**DATA MINING LAB ASSIGNMENT 1**

**Load each dataset and observe the following:**

Observations of Iris dataset:

**1.Loading Iris dataset:**

**i. List the attribute names and its types:**

|  |  |  |
| --- | --- | --- |
| ATTRIBUTE NAME | | ATTRIBUTE TYPE |
| sepallength | | Numeric |
| sepalwidth | | Numeric |
| petallength | | Numeric |
| petalwidth | | Numeric |
| **class** | Nominal | |

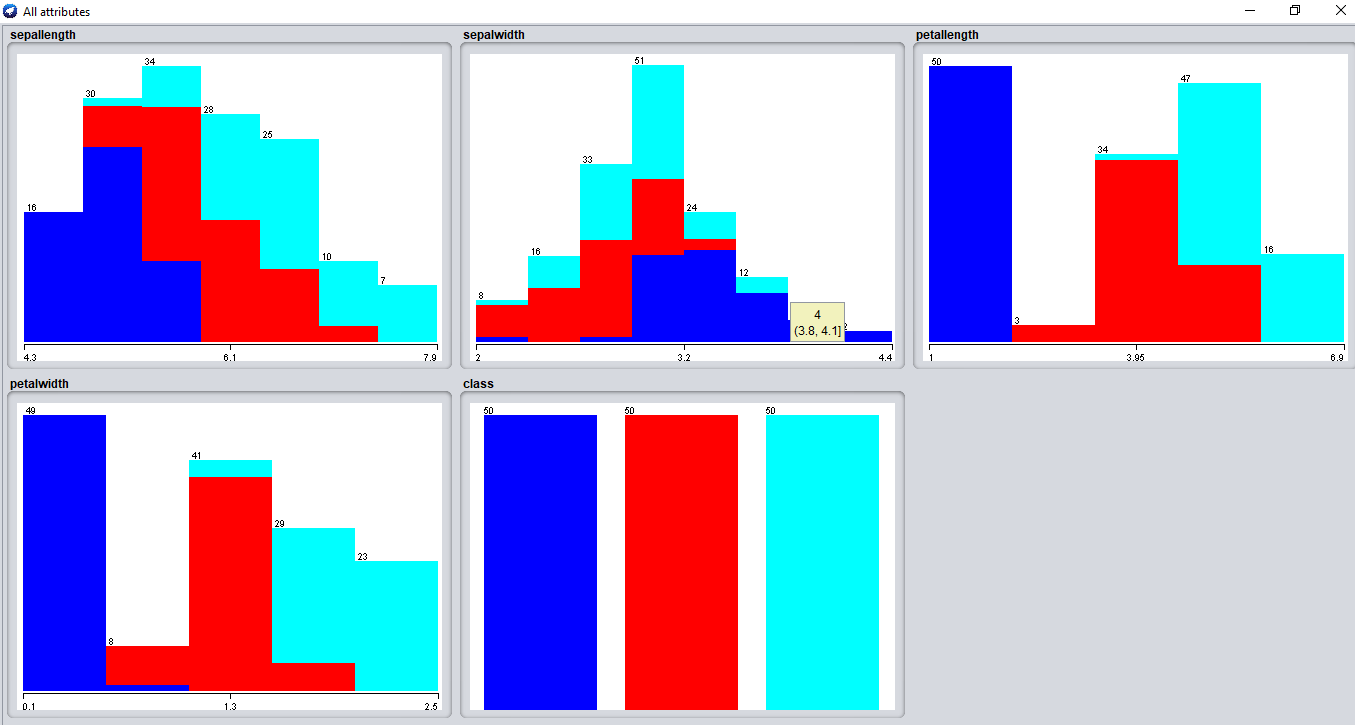
**ii. Number of records in each dataset:** 150(instances)

**iii. Identify the class attribute(if any):**

Only one class attribute is identified in the iris dataset

|  |  |
| --- | --- |
| Class Attribute | Its Type |
| class | Nominal |

**iv. Plot Histogram**:



**V. Determine the number of records in each class:**

Class attribute has three labels. They are:

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

**Load each dataset and observe the following**:

Observations of GermanCredit Dataset:

**1.** **Loading Germancredit dataset**:

**i.** **List the attribute names and its types**:

|  |  |
| --- | --- |
| **Attribute name** | **Its type** |
| checking\_status | Nominal |
| duration | Numeric |
| credit\_history | Nominal |
| purpose | Nominal |
| credit\_amount | Numeric |
| savings\_status | Nominal |
| employment | Nominal |
| installment\_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 |
| job | Nominal |
| num\_dependents | Numeric |
| own\_telephone | Nominal |
| foreign\_worker | Nominal |
| class | Nominal |

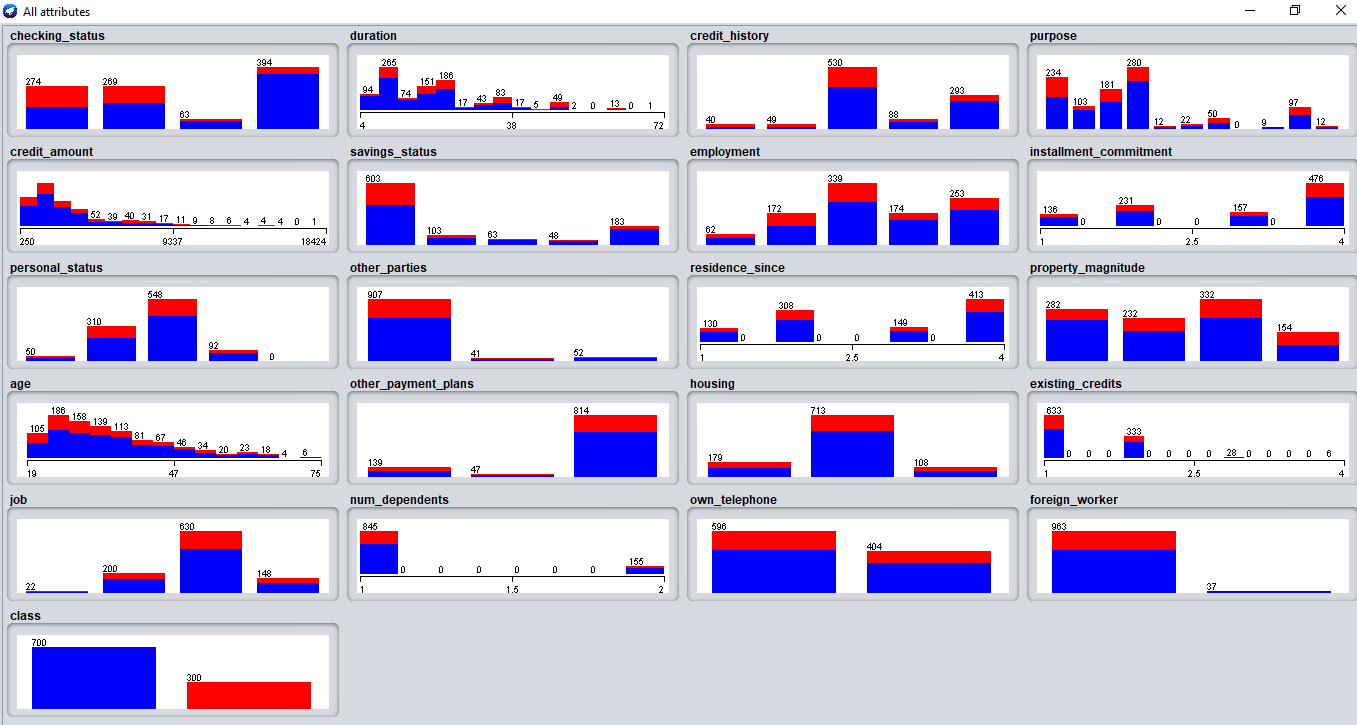
**ii. Number of records in each dataset**: 1000(instances)

**iii. Identify the class attribute(if any):**

The class attributes identified are:

class - Nominal

**iv. Plot Histogram:**



**V. Determine the number of records in each class**:

Number of records for class:

i. good - 700

**ii.** bad - 300