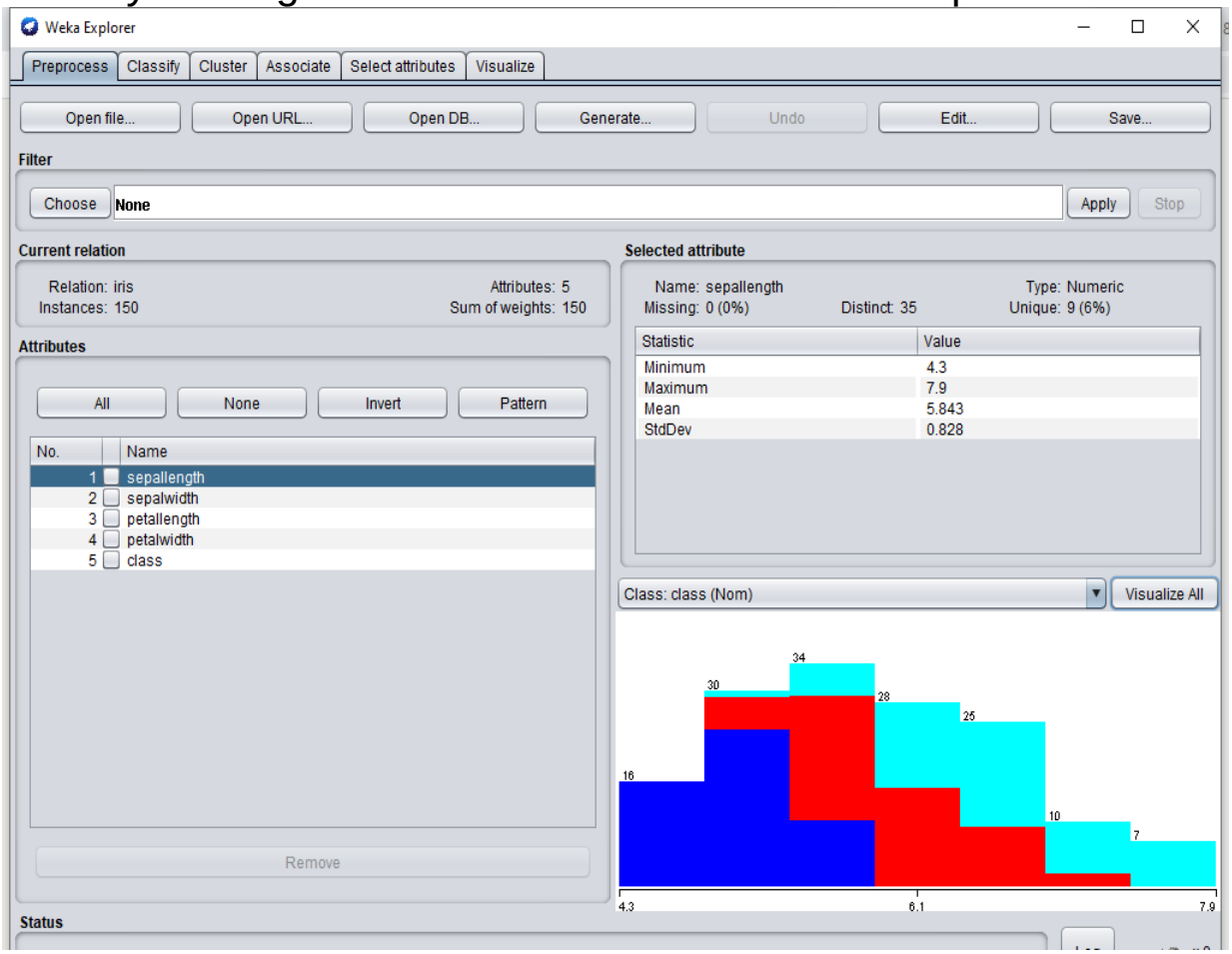
ASSIGNMENT-2

FOR IRIS DATASET:

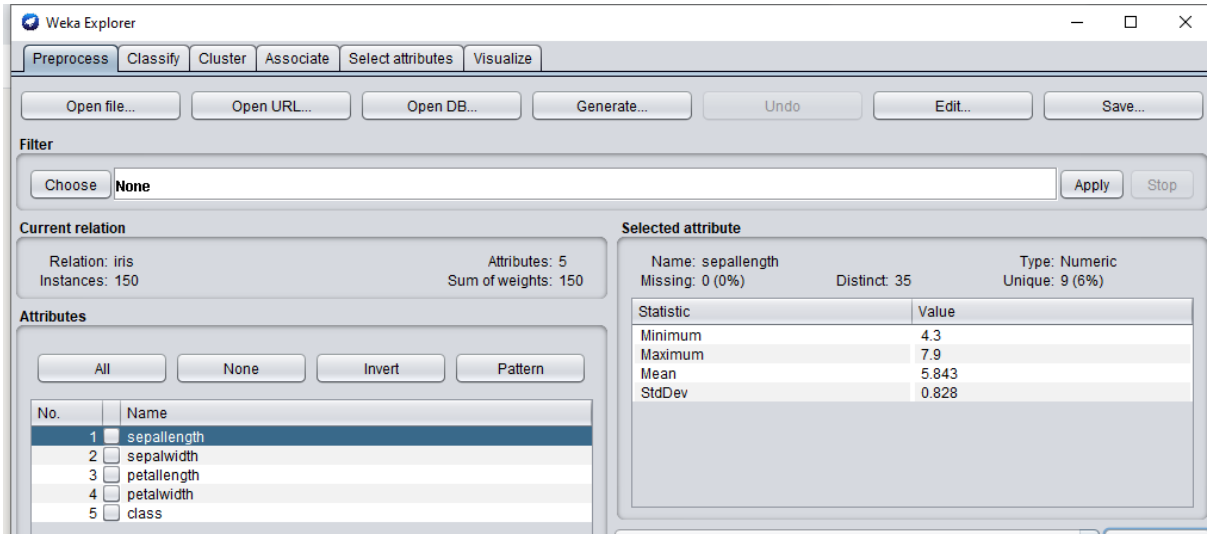
- open the iris arff file from Weka
- Below are the attribute names and their types

Attribute Name	Attribute Type
sepallength	Numeric
sepalwidth	Numeric
petallength	Numeric
petalwidth	Numeric
class	Nominal

- By clicking on the attributes we can see their pie charts.

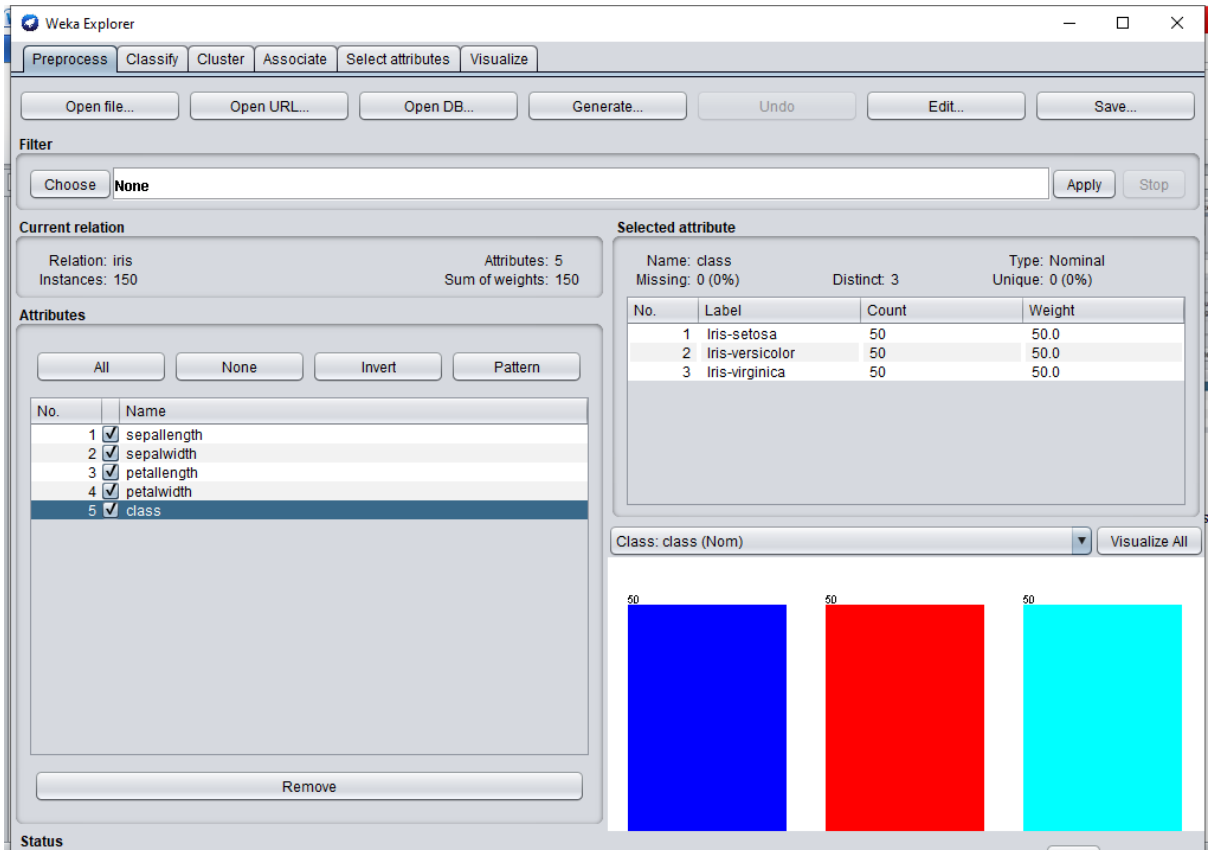


1) Number of records in the dataset



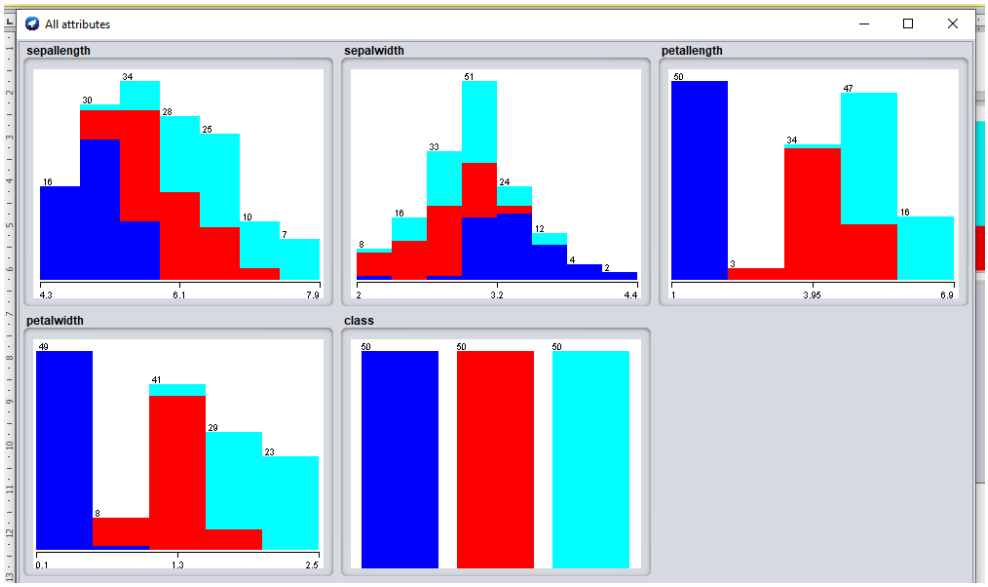
- There are 150 records.

2) Identify the class attribute in the dataset : class



3) Graphical histogram representation of all attributes

against class attribute



4) Determine the number of records for each class:

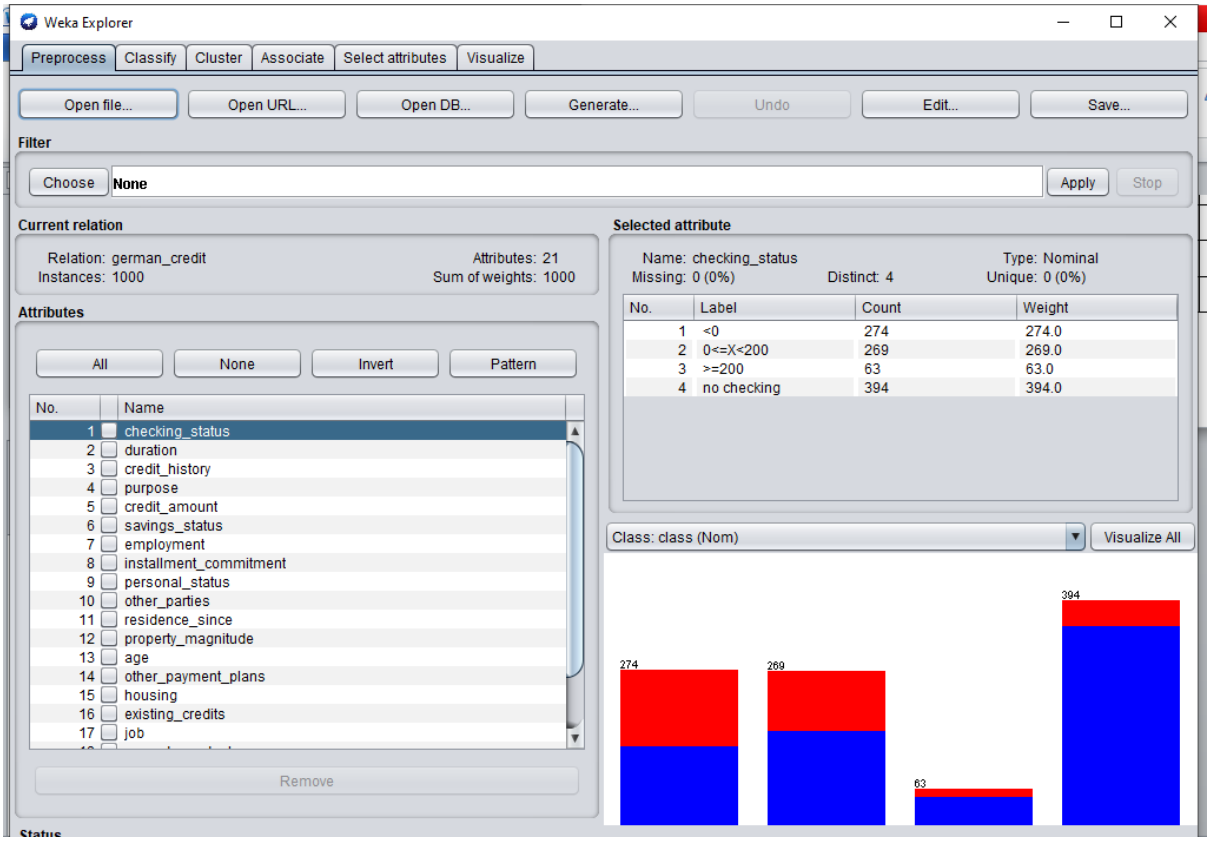
- Iris-setosa – 50 records
- Iris-virginica – 50 records
- Iris-versicolor – 50 records

FOR GERMAN CREDIT DATASET :

1) Below are the attribute names and their types:

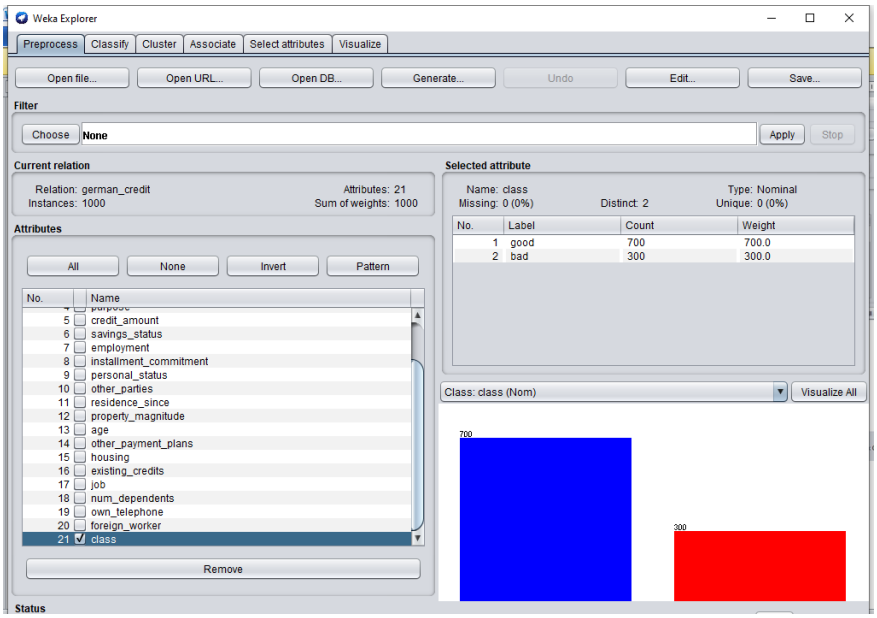
Attribute Name	Attribute Type
credit_history	Nominal
purpose	Nominal
credit_amount	Numeric
serving_status	Nominal
employment	Nominal
installment_commitment	Numeric
personal_status	Nominal
other_parties	Nominal
residence_since	Numeric
property_magnitude	Nominal
age	Numeric
duration	Numeric
housing	Nominal
existing_credits	Numeric
job	Nominal
other_payment_plans	Nominal
checking_status	Nominal
num_dependents	Numeric
own_telephone	Nominal
foreign_worker	Nominal
class	Nominal

2) Number of records in the dataset

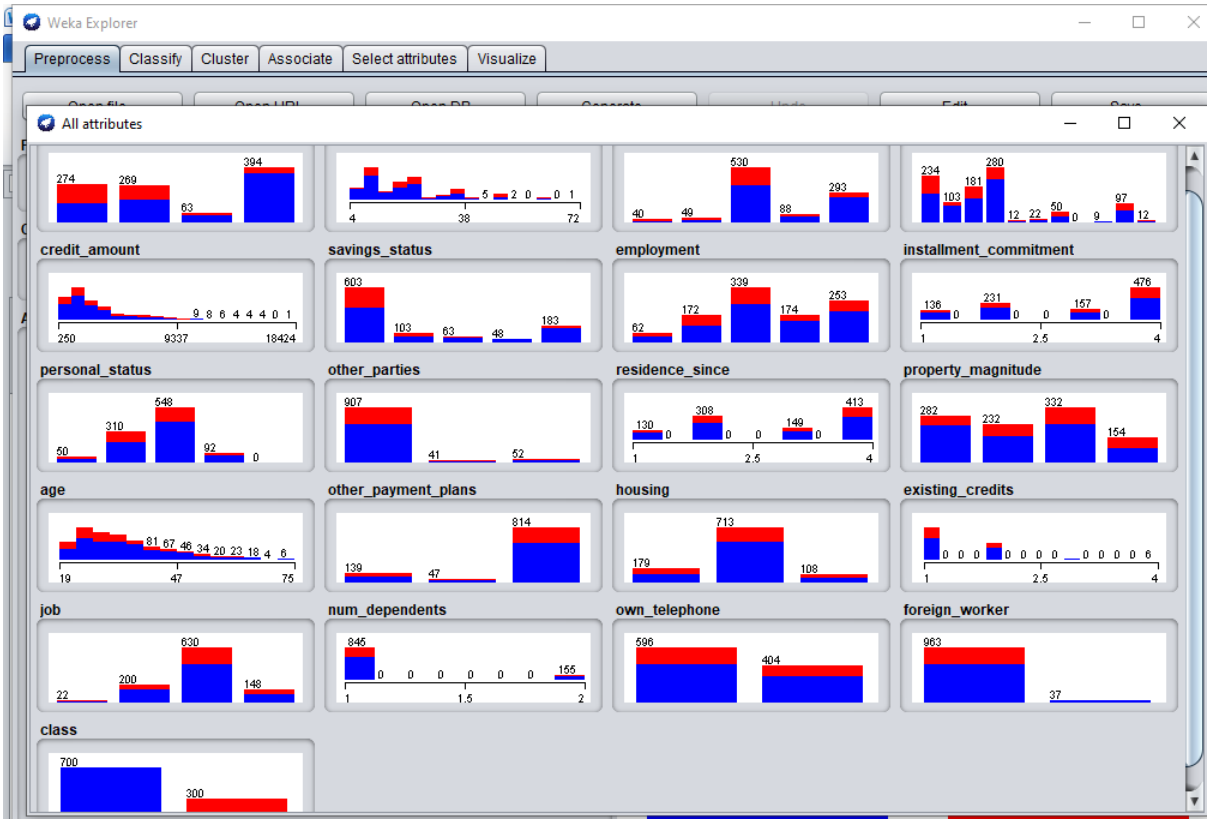


- Therefore there are 1000 Records

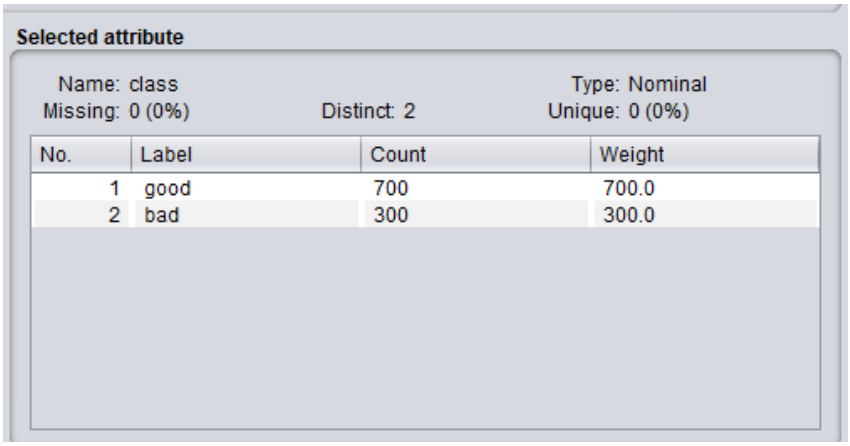
3) Identify the class attribute in the dataset : class



4) Graphical histogram representation of all attributes against class attribute



5) Determine the number of records for each class:



- The above attribute shows the labels and their respected record count.
  - Good – 700 records
  - Bad – 300 records