DAV- Practical 1

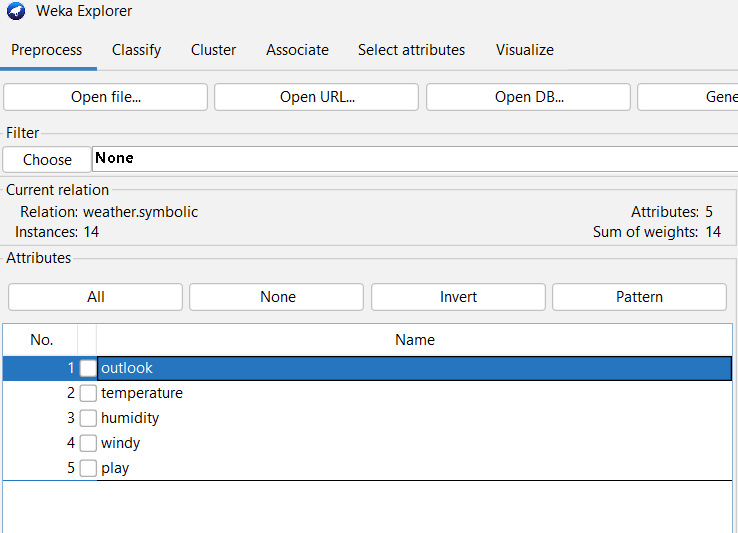
Name:- Krishna Mundada

Roll no:- 45  
Batch:- E3  
Subject:- DAV Lab

# AIM: Introduction to weka and data preprocessing on the given data set in Weka

## 1.Press the Explorer button on the main panel and load the weather dataset and answer the following questions

## Output:

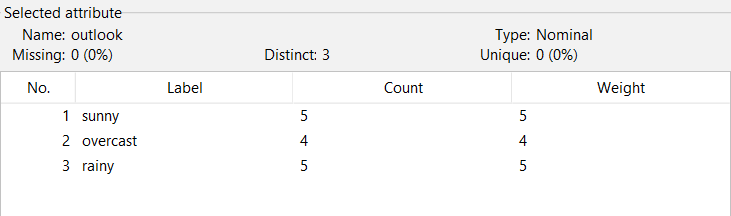


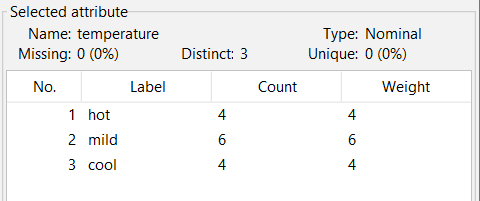
**(1) How many instances are there in the dataset?**

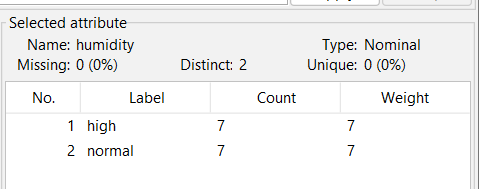
-> 14

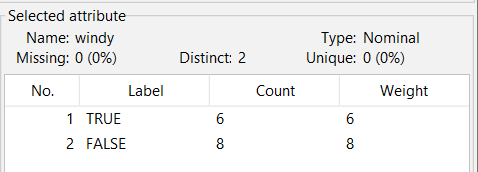
**(2) State the names of the attributes along with their types and values.**

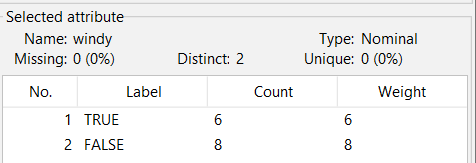
->







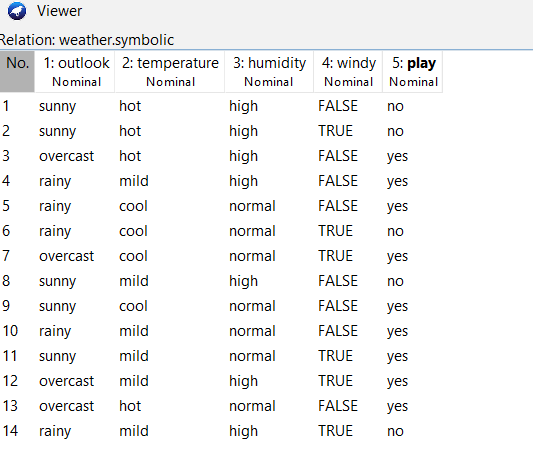




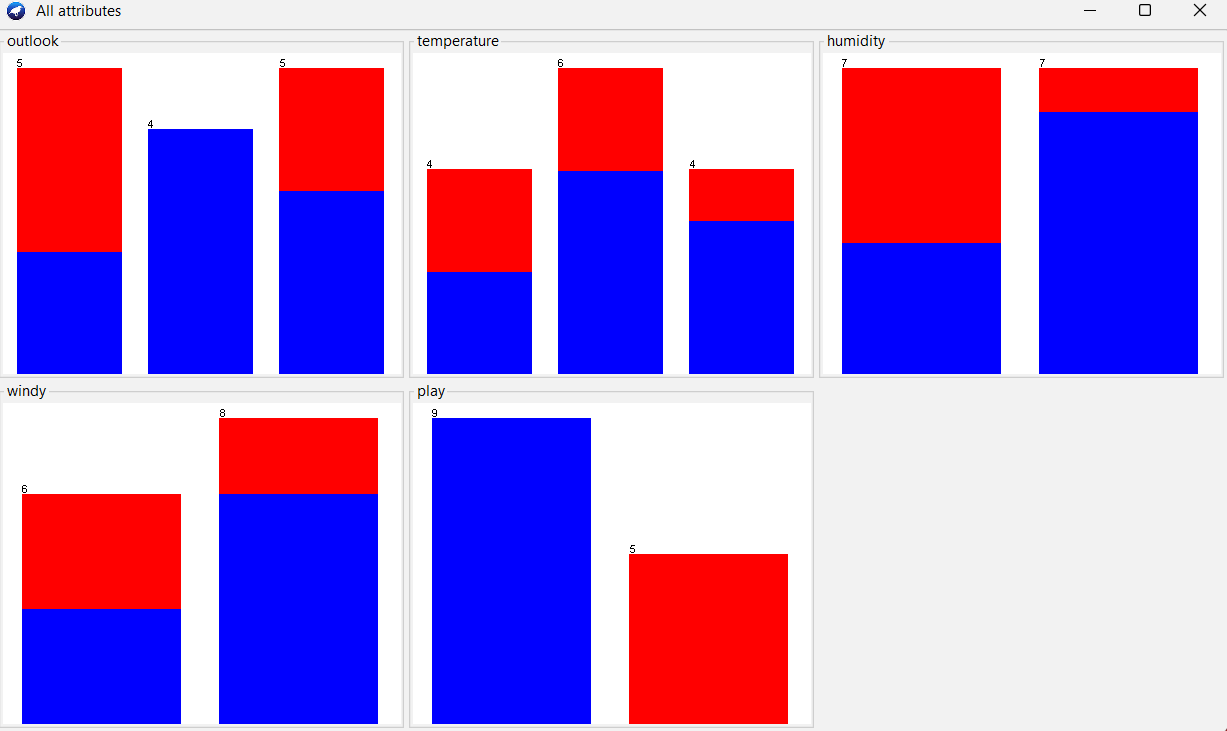
**(3) What is the class attribute?**

-> Play

**(4) How will you determine how many instances of each class are present in the data**

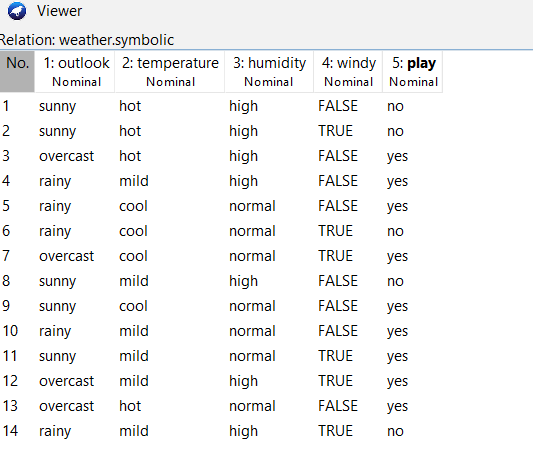


**(5) What happens with the Visualize All button is pressed?**

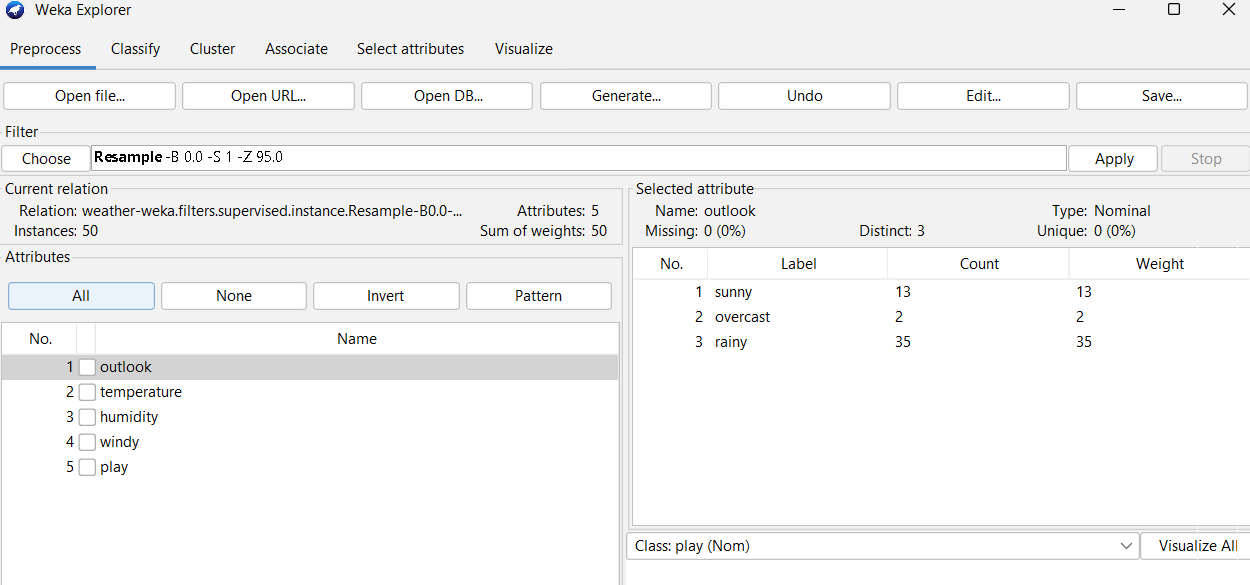


**(6) How will you view the instances in the dataset? How will you save the changes?**

-> Click edit button

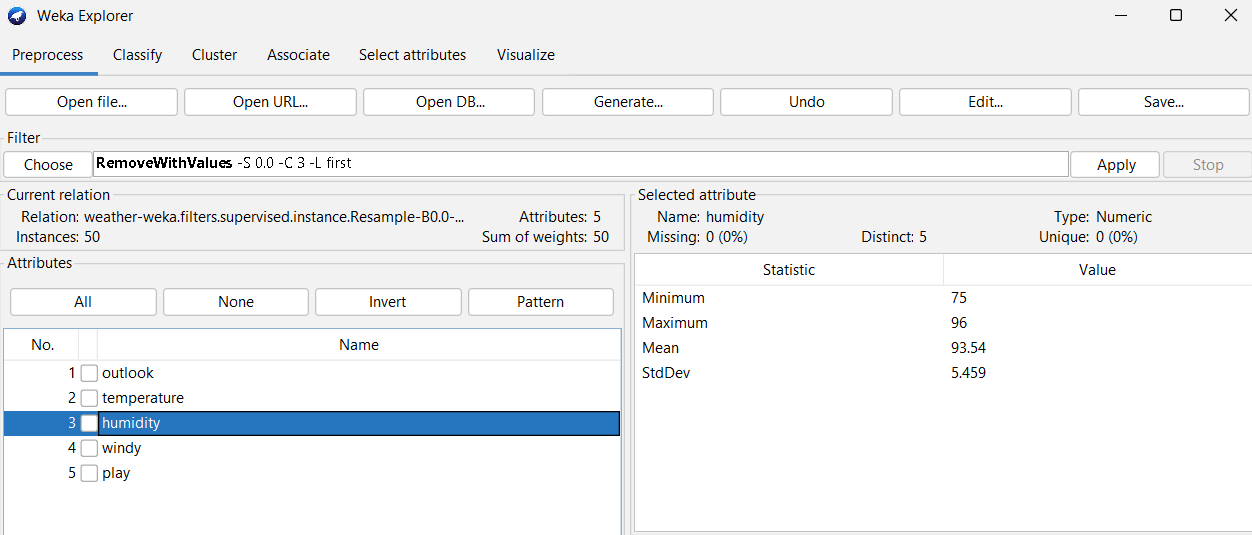


**(7) Now, extend the dataset to include 50 instances in total.**

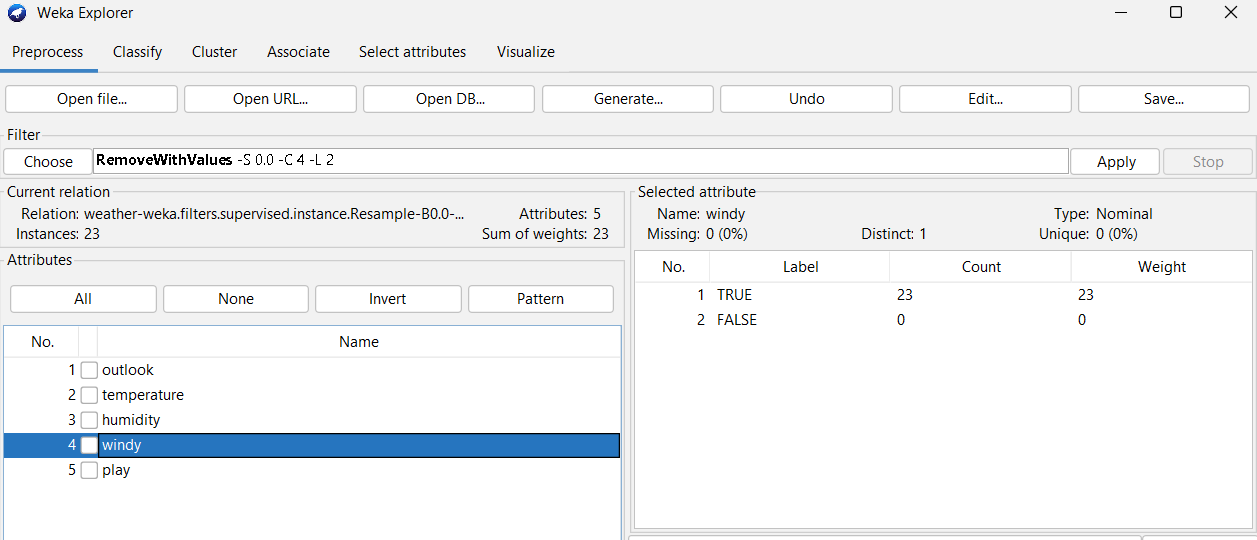
****

## 2. Do as directed to apply Filter

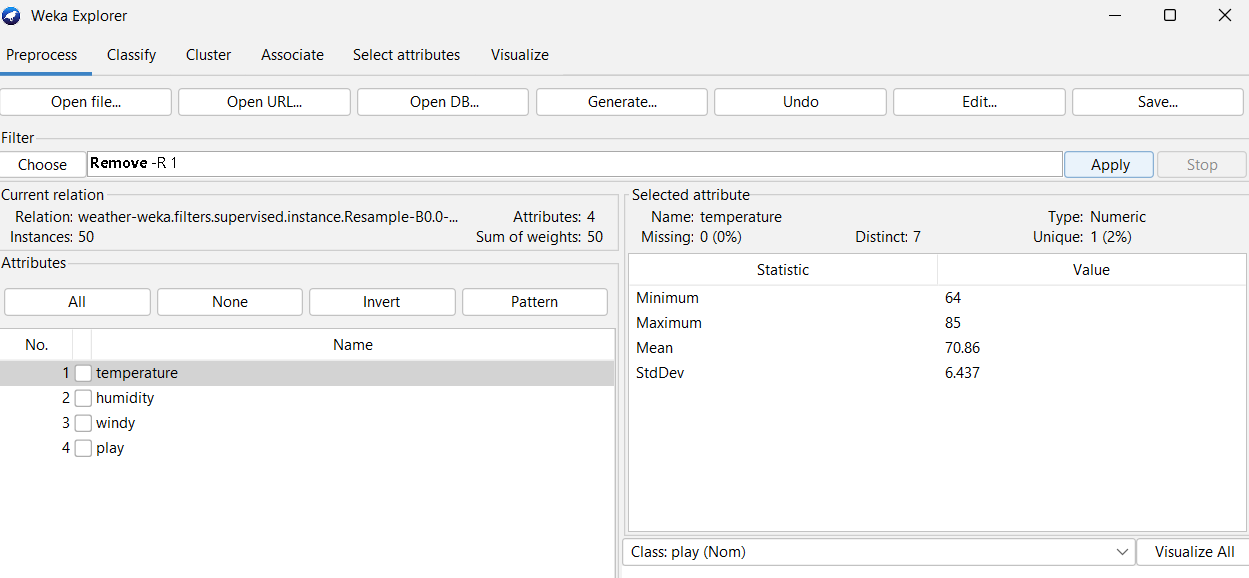
1. **Use the unsupervised filter RemoveWithValues to remove all instances where the Attribute ‘humidity’ has the value ‘high’? Undo the effect of the filter.**

****

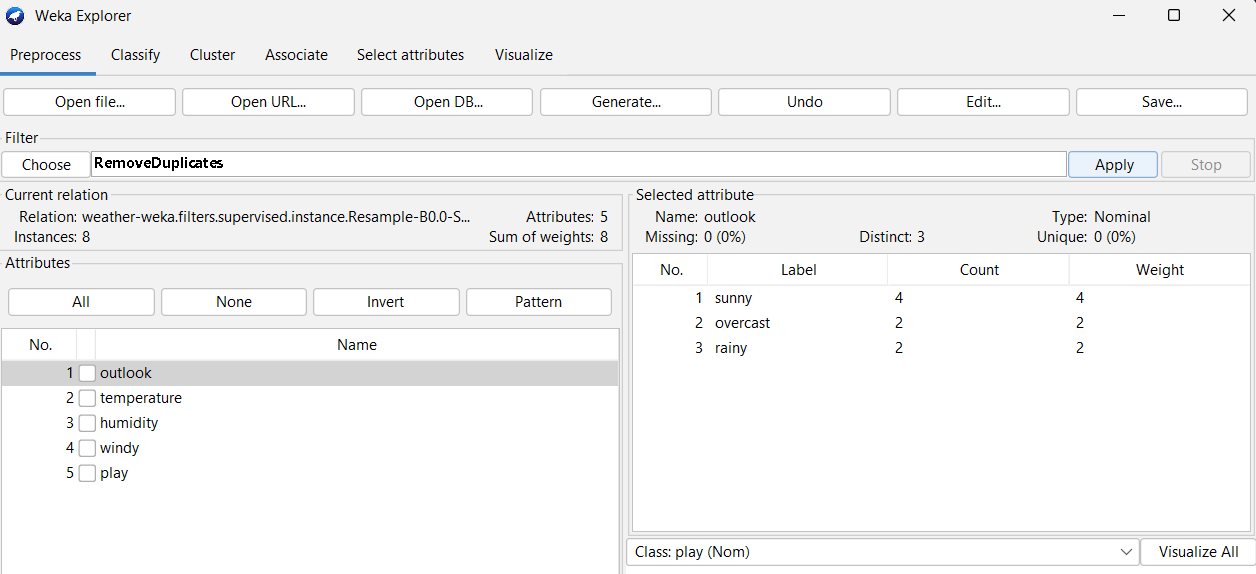
1. **Remove the ‘FALSE’ instances of windy attribute and undo the effect.**

****

1. **Remove the attribute outlook and undo the effect.**

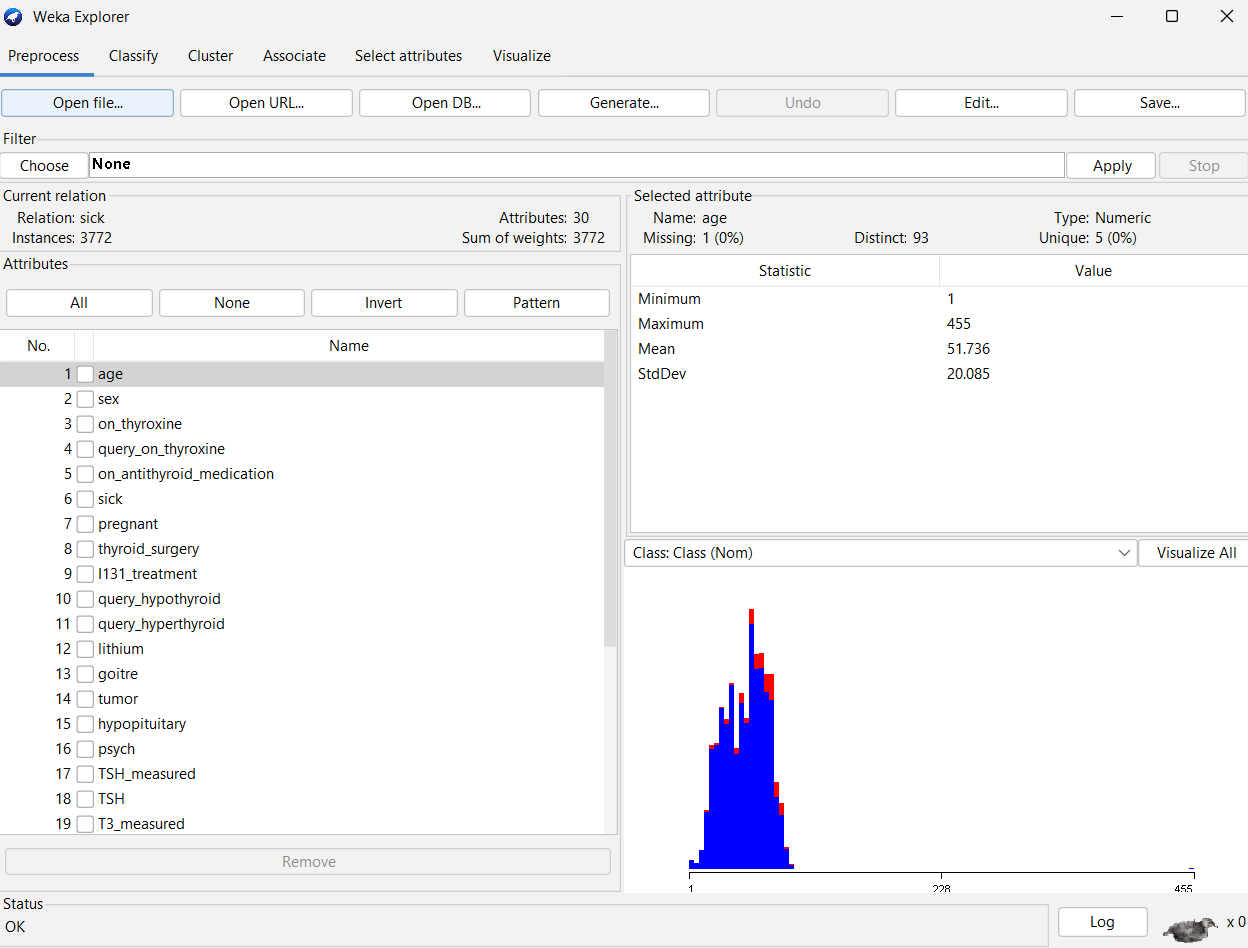
****

1. **Experiment with different filters and report their effects.**

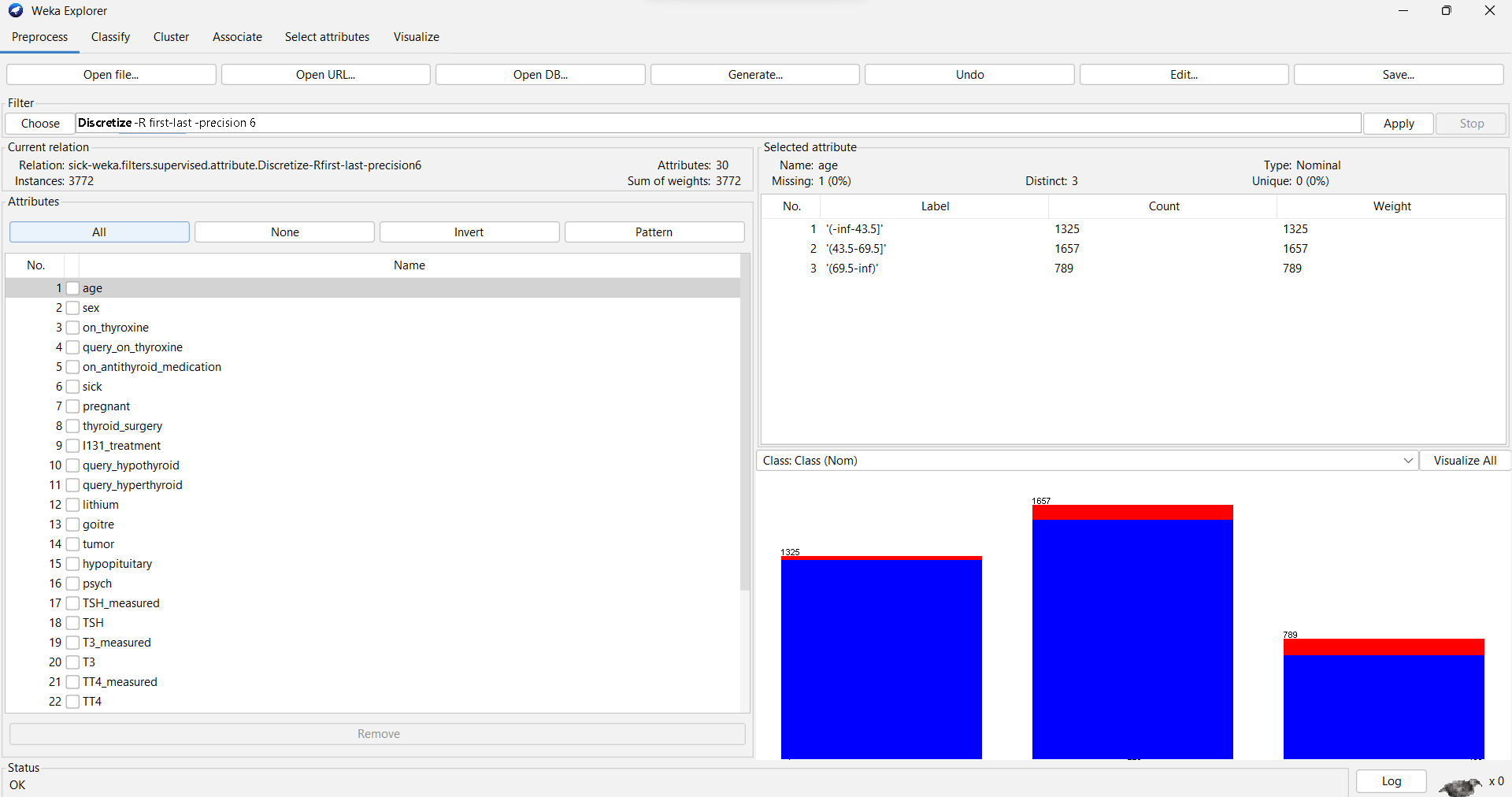
****

## 3. Application of Discretization Filters [use sick.arff dataset]

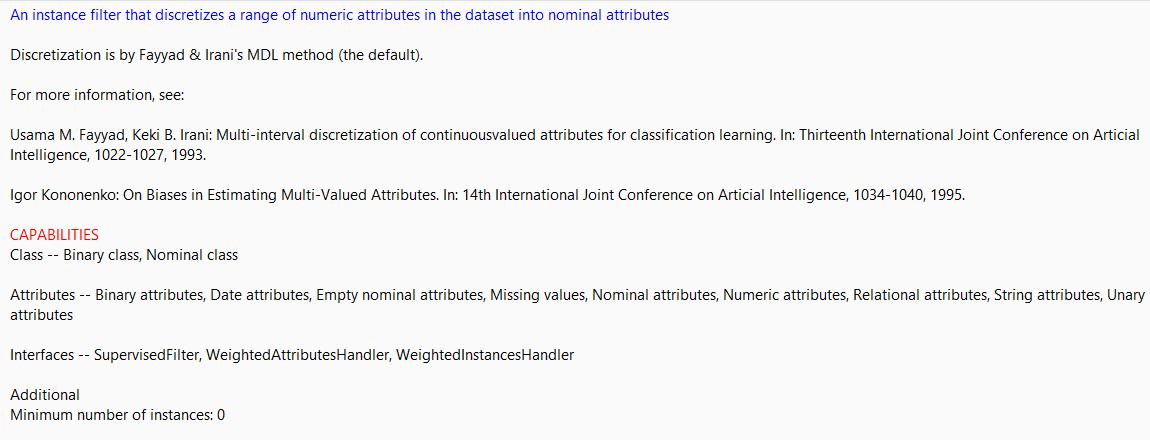
**1. Load the sick.arff dataset.**

****

**2. Apply the supervised discretization filter on different attributes.**

****

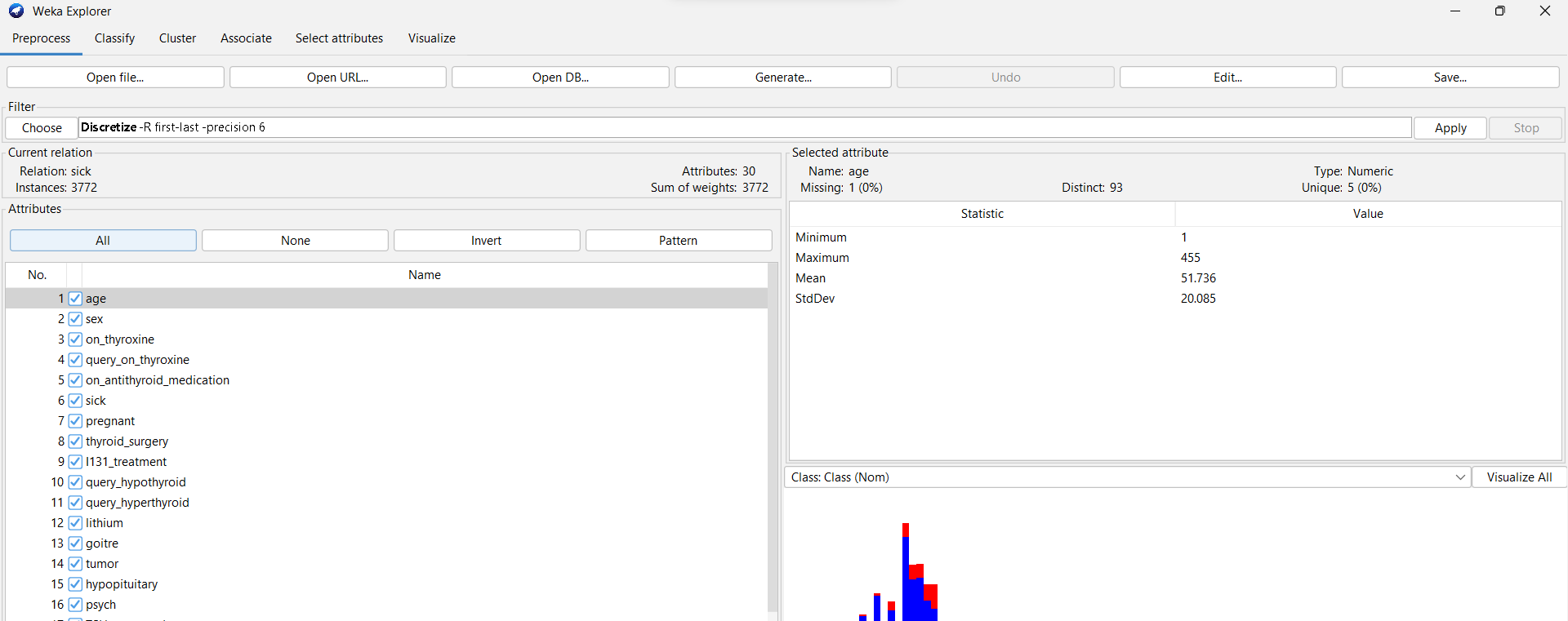
**3. What is the effect of this filter on the attributes?**

****

**4. How many distinct ranges have been created for each attribute?**

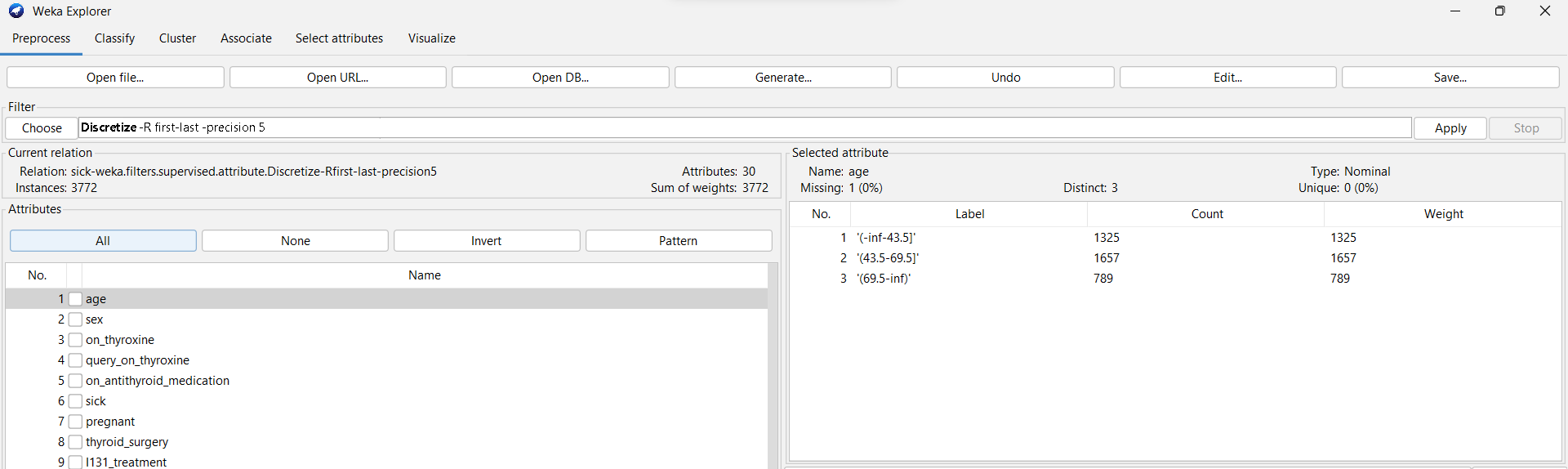
**-> 2**

**5. Undo the filter applied in the previous step.**

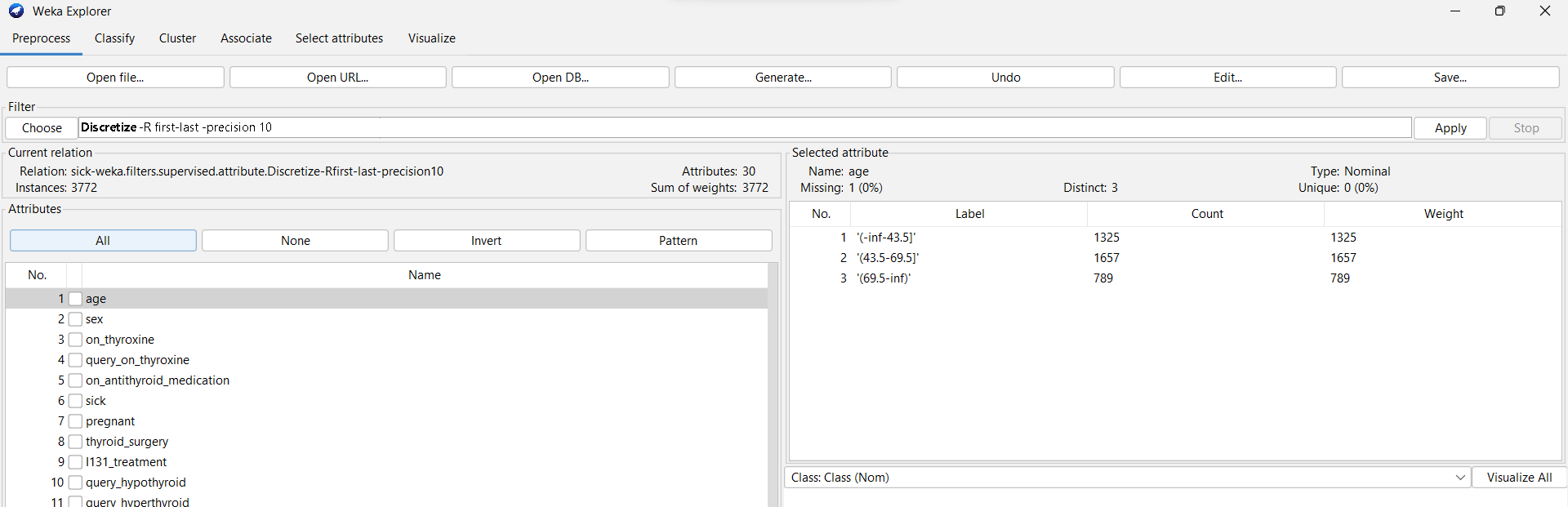
****

## 6. Apply the unsupervised discretization filter. Do this twice:

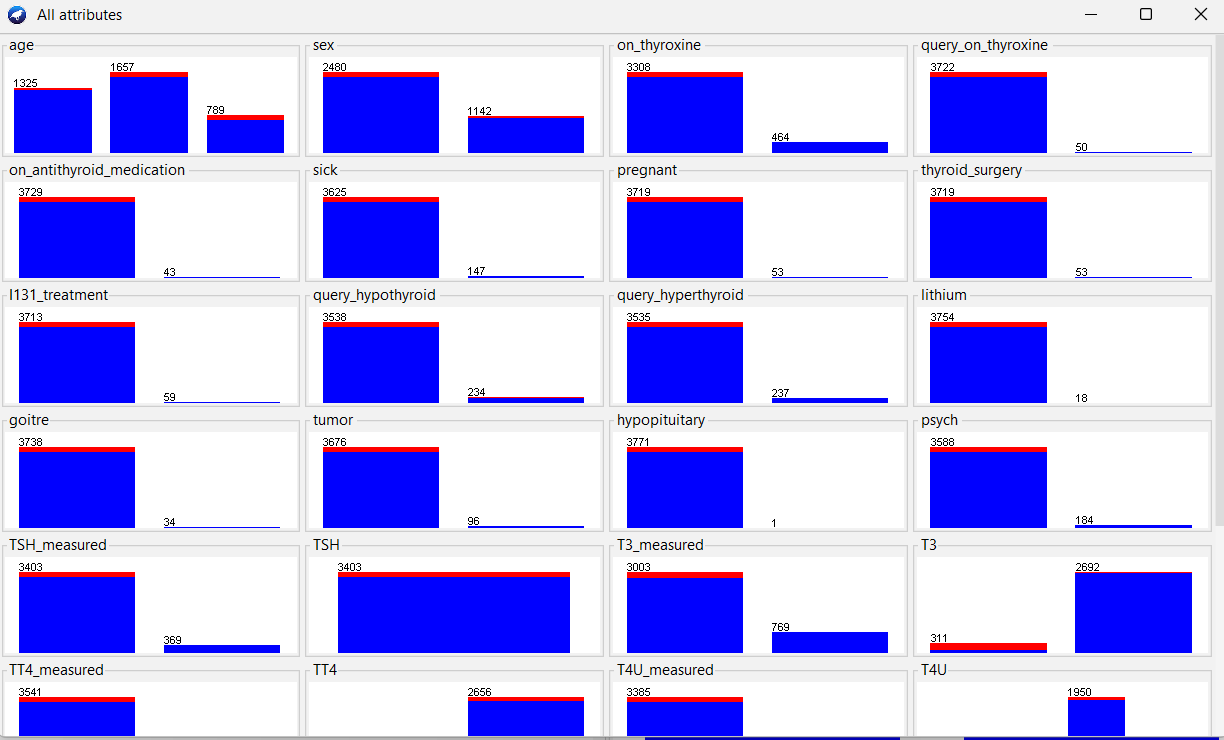
**1. In this step, set bins=5**

****

**2. In this step, set bins=10**

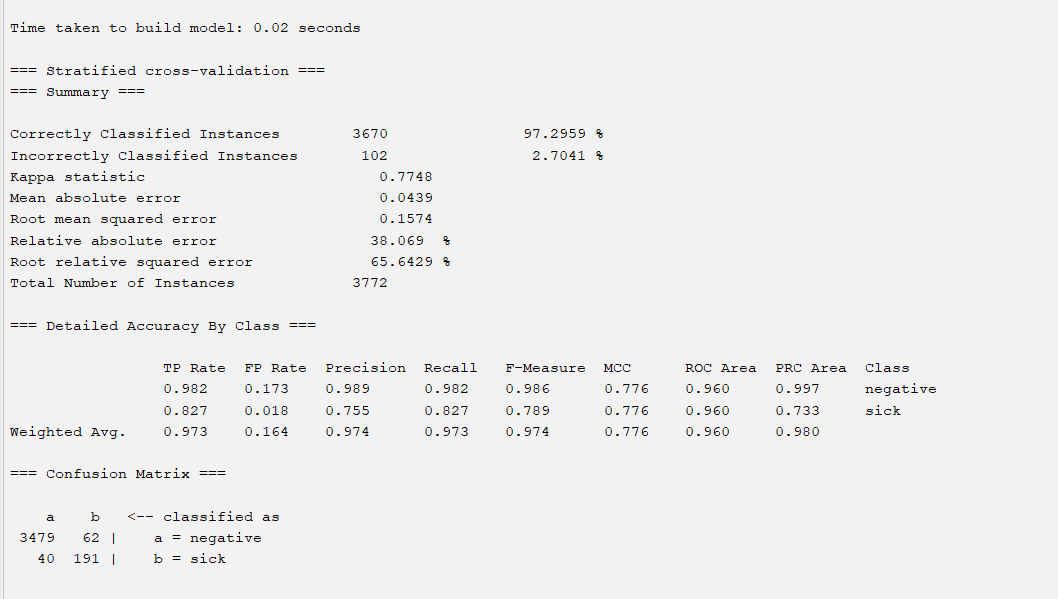
****

**3. What is the effect of the unsupervised filter on the dataset?**

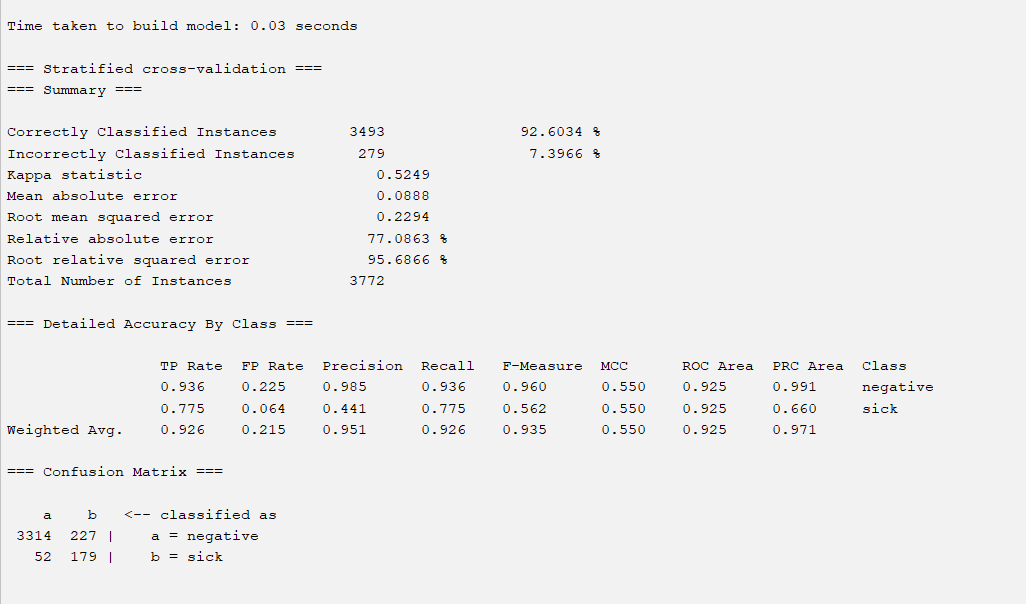
****

## 7. Run the the Naive Bayes classifier after apply the following filters

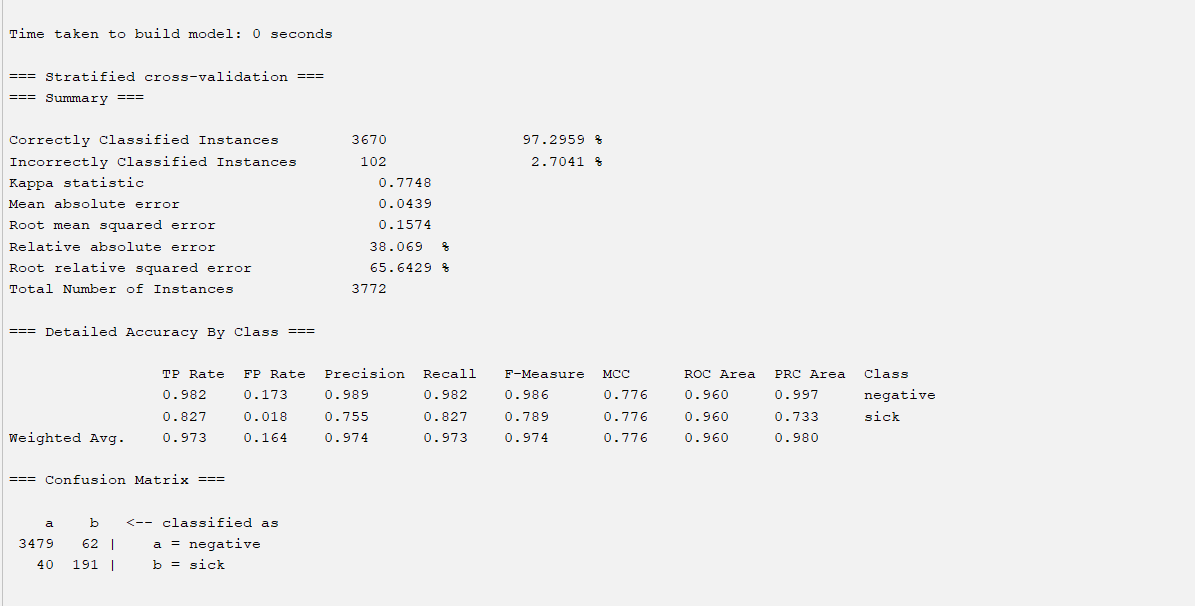
**1. Unsupervised discretized with bins=5**

****

**2. Unsupervised discretized with bins =10**

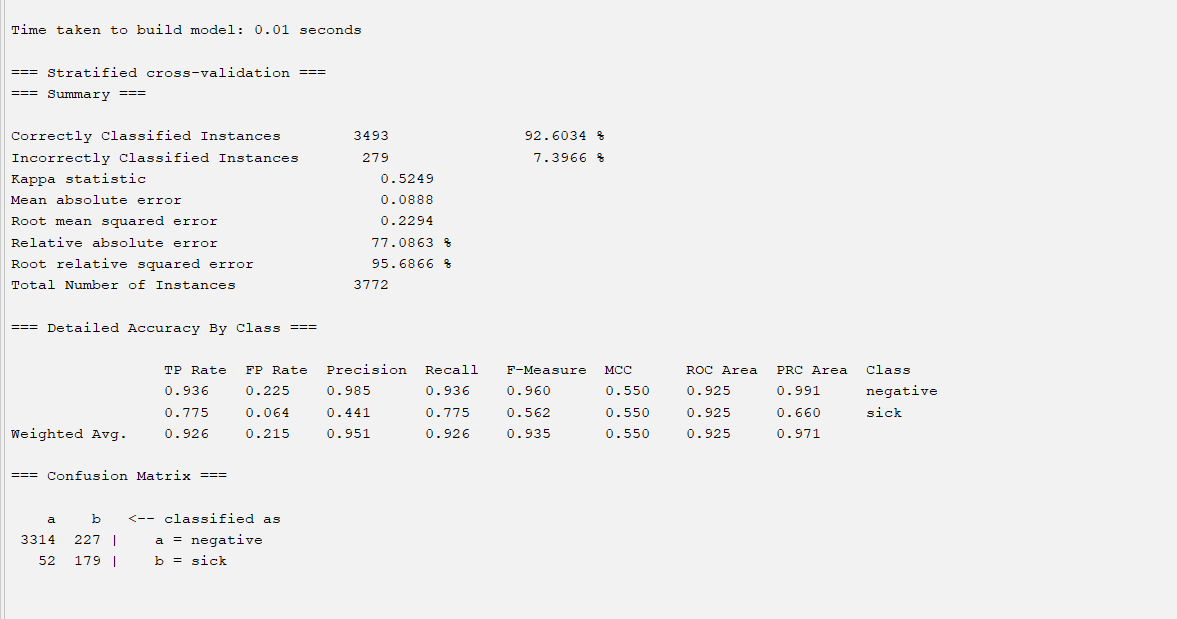
****

**3. Unsupervised discretized with bins=20.**

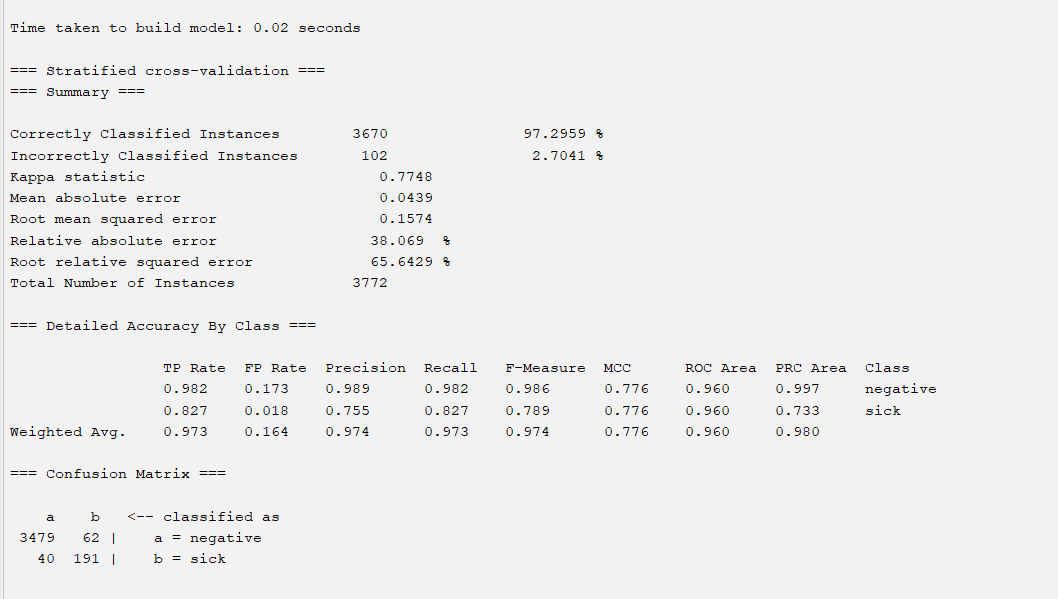
****

## 8. Compare the accuracy of the following cases

**1. Naive Bayes without discretization filters**

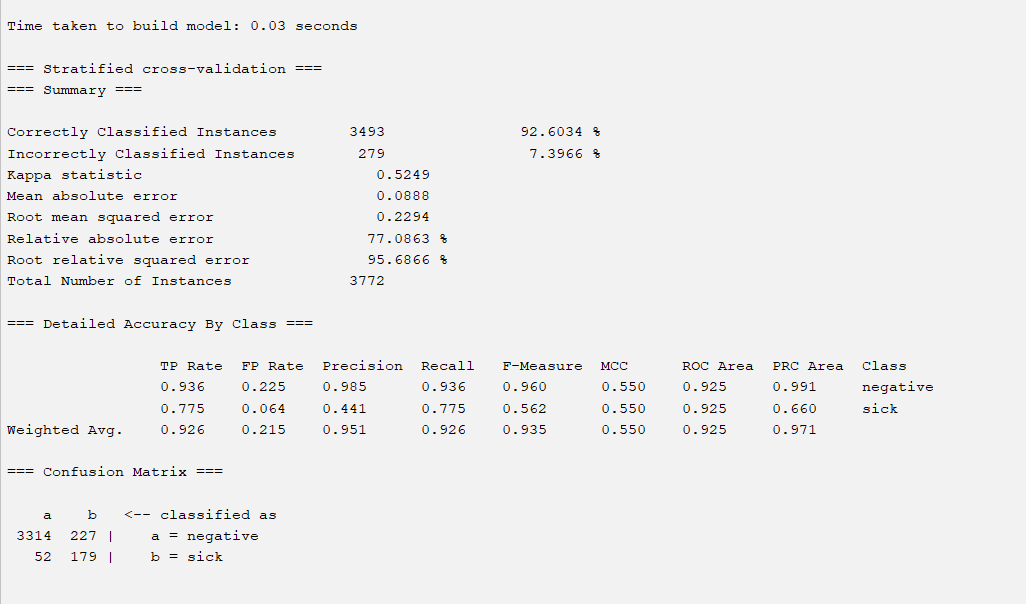
****

**2. Naive Bayes with a supervised discretization filter**

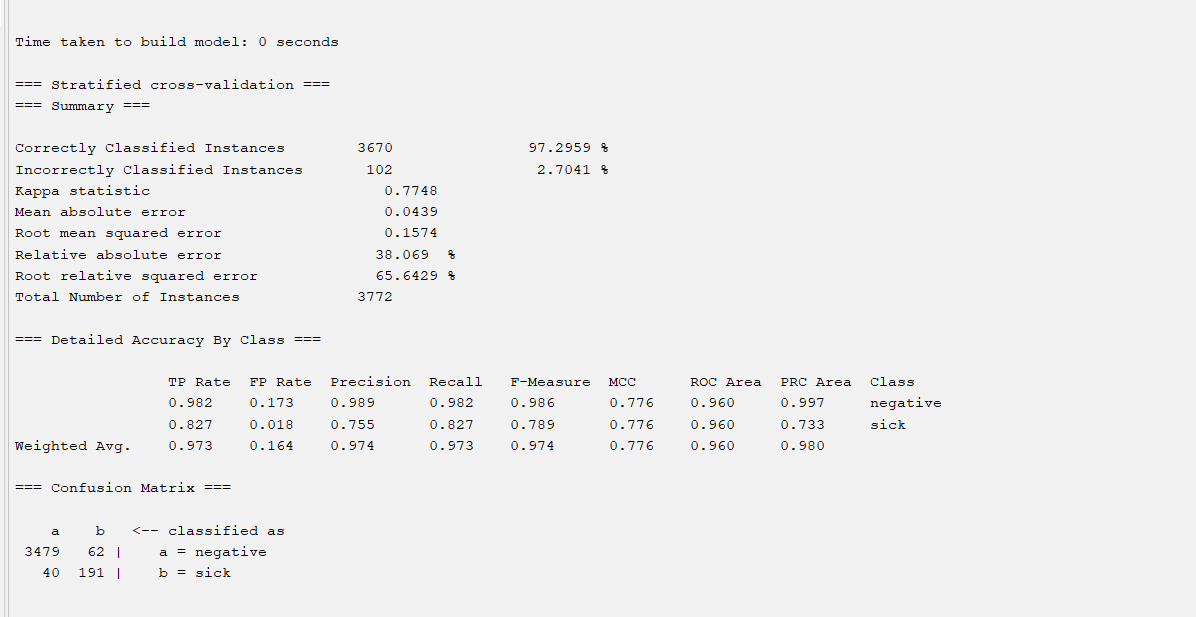
****

**3. Naive Bayes with an unsupervised discretization filter with different values for the bins attributes.**

**Bin = 10**

****

**Bin = 20**

****