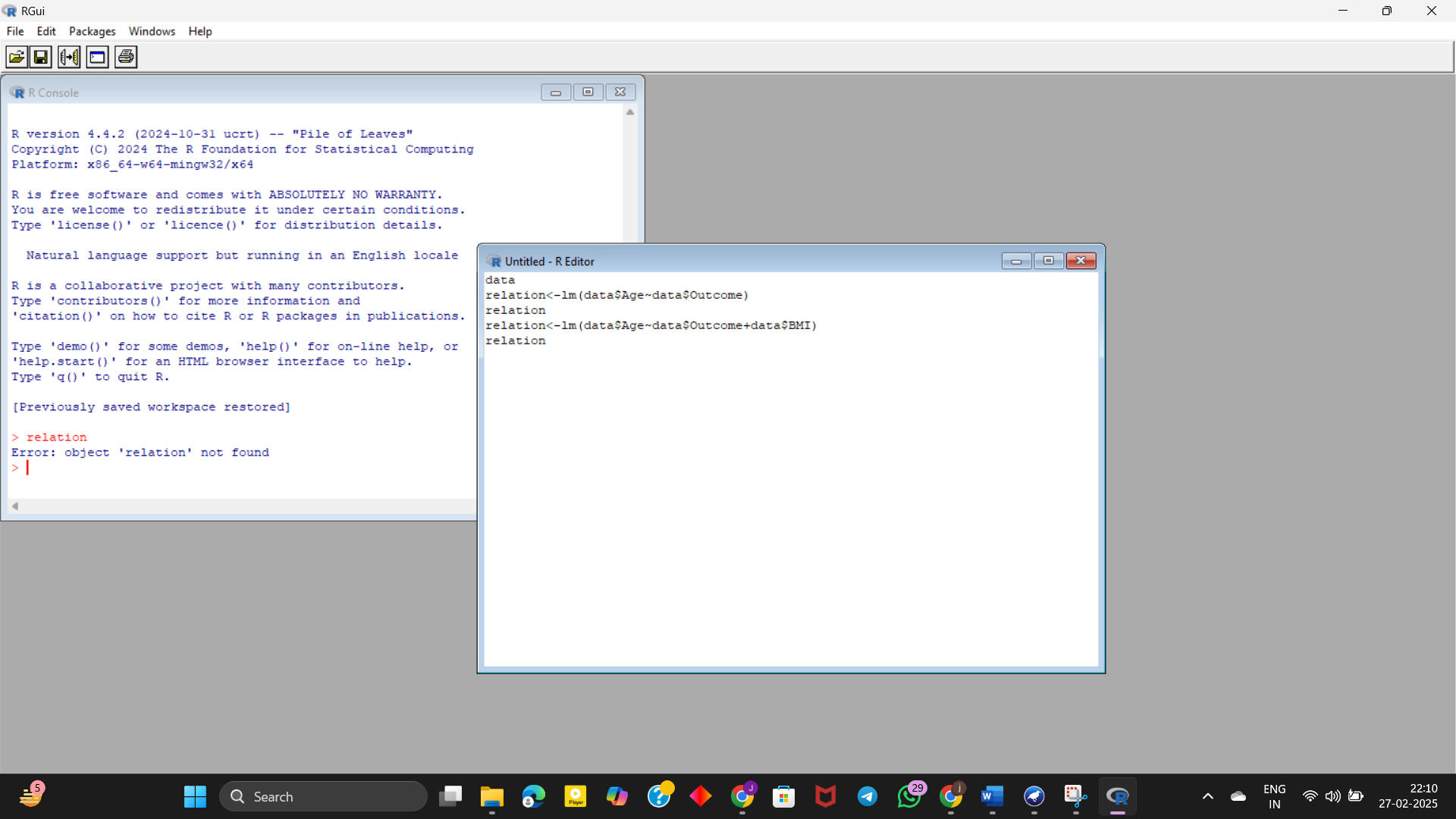
relation<-lm(data$Age~data$Outcome)

relation

relation<-lm(data$Age~data$Outcome+data$BMI)

relation

output:



26.Implement using WEKA for the given Suppose a database has five transactions. Let min sup= 50%(2) and min con f = 80%.

**Transactions** **Items**

T1 (M, O, N, K, E, Y)

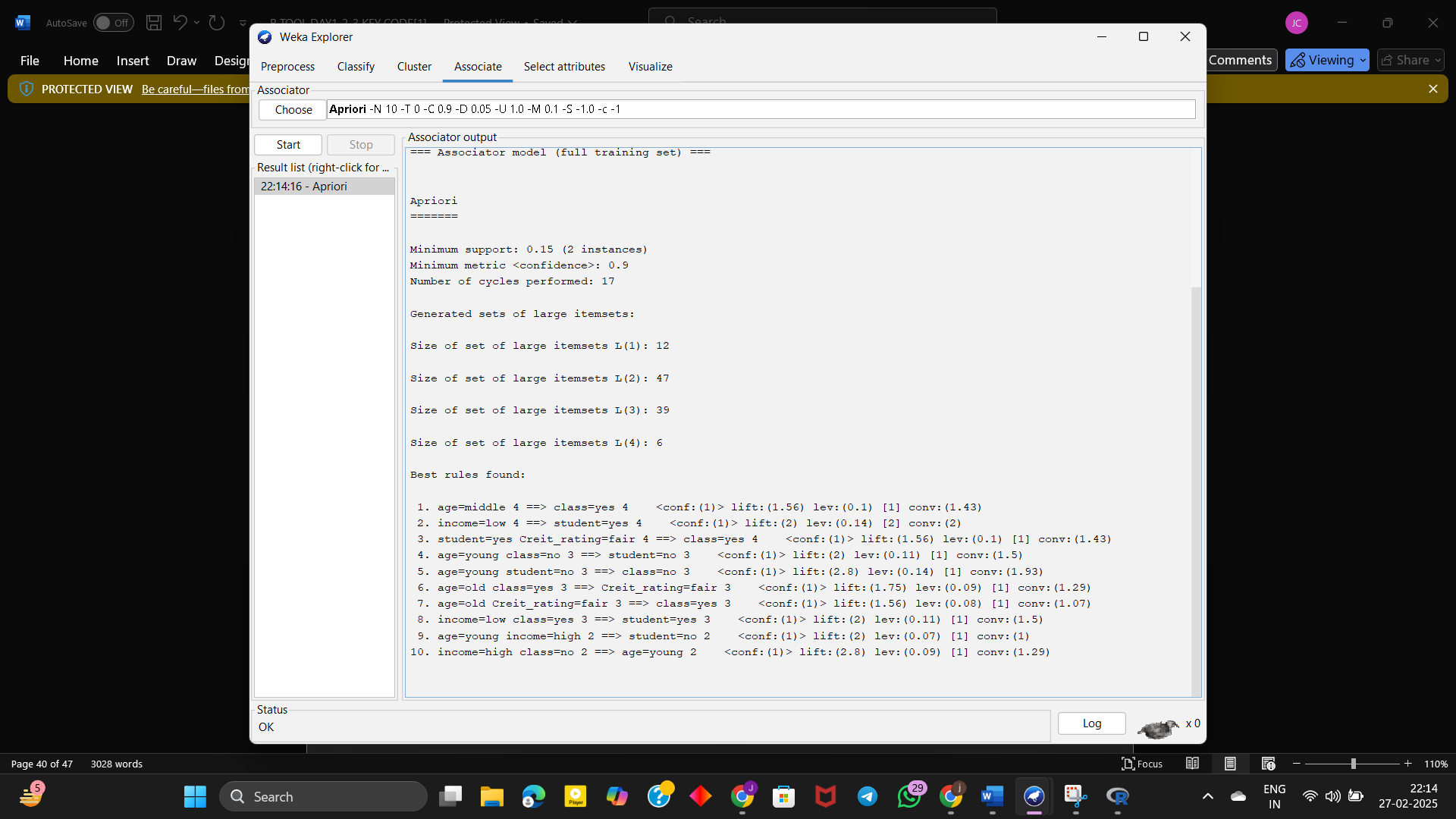
T2 (D, O, N, K, E, Y)

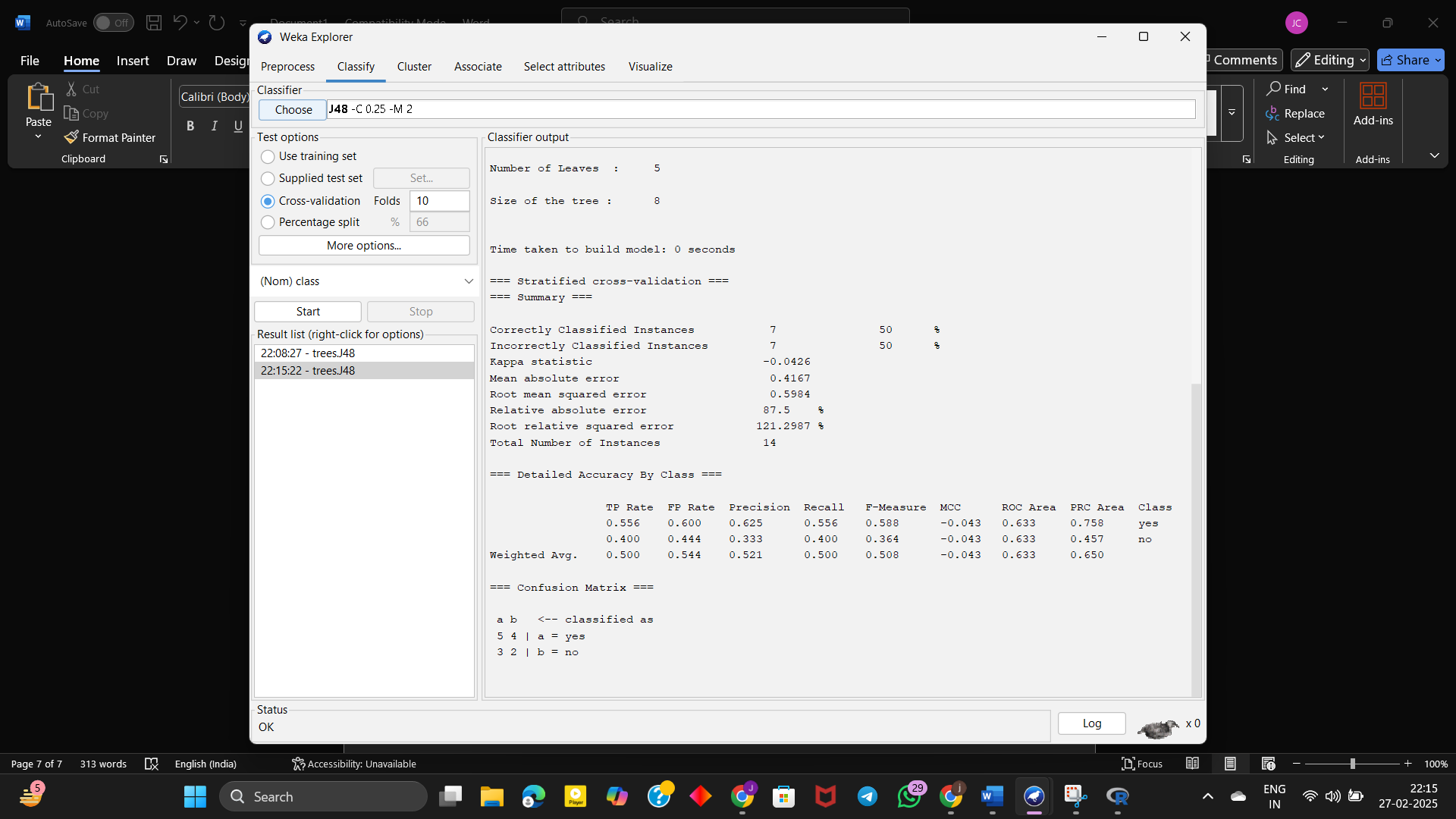
T3 (M, A, K, E)

T4 (M, U, C, K, Y)

T5 (C,O, O, K, I ,E)

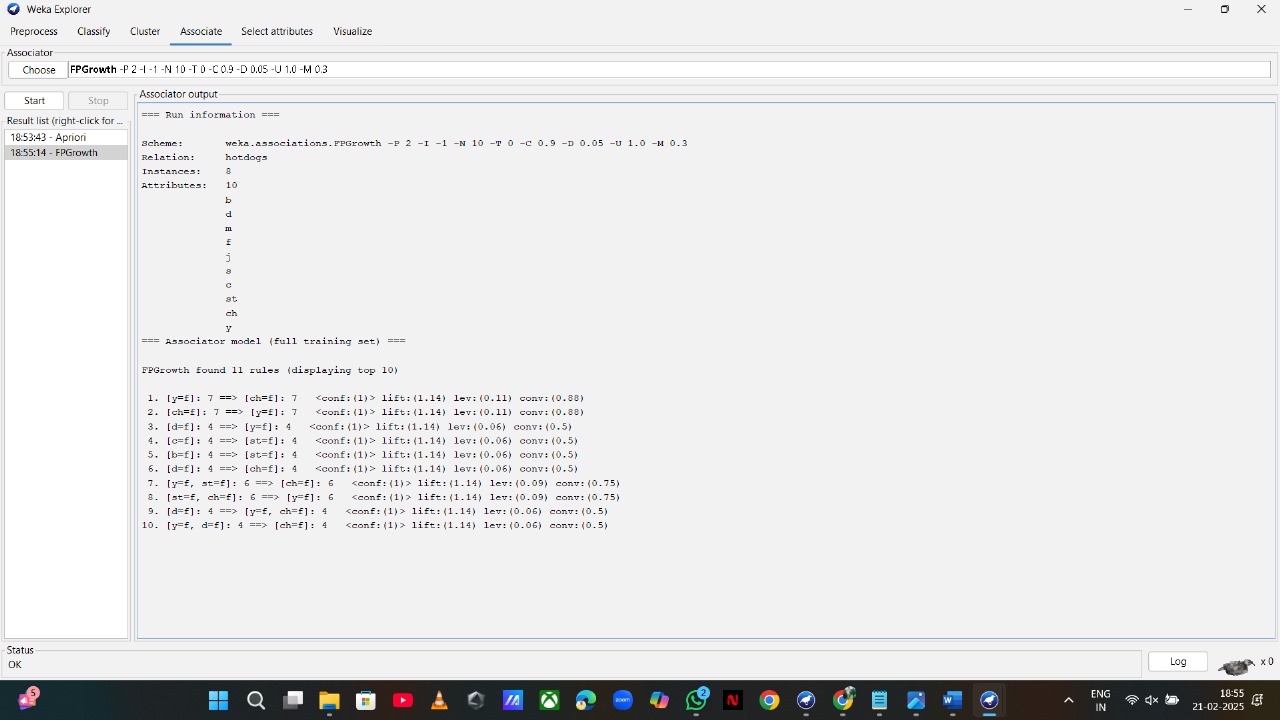
* Find all frequent item sets using Apriori algorithm
* Also draw FP-Growth Tree

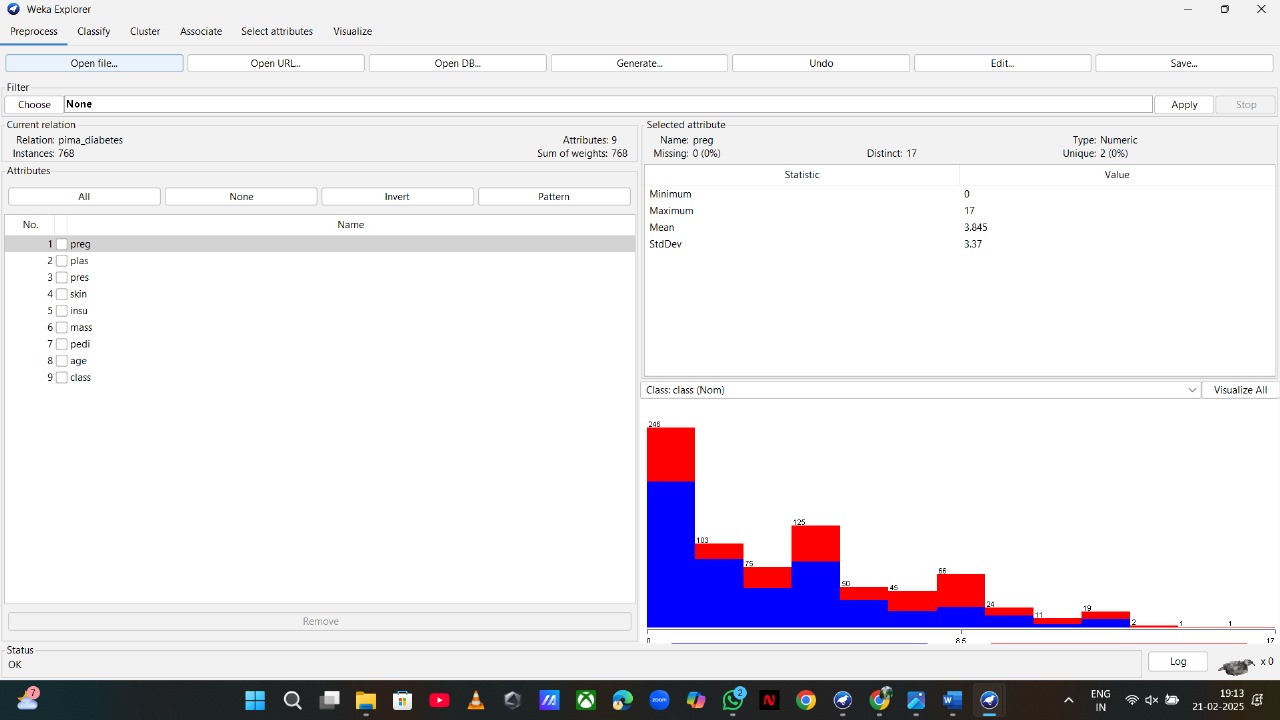


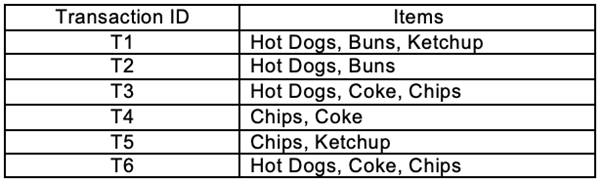


27. Prediction of Categorical Data using Decision Tree Algorithm through WEKA using any datasets. a) Tree b) Preprocess c) Logistic

Output:





28.Create the dataset using ARFF file format:

a.Find the **frequent itemsets** and generate **association rules** on this. Assume that minimum support threshold (s = 33.33%) and minimum confident threshold (c = 60%).

b.List the various rule generated by apriori and FP tree algorthim ,mention wheather accepted or rejcted.

Input:

@relation hotdogs

@attribute hotdogs{t,f}

@attribute buns{t,f}

@attribute ketchup{t,f}

@attribute coke{t,f}

@attribute chips{t,f}

@data

t t t f f

t t f f f

t f f t t

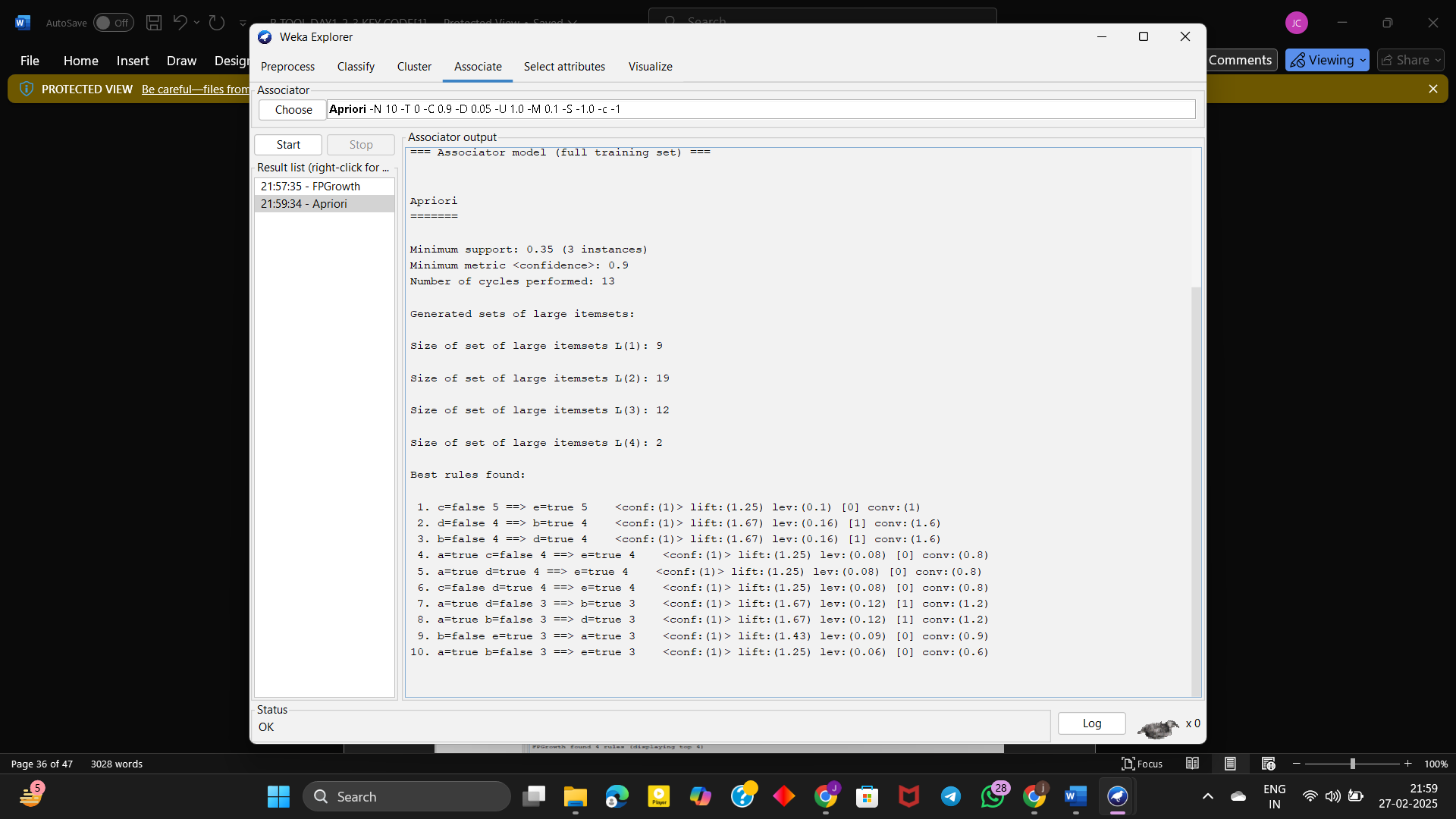
f f f t t

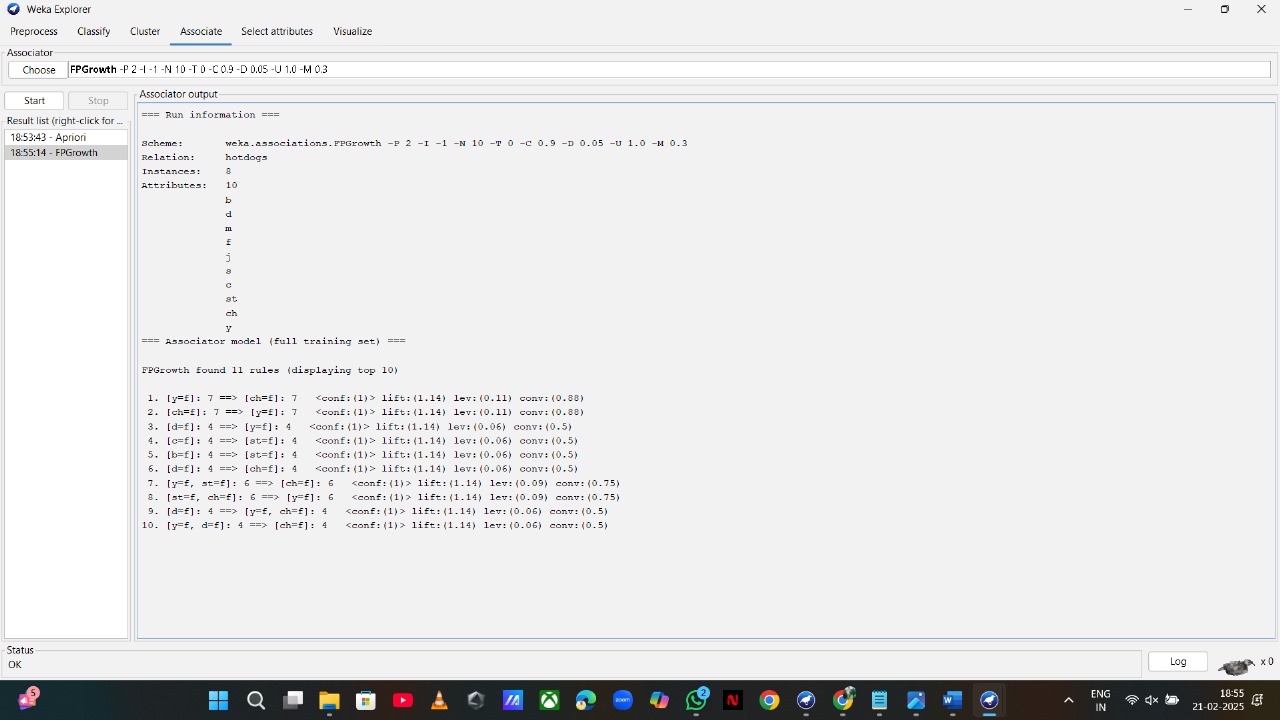
f f t f t

t f f t t

output:

apriori algorithm:





29.Prediction of Categorical Data using Rule base classification and decision tree classification through WEKA using any datasets. Compare the accuracy using two algorithm and plot the graph

Input:

@relation decision\_tree

@attribute age{young,middle,old}

@attribute income{low,medium,high}

@attribute student{yes,no}

@attribute Creit\_rating{fair,excellent}

@attribute class{yes,no}

@data

young high no fair no

young high no excellent no

middle high no fair yes

old medium no fair yes

old low yes fair yes

old low yes excellent no

middle low yes excellent yes

young medium no fair no

young low yes fair yes

old medium yes fair yes

young medium yes excellent yes

middle medium no excellent yes

middle high yes fair yes

old medium no  excellent no

Output:

Rule based classification:

