WEEK-2

Creating new Arff File and CSV file using weather dataset, load the dataset and observe.

WEATHER.NOMINAL

Step 1: List the attribute names and their types.

Name	Type
outlook	nominal
temperature	nominal
humidity	nominal
windy	nominal
play	nominal

Step 2: Number of records in each dataset

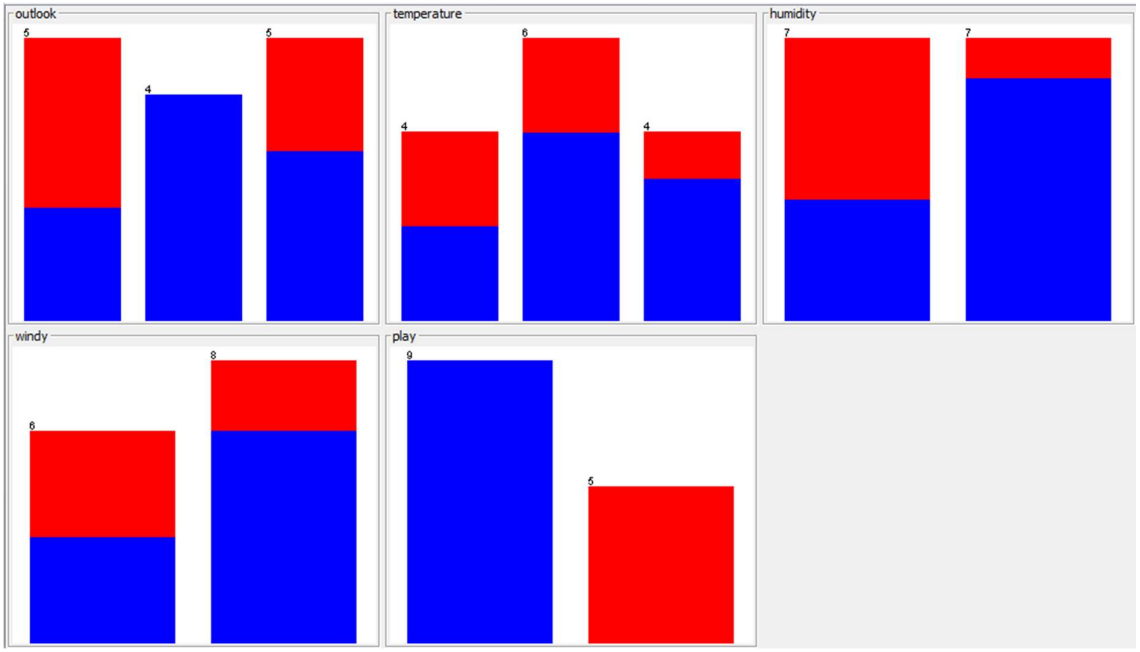
Current relation
Relation: weather.symbolic Attributes: 5
Instances: 14 Sum of weights: 14

Step 3: Identify the class attribute (if any)

Selected attribute
Name: play Type: Nominal
Missing: 0 (0%) Distinct: 2 Unique: 0 (0%)

No.	Label	Count	Weight
1	yes	9	9
2	no	5	5

Step 4: Plot Histogram



Step 5: Determine the number of records for each class

Selected attribute

Name: play Type: Nominal
Missing: 0 (0%) Distinct: 2 Unique: 0 (0%)

No.	Label	Count	Weight
1	yes	9	9
2	no	5	5

Step 6: Visualize the data in various dimensions

DIABETES.NOMINAL

Step 1: List the attribute names and their types.

Name	Type
preg	numeric
plas	numeric
age	numeric
press	numeric
skin	numeric
insu	numeric
mass	numeric
pedi	numeric
age	numeric
class	nominal

Step 2: Number of records in each dataset.

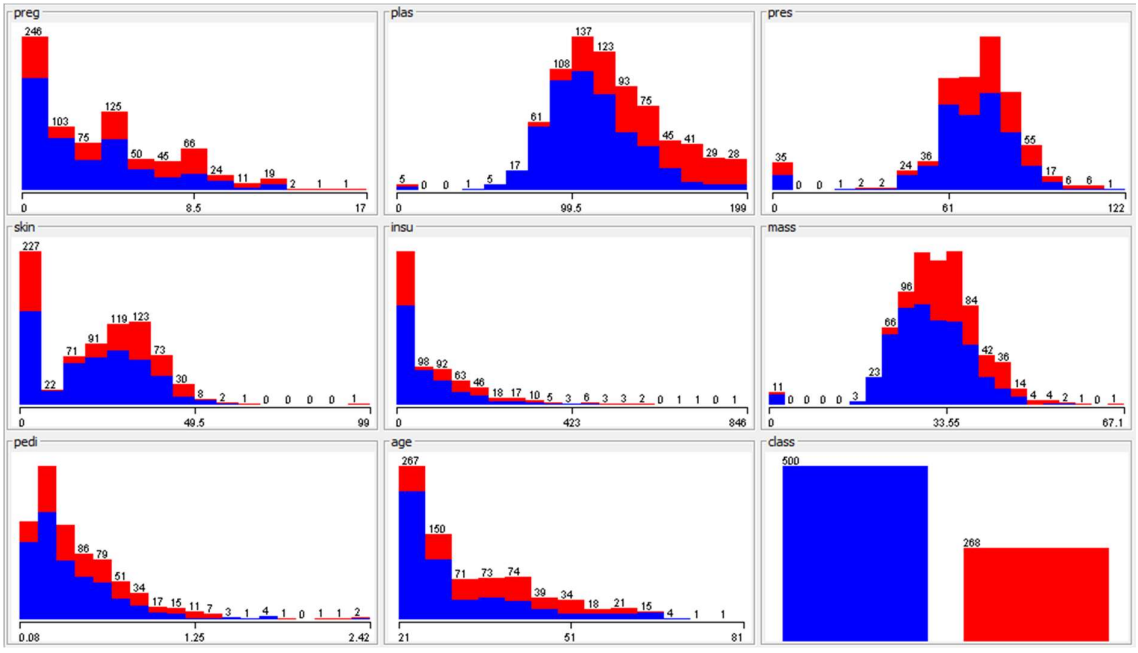
Current relation
Relation: pima_diabetes
Instances: 768
Attributes: 9
Sum of weights: 768

Step 3: Identify the class attribute (if any)

Selected attribute
Name: class
Missing: 0 (0%)
Distinct: 2
Type: Nominal
Unique: 0 (0%)

No.	Label	Count	Weight
1	tested_negative	500	500
2	tested_positive	268	268

Step 4: Plot Histogram



Step 5: Determine the number of records for each class

Current relation

Relation: pima_diabetes
Instances: 768

Attributes: 9
Sum of weights: 768

Step 6: Visualize the data in various dimensions

IRIS.NOMINAL

Step 1: List the attribute names and their types.

Name	Type
Sepallenght	numeric
sepalwidth	numeric
petallength	numeric
petalwidth	numeric
class	nominal

Step 2: Number of records in each dataset.

Current relation

Relation: iris — Attributes: 5

Instances: 150 — Sum of weights: 150

Step 3: Identify the class attribute (if any)

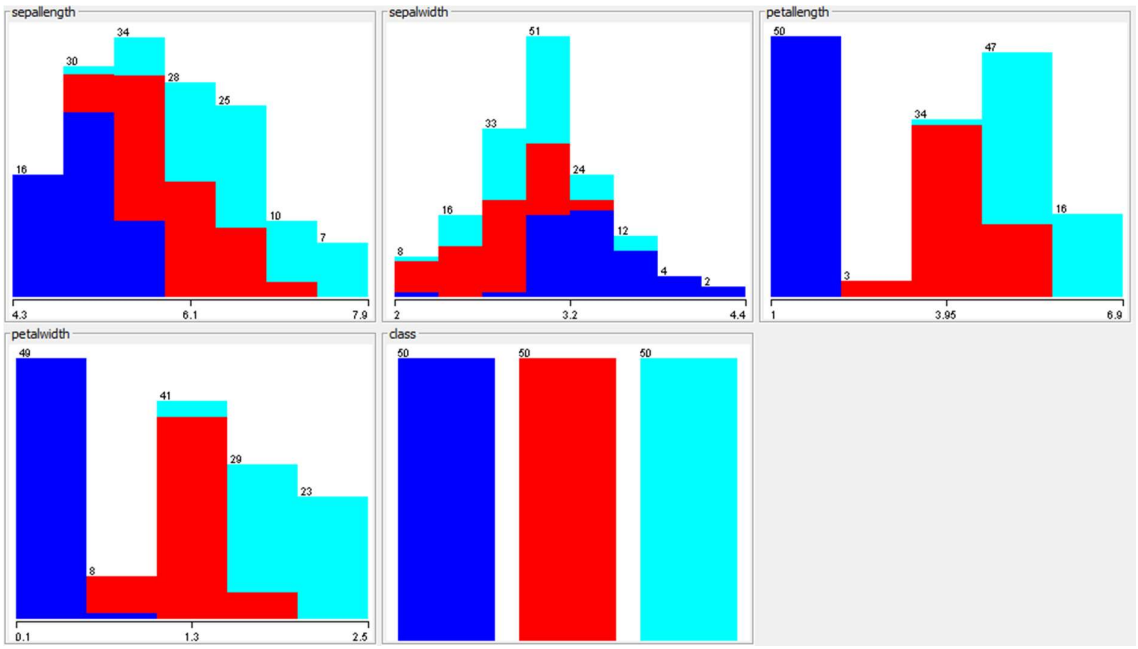
Selected attribute

Name: class — Type: Nominal

Missing: 0 (0%) — Distinct: 3 — Unique: 0 (0%)

No.	Label	Count	Weight
1	Iris-setosa	50	50
2	Iris-versicolor	50	50
3	Iris-virginica	50	50

Step 4: Plot Histogram



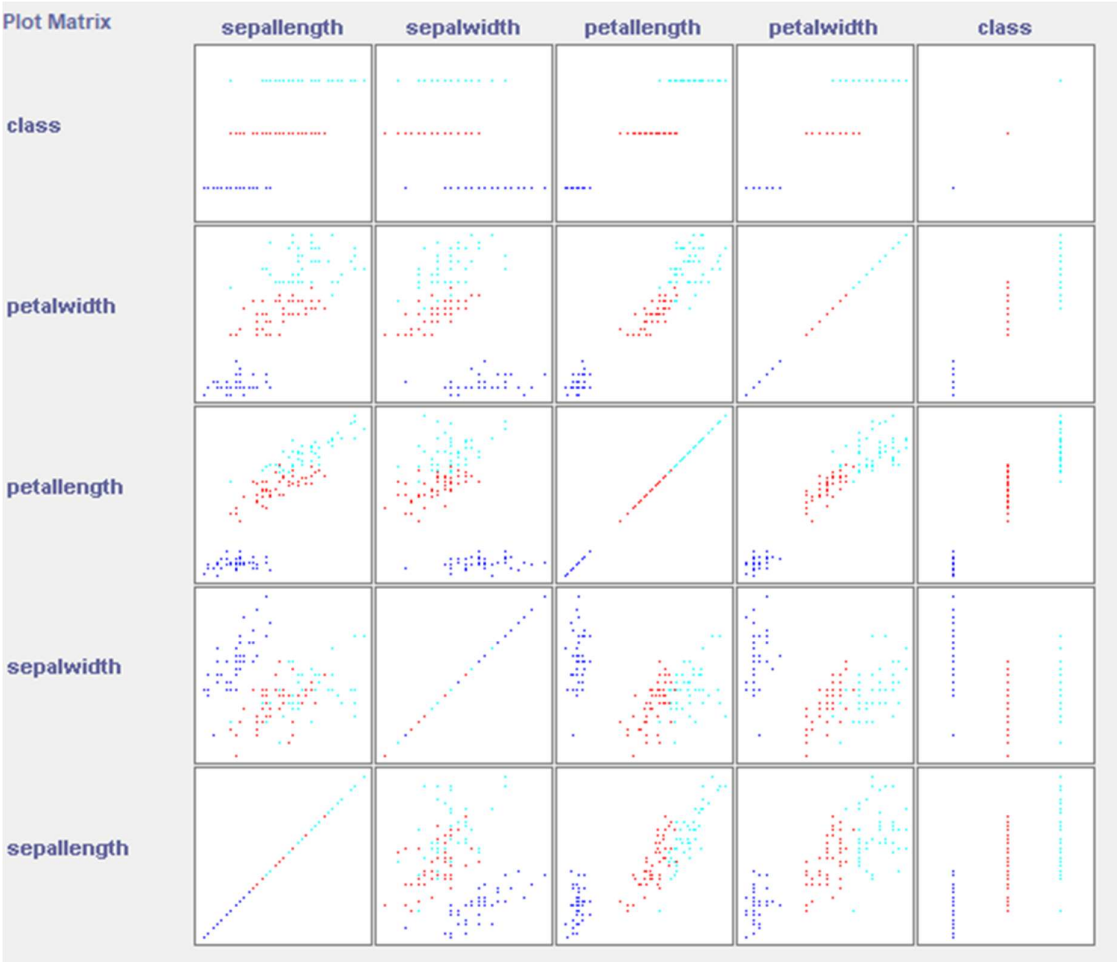
Step 5: Determine the number of records for each class

Selected attribute

Name: class Type: Nominal
Missing: 0 (0%) Distinct: 3 Unique: 0 (0%)

No.	Label	Count	Weight
1	Iris-setosa	50	50
2	Iris-versicolor	50	50
3	Iris-virginica	50	50

Step 6: Visualize the data in various dimensions



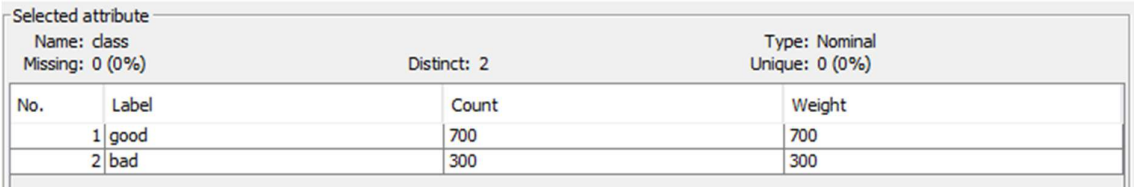
CREDIT-G.NOMINAL

Step 1: List the attribute names and their types.

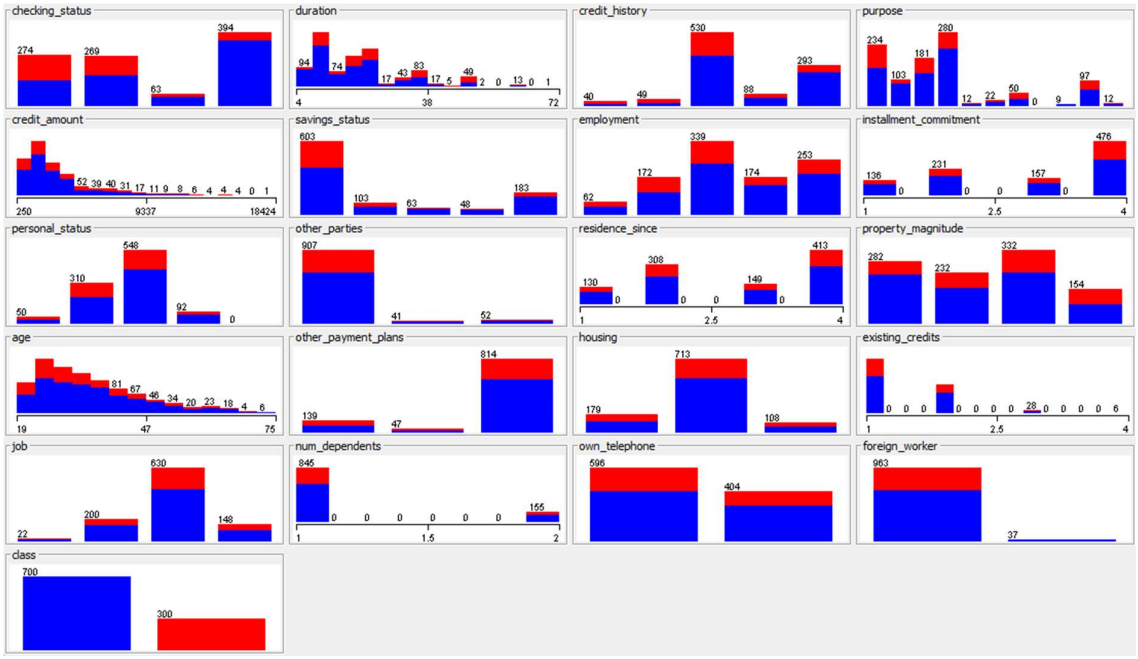
Name	Type
checking_status	nominal
duration	numeric
credit_history	nominal
purpose	nominal
credit_amount	numeric
savings_status	nominal
employment	nominal
instalment_commitment	numeric
personal_status	nominal
other_parties	nominal
residence_since	numeric
property_magnitude	nominal
age	numeric
other_payment_plans	nominal
housing	nominal
existing_credits	numeric
num_dependents	nominal
own_telephone	nominal
foreign_worker	nominal
class	nominal

Step 2: Number of records in each dataset.

Step 3: Identify the class attribute (if any)



Step 4: Plot Histogram



Step 5: Determine the number of records for each class

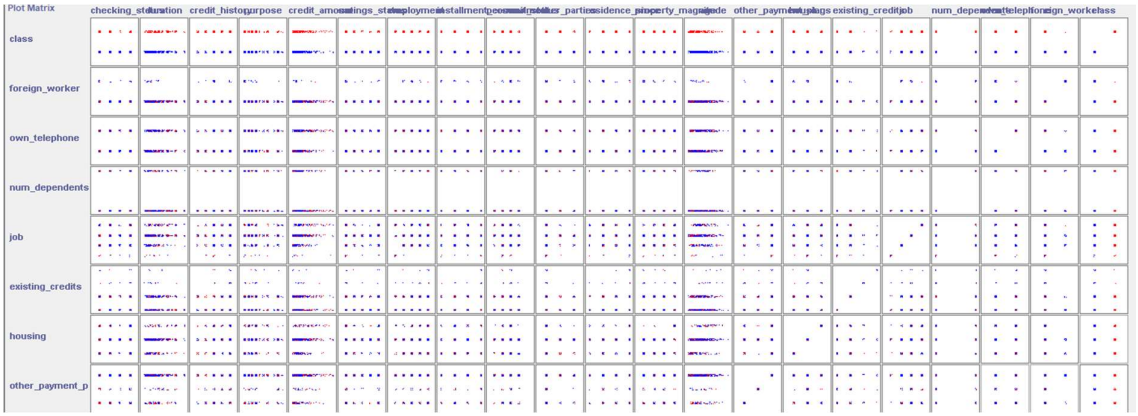
Selected attribute

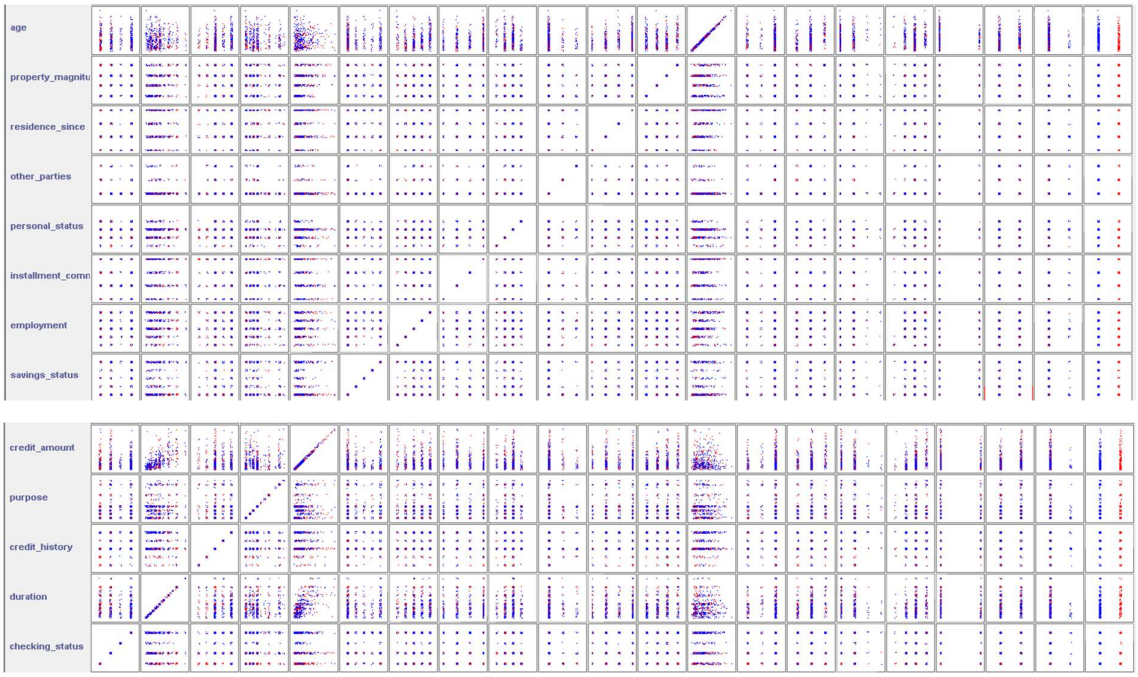
Name: class | Type: Nominal

Missing: 0 (0%) | Distinct: 2 | Unique: 0 (0%)

No.	Label	Count	Weight
1	good	700	700
2	bad	300	300

Step 6: Visualize the data in various dimensions




age
property_magnitu
residence_since
other_parties
personal_status
installment_comm
employment
savings_status
credit_amount
purpose
credit_history
duration
checking_status